Title
Doob–Martin compactification of a Markov chain for growing random words sequentially

Journal
Stochastic Processes and their Applications, 127(7)

ISSN
0304-4149

Authors
Choi, Hye Soo
Evans, Steven N

Publication Date
2017-07-01

DOI
10.1016/j.spa.2016.11.006

Peer reviewed

DOOB-MARTIN COMPACTIFICATION OF A MARKOV CHAIN FOR GROWING RANDOM WORDS SEQUENTIALLY

HYE SOO CHOI AND STEVEN N. EVANS

Abstract. We consider a Markov chain that iteratively generates a sequence of random finite words in such a way that the n^{th} word is uniformly distributed over the set of words of length $2n$ in which n letters are a and n letters are b: at each step an a and a b are shuffled in uniformly at random among the letters of the current word. We obtain a concrete characterization of the Doob-Martin boundary of this Markov chain and thereby delineate all the ways in which the Markov chain can be conditioned to behave at large times. Writing $N(u)$ for the number of letters a (equivalently, b) in the finite word u, we show that a sequence $(u_n)_{n \in \mathbb{N}}$ of finite words converges to a point in the boundary if, for an arbitrary word v, there is convergence as n tends to infinity of the probability that the selection of $N(v)$ letters a and $N(v)$ letters b uniformly at random from u_n and maintaining their relative order results in v. We exhibit a bijective correspondence between the points in the boundary and ergodic random total orders on the set $\{a_1, b_1, a_2, b_2, \ldots\}$ that have distributions which are separately invariant under finite permutations of the indices of the a's and those of the b's. We establish a further bijective correspondence between the set of such random total orders and the set of pairs (μ, ν) of diffuse probability measures on $[0, 1]$ such that $\frac{1}{2}(\mu + \nu)$ is Lebesgue measure: the restriction of the random total order to $\{a_1, b_1, \ldots, a_n, b_n\}$ is obtained by taking $X_1, \ldots, X_n$ (resp. $Y_1, \ldots, Y_n$) i.i.d. with common distribution μ (resp. ν), letting $(Z_1, \ldots, Z_{2n})$ be $\{X_1, Y_1, \ldots, X_n, Y_n\}$ in increasing order, and declaring that the k^{th} smallest element in the restricted total order is a_i (resp. b_j) if $Z_k = X_i$ (resp. $Z_k = Y_j$).

2010 *Mathematics Subject Classification.* Primary: 05A05, 60J10, 68R15.

Key words and phrases. harmonic function, exchangeability, bridge, shuffle, subword counting, binomial coefficient, Plackett-Luce model, vase model.

S.N.E. was supported in part by NSF grants DMS-0907639 and DMS-1512933, and NIH grant 1R01GM109454-01.

1. Introduction

There is a very simple way of producing a uniformly distributed random permutation of a set with n objects, say $[n] := \{1, \ldots, n\}$: we take the elements of $[n]$ in order and lay them down successively so that the k^{th} element goes into a uniformly chosen one of the k "slots" defined by the $k-1$ elements that have already been laid down (the slot before the first element, the slot after the last element, or one of the $k-2$ slots between elements). This sequential algorithm has the attractive feature that when the first k elements have been laid down they are in uniform random order; that is, the algorithm builds uniformly distributed random permutations of $[1], [2], \ldots, [n]$ in a sequential manner.

Suppose that we enumerate a standard deck of cards with the elements of the set [52]. If the deck is in some order, then the colors of the successive cards (**R**ed or **B**lack) define a word of length 52 from the two-letter alphabet $\{R, B\}$ in which 26 letters are R and 26 letters are B (recall that a word of length k from a finite alphabet $\mathcal{A}$ is just an element of the Cartesian product $\mathcal{A}^k$, although it is usual to write the word $(a_1, \ldots, a_k)$ more succinctly as $a_1 \cdots a_k$). Moreover, if the order of the deck is random and uniformly distributed, then the resulting word is uniformly distributed over the set of $\frac{52!}{26!26!}$ such words.

Unfortunately, our sequential randomization algorithm doesn't have the feature that at the $(2k)^{th}$ step for $1 \leq k \leq 26$ we have a random word from the alphabet $\{R, B\}$ that is uniformly distributed over the set of $\binom{2k}{k}$ words in which k letters are R and k letters are B.

However, there is a simple way of modifying our algorithm to produce the latter type of random words sequentially. We begin at step 0 with the empty word. Suppose that we have completed k steps and a word of length $2k$ has been produced. The first sub-step of step $k+1$ inserts the letter R uniformly at random into one of the $2k+1$ slots defined by these $2k$ letters to produce a word of length $2k+1$. The second sub-step inserts the letter B uniformly at random into one of the $2k+2$ slots defined by these $2k+1$ letters to produce a word of length $2k+2$ and thereby complete step $k+1$. It is not difficult to see that, despite the apparent dependence of this procedure on the ordering of the letters R and B, this procedure does indeed achieve what it is claimed to achieve.

From now on we will replace the alphabet $\{R, B\}$ by the alphabet $\{a, b\}$ and write $(U_n)_{n \in \mathbb{N}_0}$ for the Markov chain that arises from our random insertion procedure. Thus, $U_n \in \mathbb{W}_n$, where $\mathbb{W}_n$ is the set words drawn from the alphabet $\{a, b\}$ that consist of n letters a and n letters b. Set $\mathbb{W} := \bigsqcup_{n \in \mathbb{N}_0} \mathbb{W}_n$ and put $N(w) = n$ for $w \in \mathbb{W}_n$, $n \in \mathbb{N}_0$.

We investigate the infinite bridges (equivalently, the Doob h-transforms) for the Markov chain $(U_n)_{n\in\mathbb{N}_0}$; that is, the Markov chains that have the same backwards-in-time transition dynamics as $(U_n)_{n\in\mathbb{N}_0}$. We thereby identify the Doob-Martin compactification of the state space $\mathbb{W}$ of the Markov chain. This enables us to characterize the nonnegative harmonic functions for the Markov chain and hence delineate all the ways that the Markov chain can be conditioned to "behave at infinity".

More specifically, we show that a $\mathbb{W}$-valued Markov chain is an infinite bridge for the Markov chain $(U_n)_{n\in\mathbb{N}_0}$ if and only if the backwards dynamics are given by removing one letter a and one letter b uniformly at random from the current word. We can enrich the state space of the Markov chain $(U_n)_{n\in\mathbb{N}_0}$ by replacing $\mathbb{W}_n$ with the set $\tilde{\mathbb{W}}_n$ that consists of words made up from the letters $a_1, b_1, \ldots, a_n, b_n$ written down in some order (each letter appearing once); that is, a word such as $aababb$ will be associated with a word such as $a_3a_1b_2a_2b_1b_3$ – a given $w \in \mathbb{W}_n$ has $(n!)^2$ associated words in $\tilde{\mathbb{W}}_n$. We can then enhance an infinite bridge $(U_n^\infty)_{n\in\mathbb{N}_0}$ to produce a Markov chain $(\tilde{U}_n^\infty)_{n\in\mathbb{N}_0}$ with values in $\tilde{\mathbb{W}} := \bigsqcup_{n\in\mathbb{N}_0} \tilde{\mathbb{W}}_n$ such that given $U_n^\infty = u$ the value of $\tilde{U}_n^\infty$ is uniformly distributed over all ways of "subscripting" the letters in u; for example, if $U_2^\infty = abba$, then $\tilde{U}_2^\infty$ is uniformly distributed over the four words $a_1b_1b_2a_2$, $a_2b_1b_2a_1$, $a_1b_2b_1a_2$, $a_2b_2b_1a_1$. Moreover, in going from $\tilde{U}_n^\infty$ to $\tilde{U}_{n-1}^\infty$ the letters a_n and b_n are deleted. We may view $\tilde{U}_n^\infty$ as a random total (that is, linear) order on the set $\{a_1, b_1, \ldots, a_n, b_n\}$. As n varies, these orders are consistent in the sense that the order $\tilde{U}_n^\infty$ induces on $\{a_1, b_1, \ldots, a_{n-1}, b_{n-1}\}$ is just the order given by $\tilde{U}_{n-1}^\infty$. Consequently, there is a total order on $\{a_1, b_1, a_2, b_2, \ldots\}$ that induces each of the orders given by the $\tilde{U}_n^\infty$. This total order is exchangeable in the sense that finite permutations of the subscripts of the a's and b's separately leave its distribution unchanged. The infinite bridge $(U_n^\infty)_{n\in\mathbb{N}_0}$ is extremal (that is, not a mixture of infinite bridges or, equivalently, has an almost surely trivial tail σ-field) if and only if the exchangeable random total order on $\{a_1, b_1, a_2, b_2, \ldots\}$ is ergodic in the sense that if an event is unchanged by finite permutations of the subscripts of the a's and b's separately, then it has probability zero or one. By general Doob–Martin theory, extremal bridges correspond to extremal elements of the Doob–Martin boundary and, in general, some elements of the Doob–Martin boundary may not be extremal. We show that the latter phenomenon does not occur in our setting – all Doob–Martin boundary points are extremal.

