RANDOM WALKS WITH LOCAL MEMORY

SWEE HONG CHAN, LILA GRECO, LIONEL LEVINE, AND PETER LI

Abstract. We prove a quenched invariance principle for a class of random walks in random environment on $\mathbb{Z}^d$, where the walker alters its own environment. The environment consists of an outgoing edge from each vertex. The walker updates the edge e at its current location to a new random edge e' (whose law depends on e) and then steps to the other endpoint of e'. We show that a native environment for these walks (i.e., an environment that is stationary in time from the perspective of the walker) consists of the wired uniform spanning forest oriented toward the walker, plus an independent outgoing edge from the walker.

1. A random environment altered by the walker

Label each site of $\mathbb{Z}^2$ with either 'H' or 'V'. A walker starts at the origin. At each discrete time step the walker resamples the label at its current location (changing 'H' to 'V' and 'V' to 'H' with probability q, independent of the past) and then takes a mean zero horizontal step if the new label is 'H' and a mean zero vertical step if the new label is 'V'. We will show (see Theorem 1.3 below) that, for a certain distribution on initial labels, the scaling limit of the walk is a standard planar Brownian motion.

The walk just described is an example of a *random walk with local memory* on a graph G. Each vertex of G stores one bit of information in its label. For each vertex x that the walk visits, the label of x remembers whether the most recently traversed outgoing edge from x was horizontal or vertical. This memory in turn affects the distribution of the edge traversed the next time the walker returns to x. One can consider more complicated forms of local memory (e.g., that remember several past visits) but they all essentially reduce to the standard *retrospective form*, i.e., each vertex x is labeled by an outgoing edge from x (see Appendix A for the reduction). At each discrete time step, the walker updates the label e of its current location to a new random edge e' (whose law depends on e) and then steps to the other endpoint of e'.

Pinsky and Travers [PT17] and Kosygina and Peterson [KP17] study random walks with local memory in one dimension under the name "Markovian cookie stacks", where the labels evolve following the transition rules of a prescribed Markov chain for each each vertex. (These Markov chains are assumed to be independent; see Travers [Tra18] for the case when they are not independent.) In particular, the latter characterizes when such a walk is recurrent, transient non-ballistic, or ballistic; and derives a central limit theorem for the transient case. The methods used in [KP17] are based on the generalized Ray-Knight theory developed by Tóth (see [Tót99] and references therein) for generalized reinforced/repelling random walks, and are limited only to one dimension. The aim of this paper is to begin the study of these walks in higher dimensions, by identifying a native environment and proving an invariance principle.

In analyzing random walks with local memory in higher dimensions, we take our inspiration from the theory of random walk in random environment [Zei04, Szn04], in which the environment affects the motion of the walker but the walker does not affect the environment. In our walks, a new difficulty is that the walker alters its own environment.

(Swee Hong Chan) Department of Mathematics, UCLA, Los Angeles CA, corresponding author.
(Lila Greco) Berkshire Hathaway Specialty Insurance, Boston MA.
(Lionel Levine) Department of Mathematics, Cornell University, Ithaca NY.
(Peter Li) Department of Economics, New York University, New York NY.
E-mail addresses: sweehong@math.ucla.edu, ecg83@cornell.edu, levine@math.cornell.edu, bl2403@nyu.edu.
Date: May 7, 2021.
2010 *Mathematics Subject Classification.* 60G42, 60F17, 60G10, 60J10, 60J65, 60K37, 82C41.
Key words and phrases. random walk, random environment, rotor walk, rotor-router, uniform spanning forest, wired spanning forest, stationary distribution, scaling limit, Brownian motion.
Lila Greco was partially supported by NSF grant DGE-1650441, Lionel Levine was partially supported by NSF grant DMS-1455272.

1.1. **Main results.** An interesting feature of random walk with local memory is that the walker organizes its environment to form a tree. Indeed, when the walk is expressed in retrospective form, the local state at each previously visited vertex is the last exit edge, so the edges at visited vertices form a tree oriented toward the walker. From this observation, it is natural to use the wired spanning forest (defined below) to construct a *native environment* (i.e., an invariant measure for the environment viewed from the perspective of the walker; see Definition 5.2).

Let G be a simple connected graph that is locally finite (i.e., each vertex has a finite degree). Let $W_1 \subseteq W_2 \subseteq \dots$ be finite connected subsets of $V(G)$ such that $\bigcup_{n=1}^{\infty} W_n = V(G)$. Let G_n be obtained from G by identifying all vertices outside W_n to one new vertex, and let μ_n be the uniform measure on spanning trees of G_n. The *wired uniform spanning forest*, denoted by WUSF, is then the unique infinite-volume limit of μ_n.

Fix a vertex o of G as the initial location of the walker. To build an initial environment from WUSF, we orient the connected component of o in the WUSF toward o, orient all other components toward infinity, and add an independent outgoing edge from the o. Note that there might be more than one way to orient a component toward infinity if it has more than one end; we will orient them using the orientation given by Wilson's method rooted toward infinity [BLPS01]. We denote by $\overrightarrow{\mathsf{WUSF}}^+$ the resulting initial environment. This environment is a native environment under the following assumptions.

Assume that G is a simple (undirected) Cayley graph of a finitely generated group. A random walk with localy memory is *transitive* if every vertex follows the same rule in updating its local memory; see (Tran). A random walk with local memory is *uniform* if, averaging over all initial labels, every outgoing edge of the current location is equally likely to be the next label. We remark that we actually prove the main results under a weaker uniformity assumption called *c-stationarity* (see (cSta)), and we only use the uniformity assumption in this section to simplify the notation.

Theorem 1.1. *Consider a random walk with local memory on a simple Cayley graph that is transitive and uniform. Then* $\overrightarrow{\mathsf{WUSF}}^+$ *is a native environment.*

We prove Theorem 1.1 in §5 by proving an analogous statement for finite graphs, and then passing to a limit.

It turns out that $\overrightarrow{\mathsf{WUSF}}^+$ satisfies a stronger property, namely that it is an *ergodic native environment* (i.e., an ergodic measure for the environment viewed from the perspective of the walker; see Definition 6.3), under the additional assumption that the random walk with local memory is *elliptic* (i.e., every neighbor of the current location is visited next with positive probability; see (Ell)).

Theorem 1.2. *Consider a random walk with local memory on a simple Cayley graph that is transitive, uniform, and elliptic. Then* $\overrightarrow{\mathsf{WUSF}}^+$ *is an ergodic native environment.*

We prove Theorem 1.2 in §6.2 through a delicate combinatorial argument that makes use of the tail triviality of WUSF. We believe that the ellipticity assumption is not necessary for the conclusion of Theorem 1.2; see §8.2.

Our next result is the following functional CLT for when G is a *lattice graph* in $\mathbb{R}^d$ (i.e., a Cayley graph such that $V(G)$ is a subgroup of $\mathbb{R}^d$ with vector addition as the group operation). A random walk with local memory is a *martingale* if, conditioned on the present location and label, the expected next location of the walker is equal to the present location; see (Mtgl).

For every outgoing edge e of the initial location o, let $Y_e \in \mathbb{R}^d$ be the location of the walker after one step of the walk, assuming e is the initial label at o. We denote by Γ_e the $d \times d$ covariance matrix $\mathbb{E}\big[Y_e Y_e^\top\big]$, and by Γ the average of covariance matrices of outgoing edges of o.

Theorem 1.3. *Consider a random walk with local memory on a simple lattice graph in* $\mathbb{R}^d$ *that is transitive, uniform, and is a martingale. Suppose that the initial environment is an ergodic native environment* π*. Then, for almost every environment sampled from* π*, the trajectory of the walker scales to a Brownian motion in* $\mathbb{R}^d$*. That is to say,*

$$\frac{1}{\sqrt{n}}(X_{\lfloor nt \rfloor})_{t\geq 0} \overset{n\to\infty}{\Longrightarrow} B(t).$$

Here $X_{\lfloor nt \rfloor}$ *is the location of the walker at the* $\lfloor nt \rfloor$*-th step of the walk,* $B(t)$ *is a Brownian motion in* $\mathbb{R}^d$ *with diffusion matrix* Γ*, and the convergence is weak convergence in the Skorohod space* $D_{\mathbb{R}^d}[0,\infty)$*.*

In particular, Theorem 1.3 applies to the 'H,V'-walk described in the beginning with q strictly between 0 and 1. We prove Theorem 1.3 in §7 by using standard tools in random walks in random environments, namely the martingale CLT and the pointwise ergodic theorem, and we illustrate the flavor with the 'H,V'-walk here. By the martingale CLT, the problem reduces to showing that the walker encounters the label 'V' half of the time, i.e.,

$$\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{1}\{\text{the label used by the walker at the } i\text{-th step is 'V'}\} \longrightarrow \frac{1}{2}, \tag{1}$$

in probability as $n \to \infty$. The convergence in (1) in turn follows from the pointwise ergodic theorem. Note that, in order to apply the pointwise ergodic theorem, the initial environment needs to be native and ergodic, and $\overrightarrow{\mathsf{WUSF}}^{+}$ is such an environment by Theorem 1.2.

Our final result is the following functional CLT, assuming a stronger regularity condition on the RWLM but requiring no condition on the initial environment. An RWLM has *identical local covariances* if $\Gamma_e = \Gamma_{e'}$ for every outgoing edge e of o.

Proposition 1.4. *Consider a random walk with local memory on a simple lattice graph in $\mathbb{R}^d$ that is transitive, is a martingale, and has identical local covariances. Then, for every initial environment,*

$$\frac{1}{\sqrt{n}}\,(X_{\lfloor nt\rfloor})_{t\geq 0} \overset{n\to\infty}{\Longrightarrow} B(t),$$

where $X_{\lfloor nt\rfloor}$ is the location of the walker at the $\lfloor nt\rfloor$-th step of the walk, and $B(t)$ is a Brownian motion in $\mathbb{R}^d$ with diffusion matrix Γ.

We prove Proposition 1.4 (under slightly weaker assumptions) in §3 as a direct application of the martingale CLT. In particular, Proposition 1.4 applies to the random walk with local memory on the triangular lattice where the mechanism is rotating the current outgoing edge by 60 degrees, 180 degrees, or 300 degrees, each with probability $\frac{1}{3}$; see Example 2.6. On the other hand, Proposition 1.4 does *not* apply to 'H,V'-walk if $q \neq \frac{1}{2}$ (since Γ_e is equal to $\begin{bmatrix} 1-q & 0 \\ 0 & q \end{bmatrix}$ if e is a horizontal edge, and is equal to $\begin{bmatrix} q & 0 \\ 0 & 1-q \end{bmatrix}$ if e is a vertical edge). This necessitates results such as Theorem 1.3 that has weaker assumptions and does apply to a family of models that include 'H,V'-walk.

1.2. **Other related work.**

1.2.1. When each vertex uses a deterministic rule to update its local memory, the random walk with local memory is known as *rotor walk* (discovered independently by [WLB96, PDDK96, Pro03]). In this model, each vertex is given a prescribed cyclic ordering on its outgoing edges, and for every update the vertex changes the current edge to the next edge in the cyclic order. A fundamental difficulty with rotor walk is its lack of randomness: For example, it is an open problem to prove that the rotor walk in $\mathbb{Z}^2$ with i.i.d. uniform initial rotors is recurrent; see [HLM+08, FLP16] for an exposition of this and related problems.

1.2.2. One dimensional random walk with local memories are more commonly studied in the literature under the name *excited random walks* (introduced by Benjamini and Wilson [BW03]): A pile of cookies is initially placed at each vertex of $\mathbb{Z}^d$ ($d \geq 1$). Upon visiting a vertex, the walker consumes the topmost cookie from the pile and moves to the neighboring vertex according to probabilities prescribed by that cookie. If there are no cookies left at the current vertex, the walker chooses a neighbor uniformly at random and moves there.

The functional limit theorem for excited random walks on $\mathbb{Z}$ have been studied for the case of bounded number of i.i.d. cookies [KM11, DK12], periodic cookies [KP16], and Markovian cookies [KP17, HLSH18], among others. The functional limit theorem for higher-dimensional walks are much rarer in comparison. Nevertheless, it has been studied for the case of a single cookie with drift to a specific direction by [VdHH12] (for dimensions $d > 8$ and a specific drift intensity), by [BR07] (for all dimensions), and by [MPRV12] (for all dimensions under more general assumptions). We refer the reader to [KZ13] for an excellent survey on excited random walks. Finally, in the direction of non-Markovian walks, the most relevant recent work is [BL19], which applies martingale theory to higher-dimensional elephant random walks.

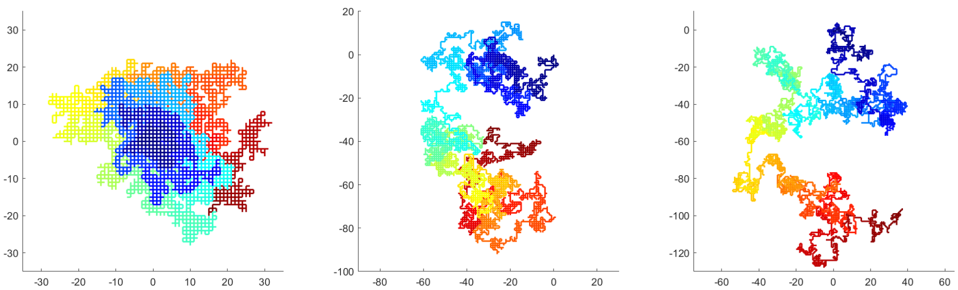

FIGURE 1. The vertices visited by a 10,000-step rotor walk (left), 'H'-'V' walk with $q = 1$ (middle), and simple random walk (right) on $\mathbb{Z}^2$; these processes are ordered in increasing amount of randomness. Each edge is colored according to the time of its first visit by the walker.

