

# SCALING LIMIT OF DYNAMICAL PERCOLATION ON CRITICAL ERDŐS–RÉNYI RANDOM GRAPHS

BY RAPHAËL ROSSIGNOL

*Institut Fourier, CNRS, Univ. Grenoble Alpes, [raphael.rossignol@univ-grenoble-alpes.fr](mailto:raphael.rossignol@univ-grenoble-alpes.fr)*

Consider a critical Erdős–Rényi random graph:  $n$  is the number of vertices, each one of the  $\binom{n}{2}$  possible edges is kept in the graph independently from the others with probability  $n^{-1} + \lambda n^{-4/3}$ ,  $\lambda$  being a fixed real number. When  $n$  goes to infinity, Addario-Berry, Broutin and Goldschmidt (*Probab. Theory Related Fields* **152** (2012) 367–406) have shown that the collection of connected components, viewed as suitably normalized measured compact metric spaces, converges in distribution to a continuous limit  $\mathcal{G}_\lambda$  made of random real graphs. In this paper, we consider notably the dynamical percolation on critical Erdős–Rényi random graphs. To each pair of vertices is attached a Poisson process of intensity  $n^{-1/3}$ , and every time it rings, one resamples the corresponding edge. Under this process, the collection of connected components undergoes coalescence and fragmentation. We prove that this process converges in distribution, as  $n$  goes to infinity, toward a fragmentation-coalescence process on the continuous limit  $\mathcal{G}_\lambda$ . We also prove convergence of discrete coalescence and fragmentation processes and provide general Feller-type properties associated to fragmentation and coalescence.

## CONTENTS

1. Introduction . . . . .	323
Background . . . . .	324
2. Notation and background . . . . .	325
2.1. General notation . . . . .	325
2.2. Discrete graphs and dynamical percolation . . . . .	326
2.3. Measured semimetric spaces . . . . .	327
2.4. The Gromov–Hausdorff–Prokhorov distance . . . . .	328
2.5. Gluing and coalescence . . . . .	331
2.5.1. Gluing and $\delta$ -gluing . . . . .	331
2.5.2. The coalescence processes . . . . .	333
2.6. Length spaces and $\mathbb{R}$ -graphs . . . . .	334
2.7. Cutting, fragmentation and dynamical percolation . . . . .	336
2.8. The scaling limit of critical Erdős–Rényi random graphs . . . . .	338
3. Main results . . . . .	338
3.1. General results: Feller and almost Feller properties . . . . .	338
3.2. Main results for Erdős–Rényi random graphs . . . . .	340
3.3. Description of the rest of the article . . . . .	340
4. Preliminary results and tools . . . . .	341
4.1. The multiplicative coalescent . . . . .	341
4.2. Tools to handle $L_{\text{GHP}}$ . . . . .	343
5. Proofs of the main results for coalescence . . . . .	347
5.1. The coalescent on $\mathcal{N}_1$ . . . . .	347
5.2. Structural result for Aldous’ multiplicative coalescent . . . . .	351
5.3. The coalescent on $\mathcal{S}$ . . . . .	355
5.4. Convergence of the coalescent on Erdős–Rényi random graphs . . . . .	362

Received February 2018; revised September 2020.

*MSC2020 subject classifications.* Primary 60K35; secondary 05C80, 60F05.

*Key words and phrases.* Erdős–Rényi, random graph, coalescence, fragmentation, dynamical percolation, scaling limit, Gromov–Hausdorff–Prokhorov distance, Feller property.

6. Proofs of the results for fragmentation . . . . . 364

    6.1. Notation . . . . . 365

    6.2. Reduction to finite  $\mathbb{R}$ -graphs . . . . . 366

    6.3. The Feller property for trees . . . . . 368

    6.4. The Feller property for graphs, Theorem 3.2 . . . . . 374

    6.5. Application to Erdős–Rényi random graphs: Proofs of Theorem 3.6 and Proposition 3.8 . . . . . 385

7. Combining fragmentation and coalescence: Dynamical percolation . . . . . 386

    7.1. Almost Feller property: Proof of Theorem 3.3 . . . . . 386

    7.2. Application to Erdős–Rényi random graphs . . . . . 389

8. Perspectives . . . . . 390

Appendix A: Topologies for processes . . . . . 391

Appendix B: The set of isometry classes of finite measured semimetric spaces . . . . . 393

Appendix C: Commutation relations for coalescence and fragmentation . . . . . 394

Appendix D:  $\mathbb{R}$ -Graphs as  $\mathbb{R}$ -trees with shortcuts . . . . . 397

Appendix E: Total variation distance between Poisson random measures . . . . . 397

Acknowledgments . . . . . 398

References . . . . . 398

**1. Introduction.** Starting with the complete graph with  $n$  vertices,  $K_n$ , the Erdős–Rényi random graph  $\mathcal{G}(n, p)$  is the graph obtained from  $K_n$  by deleting its edges independently with probability  $1 - p$ . A well-known phase transition occurs around  $p_c = \frac{1}{n}$ : when  $p = \frac{c}{n}$  with  $c < 1$ , the largest connected component is of order  $\log n$ , as  $n$  goes to infinity, while if  $p = \frac{c}{n}$  with  $c > 1$ , the largest component is of order  $n$ . This is the so-called *appearance of the giant component*; cf. [12], Section 6. A even more precise *critical window* was discovered, as early as in the seminal work of Erdős and Rényi [20]: when  $p(\lambda, n) := n^{-1} + \lambda n^{-4/3}$ , the largest components are of order  $n^{2/3}$ , and their diameter is of order  $n^{1/3}$ . Then, when  $\lambda$  goes to infinity with  $n$ , a single component starts to dominate the others, and swallows them step by step. Inside this scaling window, that is, for fixed  $\lambda$  and large  $n$ , there is a clean procedure, due to [3], to capture the metric structure of those components: if one assigns mass  $n^{-2/3}$  to each vertex and length  $n^{-1/3}$  to each edge, the largest components, seen as measured metric spaces, converge to a collection of random  $\mathbb{R}$ -graphs  $\mathcal{G}_\lambda$  (see Theorem 2.26 below, or [3], Theorem 24, for a more precise statement). Let us mention that we are particularly interested in the large components because in some sense, they contain all the complexity of the graph: small components are with high probability either trees or unicyclic components; cf. [12], Sections 4–6. Subsequently, the last decade has seen similar results for critical percolation on other random graphs of mean-field type (see notably [8, 11, 18, 19] and Section 8 for further references). Despite being interesting on its own, this phase transition is also related to the study of the minimal spanning tree on the graph on which percolation is performed; cf. [4, 13].

In this article, we shall be interested in dynamical versions of the scaling limits just described. To be more precise, put the following dynamic on  $\mathcal{G}(n, p(\lambda, n))$ : each pair of vertices is equipped with an independent Poisson process with rate  $\gamma_n$  and every time it rings, one *refreshes* the corresponding edge, meaning that one replaces its state by a new independent state: present with probability  $p(\lambda, n)$ , absent with probability  $1 - p(\lambda, n)$ . This procedure corresponds to *dynamical percolation* on the complete graph with  $n$  vertices, at rate  $\gamma_n$ . A natural question now is “At which rate should we refresh the edges in order to see a nontrivial process in the large  $n$  limit?” In this question, it is understood that one remains interested in the same scaling as before concerning masses and lengths: each vertex is assigned mass  $n^{-2/3}$  and each edge is assigned length  $n^{-1/3}$ .

A moment of thought suggests that a good choice should be  $\gamma_n = n^{-1/3}$ . Indeed, since large components are of size  $\Theta(n^{2/3})$ , in a pair of components there are  $\Theta(n^{4/3})$  pairs of vertices which after refreshment lead to  $\Theta(n^{4/3} p(\lambda, n)) = \Theta(n^{1/3})$  edges added. Thus, choosing

$\gamma_n = \Theta(n^{-1/3})$  will lead large components to coalesce at rate  $\Theta(1)$ . Furthermore, the main result of [3] implies that those large components can be seen, once the graph metric has been divided by  $n^{1/3}$ , as compact continuous trees with a finite number of additional cycles. Furthermore, on this limit, typical distances between two (random) vertices inside a large component are of order one, which translates thus to a geodesic of length  $\Theta(n^{1/3})$  in the discrete graph. An edge will be destroyed at rate  $\gamma_n(1 - p(\lambda, n))$  so the geometry inside such a component (at least between typical vertices or even in the sense of the diameter) will be affected at rate  $\Theta(n^{1/3}\gamma_n(1 - p(\lambda, n)))$ , which is again of order  $\Theta(1)$  when  $\gamma_n = \Theta(n^{-1/3})$ . Of course, instead of refreshing the edges, one may decide to only add edges, or to only destroy edges. In the first case, one will observe coalescence of components and in the second case, fragmentation. Once again, one may ask the same question as before: what is the right rate in order to obtain a nontrivial process in the large  $n$  limit, and what is this limit process? One of the main purposes of this article is to give an answer to these questions for the three cases that we just defined informally: dynamical percolation, coalescence and fragmentation. The limit processes will be dynamical percolation, coalescence and fragmentation processes acting on the limit  $\mathcal{G}_\lambda$  obtained in [3]. Furthermore, we will show that coalescence is the time-reversal of fragmentation on this limit. Our approach is to provide Feller-type properties for coalescence and fragmentation, which we hope will be useful in the future to study scaling limits of similar dynamics on other critical random graphs. Notice that the study of coalescence of graphs is a central tool in the work of [7] to show convergence of a number of critical random graphs to  $\mathcal{G}_\lambda$  (configuration models, inhomogeneous random graphs, etc.).

Since a large amount of notation is needed in order to make such statements precise, we will switch to the presentation of notation in Section 2 and then announce the main results and outline the plan of the rest of the article in Section 3. We finish this section by describing informally some works related to coalescence or dynamical percolation.

*Background.* The most important work for the present study is that of Aldous [5]. First, note that there is a natural coupling of  $\{\mathcal{G}(n, p(\lambda, n)), \lambda \in \mathbb{R}\}$  obtained by assigning i.i.d. random variables  $(U_e)$  to the edges of the complete graph on  $n$  vertices and putting edge  $e$  in  $\mathcal{G}(n, p(\lambda, n))$  if and only if  $U_e \leq p(\lambda, n)$ . Studying the coalescence of  $\mathcal{G}(n, p(\lambda, n))$  at rate  $n^{-1/3}$  or the coupled collection  $(\mathcal{G}(n, p(\lambda, n)))_{\lambda \geq 0}$  is essentially equivalent for our purposes. To describe the first major contribution of [5], let us introduce the surplus of a connected graph, which is the minimal number of edges which need to be deleted in order to get a tree. When  $\lambda$  is fixed, Aldous proved convergence in distribution of the rescaled masses of the largest components of  $\mathcal{G}(n, p(\lambda, n))$ , jointly with their surplus. Aldous' approach consists in studying the Łukasiewicz walk associated to an exploration process of the graph—a sequential revelation of the states of the edges following the graph structure—and to show convergence in distribution of this walk to a Brownian motion with parabolic drift. The second main contribution of [5] deals with the dynamics when  $\lambda$  increases. Noting that connected components merge at rate proportional to the product of their (rescaled) sizes when  $\lambda$  grows, Aldous defined an abstract version of this process, which he called the *multiplicative coalescent*: weighted points merge two by two at a rate proportional to the product of their weights. He managed to define this process when the weights are in  $\ell^2$ , and proved that the process then satisfies the Feller property for the  $\ell^2$ -topology. Together with his first main contribution, this implies the convergence of the finite dimensional marginals of the process of rescaled sizes of  $\mathcal{G}(n, p(\lambda, n))$ , when  $\lambda$  increases. Aldous' work had a large legacy. We shall only mention two works which go in the direction of tracking the dynamics of the structure of the graph during coalescence, in that they track the dynamics of the surplus (but not of the whole graph structure). Indeed, between the two contributions described above, Aldous has lost the dynamics of the surplus. In [8], the authors enrich Aldous' multiplicative coalescent by taking into account the dynamics of the surplus of the connected components. This leads to what they call

the *augmented coalescent*, which they prove to satisfy the Feller property for a topology that we will not use in this paper (see however Section 8 for more comments). This allows them to show convergence of the finite dimensional distributions of the processes of rescaled sizes and surplus of  $\{\mathcal{G}(n, p(\lambda, n)), \lambda \geq 0\}$ , and even of other random graphs, namely Achlioptas processes with a bounded-size rule. Finally, in [15], the authors manage to prove the convergence of a sequence of two-parameter processes, where the first parameter is the exploration parameter of  $\mathcal{G}(n, p(\lambda, n))$  and the second is  $\lambda$ . This allows them to obtain the convergence of the process (in  $\lambda$ ) of rescaled sizes of  $\mathcal{G}(n, p(\lambda, n))$  together with their surplus, this time in the Skorokhod sense (not only in the sense of finite dimensional marginals). In order to get this convergence, they define the exploration process using the *Prim order* on vertices of the complete graph. This order is consistent in  $\lambda$ , in the sense that connected components are always intervals of the Prim order, and when  $\lambda$  increases, only adjacent intervals coalesce. Unfortunately, the Prim order seems to be inconsistent with the internal structure of the graph (see [15], Section 4.1, and notably the remarks after Theorem 4). This approach is therefore not adapted for the purpose of the present article.

Let us finish this short review of related works by focusing on dynamical percolation. This theme was introduced in [23], and studied in a number of subsequent works by various authors. In the context of [23], only the edges of some fixed infinite graph are resampled while in the definition above, we resample the edges of a finite complete graph. The scaling limit of dynamical percolation for critical percolation on the two-dimensional triangular lattice was obtained in [22], with techniques quite different from the ones used in the present paper. More related to the present paper is the work [28], where dynamical percolation on critical Erdős–Rényi random graphs, as introduced above, is studied notably at rate 1. The authors show that the size of the largest connected component that appears during the time interval  $[0, 1]$  is of order  $n^{2/3} \log^{1/3} n$  with probability tending to one as  $n$  goes to infinity. They also study “quantitative noise-sensitivity” of the event  $A_n$  that the largest component of  $\mathcal{G}(n, p(\lambda, n))$  is of size at least  $an^{2/3}$  for some fixed  $a > 0$  (see [28], Proposition 2.2). The results in the present paper can be used to find the precise scaling of quantitative noise-sensitivity for events concerning the sizes of the largest components (like  $A_n$  for instance). However, we leave this question and precise statements for future work.

## 2. Notation and background.

2.1. *General notation.* If  $(X, \tau)$  is a topological space, we denote by  $\mathcal{B}(X)$  the Borel  $\sigma$ -field on  $X$ .

If  $\psi$  is a measurable map between  $(E, \mathcal{E})$  and  $(F, \mathcal{F})$ , and  $\mu$  is a measure on  $(E, \mathcal{E})$ , then we denote by  $\psi\#\mu$  the push-forward of  $\mu$  by  $\psi$ :  $\psi\#\mu(A) = \mu(\psi^{-1}(A))$  for any  $A \in \mathcal{F}$ .

We shall frequently use Poisson processes. Let  $(E, \mathcal{E}, \mu)$  be a measurable set with  $\mu$  a  $\sigma$ -finite measure. Denote by  $\text{Leb}(\mathbb{R}^+)$  the Lebesgue  $\sigma$ -field on  $\mathbb{R}^+$ ,  $\text{leb}_{\mathbb{R}^+}$  the Lebesgue measure on  $\mathbb{R}^+$  and let  $\gamma \geq 0$ . If  $\mathcal{P}$  is a Poisson random set with intensity  $\gamma$  on  $(E \times \mathbb{R}^+, \mathcal{E} \times \text{Leb}(\mathbb{R}^+), \mu \times \text{leb}_{\mathbb{R}^+})$  (i.e., with intensity measure  $\gamma\mu \otimes \text{leb}_{\mathbb{R}^+}$ ) we shall denote by  $\mathcal{P}_t$  the multiset containing the points of  $\mathcal{P}$  with birthtime at most  $t$ , with multiplicity the number of times they appear before  $t$ :

$$\mathcal{P}_t := \{ \{x \in E : \exists s \leq t, (x, s) \in \mathcal{P}\} \}.$$

Notice that  $\mathcal{P}_t$  is in general a multiset, not a set, but is a set if  $\mu$  is diffuse. One can equivalently see  $\mathcal{P}_t$  as a counting measure on  $E$ . The disjoint union of two multisets  $A$  and  $B$  will be denoted by  $A \sqcup B$ .

When  $(M, d)$  is a Polish space, let  $\mathcal{F}([0, \infty), M)$  (resp.,  $\mathcal{D}([0, \infty), M)$ ) be the set of functions (resp., càdlàg functions) from  $\mathbb{R}^+$  to  $M$ . We shall use two topologies on  $\mathcal{F}([0, \infty), M)$

and  $\mathcal{D}([0, \infty), M)$ : the Skorokhod topology and the *topology of compact convergence* (also known as topology of uniform convergence on compact sets) which is finer than Skorokhod's topology. Although it is not crucial to use this topology, it turns out that it is more natural, in our setting, to make approximations with this topology, which has furthermore the advantage that limits of càdlàg functions are càdlàg. Convergence in distribution for random processes will however be obtained for the Skorokhod topology. Everything (very little in fact) needed for these topologies is gathered in Appendix A. Furthermore, we shall always suppose that our processes are defined on a complete probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  (completing the original space if necessary). We shall occasionally need  $\mathbb{P}^*$ , the outer measure associated to  $\mathbb{P}$  defined on  $\mathcal{P}(\Omega)$  by

$$\mathbb{P}^*(A) := \inf\{\mathbb{P}(B) : A \subset B \text{ and } B \in \mathcal{F}\}.$$

We shall use the notation  $\mathbb{N}$  for the natural numbers (including 0),  $\mathbb{N}^*$  for the positive natural numbers and  $\overline{\mathbb{N}} := \mathbb{N} \cup \{+\infty\}$ . We shall also adopt the convention that the infimum over an empty set equals  $+\infty$ .

Finally, let us define, for  $p \geq 1$ :

$$\ell_+^p := \left\{ x \in (\mathbb{R}^+)^{\mathbb{N}^*} : \sum_{i \geq 1} |x_i|^p < \infty \right\},$$

and

$$\ell_{\searrow}^p := \{x \in \ell_+^p : x_1 \geq x_2 \geq \dots\}.$$

**2.2. Discrete graphs and dynamical percolation.** We will talk of a *discrete graph* to mean the usual graph-theoretic notion of an unoriented graph, that is, a pair  $G = (V, E)$  with  $V$  a finite set and  $E$  a subset of  $\binom{V}{2} := \{\{u, v\} : u \neq v \in V\}$ . Often,  $E$  is seen as a point in  $\{0, 1\}^{\binom{V}{2}}$ , where 0 codes for the absence of the corresponding edge and 1 for its membership of  $E$ .

For a positive integer  $n$  and  $p \in [0, 1]$ , the *Erdős–Rényi random graph* (or Gilbert random graph)  $\mathcal{G}(n, p)$  is the random graph with vertices  $[n] := \{1, \dots, n\}$  such that each edge is present with probability  $p$ , independently from the others. Alternatively, one may see it as a Bernoulli bond percolation with parameter  $p$  on the complete graph with  $n$  vertices  $K_n = ([n], \binom{[n]}{2})$ .

Let  $\gamma_+$  and  $\gamma_-$  be nonnegative real numbers. If  $G = ([n], E)$  is a discrete graph on  $n$  vertices, define a random process  $N_{\gamma_+, \gamma_-}(G, t) = (G, E_t)$ ,  $t \geq 0$  as follows on the set of subgraphs of the complete graph  $K_n$ . To each pair  $e \in \binom{[n]}{2}$ , we attach two Poisson processes on  $\mathbb{R}^+$ :  $\mathcal{P}_e^+$  of intensity  $\gamma_+$  and  $\mathcal{P}_e^-$  of intensity  $\gamma_-$ . We suppose that all the  $2\binom{[n]}{2}$  Poisson processes are independent. Each time  $\mathcal{P}_e^+$  rings, we replace  $E_{t-}$  by  $E_{t-} \cup \{e\}$  (nothing changes if  $e$  already belongs to  $E_{t-}$ ), and each time  $\mathcal{P}_e^-$  rings, we replace  $E_{t-}$  by  $E_{t-} \setminus \{e\}$  (nothing changes if  $e$  does not belong to  $E_{t-}$ ). The letter  $N$  is intended to be mnemonic for “noise.” If we want to emphasize the Poisson processes, we shall write  $N(G, (\mathcal{P}^+, \mathcal{P}^-)_t)$  instead of  $N_{\gamma_+, \gamma_-}(G, t)$ , with an implicit definition for the map  $N$ .

One may take only  $\mathcal{P}^+$  or only  $\mathcal{P}^-$  into account: write  $N^+(G, \mathcal{P}_t^+)$  for  $N(G, (\mathcal{P}^+, \emptyset)_t)$  and  $N^-(G, \mathcal{P}_t^-)$  for  $N(G, (\emptyset, \mathcal{P}^-)_t)$ . Then  $N^+(G, \mathcal{P}_t^+)$  will be referred as *the discrete coalescent process of intensity  $\gamma^+$  started at  $G$*  and  $N^-(G, \mathcal{P}_t^-)$  as *the discrete fragmentation process of intensity  $\gamma^-$  started at  $G$* .

Now, *dynamical percolation of parameter  $p$  and intensity  $\gamma$* , as described in the **Introduction**, corresponds to the process  $N_{\gamma p, \gamma(1-p)}$ , and is in its stationary state when started with  $\mathcal{G}(n, p)$  (independently of the Poisson processes used to define the dynamical percolation).

All these processes will have continuous counterparts, which will be defined in Sections 2.5 and 2.7.

2.3. *Measured semimetric spaces.* The main characters in this article are the connected components of Erdős–Rényi random graphs and their continuum limit, each one undergoing the updates due to dynamical percolation. One task is therefore to define a proper space where those characters can live, and first to state precisely what we mean by “the connected components of a graph” seen as a single object. One option is to order the components by decreasing order of size,<sup>1</sup> as in [3], or in a size-biased way, as in [5], and thus see the collected components of a graph as a sequence of graphs. However, this order is not preserved under the process of dynamical percolation. Also, looking only at the mass to impose which graphs are pairwise compared between two collections of graphs might lead to a larger distance than what one would expect. Indeed, suppose that  $(G_1, G_2)$  and  $(G'_1, G'_2)$  are two pairs of graphs, with  $G_1$  (resp.,  $G'_1$ ) having slightly larger mass than  $G_2$  (resp.,  $G'_2$ ). One might have  $G_1$  close to  $G'_2$  and  $G_2$  close to  $G'_1$  in some topology (the Gromov–Hausdorff–Prokhorov topology to be defined later), but  $G_1$  far from  $G'_1$  in this topology. For all these reasons, I found it somewhat uncomfortable to work with such a topology in the dynamical context. The topology we will use will be defined in Section 2.4, and the story begins with the definition of a semimetric space.

One way to present the connected components of a graph is to consider the graph as a metric space using the usual graph distance, allowing the metric to take the value  $+\infty$  between points which are not in the same connected component, as in [16], page 1. In addition, the main difficulty in defining dynamical percolation on the continuum limit will be in defining coalescence. In this process some points will be identified, and one clear way to present this is to modify the metric and allow it to be equal to zero between different points rather than performing the corresponding quotient operation. This type of space is called a semimetric space in [16], Definition 1.1.4, and we shall stick to this terminology.

DEFINITION 2.1. A *semimetric space* is a pair  $(X, d)$  where  $X$  is a nonempty set and  $d$  is a function from  $X \times X$  to  $\mathbb{R}^+ \cup \{+\infty\}$  such that for all  $x, y$  and  $z$  in  $X$ :

- $d(x, z) \leq d(x, y) + d(y, z)$ ,
- $d(x, x) = 0$ ,
- $d(x, y) = d(y, x)$ .

A semimetric space  $(X, d)$  is a *metric space* if in addition

- $d(x, y) = 0 \Rightarrow x = y$ .

A metric or semimetric space  $(X, d)$  is said to be *finite* if  $d$  is finite.

Of course, when thinking about a semimetric space  $(X, d)$ , one may visualize the *quotient metric space*  $(X/d, d)$  where points at null distance are identified.  $X$  and  $X/d$  are at zero Gromov–Hausdorff distance (we shall use a version of Gromov–Hausdorff distance extended to semimetric spaces, defined in Section 2.4 below). Notice that  $(X, d)$  is not necessarily a Hausdorff space (different points cannot always be separated by disjoint neighborhoods), but  $(X/d, d)$  always is. Furthermore,  $(X, d)$  is separable if and only if  $(X/d, d)$  is separable.

DEFINITION 2.2. If  $(X, d)$  is a semimetric space, the relation  $\mathcal{R}$  defined by

$$x\mathcal{R}y \iff d(x, y) < \infty$$

---

<sup>1</sup>It requires some device to break ties, but those disappear in the continuum limit, for the Erdős–Rényi random graphs at least.



is an equivalence relation. Each equivalence class is called a *component* of  $(X, d)$  and  $\text{comp}(X, d)$  denotes the set of components. We denote by  $\text{diam}(X)$  the diameter of  $(X, d)$ :

$$\text{diam}(X) = \sup_{x, y \in X} d(x, y)$$

and by  $\text{supdiam}(X)$  the supremum of the diameters of its components:

$$\text{supdiam}(X) = \sup_{m \in \text{comp}(X, d)} \text{diam}(m).$$

DEFINITION 2.3. A *measured semimetric space* (m.s-m.s) is a triple  $X = (X, d, \mu)$  where  $(X, d)$  is a semimetric space and  $\mu$  is a measure on  $X$  defined on a  $\sigma$ -field containing the Borel  $\sigma$ -field for the topology induced by  $d$ .

An m.s-m.s  $(X, d, \mu)$  is said to be *finite* if  $(X, d)$  is a finite totally bounded semimetric space and  $\mu$  is a finite measure.

Finally, we define  $\text{comp}(X) := \text{comp}(X, d)$  and

$$\text{masses}(X) := (\mu(m))_{m \in \text{comp}(X)}.$$

Notice that a finite m.s-m.s has only one component.

REMARK 2.4.

(i) The reason why we allow  $\mu$  to be defined on a larger field than the Borel  $\sigma$ -field is the following. We want to keep  $X$  and  $\mu$  unchanged during coalescence, only the semimetric will change, say to a new semimetric  $d'$ , by performing various identifications. When one performs identifications, the topology shrinks: there are fewer and fewer open sets. Thus, the Borel  $\sigma$ -field shrinks too, and the original measure  $\mu$ , defined on the original Borel  $\sigma$ -field associated to  $d$ , is now defined on a  $\sigma$ -field which is larger than the Borel  $\sigma$ -field associated to  $d'$ .

(ii) One might feel more comfortable after realizing the following. Let  $\pi$  denote the projection from  $(X, d)$  to  $X' := X/d$ ,  $\mathcal{B}'$  the Borel  $\sigma$ -field on  $X'$  and  $\mathcal{B}$  the Borel  $\sigma$ -field on  $X$ . Then  $\pi^{-1}(\mathcal{B}') = \mathcal{B}$  and the image measure  $\pi\#\mu$  on  $X'$  is a Borel measure.

2.4. *The Gromov–Hausdorff–Prokhorov distance.* In the [Introduction](#), we mentioned that  $\mathcal{G}(n, p(\lambda, n))$  converges in distribution, but we did not mention precisely the underlying topology. The main topological ingredient in [4] is the Gromov–Hausdorff–Prokhorov distance between two components of the graph, and we shall use this repeatedly. To define it, we need to recall some definitions from [4].

If  $X = (X, d, \mu)$  and  $X' = (X', d', \mu')$  are two measured semimetric spaces a *correspondence*  $\mathcal{R}$  between  $X$  and  $X'$  is a measurable subset of  $X \times X'$  such that

$$\forall x \in X, \exists x' \in X' : (x, x') \in \mathcal{R}$$

and

$$\forall x' \in X', \exists x \in X : (x, x') \in \mathcal{R}.$$

We let  $C(X, X')$  denote the set of correspondences between  $X$  and  $X'$ . The *distortion* of a correspondence  $\mathcal{R}$  is defined as

$$\text{dis}(\mathcal{R}) := \inf \left\{ \varepsilon > 0 : \forall (x, x'), (y, y') \in \mathcal{R}, \left( \begin{array}{c} d(x, y) \leq d'(x', y') + \varepsilon \\ \text{and} \\ d'(x', y') \leq d(x, y) + \varepsilon \end{array} \right) \right\}.$$

The Gromov–Hausdorff distance between two semimetric spaces  $(X, d)$  and  $(X', d')$  is defined as

$$d_{GH}((X, d), (X', d')) := \inf_{\mathcal{R} \in \mathcal{C}(X, X')} \frac{1}{2} \text{dis}(\mathcal{R}).$$

We denote by  $M(X, X')$  the set of finite Borel measures on  $X \times X'$ . For  $\pi$  in  $M(X, X')$ , we denote by  $\pi_1$  (resp.,  $\pi_2$ ) the first (resp., the second) marginal of  $\pi$ . For any  $\pi \in M(X, X')$ , and any finite measures  $\mu$  on  $X$  and  $\mu'$  on  $X'$  one defines

$$D(\pi; \mu, \mu') = \|\pi_1 - \mu\| + \|\pi_2 - \mu'\|,$$

where  $\|v\|$  is the total variation of a signed measure  $v$ .

The *Gromov–Hausdorff–Prokhorov distance* is defined as follows in [4].

DEFINITION 2.5. If  $X = (X, d, \mu)$  and  $X' = (X', d', \mu')$  are two m.s-m.s, the Gromov–Hausdorff–Prokhorov distance between them is defined as

$$d_{GHP}(X, X') = \inf_{\substack{\pi \in M(X, X') \\ \mathcal{R} \in \mathcal{C}(X, X')}} \left\{ D(\pi; \mu, \mu') \vee \frac{1}{2} \text{dis}(\mathcal{R}) \vee \pi(\mathcal{R}^c) \right\}.$$

It is not difficult to show that  $d_{GHP}$  satisfies the axioms of a semimetric. Let us give a bit more intuition to what the Gromov–Hausdorff–Prokhorov distance measures. On a semimetric space  $(X, d)$ , let us denote by  $\delta_H$  the Hausdorff distance and by  $\delta_{LP}$  the Lévy–Prokhorov distance. Let us recall their definition. For  $B \subset X$  and  $\varepsilon > 0$ , let

$$B^\varepsilon := \{x \in X : \exists y \in B, d(x, y) < \varepsilon\}.$$

Now, for  $A$  and  $B$  subsets of  $X$ ,

$$\delta_H(A, B) := \inf\{\varepsilon > 0 : A \subset B^\varepsilon \text{ and } B \subset A^\varepsilon\}$$

and for finite measures  $\mu$  and  $\nu$  on  $X$ ,

$$(2.1) \quad \delta_{LP}(\mu, \nu) := \inf \left\{ \varepsilon > 0 : \forall B \in \mathcal{B}(X), \begin{pmatrix} \mu(B) \leq \nu(B^\varepsilon) + \varepsilon \\ \text{and} \\ \nu(B) \leq \mu(B^\varepsilon) + \varepsilon \end{pmatrix} \right\}.$$

The following lemma shows that the Gromov–Hausdorff–Prokhorov distance measures how well two measured semimetric spaces can be put in the same ambient space so that simultaneously their measures are close in Prokhorov distance and their geometries are close in Hausdorff distance. It shows that the definitions of [4] and [1] are equivalent. Its proof is a small variation on the proof of [26], Proposition 6, where only probability measures were considered, so we leave it to the reader.

LEMMA 2.6. If  $X = (X, d, \mu)$  and  $X' = (X', d', \mu')$  be two measured separable semimetric spaces, let

$$\tilde{d}_{GHP}(X, X') := \inf_{d''} \{\delta_H(X, X') \vee \delta_{LP}(\mu, \mu')\},$$

where the infimum is over all semimetric  $d''$  on the disjoint union  $X \cup X'$  extending  $d$  and  $d'$ . Then

$$\frac{1}{2} \tilde{d}_{GHP}(X, X') \leq d_{GHP}(X, X') \leq \tilde{d}_{GHP}(X, X').$$



It is easy to see that two finite m.s-m.s  $X$  and  $X'$  are at zero  $d_{\text{GHP}}$ -distance if and only if there are two distance and measure preserving maps  $\phi$  and  $\phi'$  such that  $\phi$  is a map from  $X$  to  $X'$  and  $\phi'$  a map from  $X'$  to  $X$ . Let  $\mathcal{C}$  denote the class of finite measured semimetric spaces and  $\mathcal{R}$  the equivalence relation on  $\mathcal{C}$  defined by  $X\mathcal{R}X' \Leftrightarrow d_{\text{GHP}}(X, X') = 0$ . The quotient  $\mathcal{C}/\mathcal{R}$  can be seen as a set (cf. Appendix B), and we denote this set by the letter  $\mathcal{M}$ , which stands thus for *the set of isometry classes of finite measured semimetric spaces*. The following is shown in [1].

**THEOREM 2.7.** *( $\mathcal{M}, d_{\text{GHP}}$ ) is a complete separable metric space.*

Now, the Gromov–Hausdorff–Prokhorov distance in Definition 2.5 is too strong for our purposes when applied to m.s-m.s which have an infinite number of components: it essentially amounts to a uniform control of the  $d_{\text{GHP}}$ -distance between paired components. We are interested in a weaker distance which localizes around the largest components. We shall restrict to countable unions of finite semimetric spaces with the additional property that for any  $\varepsilon > 0$ , there are only a finite number of components whose size exceeds  $\varepsilon$ . To formulate the distance, it will be convenient to view those semimetric spaces as a set of counting measures on  $\mathcal{M}$ .

**DEFINITION 2.8.** For any  $\varepsilon > 0$ , let

$$\mathcal{M}_{>\varepsilon} = \{(X, d, \mu) \in \mathcal{M} \text{ s.t. } \mu(X) > \varepsilon\}.$$

For any counting measure  $\nu$  on  $\mathcal{M}$ , denote by  $\nu_{>\varepsilon}$  the restriction of  $\nu$  to  $\mathcal{M}_\varepsilon$ . Denote by  $\mathcal{N}$  the set of counting measures  $\nu$  on  $\mathcal{M}$  such that for any  $\varepsilon > 0$ ,  $\nu_{>\varepsilon}$  is a finite measure and such that  $\nu$  does not have atoms of mass 0, that is,

$$\nu(\{(X, d, \mu) \in \mathcal{M} \text{ s.t. } \mu(X) = 0\}) = 0.$$

When  $X$  is a measured semimetric space whose components are finite, we denote by  $\nu_X$  the counting measure on  $\mathcal{M}$  defined by

$$\nu_X := \sum_{m \in \text{comp}(X)} \delta_{[m]}.$$

Abusing notation, we shall say that  $X \in \mathcal{N}$  if  $\nu_X$  belongs to  $\mathcal{N}$ , and we shall denote by  $X_{>\varepsilon}$  (resp.  $X_{\leq\varepsilon}$ ) the disjoint union of components of  $X$  whose masses are larger than  $\varepsilon$  (resp., at most  $\varepsilon$ ).

Notice that  $X$  is in  $\mathcal{N}$  if and only if it has an at most countable number of components, each one of its components has positive mass and for any  $\varepsilon > 0$ ,  $X_{>\varepsilon}$  is the disjoint union of a finite number of components, each one being totally bounded and equipped with a finite measure.

Essentially, we want to define a metric on  $\mathcal{N}$  such that a sequence  $\nu_n$  converges to  $\nu$  if and only if for any  $\varepsilon > 0$ , the components of  $\nu$  of mass larger than  $\varepsilon$  are close (for  $d_{\text{GHP}}$ ) to the components of  $\nu_n$  of mass larger than  $\varepsilon$ , for  $n$  large enough. The idea is similar to the metrization of vague convergence for measures on  $\mathbb{R}^d$ , with components of mass zero playing the role of infinity. In this vein, there is an abstract notion of locally finite measure in [25], but we shall not use it.

First, let  $\delta_{\text{LP}}$  be the Lévy–Prokhorov distance on the set of finite measures on the metric space  $(\mathcal{M}, d_{\text{GHP}})$ . Recall the definition of this distance in (2.1). Now, for  $X = [(X, d, \mu)] \in$

$\mathcal{M}$  and  $k \geq 1$ , define a function  $f_k$  by

$$f_k(X) := \begin{cases} 1 & \text{if } \mu(X) \geq \frac{1}{k}, \\ k(k+1)\left(\mu(X) - \frac{1}{k+1}\right) & \text{if } \mu(X) \in \left[\frac{1}{k+1}, \frac{1}{k}\right], \\ 0 & \text{if } \mu(X) < \frac{1}{k+1}. \end{cases}$$

The following distance is an analogue of the distance in [25], Lemma 4.6, where it is used to metrize the vague topology on locally finite measures.

DEFINITION 2.9. If  $\nu$  and  $\nu'$  are counting measures on  $\mathcal{M}$ , then we define

$$L_{\text{GHP}}(\nu, \nu') := \sum_{k \geq 1} 2^{-k} \{1 \wedge \delta_{\text{LP}}(f_k \nu, f_k \nu')\},$$

where  $f_k \nu$  is defined as follows for  $\nu = \sum_{i \in I} \delta_{x_i}$ :

$$f_k \nu := \sum_{i \in I} f_k(x_i) \delta_{x_i}.$$

We shall prove later, in Proposition 4.6 that  $(\mathcal{N}, L_{\text{GHP}})$  is a complete separable metric space. Notice that any m.s-m.s of  $\mathcal{N}$  is at zero  $L_{\text{GHP}}$ -distance from a m.s-m.s whose components are measured compact metric spaces. In this article, we really are interested in equivalence classes of m.s-m.s for the equivalence relation “being at zero  $L_{\text{GHP}}$ -distance,” although in order to define random processes such as coalescence and fragmentation, it will be convenient to have in mind a particular representative of such a class.

### 2.5. Gluing and coalescence.

2.5.1. *Gluing and  $\delta$ -gluing.* Gluing corresponds to identification of points which can belong to the same semimetric space or to different semimetric spaces. A formal definition is as follows, for a single semimetric space (see also [16], pages 62–64).

DEFINITION 2.10. Let  $(X, d)$  be a semimetric space and  $\mathcal{R}$  be an equivalence relation on  $X$ . The *gluing of  $(X, d)$  along  $\mathcal{R}$* , is the semimetric space  $(X, d_{\mathcal{R}})$  with semimetric defined on  $X^2$  by

$$d_{\mathcal{R}}(x, y) := \inf \left\{ \sum_{i=0}^k d(p_i, q_i) : p_0 = x, q_k = y, k \in \mathbb{N} \right\},$$

where the infimum is taken over all choices of  $\{p_i\}_{0 \leq i \leq k}$  and  $\{q_i\}_{0 \leq i \leq k}$  such that  $(q_i, p_{i+1}) \in \mathcal{R}$  for all  $i = 0, \dots, k - 1$ .

When performing dynamical percolation on a discrete graph, edges appear, and these are not of length zero. Thus one needs a definition of gluing which leaves the possibility to add those edges. We shall define the  $\delta$ -gluing of a semimetric space  $X$  along a multiset  $\tilde{\mathcal{R}}$  with elements in  $X^2$  as the operation of joining every pair  $(x, x') \in \tilde{\mathcal{R}}$  by an isometric copy of the interval  $[0, \delta]$ . If  $(X, d)$  and  $(X', d')$  are two semimetric spaces, let us denote by  $d_{X \sqcup X'}$  the *disjoint union semimetric* on the disjoint union  $X \sqcup X'$ , which is the semimetric equal to  $d$  on  $X \times X$ , to  $d'$  on  $X' \times X'$  and to  $+\infty$  on  $(X \times X') \cup (X' \times X)$ . This notion extends trivially to the disjoint union of a collection of metric spaces.

DEFINITION 2.11. Let  $(X, d)$  be a semimetric space,  $\tilde{\mathcal{R}}$  be a multiset of elements in  $X^2$  and  $\delta \geq 0$ .

- If  $\delta = 0$ , let  $\mathcal{R}$  denote the equivalence relation generated by  $\tilde{\mathcal{R}}$  and let  $X_{\tilde{\mathcal{R}},\delta} = X$  and  $d' = d$ .
- If  $\delta > 0$ , for every pair  $(x, x') \in \tilde{\mathcal{R}}$ , let  $I_{x,x'} = [a_{x,x'}, b_{x,x'}]$  be an isometric copy of  $[0, \delta]$ . Let us denote by  $X_{\tilde{\mathcal{R}},\delta}$  the disjoint union of  $X$  and all the  $I_{x,x'}$ , for  $(x, x')$  in  $\tilde{\mathcal{R}}$  and let  $d'$  denote the disjoint union distance on  $X_{\tilde{\mathcal{R}},\delta}$ . Let  $\mathcal{R}$  denote the equivalence relation on  $X_{\tilde{\mathcal{R}},\delta}$  generated by

$$\bigcup_{(x,x') \in \tilde{\mathcal{R}}} \{(x, a_{x,x'}), (x', b_{x,x'})\}.$$

Then the  $\delta$ -gluing of  $(X, d)$  is the metric space  $(X_{\tilde{\mathcal{R}},\delta}, d_{\tilde{\mathcal{R}},\delta})$  which is the gluing of  $(X_{\tilde{\mathcal{R}},\delta}, d')$  along  $\mathcal{R}$ .

When  $\delta > 0$ , let us denote by  $\mathcal{F}$  the following  $\sigma$ -field on  $X_{\tilde{\mathcal{R}},\delta}$ :

$$\mathcal{F} := \left\{ A \cup B : A \in \mathcal{B}(X), B \subset \bigcup_{(x,x') \in \tilde{\mathcal{R}}} I_{x,x'} \right\}.$$

We can lift trivially any measure  $\mu$  on  $\mathcal{B}(X)$  to a measure  $\tilde{\mu}$  on  $\mathcal{F}$  by

$$\tilde{\mu}(C) := \mu(C \cap X).$$

When  $X = (X, d, \mu)$  is a measured semimetric space, we equip the  $\delta$ -gluing of  $(X, d)$  along  $\tilde{\mathcal{R}}$  with this lifted measure, but we shall still denote this measure by  $\mu$  and denote the resulting semimetric space by  $\text{Coal}_\delta(X, \tilde{\mathcal{R}}) = (X_{\tilde{\mathcal{R}},\delta}, d_{\tilde{\mathcal{R}},\delta}, \mu)$ .

REMARK 2.12.

- (i) For any  $(x, y) \in X^2$ , if  $\delta > 0$ , one may see that

$$d_{\tilde{\mathcal{R}},\delta}(x, y) = \inf \left\{ k\delta + \sum_{i=0}^k d(p_i, q_i) : p_0 = x, q_k = y, k \in \mathbb{N} \right\},$$

where the infimum is taken over all choices of  $\{p_i\}_{0 \leq i \leq k}$  and  $\{q_i\}_{0 \leq i \leq k}$  such that  $(q_i, p_{i+1}) \in \tilde{\mathcal{R}}$  for all  $i = 0, \dots, k - 1$ . Furthermore, the same is true for  $\delta = 0$  if  $\tilde{\mathcal{R}}$  is already an equivalence relation.

- (ii) The previous expressions allow us to prove that for  $A, B \subset X^2$ ,

$$\text{Coal}_0(X, A \cup B) = \text{Coal}_0(\text{Coal}_0(X, A), B).$$

Also, for any multisets  $A, B$  and for any  $\delta > 0$ ,

$$\text{Coal}_\delta(X, A \sqcup B) = \text{Coal}_\delta(\text{Coal}_\delta(X, A), B).$$

See Appendix C.

(iii) If  $\delta > 0$ , one may like to consider the space  $X$  with (the restriction of) the metric  $\delta_{\tilde{\mathcal{R}},\delta}$ : it can be seen as forgetting the interiors of the intervals  $I_{x,x'}$  that have been added in  $X_{\tilde{\mathcal{R}},\delta}$  (while keeping the same metric). If  $X = (X, d, \mu) \in \mathcal{N}$ ,

$$(2.2) \quad LGHP((X, d_{\tilde{\mathcal{R}},\delta}, \mu), \text{Coal}_\delta(X, \tilde{\mathcal{R}})) \leq \delta.$$

Indeed, using the notation of Definition 2.11, one may use the following correspondence on  $X \times X_{\tilde{\mathcal{R}},\delta}$ :

$$C = X^2 \cup \bigcup_{(x,x') \in \tilde{\mathcal{R}}} \{(x, y) : y \in I_{x,x'}\}$$

which has distortion  $\delta$  when  $X$  is equipped with  $d_{\tilde{R},\delta}$  (because the only distances that get distorted are those within the intervals of length  $\delta$ ), and then use  $\pi$  as the trivial coupling of  $\mu$  on  $X$  and  $\mu$  on  $X_{\tilde{R},\delta}$ , which satisfies

$$D(\pi; \mu, \mu) = \pi(C^c) = 0.$$

This can be done on each component of  $(X, d_{\tilde{R},\delta})$  separately to show that (2.2) holds.

(iv) It is easy to see that for any  $X = (X, d, \mu)$  and  $\tilde{R} \subset X^2$ ,

$$\text{Coal}_0(X/d, \tilde{R}/d)/d_{\tilde{R},0} = \text{Coal}_0(X, \tilde{R})/d_{\tilde{R},0}$$

thus one may without harm identify the semimetric spaces with their metric quotients before or after coalescence (in fact the collection of the metric quotients of its components).

2.5.2. *The coalescence processes.* When  $(X, d, \mu)$  is a measured semimetric space, there is a natural coalescence process (of mean-field type) which draws pairs of points  $(x, y)$  with intensity  $\mu(dx)\mu(dy)$  (and unit intensity in time) and identifies points  $x$  and  $y$ , changing the metric accordingly. To describe the process of addition of edges during dynamical percolation on the Erdős–Rényi random graph, one needs to replace the identification of  $x$  and  $y$  by the fact that the distance between  $x$  and  $y$  drops to  $n^{-1/3}$  (if  $x \neq y$ ). This leads to the following definition.

