

Diplomarbeit

Ein Grenzwertsatz für die letzte
Verschmelzung im
Bolthausen-Sznitman Koaleszenten

Michael Messer

Goethe-Universität
Fachbereich Informatik und Mathematik
Institut für Mathematik
Frankfurt am Main

Betreuer: Prof. Dr. Ralph Neininger

Dezember 2008

Erklärung

Hiermit erkläre ich, Michael Messer, die vorliegende Diplomarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt zu haben.

Inhaltsverzeichnis

1	Einleitung	i
2	Vorbetrachtungen	1
2.1	Notationen	1
2.2	Die Exponentialverteilung und Yule-Prozesse	2
3	BSC - RRT - CRP	12
3.1	Der Bolthausen-Sznitman Koaleszent	12
3.1.1	Der restringierte Koaleszent	12
3.1.2	Der unrestringierte Koaleszent	15
3.1.3	Eine Konstruktion des Koaleszenten	17
3.2	Zufällige rekursive Bäume	19
3.3	Eine Größenbetrachtung der Abkömmlinge	26
3.4	Der Bonsai im Chinarestaurant	34
4	Die letzte Verschmelzung im BSC	40
4.1	Der Hauptsatz	40
4.1.1	Notation und Konstruktion im Hauptbeweis	42
4.1.2	Hilfslemmata zum Hauptbeweis	47
4.1.3	Der Hauptbeweis	48
4.1.4	Beweise der Hilfslemmata	52
4.2	Folgerungen	61
5	Literaturverzeichnis	64

1 Einleitung

Der Bolthausen-Sznitman Koaleszent ist ein Spezialfall der von Pitman und von Sagitov im Jahre 1999 eingeführten Λ -Koaleszenten (vgl.[7] und [15]). Für jedes endliche Maß Λ – im Falle des Bolthausen-Sznitman Koaleszenten ist Λ gleich der uniformen Verteilung auf $[0, 1]$ – ist ein solcher Koaleszent ein Markovprozess mit Werten in der Menge der Partitionen der natürlichen Zahlen, dessen Dynamik durch Λ festgelegt wird. Die Besonderheit eines Koaleszenten gegenüber anderen Prozessen liegt darin, dass nur Übergänge in gröbere Partitionen möglich sind. Verstehen wir die Teilmengen der Partition als Blöcke, so beschreibt ein Übergang die Abnahme der Blockanzahl durch "Verschmelzen" gewisser Blöcke zu einem gemeinsamen größeren Block. Dies wirft die Frage auf, *ob* und – wenn ja – *wann* der Koaleszent in den finalen Zustand von nur noch einem Block übergeht.

Wir werden den Bolthausen-Sznitman Koaleszenten auf die ersten n natürlichen Zahlen restringieren und unser Augenmerk auf den Zeitpunkt der letzten Kollision richten. Das Hauptresultat ist ein Grenzwertsatz, der zum einen die Konvergenz der Verteilung der Blockanzahl beschreibt, und zum anderen, in Abhängigkeit von der Blockanzahl, in gewisser Weise eine Aussage über die Verteilung der Blockgrößen macht. Die betrachteten Größen beziehen sich dabei auf den Zeitpunkt der finalen Kollision. Für großes n wird sich dabei typischerweise ein Block besonders auszeichnen und zwar insofern, dass er im zeitlichen Verlauf fast die ganze Masse in sich vereint. Neben diesem riesigen Block werden dann nur noch wenige kleine Blöcke in der letzten Kollision involviert sein.

Wir werden nicht mit dem Bolthausen-Sznitman Koaleszenten selbst arbeiten, sondern diesen durch ein gewisses "Abholzverfahren" zufälliger rekursiver Bäume modellieren. Dieses Modell wird speziell für die Betrachtung der letzten Kollision geeignet sein.

Hinsichtlich unseres Hauptsatzes werden wir in einem weiteren Schritt den "Aufbau" der zufälligen rekursiven Bäume durch Yule Prozesse modellieren.

Die Arbeit ist in drei Abschnitte unterteilt. Zunächst werden wir grundlegende Notationen einführen und Eigenschaften von Yule-Prozessen diskutieren. Im zweiten Teil werden wir den Bolthausen-Sznitman Koaleszent definieren und dann die angesprochene Baumstruktur einführen und diskutieren. Im Anschluß gehen wir noch auf den Zusammenhang zwischen zufälligen rekursiven Bäumen und dem sogenannten China-Restaurant-Prozess ein. Der letzte Teil dient dem Hauptsatz dieser Arbeit. In Anlehnung an die Baumstruktur werden wir mit Hilfe der Yule-Prozesse ein neues Modell konstruieren, mit dem wir dann den Hauptsatz beweisen. Abschließend geben wir noch einige Folgerungen an.

Im Grundlegenden folgen wir bei dieser Arbeit Teilen von [12], [7] und [8], ohne darauf immer wieder explizit zu verweisen.

2 Vorbetrachtungen

2.1 Notationen

Wir setzen die üblichen in der Stochastik verwendeten Begriffe, wie etwa die der *Dichte* einer Zufallsvariablen (ZV) oder die *Unabhängigkeit von Zufallsvariablen*, wie auch die gängigen *Konvergenzarten von Folgen von ZVn* als bekannt voraus.

Seien X, Y Zufallsvariablen mit beliebigem Grund-W-Raum (meist mit (Ω, \mathcal{A}, P) bezeichnet) und gemeinsamen (messbaren) Bildraum (Ω', \mathcal{A}') . Dann bezeichnen wir das von X induzierte Bildmaß mit P_X .

Wir nennen X und Y in Verteilung gleich, wenn gilt $P_X(A') = P_Y(A')$, $\forall A' \in \mathcal{A}'$ und schreiben dafür $X \sim Y$.

Für die Arbeit mit stochastischen Prozessen, die Werte im gemeinsamen separablen metrischen Raum S annehmen, erwähnen wir den Raum der càdlàg-Funktionen $D_S[0, \infty)$, welchen wir durch den bekannten Skorohod-Abstand metrisieren. Wir bemerken, dass die von den offenen Mengen in $D_S[0, \infty)$ erzeugte \mathfrak{G} -Algebra bekanntlich mit der kanonischen (von den Projektionsabbildungen erzeugten) Produkt- \mathfrak{G} -Algebra übereinstimmt.

Zum Begriff der Dichte sei noch erwähnt, dass aus dem Kontext hervorgehen sollte, bezüglich welchen Maßes sie definiert ist. Üblicherweise wird das aber das ein- (oder mehr-) dimensionale Lebesguemaß sein.

Für die auftretenden Verteilungen nutzen wir abkürzend folgende Notationen:

- Für $n \in \mathbb{N}$ heißt X uniform verteilt auf $\{1, \dots, n\}$, kurz $X \sim \text{unif}\{1, \dots, n\}$, falls für alle $k \in \{1, \dots, n\}$ gilt: $P(X = k) = \frac{1}{n}$.
- Sei $p \in [0, 1]$. Dann heißt X Bernoulli verteilt zum Parameter p , kurz $X \sim \text{ber}(p)$, falls $P(X = 1) = p$ und $P(X = 0) = 1 - p$.
- X heißt poisson verteilt zum Parameter $\lambda > 0$, kurz $X \sim \text{Poi}(\lambda)$, falls für alle $k \in \mathbb{N}$ gilt: $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$.
- Sei $0 < \lambda < 1$. Dann heißt X geometrisch verteilt zum Parameter λ , kurz $X \sim \text{geo}(\lambda)$, falls für alle $k \in \mathbb{N}$ gilt: $P(X = k) = \lambda(1 - \lambda)^{k-1}$.
- Sei $0 < \lambda < 1$ und $n \in \mathbb{N}$. Dann heißt X negativ binomial verteilt zu den Parametern n und λ , kurz $X \sim \text{negbin}(n, \lambda)$, falls für alle natürlichen Zahlen k mit $k \geq n$ gilt: $P(X = k) = \binom{k-1}{n-1} \lambda^n (1 - \lambda)^{k-n}$.

- Seien $a, b \in \mathbb{R}, a < b$. Dann heißt eine Zufallsvariable X uniform auf $[a, b]$ verteilt, kurz $X \sim \text{unif}[a, b]$, falls $f(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x)$ ihre Dichte ist.
- Sei $\lambda > 0$. Dann heißt X exponential verteilt zum Parameter λ , kurz $X \sim \text{exp}(\lambda)$, falls $f_\lambda(x) = \mathbb{1}_{[0,\infty)}(x) \lambda e^{-\lambda x}$ ihre Dichte ist.
- Seien $\alpha > 0$ und $\nu > 0$. Weiter bezeichne $\Gamma : \mathbb{R}^+ \rightarrow \mathbb{R}$ die Gammafunktion. Dann heißt X gamma verteilt zu den Parameter α und ν , kurz $X \sim \Gamma(\alpha, \nu)$, falls $f_{\alpha,\nu} = \frac{1}{\Gamma(\nu)} \alpha^\nu x^{\nu-1} e^{-\alpha x} \mathbb{1}_{[0,1]}(x)$ ihre Dichte ist.
- Seien $\nu > 0, \mu > 0$ und $B(\nu, \mu) := \int_0^1 (1-t)^{\mu-1} t^{\nu-1} d\lambda(t)$. Dann heißt X beta verteilt zu den Parametern ν und μ , kurz $X \sim \text{beta}(\nu, \mu)$, falls $f_{(\nu,\mu)}(x) = \frac{1}{B(\nu,\mu)} (1-x)^{\mu-1} x^{\nu-1} \mathbb{1}_{[0,1]}(x)$ ihre Dichte ist.
- Seien $\mu \in \mathbb{R}$ und $\sigma \in \mathbb{R}^+$. Dann heißt X normalverteilt zu den Parametern μ und σ^2 , kurz $X \sim N(\mu, \sigma^2)$, falls $f_{(\mu,\sigma^2)}(x) = (2\pi\sigma^2)^{-\frac{1}{2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ ihre Dichte ist.

Für asymptotische Betrachtungen benutzen wir die bekannten Landausymbole "o" und "O".

Schließlich setzen wir $\tilde{\mathbb{N}} := \mathbb{N} \setminus \{1\}$ und $[n] := \{1, \dots, n\}$ für alle $n \in \mathbb{N}$.

2.2 Die Exponentialverteilung und Yule-Prozesse

Da die Konstruktion des Hauptbeweises auf Yule-Prozessen basiert, wollen wir vorab einige elementare Eigenschaften dieser Prozessen erarbeiten. Sie verlangen - wie in der Definition klar werden wird - die Kenntnis im Umgang mit der Exponentialverteilung. Auch hierauf wollen wir noch kurz eingehen. Häufig werden wir Wartezeiten bis zum Eintritt gewisser Ereignisse durch exponential verteilte Zufallsvariablen modellieren und uns deren Eigenschaft der Gedächtnislosigkeit zu Nutze machen. Wir erinnern:

Sei X exponential verteilt, dann gilt $\forall x \geq 0, t \geq 0$

$$P(X > x + t | X > t) = P(X > x). \quad (1)$$

Lemma 2.2.1 (Minimum von exponential verteilten Zufallsvariablen). *Sei I abzählbar, $I \neq \emptyset$ und $(T_k)_{k \in I}$ eine Folge unabhängiger ZVn, mit $T_k \sim \text{exp}(q_k) \forall k \in I$.*

Sei weiter $q := \sum_{k \in I} q_k < \infty$, $T := \inf_{k \in I} T_k$ und $K := \arg \left(\inf_{k \in I} T_k \right)$.

Dann ist K fast sicher eindeutig definiert und das Infimum wird von T_K angenommen, wobei gilt

(i) T und K sind unabhängig.

(ii) $P(K = k) = \frac{q_k}{q}, \quad \forall k \in I.$

(iii) $T \sim \exp(q).$

Beweis: Nach dem Satz von der totalen Wahrscheinlichkeit und aus der Tatsache, dass die Verteilung von T_k per definitionem ein Maß mit λ -Dichte $f(x) = \mathbb{1}_{[0, \infty)}(x)q_k e^{-q_k x}$ ist, folgt für alle $t \geq 0$ und $k \in I$

$$\begin{aligned}
 P(K = k, T \geq t) &= P(T_k \geq t, T_j > T_k \forall j \neq k) \\
 &= \int_t^\infty P(T_k \geq t, T_j > T_k \forall j \neq k | T_k = s) dP_{T_k}(s) \\
 &= \int_t^\infty P(T_j > s \forall j \neq k) dP_{T_k}(s) \\
 &= \int_t^\infty P(T_j > s \forall j \neq k) q_k e^{-q_k s} d\lambda(s) \\
 &= \int_t^\infty q_k e^{-q_k s} \prod_{j \neq k} e^{-q_j s} d\lambda(s) \\
 &= \int_t^\infty q_k e^{-qs} d\lambda(s) \\
 &= \frac{q_k}{q} e^{-qt}.
 \end{aligned}$$

Betrachten wir das sichere Ereignis $\{T \geq 0\}$ so folgt direkt $P(K = k) = \frac{q_k}{q}$. Weiter erhalten wir die geforderte Verteilung von T durch

$$P(T \geq t) = \sum_{k \in I} P(K = k, T \geq t) = \sum_{k \in I} \frac{q_k}{q} e^{-qt} = e^{-qt}.$$

Damit folgt dann auch

$$P(K = k, T \geq t) = P(K = k)P(T \geq t)$$

zunächst für $k \in I$, und $t \geq 0$. Da $\{[t, \infty) | t \geq 0\}$ und trivialerweise auch die einelementigen Mengen aus I \cap -stabile Erzeuger der entsprechenden borelschen \mathfrak{G} -Algebren $\mathcal{B}([0, \infty))$ bzw. $\mathcal{B}(I)$ sind, folgt die Gleichheit auch für eben genau diese borelschen Mengen, was die Unabhängigkeit von K und T bedeutet. \square

Definition 2.2.2 (Yule-Prozess). Ein **Yule-Prozess** $Y = (Y_t)_{t \geq 0}$ ist ein reiner Geburtsprozess, für den folgende zwei Eigenschaften gelten:

- (i) Die Population startet fast sicher mit einem einzelnen Individuum zur Zeit $t = 0$.
- (ii) Jedes Individuum erzeugt Nachkommen gemäß eines homogenen Poissonprozesses mit Intensität $\lambda > 0$, unabhängig von den anderen Individuen.

Bemerkungen: Einen Yule-Prozess Y zum Parameter λ können wir als $D_{\mathbb{R}^+}[0, \infty)$ -wertige ZV verstehen und schreiben dafür kurz $Y \sim \text{Yule}(\lambda)$. Für $\lambda = 1$ sprechen wir auch von einem standard Yule-Prozess. Als Sprungprozess ist Y insbesondere ein Markovprozess.

Lemma 2.2.3. Sei $Y \sim \text{Yule}(\lambda)$. Dann gilt

- (i) Die Zeit S_n des n -ten Sprungs ist verteilungsgleich zu $T_1 + \dots + T_n$ mit T_1, \dots, T_n unabhängig und $T_i \sim \exp(i\lambda) \forall i \in \{1, \dots, n\}$.
- (ii) S_n ist verteilungsgleich zu $\max_{1 \leq i \leq n} V_i$, mit V_1, \dots, V_n unabhängig und $V_i \sim \exp(\lambda) \forall i \in \{1, \dots, n\}$.

Beweis: Zu (i): Wir setzen $S_0 := 0$ und $T_i := S_i - S_{i-1}$ für $1 \leq i \leq n$. Damit ist $S_n = \sum_{i=1}^n T_i$. Nach Konstruktion sind von Zeit S_{i-1} bis S_i genau i unabhängige $\exp(\lambda)$ -verteilte ZVn involviert, wobei sich deren Geburtszeitpunkte auf Grund der Gedächtnislosigkeit gemeinsam auf S_{i-1} festlegen lässt. T_i wird also gerade durch das Minimum von i $\exp(\lambda)$ -verteilten, unabhängigen ZVn beschrieben, womit nach (2.2.1) folgt: $T_i \sim \exp(i\lambda)$. Außerdem überträgt sich die Unabhängigkeit der einzelnen Individuen auf die T_1, \dots, T_n , was die Behauptung ergibt.

Zu (ii): Die Idee ist, den obigen Prozess rückläufig zu betrachten. Genauer: Wir starten zur Zeit $S_n^- := 0$ mit n Individuen, deren Lebenszeit jeweils durch eine $\exp(\lambda)$ verteilte ZV beschrieben wird. Dabei seien die n ZVn unabhängig. Für $0 \leq i \leq n-1$ sei S_i^- der erste Zeitpunkt an dem noch genau i Individuen leben. Ähnlich wie oben setzen wir $T_i^- := S_{i-1}^- - S_i^-$ für $1 \leq i \leq n$ und nutzen wieder die Gedächtnislosigkeit und Unabhängigkeit der Individuen, um $T_i^- \sim \exp(i\lambda)$ für $1 \leq i \leq n$ und deren Unabhängigkeit zu erhalten. Also gilt

$$S_0^- = (T_n^- + \dots + T_1^-) \sim (T_1 + \dots + T_n) = S_n. \quad (2)$$

Andererseits ist S_0^- nach Konstruktion der Todeszeitpunkt des letzten aller zur Zeit $S_n^- = 0$ gestarteten Individuen und beschreibt deswegen das Maximum von n unabhängig $\exp(\lambda)$ verteilten ZVn. \square

Korollar 2.2.4. Seien X_1, \dots, X_n unabhängig und $\exp(\lambda)$ verteilt. Dann gilt

$$\max_{1 \leq i \leq n} X_i \sim \sum_{i=1}^n \frac{X_i}{i}. \quad (3)$$

Beweis: Da $\frac{X_i}{i} \sim \exp(i\lambda)$ folgt die Behauptung.

Lemma 2.2.5. Sei $Y \sim \text{Yule}(\lambda)$. Dann gilt

(i) Zum festen Zeitpunkt $t > 0$ ist die Anzahl der im Yule-Prozess existierenden Individuen geometrisch verteilt zum Parameter $e^{-\lambda t}$, d.h.

$$Y_t \sim \text{geo}(e^{-\lambda t}). \quad (4)$$

(ii) Für $s, t \geq 0$ und $j \geq i$ gilt

$$P(Y_{s+t} = j \mid Y_s = i) = \binom{j-1}{i-1} e^{-\lambda t i} (1 - e^{-\lambda t})^{j-i}. \quad (5)$$

Beweis: Zu (i): Für alle $k \in \mathbb{N}$ bezeichne S_k wie oben den Zeitpunkt der k -ten Geburt im Yule-Prozess.

Für diesen gilt nach 2.2.3(ii): $P(S_k \leq t) = (1 - e^{-\lambda t})^k$. Trivialerweise gilt weiter: $\{S_{k-1} \leq t\} = \{S_{k-1} \leq t < S_k\} \cup \{S_k \leq t\}$. Also ist

$$\begin{aligned} P(Y_t = k) &= P(S_{k-1} \leq t < S_k) \\ &= P(S_{k-1} \leq t) - P(S_{k-1} \leq S_k \leq t) \\ &= P(S_{k-1} \leq t) - P(S_k \leq t) \\ &= (1 - e^{-\lambda t})^{k-1} - (1 - e^{-\lambda t})^k \\ &= e^{-\lambda t} (1 - e^{-\lambda t})^{k-1}. \end{aligned}$$

Zu (ii): Der Fall $t = 0$ ist trivial. Sei also im Folgenden $t > 0$. Wir nutzen die Kenntnis von $\{Y_s = i\}$, um den Prozess mit Hilfe der Gedächtnislosigkeit neu zu modellieren. Die Konstruktion von Y erlaubt, jedes zur Zeit s existierende Individuum als Initiator eines eigenen Yule-Prozesses - mit Startzeit s - zu interpretieren, der von der Entwicklung der anderen Prozesse unabhängig ist. Sei dazu Y^1, Y^2, \dots eine Folge unabhängiger Yule-Prozesse zum Parameter λ . Dann gilt

$$P(Y_{s+t} = j \mid Y_s = i) = P\left(\sum_{k=1}^i Y_t^k = j\right). \quad (6)$$

Nach dem ersten Teil gilt $Y_t^k \sim \text{geo}(e^{-\lambda t})$ für $1 \leq k \leq i$. Bekanntlich ist die Summe unabhängiger, geometrisch verteilter ZVn negativ binomial verteilt. Genauer gilt

$$\sum_{k=1}^i Y_t^k \sim \text{negbin}(i, e^{-\lambda t}), \quad (7)$$

d.h. $\forall j \geq i$ folgt direkt

$$P\left(\sum_{k=1}^i Y_t^k = j\right) = \binom{j-1}{i-1} (e^{-\lambda t})^i (1 - e^{-\lambda t})^{j-i}. \quad \square \quad (8)$$

In einem späteren Beweis werden wir hinsichtlich schwacher Konvergenz die folgende Stetigkeitsaussage benötigen.

Lemma 2.2.6. *Für $\lambda > 0$ sei $Y \sim \text{Yule}(\lambda)$, sowie $s \geq 0$. Weiter sei*

$$f : [1, \infty) \rightarrow [0, 1], \quad \text{definiert durch} \quad f(t) := P(Y_s = Y_{st}). \quad (9)$$

Dann ist $f \in \mathcal{C}_b$.

Beweis: Wir haben nur die Stetigkeit zu zeigen. Der Fall $s = 0$ ist trivial. Sei also $s > 0$ und $t \geq 1$. Wieder argumentieren wir mit der unabhängigen Entwicklung der Individuen und der Gedächtnislosigkeit der Exponentialverteilung. Mit der geometrischen Reihe folgt

$$\begin{aligned} f(t) = P(Y_s = Y_{st}) &= \sum_{m=1}^{\infty} P(Y_s = Y_{st} | Y_s = m) P(Y_s = m) \\ &= \sum_{m=1}^{\infty} P(Y_{st} = m | Y_s = m) P(Y_s = m) \\ \text{keine Geburt in } [s, st] : &\stackrel{(1)}{=} \sum_{m=1}^{\infty} (\lambda e^{-\lambda s(t-1)})^m e^{-\lambda s} (1 - e^{-\lambda s})^{m-1} \\ &= \sum_{m=1}^{\infty} \lambda e^{-\lambda st} (\lambda e^{-\lambda s(t-1)} (1 - e^{-\lambda s}))^{m-1} \\ &= \lambda e^{-\lambda ts} \sum_{m=0}^{\infty} (\lambda e^{-\lambda s(t-1)} (1 - e^{-\lambda s}))^m \\ &= \lambda e^{-\lambda ts} \frac{1}{1 - \lambda e^{-\lambda s(t-1)} (1 - e^{-\lambda s})} \end{aligned}$$

Wegen der Stetigkeit der Exponentialfunktion folgt die Behauptung. \square

Wir wollen nun noch eine Aussage über die Verteilung eines Yule-Prozesses zu zufälligen Zeipunkten machen. Dafür benötigen wir folgende Identität.

Hilfslemma 2.2.7. Für $m \in \tilde{\mathbb{N}}$ und $r \in \mathbb{R}$ gilt

$$(1 - e^{-r})^{m-1} - (1 - e^{-r})^m = \sum_{k=2}^m \binom{m-2}{k-2} (-1)^k (e^{-(k-1)r} - e^{-kr}). \quad (10)$$

Beweis: Via vollständiger Induktion über m .

Induktionsanfang: $m = 2$

$$(1 - e^{-r})^1 - (1 - e^{-r})^2 = e^{-r} - e^{-2r} = \sum_{k=2}^2 \binom{m-2}{k-2} (-1)^k (e^{-(k-1)r} - e^{-kr}).$$

Induktionsschritt: Sei die Behauptung für m bewiesen, d.h. es gilt (10). Wir haben zu zeigen, dass die Behauptung auch für $m + 1$ gilt.

Also z.Z.

$$(1 - e^{-r})^m - (1 - e^{-r})^{m+1} = \sum_{k=2}^{m+1} \binom{m-1}{k-2} (-1)^k (e^{-(k-1)r} - e^{-kr}). \quad (11)$$

Wir verwenden dazu die bekannte Identität $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$, für $n \in \mathbb{N}$ und $k \in \mathbb{Z}$ und nutzen aus, dass $\binom{n}{k} = 0$, für $k > n$ und $k < 0$ gilt. Es folgt

$$\begin{aligned} & \sum_{k=2}^{m+1} \binom{m-1}{k-2} (-1)^k (e^{-(k-1)r} - e^{-kr}) \\ &= \sum_{k=2}^{m+1} \binom{m-2}{k-3} (-1)^k (e^{-(k-1)r} - e^{-kr}) + \sum_{k=2}^{m+1} \binom{m-2}{k-2} (-1)^k (e^{-(k-1)r} - e^{-kr}) \\ &= \sum_{k=1}^m \binom{m-2}{k-2} (-1)^{k+1} (e^{-kr} - e^{-(k+1)r}) + \sum_{k=2}^{m+1} \binom{m-2}{k-2} (-1)^k (e^{-(k-1)r} - e^{-kr}) \\ &= -e^{-r} \sum_{k=2}^m \binom{m-2}{k-2} (-1)^k (e^{-(k-1)r} - e^{-kr}) + \sum_{k=2}^m \binom{m-2}{k-2} (-1)^k (e^{-(k-1)r} - e^{-kr}) \\ &= (1 - e^{-r}) \sum_{k=2}^m \binom{m-2}{k-2} (-1)^k (e^{-(k-1)r} - e^{-kr}) \\ &\stackrel{(10)}{=} (1 - e^{-r}) [(1 - e^{-r})^{m-1} - (1 - e^{-r})^m] = (1 - e^{-r})^m - (1 - e^{-r})^{m+1}. \quad \square \end{aligned}$$

Hiermit können wir nun zeigen:

Lemma 2.2.8. *Sei Y ein standard Yule-Prozess und U, E ZVn, mit $U \sim \text{unif}[0, 1]$ und $E \sim \exp(1)$, wobei Y, U und E unabhängig seien. Dann gilt für $m \in \mathbb{N}$:*

$$P(Y_{UE} = m) = \frac{1}{m} \sum_{k=1}^m \binom{m}{k} (-1)^{k+1} \log(k+1). \quad (12)$$

Zudem hat Y_{UE} unendliche Erwartung.

