On a conjecture of David Aldous: river network and the continuum random tree

Kumarjit Saha∗†

June 13, 2017

Abstract

Consider the graph with vertex set \( \mathbb{Z}_2 \) even := \{ (x, t) : x + t even \} and a random edge set \( E \) which, for each \( (x, t) \in \mathbb{Z}_2 \) even, contains either the edge \( \langle (x, t), (x + 1, t + 1) \rangle \) or \( \langle (x, t), (x - 1, t + 1) \rangle \) with equal probability and independent of other vertices in \( \mathbb{Z}_2 \) even. This graph may be viewed as a graphical representation of a system of 1-dimensional coalescing simple symmetric random walks starting at every space time point \( (x, t) \in \mathbb{Z}_2 \) even. Let \( T(0, 0) \) be the subtree obtained from the vertices \( \{ (x, t) \in \mathbb{Z}_2 \) even : t ≤ 0 \} and their associated edges lying in the connected component of the origin \( (0, 0) \) of this random graph. With the usual graph distance, this gives a tree-like metric space. For this subtree \( T(0, 0) \), Aldous [2] conjectured that conditioned on the event that the farthest leaf of this tree is at a graph distance \( n \), the scaled tree obtained by scaling the distances by \( 1/n \) converges in distribution to a continuum random tree (CRT) in Gromov-Hausdorff topology. We prove the conjecture conditioned on the event that the distance of the farthest leaf is at least \( n \). The limiting continuum random tree is slightly different from what was surmised in [2]. For each \( n \geq 1 \), we define a scaled ‘dual’ tree and show that under a diffusive scaling the original tree and the dual tree converges in distribution to a pair of continuum random trees. Both these limiting continuum trees appear to be new. We expect that the same limit holds for other drainage network models also in the basin of attraction of the Brownian web.

Keywords: Continuum random tree, Gromov-Hausdorff distance, coalescing random walk, Brownian web.

AMS Classification: 60D05

1 Introduction

Trees are often used in representing genealogical structures, communication network modelling, optimisation and river basin study. During the last three decades there has been an interest in understanding the scaling limits of various discrete random trees. Most notably in [1], Aldous showed that an appropriately scaled Galton-Watson tree with a finite variance critical offspring distribution and conditioned to have the total population size \( n \), as \( n \to \infty \), converges to a continuum random tree (CRT) famously known as the Brownian CRT. In the recent years, these results have been extended for more general continuum random trees (e.g., see [7], [9], [21]). These concepts and results have found many other applications also, e.g., understanding the sequence of components of critical Erdos-Renyi random graph ([3], [6]). Further this limiting behaviour appear to be universal in the sense that Aldous’ limiting picture for critical random graph has since been extended to “immigration” models of random graphs [4], hypergraphs [14] and to random graphs with fixed degree [23].

In [2], Aldous conjectured that a sequence of appropriately scaled discrete trees obtained from a drainage network model should converge to a continuum random tree (see Subsection 4.2 Page 275 of

∗TIFR Center for Applicable Mathematics, Bangalore.
†This work is benefited from the support of the AIRBUS Group Corporate Foundation Chair in Mathematics of Complex Systems established in TIFR CAM and supported in part by ISF-UGC grant.
In this paper we prove this conjecture under the “non-degenerate” conditional setting and obtain a slightly modified limit than the one conjectured in [2]. To the best of our knowledge, in the context of random real trees, this limiting continuum random tree has not been studied in the literature earlier. A slightly related random tree has been studied in [15]. In Section 5 we mention about the work in [15] in more detail. We first present the conjecture in [2] and for this purpose, below we describe the drainage network model introduced by Scheidegger [27] formally.

Let \( \mathbb{Z}_{even}^2 := \{(x, t) : x, t \in \mathbb{Z}, x + t \text{ even}\} \) be the oriented lattice. We construct system of coalescing simple symmetric random walks starting from every lattice point as follows. For each \((x, t) \in \mathbb{Z}_{even}^2\), the walker at spatial location \(x, t\) moves to location \((x + 1)\) or \((x - 1)\) at time \(t + 1\) with equal probability. Formally, consider \(\mathcal{B}_{(x,t)} := \{b(x,t) : (x,t) \in \mathbb{Z}_{even}^2\}\), a collection of i.i.d. random variables such that

\[
\begin{align*}
\mathbb{P}(b(x,t) = +1) &= \frac{1}{2}, \\
\mathbb{P}(b(x,t) = -1) &= \frac{1}{2}.
\end{align*}
\]

For \((x, t) \in \mathbb{Z}_{even}^2\), let \(h(x, t) := (x + b(x,t), t + 1)\) and for \(k \geq 1\), let \(h^k(x, t) := h(h^{k-1}(x, t))\) with \(h(x, t) = (x, t)\). For technical reasons for a general point \((y, s)\) in \(\mathbb{R}^2\), let \((y, s)(i)\) represent the \(i\)-th co-ordinate of \((y, s)\) for \(i = 1, 2\), i.e., \((y, s)(1) := y\) and \((y, s)(2) := s\). For \((x, t) \in \mathbb{Z}_{even}^2\), we observe that the process \(\{h^k(x, t)(1) : k \geq 0\}\) is a one-dimensional simple symmetric random walk starting from location \(x\) at time \(t\). Consider the random graph \(G\) formed by the vertex set \(\mathbb{Z}_{even}^2\) and the edge set \(E := \{((x, t), h(x, t)) : (x, t) \in \mathbb{Z}_{even}^2\}\). Every vertex in \(\mathbb{Z}_{even}^2\) represents a water source. From each source vertex \((x, t)\), water flows to another source \(h(x, t)\) along the channel \(\langle(x, t), h(x, t)\rangle\).

For a source vertex \((x, t) \in \mathbb{Z}_{even}^2\), we now define the cluster or the river delta at \((x, t)\), consisting of all the source vertices whose water flow through \((x, t)\), as

\[
C(x, t) := \{(y, s) \in \mathbb{Z}_{even}^2 : h^l(y, s) = (x, t) \text{ for some } l \geq 0\}.
\]

The set of edges in cluster \(C(x, t)\) is denoted by \(E(x, t) := \{\langle(y, s), h(y, s)\rangle : (y, s) \in C(x, t), (y, s) \neq (x, t)\}\). The ‘depth’ of the cluster \(C(x, t)\), i.e., the time length that the cluster survived, is defined as

\[
L(x, t) := \max\{l \geq 0 : h^l(y, s) = (x, t) \text{ for some } (y, s) \in C(x, t)\}.
\]

For any set \(A\), we define \(\#A\) to be the cardinality of \(A\). We shall explain later that both the random variables, \(\#C(x, t)\) and \(L(x, t)\), are finite almost surely.

Corresponding to the random graph defined above, we now define the weight function \(1\) as a constant mapping \(1 : E(x, t) \to \{1\}\).

\[
T(x, t) := (C(x, t), E(x, t), 1)
\]

denote the discrete weighted tree formed by the cluster \(C(x, t)\), the edge set \(E(x, t)\) and the weight function \(1\). This gives a rooted tree with root \((x, t)\) (see Figure 1). Let \(T_n(x, t) := (C(x, t), E(x, t), 1_n)\) denote the scaled tree where \(1_n(e) = 1/n\) for all \(e \in E(x, t)\). For each \(n \geq 1\), the random object \(T_n(x, t)\) can be regarded as a random metric space where the distance between any two vertices is given by the sum of the edge weights along the unique path between them. Note that the distribution of \(T_n(x, t)\) does not depend on the vertex \((x, t) \in \mathbb{Z}_{even}^2\).

Albous’ conjecture in [2] can then be stated as:

**Conjecture 1.1.** The conditional distribution of \(T_n(0, 0)\) given that \(\{L(0, 0) = n\}\), converges in distribution to a continuum random tree as \(n \to \infty\).

The main result of this paper proves the above conjecture in the non-degenerate conditional setting \(\{L(0, 0) \geq n\}\). The limiting continuum random tree is slightly different than that was surmised in [2] (see Section 2). In the non-degenerate conditional setting, we prove a stronger version of the above conjecture in the sense that for each \(n \geq 1\), we construct a “dual” tree \(\hat{T}_n(0, 0)\), and show that as a sequence of pair of metric spaces, the pair \((T_n(0, 0), \hat{T}_n(0, 0))\) \(\{L(0, 0) \geq n\}\) jointly converges in distribution to a pair of continuum random trees. In order to describe \(\hat{T}(x, t)\), following Arratia [3] we construct the
dual graph. Before giving a formal definition of the dual graph, we direct the attention of the reader to Figure 1. Let $\hat{h}(y,s) := (y - b(y,s-1), s-1)$. The dual graph $\hat{G}$ is the random graph with the vertex set $\mathbb{Z}^2_{\text{odd}}$ and the edge set $\hat{E} := \{(y,s), \hat{h}(y,s) : (y,s) \in \hat{V}\}$. For $k \geq 1$ let $\hat{h}^k(y,s) := \hat{h}(\hat{h}^{k-1}(y,s))$ with $\hat{h}^0(y,s) = (y,s)$.

Each $(y,s) \in \mathbb{Z}^2_{\text{odd}}$ is a ‘dual’ source vertex and the ‘dual’ flow starting from $(y,s)$ goes to $\hat{h}^{(y,s)}$ through the dual edge $(y,s), \hat{h}(y,s))$. For $(x,t) \in \mathbb{Z}^2_{\text{even}}$, let $\hat{t}_{(x,t)} := \inf\{l \in \mathbb{N} : \hat{h}^l(x-1,t) = \hat{h}^l(x+1,t)\}$. Since both the processes, $\{\hat{h}^l(x-1,t)(1) : l \geq 0\}$ and $\{\hat{h}^l(x+1,t)(1) : l \geq 0\}$, are independent simple symmetric random walks till the time they meet, the random variable $\hat{t}_{(x,t)}$ is finite almost surely.

**Remark 1.2.** Before we proceed, we list some properties of the graph $G$ and its dual $\hat{G}$.

1. $G$ uniquely specifies the dual graph $\hat{G}$ and the dual edges do not intersect the original edges. The construction also ensures that $\hat{G}$ does not contain any circuit.

2. $G$ is a single tree, i.e., the paths starting from any two vertices in $\mathbb{Z}^2_{\text{even}}$ coalesce and there is no bi-infinite path in $G$.

3. For $(x,t) \in \mathbb{Z}^2_{\text{even}}$, the cluster $C(x,t)$ is enclosed within the dual paths starting from the right dual neighbour $(x+1,t)$ and the left dual neighbour $(x-1,t)$. The boundedness of $C(x,t)$ for every $(x,t) \in \mathbb{Z}^2_{\text{even}}$ implies that these two dual paths coalesce, thus $\hat{G}$ is a single tree.

4. Since paths starting from any two vertices in $G$ coalesce and the dual edges do not cross the original edges, there is no bi-infinite path in $\hat{G}$.

Having described the dual graph, we now construct the “dual” tree $\hat{T}(x,t)$. For $(x,t) \in \mathbb{Z}^2_{\text{even}}$, the dual cluster $\hat{C}(x,t)$ and its dual edge set $\hat{E}(x,t)$ are defined as follows:

$$
\hat{C}(x,t) := \{(y,t-s) \in \mathbb{Z}^2_{\text{odd}} : 0 \leq s \leq \hat{t}_{(x,t)}, \hat{h}^s(x-1,t)(1) \leq y \leq \hat{h}^s(x+1,t)(1)\} \\
\hat{E}(x,t) := \{(y,s), \hat{h}(y,s)) : (y,s) \in \hat{C}(x,t), (y,s) \neq \hat{h}^{(y,s)}(x-1,t)\}.
$$

Figure 1: Coalescing simple symmetric random walk on $\mathbb{Z}^2_{\text{even}}$ and its dual. The gray shaded region represents the rooted tree $T(x,t)$ with root at $(x,t)$. The bold blue shaded region represents $\hat{T}(x,t)$. 
The constant weight function $1$ attaches weight $1$ to each edge in $\hat{E}(x,t)$. Our dual tree $\hat{T}(x,t)$ is denoted by $\hat{T}(x,t) := (\hat{C}(x,t), \hat{E}(x,t), 1)$. For $n \geq 1$, the dual tree scaled by $1/n$ is given by

$$\hat{T}_n(x,t) := (\hat{C}(x,t), \hat{E}(x,t), 1_n)$$

where $1_n(e) = 1/n$ for all $e \in \hat{E}(x,t)$.

For each $(x,t) \in Z^2_{\text{even}}$, $\{\hat{T}_n(x,t) : n \in \mathbb{N}\}$ gives a sequence of random metric spaces. For each $n \geq 1$, the distribution of $(T_n(x,t), \hat{T}_n(x,t))|\{L(x,t) > n\}$ does not depend on the vertex $(x,t) \in Z^2_{\text{even}}$. We are interested in the weak limit of $(T_n(0,0), \hat{T}_n(0,0))|\{L(0,0) > n\}$ as $n \to \infty$, where limit is taken in the Gromov-Hausdorff topology.

Gromov-Hausdorff topology is a common way to define a topology (even a metric) on a space of compact metric spaces. This had been introduced by Gromov [16]. For any two compact subsets $K$ and $K'$ of a metric space $(X,d)$, the Hausdorff distance $d_H(K,K')$ is given by:

$$d_H(K,K') := \inf\{\epsilon > 0 : K \subseteq U_{\epsilon}(K') \text{ and } K' \subseteq U_{\epsilon}(K)\}$$

where for $\epsilon > 0$, the set $K_{\epsilon}$ is defined as $K_{\epsilon} := \{x \in X : d(x,K) \leq \epsilon\}$. The Gromov-Hausdorff distance between two compact metric spaces $(X,d)$ and $(X',d')$ is defined by

$$d_{GH}((X,d),(X',d')) := \inf\{d_H(\phi(X),\phi'(X'))\}$$

where the infimum is taken over all possible choices of isometric embeddings $\phi, \phi'$ of the metric spaces $(X,d)$ and $(X',d')$ into a common metric space. Let $\mathcal{M}$ be the set of all isometry equivalence classes of compact metric spaces endowed with the Gromov-Hausdorff metric. It is known that this distance turns $\mathcal{M}$ into a Polish space. We mention here that given a metric space $(X,d)$, we write $[X,d]$ for the isometry equivalence class of $(X,d)$, and frequently use the notation $X$ for either $(X,d)$ or $[X,d]$ when there is no risk of ambiguity.

We show that as $n \to \infty$, the conditional distribution of $(T_n(0,0), \hat{T}_n(0,0))$ given $\{L(0,0) > n\}$ converges to a pair of continuum random trees, which we denote by $(\mathcal{T}, \hat{\mathcal{T}})$. Throughout this paper the notation $\Rightarrow$ is used to denote convergence in distribution.

**Theorem 1.3. As $n \to \infty$, we have**

$$(T_n(0,0), \hat{T}_n(0,0))|\{L(0,0) \geq n\} \Rightarrow (\mathcal{T}, \hat{\mathcal{T}}),$$

where convergence in distribution holds with respect to the Gromov-Hausdorff topology on $\mathcal{M}$.

In the next section, we construct these two continuum trees, $\mathcal{T}$ and $\hat{\mathcal{T}}$ and explain the difference of $\mathcal{T}$ from that conjectured in [22].

This work has at least one other motivation. It has been empirically observed that river networks are structurally self similar and satisfy various scaling laws. The study of these laws and understanding the reasons behind their existence are at the core of hydrology. Horton’s law for river network is a scaling relation which represents the scale invariance of the natural dendritic structure. For a binary tree this law can be described using pruning operations. Pruning of a finite rooted full binary tree $T$ cuts its leaves (vertices of degree one) and their parental edges, and removes the resulting chains of degree-two vertices and their parental edges. The Horton-Strahler order $k$ of a vertex $v$ is the minimal number of prunings necessary to eliminate the subtree rooted at $v$. A branch is a sequence of neighbouring vertices of the same order. Let $N_k$ denote the total number of branches of order $k$ in a tree. While studying river streams, Horton [17] observed that $N_{k+1}/N_k \approx R$, $3 \leq R \leq 5$. This regularity has been strongly corroborated in hydrology [29, 19, 33, 24] and referred as Horton’s law. The only rigorous result on validity of this law for drainage networks was that of Shreve [30], who demonstrated that for a uniform distribution of rooted binary trees with $n$ leaves, the ratio $N_{k+1}/N_k$ converges to $4$ as $n \to \infty$. Studying such relation involves the complete understanding of the complex branching structure. Since $T(0,0)$, the discrete tree obtained from the river delta $C(0,0)$, contains all branching information of the channels which flow through (0, 0), it is hoped that finding this scaling limit may help in understanding Horton’s law for Scheidegger network.
Contour functions have been used in (see, e.g., [1], [22]) to obtain limiting distributions of scaled trees, however we use the fact that that under diffusive scaling the network and its dual (observed as collections of forward and backward paths respectively) jointly converge to the Brownian web and its dual in the path-space topology (for more details see Subsection 3.1). Using this joint convergence, we show that the coalescing time of the scaled paths also converge to the coalescing time of the limiting Brownian paths. This helps us in obtaining the metric convergence.

Brownian web appears as a universal scaling limit for various drainage network models (see [11], [10], [23]). It is reasonable to expect that with suitable modifications our method will give the same scaling limit for other network models also. Our result will hold for any network model which admits a dual and satisfies (i) conditions listed in Remark 1.2, (ii) the scaled model and its dual converge in distribution to the Brownian web and its dual. In fact it was proved in [20] that the network model studied in [10] satisfies conditions (i) and (ii) and hence the same scaling limit holds. In this sense our result can be regarded as a universality class result.

This paper is organized as follows. In the next section we construct the limiting continuum random trees \( \mathcal{T} \) and \( \mathcal{T}^* \) and explain the difference of \( \mathcal{T} \) than the one conjectured in [2]. In Section 3 we introduce the Brownian web and its dual and use them to obtain compactness and other properties of these metric spaces. In Section 4 we prove Theorem 1.3. In the concluding section, i.e., Section 5 we present some further questions on properties of \( \mathcal{T} \) and \( \mathcal{T}^* \).

### 2 Construction of the limiting CRT’s

In this section we construct the continuum random trees \( \mathcal{T} \) and \( \mathcal{T}^* \). We first observe that for \((x, t) \in \mathbb{Z}_+^2\), joining the successive steps \(h^k(x, t), h^{k+1}(x, t) : k \geq 0\) by linear segments gives a continuous path \(\pi^{(x,t)}\) starting from \(x\) at time \(t\) and moving in the forward direction of time. In other words \(\pi^{(x,t)} \in C[t, \infty)\) is such that \(\pi^{(x,t)}(t+k) = h^k(x, t)(1)\) for all \(k \in \mathbb{N} \cup \{0\}\). Here and subsequently for \((y, s) \in \mathbb{R}^2\), \(\pi^{(y,s)}\) denotes an element in \(C[s, \infty)\) with \(\pi^{(y,s)}(s) = y\). Similarly using the dual vertices \(\hat{h}^k(x, t)\) we construct the dual path \(\hat{\pi}^{(y,s)} \in C(-\infty, s]\) with \(\hat{\pi}^{(y,s)}(s) = y\). The graph distance between any two vertices in the cluster \(C(x, t)\) can then be interpreted as the sum of the times taken by the paths starting from each of these two vertices to coalesce. In the following, we generalize this notion to obtain a tree-like metric space from a given collection of paths satisfying a certain condition (described below in Definition 2.2).

This notion will be used in the construction of \(\mathcal{T}\) and \(\mathcal{T}^*\). We start with the definition of a tree-like metric space or real tree.

**Definition 2.1.** A metric space \((X, d)\) is called a real tree or an \(\mathbb{R}\)-tree if for all \(x, y \in X\)

(i) there exists a unique geodesic from \(x\) to \(y\), i.e., there exists a unique isometry \(f_{x,y} : [0, d(x, y)] \to X\) such that \(f_{x,y}(0) = x\) and \(f_{x,y}(d(x, y)) = y\). The image of \(f_{x,y}\) is denoted by \([x, y]\);

(ii) the only non-self-intersecting path from \(x\) to \(y\) is \([x, y]\), i.e., if \(q : [0, 1] \to X\) is continuous and injective such that \(q(0) = x\) and \(q(1) = y\), then \(q([0, 1]) = [xy]\).

The first condition says that \(X\) is a geodesic space, and the second is a tree property that there is a unique way to travel between two points without backtracking. Often it is assumed that a real tree is compact. However for many purposes this assumption is unnecessary and in this paper we do not make such assumption. We mention here that the metric completion of an \(\mathbb{R}\)-tree is also an \(\mathbb{R}\)-tree (see [15]). Moreover it is known that the space of all isometry equivalence classes of compact real trees is closed in \(\mathbb{M}\) (Theorem 2.1 of [20]).