The main motivation of this paper is to begin extending the results of [KP17, HLSH18] from dimension one to higher dimensions, which we partly achieve in Theorem 1.3. In particular, it is shown in [HLSH18] that the scaling limit for p-rotor walk in $\mathbb{Z}$ (where the next edge points in the same direction as the current edge with probability p, and points in the opposite direction with probability $1-p$) is a Brownian motion perturbed at extrema. This perturbation is caused by the initial environment in [HLSH18] not being a native environment. We expect that proving a scaling limit for any higher-dimensional random walk with local memory in a non-native environment will require major new ideas (for example, what are the planar and higher-dimensional analogues of the one-dimensional Brownian motion perturbed at extrema?).

1.2.3. A *self-interacting random walk* (SIRW) is a nearest-neighbour walk on $\mathbb{Z}^d$, where at each step the probability of the walker to jump along a certain direction α is proportional to $w(n_\alpha)$, where $w : \mathbb{N} \to \mathbb{R}_{>0}$ is a monotone weight function and n_α is the number of previous jumps along the direction α. Unlike random walks with local memory, the transition probabilities for SIRW depend on all of the previous visits to the current location rather than just the most recent visit. Various limit theorems for various one-dimensional SIRWs were studied by Tóth (see e.g., [Tót95, Tót96]), and we refer to the survey [Tót99] for references on this subject. It remains to be seen if the methods of this paper can be applied to SIRWs in higher dimensions.

1.2.4. The idea of viewing the environment from the perspective of the walker dates back to the work of Kozlov [KOZ85] and Papanicolaou-Varadhan [PV82]. It was pointed to us by the anonymous referee that it would be interesting to find out what are the native environments for SIRWs. We refer the reader to [BS02, Lecture 1] for references on this subject.

1.2.5. Random walk with local memory is a special case of the *stochastic abelian networks* defined in [BL16]. More precisely, a random walk with local memory is a unary network in which every processor sends exactly one letter of output for each letter of input. From this perspective, a general stochastic abelian network can then be viewed as a branching random walk with local memory with multiple types of walkers.

1.3. **Outline.** In §2 we give the rigorous definition of random walks with local memory. In §3 we prove Proposition 1.4. In §4 we construct the wired spanning forest oriented toward a fixed vertex, which is a simple modification of the construction in [BLPS01]. In §5 we use the oriented wired spanning forest from §4 to construct a native environment for random walk with local memory, and proves Theorem 1.1. In §6 we prove Theorem 1.2. In §7 we prove Theorem 1.3. In §8 we conclude with a list of open problems. In Appendix A we show the reduction that converts random walks with more complicated forms of local memory to the standard retrospective form, at the cost of changing the underlying graph to a larger graph that might have multiple edges.

2. Random walks with local memory

Throughout this paper $G := (V(G), E(G))$ denotes a connected, undirected graph that is locally finite (every vertex has finite degree) and simple (no loops, no multiple edges). We remark that all the results in this paper can be extended to non-simple graphs verbatim; and we simply restrict to the case of simple graphs to simplify the notation. When the graph G is evident from context, we will omit G from the notation and write V and E instead.

A *neighbor* of a vertex x is a vertex y such that $\{x, y\} \in E$. We denote by $N(x)$ the set of all neighbors of x. An *oriented edge* of G is a pair $(x, y) \in V \times V$ such that $\{x, y\}$ is an (unoriented) edge of G. We call (x, y) an *outgoing* edge of x and an *incoming* edge of y. In an oriented subgraph of G, the *outdegree* (respectively, *indegree*) of x is the number of outgoing (respectively, *incoming*) edges of x in the oriented subgraph. We denote by $\vec{E}$ the set of oriented edges of G. The running example for a graph in this paper is the integer lattice $\mathbb{Z}^d$ of dimension d, i.e., the graph given by

$$V := \{\mathbf{x} \mid \mathbf{x} \in \mathbb{Z}^d\}; \qquad E := \{\{\mathbf{x}, \mathbf{y}\} \in \mathbb{Z}^d \times \mathbb{Z}^d \mid ||\mathbf{x} - \mathbf{y}|| = 1\},$$

where $||\cdot||$ denotes the Euclidean norm.

Definition 2.1 (Mechanism). A *mechanism* of a random walk with local memory is a collection of independent Markov chains $\{M_x\}_{x \in V}$ indexed by the vertices of G, such that the state space of M_x is $N(x)$, the set of neighbors of x. We denote by $p_x(\cdot, \cdot)$ the *probability transition function* of the chain M_x.

A *rotor configuration* of G is a map $\rho : V \to V$ such that $\rho(x)$ is a neighbor of x for all $x \in V$. This should be thought of as assigning to each vertex x of G a *rotor* which points to a neighbor of x via an oriented edge of G. A *walker-and-rotor configuration* is a pair (x, ρ), where x is a vertex of G and ρ is a rotor configuration of G.

Remark 2.2. A rotor configuration can be interpreted as either:

- A function $\rho : V \to V$ such that $\rho(x) \in N(x)$ for all $x \in V$; or
- An oriented subgraph of G that has exactly one outgoing edge of each vertex of G.

These two objects are identified with each other by the map $\rho \mapsto (V(\rho), E(\rho))$, where

$$V(\rho) := V, \qquad E(\rho) := \{(x, \rho(x)) \mid x \in V\}.$$

We would like to warn the reader that both interpretations will be used interchangeably starting from §5.

Definition 2.3 (Random walk with local memory). A *random walk with local memory*, or RWLM for short, is a sequence $(X_n, \rho_n)_{n \geq 0}$ of walker-and-rotor configurations satisfying the following transition rules:

$$\begin{aligned} \rho_{n+1}(x) :=& \begin{cases} Y_n & \text{if } x = X_n; \\ \rho_n(x) & \text{if } x \neq X_n. \end{cases}; \text{ and} \\ X_{n+1} :=& Y_n, \end{aligned} \tag{2}$$

where Y_n is a random neighbor of X_n sampled from $p_{X_n}(\rho_n(X_n), \cdot)$ independent of the past.

Described in words, X_n records the location of the walker and ρ_n records the rotor configuration at time n of the RWLM. At time n, the walker updates the rotor of X_n using the Markov chain M_{X_n} (which depends only on X_n and $\rho_n(X_n)$), and then moves to the vertex to which the new rotor is pointing. The *local memory* in the name refers to the fact that the walker records the last exit from each vertex that it visits via the rotor configuration. See Figure 2 for an illustration of an RWLM on $\mathbb{Z}^2$.

Naturally, the dynamics of the RWLM depend on the choice of the mechanism. The following are three examples of RWLMs that have appeared in the literature:

(i) Aldous-Broder walk, in which the walker performs a simple random walk on G and the rotor configuration never influences the decision of the walker. That is to say, for every $x \in V$ and $y \in N(x)$ the measure $p_x(y, \cdot)$ is the uniform distribution on the neighbors of x. Our name for this walk comes from the algorithm of Aldous [Ald90] and Broder [Bro89] that generates the uniform spanning tree as a tree of first entrances of this walk.

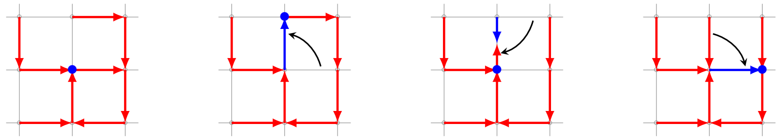

FIGURE 2. Three steps of a random walk with local memory on $\mathbb{Z}^2$. The location of the walker is given by $\bullet$, and the rotor of each vertex is given by the arrow pointing out from the vertex.

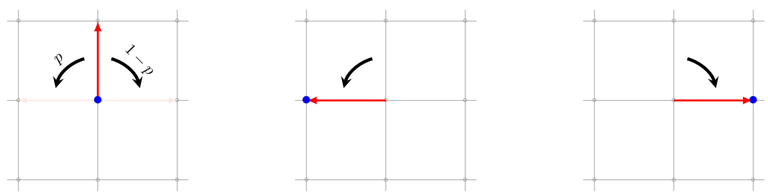

FIGURE 3. (a) The mechanism for p-rotor walk on $\mathbb{Z}^2$, in which the rotor rotates counterclockwise with probability p, and clockwise with probability $1-p$. The location of the walker and the rotor after one step of the RWLM is given by (b) if the walker chooses to rotate the rotor counterclockwise, and by (c) if the walker chooses to rotate the rotor clockwise.

(ii) Rotor walk [WLB96, PDDK96, Pro03], in which the Markov chain M_x is given by a deterministic permutation τ_x of the neighbors of x. That is, the chain M_x in state y will transition to $\tau_x(y)$ with probability 1. We refer to [HLM+08, FLP16] for more details.

(iii) p-rotor walk on $\mathbb{Z}$ [HLSH18] for $p \in [0,1]$, in which the probability transition function p_x ($x \in \mathbb{Z}$) is given by

$$p_x(x \pm 1, x \mp 1) = 1 - p; \qquad p_x(x \pm 1, x \pm 1) = p.$$

We now present three other examples of RWLMs.

Example 2.4 (p-rotor walk on $\mathbb{Z}^d$). Fix $d \geq 2$ and $p \in [0,1]$. Denote by $\mathbf{e}_1, \ldots, \mathbf{e}_d$ the canonical basis of $\mathbb{R}^d$. The Markov chain $M_{\mathbf{x}}$ ($\mathbf{x} \in \mathbb{Z}^d$) has state space $\{\mathbf{x} \pm \mathbf{e}_i \mid 1 \leq i \leq d\}$ and has the following transition rule:

$$\mathbf{x} \pm \mathbf{e}_i \quad \text{transitions to} \quad \begin{cases} \mathbf{x} \pm \mathbf{e}_j & \text{with probability } \frac{p}{d-1} \text{ if } i < j; \\ \mathbf{x} \mp \mathbf{e}_j & \text{with probability } \frac{1-p}{d-1} \text{ if } i < j; \\ \mathbf{x} \pm \mathbf{e}_j & \text{with probability } \frac{1-p}{d-1} \text{ if } i > j; \\ \mathbf{x} \mp \mathbf{e}_j & \text{with probability } \frac{p}{d-1} \text{ if } i > j. \end{cases}$$

Described in words, if the rotor at the particle's current location is parallel to $\mathbf{e}_i$, the walker first picks j uniformly from $\{1, \ldots, d\} \setminus \{i\}$. Then, the walker rotates the current rotor counterclockwise in the $\{\min(i,j), \max(i,j)\}$-plane with probability p, and rotates clockwise with probability $1-p$. See Figure 3 for an illustration of this mechanism on $\mathbb{Z}^2$.

Example 2.5 (p,r-rotor walk on $\mathbb{Z}^d$). Fix $d \geq 2$, $p \in [0,1]$, and $r \in [0,1]$. For each visit to $\mathbf{x} \in \mathbb{Z}^d$, the mechanism at $\mathbf{x}$ transitions according to the mechanism of Aldous-Broder walk with probability $1-r$, and transitions according to the mechanism of p-rotor walk with probability r, independent of the past visits. Note that we recover 'H,V'-walk on $\mathbb{Z}^2$ for $q \leq \frac{1}{2}$ in §1 by taking $p = \frac{1}{2}$ and $r = 1 - 2q$. Also note that, unlike p-rotor walks, in this model every neighbor of the current location of the walker (all $2d$ of them) is visited next with positive probability provided that $r < 1$ (i.e., the walk is elliptic). See Figure 4 for an illustration of this mechanism.

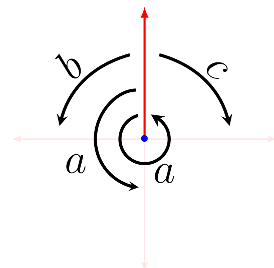

FIGURE 4. The mechanism for p,r-rotor walk on $\mathbb{Z}^2$, which stays at the current rotor with probability $a := \frac{1-r}{4}$, rotates 180 degrees with probability a, rotates 90 degrees counterclockwise with probability $b := \frac{1-r}{4} + pr$, and rotates 90 degrees clockwise with probability $c := \frac{1-r}{4} + (1-p)r$.

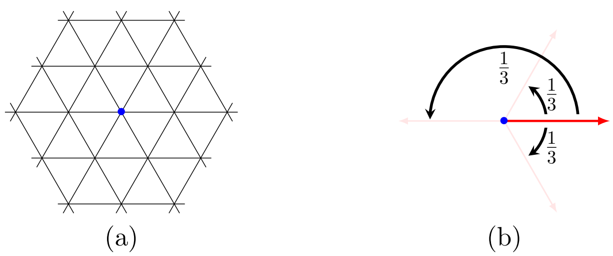

FIGURE 5. (a) The triangular lattice. (b) The mechanism for the triangular lattice, which rotates either 60 degrees counterclockwise, 180 degrees counterclockwise, or 300 degrees counterclockwise, each with probability $\frac{1}{3}$.

Example 2.6 (Triangular walk). The triangular lattice is the graph embedded in $\mathbb{R}^2$ given by:

$$V := \left\{ a\begin{pmatrix}1\\0\end{pmatrix} + b\begin{pmatrix}1/2\\ \sqrt{3}/2\end{pmatrix} \middle| \, a, b \in \mathbb{Z} \right\};$$
$$E := \{\{\mathbf{x}, \mathbf{y}\} \in V \times V \mid \|\mathbf{x} - \mathbf{y}\| = 1\}.$$

In this RWLM, the walker updates the current rotor by applying a counterclockwise rotation by either 60 degrees, 180 degrees, or 300 degrees, each with probability $\frac{1}{3}$. See Figure 5 for an illustration of this mechanism.

