CONSTRUCTIVE APPROACH TO LIMIT THEOREMS FOR RECURRENT DIFFUSIVE RANDOM WALKS ON A STRIP.

D. DOLGOPYAT AND I. GOLDSHEID

Abstract. We consider recurrent diffusive random walks on a strip. We present constructive conditions on Green functions of finite sub-domains which imply a Central Limit Theorem with polynomial error bound, a Local Limit Theorem, and mixing of environment viewed by the particle process. Our conditions can be verified for a wide class of environments including independent environments, quasiperiodic environments, and environments which are asymptotically constant at infinity. The conditions presented deal with a fixed environment, in particular, no stationarity conditions are imposed.

1. Introduction.

The one-dimensional random walk in random environment (RWRE) is a classical model in probability which was first considered in [48] and [31] in 1975. Remarkably, the behavior of the RWRE turns out to be quite different from that of the simple random walk. Perhaps the most famous example of that is the theorem of Sinai ([46]) which states that for the nearest neighbor random walks in the i.i.d. environment in the recurrent case the walker typically is in a $\mathcal{O}(\ln^2 n)$ neighborhood of the origin after n steps.

For walks on $\mathbb{Z}$ with bounded jumps, it was shown that in the recurrent case the Sinai behavior and the classical CLT are the only possible scenarios for two important classes of environments. Namely, [6] proves this for independent environments and [11] considers quasiperiodic Diophantine environments and proves that the CLT holds with probability one. Recently this result from [11] was extended in [20] to RWRE on a strip, a natural generalization of a random walk on $\mathbb{Z}$ with bounded jumps which was introduced in [5]. In fact, it is shown in [20] that in both the i.i.d and the quasiperiodic Diophantine environments the CLT holds in the recurrent case if and only if the potential is bounded (the precise definition of the potential is given in Section 2.4, see equation (2.22)). This is why it is natural and important to study recurrent RWs in a bounded potential. The recurrent RWs in bounded potential are the main object studied in this paper. However, in contrast to [20] we deal with a fixed environment. We develop a constructive approach which relates directly the rate of convergence of ergodic averages for some specific observables to the CLT. For a typical realization of a random environment our results recover the previously known results and, moreover, we obtain new information also for RWRE. Namely, in the quasiperiodic Diophantine case, the CLT is proven in [20] only for a set of environments of full measure, while our present methods imply that the CLT holds for all such environments without exception.

2010 *Mathematics Subject Classification.* Primary: 60K37; Secondary: 60F15.

Key words and phrases. random walks on a strip, Green function, RWRE, CLT, Local Limit Theorem, environment viewed from the particle.

Our approach has several additional benefits. It allows us to

- obtain explicit rate of convergence in the CLT;
- establish the almost sure mixing estimates for environment seen by the particle thus extending the results of [34] and [47];
- prove local limit theorems for several classes of environments;
- apply our method to non stationary environments.

Let us describe the main novel techniques of the present paper which are crucial for our approach. The first one is the asymptotic formula for the Green function in a large finite domain obtained in Section 4. The derivation of this formula relies on an entirely new approach to the analysis of the martingale and the invariant measure equations which was recently discovered in [21]. This approach is further developed in this work and leads to new algebraic properties of the solutions to these equations. The second key ingredient is the weak law of large numbers for the environment viewed by the particle process. Our proof of the law of large numbers relies on the Green function estimates and because of that is more transparent and applicable to a much wider class of observables (the so called self-averaging observables) than the traditional approach based on the ergodic theorem (see e.g. [7, 13]).

The layout of the paper is as follows. In Section 2 we define the model, introduce some notation, and provide the necessary background for RW on a strip. In particular we introduce RW in a bounded potential studied in [20]. In Section 3 we illustrate the main results of our paper by applying them to several important classes of environments. The precise formulation of the main results in the general case are given later since they are of a more technical nature. Section 4 contains bounds for the Green function of the walks with bounded potential. In Section 5 these bounds are used to give a constructive proof of ergodicity for the environment viewed by the particle process, giving a rate of convergence of time averages seen by the walker to the space averages. In particular, this allows us to control the drift and the variance of the increments of the walker on a mesoscopic scale. This allows us, in Section 6, to obtain the Central Limit Theorem by the martingale method and gives an estimate on the rate of convergence for several important classes of environments. In Section 7 we consider environments which have different asymptotic behaviors at $+\infty$ and $-\infty$ and show how the arguments of the previous section could be modified to obtain convergence to the skew Brownian Motion type processes. In Section 8 we use a bootstrap argument to show that the distribution of the walker's position is the same as for the Brownian Motion on a scale which is slightly larger than $\mathcal{O}(1)$. In Section 9 a local limit theorem for hitting times is used to obtain mixing of the environment seen by the particle. In Section 10 the results of Sections 8 and 9 are combined to obtain the local central limit theorem for the walker's position.

2. Definition of the model and some preparatory results.

2.1. Conventions and notation.

2.1. Conventions and notation. The following notations and definitions are used throughout the paper.

All vectors and matrices below will be m-dimensional where m is the width of the strip. The dot product of vectors x and y will be denoted by xy.

$\mathbf{1}$ is a column vector whose components are all equal to 1.

e_i is the vector whose i-th component is 1 and all other components are 0.

For a vector $x = (x_i)$ and a matrix $A = (a(i,j))$ we set

$$\|x\| \stackrel{\text{def}}{=} \max_i |x_i| \text{ which implies } \|A\| = \sup_{\|x\|=1} \|Ax\| = \max_i \sum_j |a(i,j)|.$$

We say that A is strictly positive (and we write $A > 0$), if all its matrix elements satisfy $a(i,j) > 0$. A is called non-negative (and we write $A \geq 0$), if all $a(i,j)$ are non negative. A similar convention applies to vectors. Note that if A is a non-negative matrix then $\|A\| = \|A\mathbf{1}\|$

$\mathbb{S}$ denotes the strip, $\mathbb{S} = \mathbb{Z} \times \{1, \ldots, m\}$. Given a function $h : \mathbb{S} \mapsto \mathbb{R}$, we can define a sequence of $\mathbb{R}^m$-vectors h_n with components $h(n,i)$. Vice versa, given a sequence of vectors h_n, $-\infty < n < \infty$, we define a function $h : \mathbb{S} \mapsto \mathbb{R}$ by setting $h(n,i) = h_n(i)$, where $h_n(i)$ is the i^{th} component of h_n.

2.2. **The Model.** We recall the definition of the RW on a strip from [5]. Let $\mathbb{L}_n = \{(n,i) : 1 \leq i \leq m\}$ be layer n of the strip, $\mathbb{L}_n \subset \mathbb{S}$. In our model, the walker is allowed to jump from a point $(n,i) \in \mathbb{L}_n$ only to points in $\mathbb{L}_{n-1}$, or $\mathbb{L}_n$, or $\mathbb{L}_{n+1}$. To define the corresponding transition kernel consider a sequence ω of triples (P_n, Q_n, R_n), $-\infty < n < \infty$, of $m \times m$ non-negative matrices such that for all $n \in \mathbb{Z}$ the sum $P_n + Q_n + R_n$ is a stochastic matrix:

$$(P_n + Q_n + R_n)\mathbf{1} = \mathbf{1}, \tag{2.1}$$

We say that the sequence ω is the *environment* on the strip $\mathbb{S}$.

The matrix elements of P_n are denoted $P_n(i,j)$, $1 \leq i,j \leq m$, and similar notations are used for Q_n and R_n. We now set

$$\mathfrak{P}(z, z_1) \stackrel{\text{def}}{=} \begin{cases} P_n(i,j) & \text{if} \quad z = (n,i),\ z_1 = (n+1,j), \\ R_n(i,j) & \text{if} \quad z = (n,i),\ z_1 = (n,j), \\ Q_n(i,j) & \text{if} \quad z = (n,i),\ z_1 = (n-1,j), \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

Remark 2.1. The study of one-dimensional RW on $\mathbb{Z}$ with jumps of length $\leq m$ can be reduced to the study of the above model by mapping $n \in \mathbb{Z}$ to $\left(\lfloor \frac{n}{m} \rfloor, n - m\lfloor \frac{n}{m} \rfloor\right) \in \mathbb{S}$, where $\lfloor \cdot \rfloor$ denotes the integer part. We refer the reader to [5] for the formulas for transition matrices in that case and to [6] for more comments concerning this relationship.

For a fixed ω we define a random walk $\xi(t) = (X(t), Y(t))$, $t \geq 0$, on $\mathbb{S}$ in the usual way: for any starting point $z = (n,i) \in \mathbb{S}$ and fixed ω the law $\mathbb{P}_{\omega,z}$ for the Markov chain $\xi(\cdot)$ is given by

$$\mathbb{P}_{\omega,z}\left(\xi(1) = z_1, \ldots, \xi(t) = z_t\right) \stackrel{\text{def}}{=} \mathfrak{P}_\omega(z, z_1)\mathfrak{P}_\omega(z_1, z_2) \cdots \mathfrak{P}_\omega(z_{t-1}, z_t). \tag{2.3}$$

Let Ξ_z be the set of trajectories $\xi(\cdot)$ starting at z. The just defined $\mathbb{P}_{\omega,z}$ is a probability measure on Ξ_z; we denote by $\mathbb{E}_{\omega,z}$ the expectation with respect to this measure.

Remark 2.2. Fix $\mathfrak{i} \in \{1, \ldots, m\}$. With a slight abuse of notation, we shall often write $\mathbb{P}$ and $\mathbb{E}$ for $\mathbb{P}_{\omega,(0,\mathfrak{i})}$ and $\mathbb{E}_{\omega,(0,\mathfrak{i})}$. We shall do that if the environment ω is fixed. Since the strip has finite width, all the results proved in the paper will be uniform with respect to $\mathfrak{i}$.

Also it will often be convenient to write ξ_t for $\xi(t)$ and X_t and Y_t for $X(t)$ and $Y(t)$ respectively.

Throughout the paper we suppose that the following ellipticity conditions are satisfied: there is an $\bar{\varepsilon} > 0$ and a positive integer number $k_0 < \infty$ such that for any $n \in \mathbb{Z}$ and any $i,\, j \in [1, m]$

$$||R_n^{k_0}|| \le 1 - \bar{\varepsilon}, \quad ((I - R_n)^{-1}P_n)(i,j) \ge \bar{\varepsilon}, \text{ and } ((I - R_n)^{-1}Q_n)(i,j) \ge \bar{\varepsilon}. \tag{2.4}$$

Note that $((I - R_n)^{-1}P_n)(i,j)$ (respectively $((I - R_n)^{-1}Q_n)(i,j)$) is the probability that the walker starting from (n, i) reaches $(n + 1, j)$ (respectively $(n - 1, j)$) at the first exit from the layer $\mathbb{L}_n$.

Remark 2.3. Most of our results will be proved for environments which are not random but rather satisfy certain properties (which will be listed in due time). We shall show that it is possible to apply these results to certain important classes of random environments. More precisely, denote by $(\Omega, \mathcal{F}, \mathrm{P}, T)$ the dynamical system where Ω is the space of all sequences $\omega = ((P_n, Q_n, R_n))_{n=-\infty}^{\infty}$ of triples described above, $\mathcal{F}$ is the corresponding natural σ-algebra, P denotes the probability measure on $(\Omega, \mathcal{F})$, and T is the shift operator on Ω defined by $T(P_n, Q_n, R_n) = (P_{n+1}, Q_{n+1}, R_{n+1})$. We shall always suppose that T preserves measure P and is ergodic. The expectation with respect to P will be denoted by E.

To be able to apply the result obtained for deterministic environments in the context of random environment, we shall check that the conditions we need are satisfied by P-almost all random environments.

Remark 2.4. Apart of the probability measures $\mathbb{P}$ and P defined above, we shall quite often use measures which will be denoted by $\mathbf{P}$, with related expectations denoted $\mathbf{E}$, and which describe 'reference' probabilities and expectations related to, e. g., standard normal distribution, standard Wiener processe, well known results concerning martingales, etc. Theorems 3.6, 3.8, Corollary 6.5, Propositions 6.6, 6.7 are examples where this notation is used. In each such case, the precise meaning of $\mathbf{P}(\cdot)$ is obvious from the context.

Denote by $\mathcal{J}$ the following set of triples of $m \times m$ matrices:

$$\mathcal{J} \stackrel{\text{def}}{=} \{(P, Q, R) \,:\, P \ge 0,\, Q \ge 0,\, R \ge 0 \;\text{ and }\; (P + Q + R)\mathbf{1} = \mathbf{1}\}\,.$$

We shall use the following metric on $\Omega = \mathcal{J}^{\mathbb{Z}}$. For $\omega' = \{(P'_n, Q'_n, R'_n)\}$, $\omega'' = \{(P''_n, Q''_n, R''_n)\}$ set

$$\mathbf{d}(\omega', \omega'') = \sum_{n \in \mathbb{Z}} \frac{\|P'_n - P''_n\| + \|Q'_n - Q''_n\| + \|R'_n - R''_n\|}{2^{|n|}}. \tag{2.5}$$

Below, whenever we say that a function defined on Ω is continuous we mean that it is continuous with respect to the topology induced by the metric $\mathbf{d}(\cdot, \cdot)$.

2.3. **Matrices ζ_n, A_n, α_n and some related quantities.** We are now in a position to recall the definitions of several objects most of which were first introduced and studied in [5], [6] and which will play a crucial role in this work.

For a given $\omega \in \Omega$, define a sequence of $m \times m$ stochastic matrices ζ_n as follows. For an integer a let ψ_a be a stochastic matrix. For $n > a$ define matrices $\psi_{n,a}$ recurrently as follows: $\psi_{a,a} = \psi_a$ and

$$\psi_{n,a} = (I - R_n - Q_n\psi_{n-1,a})^{-1}P_n, \quad n = a + 1,\, a + 2,\, \dots. \tag{2.6}$$

It is easy to show (see [5]) that matrices $\psi_{n,a}$ are stochastic. Now for a fixed n define

$$\zeta_n = \lim_{a\to-\infty} \psi_{n,a}. \tag{2.7}$$

As shown in [5, Theorem 1] the limit (2.7) exists and is independent of the choice of the sequence $\{\psi_a\}$.

Next, we define probability row-vectors $\sigma_n = \sigma_n(\omega) = (\sigma_n(\omega,1),\dots,\sigma_n(\omega,m))$ which are associated with the matrices ζ_n. Let $\tilde\sigma_a$ be an arbitrary sequence of probability row-vectors (by which we mean that $\tilde\sigma_a \ge 0$ and $\sum_{i=1}^m \tilde\sigma_a(i) = 1$). Set

$$\sigma_n \overset{\text{def}}{=} \lim_{a\to-\infty} \tilde\sigma_a\zeta_a\dots\zeta_{n-1}. \tag{2.8}$$

By the standard contraction property of the product of stochastic matrices, this limit exists and does not depend on the choice of the sequence $\tilde\sigma_a$ (see [25, Lemma 1]). Vectors σ_n could be equivalently defined as the unique sequence of probability vectors satisfying the infinite system of equations

$$\sigma_n = \sigma_{n-1}\zeta_{n-1}, \quad n\in\mathbb{Z}. \tag{2.9}$$

Combining (2.8) with standard contracting properties of stochastic matrices ζ we obtain for $k>n$ that

$$\|\zeta_n\dots\zeta_{k-1} - (\sigma_k(1)\mathbf{1},\dots,\sigma_k(m)\mathbf{1})\| \le C\theta^{k-n}, \tag{2.10}$$

where $0\le\theta<1$ and C depend only on the $\bar\varepsilon$ from (2.4).

Define

$$\alpha_n = Q_{n+1}(I-R_n-Q_n\zeta_{n-1})^{-1}, \quad A_n = (I-R_n-Q_n\zeta_{n-1})^{-1}Q_n. \tag{2.11}$$

Note that $\alpha_nP_n = Q_{n+1}\zeta_n$ and hence

$$\alpha_n = Q_{n+1}(I-R_n-\alpha_{n-1}P_{n-1})^{-1}. \tag{2.12}$$

Conditions (2.4) imply (see [20, Remark 2.2]) that matrices A_n have the following properties:

$$||A_n|| \le (m\bar\varepsilon)^{-1} \text{ and } A_n(i,j)\ge\bar\varepsilon \tag{2.13}$$

In turn, inequalities (2.13) imply the well known contracting property of the action of matrices A_n. We shall make use of the following version of this property: the limit

$$v_n = \lim_{a\to-\infty} \frac{A_nA_{n-1}\dots A_{a+1}\tilde v_a}{\|A_nA_{n-1}\dots A_{a+1}\tilde v_a\|} \tag{2.14}$$

exists and does not depend on the choice of the sequence of vectors $\tilde v_a\ge 0, ||\tilde v_a|| = 1$. Moreover, there is a θ, $0<\theta<1$, such that

$$\left\| v_n - \frac{A_nA_{n-1}\dots A_{a+1}\tilde v_a}{\|A_nA_{n-1}\dots A_{a+1}\tilde v_a\|}\right\| = O(\theta^{n-a}). \tag{2.15}$$

For the sake of completeness, we prove (2.14) and (2.15) in Appendix B.

Similarly, for any sequence of row-vectors $\tilde l_a\ge 0$, $\|\tilde l_a\| = 1$, define

$$l_n = \lim_{a\to\infty} \frac{\tilde l_a\alpha_{a-1}\dots\alpha_n}{\left\|\tilde l_a\alpha_{a-1}\dots\alpha_n\right\|}. \tag{2.16}$$

Set

$$\lambda_k = \|A_kv_{k-1}\| \text{ and } \tilde\lambda_k = \|l_{k+1}\alpha_k\|. \tag{2.17}$$

Then obviously

$$l_{k+1}\alpha_k = \tilde{\lambda}_k l_k, \quad A_k v_{k-1} = \lambda_k v_k \tag{2.18}$$

and for any $n \geq k$ we have

$$\|A_n A_{n-1} \dots A_k v_{k-1}\| = \lambda_n \dots \lambda_k, \quad \|l_{n+1}\alpha_n \alpha_{n-1} \dots \alpha_k\| = \tilde{\lambda}_n \dots \tilde{\lambda}_k. \tag{2.19}$$

Remark 2.5. It should be emphasized that even though [5, 6] dealt with stationary ergodic environments, the proofs provided in [5], [6] of the existence of the limits (2.7) and (2.14) are in fact working for *all* (and not just P - almost all) sequences ω satisfying (2.4).

Remark 2.6. Note that $m = 1$ corresponds to the random walks on $\mathbb{Z}$ with jumps to the nearest neighbours. In this case $p_n = P_\omega(\xi(t+1) = n+1|\xi(t) = n)$ and $q_n = 1 - p_n$. The above formulae now become very simple, namely $\psi_n = \zeta_n = 1$, $v_n = l_n = 1$, $A_n = \lambda_n = \frac{q_n}{p_n}$, $\alpha_n = \tilde{\lambda}_n = \frac{q_{n+1}}{p_n}$.

In the above considerations, matrices P_n and Q_n play asymmetric roles and it turns out to be useful to 'symmetrize' the situation. Namely, let us introduce stochastic matrices ζ_n^- as the unique sequence of stochastic matrices satisfying the system of equations which is symmetric to (2.6), (2.7)

$$\zeta_n^- = (I - R_n - P_n\zeta_{n+1}^-)^{-1} Q_n, \ -\infty < n < +\infty. \tag{2.20}$$

Next we set

$$A_n^- \stackrel{\text{def}}{=} (I - R_n - P_n\zeta_{n+1}^-)^{-1} P_n, \quad \alpha_n^- = P_{n-1}(I - R_n - P_n\zeta_{n+1}^-)^{-1}. \tag{2.21}$$

All other related objects are introduced similarly.

Matrices ζ_n^-, α_n^-, A_n^-, etc have properties which are similar to those of matrices ζ_n, α_n, A_n etc listed above. All these objects will be used below without further explanations.

2.4. **Walks in bounded potential.** In the context of random walks in random environments, the notion of potential was introduced in [46] in the case of the walks on $\mathbb{Z}$ with jumps to nearest neighbors. The following extension of this definition to the case of random walks on a strip was given in [6].

Definition. A *potential* is a function of n (and ω) defined by

$$\mathcal{U}_n(\omega) \equiv \mathcal{U}_n \stackrel{\text{def}}{=} \begin{cases} \log ||A_n...A_1|| & \text{if } n \geq 1 \\ 0 & \text{if } n = 0 \\ -\log ||A_0...A_{n+1}|| & \text{if } n \leq -1 \end{cases} \tag{2.22}$$

We say that a potential is bounded if there is a constant C_P such that

$$\big|\mathcal{U}_n\big| \leq C_P \text{ for all } n. \tag{2.23}$$

Bounded potentials appear naturally in the study of the following two classes of environments. First, it has been proved in [6] that the recurrence of a random walk in an i.i.d. environment on a strip is equivalent to exactly one of two options: either the potential is bounded or it converges, after the diffusive rescaling, to the Wiener process. In the second case the walk exhibits the Sinai behavior ([6]). Next, in [20] it was shown that for quasiperiodic environments with Diophantine frequencies the potential is bounded if and only if the random walk is recurrent.

2.5. **One useful property of a bounded potential.** Properties (2.23) and (2.13) imply that there is a constant $\tilde{C}_P > 0$ such that for any vector $x \in \mathbb{R}^m$, $x \geq 0$, $(x \neq 0)$, and for any $n > k$

$$e^{-\tilde{C}_P}\|x\|\mathbf{1} \leq A_n \dots A_{k+1}x \leq e^{\tilde{C}_P}\|x\|\mathbf{1}. \tag{2.24}$$

We shall check this statement for the case $k \geq 1$ (other cases are similar).

Note that for any k and $x \geq 0$ the second inequality in (2.13) implies $(A_k x)(i) \geq \bar{\varepsilon}\|x\|$ for all $i, 1 \leq i \leq m$, and so $A_k x \geq \bar{\varepsilon}\|x\|\mathbf{1}$.

By (2.23), $\|A_{k-1}...A_1\| = \|A_{k-1}...A_1\mathbf{1}\| \geq e^{-C_P}$ which is equivalent to saying that there is e_i such that $A_{k-1}...A_1\mathbf{1} \geq e^{-C_P}e_i$. But then

$$A_k A_{k-1}...A_1\mathbf{1} \geq e^{-C_P}A_k e_i \geq e^{-C_P}\bar{\varepsilon}\mathbf{1}.$$

So $\mathbf{1} \leq \bar{\varepsilon}^{-1}e^{C_P}A_k...A_1\mathbf{1}$ and hence

$$A_n...A_{k+1}x \leq \|x\|A_n...A_{k+1}\mathbf{1} \leq \bar{\varepsilon}^{-1}e^{C_P}\|x\|A_n...A_{k+1}A_k...A_1\mathbf{1} \leq \bar{\varepsilon}^{-1}e^{2C_P}\|x\|\mathbf{1}$$

proving the second inequality in (2.24).