Next we describe a “tree-like” path space and how to obtain a real tree out of it. Let \(\Pi\) be the collection of all continuous real valued paths moving in the forward direction of time with all possible starting times (for a formal definition see Subsection 3.1). In other words, a path \(\pi \in \Pi\) with starting time \(\sigma_\pi \in \mathbb{R}\) is a continuous mapping \(\pi : [\sigma_\pi, \infty) \to \mathbb{R}\). Similarly \(\Pi^*\) denotes the collection of all continuous real valued paths moving in the backward direction of time with all possible starting times. In what follows, paths moving in the forward direction of time, will be referred to as forward paths and paths
moving in the backward direction of time, will be referred to as backward or dual paths. The mapping \( \gamma: \Pi \times \Pi \rightarrow \mathbb{R} \cup \{\infty\} \) is defined as \( \gamma(\pi_1, \pi_2) := \inf\{s > \sigma_{\pi_1} \lor \sigma_{\pi_2} : \pi_1(s) = \pi_2(s)\} \). So \( \gamma(\pi_1, \pi_2) \) represents the first intersection time of the two forward paths \( \pi_1 \) and \( \pi_2 \) strictly after time \( \sigma_{\pi_1} \lor \sigma_{\pi_2} \). Similarly, for two backward paths \( \hat{\pi}_1, \hat{\pi}_2 \in \hat{\Pi} \), the first meeting time is given by \( \hat{\gamma}(\hat{\pi}_1, \hat{\pi}_2) := \sup\{s < \sigma_{\hat{\pi}_1} \land \sigma_{\hat{\pi}_2} : \hat{\pi}_1(s) = \hat{\pi}_2(s)\} \).

**Definition 2.2** (Tree-like path space). \( K \subset \Pi \) is said to be tree-like if the following conditions are satisfied:

(i) \( \gamma(K) := \sup\{\gamma(\pi_1, \pi_2) : \pi_1, \pi_2 \in K\} \) is finite;

(ii) for all \( \pi_1, \pi_2 \in K \), we have \( (\pi_1(\sigma_{\pi_1}), \sigma_{\pi_1}) \neq (\pi_2(\sigma_{\pi_2}), \sigma_{\pi_2}) \) and \( \pi_1(s) = \pi_2(s) \) for all \( s \geq \gamma(\pi_1, \pi_2) \).

The set \( \{(\pi(s), s) : \pi \in K, s \in [\sigma_{\pi}, \gamma(K)]\} \) consists of the images of \( \pi \) in \( K \) up to time \( \gamma(K) \). We assume that for any \( \pi \in K \) and any \( t \geq \sigma_{\pi} \), the path \( \{\pi(s) : s \geq t\} \) with starting time \( t \) is also in \( K \).

We first observe that for a tree-like \( K \subset \Pi \) and for any \( (y_1, s_1) \) in the image set, i.e., for \( (y_1, s_1) \in \{(\pi(s), s) : \pi \in K, s \in [\sigma_{\pi}, \gamma(K)]\} \), there exists a unique path \( \pi^{(y_1, s_1)} \) starting at \( (y_1, s_1) \) in \( K \). For any two points \( (y_1, s_1) \) and \( (y_2, s_2) \) in \( \{(\pi(s), s) : \pi \in K, s \in [\sigma_{\pi}, \gamma(K)]\} \), we define the ancestor metric, \( d_A((y_1, s_1), (y_2, s_2)) \),

\[
d_A((y_1, s_1), (y_2, s_2)) := 2\gamma(\pi^{(y_1, s_1)}, \pi^{(y_2, s_2)}) - (s_1 + s_2). \tag{4}
\]

As time is measured along the \( Y \) axis, \( d_A((y_1, s_1), (y_2, s_2)) \) adds the time taken by each of the two paths from their starting points till they meet. With a slight abuse of notation, let \( \mathcal{M}(K) \) denote the completion of the metric space \( \{(\pi(s), s) : \pi \in K, s \in [\sigma_{\pi}, \gamma(K)]\}, d_A \). If \( \gamma(K) \) is finite, then by construction the metric space \( \{(\pi(s), s) : \pi \in K, s \in [\sigma_{\pi}, \gamma(K)]\}, d_A \) is tree-like and hence its completion \( \mathcal{M}(K) \) is also an \( \mathbb{R} \)-tree. In exactly the same way, this notion can be extended for a collection of backward paths \( \tilde{K} \subset \tilde{\Pi} \) to obtain a complete \( \mathbb{R} \)-tree denoted by \( \mathcal{M}(\tilde{K}) \).

Now we construct our limiting object. Consider two independent standard Brownian motions, \( B_1, B_2 \in C[0, \infty) \) with \( B_1(0) = B_2(0) = 0 \). The random times \( \tau_1 \) and \( \tau_2 \) are defined as

\[
\tau_1 := \inf\{t \geq 0 : \text{for all } s_1, s_2 \in [t, t + 1], (B_1(s_1) - B_2(s_1))(B_1(s_2) - B_2(s_2)) \geq 0\} \quad \text{and} \quad \tau_2 := \inf\{t > \tau_1 : B_1(t) = B_2(t)\}.
\]

Almost surely both \( \tau_1 \) and \( \tau_2 \) are finite with \( \tau_2 > \tau_1 + 1 \). We set \( B_1(\tau_1) = B_2(\tau_1) = x_0 \) and define \( \tilde{B}^+, \tilde{B}^- \subset C(\infty, 0] \) as,

\[
(\tilde{B}^+, \tilde{B}^-)(-t) := \begin{cases} (B_1(\tau_1 + t) \lor B_2(\tau_1 + t) - x_0, B_1(\tau_1 + t) \land B_2(\tau_1 + t) - x_0) & \text{for } t \in [0, \tau_2 - \tau_1] \\ (B_1(\tau_1 + t) - x_0, B_1(\tau_1 + t) - x_0) & \text{for } t \geq \tau_2 - \tau_1. \end{cases}
\]

Let \( \Delta = \Delta(\tilde{B}^+, \tilde{B}^-) \) be the a.s. bounded region in \( \mathbb{R}^2 \) enclosed between the backward paths \( \tilde{B}^+ \) and \( \tilde{B}^- \). More formally

\[
\Delta = \Delta(\tilde{B}^+, \tilde{B}^-) := \{(x, -t) : t \in (0, \tau_2 - \tau_1), \tilde{B}^-(-t) < x < \tilde{B}^+(-t)\}.
\]

For \( k \geq 1 \), let \( (x_1, t_1), \ldots, (x_k, t_k) \) be the first \( k \) points of \( \Delta(\tilde{B}^+, \tilde{B}^-) \cap \mathbb{Q}^2 \) with respect to some ordering. Let \( \{Z_n : n \in \mathbb{N}\} \) be a collection of i.i.d. standard Brownian motions starting from the origin at time 0 and independent of the Brownian motions \( B_1, B_2 \). For \( 1 \leq i \leq k \) let \( Y_i \) be a Brownian motion running in the backward direction of time starting from \( x_i \) at time \( t_i \) and defined as \( Y_i(t_i - t) := x_i + Z_i(t) \) for \( t \geq 0 \).

Next, from the family of backward Brownian paths \( \{Y_i : 1 \leq i \leq k\} \), we construct coalescing paths by specifying the coalescing rules and denote the resulting backward paths by \( \{\hat{W}_i : 1 \leq i \leq k\} \). The formal
construction is described inductively below. We define \( \overline{W}_1(t) := Y_1(t) \) for \( t \in [\overline{\gamma}(Y_1, \hat{B}^+) \vee \overline{\gamma}(Y_1, \hat{B}^-), t_1] \).

For \( t \leq \overline{\gamma}(Y_1, \hat{B}^+) \vee \overline{\gamma}(Y_1, \hat{B}^-) \) we define it as

\[
\overline{W}_1(t) := \begin{cases} 
\hat{B}^+(t) & \text{if } \overline{\gamma}(Y_1, \hat{B}^+) > \overline{\gamma}(Y_1, \hat{B}^-) \\
\hat{B}^-(t) & \text{otherwise.}
\end{cases}
\]

For \( i \geq 2 \), let \( s_i := \overline{\gamma}(Y_i, \hat{B}^+) \vee \overline{\gamma}(Y_i, \hat{B}^-) \vee \max \{\overline{\gamma}(Y_i, \overline{W}_j) : 1 \leq j \leq i - 1\} \). Define \( \overline{W}_i(t) := Y_i(t) \) for \( t \in [s_i, t_1] \) and for \( t \leq s_i \), let

\[
\overline{W}_i(t) := \begin{cases} 
\hat{B}^+(t) & \text{if } \overline{\gamma}(Y_i, \hat{B}^+) = s_i \\
\hat{B}^-(t) & \text{if } \overline{\gamma}(Y_i, \hat{B}^-) = s_i \\
\overline{W}_j(t) & \text{if } \overline{\gamma}(Y_i, \overline{W}_j) = s_i \text{ for some } 1 \leq i \leq j - 1.
\end{cases}
\]

It follows that for any \( k \geq 1 \), this family of backward paths \( \{\overline{W}_i : 1 \leq i \leq k\} \) satisfies Kolmogorov consistency conditions and hence this system can be extended to obtain a collection of coalescing backward paths \( \overline{S}(\hat{B}^+, \hat{B}^-) := \{\overline{W}_n : n \in \mathbb{N}\} \) starting from all the points of \( \Delta(\hat{B}^+, \hat{B}^-) \cap \mathbb{Q}^2 \). From the construction it follows that, almost surely \( \overline{S}(\hat{B}^+, \hat{B}^-) \) is tree-like in the sense of Definition 2.2. We define \( \overline{T} \) to be the complete \( \mathbb{R} \)-tree \( M(\overline{S}(\hat{B}^+, \hat{B}^-)) \).

In order to construct the continuum tree \( T \), we construct coalescing forward paths withing the region \( \Delta(\hat{B}^+, \hat{B}^-) \). Let \( B_3 \) be another standard Brownian motion starting from origin independent of \( B_1, B_2 \) and the collection \( \{Z_n : n \in \mathbb{N}\} \). It was observed in \([28]\), (see the proof the Theorem 2.4) that a collection of backward coalescing Brownian paths starting from a dense set in \( \mathbb{R}^2 \), almost surely uniquely determines a collection of coalescing forward paths starting from a dense set such that the forward paths do not cross the backward paths. Using the observation made in \([28]\), it follows that for all \((x, t) \in \Delta(\hat{B}^+, \hat{B}^-) \cap \mathbb{Q}^2 \) there exists a continuous path \( \pi(x, t) \) starting from \( x \) at time \( t \) and given as

\[
\pi(x, t)(s) := \begin{cases} 
\sup \{y : y \in Q, \overline{\pi}(y, s) \in \overline{S}(\hat{B}^+, \hat{B}^-), \overline{\pi}(y, s)(t) < x\} = \\
\inf \{y : y \in Q, \overline{\pi}(y, s) \in \overline{S}(\hat{B}^+, \hat{B}^-), \overline{\pi}(y, s)(t) > x\} & \text{if } s \in [t, 0) \cap \mathbb{Q} \\
B_3(s) & \text{for } s \geq 0.
\end{cases}
\]

If equality does not hold then for any rational \( y \) in the interval

\[
\{y : y \in \mathbb{Q}, \overline{\pi}(y, s) \in \overline{S}(\hat{B}^+, \hat{B}^-), \overline{\pi}(y, s)(t) < x\}, \{y : y \in \mathbb{Q}, \overline{\pi}(y, s) \in \overline{S}(\hat{B}^+, \hat{B}^-), \overline{\pi}(y, s)(t) > x\},
\]

the dual Brownian path \( \overline{\pi}(y, s) \) in \( \overline{S}(\hat{B}^+, \hat{B}^-) \) starting from \((y, s)\) hits \((x, t)\) with zero probability. Hence this equality holds almost surely. This gives a collection of coalescing forward paths \( S(\hat{B}^+, \hat{B}^-) := \{\pi(x, t) : (x, t) \in \Delta(\hat{B}^+, \hat{B}^-) \cap \mathbb{Q}^2 \} \) starting from all the points of \( \Delta(\hat{B}^+, \hat{B}^-) \cap \mathbb{Q}^2 \). As subset of \( \Pi \), the collection \( S(\hat{B}^+, \hat{B}^-) \) is tree-like almost surely. \( T \) is defined to be the complete real tree \( M(S(\hat{B}^+, \hat{B}^-)) \).

\( T \) as conjectured in \([2]\) : We remark here that a slightly different continuum random tree was surmised in \([2]\) as the scaling limit. Since Aldous was interested about the scaling limit of the discrete tree conditioned to have time-length exactly \( n \), i.e., \( T_n(0, 0)\{L(0, 0) = n\} \), the random region \( \Delta \) was enclosed by two independent backward Brownian motions both starting at the origin and conditioned to meet for the first time exactly at time \(-1\). Since we are working with the “non-degenerate” conditioning \( \{L(0, 0) \geq n\} \), in our case the region \( \Delta = \Delta(\hat{B}^+, \hat{B}^-) \) is enclosed by two independent backward Brownian motions both starting at the origin and conditioned not to meet within time \([-1, 0)\). It is important to observe that in the continuum tree as described in \([2]\), the forward coalescing Brownian paths coalesce with the boundary of \( \Delta(\hat{B}^+, \hat{B}^-) \) as soon as they hit the boundary (see Page 276 of \([2]\) ). On the other hand, from the work of Sonciniuc et. al. \([31]\) it follows that the forward coalescing Brownian paths follow Skorohod reflection at the boundary of \( \Delta(\hat{B}^+, \hat{B}^-) \) (for a detail definition of Skorohod reflection see \([31]\)). This explains the difference of the limiting object from the conjectured one in \([2]\). We mention here that for \( \overline{T} \), the backward coalescing Brownian paths coalesce with the boundary of \( \Delta(\hat{B}^+, \hat{B}^-) \) once they hit it.
3 Double Brownian web and compactness of $\mathcal{T}$, $\hat{\mathcal{T}}$

In this section we prove that both the random metric spaces $\mathcal{T}$ and $\hat{\mathcal{T}}$ are compact almost surely. Towards this we introduce a related random object, called the double Brownian web, i.e., the Brownian web and its dual denoted by $(\mathcal{W}, \hat{\mathcal{W}})$ and show that a version of $(S(\hat{B}^+, \hat{B}^-), S(\hat{B}^+, \hat{B}^-))$ can be embedded in $(\mathcal{W}, \hat{\mathcal{W}})$. This allows us to use properties of $(\mathcal{W}, \hat{\mathcal{W}})$ to prove compactness. The proof of Theorem 1.3 also uses properties of $(\mathcal{W}, \hat{\mathcal{W}})$.

3.1 Brownian web and its dual

The Brownian web originated in the work of Arratia (see [5]). Later Fontes et al. [13] studied the Brownian web as a random variable in an appropriate Polish space. We recall the relevant details from [13].

Let $\mathbb{R}_+^2$ denote the completion of the space time plane $\mathbb{R}^2$ with respect to the metric

$$d((x_1, t_1), (x_2, t_2)) := |\tanh(t_1) - \tanh(t_2)| \sqrt{\frac{\tanh(x_1)}{1 + |t_1|} - \frac{\tanh(x_2)}{1 + |t_2|}}$$

As a topological space $\mathbb{R}_+^2$ can be identified with the continuous image of $[-\infty, \infty]^2$ under a map that identifies the line $[-\infty, \infty] \times \{\infty\}$ with the point $(*, \infty)$, and the line $[-\infty, \infty] \times \{-\infty\}$ with the point $(*, -\infty)$. A path $\pi$ in $\mathbb{R}_+^2$ with starting time $\sigma_\pi \in [-\infty, \infty]$ is a mapping $\pi : [\sigma_\pi, \infty] \to [-\infty, \infty] \cup \{*\}$ such that $\pi(\infty) = *$ and, when $\sigma_\pi = -\infty$, $\pi(-\infty) = *$. Also $t \mapsto (\pi(t), t)$ is a continuous map from $[\sigma_\pi, \infty]$ to $(\mathbb{R}_+^2, \rho)$. We then define $\Pi$ to be the space of all paths in $\mathbb{R}_+^2$ with all possible starting times in $[-\infty, \infty]$. The following metric, for $\pi_1, \pi_2 \in \Pi$

$$d_\Pi(\pi_1, \pi_2) := |\tanh(\sigma_{\pi_1}) - \tanh(\sigma_{\pi_2})| \sup_{t \geq \sigma_{\pi_1} \wedge \sigma_{\pi_2}} \frac{\tanh(\pi_1(t \lor \sigma_{\pi_1}))}{1 + |t|} - \frac{\tanh(\pi_2(t \lor \sigma_{\pi_2}))}{1 + |t|}$$

makes $\Pi$ a complete, separable metric space. The metric $d_\Pi$ is slightly different from the original choice in [13] which is somewhat less natural as explained in [32].

Remark 3.1. Convergence in the $(\Pi, d_\Pi)$ metric can be described as locally uniform convergence of paths as well as convergence of starting times. Therefore, for any $\epsilon > 0$ and $m > 0$, we can choose $\epsilon_1(= g(\epsilon, m)) > 0$ such that for $\pi_1, \pi_2 \in \Pi$ with $\{(\pi_i(t), t) : t \in [\sigma_{\pi_i}, m]\} \subseteq [-m, m] \times [-m, m]$ for $i = 1, 2$ and $d_\Pi(\pi_1, \pi_2) < \epsilon_1$ imply that $|\sigma_{\pi_1} - \sigma_{\pi_2}| \leq |\sigma_{\pi_1} - t - \pi_2(\sigma_{\pi_2})| < \epsilon$ and $\sup\{|\pi_1(t) - \pi_2(t)| : t \in [\sigma_{\pi_1} \lor \sigma_{\pi_2}, m]\} < \epsilon$.

Let $\mathcal{H}$ denote the space of compact subsets of $(\Pi, d_\Pi)$ equipped with the Hausdorff metric $d_{\mathcal{H}}$. As $(\Pi, d_\Pi)$ is Polish, it follows that $(\mathcal{H}, d_{\mathcal{H}})$ is also Polish. Let $B_{\mathcal{H}}$ be the Borel $\sigma$-algebra on the metric space $(\mathcal{H}, d_{\mathcal{H}})$. The Brownian web $\mathcal{W}$ is an $(\mathcal{H}, B_{\mathcal{H}})$ valued random variable characterized as (Theorem 2.1 of [13]):

Theorem 3.2. There exists an $(\mathcal{H}, B_{\mathcal{H}})$ valued random variable $\mathcal{W}$ such that whose distribution is uniquely determined by the following properties:

(a) for each deterministic point $z \in \mathbb{R}^2$ there is a unique path $\pi^z \in \mathcal{W}$ starting from $z$ almost surely;

(b) for a finite set of deterministic points $z^1, \ldots, z^k \in \mathbb{R}^2$, the collection $(\pi^{z^1}, \ldots, \pi^{z^k})$ is distributed as coalescing Brownian motions starting from $z^1, \ldots, z^k$;

(c) for any countable deterministic dense set $\mathcal{D} \subseteq \mathbb{R}^2$, $\mathcal{W}$ is the closure of $\{\pi^z : z \in \mathcal{D}\}$ in $(\Pi, d_\Pi)$ almost surely.
The above theorem shows that the collection is almost surely determined by countably many coalescing Brownian motions.

The metric space of compact sets of backward paths is defined similarly and denoted by \((\hat{H}, d_{\hat{H}})\). Let \(B_{\hat{H}}\) be the corresponding Borel \(\sigma\) field. \((W, \hat{W})\) is a \((\hat{H} \times \hat{H}, B_{\hat{H} \times \hat{H}})\) valued random variable such that \(W\) and \(\hat{W}\) uniquely determine each other with \(\hat{W}\) being equally distributed as \(-W\), the Brownian web rotated \(180^\circ\) about the origin. Theorems 5.8 of [12] and Theorem 2.4 of [28] give characterizations of the double Brownian web \((W, \hat{W})\). Before ending this subsection we list some properties of \((W, \hat{W})\) which hold almost surely and which will be used later.

