

Appendix

[App. A](#) contains all proofs for our new contributions and observations regarding AST. Afterwards, in [App. B](#), we introduce the notion of *strong* almost-sure termination (SAST) and prove our results for PAST and SAST as well.

A Proofs

In this section, we prove all our new contributions and observations concerning AST. Most of our proofs do not handle rewrite sequences with $\Rightarrow$, but instead we work with rewrite sequence trees that were first described in [30]. We already gave an example for such a tree in [Ex. 16](#). We start by recapitulating this notion.

Definition 39 (Rewrite Sequence Tree (RST)). *Let $\mathcal{S}$ be a PTRS. $\mathfrak{T} = (V, E, L)$ is an $\mathcal{S}$ -rewrite sequence tree (RST) if*

- (1) $V \neq \emptyset$ is a possibly infinite set of nodes and $E \subseteq V \times V$ is a set of directed edges, such that (V, E) is a (possibly infinite) directed tree where $vE = \{w \mid (v, w) \in E\}$ is finite for every $v \in V$.
- (2) $L : V \rightarrow (0, 1] \times \mathcal{T}(\Sigma, \mathcal{V})$ labels every node v by a probability p_v and a term t_v . For the root $v \in V$ of the tree, we have $p_v = 1$.
- (3) For all $v \in V$: If $vE = \{w_1, \dots, w_k\}$, then $t_v \rightarrow_{\mathcal{S}} \{\frac{p_{w_1}}{p_v} : t_{w_1}, \dots, \frac{p_{w_k}}{p_v} : t_{w_k}\}$.

If we restrict condition (3) to innermost / leftmost-innermost steps, then $\mathfrak{T}$ is an innermost / leftmost-innermost $\mathcal{S}$ -rewrite sequence tree, respectively.

When it is not clear about which RST we are talking, we will always explicitly indicate the tree. For instance, for the probability p_v of the node $v \in V$ of some RST $\mathfrak{T} = (V, E, L)$, we may also write $p_v^{\mathfrak{T}}$. $\text{Leaf}^{\mathfrak{T}}$ denotes the set of leaves of the RST $\mathfrak{T}$ and for a node x in $\mathfrak{T}$, $d^{\mathfrak{T}}(x)$ denotes the depth of node x in $\mathfrak{T}$.

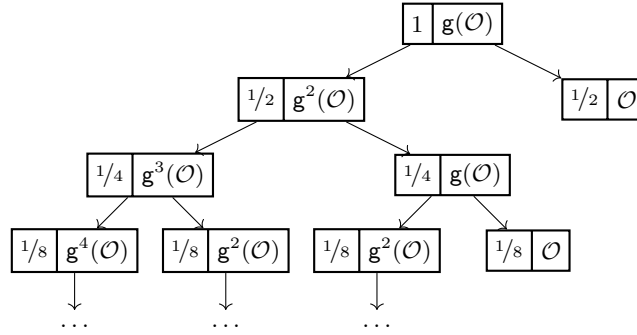
Definition 40 ($|\mathfrak{T}|_{\text{Leaf}}$, Convergence Probability). *Let $\mathcal{S}$ be a PTRS. For any $\mathcal{S}$ -RST $\mathfrak{T}$ we define $|\mathfrak{T}|_{\text{Leaf}} = \sum_{v \in \text{Leaf}} p_v$. We say that an RST $\mathfrak{T}$ converges with probability $p \in \mathbb{R}$ if we have $|\mathfrak{T}|_{\text{Leaf}} = p$.*

It is now easy to observe that a PTRS $\mathcal{S}$ is AST (i.e., for all $\Rightarrow_{\mathcal{S}}$ -rewrite sequences $(\mu_n)_{n \in \mathbb{N}}$ we have $\lim_{n \rightarrow \infty} |\mu_n|_{\mathcal{S}} = 1$) iff for all $\mathcal{S}$ -RSTs $\mathfrak{T}$ we have $|\mathfrak{T}|_{\text{Leaf}} = 1$. To see this, note that every infinite $\Rightarrow_{\mathcal{S}}$ -rewrite sequence $(\mu_n)_{n \in \mathbb{N}}$ that begins with a single start term (i.e., $\mu_0 = \{1 : t\}$) can be represented by an infinite $\mathcal{S}$ -RST $\mathfrak{T}$ that is fully evaluated (i.e., for every leaf v , t_v is an $\mathcal{S}$ -normal form) such that $\lim_{n \rightarrow \infty} |\mu_n|_{\mathcal{S}} = |\mathfrak{T}|_{\text{Leaf}}$ and vice versa. So $\mathcal{S}$ is AST iff all fully evaluated $\mathcal{S}$ -RSTs with a single start term converge with probability 1.

Example 41. Let $\mathcal{S}_{\text{rw2}}$ be the PTRS with the rule $\mathbf{g}(x) \rightarrow \{1/2 : \mathbf{g}^2(x), 1/2 : x\}$. Here, we write $\mathbf{g}^2(x)$ to abbreviate $\mathbf{g}(\mathbf{g}(x))$, etc. Consider the infinite rewrite sequence:

$$\begin{aligned} & \{1 : \mathbf{g}(\mathcal{O})\} \\ \Rightarrow_{\mathcal{S}_{\text{rw2}}} & \{1/2 : \mathbf{g}^2(\mathcal{O}), 1/2 : \mathcal{O}\} \\ \Rightarrow_{\mathcal{S}_{\text{rw2}}} & \{1/4 : \mathbf{g}^3(\mathcal{O}), 1/4 : \mathbf{g}(\mathcal{O}), 1/2 : \mathcal{O}\} \\ \Rightarrow_{\mathcal{S}_{\text{rw2}}} & \{1/8 : \mathbf{g}^4(\mathcal{O}), 1/8 : \mathbf{g}^2(\mathcal{O}), 1/8 : \mathbf{g}^2(\mathcal{O}), 1/8 : \mathcal{O}, 1/2 : \mathcal{O}\} \\ \Rightarrow_{\mathcal{S}_{\text{rw2}}} & \dots \end{aligned}$$

This sequence can be represented by the following $\mathcal{S}_{\text{rw2}}$ -RST:



Furthermore, for every $\mathcal{S}$ -RST $\mathfrak{T}$ that is not fully evaluated, there exists an $\mathcal{S}$ -RST $\mathfrak{T}'$ that is fully evaluated such that $|\mathfrak{T}|_{\text{Leaf}} \geq |\mathfrak{T}'|_{\text{Leaf}}$. To get from $\mathfrak{T}$ to $\mathfrak{T}'$ we can simply perform arbitrary (possibly infinitely many) rewrite steps at the leaves that are not in normal form to fully evaluate the tree. So $\mathcal{S}$ is AST iff all $\mathcal{S}$ -RSTs with a single start term converge with probability 1.

It remains to prove that it suffices to only regard $\Rightarrow$ -rewrite sequences that start with a single start term.

Lemma 42 (Single Start Terms Suffice for AST). *Suppose that there exists an infinite $\Rightarrow_{\mathcal{S}}$ -rewrite sequence $(\mu_n)_{n \in \mathbb{N}}$ that converges with probability < 1 . Then there is also an infinite $\Rightarrow_{\mathcal{S}}$ -rewrite sequence $(\mu'_n)_{n \in \mathbb{N}}$ with a single start term, i.e., $\mu'_0 = \{1 : t\}$, that converges with probability < 1 .*

Proof. Let $(\mu_n)_{n \in \mathbb{N}}$ be an $\Rightarrow_{\mathcal{S}}$ -rewrite sequence that converges with probability < 1 . Suppose that we have $\mu_0 = \{p_1 : t_1, \dots, p_k : t_k\}$. Let $(\mu_n^j)_{n \in \mathbb{N}}$ with $\mu_0^j = \{1 : t_j\}$ denote the infinite $\Rightarrow_{\mathcal{S}}$ -rewrite sequence that uses the same rules as $(\mu_n)_{n \in \mathbb{N}}$ does on the term t_j for every $1 \leq j \leq k$. Assume for a contradiction that for every $1 \leq j \leq k$ the $\Rightarrow_{\mathcal{S}}$ -rewrite sequence $(\mu_n^j)_{n \in \mathbb{N}}$ converges with probability 1.

Then we would have

$$\begin{aligned}
& \lim_{n \rightarrow \infty} |\mu_n|_{\mathcal{S}} \\
&= \lim_{n \rightarrow \infty} \sum_{(p:t) \in \mu_n, t \in \text{NF}_{\mathcal{S}}} p \\
&= \lim_{n \rightarrow \infty} \sum_{1 \leq j \leq k} p_j \cdot \sum_{(p:t) \in (\mu_n^j)_{n \in \mathbb{N}}, t \in \text{NF}_{\mathcal{S}}} p \\
&= \sum_{1 \leq j \leq k} p_j \cdot \lim_{n \rightarrow \infty} \sum_{(p:t) \in (\mu_n^j)_{n \in \mathbb{N}}, t \in \text{NF}_{\mathcal{S}}} p \\
&= \sum_{1 \leq j \leq k} p_j \cdot \lim_{n \rightarrow \infty} |\mu_n^j|_{\mathcal{S}} \\
&= \sum_{1 \leq j \leq k} p_j \cdot 1 \\
&= \sum_{1 \leq j \leq k} p_j \\
&= 1
\end{aligned}$$

which is a contradiction to our assumption that we have $\lim_{n \rightarrow \infty} |\mu_n|_{\mathcal{S}} < 1$. Therefore, we have at least one $1 \leq j \leq k$ such that $(\mu_n^j)_{n \in \mathbb{N}}$ converges with probability < 1 . $\square$

Obviously, all of these observations also hold for iAST and for liAST. We now obtain the following corollary.

Corollary 43 (Characterizing AST with RSTs). *Let $\mathcal{S}$ be a PTRS. Then $\mathcal{S}$ is fAST (iAST/liAST) iff for all (innermost/leftmost-innermost) $\mathcal{S}$ -RSTs $\mathfrak{T}$ we have $|\mathfrak{T}|_{\text{Leaf}} = 1$.*

Next, we prove all new results from Sect. 4. For this, we first recapitulate the notion of a *sub-rewrite sequence tree* (sub-RST) from [30].

Definition 44 (Sub-RST). *Let $\mathcal{S}$ be a PTRS and $\mathfrak{T} = (V, E, L)$ be an $\mathcal{S}$ -RST. Let $W \subseteq V$ be non-empty, weakly connected, and for all $x \in W$ we have $xE \cap W = \emptyset$ or $xE \cap W = xE$. Then, we define the sub-rewrite sequence tree (or sub-RST) $\mathfrak{T}[W]$ by $\mathfrak{T}[W] = (W, E \cap (W \times W), L^W)$. Let $w \in W$ be the root of $\mathfrak{T}[W]$. To ensure that the root of our sub-RST has the probability 1 again, we use the labeling $L^W(x) = (\frac{p_x^{\mathfrak{T}}}{p_w^{\mathfrak{T}}} : t_x^{\mathfrak{T}})$ for all nodes $x \in W$.*

The property of being non-empty and weakly connected ensures that the resulting graph $G^{\mathfrak{T}[W]} = (W, E \cap (W \times W))$ is a tree again. The property that we either have $xE \cap W = \emptyset$ or $xE \cap W = xE$ ensures that the sum of probabilities for the successors of a node x is equal to the probability for the node x itself.

Theorem 17 (From iAST to fAST (1)). *If a PTRS $\mathcal{S}$ is orthogonal and right-linear (i.e., non-overlapping and linear), then:*

$$\mathcal{S} \text{ is fAST} \iff \mathcal{S} \text{ is iAST}$$

Proof. We only have to prove the non-trivial direction “ $\Leftarrow$ ”. Let $\mathcal{S}$ be a PTRS that is non-overlapping and linear. Furthermore, assume that $\mathcal{S}$ is not fAST. This means that there exists an RST $\mathfrak{T}$ such that $|\mathfrak{T}|_{\text{Leaf}} = c$ for some $0 \leq c < 1$. We create a new *innermost* RST $\mathfrak{T}^{(\infty)}$ such that $|\mathfrak{T}^{(\infty)}|_{\text{Leaf}} \leq |\mathfrak{T}|_{\text{Leaf}} = c < 1$, which shows that $\mathcal{S}$ is not iAST either. W.l.o.g., at least one rewrite step in $\mathfrak{T}$ is performed at some node x with a redex that is not an innermost redex (otherwise we can use $\mathfrak{T}^{(\infty)} = \mathfrak{T}$). The core steps of the proof are the following:

1. We iteratively move innermost rewrite steps to a higher position in the tree using a construction $\Phi(\circ)$. The limit of this iteration, namely $\mathfrak{T}^{(\infty)}$, is an *innermost* $\mathcal{S}$ -RST that converges with probability at most $c < 1$.
 - 1.1. We formally define the construction $\Phi(\circ)$ that replaces a certain subtree $\mathfrak{T}_x$ by a new subtree $\Phi(\mathfrak{T}_x)$, by moving an innermost rewrite step to the root.
 - 1.2. We show $|\Phi(\mathfrak{T}_x)|_{\text{Leaf}} \leq |\mathfrak{T}_x|_{\text{Leaf}}$.
 - 1.3. We show that $\Phi(\mathfrak{T}_x)$ is indeed a valid RST.

1. We iteratively move innermost rewrite steps to a higher position.

In $\mathfrak{T}$ there exists at least one rewrite step performed at some node x , which is not an innermost rewrite step. Furthermore, we can assume that this is the first such rewrite step in the path from the root to the node x and that x is a node of minimum depth with this property. Let $\mathfrak{T}_x$ be the sub-RST that starts at node x , i.e., $\mathfrak{T}_x = \mathfrak{T}[xE^*]$. We then construct a new tree $\Phi(\mathfrak{T}_x)$ such that $|\Phi(\mathfrak{T}_x)|_{\text{Leaf}} = |\mathfrak{T}_x|_{\text{Leaf}}$, and where we use an innermost rewrite step at the root node x instead of the old one (i.e., we pushed the first non-innermost rewrite step deeper into the tree). This construction only works because $\mathcal{S}$ is non-overlapping and linear.⁵ Then, by replacing the subtree $\mathfrak{T}_x$ with the new tree $\Phi(\mathfrak{T}_x)$ in $\mathfrak{T}$, we get an $\mathcal{S}$ -RST $\mathfrak{T}^{(1)}$, with $|\mathfrak{T}^{(1)}|_{\text{Leaf}} = |\mathfrak{T}|_{\text{Leaf}}$, and where we use an innermost rewrite step at node x instead of the old rewrite step, as desired. We can then do such a replacement iteratively for every use of a non-innermost rewrite step, i.e., we again replace the first non-innermost rewrite step in $\mathfrak{T}^{(1)}$ to get $\mathfrak{T}^{(2)}$ with $|\mathfrak{T}^{(2)}|_{\text{Leaf}} = |\mathfrak{T}^{(1)}|_{\text{Leaf}}$, and so on. In the end, the limit of all these RSTs $\lim_{i \rightarrow \infty} \mathfrak{T}^{(i)}$ is an innermost $\mathcal{S}$ -RST, that we denote by $\mathfrak{T}^{(\infty)}$ and that converges with probability at most $c < 1$. So while the termination probability remains the same in each step, it can decrease in the limit.⁶ Hence, $\mathcal{S}$ is not iAST either.

