

# Universality of the Brownian net

Rongfeng Sun <sup>\*</sup>      Jan M. Swart <sup>†</sup>      Jinjiong Yu <sup>‡§</sup>

January 18, 2024

## Abstract

The Brownian web is a collection of one-dimensional coalescing Brownian motions starting from every point in space and time, while the Brownian net is an extension that also allows branching. We show here that the Brownian net is the universal scaling limit of one-dimensional branching-coalescing random walks with weak binary branching and arbitrary increment distributions that have finite  $(3 + \varepsilon)$ -th moment. This gives the first example in the domain of attraction of the Brownian net where paths can cross without coalescing.

*MSC 2010.* Primary: 82C22, Secondary: 82C41, 60K35.

*Keywords.* Branching-coalescing random walks, Brownian web, Brownian net, invariance principle.

## Contents

<b>1</b>	<b>Introduction and main result</b>	<b>2</b>
1.1	Model and Main Result . . . . .	3
1.2	Proof strategy . . . . .	5
1.3	Outline . . . . .	7
<b>2</b>	<b>Background</b>	<b>7</b>
2.1	Space of compact sets of paths . . . . .	7
2.2	The Brownian web . . . . .	8
2.3	Sticky Brownian webs . . . . .	9
2.4	The Brownian net . . . . .	11
<b>3</b>	<b>The lower bound</b>	<b>13</b>
3.1	Convergence of discrete sticky webs . . . . .	13
3.2	Proof of the lower bound . . . . .	18
<b>4</b>	<b>Auxiliary Bernoulli net</b>	<b>19</b>
4.1	Bernoulli net and coupling with discrete net . . . . .	19
4.2	Branching-coalescing point sets and their duals . . . . .	20
4.3	Product invariant law and the backbone . . . . .	21
4.4	Density estimates . . . . .	23

---

<sup>\*</sup>Department of Mathematics, National University of Singapore, 10 Lower Kent Ridge Road, 119076 Singapore. Email: matsr@nus.edu.sg

<sup>†</sup>The Czech Academy of Sciences, Institute of Information Theory and Automation, Pod vodárenskou věží 4, 18208 Praha 8, Czech Republic. Email: swart@utia.cas.cz

<sup>‡</sup>KLATASDS-MOE, School of Statistics, East China Normal University, 3663 North Zhongshan Road, Shanghai 200062, China. Email: jjyu@sfs.ecnu.edu.cn

<sup>§</sup>NYU-ECNU Institute of Mathematical Sciences at NYU Shanghai, Shanghai 200062, China.

<b>5</b>	<b>The upper bound</b>	<b>26</b>
5.1	An upper bound criterion . . . . .	26
5.2	Convergence of a stationary set of paths . . . . .	27
5.3	Convergence of finite dimensional distributions . . . . .	28
5.4	Proof of the upper bound . . . . .	29
<b>6</b>	<b>Tightness</b>	<b>30</b>
6.1	Tightness criterion . . . . .	31
6.2	Relevant branching points . . . . .	31
6.3	Probability of having at least 1 relevant branching point . . . . .	34
6.4	Probability of having more than $K$ relevant branching points . . . . .	37
6.5	A multiscale argument for tightness . . . . .	40
<b>A</b>	<b>Proof of some facts for the Brownian web and net</b>	<b>48</b>

## 1 Introduction and main result

The Brownian web  $\mathcal{W}$  is a collection of coalescing Brownian motions on  $\mathbb{R}$ , starting from every point in the space-time plane  $\mathbb{R}^2$ . It originated in Arratia’s Ph.D. thesis [A79], where he constructed coalescing Brownian motions starting from every point in  $\mathbb{R}$  at time 0. In a subsequent unpublished manuscript [A81], his attempt to construct coalescing Brownian motions starting from every point in the space-time plane  $\mathbb{R}^2$  was incomplete due to the presence of special points in  $\mathbb{R}^2$  where multiple Brownian paths can emanate. The construction was completed by Tóth and Werner in [TW98], where a variant of Arratia’s coalescing Brownian motions played a crucial role in the study of the so-called true self-repelling motion. They established detailed properties for Arratia’s coalescing Brownian motions, including a complete classification of its special points. Subsequently, Fontes, Isopi, Newman and Ravishankar [FINR02, FINR04] came across the same object in the study of aging in the zero-temperature dynamics of Ising and Potts models on  $\mathbb{Z}$ . They introduced a topology for Arratia’s coalescing Brownian motions so that it becomes a random variable taking values in a Polish space, and they named it the *Brownian web*  $\mathcal{W}$ . They also provided a characterisation for the Brownian web, which they used to prove the weak convergence of systems of coalescing simple random walks to the Brownian web.

In [SS08], Sun and Swart made an important generalisation of the Brownian web  $\mathcal{W}$  by allowing paths to branch, which they call the *Brownian net*  $\mathcal{N}$ . To counter the instantaneous coalescence between paths, the branching rate is effectively “infinite”. Independently, a different construction of the Brownian net was given by Newman, Ravishankar, and Schertzer in [NRS10].

Since their inception, the Brownian web and the Brownian net have found connections to many other topics of research, including population genetic models [AS11, GSW16, EFS17], drainage network type models [CDF09, CV14, RSS16, PPR22, SVZ22, RSS23, VZ23], Poisson webs [FFW05, FVV15], the directed spanning forest [CSST21], oriented percolation [SS11, BGS19, SS19], planar aggregation models [NT12, NT15], random walks in i.i.d. space-time random environments [SSS14], interface growth models [Y16, CH23, VV23, DDP23], traffic models [A17], etc. We refer to the survey article [SSS17] for a more comprehensive list of earlier results.

The Brownian web  $\mathcal{W}$  and the Brownian net  $\mathcal{N}$  are expected to be the universal scaling limits of one dimensional coalescing systems, and branching-coalescing systems with weak branching, respectively. The universality of the Brownian web  $\mathcal{W}$  has been studied extensively. Models which have been shown to converge to the Brownian web fall under two classes: either paths must coalesce whenever they attempt to cross each other, called the *non-crossing case*; or paths can cross each other without coalescing, called the *crossing case*. For the non-crossing case, convergence criteria were formulated in [FINR04] and has been verified for many models,

including coalescing simple random walks [FINR04], extremal paths in supercritical oriented percolation [SS11, SS19], paths in drainage network models [CDF09, RSS16], Poisson webs [FFW05, FVV15], the directed spanning forest [CSST21], etc. The crossing case is much more challenging due to the lack of ordering among paths. The first example in the crossing case that was shown to converge to the Brownian web is the collection of coalescing random walks on  $\mathbb{Z}$  with general increment distributions, provided the increments have finite  $(3 + \eta)$ -th moment for some  $\eta > 0$  [NRS05, BMSV06]. In [NRS05], the authors formulated a set of convergence criteria tailored for the crossing case, which has also been verified for some drainage network models in [CV14, VZ23].

The universality of the Brownian net  $\mathcal{N}$  has been much less studied. For branching-coalescing simple symmetric random walks on  $\mathbb{Z}$  with weak branching, i.e., the branching probability is proportional to the space-time diffusive scaling parameter  $\varepsilon$ , convergence to the Brownian net  $\mathcal{N}$  was established in [SS08], which relies heavily on the *non-crossing* property that paths must meet before crossing. In particular, the set of leftmost (resp. rightmost) paths starting from each point in the space-time lattice forms a non-crossing system of coalescing random walk paths. The only other model in the non-crossing case that has been shown to converge to the Brownian net is the set of genealogical paths of a spatial  $\Lambda$ -Fleming-Viot process considered in [EFS17]. There has also been partial progress for a drainage network model with weak branching [SVZ22], which also belongs to the non-crossing case.

For the crossing case, the analogue of the convergence result in [NRS05, BMSV06] for the Brownian web  $\mathcal{W}$  is long believed to hold also for the Brownian net  $\mathcal{N}$ . The fact that such a result is only proved now is due to fundamental difficulties that arise from branching together with the lack of ordering among paths. The main contribution of this paper is to overcome these difficulties and provide the first example in the crossing case that is shown to converge to the Brownian net  $\mathcal{N}$ . More precisely, we consider the system of *branching-coalescing random walks* on  $\mathbb{Z}$  with weak binary branching and increments that have finite  $(3 + \eta)$ -th moment for some  $\eta > 0$ . The ideas we develop here should also be useful for proving convergence to the Brownian net for other models in the crossing case.

## 1.1 Model and Main Result

We now introduce the collection of branching-coalescing random walk paths on  $\mathbb{Z}$ , which are discrete analogues of the Brownian net, and hence called *discrete nets*.

Let  $a(\cdot)$  be a probability kernel on  $\mathbb{Z}$ , which specifies the increment distribution for a random walk on  $\mathbb{Z}$ . We assume that

$$\begin{aligned}
& \text{(i)} \quad a(\cdot) \text{ defines an irreducible and aperiodic random walk on } \mathbb{Z}, \\
& \text{(ii)} \quad \sum_{x \in \mathbb{Z}} xa(x) = 0, \\
& \text{(iii)} \quad \sum_{x \in \mathbb{Z}} |x|^{3+\eta} a(x) < \infty \text{ for some } \eta > 0.
\end{aligned} \tag{1.1}$$

We denote  $\sigma^2 := \sum_{x \in \mathbb{Z}} x^2 a(x)$  throughout the rest of the paper.

Before introducing branching, we first give a graphical construction of coalescing random walks starting from every point of the space-time lattice  $\mathbb{Z}^2 = \{(x, t) : x, t \in \mathbb{Z}\}$ , which we call a *discrete web*. Let  $\omega = (\omega(x, t))_{(x, t) \in \mathbb{Z}^2}$  be i.i.d. random variables with common distribution  $a(\cdot)$ . From each  $(x, t) \in \mathbb{Z}^2$ , we then draw an arrow from  $(x, t)$  to  $(x + \omega(x, t), t + 1)$ . Note that for each  $(x, t) \in \mathbb{Z}^2$ , there is a unique path  $\pi$  starting from  $(x, t)$  by following the arrows, and two paths coalesce when they meet at the space-time lattice point. We will denote by  $\sigma_\pi := t$  the starting time of the path  $\pi$ , and identify  $\pi : [\sigma_\pi, \infty) \rightarrow \mathbb{R}$  with a continuous function by linearly interpolating between consecutive integer times. Given  $\omega$ , let  $\pi_{(x, t)}$  denote the random

walk path starting from  $(x, t) \in \mathbb{Z}^2$ . We then call the random set of paths

$$W := \{\pi_{(x,t)} : (x, t) \in \mathbb{Z}^2\} \quad (1.2)$$

a *discrete web with jump distribution*  $a(\cdot)$ .

To allow for binary branching, we need to allow the possibility of two arrows coming out of each  $(x, t) \in \mathbb{Z}^2$ . Let  $(\omega^1(x, t), \omega^2(x, t))_{(x,t) \in \mathbb{Z}^2}$  be an i.i.d. family of  $\mathbb{Z} \times \mathbb{Z}$ -valued random variables with common law  $a^{(2)}(\cdot, \cdot)$ , to be specified in (1.3). Note that each of the two families  $\omega^1 = (\omega^1(x, t))_{(x,t) \in \mathbb{Z}^2}$  and  $\omega^2 = (\omega^2(x, t))_{(x,t) \in \mathbb{Z}^2}$  induces a set of arrows and a corresponding discrete web, which we denote by  $W^1$  and  $W^2$  respectively. The paths between  $W^1$  and  $W^2$  undergo sticky interaction, where stickiness comes from  $(x, t)$  with  $\omega^1(x, t) = \omega^2(x, t)$ . If we consider all paths that can follow both types of arrows, induced by either  $\omega^1$  or  $\omega^2$ , then paths still coalesce when they meet at the same space-time in  $\mathbb{Z}^2$ , but a path can also branch into two at  $(x, t)$  with  $\omega^1(x, t) \neq \omega^2(x, t)$ . We call the set of all paths obtained by following arrows induced by either  $\omega^1$  or  $\omega^2$  a *discrete net*.

From now on, we assume that the law  $a^{(2)}$  is of the form

$$a^{(2)}(x_1, x_2) := (1 - \varepsilon)1_{\{x_1=x_2\}}a(x_1) + \varepsilon a(x_1)a(x_2), \quad (1.3)$$

which means that with probability  $1 - \varepsilon$ , a single arrow comes out of  $(x, t)$  with law  $a(\cdot)$ ; and with probability  $\varepsilon$ , two independent arrows with common law  $a(\cdot)$  are drawn from  $(x, t)$ , which coalesce at time  $t + 1$  if they coincide. Let us denote the resulting discrete net by  $N_\varepsilon$ , which we call a *discrete net with (binary) branching probability  $\varepsilon$  and jump distribution  $a(\cdot)$* . The branching is weak in the sense that we will consider a sequence of such models with  $\varepsilon \downarrow 0$ .

Our main result is to show that under diffusive scaling of space-time (recall  $\sigma$  from (1.1))

$$S_{\varepsilon, \sigma} : (x, t) \rightarrow (\varepsilon x, \sigma^2 \varepsilon^2 t) \quad (1.4)$$

with the same  $\varepsilon$  as the branching probability in  $N_\varepsilon$ , the discrete net  $N_\varepsilon$  converges to a Brownian net  $\mathcal{N}$ . With a slight overload of notation, for each path  $\pi : [\sigma_\pi, \infty) \rightarrow \mathbb{R}$ , we will use  $S_\varepsilon \pi$  to denote the path whose graph is the image of the graph of  $\pi$  under  $S_\varepsilon$ , and for each set of paths  $K$ , we will use  $S_\varepsilon K$  to denote the set of paths obtained by applying  $S_\varepsilon$  to each path in  $K$ . In Section 2.1, we will recall the topology introduced in [FINR04] for the Brownian web  $\mathcal{W}$ . In particular,  $N_\varepsilon$  and  $\mathcal{N}$  are both random subsets of the path space  $(\Pi, d)$ . Furthermore,  $\mathcal{N}$  and  $\overline{N}_\varepsilon$ , the closure of  $N_\varepsilon$  in  $(\Pi, d)$  (see Remark 2.1), are both compact so that they are random variables taking values in the complete separable metric space  $(\mathcal{H}, d_{\mathcal{H}})$  of compact subsets of  $(\Pi, d)$ . To keep the notation simple, we will often write  $N_\varepsilon$  instead of  $\overline{N}_\varepsilon$ . Our main convergence result can then be formulated as follows.

**Theorem 1.1 (Convergence to the Brownian net)** *For  $\varepsilon > 0$ , let  $N_\varepsilon$  be the discrete net with (binary) branching probability  $\varepsilon$  and jump distribution  $a(\cdot)$  satisfying (1.1). Then*

$$S_{\varepsilon, \sigma} N_\varepsilon \xrightarrow[\varepsilon \rightarrow 0]{} \mathcal{N}, \quad (1.5)$$

where  $\mathcal{N}$  is the standard Brownian net, and  $\Rightarrow$  denotes weak convergence of  $\mathcal{H}$ -valued random variables.

**Remark 1.2** *In Theorem 1.1, the finite  $(3 + \eta)$ -th moment assumption on  $a(\cdot)$  in (1.1) is almost optimal. It was shown in [BMSV06, Theorem 1.1 (ii)] that if  $\sum_{x \in \mathbb{Z}} \frac{|x|^\beta}{\log^\beta(|x|+2)} a(x) = \infty$  for some  $\beta > 1$ , then the laws of the rescaled discrete web  $(S_{\varepsilon, \sigma} W)_{\varepsilon \in (0,1)}$  fails to be tight due to the presence of arbitrarily large jumps originating from any finite space-time domain. Therefore the laws of  $(S_{\varepsilon, \sigma} N_\varepsilon)_{\varepsilon \in (0,1)}$  also fail to be tight for such  $a(\cdot)$ .*

**Remark 1.3** *In Theorem 1.1, if we used the scaling map  $\tilde{S}_{\varepsilon, \sigma} : (x, t) \rightarrow (\varepsilon x / \sigma, \varepsilon^2 t)$  instead of  $S_{\varepsilon, \sigma}$  defined in (1.4), then the limit in (1.5) will be the image of  $\mathcal{N}$  under the map  $(x, t) \rightarrow (x/\sigma, t/\sigma^2)$ , which is a Brownian net  $\mathcal{N}_\sigma$  with branching parameter  $\sigma$ , see Theorem 2.7.*

## 1.2 Proof strategy

We first recall from [SS08] the proof strategy for branching-coalescing simple random walks on  $\mathbb{Z}$ , which is a non-crossing system where paths must meet before crossing (see also the survey article [SSS17]). We will then explain the fundamental difficulties posed by crossing systems and how our proof strategy overcomes these difficulties for the system of branching-coalescing random walks we consider.

**Non-crossing Case.** In [SS08], the authors considered the discrete net  $N_\varepsilon$  on the even space-time lattice  $\mathbb{Z}_{\text{even}}^2 := \{(x, t) \in \mathbb{Z}^2 : x + t \text{ is even}\}$ , where the jump distribution is given by  $a(\pm 1) = 1/2$  and each  $(x, t) \in \mathbb{Z}_{\text{even}}^2$  is a branching point with probability  $\varepsilon$ . The key observation behind the construction of the Brownian net  $\mathcal{N}$  in [SS08] is that, if we consider the discrete webs  $W_\varepsilon^l, W_\varepsilon^r \subset N_\varepsilon$ , which consist of the leftmost (reps. rightmost) paths in  $N_\varepsilon$  starting from every  $(x, t) \in \mathbb{Z}_{\text{even}}^2$ , then under the diffusive scaling map  $S_\varepsilon := S_{\varepsilon, 1}$  (with  $\sigma = 1$ ),  $S_\varepsilon W_\varepsilon^l$  (resp.  $S_\varepsilon W_\varepsilon^r$ ) converges in distribution to a Brownian web  $\mathcal{W}^l$  (resp.  $\mathcal{W}^r$ ) consisting of coalescing Brownian motions with drift  $-1$  (resp.  $+1$ ). Furthermore,  $(S_\varepsilon W_\varepsilon^l, S_\varepsilon W_\varepsilon^r)$  converges jointly to a pair of coupled Brownian webs  $(\mathcal{W}^l, \mathcal{W}^r)$ , called the *left-right web*, where paths in  $\mathcal{W}^l$  and  $\mathcal{W}^r$  undergo sticky interaction. The fact that paths in  $N_\varepsilon$  cannot cross without meeting ensures that there is a well-defined leftmost (resp. rightmost) path in  $N_\varepsilon$  starting from each  $(x, t) \in \mathbb{Z}_{\text{even}}^2$ .

Note that paths in the discrete net  $N_\varepsilon$  can always be constructed by concatenating paths in the pair of discrete webs  $(W_\varepsilon^l, W_\varepsilon^r)$ . This motivated the construction of the Brownian net  $\mathcal{N}$  in [SS08] by concatenating paths in the left-right web  $(\mathcal{W}^l, \mathcal{W}^r)$  and then taking closure in the path space  $(\Pi, d)$  (concatenation between two paths is only allowed at a time of coincidence between the two paths that is strictly larger than their starting times). This is called the *hopping construction* of the *standard Brownian net*  $\mathcal{N}$  from the left-right web  $(\mathcal{W}^l, \mathcal{W}^r)$ , where standard refers to the fact that the Brownian paths in  $\mathcal{W}^l$  and  $\mathcal{W}^r$  have drift  $-1$  and  $+1$  respectively.

In [SS08], the proof of convergence of the rescaled discrete nets  $(S_\varepsilon N_\varepsilon)_{\varepsilon \in (0, 1)}$  to the Brownian net  $\mathcal{N}$  consists of three steps: a lower bound via hopping, an upper bound via wedges, and tightness via leftmost and rightmost paths.

To show the lower bound that any subsequential weak limit of  $(S_\varepsilon N_\varepsilon)_{\varepsilon \in (0, 1)}$  contains a copy of  $\mathcal{N}$ , thanks to the hopping construction of the Brownian net, it suffices to show that the concatenation between any two paths in  $(\mathcal{W}^l, \mathcal{W}^r)$  can be approximated by the concatenation of paths in the rescaled discrete webs  $(S_\varepsilon W_\varepsilon^l, S_\varepsilon W_\varepsilon^r)$ . This is an easy consequence of the convergence  $(S_\varepsilon W_\varepsilon^l, S_\varepsilon W_\varepsilon^r) \Rightarrow (\mathcal{W}^l, \mathcal{W}^r)$ .

To show the upper bound that any subsequential weak limit of  $(S_\varepsilon N_\varepsilon)_{\varepsilon \in (0, 1)}$  contains no additional paths besides  $\mathcal{N}$ , the proof in [SS08] (also in [EFS17]) relied on the so-called *wedge characterization* of the Brownian net given in [SS08] (see also the survey [SSS17]), which uses the dual of the left-right web  $(\mathcal{W}^l, \mathcal{W}^r)$  to identify forbidden regions (called wedges) that cannot be entered by any path in  $\mathcal{N}$ . But the definition of wedges also requires  $N_\varepsilon$  to have well-defined leftmost and rightmost paths.

Proving tightness for the family of rescaled discrete nets  $(S_\varepsilon N_\varepsilon)_{\varepsilon \in (0, 1)}$  amounts to controlling the uniform modulus of continuity of paths in  $S_\varepsilon N_\varepsilon$ , see [FINR04, Prop. B.1] or Prop. 6.1 below. Since all paths in  $N_\varepsilon$  starting from each  $(x, t)$  are bounded between the unique path in  $W_\varepsilon^l$ , resp.  $W_\varepsilon^r$ , starting from  $(x, t)$ , the uniform modulus of continuity of paths in  $N_\varepsilon$  is controlled by the uniform modulus of continuity of paths in  $W_\varepsilon^l \cup W_\varepsilon^r$ . The tightness of  $(S_\varepsilon N_\varepsilon)_{\varepsilon \in (0, 1)}$  then follows directly from the tightness of the rescaled discrete webs  $(S_\varepsilon W_\varepsilon^l)_{\varepsilon \in (0, 1)}$  and  $(S_\varepsilon W_\varepsilon^r)_{\varepsilon \in (0, 1)}$ , which follows from the convergence of  $S_\varepsilon W_\varepsilon^l \Rightarrow \mathcal{W}^l$  and  $S_\varepsilon W_\varepsilon^r \Rightarrow \mathcal{W}^r$ .

Of the three steps above, the upper bound and tightness proof depend crucially on the non-crossing property, which implies that every path in  $N_\varepsilon$  is bounded between the leftmost and rightmost path in  $N_\varepsilon$  starting from the same point. These arguments have to be replaced

by completely new ideas in the crossing case. Only the idea of giving a lower bound via hopping has a natural extension in the crossing case.

We now explain our proof strategy for the branching-coalescing systems we consider, which belongs to the crossing case.

**Lower Bound.** Since there is no longer a discrete left-right web  $(W_\varepsilon^l, W_\varepsilon^r)$ , we will sample two discrete webs  $(W^1, W^2)$  as defined after (1.2), so that the discrete net  $N_\varepsilon$  can still be constructed by concatenating paths in  $W^1$  and  $W^2$ . We will show in Theorem 3.1 that  $(S_{\varepsilon,\sigma}W^1, S_{\varepsilon,\sigma}W^2)$  converges in distribution to a pair of sticky Brownian webs  $(\mathcal{W}^1, \mathcal{W}^2)$ , whose construction and characterisation will be recalled in Section 2.3. Similar to the hopping construction of the Brownian net  $\mathcal{N}$  from the left right web  $(\mathcal{W}^l, \mathcal{W}^r)$ , we can also construct a Brownian net by concatenating paths in  $\mathcal{W}^1 \cup \mathcal{W}^2$  and then taking closure. The lower bound is then obtained by showing that every subsequential weak limit of  $(S_{\varepsilon,\sigma}N_\varepsilon)_{\varepsilon \in (0,1)}$  contains a copy of  $\mathcal{N}$  constructed from  $(\mathcal{W}^1, \mathcal{W}^2)$  via the hopping construction. The key step in the proof is the convergence  $(S_{\varepsilon,\sigma}W^1, S_{\varepsilon,\sigma}W^2) \Rightarrow (\mathcal{W}^1, \mathcal{W}^2)$ , which requires finding a suitable harmonic function for a random walk with sticky interaction at the origin (see Prop. 3.4).

**Upper Bound.** To show that every subsequential weak limit of  $(S_{\varepsilon,\sigma}N_\varepsilon)_{\varepsilon \in (0,1)}$  contains no additional paths besides the Brownian net  $\mathcal{N}$  constructed in the lower bound, we use a different density argument than in [SS08]. The key observation is that we can dominate the discrete net  $N_\varepsilon$  by an auxiliary branching-coalescing system which we call the *Bernoulli net*  $\tilde{N}_\varepsilon$ , see Section 4. In particular, the Bernoulli net allows branching with arbitrary number of offsprings. What is special about the Bernoulli net  $\tilde{N}_\varepsilon$  is that we can identify explicitly a product invariant law for the branching-coalescing random walks in  $\tilde{N}_\varepsilon$ , which converges under diffusive scaling to a Poisson point process with intensity 2 on  $\mathbb{R}$  and is precisely the invariant law for the branching-coalescing Brownian motions in the Brownian net  $\mathcal{N}$ . This convergence can then be used to rule out the existence of additional paths in the limit besides paths in  $\mathcal{N}$ . A key ingredient in our proof is Theorem 4.6, which shows that for the set of paths in the Bernoulli net  $\tilde{N}_\varepsilon$  started from every point in  $\mathbb{Z}$  at time 0, their density at large time  $t > 0$  still decays at the rate of  $1/\sqrt{t}$ , which is comparable to coalescing random walks without branching.

**Tightness.** In contrast to the non-crossing case, where tightness comes for free as a result of tightness of the rescaled left right webs  $(S_{\varepsilon,\sigma}W_\varepsilon^l, S_{\varepsilon,\sigma}W_\varepsilon^r)_{\varepsilon \in (0,1)}$ , in our case, tightness is the most technical part of the proof. The overall strategy follows the multiscale argument developed in [BMSV06], which proves tightness for the rescaled discrete webs (coalescing random walks)  $(S_{\varepsilon,\sigma}W)_{\varepsilon \in (0,1)}$  under the same assumption on  $a(\cdot)$  as in (1.1). We go through an exponentially increasing sequence of space-time scales. From one scale to the next, we exploit the coalescence to reduce the density of walks as the scale increases, and at the same time, we control the probability that at least one of the walks makes a large excursion that leads to a poor modulus of continuity. Although this overall strategy is not new, the technical difficulties of incorporating branching are serious and we find it the most difficult part of the proof and took us the longest to complete.

For branching-coalescing random walks with weak branching as we consider, coalescence dominates on small scales. So the density reduction argument can still be applied. However, the difficulty is that even though the branching is weak, a single random walk could still have an arbitrarily large number of descendants on large time scales. Controlling how far a single random walk and any of its descendants can travel within a short time becomes the key challenge in the crossing case, while in the non-crossing case, it suffices to consider the large deviation of the leftmost and the rightmost random walk among all the descendants. We overcome this difficulty by controlling the number of *relevant branching points* in the discrete net  $N_\varepsilon$  (see Theorem 6.9), which are space-time points where a descendant of the

original random walk (starting at some initial time  $S$ ) branches into two random walks, and each random walk has a descendant such that the two descendants do not meet before some terminal time  $T$ . The relevant branching points are the branching points that lead to new random walk positions at time  $T$ . To control the large deviation of all descendants of a single random walk, it suffices to group the descendants according to the choice made at each relevant branching point and then control the large deviation of a random walk sampled from each group.

**Remark 1.4** *The universality of the Brownian net  $\mathcal{N}$  should hold beyond the family of branching-coalescing random walks considered in Theorem 1.1. For example, we could consider branching-coalescing random walks in continuous time that arise naturally as the dual of biased voter models [SSY19, SSY21]. In particular, consider the setting where each random walk jumps at rate 1 with kernel  $a(\cdot)$  satisfying the assumptions in (1.1), and at rate  $\varepsilon > 0$ , each random walk gives birth to a new particle which immediately makes a jump with kernel  $a(\cdot)$ , while particles coalesce when they meet. We believe our proof can be extended without much difficulty to this continuous time model. One element becomes even simpler, namely that by a generator calculation, it can be shown that this branching-coalescing random walk admits a product invariant measure. Therefore we no longer need the Bernoulli net introduced in Section 4. We could also consider other systems such as branching-coalescing random walks with non-binary branching, although these lie beyond our current techniques because they do not admit explicit invariant measures and cannot be embedded in a Bernoulli net. Nevertheless, we hope that our methods can be extended in the future to cover such systems as well.*

### 1.3 Outline

The rest of the paper is organized as follows. In Section 2, we recall some necessary background on the Brownian web, the Brownian net, and sticky Brownian webs. Assuming tightness of the family of rescaled discrete nets  $(S_{\varepsilon,\sigma}\overline{N}_\varepsilon)_{\varepsilon \in (0,1)}$ , we show in Section 3 that every subsequential weak limit of  $S_{\varepsilon,\sigma}\overline{N}_\varepsilon$  contains a copy of the standard Brownian net  $\mathcal{N}$ . In Section 4, we introduce an auxiliary Bernoulli net which admits an explicit invariant measure and dominates the discrete net. In Section 5, we use the domination by the Bernoulli net to show that every subsequential weak limit of  $S_{\varepsilon,\sigma}\overline{N}_\varepsilon$  contains no additional paths besides the Brownian net  $\mathcal{N}$ . In Section 6, we prove the tightness for the family of rescaled discrete nets  $(S_{\varepsilon,\sigma}\overline{N}_\varepsilon)_{\varepsilon \in (0,1)}$ . Lastly, in Appendix A, we sketch the proof of the equivalence between two alternative constructions of a pair of sticky Brownian webs, and the construction of the Brownian net from sticky Brownian webs.

## 2 Background

We first recall in Subsection 2.1 the space of compact sets of paths in which the Brownian web  $\mathcal{W}$  and the Brownian net  $\mathcal{N}$  take their values. Then in Subsections 2.2, 2.3 and 2.4, we introduce necessary results on the Brownian web, sticky Brownian webs, and the Brownian net, respectively. Proposition 2.6 below on the equivalence of two alternative constructions of a pair of sticky Brownian webs and Theorem 2.7 below on the construction of the Brownian net from sticky Brownian webs have not been proved explicitly in the literature. We will sketch their proof in Appendix A.

### 2.1 Space of compact sets of paths

We recall here the Polish space  $\mathcal{H}$  of compact sets of paths, where the Brownian web and the Brownian net take their values. This space was first introduced in [FINR04], and we follow the variant defined in [SS08, Appendix].

Let  $R_c^2 := \mathbb{R}^2 \cup \{(\pm\infty, t) : t \in \mathbb{R}\} \cup \{(*, \pm\infty)\}$  be the compactification of  $\mathbb{R}^2$  obtained by equipping the set  $R_c^2$  with a topology such that  $(x_n, t_n) \rightarrow (\pm\infty, t)$  if  $x_n \rightarrow \pm\infty$  and  $t_n \rightarrow t \in \mathbb{R}$ , and  $(x_n, t_n) \rightarrow (*, \pm\infty)$  if  $t_n \rightarrow \pm\infty$  (regardless of the behavior of  $x_n$ ). Such a compactification can be achieved by taking the completion of  $\mathbb{R}^2$  with respect to the metric

$$\rho((x_1, t_1), (x_2, t_2)) = |\Theta_1(x_1, t_1) - \Theta_1(x_2, t_2)| \vee |\Theta_2(t_1) - \Theta_2(t_2)|, \quad (2.1)$$

where the map  $\Theta = (\Theta_1, \Theta_2)$  is defined by

$$\Theta(x, t) = (\Theta_1(x, t), \Theta_2(t)) := \left( \frac{\tanh(x)}{1 + |t|}, \tanh(t) \right). \quad (2.2)$$

We can think of  $R_c^2$  as the image of  $[-\infty, \infty]^2$  under the map  $\Theta$ .

A *path* in  $R_c^2$  is any continuous function  $\pi : [\sigma_\pi, \infty] \rightarrow [-\infty, \infty] \cup \{*\}$ , with  $\sigma_\pi \in [-\infty, \infty]$ , such that  $\pi : [\sigma_\pi, \infty] \cap \mathbb{R} \rightarrow [-\infty, \infty]$  is continuous, and  $\pi(\pm\infty) = *$  whenever  $\pm\infty \in [\sigma_\pi, \infty]$ . Equivalently, if we identify  $R_c^2$  with the image of  $[-\infty, \infty]^2$  under the map  $\Theta$ , then  $\pi$  can be identified with its graph, which is a compact subset of  $\Theta([- \infty, \infty]^2)$  with the property that for each  $t \in [\Theta_2(\sigma_\pi), \Theta_2(\infty)]$ , there is exactly one  $x$  with  $(x, t)$  on the graph of  $\pi$ . We will often identify  $\pi$  with its graph in the rest of the paper.

Let  $\Pi$  denote the space of all paths in  $R_c^2$ , equipped with the metric

$$d(\pi_1, \pi_2) := |\Theta_2(\sigma_{\pi_1}) - \Theta_2(\sigma_{\pi_2})| \vee \sup_{t \geq \sigma_{\pi_1} \wedge \sigma_{\pi_2}} |\Theta_1(\pi_1(t \vee \sigma_{\pi_1}), t) - \Theta_1(\pi_2(t \vee \sigma_{\pi_2}), t)|. \quad (2.3)$$

The space  $(\Pi, d)$  is complete and separable. Note that paths converge in  $(\Pi, d)$  if and only if their starting times converge and the functions converge locally uniformly on  $\mathbb{R}$ .