**Properties of \((W, \hat{W})\)**:

(a) As in [12], for \((W, \hat{W})\) and \((x, t) \in \mathbb{R}^2\), we define

\[
\begin{align*}
m_{in}(x, t) := & \lim_{\epsilon \downarrow 0} \text{number of paths in } W \text{ starting at } (y, t - \epsilon) \text{ for some } y \text{ that pass} \nonumber \\
& \quad \text{through } (x, t) \text{ and are disjoint in the time interval } (t - \epsilon, t) \}; \\
m_{out}(x, t) := & \lim_{\epsilon \downarrow 0} \text{number of paths in } W \text{ starting at } (x, t) \text{ that are} \nonumber \\
& \quad \text{disjoint in the time interval } (t, t + \epsilon). 
\end{align*}
\]

(b) Let \(D \subseteq \mathbb{R}^2\) be a deterministic countable dense set. For any \((x, t) \in D\), there exists a unique path \(\pi^{(x, t)} \in W\) distributed as a Brownian motion starting from \((x, t)\).

(c) Paths in \(W\) coalesce when they meet, i.e., for each \(\pi, \pi' \in W\) we have \(\gamma(\pi, \pi') < \infty\) and \(\pi(s) = \pi'(s)\) for all \(s \geq \gamma(\pi, \pi')\) (see Proposition 3.2 of [28]).

(d) For every \(\pi \in W\) and \(\tilde{\pi} \in \hat{W}\)

(i) there does not exist any \(s, t \in [\sigma_{\pi, \pi}, \sigma_{\tilde{\pi}}]\) with \((\pi(s) - \tilde{\pi}(s))(\pi(t) - \tilde{\pi}(t)) < 0\), i.e., no forward path of \(W\) crosses a dual path of \(\hat{W}\);

(ii) \(\int_{\sigma_{\pi}}^{\gamma(\pi, \pi')} 1_{\{\pi(s) = \tilde{\pi}(s)\}} ds = 0\), i.e., forward paths of \(W\) and dual paths of \(\hat{W}\) “spend zero Lebesgue time together” (see Proposition 3.2 in [22]).

(e) For any \(t > 0, t_0 \in \mathbb{R}\)

(i) the set \(\{\pi(t_0 + t) : \pi \in W, \sigma_{\pi} \leq t_0\}\) is locally finite (see Proposition 4.3 in [13]);

(ii) for any \(x \in \{\pi(t_0 + t) : \pi \in W, \sigma_{\pi} \leq t_0\}\), there exist \(\pi_1, \pi_2 \in \hat{W}\) both starting from the point \((x, t_0 + t)\) such that \(\pi_1(s) \neq \pi_2(s)\) for all \(s \in [t_0, t_0 + t)\) (follows from (a));

(iii) for any \(x \in \{\pi(t_0 + t) : \pi \in W, \sigma_{\pi} \leq t_0\}\), there exist \(\pi_1, \pi_2 \in W\) with \(\sigma_{\pi_1} < t_0 < \sigma_{\pi_2}\) such that \(\pi_1(t_0 + t) = \pi_2(t_0 + t) = x\) (follows from (d) (ii), (b) and (c)).

(f) For each point \((x, t) \in \mathbb{R}^2\) and any two sequences \(\{x^n_-, n \geq 1\}, \{x^n_+, n \geq 1\} \subset \mathbb{R}\) such that \(x^n_- \uparrow x\) and \(x^n_+ \downarrow x\) consider the paths \(\pi^{(x^n_-, t)}\) \(\pi^{(x^n_+, t)} \in W\), starting from \((x^n_-, t), (x^n_+, t)\) respectively. The limits \(\lim_{n \to \infty} \pi^{(x^n_-, t)}\) \(\lim_{n \to \infty} \pi^{(x^n_+, t)} \in W\) exist and do not depend on the choice of the sequences \(\{x^n_- : n \geq 1\}, \{x^n_+ : n \geq 1\}\) (see Proposition 3.2 (e) of [22]).

(g) For \(\{\pi^n : n \geq 1\} \subset W\) and \(\tilde{\pi} \in W\) with \(d_H(\pi^n, \tilde{\pi}) \to 0\), we have that \(\gamma(\pi^n, \tilde{\pi}) \to \sigma_{\tilde{\pi}}\) as \(n \to \infty\) (see Lemma 3.4 of [22]).
3.2 Compactness of $\mathcal{T}$ and $\tilde{T}$

We first show that a version of $(S(\hat{\pi}^+, \hat{\pi}^+), \hat{S}(\hat{\pi}^+, \hat{\pi}^+))$ can be embedded in $(\mathcal{W}, \tilde{\mathcal{W}})$. We need to introduce some notation. Consider $\hat{\pi}^{(0,0)}$, the backward Brownian path in $\mathcal{W}$ starting at $(0,0)$. For ease of notation we take $\hat{\pi}_0 = \hat{\pi}^{(0,0)}$. Let $\theta_0 \leq 0$ be defined as,

$$\theta_0 := \sup \{ t \leq 0 : \text{there exists } \hat{\pi}_1 \in \tilde{\mathcal{W}} \text{ such that } (\hat{\pi}_1(\sigma_{\hat{\pi}_1}), (\sigma_{\hat{\pi}_1})) = (\hat{\pi}_0(t), t) \text{ and } \hat{\pi}_1(t-1) \neq \hat{\pi}_0(t-1) \}.$$  

(8)

In other words, the point $(\hat{\pi}_0(\theta_0), \theta_0)$ is a point of $(1,2)$ type, i.e., with 2 outgoing dual paths and $\hat{\pi}_1$ is the dual path in $\tilde{\mathcal{W}}$, which starts at $(\hat{\pi}_0(\theta_0), \theta_0)$ and do not coalesce with $\hat{\pi}_0$ within the time $(\theta_0 - 1, \theta_0)$. The coalescing time of these two dual paths, is given by $\theta_1$ and almost surely $\theta_0 > \theta_1 + 1$.

Let $\Delta(\hat{\pi}_0, \hat{\pi}_1) \subset \mathbb{R}^2$, the region enclosed between these two dual paths $\hat{\pi}_0$ and $\hat{\pi}_1$, be defined as

$$\Delta(\hat{\pi}_0, \hat{\pi}_1) := \{ (y, s) \in \mathbb{R}^2 : s \in (\theta_1, \theta_0), y \in (\hat{\pi}_0(s) \land \hat{\pi}_1(s), \hat{\pi}_0(s) \lor \hat{\pi}_1(s)) \}.$$  

(9)

Let $\hat{S}(\hat{\pi}_0, \hat{\pi}_1) := \{ \hat{\pi} \in \tilde{\mathcal{W}} : (\hat{\pi}(\sigma_\pi), (\sigma_{\hat{\pi}})) \in \Delta(\hat{\pi}_0, \hat{\pi}_1) \cap \mathbb{Q}^2 \}$ denote the collection of dual paths in $\tilde{\mathcal{W}}$, starting from all the rational points in the region $\Delta(\hat{\pi}_0, \hat{\pi}_1)$. Similarly we define $\hat{S}(\hat{\pi}_0, \hat{\pi}_1) := \{ \pi \in \mathcal{W} : (\pi(\sigma_\pi), (\sigma_{\pi})) \in \Delta(\hat{\pi}_0, \hat{\pi}_1) \cap \mathbb{Q}^2 \}$, as the collection of paths in $\mathcal{W}$ starting from all the rational points in the region $\Delta(\hat{\pi}_0, \hat{\pi}_1)$.

From property (e) of $(\mathcal{W}, \tilde{\mathcal{W}})$, it follows that both the path spaces $\hat{S}(\hat{\pi}_0, \hat{\pi}_1)$ and $\hat{S}(\hat{\pi}_0, \hat{\pi}_1)$ are tree-like in the sense of Definition 2.2. We consider the complete metric spaces $M(\hat{S}(\hat{\pi}_0, \hat{\pi}_1))$ and $M(\hat{S}(\hat{\pi}_0, \hat{\pi}_1))$ and show that they have the same distribution as $\mathcal{T}$ and $\tilde{T}$ respectively.

**Proposition 3.3.** We have

$$(\mathcal{T}, \tilde{T}) \overset{d}{=} (M(\hat{S}(\hat{\pi}_0, \hat{\pi}_1)), M(\hat{S}(\hat{\pi}_0, \hat{\pi}_1))).$$

Assuming the above proposition we first prove that both $\mathcal{T}$ and $\tilde{T}$ are compact almost surely.

**Proposition 3.4.** Both $\mathcal{T}$ and $\tilde{T}$ are compact almost surely.

**Proof:** Because of Proposition 3.3, it suffices to show that both the metric spaces, $M(\hat{S}(\hat{\pi}_0, \hat{\pi}_1))$ and $M(\hat{S}(\hat{\pi}_0, \hat{\pi}_1))$ are compact almost surely. Being complete, it is enough to show that both the metric spaces are totally bounded as well. We prove it for $M(\hat{S}(\hat{\pi}_0, \hat{\pi}_1))$. The argument is exactly same for $M(\hat{S}(\hat{\pi}_0, \hat{\pi}_1))$. Fix $\epsilon > 0$ and set $\delta = \delta(\epsilon) \in (0, ((\theta_0 - \theta_1)^{-1} \wedge \epsilon)/4)$. For $j \geq 1$, let $t_{j}^\epsilon := \theta_1 + j \delta$ and $j_{\max} := \min \{ j \geq 1 : t_{j+1}^\epsilon > \theta_0 \}$. Because of property (e) of $(\mathcal{W}, \tilde{\mathcal{W}})$, for any $1 \leq j \leq j_{\max} - 1$ the set $\{ \pi(t_{j+1}^\epsilon) : \sigma_\pi \leq t_{j+1}^\epsilon, \pi \in \hat{S}(\hat{\pi}_0, \hat{\pi}_1) \}$ is finite. Since $M(\hat{S}(\hat{\pi}_0, \hat{\pi}_1))$ is the completion of $\{ \{ (\pi(s), s) : \pi \in \hat{S}(\hat{\pi}_0, \hat{\pi}_1), s \in [\sigma_\pi, \theta_0] \}, d_{A} \}$, the choice of $\delta$ ensures that the collection

$$\bigcup_{j=1}^{j_{\max}-1} \{ \pi(t_{j+1}^\epsilon) : \sigma_\pi \leq t_{j+1}^\epsilon, \pi \in \hat{S}(\hat{\pi}_0, \hat{\pi}_1) \} \cup \{ (\hat{\pi}_0(\theta_0), \theta_0) \}$$

forms a finite $\epsilon$-cover for $M(\hat{S}(\hat{\pi}_0, \hat{\pi}_1))$. This completes the proof.

Now we proceed to prove Proposition 3.3 through a sequence of lemmas. We need to introduce some notation. Set $x_0 = x_0(\omega) := \hat{\pi}_0(\theta_0) \in \mathbb{R}$. Let $\hat{\pi}^+, \hat{\pi}^- \in C(\mathbb{R}, \mathbb{R})$ be defined as

$$(\hat{\pi}^+(s), (\hat{\pi}^-)(s)) := (\hat{\pi}_0(\theta_0 - s) \lor \hat{\pi}_1(\theta_0 - s) - x_0, \hat{\pi}_0(\theta_0 - s) \land \hat{\pi}_1(\theta_0 - s) - x_0)$$

for $s \geq 0$. (10)

Since the forward paths in $\mathcal{W}$ do not cross the dual paths in $\tilde{\mathcal{W}}$ almost surely, as in [4] the dual paths in $\hat{S}(\hat{\pi}_0, \hat{\pi}_1)$ almost surely determine the forward paths in $S(\hat{\pi}_0, \hat{\pi}_1)$ within the region $\Delta(\hat{\pi}_0, \hat{\pi}_1)$, i.e., upto
time \( \theta_0 \) only. Hence from the construction of \((\mathcal{T}, \tilde{T})\) it follows that to prove Proposition 3.3 it suffices to show that for \((\hat{\pi}^+, \hat{\pi}^-)\) as an element in \(C(-\infty, 0] \times C(-\infty, 0]\), we have

\[
(\hat{\pi}^+, \hat{\pi}^-) \overset{d}{=} (\hat{\mathcal{B}}^+, \hat{\mathcal{B}}^-),
\]

where \((\hat{\mathcal{B}}^+, \hat{\mathcal{B}}^-)\) is as defined in (10).

Following the construction of \((\hat{\mathcal{B}}^+, \hat{\mathcal{B}}^-)\), for general \(f_1, f_2 \in C(-\infty, 0]\) with \(f_1(0) = f_2(0) = 0\), we define

\[
\lambda_0(f_1, f_2) := \sup \{t \leq 0 : (f_1(s_1) - f_2(s_1))(f_1(s_2) - f_2(s_2)) \geq 0 \text{ for all } s_1, s_2 \in [t-1, t]\} \text{ and }
\]

\[
\lambda_1(f_1, f_2) := \sup \{t < \lambda_0 : f_1(t) = f_2(t)\}.
\]

For \(n \geq 1\) we define

\[
\lambda_n^0(f_1, f_2) := \sup \{t \leq 0 : |f_1(t) - f_2(t)| = 1/n, (f_1(s_1) - f_2(s_1))(f_1(s_2) - f_2(s_2)) \geq 0 \text{ for all } s_1, s_2 \in [t-1, t]\} \text{ and }
\]

\[
\lambda_n^1(f_1, f_2) := \sup \{t < \lambda_n^0 : f_1(t) = f_2(t)\}.
\]

Let \(A, A^n \subset C(-\infty, 0]\times C(-\infty, 0]\) be such that

\[
A := \{(f_1, f_2) : f_i \in C(-\infty, 0]\text{ with } f_i(0) = 0 \text{ for } i = 1, 2 \text{ and } \lambda_1(f_1, f_2) < \lambda_0(f_1, f_2) - 1\}
\]

\[
A^n := \{(f_1, f_2) : f_i \in C(-\infty, 0]\text{ with } f_i(0) = 0 \text{ for } i = 1, 2 \text{ and } \lambda_n^1(f_1, f_2) < \lambda_n^0(f_1, f_2) - 1\}.
\]

Consider the mapping \(\Gamma : C(-\infty, 0]\times C(-\infty, 0]\rightarrow C(-\infty, 0]\times C(-\infty, 0]\) where \(\Gamma(f_1, f_2) := (f_1, f_2)\) for \((f_1, f_2) \notin A\). For \((f_1, f_2) \in A\) and for \(f_1(\lambda_0) = f_2(\lambda_0) = y_0\), we define

\[
\Gamma(f_1, f_2)(s) :=
\]

\[
\begin{cases}
(f_1(\lambda_0 + s) - y_0, f_2(\lambda_0 + s) - y_0) & \text{if } s \in [\lambda_1 - \lambda_0, 0] \text{ and } f_1(\lambda_0 + s) \geq f_2(\lambda_0 + s) \\
(f_2(\lambda_0 + s) - y_0, f_1(\lambda_0 + s) - y_0) & \text{if } s \in [\lambda_1 - \lambda_0, 0] \text{ and } f_2(\lambda_0 + s) > f_1(\lambda_0 + s) \\
(f_1(\lambda_0 + s) - y_0, f_2(\lambda_0 + s) - y_0) & \text{if } s \leq \lambda_1 - \lambda_0.
\end{cases}
\]

Similarly for \(n \geq 1\), the mapping \(\Gamma^n : C(-\infty, 0]\times C(-\infty, 0]\rightarrow C(-\infty, 0]\times C(-\infty, 0]\) is defined as \(\Gamma^n(f_1, f_2) := (f_1, f_2)\) for \((f_1, f_2) \notin A^n\). For \((f_1, f_2) \in A^n\) and for \(f_1(\lambda_0^n) = f_2(\lambda_0^n) = y_0^n\), we define

\[
\Gamma^n(f_1, f_2)(s) :=
\]

\[
\begin{cases}
(f_1(\lambda_0^n + s) - y_0^n, f_2(\lambda_0^n + s) - y_0^n) & \text{if } s \in [\lambda_1^n - \lambda_0^n, 0] \text{ and } f_1(\lambda_0^n + s) \geq f_2(\lambda_0^n + s) \\
(f_2(\lambda_0^n + s) - y_0^n, f_1(\lambda_0^n + s) - y_0^n) & \text{if } s \in [\lambda_1^n - \lambda_0^n, 0] \text{ and } f_2(\lambda_0^n + s) > f_1(\lambda_0^n + s) \\
(f_1(\lambda_0^n + s) - y_0^n, f_2(\lambda_0^n + s) - y_0^n) & \text{if } s \leq \lambda_1^n - \lambda_0^n.
\end{cases}
\]

Before proceeding further, we have the following deterministic corollary which will be used in the proof of Proposition 3.3.

**Corollary 3.5.** For any \((f_1, f_2) \in C(-\infty, 0]\times C(-\infty, 0]\) we have

\[
\lim_{n \to \infty} \Gamma^n(f_1, f_2) = \Gamma(f_1, f_2)
\]

under the product metric.

**Proof:** Take any \((f_1, f_2) \in A\) and we first observe that \((f_1, f_2) \in A_n\) also for all large \(n\). On the other hand since \(f_1(0) = f_2(0)\), if \((f_1, f_2) \in A_m\) for some \(m\) then \((f_1, f_2) \in A_n\) for all \(n \geq m\) as well as \((f_1, f_2) \in A\). This gives us that, if \(f_1, f_2 \notin A\) then \(f_1, f_2 \notin A_n\) for all \(n \geq 1\). Hence it suffices to consider \((f_1, f_2) \in A\) only. Fix \(\epsilon > 0\). Since \(f_1, f_2\) are both continuous with \(f_1(0) = f_2(0) = 0\), we have \(\lambda_0^n \leq \lambda_0\) and \(\lambda_1^n \leq \lambda_1\) for all \(n\). Choose \(\delta > 0\) such that the following conditions hold:
Since $\epsilon > 0$ consequently $\lambda$.

For all $n > 1/\beta$ we have $\lambda_n^1 \in (\lambda_0 - \delta, \lambda_0)$ and consequently $\lambda_n^0 = \lambda_1$. Further the conditions (ii) and (iii) mentioned above imply that

$$\sup\{|f_1(\lambda_n^0 - s) - f_1(\lambda_0 - s) - \lambda_2(\lambda_0 - s)| : s \in [0, \lambda_0 - \lambda_1]\} < \epsilon$$

$$\Gamma_n(f_1, f_2)(-s) = \Gamma(f_1, f_2)(-s) \text{ for all } s \geq \lambda_0 - \lambda_1.$$  \(13\)

Since $\epsilon > 0$ is chosen arbitrarily, \(13\) completes the proof.

Let $\hat{W}^{(0,0)}$ and $\hat{W}^{(1/n,0)}$ denote two independent coalescing backward Brownian motions starting from the points $(0,0)$ and $(1/n,0)$ respectively.

Let $\hat{\gamma}_n := \hat{\gamma}(\hat{W}^{(0,0)}, \hat{W}^{(1/n,0)})$ be the coalescing time of these two coalescing Brownian motions. Recall that $\hat{B}_1$ and $\hat{B}_2$ are two independent standard Brownian motions starting from the origin. $\hat{B}_1$ and $\hat{B}_2$ are two independent backward Brownian motions defined as $\hat{B}_1(-s), \hat{B}_2(-s)) := (B_1(s), B_2(s))$ for $s \geq 0$.

The following lemma and its proof is a modification of Lemma 3.1 of [8].

**Lemma 3.6.** For all $n \geq 1$ we have

$$(\hat{W}^{(1/n,0)}, \hat{W}^{(0,0)}): \hat{\gamma}_n < -1 \overset{d}{=} \Gamma_n(\hat{B}_1, \hat{B}_2).$$

**Proof:** For $t > 0$ and $n \geq 1$, let $E_n^t$ denote the event that

$$\bigcup_{s \in [0,t]} \bigcap_{m > 1/(t-s)} \{ \text{there exists } s_1 \in [s, (s + 1/m) \wedge t] \cap Q \text{ such that} \}

((B_1 - B_2)(s) - 1/n)((B_1 - B_2)(s_1) - 1/n) < 0 \text{ and one of the following holds:}

(i) $B_1(s + i/m) \geq B_2(s + i/m) : 0 \leq i < \lfloor m((t-s) \wedge 1) \rfloor$,

(ii) $B_2(s + i/m) \geq B_1(s + i/m) : 0 \leq i < \lfloor m((t-s) \wedge 1) \rfloor$.