To see that $\mathfrak{T}^{(\infty)}$ is indeed a valid innermost $\mathcal{S}$ -RST, note that in every iteration of the construction we turn a non-innermost rewrite step at minimum depth into an innermost one. Hence, for every depth H of the tree, we eventually turned every non-innermost rewrite step up to depth H into an innermost one so that the construction will not change the tree above depth H anymore, i.e., there exists an m_H such that $\mathfrak{T}^{(\infty)}$ and $\mathfrak{T}^{(i)}$ are the same trees up to depth H for all $i \geq m_H$. This means that the sequence $\lim_{i \rightarrow \infty} \mathfrak{T}^{(i)}$ really converges into an innermost $\mathcal{S}$ -RST.

Next, we want to prove that we have $|\mathfrak{T}^{(\infty)}|_{\text{Leaf}} \leq c$. By induction on n one can prove that $|\mathfrak{T}^{(i)}|_{\text{Leaf}} = c$ for all $1 \leq i \leq n$, since we have $|\mathfrak{T}^{(i)}|_{\text{Leaf}} = |\mathfrak{T}^{(i-1)}|_{\text{Leaf}}$ for all $i \geq 2$ and $|\mathfrak{T}^{(1)}|_{\text{Leaf}} = |\mathfrak{T}|_{\text{Leaf}} = c$. Assume for a contradiction that $\mathfrak{T}^{(\infty)}$ converges with probability greater than c , i.e., that $|\mathfrak{T}^{(\infty)}|_{\text{Leaf}} > c$. Then there exists a depth $H \in \mathbb{N}$ such that $\sum_{x \in \text{Leaf}^{\mathfrak{T}^{(\infty)}}, d^{\mathfrak{T}^{(\infty)}}(x) \leq H} p_x > c$.

⁵ For the construction, non-overlappingness is essential (while a related construction could also be defined without linearity). However, linearity is needed to ensure that the probability of termination in the new tree is not larger than in the original one.

⁶ As an example, consider a tree $\mathfrak{T}$ which is just a finite path and its path length increases in each iteration by one. Then the limit $\mathfrak{T}^{(\infty)}$ is an infinite path and converges with probability 0, while all the $\mathfrak{T}, \mathfrak{T}^{(1)}, \dots$ converge with probability 1.

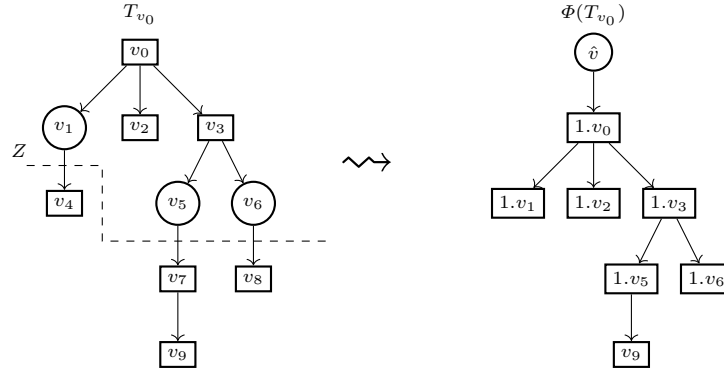
Again, let $m_H \in \mathbb{N}$ such that $\mathfrak{T}^{(\infty)}$ and $\mathfrak{T}^{(m_H)}$ are the same trees up to depth H . But this would mean that $|\mathfrak{T}^{(m_H)}|_{\text{Leaf}} \geq \sum_{x \in \text{Leaf}^{\mathfrak{T}^{(m_H)}}, d^{\mathfrak{T}^{(m_H)}}(x) \leq H} p_x = \sum_{x \in \text{Leaf}^{\mathfrak{T}^{(\infty)}}, d^{\mathfrak{T}^{(\infty)}}(x) \leq H} p_x > c$, which is a contradiction to $|\mathfrak{T}^{(m_H)}|_{\text{Leaf}} = c$.

1.1 Construction of $\Phi(\circ)$

It remains to define the mentioned construction $\Phi(\circ)$. Let $\mathfrak{T}_x$ be an $\mathcal{S}$ -RST that performs a non-innermost rewrite step at the root node x . This step has the form $t_x^{\mathfrak{T}_x} \rightarrow_{\mathcal{S}} \{p_{y_1}^{\mathfrak{T}_x} : t_{y_1}^{\mathfrak{T}_x}, \dots, p_{y_k}^{\mathfrak{T}_x} : t_{y_k}^{\mathfrak{T}_x}\}$ using the rule $\bar{\ell} \rightarrow \{\bar{p}_1 : \bar{r}_1, \dots, \bar{p}_k : \bar{r}_k\}$, the substitution $\bar{\sigma}$, and the position $\bar{\pi}$ such that $t_x^{\mathfrak{T}_x}|_{\bar{\pi}} = \bar{\ell}\bar{\sigma}$. Then we have $t_{y_j}^{\mathfrak{T}_x} = t_x^{\mathfrak{T}_x}[\bar{r}_j\bar{\sigma}]_{\bar{\pi}}$ for all $1 \leq j \leq k$. Instead of applying a non-innermost rewrite step at the root x we want to directly apply an innermost rewrite step. Let τ be the position of some innermost redex in $t_x^{\mathfrak{T}_x}$ below $\bar{\pi}$.

The construction creates a new $\mathcal{S}$ -RST $\Phi(\mathfrak{T}_x) = (V', E', L')$ whose root is labeled with $(1 : t_x^{\mathfrak{T}_x})$ such that $|\Phi(\mathfrak{T}_x)|_{\text{Leaf}} = |\mathfrak{T}_x|_{\text{Leaf}}$, and that directly performs the first rewrite step at position τ in the original tree $\mathfrak{T}_x$ (which is an innermost rewrite step) at the root of the tree, by pushing it from the original nodes in the tree $\mathfrak{T}_x$ to the root of the new tree $\Phi(\mathfrak{T}_x)$. (It could also be that this innermost redex was never reduced in $\mathfrak{T}_x$.) This can be seen in the diagram above. This push only results in the same termination probability due to our restriction that $\mathcal{S}$ is linear.

In the figure below, we illustrate the effect of Φ , where the original tree $\mathfrak{T}_x$ with $x = v_0$ is on the left and $\Phi(\mathfrak{T}_x)$ is on the right.



However, since we are allowed to rewrite above τ in the original tree $\mathfrak{T}_x$, the actual position of the innermost redex that was originally at position τ might change during the application of a rewrite step. Hence, we recursively define the position $\varphi_\tau(v)$ that contains precisely this redex for each node v in $\mathfrak{T}_x$ until we rewrite at this position. Initially, we have $\varphi_\tau(x) = \tau$. Whenever we have defined $\varphi_\tau(v)$ for some node v , and we have $t_v^{\mathfrak{T}_x} \rightarrow_{\mathcal{S}} \{p_{w_1}^{\mathfrak{T}_x} : t_{w_1}^{\mathfrak{T}_x}, \dots, p_{w_m}^{\mathfrak{T}_x} : t_{w_m}^{\mathfrak{T}_x}\}$ for the direct successors $vE = \{w_1, \dots, w_m\}$, using the rule $\ell \rightarrow \{p_1 : r_1, \dots, p_m : r_m\}$, the substitution σ , and the position π , we do the following: If $\varphi_\tau(v) = \pi$, meaning

that we rewrite this innermost redex, then we set $\varphi_\tau(w_j) = \perp$ for all $1 \leq j \leq m$ to indicate that we have rewritten the innermost redex. If we have $\varphi_\tau(v) \perp \pi$, meaning that the rewrite step takes place on a position that is parallel to $\varphi_\tau(v)$, then we set $\varphi_\tau(w_j) = \varphi_\tau(v)$ for all $1 \leq j \leq m$, as the position of the innermost redex did not change. Otherwise, we have $\pi < \varphi_\tau(v)$ (since we cannot rewrite below $\varphi_\tau(v)$ as it is an innermost redex), and thus there exists a $\chi \in \mathbb{N}^+$ such that $\pi \cdot \chi = \varphi_\tau(v)$. Since the rules of $\mathcal{S}$ are non-overlapping, the redex must be completely “inside” the used substitution σ , and we can find a position α_q of a variable q in ℓ and another position β such that $\chi = \alpha_q \cdot \beta$. Furthermore, since the rule is linear, q only occurs once in ℓ and at most once in r_j for all $1 \leq j \leq m$. If q occurs in r_j at a position ρ_q^j , then we set $\varphi_\tau(w_j) = \rho_q^j \cdot \beta$. Otherwise, we set $\varphi_\tau(w_j) = \top$ to indicate that the innermost redex was erased during the computation. Finally, if $\varphi_\tau(v) \in \{\perp, \top\}$, then we set $\varphi_\tau(w_j) = \varphi_\tau(v)$ for all $1 \leq j \leq m$ as well. So to summarize, $\varphi_\tau(v)$ is now either the position of the innermost redex in t_v , $\top$ to indicate that the redex was erased, or $\perp$ to indicate that we have rewritten the redex.

In the figure above, the circled nodes represent the nodes where we perform a rewrite step at position $\varphi_\tau(v)$. We now define the $\mathcal{S}$ -RST $\Phi(\mathfrak{T}_x)$ whose root $\hat{v}$ is labeled with $(1 : t_x^{\mathfrak{T}_x})$ and that directly performs the rewrite step $t_x^{\mathfrak{T}_x} = t_{\hat{v}}^{\Phi(\mathfrak{T}_x)} \xrightarrow{i}_{\mathcal{S}, \tau} \{\hat{p}_1 : t_{1.x}^{\Phi(\mathfrak{T}_x)}, \dots, \hat{p}_h : t_{h.x}^{\Phi(\mathfrak{T}_x)}\}$, with the rule $\hat{\ell} \rightarrow \{\hat{p}_1 : \hat{r}_1, \dots, \hat{p}_h : \hat{r}_h\} \in \mathcal{S}$, a substitution $\hat{\sigma}$, and the position τ , at the new root $\hat{v}$. Here, we have $t_{\hat{v}}^{\Phi(\mathfrak{T}_x)}|_\tau = \hat{\ell}\hat{\sigma}$. Let Z be the set of all nodes z such that $\varphi_\tau(z) \neq \perp$, i.e., at node z we did not yet perform the rewrite step with the innermost redex. In the example we have $Z = \{v_0, \dots, v_6\} \setminus \{v_4\}$. For each of these nodes $z \in Z$ and each $1 \leq e \leq h$, we create a new node $e.z \in V'$ with edges as in $\mathfrak{T}_x$ for the nodes in Z , e.g., for the node $1.v_3$ we create edges to $1.v_5$ and $1.v_6$. Furthermore, we add the edges from the new root $\hat{v}$ to the nodes $e.x$ for all $1 \leq e \leq h$. Note that x was the root in the tree $\mathfrak{T}_x$ and has to be contained in Z . For example, for the node $\hat{v}$ we create an edge to $1.v_0$. We define the labeling of the nodes in $\Phi(\mathfrak{T}_x)$ as follows for all nodes z in Z :

$$\begin{aligned} \text{(T-1)} \quad t_{e.z}^{\Phi(\mathfrak{T}_x)} &= t_z^{\mathfrak{T}_x}[\hat{r}_e \hat{\sigma}]_{\varphi_\tau(z)} \text{ if } \varphi_\tau(z) \in \mathbb{N}^* \text{ and } t_{e.z}^{\Phi(\mathfrak{T}_x)} = t_z^{\mathfrak{T}_x} \text{ if } \varphi_\tau(z) = \top. \\ \text{(T-2)} \quad p_{e.z}^{\Phi(\mathfrak{T}_x)} &= p_z^{\mathfrak{T}_x} \cdot \hat{p}_e \end{aligned}$$

Now, for a leaf $e.z' \in V'$ either $z' \in V$ is also a leaf (e.g., node v_2 in our example) or we rewrite the innermost redex at position $\varphi_\tau(z')$ at node z' in $\mathfrak{T}_x$ (e.g., node v_1 in our example). In the latter case, due to non-overlappingness, the same rule $\hat{\ell} \rightarrow \{\hat{p}_1 : \hat{r}_1, \dots, \hat{p}_h : \hat{r}_h\}$ and the same substitution $\hat{\sigma}$ were used. If we rewrite $t_{z'}^{\mathfrak{T}_x} \xrightarrow{i}_{\mathcal{S}, \varphi_\tau(z')} \{\hat{p}_1 : t_{w'_1}^{\mathfrak{T}_x}, \dots, \hat{p}_h : t_{w'_h}^{\mathfrak{T}_x}\}$, then we have $t_{w'_e}^{\mathfrak{T}_x} = t_{z'}^{\mathfrak{T}_x}[\hat{r}_e \hat{\sigma}]_{\varphi_\tau(z')} \stackrel{(\text{T-1})}{=} t_{e.z'}^{\Phi(\mathfrak{T}_x)}$ and $p_{w'_e}^{\mathfrak{T}_x} \stackrel{(\text{T-2})}{=} p_{z'}^{\mathfrak{T}_x} \cdot \hat{p}_e = p_{e.z'}^{\Phi(\mathfrak{T}_x)}$. Thus, we can copy the rest of this subtree of $\mathfrak{T}_x$ to our newly generated tree $\Phi(\mathfrak{T}_x)$. In our example, v_1 has the only successor v_4 , hence we can copy the subtree starting at node v_4 , which is only the node itself, to the node $1.v_1$ in $\Phi(\mathfrak{T}_x)$. For v_5 , we have the only successor v_7 , hence we can copy the subtree starting at node v_7 , which is the node itself together with its successor v_9 , to the node $1.v_5$ in $\Phi(\mathfrak{T}_x)$. So essentially, we just had to

define how to construct $\Phi(\mathfrak{T}_x)$ for the part of the tree before we reach nodes v with $\varphi_\tau(v) = \perp$ in $\mathfrak{T}_x$. Now we have to show that $|\Phi(\mathfrak{T}_x)|_{\text{Leaf}} = |\mathfrak{T}_x|_{\text{Leaf}}$ and that $\Phi(\mathfrak{T}_x)$ is indeed a valid $\mathcal{S}$ -RST (i.e., that the edges between nodes $e.z$ with $z \in Z$ and its successors correspond to rewrite steps with $\mathcal{S}$).