DEFINITION 2.13. Let  $X = (X, d, \mu)$  be an m.s-m.s with  $\mu$  sigma-finite and  $\delta \geq 0$ . Let  $\mathcal{P}^+$  be a Poisson random set on  $X^2 \times \mathbb{R}^+$  of intensity measure  $\frac{1}{2}\mu^2 \times \text{leb}_{\mathbb{R}^+}$ . The *coalescence process with edge-lengths  $\delta$  started from  $X$* , denoted by  $(\text{Coal}_\delta(X, t))_{t \geq 0}$ , is the random process of m.s-m.s  $(\text{Coal}_\delta(X, \mathcal{P}_t^+))_{t \geq 0}$ .

Notice that this process inherits the strong Markov property from the strong Markov property of the Poisson process, and the fact that for multisets  $A$  and  $B$  of elements of  $X^2$ ,  $\text{Coal}_\delta(X, A \sqcup B) = \text{Coal}_\delta(\text{Coal}_\delta(X, A), B)$ ; cf. Remark 2.12(ii).

When  $\delta > 0$ , if one wants to keep the space fixed and change only the metric, Remark 2.12(iii) shows that one can do so at the price of an  $L_{\text{GHP}}$ -distance at most  $\delta$ . In this paper, one wants typically to understand scaling limits of  $N^+(G_n, \mathcal{P}_t^+)$  as defined in Section 2.2 with  $\mathcal{P}^+$  of intensity  $\gamma_n$  and  $G_n$  a discrete graph equipped with the distance  $d_n$  which is the graph distance multiplied by some  $\delta_n > 0$  going to zero as  $n$  goes to infinity; see, for instance, Theorem 3.5 below. If one equips  $G_n$  and  $N^+(G_n, \mathcal{P}_t^+)$  with their counting measures multiplied by  $\sqrt{\gamma_n}$ ,  $N_{\gamma_n}^+(G_n, t)$  is at  $L_{\text{GHP}}$ -distance at most  $\delta_n$  from  $(\text{Coal}_{\delta_n}((G_n, d_n, \mu_n), t))_{t \geq 0}$  (under a natural coupling), so the scaling limits will be the same. We shall want to identify the limit itself as  $(\text{Coal}_0(\mathcal{G}_\lambda, t))_{t \geq 0}$ , and part of our work will consist in showing that it is a nicely behaved process. In order to accomplish this task, we need to define some subsets of  $\mathcal{N}$ .

DEFINITION 2.14. For  $p > 0$ , we define  $\mathcal{N}_p$  to be the set of elements  $\nu$  of  $\mathcal{N}$  such that if  $\nu$  is written as  $\sum_{m \in I} \delta_m$  for some countable index set  $I$ ,

$$\sum_{m \in I} \mu(m)^p < \infty.$$

For  $\nu$  in  $\mathcal{N}_p$ , we let  $\text{masses}(\nu)$  to be the sequence in  $\ell^p_{\searrow}$  of masses  $\mu(m)$  listed in decreasing order and define, for  $\nu$  and  $\nu'$  in  $\mathcal{N}_p$ ,

$$L_{p,\text{GHP}}(\nu, \nu') = L_{\text{GHP}}(\nu, \nu') \vee \|\text{masses}(\nu) - \text{masses}(\nu')\|_p.$$

Again we shall abuse language, saying that  $X = (X, d, \mu)$  is in  $\mathcal{N}_p$  when  $\nu_X \in \mathcal{N}_p$  and write  $L_{p,\text{GHP}}(X, X')$  for  $L_{p,\text{GHP}}(\nu_X, \nu_{X'})$ . We let the reader check that  $(\mathcal{N}_p, L_{p,\text{GHP}})$  is a complete separable metric space.

It is easy to see that if  $X = (X, d, \mu)$  belongs to  $\mathcal{N}_1$ , then almost surely, for every  $t \geq 0$ ,  $\text{Coal}_{t,\delta}(X, d, \mu)$  is in  $\mathcal{N}_1$ . We even have the Feller property on  $\mathcal{N}_1$ , which will be proved in Section 5.1. A consequence of the Feller property of the multiplicative coalescent in  $\ell^2$  is that if  $X = (X, d, \mu)$  belongs to  $\mathcal{N}_2$ , then almost surely for every  $t \geq 0$   $\sum_{m \in \text{comp}(\text{Coal}_\delta(X,t))} \mu(m)^2 < \infty$ . However, one cannot guarantee that components stay totally bounded, and thus that  $\text{Coal}_\delta(X, t)$  or even  $\text{Coal}_0(X, t)$  belongs to  $\mathcal{N}_2$ . One will thus have to restrict to a subclass of  $\mathcal{N}_2$ , which will fortunately contain  $\mathcal{G}_\lambda$  with probability one.

DEFINITION 2.15. We define  $\mathcal{S}$  to be the class of m.s-m.s  $X = (X, d, \mu)$  in  $\mathcal{N}_2$  such that

$$(2.3) \quad \forall t \geq 0, \quad \text{supdiam}(\text{Coal}_0(X_{\leq \eta}, t)) \xrightarrow[\eta \rightarrow 0]{\mathbb{P}} 0.$$

It will be shown in Lemmas 5.9 and 5.11 that if  $X \in \mathcal{S}$ , then almost surely, for any  $t \geq 0$ ,  $\text{Coal}_0(X, t) \in \mathcal{S}$ . Of course, I suspect that  $\mathcal{S}$  has a more intrinsic definition, at least when the components are  $\mathbb{R}$ -graphs, and that there is a convenient topology which turns it into a Polish space, however I could not prove this for the moment. Let us mention that there are elements in  $\mathcal{N}_2 \setminus \mathcal{S}$ ; see Remark 5.8.

Let us finish this section by a description of the coalescence at the level of components. When  $X$  and  $\mathcal{P}^+$  are as in Definition 2.13, one may associate to them a process of multigraphs with vertices  $\text{comp}(X)$  which we denote by  $\mathbb{M}\mathbb{G}(X, t)$ . It is defined as follows: there is an edge in  $\mathbb{M}\mathbb{G}(X, t)$  between  $m$  and  $m'$  if there is a point  $(a, b, s)$  of  $\mathcal{P}^+$  with  $s \leq t$ ,  $a \in m$  and  $b \in m'$ .

When  $A$  is a measurable subset of  $X$ , let  $A := (A, d|_{A \times A}, \mu|_A)$ . There is an obvious coupling between  $(\text{Coal}_\delta(X, t))_{t \geq 0}$  and  $(\text{Coal}_\delta(A, t))_{t \geq 0}$ : just take the restriction of the Poisson random set  $\mathcal{P}$  to  $A^2 \times \mathbb{R}$ . We shall call it the *obvious coupling*. We shall use several times the following easy fact.

LEMMA 2.16. Suppose that  $A$  is a union of components of  $X$ . Under the obvious coupling, if  $\mathbb{M}\mathbb{G}(X, t)$  is a forest, then for every  $s \leq t$  and every  $x, y$  in  $A$ , one has the following:

if the distance between  $x$  and  $y$  in  $(\text{Coal}_\delta(A, s))$  is finite,  
then it is equal to the distance between  $x$  and  $y$  in  $(\text{Coal}_\delta(X, s))$ .

2.6. Length spaces and  $\mathbb{R}$ -graphs. We refer to [4] for background on the definitions and statements in this section. Let us however recall the definition of a *length space* and a *geodesic space* for semimetric spaces (cf. [16] for general background on length spaces).

$(X, d)$  is a length space if and only if for any  $x$  and  $y$  in  $X$ , the distance between  $x$  and  $y$  is the infimum of the lengths of paths  $\gamma$  between  $x$  and  $y$ . The *length* of a path  $\gamma$  from  $[a, b]$  to  $X$  being defined as

$$\sup \sum_{i=0}^{r-1} d(\gamma(t_i), \gamma(t_{i+1})),$$

where the supremum is over all  $a \leq t_0 < t_1 \leq \dots \leq t_r = b$ . We say that a length space  $(X, d)$  is *geodesic* if for any  $x$  and  $y$ , there is a path from  $x$  to  $y$  with length  $d(x, y)$ .

DEFINITION 2.17. An  $\mathbb{R}$ -tree is a totally bounded geodesic and acyclic finite metric space. An  $\mathbb{R}$ -graph is a totally bounded geodesic finite metric space  $(G, d)$  such that there

exists  $R > 0$  such that for any  $x \in G$ ,  $(B_R(x), d|_{B_R(x)})$  is an  $\mathbb{R}$ -tree, where  $B_R(x)$  is the ball of radius  $R$  and center  $x$ .

For a semimetric space  $(X, d)$ , we shall say that it is a *semimetric  $\mathbb{R}$ -graph* if the quotient metric space  $(X/d, d)$  is an  $\mathbb{R}$ -graph.

REMARK 2.18.

(i) The definition above differs slightly from [4], Definition 2.2, where an  $\mathbb{R}$ -graph  $(X, d)$  is defined as a compact geodesic metric space such that for any  $x \in G$ , there exists  $\varepsilon = \varepsilon(x) > 0$  such that  $(B_\varepsilon(x), d|_{B_\varepsilon(x)})$  is an  $\mathbb{R}$ -tree, where  $B_\varepsilon(x)$  is the ball of radius  $\varepsilon$  and center  $x$ . When  $(X, d)$  is compact, the two definitions agree: one direction is obvious, whereas the other follows from the arguments at the beginning of [4], Section 6.1. One advantage of working with precompact spaces instead of compact ones is that one may avoid having to take the completion in order to recover an  $\mathbb{R}$ -graph after fragmentation (notice that after fragmentation, the space is not complete anymore).

(ii) A semimetric space  $(X, d)$  is a length space (resp., a geodesic space) if and only if its quotient metric space  $(X/d, d)$  is a length space (resp., a geodesic space). Thus, a semi-metric  $\mathbb{R}$ -graph is notably a geodesic semimetric space, and thus a length semimetric space.

The *degree*  $d_G(x)$  of a point  $x$  in an  $\mathbb{R}$ -graph  $(G, d)$  is the number of connected components of  $B_R(x) \setminus \{x\}$ , where  $R$  is any positive real number such that  $(B_R(x), d|_{B_R(x)})$  is an  $\mathbb{R}$ -tree. A *branchpoint*  $x$  is a point with  $\text{deg}_G(x) \geq 3$ . A *leaf*  $x$  is a point with degree one. We denote by  $\text{leaves}(G)$  the set of leaves of  $G$ . An  $\mathbb{R}$ -tree or an  $\mathbb{R}$ -graph is said to be *finite* if it is compact and has a finite number of leaves.

An  $\mathbb{R}$ -graph  $(G, d)$ , and more generally a length space, is naturally equipped with a *length measure*, which assigns notably its length to the image of a simple path. Formally, it is defined as the 1-dimensional Hausdorff measure on  $(G, d)$  (see [16], Sections 1.7 and 2.6). We shall denote it by  $\ell_G$ , and when  $G$  is an  $\mathbb{R}$ -graph, it is a  $\sigma$ -finite diffuse measure. If  $(X, d)$  is a semimetric length space, the length measure of  $X/d$  can be naturally carried over to  $X$ , since Borel sets of  $X/d$  are in bijection with Borel sets of  $X$ . We shall call this measure the length measure of  $X$  and denote it by  $\ell_X$ . For a semimetric space  $(X, d)$ , we shall say that a Borel measure  $\ell$  is *diffuse* if for any  $x \in X$ ,  $\ell(\{y \in X : d(x, y) = 0\}) = 0$  (or equivalently that its image on  $X/d$  is diffuse). When  $X$  is a semimetric length space,  $\ell_X$  is thus a diffuse measure.

The structure of an  $\mathbb{R}$ -graph is explained thoroughly<sup>2</sup> in [4], Section 6.2. The *core* of  $(G, d)$ , denoted by  $\text{core}(G)$  is the union of all simple arcs with both endpoints in embedded cycles of  $G$ . It is also the maximal compact subset of  $G$  having only points of degree at least 2 (cf. [4], Corollary 2.5, where one needs to replace “closed” by “compact” in our precompact setting). The core of a tree is empty, that of a unicyclic graph is a cycle. When  $G$  is neither a tree nor unicyclic, there is a finite connected multigraph  $\text{ker}(G) = (k(G), e(G))$  called the *kernel* of  $G$  such that the core of  $G$  may be obtained from  $\text{ker}(G)$  by gluing along each edge  $e$  an isometric copy of the interval  $[0, l(e)]$ , for some  $l(e) > 0$ . The *surplus* of  $G$  is defined as 0 when  $G$  is a tree, 1 when  $G$  is unicyclic, and otherwise as

$$\text{surplus}(G) = |e(G)| - |k(G)| + 1,$$

which is then at least two. If  $(X, d)$  is a semimetric  $\mathbb{R}$ -graph, we shall define its surplus as the surplus of  $(X/d, d)$ .

Using the existence of the core, one gets the following equivalent definition of an  $\mathbb{R}$ -graph, where an  $\mathbb{R}$ -graph is obtained as a “tree with shortcuts,” to employ the expression of [7]. A sketch of proof is given in Appendix D.

<sup>2</sup>Although our definition differs slightly, the proof of [4], Proposition 6.2, can be adapted straightforwardly.



LEMMA 2.19. *A metric space  $(X, d)$  is an  $\mathbb{R}$ -graph if and only if there exists an  $\mathbb{R}$ -tree  $(T, d)$  and a finite set  $A \subset T^2$  such that  $(X, d)$  is isomorphic to the quotient metric space obtained from  $\text{Coal}_0((T, d), A)$ .*

Now, let us introduce a quantity that will be useful to control the diameters of components during dynamical percolation, notably because it is monotone under fragmentation (contrarily to the diameter). In a semimetric space  $(X, d)$ , we say that a path is injective if its projection on  $X/d$  is injective. Then, for a semimetric length space  $X$  with length measure  $\ell$ , define

$$\text{suplength}(X) := \sup\{\ell(\gamma) : \gamma \text{ is a rectifiable injective path in } X\}.$$

Notice that

$$\text{suplength}(X) = \sup_{m \in \text{comp}(X)} \text{suplength}(m).$$

Now, we can define the various spaces on which we shall study fragmentation and dynamical percolation.

DEFINITION 2.20. Define  $\mathcal{S}^{\text{length}}$  as the class of length semimetric spaces  $X$  in  $\mathcal{N}_2$  whose length measure is  $\sigma$ -finite and such that

$$(2.4) \quad \forall t \geq 0, \quad \text{suplength}(\text{Coal}_0(X_{\leq \eta}, t)) \xrightarrow[\eta \rightarrow 0]{\mathbb{P}} 0.$$

Let  $\mathcal{M}^{\text{graph}}$  denote the set of equivalence classes of  $\mathbb{R}$ -graphs under  $d_{\text{GHP}}$ . Define  $\mathcal{N}^{\text{graph}}$  (resp.,  $\mathcal{N}_p^{\text{graph}}$ ) from  $\mathcal{M}^{\text{graph}}$  in the same way that  $\mathcal{N}$  (resp.,  $\mathcal{N}_p$ ) was defined from  $\mathcal{M}$ .

Define  $\mathcal{S}^{\text{graph}}$  as the class of semimetric spaces  $X$  in  $\mathcal{N}_2^{\text{graph}}$  such that (2.4) holds.

If  $X = (X, d, \mu)$  and  $X' = (X', d', \mu')$  are measured semimetric  $\mathbb{R}$ -graphs, define

$$d_{\text{GHP}}^{\text{surplus}}(X, X') := d_{\text{GHP}}(X, X') \vee |\text{surplus}(X) - \text{surplus}(X')|.$$

Finally, define  $L_{\text{GHP}}^{\text{surplus}}$  and  $L_{p, \text{GHP}}^{\text{surplus}}$  in the same way that  $L_{\text{GHP}}$  and  $L_{p, \text{GHP}}$  were defined, but replacing  $d_{\text{GHP}}$  by  $d_{\text{GHP}}^{\text{surplus}}$ .

Notice that  $\mathcal{S}^{\text{graph}}$  is a subclass of  $\mathcal{S}^{\text{length}}$ . A rigorous definition of  $\mathcal{M}^{\text{graph}}$  as a set is given in Appendix B. Thanks to Lemma 2.19, it is clear that if  $X \in \mathcal{N}^{\text{graph}}$  and  $\mathcal{P} \subset X^2$  is finite, then for any  $\delta \geq 0$ ,  $\text{Coal}_\delta(X, \mathcal{P})$  still belongs to  $\mathcal{N}^{\text{graph}}$ .

Additional notation concerning  $\mathbb{R}$ -graphs will be introduced when needed, in Section 6.1.

### 2.7. Cutting, fragmentation and dynamical percolation.

DEFINITION 2.21. Suppose that  $X = (X, d, \mu)$  is an m.s-m.s which is a length space and  $\mathcal{P}^-$  is a subset of  $X$ . Let  $\mathcal{P}^{-,d}$  denote

$$\{x \in X : \exists y \in \mathcal{P}^-, d(x, y) = 0\}.$$

Then the *cut of  $X$  along  $\mathcal{P}^-$* , denoted by  $\text{Frag}(X, \mathcal{P}^-)$  is the m.s-m.s  $(X \setminus \mathcal{P}^{-,d}, d_{\mathcal{P}^-}^{\text{Frag}}, \mu|_{X \setminus \mathcal{P}^{-,d}})$  where

$$d_{\mathcal{P}^-}^{\text{Frag}}(x, y) := \inf_{\gamma} \{\ell_X(\gamma)\}$$

and the infimum is over all paths  $\gamma$  from  $x$  to  $y$  disjoint from  $\mathcal{P}^{-,d}$ .

REMARK 2.22.

(i)  $\text{Frag}(X, \emptyset)$  is the same as  $X$  precisely because the components of  $(X, d)$  are length spaces. Furthermore,  $\text{Frag}(X, \mathcal{P}^-)$  is still a length space; cf. Lemma C.3.

(ii) If  $(X, d)$  is complete,  $\text{Frag}((X, d, \mu), \mathcal{P}^-)$  is generally not complete anymore. However, notice that a measured semimetric space is always at zero  $L_{\text{GHP}}$ -distance from its completion. Indeed, denoting  $(\bar{X}, \bar{d}, \bar{\mu})$  the completion of  $(X, d, \mu)$ , recall that  $X$  is canonically mapped isometrically to a subset of  $\bar{X}$  and  $\bar{\mu}$  is the image of  $\mu$  under this mapping. Then, for any  $\varepsilon > 0$ , the following set is a correspondence of distortion at most  $2\varepsilon$ :

$$\mathcal{R} := \{(x, y) \in X \times \bar{X} \text{ s.t. } \bar{d}(x, y) \leq \varepsilon\}.$$

(iii) It is easy to see that

$$\text{Frag}(X/d, \mathcal{P}^-/d)/d_{\mathcal{P}^-}^{\text{Frag}} = \text{Frag}(X, \mathcal{P}^-)/d_{\mathcal{P}^-}^{\text{Frag}}.$$

Thus, one may without harm identify a semimetric space with its metric quotient before or after fragmentation (in fact the collection of the metric quotient of its components).

(iv) It is easy to see that if  $(X, d)$  is a length semimetric space,  $\delta \geq 0$  and  $\tilde{\mathcal{R}}$  is a multiset of elements of  $X^2$ , then  $\text{Coal}_\delta(X, \tilde{\mathcal{R}})$  is still a length semimetric space; see, for instance, [16], pages 62–63. Thus, cutting  $\text{Coal}_\delta(X, \tilde{\mathcal{R}})$  through Definition 2.21 is well-defined.

(v) One could have defined the cut a bit differently in order to keep the base set unchanged: one could have defined  $\mathcal{P}^{-.d}$  to be a new component of  $X$ , defining distance on it via an intrinsic formula. In the sequel,  $\mathcal{P}^{-.d}$  will have  $\mu$ -measure zero, so these two definitions lead to measured semimetric spaces which are at  $L_{\text{GHP}}$ -distance zero, and when coalescence subsequently occurs, it ignores  $\mathcal{P}^{-.d}$ .

DEFINITION 2.23. Let  $X = (X, d, \mu)$  be an m.s-m.s which is a length space. Let  $\ell$  be a diffuse  $\sigma$ -finite Borel measure on  $X$ . Let  $\mathcal{P}^-$  be a Poisson random set on  $X \times \mathbb{R}^+$  of intensity measure  $\ell \otimes \text{leb}_{\mathbb{R}^+}$ . The *fragmentation process started from  $X$* , denoted by  $(\text{Frag}(X, t))_{t \geq 0}$ , is the random process of m.s-m.s  $(\text{Frag}(X, \mathcal{P}_t^-))_{t \geq 0}$ . When  $X \in \mathcal{N}^{\text{graph}}$ , we shall always take  $\ell$  to be  $\ell_X$ , the length-measure on  $X$ .

REMARK 2.24.

(i) A similar fragmentation on the CRT is considered in [6].

(ii) Since  $\ell$  is a diffuse measure, almost surely  $\mu(\mathcal{P}_t^-) = 0$  for any  $t \geq 0$ . Thus we shall abuse notation and consider that  $\text{Frag}(X, \mathcal{P}_t^-)$  is still equipped with  $\mu$ , instead of  $\mu|_{X \setminus \mathcal{P}_t^-}$ .

(iii) Notice that this process inherits the strong Markov property from the strong Markov property of the Poisson process, and the fact that for  $A, B \subset X$ ,  $\text{Frag}(X, A \cup B) = \text{Frag}(\text{Frag}(X, A), B)$ ; cf. Lemma C.3.

Now, one wants to define dynamical percolation on measured length spaces by performing independently and simultaneously coalescence and fragmentation. One needs to be a bit careful here: when  $(X, d)$  is a geodesic space,  $A \subset X^2$  and  $B \subset X$ , even if  $B \cap \{x \in X : \exists y \in X, (x, y) \text{ or } (y, x) \in A\} = \emptyset$ , one cannot guarantee that  $\text{Coal}_0(\text{Frag}(X, B), A)$  is the same as  $\text{Frag}(\text{Coal}_0(X, A), B)$ . Indeed, let  $X = [0, 1]$  with the usual metric, let  $B = \{\frac{3}{2^n}, n \geq 2\}$  and  $A = \{(\frac{1}{2^{n+1}}, \frac{1}{2^n}), n \geq 0\}$ . Then there are two components in  $\text{Coal}_0(\text{Frag}(X, B), A)$ :  $\{0\}$  and  $]0, 1] \setminus B$ , whereas there is only one component in  $\text{Frag}(\text{Coal}_0(X, A), B)$ :  $[0, 1] \setminus B$ . However, it will be shown in Lemma 5.9 that if  $X \in \mathcal{S}^{\text{graph}}$ ,  $\mathcal{P}^+$  is as in Definition 2.13,  $\mathcal{P}^-$  as in Definition 2.23 then almost surely,

$$(2.5) \quad \forall t \geq 0, \quad \text{Coal}_0(\text{Frag}(X, \mathcal{P}_t^-), \mathcal{P}_t^+) = \text{Frag}(\text{Coal}_0(X, \mathcal{P}_t^+), \mathcal{P}_t^-).$$

This will rely on Lemma C.4, proved in Appendix C. Now, let us define dynamical percolation.

DEFINITION 2.25. Let  $X = (X, d, \mu)$  be an m.s-m.s which is a length space. Let  $\ell$  be a diffuse  $\sigma$ -finite Borel measure on  $X$ . Let  $\mathcal{P}^-$  be a Poisson random set on  $X \times \mathbb{R}^+$  of intensity measure  $\ell \otimes \text{leb}_{\mathbb{R}^+}$  and  $\mathcal{P}^+$  be a Poisson random set on  $X^2 \times \mathbb{R}^+$  of intensity measure  $\frac{1}{2}\mu^2 \times \text{leb}_{\mathbb{R}^+}$ . The *dynamical percolation process started from  $X$* , denoted by  $(\text{CoalFrag}(X, t))_{t \geq 0}$ , is the stochastic process  $(\text{Coal}_0(\text{Frag}(X, \mathcal{P}_t^-), \mathcal{P}_t^+))_{t \geq 0}$ .

Property (2.5) (when it holds!) shows that  $(\text{CoalFrag}(X, t))_{t \geq 0}$  inherits the strong Markov property from that of the Poisson processes.

2.8. *The scaling limit of critical Erdős–Rényi random graphs.* The scaling limit of critical Erdős–Rényi random graphs was obtained in [3], Theorem 24, for the Gromov–Hausdorff topology, and the result is extended to Gromov–Hausdorff–Prokhorov topology in [4], Theorem 4.1. Let  $\bar{\mathcal{G}}_{n,\lambda}$  denote the element of  $\mathcal{N}_2^{\text{graph}}$  obtained by replacing each edge of  $\mathcal{G}(n, p(\lambda, n))$  by an isometric copy of a segment of length  $n^{-1/3}$  (notably, the distance is the graph distance divided by  $n^{1/3}$ ) and choosing as measure the counting measure on vertices divided by  $n^{2/3}$ . [4], Theorem 4.1, and [5], Corollary 2, easily imply the following.

THEOREM 2.26 ([3, 4]). Let  $\lambda \in \mathbb{R}$  and  $p(\lambda, n) = \frac{1}{n} + \frac{\lambda}{n^{4/3}}$ . There is a random element  $\mathcal{G}_\lambda$  of  $\mathcal{N}_2^{\text{graph}}$  such that

$$\bar{\mathcal{G}}_{n,\lambda} \xrightarrow[n \rightarrow \infty]{(d)} \mathcal{G}_\lambda,$$

where the convergence in distribution is with respect to the  $L_{2,\text{GHP}}^{\text{surplus}}$ -topology.

We refer to [3] for the precise definition of the limit  $\mathcal{G}_\lambda$  and to [2] for various properties of  $\mathcal{G}_\lambda$ .

**3. Main results.** We shall distinguish between two types of results: general ones, and those emphasized in the **Introduction**, which are applications (usually not trivial ones) of the general results to Erdős–Rényi random graphs.

3.1. *General results: Feller and almost Feller properties.* In the course of proving the results for Erdős–Rényi random graphs, I tried to obtain more general results, such that one could apply the same technology to other sequences of random graphs, for instance those belonging to the basin of attraction of  $\mathcal{G}_\lambda$  (see [7] for this notion and Section 8 in the present paper for a more detailed discussion). This is reflected in what I call below Feller or almost Feller properties for coalescence, fragmentation, and dynamical percolation. Recall that one says that a Markov process has *the Feller property*<sup>3</sup> if the distribution at a fixed time  $t > 0$  is a continuous function of the distribution at time 0, where continuity is with respect to weak convergence of probability measures. The almost Feller properties below are variations on the Feller property, some of them weaker than a true Feller property in the sense that I need to add some condition in order to ensure convergence, but also a bit stronger in the sense that I added to the results the convergence of the whole process in the sense of the Skorokhod topology.

THEOREM 3.1 (Almost Feller property for coalescence). Let  $X^n = (X^n, d^n, \mu^n)$ ,  $n \geq 0$  be a sequence of random variables in  $\mathcal{S}$  and  $(\delta^n)_{n \geq 0}$  a sequence of nonnegative real numbers. Suppose that:

<sup>3</sup>Not to be confused with being a Feller process.

- (a)  $(X^n)$  converges in distribution (for  $L_{2,\text{GHP}}$ ) to  $X^\infty = (X^\infty, d^\infty, \mu^\infty)$  as  $n$  goes to infinity,
- (b)  $\delta^n \xrightarrow[n \rightarrow \infty]{} 0$ ,
- (c) For any  $\alpha > 0$  and any  $T > 0$ ,

$$(3.1) \quad \limsup_{n \in \mathbb{N}} \mathbb{P}(\text{supdiam}(\text{Coal}_{\delta^n}(X^n_{\leq \varepsilon}, T)) > \alpha) \xrightarrow[\varepsilon \rightarrow 0]{} 0.$$

Then,

- (i)  $(\text{Coal}_0(X^\infty, t))_{t \geq 0}$  is strong Markov with càdlàg trajectories in  $\mathcal{S}$ ,
- (ii)  $(\text{Coal}_{\delta^n}(X^n, t))_{t \geq 0}$  converges in distribution to  $(\text{Coal}_0(X^\infty, t))_{t \geq 0}$  (for  $L_{2,\text{GHP}}$ ),
- (iii) if  $t^n \xrightarrow[n \rightarrow \infty]{} t$ ,  $\text{Coal}_{\delta^n}(X^n, t^n)$  converges in distribution to  $\text{Coal}_0(X^\infty, t)$  (for  $L_{2,\text{GHP}}$ ).

Let us make two comments. The first is that there is a full Feller property on  $\mathcal{N}_1$ ; cf. Proposition 5.3. The second is that in the case of Erdős–Rényi random graphs, condition (3.1) will be handled through a general technical lemma, Lemma 5.13.

**THEOREM 3.2** (Feller property for fragmentation). *Let  $(G^n)_{n \geq 0}$  be a sequence of random variables in  $\mathcal{N}_2^{\text{graph}}$  converging in distribution to  $G$  in the  $L_{2,\text{GHP}}^{\text{surplus}}$  metric. Then:*

- (i)  $(\text{Frag}(G, t))_{t \geq 0}$  is strong Markov with càdlàg trajectories (for  $L_{2,\text{GHP}}^{\text{surplus}}$ ) in  $\mathcal{N}_2^{\text{graph}}$ ,
- (ii)  $(\text{Frag}(G^n, t))_{t \geq 0}$  converges in distribution to  $(\text{Frag}(G, t))_{t \geq 0}$  (for  $L_{2,\text{GHP}}^{\text{surplus}}$ ),
- (iii) if  $t^n \xrightarrow[n \rightarrow \infty]{} t$ , then  $\text{Frag}(G^n, t^n)$  converges in distribution to  $\text{Frag}(G, t)$  (for  $L_{2,\text{GHP}}^{\text{surplus}}$ ).

**THEOREM 3.3** (Almost Feller property for dynamical percolation). *Let  $(X^n)_{n \geq 0}$  be a sequence of random variables in  $\mathcal{S}^{\text{graph}}$  converging in distribution to  $X^\infty$  in the  $L_{2,\text{GHP}}^{\text{surplus}}$  metric. Suppose also that for any  $\alpha > 0$  and any  $T \geq 0$ ,*

$$(3.2) \quad \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow +\infty} \mathbb{P}(\text{suplength}(\text{Coal}_0(X^n_{\leq \varepsilon}, T)) > \alpha) = 0.$$

Then,

- (i)  $(\text{CoalFrag}(X^\infty, t))_{t \geq 0}$  is strong Markov with càdlàg trajectories (for  $L_{2,\text{GHP}}$ ) in  $\mathcal{S}^{\text{length}}$ ,
- (ii)  $(\text{CoalFrag}(X^n, t))_{t \geq 0}$  converges in distribution to  $(\text{CoalFrag}(X^\infty, t))_{t \geq 0}$  (for  $L_{2,\text{GHP}}$ ),
- (iii) if  $t^n \xrightarrow[n \rightarrow \infty]{} t$ , then  $\text{CoalFrag}(X^n, t^n)$  converges in distribution to  $\text{CoalFrag}(X^\infty, t)$  for  $L_{2,\text{GHP}}$ .

A caveat is in order here: if  $X$  belongs to  $\mathcal{S}^{\text{graph}}$ ,  $\text{Coal}_0(X, t)$  does not necessarily belong to  $\mathcal{S}^{\text{graph}}$  (but it does belong to  $\mathcal{S}^{\text{length}}$ ) since a component with an infinite surplus might form. In order to stay with components which are real graphs, one needs, and it is sufficient, the total mass of the components with positive surplus to be finite; cf. [8]. See Section 8 for more details. Consequently,  $\text{CoalFrag}$  does not define a Markov semigroup on  $\mathcal{S}^{\text{graph}}$ .

In the same vein, notice that in Theorem 3.3, the initial convergence is in  $L_{2,\text{GHP}}^{\text{surplus}}$  and the conclusion is in  $L_{2,\text{GHP}}$ . This is unavoidable, for the same reason as above: convergence in  $L_{2,\text{GHP}}^{\text{surplus}}$  does not prevent the sequence  $X^n$  from having an infinite number of components with positive surplus whose total mass diverges when  $n$  goes to infinity. These components can at positive time be glued to large components, augmenting their surplus indefinitely. See Section 8 for more details.

Finally, we shall prove a structural lemma for the multiplicative coalescent, Lemma 5.5. It is at the base of Theorems 3.1 and 3.3. Since its statement is too technical, we do not state it here, but I see it as one of the main results of the article.

3.2. *Main results for Erdős–Rényi random graphs.* The main results for Erdős–Rényi random graphs are the following.

**THEOREM 3.4.**  *$(\text{Coal}_0(\mathcal{G}_\lambda, t))_{t \geq 0}$ ,  $(\text{Frag}(\mathcal{G}_\lambda, t))_{t \geq 0}$  and  $(\text{CoalFrag}(\mathcal{G}_\lambda, t))_{t \geq 0}$  are strong Markov processes with càdlàg trajectories (for  $L_{2, \text{GHP}}$ ) in  $\mathcal{S}^{\text{graph}}$ .*

**THEOREM 3.5.** *Let  $(\mathbf{G}^{n, \lambda, +}(t), t \geq 0)$  be the discrete coalescence process of intensity  $n^{-4/3}$ , started at  $\mathcal{G}(n, p(\lambda, n))$ , equipped with the graph distance multiplied by  $n^{-1/3}$  and the counting measure on vertices multiplied by  $n^{-2/3}$ . Then the sequence of processes  $(\mathbf{G}^{n, \lambda, +}(t), t \geq 0)$  converges to  $(\text{Coal}_0(\mathcal{G}_\lambda, t), t \geq 0)$  for  $L_{2, \text{GHP}}$  as  $n$  goes to infinity.*

**THEOREM 3.6.** *Let  $(\mathbf{G}^{n, \lambda, -}(t), t \geq 0)$  be the discrete fragmentation process of intensity  $n^{-1/3}$  started at  $\mathcal{G}(n, p(\lambda, n))$ , equipped with the graph distance multiplied by  $n^{-1/3}$  and the counting measure on vertices multiplied by  $n^{-2/3}$ . Then the sequence of processes  $(\mathbf{G}^{n, \lambda, -}(t), t \geq 0)$  converges to  $(\text{Frag}(\mathcal{G}_\lambda, t), t \geq 0)$  for  $L_{2, \text{GHP}}^{\text{surplus}}$  as  $n$  goes to infinity.*

**THEOREM 3.7.** *Let  $(\mathbf{G}^{n, \lambda}(t), t \geq 0)$  be the dynamical percolation processes of parameter  $p(\lambda, n)$  and intensity  $n^{-1/3}$  started with  $\mathcal{G}(n, p(\lambda, n))$ , equipped with the graph distance multiplied by  $n^{-1/3}$  and the counting measure on vertices multiplied by  $n^{-2/3}$ . Then the sequence of processes  $(\mathbf{G}^{n, \lambda}(t), t \geq 0)$  converges to  $(\text{CoalFrag}(\mathcal{G}_\lambda, t), t \geq 0)$  for  $L_{2, \text{GHP}}$  as  $n$  goes to infinity.*

The rate  $n^{-4/3}$  for discrete coalescence compared to  $n^{-1/3}$  for discrete fragmentation and dynamical percolation process might be disturbing. This is due to the fact that when one performs dynamical percolation, one resamples the state of the edges. Thus, a specific edge appears at rate  $n^{-1/3}p$ , which is essentially  $n^{-4/3}$ , and disappears at rate  $n^{-1/3}(1 - p)$ , which is essentially  $n^{-1/3}$ .

The last result we shall mention is the important fact that on  $(\mathcal{G}_\lambda)_{\lambda \in \mathbb{R}}$ , fragmentation is the time-reversal of coalescence.

**PROPOSITION 3.8.** *For any  $\lambda \in \mathbb{R}$  and  $s \in \mathbb{R}^+$ ,  $(\mathcal{G}_\lambda, \text{Coal}_0(\mathcal{G}_\lambda, s))$  and  $(\text{Frag}(\mathcal{G}_{\lambda+s}, s), \mathcal{G}_{\lambda+s})$  have the same distribution.*

3.3. *Description of the rest of the article.* Section 4 contains some preliminary results: lemmas from [5] and technical results about the distance  $L_{\text{GHP}}$ .

Section 5 is devoted to the proofs of the main results for coalescence: Theorem 3.1 (the almost Feller property) and Theorem 3.5. Probably the most important work lies inside Lemma 5.5, which is a statement about the structure of a (multi)graph, denoted below by  $\mathcal{W}(x, t)$ , which is central in Aldous’ study of the multiplicative coalescent. It allows notably to reduce the proof of the Feller property from  $\mathcal{N}_2$  to  $\mathcal{N}_1$ , where it is much easier to prove.

Section 6 is devoted to the main results for fragmentation: Theorem 3.2 (the Feller property) and Theorem 3.6. We shall also show that for  $\mathcal{G}_\lambda$ , coalescence is the time-reversal of fragmentation, which is Proposition 3.8.

Section 7 is devoted to the proofs of the main results for dynamical percolation, Theorem 3.3 (the almost Feller property) and Theorem 3.7. We also prove Theorem 3.4 there, that is, the fact that the coalescence, fragmentation and dynamical percolation on  $\mathcal{G}_\lambda$  define processes in  $\mathcal{S}^{\text{graph}}$ .

We finish the article by some perspectives in Section 8, and some technical tools are gathered in the Appendix.

**4. Preliminary results and tools.**

4.1. *The multiplicative coalescent.* The main tool to analyze our coalescent and fragmentation processes will be a refinement of Aldous’ work [5] on the multiplicative coalescent. In this section, we recall what we will use of his work.

Let  $(N_{i,j})_{i,j \in \mathbb{N}^*}$  be independent Poisson point processes on the real line with intensity 1. Denote by  $T_{i,j,n}$  the  $n$ th jump-time of  $N_{i,j}$ . For  $x \in \ell^2_+$ , let  $\mathbb{M}\mathbb{G}(x, t)$  denote the weighted multigraph (with loops) with vertex set  $\mathbb{N}^*$ , weight  $x_i$  on vertex  $i$  and edge multiset:

$$\bigsqcup_{n,i,j \in \mathbb{N}^*} \left\{ \left\{ \{i, j\} \text{ s.t. } T_{i,j,n} \leq \frac{t}{2} x_i x_j \right\} \right\}.$$

If  $x := \text{masses}(\mathbf{X})$ , this multigraph is of course closely related to the multigraph  $\mathbb{M}\mathbb{G}(\mathbf{X}, t)$  which was defined in Section 2.5.2, in the sense that collapsing each component of  $\mathbf{X}$  to a single point and enumerating the components in decreasing order of size transforms  $\mathbb{M}\mathbb{G}(\mathbf{X}, \cdot)$  into a process with the same distribution as  $\mathbb{M}\mathbb{G}(x, \cdot)$ . Furthermore, if one forgets loops and transforms any multiple edge into a single edge,  $\mathbb{M}\mathbb{G}(x, t)$  becomes  $\mathcal{W}(x, t)$ , the nonuniform random graph of [5], Section 1.4. Indeed, the graph  $\mathcal{W}(x, t)$  has set of vertices  $\mathbb{N}^*$  and each pair  $\{i, j\}$  is an edge with probability  $1 - \exp(-tx_i x_j)$ , independently for distinct pairs. The size, or mass, of a connected component of those (multi-)graphs is defined to be the sum of weights  $x_i$  for  $i$  in this component. Denoting by  $X(x, t)$  the sequence of sizes, listed in decreasing order, of the connected components of  $\mathcal{W}(x, t)$ , Aldous proved in [5], Proposition 5, that  $\{X(x, t) : t \geq 0\}$  defines a Markov process on  $\ell^2_{\searrow}$  which possesses the Feller property. In words, the dynamics is as follows: two distinct connected components  $c$  and  $c'$  with sizes  $m$  and  $m'$  merge at a rate proportional to  $mm'$ , the product of their sizes, whence the term *multiplicative coalescent*.

Following Aldous, we denote by  $S(x, t)$  the sum of squares of the sizes of the components of  $\mathbb{M}\mathbb{G}(x, t)$ . We shall use later the following lemmas.

LEMMA 4.1 ([5], Lemma 20). For  $x$  in  $\ell^2_{\searrow}$ ,

$$\mathbb{P}(S(x, t) > s) \leq \frac{tsS(x, 0)}{s - S(x, 0)}, \quad s > S(x, 0).$$

LEMMA 4.2 ([5], Lemma 23). Let  $(z_i, 1 \leq i \leq n)$  be strictly positive vertex weights, and let  $1 \leq m < n$ . Consider the bipartite random graph  $\mathcal{B}$  on vertices  $\{1, 2, \dots, m\} \cup \{m + 1, \dots, n\}$  defined by: for each pair  $(i, j)$  with  $1 \leq i \leq m < j \leq n$ , the edge  $\{i, j\}$  is present with probability  $1 - \exp(-tz_i z_j)$ , independently for different pairs. Write  $\alpha_1 = \sum_{i=1}^m z_i^2, \alpha_2 = \sum_{i=m+1}^n z_i^2$ . Let  $(Z_i)$  be the sizes of the components of  $\mathcal{B}$ . Then

$$\varepsilon \mathbb{P}\left(\sum_i Z_i^2 \geq \alpha_1 + \varepsilon\right) \leq (1 + t(\alpha_1 + \varepsilon))^2 \alpha_2, \quad \varepsilon > 0.$$

REMARK 4.3. As noticed in [5], page 842, Lemma 4.2 extends to  $z \in \ell^2$ . Notice that in [5], Lemma 23, the upper bound was stated as  $(2t(\alpha_1 + \varepsilon) + (t(\alpha_1 + \varepsilon))^2)\alpha_2$ , but this cannot be true, as can be seen by taking  $t = 0$ . In fact a term  $z_{m+1}^2$  is missing in the right-hand side of line –3 in [5], page 841, which impacts all subsequent inequalities. Correcting this slight mistake leads to the bound above.

In [5], this lemma is used in conjunction with the following one.



LEMMA 4.4 ([5], Lemma 17). *Let  $\tilde{G}$  be a graph with vertex weights  $(\tilde{x}_i)$ . Let  $G$  be a subgraph of  $\tilde{G}$  (i.e., they have the same vertex-sets and each edge of  $G$  is an edge of  $\tilde{G}$ ) with vertex weights  $x_i \leq \tilde{x}_i$ . Let  $\tilde{a}$  and  $a$  be the decreasing orderings of the component sizes of  $\tilde{G}$  and  $G$ . Then*

$$\|\tilde{a} - a\|_2 \leq \sum_i \tilde{a}_i^2 - \sum_i a_i^2$$

provided  $\sum_i a_i^2 < \infty$ .

Finally, the following lemma will be useful to prove convergence in the sense of Skorokhod: it will allow us to control the  $\ell^2$  distance between  $X(x, t')$  and  $X(x_{\leq \varepsilon}, t')$  for  $t' \leq t$  by the  $\ell^2$ -distance at time  $t$ .

LEMMA 4.5. *Let  $G = (V, E)$  be a multigraph whose vertices have weights  $(x_i)_{i \in V}$ . If  $W \subset V$  and  $E' \subset E$ , let  $\text{comp}(W, E')$  denote the set of connected components of the graph  $(W, E' \cap \binom{W}{2})$  and define*

$$S(W, E') := \sum_{m \in \text{comp}(W, E')} \left( \sum_{i \in m} x_i \right)^2.$$

Now, let  $W \subset V$  be such that no point of  $V \setminus W$  belongs to a cycle in  $(V, E)$ . Then, for any  $E' \subset E$ ,

$$S(V, E) - S(W, E) \geq S(V, E') - S(W, E'),$$

provided  $S(W, E) < \infty$ .

PROOF. For  $i$  and  $j$  in  $V$  and  $E' \subset E$ , we denote by “ $i \sim j \in E'$ ” the fact that  $i$  and  $j$  are distinct and connected by a path in  $(V, E')$  and by “ $i \sim j \notin E'$ ” the negation of the previous statement, that is, that  $i = j$  or  $i$  and  $j$  are not connected by a path in  $(V, E')$ . The hypothesis on  $W$  implies that for any  $E' \subset E$ ,

$$(4.1) \quad \forall i, j \in W, \left( \begin{array}{c} i \sim j \in E' \\ \text{and} \\ i \sim j \notin E' \cap \binom{W}{2} \end{array} \right) \Rightarrow \left( \begin{array}{c} i \sim j \in E \\ \text{and} \\ i \sim j \notin E \cap \binom{W}{2} \end{array} \right).$$

Indeed, suppose that  $i$  and  $j$  satisfy the left-hand side of the implication above. Since  $E' \subset E$ , we have  $i \sim j \in E$ . From the hypothesis on  $i$  and  $j$ , there is a simple path  $\gamma$  from  $i$  to  $j$  in  $E'$  with a point  $z \in \gamma \cap (V \setminus W)$ . Denote by  $i'$  the point on  $\gamma$  just before  $z$  and by  $j'$  the point on  $\gamma$  just after  $z$ . Now, suppose that we had  $i \sim j \in E \cap \binom{W}{2}$ . We would have a path  $\gamma'$  in  $(W, E \cap \binom{W}{2})$  from  $i$  to  $j$ . Concatenating the portion of  $\gamma$  from  $i$  to  $i'$ , the portion of  $\gamma$  from  $j$  to  $j'$  and the path  $\gamma'$ , we see that there is a path from  $i'$  to  $j'$  in  $(V, E)$  which avoids  $z$ . From this path, we can extract a simple path in  $(V, E)$  connecting  $i'$  to  $j'$ , still avoiding  $z$ . Now,  $z$  is a neighbor of  $i'$  and  $j'$  in  $(V, E)$ , and this gives a cycle going through  $z$ . Since  $z$  is a point of  $V \setminus W$ , this contradicts the hypothesis on  $W$ . Thus,  $i \sim j \notin E \cap \binom{W}{2}$  holds and (4.1) is true.