Next, by the definition of the norm (and since matrices are positive) we have that

$$A_k...A_1\mathbf{1} \leq e^{C_P}\mathbf{1} \text{ and hence } \mathbf{1} \geq e^{-C_P}A_k...A_1\mathbf{1}.$$

Since $A_n...A_{k+1}x \geq \bar{\varepsilon}\|x\|A_n...A_{k+1}\mathbf{1}$, we obtain

$$A_n...A_{k+1}x \geq \bar{\varepsilon}\|x\|A_n...A_{k+1}\mathbf{1} \geq \bar{\varepsilon}e^{-C_P}\|x\|A_n...A_{k+1}A_k...A_1\mathbf{1} \geq \bar{\varepsilon}e^{-2C_P}\|x\|\mathbf{1}$$

which proves the first inequality in (2.24).

From now on we always suppose that the potential is bounded and *we assume for the rest of the paper that* (2.24) *is satisfied.*

In our previous work we have shown that walks in a bounded potential satisfy the following properties.

(I) There exists a non-constant sequence of column vectors $\mathfrak{m}_n \in \mathbb{R}^m$ (with components $\mathfrak{m}_n(i)$) and a constant K such that

$$|\mathfrak{m}_{n'}(i') - \mathfrak{m}_{n''}(i'')| \leq K \text{ if } |n' - n''| \leq 1, \tag{2.25}$$

and for all n

$$\mathfrak{m}_n = P_n\mathfrak{m}_{n+1} + R_n\mathfrak{m}_n + Q_n\mathfrak{m}_{n-1}. \tag{2.26}$$

The construction of the sequence $\mathfrak{m}_n$ is presented in [20, Section 7]. We recall the probabilistic meaning of (2.26). Let $\mathfrak{m} : \mathbb{S} \mapsto \mathbb{R}$ be a function on a strip and $\mathfrak{m}_n \in \mathbb{R}^m$ be a sequence of column vectors with components $\mathfrak{m}_n(i) = \mathfrak{m}(n, i)$. If $\xi(t) = (X_t, Y_t),\ t \geq 0$, is the RW defined in Subsection 2.2 then the process $M(t) \stackrel{\text{def}}{=} \mathfrak{m}(\xi_t) \equiv \mathfrak{m}(X_t, Y_t) \equiv \mathfrak{m}_{X_t}(Y_t)$ is a martingale if and only if the vectors $\mathfrak{m}_n$ satisfy (2.26).

(II) There exists a positive bounded solution $\rho_n = (\rho_n(1), ..., \rho_n(m)),\ -\infty < n < \infty$, to the equation

$$\rho_n = \rho_{n-1}P_{n-1} + \rho_n R_n + \rho_{n+1}Q_{n+1} \tag{2.27}$$

which also satisfies $\rho_n = \rho_{n+1}\alpha_n$, $\rho_{n+1} = \rho_n\alpha^-_{n+1}$.

Equation (2.27) appears in several contexts. First, for a fixed environment, it describes the invariant measure for the walker. Second, in the case when we deal with stationary environment the solution to (2.27) provides invariant densities for the environment viewed from the particle process. We refer the reader to [21] for a comprehensive analysis of this equation on the strip. The invariant measure equation for the stationary walks on $\mathbb{Z}$ with bounded jumps was studied in [10].

In accordance with conventions of §2.1 we will often write $\mathfrak{m}(\xi_t)$ instead of $\mathfrak{m}_{X_t}(Y_t)$ and $\rho(\xi_t)$ instead of $\rho(X_t, Y_t) = \rho_{X_t}(Y_t)$.

3. Application of results to some classes of environments.

In this section we first discuss examples of important classes of environments. We then state the results which we obtained for these environments as corollaries of our main and more general (but also more technical) theorems proved in this paper.

3.1. **Classes of environments.**

Example 3.1. Quasiperiodic systems. Consider the environment given by

$$(P, Q, R)_n = (\mathcal{P}, \mathcal{Q}, \mathcal{R})(\omega + n\gamma),$$

where $\omega, \gamma \in \mathbb{T}^d$, $\mathbb{T}^d$ is a d-dimensional torus, and $\mathcal{P}, \mathcal{Q}, \mathcal{R} : \mathbb{T}^d \to \mathbb{R}$ are C^∞ functions. γ is called the rotation vector.

RWs in quasiperiodic environments received less attention than the walks discussed in the two other examples below, the main references relevant for our work being [1, 11, 47, 29]. However, its continuous space analogue, the quasiperiodic diffusion, is a classical object in the PDE literature, see [28, 32] and references therein.

For quasiperiodic environment there exists a continuous function $\lambda : \mathbb{T}^d \to \mathbb{R}$ such that $\lambda_n(\omega) = \lambda(\omega + n\gamma)$ (λ_n is defined in (2.17)). We say that γ is *Diophantine* if there are constants K, τ such that for each $k \in \mathbb{Z}^d \setminus 0$ we have

$$d(\gamma k, 2\pi\mathbb{Z}) \geq \frac{K}{|k|^\tau}, \tag{3.1}$$

where d denotes the distance on the line. If γ is Diophantine then $\lambda \in C^\infty(\mathbb{T}^d)$, see Appendix C. The recurrence condition [5] amounts to

$$\int_{\mathbb{T}^d} \ln \lambda(\omega) d\omega = 0, \tag{3.2}$$

where $d\omega$ is normalized Lebesgue measure on the torus. It is proven in [20] that if γ is Diophantine then for every triple $(\mathcal{P}, \mathcal{Q}, \mathcal{R})$ the CLT holds for almost all ω. We note that the Diophantine assumption (3.1) is necessary, since [17] gives examples showing that the CLT need not hold if (3.1) fails. In this paper we obtain additional information in the case when (3.1) and (3.2) hold.

Example 3.2. Independent environments. Here we suppose that $(P, Q, R)_n$ for different n are independent and identically distributed.

The study of RWRE on $\mathbb{Z}$ goes back to [31, 46, 48]. We refer the reader to [51] for a good overview of this subject. The papers most relevant to the present work are also described below after the formulations of Theorems 3.6, 3.8, 3.10. The study of the walks on the strip was initiated in [5], the main references for limit theorems in this setting are [6, 25, 18, 20].

In particular, for independent environments it was shown in [6] that in the recurrent case the Sinai behavior is observed unless $(P, Q, R)_n$ belong to a proper algebraic subvariety in the space of transition matrices. The behavior of the walker on this subvariety was investigated in [20] where it was proven that the solutions to (2.26) and (2.27) with properties (I) and (II) exist.

Example 3.3. Small perturbations of the simple random walk on $\mathbb{Z}$. Consider a random walk on $\mathbb{Z}$ with $p_n = \frac{1}{2} - a_n$, $q_n = \frac{1}{2} + a_n$ where a_n satisfy

$$|a_n| \leq \frac{K}{|n|^\kappa + 1}, \text{ where } \kappa > 1.$$

The condition $\kappa > 1$ is sufficient for recurrence (see e.g. [22]). However, for our results to apply we need one more condition, namely

$$\boldsymbol{\upsilon} = 1 \quad \text{where} \quad \boldsymbol{\upsilon} = \prod_{n \in \mathbb{Z}} \left(\frac{p_n}{q_n}\right). \tag{3.3}$$

Condition (3.3) appears to be restrictive. However, we will show in Corollary 7.3 that it is in fact necessary for the CLT to hold.

The study of environments where the limit $\lim\limits_{n\to\pm\infty} p_n$ exists has a long history. The limit theorems for such walks go back to [35, 50]. The setting which perhaps is the closest to ours can be found in [42] where the Central Limit Theorem is obtained in the *transient* case. Small perturbations of RWRE were studied in [23, 40, 41]. We refer the reader to [43] and references therein for a review of more recent developments. In the present paper we show that such walks fit into the more general framework that we consider.

3.2. **The results.** Next, we describe applications of the general theory developed in this paper to the classes of environments described above. We assume throughout this section that the ellipticity condition (2.4) holds and that the walk is recurrent. In addition, we assume (2.24) (this assumption is only non-trivial in Example 3.2 while in Examples 3.1 and 3.3 it follows from recurrence and ellipticity).

We would like to emphasize that our results by no means are limited to Examples 3.1, 3.2, and 3.3. In fact, Theorems 3.4, 3.5, 3.6, 3.8, and 3.10 below will be obtained as corollaries of more general results, namely, Theorems 6.1, 7.1, 6.8, 9.1, and 10.1 respectively. The statements of these general theorems are more technical and will be given in a due course, after we introduce the necessary background.

Theorems 3.4, 3.6, 3.5, 3.8, and 3.10 below are valid for all environments in Examples 3.1 and 3.3 and for almost all environments in Example 3.2. However in that last case we provide explicit conditions on environment (see equations (6.11), (6.12), (6.13)) which guarantee the validity of these theorems.

Let $\mathcal{N}$ be the standard normal random variable and Φ be the cumulative distribution function of $\mathcal{N}$.

Theorem 3.4. (Functional CLT) *There is a constant $D > 0$ such that the process $W_N(t) = \frac{X(tN)}{\sqrt{N}}$ converges in law as $N \to \infty$ to $\mathcal{W}(t)$-the Brownian Motion with zero mean and variance Dt.*

In fact, we can obtain the functional CLT also for perturbations of our environments which decay at infinity sufficiently fast. Namely, consider a perturbation $\bar{\mathfrak{P}}$ of $\mathfrak{P}$[1] such that

$$\left|\bar{\mathfrak{P}}(z, z') - \mathfrak{P}(z, z')\right| \leq \frac{C}{|n|^\kappa + 1} \text{ where } z = (n, j) \text{ and } \kappa > 1.$$

[1]In the setting of Example 3.3 this means that we allow the environments which do not satisfy (3.3). In fact, it follows from the explicit expression for $\mathfrak{p}$ in terms of $\boldsymbol{\upsilon}$ (see equation (7.16)) that in Example 3.3 $\mathfrak{p} = \frac{1}{2}$ iff $\boldsymbol{\upsilon} = 1$.

Let $\bar{\xi}(t) = (\bar{X}(t), \bar{Y}(t))$ denote the walk in the perturbed environment.

Theorem 3.5. (Functional CLT for the perturbed walk) *There exist constants $\mathfrak{p}$ and $D > 0$ such that the process $\frac{\bar{X}(tN)}{\sqrt{N}}$ converges in law as $N \to \infty$ to the skew Brownian Motion with zero mean, variance Dt, and skewness parameter $\mathfrak{p}$.*

The definition and basic properties of the skew Brownian Motion will be discussed in Section 7. Here we just mention that one way to construct the skew Brownian Motion with skewness parameter $\mathfrak{p}$ is to take the scaling limit for the random walk which is symmetric everywhere except the origin, and which moves to the right from the origin with probability $\mathfrak{p}$ and to the left with probability $1-\mathfrak{p}$ (see [27]).

Theorem 3.6. (Effective CLT) *There are constants D, υ such that for each ε there is a constant C_ε such that*

$$\sup_x \left| \mathbb{P}\left(\frac{X(N)}{\sqrt{DN}} \leq x\right) - \Phi(x) \right| \leq C_\varepsilon N^{-(\upsilon-\varepsilon)}.$$

The exponent υ is explicit. Namely, $\upsilon = \frac{1}{8}$ in Examples 3.1, 3.2, and $\upsilon = \min\left(\frac{\kappa-1}{2}, \frac{1}{8}\right)$ in Example 3.3.

For the next two theorems we assume for Examples 3.1 and 3.2 that the random walk is lazy in the sense that

$$R_n(i,i) \geq \bar{\varepsilon} > 0. \tag{3.4}$$

Remark 3.7. Assumption (3.4) is made for convenience only in order to simplify the statements. Indeed assumption (2.4) implies that the walker can reach all points at the neighboring layer by the time it changes layers. Therefore if we define the stopping times $\tau(n)$ by the conditions $\tau(0) = 0, \quad \tau(n+1) = \min(\tau > \tau(n) : X_\tau \neq X_{\tau(n)})$ then the accelerated walk $\xi^*(n) = \xi(\tau(2n))$ satisfies (3.4). However the natural objects associated with ξ^* (such as solutions to (2.27) etc) have a more complicated form than for ξ so we prefer to impose (3.4).

Theorem 3.8. (Local Limit Theorem) *(a) In Examples 3.1 and 3.2 there are constants a, b such that uniformly for $k_N/\sqrt{N}$ in a compact set for each $y \in \{1, \dots, m\}$ we have*

$$\lim_{N\to\infty} \frac{\mathbb{P}\left(\xi(N) = (k_N, y)\right)}{\mathbf{P}\left(\sqrt{\frac{bN}{a}}\mathcal{N} \in \left[k_N - \frac{1}{2},\, k_N + \frac{1}{2}\right]\right) \rho(k_N, y)} = \frac{1}{a}. \tag{3.5}$$

(b) In Example 3.3 uniformly for $k_N/\sqrt{N}$ in a compact set if k_N and N have the same parity then

$$\lim_{N\to\infty} \frac{\mathbb{P}\left(\xi(N) = k_N\right)}{\mathbf{P}\left(\sqrt{N}\mathcal{N} \in \left[k_N - \frac{1}{2},\, k_N + \frac{1}{2}\right]\right) \rho(k_N)} = 2. \tag{3.6}$$

Remark 3.9. Equation (2.27) defines ρ up to a multiplicative constant. So to complete the statement of Theorem 3.8 one needs to explain how to normalize ρ. This will be done in Section 4 (see equations (4.6) and (4.7)).

It will be shown in Section 6 (see equation (6.35)) that with this choice of normalization, we have in Example 3.3 that $\lim_{|k|\to\infty} \rho(k) = 1$. Thus if in Theorem 3.8(b)

$|k_N| \to \infty$ then (3.6) can be simplified to read

$$\lim_{N\to\infty} \frac{\mathbb{P}\left(\xi(N)=k_N\right)}{\mathbf{P}\left(\sqrt{N}\mathcal{N} \in \left[k_N-\frac{1}{2},\, k_N+\frac{1}{2}\right]\right)} = 2.$$

That is in that case the Local Limit Theorem takes the same form as for the simple random walk.

While the Central Limit Theorem was studied for many classes of RWRE, the Local Limit Theorem is less well understood. We note that there are two different classes of walks where the CLT is known and so it makes sense to study the LLT: ballistic walks are investigated in [2, 19, 38, 47] and balanced walks in [14, 47, 49]. In both cases the Local Limit Theorem takes the same form (3.5), but the meaning of ρ is different: for ballistic walks ρ_z is the expected number of visits to the site z while for recurrent walks it is proportional to the invariant measure of the walk restricted to a finite domain. For this reason different methods are usually employed to study these two cases. In the present paper we adapt the method used in [19] to study the ballistic walks to the recurrent case (our approach is a modification of the method of [25] and is related to extraction of a binomial component approach used in [16]). The universality of this method makes it promising in other problems where the Local Limit Theorem can be expected.

To formulate our last result we need one more definition. In Examples 3.1 and 3.2 a bounded function $h : \mathbb{S} \to \mathbb{R}$ will be called *self-averaging* if there is a constant $\mathfrak{h}$ (the average of h) and a sequence δ_N converging to 0 as $N \to \infty$ such that for each ε, K for each k with $|k| \le K\sqrt{N}$

$$\left| \frac{1}{2\delta_N N^{1/4}} \sum_{l=k-\delta_N N^{1/4}}^{k+\delta_N N^{1/4}} \rho_l h_l - \mathfrak{h} \right| \le \varepsilon \tag{3.7}$$

where h_l is a vector with components $h_l(j) = h(l, j)$, ρ_l is the vector defined in (2.27), whose components are denoted by $\rho_l(j)$, and

$$\rho_l h_l = \sum_{j=1}^{m} \rho_l(j) h(l, j).$$

In Example 3.3 the walk is periodic with period 2, so (3.7) has to be replaced by

$$\left| \frac{1}{2\delta_N N^{1/4}} \sum_{l=-\delta_N N^{1/4}}^{\delta_N N^{1/4}} \rho_{k+2l} h_{k+2l} - \mathfrak{h} \right| \le \varepsilon \text{ and } \left| \frac{1}{2\delta_N N^{1/4}} \sum_{l=-\delta_N N^{1/4}}^{\delta_N N^{1/4}} \rho_{k+1+2l} h_{k+1+2l} - \mathfrak{h} \right| \le \varepsilon. \tag{3.8}$$

The meaning of the notion of the self-averaging will be explained later (see Remark 5.2).

Theorem 3.10. (Mixing of environment viewed by the particle) *If $h : \mathbb{S} \to \mathbb{R}$ is self-averaging then*

$$\lim_{N\to\infty} \mathbb{E}(h(\xi(N))) = \frac{\mathfrak{h}}{a}.$$

where a is the same as in (3.5)

Remark 3.11. The term *mixing* here refers to the fact that the expectation above is asymptotically independent of N. It follows from our proof that it is also independent of the starting point of the walk. Therefore Theorem 3.10 shows that the environment seen by the walker does not remember the remote past of the walker. Results similar to Theorem 3.10 are sometimes called *renewal theorems* since mixing for certain systems allows us to recover the classical renewal theorems. We prefer the term *mixing* since it appears to describe the phenomenon more precisely.

The environment viewed by the particle process is a standard tool in studying the random walk [33, 7, 13]. For ballistic nearest neighbor random walks on $\mathbb{Z}$ in independent environments the mixing of this process was obtained in [30] in the annealed setting. The quenched result was proven in [34] for independent walks under the additional assumption that the fluctuations are diffusive (see [19] for a simple proof). [29, 47] prove mixing for quasiperiodic walks. The results of [30] have been extended to walks on the strip in [45]. In this paper we obtain quenched mixing in both independent and quasiperiodic environments. In fact, the novel feature of our results is that they are applicable to the environments satisfying explicit estimates, so no stationarity is required in our approach.

4. The Green Function.

The main result of this section is the asymptotic expansion of the Green function for the exit from a large interval (see Lemma 4.3). This asymptotic expansion plays a major role in the proofs of our main results: it allows us to compute limits of ratios of various additive functionals of our random walk using moderate deviation estimates from Appendix A.

We begin with a preliminary fact, establishing a relation between two key quantities $\mathfrak{m}_n$ and ρ_n which appear in the expansion of the Green function

Lemma 4.1. *If* $\mathfrak{m}_n$ *satisfies* (2.26) *and* ρ_n *satisfies* (2.27) *then there exist a constant* c *such that for all* n

$$\rho_{n+1}Q_{n+1}(\mathfrak{m}_n - \zeta_n\mathfrak{m}_{n+1}) = c,$$
$$\rho_n P_n(\mathfrak{m}_{n+1} - \zeta^-_{n+1}\mathfrak{m}_n) = -c.$$

This lemma complements [21, Lemmas 4.5 and 4 .6] where other relations between ρ_n and $\mathfrak{m}_n$ are described.

Proof. Let

$$u_n = \mathfrak{m}_n - \zeta_n\mathfrak{m}_{n+1}. \tag{4.1}$$

Then

$$u_n = A_n u_{n-1}. \tag{4.2}$$

Indeed (2.26) can be rewritten as

$$(I - R_n)\mathfrak{m}_n = P_n\mathfrak{m}_{n+1} + Q_n\mathfrak{m}_{n-1} \tag{4.3}$$

Since $\zeta_n = (I - R_n - Q_n\zeta_{n-1})^{-1}P_n$ we have

$$P_n = (I - R_n - Q_n\zeta_{n-1})\zeta_n.$$

Plugging this into (4.3) we get

$$(I - R_n)\mathfrak{m}_n = (I - R_n - Q_n\zeta_{n-1})\zeta_n\mathfrak{m}_{n+1} + Q_n\mathfrak{m}_{n-1}.$$

Subtracting $Q_n\zeta_{n-1}\mathfrak{m}_n$ from both sides we get

$$(I - R_n - Q_n\zeta_{n-1})\mathfrak{m}_n = (I - R_n - Q_n\zeta_{n-1})\zeta_n\mathfrak{m}_{n+1} + Q_n(\mathfrak{m}_{n-1} - \zeta_{n-1}\mathfrak{m}_n)$$

Multiplying both sides by $(I - R_n - Q_n\zeta_{n-1})^{-1}$ and remembering that

$$A_n = (I - R_n - Q_n\zeta_{n-1})^{-1}Q_n$$

we obtain (4.2).

Observe that (2.11) implies that $Q_{n+1}A_n = \alpha_n Q_n$. Hence (4.2) gives

$$\rho_{n+1}Q_{n+1}u_n = \rho_{n+1}Q_{n+1}A_n u_{n-1} = \rho_{n+1}\alpha_n Q_n u_{n-1} = \rho_n Q_n u_{n-1}$$

proving the first claim of the lemma. A similar computation shows that

$$\rho_n P_n(\mathfrak{m}_{n+1} - \zeta^-_{n+1}\mathfrak{m}_n) = c^-.$$

It remains to relate c to c^-. To this end note that

$$\begin{aligned} c &= \rho_{n+1}Q_{n+1}(\mathfrak{m}_n - \zeta_n\mathfrak{m}_{n+1}) \\ &= \rho_{n+1}Q_{n+1}\mathfrak{m}_n - \rho_{n+1}Q_{n+1}(I - R_n - Q_n\zeta_{n-1})^{-1}P_n\mathfrak{m}_{n+1} \\ &= \rho_{n+1}Q_{n+1}\mathfrak{m}_n - \rho_{n+1}\alpha_n P_n\mathfrak{m}_{n+1} = \rho_{n+1}Q_{n+1}\mathfrak{m}_n - \rho_n P_n\mathfrak{m}_{n+1}. \end{aligned}$$

Likewise

$$c^- = \rho_n P_n\mathfrak{m}_{n+1} - \rho_{n+1}Q_{n+1}\mathfrak{m}_n = -c$$

finishing the proof. □

Next, (2.24) and property (2.13) imply the following Lemma.

Lemma 4.2. *Suppose that* (2.4) *is satisfied and the potential* (2.22) *is bounded. Then:*
(i) *there is a bounded solution u_n to $u_n = A_n u_{n-1}$, $\infty < n < \infty$, and such that all entries of all vectors u_n have the same sign.*
(b) *if a sequence of vectors $\tilde u_n$ satisfies $\tilde u_n = A_n\tilde u_{n-1}$, $\infty < n < \infty$, and $\frac{1}{n}\ln\|\tilde u_n\| \to 0$ as $n \to -\infty$ then $\tilde u_n$ is bounded and proportional to u_n.*

Proof. It will be convenient to use the following notation: for $k \le n$

$$\mathbf{A}_n^k = A_n ... A_{k+1}, \text{ with the convention } \mathbf{A}_n^n = I, \quad \mathbf{A}_n^{n-1} = A_n.$$

We make use of v_n from (2.14). To construct u_n for $n \ge 0$ set $u_0 = v_0$ and define $u_n = \mathbf{A}_n^0 u_0$ for $n \ge 1$. Obviously, $u_n = A_n u_{n-1}$ if $n \ge 1$.