Let  $\mathcal{H}$  be the space of compact subsets of  $(\Pi, d)$ , equipped with the Hausdorff metric  $d_{\mathcal{H}}$ . Namely, for  $\mathcal{X}_1, \mathcal{X}_2 \in \mathcal{H}$ ,

$$d_{\mathcal{H}}(\mathcal{X}_1, \mathcal{X}_2) = \sup_{x_1 \in \mathcal{X}_1} \inf_{x_2 \in \mathcal{X}_2} d(x_1, x_2) \vee \sup_{x_2 \in \mathcal{X}_2} \inf_{x_1 \in \mathcal{X}_1} d(x_1, x_2). \quad (2.4)$$

The metric space  $(\mathcal{H}, d_{\mathcal{H}})$  is also complete and separable. For more properties, see e.g. [SSS14, Appendix B]. Let  $(\widehat{\mathcal{H}}, d_{\widehat{\mathcal{H}}})$  be defined similarly to  $(\mathcal{H}, d_{\mathcal{H}})$ , except that all paths run backwards in time.

**Remark 2.1** *To view the discrete nets  $N_\varepsilon$  as random variables taking values in  $\mathcal{H}$ , we need to take the closure of  $N_\varepsilon$  in  $(\Pi, d)$  and show that  $\overline{N}_\varepsilon$  is compact. This is easily verified if the jump distribution  $a(\cdot)$  in (1.1) has a finite first moment, see e.g. [NRS05, Lemma 1.1]. It is then easily seen that  $\overline{N}_\varepsilon \setminus N_\varepsilon$  only contains the trivial paths  $\pi$  with  $\sigma_\pi \in \{\pm\infty\} \cup \mathbb{Z}$ , and either  $\pi(t) \equiv -\infty$  or  $\pi(t) \equiv \infty$  for all  $t \in [\sigma_\pi, \infty)$ . To keep the notation simple, we will write  $N_\varepsilon$  instead of  $\overline{N}_\varepsilon$  in the rest of the paper.*

We will adopt the following notational convention:

- For  $\mathcal{X} \subset \mathcal{H}$  and  $A \subset R_c^2$ , let  $\mathcal{X}(A)$  denote the subset of paths in  $\mathcal{X}$  with starting points in  $A$ . Sometimes we will also write  $\mathcal{X}(A)$  as  $\mathcal{X}^A$ .
- For  $A = \{z\}$ , we simply write  $\mathcal{X}(z)$  instead of  $\mathcal{X}(\{z\})$ .

## 2.2 The Brownian web

We recall from [FINR04] the characterization of the Brownian web  $\mathcal{W}$  as a random variable taking values in the complete separable metric space  $(\mathcal{H}, d_{\mathcal{H}})$ .



**Theorem 2.2 (Characterization of the Brownian web)** *There exists an  $\mathcal{H}$ -valued random variable  $\mathcal{W}$ , called the standard Brownian web, whose distribution is uniquely determined by the following properties:*

- (i) *For each deterministic  $z \in \mathbb{R}^2$ , almost surely there is a unique path  $\pi_z \in \mathcal{W}(z)$ .*
- (ii) *For any finite deterministic set of points  $z_1, \dots, z_k \in \mathbb{R}^2$ , the paths  $(\pi_{z_1}, \dots, \pi_{z_k})$  are distributed as a collection of coalescing standard Brownian motions.*
- (iii) *For any deterministic countable dense subset  $\mathcal{D} \subset \mathbb{R}^2$ , almost surely,  $\mathcal{W}$  is the closure of  $\{\pi_z : z \in \mathcal{D}\}$  in  $(\Pi, d)$ .*

The Brownian web  $\mathcal{W}$  can be coupled to a dual Brownian web  $\widehat{\mathcal{W}} \in \widehat{\mathcal{H}}$ , which consists of coalescing Brownian motions running backwards in time and has the same distribution as  $\mathcal{W}$  if space-time is rotated by  $180^\circ$  around the origin. Almost surely,  $\mathcal{W}$  and  $\widehat{\mathcal{W}}$  determine each other by the condition that paths in  $\mathcal{W}$  cannot cross paths in  $\widehat{\mathcal{W}}$ , i.e., if  $\pi \in \mathcal{W}$  and  $\hat{\pi} \in \widehat{\mathcal{W}}$ , then we cannot find  $s < t$  with  $(\pi(s) - \hat{\pi}(s))(\pi(t) - \hat{\pi}(t)) < 0$ . Furthermore, paths in  $\mathcal{W}$  and  $\widehat{\mathcal{W}}$  interact through Skorohod reflection. For further properties and alternative characterizations of  $(\mathcal{W}, \widehat{\mathcal{W}})$ , see the survey article [SSS17] and the references therein.

The following analogue of Theorem 1.1 for the Brownian web  $\mathcal{W}$  is due to [NRS05], which assumed finite 5-th moment for the random walk increments in the proof of tightness, and it was then improved in [BMSV06] to the nearly optimal assumption of a finite  $(3 + \eta)$ -th moment.

**Theorem 2.3 (Convergence to the Brownian web)** *Let  $W$  be the discrete web defined in (1.2) with jump distribution  $a(\cdot)$  satisfying (1.1). Then*

$$S_{\varepsilon, \sigma} W \xrightarrow[\varepsilon \rightarrow 0]{} \mathcal{W}, \quad (2.5)$$

where  $\mathcal{W}$  is the standard Brownian web, and  $\Rightarrow$  denotes weak convergence of  $\mathcal{H}$ -valued random variables.

### 2.3 Sticky Brownian webs

The Brownian net  $\mathcal{N}$  was first constructed in [SS08] (see [NRS10] for a later alternative construction), which generalises the Brownian web  $\mathcal{W}$  by allowing paths to branch. An intermediate object in the construction of the Brownian net  $\mathcal{N}$  in [SS08] is a pair of sticky Brownian webs  $(\mathcal{W}^1, \mathcal{W}^2)$  with drift  $-1$  and  $+1$  respectively. For our purpose, we will give an alternative construction using a pair of *sticky Brownian webs*  $(\mathcal{W}^1, \mathcal{W}^2)$  with the same drift 0.

We first motivate the sticky Brownian webs (with drift 0) through discrete approximation. Recall the definition of the discrete net  $N_\varepsilon$  in Section 1.1, which is constructed from an i.i.d. family of random pairs  $(\omega^1(x, t), \omega^2(x, t))_{(x, t) \in \mathbb{Z}^2}$  with common law  $a^{(2)}$  that satisfies (1.3). Each family  $(\omega^i(x, t))_{(x, t) \in \mathbb{Z}^2}$ ,  $i = 1, 2$ , determines a discrete web  $W^i$ , and by Theorem 2.3,  $S_{\varepsilon, \sigma} W^i$  converges in distribution to a Brownian web  $\mathcal{W}^i$ . In the special case when  $\omega^i(x, t) \in \{\pm 1\}$  so that the paths make nearest neighbour jumps, and when we only consider paths starting from the even lattice sites  $\mathbb{Z}_{\text{even}}^2 := \{(x, t) \in \mathbb{Z}^2 : x + t \text{ is even}\}$  to avoid parity issues, results from [SSS14] imply that the pair  $(S_{\varepsilon, \sigma} W^1, S_{\varepsilon, \sigma} W^2)$  converges to a pair of coupled Brownian webs  $(\mathcal{W}^1, \mathcal{W}^2)$ , called *sticky Brownian webs* because paths between  $\mathcal{W}^1$  and  $\mathcal{W}^2$  undergo sticky interaction.

Instead of the discrete webs  $(W^1, W^2)$  defined above, [SS08] considered the discrete webs  $W_\varepsilon^1$  and  $W_\varepsilon^2$ , induced respectively by

$$\omega^1(x, t) = \omega^1(x, t) \wedge \omega^2(x, t), \quad \omega^2(x, t) = \omega^1(x, t) \vee \omega^2(x, t), \quad (2.6)$$

so that whenever two (nearest neighbour) arrows originate from  $(x, t)$ , paths in  $W_\varepsilon^l$  follow the arrow to the left and paths in  $W_\varepsilon^r$  follow the arrow to the right. It was shown in [SS08] that  $(S_{\varepsilon, \sigma} W_\varepsilon^l, S_{\varepsilon, \sigma} W_\varepsilon^r)$  converges to a pair of sticky Brownian webs  $(\mathcal{W}^l, \mathcal{W}^r)$  with drift  $-1$  and  $+1$  respectively, called the *left-right Brownian web*. The key observation of [SS08] is that, since the discrete net  $N_\varepsilon$  can be constructed by concatenating paths in the two discrete webs  $W_\varepsilon^l$  and  $W_\varepsilon^r$  (i.e., switching back and forth between arrows induced by either  $\omega^l$  and  $\omega^r$ ), the limit of  $S_{\varepsilon, \sigma} N_\varepsilon$  can also be recovered by concatenating paths in the pair of sticky Brownian webs  $(\mathcal{W}^l, \mathcal{W}^r)$ . This was called the *hopping construction* of the Brownian net  $\mathcal{N}$ . Since the discrete net  $N_\varepsilon$  can also be constructed by concatenating paths in the two discrete webs  $W^l$  and  $W^r$  (i.e., switching back and forth between arrows induced by either  $\omega^l$  and  $\omega^r$ ), we can also recover  $\mathcal{N}$  by concatenating paths in the pair of sticky Brownian webs  $(\mathcal{W}^l, \mathcal{W}^r)$  with drift 0. This is the content of Theorem 2.7 below.

We will mostly work with the sticky Brownian webs  $(\mathcal{W}^1, \mathcal{W}^2)$  with drift 0, for which we will need two equivalent characterizations. The first was introduced in [HW09b], which we call the *martingale characterization*. We will use it to prove  $(S_{\varepsilon, \sigma} W^1, S_{\varepsilon, \sigma} W^2) \Rightarrow (\mathcal{W}^1, \mathcal{W}^2)$  for general jump distribution  $a(\cdot)$  satisfying (1.1). The second was introduced in [SSS14], which we call the *marking construction*. It constructs the pair of sticky Brownian webs  $(\mathcal{W}^1, \mathcal{W}^2)$  and the Brownian net  $\mathcal{N}$  simultaneously through a Poisson marking procedure that was first developed in [NRS10]. This allows us to recover the Brownian net  $\mathcal{N}$  from  $(\mathcal{W}^1, \mathcal{W}^2)$  through the *hopping construction* sketched above, stated in Theorem 2.7 below. We will show that these two characterizations of the sticky Brownian webs give the same object, so that we can choose the most convenient characterization depending on the context.

**Martingale Characterization.** Following [HW09a],  $(Y^1, Y^2) = (Y_t^1, Y_t^2)_{t \geq 0}$  is called a pair of *sticky Brownian motions with parameter  $\theta \geq 0$*  if  $Y^1$  and  $Y^2$  are standard Brownian motions w.r.t. some common filtration  $(\mathcal{F}_t)_{t \geq 0}$ , and the following processes are martingales w.r.t. this filtration

$$\begin{aligned} \text{(i)} \quad & Y_t^1 Y_t^2 - \int_0^t 1_{\{Y_s^1 = Y_s^2\}} ds, \\ \text{(ii)} \quad & |Y_t^1 - Y_t^2| - 2\theta \int_0^t 1_{\{Y_s^1 = Y_s^2\}} ds. \end{aligned} \tag{2.7}$$

By [HW09b, Prop. 3], we have the following fact.

**Proposition 2.4 (Characterization of sticky Brownian motions)** *For each  $\theta \geq 0$  and  $(y_1, y_2) \in \mathbb{R}^2$ , there exists a pair  $(Y^1, Y^2)$  of sticky Brownian motions with parameter  $\theta \geq 0$  and initial state  $(Y_0^1, Y_0^2) = (y_1, y_2)$ , and such a pair is unique in law.*

We can readily extend the definition of sticky Brownian motions to the case where  $Y^1$  and  $Y^2$  start at different times. Note that the smaller the value of  $\theta$  is, the stronger is the stickiness. When  $\theta = 0$ ,  $(Y^1, Y^2)$  is a pair of coalescing Brownian motions. The case where  $Y^1, Y^2$  have nonzero drifts is treated in [HW09b, Prop. 14]. In particular, when  $Y^1$  has a smaller drift than  $Y^2$  and the pair  $(Y^1, Y^2)$  stays ordered once they meet,  $(Y^1, Y^2)$  is equal in distribution to the unique solution of the *left-right SDE* in [SS08, Sec. 2].

Note that each Brownian web  $\mathcal{W}^i$  defines in a natural way a filtration  $\mathcal{F}_t^i$ , which is the  $\sigma$ -field generated by the restriction of all paths in  $\mathcal{W}^i$  to times  $\leq t$ . Let  $\mathcal{F}_t := \mathcal{F}_t^1 \vee \mathcal{F}_t^2 := \sigma(\mathcal{F}^1 \cup \mathcal{F}^2)$  be the smallest  $\sigma$ -field containing  $\mathcal{F}_t^1$  and  $\mathcal{F}_t^2$ . Then  $\mathcal{F}_t$  also forms a filtration. We can now state the martingale characterization of a pair of sticky Brownian webs  $(\mathcal{W}^1, \mathcal{W}^2)$ , which is a variant of [HW09b, Thm. 4] and [SS19, Thm. 1].

**Theorem 2.5 (Martingale characterization of sticky Brownian webs)** *There exists an  $\mathcal{H} \times \mathcal{H}$ -valued random variable  $(\mathcal{W}^1, \mathcal{W}^2)$ , called a pair of sticky Brownian webs with parameter  $\theta \geq 0$ , whose distribution is uniquely determined by the following properties:*

- (i) Both  $\mathcal{W}^1$  and  $\mathcal{W}^2$  are distributed as the standard Brownian web;
- (ii) For each deterministic pair  $z = (x, t), z' = (x', t') \in \mathbb{R}^2$ , the pair of paths  $(X_z^1, X_{z'}^2)$ , with  $X_z^1 \in \mathcal{W}^1(z)$  and  $X_{z'}^2 \in \mathcal{W}^2(z')$ , is distributed as a pair of sticky Brownian motions with parameter  $\theta$  relative to the filtration  $(\mathcal{F}_u)_{u \in \mathbb{R}}$ ;
- (iii) (Co-adaptedness) For each deterministic pair  $z = (x, t), z' = (x', t') \in \mathbb{R}^2$ , if  $X_z^i \in \mathcal{W}^i(z)$  and  $X_{z'}^j \in \mathcal{W}^j(z')$  for  $i, j = 1, 2$ , then the pair  $(X_z^i, X_{z'}^j)$  is a Markov process relative to the filtration  $(\mathcal{F}_u)_{u \in \mathbb{R}}$ .

This formulation follows that of [SS19, Thm. 1] for the *left-right Brownian web*, cf. (2.6). The formulation for sticky Brownian webs was discussed in the proof of [SS19, Thm. 1]. Here, co-adaptedness is formulated differently than in [HW09b, Thm. 4] in order to make it more clear how [H07, Thm. 76] can be applied to show that the law of  $(\mathcal{W}^1, \mathcal{W}^2)$  is uniquely determined by conditions (i)-(ii) plus co-adaptedness (see the proof of [SS19, Thm. 1] for more details).

**Marking Construction.** We now sketch the marking construction of the sticky Brownian webs  $(\mathcal{W}^1, \mathcal{W}^2)$  from [SSS14, Section 3], which is an extension of the marking construction of the Brownian net in [NRS10]. We refer the reader to these references for details.

Given a Brownian web  $\mathcal{W}$  and its dual  $\widehat{\mathcal{W}}$ , each pair  $(\pi, \hat{\pi}) \in (\mathcal{W}, \widehat{\mathcal{W}})$  interacts through Skorohod reflection. In particular, there is a well-defined *intersection local time measure*  $\ell$  on the intersection between the graphs of  $\pi$  and  $\hat{\pi}$ , for  $\pi \in \mathcal{W}$  and  $\hat{\pi} \in \widehat{\mathcal{W}}$  (see [SSS14, Prop. 3.4]). Each point of intersection  $(x, t)$  between a path  $\pi \in \mathcal{W}$  with  $\sigma_\pi < t$ , and  $\hat{\pi} \in \widehat{\mathcal{W}}$  with  $\sigma_{\hat{\pi}} > t$ , is called a *special point of  $\mathcal{W}$*  of type  $(1, 2)$  because  $\pi$  is the unique path in  $\mathcal{W}$  passing through  $(x, t)$ , and besides  $\pi$ , there is a second path  $\pi' \in \mathcal{W}$  that starts from  $(x, t)$ . We then define a Poisson point set  $S$  on all such  $(1, 2)$  points with intensity measure  $\theta \ell$ . Given  $\mathcal{W}$  and  $S$ , we then construct a second set of paths  $\mathcal{W}'$  such that paths in  $\mathcal{W}'$  follow paths in  $\mathcal{W}$ , except when a path  $p \in \mathcal{W}'$  enters a point  $(x, t) \in S$ , instead of continuing along  $\pi \in \mathcal{W}$  that passes through  $(x, t)$ ,  $p$  switches to the second path  $\pi' \in \mathcal{W}$  that starts from  $(x, t)$ . It was shown in [SSS14, Theorem 3.5] that  $\mathcal{W}'$  is well-defined, and the pair  $(\mathcal{W}, \mathcal{W}')$  is called a *pair of sticky Brownian webs with coupling parameter  $\theta$* . If we include both the path that continues along  $\pi$  and the path that continues along  $\pi'$ , then we obtain a Brownian net, see [SSS14, Theorem 3.5 and 6.15].

We will show in Appendix A that the two definitions of sticky Brownian webs above are equivalent.

**Proposition 2.6 (Equivalence)** *Let  $(\mathcal{W}^1, \mathcal{W}^2)$  be a pair of sticky Brownian webs with parameter  $\theta \geq 0$  defined in Theorem 2.5, and let  $(\mathcal{W}, \mathcal{W}')$  a pair of sticky Brownian webs with coupling parameter  $\theta$  defined via the marking construction above. Then  $(\mathcal{W}^1, \mathcal{W}^2)$  and  $(\mathcal{W}, \mathcal{W}')$  have the same distribution.*

Both the martingale characterization and the marking construction of sticky Brownian webs can be extended to the case with non-zero drifts and parameter  $\theta \geq 0$ . For simplicity, we have restricted to the case of zero drift. When  $(\mathcal{W}^1, \mathcal{W}^2)$  is a *left-right Brownian web*, where  $\mathcal{W}^1$  is a Brownian web with a smaller drift than  $\mathcal{W}^2$  and any pair  $(\pi_1, \pi_2) \in (\mathcal{W}^1, \mathcal{W}^2)$  stay ordered once they meet, it was shown in [SSS14, Lemma 6.18] that the marking construction of the left-right web is equivalent to a characterization similar to Theorem 2.5.

## 2.4 The Brownian net

The Brownian net  $\mathcal{N}$  was first defined in [SS08] as an extension of the Brownian web  $\mathcal{W}$  by allowing branching. To counter the instantaneous coalescence between paths, the *branching rate* is effectively infinite. The actual construction, called *hopping construction* in [SS08,

Thm. 1.3], was motivated by the discrete analogue discussed after (2.6), where  $\mathcal{N}$  was constructed by concatenating paths in the left-right web  $(\mathcal{W}^l, \mathcal{W}^r)$ .

The left-right web  $(\mathcal{W}^l, \mathcal{W}^r)$  is characterized by the properties that  $\mathcal{W}^l$ , resp.  $\mathcal{W}^r$ , is a Brownian web with drift  $-1$ , resp.  $+1$ , and any pair of paths  $(l, r) \in (\mathcal{W}^l, \mathcal{W}^r)$  with deterministic starting points evolve as a pair of sticky Brownian motions with drifts  $(-1, 1)$  and stays ordered once they meet. More generally, the left-right web  $(\mathcal{W}^l, \mathcal{W}^r)$  could have arbitrary drifts  $(a, b)$  with  $a < b$ . We refer the reader to [SS08] for the precise characterization of  $(\mathcal{W}^l, \mathcal{W}^r)$ . Our goal here is to show that  $\mathcal{N}$  can also be constructed by concatenation (hopping) between paths in the pair of sticky Brownian webs  $(\mathcal{W}^l, \mathcal{W}^r)$ .

First, we give the precise definition of hopping between paths in  $(\mathcal{W}^l, \mathcal{W}^r)$ , which also applies to  $(\mathcal{W}^l, \mathcal{W}^r)$ . Given two paths  $\pi, \pi' \in \Pi$ , we call  $t \in \mathbb{R}$  an *intersection time* if  $\sigma_\pi, \sigma_{\pi'} < t < \infty$  and  $\pi(t) = \pi'(t)$ . At such an intersection time  $t$ , we can then define a new path  $\pi''$  by concatenating the piece of  $\pi$  before  $t$  with the piece of  $\pi'$  after  $t$ , i.e., by setting  $\pi''(s) := \pi(s)$  for  $s \in [\sigma_\pi, t]$  and  $\pi''(s) = \pi'(s)$  for  $s \in [t, \infty)$ . For any collection of paths  $\mathcal{A} \subset \Pi$ , we let  $\mathcal{H}_{\text{int}}(\mathcal{A})$  denote the smallest set of paths containing  $\mathcal{A}$  that is closed under such ‘hopping’ from one path onto another at intersection times, i.e.,  $\mathcal{H}_{\text{int}}(\mathcal{A})$  is the set of all paths  $\pi \in \Pi$  of the form

$$\pi(s) := \pi_k(s) \text{ for } s \in [t_{k-1}, t_k], \quad 1 \leq k \leq m, \quad (2.8)$$

where  $\pi_1, \dots, \pi_m \in \mathcal{A}$ ,  $\sigma_{\pi_1} = t_0 < \dots < t_m = \infty$ , and  $t_k$  is an intersection time of  $\pi_k$  and  $\pi_{k+1}$  for each  $k = 1, \dots, m-1$ .

In [SS08, Thm. 1.3 and Prop. 1.4], the Brownian net  $\mathcal{N}$  was constructed by setting it to be the closure

$$\mathcal{N} := \overline{\mathcal{H}_{\text{int}}(\mathcal{W}^l \cup \mathcal{W}^r)}, \quad (2.9)$$

where  $(\mathcal{W}^l, \mathcal{W}^r)$  is a left-right Brownian web with drifts  $(-1, 1)$ . If  $(\mathcal{W}^l, \mathcal{W}^r)$  is a left-right web with drifts  $(-\theta, \theta)$ , then this construction defines the *Brownian net  $\mathcal{N}_\theta$  with branching parameter  $\theta$* , which we will simply denote by  $\mathcal{N}$  when  $\theta = 1$ . We call  $\mathcal{N}$  the *standard Brownian net*. It is easily seen that  $\mathcal{N}_\theta$  is the image of  $\mathcal{N}$  under the diffusive scaling map  $(x, t) \rightarrow (x/\theta, t/\theta^2)$ .

We have the following alternative hopping construction of the Brownian net, whose proof is deferred to Appendix A.

**Theorem 2.7 (Hopping construction of a Brownian net)** *Let  $(\mathcal{W}^l, \mathcal{W}^r)$  be a pair of sticky Brownian webs with parameter  $\theta$ . Then*

$$\mathcal{N}_\theta := \overline{\mathcal{H}_{\text{int}}(\mathcal{W}^l \cup \mathcal{W}^r)} \quad (2.10)$$

*is distributed as a Brownian net with branching parameter  $\theta$ .*

*If  $K$  is a compact subset of  $\mathbb{R}_c^2$  (see definition above (2.1)), then an analogue of (2.10) holds, i.e.,*

$$\mathcal{N}_\theta(K) = \overline{\mathcal{H}_{\text{int}}(\mathcal{W}^l \cup \mathcal{W}^r)(K)}. \quad (2.11)$$

As mentioned before, [NRS10] gave an alternative *marking construction* of  $\mathcal{N}$ . There are also alternative characterizations of the Brownian net, see [SS08] and the review article [SSS17]. For more details on the coupling between  $(\mathcal{W}^l, \mathcal{W}^r)$  and  $\mathcal{N}$  through Poisson marking of  $(1, 2)$  points of  $\mathcal{W}^l$ , which will be relevant in the proof of Theorem 2.7, see [SSS14, Theorem 3.5, 4.4, and 6.15].

We will need later the fact that the Brownian net is a maximal collection of paths in the sense that it contains every path in its graph (or image set). Let  $\Pi_t := \{\pi \in \Pi : \sigma_\pi = t\}$  denote the space of all paths starting at time  $t$ , and for any  $\mathcal{X} \in \mathcal{H}$ , let

$$\mathcal{X}_t := \mathcal{X} \cap \Pi_t \quad (2.12)$$

denote the subset of paths in  $\mathcal{X}$  starting at time  $t$ . Then the following result says that every path  $\pi \in \Pi_t$  (which is identified with its graph as a subset of  $\mathbb{R}^2$ ) that is contained in the union of the graphs of  $p \in \mathcal{N}_t$  must also be a path in  $\mathcal{N}_t$ , see [SS08, Prop. 1.13].

**Proposition 2.8 (Image set property)** *Let  $\mathcal{N}$  be the standard Brownian net. Then almost surely for all  $t \in [-\infty, \infty]$ ,*

$$\mathcal{N}_t = \left\{ \pi \in \Pi_t : \pi \subset \bigcup_{p \in \mathcal{N}_t} p \right\}. \quad (2.13)$$

### 3 The lower bound

The main result of this section is Theorem 3.5, which shows that every subsequential weak limit of the rescaled discrete net  $S_{\varepsilon, \sigma} N_\varepsilon$  contains a copy of the Brownian net  $\mathcal{N}$  constructed as in Theorem 2.7 by hopping between paths in a pair of sticky Brownian webs  $(\mathcal{W}^1, \mathcal{W}^2)$ . Our strategy is to first show that the pair of discrete sticky webs  $(W_\varepsilon^1, W_\varepsilon^2)$  defined as in (1.3) converges to a pair of sticky Brownian webs  $(\mathcal{W}^1, \mathcal{W}^2)$  as characterised in Theorem 2.5. We then show that the hopping between paths in  $\mathcal{W}^1 \cup \mathcal{W}^2$  can always be approximated by their discrete counterparts.

#### 3.1 Convergence of discrete sticky webs

The main result of this subsection is the convergence of a pair of discrete sticky webs and the intersection times between paths in the two discrete webs.

**Theorem 3.1 (Convergence of discrete sticky webs)** *Let  $\varepsilon > 0$ , and let  $(W_\varepsilon^1, W_\varepsilon^2)$  be the pair of discrete sticky webs defined as in (1.3) with branching probability  $\varepsilon > 0$  and jump distribution  $a(\cdot)$  satisfying (1.1). Then,*

$$(S_{\varepsilon, \sigma} W_\varepsilon^1, S_{\varepsilon, \sigma} W_\varepsilon^2) \xrightarrow[\varepsilon \downarrow 0]{} (\mathcal{W}^1, \mathcal{W}^2), \quad (3.1)$$

where  $(\mathcal{W}^1, \mathcal{W}^2)$  is a pair of sticky Brownian webs with parameter 1 as in Theorem 2.5.

**Proof** The convergence of the marginals  $S_{\varepsilon, \sigma} W_\varepsilon^1 \Rightarrow \mathcal{W}^1$  and  $S_{\varepsilon, \sigma} W_\varepsilon^2 \Rightarrow \mathcal{W}^2$  follows from Theorem 2.3. To prove their joint convergence, it suffices to show that every subsequential weak limit of  $(S_{\varepsilon, \sigma} W_\varepsilon^1, S_{\varepsilon, \sigma} W_\varepsilon^2)$  satisfies conditions (i)-(iii) in Theorem 2.5. Condition (i) follows from the convergence of  $S_{\varepsilon, \sigma} W_\varepsilon^1$  and  $S_{\varepsilon, \sigma} W_\varepsilon^2$ .

To verify condition (ii), it suffices to show that a pair of paths  $(X^1, X^2) \in (W_\varepsilon^1, W_\varepsilon^2)$ , whose starting points converge under diffusive scaling, must converge to a pair of sticky Brownian motions in the sense of Proposition 2.4. By waiting for the earlier path to reach the starting time of the later path, and by translation invariance, it suffices to consider paths starting at the same time 0. A stronger form of such a convergence is established in Theorem 3.2 below.

The verification of condition (iii) follows the same argument as in the proof of Theorem 2 in [SS19, Section 4]. Recall that the filtration  $(\mathcal{F}_u)_{u \in \mathbb{R}}$  is generated by paths in  $\mathcal{W}^1$  and  $\mathcal{W}^2$ , where we may further restrict to  $\mathcal{W}^1(\mathcal{D})$  and  $\mathcal{W}^2(\mathcal{D})$  for a deterministic countable dense set  $\mathcal{D} \subset \mathbb{R}^2$ . For any deterministic  $z = (x, t), z' = (x', t') \in \mathbb{R}^2$ , to prove that  $(X_z^i, X_{z'}^j)$  with  $X_z^i \in \mathcal{W}^i(z)$  and  $X_{z'}^j \in \mathcal{W}^j(z')$  is a Markov process w.r.t.  $(\mathcal{F}_u)_{u \in \mathbb{R}}$ , it suffices to show that for any  $s \geq \max\{t, t'\}$  and  $z_1, \dots, z_m \in \mathcal{D}$  with time coordinates less than  $s$ , conditioned on  $(X_{z_i}^i(s), X_{z'_j}^j(s))$  for some  $s \geq \max\{t, t'\}$ , the law of  $(X_{z_i}^i(u), X_{z'_j}^j(u))_{u \geq s}$  in the future does not depend on the past realisation  $(X_{z_i}^i(u), X_{z'_j}^j(u), (X_{z_k}^1(u), X_{z_k}^2(u))_{1 \leq k \leq m})_{u < s}$ . This property follows readily from approximation by paths in the discrete sticky webs, for which an analogous property holds. This verifies condition (iii).  $\blacksquare$

We now state a strong form of convergence of a pair of sticky random walks to sticky Brownian motions, which also includes the convergence of their intersection times. For techni-

cal convenience, the random walk paths are defined on  $[0, \infty)$  by piecewise constant extension to non-integer times and regarded as random variables taking values in the Skorohod space  $\mathcal{D}([0, \infty), \mathbb{R})$ .

**Theorem 3.2 (Pair convergence including intersection times)** *Let  $X^1$  and  $X^2$  be the unique paths in  $W_\varepsilon^1$  and  $W_\varepsilon^2$  starting at time zero at positions  $x_\varepsilon^1$  and  $x_\varepsilon^2$ , respectively, such that  $(\varepsilon x_\varepsilon^1, \varepsilon x_\varepsilon^2) \rightarrow (y^1, y^2) \in \mathbb{R}^2$  as  $\varepsilon \downarrow 0$ . Set*

$$Y_\varepsilon^i(t) := \varepsilon X^i(\lfloor t/\sigma^2\varepsilon^2 \rfloor) \quad (i = 1, 2) \quad \text{and} \quad Z_\varepsilon(t) := \sigma^2\varepsilon^2 \sum_{s=0}^{\lfloor t/\sigma^2\varepsilon^2 \rfloor} 1_{\{Y_\varepsilon^1(s)=Y_\varepsilon^2(s)\}}. \quad (3.2)$$

Then as  $\varepsilon \downarrow 0$ , the process  $(Y_\varepsilon^1, Y_\varepsilon^2, Z_\varepsilon)$  converges in distribution on the Skorohod space  $\mathcal{D}([0, \infty), \mathbb{R}^3)$  to a limiting process  $(Y^1, Y^2, Z)$ , where  $(Y^1, Y^2)$  is a pair of sticky Brownian motions with parameter 1 and starting point  $(y^1, y^2)$  characterised in Proposition 2.4, and

$$Z(t) = \int_0^t 1_{\{Y^1(s)=Y^2(s)\}} ds. \quad (3.3)$$

Thanks to Theorem 3.2, we can strengthen Theorem 3.1 to include the convergence of intersection times between paths in  $W_\varepsilon^1$  and  $W_\varepsilon^2$ .

**Corollary 3.3** *Let  $\mathcal{D}$  be a countable dense subset of  $\mathbb{R}^2$ . For each  $z = (x, t) \in \mathcal{D}$  and  $i = 1, 2$ , let  $X_{\varepsilon, z}^i$  be the unique path in  $W_\varepsilon^i$  starting from position  $\varepsilon \lfloor x\varepsilon^{-1} \rfloor$  at time  $\sigma^2\varepsilon^2 \lfloor t/\sigma^2\varepsilon^2 \rfloor$ . For  $z, z' \in \mathcal{D}$ , let  $Z_{\varepsilon; z, z'}$  be the rescaled intersection time between  $X_{\varepsilon, z}^1$  and  $X_{\varepsilon, z'}^2$  defined as in (3.2). Then we have*

$$(S_{\varepsilon, \sigma} W_\varepsilon^1, S_{\varepsilon, \sigma} W_\varepsilon^2, (Z_{\varepsilon; z, z'})_{z, z' \in \mathcal{D}}) \xrightarrow[\varepsilon \downarrow 0]{} (\mathcal{W}^1, \mathcal{W}^2, (Z_{z, z'})_{z, z' \in \mathcal{D}}), \quad (3.4)$$

where  $Z_{z, z'}(u) = \int_{\max\{t, t'\}}^u 1_{\{Y_z^1(s)=Y_{z'}^2(s)\}} ds$  is the intersection time between the unique path  $Y_z^1 \in \mathcal{W}^1(z)$  and  $Y_{z'}^2 \in \mathcal{W}^2(z')$ .

This corollary follows readily from Theorem 3.1 and 3.2 and the fact that for  $z, z' \in \mathcal{D}$ ,  $Z_{z, z'}$  is determined by  $Y_z^1 \in \mathcal{W}^1(z)$  and  $Y_{z'}^2 \in \mathcal{W}^2(z')$ .