1.2 We show $|\Phi(\mathfrak{T}_x)|_{\text{Leaf}} = |\mathfrak{T}_x|_{\text{Leaf}}$.

Let v be a leaf in $\Phi(\mathfrak{T}_x)$. If $v = e'.z$ for some node $z \in Z$ that is a leaf in $\mathfrak{T}_x$ (e.g., node $1.v_2$), then also $e.z$ must be a leaf in $\Phi(\mathfrak{T}_x)$ for every $1 \leq e \leq h$. Here, we get $\sum_{1 \leq e \leq h} p_{e.z}^{\Phi(\mathfrak{T}_x)} \stackrel{(\text{T-2})}{=} \sum_{1 \leq e \leq h} p_z^{\mathfrak{T}_x} \cdot \hat{p}_e = p_z^{\mathfrak{T}_x} \cdot \sum_{1 \leq e \leq h} \hat{p}_e = p_z^{\mathfrak{T}_x} \cdot 1 = p_z^{\mathfrak{T}_x}$, and thus

$$\sum_{\substack{e.z \in \text{Leaf}^{\Phi(\mathfrak{T}_x)} \\ z \in Z, z \in \text{Leaf}^{\mathfrak{T}_x}}} p_{e.z}^{\Phi(\mathfrak{T}_x)} = \sum_{\substack{z \in \text{Leaf}^{\mathfrak{T}_x} \\ z \in Z}} \left(\sum_{1 \leq e \leq h} p_{e.z}^{\Phi(\mathfrak{T}_x)} \right) = \sum_{\substack{z \in \text{Leaf}^{\mathfrak{T}_x} \\ z \in Z}} p_z^{\mathfrak{T}_x}$$

If $v = e.z$ for some node $z \in Z$ that is not a leaf in $\mathfrak{T}_x$ (e.g., node $1.v_1$), then we know by construction that the e -th successor w_e of z in $\mathfrak{T}_x$ is not contained in Z and is a leaf of $\mathfrak{T}_x$. Here, we get $p_{e.z}^{\Phi(\mathfrak{T}_x)} \stackrel{(\text{T-2})}{=} p_z^{\mathfrak{T}_x} \cdot \hat{p}_e = p_{w_e}^{\mathfrak{T}_x}$, and thus

$$\sum_{\substack{e.z \in \text{Leaf}^{\Phi(\mathfrak{T}_x)} \\ z \in Z, z \notin \text{Leaf}^{\mathfrak{T}_x}}} p_{e.z}^{\Phi(\mathfrak{T}_x)} = \sum_{\substack{w_e \in \text{Leaf}^{\mathfrak{T}_x} \\ w_e = (zE)_e, w_e \notin Z, z \in Z}} p_{e.z}^{\Phi(\mathfrak{T}_x)} = \sum_{\substack{w_e \in \text{Leaf}^{\mathfrak{T}_x} \\ w_e = (zE)_e, w_e \notin Z, z \in Z}} p_{w_e}^{\mathfrak{T}_x}$$

Here, $w_e = (zE)_e$ denotes that w_e is the e -th successor of z .

Finally, if v does not have the form $v = e.z$, then v is also a leaf in $\mathfrak{T}_x$ with $p_v^{\Phi(\mathfrak{T}_x)} = p_v^{\mathfrak{T}_x}$ and for both v and its predecessor u we have $v, u \notin Z$, and thus

$$\sum_{\substack{v \in \text{Leaf}^{\Phi(\mathfrak{T}_x)} \\ v \in \text{Leaf}^{\mathfrak{T}_x}}} p_v^{\Phi(\mathfrak{T}_x)} = \sum_{\substack{v \in \text{Leaf}^{\mathfrak{T}_x} \\ v \in uE, u \notin Z}} p_v^{\Phi(\mathfrak{T}_x)} = \sum_{\substack{v \in \text{Leaf}^{\mathfrak{T}_x} \\ v \in uE, u \notin Z}} p_v^{\mathfrak{T}_x}$$

Note that these cases cover each leaf of $\mathfrak{T}_x$ exactly once. These three equations imply:

$$\begin{aligned} & |\Phi(\mathfrak{T}_x)|_{\text{Leaf}} \\ &= \sum_{z \in \text{Leaf}^{\Phi(\mathfrak{T}_x)}} p_z^{\Phi(\mathfrak{T}_x)} \\ &= \sum_{\substack{e.z \in \text{Leaf}^{\Phi(\mathfrak{T}_x)} \\ z \in Z, z \in \text{Leaf}^{\mathfrak{T}_x}}} p_{e.z}^{\Phi(\mathfrak{T}_x)} + \sum_{\substack{e.z \in \text{Leaf}^{\Phi(\mathfrak{T}_x)} \\ z \in Z, z \notin \text{Leaf}^{\mathfrak{T}_x}}} p_{e.z}^{\Phi(\mathfrak{T}_x)} + \sum_{\substack{v \in \text{Leaf}^{\Phi(\mathfrak{T}_x)} \\ v \in \text{Leaf}^{\mathfrak{T}_x}}} p_v^{\Phi(\mathfrak{T}_x)} \\ &= \sum_{\substack{z \in \text{Leaf}^{\mathfrak{T}_x} \\ z \in Z}} p_z^{\mathfrak{T}_x} + \sum_{\substack{w_e \in \text{Leaf}^{\mathfrak{T}_x} \\ w_e = (zE)_e, w_e \notin Z, z \in Z}} p_{w_e}^{\mathfrak{T}_x} + \sum_{\substack{v \in \text{Leaf}^{\mathfrak{T}_x} \\ v \in uE, u \notin Z}} p_v^{\mathfrak{T}_x} \\ &= \sum_{z \in \text{Leaf}^{\mathfrak{T}_x}} p_z^{\mathfrak{T}_x} \\ &= |\mathfrak{T}_x|_{\text{Leaf}} \end{aligned}$$

1.3 We show that $\Phi(\mathfrak{T}_x)$ is indeed a valid RST.

Finally, we prove that $\Phi(\mathfrak{T}_x)$ is a valid $\mathcal{S}$ -RST. Here, we only need to show that $t_{e.z}^{\Phi(\mathfrak{T}_x)} \rightarrow_{\mathcal{S}} \{p_{e.w_1}^{\Phi(\mathfrak{T}_x)} : t_{e.w_1}^{\Phi(\mathfrak{T}_x)}, \dots, p_{e.w_m}^{\Phi(\mathfrak{T}_x)} : t_{e.w_m}^{\Phi(\mathfrak{T}_x)}\}$ for all $z \in Z$ as all the other edges and labelings were already present in $\mathfrak{T}_x$, which is a valid $\mathcal{S}$ -RST, and we have already seen that we have a valid innermost rewrite step at the new root $\hat{v}$.

Let $z \in Z$. In the following, we distinguish between two different cases for a rewrite step at a node z of $\mathfrak{T}_x$:

- (A) We rewrite at a position parallel to $\varphi_\tau(z)$ or $\varphi_\tau(z) = \top$.
- (B) We rewrite at a position above $\varphi_\tau(z)$. Note that this is the most interesting case, where we need the properties L (i.e., linearity) and NO.

Let $t_z^{\mathfrak{T}_x} \rightarrow_{\mathcal{S}} \{p_z^{\mathfrak{T}_x} \cdot p_1 : t_{w_1}^{\mathfrak{T}_x}, \dots, p_z^{\mathfrak{T}_x} \cdot p_m : t_{w_m}^{\mathfrak{T}_x}\}$, with a rule $\ell \rightarrow \{p_1 : r_1, \dots, p_m : r_m\} \in \mathcal{S}$, a substitution σ , and a position π such that $t_z^{\mathfrak{T}_x}|_\pi = \ell\sigma$. We have $t_{w_j}^{\mathfrak{T}_x} = t_z^{\mathfrak{T}_x}[r_j\sigma]_\pi$ for all $1 \leq j \leq m$.

(A) We start with the case where we have $\pi \perp \varphi_\tau(z)$ or $\varphi_\tau(z) = \top$. By (T-1), we get $t_{e.z}^{\Phi(\mathfrak{T}_x)} = t_z^{\mathfrak{T}_x}[\hat{r}_e\hat{\sigma}]_{\varphi_\tau(z)}$ if $\varphi_\tau(z) \in \mathbb{N}^*$ and $t_{e.z}^{\Phi(\mathfrak{T}_x)} = t_z^{\mathfrak{T}_x}$ if $\varphi_\tau(z) = \top$. In both cases, we can rewrite $t_{e.z}^{\Phi(\mathfrak{T}_x)}$ using the same rule, the same substitution, and the same position, as we have $t_{e.z}^{\Phi(\mathfrak{T}_x)}|_\pi = t_z^{\mathfrak{T}_x}[\hat{r}_e\hat{\sigma}]_{\varphi_\tau(z)}|_\pi = t_z^{\mathfrak{T}_x}|_\pi = \ell\sigma$ or directly $t_{e.z}^{\Phi(\mathfrak{T}_x)}|_\pi = t_z^{\mathfrak{T}_x}|_\pi = \ell\sigma$.

It remains to show that $t_{e.w_j}^{\Phi(\mathfrak{T}_x)} = t_{e.z}^{\Phi(\mathfrak{T}_x)}[r_j\sigma]_\pi$ for all $1 \leq j \leq m$, i.e., that the labeling we defined for $\Phi(\mathfrak{T}_x)$ corresponds to this rewrite step. Let $1 \leq j \leq m$. If $\varphi_\tau(z) \in \mathbb{N}^*$, then we have $t_{e.z}^{\Phi(\mathfrak{T}_x)}[r_j\sigma]_\pi = t_z^{\mathfrak{T}_x}[\hat{r}_e\hat{\sigma}]_{\varphi_\tau(z)}[r_j\sigma]_\pi \stackrel{\varphi_\tau(z) \perp \pi}{=} t_z^{\mathfrak{T}_x}[r_j\sigma]_\pi[\hat{r}_e\hat{\sigma}]_{\varphi_\tau(z)} = t_{w_j}^{\mathfrak{T}_x}[\hat{r}_e\hat{\sigma}]_{\varphi_\tau(z)} = t_{e.w_j}^{\Phi(\mathfrak{T}_x)}$. If $\varphi_\tau(z) = \top$, then $t_{e.z}^{\Phi(\mathfrak{T}_x)}[r_j\sigma]_\pi = t_z^{\mathfrak{T}_x}[r_j\sigma]_\pi = t_{w_j}^{\mathfrak{T}_x} = t_{e.w_j}^{\Phi(\mathfrak{T}_x)}$. Finally, note that the probabilities of the labeling are correct, as we are using the same rule with the same probabilities in both trees.

(B) If we have $\pi < \varphi_\tau(z)$, then there exists a $\chi \in \mathbb{N}^+$ such that $\pi \cdot \chi = \varphi_\tau(z)$. Since the rules of $\mathcal{S}$ are non-overlapping, the redex must be completely “inside” the used substitution σ , and we can find a position α_q of a variable q in ℓ and another position β such that $\chi = \alpha_q \cdot \beta$. Furthermore, since the rule is linear, q only occurs once in ℓ and at most once in r_j for all $1 \leq j \leq m$. Let ρ_q^j be the position of q in r_j if it exists. By (T-1), we get $t_z^{\mathfrak{T}_x}[\hat{r}_e\hat{\sigma}]_{\varphi_\tau(z)} = t_{e.z}^{\Phi(\mathfrak{T}_x)}$. We can rewrite $t_{e.z}^{\Phi(\mathfrak{T}_x)}$ using the same rule, the same substitution, and the same position, as we have $t_{e.z}^{\Phi(\mathfrak{T}_x)}|_\pi = t_z^{\mathfrak{T}_x}[\hat{r}_e\hat{\sigma}]_{\varphi_\tau(z)}|_\pi = t_z^{\mathfrak{T}_x}|_\pi[\hat{r}_e\hat{\sigma}]_\chi = \ell\sigma'$, where $\sigma'(q) = \sigma(q)[\hat{r}_e\hat{\sigma}]_\beta$ and $\sigma'(q') = \sigma(q')$ for all other variables $q' \neq q$.

It remains to show that $t_{e.w_j}^{\Phi(\mathfrak{T}_x)} = t_{e.z}^{\Phi(\mathfrak{T}_x)}[r_j\sigma']_\pi$ for all $1 \leq j \leq m$, i.e., that the labeling we defined for $\Phi(\mathfrak{T}_x)$ corresponds to this rewrite step. Let $1 \leq j \leq m$. If ρ_q^j exists, then we have $t_{e.z}^{\Phi(\mathfrak{T}_x)}[r_j\sigma']_\pi = t_z^{\mathfrak{T}_x}[\hat{r}_e\hat{\sigma}]_{\varphi_\tau(z)}[r_j\sigma']_\pi = t_z^{\mathfrak{T}_x}[\hat{r}_e\hat{\sigma}]_{\pi \cdot \alpha_q \cdot \beta}[r_j\sigma']_\pi = t_z^{\mathfrak{T}_x}[r_j\sigma]_\pi[\hat{r}_e\hat{\sigma}]_{\rho_q^j \cdot \beta} = t_{w_j}^{\mathfrak{T}_x}[\hat{r}_e\hat{\sigma}]_{\varphi_\tau(z)} = t_{e.w_j}^{\Phi(\mathfrak{T}_x)}$. Otherwise, we erase the precise redex and get $t_{e.z}^{\Phi(\mathfrak{T}_x)}[r_j\sigma]_\pi = t_z^{\mathfrak{T}_x}[\hat{r}_e\hat{\sigma}]_{\varphi_\tau(z)}[r_j\sigma']_\pi = t_z^{\mathfrak{T}_x}[\hat{r}_e\hat{\sigma}]_{\pi \cdot \alpha_q \cdot \beta}[r_j\sigma']_\pi = t_z^{\mathfrak{T}_x}[r_j\sigma]_\pi = t_{w_j}^{\mathfrak{T}_x} = t_{e.w_j}^{\Phi(\mathfrak{T}_x)}$. Finally, note that the probabilities of the labeling are correct, as we are using the same rule with the same probabilities in both trees. $\square$

Theorem 20 (From wAST to fAST). *If a PTRS $\mathcal{S}$ is non-overlapping, linear, and non-erasing, then*

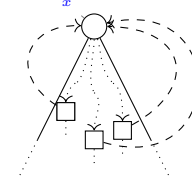
$$\mathcal{S} \text{ is fAST} \iff \mathcal{S} \text{ is wAST}$$

Proof. We only have to prove the non-trivial direction “ $\Leftarrow$ ”. Let $\mathcal{S}$ be a PTRS that is non-overlapping, left-linear, and non-erasing. Furthermore, assume for a contradiction that $\mathcal{S}$ is wAST but not fAST. This means that there exists an RST $\mathfrak{T}$ such that $|\mathfrak{T}|_{\text{Leaf}} = c$ for some $0 \leq c < 1$. Let $t \in \mathcal{T}(\Sigma, \mathcal{V})$ such that the root of $\mathfrak{T}$ is labeled with $(1 : t)$. Since $\mathcal{S}$ is wAST, there exists another RST $\tilde{\mathfrak{T}} = (\tilde{V}, \tilde{E}, \tilde{L})$ such that $|\tilde{\mathfrak{T}}|_{\text{Leaf}} = 1$ and the root of $\tilde{\mathfrak{T}}$ is also labeled with $(1 : t)$. Hence, in $\mathfrak{T}$ there is at least one rewrite step performed at some node x that is different to the rewrite step performed in $\tilde{\mathfrak{T}}$. The core steps of the proof are the same as for the proof of Thm. 17. We iteratively push the rewrite steps that would be performed in $\tilde{\mathfrak{T}}$ at node x to this node in $\mathfrak{T}$. Then, the limit of this construction is exactly $\tilde{\mathfrak{T}}$, which would mean that $|\mathfrak{T}|_{\text{Leaf}} \leq c < 1$, which is the desired contradiction. For this, we have to adjust the construction $\Phi(\circ)$. The rest of the proof remains completely the same.