Since two independent Brownian motions immediately cross each other once they intersect, occurrence of the event $E_n^t$ implies that there exists some rational $s < t$ such that $(B_1 - B_2)(s_0) = 1/n$ for some $s_0 \in (s, t)$ and $(B_1 - B_2)(s_1) > 0$ for all $s_1 \in (s, t \wedge (s + 1))$. For all $t > 0$ the event $E_n^t$ is $F_t := \sigma((B_1(s), B_2(s) : 0 \leq s \leq t))$ measurable. Note that almost surely $(\hat{B}_1, \hat{B}_2) \in A_n$ for all $n \geq 1$. For any $t > 0$ and for $n \geq 1$ we have

$$\{\lambda_n^0(\hat{B}_1, \hat{B}_2) = -t\} = \{(B_1(s_1) - B_2(s_1))(B_1(s_2) - B_2(s_2)) \geq 0 \text{ for all } s_1, s_2 \in [t, t + 1]\}

\cap \{|B_1(t) - B_2(t)| = 1/n\} \cap (E_n^t)^c.$$  \(14\)

The mapping $\Theta : C(\neg \infty, 0] \times C(\neg \infty, 0] \mapsto C(\neg \infty, 0] \times C(\neg \infty, 0]$ is defined as

$$\Theta((\pi_1, \pi_2))(s) := \begin{cases} (\pi_1(s), \pi_2(s)) & \text{for } s \in [\gamma(\pi_1, \pi_2), 0] \\ (\pi_1(s), \pi_1(s)) & \text{for } s \leq \gamma(\pi_1, \pi_2). \end{cases}$$

For any Borel set $B \subseteq C(\neg \infty, 0] \times C(\neg \infty, 0]$ and for $y_n^0 = \hat{B}_1(\lambda_n^0) \wedge \hat{B}_2(\lambda_n^0)$, we obtain

$$\mathbb{P}(\{(\hat{B}_1, \hat{B}_2) \in B\})

= \mathbb{P}(\Theta((\{\hat{B}_1(\lambda_n^0 - s) \vee \hat{B}_2(\lambda_n^0 - s) - y_n^0, \hat{B}_1(\lambda_n^0 - s) \wedge \hat{B}_2(\lambda_n^0 - s) - y_n^0 : s \geq 0\}) \in B))$$  \(15\).
Without loss of generality we assume $\hat{B}_1(\lambda_0^0) < \hat{B}_2(\lambda_0^0)$, from (15) and obtain
\[
\mathbb{E}\left( \mathbb{P}\left( \{ (\hat{B}_2(-t-s) - \hat{B}_1(-t), \hat{B}_1(-t-s) - \hat{B}_1(-t)) : s \geq 0 \} \right) \in B | \lambda_0^0 = -t \right)
\]
\[
= \mathbb{E}\left( \mathbb{P}\left( \{ (\hat{B}_2(-t-s) - \hat{B}_1(-t), \hat{B}_1(-t-s) - \hat{B}_1(-t)) : s \geq 0 \} \right) | \hat{B}_1(-t) - \hat{B}_2(-t) \right) = 1/n,
\]
\[
(\hat{B}_1(s_1) - \hat{B}_2(s_1))(\hat{B}_1(s_2) - \hat{B}_2(s_2)) > 0 \text{ for all } s_1, s_2 \in [-t, -1, -t], (E_n^{\gamma_1}'). \tag{16}
\]
The last step follows from (14) and the fact that for $\hat{B}_1, \hat{B}_2$ almost surely we have
\[
(\hat{B}_1(s_1) - \hat{B}_2(s_1))(\hat{B}_1(s_2) - \hat{B}_2(s_2)) > 0 \text{ for all } s_1, s_2 \in [\lambda_0^0, -1, \lambda_0^0].
\]
From spatial homogeneity and from Markov property of $B_1, B_2$, it follows that given $\{ |\hat{B}_2(-t-s) - \hat{B}_1(-t)| = 1/n \}$, the distribution of the process
\[
\{ (\hat{B}_2(-t-s) \vee \hat{B}_2(-t-s) - \hat{B}_1(-t) \land \hat{B}_2(-t-s) \land \hat{B}_2(-t-s) - \hat{B}_1(-t) \land \hat{B}_2(-t) : s \geq 0 \}
\]
does not depend on $\mathcal{F}_t$ and has the same distribution as two independent backward Brownian motions starting from the points $(1/n, 0)$ and $(0, 0)$ respectively. Since the event $E_n^{\gamma_1}$ is $\mathcal{F}_t$ measurable, from (16) we obtain
\[
= \mathbb{E}\left( \mathbb{P}\left( \{ (\hat{B}_2(-t-s) - \hat{B}_1(-t), \hat{B}_1(-t-s) - \hat{B}_1(-t)) : s \geq 0 \} \right) | \hat{B}_1(t) - \hat{B}_2(t) \right) = 1/n, (\hat{B}_1(s_1) - \hat{B}_2(s_1))(\hat{B}_1(s_2) - \hat{B}_2(s_2)) > 0 \text{ for all } s_1, s_2 \in [-t, -1, -t])
\]
\[
= \mathbb{P}\left( (\hat{W}^{(1/n, 0)}, \hat{W}^{(0, 0)}) \in B | \gamma_n < -1 \right) \]
\[
= \mathbb{P}\left( (\hat{W}^{(1/n, 0)}, \hat{W}^{(0, 0)}) \in B | \gamma_n < -1 \right).
\]
The Borel set $B$ being arbitrary, the lemma holds.

Now we use the above lemma to prove Proposition 3.3

**Proof of Proposition 3.3** As observed earlier it suffices to prove (11). Recall that $\hat{\pi}_0$ is the dual path in $\hat{W}$ starting from the point $(0, 0)$. For $n \geq 1$ let $l_n$ and $r_n$ be rationals given by $l_n := ([n\hat{\pi}_0(\theta_0)])/n$ and $r_n := l_n + 1/n$ where $\theta_0$ is as defined as in (5). We choose $\theta_0^0 \in (\theta_0 - 1/n, \theta_0) \cap Q$ such that

(i) $\hat{\pi}_0(s), \hat{\pi}_1(s)$ for all $s \in [\theta_0^0, \theta_0]$ where $\hat{\pi}_1$ is the dual path in $\hat{W}$ starting from $(\hat{\pi}_0(\theta_0), \theta_0)$,

(ii) $\theta_0^0 > \theta_0 + 1$.

Since $\theta_0 > \theta_0 + 1$ almost surely and both the backward continuous paths $\hat{\pi}_0$ and $\hat{\pi}_1$ pass through the point $(\hat{\pi}_0(\theta_0), \theta_0)$, such $\theta_0^0$ always exists. Clearly $\theta_0^0 \to \theta_0$ as $n \to \infty$. For all $n \geq 1$ both $(l_n, \theta_0^0)$ and $(r_n, \theta_0^0)$ are in $\mathbb{Q}$ and let $\hat{\pi}^{(l_n, \theta_0^0)}$ and $\hat{\pi}^{(r_n, \theta_0^0)}$ denote the dual paths in $\hat{W}$ starting from the points $(l_n, \theta_0^0)$ and $(r_n, \theta_0^0)$ respectively. Since the dual paths in $\hat{W}$ are non-crossing, condition (ii) implies that $\hat{\pi}^{(r_n, \theta_0^0)}(\theta_0^0 - 1) > \hat{\pi}^{(l_n, \theta_0^0)}(\theta_0^0 - 1)$. The backward paths $\hat{\pi}_n^+, \hat{\pi}_n^- \in C(\infty, 0]$ are defined as
\[
\hat{\pi}_n^+(s) := \hat{\pi}^{(r_n, \theta_0^0)}(\theta_0^0 - s) - l_n \text{ and } \hat{\pi}_n^-(s) := \hat{\pi}^{(l_n, \theta_0^0)}(\theta_0^0 - s) - l_n \text{ for } s \geq 0.
\]
By construction we have $\hat{\pi}_n^+(0) - \hat{\pi}_n^-(0) = 1/n$ and $\hat{\pi}_n^+(s) > \hat{\pi}_n^-(s)$ for all $s \in [-1, 0]$ almost surely.

Since $\hat{W}$ is compact, both the sequences $\{\hat{\pi}^{(l_n, \theta_0^0)} : n \in \mathbb{N}\}$ and $\{\hat{\pi}^{(r_n, \theta_0^0)} : n \in \mathbb{N}\}$ must have convergent subsequences. As convergence of paths implies convergence of starting times as well, subsequential limits of each of these two sequences must be backward paths in $\hat{W}$ starting at $(\hat{\pi}_0(\theta_0), \theta_0)$. From the properties of $(W, \hat{W})$ we have that there are exactly two dual paths starting from the point $(\hat{\pi}_0(\theta_0), \theta_0)$. Hence it follows that
\[
\lim_{n \to \infty} (\hat{\pi}^{(l_n, \theta_0^0)}, \hat{\pi}^{(r_n, \theta_0^0)}) = \begin{cases} \{\hat{\pi}_0(s) : s \leq \theta_0\}, & \hat{\pi}_1 \\
(\hat{\pi}_1, \{\hat{\pi}_0(s) : s \leq \theta_0\}) & \text{otherwise}, \end{cases} \tag{17}
\]
where the limit is taken in \((\hat{\Pi},d_{\mathcal{H}})\). This shows that the sequences of dual paths, \(\{\hat{\pi}^+_n : n \in \mathbb{N}\}\) and \(\{\hat{\pi}^-_n : n \in \mathbb{N}\}\), almost surely converge to the dual paths \(\hat{\pi}^+\) and \(\hat{\pi}^-\) respectively. Fix any bounded continuous function \(g : C(-\infty,0] \times C(-\infty,0] \to \mathbb{R}\). We claim that
\[
\lim_{n \to \infty} E(g(\hat{\pi}^+_n, \hat{\pi}^-_n)) = \lim_{n \to \infty} E(g(\hat{\pi}^+_{(1/n,0)}, \hat{\pi}^-_{(0,0)})) = \gamma_n < -1.
\] (18)

The proof is similar to the proof of Lemma 3.6. For completeness we present here full details. For each \(n \geq 1\), let \(O_n\) denote the event that the set of the points \(\{\hat{\pi}_0(\theta_0 - 1), \hat{\pi}_1(\theta_0 - 1)\}\) equals the set \(\{\hat{\pi}((\sigma_0, \sigma_0)) (\theta_0 - 1), \hat{\pi}((\sigma_n, \sigma_0)) (\theta_0 - 1)\}\). From (17) and from property (g) of \((\mathcal{W}, \hat{\mathcal{W}})\), we have that \(\mathbb{P}(O_n) \to 1\) as \(n \to \infty\). Hence we have
\[
\lim_{n \to \infty} E(g(\hat{\pi}^+_n, \hat{\pi}^-_n)) = \lim_{n \to \infty} E(g(\hat{\pi}^+_n, \hat{\pi}^-_n)1_{O_n}),
\] (19)
where \(1_{O_n}\) denotes indicator function of the event \(O_n\).

We need to define some more events. For \(t > 0\) and \(n \geq 1\), let \(F_t\) denote the event that
\[
\bigcup_{s \in (0,t) \cap \mathbb{Q}} \bigg\{ \text{there exists } \hat{\pi}'_m \in \hat{\mathcal{W}} \text{ with } \sigma_{\hat{\pi}'_m} \geq -s_m \text{ for some } s_m \in (0,s) \cap \mathbb{Q} \text{ such that}
\]
\[
|\hat{\pi}_0(-s_m) - \hat{\pi}'_m(-s_m)| < 1/m \text{ and } \hat{\pi}_0((s - 1) \vee (-t)) \neq \hat{\pi}'_m((s - 1) \vee (-t)).
\]
As \(\hat{\mathcal{W}}\) is compact, the sequence of dual paths \(\{\hat{\pi}'_m : m \in \mathbb{N}\}\) considered in the event \(F_t\) must have a subsequential limit. Hence occurrence of the event \(F_t\) ensures existence of a dual path \(\hat{\pi}_1\) in \(\hat{\mathcal{W}}\) starting from \((\hat{\pi}_0(s_0) - s_0)\) for some \(s_0 < t\) such that \(\hat{\pi}_0((s_0 - 1) \vee (-t)) \neq \hat{\pi}_1((s_0 - 1) \vee (-t))\).

Next for \(t > 0, i \in \mathbb{Z}\) and \(s \in \mathbb{Q}\), we define the following events:
\[
E_t(i/n, s) := \{\hat{\pi}_0(-t) \in [i/n, (i + 1)/n]\};
\]
\[
E_t(i/n, s) := \{\hat{\pi}((\sigma_0, \sigma_n)/n) (s - 1) > \hat{\pi}((\sigma_0, \sigma_n)/(s - 1))\},
\]
where \(\hat{\pi}((\sigma_0, \sigma_n)/n)\) and \(\hat{\pi}((\sigma_0, \sigma_n)/(s - 1))\) are the dual paths in \(\hat{\mathcal{W}}\) starting from the points \((\sigma_0, \sigma_n)/n\) and \((\sigma_0, \sigma_n)/(s - 1))\), respectively.

For \(t > 0\) with a slight abuse of notation let \(\mathcal{G}_t\) denote the \(\sigma\)-field given by
\[
\mathcal{G}_t := \sigma\{\{\hat{\pi}(s) : \hat{\pi} \in \hat{\mathcal{W}} \text{ with } \sigma_{\hat{\pi}} \geq -t \text{ and } s \in [-t, \sigma_{\hat{\pi}}]\}\}
\]
By definition for all \(t > 0\), both the events \(F_t\) and \(E_t(i/n, s)\) are \(\mathcal{G}_t\) measurable. Hence from non-crossing nature of paths in \(\hat{\mathcal{W}}\) it follows that for all \(t > 0\) and for \(t_0^n \in (t, t + 1/n) \cap \mathbb{Q}\), we have the following equality of events
\[
\{\theta_0 = -t\} \cap \{\theta_0^n = -t_0^n\} \cap E_t(i,n) \cap O_n = (F_t)^c \cap \{\hat{\pi}_0(s), \hat{\pi}_1(s) \in [i/n, (i + 1)/n) \text{ for all } s \in [-t_0^n, -t]\} \cap E_t(i,n, -t_0^n) \cap O_n.
\] (20)

For \(t > 0\) and for \(t_0^n \in (t, t + 1/n) \cap \mathbb{Q}\) using (19) and (20) we obtain
\[
\lim_{n \to \infty} E(\hat{\pi}^+_n, \hat{\pi}^-_n)) = \lim_{n \to \infty} E(\hat{\pi}^+_{(1/n,0)}, \hat{\pi}^-_{(0,0)})) = \gamma_n < -1.
\] (21)
Next we observe that conditioned on the event that \((F_i) \cap E_t(i, n, -t_0^n) \cap \{\tilde{\pi}_0(s), \tilde{\pi}_1(s) \in [i/n, (i + 1)/n]\} \) for all \(s \in [-t_0^n, -t]\), the process \(\{(\tilde{\pi}_r^{t_0^n}(s) - s - i/n, \tilde{\pi}_r^{t_0^n}(s) - s - i/n) : s \geq 0\}\) is independent of the \(\sigma\)-field \(\mathcal{G}_n\) and has the same distribution as \((\hat{W}^{(1/n, 0)}, \hat{W}^{(0, 0)})\). Since \(\tilde{\pi}_0(s), \tilde{\pi}_1(s) \in [i/n, (i + 1)/n]\) for all \(s \in [-t_0^n, -t]\), \((F_i) \cap (F_i)\) is \(\mathcal{G}_n\) measurable, from [21] we obtain

\[
= E\left(\sum_{i=-\infty}^{\infty} \lim_{n \to \infty} E\left(g\left(\{\tilde{\pi}_r^{t_0^n}(s) - s - i/n, \tilde{\pi}_r^{t_0^n}(s) - s - i/n) : s \geq 0\}\right)\right)
= E(\sum_{i=-\infty}^{\infty} E\left(g(\hat{W}^{(1/n, 0)}, \hat{W}^{(0, 0)})|\gamma_n < -1)\right) E(\mathbb{P}(E_t(i, n)))
= E(g(\hat{W}^{(1/n, 0)}, \hat{W}^{(0, 0)})|\gamma_n < -1) E(\mathbb{P}(E_t(i, n)))
\]

This completes the proof of [18]. Since \((\tilde{\pi}_n^+, \tilde{\pi}_n^-) \to (\tilde{\pi}_+^+, \tilde{\pi}_-^-)\) almost surely as \(n \to \infty\), from Lemma 4.3 and Corollary 3.5 we have

\[
E(g(\tilde{\pi}_n^+, \tilde{\pi}_n^-)) = \lim_{n \to \infty} E(g(\tilde{\pi}_n^+, \tilde{\pi}_n^-))
= \lim_{n \to \infty} E(g(\hat{W}^{(1/n, 0)}, \hat{W}^{(0, 0)})|\gamma_n < -1)
= \lim_{n \to \infty} E(g(\Gamma_n(\hat{B}_1, \hat{B}_2)))
= E(g(\tilde{\pi}_n^+, \tilde{\pi}_n^-)) = E(g(\Gamma_n(\hat{B}_1, \hat{B}_2))) = E(g(\hat{B}_n^+, \hat{B}_n^-)).
\]

This completes the proof.

The embedding discussed in Proposition 3.3 proves the following corollary:

**Corollary 3.7.** For any \(\epsilon \in (0, 1)\) and for \(\pi_1, \pi_2 \in \mathcal{S}(\hat{B}_n^+, \hat{B}_n^-)\) with \(\sigma_{\pi_1} \sqcup \sigma_{\pi_2} \leq -\epsilon\), almost surely there exists \(s_0 = s_0(\omega) \in (-\epsilon, 0)\) such that \(\pi_1(s) = \pi_2(s)\) for all \(s \geq s_0\).

**Proof:** The proof follows from Proposition 3.3 and from property (a) of \((W, \hat{W})\) which ensures that \(m_n(\tilde{\pi}_0(\theta_0), \theta_0) = 1\) almost surely.

We end this section with the following remark that the construction of \((\mathcal{T}, \hat{\mathcal{T}})\) does not depend on the choice of the deterministic countable dense set that we have started with.

**Remark 3.8.** Regarding the construction of \((\mathcal{T}, \hat{\mathcal{T}})\), it is useful to mention here that instead of taking closure with respect to the ancestor metric, one can take closure of \(\mathcal{S}(\hat{B}_n^+, \hat{B}_n^-)\) in \((\Pi, d_{\Pi})\). The embedding explained in Proposition 3.3 shows that in that case the resulting collection of paths will no longer satisfy the condition of Definition 2.2 as almost surely there will be points with multiple outgoing paths. Nevertheless, the resulting objects almost surely will be \(\mathbb{R}\)-graphs (for a formal definition and basic properties of an \(\mathbb{R}\)-graph, see [7]). For constructing the limiting real tree, one could also have followed the cutting operation of this \(\mathbb{R}\)-graph as in [7]. Similar comment holds true for \(\hat{S}(\hat{B}_n^+, \hat{B}_n^-)\). This further shows that the distribution of \((\mathcal{T}, \hat{\mathcal{T}})\) does not depend on the choice of the deterministic countable dense set of \(\mathbb{R}^2\).

4 Proof of Theorem 1.3

In this section we prove Theorem 1.3. Before embarking on the details, we sketch the idea of the proof. Fontes et al. [12] showed that Schieidgger network and its dual, observed as collection of paths, jointly converge in distribution to the Brownian web and its dual (see Theorem 4.1). We use this joint convergence in the path-space topology, and prove that the coalescing time of the scaled paths also converges to the coalescing time of the limiting coalescing Brownian paths. This enables us to obtain convergence of metric spaces in the Gromov-Hausdorff topology.
We need to introduce some notation. For $(x,t) \in \mathbb{Z}^2_{even}$, taking the edges $(h^{k-1}(x,t), h^k(x,t))$ to be straight line segments for $k \geq 1$, we parametrize the path formed by these edges as the piecewise linear function $\pi(x,t) : [t, \infty) \to \mathbb{R}$ such that $\pi(x,t)(t + k) = h^k(x,t)(1)$ for every integer $k \geq 0$ and and linear in between. Let $\mathcal{X} := \{\pi(x,t) : (x,t) \in \mathbb{Z}^2_{even}\}$ denote the collection of all (forward) paths obtained from $G$. For each $n \geq 1$ and for $\pi \in \mathcal{X}$, the scaled path $\pi_n \in [\sigma_\pi/n, \infty) \to [-\infty, \infty]$ is given by $\pi_n(t) := \pi(nt)/\sqrt{n}$. For each $n \geq 1$, let $\mathcal{X}_n := \{\pi_n(x,t) : (x,t) \in \mathbb{Z}^2_{even}\}$ be the collection of the scaled paths and $\mathcal{X}_n$ is the closure of $\mathcal{X}_n$ in $(\Pi, d_{d1})$.