Beweis: Nach Lemma 2.2.5 gilt $Y_t \sim \text{geo}(e^{-t})$ für alle $t > 0$. Also ist $P(Y_t = m) = e^{-t}(1 - e^{-t})^{m-1} = (1 - e^{-t})^{m-1} - (1 - e^{-t})^m$ und damit nach dem obigen Hilfslemma 2.2.7

$$P(Y_t = m) = \begin{cases} e^{-t} & \text{für } m = 1, \\ \sum_{k=2}^m \binom{m-2}{k-2} (-1)^k (e^{-(k-1)t} - e^{-kt}) & \text{für } m \geq 2. \end{cases} \quad (13)$$

Wir halten m fest und betrachten die deterministische Funktion $f_m(t) := P(Y_t = m)$. Dann folgt aufgrund der Unabhängigkeit der involvierten ZVn

$$\begin{aligned} P(Y_{UE} = m) &= \int_{\mathbb{R}} P(Y_{UE} = m \mid UE = t) dP_{UE}(t) \\ &= \int_{\mathbb{R}} P(Y_t = m) dP_{UE}(t) \\ &= \int_{\mathbb{R}} f_m(t) dP_{UE}(t) \\ &\stackrel{(13)}{=} \begin{cases} \int_{\mathbb{R}} e^{-t} dP_{UE}(t) & \text{für } m = 1, \\ \int_{\mathbb{R}} \sum_{k=2}^m \binom{m-2}{k-2} (-1)^k (e^{-(k-1)t} - e^{-kt}) dP_{UE}(t) & \text{für } m \geq 2. \end{cases} \\ &= \begin{cases} \mathbb{E}[e^{-UE}] & \text{für } m = 1, \\ \sum_{k=2}^m \binom{m-2}{k-2} (-1)^k (\mathbb{E}[e^{-(k-1)UE}] - \mathbb{E}[e^{-kUE}]) & \text{für } m \geq 2. \end{cases} \end{aligned}$$

Für $\vartheta > -1$ berechnen wir mit Hilfe des Satzes von Fubini den Erwartungswert $\mathbb{E}[e^{-\vartheta UE}]$

$$\begin{aligned} \mathbb{E}[e^{-\vartheta UE}] &= \int_{\mathbb{R}^2} e^{-\vartheta ux} dP_{(U,E)}(u, x) \\ &= \int_{\mathbb{R}^2} e^{-\vartheta ux} \mathbb{1}_{[0,1]}(u) \mathbb{1}_{[0,\infty)}(x) e^{-x} d\lambda^2(u, x) \end{aligned}$$

$$\begin{aligned}
&= \int_0^1 \left(\int_0^\infty e^{-(\vartheta u+1)x} d\lambda(x) \right) d\lambda(u) \\
&= \int_0^1 \left(-\frac{1}{\vartheta u+1} e^{-(\vartheta u+1)x} \Big|_0^\infty \right) d\lambda(u) = \int_0^1 \frac{1}{\vartheta u+1} d\lambda(u) \\
&= \frac{1}{\vartheta} \log(\vartheta u+1) \Big|_0^\infty = \frac{1}{\vartheta} \log(\vartheta+1).
\end{aligned}$$

Für $\vartheta = 1$ folgt damit direkt

$$P(Y_{UE} = 1) = \log 2, \quad (14)$$

und für $m \geq 2$ schließen wir leicht mit Hilfe der üblichen Rechenregeln für den Binomialkoeffizienten

$$\begin{aligned}
P(Y_{UE} = m) &= \sum_{k=2}^m \binom{m-2}{k-2} (-1)^k \left(\frac{1}{k-1} \log k - \frac{1}{k} \log(k+1) \right) \\
&= \sum_{k=1}^{m-1} \binom{m-2}{k-1} (-1)^{k+1} \left(\frac{1}{k} \log(k+1) \right) + \sum_{k=2}^m \binom{m-2}{k-2} (-1)^{k+1} \left(\frac{1}{k} \log(k+1) \right) \\
&= \sum_{k=2}^{m-1} \binom{m-2}{k-1} (-1)^{k+1} \left(\frac{1}{k} \log(k+1) \right) + \sum_{k=2}^{m-1} \binom{m-2}{k-2} (-1)^{k+1} \left(\frac{1}{k} \log(k+1) \right) \\
&\quad + \log 2 + (-1)^{m+1} \frac{\log(m+1)}{m} \\
&= \sum_{k=2}^{m-1} \left[\binom{m-2}{k-1} + \binom{m-2}{k-2} \right] (-1)^{k+1} \frac{\log(k+1)}{k} + \log 2 + (-1)^{m+1} \frac{\log(m+1)}{m} \\
&= \sum_{k=1}^m \binom{m-1}{k-1} (-1)^{k+1} \frac{\log(k+1)}{k} = \sum_{k=1}^m \binom{m}{k} \frac{k}{m} (-1)^{k+1} \frac{\log(k+1)}{k} \\
&= \frac{1}{m} \sum_{k=1}^m \binom{m}{k} (-1)^{k+1} \log(k+1).
\end{aligned}$$

Also gilt die Behauptung für alle $m \in \mathbb{N}$.

Es ist noch der Erwartungswert zu berechnen. Nach Lemma 2.2.5 gilt $\mathbb{E}[Y_t] = e^t$, $\forall t \geq 0$. Mit der Glättungsregel für den bedingten Erwartungswert erhalten wir

$$\mathbb{E}[Y_{UE}] = \mathbb{E}[\mathbb{E}[Y_{UE}|UE]] = \mathbb{E}[e^{UE}]. \quad (15)$$

Sei weiter $(\vartheta_n)_{n \in \mathbb{N}}$ eine monoton fallende Folge mit $\lim_{n \rightarrow \infty} \vartheta_n = -1$, dann steigt $(e^{-\vartheta_n UE})_{n \in \mathbb{N}}$ punktweise gegen e^{UE} auf und mit dem Satz von der monotonen Konvergenz folgt

$$\mathbb{E}[e^{UE}] = \lim_{n \rightarrow \infty} \mathbb{E}[e^{-\vartheta_n UE}] = \lim_{n \rightarrow \infty} \frac{1}{\vartheta_n} \log(1 + \vartheta_n) = \infty. \quad \square$$

Wir betrachten nun noch den Prozess $(\frac{Y_t}{e^t})_{t \geq 0}$, wobei Y wieder ein standard Yule-Prozess sei.

Lemma 2.2.9. *Sei Y ein standard Yule-Prozess und $(\mathcal{F}_t)_{t \geq 0}$ die kanonische Filtrierung von $(Y_t)_{t \geq 0}$, d.h. $\forall t \in \mathbb{R}_0^+$ ist $\mathcal{F}_t = \mathfrak{G}(Y_s \mid s \leq t)$. Dann gilt*

(i) $(\frac{Y_t}{e^t})_{t \geq 0}$ ist ein Martingal bzgl. $(\mathcal{F}_t)_{t \geq 0}$.

(ii) $\lim_{t \rightarrow \infty} \frac{Y_t}{e^t} = E$ f.s., mit $E \sim \exp(1)$.

Beweis: Zu (i): Für $0 \leq s < t$ ist zu zeigen

$$\mathbb{E} \left[\frac{Y_t}{e^t} \mid \mathcal{F}_s \right] = \frac{Y_s}{e^s} \quad f.s. \quad (16)$$

Wir gehen wie im Beweis von (2.2.5)(ii) vor und wählen zunächst eine andere Darstellung für Y . Dazu verstehen wir jedes zum Zeitpunkt s existierende Individuum als Initiator eines eigenen standard Yule-Prozesses, der sich unabhängig von den anderen Prozessen entwickelt. D.h.

$$Y_t = \sum_{i=1}^{Y_s} Y_{t-s}^i \quad (17)$$

wobei dann Y^1, \dots, Y^{Y_s} unabhängige standard Yule-Prozesse sind. Außerdem liefert Y_s keine Information über die Entwicklung der zum Zeitpunkt s gestarteten Prozesse, d.h. auch $\mathcal{F}_s, Y^1, \dots, Y^{Y_s}$ sind unabhängig.

Unter Beachtung von $Y_{t-s}^1 \sim \text{geo}(e^{-(t-s)})$ folgt mit den üblichen Rechenregeln für bedingte Erwartungen

$$\begin{aligned} \mathbb{E}[Y_t \mid \mathcal{F}_s] &= \mathbb{E} \left[\sum_{i=1}^{Y_s} Y_{t-s}^i \mid \mathcal{F}_s \right] \\ &= \mathbb{E} \left[\sum_{j=1}^{\infty} \left(\sum_{i=1}^j Y_{t-s}^i \right) \mathbb{1}_{\{Y_s=j\}} \mid \mathcal{F}_s \right] \\ \text{monotone Konvergenz:} &= \sum_{j=1}^{\infty} \mathbb{E} \left[\left(\sum_{i=1}^j Y_{t-s}^i \right) \mathbb{1}_{\{Y_s=j\}} \mid \mathcal{F}_s \right] \\ \mathbb{1}_{\{Y_s=j\}} \text{ ist } \mathcal{F}_s\text{-m.b.:} &= \sum_{j=1}^{\infty} \mathbb{1}_{\{Y_s=j\}} \mathbb{E} \left[\left(\sum_{i=1}^j Y_{t-s}^i \right) \mid \mathcal{F}_s \right] \\ (Y^i)^{i \in \mathbb{N}} \text{ sind i.i.d.:} &= \sum_{j=1}^{\infty} \mathbb{1}_{\{Y_s=j\}} (j \mathbb{E}[(Y_{t-s}^1) \mid \mathcal{F}_s]) \\ \mathcal{F}_s, Y^1 \text{ sind unabhängig:} &= \sum_{j=1}^{\infty} \mathbb{1}_{\{Y_s=j\}} j \mathbb{E}[Y_{t-s}^1] \\ &= Y_s \mathbb{E}[Y_{t-s}^1] = Y_s e^t e^{-s} \quad f.s. \end{aligned}$$

Dies liefert die erste Behauptung.

Zu (ii): Da Y_t nicht negativ ist, folgt nach dem bekannten Konvergenzsatz für nicht-negative Martingale, dass $(\frac{Y_t}{e^t})_{t \geq 0}$ f.s. konvergiert.

Sei also weiter $E = \lim_{t \rightarrow \infty} \frac{Y_t}{e^t}$. Für $t > 0$ folgt mit der geometrischen Reihe

$$\sum_{i=0}^{\infty} (1 - e^{-t})^i = e^t, \quad (18)$$

und somit für $x \geq 0$

$$\begin{aligned} P\left(\frac{Y_t}{e^t} < x\right) &= 1 - P(Y_t \geq \lceil xe^t \rceil) \\ &= 1 - \sum_{i=\lceil xe^t \rceil}^{\infty} e^{-t}(1 - e^{-t})^{i-1} \\ &= 1 - e^{-t}(1 - e^{-t})^{\lceil xe^t \rceil - 1} \sum_{i=0}^{\infty} (1 - e^{-t})^i \\ &= 1 - (1 - e^{-t})^{xe^t + O(1)} \\ &= 1 - (1 + o(1)) \left(1 - \frac{1}{e^t}\right)^{xe^t} \xrightarrow{t \rightarrow \infty} 1 - e^{-x}. \quad \square \end{aligned}$$

3 BSC - RRT - CRP

3.1 Der Bolthausen-Sznitman Koaleszent

3.1.1 Der restringierte Koaleszent

Definition 3.1.1. Sei M eine nichtleere Menge und π ein System von Teilmengen von M . Dann heißt π eine **Partition** von M , wenn gilt

(i) $\emptyset \notin \pi$.

(ii) M ist disjunkte Vereinigung aller Elemente aus π .

Im folgenden sei $M \neq \emptyset$.

Mit $\mathcal{P}(M)$ bezeichnen wir die Menge der Partitionen von M .

Definition 3.1.2. Seien $\pi_g, \pi_f \in \mathcal{P}(M)$. Dann heißt π_g **größer als** π_f , falls jedes Element aus π_f Teilmenge eines Elementes aus π_g ist.

Sei $\pi_s(M)$ die feinste Partition von M . Ihre Elemente sind die einelementigen Teilmengen von M , die wir auch als Singletons bezeichnen. Beispielsweise ist $\pi_s(\mathbb{N}) = \{\{1\}, \{2\}, \dots\}$.

Wir betrachten zunächst einen $\mathcal{P}([n])$ -wertigen Sprungprozess $\Pi^{[n]}$, mit der Eigenschaft

$$\Pi_t^{[n]} \text{ ist größer als } \Pi_{t+s}^{[n]}, \text{ für alle } s, t \geq 0.$$

Die Wartezeiten zwischen den Übergängen sind bekanntlich exponential verteilt.

Wir betrachten die Teilmengen der Partition, die den Zustand des Prozesses zur festen Zeit t beschreibt. Eine solche Teilmenge bezeichnen wir auch als **Block** und verstehen die Elemente eines Blocks als angefügte **Marke**. Anschaulich bedeutet der Übergang in eine gröbere Partition eine Abnahme der Blockanzahl durch "Verschmelzen" gewisser Blöcke, d.h. Vereinigen ihrer Marken in einem neuen Block.

Mit diesen Begriffen können wir nun den restringierten Bolthausen-Sznitman Koaleszenten definieren. Dies ist ein Spezialfall des Prozesses $\Pi^{[n]}$, für den dann noch der Startzustand und die Koaleszenzraten festzulegen sind.

Definition 3.1.3. Der restringierte Bolthausen-Sznitman Koaleszent $BSC^{[n]} := (BSC_t^{[n]})_{t \geq 0}$ ist ein $\mathcal{P}([n])$ -wertiger rechtsstetiger Sprungprozess mit folgenden Eigenschaften:

(i) Der Prozess startet in der feinsten Partition von $[n]$, d.h.

$$BSC_0^{[n]} = \{\{1\}, \{2\}, \dots, \{n\}\} \quad f.s. \quad (19)$$

(ii) In der Situation von b Blöcken beträgt die Koaleszenzrate von k dieser Blöcke

$$\lambda_{b,k} = \frac{(k-2)!(b-k)!}{(b-1)!}, \quad 2 \leq k \leq b \leq n. \quad (20)$$

Eigenschaften des $BSC^{[n]}$:

(i): Für $2 \leq k \leq b$ gilt

$$\lambda_{b,k} = \int_{\mathbb{R}} x^{k-2}(1-x)^{b-k} dP_U(x), \quad (21)$$

mit $U \sim \text{unif}[0, 1]$, d.h. $\lambda_{b,k} = \int_0^1 x^{k-2}(1-x)^{b-k} dx$. Dies folgt direkt durch $(b-k)$ -maliges partielles Integrieren. Allgemeiner ist der Λ -Koaleszent für ein beliebiges endliches Maß Λ auf $[0, 1]$ definiert, dessen Koaleszenzraten sich dann durch (21) ergeben, wenn man die Verteilung von U durch die von Λ ersetzt.

(ii): Offenbar wird die Dynamik nicht von der Blockgröße - d.h. der dem Block angefügten Marke - beeinflusst.

(iii): In der Situation von b Blöcken fragen wir nach der Anzahl der Zustände W_b , in die der Prozess übergehen kann. Sei $2 \leq k \leq b$. Da es genau $\binom{b}{k}$ Möglichkeiten gibt, k Blöcke zu verschmelzen, erhalten wir einfach

$$W_b = \sum_{k=2}^b \binom{b}{k} = \sum_{k=0}^b \binom{b}{k} - \binom{b}{1} - \binom{b}{0} = 2^b - (b+1). \quad (22)$$

(iv): Die Wartezeit bis zur nächsten Verschmelzung ist dann das Minimum aller beteidigten, unabhängig exponential verteilten Wartezeiten. Nach 2.2.1 ist dieses $\exp(R_b)$ verteilt, wobei R_b die Summe der Raten aller Wartezeiten bezeichne. Eine einfache Rechnung ergibt

$$\begin{aligned} R_b &= \sum_{k=2}^b \binom{b}{k} \frac{(k-2)!(b-k)!}{(b-1)!} = \sum_{k=2}^b \frac{b!(k-2)!(b-k)!}{k!(b-k)!(b-1)!} = b \sum_{k=2}^b \frac{1}{k(k-1)} \\ &= b \sum_{k=1}^{b-1} \frac{1}{k(k+1)} = b \sum_{k=1}^{b-1} \left(\frac{k}{k+1} - \frac{k-1}{k} \right) = b-1. \end{aligned} \quad (23)$$

Dies liefert uns noch folgende Charakterisierung des $BSC^{[n]}$:

Korollar 3.1.4. *Der $BSC^{[n]}$ ist ein $\mathcal{P}([n])$ -wertiger Markovprozess mit den Eigenschaften*

(i)

$$BSC_0^{[n]} = \{\{1\}, \{2\}, \dots, \{n\}\} \quad f.s.$$

(ii) *In der Situation von b Blöcken ist die Wartezeit auf die nächste Verschmelzung $\exp(b-1)$ verteilt, für $2 \leq b \leq n$. (Falls nur noch ein Block vorhanden ist, verweilt der Prozess auf ewig in diesem Zustand).*

(iii) *Die Übergangswahrscheinlichkeiten, die das Verschmelzen von beliebigen k der vorhandenen b Blöcke beschreiben, sind gegeben durch*

$$P_{b,b-k+1} = \frac{\lambda_{b,k}}{b-1}, \quad \text{für } 2 \leq k \leq b. \quad (24)$$

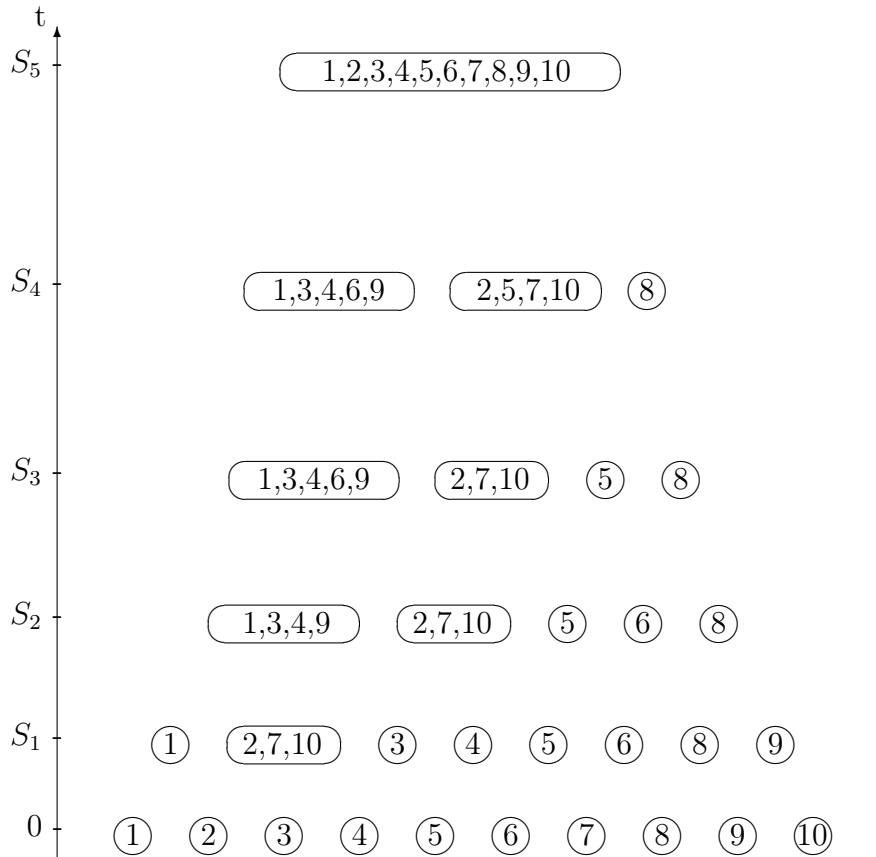


Abbildung 1: Eine Realisierung des $BSC^{[10]}$. S_1 bis S_5 bezeichnen die Zeitpunkte zu denen Verschmelzungen stattfinden. S_5 ist der Zeitpunkt der letzten Kollision.

3.1.2 Der unrestringierte Koaleszent

Der Vollständigkeit halber definieren wir noch den $\mathcal{P}(\mathbb{N})$ -wertigen Bolthausen-Sznitman Koaleszenten, kurz *BSC*.

Für alle $n \in \mathbb{N}$ betrachten wir die Restriktionsabbildungen

$$R_n : \mathcal{P}(\mathbb{N}) \rightarrow \mathcal{P}([n]), \quad \text{gegeben durch} \quad R_n(\pi) = \pi|_{[n]} \quad (25)$$

und damit die Menge von Folgen von Restriktionsabbildungen für jeweils festes π

$$F_{\mathcal{P}(\mathbb{N})} := \{(\pi_1, \pi_2, \dots) \mid (\pi_1, \pi_2, \dots) = (\pi|_{[1]}, \pi|_{[2]}, \dots), \pi \in \mathcal{P}(\mathbb{N})\}. \quad (26)$$

Dies liefert eine Bijektion

$$\phi : \mathcal{P}(\mathbb{N}) \rightarrow F_{\mathcal{P}(\mathbb{N})} \quad \text{via} \quad \pi \xrightarrow{\phi} (\pi|_{[1]}, \pi|_{[2]}, \dots). \quad (27)$$

Damit ist eine Partition von \mathbb{N} durch die Kenntnis aller Restriktionsabbildungen eindeutig bestimmt.

Definition 3.1.5. *Der Bolthausen-Sznitman Koaleszent $BSC := (BSC_t)_{t \geq 0}$ ist ein $\mathcal{P}(\mathbb{N})$ -wertiger C\`adl\`ag-Prozess, so dass für alle $n \in \mathbb{N}$ die Restriktion $BSC|_{[n]} := (BSC_t|_{[n]})_{t \geq 0}$ von BSC auf $[n]$ ein restringierter Bolthausen-Sznitman Koaleszent (gemäß Definition 3.1.3) ist.*

Wir haben den Raum $\mathcal{P}(\mathbb{N})$ noch mit einer geeigneten Metrik zu versehen. Für $\pi_1, \pi_2 \in \mathcal{P}(\mathbb{N})$ definieren wir

$$d(\pi_1, \pi_2) := \begin{cases} 0 & \text{falls } \pi_1 = \pi_2, \\ \frac{1}{n+1} & \text{falls } \pi_1|_{[n]} = \pi_2|_{[n]}, \pi_1|_{[n+1]} \neq \pi_2|_{[n+1]}, \\ 1 & \text{sonst.} \end{cases} \quad (28)$$

für alle $n \in \mathbb{N}$. D.h. der Abstand zweier Elemente ist um so kleiner, je mehr ihrer Restriktionsabbildungen übereinstimmen. Offenbar besitzt d die Eigenschaften einer Metrik. Die Dreiecksungleichung folgt dabei direkt wegen

$$d(\pi_1, \pi_2) \leq \max(d(\pi_1, \pi_3), d(\pi_2, \pi_3)), \quad \forall \pi_1, \pi_2, \pi_3 \in \mathcal{P}(\mathbb{N}).$$

Wir erklären warum diese Metrik in natürlicher Weise sinnvoll ist. Nach (27) gilt

$$\mathcal{P}(\mathbb{N}) \xrightarrow{\phi} F_{\mathcal{P}(\mathbb{N})} \subset \mathcal{P}([1]) \times \mathcal{P}([2]) \times \dots \quad (29)$$

Im Folgenden wollen wir eine Metrik für den Raum $\mathcal{P}([1]) \times \mathcal{P}([2]) \times \dots$ angeben, die dann gemäß obiger Bijektion ϕ auf den Raum $\mathcal{P}(\mathbb{N})$ vererbt wird.

Für alle $i \in \mathbb{N}$ sei $\mathcal{P}([i])$ mit der diskreten Topologie versehen. Dann ist eine die offenen Mengen (Potenzmenge) erzeugende Metrik durch

$$d_{dis}^i(\pi_1^i, \pi_2^i) := \begin{cases} 0 & \text{falls } \pi_1^i = \pi_2^i, \\ 1 & \text{sonst.} \end{cases} \quad (30)$$

gegeben, $(\pi_1^i, \pi_2^i \in \mathcal{P}([i])$.

Für $m \in \mathbb{N}$ betrachten wir dann zunächst den endlichen Produktraum $\mathcal{P}([1]) \times \cdots \times \mathcal{P}([m])$, versehen mit der Produkttopologie. Auch hier ist das System der offenen Mengen die Potenzmenge. Eine entsprechende Metrik ist dann gegeben durch

$$d_{dis(m)}((\pi_1^1, \dots, \pi_1^m), (\pi_2^1, \dots, \pi_2^m)) := \begin{cases} 0 & \text{falls } \pi_1^k = \pi_2^k \forall k \in [m], \\ 1 & \text{sonst.} \end{cases} \quad (31)$$

Dabei sind $\pi_1^i, \pi_2^i \in \mathcal{P}([i])$ für alle $i \in [m]$. Wir bemerken noch, dass (31) die gleiche Topologie beschreibt wie

$$\tilde{d}_{dis(m)}((\pi_1^1, \dots, \pi_1^m), (\pi_2^1, \dots, \pi_2^m)) := \begin{cases} 0 & \text{falls } \pi_1^k = \pi_2^k \forall k \in [m], \\ \frac{1}{n+1} & \text{falls } \pi_1^k = \pi_2^k \forall k \in [n], \pi_1^{n+1} \neq \pi_2^{n+1}, \\ 1 & \text{sonst.} \end{cases} \quad (32)$$

für $1 \leq n < m$.

