GOETHE UNIVERSITÄT

Frankfurt am Main

Fachbereich 12
Institut für Mathematik

Masterarbeit

Random Split Trees mit unbeschränktem Verzweigungsgrad

Autor: Thomas Fischer
5708132
s1077880@stud.uni-frankfurt.de (nur bis 30.09.2022)
thomas95fischer@aol.com

Erstgutachter: Prof. Dr. Ralph Neininger

Zweitgutachter: Prof. Dr. Anton Wakolbinger

Abgabedatum: 30.05.2022

Inhaltsverzeichnis

1 Einleitung

Random split trees mit beschränktem Verzweigungsgrad wurden vor allem in [2] von Luc Devroye untersucht. Allerdings sind auch random split trees mit unbeschränktem Verzweigungsgrad nützlich, wie Svante Janson in [1] aufzeigte. In dieser Arbeit wird daher das Modell der random split trees mit unbeschränktem Verzweigungsgrad aufbauend auf beiden genannten Quellen genauer untersucht.

Definition 1 (random split trees)**.**
Gegeben seien $b \in \mathbb{N}^+ \cup \{\infty\}$ und ein zufälliger b-dimensionaler Wahrscheinlichkeitsvektor S (also $S = (S_i)_{i=1}^b \in [0,1]^b$ für zufällige S_i, sodass $\sum_{i=1}^b S_i = 1$). S heißt der Split-Vektor. Sei $\mathcal{T}$ der unendliche Baum mit Wurzel w, bei dem jeder Knoten genau b Kinder hat. Hierbei seien die Kinder jedes Knotens geordnet, sodass man für jedes $i \in [b]$ vom i-ten Kind eines Knoten reden kann. Ferner seien $(S^v)_{v \in V(\mathcal{T})}$ unabhängige Kopien von S.
Zunächst betrachten wir alle Knoten von $\mathcal{T}$ als leer.
Nun werden die Knoten durch folgenden Prozess gefüllt:
1. Die Wurzel w wird gefüllt.
2. Der nächste zu füllende Knoten wird induktiv wie folgt bestimmt:
2.1 Setze $v = w$.
2.2 Solange v gefüllt ist, ziehe ein zufälliges Kind u von v, wobei das i-te Kind von v die Wahrscheinlichkeit S_i^v habe gezogen zu werden, und setze $v = u$.
2.3 v wird gefüllt (die Reihenfolge der gefüllten Knoten ist wichtig).

Nun ist der random split tree $(T_n)_{n \in \mathbb{N}^+}$ die Folge an Bäumen, bei der für jedes $n \in \mathbb{N}^+$ der Baum T_n genau der von den ersten n gefüllten Knoten induzierte Teilbaum von $\mathcal{T}$ ist.

Anmerkung 1.
1. In dieser Arbeit wird $\mathbb{N}_0$ verwendet, wenn betont werden soll, dass die 0 mit eingeschlossen ist, $\mathbb{N}^+$, wenn betont werden soll, dass die 0 nicht mit eingeschlossen ist, und $\mathbb{N}$, wenn dies egal ist.
2. Wie auch in anderer Literatur üblich sei $[n]$ eine Kurzschreibweise für die Menge der ersten n natürlichen Zahlen. Formal sei dies wie folgt definiert:

$$[n] := \left\{ m \in \mathbb{N}^+ : m \leq n \right\}$$

Insbesondere sei also $[\infty] := \mathbb{N}^+$.

Anmerkung 2.
Beliebiges (auch zufälliges und von S abhängiges!) Permutieren der Einträge des Split-Vektors S führt zu einem identisch verteilten random split tree, wenn man die Kinder entweder ungeordnet betrachtet oder in Reihenfolge ihres Erscheinens neu sortiert, da in diesem Fall die Reihenfolge der Komponenten von S keine Rolle spielt.

Ferner wird das Modell der random split trees mit folgendem Modell verglichen werden:

Definition 2 (preferential attachement trees)**.**
Gegeben seien Gewichte $(w_k)_{k\in\mathbb{N}_0}$ *mit* $w_k \in \mathbb{R}_0^+$ *für alle* $k \in \mathbb{N}_0$ *und* $w_0 > 0$.
Dann entsteht der preferential attachment tree $(B_n)_{n\in\mathbb{N}^+}$ *induktiv:*
B_1 *ist der Baum, der nur aus der Wurzel besteht.*
Gegeben B_n *entsteht* B_{n+1} *indem ein zufälliger Knoten* $v \in V(B_n)$ *ausgewählt und ein neuer Knoten an* v *als Kind angehängt wird. Die Wahrscheinlichkeit einen bestimmten Knoten* $v \in V(B_n)$ *zu wählen ist hierbei proportional zu* $w_{d(v)}$*, wobei* $d(v)$ *der Ausgangsgrad (also die Anzahl Kinder) von* v *in* B_n *ist.*

Anmerkung 3.
1. Nach Konstruktion gilt für alle $n \in \mathbb{N}^+$*, dass* $B_n \subsetneq B_{n+1}$ *und* $T_n \subsetneq T_{n+1}$.
2. Für die Gewichte $(w_k)_{k\in\mathbb{N}_0}$ *des preferential attachement trees gilt, dass eine lineare Skalierung dieser mit einer festen positiven Konstante zu einem identisch verteiltem preferential attachment tree führt, da die Gewichte nur proportional betrachtet werden.*
3. Per Konstruktion ist der preferential attachement tree ein Markov-Prozess, da es zu jedem Zeitpunkt nur relevant ist, wie der Baum gerade aussieht.

1.1 Hintergrund

In [2] führte Luc Devroye random split trees für $b \in \mathbb{N}^+$ (also mit beschränktem Verzweigungsgrad) ein, wobei er hierbei ein allgemeineres Modell aufbaute, welches allerdings über den Umfang dieser Arbeit hinausgehen würde, und untersuchte die Verteilung der Tiefe D_n des n-ten hinzugefügten Knotens. Hierbei wies er in [2, Theorem 2] nach, dass für eine größenverzerrte zufällige Wahl W aus den Komponenten von S gilt, dass $\frac{D_n}{\ln(n)} \xrightarrow[n\to\infty]{\mathbb{P}} \frac{1}{\mathbb{E}[\ln(W)]}$ sofern $W \overset{f.s.}{<} 1$. Zusätzlich wies er auch nach, dass D_n in diesem Fall mit entsprechender Skalierung in Verteilung gegen eine Standard-Normalverteilung konvergiert.

In [1, Theorem 1.5] wies Svante Janson nach, dass preferential attachement trees, bei denen w_k eine affin lineare Funktion in k ist (mit leichten Einschränkungen an die Parameter), genauso verteilt sind wie passende random split trees, wenn man die Reihenfolge der Kinder ignoriert oder in der Reihenfolge des Erscheinens neu sortiert. Hierfür ist es allerdings nötig auch den Wert $b = \infty$ (also einen unbeschränkten Verzweigungsgrad) zuzulassen.

1.2 Inhalt der Arbeit

Folgendes Theorem ist eine äquivalente Übersetzung und Umformulierung des von Svante Janson in [1] gezeigten Theorems 1.5:

Theorem 1.
Seien $\alpha \in \mathbb{R}$ und $\beta \in \mathbb{R}^+$, sodass entweder $\alpha \geq 0$ oder $\alpha \cdot n + \beta = 0$ für ein $n \in \mathbb{N}_{\geq 2}$.
Sei $(B_n)_{n\in\mathbb{N}^+}$ der preferential attachment tree mit $w_k = (\alpha \cdot k + \beta) \vee 0$ für jedes $k \in \mathbb{N}_0$.
Sei $(T_n)_{n\in\mathbb{N}^+}$ der random split tree mit $b = \infty$ und GEM$\left(\frac{\alpha}{\alpha+\beta}, \frac{\beta}{\alpha+\beta}\right)$-verteiltem S (siehe Definition 6).

Dann sind $(B_n)_{n\in\mathbb{N}^+}$ und $(T_n)_{n\in\mathbb{N}^+}$ identisch verteilt, wenn man die Kinder jeweils nach der Reihenfolge des Erscheinens ordnet.

Anmerkung 4.
Wie auch in anderer Literatur üblich, gilt folgende Notation für $a, b \in \mathbb{R}$:

$$a \vee b := \max(a, b); \qquad a \wedge b := \min(a, b)$$

Anmerkung 5.
In [1] enthält Theorem 1.5 noch folgende Aussagen:
1. Statt der GEM$\left(\frac{\alpha}{\alpha+\beta}, \frac{\beta}{\alpha+\beta}\right)$-Verteilung, lässt sich auch die PD$\left(\frac{\alpha}{\alpha+\beta}, \frac{\beta}{\alpha+\beta}\right)$-Verteilung nutzen (siehe Definition 7).
2. Statt die Kinder nach der Reihenfolge des Erscheinens zu ordnen, kann man auch ungeordnete Bäume betrachten.
Allerdings ist beides eine direkte Folgerung aus dieser Formulierung, da 1. nur eine Permutation der Komponenten von S darstellt (vergleiche Anmerkung 2) und 2. das Ergebnis einer determinierten messbaren Abbildung von dem Raum der geordneten Bäume in den Raum der ungeordneten Bäume ist.

Ferner wird analog zu [1] der Fall $\alpha + \beta = 0$ bewusst weggelassen, da in diesem Fall beide Bäume determinierte Pfade der Länge n sind. Somit wird nur $\alpha + \beta > 0$ betrachtet.

In dieser Arbeit wird der Beweis aus [1] von Svante Janson für Theorem 1 aufgearbeitet. Hierfür wird das folgende Lemma benötigt, für welches im Gegensatz zur ursprünglich eingereichten Arbeit ein kurzer Beweis vorgelegt wird, welcher von Professor Anton Wakolbinger bei der Lektüre des ursprünglichen Beweises entdeckt wurde (für die Notation siehe Definition 8):

Lemma 1.
Gegeben sei eine Polya(Θ_r, Θ_b)-Urne.
Dann existiert eine Beta(Θ_r, Θ_b)-verteilte Zufallsvariable R_∞, sodass $R_n \xrightarrow[n\to\infty]{f.s.} R_\infty$.

Ferner wird folgendes Theorem, welches eine Erweiterung von [2, Theorem 2] auf den Fall $b = \infty$ ist, bewiesen:

Theorem 2.
Gegeben sei ein random split tree $(T_n)_{n\in\mathbb{N}^+}$ mit $b \in \mathbb{N}^+ \cup \{\infty\}$ und Split-Vektor S. Sei D_n die Tiefe (also der Abstand zur Wurzel) des n-ten eingefügten Knotens.
Sei W die größenverzerrte Wahl aus S (also $\mathbb{P}(W = x|S) = \sum_{i\in[b]:S_i=x} S_i$).
Sei ferner $\mathbb{E}\left[\ln^2(W)\right] < \infty$ und

$$\mu := -\mathbb{E}\left[\ln\left(W\right)\right] > 0; \qquad \sigma := \sqrt{\mathrm{Var}\left(\ln(W)\right)} > 0.$$

Dann gilt:

$$\frac{D_{n+1} - \frac{\ln(n)}{\mu}}{\sigma\sqrt{\frac{\ln(n)}{\mu^3}}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1)$$

Anmerkung 6.
Da die Notation sonst in einem späteren Kapitel missverständlich sein würde, wird in dieser Arbeit folgende Notation für Potenzen von Logarithmen verwendet:

$$\ln^c(x) := (\ln(x))^c$$

Insbesondere ist mit $\ln^{-1}(x)$ nicht die Umkehrfunktion, sondern $\frac{1}{\ln(x)}$ gemeint.

Anmerkung 7.
In [2, Theorem 2] ist zusätzlich noch die Aussage $\frac{D_n}{\ln(n)} \xrightarrow[n\to\infty]{\mathbb{P}} \frac{1}{\mu}$ enthalten, welche allerdings von der obigen Formulierung bereits impliziert wird, da $\sigma\sqrt{\frac{\ln(n)}{\mu^3}} \in o\left(\ln(n)\right)$.

Anmerkung 8.
Die Bedingung $\sigma > 0$ ist äquivalent dazu, dass W nicht Dirac-verteilt ist. W ist genau dann Dirac-verteilt, wenn es ein $k \in \mathbb{N}^+$ gibt, sodass S f.s. genau k nicht-0-Einträge hat, die alle mit $\frac{1}{k}$ übereinstimmen. Dies entspricht allerdings einem einfacheren Modell, welches für den Fall $k = 2$ in [6] unter dem Namen random symmetric digital search trees untersucht wird. (Für $k > 2$ sind hierbei ähnliche Resultate wie für $k = 2$ zu erwarten.)
Ferner ist die Bedingung $\mu > 0$ äquivalent dazu, dass $\mathbb{P}(W = 1) < 1$. Dies impliziert insbesondere $\mathbb{E}[W] < 1$, was später benötigt wird.
Insbesondere impliziert $\sigma > 0$ also $\mu > 0$.

Anmerkung 9.
Die Bedingung $\mathbb{E}\left[\ln^2(W)\right] < \infty$ ist offensichtlich nötig, damit μ und σ überhaupt definiert sind. Dies verlangt also, dass die Schwanzverteilung von W Richtung 0 hinreichend schnell abfällt.

Auf Grund von Anmerkung 9 wird zusätzlich noch folgendes Lemma gezeigt:

Lemma 2.
Sei W die größenverzerrte Wahl aus einem GEM-verteiltem Wahrscheinlichkeitsvektor P. Dann erfüllt W die in Anmerkung 9 erwähnte Eigenschaft:

$$\mathbb{E}\left[\ln^2(W)\right] < \infty$$

Theorem 1 und 2 implizieren zusammen mit Lemma 2 insbesondere, dass auch ein preferential attachment tree mit den in Theorem 1 gewählten Gewichten $w_k = (\alpha \cdot k + \beta) \vee 0$ die in Theorem 2 beschriebene Eigenschaft bezüglich der Tiefe der neu hinzugefügten Knoten erfüllt. Dies ist allerdings wie in [1, Kapitel 5] erwähnt eine bereits bekannte Aussage.

Der Beweis von Lemma 2 führt auf natürliche Weise zu folgendem Resultat:

Lemma 3.
Seien $0 \leq \alpha < 1$ und $\alpha + \beta \geq 0$, dann gilt:

$$\sum_{i=1}^{\infty} \prod_{j=1}^{i-1} \frac{\alpha \cdot j + \beta}{\alpha \cdot j + \beta + 1} = \frac{1+\beta}{1-\alpha}$$

Da dem Autor sonst nur ein nicht intuitiver Beweis dieser Aussage über die hypergeometrische Funktion bekannt ist (was aufgrund der elementaren Struktur der Aussage überrascht), erscheint diese von eigenständigem Interesse.

2 Definitionen

In diesem Kapitel werden Definitionen und Notationen aufgelistet, die im Laufe dieser Arbeit häufiger gebraucht werden, wobei Notationen, die speziell für einzelne Beweise verwendet werden, erst am Beginn der jeweiligen Unterkapitel eingeführt werden, um den Überblick zu erleichtern.