Now,

$$S(V, E') = \sum_{i \in V} x_i^2 + \sum_{\substack{i, j \in V \\ i \sim j \in E'}} x_i x_j$$

$$\begin{aligned}
 &= S(W, E') + \sum_{i \in V \setminus W} x_i^2 + \sum_{\substack{i, j \in V \\ i \sim j \in E'}} x_i x_j - \sum_{\substack{i, j \in W \\ i \sim j \in E' \cap \binom{W}{2}}} x_i x_j \\
 &= S(W, E') + \sum_{i \in V \setminus W} x_i^2 + \sum_{\substack{i, j \in V \\ i \sim j \in E' \text{ and } i \sim j \notin E' \cap \binom{W}{2}}} x_i x_j.
 \end{aligned}$$

Now, (4.1) shows that the last sum on the right of the last equation is increasing in  $E' \subset E$ , and this shows the result.  $\square$

4.2. *Tools to handle  $L_{GHP}$ .* The following proposition is analogous to similar results concerning vague convergence of locally finite measures (see [25], Section 4). With this proposition in hand, it is easy to show, for instance, that if  $\nu$  is made of a countable number of components with strictly distinct positive masses, then  $\nu_n$  converges to  $\nu$  is equivalent to the fact that for every  $k$ , the  $k$ th largest component of  $\nu_n$  converges, for  $d_{GHP}$ , to the  $k$ th largest component of  $\nu$ .

PROPOSITION 4.6. *( $\mathcal{N}, L_{GHP}$ ) is a complete separable metric space, and if  $\nu_n, n \geq 0$  and  $\nu$  are elements of  $\mathcal{N}$ ,  $(\nu_n)_{n \geq 0}$  converges to  $\nu$  if and only if for every  $\varepsilon > 0$  such that  $\nu(\{(X, d, \mu) \in \mathcal{M} : \mu(X) = \varepsilon\}) = 0$ ,  $\delta_{LP}(\nu_n, \nu) \rightarrow 0$  as  $n$  goes to infinity.*

PROOF. The fact that  $L_{GHP}$  is a metric is left to the reader. Let  $D$  be a countable set dense in  $\{(X, d, \mu) \in \mathcal{M} : \mu(X) \neq 0\}$ . Let  $\mathcal{D} := \{\sum_{k=1}^m \delta_{X_k} : m \in \mathbb{N}, X_1, \dots, X_m \in D\}$ . It is easy to show that  $\mathcal{D}$  is dense in  $(\mathcal{N}, L_{GHP})$ . This shows separability.

Now, suppose that  $\nu^n$  is a Cauchy sequence for  $L_{GHP}$ . Then, for any  $k \geq 1$ ,  $f_k \nu^n$  is a Cauchy sequence of finite measures for  $\delta_{LP}$ . From the completeness of the Lévy–Prokhorov distance on finite measures on a Polish space we get that for each  $k$ , there is some measure  $\nu_k$  such that

$$\delta_{LP}(f_k \nu^n, \nu_k) \xrightarrow{n \rightarrow \infty} 0.$$

Notice that if  $2 \leq k \leq l$ ,  $\nu_k = \nu_l$  on  $\mathcal{M}_{> \frac{1}{k-1}}$ . Define

$$\nu = \sup_{k \geq 0} \mathbf{1}_{\mathcal{M}_{> \frac{1}{k+1}}} \nu_{k+2}$$

so that for any  $k \geq 1$ ,  $f_k \nu = \nu_k$ . Then,  $\nu$  is an element of  $\mathcal{N}$  and

$$L_{GHP}(\nu^n, \nu) \xrightarrow{n \rightarrow \infty} 0,$$

showing the completeness of  $(\mathcal{N}, L_{GHP})$ .

Finally, suppose that  $L_{GHP}(\nu^n, \nu)$  goes to zero as  $n$  goes to infinity and let  $\varepsilon > 0$  be such that  $\nu(\{(X, d, \mu) \in \mathcal{M} : \mu(X) = \varepsilon\}) = 0$ . Then, for any  $\alpha > 0$ , let  $k$  be such that

$$\frac{1}{k} \leq \varepsilon$$

let  $N$  be such that

$$\forall n \geq N, \forall A \in \mathcal{B}(\mathcal{M}), \quad f_k \nu^n(A) \leq f_k \nu^n(A^\alpha) + \alpha \quad \text{and} \quad f_k \nu^n(A) \leq f_k \nu(A^\alpha) + \alpha.$$

Then, for  $n \geq N$  and  $B \in \mathcal{B}(\mathcal{M}_{> \varepsilon})$ ,

$$\begin{aligned}
 \nu_{> \varepsilon}(B) &\leq \nu(B \cap \mathcal{M}_{> \varepsilon + \alpha}) + \nu(\mathcal{M}_{> \varepsilon} \setminus \mathcal{M}_{> \varepsilon + \alpha}) \\
 &= f_k \nu(B \cap \mathcal{M}_{> \varepsilon + \alpha}) + \nu(\mathcal{M}_{> \varepsilon} \setminus \mathcal{M}_{> \varepsilon + \alpha})
 \end{aligned}$$

$$\begin{aligned} &\leq f_k v^n((B \cap \mathcal{M}_{>\varepsilon+\alpha})^\alpha) + \alpha + v(\mathcal{M}_{>\varepsilon} \setminus \mathcal{M}_{>\varepsilon+\alpha}) \\ &= v_{>\varepsilon}^n((B \cap \mathcal{M}_{>\varepsilon+\alpha})^\alpha) + \alpha + v(\mathcal{M}_{>\varepsilon} \setminus \mathcal{M}_{>\varepsilon+\alpha}) \\ &\leq v_{>\varepsilon}^n(B^\alpha) + \alpha + v(\mathcal{M}_{>\varepsilon} \setminus \mathcal{M}_{>\varepsilon+\alpha}), \end{aligned}$$

where we used the fact that  $(B \cap \mathcal{M}_{>\varepsilon+\alpha})^\alpha \subset \mathcal{M}_{>\varepsilon}$  and  $f_k$  equals 1 on  $\mathcal{M}_{>\varepsilon}$ . Also, for  $n \geq N$  and  $B \in \mathcal{B}(\mathcal{M}_{>\varepsilon})$ ,

$$\begin{aligned} v_{>\varepsilon}^n(B) &= f_k v^n(B) \\ &\leq f_k v(B^\alpha) + \alpha \\ &\leq v(B^\alpha) + \alpha \\ &\leq v(B^\alpha \cap \mathcal{M}_{>\varepsilon}) + v(\mathcal{M}_{>\varepsilon-\alpha} \setminus \mathcal{M}_{>\varepsilon}) + \alpha \\ &= v_{>\varepsilon}(B^\alpha) + v(\mathcal{M}_{>\varepsilon-\alpha} \setminus \mathcal{M}_{>\varepsilon}) + \alpha. \end{aligned}$$

To complete the proof, note that since  $v(\{(X, d, \mu) \in \mathcal{M} : \mu(X) = \varepsilon\}) = 0$ , then

$$v(\mathcal{M}_{>\varepsilon} \setminus \mathcal{M}_{>\varepsilon+\alpha}) + v(\mathcal{M}_{>\varepsilon-\alpha} \setminus \mathcal{M}_{>\varepsilon}) \xrightarrow{\alpha \rightarrow 0} 0. \quad \square$$

We shall always use the following lemmas to bound  $L_{\text{GHP}}$  from above.

LEMMA 4.7. *Let  $X = (X, d, \mu)$  and  $X' = (X', d', \mu')$  belong to  $\mathcal{N}$  and fix  $\varepsilon \in ]0, \frac{1}{2}]$ . Suppose there exist two injective maps*

$$\sigma : \text{comp}(X_{>\varepsilon}) \rightarrow \text{comp}(X') \quad \text{and} \quad \sigma' : \text{comp}(X'_{>\varepsilon}) \rightarrow \text{comp}(X)$$

such that

$$\forall m \in \text{comp}(X_{>\varepsilon}), \quad d_{\text{GHP}}(m, \sigma(m)) \leq \alpha$$

and

$$\forall m' \in \text{comp}(X'_{>\varepsilon}), \quad d_{\text{GHP}}(m', \sigma'(m')) \leq \alpha.$$

Then

$$L_{\text{GHP}}(v_X, v_{X'}) \leq 8\alpha \# \text{comp}(X_{>\varepsilon-\alpha}) + 16\varepsilon,$$

and if  $\varepsilon > \alpha$ , for  $p > 0$ ,

$$L_{\text{GHP}}(v_X, v_{X'}) \leq 8\alpha \frac{\sum_{m \in \text{comp}(X)} \mu(m)^p}{(\varepsilon - \alpha)^p} + 16\varepsilon.$$

PROOF. Consider any  $\varepsilon_0 \geq \varepsilon$ . For a component  $m$  in  $\text{comp}(X_{>\varepsilon_0})$ , the difference between the masses  $\mu(m)$  and  $\mu'(\sigma(m))$  is at most  $\alpha$ , and the same holds between  $m'$  and  $\sigma'(m')$  when  $m' \in \text{comp}(X'_{>\varepsilon_0})$ . Thus  $\sigma'$  sends  $\text{comp}(X'_{>\varepsilon_0})$  into  $\text{comp}(X_{>\varepsilon_0-\alpha})$  and

$$\#\{m \in \text{comp}(X'_{>\varepsilon_0})\} \leq \#\{m \in \text{comp}(X_{>\varepsilon_0-\alpha})\}.$$

Now, let  $k \geq 1$  be such that

$$\frac{1}{k+1} \geq \varepsilon.$$

Let  $B \in \mathcal{B}(\mathcal{M}_{>\frac{1}{k+1}})$ . Then, for any  $m$  in  $\text{comp}(X_{>\frac{1}{k+1}}) \cap B$ ,  $\sigma(m)$  belongs to  $\text{comp}(X') \cap B^\alpha$ . Then notice that  $f_k$  is  $k(k+1)$ -Lipschitz.

$$\begin{aligned} f_k \nu_X(B) &= \sum_{m \in \text{comp}(X_{>\frac{1}{k+1}}) \cap B} f_k(m) \\ &\leq \sum_{m \in \text{comp}(X_{>\frac{1}{k+1}}) \cap B} f_k(\sigma(m)) + \alpha k(k+1) \# \text{comp}(X_{>\frac{1}{k+1}}) \\ &\leq \sum_{\substack{m' \in \text{comp}(X') \\ m' \in B^\alpha}} f_k(m') + \alpha k(k+1) \# \text{comp}(X_{>\frac{1}{k+1}}) \\ &= f_k \nu_{X'}(B^\alpha) + \alpha k(k+1) \# \text{comp}(X_{>\frac{1}{k+1}}) \end{aligned}$$

and symmetrically

$$f_k \nu_{X'}(B) \leq f_k \nu_X(B^\alpha) + \alpha k(k+1) \# \text{comp}(X'_{>\frac{1}{k+1}}).$$

Thus, for any  $k \geq 1$  such that  $\frac{1}{k+1} \geq \varepsilon$ ,

$$\begin{aligned} \delta_{\text{LP}}(f_k \nu_X, f_k \nu_{X'}) &\leq \alpha k(k+1) \# \text{comp}(X_{>\frac{1}{k+1}}) \vee \# \text{comp}(X'_{>\frac{1}{k+1}}) \\ &\leq \alpha k(k+1) \# \text{comp}(X_{>\varepsilon-\alpha}). \end{aligned}$$

Thus,

$$\begin{aligned} L_{\text{GHP}}(\nu_X, \nu_{X'}) &\leq \alpha \# \text{comp}(X_{>\varepsilon-\alpha}) \sum_{k < \frac{1}{\varepsilon}-1} 2^{-k} k(k+1) + \sum_{k \geq \frac{1}{\varepsilon}-1} 2^{-k} \\ &\leq 8\alpha \# \text{comp}(X_{>\varepsilon-\alpha}) + 16\varepsilon. \end{aligned} \quad \square$$

If  $X$  and  $X'$  are two m.s-m.s with a finite number of finite components, one may measure their distance with  $d_{\text{GHP}}$  (using Definition 2.5), with  $L_{\text{GHP}}$  (using Definition 2.9) or with

$$1 \wedge \inf_{\sigma} \sup_{m \in \text{comp}(X)} d_{\text{GHP}}(m, \sigma(m)) = 1 \wedge \delta_{\text{LP}}(\nu_X, \nu_{X'}),$$

where the infimum is over bijections  $\sigma$  between  $\text{comp}(X)$  and  $\text{comp}(X')$ . Those three distances do not necessarily coincide, and the following lemma clarifies the links between them.

LEMMA 4.8. *Let  $X = (X, d, \mu)$  and  $X' = (X', d', \mu')$  be two m.s-m.s in  $\mathcal{N}$  with a finite number of components.*

(i) *If  $d_{\text{GHP}}(X, X') < \infty$ , then there is a bijection  $\sigma$  from  $\text{comp}(X)$  to  $\text{comp}(X')$  such that:*

$$\forall m \in \text{comp}(X), \quad d_{\text{GHP}}(m, \sigma(m)) \leq 2d_{\text{GHP}}(X, X'),$$

and thus,

$$L_{\text{GHP}}(X, X') \leq 16d_{\text{GHP}}(X, X') \# \text{comp}(X).$$

(ii) *If there exists a bijection  $\sigma$  from  $\text{comp}(X)$  to  $\text{comp}(X')$  such that*

$$\sup_{m \in \text{comp}(X)} d_{\text{GHP}}(m, \sigma(m)) < \infty,$$

then

$$d_{\text{GHP}}(X, X') \leq \sup_{m \in \text{comp}(X)} d_{\text{GHP}}(m, \sigma(m)) \# \text{comp}(X),$$

and

$$L_{\text{GHP}}(X, X') \leq 8 \sup_{m \in \text{comp}(X)} d_{\text{GHP}}(m, \sigma(m)) \# \text{comp}(X).$$

PROOF. *Proof of (i).* Suppose that  $d_{\text{GHP}}(X, X') < \varepsilon < \infty$ . Let  $\mathcal{R} \in C(X, X')$  and  $\pi \in M(X, X')$  be such that

$$D(\pi; \mu, \mu') \vee \frac{1}{2} \text{dis}(\mathcal{R}) \vee \pi(\mathcal{R}^c) \leq \varepsilon.$$

Since  $\mathcal{R}$  has finite distortion,

$$\forall (x, x'), (y, y') \in \mathcal{R}, \quad d(x, y) = +\infty \Leftrightarrow d'(x', y') = +\infty$$

which shows that each component  $m$  of  $X$  (resp.,  $X'$ ) is in correspondence through  $\mathcal{R}$  with exactly one component  $\sigma(m)$  of  $X'$  (resp.,  $X$ ).  $\sigma$  is thus a bijection, and  $\mathcal{R} \cap m \times \sigma(m) \in C(m, \sigma(m))$  and has distortion at most  $2\varepsilon$ . Furthermore,

$$\pi|_{m \times \sigma(m)}((\mathcal{R} \cap m \times \sigma(m))^c) = \pi(\mathcal{R}^c \cap m \times \sigma(m)) \leq \varepsilon.$$

Finally, for any  $A \in \mathcal{B}(m)$ ,

$$\begin{aligned} |\pi|_{m \times \sigma(m)}(A \times \sigma(m)) - \mu|_m(A)| &\leq |\pi(A \times X') - \mu(A)| + \pi(A \times \sigma(m)^c) \\ &\leq \varepsilon + \pi(\mathcal{R}^c) \\ &\leq 2\varepsilon \end{aligned}$$

and similarly, for any  $A' \in \mathcal{B}(\sigma(m))$ ,

$$|\pi|_{m \times \sigma(m)}(m \times A') - \mu'|_{\sigma(m)}(A')| \leq 2\varepsilon.$$

Thus,

$$\forall m \in \text{comp}(X), \quad d_{\text{GHP}}(m, \sigma(m)) \leq 2\varepsilon,$$

and the consequence on  $L_{\text{GHP}}(X, X')$  comes from Lemma 4.7 applied for any  $\varepsilon > 0$  small enough, and letting  $\varepsilon$  go to zero.

*Proof of (ii).* Suppose that there exists a bijection  $\sigma$  from  $\text{comp}(X)$  to  $\text{comp}(X')$  such that

$$\sup_{m \in \text{comp}(X)} d_{\text{GHP}}(m, \sigma(m)) \leq \varepsilon < \infty.$$

Then, for any  $m$ , let  $\mathcal{R}_m \in C(m, \sigma(m))$  and  $\pi_m \in M(m, \sigma(m))$  be such that

$$D(\pi_m; \mu|_m, \mu'|_{\sigma(m)}) \vee \frac{1}{2} \text{dis}(\mathcal{R}_m) \vee \pi_m(\mathcal{R}_m^c) \leq \varepsilon.$$

Let  $\pi = \sum_{m \in \text{comp}(X)} \pi_m$  and  $\mathcal{R} = \bigcup_{m \in \text{comp}(X)} \mathcal{R}_m$ . Then  $\mathcal{R}$  is a correspondence between  $X$  and  $X'$ ,

$$\begin{aligned} \frac{1}{2} \text{dis}(\mathcal{R}) &\leq \frac{1}{2} \sup_m \text{dis}(\mathcal{R}_m) \leq \varepsilon, \\ \pi(\mathcal{R}^c) &= \sum_m \pi_m(\mathcal{R}_m^c) \leq \# \text{comp}(X) \varepsilon. \end{aligned}$$

Furthermore, for any  $A \in \mathcal{B}(X)$ ,

$$\begin{aligned} |\pi(A \times X') - \mu(A)| &\leq \sum_{m \in \text{comp}(X)} |\pi((A \cap m) \times X') - \mu(A \cap m)| \\ &= \sum_{m \in \text{comp}(X)} |\pi((A \cap m) \times \sigma(m)) - \mu(A \cap m)| \\ &\leq \#\text{comp}(X)\varepsilon \end{aligned}$$

and symmetrically, for any  $A' \in \mathcal{B}(X')$ ,

$$|\pi(X \times A') - \mu'(A')| \leq \#\text{comp}(X)\varepsilon.$$

Thus

$$D(\pi; \mu, \mu') \leq \#\text{comp}(X)\varepsilon,$$

and we get

$$d_{\text{GHP}}(X, X') \leq \#\text{comp}(X)\varepsilon.$$

The statement on  $L_{\text{GHP}}(X, X')$  follows from the hypothesis and Lemma 4.7 applied for any  $\varepsilon > 0$  small enough, and letting  $\varepsilon$  go to zero.  $\square$

Let us end this section with two remarks. First, notice that  $L_{\text{GHP}}$  makes sense even between counting measures whose atoms are semimetric spaces. Notice also that  $\delta_{\text{LP}}(\nu_X, 0)$  is at most the number of connected components of  $X$ . Thus

$$L_{\text{GHP}}(\nu_X, 0) \leq 2^{2 - (\sup_{m \in \text{comp}(X)} \mu(m))^{-1}}.$$

One sees thus that if  $X^n$  is a sequence of m.s-m.s such that

$$\sup_{m \in \text{comp}(X^n)} \mu(m) \xrightarrow{n \rightarrow +\infty} 0$$

and whatever the diameters of the components are, then  $\nu_{X^n}$  converges to zero for  $L_{\text{GHP}}$ , which can be seen as an empty collection of measured metric spaces. Notably,  $\text{supdiam}$  is not continuous with respect to the  $L_{\text{GHP}}$ -distance.

**5. Proofs of the main results for coalescence.**

5.1. *The coalescent on  $\mathcal{N}_1$ .* On  $\mathcal{N}_1$ , coalescence behaves very gently since there is a finite number of coalescence events in any finite time interval. Notably, for  $X \in \mathcal{N}_1$ ,  $(\text{Coal}_\delta(X, t))_{t \geq 0}$  is clearly càdlàg. The aim of this section is to prove Proposition 5.3, which is essentially a Feller property, together with a variant, Proposition 5.4.

LEMMA 5.1. *Let  $\varepsilon \in ]0, 1[$ ,  $X = (X, d, \mu)$  and  $X' = (X', d', \mu')$  be two m.s-m.s with a finite number of finite components.*

*Suppose that  $P = \{(x_i, y_i), 1 \leq i \leq k\}$  are pairs of points in  $X$  and  $P' = \{(x'_i, y'_i), 1 \leq i \leq k\}$  are pairs of points in  $X'$ . Suppose that there exists  $\pi \in M(X, X')$  and  $\mathcal{R} \in \mathcal{C}(X, X')$  such that*

$$D(\pi; \mu, \mu') \vee \pi(\mathcal{R}^c) \vee \frac{1}{2} \text{dis}(\mathcal{R}) \leq \varepsilon$$

*and that for any  $i \leq k$ ,  $(x_i, x'_i) \in \mathcal{R}$  and  $(y_i, y'_i) \in \mathcal{R}$ . Then, for any  $\delta, \delta' > 0$ ,*

$$d_{\text{GHP}}(\text{Coal}_\delta(X, P), \text{Coal}_{\delta'}(X', P')) \leq (2\varepsilon + |\delta - \delta'|)(k + 1).$$



PROOF. This is essentially<sup>4</sup> Lemma 21 in [3] and Lemma 4.2 in [4], thus we leave the details to the reader.  $\square$

LEMMA 5.2. *Let  $\varepsilon \in ]0, 1[$  and  $\delta > 0$ ,  $\mathbf{X} = (X, d, \mu)$  and  $\mathbf{X}' = (X', d', \mu')$  be two m.s.-m.s with a finite number of finite components. If there exists  $\pi \in M(X, X')$  and  $\mathcal{R} \in C(X, X')$  such that*

$$D(\pi; \mu, \mu') \vee \pi(\mathcal{R}^c) \vee \frac{1}{2} \text{dis}(\mathcal{R}) \leq \varepsilon$$

*then one may couple two Poisson random sets:  $\mathcal{P}$  of intensity  $\frac{1}{2}\mu^{\otimes 2} \otimes \text{leb}_{[0,T]}$  and  $\mathcal{P}'$  of intensity  $\frac{1}{2}(\mu')^{\otimes 2} \otimes \text{leb}_{[0,T]}$ , such that with probability larger than  $1 - T\varepsilon(10 + 8\mu(X) + 8\mu'(X')) - \sqrt{2\varepsilon + |\delta - \delta'|}$ , for any  $t \leq T$ ,*

$$d_{\text{GHP}}(\text{Coal}_\delta(\mathbf{X}, \mathcal{P}_t), \text{Coal}_{\delta'}(\mathbf{X}', \mathcal{P}'_t)) \leq (T\mu(X)^2 + 1)\sqrt{2\varepsilon + |\delta - \delta'|}.$$

PROOF. Let  $\mathcal{P}(\mu)$  denote the distribution of a Poisson random set of intensity measure  $\mu$ . Using the coupling characterization of total variation distance and the gluing lemma (cf. [29], page 23), one may construct three Poisson random sets on the same probability space,  $P$ ,  $\tilde{P}$  and  $P'$  such that:

- (i)  $P = (X_i, Y_i, t_i)_{i=1, \dots, N}$  has distribution  $\mathcal{P}(\frac{1}{2}\mu^{\otimes 2} \times \text{leb}_{[0,T]})$ ,
- (ii)  $P' = (X'_i, Y'_i, t'_i)_{i=1, \dots, N'}$  has distribution  $\mathcal{P}(\frac{1}{2}(\mu')^{\otimes 2} \times \text{leb}_{[0,T]})$ ,
- (iii)  $\tilde{P} = (\tilde{X}_i, \tilde{X}'_i, \tilde{Y}_i, \tilde{Y}'_i, \tilde{t}_i)_{i=1, \dots, \tilde{N}}$  has distribution  $\mathcal{P}(\frac{1}{2}\pi^{\otimes 2} \times \text{leb}_{[0,T]})$

and furthermore

$$\begin{aligned} & \mathbb{P}[(X_i, Y_i, t_i)_{i=1, \dots, N} \neq (\tilde{X}_i, \tilde{Y}_i, \tilde{t}_i)_{i=1, \dots, \tilde{N}}] \\ & \leq \left\| \mathcal{P}\left(\frac{1}{2}\mu^{\otimes 2} \times \text{leb}_{[0,T]}\right) - \mathcal{P}\left(\frac{1}{2}\pi_1^{\otimes 2} \times \text{leb}_{[0,T]}\right) \right\| \end{aligned}$$

and

$$\begin{aligned} & \mathbb{P}[(X'_i, Y'_i, t'_i)_{i=1, \dots, N'} \neq (\tilde{X}'_i, \tilde{Y}'_i, \tilde{t}_i)_{i=1, \dots, \tilde{N}}] \\ & \leq \left\| \mathcal{P}\left(\frac{1}{2}(\mu')^{\otimes 2} \times \text{leb}_{[0,T]}\right) - \mathcal{P}\left(\frac{1}{2}\pi_2^{\otimes 2} \times \text{leb}_{[0,T]}\right) \right\|. \end{aligned}$$

Now, for any  $T > 0$ , using Lemma E.1 in the Appendix,

$$\begin{aligned} & \left\| \mathcal{P}\left(\frac{1}{2}\mu^{\otimes 2} \times \text{leb}_{[0,T]}\right) - \mathcal{P}\left(\frac{1}{2}\pi_1^{\otimes 2} \times \text{leb}_{[0,T]}\right) \right\| \\ & \leq 2 \left\| \frac{1}{2}\mu^{\otimes 2} \times \text{leb}_{[0,T]} - \frac{1}{2}\pi_1^{\otimes 2} \times \text{leb}_{[0,T]} \right\| \\ & = T \|\mu^{\otimes 2} - \pi_1^{\otimes 2}\| \\ & \leq 2T(\mu(X) + \pi_1(X))\|\mu - \pi_1\| \\ & \leq 2T\varepsilon(2\mu(X) + \varepsilon) \end{aligned}$$

by hypothesis. Similarly,

$$\left\| \mathcal{P}\left(\frac{1}{2}(\mu')^{\otimes 2} \times \text{leb}_{[0,T]}\right) - \mathcal{P}\left(\frac{1}{2}\pi_2^{\otimes 2} \times \text{leb}_{[0,T]}\right) \right\| \leq 2T\varepsilon(2\mu'(X') + \varepsilon).$$

<sup>4</sup>[4], Lemma 4.2, is stated for trees and for  $\delta = 0$ .

Furthermore,  $(\tilde{X}_i, \tilde{X}'_i, \tilde{t}_i)_{i=1, \dots, \tilde{N}}$  and  $(\tilde{Y}_i, \tilde{Y}'_i, \tilde{t}_i)_{i=1, \dots, \tilde{N}}$  both have distribution  $\mathcal{P}(\frac{1}{2}\pi \otimes \text{leb}_{[0, T]})$ . Thus,

$$\mathbb{P}(\exists i \leq \tilde{N}, (\tilde{X}_i, \tilde{X}'_i) \notin \mathcal{R}) \leq \frac{1}{2}T\pi(\mathcal{R}^c) \leq T\varepsilon$$

and

$$\mathbb{P}(\exists i \leq \tilde{N}, (\tilde{Y}_i, \tilde{Y}'_i) \notin \mathcal{R}) \leq T\varepsilon.$$

Let  $\mathcal{E}$  be the event that  $N = N'$  and for any  $i$ ,  $(\tilde{X}_i, \tilde{X}'_i) \in \mathcal{R}$  and  $(\tilde{Y}_i, \tilde{Y}'_i) \in \mathcal{R}$ . Altogether, we get that  $\mathcal{E}$  has probability at least  $1 - T\varepsilon(10 + 8\mu(X) + 8\mu'(X'))$ .

Since the distortion of  $\mathcal{R}$  is at most  $2\varepsilon$ , we get using Lemma 5.1 that on the event  $\mathcal{E}$ , for any  $t \leq T$ ,

$$d_{\text{GHP}}(\text{Coal}_\delta(X, \mathcal{P}_t), \text{Coal}_{\delta'}(X', \mathcal{P}'_t)) \leq (N + 1)(2\varepsilon + |\delta - \delta'|).$$

Since  $N$  has distribution  $\mathcal{P}(\frac{1}{2}\mu(X)^2T)$ , Markov's inequality implies

$$\mathbb{P}\left(N \geq \frac{T\mu(X)^2}{\sqrt{2\varepsilon + |\delta - \delta'|}}\right) \leq \sqrt{2\varepsilon + |\delta - \delta'|}$$

this gives the result.  $\square$

In the proposition below, recall from Section 2.1 that convergence of processes uses the Skorokhod topology (here for the metric space  $(\mathcal{N}_1, L_{1, \text{GHP}})$ ), which we shall always prove using Lemma A.2.

**PROPOSITION 5.3.** *Let  $\mathbf{X}^n = (X^n, d^n, \mu^n)$ ,  $n \geq 0$  be a sequence of elements in  $\mathcal{N}_1$  and  $(\delta^n)_{n \geq 0}$  a sequence of nonnegative real numbers. Suppose that:*

- (a)  $(\mathbf{X}^n)_{n \geq 0}$  converges (for  $L_{1, \text{GHP}}$ ) to  $\mathbf{X}^\infty = (X^\infty, d^\infty, \mu^\infty)$  as  $n$  goes to infinity,
- (b)  $\delta^n \xrightarrow[n \rightarrow \infty]{} \delta^\infty$ .

Then

- (i)  $(\text{Coal}_{\delta^n}(\mathbf{X}^n, t))_{t \geq 0}$  converges in distribution (for  $L_{1, \text{GHP}}$ ) to  $(\text{Coal}_{\delta^\infty}(\mathbf{X}^\infty, t))_{t \geq 0}$ ,
- (ii) if  $t^n \xrightarrow[n \rightarrow \infty]{} t$ ,  $\text{Coal}_{\delta^n}(\mathbf{X}^n, t^n)$  converges in distribution (for  $L_{1, \text{GHP}}$ ) to  $\text{Coal}_{\delta^\infty}(\mathbf{X}^\infty, t)$ .

**PROOF.** Let us fix  $\varepsilon \in ]0, 1[$ . Let  $\varepsilon_0 \in ]0, \varepsilon/2[$  be such that  $\varepsilon_0 \notin \text{masses}(\mathbf{X}^\infty)$  and

$$\mu(X^\infty_{\leq \varepsilon_0}) \leq \varepsilon.$$

Proposition 4.6 shows that  $\delta_{\text{LP}}(\mathbf{X}^n_{> \varepsilon_0}, \mathbf{X}^\infty_{> \varepsilon_0})$  goes to zero as  $n$  goes to infinity. Let  $\tilde{\varepsilon}$  be a positive real number to be chosen later, depending only on  $\varepsilon, \varepsilon_0, \mu^\infty(X^\infty)$  and  $T$ . Let  $n$  be large enough so that

$$\begin{aligned} \delta_{\text{LP}}(\mathbf{X}^n_{> \varepsilon_0}, \mathbf{X}^\infty_{> \varepsilon_0}) &\leq \tilde{\varepsilon}, \\ |\delta^n - \delta^\infty| &\leq \tilde{\varepsilon} \end{aligned}$$

and

$$\|\text{masses}(\mathbf{X}^n) - \text{masses}(\mathbf{X}^\infty)\|_1 \leq \varepsilon.$$

Let  $k := \#\text{comp}(X^\infty_{> \varepsilon_0})$  and notice that

$$k \leq \frac{\mu^\infty(X^\infty)}{\varepsilon_0}.$$

Lemma 4.8 shows that

$$d_{\text{GHP}}(\mathbf{X}_{>\varepsilon_0}^n, \mathbf{X}_{>\varepsilon_0}^\infty) \leq \tilde{\varepsilon} \frac{\mu^\infty(X^\infty)}{\varepsilon_0}.$$

Notice also that

$$\mu^n(X_{\leq\varepsilon_0}^n) \leq \mu^\infty(X_{\leq\varepsilon_0}^\infty) + \|\text{masses}(X^n) - \text{masses}(X^\infty)\|_1 \leq 2\varepsilon,$$

and that

$$\mu^n(X^n) \leq \mu^\infty(X^\infty) + \|\text{masses}(X^n) - \text{masses}(X^\infty)\|_1 \leq \mu^\infty(X^\infty) + \varepsilon.$$

Thus, using Lemma 5.2, one may couple the coalescence on  $X_{>\varepsilon_0}^n$  and  $X_{>\varepsilon_0}^\infty$  in such a way that there is an event  $A$  satisfying the following. On  $A$ , for any  $t \leq T$ ,

$$d_{\text{GHP}}(\text{Coal}_{\delta^n}(\mathbf{X}_{>\varepsilon_0}^n, t), \text{Coal}_{\delta^\infty}(\mathbf{X}_{>\varepsilon_0}^\infty, t)) \leq \alpha(\tilde{\varepsilon}, \varepsilon_0, T, \mu^\infty(X^\infty))$$

with

$$\alpha(\tilde{\varepsilon}, \varepsilon_0, T, \mu^\infty(X^\infty)) := (T\mu^\infty(X^\infty)^2 + 1) \sqrt{2\tilde{\varepsilon} \frac{\mu^\infty(X^\infty)}{\varepsilon_0}} + \tilde{\varepsilon}$$

and furthermore,

$$\mathbb{P}(A^c) \leq \beta(\tilde{\varepsilon}, \varepsilon_0, T, \mu^\infty(X^\infty))$$

with

$$\beta(\tilde{\varepsilon}, \varepsilon_0, T, \mu^\infty(X^\infty)) := T\tilde{\varepsilon} \frac{\mu^\infty(X^\infty)}{\varepsilon_0} (10 + 16\mu^\infty(X^\infty)) + 2\varepsilon + 2\sqrt{2\tilde{\varepsilon} \frac{\mu^\infty(X^\infty)}{\varepsilon_0}} + \tilde{\varepsilon}.$$

Using Lemma 4.8, we obtain that on  $A$ , for any  $t \leq T$ ,

$$L_{\text{GHP}}(\text{Coal}_{\delta^n}(\mathbf{X}_{>\varepsilon_0}^n, t), \text{Coal}_{\delta^\infty}(\mathbf{X}_{>\varepsilon_0}^\infty, t)) \leq 16 \frac{\mu^\infty(X^\infty)}{\varepsilon_0} \alpha(\tilde{\varepsilon}, \varepsilon_0, T, \mu^\infty(X^\infty)).$$

Now, independently from this coupling, let us couple two independent Poisson random sets on  $(X^n)^2 \setminus (X_{>\varepsilon_0}^n)^2 \times [0, T]$  and  $(X^\infty)^2 \setminus (X_{>\varepsilon_0}^\infty)^2 \times [0, T]$  (with intensities given by the restrictions of  $\mu^n$  and  $\mu^\infty$ ). Let  $B$  denote the event that these Poisson random sets are both empty. Then

$$\begin{aligned} \mathbb{P}(B^c) &\leq 1 - e^{-\frac{T}{2}(\mu^n(X^n)^2 - \mu^n(X_{>\varepsilon_0}^n)^2)} + 1 - e^{-\frac{T}{2}(\mu^\infty(X^\infty)^2 - \mu^\infty(X_{>\varepsilon_0}^\infty)^2)} \\ &\leq T\mu^n(X^n)\mu^n(X_{\leq\varepsilon_0}^n) + T\mu^\infty(X^\infty)\mu^\infty(X_{\leq\varepsilon_0}^\infty) \\ &\leq 6T\varepsilon(\mu^\infty(X^\infty) + \varepsilon) =: \gamma(\varepsilon, \mu^\infty(X^\infty), T). \end{aligned}$$

On  $A \cap B$ , we obtain that for any  $t \leq T$ ,

$$L_{\text{GHP}}(\text{Coal}_{\delta^n}(\mathbf{X}^n, t), \text{Coal}_{\delta^\infty}(\mathbf{X}^\infty, t)) \leq 2 \left( 1 + 8 \frac{\mu^\infty(X^\infty)}{\varepsilon_0} \right) \alpha(\tilde{\varepsilon}, \varepsilon_0, T, \mu^\infty(X^\infty)).$$

Furthermore,

$$\mathbb{P}((A \cap B)^c) \leq \beta(\tilde{\varepsilon}, \varepsilon_0, T, \mu^\infty(X^\infty)) + \gamma(\varepsilon, \mu^\infty(X^\infty), T).$$

Thus, one may choose  $\tilde{\varepsilon}$  as a function of  $T, \varepsilon, \varepsilon_0$  and  $\mu^\infty(X^\infty)$  such that with probability at least  $1 - C\varepsilon$ ,

$$L_{\text{GHP}}(\text{Coal}_{\delta^n}(\mathbf{X}^n, t), \text{Coal}_{\delta^\infty}(\mathbf{X}^\infty, t)) \leq \varepsilon$$

for some finite constant  $C$  depending only on  $\mu(X^\infty)$ .

Furthermore, since the multigraphs  $\mathbb{M}\mathbb{G}(X^\infty, t)$  and  $\mathbb{M}\mathbb{G}(X^n, t)$  are the same for any  $t \leq T$  in this coupling,

$$\begin{aligned} & \|\text{masses}(\text{Coal}_{\delta^n}(X^n, t)) - \text{masses}(\text{Coal}_\delta(X^\infty, t))\|_1 \\ & \leq \|\text{masses}(\text{Coal}_{\delta^n}(X^n_{>\varepsilon_0}, t)) - \text{masses}(\text{Coal}_\delta(X^\infty_{>\varepsilon_0}, t))\|_1 \\ & \quad + \|\text{masses}(\text{Coal}_{\delta^n}(X^n_{\leq\varepsilon_0}, t)) - \text{masses}(\text{Coal}_\delta(X^\infty_{\leq\varepsilon_0}, t))\|_1 \\ & \leq \|\text{masses}(X^n_{>\varepsilon_0}, t) - \text{masses}(X^\infty_{>\varepsilon_0}, t)\|_1 + \mu(X^n_{\leq\varepsilon_0}) + \mu(X^\infty_{\leq\varepsilon_0}) \\ & \leq 4\varepsilon. \end{aligned}$$

This shows (i). To obtain (ii), notice that for any  $s$  and  $\eta > 0$ ,

$$\mathbb{P}(\exists t \in [s, s + \eta] : \text{Coal}_{\delta^\infty}(X^\infty, t) \neq \text{Coal}_{\delta^\infty}(X^\infty, s)) \leq \mu(X^\infty)^2 \eta.$$

Thus, (ii) is a simple consequence of (i).  $\square$

We shall need the following variation of Proposition 5.3 when studying simultaneous coalescence and fragmentation in Section 7.

**PROPOSITION 5.4.** *Let  $X^n = (X^n, d^n, \mu^n)$ ,  $n \in \overline{\mathbb{N}}$  be a sequence of random variables in  $\mathcal{N}_1^{\text{graph}}$  and  $(\delta^n)_{n \geq 0}$  a sequence of non-negative real numbers. Suppose that:*

- (a)  $(X^n)_{n \geq 0}$  converges in distribution for  $L_{1, \text{GHP}}^{\text{surplus}}$  to  $X^\infty$  as  $n$  goes to infinity,
- (b)  $\delta^n \xrightarrow[n \rightarrow \infty]{} \delta$ .

Then:

- (i)  $(\text{Coal}_{\delta^n}(X^n, t))_{t \geq 0}$  converges in distribution (for the Skorokhod topology associated to  $L_{1, \text{GHP}}^{\text{surplus}}$ ) to  $(\text{Coal}_\delta(X^\infty, t))_{t \geq 0}$ ,
- (ii) if  $t^n \xrightarrow[n \rightarrow \infty]{} t$ ,  $\text{Coal}_{\delta^n}(X^n, t^n)$  converges in distribution to  $\text{Coal}_\delta(X^\infty, t)$  for  $L_{1, \text{GHP}}^{\text{surplus}}$ .

**PROOF.** Notice first that when  $X$  belongs to  $\mathcal{N}_1^{\text{graph}}$ , then with probability one,  $\text{Coal}_\delta(X^n, t)$  is in  $\mathcal{N}_1^{\text{graph}}$  for any  $t \geq 0$ . Indeed, since  $X$  has finite mass, there is with probability one a finite number of points in the Poisson process  $\mathcal{P}_t^+$  on  $X^2$  for any  $t \geq 0$ .

Now, the proof is essentially the same as the one of Proposition 5.3, except that since  $(X^n)_{n \geq 0}$  converges to  $X^\infty$  for  $L_{1, \text{GHP}}^{\text{surplus}}$ , one may use  $d_{\text{GHP}}^{\text{surplus}}$  instead of  $d_{\text{GHP}}$ . The fact that the multigraphs  $\mathbb{M}\mathbb{G}(X^n, s)$  and  $\mathbb{M}\mathbb{G}(X^\infty, s)$  are the same for any  $s \leq T$ , and that in the coupling no point of the Poisson processes touches  $X^n_{\leq\varepsilon_0}$  or  $X^\infty_{\leq\varepsilon_0}$  implies that for each component of  $\text{Coal}_{\delta^n}(X^n_{>\varepsilon_0}, s)$ , its surplus is the same as the surplus of the corresponding component in  $\text{Coal}_\delta(X^\infty_{>\varepsilon_0}, s)$ .  $\square$

**5.2. Structural result for Aldous' multiplicative coalescent.** Recall the definition of the multigraph  $\mathbb{M}\mathbb{G}(x, t)$  for  $x \in \ell^2$  in Section 4.1. We shall use notation analogous to those in Definition 2.8. For instance, for  $x \in \ell^2_+$ ,  $x_{\leq\varepsilon}$  denotes the element in  $\ell^2_+$  defined by

$$\forall i \in \mathbb{N}, \quad x_{\leq\varepsilon}(i) = x(i)\mathbf{1}_{x(i) \leq \varepsilon}.$$

Also, for  $i \in \mathbb{N}$ ,  $x \setminus \{i\}$  denotes the element in  $\ell^2$  defined by

$$\forall j \in \mathbb{N}, \quad (x \setminus \{i\})(j) = x(j)\mathbf{1}_{j \neq i}.$$

Notice that at time 0, the components of  $\mathbb{M}\mathbb{G}(x, 0)$  are the singletons  $\{i\}$  for  $i \in \mathbb{N}$ . Let us fix some  $\varepsilon > 0$  and say that a component of  $\mathbb{M}\mathbb{G}(x, t)$  is significant if it has weight larger than

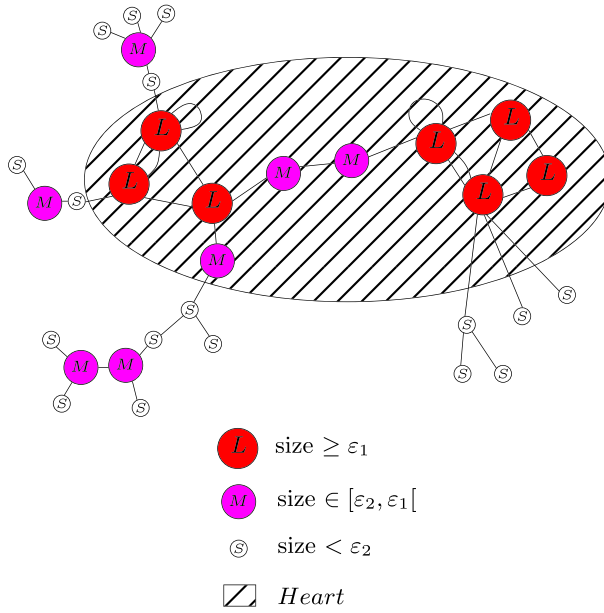


FIG. 1. The structure of a significant component.

$\varepsilon$ , that is, the sum of the weights  $x_i$  of its vertices  $i$  is larger than  $\varepsilon$ . In Lemma 5.5 below and its corollary, we shall derive three scales (at time 0), namely, Large, Medium and Small such that with high probability (as  $\varepsilon$  goes to zero), every significant component of  $\mathbb{MG}(x, t)$  is made of a *heart*<sup>5</sup> made of Large or Medium components of  $\mathbb{MG}(x, 0)$  to which are attached *hanging trees* of small or medium components of  $\mathbb{MG}(x, 0)$  such that the component of the trees attached to the heart are small components (see Figure 1) and the mass contained in the hanging trees is at most medium. Furthermore, these scales depend on  $x, \varepsilon$  and  $t$  through the functions  $\alpha \mapsto \|x_{\leq \alpha}\|_2$  and  $K \mapsto \mathbb{P}(S(x, t) \geq K)$ . This picture will be fundamental in the proof of the Feller properties of the coalescent and dynamical percolation processes. Indeed, it will imply that small components of  $\mathbf{X} \in \mathcal{N}_2$  have small influence on the geometry of the large components of  $\text{Coal}(\mathbf{X}, t)$ . This will allow us to approximate  $\text{Coal}(\mathbf{X}, t)$  by a truncated version,  $\text{Coal}(\mathbf{X}_{\geq \varepsilon}, t)$  for  $\varepsilon$  small enough, and to reduce the Feller property on  $\mathcal{N}_2$  to the Feller property on  $\mathcal{N}_1$ , where a finite number of identifications occur on every finite interval of time.

LEMMA 5.5. Let  $x \in \ell^2(\mathbb{N})$ ,  $T \geq 0$  and  $0 < \varepsilon < 1$ . Suppose that:

(i)  $K \geq 1$  is such that:

$$\mathbb{P}(S(x, T) \geq K) \leq \frac{\varepsilon}{100},$$

(ii)  $\varepsilon_1 \in (0, \varepsilon)$  is such that

$$S(x_{\leq \varepsilon_1}, 0) \leq \frac{\varepsilon^2}{100(1 + T + KT^2)},$$

<sup>5</sup>This term is not standard, but as will be apparent in the subsequent proofs, the heart of a significant component of  $\mathbb{MG}(x, t)$  is uniquely defined when  $x, t$  and the weight of a small component are defined. This notion has nothing to do with the core of a graph.