For $n \le -1$, set $u_n \overset{\text{def}}{=} \lambda_0^{-1}...\lambda_{n+1}^{-1}v_n$. Then, taking into account (2.18), we have for $n \le 0$:

$$A_n u_{n-1} = \lambda_0^{-1}...\lambda_{n+1}^{-1}\lambda_n^{-1}A_n v_{n-1} = \lambda_0^{-1}...\lambda_{n+1}^{-1}v_n = u_n.$$

The vectors u_n are strictly positive since $v_n > 0$ for all n. Finally, since $u_n = A_n \dots A_{k+1}u_k$ for any $n > k$, the inequalities (2.24) imply that $e^{-\tilde C_P}\|u_k\| \le \|u_n\| \le e^{\tilde C_P}\|u_k\|$ and so this solution is bounded. This completes the proof of (i).

Proof of (ii). We shall show that for any fixed n there is a c such that $\tilde u_n = cu_n$.

For $k \le n$, present $\tilde u_k = \tilde u_k^+ - \tilde u_k^-$, where $\tilde u_k^+ \ge 0$ and $\tilde u_k^- \ge 0$ are, respectively, the positive and the negative part of $\tilde u_k$. Then $\|\tilde u_k^+\| \le \|\tilde u_k\|$ and $\|\tilde u_k^-\| \le \|\tilde u_k\|$. It follows from (2.24) that $\|\mathbf{A}_n^k\tilde u_k^+\| = \mathcal{O}(\|\tilde u_k^+\|) = \mathcal{O}(\|\tilde u_k\|)$. This, together with (2.15) implies $\mathbf{A}_n^k\tilde u_k^+ = \|\mathbf{A}_n^k\tilde u_k^+\|(v_n + \mathcal{O}(\theta^{n-k})) = \|\mathbf{A}_n^k\tilde u_k^+\|v_n + \mathcal{O}(\|\tilde u_k\|\theta^{n-k})$. Finally,

$$\mathbf{A}_n^k\tilde u_k^+ = \|\mathbf{A}_n^k\tilde u_k^+\|\|u_n\|^{-1}u_n + \mathcal{O}(\|\tilde u_k\|\theta^{k-n})$$

since $v_n = u_n/\|u_n\|$. Similarly, $\mathbf{A}_n^k\tilde u_k^- = \|\mathbf{A}_n^k\tilde u_k^+\|\|u_n\|^{-1}u_n + \mathcal{O}(\|\tilde u_k\|\theta^{n-k})$. But $\tilde u_n = \mathbf{A}_n^k\tilde u_k^+ - \mathbf{A}_n^k\tilde u_k^-$ we have

$$\tilde u_n = (\|\mathbf{A}_n^k\tilde u_k^+\| - \|\mathbf{A}_n^k\tilde u_k^-\|)\|u_n\|^{-1}u_n + \mathcal{O}(\|\tilde u_k\|\theta^{n-k}).$$

The growth of $\|\tilde{u}_k^-\|$ is sub-exponential and therefore, sending $k \to -\infty$, we see that $c_n \overset{\text{def}}{=} \lim_{k\to-\infty}(\|\mathbf{A}_n^k \tilde{u}_k^+\| - \|\mathbf{A}_n^k \tilde{u}_k^-\|)\|u_n\|^{-1}$ exists and $\tilde{u}_n = c_n u_n$. But then $c_n u_n = c_{n-1} A_n u_{n-1}$ by the definition of $\tilde{u}_n$ which, together with (4.2), implies that $(c_n - c_{n-1})u_n = 0$. Hence $c_n = c_{n-1} = \text{const}$. □

Returning to the main equation (2.26) for $\mathfrak{m}_n$, we have now two possibilities.

(1) $u_n \equiv 0$. In this case, (4.1) implies that for each $k \geq 1$, $\mathfrak{m}_n = \zeta_n \zeta_{n+1} \dots \zeta_{n+k} \mathfrak{m}_{n+k}$. Sending k to infinity and using contracting properties of stochastic matrices we see that $\mathfrak{m}_n \equiv c\mathbf{1}$ for some constant c.

(2) u_n is non zero, so its entries have the same sign. Then $c = \rho_{n+1} Q_{n+1} u_n \neq 0$.

We note that in the second case the martingale increases faster than some linear function. Namely there are constants C_1, C_2 such that for any $n, k \in \mathbb{Z}$, $k \geq 0$ and for any $i, j \in \{1 \dots m\}$ we have

$$\frac{k}{C_1} - C_2 \leq \mathfrak{m}_{n+k}(j) - \mathfrak{m}_n(i) \leq C_1 k + C_2. \tag{4.4}$$

Indeed, iterating (4.1) we obtain

$$\mathfrak{m}_n = \sum_{r=0}^{k-1} u_{n,r} + \zeta_n \dots \zeta_{n+k-1} \mathfrak{m}_{n+k} \tag{4.5}$$

where $u_{n,r} = \zeta_n \dots \zeta_{n+r-1} u_{n+r}$. Note that by (2.24) the components of u_n are uniformly bounded from above and bounded away from zero. Since ζ's are stochastic, we conclude that the components of $u_{n,r}$ are uniformly bounded from above and bounded away from zero. Since $\mathfrak{m}$ has bounded increments we have that for each $j = 1, \dots, m$, $\mathfrak{m}_{n+k} = \mathfrak{m}_{n+k}(j)\mathbf{1} + \mathcal{O}(1)$. Plugging this into (4.5) and using that $\zeta_n \dots \zeta_{n+k-1}\mathbf{1} = \mathbf{1}$ we obtain (4.4).

In this paper we deal with the case where the martingale $\mathfrak{m}$ is non-trivial. Moreover, for the rest of the paper (except for the Section 7) we assume that $\mathfrak{m}$ is asymptotically linear and that $\mathfrak{m}$ and ρ are normalized so that

$$\lim_{n\to\infty} \frac{\mathcal{M}_n}{n} = m, \tag{4.6}$$

where $\mathcal{M}_n = \sum_{j=1}^m \mathfrak{m}_n(j)$ and

$$\rho_n P_n (\mathfrak{m}_{n+1} - \zeta_{n+1}^- \mathfrak{m}_n) = \frac{1}{2m}. \tag{4.7}$$

Note that (4.6) and (2.25) imply that

$$\lim_{n\to\infty} \frac{\mathfrak{m}_n(y)}{n} = 1 \tag{4.8}$$

uniformly in $y \in \{1, \dots, m\}$.

Asymptotically linear martingales exist in Examples 3.1–3.3. In fact, in Section 6.4 we will establish a stronger result (6.11).

Consider the Green function

$$G_{a,b}((k,i);(n,j)) = \mathbb{E}(\eta_{(n,j)} | \xi(0) = (k,i)),$$

where $\eta_{(n,j)}$ is the number of visits to (n,j) by the walk starting at (k,i) before it hits the segment $[a,b]$.

Lemma 4.3. *For $x,\ y \in (a,b)$*

$$G_{a,b}((k,i);(n,j)) = \mathcal{G}(\mathcal{M}_k, \mathcal{M}_n, \mathcal{M}_a, \mathcal{M}_b)\rho_n(j) + \mathcal{O}(1)$$

where

$$\mathcal{G}(x,y,a,b) = \frac{2(\min(x,y)-a)(b-\max(x,y))}{(b-a)}.$$

Proof. Let

$$\tilde{G}_{a,b}((k,i);n) = \sum_j G_{a,b}((k,i),(n,j)). \tag{4.9}$$

A simple computation with Markov chains [18, Appendix A] shows that

$$G_{a,b}((k,i),(n,j)) = \tilde{G}_{a,b}((k,i),n)\frac{\rho_n(j)}{\rho_n \mathbf{1}} + \mathcal{O}(1)$$

so it suffices to show that

$$\tilde{G}_{a,b}((k,i);n) = \mathcal{G}(\mathcal{M}_k, \mathcal{M}_n, \mathcal{M}_a, \mathcal{M}_j)\rho_n \mathbf{1} + \mathcal{O}(1). \tag{4.10}$$

We consider first the case where $k=n$. Let $d = \min(n-a, b-n)$. Denote $\boldsymbol{\eta}_n = \eta_n \mathbf{1}$. Let $\mathfrak{p}_{n,a,b}$ be the vector

$$\mathfrak{p}_{n,a,b}(i) = \mathbb{P}(\boldsymbol{\eta}_n = 1 | \xi(0) = (n,i)).$$

We claim that

$$\mathfrak{p}_{n,a,b} = \mathcal{O}(1/d). \tag{4.11}$$

Without loss of generality we may assume that $d = b-n$. By (4.4) there exists k such that for any n, i, j $\mathfrak{m}_{n+k}(j) \geq \mathfrak{m}_n(i)+1$. Consider our walk started from $(n+k, j)$. Let $\mathfrak{s}$ be the first time this walk reaches either $\mathbb{L}_n$ or $\mathbb{L}_b$. Applying the Optional Stopping Theorem to this stopping time gives

$$\mathfrak{m}_{n+k}(j) = \sum_{i=1}^m \mathbb{P}(\xi_{\mathfrak{s}} = (n,i))\mathfrak{m}_n(i) + \sum_{i=1}^m \mathbb{P}(\xi_{\mathfrak{s}} = (b,i))\mathfrak{m}_b(i).$$

Rewriting this identity as

$$\sum_{i=1}^m \mathbb{P}(\xi_{\mathfrak{s}} = (n,i))[\mathfrak{m}_{n+k}(j) - \mathfrak{m}_n(i)] = \sum_{i=1}^m \mathbb{P}(\xi_{\mathfrak{s}} = (b,i))[\mathfrak{m}_b(i) - \mathfrak{m}_{n+k}(j)] \tag{4.12}$$

By our choice of k the LHS is at least $1 - \mathbb{P}(\xi_{\mathfrak{s}} \in \mathbb{L}_b)$ while by (4.4) the RHS is at most $2C_1 d\, \mathbb{P}(\xi_{\mathfrak{s}} \in \mathbb{L})$. Thus

$$1 \geq 2C_1 d \mathbb{P}(\xi_{\mathfrak{s}} \in \mathbb{L}) \text{ or } \mathbb{P}(\xi_{\mathfrak{s}} \in \mathbb{L}_b) = \mathcal{O}(1/d).$$

In other words, if the walker starts from the layer $n+k$ then the probability that it would not visit $\mathbb{L}_n$ before reaching $\mathbb{L}_b$ is $\mathcal{O}(1/d)$. By the same argument, if the walker starts from the layer $n-k$ then the probability that it would not visit $\mathbb{L}_n$ before reaching $\mathbb{L}_a$ is $\mathcal{O}(1/d)$. Since the walker starting from $\mathbb{L}_n$ should visit $\mathbb{L}_{n-k}$ before reaching $\mathbb{L}_a$ and it should visit $\mathbb{L}_{n+k}$ before reaching $\mathbb{L}_b$, (4.11) follows.

Let $\tilde{S}_n$ be the matrix with components $\tilde{S}_n(i,j)$ where $\tilde{S}_n(i,j)$ is the probability that the walker starting from (n,i) returns to $\mathbb{L}_n$ for the first time at (n,j) given that it does not visit $\mathbb{L}_a$ or $\mathbb{L}_b$ in between. Let $\tilde{\pi}_n$ be the stationary distribution for $\tilde{S}_n$.

Also let S_n be the matrix with components $S_n(i,j)$, where $S_n(i,j)$ is the probability that the walker starting from (n,i) returns to $\mathbb{L}_n$ for the first time at (n,j). Thus

$$S_n = Q_n\zeta_{n-1} + R_n + P_n\zeta^-_{n+1}.$$

Let π_n be the stationary distribution of S_n. Note that π_n can be expressed in terms of ρ_n as $\pi_n(j) = \frac{\rho_n(j)}{\rho_n\mathbf{1}}$. Since, by (4.11), S_n and $\tilde{S}_n$ differ by conditioning on a set of measure $1-\mathcal{O}\left(\frac{1}{d}\right)$ we have $\tilde{S}_n = S_n + \mathcal{O}\left(\frac{1}{d}\right)$. Moreover we can write

$$\tilde{S}_n = Q_n\tilde{\zeta}_{n-1} + R_n + P_n\tilde{\zeta}^-_{n+1}$$

where $\tilde{\zeta}_{n-1}(i,j)$ and $\zeta^-_{n+1}(i,j)$ are the probabilities that the walker starting from $(n-1,i)$ (respectively $(n+1,i)$) returns to $\mathbb{L}_n$ for the first time at (n,j) given that it does not visit $\mathbb{L}_a$ or $\mathbb{L}_b$ in between. Then

$$\tilde{\zeta}_{n-1} = \zeta_{n-1} + \mathcal{O}\left(\frac{1}{d}\right), \quad \tilde{\zeta}^-_{n+1} = \zeta^-_{n+1} + \mathcal{O}\left(\frac{1}{d}\right). \tag{4.13}$$

Due to exponential mixing of both S_n and $\tilde{S}_n$ (which is guaranteed by condition (2.4)) we have that

$$\tilde{\pi}_n = \pi_n + \mathcal{O}\left(\frac{1}{d}\right) = \frac{\rho_n}{\rho_n\mathbf{1}} + \mathcal{O}\left(\frac{1}{d}\right).$$

We have

$$\mathbb{P}(\boldsymbol{\eta}_n = N+1|\boldsymbol{\eta}_n > N, \xi(0) = (n,i)) = e_i\tilde{S}_n^N\mathfrak{p}_{n,a,b}.$$

By the first step analysis

$$\mathfrak{p}_{n,a,b}(i) = \sum_{j=1}^m Q_n(i,j)\mathfrak{p}^-_n(j) + \sum_{j=1}^m P_n(i,j)\mathfrak{p}^+_n(j) \tag{4.14}$$

where $\mathfrak{p}^\pm_n(j)$ is the probability that the walker starting from $(n\pm1,j)$ does not return to $\mathbb{L}_n$ before visiting the boundary of the segment a, b. By the Optional Stopping Theorem and (2.25)

$$\mathfrak{p}^-_n = \frac{\tilde{\zeta}_{n-1}\mathfrak{m}_n - \mathfrak{m}_{n-1}}{\frac{1}{m}(\mathcal{M}_n - \mathcal{M}_a)} + \mathcal{O}\left(\frac{1}{d^2}\right), \quad \mathfrak{p}^+_n = \frac{\tilde{\zeta}^-_{n+1}\mathfrak{m}_n - \mathfrak{m}_{n+1}}{\frac{1}{m}(\mathcal{M}_b - \mathcal{M}_n)} + \mathcal{O}\left(\frac{1}{d^2}\right). \tag{4.15}$$

From (4.14) and (4.15) we get using (4.13) that

$$\mathfrak{p}_{n,a,b}(i) = me_i\left(Q_n\frac{\zeta_{n-1}\mathfrak{m}_n - \mathfrak{m}_{n-1}}{\mathcal{M}_n - \mathcal{M}_a} + P_n\frac{\zeta^-_{n+1}\mathfrak{m}_n - \mathfrak{m}_{n+1}}{\mathcal{M}_b - \mathcal{M}_n}\right) + \mathcal{O}\left(\frac{1}{d^2}\right). \tag{4.16}$$

Since $e_i\tilde{S}_n^N = \tilde{\pi}_n(1+\mathcal{O}(\theta^N))$ for some $\theta<1$, it follows that

$$e_i\tilde{S}_n^N\mathfrak{p}_{n,a,b} = \tilde{\pi}_n\mathfrak{p}_{n,a,b} + \mathcal{O}\left(\theta^N\right).$$

From (4.16), Lemma 4.1 and (4.7) we get

$$\begin{aligned}\tilde{\pi}_n\mathfrak{p}_{n,a,b} &= \frac{m}{\rho_n\mathbf{1}}\left(\frac{\rho_nQ_n(\mathfrak{m}_{n-1} - \zeta_{n-1}\mathfrak{m}_n)}{\mathcal{M}_n - \mathcal{M}_a} + \frac{\rho_nP_n(\mathfrak{m}_{n+1} - \zeta^-_{n+1}\mathfrak{m}_n)}{\mathcal{M}_b - \mathcal{M}_n} + \mathcal{O}\left(\frac{1}{d^2}\right)\right)\\ &= \frac{1}{2\rho_n\mathbf{1}}\left(\frac{1}{\mathcal{M}_n - \mathcal{M}_a} + \frac{1}{\mathcal{M}_b - \mathcal{M}_n} + \mathcal{O}\left(\frac{1}{d^2}\right)\right).\end{aligned}$$

It follows that $\tilde{\pi}_n \mathfrak{p}_{n,a,b} \boldsymbol{\eta}_n$ has asymptotically exponential distribution with parameter 1 and hence

$$\tilde{G}_{a,b}((n,i);n) = \frac{1}{\tilde{\pi}_n \mathfrak{p}_{n,a,b}} + \mathcal{O}(1) = \frac{2(\mathcal{M}_n - \mathcal{M}_a)(\mathcal{M}_b - \mathcal{M}_n)}{(\mathcal{M}_b - \mathcal{M}_a)}(\rho_n \mathbf{1}) + \mathcal{O}(1).$$

Next for $k < n$ let

$$\mathcal{P}_a((k,i);(n,j)) = \mathbb{P}(\xi \text{ reaches } \mathbb{L}_n \text{ at } (n,j) \text{ before reaching } \mathbb{L}_a | \xi(0) = (k,i)).$$

Then

$$\tilde{G}_{a,b}((k,i);n) = \sum_j \mathcal{P}_a((k,i);(n,j)) \tilde{G}_{a,b}((n,j);n)$$
$$= \left[\frac{2(\mathcal{M}_n - \mathcal{M}_a)(\mathcal{M}_b - \mathcal{M}_n)}{(\mathcal{M}_b - \mathcal{M}_a)}(\rho_n \mathbf{1}) + \mathcal{O}(1)\right] \sum_j \mathcal{P}_a((k,i);(n,j))$$

Note that

$$\mathcal{P}_a((k,i);n) := \sum_j \mathcal{P}_a((k,i);(n,j)) = \mathbb{P}(\xi \text{ reaches } \mathbb{L}_n \text{ before reaching } \mathbb{L}_a | \xi(0) = (k,i)).$$

Applying again the Optional Stopping Theorem to the stopping time $\mathfrak{s}$ which is the first time the walker reaches either $\mathbb{L}_a$ or $\mathbb{L}_b$ we we get

$$(\mathcal{M}_n + \mathcal{O}(1))\mathcal{P}_a((k,i);n) + (\mathcal{M}_a + \mathcal{O}(1))(1 - \mathcal{P}_a((k,i);n)) = \mathcal{M}_k + \mathcal{O}(1).$$

Hence

$$\mathcal{P}_a((k,i);n) = \frac{\mathcal{M}_k - \mathcal{M}_a}{\mathcal{M}_n - \mathcal{M}_a} + \mathcal{O}\left(\frac{1}{d}\right).$$

This proves (4.10) for $k < n$. The case $k > n$ is analyzed similarly. □

Remark 4.4. (4.12) also shows that there is a constant c such that for each $y \in \{1, \dots, m\}$

$$\mathbb{P}(\boldsymbol{\eta}_n = 1 | \xi(0) = (n,y)) \leq \frac{c}{d}. \tag{4.17}$$

This bound will be useful in Section 6.

5. Environment viewed by the particle: the Law of Large Numbers.

From now on we consider only those environments which, in addition to (2.24), (4.6), (4.7), satisfy the following assumption: there exists a constant a such that

$$\lim_{N \to \pm\infty} \frac{1}{|N|} \sum_{n=0}^{N-1} \rho_n \mathbf{1} = a. \tag{5.1}$$

Examples 3.1–3.3 satisfy (5.1). In fact, in Section 6.4 we prove a stronger result (6.12).

Let $h : \mathbb{S} \to \mathbb{R}$ be a bounded function and h_n be a sequence of vectors with components $h_n(i) = h(n,i)$. Let

$$H_N = \sum_{n=0}^{N-1} h(\xi(n)). \tag{5.2}$$

In this section we establish the following result.

Lemma 5.1. *Suppose that h_n is such that*

$$\lim_{N\to\pm\infty} \frac{1}{|N|}\sum_{n=0}^{N-1} \rho_n h_n = \mathfrak{h} \tag{5.3}$$

for some constant $\mathfrak{h}$. Then $\frac{H_N}{N}$ converges in probability, as $N\to\infty$, to $\frac{\mathfrak{h}}{a}$.

Remark 5.2. Assumptions (5.1) and (5.3) imply that h is an extensive observable, that is there exists the finite volume limit

$$\lim_{R\to\pm\infty} \frac{\int_{\mathbb{S}_R} h(z)d\mu(z)}{\mu(\mathbb{S}_R)} = \frac{\mathfrak{h}}{a}$$

where $\mathbb{S}_R$ is the set of points in the strip $\mathbb{S}$ such that the x coordinate is between 0 and R and μ is the invariant measure for our walk: for $A\in\mathbb{S}$

$$\mu(A) = \sum_{z\in A}\rho(z).$$

We refer the reader to [37] for a discussion of the ergodic properties of extensive observables.

A typical application of Lemma 5.1 is the following. Suppose that the environment $\omega = ((P_n,Q_n,R_n))_{n=-\infty}^{\infty}$ is as in Remark 2.3 and set $h_n = 1_{(T^n\omega,Y_n)\in\mathcal{A}}\mathbf{1}$, where $\mathcal{A}\subset\Omega\times\{1,\dots,m\}$ (this defines a function h - see Section 2.1). Then $\frac{H_N}{N}$ describes how often the walker sees the environment from $\mathcal{A}$. For example, one can ask how often the drift or the variance of the walker's increment are of a certain size. Quite often the law of large numbers for H_N is obtained as a consequence of ergodicity of the *environment viewed by the particle* process, see e.g. [7]. This approach, however, makes it difficult to control the exceptional zero measure set in the ergodic theorem. In this section we present a different argument which allows one to obtain explicit sufficient conditions for the law of large numbers (namely, (5.3)).

Proof of Lemma 5.1. Let us first describe the idea of the proof. Fix $\varepsilon>0$. We need to show that $H_N - \left(\frac{\mathfrak{h}}{a}-\varepsilon\right)N$ is positive for large N while $H_N - \left(\frac{\mathfrak{h}}{a}+\varepsilon\right)N$ is negative for large N with probability close to 1. To this end we divide the sum (5.2) into blocks. Choose a small constant δ (the exact requirements on δ will be explained later, see the proof of (5.9) below) and let $L_N = \lfloor\delta\sqrt{N}\rfloor$. We will consider our random walk only at the moments when $X(t)$ visits the nodes of the lattice $L_N\mathbb{Z}$, more precisely, when X moves from one node to the next. That is, define $\tau_0=0$, and for $k>0$ let

$$\tau_k = \min(j\geq\tau_{k-1} : X_j = X_{\tau_{k-1}} - L_N \text{ or } X_j = X_{\tau_{k-1}} + L_N). \tag{5.4}$$

We would like to use the results of Section 4 to show $H^+_{(k,\varepsilon)} := \sum_{n=1}^{\tau_k}\left[h(\xi(n)) - \left(\frac{\mathfrak{h}}{a}-\varepsilon\right)\right]$ is a submartingale and $H^-_{(k,\varepsilon)} := \sum_{n=1}^{\tau_k}\left[h(\xi(n)) - \left(\frac{\mathfrak{h}}{a}+\varepsilon\right)\right]$ is a supermartingale with respect to the natural filtration and then use the large deviation estimates for supermartgales from Appendnix A. However, for a fixed δ, (4.6) and (5.3) only allow us to control the nodes of $L_N\mathbb{Z}$ which are not too far from the origin, so an additional cut off is required.