Definition 3 (T^v)**.**
Sei T ein gewurzelter Baum und $v \in V(T)$. Dann ist T^v der Baum mit Wurzel v, welcher aus v und allen Nachfahren von v in T besteht.

2.1 Die verwendeten Verteilungen

Definition 4 (Gamma-Verteilung)**.**
Gegeben seien $\alpha \geq 0, \beta > 0$.
Falls $\alpha > 0$, dann ist die Gamma(α, β)-Verteilung die reellwertige Verteilung, welche folgende Dichte hat:

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \mathbb{1}_{\{x>0\}}$$

Sonst ($\alpha = 0$) ist die Gamma(α, β)-Verteilung die δ_0-Verteilung.

Anmerkung 10.
Die Gamma-Verteilung ist eine Verallgemeinerung der Exponential-Verteilung, weil die Summe von n unabhängigen Exponential(λ)-verteilten Zufallsvariablen Gamma(n, λ)-verteilt ist.

Definition 5 (Beta-Verteilung)**.**
Gegeben seien $\alpha, \beta \geq 0$ mit $\alpha + \beta > 0$. Seien ferner X_1, X_2 unabhängige Zufallsvariablen, wobei X_1 Gamma$(\alpha, 1)$-verteilt und X_2 Gamma$(\beta, 1)$-verteilt sei.
Dann ist die Beta(α, β)-Verteilung die $[0, 1]$-wertige Verteilung von:

$$\frac{X_1}{X_1 + X_2}$$

Anmerkung 11.
Statt Gamma$(\alpha, 1)$- und Gamma$(\beta, 1)$-verteilten Zufallsvariablen, kann man bekanntlich auch Gamma(α, γ)- und Gamma(β, γ)-verteilte Zufallsvariablen für beliebiges $\gamma > 0$ nutzen ohne etwas an der definierten Verteilung zu ändern.

Definition 6 (GEM-Verteilung)**.**
Gegeben seien $\alpha, \beta \in \mathbb{R}$, die die folgenden drei Eigenschaften erfüllen:
1.) $\alpha < 1$ 2.) $\alpha + \beta \geq 0$ 3.) $\alpha \geq 0$ oder $\alpha \cdot n + \beta = 0$ für ein $n \in \mathbb{N}^+$.
Seien ferner $(Z_j)_{j=1}^\infty$ unabhängige Zufallsvariablen, sodass für jedes $j \in \mathbb{N}^+$ die Zufallsvariable Z_j Beta$(1 - \alpha, (\alpha \cdot j + \beta) \vee 0)$-verteilt ist.

Dann ist die GEM(α, β)-Verteilung die Verteilung des folgenden zufälligen Wahrscheinlichkeitsvektors:

$$P = (P_i)_{i=1}^{\infty} \text{ mit } P_i := Z_i \cdot \prod_{j=1}^{i-1} (1 - Z_j)$$

Definition 7 (PD-Verteilung)**.**
Gegeben seien $\alpha, \beta \in \mathbb{R}$, die die folgenden drei Eigenschaften erfüllen:
1.) $\alpha < 1$ 2.) $\alpha + \beta \geq 0$ 3.) $\alpha \geq 0$ oder $\alpha \cdot n + \beta = 0$ für ein $n \in \mathbb{N}^+$.
Sei ferner $P = (P_i)_{i=1}^{\infty}$ GEM(α, β)-verteilt und $\hat{P}$ der zufällige Wahrscheinlichkeitsvektor, den man erhält, wenn man die $(P_i)_{i=1}^{\infty}$ der Größe nach absteigend sortiert.
Dann ist die PD(α, β)-Verteilung die Verteilung von $\hat{P}$.

2.2 Polya-Urne

Die Polya-Urne ist ein wichtiges Modell in der Stochastik. Das elementare Modell startet mit einer Urne in der sich eine einzelne rote und eine einzelne blaue Kugel befinden. Nun wird nacheinander eine rein zufällige Kugel aus der Urne gezogen und diese zusammen mit einer Kugel der identischen Farbe wieder in die Urne gelegt.
Anschaulich kann man sich dies durch folgendes Beispiel vorstellen:
Bei einer Sportart gibt es das rote und das blaue Team. Angenommen ein neuer Fan wird immer durch einen bestehenden Fan angeworben und wird dann Fan der gleichen Mannschaft. Wenn jetzt jede Mannschaft mit genau einem Fan startet, dann lässt sich die Entwicklung der Anzahl an Fans beider Mannschaften durch eine Polya-Urne modellieren.

Für diese Arbeit wird aber ein allgemeineres Modell der Polya-Urne benötigt:

Definition 8 (Polya-Urne mit reellen Parametern)**.**
Gegeben seien $\Theta_r, \Theta_b > 0$, dann ist die Polya(Θ_r, Θ_b)-Urne der folgende stochastische Prozess:

- $R_0 := \frac{\Theta_r}{\Theta_r + \Theta_b}$.
- I_1 *sei* Be(R_0)*-verteilt.*
- *Gegeben* I_1 *bis* I_n *sei* $R_n := \frac{\Theta_r + \sum_{i=1}^{n} I_i}{\Theta_r + \Theta_b + n}$ *und* I_{n+1} *sei* Be(R_n)*-verteilt.*

Anmerkung 12.
Hierbei kann man sich R_n als den relativen Anteil der roten Kugeln in der Urne nach n-Zügen vorstellen, während I_n die Indikatorvariable ist, ob im n-ten Zug eine rote Kugel hinzugefügt wurde. Allerdings startet man nicht mehr unbedingt mit einer ganzzahligen Anzahl an Kugeln.

3 Intuition und Ideen der Beweise

In diesem Kapitel werden die Ideen hinter den in dieser Arbeit vollzogenen Beweisen dargelegt, um eine leichtere Nachvollziehbarkeit dieser zu ermöglichen. Dazu wird hier die Reihenfolge der Argumente allerdings von der Reihenfolge der Argumente, wie sie in den Beweisen in Kapitel 5 vorkommen, abweichen.

3.1 Idee hinter Lemma 1

Die Idee dieses kürzeren Beweises ist die Aussage von hinten aufzurollen:
Wir starten mit einer Beta (Θ_r, Θ_b)-verteilten Zufallsvariable R_∞. Ferner seien $(U_n)_{n\in\mathbb{N}^+}$ unabhängig uniform auf $[0,1]$ verteilte Zufallsvariablen.
Dann lässt sich zeigen, dass es sich mit $I_n := \mathbb{1}_{\{U_n<R_\infty\}}$ um eine Polya(Θ_r, Θ_b)-Urne handelt und auf Grund des Starken Gesetzes der großen Zahlen $R_n \xrightarrow[n\to\infty]{f.s.} R_\infty$.

3.2 Idee hinter Theorem 1

Um Theorem 1 zu zeigen, ist es praktisch den Entstehungsprozess des preferential attachment trees umzuformulieren, um eine größere Ähnlichkeit zum random split tree zu schaffen. Sei hierfür V_{n+1} der Knoten aus $V(B_n)$, an den der $(n+1)$-te Knoten im preferential attachment tree angehängt wird (für $n \geq 1$, da die Wurzel als erster Knoten keinem Knoten angehängt wird).
Ferner wollen wir nun iterativ die Ereignisse $\{V_{n+1} \in V(B_n^v)\}$ und $\{V_{n+1} = v\}$ für einige der $v \in V(B_n)$ betrachten. Hierfür sind folgende Vorüberlegungen wichtig:

Anmerkung 13.
1. Sei u die Wurzel von B_n. Dann ist $\{V_{n+1} \in V(B_n^u)\} = \{V_{n+1} \in V(B_n)\}$ das sichere Ereignis.
2. Sei $v \in V(B_n)$ beliebig und seien $(v_i)_{i=1}^{d(v)}$ die Kinder von v. Dann gilt $\{V_{n+1} \in V(B_n^v)\} = \{V_{n+1} = v\} \dot\cup \dot\bigcup_{i=1}^{d(v)} \{V_{n+1} \in V(B_n^{v_i})\}$.

Dies ermöglicht es die Wahl von V_{n+1} ähnlich wie beim random split tree auf folgende Weise durchzuführen:
1. Sei $v := u$ (wie oben sei u die Wurzel von B_n).
2. Seien wieder $(v_i)_{i=1}^{d(v)}$ die Kinder von v. Entscheide nun, ob das Ereignis $\{V_{n+1} = v\}$ eintritt (mit passender Wahrscheinlichkeit). Wenn dies nicht eintritt, entscheide für welches $i \in [d(v)]$ das Ereignis $\{V_{n+1} \in V(B_n^{v_i})\}$ eintritt. (Dies können wir auf Grund von Anmerkung 13 so machen.)
3. Falls $V_{n+1} = v$, so wurde V_{n+1} festgelegt und man ist hiermit fertig. Sonst sei i so gewählt, dass $V_{n+1} \in V(B_n^{v_i})$, und es werde zu 2. gesprungen mit $v := v_i$.

In Lemma 20 wird dies aufgegriffen und nachgewiesen, dass der preferential attachment tree in dieser Betrachtungsweise nicht nur dem random split tree ähnelt, sondern im Fall von Theorem 1 sogar wie einer verteilt ist.

Da sich zusätzlich noch zeigen lässt, dass die relativen Größen der an der Wurzel hängenden Teilbäume in Verteilung gegen eine GEM-Verteilung konvergieren, lässt sich folgern, dass der entsprechende Splitvektor höchstens eine (zufällige) Permutation eines GEM-verteilten Splitvektors ist, was mit Hilfe von Anmerkung 2 Theorem 1 liefert.

3.3 Idee hinter Theorem 2

Um D_{n+1} zu untersuchen, ist es zunächst wichtig folgende Idee zu verstehen:
In jedem Schritt, in dem ein Kind anhand des Splitvektors ausgewählt wird, findet eine von den anderen Wahlen unabhängige größenverzerrte Wahl aus den Komponenten des Splitvektors statt (bzw. aus der unabhängigen Kopie von diesem für den entsprechenden Knoten). Die Größe dieser gewählten Komponente ist folglich eine unabhängige Kopie von W (entsprechend bezeichne W_d die Größe der gewählten Komponente in Tiefe $d-1$).
Hierbei stellen wir uns für jeden hinzugefügten Knoten eine unendlich lange Folge von Komponenten vor, die nach dem selben Schema wie beim random split tree gewählt werden, als wären alle Knoten bereits gefüllt.
Die Wahrscheinlichkeit, dass (gegeben W_d) ein anderer Durchlauf, der bis zu dem selben Knoten gekommen ist, genau diese Komponente auch auswählt, ist per Definition des random split trees genau die Größe W_d dieser Komponente.
Folglich ist (gegeben $(W_i)_{i=1}^d$) die Anzahl der Durchläufe unter den ersten n Durchläufen, die bis zur Tiefe d mit dem $n+1$-sten Durchlauf übereinstimmen, binomial-verteilt mit $p = \prod_{i=1}^d W_i$ (die verschiedenen Wahlen geschehen unabhängig voneinander).
Nun sucht man nach der Tiefe, in der es keinen vorherigen Durchlauf mehr gibt, der bis zu dieser Tiefe übereinstimmte. (Hierbei muss man allerdings aufpassen, dass der erste Durchlauf, der bei einem bestimmten Knoten ankommt, diesen füllt, und dieser Durchlauf somit eigentlich nicht weiter läuft. Allerdings lässt sich zeigen, dass dies vernachlässigbar ist.)
Die Idee des Beweises besteht nun darin $\prod_{i=1}^d W_i$ bzw. eigentlich $\sum_{i=1}^d \ln(W_i)$ mit Hilfe des Zentralen Grenzwertsatzes genauer zu untersuchen und zu analysieren, wann diese Summe einen bestimmten Wert erreicht (also das Produkt hinreichend klein wird). Es lässt sich dann feststellen, dass D_{n+1} im Grenzwert für $n \to \infty$ f.s. größer ist als die Tiefe, in der $\prod_{i=1}^d W_i$ den Wert $\frac{\ln(n)}{n}$ unterschreitet, aber auf der anderen Seite f.s. kleiner ist als die Tiefe, in der $\prod_{i=1}^d W_i$ den Wert $\frac{\ln^{-1}(n)}{n}$ unterschreitet. Da sich sogar nachweisen lässt, dass diese beiden Unterschreitungstiefen die gewünschte Eigenschaft haben, folgt so Theorem 2.

3.4 Idee hinter Lemma 3

Der Trick hier ist es $\mathbb{E}\left[\sum_{i=1}^{\infty} P_i\right]$ für $\mathrm{GEM}\,(\alpha, \beta)$-verteiltes $P = (P_i)_{i=1}^{\infty}$ zu berechnen. Hierbei kann man zum einen mit ein wenig Rechenaufwand folgendes nachweisen:

$$\mathbb{E}\left[\sum_{i=1}^{\infty} P_i\right] = \frac{1-\alpha}{1+\beta} \cdot \sum_{i=1}^{\infty} \prod_{j=1}^{i-1} \frac{\alpha \cdot j + \beta}{\alpha \cdot j + \beta + 1}$$

Zum anderen ist aber auch $\sum_{i=1}^{\infty} P_i \overset{f.s.}{=} 1$ per Definition. Somit gilt insbesondere:

$$1 = \frac{1-\alpha}{1+\beta} \cdot \sum_{i=1}^{\infty} \prod_{j=1}^{i-1} \frac{\alpha \cdot j + \beta}{\alpha \cdot j + \beta + 1}$$

Lemma 3 folgt dann direkt durch Multiplizieren beider Seiten der Gleichung mit $\frac{1+\beta}{1-\alpha}$.

3.5 Idee hinter Lemma 2

Die erste entscheidende Idee ist es hier, dass man für $\mathrm{GEM}\,(\alpha, \beta)$-verteiltes $P = (P_i)_{i=1}^{\infty}$ mit Hilfe der Turm-Eigenschaft $\mathbb{E}\left[\ln^2(W)\right]$ zu $\sum_{i=1}^{\infty} \mathbb{E}\left[P_i \ln^2(P_i)\right]$ umformen kann.
Ferner lässt sich zeigen, dass sich $x \ln^2(x)$ auf $[0, 1]$ für alle $\varepsilon > 0$ nach oben gegen ein festes Vielfaches von $x^{1-\varepsilon}$ abschätzen lässt, da $\frac{x \ln^2(x)}{x^{1-\varepsilon}}$ für $x \to 0$ gegen 0 geht.
Abschließend lässt sich $\mathbb{E}\left[\ln^2(W)\right] \leq \mathbb{E}\left[(P_i)^{1-\varepsilon}\right] < \infty$ mit Hilfe von Lemma 3 nachrechnen, was bereits Lemma 2 impliziert.

