Exploration on and of networks

Güldas, H.

CHAPTER 2

Mixing times of random walks on dynamic configuration models

This chapter is based on a joint article with Luca Avena, Remco van der Hofstad and Frank den Hollander [12].

Abstract

The mixing time of a random walk, with or without backtracking, on a random graph generated according to the configuration model on n vertices, is known to be of order $\log n$. In this paper we investigate what happens when the random graph becomes *dynamic*, namely, at each unit of time a fraction α_n of the edges is randomly rewired. Under mild conditions on the degree sequence, guaranteeing that the graph is locally tree-like, we show that for every $\varepsilon \in (0,1)$ the ε-mixing time of random walk without backtracking grows like $\sqrt{2\log(1/\varepsilon)/\log(1/(1-\alpha_n))}$ as $n \to \infty$, provided that $\lim_{n\to\infty} \alpha_n (\log n)^2 = \infty$. The latter condition corresponds to a regime of fast enough graph dynamics. Our proof is based on a randomised stopping time argument, in combination with coupling techniques and combinatorial estimates. The stopping time of interest is the first time that the walk moves along an edge that was rewired before, which turns out to be close to a strong stationary time.

§2.1 Introduction and main result

§2.1.1 Motivation and background

The *mixing time* of a Markov chain is the time it needs to approach its stationary distribution. For random walks on *finite graphs*, the characterisation of the mixing time has been the subject of intensive study. One of the main motivations is the fact that the mixing time gives information about the geometry of the graph (see the books by Aldous and Fill [4] and by Levin, Peres and Wilmer [65] for an overview and for applications). Typically, the random walk is assumed to be ‘simple’, meaning that steps are along edges and are drawn uniformly at random from a set of allowed edges, e.g. with or without backtracking.

In the last decade, much attention has been devoted to the analysis of mixing times for random walks on *finite random graphs*. Random graphs are used as models for real-world networks. Three main models have been in the focus of attention: (1) the Erdős-Rényi random graph (Benjamini, Kozma and Wormald [18], Ding, Lubetzky and Peres [37], Fountoulakis and Reed [50], Nachmias and Peres [75]); (2) the configuration model (Ben-Hamou and Salez [16], Berestycki, Lubetzky, Peres and Sly [21], Bordenave, Caputo and Salez [28], Lubetzky and Sly [67]); (3) percolation clusters (Benjamini and Mossel [19]).

Many real-world networks are dynamic in nature. It is therefore natural to study random walks on *dynamic finite random graphs*. This line of research was initiated recently by Peres, Stauffer and Steif [83] and by Peres, Sousi and Steif [82], who characterised the mixing time of a simple random walk on a dynamical percolation cluster on a d-dimensional discrete torus, in various regimes. The goal of the present paper is to study the mixing time of a random walk *without backtracking* on a dynamic version of the configuration model.

The *static* configuration model is a random graph with a prescribed degree sequence (possibly random). It is popular because of its mathematical tractability and its flexibility in modeling real-world networks (see van der Hofstad [93, Chapter 7] for an overview). For random walk on the static configuration model, with or without backtracking, the asymptotics of the associated mixing time, and related properties such as the presence of the so-called cutoff phenomenon, were derived recently by Berestycki, Lubetzky, Peres and Sly [21], and by Ben-Hamou and Salez [16]. In particular, under mild assumptions on the degree sequence, guaranteeing that the graph is an *expander* with high probability, the mixing time was shown to be of order $\log n$, with n the number of vertices.

In the present paper we consider a *discrete-time dynamic* version of the configuration model, where at each unit of time a fraction α_n of the edges is sampled and rewired uniformly at random. [A different dynamic version of the configuration model was considered in the context of *graph sampling*. See Greenhill [54] and references therein.] Our dynamics *preserves the degrees* of the vertices. Consequently, when considering a random walk on this dynamic configuration model, its *stationary distribution remains constant over time* and the analysis of its mixing time is a well-posed question. It is natural to expect that, due to the graph dynamics, the random walk *mixes faster* than the $\log n$ order known for the static model. In our main theorem we will make this precise under mild assumptions on the prescribed degree sequence stated in Condition 2.1.2 and Remark 2.1.3 below. By requiring that $\lim_{n\to\infty} \alpha_n (\log n)^2 = \infty$, which corresponds to a regime of fast enough graph dynamics, we find in Theorem 2.1.7 below that for every $\varepsilon \in (0,1)$ the ε-mixing time for random walk *without backtracking* grows like $\sqrt{2\log(1/\varepsilon)/\log(1/(1-\alpha_n))}$ as $n \to \infty$, with high probability in the sense of Definition 2.1.5 below. Note that this mixing time is $o(\log n)$, so that the dynamics indeed speeds up the mixing.

§2.1.2 Model

We start by defining the model and setting up the notation. The set of vertices is denoted by V and the degree of a vertex $v \in V$ by $d(v)$. Each vertex $v \in V$ is thought of as being incident to $d(v)$ *half-edges* (see Fig. 2.1). We write H for the set of half-edges, and assume that each half-edge is associated to a vertex via incidence. We denote by $v(x) \in V$ the vertex to which $x \in H$ is incident and by $H(v) := \{x \in H\colon v(x) = v\} \subset H$ the set of half-edges incident to $v \in V$. If $x, y \in H(v)$ with $x \neq y$, then we write $x \sim y$ and say that x and y are siblings of each other. The degree of a half-edge $x \in H$ is defined as

$$\deg(x) := d(v(x)) - 1. \tag{2.1}$$

We consider graphs on n vertices, i.e., $|V| = n$, with m edges, so that $|H| = \sum_{v\in V} \deg(v) = 2m =: \ell$.

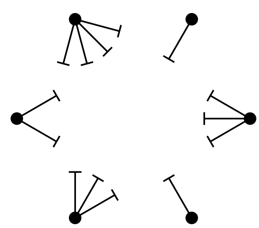

Figure 2.1: Vertices with half-edges.

The *edges* of the graph will be given by a *configuration* that is a *pairing of half-edges*. We denote by $\eta(x)$ the half-edge paired to $x \in H$ in the configuration η. A configuration η will be viewed as a bijection of H without fixed points and with the property that $\eta(\eta(x)) = x$ for all $x \in H$ (also called an involution). With a slight abuse of notation, we will use the same symbol η to denote the set of pairs of half-edges in η, so $\{x, y\} \in \eta$ means that $\eta(x) = y$ and $\eta(y) = x$. Each pair of half-edges in η will also be called an edge. The set of all configurations on H will be denoted by $Conf_H$.

We note that each configuration gives rise to a graph that may contain self-loops (edges having the same vertex on both ends) or multiple edges (between the same pair of vertices). On the other hand, a graph can be obtained via several distinct configurations.

We will consider asymptotic statements in the sense of $|V| = n \to \infty$. Thus, quantities like $V, H, d, \deg$ and ℓ all depend on n. In order to lighten the notation, we often suppress n from the notation.

Configuration model

We recall the definition of the configuration model, phrased in our notation. Inspired by Bender and Canfield [17], the configuration model was introduced by Bollobás [24] to study the number of regular graphs of a given size (see also Bollobás [25]). Molloy and Reed [72], [73] introduced the configuration model with general prescribed degrees.

The configuration model on V with degree sequence $(d(v))_{v\in V}$ is the uniform distribution on $Conf_H$. We sometimes write $d_n = (d(v))_{v\in V}$ when we wish to stress the n-dependence of the degree sequence. Identify H with the set

$$[1, \ell] \coloneqq \{1, \dots, \ell\}.$$

A sample η from the configuration model can be generated by the following *sampling algorithm*:

1. Initialize $U = H, \eta = \varnothing$, where U denotes the set of unpaired half-edges.
2. Pick a half-edge, say x, uniformly at random from $U \setminus \{\min U\}$.
3. Update $\eta \to \eta \cup \{\{x, \min U\}\}$ and $U \to U \setminus \{x, \min U\}$.
4. If $U \neq \varnothing$, then continue from step 2. Else return η.

The resulting configuration η gives rise to a graph on V with degree sequence $(d(v))_{v\in V}$.

Remark 2.1.1. Note that in the above algorithm two half-edges that belong to the same vertex can be paired, which creates a self-loop, or two half-edges that belong to vertices that already have an edge between them can be paired, which creates multiple edges. However, if the degrees are not too large (as in Condition 2.1.2 below), then as $n \to \infty$ the number of self-loops and the number of multiple edges converge to two independent Poisson random variables (see Janson [58], [59], Angel, van der Hofstad and Holmgren [10]). Consequently, convergence in probability for the configuration model implies convergence in probability for the configuration model conditioned on being simple.

Let U_n be uniformly distributed on $[1, n]$. Then

$$D_n = d(U_n) \tag{2.2}$$

is the degree of a random vertex on the graph of size n. Write $\mathbb{P}_n$ to denote the law of D_n. Throughout the sequel, we impose the following mild regularity conditions on the degree sequence:

Condition 2.1.2. ***(Regularity of degrees)***

(R1) *Let* $\ell = |H|$*. Then* ℓ *is even and of order* n*, i.e.,* $\ell = \Theta(n)$ *as* $n \to \infty$*.*

(R2) *Let*

$$\nu_n := \frac{\sum_{z\in H}\deg(z)}{\ell} = \frac{\sum_{v\in V} d(v)[d(v)-1]}{\sum_{v\in V} d(v)} = \frac{\mathbb{E}_n(D_n(D_n-1))}{\mathbb{E}_n(D_n)} \tag{2.3}$$

denote the expected degree of a uniformly chosen half-edge. Then $\limsup_{n\to\infty}\nu_n < \infty$*.*

(R3) $\mathbb{P}_n(D_n \geq 2) = 1$ *for all* $n \in \mathbb{N}$*.*

Remark 2.1.3. Conditions (R1) and (R2) are minimal requirements to guarantee that the graph is locally tree-like (in the sense of Lemma 2.4.2 below). They also ensure that the probability of the graph being simple has a strictly positive limit. Conditioned on being simple, the configuration model generates a random graph that is uniformly distributed among all the simple graphs with the given degree sequence (see van der Hofstad [93, Chapter 7], [94, Chapters 3 and 6]). Condition (R3) ensures that the random walk without backtracking is well-defined because it cannot get stuck on a dead-end.