Using the maximal inequality for martingales we can find a constant K such that

$$\mathbb{P}\left(\max_{t\in[0,N]}|\mathfrak{m}(\xi(t))|\geq\frac{K\sqrt{N}}{2}\right)\leq\frac{\varepsilon}{2}.$$

Now (4.8) gives

$$\mathbb{P}\left(\max_{t\in[0,N]}|X(t)|\geq K\sqrt{N}\right)\leq\frac{\varepsilon}{2}. \tag{5.5}$$

Denoting $a_k=X_{\tau_{k-1}}-L_N$, $b_k=X_{\tau_{k-1}}+L_N$ we have

$$\mathbb{E}(\tau_k-\tau_{k-1}|\xi_{\tau_{k-1}})=\sum_{n=a_k}^{b_k}\sum_{j=1}^{m}G_{a_k,b_k}(\xi_{\tau_{k-1}};(n,j)). \tag{5.6}$$

We claim that for each K we have

$$\mathbb{E}(\tau_k-\tau_{k-1}|\xi_{\tau_{k-1}})\sim aL_N^2 \tag{5.7}$$

provided that N is large enough and $|X_{\tau_{k-1}}|\leq K\sqrt{N}$.

To prove (5.7) divide the segment $[a_k,b_k]$ into subsegments $I_j=[s_j,s_{j+1}]$ of length $[\tilde{\delta}\sqrt{N}]$ where $\tilde{\delta}\ll\delta$. (4.6), Lemma 4.3, and (5.1) show that the contribution to (5.6) of terms with $n\in I_j$ is asymptotic to

$$\tilde{\delta}_N a\mathcal{G}(0,s_j-X_{\tau_{k-1}};-L_N,L_N).$$

Summing over the intervals I_j we obtain (5.7).

Let

$$\hat{k}=a^{-1}\delta^{-2}, \tag{5.8}$$

$T_N=\tau_{\hat{k}}$. We claim that if δ is sufficiently small then

$$\mathbb{P}\left(\left|\frac{T_N}{N}-1\right|>\varepsilon\right)<\varepsilon. \tag{5.9}$$

Indeed define a sequence $\tilde{\tau}_k$ such that $\tilde{\tau}_0=0$ and

$$\tilde{\tau}_k-\tilde{\tau}_{k-1}=\begin{cases}\tau_k-\tau_{k-1} & \text{if } |\xi_{\tau_{k-1}}|\leq K\sqrt{N}\\ aL_N^2 & \text{otherwise.}\end{cases}$$

We want to estimate $\mathbb{P}(\tilde{\tau}_{\hat{k}}\geq(1+\varepsilon)N)$. To this end we apply Proposition A.1 from Appendix A with

$$\Delta_k=\frac{(\tilde{\tau}_k-\tilde{\tau}_{k-1})}{aL_N^2}-(1+\varepsilon).$$

To apply this proposition we need to check conditions (A.1) and (A.2). For the case at hand, (A.1) follows from (5.7). To prove (A.2) we use that there exists a constant $\theta<1$ such that for each $K\in\mathbb{R}$ there is a constant $N_0=N_0(K)$ such that if $N\geq N_0$ and $|\xi_{\tau_{k-1}}|\leq K\sqrt{N}$ then for all $l\in\mathbb{N}$

$$\mathbb{P}(\tau_k-\tau_{k-1}>2alL_N^2)<\theta^l. \tag{5.10}$$

Indeed similarly to (5.7) one can show that if $N\geq N_0(K)$ then for each (x,y) such that $|x-x(\xi_{\tau_{k-1}})|\leq L_N$ for each $s\in\mathbb{Z}$ we have

$$\mathbb{E}(\tau_k-s|\xi(s)=(x,y))\leq 1.1aL_N^2.$$

Combining this with the Markov inequality we see that for any stopping time $\mathfrak{s}$

$$\mathbb{P}(\tau_k > \mathfrak{s} + 2aL_N^2 | \tau_k \geq \mathfrak{s}) \leq \frac{1.1}{2}. \tag{5.11}$$

Applying (5.11) with $\mathfrak{s} = \tau_{k-1}$ we obtain (5.10) with $l = 1$. (5.10) for $l > 1$ follows by induction on l by applying (5.11) with $\mathfrak{s} = \tau_{k-1} + 2a(l-1)$.

Now Proposition A.1 gives

$$\mathbb{P}(\tilde{\tau}_{\hat{k}} \geq (1+\varepsilon)N) \leq e^{-\bar{c}\sqrt{\varepsilon \hat{k}}}.$$

Likewise

$$\mathbb{P}(\tilde{\tau}_{\hat{k}} \leq (1-\varepsilon)N) \leq e^{-\bar{c}\sqrt{\varepsilon \hat{k}}}.$$

Combining the last two displays with (5.8) we see that for large N

$$\mathbb{P}\left(\left|\frac{\tilde{\tau}_{\hat{k}}}{N} - 1\right| > \varepsilon\right) < \frac{\varepsilon}{2}. \tag{5.12}$$

By our choice of K (see (5.5)), for large N we have

$$\mathbb{P}(\tilde{\tau}_{\hat{k}} \neq T_N) < \frac{\varepsilon}{2}. \tag{5.13}$$

Combining (5.12) and (5.13) we obtain (5.9).

Next, similarly to (5.7) we get

$$\mathbb{E}(H_{\tau_k} - H_{\tau_{k-1}} | \xi_{\tau_{k-1}}) \sim \frac{\mathfrak{h} L_N^2}{2} \tag{5.14}$$

and similarly to (5.9) we get (possibly, after decreasing δ) that

$$\mathbb{P}\left(\left|\frac{H_{\tau_{\bar{k}}}}{N} - \frac{\mathfrak{h}}{a}\right| > \varepsilon\right) < \varepsilon.$$

Indeed we can apply Proposition A.1 since (A.1) follows by (5.14) while (A.2) follows from (5.10) since h is bounded so for some constant C

$$|H_{\tau_k} - H_{\tau_{k-1}}| \leq C(\tau_k - \tau_{k-1}).$$

Also since h is bounded, $\dfrac{|H_{\tau_{\bar{k}}} - H_N|}{N} \leq C\left|\dfrac{\tau_{\bar{k}}}{N} - 1\right|$ and so (5.12) and (5.13) give

$$\mathbb{P}\left(\frac{|H_{\tau_{\bar{k}}} - H_N|}{N} > C\varepsilon\right) < \varepsilon. \tag{5.15}$$

Since ε is arbitrary, (5.14) and (5.15) prove the lemma. □

Remark 5.3. We note that the information on X_{τ_k} obtained in the proof of Lemma 5.1, especially (5.14), will play a crucial role in the sequel. In particular, it will be used in Section 6 to show that, under appropriate assumptions, X_{τ_k}/L_N is well approximated by the simple random walk. Passing to the limit as $N \to \infty$, $\delta_N \to 0$ we shall obtain the CLT for $X(t)$.

6. The Central Limit Theorem

6.1. **Sufficient conditions for the CLT.** In this section, with a slight abuse of notation, we write ξ_{Nt}, X_{Nt}, and Y_{Nt} for $\xi_{\lfloor Nt \rfloor}$, $X_{\lfloor Nt \rfloor}$, $Y_{\lfloor Nt \rfloor}$ respectively.

Denote $W_N(t) = \frac{X_{Nt}}{\sqrt{N}}$, where $t \in [0, 1]$. Let $\mathfrak{q}_n$ be a column vector with components

$$\begin{aligned}\mathfrak{q}_n(i) &= \mathbb{E}\left((\mathfrak{m}(\xi_{k+1}) - \mathfrak{m}(\xi_k))^2 \middle| \xi_k = (n, i)\right) \\ &= \sum_{j' \in \{-1,0,1\}, 1 \le i' \le m} \mathfrak{P}\left((n, i), (n + j', i')\right)\left(\mathfrak{m}_{n+j'}(i') - \mathfrak{m}_n(i)\right)^2,\end{aligned}$$

where $\mathfrak{P}(\cdot, \cdot)$ are the transition probabilities (2.2) and $\mathfrak{m}_n(i)$, $\mathfrak{m}(\xi_k)$, etc are as in (2.26) and Remark 2.1.

Theorem 6.1. *If* (5.1) *holds and there is a constant b such that*

$$\lim_{N \to \pm\infty} \frac{1}{|N|} \sum_{n=0}^{N-1} \rho_n \mathfrak{q}_n = b \tag{6.1}$$

then $W_N(t)$ converges in law as $N \to \infty$ to $\mathcal{W}(t)$-the Brownian Motion with zero mean and variance Dt, where $D = \frac{b}{a}$ with b as in (6.1) *and a as in* (5.1).

Proof. In view of (4.8) it suffices to show that

$$\hat{W}_N(t) \Rightarrow \mathcal{W} \tag{6.2}$$

where

$$\hat{W}_N(t) = \frac{\mathfrak{m}(\xi_{Nt})}{\sqrt{N}}. \tag{6.3}$$

Let $\mathfrak{Q}_N = \sum_{n=0}^{N-1} \mathfrak{q}(\xi_n)$, where $\mathfrak{q}(\xi_n) = \mathfrak{q}_{X_n}(Y_n)$. By [12, Theorem 3] to prove (6.2) it suffices to check that

$$\frac{\mathfrak{Q}_N}{N} \Rightarrow D, \tag{6.4}$$

but this follows from (6.1) and Lemma 5.1. □

Corollary 6.2. *For uniquely ergodic environments with bounded potential the Central Limit Theorem holds for* **all** *ω.*

In [20], the Central Limit Theorem was proved for almost all ω for a wide class of environments which includes the uniquely ergodic ones as a particular case. Here, for uniquely ergodic environments, we prove that this result holds for all (rather than almost all) ω.

6.2. **Expectation of the local time.** Here we discuss the distribution of the local time of the walk. Let $V((k, y), N)$ be the number of visits to the site (k, y) by our walk before time N.

Lemma 6.3. *Under the assumptions of Section 4 for each $K > 0$ the collection of random variables*

$$\left\{\frac{V((k, y), N)|\xi(0) = z)}{\sqrt{N}}\right\}$$

is uniformly integrable where the uniformity is with respect to $N \in \mathbb{N}$, $z \in \mathbb{S}$, and $(k, y) \in \mathbb{S}$ such that $|k| \le K\sqrt{N}$.

In the proof we will use the following notion. Let $\mathcal{X}$ and $\mathcal{Y}$ be non-negative random variables. We say that $\mathcal{Y}$ *stochastically dominates* $\mathcal{X}$ if for each $t > 0$ $\mathbb{P}(\mathcal{X} \geq t) \leq \mathbb{P}(\mathcal{Y} \geq t)$. Clearly if $\mathcal{Y}$ stochastically dominates $\mathcal{X}$ then $\mathbb{E}(\mathcal{Y}) \geq \mathbb{E}(\mathcal{X})$.

Proof of Lemma 6.3. It suffices to prove the result for the walk starting from $z = (k, y)$ since the local time does not accumulate before the first visit to the site (k, y).

By (5.7) and the maximal inequality for martingales, there is a constant $\hat{p} < 1$ such that for each K there exists $N_0(K)$ such that if $N \geq N_0(K)$ then for any $(k', y') \in \mathbb{S}$ with $|k'| \leq (K+1)\sqrt{N}$, the probability that the random walk exits the segment $[k' - \sqrt{N}, k' + \sqrt{N}]$ before time N is less than $\hat{p}$ (In fact, $\hat{p}$ can be any number which is greater than the probability that the Brownian motion with zero mean and with variance Dt exits the interval $[-1, 1]$ before time 1).

Let η be the total number of visits to (k, y) before the walk exits from the segment $(k - \sqrt{N}, k + \sqrt{N})$. By the foregoing discussion, the probability that $V((k, y), N) \leq \eta$ is greater than $1 - \hat{p}$. Therefore, for large N, $V((k, y), N)$ is stochastically dominated by $\eta + \hat{p}V((k, y), N)$. Iterating this estimate we conclude that $V((k, y), N)$ is stochastically dominated by $\mathcal{V} := \sum_{r=1}^{\hat{G}} \eta_r$ where $\hat{G}$ is has geometric distribution with parameter $1 - \hat{p}$ and η_r are i.i.d random variables independent of $\hat{G}$ and having the same distribution as η. Since $\mathbb{E}(\mathcal{V}) = \dfrac{\mathbf{E}(\eta)}{1 - \hat{p}}$ it suffices to show the uniform integrability of $\eta/\sqrt{N}$ (with respect to time and the initial position of the walk). However the fact that $\{\eta/\sqrt{N}\}$ is uniformly integrable follows from (4.17) $\square$

Let $\mathfrak{l}_{x,t}$ denote the local time of the standard Brownian motion.

Theorem 6.4. *Suppose that* (5.1) *and* (6.1) *hold. Let* $(k_N, y_N) \in \mathbb{S}$ *be a sequences such that* $\frac{k_N}{\sqrt{N}} \to x$ *as* $N \to \infty$. *Then, as* $N \to \infty$

$$\frac{V((k_N, y_N), N)}{\rho_{k_N}(y_N)\sqrt{N}} \Rightarrow \mathfrak{l}_{x,1/a}.$$

Combining Theorem 6.4 with Lemma 6.3 we obtain

Corollary 6.5. *Suppose that* (k_N, y_N) *is a sequence of points in* $\mathbb{S}$ *such that* $\frac{k_N}{\sqrt{N}} \to x$. *Then uniformly for* x *in a compact set we have*

$$\lim_{N\to\infty} \mathbb{E}\left(\frac{V((k_N, y_N), N)}{\rho_{k_N}(y_N)\sqrt{N}}\right) = \mathbf{E}(\mathfrak{l}_{x,1/a}). \tag{6.5}$$

Proof of the theorem. Consider first the case $k_N \equiv 0$. We use the same notation as in the proof of Lemma 5.1. In particular we let $L_N = \lfloor \delta\sqrt{N} \rfloor$ for a small constant δ.

Fix $\varepsilon > 0$. We show that if δ is sufficiently small then for large N the following estimates hold:

$$\max_{u\in\mathbb{R}} \left| \mathbb{P}\left(\frac{V((0, y_N), \tau_{\hat{k}})}{\sqrt{N}} \leq u\right) - \mathbf{P}\left(\rho(0, y_N)\, \mathfrak{l}_{0,1/a} \leq u\right) \right| \leq \varepsilon, \tag{6.6}$$

where $\hat{k}$ is defined by (5.8), τ_k is defined by (5.4) and

$$\mathbb{P}\left(\frac{|V((0, y_N), N) - V((0, y_N), \tau_{\hat{k}})|}{\sqrt{N}} > \sqrt{\varepsilon}\right) \leq C\sqrt{\varepsilon}. \tag{6.7}$$

To prove (6.7) we note that by (5.9), if N is large enough, then $\mathbb{P}(|\tau_{\hat{k}} - N| \geq \varepsilon N) \leq \varepsilon$. On the other hand if $|\tau_{\hat{k}} - N| \leq \varepsilon N$ then

$$|V((0,y_N),N) - V((0,y_N),\tau_{\hat{k}})| \leq V((0,y_N),(1+\varepsilon)N) - V((0,y_N),(1-\varepsilon)N).$$

By Lemma 6.3 the expectation of the RHS is less than $\bar{C}\varepsilon\sqrt{N}$ so by the Markov inequality

$$\mathbb{P}(V((0,y_N),(1+\varepsilon)N) - V((0,y_N),(1-\varepsilon)N) \geq \sqrt{\varepsilon}\,\sqrt{N}) \leq \bar{C}\sqrt{\varepsilon}. \tag{6.8}$$

Combining (5.9) and (6.8) we obtain (6.7).

To prove (6.6) let U_j be the number of visits to $(0, y_N)$ during the time interval $[\tau_{j-1}, \tau_j]$. Note that $U_j = 0$ unless $\xi(\tau_{j-1}) \in \mathbb{L}_0$. In case $\xi(\tau_{j-1}) \in \mathbb{L}_0$, (4.17) shows that

$$\mathbb{P}\left(U_j = 0 \,|\, \xi(\tau_{j-1}) \in \mathbb{L}_0\right) \leq \frac{C}{L_N}.$$

On the other hand, the general theory of Markov chains shows that, conditioned on $U_j \neq 0$, U_j has geometric distribution with the mean $G_{-L_N,L_N}((0,y_N);(0,y_N))$ and moreover it is independent of $\xi(\tau_j)$. By Lemma 4.3

$$G_{-L_N,L_N}((0,y_N);(0,y_N)) = L_N\rho(0,y_N)(1+o_{N\to\infty}(1)).$$

Now it is easy to show using, for example, Proposition A.1, that

$$\mathbb{P}\left(\left|V((0,y_N),\tau_{\hat{k}}) - L_N\mathfrak{n}(\hat{k})\rho(0,y_N)|\right) \geq \frac{\varepsilon\sqrt{N}}{3}\right) \to 0 \text{ as } N \to \infty$$

where $\mathfrak{n}(k) = \text{Card}(j < k : \xi(\tau_j) \in \mathbb{L}_0)$.

Since the local time of the simple random walk converges after the diffusive rescaling to a local time of the Brownian Motion ([9]), we can take δ so small that

$$\max_{u\in\mathbb{R}} \left|\mathbf{P}(\delta\tilde{\mathfrak{n}}(\hat{k}) \leq u) - \mathbf{P}(\mathfrak{l}_{0,1/a} \leq u)\right| \leq \frac{\varepsilon}{3}$$

where $\tilde{\mathfrak{n}}(\hat{k})$ is the number of times the simple symmetric random walk returns to 0 before time $\hat{k}$. On the other hand, (4.6) and the Optional Stopping Theorem for martingales show that $\left\{\frac{\xi(\tau_j)}{L_N}\right\}_{j\in\mathbb{N}}$ converges as $N \to \infty$ to the simple random walk on $\mathbb{Z}$. Hence for each δ we have

$$\max_{u\in\mathbb{R}} \left|\mathbf{P}(\delta\tilde{\mathfrak{n}}(\hat{k}) \leq u) - \mathbf{P}(\delta\mathfrak{n}(\hat{k}) \leq u)\right| \leq \frac{\varepsilon}{3}$$

provided that N is large enough. Combining the last three displays we obtain (6.6).

This completes the proof of the Theorem in the case $k_N \equiv 0$. The same argument shows that for each k_N, y_N, y'_N, if the walk starts from (k_N, y'_N) then $\frac{V((k_N,y_N),N)}{\sqrt{N}\rho(k_N,y_N)}$ converges to $\mathfrak{l}_{0,1/a}$. Let $\mathfrak{t}_k$ be the first time the walk reaches layer $\mathbb{L}_k$. Divide $[0,1]$ into intervals I_j of small length h and let t_j be the center of I_j. By Theorem 6.1, the probability that $\frac{\mathfrak{t}_{k_N}}{N} \in I_j$ converges as $N \to \infty$ to $\mathbf{P}(\mathcal{T}_x \in I_j)$ where $\mathcal{T}_x$ is the first time the standard Brownian Motion reaches x. On the other hand conditioned on $\frac{\mathfrak{t}_N}{N} \in I_j$ we have that the distribution of $\frac{V((k_N,y_N),N)}{\sqrt{N}}$ is close to the distribution of $\mathfrak{l}_{0,(1-t_j)/a}$ (the closeness means that the error goes to 0 when $h \to 0$ and $N \to \infty$). Therefore for each s

$$\lim_{N\to\infty} \mathbb{P}\left(\frac{V((k_n,y_N),N)}{\sqrt{N}\rho(k_N,y_N)} \geq s\right) = \int_0^1 f_{\mathcal{T}}(t)\mathbf{P}(\mathfrak{l}_{0,(1-t)/a} \geq s)dt$$

where $f_{\mathcal{T}_x}$ is the density of $\mathcal{T}_x$. The last integral is equal to $\mathbf{P}(\mathfrak{l}_{x,1/a} \geq s)$ completing the proof of Theorem 6.4. □

6.3. **Rate of convergence.** Here we estimate the rate of convergence in Theorem 6.1 assuming that we have a good control of error rates in (4.6), (5.1), and (6.1).

Let $\Phi(x)$ denote the distribution function of a standard normal random variable.

We will use the following two results.

Proposition 6.6. ([26, Theorem 3.7]) *Given constants C_1, C_2, C_3 there is a constant C_4 such that the following holds. Let Z_n be a martingale difference sequence such that for $n \leq N$*

$$|Z_n| \leq C_1 \tag{6.9}$$

and $\mathcal{Q}_N = \sum_{n=1}^N \mathbf{E}(Z_n^2|\mathcal{F}_{n-1})$ satisfies

$$\mathbf{P}\left(|\mathcal{Q}_N - N| \geq C_2\sqrt{N}\ln^2 N\right) \leq \frac{C_3 \ln N}{N^{1/4}} \tag{6.10}$$

Then

$$\sup_x \left|\mathbf{P}\left(\frac{\sum_{n=1}^N Z_n}{\sqrt{N}} \leq x\right) - \Phi(x)\right| \leq \frac{C_4 \ln N}{N^{1/4}}.$$

Proposition 6.7. *Let S, Z be random variables and set*

$$\delta = \sup_x |\mathbf{P}(S \leq x) - \Phi(x)|, \quad \delta^* = \sup_x |\mathbf{P}(S + Z \leq x) - \Phi(x)|.$$

Then:
(a) There exists a constant C (independent of S and Z), such that

$$\delta^* \leq 2\delta + C||\sqrt{\mathbf{E}(Z^2|S)}||_\infty,$$

(b) $\delta^* \leq \delta + \mathbf{P}(Z \neq 0)$.