Schließlich betrachten wir den unendlichen Produktraum $\mathcal{P}_\infty := \mathcal{P}([1]) \times \mathcal{P}([2]) \times \cdots$, versehen mit der Produkttopologie. Die offenen Mengen sind gerade jene, die in nur endlich vielen Komponenten eine beliebige Teilmenge der Potenzmenge des entsprechenden Raumes besitzen und deren restliche Komponenten durch den vollen Raum beschrieben werden, d.h.

$$\tau_\infty = \{A^1 \times A^2 \times \cdots \times A^i \times \mathcal{P}([i+1]) \times \mathcal{P}([i+2]) \times \cdots \mid i \in \mathbb{N}, A^j \subset \mathcal{P}([j]) \forall j \in [i]\}.$$

Eine die offenen Mengen charakterisierende Metrik ist dann gegeben durch

$$d_\infty((\pi_1^1, \pi_1^2, \dots), (\pi_2^1, \pi_2^2, \dots)) := \begin{cases} 0 & \text{falls } \pi_1^k = \pi_2^k \forall k \in \mathbb{N}, \\ \frac{1}{n+1} & \text{falls } \pi_1^k = \pi_2^k \forall k \in [n], \pi_1^{n+1} \neq \pi_2^{n+1}, \\ 1 & \text{sonst.} \end{cases} \quad (33)$$

mit $n \in \mathbb{N}$ und $\pi_1^i, \pi_2^i \in \mathcal{P}([i])$, für alle $i \in \mathbb{N}$.

Der Vergleich mit (32) zeigt, dass die Metrik d_∞ in gewisser Weise die Fortführung der Metrik des endlichen Produktraumes ist.

Vererben wir nun wie angesprochen die Metrik d_∞ anhand der Bijektion ϕ auf den Raum $\mathcal{P}(\mathbb{N})$, so erhalten wir direkt die Metrik d aus (28).

Schließlich wollen wir noch die Separabilität von $(\mathcal{P}(\mathbb{N}), d)$ folgern.

Lemma 3.1.6. *$(\mathcal{P}(\mathbb{N}), d)$ ist ein kompakter, separabler Raum.*

Beweis: Es ist zu zeigen, dass $\mathcal{P}(\mathbb{N})$ ein kompakter Raum ist. Die Separabilität folgt dann bekanntlich aus der Tatsache, dass $\mathcal{P}(\mathbb{N})$ ein metrischer Raum ist.

Wir gehen ähnlich wie oben vor und zeigen zunächst die Kompaktheit von $F_{\mathcal{P}(\mathbb{N})}$. Offenbar ist $\mathcal{P}([i]) = (\mathcal{P}([i]), d_{dis}^i)$ kompakt, für alle $i \in \mathbb{N}$. Nach dem Satz von Tychonoff ist dann aber auch der Produktraum $\mathcal{P}_\infty = (\mathcal{P}_\infty, d_\infty)$ kompakt.

Wir überlegen uns nun, warum der Teilraum $F_{\mathcal{P}(\mathbb{N})} = (F_{\mathcal{P}(\mathbb{N})}, d_\infty)$ abgeschlossen in \mathcal{P}_∞ ist. Dafür zeigen wir, dass $F_{\mathcal{P}(\mathbb{N})}$ alle seine Häufungspunkte enthält. Sei $(\pi_n)_{n \in \mathbb{N}}$ eine Folge in $F_{\mathcal{P}(\mathbb{N})}$.

Angenommen sie häufte sich in einem Punkt $\xi = (\xi^1, \xi^2, \dots) \in \mathcal{P}_\infty$ wobei $\xi \notin F_{\mathcal{P}(\mathbb{N})}$. Offenbar gilt

$$\xi \notin F_{\mathcal{P}(\mathbb{N})} \Leftrightarrow \exists i \in \mathbb{N}, \text{ mit } \xi^{i+1}|_{[i]} \neq \xi^i.$$

Andererseits gilt für jedes Folgenglied π_n per definitionem $\pi_n^{i+1}|_{[i]} = \pi_n^i$, $\forall i \in \mathbb{N}$, weil $\pi_n \in F_{\mathcal{P}(\mathbb{N})}$. Damit folgt aber

$$d_\infty(\pi_n, \xi) \geq \frac{1}{i+1}, \quad \forall n \in \mathbb{N}, \quad (34)$$

was im Widerspruch dazu steht, dass ξ ein Häufungspunkt ist. Also ist $F_{\mathcal{P}(\mathbb{N})}$ abgeschlossen.

Als abgeschlossene Teilmenge des kompakten Raums \mathcal{P}_∞ ist aber $F_{\mathcal{P}(\mathbb{N})}$ selbst kompakt und damit separabel. Schließlich gehen die Metriken d und d_∞ gerade so auseinander hervor, dass $\phi = \phi(d, d_\infty)$ eine Isometrie bezüglich dieser Metriken darstellt, d.h. die Kompaktheit bleibt auch beim Übergang von $F_{\mathcal{P}(\mathbb{N})}$ durch ϕ nach $\mathcal{P}(\mathbb{N})$ erhalten, was die Separabilität von $\mathcal{P}(\mathbb{N})$ liefert. \square

3.1.3 Eine Konstruktion des Koaleszenten

Schließlich wollen wir noch eine explizite Konstruktion des Bolthausen-Sznitman Koaleszenten angeben.

Wir führen zunächst für die Blöcke einer Partition von \mathbb{N} oder $[n]$, mit $n \in \mathbb{N}$, eine Ordnungsrelation ein. Sei b die Blockanzahl der Partition, dann ist $b \in \mathbb{N} \cup \{\infty\}$. Seien die Blöcke mit $(\ell_1), (\ell_2), \dots$ bezeichnet. Dann nennen wir sie **gemäß der „ $\overset{kE}{<}$ “-Relation angeordnet**, falls für $1 \leq i < j \leq b$ gilt: Das kleinste Element von (ℓ_i) ist kleiner als das kleinste Element von (ℓ_j) , kurz

$$(\ell_i) \overset{kE}{<} (\ell_j), \quad \text{für } 1 \leq i < j \leq b. \quad (35)$$

„ $\overset{kE}{<}$ “ beschreibt also die von den jeweils kleinsten Elementen induzierte Totalordnung.

Konstruktion:

Für festes $x \in (0, 1]$ sei $(Y_i)_{i \in \mathbb{N}}$ eine Folge unabhängiger ZVn mit $Y_i \sim \text{ber}(x)$, $\forall i \in \mathbb{N}$. (D.h. $P(Y_i = 1) = x$ und $P(Y_i = 0) = 1 - x$).

Wir setzen $Y := (Y_1, Y_2, \dots)$. Wegen der Unabhängigkeit der Folge (Y_i) ist die Verteilung von Y durch das Produktmaß auf $\prod_{i=1}^{\infty} \{0, 1\} = \{0, 1\}^{\infty}$ gegeben, welches wir für festes x mit P^x bezeichnen.

Weiter sei \mathcal{B} die Produkt- \mathfrak{G} -Algebra auf $\{0, 1\}^{\infty}$ und wir setzen

$$L(B) := \int_0^1 P^x(B) x^{-2} dx, \quad \forall B \in \mathcal{B}. \quad (36)$$

Dies ist gerade so konstruiert, dass es später mit den Koaleszenzraten $\lambda_{b,k}$ zusammenpassen wird.

Wir konstruieren nun eine Folge von Koaleszenten $(\Pi^{[n]})^{n \in \mathbb{N}}$, so dass $\Pi^{[n]}$ die Restriktion von $\Pi^{[n+1]}$ auf $[n]$ ist, für alle $n \in \mathbb{N}$. Dabei wollen wir stets fast sicher in Singletons starten.

Wir betrachten die Punkte eines Poisson-Prozesses \mathcal{N} auf $(0, \infty) \times \{0, 1\}^{\infty}$ mit Intensitätsmaß $\mu = \lambda \otimes L$ (λ bezeichnet das Lebesguemaß), d.h.

$$\mu(A \times B) = \lambda(A)L(B) \quad (37)$$

für alle $A \in \mathfrak{B}((0, \infty))$ und $B \in \mathcal{B}$. Dies legt dann aber schon das Maß aller messbaren Mengen des Produktraumes fest.

Als Sprungzeitpunkte von $\Pi^{[n]}$ lassen wir nur die Zeitpunkte t zu, für die ein $Y \in \{0, 1\}^{\infty}$ existiert mit $(t, Y) \in \mathcal{N}$ und zusätzlich $\sum_{i=1}^n Y_i \geq 2$ ist.

(Falls $\sum_{i=1}^n Y_i \geq 2$, so existiert das Integral aus Gleichung (36), da der Faktor x^2 aus dem Nenner verschwindet. Damit bleibt die Intensität beschränkt, $\mu(A \times \{Y \mid \sum_{i=1}^n Y_i \geq 2\}) < \infty$, falls $\lambda(A) < \infty$, d.h. die Sprünge können sich nicht häufen).

Sei die (diskrete) Menge der zulässigen Sprungzeitpunkte mit \mathcal{T} bezeichnet. Angenommen zur Zeit t_0 haben sich b Blöcke $(\ell_1), \dots, (\ell_b)$ gebildet, mit $2 \leq b \leq n$, d.h. es gilt $\Pi_{t_0}^{[n]} = \{(\ell_1), \dots, (\ell_b)\} =: \mathcal{L}$. Die Blöcke seien gemäß der " \prec "-Relation angereiht und wir ordnen dem i -ten Block die ZV Y_i zu.

Die Idee ist nun zur nächsten potentiellen Sprungzeit $\tau \in \mathcal{T}$ mit $\tau \geq t_0$ alle Blöcke (ℓ_i) , $1 \leq i \leq b$, zu vereinigen, für die die entsprechende ZV Y_i den Wert 1 angenommen hat. Wir werden also zur Zeit τ genau dann eine Koaleszenz erleben, wenn $\sum_{i=1}^b Y_i \geq 2$ gilt.

Wir benötigen noch eine Interpretation des Poisson-Prozesses (vgl. Gleichung (37)). Für jede feste Menge $B \in \mathcal{B}$ betrachten wir die Restriktion $\mathcal{N}_B := \mathcal{N} \cap \{(0, \infty) \times B\}$. Dann beschreibt \mathcal{N}_B einen homogenen P.P. mit Intensität $L(B)$.

(Denn das Intensitätsmaß ist offenbar das $L(B)$ -fache des Lebesguemaßes). Außerdem sind dann für paarweise disjunkte Mengen $(B_i)_{i \in \mathbb{N}}$ (mit $B_i \in \mathcal{B}$) die P.P. $(\mathcal{N}_{B_i})_{i \in \mathbb{N}}$ unabhängig.

In der Situation $\Pi_{t_0}^{[n]} = \mathcal{L}$ fragen wir nach der Rate $L(\text{melt}(\mathcal{S}))$ desjenigen eingeschränkten Poisson-Prozesses, der das Verschmelzen der k -elementigen Menge $\mathcal{S} := \{\ell_{i_1}, \dots, \ell_{i_k}\} \subset \mathcal{L}$ verursacht, $2 \leq k \leq b$.

Dies bedeutet, dass $(\tau, Y) \in \mathcal{N}$ mit $Y_{i_1} = \dots = Y_{i_k} = 1$. Da die Y_i mit $i > b$ keine Bedeutung für die Verschmelzung der b Blöcke haben, wird dieser Folge die Wahrscheinlichkeit $P^x(\text{melt}\mathcal{S}) = x^k(1-x)^{b-k}$ zugeordnet. (Wir bemerken, dass $Y_j = 0 \forall j \in [b] \setminus \{i_1, \dots, i_k\}$ gilt). Nach Konstruktion ergibt sich (vgl. (36))

$$L(\text{melt}\mathcal{S}) = \int_0^1 x^{k-2}(1-x)^{b-k} dx = \lambda_{b,k}. \quad (38)$$

D.h. gegeben der Information des Zustandes \mathcal{L} zu einer gewissen Zeit t_0 ist die Wartezeit auf die Verschmelzung von \mathcal{S} exponential verteilt (Eigenschaft eines homogenen P.P.) mit den benötigten Raten $\lambda_{b,k}$.

Außerdem sind die eingeschränkten Prozesse \mathcal{N}_B , die die Wartezeit auf die nächste Verschmelzung beschreiben, unabhängig, denn die entsprechenden Mengen B sind nach Konstruktion paarweise disjunkt.

Damit gilt $\Pi^{[n]} \sim BSC^{[n]}$, für alle $n \in \mathbb{N}$. Außerdem folgt nach Konstruktion direkt, dass $\Pi^{[n]}$ die Restriktion von $\Pi^{[n+1]}$ auf $[n]$ ist. Wir haben uns nur klar zu machen, welche Punkte des P.P. \mathcal{N} Übergänge der Prozesse $\Pi^{[n]}$ bzw. $\Pi^{[n+1]}$ beschreiben.

Insbesondere gilt $(\pi^{[1]}, \pi^{[2]}, \dots) \in F_{\mathcal{P}(\mathbb{N})}$ und wir können diese Folge anhand der Abbildung ϕ^{-1} eindeutig mit einem $\pi \in \mathcal{P}(\mathbb{N})$ identifizieren (vgl. (27)). Dann ist π der gesuchte unrestringierte Bolthausen-Sznitman Koaleszent.

3.2 Zufällige rekursive Bäume

Wie schon in der Einleitung erwähnt wollen wir Aussagen über den restringierten Bolthausen-Sznitman Koaleszenten mit Hilfe zufälliger rekursiver Bäume gewinnen. Dieser Abschnitt dient der Modellierung des Koaleszenten durch eine solche Baumstruktur. Begriffe wie *Baum*, *Knoten*, *Kante etc.* setzen wir wieder als bekannt voraus. Wir betrachten Bäume mit n Knoten, $n \in \mathbb{N}$. Dabei sei jeder Knoten mit einer Marke (k) versehen, $k \in \{1, \dots, n\}$, so dass die Menge der Marken gerade die Partition der ersten n natürlichen Zahlen in Singletons ergibt. Später werden wir die Markenmenge verallgemeinern. Sollte bei der Betrachtung von Bäumen künftig keine Angabe zur Markenmenge gemacht werden, gehen wir von Bäumen in Singletons aus.

Definition 3.2.1. Für $n \in \mathbb{N}$ sei T^n ein Baum mit n Knoten. T^n heißt *rekursiv*, wenn folgende zwei Eigenschaften erfüllt sind:

- (i) Die Wurzel von T^n ist mit (1) markiert.
- (ii) Für jeden Knoten in T^n bilden die Marken der Knoten, die man auf dem Pfad von der Wurzel bis zu dem Knoten durchläuft, eine aufsteigende Folge.

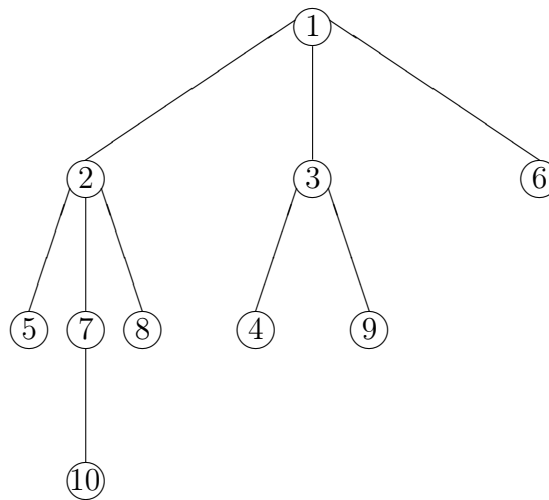


Abbildung 2: Ein rekursiver Baum bestehend aus 10 Knoten.

Ein solcher rekursiver Baum lässt sich explizit konstruieren: Zunächst ist die Wurzel festzulegen. Induktiv wird im k -ten Schritt der Knoten mit der Marke (k) an einen beliebigen bereits im Baum vorhandenen Knoten angefügt, $2 \leq k \leq n$.

Für $n \in \mathbb{N}$ existieren also genau $(n - 1)!$ rekursive Bäume.

Definition 3.2.2. Für $n \in \mathbb{N}$ heißt ein Baum T^n *zufälliger rekursiver Baum*, kurz RRT^n , falls er uniform aus der Menge der rekursiven Bäume gewählt ist.

Für unsere Konstruktion bedeutet dies wieder die (1) als Wurzel festzulegen und den Knoten, an dem im k -ten Schritt angefügt wird, uniform aus der Menge aller $k - 1$ bislang im Baum vorhandenen Knoten zu wählen, $2 \leq k \leq n$. So wird keiner der rekursiven Bäume bevorzugt.

Unser Ziel ist es den $BSC^{[n]}$ durch einen RRT^n zu modellieren: Für alle $1 \leq k \leq n$ ordnen wir dem Block mit der Marke (k) im Koaleszenten den mit (k) markierten Knoten des Baums zu. Ebenso werden wir die Koaleszenz im $BSC^{[n]}$ durch Verschmelzen der entsprechenden Knoten im Baum beschreiben. Dieses Vorhaben legt nun nahe, die Markenmenge eines zufälligen rekursiven Baumes zu verallgemeinern (vgl. (35)).

Definition 3.2.3. Sei $n \in \mathbb{N}$ und $\{(\ell_1), \dots, (\ell_b)\}$ eine Partition von $[n]$ mit $(\ell_i) \stackrel{kE}{<} (\ell_j)$, für $1 \leq i < j \leq b$. Ein Baum bestehend aus den Knoten $(\ell_1), \dots, (\ell_b)$ heißt ein zufälliger rekursiver Baum (RRT^b), falls er aus folgender Konstruktion hervorgeht:

- (i) Die Wurzel ist mit (ℓ_1) bezeichnet.
- (ii) (ℓ_k) wird uniform an einen der $k - 1$ im Baum enthaltenen Knoten $(\ell_1), \dots, (\ell_{k-1})$ angefügt, $2 \leq k \leq b$.

Die Grundidee für ein Analogon zur Entwicklung des $BSC^{[n]}$ - d.h. zur sukzessiven Verschmelzung der Blöcke - ist ein bekanntes Abholzverfahren eines zufälligen rekursiven Baumes, das auf Meir und Moon (vgl. [10]) zurückgeht: Wir betrachten einen RRT^n mit $n - 1$ Kanten. Wir wählen uniform eine Kante e , schneiden diese ab und lassen den Teilbaum, dessen Wurzel der Knoten unterhalb der geschnittenen Kante ist, wegfallen. Dadurch entsteht ein neuer Baum mit weniger Knoten. Das gleiche Verfahren führen wir solange durch bis die Wurzel schließlich isoliert ist.

Für unsere Belange ist es sinnvoll, dieses Verfahren etwas abzuändern: Dabei lassen wir den abgeschnittenen Teilbaum nicht wegfallen, sondern fügen die Marken aller seiner Knoten zu dem Knoten hinzu, der direkt oberhalb der geschnittenen Kante liegt. Dies kommt der Vorstellung der Entwicklung des Koaleszenten schon sehr nahe.

Die folgende Darstellung zeigt den Abholzprozess eines rekursiven Baumes der gemäß unseres entwickelten Modells mit dem in Abbildung 1 auf Seite 14 betrachteten Bolthausen-Sznitman Koaleszenten korrespondiert. S_1, \dots, S_5 bezeichnen dabei die Zeitpunkte, zu denen die angedeuteten Kanten geschnitten werden.

Hilfssatz 3.2.4. Sei $n \in \mathbb{N}$ und RRT^b ein zufälliger rekursiver Baum auf der Markenmenge $\mathcal{L} = \{(\ell_1), \dots, (\ell_b)\} \in \mathcal{P}([n])$. Sei weiter X eine ZV mit $X \sim \text{unif}\{(\ell_2), \dots, (\ell_b)\}$ und X, RRT^b unabhängig. Schneidet man die unmittelbar oberhalb von X liegende Kante und fügt alle Marken des Teilbaumes, dessen Wurzel X ist, zum direkt oberhalb von X liegenden Knoten hinzu, so ist der dadurch entstandene Baum ein zufälliger rekursiver Baum auf der neuen Markenmenge.

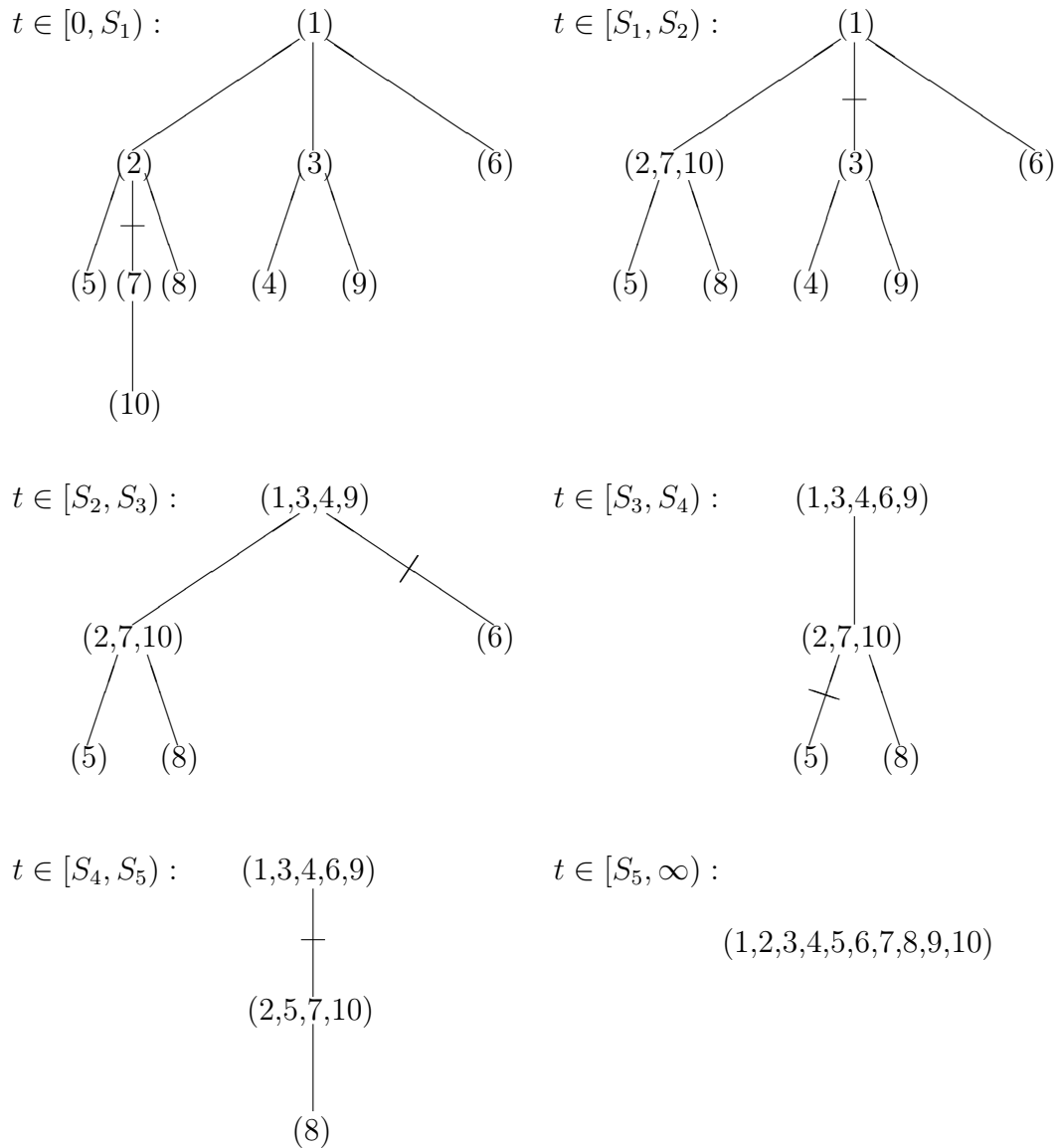


Abbildung 3: Eine mögliche Abholzung eines RRT^{10} . Das Prozedere korrespondiert mit dem $BSC^{[10]}$ von Abbildung 1.

Beweis:

(i): z.Z.: Der neue Baum ist rekursiv. Wir haben nur zu zeigen, dass für jeden Knoten K im neuen Baum, die Folge der Marken von der Wurzel bis K aufsteigend ist, im Sinne der oben eingeführten " $\stackrel{kE}{<}$ "-Relation. Sei X^o der Knoten der unmittelbar oberhalb von X liegt und (ℓ^o) dessen Marke,

$(\ell^o) \in \mathcal{L}$. Nach Konstruktion des Baumes sind alle Knoten des Teilbaumes, dessen Wurzel X ist, mit Marken versehen, die größer sind als (ℓ^o) . Sei k^o das kleinste Element von (ℓ^o) . Nach der Vereinigung der Marken des abgetrennten Baumes mit (ℓ^o) bleibt k^o also das kleinste Element dieser neuen Marke. Andererseits sind bis auf X^o alle Knoten im neuen Baum von der Verschmelzung unbeeinflusst. Damit bleibt für alle neuen Marken die Rekursivitätseigenschaft erhalten.

(ii): Es ist noch zu zeigen, dass der durch Abschneiden und Verschmelzen entstandene Baum $T^{(neu)}$ uniform verteilt ist auf der neu entstandenen Markenmenge $\mathcal{L}^{(neu)}$. Sei $X = (\ell_{i_2})$ und $\{(\ell_{i_2}), \dots, (\ell_{i_k})\}$ die Knoten des Teilbaums dessen Wurzel (ℓ_{i_2}) ist. Sei weiter (ℓ_{i_1}) der direkte Vorfahr von (ℓ_{i_2}) . Dann ist $\mathcal{S} := \{(\ell_{i_1}), \dots, (\ell_{i_k})\} \subset \mathcal{L}$ die Markenmenge, die bei diesem Vorgang vereinigt wird.