1.1 Construction of $\Phi(\circ)$

Let $\mathfrak{T}_x = \mathfrak{T}[xE^*]$ be an $\mathcal{S}$ -RST that performs a rewrite step at position ζ at the root node x , i.e., $t_x^{\mathfrak{T}_x} \rightarrow_{\mathcal{S}, \zeta} \{p_{y_1}^{\mathfrak{T}_x} : t_{y_1}^{\mathfrak{T}_x}, \dots, p_{y_k}^{\mathfrak{T}_x} : t_{y_k}^{\mathfrak{T}_x}\}$ using the rule $\bar{\ell} \rightarrow \{\bar{p}_1 : \bar{r}_1, \dots, \bar{p}_k : \bar{r}_k\}$, and the substitution $\bar{\sigma}$ such that $t_x^{\mathfrak{T}_x}|_{\zeta} = \bar{\ell}\bar{\sigma}$. Then $t_{y_j}^{\mathfrak{T}_x} = t_x^{\mathfrak{T}_x}[\bar{r}_j\bar{\sigma}]_{\zeta}$ for all $1 \leq j \leq k$. Furthermore, assume that $\tilde{T}_x = \tilde{\mathfrak{T}}[xE^*]$ rewrites at position τ (using a rule $\hat{\ell} \rightarrow \{\hat{p}_1 : \hat{r}_1, \dots, \hat{p}_h : \hat{r}_h\} \in \mathcal{S}$ and the substitution $\hat{\sigma}$) with $\tau \neq \zeta$. (Note that if $\tau = \zeta$, then by non-overlappingness the rewrite step would be the same.) Instead of applying the rewrite step at position ζ at the root x we want to directly apply the rewrite step at position τ .

The construction creates a new RST $\Phi(\mathfrak{T}_x) = (V', E', L')$ such that $|\Phi(\mathfrak{T}_x)|_{\text{Leaf}} = |\mathfrak{T}_x|_{\text{Leaf}}$, and that directly performs a rewrite step at position τ at the root of the tree, by pushing it from the original nodes in the tree $\mathfrak{T}_x$ to the root. This push only results in the same termination probability due to our restriction that $\mathcal{S}$ is linear and non-erasing. We need the non-erasing property now, because τ may be above ζ , which was not possible in the proof of Thm. 17.



Again, since we are allowed to rewrite above τ in the original tree $\mathfrak{T}_x$, the actual position of the redex that was originally at position τ might change during the application of a rewrite step. Hence, we recursively define the position $\varphi_\tau(v)$ that contains precisely this redex for each node v in $\mathfrak{T}_x$ until we rewrite at this position. Compared to the proof of Thm. 17, since the rules in $\mathcal{S}$ are non-erasing, we only have $\varphi_\tau(v) \in \mathbb{N}^*$ or $\varphi_\tau(v) = \perp$. The option $\varphi_\tau(v) = \top$ is not possible anymore. Initially, we have $\varphi_\tau(x) = \tau$. Whenever we have defined $\varphi_\tau(v)$ for some node v , and we have $t_v^{\mathfrak{T}_x} \rightarrow_{\mathcal{S}} \{p_{w_1}^{\mathfrak{T}_x} : t_{w_1}^{\mathfrak{T}_x}, \dots, p_{w_m}^{\mathfrak{T}_x} : t_{w_m}^{\mathfrak{T}_x}\}$ for the direct successors $vE = \{w_1, \dots, w_m\}$, using the rule $\ell \rightarrow \{p_1 : r_1, \dots, p_m : r_m\}$, the substitution σ , and position π , we do the following: If $\varphi_\tau(v) = \pi$, then we set $\varphi_\tau(w_j) = \perp$

for all $1 \leq j \leq m$ to indicate that we have rewritten the redex. If we have $\varphi_\tau(v) \perp \pi$, meaning that the rewrite step takes place parallel to $\varphi_\tau(v)$, then we set $\varphi_\tau(w_j) = \varphi_\tau(v)$ for all $1 \leq j \leq m$, as the position did not change. If we have $\varphi_\tau(v) < \pi$, then we set $\varphi_\tau(w_j) = \varphi_\tau(v)$ for all $1 \leq j \leq m$ as well, as the position did not change either. If we have $\pi < \varphi_\tau(v)$, then there exists a $\chi \in \mathbb{N}^+$ such that $\pi \cdot \chi = \varphi_\tau(v)$. Since the rules of $\mathcal{S}$ are non-overlapping, the redex must be completely inside the used substitution σ , and we can find a position α_q of a variable q in ℓ and another position β such that $\chi = \alpha_q \cdot \beta$. Furthermore, since the rule is linear and non-erasing, q only occurs once in ℓ and once in r_j for all $1 \leq j \leq m$. Let ρ_q^j be the position of q in r_j . Here, we set $\varphi_\tau(w_j) = \rho_q^j \cdot \beta$. Finally, if $\varphi_\tau(v) = \perp$, then we set $\varphi_\tau(w_j) = \varphi_\tau(v) = \perp$ for all $1 \leq j \leq m$ as well.

Again, we now define the $\mathcal{S}$ -RST $\Phi(\mathfrak{T}_x)$ whose root is labeled with $(1 : t_x^{\mathfrak{T}_x})$ such that $|\Phi(\mathfrak{T}_x)|_{\text{Leaf}} = |\mathfrak{T}_x|_{\text{Leaf}}$, and that directly performs the rewrite step $t_x^{\mathfrak{T}_x} \rightarrow_{\mathcal{S}, \tau} \{\hat{p}_1 : t_{1.x}^{\Phi(\mathfrak{T}_x)}, \dots, \hat{p}_h : t_{h.x}^{\Phi(\mathfrak{T}_x)}\}$, with the rule $\hat{\ell} \rightarrow \{\hat{p}_1 : \hat{r}_1, \dots, \hat{p}_h : \hat{r}_h\} \in \mathcal{S}$, the substitution $\hat{\sigma}$, and the position τ , at the new root $\hat{v}$. Here, we have $t_x^{\mathfrak{T}_x}|_\tau = \hat{\ell}\hat{\sigma}$. Let Z be the set of all nodes v such that $\varphi_\tau(v) \neq \perp$. For each of these nodes $z \in Z$ and each $1 \leq e \leq h$, we create a new node $e.z \in V'$ with edges as in $\mathfrak{T}_x$ for the nodes in Z . Furthermore, we add the edges from the new root $\hat{v}$ to the nodes $e.x$ for all $1 \leq e \leq h$. Remember that x was the root in the tree $\mathfrak{T}_x$ and has to be contained in Z . We define the labeling of the nodes in $\Phi(\mathfrak{T}_x)$ as follows for all nodes z in Z :

$$\begin{aligned} \text{(T-1)} \quad t_{e.z}^{\Phi(\mathfrak{T}_x)} &= t_z^{\mathfrak{T}_x}[\hat{r}_e\delta]_{\varphi_\tau(z)} \text{ for the substitution } \delta \text{ such that } t_z^{\mathfrak{T}_x}|_{\varphi_\tau(z)} = \hat{\ell}\delta \\ \text{(T-2)} \quad p_{e.z}^{\Phi(\mathfrak{T}_x)} &= p_z^{\mathfrak{T}_x} \cdot \hat{p}_e \end{aligned}$$

Now, for a leaf $e.z' \in V'$ either $z' \in V$ is also a leaf or we rewrite at the position $\varphi_\tau(z')$ in node z' in $\mathfrak{T}_x$. If we rewrite $t_{z'}^{\mathfrak{T}_x} \xrightarrow{i}_{\mathcal{S}, \varphi_\tau(z')} \{p_{w'_1}^{\mathfrak{T}_x} : t_{w'_1}^{\mathfrak{T}_x}, \dots, p_{w'_h}^{\mathfrak{T}_x} : t_{w'_h}^{\mathfrak{T}_x}\}$, then we have $t_{w'_e}^{\mathfrak{T}_x} = t_{z'}^{\mathfrak{T}_x}[\hat{r}_e\delta]_{\varphi_\tau(z')} \stackrel{\text{(T-1)}}{=} t_{e.z'}^{\Phi(\mathfrak{T}_x)}$ for the substitution δ such that $t_{z'}^{\mathfrak{T}_x}|_{\varphi_\tau(z')} = \hat{\ell}\delta$ and $p_{w'_e}^{\mathfrak{T}_x} \stackrel{\text{(T-2)}}{=} p_{z'}^{\mathfrak{T}_x} \cdot \hat{p}_e = p_{e.z'}^{\Phi(\mathfrak{T}_x)}$. Thus, we can copy the rest of this subtree of $\mathfrak{T}_x$ in our newly generated tree $\Phi(\mathfrak{T}_x)$. So essentially, we just had to define how to construct $\Phi(\mathfrak{T}_x)$ for the part of the tree before we reach the nodes v with $\varphi_\tau(v) = \perp$ in $\mathfrak{T}_x$. As in the proof of [Thm. 17](#), we get $|\Phi(\mathfrak{T}_x)|_{\text{Leaf}} = |\mathfrak{T}_x|_{\text{Leaf}}$. We only need to show that $\Phi(\mathfrak{T}_x)$ is a valid $\mathcal{S}$ -RST, i.e., that $t_{e.z}^{\Phi(\mathfrak{T}_x)} \rightarrow_{\mathcal{S}} \{p_{e.w_1}^{\Phi(\mathfrak{T}_x)} : t_{e.w_1}^{\Phi(\mathfrak{T}_x)}, \dots, p_{e.w_m}^{\Phi(\mathfrak{T}_x)} : t_{e.w_m}^{\Phi(\mathfrak{T}_x)}\}$ for all $z \in Z$ as in the proof of [Thm. 17](#).

1.3 We show that $\Phi(\mathfrak{T}_x)$ is indeed a valid RST.

Let $z \in Z$. In the following, we distinguish between three different cases for a rewrite step at node z of $\mathfrak{T}_x$:

- (A) We rewrite at a position parallel to $\varphi_\tau(z)$.
- (B) We rewrite at a position above $\varphi_\tau(z)$.
- (C) We rewrite at a position below $\varphi_\tau(z)$.

The cases (A) and (B) are analogous as before but in Case (A) we cannot erase the redex anymore. We only need to look at the new Case (C).

(C) If we have $t_z^{\mathfrak{T}_x} \rightarrow_{\mathcal{S}} \{p_z^{\mathfrak{T}_x} \cdot p_1 : t_{w_1}^{\mathfrak{T}_x}, \dots, p_z^{\mathfrak{T}_x} \cdot p_m : t_{w_m}^{\mathfrak{T}_x}\}$, then there is a rule $\ell \rightarrow \{p_1 : r_1, \dots, p_m : r_m\} \in \mathcal{S}$, a substitution σ , and a position π with $t_z^{\mathfrak{T}_x}|_{\pi} = \ell\sigma$. Then $t_{w_j}^{\mathfrak{T}_x} = t_z^{\mathfrak{T}_x}[r_j\sigma]_{\pi}$ for all $1 \leq j \leq m$. Additionally, we assume that $\varphi_{\tau}(z) < \pi$, and thus there exists a $\chi \in \mathbb{N}^+$ such that $\pi = \varphi_{\tau}(z) \cdot \chi$. By (T-1), we get $t_z^{\mathfrak{T}_x}[\hat{r}_e\delta]_{\varphi_{\tau}(z)} = t_{e,z}^{\Phi(\mathfrak{T}_x)}$ for the substitution δ such that $t_z^{\mathfrak{T}_x}|_{\varphi_{\tau}(v)} = \hat{\ell}\delta$. Since the rules of $\mathcal{S}$ are non-overlapping, including the rule $\hat{\ell} \rightarrow \{\hat{p}_1 : \hat{r}_1, \dots, \hat{p}_h : \hat{r}_h\}$ that we use at the root of $\Phi(\mathfrak{T}_x)$, the redex for the current rewrite step must be completely “inside” the substitution δ , and we can find a variable position α_q of a variable q in $\hat{\ell}$ and another position β such that $\chi = \alpha_q \cdot \beta$. Furthermore, since the rule is also linear and non-erasing, q occurs exactly once in $\hat{\ell}$ and exactly once in $\hat{r}_j$ for all $1 \leq j \leq m$. Let ρ_q^j be the position of q in r_j . We can rewrite $t_{e,z}^{\Phi(\mathfrak{T}_x)}$ using the same rule, the same substitution, and the same position, as we have $t_{e,z}^{\Phi(\mathfrak{T}_x)}|_{\pi} = t_z^{\mathfrak{T}_x}[\hat{r}_e\delta]_{\varphi_{\tau}(z)}|_{\pi} = \hat{r}_e\delta|_{\chi} = \delta(z)|_{\beta} = t_z^{\mathfrak{T}_x}|_{\varphi_{\tau}(z)}|_{\alpha_q\beta} = t_z^{\mathfrak{T}_x}|_{\varphi_{\tau}(z) \cdot \alpha_q\beta} = t_z^{\mathfrak{T}_x}|_{\varphi_{\tau}(z) \cdot \chi} = t_z^{\mathfrak{T}_x}|_{\pi} = \ell\sigma$.