Dynamic configuration model

We begin by describing the random graph process. It is convenient to take as the state space the set of configurations Conf_H. For a fixed initial configuration η and fixed $2 \leq k \leq m = \ell/2$, the graph evolves as follows (see Fig. 2.2):

(a) At each time $t \in \mathbb{N}$, pick k edges (pairs of half-edges) from C_{t-1} uniformly at random without replacement. Cut these edges to get $2k$ half-edges and denote this set of half-edges by R_t.

(b) Generate a uniform pairing of these half-edges to obtain k new edges. Replace the k edges chosen in step 1 by the k new edges to get the configuration C_t at time t.

This process rewires k edges at each step by applying the configuration model sampling algorithm in Section 2.1.2 restriced to k uniformly chosen edges. Since half-edges are not created or destroyed, the degree sequence of the graph given by C_t is the same for all $t \in \mathbb{N}_0$. This gives us a Markov chain on the set of configurations Conf_H. For $\eta, \zeta \in \mathit{Conf}_H$, the *transition probabilities* for this Markov chain are given by

$$Q(\eta,\zeta) = Q(\zeta,\eta) := \begin{cases} \frac{1}{(2k-1)!!}\frac{\binom{m-d_{\mathrm{Ham}}(\eta,\zeta)}{k-d_{\mathrm{Ham}}(\eta,\zeta)}}{\binom{m}{k}} & \text{if } d_{\mathrm{Ham}}(\eta,\zeta) \leq k, \\ 0 & \text{otherwise,} \end{cases} \tag{2.4}$$

where $d_{\mathrm{Ham}}(\eta,\zeta) := |\eta \setminus \zeta| = |\zeta \setminus \eta|$ is the Hamming distance between configurations η and ζ, which is the number of edges that appear in η but not in ζ. The factor $1/(2k-1)!!$ comes from the uniform pairing of the half-edges, while the factor

$\binom{m-d_{\mathrm{Ham}}(\eta,\zeta)}{k-d_{\mathrm{Ham}}(\eta,\zeta)}/\binom{m}{k}$ comes from choosing uniformly at random a set of k edges in η that contains the edges in $\eta \setminus \zeta$. It is easy to see that this Markov chain is irreducible and aperiodic, with stationary distribution the uniform distribution on $Conf_H$, denoted by Conf_H, which is the distribution of the configuration model.

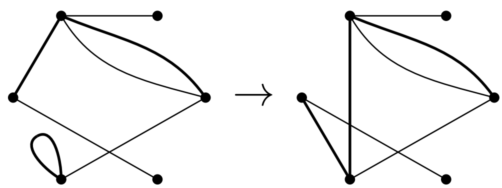

Figure 2.2: One move of the dynamic configuration model. Bold edges on the left are the ones chosen to be rewired. Bold edges on the right are the newly formed edges.

Random walk without backtracking

On top of the random graph process we define the random walk without backtracking, i.e., the walk cannot traverse the same edge twice in a row. As in Ben-Hamou and Salez [16], we define it as a random walk on the set of half-edges H, which is more convenient in the dynamic setting because the edges change over time while the half-edges do not. For a fixed configuration η and half-edges $x, y \in H$, the transition probabilities of the random walk are given by (recall (2.1))

$$P_\eta(x,y) := \begin{cases} \frac{1}{\deg(y)} & \text{if } \eta(x) \sim y \text{ and } \eta(x) \neq y, \\ 0 & \text{otherwise.} \end{cases} \tag{2.5}$$

When the random walk is at half-edge x in configuration η, it jumps to one of the siblings of the half-edge it is paired to uniformly at random (see Fig. 2.3). The transition probabilities are symmetric with respect to the pairing given by η, i.e., $P_\eta(x,y) = P_\eta(\eta(y), \eta(x))$, in particular, they are doubly stochastic, and so the uniform distribution on H, denoted by U_H, is stationary for P_η for any $\eta \in Conf_H$.

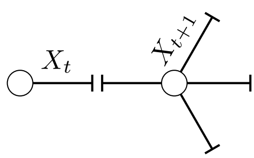

Figure 2.3: The random walk moves from half-edge X_t to half-edge X_{t+1}, one of the siblings of the half-edge that X_t is paired to.

Random walk on dynamic configuration model

The random walk without backtracking on the dynamic configuration model is the joint Markov chain $(M_t)_{t\in\mathbb{N}_0} = (C_t, X_t)_{t\in\mathbb{N}_0}$ in which $(C_t)_{t\in\mathbb{N}_0}$ is the Markov chain on the set of configurations $Conf_H$ as described in (2.4), and $(X_t)_{t\in\mathbb{N}_0}$ is the random walk that at each time step t jumps according to the transition probabilities $P_{C_t}(\cdot,\cdot)$ as in (2.5).

Formally, for initial configuration η and half-edge x, the one-step evolution of the joint Markov chain is given by the conditional probabilities

$$\mathbb{P}_{\eta,x}\big(C_t=\zeta, X_t=z \mid C_{t-1}=\xi, X_{t-1}=y\big) = Q(\xi,\zeta)\, P_\zeta(y,z), \qquad t\in\mathbb{N}, \tag{2.6}$$

with

$$\mathbb{P}_{\eta,x}(C_0=\eta, X_0=x)=1. \tag{2.7}$$

It is easy to see that if $d(v)>1$ for all $v\in V$, then this Markov chain is irreducible and aperiodic, and has the unique stationary distribution $\mathrm{Conf}_H \times U_H$.

While the graph process $(C_t)_{t\in\mathbb{N}_0}$ and the joint process $(M_t)_{t\in\mathbb{N}_0}$ are Markovian, the random walk $(X_t)_{t\in\mathbb{N}_0}$ is not. However, U_H is still the stationary distribution of $(X_t)_{t\in\mathbb{N}_0}$. Indeed, for any $\eta\in Conf_H$ and $y\in H$ we have

$$\sum_{x\in H} U_H(x)\,\mathbb{P}_{\eta,x}(X_t=y) = \sum_{x\in H}\frac{1}{\ell}\,\mathbb{P}_{\eta,x}(X_t=y) = \frac{1}{\ell} = U_H(y). \tag{2.8}$$

The next to last equality uses that $\sum_{x\in H}\mathbb{P}_{\eta,x}(X_t=y)=1$ for every $y\in H$, which can be seen by conditioning on the graph process and using that the space-time inhomogeneous random walk has a doubly stochastic transition matrix (recall the remarks made below (2.5)).

§2.1.3 Main theorem

We are interested in the behaviour of the total variation distance between the distribution of X_t and the uniform distribution

$$\mathcal{D}_{\eta,x}(t) := \|\mathbb{P}_{\eta,x}(X_t\in\cdot\,) - U_H(\cdot)\|_{\mathrm{TV}}. \tag{2.9}$$

[We recall that the total variation distance of two probability measures μ_1,μ_2 on a finite state space S is given by the following equivalent expressions:

$$\|\mu_1-\mu_2\|_{\mathrm{TV}} := \sum_{x\in S}|\mu_1(x)-\mu_2(x)| = \sum_{x\in S}[\mu_1(x)-\mu_2(x)]_+ = \sup_{A\subseteq S}[\mu_1(A)-\mu_2(A)], \tag{2.10}$$

where $[a]_+ := \max\{a,0\}$ for $a\in\mathbb{R}$.] Since $(X_t)_{t\in\mathbb{N}_0}$ is not Markovian, it is not clear whether $t\mapsto\mathcal{D}_{\eta,x}(t)$ is decreasing or not. On the other hand,

$$\mathcal{D}_{\eta,x}(t) \le \|\mathbb{P}_{\eta,x}(M_t\in\cdot\,) - (U_H\times\mathrm{Conf}_H)(\cdot)\|_{\mathrm{TV}}, \tag{2.11}$$

and since the right-hand side converges to 0 as $t\to\infty$, so does $\mathcal{D}_{\eta,x}(t)$. Therefore the following definition is well-posed:

Definition 2.1.4 (Mixing time of the random walk). *For* $\varepsilon\in(0,1)$*, the* ε-mixing time of the random walk *is defined as*

$$t^n_{\mathrm{mix}}(\varepsilon;\eta,x) := \inf\big\{t\in\mathbb{N}_0\colon\ \mathcal{D}_{\eta,x}(t)\le\varepsilon\big\}. \tag{2.12}$$

Note that $t^n_{\mathrm{mix}}(\varepsilon;\eta,x)$ depends on the initial configuration η and half-edge x. We will prove statements that hold for *typical* choices of (η,x) under the uniform distribution μ_n (recall that H depends on the number of vertices n) given by

$$\mu_n := \mathrm{Conf}_H \times U_H \quad \text{on } \mathit{Conf}_H \times H, \tag{2.13}$$

where *typical* is made precise through the following definition:

Definition 2.1.5 (**With high probability**). *A statement that depends on the initial configuration η and half-edge x is said to hold* with high probability (whp) *in η and x if the μ_n-measure of the set of pairs (η,x) for which the statement holds tends to* 1 *as $n\to\infty$.*

Below we sometimes write whp with respect to some probability measure other than μ_n, but it will always be clear from the context which probability measure we are referring to.