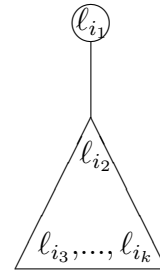


Abbildung 4: Teilstruktur eines zulässigen rekursiven Baumes.

Diese sei wieder gemäß der „ \leq^{kE} “-Relation geordnet. Wir bemerken zunächst, dass ein beliebiger rekursiver Baum auf \mathcal{L} aufgrund der Anordnung der Marken die Verschmelzung von \mathcal{S} im Allgemeinen gar nicht ermöglicht. Bäume für die dies doch möglich ist, bezeichnen wir als zulässig. Genauer haben wir also zu zeigen, dass, wenn wir durch den Verschmelzungsvorgang einen Baum mit Markenmenge $\mathcal{L}^{(neu)}$ erhalten, dieser nicht vor anderen Bäumen auf $\mathcal{L}^{(neu)}$ bevorzugt wird. Dies folgt aber leicht, denn notwendig und hinreichend für einen zulässigen Baum ist ein rekursiver Unterbaum bestehend aus Knoten mit den Marken $(\ell_{i_2}), \dots, (\ell_{i_k})$, der an den Knoten mit der Marke (ℓ_{i_1}) angefügt ist (vgl. Abbildung 4). Für den rekursiven Unterbaum gibt es damit $(k - 2)!$ mögliche Anordnungen und damit für jeden Baum auf $\mathcal{L}^{(neu)}$ also genau $(k - 2)!$ zulässige Bäume auf \mathcal{L} . Weil wir von einem zufälligen rekursiven Baum ausgegangen sind, ist aber auf \mathcal{L} kein Baum bevorzugt. \square

Anhand des folgenden Satzes lässt sich nun die Entwicklung des $BSC^{[n]}$ durch einen rekursiven Baum mit Hilfe des oben erwähnten Abholzverfahrens beschreiben.

Aus technischen Gründen betrachten wir im folgenden Satz nur Bolthausen-Sznitman Koaleszenten mit wenigsten zwei Blöcken. Man beachte aber, dass der Fall $n = 1$ trivialerweise richtig ist, auch wenn er für unsere asymptotischen Betrachtungen keine Bedeutung hat.

Konstruktion 3.2.5. Für $n \in \tilde{\mathbb{N}}$ sei zur Zeit $t = 0$ ein zufälliger rekursiver Baum in Singletons gegeben. Weiter $(E_k)_{2 \leq k \leq n}$ eine Folge standard exponential verteilter ZVn, so dass E_2, \dots, E_n, RRT^n unabhängig sind. Dabei gibt

E_k die Lebenszeit der Kante an, mit der der Knoten (k) angefügt wurde, d.h. E_k beschreibt den Zeitpunkt, zu dem der Teilbaum dessen Wurzel (k) ist, instantan in den Vater von (k) verschmilzt.

Satz 3.2.6. Die Markenmenge des Baumes aus Konstruktion 3.2.5 beschreibt die Blöcke einer Partition der ersten n natürlichen Zahlen, die sich gemäß der Dynamik des BSC^[n] entwickeln.

Beweis: Wir weisen die Eigenschaften aus Korollar 3.1.4 nach. Sei dafür zur Zeit $t \geq 0$ die Markenmenge des betrachteten Baumes durch $\Pi_t^{[n]}$ bezeichnet. Zunächst starten wir nach Konstruktion fast sicher in Singletons.

Angenommen es wurden mit der Zeit gewisse Kanten geschnitten, so dass zum Zeitpunkt t_0 noch b Knoten vorhanden sind ($2 \leq b \leq n$), deren Markenmenge durch $\mathcal{L} = \{(l_1), \dots, (l_b)\}$ gegeben ist (welche wiederum nach der obigen " \prec^{kE} "-Relation angeordnet ist), d.h. $\Pi_{t_0}^{[n]} = \mathcal{L}$.

Mit b Knoten sind noch $b - 1$ Kanten vorhanden. Also wird die Wartezeit auf die nächste Verschmelzung durch das Minimum von $b - 1$ unabhängigen $\exp(1)$ verteilte ZVn beschrieben und ist somit $\exp(b - 1)$ verteilt.

Sei τ der Zeitpunkt, an dem die nächste Verschmelzung stattfindet. Betrachten wir dann den Baum zur Zeit t_0 so ist die künftige Entwicklung nach Konstruktion vom vorherigen Verlauf unabhängig. Außerdem ist der zur Zeit τ entstandene Baum nach Hilfssatz 3.2.4 wieder zufällig rekursiv. Damit ist der Prozess markovsch.

Schließlich haben wir noch die Übergangswahrscheinlichkeiten zu berechnen. Gegeben der Information $\Pi_{t_0}^{[n]} = \mathcal{L}$ fragen wir nach der Wahrscheinlichkeit, dass im nächsten Schritt genau die k Knoten zusammenfallen, die durch die Markenmenge $\mathcal{S} = \{(l_{i_1}), \dots, (l_{i_k})\} \subset \mathcal{L}$ beschrieben werden. Die neu entstandene Markenmenge wollen wir wieder mit $\mathcal{L}^{(neu)}$ bezeichnen. Wir benötigen die Anzahl der rekursiven Bäume, die das Verschmelzen von \mathcal{S} überhaupt erst begünstigen. Diese Bäume hatten wir im vorigen Hilfssatz 3.2.4 als zulässig bezeichnet. Sie müssen also die Struktur aus Abbildung 4 aufweisen. Wir wählen erstens einen der $(k - 2)!$ möglichen rekursiven Bäume aus den Knoten der Marken $(l_{i_2}), \dots, (l_{i_k})$. Und zweitens einen der $(b - k)!$ rekursiven Bäume aus dem Knoten mit der Marke (l_{i_1}) und den übrigen $b - k$ Knoten. Dann fügen wir die beiden Bäume zusammen indem (l_{i_1}) und (l_{i_2}) durch eine Kante e verbunden werden. Schließlich haben wir noch eine Kante zu wählen, die das nächste Ereignis beschreibt. Dafür haben wir nur eine Möglichkeit, denn es muss e geschnitten werden, damit \mathcal{S} verschmilzt. Insgesamt erhalten wir also $(b - k)!(k - 2)!$ "günstige" Bäume.

Andererseits beträgt die Anzahl aller rekursiver Bäume mit einer markierten Kante, die das Eintreten des nächsten Ereignisses beschreibt, gerade

$(b-1)!(b-1)$.

Nach dem Hilfssatz 3.2.4 ist der zur Zeit t_0 betrachtete Baum zufällig rekursiv, d.h. er ist unter allen Bäumen auf \mathcal{L} nicht bevorzugt. Da auch die geschnittene Kante uniform gewählt wird, ergibt sich die gesuchte Wahrscheinlichkeit leicht als Quotient

$$P(\Pi_\tau^{[n]} = \mathcal{L}^{(neu)} \mid \Pi_{t_0}^{[n]} = \mathcal{L}) = \frac{(b-k)!(k-2)!}{(b-1)!(b-1)} = \frac{\lambda_{b,k}}{b-1}. \quad (39)$$

dies ist aber gerade die passende Übergangswahrscheinlichkeit, also gilt die Behauptung. \square

Bemerkung: Sei S_n die Anzahl der Schnitte, die nötig sind, um die Wurzel eines RRT^n zu isolieren. Dann beschreibt S_n nach Satz 3.2.6 auch die Anzahl der Verschmelzungen eines $BSC^{[n]}$ bis zum finalen Zustand. Meir und Moon (vgl. [10]) haben gezeigt, dass für $n \rightarrow \infty$ gilt

$$\mathbb{E}[S_n] = \frac{n}{\log n}(1 + o(1)) \quad \text{und} \quad \text{Var}[S_n] = o\left(\frac{n^2}{(\log n)^2}\right). \quad (40)$$

Ein genaueres Resultat über die asymptotische Entwicklung der Momente, sowie der zentrierten Momente von S_n lieferte später Panholzer (vgl. [14, Theorem 21 mit $\alpha = 1$]). Er bewies

Satz 3.2.7. *Für $k \in \mathbb{N}$ existieren sowohl die k -ten Momente als auch die k -ten zentrierten Momente von S_n und für $n \rightarrow \infty$ gilt*

$$\mathbb{E}[S_n^k] = \frac{n^k}{(\log n)^k} + \left(H_k + k + \sum_{\ell=1}^k \Psi(\ell)\right) \frac{n^k}{(\log n)^{k+1}} + O\left(\frac{n^k}{(\log n)^{k+2}}\right) \quad (41)$$

und

$$\mathbb{E}\left[|S_n - \mathbb{E}[S_n]|^k\right] = \left((-1)^k + \sum_{\ell=0}^{k-1} \binom{k-1}{\ell} (-1)^{k-\ell-1} \Psi(\ell+1)\right) \frac{n^k}{(\log n)^{k+1}} + O\left(\frac{n^k}{(\log n)^{k+2}}\right), \quad (42)$$

mit $H_n = \sum_{i=1}^n \frac{1}{i}$ und $\Psi(x) = \frac{d}{dx} \log \Gamma(x)$. Folglich gilt für $n \rightarrow \infty$

$$\frac{\log n}{n} S_n \xrightarrow{P} 1. \quad (43)$$

Zudem haben Möhle und Iksanov (vgl. [13]) folgenden Grenzwertsatz geliefert:

Satz 3.2.8. Für $n \rightarrow \infty$ gilt

$$\frac{(\log n)^2}{n} S_n - \log n - \log \log n \xrightarrow{d} S, \quad (44)$$

wobei S eine stabile Zufallsvariable bezeichne, deren charakteristische Funktion $E(e^{itX}) = \exp(it \log |t| - \frac{\pi}{2}|t|)$, $\forall t \in \mathbb{R}$, ist.

3.3 Eine Größenbetrachtung der Abkömmlinge

Wie angekündigt betrachten wir gemäß Satz 3.2.6 einen zufälligen rekursiven Baum auf $[n]$. Der $BSC^{[n]}$ ist in gewisser Weise symmetrisch. Hinsichtlich der letzten Verschmelzung kommt dagegen der Wurzel im Baum - also dem Knoten, dessen Marke die (1) enthält - eine besondere Rolle zu: Der letzte Schnitt betrifft eine der Kanten, die von der Wurzel abgehen.

Für Betrachtungen, die die Blockgrößen in der letzten Verschmelzung betreffen, benötigen wir nicht die volle Information über den entsprechenden zufälligen rekursiven Baum, sondern lediglich die Größen der Teilbäume, die von den Kindern der Wurzel erzeugt werden, und die Zeiten, zu denen ihre Verbindungskante mit der Wurzel geschnitten wird. Letztere werden nach Konstruktion durch unabhängige standard exponential verteilte ZVn beschrieben (vgl. Abbildung 5).

Aus technischen Gründen sei im folgenden wieder $n \in \tilde{\mathbb{N}}$. Wir führen noch einige Notationen ein.

Definition 3.3.1. Sei RRT^n ein zufälliger rekursiver Baum in Singletons und K_n die Anzahl der Kinder der mit (1) markierten Wurzel. Für $i \in \{1, \dots, K_n\}$ bezeichnet der **i -te Abkömmling**, kurz $A_{n,i}$, den Teilbaum, dessen Wurzel das i -te Kind der (1) ist.

Sei $A_{n,i} := \emptyset$, für $i > K_n$. Für alle $i \in \mathbb{N}$ bezeichnet $I_{n,i}$ die **Größe des Abkömmlings** $A_{n,i}$, d.h. die Anzahl seiner enthaltenen Knoten bzw. Marken.

Für alle $i \in \mathbb{N}$ ordnen wir dem i -ten Abkömmling den **Rang** i zu.

Bezüglich der Abschneideprozedur bezeichnet E_i nun die standard exponential verteilte Lebenszeit des i -ten Abkömmlings, $\forall i \in \mathbb{N}$.

Gesucht sind die Größen $I_{n,i}$ der Abkömmlinge eines zufälligen rekursiven Baumes RRT^n . Es wird sich herausstellen, dass sie erwartungsgemäß nicht gleich sein werden. Für zwei Abkömmlinge wird derjenige kleineren Ranges typischerweise größer sein. (vgl. Abbildung 5)

Die Konstruktion der zufälligen rekursiven Bäume erlaubt es uns, eine Folge $(RRT^n)_{n \in \mathbb{N}}$ als Markovkette auf dem Raum der rekursiven Bäume zu verstehen. Beim Übergang von einem rekursiven Baum zum nächsten wird dazu

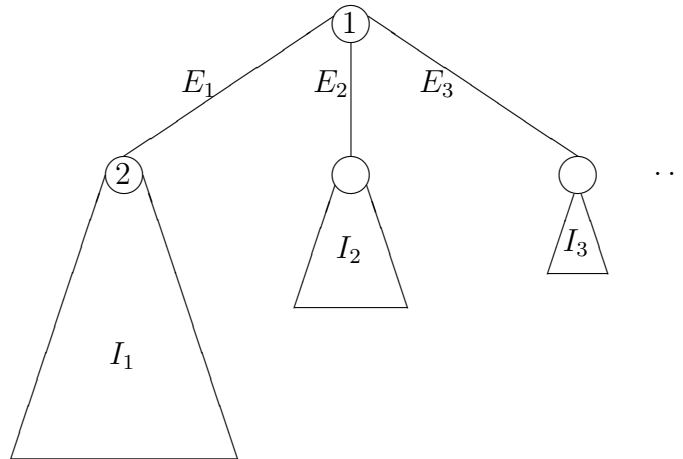


Abbildung 5: Das Modell eines rekursiven Baumes bzgl. der Größenbetrachtung der Abkömmlinge, d.h. der von den Kindern der Wurzel erzeugten Teilbäume. I_1, I_2, I_3 bezeichnen die Größen der ersten drei Abkömmlinge, E_1, E_2, E_3 sind unabhängige standard exponential verteilte ZVn, die die Lebenszeiten entsprechenden Abkömmlinge beschreiben.

ein neuer Knoten uniform an einen bereits im Baum vorhandenen Knoten angefügt.

Analog erhalten wir durch Weglassen der letzten m Knoten eines RRT^{n+m} einen zufälligen rekursiven Baum auf $\{\{1\}, \dots, \{n\}\}$. (Wäre dieser Baum einem anderen RRT^n bevorzugt oder benachteiligt, so würde sich dieser Vor- (bzw. Nach)teil durch uniformes Anfügen von m Knoten vererben. Dies widerspricht aber der Tatsache, dass keiner der ursprünglichen Bäume auf $\{\{1\}, \dots, \{n+m\}\}$ bevorzugt war).

Wir bezeichnen diese Eigenschaft als **Konsistenz** der zufälligen rekursiven Bäume.

Für eine Folge $(RRT^n)^{n \in \mathbb{N}}$ fragen wir noch nach der Anzahl der Abkömmlinge K_n des entsprechenden RRT^n . Es stellt sich heraus, dass für $n \rightarrow \infty$ gilt

$$K_n \longrightarrow \infty \quad f.s. \quad (45)$$

Dazu betrachten wir für alle $i \in \tilde{\mathbb{N}}$ das Ereignis

$$B_i := \{i\text{-ter Knoten ist Kind der Wurzel}\}. \quad (46)$$

Dann sind die Ereignisse $(B_i)_{i \in \tilde{\mathbb{N}}}$ nach Konstruktion unabhängig und es gilt $P(B_i) = \frac{1}{i-1}$ für alle $i \in \tilde{\mathbb{N}}$. Weiter ist $K_n = \sum_{i=2}^{\infty} \mathbb{1}_{B_i}$. Offenbar ist K_n monoton wachsend und wir erhalten

$$\mathbb{E}[K_n] = \sum_{i=2}^n \mathbb{E}[\mathbb{1}_{B_i}] = \log n + O(1), \quad (47)$$

$$\mathbb{V}\text{ar}[K_n] = \sum_{i=2}^n \mathbb{V}\text{ar}[\mathbb{1}_{B_i}] = \log n + O(1). \quad (48)$$

Mit der Chebyshev-Ungleichung folgt $\frac{K_n}{\log n} \xrightarrow{\mathbb{P}} 1$, für $n \rightarrow \infty$. Also geht K_n in Wahrscheinlichkeit gegen unendlich und da es monoton wächst folgt auch (45).

Wir bemerken noch: Offenbar ist $(\mathbb{1}_{B_i})_{i \in \mathbb{N}}$ eine Folge unabhängiger, quadratisch integrierbarer, fast sicher gleichmäßig beschränkter ZVn und es gilt $\mathbb{V}\text{ar}[\sum_{i=2}^n \mathbb{1}_{B_i}] \rightarrow \infty$, für $n \rightarrow \infty$. Bekanntlich implizieren diese Eigenschaften die Lyapunov-Bedingung, d.h. für K_n gilt der zentrale Grenzwertsatz

$$\frac{K_n - \mathbb{E}[K_n]}{\sqrt{\mathbb{V}\text{ar}[K_n]}} \xrightarrow{d} N(0, 1), \quad \text{für } n \rightarrow \infty. \quad (49)$$

Für einen RRT^n betrachten wir hinsichtlich asymptotischer Überlegungen für alle $i \in \mathbb{N}$ den Quotienten

$$VAF_i^n := \frac{I_{n,i}}{n-1}. \quad (50)$$

Sei nun eine Folge $(RRT^n)_{n \in \mathbb{N}}$ gegeben. Es wird sich zeigen, dass der Grenzwert $\lim_{n \rightarrow \infty} VAF_i^n$ fast sicher existiert. Dies motiviert die folgende Definition.

Definition 3.3.2. Sei $(RRT^n)_{n \in \mathbb{N}}$ eine Folge zufälliger rekursiver Bäume in Singletons. Die **asymptotische Häufigkeit** des i -ten Abkömmling ist gegeben durch

$$VAF_i := \lim_{n \rightarrow \infty} VAF_i^n. \quad (51)$$

Der **Vektor der asymptotischen Häufigkeiten**, kurz VAF , ist ein unendlich dimensionaler Vektor, dessen i -te Komponente gerade VAF_i ist.

Lemma 3.3.3. Seien $I_{n,i}$ die Größen der Abkömmlinge einer Folge $(RRT^n)_{n \in \mathbb{N}}$ in Singletons.

Für festes i bezeichne $\mathcal{F}_n := \mathfrak{S}(I_{k,i} \mid 1 \leq k \leq n)$ die kanonische Filtrierung von $(I_{n,i})_{n \in \mathbb{N}}$. Dann gilt für alle $i \in \mathbb{N}$:

(i) $(\frac{I_{n,i}}{n})_{n \in \mathbb{N}}$ ist ein Martingal bzgl. $(\mathcal{F}_n)_{n \in \mathbb{N}}$.

(ii) $(VAF_i^n)_{n \in \tilde{\mathbb{N}}} = (\frac{I_{n,i}}{n-1})_{n \in \tilde{\mathbb{N}}}$ ist ein Supermartingal bzgl. $(\mathcal{F}_n)_{n \in \tilde{\mathbb{N}}}$.

Insbesondere konvergieren die beiden Folgen fast sicher und in \mathcal{L}^p (für $p > 0$) gegen dieselbe ZV VAF_i .

Beweis: Für die Martingaleigenschaft aus (i) ist zu zeigen, dass $\forall n \in \mathbb{N}$ gilt

$$\mathbb{E} \left[\frac{I_{n+1,i}}{n+1} \mid \mathcal{F}_n \right] = \frac{I_{n,i}}{n} \quad f.s. \quad (52)$$

Wir betrachten das Ereignis $A := \{I_{n+1,i} = I_{n,i} + 1\}$, dass also der $(n+1)$ -te Knoten an den Abkömmling $A_{n,i}$ angefügt wird, d.h.

$$I_{n+1,i} = I_{n,i} + \mathbb{1}_A \quad (53)$$

Da keiner der bislang vorhandenen Knoten als Vater bevorzugt wird, folgt unter Ausnutzung der \mathcal{F}_n -Messbarkeit von $I_{n,i}$

$$\begin{aligned} \mathbb{E}[I_{(n+1),i} \mid \mathcal{F}_n] &= I_{n,i} + \mathbb{E}[\mathbb{1}_A \mid \mathcal{F}_n] \\ &= I_{n,i} + P(A \mid \mathcal{F}_n) \\ &= I_{n,i} + \frac{I_{n,i}}{n} = \frac{n+1}{n} I_{n,i} \quad f.s. \end{aligned}$$

Damit folgt (i).

Weiter folgt (ii) nach Division durch $n \in \tilde{\mathbb{N}}$

$$\mathbb{E} \left[\frac{I_{n+1,i}}{n} \mid \mathcal{F}_n \right] = \frac{n+1}{n^2} I_{n,i} = \frac{n^2-1}{n^2} \frac{I_{n,i}}{n-1} \leq \frac{I_{n,i}}{n-1} \quad f.s.$$

Da $0 \leq \frac{I_{n,i}}{n} \leq \frac{I_{n,i}}{n-1} \leq 1$ für alle $n \in \tilde{\mathbb{N}}$, konvergieren die beiden Folgen nach dem bekannten Martingalkonvergenzsatz für beschränkte Martingale fast sicher und aufgrund der Beschränktheit auch in \mathcal{L}^p , für $p > 0$. Wegen $\frac{I_{n,i}}{n} = (1 + o(1)) \frac{I_{n,i}}{n-1}$ ist der Grenzwert beider Folgen identisch. \square

Für eine genauere Größenbetrachtung der Abkömmlinge eines zufälligen rekursiven Baumes wollen wir nun noch die Verteilung des Vektors der asymptotischen Häufigkeiten VAF beschreiben.

Definition 3.3.4. Sei $(U_i)_{i \in \mathbb{N}}$ eine Folge unabhängig identisch unif $[0, 1]$ verteilter ZVn. Dann heißt ein unendlich dimensionaler Vektor V **Griffiths-Engen-McCloskey** verteilt zu den Parametern 0 und 1, kurz $V \sim GEM(0, 1)$, wenn für ihn gilt

$$V = (V_1, V_2, \dots, V_i, \dots) \sim \left(U_1, U_2(1 - U_1), \dots, U_i \prod_{k=1}^{i-1} (1 - U_k), \dots \right). \quad (54)$$

Der Ausdruck (54) ist dabei folgendermaßen zu lesen: Die erste Komponente ist uniform auf $[0, 1]$ verteilt. Die zweite Komponente ist uniform auf dem Rest $R_1 := 1 - U_1$ verteilt, der von der ersten Komponente erzeugt wurde. Sukzessive ist die $(i + 1)$ -te Komponente uniform auf dem Rest R_i der ersten i Komponenten verteilt.

Satz 3.3.5. *Für eine Folge $(RRT^n)^{n \in \mathbb{N}}$ zufälliger rekursiver Bäume in Singletons gilt*

$$VAF \sim GEM(0, 1). \quad (55)$$

Beweis: Es ist zu zeigen

$$(VAF_1, VAF_2, \dots, VAF_i, \dots) \sim \left(U_1, U_2(1 - U_1), \dots, U_i \prod_{k=1}^{i-1} (1 - U_k), \dots \right), \quad (56)$$

wobei $(U_i)_{i \in \mathbb{N}}$ eine Folge unabhängiger identisch unif $[0, 1]$ -verteilter ZVn bezeichne.

Nach Lemma 3.3.3 gilt für alle $i \in \mathbb{N}$, dass der Quotient $\frac{I_{n,i}}{n-1}$ für $n \rightarrow \infty$ fast sicher konvergiert. Damit folgt auch (für $n \rightarrow \infty$) direkt die fast sichere Konvergenz von

$$\frac{I_{n,i}}{(n-1) - \sum_{j=1}^{i-1} I_{n,j}} = \frac{I_{n,i}}{n-1} \frac{1}{\frac{(n-1) - \sum_{j=1}^{i-1} I_{n,j}}{n-1}} \quad (57)$$

und wir setzen für alle $i \in \mathbb{N}$

$$U_i := \lim_{n \rightarrow \infty} \frac{I_{n,i}}{(n-1) - \sum_{j=1}^{i-1} I_{n,j}}. \quad (58)$$

U_i ist also der Grenzwert des Quotienten aus der Größe des i -ten Abkömmlings und der Anzahl aller Knoten, die nicht in den ersten $i-1$ Abkömmlingen enthalten sind.