Similarly we construct a scaled family of backward dual paths obtained from $\overline{G}$. For $(x,t) \in \mathbb{Z}^2_{odd}$, the dual path $\tilde{\pi}(x,t)$ is the piecewise linear function $\tilde{\pi}(x,t) : (-\infty, t] \to \mathbb{R}$ with $\tilde{\pi}(x,t)(t - k) = \tilde{h}^k(x,t)(1)$ for every integer $k \geq 0$. Let $\tilde{\mathcal{X}} := \{\tilde{\pi}(x,t) : (x,t) \in \mathbb{Z}^2_{odd}\}$ be the collection of all possible dual paths admitted by $\overline{G}$. For the backward path $\tilde{\pi}$, the scaled version is $\tilde{\pi}_n : [-\infty, \sigma_{\tilde{\pi}}/n] \to [-\infty, \infty]$ given by $\tilde{\pi}_n(t) = \tilde{\pi}(nt)/\sqrt{n}$ for each $n \geq 1$. Let $\tilde{\mathcal{X}}_n := \{\tilde{\pi}_n(x,t) : (x,t) \in \mathbb{Z}^2_{odd}\}$ be the collection of all the $n$th order diffusively scaled dual paths. The following theorem regarding the convergence of $(\mathcal{X}_n, \mathcal{X}_n)$ to $(\mathcal{W}, \mathcal{W})$ is due to Fontes et. al. \cite{13}.

Theorem 4.1 (Theorem 6.1 of [13]). As $n \to \infty$, $(\mathcal{X}_n, \mathcal{X}_n)$ converges in distribution to the double Brownian web $(\mathcal{W}, \mathcal{W})$ as $(\mathcal{H} \times \bar{\mathcal{H}}, \mathcal{B}_{\mathcal{H} \times \bar{\mathcal{H}}})$ valued random variables.

We will use Theorem 4.1 to prove Theorem 1.3. First we introduce some notation. For $t \in \mathbb{R}$ and $K \subseteq \Pi$, we denote the collection of paths in $\mathcal{K}$ which start before time $t$. Similarly for $K \subseteq \Pi$, let $K^{+t} := \{\pi \in \mathcal{K} : \sigma_{\pi} \geq t\}$ denote the collection of dual paths in $\mathcal{K}$ which start after time $t$. Let $\xi_K \subset \mathbb{R}$ be defined as

$$\xi_K := \{\pi(0) : \pi \in K^{(-1)} \} \subset [0, 1].$$

For $x \in \xi_K$, for $K \subseteq \Pi$ and $\tilde{K} \subseteq \Pi$, let $\tilde{\pi}_r(x,0) = \tilde{\pi}(r)(K \times \tilde{K})$ be defined to be the path $\tilde{\pi} \in \tilde{K}$ with $\sigma_{\tilde{\pi}} = 0$ and such that there is no other path $\tilde{\pi}' \in \tilde{K}^{0+}$ with $x < \tilde{\pi}'(0) < \tilde{\pi}(0)$; if no such $\tilde{\pi} \in \tilde{K}$ exists then we define $\tilde{\pi}_r(x,0)$ to be the backward constant zero function $\tilde{\phi}$ starting at $\sigma_{\tilde{\phi}} = 0$. In other words $\tilde{\pi}_r(x,0)$ is the path of $\tilde{K}^{0+}$ which intersects the $x$-axis to the right of $(x,0)$ and is the closest such path to do so. Similarly we define $\pi_l(x,0)$ as the path of $\pi_l^{0+}$ which intersects the $x$-axis to the left of $(x,0)$ and is the closest such path to do so.

Let $\hat{\gamma}(l, \tau) := \hat{\gamma}(\pi_l(x,0), \pi_r(x,0))$ be the first meeting time of the two backward paths $\pi_l(x,0)$ and $\pi_r(x,0)$ (which can possibly be $-\infty$). Let $\Delta = \Delta_{r,l}(x,0) \subset \mathbb{R}^2$ denote the region enclosed between these two backward paths $\pi_l(x,0)$ and $\pi_r(x,0)$, and formally defined as

$$\Delta_{r,l} := \{(y,s) : s \in \hat{\gamma}(l, \tau), y \in (\pi_l(x,0)(s), \pi_r(x,0)(s))\}.$$ (23)

If $\hat{\gamma}(l, \tau)$ is finite then this region is also bounded.

Let $\mathcal{D} := \{(x,t,n) : (x,t) \in \mathbb{Z}^2_{even}, n \in \mathbb{N}\}$ and

$$\begin{align*}
\mathcal{S}(\pi_l(x,0), \pi_r(x,0)) &= \mathcal{S}(\pi_l(x,0), \pi_r(x,0))_{(K, \tilde{K})} := \{\pi \in K : (\pi, \sigma_{\pi}) \in \Delta_{r,l} \cap \mathcal{D}\}, \\
\mathcal{S}(\pi_l(x,0), \pi_r(x,0)) &= \mathcal{S}(\pi_l(x,0), \pi_r(x,0))_{(K, \tilde{K})} := \{\tilde{\pi} \in \tilde{K} : (\tilde{\pi}, \sigma_{\tilde{\pi}}) \in \Delta_{r,l} \cap \mathcal{D}\}, \\
\mathcal{S}(\pi_l(x,0), \pi_r(x,0)) &= \mathcal{S}(\pi_l(x,0), \pi_r(x,0))_{(K, \tilde{K})} := \mathcal{S}(\pi_l(x,0), \pi_r(x,0))_{(K, \tilde{K})}, \\
\mathcal{S}(\pi_l(x,0), \pi_r(x,0)) &= \mathcal{S}(\pi_l(x,0), \pi_r(x,0))_{(K, \tilde{K})} := \mathcal{S}(\pi_l(x,0), \pi_r(x,0))_{(K, \tilde{K})}. 
\end{align*}$$

the closures being taken in $(\Pi, d_{d1})$ and $(\Pi, d_{d1})$ respectively. We further assume that $(K, \tilde{K})$ is such that, for each $x \in \xi_K$ both the collections, $\mathcal{S}(\pi_l(x,0), \pi_r(x,0))$ and $\mathcal{S}(\pi_l(x,0), \pi_r(x,0))$, are tree-like in the sense of Definition 2.2. The completions of these spaces with respect to the ancestor metric are defined as

$$\mathcal{M}(x,0) := \mathcal{M}(x,0)_{(K, \tilde{K})} := \mathcal{M}(\mathcal{S}(\pi_l(x,0), \pi_r(x,0)))$$ and

$$\mathcal{M}(x,0) := \mathcal{M}(x,0)_{(K, \tilde{K})} := \mathcal{M}(\mathcal{S}(\pi_l(x,0), \pi_r(x,0))).$$
We further assume that for each \( x \in \xi_K \), both these metric spaces, \( \mathcal{M}^{(x,0)} \) and \( \widetilde{\mathcal{M}}^{(x,0)} \), are compact and let \( \phi(\mathcal{M}^{(x,0)}) \) and \( \phi'(\widetilde{\mathcal{M}}^{(x,0)}) \) denote the isometric embeddings of these metric spaces \( \mathcal{M}^{(x,0)} \) and \( \widetilde{\mathcal{M}}^{(x,0)} \) into \( \mathcal{M} \), the metric space of isometric equivalence classes of compact metric spaces endowed with the Gromov-Hausdorff metric. Fix \( f : \mathcal{M} \times \mathcal{M} \to \mathbb{R} \), a bounded continuous real valued function and define

\[
\kappa_{(K, \hat{K})}(f) := \begin{cases} 
0 & \text{if } \xi_K = \emptyset \\
\sum_{x \in \xi_K \cap [0,1]} f(\phi(\mathcal{M}^{(x,0)}), \phi'(\widetilde{\mathcal{M}}^{(x,0)})) & \text{if } 0 < \#\xi_K < \infty \\
\infty & \text{otherwise.}
\end{cases}
\]

For the rest of the section we make two specific choices of \((K, \hat{K})\), namely, \((\mathcal{W}, \hat{\mathcal{W}})\) and \((\mathcal{X}_n, \hat{\mathcal{X}}_n)\). Also to simplify notation we write \( \xi := \xi_{\mathcal{W}} \), \( \xi_n := \xi_{\mathcal{X}_n} \), \( \eta := \#\xi \) and \( \eta_n := \#\xi_n \).

It is well known that \( E(\eta) = 1/\sqrt{\tau} \) (see, e.g., Proposition 2.7 of \[28\]).

From the properties of \((\mathcal{W}, \hat{\mathcal{W}})\) it follows that, for each \( x \in \xi \), there exist two dual paths \( \hat{\pi}^{(x,0)}_t, \hat{\pi}^{(x,0)}_l \in \hat{\mathcal{W}} \) both starting at \((x,0)\) such that \( \hat{\pi}^{(x,0)}(1) > \hat{\pi}^{(x,0)}(1) \) with \( \hat{\gamma}(r,l) \) finite. Hence the region \( \Delta_{x,l}(x,0) \) is almost surely nonempty and bounded. From the properties of \((\mathcal{W}, \hat{\mathcal{W}})\) we have that both the path spaces, \( \mathcal{S}((\hat{\pi}^{(x,0)}_t, \hat{\pi}^{(x,0)}_l)) \) and \( \hat{\mathcal{S}}((\hat{\pi}^{(x,0)}_t, \hat{\pi}^{(x,0)}_l)) \), are tree-like in the sense of Definition 2.2. Same argument as in Proposition 3.4 shows that both the metric spaces, \( \mathcal{M}^{(x,0)}_{(\mathcal{W}, \hat{\mathcal{W}})} \) and \( \widetilde{\mathcal{M}}^{(x,0)}_{(\mathcal{W}, \hat{\mathcal{W}})} \), are compact as well. Similarly for \( n \geq 1 \) and for each \( x_n \in \xi_n \), the region \( \Delta_{x,l}(x_n,0) \) is almost surely nonempty and bounded. From the construction it follows that, both the families, \( \mathcal{S}((\hat{\pi}^{(x_n,0)}_t, \hat{\pi}^{(x_n,0)}_l)) \) and \( \hat{\mathcal{S}}((\hat{\pi}^{(x_n,0)}_t, \hat{\pi}^{(x_n,0)}_l)) \), satisfy the condition of Definition 2.2 and both these metric spaces, \( \mathcal{M}^{(x_n,0)} \) and \( \widetilde{\mathcal{M}}^{(x_n,0)} \), are compact.

Hence the random variables \( \kappa(f) := \kappa_{(\mathcal{W}, \hat{\mathcal{W}})}(f) \), and \( \kappa_n(f) := \kappa_{(\mathcal{X}_n, \hat{\mathcal{X}}_n)}(f) \) are well defined almost surely. In the next subsection we calculate \( E(\kappa(f)) \).

### 4.1 Calculation of \( E(\kappa(f)) \)

The aim of this subsection is to prove the following proposition which calculates \( E(\kappa(f)) \).

**Proposition 4.2.** \( E(\kappa(f)) = E(f(\phi(T), \phi'(\hat{T}))) / \sqrt{\tau} \).

To prove this proposition we need to prove a limit result for conditional continuum random trees. Recall that \( \widehat{\mathcal{W}}^{(0,0)} \) and \( \widehat{\mathcal{W}}^{(1/n,0)} \), are two independent coalescing backward Brownian motions starting from the points \((0,0)\) and \((1/n,0)\) respectively with their coalescing time \( \hat{\gamma}_n = \hat{\gamma}(\widehat{\mathcal{W}}^{(1/n,0)}, \widehat{\mathcal{W}}^{(0,0)}) \). Let \( \Delta(\widehat{\mathcal{W}}^{(0,0)}, \widehat{\mathcal{W}}^{(1/n,0)}) \) denote the random region enclosed between these two backward Brownian paths, defined as

\[
\Delta(\widehat{\mathcal{W}}^{(1/n,0)}, \widehat{\mathcal{W}}^{(0,0)}) := \{(x,t) : t \in [\hat{\gamma}_n, 0), x \in (\widehat{\mathcal{W}}^{(0,0)}(t), \widehat{\mathcal{W}}^{(1/n,0)}(t)]\}.
\]

Following the construction of \((T, \hat{T})\), we do a similar construction here. We start coalescing backward Brownian motions from all points of \( \Delta(\widehat{\mathcal{W}}^{(1/n,0)}, \widehat{\mathcal{W}}^{(0,0)}) \cap \mathbb{Q}^2 \) such that these backwards paths coalesce with the boundary of \( \Delta(\widehat{\mathcal{W}}^{(1/n,0)}, \widehat{\mathcal{W}}^{(0,0)}) \) as soon as they intersect the boundary. This collection of coalescing backward paths is denoted by \( \hat{\mathcal{S}}(\widehat{\mathcal{W}}^{(1/n,0)}, \widehat{\mathcal{W}}^{(0,0)}) \). From earlier discussions it follows that for each \( (x,t) \in \Delta(\widehat{\mathcal{W}}^{(1/n,0)}, \widehat{\mathcal{W}}^{(0,0)}) \cap \mathbb{Q}^2 \), there exists a forward path \( \pi^{(s,t)} \in \Pi \) which does not cross any dual path in \( \hat{\mathcal{S}}(\widehat{\mathcal{W}}^{(1/n,0)}, \widehat{\mathcal{W}}^{(0,0)}) \) and defined as

\[
\pi^{(s,t)}(s) := \begin{cases} 
\sup\{y : y \in \mathbb{Q}, \hat{\pi}^{(s,t)}(s) \in \hat{\mathcal{S}}(\widehat{\mathcal{W}}^{(1/n,0)}, \widehat{\mathcal{W}}^{(0,0)}), \hat{\pi}^{(s,t)}(t) < x\} &= \infty \\
\inf\{y : y \in \mathbb{Q}, \hat{\pi}^{(s,t)}(s) \in \hat{\mathcal{S}}(\widehat{\mathcal{W}}^{(1/n,0)}, \widehat{\mathcal{W}}^{(0,0)}), \hat{\pi}^{(s,t)}(t) > x\} &= \pi^{(s,t)}(0) \text{ for } s \in [t,0) \cap \mathbb{Q} \text{ for } s \geq 0.
\end{cases}
\]
We denote this collection of forward paths starting from all the points of $\Delta(\hat{x})$ by $S(\hat{W}^{(1/n,0)}, \hat{W}^{(0,0)})$. Because of property (d) of $(W, \hat{W})$, the set $\{\pi(0) : \pi \in S(\hat{W}^{(1/n,0)}, \hat{W}^{(0,0)}), \sigma_\pi \leq -1\}$ is almost surely finite. Let $x^n = x^n(\omega) \in (0,1/n)$ be defined as
\[ x^n := \min\{x \in (0,1/n) : \text{there exists } \pi \in S(\hat{W}^{(1/n,0)}, \hat{W}^{(0,0)}) \text{ with } \sigma_\pi \leq -1 \text{ and } \pi(0) = x\}. \]
Next we consider all the paths in $S(\hat{W}^{(1/n,0)}, \hat{W}^{(0,0)})$ which pass through $x^n$. In other words we consider a new collection of coalescing paths $S'(\hat{W}^{(1/n,0)}, \hat{W}^{(0,0)})$ defined as $S'(\hat{W}^{(1/n,0)}, \hat{W}^{(0,0)}) := \{\pi \in S(\hat{W}^{(1/n,0)}, \hat{W}^{(0,0)}) : \pi(0) = x^n\}$. By construction, both $S'(\hat{W}^{(1/n,0)}, \hat{W}^{(0,0)})$ and $\hat{S}(\hat{W}^{(1/n,0)}, \hat{W}^{(0,0)})$ are tree-like in the sense of Definition 2.2. For $n \geq 1$ we consider the pair of complete metric spaces
\[ (T_n, \hat{T}_n) := (\mathcal{M}(S'(\hat{W}^{(1/n,0)}, \hat{W}^{(0,0)})), \mathcal{M}(\hat{S}(\hat{W}^{(1/n,0)}, \hat{W}^{(0,0)}))). \]
Similar argument as in the proof of Proposition 3.4 shows that these metric spaces are also compact almost surely. It is useful to observe that for each $n \geq 1$, the metric space $T_n$ does not depend on how the forward paths evolve after time 0.

**Lemma 4.3.**
\[ (T_n, \hat{T}_n)\{\gamma_n < -1\} \Rightarrow (\mathcal{T}, \hat{\mathcal{T}}) \text{ as } n \to \infty, \]
where convergence in distribution holds with respect to the Gromov-Hausdorff topology.

Assuming the above lemma we first calculate $\mathbb{E}(\kappa(f))$ and complete the proof of Proposition 4.2.

**Proof of Proposition 4.2:** Let $I_n \subset \{0,1,\ldots,n−1\}$ be given by
\[ I_n := \{i : 0 \leq i \leq n−1 \text{ such that } \pi(i/n,0), \pi((i+1)/n,0) \in \hat{W} \text{ with } \pi(i/n,0)(−1) < \pi((i+1)/n,0)(−1)\}. \]
For $i \in I_n$, let $\Delta(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0))$ denote the random region enclosed by the two backward Brownian paths $\hat{\pi}(i+1/n,0)$ and $\hat{\pi}(i/n,0)$ and let
\[ S(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0)) := \{\pi \in W : (\pi(\sigma_\pi), \sigma_\pi) \in \Delta(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0)) \cap \mathbb{Q}^2\} \quad \text{and} \quad \hat{S}(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0)) := \{\hat{\pi} \in \hat{W} : (\hat{\pi}(\sigma_\hat{\pi}), \sigma_\hat{\pi}) \in \Delta(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0)) \cap \mathbb{Q}^2\}. \]
For $i \in I_n$, let $x^n_i = x^n_i(\omega) \in (i/n, (i+1)/n)$ be defined as
\[ x^n_i := \min\{x \in (i/n, (i+1)/n) : \text{there exists } \pi \in S(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0)) \text{ with } \sigma_\pi \leq -1 \text{ and } \pi(0) = x\}. \]
We recall that for each $i \in I_n$, the set $\{\pi(0) : \pi \in W^{(i+1,n−1)}, \pi(0) \in (i/n, (i+1)/n)\}$ is nonempty and finite and hence $x^n_i$ is well defined. Next we define the coalescing family of forward paths passing through $x^n_i$ given as
\[ S'(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0)) := \{\pi \in S(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0)) : \pi(0) = x^n_i\}. \]
From the properties of $(W, \hat{W})$ it follows that both these collections of paths, $S'(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0))$ and $\hat{S}(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0))$, are tree-like and we consider the complete metric spaces
\[ \mathcal{M}(S'(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0))) \quad \text{and} \quad \mathcal{M}(\hat{S}(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0))). \]
For $i \in I_n$, the paths in $S'(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0))$ do not cross the dual paths in $\hat{S}(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0))$ and $(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0))$ is distributed as coalescing backward Brownian motions starting from $(i+1)/n, 0)$ and $(i/n, 0)$ conditioned to meet after time $-1$. Hence for each $i \in I_n$ we have
\[ (T_n, \hat{T}_n)\{\gamma_n < -1\} \overset{d}{=} (\mathcal{M}(S'(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0))), \mathcal{M}(\hat{S}(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0)))), \]
Fix a bounded continuous function $g : \mathbb{M} \times \mathbb{M} \to \mathbb{R}$. We define
\[ \mathcal{R}_n(g) := \sum_{i \in I_n} g(\phi_i/n(\mathcal{M}(S'(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0)))), \phi_i/n(\mathcal{M}(\hat{S}(\hat{\pi}(i+1/n,0), \hat{\pi}(i/n,0)))). \]
where \( \phi_{ij/n} \) and \( \phi_{ij/n}' \) are isometric embeddings of the respective metric spaces into \( \mathcal{M} \), the space of all isometry equivalence classes of compact metric spaces. We want to show that

\[
\mathcal{R}_n(g) \to \sum_{x \in \xi} g(\phi_x(M^{(x,0)}_{(W,\tilde{W})}), \phi_x'(\tilde{M}^{(x,0)}_{(W,\tilde{W})})) \text{ as } n \to \infty \text{ almost surely},
\]

where \( \phi_x, \phi_x' \) are the corresponding embeddings of the respective metric spaces into \( \mathcal{M} \).