Throughout the paper we assume the following condition on

$$\alpha_n := k/m, \qquad n\in\mathbb{N}, \tag{2.14}$$

denoting *the proportion of edges involved in the rewiring* at each time step of the graph dynamics defined in Section 2.1.2:

Condition 2.1.6 (**Fast graph dynamics**). *The ratio α_n in* (2.14) *is subject to the constraint*

$$\lim_{n\to\infty}\alpha_n(\log n)^2 = \infty. \tag{2.15}$$

We can now state our main result.

Theorem 2.1.7 (**Sharp mixing time asymptotics**). *Suppose that Conditions* 2.1.2 *and* 2.1.6 *hold. Then, for every $\varepsilon>0$, whp in η and x,*

$$t^n_{\mathrm{mix}}(\varepsilon;\eta,x) = [1+o(1)]\sqrt{\frac{2\log(1/\varepsilon)}{\log(1/(1-\alpha_n))}}. \tag{2.16}$$

Note that Condition 2.1.6 allows for $\lim_{n\to\infty}\alpha_n = 0$. In that case (2.16) simplifies to

$$t^n_{\mathrm{mix}}(\varepsilon;\eta,x) = [1+o(1)]\sqrt{\frac{2\log(1/\varepsilon)}{\alpha_n}}. \tag{2.17}$$

§2.1.4 Discussion

1. Theorem 2.1.7 gives the sharp asymptotics of the mixing time in the regime where the dynamics is fast enough (as specified by Condition 2.1.6). Note that if $\lim_{n\to\infty}\alpha_n = \alpha\in(0,1]$, then $t^n_{\mathrm{mix}}(\varepsilon;\eta,x)$ is of order one: at every step the random walk has a non-vanishing probability to traverse a rewired edge, and so it is qualitatively similar to a random walk on a complete graph. On the other hand, when $\lim_{n\to\infty}\alpha_n = 0$ the mixing time is of order $1/\sqrt{\alpha_n} = o(\log n)$, which shows that the dynamics still *speeds up* the mixing. The regime $\alpha_n = \Theta(1/(\log n)^2)$, which is not captured by Theorem 2.1.7, corresponds to $1/\sqrt{\alpha_n} = \Theta(\log n)$, and we expect the mixing time to be *comparable* to that of the static configuration model. In the regime $\alpha_n = o(1/(\log n)^2)$ we expect the mixing time to be the *same* as that of the static configuration model. In a future paper we plan to provide a comparative analysis of the three regimes.

2. In the static model the ε-mixing time is known to scale like $[1 + o(1)]\, c \log n$ for some $c \in (0, \infty)$ that is independent of $\varepsilon \in (0, 1)$ (Ben-Hamou and Salez [16]). Consequently, there is *cutoff*, i.e., the total variation distance drops from 1 to 0 in a time window of width $o(\log n)$. In contrast, in the regime of fast graph dynamics there is *no cutoff*, i.e., the total variation distance drops from 1 to 0 gradually on scale $1/\sqrt{\alpha_n}$.

3. Our proof is robust and can be easily extended to variants of our model where, for example, $(k_n)_{n\in\mathbb{N}}$ is random with k_n having a first moment that tends to infinity as $n \to \infty$, or where time is continuous and pairs of edges are randomly rewired at rate α_n.

4. Theorem 2.1.7 can be compared to the analogous result for the static configuration model only when $\mathbb{P}_n(D_n \geq 3) = 1$ for all $n \in \mathbb{N}$. In fact, only under the latter condition does the probability of having a connected graph tend to one (see Luczak [68], Federico and van der Hofstad [47]). If (R3) holds, then on the dynamic graph the walk mixes on the whole of H, while on the static graph it mixes on the subset of H corresponding to the giant component.

5. We are not able to characterise the mixing time of the joint process of dynamic random graph and random walk. Clearly, the mixing time of the joint process is at least as large as the mixing time of each process separately. While the graph process helps the random walk to mix, the converse is not true because the graph process does not depend on the random walk. Observe that once the graph process has mixed it has an almost uniform configuration, and the random walk ought to have mixed already. This observation suggests that if the mixing times of the graph process and the random walk are not of the same order, then the mixing time of the joint process will have the same order as the mixing time of the graph process. Intuitively, we may expect that the mixing time of the graph corresponds to the time at which all edges are rewired at least once, which should be of order $(n/k) \log n = (1/\alpha_n) \log n$ by a coupon collector argument. In our setting the latter is much larger than $1/\sqrt{\alpha_n}$.

6. We emphasize that we look at the mixing times for 'typical' initial conditions and we look at the distribution of the random walk averaged over the trajectories of the graph process: the 'annealed' model. It would be interesting to look at different setups, such as 'worst-case' mixing, in which the maximum of the mixing times over all initial conditions is considered, or the 'quenched' model, in which the entire trajectory of the graph process is fixed instead of just the initial configuration. In such setups the results can be drastically different. For example, if we consider the quenched model for d-regular graphs, then we see that for any time t and any fixed realization of configurations up to time t, the walk without backtracking can reach at most $(d-1)^t$

half-edges. This gives us a lower bound of order $\log n$ for the mixing time in the quenched model, which contrasts with the $o(\log n)$ mixing time in our setup.

7. It would be of interest to extend our results to random walk with backtracking, which is harder. Indeed, because the configuration model is locally tree-like and random walk without backtracking on a tree is the same as self-avoiding walk, in our proof we can exploit the fact that typical walk trajectories are self-avoiding. In contrast, for the random walk with backtracking, after it jumps over a rewired edge, which in our model serves as a randomized stopping time, it may jump back over the same edge, in which case it has not mixed. This problem remains to be resolved.

§2.1.5 Outline

The remainder of this paper is organised as follows. Section 2.2 gives the main idea behind the proof, namely, we introduce a randomised stopping time $\tau = \tau_n$, the first time the walk moves along an edge that was rewired before, and we state a key proposition, Proposition 2.2.1 below, which says that this time is close to a strong stationary time and characterises its tail distribution. As shown at the end of Section 2.2, Theorem 2.1.7 follows from Proposition 2.2.1, whose proof consists of three main steps. The first step in Section 2.3 consists of a careful combinatorial analysis of the distribution of the walk given the history of the rewiring of the half-edges in the underlying evolving graph. The second step in Section 2.4 uses a classical exploration procedure of the static random graph from a uniform vertex to unveil the locally tree-like structure in large enough balls. The third step in Section 2.5 settles the closeness to stationarity and provides control on the tail of the randomized stopping time τ.

§2.2 Stopping time decomposition

We employ a *randomised stopping time* argument to get bounds on the total variation distance. We define the randomised stopping time $\tau = \tau_n$ to be the first time the walker makes a move through an edge that was rewired before. Recall from Section 2.1.2 that R_t is the set of half-edges involved in the rewiring at time step t. Letting $R_{\leq t} = \cup_{s=1}^{t} R_s$, we set

$$\tau := \min\{t \in \mathbb{N}\colon\ X_{t-1} \in R_{\leq t}\}. \tag{2.18}$$

As we will see later, τ behaves like a strong stationary time. We obtain our main result by deriving bounds on $\mathcal{D}_{\eta,x}(t)$ in terms of conditional distributions of the random walk involving τ and in terms of tail probabilities of τ. In particular, by the triangle inequality, for any $t \in \mathbb{N}_0$, $\eta \in \mathit{Conf}_H$ and $x \in H$,

$$\begin{aligned}\mathcal{D}_{\eta,x}(t) \leq \mathbb{P}_{\eta,x}(\tau > t)\, \|\mathbb{P}_{\eta,x}(X_t \in \cdot \mid \tau > t) - U_H(\cdot)\|_{\mathrm{TV}} \\ + \mathbb{P}_{\eta,x}(\tau \leq t)\, \|\mathbb{P}_{\eta,x}(X_t \in \cdot \mid \tau \leq t) - U_H(\cdot)\|_{\mathrm{TV}}\end{aligned} \tag{2.19}$$

and

$$\mathcal{D}_{\eta,x}(t) \geq \mathbb{P}_{\eta,x}(\tau > t) \, \|\mathbb{P}_{\eta,x}(X_t \in \cdot \mid \tau > t) - U_H(\cdot)\|_{\mathrm{TV}} \\ - \mathbb{P}_{\eta,x}(\tau \leq t) \, \|\mathbb{P}_{\eta,x}(X_t \in \cdot \mid \tau \leq t) - U_H(\cdot)\|_{\mathrm{TV}}. \tag{2.20}$$

With these in hand, we only need to find bounds for $\mathbb{P}_{\eta,x}(\tau > t)$, $\|\mathbb{P}_{\eta,x}(X_t \in \cdot \mid \tau > t) - U_H(\cdot)\|_{\mathrm{TV}}$ and $\|\mathbb{P}_{\eta,x}(X_t \in \cdot \mid \tau \leq t) - U_H(\cdot)\|_{\mathrm{TV}}$.