Wir zeigen nun via Vollständiger Induktion über i

$$\frac{(n-1) - \sum_{j=1}^i I_{n,j}}{n-1} \xrightarrow{n \rightarrow \infty} \prod_{j=1}^i (1 - U_j) \quad \text{f.s.} \quad \forall i \in \mathbb{N}. \quad (59)$$

Der Induktionsanfang $i = 1$ folgt direkt aus (58), denn

$$1 - \frac{I_{n,1}}{n-1} \xrightarrow{n \rightarrow \infty} 1 - U_1 \quad \text{f.s.}$$

Sei nun die Behauptung für i bewiesen, d.h. es gilt (59). Dann folgt für $i + 1$

$$\begin{aligned}
& \frac{(n-1) - \sum_{j=1}^{i+1} I_{n,j}}{n-1} \\
= & \frac{(n-1) - \sum_{j=1}^i I_{n,j}}{n-1} - \frac{I_{n,i+1}}{n-1} \\
= & \frac{(n-1) - \sum_{j=1}^i I_{n,j}}{n-1} \left(\frac{I_{n,i+1}}{(n-1) - \sum_{j=1}^i I_{n,j}} - \frac{(n-1) - \sum_{j=1}^i I_{n,j}}{n-1} \right) \\
= & \frac{(n-1) - \sum_{j=1}^i I_{n,j}}{n-1} \left(1 - \frac{I_{n,i+1}}{(n-1) - \sum_{j=1}^i I_{n,j}} \right) \\
\stackrel{n \rightarrow \infty}{\longrightarrow} & \prod_{j=1}^i (1 - U_j)(1 - U_{i+1}) = \prod_{j=1}^{i+1} (1 - U_j) \quad \text{f.s.}
\end{aligned}$$

Somit gilt die Behauptung (59) für alle $i \in \mathbb{N}$ und es folgt

$$\frac{I_{n,i}}{n-1} = \frac{I_{n,i}}{(n-1) - \sum_{j=1}^{i-1} I_{n,j}} \frac{(n-1) - \sum_{j=1}^{i-1} I_{n,j}}{n-1} \stackrel{n \rightarrow \infty}{\longrightarrow} U_i \prod_{j=1}^{i-1} (1 - U_j) \quad \text{f.s.} \quad (60)$$

für alle $i \in \mathbb{N}$, was die gemeinsame Verteilung

$$(VAF_1, VAF_2, \dots, VAF_i, \dots) \sim \left(U_1, U_2(1 - U_1), \dots, U_i \prod_{k=1}^{i-1} (1 - U_k), \dots \right)$$

liefert. Es bleibt noch zu zeigen, dass $(U_i)_{i \in \mathbb{N}}$ unabhängig identisch unif $[0, 1]$ verteilt sind.

Zunächst zeigen wir $U_1 \sim \text{unif}[0, 1]$ und dafür

$$I_{n,1} \sim \text{unif}\{1, \dots, n-1\}. \quad (61)$$

Dazu betrachten wir das Ereignis $\{I_{n,1} = k\}$, dass also der erste Abkömmling $A_{n,1}$ aus genau k Knoten besteht, $k \in [n-1]$. Da per definitionem gilt: (1) $\notin A_{n,1}$ und (2) $\in A_{n,1}$ haben wir genau $\binom{n-2}{k-1}$ Möglichkeiten $k-1$ Knoten aus der Menge $\{(3), (4), \dots, (n)\}$ zu wählen.

Aus der (2) und den $k-1$ gewählten Knoten lässt sich auf $(k-1)!$ Weisen ein rekursiver Baum bilden, der dann den ersten Abkömmling beschreibt.

Für die übrigen $n-k$ Elemente gibt es genau $(n-k-1)!$ Möglichkeiten einen rekursiven "Restbaum" zu bilden. Dieser enthält insbesondere die (1).

Schließlich fügen wir den Abkömmling (\ni (2)) mit dem Restbaum (\ni (1)) zusammen, indem wir die Knoten (1) und (2) durch eine Kante verbinden.

Insgesamt erhalten wir so $\binom{n-2}{k-1}(k-1)!(n-k-1)!$ mögliche rekursive Bäume, die das Ereignis $\{I_{n,1} = k\}$ begünstigen.

Da die Anzahl aller rekursiven Bäume mit n Knoten genau $(n-1)!$ beträgt, ergibt sich

$$P(I_{n,1} = k) = \frac{\binom{n-2}{k-1}(k-1)!(n-k-1)!}{(n-1)!} = \frac{1}{n-1}, \quad \forall k \in [n-1].$$

Also gilt (61). Dann folgt aber direkt

$$I_{n,1} \sim [U(n-1)] = (n-1)U + O(1), \quad (62)$$

mit $U \sim \text{unif}[0, 1]$. Skalieren mit $n-1$ und anschließende Grenzwertbildung liefert die gewünschte Verteilung von U_1 .

(Betrachten wir eine Pólya Urne, die zu Beginn eine schwarze und eine weiße Kugel enthält, so bemerken wir, dass sich die Größe des ersten Abkömmlings wie die Anzahl der (o.B.d.A.) schwarzen Kugeln der Pólya Urne entwickelt. Denn seien die Knoten des ersten Abkömmlings schwarz und die des Restbaumes weiß eingefärbt, so ist nach Konstruktion die (1) weiß und die (2) schwarz fast sicher. Als Vater für die (3) ist keiner der beiden Knoten bevorzugt. Wird die (2) der Vater, so wird die (3) schwarz gefärbt, ist (1) der Vater, dann wird die (3) weiß, etc. Sukzessive erhalten wir (61) aufgrund der Konstruktion der zufälligen rekursiven Bäume).

Sei nun $i \in \tilde{\mathbb{N}}$. Betrachten wir beim Aufbau unseres Baumes nur die Knoten, welche an die Abkömmlinge A_i, A_{i+1}, \dots angebracht werden, und "ignorieren" gewissermaßen die ersten $i-1$ Abkömmlinge, so entwickelt sich die Größe des i -ten Abkömmlings A_i nach der gleichen Dynamik wie die des ersten Abkömmlings bei der "konventionellen" Baumkonstruktion. Da schließlich genau $N := (n-1) - \sum_{j=1}^{i-1} I_{n,j}$ Knoten an die Abkömmlinge A_i, A_{i+1}, \dots angebracht werden, folgt

$$\frac{I_{n,i}}{(n-1) - \sum_{j=1}^{i-1} I_{n,j}} \sim \frac{I_{N,1}}{N-1}, \quad \forall i \in \mathbb{N}. \quad (63)$$

Da diese Größen aufgrund der Gestalt der zufälligen rekursiven Bäume unabhängig sind, folgt die Behauptung. \square

Bemerkung: Da insbesondere $VAF_1 \sim \text{unif}[0, 1]$ wird sich im Mittel schon im ersten Abkömmling die Hälfte der Masse des gesamten Baumes befinden. Genauer gilt $\forall i \in \mathbb{N}$

$$\mathbb{E}[VAF_i] = \mathbb{E} \left[U_i \prod_{k=1}^{i-1} (1 - U_k) \right] = \mathbb{E}[U_i] \prod_{k=1}^{i-1} \mathbb{E}[(1 - U_k)] = \frac{1}{2^i}. \quad (64)$$

Also für $m < n$, mit $m, n \in \mathbb{N}$

$$\mathbb{E}[VAF_m] = \frac{1}{2^m} > \frac{1}{2^n} = \mathbb{E}[VAF_n]. \quad (65)$$

Daraus folgt natürlich nicht, dass ein Abkömmling fast sicher kleiner sein wird als alle Abkömmlinge kleineren Rangs. Mit anderen Worten: Für einen $GEM(0, 1)$ -verteilten Vektor V wird die Folge seiner Komponenten $(V_i)_{i \in \mathbb{N}}$ i.a. nicht monoton fallend sein. Da in der Baumkonstruktion die Zerschneidezeitpunkte aller Kanten durch unabhängige standard exponentialverteilte ZVn beschrieben werden, also kein Abkömmling bzgl. seiner Lebenszeit bevorzugt ist, wollen wir die Ränge der Abkömmlinge in unserer Betrachtung vernachlässigen und sie gemäß der folgenden Definition anordnen.

Definition 3.3.6. Sei $V \sim GEM(0, 1)$ und $J_i := V_{(i)} \forall i \in \mathbb{N}$, wobei $V_{(i)}$ die i -t größte Komponente von V bezeichne. Dann heißt der unendlich dimensionale Vektor $J = (J_1, J_2, \dots)$ **Poisson-Dirichlet** verteilt zu den Parametern 0 und 1, kurz $J \sim PD(0, 1)$.

Bezeichnen wir mit $F_t(1)$ die relative Größe des Blockes eines BSC , der zur Zeit $t \geq 0$ die (1) enthält, so erhalten wir bei Betrachtung des BSC anhand der Baumkonstruktion als Folgerung von Satz 3.3.5:

Korollar 3.3.7. Seien $J_1 \geq J_2 \geq \dots$ die der Größe nach abfallend geordneten Sprungrößen des Prozesses $(F_t(1))_{t \geq 0}$, und sei für alle $i \in \mathbb{N}$ mit T_i der Zeitpunkt bezeichnet zu dem der Sprung der Größe J_i stattfindet. Dann gilt

$$J = (J_1, J_2, \dots) \sim PD(0, 1),$$

mit J, T_1, T_2, \dots unabhängig und $T_i \sim \exp(1) \forall i \in \mathbb{N}$.

Bemerkung: Am Anfang dieses Kapitels hatten wir gesehen, dass die Anzahl der Abkömmlinge fast sicher beliebig groß wird (vgl. (45)). Dabei sollte sich ein "Großteil" der Knoten auf relativ "wenige" Abkömmlinge verteilen. (vgl. (64)). Da im Hinblick auf die letzte Kollision keiner der Abkömmlinge bevorzugt ist, bedeutet dies, dass tendenziell eher ein kleiner Abkömmling in der letzten Verschmelzung involviert ist. Im BSC sollten also unmittelbar vor der letzten Kollision sehr viele der Singletons zu einem riesigen Block verschmolzen sein, neben dem nur noch wenige Mini-Blöcke existieren.

Diese Überlegung werden wir in Abschnitt 4 - im Hauptsatz dieser Arbeit - anhand einer genaueren Analyse der Asymptotik gewisser Blockgrößen verifizieren können.

3.4 Der Bonsai im Chinarestaurant

Hinsichtlich unserer Betrachtung waren lediglich die Größen $I_{n,i}$ der Abkömmlinge eines Baumes von Interesse. Diese erhalten wir als Spezialfall folgender Klasse von Prozessen zurück.

Definition 3.4.1. Sei $\alpha \in [0, 1)$, $\theta > -\alpha$ und $\mathcal{P}([n])$ die Menge der Partitionen von $[n]$. Ein **China-Restaurant-Prozess** zu den Parametern α und θ , kurz $(CRP^{[n]}(\alpha, \theta))_{n \in \mathbb{N}}$, ist ein diskreter - im n -ten Schritt $\mathcal{P}([n])$ -wertiger - stochastischer Prozess, dessen Dynamik gegeben ist durch

(i) Es gilt $CRP^{[1]}(\alpha, \theta) = (1)$.

(ii) Angenommen die ersten m Elemente ($m \in [n - 1]$) haben k Blöcke gebildet, wobei der i -te Block aus $|B_i|$ Elementen besteht, mit $m = \sum_{i=1}^k |B_i|$. Dann gilt für das $(m + 1)$ -te Element:

$$P((m + 1) - \text{tes Element bildet einen neuen Block}) = \frac{\theta + \alpha k}{m + \theta},$$

$$P((m + 1) - \text{tes Element wird an Block } i \text{ angefügt}) = \frac{|B_i| - \alpha}{m + \theta},$$

für $1 \leq i \leq k$.

Wir bemerken den Spezialfall $\alpha = 0$ und $\theta = 1$. Dann entspricht die Anfügewahrscheinlichkeit des m -ten Elements im China-Restaurant-Prozess genau der des $(m + 1)$ -ten Knoten eines zufälligen rekursiven Baums, wenn die Blöcke durch die Abkömmlinge des Baumes beschrieben werden. M.a.W.: Die Entwicklung der Größe des i -ten Abkömmlings im Baum erfolgt nach der gleichen Dynamik mit der der i -te Block (Block vom Rang i) im China-Restaurant-Prozess wächst. Die Anordnung der Blöcke eines $CRP^{[n]}(\alpha, \theta)$ erfolgt dabei nach der bekannten " kE'' "-Relation.

Wir betrachten die Blockgröße $|B_i^{[n]}|$ des i -ten Blocks eines $CRP^{[n]}(\alpha, \theta)$, sofern der Block existiert und weiter die Größe des i -ten Abkömmlings eines RRT^{n+1} . Dann gilt

$$|B_i^{[n]}| \sim I_{n+1,i}, \quad \forall n \in \mathbb{N}. \quad (66)$$

Die Analogie zum Baum setzt sich damit auch in der Asymptotik fort. Nach Lemma 3.3.3(i) folgt, dass $(\frac{|B_i^{[n]}|}{n+1})_{n \in \mathbb{N}}$ ein Martingal ist und für $n \rightarrow \infty$ fast sicher konvergiert. Es liegt nun nahe auch hier von den asymptotischen Häufigkeiten der Blöcke zu sprechen.

Entscheidend ist nun für uns, dass entsprechende Resultate auch auf allgemeines α und θ übertragbar sind. Es gilt

Satz 3.4.2. Sei $(CRP^{[n]}(\alpha, \theta))_{n \in \mathbb{N}}$ ein Chinesischer-Restaurant-Prozess mit $\alpha \in [0, 1)$ und $\theta > -\alpha$. Dann gilt:

Für alle $i \in \mathbb{N}$ existiert die asymptotische Häufigkeit

$VAF_i(\alpha, \theta) := \lim_{n \rightarrow \infty} \frac{|B_i^{[n]}|}{n+1} = \lim_{n \rightarrow \infty} \frac{|B_i^{[n]}|}{n}$ f.s. und ihre gemeinsame Verteilung ist gegeben durch

$$(VAF_1(\alpha, \theta), VAF_2(\alpha, \theta), \dots, VAF_i(\alpha, \theta), \dots) \\ \sim \left(W_1, W_2(1 - W_1), \dots, W_i \prod_{j=1}^{i-1} (1 - W_j), \dots \right), \quad (67)$$

wobei $(W_i)_{i \in \mathbb{N}}$ eine Folge unabhängiger ZVn bezeichne, mit $W_i \sim \text{beta}(1 - \alpha, \theta + i\alpha)$.

Ein ausführlicher Beweis von Satz 3.4.2 ist in [6] gegeben. (Pitman wählt einen Zugang über zufällige austauschbare Partitionen, um Aussagen über die Asymptotik der relativen Blockgrößen von Partitionen zu formulieren. Dabei erfüllt ein *CRP* gerade die nötigen Austauschbarkeitsbedingungen, was schließlich die Behauptungen liefert.)

Wir bemerken die Analogie zur Verteilung der asymptotischen Häufigkeiten eines zufälligen rekursiven Baumes, vgl. Definition 3.3.4. Dies ist in gewisser Weise ersichtlich, da im *CRP* zwar die Anfügewahrscheinlichkeiten neu zu setzender Elemente verallgemeinert wurden, die Blöcke aber prinzipiell immer noch anhand des gleichen Verfahrens aufgebaut werden. Auch hier erhalten wir sofort den Spezialfall der zufälligen rekursiven Bäumen zurück, da für alle $i \in \mathbb{N}$ mit $\alpha = 0$ und $\theta = 1$ gilt: $W_i \sim \text{beta}(1, 1)$, was der uniformen Verteilung auf $[0, 1]$ entspricht. Ein anderer Hinweis auf die Gültigkeit von (67) liefert eine genauere Analyse der Beta-Verteilung zum Parameter $(1 - \alpha, \theta + i\alpha)$. Die Änderung der Parameter α und θ wird dabei etwa den gleichen Effekt auf die asymptotische Größe der Blöcke haben, den wir auch intuitiv nach entsprechender Veränderung der Übergangswahrscheinlichkeiten für den CRP "im Kleinen" erwarten würden.

(Betrachten wir beispielsweise für den Fall $\alpha = 0$ den ersten Abkömmling: Die Dichte von $VAF_1(0, \theta)$ ist dann durch $f(x) = \frac{1}{B(1, \theta)}(1 - x)^{\theta-1} \mathbb{1}_{[0,1]}(x)$ gegeben, wobei $B(1, \theta)$ lediglich der Skalierung dient. Den Parameter θ wachsen zu lassen bedeutet dabei für die Dichte anschaulich, dass die Masse immer weiter Richtung 0 verschoben wird. Für großes θ wird die asymptotische Häufigkeit des ersten Blocks also erwartungsgemäß nur unwesentlich größer sein als die der Folgenden. Ähnliches erwarten wir auch bei der Betrachtung der Übergangswahrscheinlichkeiten beim Aufbau des *CRP*. Für wachsendes

θ wird insbesondere für die ersten zu involvierenden Elemente die Wahrscheinlichkeit groß, immer einen neuen Block zu bilden, was nachhaltig nur zu einem geringen Größenunterschied der ersten Blöcke führen sollte.) Es liegt nun nahe auch die Definitionen 3.3.4 und 3.3.6 zu verallgemeinern.

Definition 3.4.3. Sei $\alpha \in [0, 1)$ und $\theta > -\alpha$ und weiter $(W_i)_{i \in \mathbb{N}}$ eine Folge unabhängiger ZVn, mit $W_i \sim \text{beta}(1 - \alpha, \theta + i\alpha)$, $\forall i \in \mathbb{N}$.

- (i) Ein unendlich dimensionaler Vektor V heißt **Griffiths-Engen-McCloskey** verteilt zu den Parametern α und θ , kurz $V \sim GEM(\alpha, \theta)$, falls gilt

$$(V_1, V_2, \dots, V_i, \dots) \sim \left(W_1, W_2(1 - W_1), \dots, W_i \prod_{j=1}^{i-1} (1 - W_j), \dots \right). \quad (68)$$

- (ii) Sei $V \sim GEM(\alpha, \theta)$ und $J_i := V_{(i)} \forall i \in \mathbb{N}$, wobei $V_{(i)}$ die i -t größte Komponente von V bezeichne. Dann heißt der unendlich dimensionale Vektor $J = (J_1, J_2, \dots)$ **Poisson-Dirichlet** verteilt zu den Parametern α und θ , kurz $J \sim PD(\alpha, \theta)$.

Mit dieser Definition und den Aussagen von Satz 3.4.2 folgt nun direkt

Korollar 3.4.4. Sei ein CRP(α, θ) gegeben. Der Vektor der asymptotischen Häufigkeiten der Blöcke ist $GEM(\alpha, \theta)$ -verteilt. Ordnet man die Komponenten der Größe nach abfallend an, so ist der dadurch entstandene Vektor $PD(\alpha, \theta)$ -verteilt.

Mit dem erlangten Begriffsapparat können wir nun anhand der rekursiven Bäume eine Aussage über die Verteilung der Blockgrößen eines BSC zur festen Zeit $t > 0$ machen. Seien zur Zeit $t > 0$ die Blöcke B_1, B_2, \dots wieder nach der " $\stackrel{kE''}{<}$ "-Relation angeordnet. Dann betrachten wir für $i \in \mathbb{N}$ die Quotienten

$$\frac{|B_i \cap [n]|}{n}. \quad (69)$$

Es wird sich herausstellen, dass auch diese Größen für wachsendes n fast sicher konvergieren. Den Grenzwert bezeichnen wir dann als **asymptotische Häufigkeit** von B_i zur Zeit t .

Von der Vorstellung des Baumes als Ansammlung von Abkömmlingen und einer Wurzel wollen wir im Folgenden wieder Abstand nehmen. Zwar werden wir sowohl mit einem zufälligen rekursiven Baum als auch dem China-Restaurant-Prozess arbeiten, dabei sind die Blöcke des CRP aber nicht mehr

mit den Abkömmlingen des Baumes zu identifizieren. Sie werden nun kanonisch den Blöcken des BSC - und damit auch den entsprechenden Marken im Baum - zugeordnet.

Satz 3.4.5. (*J. Schweinsberg*). Sei $(BSC)_{t \geq 0}$ ein unrestringierter Bolthausen-Sznitman Koaleszent. Für alle Zeiten $t \geq 0$ existieren für alle Blöcke die asymptotischen Häufigkeiten. Sind die Blöcke gemäß der " $\stackrel{kE''}{<}$ "-Relation angeordnet, so ist der dadurch entstandene Vektor $GEM(e^{-t}, 0)$ verteilt. Ordnet man sie der Größe nach abfallend an, so ist der entsprechende Vektor $PD(e^{-t}, 0)$ verteilt.

Beweis: Anstatt den Bolthausen-Sznitman Koaleszenten zu untersuchen betrachten wir gemäß Satz 3.2.6 einen zufälligen rekursiven Baum in Singletons. Dabei ist die Kante mit der der i -te Knoten angefügt ist wieder mit einer Lebenszeit E_i versehen, $i \in \tilde{\mathbb{N}}$. $((E_i)_{i \in \tilde{\mathbb{N}}}$ ist wie früher gewählt). Wir markieren nun genau jene Kanten, deren Lebenszeit kleiner t ist. Die Idee ist nun, die Knoten des Baumes zwar nicht sukzessive zu verschmelzen, aber trotzdem - anhand der Markierungen - die entsprechende Blockbildung im Bolthausen-Sznitman Koaleszenten zu betrachten. (vgl. Abbildung 6)

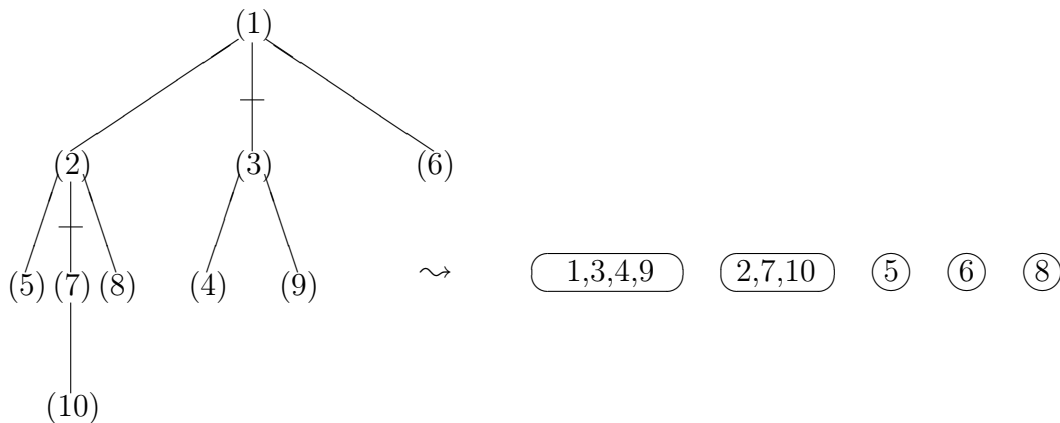


Abbildung 6: Links: ein ungeschnittener RRT^{10} . Die Markierungen der Kanten deuten an, dass deren Lebenszeit kleiner einem gegebenem $t > 0$ ist. Rechts: die dem Baum entsprechende Blockbildung im Bolthausen-Sznitman Koaleszenten bzw. im China-Restaurant-Prozess.

Damit ist es unser Ziel zu zeigen, dass der nächste Knoten mit den Übergangswahrscheinlichkeiten eines $CRP(e^{-t}, 0)$ angebracht wird. Nach dem vorherigen Korollar 3.4.4 folgt dann die Behauptung.

Zu zeigen ist also: Angenommen die n Knoten im Baum haben entsprechend

der Markierungen der Kanten genau k Blöcke B_1, \dots, B_k gebildet, wobei der i -te Block B_i aus $|B_i|$ Elementen besteht, mit $n := \sum_{i=1}^k |B_i|$. Dann gilt für das $(n+1)$ -te Element

$$(a) \quad P(\text{Das } (n+1)\text{-te Element bildet einen neuen Block}) = \frac{ke^{-t}}{n},$$

$$(b) \quad P(\text{Das } (n+1)\text{-te Element wird an } B_i \text{ angefügt}) = \frac{|B_i| - e^{-t}}{n}.$$

Zu (a): Sei \mathcal{A} das Ereignis, dass der $(n+1)$ -te Knoten an das kleinste Element einer der k Blöcke angebracht wird. Und \mathcal{B} ist das Ereignis, dass die Kante mit der er angefügt wird nicht markiert ist. Dann gilt trivialerweise

$$\{\text{Das } (n+1)\text{-te Element bildet einen neuen Block}\} \Leftrightarrow \mathcal{A} \cap \mathcal{B} \text{ tritt ein,}$$

denn wird es nicht an das kleinste Element einer der k Blöcke angebracht, dann ist es nach Konstruktion (unabhängig vom Wert der angebrachten ZV) schon Teil des Blocks, in dem auch dasjenige Element integriert ist, an das das $(n+1)$ -te schließlich angefügt wird.

Da sich \mathcal{A} auf die Struktur des Baumes bezieht und \mathcal{B} auf den Wert der angebrachten Kante, sind \mathcal{A} und \mathcal{B} unabhängig.

Es gilt nun $P(\mathcal{A}) = \frac{k}{n}$, weil als Vater des neuen Knotens keiner der bislang vorhandenen Knoten bevorzugt wird.

Sei E_{n+1} die Lebenszeit der Kante, mit der der $(n+1)$ -te Knoten angefügt wird.

Dann gilt $P(\mathcal{B}) = P(E_{n+1} \geq t) = e^{-t}$. Insgesamt folgt damit

$$P(\mathcal{A} \cap \mathcal{B}) = P(\mathcal{A})P(\mathcal{B}) = \frac{ke^{-t}}{n},$$

somit gilt (a).

Zu (b): Wir betrachten jetzt einen der vorhandenen Blöcke B_i . Dieser habe $|B_i|$ Elemente. Dann gilt

$$\{\text{Das } (n+1)\text{-te Element wird an } B_i \text{ angefügt}\} \Leftrightarrow \mathcal{C} \text{ oder } \mathcal{D} \text{ tritt ein,}$$

mit $\mathcal{C} := \{\text{Der } (n+1)\text{-te Knoten wird mit einer markierten Kante an das kleinste Element von } B_i \text{ angebracht.}\}$

und $\mathcal{D} := \{\text{Der } (n+1)\text{-te Knoten wird (durch eine markierte oder unmarkierte Kante) an eines der } (|B_i| - 1) \text{ nicht-kleinsten Elemente von } B_i \text{ angefügt.}\}$.

Da die Markierung unabhängig vom Knoten ist, an den angefügt wird, folgt $P(\mathcal{C}) = P(\mathcal{A})P(\mathcal{B}^c) = \frac{1-e^{-t}}{n}$.