4 Verwendete Aussagen

In diesem Kapitel werden allgemein bekannte Aussagen aufgelistet, die in dieser Arbeit verwendet werden. Mit Namen bekannte Aussagen werden dabei nicht bewiesen:

Lemma 4 (Lemma von Borel-Cantelli)**.**
Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge von Ereignissen, dann gilt:

$$\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) < \infty \Rightarrow \sum_{n\in\mathbb{N}} \mathbb{1}_{A_n} \overset{f.s.}{<} \infty \tag{1}$$

Sind die Ereignisse zusätzlich noch wenigstens paarweise unabhängig, dann gilt zusätzlich:

$$\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = \infty \Rightarrow \sum_{n\in\mathbb{N}} \mathbb{1}_{A_n} \overset{f.s.}{=} \infty \tag{2}$$

Lemma 5 (de Finetti's Theorem)**.**
Sei $(X_n)_{n\in\mathbb{N}}$ eine austauschbare reellwertige borelmessbare Folge von Zufallsvariablen. Dann existiert eine Zufallsvariable Y, sodass $(X_n)_{n\in\mathbb{N}}$ gegeben Y f.s. u.i.v. ist.

Anmerkung 14.
Unter „$(X_n)_{n\in\mathbb{N}}$ gegeben Y f.s. u.i.v.“ kann man sich folgendes zweistufiges Zufallsexperiment vorstellen:
1. Zuerst werde Y gezogen. Sei also $Y = y$.
2. Danach werden Zufallsvariablen $\left(\hat{X}_n\right)_{n\in\mathbb{N}}$ unabhängig voneinander anhand der f.s. eindeutigen Verteilung $X_1|Y = y$ gezogen.
Dann sind die Folgen $(X_n)_{n\in\mathbb{N}}$ und $\left(\hat{X}_n\right)_{n\in\mathbb{N}}$ identisch verteilt.

Lemma 6 (Starkes Gesetz der großen Zahlen)**.**
Sei $(X_i)_{i\in\mathbb{N}}$ eine Folge u.i.v. Zufallsvariablen mit $\mathbb{E}[|X_1|] < \infty$, dann gilt:

$$\frac{\sum_{i=1}^n X_i}{n} \xrightarrow[n\to\infty]{f.s.} \mathbb{E}[X_1]$$

Anmerkung 15.
Für $(X_i)_{i\in\mathbb{N}}$ gegeben Y f.s. u.i.v. (statt $(X_i)_{i\in\mathbb{N}}$ u.i.v.) gilt analog:

$$\frac{\sum_{i=1}^n X_i}{n} \xrightarrow[n\to\infty]{f.s.} \mathbb{E}[X_1|Y]$$

Lemma 7.
Gegeben sei ein random split tree mit Splitvektor $S = (S_j)_{j=1}^b$. Sei $\mathcal{T}^j$ der Teilbaum von $\mathcal{T}$, der aus dem j-ten Kind der Wurzel und allen seinen Nachfahren besteht.
Sei $V_{j,n}$ der relative Anteil der Knoten in $\mathcal{T}^j$ an den ersten n gefüllten Knoten.
Dann gilt:

$$V_{j,n} \xrightarrow[n\to\infty]{f.s.} S_j$$

Beweis.
Betrachte $A_{j,n} := \{\text{Der } n\text{-te gefüllte Knoten liegt in } \mathcal{T}^j\}$. Dann gilt $V_{j,n} = \frac{\sum_{i=1}^n \mathbb{1}_{A_{j,i}}}{n}$ und ferner $\mathbb{1}_{A_{j,1}} = 0$ (die Wurzel wird als erstes gefüllt) sowie, dass die $\mathbb{1}_{A_{j,n}}$ für $n \geq 2$ gegeben S unabhängig und $\mathrm{Be}(S_j)$-verteilt sind. Folglich lässt sich das Starke Gesetz der großen Zahlen anwenden ($\mathbb{E}\left[\left|\mathbb{1}_{A_{j,n}}\right|\right] \leq 1$, $\mathbb{E}\left[\mathbb{1}_{A_{j,n}}|S\right] = S_j$):

$$V_{j,n} = \frac{\sum_{i=1}^n \mathbb{1}_{A_{j,i}}}{n} = \frac{\sum_{i=2}^n \mathbb{1}_{A_{j,i}}}{n} = \frac{n-1}{n} \cdot \frac{\sum_{i=2}^n \mathbb{1}_{A_{j,i}}}{n-1} \xrightarrow[n\to\infty]{f.s.} 1 \cdot S_j = S_j$$

□

Lemma 8 (Zentraler Grenzwertsatz)**.**
Sei $(X_i)_{i=1}^\infty$ eine Folge von u.i.v. Zufallsvariablen mit $\mathbb{E}[X_1^2] < \infty$. Dann gilt:

$$\frac{\sum_{i=1}^n X_i - n \cdot \mathbb{E}[X_1]}{\sqrt{n \cdot \mathrm{Var}(X_1)}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1)$$

Lemma 9 (Markov-Ungleichung)**.**
Sei $X \geq 0$ eine ZV mit $\mathbb{E}[X] \leq l < \infty$ und $\varepsilon > 0$, dann gilt:

$$\mathbb{P}(X \geq \varepsilon) \leq \frac{\mathbb{E}[X]}{\varepsilon} \leq \frac{l}{\varepsilon}$$

Lemma 10 (Chebyshev-Ungleichung)**.**
Sei X eine ZV mit $E[|X|^2] < \infty$, $\mathrm{Var}(X) \leq l < \infty$ und $\varepsilon > 0$, dann gilt:

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq \varepsilon) \leq \frac{\mathrm{Var}(X)}{\varepsilon^2} \leq \frac{l}{\varepsilon^2}$$

Lemma 11 (Turm-Eigenschaft)**.**
Seien X, Y Zufallsvariable mit $E[|X|] < \infty$, dann gilt:

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]]$$

Lemma 12 (Regel von de L'Hospital)**.**
Sei $c \in \mathbb{R}$. Seien ferner f, g differenzierbare Funktionen von $(c, c+\varepsilon)$ nach $\mathbb{R}$ für ein $\varepsilon > 0$. Wenn nun zusätzlich $\lim_{x\downarrow c}(f(x)) = \lim_{x\downarrow c}(g(x)) \in \{-\infty, 0, \infty\}$, dann gilt:

$$\lim_{x\downarrow c}\left(\frac{f(x)}{g(x)}\right) = \lim_{x\downarrow c}\left(\frac{f'(x)}{g'(x)}\right)$$

Sofern die rechte Seite wohldefiniert ist.

Lemma 13 (Dichte der Beta-Verteilung)**.**
Seien $\alpha, \beta > 0$, dann hat die Beta(α, β)-Verteilung folgende Dichte:

$$\frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} \mathbb{1}_{\{x\in(0,1)\}}$$

Wobei $B(\alpha,\beta) = \frac{\Gamma(\alpha)\cdot\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \int_{x\in(0,1)} x^{\alpha-1}(1-x)^{\beta-1}dx$ die Beta-Funktion ist.

Lemma 14.
Sei X Beta(α, β)-verteilt mit $\alpha, \beta, c > 0$, dann gilt:

$$\begin{aligned}
&1. \qquad \mathbb{E}\left[X^c\right] = \frac{B(\alpha+c,\beta)}{B(\alpha,\beta)} = \frac{\Gamma(\alpha+c)\cdot\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\alpha+\beta+c)} \\
&2. \ \mathbb{E}\left[(1-X)^c\right] = \frac{B(\alpha,\beta+c)}{B(\alpha,\beta)} = \frac{\Gamma(\beta+c)\cdot\Gamma(\alpha+\beta)}{\Gamma(\beta)\cdot\Gamma(\alpha+\beta+c)}
\end{aligned}$$

Beweis.

$$\begin{aligned}
1. \qquad \mathbb{E}\left[X^c\right] &= \int_{x\in(0,1)} x^c \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}dx \\
&= \frac{1}{B(\alpha,\beta)} \int_{x\in(0,1)} x^{\alpha+c-1}(1-x)^{\beta-1}dx \\
&= \frac{B(\alpha+c,\beta)}{B(\alpha,\beta)} = \frac{\frac{\Gamma(\alpha+c)\cdot\Gamma(\beta)}{\Gamma(\alpha+\beta+c)}}{\frac{\Gamma(\alpha)\cdot\Gamma(\beta)}{\Gamma(\alpha+\beta)}} \\
&= \frac{\Gamma(\alpha+c)\cdot\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\alpha+\beta+c)}
\end{aligned}$$

$$\begin{aligned}
2. \quad \mathbb{E}\left[(1-X)^c\right] &= \int_{x\in(0,1)} (1-x)^c \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}dx \\
&= \frac{1}{B(\alpha,\beta)} \int_{x\in(0,1)} x^{\alpha-1}(1-x)^{\beta+c-1}dx \\
&= \frac{B(\alpha,\beta+c)}{B(\alpha,\beta)} = \frac{\frac{\Gamma(\alpha)\cdot\Gamma(\beta+c)}{\Gamma(\alpha+\beta+c)}}{\frac{\Gamma(\alpha)\cdot\Gamma(\beta)}{\Gamma(\alpha+\beta)}} \\
&= \frac{\Gamma(\beta+c)\cdot\Gamma(\alpha+\beta)}{\Gamma(\beta)\cdot\Gamma(\alpha+\beta+c)}
\end{aligned}$$

□

5 Beweise

5.1 Beweis von Lemma 1

Für dieses Unterkapitel gelten die folgenden Notationen:

- Wie in Definition 8 sei $\Theta_r, \Theta_b > 0$.
- R_∞ sei Beta (Θ_r, Θ_b)-verteilt.
- Seien $(U_n)_{n\in\mathbb{N}^+}$ unabhängig uniform auf $[0,1]$ verteilt.
- Sei $I_n := \mathbb{1}_{\{U_n<R_\infty\}}$ für alle $n \in \mathbb{N}^+$.
- Wie in Definition 8 sei $R_n := \frac{\Theta_r+\sum_{i=1}^n I_i}{\Theta_r+\Theta_b+n}$ und $R_0 := \frac{\Theta_r}{\Theta_r+\Theta_b}$.

Lemma 15.
R_∞ gegeben I_1 bis I_n ist Beta $(\Theta_r + \sum_{i=1}^n I_i, \Theta_b + n - \sum_{i=1}^n I_i)$-verteilt.

Beweis.
Die bedingte Dichte von R_∞ gegeben I_1 bis I_n lässt sich wie folgt berechnen (für $x \in (0,1)$):

$$\begin{aligned}
&\frac{\frac{1}{B(\Theta_r,\Theta_b)}x^{\Theta_r-1}(1-x)^{\Theta_b-1}\cdot x^{\sum_{i=1}^n I_i}\cdot(1-x)^{n-\sum_{i=1}^n I_i}}{\int_{y\in(0,1)}\frac{1}{B(\Theta_r,\Theta_b)}y^{\Theta_r-1}(1-y)^{\Theta_b-1}\cdot y^{\sum_{i=1}^n I_i}\cdot(1-y)^{n-\sum_{i=1}^n I_i}dy}\\
&= \frac{x^{\Theta_r+\sum_{i=1}^n I_i\ -1}(1-x)^{\Theta_b+n-\sum_{i=1}^n I_i\ -1}}{\int_{y\in(0,1)} y^{\Theta_r+\sum_{i=1}^n I_i\ -1}(1-y)^{\Theta_b+n-\sum_{i=1}^n I_i\ -1}dy}\\
&= \frac{x^{\Theta_r+\sum_{i=1}^n I_i\ -1}(1-x)^{\Theta_b+n-\sum_{i=1}^n I_i\ -1}}{B(\Theta_r+\sum_{i=1}^n I_i, \Theta_b+n-\sum_{i=1}^n)}
\end{aligned}$$

Nach Lemma 13 ist dies die Dichte der Beta $(\Theta_r + \sum_{i=1}^n I_i, \Theta_b + n - \sum_{i=1}^n I_i)$-Verteilung. □

Beweis von Lemma 1.
$I_{n+1} = \mathbb{1}_{\{U_{n+1}<R_\infty\}}$ gegeben I_1 bis I_n ist per Definition Bernoulli-verteilt. Ferner lässt sich der Parameter mit Lemma 15 folgendermaßen nachrechnen:

$$\mathbb{P}\left(I_{n+1}|I_1,\ldots,I_n\right) = \mathbb{E}\left[\mathbb{1}_{\{U_{n+1}<R_\infty\}}|I_1,\ldots,I_n\right] = \mathbb{E}\left[R_\infty|I_1,\ldots,I_n\right] = \frac{\Theta_r+\sum_{i=1}^n I_i}{\Theta_r+\Theta_b+n} = R_n$$

Folglich handelt es sich bei der obigen Konstruktion um eine Polya-Urne.
Ferner folgt aus dem Starken Gesetzt der großen Zahlen direkt, dass $R_n \xrightarrow[n\to\infty]{f.s.} R_\infty$. □

5.2 Beweis von Theorem 1

Für dieses Unterkapitel seien immer die Voraussetzungen aus Theorem 1 ohne explizite Nennung gegeben.
Ferner gelten folgende Notationen:

- Sei V_{n+1} der Knoten aus $V(B_n)$, an den der $(n+1)$-te Knoten angehängt wird (für $n \geq 1$, da die Wurzel als erster Knoten keinem Knoten angehängt wird).
- Sei u die Wurzel von B_n (stimmt für alle n überein).
- Sei $K_n := d(u)$ die Anzahl der Kinder von u in B_n.
- Seien $(v_i)_{i=1}^{K_n}$ die Kinder von u in B_n in Reihenfolge ihres Erscheinens (die v_i brauchen n nicht als Index, da sie für alle n, für die $K_n \geq i$ ist, übereinstimmen).
- Das Gewicht einer Knotenmenge $K \subseteq V(B_n)$ sei $w(K) := \sum_{v \in K} w(v) := \sum_{v \in K} w_{d(v)}$.
- Das Gewicht eines Teilbaumes $T \subseteq B_n$ sei $w(T) := w(V(T)) = \sum_{v \in V(T)} w_{d(v)}$.

Anmerkung 16.
1. Wie anfangs bereits erwähnt, ist $d(v)$ der Ausgangsgrad von v, also die Anzahl seiner Kinder (in B_n). Da aus dem Kontext immer klar ist, auf welchen Baum sich $d(v)$ bezieht, wird zur Übersichtlichkeit auf eine genauere Notation verzichtet.
2. Die Wahrscheinlichkeit, dass von B_n nach B_{n+1} der neue Knoten an einen Knoten aus $K \subseteq V(B_n)$ angehängt wird, ist auf Grund der zu $w_{d(v)}$ proportionalen Wahrscheinlichkeit bei der Auswahl des Knotens genau $\frac{w(K)}{w(B_n)}$.