It remains to show that $t_{e,w_j}^{\Phi(\mathfrak{T}_x)} = t_{e,z}^{\Phi(\mathfrak{T}_x)}[r_j\sigma']_{\pi}$ for all $1 \leq j \leq m$, i.e., that the labeling we defined for $\Phi(\mathfrak{T}_x)$ corresponds to this rewrite step. Let $1 \leq j \leq m$. We have $t_{e,z}^{\Phi(\mathfrak{T}_x)}[r_j\sigma]_{\pi} = t_z^{\mathfrak{T}_x}[\hat{r}_e\delta]_{\varphi_{\tau}(z)}[r_j\sigma]_{\pi} = t_z^{\mathfrak{T}_x}[\hat{r}_e\delta']_{\varphi_{\tau}(z)} = t_{e,w_j}^{\Phi(\mathfrak{T}_x)}$ for the substitution δ' with $\delta'(q) = \delta(q)[r_j\sigma]_{\beta}$ and $\delta'(q') = \delta(q')$ for all other variables $q' \neq q$. With this new substitution, we get $t_{w_j}^{\mathfrak{T}_x}|_{\varphi_{\tau}(w_j)} = t_z^{\mathfrak{T}_x}[r_j\sigma]_{\varphi_{\tau}(z) \cdot \alpha_q \cdot \beta}|_{\varphi_{\tau}(z)} = t_z^{\mathfrak{T}_x}|_{\varphi_{\tau}(z)}[r_j\sigma]_{\alpha_q \cdot \beta} = (\hat{\ell}\delta)[r_j\sigma]_{\alpha_q \cdot \beta} = \hat{\ell}\delta'$. Finally, note that the probabilities of the labeling are correct, as we are using the same rule with the same probabilities in both trees. $\square$

Theorem 22 (From liAST to iAST). *If a PTRS $\mathcal{S}$ is non-overlapping, then*

$$\mathcal{S} \text{ is iAST} \iff \mathcal{S} \text{ is liAST}$$

Proof. The idea and the construction of this proof are completely analogous to the one of [Thm. 17](#). We iteratively move leftmost innermost rewrite steps to a higher position in the innermost RST. Hence, the resulting tree is a leftmost innermost $\mathcal{S}$ -RST.

The construction of $\Phi(\circ)$ is also analogous to the one in [Thm. 17](#). The only difference to the proof of [Thm. 17](#) is that the original tree $\mathfrak{T}_x$ is already an innermost $\mathcal{S}$ -RST. This means that during our construction only Case (A) can occur, as we cannot rewrite above a redex in an innermost $\mathcal{S}$ -RST. And for Case (A), we only need the property of being non-overlapping. $\square$

Next, we prove the new results from [Sect. 5](#) regarding AST.

Theorem 27 (From iAST to fAST (2)). *If a PTRS $\mathcal{S}$ is non-overlapping and right-linear, then*

$$\mathcal{S} \text{ is fAST} \iff \mathcal{S} \text{ is iAST w.r.t. } \rightarrow_{\mathcal{S}}$$

Proof. Once again, we use the idea and the construction of the proof for [Thm. 17](#). We iteratively move innermost rewrite steps to a higher position in the tree. But

in this case, for moving these innermost rewrite steps, we also allow using rewrite steps with $\xrightarrow{i}_{\mathcal{S}}$ instead of $\xrightarrow{\cdot}_{\mathcal{S}}$. Hence, the resulting tree is an innermost $\mathcal{S}$ -RST w.r.t. $\xrightarrow{i}_{\mathcal{S}}$. The construction of $\Phi(\circ)$ also remains similar to the one in [Thm. 17](#).

1.1 Construction of $\Phi(\circ)$

Let $\mathfrak{T}_x$ be an $\mathcal{S}$ -RST that performs a non-innermost rewrite step at the root node x , i.e., $t_x^{\mathfrak{T}_x} \rightarrow_{\mathcal{S}} \{p_{y_1}^{\mathfrak{T}_x} : t_{y_1}^{\mathfrak{T}_x}, \dots, p_{y_k}^{\mathfrak{T}_x} : t_{y_k}^{\mathfrak{T}_x}\}$ using the rule $\bar{\ell} \rightarrow \{\bar{p}_1 : \bar{r}_1, \dots, \bar{p}_k : \bar{r}_k\}$, the substitution $\bar{\sigma}$, and the position $\bar{\pi}$ such that $t_x^{\mathfrak{T}_x}|_{\bar{\pi}} = \bar{\ell}\bar{\sigma}$. Then we have $t_{y_j}^{\mathfrak{T}_x} = t_x^{\mathfrak{T}_x}[\bar{r}_j\bar{\sigma}]_{\bar{\pi}}$ for all $1 \leq j \leq k$. Instead of applying a non-innermost rewrite step at the root x we want to directly apply an innermost rewrite step. Let τ be the position of some innermost redex in t_x that is below $\bar{\pi}$.

The construction creates a new RST $\Phi(\mathfrak{T}_x) = (V', E', L')$ such that $|\Phi(\mathfrak{T}_x)|_{\text{Leaf}} = |\mathfrak{T}_x|_{\text{Leaf}}$. In $\Phi(\mathfrak{T}_x)$, directly at the root one performs the first rewrite step at position τ (which is an innermost rewrite step) and possibly simultaneously at some other innermost positions using the relation $\xrightarrow{i}_{\mathcal{S}}$, by pushing it from the original nodes in the tree $\mathfrak{T}_x$ to the root. This push only works due to our restriction that $\mathcal{S}$ is right-linear. We can now remove the left-linearity requirement by using $\xrightarrow{i}_{\mathcal{S}}$ at the root instead of just $\xrightarrow{\cdot}_{\mathcal{S}}$.

Again, we recursively define the position $\varphi_{\tau}(v)$ that contains precisely this redex for each node v in $\mathfrak{T}_x$ until we rewrite at this position. This works exactly as for [Thm. 17](#), again, due to our restriction of right-linearity. So to recapitulate, $\varphi_{\tau}(v)$ for a node v is either the position of the redex in t_v , $\top$ to indicate that the redex was erased, or $\perp$ to indicate that we have rewritten the redex. Also, the construction of the tree is similar to the one for [Thm. 17](#). Let Z be the set of all nodes z such that $\varphi_{\tau}(z) \neq \perp$. For each of these nodes $z \in Z$ and each $1 \leq e \leq h$, we create a new node $e.z \in V'$ with edges as in $\mathfrak{T}_x$ for the nodes in Z . Furthermore, we add the edges from the new root $\hat{v}$ to the nodes $e.x$ for all $1 \leq e \leq h$.

Next, we define the labeling for each of the new nodes. This is the part that differs compared to [Thm. 17](#). Since $\mathcal{S}$ is non-overlapping, the position τ must be completely “inside” the substitution $\bar{\sigma}$, and we can find a variable q of $\bar{\ell}$ such that $\tau = \bar{\pi}.\alpha_q.\beta$ for some variable position α_q of q in $\bar{\ell}$ and some other position β . However, since the left-hand side $\bar{\ell}$ may contain the same variable several times (as $\mathcal{S}$ does not have to be left-linear), there may exist multiple occurrences of q in $\bar{\ell}$. Let $\{\alpha_1, \dots, \alpha_n\}$ be the set of all positions α of $\bar{\ell}$ such that $\bar{\ell}|_{\alpha} = q$. The root $\hat{v}$ of $\Phi(\mathfrak{T}_x)$ is labeled with $(1 : t_x^{\mathfrak{T}_x})$, and we perform the rewrite step $t_x^{\mathfrak{T}_x} \xrightarrow{i}_{\mathcal{S}, \Gamma} \{\hat{p}_1 : t_{1.x}^{\Phi(\mathfrak{T}_x)}, \dots, \hat{p}_h : t_{h.x}^{\Phi(\mathfrak{T}_x)}\}$, with the rule $\bar{\ell} \rightarrow \{\hat{p}_1 : \hat{r}_1, \dots, \hat{p}_h : \hat{r}_h\} \in \mathcal{S}$, a substitution $\hat{\sigma}$, and the set of positions $\Gamma = \{\bar{\pi}.\alpha_1.\beta, \dots, \bar{\pi}.\alpha_n.\beta\}$, at the new root $\hat{v}$. Here, we have $t_x^{\Phi(\mathfrak{T}_x)}|_{\gamma} = \hat{\ell}\hat{\sigma}$ for all $\gamma \in \Gamma$ and get $t_{e.x}^{\Phi(\mathfrak{T}_x)} = t_x[\hat{r}_e\hat{\sigma}]_{\bar{\pi}.\alpha_1.\beta} \dots [\hat{r}_e\hat{\sigma}]_{\bar{\pi}.\alpha_n.\beta}$ for all $1 \leq e \leq h$.

After this rewrite step, we mirror the rewrite step from the root x at each node $e.x$ for all $1 \leq e \leq h$. This is possible, since we have $t_{e.x}^{\Phi(\mathfrak{T}_x)}|_{\bar{\pi}} = \bar{\ell}\bar{\sigma}'$ using the substitution $\bar{\sigma}'$ with $\bar{\sigma}'(q) = \bar{\sigma}(q)[\hat{r}_e\hat{\sigma}]_{\beta}$ and $\bar{\sigma}'(q') = \bar{\sigma}(q')$ for all other variables $q' \neq q$. Note that this is only possible because we have rewritten all occurrences of the same redex at a position $\gamma \in \Gamma$ simultaneously. Otherwise, we would not

be able to define $\bar{\sigma}'$ like this, because the matching might fail in some cases (see [Ex. 18](#)). With this definition of $\bar{\sigma}'$ we really have $\bar{\ell}\bar{\sigma}' = \bar{\ell}\bar{\sigma}[\hat{r}_e\hat{\sigma}]_{\alpha_1.\beta} \dots [\hat{r}_e\hat{\sigma}]_{\alpha_n.\beta} = t_x^{\mathfrak{T}_x}|_{\bar{\pi}}[\hat{r}_e\hat{\sigma}]_{\alpha_1.\beta} \dots [\hat{r}_e\hat{\sigma}]_{\alpha_n.\beta} = t_x^{\mathfrak{T}_x}[\hat{r}_e\hat{\sigma}]_{\bar{\pi}.\alpha_1.\beta} \dots [\hat{r}_e\hat{\sigma}]_{\bar{\pi}.\alpha_n.\beta}|_{\bar{\pi}} = t_{e.x}^{\Phi(\mathfrak{T}_x)}|_{\bar{\pi}}$. Since $\mathcal{S}$ is right-linear, we know that q can occur at most once in every $\bar{r}_j$. In this case, let ψ_j be the position of q in $\bar{r}_j$ for all $1 \leq j \leq k$. Then initially, we have $\varphi_{y_j}(\tau) = \bar{\pi}.\psi_j.\beta$ for all $1 \leq j \leq k$. Now, for all $1 \leq e \leq h$ and $1 \leq j \leq k$, if q exists in $\bar{r}_j$, then we get $t_{e.y_j}^{\Phi(\mathfrak{T}_x)} = t_{e.x}^{\Phi(\mathfrak{T}_x)}[\bar{r}_j\bar{\sigma}']_{\bar{\pi}} = t_{e.x}^{\Phi(\mathfrak{T}_x)}[\bar{r}_j\bar{\sigma}]_{\bar{\pi}}[\hat{r}_e\hat{\sigma}]_{\bar{\pi}.\psi_j.\beta} = t_{y_j}^{\mathfrak{T}_x}[\hat{r}_e\hat{\sigma}]_{\varphi_{y_j}(\tau)}$, and if q does not exist in $\bar{r}_j$, then we get $t_{e.y_j}^{\Phi(\mathfrak{T}_x)} = t_{e.x}^{\Phi(\mathfrak{T}_x)}[\bar{r}_j\bar{\sigma}']_{\bar{\pi}} = t_{e.x}^{\Phi(\mathfrak{T}_x)}[\bar{r}_j\bar{\sigma}]_{\bar{\pi}} = t_{y_j}^{\mathfrak{T}_x}$ and $\varphi_{y_j}(\tau) = \top$.

The rest of this construction is now completely analogous to the one for [Thm. 17](#). The labeling of the nodes $e.z$ for $z \in Z$ with $z \neq x$ is defined by:

$$\begin{aligned} \text{(T-1)} \quad t_{e.z}^{\Phi(\mathfrak{T}_x)} &= t_z^{\mathfrak{T}_x}[\hat{r}_e\hat{\sigma}]_{\varphi_\tau(z)} \text{ if } \varphi_\tau(z) \in \mathbb{N}^* \text{ and } t_{e.z}^{\Phi(\mathfrak{T}_x)} = t_z^{\mathfrak{T}_x} \text{ if } \varphi_\tau(z) = \top. \\ \text{(T-2)} \quad p_{e.z}^{\Phi(\mathfrak{T}_x)} &= p_z^{\mathfrak{T}_x} \cdot \hat{p}_e \end{aligned}$$

Now, for a leaf $e.z' \in V'$ either $z' \in V$ is also a leaf or we rewrite the innermost redex at position $\varphi_\tau(z')$ at node z' in $\mathfrak{T}_x$. If we rewrite $t_{z'}^{\mathfrak{T}_x} \xrightarrow{i}_{\mathcal{S}, \varphi_\tau(z')} \{\hat{p}_1 : t_{w'_1}^{\mathfrak{T}_x}, \dots, \hat{p}_h : t_{w'_h}^{\mathfrak{T}_x}\}$, then we have $t_{w'_e}^{\mathfrak{T}_x} = t_{z'}^{\mathfrak{T}_x}[\hat{r}_e\hat{\sigma}]_{\varphi_\tau(z')} \stackrel{(\text{T-1})}{=} t_{e.z'}^{\Phi(\mathfrak{T}_x)}$ and $p_{w'_e}^{\mathfrak{T}_x} = p_{z'}^{\mathfrak{T}_x} \cdot \hat{p}_e \stackrel{(\text{T-2})}{=} p_{e.z'}^{\Phi(\mathfrak{T}_x)}$. Thus, we can again copy the rest of this subtree of $\mathfrak{T}_x$ to our newly generated tree $\Phi(\mathfrak{T}_x)$. As for [Thm. 17](#), one can now prove that $|\Phi(\mathfrak{T}_x)|_{\text{Leaf}} = |\mathfrak{T}_x|_{\text{Leaf}}$, and that all other rewrite steps between a node $e.z$ and its successors are valid $\rightarrow_{\mathcal{S}}$ -rewrite steps. $\square$

Theorem 33 (From iAST to fAST (3)). *If a PTRS $\mathcal{S}$ is orthogonal and spare, then*

$$\mathcal{S} \text{ is fAST on basic terms} \iff \mathcal{S} \text{ is iAST on basic terms}$$

Proof. We only consider the non-trivial direction “ $\Leftarrow$ ”. Let $\mathcal{S}$ be iAST on basic terms. Then we have to show that all $\mathcal{S}$ -RSTs that start with $(1 : t)$ converge with probability 1 if t is a basic term. The proof for this part is completely analogous to the one of [Thm. 17](#). We iteratively move the innermost rewrite steps to a higher position using the construction $\Phi(\circ)$. Note that since $\mathcal{S}$ is spare and left-linear, in the construction of $\Phi(\circ)$ and in the proof of Case (B), if the innermost redex is below a redex $\ell\sigma$ that is reduced next via a rule $\ell \rightarrow \{p_1 : r_1, \dots, p_m : r_m\}$, then the innermost redex is completely “inside” the used substitution σ , and it corresponds to a variable q which occurs only once in ℓ and at most once in r_j for all $1 \leq j \leq m$, due to sparseness of $\mathcal{S}$. Hence, we can use the same construction as in [Thm. 17](#). $\square$

As mentioned in [Sect. 5](#), iAST on basic terms is equivalent to iAST on arbitrary terms. This is a consequence of the *starting lemma* (Lemma 53) from [\[30\]](#).