3. Martingale central limit theorem

In this section we show that, under strong regularity assumptions on the RWLM, we can directly prove functional CLT from the vector-valued martingale CLT proved in [RAS05]. We denote by $D_{\mathbb{R}^d}[0,\infty)$ the Skorohod space of $\mathbb{R}^d$-valued càdlàg paths on $[0,\infty)$. Recall that $||\cdot||$ denotes the Euclidean metric.

Theorem 3.1 (Martingale CLT [RAS05, Theorem 3]). *Let $(X_n)_{n\geq 0}$ be an $\mathbb{R}^d$-valued square-integrable martingale process w.r.t. a filtration $(\mathscr{F}_n)_{n\geq 0}$, and let $V_n := X_{n+1} - X_n$ be the corresponding martingale difference sequence. Suppose that:*

(i) *There exists a symmetric, nonnegative definite $d \times d$ matrix Γ such that*

$$\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\big[V_i V_i^\top \mid \mathscr{F}_i\big] \to \Gamma \qquad \text{in probability as } n \to \infty; \tag{CLT1}$$

(ii) *For any $\epsilon > 0$,*

$$\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\big[\|V_i\|^2 \, \mathbb{1}\{\|V_i\| \geq \epsilon\sqrt{n}\} \mid \mathscr{F}_i\big] \to 0 \qquad \text{in probability as } n \to \infty. \tag{CLT2}$$

Then $\left\{\frac{1}{\sqrt{n}}X_{\lfloor nt\rfloor}, t\geq 0\right\}$ *converges weakly on* $D_{\mathbb{R}^d}[0,\infty)$ *to a Brownian motion with diffusion matrix* Γ. □

We now apply Theorem 3.1 to RWLMs under the following assumptions. Let $G=(V,E)$ be a simple connected graph such that V is a subset of $\mathbb{R}^d$. An RWLM is *bounded* if

$$\sup_{\{\mathbf{x},\mathbf{y}\}\in E} ||\mathbf{x}-\mathbf{y}|| < \infty; \tag{Bdd}$$

All the RWLMs described in §2 are bounded.

Recall the definition of probability transition functions $p_{\mathbf{x}}$ from Definition 2.1. Let $\mathbf{x}$ be a vertex of G, and let $\mathbf{y}$ be a neighbor of $\mathbf{x}$. We denote by $Y_{\mathbf{x},\mathbf{y}}$ the random variable sampled from $p_{\mathbf{x}}(\mathbf{y},\cdot)$. The *local covariance matrix* of $\mathbf{x},\mathbf{y}$ is the $d\times d$ matrix $\Gamma_{\mathbf{x},\mathbf{y}} := \mathbb{E}\big[(Y-\mathbf{x})(Y-\mathbf{x})^\top\big]$.

We say that an RWLM is a *martingale* if

$$\mathbb{E}[Y_{\mathbf{x},\mathbf{y}}] \;=\; \mathbf{x} \qquad \text{for every } \mathbf{x}\in V \text{ and } \mathbf{y}\in N(\mathbf{x}). \tag{Mtgl}$$

Note that this condition is equivalent to requiring the sequence $(X_n)_{n\geq 0}$ of locations of walker of the RWLM to be a martingale. The Aldous-Broder walk on $\mathbb{Z}^d$ and the triangular walk (Example 2.6) is a martingale, the deterministic rotor walk is *not* a martingale, and the p-rotor walk (Example 2.4) and p,r-rotor walk (Example 2.5) are martingales only if $p=\frac{1}{2}$.

We say that an RWLM has *identical local covariances* if

$$\Gamma_{\mathbf{x},\mathbf{y}} \;=\; \Gamma_{\mathbf{x}',\mathbf{y}'} \qquad \text{for every } x,x'\in V \text{ and } \mathbf{y}\in N(\mathbf{x}), \mathbf{y}'\in N(\mathbf{x}'), \tag{ILC}$$

and in this case we write $\Gamma := \Gamma_{\mathbf{x},\mathbf{y}}$. Aldous-Brouder walk on $\mathbb{Z}^d$ and triangular walk are the only RWLMs from §2 for which (ILC) holds. The matrix Γ is equal to $\frac{1}{d}I_d$ (where I_d is the $d\times d$ identity matrix) in the former case, and is equal to $\begin{bmatrix}\frac{1}{2} & 0\\ 0 & \frac{1}{6}\end{bmatrix}$ in the latter case. The p-rotor walk does *not* satisfy (ILC) as the covariance matrix $\Gamma_{\mathbf{x},\mathbf{y}}$ is equal to $\frac{1}{d-1}(I_d-\mathbf{e}_i\mathbf{e}_i^\top)$, where $\mathbf{e}_i$ is the standard unit vector parallel to the edge $(\mathbf{x},\mathbf{y})$. The p,r-rotor walk (with $r>0$) does not satisfy (ILC) either by an analogous calculation.

We now restate Proposition 1.4 from the introduction in a slightly more general form.

Proposition 1.4. Let G be a simple, connected graph with its vertex set being a subset of $\mathbb{R}^d$. Consider an RWLM on G that satisfies (Bdd), (Mtgl), and (ILC). Then the scaled walk $\left\{\frac{1}{\sqrt{n}}X_{\lfloor nt\rfloor}, t\geq 0\right\}$ converges weakly on $D_{\mathbb{R}^d}[0,\infty)$ to a Brownian motion with diffusion matrix Γ.

The remarkable part of Proposition 1.4 is that the conditions involve only the mechanism of the RWLM, and hence we can derive a scaling limit result *regardless* of the initial walker-and-rotor configuration. In particular, it follows from Proposition 1.4 that, for every initial walker-and-rotor configuration, the triangular walk from Example 2.6 satisfies a functional CLT.

Naturally, Proposition 1.4 does not apply to p-rotor walk and p,r-rotor walk even when $p=\frac{1}{2}$, as (ILC) is never satisfied. Thus we need a different approach to prove a scaling limit for these RWLMs, which we partially achieve at the cost of starting the walk with a specific rotor configuration; see Theorem 1.3.

Proof of Proposition 1.4. It suffices to check that all conditions of Theorem 3.1 are satisfied. Write $C := \sup_{\{\mathbf{x},\mathbf{y}\}\in E}||\mathbf{x}-\mathbf{y}||$. Note that C is finite by (Bdd). This implies that $||X_n||\leq Cn+||X_0||$ for all $n\geq 0$, and it then follows that $(X_n)_{n\geq 0}$ is square-integrable.

We now check that $(X_n)_{n\geq 0}$ is a martingale process with respect to the filtration $\mathscr{F}_n := \sigma(X_0,\ldots,X_n,\rho_0,\ldots,\rho_n)$. It then follows from the transition rule of RWLM (see (2)) that, for any $n\geq 0$:

$$\begin{aligned}
\mathbb{E}[X_{n+1}\mid \mathscr{F}_n] &= \sum_{\mathbf{x}\in V}\sum_{\mathbf{y}\in N(\mathbf{x})} \mathbb{E}[Y_{\mathbf{x},\mathbf{y}}\,\mathbb{1}\{X_n=\mathbf{x},\rho_n(\mathbf{x})=\mathbf{y}\}\mid \mathscr{F}_n] \qquad \text{(by Definition 2.3)}\\
&= \sum_{\mathbf{x}\in V}\sum_{\mathbf{y}\in N(\mathbf{x})} \mathbb{E}[Y_{\mathbf{x},\mathbf{y}}]\,\mathbb{1}\{X_n=\mathbf{x},\rho_n(\mathbf{x})=\mathbf{y}\}\\
&= \sum_{\mathbf{x}\in V}\sum_{\mathbf{y}\in N(\mathbf{x})} \mathbf{x}\,\mathbb{1}\{X_n=\mathbf{x},\rho_n(\mathbf{x})=\mathbf{y}\} \qquad \text{(by (Mtgl))}\\
&= X_n.
\end{aligned}$$

This shows that $(X_n)_{n\geq 0}$ is a martingale.

We now check the condition (CLT1). It follows from the the transition rule of RWLM that, for any $n \geq 0$:

$$\begin{aligned}\mathbb{E}\big[V_n V_n^\top \mid \mathscr{F}_n\big] &= \sum_{\mathbf{x}\in V}\sum_{\mathbf{y}\in N(\mathbf{x})} \mathbb{E}\big[(Y_{\mathbf{x},\mathbf{y}}-\mathbf{x})(Y_{\mathbf{x},\mathbf{y}}-\mathbf{x})^\top \mathbb{1}\{X_n=\mathbf{x},\rho_n(\mathbf{x})=\mathbf{y}\} \mid \mathscr{F}_n\big]\\ &= \sum_{\mathbf{x}\in V}\sum_{\mathbf{y}\in N(\mathbf{x})} \mathbb{E}\big[(Y_{\mathbf{x},\mathbf{y}}-\mathbf{x})(Y_{\mathbf{x},\mathbf{y}}-\mathbf{x})^\top\big] \mathbb{1}\{X_n=\mathbf{x},\rho_n(\mathbf{x})=\mathbf{y}\}\\ &= \sum_{\mathbf{x}\in V}\sum_{\mathbf{y}\in N(\mathbf{x})} \Gamma\,\mathbb{1}\{X_n=\mathbf{x},\rho_n(\mathbf{x})=\mathbf{y}\} \qquad \text{(by (ILC))}\\ &= \Gamma.\end{aligned}$$

It then follows that $\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}\big[V_iV_i^\top \mid \mathscr{F}_i\big] = \Gamma$, which proves (CLT1).

We now check the condition (CLT2). Note that

$$||V_n|| = ||X_{n+1}-X_n|| \leq \sup_{\{\mathbf{x},\mathbf{y}\}\in E} ||\mathbf{x}-\mathbf{y}|| < \infty,$$

where the last inequality is due to (Bdd). Hence for any $\epsilon > 0$, for sufficiently large n we have that $\mathbb{1}\{\|V_i\| \geq \epsilon\sqrt{n}\} = 0$ for every $i \geq 0$. This implies that

$$\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}\big[\|V_i\|^2\mathbb{1}\{\|V_i\|\geq \epsilon\sqrt{n}\} \mid \mathscr{F}_i\big] = 0,$$

which proves (CLT2). The proof is now complete. □

4. Wired spanning forest oriented toward a root

In this section we present two methods to generate the wired spanning forest oriented toward a chosen root vertex, which we will use to construct an initial rotor configuration for random walks with local memory in §5 and §7. Most of the material in this section is not new. Indeed, the material in §4.1 and §4.2 is taken from the relevant part of [LP16], and the material in §4.3 and §4.4 is a straightforward modification of Wilson's method [Wil96, BLPS01], which we spell out for completeness.

4.1. **Unoriented wired spanning forest.** We begin by defining the unoriented wired spanning forest, and we refer to [BLPS01] and [LP16, Chapters 4 & 10] for a detailed discussion on this topic.

Recall that $G := (V(G), E(G))$ is a simple, connected, undirected graph that is locally finite. Let $\mathscr{F} := \mathscr{F}(G)$ be the σ-algebra on the set of subgraphs of G generated by sets of the form $\{H \in 2^{E(G)} \mid B \subseteq H\}$, where B is a finite subset of $E(G)$. The unoriented wired spanning forest will be a probability distribution on the measurable space $(2^{E(G)}, \mathscr{F}(G))$.

An *electrical network* is a pair (G, c), where G is a locally finite, simple, connected graph, and the *conductance* $c : E \to \mathbb{R}_{>0}$ is a function that sends each unoriented edge of G to a positive real number. We denote by $c\{x, y\}$ the conductance of the unoriented edge $\{x, y\}$. (We emphasize that G is always an **unoriented** graph, and $c\{x, y\} = c\{y, x\}$.)

We associate to each (G, c) the Markov chain with state space $V(G)$ and such that, for every adjacent vertices x, y, the probability to transition from x to y is proportional to $c\{x, y\}$. This Markov chain is called the *network random walk* on (G, c). The network (G, c) is *recurrent* if the network random walk eventually returns to its starting point with probability 1, and is *transient* otherwise.

We start by defining the wired spanning forest for the network (G, c) when G is a finite graph, in which case the distribution is concentrated on the spanning trees of G. The *weight* of a finite subgraph H of G is

$$\Xi(H) := \prod_{\{x,y\}\in E(H)} c\{x, y\}.$$

Definition 4.1 (Unoriented spanning forest for finite graphs). For a finite graph G, the *unoriented wired spanning forest* $\mathsf{WSF} := \mathsf{WSF}(G, c)$ is the probability distribution on spanning trees of G in which each tree T is picked with probability proportional to $\Xi(T)$.

Note that the term "wired" is not usually present in Definition 4.1 when G is finite, as wired exhaustion (see Definition 4.2 below) is not a required concept here. In fact, in this case the wired spanning forest will always be a tree. However, using the terms "wired" and "forest" will significantly simplify the notation in this paper, as our results apply to both finite and infinite graphs.

Definition 4.2 (Wired exhaustion). Let $(W_n)_{n\geq 0}$ be a sequence of finite, connected subsets of $V(G)$ such that

- $\bigcup_{n\geq 0} W_n = V(G)$; and
- $W_n \subseteq W_{n+1}$ for all $n \geq 0$.