Lemma 16.
Sei $v \in V(B_n)$.
Dann gelten gegeben B_n folgende Gleichungen:

$$\begin{aligned} w(B_n) &= (\alpha+\beta)\cdot n - \alpha \\ w(B_n^v) &= (\alpha+\beta)\cdot |V(B_n^v)| - \alpha \\ \mathbb{P}(V_{n+1} \in V(B_n^v) \,|B_n) &= \frac{(\alpha+\beta)\cdot |V(B_n^v)| - \alpha}{(\alpha+\beta)\cdot n - \alpha} \end{aligned}$$

Beweis.
B_n hat genau n Knoten. Folglich ist $\sum_{v \in V(B_n)} d(v) = n-1$, da außer der Wurzel jeder Knoten Kind genau eines anderen Knotens ist. Somit lässt sich nachrechnen:

$$\begin{aligned} w(B_n) = \sum_{v \in V(B_n)} w_{d(v)} &= \sum_{v \in V(B_n)} (\alpha \cdot d(v) + \beta) = \alpha \cdot \left(\sum_{v \in V(B_n)} d(v) \right) + n \cdot \beta \\ &= \alpha \cdot (n-1) + \beta \cdot n = (\alpha+\beta)\cdot n - \alpha \end{aligned}$$

Analog folgt die zweite Aussage, da per Definition B_n^v selbst ein (Teil-)Baum ist und genau $|V(B_n^v)|$ Knoten hat.
Die letzte Aussage ergibt sich nun wie folgt aus Anmerkung 16:

$$\mathbb{P}(V_{n+1} \in V(B_n^v) | B_n) = \frac{w(B_n^v)}{w(B_n)} = \frac{(\alpha+\beta) \cdot |V(B_n^v)| - \alpha}{(\alpha+\beta) \cdot n - \alpha}$$

□

Lemma 17.

$$K_n \xrightarrow[n\to\infty]{f.s.} \begin{cases} \infty & \textit{falls } \alpha \geq 0 \\ -\frac{\beta}{\alpha} & \textit{falls } \alpha < 0 \end{cases}$$

Beweis.
Zunächst lässt sich festhalten, dass K_n realisierungsweise monoton wachsend ist und somit realisierungsweise der Limes existiert.

Fall 1: $\alpha \geq 0$
Dann gilt analog zu Lemma 16:

$$\mathbb{P}(V_{n+1} = u | B_n) = \frac{w(u)}{w(B_n)} = \frac{\alpha \cdot K_n + \beta}{(\alpha+\beta) \cdot n - \alpha} \geq \frac{\alpha \cdot 0 + \beta}{(\alpha+\beta) \cdot n} = \frac{\beta}{(\alpha+\beta)} \cdot \frac{1}{n}$$

Dies bedeutet, dass sich die Folge $\left(\mathbb{1}_{\{V_{n+1}=u\}}\right)_{n=1}^{\infty}$ eintrags- und realisierungsweise nach unten durch eine entsprechend gekoppelte Folge $(X_n)_{n=1}^{\infty}$ abschätzen lässt, wobei die X_n unabhängig und jeweils $\mathrm{Be}\left(\frac{\beta}{(\alpha+\beta)} \cdot \frac{1}{n}\right)$-verteilt sind ((*)formale Definition der Kopplung am Ende von Fall 1).
Somit gilt realisierungsweise:

$$K_n = \sum_{i=1}^{n-1} \mathbb{1}_{\{V_{i+1}=u\}} \geq \sum_{i=1}^{n-1} X_i$$

Und ferner, da $\beta > 0$:

$$\sum_{i=1}^{\infty} \mathbb{P}(X_i = 1) = \sum_{i=1}^{\infty} \frac{\beta}{(\alpha+\beta)} \cdot \frac{1}{n} = \infty$$

Somit folgt mit dem Lemma von Borel-Cantelli (2), dass:

$$\lim_{n\to\infty}(K_n) \geq \lim_{n\to\infty}\left(\sum_{i=1}^{n-1} X_i\right) \overset{f.s.}{=} \infty$$

Woraus direkt die Aussage für diesen Fall folgt.

(*) zur Vollständigkeit hier die erwähnte Kopplung der X_n an V_{n+1}:

Seien $(U_n)_{n\in\mathbb{N}^+}$ unabhängig uniform($[0,1]$)-verteilte Zufallsvariablen mit der Eigenschaft, dass $\{V_{n+1}=u\} = \left\{U_n \leq \frac{\alpha\cdot K_n+\beta}{(\alpha+\beta)\cdot n-\alpha}\right\}$ (solche existieren bekanntlich, da beide Ereignisse die identische Wahrscheinlichkeit haben).
Dann sei:

$$X_n := \mathbb{1}_{\left\{U_n \leq \frac{\beta}{(\alpha+\beta)}\cdot\frac{1}{n}\right\}}$$

Nun folgt aus $\frac{\beta}{(\alpha+\beta)}\cdot\frac{1}{n} \leq \frac{\alpha\cdot K_n+\beta}{(\alpha+\beta)\cdot n-\alpha}$ direkt:

$$X_n = \mathbb{1}_{\left\{U_n \leq \frac{\beta}{(\alpha+\beta)}\cdot\frac{1}{n}\right\}} \leq \mathbb{1}_{\left\{U_n \leq \frac{\alpha\cdot K_n+\beta}{(\alpha+\beta)\cdot n-\alpha}\right\}} = \mathbb{1}_{\{V_{n+1}=u\}}$$

Fall 2: $\alpha < 0$
Nach Voraussetzung ist $l := -\frac{\beta}{\alpha} \in \mathbb{N}^+$ und $w_l = \alpha\cdot\left(-\frac{\beta}{\alpha}\right)+\beta = 0$. Somit kann f.s. kein Knoten mehr als l Kinder in B_n haben und es gilt insbesondere f.s. $K_n \leq l$ für alle $n \in \mathbb{N}^+$. Folglich gilt auch $\lim_{n\to\infty} K_n \overset{f.s.}{\leq} l$.
Angenommen: $\mathbb{P}\left(\lim_{n\to\infty} K_n < l\right) > 0$
Da K_n für $n > 1$ nur Werte aus $\mathbb{N}^+$ annehmen kann, folgt, dass es ein $m \in [l-1]$ gibt mit:

$$\mathbb{P}\left(\lim_{n\to\infty} K_n = m\right) > 0$$

Sei $N_0 := \min\{n \in \mathbb{N}^+ : K_n = m\}$. Da K_n für $n > 1$ immer noch nur Werte aus $\mathbb{N}^+$ annehmen kann, gilt $\{\lim_{n\to\infty} K_n = m\} \subseteq \{N_0 < \infty\}$. Somit folgt insbesondere:

$$\begin{aligned}
&\mathbb{P}\left(N_0 < \infty \middle| \lim_{n\to\infty} K_n = m\right) = 1\\
\Rightarrow \exists\, n_0 \in \mathbb{N}: \quad &\mathbb{P}\left(N_0 \leq n_0 \middle| \lim_{n\to\infty} K_n = m\right) > 0\\
\Rightarrow \quad &\mathbb{P}\left(N_0 \leq n_0 \middle| \lim_{n\to\infty} K_n = m\right)\cdot\mathbb{P}\left(\lim_{n\to\infty} K_n = m\right) > 0\\
\Rightarrow \quad &\mathbb{P}\left(N_0 \leq n_0 \wedge \lim_{n\to\infty} K_n = m\right) > 0\\
\Rightarrow \quad &\mathbb{P}\left(K_{n_0} = K_{n_0+1} = K_{n_0+2} = \ldots = m\right) > 0
\end{aligned}$$

Andererseits gilt aber (für die gesamte Rechnung ist $\alpha < 0$ im Hinterkopf zu behalten):

$$\begin{aligned}
& \mathbb{P}\left(K_{n_0} = K_{n_0+1} = K_{n_0+2} = \ldots = m\right) \\
= \; & \mathbb{P}\left(K_{n_0+1} = K_{n_0+2} = \ldots = m | K_{n_0} = m\right) \cdot \mathbb{P}\left(K_{n_0} = m\right) \\
\leq \; & \mathbb{P}\left(K_{n_0+1} = K_{n_0+2} = \ldots = m | K_{n_0} = m\right) = \prod_{j=n_0}^{\infty} \mathbb{P}\left(K_{j+1} = m | K_j = m\right) \\
= \; & \prod_{j=n_0}^{\infty} \mathbb{P}\left(V_{j+1} \neq u | K_j = m\right) = \prod_{j=n_0}^{\infty} \left(1 - \frac{\alpha \cdot m + \beta}{(\alpha+\beta) \cdot j - \alpha}\right) \\
\overset{m \leq l-1}{\leq} & \prod_{j=n_0}^{\infty} \left(1 - \frac{\alpha \cdot (l-1) + \beta}{(\alpha+\beta) \cdot j - \alpha}\right) \overset{\alpha \cdot l + \beta = 0}{=} \prod_{j=n_0}^{\infty} \left(1 - \frac{-\alpha}{(\alpha+\beta) \cdot j - \alpha}\right) \\
= \; & \prod_{j=n_0}^{\infty} \left(1 - \frac{\frac{-\alpha}{\alpha+\beta}}{j + \frac{-\alpha}{\alpha+\beta}}\right) = \prod_{j=n_0}^{\infty} \frac{j}{j + \frac{-\alpha}{\alpha+\beta}} = \prod_{j=n_0}^{\infty} \frac{1}{1 + \frac{-\alpha}{\alpha+\beta} \cdot \frac{1}{j}} \\
= \; & \frac{1}{\prod_{j=n_0}^{\infty} \left(1 + \frac{-\alpha}{\alpha+\beta} \cdot \frac{1}{j}\right)} \leq \frac{1}{\sum_{j=n_0}^{\infty} \frac{-\alpha}{\alpha+\beta} \cdot \frac{1}{j}} = \frac{1}{\infty} = 0
\end{aligned}$$

Dies ist ein Widerspruch.
Folglich ist $\mathbb{P}\left(\lim_{n\to\infty} K_n < l\right) = 0$ und somit $\lim_{n\to\infty} K_n \overset{f.s.}{=} l$. □

Lemma 18.
Sei $n \in \mathbb{N}^+$ und $v \in V\left(B_n\right)$, dann gilt:

$$\left|V\left(B_{n+i}^v\right)\right| \xrightarrow[i\to\infty]{f.s.} \infty$$

Beweis.
Dies lässt sich analog zum Fall 1 von Lemma 17 beweisen:
OE sei $n \geq 2$ (sonst ist $v = u$ und wir würden nach der Größe des gesamten Baums fragen, welche natürlich gegen ∞ geht). Mit Lemma 16 gilt:

$$\begin{aligned}
\mathbb{P}\left(V_{n+j+1} \in V\left(B_{n+j}^v\right) | B_{n+j}\right) &= \frac{(\alpha+\beta) \cdot \left|V\left(B_{n+j}^v\right)\right| - \alpha}{(\alpha+\beta) \cdot (n+j) - \alpha} = \frac{(\alpha+\beta) \cdot \left|V\left(B_{n+j}^v\right)\right| - \alpha}{(\alpha+\beta) \cdot (n-1+j) + \beta} \\
&\overset{*}{\geq} \frac{(\alpha+\beta) \cdot 1 - \alpha}{(\alpha+\beta) \cdot (n-1+j) + \beta \cdot (n-1+j)} \\
&= \frac{\beta}{(\alpha+2\beta) \cdot (n-1+j)} = \frac{\beta}{\alpha+2\beta} \cdot \frac{1}{n-1+j}
\end{aligned}$$

$*$ Es gilt $\left|V\left(B_{n+j}^v\right)\right| \geq 1$, $\beta > 0$ und $n - 1 + j \geq 1$.

Dies bedeutet, dass sich die Folge $\left(\mathbb{1}_{\left\{V_{n+j+1} \in V\left(B_{n+j}^v\right)\right\}}\right)_{j=1}^{\infty}$ eintrags- und realisierungsweise nach unten durch eine entsprechend gekoppelte Folge $(X_j)_{j=1}^{\infty}$ abschätzen lässt, wobei die X_j unabhängig und jeweils $\mathrm{Be}\left(\frac{\beta}{\alpha+2\beta} \cdot \frac{1}{n-1+i}\right)$-verteilt sind ((*)formale Definition der Kopplung am Ende des Beweises).

Somit gilt realisierungsweise:

$$\left|V\left(B_{n+i}^{v}\right)\right| = \sum_{j=1}^{i-1} \mathbb{1}_{\left\{V_{n+j+1} \in V\left(B_{n+j}^{v}\right)\right\}} \geq \sum_{i=1}^{n-1} X_j$$

Und ferner, da $\beta > 0$:

$$\sum_{j=1}^{\infty} \mathbb{P}\left(X_j = 1\right) = \sum_{j=1}^{\infty} \frac{\beta}{\alpha + 2\beta} \cdot \frac{1}{n-1+j} = \sum_{m=n}^{\infty} \frac{\beta}{\alpha+2\beta} \cdot \frac{1}{m} = \infty$$

Somit folgt mit dem Lemma von Borel-Cantelli (2) (die Grenzwerte existieren, da es sich um monoton wachsende Folgen handelt):

$$\lim_{i\to\infty}\left(\left|V\left(B_{n+i}^{v}\right)\right|\right) \geq \lim_{i\to\infty}\left(\sum_{j=1}^{i-1} X_j\right) \overset{f.s.}{=} \infty$$

(*) zur Vollständigkeit hier die erwähnte Kopplung der X_j an V_{n+j+1}:
Seien $(U_j)_{j\in\mathbb{N}^+}$ unabhängig uniform($[0,1]$)-verteilte Zufallsvariablen mit der Eigenschaft, dass $\left\{V_{n+j+1} \in V\left(B_{n+i}^{v}\right)\right\} = \left\{U_j \leq \frac{(\alpha+\beta)\cdot\left|V\left(B_{n+j}^{v}\right)\right|-\alpha}{(\alpha+\beta)\cdot(n+j)-\alpha}\right\}$ (solche existieren bekanntlich, da beide Ereignisse die identische Wahrscheinlichkeit haben).
Dann sei:

$$X_j := \mathbb{1}_{\left\{U_n \leq \frac{\beta}{\alpha+2\beta}\cdot\frac{1}{n-1+j}\right\}}$$

Nun folgt aus $\frac{\beta}{\alpha+2\beta}\cdot\frac{1}{n-1+j} \leq \frac{(\alpha+\beta)\cdot\left|V\left(B_{n+j}^{v}\right)\right|-\alpha}{(\alpha+\beta)\cdot(n+j)-\alpha}$ direkt:

$$X_j = \mathbb{1}_{\left\{U_n \leq \frac{\beta}{\alpha+2\beta}\cdot\frac{1}{n-1+j}\right\}} \leq \mathbb{1}_{\left\{U_n \leq \frac{(\alpha+\beta)\cdot\left|V\left(B_{n+j}^{v}\right)\right|-\alpha}{(\alpha+\beta)\cdot(n+j)-\alpha}\right\}} = \mathbb{1}_{\{V_{n+j+1}=u\}}$$

□

Lemma 19.
Sei $N_{i,n} := \frac{\left|V\left(B_n^{v_i}\right)\right|}{n}$ *für* $i \leq K_n$ *und* $N_{i,n} := 0$ *sonst.*

Dann existiert eine GEM$\left(\frac{\alpha}{\alpha+\beta}, \frac{\beta}{\alpha+\beta}\right)$*-verteilte Zufallsvariable* $N = (N_i)_{i=1}^{\infty}$*, sodass für alle* $i \in \mathbb{N}^+$ *gilt:*

$$N_{i,n} \xrightarrow[n\to\infty]{f.s.} N_i$$

Beweis.
Da $\alpha + \beta > 0$, sind $\alpha' := \frac{\alpha}{\alpha+\beta}$ und $\beta' := \frac{\beta}{\alpha+\beta}$ wohldefiniert und $\alpha' + \beta' = 1$.
Sei $(B'_n)_{n\in\mathbb{N}^+}$ der preferential attachment tree mit den Gewichten $w'_k = (\alpha' \cdot k + \beta') \vee 0$. Dann handelt es sich bei $(B_n)_{n\in\mathbb{N}^+}$ und $(B'_n)_{n\in\mathbb{N}^+}$ nach Anmerkung 3 um identisch verteilte preferential attachment trees, da $w_k = (\alpha + \beta) \cdot w'_k$.
Also genügt es die Aussage für $w_k = (\alpha' \cdot k + \beta') \vee 0$ und $\mathrm{GEM}\,(\alpha', \beta')$-verteiltes N zu zeigen.