We demonstrate that there is a bijective correspondence between ergodic exchangeable random total orders on $\{a_1, b_1, a_2, b_2, \ldots\}$ and pairs (μ, ν) of diffuse probability measures on the unit interval $[0, 1]$ such that $\frac{\mu+\nu}{2} = \lambda$, where λ is Lebesgue measure on $[0, 1]$: let $V_1, V_2, \ldots$ be i.i.d. with distribution μ and $W_1, W_2, \ldots$ be independent and i.i.d. with distribution ν, then, writing $\prec$ for the total order, we have $a_i \prec a_j$ (resp. $a_i \prec b_j$, $b_i \prec a_j$, $b_i \prec b_j$) if $V_i < V_j$ (resp. $V_i < W_j$, $W_i < V_j$, $W_i < W_j$). Another way of describing this construction is the following. We only need to describe the restriction of the random total order to $\{a_1, b_1, \ldots, a_n, b_n\}$ for each $n \in \mathbb{N}_0$. Let $(Z_1, \ldots, Z_{2n})$ be $\{V_1, W_1, \ldots, V_n, W_n\}$ in increasing order and declare that the k^{th} smallest element of $\{a_1, b_1, \ldots, a_n, b_n\}$ in the restricted total order is a_i (resp. b_j) if $Z_k = X_i$ (resp. $Z_k = Y_j$).

We remark that, due to the relationship $\frac{\mu+\nu}{2} = \lambda$, the probability measure ν is uniquely determined by the probability measure μ and *vice versa* and hence we could have said that the ergodic exchangeable random total orders are in bijective correspondence with the probability measures μ on $[0, 1]$ that satisfy $\mu \le 2\lambda$. However, we find the more symmetric description to be preferable.

In terms of the Doob–Martin topology, we show that a sequence $(y_k)_{k\in\mathbb{N}}$ with $y_k \in \mathbb{W}_{N(y_k)}$ and $N(y_k) \to \infty$ as $k \to \infty$ converges to the point in the Doob–Martin boundary corresponding to the pair of measures (μ, ν) if and only if for each $m \in \mathbb{N}$ the the sequence of random words obtained by selecting m letters a and m letters b uniformly at random from y_k and maintaining their relative order converges in distribution as $k \to \infty$ to the random word that is obtained by writing $V_1, \ldots, V_m, W_1, \ldots, W_m$ in increasing order to make a list $(Z_1, \ldots, Z_{2m})$ as above and then putting a letter a (resp. b) in position ℓ of the word when $Z_\ell \in \{V_1, \ldots, V_m\}$ (resp. $Z_\ell \in \{W_1, \ldots, W_m\}$). Moreover, the convergence of $(y_k)_{k\in\mathbb{N}}$ to y is equivalent to the weak convergence of μ_k to μ and ν_k to ν, where μ_k (resp. ν_k) is the probability measure that places mass $\frac{1}{N(y_k)}$ at the point $\frac{\ell}{2N(y_k)}$ $1 \le \ell \le 2N(y_k)$, if the ℓ^{th} letter of the word y_k is the letter a (resp. b).

2. Background on the Doob–Martin compactification

The primary reference on the Doob–Martin compactification theory for discrete time Markov chains is [Doo59], but useful reviews may be found in [KSK76, Chapter 10], [Rev75, Chapter 7], [Saw97], [Woe00, Chapter IV], [RW00, Chapter III]. We restrict the following sketch to the setting that is of interest to us.

Suppose that $(X_n)_{n\in\mathbb{N}_0}$ is a discrete time Markov chain with countable state space E and transition matrix P. Suppose in addition that E can be partitioned as $E = \bigsqcup_{n\in\mathbb{N}_0} E_n$, where $E_0 = \{e\}$ for some distinguished state e, each set E_n, $n \in \mathbb{N}_0$ is finite, and the transition matrix P is such that $P(k,\ell) = 0$ unless $k \in E_n$ and $\ell \in E_{n+1}$ for some $n \in \mathbb{N}_0$. Define the *Green kernel* or *potential kernel* G of P by

$$G(i,j) := \sum_{n=0}^{\infty} P^n(i,j) = \mathbb{P}^i\{X_n = j \text{ for some } n \in \mathbb{N}_0\} =: \mathbb{P}^i\{X \text{ hits } j\},$$

$i, j \in E$, and assume that $G(e,j) > 0$ for all $j \in E$, so that any state can be reached with positive probability starting from e.

The *Doob–Martin kernel with reference state* e is

$$K(i,j) := \frac{G(i,j)}{G(e,j)} = \frac{\mathbb{P}^i\{X \text{ hits } j\}}{\mathbb{P}^e\{X \text{ hits } j\}}.$$

If $j, k \in E$ with $j \neq k$, then $K(\cdot, j) \neq K(\cdot, k)$ and so E can be identified with the collection of functions $(K(\cdot, j))_{j\in E}$. Note that

$$0 \le K(i,j) \le \frac{1}{\mathbb{P}^e\{X \text{ hits } i\}},$$

and so the set of functions $(K(\cdot, j))_{j\in E}$ is a pre-compact subset of $\mathbb{R}_+^E$. Its closure $\bar{E}$ is the *Doob–Martin compactification* of E. The set $\partial E := \bar{E} \setminus E$ is the *Doob–Martin boundary* of E.

By definition, a sequence $(j_n)_{n\in\mathbb{N}}$ in E converges to a point in $\bar{E}$ if and only if the sequence of real numbers $(K(i, j_n))_{n\in\mathbb{N}}$ converges for all $i \in E$. Each function $K(i, \cdot)$ extends continuously to $\bar{E}$. The resulting function $K : E \times \bar{E} \to \mathbb{R}$ is the *extended Martin kernel.* For $y \in \partial E$ the nonnegative function $K(\cdot, y)$ is harmonic and any nonnegative harmonic function can be represented as $\int K(\cdot, y)\, \mu(dy)$ for a suitable finite measure μ on ∂E.

If Z is a $\mathbb{P}^e$-a.s. bounded random variable that is measurable with respect to the tail σ-field of $(X_n)_{n\in\mathbb{N}_0}$, then $\mathbb{E}^e[Z \mid X_0, \ldots, X_n] = h(X_n)$ for some bounded harmonic function h and, by the martingale convergence theorem, $\lim_{n\to\infty} h(X_n) = Z$ $\mathbb{P}^e$-a.s. Conversely, if h is a bounded harmonic function, then $\lim_{n\to\infty} h(X_n)$ exists $\mathbb{P}^e$-a.s. and the limit random variable is $\mathbb{P}^e$-a.s. equal to a random variable that is measurable with respect to the tail σ-field of $(X_n)_{n\in\mathbb{N}_0}$.

The limit $X_\infty := \lim_{n\to\infty} X_n$ exists $\mathbb{P}^e$-almost surely in the topology of $\bar{E}$ and the limit belongs to ∂E $\mathbb{P}^e$-almost surely. The tail σ-field of $(X_n)_{n\in\mathbb{N}_0}$ coincides $\mathbb{P}^e$-almost surely with the σ-field generated by X_∞.

Each $j \in E = \bigsqcup_{n\in\mathbb{N}_0} E_n$ belongs to a unique E_n whose index n we denote by $N(j)$. If the Markov chain starts in state e, then $N(j)$ is the only time that there is positive probability the Markov chain will be in state j. Write $(X_0^j, \ldots, X_{N(j)}^j)$ for the *bridge* obtained by starting the Markov chain in state e and conditioning it to be in state j at time $N(j)$. This process is a Markov chain with transition probabilities

$$\begin{aligned}
\mathbb{P}\{X_{n+1}^j = i'' \,|\, X_n^j = i'\} &= \frac{\mathbb{P}^e\{X_n = i',\, X_{n+1} = i'',\, X_{N(j)} = j\}}{\mathbb{P}^e\{X_n = i',\, X_{N(j)} = j\}} \\
&= \frac{\mathbb{P}^e\{X \text{ hits } i'\}P(i',i'')\mathbb{P}^{i''}\{X \text{ hits } j\}}{\mathbb{P}^e\{X \text{ hits } i'\}\mathbb{P}^{i'}\{X \text{ hits } j\}} \\
&= \frac{P(i',i'')\mathbb{P}^{i''}\{X \text{ hits } j\}/\mathbb{P}^e\{X \text{ hits } j\}}{\mathbb{P}^{i'}\{X \text{ hits } j\}/\mathbb{P}^e\{X \text{ hits } j\}} \\
&= K(i',j)^{-1}P(i',i'')K(i'',j).
\end{aligned}$$

The backward transition probabilities of $(X_0^j, \ldots, X_{N(j)}^j)$ are given by

$$\begin{aligned}
\mathbb{P}\{X_n^j = i' \,|\, X_{n+1}^j = i''\} &= \frac{\mathbb{P}^e\{X \text{ hits } i'\}P(i',i'')\mathbb{P}^{i''}\{X \text{ hits } j\}}{\mathbb{P}^e\{X \text{ hits } i''\}\mathbb{P}^{i''}\{X \text{ hits } j\}} \\
&= \frac{\mathbb{P}^e\{X \text{ hits } i'\}P(i',i'')}{\mathbb{P}^e\{X \text{ hits } i''\}},
\end{aligned}$$

so that all bridges have the same backward transition probabilities. An *infinite bridge* for $(X_n)_{n\in\mathbb{N}_0}$ is a Markov chain $(X_n^\infty)_{n\in\mathbb{N}_0}$ with these backward transition probabilities. If $(X_n^\infty)_{n\in\mathbb{N}_0}$ is an infinite bridge, then

$$\begin{aligned}
\mathbb{P}\{X_{n+1}^\infty = i'' \,|\, X_n^\infty = i'\} &= \frac{\mathbb{P}^e\{X^\infty \text{ hits } i''\}\mathbb{P}\{X_n^\infty = i' \,|\, X_{n+1}^\infty = i''\}}{\mathbb{P}^e\{X^\infty \text{ hits } i'\}} \\
&= h(i')^{-1}P(i',i'')h(i''),
\end{aligned}$$

where

$$h(i) = \frac{\mathbb{P}^e\{X^\infty \text{ hits } i\}}{\mathbb{P}^e\{X \text{ hits } i\}}.$$

Thus an infinite bridge is a Doob h-transform of $(X_n)_{n\in\mathbb{N}_0}$ with a particular harmonic function h. Conversely, any Doob h-transform is an infinite bridge.