To prove Theorem 3.2, we need a preliminary result for  $S(t) := X^1(t) - X^2(t)$ , which is a random walk on  $\mathbb{Z}$  with stickiness at 0 and transition kernel

$$P(x, y) = \begin{cases} (1 - \varepsilon)1_{\{y=0\}} + \varepsilon\bar{P}(x, y) & \text{if } x = 0, \\ \bar{P}(x, y) & \text{if } x \neq 0, \end{cases} \quad (3.5)$$

where

$$\bar{P}(x, x+y) := \sum_{z \in \mathbb{Z}} a(z+y)a(z) \quad (x, y \in \mathbb{Z}). \quad (3.6)$$

For  $t \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ , let  $\bar{P}^t$  denote the  $t$ -th power of the transition matrix  $\bar{P}$ , with  $\bar{P}^0(x, y) = 1_{\{x=y\}}$ , and set  $\bar{G}_t(x, y) := \sum_{s=0}^t \bar{P}^s(x, y)$ . By [S01, Prop. 28.8], the limit

$$\bar{A}(x) := \lim_{t \rightarrow \infty} [\bar{G}_t(0, 0) - \bar{G}_t(x, 0)] < \infty \quad (x \in \mathbb{Z}) \quad (3.7)$$

exists, and  $\bar{A}$  is called the *potential kernel* of the random walk with transition kernel  $\bar{P}$ .

**Proposition 3.4 (Compensator of potential kernel)** *Let  $S$  denote the difference random walk with transition kernel  $P$  as in (3.5). Then*

$$\bar{A}(S(t)) - \varepsilon \sum_{s=0}^{t-1} 1_{\{S(s)=0\}} \quad (3.8)$$

is a martingale. Moreover

$$\lim_{|x| \rightarrow \infty} \frac{\bar{A}(x)}{|x|} = \frac{1}{2\sigma^2}. \quad (3.9)$$

**Proof** Formula (3.9) follows from [S01, Prop. 29.2]. To prove (3.8), first denote  $\bar{A}_t(x) := \bar{G}_t(0,0) - \bar{G}_t(x,0)$ . Then we have

$$\begin{aligned} \bar{P}\bar{A}_t(x) &= \sum_{y \in \mathbb{Z}} \bar{P}(x,y) \bar{A}_t(y) = \sum_{y \in \mathbb{Z}} \bar{P}(x,y) \sum_{s=0}^t (\bar{P}^s(0,0) - \bar{P}^s(y,0)) \\ &= \sum_{s=0}^t (\bar{P}^s(0,0) - \bar{P}^{s+1}(x,0)) = \bar{A}_t(x) + \bar{P}^0(x,0) - \bar{P}^{t+1}(x,0). \end{aligned} \quad (3.10)$$

Here  $\lim_{t \rightarrow \infty} \bar{P}^{t+1}(x,0) = 0$  by the null-recurrence of the random walk with kernel  $\bar{P}$ , while  $\lim_{t \rightarrow \infty} \bar{P}\bar{A}_t(x) = \bar{P}\bar{A}(x)$  by [S01, Prop. 13.3]. Therefore letting  $t \rightarrow \infty$  in (3.10) gives

$$\bar{P}\bar{A}(x) - \bar{A}(x) = 1_{\{x=0\}} \quad (x \in \mathbb{Z}). \quad (3.11)$$

We claim that for the transition kernel  $P$  of the difference process  $S$ ,

$$P\bar{A}(x) - \bar{A}(x) = \varepsilon 1_{\{x=0\}}. \quad (3.12)$$

Indeed, when  $x \neq 0$ , by the definition (3.5), we have  $P(x,y) = \bar{P}(x,y)$  for all  $y \in \mathbb{Z}$ . Hence (3.11) is the same as (3.12). When  $x = 0$ , recalling (3.5) and using  $\bar{A}(0) = 0$ , we have

$$P\bar{A}(0) - \bar{A}(0) = (1 - \varepsilon) \sum_{y \in \mathbb{Z}} 1_{\{y=0\}} \bar{A}(y) + \varepsilon \bar{P}\bar{A}(0) = \varepsilon \bar{P}\bar{A}(0) = \varepsilon 1_{\{x=0\}}, \quad (3.13)$$

where in the last equality we used (3.11).

In general, if  $S(t)$  is a Markov chain with countable state space and transition kernel  $P$ , then for all real functions  $f$ ,

$$M(t) := f(S(t)) - \sum_{s=0}^{t-1} (Pf(S(s)) - f(S(s))) \quad (t \geq 0) \quad (3.14)$$

is a martingale as long as each term in this expression is in  $L^1$ . Applying this to the Markov chain  $S$  and the function  $\bar{A}$ , and using (3.12), we see that the process in (3.8) is a martingale. Here, to check integrability, we have used that  $\mathbb{E}[|S(t)|] < \infty$  and hence by (3.9), we also have  $\mathbb{E}[\bar{A}(S(t))] < \infty$ .  $\blacksquare$

To prove Theorem 3.2, we also need the following facts. Recall that for continuous local martingales  $M_1$  and  $M_2$ , there exists a unique process  $\langle M_1, M_2 \rangle$  of bounded variation such that

$$M_1(t)M_2(t) - \langle M_1, M_2 \rangle(t) \quad (3.15)$$

is a local martingale (see [KS91, Def. 1.5.18]). Moreover, by [KS91, Thm. 3.7.1] one can define a local time  $L(t,x)$  for any real square integrable continuous semimartingale  $M$  such that  $L(t,x)$  is a measurable function, continuous and nondecreasing in  $t$ , and for any measurable function  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,

$$\int_0^t f(M(s)) d\langle M, M \rangle(s) = \int_{\mathbb{R}} f(x) L(t,x) dx. \quad (3.16)$$

**Proof of Theorem 3.2** Since  $X^1$  and  $X^2$  are both random walks with the same jump distribution  $a(\cdot)$ , by Donsker's invariance principle,  $Y_\varepsilon^1$  and  $Y_\varepsilon^2$  converge to standard Brownian motions starting from  $y^1$  and  $y^2$ , respectively. This also implies the tightness of  $(Y_\varepsilon^1, Y_\varepsilon^2)$ . Consequently, to show that  $(Y_\varepsilon^1, Y_\varepsilon^2)$  converges to a pair of sticky Brownian motions, it suffices to show that for any subsequential weak limit  $(Y^1, Y^2)$ , the processes in (i) and (ii) of (2.7)

are martingales. Along the way, we will see that  $Z_\varepsilon$ , as a function of  $(Y_\varepsilon^1, Y_\varepsilon^2)$ , also converges weakly to the process  $Z$  defined from  $(Y^1, Y^2)$  as in (3.3).

Note that the transition kernel of  $(X^1(t), X^2(t))_{t \in \mathbb{N}}$  is given by

$$P^{(2)}(x_1, x_2; x_1 + z_1, x_2 + z_2) = \begin{cases} (1 - \varepsilon)1_{\{z_1=z_2\}}a(z_1) + \varepsilon a(z_1)a(z_2) & \text{if } x_1 = x_2, \\ a(z_1)a(z_2) & \text{if } x_1 \neq x_2. \end{cases} \quad (3.17)$$

Therefore by standard theory (see (3.14)), the discrete time process

$$X^1(t)X^2(t) - \sum_{k=0}^{t-1} \Gamma(X^1(k), X^2(k)) \quad (t \geq 0) \quad (3.18)$$

is a martingale, where

$$\begin{aligned} \Gamma(x_1, x_2) &= \sum_{z_1, z_2 \in \mathbb{Z}} P^{(2)}(x_1, x_2; x_1 + z_1, x_2 + z_2)(x_1 + z_1)(x_2 + z_2) - x_1 x_2 \\ &= (1 - \varepsilon)\sigma^2 1_{\{x_1=x_2\}}. \end{aligned} \quad (3.19)$$

In terms of  $Y_\varepsilon^1$  and  $Y_\varepsilon^2$ , whose trajectories are piecewise constant, this implies that

$$Y_\varepsilon^1(t)Y_\varepsilon^2(t) - (1 - \varepsilon) \int_0^{\lfloor t \rfloor_\varepsilon} 1_{\{Y_\varepsilon^1(s)=Y_\varepsilon^2(s)\}} ds \quad (t \geq 0) \quad (3.20)$$

is a martingale, where  $\lfloor t \rfloor_\varepsilon := \sigma^2 \varepsilon^2 \lfloor t / \sigma^2 \varepsilon^2 \rfloor$ .

On the other hand, since the transition kernel of the difference process  $X^1(t) - X^2(t)$  is given by (3.5), by Proposition 3.4,

$$\varepsilon \bar{A}(\varepsilon^{-1}[Y_\varepsilon^1(t) - Y_\varepsilon^2(t)]) - \frac{1}{\sigma^2} \int_0^{\lfloor t \rfloor_\varepsilon} 1_{\{Y_\varepsilon^1(s)=Y_\varepsilon^2(s)\}} ds \quad (3.21)$$

is also a martingale. It will be more convenient to work with

$$\tilde{Z}_\varepsilon(t) := \int_0^t 1_{\{Y_\varepsilon^1(s)=Y_\varepsilon^2(s)\}} ds, \quad (3.22)$$

for which we note that  $Z_\varepsilon(t) = \tilde{Z}_\varepsilon(\lfloor t \rfloor_\varepsilon)$  for all  $t \geq 0$ .

To prove the convergence of the discrete martingale (3.20) to a continuous one, we adopt an argument in [SSS14, Prop. A.6] (which is adapted from [HW09a]). First of all, recall that the laws of the processes  $(Y_\varepsilon^1, Y_\varepsilon^2)$  on  $\mathcal{D}([0, \infty), \mathbb{R}^2)$  are tight. Secondly, since the slope of  $\tilde{Z}_\varepsilon$  is between zero and one, by the Arzelà-Ascoli Theorem, tightness holds also for  $(\tilde{Z}_\varepsilon)_{\varepsilon > 0}$ . As a result, by going to a subsequence if necessary, we may assume that the joint processes  $(Y_\varepsilon^1, Y_\varepsilon^2, \tilde{Z}_\varepsilon)$  converge weakly in law as  $\varepsilon \downarrow 0$ , and by Skorohod's representation Theorem (see e.g. [B99, Thm. 6.7]), we can couple  $(Y_\varepsilon^1, Y_\varepsilon^2, \tilde{Z}_\varepsilon)$  along any sequence  $\varepsilon \downarrow 0$  such that the convergence is almost sure. Moreover, since for all  $t \geq 0$ ,  $(Y_\varepsilon^1(t), Y_\varepsilon^2(t), \tilde{Z}_\varepsilon(t))_{\varepsilon > 0}$  are uniformly integrable, their convergence is also in  $L^1$ . Let  $(Y^1, Y^2, Z_*)$  denote the limiting process. Then, taking the limit in (3.20) and using the Markov property of  $(Y_\varepsilon^1, Y_\varepsilon^2)$ , we see that

$$Y^1(t)Y^2(t) - Z^*(t) \quad (3.23)$$

is a martingale, and hence

$$\langle Y^1, Y^2 \rangle(t) = Z^*(t) = \lim_{\varepsilon \downarrow 0} \int_0^t 1_{\{Y_\varepsilon^1(s)=Y_\varepsilon^2(s)\}} ds \quad \forall t \geq 0 \quad \text{a.s.} \quad (3.24)$$



On one hand, since for a given  $t > 0$ , the function  $(w_1, w_2) \mapsto \int_0^t 1_{\{w_1(s)=w_2(s)\}} ds$  is upper semicontinuous with respect to the topology on  $\mathcal{D}([0, \infty), \mathbb{R}^2)$ , formula (3.24) implies that

$$\langle Y^1, Y^2 \rangle(t) \leq \int_0^t 1_{\{Y^1(s)=Y^2(s)\}} ds \quad (t \geq 0). \quad (3.25)$$

On the other hand, since  $\langle Y^1, Y^2 \rangle(t)$  is nondecreasing by (3.24), almost surely  $d\langle Y^1, Y^2 \rangle(t)$  defines a nonnegative measure on  $[0, \infty)$ , which implies that

$$\langle Y^1, Y^2 \rangle(t) \geq \int_0^t 1_{\{Y^1(s)=Y^2(s)\}} d\langle Y^1, Y^2 \rangle(s) \quad (t \geq 0). \quad (3.26)$$

To prove the opposite inequality to the one in (3.25), let  $L(t, x)$  be the local time of the semimartingale  $Y^1 - Y^2$ . Applying (3.16) to the function  $f = 1_{\{0\}}$ , we find that

$$\int_0^t 1_{\{Y^1(s)=Y^2(s)\}} d\langle Y^1 - Y^2, Y^1 - Y^2 \rangle(s) = \int_{\mathbb{R}} 1_{\{0\}}(x) L(t, x) dx = 0. \quad (3.27)$$

Since  $Y^1, Y^2$  are standard Brownian motions, we have

$$\begin{aligned} \langle Y^1 - Y^2, Y^1 - Y^2 \rangle(t) &= \langle Y^1, Y^1 \rangle(t) + \langle Y^2, Y^2 \rangle(t) - 2\langle Y^1, Y^2 \rangle(t) \\ &= 2t - 2\langle Y^1, Y^2 \rangle(t). \end{aligned} \quad (3.28)$$

Inserting this into (3.27) yields

$$\int_0^t 1_{\{Y^1(s)=Y^2(s)\}} d\langle Y^1, Y^2 \rangle(s) = \int_0^t 1_{\{Y^1(s)=Y^2(s)\}} ds. \quad (3.29)$$

Combining the above equality with (3.26) and (3.25) shows that

$$Z^*(t) = \langle Y^1, Y^2 \rangle(t) = \int_0^t 1_{\{Y^1(s)=Y^2(s)\}} ds = Z(t) \quad \forall t \geq 0 \quad \text{a.s.} \quad (3.30)$$

Therefore,  $Y^1(t)Y^2(t) - \int_0^t 1_{\{Y^1(s)=Y^2(s)\}} ds$  is a martingale. Moreover, it is easy to see that the almost sure convergence of  $\tilde{Z}_\varepsilon$  to  $Z$  implies the almost sure convergence of  $Z_\varepsilon$  to  $Z$ .

To check that the process in (ii) of (2.7) is a martingale, we take the limit  $\varepsilon \downarrow 0$  in (3.21). Since  $\bar{P}$  defined in (3.6) has second moment  $2\sigma^2$ , by Proposition 3.4, for any  $\delta > 0$ , there exists  $N > 0$  such that for all integers  $|x| > N$ ,

$$\Delta(x) := |2\sigma^2 \bar{A}(x) - |x|| < \delta |x|. \quad (3.31)$$

Recall that  $S = X^1 - X^2$  and denote  $C_N := \sup_{|x| \leq N} |2\sigma^2 \bar{A}(x) - |x||$ . We see that

$$\begin{aligned} \varepsilon \Delta(S(t\sigma^{-2}\varepsilon^{-2})) &= \varepsilon \Delta(S(t\sigma^{-2}\varepsilon^{-2})) 1_{\{S(t\sigma^{-2}\varepsilon^{-2}) > N\}} + \varepsilon \Delta(S(t\sigma^{-2}\varepsilon^{-2})) 1_{\{S(t\sigma^{-2}\varepsilon^{-2}) \leq N\}} \\ &\leq \varepsilon \delta |S(t\sigma^{-2}\varepsilon^{-2})| + \varepsilon C_N = \delta |Y_\varepsilon^1(t) - Y_\varepsilon^2(t)| + \varepsilon C_N. \end{aligned} \quad (3.32)$$

Sending  $\varepsilon \downarrow 0$  first and then  $\delta \downarrow 0$ , and using the almost sure convergence of  $Y_\varepsilon^1(t) - Y_\varepsilon^2(t)$  to  $Y^1(t) - Y^2(t)$ , we obtain that

$$\limsup_{\varepsilon \downarrow 0} \varepsilon \Delta(\varepsilon^{-1}[Y_\varepsilon^1(t) - Y_\varepsilon^2(t)]) = 0 \quad \forall t \geq 0 \quad \text{a.s.} \quad (3.33)$$

Therefore, almost surely, for all  $t \geq 0$ ,

$$\lim_{\varepsilon \downarrow 0} \varepsilon \bar{A}(\varepsilon^{-1}[Y_\varepsilon^1(t) - Y_\varepsilon^2(t)]) = \lim_{\varepsilon \downarrow 0} \frac{1}{2\sigma^2} |Y_\varepsilon^1(t) - Y_\varepsilon^2(t)| = \frac{1}{2\sigma^2} |Y^1(t) - Y^2(t)|. \quad (3.34)$$

Recalling (3.24) and (3.30), the almost sure limit (also  $L^1$  limit) of (3.21) is

$$\frac{1}{2\sigma^2}|Y^1(t) - Y^2(t)| - \frac{1}{\sigma^2} \int_0^t 1_{\{Y^1(s)=Y^2(s)\}} ds. \quad (3.35)$$

Together with the Markov property of  $(Y_\varepsilon^1, Y_\varepsilon^2)$ , it follows that

$$|Y^1(t) - Y^2(t)| - 2 \int_0^t 1_{\{Y^1(s)=Y^2(s)\}} ds \quad (3.36)$$

is a martingale.

Therefore, the process  $(Y_\varepsilon^1, Y_\varepsilon^2, Z_\varepsilon)$  converges in distribution on  $\mathcal{D}([0, \infty), \mathbb{R}^3)$  to  $(Y^1, Y^2, Z)$ , where  $(Y^1, Y^2)$  is a pair of sticky Brownian motions with parameter 1 and starting point  $(y^1, y^2)$  as characterized in Proposition 2.4, and  $Z$  is defined from  $(Y^1, Y^2)$  as in (3.3).  $\blacksquare$

### 3.2 Proof of the lower bound

We are now ready to complete the lower bound on any subsequential weak limit of  $S_{\varepsilon, \sigma} N_\varepsilon$ .

**Theorem 3.5 (Lower bound)** *For  $\varepsilon > 0$ , let  $N_\varepsilon$  be the discrete net with branching probability  $\varepsilon$  and jump distribution  $a(\cdot)$  satisfying (1.1). Then for any subsequential weak limit  $\mathcal{N}^*$  of  $S_{\varepsilon, \sigma} N_\varepsilon$  as  $\varepsilon \downarrow 0$ , there exists a standard Brownian net  $\mathcal{N}$  such that almost surely,*

$$\mathcal{N} \subseteq \mathcal{N}^*. \quad (3.37)$$

**Proof** Let  $(W_\varepsilon^1, W_\varepsilon^2)$  be the pair of discrete sticky webs defined before (1.3), which is contained in the discrete net  $N_\varepsilon$ . Thanks to Corollary 3.3, if  $\mathcal{N}^*$  is a subsequential limit of  $S_{\varepsilon, \sigma} N_\varepsilon$  as  $\varepsilon \downarrow 0$ , then by going to a subsequence if necessary and by Skorohod's representation theorem, we can find a coupling such that

$$(S_{\varepsilon, \sigma} W_\varepsilon^1, S_{\varepsilon, \sigma} W_\varepsilon^2, S_{\varepsilon, \sigma} N_\varepsilon, (Z_{\varepsilon; z, z'})_{z, z' \in \mathcal{D}}) \xrightarrow{\varepsilon \downarrow 0} (\mathcal{W}^1, \mathcal{W}^2, \mathcal{N}^*, (Z_{z, z'})_{z, z' \in \mathcal{D}}) \quad \text{a.s.}, \quad (3.38)$$

where  $\mathcal{W}^1 \cup \mathcal{W}^2 \subseteq \mathcal{N}^*$ ,  $\mathcal{D}$  is a deterministic countable subset of  $\mathbb{R}^2$ , and  $Z_{z, z'}$  is the intersection time process between  $Y_z^1 \in \mathcal{W}^1(z)$  and  $Y_{z'}^2 \in \mathcal{W}^2(z')$  while  $Z_{\varepsilon; z, z'}$  is the discrete counterpart. By Theorem 2.7, we can construct a standard Brownian net via hopping between paths in  $\mathcal{W}^1$  and  $\mathcal{W}^2$ , i.e.,

$$\mathcal{N} = \overline{\mathcal{H}_{\text{int}}(\mathcal{W}^1 \cup \mathcal{W}^2)}. \quad (3.39)$$

To show  $\mathcal{N} \subseteq \mathcal{N}^*$ , it suffices to show that

$$\mathcal{H}_{\text{int}}(\mathcal{W}^1 \cup \mathcal{W}^2) \subseteq \mathcal{N}^*, \quad (3.40)$$

since  $\mathcal{N}^*$  is a compact subset of the path space  $\Pi$ . Let  $\mathcal{D}$  be a countable dense subset of  $\mathbb{R}^2$ . For any path  $\pi \in \mathcal{W}^i$ , we can find a sequence of  $\pi_n \in \mathcal{W}^i(\mathcal{D})$  such that  $\pi_n \rightarrow \pi$  and the time of coalescence between  $\pi_n$  and  $\pi$  converges to the starting time of  $\pi$ , see [SS08, Lemma 3.4]. Therefore to prove (3.40), it suffices to show  $\mathcal{H}_{\text{int}}(\mathcal{W}^1(\mathcal{D}) \cup \mathcal{W}^2(\mathcal{D})) \subseteq \mathcal{N}^*$ . In particular, by induction, it suffices to show that for any  $Y_z^1 \in \mathcal{W}^1(z)$  and  $Y_{z'}^2 \in \mathcal{W}^2(z')$  with  $z = (x, t)$ ,  $z' = (x', t') \in \mathcal{D}$ , if  $u > \max\{t, t'\}$  is an intersection time of  $Y^1$  and  $Y^2$ , then the concatenated path with graph

$$Y := \{(Y_z^1(s), s) : s \in [t, u]\} \cup \{(Y_{z'}^2(s), s) : s \in [u, \infty)\} \in \mathcal{N}^*. \quad (3.41)$$

By (3.38), we can find  $Y_{\varepsilon, z}^1 \in S_{\varepsilon, \sigma} W_\varepsilon^1$  and  $Y_{\varepsilon, z'}^2 \in S_{\varepsilon, \sigma} W_\varepsilon^2$  such that

$$(Y_{\varepsilon, z}^1, Y_{\varepsilon, z'}^2, Z_{\varepsilon; z, z'}) \rightarrow (Y_z^1, Y_{z'}^2, Z_{z, z'}),$$

where  $Z_{\varepsilon; z, z'}$  and  $Z_{z, z'}$  are the intersection time processes for  $(Y_{\varepsilon, z}^1, Y_{\varepsilon, z'}^2)$  and  $(Y_z^1, Y_{z'}^2)$ , respectively. Note that  $Y_z^1(s) - Y_{z'}^2(s)$  is a sticky Brownian motion with stickiness at 0, which can be obtained from a Brownian motion via a random time change that changes its local time at the origin into real time, see e.g. [SS08, Section 2.1] and [Y16, Section 2.1]. This ensures that any intersection time  $u$  of  $Y_z^1$  and  $Y_{z'}^2$  lies in the support of the measure  $dZ_{z, z'}(s)$ . The convergence  $Z_{\varepsilon; z, z'} \rightarrow Z_{z, z'}$  then implies that we can find intersection times  $u_\varepsilon$  of  $Y_{\varepsilon, z}^1$  and  $Y_{\varepsilon, z'}^2$  such that  $u_\varepsilon \rightarrow u$ . Therefore we can concatenate  $Y_{\varepsilon, z}^1$  and  $Y_{\varepsilon, z'}^2$  at time  $u_\varepsilon$ , which gives a path in  $S_{\varepsilon, \sigma} N_\varepsilon$  that clearly converges to  $Y$ . This shows that  $Y$  is an element of  $\mathcal{N}^*$  and completes the proof.  $\blacksquare$

## 4 Auxiliary Bernoulli net

In this section, we introduce an auxiliary system of branching-coalescing random walks which we call the *Bernoulli net*  $\tilde{N}$ . In Subsection 4.1, we show that  $\tilde{N}$  can be coupled to the discrete net  $N_\varepsilon$  such that  $N_\varepsilon \subset \tilde{N}$  almost surely. In Subsection 4.2, we introduce the branching-coalescing point set  $\tilde{\xi}$  generated by  $\tilde{N}$ , and its dual process  $\tilde{\phi}$  which is a discrete time biased voter model. In Subsection 4.3, we show that  $\tilde{\xi}$  admits an invariant Bernoulli product law, which converges to the invariant law of the branching-coalescing point set  $\xi$  generated by the Brownian net  $\mathcal{N}$ . This fact is the key to the upper bound proof in Section 5. Lastly in Subsection 4.4, we give an upper bound on the density of  $\tilde{\xi}$  with maximal initial configuration, which is comparable to the density of coalescing random walks up to the diffusive time scale. The results of this section will be crucial in the upper bound proof in Section 5 and in the tightness proof in Section 6.

### 4.1 Bernoulli net and coupling with discrete net

Recall that the discrete net  $N_\varepsilon$  was defined before (1.3) by specifying the set of outgoing arrows from each  $(x, t) \in \mathbb{Z}^2$ . To define a more general branching-coalescing system, we can specify a probability law  $\mu$  on  $\mathcal{P}_{\text{fin},+}(\mathbb{Z})$ , the space of non-empty finite subsets of  $\mathbb{Z}$ . We then assign i.i.d.  $\mathcal{P}_{\text{fin},+}(\mathbb{Z})$ -valued random variables  $\mathcal{A}(x, t)$  with common law  $\mu$  to each  $(x, t) \in \mathbb{Z}^2$ . For every  $y \in \mathcal{A}(x, t)$ , we then draw an arrow from  $(x, t)$  to  $(x + y, t + 1)$ . The set of arrows then defines a branching-coalescing system similar to the discrete net  $N_\varepsilon$ , with  $\mu$  specifying the offspring distribution.

We now introduce the Bernoulli net  $\tilde{N}$ , which corresponds to a specific class of  $\mu$ . Given  $\psi : \mathbb{Z} \rightarrow [0, 1]$  with  $\sum_{x \in \mathbb{Z}} \psi(x) < \infty$ , let  $(V_x)_{x \in \mathbb{Z}}$  be independent Bernoulli random variables with parameters  $(\psi(x))_{x \in \mathbb{Z}}$  respectively, and let  $\text{Supp}(V) := \{x \in \mathbb{Z} : V(x) = 1\}$ . We then choose the offspring distribution  $\mu$  such that

$$1 - \mu(A) = \mathbb{P}(\text{Supp}(V) = A \mid \text{Supp}(V) \neq \emptyset) \quad (A \in \mathcal{P}_{\text{fin},+}(\mathbb{Z})). \quad (4.1)$$

We call the associated branching-coalescing system  $\tilde{N}_\psi$  the *Bernoulli net with kernel  $\psi$* . We say that  $\psi$  is *irreducible* if paths in  $\tilde{N}_\psi(0, 0)$  (the set of paths in  $\tilde{N}_\psi$  starting at  $(0, 0)$ ) can reach any point in  $\mathbb{Z}$  with positive probability; and we say  $\psi$  is *aperiodic* if the set of times at which 0 can be reached by paths in  $\tilde{N}_\psi(0, 0)$  with positive probability has greatest common divisor 1.

We now show how the discrete net  $N_\varepsilon$  with jump distribution  $a(\cdot)$  and branching probability  $\varepsilon$ , defined as in (1.3), can be coupled to a Bernoulli net  $\tilde{N}_\psi$  with kernel  $\psi$ , such that  $N_\varepsilon \subset \tilde{N}_\psi$  almost surely and the difference between them vanishes as  $\varepsilon \downarrow 0$ .

**Lemma 4.1 (Coupling of Bernoulli and discrete nets)** *Let  $N_\varepsilon$  be the discrete net with jump distribution  $a(\cdot)$  and branching probability  $\varepsilon > 0$ . Let  $r > 0$  be a solution of the equation*

$$\varepsilon = 1 - \frac{r}{e^r - 1}. \quad (4.2)$$

Define

$$\psi_\varepsilon(x) := 1 - e^{-ra(x)} \quad (x \in \mathbb{Z}). \quad (4.3)$$

Then the Bernoulli net  $\tilde{N}_{\psi_\varepsilon}$  with kernel  $\psi_\varepsilon$  can be coupled to  $N_\varepsilon$  such that  $N_\varepsilon \subset \tilde{N}_{\psi_\varepsilon}$  almost surely.

**Proof** It suffices to couple the configuration of arrows originating from  $(0,0)$  in  $N_\varepsilon$  and  $\tilde{N}_{\psi_\varepsilon}$ . The basic idea is to construct the arrows in  $\tilde{N}_{\psi_\varepsilon}$  through a Poisson point process. Let  $(M_x)_{x \in \mathbb{Z}}$  be independent Poisson random variables with mean  $ra(x)$ . We then draw  $M_x$  arrows from  $(0,0)$  to  $(1,x)$ , where multiple arrows coalesce with each other at  $(1,x)$ . Note that this is equivalent to assigning a single arrow from  $(0,0)$  to  $(1,x)$  with probability  $\psi_\varepsilon(x) = 1 - e^{-ra(x)}$ . To construct the configuration of arrows in  $\tilde{N}_{\psi_\varepsilon}$ , we are simply conditioning on the event  $M := \sum_x M_x \geq 1$ . With our choice of  $r$ , we have

$$\mathbb{P}(M = 1 | M \geq 1) = \frac{re^{-r}}{1 - e^{-r}} = \frac{r}{e^r - 1} = 1 - \varepsilon.$$

Restricted to the event  $\{M = 1\}$ , it is clear that the single arrow from  $(0,0)$  is drawn according to distribution  $a(\cdot)$ , which coincides with the definition of  $a^{(2)}$  in (1.3). On the event  $\{M \geq 2\}$ , we draw  $M$  arrows independently, each following the distribution  $a(\cdot)$ . Retaining only the first 2 arrows already reproduces the arrow configuration in the definition of  $a^{(2)}$  in (1.3). This establishes the desired coupling between  $N_\varepsilon$  and  $\tilde{N}_{\psi_\varepsilon}$ .  $\blacksquare$

**Remark 4.2** *We are interested in applying Lemma 4.1 as  $\varepsilon \downarrow 0$ . By Taylor expansion around  $r = 0$  in the r.h.s. of (4.2), it is easy to see that for  $\varepsilon > 0$  sufficiently small, there is a unique solution  $r(\varepsilon)$  to (4.2), and  $r(\varepsilon) = (2 + o(1))\varepsilon$  as  $\varepsilon \downarrow 0$ .*

## 4.2 Branching-coalescing point sets and their duals

Let  $\tilde{N} := \tilde{N}_\psi$  be a Bernoulli net with kernel  $\psi$ . We now introduce two associated Markov chains. One is the so-called branching-coalescing point set  $\tilde{\xi}_t$ , the other is its dual  $\tilde{\phi}_t$  which can be interpreted as a biased voter model (see e.g. [SSY19, SSY21]).

Let  $\mathcal{P}(\mathbb{Z})$  denote the collection of subsets of  $\mathbb{Z}$ . Given a Bernoulli net  $\tilde{N}$ , for each  $s, t \in \mathbb{Z}$  with  $s \leq t$ , we define two random maps  $\Xi_{s,t} : \mathcal{P}(\mathbb{Z}) \rightarrow \mathcal{P}(\mathbb{Z})$  and  $\Phi_{t,s} : \mathcal{P}(\mathbb{Z}) \rightarrow \mathcal{P}(\mathbb{Z})$  by

$$\begin{aligned} \Xi_{s,t}(A) &:= \{y \in \mathbb{Z} : \exists x \in A, \pi \in \tilde{N}(x,s) \text{ s.t. } \pi(t) = y\}, \\ \Phi_{t,s}(B) &:= \{x \in \mathbb{Z} : \exists y \in B, \pi \in \tilde{N}(x,s) \text{ s.t. } \pi(t) = y\}, \end{aligned} \quad (A, B \in \mathcal{P}(\mathbb{Z})), \quad (4.4)$$

where  $\tilde{N}(x,s)$  is the set of paths in  $\tilde{N}$  starting from  $(x,s)$ . In words,  $\Xi_{s,t}(A) \subset \mathbb{Z}$  is the set of positions generated at time  $t$  by paths in  $\tilde{N}$  that start from  $A$  at time  $s$ , while  $\Phi_{t,s}(B)$  is the set of positions at time  $s$  that can lead to some point in  $B$  by following paths in  $\tilde{N}$ .

Fix two times  $s, u \in \mathbb{Z}$ . We can then define two  $\mathcal{P}(\mathbb{Z})$ -valued Markov chains

$$\tilde{\xi}_t := \Xi_{s,s+t}(A) \quad \text{and} \quad \tilde{\phi}_t := \Phi_{u,u-t}(B) \quad (t \geq 0), \quad (4.5)$$

where  $\tilde{\xi}_t$  evolves forward in time and  $\tilde{\phi}_t$  evolves backward in time. We call  $\tilde{\xi}$  the branching-coalescing point set generated by  $\tilde{N}$ , and  $\tilde{\phi}$  its dual. More precisely, they are dual in the sense that

$$\mathbb{P}(\tilde{\xi}_t \cap B \neq \emptyset | \tilde{\xi}_0 = A) = \mathbb{P}(A \cap \tilde{\phi}_t \neq \emptyset | \tilde{\phi}_0 = B), \quad (A, B \in \mathcal{P}(\mathbb{Z}), t \geq 0), \quad (4.6)$$

which holds since

$$\{\Xi_{0,t}(A) \cap B \neq \emptyset\} = \{\exists x \in A, y \in B, \pi \in \tilde{N}(x,0) \text{ s.t. } \pi(t) = y\} = \{A \cap \Phi_{t,0}(B) \neq \emptyset\}.$$

It will be convenient to define the Markov chain  $\tilde{\xi}$  also at half-integer times in order to keep track of the arrows that are used in each time step. To this aim, for each  $s, t \in \mathbb{Z}$  with  $s \leq t$ , we define  $\Xi_{s,t+1/2} : \mathcal{P}(\mathbb{Z}) \rightarrow \mathcal{P}(\mathbb{Z}^2)$  by

$$\Xi_{s,t+1/2}(A) := \{(\pi(t), \pi(t+1)) \in \mathbb{Z}^2 : \pi \in \tilde{N}(s, x) \text{ for some } x \in A\}, \quad (4.7)$$

and we define  $\xi_{t+1/2}$  as in (4.5). Then  $(\xi_0, \xi_{1/2}, \xi_1, \dots)$  is a Markov chain taking values alternatively in  $\mathcal{P}(\mathbb{Z})$  and  $\mathcal{P}(\mathbb{Z}^2)$ , where in fact  $\xi_{t+1}$  is a deterministic function of  $\xi_{t+1/2}$ .