Proof. Part (a) is proven in [4, Lemma 1]. To prove part (b) it suffices to observe that by the triangle inequality $|\delta - \delta^*| \leq \sup_x |\mathbf{P}(S \leq x) - \mathbf{P}(S + Z \leq x)|$. □

In this section, in order to bound the error rate in the CLT, we assume that there is $\beta_1 < 1$ such that for each $L \geq N^{0.01}$ and each $|k| \leq N$

$$|\mathfrak{m}_{k+L}(1) - \mathfrak{m}_{k-L}(1) - 2L| \leq CL^{1-\beta_1}, \tag{6.11}$$

$$\left|\sum_{j=k-L}^{k+L} \rho_j \mathbf{1} - 2La\right| \leq CL^{1-\beta_1}, \tag{6.12}$$

$$\left|\sum_{j=k-L}^{k+L} \rho_j \mathfrak{q}_j - 2Lb\right| \leq CL^{1-\beta_1}. \tag{6.13}$$

Recall the notation of Section 5. Define τ_j as in (5.4) with $L_N = N^{1/4}$. Note that (6.12), (6.13) implies that

$$\mathbb{E}(\tau_k - \tau_{k-1}|\xi_{\tau_{k-1}}) = aL_N^2\left(1 + \mathcal{O}\left(L_N^{-\beta_1}\right)\right), \tag{6.14}$$

$$\mathbb{E}\left(\sum_{n=\tau_{k-1}}^{\tau_k} \mathfrak{q}(\xi_n)|\xi_{\tau_{k-1}}\right) = bL_N^2\left(1+\mathcal{O}\left(L_N^{-\beta_1}\right)\right) \tag{6.15}$$

provided that $|X(\tau_{k-1})| \le N$.

To establish (6.14) we temporarily denote

$$\xi = \xi_{\tau_{k-1}} = (x, y), \quad c = X(\tau_{k-1}) - L_N, \quad d = X(\tau_{k-1}) + L_N.$$

Then Lemma 4.3 gives

$$\mathbb{E}(\tau_k - \tau_{k-1}|\xi_{\tau_{k-1}}) = \sum_{n=c}^{d} \tilde{G}_{c,d}(\xi; n).$$

In view of (4.10) and (6.11)

$$\sum_{n=c}^{d} \tilde{G}_{c,d}(\xi; n) = \left[\sum_{n=c}^{d} \mathcal{G}(x, n, c, d)\rho_n \mathbf{1}\right] + \mathcal{O}\left(L_N^{2-\beta_1}\right).$$

The main term equals to

$$\sum_{n=c}^{d} \mathcal{G}(x, n, c, d)\rho_n \mathbf{1} = \sum_{n=c}^{d} \mathcal{G}(x, n, c, d)a + \sum_{n=c}^{d} \mathcal{G}(x, n, c, d)(\rho_n \mathbf{1} - a)$$

$$= aL_N^2 + \mathcal{O}(L_N) + \sum_{n=c}^{d} \mathcal{G}(x, n, c, d)(\rho_n \mathbf{1} - a).$$

To estimate the last term denote $\mathcal{I}_n = \sum_{k=x}^{n}(\rho_n \mathbf{1} - a)$. Summation by parts gives

$$\sum_{n=c}^{d} \mathcal{G}(x, n, c, d)(\rho_n \mathbf{1} - a) = -\sum_{n=c}^{d} \mathcal{I}_n \nabla \mathcal{G}(x, n, c, d)$$

where ∇ is the difference operator, $\nabla H = H_n - H_{n-1}$. The first term in the last sum is $\mathcal{O}\left(L_N^{1-\beta_1}\right)$ and the second term is bounded. Whence the last sum is $\mathcal{O}\left(L_N^{2-\beta_1}\right)$ proving (6.14). The proof of (6.15) is similar.

Theorem 6.8. *If* (6.11), (6.12) *and* (6.13) *hold then for each* $\varepsilon > 0$ *there is a constant* $C = C_\varepsilon$ *such that*

$$\sup_x \left|\mathbb{P}\left(\frac{X_N}{\sqrt{DN}} \le x\right) - \Phi(x)\right| \le CN^{-(\upsilon-\varepsilon)} \tag{6.16}$$

where

$$\upsilon = \frac{1}{2}\min\left(\frac{1}{4}, \beta_1\right).$$

Proof. To establish the theorem it suffices to show that

$$\sup_x \left|\mathbb{P}\left(\frac{\mathfrak{m}(\xi_N)}{\sqrt{DN}} \le x\right) - \Phi(x)\right| \le CN^{-\bar{\upsilon}}. \tag{6.17}$$

where

$$\bar{\upsilon} = \frac{1}{2}\min\left(\frac{1}{4} - \varepsilon, \beta_1\right).$$

Indeed suppose that (6.17) holds. Let

$$\tilde{X}_N = \begin{cases} X_N & \text{if } |\mathfrak{m}(\xi_N)| < N^{\frac{1+\hat{\varepsilon}}{2}} \\ \mathfrak{m}(\xi_N) & \text{otherwise} \end{cases}$$

where $\hat{\varepsilon}$ is a sufficiently small number. Then due to (6.11) there is a constant K such that $\left|\tilde{X}_N - \mathfrak{m}(\xi_N)\right| \leq KN^{\frac{(1+\hat{\varepsilon})(1-\beta_1)}{2}}$. Therefore

$$\mathbb{P}\left(\frac{\mathfrak{m}(\xi_N)}{\sqrt{DN}} \leq x - \frac{K}{\sqrt{D}}N^{\frac{(1-\beta_1)(1+\hat{\varepsilon})-1}{2}}\right) \leq \mathbb{P}\left(\frac{\tilde{X}_N}{\sqrt{DN}} \leq x\right) \tag{6.18}$$

$$\leq \mathbb{P}\left(\frac{\mathfrak{m}(\xi_N)}{\sqrt{DN}} \leq x + \frac{K}{\sqrt{D}}N^{\frac{(1-\beta_1)(1+\hat{\varepsilon})-1}{2}}\right).$$

Combining (6.18) with (6.17) we obtain

$$\sup_x \left|\mathbb{P}\left(\frac{\tilde{X}_N}{\sqrt{DN}} \leq x\right) - \Phi(x)\right| \leq CN^{-\bar{v}} + \bar{K}N^{\frac{\hat{\varepsilon}-\beta_1(1+\hat{\varepsilon})}{2}} \leq \bar{C}N^{-(v-\varepsilon)} \tag{6.19}$$

provided that $\hat{\varepsilon}$ is small enough,

On the other hand, by Azuma inequality, there are constants $\tilde{c}_1, \tilde{c}_2$ such that

$$\mathbb{P}(\tilde{X}_N \neq X_N) = P\left(\mathfrak{m}(\xi_N) > N^{\frac{1+\hat{\varepsilon}}{2}}\right) \leq c_1 e^{-c_2 N^{\hat{\varepsilon}}}. \tag{6.20}$$

Combining (6.19) with (6.20) we obtain (6.16).

It remains to obtain (6.17). Let $Z_j = \frac{\mathfrak{m}(\xi(\tau_j)) - \mathfrak{m}(\xi(\tau_{j-1}))}{L_N}$ and

$$\mathcal{Q}_j = \sum_{n=1}^{\tau_j} \mathfrak{q}(\xi(n)), \quad j^* = \min\left(j : \mathcal{Q}_{\tau_j} > DN\right), \quad \tau^* = \tau_{j^*}, \quad \mathcal{Q}^* = \mathcal{Q}_{\tau^*}, \quad \mathfrak{m}^* = \mathfrak{m}(\xi_{\tau^*}).$$

Note that $Z_n = \pm 1 + \mathcal{O}\left(N^{-\beta_1}\right)$ due to (6.11) and

$$\mathbb{P}\left(DN \leq Q^* < DN + L_N^2 \ln^2 L_N\right) \leq \theta^{\ln^2 L_N} \tag{6.21}$$

due to (5.10).

Next, we show that if R_1 is a large constant then for each $j \geq \frac{N}{10aL_N^2}$ we have

$$\mathbf{P}\left(\left|\mathcal{Q}_{\tau_j} - aL_N^2 j\right| > R_1 j L_N^{2-\beta_2}\right) \leq c_1 e^{-c_2 N^{\beta_3}} \tag{6.22}$$

where $\beta_2 = \min(\beta_1, \frac{1}{4} - \varepsilon)$, and c_1, c_2, and β_3 are positive constants. We will prove that

$$\mathbf{P}\left(\mathcal{Q}_{\tau_j} - aL_N^2 j > R_1 j L_N^{2-\beta_2}\right) \leq \bar{c}_1 e^{-c_2 N^{\beta_3}}, \tag{6.23}$$

the estimate of $\mathbf{P}\left(\mathcal{Q}_{\tau_j} - aL_N^2 j < -R_1 j L_N^{2-\beta_2}\right)$ being similar.

To prove (6.23) we apply the results of Appendix A, specifically (A.11) with

$$\Delta_n = \frac{\sum_{k=\tau_{n-1}}^{\tau_n} \mathfrak{q}(\xi(k))}{L_N^2} - \frac{R_1}{2}L_N^{-\beta_2}, \quad \varepsilon = \frac{R_1}{2}L_N^{-\beta_2}$$

and the number of summands equal to j. Observe that (A.11) is applicable, because (A.1) follows from (6.15) since $\beta_2 \leq \beta_1$, (A.2) holds by (5.10) and (A.10) holds because $\beta_2 \leq \frac{1}{4} - \varepsilon$.

(6.22) implies that

$$\mathbb{P}\left(j^* > \frac{2}{a}\sqrt{N}\right) = \mathbb{P}\left(j^* > \frac{2N}{aL_N^2}\right) \leq C_1 e^{-C_2 N^{\tilde{\delta}}}. \tag{6.24}$$

Let $\mathfrak{z} = \sum_{j=1}^{\min(\frac{2}{a}\sqrt{N}, j^*)} Z_j$. By (6.24)

$$\mathbb{P}\left(\mathfrak{z} \neq \frac{\mathfrak{m}}{\sqrt{N}}\right) \leq C_1 e^{-C_2 N^{\tilde{\delta}}}. \tag{6.25}$$

(6.25) and (6.21) allow us to apply Proposition 6.6 to $\mathfrak{z}$ obtaining

$$\sup_x |\mathbb{P}(\mathfrak{z} \leq x) - \Phi(x)| \leq \frac{C \ln N}{N^{1/8}}$$

(note that $N^{1/8}$ appears in the denominator since we apply the proposition with $\frac{2}{a}\sqrt{N}$ instead of N). Using (6.25) once more we get

$$\sup_x \left|\mathbb{P}\left(\frac{\mathfrak{m}^*}{\sqrt{DN}} \leq x\right) - \Phi(x)\right| \leq \frac{C \ln N}{N^{1/8}}. \tag{6.26}$$

Next, similarly to (6.22), one can show that there is a constant C_2 such that for each $j \geq \frac{N}{10bL_N^2}$ we have

$$\mathbf{P}\left(\left|\tau_j - bL_N^2 j\right| > R_2 j L_N^{2-\beta_2}\right) \leq c_3 e^{-c_4 N^{\beta_3}} \tag{6.27}$$

Combining (6.22) with (6.27) we conclude that for sufficiently large R_3

$$\mathbb{P}\left(|\tau^* - N| \geq \frac{R_3 N}{L_N^{\beta_2}}\right) \leq c_5 e^{-c_6 L_N^{\beta_3}}. \tag{6.28}$$

Letting

$$\tilde{\mathfrak{m}} = \begin{cases} \mathfrak{m}(\xi_N) & \text{if } |\tau^* - N| \leq \frac{N}{L_N^{\beta_2}} \\ \mathfrak{m}^* & \text{otherwise} \end{cases}$$

we get that with probability 1

$$\mathbb{E}\left((\tilde{\mathfrak{m}} - \mathfrak{m}^*)^2 | \mathfrak{m}^*\right) \leq \frac{R_4 N}{L_N^{\beta_2}}$$

or, equivalently,

$$\mathbb{E}\left(\left(\frac{\tilde{\mathfrak{m}} - \mathfrak{m}^*}{\sqrt{N}}\right)^2 \Big| \frac{\mathfrak{m}^*}{\sqrt{N}}\right) \leq \frac{R_4}{L_N^{\beta_2}}. \tag{6.29}$$

Therefore combining Proposition 6.7(a) and (6.26) we obtain

$$\sup_x \left|\mathbb{P}\left(\frac{\tilde{\mathfrak{m}}}{\sqrt{DN}} \leq x\right) - \Phi(x)\right| \leq C L_N^{-\bar{\upsilon}}$$

(note that $\bar{\upsilon} = \frac{\beta_2}{2} < \frac{1}{8}$, so the main contribution to the error comes from (6.29) rather than from (6.26)).

Next, (6.28) shows that

$$\mathbb{P}(\tilde{\mathfrak{m}} \neq \mathfrak{m}(\xi_N)) \leq c_5 e^{-c_6 L_N^{\beta_3}}.$$

(6.17) follows from the last two displays and Proposition 6.7(b). □

6.4. **Examples.** Here we show that the examples of Section 3.1 satisfy (6.11), (6.12), and (6.13). It is convenient to denote $\Delta_n = \mathfrak{m}_n - \mathfrak{m}_{n-1}$.

We begin with quasiperiodic systems from Example 3.1.

Proposition 6.9. *For quasiperiodic environments of Example 3.1 if γ is Diophantine then* (6.11), (6.12), *and* (6.13) *hold.*

Proof. It is proven in [20] that for quasiperiodic environments with Diophantine frequency γ

$$\Delta_n = \boldsymbol{\Delta}(\omega + n\gamma), \quad \rho_n = \boldsymbol{\rho}(\omega + n\gamma)$$

where $\boldsymbol{\Delta}, \boldsymbol{\rho} : \mathbb{T}^d \to \mathbb{R}$ are continuous functions. In Appendix C of the present paper we obtain a stronger result.

Lemma 6.10. $\boldsymbol{\Delta}, \boldsymbol{\rho}$ *are C^∞.*

Lemma 6.10 implies (6.11), (6.12), and (6.13) with $\beta_1 = 1$. For example to check (6.11) we use the fact that for Diophantine γ there is a constant c and a function u such that

$$\boldsymbol{\Delta}(\omega)\mathbf{1} = c + u(\omega + \gamma) - u(\omega).$$

It follows that

$$\mathcal{M}_{k+L} - \mathcal{M}_{k-L} = 2Lc + u(\omega + L\gamma) - u(\omega - L\gamma).$$

Now (4.6) implies that $c = m$ proving (6.11). Estimates (6.12) and (6.13) are verified similarly. □

Since quasiperiodic environments satisfy (6.11), (6.12), and (6.13) with $\beta_1 = 1$, Theorem 6.8 holds for those environments with $\upsilon = \frac{1}{8}$.

Next, we consider independent environments from Example 3.2.

Proposition 6.11. (6.11) (6.12), *and* (6.13) *hold for independent environments.*

Proof. Let $\mathbb{F}_{a,b}$ be the σ algebra generated by $\{(P, Q, R)_n\}_{a \le n \le b}$. We use the following fact from Appendix C.

Lemma 6.12. *$\rho_n = \boldsymbol{\rho}(T^n\omega)$ and $\Delta_n = \boldsymbol{\Delta}(T^n\omega)$ where $\boldsymbol{\rho} : \Omega \to \mathbb{R}^m$ is Holder continuous with respect to the metric $\mathbf{d}$ defined by* (2.5).

By Lemma 6.12, there is $\theta < 1$ such that for each l there is $\mathbb{F}_{-l,l}$ measurable random vector $\boldsymbol{\rho}^{(l)}$ such that $|\boldsymbol{\rho}(\omega) - \boldsymbol{\rho}^{(l)}(\omega)| \le \theta^l$. Hence

$$|\mathrm{E}(\rho_n\mathbf{1}|\mathcal{F}_{-\infty,n-l}) - \mathrm{E}(\boldsymbol{\rho}\mathbf{1})| \le C\theta^l.$$

Now [24] tells us that for almost every ω there exists $N_0 = N_0(\omega)$ such that for all $|k| < N$ for all $L > N^{0.01}$ we have

$$\left| \sum_{n=k-L}^{k+L} \rho_n\mathbf{1} - 2La \right| \le \sqrt{L \ln^3 L}. \tag{6.30}$$

This proves (6.12). Estimates (6.11) and (6.12) can be established similarly. □

The foregoing discussion shows that (6.11) (6.12), and (6.13) hold with $\beta_1 = \frac{1}{2} - \varepsilon$ (cf. (6.30)). Accordingly, Theorem 6.8 holds with $\upsilon = \frac{1}{8}$.

Finally we consider small perturbations of the simple random walk on $\mathbb{Z}$ from Example 3.3.

Then the invariant measure equation (2.27) reduces to a zero flux condition (see e.g. [22, §5.5])

$$p_n \rho_n = q_{n+1} \rho_{n+1}$$

which gives

$$\frac{\rho_{n+1}}{\rho_n} = \frac{1 - 2a_n}{1 + 2a_{n+1}}.$$

Considering first the case $n > 0$ we obtain

$$\rho_n = \rho_0 \left[\prod_{j=0}^{n-1} \left(\frac{1 - 2a_j}{1 + 2a_{j+1}} \right) \right].$$

Therefore the limit $\rho_+ = \lim_{n\to\infty} \rho_n$ exists and

$$\rho_n = \rho_+ + O\left(\frac{1}{n^{\kappa - 1}} \right). \tag{6.31}$$

Likewise the limit $\rho_- = \lim_{n\to\infty} \rho_{-n}$ exists and

$$\rho_{-n} = \rho_- + O\left(\frac{1}{n^{\kappa - 1}} \right). \tag{6.32}$$

Next, recall a formula for $\mathfrak{m}_n$ ([22]). Let $\Delta_n = \mathfrak{m}_{n+1} - \mathfrak{m}_n$. Then

$$\Delta_{n+1} = \Delta_n \frac{1 + 2a_n}{1 - 2a_n}$$

Thus the limit $\Delta_+ = \lim_{n\to\infty} \Delta_n$ exists and

$$\Delta_n = \Delta_0 \left[\prod_{j=0}^{n-1} \left(\frac{1 + 2a_j}{1 - 2a_j} \right) \right] = \Delta_+ + O\left(\frac{1}{n^{\kappa - 1}} \right). \tag{6.33}$$

Likewise the limit $\Delta_- = \lim_{n\to\infty} \Delta_{-n}$ exists and

$$\Delta_{-n} = \Delta_0 \left[\prod_{j=0}^{n-1} \left(\frac{1 - 2a_{-j}}{1 + 2a_{-j}} \right) \right] = \Delta_- + O\left(\frac{1}{n^{\kappa - 1}} \right). \tag{6.34}$$

Accordingly (3.3) is equivalent to the condition $\Delta_+ = \Delta_-$. Hence if (3.3) holds we can normalize $\{\Delta_n\}$ in such a way that $\lim_{n\to\pm\infty} \Delta_n = 1$. In this case (4.6) holds and (4.7) gives $\lim_{n\to\pm\infty} \rho_n = 1$.

Now (6.33), (6.34), (6.31), and (6.32) show that

$$\Delta_n = 1 + O\left(\frac{1}{|n|^{\kappa - 1}} \right) \quad \text{and} \quad \rho_n = 1 + O\left(\frac{1}{|n|^{\kappa - 1}} \right). \tag{6.35}$$

It follows that (6.11), (6.12) and (6.13) hold with $\beta_1 = \min(\kappa - 1, 1)$. Hence Theorem 6.8 holds in Example 3.3 with $\upsilon = \min\left(\frac{\kappa - 1}{2}, \frac{1}{8} \right)$.

7. Different growth rates.

7.1. **Notation.** In this section we consider the case where $\mathfrak{m}$, ρ, and $\mathfrak{q}$ have different growth rates at $-\infty$ and $+\infty$. Thus we assume that instead of (4.6), (5.1) and (6.1) we have

$$\lim_{n\to-\infty} \frac{\mathcal{M}_n}{n} = m\mu_-, \quad \lim_{n\to+\infty} \frac{\mathcal{M}_n}{n} = m\mu_+; \tag{7.1}$$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=-N+1}^{0} \rho_n \mathbf{1} = a_-, \quad \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \rho_n \mathbf{1} = a_+; \tag{7.2}$$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=-N+1}^{0} \rho_n \mathfrak{q}_n = b_-, \quad \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \rho_n \mathfrak{q}_n = b_+. \tag{7.3}$$

We denote $D_\pm = \frac{b_\pm}{a_\pm}$.

Given μ_1, μ_2 let

$$\mathcal{S}_{\mu_1,\mu_2}(w) = \begin{cases} \frac{w}{\mu_1} & \text{if } w \geq 0 \\ \frac{w}{\mu_2} & \text{if } w \leq 0. \end{cases}$$

Given θ, γ and D we consider the following Markov process. Let $\mathcal{W}_D(u)$ be the Brownian motion with zero mean and variance Du. Denote by $u^+(u)$ the total time on $[0, u]$ when $\mathcal{W}$ is positive and $u^-(u)$ the total time on $[0, u]$ when $\mathcal{W}$ is negative. Given t let $u_\gamma(t)$ be the solution of

$$u^+(u_\gamma) + \frac{u^-(u_\gamma)}{\gamma} = t.$$

Set

$$\mathcal{W}_{\gamma,\theta,D}(t) = \mathcal{S}_{\theta,1}(\mathcal{W}_D(u_\gamma(t))).$$

Note that this process is defined using the function $\mathcal{S}$ with parameters θ and 1. Allowing more general parameters does not increase the generality since $\mu_2 = 1$ can always be achieved by rescaling because $\mathcal{S}_{\mu_1,\mu_2}(\mathcal{W}_D(u_\gamma(\cdot)))$ has the same law as $\mathcal{W}_{\gamma,\frac{\mu_2}{\mu_1},\frac{D}{\mu_1^2}}$.

In the case where

$$\gamma = \left(\frac{1-\mathfrak{p}}{\mathfrak{p}}\right)^2, \quad \theta = \frac{1-\mathfrak{p}}{\mathfrak{p}} \tag{7.4}$$

the process $\mathcal{W}_{\gamma,\theta,D}$ is referred to as the *skew Brownian Motion with parameter* $\mathfrak{p}$. We will thus abbreviate $\mathcal{W}_{\left(\frac{1-\mathfrak{p}}{\mathfrak{p}}\right)^2,\frac{1-\mathfrak{p}}{\mathfrak{p}},D}$ as $\mathcal{B}_{\mathfrak{p},D}$. Note that $\mathfrak{p}$ in (7.4) is given by

$$\mathfrak{p} = \frac{1}{\theta+1}. \tag{7.5}$$

We refer the reader to [36] for description of various equivalent definitions of the skew Brownian Motion as well as its numerous applications. Of these definitions, the most relevant for us is the following one [27]: $\mathcal{B}_{\mathfrak{p},1}(t)$ is the scaling limit of $\frac{\mathfrak{X}(tN)}{\sqrt{N}}$ where $\mathfrak{X}$ is the random walk on $\mathbb{Z}$ which moves to the left and to the right with probability 1/2 everywhere except the origin; at the origin $\mathfrak{X}$ moves to the right with probability $\mathfrak{p}$ and to the left with probability $1-\mathfrak{p}$.