For [Lemma 35](#) we need some more auxiliary functions from [\[19\]](#) to decode a basic term over $\Sigma \cup \Sigma_{\mathcal{G}(\mathcal{S})}$ into the original term over Σ .

Definition 45 (Constructor Variant, Basic Variant, Decoded Variant).

Let $\mathcal{S}$ be a PTRS over the signature Σ . For a term $t \in \mathcal{T}(\Sigma, \mathcal{V})$, we define its constructor variant $\text{cv}(t)$ inductively as follows:

- $\text{cv}(x) = x$ for $x \in \mathcal{V}$
- $\text{cv}(f(t_1, \dots, t_n)) = f(\text{cv}(t_1), \dots, \text{cv}(t_n))$ for $f \in \Sigma_C$
- $\text{cv}(f(t_1, \dots, t_n)) = \text{cons}_f(\text{cv}(t_1), \dots, \text{cv}(t_n))$ for $f \in \Sigma_D$

For a term $t \in \mathcal{T}(\Sigma, \mathcal{V})$ with $t = f(t_1, \dots, t_n)$, we define its basic variant $\text{bv}(f(t_1, \dots, t_n)) = \text{enc}_f(\text{cv}(t_1), \dots, \text{cv}(t_n))$. For a term $t \in \mathcal{T}(\Sigma \cup \Sigma_{\mathcal{G}(\mathcal{S})}, \mathcal{V})$, we define its decoded variant $\text{dv}(t) \in \mathcal{T}(\Sigma, \mathcal{V})$ as follows:

- $\text{dv}(x) = x$ for $x \in \mathcal{V}$
- $\text{dv}(\text{argenc}(t)) = \text{dv}(t)$
- $\text{dv}(f(t_1, \dots, t_n)) = g(\text{dv}(t_1), \dots, \text{dv}(t_n))$ for $f \in \{g, \text{cons}_g, \text{enc}_g\}$ with $g \in \Sigma_D$
- $\text{dv}(f(t_1, \dots, t_n)) = f(\text{dv}(t_1), \dots, \text{dv}(t_n))$ for $f \in \Sigma_C$

The only difference to the auxiliary functions from [19] is that dv now also removes the **argenc** symbols from a term. This was handled differently in [19] in order to ensure $|\text{dv}(t)| = |t|$, but this is irrelevant for our proofs regarding AST. Here, $|t|$ denotes the size of a term $t \in \mathcal{T}(\Sigma, \mathcal{V})$ and is recursively defined by $|t| = 1$ if $t \in \mathcal{V}$ and $|t| = 1 + |t_1| + \dots + |t_n|$ if $t = f(t_1, \dots, t_n)$ with $f \in \Sigma \cup \Sigma_{\mathcal{G}(\mathcal{S})}$.

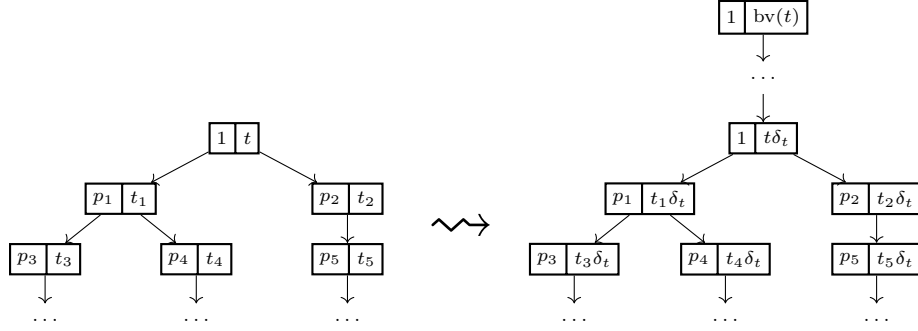
Lemma 35. A PTRS $\mathcal{S}$ is fAST iff $\mathcal{S} \cup \mathcal{G}(\mathcal{S})$ is fAST on basic terms.

Proof.

“ $\Leftarrow$ ”

Assume that $\mathcal{S}$ is not fAST. Then there exists an $\mathcal{S}$ -RST $\mathfrak{T} = (V, E, L)$ that converges with probability < 1 whose root is labeled with $(1 : t)$ for some $t \in \mathcal{T}(\Sigma, \mathcal{V})$. We construct an $(\mathcal{S} \cup \mathcal{G}(\mathcal{S}))$ -RST $\mathfrak{T}' = (V', E', L')$ whose root is labeled with $(1 : \text{bv}(t))$ and with $|\mathfrak{T}'|_{\text{Leaf}} = |\mathfrak{T}|_{\text{Leaf}}$. Therefore, there exists an $(\mathcal{S} \cup \mathcal{G}(\mathcal{S}))$ -RST $\mathfrak{T}'$ that converges with probability < 1 and starts with the basic term $\text{bv}(t)$, i.e., $\mathcal{S} \cup \mathcal{G}(\mathcal{S})$ is not fAST on basic terms.

In [19] it was shown that $\{1 : \text{bv}(t)\} \xrightarrow{\text{enc}}^*_{\mathcal{G}(\mathcal{S})} \{1 : t\delta_t\}$ ([19] did not consider PTRSs but since all the probabilities in the rules of $\mathcal{G}(\mathcal{S})$ are trivial, the proof in [19] directly translates to the probabilistic setting). Here, for any term $t \in \mathcal{T}(\Sigma, \mathcal{V})$ the substitution δ_t is defined by $\delta_t(x) = \text{argenc}(x)$ if $x \in \mathcal{V}(t)$ and $\delta_t(x) = x$ otherwise. The $(\mathcal{S} \cup \mathcal{G}(\mathcal{S}))$ -RST $\mathfrak{T}'$ first performs these innermost rewrite steps to get from $\text{bv}(t)$ to $t\delta_t$, and then we can mirror the rewrite steps from $\mathfrak{T}$. To be precise, we use the same underlying tree structure and an adjusted labeling such that $p_x^{\mathfrak{T}} = p_x^{\mathfrak{T}'}$ and $t_x^{\mathfrak{T}} \delta_t = t_x^{\mathfrak{T}'}$ for all $x \in V$. Since the tree structure and the probabilities are the same, we then get $|\mathfrak{T}|_{\text{Leaf}} = |\mathfrak{T}'|_{\text{Leaf}}$. To be precise, the set of leaves in $\mathfrak{T}$ is equal to the set of leaves in $\mathfrak{T}'$, and they have the same probabilities. Since $|\mathfrak{T}|_{\text{Leaf}} < 1$, we thus have $|\mathfrak{T}'|_{\text{Leaf}} < 1$.



“ $\Rightarrow$ ”

Let $\mathcal{S} \cup \mathcal{G}(\mathcal{S})$ be not FAST on basic terms. Then there exists an $(\mathcal{S} \cup \mathcal{G}(\mathcal{S}))$ -RST $\mathfrak{T} = (V, E, L)$ that converges with probability < 1 whose root is labeled with $(1 : t)$ for some term $t \in \mathcal{T}(\Sigma \cup \Sigma_{\mathcal{G}(\mathcal{S})}, \mathcal{V})$. We construct an $\mathcal{S}$ -RST $\mathfrak{T}' = (V', E', L')$ inductively such that for all leaves x of $\mathfrak{T}'$ during the construction there exists a node $\varphi(x)$ of $\mathfrak{T}$ such that

$$t_x^{\mathfrak{T}'} = \text{dv}(t_{\varphi(x)}^{\mathfrak{T}}) \quad \text{and} \quad p_x^{\mathfrak{T}'} = p_{\varphi(x)}^{\mathfrak{T}}. \quad (1)$$

Here, φ is injective, i.e., every leaf x of $\mathfrak{T}'$ is mapped to a (unique) node $\varphi(x)$ of $\mathfrak{T}$. Furthermore, after this construction, if x is still a leaf in $\mathfrak{T}'$, then $\varphi(x)$ is also a leaf in $\mathfrak{T}$. Hence, we get $|\mathfrak{T}|_{\text{Leaf}} = \sum_{x \in \text{Leaf}^{\mathfrak{T}}} p_x^{\mathfrak{T}} \geq \sum_{x \in \text{Leaf}^{\mathfrak{T}'}} p_x^{\mathfrak{T}'} = |\mathfrak{T}'|_{\text{Leaf}}$. Since $|\mathfrak{T}|_{\text{Leaf}} < 1$, we thus have $|\mathfrak{T}'|_{\text{Leaf}} < 1$. This means that $\mathcal{S}$ is also not FAST.

We label the root of $\mathfrak{T}'$ with $(1 : \text{dv}(t))$. By letting φ map the root of $\mathfrak{T}'$ to the root of $\mathfrak{T}$, the claim (1) is clearly satisfied. As long as there is still a node x in $\mathfrak{T}'$ such that $\varphi(x)$ is not a leaf in $\mathfrak{T}$, we do the following. If we perform a rewrite step with $\mathcal{G}(\mathcal{S})$ at node $\varphi(x)$, i.e., $t_{\varphi(x)}^{\mathfrak{T}} \rightarrow_{\mathcal{G}(\mathcal{S})} \{1 : t_y^{\mathfrak{T}}\}$ for the only successor y of $\varphi(x)$, then we have $\text{dv}(t_{\varphi(x)}^{\mathfrak{T}}) = \text{dv}(t_y^{\mathfrak{T}})$, i.e., we do nothing in $\mathfrak{T}'$ but simply change the definition of φ such that $\varphi(x)$ is now y . To see why $\text{dv}(t_{\varphi(x)}^{\mathfrak{T}}) = \text{dv}(t_y^{\mathfrak{T}})$ holds, note that we have $\text{dv}(\ell) = \text{dv}(r)$ for all rules $\ell \rightarrow \{1 : r\} \in \mathcal{G}(\mathcal{S})$, since $\text{dv}(\text{enc}_f(x_1, \dots, x_n)) = f(x_1, \dots, x_n) = \text{dv}(f(\text{argenc}(x_1), \dots, \text{argenc}(x_n)))$, and analogously $\text{dv}(\text{argenc}(\text{cons}_f(x_1, \dots, x_n))) = f(x_1, \dots, x_n) = \text{dv}(f(\text{argenc}(x_1), \dots, \text{argenc}(x_n)))$. This can then be lifted to arbitrary rewrite steps.

Otherwise, let $\varphi(x)E = \{y_1, \dots, y_k\}$ be the set of successors of $\varphi(x)$ in $\mathfrak{T}$, and we have $t_{\varphi(x)}^{\mathfrak{T}} \rightarrow_{\mathcal{S}} \{\frac{p_{y_1}^{\mathfrak{T}}}{p_{\varphi(x)}^{\mathfrak{T}}} : t_{y_1}^{\mathfrak{T}}, \dots, \frac{p_{y_k}^{\mathfrak{T}}}{p_{\varphi(x)}^{\mathfrak{T}}} : t_{y_k}^{\mathfrak{T}}\}$. Since $t_x^{\mathfrak{T}'} = \text{dv}(t_{\varphi(x)}^{\mathfrak{T}})$ and $p_x^{\mathfrak{T}'} = p_{\varphi(x)}^{\mathfrak{T}}$ by the induction hypothesis, then we also have $t_x^{\mathfrak{T}'} \rightarrow_{\mathcal{S}} \{\frac{p_{y_1}^{\mathfrak{T}'}}{p_x^{\mathfrak{T}'}} : t_{y_1}^{\mathfrak{T}'}, \dots, \frac{p_{y_k}^{\mathfrak{T}'}}{p_x^{\mathfrak{T}'}} : t_{y_k}^{\mathfrak{T}'}\}$ for terms $t_{y_j}^{\mathfrak{T}'}$ with $t_{y_j}^{\mathfrak{T}'} = \text{dv}(t_{y_j}^{\mathfrak{T}})$ and for $p_{y_j}^{\mathfrak{T}'} = p_{y_j}^{\mathfrak{T}}$. Thus, when defining $\varphi(y_j) = y_j$ for all $1 \leq j \leq k$, (1) is satisfied for the new leaves $y_1, \dots, y_k$ of $\mathfrak{T}'$. $\square$

Theorem 38 (From iAST to fAST (4)). *If a PTRS $\mathcal{S}$ is non-overlapping and spare, then*

$$\mathcal{S} \text{ is fAST on basic terms} \iff \mathcal{S} \text{ is iAST w.r.t. } \rightarrow_{\mathcal{S}} \text{ on basic terms}$$

Proof. The proof is completely analogous to the one of [Thm. 27](#). We iteratively move the innermost rewrite steps to a higher position using the construction $\Phi(\circ)$. Note that since $\mathcal{S}$ is spare, in the construction of $\Phi(\circ)$ and in the proof of Case (B) (i.e., rewriting at a position above $\varphi_{\tau}(z)$, see the proof of [Thm. 17](#)) and the construction of the label at the nodes $e.y_1, \dots, e.y_k$ (the second rewrite step after the root), if the innermost redex is below a redex $\ell\sigma$ that is reduced next via a rule $\ell \rightarrow \{p_1 : r_1, \dots, p_m : r_m\}$, then the innermost redex is completely “inside” the used substitution σ , and it corresponds to a variable q which occurs only once in ℓ and at most once in r_j for all $1 \leq j \leq m$, due to sparseness of $\mathcal{S}$. Hence, we can use the same construction as in [Thm. 27](#). $\square$

B PAST & SAST

In this section, we prove all our new contributions and observations regarding PAST. But first, we define a stricter notion than PAST, namely *strong almost-sure termination* [\[4, 18\]](#), which requires a finite bound on the expected derivation lengths of all rewrite sequences with the same starting term in case of non-determinism. Thus, for a term t , the *expected derivation height* $\text{edh}(t)$ considers all possible rewrite sequences that start with $\{1 : t\}$.