Suppose now that $(j_k)_{k\in\mathbb{N}}$ is a sequence of elements of the state space E such that $N(j_k) \to \infty$ as $k \to \infty$. As observed in [Föl75], such a sequence $(j_k)_{k\in\mathbb{N}}$ converges in the Doob–Martin topology if and only if finite initial segments of the corresponding bridges converge in distribution. Moreover, two sequences of states converge to the same limit if and only if the limiting distributions of finite initial segments are the same. For a sequence $(j_k)_{k\in\mathbb{N}}$ that converges to a point in the

Doob–Martin boundary, the limiting distributions of the initial segments define the distribution of an E-valued Markov chain $(X_n^{(h)})_{n\in\mathbb{N}_0}$ with transition probabilities $P^{(h)}$ given by

$$P^{(h)}(i,j) := h(i)^{-1}P(i,j)h(j), \quad i,j \in E^{(h)},$$

where $h(i) = \lim_{k\to\infty} K(i,j_k)$ and

$$\begin{aligned} E^{(h)} &:= \{i \in E : h(i) > 0\} \\ &= \{i \in E : \lim_{k\to\infty} \mathbb{P}\{X_{N(i)} = i \,|\, X_{N(j_k)} = j_k\} > 0\}. \end{aligned}$$

This Markov chain $(X_n^{(h)})_{n\in\mathbb{N}_0}$ is an infinite bridge. A necessary condition for an infinite bridge to be extremal (that is, having a distribution that is not a nontrivial mixture of infinite bridge distributions) is that it is of this form.

3. Transition probabilities and the Doob–Martin kernel for the growing word chain

Definition 3.1. For $n \in \mathbb{N}_0$ write $\mathbb{W}_n$ for the set of words from the alphabet $\{a,b\}$ that have n letters a and n letters b and put $\mathbb{W} := \bigsqcup_{n\in\mathbb{N}_0} \mathbb{W}_n$.

By definition, the Markov chain $(U_n)_{n\in\mathbb{N}_0}$ has state space $\mathbb{W}$ and one-step transition probabilities

$$\mathbb{P}\{U_{m+1} = w \,|\, U_m = v\} = \frac{M(v,w)}{(2m+2)(2m+1)}$$

for $v \in \mathbb{W}_n$ and $w \in \mathbb{W}_{n+1}$, where $M(v,w)$ is the number of ways to write $w = v_1 x v_2 y v_3$ in such a way that $\{x,y\} = \{a,b\}$ and v_1, v_2, v_3 are (possibly empty) words such that $v = v_1v_2v_3$. That is, $M(v,w)$ is the number of times that v appears inside w as a *sub-word*. (We recall that, in general, a word $c_1 \cdots c_p$ is a sub-word of a word $d_1 \cdots d_q$ if there is a map $f : [p] \to [q]$ such that $f(i) < f(j)$ for $1 \le i < j \le p$ and $d_{f(k)} = c_k$ for $1 \le k \le p$.)

In order to write down multi-step transition probabilities for the Markov chain $(U_n)_{n\in\mathbb{N}_0}$, it is convenient to introduce the following standard notation (see, for example, [Lot97]).

Definition 3.2. Given two words w and v drawn from some finite alphabet, write $\binom{w}{v}$ for the number of times that v appears as a sub-word of w.

Example 3.3. For example, $\binom{abbaba}{bba} = 4$ because bba appears inside $abbaba$ as a sub-word four times:

a**bba**ba a**bb**ab**a** a**b**ba**ba** ab**b**a**ba**.

Remark 3.4. Note that if our alphabet has only one letter, then $\binom{w}{v}$ is just the usual binomial coefficient $\binom{|w|}{|v|}$, where we use the notation $|u|$ for the length of the word u.

For a general finite alphabet $\mathcal{A}$, $\binom{w}{v}$ is uniquely determined by the following three properties, where we write $\mathcal{A}^*$ for the set of finite words with letters drawn from the alphabet $\mathcal{A}$ (see [Lot97, Proposition 6.3.3]):

- $\binom{w}{\emptyset} = 1$ for all $w \in \mathcal{A}^*$, where $\emptyset$ is the empty word,
- $\binom{w}{v} = 0$ for all $v, w \in \mathcal{A}^*$ with $|w| < |v|$,
- $\binom{wy}{vx} = \binom{w}{vx} + \delta_{x,y}\binom{w}{v}$, for all $v, w \in \mathcal{A}^*$ and $x, y \in \mathcal{A}$, where δ is the usual Kronecker delta.

The counting involved in determining $\binom{w}{v}$ for general $v, w \in \mathcal{A}^*$ is handled by the following result from [Cla15]. Define an infinite matrix $\mathcal{P}$ with entries indexed by $\mathcal{A}^*$ by setting the (v, w) entry to be $\binom{w}{v}$. If the row and column indices are ordered so that they are nondecreasing in word length, then $\mathcal{P}$ is an upper triangular matrix with 1 in every position on the diagonal. Define another infinite matrix $\mathcal{H}$ indexed by $\mathcal{A}^*$ by setting the (v, w) entry to be $\binom{w}{v}$ if $|w| = |v| + 1$ and 0 otherwise. With the same ordering of the indices as for $\mathcal{P}$, the matrix $\mathcal{H}$ is upper triangular with 0 in every position on the diagonal. The matrix exponential $\exp(\mathcal{H})$ is well-defined and is equal to $\mathcal{P}$.

Using the above notation, we can express the transition probabilities of $(U_n)_{n \in \mathbb{N}_0}$ as follows.

Lemma 3.5. *For words $v \in \mathbb{W}_m$ and $w \in \mathbb{W}_{m+n}$*

$$\mathbb{P}\{U_{m+n} = w \,|\, U_m = v\} = \binom{w}{v} \frac{n!n!}{(2m+1)(2m+2)\cdots(2(m+n))}.$$

Proof. We proceed by induction. The result is certainly true when $n = 1$. Supposing it is true for some value of n, in order to show it is true for $n + 1$, we need to show that for $u \in \mathbb{W}_m$ and $w \in \mathbb{W}_{m+n+1}$ we have

$$\sum_{v \in \mathbb{W}_{m+1}} \binom{v}{u} \frac{1}{(2m+1)(2m+2)} \binom{w}{v} \frac{n!n!}{(2m+3)(2m+4)\cdots(2(m+n+1))}$$
$$= \binom{w}{u} \frac{(n+1)!(n+1)!}{(2m+1)(2m+2)\cdots(2(m+n+1))},$$

or, equivalently, that

$$\sum_{v \in \mathbb{W}_{m+1}} \binom{v}{u} \binom{w}{v} = \binom{w}{u} (n+1)^2.$$

This, however, is clear. The lefthand side counts the number of words $v \in \mathbb{W}_{m+1}$ such that u is subword of v and v is a subword of w. Any such v and its embedding in w arises by taking an embedding of u in w and then specifying which of the remaining $n+1$ letters a in w and which of the remaining $n+1$ letters b in w are used to build the word with its particular embedding, and this is what the righthand side counts. □

Corollary 3.6. *The Doob–Martin kernel of* $(U_n)_{n \in \mathbb{N}_0}$ *with distinguished state the empty word is, for* $v \in \mathbb{W}_m$ *and* $w \in \mathbb{W}_{m+n}$*,*

$$K(v, w) = \binom{w}{v} \frac{\binom{2m}{m}}{\binom{m+n}{m}^2}.$$

Proof. We have

$$\begin{aligned}
K(v,w) & \\
&= \frac{\mathbb{P}\{U_{m+n} = w \,|\, U_m = v\}}{\mathbb{P}\{U_{m+n} = w \,|\, U_0 = \emptyset\}} \\
&= \frac{\binom{w}{v} \frac{n!n!}{(2m+1)(2m+2)\cdots(2(m+n))}}{\binom{w}{\emptyset} \frac{(m+n)!(m+n)!}{(2(m+n))!}} \\
&= \binom{w}{v} \frac{n!n!(2(m+n))!}{(m+n)!(m+n)!(2m+1)(2m+2)\cdots(2(m+n))} \\
&= \binom{w}{v} \frac{\binom{2m}{m}}{\binom{m+n}{n}\binom{m+n}{n}}.
\end{aligned}$$

□

Remark 3.7. Up to the factor $\binom{2m}{m}$, the Doob–Martin kernel $K(v, w)$ is the probability that if we select m of the letters a and m of the letters b uniformly at random from w and list these letters in the same relative order that they appear in w, then the resulting word is v. Therefore, a sequence $(w_k)_{k \in \mathbb{N}}$ in $\mathbb{W}$ with $N(w_k) \to \infty$ as $k \to \infty$ converges in the Doob–Martin topology if and only if for every $m \in \mathbb{N}$ the sequence of random words in $\mathbb{W}_m$ obtained by selecting m letters a and m letters b from w_k (and maintaining their relative order) converges in distribution as $k \to \infty$.