In the continuum setting, if  $\mathcal{N}$  is the standard Brownian net, then for each real  $s \leq t$ , we can similarly define a random map  $\Xi_{s,t}^{\mathcal{N}}$  acting on closed subsets of the real line by

$$\Xi_{s,t}(A) := \{\pi(t) : \pi \in \mathcal{N}(x, t) \text{ for some } x \in A\}. \quad (4.8)$$

Setting  $\xi_t^{\mathcal{N}} := \Xi_{s,s+t}^{\mathcal{N}}(A)$  ( $t \geq 0$ ) then defines a Markov process taking values in the space of closed subsets of the real line. This process is known as the *branching-coalescing point set* of the Brownian net  $\mathcal{N}$  [SS08, Thm 1.11].

### 4.3 Product invariant law and the backbone

For the standard Brownian net  $\mathcal{N}$ , the branching-coalescing point set  $\xi_t^{\mathcal{N}}$  is known to have a unique invariant law given by the law of the Poisson point process on  $\mathbb{R}$  with intensity 2, see [SS08, Prop. 1.15]. Furthermore, the stationary process can be constructed by considering  $\mathcal{N}(*, -\infty)$ , the set of paths in  $\mathcal{N}$  starting at  $t = -\infty$ , which is called the *backbone* of  $\mathcal{N}$ . Similar results hold for the discrete net  $N_\varepsilon$  with jump distribution  $a(\cdot)$  supported on  $\{\pm 1\}$ , see [SS08, Prop. 1.14]. In this subsection, we show that analogous results also hold for the Bernoulli nets  $\tilde{N}$ , and we identify explicitly the product invariant law of the branching-coalescing point set  $\tilde{\xi}$  generated by  $\tilde{N}$ .

In what follows, a *Bernoulli subset of  $\mathbb{Z}$  with intensity  $\alpha(\cdot)$*  refers to a random subset of  $\Lambda \subset \mathbb{Z}$ , where independently each  $x \in \mathbb{Z}$  belongs to  $\Lambda$  with probability  $\alpha(x)$ . A Bernoulli subset of  $\mathbb{Z}^2$  with intensity  $\alpha(\cdot)$  is defined similarly. We will write  $\mathbb{N}_0/2 := \{x/2 : x \in \mathbb{N}_0\}$  where  $\mathbb{N}_0 := \{0\} \cup \mathbb{N}$ .

**Lemma 4.3 (Product invariant law)** *Let  $\tilde{N}$  be a Bernoulli net with kernel  $\psi : \mathbb{Z} \rightarrow [0, 1)$ , which is irreducible and aperiodic. Let  $\Lambda$  be a Bernoulli subset of  $\mathbb{Z}$  with constant intensity*

$$\rho := 1 - \prod_{x \in \mathbb{Z}} (1 - \psi(x)). \quad (4.9)$$

*Let  $(\tilde{\xi}_t)_{t \in \mathbb{N}_0/2}$  be the branching-coalescing point set defined as in (4.5) and (4.7) from  $\tilde{N}$  with initial condition  $\tilde{\xi}_0 = \Lambda$ . Then for each  $t \in \mathbb{N}$ ,*

- (i)  $\tilde{\xi}_t$  is a Bernoulli subset of  $\mathbb{Z}$  with intensity  $\rho$ ;
- (ii)  $\tilde{\xi}_{t+1/2}$  is a Bernoulli subset of  $\mathbb{Z}^2$  with intensity  $\alpha(x, y) = \psi(y - x)$ .

**Proof** Let  $\xi_{1/2}$  be a Bernoulli subset of  $\mathbb{Z}^2$  with intensity  $\alpha(x, y) = \psi(y - x)$ . Define random sets  $\xi_0, \xi_1 \subset \mathbb{Z}$  by

$$\begin{aligned} \xi_0 &:= \{x \in \mathbb{Z} : (x, y) \in \xi_{1/2} \text{ for some } y \in \mathbb{Z}\}, \\ \xi_1 &:= \{y \in \mathbb{Z} : (x, y) \in \xi_{1/2} \text{ for some } x \in \mathbb{Z}\}. \end{aligned} \quad (4.10)$$

Then it is easy to see that both  $\xi_0$  and  $\xi_1$  are Bernoulli subset of  $\mathbb{Z}$  with constant intensity  $\rho$ . Moreover, for  $x \in \xi_0$ , if we define

$$\mathcal{A}(x, 0) := \{y \in \mathbb{Z} : (x, x + y) \in \xi_{1/2}\}, \quad (4.11)$$

then we note that conditioned on  $\xi_0$ , the random variables  $(\mathcal{A}(x, 0))_{x \in \xi_0}$  are i.i.d. with common law  $\mu$  defined from  $\psi$  as in (4.1). In view of this, the conditional law of  $\xi_{1/2}$  given  $\xi_0$  is the same as that for the branching-coalescing point set  $\tilde{\xi}$  defined from the Bernoulli net  $\tilde{N}$  with kernel  $\psi$ , and also  $\xi_1$  is a function of  $\xi_{1/2}$  in exactly the same way as  $\tilde{\xi}$ . The statements (i) and (ii) then follow immediately.  $\blacksquare$

Next we show that the backbone of the Bernoulli net  $\tilde{N}$  with kernel  $\psi$ , defined by

$$\tilde{\xi}_t^* := \{\pi(t) : \pi \in \tilde{N}(*, -\infty)\}, \quad (4.12)$$

is a stationary process with the law of  $\tilde{\xi}_t^*$  for each  $t \in \mathbb{R}$  given by the product invariant law identified in Lemma 4.3. In [SS08], such a result has been proved for the Brownian net  $\mathcal{N}$  as well as for the nearest neighbour discrete net. The proof for the latter, see [SS08, Prop. 1.14 (i)], makes essential use of the nearest neighbour property. Here we argue differently by identifying a suitable martingale for the process  $\tilde{\phi}$  dual to the branching-coalescing point set  $\tilde{\xi}$ . We start with a preliminary lemma for  $\tilde{\phi}$ .

**Lemma 4.4 (Positive fluctuations)** *Let  $\tilde{N}$  be a Bernoulli net with kernel  $\psi : \mathbb{Z} \rightarrow [0, 1)$ , which is irreducible and aperiodic. Let  $\tilde{\phi} = (\tilde{\phi}_t)_{t \geq 0}$  be the dual Markov chain defined in (4.5) with initial state  $\tilde{\phi}_0 \subset \mathbb{Z}$ . Then for all initial states  $\tilde{\phi}_0 \subset \mathbb{Z}$  with  $0 < |\tilde{\phi}_0| < \infty$ , we have*

$$\mathbb{P}(|\tilde{\phi}_1| \neq |\tilde{\phi}_0|) \geq \delta(\psi) := \frac{1}{\rho} \min \left\{ \psi(y_1), \psi(y_2) \prod_{w \leq y_1} (1 - \psi(w)) \right\} > 0, \quad (4.13)$$

where  $\rho$  is defined as in (4.9), and  $y_1 < y_2 \in \mathbb{Z}$  are any two points with  $\psi(y_1), \psi(y_2) \in (0, 1)$ .

In particular, if  $\tilde{N}_{\psi_\varepsilon}$  is the Bernoulli net coupled to the discrete net  $N_\varepsilon$  with jump distribution  $a(\cdot)$  and branching probability  $\varepsilon$  as in Lemma 4.1, then  $\liminf_{\varepsilon \downarrow 0} \delta(\psi_\varepsilon) > 0$ .

**Proof** By assumption, there exist  $y_1 < y_2 \in \mathbb{Z}$  with  $\psi(y_1), \psi(y_2) \in (0, 1)$ . Denote  $z := \max \tilde{\phi}_0$ , and let  $x = z - y_1$ . Then by the definition of  $\tilde{\phi}$  in (4.4) and (4.5) and the definition of  $\tilde{N}$  with kernel  $\psi$  in (4.1), we have the following bounds:

$$\mathbb{P}(x \in \tilde{\phi}_1) \geq \frac{\psi(y_1)}{\rho} \quad \text{and} \quad \mathbb{P}(x \notin \tilde{\phi}_1) \geq \frac{1}{\rho} \cdot \psi(y_2) \prod_{w \leq y_1} (1 - \psi(w)), \quad (4.14)$$

where the second bound is based on the event that in the random map  $\Phi_{1,0}(\tilde{\phi}_0)$  defined as in (4.4), there is no arrow in  $\tilde{N}$  from  $x$  to any  $z' \leq z$ , and there is an arrow from  $x$  to  $x + y_2$ . Note that conditioned on  $\tilde{\phi}_1 \setminus \{x\}$ , either  $\{x \in \tilde{\phi}_1\} \subset \{|\tilde{\phi}_1| \neq |\tilde{\phi}_0|\}$  or  $\{x \notin \tilde{\phi}_1\} \subset \{|\tilde{\phi}_1| \neq |\tilde{\phi}_0|\}$ . Furthermore, the event  $\{x \in \tilde{\phi}_1\}$  is independent of  $\tilde{\phi}_1 \setminus \{x\}$ . The claim (4.13) then follows readily from the bounds in (4.14).

When  $\psi = \psi_\varepsilon = 1 - e^{-ra(x)}$  as in (4.3), where  $r = r(\varepsilon) = (2 + o(1))\varepsilon$  as  $\varepsilon \downarrow 0$  by Remark 4.2, we have

$$\rho = 1 - e^{-r} = (2 + o(1))\varepsilon. \quad (4.15)$$

If we choose  $y_1 < y_2$  with  $a(y_1), a(y_2) > 0$ , then we have

$$\lim_{\varepsilon \downarrow 0} \psi_\varepsilon(y_1)/\rho = a(y_1) \quad \text{and} \quad \lim_{\varepsilon \downarrow 0} \psi_\varepsilon(y_2) \prod_{w \leq y_1} (1 - \psi_\varepsilon(y_1))/\rho = a(y_2).$$

Therefore  $\liminf_{\varepsilon \downarrow 0} \delta(\psi_\varepsilon) \geq \min\{a(y_1), a(y_2)\} > 0$ .  $\blacksquare$

**Proposition 4.5 (The backbone and the product invariant law)** *Let  $\tilde{N}$  be a Bernoulli net with kernel  $\psi$ , which is irreducible and aperiodic. Let  $(\tilde{\xi}_t^*)_{t \in \mathbb{Z}}$  be the backbone of  $\tilde{N}$  defined in (4.12). Then for each  $t \in \mathbb{Z}$ ,  $\tilde{\xi}_t^*$  is a Bernoulli subset of  $\mathbb{Z}$  with constant intensity  $\rho$ , defined as in Lemma 4.3.*

**Proof** Let  $\tilde{\phi} = (\tilde{\phi}_t)_{t \geq 0}$  be the Markov chain defined from  $\tilde{N}$  as in (4.5), starting from a deterministic initial state  $\tilde{\phi}_0$  with  $|\tilde{\phi}_0| < \infty$ . Let  $\rho$  be defined from  $\psi$  as in (4.9). We claim that the process

$$M_t := (1 - \rho)^{|\tilde{\phi}_t|} \quad (t \in \mathbb{N}) \quad (4.16)$$

is a martingale. Indeed, the duality between the branching-coalescing point set  $\tilde{\xi}$  and  $\tilde{\phi}$  in (4.6) gives

$$\mathbb{E}[M_1] = \mathbb{E}[(1 - \rho)^{|\tilde{\phi}_1|}] = \mathbb{P}(\tilde{\xi}_0 \cap \tilde{\phi}_1 = \emptyset) = \mathbb{P}(\tilde{\xi}_1 \cap \tilde{\phi}_0 = \emptyset) = (1 - \rho)^{|\tilde{\phi}_0|} = M_0, \quad (4.17)$$

where we have chosen  $\tilde{\xi}_0$  according to the product invariant law identified in Lemma 4.3. Since  $\tilde{\phi}$  is a Markov chain, it follows that  $M = (M_t)_{t \geq 0}$  is a martingale. Since  $|M_t| \leq 1$ , it must converge a.s. to a limit in  $[0, 1]$ , and hence  $|\tilde{\phi}_t|$  converges a.s. to a limit in  $\mathbb{N} \cup \{0, \infty\}$ . By Lemma 4.4, a finite nonzero limit can be ruled out. Therefore almost surely,

$$(1 - \rho)^{|\tilde{\phi}_t|} \xrightarrow[t \rightarrow \infty]{} 1_{\{\tilde{\phi}_u = \emptyset \text{ for some } u \geq 0\}}. \quad (4.18)$$

Using the martingale property, it follows that for any finite  $B \subset \mathbb{Z}$ ,

$$M_0 = (1 - \rho)^{|B|} = \mathbb{E}[M_\infty] = \mathbb{P}(\tilde{\phi}_u = \emptyset \text{ for some } u \geq 0 \mid \tilde{\phi}_0 = B). \quad (4.19)$$

Recalling the almost sure construction of the branching-coalescing point set  $\tilde{\xi}$  and the dual process  $\tilde{\phi}$  from the Bernoulli net  $\tilde{N}$  in Section 4.2, we note that almost surely,

$$\{\pi(t) \in B \text{ for some } \pi \in \tilde{N}(*, -\infty)\} = \{\Phi_{t,s}(B) \neq \emptyset \forall s \leq t\}. \quad (4.20)$$

It follows that the backbone  $\tilde{\xi}^*$  satisfies

$$\mathbb{P}(\tilde{\xi}_t^* \cap B = \emptyset) = \mathbb{P}(\Phi_{t,s}(B) = \emptyset \text{ for some } s \leq t) = \mathbb{P}(|\tilde{\phi}_t| \rightarrow 0 \mid \tilde{\phi}_0 = B) = (1 - \rho)^{|B|}. \quad (4.21)$$

Since this holds for each finite  $B \subset \mathbb{Z}$ , the claim follows.  $\blacksquare$

#### 4.4 Density estimates

For the discrete net  $N_\varepsilon$  with jump distribution  $a(\cdot)$  and branching probability  $\varepsilon$ , let  $\tilde{N}_\varepsilon := \tilde{N}_{\psi_\varepsilon}$  be the Bernoulli net with kernel  $\psi_\varepsilon$  as in Lemma 4.1, coupled to  $N_\varepsilon$ . The goal of this subsection is to bound the density of the branching-coalescing point set  $(\tilde{\xi}_t^\varepsilon)_{t \in \mathbb{N}_0}$  defined from  $\tilde{N}_\varepsilon$  as in (4.5), starting from the maximal configuration  $\tilde{\xi}_0^\varepsilon = \mathbb{Z}$ . We will show that the density of  $\tilde{\xi}_t^\varepsilon$  decays at the rate of  $1/\sqrt{t}$  up to time of order  $\varepsilon^{-2}$ , see Theorem 4.6, which is comparable to that of coalescing random walks. By the coupling between  $N_\varepsilon$  and  $\tilde{N}_\varepsilon$ , this gives an upper bound on the density of the branching-coalescing point set  $\xi_t^\varepsilon$  generated by the discrete net  $N_\varepsilon$  with  $\xi_0 = \mathbb{Z}$ . In particular, this implies that diffusive scaling limits of  $\xi^\varepsilon$  are locally finite subsets of  $\mathbb{R}$ , which will play an important role in the proof of the upper bound in Section 5.

To simplify notation, from now on, we will omit  $\varepsilon$  from  $\tilde{\xi}^\varepsilon$  and  $\tilde{\phi}^\varepsilon$ .

**Theorem 4.6 (Density decay)** *Let  $\tilde{\xi}$  be the branching-coalescing point set defined from the Bernoulli net  $\tilde{N}_\varepsilon$  with  $\tilde{\xi}_0 = \mathbb{Z}$ , and let  $\tilde{\phi}$  be the dual of  $\tilde{\xi}$  with  $\tilde{\phi}_0 = \{0\}$ . Then there exists  $C > 0$  such that for all  $\varepsilon > 0$  sufficiently small and all  $T \leq \varepsilon^{-2}$ , we have*

$$\mathbb{P}(0 \in \tilde{\xi}_T) = \mathbb{P}(|\tilde{\phi}_T| \neq 0) \leq \frac{C}{\sqrt{T}}. \quad (4.22)$$

*In particular, the same bound holds for the branching-coalescing point set  $\xi$  defined from the discrete net  $N_\varepsilon$  with  $\xi_0 = \mathbb{Z}$ .*

We now sketch the key ideas behind the proof of Theorem 4.6. The first observation is that we can relate the process  $|\tilde{\phi}_t|$  to a martingale. Instead of working with the martingale  $M_t$  defined in (4.16), it is more convenient to work with

$$\widetilde{M}_t := \frac{1 - M_t}{\rho} = \frac{1 - (1 - \rho)^{|\tilde{\phi}_t|}}{\rho}. \quad (4.23)$$

Note that  $|\tilde{\phi}_t| = 0$  if and only if  $\widetilde{M}_t = 0$ , and the time  $\tilde{\phi}$  hits the empty set is exactly the time  $\widetilde{M}$  hits 0. For  $\varepsilon$  small, by (4.15), we have  $\rho = 1 - e^{-r} = (2 + o(1))\varepsilon$ . From (4.23), we see that  $\widetilde{M}_t$  is close to  $|\tilde{\phi}_t|$  as long as  $|\tilde{\phi}_t| \ll 1/\varepsilon$ , and the two remain comparable as long as  $|\tilde{\phi}_t| \leq c/\varepsilon$  for some fixed constant  $c > 0$ . We will show that for  $T = O(\varepsilon^{-2})$ , the probability of the event  $\{\max_{1 \leq t \leq T} |\tilde{\phi}_t| \geq c/\varepsilon\}$  is of order  $O(1/\sqrt{T})$  and hence constitutes part of the bound in (4.22). On the complementary event,  $\widetilde{M}_t$  and  $|\tilde{\phi}_t|$  are comparable up to time  $T$ . Lemma 4.4 gives a lower bound on the fluctuations of  $|\tilde{\phi}_t|$ , which translates into a lower bound on the fluctuations of  $\widetilde{M}_t$ . This in turn implies that the probability that  $\widetilde{M}$  does not hit 0 before time  $T$  is comparable to the probability of the same event for a Brownian motion, which is also of order  $O(1/\sqrt{T})$ . To make this last comparison rigorous, we need to Skorohod embed  $\widetilde{M}_t$  in a Brownian motion (see e.g. [K97, Thm 12.16]), which requires an additional estimate due to the randomness in the embedding.

**Theorem 4.7 (Skorohod embedding)** *Let  $(\widetilde{M}_t)_{t \in \mathbb{N}_0}$  be a martingale with  $\widetilde{M}_0 = 1$  and induced filtration  $(\mathcal{G}_t)_{t \in \mathbb{N}_0}$ . Then there exist a Brownian motion  $B$  with  $B_0 = 1$  and stopping times  $0 = \tau_0 \leq \tau_1 \leq \dots$ , such that  $(\widetilde{M}_t)_{t \in \mathbb{N}_0} \stackrel{d}{=} (B_{\tau_t})_{t \in \mathbb{N}_0}$  and*

$$\mathbb{E}[(\widetilde{M}_t - \widetilde{M}_{t-1})^2 | \mathcal{G}_{t-1}] = \mathbb{E}[\tau_t - \tau_{t-1} | \mathcal{F}_{t-1}], \quad (4.24)$$

where  $(\mathcal{F}_t)_{t \in \mathbb{N}_0}$  is the filtration induced by the process  $(\tau_t, B_{\tau_t})_{t \in \mathbb{N}_0}$ . Moreover, conditioned on  $\mathcal{F}_{t-1}$ , there exists a random pair  $\alpha_t \leq 0 \leq \beta_t$  independent of  $(B_s - B_{\tau_{t-1}})_{s \geq \tau_{t-1}}$  such that

$$\tau_t = \inf \{s \geq \tau_{t-1} : B_s - B_{\tau_{t-1}} \in \{\alpha_t, \beta_t\}\}, \quad (4.25)$$

**Proof of Theorem 4.6.** The equality in (4.22) follows from the duality relation (4.6). Since  $\tilde{\phi}_t = 0$  if and only if  $\widetilde{M}_t = 0$ , to prove the bound in (4.22), it suffices to bound  $\mathbb{P}(\widetilde{M}_T \neq 0)$ , for which we need Theorem 4.7.

The idea behind the construction of the stopping times  $\tau_t$  in (4.25) is that conditioned on  $\widetilde{M}_{t-1} = B_{\tau_{t-1}}$ , the distribution of the martingale increment  $\widetilde{M}_t - \widetilde{M}_{t-1}$  can be decomposed into a mixture of two-point distributions, where the mixture is given by the law of  $(\alpha_t, \beta_t)$ . An  $\{\alpha_t, \beta_t\}$ -valued random variable with mean zero can then be embedded into the Brownian motion  $(B_{\tau_{t-1}+s} - B_{\tau_{t-1}})_{s \geq 0}$  by letting  $\tau_t$  be the first time that  $B_t$  hits  $\{\alpha_t, \beta_t\}$ . Furthermore, since the martingale  $(\widetilde{M}_t)_{t \in \mathbb{N}_0}$  is non-negative and will stay at 0 once it hits 0, we can conclude that  $\widetilde{M}$  first hits 0 at time  $t$  if and only if the Brownian motion  $B$  (with  $B_0 = 1$ ) hits 0 at time  $\tau_t$ . Therefore

$$\mathbb{P}(\widetilde{M}_T \neq 0) = \mathbb{P}(B_s > 0 \forall 0 < s < \tau_T). \quad (4.26)$$

To bound the r.h.s. of (4.26), we choose a constant  $C_1 > 0$  and then define a stopping time  $\sigma$  (w.r.t. the filtration of  $\tilde{\phi}$  and  $\widetilde{M}$ ) by

$$\sigma := \inf \{t \in \mathbb{Z} : |\tilde{\phi}_t| \geq C_1/\rho\}, \quad (4.27)$$

where

$$\rho = 1 - \prod_{x \in \mathbb{Z}} (1 - \psi_\varepsilon(x)) = 1 - e^{-r} = (2 + o(1))\varepsilon \quad (4.28)$$



by (4.3) and Remark 4.2. Consider three events  $E_1, E_2, E_3$ , which are respectively the intersection of the event in the r.h.s. of (4.26) with each of the following three events

$$\{\tau_T \geq c_1 T\}, \{\sigma \leq T\}, \text{ and } \{\tau_T < c_1 T, \sigma > T\}, \quad (4.29)$$

where  $c_1 > 0$  is a small constant to be determined in Lemma 4.9 below. To prove (4.22), it then suffices to show that  $\mathbb{P}(E_i) \leq C/\sqrt{T}$  for each  $i$ .

For the first event, we have

$$\mathbb{P}(E_1) \leq \mathbb{P}(B_s > 0 \forall 0 < s < c_1 T) \leq \frac{1}{\sqrt{c_1 T}}, \quad (4.30)$$

while the desired bounds on  $\mathbb{P}(E_2)$  and  $\mathbb{P}(E_3)$  will follow from Lemmas 4.8 and 4.9 below.  $\blacksquare$

**Lemma 4.8** *Let  $\tilde{N}_\varepsilon$  and  $(\tilde{\phi}_t)_{t \in \mathbb{N}_0}$  be as in Theorem 4.6, and let  $\sigma$  be as in (4.27). Then there exists  $C > 0$  such that for all  $\varepsilon > 0$  sufficiently small and all  $T \leq \varepsilon^{-2}$ , we have*

$$\mathbb{P}(\sigma \leq T) \leq C\rho \leq \frac{3C}{\sqrt{T}}. \quad (4.31)$$

**Proof** By the definition of  $\sigma$  in (4.27) and  $\tilde{M}$  in (4.23), we have

$$\tilde{M}_\sigma \geq \rho^{-1}[1 - (1 - \rho)^{C_1/\rho}].$$

Since  $\rho = 2\varepsilon + o(\varepsilon)$  by (4.28), we can find  $C > 0$  such that  $\tilde{M}_\sigma \geq 1/C\rho$  for all  $\varepsilon > 0$  sufficiently small. Since  $\tilde{M}$  is a non-negative martingale, we then have

$$1 = \tilde{M}_0 = \mathbb{E}[\tilde{M}_{\sigma \wedge T}] \geq \frac{1}{C\rho} \mathbb{P}(\sigma \leq T), \quad (4.32)$$

which implies (4.31).  $\blacksquare$

**Lemma 4.9** *Let  $\tilde{N}_\varepsilon$  and  $(\tilde{\phi}_t)_{t \in \mathbb{N}_0}$  be as in Theorem 4.6, and let  $\tau_t$  and  $\sigma$  be defined as in (4.25) and (4.27), respectively. Then there exist  $c_1, C > 0$  such that for all  $\varepsilon > 0$  sufficiently small and all  $T \leq \varepsilon^{-2}$ , we have*

$$\mathbb{P}(\tau_T < c_1 T, \sigma > T) < \frac{C}{\sqrt{T}}. \quad (4.33)$$

**Proof** Note that the event  $\{\sigma > T\}$  implies that  $|\tilde{\phi}_t| \leq C_1/\rho$  for all  $t \leq T$ . Therefore by the definition of  $\tilde{M}$  in (4.23), the increments of the martingale  $\tilde{M}$  satisfy

$$|\tilde{M}_t - \tilde{M}_{t-1}| \geq C_2 ||\tilde{\phi}_t| - |\tilde{\phi}_{t-1}|| \quad (1 \leq t \leq T)$$

for some constant  $C_2 > 0$ .

By Lemma 4.4, with probability at least  $\delta > 0$  for some  $\delta$  that is uniform in  $\varepsilon > 0$  sufficiently small and uniform in  $\tilde{\phi}_{t-1}$ , we have  $||\tilde{\phi}_t| - |\tilde{\phi}_{t-1}|| \geq 1$ , and hence  $|\tilde{M}_t - \tilde{M}_{t-1}| \geq C_2$ . Recall the construction of the stopping times  $\tau_t$  in the Skorohod embedding in (4.25), where conditioned on the  $\sigma$ -algebra  $\mathcal{G}_{t-1}$  generated by  $(\tilde{M}_s)_{0 \leq s \leq t-1}$ , the law of  $\tilde{M}_t - \tilde{M}_{t-1}$  is given by a mixture of the exit distribution of a Brownian motion  $B'$  with  $B'_0 = 0$  from  $(\alpha_t, \beta_t)$ , for some random vector  $(\alpha_t, \beta_t)$ . We note that the law of  $(\alpha_t, \beta_t)$  must satisfy

$$\mathbb{P}(\min\{|\alpha_t|, \beta_t\} \geq \alpha C_2/2) \geq \alpha/2, \quad (4.34)$$

since on the event  $\{\min\{|\alpha_t|, \beta_t\} \leq \alpha C_2/2\}$ , we have  $\mathbb{P}(|B'_\tau| \geq C_2 | (\alpha_t, \beta_t)) \leq \alpha/2$ , where  $\tau$  is the exit time of  $B'$ , with  $B'_0 = 0$ , from  $[\alpha_t, \beta_t]$ . Therefore (4.34) must hold in order to have  $\mathbb{P}(|\widetilde{M}_t - \widetilde{M}_{t-1}| \geq C_2 | \mathcal{G}_{t-1}) \geq \delta$ .

It is then easily seen that  $(\tau_t - \tau_{t-1})_{t \in \mathbb{N}}$  is stochastically dominated from below by a sequence of i.i.d. non-negative random variables  $(\gamma_t)_{t \in \mathbb{N}}$  with positive mean, where the law of  $\gamma_t$  depends only on  $\alpha$  and  $C_2$ . In particular, we can couple  $(\tau_t)_{t \in \mathbb{N}_0}$  and  $(\gamma_t)_{t \in \mathbb{N}}$  such that almost surely,  $\tau_t \geq \sum_{s=1}^t \gamma_s$  for all  $t \in \mathbb{N}$ . Choose any  $0 < c_1 < \mathbb{E}[\gamma_1]$ . We then have

$$\mathbb{P}(\tau_T \leq c_1 T) \leq \mathbb{P}\left(\sum_{t=1}^T \gamma_t \leq c_1 T\right) \leq C' e^{-T/C'} \leq \frac{C}{\sqrt{T}}$$

uniformly in  $T \in \mathbb{N}$  by elementary large deviation bounds. Furthermore, this bound is also uniform in  $\varepsilon > 0$  small, thanks to Lemma 4.4. This concludes the proof.  $\blacksquare$

## 5 The upper bound

Let  $N_\varepsilon$  be the discrete net as in Theorem 1.1, let  $\widetilde{N}_\varepsilon$  be the Bernoulli net coupled to  $N_\varepsilon$  as in Lemma 4.1, and let  $\mathcal{N}$  be the standard Brownian net. In Section 3, it was shown that every subsequential weak limit of the rescaled discrete net  $S_{\varepsilon, \sigma} N_\varepsilon$  contains a copy of the Brownian net  $\mathcal{N}$ . In this section, we show the matching upper bound that every subsequential weak limit of  $S_{\varepsilon, \sigma} N_\varepsilon$  contains no additional paths besides  $\mathcal{N}$ . We will start with a sufficient criterion for this last statement in Section 5.1, which is based on a family of counting random variables. In Section 5.2, we will show that the subset of  $N_\varepsilon$  started at time 0 from a random subset of  $\mathbb{Z}$ , distributed according to the product invariant law of the branching-coalescing point set generated by the Bernoulli net  $\widetilde{N}_\varepsilon$  (see Lemma 4.3), converges under the scaling map  $S_{\varepsilon, \sigma}$  to the subset of  $\mathcal{N}$  started at time 0 according to the stationary law of the branching-coalescing point set generated by  $\mathcal{N}$ . This is then used in Section 5.3 to deduce the convergence of paths in  $S_{\varepsilon, \sigma} N_\varepsilon$  starting from finitely many space-time points. Finally, the desired upper bound on subsequential weak limits of  $S_{\varepsilon, \sigma} N_\varepsilon$  is established in Section 5.4 by verifying the criterion in Section 5.1.

### 5.1 An upper bound criterion

Recall from Section 2.1 the space of paths  $\Pi$ , and the space  $\mathcal{H}$  of compact subsets of  $\Pi$ . Let  $\Pi_t \subset \Pi$  denote the set of paths starting at time  $t$ , and let  $\mathcal{N}_t := \mathcal{N} \cap \Pi_t$ . For a path  $\pi \in \Pi$  and any  $t \geq \sigma_\pi$ , let  $\pi^t := \{(\pi(s), s) : s \geq t\}$  denote the truncation of  $\pi$  at time  $t$  so that  $\sigma_{\pi^t} = t$ . Recall that the Brownian net  $\mathcal{N}$  satisfies the image set property stated in Proposition 2.8, which implies that for any  $\mathcal{H}$ -valued random variable  $\mathcal{X}$ , if

$$\mathcal{X}^t := \{\pi^t : \pi \in \mathcal{X}, \sigma_\pi \leq t\}$$

denotes the set of paths in  $\mathcal{X}$  starting before or at time  $t$  and truncated at time  $t$ , and almost surely  $\cup_{p \in \mathcal{X}^t} p \subset \cup_{p \in \mathcal{N}_t} p$  for all  $t \in \mathbb{R}$ , then  $\mathcal{X} \setminus \mathcal{N} = \emptyset$ . Motivated by this observation, we will formulate an upper bound criterion for the Brownian net using a family of *counting random variables*  $\eta$  to be defined below. Analogous counting random variables were used in the upper bound criteria for the Brownian web  $\mathcal{W}$  in [NRS05], which were in turn adapted from [FINR04].

Similar to the definition of the branching-coalescing point set from  $\mathcal{N}$ , for  $\mathcal{X} \in \mathcal{H}$ , we define

$$\xi_{\mathcal{X}}(t, h; a, b) := \{\pi(t+h) \cap (a, b) : \pi \in \mathcal{X}, \sigma_\pi \leq t\}, \quad (t \in \mathbb{R}, h > 0, a < b), \quad (5.1)$$

which is the point set generated at time  $t + h$  by paths in  $\mathcal{X}$  that start before or at time  $t$ . We then define the family of counting random variables

$$\eta_{\mathcal{X}}(t, h; a, b) := |\xi_{\mathcal{X}}(t, h; a, b)|, \quad (t \in \mathbb{R}, h > 0, a < b). \quad (5.2)$$

Here is our upper bound criterion, which is in fact also a convergence criterion since it already assumes tightness and the desired lower bound.

**Theorem 5.1 (Upper bound criterion)** *Let  $(\mathcal{X}_n)_{n \in \mathbb{N}}$  be a sequence of  $\mathcal{H}$ -valued random variables whose laws are tight. If any of its subsequential weak limits  $\mathcal{X}$  contains a copy of the Brownian net  $\mathcal{N}$ , and moreover for all  $t \in \mathbb{R}, h > 0, a < b$ ,*

$$\mathbb{E}[\eta_{\mathcal{X}}(t, h; a, b)] \leq \mathbb{E}[\eta_{\mathcal{N}}(t, h; a, b)], \quad (5.3)$$

then  $\mathcal{X}_n$  converges weakly to the Brownian net  $\mathcal{N}$  as  $n \rightarrow \infty$ .

**Remark 5.2** Due to the assumption that any subsequential weak limit  $\mathcal{X}$  contains a copy of the Brownian net  $\mathcal{N}$ , it is easily seen that the inequality in (5.3) actually implies equality.