Sei $M_{i,n} := \frac{\left|V\left(B_n^{v_i}\right)\right|}{\sum_{j=i}^{K_n}\left|V\left(B_n^{v_j}\right)\right|}$ für $i \leq K_n$ und $M_{i,n} := 1$ für $i > K_n$ der relative Anteil der Knoten des Teilbaumes mit Wurzel v_i an allen Knoten, die weder u sind noch v_j oder Nachfahren von v_j für irgendein $j < i$ sind.
Nun wird der Grenzwert von $M_{i,n}$ für $n \to \infty$ untersucht:

Fall 1: $\alpha' < 0$ und $i \geq -\frac{\beta'}{\alpha'}$
In diesem Fall folgt aus Lemma 17, dass $K_n \leq i$ f.s. für alle $n \in \mathbb{N}^+$ und somit $M_{i,n} = 1$ f.s. für alle $n \in \mathbb{N}^+$. (Für $K_n < i$ folgt dies direkt aus der Definition von $M_{i,n}$ und für $K_n = 1$ gilt dies, weil dann $\sum_{j=i}^{K_n}\left|V\left(B_n^{v_j}\right)\right| = \left|V\left(B_n^{v_i}\right)\right|$.)
Folglich konvergiert in diesem Fall $M_{i,n}$ f.s. gegen 1, was sich auch als $\mathrm{Beta}\,(1-\alpha', 0)$-verteilte Zufallsvariable betrachten lässt.
Fall 2: sonst
In diesem Fall folgt aus Lemma 17, dass es f.s. ein $n_0 \in \mathbb{N}^+$ gibt, sodass $K_n \geq i$ für alle $n \geq n_0$ (n_0 hängt natürlich von der Realisierung ab). Sei n_0 kleinstmöglich gewählt hierfür (also der Zeitpunkt zu dem v_i hinzugefügt wird). Zu diesem Zeitpunkt sieht die Situation wie folgt aus:
Es gilt $K_{n_0} = i$, $\left|V\left(B_{n_0}^{v_i}\right)\right| = 1$, $w(u) = w_{K_{n_0}} = \alpha' \cdot i + \beta'$ und $w\left(B_{n_0}^{v_i}\right) = \beta'$.
Lemma 18 liefert, dass $B_n^{v_i}$ für $n \to \infty$ f.s. unendlich groß wird. Dies impliziert insbesondere, dass es f.s. unendlich viele Zeitpunkte gibt, zu denen ein Knoten hinzugefügt wird, der nicht Nachfahre von v_1 bis v_{i-1} ist, was folgende Betrachtungsweise ermöglicht:
Sei $(n_m)_{m\in\mathbb{N}^+}$ die Folge der Zeitpunkte nach n_0 zu denen ein Knoten hinzugefügt wird, der nicht Nachfahre von v_1 bis v_{i-1} ist, mit $n_0 < n_1 < \ldots < n_m < n_{m+1} \ldots$. (Beim Betrachten der Zeitpunkt bedingen wir also indirekt darauf, dass der neue Knoten in ein $B_n^{v_j}$ mit $j \geq i$ fällt.)(Da es f.s. unendlich viele solche Zeitpunkte gibt, sei hierauf im Folgenden immer implizit bedingt.)
Folglich ist $\sum_{j=i}^{K_{n_m}}\left|V\left(B_{n_m}^{v_j}\right)\right| = m + 1$, da genau zu diesen Zeitpunkten Knoten in diesem Bereich hinzugefügt werden. Ferner ist $M_{i,n} = M_{i,n_m}$ für $n_m \leq n < n_{m+1}$, da zu $B_n^{v_j}$ für $j < i$ hinzugefügte Knoten nichts an $M_{i,n}$ ändern.
Dementsprechend wird nun die Teilfolge $(B_{n_m})_{m\in\mathbb{N}_0}$ untersucht:
Alle Nachfahren von v_i seien rot und alle v_j für $j > i$ und alle ihre jeweiligen Nachfahren blau gefärbt. Sei r_m die Anzahl an roten Knoten in B_{n_m} und b_m die Anzahl an blauen Knoten in B_{n_m}. Es ist also $r_m + b_m = m$, $\left|V\left(B_{n_m}^{v_i}\right)\right| = r_m + 1$, $\sum_{j=i}^{K_{n_m}}\left|V\left(B_{n_m}^{v_j}\right)\right| = m + 1$

und $M_{i,n_m} = \frac{r_m+1}{r_m+b_m+2} = \frac{r_m+1}{m+1}$.
Die Wahrscheinlichkeit, dass von B_{n_m} nach $B_{n_{m+1}}$ der neue Knoten rot ist, ist folglich:

$$\begin{aligned}
\mathbb{P}\left(V_{n_{m+1}} \in V\left(B_{n_m}^{v_i}\right) \middle| B_{n_m}\right) &\overset{*1}{=} \frac{w\left(B_{n_m}^{v_i}\right)}{\sum_{j=i}^{K_{n_m}} w\left(B_{n_m}^{v_j}\right) + w(u)} \\
&\overset{*2}{=} \frac{(\alpha'+\beta') \cdot \left|V\left(B_{n_m}^{v_i}\right)\right| - \alpha'}{\sum_{j=i}^{K_{n_m}} \left((\alpha'+\beta') \cdot \left|V\left(B_{n_m}^{v_j}\right)\right| - \alpha'\right) + \alpha' \cdot K_{n_m} + \beta'} \\
&= \frac{1 \cdot (r_m+1) - \alpha'}{\sum_{j=i}^{K_{n_m}} \left|V\left(B_{n_m}^{v_j}\right)\right| - (K_{n_m} - (i-1)) \cdot \alpha' + \alpha' \cdot K_{n_m} + \beta'} \\
&= \frac{r_m + 1 - \alpha'}{(m+1) - K_{n_m} \cdot \alpha' + (i-1) \cdot \alpha' + \alpha' \cdot K_{n_m} + \beta'} \\
&= \frac{r_m + 1 - \alpha'}{m + 1 + (i-1) \cdot \alpha' + \beta'} \\
&= \frac{1 - \alpha' + r_m}{1 - \alpha' + i\alpha' + \beta' + m}
\end{aligned}$$

$*1$ Analoges Vorgehen zum Beweis von Lemma 16, wobei hier die $B_{n_m}^{v_j}$ für $j < i$ ausgeschlossen sind.
$*2$ Anwendung von Lemma 16.

Folglich lässt sich dieser Prozess als Polya$(1-\alpha', i\alpha'+\beta')$-Urne mit $R_m := \frac{1-\alpha'+r_m}{1-\alpha'+i\alpha'+\beta'+m}$ betrachten (es gilt $1-\alpha' = \beta' > 0$ und $i\alpha'+\beta' > 0$ nach Voraussetzung) und Lemma 1 liefert, dass es eine Beta $(1-\alpha', i\alpha'+\beta')$-verteilte Zufallsvariable R_∞ gibt, sodass:

$$\begin{aligned}
\lim_{m\to\infty} M_{i,n_m} &= \lim_{m\to\infty} \frac{r_m+1}{m+1} = \lim_{m\to\infty} \frac{1+r_m}{1+m} \\
&= \lim_{m\to\infty} \frac{1-\alpha'+r_m}{1-\alpha'+i\alpha'+\beta'+m} \\
&= \lim_{m\to\infty} R_m \overset{f.s.}{=} R_\infty
\end{aligned}$$

Folglich konvergiert auch $M_{i,n}$ f.s. gegen R_∞, da $M_{i,n} = M_{i,n_m}$ für $n_m \leq n < n_{m+1}$.

Aus beiden Fällen zusammen folgt also, dass es eine Beta $(1-\alpha', (i\alpha'+\beta') \vee 0)$-verteilte Zufallsvariable M_i gibt, sodass $M_{i,n} \overset{f.s.}{\to} M_i$ für $n \to \infty$.

Es gilt $\sum_{j=1}^{K_n} \left|V\left(B_n^{v_j}\right)\right| = n-1$, da $V(B_n) = \left(\dot{\bigcup}_{j=1}^{K_n} V\left(B_n^{v_j}\right)\right) \dot{\cup} \{u\}$ und somit folgt:

$$\begin{aligned}
N_{i,n} &= \frac{|V(B_n^{v_i})|}{n} \\
&\overset{*}{=} \frac{\sum_{j=1}^{K_n}\left|V\left(B_n^{v_j}\right)\right|}{n} \cdot \prod_{k=1}^{i-1} \left(\frac{\sum_{j=k+1}^{K_n}\left|V\left(B_n^{v_j}\right)\right|}{\sum_{j=k}^{K_n}\left|V\left(B_n^{v_j}\right)\right|} \right) \cdot \frac{|V(B_n^{v_i})|}{\sum_{j=i}^{K_n}\left|V\left(B_n^{v_j}\right)\right|} \\
&= \frac{n-1}{n} \cdot \prod_{k=1}^{i-1} \left(1 - \frac{|V(B_n^{v_k})|}{\sum_{j=k}^{K_n}\left|V\left(B_n^{v_j}\right)\right|} \right) \cdot M_{i,n} \\
&= \left(1 - \frac{1}{n}\right) \cdot \prod_{k=1}^{i-1} (1 - M_{k,n}) \cdot M_{i,n}
\end{aligned}$$

$*$ Teleskop-Produkt, wobei sich die untere Summationsgrenze jeweils immer um eins verschiebt.

Somit konvergiert $(N_{i,n})_{i=1}^{\infty}$ f.s. gegen $N = (N_i)_{i=1}^{\infty}$ für $N_i := \prod_{k=1}^{i-1}(1-M_k) \cdot M_i$. Ferner ist N nach Definition 6 GEM (α', β')-verteilt. □

Lemma 20.
Die Betrachtungsweise aus Kapitel 3.2 des preferential attachment trees stimmt in Verteilung mit einem random split tree für einen Splitvektor S überein.

Beweis.
Wir zeigen die Aussage zunächst für den Schritt an der Wurzel (also mit $v = u$):
Sei w_n jeweils der n-te hinzugefügte Knoten (w_1 ist also die Wurzel u). Ferner sei $(U_{w_n})_{n=2}^{\infty}$ eine Folge von Zufallsvariablen, die wie folgt konstruiert werde:
Für alle $n \in \mathbb{N}^+$: Wenn $V_n = u$, dann sei $U_{w_{n+1}}$ unabhängig von allem bisherigen uniform($[0,1]$)-verteilt. Sonst sei $U_{w_{n+1}} = U_{V_n}$.
Also bekommt jeder an der Wurzel hängende Teilbaum eine uniform aus $[0,1]$ gewählte Bezeichnung und jeder Knoten in diesem Teilbaum erhält diese Bezeichnung. Es sollte klar sein, dass sich aus der Folge der $(U_{w_n})_{n=2}^{\infty}$ f.s. eindeutig rekonstruieren lässt, welcher Knoten in welchem an der Wurzel hängenden Teilbaum unterkommt (Das einzige Problem wäre es, wenn zwei verschiedene Teilbäume die identische Bezeichnung bekommen würden, was allerdings Wahrscheinlichkeit 0 hat; daher f.s.). Somit lässt sich auch der Schritt an der Wurzel f.s. eindeutig hieraus rekonstruieren.
Nun wird nachgewiesen, dass die Folge der $(U_{w_n})_{n=2}^{\infty}$ austauschbar ist, um de Finetti's Theorem anwenden zu können. Hierfür ist zunächst wichtig einzusehen, dass die Verteilung von $U_{w_{n+1}}$ nicht von der Reihenfolge der U_{w_i} für $i \leq n$ abhängt, sondern nur von den bisher aufgetretenen Werten und ihren Häufigkeiten, da es sich nach Anmerkung 3 bei dem preferential attachment tree um einen Markov-Prozess handelt und die Wahl der Teilbäume somit lediglich von deren Größe abhängt. Folglich genügt es zu zeigen, dass für beliebiges festes m die ersten m der U_{w_n} austauschbar sind (zur Erinnerung: Austauschbarkeit verlangt nur, dass endliche Permutationen die Verteilung nicht verändern).

Zunächst lässt sich festhalten, dass die unabhängigen uniformen Wahlen aus $[0,1]$ die Austauschbarkeit nicht verhindern, da sie sogar u.i.v. sind. Also lassen sich hiervon m Werte „auf Reserve“ ziehen, aus welchen dann bei jedem neuen Kind der Wurzel ein Wert ohne Zurücklegen gezogen wird (es werden höchstens $m-1$ dieser Werte gebraucht). Um die Notation zu vereinfachen nehmen wir für diese „Reservewerte“ als Vertretung die Zahlen $\{1,2,\dots,m\}$ (also steht jede Zahl für einen vorher unabhängig uniform aus $[0,1]$ gezogenen Wert).
Dann lässt sich die Wahrscheinlichkeit für einen konkreten Ausgang wie folgt berechnen (Seien $i_n \in [m]$ für $2 \leq n \leq m$. Ferner sei $h_{i,n} := |\{2 \leq j \leq n : i_j = i\}|$ und $h_i := h_{i,m}$ die Anzahl an i_n, die gleich i sind.):

$$\begin{aligned}
\mathbb{P}\left(U_{w_2} = i_2, U_{w_3} = i_3, \dots, U_{w_m} = i_m\right) &= \prod_{n=1}^{m-1} \mathbb{P}\left(U_{w_{n+1}} = i_{n+1} | U_{w_2} = i_2, \dots, U_{w_n} = i_n\right) \\
&\overset{*1}{=} \prod_{n=1}^{m-1} \frac{\frac{\alpha \cdot K_n + \beta}{m - K_n} \cdot \mathbb{1}_{\{h_{i_{n+1},n}=0\}} + \left((\alpha+\beta)\cdot h_{i_{n+1},n} - \alpha\right) \cdot \mathbb{1}_{\{h_{i_{n+1},n} \geq 1\}}}{(\alpha+\beta)\cdot n - \alpha} \\
&\overset{*2}{=} \frac{\prod_{j=1}^{K_m} \frac{\alpha \cdot j + \beta}{m-j} \cdot \prod_{i=1}^{m} \left(\prod_{j=1}^{h_i - 1} (\alpha+\beta)\cdot j - \alpha\right)}{\prod_{n=1}^{m-1} (\alpha+\beta)\cdot n - \alpha}
\end{aligned}$$

$*1$ Anwendung von Lemma 16 und der Tatsache, dass $w(u) = \alpha \cdot K_n + \beta$, sowie, dass K_n der m Werte bereits gezogen wurden.
$*2$ Umsortierung der Faktoren im Zähler.