The key result for the proof of our main theorem is the following proposition:

Proposition 2.2.1 (Closeness to stationarity and tail behavior of stopping time).
Suppose that Conditions 2.1.2 *and* 2.1.6 *hold. For $t = t(n) = o(\log n)$, whp in x and η,*

$$\|\mathbb{P}_{\eta,x}(X_t \in \cdot \mid \tau \leq t) - U_H(\cdot)\|_{\mathrm{TV}} = o(1), \tag{2.21}$$

$$\|\mathbb{P}_{\eta,x}(X_t \in \cdot \mid \tau > t) - U_H(\cdot)\|_{\mathrm{TV}} = 1 - o(1), \tag{2.22}$$

$$\mathbb{P}_{\eta,x}(\tau > t) = (1 - \alpha_n)^{t(t+1)/2} + o(1). \tag{2.23}$$

We close this section by showing how Theorem 2.1.7 follows from Proposition 2.2.1:

Proof. By Condition 2.1.6,

$$\sqrt{\frac{2\log(1/\varepsilon)}{\log(1/(1-\alpha_n))}} = O(\alpha_n^{-1/2}) = o(\log n). \tag{2.24}$$

Using the bounds in (2.19)–(2.20), together with (2.21)–(2.23) in Proposition 2.2.1, we see that for $t = o(\log n)$,

$$(1 - \alpha_n)^{t(t+1)/2} + o(1) \leq \mathcal{D}_{\eta,x}(t) \leq (1 - \alpha_n)^{t(t+1)/2} + o(1). \tag{2.25}$$

Choosing t as in (2.16) we obtain $\mathcal{D}_{\eta,x}(t) = \varepsilon + o(1)$, which is the desired result. □

The remainder of the paper is devoted to the proof of Proposition 2.2.1.

§2.3 Pathwise probabilities

In order to prove (2.21) of Proposition 2.2.1, we will show in (2.69) in Section 2.5 that the following crucial bound holds for *most* $y \in H$:

$$\mathbb{P}_{\eta,x}(X_t = y \mid \tau \leq t) \geq \frac{1 - o(1)}{\ell}. \tag{2.26}$$

By *most* we mean that the number of y such that this inequality holds is $\ell - o(\ell)$ whp in η and x. To prove (2.26) we will look at $\mathbb{P}_{\eta,x}(X_t = y, \tau \leq t)$ by partitioning according to all possible paths taken by the walk and all possible rewiring patterns that occur on these paths. For a time interval $[s,t] \coloneqq \{s, s+1, \ldots, t\}$ with $s \leq t$, we define

$$x_{[s,t]} \coloneqq x_s \cdots x_t. \tag{2.27}$$

In particular, for any $y \in H$,

$$\mathbb{P}_{\eta,x}(X_t = y, \tau \le t) = \sum_{T \subseteq [1,t]} \sum_{x_1,\dots,x_{t-1} \in H} \mathbb{P}_{\eta,x}\Big(X_{[1,t]} = x_{[1,t]},\ x_{i-1} \in R_{\le i}\ \forall i \in T,\ x_{j-1} \notin R_{\le j}\ \forall j \in [1,t] \setminus T\Big) \tag{2.28}$$

with $x_0 = x$ and $x_t = y$. Here, r is the number of steps at which the walk moves along a previously rewired edge, and T is the set of times at which this occurs.

For a fixed sequence of half-edges $x_{[0,t]}$ with $x_0 = x$ and a fixed set of times $T \subseteq [1,t]$ with $|T| = r$, we will use the short-hand notation

$$A(x_{[0,t]}; T) := \big\{x_{i-1} \in R_{\le i}\ \forall i \in T,\ x_{j-1} \notin R_{\le j}\ \forall j \in [1,t] \setminus T\big\}. \tag{2.29}$$

Writing $T = \{t_1, \dots, t_r\}$ with $1 \le t_1 < t_2 < \dots < t_r \le t$, we note that the conditional probability $\mathbb{P}_{\eta,x}(X_{[1,t]} = x_{[1,t]} \mid A(x_{[0,t]}; T))$ can be non-zero only if each subsequence $x_{[t_{i-1}, t_i - 1]}$ induces a non-backtracking path in η for $i \in [2, r+1]$ with $t_0 = 0$ and $t_{r+1} = t+1$. The last sum in (2.28) is taken over such sequences in H, which we call *segmented paths* (see Fig. 2.4). For each $i \in [1, r+1]$ the subsequence $x_{[t_{i-1}, t_i - 1]}$ of length $t_i - t_{i-1}$ that forms a non-backtracking path in η is called a *segment*.

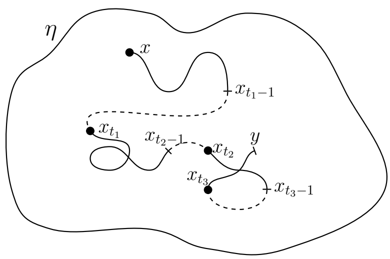

Figure 2.4: An example of a segmented path with 4 segments. Solid lines represent the segments, consisting of a path of half-edges in η, dashed lines indicate the succession of the segments. The latter do not necessarily correspond to a pair in η, and will later correspond to rewired edges in the graph dynamics.

We will restrict the last sum in (2.28) to the set of *self-avoiding segmented paths*. These are the paths where no two half-edges are siblings, which means that the vertices $v(x_i)$ visited by the half-edges x_i are distinct for all $i \in [0,t]$, so that if the random walk takes this path, then it does not see the same vertex twice. We will denote by $\mathsf{SP}_t^\eta(x, y; T)$ *the set of self-avoiding segmented paths* in η of length $t+1$ that start at x and end at y, where T gives the positions of the ends of the segments (see Fig. 2.5). Segmented paths $x_{[0,t]}$ have the nice property that the probability $\mathbb{P}_{\eta,x}(A(x_{[0,t]}; T))$ is the same for all $x_{[0,t]}$ that are isomorphic, as stated in the next lemma:

Lemma 2.3.1 (Isomorphic segmented path are equally likely). *Fix $t \in \mathbb{N}$, $T \subseteq [1,t]$ and $\eta \in \mathrm{Conf}_H$. Suppose that $x_{[0,t]}$ and $y_{[0,t]}$ are two segmented paths in η of length $t+1$ with $|x_{[s,s']}| = |y_{[s,s']}|$ for any $0 \le s < s' \le t$, where $|x_{[s,s']}|$ denotes the number of distinct half-edges in $x_{[s,s']}$. Then*

$$\mathbb{P}_{\eta,x}\big(A(x_{[0,t]}; T)\big) = \mathbb{P}_{\eta,x}\big(A(y_{[0,t]}; T)\big). \tag{2.30}$$

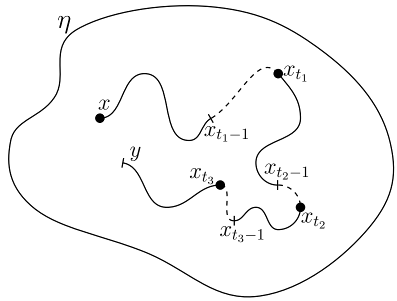

Figure 2.5: An element of $\mathsf{SP}^{\eta}_t(x, y; T)$ *with* $T = \{t_1, t_2, t_3\}$.

Proof. Fix $x, y \in H$. Consider the coupling $((C^x_t)_{t\in\mathbb{N}_0}, (C^y_t)_{t\in\mathbb{N}_0})$ of two dynamic configuration models with parameter k starting from η, defined as follows. Let $f\colon H \to H$ be such that

$$f(x) = \begin{cases} y_i & \text{if } x = x_i \text{ for some } i \in [0, t], \\ x_i & \text{if } x = y_i \text{ for some } i \in [0, t], \\ \eta(y_i) & \text{if } x = \eta(x_i) \text{ for some } i \in [0, t], \\ \eta(x_i) & \text{if } x = \eta(y_i) \text{ for some } i \in [0, t], \\ x & \text{otherwise.} \end{cases} \tag{2.31}$$

This is a one-to-one function because $|x_{[s,s']}| = |y_{[s,s']}|$ for any $0 \le s < s' \le t$. What f does is to map the half-edges of $x_{[0,t]}$ and their pairs in η to the half-edges of $y_{[0,t]}$ and their pairs in η, and vice versa, while preserving the order in the path. For the coupling, at each time $t \in \mathbb{N}$ we rewire the edges of C^x_{t-1} and C^y_{t-1} as follows:

(a) Choose k edges from C^x_{t-1} uniformly at random without replacement, say $\{z_1, z_2\}$, $\dots$, $\{z_{2k-1}, z_{2k}\}$. Choose the edges $\{f(z_1), f(z_2)\}, \dots, \{f(z_{2k-1}), f(z_{2k})\}$ from C^y_{t-1}.

(b) Rewire the half-edges $z_1, \dots, z_{2k}$ uniformly at random to obtain C^x_t. Set $C^y_t(f(z_i)) = f(C^x_t(z_i))$.

Step 2 and the definition of f ensure that in Step 1 $\{f(z_1), f(z_2)\}, \dots, \{f(z_{2k-1}), f(z_{2k})\}$ are in C^y_{t-1}. Since under the coupling the event $A(x_{[0,t]}; T)$ is the same as the event $A(y_{[0,t]}; T)$, we get the desired result. □

In order to prove the lower bound in (2.26), we will need two key facts. The first, stated in Lemma 2.3.2 below, gives a lower bound on the probability of a walk trajectory given the rewiring history. The second, stated in Lemma 2.4.3 below, is a lower bound on the number of relevant self-avoiding segmented paths, and exploits the locally tree-like structure of the configuration model.