Definition 3.8. For $w \in \mathbb{W}_k$, $k \in \mathbb{N}_0$, let $(U_0^w, \dots, U_k^w)$ be the bridge from the empty word to w.

Theorem 3.9. *The backward transition dynamics for all bridges from the empty word are the same and consist of removing at each step one letter a and one letter b uniformly at random.*

Proof. Consider the bridge from the empty word to $w \in \mathbb{W}_k$.

For $0 \le m \le k-1$, $v \in \mathbb{W}_{m+1}$, and $u \in \mathbb{W}_m$ we have

$$
\begin{aligned}
&\mathbb{P}\{U_m^w = u \,|\, U_{m+1}^w = v\} \\
&= \frac{\mathbb{P}\{U_m = u,\, U_{m+1} = v \,|\, U_k = w\}}{\mathbb{P}\{U_{m+1} = v \,|\, U_k = w\}} \\
&= \frac{\mathbb{P}\{U_m = u,\, U_{m+1} = v,\, U_k = w\}}{\mathbb{P}\{U_{m+1} = v,\, U_k = w\}} \\
&= \frac{\mathbb{P}\{U_m = u\}\mathbb{P}\{U_{m+1} = v \,|\, U_m = u\}\mathbb{P}\{U_k = w \,|U_{m+1} = v\}}{\mathbb{P}\{U_{m+1} = v\}\mathbb{P}\{U_k = w \,|U_{m+1} = v\}} \\
&= \binom{v}{u} \frac{\frac{1}{(2m+1)(2m+2)} \times \frac{m!m!}{(2m)!}}{\frac{(m+1)!(m+1)!}{(2m+2)!}} \\
&= \frac{\binom{v}{u}}{(m+1)^2}.
\end{aligned}
$$

In order to go backward from the word v of length $2(m+1)$ to the word u of length $2m$, we have to remove one a and one b. There are $\binom{v}{u}$ pairs of a and b such that the removal of the pair from v results in u, and there are a total of $(m+1)^2$ pairs of a and b in v, and so the result follows from the calculation above. ☐

4. Labeled infinite bridges

Suppose that $(y_n)_{n\in\mathbb{N}}$ is a sequence of words in $\mathbb{W} := \bigsqcup_{n\in\mathbb{N}_0} \mathbb{W}_n$ that converges in the Doob–Martin topology and is such that $N(y_n) \to \infty$ as $n \to \infty$. Recall that $(U_0^{y_n}, \dots, U_{N(y_n)}^{y_n})$, $n \in \mathbb{N}$, is the associated bridge that starts from the empty word and is tied to being in state y_n at time $N(y_n)$. The finite dimensional distributions of $(U_0^{y_n}, \dots, U_{N(y_n)}^{y_n})$ converge as $n \to \infty$. Thus, there exists a process $(U_n^\infty)_{n\in\mathbb{N}_0}$ such that for every $k \in \mathbb{N}_0$ the random $(k+1)$-tuple $(U_0^{y_n}, \dots, U_k^{y_n})$ converges in distribution to $(U_0^\infty, \dots, U_k^\infty)$.

The forward evolution dynamics of the Markov chain $(U_n^\infty)_{n\in\mathbb{N}}$ depend on the sequence $(y_n)_{n\in\mathbb{N}}$, whereas from Section 2 and Theorem 3.9

the backward evolution is Markovian and doesn't depend on the sequence $(y_n)_{n\in\mathbb{N}}$; given U_{k+1}^∞, the word U_k^∞ is obtained by removing one letter a and one letter b uniformly at random from U_{k+1}^∞.

For each $n \in \mathbb{N}_0$ the distribution of U_n^∞ defines the distribution of a random element $\tilde{U}_{n,n}^\infty$ of the set $\tilde{\mathbb{W}}_n$ of words of length $2n$ drawn from the alphabet $\{a_1, b_1, \dots, a_n, b_n\}$ with each letter appearing once by assigning the labels $[n]$ uniformly at random to the letters a and to the letters b. More precisely, for $U_n^\infty = c_1 \dots c_{2n}$, let $A_n := \{i \in [n] : c_i = a\}$ and $B_n := \{j \in [n] : c_j = b\}$, let $\Sigma : A_n \to [n]$ and $T : B_n \to [n]$ be random bijections that are conditionally independent and uniformly distributed given U_n^∞, and define $\tilde{U}_{n,n}^\infty := \tilde{c}_1 \dots \tilde{c}_{2n}$ by

$$\tilde{c}_k := \begin{cases} a_{\Sigma(k)}, & k \in A_n, \\ b_{T(k)}, & k \in B_n. \end{cases}$$

For $0 \le p \le n$, define $\tilde{U}_{n,p}^\infty$ to be the word obtained by deleting $\{a_{p+1}, b_{p+1}, \dots, a_n, b_n\}$ from $\tilde{U}_{n,n}^\infty$. Observe that if $0 \le p \le m \wedge n$, then $\tilde{U}_{m,p}^\infty$, $\tilde{U}_{n,p}^\infty$ and $\tilde{U}_{p,p}^\infty$ have the same distribution. Moreover, if for $0 \le p \le n$ we let $U_{n,p}^\infty$ be the result of removing the labels from $\tilde{U}_{n,p}^\infty$ (that is, $U_{n,p}^\infty$ is the element of $\mathbb{W}_p$ obtained by replacing the letters a_k, $1 \le k \le p$, by the letter a and the letters b_k, $1 \le k \le p$, by b), then $(U_{n,0}^\infty, \dots, U_{n,n}^\infty)$ has the same distribution as $(U_0^\infty, \dots, U_n^\infty)$.

By Kolmogorov's consistency theorem, there is a process $(\tilde{U}_n^\infty)_{n\in\mathbb{N}_0}$ such that $(\tilde{U}_0^\infty, \dots, \tilde{U}_m^\infty)$ has the same distribution as $(\tilde{U}_{n,0}^\infty, \dots, \tilde{U}_{n,m}^\infty)$ for any $m \le n$ and the result of removing the labels from $(\tilde{U}_n^\infty)_{n\in\mathbb{N}_0}$ has the same distribution as $(U_n^\infty)_{n\in\mathbb{N}_0}$. By the transfer theorem [Kal02, Theorem 6.10], we may even suppose that $(\tilde{U}_n^\infty)_{n\in\mathbb{N}_0}$ is defined on an extension of the probability space on which $(U_n^\infty)_{n\in\mathbb{N}_0}$ is defined in such a way that $(U_n^\infty)_{n\in\mathbb{N}_0}$ is the result of removing the labels from $(\tilde{U}_n^\infty)_{n\in\mathbb{N}_0}$.

5. The exchangeable random total order associated with an infinite bridge

A state of a labeled infinite bridge is a word of length $2n$ from the alphabet $\{a_1, b_1, \dots, a_n, b_n\}$ in which each letter appears once. Another way to think of such an object is as a total order on the set $\bigcup_{k=1}^n \{a_k, b_k\}$. Because the labeled infinite bridge evolves by slotting in the letters a_{n+1} and b_{n+1} at the $(n+1)^{\text{th}}$ step while leaving the relative positions of $\{a_1, b_1, \dots, a_n, b_n\}$ unchanged, these successive total orders are consistent: the total order on $\{a_1, b_1, \dots, a_n, b_n\}$ given by the state of the infinite bridge at step n is the same as the total order obtained by taking the state of the infinite bridge at step $n+1$ (a total order on $\{a_1, b_1, \ldots, a_n, b_n, a_{n+1}, b_{n+1}\}$) and looking at the corresponding induced total order on $\{a_1, b_1, \ldots, a_n, b_n\}$.

This projective structure means that we can associate any path of a labeled infinite bridge with a unique total order on $\mathbb{I}_0 := \bigcup_{n \in \mathbb{N}} \{a_n, b_n\}$ such that the induced total order on $\{a_1, b_1, \ldots, a_n, b_n\}$ coincides with the state of the labeled infinite bridge at step n.

We now introduce some general notions about random total orders.

Definition 5.1. A *random total order* $\prec$ on $\mathbb{I}_0$ is a map from the underlying probability space to the collection of total orders on $\mathbb{I}_0$ such that the indicator $\mathbb{1}\{x \prec y\}$ is a random variable for every $x, y \in \mathbb{I}_0$. A random total order $\prec$ is *exchangeable* if for every $n \in \mathbb{N}$ the induced total order $\prec^n$ on $\bigcup_{k=1}^n \{a_k, b_k\}$ has the same distribution as the random total order $\prec^n_{\sigma,\tau}$ for any permutations σ, τ of $\{1, 2, \ldots, n\}$, where $\prec^n_{\sigma,\tau}$ is defined as follows:

- $a_{\sigma(i)} \prec^n_{\sigma,\tau} b_{\tau(j)}$ iff $a_i \prec^n b_j$
- $b_{\tau(i)} \prec^n_{\sigma,\tau} a_{\sigma(j)}$ iff $b_i \prec^n a_j$
- $a_{\sigma(i)} \prec^n_{\sigma,\tau} a_{\sigma(j)}$ iff $a_i \prec^n a_j$
- $b_{\tau(i)} \prec^n_{\sigma,\tau} b_{\tau(j)}$ iff $b_i \prec^n b_j$.