(iii)  $\varepsilon_2 \in (0, \varepsilon_1)$  is such that

$$S(x_{\leq \varepsilon_2}, 0) \leq \frac{2\varepsilon_1^2 \varepsilon^2}{100(1 + T(K + 2))^2}.$$

Then with probability larger than  $1 - \varepsilon$ , the following holds for any  $t \leq T$ :

- (a) every component of  $\text{MG}(x, t)$  of size larger than  $\varepsilon$  contains a component of  $\text{MG}(x, 0)$  of size larger than  $\varepsilon_1$ .
- (b) no component of  $\text{MG}(x_{\leq \varepsilon_1}, 0)$  is contained in a cycle in  $\text{MG}(x, t)$ .
- (c) for each component  $m$  of  $\text{MG}(x_{\leq \varepsilon_1}, t)$  and each component  $m'$  of  $\text{MG}(x_{> \varepsilon_1}, t)$ , there is at most one edge between  $m$  and  $m'$  in  $\text{MG}(x, t)$ .
- (d)  $S(x, t) - S(x_{> \varepsilon_2}, t) \leq 2\varepsilon_1^2$ .
- (e) for any component  $\{i\}$  of  $\text{MG}(x, 0)$  of size larger than  $\varepsilon_1$ , the difference between the sizes of the component containing  $i$  in  $\text{MG}(x, t)$  and the one containing  $i$  in  $\text{MG}(x_{> \varepsilon_2}, t)$  is less than  $\varepsilon_1$ .

We may now state a precise version of the structural decomposition sketched just before Lemma 5.5. We state it uniformly on a convergent sequence in  $\ell^2_{\searrow}$  because it will be convenient to prove the almost Feller property on  $\mathcal{S}$ , Theorem 3.1.

When  $\varepsilon_2 \leq \varepsilon_1 \leq \varepsilon$  are three positive real numbers, we shall say that a component is *significant* if it has size larger than  $\varepsilon$ , *Large* if it has size larger than  $\varepsilon_1$ , *Medium* for a size in  $(\varepsilon_2, \varepsilon_1]$  and *Small* for a size not larger than  $\varepsilon_2$ .

**COROLLARY 5.6.** *Let  $x^n$  be a sequence in  $\ell^2_{\searrow}$  converging to  $x^\infty$  in  $\ell^2$ . Then, for any  $\varepsilon > 0$  and any  $T > 0$ , there exist  $\varepsilon_2 \leq \varepsilon_1 \leq \varepsilon$  such that for any  $n \in \overline{\mathbb{N}}$ , with probability at least  $1 - \varepsilon$ . The following holds for any  $t \in [0, T]$ :*

- (a) every significant component of  $\text{MG}(x^n, t)$  is made of a connected heart made of Large or Medium components of  $\text{MG}(x^n, 0)$  to which are attached hanging trees (each one attached by a single edge to the heart) of Small or Medium components of  $\text{MG}(x^n, 0)$  such that the components of the trees attached to the heart are Small components and the mass contained in the hanging trees is less than  $\varepsilon_1$ ,
- (b) no Medium or Small component of  $\text{MG}(x^n, 0)$  belongs to a cycle in  $\text{MG}(x^n, t)$ ,
- (c)  $S(x^n, t) - S(x^n_{> \varepsilon_2}, t) \leq 2\varepsilon_1^2$ .

The proof of Lemma 5.5 relies essentially on Aldous’ analysis of the multiplicative coalescent.

**PROOF OF LEMMA 5.5.** If for some  $t \leq T$ , there exists a significant component of  $\text{MG}(x, t)$  which does not contain any large component of  $x$ , then  $S(x_{\leq \varepsilon_1}, t) > \varepsilon^2$  and thus  $S(x_{\leq \varepsilon_1}, T) > \varepsilon^2$ , since  $S(x, \cdot)$  is nondecreasing. Thus Lemma 4.1 shows that the probability of (a) is larger than  $1 - \varepsilon/4$ , as soon as hypothesis (ii) of Lemma 5.5 holds.

Let  $\{i\}$  be a component of  $\text{MG}(x, 0)$  and define the event

$$A_i = \{\text{there exist at least two edges of } \text{MG}(x, T) \text{ connecting } i \text{ to the same component of } \text{MG}(x \setminus \{i\}, T)\}.$$

Due to the properties of the Poisson process defining  $\text{MG}(x, T)$ , one sees that conditionally on  $\text{MG}(x \setminus \{i\}, T)$ , the number of edges of  $\text{MG}(x, T)$  connecting a fixed component  $m$  of  $\text{MG}(x \setminus \{i\}, T)$  to  $i$  is a Poisson random variable with parameter  $\sum_{j \in m} T x_i x_j$ . Thus,

$$\begin{aligned} \mathbb{P}(A_i | \text{MG}(x \setminus \{i\}, T)) &\leq \sum_{m \text{ c.c. of } \text{MG}(x \setminus \{i\}, T)} \left( T x_i \sum_{j \in m} x_j \right)^2 \\ &= T^2 x_i^2 S(x \setminus \{i\}, T), \end{aligned}$$

thus,

$$(5.1) \quad \begin{aligned} \mathbb{P}(A_i \cap \{S(x, T) \leq K\}) &\leq \mathbb{E}[\mathbb{P}(A_i | \text{MG}(x \setminus \{i\}, T)) \mathbf{1}_{S(x \setminus \{i\}, T) \leq K}] \\ &\leq KT^2 x_i^2. \end{aligned}$$

We thus obtain

$$\mathbb{P}\left(\bigcup_{i \in \mathbb{N} \text{ s.t. } x_i \leq \varepsilon_1} A_i\right) \leq KT^2 S(x_{\leq \varepsilon_1}, 0) + \mathbb{P}(S(x, T) > K),$$

which shows that the probability of (b) is at least  $1 - \varepsilon/4$  as soon as hypotheses (i) and (ii) of Lemma 5.5 hold.

The proof of (c) is similar. Let

$$B_T := \{\text{there exist at least two edges of } \text{MG}(x, T) \text{ connecting a component } m \text{ of } \text{MG}(x_{\leq \varepsilon_1}, T) \text{ to a component } m' \text{ of } \text{MG}(x_{> \varepsilon_1}, T)\}.$$

Then

$$\begin{aligned} &\mathbb{P}(B_T | \text{MG}(x_{> \varepsilon_1}, T), \text{MG}(x_{\leq \varepsilon_1}, T)) \\ &\leq \sum_{m \text{ c.c. of } \text{MG}(x_{\leq \varepsilon_1}, T)} \sum_{m' \text{ c.c. of } \text{MG}(x_{> \varepsilon_1}, T)} \left(\sum_{i \in m} x_i \sum_{j \in m'} x_j T\right)^2 \\ &= T^2 S(x_{\leq \varepsilon_1}, T) S(x_{> \varepsilon_1}, T). \end{aligned}$$

Thus,

$$\mathbb{P}(B_T) \leq \frac{\varepsilon}{5} + \mathbb{P}(S(x, T) \geq K) + \mathbb{P}\left(S(x_{\leq \varepsilon_1}, T) \geq \frac{\varepsilon}{5KT^2}\right),$$

which shows using Lemma 4.1 that the probability of (c) is at least  $1 - \varepsilon/4$  as soon as hypotheses (i) and (ii) of Lemma 5.5 hold (notice that  $B_t \subset B_T$  if  $t \leq T$ ).

Now, let  $Y$  be the supremum, over Large components  $\{i\}$  of  $\text{MG}(x, 0)$ , of the difference between the sizes of the component containing  $i$  in  $\text{MG}(x, t)$  and the one containing  $i$  in  $\text{MG}(x_{> \varepsilon_2}, t)$ . Notice that if  $Y \geq \alpha$ , then  $S(x, t) \geq S(x_{> \varepsilon_2}, t) + 2\varepsilon_1 \alpha$ , which implies  $S(x, T) \geq S(x_{> \varepsilon_2}, T) + 2\varepsilon_1 \alpha$  when (b) holds, thanks to Lemma 4.5. Thus points (e) and (d) will be proved if we show that  $S(x, T) > S(x_{> \varepsilon_2}, T) + 2\varepsilon_1^2$  with probability at most  $\varepsilon/4$ .

Define the event

$$C = \{S(x, T) > S(x_{> \varepsilon_2}, T) + 2\varepsilon_1^2\}.$$

Lemma 4.2 shows that

$$\begin{aligned} &\mathbb{P}(C \text{ and } S(x, T) \leq K \text{ and } S(x_{\leq \varepsilon_2}, T) \leq \beta) \\ &\leq (1 + T(K + 2\varepsilon_1^2))^2 \frac{\beta}{2\varepsilon_1^2}. \end{aligned}$$

Thus,

$$\mathbb{P}(C) \leq (1 + T(K + 2))^2 \frac{\beta}{2\varepsilon_1^2} + \mathbb{P}(S(x_{\leq \varepsilon_2}, T) > \beta) + \mathbb{P}(S(x, T) > K),$$

which is less than  $\varepsilon/4$  if we take

$$\beta = \frac{2\varepsilon_1^2 \varepsilon}{100(1 + T(K + 2))^2}$$

and if hypotheses (i), (ii), and (iii) of Lemma 5.5 hold.  $\square$



PROOF OF COROLLARY 5.6. Since  $x^n$  converges to  $x^\infty$  in  $\ell^2$ ,

$$\sup_{n \in \mathbb{N}} \|x^n_{\leq \varepsilon}\|_2 \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Also, the Feller property of the multiplicative coalescent in  $\ell^2$ , cf. [5] implies that the distributions of the sizes of  $\text{MG}(x^n, T)$  for  $n \in \mathbb{N}$  form a compact family of probability measures on  $\ell^2$ . Thus,

$$(5.2) \quad \sup_{n \in \mathbb{N}} \mathbb{P}(S(x^n, T) \geq K) \xrightarrow{K \rightarrow +\infty} 0.$$

This shows that for any  $T$  and  $\varepsilon$ , one may find  $K, \varepsilon_1$  and  $\varepsilon_2$  such that the three hypotheses of Lemma 5.5 hold for  $x^n$  uniformly over  $n \in \mathbb{N}$ .

Now suppose that  $x, t, \varepsilon, \varepsilon_1$  and  $\varepsilon_2$  are such that (a), (b), (e) and (c) of Lemma 5.5 hold. Let  $m$  be a significant component of  $\text{MG}(x, t)$ . It contains a large component  $\{i\}$  of  $\text{MG}(x, 0)$  by point (a). Let  $\sigma(m)$  denote the component of  $\text{MG}(x_{>\varepsilon_2}, t)$  containing  $\{i\}$ . Point (e) implies that two large components of  $\text{MG}(x, 0)$  are connected in  $\text{MG}(x, t)$  if and only if they are connected in  $\text{MG}(x_{>\varepsilon_2}, t)$ , that is, only through Large or Medium components. This shows that  $\sigma(m)$  does not depend on the choice of the large component  $\{i\}$  included in  $m$  and that there cannot be any large component in  $m \setminus \sigma(m)$ . Let us define the heart of  $m$  as  $\sigma(m)$ . It is made of Large or Medium components, and  $m \setminus \sigma(m)$  is a graph of medium or small components. Now if a medium component in  $m \setminus \sigma(m)$  was directly connected to some component of  $\sigma(m)$ , it would be connected to  $m$  in  $\text{MG}(x_{>\varepsilon_2}, t)$ , and thus would belong to  $\sigma(m)$ . Thus, the exterior boundary of  $\sigma(m)$  in  $m$  is made of small components. Point (b) shows that no Medium or Small component of  $\text{MG}(x, 0)$  belongs to a cycle in  $\text{MG}(x, t)$ , which implies notably that  $m \setminus \sigma(m)$  is a forest, and point (c) shows that each tree of this forest is attached by a single edge to  $\sigma(m)$ .  $\square$

A useful by-product of the proof of Lemma 5.5 and Corollary 5.6 is the following simple lemma.

LEMMA 5.7. For  $x \in \ell^2, \varepsilon > 0$  and  $T > 0$  let  $A(x, \varepsilon, T)$  be the event that for any  $t \leq T$ :

- $\text{MG}(x_{\leq \varepsilon}, t)$  is a forest and
- there is at most one edge between any connected component of  $\text{MG}(x_{\leq \varepsilon}, t)$  and any component of  $\text{MG}(x_{>\varepsilon}, t)$ .

Suppose that  $x^n$  converges to  $x^\infty$  in  $\ell^2$  as  $n$  goes to infinity. Then, for any  $T > 0$ ,

$$\inf_{n \in \mathbb{N}} \mathbb{P}(A(x^n, \varepsilon, T)) \xrightarrow{\varepsilon \rightarrow 0} 1.$$

REMARK 5.8. Let us give an example of an m.s-m.s which is in  $\mathcal{N}_2$  but not in  $\mathcal{S}$ . Let  $I_i, i \geq 1$  be disjoint copies of the interval  $[0, 1]$ , with its usual metric, and equip  $I_i$  with the measure  $\frac{1}{i}(\delta_0 + \delta_1)$ . Then  $X \in \mathcal{N}_2 \setminus \mathcal{S}$ . In fact, thanks to Lemma 5.5, for any  $\varepsilon > 0$  and  $t > 0$  every component of  $\text{Coal}_0((X)_{>\varepsilon}, t)$  is unbounded since it contains a forest of an infinite number (since the sizes are not in  $\ell^1$ ) of components of diameter 1.

5.3. The coalescent on  $\mathcal{S}$ . The aim of this section is to prove Theorem 3.1. We shall first prove two lemmas.

LEMMA 5.9. Let  $X$  be an m.s-m.s.

- (i) If  $X \in \mathcal{S}$  then, almost surely, for any  $t \geq 0, \text{Coal}_0(X, t)$  belongs to  $\mathcal{N}_2$ .

(ii) If  $X \in \mathcal{N}_2$  and is a length space, then, almost surely, for any  $t \geq 0$ ,  $\text{Coal}_0(X, t)$  is a length space and the commutation relation (2.5) holds.

(iii) If the components of  $X$  are  $\mathbb{R}$ -graphs then, almost surely, for any  $t \geq 0$ , the components of  $\text{Frag}(X, t)$  are  $\mathbb{R}$ -graphs. Consequently, if  $X$  belongs to  $\mathcal{S}^{\text{graph}}$ , then, almost surely, for any  $t \geq 0$ ,  $\text{Frag}(X, t)$  belongs to  $\mathcal{S}^{\text{graph}}$ .

PROOF.

(i) Suppose that  $X$  belongs to  $\mathcal{S}$  and let us show that with probability one,  $\text{Coal}_0(X, t)$  has totally bounded components for any  $t \geq 0$ . Let  $\alpha > 0$  and  $\varepsilon \in ]0, 1[$  be fixed and let

$$B(\varepsilon) := \{\text{supdiam}(\text{Coal}_0(X_{\leq \varepsilon}, T)) \leq \alpha\}.$$

Since  $X$  satisfies (2.3),

$$\mathbb{P}(B(\varepsilon)^c) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Now, we perform coalescence and use the obvious coupling between  $\text{Coal}_0(X, t)$  and  $\text{Coal}_0(X_{>\varepsilon_2}, t)$ . Recall the multigraph  $\mathbb{M}\mathbb{G}(X, t)$  introduced in Section 2.5.2. We let  $S(X, t)$  denote the sum of the squares of the masses of the components in  $\text{Coal}_0(X, t)$  (or  $\mathbb{M}\mathbb{G}(X, t)$ ). Let  $\varepsilon_1$  and  $\varepsilon_2$  be positive numbers to be chosen soon and let  $A(\varepsilon)$  be the event that for any  $t \leq T$ :

(a) every significant component of  $\mathbb{M}\mathbb{G}(X, t)$  is made of a connected *heart* made of Large or Medium components of  $\mathbb{M}\mathbb{G}(X, 0)$  to which are attached *hanging trees* (each one attached by a single edge to the heart) of Small or Medium components of  $\mathbb{M}\mathbb{G}(X, 0)$  such that the components of the trees attached to the heart are Small components and the mass contained in the hanging trees is less than  $\varepsilon_1$ ,

(b) no Medium or Small component of  $\mathbb{M}\mathbb{G}(X, 0)$  belongs to a cycle in  $\mathbb{M}\mathbb{G}(X, t)$ ,

(c)  $S(X, t) - S(X_{>\varepsilon_2}, t) \leq 2\varepsilon_1^2$ .

Then Corollary 5.6 (with  $x^n = x^\infty = \text{masses}(X)$ ) shows that one can choose  $\varepsilon_1$  and  $\varepsilon_2$  (as functions of  $\varepsilon, T$  and  $\text{masses}(X)$ ) such that

$$\mathbb{P}(A(\varepsilon)^c) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Then, on  $A(\varepsilon) \cap B(\varepsilon)$ , we have that for any  $t \in [0, T]$ , any component of size larger than  $\varepsilon$  of  $\text{Coal}_0(X, t)$  can be covered with a finite number of balls of radius  $2\alpha$ . Indeed, if  $m$  is such a component, one may first cover the heart with a finite number of balls of radius  $\alpha$  since the heart of  $m$  is composed of a finite number of totally bounded components of  $X$  glued together, and then if we increase the radius to  $2\alpha$ , those balls will cover the whole component  $m$  because we are on  $B(\varepsilon)$ . Making  $\varepsilon$  go to zero, we see that with probability one, for any  $t \in [0, T]$  every component of  $\text{Coal}_0(X, t)$  can be covered with a finite number of balls of radius  $2\alpha$ . Then, letting  $\alpha$  go to zero, we see that with probability one, for any  $t \in [0, T]$  every component of  $\text{Coal}_0(X, t)$  is totally bounded, so  $\text{Coal}_0(X, t) \in \mathcal{N}_2$ .

(ii) The fact that  $\text{Coal}_0(X, t)$  is a length space is an immediate consequence of Remark 2.22(iv). If  $X = (X, d, \mu) \in \mathcal{N}_2$ , using the same notation as above, one can still guarantee that

$$\mathbb{P}(A(\varepsilon)^c) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

On  $A(\varepsilon)$ , for any  $x$  and  $y$  in a component of  $\text{Coal}_0(X, t)$  of mass larger than  $\varepsilon$  (i.e., a significant component), there is only a finite number of simple paths from  $x$  to  $y$ ,

and every such simple path takes a finite number of shortcuts of the Poisson process  $\mathcal{P}_t^+$ . Letting  $\varepsilon$  go to zero, this holds almost surely for any component of  $\text{Coal}_0(\mathbf{X}, \mathcal{P}_t^+)$ . Furthermore, since  $\ell_X$  is diffuse and  $\mathcal{P}^-$  and  $\mathcal{P}^+$  are independent, almost surely one has, for any  $t$ , and using the notation of Lemma C.4,

$$\mathcal{P}_t^{-,d} \cap \{x \in X : \exists y \in X, (x, y) \text{ or } (y, x) \in \mathcal{P}_t^+\} = \emptyset.$$

Thus, Lemma C.4 shows that (ii) holds.

(iii) Using Lemma 2.19,  $X$  is isometric to  $\text{Coal}_0(X', A)$  where  $X'$  is an m.s-m.s whose components are real trees, and  $A \subset \bigcup_{m \in \text{comp}(X')} m^2$  is finite on any  $m^2$ . Again, since  $\ell_X$  is diffuse, almost surely, for any  $t$ , and using the notation of Lemma C.4,

$$\mathcal{P}_t^{-,d} \cap \{x \in X : \exists y \in X, (x, y) \text{ or } (y, x) \in A\} = \emptyset.$$

Thus, Lemma C.4 shows that

$$\text{Frag}(X, \mathcal{P}_t^-) = \text{Frag}(\text{Coal}_0(X', A), \mathcal{P}_t^-) = \text{Coal}_0(\text{Frag}(X', \mathcal{P}_t^-), A).$$

The components of  $\text{Frag}(X', \mathcal{P}_t^-)$  are  $\mathbb{R}$ -trees, thus the components of  $\text{Coal}_0(\text{Frag}(X', \mathcal{P}_t^-), A)$  are  $\mathbb{R}$ -graphs. The last part of (iii) follows from the fact that  $\mathcal{S}^{\text{graph}}$  is clearly stable by fragmentation.  $\square$

REMARK 5.10.

(i) If  $\mathbf{X} \in \mathcal{N}$  and  $\mathcal{P}$  is as in Definition 2.21, it may happen that  $\text{Frag}(\mathbf{X}, \mathcal{P})$  has a component of mass zero. In this case,  $\text{Frag}(\mathbf{X}, \mathcal{P})$  does not belong to  $\mathcal{N}$ , strictly speaking. However,  $\text{Frag}(\mathbf{X}, \mathcal{P})$  is at zero  $L_{\text{GHP}}$ -distance from an element of  $\mathcal{N}$ , which is  $\text{Frag}(\mathbf{X}, \mathcal{P})|_{\bigcup_{\varepsilon>0} \mathcal{M}_{>\varepsilon}}$ . In fact, we could have defined  $\mathcal{N}$  as the quotient of the set of counting measures on  $\mathcal{M}$  with respect to the equivalence relation defined by being at zero  $L_{\text{GHP}}$ -distance. This space is isometric to  $\mathcal{N}$  modulo the addition of components of null masses. Then  $\text{Frag}(\mathbf{X}, \mathcal{P})$  would have always belonged to  $\mathcal{N}$ . But I feel that it would have obscured the definition of  $\mathcal{N}$ . In the sequel, we shall keep in mind that components of null masses are neglected.

(ii) It is apparent from the proof of point (ii) above that when  $\mathbf{X}$  belongs to  $\mathcal{S}^{\text{graph}}$ , then  $\text{Coal}_0(\mathbf{X}, t)$  has components which are precompact  $\mathbb{R}$ -trees with a countable number of identifications. Thus,  $\text{Coal}_0(\mathbf{X}, t)$  is in  $\mathcal{S}^{\text{graph}}$  if and only if the numbers of identifications on any component is finite. One consequence of this is that if  $\text{Coal}_0(\mathbf{X}, t)$  is in  $\mathcal{S}^{\text{graph}}$  for some  $t \geq 0$ , then  $\text{Coal}_0(\mathbf{X}, s)$  and  $\text{CoalFrag}(\mathbf{X}, s)$  are in  $\mathcal{S}^{\text{graph}}$  for every  $s \in [0, t]$ .

LEMMA 5.11. *Let  $\mathbf{X}^n = (X^n, d^n, \mu^n)$ ,  $n \geq 0$  be a sequence of random variables in  $\mathcal{S}$  and  $(\delta^n)_{n \geq 0}$  be a sequence of nonnegative real numbers. Suppose that:*

(i)  $(\mathbf{X}^n)$  converges in distribution for  $L_{2,\text{GHP}}$  to  $\mathbf{X}^\infty = (X^\infty, d^\infty, \mu^\infty)$  as  $n$  goes to infinity,

(ii)  $\delta^n \xrightarrow{n \rightarrow \infty} 0$ ,

(iii) For any  $\alpha > 0$  and any  $T > 0$ ,

$$\limsup_{n \in \mathbb{N}} \mathbb{P}(\text{supdiam}(\text{Coal}_{\delta^n}(\mathbf{X}_{\leq \varepsilon}^n, T)) > \alpha) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Then, with probability 1,  $\text{Coal}_0(\mathbf{X}^\infty, t)$  belongs to  $\mathcal{S}$  for any  $t \geq 0$ .

PROOF. First, the Feller property of the multiplicative coalescent, [5], Proposition 5, shows that  $\text{masses}(\text{Coal}_{\delta^n}(\mathbf{X}^n, T))$  converges in distribution (in  $\ell_{\downarrow}^2$ ) to  $\text{masses}(\text{Coal}_0(\mathbf{X}^\infty, T))$ . Together with Skorokhod's representation theorem and Lemma 5.7, this implies that

$$(5.3) \quad \mathbb{P}[\text{MG}(\mathbf{X}_{\leq \varepsilon}^\infty, T) \text{ is not a forest}] \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Notice that under the obvious coupling, when  $\text{MG}(\mathbf{X}_{\leq \varepsilon}^\infty, t + s)$  is a forest, Lemma 2.16 implies that

$$\text{supdiam}(\text{Coal}_0(\text{Coal}_0(\mathbf{X}^\infty, t)_{\leq \varepsilon}, s)) \leq \text{supdiam}(\text{Coal}_0(\mathbf{X}_{\leq \varepsilon}^\infty, t + s)).$$

Thus, thanks to Lemmas 5.7 and 5.9 it is enough to show that with probability one,  $\mathbf{X}^\infty$  satisfies (2.3) for any  $t \geq 0$ .

Let  $P_{\mathbf{X}^\infty}$  be the distribution of  $\mathbf{X}^\infty$ . Then for  $P_{\mathbf{X}^\infty}$ -almost every  $\mathbf{X}$  and every  $t \in [0, T]$  and  $\alpha > 0$ ,

$$\begin{aligned} & \limsup_{\varepsilon \rightarrow 0} \mathbb{P}[\text{supdiam}(\text{Coal}_0(\mathbf{X}_{\leq \varepsilon}, t)) > \alpha] \\ &= \limsup_{\varepsilon \rightarrow 0} \mathbb{P}[\text{supdiam}(\text{Coal}_0(\mathbf{X}_{\leq \varepsilon}, t)) > \alpha \text{ and } \text{MG}(\mathbf{X}_{\leq \varepsilon}, T) \text{ is a forest}] \\ &\leq \limsup_{\varepsilon \rightarrow 0} \mathbb{P}[\text{supdiam}(\text{Coal}_0(\mathbf{X}_{\leq \varepsilon}, T)) > \alpha \text{ and } \text{MG}(\mathbf{X}_{\leq \varepsilon}, T) \text{ is a forest}]. \end{aligned}$$

Thus,

$$\begin{aligned} & P_{\mathbf{X}^\infty} \left\{ \mathbf{X} \in \mathcal{N}_2 : \sup_{\substack{t \leq T \\ \alpha > 0}} \limsup_{\varepsilon \rightarrow 0} \mathbb{P}[\text{supdiam}(\text{Coal}_0(\mathbf{X}_{\leq \varepsilon}, t)) > \alpha] > 0 \right\} \\ &= \sup_{\alpha > 0} P_{\mathbf{X}^\infty} \left\{ \mathbf{X} : \limsup_{\varepsilon \rightarrow 0} \mathbb{P}[\text{supdiam}(\text{Coal}_0(\mathbf{X}_{\leq \varepsilon}, T)) > \alpha] > 0 \right\} \\ &= \sup_{\alpha > 0} \sup_{\eta > 0} \lim_{\varepsilon \rightarrow 0} P_{\mathbf{X}^\infty} \left\{ \mathbf{X} : \mathbb{P}[\exists \varepsilon' \in ]0, \varepsilon], \text{supdiam}(\text{Coal}_0(\mathbf{X}_{\leq \varepsilon'}, T)) > \alpha] > \eta \right\} \\ &\leq \sup_{\alpha > 0} \sup_{\eta > 0} \lim_{\varepsilon \rightarrow 0} \frac{1}{\eta} \mathbb{P}[\exists \varepsilon' \in ]0, \varepsilon] : \text{supdiam}(\text{Coal}_0(\mathbf{X}_{\leq \varepsilon'}, T)) > \alpha]. \end{aligned}$$

Thus, using (5.3), and denoting  $x^\infty := \text{masses}(\mathbf{X}^\infty)$ , it is sufficient to prove that for any  $T \geq 0$ ,  $\alpha > 0$  and  $\tilde{\varepsilon} > 0$ ,

$$(5.4) \quad \mathbb{P} \left[ \begin{array}{c} \sup_{\varepsilon' \in ]0, \varepsilon]} \text{supdiam}(\text{Coal}_0(\mathbf{X}_{\leq \varepsilon'}^\infty, T)) > \alpha \\ \text{and} \\ \text{MG}(x_{\leq \tilde{\varepsilon}}^\infty, T) \text{ is a forest} \end{array} \right] \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Notice first that there exists a decreasing sequence of positive numbers  $(\varepsilon_p)_{p \geq 0}$  going to zero and such that

$$\forall p \in \mathbb{N}, \quad \mathbb{P}[\varepsilon_p \in x^\infty] = 0.$$

Fix  $\tilde{\varepsilon} \geq \varepsilon$  and choose the sequence so that  $\varepsilon_0 \leq \tilde{\varepsilon}$ . Then, using the obvious coupling and Lemma 2.16,

$$\begin{aligned} & \mathbb{P} \left[ \begin{array}{c} \sup_{\varepsilon' \in ]0, \varepsilon]} \text{supdiam}(\text{Coal}_0(\mathbf{X}_{\leq \varepsilon'}^\infty, T)) > \alpha \\ \text{and} \\ \text{MG}(x_{\leq \tilde{\varepsilon}}^\infty, T) \text{ is a forest} \end{array} \right] \\ &\leq \mathbb{P}[\text{supdiam}(\text{Coal}_0(\mathbf{X}_{\leq \varepsilon}^\infty, T)) > \alpha \text{ and } \text{MG}(x_{\leq \tilde{\varepsilon}}^\infty, T) \text{ is a forest}]. \end{aligned}$$

Then, we have

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \mathbb{P}(\text{supdiam}(\text{Coal}_0(\mathbf{X}_{\leq \varepsilon}^\infty, T)) > \alpha \text{ and } \text{MG}(x_{\leq \tilde{\varepsilon}}^\infty, T) \text{ is a forest}) \\ &= \lim_{m \rightarrow \infty} \mathbb{P}(\text{supdiam}(\text{Coal}_0(\mathbf{X}_{\leq \varepsilon_m}^\infty, T)) > \alpha). \end{aligned}$$

Furthermore, define  $X_{m,p} := (X_{\leq \varepsilon_m})_{>\varepsilon_p}$  for  $m \leq p$ . Then

$$\mathbb{P}(\text{supdiam}(\text{Coal}_0(X_{\leq \varepsilon_m}^\infty, T)) > \alpha) = \lim_{p \rightarrow \infty} \mathbb{P}(\text{supdiam}(\text{Coal}_0(X_{m,p}^\infty, T)) > \alpha).$$

Now, Proposition 5.3 implies that  $(\text{Coal}_{\delta^n}(X_{m,p}^n, T))$  converges in distribution to  $(\text{Coal}_0(X_{m,p}^\infty, T))$  for any  $m \leq p$ . Since we are dealing here with finite collections of m.s.m.s with positive masses, this entails that for any  $m \leq p$ ,

$$\begin{aligned} &\mathbb{P}(\text{supdiam}(\text{Coal}_0(X_{m,p}^\infty, T)) > \alpha) \\ &\leq \limsup_{n \rightarrow \infty} \mathbb{P}(\text{supdiam}(\text{Coal}_{\delta^n}(X_{m,p}^n, T)) > \alpha) \\ &\leq \limsup_{n \rightarrow \infty} \mathbb{P}(\text{supdiam}(\text{Coal}_{\delta^n}(X_{m,p}^n, T)) > \alpha \text{ and } \text{MG}(X_{\leq \tilde{\varepsilon}}^n, T) \text{ is a forest}) \\ &\quad + \limsup_{n \rightarrow \infty} \mathbb{P}(\text{MG}(X_{\leq \tilde{\varepsilon}}^n, T) \text{ is not a forest}) \\ &\leq \limsup_{n \rightarrow \infty} \mathbb{P}(\text{supdiam}(\text{Coal}_{\delta^n}(X_{\leq \varepsilon_m}^n, T)) > \alpha \text{ and } \text{MG}(X_{\leq \tilde{\varepsilon}}^n, T) \text{ is a forest}) \\ &\quad + \limsup_{n \rightarrow \infty} \mathbb{P}(\text{MG}(X_{\leq \tilde{\varepsilon}}^n, T) \text{ is not a forest}) \\ &\leq \limsup_{n \rightarrow \infty} \mathbb{P}(\text{supdiam}(\text{Coal}_{\delta^n}(X_{\varepsilon_m}^n, T)) > \alpha) \\ &\quad + \limsup_{n \rightarrow \infty} \mathbb{P}(\text{MG}(X_{\leq \tilde{\varepsilon}}^n, T) \text{ is not a forest}). \end{aligned}$$

Then, using Lemma 5.7 and the hypothesis on  $\text{supdiam}(\text{Coal}_{\delta^n}(X_\varepsilon^n, T))$  one sees that the right-hand side above goes to zero when we make  $m$  and then  $\tilde{\varepsilon}$  go to zero. This completes the proof of (5.4).  $\square$

We are now in position to prove Theorem 3.1.

**PROOF OF THEOREM 3.1.** The fact that the trajectories stay in  $\mathcal{S}$  has been proved in Lemma 5.11. The strong Markov property follows from the commutation property for coalescence, cf. Remark 2.12(ii). So to prove (i), it remains to prove that the trajectories are càdlàg (almost surely). We shall in fact prove this in the course of proving point (ii).

Let  $t^n \xrightarrow{n \rightarrow \infty} t$  and let  $T = \sup_n t^n$ . Let us fix  $\varepsilon \in ]0, 1[$  and let  $(x^n)_{n \in \mathbb{N} \cup \{\infty\}} := (\text{masses}(X^n))_{n \in \mathbb{N} \cup \{\infty\}}$ . We know that  $x^n$  converges in distribution to  $x^\infty$ . Using Skorokhod’s representation theorem, Corollary 5.6 and (5.2), we obtain that there exists  $K(\varepsilon) \in ]0, +\infty[$ ,  $\varepsilon_1 \in ]0, \varepsilon[$  and  $\varepsilon_2 \in ]0, \varepsilon_1[$  such that for every  $n \in \mathbb{N}$ , with probability larger than  $1 - \varepsilon$  the event  $\mathcal{A}_n$  holds, where  $\mathcal{A}_n$  is the event that points (a), (b) and (c) of Corollary 5.6 hold for any  $t \in [0, T]$  and  $S(x^n, T) \leq K(\varepsilon)$ .

Let  $\delta^\infty := 0$ . On this event  $\mathcal{A}_n$ , the Gromov–Hausdorff–Prokhorov distance<sup>6</sup> between a significant component of  $\text{Coal}_{\delta^n}(X^n, t)$  (at any time  $t \leq T$ ) and its heart is at most  $\alpha := \delta^n + \text{supdiam}(\text{Coal}_{\delta^n}(X_{\leq \varepsilon_1}^n, T)) + \varepsilon_1$ . Let  $\sigma$  be the function from  $\text{comp}(\text{Coal}_{\delta^n}(X^n, t)_{>\varepsilon+\alpha})$  to  $\text{comp}(\text{Coal}_{\delta^n}(X_{>\varepsilon_2}^n, t))$  which maps a component to its heart, and let  $\sigma'$  denote the function from  $\text{comp}(\text{Coal}_{\delta^n}(X_{>\varepsilon_2}^n, t)_{>\varepsilon+\alpha})$  to  $\text{comp}(\text{Coal}_{\delta^n}(X^n, t))$  which maps a component to the (unique, on  $\mathcal{A}_n$ ) component of  $\text{comp}(\text{Coal}_{\delta^n}(X^n, t))$  which contains it. These functions satisfy the hypotheses of Lemma 4.7, with  $\varepsilon$  replaced by  $\varepsilon + \alpha$ . This shows that on  $\mathcal{A}_n$ , we

<sup>6</sup>In fact, here we could talk simply of Hausdorff–Prokhorov distance since there is a trivial embedding of one measured semimetric space into the other.

have for every time  $t \leq T$  and every  $\varepsilon'_2 \leq \varepsilon_2$ :

$$\begin{aligned} &L_{\text{GHP}}(\text{Coal}_{\delta^n}(\mathbf{X}^n, t), \text{Coal}_{\delta^n}(\mathbf{X}^n_{>\varepsilon'_2}, t)) \\ &\leq 8\alpha \frac{S(x^n, t)}{\varepsilon^2} + 16(\varepsilon + \alpha) \\ &\leq 17(\delta^n + \text{supdiam}(\text{Coal}_{\delta^n}(\mathbf{X}^n_{\leq \varepsilon_1}, T)) + \varepsilon_1) \frac{8K(\varepsilon)}{\varepsilon^2} + 16\varepsilon. \end{aligned}$$

Now, recall from Lemma 5.11 that

$$\mathbb{P}(\text{supdiam}(\text{Coal}_{\delta^\infty}(\mathbf{X}^\infty_{\leq \varepsilon}, T)) > \alpha) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Thus, using the hypothesis on  $\text{supdiam}(\text{Coal}_{\delta^\infty}(\mathbf{X}^\infty_{\leq \varepsilon}, T))$ , one may choose  $\varepsilon_1$  small enough (and thus  $\varepsilon_2$  small enough) to get that for every  $n$  large enough (possibly infinite), with probability larger than  $1 - 2\varepsilon$ , we have for every time  $t \leq T$  and every  $\varepsilon'_2 \leq \varepsilon_2$ :

$$L_{\text{GHP}}(\text{Coal}_{\delta^n}(\mathbf{X}^n, t), \text{Coal}_{\delta^n}(\mathbf{X}^n_{>\varepsilon'_2}, t)) \leq 40\varepsilon.$$

Furthermore, since (c) of Corollary 5.6 holds on  $\mathcal{A}_n$ ,

$$\begin{aligned} &\|\text{masses}(\text{Coal}_{\delta^n}(\mathbf{X}^n, t)) - \text{masses}(\text{Coal}_{\delta^n}(\mathbf{X}^n_{>\varepsilon_2}, t))\|_2^2 \\ &\leq S(x^n, t) - S(x^n_{>\varepsilon_2}, t) \leq \varepsilon, \end{aligned}$$

where the first inequality comes from Lemma 4.4. This shows that<sup>7</sup>

$$(5.5) \quad \lim_{\varepsilon_2 \rightarrow 0} \lim_{N \rightarrow \infty} \sup_{\substack{n \geq N \\ n \in \mathbb{N}}} \mathbb{P}^* \left( \sup_{t \leq T} L_{2,\text{GHP}}(\text{Coal}_{\delta^n}(\mathbf{X}^n, t), \text{Coal}_{\delta^n}(\mathbf{X}^n_{>\varepsilon_2}, t)) > \varepsilon \right) = 0.$$

Let us prove that the trajectories of  $\text{Coal}(\mathbf{X}^\infty, \cdot)$  are almost surely càdlàg if  $\mathbf{X}^\infty$  belongs to  $\mathcal{S}$ . Let  $\mathbf{Y}^n := \mathbf{X}^\infty_{>\frac{1}{n}}$ . Notice that  $(\text{Coal}_0(\mathbf{Y}^n, t))_{t \geq 0}$  is càdlàg: it is right continuous and piecewise constant, with a finite number of jumps. Then equation (5.5) applied with  $\mathbf{X}^n = \mathbf{X}^\infty$  shows that the hypotheses of Lemma A.1 are satisfied with  $\omega^n = \text{Coal}(\mathbf{Y}^n, \cdot)$  and  $\omega^\infty = \text{Coal}(\mathbf{X}^\infty, \cdot)$ . This shows that the trajectories of  $\text{Coal}(\mathbf{X}^\infty, \cdot)$  are almost surely càdlàg, thus finishing to prove point (i) of the theorem.

Now, let  $(\alpha_p)_{p \geq 0}$  be a decreasing sequence of positive numbers going to zero such that

$$\forall p \in \mathbb{N}, \quad \mathbb{P}[\alpha_p \in x^\infty] = 0.$$

For any  $p$ ,  $\mathbf{X}^n_{>\alpha_p}$  converges to  $\mathbf{X}^\infty_{>\alpha_p}$  in distribution for  $L_{1,\text{GHP}}$ . Proposition 5.3 implies that  $(\text{Coal}_{\delta^n}(\mathbf{X}^n_{>\alpha_p}, t))_{t \leq T}$  converges to  $(\text{Coal}_0(\mathbf{X}^\infty_{>\alpha_p}, t))_{t \leq T}$  for the Skorokhod topology associated to  $L_{2,\text{GHP}}$ . Together with (5.5), Lemma A.2 and inequality (A.1), this proves (ii).

Furthermore, Proposition 5.3 implies that  $\text{Coal}_{\delta^n}(\mathbf{X}^n_{>\alpha_p}, t^n)$  converges in distribution to  $\text{Coal}_0(\mathbf{X}^\infty_{>\alpha_p}, t)$  as  $n$  goes to infinity for  $L_{1,\text{GHP}}$ , and thus for  $L_{2,\text{GHP}}$ . Together with (5.5) we obtain that  $\text{Coal}_{\delta^n}(\mathbf{X}^n, t^n)$  converges to  $\text{Coal}_0(\mathbf{X}^\infty, t)$  in distribution for  $L_{2,\text{GHP}}$ . This proves (iii).  $\square$

REMARK 5.12. Notice that if  $\delta_n > 0$  and  $\mathbf{X}^n \in \mathcal{N}_2 \setminus \mathcal{N}_1$ ,  $\text{Coal}_{\delta_n}(\mathbf{X}^n, t)$  is not in  $\mathcal{N}_2$  for  $t > 0$  since the components are not totally bounded. Thus, the terms “convergence in distribu-

<sup>7</sup>At this stage, since we did not proved yet the càdlàg property, it is not guaranteed that the event in the probability is measurable, due to the uncountable supremum. This is why we use the outer measure  $\mathbb{P}^*$ . However, once the càdlàg property is proved, one may remove the star.

tion for  $L_{2,GHP}$ ” in Theorem 3.1 should be understood in a larger space, where components are allowed not to be totally bounded.

However, we do not insist on this because when  $\delta^n > 0$ , we shall always use Theorem 3.1 with  $X^n \in \mathcal{N}_1$  for any  $n \in \mathbb{N}$ , in which case  $\text{Coal}_{\delta_n}(X^n, t)$  is in  $\mathcal{N}_1$  for any  $t$ .

Finally, we finish this section by exhibiting a sufficient condition for (3.1) which will be useful in Section 5.4. In words, it says that when the diameter of a component of  $\text{Coal}_{\delta_n}(X^n, T)$  goes to zero when its size goes to zero, uniformly in  $n$ , then one may restrict in (3.1) to components which are attached to a significant component by at most one edge of the multigraph  $\text{MG}(X^n, T)$ . In Lemma 5.13, we consider the obvious coupling between  $\text{Coal}_{\delta_n}(X^n, T)$ ,  $\text{Coal}_{\delta_n}(X^n_{\leq \varepsilon_1}, T)$ ,  $\text{MG}(X^n_{> \varepsilon_1}, T)$  and  $\text{MG}(X^n, T)$ .

LEMMA 5.13. *Let  $X^n = (X^n, d^n, \mu^n)$ ,  $n \geq 0$  be a sequence of random variables in  $\mathcal{N}_2$  and  $(\delta^n)_{n \geq 0}$  a sequence of nonnegative real numbers. For  $T, \varepsilon$  and  $\varepsilon_1 > 0$ , let  $C_n(\varepsilon, \varepsilon_1, T)$  denote the set of components  $m$  of  $\text{Coal}_{\delta_n}(X^n_{\leq \varepsilon_1}, T)$  which are included in a component of  $\text{Coal}_{\delta_n}(X^n, T)$  of size at least  $\varepsilon$  and such that there is exactly one edge between  $m$  and  $\text{MG}(X^n_{> \varepsilon_1}, T)$  in  $\text{MG}(X^n, T)$ . Let  $d_n(\varepsilon, \varepsilon_1, T)$  denote the supremum of the diameters of elements of  $C_n(\varepsilon, \varepsilon_1, T)$ . Suppose that:*

(i)  $(\text{masses}(X^n))_{n \in \mathbb{N}}$  converges in distribution, in  $\ell^2$  to a random variable  $x^\infty$  as  $n$  goes to infinity,

(ii) for any  $\alpha > 0$

$$\limsup_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(\text{supdiam}(\text{Coal}_{\delta_n}(X^n, T)_{< \varepsilon}) > \alpha) = 0,$$

(iii) for any  $\alpha > 0$  and  $\varepsilon > 0$ ,

$$\limsup_{\varepsilon_1 \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(d_n(\varepsilon, \varepsilon_1, T) > \alpha) = 0.$$

Then

$$\limsup_{n \in \mathbb{N}} \mathbb{P}(\text{supdiam}(\text{Coal}_{\delta^n}(X^n_{\leq \varepsilon}, T)) > \alpha) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

PROOF. Let  $A_n(\varepsilon, \varepsilon_1)$  denote the event that there is a component of  $\text{Coal}_{\delta_n}(X^n_{\leq \varepsilon_1}, T)$  which is included in a component of  $\text{Coal}_{\delta_n}(X^n, T)$  of size at least  $\varepsilon$  and such that there is more than one edge between  $m$  and  $\text{MG}(X^n_{> \varepsilon_1}, T)$  in  $\text{MG}(X^n, T)$ . Using the Skorokhod representation theorem, Corollary 5.6 and (5.2), one sees that for any  $\varepsilon > 0$ ,

$$\limsup_{n \rightarrow \infty} \mathbb{P}(A_n(\varepsilon, \varepsilon_1)) \xrightarrow{\varepsilon_1 \rightarrow 0} 0.$$

Recall also Lemma 5.7. Letting  $B_n(\varepsilon_1)$  denote the event that there is more than one edge between a connected component of  $\text{Coal}_{\delta^n}(X^n_{\leq \varepsilon_1}, T)$  and  $\text{Coal}_{\delta^n}(X^n_{> \varepsilon_1}, T)$ , Lemma 5.7 shows that

$$\limsup_{n \rightarrow \infty} \mathbb{P}(B_n(\varepsilon_1)) \xrightarrow{\varepsilon_1 \rightarrow 0} 0.$$

On  $B_n(\varepsilon_1)^c$ , the diameter of a component of  $\text{Coal}_{\delta^n}(X^n_{\leq \varepsilon_1}, T)$  is at most the diameter of the component of  $\text{Coal}_{\delta^n}(X^n, T)$  which contains it. Thus, for any  $\varepsilon_1$  and  $\varepsilon$ ,

$$\begin{aligned} & \mathbb{P}(\text{supdiam}(\text{Coal}_{\delta^n}(X^n_{\leq \varepsilon_1}, T)) > \alpha) \\ & \leq \mathbb{P}(d_n(\varepsilon, \varepsilon_1, T) > \alpha) + \mathbb{P}(\text{supdiam}(\text{Coal}_{\delta_n}(X^n, T)_{< \varepsilon}) > \alpha) \\ & \quad + \mathbb{P}(A_n(\varepsilon, \varepsilon_1)) + \mathbb{P}(B_n(\varepsilon_1)), \end{aligned}$$

this gives the result.  $\square$



5.4. *Convergence of the coalescent on Erdős–Rényi random graphs.* In this section, we prove Theorem 3.5. Recall that  $\overline{\mathcal{G}}_{n,\lambda}$  is the element of  $\mathcal{N}_2^{\text{graph}}$  obtained from  $\mathcal{G}(n, p(\lambda, n))$  by assigning to each edge a length  $n^{-1/3}$  and to each vertex a mass  $n^{-2/3}$ . We know by Theorem 2.26 that  $\overline{\mathcal{G}}_{n,\lambda}$  converges in distribution (for  $L_{2,\text{GHP}}$ ) to  $\mathcal{G}_\lambda$ . In view of Theorem 3.1, it is sufficient to prove that for any  $T$  and  $\alpha > 0$ :

$$(5.6) \quad \limsup_{n \rightarrow \infty} \mathbb{P}(\text{supdiam}(\text{Coal}_{n^{-1/3}}((\overline{\mathcal{G}}_{n,\lambda})_{\leq \varepsilon}, T)) > \alpha) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

To this end, we shall use Lemma 5.13. The notion of depth-first exploration process on a finite graph  $G = (V, E)$ , as defined in [3], Sections 1 and 2, will be useful. This depth-first exploration process defines an order  $\sigma$  on  $V$  (a bijection from  $V$  to  $\{0, \dots, n - 1\}$  with  $n = |V|$ ), a height process  $h$  (from  $\{1, \dots, |V|\}$  to  $\mathbb{N}$ ) and, for each connected component  $C$ , a rooted tree  $(\rho_C, \mathcal{T}_C)$  such that  $\rho_C$  is the first vertex visited in  $C$  and  $\mathcal{T}_C$  spans  $C$ . Furthermore, denoting by  $d_{\mathcal{T}_C}$  the graph metric on  $\mathcal{T}_C$ , one has, for any connected component  $C$  of  $G$ :

$$\forall i \in C, \quad h(\sigma(i)) = d_{\mathcal{T}_C}(\rho_C, i).$$

We shall need the following lemma.