Lemma 2.3.2 (Paths estimate given rewiring history). *Suppose that* $t = t(n) = o(\log n)$ *and* $T = \{t_1, \dots, t_r\} \subseteq [1, t]$. *Let* $x_0 \cdots x_t \in \mathsf{SP}^{\eta}_t(x, y; T)$ *be a self-avoiding* *segmented path in* η *that starts at* x *and ends at* y*. Then*

$$\mathbb{P}_{\eta,x}\big(X_{[1,t]} = x_{[1,t]} \mid A(x_{[0,t]};T)\big) \geq \frac{1-o(1)}{\ell^r} \prod_{i\in[1,t]\setminus T} \frac{1}{\deg(x_i)}. \tag{2.32}$$

Proof. In order to deal with the dependencies introduced by conditioning on the event $A(x_{[0,t]};T))$, we will go through a series of conditionings. First we note that for the random walk to follow a specific path, the half-edges it traverses should be rewired *correctly* at the *right times*. Conditioning on $A(x_{[0,t]};T)$ accomplishes part of the job: since we have $x_{i-1} \notin R_{\leq i}$ for $i \in [1,t] \setminus T$ and $x_{[0,t]} \in \mathsf{SP}_t^\eta(x,y;T)$, we know that, at time i, x_{i-1} is paired to a sibling of x_i in C_i, and so the random walk can jump from x_{i-1} to x_i with probability $1/\deg(x_i)$ at time i for $i \in [1,t] \setminus T$.

Let us call the path $x_{[0,t]}$ *open* if $C_i(x_{i-1}) \sim x_i$ for $i \in [1,t]$, i.e., if x_{i-1} is paired to a sibling of x_i in C_i for $i \in [1,t]$. Then

$$\mathbb{P}_{\eta,x}(X_{[1,t]} = x_{[1,t]} \mid x_{[0,t]} \text{ is open}) = \prod_{i=1}^{t} \frac{1}{\deg(x_i)}, \tag{2.33}$$

and

$$\mathbb{P}_{\eta,x}\big(X_{[1,t]} = x_{[1,t]} \mid x_{[0,t]} \text{ is not open}\big) = 0. \tag{2.34}$$

Using these observations, we can treat the random walk and the rewiring process separately, since the event $\{x_{[0,t]} \text{ is open}\}$ depends only on the rewirings. Our goal is to compute the probability

$$\mathbb{P}_{\eta,x}\big(x_{[0,t]} \text{ is open} \mid A(x_{[0,t]};T)\big). \tag{2.35}$$

Note that, by conditioning on $A(x_{[0,t]};T)$, the part of the path within segments is already open, so we only need to deal with the times the walk jumps from one segment to another. To have $x_{[0,t]}$ open, each x_{t_j-1} should be paired to one of the siblings of x_{t_j} for $j \in [1,r]$. Hence

$$\begin{aligned}&\mathbb{P}_{\eta,x}\big(x_{[0,t]} \text{ is open} \mid A(x_{[0,t]};T)\big)\\ &\qquad= \sum_{\substack{z_1,\dots,z_r\in H\\ z_j\sim x_{t_j}\ \forall j\in[1,r]}} \mathbb{P}_{\eta,x}\big(C_{t_j}(x_{t_j-1}) = z_j\ \forall j \in [1,r] \mid A(x_{[0,t]};T)\big).\end{aligned} \tag{2.36}$$

Fix $z_1,\dots,z_r \in H$ with $z_j \sim x_{t_j}$, and let $y_j = x_{t_j-1}$ for $j \in [1,r]$. We will look at the probability

$$\mathbb{P}_{\eta,x}\big(C_{t_j}(y_j) = z_j\ \forall j \in [1,r] \mid A(x_{[0,t]};T)\big). \tag{2.37}$$

Conditioning on the event $A(x_{[0,t]};T)$ we impose that each y_j is rewired at some time before t_j, but do not specify at which time this happens. Let us refine our conditioning one step further by specifying these times. Fix $s_1,\dots,s_r \in [1,t]$ such that $s_j \leq t_j$ for each $j \in [1,r]$ (the s_j need not be distinct). Let $\widehat{A}$ be the event that $x_{i-1} \notin R_{\leq i}$ for $i \in [1,t] \setminus T$ and y_j is rewired at time s_j for the last time before time t_j for $j \in [1,r]$. Then $\widehat{A} \subseteq A(x_{[0,t]};T)$. Since s_j is the last time before t_j at which

y_j is rewired, the event $C_{t_j}(y_j) = z_j$ is the same as the event $C_{s_j}(y_j) = z_j$ when we condition on $\widehat{A}$. We look at the probability

$$\mathbb{P}_{\eta,x}\big(C_{s_j}(y_j) = z_j \ \forall\, j \in [1,r] \mid \widehat{A}\big). \tag{2.38}$$

Let $s'_1 < \cdots < s'_{r'} \in [1,t]$ be the distinct times such that $s'_i = s_j$ for some $j \in [1,r]$, and n_i^y the number of j's for which $s_j = s'_i$ for $i \in [1,r']$, so that by conditioning on $\widehat{A}$ we rewire n_i^y half-edges y_j at time s'_i. Letting also $D_i = \{C_{s'_i}(y_j) = z_j,$ for j such that $s_j = s'_i\}$, we can write the above conditional probability as

$$\prod_{i=1}^{r'} \mathbb{P}_{\eta,x}\big(D_i \mid \widehat{A},\, \cap_{j=1}^{i-1} D_j\big). \tag{2.39}$$

We next compute these conditional probabilities.

Fix $i \in [1,r']$ and $\eta' \in \mathit{Conf}_H$. We do one more conditioning and look at the probability

$$\mathbb{P}_{\eta,x}\big(D_i \mid \widehat{A},\, \cap_{j=1}^{i-1} D_j,\, C_{s'_i-1} = \eta'\big). \tag{2.40}$$

The rewiring process at time s'_i consists of two steps: (1) pick k edges uniformly at random; (2) do a uniform rewiring. Concerning (1), by conditioning on $\widehat{A}$, we see that the y_j's for which $s_j = s'_i$ are already chosen. In order to pair these to z_j's with $s_j = s'_i$, the z_j's should be chosen as well. If some of the z_j's are already paired to some y_j's already chosen, then they will be automatically included in the rewiring process. Let m'_i be m minus the number of half-edges in $\{x_0, \dots, x_t\} \cup \{z_1, \dots, z_r\}$, for which the conditioning on $\widehat{A}$ implies that they cannot be in $R_{s'_i}$. Then

$$\begin{aligned} &\mathbb{P}_{\eta,x}\Big(z_j \in R_{s'_i} \text{ for } j \text{ such that } s_j = s'_i \;\Big|\; \widehat{A},\, \cap_{j=1}^{i-1} D_j,\, C_{s'_i-1} = \eta'\Big) \\ &\geq \frac{\binom{m'_i - 2n_i^y}{k - 2n_i^y}}{\binom{m'_i - n_i^y}{k - n_i^y}} = \frac{\prod_{j=0}^{n_i^y - 1}(k - n_i^y - j)}{\prod_{j=0}^{n_i^y - 1}(m'_i - n_i^y - j)} \geq \frac{\prod_{j=0}^{n_i^y - 1}(k - n_i^y - j)}{m^{n_i^y}}. \end{aligned} \tag{2.41}$$

Concerning (2), conditioned on the relevant z_j's already chosen in (1), the probability that they will be paired to correct y_j's is

$$\frac{1}{\prod_{j=1}^{n_i^y}(2k - 2j + 1)}. \tag{2.42}$$

Since the last two statements hold for any η' with $\mathbb{P}_{\eta,x}(C_{s'_i-1} = \eta' \mid \widehat{A},\, \cap_{j=1}^{i-1} D_j) > 0$, combining these we get

$$\mathbb{P}_{\eta,x}\big(D_i \mid \widehat{A},\, \cap_{j=1}^{i-1} D_j\big) \geq \frac{\prod_{j=0}^{n_i^y - 1}(k - n_i^y - j)}{m^{n_i^y} \prod_{j=1}^{n_i^y}(2k - 2j + 1)} = \left(\frac{1 - O(n_i^y/k)}{2m}\right)^{n_i^y}. \tag{2.43}$$

Since $\sum_{i=1}^{r'} n_i^y = r$, substituting (2.43) into (2.39) and rolling back all the conditionings we did so far, we get

$$\mathbb{P}_{\eta,x}\big(C_{t_j}(x_{t_j-1}) = z_j \ \forall\, j \in [1,r] \mid A(x_{[0,t]}; T)\big) \geq \frac{1 - O(r^2/k)}{\ell^r} = \frac{1 - o(1)}{\ell^r}, \tag{2.44}$$

where we use that $r^2/k \to 0$ since $r = o(\log n)$ and $k = \alpha_n n$ with $(\log n)^2 \alpha_n \to \infty$. Now sum over $z_1, \dots, z_r$ in (2.36), to obtain

$$\mathbb{P}_{\eta,x}\big(x_{[0,t]} \text{ is open} \mid A(x_{[0,t]}; T)\big) \geq \frac{(1-o(1)) \prod_{j=1}^{r} \deg(x_{t_j})}{\ell^r}, \tag{2.45}$$

and multiply with (2.33) to get the desired result. □

§2.4 Tree-like structure of the configuration model

In this section we look at the structure of the neighborhood of a half-edge chosen uniformly at random in the configuration model. Since we will work with different probability spaces, we will denote by $\mathbb{P}$ a generic probability measure whose meaning will be clear from the context.