Remark 5.2. The distribution of a random total order $\prec$ is determined by the joint distribution of the random variables $\{\mathbb{1}\{x \prec y\} : x, y \in \bigcup_{k=1}^n \{a_k, b_k\}\}$ for arbitrary $n \in \mathbb{N}$.

Remark 5.3. If $\prec$ is an exchangeable random total order, then the induced random total orders $\prec^n$, $n \in \mathbb{N}$, are consistent in the sense that if we take the random total order $\prec^{n+1}$ on $\bigcup_{k=1}^{n+1} \{a_k, b_k\}$ and remove $\{a_{n+1}, b_{n+1}\}$, then the induced random total order on $\bigcup_{k=1}^n \{a_k, b_k\}$ is $\prec^n$.

Conversely, suppose for each $n \in \mathbb{N}$ that there is a random total order $\prec^n$ on $\bigcup_{k=1}^n \{a_k, b_k\}$, these random total orders have the property that $\prec^n$ has the same distribution as $\prec^n_{\sigma,\tau}$ for any permutations σ, τ of $[n]$ for all $n \in \mathbb{N}$, and these total orders are consistent. Then there is an exchangeable random order $\prec$ on $\mathbb{I}_0$ such that $\prec^n$ is the corresponding induced total order on $\bigcup_{k=1}^n \{a_k, b_k\}$.

In terms of these general notions, if we let $\prec^n$, $n \in \mathbb{N}$, be the random total order on $\bigcup_{k=1}^n \{a_k, b_k\}$ corresponding to $\tilde{U}_n^\infty$, then these total orders are consistent and there is an exchangeable random total order $\prec$ on $\mathbb{I}_0$ such that the restriction of $\prec$ to $\bigcup_{k=1}^n \{a_k, b_k\}$ is $\prec^n$.

6. Characterization of exchangeable random total orders

The results of the previous sections indicate that if we want to understand the Doob–Martin compactification, then we need to understand infinite bridges, and this boils down to understanding exchangeable random total orders on $\mathbb{I}_0$.

A mixture of two exchangeable random total orders is also an exchangeable random total order, so we are interested in exchangeable random total orders $\prec$ that are extremal in the sense that their distributions cannot be written as a nontrivial mixture of the distributions of two other exchangeable random total orders. This is equivalent to requiring that if A is a measurable subset of the space of total orders on $\mathbb{I}_0$ with the property that $\prec \in A$ if and only if $\prec^{\sigma,\tau} \in A$ for all finite permutations σ, τ, then $\mathbb{P}\{\prec \in A\} \in \{0,1\}$. We say that an exchangeable random total order with this property is *ergodic*.

The following result can be established using essentially the same argument as in Proposition 5.19 (see also the subsequent Remark 5.20) of [EGW15], and we omit the details.

Lemma 6.1. *The tail σ-field of an infinite bridge $(U_n^\infty)_{n\in\mathbb{N}_0}$ is almost surely trivial if and only if the exchangeable random total order induced by the corresponding labeled infinite bridge $(\tilde{U}_n^\infty)_{n\in\mathbb{N}_0}$ is ergodic.*

Remark 6.2. There is one obvious way to produce an ergodic exchangeable random total order. Let ζ and η be two diffuse probability measures on $\mathbb{R}$. Let $(V_n)_{n\in\mathbb{N}}$ be i.i.d. with common distribution ζ, let $(W_n)_{n\in\mathbb{N}}$ be i.i.d. with common distribution η, and suppose that these two sequences are independent. The total order $\prec$ on $\mathbb{I}_0$ defined by declaring that

- $a_i \prec a_j$ if $V_i < V_j$,
- $b_i \prec b_j$ if $W_i < W_j$,
- $a_i \prec b_j$ if $V_i < W_j$,
- $b_i \prec a_j$ if $W_i < V_j$,

is exchangeable and ergodic; exchangeability is obvious and ergodicity is immediate from the Hewitt–Savage zero–one law applied to the i.i.d. sequence $((V_n, W_n))_{n\in\mathbb{N}}$ (indeed, it follows from the Hewitt–Savage zero–one law that if A is a measurable subset of the space of total orders on $\mathbb{I}_0$ with the property that $\prec \in A$ if and only if $\prec^{\rho,\rho} \in A$ for all finite permutations ρ, then $\mathbb{P}\{\prec \in A\} \in \{0,1\}$).

We will show that all ergodic exchangeable random total orders arise this way. Note that many pairs of probability measures can give rise to random total orders with the same distribution: replacing ζ and η by their push-forwards by some common strictly increasing function does not change the distribution of the resulting random total order.

Definition 6.3. Given an exchangeable random total order $\prec$ on $\mathbb{I}_0$, define $d : \mathbb{I}_0 \times \mathbb{I}_0 \to [0, 1]$ by requiring that $d(x, x) = 0$ for all $x \in \mathbb{I}_0$, $d(x, y) = d(y, x)$ for all $x, y \in \mathbb{I}_0$, and

$$\begin{aligned} d(x,y) := \limsup_{n\to\infty} \frac{1}{2n} \#\{1 \le k \le n : x \prec a_k \prec y\} \\ + \limsup_{n\to\infty} \frac{1}{2n} \#\{1 \le \ell \le n : x \prec b_\ell \prec y\} \end{aligned}$$

for $x \prec y$. It follows from exchangeability, de Finetti's theorem, and the strong law of large numbers that in the above the superior limits are actually limits almost surely.

Remark 6.4. It is clear that by redefining d on a $\mathbb{P}$-null set we may assume for every $x, y, z \in \mathbb{I}_0$ that

- $d(x, y) \ge 0$,
- $d(x, y) = d(y, x)$,
- $d(x, z) \le d(x, y) + d(y, z)$,
- $d(x, y) = 0$ if $x = y$.

Remark 6.5. For distinct $x, y, z \in \mathbb{I}_0$ the triangle inequality $d(x, z) \le d(x, y) + d(y, z)$ can be sharpened to a statement that for all x, y, z

- $d(x, z) = d(x, y) + d(y, z)$ if $x \prec y \prec z$,
- $d(x, z) = d(x, y) - d(y, z)$ if $x \prec z \prec y$,
- $d(x, z) = d(y, z) - d(x, y)$ if $y \prec x \prec z$,

and three analogous equalities when $z \prec x$.

Proposition 6.6. *If $x, y \in \mathbb{I}_0$ with $x \ne y$, then $d(x, y) > 0$ almost surely. Therefore almost surely d is a metric.*

Proof. We need to show for $k, \ell \in \mathbb{N}$ with $k \ne \ell$ that $d(a_k, a_\ell) > 0$ and $d(b_k, b_\ell) > 0$, and, furthermore, for arbitrary $k, \ell \in \mathbb{N}$ that $d(a_k, b_\ell) > 0$.

Consider $d(a_k, a_\ell)$. Set

$$I_m := \mathbb{1}(\{a_k \prec a_m \prec a_\ell\} \cup \{a_\ell \prec a_m \prec a_k\}), \quad m \notin \{k, \ell\}.$$

Suppose that Π_n, $n \in \mathbb{N}$ is a uniform random permutation of $[n]$. By exchangeability of the total order, if $k \vee \ell \le n$, then

$$\begin{aligned}
&\mathbb{P}\{I_m = 0,\, 1 \le m \le n,\, m \notin \{k,\ell\}\} \\
&\quad = \mathbb{P}(\{\Pi_n(\ell) = \Pi_n(k)+1\} \cup \{\Pi_n(k) = \Pi_n(\ell)+1\}) \\
&\quad = 2(n-1)\frac{1}{n(n-1)} \\
&\quad = \frac{2}{n}
\end{aligned}$$

and the random variables $\{I_m : m \in \mathbb{N},\, m \notin \{k,\ell\}\}$ are exchangeable. It follows from de Finetti's theorem and the strong law of large numbers that

$$\lim_{n\to\infty} \frac{1}{n} \#\{1 \le m \le n : a_k \prec a_m \prec a_\ell\} = \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} I_m > 0$$

almost surely and hence $d(a_k, a_\ell) > 0$. A similar argument shows that $d(b_k, b_\ell) > 0$.

It remains to show that $d(a_k, b_\ell) > 0$. Set $M := \{m \in \mathbb{N} : a_k \prec b_m\}$. It follows from exchangeability that on the event $\{M \neq \emptyset\} \supseteq \{a_k \prec b_\ell\}$ we have $\#M = \infty$ almost surely and indeed that $\lim_{n\to\infty} \frac{1}{n}\#(M \cap [n]) > 0$. Write $M = \{m_1, m_2, \ldots\}$ with $m_1 < m_2 < \ldots$. Fix $p \in \mathbb{N}$ and set

$$J_q := \mathbb{1}\{b_{m_q} \prec b_{m_p}\}, \quad q \neq p.$$

By exchangeability of the total order, if $p \vee q \le r$, then

$$\mathbb{P}\{J_q = 0,\, 1 \le q \le r,\, q \neq p \mid M \neq \emptyset\} = \mathbb{P}\{\Pi_r(p) = 1\} = \frac{1}{r}$$

and the random variables $\{J_q : q \in \mathbb{N},\, q \neq p\}$ are conditionally exchangeable given $\{M \neq \emptyset\}$. It follows from de Finetti's theorem that on the event $\{M \neq \emptyset\}$

$$\lim_{n\to\infty} \frac{1}{n} \#\{q : m_q \in [n],\, a_k \prec b_{m_q} \prec b_{m_p}\} > 0$$

almost surely and hence $d(a_k, b_\ell) > 0$ almost surely on the event $\{a_k \prec b_\ell\}$. A similar argument shows that $d(a_k, b_\ell) > 0$ almost surely on the event $\{b_\ell \prec a_k\}$. □

Definition 6.7. Given an ergodic exchangeable random total order $\prec$ on $\mathbb{I}_0$, denote by $\mathbb{I}$ the completion of $\mathbb{I}_0$ with respect to the metric d.