LEMMA 5.14. *Let  $G$  be a finite graph with vertex set  $V$ . Let  $h$  be the height process associated to the depth-first exploration process on  $G$  and denote by  $\sigma$  the order induced by the depth-first exploration on  $V$ . Let  $I$  denote a subset of  $V$  such that the subgraph induced by  $I$  in  $G$  is connected and such that, denoting by  $C$  the connected component of  $G$  containing  $I$ , either  $C = I$  or there is exactly one edge connecting  $I$  to  $C \setminus I$ . Then*

$$\text{diam}(I) \leq 2 + 3 \max_{i,j \in C: |\sigma(i) - \sigma(j)| \leq |I|} |h(\sigma(i)) - h(\sigma(j))|,$$

where the diameter is computed either for the distance  $d_{\mathcal{T}_C}$  or for the graph distance on  $G$ .

PROOF. If  $C = I$ , the result is trivial, so let us suppose that  $I \neq C$  and let  $x$  denote the unique vertex of  $I$  connected to  $C \setminus I$ . We consider two cases.

1. Suppose that  $\rho_C \in C \setminus I$ . In this case, the vertices of  $I$  are explored consecutively, that is,  $\sigma(I)$  is an interval of length  $|I|$ , and the first vertex explored in  $I$  is  $x$ . Thus, for any  $z \in I$ ,

$$d_{\mathcal{T}_C}(z, \rho_C) = d_{\mathcal{T}_C}(z, x) + d_{\mathcal{T}_C}(x, \rho_C).$$

Thus, if  $z$  and  $z'$  denote two vertices of  $I$ ,

$$\begin{aligned} d_{\mathcal{T}_C}(z, z') &\leq d_{\mathcal{T}_C}(z, x) + d_{\mathcal{T}_C}(z', x) \\ &= d_{\mathcal{T}_C}(z, \rho_C) - d_{\mathcal{T}_C}(x, \rho_C) + d_{\mathcal{T}_C}(z', \rho_C) - d_{\mathcal{T}_C}(x, \rho_C) \\ &= h(\sigma(z)) - h(\sigma(x)) + h(\sigma(z')) - h(\sigma(x)) \\ &\leq 2 \max_{i,j \in C: |\sigma(i) - \sigma(j)| \leq |I|} |h(\sigma(i)) - h(\sigma(j))| \end{aligned}$$

since  $\sigma(I)$  is an interval of length  $|I|$ .

2. Suppose that  $\rho_C \in I$ . Then, the vertices of  $C \setminus I$  are explored consecutively. Let  $J$  denote the subset of  $I$  composed of the vertices of  $I$  which are explored before those of  $C \setminus I$ .  $J$  contains  $\rho_C$  and  $x$  (which might be the same vertex). Notice that the vertices of  $J$  are explored consecutively (i.e.,  $\sigma(J)$  is an interval of length  $|J|$ ), and those of  $I \setminus J$  also (i.e.,  $\sigma(I \setminus J)$  is an interval of length  $|I| - |J|$ ). Notice also that  $J$  forms a subtree of  $\mathcal{T}_C$ . When  $z$  and  $z'$  belong to  $C$ , we shall denote by  $[z, z']$  the unique path from  $z$  to  $z'$  in  $\mathcal{T}_C$  and by  $T(z)$  the subtree above  $z$ , that is, the set of vertices  $u$  such that  $z \in [\rho_C, u]$ . Notice that if  $z \in I \setminus J$ , then  $T(z)$  is included in  $I \setminus J$ , and that if  $u \in T(z)$ ,

$$(5.7) \quad d_{\mathcal{T}_C}(z, u) = h(\sigma(u)) - h(\sigma(z)).$$

Now, let  $z$  and  $z'$  belong to  $I$  and consider the following cases:

(a) If  $z$  and  $z'$  belong to  $J$ , then

$$\begin{aligned} d_{\mathcal{T}_C}(z, z') &\leq d_{\mathcal{T}_C}(z, \rho_C) + d_{\mathcal{T}_C}(z', \rho_C) \\ &= h(\sigma(z)) - h(\sigma(\rho_C)) + h(\sigma(z')) - h(\sigma(\rho_C)) \\ &\leq 2 \max_{i, j \in C: |\sigma(i) - \sigma(j)| \leq |J|} |h(\sigma(i)) - h(\sigma(j))| \end{aligned}$$

since  $\rho_C$  belongs to  $J$ .

(b) If  $z \in J$  and  $z' \in I \setminus J$ , let  $y$  denote the vertex of  $[\rho_C, z'] \cap J$  closest (for  $d_{\mathcal{T}_C}$ ) to  $z'$  and  $y'$  the vertex of  $[\rho_C, z'] \cap I \setminus J$  closest to  $\rho_C$ . Since  $J$  is explored consecutively and  $\mathcal{T}_C$  is the depth-search tree,  $y$  and  $y'$  are neighbors in  $\mathcal{T}_C$ ,  $[\rho_C, y]$  is included in  $J$  and  $z' \in T(y')$ . We already proved in point (a) above that

$$d_{\mathcal{T}_C}(z, y) \leq 2 \max_{i, j \in C: |\sigma(i) - \sigma(j)| \leq |J|} |h(\sigma(i)) - h(\sigma(j))|.$$

Thus, using (5.7):

$$\begin{aligned} d_{\mathcal{T}_C}(z, z') &\leq d_{\mathcal{T}_C}(z, y) + d_{\mathcal{T}_C}(y, y') + d_{\mathcal{T}_C}(y', z') \\ &\leq 2 \max_{i, j \in C: |\sigma(i) - \sigma(j)| \leq |J|} |h(\sigma(i)) - h(\sigma(j))| \\ &\quad + 1 + h(\sigma(z')) - h(\sigma(y')) \\ &\leq 2 \max_{i, j \in C: |\sigma(i) - \sigma(j)| \leq |J|} |h(\sigma(i)) - h(\sigma(j))| \\ &\quad + 1 + \max_{i, j \in C: |\sigma(i) - \sigma(j)| \leq |I| - |J|} |h(\sigma(i)) - h(\sigma(j))|. \end{aligned}$$

(c) If  $z$  and  $z'$  belong to  $I \setminus J$ , the arguments are similar: one finds  $y$  and  $y'$  in  $I \setminus J$ ,  $t$  and  $t'$  in  $J$  such that  $z \in T(y)$ ,  $z' \in T(y')$ ,  $t$  is a neighbor of  $y$  and  $t'$  is a neighbor of  $y'$ . Notice that we already proved that

$$d_{\mathcal{T}_C}(t, t') \leq 2 \max_{i, j \in C: |\sigma(i) - \sigma(j)| \leq |J|} |h(\sigma(i)) - h(\sigma(j))|.$$

Then, using (5.7),

$$\begin{aligned} d_{\mathcal{T}_C}(z, z') &\leq d_{\mathcal{T}_C}(z, y) + d_{\mathcal{T}_C}(y, t) + d_{\mathcal{T}_C}(t, t') \\ &\quad + d_{\mathcal{T}_C}(t', y') + d_{\mathcal{T}_C}(y', z') \\ &\leq h(\sigma(z)) - h(\sigma(y)) + 1 + d_{\mathcal{T}_C}(t, t') + 1 \\ &\quad + h(\sigma(z')) - h(\sigma(y')) \\ &\leq 2 + 3 \max_{i, j \in C: |\sigma(i) - \sigma(j)| \leq |I|} |h(\sigma(i)) - h(\sigma(j))|. \quad \square \end{aligned}$$

Let us denote by  $h_{n,\lambda}$  the height process associated to the depth-first exploration process on  $\overline{\mathcal{G}}_{n,\lambda}$ , and let  $\overline{h}_{n,\lambda}$  be its rescaled version:

$$\overline{h}_{n,\lambda}(x) := \frac{1}{n^{1/3}} h_{n,\lambda}(xn^{2/3}).$$

Now let us consider the depth-first exploration process of  $\text{Coal}_{n^{-1/3}}(\overline{\mathcal{G}}_{n,\lambda}, T)$ . When  $\mathcal{P}_T^+$  has intensity  $\gamma$ ,  $N^+(\mathcal{G}(n, p), \mathcal{P}_T^+)$  is equal in distribution to  $\mathcal{G}(n, p')$  with

$$p' = p + (1 - p)(1 - e^{-\gamma T}).$$

When  $\gamma = n^{-4/3}$ ,  $p' = p(\lambda'_n, n)$  with

$$\lambda'_n \xrightarrow[n \rightarrow \infty]{} \lambda + T$$

and  $\text{Coal}_{n^{-1/3}}(\overline{\mathcal{G}}_{n,\lambda}, T)$  is equal, in distribution, to  $\overline{\mathcal{G}}_{n,\lambda'_n}$ . The difference between  $\lambda'_n$  and  $\lambda_n + T$  is unimportant for us (for instance using Lemma 5.7,  $\text{supdiam}(\text{Coal}_{n^{-1/3}}((\overline{\mathcal{G}}_{n,\lambda})_{\leq \varepsilon}, T))$  is essentially nondecreasing in  $T$ ), so we shall continue as if  $\lambda'_n = \lambda + T$ . Under this approximation, the rescaled version of the height process associated to the exploration of  $\text{Coal}_{n^{-1/3}}(\overline{\mathcal{G}}_{n,\lambda}, T)$  has the same distribution as  $\overline{h}^{n,\lambda+T}$ , and we shall adopt the same notation to keep things simple.

Now, we shall use Lemma 5.13 to prove (5.6). Let us use the notation of Lemma 5.13, with  $X^n = \overline{\mathcal{G}}_{n,\lambda}$  and  $\delta^n = n^{-1/3}$ . Hypothesis (i) of Lemma 5.13 is satisfied, so we only need to prove hypotheses (ii) and (iii). Lemma 5.14 ensures that

$$d_n(\varepsilon, \varepsilon_1) \leq \frac{2}{n^{1/3}} + 3 \sup_{\substack{x,y \in \mathbb{R}^+ \\ |x-y| \leq \varepsilon_1}} |\overline{h}^{n,\lambda+T}(x) - \overline{h}^{n,\lambda+T}(y)|,$$

where the supremum is restricted to pairs  $(x, y)$  such that  $x$  and  $y$  belong to a same excursion of  $\overline{h}^{n,\lambda+T}$  above zero, of length at least  $\varepsilon$ . Thus, we see that hypotheses (ii) and (iii) of Lemma 5.13 will be established if we can prove that for any  $\alpha > 0$ ,

$$(5.8) \quad \limsup_{n \rightarrow \infty} \mathbb{P}\left( \sup_{\substack{x,y \in \mathbb{R}^+ \\ |x-y| \leq \varepsilon_1}} |\overline{h}^{n,\lambda+T}(x) - \overline{h}^{n,\lambda+T}(y)| > \alpha \right) \xrightarrow{\varepsilon_1 \rightarrow 0} 0,$$

with the supremum restricted to pairs  $(x, y)$  such that  $x$  and  $y$  belong to the same excursion of  $\overline{h}^{n,\lambda+T}$  above zero.

Let  $B^\lambda$  be a Brownian motion with quadratic drift, defined by  $B_t^\lambda = B_t + \lambda t - \frac{t^2}{2}$  with  $B$  a standard Brownian motion. Let  $W^\lambda$  be  $B^\lambda$  reflected above its current minimum:

$$W_t^\lambda := B_t^\lambda - \min_{0 \leq s \leq t} B_s^\lambda.$$

Then (5.8) is a consequence of the fact that  $\overline{h}^{n,\lambda+T}$  converges in distribution to  $2W^\lambda$  for the sup norm on  $\mathbb{R}^+$ . It seems however that this convergence is not written in the literature, so in order to use only available sources, one may rely on the work done in [3] as follows. One may separate the analysis of the supremum on the  $N$  largest excursions and on the others. Let  $B_{N,n}(\varepsilon)$  be the event that the maximal height of the  $i$ th largest component in  $\overline{\mathcal{G}}_{n,\lambda+T}$  exceeds  $\varepsilon$  for some  $i > N$ . The equation page 402 below equation (24) in [3] shows that for any  $\varepsilon > 0$ ,

$$(5.9) \quad \lim_{N \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(B_{N,n}(\varepsilon)) = 0.$$

Then, for a fixed  $N$ , one may argue as in the proof of [3], Theorem 24, page 398: conditionally on the sizes, the rescaled height processes associated to those components are independent and each one converges in distribution (for the uniform topology) to a continuous excursion (a tilted Brownian excursion). Together with the convergence of the sizes and Skorokhod's representation theorem, this proves that the  $N$  largest excursions of  $\overline{h}^{n,\lambda+T}$  converges as a vector in  $C([0, +\infty[)^N$  to a random vector of continuous functions with bounded support. This implies that (5.8) holds when the supremum is restricted to pairs  $(x, y)$  such that  $x$  and  $y$  belong to one of the  $N$  largest components of  $\overline{h}^{n,\lambda+T}$ . Together with (5.9), this shows (5.8) and this completes the proof of Theorem 3.5.

**6. Proofs of the results for fragmentation.** The main goal of this section is to prove the Feller property for fragmentation on  $\mathcal{N}_2^{\text{graph}}$ , Theorem 3.2 and to apply it to prove Theorem 3.6. It is very close to the work performed in [4], which proves a continuity result for a fragmentation restricted to the core of a graph (and stopped when you get a tree). The main difference is that we want in addition to perform fragmentation on the tree part of the graphs. Another technical difference will be detailed at the beginning of Section 6.4. Unfortunately, those differences force us to make substantial modifications to the arguments of [4].

6.1. *Notation.* We need to introduce a few more definitions to deal with fragmentation of  $\mathbb{R}$ -graphs. For more details, we refer to [4].

A *multigraph with edge-lengths* is a triple  $(V, E, (\ell(e))_{e \in E})$  where  $(V, E)$  is a finite connected multigraph and for every  $e \in E$ ,  $\ell(e)$  is a strictly positive number. One may associate to such a multigraph with edge-lengths a compact  $\mathbb{R}$ -graph with a finite number of leaves by performing on  $V$  (seen as a metric space as the disjoint union of its elements) the  $\ell(e)$ -gluing along  $e$  for each edge  $e \in E$ .

Until the end of the article, we shall say that an  $\mathbb{R}$ -graph is finite if it is compact and has a finite number of leaves. Equivalently, it can be associated to a multigraph with edge-lengths as above. This terminology, applied to trees, comes from [21].

Let  $G$  be an  $\mathbb{R}$ -graph. When there is only one geodesic between  $x$  and  $y$  in  $G$ , we denote by  $[x, y]$  its image. Recall the notion of the core of  $G$  defined in Section 2.6. If  $S$  is a closed connected subset of  $G$  containing  $\text{core}(G)$ , then for any  $x \in G$ , there is a unique shortest path  $\gamma_x$  going from  $x$  to  $S$ . We denote by  $p_S(x)$  the unique point belonging to  $\gamma_x \cap S$ . When  $G$  is not a tree and  $S = \text{core}(G)$ , we let  $\alpha_G(x) := p_S(x)$ .

For any  $\eta > 0$  and any  $\mathbb{R}$ -graph  $G$  which is not a tree, let

$$R_\eta(G) := \text{core}(G) \cup \{x \in G \text{ s.t. } \exists y : x \in [y, \alpha_G(y)] \text{ and } d(y, x) \geq \eta\}.$$

In words, seeing  $G$  as rooted at the core, in  $R_\eta(G)$  we keep the root (i.e., the core) and all the points whose subtree has height at least  $\eta$ . When  $(T, \rho)$  is a rooted  $\mathbb{R}$ -tree, we let  $R_\eta(T) := R_\eta(T, \rho)$  be defined as above, with  $\alpha_G(y)$  replaced by the root  $\rho$  and  $\text{core}(G)$  replaced by  $\{\rho\}$ . Thus the definition of  $R_\eta(G)$  extends the definition of  $R_\eta(T)$  for a rooted  $\mathbb{R}$ -tree  $(T, \rho)$  in [21]. Notably, [21], Lemma 2.6(i), shows that for any  $\eta > 0$ ,  $R_\eta(G)$  is a finite  $\mathbb{R}$ -graph. For a (nonrooted) measured tree  $(T, \mu)$ , with  $\mu$  a positive finite measure, we let  $R_\eta(T)$  denote  $R_\eta(T, \rho)$  where  $\rho$  is a random root, sampled according to  $\frac{\mu(\cdot)}{\mu(T)}$ . Finally, if  $G$  belongs to  $\mathcal{N}_2^{\text{graph}}$ , we let

$$R_\eta(G) := \bigcup_{m \in \text{comp}(G)} R_\eta(m),$$

where the random roots of components which are trees are sampled independently.

The  $\varepsilon$ -enlargement of a correspondence  $\mathcal{R} \in C(X, X')$  is defined as

$$\mathcal{R}^\varepsilon := \{(x, x') \in X \times X' : \exists (y, y') \in \mathcal{R}, d(x, y) \vee d(x', y') \leq \varepsilon\}.$$

It is a correspondence containing  $\mathcal{R}$  with distortion at most  $\text{dis}(\mathcal{R}) + 4\varepsilon$ .

If  $\mathcal{R} \in C(X, X')$ , two Borel subsets  $A \subset X$  and  $B \subset X'$  are said to be in correspondence through  $\mathcal{R}$  if  $\mathcal{R} \cap (A \times B) \in C(A, B)$ .

Let  $\varepsilon > 0$ . If  $X$  and  $X'$  are  $\mathbb{R}$ -graphs with surplus at least 2, an  $\varepsilon$ -overlay is a correspondence  $\mathcal{R} \in C(X, X')$  with distortion less than  $\varepsilon$  and such that there exists a multigraph isomorphism  $\chi$  between the kernels  $\ker(X)$  and  $\ker(X')$  satisfying:

1.  $\forall v \in k(X), (v, \chi(v)) \in \mathcal{R}$ .
2. For every  $e \in e(X)$ ,  $e$  and  $\chi(e)$  are in correspondence through  $\mathcal{R}$  and  $|\ell_X(e) - \ell_{X'}(\chi(e))| \leq \varepsilon$ .

If  $X$  and  $X'$  have surplus one, an  $\varepsilon$ -overlay is a correspondence with distortion less than  $\varepsilon$  such that the unique cycles of  $X$  and  $X'$  are in correspondence and the difference of their lengths is at most  $\varepsilon$ . If  $X$  and  $X'$  are trees, an  $\varepsilon$ -overlay is simply a correspondence with distortion less than  $\varepsilon$ .

We let  $\mathcal{N}_2^{\text{tree}}$  be the set of elements  $X \in \mathcal{N}_2^{\text{graph}}$  whose components are trees.

6.2. *Reduction to finite  $\mathbb{R}$ -graphs.* The following lemmas will be useful to reduce the proof of the Feller property to finite  $\mathbb{R}$ -graphs, notably to adapt the arguments of [21].

LEMMA 6.1. *Let  $\eta \in (0, 1]$  and  $T > 0$ . Let  $\mathbf{G}$  belong to  $\mathcal{N}_2^{\text{graph}}$ . Let  $S$  be a closed connected subset of  $\mathbf{G}$  such that  $R_\eta(\mathbf{G}) \subset S \subset \mathbf{G}$ . Suppose that for each component  $H$  of  $\mathbf{G}$ ,  $S \cap H$  is a connected  $\mathbb{R}$ -graph. Let  $\mathcal{S} := (S, d|_{S \times S}, p_S \sharp \mu)$ . Then, with probability at least  $1 - T\eta^{1/7}$ , for any  $t \in [0, T]$ , under the obvious coupling,*

$$L_{2,\text{GHP}}^{\text{surplus}}(\text{Frag}(\mathbf{G}, t), \text{Frag}(\mathcal{S}, t)) \leq 34\eta^{1/7} \left( 1 + \sum_{H \in \text{comp}(\mathbf{G})} \mu(H)^2 \right)^2.$$

PROOF. Let  $\mathcal{P}$  be a Poisson random set of intensity measure  $\ell_G \otimes \text{leb}_{\mathbb{R}}^+$  on  $G \times \mathbb{R}^+$  and let us use it to perform the fragmentation on  $S$  and  $G$ . Define

$$G_t^\eta := \{x \in G \setminus S \text{ s.t. } \exists y \in \mathcal{P}_t \cap (G \setminus S) \cap [x, \alpha_G(x)]\}.$$

Notice that a component  $m$  of  $\text{Frag}(\mathcal{S}, t)$  is endowed with the distance  $d|_{m \times m}$  and the measure  $(p_S \sharp \mu)|_m$ , while a component  $m$  of  $\text{Frag}(\mathbf{G}, t)$  is endowed with the distance  $d|_{m \times m}$  and the measure  $\mu|_m$ .

If  $m$  is a component of  $\text{Frag}(\mathbf{G}, \mathcal{P}_t)$  such that  $m \cap S = \emptyset$  then  $m \subset G_t^\eta$ . Notably, if  $t \in [0, T]$ ,  $H \in \text{comp}(\mathbf{G})$ ,  $m$  is a component of  $\text{Frag}(\mathbf{G}, \mathcal{P}_t)$  included in  $H$  and  $\mu(m) > \mu(G_T^\eta \cap H)$ , then  $m$  must intersect  $S$ . Furthermore, if  $m \cap S \neq \emptyset$  then  $m \cap S$  is a component of  $\text{Frag}(\mathcal{S}, \mathcal{P}_t)$ .

Let  $H \in \text{comp}(\mathbf{G})$ . For any component  $m$  of  $\text{Frag}(H, \mathcal{P}_t)$  such that  $m \cap S \neq \emptyset$ , we claim that

$$(6.1) \quad d_{\text{GHP}}(m, m \cap S) \leq \eta \vee \mu(G_T^\eta \cap H).$$

Indeed, let  $\mathcal{R} := \{(x, p_S(x)) : x \in m\}$ , which has distortion at most  $2\eta$ , and define  $\pi := ((Id, p_S) \sharp \mu)|_{m \times m \cap S}$ . Then  $\pi(\mathcal{R}^c) = 0$  and

$$\begin{aligned} D(\pi; \mu|_m, (p_S \sharp \mu)|_m) &= \sup_{A \in \mathcal{B}(m \cap S)} \mu(p_S^{-1}(A) \setminus m) \\ &= \mu(p_S^{-1}(m) \setminus m) \\ &\leq \mu(G_t^\eta \cap H) \\ &\leq \mu(G_T^\eta \cap H). \end{aligned}$$

This shows (6.1). Furthermore,

$$\begin{aligned} &\|\text{masses}(\text{Frag}(\mathbf{G}, t)) - \text{masses}(\text{Frag}(\mathcal{S}, t))\|_2^2 \\ (6.2) \quad &\leq \sum_{\substack{m \in \text{Frag}(\mathbf{G}, t) \\ m \cap S \neq \emptyset}} \mu(p_S^{-1}(m) \setminus m)^2 + \sum_{\substack{m \in \text{Frag}(\mathbf{G}, t) \\ m \cap S = \emptyset}} \mu(m)^2 \\ &\leq 2 \sum_{H \in \text{comp}(\mathbf{G})} \mu(G_t^\eta \cap H)^2 \end{aligned}$$

$$(6.3) \quad \leq 2 \sum_{H \in \text{comp}(\mathbf{G})} \mu(G_T^\eta \cap H)^2.$$

Using Fubini's theorem,

$$\mathbb{E}[\mu(G_T^\eta \cap H)^2] \leq \mu(H)^2(1 - e^{-\eta T}) \leq \mu(H)^2 \eta T.$$

Thus,

$$\mathbb{P}\left(\sum_{H \in \text{comp}(\mathbf{G})} \mu(G_T^\eta \cap H)^2 \geq \eta^{6/7} \sum_{H \in \text{comp}(\mathbf{G})} \mu(H)^2\right) \leq T\eta^{1/7}.$$

Now, let us place ourselves on the event

$$\mathcal{E} := \left\{ \sum_{H \in \text{comp}(\mathbf{G})} \mu(G_T^\eta \cap H)^2 < \eta^{6/7} \sum_{H \in \text{comp}(\mathbf{G})} \mu(H)^2 \right\}$$

and define  $\alpha := \eta^{3/7} \sqrt{1 + \sum_{H \in \text{comp}(\mathbf{G})} \mu(H)^2}$ . Notice that on  $\mathcal{E}$ , we have for any  $H \in \text{comp}(\mathbf{G})$ :

$$\mu(G_T^\eta \cap H) \leq \sqrt{\sum_{H \in \text{comp}(\mathbf{G})} \mu(G_T^\eta \cap H)^2} \leq \alpha.$$

Let  $\sigma$  assign to each component of  $\text{Frag}(\mathbf{S}, t)$  the component of  $\text{Frag}(\mathbf{G}, t)$  which contains it, and let  $\sigma'$  assign to a component  $m$  of  $\text{comp}((\text{Frag}(\mathbf{G}, t))_{>\alpha^{1/3}+\alpha})$  the component  $m \cap \mathbf{S}$  of  $\text{comp}(\text{Frag}(\mathbf{S}, t))$ . From (6.1), we deduce that for any component  $m$  of  $\text{Frag}(\mathbf{S}, t)$ ,

$$d_{\text{GHP}}(m, \sigma(m)) \leq \alpha$$

and notice that  $m$  and  $\sigma(m)$  have the same surplus. Also, for any component  $m'$  of  $\text{Frag}(\mathbf{G}, t)$ ,

$$d_{\text{GHP}}(m', \sigma'(m')) \leq \alpha,$$

and  $m'$  and  $\sigma'(m')$  have the same surplus. According to Lemma 4.7, this shows that on the event  $\mathcal{E}$ :

$$\begin{aligned} &L_{\text{GHP}}^{\text{surplus}}(\text{Frag}(\mathbf{G}, t), \text{Frag}(\mathbf{S}, t)) \\ &\leq 8\alpha \frac{\sum_{m \in \text{comp}(\text{Frag}(\mathbf{G}, t))} \mu(m)^2}{\alpha^{2/3}} + 16(\alpha^{1/3} + \alpha) \\ &\leq 16\alpha + \alpha^{1/3} \left( 16 + 8 \sum_{H \in \text{comp}(\mathbf{G})} \mu(H)^2 \right). \end{aligned}$$

And, thanks to (6.3),

$$\| \text{masses}(\text{Frag}(\mathbf{G}, t)) - \text{masses}(\text{Frag}(\mathbf{S}, t)) \|_2 \leq \eta^{6/7} \sum_{H \in \text{comp}(\mathbf{G})} \mu(H)^2$$

which shows the result.  $\square$

We shall need a slight variation of the preceding lemma at time zero for rooted trees. When  $(X, \rho)$  and  $(X', \rho')$  are two rooted m.s.m.s, rooted respectively at  $\rho$  and  $\rho'$ , we define

$$d_{\text{GHP}}^{\text{root}}((X, \rho), (X', \rho')) = \inf_{\substack{\pi \in M(X, X') \\ \mathcal{R} \in C_{\rho, \rho'}(X, X')}} \left\{ D(\pi; \mu, \mu') \vee \frac{1}{2} \text{dis}(\mathcal{R}) \vee \pi(\mathcal{R}^c) \right\},$$

where  $C_{\rho, \rho'}(X, X')$  is the set of correspondences between  $X$  and  $X'$  which contain  $(\rho, \rho')$ .

Using natural correspondences and couplings, one may see that  $R_\eta(T)$  approximates nicely a rooted tree.

LEMMA 6.2. *Let  $\mathbf{T} = (T, d, \mu)$  be a measured real tree and  $\rho \in T$  a root. Let  $\mathbf{R}_\eta(T)$  be the measured real tree  $R_\eta(T)$  equipped with the measure  $p_{R_\eta(T)} \sharp \mu$ . Then*

$$d_{\text{GHP}}^{\text{root}}((\mathbf{T}, \rho), (\mathbf{R}_\eta(T), \rho)) \leq \eta.$$

PROOF. Let  $p := p_{R_\eta(T)}$ . Take

$$\mathcal{R} := \{(x, p(x)) \in T \times R_\eta(T) : x \in T\}$$

which is a correspondence containing  $(\rho, \rho)$  of distortion at most  $2\eta$ . Then take

$$\pi := (Id, p)\sharp\mu$$

that is,  $\pi(C) = \mu(\{x \in T : (x, p(x)) \in C\})$ , which verifies

$$D(\pi; \mu, p\sharp\mu) = 0$$

and  $\pi(\mathcal{R}^c) = 0$ .  $\square$

6.3. *The Feller property for trees.* If  $x$  and  $y$  belong to a rooted tree  $(T, \rho)$ , we denote by  $[x, y]$  the unique geodesic between  $x$  and  $y$ , and we say that  $x \leq y$  if  $x$  belongs to  $[\rho, y]$ . The subtree above  $x$  is then defined as

$$\{y \in T : x \leq y\}.$$

The following lemma is a slight extension of [21], Lemma 6.3, designed to take measures into account.

LEMMA 6.3. *Let  $\mathbf{T} = (T, d, \mu)$  be a measured real tree,  $\rho \in T$  and  $\varepsilon > 0$ . There exists  $\eta > 0$  (depending only on  $\varepsilon$ ), and  $\delta > 0$  (depending on  $\mathbf{T}$ ,  $\rho$  and  $\varepsilon$ ) such that if  $\mathbf{T}' = (T', d', \mu')$  is a measured  $\mathbb{R}$ -tree rooted at  $\rho'$  and  $d_{\text{GHP}}^{\text{root}}((\mathbf{T}, \rho), (\mathbf{T}', \rho')) < \delta$ , then there exist finite subtrees  $S \subset R_\eta(T)$  and  $S' \subset T'$  such that  $\rho \in S$ ,  $\rho' \in S'$  and:*

- (i)  $\delta_H(S, T) < \varepsilon$  and  $\delta_H(S', T') < \varepsilon$ ,
- (ii) *there is a bijective measurable map  $\psi : S \rightarrow S'$  that preserves length measure and has distortion at most  $\varepsilon$ ,*
- (iii)  $\psi(\rho) = \rho'$ ,
- (iv) *the length measure of the set of points  $a \in S$  such that  $\{b \in S : \psi(a) \leq b\} \neq \psi(\{b \in S : a \leq b\})$  (that is, the set of points  $a$  such that the subtree above  $\psi(a)$  is not the image under  $\psi$  of the subtree above  $a$ ) is less than  $\varepsilon$ ,*
- (v) *there is a correspondence  $\mathcal{R} \in C(S, S')$  and a measure  $\pi \in M(S, S')$  such that:*
  - (a)  $\forall x \in S (x, \psi(x)) \in \mathcal{R}$ ,
  - (b)  $\pi(\mathcal{R}^c) \leq \varepsilon$ ,
  - (c)  $D(\pi; p_S\sharp\mu, p_{S'}\sharp\mu') \leq \varepsilon$ ,
  - (d)  $\text{dis}(\mathcal{R}) \leq 2\varepsilon$ .

PROOF. Notice that in [21], Lemma 6.3,  $\mathbf{T}$  and  $\mathbf{T}'$  are supposed to be finite trees, but we shall soon be back to this case.

Suppose that  $\eta > 0$  and  $d_{\text{GHP}}^{\text{root}}((\mathbf{T}, \rho), (\mathbf{T}', \rho')) < \delta$  ( $\delta$  and  $\eta$  will be chosen small enough later). Define, to lighten notation:

$$\mathbf{T}_\eta := R_\eta(T) \quad \text{and} \quad \mathbf{T}'_\eta := R_\eta(T').$$

Using Lemma 6.2,

$$d_{\text{GHP}}^{\text{root}}((\mathbf{T}_\eta, \rho), (\mathbf{T}'_\eta, \rho')) < \delta + 2\eta =: \tilde{\delta}.$$

Then there exists a correspondence  $\mathcal{R}_0 \in C(\mathbf{T}_\eta, \mathbf{T}'_\eta)$  and a measure  $\pi_0 \in M(\mathbf{T}_\eta, \mathbf{T}'_\eta)$  such that:

- (a)  $(\rho, \rho') \in \mathcal{R}_0$ ,
- (b)  $\pi_0(\mathcal{R}_0^c) \leq \tilde{\delta}$ ,



- (c)  $D(\pi_0; p_{T_\eta}\mu, p_{T'_\eta}\mu') \leq \tilde{\delta}$ ,
- (d)  $\text{dis}(\mathcal{R}) \leq 2\tilde{\delta}$ .

Now, we perform the proof of [21], Lemma 6.3, and we shall use their notation. We introduce a function  $f$  from  $T_\eta$  to  $T'_\eta$ . First, let  $f(\rho) := \rho'$  and then for each  $x \in T_\eta$ , one chooses  $f(x) \in T'_\eta$  such that  $(x, f(x)) \in \mathcal{R}_0$  (notice that this can be done in a measurable way). Then, letting  $x_1, \dots, x_n$  to be the leaves of  $T_\eta$  one defines  $x'_i = f(x_i)$ , and let  $T''$  be the subtree of  $T'_\eta$  spanned by  $\rho', x'_1, \dots, x'_n$ . Finally,  $\bar{f}(x)$  is defined to be the closest point to  $f(x)$  on  $T''$ . Notice that  $x'_i = \bar{f}(x_i)$ . [21], Lemma 6.3, shows that  $T''$  has leaves  $x'_1, \dots, x'_n$  (and root  $\rho' = \bar{f}(\rho)$ ), that  $\delta_H(T_\eta, T'') < 3\tilde{\delta}$  and that the function  $\bar{f}$  from  $T_\eta$  to  $T''$  has distortion at most  $8\tilde{\delta}$ . It is easy to see that

$$(6.4) \quad \forall x \in T, \quad d'(\bar{f}(x), f(x)) \leq 4\tilde{\delta}.$$

In the proof of [21], Lemma 6.3, they then take  $y_1 \in [\rho, x_1]$  and  $y'_1 \in [\rho', x'_1]$  such that  $d(\rho, y_1) = d'(\rho', y'_1) = d(\rho, x_1) \wedge d'(\rho', x'_1)$  and define  $\psi$  from  $S_1 := [\rho, y_1]$  to  $S'_1 := [\rho', y'_1]$  in the obvious way. The proof then proceeds inductively, defining  $z_{k+1}$  (resp.,  $z'_{k+1}$ ) as the closest point to  $x_{k+1}$  on  $S_k$  (resp., to  $x'_{k+1}$  on  $S'_k$ ), letting  $y_{k+1} \in ]z_{k+1}, x_{k+1}]$  and  $y'_{k+1} \in ]z'_{k+1}, x'_{k+1}]$  be such that

$$d(z_{k+1}, y_{k+1}) = d'(z'_{k+1}, y'_{k+1}) = d(z_{k+1}, x_{k+1}) \wedge d'(z'_{k+1}, x'_{k+1})$$

defining  $\psi$  from  $]z_{k+1}, y_{k+1}]$  to  $]z'_{k+1}, y'_{k+1}]$  in the obvious way and gluing  $]z_{k+1}, y_{k+1}]$  to  $S_k$  to get  $S_{k+1}$  (resp.,  $]z'_{k+1}, y'_{k+1}]$  to  $S'_k$  to get  $S'_{k+1}$ ). Finally, let  $S := S_n$  and  $S' := S'_n$ . They prove then that

$$\text{dis}(\psi) < 280\tilde{\delta}, \quad \delta_H(S, T_\eta) < 56\tilde{\delta} \quad \text{and} \quad \delta_H(S', T'_\eta) < 58\tilde{\delta}.$$

They also prove in [21], page 113, that the length measure of the set of points mentioned in (iv) is at most  $224\tilde{\delta}n$  where  $n$  is the number of leaves of  $R_\eta(T)$ . Notice that  $n$  depends only on  $T$  and  $\eta$ , let us call it  $n(\eta, T)$ . This shows that  $\psi, S$  and  $S'$  satisfy (i) – (iv) above if  $\delta$  and  $\eta$  are chosen small enough: first fix a positive  $\eta \leq \varepsilon/2000$ , and then, choose  $\delta < \frac{\varepsilon}{2000n(\eta, T)}$ .

Also, it is shown in [21], inequality (6.28), that for any  $k$ ,  $d(x_k, y_k) \vee d'(x'_k, y'_k) \leq 12\tilde{\delta}$ .

Now, let us show that:

$$(6.5) \quad \forall x \in S, \quad d'(\bar{f}(x), \psi(x)) \leq 56\tilde{\delta}.$$

Let  $x \in ]z_k, y_k]$ , then  $d'(\psi(x), y'_k) = d(x, y_k)$  (recall that  $\psi(y_k) = y'_k$ ). Then  $|d(x, y_k) - d(x, x_k)| \leq 12\tilde{\delta}$  and  $|d'(\psi(x), y'_k) - d'(\psi(x), x'_k)| \leq 12\tilde{\delta}$ . Since  $\bar{f}$  has distortion at most  $8\tilde{\delta}$ ,  $|d(x, x_k) - d'(\bar{f}(x), x'_k)| \leq 8\tilde{\delta}$ . We get

$$|d'(\psi(x), x'_k) - d'(\bar{f}(x), x'_k)| \leq 32\tilde{\delta}.$$

Let  $z$  be the closest point to  $\bar{f}(x)$  on  $[\rho', x'_k]$ . Then

$$\begin{aligned} d'(\bar{f}(x), z) &= \frac{1}{2}[d'(\bar{f}(x), \rho') + d'(\bar{f}(x), x'_k) - d'(\rho', x'_k)] \\ &\leq \frac{3}{2} \text{dis}(\bar{f}) + \frac{1}{2}[d(x, \rho) + d(x, x_k) - d(\rho, x_k)] \\ &\leq 12\tilde{\delta}, \end{aligned}$$

since  $x \in [\rho, x_k]$ . Finally, since  $\psi(x) \in [\rho', x'_k]$ ,

$$\begin{aligned} d'(\bar{f}(x), \psi(x)) &= d'(\bar{f}(x), z) + d'(z, \psi(x)) \\ &= d'(\bar{f}(x), z) + |d'(x'_k, \psi(x)) - d'(x'_k, z)| \\ &\leq 2d'(\bar{f}(x), z) + |d'(x'_k, \psi(x)) - d'(x'_k, \bar{f}(x))| \\ &\leq 24\tilde{\delta} + 32\tilde{\delta}. \end{aligned}$$

This shows (6.5).

Now, let  $\mathcal{R}$  be defined by

$$\mathcal{R} := \left\{ (x, x') \in S \times S' : \exists (y, y') \in \mathcal{R}_0, \begin{pmatrix} d(x, y) \leq 100\tilde{\delta} \\ \text{and} \\ d'(x', y') \leq 100\tilde{\delta} \end{pmatrix} \right\},$$

and define  $\pi := (p_S \otimes p_{S'}) \sharp \pi_0$ . It remains to prove point (v). First, recall that  $(x, f(x)) \in \mathcal{R}_0$  for any  $x \in T_\eta$ . Thus (v)(a) is satisfied thanks to (6.5) and (6.4). This shows also that  $\mathcal{R}$  is a correspondence on  $S \times S'$ .

Then

$$\text{dis}(\mathcal{R}) \leq \text{dis}(\mathcal{R}_0) + 400\tilde{\delta}$$

which is less than  $\varepsilon$  and shows (v)(d) if  $\tilde{\delta}$  is chosen small enough. Since  $\tilde{\delta}_H(S, T_\eta) \vee \tilde{\delta}_H(S', T'_\eta) < 58\tilde{\delta}$ , we see that for any  $x \in T_\eta$  and  $x' \in T'_\eta$ ,

$$d(x, p_S(x)) < 58\tilde{\delta} \quad \text{and} \quad d(x', p_{S'}(x')) < 58\tilde{\delta}.$$

Thus, if  $(x, x') \in \mathcal{R}_0$ , then  $(p_S(x), p_{S'}(x')) \in \mathcal{R}$  and this gives

$$\pi(\mathcal{R}^c) \leq \pi_0(\mathcal{R}_0^c) \leq \tilde{\delta},$$

which shows (v)(b) if  $\tilde{\delta}$  is chosen small enough. Finally, since  $\pi = (p_S \otimes p_{S'}) \sharp \pi_0$  one sees that

$$D(\pi; p_S \sharp \mu, p_{S'} \sharp \mu') \leq D(\pi_0, \mu, \mu') < \tilde{\delta}.$$

This completes the proof.  $\square$

Now, let us prove the Feller property for trees.

**PROPOSITION 6.4.** *Let  $(X^n)_{n \geq 0}$  be a sequence in  $\mathcal{N}_2^{\text{tree}}$  converging to  $X$  (in the  $L_{2,\text{GHP}}$  metric). Then:*

- (i)  $(\text{Frag}(X, t))_{t \geq 0}$  is strong Markov with càdlàg trajectories (for  $L_{2,\text{GHP}}$ ) in  $\mathcal{N}_2^{\text{tree}}$ ,
- (ii)  $(\text{Frag}(X^n, t))_{t \geq 0}$  converges in distribution to  $(\text{Frag}(X, t))_{t \geq 0}$  (for the Skorokhod topology associated to  $L_{2,\text{GHP}}$ ),
- (iii) if  $t^n \xrightarrow[n \rightarrow \infty]{} t$ , then  $\text{Frag}(X^n, t^n)$  converges in distribution to  $\text{Frag}(X, t)$  (for  $L_{2,\text{GHP}}$ ).

**PROOF.** First, we argue that one may without loss of generality suppose that  $X^n$  and  $X$  contain a single component. Indeed, fix  $\varepsilon > 0$ . Since  $\text{masses}(X^n)$  converges to  $\text{masses}(X)$  in  $\ell^2$ , one may choose  $\varepsilon' \notin \text{masses}(X)$  such that

$$\|\text{masses}(X_{\leq \varepsilon'})\|_2^2 \vee \sup_{n \in \mathbb{N}} \|\text{masses}(X^n_{\leq \varepsilon'})\|_2^2 \leq \varepsilon.$$

Then, since  $X^n_{> \varepsilon'}$  converges to  $X_{> \varepsilon'}$  as  $n$  goes to infinity, they have the same number of components for  $n$  large enough. Call this number  $K$ . One may list them as follows: let  $T_i^n$  (resp.,

$T_i$ ),  $i = 1, \dots, K$  be the components of  $X^n$  (resp., of  $X$ ) such that for any  $i$ ,  $T_i^n$  converges to  $T^n$ . Fix  $t > 0$ . Then, for any coupling between  $(\text{Frag}(X, s))_{s \in [0, t]}$  and  $(\text{Frag}(X^n, s))_{s \in [0, t]}$ , one has

$$\begin{aligned} & \| \text{masses}(\text{Frag}(X, s)) - \text{masses}(\text{Frag}(X^n, s')) \|_2^2 \\ & \leq \sum_{i=1}^K \| \text{masses}(\text{Frag}(T_i, s)) - \text{masses}(\text{Frag}(T_i^n, s')) \|_2^2 \\ & \quad + \| \text{masses}(X_{\leq \varepsilon'}) \|_2^2 + \| \text{masses}(X^n_{\leq \varepsilon'}) \|_2^2 \end{aligned}$$

and

$$L_{\text{GHP}}(\text{Frag}(X, s), \text{Frag}(X^n, s')) \leq \sum_{i=1}^K L_{\text{GHP}}(\text{Frag}(T_i, s), \text{Frag}(T_i^n, s')) + 16\varepsilon.$$

This shows that to prove (i) and (iii); one may suppose that  $X^n$  has a single component. Also, to prove (ii), it is sufficient to prove that for any fixed  $i$  and  $n$ , one may find a coupling such that

$$\sup_{s \in [0, t]} L_{\text{GHP}}(\text{Frag}(T_i^n, s), \text{Frag}(T_i, s)) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0.$$

In the sequel, we suppose that  $X^n =: T^n$  (resp.  $X =: T$ ) contains a single component.

Let us first prove (i). The strong Markov property was already noticed (see Remark 2.24) so let us prove that the trajectories are almost surely càdlàg. Then, for any  $\eta > 0$ ,  $\text{Frag}(R_\eta(X), \cdot)$  clearly has càdlàg trajectories: it is right continuous and piecewise constant with a finite number of jumps, since  $R_\eta(X)$  has finite length. Then, Lemma 6.1 and Lemma A.1 show that  $\text{Frag}(X, \cdot)$  also has càdlàg trajectories.