7.2. **Functional CLT.**

Theorem 7.1. *$W_N(t)=\frac{X_{Nt}}{\sqrt{N}}$ converges in law as $N\to\infty$ to $\mathcal{W}_{\gamma,\theta,D}$ where*

$$\gamma=\frac{D_-}{D_+},\quad \theta=\frac{\mu_-}{\mu_+},\quad D=\frac{D^+}{\mu_+^2}.$$

Proof. The proof of Theorem 7.1 is very similar to the proof of Theorem 6.1 so we just sketch the argument. As in Theorem 6.1 it suffices to show that $\hat{W}_N$ defined by (6.3) converges to $\mathcal{W}_{D^+}(u_\gamma(t))$. We may assume without loss of generality that $D^-<D^+$ and so $\gamma<1$. If this is not the case we could consider the reflected walk $(-X(N),Y(N))$. Let $\tilde{\xi}(u)=(\tilde{X}(u),\tilde{Y}(u))$ be the following lazy walk. If $\tilde{X}\geq 0$ then its transition probability coincides with $\mathfrak{P}$. If $\tilde{X}<0$ then, with probability $1-\gamma$, $\tilde{\xi}$ stays at its present location and with probability γ it moves according to $\mathfrak{P}$. There is a natural coupling between ξ and $\tilde{\xi}$ such that $\tilde{\xi}(u)=\xi(t(u))$. Let $u(t)$ be the left inverse to $t(u)$. It is clear from the law of large numbers for sums of geometric random variables that with probability 1

$$\lim_{t\to\infty}\frac{t}{u^+(u(t))+u^-(u(t))/\gamma}=1$$

where $u^+(u)$ and $u^-(u)$ are occupation times of positive and negative semi-axis. It therefore suffices to show that

$$\tilde{W}_N\Rightarrow\mathcal{W}_{D^+} \tag{7.6}$$

where $\tilde{W}_N(u)=\frac{\mathfrak{m}(\tilde{\xi}(Nu))}{\sqrt{N}}$ and $\mathcal{W}_{D^+}$ is the Brownian Motion with zero drift and variance D^+u. Note that $\mathfrak{m}(\tilde{\xi}(u))$ is a martingale with quadratic variation $\sum_{n=1}^{u}\widetilde{\mathfrak{q}}(\tilde{\xi}(n))$ where

$$\widetilde{\mathfrak{q}}(x,y)=\begin{cases}\mathfrak{q}(x,y) & \text{if } x\geq 0\\ \gamma\mathfrak{q}(x,y) & \text{if } x<0.\end{cases}$$

According to [12] it suffices to show that

$$\lim_{N\to\infty}\frac{\sum_{j=1}^{N}\widetilde{\mathfrak{q}}(\tilde{\xi}(j))}{N}=D^+. \tag{7.7}$$

The proof of (7.7) is the same as the proof of Lemma 5.1. The key step is to show that if L_N is the same as in the lemma, $|x|\leq K\sqrt{N}$ and $\tilde{\tau}$ is the exit time from $[x-L_N,x+L_N]$ by $\tilde{\xi}$ then we have that for each $y\in\{1\dots m\}$

$$\tilde{D}_N(x,y)\approx D^+\quad\text{where}\quad \tilde{D}_N(x,y)=\frac{\mathbb{E}(\sum_{u=0}^{\tilde{\tau}}\widetilde{\mathfrak{q}}(\tilde{\xi}(u))|\tilde{\xi}(0)=(x,y))}{\mathbb{E}(\tilde{\tau}|\tilde{\xi}(0)=(x,y))}. \tag{7.8}$$

To fix the ideas, suppose that $[x-L_N,x+L_N]\subset(-\infty,0)$ then

$$\tilde{D}_N(x,y)=\gamma\frac{\mathbb{E}(\sum_{u=0}^{\tilde{\tau}}\mathfrak{q}(\tilde{\xi}(u))|\tilde{\xi}(0)=(x,y))}{\mathbb{E}(\tilde{\tau}|\tilde{\xi}(0)=(x,y))}. \tag{7.9}$$

Note that $\tilde{\xi}_N$ satisfies (2.26), (2.27) with

$$\tilde{\mathfrak{m}}(x,y)=\mathfrak{m}(x,y),\quad \tilde{\rho}(x,y)=\begin{cases}\rho(x,y) & \text{if } x\geq 0\\ \frac{\rho(x,y)}{\gamma} & \text{if } x<0.\end{cases}$$

The computations in Section 5, in particular, (5.6) and (5.14), applied to $\tilde{\xi}$, show that the second factor in the RHS of (7.9) is asymptotic to D_- so that

$$\tilde{D}_N(x,y) \approx \gamma D^- = D^+$$

as claimed. Once (7.8) is established the proof of (7.7) proceeds as in Section 5. □

7.3. **Small perturbations of the environment.** Consider an environment on $\mathbb{S}$ satisfying conditions (4.6), (5.1), and (6.1). Let $\mathfrak{P}$ be defined by (2.2). Consider a perturbation $\tilde{\mathfrak{P}}$ of $\mathfrak{P}$ such that

$$\left|\bar{\mathfrak{P}}(z,z') - \mathfrak{P}(z,z')\right| \leq \frac{C}{|n|^\kappa + 1} \text{ where } z = (n,j) \text{ and } \kappa > 1.$$

Let $\beta_n, \tilde{\beta}_n$ be sequences such that

$$\lambda_n = \frac{\beta_{n+1}}{\beta_n}, \quad \tilde{\lambda}_n = \frac{\tilde{\beta}_{n+1}}{\tilde{\beta}_n}. \tag{7.10}$$

The following result is proven in Appendix C.

Lemma 7.2. *(a) The following estimates hold*

$$||\zeta_n - \bar{\zeta}_n|| \leq \frac{C}{|n|^\kappa + 1}, \quad ||l_n - \bar{l}_n|| \leq \frac{C}{|n|^\kappa + 1}, \quad ||v_n - \bar{v}_n|| \leq \frac{C}{|n|^\kappa + 1}, \tag{7.11}$$

$$||A_n - \bar{A}_n|| \leq \frac{C}{|n|^\kappa + 1}, \quad ||\lambda_n - \bar{\lambda}_n|| \leq \frac{C}{|n|^\kappa + 1}, \quad ||\tilde{\lambda}_n - \bar{\tilde{\lambda}}_n|| \leq \frac{C}{|n|^\kappa + 1}. \tag{7.12}$$

(b) The following limits exist

$$\beta_\pm = \lim_{n\to\pm\infty} \frac{\bar{\beta}_n}{\beta_n} = \lim_{n\to\pm\infty} \frac{\bar{\tilde{\beta}}_n}{\tilde{\beta}_n}. \tag{7.13}$$

(c) The perturbed walk satisfies (7.1), (7.2) *and* (7.3) *with*

$$\mu_\pm = \beta_\pm, \quad a_\pm = a/\beta_\pm, \quad b_\pm = b\beta_\pm$$

where a and b are the limits of (5.1) *and* (6.1) *respectively for the unperturbed walk.*

For random walks on $\mathbb{Z}$ the lemma follows easily from the explicit expressions for the objects involved. Namely

$$A_n = \frac{q_n}{p_n} = \lambda_n = \frac{\beta_{n+1}}{\beta_n}, \quad \text{and} \quad \Delta_n = \beta_n, \quad \rho_n = \frac{1}{\beta_n q_n}, \quad \mathfrak{q}_n = \beta_{n+1}\beta_n \tag{7.14}$$

(see [20, Section 5]). The case of the strip is more complicated and will be considered in Appendix C.

Combining Theorem 7.1 with Lemma 7.2 we obtain the following result.

Corollary 7.3. *Let $\bar{\xi}(t) = (\bar{X}(t), \bar{Y}(t))$ denote the walk in the perturbed environment $\overline{\mathfrak{P}}$.*

$$\frac{\bar{X}(tN)}{\sqrt{N}} \Rightarrow \mathcal{B}_{\mathfrak{p},\, D}$$

where D is the limiting variance of the walk in the unperturbed environment and

$$\mathfrak{p} = \frac{\beta_+}{\beta_+ + \beta_-}. \tag{7.15}$$

Remark 7.4. Note that (7.10) does not define β_n and $\tilde{\beta}_n$ uniquely. Namely, if we replace β_n by $c\beta_n$ and $\tilde{\beta}_n$ by $\tilde{c}\beta_n$ for any constants $c, \tilde{c}$ then (7.10) remains valid. In this case $\beta_\pm$ get replaced by $\frac{\tilde{c}}{c}\beta_\pm$ but expression of $\mathfrak{p}$ does not depend on the arbitrariness involved in the choice of c and $\tilde{c}$.

For random walks on $\mathbb{Z}$ using the explicit expression for λ_n in terms of p_n and q_n (see (7.14)) we obtain

$$\mathfrak{p} = \frac{\boldsymbol{\upsilon}}{\boldsymbol{\upsilon}+1} \quad \text{where} \quad \boldsymbol{\upsilon} = \prod_{n=-\infty}^{\infty} \left(\frac{\tilde{q}_n p_n}{\tilde{p}_n q_n}\right). \tag{7.16}$$

8. Semilocal Limit Theorem

We say that X_N satisfies the semilocal limit theorem at the scale L_N with $1 \ll L_N \ll \sqrt{N}$ if there exists a constant $\beta > 0$ such that for each interval I of length L_N, for each (x, y) with $|x| \leq N$

$$\mathbb{P}(X_N - x \in I | \xi(0) = (x, y)) = \mathbb{P}\left(\sqrt{DN}\mathcal{N} \in I\right) + \mathcal{O}\left(\frac{L_N^{1-\beta}}{\sqrt{N}}\right), \tag{8.1}$$

where $\mathcal{N}$ is the standard normal random variable and D is a positive number (in our case D comes from Theorem 6.1).

Clearly if for each (x, y) with $|x| \leq N$ we have

$$\sup_z \left|\mathbb{P}\left(\frac{X_N - x}{\sqrt{DN}} \leq z \middle| \xi(0) = (x, y)\right) - \Phi(z)\right| \leq N^{-\upsilon}$$

then X satisfies the semilocal limit theorem at the scale N^γ for each $\gamma > \frac{1}{2} - \upsilon$. The next lemma allows us to decrease the scale in the semilocal limit theorem.

Lemma 8.1. *Let $\varepsilon, \varepsilon_1 < \varepsilon_2$ be small positive constants. If N is sufficiently large and for each $\tilde{N}$ such that $N^\varepsilon \leq \tilde{N} \leq N$, for each (x, y) such that $|x| \leq N(1 + \varepsilon_2)$, for each interval I of length $L = \tilde{N}^\gamma$ where*

$$\gamma < \frac{1}{2} \quad \textit{and} \quad \gamma\left(\frac{1}{2} + \gamma\right) > \varepsilon$$

we have

$$\mathbb{P}(X_{\tilde{N}} - x \in I | \xi(0) = (x, y)) = \mathbf{P}\left(\sqrt{D\tilde{N}}\mathcal{N} \in I\right) + \mathcal{O}\left(\frac{L^{1-\beta}}{\sqrt{\tilde{N}}}\right) \tag{8.2}$$

then (8.1) *holds for all (x, y) with $|x| < (1 + \varepsilon_1)N$ and $L_N = N^{(\frac{1}{2}+\gamma)\gamma}$.*

Applying this lemma several times we obtain the following

Corollary 8.2. *Suppose that there exits $\gamma < \frac{1}{2}$ such that for each ε there are constants $\varepsilon_1, \varepsilon_2, N_0$ such that the conditions of Lemma 8.1 are satisfied for $N \geq N_0$. Then, for arbitrarily small $\tilde{\gamma} > 0$, X satisfies the semilocal limit theorem at scale $N^{\tilde{\gamma}}$.*

Proof of Lemma 8.1. Throughout this proof we fix (x, y) and let $\hat{\mathbb{P}}$ denote the distribution of ξ under the condition that $\xi(0) = (x, y)$.

Let $s = \frac{1}{2} + \gamma$. Note that $s < 1$. Consider an interval I of length $N^{\gamma s}$. Let $N_1 = N - N^s$, $N_2 = N^s$. Divide $\mathbb{Z}$ into intervals I_p of size N^γ. Let $\bar{x}$ be the center of I and x_p be the centers of I_p. Call p *feasible* if

$$|x_p - x| \leq N^{1/2+\varepsilon} \quad \text{and} \quad |\bar{x} - x_p| \leq N_2^{1/2+\varepsilon}.$$

By the Azuma inequality, if p is not feasible, then

$$\hat{\mathbb{P}}(X_{N_1} \in I_p, \quad X_N \in I) \leq \exp(-N_2^{2\varepsilon}).$$

Accordingly

$$\hat{\mathbb{P}}(X_N \in I) = \sum_{p-feasible} \hat{\mathbb{P}}(X_{N_1} \in I_p)\, \hat{\mathbb{P}}(X_N \in I | X_{N_1} \in I_p) + \mathcal{O}\left(e^{-N_2^{2\varepsilon}}\right). \tag{8.3}$$

By (8.2) each individual term in this sum is

$$\frac{N^\gamma}{\sqrt{2\pi D N_1}} e^{-(x_p-x)^2/(2DN_1)} \times \frac{N^{\gamma s}}{\sqrt{2\pi D N_2}} e^{-(\bar{x}-x_p)^2/(2DN_2)} + \mathcal{O}\left(N^{(\gamma-\frac{1}{2})(1+s)-\beta s}\right).$$

Since p is feasible we can replace

$$e^{-(x_p-x)^2/(2DN_1)} \quad \text{by} \quad e^{-(\bar{x}-x)^2/(2DN_1)}$$

with an error of order $\mathcal{O}(N^{-\varepsilon})$. Accordingly the main contribution to (8.3) comes from

$$\frac{N^{\gamma s}}{\sqrt{2\pi D N_1}} e^{-(x-\bar{x})^2/(2DN_1)} \sum_p \frac{N^\gamma}{\sqrt{2\pi D N_2}} e^{-(\bar{x}-x_p)^2/(2DN_2)}$$

$$= \frac{N^{\gamma s}}{\sqrt{2\pi D N_1}} e^{-(\bar{x}-x)^2/(2DN_1)} \left[\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi D}} e^{-(\bar{x}-z)^2/(2D)}\, dz + \mathcal{O}\left(\frac{N^\gamma}{\sqrt{N_2}}\right)\right]$$

$$= \frac{N^{\gamma s}}{\sqrt{2\pi D N_1}} e^{-(\bar{x}-x)^2/(2DN_1)} \left[1 + \mathcal{O}\left(N^{\frac{\gamma-1/2}{2}}\right)\right]$$

where the first equality is obtained by replacing the Riemann sum with step $\hbar = \frac{N^\gamma}{\sqrt{N_2}}$ with the Riemann integral with mistake $\mathcal{O}(\hbar)$. The result follows. □

9. Environment viewed by the particle: Mixing.

9.1. **General result.** Here we provide sufficient conditions for mixing of the environment viewed by the particle process. Namely we assume that there is a sequence δ_N converging to 0 as $N \to \infty$, such that for each ε, K there exists N_0 such that for $N \geq N_0$ for each k with $|k| \leq K\sqrt{N}$

$$\left| \frac{1}{2\delta_N N^{1/4}} \sum_{j=k-\delta_N N^{1/4}}^{k+\delta_N N^{1/4}} \rho_j \mathbf{1} - a \right| \leq \varepsilon. \tag{9.1}$$

We consider functions $h : \mathbb{S} \to \mathbb{R}$ satisfying (3.7).

Theorem 9.1. *If* (3.7), (3.4) *and* (9.1) *hold then* $\mathbb{E}(h(\xi(N))) \to \frac{\mathfrak{h}}{a}$ *as* $N \to \infty$.

Proof. If (3.7), and (9.1) hold then the argument of Section 5 shows that for each ε, δ, K there exists N_1 such that for $N \geq N_1$ and for each $(k, y) \in \mathbb{S}$ such that $|k| \leq K\sqrt{N}$ we have

$$\mathbb{P}\left(\left| \frac{1}{\delta\sqrt{N}} \sum_{j=0}^{\delta\sqrt{N}-1} h(\xi(j)) - \frac{\mathfrak{h}}{a} \right| > \varepsilon \;\middle|\; \xi(0) = (k, y)\right) < \frac{\varepsilon}{||h||_\infty}. \tag{9.2}$$

That is, the conclusion of Lemma 5.1 holds uniformly for initial conditions (k, y) satisfying $|k| < K\sqrt{N}$.

Given a trajectory ξ we denote by $\tilde{\xi}$ the accelerated trajectory which skips all steps where ξ stays at the same place. That is, $\tilde{\xi}(n) = \xi(t(n))$ where $t(0) = 0$ and for $n > 0$, $t(n) = \min(t > t(n-1) : \xi(t) \neq \xi(t(n-1)))$. We denote by $s(n) = t(n+1) - t(n)$ the time the walker spends at $\tilde{\xi}(n)$.

A *path* is a finite set of points $W = \{z_0, z_1 \dots z_l\}$ such that $z_j \neq z_{j+1}$ for $j = 0, \dots, l-1$. The number $l = l(W)$ is called the *length of the path*. A path is called *admissible* if there is an accelerated trajectory $\tilde{\xi}$ such that $\tilde{\xi}(n) = z_n$ for $0 \leq n \leq l(W)$. Given an admissible path W and a trajectory ξ following this path let $\tau(W, \xi) = t(l(W))$ be the number of steps it takes ξ to traverse this path. Let $T(W)$ be the expectation of $\tau(W, \xi)$ conditioned on the event that W is the beginning part of $\tilde{\xi}$. Observe that

$$\tau(W, \xi) = \sum_{n=0}^{l-1} s(n) \tag{9.3}$$

where, for a fixed W, $s(n)$ are independent random variables having geometric distributions with parameter $1 - \mathfrak{P}(\tilde{\xi}(n), \tilde{\xi}(n))$.

Let $\mathcal{S}(N)$ be the set of (admissible) paths such that $T(W) \geq \frac{N}{2}$ but $T(W^-) < \frac{N}{2}$ where W^- is the path obtained by removing the last edge from W. Given $W \in \mathcal{S}(N)$, δ, j let

$$A_{W,\delta,j} = \left\{ \xi : W = \tilde{\xi}([0, l(W)]) \text{ and } \tau(W, \xi) \in \left[\frac{N}{2} + \delta j \sqrt{N}, \frac{N}{2} + \delta(j+1)\sqrt{N} \right) \right\}.$$

By the Central Limit Theorem for $\tau(W, \xi)$, (see (9.3)), given $\varepsilon > 0$ we can find R such that

$$\mathbb{P}\left(\bigcup_{W \in \mathcal{S}(N)} \left\{ \xi : \left| \tau(W, \xi) - \frac{N}{2} \right| > R\sqrt{N} \right\} \right) \leq \varepsilon.$$

Accordingly

$$\mathbb{E}(h(\xi(N))) = \left[\sum_{W \in \mathcal{S}(N)} \sum_{|j| \leq R} \mathbb{P}(A_{W,\delta,j})\, \mathbb{E}\Big(h(\xi(N)) \big| A_{W,\delta,j} \Big) \right] + \tilde{\varepsilon}$$

where $\tilde{\varepsilon} \leq (\sup |h|)\varepsilon$.

We claim that

$$\left| \sum_{W \in \mathcal{S}(N)} \sum_{|j| \leq R/\delta} \mathbb{P}(A_{W,\delta,j})\, \mathbb{E}\Big(h(\xi(N)) | A_{W,\delta,j} \Big) - \frac{\mathfrak{h}}{a} \right| \leq 3\varepsilon \tag{9.4}$$

provided that δ is small enough. Indeed

$$\mathbb{E}\Big(h(\xi(N)) | A_{W,\delta,j} \Big) = \sum_{l=1}^{\delta\sqrt{N}} \mathbb{E}\left(h\left(\xi\left(\frac{N}{2} - \delta(j+1)\sqrt{N} + l \right) \right) \Big| \, \xi(0) = e(W) \right) \tag{9.5}$$
$$\times \mathbb{P}\left(\tau(W, \xi) = \frac{N}{2} + \delta(j+1)\sqrt{N} - l \,\Big|\, A_{W,\delta,j} \right)$$

where $e(W) = (x(W), y(W))$ is the endpoint of W. By the Local Limit Theorem for the sum (9.3) ([44, 16])

$$\left|\mathbb{P}\left(\tau(W,\xi) = \frac{N}{2} + \delta(j+1)\sqrt{N} - l \,\Big|\, A_{W,\delta,j}\right) - \frac{1}{\delta\sqrt{N}}\right| \leq \frac{\varepsilon}{||h||_\infty} \tag{9.6}$$

uniformly in $l \leq \delta\sqrt{N}$ provided that δ is small enough. This allows us to replace $\mathbb{E}\Big(h(\xi(N))|A_{W,\delta,j}\Big)$ by

$$\frac{1}{\delta\sqrt{N}} \sum_{l=1}^{\delta\sqrt{N}} \mathbb{E}\left(h\left(\xi\left(b_{N,j,l}\right)\right) \,\big|\, \xi(0) = e(W)\right),$$

where

$$b_{N,j,l} = \frac{N}{2} - \delta(j+1)\sqrt{N} + l. \tag{9.7}$$

To control this sum we consider two cases.

(I) The terms where $|x(W)|$ is large can be controlled as follows. By Theorem 6.1

$$\frac{1}{\delta\sqrt{N}} \sum_{W\in\mathcal{S}(N)} \sum_{|j|\leq R/\delta} \sum_{|x(W)|>K\sqrt{N}} \sum_{l=1}^{\delta j\sqrt{N}} \mathbb{P}(A_{W,\delta,j})\mathbb{P}\left(\xi\left(b_{N,j,l}\right) = e(W) \Big|\, \xi(0) = e(W)\right)$$

$$\leq \mathbb{P}\left(|\xi(N)| > \frac{K\sqrt{N}}{2}\right) \leq \frac{\varepsilon}{||h||_\infty}$$

provided that K is sufficiently large and $N \geq N_2(K)$.

(II) On the other hand if $|x(W)| \leq K\sqrt{N}$ then in view of (9.2)

$$\left|\frac{1}{\delta\sqrt{N}} \sum_{l=1}^{\delta\sqrt{N}} \mathbb{E}\left(h\left(\xi\left(b_{N,j,l}\right)\right) \Big|\, \xi\left(0\right) = e(W)\right) - \frac{\mathfrak{h}}{a}\right| \leq \varepsilon$$

provided that N is large enough.

Combining the estimates for the cases (I) and (II) above with (9.6) we obtain (9.4). Since ε is arbitrary Theorem 9.1 follows. □

9.2. **Examples.** Examples presented in Section 3.1 also satisfy (3.7) and (9.1).

In fact, in Example 3.1 we can replace quasiperiodic environments by more general environments generated by uniquely ergodic transformation (we refer the reader to [15, §1.8] (for background on uniquely ergodic transformations). That is, let T be a uniquely ergodic transformation of a compact metric space Ω, $(P, Q, R)_n(\omega) = (\mathcal{P}, \mathcal{Q}, \mathcal{R})(T^n\omega)$ and $h_n(\omega) = \mathcal{H}(T^n\omega)$.