For fixed $t \in \mathbb{N}$, $x \in H$ and $\eta \in Conf_H$, we denote by $B_t^\eta(x) := \{y \in H\colon \operatorname{dist}_\eta(x,y) \leq t\}$ the t-neighborhood of x in η, where $\operatorname{dist}_\eta(x,y)$ is *the length of the shortest non-backtracking path from x to y.* We start by estimating the mean of $|B_t^\eta(x)|$, the number of half-edges in $B_t^\eta(x)$.

Lemma 2.4.1 (Average size of balls of relevant radius). *Let ν_n be as in Condition* 2.1.2 *and suppose that $t = t(n) = o(\log n)$. Then, for any $\delta > 0$,*

$$\mathbb{E}(|B_t^\eta(x)|) = [1 + o(1)]\, \nu_n^{t+1} = o(n^\delta), \tag{2.46}$$

where the expectation is w.r.t. μ_n in (2.13).

Proof. We have

$$|B_t^\eta(x)| = \sum_{y \in H} \mathbb{1}_{\{\operatorname{dist}_\eta(x,y) \leq t\}}. \tag{2.47}$$

Putting this into the expectation, we get

$$\mathbb{E}(|B_t^\eta(x)|) = \frac{1}{\ell} \sum_{x,y \in H} \mathbb{P}(\operatorname{dist}_\eta(x,y) \leq t). \tag{2.48}$$

For fixed $x, y \in H$,

$$\begin{aligned}
\mathbb{P}(\operatorname{dist}_\eta(x,y) \leq t) &\leq \sum_{d=1}^{t} \sum_{x_1,\dots,x_{d-1} \in H} \mathbb{P}(x x_1 \cdots x_{d-1} y \text{ forms a self-avoiding path in } \eta) \\
&\leq \sum_{d=1}^{t} \sum_{x_1,\dots,x_{d-1} \in H} \left(\prod_{j=1}^{d-1} \frac{\deg(x_j)}{\ell - 2j + 1} \right) \frac{\deg(y)}{\ell - 2d + 1} \\
&= \frac{\deg(y)}{\ell} \sum_{d=1}^{t} \left(\prod_{i=1}^{d} \frac{\ell}{\ell - 2i + 1} \right) \sum_{x_1,\dots,x_{d-1} \in H} \left(\prod_{i=1}^{d-1} \frac{\deg(x_i)}{\ell} \right) \\
&= \frac{\deg(y)}{\ell} \sum_{d=1}^{t} \left(\prod_{i=1}^{d} \frac{\ell}{\ell - 2i + 1} \right) \left(\sum_{z \in H} \frac{\deg(z)}{\ell} \right)^{d-1}.
\end{aligned} \tag{2.49}$$

Since $t = o(\log n)$ and $\ell = \Theta(n)$, we have

$$\mathbb{P}(\mathrm{dist}_\eta(x, y) \le t) \le [1 + o(1)]\, \frac{\deg(y)}{\ell}\, (\nu_n)^t. \tag{2.50}$$

Substituting this into (2.48), we get

$$\mathbb{E}(|B_t^\eta(x)|) \le \frac{1 + o(1)}{\ell} \sum_{x,y\in H} \frac{\deg(y)}{\ell}\, (\nu_n)^t = [1 + o(1)]\, (\nu_n)^{t+1} = o(n^\delta), \tag{2.51}$$

where the last equality follows from (R2) in Condition 2.1.2 and the fact that $t = o(\log n)$. □

For the next result we will use an *exploration process* to build the neighborhood of a uniformly chosen half-edge. (Similar exploration processes have been used in [16],[21] and [67].) We explore the graph by starting from a uniformly chosen half-edge x and building up the graph by successive uniform pairings, as explained in the procedure below. Let $\mathsf{G}(s)$ denote the *thorny graph* obtained after s pairings as follows (in our context, a thorny graph is a graph in which half-edges are not necessarily paired to form edges, as shown in Fig. 2.6). We set $\mathsf{G}(0)$ to consist of x, its siblings, and the incident vertex $v(x)$. Along the way we keep track of *the set of unpaired half-edges at each time* s, denoted by $U(s) \subset H$, and the so-called *active* half-edges, $A(s) \subset U(s)$. We initialize $U(0) = H$ and $A(0) = \{x\}$. We build up the sequence of graphs $(\mathsf{G}(s))_{s\in\mathbb{N}_0}$ as follows:

(a) At each time $s \in \mathbb{N}$, take the *next* unpaired half-edge in $A(s-1)$, say y. Sample a half-edge uniformly at random from H, say z. If z is already paired or $z = y$, then reject and sample again. Pair y and z.

(b) Add the newly formed edge $\{y, z\}$, the incident vertex $v(z)$ of z, and its siblings to $\mathsf{G}(s-1)$, to obtain $\mathsf{G}(s)$.

(c) Set $U(s) = U(s-1) \setminus \{y, z\}$, i.e., remove y, z from the set of unpaired half-edges, and set $A(s) = A(s-1) \cup \{H(v(z))\} \setminus \{y, z\}$, i.e., add siblings of z to the set of active half-edges and remove the active half-edges just paired.

This procedure stops when $A(s)$ is empty. We think of $A(s)$ as a first-in first-out queue. So, when we say that we pick the *next* half-edge in Step 1, we refer to the half-edge on top of the queue, which ensures that we maintain the breadth-first order. The rejection sampling used in Step 1 ensures that the resulting graph is distributed according to the configuration model. This procedure eventually gives us the connected component of x in η, the part of the graph that can be reached from x by a non-backtracking walk, where η is distributed uniformly on Conf_H.

Lemma 2.4.2 (Tree-like neighborhoods). *Suppose that $s = s(n) = o(n^{(1-2\delta)/2})$ for some $\delta \in (0, \frac{1}{2})$. Then $\mathsf{G}(s)$ is a tree with probability $1 - o(n^{-\delta})$.*

Proof. Let F be the first time the uniform sampling of z in Step 1 fails at the first attempt, or z is a sibling of x, or z is in $A(s-1)$. Thus, at time F we either choose

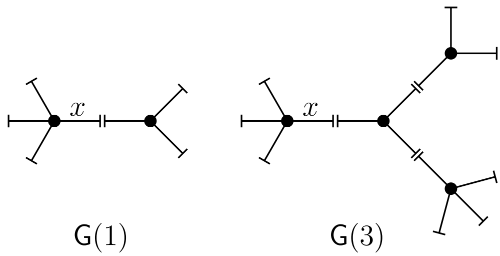

Figure 2.6: Example snapshots of $\mathsf{G}(s)$ *at times* $s = 1$ *and* $s = 3$.

an already paired half-edge or we form a cycle by pairing to some half-edge already present in the graph. We have

$$\mathbb{P}(\mathsf{G}(s) \text{ is not a tree}) \leq \mathbb{P}(F \leq s). \tag{2.52}$$

Let Y_i, $i \in \mathbb{N}$, be i.i.d. random variables whose distribution is the same as the distribution of the degree of a uniformly chosen half-edge. When we form an edge before time F, we use one of the unpaired half-edges of the graph, and add new unpaired half-edges whose number is distributed as Y_1. Hence the number of unpaired half-edges in $\mathsf{G}(u)$ is stochastically dominated by $\sum_{i=1}^{u+1} Y_i - u$, with one of the Y_i's coming from x and the other ones coming from the formation of each edge. Therefore the probability that one of the conditions of F will be met at step u is stochastically dominated by $(\sum_{i=1}^{u} Y_i + u - 2)/\ell$. We either choose an unpaired half-edge in $\mathsf{G}(u)$ or we choose a half-edge belonging to an edge in $\mathsf{G}(u)$, and by the union bound we have

$$\begin{aligned} &\mathbb{P}(\mathsf{G}(s) \text{ is not a tree} \mid (Y_i)_{i\in[1,s]}) \leq \mathbb{P}(F \leq s \mid (Y_i)_{i\in[1,s]}) \\ &\leq \frac{\sum_{u=1}^{s}\sum_{i=1}^{u}(Y_i + u - 2)}{\ell} = \frac{\sum_{i=1}^{s}(s-i+1)Y_i + s(s-1)/2}{\ell}. \end{aligned} \tag{2.53}$$

Since $\mathbb{E}(Y_1) = \nu_n = O(1)$ and $s = o(n^{(1-2\delta)/2})$, via the Markov inequality we get that, with probability at least $1 - o(n^{-\delta})$,

$$s \sum_{i=1}^{s} Y_i < n^{1-\delta}. \tag{2.54}$$

Combining this with the bound given above and the fact that $\ell = \Theta(n)$, we arrive at

$$\mathbb{P}(\mathsf{G}(s) \text{ is not a tree}) = o(n^{-\delta}). \tag{2.55}$$

□

To further prepare for the proof of the lower bound in (2.26) and Proposition 2.2.1 in Section 2.5, we introduce one last ingredient. For $x \in H$ and $\eta \in Conf_H$, we denote by $\bar{B}_t^{\eta}(x)$ the set of half-edges from which there is a non-backtracking path to x of length at most t. For fixed $t \in \mathbb{N}$, $T = \{t_1, \ldots, t_r\} \subseteq [1, t]$ and $\eta \in Conf_H$, we say that an $(r+1)$-tuple $(x_0, x_1, \ldots, x_r)$ is *good* for T in η if it satisfies the following two properties:

(a) $B^{\eta}_{t_j-t_{j-1}}(x_j)$ is a tree for $j \in [1,r]$ with $t_0 = 0$, and $\bar{B}^{\eta}_{t-t_r}(x_r)$ is a tree.

(b) The trees $B^{\eta}_{t_j-t_{j-1}}(x_j)$ for $j \in [1,r]$ and $\bar{B}^{\eta}_{t-t_r}(x_r)$ are all disjoint.