We first show that \( \mathcal{R}_n \to \mathcal{R} = \eta \) as \( n \to \infty \). Choose \( n_0 = n_0(\omega) \) such that for all \( x, y \in \xi \) and for \( n \geq n_0 \) we have \( |x - y| > 2n \). For each \( x \in \xi \), set \( l^+_{n} = \lfloor nx \rfloor/n \) and \( r^+_{n} = l^+_{n} + (1/n) \). From the non-crossing nature of paths in \( \tilde{W} \) it follows that for each \( x \in \xi \), we have \( \hat{\pi}^{(r^+_{n},0)}(1) > \hat{\pi}^{(r^+_n,0)}(1) \) and hence \( l^+_{n} \in I_n \). Hence \( \#I_n \leq \xi \) for all \( n \geq n_0 \). On the other hand for each \( i \in I_n \), choose \( y_i \) such that \( y_i \in (\hat{\pi}^{(r^+_{n},0)}(-1), \hat{\pi}^{(r^+_n,0)}(-1)) \) and consider the path \( \pi^{y_i,-1} \) in \( \tilde{W} \). Then we must have \( \hat{\pi}(y_i,1) = x_i \in \xi \). This implies that \( \mathcal{R}_n \leq \mathcal{R} \) for all \( n \). Hence for all \( n \geq n_0 \) we have \( \mathcal{R}_n = \mathcal{R} \).

This ensures that for all large \( n \),

\[
\mathcal{R}_n(g) = \sum_{x \in \xi} g(\phi_x(\mathcal{M}(\hat{\pi}^{(r^+_n,0)}, \hat{\pi}^{(l^+_{n},0)})), \phi_x'(\hat{\mathcal{M}}(\hat{\pi}^{(r^+_n,0)}, \hat{\pi}^{(l^+_{n},0)}))).
\]  

(26)

Choose \( n_1 = n_1(\omega) \geq n_0 \) large enough such that for all \( n \geq n_1 \) we have \( \mathcal{R}_n = \eta \) and consequently for all \( n \geq n_1 \) and for each \( x \in \xi \) we have

\[
\mathcal{M}(\hat{\pi}^{(r^+_n,0)}, \hat{\pi}^{(l^+_{n},0)}) = \mathcal{M}^{(x,0)}_{(W,\tilde{W})}.
\]  

(27)

We recall that there are exactly two dual paths \( \hat{\pi}_i^{(r^+_n,0)} \) and \( \hat{\pi}_r^{(r^+_n,0)} \) starting from \( (x,0) \) with \( \hat{\pi}_i^{(r^+_n,0)}(-1) > \hat{\pi}_r^{(r^+_n,0)}(-1) \). Hence from property (f) of \( (W,\tilde{W}) \) it follows that, \( \{ \hat{\pi}_i^{(l^+_{n},0)} : n \in \mathbb{N} \} \) and \( \{ \hat{\pi}_r^{(r^+_n,0)} : n \in \mathbb{N} \} \) converge to \( \hat{\pi}_i^{(x,0)} \) and \( \hat{\pi}_r^{(x,0)} \) respectively in \( (\hat{\Pi}, d_{\hat{\Pi}}) \) as \( n \to \infty \). Let \( \hat{\gamma}(\hat{\pi}^{(x,0)}, \hat{\pi}^{(l^+_{n},0)}) \) denote the coaliscent time of the two dual paths \( \hat{\pi}_i^{(l^+_{n},0)} \) and \( \hat{\pi}_r^{(l^+_{n},0)} \). Similarly the coaliscent time of the two dual paths \( \hat{\pi}_i^{(r^+_n,0)} \) and \( \hat{\pi}_r^{(r^+_n,0)} \) is given by \( \hat{\gamma}(\hat{\pi}_i^{(r^+_n,0)}, \hat{\pi}_r^{(r^+_n,0)}) \). For all \( n \geq n_0 \) and for each \( x \in \xi \), the metric space \( \hat{\mathcal{M}}^{(x,0)} \) is naturally embedded into \( \mathcal{M}(\hat{\pi}^{(r^+_n,0)}, \hat{\pi}^{(l^+_{n},0)}) \). Using this embedding we obtain that for all \( n \geq n_1 \) and for each \( x \in \xi \) we have

\[
d_{GH}(\mathcal{M}(\hat{\pi}^{(r^+_n,0)}, \hat{\pi}^{(l^+_{n},0)}), \hat{\mathcal{M}}^{(x,0)}_{(W,\tilde{W})}) \leq 2(\hat{\gamma}(\hat{\pi}_i^{(x,0)}, \hat{\pi}^{(l^+_{n},0)}) \vee \hat{\gamma}(\hat{\pi}_r^{(x,0)}, \hat{\pi}^{(r^+_n,0)})).
\]  

(28)

Now because of property (g) of \( (W,\tilde{W}) \), we have that \( \hat{\gamma}(\hat{\pi}_i^{(x,0)}, \hat{\pi}^{(l^+_{n},0)}), \hat{\gamma}(\hat{\pi}_r^{(x,0)}, \hat{\pi}^{(r^+_n,0)}) \to (0,0) \) as \( n \to \infty \). Since \( g \) is continuous on the product metric space and the set \( \xi \) is finite, \( (26), (27) \) and \( (28) \) show that

\[
\mathcal{R}_n(g) \to \sum_{x \in \xi} g(\phi_x(M^{(x,0)}_{(W,\tilde{W})}), \phi_x'(\tilde{M}^{(x,0)}_{(W,\tilde{W})})) \text{ as } n \to \infty.
\]

As \( g \) is bounded and \( \mathbb{E}[\eta] < \infty \), the family \( \{ \mathcal{R}_n(g) : n \in \mathbb{N} \} \) is uniformly integrable and hence we have

\[
\lim_{n \to \infty} \mathbb{E}[\mathcal{R}_n(g)] = \mathbb{E}\left(\sum_{x \in \xi} g(\phi_x(M^{(x,0)}_{(W,\tilde{W})}), \phi_x'(\tilde{M}^{(x,0)}_{(W,\tilde{W})}))\right).
\]

From \( (25) \) and for \( B \in [0, \infty) \) denoting a standard Brownian motion we have,

\[
\lim_{n \to \infty} \mathbb{E}(\mathcal{R}_n(g)) = \mathbb{E}(g(T, \tilde{T})) = \mathbb{E}(g(T, \tilde{T})|\hat{\gamma}_n < 1)\mathbb{P}(\hat{\gamma}_n < 1)
\]

\[
= \lim_{n \to \infty} \mathbb{E}(g(T_n, \tilde{T}_n)|\hat{\gamma}_n < 1)\mathbb{P}(\hat{\gamma}_n < 1) + \lim_{n \to \infty} \mathbb{E}(g(T_n, \tilde{T}_n)|\hat{\gamma}_n < 1)\mathbb{P}(1/n + \min_{t \in [0,1]} \sqrt{2}B(t) > 0)
\]

\[
= \lim_{n \to \infty} \mathbb{E}(g(T_n, \tilde{T}_n)|\hat{\gamma}_n < 1)n(2\Phi(1/\sqrt{2n}) - 1)
\]

\[
= \mathbb{E}(g(T, \tilde{T}))/\sqrt{\pi}
\]
where \( \Phi \) is the distribution function of a standard normal random variable. The last step follows from Lemma 4.3. This completes the proof. \( \square \)

**Proof of Lemma 4.3:** Recall that \( \hat{B}_1 \) and \( \hat{B}_2 \) are two independent backward Brownian motions both starting from the origin. We observe that \((\hat{B}_1,\hat{B}_2) \in \mathbb{A}_n \) almost surely for all \( n \geq 1 \), where \( \mathbb{A}_n \) is as in (12). The proof of Corollary 3.5 shows that \( (\lambda_1^n, \lambda_2^n) \to (\lambda_1, \lambda_0) \) as \( n \to \infty \) almost surely. Choose \( n_0 = n_0(\omega) \) so that for all \( n \geq n_0 \) we have \( \lambda_i^n \in (\lambda_1, \lambda_0) \) and consequently \( \lambda_1^n = \lambda_1 \).

For all \( n \geq n_0 \), the set \( \{ \pi(\lambda_0^n - \lambda_0) : \pi \in S(\hat{B}_1^+, \hat{B}_2^-), \sigma_\pi \leq (\lambda_0^n - \lambda_0) - 1 \} \) is nonempty and finite. We take \( y^n \in \mathbb{R} \) as \( y^n := \min \{ \pi(\lambda_0^n - \lambda_0) : \pi \in S(\hat{B}_1^+, \hat{B}_2^-), \sigma_\pi \leq (\lambda_0^n - \lambda_0) - 1 \} \). Next we consider the coalescing families of forward and backward paths defined by

\[
S_n(\hat{B}_1^+, \hat{B}_2^-) := \{ \pi \in S_n(\hat{B}_1^+, \hat{B}_2^-) : (\lambda_0^n - \lambda_0) - 1 \}
\]
and

\[
\hat{S}_n(\hat{B}_1^+, \hat{B}_2^-) := \{ \bar{\pi} \in \hat{S}_n(\hat{B}_1^+, \hat{B}_2^-) : \sigma_{\bar{\pi}} \leq (\lambda_0^n - \lambda_0) \}.
\]

Let \( \mathcal{M}(S_n(\hat{B}_1^+, \hat{B}_2^-)) \) and \( \mathcal{M}(\hat{S}_n(\hat{B}_1^+, \hat{B}_2^-)) \) denote the corresponding complete metric spaces where completions are taken with respect to the ancestor metric. We observe that both these metric spaces are naturally embedded into the metric spaces \( \mathcal{M}(S(\hat{B}_1, \hat{B}_2^+)) = T \) and \( \mathcal{M}(\hat{S}(\hat{B}_1^+, \hat{B}_2^-)) = \hat{T} \) respectively. We show that almost surely

\[
d_{G\hat{H}}(\mathcal{M}(S_n(\hat{B}_1^+, \hat{B}_2^-)), T) \lor d_{G\hat{H}}(\mathcal{M}(\hat{S}_n(\hat{B}_1^+, \hat{B}_2^-)), \hat{T}) \to 0, \quad (29)
\]
as \( n \to \infty \). Fix \( \epsilon > 0 \) and choose \( n_1 = n_1(\omega) \geq n_0 \) such that \( (\lambda_0^n - \lambda_0) < \epsilon/2 \) for all \( n \geq n_1 \). We first observe that the natural embedding shows that \( d_{G\hat{H}}(\mathcal{M}(S_n(\hat{B}_1^+, \hat{B}_2^-)), \hat{T}) \leq 2(\lambda_0^n - \lambda_0) < \epsilon \) for all \( n \geq n_1 \).

In order to get an upper bound for \( d_{G\hat{H}}(\mathcal{M}(S_n(\hat{B}_1^+, \hat{B}_2^-)), T) \), from Remark 3.8 we obtain that for all \( \pi_1, \pi_2 \in S(\hat{B}_1^+, \hat{B}_2^-) \) with \( \sigma_{\pi_1} \lor \sigma_{\pi_2} \leq -\epsilon/2 \) there exists \( s_0 = s_0(\omega) \in (0, \epsilon/2) \) which does not depend on the choice of \( \pi_1, \pi_2 \) such that for all \( s \geq -s_0 \) we have \( \pi_1(s) = \pi_2(s) \). Choose \( \bar{n}_2 = \bar{n}_2(\omega) \geq n_1 \) such that \( 1/n_2 < s_0 \). Then for all \( \pi \in S(\hat{B}_1^+, \hat{B}_2^-) \) with \( \sigma_{\pi} \leq -\epsilon/2 \) we have \( \pi \in S_n(\hat{B}_1^+, \hat{B}_2^-) \) as well for all \( n \geq n_2 \). Hence we have \( d_{G\hat{H}}(\mathcal{M}(S_n(\hat{B}_1^+, \hat{B}_2^-)), T) \leq \epsilon \) for all \( n \geq n_2 \). This proves (29).

Finally same argument as in Lemma 3.6 shows that \( (T_n, \hat{T}_n) |_{\{ \hat{\gamma}_n < -1 \}} \) converges in distribution to \( (T, \hat{T}) \) as \( n \to \infty \). \( \square \)

### 4.2 Convergence of \( \mathbb{E}(\kappa_n(f)) \)

In this subsection we prove that \( \mathbb{E}(\kappa_n(f)) \to \mathbb{E}(\kappa(f)) \) as \( n \to \infty \) and use it to prove Theorem 4.3. In order to prove the above mentioned convergence, we first need to show that \( \mathbb{E}(\eta_n) \to \mathbb{E}(\eta) \) as \( n \to \infty \). For \( t_1 < t_2 \) let

\[
\xi(t_1, t_2) := (\pi(t_2) : \pi \in \mathcal{W}^{t_1 t_2}, \pi(t_2) \in [0, 1]) \quad and \quad \eta(t_1, t_2) := \#(\xi, t_2),
\]

\[
\xi_n(t_1, t_2) := (\pi(t_2) : \pi \in \mathcal{X}^{t_1 t_2}_n, \pi(t_2) \in [0, 1]) \quad and \quad \eta_n(t_1, t_2) := \#(\xi_n, t_2).
\]

Below we prove a continuity property of \( \eta(s, 0) : s < 0 \), which will be used in proving convergence of \( \mathbb{E}(\eta_n) \).

**Corollary 4.4:** We have \( \lim_{s \to -1} \eta(s, 0) = \eta \) a.s.

**Proof:** For \(-1 \leq s_1 < s_2 < 0 \) we have \( \xi(s_1, 0) \subseteq \xi(s_2, 0) \) and consequently \( \eta(s_1, 0) \leq \eta(s_2, 0) \). Because of this monotonicity \( \lim_{s \to -1} \eta(s, 0) \geq \eta \). If \( \lim_{s \to -1} \eta(s, 0) > \eta \), then there exists \( x \in \xi((1-1/2, 0) \setminus \xi \) and a sequence of paths \( \{ \pi^n \in \mathcal{W} \} \) with \( \pi^n(0) = x \) for all \( n \) such that \( -1 < \sigma_{\pi^n} \leq -1/2 \) and \( \sigma_{\pi^n} \) decreases to \(-1 \) as \( n \to \infty \). By compactness of \( \mathcal{W} \), it follows that there exists a convergent subsequence \( \{ \pi^n_k : k \in \mathbb{N} \} \) and \( \pi \in \mathcal{W} \) such that \( \pi^n_k \to \pi \) in \( (\Omega, d_H) \) as \( k \to \infty \). Since convergence in \( (\Omega, d_H) \) implies convergence of starting times also, we must have \( \sigma_{\pi} = -1 \) and \( \pi(0) = x \) which contradicts the choice of \( x \) and the proof follows. \( \square \)

Next we use Corollary 4.4 to prove convergence of \( \mathbb{E}(\eta_n) \).
Lemma 4.5. $\mathbb{E}[\eta_n] \to \mathbb{E}[\eta]$ as $n \to \infty$.

Proof: We show that $\eta_n$ converges in distribution to $\eta$ as $n \to \infty$ and the sequence $\{\eta_n : n \in \mathbb{N}\}$ is uniformly integrable as well. Recall from Theorem [13] that $(\tilde{\mathcal{X}}_n, \tilde{\mathcal{X}}_n) \Rightarrow (\mathcal{W}, \tilde{\mathcal{W}})$ as $n \to \infty$. Using Skorohod's representation theorem we assume that we are working on a probability space such that

$$d_{\mathcal{H} \times \tilde{\mathcal{H}}}(\tilde{\mathcal{X}}_n, \tilde{\mathcal{X}}_n, (\mathcal{W}, \tilde{\mathcal{W}})) \to 0,$$

almost surely as $n \to \infty$. Same argument as in Lemma 3.2 of [25] shows that $\eta_n$ converges to $\eta$ almost surely as $n \to \infty$. For completeness we present the proof here also. First we show that, for all $k \geq 0$,

$$\liminf_{n \to \infty} \mathbf{1}_{\{\eta_n \geq k\}} \geq \mathbf{1}_{\{\eta \geq k\}}$$

almost surely. Indeed, for $k = 0$, both $\mathbf{1}_{\{\eta_n \geq k\}}$ and $\mathbf{1}_{\{\eta \geq k\}}$ equal 1. We need to show [30] for $k \geq 1$.

From Properties of $(\mathcal{W}, \tilde{\mathcal{W}})$ it follows that $\eta$ is finite and $\xi \subset (0,1)$ almost surely. Hence we can choose $0 < \epsilon = \epsilon(\omega)$ such that

(i) for all $x, y \in \xi$ with $x \neq y$ we have $(x - \epsilon, x + \epsilon) \subset (0,1)$ and $|x - y| > 2\epsilon$;

(ii) for all $x \in \xi$ there exists $\pi \in \mathcal{W}$ with $\sigma_{\pi} \leq -1 - \epsilon$ and $\pi(0) = x$.

Let $m = m(\omega)$ be such that

$$(y, t) : \pi(0,0)(t) \leq y \leq \pi(1,0)(t), t \in [-1,0], \pi(0,0), \pi(1,0) \in \tilde{\mathcal{W}} \subset [-m, m]^2. \quad (31)$$

Choose $n_0 = n_0(\omega)$ such that for all $n \geq n_0$, $d_{\mathcal{H} \times \tilde{\mathcal{H}}}(\mathcal{W}, \tilde{\mathcal{W}}, (\tilde{\mathcal{X}}_n, \tilde{\mathcal{X}}_n)) < g(\epsilon, m)$ where $g(\epsilon, m)$ is as in Remark 3.1. By the choice of $n_0$, it follows that for each $\pi \in \mathcal{W}^{(1-\epsilon)}$ with $\pi(0) \in [0,1]$ and for all $n \geq n_0$, there exists $\pi^n \in \mathcal{X}_n$ with $\sigma_{\pi^n} \leq -1$ and $d_{\mathcal{H}}(\pi, \pi^n) < g(\epsilon, m)$. Since $|\pi(0) - \pi^n(0)| < \epsilon$, from the choice of $\epsilon$ we must have $\pi^n(0) \in (0,1)$. Thus $\eta_n(\omega) \geq k$ for all $n \geq n_0$. This proves [30].

Finally we need to show that $\mathbb{P}(\limsup_{n \to \infty} \{\eta_n > \eta\}) = 0$. This is equivalent to showing that $\mathbb{P}(\Omega_0^k) = 0$ for all $k \geq 0$, where

$$\Omega_0^k := \{\omega : \eta_n(\omega) > \eta(\omega) = k \text{ for infinitely many } n\}.$$

Consider $k = 0$ first. Using Corollary 4.4 we can obtain $l_0 := l_0(\omega) \in (-1,0)$ such that $\eta(l_0, 0)(\omega) = \eta(\omega)$. Since both the points $(0,0)$ and $(1,0)$ are of type $(0,1)$ and on the event $\{\eta = 0\}$ the set $\xi(l_0, 0) = \xi = 0$, we can obtain $\epsilon := \epsilon(\omega) \in (0,1+l_0)$ such that for all $\pi \in \mathcal{W}(\omega)$ with $\sigma_{\pi} \leq -1 + l_0$, $\pi(0) \notin (-\epsilon, 1+\epsilon)$. On the event $\Omega_0^0$, choose large enough $n$ such that $\eta_n(\omega) > 0$ and $d_{\mathcal{H} \times \tilde{\mathcal{H}}}(\mathcal{W}, \tilde{\mathcal{W}}, (\tilde{\mathcal{X}}_n, \tilde{\mathcal{X}}_n)) < g(\epsilon, m)$ where $m = m(\omega) > 0$ is chosen as in [31] and $g(\epsilon, m)$ is as in Remark 3.1. From the choice of $\epsilon$, it follows that there exists $\pi \in \mathcal{W}(\omega)$ with $\sigma_{\pi} \leq -1 + \epsilon$ and $\pi(0) \in (-\epsilon, 1+\epsilon)$. This gives a contradiction. Hence we have $\mathbb{P}(\Omega_0^0) = 0$.