Proposition 9.2. *If* $(\mathcal{P}, \mathcal{Q}, \mathcal{R})$ *and* $\mathcal{H}$ *are continuous then* (3.7) *and* (9.1) *hold.*

Proof. By Section 6 and Appendix A of [20], $\rho_n = \boldsymbol{\rho}(T^n\omega)$, where $\boldsymbol{\rho} : X \to \mathbb{R}^m$ is continuous. Therefore (3.7) and (9.1) follow from the fact that the convergence in ergodic theorem for uniquely ergodic systems is uniform with respect to ω ([15, Theorem 1.8.2]). □

In the case of independent environments we suppose that $h_n = \mathcal{H}(P_n, Q_n, R_n)$ where $\mathcal{H}$ is a bounded continuous function.

Proposition 9.3. (3.7) *and* (9.1) *hold for independent environments.*

Proof. The proof of (9.1) is very similar to the proof of Proposition 6.11 so it can be left to the reader. The proof of (3.7) in case $\mathcal{H}$ is a local function (that is there exists R such that $\mathcal{H}(\omega)$ depends only on $(P_n, Q_n, R_n)_n$ with $|n| \leq R$) is also similar to Proposition 6.11. To prove (3.7) for general continuos function, it suffices to approximate it by a local function with error less than $\varepsilon/2$. □

Propositions 9.2 and 9.3 complete the proof of Theorem 3.10 for Examples 3.1 and 3.2. To prove Theorem 3.10 for Example 3.3 we need to take into account that the walk is not allowed to remain at the same site at two consecutive moments of time. Because of that, we consider ξ at odd and at even times separately and note that (3.8) implies (3.7) for both odd and even sublattices.

10. Local Limit Theorem

Theorem 10.1. *If* (6.11), (6.12), *and* (6.13) *hold then for each sequence* (k_N, y_N) *such that* $k_N/\sqrt{N}$ *is bounded we have*

$$\lim_{N\to\infty} \frac{\mathbb{P}\left(\xi(N) = (k_N, y_N)\right)}{\mathbf{P}\left(\sqrt{\frac{bN}{a}}\mathcal{N} \in \left[k_N - \frac{1}{2}, k_N + \frac{1}{2}\right]\right)\rho(k_N, y_N)} = \frac{1}{a} \tag{10.1}$$

where a and b are the constants from (6.12) *and* (6.13) *respectively.*

Proof. We use the same notation as in Section 9. In particular we choose a small constant ε_2 and let δ be as in the proof of Theorem 9.1. We have

$$\mathbb{P}(\xi(N) = (k_N, y_N)) = \sum_W \sum_j \mathbb{P}(A_{W,\delta,j})\mathbb{P}\Big(\xi(N) = (k_N, y_N)|A_{W,\delta,j}\Big), \tag{10.2}$$

where the sum is over all admissible paths W.

Given $\bar{R}$ denote by $\mathcal{S}_{\bar{R}}(N)$ the set of paths in $\mathcal{S}(N)$ whose endpoint $e(W) = (x(W), y(W))$ satisfies $|x(W)| \leq \bar{R}\sqrt{N}$. Pick $\bar{R} \gg 1$ and divide the sum (10.2) into three parts.

(I) If $W \in \mathcal{S}_{\bar{R}}(N)$ and $|j| < N^{\bar{\varepsilon}}$ then (9.6) allows us to replace

$$\mathbb{P}(\xi(N) = (k_N, y_N)|A_{W,\delta,j})$$

by

$$\frac{1}{\delta\sqrt{N}} \sum_{l=1}^{\delta\sqrt{N}} \mathbb{P}\Big(\xi(b_{N,j,l}) = (k_N, y_N)|\xi(0) = e(W)\Big).$$

where $b_{N,j,0} = \frac{N}{2} - \delta(j+1)\sqrt{N}$, $b_{N,j,l} = b_{N,j,0} + l$ (see (9.7)). Divide $\mathbb{Z}$ into segments I_p of length $L_N = N^{1/5}$. Let $\tilde{k}_p$ be the center of I_p. We split the above sum as

$$\sum_p \sum_{\substack{\tilde{k}\in I_p \\ \tilde{y}\in\{1,\dots,m\}}} \mathbb{P}(\xi(b_{N,j,0}) = (\tilde{k}, \tilde{y})|\xi(0) = e(W)) \;\; \frac{\sum_{l=1}^{\delta\sqrt{N}} \mathbb{P}(\xi(l) = (k_N, y_N)|\xi(0) = (\tilde{k}, \tilde{y}))}{\delta\sqrt{N}}$$

Denote $\bar{N} = \delta^{1/2}N^{1/4}$. By Corollary 6.5 if $|k_N - \tilde{k}_p| \leq \bar{R}\sqrt{\bar{N}}$ and $\tilde{k} \in I_p$ then

$$\frac{\sum_{l=1}^{\delta\sqrt{N}} \mathbb{P}(\xi(l) = (k_N, y_N)|\xi(0) = (\tilde{k}, \tilde{y}))}{\delta\sqrt{N}} \sim \frac{\mathbf{E}(\mathfrak{l}_{(k_N-\tilde{k})/\bar{N}),1/a})}{\bar{N}}\rho_{k_N,y_N}$$

$$\sim \frac{\mathbf{E}(\mathfrak{l}_{(k_N-\tilde{k}_p)/\bar{N}),1/a})}{\bar{N}}\rho_{k_N,y_N} \tag{10.3}$$

where the last step uses that $|k_N - \tilde{k}_p| \ll \sqrt{N}$.

On the other hand Corollary 8.2 and Theorem 6.8 show that

$$\sum_{\substack{\tilde{k}\in I_p \\ \tilde{y}\in\{1,\dots,m\}}} \mathbb{P}(\xi(b_{N,j,0}) = (\tilde{k},\tilde{y})|\xi(0) = e(W)) = \frac{L_N(1+o_{N\to\infty}(1))}{\sqrt{\pi DN}} \exp\left(-\frac{(x(W)-\tilde{k}_p)^2}{DN}\right)$$

$$= \frac{L_N(1+o_{N\to\infty}(1))}{\sqrt{\pi DN}} \exp\left(-\frac{(x(W)-k_N)^2}{DN}\right) \tag{10.4}$$

where $DN = 2D(\frac{N}{2})$ appears in the above expression since $b_{N,j,0} = \frac{N}{2} + \mathcal{O}\left(N^{(1/2)+\varepsilon}\right)$.

Next, if

$$|k_N - \tilde{k}_p| \geq \bar{R}\sqrt{\bar{N}} \tag{10.5}$$

then

$$\frac{\sum_{l=1}^{\delta\sqrt{N}} \mathbb{P}(\xi(l) = (k_N, y_N)|\xi(0) = (\tilde{k},\tilde{y}))}{\delta\sqrt{N}} \leq \tag{10.6}$$

$$\mathbb{P}\left(\xi \text{ visits } (k_N, y_N) \text{ before time } \delta\sqrt{N}|\xi(0) = (\tilde{k},\tilde{y})\right) \times$$
$$\mathbb{E}(\mathrm{Card}(l \leq \sqrt{N} : \xi(l) = (k_N, y_N))|\xi(0) = (k_N, y_N)).$$

The first factor is $\mathcal{O}\left(e^{-c(k_N-\tilde{k})^2/[\bar{N}^2]}\right)$ by the Azuma inequality and the second factor is $\mathcal{O}(\bar{N})$ by Lemma 6.3, so in case (10.5) we have

$$\frac{\sum_{l=1}^{\delta\sqrt{N}} \mathbb{P}(\xi(l) = (k_N, y_N)|\xi(0) = (\tilde{k},\tilde{y}))}{\delta\sqrt{N}} \leq \frac{C}{\bar{N}} \exp\left(-\frac{c(k_N-\tilde{k}_p)^2}{\bar{N}^2}\right). \tag{10.7}$$

Hence (see (10.4))

$$\sum_{\substack{\tilde{k}\in I_p \\ \tilde{y}\in\{1,\dots,m\}}} \mathbb{P}(\xi(b_{N,j,0}) = (\tilde{k},\tilde{y})|\xi(0) = e(W)) \; \frac{\sum_{l=1}^{\delta\sqrt{N}} \mathbb{P}(\xi(l) = (k_N, y_N)|\xi(0) = (\tilde{k},\tilde{y}))}{\delta\sqrt{N}}$$

$$\leq \frac{CL_N}{\bar{N}\sqrt{N}} \exp\left(-\frac{c(k_N-\tilde{k}_p)^2}{\bar{N}^2}\right). \tag{10.8}$$

Next, we perform the summation over p. Equations (10.3), (10.4), (10.8) show that in case (I)

$$\frac{1}{\delta\sqrt{N}} \sum_{l=1}^{\delta\sqrt{N}} \mathbb{P}(\xi(b_{N,j,l}) = (k_N, y_N)|\xi(0) = e(W))$$

$$= \sum_{|k_p-\tilde{k}_N|<\bar{R}\sqrt{\bar{N}}} \frac{L_N(1+o_{N\to\infty,\bar{R}\to\infty}(1))}{\sqrt{\pi DN}} \frac{\mathbf{E}(\mathfrak{l}_{(k_N-\tilde{k}_p)/\bar{N}),1/a})}{\bar{N}} \rho_{k_N,y_N} \exp\left(-\frac{(x(W)-k_N)^2}{DN}\right)$$

$$= \frac{1}{\sqrt{\pi DN}} \frac{\rho_{k_N,y_N}}{a} \exp\left[-\frac{(x(W)-k_N)^2}{DN}\right] \left(1+o_{N\to\infty,\bar{R}\to\infty}(1)\right) \tag{10.9}$$

where the last step relies on the fact that

$$\int_{-\infty}^{\infty} \mathfrak{l}_{x,t}\, dx = t.$$

(II) $W \notin \mathcal{S}_{\bar{R}}(N)$ and $|j| < N^{\bar{\varepsilon}}$. In this case the same argument as in the proof of (10.7) shows that

$$\frac{1}{\delta\sqrt{N}} \sum_{l=1}^{\delta\sqrt{N}} \mathbb{P}\Big(\xi(b_{N,j,l}) = (k_N, y_N)|\xi(0) = e(W)\Big) \leq \frac{\varepsilon(\bar{R})}{\sqrt{N}},$$

where $\varepsilon(\bar{R}) \to 0$ as $\bar{R} \to \infty$.

(III) $|j| \geq N^{\bar{\varepsilon}}$. Due to moderate deviation estimate for sums of independent random variables applied to the sum (9.3).

$$\mathbb{P}\left(\bigcup_W \bigcup_{j \geq N^{\bar{\varepsilon}}} A_{W,\delta,j}\right) \leq Ce^{-cN^{2\bar{\varepsilon}}}.$$

Thus the main contribution to (10.2) comes from case (I). Performing the summation over $W \in \mathcal{S}_{\bar{R}}(N)$ and $j \in [-N^{\bar{\varepsilon}}, N^{\bar{\varepsilon}}]$ and using (10.9) and the CLT for $x(W)$ we obtain (10.1). □

Theorem 10.1 implies Theorem 3.8(a). To prove Theorem 3.8(b) we need to consider $\xi(2N)$ and $\xi(2N+1)$ separately (see the discussion at the end of Section 9) and note that in Example 3.3 $D = 1$ since, due to equation (6.35), ξ is a small perturbation of the simple random walk away from the origin.

Appendix A. A rough bound on large and moderate deviations.

Proposition A.1. *Let $\{\mathcal{F}_n\}$, $n \geq 0$, be a filtration and B_n be a sequence of $\mathcal{F}_n$-measurable random variables such that $B_0 = 0$ and $\Delta_n = B_n - B_{n-1}$ satisfies for $n \leq N$ the following estimates:*

$$\mathbf{E}(\Delta_n|\mathcal{F}_{n-1}) \leq -\varepsilon \quad \textit{where} \quad \varepsilon \geq N^{-1/2} \tag{A.1}$$

and for some positive constants c, K

$$\mathbf{E}(e^{c|\Delta_n|}|\mathcal{F}_{n-1}) \leq K. \tag{A.2}$$

Then there is a constant $\bar{c} = \bar{c}(c, K) > 0$ such that

$$\mathbf{P}(B_N \geq 0) \leq \begin{cases} e^{-\bar{c}\sqrt{\varepsilon N}} & \textit{if } \varepsilon \geq N^{-1/3} \\ Ne^{-\bar{c}\varepsilon^2 N} & \textit{if } \varepsilon < N^{-1/3}. \end{cases}$$

Remark A.2. The first case ($\varepsilon \geq N^{-1/3}$) is sufficient for all the applications given in this paper except that one would get worse constants in Section 6.3.

Proof. Suppose first that $\varepsilon \geq N^{-1/3}$. Let $s = \frac{c_1}{\sqrt{\varepsilon N}}$ for a sufficiently small constant c_1 (see (A.3) and (A.7) below for the precise conditions on c_1.) Set $A = \sqrt{\varepsilon N}$ and define

$$\tilde{\mathbf{\Delta}}_k = \Delta_k 1_{\Delta_k < A}, \quad \tilde{B}_n = \sum_{k=1}^n \tilde{\mathbf{\Delta}}_k, \quad \phi_n = \mathbf{E}\left(e^{s\tilde{B}_n}\right).$$

Then

$$\phi_k = \mathbf{E}\left(e^{s\tilde{B}_{k-1}}\mathbf{E}\left(e^{s\tilde{\mathbf{\Delta}}_k}|\mathcal{F}_{k-1}\right)\right).$$

Note that $s\tilde{\mathbf{\Delta}}_k \leq sA = c_1$ and so we can choose c_1 so small that

$$e^{s\tilde{\mathbf{\Delta}}_k} \leq 1 + s\tilde{\mathbf{\Delta}}_k + (s\tilde{\mathbf{\Delta}}_k)^2, \tag{A.3}$$

and so

$$\mathbf{E}\left(e^{s\tilde{\mathbf{\Delta}}_k}|\mathcal{F}_{k-1}\right) \leq 1 + s\mathbf{E}\left(\tilde{\mathbf{\Delta}}_k|\mathcal{F}_{k-1}\right) + s^2\mathbf{E}\left(\tilde{\mathbf{\Delta}}_k^2|\mathcal{F}_{k-1}\right). \tag{A.4}$$

In view of (A.2)

$$\mathbf{E}(\tilde{\mathbf{\Delta}}^2|\mathcal{F}_{k-1}) \leq \mathbf{E}(\Delta^2|\mathcal{F}_{k-1}) \leq \text{Const} \tag{A.5}$$

and since $A = \sqrt{\varepsilon N} \geq N^{2/3}$ we have that for large N

$$\mathbb{E}(\tilde{\mathbf{\Delta}}_k|\mathcal{F}_{k-1}) \leq -\frac{2\varepsilon}{3}. \tag{A.6}$$

Note that $\dfrac{s^2}{\varepsilon s} = \dfrac{c_1}{\sqrt{\varepsilon^3 N}} \leq c_1$ can be made as small as we wish by choosing c_1 small. Hence (A.4), (A.5) and (A.6) show that we can choose c_1 so small that

$$\mathbf{E}\left(e^{s\tilde{\mathbf{\Delta}}_k}|\mathcal{F}_{k-1}\right) \leq 1 - \frac{\varepsilon s}{2}. \tag{A.7}$$

Accordingly

$$\mathbf{E}\left(e^{s\tilde{B}_N}\right) \leq \left(1 - \frac{\varepsilon s}{2}\right)^N.$$

Thus for large N

$$\mathbf{P}(\tilde{B}_N \geq 0) \leq e^{-s\varepsilon N/4}. \tag{A.8}$$

Next for each n

$$\mathbb{P}(\Delta_n \geq A) \leq \mathbb{P}\left(e^{c\Delta_n} \geq e^{cA}\right) \leq Ke^{-cA}.$$

Hence

$$\mathbf{P}(\tilde{B}_N \neq B_N) \leq N \max_{n\leq N} \mathbb{P}(\Delta_n \geq A) \leq Ne^{-cA} \tag{A.9}$$

where the last inequality follows by (A.2). Combining (A.8) with (A.9) and using that $\varepsilon s N = c_1\sqrt{\varepsilon N}$, $A = \sqrt{\varepsilon N}$ we obtain the required estimate in case $\varepsilon \leq N^{-1/3}$.

Next consider the case where $\varepsilon < N^{1/3}$. We can also assume that $\varepsilon \geq \sqrt{\dfrac{\ln\ln N}{N}}$ since the result is trivial (and useless) if $\varepsilon^2 N < \ln\ln N$ because the RHS is greater than 1. The argument in the case where $\sqrt{\dfrac{\ln\ln N}{N}} \leq \varepsilon < N^{-1/3}$ is the same as in the case where $\varepsilon > N^{-1/3}$ except that the parameters are chosen differently. Namely, we let $s = c_1\varepsilon$ where c_1 is appropriately small and $A = N\varepsilon^2$. With this choice of parameters both $sA = c_1 N\varepsilon^3$ and $\frac{s}{\varepsilon} = c_1$ still can be made as small as needed. Accordingly we still have

$$\mathbf{P}(B_N \geq 0) \leq e^{-s\varepsilon N/4} + Ne^{-cA}$$

giving the required bound. □

Remark A.3. We will often use the following consequence of Proposition A.1: for any δ_1 there are positive constants C_1, C_2 and δ_2 such that if (A.1) and (A.2) hold and

$$\varepsilon > N^{\delta_1 - \frac{1}{2}} \tag{A.10}$$

then

$$\mathbf{P}(B_N \geq 0) \leq C_1 e^{-C_2 N^{\delta_2}}. \tag{A.11}$$

Appendix B. Contraction properties of products of positive matrices.

The proof of relations (2.14), (2.15) follows from very general and well known contracting properties of positive matrices which we now recall.

Let $\mathbb{A}_\delta$ be the set of positive $m \times m$ matrices such that for any $A = (A(i,j)) \in \mathbb{A}_\delta$ one has $\min_{i,j,k} A(i,k)/A(j,k) \geq \delta$, where $\delta > 0$ does not depend on A. Let $\mathbb{R}_+^m$ be the cone of non-negative vectors in $\mathbb{R}^m$ and $\mathbb{R}_{+,\delta}^m$ its sub-cone of positive column vectors with $\min_{i,j} x_i/x_j \geq \delta$. Then $A\mathbb{R}_+^m \subset \mathbb{R}_{+,\delta}^m$ for any $A \in \mathbb{A}_\delta$. Indeed, for any vector $x \geq 0$ ($x \neq 0$) we have

$$\frac{(Ax)_i}{(Ax)_j} = \frac{\sum_{k=1}^m A(i,k)x_k}{\sum_{k=1}^m A(j,k)x_k} \geq \min_k \frac{A(i,k)}{A(j,k)} \geq \delta. \tag{B.1}$$

Next denote by $\mathcal{C}^\delta$ the set of rays generated by vectors from $\mathbb{R}_{+,\delta}^m$. Also, we introduce the convention that $\mathcal{C}^0$ is the set of rays generated by vectors from $\mathbb{R}_+^m$. If $\mathbf{x}, \mathbf{y} \in \mathcal{C}^\delta$ are two rays generated by vectors $x, y \in \mathbb{R}_{+,\delta}^m$ then the Hilbert's projective distance between them is defined by

$$\mathfrak{r}(\mathbf{x}, \mathbf{y}) = \max_{i,j} \ln \frac{x_i y_j}{x_j y_i}.$$

The set $\mathcal{C}^\delta$ equipped with this metric is a compact metric space. The action of a matrix $A \in \mathbb{A}_\delta$ on $\mathcal{C}^0$ is naturally defined by its action on $\mathbb{R}_+^m$ and for $\mathbf{x} \in \mathcal{C}^0$ we write $A\mathbf{x}$ for the image of $\mathbf{x}$ under the action of A. (B.1) shows that in fact $A\mathcal{C}^0 \subset \mathcal{C}^\delta$.

We need the following version of a (stronger) result from [3, Chapter XVI, Theorem 3]: for all $A \in \mathbb{A}_\delta$ and all $\mathbf{x}, \mathbf{y} \in \mathcal{C}^\delta$

$$\mathfrak{r}(A\mathbf{x}, A\mathbf{y}) \leq c\, \mathfrak{r}(\mathbf{x}, \mathbf{y}), \quad \text{where} \quad c = \frac{1-\delta}{1+\delta}. \tag{B.2}$$

We are now in a position to prove (2.14) and (2.15) from Section 2.3. To this end, note first that (2.13) implies that $A_n \in \mathbb{A}_\delta$ with $\delta = m\bar{\varepsilon}^2$.

Next, for $a \leq n$, the sets $\mathcal{C}_a \stackrel{\text{def}}{=} A_n ... A_a \mathcal{C}^0$ form a decreasing sequence, $\mathcal{C}_a \supset \mathcal{C}_{a-1}$, of compact subsets of $\mathcal{C}^\delta$ and therefore $\bigcap_{a \leq n} \mathcal{C}_a \neq \emptyset$. Due to (B.2), for any two rays $\mathbf{x}, \mathbf{y} \in \mathcal{C}^\delta$ the projective distance between their images in $\mathcal{C}_a$ decays exponentially as $a \to -\infty$:

$$\mathfrak{r}(A_n ... A_a \mathbf{x}, A_n ... A_a \mathbf{y}) \leq c^{n-a} \mathfrak{r}(\mathbf{x}, \mathbf{y}). \tag{B.3}$$

(There is no loss of generality in assuming that $\mathbf{x}, \mathbf{y} \in \mathcal{C}^\delta$ since $A_a \mathcal{C}^0 \subset \mathcal{C}^\delta$.)

Therefore there is a unique ray $\mathbf{v}_n = \bigcap_{a \leq n} \mathcal{C}_a$ and v_n in (2.14) is the unit vector corresponding to $\mathbf{v}_n$ which proves (2.14). It remains to note that at a small scale the standard distance between unit vectors (as in (2.15)) is equivalent to the distance between rays generated by these vectors which means that (B.3) is equivalent to (2.15).

Appendix C. Regularity of ρ and Δ.

Here we discuss the regularity of ρ and Δ which plays a key role in our analysis. To this end we recall the formulas for these expressions obtained in [20].

Let Ω be a compact metric space and $T : \Omega \to \Omega$ be a continuous map. (This meaning for the letter T is reserved for Appendix C only.)