Definition 6.8. Define $f : \mathbb{I}_0 \to [0,1]$ by

$$f(y) := \sup\{d(x,y) : x \in \mathbb{I}_0,\, x \prec y\}.$$

Remark 6.9. It follows from Remark 6.5 that

$$f(y) = \limsup_{n\to\infty} \frac{1}{2n} \#\{1 \le k \le n : a_k \prec y\} + \limsup_{n\to\infty} \frac{1}{2n} \#\{1 \le \ell \le n : b_\ell \prec y\},$$

$$|f(x) - f(y)| = d(x,y), \quad x, y \in \mathbb{I}_0,$$

and

$$f(x) < f(y) \Longleftrightarrow x \prec y, \quad x, y \in \mathbb{I}_0,$$

so that f is an order-preserving isometry from $\mathbb{I}_0$ into $[0,1]$. Thus the function f extends by continuity to an isometry from $\mathbb{I}$ into $[0,1]$ and if $\prec$ is extended to $\mathbb{I}$ by declaring that $x \prec y \Longleftrightarrow f(x) < f(y)$, then $\prec$ is a total order on $\mathbb{I}$ and f is an order-preserving isometry from $\mathbb{I}$ into $[0,1]$ and hence an order-preserving isometric bijection from $\mathbb{I}$ to the image set $\mathbb{J} := f(\mathbb{I}) \subseteq [0,1]$. Because $\mathbb{I}$ is complete, $\mathbb{J}$ is complete. Because $\mathbb{J}$ is a complete subset of $[0,1]$ it is closed and hence compact, and therefore $\mathbb{I}$ itself is compact. It follows from the ergodicity of $\prec$ that $\mathbb{J}$ is almost surely constant. We will see below that $\mathbb{J} = [0,1]$.

Remark 6.10. Define a sequence $((X_n, Y_n))_{n\in\mathbb{N}}$ of $\mathbb{J}^2$-valued random variables by setting $X_n := f(a_n)$ and $Y_n := f(b_n)$. The exchangeability of $\prec$ implies that if σ and τ are two finite permutations of $\mathbb{N}$, then $((X_{\sigma(n)}, Y_{\tau(n)}))_{n\in\mathbb{N}}$ has the same distribution as $((X_n, Y_n))_{n\in\mathbb{N}}$. In particular, the sequence $((X_n, Y_n))_{n\in\mathbb{N}}$ is exchangeable. It is a consequence of de Finetti's theorem and the ergodicity of $\prec$ that this sequence is i.i.d. with common distribution some probability measure π on $\mathbb{J}^2$. It follows from the next result that $\pi = \mu \otimes \nu$ for two probability measures μ and ν on $\mathbb{J}$ that we call the *canonical pair*. Because $X_m \neq X_n$ and $Y_m \neq Y_n$ almost surely for $m \neq n$, the probability measures μ and ν must be diffuse.

Lemma 6.11. *Suppose that the random variables X', Y', X'', Y'' are such that*

(1) $(X', Y') \stackrel{d}{=} (X'', Y'')$
(2) $((X', Y'), (X'', Y'')) \stackrel{d}{=} ((X', Y''), (X'', Y'))$
(3) $(X', Y') \perp\!\!\!\perp (X'', Y'')$.

Then X', X'', Y', Y'' are independent.

Proof. For Borel sets A', A'', B', B'' we have

$$\begin{aligned}
&\mathbb{P}\{X' \in A',\, X'' \in A'',\, Y' \in B',\, Y'' \in B''\} \\
&= \mathbb{P}\{X' \in A',\, Y' \in B'\}\mathbb{P}\{X'' \in A'',\, Y'' \in B''\} && \text{by (3)} \\
&= \mathbb{P}\{X' \in A',\, Y'' \in B'\}\mathbb{P}\{X'' \in A'',\, Y' \in B''\} && \text{by (2)} \\
&= \mathbb{P}\{X' \in A'\}\mathbb{P}\{Y'' \in B'\}\mathbb{P}\{X'' \in A''\}\mathbb{P}\{Y' \in B''\} && \text{by (3)} \\
&= \mathbb{P}\{X' \in A'\}\mathbb{P}\{Y' \in B'\}\mathbb{P}\{X'' \in A''\}\mathbb{P}\{Y'' \in B''\} && \text{by (1).}
\end{aligned}$$

□

Theorem 6.12. *Any ergodic exchangeable random total order $\prec$ has the same distribution as one given by the construction in Remark 6.2 for some pair of diffuse probability measures (ζ, η) on $\mathbb{R}$. The canonical pair of diffuse probability measures (μ, ν) on $[0,1]$ is uniquely determined by the moment formulae*

$$\begin{aligned}
&\int_{[0,1]} x^n\, \mu(dx) \\
&\quad = \left(\frac{1}{2}\right)^n \sum_{c \in \prod_{k=1}^n \{a_k, b_k\}} \mathbb{P}\{c_1 \prec a_{n+1},\, \ldots,\, c_n \prec a_{n+1}\}
\end{aligned}$$

and

$$\begin{aligned}
&\int_{[0,1]} y^n\, \nu(dy) \\
&\quad = \left(\frac{1}{2}\right)^n \sum_{c \in \prod_{k=1}^n \{a_k, b_k\}} \mathbb{P}\{c_1 \prec b_{n+1},\, \ldots,\, c_n \prec b_{n+1}\}.
\end{aligned}$$

The probability measure $\frac{1}{2}(\mu + \nu)$ is Lebesgue measure on $[0,1]$ and, in particular, $\mathbb{J} = [0,1]$. Moreover, μ and ν are the respective push-forwards of ζ and η by the function $z \mapsto \frac{1}{2}(\zeta + \eta)((-\infty, z])$

Proof. We have already shown that an ergodic exchangeable random total order has the same distribution as one built from an arbitrary pair (ζ, η) of diffuse probability measures on $\mathbb{R}$ using the construction in Remark 6.2.

Define $((X_n, Y_n))_{n \in \mathbb{N}}$ as in Remark 6.10. It follows from Remark 6.9 that

$$X_n = \frac{1}{2}\mu((-\infty, X_n]) + \frac{1}{2}\nu((-\infty, X_n])$$

and

$$Y_n = \frac{1}{2}\mu((-\infty, Y_n]) + \frac{1}{2}\nu((-\infty, Y_n])$$

for any $n \in \mathbb{N}$. Let $(I_n)_{n\in\mathbb{N}}$ be a sequence of i.i.d. random variables that is independent of $((X_n, Y_n))_{n\in\mathbb{N}}$ with $\mathbb{P}\{I_n = 0\} = \mathbb{P}\{I_n = 1\} = \frac{1}{2}$ and set $Z_n := I_n X_n + (1 - I_n)Y_n$ so that the sequence $(Z_n)_{n\in\mathbb{N}}$ is i.i.d. with common distribution $\frac{1}{2}(\mu + \nu)$. We have

$$Z_n = \frac{1}{2}(\mu + \nu)((-\infty, Z_n]),$$

and so $\frac{1}{2}(\mu + \nu)$ is Lebesgue measure on $[0, 1]$. Thus, for any $n \in \mathbb{N}$

$$\begin{aligned}\int_{[0,1]} x^n \, \mu(dx) &= \mathbb{P}\{Z_1 < X_{n+1}, \, \ldots, \, Z_n < X_{n+1}\} \\ &= \left(\frac{1}{2}\right)^n \sum_{c \in \prod_{k=1}^n \{a_k, b_k\}} \mathbb{P}\{c_1 \prec a_{n+1}, \, \ldots, \, c_n \prec a_{n+1}\}\end{aligned}$$

and

$$\begin{aligned}\int_{[0,1]} y^n \, \nu(dy) &= \mathbb{P}\{Z_1 < Y_{n+1}, \, \ldots, \, Z_n < Y_{n+1}\} \\ &= \left(\frac{1}{2}\right)^n \sum_{c \in \prod_{k=1}^n \{a_k, b_k\}} \mathbb{P}\{c_1 \prec b_{n+1}, \, \ldots, \, c_n \prec b_{n+1}\},\end{aligned}$$

as claimed.

The proof of the final claim is straightforward and we omit it. □

Remark 6.13. We haven't shown that if $(y_k)_{k\in\mathbb{N}}$ is a sequence of points of $\mathbb{W}$, where $y_k \in \mathbb{W}_{N(y_k)}$, $N(y_k) \to \infty$ as $k \to \infty$, and $\lim_{k\to\infty} y_k = y$ in the Doob–Martin topology for some arbitrary y in the Doob–Martin boundary, then the harmonic function $K(\cdot, y)$ is extremal. This is equivalent to showing that if the infinite bridge $(U_n^\infty)_{n\in\mathbb{N}_0}$ is the limit of the bridges $(U_0^{y_k}, \ldots, U_{N(y_k)}^{y_k})$, then $(U_n^\infty)_{n\in\mathbb{N}_0}$ has an almost surely trivial tail σ-field. This is, in turn, equivalent to showing that the corresponding labeled infinite bridge induces an ergodic exchangeable random order. The latter, however, can be established along the lines of [EGW15, Corollary 5.21] and [EW16, Corollary 7.2], so we omit the details.