Definition 46 (edh, SAST). *Let $\mathcal{S}$ be a PTRS, $t \in \mathcal{T}(\Sigma, \mathcal{V})$ be a term, and let $s \in \{\mathbf{f}, \mathbf{i}, \mathbf{li}\}$ be a rewrite strategy where $\mathbf{f}$, $\mathbf{i}$, and $\mathbf{li}$ stand for full, innermost, and leftmost-innermost rewriting, respectively. By $\text{edh}_s(t) = \sup\{\text{edl}(\vec{\mu}) \mid \vec{\mu} \text{ is a } \xrightarrow{\mathcal{S}}\text{-rewrite sequence with } \mu_0 = \{1 : t\}\}$ we denote the expected derivation height of t . $\mathcal{S}$ is strongly almost-surely terminating (SAST) (innermost SAST (iSAST) / leftmost-innermost SAST (liSAST)) if $\text{edh}_{\mathbf{f}}(t)$ ($\text{edh}_{\mathbf{i}}(t)$ / $\text{edh}_{\mathbf{li}}(t)$) is finite for all $t \in \mathcal{T}(\Sigma, \mathcal{V})$.*

Note that in contrast to wAST and wPAST, we did not define any notion of weak SAST. The reason is that the definition of weak forms of termination always only requires the *existence* of some suitable rewrite sequence. But the definition of SAST imposes a requirement on *all* rewrite sequences. Thus, it is not clear how to obtain a useful definition for “weak SAST” that differs from wPAST.

It is well known that SAST implies PAST [\[4\]](#). However, even for PTRSs with *finitely* many rules, SAST and PAST are *not* equivalent. A counterexample is a PTRS with the rules $\mathbf{f}(x) \rightarrow \{1/2 : \mathbf{f}(\mathbf{s}(x)), 1/2 : \mathcal{O}\}$, $\mathbf{f}(x) \rightarrow \{1 : \mathbf{g}(x)\}$, and terminating non-probabilistic rules for $\mathbf{g}$ such that $\mathbf{g}(\mathbf{s}^n(\mathcal{O}))$ needs least $4^{n+1} - 1$ steps. The PTRS is PAST. To see this, consider the possible derivations of $\mathbf{f}(\mathbf{s}^n(\mathcal{O}))$. If we never use the rule from $\mathbf{f}$ to $\mathbf{g}$, then the expected derivation length is $\frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 + \dots = \sum_{i=1}^{\infty} \frac{i}{2^i} = 2$ and if we never use the probabilistic

f-rule, then the derivation length is 4^{n+1} . Otherwise, if we use the probabilistic f-rule in the first step and the rule from f to g in the second step, then the expected derivation length is $\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 4^{n+1} = (\sum_{i=1}^1 \frac{i}{2^i}) + \frac{4^{n+1}}{2^1}$. If we use the probabilistic f-rule in the first two steps and the rule from f to g in the third step, then the expected derivation length is $\frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{4} \cdot 4^{n+2} = (\sum_{i=1}^2 \frac{i}{2^i}) + \frac{4^{n+2}}{2^2}$. So to summarize, for any $k \geq 1$, if we use the probabilistic f-rule in the first k steps and the rule from f to g in the $(k+1)$ -th step, then the expected derivation length is $(\sum_{i=1}^k \frac{i}{2^i}) + \frac{4^{n+k}}{2^k} = (\sum_{i=1}^k \frac{i}{2^i}) + 2^{2 \cdot n+k}$.

Thus, for every rewrite sequence, the expected derivation length is finite. However, since k can be any number, for any $k \geq 1$, there is a rewrite sequence starting with $f(s^n(\mathcal{O}))$ whose expected derivation length is $(\sum_{i=1}^k \frac{i}{2^i}) + \frac{4^{n+k}}{2^k} = (\sum_{i=1}^k \frac{i}{2^i}) + 2^{2 \cdot n+k} \geq 2^{2 \cdot n+k}$. In other words, although every rewrite sequence starting with $f(s^n(\mathcal{O}))$ has finite expected derivation length, the supremum over the lengths for all such rewrite sequences is infinite. Hence, this PTRS is not SAST.

Now we present our new results when regarding PAST and SAST instead of AST. So now we investigate restricted forms like iPAST, liPAST, etc. (and similar for SAST). While the theorems are the same as for AST, except for [Thm. 33](#) and [38](#), they need a slight change in their proofs as we now need to argue about expected derivation length/height and not just about the termination probability.

We first show that PAST and SAST can also be formulated in terms of rewrite sequence trees.

Theorem 47 (PAST via RSTs). *Let $\mathcal{S}$ be a PTRS and $\mathfrak{T}$ be an $\mathcal{S}$ -RST. By*

$$\text{edl}(\mathfrak{T}) = \sum_{x \in V \setminus \text{Leaf}} p_x^{\mathfrak{T}} = \sum_{n=0}^{\infty} \sum_{\substack{x \in V \setminus \text{Leaf} \\ d^{\mathfrak{T}}(x)=n}} p_x^{\mathfrak{T}}$$

we define the expected derivation length of $\mathfrak{T}$, where $d^{\mathfrak{T}}(x)$ again denotes the depth of a node x in the tree $\mathfrak{T}$. Here, the root has depth 0. Then $\mathcal{S}$ is PAST (iPAST / liPAST) iff $\text{edl}(\mathfrak{T})$ is finite for every (innermost / leftmost-innermost) $\mathcal{S}$ -RST $\mathfrak{T}$. Similarly, $\mathcal{S}$ is wPAST iff for every term t there exists an $\mathcal{S}$ -RST $\mathfrak{T}$ whose root is labeled $(1 : t)$ such that $\text{edl}(\mathfrak{T})$ is finite.

Theorem 48 (SAST via RSTs). *Let $\mathcal{S}$ be a PTRS, $\mathfrak{T}$ be an $\mathcal{S}$ -RST, and let $s \in \{f, i, li\}$ be a rewrite strategy. Then $\mathcal{S}$ is SAST (iSAST / liSAST) iff $\sup\{\text{edl}(\mathfrak{T}) \mid \mathfrak{T} \text{ is an } \mathcal{S}\text{-RST w.r.t. the strategy } s \text{ whose root is labeled with } (1 : t)\}$ is finite for all $t \in \mathcal{T}(\Sigma, \mathcal{V})$ where $s = f$ ($s = i$ / $s = li$).*

Both of these theorems are again easy to prove, similar to the characterization of AST with RSTs in [Cor. 43](#). First note that every infinite $\Rightarrow_{\mathcal{S}}$ -rewrite sequence $\vec{\mu} = (\mu_n)_{n \in \mathbb{N}}$ that begins with a single start term (i.e., $\mu_0 = \{1 : t\}$) can be represented by an infinite $\mathcal{S}$ -RST $\mathfrak{T}$ that is fully evaluated such that $|\mathfrak{T}|_{\text{Leaf}} = 1$ and $\text{edl}(\vec{\mu}) = \sum_{n=0}^{\infty} (1 - |\mu_n|_s) = \sum_{n=0}^{\infty} \sum_{\substack{x \in V \setminus \text{Leaf} \\ d^{\mathfrak{T}}(x)=n}} p_x^{\mathfrak{T}} = \text{edl}(\mathfrak{T})$ and vice versa,

since whenever we reach a normal form after n steps, this is both a normal form in μ_n and a leaf at depth n of the RST with the same probability, and otherwise, if we have a term t in $\text{Supp}(\mu_n)$ that is not in normal form, then there exists a (unique) node in $\mathfrak{T}$ at depth n with the same probability that is also not a leaf. So $\mathcal{S}$ is PAST for rewrite sequences that begin with a single start term iff all fully evaluated $\mathcal{S}$ -RSTs have finite expected derivation lengths.

To show that this is equivalent to PAST, it remains to show that it suffices to only regard $\Rightarrow$ -rewrite sequences that begin with a single start term. Hence, the following lemma adapts [Lemma 42](#) from AST to PAST.

Lemma 49 (Single Start Terms Suffice for PAST). *Suppose that there exists an infinite $\Rightarrow_{\mathcal{S}}$ -rewrite sequence $(\mu_n)_{n \in \mathbb{N}}$ with infinite expected derivation length. Then there is also an infinite $\Rightarrow_{\mathcal{S}}$ -rewrite sequence $(\mu'_n)_{n \in \mathbb{N}}$ with a single start term, i.e., $\mu'_0 = \{1 : t\}$, that has infinite expected derivation length.*

Proof. Let $\vec{\mu} = (\mu_n)_{n \in \mathbb{N}}$ be an $\Rightarrow_{\mathcal{S}}$ -rewrite sequence that has infinite expected derivation length. Suppose that we have $\mu_0 = \{p_1 : t_1, \dots, p_k : t_k\}$. Let $\vec{\mu}^j = (\mu_n^j)_{n \in \mathbb{N}}$ with $\mu_0^j = \{1 : t_j\}$ denote the infinite $\Rightarrow_{\mathcal{S}}$ -rewrite sequence that uses the same rules as $(\mu_n)_{n \in \mathbb{N}}$ does on the term t_j for every $1 \leq j \leq k$. Assume for a contradiction that for every $1 \leq j \leq k$ the $\Rightarrow_{\mathcal{S}}$ -rewrite sequence $(\mu_n^j)_{n \in \mathbb{N}}$ has finite expected derivation length. Then we would have

$$\begin{aligned} \text{edl}(\vec{\mu}) &= \sum_{n=0}^{\infty} (1 - |\mu_n|_{\mathcal{S}}) \\ &= \sum_{n=0}^{\infty} \sum_{1 \leq j \leq k} p_j \cdot (1 - |\mu_n^j|_{\mathcal{S}}) \\ &= \sum_{1 \leq j \leq k} p_j \cdot \sum_{n=0}^{\infty} (1 - |\mu_n^j|_{\mathcal{S}}) \\ &= \sum_{1 \leq j \leq k} p_j \cdot \text{edl}(\vec{\mu}^j) \\ &< \infty \end{aligned}$$

which is a contradiction to our assumption that we have $\text{edl}(\vec{\mu}) = \infty$. Therefore, we have at least one $1 \leq j \leq k$ such that $\text{edl}(\vec{\mu}^j) = \infty$. $\square$

[Thm. 48](#) is then a direct consequence of [Thm. 47](#). Note that we do not need a version of [Lemma 49](#) for SAST, since SAST already considers single start terms only.

Now we show that [Thm. 17](#) also holds for PAST and SAST.

Theorem 17 (From iPAST/iSAST to fPAST/fSAST (1)). *If a PTRS $\mathcal{S}$ is orthogonal and right-linear (i.e., non-overlapping and linear), then:*

$$\begin{aligned} \mathcal{S} \text{ is fPAST} &\iff \mathcal{S} \text{ is iPAST} \\ \mathcal{S} \text{ is fSAST} &\iff \mathcal{S} \text{ is iSAST} \end{aligned}$$

Proof. Again, we only prove the non-trivial direction “ $\Leftarrow$ ”.

1. From iPAST to fPAST:

Let $\mathcal{S}$ be a PTRS that is non-overlapping and linear. Furthermore, assume that $\mathcal{S}$ is not fPAST. This means that there exists an RST $\mathfrak{T}$ such that $\text{edl}(\mathfrak{T}) = \infty$. We

create a new *innermost* RST $\mathfrak{T}^{(\infty)}$, using the same construction as in the proof of [Thm. 17](#) for AST. If we can show that this construction yields $\text{edl}(\mathfrak{T}^{(\infty)}) \geq \text{edl}(\mathfrak{T}) = \infty$, then this shows that $\mathcal{S}$ is also not iPAST.

Let x again be a topmost node in $\mathfrak{T}$ where a non-innermost step is performed and let $\mathfrak{T}^{(1)}, \mathfrak{T}^{(2)}, \dots$ be defined via the transformation Φ as in the proof of [Thm. 17](#) for AST. So $\mathfrak{T}^{(1)}$ results from replacing the subtree $\mathfrak{T}_x$ by $\Phi(\mathfrak{T}) = (V', E', L')$. Our goal is to show that in the construction of the trees $\mathfrak{T}^{(1)}, \mathfrak{T}^{(2)}, \dots$, every leaf v of $\mathfrak{T}$ is turned into leaves of $\mathfrak{T}^{(1)}, \mathfrak{T}^{(2)}, \dots$ whose probabilities sum up to $p^{\mathfrak{T}}(v)$, and whose depths are greater or equal than the original depth $d^{\mathfrak{T}}(v)$ of v in $\mathfrak{T}$. This then implies $\infty = \text{edl}(\mathfrak{T}) \leq \text{edl}(\mathfrak{T}^{(1)}) \leq \text{edl}(\mathfrak{T}^{(2)}) \leq \dots$. From this observation, we then conclude $\infty = \text{edl}(\mathfrak{T}) \leq \text{edl}(\mathfrak{T}^{(\infty)})$.

1.1 We prove $\infty = \text{edl}(\mathfrak{T}) \leq \text{edl}(\mathfrak{T}^{(1)}) \leq \text{edl}(\mathfrak{T}^{(2)}) \leq \dots$

We start by considering how the leaves of $\mathfrak{T}$ correspond to the leaves of $\mathfrak{T}^{(1)}$. Let v be a leaf in $\mathfrak{T}$. If $v \notin xE^*$, then v is also a leaf in $\mathfrak{T}^{(1)}$, labeled with the same probability, and at the same depth. Otherwise, $v \in xE^*$, which means that either $v \in Z$, $v \notin Z$ but its predecessor is in Z , or neither v nor its predecessor is in Z .

If $v \in Z$ (like, e.g., the node v_2 in the example of Step 1.1 in the proof of [Thm. 17](#) for AST), then also $e.v$ must be a leaf in $\Phi(\mathfrak{T}_x)$ for every $1 \leq e \leq h$. Here, we get $\sum_{1 \leq e \leq h} p_{e.v}^{\Phi(\mathfrak{T}_x)} = p_v^{\mathfrak{T}_x}$ as in the proof of [Thm. 17](#) for AST. In addition, we also know that if $e.v$ is at depth m in $\Phi(\mathfrak{T}_x)$, then v is at depth $m - 1$ in $\mathfrak{T}_x$.

If $v \notin Z$ but v is the e -th successor of a node $z \in Z$ (like the node v_8 in the example in the proof of [Thm. 17](#) for AST), then $e.z$ is a leaf in $\Phi(\mathfrak{T}_x)$. Here, we get $p_{e.z}^{\Phi(\mathfrak{T}_x)} = p_v^{\mathfrak{T}_x}$ as in the proof of [Thm. 17](#) for AST. In addition, we also know that if $e.z$ is at depth m in $\Phi(\mathfrak{T}_x)$, then v is at depth m in $\mathfrak{T}_x$.

Finally, if neither v nor its predecessor is in Z (like the node v_9 in the example in the proof of [Thm. 17](#) for AST), then v is also a leaf in $\Phi(\mathfrak{T}_x)$ with $p_v^{\Phi(\mathfrak{T}_x)} = p_v^{\mathfrak{T}_x}$ and at the same depth.