The *wired exhaustion* of G is the sequence of electrical networks $(G_n, c_n)_{n\geq 0}$ defined as follows. The graph G_n is the undirected graph obtained from G by identifying all the vertices of $V(G) \setminus W_n$ to a single vertex z_n and removing loops and extra multiple edges that are formed. The conductance $c_n : E(G_n) \to \mathbb{R}_{>0}$ is defined by

$$c_n\{x, y\} := \begin{cases} c\{x, y\} & \text{if } x, y \in W_n; \\ \sum_{y' \notin W_n} c\{x, y'\} & \text{if } x \in W_n \text{ and } y = z_n. \end{cases}$$

We denote by μ_n the probability distribution $\mathsf{WSF}(G_n, c_n)$ on the subgraphs of G_n. We can now define the wired spanning forest for infinite graphs using the concept of wired exhaustion.

Definition 4.3 (Unoriented wired spanning forest for infinite graphs). The *wired spanning forest* $\mathsf{WSF} := \mathsf{WSF}(G, c)$ is a probability distribution on subgraphs of G such that, for any wired exhaustion and any finite $B \subseteq E(G)$,

$$\mathsf{WSF}[B \subseteq F] = \lim_{n\to\infty} \mu_n[B \subseteq T_n], \tag{3}$$

where F is a random subgraph of G distributed according to WSF, and T_n is a random spanning tree of G_n distributed according to μ_n.

The quantity $\mu_n[B \subseteq T_n]$ decreases as $n \to \infty$ [LP16, Chapter 10], and hence the limit in (3) exists and does not depend on the choice of the wired exhaustion. By the Kolmogorov extension theorem, there exists a unique probability distribution on $(2^{E(G)}, \mathscr{F}(G))$ that satisfies (3).

The random subgraph sampled from WSF is always a spanning forest but not necessarily a spanning tree. It is well-known that, for the graph $\mathbb{Z}^d$ with a constant conductance, this random subgraph has one connected component a.s. if $d \leq 4$, and infinitely many connected components a.s. if $d \geq 5$ [Pem91, Theorem 4.2]. For more on the geometry of the WSF and its dependence on dimension, see [BKPS04, HP19].

4.2. **Wired spanning forest oriented toward a root.** We now define the wired spanning forest oriented toward a chosen root vertex. Denote by

$$\overrightarrow{E}(G) := \bigcup_{\{x,y\}\in E(G)} \{(x, y), (y, x)\}$$

the set of oriented edges of G. Let $\overrightarrow{\mathscr{F}} := \overrightarrow{\mathscr{F}}(G)$ be the σ-algebra on the set of oriented subgraphs of G generated by sets of the form $\{\overrightarrow{H} \in 2^{\overrightarrow{E}(G)} \mid \overrightarrow{B} \subseteq \overrightarrow{H}\}$, where $\overrightarrow{B}$ is a finite subset of $\overrightarrow{E}(G)$. The oriented wired spanning forest will be a probability distribution on the measurable space $(2^{\overrightarrow{E}(G)}, \overrightarrow{\mathscr{F}}(G))$.

We start by defining the oriented wired spanning forest when G is a finite graph, in which case the distribution is concentrated on the oriented spanning trees of G. Fix a root vertex $r \in V(G)$ for the rest of this section.

Definition 4.4 (Oriented spanning tree). An *r-oriented spanning tree* $\overrightarrow{T}$ of G is an oriented subgraph of G such that, for any $x \in V(G)$, there exists a unique directed path in $\overrightarrow{T}$ that starts at x and ends at r.

Note that in an r-oriented spanning tree $\overrightarrow{T}$, every vertex in $V(G) \setminus \{r\}$ has outdegree 1 in $\overrightarrow{T}$, and the root vertex r has outdegree 0 in $\overrightarrow{T}$. Also note that given an unoriented spanning tree of a finite graph and a root vertex r, there is a unique way to orient the tree to become an r-oriented spanning tree. The *weight* of a finite oriented subgraph $\overrightarrow{H}$ of G is

$$\Xi(\overrightarrow{H}) := \prod_{(x,y)\in\overrightarrow{E}(\overrightarrow{H})} c\{x, y\}.$$

Definition 4.5 (Rooted oriented wired spanning forest for finite graphs). Let G be a finite graph. The *r-oriented wired spanning forest*, denoted $\overrightarrow{\mathsf{WSF}}_r := \overrightarrow{\mathsf{WSF}}_r(G, c)$, is the probability distribution on r-oriented spanning trees of G in which each tree $\overrightarrow{T}$ is picked with probability proportional to $\Xi(\overrightarrow{T})$.

We now define the r-oriented wired spanning forest for infinite graphs G. Let $(G_n, c_n)_{n\geq 0}$ be a wired exhaustion of G. We denote by $\overrightarrow{\mu_{r,n}}$ the probability distribution $\overrightarrow{\mathsf{WSF}}_r(G_n, c_n)$ on the oriented subgraphs of G_n.

Definition 4.6 (Rooted oriented wired spanning forest for infinite graphs). The *r-oriented wired spanning forest*, denoted $\overrightarrow{\mathsf{WSF}}_r := \overrightarrow{\mathsf{WSF}}_r(G, c)$, is a probability distribution on oriented subgraphs of G such that, for any wired exhaustion and any finite $\overrightarrow{B} \subseteq \overrightarrow{E}(G)$,

$$\overrightarrow{\mathsf{WSF}}_r[\overrightarrow{B} \subseteq \overrightarrow{F}] = \lim_{n\to 0} \overrightarrow{\mu_{r,n}}[\overrightarrow{B} \subseteq \overrightarrow{T_n}], \tag{4}$$

where $\overrightarrow{F}$ is a random oriented subgraph of G distributed according to $\overrightarrow{\mathsf{WSF}}_r$ and $\overrightarrow{T_n}$ is a random r-oriented tree of G_n distributed according to $\overrightarrow{\mu_{r,n}}$.

The limit in Definition 4.6 exists and does not depend on the choice of the wired exhaustion as we will see in §4.3 (for recurrent networks) and §4.4 (for transient networks). By the Kolmogorov extension theorem, there exists a unique probability distribution on $(2^{\overrightarrow{E}(G)}, \overrightarrow{\mathscr{F}}(G))$ that satisfies Definition 4.6.

The underlying graph of the r-oriented wired spanning forest is the unoriented wired spanning forest, in the following sense.

Lemma 4.7. *Let $f : 2^{\overrightarrow{E}(G)} \to 2^{E(G)}$ be the map that takes an oriented subgraph and erases the orientation of every edge. If $\overrightarrow{F}$ is an oriented subgraph of G sampled from $\overrightarrow{\mathsf{WSF}}_r$, then $f(\overrightarrow{F})$ is an unoriented subgraph of G that has the law of* WSF.

Proof. Note that, for any finite subset B of $E(G)$, the event $\{B \subseteq f(\overrightarrow{F})\}$ depends only on finitely many oriented edges. Therefore, it suffices to consider the case when G is a finite graph, as the case of infinite graphs follows by taking the limit over a wired exhaustion and then verifying the lemma for all events of the form $\{B \subseteq f(\overrightarrow{F})\}$ for some finite B.

When G is finite, note that f is a bijection between r-oriented spanning trees of G and unoriented spanning trees of G that preserves the weight of spanning trees. The lemma now follows from Definition 4.1 and Definition 4.5, and the proof is complete. ☐

As in the unoriented case, a random oriented subgraph $\overrightarrow{F}$ sampled from $\overrightarrow{\mathsf{WSF}}_r$ is not necessarily an oriented spanning tree. However, it is always an *r-oriented spanning forest* of G: the underlying graph of $\overrightarrow{F}$ is a spanning forest, every vertex in $V(G) \setminus \{r\}$ has outdegree 1 in $\overrightarrow{F}$, and r has outdegree 0 in $\overrightarrow{F}$. The first condition follows from Lemma 4.7, and the others can be verified directly from the limit in Definition 4.6 as these events only depend on finitely many edges.

4.3. **Wilson's method oriented toward a root: recurrent case.** In this subsection we describe an algorithm due to Wilson [Wil96] that generates $\mathsf{WSF}(G, c)$ and $\overrightarrow{\mathsf{WSF}}_r(G, c)$ for recurrent networks without using the weak limit in Definition 4.6.

A *(finite) directed walk* in G is a sequence $\langle x_0, \dots, x_n\rangle$ such that $\{x_i, x_{i+1}\} \in E(G)$ for $i \in \{0, \dots, n-1\}$. The *loop erasure* of a directed walk $\langle x_0, \dots, x_n\rangle$, denoted by $\mathsf{LE}\langle x_0, \dots, x_n\rangle$, is obtained by erasing cycles in the directed walk in the order they appear, i.e., it is the directed walk given by the following recursive definition. Let $y_0 := x_0$. Suppose that y_i has been defined, and let j be the largest element of $\{0, \dots, n\}$ such that $x_j = y_i$. Set $y_{i+1} := x_{j+1}$ if $j < n$; otherwise, define $\mathsf{LE}\langle x_0, \dots, x_n\rangle := \langle y_0, \dots y_i\rangle$. Note that even if the directed walk is infinite, its loop erasure is still well-defined provided that the walk is *locally finite*, i.e., every vertex is visited at most finitely many times in the walk.

Definition 4.8 (Wilson's method for recurrent networks). Let (G, c) be a recurrent network. Let $x_1, x_2, \dots$ be an ordering of elements of the $V(G) \setminus \{r\}$. Define a growing sequence $(\overrightarrow{T}(i))_{i\geq 0}$ of oriented trees recursively as follows:

- Set $\overrightarrow{T}(0)$ to be the tree with the single vertex r and with no edges.

- Suppose that $\overrightarrow{T}(i)$ has been generated. Start an independent network random walk at x_{i+1} and stop it at the first time it hits $\overrightarrow{T}(i)$ (note that the random walk hits $\overrightarrow{T}(i)$ a.s. by recurrence). Let $\langle y_0, \ldots, y_m \rangle$ be the loop erasure of this random walk.
- Set $\overrightarrow{T}(i+1)$ to be the oriented tree obtained by adding the oriented edges (y_0, y_1), (y_1, y_2), $\ldots$, (y_{m-1}, y_m) to $\overrightarrow{T}(i)$.
- The output of this algorithm is $\overrightarrow{T} := \bigcup_{i \geq 0} \overrightarrow{T}(i)$.

The oriented spanning forest sampled using Wilson's method has the law of the r-oriented wired spanning forest, due to the following theorems.

Theorem 4.9 ([Wil96, Theorem 1]). *Let G be a finite graph. Then, regardless of the ordering of $V(G) \setminus \{r\}$, the oriented tree $\overrightarrow{T}$ sampled using Wilson's method has the law of $\overrightarrow{\mathsf{WSF}}_r(G, c)$.* □

Theorem 4.10 ([BLPS01, Proposition 5.6]). *Let (G, c) be a recurrent network. Then for any finite subset $\overrightarrow{B}$ of $\overrightarrow{E}(G)$, any ordering of $V(G) \setminus \{r\}$, and any wired exhaustion of G,*

$$\mathbb{P}[\overrightarrow{B} \subseteq \overrightarrow{T}] = \lim_{n \to 0} \overrightarrow{\mu_{r,n}}[\overrightarrow{B} \subseteq \overrightarrow{T}_n],$$

where $\overrightarrow{T}$ is a random tree of G generated using Wilson's method, and $\overrightarrow{T}_n$ is a random tree of G_n distributed according to $\overrightarrow{\mu_{r,n}}$. □

We remark that [BLPS01] stated only the unoriented version of Theorem 4.10, but their argument in fact proves the oriented version as well. As a consequence of Theorem 4.10, we have that, for every recurrent network, the limit in (4) exists and does not depend on the choice of the wired exhaustion.

4.4. **Wilson's method oriented toward a root: transient case.** In this subsection we describe an algorithm that generates $\overrightarrow{\mathsf{WSF}}_r(G, c)$ for transient networks without using the weak limit in Definition 4.6.

For any walk $\langle x_i \mid 0 \leq i < I \rangle$ (including the case $I = \infty$), we denote by $\overrightarrow{E}(\langle x_i \mid 0 \leq i < I \rangle)$ the set of oriented edges $\{(x_i, x_{i+1}) \mid 0 \leq i < I - 1\}$, and we denote by $\overrightarrow{E}(\mathsf{R}(\langle x_i \mid 0 \leq i < I \rangle))$ the set of oriented edges $\{(x_{i+1}, x_i) \mid 0 \leq i < I - 1\}$.

Definition 4.11 (Wilson's method for transient networks). Let (G, c) be a transient network. Let $x_1, x_2, \ldots$ be an ordering of elements of $V(G) \setminus \{r\}$. Define a growing sequence $(\overrightarrow{F}(i))_{i \geq 0}$ of oriented forests recursively as follows:

- Start a network random walk at r that runs indefinitely. This random walk is locally finite a.s. by transience. Let $\langle y_0, y_1, \ldots \rangle$ be the loop erasure of this random walk. Set $\overrightarrow{F}(0)$ to be the tree oriented toward r given by
$$V(\overrightarrow{F}(0)) := \{y_i \mid i \geq 0\}; \qquad \overrightarrow{E}(\overrightarrow{F}(0)) := \overrightarrow{E}(\mathsf{R}(\langle y_i \mid i \geq 0 \rangle)).$$
- Suppose that $\overrightarrow{F}(i)$ has been generated. Start a network random walk at x_{i+1}. Stop the walk the first time it hits $\overrightarrow{F}(i)$; if it never hits $\overrightarrow{F}(i)$ then let it run indefinitely. This walk is locally finite a.s. by transience. Let $\langle y'_0, y'_1, \ldots \rangle$ be the loop erasure of this random walk.
- Set $\overrightarrow{F}(i+1)$ to be the oriented forest obtained by adding the edges in $\overrightarrow{E}(\langle y'_i \mid i \geq 0 \rangle)$ to $\overrightarrow{F}(i)$.
- The output of this algorithm is $\overrightarrow{F} := \bigcup_{i \geq 0} \overrightarrow{F}(i)$.