For $k > 0$, on the event $\Omega_0^k$, we show that a forward path $\pi \in \mathcal{W}$ coincides with a dual path $\hat{\pi} \in \tilde{\mathcal{W}}$ for a positive time which leads to a contradiction. From Corollary 4.4 it follows that, we can choose $-1 < l_0(\omega) < -1/2$ such that $\xi(l_0, 0)(\omega) = \xi(\omega)$. From property (a) of $(\mathcal{W}, \tilde{\mathcal{W}})$ it further follows that for any $x \in \xi$, the type of the point $(x, 0)$ is $(1,1)$. Hence for each $x \in \xi$ there exists $s_x = s_x(\omega) \in (-1,0)$ such that for any $\pi_1, \pi_2 \in \mathcal{W}^{l_0}$ with $\pi_1(0) = \pi_2(0) = x$, we have $\pi_1(s) = \pi_2(s)$ for all $s \geq s_x$. Set $s_0 := \max\{s_x : x \in \xi\}$ and observe that $s_0 \geq 0$ almost surely. By definition, for all $s \geq s_0$ we have $\#\{\pi(s_0) : \pi \in \mathcal{W}^{l_0}, \pi(1) \in [0,1]\} = k$. This gives us that $\xi(l_0, 0)(\omega) = \xi(\omega) = \xi(l_0, s_0)(\omega)$, i.e., the paths leading to any single point considered in $\xi(l_0, 0)$ have coalesced before time $s_0$. We consider $k$ paths contributing to $\xi(l_0, 0)$, viz., $\pi_1, \ldots, \pi_k$ such that $\pi_1(0), \ldots, \pi_k(0) = \xi(l_0, 0)$. Choose $0 < \epsilon = \epsilon(\omega) < \min\{s_0, 1 + l_0\}/3$ such that

(i) $(x - \epsilon, x + \epsilon) \subset (0,1)$ for all $x \in \xi(l_0, 0)$;
(ii) The $\epsilon$-tubes around $\pi_1, \ldots, \pi_k$, given by

$$T_i^\epsilon := \{(x, t) : |x - \pi_i(t)| < \epsilon, s_0 \leq t \leq 0\}$$

are disjoint.

Choose $m = m(\omega) > 0$ as in (31). Set $n_0 = n_0(\omega)$ such that for all $n \geq n_0$, $\xi_{n_0}(\omega) > k$ and

$$d_{\mathcal{H} \times \hat{R}}((W, \hat{W})(\omega), (\hat{X}_n, \hat{\pi}_n)(\omega)) < g(\epsilon, m)$$

where $g(\epsilon, m)$ is as in Remark 3.1. By the choice of $n_0$, it follows that one of the $k$ tubes must contain at least two paths, $\pi_1^{n_0}, \pi_2^{n_0}$ (say) of $\mathcal{A}_n^{(-1)-}$ which do not coalesce by time 0. From the construction of dual paths it follows that there exist at least one dual path $\hat{\pi}^{n_0} \in \hat{\pi}_{n_0}$ lying between $\pi_1^{n_0}$ and $\pi_2^{n_0}$ for $t \in [-1, 0]$ and hence we must have an approximating $\hat{\pi} \in \hat{W}$ close to $\hat{\pi}^{n_0}$ for $t \in [-1, -\epsilon]$. More formally there exists a dual path $\hat{\pi} \in \hat{W}^{(-\epsilon)+}$ such that $\sup_{t \in [-1, -\epsilon]} |\hat{\pi}(t) - \hat{\pi}^{n_0}(t)| < \epsilon$ and thus $\sup_{t \in [-1, -\epsilon]} |\hat{\pi}(t) - \hat{\pi}^{n_0}(t)| < 4\epsilon$.

Now select any sequence $\epsilon_l \downarrow 0$ with $\epsilon_l < \epsilon$ for all $l \geq 1$. By the above procedure we can select $\hat{\pi}_l \in \hat{W}$ with $\sigma_{\hat{\pi}_l} \geq -\epsilon_l > -\epsilon$ such that $\sup_{t \in [s_l, -\epsilon_l]} |\hat{\pi}_l(t) - \pi(t)| < 4\epsilon_l$ for some $\pi \in \{\pi_1, \ldots, \pi_k\}$. Since we have only finitely many paths $\pi_1, \ldots, \pi_k$, we can choose $\pi_i$ for some $1 \leq i \leq k$ and a subsequence $\hat{\pi}_l$ so that $\sup_{t \in [s_l, -\epsilon]} |\hat{\pi}_l(t) - \pi_i(t)| < 4\epsilon_l$ for all $j \geq 1$. By the compactness of $\hat{W}$, there exists $\hat{\pi} \in \hat{W}$ with $\sigma_{\hat{\pi}} \geq -\epsilon$ such that $\hat{\pi}(t) = \pi(t)$ for $t \in [s_l, -\epsilon]$. This violates property (c) of Brownian web and its dual listed earlier. Hence $\mathbb{P}(\Omega_{n_0}^d) = 0$ for all $k \geq 0$ and this completes the proof that $\eta_n \to \eta$ almost surely as $n \to \infty$.

Now it suffices to show that the sequence $\{\eta_n : n \in \mathbb{N}\}$ is uniformly integrable. As $\eta_n \to \eta$ almost surely, using Proposition 4.1 of [13] for any $k \geq 1$ we have,

$$\lim_{n \to \infty} \mathbb{P}(\eta_n \geq k) = \mathbb{P}(\eta \geq k) \leq (\mathbb{P}(\eta \geq 1))^k < 1.$$  

We have used the fact that $\mathbb{P}(\eta \geq 1)$ is same as the probability that two independent Brownian motions starting at unit distance do not meet by time 1, which is strictly smaller than 1. This completes the proof.

**Remark 4.6.** It is important to observe that the argument for uniform integrability presented here depends only on the fact that $\eta_n \to \eta$ almost surely whereas the argument of Lemma 3.3 of [22] uses some model specific assumptions. We comment here that similar arguments show that $\eta_n(t_0, t_1) \to \eta(t_0, t_1)$ almost surely for any $t_0 < t_1$.

The following lemma proves Theorem 1.3.

**Lemma 4.7.** As $n \to \infty$, we have $E[\kappa_n(f)] \to E[\kappa(f)]$.

We first prove Theorem 1.3 assuming Lemma 4.7 holds.

**Proof of Theorem 1.3**. Recall that for $(x, t) \in \mathbb{Z}_{\text{even}}^2$ and for $\pi^{(x,t)} \in \mathcal{X}$, i.e., the path in $\mathcal{X}$ starting from $(x, t)$, the $n$-th order diffusively scaled path is denoted by $\pi^n(x, t)$. For $(x, t) \in \mathbb{Z}_{\text{even}}^2$ and for $n \geq 1$, let

$$S_n(x, t) := \{\pi^n(y, s) \in \mathcal{X}_n : (y, s) \in C(x, t)\}$$

and

$$\hat{S}_n(x, t) := \{\hat{\pi}_n : \hat{\pi}_n \text{ is the } n\text{-th order scaled version of } \hat{\pi} \in \hat{X}\}.$$  

Both the path spaces, $S_n(x, t)$ and $\hat{S}_n(x, t)$, are tree-like in the sense of Definition 2.2 and we consider the complete metric spaces $\mathcal{M}(S_n(x, t))$ and $\mathcal{M}(\hat{S}_n(x, t))$. It follows that for all $(x, t) \in \mathbb{Z}_{\text{even}}^2$ and for all $n \geq 1$, we have that both the metric spaces $T_n(x, t)$ and $\hat{T}_n(x, t)$ are naturally embedded into the metric spaces $\mathcal{M}(S_n(x, t))$ and $\mathcal{M}(\hat{S}_n(x, t))$ respectively. This embedding ensures that for all $n \geq 1$ we have

$$d_{GH}(\mathcal{M}(S_n(x, t)), T_n(x, t)) \vee d_{GH}(\mathcal{M}(\hat{S}_n(x, t)), \hat{T}_n(x, t)) \leq 1/n.$$  

22
Hence to prove Theorem 1.3 it suffices to show that

$$(\mathcal{M}(S_n(0,0)), \mathcal{M}(\bar{S}_n(0,0))) \subset \{L(0,0) \geq n\} \Rightarrow (\mathcal{T}, \bar{\mathcal{T}})$$ as $n \to \infty$.

Using translation invariance of our model, we have

$$E[\kappa_n(f)] = E\left[\sum_{k=0}^{\lfloor \sqrt{n} \rfloor} 1_{\{L(k,0) \geq n\}} f(\phi_k(\mathcal{M}(S_n(0,0))), \phi'_k(\mathcal{M}(\bar{S}_n(0,0))))\right]$$

where $\phi_k, \phi'_k$ are isometric embeddings of the respective metric spaces into $\mathbb{M}$.

$$= ([\sqrt{n}] + 1) E\left[1_{\{L(0,0) \geq n\}} f(\phi_0(\mathcal{M}(S_n(0,0))), \phi'_0(\mathcal{M}(\bar{S}_n(0,0))))\right]$$

$$= ([\sqrt{n}] + 1) P(L(0,0) \geq n) E\left[f(\phi_0(\mathcal{M}(S_n(0,0))), \phi'_0(\mathcal{M}(\bar{S}_n(0,0))))\right] \bigg| L(0,0) \geq n$$

$$= E\left[\sum_{k=0}^{\lfloor \sqrt{n} \rfloor} 1_{\{L(k,0) \geq n\}} f(\phi_0(\mathcal{M}(S_n(0,0))), \phi'_0(\mathcal{M}(\bar{S}_n(0,0))))\right] \bigg| L(0,0) \geq n$$

$$= E[\eta_n] E\left[f(\phi_0(\mathcal{M}(S_n(0,0))), \phi'_0(\mathcal{M}(\bar{S}_n(0,0))))\right] \bigg| L(0,0) \geq n$$

Using Lemma 4.5 and Lemma 4.7 we obtain

$$\lim_{n \to \infty} E\left[f(\phi_0(\mathcal{M}(S_n(0,0))), \phi'_0(\mathcal{M}(\bar{S}_n(0,0))))\right] \bigg| L(0,0) \geq n = \lim_{n \to \infty} \frac{E[\kappa_n(f)]}{E[\eta_n]}$$

$$= \frac{E[\kappa(f)]}{E[\eta]} = E(f(\phi(\mathcal{T})), \phi'(\mathcal{T})).$$

Since $f$ is chosen arbitrarily, this completes the proof. □

We now prove Lemma 4.7. Since $f$ is bounded, uniform integrability of $\{\kappa_n(f)\}$ follows from uniform integrability of $\{\eta_n : n \in \mathbb{N}\}$. Hence to prove Lemma 4.7, it suffices to show that $\kappa_n(f) \to \kappa(f)$ almost surely as $n \to \infty$. We introduce some notation now. On the event $\{\eta = 1\}$, Lemma 4.5 ensures that there exists $n_0 = n_0(\omega)$ such that $\eta_n = 1$ for all $n \geq n_0$. For $n \geq n_0$ consider $x = x(\omega), x_n = x_n(\omega) \in \mathbb{R}$ such that $\xi = \{x\}$ and $\xi_n = \{x_n\}$. For ease of notation on the event $\{\eta = 1\}$ and for $n \geq n_0$ we take

$$(S, \bar{S}) := \left(S(\pi_{r}^{(x,0)}, \tilde{\pi}_{l}^{(x,0)})(W, \tilde{W}), \bar{S}(\pi_{r}^{(x,0)}, \tilde{\pi}_{l}^{(x,0)})(W, \tilde{W})\right),$$

$$(\mathcal{S}, \mathcal{\bar{S}}) := \left(\mathcal{S}(\pi_{r}^{(x,0)}, \tilde{\pi}_{l}^{(x,0)})(W, \tilde{W}), \mathcal{\bar{S}}(\pi_{r}^{(x,0)}, \tilde{\pi}_{l}^{(x,0)})(W, \tilde{W})\right),$$

$$(\mathcal{S}_n, \mathcal{\bar{S}}_n) := \left(\mathcal{S}(\pi_{r}^{(x_n,0)}, \tilde{\pi}_{l}^{(x_n,0)})(\bar{X}_n, \tilde{X}_n), \mathcal{\bar{S}}(\pi_{r}^{(x_n,0)}, \tilde{\pi}_{l}^{(x_n,0)})(\bar{X}_n, \tilde{X}_n)\right),$$

$$(\mathcal{M}, \mathcal{\bar{M}}) := \left(M(\mathcal{S}, \mathcal{\bar{S}})(W, \tilde{W}), \mathcal{\bar{M}}(\mathcal{S}, \mathcal{\bar{S}})(W, \tilde{W})\right)$$ and

$$(\mathcal{M}_n, \mathcal{\bar{M}}_n) := \left(M(\mathcal{S}_n, \mathcal{\bar{S}}_n)(\bar{X}_n, \tilde{X}_n), \mathcal{\bar{M}}(\mathcal{S}_n, \mathcal{\bar{S}}_n)(\bar{X}_n, \tilde{X}_n)\right).$$

In what follows, we assume that we are working on a probability space such that $(\bar{X}_n, \tilde{\bar{X}}_n)$ converges to $(W, \tilde{W})$ almost surely in $(\mathcal{H} \times \mathcal{H}, d_{\mathcal{H} \times \mathcal{H}})$.

**Lemma 4.8.** On the event $\{\eta = 1\}$, we have $d_{\mathcal{H} \times \mathcal{H}}((\mathcal{S}_n, \mathcal{\bar{S}}_n), (\mathcal{S}, \mathcal{\bar{S}})) \to 0$ almost surely as $n \to \infty$.

**Proof:** We prove that $d_{\mathcal{H}}(S_n, S) \to 0$ as $n \to \infty$ almost surely. The argument for $d_{\mathcal{H}}(\bar{S}_n, \bar{S})$ is exactly the same and hence omitted. Fix $\epsilon > 0$. Choose $\delta = \delta(\omega) \in (0, \epsilon) \cap \mathbb{Q}$ such that

$$\sup\{\hat{\pi}(s_1) - \hat{\pi}(s_2) : s_1, s_2 \in [-2\delta, 0], \hat{\pi} \in \tilde{\pi}^{0+}, \hat{\pi}(0) \in [0, 1]\} < \epsilon/8.$$
Assume that \(\xi(-\delta,0) = \{x, x_1, \ldots, x_k\}\) where \(\xi = \{x\}\). We choose \(\gamma_\delta = \gamma_\delta(\omega) \in (0, \epsilon)\) such that (a) the points in \(\xi(-\delta,0)\) are at least \(2\gamma_\delta\) distance apart, i.e., \(|x - x_1| \wedge |x_i - x_j| > 2\gamma_\delta\) for all \(1 \leq i < j \leq k\) and (b) \((x - \gamma_\delta, x + \gamma_\delta)\) as well as \((x_i - \gamma_\delta, x_i + \gamma_\delta)\) are both included in the interval \((0,1)\) for all \(1 \leq i \leq k\).

We mentioned earlier that there exist dual paths \(\hat{\pi}_r^{(x,0)}\), \(\hat{\pi}_r^{(x',0)}\) in \(\hat{W}\) both starting at \((x,0)\) such that \(\hat{\pi}_r^{(x,0)}(1) > \hat{\pi}_r^{(x',0)}(1)\). Similarly for each \(x_i \in \xi(-\delta,0)\) there exist dual paths \(\hat{\pi}_r^{(x,0)}\), \(\hat{\pi}_r^{(x',0)}\) in \(\hat{W}\) both starting at \((x_i,0)\) such that \(\hat{\pi}_r^{(x,0)}(1) > \hat{\pi}_r^{(x',0)}(1)\) for all \(1 \leq i \leq k\). Because of continuity, there exists \(\nu_0 = \nu_0(\omega) > 0\) such that \(\pi_r^{(x,0)}(1-\epsilon) > \hat{\pi}_r^{(x,0)}(1-\nu_0)\) and \(\pi_r^{(x,0)}(1-\epsilon) - \pi_r^{(x,0)}(1-\nu_0) > \hat{\pi}_r^{(x,0)}(1-\nu_0)\) for \(1 \leq i \leq k\). From properties (f) and (g) of \((W_r, \hat{W})\) (see Subsection \[\text{B.1}\]) it follows that there exist \(\hat{\pi}_r, \hat{\pi}_l, \hat{\pi}_r^{(x,0)}\) in \(\hat{W}\) such that all these dual paths have starting times strictly larger than 0 and

\[
\begin{align*}
\text{(i)} & \quad \max\{\{\hat{\pi}_r - \hat{\pi}_r^{(x,0)}(0)\}, \{|\hat{\pi}_r^{(x,0)}(0)|, |\hat{\pi}_r^{(x,0)}(0)|, |\hat{\pi}_r^{(x,0)}(0)|, |\hat{\pi}_r^{(x,0)}(0)| : 1 \leq i \leq k\} < \gamma_\delta/4, \\
\text{(ii)} & \quad \max\{\{\hat{\pi}_r - \hat{\pi}_r^{(x,0)}(\delta)\}, |\hat{\pi}_r^{(x,0)}(\delta)|, |\hat{\pi}_r^{(x,0)}(\delta)|, |\hat{\pi}_r^{(x,0)}(\delta)| : 1 \leq i \leq k\} = 0. 
\end{align*}
\]  

(32)

This gives that

\[
\max\{\{\hat{\pi}_r - \hat{\pi}_r^{(x,0)}(s)\} : s \in [-2\delta,0]\} \text{ and } \max\{\{\hat{\pi}_r - \hat{\pi}_r^{(x,0)}(s)\} : s \in [-2\delta,0]\} < \epsilon/2. 
\]  

(33)

From Remark \[\text{4.6}\] we have \(\mathbb{P}\{\bigcap_{t \in (-1,0)} \eta_n(t,0) = \eta(t,0)\}\} = 1. \) Set \(\zeta_0 > 0\) such that

\[
\zeta_0 < \min\{\{\hat{\pi}_r^{(x,0)}(1) - \nu_0, \sigma_{\hat{\pi}_l}, \sigma_{\hat{\pi}_r}\} \wedge \min\{\{\hat{\pi}_r^{(x,0)}(1) - \nu_0, \sigma_{\hat{\pi}_l}, \sigma_{\hat{\pi}_r}: 1 \leq i \leq k\}.
\]

Let \(n_1 = n_1(\omega) \geq \max\{4/\nu_0, 1/\delta, n_0\}\) be such that, for all \(n \geq n_1\)

\[
\begin{align*}
\text{(i)} & \quad \eta_n(-\delta,0) = \eta(-\delta,0)\text{ and } \eta_n = \eta, \\
\text{(ii)} & \quad d_{\hat{\pi}_r}(\hat{W}_r, \hat{X}_n, \hat{X}_n) < g(\min\{\gamma_\delta, \zeta_0, \epsilon\})/m \text{ where } m \text{ is defined as in } \text{[31]} \text{ and the value } \text{g} = \text{g}(\min\{\gamma_\delta, \zeta_0, \epsilon\})/m \text{ is taken as in Remark } \text{3.1}.
\end{align*}
\]

The choice of \(n_1\) and \(\zeta_0\) ensure that there exist \(\hat{\pi}_r^n, \hat{\pi}_l^n \in \hat{X}_n^{0+}\) and for all \(1 \leq i \leq k\) there exist \(\hat{\pi}_r^n, \hat{\pi}_l^n \in \hat{X}_n^{0+}\) such that

\[
\sup\{\{\hat{\pi}_r^n - \hat{\pi}_r^n(1)\} : s \in [-1,0]\} < \min\{\gamma_\delta, \zeta_0, \epsilon\}/4, 
\]

(34)

as well as \(\hat{\pi}_r^n(-1) > \hat{\pi}_r^n(-1)\) and \(\hat{\pi}_l^n(-\delta - \nu_0) > \hat{\pi}_l^n(-\delta - \nu_0)\) for all \(1 \leq i \leq k\). Therefore, there exists a forward path \(\hat{\phi}^n \in \hat{X}_n^{(1-\delta)}\) such that \(\hat{\pi}_r^n(-1) < \hat{\phi}^n(-1) < \hat{\phi}^n(1)\). Using \(\text{[32]}\) we further have that

\[
\begin{align*}
\text{as well as } \hat{\pi}_r^n(-1) > \hat{\pi}_r^n(-1) \text{ and } \hat{\pi}_r^n(-\delta - \nu_0) > \hat{\pi}_l^n(-\delta - \nu_0)\text{ for all } 1 \leq i \leq k. \text{ Therefore, there exists a forward path } \hat{\phi}^n \in \hat{X}_n^{(1-\delta)}\text{ such that } (a) \hat{\pi}_r^n(-\delta) < \hat{\phi}^n(-\delta) < \hat{\phi}^n(-\delta) \text{ and } (b) x_i - \gamma_\delta < \hat{\phi}^n(0) < \hat{\phi}^n(0) \text{ and } \hat{\phi}^n(0) < x_i + \gamma_\delta. \text{ Since for all } n \geq n_1 \text{ we have } \#\xi_n(-\delta,0) = k + 1, \text{ hence } \xi_n(-\delta,0) = \{\hat{\phi}^n(0) : 1 \leq i \leq k\} \cup \{\hat{\phi}^n(0)\}. \text{ Thus, there exist a unique } x^0 = \hat{\phi}^n(0) \in \xi_n(-\delta,0) \text{ with } |x^0 - x| < \gamma_\delta \text{ and unique } x^0 = \hat{\phi}^n(0) \in \xi_n(-\delta,0) \text{ with } |x^0 - x| < \gamma_\delta \text{ for each } i.
\end{align*}
\]

(35)

We observe that \((y^n,0) = (\sqrt{n}\hat{\phi}^n(0),0)\) is in \(Z^{2,\text{even}}_{y^n}\) and the vertices \((y^n + 1,0), (y^n - 1,0)\) are the right and left dual neighbours of \((y^n,0)\) respectively. Similarly \((y^n,0) = (\sqrt{n}\hat{\phi}^n(0),0)\) is in \(Z^{2,\text{even}}_{y^n}\) for \(1 \leq i \leq k\).