7. Identification of extremal harmonic functions

Any extremal infinite bridge $(U_n^\infty)_{n\in\mathbb{N}_0}$ is the h-transform of our original Markov chain with an extreme harmonic function h. We know from the above that such a process arises as follows in terms of the canonical pair (μ, ν) of diffuse probability measures associated with the corresponding point in the Doob–Martin boundary.

We first require some notation. Given $(x_1, \dots, x_n, y_1, \dots, y_n) \in \mathbb{R}^{2n}$ with distinct entries, let $z_1 < \dots < z_{2n}$ be a listing of $\{x_1, \dots, x_n, y_1, \dots, y_n\}$ in increasing order. Define

$$\mathcal{W}((x_1, \dots, x_n, y_1, \dots, y_n)) = u_1 \dots u_{2n} \in \mathbb{W}_n$$

by

$$u_i = \begin{cases} a, & \text{if } z_i \in \{x_1, \dots, x_n\}, \\ b, & \text{if } z_i \in \{y_1, \dots, y_n\}. \end{cases}$$

Given $v \in \mathbb{W}_n$, set

$$\mathcal{S}(v) := \mathcal{W}^{-1}(\{v\}) \subset \mathbb{R}^{2n}.$$

For example,

$$\mathcal{S}(abba) = \bigsqcup_{\sigma,\tau} \{(x_1, x_2, y_1, y_2) \in \mathbb{R}^4 : x_{\sigma(1)} < y_{\tau(1)} < y_{\tau(2)} < x_{\sigma(2)}\},$$

where the union is over all pairs of permutations σ, τ of the set $\{1, 2\}$. In general, $\mathcal{S}(v)$ is the disjoint union of $(n!)^2$ connected open sets that all have boundaries of zero Lebesgue measure.

Now take independent sequences of real-valued random variables $(X_k)_{k \in \mathbb{N}}$ and $(Y_k)_{k \in \mathbb{N}}$, where the X_k are i.i.d. with common distribution μ and the Y_k are i.i.d. with common distribution ν and set

$$U_n^\infty = \mathcal{W}((X_1, \dots, X_n, Y_1, \dots, Y_n)).$$

We have

$$\mathbb{P}\{U_n^\infty = u\} = \mu^{\otimes n} \otimes \nu^{\otimes n}(\mathcal{S}(u))$$

We also know that

$$\mathbb{P}\{U_n^\infty = u \,|\, U_{n+1}^\infty = v\} = \frac{\binom{v}{u}}{(n+1)^2}.$$

It follows that

$$\begin{aligned} &\mathbb{P}\{U_{n+1}^\infty = v \,|\, U_n^\infty = u\} \\ &\quad = \mu^{\otimes(n+1)} \otimes \nu^{\otimes(n+1)}(\mathcal{S}(v)) \frac{\binom{v}{u}}{(n+1)^2} \Big/ \mu^{\otimes n} \otimes \nu^{\otimes n}(\mathcal{S}(u)). \end{aligned}$$

On the other hand,

$$\begin{aligned} \mathbb{P}\{U_{n+1}^\infty = v \,|\, U_n^\infty = u\} &= \frac{1}{h(u)} \mathbb{P}\{U_{n+1} = v \,|\, U_n = u\} h(v) \\ &= \frac{h(v)}{h(u)} \frac{\binom{v}{u}}{(2n+2)(2n+1)}. \end{aligned}$$

Thus,

$$\frac{h(v)}{h(u)} = \frac{\mu^{\otimes(n+1)} \otimes \nu^{\otimes(n+1)}(\mathcal{S}(v))}{\mu^{\otimes n} \otimes \nu^{\otimes n}(\mathcal{S}(u))} \frac{(2n+2)(2n+1)}{(n+1)^2}$$

and, up to an arbitrary multiplicative constant,

$$h(w) = \binom{2m}{m} \mu^{\otimes m} \otimes \nu^{\otimes m}(\mathcal{S}(w))$$

for $w \in \mathbb{W}_m$.

Since $h(\emptyset) = 1$, this normalization is the extended Doob–Martin kernel $w \mapsto K(w, y)$, where y is the point in the Doob–Martin boundary that corresponds to the pair of diffuse probability measures (μ, ν).

Remark 7.1. The constant harmonic function $h \equiv 1$ arises from the above construction with μ and ν both being the Lebesgue measure λ on $[0, 1]$. Therefore the process $(U_n)_{n\in\mathbb{N}_0}$ is itself the extremal bridge associated with the canonical pair (λ, λ). In particular, $(U_n)_{n\in\mathbb{N}_0}$ converges almost surely to the point in the Doob–Martin boundary associated with this pair.

We observed in Remark 3.7 that a sequence $(y_k)_{k\in\mathbb{N}}$ with $N(y_k) \to \infty$ as $k \to \infty$ converges in the Doob–Martin topology if and only if for every $m \in \mathbb{N}$ the sequence of random words in $\mathbb{W}_m$ obtained by selecting m letters a and m letters b uniformly at random from y_k and maintaining their relative order converges in distribution as $k \to \infty$. We can now enhance that result as follows.

Proposition 7.2. *Consider a sequence $(y_k)_{k\in\mathbb{N}}$ in $\mathbb{W}$, where $y_k \in \mathbb{W}_{N(y_k)}$, $k \in \mathbb{N}$, and $N(y_k) \to \infty$ as $k \to \infty$. If y is the point in the Doob–Martin boundary that corresponds to the pair of (diffuse) probability measures (μ, ν) with $\frac{1}{2}(\mu + \nu) = \lambda$, then $\lim_{k\to\infty} y_k = y$ in the Doob–Martin topology if and only if*

$$\lim_{k\to\infty} \frac{\binom{y_k}{w}}{\binom{N(y_k)}{m}^2} = \mu^{\otimes m} \otimes \nu^{\otimes m}(\mathcal{S}(w))$$

for all $w \in \mathbb{W}_m$ for all $m \in \mathbb{N}$. That is, $\lim_{k\to\infty} y_k = y$ if and only if for each $m \in \mathbb{N}$ the sequence of random words in $\mathbb{W}_m$ obtained by selecting m letters a and m letters b uniformly at random from y_k and maintaining their relative order converges in distribution as $k \to \infty$ to the random word $U_m^\infty = \mathcal{W}(X_1, \ldots, X_m, Y_1, \ldots, Y_m)$ defined above.

Given a sequence $(y_k)_{k\in\mathbb{N}}$ in $\mathbb{W}$, where $y_k \in \mathbb{W}_{N(y_k)}$, $k \in \mathbb{N}$, and $N(y_k) \to \infty$ as $k \to \infty$, define a sequence of pairs of discrete probability measures $((\mu_k, \nu_k))_{k\in\mathbb{N}}$ on $[0, 1]$ as follows. For $k \in \mathbb{N}$ the two probability measure μ_k and ν_k both assign all of their mass to the set

$\{\frac{\ell}{2N(y_k)} : 1 \le \ell \le 2N(y_k)\}$. For $1 \le i \le 2N(y_k)$, $\mu_k(\frac{i}{2N(y_k)}) = \frac{1}{N(y_k)}$ if the i^{th} letter of y_k is the letter a, otherwise $\mu_k(\frac{i}{2N(y_k)}) = 0$. Similarly, for $1 \le j \le 2N(y_k)$, $\nu_k(\frac{j}{2N(y_k)}) = \frac{1}{N(y_k)}$ if the j^{th} letter of y_k is the letter b, otherwise $\nu_k(\frac{j}{2N(y_k)}) = 0$. In particular, $\frac{1}{2}(\mu_k + \nu_k)$ is the uniform probability measure on $\{\frac{\ell}{2N(y_k)} : 1 \le \ell \le 2N(y_k)\}$. Observe that if $w \in \mathbb{W}_m$, then, for $w \in \mathbb{W}_m$,

$$(N(y_k)^m)^2 \mu_k^{\otimes m} \otimes \nu_k^{\otimes m}(\mathcal{S}(w)) = (m!)^2 \binom{y_k}{w}$$

so that

$$\frac{\binom{y_k}{w}}{\binom{N(y_k)}{m}^2} = \left(\frac{N(y_k)^m}{N(y_k)(N(y_k)-1)\cdots(N(y_k)-m+1)}\right)^2 \mu_k^{\otimes m} \otimes \nu_k^{\otimes m}(\mathcal{S}(w)).$$

One direction of the following corollary is now immediate.