Außerdem ist $P(\mathcal{D}) = \frac{|B_i|-1}{n}$, wegen des uniformen Anfügens neuer Knoten. Da \mathcal{C} und \mathcal{D} disjunkte Ereignisse sind, folgt schließlich

$$P(\mathcal{C} \cup \mathcal{D}) = P(\mathcal{C}) + P(\mathcal{D}) = \frac{|B_i| - e^{-t}}{n},$$

d.h. es gilt auch (b).

□

4 Die letzte Verschmelzung im BSC

Hinsichtlich des letzten Übergangs im Bolthausen-Sznitman Koaleszenten hatten wir schon in Abschnitt 3.3 auf die Besonderheit der Wurzel im Baum - also den Knoten dessen Marke die (1) enthält - hingewiesen. (vgl. Korollar 3.3.7). Entsprechend wurde der Begriffsapparat schon dort im Hinblick auf die letzte Verschmelzung gewählt. Wir werden deshalb in der Begriffsbildung ähnlich vorgehen.

4.1 Der Hauptsatz

Für $n \in \mathbb{N}$ sei ein auf $[n]$ restringierter $BSC^{[n]}$ gegeben. Wir hatten jedem Block eine Marke angefügt, die die Information über die in ihm vereinigten Singletons enthält. Als **Blockgröße** oder **Masse** eines Blocks bezeichnen wir dann die Anzahl der Singletons, die in ihm zusammengefasst sind.

Zum Zeitpunkt der letzten Kollision betrachten wir folgende zwei Größen:

- $|M_{(\neq 1)}^{[n]}|$ bezeichnet die Summe der Massen aller Blöcke, die nicht das Element (1) enthalten.
- $\#B^{[n]}$ beschreibt die Anzahl der involvierten Blöcke.

Mit diesen Notationen gilt beispielsweise für den BSC der Abbildung 1 (auf Seite 14): $n = 10$, $|M_{(\neq 1)}^{[n]}| = 5$ und $\#B^{[n]} = 3$.

Wir formulieren nun die Hauptaussage dieser Arbeit.

Hauptsatz 4.1.1. *Sei ein auf $[n]$ restringierter $BSC^{[n]}$ gegeben. Für $n \rightarrow \infty$ gilt*

$$\left(\frac{\log |M_{(\neq 1)}^{[n]}|}{\log n}, \#B^{[n]} \right) \xrightarrow{d} (U, 1 + Y_{UE}). \quad (70)$$

Dabei ist $Y = (Y_t)_{t \geq 0}$ ein standard Yule Prozess, U und E sind ZVn mit $U \sim \text{unif}[0, 1]$ und $E \sim \exp(1)$. Zudem sind Y, U und E unabhängig.

Bevor wir mit dem Beweis dieses Satzes beginnen, stellen wir einige Betrachtungen voran.

Bemerkungen: (i): Die Konvergenz $\frac{\log |M_{(\neq 1)}^{[n]}|}{\log n} \xrightarrow{d} U$ der ersten Komponente liefert für großes n näherungsweise

$$|M_{(\neq 1)}^{[n]}| \approx n^U. \quad (71)$$

d.h. die Masse aller Blöcke, die in der letzten Verschmelzung nicht die (1) enthalten, sollten in Bezug auf die Gesamtmasse sehr klein sein. Für großes

n werden also die Singletons im Bolthausen-Sznitman Koaleszenten mit der Zeit zu einem großen Block verschmelzen, der fast alle Masse enthält. Man beachte, dass genauer gilt $|M_{(\neq 1)}^{[n]}| = n^{U+o(1)}$. Der Fehler steht also im Exponenten. Das Ergebnis ist damit in gewisser Weise sehr grob. Beispielsweise sind Aussagen über den Erwartungswert ohne Weiteres nicht möglich.

(ii): Die zweite Komponente gibt uns Auskunft über die Blockbildung der übrigen Elemente, die nicht in dem eben betrachteten großen Block verschmolzen sind. Intuitiv würden wir erwarten, dass sie sich auf wenige Blöcke verteilen, also die betrachtete Blockanzahl $\#B^{[n]}$ eher klein vorzufinden ist. Denn wäre die Anzahl der Blöcke groß, dann hieße das für die Baumkonstruktion, dass speziell in dem Abkömmling, der für die finale Kollision übrig bleibt, selbst nur "wenige" Kanten geschnitten wurden, aber alle übrigen Abkömmlinge notwendigerweise schon in die Wurzel verschmolzen sind. Dies ist eher untypisch, wenn man bedenkt, dass keine der Kanten bzgl. ihres Zerschneidezeitpunktes bevorzugt ist.

Wir erinnern an Lemma 2.2.8. Es gilt für alle $m \in \mathbb{N}$

$$P(Y_{UE} = m) = \frac{1}{m} \sum_{k=1}^m \binom{m}{k} (-1)^{k+1} \log(k+1),$$

also insbesondere $P(\#B^{[n]} = 2) \xrightarrow{d} P(Y_{UE} = 1) = \log 2 \approx 0,69$ und $P(\#B^{[n]} \in \{2, 3\}) \xrightarrow{d} P(Y_{UE} \in \{1, 2\}) \approx 0,84$, für $n \rightarrow \infty$. Somit ist die Wahrscheinlichkeit in der letzten Verschmelzung neben dem großen Block nur noch einen bzw. zwei andere Blöcke vorzufinden tatsächlich recht hoch, wenn wir n nur groß genug werden lassen. Weiter ist der Erwartungswert von Y_{UE} unendlich (vgl. wieder Lemma 2.2.8). Somit werden die betrachteten Wahrscheinlichkeitsgewichte $P(Y_{UE} = m)$ für wachsendes m zunächst stark abfallen (weil schon viel Masse auf den ersten Knoten lastet), dann aber relativ langsam kleiner werden.

(iii): Schließlich gehen wir noch auf den Aspekt der gemeinsamen Verteilung ein. In beiden Komponenten taucht die gleiche uniform verteilte ZV U auf. Dies lässt sich wieder mit unserem Baummodell erklären. In der finalen Kollision muss eine der Kanten der Wurzel geschnitten werden. Dafür ist aber nach Konstruktion keine der Wurzelkanten bevorzugt. Im entsprechenden Abkömmling erkennen wir beide betrachteten Größen wieder. $|M_{(\neq 1)}^{[n]}|$ entspricht gerade dessen Gesamtmasse und $\#B^{[n]}$ beschreibt die Anzahl der Blöcke, die sich innerhalb des Abkömmlings bis zur finalen Kollision gebildet haben.

Außerdem werden wir später die ZV E mit der Lebenszeit dieses Abkömmlings in Verbindung bringen. Wir bemerken jetzt schon, dass diese Lebenszeit keine Bedeutung für die Anzahl der Elemente des Abkömmlings hat, wohl

aber für die interne Blockbildung. Deshalb taucht E nur in der zweiten Komponente auf. Ist die Lebenszeit groß, so sollten innerhalb des Abkömmlings tendenziell schon viele Schnitte stattgefunden haben, d.h. $\#B^{[n]}$ wird eher klein sein. (Für großes E wird aber Y_{UE} selbst groß (bei festem U), d.h. E sollte um so größer werden, je kleiner die Lebenszeit des ausgezeichneten Abkömmlings ist.)

Für wachsendes U (1. Komponente) wird auch, wenn man E fest hält, das Argument des Yule-Prozesses (2. Komponente) größer und damit wird der Prozess selbst tendenziell einen höheren Wert annehmen. Auch dies entspricht unserer Vorstellung: Wenn die Masse der kleinen Blöcke zunimmt, so sollte sie sich auch auf mehr und mehr Blöcke verteilen.

4.1.1 Notation und Konstruktion im Hauptbeweis

Wir wollen auch hier wieder gemäß unserer Korrespondenz zu den zufälligen rekursiven Bäumen vorgehen. (vgl. Satz 3.2.6.) Dabei lässt sich der Aufbau eines solchen Baumes in einfacher Weise durch einen standard Yule-Prozess simulieren. (Das Abholz-Prozedere bleibt zunächst außen vor.) Für alle $i \in \mathbb{N}$ entspricht der i -te Knoten im Baum dem i -ten Individuum, das im Yule-Prozess geboren wird. Nach Konstruktion der zufälligen rekursiven Bäume wird als Vater eines neu ankommenden Knotens keiner der bislang vorhandenen Knoten bevorzugt. Diese Eigenschaft finden wir auch im Yule-Prozess wieder. Denn per definitionem sind die Wartezeiten auf die nächste Geburt standard exponential verteilt und voneinander unabhängig. Angenommen es seien zur Zeit t (mit $t \geq 0$) genau i Individuen vorhanden, $i \in \mathbb{N}$. Wegen der Gedächtnislosigkeit der Exponentialverteilung können wir den Startzeitpunkt der zufälligen Wartezeit auf die nächste Geburt für jedes existierende Individuum auf den Wert t setzen. Da alle Wartezeiten unabhängig und identisch verteilt sind, wird auch hier keines der Individuen für die nächste Geburt präferiert.

Dieses Grundmodell wollen wir noch etwas erweitern. In Abschnitt 3.3 hatten wir auf die Besonderheit der Wurzel (1) im Baum hingewiesen und deren Kinder selbst als Wurzel eines eigenen Baumes betrachtet. Den Teilbaum dessen Wurzel das i -te Kind der Wurzel (1) ist, hatten wir als i -ten Abkömmling bezeichnet. Auch hierbei wollen wir die Analogie wahren. Zunächst wird das erste, seit dem Zeitpunkt $t = 0$ existierende Individuum mit (1) bezeichnet. Den zugrunde liegenden Yule-Prozess nennen wir kurz Y .

Für alle $i \in \mathbb{N}$ bezeichnet dann der i -te **Immigrant (von Y)** das i -te Kind der (1). Die Geburten der (1) heißen also speziell **Immigrationen**.

Die Struktur des Prozesses gestattet uns nun (nach Umbenennung der Marken),

den i -ten Immigranten als Initiator eines eigenen standard Yule-Prozesses Y^i zu betrachten, welcher eben genau zu dessen Geburtszeitpunkt startet. Da im Prozess Y die Entwicklung eines Individuums unabhängig von den anderen Individuen verläuft, ist auch die Folge der neu erklärten Prozesse $(Y^i)_{i \in \mathbb{N}}$ unabhängig. (Die Immigranten bilden einen homogenen Poisson-Prozess mit Rate 1.) Mit dieser Notation korrespondiert Y^i also mit dem i -ten Abkömmling im zufälligen rekursiven Baum. Entsprechend wollen wir auch hier vom **i -ten Abkömmling** Y^i sprechen.

Für $i \in \mathbb{N}$ und $t \geq 0$ bezeichnen wir im Yule-Prozess Y

$R_i :=$ **Zeitpunkt der i -ten Immigration.**

$P_t :=$ **Anzahl der Immigrationen bis t .**

Dann folgt direkt $R_i \sim \Gamma(1, 1)$ und $P_t \sim \text{Poi}(t)$.

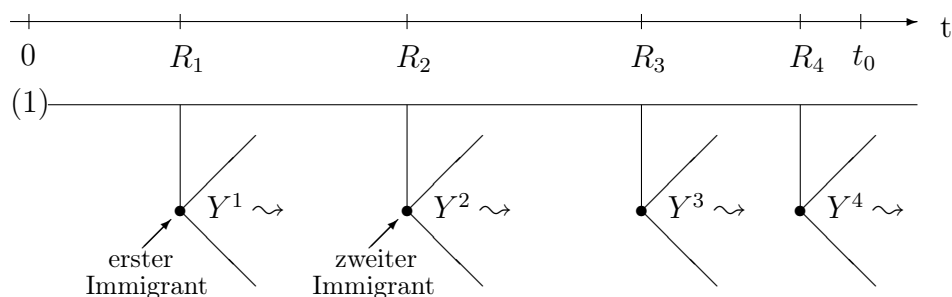


Abbildung 7: Das erweiterte Modell eines Yule-Prozesses. R_1, R_2, \dots beschreiben die Zeitpunkte zu denen die (1) immigriert. Diese Immigranten initiieren dann jeweils selbst instantan einen eigenen Yule-Prozess. Zur Zeit t_0 wurden $P_{t_0} = 4$ Immigranten geboren.

Wir bemerken noch, dass der i -te Abkömmling Y^i zum Zeitpunkt R_i gestartet wird. Wir gönnen zwar jedem der Yule-Prozess einen eigenen Zeithorizont, wollen aber trotzdem unsere Ursprungsskala als absolut ansehen. Damit beschreibt Y_t^i den Zustand des i -ten Abkömmlings zum absoluten Zeitpunkt $t + R_i$.

Weiter betrachten wir den ersten Zeitpunkt an dem genau n Individuen vorhanden sind. Für $n \in \mathbb{N}$ setzen wir

$$T_n := \inf \left\{ t \geq 0 \mid \sum_{i=1}^{P_t} Y_{t-R_i}^i = n - 1 \right\}. \quad (72)$$

In der Summe addieren wir also die Individuen aller bis zum Zeitpunkt $t \geq 0$ gestarteten ("Unter")-Yule-Prozesse. Die (1) wird hierbei nicht mitgezählt,

ist aber natürlich im Ursprungsprozess Y vorhanden. T_n ist eine Stoppzeit (bzgl. der kanonischen Filtration), denn zu jedem Zeitpunkt $t \geq 0$ wissen wir, ob bislang schon n Individuen vorhanden sind, oder noch nicht. Ein bei T_n gestoppter standard Yule-Prozess liefert also anhand dieser Konstruktion ein Analogon zum RRT^n . Der i -te Abkömmling wird beschrieben durch $Y_{T_n - R_i}^i$.

Wir haben noch ein Äquivalent zum Abschneideprozedere des RRT zu bilden und gehen dabei ähnlich zur Baumkonstruktion vor, indem wir jedes Individuum bei seiner Geburt mit einer standard exponential verteilten Lebenszeit versehen, wobei alle diese Lebenszeiten und der Prozess unabhängig seien.

Weiter definieren wir wie folgt für festes $i \in \mathbb{N}$ eine Familie von Prozessen. Dazu sei Y_s^i der i -te Abkömmling des Yule-Prozesses Y zur Zeit s . Wir betrachten zunächst eines der vorhandenen Individuen I und dessen Lebenszeit E . Sei $u \geq 0$. Wir lassen I und alle seine Nachkommen wegfallen, falls $E \leq u$ gilt. Verfahren wir entsprechend mit allen involvierten Individuen, so sprechen wir auch vom **zur Zeit s bei u ausgedünnten Prozess**, kurz $Y_s^i(u)$. Offenbar macht es keinen Unterschied ob wir einen Yule-Prozess zur Zeit s betrachten und dann bei u ausdünnen, oder ob wir einen Yule-Prozess bei s auswerten, für den direkt bei der "potentiellen" Geburt eines neuen Individuums entschieden wird, ob es tatsächlich geboren wird ($E > u$) oder nicht ($E \leq u$). Entsprechend erhalten wir eine Familie

$$((Y_s^i(u))_{s \geq 0})_{u \geq 0} \quad (73)$$

ausgedünnter Yule-Prozesse.

Schneiden wir also im i -ten Abkömmling eines RRT^n alle Kanten, deren Lebenszeit kleiner (oder gleich) einem gegebenem Wert u ist ($u \geq 0$) und lassen die anhängenden Unterbäume wegfallen, so korrespondiert dies in der Yule-Prozess Konstruktion mit $Y_{T_n - R_i}^i(u)$.

Nun noch einige Notationen: Wir betrachten die Anzahl der Immigrationen P_{T_n} bis zum Zeitpunkt T_n , zu welchem zum ersten mal n Individuen vorhanden sind. Speziell bezeichnen wir mit E_i die Lebenszeit des i -ten Immigrantanten, für $1 \leq i \leq P_{T_n}$. Lassen wir in unserer Ausdünnkonstruktion die "Intensität" u sukzessive anwachsen, so fällt zuletzt der Immigrant C_* weg, für dessen Lebenszeit

$$E_* := \max_{1 \leq i \leq P_{T_n}} E_i \quad (74)$$

gilt. Es ist dann

$$C_* = \arg \left(\max_{1 \leq i \leq P_{T_n}} E_i \right). \quad (75)$$

Für den Geburtszeitpunkt von C_* setzen wir weiter $R_* := R_{C_*}$.

Wir bemerken, dass diese Größen von n abhängen, ohne dass dieser Parameter explizit angegeben ist. Mit diesem Begriffsapparat folgt schließlich direkt

$$\left(|M_{(\neq 1)}^{[n]}|, \#B^{[n]} - 1\right) \sim \left(Y_{T_n - R_*}^{C_*}, Y_{T_n - R_*}^{C_*}(E_*)\right). \quad (76)$$

Denn $Y_{T_n - R_*}^{C_*}$ beschreibt die Größe des am längsten überlebenden Abkömmlings und $Y_{T_n - R_*}^{C_*}(E_*)$ gibt gerade an, wieviele Individuen zum Zeitpunkt der letzten Kollision in diesem Abkömmling enthalten sind. Jedes dieser Individuen steht dabei für einen Block. Das entspricht aber im wesentlichen der Struktur des letzten überlebenden Abkömmlings unserer Baumkonstruktion unmittelbar vor der letzten Verschmelzung. Die Modelle unterscheiden sich nur dahingehend, dass die "geschnittenen" Elemente im Yule-Prozess wegfallen und nicht in ihren Vater verschmelzen. Dies hat aber keine Auswirkung auf die Blockanzahl.

Die eingeführten Begriffe liefern zwar einen direkten Zusammenhang zu den zufälligen rekursiven Bäumen, sind aber wegen ihrer Struktur schwer zu handhaben. So hängen die ZVn E_* , R_* und C_* von der zufälligen Größe T_n ab. Die Idee ist nun T_n zwischen bestimmten deterministischen Zeitpunkten einzuschachteln. Es ist $T_n = \sum_{i=1}^{n-1} (T_{i+1} - T_i)$, wobei die Folge der Summanden nach Konstruktion des Yule-Prozesses unabhängig ist und $(T_{i+1} - T_i) \sim \exp(i)$ für $i \in \{1, \dots, n-1\}$ gilt, denn zwischen der Geburt des i -ten und $(i+1)$ -ten Kindes sind genau i Individuen involviert. Damit gilt für wachsendes n

$$\mathbb{E}[T_n] = \sum_{i=1}^{n-1} \mathbb{E}[T_{i+1} - T_i] = \sum_{i=1}^{n-1} \frac{1}{i} = \log n + O(1), \quad (77)$$

sowie

$$\text{Var}[T_n] = \sum_{i=1}^{n-1} \text{Var}[T_{i+1} - T_i] = \sum_{i=1}^{n-1} \frac{1}{i^2} = O(1), \quad (78)$$

d.h. T_n ist sehr stark um seinen Erwartungswert konzentriert. Deshalb setzen wir für $n \in \mathbb{N}$

$$t_n^- := \log n - (\log n)^{\frac{1}{2}} \quad \text{und} \quad t_n^+ := \log n + (\log n)^{\frac{1}{2}}, \quad (79)$$

und weiter für $P_{t_n^-} \in \mathbb{N}$

$$E_*^- := \max_{1 \leq i \leq P_{t_n^-}} E_i, \quad (80)$$

$$C_*^- := \arg \left(\max_{1 \leq i \leq P_{t_n^-}} E_i \right) \quad \text{sowie} \quad R_*^- := R_{C_*^-}. \quad (81)$$

Falls bis zum Zeitpunkt t_n^- noch keine Immigration stattgefunden hat, falls also $P_{t_n^-} = 0$ gilt, setzen wir der Vollständigkeit halber $E_*^- := C_*^- := 0$ und $R_*^- := X$ mit $X \sim \text{unif}[0, t_n^-]$ und unabhängig von E_*^- und C_*^- . Für wachsendes n wird dieses Ereignis aber sehr unwahrscheinlich.

Analog definieren wir E_*^+, C_*^+ und R_*^+ .

Wir werden später zeigen, dass auch für $P_{t_n^-} \in \mathbb{N}$ gilt: $R_*^- \sim \text{unif}[0, t_n^-]$.

Schließlich bemerken wir noch die Abhängigkeit von R_*^- und C_*^- . Andererseits sind nach Konstruktion $(Y^i)^{i \in \mathbb{N}}, E_*^-$ und das Paar (R_*^-, C_*^-) unabhängig. (Entsprechende Abhängigkeiten gelten natürlich auch für E_*^+, R_*^+, C_*^+ bzw. E_*, R_*, C_* .)

Damit können wir schließlich die Menge einführen, auf der wir uns bewegen wollen. Für $n \in \mathbb{N}$ sei

$$\mathcal{G}_n := \left\{ \underbrace{T_n \in (t_n^-, t_n^+)}_{(a)}, \underbrace{C_* = C_*^-}_{(b)} \text{ und } \underbrace{Y_{t_n^- - R_*^-}^{C_*^-}(E_*^-) = Y_{t_n^+ - R_*^-}^{C_*^-}(E_*^-)}_{(c)} \right\}. \quad (82)$$

Wir schließen noch einige Bemerkungen zu \mathcal{G}_n an. Dessen genaue Bedeutung wird sich allerdings erst im Beweis entfalten.

(a): Mit $T_n \in (t_n^-, t_n^+)$ legen wir also, wie oben angesprochen, den Geburtszeitpunkt des n -ten Individuums zwischen die zwei deterministischen Werte t_n^- und t_n^+ . Darauf werden wir später noch genauer eingehen.

(b): Es ist leicht einzusehen, dass $C_* = C_*^- \Leftrightarrow E_* = E_*^- \Leftrightarrow R_* = R_*^-$ gilt. Wegen (a) ist außerdem $T_n > t_n^-$. Betrachten wir alle bis zum Zeitpunkt T_n geborenen Immigranten, so wurde diejenige $\exp(1)$ verteilte ZV, die letztlich den größten Wert annimmt, an einem Immigranten angebracht, der schon bis zum Zeitpunkt t_n^- geboren wurde. M.a.W.: Der Abkömmling, der bis zum Zeitpunkt t_n^- für die finale Kollision prädestiniert war, wird auch unter allen bis zum Zeitpunkt T_n entstandenen Abkömmlingen zuletzt sterben. (vgl. Abbildung 8)

(c): Dies beschreibt den Abkömmling, der am längsten überlebt, ausgewertet zu zwei verschiedenen Zeitpunkten $t_n^- - R_*^-$ und $t_n^+ - R_*^-$. Zudem wird bei E_*^- ausgedünnt. (Wegen (b) könnte man auf die Minuszeichen von R_*^- und E_*^- verzichten). Die Gleichheit bedeutet, dass jedes zwischen den beiden Zeitpunkten geborene Individuum entweder mit einer Lebenszeit versehen wurde, die einen Wert kleiner (oder gleich) E_*^- annimmt, oder aber einen Vater hat, der selbst schon beim Ausdünnen wegfällt. (Der Wert der eigenen Lebenszeit ist dabei nicht relevant).

Nun sollte der für den Beweis des Hauptsatzes benötigte Begriffsapparat vorhanden sein.

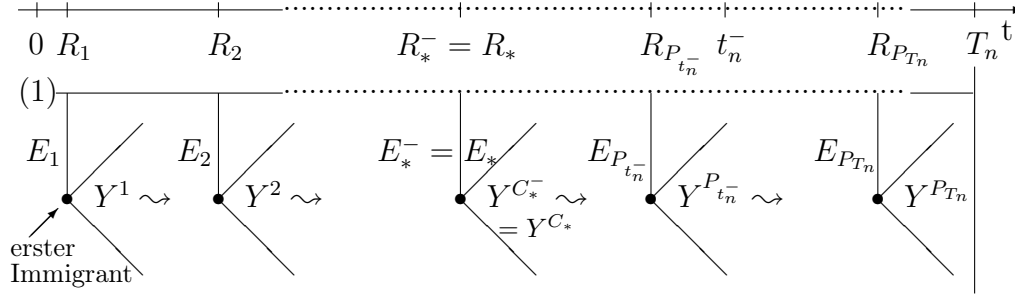


Abbildung 8: Ein auf \mathcal{G}_n verlaufender Yule-Prozess. Insbesondere ist $t_n^- < T_n$ und $E_*^- = E_*$. D.h. die größte aller Lebenszeiten bis zum Zeitpunkt T_n gehört zu einem Immigrant, der schon bis zum Zeitpunkt t_n^- geboren wurde.

4.1.2 Hilfslemmata zum Hauptbeweis

Wir geben hier einige Lemmata an, deren wir uns im Hauptbeweis bedienen. Für einen besseren Überblick werden diese erst im Anschluß an den Hauptbeweis in Abschnitt 4.1.4 bewiesen.

Sämtliche Notationen werden aus dem obigen Abschnitt 4.1.1 übernommen. Wie schon angesprochen zeigen wir zunächst

Lemma 4.1.2. *Für alle $n \in \mathbb{N}$ gilt*

$$R_*^- \sim \text{unif}[0, t_n^-] \quad \text{und} \quad R_*^+ \sim \text{unif}[0, t_n^+]. \quad (83)$$

In der eingeführten Darstellung von $|M_{(\neq 1)}^{[n]}|$ und $\#B^{[n]}$ durch die Yule-Prozesse taucht jeweils der zufällige Zeitpunkt R_*^- auf. Andererseits ist im Hauptsatz die gemeinsame Verteilungskonvergenz im wesentlichen durch eine - in beiden Komponenten auftretende - uniform verteilte ZV bestimmt. Das eben bewiesene Lemma lässt schon erahnen, dass R_*^- der Schlüssel für den Beweis der gemeinsamen Konvergenz sein wird.