We remark that this algorithm is identical to Wilson's method oriented toward infinity [BLPS01] except for the first step, where we take the oriented edges from $\overrightarrow{E}(\mathsf{R}(\langle y_i \mid i \geq 0 \rangle))$ instead of $\overrightarrow{E}(\langle y_i \mid i \geq 0 \rangle)$. This difference causes the output to be a forest oriented toward r instead of toward infinity. We refer to [BLPS01] and [Hut18] for other methods to sample wired spanning forest oriented toward infinity.

The subgraph sampled using this method has the law of the r-oriented wired spanning forest, due to the following theorem.

Theorem 4.12 (cf.[BLPS01, Theorem 5.1]). *Let (G, c) be a transient network. Then for any finite subset $\overrightarrow{B}$ of $\overrightarrow{E}(G)$, any ordering of $V(G) \setminus \{r\}$, and any wired exhaustion of G,*

$$\mathbb{P}[\overrightarrow{B} \subseteq \overrightarrow{F}] = \lim_{n \to 0} \overrightarrow{\mu_{r,n}}[\overrightarrow{B} \subseteq \overrightarrow{T_n}],$$

where $\overrightarrow{F}$ is a random oriented forest of G generated using Wilson's method oriented toward r (Definition 4.11), and $\overrightarrow{T_n}$ is a random oriented tree of G_n distributed according to $\overrightarrow{\mu_{r,n}}$.

As a consequence of Theorem 4.12, we have that for all transient networks the limit in (4) exists and does not depend on the choice of the wired exhaustion. Our proof of Theorem 4.12 is paraphrased from its counterpart in [BLPS01].

Proof of Theorem 4.12. For any locally finite walk $\langle x_i \mid i \geq 0\rangle$, we have $\mathsf{LE}\langle x_i \mid i < I\rangle \to \mathsf{LE}\langle x_i \mid i \geq 0\rangle$ as $I \to \infty$. That is, if $\mathsf{LE}\langle x_i \mid i \leq I\rangle = \langle y_{I,i} \mid i \leq m_I\rangle$ and $\mathsf{LE}\langle x_i \mid i \geq 0\rangle = \langle y_i \mid i \geq 0\rangle$, then for every i and all sufficiently large I we have $y_{I,i} = y_i$. Since G is transient, it follows that $\mathsf{LE}\langle X_i \mid i < I\rangle \to \mathsf{LE}\langle X_i \mid i \geq 0\rangle$ as $I \to \infty$ a.s., where $\langle X_i \mid i \geq 0\rangle$ is a network random walk starting from any fixed vertex of G.

Let $x_1, x_2, \ldots$ be the ordering of $V(G) \setminus \{r\}$ used in Wilson's method for G. Write $x_0 := r$. Let L be a sufficiently large integer such that the endpoints of all edges in $\overrightarrow{B}$ are contained in $x_0, x_1, \ldots, x_L$. Let $\langle X_i^j \mid i \geq 0\rangle$ be independent random walks on G that start at x_j ($j \in \{0, \ldots, L\}$).

Let n be sufficiently large so that the wired exhaustion W_n contains $x_0, \ldots, x_L$. Run Wilson's method rooted at z_n in G_n with an ordering of $V(G_n)\setminus\{z_n\}$ that starts with $x_0, \ldots, x_L$, using the walks $\langle X_i^j \mid i \geq 0\rangle$ for $j \in \{0, \ldots, L\}$. Since these walks are on G rather than G_n, we simply stop the random walks once they leave the set W_n and say that they have hit z_n. In this way, we can couple the random walk in G_n that starts at x_j with the random walk in G that starts at x_j by using the same (infinite) random walk $\langle X_i^j \mid i \geq 0\rangle$ for $j \in \{0, \ldots, L\}$.

Let $\overrightarrow{T}'_n$ be the random spanning tree of G_n oriented toward z_n picked using Wilson's method for G_n as described in the previous paragraph. Note that $\overrightarrow{T}'_n$ has the law of $\overrightarrow{\mathsf{WSF}}_{z_n}(G_n, c_n)$ by Theorem 4.9.

Let h be the map from z_n-oriented spanning trees of G_n to r-oriented spanning trees of G_n that reverses the orientation of all edges in the unique directed path from r to z_n. Note that h is a bijection that preserves the weight of spanning trees. Write $\overrightarrow{T}_n := h(\overrightarrow{T}'_n)$. It then follows from definition of oriented wired spanning forest for finite graphs (Definition 4.5) that $\overrightarrow{T}_n$ has the law of $\overrightarrow{\mathsf{WSF}}_r(G_n, c_n)$.

Let τ_n^j be the first time that $\langle X_i^j \mid i \geq 0\rangle$ reaches the portion of the spanning tree created by the preceding random walks $\langle X_i^l \mid i \geq 0\rangle$ for $(l < j)$ using Wilson's method for G_n oriented toward z_n. Note that we have:

$$\overrightarrow{\mu_{r,n}}[\overrightarrow{B} \subseteq \overrightarrow{T}_n] = \mathbb{P}\left[\overrightarrow{B} \subseteq \overrightarrow{E}(\mathsf{R}(\mathsf{LE}\langle X_i^0 \mid i \leq \tau_n^0\rangle)) \cup \bigcup_{j=1}^{L} \overrightarrow{E}(\mathsf{LE}\langle X_i^j \mid i \leq \tau_n^j\rangle)\right]. \tag{5}$$

Let τ^j be the first time that $\langle X_i^j \mid i \geq 0\rangle$ reaches the portion of the spanning tree created by the preceding random walks $\langle X_i^l \mid i \geq 0\rangle$ for $(l < j)$ using Wilson's method for G oriented toward r. Note that we have

$$\mathbb{P}[\overrightarrow{B} \subseteq \overrightarrow{F}] = \mathbb{P}\left[\overrightarrow{B} \subseteq \overrightarrow{E}(\mathsf{R}(\mathsf{LE}\langle X_i^0 \mid i \leq \tau^0\rangle)) \cup \bigcup_{j=1}^{L} \overrightarrow{E}(\mathsf{LE}\langle X_i^j \mid i \leq \tau^j\rangle)\right], \tag{6}$$

where $\overrightarrow{F}$ is the oriented spanning forest generated using Wilson's method for G. Since the random walks used in Wilson's method for G and Wilson's method for G_n are the same, it follows from induction on j that $\tau_n^j \to \tau^j$ as $n \to \infty$. Together with (5) and (6), this implies the conclusion of the theorem. ☐

4.5. **Tail triviality.** An important property of the wired spanning forest (which will be used in proving Theorem 1.2) is that it is a tail trivial measure.

We first define tail triviality for measures on unoriented subgraphs. For any subset $K \subseteq E(G)$, let $\mathscr{F}(K) \subseteq \mathscr{F}$ denote the σ-algebra of events that depend only on K. An event $\mathscr{B} \in \mathscr{F}$ is a *tail event* if $\mathscr{B} \in \mathscr{F}(E \setminus K)$ for all finite $K \subseteq E$. A measure π on $\mathscr{F}$ is *tail trivial* if, for every tail event $\mathscr{B} \in \mathscr{F}$, we have $\pi[\mathscr{B}] \in \{0, 1\}$.

Theorem 4.13 ([LP16, Theorem 10.18])**.** *For every tail event $\mathscr{B} \in \mathscr{F}$, we have* $\mathsf{WSF}[\mathscr{B}] \in \{0, 1\}$. ☐

We now define tail triviality for measures on oriented subgraphs analogously. For any subset $\overrightarrow{K} \subseteq \overrightarrow{E}(G)$, let $\overrightarrow{\mathscr{F}}(\overrightarrow{K}) \subseteq \overrightarrow{\mathscr{F}}$ denote the σ-algebra of events that depend only on $\overrightarrow{K}$. An event $\overrightarrow{\mathscr{B}} \in \overrightarrow{\mathscr{F}}$ is a *tail event* if

$\overrightarrow{\mathscr{B}} \in \overrightarrow{\mathscr{F}}(\overrightarrow{E} \setminus \overrightarrow{K})$ for all finite $\overrightarrow{K} \subseteq \overrightarrow{E}$. A measure π on $\overrightarrow{\mathscr{F}}$ is *tail trivial* if, for every tail event $\overrightarrow{\mathscr{B}} \in \overrightarrow{\mathscr{F}}$, we have $\pi[\overrightarrow{\mathscr{B}}] \in \{0,1\}$. We now show that the following oriented subgraph measure is tail trivial.

Definition 4.14 (Oriented wired spanning forest plus one edge). The *r-oriented wired spanning forest plus one edge*, denoted $\overrightarrow{\mathsf{WSF}}_r^+ := \overrightarrow{\mathsf{WSF}}_r^+(G,c)$, is the law of the random subgraph $\overrightarrow{F} \sqcup \{(r,Y)\}$, where $\overrightarrow{F}$ is a random r-oriented forest of G sampled from $\overrightarrow{\mathsf{WSF}}_r$ and Y is a random neighbor of r sampled from μ_r independently of $\overrightarrow{F}$.

Lemma 4.15. *For every tail event $\overrightarrow{\mathscr{B}} \in \overrightarrow{\mathscr{F}}$, we have $\overrightarrow{\mathsf{WSF}}_r^+[\overrightarrow{\mathscr{B}}] \in \{0,1\}$.*

Proof. Let $f : 2^{\overrightarrow{E}} \to 2^E$ be the map that takes an oriented subgraph and erases the orientation of every edge. Let $g : 2^{\overrightarrow{E}} \to 2^{\overrightarrow{E}}$ be the map that takes an oriented subgraph and removes any outgoing edges of r.

Let $\overrightarrow{\mathscr{B}}$ be a tail event in $\overrightarrow{\mathscr{F}}$. Note that $\overrightarrow{\mathsf{WSF}}_r^+[\overrightarrow{\mathscr{B}}] = \overrightarrow{\mathsf{WSF}}_r[g(\overrightarrow{\mathscr{B}})]$ by the definition of $\overrightarrow{\mathsf{WSF}}_r^+$ and by the fact that $\overrightarrow{\mathscr{B}}$ does not depend on any outgoing edges of r. Also note that $\overrightarrow{\mathsf{WSF}}_r[g(\overrightarrow{\mathscr{B}})] = \mathsf{WSF}[f \circ g(\overrightarrow{\mathscr{B}})]$ by Lemma 4.7. Finally, note that the set $f \circ g(\overrightarrow{\mathscr{B}})$ is a tail event in $\mathscr{F}$ since $\overrightarrow{\mathscr{B}}$ is a tail event in $\overrightarrow{\mathscr{F}}$. The conclusion of the lemma now follows from the tail triviality of unoriented wired spanning forest (Theorem 4.13). $\square$

5. A native environment for random walk with local memory

In this section we show that the wired spanning forest measure can be used to construct a native environment. To rigorously define the notion of native environment, the underlying RWLM needs to satisfy the conditions described below.

A graph G is a *Cayley graph* if

- $V(G)$ is a group with identity element o;
- The group $V(G)$ is generated by a finite set $\mathcal{S} \subseteq V(G) \setminus \{o\}$;
- The set $\mathcal{S}$ is symmetric, i.e., if x is in $\mathcal{S}$ then x^{-1} is also in $\mathcal{S}$; and
- $E(G) = \{\{x,y\} \mid y^{-1}x \in \mathcal{S}\}$.

The square lattice $\mathbb{Z}^2$ is an example of a Cayley graph where the generating set $\mathcal{S}$ is $\{(\pm1,0),(0,\pm1)\}$ and the group operation is vector addition. Note that a Cayley graph is locally finite (because $\mathcal{S}$ is finite), connected (because $\mathcal{S}$ is a generating set), and simple (because $\mathcal{S}$ does not contain o).

A *weighted Cayley graph* (G,c) is a Cayley graph G with a weight function $c : \mathcal{S} \to \mathbb{R}_{>0}$ such that $c(x) = c(x^{-1})$ for all $x \in \mathcal{S}$. Note that the function $c : \mathcal{S} \to \mathbb{R}_{>0}$ extends naturally to a conductance $c : E \to \mathbb{R}_{>0}$ on edges of G by setting $c\{x,y\} := c(y^{-1}x) = c(x^{-1}y)$ for all $\{x,y\} \in E$.

Recall the definition of the probability transition function $p_x(\cdot,\cdot)$ from Definition 2.1. For every vertex x of G, we denote by μ_x the probability distribution on neighbors of x given by

$$\mu_x(y) \;:=\; \frac{c\{x,y\}}{\sum_{z \in N(x)} c\{x,z\}} \qquad (y \in N(x)). \tag{7}$$

Note that the measure μ_o is symmetric (i.e., $\mu_o(x) = \mu_o(x^{-1})$) as a consequence of $c : \mathcal{S} \to \mathbb{R}_{>0}$ being symmetric.

An RWLM is *transitive* if,

$$p_x(y,y') = p_{gx}(gy,gy') \qquad \text{for every} \quad x,g,y,y' \in V. \tag{Tran}$$

An RWLM is *c-stationary* if, for every vertex x,

$$\mu_x \text{ is a stationary distribution of the local chain } M_x. \tag{cSta}$$

Intuitively, the transitivity condition requires that the RWLM's mechanism at every vertex follow the same procedure. We remark that every RWLM in §2, with c being a constant function, is transitive and c-stationary.

For the rest of this paper, every RWLM will be transitive and c-stationary, and the underlying graph will always be a weighted Cayley graph, unless stated otherwise. Recall that X_n denotes the location of the walker and ρ_n denotes the rotor configuration at the n-th step of RWLM.