Corollary 7.3. *Suppose that $(y_k)_{k\in\mathbb{N}}$ and $((\mu_k, \nu_k))_{k\in\mathbb{N}}$ are as above. If $(y_k)_{k\in\mathbb{N}}$ converges in the Doob–Martin topology to the point y in the Doob–Martin boundary that corresponds to the pair of probability measures (μ, ν), then $(\mu_k)_{k\in\mathbb{N}}$ converges weakly to μ and $(\nu_k)_{k\in\mathbb{N}}$ converges weakly to ν. Conversely, if $(\mu_k)_{k\in\mathbb{N}}$ converges weakly to μ and $(\nu_k)_{k\in\mathbb{N}}$ converges weakly to ν, then $\frac{1}{2}(\mu + \nu) = \lambda$, and if y is the point in the Doob–Martin boundary that corresponds to the pair (μ, ν), then $(y_k)_{k\in\mathbb{N}}$ converges in the Doob–Martin topology to y.*

Proof. As we have already remarked, if $(\mu_k)_{k\in\mathbb{N}}$ converges weakly to μ and $(\nu_k)_{k\in\mathbb{N}}$ converges weakly to ν then, since the boundary of $\mathcal{S}(w)$ is Lebesgue null for any word $w \in \mathbb{W}_m$, $m \in \mathbb{N}$, we have that $\mu_k^{\otimes m} \otimes \nu_k^{\otimes m}(\mathcal{S}(w))$ converges to $\mu^{\otimes m} \otimes \nu^{\otimes m}(\mathcal{S}(w))$ so that

$$\lim_{k\to\infty} \frac{\binom{y_k}{w}}{\binom{N(y_k)}{m}^2} = \mu^{\otimes m} \otimes \nu^{\otimes m}(\mathcal{S}(w)),$$

and it follows from Proposition 7.2 that $(y_k)_{k\in\mathbb{N}}$ converges to the point y in the Doob–Martin boundary that corresponds to the pair (μ, ν).

Conversely, suppose that $(y_k)_{k\in\mathbb{N}}$ converges to the point y in the Doob–Martin boundary corresponding to the pair (μ, ν). Given any subsequence of $\mathbb{N}$ there is, by the compactness in the weak topology of probability measures on $[0, 1]$, a further subsequence such that along this further subsequence μ_k converges weakly to some probability measure μ' and ν_k converges weakly to some probability measure ν'. Note that $\frac{1}{2}(\mu' + \nu') = \lambda$. From the other direection of the corollary, this implies that along the subsubsequence y_k converges to the point y' in the Doob–Martin boundary corresponding to the pair of probability measures (μ', ν'). Because $y' = y$ it must be the case $(\mu', \nu') = (\mu, \nu)$. Thus, from any subsequence of $\mathbb{N}$ we can extract a further subsequence along which μ_k converges weakly to μ and ν_k converges weakly to ν, and this implies that $(\mu_k)_{k\in\mathbb{N}}$ converges weakly to μ and $(\nu_k)_{k\in\mathbb{N}}$ converges weakly to ν. □

8. Example: the Plackett-Luce chain

In general, there is no simple closed form expression for the transition probabilities of an infinite bridge $(U_n^\infty)_{n\in\mathbb{N}_0}$ associated with a pair of (not necessarily canonical) diffuse probability measures ζ, η and hence the associated harmonic function h. However, it is possible to obtain such expressions in the special case where ζ is the exponential distribution with rate parameter α and η is the exponential distribution with rate parameter β. Given $u \in \mathbb{W}_n$ and $1 \le i \le 2n$, set

$$\mathbf{A}_i^n(u) := \#\{i \le j \le 2n : u_j = a\}$$

and

$$\mathbf{B}_i^n(u) := \#\{i \le j \le 2n : u_j = b\}.$$

By the reasoning that goes into the analysis of the Plackett-Luce or vase model of random permutations (see, for example, [Mar95]),

$$\mathbb{P}\{U_n^\infty = u\} = (n!)^2 \alpha^n \beta^n \prod_{i=1}^{2n} \frac{1}{\mathbf{A}_i^n(u)\alpha + \mathbf{B}_i^n(u)\beta}$$

– this is essentially just repeated applications of the elementary result usually called *competing exponentials*: if S and T are independent exponentially distributed random variables with rate parameters λ and θ, then the probability of the event $\{S < T\}$ is $\frac{\lambda}{\lambda+\theta}$ and conditional on this event the random variables S and $T - S$ are independent and exponentially distributed with rate parameters $\lambda + \theta$ and θ. (As a check, note that when $\alpha = \beta = \gamma$, say, this probability is, as expected, $1/\binom{2n}{n}$.) We also know that

$$\mathbb{P}\{U_n^\infty = u \,|\, U_{n+1}^\infty = v\} = \frac{\binom{v}{u}}{(n+1)^2}.$$

It follows that

$$\begin{aligned}
&\mathbb{P}\{U_{n+1}^\infty = v \,|\, U_n^\infty = u\} \\
&= \frac{\binom{v}{u}}{(n+1)^2}((n+1)!)^2\alpha^{n+1}\beta^{n+1}\prod_{i=1}^{2(n+1)}\frac{1}{\mathbf{A}_i^{n+1}(v)\alpha+\mathbf{B}_i^{n+1}(v)\beta} \\
&\quad \Big/ (n!)^2\alpha^n\beta^n\prod_{i=1}^{2n}\frac{1}{\mathbf{A}_i^n(u)\alpha+\mathbf{B}_i^n(u)\beta} \\
&= \binom{v}{u}\alpha\beta\frac{\prod_{i=1}^{2n}(\mathbf{A}_i^n(u)\alpha+\mathbf{B}_i^n(u)\beta)}{\prod_{i=1}^{2(n+1)}(\mathbf{A}_i^{n+1}(v)\alpha+\mathbf{B}_i^{n+1}(v)\beta)}.
\end{aligned}$$

As a check, when $\alpha = \beta = \gamma$, say, this transition probability is

$$\binom{v}{u}\frac{(2n)!}{(2(n+1))!} = \frac{\binom{v}{u}}{(2n+2)(2n+1)},$$

as expected.

The corresponding harmonic function h satisfies

$$\begin{aligned}
&\binom{v}{u}\alpha\beta\frac{\prod_{i=1}^{2n}(\mathbf{A}_i^n(u)\alpha+\mathbf{B}_i^n(u)\beta)}{\prod_{i=1}^{2(n+1)}(\mathbf{A}_i^{n+1}(v)\alpha+\mathbf{B}_i^{n+1}(v)\beta)} \\
&\qquad = \frac{h(v)}{h(u)}\frac{\binom{v}{u}}{(2n+2)(2n+1)}.
\end{aligned}$$

We conclude from this that, up to an arbitrary positive constant,

$$h(w) = \frac{(2m)!\alpha^m\beta^m}{\prod_{i=1}^{2m}(\mathbf{A}_i^m(w)\alpha+\mathbf{B}_i^m(w)\beta)}$$

for $w \in \mathbb{W}_n$.

Acknowledgments. We thank two anonymous referees for a number of helpful suggestions that improved the presentation considerably.

References

[Cla15] Anders Claesson, *Subword counting and the incidence algebra*, 2015, arXiv:1502.03065v3.

[Doo59] J. L. Doob, *Discrete potential theory and boundaries*, J. Math. Mech. **8** (1959), 433–458; erratum 993. MR 0107098 (21 #5825)

[EGW15] Steven N. Evans, Rudolf Grübel, and Anton Wakolbinger, *Doob–Martin boundary of Rémy's tree growth chain*, 2015, To appear in *Annals of Probability*. Available at arXiv:1411.2526 [math.PR].

[EW16] Steven N. Evans and Anton Wakolbinger, *Radix sort trees in the large*, 2016, Available at arXiv:1603.07385 [math.PR].

[Föl75] Hans Föllmer, *Phase transition and Martin boundary*, Séminaire de Probabilités, IX (Seconde Partie, Univ. Strasbourg, Strasbourg, années universitaires 1973/1974 et 1974/1975), Springer, Berlin, 1975, pp. 305–317. Lecture Notes in Math., Vol. 465. MR 0426176 (54 #14122)

[Kal02] Olav Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR 1876169 (2002m:60002)

[KSK76] John G. Kemeny, J. Laurie Snell, and Anthony W. Knapp, *Denumerable Markov chains*, second ed., Springer-Verlag, New York, 1976, With a chapter on Markov random fields, by David Griffeath, Graduate Texts in Mathematics, No. 40. MR 0407981 (53 #11748)

[Lot97] M. Lothaire, *Combinatorics on words*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1997, With a foreword by Roger Lyndon and a preface by Dominique Perrin, Corrected reprint of the 1983 original, with a new preface by Perrin. MR 1475463 (98g:68134)

[Mar95] John I. Marden, *Analyzing and modeling rank data*, Monographs on Statistics and Applied Probability, vol. 64, Chapman & Hall, London, 1995. MR 1346107 (96j:62127)

[Rev75] D. Revuz, *Markov chains*, North-Holland Publishing Co., Amsterdam, 1975, North-Holland Mathematical Library, Vol. 11. MR 0415773 (54 #3852)

[RW00] L. C. G. Rogers and David Williams, *Diffusions, Markov processes, and martingales. Vol. 1*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2000, Foundations, Reprint of the second (1994) edition. MR 1796539 (2001g:60188)

[Saw97] Stanley A. Sawyer, *Martin boundaries and random walks*, Harmonic functions on trees and buildings (New York, 1995), Contemp. Math., vol. 206, Amer. Math. Soc., Providence, RI, 1997, pp. 17–44. MR 1463727 (98k:60127)

[Woe00] Wolfgang Woess, *Random walks on infinite graphs and groups*, Cambridge Tracts in Mathematics, vol. 138, Cambridge University Press, Cambridge, 2000. MR 1743100 (2001k:60006)

E-mail address: hyeso0oa@gmail.com

Department of Statistics #3860, 367 Evans Hall, University of California, Berkeley, CA 94720-3860, USA

E-mail address: evans@stat.berkeley.edu