Wir wollen den ausgedünnten Prozess $(Y_t^{C_*^-}(E_*^-))_{t \geq 0}$ unserer Konstruktion umschreiben. Dann können wir auf den Akt des Ausdünnens verzichten, wenn wir einen zeitlich umskalierten Yule-Prozess betrachten. In mehreren Schritten zeigen wir

Lemma 4.1.3. *Es gilt*

$$\left(Y_t^{C_*^-}(E_*^-)\right)_{t \geq 0} \sim \left(Y_{e^{-E_*^-} t}\right)_{t \geq 0}, \quad (84)$$

mit $Y \sim \text{Yule}(1)$ und Y, E_*^- unabhängig. Insbesondere gilt

$$\left(Y_t^i(E_*^-)\right)_{t \geq 0} \sim \left(Y_{e^{-E_*^-} t}\right)_{t \geq 0}, \quad (85)$$

für alle $i \in \mathbb{N}$ mit Y, E_*^- wie in (84).

Lemma 4.1.4. *Für $n \rightarrow \infty$ gelten folgende Konvergenzen*

$$(i) \quad \frac{\log Y_{t_n^- - R_*^-}^{C_*^-}}{t_n^- - R_*^-} \xrightarrow{P} 1, \quad (ii) \quad \frac{\log Y_{t_n^+ - R_*^-}^{C_*^-}}{t_n^+ - R_*^-} \xrightarrow{P} 1, \quad (86)$$

$$(iii) \quad \frac{t_n^- - R_*^-}{t_n^+ - R_*^-} \xrightarrow{P} 1. \quad (87)$$

Lemma 4.1.5. *Für $n \in \mathbb{N}$ gilt*

$$t_n^- e^{-E_*^-} \sim E \wedge t_n^-, \quad (88)$$

mit $E \sim \exp(1)$.

Wir werden die exponential verteilte ZV, die im Hauptsatz auftaucht, durch (88) aufgreifen und erinnern an Bemerkung (iii) des Hauptsatzes. Wird E_*^- groß, so sollten auch im letzten Abkömmling des Baumes schon viele Schnitte stattgefunden haben, d.h. die Blockanzahl in der letzten Kollision $\#B^{[n]}$ sollte eher klein sein. Dies stellen wir nun auch mit (88) fest. Für großes E_*^- ist E klein. Für die zweite Komponente des Hauptsatzes bedeutet das dann, dass Y_{UE} (bei festem U) ein kleines Argument hat und dementsprechend einen eher kleinen Wert annehmen sollte.

Die letzte nötige Aussage bezieht sich auf die Menge \mathcal{G}_n . Wir erinnern

$$\mathcal{G}_n := \{ T_n \in (t_n^-, t_n^+), C_* = C_*^- \text{ und } Y_{t_n^- - R_*^-}^{C_*^-}(E_*^-) = Y_{t_n^+ - R_*^-}^{C_*^-}(E_*^-) \}.$$

Lemma 4.1.6. *Für $n \rightarrow \infty$ gilt*

$$P(\mathcal{G}_n) \longrightarrow 1. \quad (89)$$

4.1.3 Der Hauptbeweis

Mit den obigen Lemmata können wir nun den Hauptbeweis führen. Wie schon angedeutet, liegt die Idee darin, zunächst nur auf der Menge \mathcal{G}_n zu arbeiten, denn dort sind die betrachteten Objekte $|M_{(\neq 1)}^{[n]}|$ und $\#B^{[n]}$ von der kompliziert strukturierten Stoppzeit T_n entkoppelt. Mit dem Satz von Slutsky werden wir dann die behauptete Konvergenz folgern. Der allgemeine Fall folgt dann wegen $P(\mathcal{G}_n) \rightarrow 1$, für $n \rightarrow \infty$ direkt.

Beweis: Nach Konstruktion gilt

$$|M_{(\neq 1)}^{[n]}| = Y_{T_n - R_*}^{C_*} \quad \text{und} \quad \#B^{[n]} - 1 = Y_{T_n - R_*}^{C_*}(E_*).$$

Sofern nicht anders dargestellt, wollen wir uns fortan auf der Menge \mathcal{G}_n bewegen. Gegeben dieser Information gilt dann für alle $n \in \mathbb{N}$

$$Y_{t_n^- - R_*^-}^{C_*^-} \leq |M_{(\neq 1)}^{[n]}| \leq Y_{t_n^+ - R_*^-}^{C_*^-} \quad (90)$$

also für $n \in \tilde{\mathbb{N}}$

$$\frac{\log Y_{t_n^- - R_*^-}^{C_*^-}}{\log n} \leq \frac{\log |M_{(\neq 1)}^{[n]}|}{\log n} \leq \frac{Y_{t_n^+ - R_*^-}^{C_*^-}}{\log n} \quad (91)$$

und auch

$$\#B^{[n]} - 1 = Y_{t_n^- - R_*^-}^{C_*^-}(E_*^-). \quad (92)$$

Wir erkennen, dass die behauptete Aussage

$$\left(\frac{\log |M_{(\neq 1)}^{[n]}|}{\log n}, \#B^{[n]} \right) \xrightarrow{d} (U, 1 + Y_{UE})$$

schon folgt, wenn die folgende Konvergenz gilt

$$\left(\frac{\log Y_{t_n^- - R_*^-}^{C_*^-}}{\log n}, \frac{\log Y_{t_n^+ - R_*^-}^{C_*^-}}{\log n}, Y_{t_n^- - R_*^-}^{C_*^-}(E_*^-) \right) \xrightarrow{d} (U, U, Y_{UE}), \quad (93)$$

für $n \rightarrow \infty$. (Dabei sind Y, E und U wie im Hauptsatz gewählt).

Um den Satz von Slutsky anzuwenden, zeigen wir nun für das Tripel

$$\left(\frac{\log Y_{t_n^- - R_*^-}^{C_*^-}}{\log n}, \frac{\log Y_{t_n^+ - R_*^-}^{C_*^-}}{\log n}, Y_{t_n^- - R_*^-}^{C_*^-}(E_*^-) \right) = (A_n U, B_n U, Y_{t_n^- U}^{C_*^-}(E_*^-)), \quad (94)$$

mit $U \sim \text{unif}[0, 1]$, $A_n \xrightarrow{P} 1$ und $B_n \xrightarrow{P} 1$, für $n \rightarrow \infty$.

Hier taucht nun zum ersten Mal die ZV U auf, die der Komponente R_*^- entspringt. Nach Lemma 4.1.2 gilt $R_*^- \sim \text{unif}[0, t_n^-]$, also auch $t_n^- - R_*^- \sim \text{unif}[0, t_n^-]$ und wir setzen

$$U := \frac{t_n^- - R_*^-}{t_n^-}, \quad (95)$$

für $n \in \tilde{\mathbb{N}}$. Dann ist $U \sim \text{unif}[0, 1]$.

Wir wollen jede Komponente des Tripels in eine Darstellung bringen, in der der Quotient $\frac{t_n^- - R_*^-}{t_n^-}$ auftaucht, der dann durch U ersetzt werden kann.

1. Komponente:

$$\frac{\log Y_{t_n^- - R_*^-}^{C_*^-}}{\log n} = \frac{t_n^-}{\log n} \underbrace{\frac{t_n^- - R_*^-}{t_n^-}}_{=U} \frac{\log Y_{t_n^- - R_*^-}^{C_*^-}}{t_n^- - R_*^-} = A_n U,$$

mit

$$A_n := \frac{t_n^-}{\log n} \frac{\log Y_{t_n^- - R_*^-}^{C_*^-}}{t_n^- - R_*^-},$$

für alle $n \in \tilde{\mathbb{N}}$. Nach Lemma 4.1.4 folgt

$$\frac{t_n^-}{\log n} \longrightarrow 1, \quad \text{und} \quad \frac{\log Y_{t_n^- - R_*^-}^{C_*^-}}{t_n^- - R_*^-} \xrightarrow{P} 1,$$

für $n \rightarrow \infty$, also gilt trivialerweise auch $A_n \xrightarrow{P} 1$.

2. Komponente: Wir bemerken zunächst, dass für $n \rightarrow \infty$ gilt

$$\frac{t_n^+ - R_*^-}{t_n^+} = \frac{t_n^- - R_*^-}{t_n^-} (1 + o(1)),$$

was durch einfaches Nachrechnen folgt und letztlich an der asymptotischen Gleichheit von t_n^- und t_n^+ liegt. Damit schließen wir analog

$$\frac{\log Y_{t_n^+ - R_*^-}^{C_*^-}}{\log n} = \frac{t_n^+}{\log n} \underbrace{\frac{t_n^+ - R_*^-}{t_n^+}}_{=(1+o(1))U} \frac{\log Y_{t_n^+ - R_*^-}^{C_*^-}}{t_n^+ - R_*^-} = B_n U,$$

mit

$$B_n := \frac{t_n^+}{\log n} \frac{\log Y_{t_n^+ - R_*^-}^{C_*^-}}{t_n^+ - R_*^-} (1 + o(1)),$$

womit wieder nach Lemma 4.1.4 die Konvergenz $B_n \xrightarrow{P} 1$ folgt.

3. Komponente: Die Behauptung

$$Y_{t_n^- - R_*^-}^{C_*^-}(E_*^-) = Y_{t_n^-}^{C_*^-}(E_*^-)$$

folgt direkt. Also gilt (94).

Mit den obigen Notationen bleibt folgende Verteilungskonvergenz zu zeigen

$$\left(A_n U, B_n U, Y_{t_n^-}^{C_*^-}(E_*^-) \right) \xrightarrow{d} (U, U, Y_{UE}). \quad (96)$$

Nach dem Satz von Slutsky folgt dies bereits, wenn für $n \rightarrow \infty$ gilt

$$(a) \quad d_2 \left(\left(A_n U, B_n U, Y_{t_n^- U}^{C_*^-}(E_*^-) \right), \left(U, U, Y_{t_n^- U}^{C_*^-}(E_*^-) \right) \right) \xrightarrow{P} 0. \quad (97)$$

$$(b) \quad \left(U, U, Y_{t_n^- U}^{C_*^-}(E_*^-) \right) \xrightarrow{d} (U, U, Y_{UE}). \quad (98)$$

Hierbei bezeichnet $d_2 : \mathbb{R}^3 \rightarrow \mathbb{R}$ die gewöhnliche euklidische Metrik im \mathbb{R}^3 .

(a) folgt nach einer einfachen Abschätzung. A_n und B_n konvergieren stochastisch gegen 1. Sei dann $\varepsilon > 0$. Da $U \sim \text{unif}[0, 1]$ gilt

$$\begin{aligned} & P \left(d_2 \left[\left(A_n U, B_n U, Y_{t_n^- U}^{C_*^-}(E_*^-) \right), \left(U, U, Y_{t_n^- U}^{C_*^-}(E_*^-) \right) \right] \geq \varepsilon \right) \\ &= P \left(U \left((A_n - 1)^2 + (B_n - 1)^2 \right)^{\frac{1}{2}} \geq \varepsilon \right) \\ &\leq P \left(\left((A_n - 1)^2 + (B_n - 1)^2 \right)^{\frac{1}{2}} \geq \varepsilon \right) \\ &= P \left((A_n - 1)^2 + (B_n - 1)^2 \geq \varepsilon^2 \right) \\ &\leq P \left(\left\{ (A_n - 1)^2 \geq \frac{\varepsilon^2}{4} \right\} \cup \left\{ (B_n - 1)^2 \geq \frac{\varepsilon^2}{4} \right\} \right) \\ &\leq P \left(|A_n - 1| \geq \frac{\varepsilon}{2} \right) + P \left(|B_n - 1| \geq \frac{\varepsilon}{2} \right) \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Da $\varepsilon > 0$ beliebig gewählt war folgt (a).

Zu (b): Nach Lemma 4.1.5 gilt $t_n^- e^{-E_*^-} \sim E \wedge t_n^-$. Für einen standard Yule-Prozess Y mit $Y, (E, E_*^-)$ unabhängig folgt dann

$$Y_{t_n^- e^{-E_*^-}} \sim Y_{E \wedge t_n^-} \quad (99)$$

direkt durch Bedingen auf einen festen Zeitpunkt. Wir erinnern außerdem an Lemma 4.1.3. Für den i -ten Abkömmling ($i \in \mathbb{N}$) gilt

$$(Y_t^i(E_*^-))_{t \geq 0} \sim (Y_{e^{-E_*^-} t})_{t \geq 0},$$

mit $Y \sim \text{Yule}(1)$ und Y, E_*^- unabhängig.

Nach Konstruktion waren C_*^- und R_*^- abhängig. Also auch C_*^- und U . Allerdings bleibt $(Y^i)^{i \in \mathbb{N}}, E_*^-, (C_*^-, U)$ unabhängig.

Sind auch $Y, (E_*^-, E), (C_*^-, U)$ unabhängig, so folgt für $k \in \mathbb{N}$, $A, B \in \mathcal{B}([0, 1])$, und $D := A \cap B$.

$$\begin{aligned} & P \left(U \in A, U \in B, Y_{t_n^- U}^{C_*^-}(E_*^-) = k \right) \\ &= \int_{\mathbb{R}^2} P \left(U \in D, Y_{t_n^- U}^{C_*^-}(E_*^-) = k \mid (C_*^-, U) = (i, u) \right) dP_{(C_*^-, U)}((i, u)) \end{aligned}$$

$$\begin{aligned}
&= \int_{\mathbb{R}^2} \mathbb{1}_D(u) P\left(Y_{t_n^- u}^i(E_*^-) = k\right) dP_{(C_*^-, U)}((i, u)) \\
&= \int_{\mathbb{R}^2} \mathbb{1}_D(u) P\left(Y_{e^{-E_*^-} t_n^- u} = k\right) dP_{(C_*^-, U)}((i, u)) \\
&\stackrel{(99)}{=} \int_{\mathbb{R}^2} \mathbb{1}_D(u) P\left(Y_{(E \wedge t_n^-)u} = k\right) dP_{(C_*^-, U)}((i, u)) \\
&= \int_{\mathbb{R}^2} P\left(U \in D, Y_{(E \wedge t_n^-)U} = k \mid (C_*^-, U) = (i, u)\right) dP_{(C_*^-, U)}((i, u)) \\
&= P\left(U \in A, U \in B, Y_{(E \wedge t_n^-)U} = k\right).
\end{aligned}$$

Das Minimum lässt sich wieder leicht wegdiskutieren

$$\begin{aligned}
&P\left(U \in A, U \in B, Y_{(E \wedge t_n^-)U} = k\right) \\
&= P\left(U \in A, U \in B, Y_{EU} = k, E \leq t_n^-\right) \\
&\quad + P\left(U \in A, U \in B, Y_{t_n^- U} = k, E > t_n^-\right) \\
&\xrightarrow{n \rightarrow \infty} P\left(U \in A, U \in B, Y_{EU} = k\right),
\end{aligned}$$

da $t_n^- \rightarrow \infty$ *f.s.* für $n \rightarrow \infty$. Damit gilt auch (b).

Gegeben der Information \mathcal{G}_n folgt somit die Konvergenz

$$\left(\frac{\log |M_{(\neq 1)}^{[n]}|}{\log n}, \#B^{[n]}\right) \xrightarrow{d} (U, 1 + Y_{UE})$$

und wir bemerken, dass Y, U, E im Beweis unabhängig gewählt waren. Damit folgt aber auch direkt der allgemeine Fall, wegen

$$P(\mathcal{G}_n) \xrightarrow{n \rightarrow \infty} 1.$$

Also ist der Hauptsatz vollständig bewiesen. □

4.1.4 Beweise der Hilfslemmata

Beweis von Lemma 4.1.2: Zunächst die Verteilung von R_*^- . Gegeben $\{P_{t_n^-} = 0\}$, gilt die Behauptung nach Konvention.

Sei deshalb im Folgenden $P_{t_n^-} \in \mathbb{N}$. Wir interpretieren die Geburtszeitpunkte des ersten Individuums (1) wieder als die Erneuerungspunkte $R_1, \dots, R_{P_{t_n^-}}$ eines homogenen standard Poisson-Prozesses und erinnern an folgendes bekanntes Resultat:

- Gegeben $\{P_{t_n^-} = m\}$ ist der Zufallsvektor (R_1, \dots, R_m) verteilt die Ordnungsstatistiken $(U_{(1)}, \dots, U_{(m)})$ von m unabhängigen $\text{unif}[0, t_n^-]$ verteilten ZVn U_1, \dots, U_m .

D.h. gegeben $\{P_{t_n^-} = m\}$ ist ein uniform - vom Immigrationsprozess unabhängig - gewählter Erneuerungspunkt $\text{unif}[0, t_n^-]$ verteilt. Dieses uniforme Wählen wird nach Konstruktion durch die Lebenszeiten beschrieben. Da außerdem m beliebig gewählt war, folgt für $k \in [0, t_n^-]$

$$P(R_*^- \leq k) = \sum_{m=0}^{\infty} P(R_*^- \leq k | P_{t_n^-} = m) P(P_{t_n^-} = m) = \frac{k}{t_n^-} \sum_{m=0}^{\infty} P(P_{t_n^-} = m) = \frac{k}{t_n^-}.$$

Analog lässt sich $R_*^+ \sim \text{unif}[0, t_n^+]$ beweisen. \square

Beweis von Lemma 4.1.3: Seien $(Y^i)^{i \in \mathbb{N}}$ die Abkömmlinge unserer Konstruktion. Zunächst zeigen wir für deterministische Werte $i \in \mathbb{N}$ und $u \in [0, \infty)$, dass gilt

$$(Y_t^i(u))_{t \geq 0} \sim (Y_{e^{-u}t})_{t \geq 0}, \quad (100)$$

mit $Y \sim \text{Yule}(1)$.

D.h. ein bei u ausgedünnter standard Yule-Prozess ist verteilt wie ein (nicht ausgedünnter) standard Yule-Prozess, der anstatt bei t zur früheren Zeit $e^{-u}t$ ausgewertet wird.

Nach Konstruktion von Y^i beschreiben die Geburten jedes Individuums einen homogenen standard Poisson-Prozesses (P.P.). Dessen Intensität beträgt $\mu = 1$, d.h. das entsprechende Intensitätsmaß ist das konventionelle Lebesguemaß λ . Nach unserer Konstruktion ist wieder jedes Individuum mit einer $\exp(1)$ -verteilten Lebenszeit versehen (alle diese ZVn sind unabhängig). Da wir ausschließlich den ausgedünnten Prozess betrachten und sich alle Individuen voneinander unabhängig entwickeln, erinnern wir an folgendes Vorgehen: Anstatt den Prozess erst aufzubauen und im nachhinein auszudünnen, entscheiden wir direkt an jedem Erneuerungspunkt, ob ein Individuum tatsächlich gebiert (falls die angefügte ZV E einen Wert größer u annimmt) oder aber doch kein Nachkomme geboren wird ($E \leq u$). Wir setzen

$$p := P(E > u) = e^{-u}.$$

Interpretieren wir die Entwicklung eines Individuums als markierten P.P., so ergibt sich dessen Intensitätsmaß nach gängigen Resultaten als $p\lambda$, d.h. auch dieser P.P. ist homogen mit Intensität $\tilde{\mu} = e^{-u}$. Für die Gesamtheit der Individuen bedeutet dies

$$(Y_t^i(u))_{t \geq 0} \sim \text{Yule}(e^{-u}). \quad (101)$$

Wir können nun zeigen, dass die endlich dimensionalen Verteilungen von $(Y_t^i(u))_{t \geq 0}$ und $(Y_{e^{-ut}})_{t \geq 0}$ übereinstimmen, womit die Zwischenbehauptung (100) bewiesen wäre. Für $n \in \mathbb{N}$ seien dazu $t_1, \dots, t_n \in \mathbb{R}_0^+$ mit $t_1 < \dots < t_n$ sowie $k_1, \dots, k_n \in \mathbb{N}$ mit $k_1 < \dots < k_n$. Da Yule-Prozesse insbesondere Markov-Prozesse sind folgt unter Ausnutzung von Lemma 2.2.5

$$\begin{aligned}
& P(Y_{t_n}^i(u) = k_n, \dots, Y_{t_1}^i(u) = k_1) \\
&= P(Y_{t_n}^i(u) = k_n \mid Y_{t_{n-1}}^i(u) = k_{n-1}) \cdot \dots \cdot P(Y_{t_1}^i = k_1) \\
(101) \quad & \stackrel{=}{=} \binom{k_n - 1}{k_{n-1} - 1} e^{-e^{-u}(t_n - t_{n-1})} (1 - e^{-e^{-u}(t_n - t_{n-1})})^{k_n - k_{n-1}} \\
& \quad \dots \cdot e^{-e^{-u}t_1} (1 - e^{-e^{-u}t_1})^{k_1 - 1} \\
& \stackrel{Y \sim \text{Yule}(1)}{=} P(Y_{t_n e^{-u}} = k_n \mid Y_{t_{n-1} e^{-u}} = k_{n-1}) \cdot \dots \cdot P(Y_{t_1 e^{-u}} = k_1) \\
&= P(Y_{t_n e^{-u}} = k_n, \dots, Y_{t_1 e^{-u}} = k_1).
\end{aligned}$$

Also gilt (100). Damit können wir nun zeigen

$$\left(Y_t^{C_*^-} (E_*^-) \right)_{t \geq 0} \sim \left(Y_{e^{-E_*^-} t} \right)_{t \geq 0},$$

mit $Y \sim \text{Yule}(1)$ und Y, E_*^- unabhängig. Offensichtlich sind die Prozesse càdlàg, also geben wir uns ein $B \in \mathcal{B}(D_{\mathbb{R}}[0, \infty))$ vor.

Falls $P_{t_n^-} = 0$ ist, gilt nach Konvention $C_*^- = 0$. In diesem Fall setzen wir auch

$$Y^{C_*^-}(u) := Y^1(u).$$

Wegen der Unabhängigkeiten folgt dann

$$\begin{aligned}
P(Y^{C_*^-}(E_*^-) \in B) &= \int_{\mathbb{R}} \sum_{i=0}^{P_{t_n^-}} P(Y^{C_*^-}(E_*^-) \in B \mid E_*^- = u, C_*^- = i) P(C_*^- = i) dP_{E_*^-}(u) \\
Y^i, C_*^-, E_*^- \text{ u.a.} &= \int_{\mathbb{R}} \sum_{i=0}^{P_{t_n^-}} P(Y^i(u) \in B) P(C_*^- = i) dP_{E_*^-}(u) \\
(Y^i)^{i \in \mathbb{N}} \text{ i.i.d.} &= \int_{\mathbb{R}} P(Y^1(u) \in B) dP_{E_*^-}(u) \\
& \stackrel{(100)}{=} \int_{\mathbb{R}} P((Y_{e^{-ut}})_{t \geq 0} \in B) dP_{E_*^-}(u) \\
Y, E_*^- \text{ u.a.} &= \int_{\mathbb{R}} P\left((Y_{e^{-E_*^-} t})_{t \geq 0} \in B \mid E_*^- = u \right) dP_{E_*^-}(u) \\
&= P\left((Y_{e^{-E_*^-} t})_{t \geq 0} \in B \right).
\end{aligned}$$

Damit gilt (84). Analog folgt auch (85). \square

Beweis von Lemma 4.1.4: Wir beginnen mit dem Beweis von (i) und (ii). Dabei zeigen wir zunächst, dass für einen standard Yule-Prozess $Y = (Y_t)_{t \geq 0}$ für $t \rightarrow \infty$ gilt

$$\frac{\log Y_t}{t} \longrightarrow 1 \quad f.s. \quad (102)$$

Wegen Lemma 2.2.9 gilt

$$\frac{Y_t}{e^t} \longrightarrow E \quad f.s. \quad (103)$$

mit $E \sim \exp(1)$. Damit folgt

$$\log Y_t - t = \log \left(\frac{Y_t}{e^t} \right) \longrightarrow \log(E) \quad f.s.$$

also

$$\frac{Y_t - t}{t} \longrightarrow 0 \quad f.s.$$

Dies liefert die Zwischenbehauptung.

Im weiteren betrachten wir eine Folge $(V_n)_{n \in \mathbb{N}}$ \mathbb{R}^+ -wertiger ZVn. Diese heie (stochastisch) konvergent gegen unendlich, falls für alle $K \in \mathbb{R}$ gilt

$$P(V_n > K) \longrightarrow 0, \quad \text{für } n \rightarrow \infty. \quad (104)$$

Wir wollen nun für eine solche Folge $(V_n)_{n \in \mathbb{N}}$ zeigen

$$\frac{\log Y_{V_n}}{V_n} \xrightarrow{P} 1. \quad (105)$$

Wir setzen $X(t) := \frac{\log Y_t}{t}$.

Da $V_n \xrightarrow{P} \infty$, existiert zu jeder Teilfolge $(V_{n_k})_{k \in \mathbb{N}}$ von $(V_n)_{n \in \mathbb{N}}$ eine Teilfolge $(V_{n_{k_\ell}})_{\ell \in \mathbb{N}}$ mit

$$V_{n_{k_\ell}} \longrightarrow \infty \quad f.s. \quad \text{für } \ell \rightarrow \infty.$$

Da auch $X(t) \xrightarrow{t \rightarrow \infty} 1$ *f.s.* folgt

$$X(V_{n_{k_\ell}}) \longrightarrow 1 \quad f.s. \quad \text{für } \ell \rightarrow \infty.$$

Damit folgt aber (105) nach dem bekannten Teilfolgenprinzip.