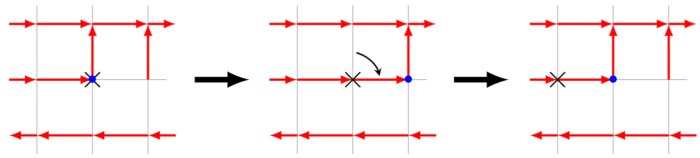

FIGURE 6. One step of the scenery process of a rotor walk on $\mathbb{Z}^2$ with clockwise rotation as its mechanism. The location of the origin in the original process is marked by the $\times$ symbol, and the location of the walker is marked by the $\bullet$ symbol.

Definition 5.1 (Scenery process). The *scenery process* is the sequence $(\widehat{\rho}_n)_{n\geq 0}$ of rotor configurations given by

$$\widehat{\rho}_n(x) := X_n^{-1}\rho_n(X_n x) \qquad (x \in V,\, n \geq 0).$$

Described in words, at each time step we apply a translation to the current rotor configuration so that the current location of the walker is mapped to the origin. In this way, $\widehat{\rho}_n$ is the rotor configuration *as viewed from the perspective of the walker* at the n-th step of the RWLM. See Figure 6 for an illustration of a scenery process.

Note that, as a consequence of (Tran), the scenery process $(\widehat{\rho}_n)_{n\geq 0}$ is a Markov chain with state space the set of rotor configurations of G and with transition rule

$$\widehat{\rho}_{n+1}(x) := \begin{cases} o & \text{if } x = Y_n^{-1}; \\ Y_n^{-1}\widehat{\rho}_n(Y_n x) & \text{if } x \neq Y_n^{-1}, \end{cases} \tag{8}$$

where Y_n is a random neighbor of o sampled from $p_o(\widehat{\rho}_n(o),\cdot)$ independently of $\sigma(\widehat{\rho}_0,\ldots,\widehat{\rho}_{n-1})$ (recall that p_o is the probability transition function of the local chain M_o).

Definition 5.2 (Native environment). A *native environment* is a probability distribution on rotor configurations of G such that, if the walker starts at o and the initial rotor configuration is sampled from the distribution, then the scenery process is a stationary sequence, i.e.,

$$(\widehat{\rho}_n)_{n\geq 0} \overset{d}{=} (\widehat{\rho}_{n+1})_{n\geq 0}.$$

Intuitively, a native environment means that, at each time step of the walk, the rotor configuration *viewed from the perspective of the walker* has the same law as the initial environment. See Figure 6 for an illustration of a native environment.

We now restate Theorem 1.1 from the introduction (also the main result of this section) in a slightly more general form. Recall the definition of $\overrightarrow{\mathsf{WSF}}_o^+ := \overrightarrow{\mathsf{WSF}}_o^+(G,c)$ from Definition 4.14.

Theorem 1.1. Consider an RWLM on a weighted Cayley graph that satisfies (Tran) and (cSta). Then $\overrightarrow{\mathsf{WSF}}_o^+$ is a native environment.

Note that $\overrightarrow{\mathsf{WSF}}_o^+$ is indeed a probability distribution on rotor configurations of G. This is because, by Wilson's method (see §4.3 and §4.4), the random subgraph sampled from $\overrightarrow{\mathsf{WSF}}_o$ has exactly one outgoing edge for every $x \in V \setminus \{o\}$ and no outgoing edge for o. Hence the random subgraph sampled from $\overrightarrow{\mathsf{WSF}}_o^+$ has exactly one outgoing edge for every vertex, and by Remark 2.2 it defines a rotor configuration of G.

We remark that, when G is a finite Cayley graph, Theorem 1.1 then specializes to the result of [Bro89, Ald90] (for Aldous-Broder walk) and [HLM+08, Lemma 3.4] (for rotor walk).

We now build toward the proof of Theorem 1.1. We will use the following identity, which is a special case of [Lev11, Lemma 2.4] if the graph G is finite.

Lemma 5.3. *Let (G,c) be an electrical network, and let r be a vertex. Let Y be a random neighbor of r sampled from μ_r, and let $\overrightarrow{F_Y}$ be a random oriented spanning forest of G sampled from $\overrightarrow{\mathsf{WSF}}_Y$. Then the random oriented subgraph $\overrightarrow{F_Y} \sqcup \{(Y,r)\}$ has the distribution $\overrightarrow{\mathsf{WSF}}_r^+$.*

Proof. It suffices to consider the case when G is a finite graph, as the case of infinite graphs follows by taking the limit over a wired exhaustion and then verifying the lemma for all events that depend on only finitely many edges.

When G is a finite graph, note that $\overrightarrow{F_Y} \sqcup \{(Y, r)\}$ is concentrated on oriented spanning unicycles rooted at r, i.e., oriented subgraphs of G with one outgoing edge for every vertex of G and one unique oriented cycle, where r is contained in that oriented cycle. Each unicycle $\overrightarrow{U}$ is picked with probability proportional to the product of the weight of its edges. This implies that $\overrightarrow{F_Y} \sqcup \{(Y, r)\}$ is distributed as $\overrightarrow{\mathsf{WSF}}_r^+$, as desired. $\square$

Proof of Theorem 1.1. Since $(\widehat{\rho}_n)_{n\geq 0}$ is a Markov chain, it suffices to show that if $\widehat{\rho}_0$ is distributed as $\overrightarrow{\mathsf{WSF}}_o^+$, then $\widehat{\rho}_1$ is also distributed as $\overrightarrow{\mathsf{WSF}}_o^+$.

Let $\overrightarrow{F}$ be the random spanning forest of G sampled from $\overrightarrow{\mathsf{WSF}}_o$. Let Y be a random neighbor of the identity sampled from μ_o independently of $\overrightarrow{F}$. For any $x \in V$, denote by $\tau_x : V(G) \to V(G)$ the network isomorphism of (G, c) given by left multiplication by x. (A network isomorphism of (G, c) is a graph isomorphism of G which also preserves the conductance c.)

Since $\rho_0(o) \overset{d}{=} Y$ and the RWLM satisfies (cSta), we have $\rho_1(o) \overset{d}{=} Y$. By the transition rule of RWLM (see (2)), we then have $\rho_1 \overset{d}{=} \overrightarrow{F} \sqcup \{(o, Y)\}$. By the transition rule of the scenery process (see (8)), we then have $\widehat{\rho}_1 \overset{d}{=} \tau_{Y^{-1}}(\overrightarrow{F}) \sqcup \{(Y^{-1}, o)\}$.

Now note that $Y \overset{d}{=} Y^{-1}$ since μ_o is symmetric, and together with the conclusion of the previous paragraph this implies that $\widehat{\rho}_1 \overset{d}{=} \tau_Y(\overrightarrow{F}) \sqcup \{(Y, o)\}$. Also note that $\tau_Y(\overrightarrow{F})$ is equal in distribution to the random spanning forest picked from $\overrightarrow{\mathsf{WSF}}_Y$ since τ_Y is a network isomorphism of (G, c). It now follows from Lemma 5.3 that $\widehat{\rho}_1$ is distributed according to $\overrightarrow{\mathsf{WSF}}_o^+$, and the proof is complete. $\square$

6. Ergodic native environments

In this section we prove Theorem 1.2 by showing that $\overrightarrow{\mathsf{WSF}}_o^+$ is an ergodic native environment. This requires tools from the ergodic theory of Markov chains, which we quickly review in the next subsection, and we refer the reader to [HLL98] for a more detailed discussion on this subject.

6.1. Ergodic theory for Markov chains.

Let $M := (\Omega, \mathscr{F}, P)$ be a Markov chain, where the state space Ω is a metric space, $\mathscr{F}$ is the Borel σ-algebra of Ω, and $P : \Omega \times \mathscr{F} \to [0, 1]$ is the probability transition function of this chain. A set $\mathscr{B} \in \mathscr{F}$ is *invariant* if $P(x, \mathscr{B}) = 1$ for all $x \in \mathscr{B}$. A stationary distribution π of M is *ergodic* if $\pi[\mathscr{B}] \in \{0, 1\}$ for any invariant set $\mathscr{B}$.

Let $\Omega^{\mathbb{N}}$ be the *trajectory space* of M,

$$\Omega^{\mathbb{N}} := \{(\omega_i)_{i\geq 0} \mid \omega_i \in \Omega\},$$

equipped with the product σ-algebra induced by $\mathscr{F}$. For any $\omega \in \Omega$ we denote by P_ω the probability distribution on $\Omega^{\mathbb{N}}$ given by:

$$P_\omega[\mathscr{A}] := \mathbb{E}[1_{\mathscr{A}}(\omega_0, \omega_1, \ldots)] \qquad (\mathscr{A} \in \mathscr{F}),$$

where $(\omega_n)_{n\geq 0}$ is the Markov chain M with initial state $\omega_0 = \omega$, and $\mathbb{E}$ is the corresponding expectation function for this chain.

Theorem 6.1 (Pointwise ergodic theorem [HLL98, Theorem 6.1(b)]**).** *Let M be a Markov chain on a compact metric space $(\Omega, \mathscr{F})$, and let π be an ergodic distribution of M. Then for every π-integrable function $f : \Omega \to \mathbb{R}$,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(\omega_i) = \int_\Omega f \, d\pi \qquad P_\omega\text{-a.s.},$$

for π-almost every $\omega \in \Omega$. $\square$

The following lemma will be useful for checking if a given stationary distribution π is ergodic. For any $n \geq 1$, we denote by $P^{(n)}$ the n-step transition function of the Markov chain M.

Lemma 6.2. *Let $M := (\Omega, \mathscr{F}, P)$ be a Markov chain, and let π be a stationary distribution of M. If $\mathscr{B}$ is an invariant set, then the set*

$$\mathscr{B}' := \{x \in \Omega \mid \exists\ n \geq 1 \text{ s.t. } P^{(n)}(x, \mathscr{B}) > 0\},$$

differs from $\mathscr{B}$ by a set of π-measure zero.

Proof. First note that $\mathscr{B} \subseteq \mathscr{B}'$ by the invariance of $\mathscr{B}$. Now note that, for any $n \geq 1$,

$$\begin{aligned} \pi[\mathscr{B}] &= \int_\Omega P^{(n)}(x, \mathscr{B})\ d\pi(x) \qquad \text{(by the stationarity of } \pi) \\ &= \int_{\mathscr{B}} P^{(n)}(x, \mathscr{B})\ d\pi(x) + \int_{\mathscr{B}' \setminus \mathscr{B}} P^{(n)}(x, \mathscr{B})\ d\pi(x) \qquad (\text{as } P^{(n)}(x, \mathscr{B}) = 0 \text{ for } x \notin \mathscr{B}') \\ &= \pi[\mathscr{B}] + \int_{\mathscr{B}' \setminus \mathscr{B}} P^{(n)}(x, \mathscr{B})\ d\pi(x) \qquad (\text{as } \mathscr{B} \text{ is invariant}). \end{aligned}$$

Hence we conclude that $\int_{\mathscr{B}' \setminus \mathscr{B}} P^{(n)}(x, \mathscr{B})\ d\pi(x) = 0$ for any $n \geq 1$. It then follows from the definition of $\mathscr{B}'$ that that $\pi[\mathscr{B}' \setminus \mathscr{B}] = 0$. This proves the lemma. □

6.2. **Proof of Theorem 1.2.** Recall the definition of scenery process $(\widehat{\rho}_n)_{n \geq 0}$ from Definition 5.1.

Definition 6.3 (Ergodic native environment). Consider an RWLM on a weighted Cayley graph that satisfies (Tran) and (cSta). An *ergodic environment* is a distribution on rotor configurations of G that is an ergodic measure for the scenery process of the RWLM.

We now restate Theorem 1.2 from the introduction (also the main result of this section) in a slightly more general form. Recall that the definition of probability transition functions p_x from Definition 2.1. We say that the RWLM is *elliptic* if,

$$p_x(y, y') \ > \ 0 \qquad \text{for every } x \in V \text{ and every } y, y' \in N(x). \tag{Ell}$$

Note that, from the RWLMs in §2, the Aldous-Broder walk and the p,r-rotor walk with $r < 1$ are elliptic, while p-rotor walk and deterministic rotor walk are not elliptic.

Theorem 1.2. Consider an RWLM on a weighted Cayley graph that satisfies (Tran), (cSta), and (Ell). Then $\overrightarrow{\mathsf{WSF}}_o^+$ is an ergodic native environment.

Proof. Let $\pi = \overrightarrow{\mathsf{WSF}}_o^+$, and let $\overrightarrow{\mathscr{B}}$ be a set of rotor configurations that is invariant w.r.t. the scenery process. Recall the definition of tail event for rotor configurations (equivalently, oriented subgraphs) from §4.5. It suffices to show that $\overrightarrow{\mathscr{B}}$ differs from a tail event by a set of π-measure zero, as it will then follow from the tail triviality of π (Lemma 4.15) that $\pi[\overrightarrow{\mathscr{B}}] \in \{0, 1\}$.

Let $\mathrm{Rot}(G)$ denote the set of rotor configurations of G. We write

$$\overrightarrow{\mathscr{C}} := \{\rho \in \mathrm{Rot}(G) \mid \exists\ \rho' \in \overrightarrow{\mathscr{B}} \text{ s.t. } \rho \text{ and } \rho' \text{ differ at finitely many vertices}\}.$$

Note that $\overrightarrow{\mathscr{C}}$ is a tail event that contains $\overrightarrow{\mathscr{B}}$. It then suffices to show that $\pi[\overrightarrow{\mathscr{C}} \setminus \overrightarrow{\mathscr{B}}] = 0$.