Hieran sieht man, dass die Reihenfolge der i_n keine Rolle spielt und somit die ersten m der U_{w_n} austauschbar sind. Folglich ist die ganze Folge der U_{w_n} austauschbar.

Nun folgt aus de Finetti's Theorem (Lemma 5), dass es eine Zufallsvariable Y gibt, sodass gegeben Y die U_{w_n} f.s. u.i.v. sind. Da die Wurzel höchstens abzählbar viele Kinder haben kann, kann U_{w_1} gegeben Y auch f.s. höchstens abzählbar viele verschiedene Werte annehmen. Folglich hat U_{w_1} gegeben Y f.s. eine diskrete Verteilung und die möglichen Werte lassen sich mit ihren jeweiligen Treffwahrscheinlichkeiten aufzählen. Wenn man nun S^u als den Vektor dieser Treffwahrscheinlichkeiten nimmt und statt von Werten aus $[0,1]$ wieder an die an u hängenden Teilbäume denkt, dann erhält man genau den Prozess, der beim random Split tree an der Wurzel durchlaufen wird.

Lemma 18 liefert uns, dass jeder Teilbaum B_n^v f.s. unbeschränkt groß wird, was die analoge Argumentation zu dem Vorgang an der Wurzel für jeden beliebigen Knoten $v \in V(B_n)$ zulässt (Teilfolge der in B_n^v fallenden Knoten betrachten und B_n^v als eigenen Baum mit Wurzel v sehen). Folglich handelt es sich hierbei insgesamt um einen random split tree, wenn man als S eine unabhängige Kopie von S^u nimmt. □

Lemma 21.
Der Splitvektor S aus Lemma 20 ist f.s. eine (eventuell zufällige) Permutation eines GEM$\left(\frac{\alpha}{\alpha+\beta}, \frac{\beta}{\alpha+\beta}\right)$-verteilten Vektors.

Beweis.
Sei $\left(S_j^u\right)_{j=1}^b := S^u$ und $N_{i,n} := \frac{|V(B_n^{v_i})|}{n}$.
Lemma 7 liefert, dass die relativen Teilbaumgrößen f.s. gegen die Komponenten des Splitvektors konvergieren. Wichtig ist es hier einzusehen, dass die Reihenfolge der v_i (also der ersten Treffer der Teilbäume von $\mathcal{T}$) nicht mit der Reihenfolge der Teilbäume in $\mathcal{T}$ übereinstimmen muss. Hier findet also eine (von der Realisierung abhängige) Permutation der Reihenfolge statt. Also gilt:
Es existiert f.s. eine (von der Realisierung abhängige) Permutaion π von $\mathbb{N}^+$, sodass für jedes $j \in [b]$ die Folge $N_{\pi(i),n}$ für $n \to \infty$ gegen S_j konvergiert.
Aus Lemma 19 wissen wir aber andererseits, dass $N_{i,n}$ für $n \to \infty$ f.s. gegen N_i konvergiert für einen GEM $\left(\frac{\alpha}{\alpha+\beta}, \frac{\beta}{\alpha+\beta}\right)$-verteilten Vektor $(N_i)_{i=1}^\infty$.
Folglich stimmen S^u und $\left(N_{\pi(j)}\right)_{j=1}^b$ f.s. überein und da S^u eine unabhängige Kopie von S ist, ist auch S f.s. eine (zufällige) Permutation eines GEM $\left(\frac{\alpha}{\alpha+\beta}, \frac{\beta}{\alpha+\beta}\right)$-verteilten Vektors. □

Beweis von Theorem 1.
Da es sich bei der Betrachtungsweise in Kapitel 3.2 lediglich um eine Umformulierung des preferential attachment trees handelt, folgt aus Lemma 20 und 21 auf Grund der Irrelevanz von Permutierungen des Splitvektors (siehe Anmerkung 2) bereits, dass $(B_n)_{n\in\mathbb{N}^+}$ und $(T_n)_{n\in\mathbb{N}^+}$ identisch verteilt sind, wenn man die Kinder jeweils nach der Reihenfolge des Erscheinens ordnet. □

5.3 Beweis von Theorem 2

Für dieses Unterkapitel seien immer die Voraussetzungen aus Theorem 2 ohne explizite Nennung gegeben. Zusätzlich sei an die in Anmerkung 6 erwähnte Notation bezüglich $\ln^{-1}(x) := \frac{1}{\ln(x)}$ erinnert.
Ferner gelten aufbauend auf der Intuition aus Kapitel 3.3 folgende Notationen:
Sei W_d die Größe der Komponente von der entsprechenden unabhängigen Kopie von S, die in Tiefe $d-1$ bei der Bestimmung des $(n+1)$-ten zu füllenden Knotens gewählt wird (wobei diese Folge zur Vereinfachung für alle $d \in \mathbb{N}^+$ weitergedacht werden soll). Folglich ist W_d eine unabhängige Kopie von W.
Sei F_d die Anzahl an Durchläufen unter den ersten n, die mit dem $(n+1)$-ten bis Tiefe d übereinstimmen würden, aber vorher bereits einen Knoten gefüllt haben.
Sei C_d die Anzahl an Durchläufen unter den ersten n, die mit dem $(n+1)$-ten bis Tiefe d übereinstimmen würden, wenn man vorzeitiges Füllen von Knoten ignorieren würde.
Nach den in Kapitel 3.3 ausgeführten Vorüberlegungen ist C_d, wenn man $(W_i)_{i=1}^d$ gegeben hat, Binomial$\left(n, \prod_{i=1}^d W_i\right)$-verteilt. Zusätzlich ist die Anzahl an Durchläufen unter den ersten n, die wirklich mit dem $(n+1)$-ten bis Tiefe d übereinstimmen, genau $C_d - F_d$ und folglich ist $D_{n+1} = \inf\{d \in \mathbb{N}^+ : C_d - F_d = 0\}$.
Abschließend sei $\hat{D}_c := \inf\left\{d \in \mathbb{N}^+ : \prod_{i=1}^d W_i \leq \frac{c}{n}\right\}$ für (möglicherweise von n abhängiges) $c \in \mathbb{R}^+$.

Anmerkung 17.
Da dies aus dem Kontext klar ist, wurde zur Übersichtlichkeit oft auf eine zusätzliche Indizierung mit n bzw. $n+1$ bei der Notation oben verzichtet. Trotzdem sei dieser Index im Folgenden immer implizit vorhanden und bei den betrachteten Grenzwerten mit berücksichtigt.

Lemma 22.
Sei f eine beliebige (deterministische oder von $(F_d)_{d=1}^\infty$ unabhängige) Funktion von $\mathbb{N}^+$ nach $\mathbb{N}^+$. Dann gilt:

$$\mathbb{P}\left(F_{f(n)} \geq \ln(\ln(n))\right) \underset{n\to\infty}{\longrightarrow} 0$$

Beweis.
$F_{f(n)}$ ist die Anzahl an Durchläufen unter den ersten n, die mit dem $(n+1)$-ten bis Tiefe $f(n)$ übereinstimmen würden, aber vorher bereits einen Knoten gefüllt haben. Oder anders formuliert: Die Anzahl an Durchläufen, die einen Knoten auf dem Weg zum $(n+1)$-ten gefüllten Knoten vor Tiefe $f(n)$ gefüllt haben und bis Tiefe $f(n)$ mit dem $(n+1)$-ten Durchlauf übereingestimmt hätten.
Also lässt sich für jeden gefüllten Knoten auf dem Pfad zum $(n+1)$-ten gefüllten Knoten vor Tiefe $f(n)$ fragen, ob der Durchgang, der diesen gefüllt hat, bis Tiefe $f(n)$ mit dem $(n+1)$-ten Durchlauf übereingestimmt hätte. Sei I_i für $1 \leq i \leq f(n)$ also der Indikator, ob der Durchlauf, der den Knoten in Tiefe $f(n)-i$ auf dem Pfad zum $(n+1)$-ten gefüllten

Knoten gefüllt hat, bis Tiefe $f(n)$ mit dem $(n+1)$-ten Durchlauf übereingestimmt hätte. Somit gilt insbesondere $F_{f(n)} = \sum_{i=1}^{f(n)} I_i$ und ferner, dass gegeben $(W_i)_{i=1}^{f(n)}$ die $(I_i)_{i=1}^{f(n)}$ unabhängig und jeweils Bernoulli$\left(\prod_{j=f(n)-i+1}^{f(n)} W_j\right)$-verteilt sind. Folglich lässt sich berechnen:

$$\begin{aligned}
\mathbb{E}\left[F_{f(n)}\right] &= \mathbb{E}\left[\sum_{i=1}^{f(n)} I_i\right] = \sum_{i=1}^{f(n)} \mathbb{E}\left[I_i\right] \overset{*1}{=} \sum_{i=1}^{f(n)} \mathbb{E}\left[\mathbb{E}\left[I_i \middle| W_1, \ldots, W_{f(n)}\right]\right] \\
&= \sum_{i=1}^{f(n)} \mathbb{E}\left[\prod_{j=f(n)-i+1}^{f(n)} W_j\right] \overset{*2}{=} \sum_{i=1}^{f(n)} \prod_{j=f(n)-i+1}^{f(n)} \mathbb{E}\left[W\right] \\
&= \sum_{i=1}^{f(n)} \mathbb{E}\left[W\right]^i \leq \sum_{i=1}^{\infty} \mathbb{E}\left[W\right]^i \overset{*3}{=} \frac{1}{1-\mathbb{E}\left[W\right]} < \infty
\end{aligned}$$

$*1$ Hier wird die Turm-Eigenschaft (Lemma 11) angewendet.
$*2$ Die W_j sind unabhängige Kopien von W.
$*3$ Nach Anmerkung 8 ist $\mathbb{E}[X] < 1$. Somit ist dies eine geometrische Reihe.

Und damit folgt mit Hilfe der Markov-Ungleichung (Lemma 9):

$$\mathbb{P}\left(F_{f(n)} \geq \ln(\ln(n))\right) \leq \frac{\frac{1}{1-\mathbb{E}[W]}}{\ln(\ln(n))} \underset{n\to\infty}{\longrightarrow} 0$$

□

Anmerkung 18.
Lemma 22 gilt auch, solange F_d und $\mathbb{1}_{\{f(n)=d\}}$ nicht positiv korreliert sind, da der Beweis nur die Erwartungswerte betrachtet.

Lemma 23.

$$\mathbb{P}\left(\hat{D}_{\ln(n)} \leq D_{n+1} \leq \hat{D}_{\ln^{-1}(n)}\right) \underset{n\to\infty}{\longrightarrow} 1$$

Beweis.
Zunächst lässt sich festhalten, dass aus der Definition von $\hat{D}_c$ direkt folgt, dass $\prod_{i=1}^{\hat{D}_{\ln^{-1}(n)}} W_i \leq \frac{1}{\ln(n)\cdot n}$ und $\prod_{i=1}^{\hat{D}_{\ln(n)}\ -1} W_i > \frac{\ln(n)}{n}$. Somit folgen insbesondere folgende Ungleichungen:

$$\mathbb{E}\left[C_{\hat{D}_{\ln^{-1}(n)}}\right] = n \cdot \prod_{i=1}^{\hat{D}_{\ln^{-1}(n)}} W_i \leq n \cdot \frac{1}{\ln(n)\cdot n} = \frac{1}{\ln(n)} \tag{3}$$

$$\mathbb{E}\left[C_{\hat{D}_{\ln(n)}\ -1}\right] = n \cdot \prod_{i=1}^{\hat{D}_{\ln(n)}\ -1} W_i > n \cdot \frac{\ln(n)}{n} = \ln(n) \tag{4}$$

$$\mathrm{Var}\left(C_{\hat{D}_{\ln(n)}\ -1}\right) = n \cdot \prod_{i=1}^{\hat{D}_{\ln(n)}\ -1} W_i \cdot \left(1 - \prod_{i=1}^{\hat{D}_{\ln(n)}\ -1} W_i\right) \leq \mathbb{E}\left[C_{\hat{D}_{\ln(n)}\ -1}\right] \tag{5}$$

Somit lässt sich nachrechnen:

$$1.\ \mathbb{P}\left(D_{n+1} > \hat{D}_{\ln^{-1}(n)}\right) \overset{*1}{\leq} \mathbb{P}\left(C_{\hat{D}_{\ln^{-1}(n)}} \geq 1\right) \overset{*0}{\leq} \frac{\mathbb{E}\left[C_{\hat{D}_{\ln^{-1}(n)}}\right]}{1} \overset{(3)}{=} \frac{\frac{1}{\ln(n)}}{1} = \frac{1}{\ln(n)} \underset{n\to\infty}{\longrightarrow} 0$$

$$\begin{aligned}
2.\ \mathbb{P}\left(D_{n+1} \geq \hat{D}_{\ln(n)}\right) &= \mathbb{P}\left(D_{n+1} > \hat{D}_{\ln(n)} - 1\right) = \mathbb{P}\left(C_{\hat{D}_{\ln(n)\ -1}} - F_{\hat{D}_{\ln(n)\ -1}} > 0\right) \\
&= \mathbb{P}\left(C_{\hat{D}_{\ln(n)\ -1}} > F_{\hat{D}_{\ln(n)\ -1}}\right) \\
&\geq \mathbb{P}\left(C_{\hat{D}_{\ln(n)\ -1}} > \ln(\ln(n)) > F_{\hat{D}_{\ln(n)\ -1}}\right) \\
&\overset{*2}{\geq} 1 - \mathbb{P}\left(C_{\hat{D}_{\ln(n)\ -1}} \leq \ln(\ln(n))\right) - \mathbb{P}\left(F_{\hat{D}_{\ln(n)\ -1}} \geq \ln(\ln(n))\right) \\
&\geq 1 - \mathbb{P}\left(\left|C_{\hat{D}_{\ln(n)\ -1}} - \mathbb{E}\left[C_{\hat{D}_{\ln(n)\ -1}}\right]\right| \geq \mathbb{E}\left[C_{\hat{D}_{\ln(n)\ -1}}\right] - \ln(\ln(n))\right) \\
&\quad - \mathbb{P}\left(F_{\hat{D}_{\ln(n)\ -1}} \geq \ln(\ln(n))\right) \\
&\overset{*0}{\geq} 1 - \frac{\mathrm{Var}\left(C_{\hat{D}_{\ln(n)\ -1}}\right)}{\left(\mathbb{E}\left[C_{\hat{D}_{\ln(n)\ -1}}\right] - \ln(\ln(n))\right)^2} - \mathbb{P}\left(F_{\hat{D}_{\ln(n)\ -1}} \geq \ln(\ln(n))\right) \\
&\overset{(5)}{\geq} 1 - \frac{\mathbb{E}\left[C_{\hat{D}_{\ln(n)\ -1}}\right]}{\left(\mathbb{E}\left[C_{\hat{D}_{\ln(n)\ -1}}\right] - \ln(\ln(n))\right)^2} - \mathbb{P}\left(F_{\hat{D}_{\ln(n)\ -1}} \geq \ln(\ln(n))\right) \\
&\overset{(4)}{\geq} 1 - \frac{\ln(n)}{\left(\ln(n) - \ln(\ln(n))\right)^2} - \mathbb{P}\left(F_{\hat{D}_{\ln(n)\ -1}} \geq \ln(\ln(n))\right) \\
&= 1 - \frac{1}{\left(\sqrt{\ln(n)} - \frac{\ln(\ln(n))}{\sqrt{\ln(n)}}\right)^2} - \mathbb{P}\left(F_{\hat{D}_{\ln(n)\ -1}} \geq \ln(\ln(n))\right) \\
&\underset{n\to\infty}{\overset{*3}{\longrightarrow}} 1 - 0 - 0 = 1
\end{aligned}$$

$*0$ Hier wird die Markov- bzw. Chebyshev-Ungleichung (Lemma 9 und 10) angewendet.
$*1$ $D_{n+1} > \hat{D}_{\ln^{-1}(n)}$ ist nach den Vorüberlegungen zu Beginn dieses Unterkapitels äquivalent zu $C_{\hat{D}_{\ln^{-1}(n)}} - F_{\hat{D}_{\ln^{-1}(n)}} \geq 1$.
$*2$ Hier wird die Gegenwahrscheinlichkeit und die Subadditivität verwendet.
$*3$ Zum Einen gilt $\ln(\ln(n)) \in o\left(\sqrt{\ln(n)}\right)$. Zum Anderen lässt sich Lemma 22 anwenden, da $\hat{D}_{\ln(n)}$ nur von $(W_i)_{i=1}^{\infty}$ abhängt und mit erhöhter Wahrscheinlichkeit nach kleineren Werten von W_i auftritt (siehe Anmerkung 18).