For a good $(r+1)$-tuple all the segmented paths, such that the ith segment starts from x_{i-1} and is of length $t_i - t_{i-1}$ for $i \in [1,r]$ and the $(r+1)$st segment ends at x_r and is of length $t - t_r$, are self-avoiding by the tree property. The next lemma states that whp in η almost all $(r+1)$-tuples are good. We denote by $N^{\eta}_t(T)$ the set of $(r+1)$-tuples that are good for T in η, and let $N^{\eta}_t(T)^c$ be the complement of $N^{\eta}_t(T)$. We have the following estimate on $|N^{\eta}_t(T)|$:

Lemma 2.4.3 (Estimate on good paths). *Suppose that $t = t(n) = o(\log n)$. Then there exist $\bar{\delta} > 0$ such that whp in η for all $T \subseteq [1,t]$,*

$$|N^{\eta}_t(T)| = (1 - n^{-\bar{\delta}})\ell^{|T|+1}. \tag{2.56}$$

Proof. Fix $\varepsilon > 0$ and $T \subseteq [1,t]$ with $|T| = r$. We want to show that whp $|N^{\eta}_t(T)^c| \leq \varepsilon \ell^{r+1}$. By the Markov inequality, we have

$$\mathbb{P}(|N^{\eta}_t(T)^c| > \varepsilon \ell^{r+1}) \leq \frac{\mathbb{E}(|N^{\eta}_t(T)^c|)}{\varepsilon \ell^{r+1}} = \frac{\mathbb{P}(Z_{[0,r]} \in N^{\eta}_t(T)^c)}{\varepsilon}, \tag{2.57}$$

where $Z_0, \dots, Z_r$ are i.i.d. uniform half-edges and we use that $1/\ell^{r+1}$ is the uniform probability over a collection of $r+1$ half-edges. Let $B_{i-1} = B^{\eta}_{t_i - t_{i-1}}(Z_{i-1})$ for $i \in [1,r]$ and $B_r = B^{\eta}_{t-t_r}(Z_r)$. By the union bound,

$$\mathbb{P}\big(Z_{[0,r]} \in N^{\eta}_t(T)^c\big) \leq \sum_{i=0}^{r} \mathbb{P}(B_i \text{ is not a tree}) + \sum_{i,j=0}^{r} \mathbb{P}(B_i \cap B_j \neq \varnothing). \tag{2.58}$$

By Lemma 2.4.1 and since $t = o(\log n)$, for any $0 < \delta < \frac{1}{2}$ we have $\mathbb{E}|B_i| = o(n^{\delta})$, and so by the Markov inequality $|B_i| = o(n^{(1-2\delta)/2})$ with probability $1 - o(n^{-\delta})$. Hence, by Lemma 2.4.2 and since $\ell = \Theta(n)$, for $i \in [1,r]$, we have

$$\mathbb{P}(B_{i-1} \text{ is not a tree}) = o(n^{-\delta}). \tag{2.59}$$

Again using Lemma 2.4.1, we see that for any $i,j \in [1,r]$,

$$\mathbb{P}(B_i \cap B_j \neq \varnothing) \leq \mathbb{P}(Z_j \in B^{\eta}_t(Z_i)) = \frac{\mathbb{E}(|B^{\eta}_t(Z_i)|)}{\ell} \leq o(n^{\delta-1}). \tag{2.60}$$

Since $r \leq t = o(\log n)$, setting $\bar{\delta} = 2\delta/3$ and taking $\varepsilon = n^{-\delta}$, we get

$$\mathbb{P}(|N^{\eta}_t(T)^c| > \varepsilon \ell^{r+1}) \leq \frac{r n^{-\bar{\delta}} + r^2 n^{\bar{\delta}-1}}{\varepsilon} = o(n^{-\delta/4}) \tag{2.61}$$

uniformly in $T \subseteq [1,t]$. Since there are 2^t different $T \subseteq [1,t]$ and $2^t = 2^{o(\log n)} = o(n^{\delta/8})$, taking the union bound we see that (2.56) holds for all $T \subseteq [1,t]$ with probability $1 - o(n^{-\delta/8})$. $\square$

§2.5 Closeness to stationarity and tail behavior of stopping time

We are now ready to prove the lower bound in (2.26) and Proposition 2.2.1. Before giving these proofs, we need one more lemma, for which we introduce some new notation. For fixed $t \in \mathbb{N}$, $T \subseteq [1,t]$ with $|T| = r > 0$, $\eta \in Conf_H$ and $x, y \in H$, let $N_t^\eta(x,y;T)$ denote the set of $(r-1)$-tuples such that $(x, x_1, \ldots, x_{r-1}, y)$ is good for T in η. Furthermore, for a given $(r+1)$-tuple $\mathbf{x} = (x, x_1, \ldots, x_{r-1}, y)$ that is good for T in η, let $\mathsf{SP}_t^\eta(\mathbf{x};T)$ denote the set of all segmented paths in which the ith segment starts at x_{i-1} and is of length $t_i - t_{i-1}$ for $i \in [1,r]$ with $x_0 = x$ and $t_0 = 0$, and the $(r+1)$st segment ends at y and is of length $t - t_r$. By the definition of a good tuple, these paths are self-avoiding, and hence $\mathsf{SP}_t^\eta(\mathbf{x};T) \subset \mathsf{SP}_t^\eta(x,y;T)$.

Lemma 2.5.1 (**Total mass of relevant paths**). *Suppose that $t = t(n) = o(\log n)$. Then whp in η and x, y for all $T \subseteq [1,t]$,*

$$\sum_{x_{[0,t]} \in \mathsf{SP}_t^\eta(x,y;T)} \mathbb{P}_{\eta,x}\big(X_{[1,t]} = x_{[1,t]} \mid A(x_{[0,t]};T)\big) \geq \frac{1-o(1)}{\ell}. \tag{2.62}$$

Proof. By Lemma 2.4.3, the number of pairs of half-edges x, y for which $|N_t^\eta(x,y;T)| \geq (1-n^{-\bar{\delta}})\ell^{|T|-1} = [1-o(1)]\,\ell^{|T|-1}$ for all $T \in [1,t]$ is at least $(1-2^t n^{-\bar{\delta}})\ell^2 = [1-o(1)]\,\ell^2$ whp in η. Take such a pair $x, y \in H$, and let $r = |T|$. By Lemma 2.3.2 and the last observation before the statement of Lemma 2.5.1, we have

$$\begin{aligned}\sum_{x_{[0,t]} \in \mathsf{SP}_t^\eta(x,y;T)} &\mathbb{P}_{\eta,x}\big(X_{[1,t]} = x_{[1,t]} \mid A(x_{[0,t]};T)\big)\\ &\geq \sum_{\mathbf{x} \in N_t^\eta(x,y;T)} \sum_{y_0 \ldots y_t \in \mathsf{SP}_t^\eta(\mathbf{x},T)} \frac{1-o(1)}{\ell^r} \prod_{i \in [1,t]\setminus T} \frac{1}{\deg(y_i)}.\end{aligned} \tag{2.63}$$

We analyze at the second sum by inspecting the contributions coming from each segment separately. For fixed $\mathbf{x} \in N_t^\eta(x,y;T)$, when we sum over the segmented paths in $\mathsf{SP}_t^\eta(\mathbf{x},T)$, we sum over all paths that go out of x_{i-1} of length $t_i - t_{i-1}$ for $i \in [1,r]$. Since $\prod_{j=t_{i-1}+1}^{t_i-1} \frac{1}{\deg(y_j)}$ is the probability that the random walk without backtracking follows this path on the static graph given by η starting from x_{i-1}, when we sum over all such paths the contribution from these terms sums up to 1 for each $i \in [1,r]$, i.e., the contributions of the first r segments coming from the products of inverse degrees sum up to 1. For the last segment we sum, over all paths going into y, the probability that the random walk without backtracking on the static graph given by η follows the path. Since the uniform distribution is stationary for this random walk, the sum over the last segment of the probabilities $\frac{1}{\ell}\prod_{j=t_r+1}^{t} \frac{1}{\deg(y_j)}$ gives us $1/\ell$. With this observation, using that $|N_t^\eta(x,y;T)| \geq (1-o(1))\ell^{r-1}$, we get

$$\begin{aligned}\sum_{x_{[0,t]} \in \mathsf{SP}_t^\eta(x,y;T)} &\mathbb{P}_{\eta,x}\big(X_{[1,t]} = x_{[1,t]} \mid A(x_{[0,t]};T)\big)\\ &\geq \frac{1-o(1)}{\ell} \sum_{\mathbf{x} \in N_t^\eta(x,y;T)} \frac{1-o(1)}{\ell^{r-1}} = \frac{1-o(1)}{\ell},\end{aligned} \tag{2.64}$$

which is the desired result. □

• *Proof of* (2.21). For any self-avoiding segmented path $x_0 \cdots x_t$, we have $|x_{[s,s']}| = s' - s + 1$ for all $1 \le s < s' \le t$. By Lemma 2.3.1, the probability $\mathbb{P}_{\eta,x}(A(x_{[0,t]};T))$ depends on η and T only, and we can write $\mathbb{P}_{\eta,x}(A(x_{[0,t]};T)) = p_t^\eta(T)$ for any $xx_1 \cdots x_{t-1}y \in \mathsf{SP}_t^\eta(x,y;T)$. Applying Lemma 2.5.1, we get

$$\begin{aligned}
&\mathbb{P}_{\eta,x}(X_t = y, \tau \le t) \qquad (2.65)\\
&\ge \sum_{r=1}^{t} \sum_{\substack{T\subseteq[1,t]\\|T|=r}} \sum_{x_{[0,t]}\in\mathsf{SP}_t^\eta(x,y;T)} \mathbb{P}_{\eta,x}\big(X_{[1,t]} = x_{[1,t]} \mid A(x_{[0,t]};T)\big)\,\mathbb{P}_{\eta,x}\big(A(x_{[0,t]};T)\big)\\
&\ge \frac{1-o(1)}{\ell} \sum_{r=1}^{t} \sum_{\substack{T\subseteq[1,t]\\|T|=r}} p_t^\eta(T).
\end{aligned}$$