Let ρ be any rotor configuration in $\overrightarrow{\mathscr{C}}$. Then there exists $\rho' \in \overrightarrow{\mathscr{B}}$ such that ρ' differs from ρ at finitely many vertices. Let $\langle x_0, \ldots, x_n \rangle$ be a directed walk in G that starts at o and such that $\{x_0, \ldots, x_{n-1}\}$ contains all the vertices for which ρ and ρ' differ.

For each $i \in \{1, \ldots, n\}$, define ρ_i to be the rotor configuration at the i-th step of the RWLM if the initial walker-and-rotor configuration is (x_0, ρ) and the trajectory of the walker for the first i steps is given by $\langle x_0, \ldots, x_i \rangle$. That is, these rotor configurations are given by the recursive definition

$$\rho_{i+1}(x) := \begin{cases} x_{i+1} & \text{if } x = x_i; \\ \rho_i(x) & \text{otherwise.} \end{cases}$$

Define ρ'_i in a similar manner, but with (x_0, ρ') as the initial walker-and-rotor configuration. Note that $\rho_n = \rho'_n$ since the walker of the RWLM that follows the directed walk $\langle x_0, \ldots, x_n \rangle$ would have visited and changed the rotors at all vertices for which ρ and ρ' differ; see Figure 7.

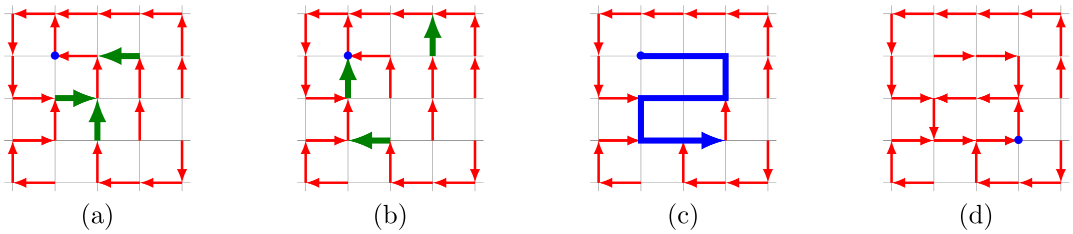

FIGURE 7. (a) and (b) Two rotor configurations that differ at finitely many vertices. The rotors at which they differ are drawn oversized in green. (c) The trajectory (drawn in blue) taken by the walker that visits every green rotor. (d) The final rotor configuration of the RWLM at the end of this process, which is the same regardless of whether the initial configuration is (a) or (b).

Write $\rho'' := \tau_{x_n^{-1}}(\rho_n) = \tau_{x_n^{-1}}(\rho'_n)$. Note that ρ'' is the rotor configuration at the n-th step of the scenery process if the walker of the RWLM follows $\langle x_0, \dots, x_n\rangle$ and the initial rotor configuration is ρ or ρ' (recall that τ_x is the network isomorphism of (G, c) given by left multiplication by x). In particular, the probability to transition from ρ to ρ'' in n steps of the scenery process satisfies the following inequality:

$$P^{(n)}(\rho, \rho'') \geq \prod_{i=0}^{n-1} p_{x_i}(\rho_i(x_i), x_{i+1}) > 0,$$

where the strict inequality is due to (Ell). Note that, by the same argument, we also have $P^{(n)}(\rho', \rho'') > 0$.

Since $\rho' \in \overrightarrow{\mathscr{B}}$ and $P^{(n)}(\rho', \rho'') > 0$, we have $\rho'' \in \overrightarrow{\mathscr{B}}$ by the invariance of $\overrightarrow{\mathscr{B}}$. This implies that ρ can transition into $\overrightarrow{\mathscr{B}}$ in n steps of the scenery process with positive probability, as

$$P^{(n)}(\rho, \overrightarrow{\mathscr{B}}) \geq P^{(n)}(\rho, \rho'') > 0.$$

As the choice of $\rho \in \overrightarrow{\mathscr{C}}$ is arbitrary, we have from the argument above that:

$$\overrightarrow{\mathscr{C}} \subseteq \{\rho \in \mathrm{Rot}(G) \mid \exists\, n \geq 1 \text{ s.t. } P^{(n)}(\rho, \overrightarrow{\mathscr{B}}) > 0\}.$$

Since π is a stationary distribution of the scenery process (Theorem 1.1), it then follows from Lemma 6.2 that the set on the right side of the equation differs from $\overrightarrow{\mathscr{B}}$ by a set of π-measure zero. Hence we conclude that $\pi[\overrightarrow{\mathscr{C}} \setminus \overrightarrow{\mathscr{B}}] = 0$, as desired. $\square$

7. FUNCTIONAL CLT FOR RWLM

In this section we present the proof of Theorem 1.3. An electrical network (G, c) is a *weighted lattice graph* in $\mathbb{R}^d$ if G is weighted Cayley graph such that $V(G)$ is a subgroup of $\mathbb{R}^d$ with vector addition as the group operation. In this section we will assume that G is a weighted lattice graph, and that the walker is initially located at the origin $\mathbf{0} := (0, \dots, 0)$, unless stated otherwise.

We now restate Theorem 1.3 from the introduction in a slightly more general form. Recall the definition the measure μ_x from (7). We denote by Γ the matrix

$$\Gamma := \sum_{\mathbf{y} \in N(\mathbf{0})} \mu_{\mathbf{0}}(\mathbf{y})\, \mathbf{y}\, \mathbf{y}^\top.$$

Recall that X_n and ρ_n $(n \geq 0)$ denotes the location of the walker and the rotor configuration at the n-th step of the RWLM, respectively, and that $D_{\mathbb{R}^d}[0, \infty)$ denotes the Skorohod space of $\mathbb{R}^d$-valued càdlàg paths on $[0, \infty)$.

Theorem 1.3. Consider an RWLM on a weighted lattice graph in $\mathbb{R}^d$ that is (Tran), (cSta), and (Mtgl). Suppose that the initial environment π is an ergodic native environment. Then, for almost every environment sampled from π, the scaled walk $(\frac{1}{\sqrt{n}} X_{\lfloor nt \rfloor})_{t\geq 0}$ converges weakly on $D_{\mathbb{R}^d}[0, \infty)$ to a Brownian motion with diffusion matrix Γ.

As a consequence of Theorem 1.3, the p,r-rotor walk on $\mathbb{Z}^d$ (Example 2.5) with constant conductance, with $p = \frac{1}{2}$ and $r < 1$, and with $\overrightarrow{\mathsf{WSF}}_o^+$ as the initial environment, converges weakly on $D_{\mathbb{R}^d}[0,\infty)$ to a Brownian motion with diffusion matrix $\frac{1}{d} I_d$.

Proof of Theorem 1.3. Let $V_n := X_{n+1} - X_n$ and $\mathscr{F}_n := \sigma(X_0, \ldots, X_n, \rho_0, \ldots, \rho_n)$. It suffices to verify that all conditions in Theorem 3.1 are satisfied. By using the same argument as in the proof of Proposition 1.4, we have that $(X_n)_{n\geq 0}$ is a square-integrable martingale process (as a consequence of (Mtgl)), and that (CLT2) is satisfied. We omit the details for brevity.

We now verify (CLT1). Let $i \geq 0$. It follows from Definition 2.3 and (Tran) that

$$V_i = X_{i+1} - X_i = \sum_{\mathbf{y} \in N(\mathbf{0})} \mathbb{1}\{\rho_i(X_i) - X_i = \mathbf{y}\}\, Y_{\mathbf{y},i},$$

where $Y_{\mathbf{y},i}$ is a random variable on neighbors of the origin sampled from $p_{\mathbf{0}}(\mathbf{y}, \cdot)$ independently of $\mathscr{F}_i$. Then, for any $n \geq 0$:

$$\frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}\big[V_i V_i^\top \mid \mathscr{F}_i\big] = \sum_{\mathbf{y} \in N(\mathbf{0})} \left(\frac{1}{n} \sum_{i=0}^{n-1} \mathbb{1}\{\rho_i(X_i) - X_i = \mathbf{y}\} \right) \mathbb{E}\big[Y_{\mathbf{y},0}\, Y_{\mathbf{y},0}^\top\big]. \tag{9}$$

Here we have used the fact that $Y_{\mathbf{y},i}$ has the same law as $Y_{\mathbf{y},0}$ for all i.

We now show that, for every $\mathbf{y} \in N(\mathbf{0})$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{1}\{\rho_i(X_i) - X_i = \mathbf{y}\} \;=\; \mu_{\mathbf{0}}(\mathbf{y}). \tag{10}$$

Fix an ordering $x_1, x_2, \ldots$ of $V(G)$. Note that the set of rotor configurations $\mathrm{Rot}(G)$ is a compact metric space with the metric $d(\rho_1, \rho_2) := \sum_{i=1}^\infty \frac{1}{2^i} \mathbb{1}\{\rho_1(x_i) \neq \rho_2(x_i)\}$. It is straightforward to check that $\overrightarrow{\mathscr{F}}(G)$ (from §4.2) restricted to $\mathrm{Rot}(G)$ is the Borel σ-algebra corresponding to this metric. Hence all conditions of Theorem 6.1 are satisfied, and (10) now follows by applying Theorem 6.1 to the function $f : \mathrm{Rot}(G) \to \mathbb{R}$ given by $f(\widehat{\rho}) := \mathbb{1}\{\widehat{\rho}(\mathbf{0}) = \mathbf{y}\}$.

Plugging (10) into (9), we get

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}\big[V_i V_i^\top \mid \mathscr{F}_i\big] &= \sum_{\mathbf{y} \in N(\mathbf{0})} \mu_{\mathbf{0}}(\mathbf{y})\, \mathbb{E}\big[Y_{\mathbf{y},0}\, Y_{\mathbf{y},0}^\top\big] \\ &= \sum_{\mathbf{y} \in N(\mathbf{0})} \mu_{\mathbf{0}}(\mathbf{y}) \sum_{\mathbf{y}' \in N(\mathbf{0})} p_{\mathbf{0}}(\mathbf{y}, \mathbf{y}')\, \mathbf{y}'\, \mathbf{y}'^\top.\end{aligned}$$

Since $\mu_{\mathbf{0}}$ is a stationary distribution of the mechanism at $\mathbf{0}$ by (cSta), it then follows that:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}\big[V_i V_i^\top \mid \mathscr{F}_i\big] = \sum_{\mathbf{y}' \in N(\mathbf{0})} \mu_{\mathbf{0}}(\mathbf{y}')\, \mathbf{y}'\, \mathbf{y}'^\top = \Gamma.$$

Hence (CLT1) is verified, and the proof is complete. ☐

8. Concluding remarks

We conclude with a few natural questions.

8.1. Theorem 1.3 allows us to derive a functional CLT, but only when the initial environment is an ergodic native environment. Does the conclusion of Theorem 1.3 still hold for other initial environments? We believe that the answer to this question is positive for the iid initial environment, and simulations suggests that there should be no quantitative difference between iid initial environment and wired spanning forest plus one edge environment eventually.

Problem 8.1. Consider an RWLM on a simple Cayley graph that is transitive, uniform, and elliptic. Let $(\widehat{\rho}_n)_{n\geq 0}$ be the scenery process of the RWLM with iid initial environment. Show that $\widehat{\rho}_n$ converges weakly to $\overrightarrow{\mathsf{WUSF}}^+$, i.e., for every edge $\{x_1, y_1\}, \ldots, \{x_m, y_m\}$ of G,

$$\mathbb{P}\big[\widehat{\rho}_n(x_1) = y_1, \ldots, \widehat{\rho}_n(x_m) = y_m\big] \overset{n\to\infty}{\longrightarrow} \mathbb{P}\big[(x_1, y_1), \ldots, (x_m, y_m) \in \overrightarrow{U}\big],$$

where $\overrightarrow{U}$ is a random subgraph sampled from $\overrightarrow{\mathsf{WUSF}}^+$.

8.2. Can we drop the ellipticity assumption from Theorem 1.2 and Theorem 1.3? In particular, a positive answer to this question will give us a scaling limit result for p-rotor walk on $\mathbb{Z}^d$ ($d \geq 2$) when $p = \frac{1}{2}$, which will be consistent with the simulation results in Figure 1.

8.3. An RWLM is *recurrent* if every vertex is visited infinitely often by the walker a.s. and is *transient* otherwise. Note that every d-dimensional RWLM on $\mathbb{Z}^d$ satisfying conditions in Theorem 1.3 is transient if $d \geq 3$ (as the transience of the scaling limit implies the transience of the original walk). Is it true that these RWLMs are recurrent if $d = 2$? We remark that a partial answer to this question has been given in [Cha20] (the sequel to this paper), namely for the 'H'–'V' walk on $\mathbb{Z}^2$ with i.i.d. initial environment and $p = \frac{1}{2}$, and remains open for other values of p.

Acknowledgement

We would like to thank Elena Kosygina, Yuval Peres and Ofer Zeitouni for inspiring discussions with the first author and the third author. We would also like to thank Timo Seppäläinen for pointing us to references for Theorem 3.1. The first author would also like to thank Igor Pak for writing advice. Last but not the least, we would like to thank the anonymous referees for the insightful comments that substantially improve the readability of the paper and for additional references.

References

[Ald90] D. J. Aldous, The random walk construction of uniform spanning trees and uniform labelled trees, *SIAM J. Discrete Math.* **3** (1990), 450—465.

[Bil06] J. A. Bilmes, What hmms can do, *IEICE Trans. Inf. Syst.* **E89-D** (2006), 869–891.

[BKPS04] I. Benjamini, H. Kesten, Y. Peres, and O. Schramm, Geometry of the uniform spanning forest: transitions in dimensions 4,8,12,. . . , *Ann. of Math. (2)* **160** (2004), 465–491.