For any \(\pi_n \in S_n\) we have \(\pi_n = \pi_n(0) = x^n\) and hence for all \(n \geq n_1,\) using \(\text{[33]}\) and \(\text{[34]}\) obtain

\[
\begin{align*}
\sup_{\|s_1, s_2\| \leq \epsilon} \{|\pi_n(s_1) - \pi_n(s_2)|\} & \leq \sup_{\|s_1, s_2\| \leq \epsilon} \{|\pi_n(\gamma^n + 1,0) - \pi_n(\gamma^n - 1,0)|\} \\
& \leq \sup_{\|s_1, s_2\| \leq \epsilon} \{|\pi_n(s_1) - \pi_n(s_2)|\} \\
& \leq \sup_{\|s_1, s_2\| \leq \epsilon} \{|\pi_n(s_1) - \pi_n(s_2)|\} + \epsilon/2 \leq \epsilon.
\end{align*}
\]

24
For any \( \pi_1, \pi_2 \in \mathbb{S} \) and \( \pi_1^1, \pi_2^2 \in S_n \) we have \( \pi_1(s) = \pi_2(s) \) and \( \pi_1^1(s) = \pi_2^2(s) \) for all \( s \geq 0 \) with \( |\pi_1^1(0) - \pi_1(0)| < \epsilon \). This observation together with the fact that \( \sup_{s_1, s_2 \in [-2\delta, 0]} |\pi_1(s_1) - \pi_1(s_2)| \leq \sup_{s_1, s_2 \in [-2\delta, 0]} |\pi_1^1(s_1) - \pi_1^1(s_2)| < \epsilon \) ensures that to prove Lemma 4.8 it suffices to show that

\[
d_H(S_n^{(-2\delta)}, \mathbb{S}^{(-2\delta)}) \to 0 \quad \text{as} \quad n \to \infty.
\]

Fix \( \pi \in \mathbb{S} \) with \( \sigma_\pi \leq -2\delta \). From the choice of \( n_1 \) it follows that for all \( n \geq n_1 \) there exists an approximating path \( \pi^n \in X_n^{(-\delta)} \) with \( |\pi^n(0) - \pi(0)| < \gamma_\delta \). As \( \sigma_{\pi^n} \leq -\delta \), we have \( \pi^n(0) \in \xi(\delta, 0) \). Since \( \pi \in \mathbb{S} \), it follows that \( \pi(0) = x \). Hence from the uniqueness as mentioned in (43), it follows that for all \( n \geq n_1 \) the approximating path \( \pi^n \) must have \( \pi^n(0) = x^n \) implying that \( \pi^n \in S_n \). Similar argument shows that for all \( n \geq n_1 \) and for \( \pi^n \in S_n \) with \( \sigma_{\pi^n} \leq -\delta \), there exists approximating path in \( \mathbb{S} \). This completes the proof.

Before stating the next lemma we give an alternate definition of \( d_{GH} \) which will be used in our proof. For two compact metric spaces \((X_1, d_1)\) and \((X_2, d_2)\), a correspondence between \(X_1\) and \(X_2\) is a subset \(\mathcal{R}\) of \(X_1 \times X_2\) such that for every \(x_1 \in X_1\) there exists at least one \(x_2 \in X_2\) such that \((x_1, x_2) \in \mathcal{R}\) and conversely for every \(y_2 \in X_2\) there exists at least one \(y_1 \in X_1\) such that \((y_1, y_2) \in \mathcal{R}\). The distortion of a correspondence \(\mathcal{R}\) is defined by

\[
dis(\mathcal{R}) := \sup \{|d_1(x_1, y_1) - d_2(x_2, y_2)| : (x_1, x_2), (y_1, y_2) \in \mathcal{R}\}.
\]

Then, if \(X_1\) and \(X_2\) are two real trees, we have

\[
d_{GH}(X_1, X_2) := 1/2 \inf_{\mathcal{R} \in \mathcal{C}(X_1, X_2)} \text{dis}(\mathcal{R})
\]

where \(\mathcal{C}(X_1, X_2)\) denotes the set of all correspondences between \(X_1\) and \(X_2\). Using Lemma 4.8 we prove the following lemma which is the main lemma of this section.

**Lemma 4.9.** On the event \( \{\gamma = 1\} \) we have \( d_{GH}(\phi_n(M_n), \phi(M)) \rightarrow d_{GH}(\phi'_n(\widetilde{M}_n), \phi'(\widetilde{M})) \rightarrow 0 \) almost surely as \( n \to \infty \), where \( \phi_n, \phi, \phi'_n, \phi' \) denote the corresponding embeddings of the respective metric spaces into \( \mathbb{M} \).

**Proof:** Before proving this lemma we first sketch the key idea. The metric \( d_{\mathcal{H} \times \mathcal{H}} \) deals with the path space topology, convergence in this metric does not ensure that the coalescing times also converge (even if they are finite). Using joint convergence of the collection of forward paths together with the collection of dual paths, we show that the coalescing times of the scaled paths converge jointly to the coalescing time of the limiting Brownian paths and this gives convergence with respect to the ancestor metric. We first show that \( d_{GH}(\phi_n(M_n), \phi(M)) \to 0 \) as \( n \to \infty \) almost surely and the argument for \( d_{GH}(\phi'_n(\widetilde{M}_n), \phi'(\widetilde{M})) \to 0 \) as \( n \to \infty \) is exactly the same.

Using Definition (36), it suffices to show that there exists a correspondence \( \mathcal{R}_n \) between \( M_n \) and \( M \) such that

\[
\lim_{n \to \infty} \text{dis}(\mathcal{R}_n) = 0.
\]

Choose any sequence \( \{\epsilon_k : k \in \mathbb{N}\} \) such that \( \epsilon_k \downarrow 0 \) as \( k \to \infty \). For \( k > 1 \), let \( n_k = n_k(\omega) \in \mathbb{N} \) be such that \( d_{\mathcal{H}}(S_{n_k}, \mathbb{S}) < g(\epsilon_k/8, m) \) almost surely where \( m \) is defined as in (51) and \( g(\epsilon, m) \) as in Remark 3.4. We define correspondences \( \mathcal{R}_n \) for large \( n \) only. For \( n_k \leq n < n_{k+1} \), we define our correspondence \( \mathcal{R}_n \) between \( M_n \) and \( M \) as

\[
\mathcal{R}_n := \left( \bigcup_{(y, s) \in M} \{(y, s), (y_n, s_n) : (y_n, s_n) \in M_n \text{ such that } d_H(\pi'(y', s'), \pi'(y_n, s_n)) < g(\epsilon_k/8, m) \} \right) \cup
\]

\[
\left( \bigcup_{(y_n, s_n) \in M_n} \{(y_n, s_n), (y, s) : (y, s) \in M \text{ with } \pi(y, s) \in S \text{ such that } d_H(\pi(y'', s''), \pi(y_n, s_n)) < g(\epsilon_k/8, m) \} \right).
\]

(38)

25
For all $n_0 \leq n < n_{k+1}$, (38) gives a valid correspondence.

Aiming for a contradiction suppose (37) does not hold. As $d_H(\pi_1, \pi_2) < \epsilon_k/8$, it implies that $|\sigma_{\pi_1} - \sigma_{\pi_2}| < \epsilon_k/8$, if (37) does not hold then there exists $\beta = \beta(\omega) > 0$ such that

$$\limsup_{k \to \infty} \sup_{\gamma(1,2) \in [\gamma(1,2)_d \cap [28], \gamma(1,2)_m]} \left|\gamma(1,2)_k - \gamma(1,2)_m\right| : \pi_1^k, \pi_2^k \in S \text{ and } \pi_1^n, \pi_2^n \in S_n \text{ for some } n_k \leq n < n_{k+1} \text{ and }$$

$$d_H(\pi_1, \pi_1^n) \lor d_H(\pi_2, \pi_2^n) < \epsilon_k/8, \text{ for all } k \leq k_0,$$

where for ease of notation we take $\gamma(1,2)_k = \gamma(\pi_1^k, \pi_2^k)$ and $\gamma(1,2)_m = \gamma(\pi_1^n, \pi_2^n)$. Choose $k_0$ such that $\epsilon_k \lor (1/n_k) < \beta/8$ for all $k \geq k_0$. From (39), it follows that there are infinitely many $k \geq k_0$ such that there exist $\pi_1^k, \pi_2^k \in S$ and $\pi_1^n, \pi_2^n \in S_n$ for some $n_k \leq n < n_{k+1}$ with $d_H(\pi_1^k, \pi_1^n) \lor d_H(\pi_2^k, \pi_2^n) < \epsilon_k/8, m)$ and $|\gamma(1,2)_k - \gamma(1,2)_m| > \beta/2$. We consider the two cases:

(a) $\gamma(1,2)_n > \gamma(1,2)_k + \beta$ and

(b) $\gamma(1,2)_n \leq \gamma(1,2)_k - \beta$.

For (a) from the choice of $k_0$, it follows that for infinitely many $k \geq k_0$ and for some $n_k \leq n < n_{k+1}$ we have $\sigma_{\pi_1^k} \lor \sigma_{\pi_2^k} \leq \gamma(1,2)_k + \beta/4$ with $\gamma(1,2)_k + \beta) \not= \gamma(\pi_1^k, \pi_2^k)$ and $\beta/4$. From the construction of the dual graph it follows that for infinitely many $k \geq k_0$ and for some $n_k \leq n < n_{k+1}$, there exists $\pi_1^k \in \hat{S}_n$ with $\sigma_{\pi_1^k} \geq \gamma(1,2)_k + 3/4$ and for $s \in [\gamma(1,2)_k, \gamma(1,2)_k + 3/4]$ we have $\pi_1^k(s) \in [\pi_1^k(s) \lor \sigma_{\pi_1^k}(s)]$. Hence we have $|\pi_1^k(s) - \pi_1^n(s)| < \epsilon_k/2$ for all $s \in [\gamma(1,2)_k, \gamma(1,2)_k + 3/4)$. Since $d_H(\hat{S}_n, \hat{S}_n, \hat{S}^k) < \epsilon_k/8$ for all $k \leq n_k$, for infinitely many $k \geq k_0$ there exists $\pi_1^k \in \hat{S}$ with $\sigma_{\pi_1^k} \geq \gamma(1,2)_k + \beta/4$ and $d_H(\pi_1^k, \pi_1^n) < \epsilon_k/8$. This gives us that $|\pi_1^k(s) - \pi_1^n(s)| < \epsilon_k$ for all $s \in [\gamma(1,2)_k, \gamma(1,2)_k + \beta/4]$. For (b) we need a slightly different argument. From the properties of $(W, \hat{W})$, discussed earlier, it follows that there exists $\pi_{\hat{k}} \in \hat{S}$ starting from the point $(\pi_1^k(\gamma(1,2)_k), \gamma(1,2)_k)$ and lying between the forward paths $\pi_1^k$ and $\pi_2^k$ in the time interval $[\sigma_{\pi_1^k} \lor \sigma_{\pi_2^k}, \gamma(1,2)_k]$. For $s \in [\gamma(1,2)_k, \gamma(1,2)_k - \beta/4, \gamma(1,2)_k]$ we have $\pi_1^k(s) = \pi_2^k(s)$. Since $d_H(\pi_1^k, \pi_1^n) < \epsilon_k/8, m)$ for $i = 1, 2$, it follows that $0 < |\pi_i^k(s) - \pi_i^k(s)| < \epsilon_k/8$ for all $s \in [\gamma(1,2)_k, \gamma(1,2)_k - \beta/4, \gamma(1,2)_k]$. Since the dual path $\pi_{\hat{k}}^k$ starting from the point $(\pi_1^k(\gamma(1,2)_k), \gamma(1,2)_k) \hat{k}$ lies between $\pi_1^k$ and $\pi_2^k$, we have $|\pi_{\hat{k}}^k(s) - \pi_{\hat{k}}^k(s)| < \epsilon_k/4$ for all $s \in [\gamma(1,2)_k - \beta/4, \gamma(1,2)_k]$. Since $\epsilon_k \downarrow 0$, because of compactness of $(W, \hat{W})$, this contradicts the fact that in the double Brownian web almost surely no two forward path and dual path spend positive Lebesgue measure time together. This completes the proof.

Finally we prove Lemma 4.7 to complete the proof of Theorem 1.3

**Proof of Lemma 4.7**: Since $f$ is bounded, the uniform integrability of the sequence $\{\kappa_n(f); n \in N\}$ follows from that of the sequence $\{\eta_n : n \in N\}$. Since $\eta_n \to \eta$ almost surely, $\kappa_n(f) \to \kappa(f)$ holds trivially on the event $\{\eta = 0\}$. On the event $\{\eta = 1\}$, using continuity of $f$ this follows from Lemma 4.9. For general $k \geq 1$ on the event $\{\eta = k\}$, the proof is similar.

5 Concluding remark

To the best of our knowledge both these continuum random trees, $T$ and $\hat{T}$, have not been studied in the literature so far. On the other hand, the system of coalescing Brownian motions starting from every space-time points on $\mathbb{R}^2$ has been extensively studied but with respect to a different topology (see [13, 23] for details). It should be mentioned that coalescing Brownian motions starting from all the points in $\mathbb{R}$ at a given time has been studied as a genealogical tree in [15] with respect to a different topology where the spatial locations of the points are also taken care of.

Further understanding of these continuum random trees $T$ and $\hat{T}$ might be useful and we present some questions towards that direction. In the above construction we have used the fact that the collection of forward paths $S(B^+, B^-)$ in the region $\Delta(B^+, B^-)$ is almost surely uniquely determined by the collection of backward or dual paths $\hat{S}(B^+, B^-)$. It should be possible to construct $T$ directly as completion of the metric space obtained from the system of coalescing forward Brownian paths starting from all the points.
of $\Delta(\tilde{B^+}, \tilde{B^-}) \cap \mathbb{Q}^2$, which follow Skorohod reflection at the boundary of $\Delta(\tilde{B^+}, \tilde{B^-})$. In order to do that construction, one has to show that starting from finitely many space-time points such a collection satisfies Kolmogorov’s consistency conditions, as it does not directly follow from [31].

From the construction of $(T, \tilde{T})$, it is reasonable to expect that $\tilde{T}$ almost surely determines $T$ and vice-versa. One possible way to prove this is to show that the contour function of $T$ is determined by $\tilde{T}$ and vice-versa and in that case $\tilde{T}$ can be regarded as the dual tree of $T$. It is natural to ask how are the contour functions of these two continuum random trees distributed? Presently we do not have any understanding about these processes. Finally in the context of drainage network scaling relations, it might be of interest to see whether these two continuum trees obey Horton’s law or not.

Acknowledgements: The author thanks Siva Athreya, Manjunath Krishnapur, Rahul Roy, Rongfeng Sun and Sreekar Vadlamani for discussion and comments. Part of the work was done when the author was a visiting scientist at Indian Statistical Institute, Bangalore center.

References
[1] Aldous, D. (1991) The continuum random tree I. Ann. Probab. 19 1–28.
[2] Aldous, D. (1993) The continuum random tree III. Ann. Probab. 21 248–289.
[3] Aldous, D. (1997) Brownian excursions, critical random graphs and the multiplicative coalescent. Ann. Probab. 25 812–854.
[4] Aldous, D. and Pittel, B. (2000) On a random graph with immigrating vertices: emergence of the giant component. Random Structures Algorithms 17 79–102.
[5] Arratia, R. (1979). Coalescing Brownian motions on the line. Ph.D. dissertation, Univ. Wisconsin, Madison.
[6] Berry, L. A., Broutin, N. and Goldschmidt, C. (2012) The continuum limit of critical random graphs. Probab. Theory Related Fields 152 367–406.
[7] Berry, L. A., Broutin, N., Goldschmidt, C. and Miermont, G. (2015) The scaling limit of the minimum spanning tree of a complete graph. arXiv:1301.1664
[8] Bolthausen, E. (1976) On a functional central limit theorem for random walks conditioned to stay positive. Ann. Probab. 4 480–485.
[9] Bhamidi, S., van der Hofstad, R. and Sen, S. The multiplicative coalescent, inhomogeneous continuum random trees, and new universality classes for critical random graphs. To appear in Probab. Theory Relat. Fields.
[10] Coletti, C. F., Dias, E. S. and Fontes, L. R. G. (2009) Scaling limit for a drainage network model. J. Appl. Probab., 46 1184–1197.
[11] Ferrari, P. A., Fontes, L. R. G. and Wu, X. Y. (2005) Two-dimensional Poisson Trees converge to the Brownian web. Ann. Inst. H. Poincaré Probab. Statist. 41 851–858.
[12] Fontes, L. R. G., Isopi, M., Newman, C. M. and Ravishankar, K. (2003) The Brownian web: characterization and convergence. arXiv:0304119.
[13] Fontes, L. R. G., Isopi, M., Newman, C. M. and Ravishankar, K. (2004) The Brownian web: characterization and convergence. Ann. Probab. 32 2857–2883.
[14] Goldschmidt, C. (2005) Critical random hypergraphs: the emergence of a giant set of identifiable vertices. Ann. Probab. 33 1573–1600.
[15] Greven, A., Sun, R. and Winter, A. (2016) Continuum space limit of the genealogies of interacting Fleming-Viot processes on Z. *Electron. J. Probab.*, 21 1–64.

[16] Gromov, M. (1999) *Metric Structures for Riemannian and Non-Riemannian Spaces*. Birkhauser.

[17] Horton, R. T. (1945) Erosional development of streams and their drainage basins: hydro-physical approach to quantitative morphology. *Geol. Soc. Am. Bull.* 56 275–370.

[18] Imrich, W. (1977) On metric properties of tree-like spaces. 129–156.

[19] Kirchner, J., W. (1993) Statistical inevitability of Horton’s laws and the apparent randomness of stream channel networks. *J. Geol.* 21(7) 591–594.

[20] Le Gall, J. F. (2005) Random trees and applications. *Probab. Surv.* 2 245–311 (electronic), ISSN 1549-5787.

[21] Duquesne1, T. and Le Gall, J. F. (2002) Random Trees, Levy Processes and Spatial Branching Processes. *Astérisque* 281.

[22] MarcKert, J. F. and Mokkadem, A. (2003) The depth first processes of Galton-Watson trees converge to the same Brownian excursion. *Ann. Probab.* 31 1655–1678.

[23] Nachmias, A. and Peres, Y. (2010) Critical percolation on random regular graphs. *Random Structures Algorithms* 36 111–148.

[24] Rodriguez-Iturbe, I. and Rinaldo, A. (1997) *Fractal river basins: chance and self-organization*. Cambridge Univ. Press, New York.

[25] Roy, R., Saha, K. and Sarkar, A. (2016) Random directed forest and the Brownian web. *Ann. Inst. H. Poincaré Probab. Statist.* 52 1106–1143.

[26] Roy, R., Saha, K. and Sarkar, A. (2016) Hack’s law in a drainage network model: a Brownian web approach. *Ann. Appl. Probab.* 26 1807–1836.

[27] Scheidegger, A. E. (1967) A stochastic model for drainage pattern into an intramontane trench. *Bull. Ass. Sci. Hydrol.* 12 15–20.

[28] Schertzer, E., Sun, R. and Swart, J. M. (2017) The Brownian web, the Brownian net and their universality. *Advances in Disordered Systems, Random Processes and Some Applications*, 270–368, Cambridge University Press.

[29] Shreve, R. L. (1966) Statistical law of stream numbers. *J. Geol.* 74 17–37.

[30] Shreve, R. L. (1967) Infinite topologically random channel networks. *J. Geol.* 75 178–186.

[31] SoucaliuC, F., Toth, B. and Werner, W. (2000) Reflection and coalescence between one-dimensional Brownian paths. *Ann. Inst. H. Poincaré Probab. Statist.* 36 509-536.

[32] Sun, R. and Swart, J. M. (2008) The Brownian net. *Ann. Probab.* 36 1153–1208.

[33] Tarboton, D. G. (1996) Fractal river networks, Horton’s laws and Tokunaga cyclicity. *Journal of hydrology* 187 105–117.