Now, let us prove (ii). Let us fix  $\varepsilon > 0$ . We know (cf. [4], Proposition 2.1) that we can take  $\rho$  (resp.  $\rho^n$ ) a root in  $T$  (resp.,  $T^n$ ) such that  $d_{\text{GHP}}^{\text{root}}((T, \rho), (T^n, \rho^n))$  goes to zero as  $n$  goes to infinity: it can be done by taking the roots sampled from the respective measures (normalized to be probability measures). For  $n$  large enough,  $d_{\text{GHP}}^{\text{root}}((T, \rho), (T^n, \rho^n))$  is small enough so that one may apply Lemma 6.3.

Let us call  $(T', \rho') = (T^n, \rho^n)$  for such a large  $n$ , in order to lighten the notation. Notice that one may suppose that  $\mu'(T') \leq \mu(T) + \varepsilon$ . Let  $\delta, \eta, S, S', \psi, \mathcal{R}$  and  $\pi$  be as in Lemma 6.3. Define

$$S := (S, d|_{S \times S}, p_S \# \mu)$$

and

$$S' := (S', d|_{S' \times S'}, p_{S'} \# \mu').$$

Let  $t > 0$ . Lemma 6.1 ensures that with probability at least  $1 - 2t\varepsilon^{1/7}$ , for any  $s \in [0, t]$ ,

$$L_{2, \text{GHP}}(\text{Frag}(T, s), \text{Frag}(S, s)) \leq 7\varepsilon^{1/7}(1 + \mu(T))^4,$$

and

$$(6.6) \quad L_{2, \text{GHP}}(\text{Frag}(T', s), \text{Frag}(S', s)) \leq 7\varepsilon^{1/7}(1 + \mu'(T'))^4 \leq 7\varepsilon^{1/7}(1 + \mu(T) + \varepsilon)^4.$$

For any  $z \in S$  (resp.,  $z' \in S'$ ) we let  $S_z$  (resp.,  $S'_{z'}$ ) be the subtree above  $z$  (resp., above  $z'$ ):

$$S_z := \{x \in S : z \in [\rho, x]\}.$$

Let us define

$$\text{Bad} := \{a \in S : S_{\psi(a)} \neq \psi(S_a)\}$$

so that Lemma 6.3 ensures that  $\ell_S(\text{Bad}) \leq \varepsilon$ .

Now, let  $\mathcal{P}$  be a Poisson random set of intensity  $\ell_S \otimes \text{leb}_{\mathbb{R}^+}$  on  $S \times \mathbb{R}^+$ . Then, for any  $s$ ,  $\psi(\mathcal{P}_s)$  is a Poisson random set of intensity  $s\ell_S$  on  $S$  (since  $\psi$  is a measure-preserving bijection), and we want to show that for any  $s \leq t$ , the fragmentation of  $S$  along  $\mathcal{P}_s$ ,  $\text{Frag}(S, \mathcal{P}_s)$ , and that of  $S'$  along  $\psi(\mathcal{P}_s)$ ,  $\text{Frag}(S', \psi(\mathcal{P}_s))$ , are close in  $L_{\text{GHP}}$ -distance with large probability.

Notice first that  $\text{Frag}(S, \mathcal{P}_s)$  and  $\text{Frag}(S', \psi(\mathcal{P}_s))$  have the same number of components on  $\mathcal{P}_t \cap \text{Bad} = \emptyset$ . If  $m$  is a component of  $\text{Frag}(S, \mathcal{P}_s)$ , it can be written as  $S_{z_s} \setminus \bigcup_{i=1}^k S_{z_{i,s}}$  for some points  $z_s, z_{1,s}, \dots, z_{k,s}$  in  $\mathcal{P}_s \cup \{\rho\}$  (we identify  $S_z \setminus \{z\}$  with  $S_z$  since it is at zero  $d_{\text{GHP}}$ -distance). If  $\mathcal{P}_t \cap \text{Bad} = \emptyset$ , then for any  $s \leq t$ ,  $\psi(m) = \psi(S_{z_s}) \setminus \bigcup_{i=1}^k \psi(S_{z_{i,s}})$  and this is a component of  $\text{Frag}(S', \psi(\mathcal{P}_s))$ .

Thus, let us place ourselves on the event  $\mathcal{E}_1 := \{\mathcal{P}_t \cap \text{Bad} = \emptyset\}$  and define  $\sigma$  (which depends on  $s$ ) to be the bijection from  $\text{Frag}(S, \mathcal{P}_s)$  to  $\text{Frag}(S', \psi(\mathcal{P}_s))$  which maps a component  $m$  to  $\psi(m)$ . Since  $\mathcal{R}$  contains the pairs  $(x, \psi(x))$  for  $x \in S$ ,  $\mathcal{R}|_{m \times \psi(m)}$  is a correspondence between  $m$  and  $\psi(m)$  with distortion at most  $\varepsilon$ . Furthermore,  $\pi|_{m \times \psi(m)}$  is a measure on  $m \times \psi(m)$  which satisfies

$$\pi|_{m \times \psi(m)}(\mathcal{R}|_{m \times \psi(m)}^c) \leq \pi(\mathcal{R}^c) \leq \varepsilon.$$

It remains to bound  $D(\pi|_{m \times \psi(m)}; (p_S \sharp \mu)|_m, (p_{S'} \sharp \mu')|_{\psi(m)})$  from above. For any Borel subset  $A$  of  $m$ ,

$$\begin{aligned} & |\pi|_{m \times \psi(m)}(A \times \psi(m)) - p_S \sharp \mu(A)| \\ & \leq |\pi(A \times S') - p_S \sharp \mu(A)| + \pi(A \times S') - \pi(A \times \psi(m)) \\ & \leq |\pi(A \times S') - p_S \sharp \mu(A)| + \pi(\{(x, x') \in S \times S' : x \in m, x' \notin \psi(m)\}). \end{aligned}$$

A symmetric inequality holds for  $A'$  a Borel subset of  $\psi(m)$ , and we get

$$D(\pi|_{m \times \psi(m)}; \mu|_m, \mu'|_{\psi(m)}) \leq D(\pi; p_S \sharp \mu, p_{S'} \sharp \mu') + \pi(m \times \psi(m)^c) + \pi(m^c \times \psi(m)).$$

Now, notice that for any  $x \in S$ ,

$$x \in m \quad \text{and} \quad x' \notin \psi(m) \quad \Rightarrow \quad [\psi(x), x'] \cap \psi(\mathcal{P}_t) \neq \emptyset$$

and

$$x \notin m \quad \text{and} \quad x' \in \psi(m) \quad \Rightarrow \quad [\psi(x), x'] \cap \psi(\mathcal{P}_t) \neq \emptyset,$$

where  $[\psi(x), x']$  is the geodesic between  $\psi(x)$  and  $x'$ . Thus,

$$\pi(m \times \psi(m)^c) + \pi(m^c \times \psi(m)) \leq \pi\{(x, x') \in S \times S' : [\psi(x), x'] \cap \psi(\mathcal{P}_t) \neq \emptyset\}.$$

Let us denote by  $\mathcal{E}_2$  the event

$$\mathcal{E}_2 := \{\pi\{(x, x') \in S \times S' : [\psi(x), x'] \cap \psi(\mathcal{P}_t) \neq \emptyset\} \leq \sqrt{\varepsilon}\}.$$

On  $\mathcal{E}_1 \cap \mathcal{E}_2$ , we get, for any  $s \leq t$  and any component  $m$  of  $\text{Frag}(S, \mathcal{P}_s)$ :

$$D(\pi|_{m \times \psi(m)}; p_S \sharp \mu|_m, p_{S'} \sharp \mu'|_{\psi(m)}) \leq \varepsilon + \sqrt{\varepsilon}.$$

Furthermore,

$$\begin{aligned} & \|\text{masses}(\text{Frag}(S, \mathcal{P}_s)) - \text{masses}(\text{Frag}(S', \psi(\mathcal{P}_s)))\|_2^2 \\ & \leq \sum_{m \in \text{comp}(\text{Frag}(S, \mathcal{P}_s))} (p_S \sharp \mu(m) - p_{S'} \sharp \mu'(\psi(m)))^2 \\ & \leq \sup_{m \in \text{comp}(\text{Frag}(S, \mathcal{P}_s))} |p_S \sharp \mu(m) - p_{S'} \sharp \mu'(\psi(m))| \end{aligned}$$

$$\begin{aligned} & \times \sum_{m \in \text{comp}(\text{Frag}(S, \mathcal{P}_s))} p_S \sharp \mu(m) + p_{S'} \sharp \mu'(\psi(m)) \\ & \leq \sup_{m \in \text{comp}(\text{Frag}(S, \mathcal{P}_s))} D(\pi|_{m \times \psi(m)}; p_S \sharp \mu|_m, p_{S'} \sharp \mu'|_{\psi(m)})(\mu(T) + \mu'(T')) \\ & \leq (\varepsilon + \sqrt{\varepsilon})(2\mu(T) + \varepsilon). \end{aligned}$$

Thus, on  $\mathcal{E}_1 \cap \mathcal{E}_2$ , we obtain, for any  $s \leq t$ :

$$L_{2,\text{GHP}}(\text{Frag}(S, \mathcal{P}_s), \text{Frag}(S', \psi(\mathcal{P}_s))) \leq (\varepsilon + \sqrt{\varepsilon}) \vee \sqrt{(\varepsilon + \sqrt{\varepsilon})(2\mu(T) + \varepsilon)}.$$

It remains to bound from above the probability of  $(\mathcal{E}_1 \cap \mathcal{E}_2)^c$ . Since *Bad* has length measure at most  $\varepsilon$ ,

$$\mathbb{P}(\mathcal{E}_1^c) \leq t\varepsilon.$$

Notice that since  $\mathcal{R}$  contains  $(x, \psi(x))$  for any  $x \in S$  and has distortion less than  $2\varepsilon$ ,

$$\pi \{(x, x') \in S \times S' : d'(\psi(x), x') > 2\varepsilon\} \leq \pi(\mathcal{R}^c) < \varepsilon.$$

Then, using Fubini's theorem,

$$\begin{aligned} & \mathbb{E}[\pi \{(x, x') \in S \times S' : [\psi(x), x'] \cap \psi(\mathcal{P}_t) \neq \emptyset\}] \\ & \leq \varepsilon + \mathbb{E}[\pi \{(x, x') \in S \times S' : [\psi(x), x'] \cap \psi(\mathcal{P}_t) \neq \emptyset \text{ and } d'(\psi(x), x') \leq 2\varepsilon\}] \\ & = \varepsilon + \int_{S \times S'} \mathbb{P}([\psi(x), x'] \cap \psi(\mathcal{P}_t) \neq \emptyset) \mathbf{1}_{d'(\psi(x), x') \leq 2\varepsilon} d\pi(x, x') \\ & \leq 2t\varepsilon\pi(S \times S') \\ & \leq 2t\varepsilon(\mu(T) + \varepsilon). \end{aligned}$$

Thus, by Markov's inequality,

$$\mathbb{P}(\mathcal{E}_2^c) \leq 2t\sqrt{\varepsilon}(\mu(T) + \varepsilon),$$

which ends the proof of point (ii) of the theorem (through Lemma A.2 and inequality (A.1)).

Finally, let us prove (iii). Suppose that  $t^n$  converges to  $t$  as  $n$  goes to infinity and let  $\tilde{t} := \sup_n t^n + 1$ . Again, it is sufficient to suppose that  $X^n$  and  $X$  have only one component, so let us suppose that  $(T, \rho)$  and  $(T^n, \rho^n)$ ,  $n \geq 0$  are rooted trees such that  $d_{\text{GHP}}^{\text{root}}((T, \rho), (T^n, \rho^n))$  goes to zero as  $n$  goes to infinity. Now recall inequality (6.6) above: for any  $\varepsilon > 0$ , for  $n$  large enough, we found finite subtrees  $S^n \subset T^n$  such that, with probability at least  $1 - 2T\varepsilon^{1/7}$ , for any  $s \in [0, \tilde{t}]$ ,

$$L_{2,\text{GHP}}(\text{Frag}(T^n, s), \text{Frag}(S^n, s)) \leq 7\varepsilon^{1/7}(1 + \mu(T) + \varepsilon)^4,$$

and furthermore,  $\ell_{T^n}(S^n) \leq \ell_T(R_\eta(T)) < \infty$ , since  $S^{(n)}$  has the same length measure as a subset of  $R_\eta(T)$ ,  $\eta$  depending only on  $\varepsilon$  (cf. Lemma 6.3). Then

$$\mathbb{P}(\text{Frag}(S^n, t^n) \neq \text{Frag}(S^n, t)) \leq 1 - e^{-|t-t^n|\ell_T(R_\eta(T))}.$$

For  $n$  large enough, this is less than  $\varepsilon$ . Then we have that with probability at least  $1 - 2\tilde{t}\varepsilon^{1/7} - \varepsilon$ ,

$$L_{2,\text{GHP}}(\text{Frag}(T^n, t^n), \text{Frag}(T^n, t)) \leq 7\varepsilon^{1/7}(1 + \mu(T) + \varepsilon)^4.$$

All in all, we proved that  $L_{2,\text{GHP}}(\text{Frag}(T^n, t^n), \text{Frag}(T^n, t))$  converges in probability to zero when  $n$  goes to infinity. But using point (i), we know that  $\text{Frag}(T^n, t)$  converges to  $\text{Frag}(T, t)$  as  $n$  goes to infinity. This completes the proof of (iii).  $\square$

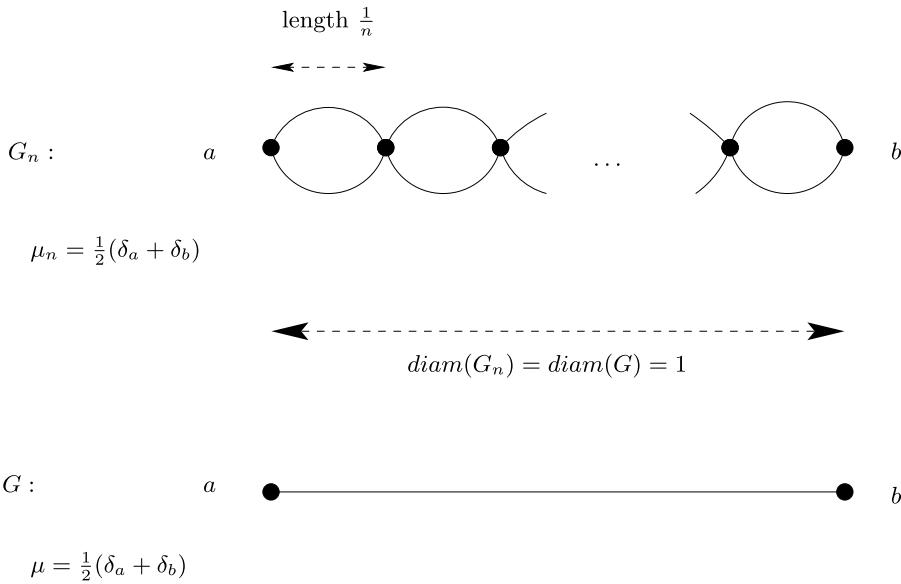


FIG. 2.  $G_n$  is composed of  $n$  graphs in series each one made of 2 intervals of length  $1/n$  in parallel.  $G_n$  converges to  $G$  for  $d_{GHP}$  when  $n$  goes to infinity, but  $Frag(G_n, t)$  will not converge to  $Frag(G, t)$  for  $t > 0$ .

6.4. *The Feller property for graphs, Theorem 3.2.* We now want to prove Theorem 3.2, which is the analog of Proposition 6.4 for graphs. However, this cannot be true without strengthening the metric  $L_{GHP}$ . For instance, consider the situation depicted in Figure 2. There  $G_n$  converges to  $G$  for  $d_{GHP}$ , but the probability that  $a$  is separated from  $b$  in  $G_n$  when fragmentation occurs (until a fixed time  $t > 0$ ) is exactly  $1 - (1 - (1 - e^{-t/n})^2)^n$ , which is asymptotically 0, whereas the probability that this event occurs in  $G$  is strictly positive. However, if we impose that the surplus of  $G_n$  converges to the surplus of  $G$ , such a situation cannot happen anymore, and one may recover the Feller property.

Let us notice that this problem was treated a bit differently in [4]: they recover continuity (in probability) of fragmentation by imposing that  $G_n$  and  $G$  live on some common subspace  $\mathcal{A}_r$  for some  $r > 0$ , where  $\mathcal{A}_r$  contains the graphs which have surplus and total length of the core bounded from above by  $1/r$  and minimal edge length of the core bounded from below by  $r$  (see Section 6.4 in [4] for a precise statement). When one wants to have Feller-type properties, this seems to us less natural than imposing convergence of the surplus. In fact, the work below shows that if  $G_n$  converges to  $G$  in the Gromov–Hausdorff topology while having the same surplus for  $n$  large enough, then there is some  $r > 0$  such that for  $n$  large enough,  $G_n$  and  $G$  belong to  $\mathcal{A}_r$ . The converse statement is also true and is a consequence of [4], Proposition 6.5.

To prove Theorem 3.2, we first notice that the proof of Section 6.3 extends to the case where one replace trees by graphs having the same core.

LEMMA 6.5. *Let  $\mathbf{G} = (G, d, \mu)$  be a measured  $\mathbb{R}$ -graph which is not a tree. Let  $(\mathbf{G}^n)_{n \geq 0}$  be a sequence of measured  $\mathbb{R}$ -graphs such that for each  $n$ , there is a correspondence  $\mathcal{R}^n \in C(G, G^n)$ , a measure  $\pi^n \in M(G, G^n)$  and a homeomorphism  $\psi^n : \text{core}(G) \rightarrow \text{core}(G^n)$  such that:*

- $\psi^n$  preserves the length-measure,
- $\forall x \in \text{core}(G^n) (x, \psi^n(x)) \in \mathcal{R}^n$ ,
- $\text{dis}(\mathcal{R}^n) \vee \pi^n((\mathcal{R}^n)^c) \vee D(\pi^n, \mu, \mu^n) \xrightarrow{n \rightarrow \infty} 0$ .

Then the sequence of processes  $\text{Frag}(\mathbf{G}^n, \cdot)$  converges in distribution to  $\text{Frag}(\mathbf{G}, \cdot)$  for  $L_{2,\text{GHP}}^{\text{surplus}}$ .

PROOF. It is a straightforward extension of the arguments of Section 6.3, replacing roots by cores and using  $\psi^n$  to map fragmentation on  $\text{core}(G^n)$  to fragmentation on  $\text{core}(G)$ .  $\square$

To prepare the proof of Theorem 3.2 we shall need a series of lemmas, but before let us explain the idea of the proof of the theorem. If  $\mathbf{G}^n$  is close enough to  $\mathbf{G}$ , Lemma 6.7 below shows that their cores are homomorphic multigraphs with edges having almost the same length. One may then shorten some edges of the core of  $G$  and other edges of the core of  $G^n$  in such a way that the two cores become homeomorphic as metric spaces with a length measure. Lemma 6.11 shows that one does not lose too much doing this. Finally, Lemma 6.5 then shows that the fragmentations on the two graphs are close to each other.

LEMMA 6.6. Let  $(G, d)$  and  $(G', d')$  be  $\mathbb{R}$ -graphs and  $\mathcal{R} \in C(G, G')$ . Let  $(a, a') \in \mathcal{R}$ ,  $(b, b') \in \mathcal{R}$  and  $(c, c') \in \mathcal{R}$ . Suppose that  $a$  belongs to a geodesic between  $b$  and  $c$ . Let  $\gamma_{a', b'}$  (resp.  $\gamma_{a', c'}$ ) be a geodesic from  $a'$  to  $b'$  (resp., from  $a'$  to  $c'$ ). Then

$$\forall a'' \in \gamma_{a', b'} \cap \gamma_{a', c'}, \quad d'(a'', a') \leq 3 \text{dis}(\mathcal{R}).$$

PROOF. Let  $a'' \in \gamma_{a', b'} \cap \gamma_{a', c'}$ . Then

$$\begin{aligned} d'(a', a'') &= d'(a', b') + d'(a', c') - d'(a'', b') - d'(a'', c') \\ &\leq d'(a', b') + d'(a', c') - d'(b', c') \\ &\leq d(a, b) + d(a, c) - d(b, c) + 3 \text{dis}(\mathcal{R}) \\ &= 3 \text{dis}(\mathcal{R}), \end{aligned}$$

where we used the triangle inequality in the second step and the fact that  $a$  belongs to a geodesic between  $b$  and  $c$  in the last step.  $\square$

The following should be compared to [4], Proposition 5.6.

LEMMA 6.7. Let  $G$  be an  $\mathbb{R}$ -graph and  $\varepsilon > 0$ . There exists  $\delta$  depending on  $\varepsilon$  and  $G$  such that if  $G'$  is an  $\mathbb{R}$ -graph with the same surplus as  $G$  and if  $\mathcal{R}_0 \in C(G, G')$  is such that  $\text{dis}(\mathcal{R}_0) < \delta$ , then there exists an  $\varepsilon$ -overlay  $\mathcal{R} \in C(G, G')$  containing  $\mathcal{R}_0$ .

PROOF. If  $G$  has surplus 0, there is nothing to prove. In the sequel, we suppose that  $G$  has surplus at least 2, the easier proof for unicyclic  $G$  is left to the reader. Furthermore, to lighten notation and make the argument clearer, we shall suppose that the vertices of  $\ker(G)$  are of degree 3, leaving the adaptation to the general case to the reader.

Let  $\eta := \min_{e \in e(G)} \ell(e)$ . Notice that one may view  $\text{core}(G)$  (and  $\text{core}(G')$ ) as a multigraph with edge-lengths, and we shall adopt this point of view in this proof. However, not all the edges of this graph correspond to geodesics in  $G$ . Divide each edge of  $\text{core}(G)$  into five pieces of equal length, introducing thus four new vertices of degree 2 for each edge (all degrees will be relative to the core). The new graph obtained satisfies the following:

- (i) all the edges remain of length larger than  $\eta/5$ ,
- (ii) every edge  $e$  is the unique geodesic between its two endpoints, and for any path  $\gamma$  between these endpoints which does not contain  $e$ ,  $\ell(\gamma) - \ell(e) > \eta/5$ ,
- (iii) for every three vertices  $a, b, c$  such that  $b \sim a$  and  $a \sim c$ ,  $a$  belongs to a geodesic between  $a$  and  $c$ .



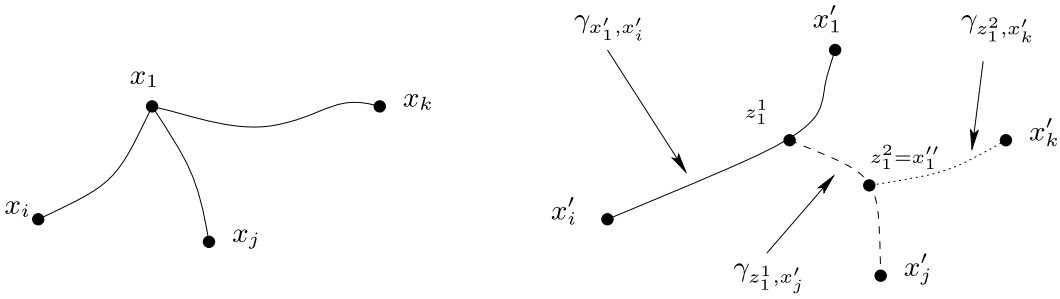


FIG. 3. One maps  $\text{core}(G)$  to  $\text{core}(G')$  by first mapping the neighborhood of each vertex of  $\text{core}(G)$  to a subset of  $G'$ . Here,  $x''_1$  is a vertex of  $\ker(G')$ .

Let us call  $\text{c\^o}re(G)$  this new graph (it is indeed a graph, not merely a multigraph), which has the same surplus as  $G$ , and write  $x_1, \dots, x_n$  for its vertices, which are of degree 2 or 3.

Let  $G'$  be an  $\mathbb{R}$ -graph with the same surplus as  $G$  and  $\mathcal{R}_0 \in C(G, G')$ . Let  $x'_1, \dots, x'_n$  be elements of  $G'$  such that  $(x_i, x'_i) \in \mathcal{R}_0$ . Now, we shall build a subgraph of  $G'$  by mapping recursively edges adjacent to a given vertex in  $\text{c\^o}re(G)$  to a geodesic in  $G'$ . Suppose for instance that  $x_1$  has degree 3 (the argument is analogous for vertices with degree 2). Let  $x_i, x_j$  and  $x_k$  be its neighbors in  $\text{c\^o}re(G)$ , with  $i < j < k$ . Choose a geodesic  $\gamma_{x'_1, x'_i}$  between  $x'_1$  and  $x'_i$ , then choose a geodesic  $\gamma$  between  $x'_1$  and  $x'_j$ , and let  $z^1_1$  be the point of  $\gamma_{x'_1, x'_i} \cap \gamma$  which is the furthest from  $x'_1$  (see Figure 3). Let us call  $\gamma_{z^1_1, x'_j}$  the subpath of  $\gamma$  from  $z^1_1$  to  $x'_j$ . Notice that the path using  $\gamma_{x'_1, x'_i}$  from  $x'_1$  to  $z^1_1$  and  $\gamma$  from  $z^1_1$  to  $x'_j$  is a geodesic. Finally, choose a geodesic  $\gamma$  between  $x'_1$  and  $x'_k$  and let  $z^2_1$  be the point of  $(\gamma_{x'_1, x'_i} \cup \gamma_{x'_1, x'_j}) \cap \gamma$  which is the furthest from  $x'_1$ . Let us call  $\gamma_{z^2_1, x'_k}$  the subpath of  $\gamma$  from  $z^2_1$  to  $x'_k$ . Let  $S'_1 := \gamma_{x'_1, x'_i} \cup \gamma_{z^1_1, x'_j} \cup \gamma_{z^2_1, x'_k}$ . Define  $x''_1$  to be the point between  $z^1_1$  and  $z^2_1$  which is the furthest from  $x'_1$ . If  $x_1$  is of degree 2, there is only one point  $z^1_1$  defined and  $x''_1$  is this one.

Then we proceed similarly for  $r = 2, \dots, n$ : we inspect the neighbors of  $x_r$ . Notice that we do not need to choose a new geodesic between  $x'_r$  and a neighbor  $x'_j$  for  $j < r$ , we just keep the one already built. Doing this, we obtain  $S'_r$  the union of the geodesics chosen going from  $x'_r$  to the points associated to the neighbors of  $x_r$ , we get two points  $z^1_r$  and  $z^2_r$  if  $x_r$  is of degree 3 and only one point  $z^1_r$  if  $x_r$  is of degree 1. We define  $x''_r$  to be the one between  $z^1_r$  and  $z^2_r$  which is the furthest from  $x'_r$ .

Finally, let  $S' = \bigcup_{i=1}^n S'_i$ , with all the vertices  $z^b_i$  and  $x'_i$  for  $1 \leq i \leq n$ . This is a graph with edge-lengths (notice that the edges have pairwise disjoint interiors). Some edge-lengths might be zero. Thanks to point (iii) above and Lemma 6.6, we know that:

$$d'(z^1_i, x'_i) \leq 3 \text{dis}(\mathcal{R}_0),$$

and when  $x'_i$  is of degree 3,

$$d'(z^2_i, x'_i) \leq 3 \text{dis}(\mathcal{R}_0).$$

Thus, for any  $b, b' \in \{1, 2\}$  and any  $i \neq j$ ,

$$\begin{aligned} d'(z^b_i, z^{b'}_j) &\geq d'(x'_i, x'_j) - 6 \text{dis}(\mathcal{R}_0) \\ &\geq \frac{\eta}{5} - 7 \text{dis}(\mathcal{R}_0). \end{aligned}$$

Thus, if  $\text{dis}(\mathcal{R}_0) < \eta/35$ , two points  $z^b_i$  and  $z^{b'}_j$  are always distinct. This shows that  $S'$  has the same surplus as  $\text{core}(G)$ . Since  $G'$  has the same surplus as  $G$ , we deduce that  $S'$  contains

core( $G'$ ). Let  $S''$  be the subgraph of  $S'$  spanned by  $x''_1, \dots, x''_n$ , in the sense that we forget the vertices  $z_i^1$  when  $x_i$  is of degree 3, and we remove the semiopen path going from  $x'_i$  to  $z_i^1$ . Notice that  $S''$  has positive edge-lengths and its edges have pairwise disjoint interiors.  $S''$  has the same surplus as  $S'$ , so it contains again core( $G'$ ). But all the vertices in  $S''$  have degree 2 or 3, so  $S'' = \text{core}(G')$  as a set.

Now, consider  $S''$  as a graph with edge-lengths and with vertices  $x''_i, i = 1, \dots, r$ . Let  $\chi_0$  be the map from core( $G$ ) to  $S''$  which maps  $x_i$  to  $x''_i$ . We shall see that it is a graph isomorphism if  $\text{dis}(\mathcal{R}_0)$  is small enough. Indeed, from the inequalities above, we get that for any  $i$  and  $j$ ,

$$|d'(x''_i, x''_j) - d(x_i, x_j)| \leq 7 \text{dis}(\mathcal{R}_0).$$

Now, for any edge  $e = (x_i, x_j)$  and any  $k$  distinct from  $i$  and  $j$ ,

$$\begin{aligned} d'(x''_i, x''_k) + d'(x''_k, x''_j) &\geq d(x_i, x_k) + d(x_k, x_j) - 14 \text{dis}(\mathcal{R}_0) \\ &\geq d(x_i, x_j) + \frac{\eta}{5} - 14 \text{dis}(\mathcal{R}_0) \\ &\geq d'(x''_i, x''_j) + \frac{\eta}{5} - 21 \text{dis}(\mathcal{R}_0), \end{aligned}$$

where we used point (ii) above in the last inequality. Thus, if  $\text{dis}(\mathcal{R}_0) < \frac{\eta}{105}$ ,  $x''_i$  and  $x''_j$  are neighbors in  $S''$  as soon as  $x_i$  and  $x_j$  are neighbors in core( $G$ ). Furthermore, from the construction of  $S''$ , one sees that the number of edges in  $S''$  is at most the number of edges of core( $G$ ). Thus,  $\chi_0$  is a graph isomorphism and we deduce from the last inequality that for any edge  $e$ ,  $\chi_0(e)$  is the unique geodesic between its endpoints. Furthermore, let  $x''_i$  and  $x''_j$  be neighbors in  $S''$ . If  $\text{dis}(\mathcal{R}_0) < \frac{\eta}{210}$ , we see from the last inequality that every path  $\gamma$  from  $x''_i$  to  $x''_j$  which does not contain  $[x''_i, x''_j]$  satisfies

$$(6.7) \quad \ell'(\gamma) > \ell'([x''_i, x''_j]) + \frac{\eta}{10}.$$

Let us define  $\mathcal{R}'$  by adding to  $\mathcal{R}_0$  the pairs  $(x_i, x''_i)$  for  $i = 1, \dots, r$ . Then  $\text{dis}(\mathcal{R}') \leq 7 \text{dis}(\mathcal{R}_0)$ . Let  $\mathcal{R}$  be the  $3 \text{dis}(\mathcal{R}_0)$ -enlargement of  $\mathcal{R}'$ . It has distortion at most  $19 \text{dis}(\mathcal{R}_0)$ . Let  $x$  belong to an edge  $[x_i, x_j]$  of core( $G$ ) and let  $x'$  be such that  $(x, x') \in \mathcal{R}_0$ . Let  $\gamma_{x', x''_i}$  (resp.  $\gamma_{x', x''_j}$ ) be a geodesic between  $x'$  and  $x''_i$  (resp., between  $x'$  and  $x''_j$ ). Then let  $\gamma$  be the path from  $x''$  to  $x''_j$  obtained by concatenating  $\gamma_{x', x''_i}$  and  $\gamma_{x', x''_j}$ . We have

$$\begin{aligned} \ell'(\gamma) &\leq d'(x''_i, x') + d'(x', x''_j) \\ &\leq d(x_i, x) + d(x, x_j) + 2 \text{dis}(\mathcal{R}') \\ &= d(x_i, x_j) + 2 \text{dis}(\mathcal{R}') \\ &\leq \ell'([x''_i, x''_j]) + 3 \text{dis}(\mathcal{R}'). \end{aligned}$$

Thus, if  $21 \text{dis}(\mathcal{R}_0) < \frac{\eta}{10}$ ,  $3 \text{dis}(\mathcal{R}') < \frac{\eta}{10}$  and we deduce from (6.7) that  $\gamma$  contains  $[x''_i, x''_j]$ . Thus, defining  $x''$  to be the furthest point from  $x'$  on  $\gamma_{x', x''_i} \cap \gamma_{x', x''_j}$ , we see that  $x''$  belongs to the geodesic  $[x''_i, x''_j]$ . Lemma 6.6 ensures that  $d'(x', x'') \leq 3 \text{dis}(\mathcal{R}')$ . Thus,  $(x, x'') \in \mathcal{R}$ . Similarly, one shows that for every  $x''$  in  $[x''_i, x''_j]$  there is an  $x$  in  $[x_i, x_j]$  such that  $x \in \mathcal{R}$ . We have shown that for each edge  $e$  of core( $G$ ),  $e$  and  $\chi_0(e)$  are in correspondence via  $\mathcal{R}$ . Now, notice that the multigraph with edge-lengths  $S''$  obtained by keeping only vertices of degree 3 is core( $G'$ ) seen as a multigraph with edge-lengths. The isomorphism  $\chi_0$  induces an isomorphism  $\chi$  between core( $G$ ) and core( $G'$ ) (by restricting  $\chi_0$  to vertices of degree 3), and we have (since every edge of core( $G$ ) was divided into five parts):

$$|\ell(e) - \ell'(\chi(e))| \leq 30 \text{dis}(\mathcal{R}_0).$$

Furthermore, the same correspondence  $\mathcal{R}$  as before is suitable to have that for each edge  $e$  of  $\text{core}(G)$ ,  $e$  and  $\chi(e)$  are in correspondence via  $\mathcal{R}$ .

This completes the proof by taking  $\text{dis}(\mathcal{R}_0) < \delta$  for  $\delta$  small enough, namely less than  $\frac{\varepsilon}{40} \wedge \frac{\eta}{210}$ .  $\square$

LEMMA 6.8. *Let  $(G, d)$  and  $(G', d')$  be  $\mathbb{R}$ -graphs and  $\mathcal{R} \in C(G, G')$ . Suppose that  $\text{core}(G)$  and  $\text{core}(G')$  are in correspondence through  $\mathcal{R}$ . Let  $(v, v')$  and  $(x, x') \in \mathcal{R}$  with  $v \in \text{core}(G)$  and  $v' \in \text{core}(G')$ . Then*

$$d(\alpha_G(x), v) \leq d'(\alpha_{G'}(x'), v') + 5 \text{dis}(\mathcal{R}).$$

PROOF. Since  $\text{core}(G)$  and  $\text{core}(G')$  are in correspondence through  $\mathcal{R}$ , one may find  $y \in \text{core}(G)$  and  $y' \in \text{core}(G')$  such that

$$(y, \alpha_{G'}(x')) \in \mathcal{R} \quad \text{and} \quad (\alpha_G(x), y') \in \mathcal{R}.$$

Let us distinguish two cases.

- $d(y, v) \geq d(\alpha_G(x), v)$ . Then

$$d(\alpha_G(x), v) \leq d(y, v) \leq d'(\alpha_{G'}(x'), v') + \text{dis}(\mathcal{R})$$

and the result follows.

- $d(y, v) < d(\alpha_G(x), v)$ . Then

$$\begin{aligned} d(x, v) &= d(x, \alpha_G(x)) + d(\alpha_G(x), v) \\ &\geq d(x, \alpha_G(x)) + d(y, v) \\ &\geq d'(x', y') + d'(\alpha_{G'}(x'), v') - 2 \text{dis}(\mathcal{R}) \\ &= d'(x', \alpha_{G'}(x')) + d'(\alpha_{G'}(x'), y') + d'(\alpha_{G'}(x'), v') - 2 \text{dis}(\mathcal{R}) \\ &= d'(x', v') + d'(\alpha_{G'}(x'), y') - 2 \text{dis}(\mathcal{R}) \\ &\geq d(x, v) + d'(\alpha_{G'}(x'), y') - 3 \text{dis}(\mathcal{R}). \end{aligned}$$

Thus,

$$d'(\alpha_{G'}(x'), y') \leq 3 \text{dis}(\mathcal{R})$$

which implies

$$d(y, \alpha_G(x)) \leq 4 \text{dis}(\mathcal{R}).$$

Finally,

$$\begin{aligned} d(\alpha_G(x), v) &\leq d(\alpha_G(x), y) + d(y, v) \\ &\leq 4 \text{dis}(\mathcal{R}) + d(y, v) \\ &\leq 5 \text{dis}(\mathcal{R}) + d'(\alpha_{G'}(x'), v'). \end{aligned} \quad \square$$

Let us introduce some notation for the following lemmas (see Figure 4).

DEFINITION 6.9. For any graph  $G$ , for each oriented edge  $e = (u, v) \in \ker(G)$  and each  $\eta \in [0, \ell(e)]$ , we denote by  $v - \eta e$  the point at distance  $\eta$  from  $v$  on the edge  $(u, v)$ , on  $\text{core}(G)$ . For  $a < b$  in  $[0, \ell(e)]$ , let  $]v - be, v - ae[$  be the open oriented arc between  $v - be$  and  $v - ae$  in  $(u, v)$ .

We define  $\mathbf{G}^{(e,a,b)}$  the  $(a, b)$ -shortening along  $e$  as the measured  $\mathbb{R}$ -graph  $(H, d_H, \mu_H)$  obtained from  $G$  as follows:

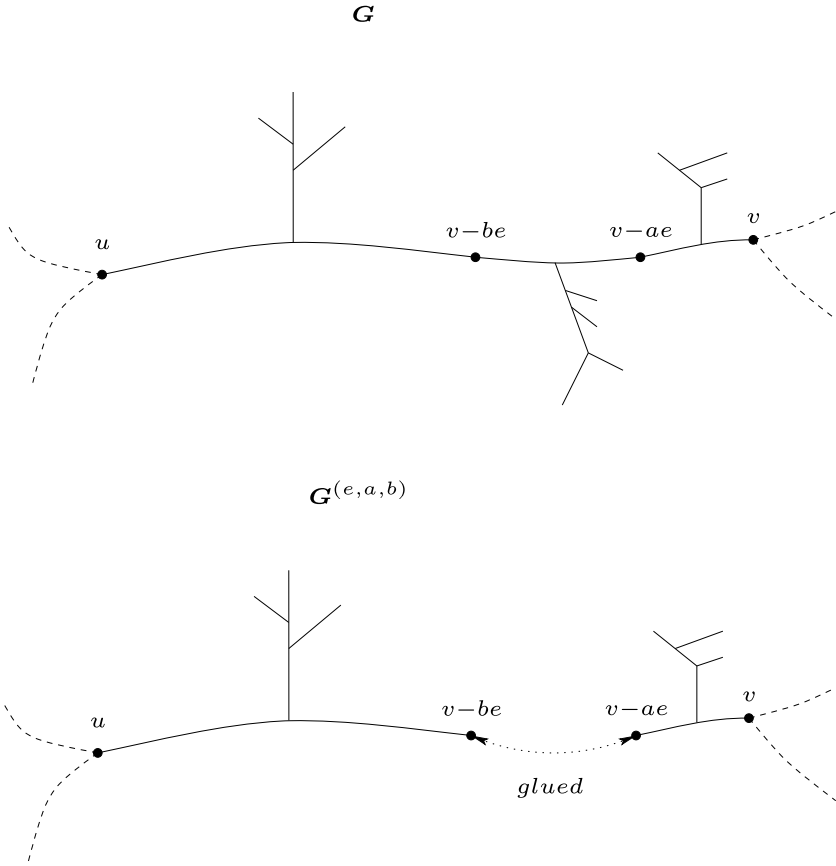


FIG. 4.  $G^{(e,a,b)}$  is the  $(a, b)$ -shortening of  $G$  along  $e = (u, v)$ .

- $H = G \setminus \alpha_G^{-1}(]v - be, v - ae[)$ ,
- $d_H$  is obtained from  $(H, d|_{H \times H})$  by gluing it along the equivalence relation generated by  $\{(v - be, v - ae)\}$  (thus  $d_H(v - be, v - ae) = 0$ ),
- $\mu_H$  is the restriction of  $\mu$  on  $H$ .

Notice that  $G^{(e,a,b)}$  has the same surplus as  $G$ .

LEMMA 6.10. *Let  $G$  be an  $\mathbb{R}$ -graph, and define:*

$$\gamma_G(\eta) := \sup_{e=(u,v) \in \ker(G)} \text{diam}(\alpha_G^{-1}(]v - \eta e, v[)).$$

Then

$$\gamma_G(\eta) \xrightarrow{\eta \rightarrow 0} 0.$$

PROOF. Suppose on the contrary that  $\gamma_G(\eta) \xrightarrow{\eta \rightarrow 0} \gamma > 0$ . Then one may find an edge  $e$  and a sequence of pairs  $(x_n, y_n)_{n \in \mathbb{N}}$  in  $\alpha_G^{-1}(]v - e, v[)$  such that

$$d(\alpha_G(x_n), v) \vee d(\alpha_G(y_n), v) \xrightarrow{n \rightarrow \infty} 0,$$

$$\forall n \in \mathbb{N}, d(x_n, y_n) \geq \gamma,$$

and

$$\forall n \in \mathbb{N}, \quad d(\alpha_G(x_n), v) \wedge d(\alpha_G(y_n), v) > 0.$$

Let  $z_n \in \{x_n, y_n\}$  be such that  $d(z_n, v) = d(x_n, v) \vee d(y_n, v)$ . Up to extracting a subsequence, one may also suppose that  $d(\alpha_G(z_n), v)$  is strictly decreasing and that for any  $n$ ,  $d(\alpha_G(z_n), v) < \gamma/4$ . This implies that for  $n \neq m$ ,

$$\begin{aligned} d(z_n, z_m) &\geq d(z_n, \alpha_G(z_n)) \\ &= d(z_n, v) - d(\alpha_G(z_n), v) \\ &\geq \frac{\gamma}{2} - d(\alpha_G(z_n), v) \\ &\geq \frac{\gamma}{4}. \end{aligned}$$

This contradicts the precompactness of  $G$ .  $\square$

LEMMA 6.11. *Let  $\mathbf{G} = (G, d, \mu)$  be a measured  $\mathbb{R}$ -graph with surplus at least one, let  $e$  be an edge of  $\text{core}(G)$  and  $a < b \in [0, \ell(e)]$ . Let*

$$\tilde{\gamma}_G(\varepsilon) := \sum_{e=(u,v) \in \ker(G)} \mu(\alpha_G^{-1}(]v - \varepsilon e, v[)).$$

*Then, under the natural coupling between  $\text{Frag}(\mathbf{G}, \cdot)$  and  $\text{Frag}(\mathbf{G}^{(e,a,b)}, \cdot)$  we have, with probability at least  $1 - t(b - a)$ , for any  $s \in [0, t]$ ,*

$$L_{2,\text{GHP}}^{\text{surplus}}(\text{Frag}(\mathbf{G}, s), \text{Frag}(\mathbf{G}^{(e,a,b)}, s)) \leq C(G) \sqrt{\tilde{\gamma}_G(b) \vee \gamma_G(b)},$$

*where  $C(G)$  is a positive and finite constant depending only on  $\text{diam}(G)$  and  $\mu(G)$ .*

PROOF. Let  $e = (u, v)$ , and  $\mathcal{P}$  be a Poisson random set of intensity  $\ell_G \times \text{leb}_{\mathbb{R}^+}$  on  $(G, d)$ . Then  $\mathcal{P}' := [\mathcal{P} \setminus \alpha_G^{-1}(]v - be, v - ae[)] \times \mathbb{R}^+$  is a Poisson random set of intensity  $\ell_{G'} \times \text{leb}_{\mathbb{R}^+}$  on  $G' \times \mathbb{R}^+$  with  $G' := G \setminus \alpha_G^{-1}(]v - be, v - ae[)$ . Let  $t > 0$  be fixed and let  $\mathcal{E}$  denote the event

$$\mathcal{E} := \{\mathcal{P}_t \cap ]v - be, v - ae[ = \emptyset\},$$

and let us suppose that  $\mathcal{E}$  holds. Let  $\varepsilon > 0$  be such that

$$(6.8) \quad \varepsilon \geq \mu(\alpha_G^{-1}(]v - be, v - ae[)).$$

Let us take  $s \leq t$  and let  $m$  be a component of  $\text{Frag}(G, \mathcal{P}_s)$ . Then:

- if  $m \subset \alpha_G^{-1}(]v - be, v - ae[)$ , then  $\mu(m) \leq \varepsilon$ ,
- if  $m \cap \alpha_G^{-1}(]v - be, v - ae[) = \emptyset$ , then  $m$  is a component of  $\text{Frag}(G', \mathcal{P}'_s)$ ,
- if  $m \cap \alpha_G^{-1}(]v - be, v - ae[) \neq \emptyset$  but  $m \not\subset \alpha_G^{-1}(]v - be, v - ae[)$ , then  $m$  is the unique component of  $\text{Frag}(G, \mathcal{P}_s)$  which intersects  $]v - be, v - ae[$ , and  $m \setminus \alpha_G^{-1}(]v - be, v - ae[)$  is a component of  $\text{Frag}(G', \mathcal{P}'_s)$ .

This shows that the function

$$\sigma : \begin{cases} \text{comp}(\text{Frag}(G, \mathcal{P}_s)_{>\varepsilon}) \rightarrow \text{comp}(\text{Frag}(G', \mathcal{P}'_s)), \\ m \mapsto m \setminus \alpha_G^{-1}(]v - be, v - ae[) \end{cases}$$

is well-defined and injective. This shows also that the function  $\sigma'$  from  $\text{comp}(\text{Frag}(G', \mathcal{P}'_s)_{>\varepsilon})$  to  $\text{comp}(\text{Frag}(G, \mathcal{P}_s)_{>\varepsilon})$  which maps  $m'$  to the unique  $m$  which contains it is well-defined and injective.