Nun folgt hieraus:

$$\mathbb{P}\left(\hat{D}_{\ln(n)} \leq D_{n+1} \leq \hat{D}_{\ln^{-1}(n)}\right) \geq \mathbb{P}\left(D_{n+1} \geq \hat{D}_{\ln(n)}\right) - \mathbb{P}\left(D_{n+1} > \hat{D}_{\ln^{-1}(n)}\right) \underset{n\to\infty}{\longrightarrow} 1 - 0 = 1$$

□

Lemma 24.
Sei f eine deterministische Funktion von $\mathbb{N}^+$ nach $\mathbb{N}^+$ mit $f(n) \xrightarrow[n\to\infty]{} \infty$. Dann gilt:

$$\frac{\sum_{d=1}^{f(n)} \ln(W_d) + f(n)\cdot\mu}{\sqrt{f(n)\cdot\sigma^2}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1)$$

Beweis.
Die $(W_d)_{d=1}^{\infty}$ sind u.i.v. und somit sind auch die $(\ln(W_d))_{d=1}^{\infty}$ u.i.v.. Ferner ist per Definition $\mathbb{E}[\ln(W_d)] = -\mu$ und $\mathrm{Var}(\ln(W_d)) = \sigma^2$ und zusätzlich gilt per Voraussetzung $\mathbb{E}\left[\ln^2(W_d)\right] < \infty$ und somit folgt mit dem Zentralen Grenzwertsatz (Lemma 8) direkt:

$$\frac{\sum_{d=1}^{n} \ln(W_d) + n\cdot\mu}{\sqrt{n\cdot\sigma^2}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1)$$

Da $f(n)$ eine Folge in $\mathbb{N}^+$ ist, welche für $n \to \infty$ gegen ∞ geht, folgt hieraus direkt die gefragte Aussage. □

Lemma 25.

1. $$\frac{\hat{D}_{\ln(n)} - \frac{\ln(n)}{\mu}}{\sigma\sqrt{\frac{\ln(n)}{\mu^3}}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1)$$
2. $$\frac{\hat{D}_{\ln^{-1}(n)} - \frac{\ln(n)}{\mu}}{\sigma\sqrt{\frac{\ln(n)}{\mu^3}}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1)$$

Beweis.
Die Konvergenz in Verteilung ist bekanntlich äquivalent zur punktweisen Konvergenz der Verteilungsfunktion. Hierfür lässt sich mit $c(n) \in \left\{\ln(n), \ln^{-1}(n)\right\}$ und $\hat{f}(n) := \left\lfloor x\cdot\sigma\sqrt{\frac{\ln(n)}{\mu^3}} + \frac{\ln(n)}{\mu} \right\rfloor$ folgendes berechnen ($\mu, \sigma > 0$):

$$\begin{aligned}
F_{\frac{\hat{D}_{c(n)} - \frac{\ln(n)}{\mu}}{\sigma\sqrt{\frac{\ln(n)}{\mu^3}}}}(x) &= \mathbb{P}\left(\frac{\hat{D}_{c(n)} - \frac{\ln(n)}{\mu}}{\sigma\sqrt{\frac{\ln(n)}{\mu^3}}} \le x\right) = \mathbb{P}\left(\hat{D}_{c(n)} \le x\cdot\sigma\sqrt{\frac{\ln(n)}{\mu^3}} + \frac{\ln(n)}{\mu}\right) \\
&= \mathbb{P}\left(\prod_{i=1}^{\left\lfloor x\cdot\sigma\sqrt{\frac{\ln(n)}{\mu^3}} + \frac{\ln(n)}{\mu}\right\rfloor} W_i \le \frac{c(n)}{n}\right) = \mathbb{P}\left(\sum_{i=1}^{\hat{f}(n)} \ln(W_i) \le \ln\left(\frac{c(n)}{n}\right)\right) \\
&= \mathbb{P}\left(\frac{\sum_{i=1}^{\hat{f}(n)} \ln(W_i) + \hat{f}(n)\cdot\mu}{\sqrt{\hat{f}(n)\cdot\sigma^2}} \le \frac{\ln\left(\frac{c(n)}{n}\right) + \hat{f}(n)\cdot\mu}{\sqrt{\hat{f}(n)\cdot\sigma^2}}\right)
\end{aligned}$$

Lemma 24 (mit $f(n) = \hat{f}(n) = \left\lfloor x \cdot \sigma\sqrt{\frac{\ln(n)}{\mu^3}} + \frac{\ln(n)}{\mu} \right\rfloor$) liefert, dass der linke Term der Ungleichung in Verteilung gegen die Standard-Normalverteilung konvergiert. Da die Verteilungsfunktion der Standard-Normalverteilung stetig ist, lässt sich der Grenzwert des rechten Terms der Ungleichung separat berechnen und dann später in die Verteilungsfunktion der Standard-Normalverteilung einsetzen:

$$
\begin{aligned}
&\lim_{n\to\infty} \left(\frac{\ln\left(\frac{c(n)}{n}\right) + \left\lfloor x \cdot \sigma\sqrt{\frac{\ln(n)}{\mu^3}} + \frac{\ln(n)}{\mu} \right\rfloor \cdot \mu}{\sqrt{\left\lfloor x \cdot \sigma\sqrt{\frac{\ln(n)}{\mu^3}} + \frac{\ln(n)}{\mu} \right\rfloor \cdot \sigma^2}} \right) \\
&\overset{*1}{=} \lim_{n\to\infty} \left(\frac{\ln\left(c(n)\right) - \ln(n) + x \cdot \sigma\sqrt{\frac{\ln(n)}{\mu}} + \ln(n)}{\sigma\sqrt{x \cdot \sigma\sqrt{\frac{\ln(n)}{\mu^3}} + \frac{\ln(n)}{\mu}}} \right) \\
&= \lim_{n\to\infty} \left(\frac{\pm\ln\left(\ln(n)\right)}{\sigma\sqrt{x \cdot \sigma\sqrt{\frac{\ln(n)}{\mu^3}} + \frac{\ln(n)}{\mu}}} + \frac{x \cdot \sigma\sqrt{\frac{\ln(n)}{\mu}}}{\sigma\sqrt{x \cdot \sigma\sqrt{\frac{\ln(n)}{\mu^3}} + \frac{\ln(n)}{\mu}}} \right) \\
&\overset{*2}{=} \lim_{n\to\infty} \left(0 + \frac{x \cdot \sqrt{\frac{\ln(n)}{\mu}}}{\sqrt{x \cdot \sigma\sqrt{\frac{\ln(n)}{\mu^3}} + \frac{\ln(n)}{\mu}}} \right) \\
&= \lim_{n\to\infty} \left(x \cdot \sqrt{\frac{\ln(n)}{x \cdot \sigma\sqrt{\frac{\ln(n)}{\mu}} + \ln(n)}} \right) \\
&= \lim_{n\to\infty} \left(x \cdot \sqrt{\frac{1}{\frac{x\cdot\sigma}{\sqrt{\mu\cdot\ln(n)}} + 1}} \right) = x
\end{aligned}
$$

$*1$ Da sowohl Zähler als auch Nenner für $n \to \infty$ gegen ∞ konvergieren, ist das Abrunden hier asymptotisch irrelevant.
$*2$ Es gilt $\ln\left(\ln(n)\right) \in o\left(\sqrt{\ln(n)}\right)$.

Durch Einsetzen folgt nun:

$$
\lim_{n\to\infty} \left(F_{\frac{\hat{D}_{c(n)} - \frac{\ln(n)}{\mu}}{\sigma\sqrt{\frac{\ln(n)}{\mu^3}}}}(x) \right) = F_{\mathcal{N}(0,1)}(x)
$$

□

Beweis von Theorem 2.
Aus Lemma 23 wissen wir, dass $\mathbb{P}\left(\hat{D}_{\ln(n)} \leq D_{n+1} \leq \hat{D}_{\ln^{-1}(n)}\right) \underset{n\to\infty}{\longrightarrow} 1$. Dies impliziert insbesondere:

$$\mathbb{P}\left(\frac{\hat{D}_{\ln(n)} - \frac{\ln(n)}{\mu}}{\sigma\sqrt{\frac{\ln(n)}{\mu^3}}} \leq \frac{D_{n+1} - \frac{\ln(n)}{\mu}}{\sigma\sqrt{\frac{\ln(n)}{\mu^3}}} \leq \frac{\hat{D}_{\ln^{-1}(n)} - \frac{\ln(n)}{\mu}}{\sigma\sqrt{\frac{\ln(n)}{\mu^3}}}\right) \underset{n\to\infty}{\longrightarrow} 1$$

Dies impliziert insbesondere, dass für jedes $x \in \mathbb{R}$ beide folgenden Ungleichungen gelten (da sonst x als Gegenbeispiel für die obige Aussage genommen werden könnte):

$$1.\ \lim_{n\to\infty}\left(\mathbb{P}\left(\frac{D_{n+1} - \frac{\ln(n)}{\mu}}{\sigma\sqrt{\frac{\ln(n)}{\mu^3}}} \leq x\right)\right) \geq \lim_{n\to\infty}\left(\mathbb{P}\left(\frac{\hat{D}_{\ln^{-1}(n)} - \frac{\ln(n)}{\mu}}{\sigma\sqrt{\frac{\ln(n)}{\mu^3}}} \leq x\right)\right)$$

$$2.\ \lim_{n\to\infty}\left(\mathbb{P}\left(\frac{D_{n+1} - \frac{\ln(n)}{\mu}}{\sigma\sqrt{\frac{\ln(n)}{\mu^3}}} \leq x\right)\right) \leq \lim_{n\to\infty}\left(\mathbb{P}\left(\frac{\hat{D}_{\ln(n)} - \frac{\ln(n)}{\mu}}{\sigma\sqrt{\frac{\ln(n)}{\mu^3}}} \leq x\right)\right)$$

Nun folgt aus Lemma 25, dass die obere und untere Schranke gegen die Verteilungsfunktion der Standard-Normalverteilung konvergieren, was dies auch für den eingeschränkten Term impliziert, welcher nichts anderes als die Verteilungsfunktion von $\frac{D_{n+1} - \frac{\ln(n)}{\mu}}{\sigma\sqrt{\frac{\ln(n)}{\mu^3}}}$ ist.
Da die Konvergenz in Verteilung bekanntlich äquivalent zur punktweisen Konvergenz der Verteilungsfunktion ist, zeigt dies bereits die zu zeigende Aussage. □

5.4 Beweis von Lemma 3

Wie in Lemma 3 sei in diesem Unterkapitel $0 \leq \alpha < 1$ und $\alpha + \beta \geq 0$. Ferner sei $P = (P_i)$ GEM (α, β)-verteilt mit $P_i := Z_i \cdot \prod_{j=1}^{i-1} (1 - Z_j)$ für Beta $(1 - \alpha, (\alpha \cdot j + \beta) \vee 0)$-verteilte Z_j wie in Definition 6.

Lemma 26.
Sei $c > 0$, dann gilt:

$$\mathbb{E}\left[(P_i)^c\right] = \frac{\Gamma(1-\alpha+c)\cdot\Gamma(\beta+1)}{\Gamma(1-\alpha)\cdot\Gamma(\beta+1+c)} \cdot \prod_{j=1}^{i-1} \frac{\alpha\cdot j+\beta}{\alpha\cdot j+\beta+c}$$

Beweis.

$$\mathbb{E}\left[(P_i)^c\right] = \mathbb{E}\left[\left(Z_i \cdot \prod_{j=1}^{i-1}(1-Z_j)\right)^c\right] = \mathbb{E}\left[(Z_i)^c \cdot \prod_{j=1}^{i-1}(1-Z_j)^c\right] \overset{*1}{=} \mathbb{E}\left[(Z_i)^c\right] \cdot \prod_{j=1}^{i-1} \mathbb{E}\left[(1-Z_j)^c\right]$$

$$\overset{*2}{=} \frac{\Gamma(1-\alpha+c)\cdot\Gamma(1-\alpha+\alpha\cdot i+\beta)}{\Gamma(1-\alpha)\cdot\Gamma(1-\alpha+\alpha\cdot i+\beta+c)} \cdot \prod_{j=1}^{i-1} \frac{\Gamma(\alpha\cdot j+\beta+c)\cdot\Gamma(1-\alpha+\alpha\cdot j+\beta)}{\Gamma(\alpha\cdot j+\beta)\cdot\Gamma(1-\alpha+\alpha\cdot j+\beta+c)}$$

$$= \frac{\Gamma(1-\alpha+c)\cdot\Gamma(\alpha\cdot(i-1)+\beta+1)}{\Gamma(1-\alpha)\cdot\Gamma(\alpha\cdot(i-1)+\beta+c+1)} \cdot \prod_{j=1}^{i-1} \frac{\Gamma(\alpha\cdot j+\beta+c)\cdot\Gamma(\alpha\cdot(j-1)+\beta+1)}{\Gamma(\alpha\cdot(j-1)+\beta+c+1)\cdot\Gamma(\alpha\cdot j+\beta)}$$

$$\overset{*3}{=} \frac{\Gamma(1-\alpha+c)\cdot\Gamma(\alpha\cdot 0+\beta+1)}{\Gamma(1-\alpha)\cdot\Gamma(\alpha\cdot 0+\beta+c+1)} \cdot \prod_{j=1}^{i-1} \frac{\Gamma(\alpha\cdot j+\beta+c)\cdot\Gamma(\alpha\cdot j+\beta+1)}{\Gamma(\alpha\cdot j+\beta+c+1)\cdot\Gamma(\alpha\cdot j+\beta)}$$

$$\overset{*4}{=} \frac{\Gamma(1-\alpha+c)\cdot\Gamma(\beta+1)}{\Gamma(1-\alpha)\cdot\Gamma(\beta+1+c)} \cdot \prod_{j=1}^{i-1} \frac{\alpha\cdot j+\beta}{\alpha\cdot j+\beta+c}$$