**Proof** By assumption,  $\mathcal{X}$  is coupled to a Brownian net  $\mathcal{N}$  such that  $\mathcal{N} \subset \mathcal{X}$  a.s., which implies that  $\cup_{p \in \mathcal{X}^t} p \supset \cup_{p \in \mathcal{N}^t} p$  for all  $t \in \mathbb{R}$ . It remains to show that  $\mathcal{X} \setminus \mathcal{N} = \emptyset$  a.s. By the image set property Proposition 2.8, it suffices to show that a.s.,

$$\bigcup_{p \in \mathcal{X}^t} p \subset \bigcup_{p \in \mathcal{N}^t} p \quad \text{for all } t \in \mathbb{R}. \quad (5.4)$$

Note that (5.3) implies that a.s., for all  $t \in \mathbb{Q}, h \in \mathbb{Q} \cap (0, \infty)$  and  $a < b$ , we have  $\eta_{\mathcal{X}}(t, h; a, b) \leq \eta_{\mathcal{N}}(t, h; a, b)$ , which can be extended to all  $t \in \mathbb{R}$  and  $h > 0$  since  $\mathcal{X}$  and  $\mathcal{N}$  are compact sets of continuous paths. This immediately implies (5.4) and concludes the proof.  $\blacksquare$

## 5.2 Convergence of a stationary set of paths

For the standard Brownian net  $\mathcal{N}$ , recall from [SS08, Prop. 1.15] that its associated branching-coalescing point set has a unique invariant law given by that of a Poisson point process on  $\mathbb{R}$  with intensity 2. Let  $\Lambda$  be such a Poisson point set independent of  $\mathcal{N}$ , and let  $\mathcal{B} := \mathcal{N}(\Lambda \times \{0\})$  denote the subset of  $\mathcal{N}$  starting from  $\Lambda$  at time 0. Similarly, let  $\Lambda_\varepsilon$  be a random subset of  $\mathbb{Z}$ , whose law is invariant for the branching-coalescing point set generated by the Bernoulli net  $\tilde{N}_\varepsilon$ , see Lemma 4.3 and (4.28). Let  $B_\varepsilon := N_\varepsilon(\Lambda_\varepsilon \times \{0\})$ , the set of paths in the discrete net  $N_\varepsilon$  (not  $\tilde{N}_\varepsilon$ ) started from  $\Lambda_\varepsilon$  at time 0. Here is the main result of this subsection.

**Theorem 5.3 (Convergence of a stationary set of paths)** *Assume that the law of the family of rescaled discrete nets  $S_{\varepsilon, \sigma} N_\varepsilon$  is tight. As defined above, let  $\varepsilon \Lambda_\varepsilon$  and  $\Lambda$  be random subsets of  $\varepsilon \mathbb{Z}$  and  $\mathbb{R}$ , respectively, which are now regarded as random point measures on  $\mathbb{R}$ , equipped with vague convergence. Then we have  $(\varepsilon \Lambda_\varepsilon, S_{\varepsilon, \sigma} B_\varepsilon) \Rightarrow (\Lambda, \mathcal{B})$  as  $\varepsilon \downarrow 0$ .*

**Remark 5.4** We need to enhance the convergence  $S_{\varepsilon, \sigma} B_\varepsilon \Rightarrow \mathcal{B}$  with the convergence of the random measures  $\varepsilon \Lambda_\varepsilon \Rightarrow \Lambda$ , so that we can find a coupling between  $(\varepsilon \Lambda_\varepsilon, S_{\varepsilon, \sigma} B_\varepsilon)$  and  $(\Lambda, \mathcal{B})$  such that not only do the locations of the atoms in  $\varepsilon \Lambda_\varepsilon$  converge to that in  $\Lambda$ , but also the number of atoms converge. This will be used to deduce the convergence of the rescaled discrete net paths starting from a single point, see Proposition 5.5 below.

**Proof** Thanks to Lemma 4.3 and (4.28), the convergence  $\varepsilon \Lambda_\varepsilon \Rightarrow \Lambda$  as random measures on  $\mathbb{R}$  is clear. By Skorohod's representation theorem, we can choose a coupling such that a.s.,  $\varepsilon \Lambda_\varepsilon \rightarrow \Lambda$  w.r.t. the vague topology. It remains to show that conditioned on a typical realisation of  $\Lambda_\varepsilon$  and  $\Lambda$ ,  $S_{\varepsilon, \sigma} B_\varepsilon \Rightarrow \mathcal{B}$ .

The tightness of  $(S_{\varepsilon,\sigma}N_\varepsilon)_{\varepsilon>0}$  implies the tightness of  $(S_{\varepsilon,\sigma}B_\varepsilon)_{\varepsilon>0}$ . It is easily seen that the analogue of Theorem 5.1 holds if  $\mathcal{N}$  is replaced by  $\mathcal{B} = \mathcal{N}(\Lambda \times \{0\})$  for any closed  $\Lambda \subset \mathbb{R}$  independent of  $\mathcal{N}$ . To show that any subsequential weak limit  $\mathcal{A}$  of  $S_{\varepsilon,\sigma}B_\varepsilon$  contains a copy of  $\mathcal{B}$ , we can apply the same discrete approximation argument as in the proof of Theorem 3.5, provided we use the fact from Theorem 2.7 that  $\mathcal{N}(\Lambda \times \{0\})$  can be constructed via the hopping construction where the first hopping path is chosen from  $\mathcal{W}^1(\Lambda \times \{0\}) \cup \mathcal{W}^2(\Lambda \times \{0\})$ . In particular, this implies that a.s.,

$$\mathbb{E}[\tilde{\eta}_{\mathcal{A}}(t, h; a, b) | \Lambda] \geq \mathbb{E}[\eta_{\mathcal{B}}(t, h; a, b) | \Lambda] \quad \text{for all } t \geq 0, h > 0, a < b \in \mathbb{R}, \quad (5.5)$$

To show the reverse inequality, it suffices to show the inequality after taking expectation, i.e., show that

$$\mathbb{E}[\eta_{\mathcal{A}}(t, h; a, b)] \leq \mathbb{E}[\eta_{\mathcal{B}}(t, h; a, b)] \quad \text{for all } t \geq 0, h > 0, a < b \in \mathbb{R}, \quad (5.6)$$

where we note that  $\mathbb{E}[\eta_{\mathcal{B}}(t, h; a, b)] = 2(b - a)$  by [SS08, Prop. 1.15]. Recall from Lemma 4.1 that  $N_\varepsilon$  can be coupled to a Bernoulli net  $\tilde{N}_\varepsilon$  such that  $N_\varepsilon \subset \tilde{N}_\varepsilon$  a.s., and hence also  $B_\varepsilon \subset \tilde{B}_\varepsilon := \tilde{N}_\varepsilon(\Lambda_\varepsilon \times \{0\})$ . By Lemma 4.3 and (4.28),  $\Lambda_\varepsilon$  is an invariant law for the branching-coalescing point set generated by the Bernoulli net  $\tilde{N}_\varepsilon$ , and  $\Lambda_\varepsilon$  has density  $\rho = (2 + o(1))\varepsilon$ . It follows by Fatou's lemma that under the scaling map  $S_{\varepsilon,\sigma} : (x, t) \rightarrow (\varepsilon x, \sigma^2 \varepsilon^2 t)$  with  $\varepsilon \downarrow 0$ , at each time  $t \geq 0$ , the point set generated at time  $t$  by paths in the subsequential weak limit  $\mathcal{A}$  has density at most 2. Therefore  $\mathbb{E}[\eta_{\mathcal{A}}(t, h; a, b)] \leq 2(b - a)$ , which concludes the proof of (5.6).  $\blacksquare$

### 5.3 Convergence of finite dimensional distributions

We now use Theorem 5.3 to deduce the convergence of paths in the discrete net  $N_\varepsilon$  starting from finitely many space-time points. We first treat the one point case. Recall that for  $\mathcal{X} \in \mathcal{H}$ ,  $A \subseteq \mathbb{R}^2$  and  $z \in \mathbb{R}^2$ ,  $\mathcal{X}(A)$  denotes the set of paths in  $\mathcal{X}$  starting from the space-time set  $A$ , and  $\mathcal{X}(z) = \mathcal{X}(\{z\})$ .

**Proposition 5.5 (Convergence of one dimensional distribution)** *Assume that the law of the family of rescaled discrete nets  $S_{\varepsilon,\sigma}N_\varepsilon$  is tight. Let  $N_\varepsilon(0, 0)$  be the subset of paths in  $N_\varepsilon$  starting from the origin at time 0. Then we have  $S_{\varepsilon,\sigma}N_\varepsilon(0, 0) \Rightarrow \mathcal{N}(0, 0)$  as  $\varepsilon \downarrow 0$ .*

**Proof** Recall from Section 2.1 the space  $\mathcal{H}$  of compact subsets of the paths space  $\Pi$ . Let  $\Pi(0, 0)$  denote the subset of  $\Pi$  starting from  $(0, 0)$ , and let  $\mathcal{H}(0, 0)$  denote the space of compact subsets of  $\Pi(0, 0)$ . We will prove  $S_{\varepsilon,\sigma}N_\varepsilon(0, 0) \Rightarrow \mathcal{N}(0, 0)$  by regarding  $S_{\varepsilon,\sigma}N_\varepsilon(0, 0)$  and  $\mathcal{N}(0, 0)$  as  $\mathcal{H}(0, 0)$ -valued random variables, which also implies weak convergence if they are regarded as  $\mathcal{H}$ -valued random variables. We need to show that for any bounded continuous  $f : \mathcal{H}(0, 0) \rightarrow \mathbb{R}$ , we have

$$\mathbb{E}[f(S_{\varepsilon,\sigma}N_\varepsilon(0, 0))] \xrightarrow{\varepsilon \downarrow 0} \mathbb{E}[f(\mathcal{N}(0, 0))]. \quad (5.7)$$

Let  $\Pi(\mathbb{R} \times \{0\})$  denote the subset of  $\Pi$  starting from  $\mathbb{R}$  at time 0, and let  $\mathcal{H}(\mathbb{R} \times \{0\})$  denote the space of compact subsets of  $\Pi(\mathbb{R} \times \{0\})$ . Let  $\mathcal{H}_{\text{fin}}(\mathbb{R} \times \{0\})$  denote the subset of  $\mathcal{H}(\mathbb{R} \times \{0\})$  such that for  $\mathcal{X} \in \mathcal{H}_{\text{fin}}(\mathbb{R} \times \{0\})$ ,  $\mathcal{X}_0 = \{\pi(0) : \pi \in \mathcal{X}\}$  is locally finite subset of  $\mathbb{R}$ . For any bounded continuous  $f : \mathcal{H}(0, 0) \rightarrow \mathbb{R}$ , we can define  $g : \mathcal{H}_{\text{fin}}(\mathbb{R} \times \{0\}) \rightarrow \mathbb{R}$  by

$$g(\mathcal{X}) = \sum_{x \in \mathcal{X}_0 \cap [0, 1]} f(\mathcal{X}(x, 0) - x) \quad (\mathcal{X} \in \mathcal{H}_{\text{fin}}(\mathbb{R} \times \{0\})), \quad (5.8)$$

where  $\mathcal{X}(x, 0) - x$  denotes a spatial shift of  $-x$  of paths in  $\mathcal{X}(x, 0)$ . By Theorem 5.3, we can find a coupling such that a.s.  $S_{\varepsilon,\sigma}B_\varepsilon \rightarrow \mathcal{B}$  in  $\mathcal{H}(\mathbb{R} \times \{0\})$  and  $\varepsilon\Lambda_\varepsilon \rightarrow \Lambda$  w.r.t. the vague

topology, i.e., the number of atoms in the random set  $\varepsilon\Lambda_\varepsilon \cap [0, 1]$  and their locations converge to that of  $\Lambda \cap [0, 1]$ , where we note that a.s.  $\Lambda$  has no atom on the boundary of  $[0, 1]$ . This is easily seen to imply

$$\mathbb{E}[g(S_{\varepsilon,\sigma}B_\varepsilon)] \xrightarrow{\varepsilon \downarrow 0} \mathbb{E}[g(\mathcal{B})]. \quad (5.9)$$

By translation invariance and the fact that  $\Lambda_\varepsilon$  has density  $\rho = (2 + o(1))\varepsilon$ , see (4.28), and  $\Lambda$  has density 2, (5.9) is equivalent to

$$\varepsilon^{-1}\rho\mathbb{E}[f(S_{\varepsilon,\sigma}N_\varepsilon(0,0))] \rightarrow 2\mathbb{E}[f(\mathcal{N}(0,0))].$$

The desired convergence in (5.7) follows immediately.  $\blacksquare$

We now treat the case with finitely many starting points in space and time.

**Theorem 5.6** *Assume that the law of the family of rescaled discrete nets  $S_{\varepsilon,\sigma}N_\varepsilon$  is tight. Let  $k \in \mathbb{N}$ . Suppose that for each  $1 \leq i \leq k$ ,  $z_i^\varepsilon \in \mathbb{Z}^2$  satisfies  $S_{\varepsilon,\sigma}z_i^\varepsilon \rightarrow z_i \in \mathbb{R}^2$  as  $\varepsilon \downarrow 0$ . Then as random variables taking values in  $\mathcal{H}^k$ , we have*

$$(S_{\varepsilon,\sigma}N_\varepsilon(z_1^\varepsilon), \dots, S_{\varepsilon,\sigma}N_\varepsilon(z_k^\varepsilon)) \xrightarrow{\varepsilon \downarrow 0} (\mathcal{N}(z_1), \dots, \mathcal{N}(z_k)). \quad (5.10)$$

**Proof** Tightness follows by assumption. By the same discrete approximation argument as in the proof of Theorem 5.3, we note that every subsequential weak limit  $(\mathcal{A}_1, \dots, \mathcal{A}_k)$  of  $(S_{\varepsilon,\sigma}N_\varepsilon(z_1^\varepsilon), \dots, S_{\varepsilon,\sigma}N_\varepsilon(z_k^\varepsilon))$  can be coupled to a Brownian net  $\mathcal{N}$  such that  $\mathcal{A}_i \supset \mathcal{N}(z_i)$  for each  $1 \leq i \leq k$ . On the other hand, the convergence in Theorem 5.5 implies that a.s., we must have  $\mathcal{A}_i = \mathcal{N}(z_i)$  for each  $i$ . This concludes the proof.  $\blacksquare$

## 5.4 Proof of the upper bound

In this subsection, we prove the desired upper bound on subsequential weak limits of  $S_{\varepsilon,\sigma}N_\varepsilon$  by verifying condition (5.3) in Theorem 5.1.

**Theorem 5.7 (Upper bound)** *Let  $N_\varepsilon$  be the discrete net as in Theorem 1.1. For any subsequential weak limit  $\mathcal{N}^*$  of  $S_{\varepsilon,\sigma}N_\varepsilon$  as  $\varepsilon \downarrow 0$ , there exists a standard Brownian net  $\mathcal{N}$  such that almost surely,  $\mathcal{N} = \mathcal{N}^*$ .*

**Proof** By going to a subsequence if necessary, we can assume that the rescaled discrete net  $S_{\varepsilon,\sigma}N_\varepsilon$  converges in distribution to a limit  $\mathcal{N}^*$ , which by Theorem 3.5, is coupled to a standard Brownian net such that  $\mathcal{N} \subseteq \mathcal{N}^*$  a.s. In order to show  $\mathcal{N}^* \subseteq \mathcal{N}$ , by Theorem 5.1, it only remains to check (5.3) for the counting random variables  $\eta_{\mathcal{N}^*}(t, h; a, b)$  and  $\eta_{\mathcal{N}}(t, h; a, b)$ . By translation invariance in time, we only need to show that for all  $h > 0$  and  $a < b$ ,

$$\mathbb{E}[\eta_{\mathcal{N}^*}(0, h; a, b)] \leq \mathbb{E}[\eta_{\mathcal{N}}(0, h; a, b)]. \quad (5.11)$$

We prove (5.11) by an argument similar to the one in [NRS05, Section 6] in the proof of convergence to the Brownian web. The strategy is to first show that

$$\xi_\delta^{\mathcal{N}^*} := \{\pi(\delta) : \pi \in \mathcal{N}^*, \sigma_\pi \leq 0\}, \quad (5.12)$$

the branching-coalescing point set generated at time  $\delta$  by paths in  $\mathcal{N}^*$  started before or at time 0, is a locally finite subset of  $\mathbb{R}$  (sometimes called coming down from infinite). We then apply Theorem 5.6 to show that the set of truncated paths  $\{\pi^\delta : \pi \in \mathcal{N}^*, \sigma_\pi \leq 0\}$ , truncated at time  $\delta$ , is distributed as the set of Brownian net paths  $\mathcal{N}(\xi_\delta^{\mathcal{N}^*} \times \{\delta\})$  starting from  $\xi_\delta^{\mathcal{N}^*}$  at time  $\delta$ , which allows us to bound

$$\mathbb{E}[\eta_{\mathcal{N}^*}(0, h; a, b)] \leq \mathbb{E}[\eta_{\mathcal{N}}(\delta, h - \delta; a, b)].$$

Sending  $\delta \downarrow 0$  then gives (5.11).

We now give the details of the proof. By the density bound in Theorem 4.6, there exists a constant  $C$  such that for all  $0 < \delta \leq h \wedge 1$  and  $a < b$ ,

$$\limsup_{\varepsilon \downarrow 0} \mathbb{E}[\eta_{S_{\varepsilon, \sigma} N_{\varepsilon}}(0, \delta; a, b)] \leq \frac{C\sigma(b-a)}{\sqrt{\delta}} < \infty. \quad (5.13)$$

Since we have assumed  $S_{\varepsilon, \sigma} N_{\varepsilon} \Rightarrow \mathcal{N}^*$ , by Fatou's Lemma, we must have  $\mathbb{E}[\eta_{\mathcal{N}^*}(0, \delta; a, b)] < \infty$ , which implies that  $\xi_{\delta}^{\mathcal{N}^*}$ , defined in (5.12) above, is a.s. locally finite. Let  $\xi_t^{\varepsilon}$  denote the branching-coalescing point set generated by  $N_{\varepsilon}$ , starting from  $\xi_0^{\varepsilon} = \mathbb{Z}$ . Then thanks to the uniform density bound in (5.13) and the convergence  $S_{\varepsilon, \sigma} N_{\varepsilon} \Rightarrow \mathcal{N}^*$  as  $\mathcal{H}$ -valued random variables, by going to a further subsequence if necessary, we have  $\varepsilon \xi_{\delta/\sigma^2 \varepsilon^2}^{\varepsilon} \Rightarrow \nu_{\delta}^{\mathcal{N}^*}$  as random variables taking values in the space of locally finite measures on  $\mathbb{R}$ , equipped with vague convergence. We can assume a coupling such that  $\varepsilon \xi_{\delta/\sigma^2 \varepsilon^2}^{\varepsilon} \rightarrow \nu_{\delta}^{\mathcal{N}^*}$  a.s., and the support of  $\nu_{\delta}^{\mathcal{N}^*}$  equals  $\xi_{\delta}^{\mathcal{N}^*}$ , where  $\nu_{\delta}^{\mathcal{N}^*}$  may assign integer mass greater than 1 to some points in  $\xi_{\delta}^{\mathcal{N}^*}$ . For any finite interval  $(c, d)$ , the a.s. vague convergence  $\varepsilon \xi_{\delta/\sigma^2 \varepsilon^2}^{\varepsilon} \rightarrow \nu_{\delta}^{\mathcal{N}^*}$  and the fact that  $\nu_{\delta}^{\mathcal{N}^*}$  assigns no mass to the boundary of  $(c, d)$  imply that the number of points in  $\varepsilon \xi_{\delta/\sigma^2 \varepsilon^2}^{\varepsilon} \cap (c, d)$  will converge, and their locations will also converge. We can now apply Theorem 5.6 to  $N_{\varepsilon}(\xi_{\delta/\sigma^2 \varepsilon^2}^{\varepsilon} \cap (c/\varepsilon, d/\varepsilon) \times \{\delta/\sigma^2 \varepsilon^2\})$ ,<sup>1</sup> which by the Markov property of paths in  $N_{\varepsilon}$ , is distributed as branching-coalescing random walks starting from the set  $\xi_{\delta/\sigma^2 \varepsilon^2}^{\varepsilon} \cap (c/\varepsilon, d/\varepsilon)$  at time  $\delta/\sigma^2 \varepsilon^2$ . In particular, conditioned on a realisation of the coupling  $\varepsilon \xi_{\delta/\sigma^2 \varepsilon^2}^{\varepsilon} \rightarrow \nu_{\delta}^{\mathcal{N}^*}$ , we have

$$S_{\varepsilon, \sigma} \left( N_{\varepsilon}(\xi_{\delta/\sigma^2 \varepsilon^2}^{\varepsilon} \cap (c/\varepsilon, d/\varepsilon) \times \{\delta/\sigma^2 \varepsilon^2\}) \right) \Rightarrow \mathcal{N}(\xi_{\delta}^{\mathcal{N}^*} \cap (c, d) \times \{\delta\}) \quad (c < d \in \mathbb{R}),$$

where  $\mathcal{N}$  is a standard Brownian net. This implies that we can couple  $\mathcal{N}^*$  with a standard Brownian net  $\mathcal{N}$  such that a.s., the set of truncated paths

$$\{\pi^{\delta} : \pi \in \mathcal{N}^*, \sigma_{\pi} \leq 0\} = \{(\pi(t), t)_{t \geq \delta} : \pi \in \mathcal{N}^*, \sigma_{\pi} \leq 0\}$$

is contained in  $\mathcal{N}(\xi_{\delta}^{\mathcal{N}^*} \times \{\delta\})$ . Furthermore, by comparing with paths in  $\mathcal{N}(\mathbb{R} \times \{\delta\})$ , we have

$$\eta_{\mathcal{N}^*}(0, h; a, b) \leq \eta_{\mathcal{N}(\xi_{\delta}^{\mathcal{N}^*} \times \{\delta\})}(\delta, h - \delta; a, b) \leq \eta_{\mathcal{N}}(\delta, h - \delta; a, b). \quad (5.14)$$

Take expectation on both sides and apply the density calculation for  $\mathcal{N}$  in [SS08, Prop. 1.12], we obtain

$$\mathbb{E}[\eta_{\mathcal{N}^*}(0, h; a, b)] \leq \mathbb{E}[\eta_{\mathcal{N}}(\delta, h - \delta; a, b)] = (b - a) \cdot \left( \frac{e^{-(h-\delta)}}{\sqrt{\pi(h-\delta)}} + 2\Phi(\sqrt{2(h-\delta)}) \right), \quad (5.15)$$

where  $\Phi$  is the distribution function of the standard normal. Sending  $\delta \downarrow 0$  then gives (5.11) and concludes the proof.  $\blacksquare$

## 6 Tightness

In this section, we prove that the laws of the rescaled discrete nets  $(S_{\varepsilon, \sigma} N_{\varepsilon})_{\varepsilon \in (0, 1)}$  are tight. In Section 6.1, we recall the standard tightness criterion for a family of  $\mathcal{H}$ -valued random variables. To address the fundamental difficulty created by branching, we introduce in Section 6.2 the so-called relevant branching points of the branching-coalescing random walks. In Sections 6.3 and 6.4, we show that such relevant branching points are rare on small time scales, which then allows us to implement in Section 6.5 the multiscale argument used in [BMSV06] to verify the tightness criterion for the rescaled discrete webs.

<sup>1</sup>The argument here differs from the one in [NRS05, Section 6] for the Brownian web. Translated into our setting, it would mean using the weaker convergence of  $\varepsilon \xi_{\delta/\sigma^2 \varepsilon^2}^{\varepsilon} \rightarrow \xi_{\delta}^{\mathcal{N}^*}$  as closed subsets of  $\mathbb{R} \cup \{\pm\infty\}$  w.r.t. the Hausdorff distance. The rest of the argument would then require a stronger version of Proposition 5.5 (cf. [NRS05, Lemma 6.5]), which would establish the convergence of  $N_{\varepsilon}(D_{\varepsilon}) \Rightarrow \mathcal{N}(0, 0)$  for any closed  $D_{\varepsilon} \subset \mathbb{R}^2$  that converges in Hausdorff distance to  $\{(0, 0)\}$ .

## 6.1 Tightness criterion

We recall here the standard tightness criterion for a sequence of  $\mathcal{H}$ -valued random variables  $(\mathcal{X}_n)_{n \in \mathbb{N}}$ , which was first formulated in [FINR04, Proposition B.1].

**Proposition 6.1 (Tightness criterion)** *A sequence of  $\mathcal{H}$ -valued random variables  $(\mathcal{X}_n)_{n \in \mathbb{N}}$  is tight if for any space-time box  $\Lambda_{L,T} := [-L, L] \times [-T, T]$  with  $L, T > 0$  and every  $M > 0$ ,*

$$\lim_{\delta \downarrow 0} \frac{1}{\delta} \limsup_{n \rightarrow \infty} \sup_{(x,t) \in \Lambda_{L,T}} \mathbb{P}(\mathcal{X}_n \in A_{M,\delta}(x,t)) = 0, \quad (6.1)$$

where  $A_{M,\delta}(x,t) \subset \mathcal{H}$  consists of compact sets of paths  $K \in \mathcal{H}$ , such that  $K$  contains some path which intersects the box  $[x - M, x + M] \times [t, t + \delta]$ , and at a later time, intersects the left or right boundary of the larger box  $[x - 2M, x + 2M] \times [t, t + 2\delta]$  (see Figure 1).

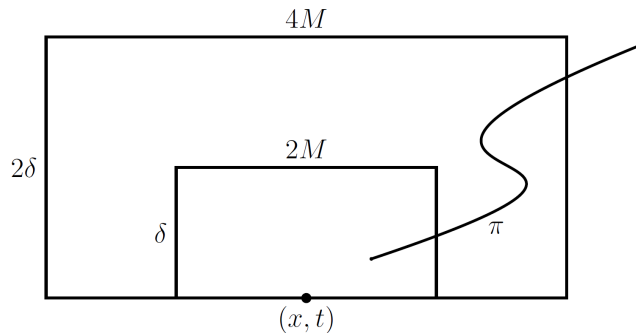


Figure 1: A path  $\pi$  causing the unlikely event  $\{\mathcal{X}_n \in A_{M,\delta}(x,t)\}$  to occur.

**Remark 6.2** *Proving tightness of  $(\mathcal{X}_n)_{n \in \mathbb{N}}$  amounts to controlling the uniform modulus of continuity of paths in  $\mathcal{X}_n$  in such a way that the probability bounds are uniform in  $n \in \mathbb{N}$ . Thanks to the compactification of space-time in Section 2.1, we can restrict to a large but finite space-time window. Dividing the time window into intervals of length  $\delta$  and controlling the uniform modulus of continuity on each time interval then leads to the condition (6.1). For more details, see e.g. [SSS17, Section 6.6].*

## 6.2 Relevant branching points

To verify the tightness condition (6.1) for the rescaled discrete nets  $S_{\varepsilon,\sigma}N_\varepsilon$ , we will adapt the multiscale argument developed in [BMSV06] to prove tightness for the rescaled discrete webs. A key ingredient is to show that the probability that some of the random walks and their descendants (due to branching) will exit the left or right boundary of the larger box  $[x - 2M, x + 2M] \times [t, t + 2\delta]$  in Figure 1 is very small. Branching creates fundamental difficulties for this estimate, which motivates us to consider the so-called *relevant branching points (RBP)* of the discrete net  $N_\varepsilon$ . Controlling the number of RBP along paths in  $N_\varepsilon(0,0)$  then allows us to show that the probability that some path in  $N_\varepsilon(0,0)$  makes a large excursion as depicted in Figure 1 is comparable to the probability that a single random walk starting from  $(0,0)$  makes a large excursion. We now give the precise definition of RBP.

Let  $A \subset \mathbb{Z}^2$  be a set of space-time points. Given  $S, U \in \mathbb{Z}$  with  $S < U$ , a space-time point  $z = (x, t) \in \mathbb{Z}^2$  with  $t \in (S, U)$  is called a  $(S, U)$ -*relevant branching point (RBP)* of  $N_\varepsilon(A)$  if there exist two paths  $\pi_1, \pi_2 \in N_\varepsilon$  starting from  $A$ , such that

$$\pi_1(s) = \pi_2(s) \quad \forall s \in [S, t], \quad \pi_1(t) = \pi_2(t) = x, \quad \text{and} \quad \pi_1(s) \neq \pi_2(s) \quad \forall s \in (t, U] \cap \mathbb{Z}. \quad (6.2)$$

We call  $t$  an  $(S, U)$ -relevant branching time of  $N_\varepsilon(A)$ . When things are clear from the context, we will simply call  $z$  (resp.  $t$ ) a relevant branching point (resp. time). We say that a path  $\pi$  passes through  $z = (x, t)$  if  $\pi$  starts before or at time  $t$  and  $\pi(t) = x$ .

We define the branching-coalescing point set generated by  $N_\varepsilon(A)$  by

$$\xi_t^A := \{\pi(t) : \pi \in N_\varepsilon(A)\}. \quad (6.3)$$

To determine how paths in  $N_\varepsilon(A)$  lead from the set  $\xi_S^A \subset \mathbb{Z}$  at time  $S$  to the set  $\xi_U^A$  at time  $U$ , we consider the set of  $(S, U)$ -RBP of  $N_\varepsilon(A)$  and how each path  $\pi \in N_\varepsilon(A)$  passes through these RBP's. More precisely, we build a directed graph  $G_{S,U}(N_\varepsilon(A)) = (V, E)$  with the vertex set  $V$  consisting of  $\xi_S^A \times \{S\}$ , the set of  $(S, U)$ -RBP of  $N_\varepsilon(A)$ , and  $\xi_U^A \times \{U\}$ . For any two vertices  $z_1 = (x_1, t_1), z_2 = (x_2, t_2) \in V$  with  $S \leq t_1 < t_2 \leq U$ , we draw a directed edge  $e$  from  $z_1$  to  $z_2$  if there exists  $\pi \in N_\varepsilon(A)$  with starting time  $\sigma_\pi \leq S$ , and  $\pi$  pass through  $z_1$  and  $z_2$ , but no  $(S, U)$ -RBP in between. Given such a directed edge  $e = (z_1, z_2)$ , we call  $e$  an outgoing edge at  $z_1$  and an incoming edge at  $z_2$ .

The definition of RBP above for the discrete net  $N_\varepsilon$  is motivated by a similar definition of *relevant separation points* for the Brownian net  $\mathcal{N}$  in [SSS09, Section 2.3] and [SSS17, Section 6.1]. The Brownian net analogue of the directed graph  $G_{S,U}(N_\varepsilon(A))$  was studied in [SSS17, Section 6.2] and was called the *finite graph representation* of the Brownian net  $\mathcal{N}$  in the time interval  $[S, U]$ , which played a crucial role in [SSS17] in showing that the coupling between the discrete sticky webs and the discrete net converge to a coupling between the sticky Brownian webs and the Brownian net.

For the discrete net  $N_\varepsilon$ , we show that the RBP and the directed graph  $G_{S,U}$  share similar properties to their continuum counterparts for the Brownian net  $\mathcal{N}$ .

Our first result shows that if  $\pi \in N_\varepsilon(A)$  starting at or before time  $S$  passes through  $z_1 = (x_1, t_1)$  and  $z_2 = (x_2, t_2)$  for some directed edge  $(z_1, z_2)$  in  $G_{S,U}(N_\varepsilon(A))$ , then descendants of  $\pi$  resulting from branching in the time interval  $(t_1, t_2)$  must all pass through  $z_2$ .

**Lemma 6.3 (No effective branching)** *For  $S < U$  and  $A \subset \mathbb{Z}^2$ , let  $G_{S,U}(N_\varepsilon(A)) =: G = (V, E)$  be defined as above. Let  $e = (z_1, z_2)$ , with  $z_i = (x_i, t_i) \in \mathbb{Z}^2$ , be a directed edge of  $G$ . If  $\pi \in N_\varepsilon(A)$  with  $\sigma_\pi \leq S$  passes through  $z_1$  and  $z_2$  and no  $(S, U)$ -RBP in between, then there is no effective branching along  $\pi$  in the time interval  $(t_1, t_2)$ , that is, for any  $\varpi \in N_\varepsilon(A)$  that passes through  $z_1$  and  $\varpi(s) = \pi(s)$  for some  $s \in (t_1, t_2) \cap \mathbb{Z}$ ,  $\varpi$  must pass through  $z_2$  and no  $(S, U)$ -RBP in the time interval  $(s, U)$ .*

**Proof** We will implicitly assume that all times are integers. By assumption, either  $t_2 = U$  or  $z_2$  is an  $(S, U)$ -RBP. If  $t_2 = U$  and  $\varpi$  does not pass through  $z_2$ , then a.s.  $t_0 := \sup\{n \leq U : \varpi(n) = \pi(n)\} < U$ . By hopping from  $\pi$  to  $\varpi$  at time  $t_0$ , which still defines a path in  $N_\varepsilon(A)$ , it is easily seen that  $z_0 = (\pi(t_0), t_0)$  is also an  $(S, U)$ -RBP for  $N_\varepsilon(A)$ . This contradicts the assumption that  $\pi$  does not pass through any RBP between  $z_1$  and  $z_2$ .

If  $z_2$  is an  $(S, U)$ -RBP, then there exist two paths  $\pi_1, \pi_2 \in N_\varepsilon(z_2)$  which do not meet on the time interval  $(t_2, U]$ , see Figure 2. If  $\varpi$  does not pass through  $z_2$ , then we have  $t_0 := \sup\{n \leq t_2 : \varpi(n) = \pi(n)\} \in (s, t_2)$ . Denote  $z_0 = (\pi(t_0), t_0)$ . We claim that  $z_0$  must be a RBP, which again leads to a contradiction. Indeed, if  $\varpi$  intersects neither  $\pi_1$  nor  $\pi_2$  on the time interval  $(t_2, U] \cap \mathbb{Z}$ , then hopping from  $\pi$  to  $\pi_1$  at time  $t_2$ , and hopping from  $\pi$  to  $\varpi$  at time  $t_0$ , identifies two paths in  $N_\varepsilon(A)$  which shows that  $z_0$  is a RBP. If  $\varpi$  intersects  $\pi_1 \cup \pi_2$  for the first time at some  $n \in (t_2, U]$  (say the intersection is with  $\pi_1$ ), then hopping from  $\pi$  to  $\varpi$  at time  $t_0$  and then from  $\varpi$  to  $\pi_1$  at time  $n$ , and hopping from  $\pi$  to  $\pi_2$  at time  $t_2$ , identifies two paths in  $N_\varepsilon(A)$  which shows that  $z_0$  is a RBP.