Throughout this section we assume that $(P, Q, R)_n(\omega) = (\mathcal{P}, \mathcal{Q}, \mathcal{R})(T^n\omega)$ where $(\mathcal{P}, \mathcal{Q}, \mathcal{R})$ are continuous functions such that (2.1) and (2.4) are satisfied. Define

$$\zeta(\omega) = \zeta_0(\omega), \quad A(\omega) = A_0(\omega), \quad \alpha(\omega) = \alpha_0(\omega), \quad \sigma(\omega) = \sigma_0(\omega)$$
$$v(\omega) = v_0(\omega), \quad l(\omega) = l_0(\omega) \quad \lambda(\omega) = \lambda_0(\omega), \quad \tilde{\lambda}(\omega) = \tilde{\lambda}_0(\omega) \tag{C.1}$$

then

$$\zeta_n = \zeta(T^n\omega), \quad A_n = A(T^n\omega), \quad \alpha_n = \alpha(T^n\omega), \quad \sigma_n(\omega) = \sigma(T^n\omega),$$
$$v_n = v(T^n\omega), \quad l_n = l(T^n\omega), \quad \lambda_n = \lambda(T^n\omega), \quad \tilde{\lambda}_n = \tilde{\lambda}(T^n\omega). \tag{C.2}$$

It is proven in [20] that RWRE in bounded potential enjoy the property that

$$\tilde{\lambda}(\omega) = \frac{\tilde{\beta}(T\omega)}{\tilde{\beta}(\omega)}, \quad \lambda(\omega) = \frac{\beta(T\omega)}{\beta(\omega)} \tag{C.3}$$

for continuous functions $\beta, \tilde{\beta}$. Moreover, the functions $\zeta(\cdot)$, $v(\cdot)$, $l(\cdot)$ are continuous in ω. The continuity of all other functions is implied by the continuity of ζ, v, and l.

It is proven in [20] that

$$\rho_n(\omega) = \boldsymbol{\rho}(T^n\omega) \text{ and } \Delta_n(\omega) = \boldsymbol{\Delta}(T^n\omega)$$

where

$$\boldsymbol{\rho}(\omega) = c\frac{l(\omega)}{\tilde{\beta}(\omega)} \quad \text{and} \quad \boldsymbol{\Delta}(\omega) = \beta(T\omega)\sigma(\omega)v(\omega) + \mathcal{B}(T\omega) - \mathcal{B}(\omega) \tag{C.4}$$

and

$$\mathcal{B}(\omega) = \sum_{k=0}^{\infty} \beta(T^{k+1}\omega)\left[\zeta_0 \dots \zeta_{k-1}v_k - (\sigma_k v_k)\mathbf{1}\right]. \tag{C.5}$$

Proof of Lemma 6.10. We claim that functions $l, \beta, \tilde{\beta}, \sigma, v$ and ζ are C^∞. The smoothness of ζ and v is proven in [20, Lemma 12.1], and the smoothness of β is proven in [20, equation (12.2)]. The smoothness of σ and l can be established similar to v and the smoothness of $\tilde{\beta}$ is similar to β. (C.4) now shows that $\boldsymbol{\rho}$ is C^∞ and moreover that the first term in the formula for $\boldsymbol{\Delta}$ is C^∞. It remains to show that $\tilde{\mathcal{B}}(\omega) := \mathcal{B}(\omega) - \mathcal{B}(\omega + \gamma)$ is C^∞. From (C.5) it follows that

$$\tilde{\mathcal{B}}(\omega) = \beta(\omega + \gamma)[v(\omega) - (\sigma(\omega)v(\omega))\mathbf{1}] + \sum_{k=0}^{\infty} \beta(\omega + (k+1)\gamma)\Lambda_k(\omega)$$

where

$$\Lambda_k(\omega) = [\zeta(\omega) - I]\zeta(\omega + \gamma) \dots \zeta(\omega + (k-1)\gamma)v(\omega + k\gamma).$$

In view of the foregoing discussion it remains to show that for each r there exist constants $C_r > 0$ and $\theta_r < 1$ such that

$$||\Lambda_k||_{C^r} \leq C_r\theta_r^k. \tag{C.6}$$

Denote

$$v_{k,l}(\omega) = \zeta(\omega + (k-1-l)\gamma) \dots \zeta(\omega + (k-1)\gamma)v(\omega + k\gamma),$$

$$w_{k,l} = \frac{v_{k,l}}{||v_{k,l}||}, \quad \eta_{k,l} = \ln\frac{||v_{k,l}||}{||v_{k,l-1}||}.$$

We have

$$\Lambda_k(\omega) = [\zeta(\omega) - I]v_{k,k-1}(\omega) = [\zeta(\omega) - I] \exp\left(\sum_{l=0}^{k-1} \eta_{k,l}\right) w_{k,k-1}(\omega)$$

$$= [\zeta(\omega) - I] \exp\left(\sum_{l=0}^{k-1} \eta_{k,l}\right) [w_{k,k-1}(\omega) - \mathbf{1}]$$

where the last equality holds since $\zeta(\omega)$ is a stochastic matrix. Using this representation we can deduce (C.6) from the following inequalities.

$$||w_{k,l} - \mathbf{1}||_{C^r} \leq C_r \theta_r^l, \tag{C.7}$$

$$||\eta_{k,l}||_{C^r} \leq C_r \theta_r^l. \tag{C.8}$$

Indeed (C.8) shows that $\left\|\sum_{l=0}^{k-1} \eta_{k,l}\right\|_{C^r} \leq \bar{C}_r$ and so $\left\|\exp\left(\sum_{l=0}^{k-1} \eta_{k,l}\right)\right\|_{C^r} \leq \bar{\bar{C}}_r$.

We note that, by the definition of $\eta_{k,l}$, (C.8) follows from (C.7), so it suffices to show the latter inequality. We shall use the following fact.

Lemma C.1. *Let $\Phi_j(x, u)$ be a family of contractions of a manifold X depending on a parameter u from an open set $D \subset \mathbb{R}^d$. That is, we assume that there exist constants $K > 0$ and $\theta < 1$ such that*

$$||D_x \Phi_j|| \leq \theta \tag{C.9}$$

and for some $r \geq 2$

$$||\Phi_j||_{C^r(X \times D)} \leq K.$$

Assume also that there exists a common fixed point for all values of the parameter, that is, there exists $p \in X$ such that for all $u \in D$

$$\Phi_j(p, u) \equiv p. \tag{C.10}$$

Then there are constants $\bar{K} > 0, \bar{\theta} < 1$ such that

$$\|\Phi_l \circ \Phi_{l-1} \circ \cdots \circ \Phi_1\|_{C^r(X \times D)} \leq \bar{K} \bar{\theta}^l. \tag{C.11}$$

To prove (C.7) we apply Lemma C.1 where X is a neighborhood of $\mathbf{1}$ in $(m-1)$-dimensional projective space and $\Phi_l(w, \omega) = \zeta(\omega - (l-1)\gamma)w$. To verify the conditions of the lemma we note that ζ contracts the Hilbert metric on the positive cone and that $\zeta(\cdot)\mathbf{1} \equiv \mathbf{1}$ since ζs are stochastic matrices. (See Appendix B for the definition of the Hilbert's metric and the related contraction properties of positive matrices.) This completes the proof of Lemma 6.10 modulo the proof of Lemma C.1 given below. □

Proof of Lemma C.1. Since the iterations of Φ converge to p exponentially fast, we may assume that we start in a small neighborhood of p. By passing to local coordinates we may further assume that X is a bounded domain in $\mathbb{R}^q$ for some q.

We prove (C.11) by induction on r. For $r = 0$ the estimate follows by contraction mapping principle. Let us now consider $r = 1$. Denoting

$$x_l = (\Phi_l \circ \cdots \circ \Phi_1)(x), \quad A_l = D_x(\Phi_l \circ \cdots \circ \Phi_1)x, \quad B_l = D_u(\Phi_l \circ \cdots \circ \Phi_1)x$$

we get

$$A_l = D_x \Phi_l(x_{l-1}) A_{l-1}, \quad B_l = D_x \Phi_l(x_{l-1}) B_{l-1} + D_u \Phi_l(x_l). \tag{C.12}$$

Now the required bound for A_l follows directly from (C.9). To estimate B_l we iterate the corresponding recurrence to get

$$B_l = \sum_{j<l} D_x\Phi_l \dots D_x\Phi_{j+1} D_u\Phi_j(x_{j-1})$$

To estimate the above sum we note that the terms with $j < l/2$ are exponetially small due to (C.9) while the terms with $j \geq l/2$ are exponentially small since (C.10) implies that $D_u\Phi_j(p,u) \equiv 0$ and so $D_u\Phi_j(x_{j-1}) = O(\theta^j)$. This proves the claim for $r = 1$ and completes the base of induction.

To perform the inductive step we assume that the Lemma holds for $r-1$. In view of the foregoing discussion to prove the result for r we need to estimate C^{r-1} norm of

$$(D_x(\Phi_l \circ \dots \circ \Phi_1), D_u(\Phi_l \circ \dots \circ \Phi_1)).$$

In view of (C.12) this reduces to studying the iterations of maps

$$\hat{\Phi}_j(x, A, B, u) = (\Phi_j(x), (D_x\Phi_j)A, (D_x\Phi_j)B + D_u\Phi_j).$$

Since $\hat{\Phi}_j$ are contractions having common fixed point $(p, 0, 0)$ the required estimate is true by inductive assumption. □

Proof of Lemma 6.12. We claim that functions l, $\beta, \tilde{\beta}, \sigma, v$ and $\mathcal{B}$ are Hölder continuous with respect to the metric $\mathbf{d}$. In fact, the Hölder continuity of λ and v is proven in [20, Appendix A]. The proof of Hölder continuity of $\tilde{\lambda}, l,$ and σ is very similar. Next the Hölder continuity of β and $\tilde{\beta}$ follows from the Hölder continuity of λ and $\tilde{\lambda}$, relation (C.3) and the Livsic Theorem [39]. To prove the Hölder continuity of $\mathcal{B}$ we note that the second factor in the sum (C.5) is exponentially small due to (2.10). Therefore the required statement is a consequence of Proposition C.2 below. □

Proposition C.2. *Given positive constants a, c_1 and c_2 there exists a constant $b = b(a, c_1, c_2)$ such that if $(X, \mathbf{d})$ is a metric space and*

$$H(x) = \sum_{k=1}^{\infty} H_k(x)$$

where

$$||H_k||_\infty \leq Ke^{-c_1 k}, \quad ||H_k||_{C^a(X)} \leq Ke^{c_2 k}$$

for some constant K. Then $H \in C^b(X)$.

Proof. For each n we have the following estimate

$$|H(x) - H(y)| \leq \left[\sum_{k=1}^{n-1} K\mathbf{d}^a(x,y)e^{c_2 k}\right] + 2\sum_{k=n}^{\infty} Ke^{-c_1 k} = K\left[\frac{e^{c_2 n} - e^{c_2}}{e^{c_2} - 1}\mathbf{d}^a(x,y) + \frac{e^{-c_1 n}}{1 - e^{-c_1}}\right].$$

Choosing n so that $e^{c_2 n}\mathbf{d}^a(x,y)$ and $e^{-c_1 n}$ are of the same order we obtain the claim. □

The proof of Lemma 7.2 relies on the following fact

Proposition C.3. *Let $\Phi'_n(x)$ and $\Phi''_n(x)$ be two families of contractions of a bounded metric space X. That is, assume that there are constants $K > 0$ and $\theta < 1$ such that* $\mathrm{diam}(X) \leq K$, *and for all $n \in \mathbb{N}$*

$$d(\Phi'_n(x_1), \Phi'_n(x_2)) \leq \theta d(x_1, x_2), \quad d(\Phi''_n(x_1), \Phi'_n(x_2)) \leq \theta d(x_1, x_2).$$

If there are constants C, σ such that for each $x \in X$ and for all $n \in \mathbb{N}$

$$d(\Phi'_n(x), \Phi''_n(x)) \leq \frac{C}{n^\sigma}$$

then there is a constant $\bar{C}$ such that for all $n \in \mathbb{N}$

$$d_n := d(\Phi'_n \dots \Phi'_1(x'), \Phi''_n \dots \Phi''_1(x'')) \leq \frac{\bar{C}}{n^\kappa}.$$

Proof. Iterating the estimate $d_n \leq \theta d_{n-1} + \dfrac{C}{n^\kappa}$ we obtain $d_n \leq \theta^{n-1} d_1 + \displaystyle\sum_{j=1}^{n-1} \frac{C\theta^{n-j}}{j^\kappa}$. Since $d_1 \leq K$ the result follows. □

Proof of Lemma 7.2. (7.11) follows from Proposition C.3 since the map relating ζ_n to ζ_{n-1} is a contraction in the total variation distance (see [18, Appendix D]) while the maps relating v_n to v_{n-1} and l_n to l_{n-1} are contractions in the Hilbert metric. (7.12) follows from (7.11) and the explicit formulas relating A_n, λ_n and $\tilde{\lambda}_n$ to ζ_n, v_n and l_n. Next

$$\frac{\bar{\beta}_n}{\beta_n} = \frac{\bar{\beta}_1}{\beta_1} \prod_{j=1}^{n-1} \left(\frac{\bar{\lambda}_j}{\lambda_j} \right).$$

Since the above series converges due to (7.12) we obtain that $\beta_+ = \lim\limits_{n\to+\infty} \dfrac{\bar{\beta}_n}{\beta_n}$ exists. The existence of $\beta_- = \lim\limits_{n\to-\infty} \dfrac{\bar{\beta}_n}{\beta_n}$ and $\tilde{\beta}_\pm = \lim\limits_{n\to\pm\infty} \dfrac{\bar{\tilde{\beta}}_n}{\tilde{\beta}_n}$ are similar.

Next the existence of $\mu_\pm$, $a_\pm$ and $b_\pm$ follows from the existence of the above limits in view of the formulae

$$\rho_n = \frac{c l_n}{\tilde{\beta}_n}, \quad \Delta_n = \beta_n \sigma_n v_n + \mathcal{B}_{n+1} - \mathcal{B}_n \tag{C.13}$$

with

$$\mathcal{B}_n = \sum_{k=n}^{\infty} \beta_{k+1} \left[\zeta_n \dots \zeta_{k-1} v_k - (\sigma_k v_k)\mathbf{1} \right].$$

proven in [20].

It remains to show that

$$\beta_+ = \tilde{\beta}^+ \quad \text{and} \quad \beta_- = \tilde{\beta}_-. \tag{C.14}$$

In view of (4.7)

$$\rho_n P_n(\mathfrak{m}_{n+1} - \zeta^-_{n+1}\mathfrak{m}_n) = \bar{\rho}_n \bar{P}_n(\bar{\mathfrak{m}}_{n+1} - \bar{\zeta}^-_{n+1}\bar{\mathfrak{m}}_n).$$

However due to (C.13) the ratio of the RHS to the LHS for $n \to \pm\infty$ equals to $\frac{\beta_\pm}{\tilde{\beta}_\pm}(1 + o_{n\to\pm\infty}(1))$ proving (C.14). □

References

[1] S. Alili: *Asymptotic behaviour for random walks in random environments,* J. Appl. Prob. **36** (1999) 334–349.

[2] N. Berger, M. Cohen, R. Rosenthal: *Local limit theorem and equivalence of dynamic and static points of view for certain ballistic random walks in i.i.d. environments,* Ann. Probab. **44** (2016) 2889–2979.

[3] G. Birkhoff: *Lattice Theory*, AMS Colloquium publications, Volume 25. Prvidence, Rhode Island, Third Edition (1973), vi+418p.
[4] E. Bolthausen: *Exact convergence rates in some martingale central limit theorems,* Ann. Probab. **10** (1982) 672–688.
[5] E. Bolthausen, I. Goldsheid: *Recurrence and transience of random walks in random environments on a strip,* Comm. Math. Phys. **214** (2000) 429–447.
[6] E. Bolthausen, I. Goldsheid: *Lingering random walks in random environment on a strip,* Comm. Math. Phys. **278** (2008) 253–288.
[7] E. Bolthausen, A.-S. Sznitman: *Ten lectures on random media,* DMV Seminar, **32** (2002) Birkhauser, Basel.
[8] E. Bolthausen, A.-S. Sznitman: *On the static and dynamic poins of view for certain random walks in random environment,* Methods and Applications of Analysis **9** (2002), 345–376.
[9] A. N. Borodin: *The asymptotic behavior of local times of recurrent random walks with finite variance,* Teor. Veroyatnost. i Primenen. 26 (1981) 769–783.
[10] J. Brémont: *One-dimensional finite range random walk in random medium and invariant measure equation,* Ann. Inst. Henri Poincaré Probab. Stat. 45 (2009) 70–103.
[11] J. Brémont: Random walk in quasi-periodic random environment. Stochastics and Dynamics, **9** (2009) 47–70.
[12] B. M. Brown: *Martingale central limit theorems,* Ann. Math. Statist. **42** (1971) 59–66.
[13] E. Bolthausen, A.-S. Sznitman: *On the static and dynamic poins of view for certain random walks in random environment,* Methods and Applications of Analysis **9** (2002), 345–376 .
[14] A. Chiarini, J.-D. Deuschel: *Local central limit theorem for diffusions in a degenerate and unbounded random medium,* Electron. J. Probab. **20** (2015), paper 112, 30 pp.
[15] I. P. Cornfeld, S. V. Fomin, Ya. G. Sinai: *Ergodic theory,* Grundlehren der Mathematischen Wissenschaften **245** (1982) Springer, New York, x+486 pp.
[16] B. Davis, D. McDonald: *An elementary proof of the local central limit theorem,* J. Theoret. Probab. **8** (1995) 693–701.
[17] D. Dolgopyat, B. Fayad, M. Saprykina: *Erratic behavior for 1-dimensional random walks in generic quasi-periodic environment,* in preparation.
[18] D. Dolgopyat, I. Goldsheid: *Limit theorems for random walks on a strip in subdiffusive regimes,* Nonlinearity **26** (2013) 1741–1782.
[19] D. Dolgopyat, I. Goldsheid: *Local Limit Theorems for random walks in a 1D random environment,* Arch. Math. **101** (2013) 191–200.
[20] D. Dolgopyat, I. Goldsheid: *Central Limit Theorem for recurrent random walks on a strip with bounded potential,* Nonlinearity **31** (2018) 3381–3412.
[21] D. Dolgopyat, I. Goldsheid: *Invariant measure for random walks on ergodic environments on a strip,* Ann. Prob. **47**, No. 4 (2019), 2494–2528.
[22] R. Durrett: *Probability: theory and examples.* 4th edition. Cambridge Univ. Press, Cambridge (2010).
[23] C. Gallesco, S. Popov, G. M. Schütz: *Localization for a random walk in slowly decreasing random potential,* J. Stat. Phys. **150** (2013) 285–298.
[24] V. P. Gaposhkin: *On the Dependence of the Convergence Rate in the Strong Law of Large Numbers for Stationary Processes on the Rate of Decay of the Correlation Function,* Th. Prob., Appl. **26** (1982) 706–720.
[25] I. Goldsheid: *Linear and Sub-linear Growth and the CLT for Hitting Times of a Random Walk in Random Environment on a Strip,* Prob. Th. Rel. Fields **141** (2008) 471–511.
[26] P. Hall, C. C. Heyde: *Martingale limit theory and its application,* Academic Press, New York-London, 1980. xii+308 pp.
[27] J. M. Harrison, L. A. Shepp: *On skew Brownian motion,* Ann. Probab. **9** (1981) 309–313.
[28] V. V. Jikov, S. M. Kozlov, O. A. Oleinik: *Homogenization of differential operators and integral functionals,* Springer, Berlin, 1994. xii+570 pp.
[29] V. Kaloshin, Ya. Sinai: *Nonsymmetric Random Walks along orbits of Ergodic Automorphisms,* AMS. Transl. Ser. 2, Vol. **198** (2000) 8 pp.
[30] H. Kesten: *A renewal theorem for random walk in a random environment,* Proc. Sympos. Pure Math., **31** (1977) 67–77.

[31] H. Kesten, M. V. Kozlov, F. Spitzer: *Limit law for random walk in a random environment.* Composito Mathematica **30**, 145–168 (1975).
[32] T. Komorowski, C. Landim, S. Olla: *Fluctuations in Markov processes. Time symmetry and martingale approximation,* Grundlehren der Mathematischen Wissenschaften **345** (2012) Springer, Heidelberg xviii+491 pp.
[33] S. Kozlov: *The method of averaging and walks in inhomogeneous environments,* Russian Math. Surveys **40** (1985) 73–145.
[34] S. Lalley: *An extension of Kesten's renewal theorem for random walk in a random environment,* Adv. in Appl. Math. **7** (1986) 80–100.
[35] J. Lamperti: *A new class of probability limit theorems,* J. Math. Mech. **11** (1962) 749–772.
[36] A. Lejay: *On the constructions of the skew Brownian motion,* Probab. Surv. **3** (2006), 413–466.
[37] M. Lenci: *On infinite-volume mixing,* Comm. Math. Phys. **298** (2010) 485–514.
[38] L. Leskela, M. Stenlund: *A local limit theorem for a transient chaotic walk in a frozen environment,* Stochastic Process. Appl. **121** (2011) 2818–2838.
[39] A. N. Livsic: *Cohomology of dynamical systems,* (Russian) Izv. Akad. Nauk SSSR Ser. Mat. **36** (1972) 1296–1320.
[40] M. Menshikov, A. Wade: *Random walk in random environment with asymptotically zero perturbation,* JEMS **8** (2006) 491–513.
[41] M. Menshikov, A. Wade: *Logarithmic speeds for one-dimensional perturbed random walks in random environments,* Stoch. Process. Appl. **118** (2008) 389–416.
[42] M. Menshikov, A. Wade: *Rate of escape and central limit theorem for the supercritical Lamperti problem,* Stochastic Process. Appl. **120** (2010) 2078–2099.
[43] M. Menshikov, S. Popov, A. Wade: *Non-homogeneous random walks: Lyapunov function methods for near-critical stochastic systems,* Cambridge Tracts in Math. **209** (2017) xviii+363 pp, Cambridge Univ. Press, Cambridge, UK.
[44] Yu. V. Prokhorov: *On the local limit theorem for lattice distributions,* Dokl. Akad. Nauk SSSR **98** (1954) 535–538.
[45] A. Roitershtein: *Tranzient random walks on a strip in a random environment,* Ann. Probab. **36** (2009) 2354–2387.
[46] Ya. G. Sinai: *The limiting behavior of a one-dimensional random walk in a random medium,* Theory Prob. Appl. **27** (1982) 256–268.
[47] Ya. G. Sinai: *Simple random walks on tori,* J. Statist. Phys. **94** (1999) 695–708.
[48] F. Solomon: *Random walks in a random environment,* Ann. Prob. **3** (1975) 1–31.
[49] M. Stenlund *A local limit theorem for random walks in balanced environments,* Electron. Commun. Probab. **18** (2013) paper 19, 13 pp.
[50] C. Stone: *Limit theorems for random walks, birth and death processes, and diffusion processes,* Illinois J. Math. **7** (1963) 638–660.
[51] O. Zeitouni: Random walks in random environment, XXXI Summer school in Probability, St. Flour (2001). Lecture notes in Math. **1837**, 193–312, Springer, Berlin, 2004.

Dmitry Dolgopyat: Department of Mathematics and Institute of Physical Science and Technology, University of Maryland, College Park, MD, 20742, USA

Ilya Goldsheid: School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, Great Britain