Now, let  $m \in \text{comp}(\text{Frag}(G, \mathcal{P}_s)_{>\varepsilon})$  and let

$$m' := \sigma(m) = m \setminus \alpha_G^{-1}(\lceil v - be, v - ae \rceil).$$

Let

$$\mathcal{R}_m := \{(x, x) : x \in m'\} \cup \{(x, v) : x \in m \cap \alpha_G^{-1}(\lceil v - be, v - ae \rceil)\}$$

and  $\pi_m$  be the measure in  $M(m, m')$  defined by

$$\pi_m(C) = \mu(\{x \in m' : (x, x) \in C\}).$$

Let  $d'$  be the distance on  $m'$ . Notice that for any  $x, y$  in  $m'$ ,

$$|d(x, y) - d'(x, y)| \leq b$$

Thus,

$$(6.9) \quad \text{dis}(\mathcal{R}_m) \leq b + 2 \text{diam}(\alpha_G^{-1}(\lceil v - be, v - ae \rceil)) \leq 3\gamma_G(b).$$

Also,

$$(6.10) \quad \pi_m(\mathcal{R}_m^c) = \mu|_{m'}(\{x \in m' : (x, x) \in \mathcal{R}_m^c\}) = 0.$$

For  $A$  a Borel subset of  $m$ ,

$$\pi(A \times m') = \mu(A \cap m')$$

and for  $A'$  a Borel subset of  $m'$ ,

$$\pi(m \times A') = \mu(A').$$

Thus,

$$(6.11) \quad D(\pi; \mu|_m, \mu|_{m'}) \leq \mu(\alpha_G^{-1}(\lceil v - be, v - ae \rceil)) \leq \tilde{\gamma}_G(b).$$

Inequalities (6.9), (6.10) and (6.11) show that for any  $m \in \text{comp}(\text{Frag}(G, \mathcal{P}_s)_{>\varepsilon})$ ,

$$d_{\text{GHP}}(m, \sigma(m)) \leq \alpha := \tilde{\gamma}_G(b) \vee 3\gamma_G(b).$$

The same inequalities show also that for any  $m' \in \text{comp}(\text{Frag}(G', \mathcal{P}'_s)_{>\varepsilon})$ ,

$$d_{\text{GHP}}(m', \sigma'(m')) \leq \alpha.$$

Now, let us fix

$$\varepsilon = \sqrt{\alpha} + \alpha$$

so that (6.8) is satisfied. Using the last inequality of Lemma 4.7 with  $p = 1$  and  $\alpha, \varepsilon, \sigma$  and  $\sigma'$  as above, we have shown that as soon as  $\mathcal{E}$  holds, for any  $s \in [0, t]$ ,

$$\begin{aligned} & L_{\text{GHP}}(\text{Frag}(G, \mathcal{P}_s), \text{Frag}(G^{(e,a,b)}, \mathcal{P}'_s)) \\ & \leq 8\alpha \frac{\mu(G)}{\sqrt{\alpha}} + 16\varepsilon \\ & = 17(\tilde{\gamma}_G(b) \vee 3\gamma_G(b)) + (16 + \mu(G))\sqrt{\tilde{\gamma}_G(b) \vee 3\gamma_G(b)}. \end{aligned}$$

Furthermore,

$$\|\text{masses}(\text{Frag}(G, \mathcal{P}_s)) - \text{masses}(\text{Frag}(G^{(e,a,b)}, \mathcal{P}'_s))\|_2^2 \leq 2\tilde{\gamma}_G(b)^2.$$

Also, for any  $m$  in  $\text{comp}(\text{Frag}(G, \mathcal{P}_s)_{>\varepsilon})$ ,  $m$  and  $\sigma(m)$  have the same surplus (recall the gluing in Definition 6.9). The same is true for  $m'$  and  $\sigma'(m')$ . Notice also that

$$\tilde{\gamma}_G(b) \leq \mu(G) \quad \text{and} \quad \gamma_G(b) \leq \text{diam}(G).$$

Thus,

$$\begin{aligned}
 &L_{2,\text{GHP}}^{\text{surplus}}(\text{Frag}(\mathbf{G}, \mathcal{P}_s), \text{Frag}(\mathbf{G}^{(e,a,b)}, \mathcal{P}'_s)) \\
 &\leq [17(\tilde{\gamma}_G(b) \vee 3\gamma_G(b)) + (16 + \mu(G))\sqrt{\tilde{\gamma}_G(b) \vee 3\gamma_G(b)}] \vee 2\tilde{\gamma}_G(b)^2 \\
 &\leq C(G)\sqrt{\tilde{\gamma}_G(b) \vee \gamma_G(b)},
 \end{aligned}$$

where  $C(G)$  is a positive constant depending only on  $\text{diam}(G)$  and  $\mu(G)$ . Finally, notice that  $\mathcal{E}$  has probability at least  $\exp(-t(b - a)) \geq 1 - t(b - a)$ .  $\square$

Now, we shall prove Theorem 3.2.

**PROOF OF THEOREM 3.2.** The proofs of (i) and (iii) are completely analogous to the proofs of (i) and (iii) in Proposition 6.4, so we leave them to the reader.

Let us prove (ii). First, we may suppose, thanks to Skorokhod representation theorem, that  $\mathbf{G}^n$  and  $\mathbf{G}$  are deterministic and that  $\mathbf{G}^n$  converges to  $\mathbf{G}$  as  $n$  goes to infinity. Thanks to Remarks 2.12(iv) and 2.22(iii), it is sufficient to prove Theorem 3.2 when the components of  $\mathbf{G}^n$  and  $\mathbf{G}$  are genuine metric spaces, that is,  $\mathbb{R}$ -graphs.

The argument at the beginning of the proof of Proposition 6.4 shows that it is sufficient to prove the result when  $\mathbf{G}^n$  and  $\mathbf{G}$  have a single component. Let  $\mathbf{G} = (G, d, \mu)$  be a measured  $\mathbb{R}$ -graph and let  $\varepsilon > 0$ . We want to show that  $\text{Frag}(\mathbf{G}^n, t)$  converges in distribution to  $\text{Frag}(\mathbf{G}, t)$  when  $\mathbf{G}^n$  is a sequence of  $\mathbb{R}$ -graphs which converges to  $\mathbf{G}$  while having the same surplus.

Let  $\delta < \delta(\varepsilon, G)$  be given by Lemma 6.7 and let  $\mathbf{G}'$  be such that  $d_{\text{GHP}}(\mathbf{G}, \mathbf{G}') < \delta$  (we will take  $\delta$  small enough later). Thus, there is a correspondence  $\mathcal{R}_0 \in C(G, G')$  and a measure  $\pi_0 \in M(G, G')$  such that

$$\text{dis}(\mathcal{R}_0) \vee \pi_0(\mathcal{R}_0^c) \vee D(\pi_0; \mu, \mu') < \delta.$$

Lemma 6.7 shows that there exists an  $\varepsilon$ -overlay  $\mathcal{R} \in C(G, G')$  containing  $\mathcal{R}_0$ . Let us denote by  $\chi$  the multigraph isomorphism from  $\ker(G)$  to  $\ker(G')$  given by this overlay. For any edge,  $e \in \ker(G)$ ,  $|\ell(e) - \ell'(\chi(e))| < \varepsilon$ .

We define two graphs  $\tilde{\mathbf{G}}$  and  $\tilde{\mathbf{G}}'$  obtained from  $\mathbf{G}$  and  $\mathbf{G}'$  as follows. For each oriented edge  $e = (u, v) \in \ker(G)$ , denoting  $(u', v') = \chi(e)$  and  $\eta_e := |\ell(e) - \ell'(\chi(e))|$ , which is less than  $\varepsilon$ :

- if  $\ell(e)$  is smaller than  $\ell'(e')$ , we replace  $G'$  by its  $(6\varepsilon - \eta_e, 6\varepsilon)$ -shortening along  $e'$  (cf. Definition 6.9),
- if  $\ell'(e')$  is smaller than  $\ell(e)$ , we replace  $G$  by its  $(6\varepsilon - \eta_e, 6\varepsilon)$ -shortening along  $e$ .

Let us denote by  $(\tilde{G}, \tilde{d})$  and  $(\tilde{G}', \tilde{d}')$  the resulting  $\mathbb{R}$ -graphs, let  $\tilde{\mu} := \mu|_{\tilde{G}}$ ,  $\tilde{\mu}' := \mu'|_{\tilde{G}'}$  and define  $\tilde{\mathbf{G}} := (\tilde{G}, \tilde{d}, \tilde{\mu})$ ,  $\tilde{\mathbf{G}}' := (\tilde{G}', \tilde{d}', \tilde{\mu}')$ .

Recalling the notation in Lemma 6.10, let

$$\kappa := \gamma_G(11\varepsilon) + 12\varepsilon$$

and define  $\mathcal{R}_1$  the  $\kappa$ -enlargement of  $\mathcal{R}$ . We will show that

$$(6.12) \quad \tilde{\mathbf{G}} \text{ and } \tilde{\mathbf{G}}' \text{ are in correspondence through } \mathcal{R}_1.$$

If  $x \in \tilde{\mathbf{G}}$  and  $(x, x') \in \mathcal{R}$  with  $x' \notin \tilde{\mathbf{G}}'$ , then  $x' \in \alpha_{G'}^{-1}([v' - 6\varepsilon e', v' - (6\varepsilon - \eta_e)e'])$  for some edge  $e' = (u', v')$  of  $\ker(G')$ . Lemma 6.8 shows that

$$(6.13) \quad 0 < 6\varepsilon - \eta_e - 5 \text{dis}(\mathcal{R}) \leq d(\alpha_G(x), v) \leq 6\varepsilon + 5 \text{dis}(\mathcal{R}) \leq 11\varepsilon$$



and thus

$$d(x, v) \leq \gamma_G(11\varepsilon) + 11\varepsilon.$$

Thus,

$$(6.14) \quad d'(x', v') \leq \gamma_G(11\varepsilon) + 12\varepsilon = \kappa.$$

This shows that  $(x, v') \in \mathcal{R}_1$ .

Now, suppose  $x' \in \tilde{G}'$  and  $(x, x') \in \mathcal{R}$  with  $x \notin \tilde{G}$ . Then,  $x \in \alpha_G^{-1}([v - 6\varepsilon e, v - (6\varepsilon - \eta_e)'])$  for some edge  $e = (u, v)$  of  $\ker(G)$  and  $\eta < \varepsilon$ . Notice that

$$d(x, v) \leq \gamma_G(6\varepsilon) + 6\varepsilon,$$

and

$$d'(x', v') \leq d(x, v) + \text{dis}(\mathcal{R}) \leq \gamma_G(6\varepsilon) + 7\varepsilon \leq \kappa.$$

Thus  $(v, x') \in \mathcal{R}_1$ . This completes the proof of (6.12).

Notice that

$$\text{dis}(\mathcal{R}_1) \leq \varepsilon + 4\kappa.$$

Let  $\mathcal{R}_2 := \mathcal{R}_1|_{\tilde{G} \times \tilde{G}'} \in C(\tilde{G}, \tilde{G}')$ . Let  $K$  be the number of edges in  $\ker(G)$ . Notice that

$$\forall (x, y) \in \tilde{G}, \quad |d(x, y) - \tilde{d}(x, y)| < K\varepsilon$$

and

$$\forall (x', y') \in \tilde{G}', \quad |d'(x', y') - \tilde{d}'(x', y')| < K\varepsilon.$$

Thus,

$$(6.15) \quad \text{dis}(\mathcal{R}_2) \leq 2K\varepsilon + \text{dis}(\mathcal{R}_1) < (2K + 49)\varepsilon + 4\gamma_G(11\varepsilon).$$

Clearly, there exists a homeomorphism  $\psi$  from  $\text{core}(\tilde{G})$  to  $\text{core}(\tilde{G}')$  which preserves the length-measure. For each oriented edge  $e = (u, v) \in \ker(G)$ , denoting  $(u', v') = \chi(e)$ ,  $\psi$  satisfies  $\psi(v) = v'$ . Furthermore, since  $e$  and  $e'$  are in correspondence through the overlay  $\mathcal{R}$ , we have, for each  $x \in [u, v]$ , that there exists  $x' \in [u', v']$  such that  $(x, x') \in \mathcal{R}$  and

$$|d(x, u) - d'(x', u')| < \varepsilon.$$

If furthermore  $x \in \text{core}(\tilde{G})$ , we know that  $d(x, u) = d'(\psi(x), u')$ , so

$$|d'(\psi(x), u') - d'(x', u')| < \varepsilon.$$

Since  $x'$  and  $\psi(x)$  belong to  $[u', v']$ ,

$$d'(\psi(x), x') = |d'(\psi(x), u') - d'(x', u')| < \varepsilon,$$

which shows that for every  $x \in \text{core}(\tilde{G})$ ,

$$(6.16) \quad (x, \psi(x)) \text{ belongs to } \mathcal{R}_2,$$

the restriction to  $\tilde{G} \times \tilde{G}'$  of the  $\kappa$ -enlargement of  $\mathcal{R}$ .

Now, let  $\pi := \pi_0|_{\tilde{G} \times \tilde{G}'} \in M(\tilde{G}, \tilde{G}')$ . First,

$$(6.17) \quad \pi(\mathcal{R}_2^c) = \pi(\mathcal{R}_1^c) \leq \pi_0(\mathcal{R}_0^c) < \varepsilon.$$

Then

$$(6.18) \quad D(\pi; \tilde{\mu}, \tilde{\mu}') \leq 2D(\pi_0; \mu, \mu') + \mu(G \setminus \tilde{G}) + \mu'(G' \setminus \tilde{G}').$$

Now, recall that

$$\tilde{\gamma}_G(\varepsilon) := \sum_{e=(u,v) \in \ker(G)} \mu(\alpha_G^{-1}([v - \varepsilon e, v]))$$

which goes to zero as  $\varepsilon$  goes to zero. We have

$$\mu(G \setminus \tilde{G}) \leq \tilde{\gamma}_G(6\varepsilon).$$

Furthermore, recall inequality (6.13) which shows that if  $x' \in G' \setminus \tilde{G}'$ , then for every  $x \in G$  such that  $(x, x') \in \mathcal{R}$ ,

$$x \in \bigcup_{e=(u,v) \in \ker(G)} \alpha_G^{-1}([v - 11\varepsilon e, v]).$$

Thus,

$$\begin{aligned} \mu'(G' \setminus \tilde{G}') &\leq \pi_0(G \times (G' \setminus \tilde{G}')) + D(\pi_0; \mu, \mu') \\ &\leq \pi_0((G \times (G' \setminus \tilde{G}')) \cap \mathcal{R}) + \pi_0(\mathcal{R}^c) + \varepsilon \\ (6.19) \quad &\leq \pi_0\left(\bigcup_{e=(u,v) \in \ker(G)} \alpha_G^{-1}([v - 11\varepsilon e, v]) \times G'\right) + 2\varepsilon \\ &\leq \mu\left(\bigcup_{e=(u,v) \in \ker(G)} \alpha_G^{-1}([v - 11\varepsilon e, v])\right) + D(\pi_0; \mu, \mu') + 2\varepsilon \\ &\leq \tilde{\gamma}_G(11\varepsilon) + 3\varepsilon. \end{aligned}$$

Thus, using (6.18),

$$(6.20) \quad D(\pi; \tilde{\mu}, \tilde{\mu}') \leq 5\varepsilon + 2\tilde{\gamma}_G(11\varepsilon).$$

Gathering (6.16), (6.15), (6.17) and (6.20) shows that one may apply Lemma 6.5, in the sense that there is a function  $f_G(\varepsilon)$  going to zero as  $\varepsilon$  goes to zero such that the Lévy–Prokhorov distance (for the Skorokhod topology associated to  $L_{2,GHP}^{\text{surplus}}$ ) between the distributions of  $(\text{Frag}(\tilde{G}, s))_{s \in [0,t]}$  and  $(\text{Frag}(\tilde{G}', s))_{s \in [0,t]}$  is less than  $f_G(\varepsilon)$ .

On the other hand, let  $v' \in \ker(G')$  and  $x' \in \alpha_{G'}^{-1}([v - \varepsilon, v])$ . Let  $v \in \ker(G)$  (resp.  $x \in G$ ) be such that  $(v, v') \in \mathcal{R}$  (resp.  $(x, x') \in \mathcal{R}$ ). Then

$$d'(x', v') \leq d(x, v) + \varepsilon.$$

But using Lemma 6.8,

$$d(\alpha_G(x), v) \leq d'(\alpha_{G'}(x'), v') + 5\varepsilon \leq 6\varepsilon,$$

which implies

$$d(x, v) \leq \gamma_G(6\varepsilon).$$

Thus,

$$d'(x', v') \leq \gamma_G(6\varepsilon) + \varepsilon,$$

which shows that

$$\gamma_{G'}(\varepsilon) \leq \gamma_G(6\varepsilon) + \varepsilon.$$

Inequality (6.19) shows that

$$\tilde{\gamma}_{G'}(\varepsilon) \leq \tilde{\gamma}_G(11\varepsilon) + 3\varepsilon.$$

Then Lemma 6.11 shows that there is a function  $f_G(\varepsilon)$  going to zero as  $\varepsilon$  goes to zero such that the Lévy–Prokhorov distance between the distributions of  $(\text{Frag}(G, s))_{s \in [0, t]}$  and  $(\text{Frag}(\tilde{G}, s))_{s \in [0, t]}$  is less than  $f_G(\varepsilon)$  and the Lévy–Prokhorov distance (for the Skorokhod topology associated to  $L_{2, \text{GHP}}^{\text{surplus}}$ ) between the distributions of  $(\text{Frag}(G', s))_{s \in [0, t]}$  and  $(\text{Frag}(\tilde{G}', s))_{s \in [0, t]}$  is less than  $f_G(\varepsilon)$ . This completes the proof of (ii) (through Lemma A.2 and inequality (A.1)).  $\square$

6.5. *Application to Erdős–Rényi random graphs: Proofs of Theorem 3.6 and Proposition 3.8.* Let us first compare the discrete fragmentation process and the continuous one. Let  $\mathcal{P}^-$  be a Poisson process driving the discrete fragmentation on  $G^n := G(n, p(\lambda, n))$ . Recall that  $N^-(G^n, \mathcal{P}_t^-)$  is the state of this process at time  $t$ , seen as an element of  $\mathcal{N}_2^{\text{graph}}$ . Let  $\mathcal{Q}^-$  be a Poisson process of intensity  $\ell_n \otimes \text{leb}_{\mathbb{R}^+}$  on  $K_n \times \mathbb{R}^+$  where  $K_n$  is the complete graph on  $n$  vertices seen as an  $\mathbb{R}$ -graph where the edge lengths are  $\delta_n = n^{-1/3}$  and  $\ell_n$  is its length measure. Then one may suppose that  $\mathcal{P}^-$  is obtained as follows:

$$\mathcal{P}^- = \{(e, t) : \exists x \in e, (x, t) \in \mathcal{Q}^-\}.$$

Then, for any  $t$ ,  $N^-(G^n, \mathcal{P}_t^-)$  is at  $L_{2, \text{GHP}}$ -distance at most  $n^{-1/3}$  from  $\text{Frag}(G^n, \mathcal{Q}_t^-)$  (cf. for instance [4], Proposition 3.4). Recall that by Theorem 2.26,  $\bar{\mathcal{G}}_{n, \lambda}$  (which is  $G^n$  with edge length  $\delta_n$  and vertex weights  $n^{-2/3}$ ) converges in distribution to  $\mathcal{G}_\lambda$  for  $L_{2, \text{GHP}}^{\text{surplus}}$ . Thus Theorem 3.2 implies that  $(\text{Frag}(G^n, \mathcal{Q}_t^-))_{t \geq 0}$ , and thus  $(N^+(G^n, \mathcal{P}_t^-))_{t \geq 0}$ , converges to  $(\text{Frag}(\mathcal{G}_\lambda, t))_{t \geq 0}$  as  $n$  goes to infinity (in the Skorokhod topology associated to  $L_{2, \text{GHP}}^{\text{surplus}}$ ). This shows Theorem 3.6.

We are now able to prove Proposition 3.8. Take  $\mathcal{P}_t^+$  of intensity  $\gamma = n^{-4/3}$ . Notice that the states of the edges are independent and identically distributed in  $(\mathcal{G}(n, p), N^+(\mathcal{G}(n, p), \mathcal{P}_t^+))$ . Let  $(X, Y)$  be the joint distribution of the state of one edge in  $(\mathcal{G}(n, p), N^+(\mathcal{G}(n, p), \mathcal{P}_t^+))$ . Denoting by 0 the state “absent” and 1 the state “present” it is easy to compute this distribution:

$$\begin{aligned} \mathbb{P}((X, Y) = (0, 0)) &= (1 - p)e^{-\gamma t}, & \mathbb{P}((X, Y) = (0, 1)) &= (1 - p)(1 - e^{-\gamma t}), \\ \mathbb{P}((X, Y) = (1, 0)) &= 0, & \mathbb{P}((X, Y) = (1, 1)) &= p. \end{aligned}$$

Now, take  $\mathcal{P}_t^-$  of intensity  $\mu = n^{-1/3}$  and let  $(X', Y')$  be the joint distribution of the state of one edge in  $(N^-(G(n, p'), \mathcal{P}_t^-), G(n, p'))$ . Then

$$\begin{aligned} \mathbb{P}((X', Y') = (0, 0)) &= (1 - p'), & \mathbb{P}((X, Y) = (0, 1)) &= p'(1 - e^{-\mu t'}), \\ \mathbb{P}((X', Y') = (1, 0)) &= 0, & \mathbb{P}((X', Y') = (1, 1)) &= p'e^{-\mu t'}. \end{aligned}$$

Thus, if one chooses

$$t = \frac{1}{\gamma} \ln \frac{1 - p}{1 - p'} \quad \text{and} \quad t' = \frac{1}{\mu} \ln \frac{p'}{p},$$

then  $(\mathcal{G}(n, p), N^+(\mathcal{G}(n, p), \mathcal{P}_t^+))$  and  $(N^-(G(n, p'), \mathcal{P}_{t'}^-), G(n, p'))$  have the same distribution. Now, take  $p = p(\lambda, n)$ ,  $p' = p(\lambda + s, n)$ . We have

$$t = n^{4/3} \ln \left( 1 + \frac{s}{n^{4/3}(1 - p')} \right) \xrightarrow{n \rightarrow \infty} s.$$

We consider that  $\mathcal{G}(n, p)$  is equipped with edge lengths  $n^{-1/3}$  and vertex weight  $n^{-2/3}$ . Thus Theorem 3.5 shows that  $(\mathcal{G}(n, p), N^+(\mathcal{G}(n, p), \mathcal{P}_t^+))$  converges in distribution to  $(\mathcal{G}_\lambda, \text{Coal}_0(\mathcal{G}_\lambda, s))$ . Also,

$$t' = n^{1/3} \ln \frac{1 + \frac{\lambda + s}{n^{1/3}}}{1 + \frac{\lambda}{n^{1/3}}} \xrightarrow{n \rightarrow \infty} s$$

thus Theorem 3.2 shows that  $(N^-(G(n, p'), \mathcal{P}_t^-), G(n, p'))$  converges in distribution to  $(\text{Frag}(\mathcal{G}_{\lambda+s}, s), \mathcal{G}_{\lambda+s})$ . Thus  $(\mathcal{G}_\lambda, \text{Coal}_0(\mathcal{G}_\lambda, s))$  and  $(\text{Frag}(\mathcal{G}_{\lambda+s}, s), \mathcal{G}_{\lambda+s})$  have the same distribution. This completes the proof of Proposition 3.8.

Notice a curious fact: in [6], Theorem 3, it is shown that the sizes of the components of a fragmentation on the CRT are the time-reversal (after an exponential time-change) of the standard *additive* coalescent. It would be interesting to make a direct link between additive and multiplicative coalescent in the context of fragmentation on  $\mathcal{G}_\lambda$ .

**7. Combining fragmentation and coalescence: Dynamical percolation.**

7.1. *Almost Feller property: Proof of Theorem 3.3.* In this section, we prove Theorem 3.3. The following lemma is a simple variation on the proof of (5.4).

LEMMA 7.1. *Let  $X^n = (X^n, d^n, \mu^n)$ ,  $n \geq 0$  be a sequence of random m.s-m.s in  $\mathcal{N}_2^{\text{graph}}$  and  $(\delta^n)_{n \geq 0}$  be a sequence of nonnegative real numbers. Suppose that:*

- (i)  $(X^n)$  converges in distribution (for  $L_{2,\text{GHP}}$ ) to  $X^\infty = (X^\infty, d^\infty, \mu^\infty)$  as  $n$  goes to infinity,
- (ii)  $\delta^n \xrightarrow{n \rightarrow \infty} 0$ ,
- (iii) for any  $\alpha > 0$  and any  $T > 0$ ,

$$\limsup_{n \in \mathbb{N}} \mathbb{P}(\text{suplength}(\text{Coal}_{\delta^n}(X^n_{\leq \varepsilon}, T)) > \alpha) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Then, for any  $\alpha > 0$  and any  $T > 0$ ,

$$\mathbb{P}(\text{suplength}(\text{Coal}_0(X^\infty_{\leq \varepsilon}, T)) > \alpha) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

PROOF. The situation is simpler than in the proof of (5.4), since *suplength* is non-decreasing under coalescence. Using the notation of the proof of (5.4),

$$\begin{aligned} &\mathbb{P}(\text{suplength}(\text{Coal}_0(X^\infty_{\leq \varepsilon_m}, T)) > \alpha) \\ &= \lim_{p \rightarrow \infty} \mathbb{P}(\text{suplength}(\text{Coal}_0(X^\infty_{m,p}, T)) > \alpha). \end{aligned}$$

Now, Proposition 5.3 implies that  $(\text{Coal}_{\delta^n}(X^n_{m,p}, T))$  converges in distribution to  $(\text{Coal}_0(X^\infty_{m,p}, T))$  for any  $m \leq p$ . Thus, for any  $m \leq p$ ,

$$\begin{aligned} &\mathbb{P}(\text{suplength}(\text{Coal}_0(X^\infty_{m,p}, T)) > \alpha) \\ &\leq \limsup_{n \rightarrow \infty} \mathbb{P}(\text{suplength}(\text{Coal}_{\delta^n}(X^n_{m,p}, T)) > \alpha) \\ &\leq \limsup_{n \rightarrow \infty} \mathbb{P}(\text{suplength}(\text{Coal}_{\delta^n}(X^n_{\leq \varepsilon_m}, T)) > \alpha), \end{aligned}$$

which goes to zero when  $m$  goes to infinity.  $\square$

Now, we are able to prove Theorem 3.3.

Notice that the strong Markov property was already noticed (see the remark after Definition 2.25). The fact that trajectories lie in  $\mathcal{S}^{\text{length}}$  is a consequence of Lemmas 5.9 and 5.11. Thus, to prove (i), we only need to prove that the trajectories are càdlàg (almost surely). This will be done in the course of proving point (ii).

We will reduce the problem to  $\mathcal{N}_1^{\text{graph}}$  using a variation on the proof of Theorem 3.1. Let us study first  $\text{Frag}(\text{Coal}_{\delta^n}(X^n, \mathcal{P}_t^+, \mathcal{P}_t^-)$  with  $\mathcal{P}^+$  and  $\mathcal{P}^-$  as in Definitions 2.13 and 2.23. Let us fix  $\varepsilon \in ]0, 1/2[$  and  $0 \leq t \leq T$ . Any component of size at least  $\varepsilon$  in

$\text{Frag}(\text{Coal}_{\delta^n}(X^n, \mathcal{P}_t^+), \mathcal{P}_t^-)$  has to belong to a component of size at least  $\varepsilon$  in  $\text{Coal}_{\delta^n}(X^n, \mathcal{P}_t^+)$ . Let  $x^n := \text{masses}(X^n)$  for  $n \in \overline{\mathbb{N}}$ .

As in the proof of Theorem 3.1, we obtain that there exists  $K(\varepsilon)$ ,  $\varepsilon_1 \in ]0, \varepsilon[$  and  $\varepsilon_2 \in ]0, \varepsilon[$  such that for every  $n \in \overline{\mathbb{N}}$ , with probability larger than  $1 - \varepsilon$  the event  $\mathcal{A}_n$  holds, where  $\mathcal{A}_n$  is the event that points (a), (b) and (c) of Corollary 5.6 hold for any  $t \in [0, T]$  and  $S(x^n, T) \leq K(\varepsilon)$ .

Let us place ourselves on  $\mathcal{A}_n$ . Then, for a significant component at time  $t$ , notice that fragmentation on the hanging trees of components does change neither the mass neither the distance in the heart of a component. Thus, the same proof as that of Theorem 3.1 shows that on  $\mathcal{A}_n$ , we have for every time  $t \leq T$  and every  $\varepsilon'_2 \leq \varepsilon_2$ :

$$\begin{aligned} &L_{\text{GHP}}(\text{Frag}(\text{Coal}_{\delta^n}(X^n, \mathcal{P}_t^+), \mathcal{P}_t^-), \text{Frag}(\text{Coal}_{\delta^n}(X^n_{>\varepsilon'_2}, \mathcal{P}_t^+), \mathcal{P}_t^-)) \\ &\leq 17(\delta^n + \text{supdiam}(\text{Frag}(\text{Coal}_{\delta^n}(X^n_{\leq \varepsilon_1}, \mathcal{P}_T^+), \mathcal{P}_T^-)) + \varepsilon_1) \left(1 + \frac{8K(\varepsilon)}{\varepsilon^2}\right) \\ &\quad + 16\varepsilon. \end{aligned}$$

A slight difference occurs here:

$$t \mapsto \text{supdiam}(\text{Frag}(\text{Coal}_{\delta^n}(X^n_{\leq \varepsilon_1}, \mathcal{P}_T^+), \mathcal{P}_T^-))$$

is not necessarily nonincreasing. However, the supremum of the lengths of injective paths clearly decreases (non-strictly) under fragmentation. Thus, on  $\mathcal{A}_n$ ,

$$\begin{aligned} &L_{\text{GHP}}(\text{Frag}(\text{Coal}_{\delta^n}(X^n, \mathcal{P}_t^+), \mathcal{P}_t^-), \text{Frag}(\text{Coal}_{\delta^n}(X^n_{>\varepsilon'_2}, \mathcal{P}_t^+), \mathcal{P}_t^-)) \\ &\leq 17(\delta^n + \text{suplength}(\text{Coal}_{\delta^n}(X^n_{\leq \varepsilon_1}, \mathcal{P}_T^+)) + \varepsilon_1) \left(1 + \frac{8K(\varepsilon)}{\varepsilon^2}\right) + 16\varepsilon. \end{aligned}$$

Let  $V = \text{comp}(\text{Frag}(X^n, \mathcal{P}_t^-))$  and  $W = \text{comp}(\text{Frag}(X^n_{>\varepsilon'_2}, \mathcal{P}_t^-)) \subset V$ . Let  $E'$  denote the set of pairs  $(i, j)$  in  $V^2$  such that  $i \sim j$  if and only if  $i$  and  $j$  are at finite distance in  $\text{Frag}(\text{Coal}_{\delta^n}(X^n, \mathcal{P}_t^+), \mathcal{P}_t^-)$ . Let  $E$  denote the set of pairs  $(i, j)$  in  $V^2$  such that  $i \sim j$  if and only if  $i$  and  $j$  are at finite distance in  $\text{Coal}_{\delta^n}(X^n, \mathcal{P}_t^+)$ . Define

$$x^n(t) := \text{masses}(\text{Frag}(\text{Coal}_{\delta^n}(X^n, \mathcal{P}_t^+), \mathcal{P}_t^-))$$

and

$$x^n_{>\varepsilon}(t) := \text{masses}(\text{Frag}(\text{Coal}_{\delta^n}(X^n_{>\varepsilon}, \mathcal{P}_t^+), \mathcal{P}_t^-)).$$

Lemma 4.4 shows that

$$\begin{aligned} &\|x^n(t) - x^n_{>\varepsilon'_2}(t)\|_2^2 \\ &\leq \|x^n(t)\|_2^2 - \|x^n_{>\varepsilon'_2}(t)\|_2^2 \end{aligned}$$

then, we can use Lemma 4.5, since (b) of Corollary 5.6 holds on  $\mathcal{A}_n$ :

$$\begin{aligned} &\|x^n(t) - x^n_{>\varepsilon'_2}(t)\|_2^2 \\ &\leq S(x^n, t) - S(x^n_{>\varepsilon_2}, t) \\ &\leq 2\varepsilon_1^2 \leq \varepsilon \end{aligned}$$

since (c) of Corollary 5.6 holds on  $\mathcal{A}_n$ . Now, let us take  $\delta_n = 0$ . Define

$$X^n(t) := \text{Frag}(\text{Coal}_0(X^n, \mathcal{P}_t^+), \mathcal{P}_t^-) = \text{CoalFrag}(X^n, t)$$

and

$$X_{>\varepsilon_1}^n(t) := \text{Frag}(\text{Coal}_0(X_{>\varepsilon_1}^n, \mathcal{P}_t^+), \mathcal{P}_t^-) = \text{CoalFrag}(X_{>\varepsilon_1}^n, t).$$

Notice that we used the commutation relation guaranteed by Lemma 5.9. Using the hypothesis (3.2) on suplength and Lemma 7.1, we conclude that for any  $\varepsilon > 0$ ,

$$(7.1) \quad \lim_{\varepsilon_1 \rightarrow 0} \sup_{n \in \overline{\mathbb{N}}} \mathbb{P}^* \left[ \sup_{t \in [0, T]} L_{2, \text{GHP}}(X^n(t), X_{>\varepsilon_1}^n(t)) > \varepsilon \right] = 0.$$

Now, let us prove that the trajectories of  $\text{CoalFrag}(X^\infty, \cdot)$  are càdlàg. Let  $Y^n := R_{\frac{1}{n}}(X_{>\frac{1}{n}}^\infty)$ .  $\text{CoalFrag}(Y^n, \cdot)$  is càdlàg (it has a finite number of jumps on any bounded interval). Now, Lemma 6.1 applied to  $X_{>\frac{1}{n}}^\infty$  and equation (7.1) applied to  $X^n := X^\infty$  show that the hypotheses of Lemma A.1 are satisfied for  $\omega^n = \text{CoalFrag}(Y^n, \cdot)$  and  $\omega^\infty = \text{CoalFrag}(X^\infty, \cdot)$ . This proves that the trajectories of  $\text{CoalFrag}(X^\infty, \cdot)$  are càdlàg and completes the proof of (i).

Now, let us prove point (ii). Equation (7.1) shows that it is sufficient to show the proposition for  $X^n$  converging to  $X$  in  $L_{1, \text{GHP}}$  with  $X^n$  and  $X$  being m.s.m.s with a finite number of components which are  $\mathbb{R}$ -graphs. We shall only sketch the proof, since it is a variation on the arguments of the proofs of Proposition 5.3 and Theorem 3.2. For any  $n$  large enough, the proof of Theorem 3.2 shows that one may couple a Poisson random set  $\mathcal{P}^{-,n}$  on  $X^n \times \mathbb{R}^+$  with intensity measure  $\ell_{X^n} \otimes \text{leb}_{\mathbb{R}^+}$  with a Poisson random set  $\mathcal{P}^-$  on  $X \times \mathbb{R}^+$  with intensity  $\ell_X \otimes \text{leb}_{\mathbb{R}^+}$  and one may find  $\pi^n \in M(X, X^n)$  and  $\mathcal{R}^n \in C(X, X^n)$  such that there is an event  $\mathcal{E}_n$  in the  $\sigma$ -algebra of  $(\mathcal{P}_t^{-,n}, \mathcal{P}_t^-)$  and a sequence  $\varepsilon_n$  such that:

- (i)  $\mathbb{P}(\mathcal{E}_n^c) \leq \varepsilon_n$ ,
- (ii)  $\varepsilon_n \xrightarrow{n \rightarrow \infty} 0$ ,
- (iii) on  $\mathcal{E}_n$ , for any  $s \leq t$ ,  $\mathcal{R}^n \cap (X \setminus \mathcal{P}_s^-) \times (X^n \setminus \mathcal{P}_s^{-,n}) \in C(X \setminus \mathcal{P}_s^-, X^n \setminus \mathcal{P}_s^{-,n})$  and

$$D(\pi|_{(X \setminus \mathcal{P}_s^-) \times (X^n \setminus \mathcal{P}_s^{-,n})}; \mu|_{X \setminus \mathcal{P}_s^-}, \mu_{X^n \setminus \mathcal{P}_s^{-,n}}) \vee \pi^n((\mathcal{R}^n)^c) \vee \text{dis}_s(\mathcal{R}^n) \leq \varepsilon_n,$$

where  $\text{dis}_s(\mathcal{R}^n)$  is the distortion of  $\mathcal{R}^n$  as a correspondence between the semimetric spaces  $\text{Frag}(X, \mathcal{P}_s^-)$  and  $\text{Frag}(X^n, \mathcal{P}_s^{-,n})$ ,

- (iv) on  $\mathcal{E}_n$ , for any  $s \leq t$ ,  $\|\text{masses}(\text{Frag}(X, \mathcal{P}_s^-)) - \text{masses}(\text{Frag}(X^n, \mathcal{P}_s^{-,n}))\|_1 \leq \varepsilon_n$ .

Then one may use the proofs of Lemma 5.2 and Proposition 5.3 to couple a Poisson random set  $\mathcal{P}^{+,n}$  on  $(X^n)^2 \times \mathbb{R}^+$  with intensity measure  $\frac{1}{2}(\mu^n)^\otimes \otimes \text{leb}_{\mathbb{R}^+}$  with a Poisson random set  $\mathcal{P}^+$  on  $(X)^2 \times \mathbb{R}^+$  with intensity  $\frac{1}{2}(\mu)^\otimes \otimes \text{leb}_{\mathbb{R}^+}$  in such a way that there is an event  $\mathcal{E}'_n$ , a sequence  $\varepsilon'_n$  such that:

- (i)  $\mathbb{P}((\mathcal{E}'_n)^c) \leq \varepsilon'_n$ ,
- (ii)  $\varepsilon'_n \xrightarrow{n \rightarrow \infty} 0$ ,
- (iii) on  $\mathcal{E}'_n$ , for any  $s \leq t$ ,

$$D(\pi|_{(X \setminus \mathcal{P}_s^-) \times (X^n \setminus \mathcal{P}_s^{-,n})}; \mu|_{X \setminus \mathcal{P}_s^-}, \mu^n|_{X^n \setminus \mathcal{P}_s^{-,n}}) \vee \pi^n((\mathcal{R}^n)^c) \vee \text{dis}'_s(\mathcal{R}^n) \leq \varepsilon'_n,$$

where  $\text{dis}'_s(\mathcal{R}^n)$  is the distortion of  $\mathcal{R}^n$  as a correspondence between the semimetric spaces  $\text{Coal}_0(\text{Frag}(X, \mathcal{P}_s^-), \mathcal{P}_s^+)$  and  $\text{Coal}_0(\text{Frag}(X^n, \mathcal{P}_s^{-,n}), \mathcal{P}_s^{+,n})$ ,

- (iv) on  $\mathcal{E}'_n$ , for any  $s \leq t$ , the multigraphs  $\text{MG}(\text{Frag}(X, \mathcal{P}_s^-), \mathcal{P}_s^+)$  and  $\text{MG}(\text{Frag}(X^n, \mathcal{P}_s^{-,n}), \mathcal{P}_s^{+,n})$  are the same.

Thanks to the properties of the multigraphs above, on  $\mathcal{E}_n \cap \mathcal{E}'_n$ , we get that for any  $s \leq t$ ,

$$\begin{aligned} & \|\text{masses}(\text{Coal}(\text{Frag}(X, \mathcal{P}_s^-), \mathcal{P}_s^+)) - \text{masses}(\text{Coal}(\text{Frag}(X^n, \mathcal{P}_s^{-,n}), \mathcal{P}_s^{+,n}))\|_1 \\ & \leq \|\text{masses}(\text{Frag}(X, \mathcal{P}_s^-)) - \text{masses}(\text{Frag}(X^n, \mathcal{P}_s^{-,n}))\|_1 \\ & \leq \varepsilon_n. \end{aligned}$$

Using Lemma 4.8, this ends the proof of the convergence in the sense of  $L_{1,\text{GHP}}$ , and thus the proof of point (i).

Finally, (iii) is a direct consequence of Theorem 3.1(iii) and Theorem 3.2(iii). This completes the proof of Theorem 3.3.

7.2. *Application to Erdős–Rényi random graphs.* Now, we want to prove Theorem 3.7 and Theorem 3.4. Intuitively, the dynamical percolation process on the complete graph  $K_n$  should be very close to the process  $\text{CoalFrag}(K_n, \cdot)$ , but such a statement needs some care, essentially because  $N^+$  and  $N^-$  do not commute: some pairs of vertices might be affected by the two Poisson processes  $\mathcal{P}^+$  and  $\mathcal{P}^-$  in a time interval  $[0, T]$ . Furthermore, the typical number of such edges is of order  $n^{1/3}$ . It turns out that these edges will not be important for the  $L_{\text{GHP}}$ -metric, but this issue requires us to adapt the proof of Theorem 3.3.

PROOF OF THEOREM 3.7. Let  $G^\infty = \mathcal{G}_\lambda$ . Let  $p = p(\lambda, n)$ , let  $G^n$  be the graph  $\mathcal{G}(n, p)$  seen as a measured  $\mathbb{R}$ -graph, with edge-lengths  $\delta_n := (1 - p)n^{-1/3} \sim n^{-1/3}$  and measure the counting measure times  $\sqrt{pn^{-1/3}} \sim n^{-2/3}$ . Let  $\mathcal{P}^+$  (of intensity  $pn^{-1/3}$ ) and  $\mathcal{P}^-$  (of intensity  $(1 - p)n^{-1/3}$ ) be the two Poisson processes driving the dynamical percolation on  $G^n$ . Let us write

$$G^n(t) := N(G^n, (\mathcal{P}^+, \mathcal{P}^-)_t)$$

and

$$G^n_{>\varepsilon_1}(t) := N(G^n_{>\varepsilon_1}, (\mathcal{P}^+, \mathcal{P}^-)_t)$$

for the state of this process at time  $t$ , seen as a member of  $\mathcal{N}_2^{\text{graph}}$ . Let us fix  $\varepsilon > 0$  and  $0 \leq t \leq T$ . Any component of size at least  $\varepsilon$  in  $N(G^n, (\mathcal{P}^+, \mathcal{P}^-)_t)$  has to belong to a component of size at least  $\varepsilon$  in  $N(G^n, (\mathcal{P}^+, \emptyset)_t)$ , which is nothing else but  $\text{Coal}_{\delta_n}(G^n, \mathcal{P}_t^+)$ . Now, we claim that

$$(7.2) \quad \limsup_{n \in \mathbb{N}} \mathbb{P}(\text{suplength}(\text{Coal}_{\delta_n}(G^n_{\leq \varepsilon}, T)) > \alpha) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Indeed, if  $G$  is a discrete graph with diameter  $D$  and surplus bounded from above by  $s$ ,

$$\text{suplength}(G) \leq 2D(1 + s).$$

Thus, (7.2) is a consequence of (5.6) and the fact that the maximal surplus in  $G^n$  forms a tight sequence (see, for instance, [24], Sections 13 and 14). Then (7.2) and Lemma 7.1 show that for any  $\alpha > 0$  and  $T > 0$ ,

$$\mathbb{P}(\text{suplength}(\text{Coal}_0(G^\infty_{\leq \varepsilon}, T)) > \alpha) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

The arguments leading to (7.1) show that

$$\lim_{\varepsilon_1 \rightarrow 0} \limsup_{n \in \mathbb{N}} \mathbb{P} \left[ \sup_{t \in [0, T]} L_{2,\text{GHP}}(G^n(t), G^n_{>\varepsilon_1}(t)) > \varepsilon \right] = 0$$

and

$$\lim_{\varepsilon_1 \rightarrow 0} \mathbb{P} \left[ \sup_{t \in [0, T]} L_{2,\text{GHP}}(\text{CoalFrag}(G^\infty, t), \text{CoalFrag}(G^\infty_{>\varepsilon_1}, t)) > \varepsilon \right] = 0.$$

Thus, it is sufficient to show that for any  $\varepsilon_1 > 0$ ,  $(G^n_{>\varepsilon_1}(t))_{t \geq 0}$  converges to  $(\text{CoalFrag}(G^\infty_{>\varepsilon_1}, t))_{t \geq 0}$  in the Skorokhod topology associated to  $L_{2,\text{GHP}}$ . Let  $Y_n$  denote the number of discrete coalescence events of  $\mathcal{P}_T^+$  occurring on  $G^n_{>\varepsilon_1}$ . Since the masses of  $G^n_{>\varepsilon_1}$  form a tight sequence,  $(Y_n)$  is a tight sequence. Since  $\delta_n$  goes to zero, the probability that  $\mathcal{P}_T^-$  touches an



edge from  $\mathcal{P}_T^+$  in  $G_{>\varepsilon_1}^n$  goes to zero as  $n$  goes to infinity. Thus, with probability going to one, for any  $t \in [0, T]$   $G_{>\varepsilon_1}^n(t) = \tilde{G}_{>\varepsilon_1}^n(t)$  where

$$\tilde{G}_{>\varepsilon_1}^n(t) = N(\text{Coal}_{\delta_n}(G_{>\varepsilon_1}^n, \mathcal{P}_t^+), (\emptyset, \mathcal{P}^-)_t).$$

Furthermore, since  $Y_n$  is a tight sequence and  $\delta_n$  goes to zero,

$$\sup_{t \in [0, T]} L_{2, \text{GHP}}(\tilde{G}_{>\varepsilon_1}^n(t), N(\text{Coal}_0(G_{>\varepsilon_1}^n, \mathcal{P}_t^+), (\emptyset, \mathcal{P}^-)_t)) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0.$$

Let  $\mathcal{Q}^-$  be a Poisson process of intensity  $\ell_n \otimes \text{leb}_{\mathbb{R}^+}$  on  $K_n \times \mathbb{R}^+$  where  $K_n$  is the complete graph on  $n$  vertices seen as an  $\mathbb{R}$ -graph where the edge lengths are  $\delta_n$  and  $\ell_n$  is its length measure. Then one may suppose that  $\mathcal{P}^-$  is obtained as follows:

$$\mathcal{P}^- = \{(e, t) : \exists x \in e, (x, t) \in \mathcal{Q}^-\}.$$

Then, for any  $t$ ,  $N(\text{Coal}_0(G_{>\varepsilon_1}^n, \mathcal{P}_t^+), (\emptyset, \mathcal{P}^-)_t)$  is at  $L_{2, \text{GHP}}$ -distance at most  $\delta_n$  from  $\text{Frag}(\text{Coal}_0(G_{>\varepsilon_1}^n, \mathcal{P}_t^+), \mathcal{Q}_t^-)$  (cf. for instance [4], Proposition 3.4). Altogether, we get

$$\sup_{t \in [0, T]} L_{2, \text{GHP}}(G_{>\varepsilon_1}^n(t), \text{Frag}(\text{Coal}_0(G_{>\varepsilon_1}^n, \mathcal{P}_t^+), \mathcal{Q}_t^-)) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0,$$

and  $(\text{Frag}(\text{Coal}_0(G_{>\varepsilon_1}^n, \mathcal{P}_t^+), \mathcal{Q}_t^-))_{t \geq 0}$  is distributed as  $\text{CoalFrag}(G_{>\varepsilon_1}^n, t)_{t \geq 0}$ . Now Theorem 3.3 shows that the sequence of processes  $\text{CoalFrag}(G_{>\varepsilon_1}^n, \cdot)$  converges to  $\text{CoalFrag}(G_{>\varepsilon_1}^\infty, \cdot)$  for the Skorokhod topology associated to  $L_{2, \text{GHP}}$ , which completes the proof.  $\square$

Finally, let us prove Theorem 3.4. First, notice that  $\mathcal{G}_\lambda \in \mathcal{S}^{\text{graph}}$  (it is a consequence of Theorem 3.1). For  $\text{Frag}(\mathcal{G}_\lambda, \cdot)$ , Theorem 3.4 is a consequence of Theorem 3.2 and Lemma 5.9. The fact that  $\text{Coal}(\mathcal{G}_\lambda, \cdot)$  and  $\text{CoalFrag}(\mathcal{G}_\lambda, \cdot)$  are strong Markov processes with càdlàg trajectories in  $\mathcal{S}^{\text{length}}$  is a consequence of the convergences already proven and of Theorems 3.1, 3.2 and 3.3. It remains to prove that almost surely, for any  $t \geq 0$ ,  $\text{Coal}(\mathcal{G}_\lambda, t)$ , and  $\text{CoalFrag}(\mathcal{G}_\lambda, t)$  belong to  $\mathcal{S}^{\text{graph}}$ .