Lastly we note that since all  $\varpi \in N_\varepsilon$  that coincides with  $\pi$  at some time  $s \in (t_1, t_2)$  must pass through  $z_2$ , it means that  $z_2$  is a bottleneck and hence  $\varpi$  cannot pass through any  $(S, U)$ -RBP before reaching  $z_2$ .  $\blacksquare$

As a corollary of Lemma 6.3, we show that the reason a branching point along some



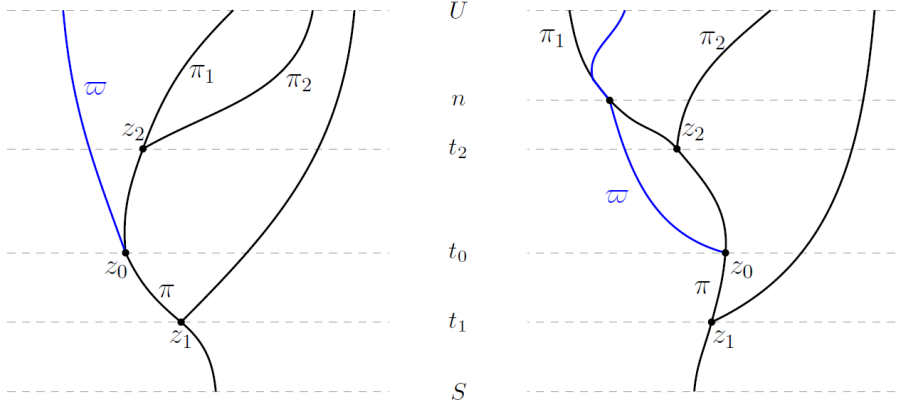


Figure 2: Two cases of  $\varpi$  branching off from  $\pi$  at  $z_0$ , but does not pass through the RBP  $z_2$ .

$\pi \in N_\varepsilon(A)$  with  $\sigma_\pi \leq S$  fails to be  $(S, U)$ -RBP is because all its descendants must pass through a common bottleneck before or at time  $U$ .

**Corollary 6.4 (Non-relevant branching points)** *Let  $S < U$ . If some  $\pi \in N_\varepsilon$  with starting time  $\sigma_\pi \leq S$  passes through  $z = (x, t) \in \mathbb{Z}^2$  with  $t \in (S, U)$ , and  $z$  is a branching point but not an  $(S, U)$ -RBP for  $N_\varepsilon$ , then there exists  $z' = (x', t')$  with  $t' \in (t, U]$  such that all paths in  $N_\varepsilon(z)$  passes through  $z'$ .*

**Proof** Let  $z_1$  be the last vertex in  $G_{S,U}(N_\varepsilon)$  that  $\pi$  passes through before time  $t$ , and let  $z_2$  be the first vertex in  $G_{S,U}(N_\varepsilon)$  that  $\pi$  passes through after time  $t$ . Then  $(z_1, z_2)$  is a directed edge in  $G_{S,U}$ , and hence Lemma 6.3 implies that all paths in  $N_\varepsilon(z)$  must pass through  $z_2$ . ■

Our next result shows that for each  $(S, U)$ -RBP  $z$ , the two branches of paths starting from  $z$  will not intersect until one branch visits the next RBP or reaches time  $U$ .

**Lemma 6.5 (Disjoint paths from relevant branching points)** *For  $S < U$  and  $A \subset \mathbb{Z}^2$ , let  $G_{S,U}(N_\varepsilon(A))$  be defined as before. Let  $z = (x, t)$  be a  $(S, U)$ -RBP. If  $\pi_1, \pi_2 \in N_\varepsilon(z)$  follow the two different branches from  $z$ , i.e.,  $\pi_1(t+1) \neq \pi_2(t+1)$ , then  $\pi_1$  and  $\pi_2$  follow two different edges  $(z, z_1)$  and  $(z, z_2)$  in the directed graph  $G_{S,U}$ , where  $z_1 = (x_1, t_1) \neq z_2 = (x_2, t_2)$ , and*

$$\pi_1(s) \neq \pi_2(s) \quad \text{for all } t < s \leq \min\{t_1, t_2\}. \quad (6.4)$$

**Proof** Without loss of generality, we assume that  $t_1 \leq t_2$ . Suppose that  $\pi_1(t_0) = \pi_2(t_0)$  for some  $t < t_0 < t_1$ . Then Lemma 6.3 implies that  $\pi_2$  must also pass through  $z_1 = (x_1, t_1)$  and follows the same directed  $(z, z_1)$  in  $G_{S,U}$ . Since there is only binary branching in the discrete net  $N_\varepsilon$ , any path in  $N_\varepsilon(z)$  must either coincide with  $\pi_1$  or  $\pi_2$  at time  $t+1$ . Again, Lemma 6.3 implies that all such paths must pass through  $z_1$ , which contradicts the assumption that  $z$  is a RBP. Therefore  $\pi_1$  and  $\pi_2$  must be disjoint until  $z_1$ , and hence  $\pi_2$  follows a different directed edge  $(z, z_2)$  in  $G_{S,U}$ . ■

Combing Lemmas 6.3 and 6.5, we obtain the following result on  $G_{S,U}$ .

**Proposition 6.6** *For  $S < U$  and  $A \subset \mathbb{Z}^2$ , let  $G_{S,U}(N_\varepsilon(A))$  be defined as before. If  $z = (x, t)$  is an  $(S, U)$ -RBP for  $N_\varepsilon(A)$ , then there are exactly two outgoing edges at  $z$  in  $G_{S,U}$ . If  $z \in \xi_S^A \times \{S\}$ , then there is either one or two outgoing edges at  $z$  in  $G_{S,U}$ .*

**Proof** If  $z = (x, t)$  is an  $(S, U)$ -RBP, then we can find  $\pi_1, \pi_2 \in N_\varepsilon$  passing through  $z$  such that  $\pi(s) \neq \pi_2(s)$  for  $s \in (t, U]$ . By Lemma 6.5,  $\pi_1$  and  $\pi_2$  follow two different directed edges  $(z, z_1)$  and  $(z, z_2)$  in  $G_{S,U}$ . Any other path  $\varpi$  passing through  $z$  coincides with either  $\pi_1$  or

$\pi_2$  at time  $t + 1$ , and hence by Lemma 6.3 must follow the same directed edge in  $G_{S,U}$ . This implies that there are exactly two outgoing edges at  $z$  in  $G_{S,U}$ . If  $z \in \xi_S^A \times \{S\}$ , the argument is similar, and there is either one or two outgoing edges at  $z$  because there is either one or two branches in  $N_\varepsilon$  coming out of  $z$ .  $\blacksquare$

### 6.3 Probability of having at least 1 relevant branching point

We now consider the branching-coalescing point set  $\xi$  generated by the discrete net  $N_\varepsilon$ , starting from  $\xi_0 = \{0\}$ . Let  $\mathcal{R}_T$  denote the number of  $(0, T)$ -relevant branching points for  $N_\varepsilon(0, 0)$ . We give here an upper bound on  $\mathbb{P}(\mathcal{R}_T \geq 1)$ , or equivalently,  $\mathbb{P}(|\xi_T| \geq 2)$ . We will bootstrap this bound to obtain bounds on  $\mathbb{P}(\mathcal{R}_T \geq K)$  for  $K \geq 2$  in the next subsection.

**Proposition 6.7 (At least 1 RBP)** *Let  $(\xi_t)_{t \geq 0}$  be the branching-coalescing point set generated by the discrete net  $N_\varepsilon$  with  $\xi_0 = \{0\}$ . Let  $\mathcal{R}_T$  be the number of  $(0, T)$ -RBP for  $N_\varepsilon(0, 0)$ . Then there exist  $C, \delta_0, \varepsilon_0 > 0$  such that for all  $\varepsilon \leq \varepsilon_0$  and  $1 \leq T \leq \delta_0 \varepsilon^{-2}$ ,*

$$\mathbb{P}(\mathcal{R}_T \geq 1) = \mathbb{P}(|\xi_T| \geq 2) \leq C\varepsilon\sqrt{T}. \quad (6.5)$$

**Proof strategy** The proof starts with the fact that the i.i.d. Bernoulli product law with parameter  $\rho_\varepsilon = 2\varepsilon + O(\varepsilon^2)$  as in (4.28) is invariant for the branching-coalescing point set  $(\tilde{\xi}_t)_{t \geq 0}$  generated by the Bernoulli net  $\tilde{N}_\varepsilon$  (see Lemma 4.3), which thanks to the coupling between  $N_\varepsilon$  and  $\tilde{N}_\varepsilon$  in Lemma 4.1, gives an upper bound on the branching-coalescing point set  $(\xi_t)_{t \geq 0}$  generated from  $N_\varepsilon$  with  $\xi_0 = \tilde{\xi}_0$ . The stationarity of  $\tilde{\xi}$  implies that for  $\xi$ , the reduction of particle density up to time  $T$  due to coalescence must be greater than the increase of density due to branching. The reduction in density can be computed for the collection of coalescing random walks starting from the same Bernoulli product law. By the reasoning above, this gives an upper bound on the number of offsprings of each particle (with branching turned on for that particle and its offsprings) that differ from all particles in the coalescing system. Restricted to the event that there is initially no other particles nearby in the coalescing system, this then leads to an upper bound on the probability that a single particle has more than 1 offspring at time  $T$ .  $\blacksquare$

We now make the above proof sketch rigorous. First we consider the coalescing random walks (CRW)  $(\eta_t)_{t \geq 0}$  with  $\eta_0$  following the i.i.d. Bernoulli product law with  $\mathbb{P}(0 \in \eta_0) = \rho_\varepsilon$  as in (4.28), and the random walk increment distribution  $a(\cdot)$  is the same as that for the discrete net  $N_\varepsilon$ . Note that there is a natural coupling between  $\eta$  and the branching-coalescing point set  $\xi$  generated by  $N_\varepsilon$ , such that almost surely,  $\eta_0 = \xi_0$ , and  $\eta_t \subset \xi_t$  for all  $t \geq 0$ . By translation invariance, the density of  $\eta_t$  equals

$$p_\eta(t) := \mathbb{P}(0 \in \eta_t) = \mathbb{P}(x \in \eta_t) \quad \text{for all } x \in \mathbb{Z}. \quad (6.6)$$

We now bound the decrease of density  $p_\eta(0) - p_\eta(t)$  due to coalescence.

**Lemma 6.8 (Density reduction for CRW)** *Let  $(\eta_t)_{t \geq 0}$  be the coalescing random walks defined above, starting from an i.i.d. Bernoulli product law with density  $\rho_\varepsilon$  as in (4.28). Then there exists a constant  $C > 0$  such that for any  $\varepsilon > 0$  and  $T > 0$ ,*

$$p_\eta(0) - p_\eta(T) \leq C\varepsilon^2\sqrt{T}. \quad (6.7)$$

**Proof** For each  $x \in \mathbb{Z}$ , let  $(\pi_t^x)_{t \geq 0}$  denote the coalescing random walk starting from  $x$  if  $x \in \eta_0$ ; otherwise let  $\pi_t^x \equiv *$  for some cemetery state  $*$ . Note that for any  $x, y \in \mathbb{Z}$ , conditioned on  $x, y \in \eta_0$ ,  $(\pi^x, \pi^y)$  is a pair of coalescing random walks, which up to the time they coalesce, has the same distribution as a pair of independent random walks  $(S^x, S^y)$  up to their first

meeting time  $\tau_{x,y} := \inf\{t \geq 0 : S_t^x = S_t^y\}$ . Let us fix a total order  $\prec$  on  $\mathbb{Z}$  with 0 being the minimal element, and let  $\sum_{x \in \mathbb{Z}}$  denote summation w.r.t. this total order. We can then bound

$$\begin{aligned}
p_\eta(T) &= \mathbb{P}(\pi_T^x = 0 \text{ for some } x \in \mathbb{Z}) = \sum_{x \in \mathbb{Z}} \mathbb{E} \left[ 1_{\{\pi_T^x=0\}} 1_{\{\pi_T^x \neq \pi_T^y \forall y \prec x\}} \right] \\
&\geq \sum_{x \in \mathbb{Z}} \left( \mathbb{E} [1_{\{\pi_T^x=0\}}] - \sum_{y \prec x} \mathbb{E} [1_{\{\pi_T^x = \pi_T^y = 0\}}] \right) \\
&\geq \rho_\varepsilon \sum_{x \in \mathbb{Z}} \mathbb{P}(S_T^x = 0) - \rho_\varepsilon^2 \sum_{x \neq y} \mathbb{P}(S_T^y = 0, \tau_{x,y} \leq T) \\
&= \rho_\varepsilon - \rho_\varepsilon^2 \sum_{z \neq 0} \mathbb{P}(\tau_{0,z} \leq T),
\end{aligned} \tag{6.8}$$

where in the last step, we made the change of variable  $z = y - x$  and used translation invariance. Let  $\tilde{S} := S - S'$  where  $S$  and  $S'$  are two independent random walks with the same jump distribution  $a(\cdot)$ . Then  $\tau_{0,z}$  has the same distribution as  $\tilde{\tau}_z$ , the first hitting time of  $z$  by  $\tilde{S}$  with  $\tilde{S}_0 = 0$ . Therefore,

$$\sum_{z \neq 0} \mathbb{P}(\tau_{0,z} \leq T) \leq \sum_{z \in \mathbb{Z}} \mathbb{P}(\tilde{\tau}_z \leq T) = \mathbb{E} [ |\cup_{t \in \mathbb{N}} \{\tilde{S}_t : t \leq T\}| ] \leq 2\mathbb{E} [ \sup_{0 \leq t \leq T} \tilde{S}_t ] \leq C\sqrt{T}, \tag{6.9}$$

where the last bound follows from Donsker's invariance principle. Substituting the bound into (6.8) and using  $\rho_\varepsilon = 2\varepsilon + O(\varepsilon^2)$  then gives (6.7).  $\blacksquare$

**Proof of Proposition 6.7** We now turn on branching along the coalescing random walks  $\eta$  in Lemma 6.8, namely, independently for each space-time point, we turn it into a branching point with probability  $\varepsilon$ , where a new random walk is born and makes an independent increment with distribution  $a(\cdot)$  satisfying (1.1). This leads to a branching-coalescing point set  $(\xi_t)_{t \geq 0}$  with  $\xi_0 = \eta_0$ , and  $\eta_t \subset \xi_t$  a.s. for all  $t \geq 0$ . Using the coupling between the discrete net  $N_\varepsilon$  and the Bernoulli net  $\tilde{N}_\varepsilon$  in Lemma 4.1, we can further couple  $(\xi_t)_{t \geq 0}$  with  $(\tilde{\xi}_t)_{t \geq 0}$ , the branching-coalescing point set generated from  $\tilde{N}_\varepsilon$  with  $\tilde{\xi}_0 = \xi_0 = \eta_0$ . Then  $(\tilde{\xi}_t)_{t \geq 0}$  is a stationary process by Lemma 4.3, and a.s.  $\xi_t \subset \tilde{\xi}_t$  for all  $t \geq 0$ .

Let  $J \in \mathbb{Z}$  denote the site in  $\eta_0$  closest to 0 (pick the positive  $J$  if there is a tie). If we replace the coalescing random walk path  $\pi^J$  starting from  $J$  by branching-coalescing random walks starting from  $J$ , and denote the resulting particle configuration by  $(\bar{\eta}_t)_{t \geq 0}$ , then there is a natural coupling such that almost surely,

$$\eta_t \subset \bar{\eta}_t \subset \xi_t \quad \text{for all } t \geq 0.$$

Since  $\xi_t \subset \tilde{\xi}_t$ , and  $\tilde{\xi}_t$  is stationary with density  $\rho_\varepsilon = p_\eta(0)$ , we can compare the mean number of particles in  $\xi_T \cap [-5\varepsilon^{-1}, 5\varepsilon^{-1}]$  and  $\eta_T \cap [-5\varepsilon^{-1}, 5\varepsilon^{-1}]$  and obtain the bound

$$\begin{aligned}
(10\varepsilon^{-1} + O(1))(p_\eta(0) - p_\eta(T)) &\geq \mathbb{E}[\xi_T \cap [-5\varepsilon^{-1}, 5\varepsilon^{-1}]] - \mathbb{E}[\eta_T \cap [-5\varepsilon^{-1}, 5\varepsilon^{-1}]] \\
&\geq \mathbb{E}[\bar{\eta}_T \cap [-5\varepsilon^{-1}, 5\varepsilon^{-1}]] - \mathbb{E}[\eta_T \cap [-5\varepsilon^{-1}, 5\varepsilon^{-1}]] \\
&\geq \mathbb{P}((\bar{\eta}_T \setminus \eta_T) \cap [-5\varepsilon^{-1}, 5\varepsilon^{-1}] \neq \emptyset) =: \mathbb{P}(A).
\end{aligned} \tag{6.10}$$

If  $\xi_t^J$  denotes the configuration of branching-coalescing random walks in the system  $\bar{\eta}$  that start from the single particle at position  $J$ , then  $\bar{\eta}_T \setminus \eta_T = \xi_T^J \setminus \eta_T$ , and  $A$  is just the event that  $\xi_t^J$  adds some new particles (compared to  $\eta_T$ ) in  $[-5\varepsilon^{-1}, 5\varepsilon^{-1}]$  at time  $T$ . We will lower bound  $\mathbb{P}(A)$  as follows.

First we restrict to the event

$$A_1 := \{ |\eta_0 \cap [-\varepsilon^{-1}, \varepsilon^{-1}]| = |\eta_0 \cap [-6\varepsilon^{-1}, 6\varepsilon^{-1}]| = 1 \},$$

namely that initially there is only one particle in  $[-\varepsilon^{-1}, \varepsilon^{-1}]$  (whose position is denoted by  $J$ ) and no particle in  $[-6\varepsilon^{-1}, 6\varepsilon^{-1}] \setminus [-\varepsilon^{-1}, \varepsilon^{-1}]$ . Note that uniformly in  $\varepsilon$  small,

$$\mathbb{P}(A_1) = (1 + o(1))2\varepsilon^{-1} \cdot \rho_\varepsilon \cdot (1 - \rho_\varepsilon)^{12\varepsilon^{-1}} \geq c_0 > 0. \quad (6.11)$$

We then further restrict to the event  $A_2$  that the coalescing random walk paths  $(\pi_t^x)_{t \geq 0}$  starting from all  $x \in \eta_0 \setminus [-6\varepsilon^{-1}, 6\varepsilon^{-1}]$  do not enter  $[-5\varepsilon^{-1}, 5\varepsilon^{-1}]$  up to time  $T$ , namely

$$A_2 := \left\{ \inf_{x \in \eta_0 \setminus [-6\varepsilon^{-1}, 6\varepsilon^{-1}]} \inf_{t \in [0, T]} |\pi_t^x| > 5\varepsilon^{-1} \right\} \subset \{ \eta_T \setminus \{ \pi_T^J \} \cap [-5\varepsilon^{-1}, 5\varepsilon^{-1}] = \emptyset \}.$$

Recall that  $T \leq \delta_0 \varepsilon^{-2}$ . Then uniformly in  $\eta_0 \in A_1$ , we have

$$\lim_{\delta_0 \downarrow 0} \lim_{\varepsilon \downarrow 0} \mathbb{P}(A_2^c | \eta_0) \leq 2 \lim_{\delta_0 \downarrow 0} \lim_{\varepsilon \downarrow 0} \mathbb{P}(\sup_{x \leq 0} \sup_{t \leq T} \pi_t^x \geq \varepsilon^{-1}) \leq 4 \lim_{\delta_0 \downarrow 0} \lim_{\varepsilon \downarrow 0} \mathbb{P}(\sup_{x \leq 0} \pi_T^x \geq \frac{1}{2}\varepsilon^{-1}) = 0, \quad (6.12)$$

where the second inequality follows by a stopping time argument (the time when one of the  $\pi^x$  exceeds  $\varepsilon^{-1}$ ) and Donsker's invariance principle, and the last equality follows from known results on the interface tightness of the dual voter model [BMV07] and a CLT for the location of the interface [AS11] (the last event in (6.12) is exactly the event that a dual voter model  $\phi_t$  starting with  $\phi_0(x) = 1_{\{x \leq 0\}}$  satisfies  $\phi_t(x) = 1$  for some  $x \geq 1/2\varepsilon$ ). Therefore uniformly in  $\eta_0 \in A_1$ , we can make  $\mathbb{P}(A_2 | \eta_0)$  arbitrarily close to 1 by choosing  $\delta_0 > 0$  and  $\varepsilon \leq \varepsilon_0$  small.

Lastly, we restrict to the event

$$A_3 := \left\{ \sup_{t \in [0, T]} |\pi_t^J| < 5\varepsilon^{-1} \right\} \cap \left\{ \exists \tilde{\pi}^J \in \xi^J \text{ s.t. } \sup_{t \in [0, T]} |\tilde{\pi}_t^J| < 5\varepsilon^{-1} \ \& \ \tilde{\pi}_T^J \neq \pi_T^J \right\},$$

where  $\tilde{\pi}^J \in \xi^J$  means  $\tilde{\pi}^J$  is a path among the collection of branching-coalescing random walks  $(\xi_t^J)_{t \geq 0}$ . Note that  $A_1 \cap A_2 \cap A_3 \subset A$ , the event defined in (6.10), and by translating  $J$  to 0, uniformly in  $\eta_0 \in A_1$ , we have

$$\mathbb{P}(A_3 | \eta_0) \geq \mathbb{P}(A_4) := \mathbb{P}\left( \sup_{t \in [0, T]} |\pi_t^0| < 4\varepsilon^{-1} \ \& \ \exists \tilde{\pi}^0 \in \xi^0 \text{ s.t. } \sup_{t \in [0, T]} |\tilde{\pi}_t^0| < 4\varepsilon^{-1} \ \& \ \tilde{\pi}_T^0 \neq \pi_T^0 \right). \quad (6.13)$$

Note that conditioned on  $\eta_0 \in A_1$ , the events  $A_2$  and  $A_3$  are independent since they depend on randomness in disjoint space-time regions. Therefore applying Lemma 6.8 to (6.10), we have

$$C\varepsilon\sqrt{T} \geq \mathbb{P}(A) \geq \mathbb{P}(A_1) \inf_{\eta_0 \in A_1} \mathbb{P}(A_2 | \eta_0) \inf_{\eta_0 \in A_1} \mathbb{P}(A_3 | \eta_0) \geq \frac{c_0}{2} \mathbb{P}(A_4),$$

and hence

$$\mathbb{P}(A_4) \leq C\varepsilon\sqrt{T}. \quad (6.14)$$

Note that  $A_4 \subset \{ |\xi_T^0| \geq 2 \}$ . To conclude the proof of Proposition 6.7, it only remains to prove the same probability bound for the event

$$A_5 := \{ |\xi_T^0| \geq 2 \} \setminus A_4 \subset A_6 \cup A_7,$$

where

$$A_6 := \left\{ \sup_{t \in [0, T]} |\pi_t^0| \geq \varepsilon^{-1} \right\},$$

$$A_7 := A_6^c \cap \left\{ \exists \tilde{\pi}^0 \text{ s.t. } \tilde{\pi}_T^0 \neq \pi_T^0 \right\} \cap \left\{ \forall \tilde{\pi}^0 \in \xi^0 \text{ s.t. } \tilde{\pi}_T^0 \neq \pi_T^0, \sup_{t \in [0, T]} |\tilde{\pi}_t^0| \geq 4\varepsilon^{-1} \right\}.$$

By a standard large deviation estimate for random walks whose increments  $a(\cdot)$  have finite  $\gamma$ -th moment, see e.g. [N79, Corollary 1.8] (the continuous-time random walk analogue is provided in [BMSV06, Lemma 5.1]),

$$\mathbb{P}(A_6) \leq C(e^{-c/(\varepsilon^2 T)} + T\varepsilon^\gamma) \leq C\varepsilon\sqrt{T}, \quad (6.15)$$

where the second inequality is easily seen to hold for  $\delta$  and  $\varepsilon$  small if  $\gamma \geq 2$ .

Lastly we show that  $\mathbb{P}(A_7) \leq C\varepsilon\sqrt{T}$ , which will imply the same bound for  $\mathbb{P}(A_5)$  and conclude the proof of Proposition 6.7. Note that on the event  $A_7$ , at least one branching-coalescing random walk in  $(\xi_t^0)_{t \geq 0}$  must exit  $[-2\varepsilon^{-1}, 2\varepsilon^{-1}]$  before time  $T$ . Let  $\tau$  denote the stopping time when this happens, and let  $A_8$  denote the event that all branching-coalescing random walks in  $(\xi_t^0)_{t \geq 0}$  that exit  $[-2\varepsilon^{-1}, 2\varepsilon^{-1}]$  at time  $\tau$  actually exit  $[-3\varepsilon^{-1}, 3\varepsilon^{-1}]$ . Using the graphical construction of  $(\xi_t^0)_{t \geq 0}$ , where we sample the realisation of the branchings and random walk jumps originating at each space-time point, we note that  $A_8$  occurs only if one of the random walk jumps originating from the space-time window  $[-2\varepsilon^{-1}, 2\varepsilon^{-1}] \times [0, T]$  jumps outside  $[-3\varepsilon^{-1}, 3\varepsilon^{-1}]$ . Let  $\Delta$  denote a random walk increment. Then by a union bound,

$$\mathbb{P}(A_8) \leq 2T \sum_{x \geq 0} \mathbb{P}(|\Delta| \geq \varepsilon^{-1} + x) \leq 2T \sum_{x \geq 0} \frac{\mathbb{E}[|\Delta|^\gamma]}{(\varepsilon^{-1} + x)^\gamma} \leq C\varepsilon^{\gamma-1}T \leq C\varepsilon\sqrt{T}, \quad (6.16)$$

where the last bound holds if  $\gamma \geq 3$ . On the other hand, we have

$$\mathbb{P}(A_7 \cap A_8^c) = \mathbb{E}[1_{A_8^c} \mathbb{P}(A_7 | (\xi_t^0)_{t \leq \tau})] \leq C\varepsilon\sqrt{T}, \quad (6.17)$$

because on the event  $A_8^c$ , there exists some  $x \in \xi_\tau^0 \cap [-3\varepsilon^{-1}, 3\varepsilon^{-1}] \setminus [-2\varepsilon^{-1}, 2\varepsilon^{-1}]$ . We can then sample a random walk  $S = (S_t)_{t \geq \tau}$  in  $(\xi_t^0)_{t \geq \tau}$  with  $S_\tau = x$ , whose law is still that of a random walk with transition kernel  $a(\cdot)$  since  $\tau$  is a stopping time. The event  $A_7$  implies that either  $S$  returns to the interval  $[-\varepsilon^{-1}, \varepsilon^{-1}]$  before time  $T$  to ensure  $S_T = \pi_T^0$ , or  $\sup_{t \in [\tau, T]} |S_t| \geq 4\varepsilon^{-1}$ . Either way, we must have  $\sup_{t \in [\tau, T]} |S_t - S_\tau| \geq \varepsilon^{-1}$ , the probability of which is bounded by  $C\varepsilon\sqrt{T}$  by (6.15). The bounds (6.16) and (6.17) imply  $\mathbb{P}(A_7) \leq C\varepsilon\sqrt{T}$ , which concludes the proof of Proposition 6.7.  $\blacksquare$

#### 6.4 Probability of having more than $K$ relevant branching points

We now bootstrap the bound on  $\mathbb{P}(\mathcal{R}_T \geq 1)$  in Proposition 6.7 to get a bound on  $\mathbb{P}(\mathcal{R}_T \geq K)$  for  $K \geq 2$ .

**Theorem 6.9 (More than  $K$  RBP)** *Let  $(\xi_t)_{t \geq 0}$  be the branching-coalescing point set generated by the discrete net  $N_\varepsilon$  with  $\xi_0 = \{0\}$ . Let  $\mathcal{R}_T$  be the number of  $(0, T)$ -RBP for  $N_\varepsilon(0, 0)$ . Then for any  $K \in \mathbb{N}$ , there exist  $C_K, \delta_0 = \delta_0(K), \varepsilon_0 = \varepsilon_0(K) > 0$  such that for all  $\varepsilon \leq \varepsilon_0$  and  $1 \leq T \leq \delta_0\varepsilon^{-2}$ ,*

$$\mathbb{P}(\mathcal{R}_T \geq K) \leq C_K (\varepsilon\sqrt{T})^K. \quad (6.18)$$

**Proof** We first outline the proof strategy, which starts with a decomposition according to the times  $t_1, \dots, t_K$  of the  $K$  earliest  $(0, T)$ -RBP's  $z_i = (x_i, t_i)$ ,  $1 \leq i \leq K$ . Assume for the moment the  $t_i$ 's are all distinct. Let us include half integer times in  $(\xi_t)_{t \geq 0}$  to keep track of the increments of the branching-coalescing random walks. We can construct  $(\xi_t)_{t \geq 0}$  through an exploration procedure as follows (see Figure 3). First sample a random walk  $S^0$  starting at  $(0, 0)$  and run till time  $t_1$ , where at each branching point of  $N_\varepsilon$  the walk encounters, it chooses one of the two outgoing jumps with equal probability. This generates a path in  $(\xi_t)_{t \geq 0}$ . Then  $z_1 = (t_1, S_{t_1}^0)$  must be the first  $(0, T)$ -RBP. In particular,  $z_1$  is a branching point, an event which we denote by  $B_1$  and has probability  $\varepsilon$ . We next sample two random walks  $S^{1,1}$  and  $S^{1,2}$  starting at  $z_1$ , whose first jumps are independent due to the binary branching at  $z_1$ , but then interact as coalescing random walks. The paths  $(S_t^{1,1})_{t_1 \leq t \leq t_2}$  and  $(S_t^{1,2})_{t_1 \leq t \leq t_2}$  are also paths in  $(\xi_t)_{t \geq 0}$ , and by Lemma 6.5,  $S^{1,1}$  and  $S^{1,2}$  cannot coalesce before or at time  $t_2$  because  $z_1$  is a  $(0, T)$ -relevant branching point. We denote this event by  $E_1$ , the probability of which can be bounded by  $C/\sqrt{t_2 - t_1}$  by standard random walk estimates. At time  $t_2$ , the next RBP  $z_2$  is either  $(t_2, S_{t_2}^{1,1})$  or  $(t_2, S_{t_2}^{1,2})$ . Once a choice has been made, that point

must be a branching point, an event which we denote by  $B_2$ . We then sample two random walks  $S^{2,1}$  and  $S^{2,2}$  starting at  $z_2$ , whose first jumps are independent and then interact as coalescing random walks (also coalesce with previously sampled random walks if they have not terminated as a result of reaching a RBP). Denote by  $E_2$  the event that  $S^{2,1}$  and  $S^{2,2}$  do not coalesce by time  $t_3$ . We then repeat this construction so that at each time  $t_k$ , there are at most  $k$  choices (the number of locations of the previously sampled coalescing random walks) for the RBP  $z_k$ , there is an event  $B_k$  for  $z_k$  to be a branching point, and there is an event  $E_k$  for two sampled random walks  $S^{k,1}$  and  $S^{k,2}$  starting from  $z_k$  not to coalesce by time  $t_{k+1}$ . The event for the time interval  $[t_K, T]$  will be different because  $t_{K+1}$  is unspecified, and we avoided a full decomposition of all  $(0, T)$ -RBP's because we are unable to control the growing combinatorial factor  $k$  associated with the choice of  $z_k$  at time  $t_k$ . Instead of restricting to the event  $E_K$  on the two sampled random walks  $S^{K,1}$  and  $S^{K,2}$  starting from  $z_K$ , we simply restrict to the event  $G(z_K; t_K, T)$  that there are two paths in  $N_\varepsilon(z_K)$  which are disjoint on the time interval  $(t_K, T]$ . Thanks to Proposition 6.7, we can control  $\mathbb{P}(G(z_K; t_K, T))$  averaged over  $t_K$ , which leads to a version of (6.18) with  $T$  therein replaced by  $T'$  averaged over  $[T, 2T]$ . A monotonicity argument then gives (6.18).

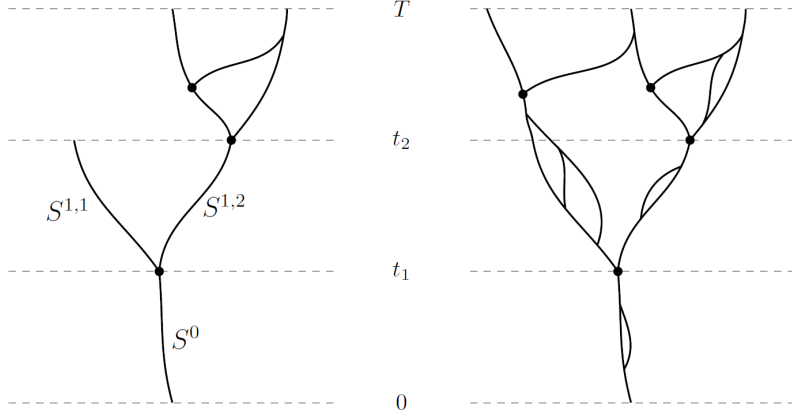


Figure 3: The image on the left illustrates the exploration procedure for  $\{\mathcal{R}_T \geq 2\}$ . The image on the right illustrates the branching-coalescing point set, with  $\bullet$  identifying the relevant branching points (RBP's).