Setzen wir $\tilde{V}_n := t_n^- - R_*^-$ für alle $n \in \mathbb{N}$, so haben wir noch zu zeigen, dass $(\tilde{V}_n)_{n \in \mathbb{N}}$ stochastisch gegen unendlich konvergiert. In Lemma 4.1.2 hatten wir $R_*^- \sim \text{unif}[0, t_n^-]$ gezeigt. Sei nun $z \in \mathbb{R}_0^+$ gegeben, dann folgt

$$\begin{aligned} P(\tilde{V}_n \geq z) &= P(R_*^- \leq t_n^- - z) = \int_0^{t_n^- - z} \mathbb{1}_{[0, t_n^-]}(x) \frac{1}{t_n^-} d\lambda(x) \\ &= \begin{cases} \frac{t_n^- - z}{t_n^-} & \text{für } t_n^- > z, \\ 0 & \text{sonst.} \end{cases} \quad n \rightarrow \infty \rightarrow 1, \end{aligned}$$

weil $t_n^- \xrightarrow{n \rightarrow \infty} \infty$. Damit ist (i) bewiesen und völlig analog folgt auch (ii), wenn wir im letzten Schritt t_n^- durch t_n^+ ersetzen.

Behauptung (iii) ist unmittelbar einsichtig. Es gilt

$$\frac{t_n^-}{t_n^+} \xrightarrow{n \rightarrow \infty} 1.$$

Sei nun $\varepsilon > 0$ gegeben, dann folgt direkt

$$\begin{aligned} P\left(\frac{t_n^- - R_*^-}{t_n^+ - R_*^-} \geq 1 - \varepsilon\right) &= P(t_n^- \geq t_n^+ - \varepsilon t_n^+ + \varepsilon R_*^-) \\ &= P(\underbrace{\varepsilon R_*^- \leq \left(\frac{t_n^-}{t_n^+} - 1 + \varepsilon\right) t_n^+}_{n \rightarrow \infty \rightarrow \varepsilon}) \xrightarrow{n \rightarrow \infty} 1, \end{aligned}$$

da $R_*^- \sim \text{unif}[0, t_n^-]$. Weil $\varepsilon > 0$ beliebig gewählt war und außerdem

$$\frac{t_n^- - R_*^-}{t_n^+ - R_*^-} \leq 1 \quad f.s.$$

gilt, folgt die Behauptung. \square

Beweis von Lemma 4.1.5: Für $n \in \mathbb{N}$ interpretieren wir wieder die Geburten der (1) im Intervall $[0, t_n^-]$ als die Erneuerungspunkte eines homogenen P.P. mit Intensität $\mu = 1$. Deren Anzahl $P_{t_n^-}$ ist dann $\text{Poi}(t_n^-)$ -verteilt. Nach Konstruktion ist jeder Immigrant mit einer $\exp(1)$ -verteilten Lebenszeit E versehen. Sei $u > 0$. Markieren wir alle Immigranten deren Lebenszeit $E \geq u$ ist, so bilden die markierten Immigranten einen homogenen P.P. mit Intensität $\tilde{\mu} = e^{-u}$. Beschreibe X die Anzahl der Erneuerungen des markierten Prozesses in $[0, t_n^-]$, dann ist $X \sim \text{Poi}(t_n^- e^{-u})$. Da nach Konvention gilt $E_*^- = 0$, falls $P_{t_n^-} = 0$ folgt

$$P(E_*^- < u) = P(X = 0) = e^{-t_n^- e^{-u}} \frac{(-t_n^- e^{-u})^0}{0!} = \exp(-t_n^- e^{-u}). \quad (106)$$

Und damit gilt für $0 < x < t_n^-$

$$\begin{aligned} P(t_n^- e^{-E_*^-} > x) &= P\left(E_*^- < -\log\left(\frac{x}{t_n^-}\right)\right) \\ &\stackrel{(106)}{=} \exp\left(-t_n^- e^{\log\left(\frac{x}{t_n^-}\right)}\right) = e^{-x}. \end{aligned}$$

(Wir müssen $x < t_n^-$ wählen, weil die Gleichung (106) nur für $u > 0$ gilt.) Andererseits ist

$$0 < t_n^- e^{-E_*^-} \leq t_n^- \quad f.s.,$$

was die Behauptung liefert. \square

Beweis von Lemma 4.1.6: Der Beweis erfolgt in drei Schritten. Wir zeigen

- (a) $P(T_n \in (t_n^-, t_n^+)) \longrightarrow 1.$
- (b) $P(T_n \in (t_n^-, t_n^+), C_* = C_*^-) \longrightarrow 1.$
- (c) $P(Y_{t_n^+ - R_*^-}^{C_*^-}(E_*^-) = Y_{t_n^+ - R_*^-}^{C_*^-}(E_*^-)) \longrightarrow 1.$

Aus (b) und (c) folgt dann direkt die Behauptung.

Zu (a): Wir hatten das Ereignis $\{T_n \in (t_n^-, t_n^+)\}$ schon bei der Einführung der Menge \mathcal{G}_n betrachtet. Dazu hatten wir $T_n = \sum_{i=1}^{n-1} (T_{i+1} - T_i)$ als Summe von unabhängigen exponential verteilten ZV angesehen und somit

$$\mathbb{E}[T_n] = \log n + O(1) \quad \text{und} \quad \text{Var}[T_n] = O(1) \quad (107)$$

gezeigt (vgl. Seite 45). D.h. es existiert ein $C > 0$, so dass für alle $n \in \mathbb{N}$ gilt

$$|\mathbb{E}[T_n] - \log n| \leq C. \quad (108)$$

Da gilt

$$|T_n - \log n| \leq |T_n - \mathbb{E}[T_n]| + |\mathbb{E}[T_n] - \log n| \leq |T_n - \mathbb{E}[T_n]| + C, \quad (109)$$

folgt mit der Chebyshev-Ungleichung wegen der beschränkten Varianz

$$\begin{aligned} P(T_n \notin (t_n^-, t_n^+)) &= P\left(|T_n - \log n| > (\log n)^{\frac{1}{2}}\right) \\ &\leq P\left(|T_n - \mathbb{E}[T_n]| > (\log n)^{\frac{1}{2}} - C\right) \\ &\leq \frac{\text{Var}[T_n]}{\left((\log n)^{\frac{1}{2}} - C\right)^2} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Also gilt (a).

Zu (b): Wir betrachten das Ereignis

$$\{R_*^+ \in [t_n^-, t_n^+]\}, \quad (110)$$

dass aus allen Kindern der Wurzel, die bis zum Zeitpunkt t_n^+ geboren werden dasjenige mit der größten Lebenszeit im Intervall $[t_n^-, t_n^+]$ geboren wird.

Dann gilt für alle $n \in \mathbb{N}$

$$\{C_* \neq C_*^-, T_n \in (t_n^-, t_n^+)\} \subset \{R_*^+ \in [t_n^-, t_n^+]\}, \quad (111)$$

denn wegen $C_* \neq C_*^-$ wird aus allen bis zur Zeit T_n geborenen Kindern dasjenige mit der größten Lebenszeit im Intervall (t_n^-, ∞) geboren. Andererseits ist die Geburtszeit R_* von C_* wegen $T_n \in (t_n^-, t_n^+)$ durch t_n^+ beschränkt, d.h. es gilt $R_* \in [t_n^-, t_n^+]$ und somit hier auch $R_*^+ \in [t_n^-, t_n^+]$.

(Die umgekehrte Inklusion gilt natürlich nicht, weil $\{R_* \in [t_n^-, t_n^+]\}$ keine Bedingung an T_n stellt.)

Da nach Lemma 4.1.2 gilt $R_*^+ \sim \text{unif}[0, t_n^+]$ folgt

$$P(C_* \neq C_*^-, T_n \in (t_n^-, t_n^+)) \leq P(R_*^+ \in [t_n^-, t_n^+]) = \frac{t_n^+ - t_n^-}{t_n^+} \xrightarrow{n \rightarrow \infty} 0.$$

Wegen (a) muss dann auch

$$P(C_* = C_*^-, T_n \in (t_n^-, t_n^+)) \xrightarrow{n \rightarrow \infty} 1$$

gelten.

Bleibt noch (c) zu zeigen. Dazu zunächst

$$P\left(Y_{(t_n^- - R_*^-)}^{C_*^-}(E_*^-) = Y_{(t_n^+ - R_*^-)}^{C_*^-}(E_*^-)\right) = P\left(Y_{e^{-E_*^-}(t_n^- - R_*^-)} = Y_{e^{-E_*^-}(t_n^+ - R_*^-)}\right). \quad (112)$$

Nach Lemma 4.1.3 gilt $(Y_t^i(E_*^-))_{t \geq 0} \sim (Y_{e^{-E_*^-}t}^i)_{t \geq 0}$ mit $Y \sim \text{Yule}(1)$ und Y, E_*^- unabhängig. Weiter fordern wir $Y, E_*^-, (R_*^-, C_*^-)$ unabhängig. Insbesondere sind die endlich dimensionalen Verteilungen gleich und weil außerdem $(Y^i)^{i \in \mathbb{N}}, E_*^-, (R_*^-, C_*^-)$ nach Konstruktion unabhängig sind, folgt

$$\begin{aligned} & P\left(Y_{(t_n^- - R_*^-)}^{C_*^-}(E_*^-) = Y_{(t_n^+ - R_*^-)}^{C_*^-}(E_*^-)\right) \\ &= \int_{\mathbb{R}^2} P\left(Y_{(t_n^- - t)}^i(E_*^-) = Y_{(t_n^+ - t)}^i(E_*^-)\right) dP_{(C_*^-, R_*^-)}((i, t)) \\ &= \int_{\mathbb{R}^2} \sum_{m=1}^{\infty} P\left(Y_{(t_n^- - t)}^i(E_*^-) = m, Y_{(t_n^+ - t)}^i(E_*^-) = m\right) dP_{(C_*^-, R_*^-)}((i, t)) \end{aligned}$$

$$\begin{aligned}
&= \int_{\mathbb{R}^2} \sum_{m=1}^{\infty} P\left(Y_{e^{-E_*^-}(t_n^- - t)} = m, Y_{e^{-E_*^-}(t_n^+ - t)} = m\right) dP_{(C_*^-, R_*^-)}((i, t)) \\
&= \int_{\mathbb{R}^2} P\left(Y_{e^{-E_*^-}(t_n^- - t)} = Y_{e^{-E_*^-}(t_n^+ - t)}\right) dP_{(C_*^-, R_*^-)}((i, t)) \\
&= P\left(Y_{e^{-E_*^-}(t_n^- - R_*^-)} = Y_{e^{-E_*^-}(t_n^+ - R_*^-)}\right).
\end{aligned}$$

Weiter ist $R_*^- \sim \text{unif}[0, t_n^-]$ und wir hatten uns schon überlegt, dass dann gilt

$$U = \frac{t_n^- - R_*^-}{t_n^-}, \quad (113)$$

mit $U \sim \text{unif}[0, 1]$, für alle $n \in \tilde{\mathbb{N}}$. (vgl.(95))

Außerdem gilt nach Lemma 4.1.5 $\forall n \in \tilde{\mathbb{N}}$

$$e^{-E_*^-} \sim \frac{E \wedge t_n^-}{t_n^-}, \quad (114)$$

mit $E \sim \exp(1)$.

Dabei bleiben Y, U und E nach Konstruktion unabhängig. Aus (113) und (114) folgt nun für $n \in \tilde{\mathbb{N}}$ direkt

$$\left(e^{-E_*^-}(t_n^- - R_*^-), e^{-E_*^-}(t_n^+ - R_*^-)\right) \sim \left(U(E \wedge t_n^-), U(E \wedge t_n^-) \frac{t_n^+ - R_*^-}{t_n^- - R_*^-}\right). \quad (115)$$

Damit können wir an die letzte Gleichung anknüpfen und erhalten mit der üblichen Unabhängigkeitsargumentation durch Bedingen

$$\begin{aligned}
P\left(Y_{(t_n^- - R_*^-)}^{C_*^-}(E_*^-) = Y_{(t_n^+ - R_*^-)}^{C_*^-}(E_*^-)\right) &= P\left(Y_{e^{-E_*^-}(t_n^- - R_*^-)} = Y_{e^{-E_*^-}(t_n^+ - R_*^-)}\right) \\
&= P\left(Y_{U(E \wedge t_n^-)} = Y_{U(E \wedge t_n^-) \frac{t_n^+ - R_*^-}{t_n^- - R_*^-}}\right).
\end{aligned}$$

Setzen wir für $n \in \tilde{\mathbb{N}}$

$$A_n := \frac{t_n^+ - R_*^-}{t_n^- - R_*^-}, \quad (116)$$

so folgt nach Lemma 4.1.4, dass $A_n \xrightarrow{P} 1$ gilt.

Wir haben noch zu zeigen

$$P\left(Y_{U(E \wedge t_n^-)} = Y_{U(E \wedge t_n^-) A_n}\right) \longrightarrow 1, \quad \text{für } n \rightarrow \infty. \quad (117)$$

Dafür zeigen wir

$$P(Y_{UE} = Y_{UEA_n}) \longrightarrow 1, \quad \text{für } n \rightarrow \infty. \quad (118)$$

Wegen $t_n^- \rightarrow \infty$ *f.s.* für $n \rightarrow \infty$ folgt dann (117) aus

$$\begin{aligned} & P\left(Y_{U(E \wedge t_n^-)} = Y_{U(E \wedge t_n^-)A_n}\right) \\ &= P\left(Y_{UE} = Y_{UEA_n}, E \leq t_n^-\right) + \underbrace{P\left(Y_{Ut_n^-} = Y_{Ut_n^-A_n}, E > t_n^-\right)}_{\xrightarrow{n \rightarrow \infty} 0} \xrightarrow{n \rightarrow \infty} 1. \end{aligned}$$

Für (117) nutzen wir nun $A_n \xrightarrow{P} 1$. Wegen der Unabhängigkeit von U, E und Y folgt durch Anwendung des Satzes von der majorisierten Konvergenz

$$\begin{aligned} & \lim_{n \rightarrow \infty} P(Y_{UE} = Y_{UEA_n}) \\ &= \lim_{n \rightarrow \infty} \int_{\mathbb{R}^2} P(Y_{UE} = Y_{UEA_n} \mid U = u, E = x) dP_{(E,U)}(x, u) \\ &= \lim_{n \rightarrow \infty} \int_{\mathbb{R}^2} P(Y_{ux} = Y_{uxA_n} \mid U = u, E = x) dP_{(E,U)}(x, u) \\ &= \int_{\mathbb{R}^2} \lim_{n \rightarrow \infty} P(Y_{ux} = Y_{uxA_n}) dP_{(E,U)}(x, u) \\ &= \int_{\mathbb{R}^2} \left[\lim_{n \rightarrow \infty} \int_{\mathbb{R}} P(Y_{ux} = Y_{uxA_n} \mid A_n = t) dP_{A_n}(t) \right] dP_{(E,U)}(x, u) \\ &= \int_{\mathbb{R}^2} \left[\lim_{n \rightarrow \infty} \int_{\mathbb{R}} \underbrace{P(Y_{ux} = Y_{uxt})}_{\in \mathcal{C}_b} dP_{A_n}(t) \right] dP_{(E,U)}(x, u) \\ &= \int_{\mathbb{R}^2} \left[\int_{\mathbb{R}} P(Y_{ux} = Y_{uxt}) dP_{\delta_1}(t) \right] dP_{(E,U)}(x, u) \\ &= \int_{\mathbb{R}^2} \underbrace{P(Y_{ux} = Y_{ux})}_{=1} dP_{(E,U)}(x, u) = 1. \end{aligned}$$

Dass $f(t) := P(Y_{ux} = Y_{uxt}) \in \mathcal{C}_b$ ist hatten wir in Lemma 2.2.6 gezeigt. Damit ist alles bewiesen. \square

4.2 Folgerungen

Folgerung 1:

Die gemeinsame Konvergenz im Hauptsatz erlaubt uns natürlich Aussagen folgender Art zu treffen:

Lemma 4.2.1. *Sei $BSC^{[n]}$ ein Bolthausen-Sznitman Koaleszent restringiert auf $[n]$. Gegeben das Ereignis $\{\#B^{[n]} = 2\}$ gilt für $n \rightarrow \infty$*

$$\frac{\log |M_{(\neq 1)}^{[n]}|}{\log n} \xrightarrow{d} V, \quad (119)$$

wobei V die Dichte $f(x) = \frac{1}{(1+x)\log 2} \mathbb{1}_{[0,1]}(x)$ besitzt.

Bemerkung: Die Dichte f ist auf dem Intervall $[0, 1]$ offenbar konvex und streng monoton fallend. Wir sollten also tendenziell einen Wert nahe bei 0 annehmen. Gegeben $\{\#B^{[n]} = 2\}$ gilt dann für großes n näherungsweise

$$|M_{(\neq 1)}^{[n]}| \approx n^V, \quad (120)$$

d.h. mit dieser Zusatzinformation sollte $|M_{(\neq 1)}^{[n]}|$ (noch) kleiner sein. Dies würden wir auch intuitiv erwarten, denn betrachten wir den letzten Abkömmling der Baumkonstruktion, so bedeutet $\{\#B^{[n]} = 2\}$, dass unmittelbar vor der finalen Kollision schon alle Kanten innerhalb dieses letzten Abkömmlings geschnitten wurden. Dies wird um so unwahrscheinlicher, je größer der Abkömmling ist, wenn man bedenkt, dass kein Kantenschnitt bevorzugt ist. Also sollte der Abkömmling eher klein sein.

Beweis: Sei $k \in [0, 1]$. Mit den Notationen des Hauptsatzes folgt zunächst nach Lemma 2.2.8, dass $P(Y_{UE} = 1) = \log 2$ gilt. Für deterministisches $z \geq 0$ ist andererseits $Y_z \sim \text{geo}(e^{-z})$. Wir verwenden also den Hauptsatz und durch kanonisches Bedingen und unter Ausnutzung des Satzes von Fubini folgt:

$$\begin{aligned} P \left(\frac{\log |M_{(\neq 1)}^{[n]}|}{\log n} \leq k \mid \#B^{[n]} = 2 \right) &= \frac{P \left(\frac{\log |M_{(\neq 1)}^{[n]}|}{\log n} \leq k, \#B^{[n]} = 2 \right)}{P(\#B^{[n]} = 2)} \\ \xrightarrow{n \rightarrow \infty} \frac{P(U \leq k, Y_{UE} = 1)}{\underbrace{P(Y_{UE} = 1)}_{\log 2}} \\ &= \frac{1}{\log 2} \int_{\mathbb{R}^2} P(U \leq k, Y_{UE} = 1 \mid U = u, E = x) dP_{(U,E)}(u, x) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\log 2} \int_0^k \left(\int_0^\infty P(Y_{ux} = 1) e^{-x} d\lambda(x) \right) d\lambda(u) \\
&= \frac{1}{\log 2} \int_0^k \underbrace{\left(\int_0^\infty \overbrace{e^{-ux} (1 - e^{-ux})^0 e^{-x}}^{e^{-x(1+u)}} d\lambda(x) \right)}_{\frac{1}{1+u}} d\lambda(u) = \int_0^k \frac{1}{(1+u) \log 2} d\lambda(u).
\end{aligned}$$

Mit dem Eindeutigkeitsatz folgt die Behauptung. \square

Folgerung 2:

Diese Aussage bezieht sich auf den Zeitpunkt der letzten Kollision eines $BSC^{[n]}$. Dieser sei $\forall n \in \mathbb{N}$ mit A_n bezeichnet. Also ist

$$A_n = \inf \left\{ t \geq 0 \mid BSC_t^{[n]} = 1 \right\}. \quad (121)$$

Dann gilt das folgende Lemma.

Lemma 4.2.2. *Für $n \rightarrow \infty$ gilt*

$$A_n - \log \log n \xrightarrow{d} -\log E, \quad (122)$$

mit $E \sim \exp(1)$.

Für eine sehr große Blockanzahl gilt also näherungsweise

$$A_n \approx \log \log n - \log E = \log \log \left(n^{\frac{1}{E}} \right). \quad (123)$$

D.h. die Zeitpunkt der letzten Verschmelzung ist asymptotisch nur von der Größenordnung $\log \log n$. Wir werden nur sehr kurze Zeit warten müssen, bis wir uns im Zustand von nur noch einem Block befinden.

Beweis: In Analogie zur Notation im Hauptbeweis gilt: $A_n = E_*$. Die Aussage ist damit lediglich eine Folgerung von Lemma 4.1.5. Wir erinnern

$$t_n^- e^{-E_*^-} \sim E \wedge t_n^-,$$

mit $E \sim \exp(1)$. Um es anwenden zu können haben wir noch E_* durch E_*^- zu ersetzen. Das ist insbesondere wieder auf unserer Menge \mathcal{G}_n möglich. Gegeben \mathcal{G}_n folgt dann aber direkt

$$\log(E \wedge t_n^-) \sim \log t_n^- e^{-E_*^-} = \log t_n^- e^{-E_*}$$

$$\begin{aligned}
&= \log t_n^- + \log e^{-E_*} = \log \left(\log n - (\log n)^{\frac{1}{2}} \right) - E_* \\
&= \log(\log n)^{\frac{1}{2}} + \underbrace{\log \left((\log n)^{\frac{1}{2}} - 1 \right)}_{\log(\log n)^{\frac{1}{2}} + o(1)} - E_* \\
&= \log \log n + o(1) - E_*.
\end{aligned}$$

$$\Rightarrow E_* - \log \log n \sim -\log(E \wedge t_n^-) + o(1) \xrightarrow{n \rightarrow \infty} -\log E \quad f.s. \quad (124)$$

Wegen $P(\mathcal{G}_n) \rightarrow 1$ für $n \rightarrow \infty$ folgt dann die allgemeine Behauptung. \square

Folgerung 3:

Wir betrachten nun noch kurz eine Konstruktion die auf Janson (vgl.[11]) zurückgeht. Sei ein zufälliger rekursiver Baum mit n Knoten gegeben und jede Kante e mit einer ZV λ_e versehen. Diese seien i.i.d. mit kontinuierlicher Verteilung und damit paarweise verschieden. Weiter nennen wir λ_e einen **Rekord**, wenn es den größten Wert aller angefügten ZVn von der Wurzel bis zu sich selbst annimmt. Dann ist die Verteilung der Anzahl der Rekorde gleich der Verteilung der Anzahl der Schnitte, die nötig sind, um in unserem konventionellen Baummodell durch das Abholzverfahren die Wurzel zu isolieren. Denn durchlaufen wir die der Größe nach abfallend angeordneten Rekorde, so entspricht das einem uniformen Auswählen der Kanten, da alle λ_e i.i.d. sind. Nichts anderes geschieht aber beim Abholzprozedere.

Insbesondere sind definitionsgemäß alle Kanten der Wurzel mit Rekorden versehen und der kleinste aller Rekorde λ_* hängt immer an einer Kante, die an die Wurzel ragt.

Sei nun die Größe $\#A^{[n]}$ des maximalen Teilbaums, dessen einzig enthaltener Rekord der Kleinste λ_* ist.

Identifizieren wir also λ_* mit der im Abholzprozess zuletzt geschnittenen Kante und alle größeren Rekorde mit den entsprechenden - bis dahin schon geschnittenen - Kanten, so entspricht die Größe $\#A^{[n]}$ genau der Anzahl der zum letzten Schnitt noch im Baum vorhandenen Knoten $\#B^{[n]}$. Mit der Notation aus dem Hauptsatz folgt dann

Korollar 4.2.3. *Für $n \rightarrow \infty$ gilt*

$$\#A^{[n]} \xrightarrow{d} 1 + Y_{UE}. \quad (125)$$

Literatur

- [1] L.C.G. Rogers & D. Williams. (2000) Diffusions, Markov Processes and Martingales. *Cambridge: Cambridge Univ. Press.*
- [2] Norris, J. R. (1998) Markov chains. Cambridge Series in Statistical and Probabilistic Mathematics, 2. *Cambridge: Cambridge Univ. Press.*
- [3] Krenzel, U. (2005) Einführung in die Wahrscheinlichkeitstheorie und Statistik. Für Studium, Berufspraxis und Lehramt. 8th expanded ed. Vieweg Studium: Aufbaukurs Mathematik. *Wiesbaden: Vieweg.*
- [4] Kersting, G., Wakolbinger, A. (2008) Elementare Stochastik. Mathematik Kompakt. *Basel: Birkhäuser.*
- [5] Arratia, R., Barbour, A. D., Tavaré, S. (2003) Logarithmic combinatorial structures: a probabilistic approach. EMS Monographs in Mathematics. *Zürich: European Mathematical Society (EMS).*
- [6] Pitman, J. (1995) Exchangeable and partially exchangeable random partitions. *Probab. Theory Related Fields* **102**, 145–158.
- [7] Pitman, J. (1999) Coalescents with multiple collisions. *Ann. Probab.* **27**, 4, 1870–1902.
- [8] Pitman, J. (2002) Combinatorial stochastic processes. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002. With a foreword by Jean Picard. Lecture Notes in Mathematics, 1875. *Berlin: Springer-Verlag.*
- [9] G. Kersting (2005) Schwache Konvergenz und Martingalprobleme. *Skript zur Vorlesung.*
- [10] Meir, A.; Moon, J. W. (1970) Cutting down random trees. *J. Austral. Math. Soc.* **11**, 313–324.
- [11] Janson, S. (2006) Random cutting and records in deterministic and random trees. *Random Structures Algorithms* **29**, 2, 139–179.
- [12] Goldschmidt, C., Martin, J. B. (2005) Random recursive trees and the Bolthausen-Sznitman coalescent. *Electron. J. Probab.* **10**, 21, 718–745.
- [13] Iksanov, A., Möhle, M. (2007) A probabilistic proof of a weak limit law for the number of cuts needed to isolate the root of a random recursive tree. *Electron. Comm. Probab.* **12**, 28–35.

- [14] Panholzer, A. (2004) Destruction of recursive trees. em Mathematics and computer science. III, 267–280, Trends Math., *Basel: Birkhäuser*.
- [15] Sagitov, S. (1999) The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.* **36**, 4, 1116–1125.