[BL16] B. Bond and L. Levine, Abelian networks I. Foundations and examples, *SIAM J. Discrete Math.* **30** (2016), 856–874.

[BL19] B. Bercu and L. Laulin, On the multi-dimensional elephant random walk, *J. Stat. Phys.* **175** (2019), 1146–1163.

[BLPS01] I. Benjamini, R. Lyons, Y. Peres, and O. Schramm, Uniform spanning forests, *Ann. Probab.* **29** (2001), 1–65.

[BR07] J. Bérard and A. F. Ramírez, Central limit theorem for the excited random walk in dimension $d \geq 2$, *Electron. Comm. Probab.* **12** (2007), 303–314.

[Bro89] A. Broder, Generating random spanning trees, in *Proc. 58th FOCS, IEEE* (2019), 442–447.

[BW03] I. Benjamini and D. B. Wilson, Excited random walk, *Electron. Comm. Probab.* **8** (2003), 86–92.

[BS02] E. Bolthausen and A.-S. Sznitman, Ten lectures on random media, in *DMV Seminar* **32**, Birkhäuser, Basel (2002), vi+116 pp.

[Cha20] S. H. Chan, Recurrence of horizontal-vertical walks, preprint (2020), 34 pp.; arXiv:2012.10811.

[DK12] D. Dolgopyat and E. Kosygina, Scaling limits of recurrent excited random walks on integers, *Electron. Commun. Probab.* **17** (2012), 14 pp.

[FLP16] L. Florescu, L. Levine, and Y. Peres, The range of a rotor walk, *Amer. Math. Monthly* **123** (2016), 627–642.

[HLL98] O. Hernández-Lerma and J. B. Lasserre, Ergodic theorems and ergodic decomposition for Markov chains, *Acta Appl. Math.* **54** (1998), 99–119.

[HLM+08] A. E. Holroyd, L. Levine, K. Meszáros, Y. Peres, J. Propp, and D. Wilson, Chip-firing and rotor-routing on directed graphs, in *In and out of equilibrium. 2, Progr. Probab.* **60**, Birkhäuser, Basel (2008), 331–364.

[HLSH18] W. Huss, L. Levine, and E. Sava-Huss, Interpolating between random walk and rotor walk, *Random Structures Algorithms* **52** (2018), 263–282.

[HP19] T. Hutchcroft and Y. Peres, The component graph of the uniform spanning forest: transitions in dimensions 9,10,11,. . . , *Probab. Theory Related Fields* **175** (2019), 141–208.

[Hut18] T. Hutchcroft, Interlacements and the wired uniform spanning forest, *Ann. Probab.* **46** (2018), 1170–1200.

[KM11] E. Kosygina and T. Mountford, Limit laws of transient excited random walks on integers, *Ann. Inst. Henri Poincaré Probab. Stat.* **47** (2011), 575–600.

[KOS16] G. Kozma, T. Orenshtein, and I. Shinkar, Excited random walk with periodic cookies, *Ann. Inst. Henri Poincaré Probab. Stat.* **52** (2016), 1023–1049.

[KOZ85] S.M. Kozlov, The method of averaging and walks in inhomogeneous environments, *Russian Math. Surveys* **40** (1985), 73–145.

[KP16] E. Kosygina and J. Peterson, Functional limit laws for recurrent excited random walks with periodic cookie stacks, *Elecrton. J. Probab* **21** (2016), 24 pp.

[KP17] E. Kosygina and J. Peterson, Excited random walks with Markovian cookie stacks, *Ann. Inst. Henri Poincaré Probab. Stat.* **53** (2017), 1458–1497.

[KZ13] E. Kosygina and M. Zerner, Excited random walks: Results, methods, open problems, *Bull. Inst. Math. Acad. Sin. (N.S.)* **8** (2013), 105–157.

[Lev11] L. Levine, Sandpile groups and spanning trees of directed line graphs, *J. Combin. Theory Ser. A* **118** (2011), 350–364.

[LP16] R. Lyons and Y. Peres, *Probability on trees and networks*, *Camb. Ser. Stat. Probab. Math.* **42**, Cambridge University Press, New York (2016).

[MPRV12] M. Menshikov, S. Popov, A. F. Ramírez, and M. Vachkovskaia, On a general many-dimensional excited random walk, *Ann. Probab.* **40** (2012), 2106–2130.

[PDDK96] V. Priezzhev, D. Dhar, A. Dhar, and S. Krishnamurthy, Eulerian walkers as a model of self-organized criticality, *Phys. Rev. Lett.* **77** (1996), 5079–5082.

[Pem91] R. Pemantle, Choosing a spanning tree for the integer lattice uniformly, *Ann. Probab.* **19** (1991), 1559–1574.

[Pro03] J. Propp, Random walk and random aggregation, derandomized, online lecture (2003), https://www.microsoft.com/en-us/research/video/random-walk-and-randomaggregation-derandomized/.

[PT17] R. Pinsky and N. Travers, Transience, recurrence and the speed of a random walk in a site-based feedback environment, *Probab. Theory Related Fields* **167** (2017), 917–978.

[PV82] G. C. Papanicolaou and S. R. S. Varadhan, Diffusions with random coefficients, in *Statistics and probability: essays in honor of C. R. Rao*, Amsterdam (1982), 547–552.

[RAS05] F. Rassoul-Agha and T. Seppäläinan, An almost sure invariance principle for random walks in a space-time random environment, *Probab. Theory Related Fields* **133** (2005), 299–314.

[Szn04] A.-S. Sznitman, Topics in random walks in random environment, in *School and Conference on Probability Theory*, *Abdus Salam Int. Cent. Theoret. Phys.*, Trieste, ICTP Lect. Notes, XVII (2004), 203–266.

[Tót95] B. Tóth, The "true" self-avoiding walk with bond repulsion on $\mathbb{Z}$: limit theorems, *Ann. Probab.* **23** (1995), 1523–1556.

[Tót96] B. Tóth, Generalized Ray-Knight theory and limit theorems for self-interacting random walks on $\mathbb{Z}$, *Ann. Probab.* **24** (1996), 1324–1367.

[Tót99] B. Tóth, Self-interacting random motions–a survey, in *Random Walks – A Collection of Surveys, Bolyai Soc. Math. Stud.* **9**, János Bolyai Math. Soc., Budapest (1999), 349–384.

[Tra18] N. F. Travers, Excited random walk in a Markovian environment, *Electron. J. Probab.* **23** (2018), 60 pp.

[VdHH12] R. van der Hofstad and M. Holmes, An expansion for self-interacting random walks, *Braz. J. Probab. Stat.* **26** (2012), 1–55.

[Wil96] D. B. Wilson, Generating random spanning trees more quickly than the cover time, in *Proc. 28th STOC*, ACM, New York (1996), 296–303.

[WLB96] I. Wagner, M. Lindenbaum, and A. Bruckstein, Smell as a computational resource–a lesson we can learn from the ant, in *Israel Symposium on Theory of Computing and Systems*, IEEE Comput. Soc. Press, Los Alamitos (1996), 219–230.

[Zei04] O. Zeitouni, Random walks in random environment, in *Lectures on probability theory and statistics*, *Lecture Notes in Math.* **1837**, Springer, Berlin (2004), 189–312.

Appendix A. Random walks with hidden local memory

In this section we present a more general version of random walk with local memory inspired by hidden Markov chains. We refer to [Bil06] for a more detailed discussion on hidden Markov chains.

For each $x \in V$, a *hidden mechanism* at x is a Markov chain M_x with finite state space S_x and probability transition function $p_x(\cdot,\cdot)$. A *jump rule* is a map $f_x : S_x \to \mathcal{P}(N(x))$ from S_x to the set of probability distributions on the set of neighbors of x. A *hidden state configuration* is a map $\kappa : V \to \sqcup_{x\in V} S_x$ such that $\kappa(x) \in S_x$ for all $x \in V$.

Definition A.1 (Random walk with hidden local memory). A *random walk with hidden local memory*, or RWHLM for short, is a sequence $(X_n, \rho_n, \kappa_n)_{n\geq 0}$ satisfying the following transition rules:

(i) $\kappa_{n+1}(x) := \begin{cases} K_n & \text{if } x = X_n; \\ \kappa_n(x) & \text{if } x \neq X_n. \end{cases}$;

(ii) $\rho_{n+1}(x) := \begin{cases} Y_n & \text{if } x = X_n; \\ \rho_n(x) & \text{if } x \neq X_n, \end{cases}$

(iii) $X_{n+1} := Y_n$,

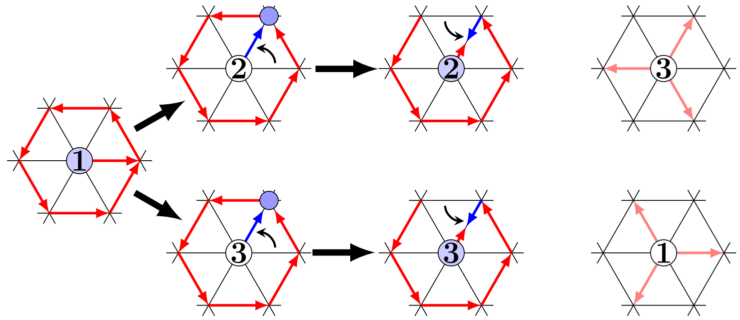

FIGURE 8. Two instances of a two-step hidden triangular walk with the same walker's trajectory and rotor configurations. The number at the origin records the hidden state of the origin. The pictures at the right side illustrate the future hidden state of the origin and the arrows point to (possible) future locations of the walker.

where K_n is a random element of S_{X_n} sampled from $p_{X_n}(\kappa_n(X_n), \cdot)$ independent of the past, and Y_n is a random neighbor of x sampled from $f_{X_n}(K_n)$ independent of the past.

Described in words, at each time step (i) the walker first updates the hidden state of its current location using the given hidden mechanism. Then, (ii) the walker updates the rotor of its current location by sampling the new rotor from the probability distribution corresponding to the new hidden state. Finally, (iii) the walker travels to the vertex specified by the new rotor.

Example A.2 (Hidden triangular walk). Let G be the triangular lattice. For each $\mathbf{x} \in V$, the hidden mechanism at $\mathbf{x} \in V$ has the following state space and transition probability:

$$S_{\mathbf{x}} := \{s_1, s_2, s_3\}; \qquad p_{\mathbf{x}} := \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

That is, s_1 transitions to either s_2 or s_3 with equal probability, s_2 transitions to s_3 with probability 1, and s_3 transitions to s_1 with probability 1.

We now describe the jump rule $f_{\mathbf{x}}$. Let $N_1 \sqcup N_2$ be the partition of the neighbors $N(\mathbf{x})$ of $\mathbf{x}$ given by:

$$N_1 := \mathbf{x} + \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix}, \frac{1}{2}\begin{pmatrix} -1 \\ -\sqrt{3} \end{pmatrix} \right\}; \qquad N_2 := \mathbf{x} + \left\{ \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{3} \end{pmatrix}, \frac{1}{2}\begin{pmatrix} 1 \\ -\sqrt{3} \end{pmatrix} \right\}.$$

The distribution $f_{\mathbf{x}}(s_1)$ is then given by the uniform distribution on N_1, while $f_{\mathbf{x}}(s_2)$ and $f_{\mathbf{x}}(s_3)$ are the uniform distribution on N_2.

Without knowing the hidden states, an outside observer will not be able to predict the future dynamics of this RWHLM even while knowing the past and present location of the walker and rotor configuration, as illustrated in Figure 8.

Note that a non-hidden RWLM is a special case of RWHLM, with S_x $(x \in V)$ being the set of neighbors of x and with $f_x(y)$ $(y \in N(x))$ being the probability distribution concentrated on y. On the other hand, every RWHLM on a simple graph G can be emulated by a non-hidden RWLM on a larger graph (with multiple edges) in the following manner.

Let $G^\times$ be the undirected graph with vertex set $V(G)$ and with an edge incident to x and y in $G^\times$ for each $\{x, y\} \in E(G)$ and each hidden state $s \in S_x$ of the RWHLM. Such an edge is labeled $e(x, y, s)$.

For any $x \in V(G^\times)$, the mechanism of this RWLM on x is the Markov chain with state space the set of edges incident to x in $G^\times$ (instead of the set of neighbors of x), and with probability transition function

$$p_x^\times(e(x, y, s), e(x, y', s')) := p_x(s, s')\,(f_x(s'))(y'),$$

where p_x and f_x are the probability transition function and the jump rule for the RWHLM, respectively.

This RWLM on $G^\times$ emulates the RWHLM on G in the following sense. Let $(X_n, \rho_n, \kappa_n)_{n \geq 0}$ be an RWHLM on G. Start an RWLM $(X_n^\times, \rho_n^\times)_{n \geq 0}$ on $G^\times$ with the following initial configuration:

$$X_0^\times := X_0; \qquad \rho_0^\times(x) := e(x, \rho_0(x), \kappa_0(x)) \quad (x \in V).$$

Then $(X_n, \rho_n)_{n\geq 0}$ is equal in distribution to $(X_n^\times, h(\rho_n^\times))_{n\geq 0}$, where $h(\rho_n^\times)$ is the rotor configuration of G given by $h(\rho_n^\times)(x) := y$ if $\rho_n^\times(x) = e(x, y, s)$ for some $s \in S_x$.

As a consequence of this reduction, we can convert the hidden triangular walk from Example A.2 to a non-hidden random walk with local memory, and then apply a version of Proposition 1.4 for non-simple graphs to conclude that the scaling limit of this hidden triangular walk is a Brownian motion in $\mathbb{R}^2$.