We now make the above proof sketch more precise. First we restrict to the event  $D$  that the first  $K$  relevant branching points all occur at distinct times. Then by successively conditioning on the sampled coalescing random walks  $(S^0, S^{1,1}, S^{1,2}, \dots)$  up to time  $t_K, t_{K-1}, \dots$  and applying the Markov property, we have

$$\begin{aligned} \mathbb{P}(D \cap \{\mathcal{R}_T \geq K\}) &\leq \sum_{0 < t_1 < \dots < t_K < T} \prod_{k=1}^{K-1} (\varepsilon^k \mathbb{P}(E_k)) \cdot K \mathbb{P}(G(z_K; t_K, T)) \\ &= \varepsilon^{K-1} K! \sum_{0 \leq t_1 < \dots < t_K < T} \prod_{i=1}^{K-1} \mathbb{P}(E_i) \cdot \phi(T - t_K), \end{aligned} \tag{6.19}$$

where  $\phi(T - t_K) := \mathbb{P}(G(z_K; t_K, T))$  depends only on  $T - t_K$  by translation invariance.

To bound  $\mathbb{P}(E_k)$ , note that if  $S$  and  $S'$  are two independent random walks with kernel  $a(\cdot)$  starting at  $(0, 0)$ , and  $\tau^{S, S'} := \min\{t \geq 1 : S_t = S'_t\}$ . Since  $S' - S$  is a symmetric random walk with finite second moments, by a standard random walk hitting time estimate

(see e.g. [NRS05, Lemma 2.2]),

$$\mathbb{P}(E_k) = \mathbb{P}(\tau^{S,S'} > t_{k+1} - t_k) \leq \mathbb{E}\left[C \frac{|S_1 - S'_1|}{\sqrt{t_{k+1} - t_k}}\right] \leq \frac{C}{\sqrt{t_{k+1} - t_k}}. \quad (6.20)$$

Substitute this bound into (6.19), replace  $T$  by  $T'$  and average over  $T \leq T' \leq 2T$ , we obtain

$$\begin{aligned} \frac{1}{T} \sum_T \mathbb{P}(D \cap \{\mathcal{R}_{T'} \geq K\}) &\leq \varepsilon^{K-1} K! \left(\sum_{t=1}^{2T} \frac{C}{\sqrt{t}}\right)^{K-1} \sum_{t=1}^{2T} \phi(t) \\ &\leq K! (C\varepsilon\sqrt{T})^{K-1} \sum_{t=1}^{2T} \phi(t), \end{aligned} \quad (6.21)$$

where in the first inequality, the sum over  $t_1 \in (0, T)$  compensates the factor  $\frac{1}{T}$ . To bound the last sum in (6.21), note that for  $K = 1$ , we have a similar decomposition according to the time of the first relevant branching point:

$$\mathbb{P}(\mathcal{R}_{2T} \geq 1) = \sum_{t=0}^{2T-1} \mathbb{P}(|\xi_t| = 1) \phi(2T - t) \geq \frac{1}{2} \sum_{t=1}^{2T} \phi(t),$$

where we applied Proposition 6.7 to obtain  $\mathbb{P}(|\xi_t| = 1) \geq 1/2$ . Again by Proposition 6.7,

$$\sum_{t=1}^{2T} \phi(t) \leq 2\mathbb{P}(\mathcal{R}_{2T} \geq 1) \leq C\varepsilon\sqrt{T}. \quad (6.22)$$

Substituting this bound into (6.21) then gives

$$\frac{1}{T} \sum_T \mathbb{P}(D \cap \{\mathcal{R}_{T'} \geq K\}) \leq K! (C\varepsilon\sqrt{T})^K \leq C_K (\varepsilon\sqrt{T})^K, \quad (6.23)$$

which is an averaged version of (6.18).

Before removing the averaging in (6.23), let us first remove the restriction  $D$  that the first  $K$  relevant branching points all occur at distinct times. Assume  $D^c$ , so that  $|\{t_1, t_2, \dots, t_K\}| = M < K$ . Let us relabel these distinct times by  $0 \leq s_1 < \dots < s_M$ , with multiplicities  $r_1, \dots, r_M$  so that  $\sum r_i = K$ . We can apply the same argument as in the bound for  $\mathbb{P}(D \cap \{\mathcal{R}_T \geq M\})$ , except that at each time  $s_k$  with  $r_k > 1$ , we only restrict to the event  $E_k$  for a pair of sampled random walks starting from one of the  $r_k$  RBP's. However, the  $r_k$  RBP's at time  $s_k$  must still all be branching points, which contribute a probability factor of  $\varepsilon^{r_k}$  instead of  $\varepsilon$ . This argument gives

$$\begin{aligned} \frac{1}{T} \sum_T \mathbb{P}(D^c \cap \{\mathcal{R}_{T'} \geq K\}) &\leq \sum_{M=1}^{K-1} C_{K,M} \varepsilon^{K-M} (C\varepsilon\sqrt{T})^M \\ &= \sum_{M=1}^{K-1} C_{K,M} (C\sqrt{T})^{M-K} (C\varepsilon\sqrt{T})^K \leq C_K (\varepsilon\sqrt{T})^K, \end{aligned} \quad (6.24)$$

where  $C_{K,M}$  are combinatorial factors depending on  $K$  and  $M$ . Therefore (6.23) can be replaced by

$$\frac{1}{T} \sum_T \mathbb{P}(\mathcal{R}_{T'} \geq K) \leq C_K (\varepsilon\sqrt{T})^K. \quad (6.25)$$

To remove the averaging in  $T'$ , note that for any  $T' \geq T$ , the event  $\{\mathcal{R}_{T'} \geq K\}$  contains the intersection of the events that  $\xi_{T'-T} =: \{x\}$  is a singleton and the branching-coalescing

random walks starting from  $x$  at time  $T' - T$  has at least  $K$  relevant branching points for the time interval  $(T' - T, T')$ . Using the Markov property of  $(\xi_t^\varepsilon)_{t \geq 0}$ , we then have

$$\mathbb{P}(\mathcal{R}_{T'} \geq K) \geq \mathbb{P}(|\xi_{T'-T}| = 1) \mathbb{P}(\mathcal{R}_T \geq K) \geq \frac{1}{2} \mathbb{P}(\mathcal{R}_T \geq K), \quad (6.26)$$

where the last bound follows from Proposition 6.7. Substituting this bound into (6.25) then gives (6.18).  $\blacksquare$

## 6.5 A multiscale argument for tightness

We are now ready to verify the tightness condition (6.1) for the rescaled discrete nets  $S_{\varepsilon, \sigma} N_\varepsilon$  by adapting a multiscale argument from [BMSV06] for coalescing random walks. More precisely, we will prove the following:

$$\lim_{\delta \downarrow 0} \frac{1}{\delta} \limsup_{\varepsilon \downarrow 0} \sup_{(\varepsilon x, \varepsilon^2 t) \in \Lambda_{L, T}} \mathbb{P}(N_\varepsilon \in A_{M\varepsilon^{-1}, \delta\varepsilon^{-2}}(x, t)) = 0, \quad \text{for all } L, T, M > 0, \quad (6.27)$$

where  $A_{M\varepsilon^{-1}, \delta\varepsilon^{-2}}(x, t)$  is defined as in (6.1). By translation invariance, it suffices to show

$$\lim_{\delta \downarrow 0} \frac{1}{\delta} \limsup_{\varepsilon \downarrow 0} \mathbb{P}(N_\varepsilon \in A_{M\varepsilon^{-1}, \delta\varepsilon^{-2}}(0, 0)) = 0 \quad \text{for all } M > 0. \quad (6.28)$$

The key ingredients of the multiscale argument used in [BMSV06] are: density reduction through coalescence, and large deviation estimates for a single random walk. In our case, we need large deviation estimates for a random walk and all its branching descendants, which can be achieved thanks to our bounds on the number of relevant branching points (RBP's) in Theorem 6.9. This is the content of the following lemma, where the three terms in (6.29) originate respectively from large deviation from the sum of many small jumps, large deviation from rare large jumps, and excessive branching.

**Lemma 6.10 (Large deviation estimate)** *Let  $(\xi_t)_{t \geq 0}$  be the branching-coalescing point set generated by the discrete net  $N_\varepsilon$  with  $\xi_0 = \{0\}$ . Then for any  $K \in \mathbb{N}$ , there exist  $c_K, \delta_0, \varepsilon_0 > 0$  such that for all  $\varepsilon \leq \varepsilon_0$ ,  $\ell > 0$ , and  $1 \leq T \leq \delta_0 \varepsilon^{-2}$ ,*

$$\mathbb{P}\left(\sup_{0 \leq t \leq T} \max \xi_t \geq \ell\right) \leq e^{-\frac{\ell^2}{c_K T}} + c_K T \ell^{-\gamma} + c_K \varepsilon^K T^{K/2}, \quad (6.29)$$

where the jump distribution  $a(\cdot)$  satisfies (1.1) with  $\sum_x |x|^\gamma a(x) < \infty$  for some  $\gamma > 3$ .

**Proof** We first prove the deviation bound for  $\xi$  at the terminal time  $T$ , i.e.,

$$\mathbb{P}(\max \xi_T \geq \ell) \leq e^{-\frac{\ell^2}{c_K T}} + c_K T \ell^{-\gamma} + c_K \varepsilon^K T^{K/2}. \quad (6.30)$$

Let  $\mathcal{R}_T$  be the number of  $(0, T)$ -RBP for  $N_\varepsilon(0, 0)$ , defined as in (6.2). Then by Theorem 6.9,

$$\mathbb{P}(\mathcal{R}_T \geq K) \leq c_K \varepsilon^K T^{K/2}, \quad (6.31)$$

which gives the last term in (6.30).

On the complementary event  $\{\mathcal{R}_T \leq K - 1\}$ , we have  $|\xi_T| \leq K$  by the properties of RBP established in Section 6.2. Let us include half integer times in  $(\xi_t)_{t \geq 0}$  to keep track of the increments of the branching-coalescing random walks. We can then sample a random walk  $S$  in  $\xi$  such that  $S_0 = 0$ , and whenever  $S$  encounters a branching point of  $N_\varepsilon$ , it chooses one of the two outgoing jumps with equal probability. Clearly  $S$  has the law of a random walk with jump distribution  $a(\cdot)$ . On the other hand, the position  $S_T$  depends only on the choices made



by  $S$  at each  $(0, T)$ -RBP it encounters along the way. Therefore on the event  $\{\mathcal{R}_T \leq K - 1\}$ , each point in  $\xi_T$  (including  $\max \xi_T$ ) will coincide with  $S_T$  with probability at least  $1/2^{K-1}$ , and hence

$$\mathbb{P}(\max \xi_T \geq \ell, \mathcal{R}_T \leq K - 1, S_T \geq \ell) \geq \frac{1}{2^{K-1}} \mathbb{P}(\max \xi_T \geq \ell, \mathcal{R}_T \leq K - 1).$$

It follows that

$$\mathbb{P}(\max \xi_T \geq \ell, \mathcal{R}_T \leq K - 1) \leq 2^{K-1} \mathbb{P}(S_T \geq \ell) \leq e^{-\frac{\ell^2}{c_K T}} + c_K T \ell^{-\gamma}, \quad (6.32)$$

where the last inequality is a standard random walk large deviation estimate (see e.g. [N79, Corollary 1.8]).

Lastly, to show (6.29), one only needs to note that

$$\tau := \inf\{t \in \mathbb{N} : \max \xi_t \geq \ell\} \quad (6.33)$$

is a stopping time. On the event  $\tau \leq T$ , we can sample a random walk  $S$  in  $\xi$  starting from position  $\max \xi_\tau$  at time  $\tau$ , which then leads to  $\xi_T \ni S_T \geq \ell$  with a probability that is bounded away from 0 uniformly in  $T - \tau$  (by the central limit theorem). The bound (6.29) then follows from (6.30) with an adjusted value of  $c_K$ .  $\blacksquare$

We are now ready to prove the tightness of the family of rescaled discrete nets.

**Theorem 6.11 (Tightness)** *The family of rescaled discrete nets  $(S_{\varepsilon, \sigma} N_\varepsilon)_{\varepsilon \in (0, 1)}$  in Theorem 1.1 is tight.*

**Proof** As discussed above, we need to verify the condition (6.28) for all  $M > 0$ . As we will see, the constant  $M$  enters the proof only through the large deviation estimate in Lemma 6.10 when we take  $\ell = cM\varepsilon^{-1}$  for constants  $c > 0$  independent of  $M$ . For any fixed  $M > 0$ , the last two terms in the right hand side of (6.29) will always be of order  $o(\varepsilon^3)$ , while the first term will always be of order  $o(1/\delta)$  when we take  $T = \delta\varepsilon^{-2}$  with  $\delta \downarrow 0$ . These are the only bounds we need. To simplify the notation, we may therefore assume that  $M = 10$  and  $\sigma^2 = 1$ . We also note that to bound  $\mathbb{P}(N_\varepsilon \in A_{M\varepsilon^{-1}, \delta\varepsilon^{-2}}(x, t))$  in (6.28), we can discard the event that there are random walk increments in  $N_\varepsilon$  that jump across the interval  $[-M\varepsilon^{-1}, M\varepsilon^{-1}]$  in the time interval  $[0, \delta\varepsilon^{-2}]$ , so that we only need to restrict our attention to random walks that originate inside the space-time box  $[-M\varepsilon^{-1}, M\varepsilon^{-1}] \times [0, \delta\varepsilon^{-2}]$ . The same issue arises in the proof of tightness for the family of rescaled coalescing random walks, and the proof is identical since the branching at most doubles the number of random walk increments, see [NRS05, Section 4].

We now apply the multiscale argument as in the proof of [BMSV06, Proposition 2.3], where for convenience we actually consider all random walks originating from a larger space-time box  $[-M\varepsilon^{-1}, M\varepsilon^{-1}] \times [0, 2\delta\varepsilon^{-2}]$ . In the first step, we choose a large fixed constant  $R > 0$  independent of  $\varepsilon > 0$  (to be determined below Lemma 6.12) and decompose the space-time box  $[-M\varepsilon^{-1}, M\varepsilon^{-1}] \times [0, 2\delta\varepsilon^{-2}]$  in the time direction into  $I_0 := 2\delta\varepsilon^{-2}/R$  boxes of the form

$$B_{0,i} := [-M\varepsilon^{-1}, M\varepsilon^{-1}] \times [(i-1)R, iR], \quad i = 1, \dots, I_0. \quad (6.34)$$

To simplify notation later, we may assume  $I_0 = 2^J$  for some  $J \in \mathbb{N}$ ; otherwise we can increase  $\delta$  and  $I_0$  by a factor of at most 2 such that this property holds. For the branching-coalescing random walks in  $N_\varepsilon$  starting from each space-time box  $B_{0,i}$ , we let them evolve up to time  $T_i^{(0)} := iR$  and discard the event  $E_{0,i}$  that some particle exits the interval

$$[-\widetilde{M}_0, \widetilde{M}_0] := [-(M+1)\varepsilon^{-1}, (M+1)\varepsilon^{-1}]$$

before time  $T_i^{(0)}$ . Indeed, by the large deviation estimate in Lemma 6.10 with  $K = 4$  and a union bound over all starting points in  $\cup_{i=1}^{I_0} B_{0,i}$ , we have

$$\mathbb{P}(\cup_{i=1}^{I_0} E_{0,i}) \leq 4M\delta\varepsilon^{-3} \left( e^{-\frac{cM^2\varepsilon^{-2}}{R}} + cRM^{-\gamma}\varepsilon^\gamma + c\varepsilon^4 R^2 \right), \quad (6.35)$$

where  $\gamma = 3 + \eta > 3$  by the assumption on the jump distribution  $a(\cdot)$  in (1.1). Denote  $E_0 := \cup_{i=1}^{I_0} E_{0,i}$ . Given  $M$  and  $R$ , clearly

$$\lim_{\delta \downarrow 0} \frac{1}{\delta} \limsup_{\varepsilon \downarrow 0} \mathbb{P}(E_0) = 0. \quad (6.36)$$

Therefore to verify (6.28), we can discard the event  $E_0$ , which ensures that no walk has exited  $[-\widetilde{M}_0, \widetilde{M}_0]$  in this first step, and at the end of this step, all remaining walks lie in  $[-\widetilde{M}_0, \widetilde{M}_0]$  at times  $T_i^{(0)} = iR$ ,  $1 \leq i \leq I_0$ , whose configurations we denote by  $\xi_i^{(0)} \subset [-\widetilde{M}_0, \widetilde{M}_0]$ .

In the second step, let  $I_1 := I_0/2$  and let  $T_i^{(1)} := T_{2i}^{(0)} = 2iR$  for  $1 \leq i \leq I_1$ . For each  $1 \leq i \leq I_1$ , we consider the remaining random walks starting from  $\xi_{2i-2}^{(0)} \times \{T_{2i-2}^{(0)}\}$  and  $\xi_{2i-1}^{(0)} \times \{T_{2i-1}^{(0)}\}$  and let them evolve together till time  $T_i^{(1)} = T_{2i}^{(0)}$ , and denote the resulting particle configuration at time  $T_i^{(1)}$  by  $\xi_i^{(1)}$ . We will show that we can also discard the events  $E_{1,i} := D_{1,i} \cup F_{1,i}$ ,  $1 \leq i \leq I_1$ , where

$D_{1,i}$ : Some random walk starting from  $\xi_{2i-2}^{(0)}$  at time  $T_{2i-2}^{(0)}$ , or from  $\xi_{2i-1}^{(0)}$  at time  $T_{2i-1}^{(0)}$ , exits the interval  $[-\widetilde{M}_1, \widetilde{M}_1]$  before time  $T_i^{(1)} = T_{2i}^{(0)}$ , where  $\widetilde{M}_1 := \widetilde{M}_0 + \beta_1\varepsilon^{-1}$  for some  $\beta_1$  to be chosen.

$F_{1,i}$ : The particle density  $\frac{1}{4M\varepsilon^{-1}} |\xi_i^{(1)}| > p := \frac{1}{\sqrt{2}}$ .

Let us denote  $E_1 := \cup_{i=1}^{I_1} E_{1,i}$ . Similar to (6.36), we will show that

$$\lim_{\delta \downarrow 0} \frac{1}{\delta} \limsup_{\varepsilon \downarrow 0} \mathbb{P}(E_1 \cap E_0^c) = 0, \quad (6.37)$$

which allows us to discard the event  $E_1$  in the verification of (6.28).

We now iterate the above argument. After the  $(j-1)$ -st iteration for some  $j \in \mathbb{N}$ , where we have discarded the events  $E_0 \cup \dots \cup E_{j-1}$ , we are left with random walk configurations  $\xi_i^{(j-1)}$  at times  $T_i^{(j-1)} = T_{2i}^{(j-2)} = i \cdot 2^{j-1}R$  for  $1 \leq i \leq I_{j-1} := 2^{-(j-1)}I_0$ , such that

$$\xi_i^{(j-1)} \subset [-\widetilde{M}_{j-1}, \widetilde{M}_{j-1}] \quad \text{and} \quad \frac{1}{4M\varepsilon^{-1}} |\xi_i^{(j-1)}| \leq p^{j-1} = 2^{-(j-1)/2}, \quad (6.38)$$

where  $\widetilde{M}_{j-1} = \widetilde{M}_0 + (\beta_1 + \dots + \beta_{j-1})\varepsilon^{-1}$ . Note that we choose  $p = 1/\sqrt{2}$  because by diffusive scaling, this is exactly the rate at which the density of coalescing random walks on  $\mathbb{Z}$  decays when time is doubled. This allows us to show inductively in  $j$  that the events in (6.38) occur with high probability.

To define the  $j$ -th iteration (see Figure 4), for each  $1 \leq i \leq I_j := 2^{-j}I_0$ , we consider the remaining random walks starting from  $\xi_{2i-2}^{(j-1)} \times \{T_{2i-2}^{(j-1)}\}$  and  $\xi_{2i-1}^{(j-1)} \times \{T_{2i-1}^{(j-1)}\}$  and let them evolve together till time  $T_i^{(j)} := T_{2i}^{(j-1)} = i \cdot 2^j R$ , and denote the resulting particle configuration at time  $T_i^{(j)}$  by  $\xi_i^{(j)}$ . For  $1 \leq i \leq I_j$ , define the events

$D_{j,i}$ : Some random walk starting from  $\xi_{2i-2}^{(j-1)}$  at time  $T_{2i-2}^{(j-1)}$ , or from  $\xi_{2i-1}^{(j-1)}$  at time  $T_{2i-1}^{(j-1)}$ , exits the interval  $[-\widetilde{M}_j, \widetilde{M}_j]$  before time  $T_i^{(j)} = T_{2i}^{(j-1)}$ , where  $\widetilde{M}_j := \widetilde{M}_{j-1} + \beta_j\varepsilon^{-1}$  for some  $\beta_j > 0$  to be chosen.

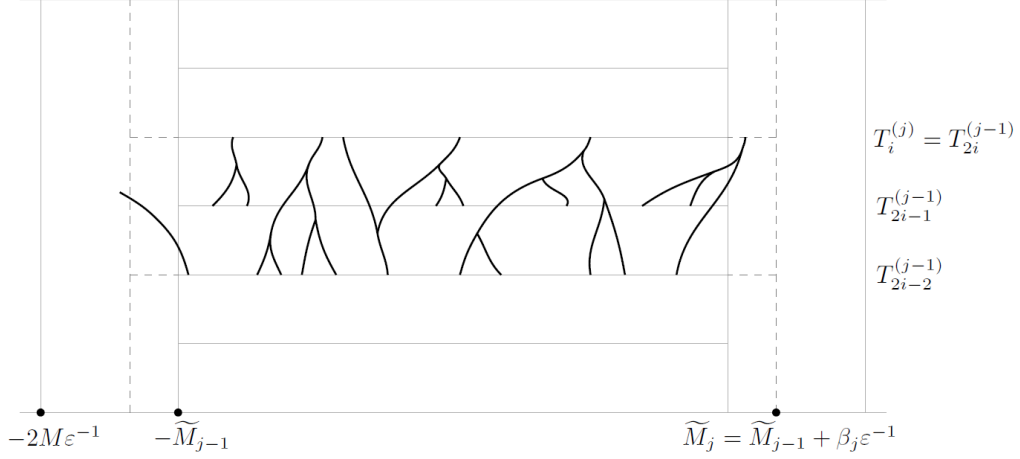


Figure 4: Illustration of the  $j$ -th iteration of the multiscale argument.

$F_{j,i}$ : The particle density  $\frac{1}{4M\varepsilon^{-1}}|\xi_i^{(j)}| > p^j = 2^{-j/2}$ .

As before, we denote  $E_{j,i} := D_{j,i} \cup F_{j,i}$ , and  $E_j := \cup_{i=1}^{I_j} E_{j,i}$ .

By the assumption after (6.34),  $I_0 = 2^J$  for some  $J \in \mathbb{N}$ . After  $J$  iterations, all walks starting from the box  $[-M\varepsilon^{-1}, M\varepsilon^{-1}] \times [0, 2\delta\varepsilon^{-2}]$  have evolved till time  $2\delta\varepsilon^{-2}$ . To verify (6.28) and bound the probability of the event that some random walk exits the interval  $[-2M\varepsilon^{-1}, 2M\varepsilon^{-1}]$  before time  $2\delta\varepsilon^{-2}$ , it suffices to choose  $(\beta_j)_{1 \leq j \leq J}$  such that  $\widetilde{M}_J < 2M\varepsilon^{-1}$ , and we can discard  $\cup_{j=0}^J E_j$  in the sense that

$$\lim_{\delta \downarrow 0} \frac{1}{\delta} \limsup_{\varepsilon \downarrow 0} \mathbb{P}(\cup_{j=0}^J E_j) = 0,$$

which would follow if we show that

$$\lim_{\delta \downarrow 0} \frac{1}{\delta} \limsup_{\varepsilon \downarrow 0} \sum_{j=0}^J \mathbb{P}(E_j \cap \cap_{k=0}^{j-1} E_k^c) = 0. \quad (6.39)$$

We will verify (6.39) for  $\beta_j := \frac{1}{(J+1-j)^2}$ , a choice that ensures

$$\widetilde{M}_J = \widetilde{M}_0 + \varepsilon^{-1} \sum_{j=1}^J \beta_j < 2M\varepsilon^{-1}.$$

For  $1 \leq j \leq J$ , we can write  $E_j = D_j \cup F_j$ , where  $D_j := \cup_{i=1}^{I_j} D_{j,i}$  and  $F_j := \cup_{i=1}^{I_j} F_{j,i}$ . Since (6.36) already gives the desired bound on  $\mathbb{P}(E_0)$ , to show (6.39), we will bound  $\mathbb{P}(D_j \cap \cap_{k=0}^{j-1} E_k^c)$  and  $\mathbb{P}(F_j \cap \cap_{k=0}^{j-1} E_k^c)$  for  $1 \leq j \leq J$ .

We first consider each event  $D_{j,i} \subset D_j = \cup_{i=1}^{I_j} D_{j,i}$ , which concerns the random walks starting from  $\xi_{2i-2}^{(j-1)} \times \{T_{2i-2}^{(j-1)}\}$  and  $\xi_{2i-1}^{(j-1)} \times \{T_{2i-1}^{(j-1)}\}$ . Restricted to the event  $\cap_{k=0}^{j-1} E_k^c$ , the particle configurations  $\xi_{2i-2}^{(j-1)}$  and  $\xi_{2i-1}^{(j-1)}$  satisfy the conditions in (6.38). For the event  $D_{j,i}$  to occur, one of the random walks starting from either  $\xi_{2i-2}^{(j-1)}$  at time  $T_{2i-2}^{(j-1)}$ , or from  $\xi_{2i-1}^{(j-1)}$  at time  $T_{2i-1}^{(j-1)}$ , must travel a distance of at least  $\beta_j \varepsilon^{-1}$  within a time interval of length  $2^j R$ . Therefore by conditioning on  $\xi_{2i-2}^{(j-1)}$ , resp.,  $\xi_{2i-1}^{(j-1)}$ , and performing a union bound and applying

Lemma 6.10 with  $\ell = \beta_j \varepsilon^{-1}$  and  $T = 2^j R$ , we obtain

$$\begin{aligned}
\mathbb{P}(D_j \cap \bigcap_{k=0}^{j-1} E_k^c) &\leq \sum_{i=1}^{I_j} \mathbb{P}(D_{j,i} \cap \bigcap_{k=0}^{j-1} E_k^c) \\
&\leq I_j 2^{2-\frac{j-1}{2}} M \varepsilon^{-1} \left( \exp \left\{ -\frac{c \beta_j^2 \varepsilon^{-2}}{2^j R} \right\} + c 2^j R (\beta_j \varepsilon^{-1})^{-\gamma} + c \varepsilon^4 2^{2j} R^2 \right) \\
&\leq 2^{3-\frac{3}{2}(j-1)} \frac{M \delta}{R \varepsilon^3} \left( \exp \left\{ -\frac{c \beta_j^2 \varepsilon^{-2}}{2^j R} \right\} + c 2^j R (\beta_j \varepsilon^{-1})^{-\gamma} + c \varepsilon^4 2^{2j} R^2 \right) \\
&=: 8(B_j^1 + B_j^2 + B_j^3), \tag{6.40}
\end{aligned}$$

where we have expanded the product.

Note that

$$\sum_{j=1}^J B_j^3 = c \sum_{j=1}^J 2^{\frac{j+3}{2}} M R \delta \varepsilon \leq C 2^{\frac{J}{2}} M R \delta \varepsilon \leq C \sqrt{\frac{\delta}{\varepsilon^2 R}} M R \delta \varepsilon = C M R^{\frac{1}{2}} \delta^{\frac{3}{2}},$$

where we used that  $I_0 = 2\delta\varepsilon^{-2}/R = 2^J$ . Therefore

$$\limsup_{\delta \downarrow 0} \frac{1}{\delta} \limsup_{\varepsilon \downarrow 0} \sum_{j=1}^J B_j^3 = 0. \tag{6.41}$$

Also note that because  $\beta_j \geq \beta_1 = J^{-2}$ ,

$$\sum_{j=1}^J B_j^2 \leq C \sum_{j=1}^J 2^{-\frac{j}{2}} \beta_j^{-\gamma} M \delta \varepsilon^{\gamma-3} \leq C M \delta \varepsilon^{\gamma-3} \beta_1^{-\gamma} \sum_{j=1}^J 2^{-\frac{j}{2}} \leq C M \delta \varepsilon^{\gamma-3} J^{2\gamma}.$$

Since  $J = \log_2 I_0 = \log_2 \frac{2\delta\varepsilon^{-2}}{R} = o(\varepsilon^{\frac{3-\gamma}{2\gamma}})$  as  $\varepsilon \downarrow 0$  because  $\gamma > 3$ , we have

$$\limsup_{\delta \downarrow 0} \frac{1}{\delta} \limsup_{\varepsilon \downarrow 0} \sum_{j=1}^J B_j^2 = 0. \tag{6.42}$$

Lastly using  $\beta_j := \frac{1}{(J+1-j)^2}$  and  $2^J = I_0 = \frac{2\delta\varepsilon^{-2}}{R}$ , we note that

$$\begin{aligned}
\sum_{j=1}^J B_j^1 &\leq C \frac{M \delta}{R \varepsilon^3} \sum_{j=1}^J 2^{-\frac{3j}{2}} e^{-\frac{c \beta_j^2}{2^j R \varepsilon^2}} \\
&= C \frac{M \delta}{R \varepsilon^3} \sum_{j=1}^J 2^{-\frac{3j}{2}} e^{-\frac{c}{2^j (J+1-j)^2 R \varepsilon^2}} = C \frac{M \delta}{R \varepsilon^3} 2^{-\frac{3(J+1)}{2}} \sum_{\tilde{j}=1}^J 2^{\frac{3\tilde{j}}{2}} e^{-\frac{c}{2^{J+1-\tilde{j}} \tilde{j}^2 R \varepsilon^2}} \\
&\leq C M R^{\frac{1}{2}} \delta^{-\frac{1}{2}} \sum_{\tilde{j}=1}^J 2^{\frac{3\tilde{j}}{2}} e^{-c \frac{2^{\tilde{j}-1}}{\tilde{j}^2} \cdot \frac{1}{\delta}} \leq C M R^{\frac{1}{2}} \delta^{-\frac{1}{2}} e^{-\frac{c}{2\delta}} \sum_{\tilde{j}=1}^J 2^{\frac{3\tilde{j}}{2}} e^{-c \left( \frac{2^{\tilde{j}-1}}{\tilde{j}^2} - \frac{1}{2} \right) \frac{1}{\delta}},
\end{aligned}$$

where we note that the last sum is uniformly bounded in  $J$  and in  $\delta > 0$  sufficiently small, because  $c$  comes from Lemma 6.10 and does not depend on  $M, R, \delta$  and  $\varepsilon$ . It follows that

$$\limsup_{\delta \downarrow 0} \frac{1}{\delta} \limsup_{\varepsilon \downarrow 0} \sum_{j=1}^J B_j^1 = 0. \tag{6.43}$$

Combining (6.41)-(6.43) gives

$$\lim_{\delta \downarrow 0} \frac{1}{\delta} \limsup_{\varepsilon \downarrow 0} \sum_{j=0}^J \mathbb{P}(D_j \cap \bigcap_{k=0}^{j-1} E_k^c) = 0, \quad (6.44)$$

which accounts for the contribution of  $D_j \subset E_j$  in (6.39).

We now consider the events  $F_j = \cup_{i=1}^{I_j} F_{j,i}$  on the density of the branching-coalescing random walks. We will need the following result, which generalises [BMSV06, Prop. 5.4] on the density of purely coalescing random walks. Roughly speaking, it says that if the initial density of the random walks is  $v \in (0, 1)$  so that the average distance between neighbouring walks is of order  $v^{-1}$ , then after evolving for a time  $Rv^{-2}$ , the density of the branching-coalescing random walks would have decreased by a fixed factor depending on  $R$ .

**Lemma 6.12 (Density decay of branching-coalescing random walks)** *Let  $p \in (1/2, 1)$  be fixed. Let  $(\xi_t)_{t \geq 0}$  be a branching-coalescing point set generated by the discrete net  $N_\varepsilon$  with initial condition  $\xi_0 \subset [0, L]$ . Then there exist  $R_0, \varepsilon_0, \delta_0, C, c > 0$  such that for all  $\varepsilon \in (0, \varepsilon_0)$ , if  $|\xi_0| \leq vL$  for some  $v \in (0, 1)$  and  $T := R_0 v^{-2} \leq 2\delta_0 \varepsilon^{-2}$ , then*

$$\mathbb{P}(|\xi_T| \geq pvL) \leq Ce^{-cvL} + CvL\varepsilon^4 T^2. \quad (6.45)$$

We are now ready to choose  $R$  in (6.34). Let  $R_0$  be as in Lemma 6.12 for  $p = 1/\sqrt{2}$ , and set  $R := 7R_0$ .

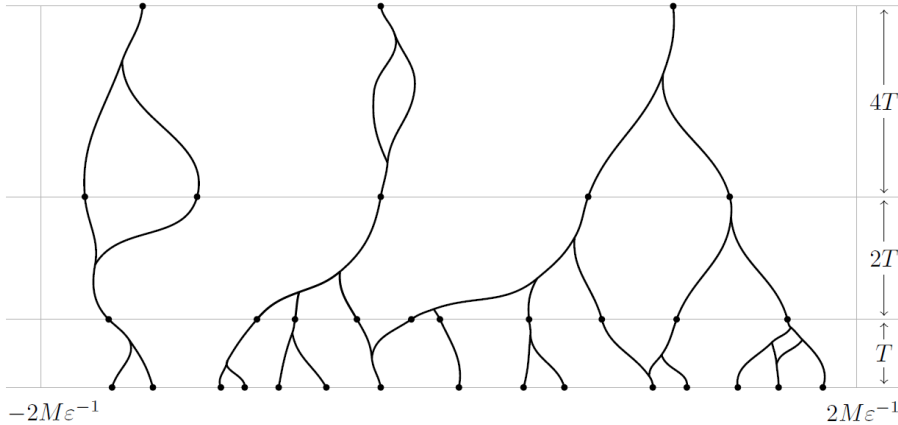


Figure 5: For random walks starting from  $\xi_{2i-1}^{(j-1)}$  at time  $T_{2i-1}^{(j-1)}$ , Lemma 6.12 is applied three times with  $p = 1/\sqrt{2}$  and successive time durations  $T = 2^{j-1}R_0$ ,  $2T$ , and  $4T$ .