These cases cover each leaf of $\mathfrak{T}^{(1)}$ exactly once and all leaves in $\Phi(\mathfrak{T}_x)$ are at a depth greater or equal than the corresponding leaves in $\mathfrak{T}_x$, implying that for each leaf u in $\mathfrak{T}$ we can find a set of leaves Ω_u^1 in $\mathfrak{T}^{(1)}$ such that $d^{\mathfrak{T}^{(1)}}(u) \leq d^{\mathfrak{T}^{(1)}}(w)$ for each $w \in \Omega_u^1$, $\sum_{w \in \Omega_u^1} p_w^{\mathfrak{T}^{(1)}} = p_u^{\mathfrak{T}}$, and $\text{Leaf}^{\mathfrak{T}^{(1)}} = \biguplus_{u \in \text{Leaf}^{\mathfrak{T}}} \Omega_u^1$.

We can now use the same observation for each leaf w in the tree $\mathfrak{T}^{(1)}$ in order to get a new set Ξ_w^2 of leaves in $\mathfrak{T}^{(2)}$ and define $\Omega_u^2 = \bigcup_{w \in \Omega_u^1} \Xi_w^2$. Again, we get $d^{\mathfrak{T}}(u) \leq d^{\mathfrak{T}^{(2)}}(w)$ for each $w \in \Omega_u^2$, $\sum_{w \in \Omega_u^2} p_w^{\mathfrak{T}^{(2)}} = p_u^{\mathfrak{T}}$, and $\text{Leaf}^{\mathfrak{T}^{(2)}} = \biguplus_{u \in \text{Leaf}^{\mathfrak{T}}} \Omega_u^2$. We now do this for each $i \in \mathbb{N}$ to define the set Ω_u^i for each leaf u in the original tree $\mathfrak{T}$. Overall, this implies $\infty = \text{edl}(\mathfrak{T}) \leq \text{edl}(\mathfrak{T}^{(1)}) \leq \text{edl}(\mathfrak{T}^{(2)}) \leq \dots$.

1.2 We prove $\infty = \text{edl}(\mathfrak{T}) \leq \text{edl}(\mathfrak{T}^{(\infty)})$.

From the construction of the Ω_u^i above, in the end, we get

$$\text{Leaf}^{\mathfrak{T}^{(\infty)}} = \biguplus_{u \in \text{Leaf}^{\mathfrak{T}}} \limsup_{i \rightarrow \infty} \Omega_u^i,$$

where $\limsup_{i \rightarrow \infty} \Omega_u^i = \bigcap_{i' \in \mathbb{N}} \bigcup_{i > i'} \Omega_u^i = \{w \mid w \text{ is contained in infinitely many } \Omega_u^i\}$. To see this, let $v \in \text{Leaf}^{\mathfrak{T}^{(\infty)}}$. Remember that for every depth H of the tree,

there exists an m_H such that $\mathfrak{T}^{(\infty)}$ and $\mathfrak{T}^{(i)}$ are the same trees up to depth H for all $i \geq m_H$. This means that the node v must be contained in all trees $\mathfrak{T}^{(m)}$ with $m \geq m_{d^{\mathfrak{T}^{(\infty)}}(v)}$, i.e., it is contained in $\biguplus_{u \in \text{Leaf}^{\mathfrak{T}}} \limsup_{i \rightarrow \infty} \Omega_u^i$. The other direction of the equality follows in the same manner.

Since the probabilities of the leaves in $\mathfrak{T}^{(i)}$ always add up to the probability of the corresponding leaf in $\mathfrak{T}$, we have $\sum_{v \in \limsup_{i \rightarrow \infty} \Omega_u^i} p_v^{\mathfrak{T}^{(\infty)}} \leq p_u^{\mathfrak{T}}$ for all $u \in \text{Leaf}^{\mathfrak{T}}$. We now show $\infty = \text{edl}(\mathfrak{T}) \leq \text{edl}(\mathfrak{T}^{(\infty)})$ by considering the two cases where $\sum_{v \in \limsup_{i \rightarrow \infty} \Omega_u^i} p_v^{\mathfrak{T}^{(\infty)}} < p_u^{\mathfrak{T}}$ for some $u \in \text{Leaf}^{\mathfrak{T}}$ and where $\sum_{v \in \limsup_{i \rightarrow \infty} \Omega_u^i} p_v^{\mathfrak{T}^{(\infty)}} = p_u^{\mathfrak{T}}$ for all $u \in \text{Leaf}^{\mathfrak{T}}$.

If we have $\sum_{v \in \limsup_{i \rightarrow \infty} \Omega_u^i} p_v^{\mathfrak{T}^{(\infty)}} < p_u^{\mathfrak{T}}$ for some $u \in \text{Leaf}^{\mathfrak{T}}$, then we obtain

$$\begin{aligned} 1 &\geq |\mathfrak{T}|_{\text{Leaf}} = \sum_{u \in \text{Leaf}^{\mathfrak{T}}} p_u^{\mathfrak{T}} \\ &> \sum_{u \in \text{Leaf}^{\mathfrak{T}}} \sum_{v \in \limsup_{i \rightarrow \infty} \Omega_u^i} p_v^{\mathfrak{T}^{(\infty)}} = \sum_{v \in \biguplus_{u \in \text{Leaf}^{\mathfrak{T}}} \limsup_{i \rightarrow \infty} \Omega_u^i} p_v^{\mathfrak{T}^{(\infty)}} \\ &= \sum_{v \in \text{Leaf}^{\mathfrak{T}^{(\infty)}}} p_v^{\mathfrak{T}^{(\infty)}} = |\mathfrak{T}^{(\infty)}|_{\text{Leaf}} \end{aligned}$$

and from $|\mathfrak{T}^{(\infty)}|_{\text{Leaf}} < 1$ we directly get $\text{edl}(\mathfrak{T}^{(\infty)}) = \infty$.

Otherwise, we have $\sum_{v \in \limsup_{i \rightarrow \infty} \Omega_u^i} p_v^{\mathfrak{T}^{(\infty)}} = p_u^{\mathfrak{T}}$ for all $u \in \text{Leaf}^{\mathfrak{T}}$, and thus we obtain

$$\begin{aligned} &\text{edl}(\mathfrak{T}) \\ &= \sum_{n=0}^{\infty} \sum_{\substack{u \in V \setminus \text{Leaf}^{\mathfrak{T}} \\ d^{\mathfrak{T}}(u)=n}} p_u^{\mathfrak{T}} \\ &= \sum_{n=0}^{\infty} (1 - \sum_{\substack{u \in \text{Leaf}^{\mathfrak{T}} \\ d^{\mathfrak{T}}(u) \leq n}} p_u^{\mathfrak{T}}) \quad (\sum_{\substack{u \in V \setminus \text{Leaf}^{\mathfrak{T}} \\ d^{\mathfrak{T}}(u)=n}} p_u^{\mathfrak{T}} = 1 - \sum_{\substack{u \in \text{Leaf}^{\mathfrak{T}} \\ d^{\mathfrak{T}}(u) \leq n}} p_u^{\mathfrak{T}}) \\ &= \sum_{n=0}^{\infty} (1 - \sum_{\substack{u \in \text{Leaf}^{\mathfrak{T}} \\ d^{\mathfrak{T}}(u) \leq n}} \sum_{v \in \limsup_{i \rightarrow \infty} \Omega_u^i} p_v^{\mathfrak{T}^{(\infty)}}) \quad (p_u^{\mathfrak{T}} = \sum_{v \in \limsup_{i \rightarrow \infty} \Omega_u^i} p_v^{\mathfrak{T}^{(\infty)}}) \\ &\leq \sum_{n=0}^{\infty} (1 - \sum_{\substack{u \in \text{Leaf}^{\mathfrak{T}} \\ d^{\mathfrak{T}}(u) \leq n}} \sum_{\substack{v \in \limsup_{i \rightarrow \infty} \Omega_u^i \\ d^{\mathfrak{T}^{(\infty)}}(v) \leq n}} p_v^{\mathfrak{T}^{(\infty)}}) \quad (\forall v \in \limsup_{i \rightarrow \infty} \Omega_u^i : d^{\mathfrak{T}}(u) \leq d^{\mathfrak{T}^{(\infty)}}(v)) \\ &= \sum_{n=0}^{\infty} (1 - \sum_{\substack{v \in \biguplus_{u \in \text{Leaf}^{\mathfrak{T}}} \limsup_{i \rightarrow \infty} \Omega_u^i \\ d^{\mathfrak{T}^{(\infty)}}(v) \leq n}} p_v^{\mathfrak{T}^{(\infty)}}) \quad (\Omega_u^i \text{ all disjoint}) \\ &= \sum_{n=0}^{\infty} (1 - \sum_{\substack{v \in \text{Leaf}^{\mathfrak{T}^{(\infty)}} \\ d^{\mathfrak{T}^{(\infty)}}(v) \leq n}} p_v^{\mathfrak{T}^{(\infty)}}) \quad (\text{Leaf}^{\mathfrak{T}^{(\infty)}} = \biguplus_{u \in \text{Leaf}^{\mathfrak{T}}} \limsup_{i \rightarrow \infty} \Omega_u^i) \\ &= \sum_{n=0}^{\infty} \sum_{\substack{u \in V \setminus \text{Leaf}^{\mathfrak{T}^{(\infty)}} \\ d^{\mathfrak{T}^{(\infty)}}(u)=n}} p_u^{\mathfrak{T}^{(\infty)}} \quad (\sum_{\substack{u \in V \setminus \text{Leaf}^{\mathfrak{T}^{(\infty)}} \\ d^{\mathfrak{T}^{(\infty)}}(u)=n}} p_u^{\mathfrak{T}^{(\infty)}} = 1 - \sum_{\substack{v \in \text{Leaf}^{\mathfrak{T}^{(\infty)}} \\ d^{\mathfrak{T}^{(\infty)}}(v) \leq n}} p_v^{\mathfrak{T}^{(\infty)}}) \\ &= \text{edl}(\mathfrak{T}^{(\infty)}) \end{aligned}$$

and therefore, $\text{edl}(\mathfrak{T}^{(\infty)}) \geq \text{edl}(\mathfrak{T}) = \infty$.

2. From iSAST to fSAST:

The proof for SAST is now a direct consequence of the proof for PAST. Let $\mathcal{S}$ be a PTRS that is non-overlapping, linear, and not fSAST. If there exists a single RST $\mathfrak{T}$ such that $\text{edl}(\mathfrak{T}) = \infty$, then we do the same as for PAST. Otherwise, there exists an infinite sequence of RSTs $(\mathfrak{T}_i)_{i \in \mathbb{N}}$ such that $\lim_{i \rightarrow \infty} \text{edl}(\mathfrak{T}_i) = \infty$. For each of these infinite RSTs we apply the construction as before such that we get

an *innermost* RST $\mathfrak{T}_i^{(\infty)}$ with $\text{edl}(\mathfrak{T}_i^{(\infty)}) \geq \text{edl}(\mathfrak{T}_i)$ for all $i \in \mathbb{N}$. Thus, $(\mathfrak{T}_i^{(\infty)})_{i \in \mathbb{N}}$ is an infinite sequence of innermost RSTs such that $\lim_{i \rightarrow \infty} \text{edl}(\mathfrak{T}_i^{(\infty)}) = \infty$, which proves that $\mathcal{S}$ is not iSAST either. $\square$

Theorem 20 (From wPAST to fPAST). *If a PTRS $\mathcal{S}$ is non-overlapping, linear, and non-erasing, then*

$$\mathcal{S} \text{ is fPAST} \iff \mathcal{S} \text{ is wPAST}$$

Proof. Same construction as in the proof of [Thm. 20](#) for AST. Again, we also get $\text{edl}(\mathfrak{T}^{(\infty)}) \geq \text{edl}(\mathfrak{T})$, which we then use as in the proof of [Thm. 17](#) regarding PAST. $\square$

Theorem 22 (From liPAST/liSAST to iPAST/iSAST). *If a PTRS $\mathcal{S}$ is non-overlapping, then*

$$\mathcal{S} \text{ is iPAST} \iff \mathcal{S} \text{ is liPAST}$$

$$\mathcal{S} \text{ is iSAST} \iff \mathcal{S} \text{ is liSAST}$$

Proof. Same construction as in the proof of [Thm. 22](#) for AST. Again, we also get $\text{edl}(\mathfrak{T}^{(\infty)}) \geq \text{edl}(\mathfrak{T})$, which we then use as in the proof of [Thm. 17](#) regarding PAST/SAST. $\square$

Theorem 27 (From iPAST/iSAST to fPAST/fSAST (2)). *If a PTRS $\mathcal{S}$ is non-overlapping and right-linear, then*

$$\mathcal{S} \text{ is fPAST} \iff \mathcal{S} \text{ is iPAST w.r.t. } \rightarrow_{\mathcal{S}}$$

$$\mathcal{S} \text{ is fSAST} \iff \mathcal{S} \text{ is iSAST w.r.t. } \rightarrow_{\mathcal{S}}$$

Proof. Same construction as in the proof of [Thm. 27](#) for AST. Again, we also get $\text{edl}(\mathfrak{T}^{(\infty)}) \geq \text{edl}(\mathfrak{T})$, which we then use as in the proof of [Thm. 17](#) regarding PAST/SAST. $\square$

Theorem 33 (From iPAST/iSAST to fPAST/fSAST (3)). *If a PTRS $\mathcal{S}$ is orthogonal and spare, then*

$$\mathcal{S} \text{ is fPAST on basic terms} \iff \mathcal{S} \text{ is iPAST on basic terms}$$

$$\mathcal{S} \text{ is fSAST on basic terms} \iff \mathcal{S} \text{ is iSAST on basic terms}$$

Proof. Same construction as in the proof of [Thm. 33](#) for AST. Again, we also get $\text{edl}(\mathfrak{T}^{(\infty)}) \geq \text{edl}(\mathfrak{T})$, which we then use as in the proof of [Thm. 17](#) regarding PAST/SAST. $\square$

Theorem 38 (From iPAST/iSAST to fPAST/fSAST (4)). *If a PTRS $\mathcal{S}$ is non-overlapping and spare, then*

$$\mathcal{S} \text{ is fPAST on basic terms} \iff \mathcal{S} \text{ is iPAST w.r.t. } \rightarrow_{\mathcal{S}} \text{ on basic terms}$$

$$\mathcal{S} \text{ is fSAST on basic terms} \iff \mathcal{S} \text{ is iSAST w.r.t. } \rightarrow_{\mathcal{S}} \text{ on basic terms}$$

Proof. Same construction as in the proof of [Thm. 38](#) for AST. Again, we also get $\text{edl}(\mathfrak{T}^{(\infty)}) \geq \text{edl}(\mathfrak{T})$, which we then use as in the proof of [Thm. 17](#) regarding PAST/SAST. $\square$