Proposition 3.8 shows that  $\text{Coal}(\mathcal{G}_\lambda, t)$  has the same distribution as  $\mathcal{G}_{\lambda+t}$ . Thus, for a fixed  $t \geq 0$ ,  $\text{Coal}(\mathcal{G}_\lambda, t)$  is in  $\mathcal{S}^{\text{graph}}$  almost surely. But if  $\text{Coal}(\mathcal{G}_\lambda, t)$  is in  $\mathcal{S}^{\text{graph}}$ , then  $\text{Coal}(\mathcal{G}_\lambda, s)$  and  $\text{CoalFrag}(\mathcal{G}_\lambda, s)$  are in  $\mathcal{S}^{\text{graph}}$  for any  $s \leq t$ ; cf. Remark 5.10. This proves that  $\text{Coal}(\mathcal{G}_\lambda, t)$  and  $\text{CoalFrag}(\mathcal{G}_\lambda, \cdot)$  have trajectories in  $\mathcal{S}^{\text{graph}}$  and completes the proof of Theorem 3.4.

**8. Perspectives.** Today, there are a lot of results concerning convergence in distribution of rescaled critical random graphs to  $\mathcal{G}_\lambda$  (see [3, 7, 10]), or to other scaling limits (see [9, 11, 14]). For a lot of them, there is a notion of critical window parametrized by a real number  $\lambda$ , and the graph components merge approximately like a multiplicative coalescent when  $\lambda$  crosses the critical window. This is particularly obvious for critical percolation on the configuration model and inhomogeneous random graphs, [7, 9], but also parametrized rank-1 inhomogeneous random graphs [10, 11]. It is natural to expect the results of the present paper to apply to those random graphs. More precisely, I expect that analogues of Theorems 3.5, 3.6 and 3.7 hold. Let me describe quickly what I believe to be straightforward and what I believe not to be.

Concerning fragmentation, in order to apply Theorem 3.2, one needs, informally, joint convergence of the graphs, their size and their surplus. For instance, [19] and [9] should be enough to prove directly through Theorem 3.2 that the discrete fragmentation process on the critically percolated heavy-tailed configuration models of [9] converge to a fragmentation process on the limit. Since discrete fragmentation is itself a coupled percolation process (with

a decreasing percolation parameter), this in fact will show convergence of the critical percolation process (with decreasing percolation parameter) to the fragmentation process on the limit. For percolation on light-tailed configuration models at criticality, the analogous result should hold by resorting to the results of [18] and [7]. For rank-one inhomogeneous random graphs, whether light or heavy-tailed, [14], Theorem 2.14, should be enough to prove through Theorem 3.2 that the discrete fragmentation process on critical rank-one inhomogeneous random graphs converge to a fragmentation process on the limit.

Concerning coalescence, it seems to me that applying Theorem 3.1 is not as straightforward, since it could require substantial work to prove the additional hypothesis (3.1). For some models where the convergence of a height process is available, as rank-one inhomogeneous random graphs [14], there is some hope to mimic the proof of Section 5.4, but not without substantial additional work, essentially to prove condition (ii) of Lemma 5.13. It could also be feasible for the configuration model with i.i.d. degrees leading to stable graphs; cf. [17].

Concerning dynamical percolation, once the difficulties explained above concerning fragmentation and coalescence will be resolved, it should be straightforward to obtain a result analogous to Theorem 3.7. Also, I believe that a duality result analogous to Proposition 3.8 should hold also in the heavy-tailed setting.

Finally, let us mention that a version of Theorem 3.3 where the convergence of initial data and of the process would be with the same topology should be true by using a stronger topology. One should be able to do this with a distance compatible with the Feller property of the augmented coalescent proved in [8].

APPENDIX A: TOPOLOGIES FOR PROCESSES

Let  $(M, d)$  denote a separable, complete metric space. Let  $\mathcal{F}([0, \infty), M)$  (resp.,  $D([0, \infty), M)$ ) denote the set of functions from  $[0, \infty)$  to  $M$  (resp., càdlàg functions from  $[0, \infty)$  to  $M$ ). For  $\omega_1$  and  $\omega_2$  in  $\mathcal{F}([0, \infty), M)$ , let us define

$$d_{c,k}(\omega_1, \omega_2) := \sup_{t \in [0,k]} d(\omega_1(t), \omega_2(t))$$

and

$$d_c(\omega_1, \omega_2) := \sum_{k \geq 1} 2^{-k} (1 \wedge d_{c,k}(\omega_1, \omega_2)).$$

It is easy to see that  $(\mathcal{F}([0, \infty), M), d_c)$  and  $(D([0, \infty), M), d_c)$  are complete metric spaces (not separable in general), and that a sequence  $\omega^n = (\omega^n(t))_{t \in \mathbb{R}^+}$  converges in this metric space to  $\omega^\infty = (\omega^\infty(t))_{t \in \mathbb{R}^+}$  if and only if for every  $T > 0$ ,

$$\sup_{t \in [0,T]} |\omega^n(t) - \omega^\infty(t)| \xrightarrow{n \rightarrow +\infty} 0,$$

whence the term “topology of compact convergence.” A reference when  $M = \mathbb{R}$ , with applications, is [27], Section V.5, but we shall not use it here.

Now, let us define  $d_S$  the Skorokhod metric as in [27], Section VI.1. For each  $k \in \mathbb{N}^*$ ,  $\omega_1$  and  $\omega_2$  in  $\mathcal{F}([0, \infty), M)$ , let  $d_{S,k}(\omega_1, \omega_2)$  be the infimum of those values  $\delta$  for which there exists grids  $0 = t_0 < t_1 < \dots < t_r$  with  $t_r \geq k$  and  $0 = s_0 < s_1 < \dots < s_r$ , with  $s_r \geq k$  such that  $|t_i - s_i| \leq \delta$  for  $i = 1, \dots, r$  and

$$d(\omega_1(t), \omega_2(s)) \leq \delta \quad \text{if } t_i \leq t < t_{i+1} \text{ and } s_i \leq s < s_{i+1}.$$

Then let

$$d_S(\omega_1, \omega_2) := \sum_{k \geq 1} 2^{-k} (1 \wedge d_{S,k}(\omega_1, \omega_2)).$$

Then  $(D([0, \infty), M), d_S)$  is a separable metric space; cf. [27], Theorem VI.6. Notice that for any  $k$ ,

$$(A.1) \quad d_{S,k}(\omega_1, \omega_2) \leq d_{c,k}(\omega_1, \omega_2).$$

Notably,  $d_S \leq d_c$ , and thus the topology induced by  $d_c$  on  $D([0, \infty), M)$  is finer than Skorokhod's topology induced by  $d_S$ . Notice also that

$$(A.2) \quad d_S(\omega_1, \omega_2) \leq d_{S,k}(\omega_1, \omega_2) + 2^{-k}.$$

Concerning measurability,  $(\mathcal{F}([0, \infty), M), d_c)$  (resp.,  $(D([0, \infty), M), d_c)$ ) will always be equipped with its projection  $\sigma$ -field  $\mathcal{P}_{\mathcal{F}}$  (resp.,  $\mathcal{P}_D$ ), the smallest  $\sigma$ -field making the projections  $\pi_t$  measurable, where  $\pi_t$  maps  $\omega$  to  $\omega(t) \in M$ . Those  $\sigma$ -fields are included in the Borel  $\sigma$ -fields for  $d_c$ . But  $\mathcal{P}_D$  coincides with the Borel  $\sigma$ -field induced by the Skorokhod topology on  $D([0, \infty), M)$ ; cf. [27], Theorem VI.6.

In this article, we prove convergence in distribution of a sequence of processes  $((X_n(t))_{t \geq 0})_{n \geq 1}$  towards  $(X(t))_{t \geq 0}$  by exhibiting couplings showing essentially that the Lévy–Prokhorov distance for  $d_c$  between the distributions of  $X_n$  and  $X$  goes to zero as  $n$  goes to infinity. This implies convergence in distribution for the Skorokhod metric. Everything needed is gathered in the following lemmas. Recall that  $\mathbb{P}^*$  denotes the outer measure associated to  $\mathbb{P}$ .

LEMMA A.1. *Let  $\omega^n, n \in \bar{\mathbb{N}}$  be random variables with values in  $\mathcal{F}([0, \infty), M)$ , defined on the same complete probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Suppose that for every  $n \in \bar{\mathbb{N}}$ ,  $\omega^n \in D([0, \infty), M)$  almost surely, and that*

$$d_c(\omega^n, \omega^\infty) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}^*} 0.$$

*Then  $\omega^\infty$  belongs to  $D([0, \infty), M)$  almost surely.*

PROOF. Since  $d_c(\omega^n, \omega^\infty)$  converges in probability to zero, one may extract a subsequence  $\omega^{n_k}, k \geq 0$  such that  $d_c(\omega^{n_k}, \omega^\infty)$  converges to zero  $\mathbb{P}^*$ -a.s., and hence (since  $\mathbb{P}$  is complete),  $\mathbb{P}$ -a.s. But since almost surely,  $\omega^n$  is càdlàg for any  $n$ , and any limit of a càdlàg sequence for  $d_c$  is càdlàg, we obtain that  $\omega^\infty$  is càdlàg  $\mathbb{P}$ -a.s.  $\square$

LEMMA A.2. *Let  $\omega^n, n \in \bar{\mathbb{N}}$  be random variables with values in  $D([0, \infty), M)$  such that for any  $\varepsilon > 0$ , any  $k \in \mathbb{N}^*$ , there exists  $N \in \bar{\mathbb{N}}$  such that for every  $n \geq N$ , there exists a coupling of  $\omega^n|_{[0,k]}$  and  $\omega^\infty|_{[0,k]}$  such that*

$$\mathbb{P}[d_{S,k}(\omega^n, \omega^\infty) > \varepsilon] \leq \varepsilon.$$

*Then  $\omega^n$  converges in distribution to  $\omega^\infty$ , for the Skorokhod topology.*

PROOF. Let  $F$  be a closed set in  $(D([0, \infty), M), d_S)$ . It is sufficient to prove that

$$\limsup_{n \rightarrow +\infty} \mathbb{P}(\omega^n \in F) \leq \mathbb{P}(\omega^\infty \in F).$$

Let  $\varepsilon > 0$ . Let us define

$$F^\varepsilon := \{x \in D([0, \infty), M) : \exists y \in F, d_S(x, y) \leq \varepsilon\},$$

which is a closed set. We have

$$\mathbb{P}(\omega^n \in F) \leq \mathbb{P}(\omega^\infty \in F^{2\varepsilon}) + \mathbb{P}(d_S(\omega^n, \omega^\infty) > \varepsilon).$$

Now, choose  $k \in \bar{\mathbb{N}}$  such that  $2^{-k} \leq \frac{\varepsilon}{2}$ . Then, using (A.1),

$$\mathbb{P}(\omega^n \in F) \leq \mathbb{P}(\omega^\infty \in F^{2\varepsilon}) + \mathbb{P}\left(d_{S,k}(\omega^n, \omega^\infty) > \frac{\varepsilon}{2}\right).$$

By hypothesis, the last term above is less than  $\varepsilon$  if  $n$  is chosen large enough. Letting  $\varepsilon$  go to zero and using the fact that  $\bigcap_{\varepsilon > 0} F^\varepsilon = F$ , we get the result.  $\square$

APPENDIX B: THE SET OF ISOMETRY CLASSES OF FINITE MEASURED SEMIMETRIC SPACES

Even though there is no *set* of finite measured semimetric spaces in the sense of Zermelo–Frankel set theory,<sup>8</sup>  $\mathcal{C}/\mathcal{R}$  can be considered as a set in the sense that there exists a true set of representatives of elements of  $\mathcal{C}$ . The idea is to consider the set of all finite measured semimetric spaces of a sufficiently large set  $\mathbb{U}$  such that  $\mathbb{U}$  contains a copy of any separable metric space, in the following (quite weak) sense.

**HYPOTHESIS B.1.** *For any separable metric space  $M$ , there exists a map  $\phi$  from  $M$  to  $\mathbb{U}$  such that  $\phi$  is a bijection from  $M$  to  $\phi(M)$ .*

In Definition B.2, we make the (quite standard) choice  $\mathbb{U} = (\mathbb{R}_+)^{\mathbb{N}}$ . Indeed, for any metric space  $(M, d)$  with a dense sequence  $(x_i)_{i \in \mathbb{N}}$ , the following function shows that Hypothesis B.1 is satisfied:

$$\phi : \begin{cases} M \rightarrow \mathbb{U}, \\ x \mapsto (d(x, x_i))_{i \in \mathbb{N}}. \end{cases}$$

**DEFINITION B.2.** Let  $\mathbb{U} = (\mathbb{R}_+)^{\mathbb{N}}$ . Let  $\mathcal{P}$  denote the set of measured semimetric spaces  $X = (X, d, \mu)$  such that:

- $X$  is a subset of  $\mathbb{U}$ ,
- $d$  is a finite semimetric on  $X$ .

We denote by  $\mathcal{M}$  the quotient  $\mathcal{P}/\mathcal{R}$  of  $\mathcal{P}$  by the equivalence relation  $\mathcal{R}$ , where

$$X \mathcal{R} X' \iff d_{\text{GHP}}(X, X') = 0.$$

By abuse of language, we may call  $\mathcal{M}$  the “set of equivalence classes of finite measured semimetric spaces, equipped with the Gromov–Hausdorff–Prokhorov distance  $d_{\text{GHP}}$ .”

$\mathcal{M}$  is a set of representatives of elements of  $\mathcal{C}$ . Indeed, since every separable metric space is isometric to a subspace of  $\mathbb{U}$  equipped with a suitable metric through the function  $\phi$  above, every member of  $\mathcal{C}$  will be at zero  $d_{\text{GHP}}$ -distance from some element of  $\mathcal{P}$ , and even at zero  $d_{\text{GHP}}$ -distance from some compact element of  $\mathcal{P}$ . Thus, for every member  $X$  of the class  $\mathcal{C}$ , there is an element  $[X']$  of  $\mathcal{M}$  such that for any  $X'' \in [X]$ ,  $d_{\text{GHP}}(X, X'') = 0$ . Abusing notation, we shall denote by  $[X]$  the member of  $\mathcal{M}$  whose elements are at zero  $d_{\text{GHP}}$ -distance from  $X$ .

For our purpose, it is in fact not crucial to have Definition B.2, and one could reformulate all the results in this article in terms of sequences of random variables, at the expense of much more heavy statements.

Finally, one may define a set of equivalence classes of  $\mathbb{R}$ -graphs as follows.

**DEFINITION B.3.** Let  $\mathcal{P}^{\text{graph}}$  denote the subset of  $\mathcal{P}$  composed of measured semimetric spaces which are semimetric  $\mathbb{R}$ -graphs. We denote by  $\mathcal{M}^{\text{graph}}$  the quotient  $\mathcal{P}^{\text{graph}}/\mathcal{R}$  of  $\mathcal{P}^{\text{graph}}$  by the equivalence relation  $\mathcal{R}$ , where

$$X \mathcal{R} X' \iff d_{\text{GHP}}(X, X') = 0.$$

---

<sup>8</sup>See, for instance, [16], Remark 7.2.5: the problem is that if this class  $\mathcal{C}$  was a set, one can assign to it a finite semimetric,  $d_{\text{GHP}}$  and a measure  $\mu$ , so that  $(\mathcal{C}, d_{\text{GHP}}, \mu)$  is a member of  $\mathcal{C}$ , leading to Russell’s paradox.

APPENDIX C: COMMUTATION RELATIONS FOR COALESCENCE AND FRAGMENTATION

LEMMA C.1. *Let  $(X, d)$  be a semimetric space. For any  $A, B \subset X^2$ ,  $\text{Coal}_0(X, A \cup B) = \text{Coal}_0(\text{Coal}_0(X, A), B)$ .*

PROOF. Let  $\text{Eq}(A)$  denote the equivalence relation generated by  $A$ . Since  $\text{Eq}(A \cup B) = \text{Eq}(\text{Eq}(A) \cup \text{Eq}(B))$ , it is sufficient to prove the lemma when  $A$  and  $B$  are equivalence relations. So suppose that  $A$  and  $B$  are equivalence relations. Let  $x$  and  $y$  be in  $X$ . Then

$$d_{A \cup B, 0}(x, y) = \inf \sum_{i=1}^k d(p_i, q_i),$$

where the infimum is over all  $k \in \mathbb{N}^*$ ,  $p_1, q_1, \dots, p_k, q_k$  such that  $p_0 = x$ ,  $q_k = y$  and  $(q_i, p_{i+1}) \in \text{Eq}(A \cup B)$  for any  $i$  in  $\{1, \dots, k - 1\}$ . Now, if  $(q_i, p_{i+1}) \in \text{Eq}(A \cup B)$ , there exists  $r_1, s_1, \dots, r_l, s_l, r_{l+1}$  such that  $r_1 = q_i$ ,  $s_l = p_{i+1}$  and

$$\forall j \in \{1, \dots, l\}, \quad (r_j, s_j) \in A \quad \text{and} \quad (s_j, r_{j+1}) \in B.$$

Then

$$d_X(p_i, q_i) = d(p_i, q_i) + \sum_{j=1}^l d_{A, 0}(r_j, s_j).$$

And we obtain, denoting by  $\tilde{d}$  the distance of  $\text{Coal}_0(\text{Coal}_0(X, A), B)$ ,

$$d_{A \cup B, 0}(x, y) \geq \tilde{d}(x, y).$$

In the other direction, let  $k \in \mathbb{N}^*$ ,  $p_1, q_1, \dots, p_k, q_k$  and  $\varepsilon > 0$  be such that

$$\tilde{d}(x, y) \leq \sum_{i=1}^k d_{A, 0}(p_i, q_i) + \varepsilon$$

with  $p_0 = x$ ,  $q_k = y$  and  $(q_i, p_{i+1}) \in B$  for any  $i$  in  $\{1, \dots, k - 1\}$ . Then let  $h_1, \dots, h_k$  and  $p_{i,j}, q_{i,j}$ ,  $i = 1, \dots, k$ ,  $j = 1, \dots, h_i$  be such that for any  $i \in \{1, \dots, k\}$ ,

$$d_{A, 0}(p_i, q_i) \leq \sum_{j=1}^{h_i} d(p_{i,j}, q_{i,j}) + \frac{\varepsilon}{k},$$

with  $p_{i,1} = p_i$ ,  $q_{i,h_i} = q_i$  and  $(q_{i,j}, p_{i,j+1}) \in A$ . Then we get

$$\tilde{d}(x, y) \leq \sum_{i=1}^k \sum_{j=1}^{h_i} d(p_{i,j}, q_{i,j}) + 2\varepsilon$$

with  $p_{1,1} = x$ ,  $q_{k,h_k} = y$  and  $(q_{i,j}, p_{i,j+1}) \in A \cup B$  with the convention  $p_{i,h_i+1} = p_{i+1,1}$ . Minimizing on  $p$  and  $q$ , we get, for any  $\varepsilon > 0$ :

$$\tilde{d}(x, y) \leq d_{A \cup B, 0}(x, y) + 2\varepsilon$$

which gives the result.  $\square$

LEMMA C.2. *Let  $(X, d)$  be a semimetric space. For any multisets  $A, B$  of elements of  $X^2$  and any  $\delta > 0$ ,  $\text{Coal}_\delta(X, A \sqcup B) = \text{Coal}_\delta(\text{Coal}_\delta(X, A), B)$ .*

PROOF. For a multiset  $A$  of elements of  $X^2$ , let us denote by  $I_A$  the multiset of intervals added when  $\delta > 0$ :

$$I_A := \bigsqcup_{(x,x') \in A} [a_{x,x'}, b_{x,x'}],$$

and let  $\tilde{A}$  denote the set of points to be identified:

$$\tilde{A} := \{(x, a_{x,x'}) : (x, x') \in A\} \cup \{(x', b_{x,x'}) : (x, x') \in A\}.$$

Then

$$\text{Coal}_\delta(X, A) = \text{Coal}_0(X \sqcup I_A, \tilde{A}).$$

Notice also that  $I_{A \sqcup B} = I_A \sqcup I_B$  and  $\widetilde{A \sqcup B} = \tilde{A} \sqcup \tilde{B}$ . Now,

$$\begin{aligned} \text{Coal}_\delta(X, A \sqcup B) &= \text{Coal}_0(X \sqcup I_{A \sqcup B}, \widetilde{A \sqcup B}) \\ &= \text{Coal}_0(X \sqcup I_A \sqcup I_B, \tilde{A} \sqcup \tilde{B}) \\ &= \text{Coal}_0(\text{Coal}_0(X \sqcup I_A \sqcup I_B, \tilde{A}), \tilde{B}) \\ &= \text{Coal}_0(\text{Coal}_0(X \sqcup I_A, \tilde{A}) \sqcup I_B, \tilde{B}) \\ &= \text{Coal}_\delta(\text{Coal}_0(X \sqcup I_A, \tilde{A}), B) \\ &= \text{Coal}_\delta(\text{Coal}_\delta(X, A), B), \end{aligned}$$

where we used Lemma C.1 at the third line.  $\square$

Recall the following notation: if  $(X, d)$  is a semimetric space and  $A \subset X$ , we let

$$A^d := \{x \in X \text{ s.t. } \exists y \in A, d(x, y) = 0\}.$$

LEMMA C.3. *Let  $X$  be a length-space. Then, for any  $A$  and  $B \subset X$ ,  $\text{Frag}(\text{Frag}(X, A), B) = \text{Frag}(X, A \cup B)$ .*

PROOF. Notice first that if  $x \in X \setminus A^d$  and  $y \in B^d \setminus A^d$ ,

$$d_A^{\text{Frag}}(x, y) = 0 \iff d(x, y) = 0.$$

Thus,

$$(A \cup B)^d = A^d \cup (B \setminus A)^d_A^{\text{Frag}}.$$

Now, let  $x$  and  $y$  belong to  $X$  and  $\gamma$  a path from  $x$  to  $y$ , indexed by  $[0, 1]$ , such that

$$\gamma \cap A^d = \emptyset,$$

then we claim that

$$(C.1) \quad \ell_{\text{Frag}(X,A)}(\gamma) = \ell_X(\gamma).$$

Indeed, trivially,

$$\ell_{\text{Frag}(X,A)}(\gamma) \geq \ell_X(\gamma).$$

On the other hand, if  $0 = t_1 \leq \dots \leq t_n = 1$  is a subdivision of  $[0, 1]$ ,  $\gamma|_{[t_i, t_{i+1}]}$  is a path from  $\gamma(t_i)$  to  $\gamma(t_{i+1})$  in  $X \setminus A^d$ . Thus,

$$\begin{aligned} \sum_{i=1}^{n-1} d_A^{\text{Frag}}(\gamma(t_i), \gamma(t_{i+1})) &\leq \sum_{i=1}^{n-1} \ell_X(\gamma|_{[t_i, t_{i+1}]}) \\ &= \ell_X(\gamma) \end{aligned}$$

which shows (C.1). Finally, denoting by  $\tilde{d}$  the distance on  $\text{Frag}(\text{Frag}(X, A), B)$ ,

$$\begin{aligned} d_{A \cup B}^{\text{Frag}}(x, y) &= \inf_{\substack{\gamma: x \rightarrow y \\ \gamma \cap (A \cup B)^d = \emptyset}} \ell_X(\gamma) \\ &= \inf_{\substack{\gamma: x \rightarrow y \\ \gamma \cap (A \cup B)^d = \emptyset}} \ell_{\text{Frag}(X, A)}(\gamma) \\ &= \inf_{\substack{\gamma: x \rightarrow y \text{ in } \text{Frag}(X, A) \\ \gamma \cap (B \setminus A)^{d_A^{\text{Frag}}} = \emptyset}} \ell_{\text{Frag}(X, A)}(\gamma) \\ &= \tilde{d}(x, y). \end{aligned} \quad \square$$

Hereafter, we say that a path  $\gamma$  in  $\text{Coal}_0(X, A)$  takes a shortcut  $(a, b)$  in  $A \subset X^2$  if  $(a, b) \in A$  and  $\gamma \cap \{a, b\} \neq \emptyset$ .

LEMMA C.4. *Let  $(X, d)$  be a semimetric length space and  $A \subset X^2$  an equivalence relation. Suppose that for any  $(x, y) \in X^2$ , every simple rectifiable path in  $\text{Coal}_0(X, A)$  from  $x$  to  $y$  takes only a finite number of shortcuts in  $A$ . Let  $B^d$  denote the set*

$$\{x \in X : \exists y \in B, d(x, y) = 0\}.$$

Then, for any  $B \subset X$  such that  $B^d \cap \{x \in X : \exists y \in X, (x, y) \text{ or } (y, x) \in A\} = \emptyset$ ,

$$\text{Coal}_0(\text{Frag}(X, B), A) = \text{Frag}(\text{Coal}_0(X, A), B).$$

PROOF. Let  $\ell_X(\gamma)$  denote the length of a path  $\gamma$  in  $X$ . Let  $d^{\text{fragcoal}}$  (resp.,  $d^{\text{coalfrag}}$ , resp.,  $d^{\text{frag}}$ ) denote the distance of  $\text{Frag}(\text{Coal}_0(X, A), B)$  (resp.,  $\text{Coal}_0(\text{Frag}(X, B), A)$ , resp.,  $\text{Frag}(X, B)$ ) on  $X \setminus B^d$ . We want to show that  $d^{\text{fragcoal}} = d^{\text{coalfrag}}$ . First, it is always true that  $d^{\text{fragcoal}} \leq d^{\text{coalfrag}}$ . Indeed, let  $\{p_i\}$  and  $\{q_i\}$ ,  $i = 0, \dots, k$  be such that  $(q_i, p_{i+1}) \in A$  for all  $i = 0, \dots, k - 1$  and  $p_0 = x, q_k = y$ . Then the concatenation of  $(k + 1)$  paths  $\gamma_i$  in  $X$ ,  $i = 0, \dots, k$  such that  $\gamma_i$  goes from  $p_i$  to  $q_i$  and each path avoids  $B^d$  gives a path in  $\text{Coal}_0(X, A)$  from  $x$  to  $y$  avoiding  $B^d$ . Thus, for any  $x$  and  $y$  in  $X \setminus B^d$ ,

$$\begin{aligned} d^{\text{fragcoal}}(x, y) &\leq \inf_{\substack{k, \{p_i\}, \{q_i\} \\ \gamma_i: p_i \rightarrow q_i \\ \gamma_i \cap B^d = \emptyset}} \left( \sum_{i=1}^k \ell_X(\gamma_i) \right) \\ &= \inf_{k, \{p_i\}, \{q_i\}} \left( \sum_{i=1}^k d^{\text{frag}}(p_i, q_i) \right) \\ &= d^{\text{coalfrag}}(x, y). \end{aligned}$$

Let us show now that  $d^{\text{coalfrag}} \leq d^{\text{fragcoal}}$ . Let  $x$  and  $y$  be in  $X \setminus B^d$  and let  $\gamma$  be a rectifiable simple path from  $x$  to  $y$  in  $\text{Coal}_0(X, A)$  such that  $\gamma \cap B^d = \emptyset$ . Then  $\gamma$  takes only a finite number of shortcuts in  $A$ . Thus, there exists  $\{p_i\}$  and  $\{q_i\}$ ,  $i = 0, \dots, k$  with  $p_0 = x$  and  $q_k = y$  and paths  $\gamma_i$ ,  $i = 0, \dots, k$  such that  $(q_i, p_{i+1}) \in A$  and  $\gamma_i$  is a path from  $p_i$  to  $q_i$  in  $X$  and  $\gamma$  is the concatenation of  $\gamma_0, \dots, \gamma_k$ . Thus,  $\gamma_i \cap B^d = \emptyset$  for any  $i$  and

$$\begin{aligned} \ell_{\text{Coal}_0(X, A)}(\gamma) &= \sum_{i=1}^k \ell_X(\gamma_i) \\ &\geq \sum_{i=1}^k d^{\text{frag}}(p_i, q_i) \\ &\geq d^{\text{coalfrag}}(x, y). \end{aligned}$$



Taking the infimum over rectifiable simple paths  $\gamma$  from  $x$  to  $y$  in  $\text{Coal}_0(X, A)$  such that  $\gamma \cap B^d = \emptyset$  gives that  $d^{\text{coalfrag}} \leq d^{\text{fragcoal}}$ .  $\square$

APPENDIX D:  $\mathbb{R}$ -GRAPHS AS  $\mathbb{R}$ -TREES WITH SHORTCUTS

Here, we sketch the proof of Lemma 2.19. Let us start with the “if” direction. By induction, it is sufficient to show that if  $(X, d)$  is an  $\mathbb{R}$ -graph and  $(x, y) \in X^2$ , then the quotient metric space obtained from  $\text{Coal}_0(X, \{(x, y)\})$  is isomorphic to an  $\mathbb{R}$ -graph. Notice that  $\text{Coal}_0(X, \{(x, y)\})$  is obviously a totally bounded and finite semimetric space. Let  $R > 0$  be such that for any  $x \in X$ ,  $(B_R(x), d|_{B_R(x)})$  is an  $\mathbb{R}$ -tree, where  $B_R(x)$  is the ball of radius  $R$  and center  $x$  in  $X$ . Then let  $R' := \min\{R/2, d(x, y)/5\}$ ,  $d'$  denote the distance on  $\text{Coal}_0(X, \{(x, y)\})$  and  $B'_{R'}(z)$  be the ball of radius  $R'$  and center  $z$  in  $(X, d')$ . If  $R' = 0$ , then clearly  $\text{Coal}_0(X, \{(x, y)\})$  is equal to  $X$ , so let us suppose that  $R' \neq 0$ . The reader can check that for any  $z \in X$ ,  $(B'_{R'}(z), d'|_{B'_{R'}(z)})$  is a totally bounded acyclic geodesic finite semimetric space, and the only pair of points at  $d'$ -distance zero is  $(x, y)$ . Its quotient metric space is thus an  $\mathbb{R}$ -graph.

Now, let us look at the “only if” direction. Let  $(X, d)$  be an  $\mathbb{R}$ -graph which is not an  $\mathbb{R}$ -tree. The set of branchpoints in an  $\mathbb{R}$ -tree is at most countable. Thus, one may find a point  $x \in \text{core}(X)$  such that  $x$  belongs to a cycle, and  $x$  is of degree 2 in  $X$ . Let  $Y_1$  and  $Y_2$  denote the two components of  $B_R(x) \setminus \{x\}$  (with  $R > 0$  small enough). Then let  $(X', d')$  denote the completion of  $\text{Frag}(X, \{x\})$ . It is shown in [4], Section 7.1, that  $X'$  adds exactly two points  $x_{(1)}$  and  $x_{(2)}$  to  $\text{Frag}(X, \{x\})$  and  $d'$  can be described as follows:

- If  $y, z \notin \{x_{(1)}, x_{(2)}\}$ , then  $d'(y, z)$  is the minimal length of a path from  $y$  to  $z$  in  $X$  not visiting  $x$ .
- If  $y \neq x_{(2)}$ , then  $d'(x_{(1)}, y)$  is the minimal length of an injective path from  $x$  to  $y$  in  $X$  which takes its values in the component  $Y_1$  on some small initial interval, and similarly for  $d'(x_{(2)}, z)$  with  $z \neq x_{(1)}$ .
- $d'(x_{(1)}, x_{(2)})$  is the minimal length of a cycle passing through  $x$ .

The reader can check that  $X'$  is an  $\mathbb{R}$ -graph, with a surplus strictly smaller than  $X$ , and that  $X$  is isomorphic to the quotient metric space associated to  $\text{Coal}(X', \{x_{(1)}, x_{(2)}\})$ . Thus, one may conclude by induction on the surplus of  $X$ .

APPENDIX E: TOTAL VARIATION DISTANCE BETWEEN POISSON RANDOM MEASURES

LEMMA E.1. *Let  $\mathcal{P}(\mu)$  denote the distribution of a Poisson random measure with a finite intensity measure  $\mu$  on a measurable space  $(E, \mathcal{E})$ . If  $\nu$  is another finite measure on  $E$ , then*

$$\|\mathcal{P}(\mu) - \mathcal{P}(\nu)\| \leq 2\|\mu - \nu\|.$$

PROOF. Suppose first that  $\mu(E) = 0$ . Then

$$\|\mathcal{P}(\mu) - \mathcal{P}(\nu)\| = \|\delta_\emptyset - \mathcal{P}(\nu)\| = 1 - e^{-\nu(E)} \leq \nu(E) \leq 2\|\mu - \nu\|,$$

and the lemma is proved. So suppose now that  $\mu(E)$  and  $\nu(E)$  are nonzero. Without loss of generality, suppose that  $\mu(E) \leq \nu(E)$ . Let  $\pi$  denote an optimal coupling between  $\frac{\mu}{\mu(E)}$  and  $\frac{\nu}{\nu(E)}$ . Let  $N$  (resp.,  $N'$ ) be a Poisson random variable with parameter  $\mu(E)$  (resp.,  $\nu(E)$ ) coupled in an optimal way, that is,

$$\mathbb{P}(N \neq N') = \|\mathcal{P}(\mu(E)) - \mathcal{P}(\nu(E))\|.$$

It is easy to see, for instance, approximating the Poisson distribution by the binomial distribution, that

$$\|\mathcal{P}(\mu(E)) - \mathcal{P}(v(E))\| \leq |\mu(E) - v(E)|.$$

Next, let  $(X_i, Y_i)_{i \geq 1}$  be a sequence of i.i.d random variables of distribution  $\pi$ , independent from  $(N, N')$ . Then  $(\sum_{i=1}^N \delta_{X_i}, \sum_{j=1}^{N'} \delta_{Y_j})$  is a coupling of  $\mathcal{P}(\mu)$  and  $\mathcal{P}(v)$ . Notice also that  $\sum_{i=1}^N \mathbf{1}_{X_i \neq Y_i}$  has Poisson distribution with parameter

$$\mu(E)\mathbb{P}(X_1 \neq Y_1) = \mu(E) \left\| \frac{\mu}{\mu(E)} - \frac{v}{v(E)} \right\|.$$

Thus,

$$\begin{aligned} \|\mathcal{P}(\mu) - \mathcal{P}(v)\| &\leq \mathbb{P}\left(\sum_{i=1}^N \delta_{X_i} \neq \sum_{j=1}^{N'} \delta_{Y_j}\right) \\ &\leq \mathbb{P}(N \neq N') + \mathbb{P}(\exists i \in \{1, \dots, N\}, X_i \neq Y_i) \\ &\leq |\mu(E) - v(E)| + 1 - e^{-\mu(E)\|\frac{\mu}{\mu(E)} - \frac{v}{v(E)}\|} \\ &\leq \|\mu - v\| + \mu(E) \left\| \frac{\mu}{\mu(E)} - \frac{v}{v(E)} \right\|. \end{aligned}$$

Now, for any  $A \in \mathcal{E}$ , since  $\mu(E) \leq v(E)$ ,

$$\frac{v(A)}{v(E)} - \frac{\mu(A)}{\mu(E)} \leq \frac{v(A) - \mu(A)}{\mu(E)} \leq \frac{\|\mu - v\|}{\mu(E)}.$$

Thus,

$$\left\| \frac{\mu}{\mu(E)} - \frac{v}{v(E)} \right\| \leq \frac{\|\mu - v\|}{\mu(E)}$$

which gives the result.  $\square$

**Acknowledgments.** I want to thank Christophe Leuridan for helpful discussions and three very conscientious referees for their careful reading of the manuscript, which led to a substantial improvement of this article.

## REFERENCES

- [1] ABRAHAM, R., DELMAS, J.-F. and HOSCHEIT, P. (2013). A note on the Gromov–Hausdorff–Prokhorov distance between (locally) compact metric measure spaces. *Electron. J. Probab.* **18** no. 14, 21. [MR3035742 https://doi.org/10.1214/EJP.v18-2116](https://doi.org/10.1214/EJP.v18-2116)
- [2] ADDARIO-BERRY, L., BROUTIN, N. and GOLDSCHMIDT, C. (2010). Critical random graphs: Limiting constructions and distributional properties. *Electron. J. Probab.* **15** 741–775. [MR2650781 https://doi.org/10.1214/EJP.v15-772](https://doi.org/10.1214/EJP.v15-772)
- [3] ADDARIO-BERRY, L., BROUTIN, N. and GOLDSCHMIDT, C. (2012). The continuum limit of critical random graphs. *Probab. Theory Related Fields* **152** 367–406. [MR2892951 https://doi.org/10.1007/s00440-010-0325-4](https://doi.org/10.1007/s00440-010-0325-4)
- [4] ADDARIO-BERRY, L., BROUTIN, N., GOLDSCHMIDT, C. and MIERMONT, G. (2017). The scaling limit of the minimum spanning tree of the complete graph. *Ann. Probab.* **45** 3075–3144. [MR3706739 https://doi.org/10.1214/16-AOP1132](https://doi.org/10.1214/16-AOP1132)
- [5] ALDOUS, D. (1997). Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.* **25** 812–854. [MR1434128 https://doi.org/10.1214/aop/1024404421](https://doi.org/10.1214/aop/1024404421)
- [6] ALDOUS, D. and PITMAN, J. (1998). The standard additive coalescent. *Ann. Probab.* **26** 1703–1726. [MR1675063 https://doi.org/10.1214/aop/1022855879](https://doi.org/10.1214/aop/1022855879)

- [7] BHAMIDI, S., BROUTIN, N., SEN, S. and WANG, X. (2014). Scaling limits of random graph models at criticality: Universality and the basin of attraction of the Erdős–Rényi random graph. ArXiv e-prints.
- [8] BHAMIDI, S., BUDHIRAJA, A. and WANG, X. (2014). The augmented multiplicative coalescent, bounded size rules and critical dynamics of random graphs. *Probab. Theory Related Fields* **160** 733–796. MR3278920 <https://doi.org/10.1007/s00440-013-0540-x>
- [9] BHAMIDI, S., DHARA, S., VAN DER HOFSTAD, R. and SEN, S. (2020). Universality for critical heavy-tailed network models: Metric structure of maximal components. *Electron. J. Probab.* **25** Paper No. 47, 57. MR4092766 <https://doi.org/10.1214/19-ejp408>
- [10] BHAMIDI, S., SEN, S. and WANG, X. (2017). Continuum limit of critical inhomogeneous random graphs. *Probab. Theory Related Fields* **169** 565–641. MR3704776 <https://doi.org/10.1007/s00440-016-0737-x>
- [11] BHAMIDI, S., VAN DER HOFSTAD, R. and SEN, S. (2018). The multiplicative coalescent, inhomogeneous continuum random trees, and new universality classes for critical random graphs. *Probab. Theory Related Fields* **170** 387–474. MR3748328 <https://doi.org/10.1007/s00440-017-0760-6>
- [12] BOLLOBÁS, B. (1985). *Random Graphs*. Academic Press, London. MR0809996
- [13] BRAUNSTEIN, L. A., WU, Z., CHEN, Y., BULDYREV, S. V., KALISKY, T., SREENIVASAN, S., COHEN, R., LÓPEZ, E., HAVLIN, S. et al. (2007). Optimal path and minimal spanning trees in random weighted networks. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* **17** 2215–2255. MR2349740 <https://doi.org/10.1142/S0218127407018361>
- [14] BROUTIN, N., DUQUESNE, T. and WANG, M. (2018). Limits of multiplicative inhomogeneous random graphs and Lévy trees: The continuum graphs. ArXiv e-prints. Available at arXiv:1804.05871.
- [15] BROUTIN, N. and MARCKERT, J.-F. (2016). A new encoding of coalescent processes: Applications to the additive and multiplicative cases. *Probab. Theory Related Fields* **166** 515–552. MR3547745 <https://doi.org/10.1007/s00440-015-0665-1>
- [16] BURAGO, D., BURAGO, Y. and IVANOV, S. (2001). *A Course in Metric Geometry. Graduate Studies in Mathematics* **33**. Amer. Math. Soc., Providence, RI. MR1835418 <https://doi.org/10.1090/gsm/033>
- [17] CONCHON-KERJAN, G. and GOLDSCHMIDT, C. (2020). The stable graph: The metric space scaling limit of a critical random graph with i.i.d. power-law degrees. ArXiv e-prints. Available at arXiv:2002.04954.
- [18] DHARA, S., VAN DER HOFSTAD, R., VAN LEEUWAARDEN, J. S. H. and SEN, S. (2017). Critical window for the configuration model: Finite third moment degrees. *Electron. J. Probab.* **22** Paper No. 16, 33. MR3622886 <https://doi.org/10.1214/17-EJP29>
- [19] DHARA, S., VAN DER HOFSTAD, R., VAN LEEUWAARDEN, J. S. H. and SEN, S. (2020). Heavy-tailed configuration models at criticality. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 1515–1558. MR4116701 <https://doi.org/10.1214/19-AIHP980>
- [20] ERDŐS, P. and RÉNYI, A. (1960). On the evolution of random graphs. *Magy. Tud. Akad. Mat. Kut. Intéz. Közl.* **5** 17–61. MR0125031
- [21] EVANS, S. N., PITMAN, J. and WINTER, A. (2006). Rayleigh processes, real trees, and root growth with re-grafting. *Probab. Theory Related Fields* **134** 81–126. MR2221786 <https://doi.org/10.1007/s00440-004-0411-6>
- [22] GARBAN, C., PETE, G. and SCHRAMM, O. (2018). The scaling limits of near-critical and dynamical percolation. *J. Eur. Math. Soc. (JEMS)* **20** 1195–1268. MR3790067 <https://doi.org/10.4171/JEMS/786>
- [23] HÄGGSTRÖM, O., PERES, Y. and STEIF, J. E. (1997). Dynamical percolation. *Ann. Inst. Henri Poincaré Probab. Stat.* **33** 497–528. MR1465800 [https://doi.org/10.1016/S0246-0203\(97\)80103-3](https://doi.org/10.1016/S0246-0203(97)80103-3)
- [24] JANSON, S., KNUTH, D. E., ŁUCZAK, T. and PITTEL, B. (1993). The birth of the giant component. *Random Structures Algorithms* **4** 231–358. With an introduction by the editors. MR1220220 <https://doi.org/10.1002/rsa.3240040303>
- [25] KALLENBERG, O. (2017). *Random Measures, Theory and Applications. Probability Theory and Stochastic Modelling* **77**. Springer, Cham. MR3642325 <https://doi.org/10.1007/978-3-319-41598-7>
- [26] MIERMONT, G. (2009). Tessellations of random maps of arbitrary genus. *Ann. Sci. Éc. Norm. Supér. (4)* **42** 725–781. MR2571957 <https://doi.org/10.24033/asens.2108>
- [27] POLLARD, D. (1984). *Convergence of Stochastic Processes. Springer Series in Statistics*. Springer, New York. MR0762984 <https://doi.org/10.1007/978-1-4612-5254-2>
- [28] ROBERTS, M. I. and ŞENGÜL, B. (2018). Exceptional times of the critical dynamical Erdős–Rényi graph. *Ann. Appl. Probab.* **28** 2275–2308. MR3843829 <https://doi.org/10.1214/17-AAP1357>
- [29] VILLANI, C. (2009). *Optimal Transport. Old and New*. Springer, Berlin.