If the t-neighborhood of x in η is a tree, then all t-step non-backtracking paths starting at x are self-avoiding. (Here is a place where the non-backtracking nature of our walk is crucially used!) In particular, for any such path $xx_1 \cdots x_t$ we have $\mathbb{P}_{\eta,x}(A(x_{[0,t]};\varnothing)) = p_t^\eta(\varnothing)$. Denoting by $\Gamma_t^\eta(x)$ the set of paths in η of length t that start from x, we also have

$$\begin{aligned}
\mathbb{P}_{\eta,x}(\tau > t) &= \sum_{x_0\cdots x_t\in\Gamma_t^\eta(x)} \mathbb{P}_{\eta,x}\big(X_{[1,t]} = x_{[1,t]}, A(x_{[0,t]};\varnothing)\big)\\
&= \sum_{x_0\cdots x_t\in\Gamma_t^\eta(x)} \prod_{i=1}^{t} \frac{1}{\deg(x_i)}\, p_t^\eta(\varnothing) = p_t^\eta(\varnothing), \qquad (2.66)
\end{aligned}$$

since the product $\prod_{i=1}^t \frac{1}{\deg(x_i)}$ is the probability that a random walk without backtracking in the static η follows the path $x_0x_1\cdots x_t$, and we take the sum over all paths going out of x.

For a fixed path $x_0x_1\cdots x_t$, we have

$$\sum_{r=1}^{t} \sum_{\substack{T\subseteq[1,t]\\|T|=r}} \mathbb{P}_{\eta,x}\big(A(x_{[0,t]};T)\big) = 1 - \mathbb{P}_{\eta,x}\big(A(x_{[0,t]};\varnothing)\big). \qquad (2.67)$$

So, when the t-neighborhood of x in η is a tree, we have

$$\sum_{r=1}^{t} \sum_{\substack{T\subseteq[1,t]\\|T|=r}} p_t^\eta(T) = 1 - p_t^\eta(\varnothing) = 1 - \mathbb{P}_{\eta,x}(\tau > t) = \mathbb{P}_{\eta,x}(\tau \le t), \qquad (2.68)$$

which gives

$$\mathbb{P}_{\eta,x}(X_t = y, \tau \le t) \ge \frac{1-o(1)}{\ell}\, \mathbb{P}_{\eta,x}(\tau \le t) \qquad (2.69)$$

and settles the lower bound (2.26). Since the latter holds whp in η and x, y, we have that the number of y for which this holds is $[1-o(1)]\,\ell$ whp in η and x. Denoting the set of $y \in H$ for which the lower bound in (2.26) holds by $N_t^\eta(x)$, we get that whp in η and x,

$$\begin{aligned}\|\mathbb{P}_{\eta,x}(X_t \in \cdot \mid \tau \le t) - U_H(\cdot)\|_{\mathrm{TV}} &= \sum_{y\in H}\left[\frac{1}{\ell} - \mathbb{P}_{\eta,x}(X_t = y \mid \tau \le t)\right]^+ \\ &\le \sum_{y\in N_t^\eta(x)}\left[\frac{1}{\ell} - \frac{1-o(1)}{\ell}\right]^+ + \sum_{y\notin N_t^\eta(x)} \frac{1}{\ell} = o(1),\end{aligned} \tag{2.70}$$

which is (2.21). □

• *Proof of* (2.22). First note that $\mathbb{P}_{\eta,x}(X_t \in B_t^\eta(x) \mid \tau > t) = 1$. On the other hand, using Lemma 2.4.1 and the Markov inequality, we see that $U_H(B_t^\eta(x)) = |B_t^\eta(x)|/\ell = o(1)$ whp in η and x, and so we get

$$\|\mathbb{P}_{\eta,x}(X_t \in \cdot \mid \tau > t) - U_H(\cdot)\|_{\mathrm{TV}} \ge \mathbb{P}_{\eta,x}(X_t \in B_t^\eta(x) \mid \tau > t) - U_H(B_t^\eta(x)) = 1 - o(1). \tag{2.71}$$

□

• *Proof of* (2.23). Taking $T = \varnothing$ in Lemma 2.4.3, we see that $B_t^\eta(x)$ is a tree whp in η and x, so each path in η of length t that goes out of x is self-avoiding. By looking at pathwise probabilities, we see that

$$\mathbb{P}_{\eta,x}(\tau > t) = \sum_{x_0\cdots x_t \in \Gamma_t^\eta(x)} \mathbb{P}_{\eta,x}\big(X_{[1,t]} = x_{[1,t]}, x_{i-1} \notin R_{\le i}\, \forall\, i \in [1,t]\big). \tag{2.72}$$

Since the event $\{x_{i-1} \notin R_{\le i}\, \forall\, i \in [1,t]\}$ implies that the edge involving x_{i-1} is open a time i,

$$\mathbb{P}_{\eta,x}\big(X_{[1,t]} = x_{[1,t]} \mid x_{i-1} \notin R_{\le i}\, \forall\, i \in [1,t]\big) = \prod_{i=1}^{t} \frac{1}{\deg(x_i)}. \tag{2.73}$$

Next, let us look at the probability $\mathbb{P}_{\eta,x}(x_i \notin R_{\le i}\, \forall\, i \in [1,t])$. By rearranging and conditioning, we get

$$\begin{aligned}\mathbb{P}_{\eta,x}\big(x_{i-1} \notin R_{\le i}\, \forall\, i \in [1,t]\big) &= \mathbb{P}_{\eta,x}\big(x_j \notin R_i\, \forall\, j \in [i-1,t-1]\, \forall\, i \in [1,t]\big) \\ &= \prod_{i=1}^{t} \mathbb{P}_{\eta,x}\Big(x_j \notin R_i\, \forall j \in [i-1,t-1] \;\Big|\; x_k \notin R_j\, \forall\, k \in [j-1,t-1]\, \forall\, j \in [1,i-1]\Big).\end{aligned} \tag{2.74}$$

Observe that, on the event $\{x_k \notin R_j\, \forall\, k \in [j-1,t-1\, \forall\, j \in [1,i-1]\}$, the path $x_{i-1}\cdots x_{t-1}$ has not rewired until time $i-1$, and so the number of edges given by these half-edges is $t-i+1$, since it was originally a self-avoiding path. With this we see that for any $i \in [1,t]$,

$$\mathbb{P}_{\eta,x}\Big(x_j \notin R_i\, \forall j \in [i-1,t-1] \;\Big|\; x_k \notin R_j\, \forall\, k \in [j-1,t-1]\, \forall\, j \in [1,i-1]\Big) = \frac{\binom{m-t+i-1}{k}}{\binom{m}{k}}, \tag{2.75}$$

and hence

$$\begin{aligned}\mathbb{P}_{\eta,x}\big(x_{i-1} \notin R_{\le i} \,\forall\, i \in [1,t]\big) &= \prod_{i=1}^{t} \frac{\binom{m-t+i-1}{k}}{\binom{m}{k}} = \prod_{i=1}^{t} \frac{\binom{m-i}{k}}{\binom{m}{k}} \\ &= \prod_{i=1}^{t}\prod_{j=0}^{i-1}\Big(1 - \frac{k}{m-j}\Big) = \prod_{j=1}^{t}\Big(1 - \frac{k}{m-j+1}\Big)^{t-j+1}.\end{aligned} \tag{2.76}$$

Since $j \le t = o(\log n)$, $m = \Theta(n)$ and $n/\log^2 n = o(k)$, we have

$$\mathbb{P}_{\eta,x}\big(x_{i-1} \notin R_{\le i} \text{ for all } i \in [1,t]\big) = [1+o(1)]\,(1-k/m)^{t(t+1)/2} = (1-\alpha_n)^{t(t+1)/2} + o(1). \tag{2.77}$$

Putting this together with (2.73) and inserting it into (2.72), we get

$$\begin{aligned}\mathbb{P}_{\eta,x}(\tau > t) &= [(1-\alpha_n)^{t(t+1)/2} + o(1)] \sum_{x_0\cdots x_t \in \Gamma_t^{\eta}(x)} \prod_{i=1}^{t} \frac{1}{\deg(x_i)} \\ &= (1-\alpha_n)^{t(t+1)/2} + o(1),\end{aligned} \tag{2.78}$$

since, for each path $x_0 \cdots x_t$, the product $\prod_{i=1}^{t} \frac{1}{\deg(x_i)}$ is the probability that the random walk without backtracking on the static graph given by η follows the path, and we sum over all paths starting from x. □

Acknowledgements

The research in this paper is supported by the Netherlands Organisation for Scientific Research (NWO Gravitation Grant NETWORKS-024.002.003). The work of RvdH is further supported by NWO through VICI grant 639.033.806. The authors are grateful to Perla Sousi and Sam Thomas for pointing out an error in an earlier draft of the paper. The authors would like to thank the referees for their careful reading and fruitful comments.