$*1$ Hier wird ausgenutzt, dass die Z_j unabhängig sind.
$*2$ Lemma 14 wird angewendet.
$*3$ Verschiebung der Faktoren im Produkt.
$*4$ Die Identität $\frac{\Gamma(x+1)}{\Gamma(x)} = x$ wird verwendet. □

Beweis von Lemma 3.
Per Konstruktion gilt $\sum_{i=1}^{\infty} P_i = 1$ f.s.. Also gilt auch:

$$\begin{aligned}
1 =& \mathbb{E}[1] = \mathbb{E}\left[\sum_{i=1}^{\infty} P_i\right] = \sum_{i=1}^{\infty} \mathbb{E}\left[(P_i)^1\right] \\
&\overset{*1}{=} \sum_{i=1}^{\infty} \frac{\Gamma(1-\alpha+1)\cdot\Gamma(\beta+1)}{\Gamma(1-\alpha)\cdot\Gamma(\beta+1+1)} \cdot \prod_{j=1}^{i-1} \frac{\alpha\cdot j+\beta}{\alpha\cdot j+\beta+1} \\
&\overset{*2}{=} \sum_{i=1}^{\infty} \frac{1-\alpha}{\beta+1} \cdot \prod_{j=1}^{i-1} \frac{\alpha\cdot j+\beta}{\alpha\cdot j+\beta+1} \\
\Leftrightarrow \quad & \frac{1+\beta}{1-\alpha} = \sum_{i=1}^{\infty} \prod_{j=1}^{i-1} \frac{\alpha\cdot j+\beta}{\alpha\cdot j+\beta+1}
\end{aligned}$$

$*1$ Lemma 26 wird angewendet.
$*2$ Die Identität $\frac{\Gamma(x+1)}{\Gamma(x)} = x$ wird verwendet. □

5.5 Beweis von Lemma 2

Wie in Definition 6 (und damit Lemma 2) seien in diesem Unterkapitel $\alpha, \beta \in \mathbb{R}$, sodass folgende drei Eigenschaften erfüllt sind:
1.) $\alpha < 1$ 2.) $\alpha + \beta \geq 0$ 3.) $\alpha \geq 0$ oder $\alpha \cdot n + \beta = 0$ für ein $n \in \mathbb{N}^+$.
Ebenfalls wie in Definition 6 sei ferner $P = (P_i)_{i=1}^{\infty}$ hier GEM (α, β)-verteilt mit:
$P_i = Z_i \cdot \prod_{j=1}^{i-1} (1 - Z_j)$ für Beta $(1 - \alpha, (\alpha \cdot j + \beta) \vee 0)$-verteilte Z_j.

Zunächst kann man $\mathbb{E}\left[\ln^2(W)\right]$ mit Hilfe der Turm-Eigenschaft (Lemma 11) wie folgt umformen:

$$\mathbb{E}\left[\ln^2(W)\right] = \mathbb{E}\left[\mathbb{E}\left[\ln^2(W)|P\right]\right] = \mathbb{E}\left[\sum_{i=1}^{\infty} P_i \ln^2(P_i)\right] = \sum_{i=1}^{\infty} \mathbb{E}\left[P_i \ln^2(P_i)\right]$$

Wobei die übliche Konvention $0 \ln(0) = 0 \ln^2(0) = 0$ gelte, da dies die stetige Fortsetzung von $x \ln(x)$ bzw. $x \ln^2(x)$ in der 0 ist. Damit genügt es für Lemma 2 zu zeigen:

$$\sum_{i=1}^{\infty} \mathbb{E}\left[P_i \ln^2(P_i)\right] < \infty$$

Da der Beweis hierfür für negative α anders funktioniert als für nichtnegative, werden nun die Fälle $\alpha < 0$ und $\alpha \geq 0$ jeweils separat abgearbeitet:

5.5.1 Fall: $\alpha < 0$

Beweis von Lemma 2 für $\alpha < 0$.
In diesem Fall gibt es ein $n \in \mathbb{N}^+$, sodass $\alpha \cdot n + \beta = 0$. Somit ist $Z_n \overset{\text{f.s.}}{=} 1$ und damit $P_i \overset{\text{f.s.}}{=} 0$ für alle $i > n$. Folglich hat P f.s. höchstens n Einträge, die $\neq 0$ sind. Also gilt:

$$\sum_{i=1}^{\infty} \mathbb{E}\left[P_i \ln^2(P_i)\right] = \sum_{i=1}^{n} \mathbb{E}\left[P_i \ln^2(P_i)\right]$$

Da $\lim_{x \downarrow 0}\left(x \cdot \ln^2(x)\right) = 0$ und ferner $x \cdot \ln^2(x)$ stetig auf $(0, 1]$ ist, ist $x \cdot \ln^2(x)$ auf $(0, 1]$ beschränkt. Somit existiert ein $k \in \mathbb{R}^+$ sodass $x \cdot \ln^2(x) \leq k$ für alle $x \in (0, 1]$. Folglich gilt:

$$\sum_{i=1}^{n} \mathbb{E}\left[P_i \ln^2(P_i)\right] \leq \sum_{i=1}^{n} \mathbb{E}\left[k\right] = n \cdot k < \infty$$

□

5.5.2 Fall: $\alpha \geq 0$

Hier gilt insbesondere $\alpha \cdot j + \beta \geq 0$ für alle $j \geq 1$, was die Z_j schlicht Beta $(1 - \alpha, \alpha \cdot j + \beta)$-verteilt macht. Zunächst möchten wir aber $x \cdot \ln^2(x)$ durch $x^{1-\varepsilon}$ abschätzen:

Lemma 27.
Sei $\varepsilon > 0$, dann gilt:

$$\lim_{x \downarrow 0} \frac{x \cdot \ln^2(x)}{x^{1-\varepsilon}} = 0$$

Beweis.
Durch wiederholtes Anwenden von der Regel von de L'Hospital (Lemma 12) (die Stellen sind jeweils mit $*$ gekennzeichnet) folgt (OE: $\varepsilon < 1$):

$$\begin{aligned}
\lim_{x \downarrow 0} \frac{x \cdot \ln^2(x)}{x^{1-\varepsilon}} &\overset{*}{=} \lim_{x \downarrow 0} \frac{\ln^2(x) + x \cdot 2 \cdot \ln(x) \cdot \frac{1}{x}}{(1-\varepsilon) \cdot x^{-\varepsilon}} = \lim_{x \downarrow 0} \frac{\ln^2(x) + 2 \cdot \ln(x)}{(1-\varepsilon) \cdot x^{-\varepsilon}} \\
&\overset{*}{=} \lim_{x \downarrow 0} \frac{2 \cdot \ln(x) \cdot \frac{1}{x} + 2 \cdot \frac{1}{x}}{-\varepsilon \cdot (1-\varepsilon) \cdot x^{-\varepsilon-1}} = \lim_{x \downarrow 0} \frac{2 \cdot \ln(x) + 2}{-\varepsilon \cdot (1-\varepsilon) \cdot x^{-\varepsilon}} \\
&\overset{*}{=} \lim_{x \downarrow 0} \frac{2 \cdot \frac{1}{x}}{\varepsilon^2 \cdot (1-\varepsilon) \cdot x^{-\varepsilon-1}} = \lim_{x \downarrow 0} \frac{2}{\varepsilon^2 \cdot (1-\varepsilon) \cdot x^{-\varepsilon}} \\
&= 0
\end{aligned}$$

□

Lemma 28.
Sei $\varepsilon > 0$, dann gilt:

$$\exists x_0 \in (0,1)\ \forall x \in (0, x_0] : x \cdot \ln^2(x) \leq x^{1-\varepsilon}$$

Beweis.
Nach Lemma 27 ist $\lim_{x \downarrow 0} \frac{x \cdot \ln^2(x)}{x^{1-\varepsilon}} = 0$. Folglich gibt es ein $x_0 \in (0,1)$, sodass für alle $x \in (0, x_0]$ gilt:

$$\begin{aligned}
&\frac{x \cdot \ln^2(x)}{x^{1-\varepsilon}} \leq 1 \\
\Leftrightarrow \quad &x \cdot \ln^2(x) \leq x^{1-\varepsilon}
\end{aligned}$$

□

Lemma 29.
Sei $\varepsilon > 0$, dann gilt:

$$\exists c \in \mathbb{R}^+\ \forall x \in (0,1] : x \cdot \ln^2(x) \leq c \cdot x^{1-\varepsilon}$$

Beweis.
Da $\lim_{x \downarrow 0} \left(x \cdot \ln^2(x) \right) = 0$ und ferner $x \cdot \ln^2(x)$ stetig auf $(0,1]$ ist, ist $x \cdot \ln^2(x)$ auf $(0,1]$

beschränkt. Somit existiert ein $k \in \mathbb{R}^+$, sodass $x \cdot \ln^2(x) < k$ für alle $x \in (0,1]$.
Ferner ist $x^{1-\varepsilon}$ monoton wachsend und echt positiv auf $(0,1]$. Sei nun x_0 entsprechend Lemma 28 und $c = \frac{k}{x_0^{1-\varepsilon}} \vee 1$, dann gilt für $x \in (x_0, 1]$:

$$x \cdot \ln^2(x) \leq k \leq c \cdot x_0^{1-\varepsilon} \leq c \cdot x^{1-\varepsilon}$$

Ferner gilt für $x \in (0, x_0]$ nach Lemma 28:

$$x \cdot \ln^2(x) \leq x^{1-\varepsilon} \leq c \cdot x^{1-\varepsilon}$$

□

Lemma 30.
Sei $c > 0$ und $\frac{\alpha}{c} < 1$, dann gilt:

$$\sum_{i=1}^{\infty} \mathbb{E}\left[(P_i)^c\right] = \frac{\Gamma(1-\alpha+c) \cdot \Gamma(\beta+1)}{\Gamma(1-\alpha) \cdot \Gamma(\beta+1+c)} \cdot \frac{1+\frac{\beta}{c}}{1-\frac{\alpha}{c}} < \infty$$

Beweis.

$$\begin{aligned}
\sum_{i=1}^{\infty} \mathbb{E}\left[(P_i)^c\right] &\overset{*1}{=} \sum_{i=1}^{\infty} \frac{\Gamma(1-\alpha+c) \cdot \Gamma(\beta+1)}{\Gamma(1-\alpha) \cdot \Gamma(\beta+1+c)} \cdot \prod_{j=1}^{i-1} \frac{\alpha \cdot j + \beta}{\alpha \cdot j + \beta + c} \\
&= \frac{\Gamma(1-\alpha+c) \cdot \Gamma(\beta+1)}{\Gamma(1-\alpha) \cdot \Gamma(\beta+1+c)} \cdot \sum_{i=1}^{\infty} \prod_{j=1}^{i-1} \frac{\frac{\alpha}{c} \cdot j + \frac{\beta}{c}}{\frac{\alpha}{c} \cdot j + \frac{\beta}{c} + 1} \\
&\overset{*2}{=} \frac{\Gamma(1-\alpha+c) \cdot \Gamma(\beta+1)}{\Gamma(1-\alpha) \cdot \Gamma(\beta+1+c)} \cdot \frac{1+\frac{\beta}{c}}{1-\frac{\alpha}{c}} < \infty
\end{aligned}$$

$*1$ Lemma 26 wird angewendet.
$*2$ Lemma 3 wird angewendet. □

Beweis von Lemma 2 für $\alpha \geq 0$.
Zunächst lässt sich $\sum_{i=1}^{\infty} \mathbb{E}\left[P_i \ln^2(P_i)\right]$ wie folgt mit Hilfe von Lemma 29 abschätzen:

$$\sum_{i=1}^{\infty} \mathbb{E}\left[P_i \ln^2(P_i)\right] \leq \sum_{i=1}^{\infty} \mathbb{E}\left[c(\varepsilon) \cdot (P_i)^{1-\varepsilon}\right] = c(\varepsilon) \cdot \sum_{i=1}^{\infty} \mathbb{E}\left[(P_i)^{1-\varepsilon}\right]$$

Wenn man nun $\varepsilon = \frac{1-\alpha}{2}$ wählt, gilt $\frac{\alpha}{1-\varepsilon} = \frac{\alpha}{\frac{1}{2}+\frac{\alpha}{2}} = \frac{2\alpha}{1+\alpha} < 1$ und man erhält mit Lemma 30:

$$\begin{aligned}
\sum_{i=1}^{\infty} \mathbb{E}\left[P_i \ln^2(P_i)\right] &\leq c(\varepsilon) \cdot \sum_{i=1}^{\infty} \mathbb{E}\left[(P_i)^{1-\varepsilon}\right] \\
&= c(\varepsilon) \cdot \frac{\Gamma(1-\alpha+1-\varepsilon) \cdot \Gamma(\beta+1)}{\Gamma(1-\alpha) \cdot \Gamma(\beta+1+1-\varepsilon)} \cdot \frac{1+\frac{\beta}{1-\varepsilon}}{1-\frac{\alpha}{1-\varepsilon}} \\
&< \infty
\end{aligned}$$

□

Literatur

[1] Svante Janson. *Random Recursive Trees and Preferential Attachment Trees are Random Split Trees*. Combinatorics, Probability and Computing **28** (2019), no. 1, 81-99.

[2] Luc Devroye. *Universal limit laws for depths in random trees*. SIAM J. Comput. **28** (1999), no. 2, 409-432.

[3] Götz Kersting, Anton Wakolbinger. *Stochastische Prozesse*. Mathematik Kompakt, Birkhäuser, Springer, Basel, 2014.

[4] Olav Kallenberg. *Foundations of Modern Probability*. 2nd ed., Springer, New York, 2002.

[5] Rick Durett. *Probability Models for DNA Sequence Evolution*. 2nd ed., Springer, New York, 2008.

[6] Hsien-Kuei Hwang, Michael Fuchs, Vytas Zacharovas. *Asymptotic variance of random symmetric digital search trees*. Discrete Mathematics and Theoretical Computer Science DMTCS vol. 12:2, 2010, 103–166.
(arXiv: `https://arxiv.org/pdf/1001.0095.pdf`)

Erklärung zur Masterarbeit:

Hiermit erkläre ich, dass ich die vorliegende Masterarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und die aus fremden Quellen direkt oder indirekt übernommenen Gedanken als solche kenntlich gemacht habe. Die Arbeit habe ich bisher keinem anderen Prüfungsamt in gleicher oder vergleichbarer Form vorgelegt. Sie wurde bisher nicht veröffentlicht.

Frankfurt, den .....

Thomas Fischer