Assuming Lemma 6.12 for the moment, we now bound  $\mathbb{P}(F_{j,i} \cap \bigcap_{k=0}^{j-1} E_k^c)$  for  $1 \leq j \leq J$  and  $1 \leq i \leq I_j$ . Note that restricted to the event  $\bigcap_{k=0}^{j-1} E_k^c$ , the particle configurations  $\xi_{2i-1}^{(j-1)}$  and  $\xi_{2i-2}^{(j-1)}$  satisfy the conditions in (6.38). To bound  $\mathbb{P}(F_{j,i} \cap \bigcap_{k=0}^{j-1} E_k^c)$ , we first apply Lemma 6.12 with

$$p = 1/\sqrt{2}, \quad L = 4M\varepsilon^{-1}, \quad v = 2^{-\frac{j-1}{2}} = p^{j-1}, \quad T = R_0 v^{-2} = 2^{j-1}R_0$$

to the random walks starting from  $\xi_{2i-2}^{(j-1)}$  at time  $T_{2i-2}^{(j-1)}$ , which implies that

$$\begin{aligned} \mathbb{P}(|\xi_{2^{j-1}R_0}| \geq p^j L \mid \xi_0 = \xi_{2i-2}^{(j-1)}) &\leq Ce^{-cp^{j-1}L} + Cp^{j-1}L\varepsilon^4 2^{2(j-1)} R_0^2 \\ &\leq Ce^{-cp^j \varepsilon^{-1}} + Cp^{-3j} \varepsilon^3, \end{aligned} \quad (6.46)$$

where we have adjusted the values of  $C$  and  $c$  to absorb factors of  $M, R_0$  and  $p$ . Thanks to (6.44), we may restrict to the event that the random walks starting from  $\xi_{2i-1}^{(j-1)}$  at time  $T_{2i-1}^{(j-1)}$  is contained in  $(-2M\varepsilon^{-1}, 2M\varepsilon^{-1})$  for a time duration of  $2^{j-1}R = 2^{j-1} \cdot 7R_0$ . Therefore, after running the walks for a time duration of  $T = 2^{j-1}R_0$ , which reduces the density by a factor of  $p$ , we can evolve the random walks further for a time duration of  $2T$  and then  $4T$  and apply Lemma 6.12 each time (see Figure 5), so that the density of the walks in  $(-2M\varepsilon^{-1}, 2M\varepsilon^{-1})$  at time  $T_{2i}^{(j-1)} = T_{2i-1}^{(j-1)} + 7T = i \cdot 2^j R$  is bounded by  $p^{j+2}$  (the probability of the complementary event is bounded by the r.h.s. of (6.46) with adjusted values of  $C$  and  $c$ ). Similarly, for the walks starting at  $\xi_{2i-2}^{(j-1)}$  at time  $T_{2i-2}^{(j-1)}$ , we can first evolve them till time  $T_{2i-1}^{(j-1)}$  and then till time  $T_i^{(j)} = T_{2i}^{(j-1)}$ , and we can apply Lemma 6.12 to conclude that the density of these walks is also bounded by  $p^{j+2}$  with high probability. Taking a union bound over the walks starting from  $\xi_{2i-2}^{(j-1)} \times T_{2i-2}^{(j-1)}$  and from  $\xi_{2i-1}^{(j-1)} \times T_{2i-1}^{(j-1)}$ , we conclude that the walks that remain at time  $T_i^{(j)}$  have density at most  $2p^{j+2} = p^j$  in the interval  $(-2M\varepsilon^{-1}, 2M\varepsilon^{-1})$ , which is exactly the event  $F_{j,i}^c$ . Therefore using the bound in (6.46), we have

$$\sum_{j=1}^J \mathbb{P}(F_j \cap \bigcap_{k=0}^{j-1} E_k^c) \leq C \sum_{j=1}^J I_j (e^{-cp^j \varepsilon^{-1}} + p^{-3j} \varepsilon^3) \leq \frac{C\delta}{R\varepsilon^2} \sum_{j=1}^J 2^{-j} (e^{-cp^j \varepsilon^{-1}} + p^{-3j} \varepsilon^3). \quad (6.47)$$

Writing the above as two sums and using  $2^J = I_0 = \frac{2\delta\varepsilon^{-2}}{R}$ , we can bound the first sum by

$$\begin{aligned} & \frac{C\delta}{R\varepsilon^2} \sum_{j=1}^J 2^{-j} e^{-cp^j \varepsilon^{-1}} = \frac{C\delta}{R\varepsilon^2} \sum_{j=1}^J p^{2j} e^{-cp^j \varepsilon^{-1}} \\ & \leq \frac{C\delta}{R\varepsilon^2} \cdot p^{2J} \sum_{\tilde{j}=1}^J p^{-2\tilde{j}} e^{-cp^{J-\tilde{j}} \varepsilon^{-1}} \leq C e^{-cp^J \varepsilon^{-1}} \sum_{\tilde{j}=1}^J 2^{\tilde{j}} e^{-cp^J \varepsilon^{-1} (p^{-\tilde{j}} - 1)} \leq C e^{-\frac{c}{\sqrt{\delta}}} \sum_{\tilde{j}=1}^J 2^{\tilde{j}} e^{-\frac{c}{\sqrt{\delta}} (2^{\tilde{j}/2} - 1)}, \end{aligned}$$

where the last sum is uniformly bounded in  $J$  and in  $\delta > 0$  small.

The second sum from (6.47) can be bounded by

$$\frac{C\delta}{R\varepsilon^2} \sum_{j=1}^J 2^{-j} p^{-3j} \varepsilon^3 = \frac{C\delta\varepsilon}{R} \sum_{j=1}^J p^{-j} \leq \frac{C\delta\varepsilon}{R} \cdot 2^{\frac{J}{2}} \leq \frac{C\delta\varepsilon}{R} \frac{\sqrt{\delta}}{\varepsilon\sqrt{R}} \leq C\delta^{3/2}.$$

Substituting the above bounds into (6.47) then gives

$$\limsup_{\delta \downarrow 0} \frac{1}{\delta} \limsup_{\varepsilon \downarrow 0} \sum_{j=1}^J \mathbb{P}(F_j \cap \bigcap_{k=0}^{j-1} E_k^c) = 0. \quad (6.48)$$

Together with (6.44), this concludes the proof of (6.39) and Theorem 6.11.  $\blacksquare$

It remains to prove Lemma 6.12.

**Proof of Lemma 6.12** By adding new walks if necessary, we can assume that we start with exactly  $\nu L$  walks at positions  $0 \leq x_1 < x_2 < \dots < x_{\nu L} \leq L$ . Let  $\mathcal{R}_T^{x_k}$  be the number of  $(0, T)$ -relevant branching points of  $N_\varepsilon(x_k, 0)$ , and let  $\mathcal{R}_T := \max_{1 \leq k \leq \nu L} \mathcal{R}_T^{x_k}$ . First note that by Theorem 6.9 with  $K = 4$ , we have

$$\mathbb{P}(\mathcal{R}_T \geq 4) \leq C\nu L \varepsilon^4 T^2, \quad (6.49)$$

which is the source of the second term in the bound in (6.45).

Note that for any  $L_0 \in \mathbb{N}$  (to be chosen later), the number of consecutive pairs  $(x_k, x_{k+1})$  with  $x_{k+1} - x_k > L_0$  is at most  $\frac{L}{L_0}$ . Dividing  $(x_i, \dots, x_j, x_{j+1}, \dots, x_k)$  into two blocks

$(x_i, \dots, x_j)$  and  $(x_{j+1}, \dots, x_k)$  whenever  $x_{j+1} - x_j > L_0$ , we can divide  $(x_1, \dots, x_{vL})$  into at most  $1 + L/L_0$  blocks. In each block, we can match adjacent  $(x_i, x_{i+1})$  into pairs, which leaves at most one unmatched particle. This gives at least  $\kappa := \frac{1}{2}(vL - 1 - L/L_0)$  disjoint pairs of particles where each pair is within distance  $L_0$  of each other. For such a pair  $(x, y)$  with  $0 < y - x \leq L_0$ , if we consider the branching-coalescing point sets  $\xi^x$  and  $\xi^y$  generated by the same discrete net  $N_\varepsilon$  with initial conditions  $\xi_0^x = \{x\}$  and  $\xi_0^y = \{y\}$ , then we have

$$\begin{aligned} \mathbb{P}(|\xi_T^x \cup \xi_T^y| = 1) &\geq \mathbb{P}(\mathcal{R}_T^x = \mathcal{R}_T^y = 0, \xi_T^x = \xi_T^y) \\ &= \mathbb{P}(\mathcal{R}_T^x = \mathcal{R}_T^y = 0) - \mathbb{P}(\mathcal{R}_T^x = \mathcal{R}_T^y = 0, \xi_T^x \neq \xi_T^y). \end{aligned} \quad (6.50)$$

On the event  $\{\mathcal{R}_T^x = \mathcal{R}_T^y = 0\}$ ,  $|\xi_T^x| = |\xi_T^y| = 1$ . Therefore by sampling the random walk jumps at branching points of  $N_\varepsilon$ , we can couple  $\xi^x$  and  $\xi^y$  with two coalescing random walks  $S^x$  and  $S^y$  such that  $\xi_T^x = \{S_T^x\}$  and  $\xi_T^y = \{S_T^y\}$ . This implies that  $\{\xi_T^x \neq \xi_T^y\} = \{S_T^x \neq S_T^y\}$  and

$$\mathbb{P}(\mathcal{R}_T^x = \mathcal{R}_T^y = 0, \xi_T^x \neq \xi_T^y) \leq \mathbb{P}(\tau_{x,y} > T) \leq C \frac{L_0}{\sqrt{T}}, \quad (6.51)$$

where  $\tau_{x,y}$  is the first collision time between  $S^x$  and  $S^y$ , and the last inequality is a standard estimate, see e.g. [NRS05, Lemma 2.2].

On the other hand, by Theorem 6.9, we have

$$\mathbb{P}(\mathcal{R}_T^x = \mathcal{R}_T^y = 0) \geq 1 - \sum_{w=x,y} \mathbb{P}(\mathcal{R}_T^w \geq 1) \geq 1 - 2C\varepsilon\sqrt{T}. \quad (6.52)$$

Substituting the above bounds in to (6.50) then gives

$$\mathbb{P}(|\xi_T^x \cup \xi_T^y| = 1) \geq 1 - 2C\varepsilon\sqrt{T} - C \frac{L_0}{\sqrt{T}} =: 1 - p_0. \quad (6.53)$$

Let us denote the  $\kappa = \frac{1}{2}(vL - 1 - L/L_0)$  pairs of particles by  $(u_i, v_i)_{1 \leq i \leq \kappa}$ , with  $v_i \in (u_i, u_i + L_0]$ . We observe that the branching coalescing point set  $\xi^{\vec{u}, \vec{v}}$  generated by the discrete net  $N_\varepsilon$  with  $\xi_0^{\vec{u}, \vec{v}} = \{u_1, v_1, \dots, u_\kappa, v_\kappa\}$  can be coupled to  $\kappa$  independent branching-coalescing point sets  $\xi^{u_i, v_i}$ ,  $1 \leq i \leq \kappa$ , with  $\xi_0^{u_i, v_i} = \{u_i, v_i\}$ , such that almost surely,

$$|\xi_T^{\vec{u}, \vec{v}}| \leq \sum_{i=1}^{\kappa} |\xi_T^{u_i, v_i}|. \quad (6.54)$$

This is an analogue of [BMSV06, Lemma 5.3] for coalescing random walks and can be shown as follows. First sample a realisation of the jointly independent branching-coalescing random walks associated with  $(\xi^{u_i, v_i})_{1 \leq i \leq \kappa}$ , keeping track of the random walk jumps by defining  $\xi_t^{u_i, v_i}$  at the half-integer times  $t \in \mathbb{N}_0 + 1/2$ . We then construct  $\xi^{\vec{u}, \vec{v}}$  by recursively defining the branching-coalescing random walks starting from  $(u_i, v_i)_{1 \leq i \leq \kappa}$ . For  $i = 1$ , we construct the branching-coalescing random walks starting from  $\{u_1, v_1\}$  by following exactly  $\xi^{u_1, v_1}$ . We then add random walk paths starting from  $\{u_2, v_2\}$  by following paths in  $\xi^{u_2, v_2}$ , except when a random walk path meets an existing walk, it coalesces with that walk and discards its own descendants. The same procedure is then iterated by adding random walk paths in  $\xi^{u_3, v_3}, \dots, \xi^{u_\kappa, v_\kappa}$ . Clearly (6.54) holds, while the  $\xi^{\vec{u}, \vec{v}}$  constructed this way has the correct distribution because our procedure recursively explores the realisation of the random walk branching and increments at different lattice sites.

For  $\xi$  with  $\xi_0 = \{x_1, \dots, x_{vL}\}$ , we can now use (6.54) to bound

$$\begin{aligned} \mathbb{P}(|\xi_T| \geq pvL) &\leq \mathbb{P}(\mathcal{R}_T \geq 4) + \mathbb{P}(|\xi_T| \geq pvL, \mathcal{R}_T \leq 3) \\ &\leq CvL\varepsilon^4 T^2 + \mathbb{P}(|\xi_T^{\vec{u}, \vec{v}}| + 4(1 + L/L_0) \geq pvL) \\ &\leq CvL\varepsilon^4 T^2 + \mathbb{P}\left(\sum_{i=1}^{\kappa} |\xi_T^{u_i, v_i}| \geq pvL - 4(1 + L/L_0)\right), \end{aligned} \quad (6.55)$$

where we used the bound (6.49) and the fact that for each  $x_i \notin \{u_1, v_1, \dots, u_\kappa, v_\kappa\}$ , the event  $\{\mathcal{R}_T^{x_i} \leq 3\}$  implies that  $|\xi_T^{x_i}| \leq 4$ , and there are at most  $1 + L/L_0$  many such  $x_i$ 's.

We now choose  $L_0 = (\alpha v)^{-1}$  with  $\alpha > 0$  small and depending only on  $p \in (1/2, 1)$  such that

$$pvL - 4(1 + L/L_0) \geq p'vL - 4, \quad \text{with } p' = \frac{1}{2}(p + 1/2) > 1/2. \quad (6.56)$$

Note that  $|\xi_T^{u_i, v_i}|$  are independent random variables, and  $|\xi_T^{u_i, v_i}| 1_{\{\mathcal{R}_T^{u_i}, \mathcal{R}_T^{v_i} \leq 3\}}$  can be dominated by i.i.d. random variables  $\zeta_i$  with

$$\mathbb{P}(\zeta_i = 1) = 1 - p_0, \quad \mathbb{P}(\zeta_i = 8) = p_0,$$

where as in (6.53),

$$p_0 := 2C\varepsilon\sqrt{T} + C\frac{L_0}{\sqrt{T}} \leq 2C\sqrt{2\delta_0} + \frac{C}{\alpha v\sqrt{T}}.$$

We now choose  $\delta_0$  sufficiently small and  $T = R_0 v^{-2} \leq 2\delta_0 \varepsilon^{-2}$  with  $R_0$  sufficiently large such that  $\mathbb{E}[\zeta_i] = 1 + 7p_0 < 2p'$  with  $p'$  defined in (6.56).

We can then continue the bound in (6.55) to obtain

$$\begin{aligned} \mathbb{P}(|\xi_T| \geq pvL) &\leq CvL\varepsilon^4 T^2 + \mathbb{P}(\cup_{i=1}^\kappa \{\mathcal{R}_T^{u_i} \vee \mathcal{R}_T^{v_i} \geq 4\}) + \mathbb{P}\left(\sum_{i=1}^{\gamma L/2} \zeta_i \geq p'\gamma L - 4\right) \\ &\leq 2CvL\varepsilon^4 T^2 + Ce^{-cvL}, \end{aligned} \quad (6.57)$$

where we used  $\kappa \leq vL/2$  and applied a standard large deviation bound for  $\sum \zeta_i$ . This is exactly (6.45), which concludes the proof of Lemma 6.12.  $\blacksquare$

## A Proof of some facts for the Brownian web and net

In this appendix, we sketch the proof of Proposition 2.6 and Theorem 2.7, which follow readily from results in the literature.

**Proof sketch for Proposition 2.6** As in (1.3), let us consider a pair jump distribution

$$a^{(2)}(x_1, x_2) := (1 - 2\theta\varepsilon)1_{\{x_1=x_2\}}a(x_1) + 2\theta\varepsilon a(x_1)a(x_2) \quad \text{with} \quad a(x) = \frac{1}{2}(1_{\{x=1\}} + 1_{\{x=-1\}}).$$

Since the simple random walk has period 2, we consider the pair of discrete webs  $(W^1, W^2)$  defined as before (1.3), but with paths starting only from the even lattice  $\mathbb{Z}_{\text{even}}^2 = \{(x, t) : x + t \text{ is even}\}$ . Since  $\sigma^2 = \sum_x x^2 a(x) = 1$ , we write  $S_\varepsilon$  for the scaling map  $S_{\varepsilon, \sigma}$ . We will show that  $(S_\varepsilon W^1, S_\varepsilon W^2)$  converges to a pair of sticky Brownian webs  $(\mathcal{W}^1, \mathcal{W}^2)$  with parameter  $\theta$  both in the sense of marking construction and in the sense of martingale characterisation, which then yields the equivalence of the two characterisations.

The convergence  $(S_\varepsilon W^1, S_\varepsilon W^2)$  to  $(\mathcal{W}^1, \mathcal{W}^2)$  with parameter  $\theta$  is a slight extension of [SSS14, Theorem 6.15] and follows the same proof. Assume  $\theta = 1$  to simplify notation. Theorem 6.15 of [SSS14] establishes the convergence of  $(S_\varepsilon W^1, S_\varepsilon N_\varepsilon)$ , the rescaling of a nearest neighbour discrete web  $W^1$  coupled to a nearest neighbour discrete net  $N_\varepsilon$ , to a coupled Brownian web and net  $(\mathcal{W}, \mathcal{N})$ . The discrete web  $W^1$  is obtained from  $N_\varepsilon$  by sampling at branching points, and the pair  $(\mathcal{W}, \mathcal{N})$  is constructed in two ways: either start with  $\mathcal{W}$  and construct  $\mathcal{N}$  by a marking construction similar to the marking construction of the sticky webs, or start with  $\mathcal{N}$  and construct  $\mathcal{W}$  by sampling at the so-called relevant separation points of  $\mathcal{N}$ . The key step in the proof of [SSS14, Theorem 6.15] is that the relevant branching points of  $N_\varepsilon$



converges to the relevant separation points of  $\mathcal{N}$ , while the sampling distribution at relevant branching points of  $N_\varepsilon$  converges to the sampling distribution at the relevant separation points of  $\mathcal{N}$ . As noted in (7.23) in [SSS14], the convergence of  $(S_\varepsilon W^1, S_\varepsilon N_\varepsilon) \Rightarrow (\mathcal{W}, \mathcal{N})$  can be extended to include multiple sticky discrete webs sampled from the same discrete net  $N_\varepsilon$ . In our case, this means  $(S_\varepsilon W^1, S_\varepsilon W^2, S_\varepsilon N_\varepsilon) \Rightarrow (\mathcal{W}^1, \mathcal{W}^2, \mathcal{N})$ , where the discrete net  $N_\varepsilon$  is generated from the pair of discrete webs  $(W^1, W^2)$  by switching between paths in  $W^1$  and  $W^2$ . This is equivalent to starting from a discrete net  $N_\varepsilon$  and then sampling the random walk increments in  $(W^1, W^2)$  with a sampling distribution such that at each branching point of  $N_\varepsilon$ ,  $W^1$  samples one of the two outgoing arrows in  $N_\varepsilon$  with equal probability while  $W^2$  always samples the other arrow. The same sampling distribution can be defined at the relevant separation points of  $\mathcal{N}$ , which in the notation of [SSS14, Lem. 6.16], corresponds to parameters  $p_{--} = p_{++} = 0$ ,  $p_{-+} = p_{+-} = 1/2$ . The limit  $(\mathcal{W}^1, \mathcal{W}^2)$  obtained this way is defined exactly through the marking construction in [SSS14, Theorem 3.5].

On the other hand, the convergence of  $(S_\varepsilon W^1, S_\varepsilon W^2)$  to a limit  $(\mathcal{W}^1, \mathcal{W}^2)$  whose law satisfies the conditions in Theorem 2.5 is a special case of our Theorem 3.1. For the special nearest neighbour case, it also follows from results in the literature. Conditions (i) and (iii) in Theorem 2.5 follow by the same argument as in the proof of Theorem 3.1, while the key condition (ii) on the convergence of a pair of nearest neighbour sticky random walks to sticky Brownian motions was proved in [SSS14, Proposition A.6] (the continuous time version was proved earlier in [HW09a, Theorem 8.1]). This concludes the proof sketch.  $\blacksquare$

**Proof sketch for Theorem 2.7** Assume  $\theta = 1$  to simplify the notation. In [SS08], the Brownian net  $\mathcal{N}$  is defined via hopping between a pair of sticky Brownian webs  $(\mathcal{W}^l, \mathcal{W}^r)$  with drifts  $-1$  and  $+1$  respectively, called the *left-right web*, where paths in  $\mathcal{W}^l$  cannot cross paths in  $\mathcal{W}^r$  from left to right and vice versa. To show that  $\mathcal{N}$  constructed via hopping from a pair of sticky Brownian webs  $(\mathcal{W}^1, \mathcal{W}^2)$  with parameter  $\theta = 1$  is the same as the  $\mathcal{N}$  constructed via hopping from a pair of left-right webs  $(\mathcal{W}^l, \mathcal{W}^r)$ , it suffices to construct a coupling  $(\mathcal{W}^1, \mathcal{W}^2, \mathcal{W}^l, \mathcal{W}^r, \mathcal{N})$  that makes this equivalence obvious. In the discrete setting, such a coupling is given precisely by  $(W^1, W^2, N_\varepsilon)$  defined around (1.3), and then define  $W_\varepsilon^l$  (resp.  $W_\varepsilon^r$ ) by always choosing the leftmost (resp. rightmost) arrow coming out of each point in  $\mathbb{Z}_{\text{even}}^2 := \{(x, t) \in \mathbb{Z}^2 : x + t \text{ is even}\}$ . Equivalently, we can start with  $N_\varepsilon$  and then sample the arrows followed by  $(W^1, W^2, W_\varepsilon^l, W_\varepsilon^r)$  at each branching point of  $N_\varepsilon$ , such that  $W^1$  (resp.  $W^r$ ) always samples the leftmost (resp. rightmost) arrow,  $W^1$  samples one of the two outgoing arrows with equal probability while  $W^2$  samples the other arrow. The scaling limit then gives the desired coupling between  $(\mathcal{W}^1, \mathcal{W}^2, \mathcal{W}^l, \mathcal{W}^r, \mathcal{N})$ , where the same sampling distribution is applied independently at each relevant separation point of  $\mathcal{N}$  to recover the quadruple  $(\mathcal{W}^1, \mathcal{W}^2, \mathcal{W}^l, \mathcal{W}^r)$ . Such a convergence follows from the proof of Theorem 6.15 and (7.23) in [SSS14], and the coupling  $(\mathcal{W}^1, \mathcal{W}^2, \mathcal{N})$  was formulated in [SSS14, Lemma 6.16], which can be easily extended to include also  $(\mathcal{W}^l, \mathcal{W}^r)$ . Lastly, to see why hopping among  $\mathcal{W}^1 \cup \mathcal{W}^2$  and hopping among  $\mathcal{W}^l \cup \mathcal{W}^r$  will both lead to  $\mathcal{N}$ , we note that this is because in both cases, the hopping construction will exhaust both branches of paths coming out of each relevant separation point of  $\mathcal{N}$ , just as in  $N_\varepsilon$ , paths in both  $W^1 \cup W^2$  and  $W^l \cup W^r$  use up all arrows coming out of each branching point of  $N_\varepsilon$ . Using the finite graph representation of paths in the Brownian net (see [SSS14, Section 6.2]), which is the continuum analogue of the directed graph connecting relevant branching points of  $N_\varepsilon$  as introduced in Section 6.2, every path  $\pi \in \mathcal{N}$  can be approximated by a hopping path  $\pi_\delta$  constructed from either  $\mathcal{W}^1 \cup \mathcal{W}^2$  or  $\mathcal{W}^l \cup \mathcal{W}^r$  that visits the same set of  $(\sigma_\pi + (n-1)\delta, \sigma_\pi + n\delta)$ -relevant separation points as  $\pi$  for every  $n \in \mathbb{N}$ . In fact, it suffices to hop between paths at the relevant separation points of  $\mathcal{N}$ . As the time spacing  $\delta \downarrow 0$ , the approximating path  $\pi_\delta$  then converges to  $\pi$  thanks to the

almost sure compactness of  $\mathcal{N}$ . ■

**Acknowledgements.** R. Sun is supported by NUS grant A-8001448-00-00 and NSFC grant 12271475. J.M. Swart is supported by grant 20-08468S of the Czech Science Foundation (GA CR). J. Yu is supported by National Key R&D Program of China (No. 2021YFA1002700) and NSFC (No. 12101238).

## References

- [A17] D. Aldous. Waves in a Spatial Queue *Stochastic Systems* 7, 197-236, 2017.
- [A79] R. Arratia. Coalescing Brownian motions on the line. Ph.D. Thesis, University of Wisconsin, Madison, 1979.
- [A81] R. Arratia. Coalescing Brownian motions and the voter model on  $\mathbb{Z}$ . Unpublished partial manuscript. Available from rarratia@math.usc.edu.
- [AS11] S.R. Athreya and R. Sun. One-dimensional voter model interface revisited. *Electron. Commun. Probab.* 16, Article 70, 792-800, 2011.
- [B99] P. Billingsley. *Convergence of probability measure*, 2nd edition, John Wiley & Sons Inc., New York, 1999.
- [BGS19] M. Birkner, N. Gantert, and S. Steiber. Coalescing directed random walks on the backbone of a 1+1-dimensional oriented percolation cluster converge to the Brownian web. *ALEA Lat. Am. J. Probab. Math. Stat.* 16, no. 2, 1029-1054, 2019.
- [BMSV06] S. Belhaouari, T. Mountford, R. Sun, and G. Valle. Convergence results and sharp estimates for the voter model interfaces. *Electron. J. Probab.* 11, no. 30, 768-801, 2006.
- [BMV07] S. Belhaouari, T. Mountford, and G. Valle. Tightness for the interfaces of one-dimensional voter models. *Proc. London Math. Soc.* 94(3), 421-442, 2007.
- [CDF09] C. Coletti, E. Dias, L.R.G. Fontes. Scaling limit for a drainage network model, *J. Appl. Prob.* 46, 1184-1197, 2009.
- [CH23] G. Cannizzaro, M. Hairer. The Brownian Castle. *Commun. Pure Appl. Math.* 76(10), 2693-2764, 2023.
- [CV14] C.F. Coletti and G. Valle. Convergence to the Brownian Web for a generalization of the drainage network model. *Ann. Inst. H. Poincaré Probab. Statist.* 50(3), 899-919, 2014.
- [CSST21] D. Coupier, K. Saha, A. Sarkar, V.C. Tran. The 2d-directed spanning forest converges to the Brownian web. *Ann. Probab.* 49(1), 435-484, 2021
- [DDP23] S. Das, H. Drillick, S. Parekh. KPZ equation limit of sticky Brownian motion. ArXiv:2304.14279, 2023.
- [EFS17] A. Etheridge, N. Freeman, D. Straulino. The Brownian net and selection in the spatial Lambda-Fleming-Viot process. *Electr. J. Probab.* 22, no. 39, 1-36, 2017.
- [FFW05] P.A. Ferrari, L.R.G. Fontes, and X.-Y. Wu. Two-dimensional Poisson Trees converge to the Brownian web. *Ann. Inst. H. Poincaré Probab. Statist.* 41, 851-858, 2005.

- [FVV15] L.R.G. Fontes, L. Valencia, and G. Valle. Scaling limit of the radial Poissonian web. *Electron. J. Probab.* 20, no. 31, 1-40, 2015.
- [FINR02] L.R.G. Fontes, M. Isopi, C.M. Newman, and K. Ravishankar. The Brownian web. *Proc. Natl. Acad. Sci. USA* 99, no. 25, 15888-15893, 2002.
- [FINR04] L.R.G. Fontes, M. Isopi, C.M. Newman, K. Ravishankar. The Brownian web: characterization and convergence. *Ann. Probab.* 32(4), 2857-2883, 2004.
- [GSW16] A. Greven, R. Sun, A. Winter. Continuum space limit of the genealogies of interacting Fleming-Viot processes on  $\mathbb{Z}$ . *Electron. J. Probab.* 21, paper no. 58, 1-64, 2016.
- [H07] C. Howitt. Stochastic Flows and Sticky Brownian motion. Ph.D. Thesis, University of Warwick, 2007.
- [HW09a] C. Howitt and J. Warren. Consistent families of Brownian motions and stochastic flows of kernels. *Ann. Probab.* 37, 1237-1272, 2009.
- [HW09b] C. Howitt and J. Warren. Dynamics for the Brownian web and the erosion flow. *Stochastic Processes Appl.* 119, 2028-2051, 2009.
- [K97] O. Kallenberg. *Foundations of modern probability*. Springer, New York, 1997.
- [KS91] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*, 2nd edition, Springer-Verlag, New York, 1991.
- [N79] S. V. Nagaev. Large deviations of sums of independent random variables. *Ann. Prob.* 7(5), 745-789, 1979.
- [NRS05] C.M. Newman, K. Ravishankar, and R. Sun. Convergence of coalescing nonsimple random walks to the Brownian web. *Electron. J. Probab.* 10, paper no. 2, 21-60, 2005.
- [NRS10] C.M. Newman, K. Ravishankar, and E. Schertzer. Marking  $(1, 2)$  points of the Brownian web and applications. *Ann. Inst. Henri Poincaré Probab. Statist.* 46, 537-574, 2010.
- [NT12] J. Norris and A. Turner. Hastings-Levitov aggregation in the small-particle limit. *Comm. Math. Phys.* 316, 809-841, 2012.
- [NT15] J. Norris and A. Turner. Weak convergence of the localized disturbance flow to the coalescing Brownian flow. *Ann. Probab.* 43(3), 935-970, 2015.
- [PPR22] A. Parvaneh, A. Parvardeh, and R. Roy. A Drainage Network with Dependence and the Brownian Web. *J. Stat. Phys.* 189(1), paper no. 14, 2022.
- [RSS16] R. Roy, K. Saha, A. Sarkar. Random directed forest and the Brownian web. *Ann. Inst. H. Poincaré Probab. Statist.* 52(3), 1106-1143, 2016.
- [RSS23] R. Roy, K. Saha, A. Sarkar. Scaling limit of a drainage network model on perturbed lattice. arXiv:2302.09489, 2023.
- [S01] F. Spitzer. *Principles of Random Walk*, 2nd edition, Springer-Verlag, New York, 2001.
- [SS08] R. Sun and J.M. Swart. The Brownian net. *Ann. Probab.* 36(3), 1153-1208, 2008.

- [SS11] A. Sarkar, R. Sun. Brownian web in the scaling limit of supercritical oriented percolation in dimension  $1+1$ . *Electron. J. Probab.* 18, no. 21, 2013.
- [SS19] E. Schertzer and R. Sun. Perturbations of supercritical oriented percolation and sticky Brownian webs. *Sojourns in Probability Theory and Statistical Physics - II: Brownian Web and Percolation, A Festschrift for Charles M. Newman*, 241-261, Springer Proceedings in Mathematics & Statistics, 2019.
- [SSS09] E. Schertzer, R. Sun, and J.M. Swart. Special points of the Brownian net. *Electron. J. Probab.* 14, 805-864, 2009.
- [SSS14] E. Schertzer, R. Sun, and J.M. Swart. Stochastic flows in the Brownian web and net. *Mem. Am. Math. Soc.* 227, Nr. 1065, 2014.
- [SSS17] E. Schertzer, R. Sun, and J.M. Swart. The Brownian web, the Brownian net, and their universality. *Advances in Disordered Systems, Random Processes and Some Applications*, 270-368, Cambridge University Press, 2017.
- [SSY19] R. Sun, J.M. Swart, and J. Yu. Equilibrium interfaces of biased voter models. *Ann. Appl. Probab.* 29, 2556-2593, 2019.
- [SSY21] R. Sun, J.M. Swart, and J. Yu. An invariance principle for biased voter model interfaces. *Bernoulli* 27, 615-636, 2021.
- [SVZ22] R. Santos, G. Valle, and L. Zuaznabar. On the convergence of the drainage network with branching. ArXiv:2206.04235, 2022.
- [TW98] B. Tóth and W. Werner. The true self-repelling motion. *Probab. Theory Related Fields* 111, 375-452, 1998.
- [VZ23] G. Valle and L. Zuaznabar. A version of the random directed forest and its convergence to the Brownian web. *J. Theoret. Probab.* 36, no. 2, 948-1002, 2023.
- [VV23] B. Vetó, B. Virág. The geometry of coalescing random walks, the Brownian web distance and KPZ universality. ArXiv:2305.15246, 2023.
- [Y16] J. Yu. Edwards-Wilkinson fluctuations in the Howitt-Warren flows. *Stoch. Process. Their Appl.* 126(3), 948-982, 2016.