# Upper Tail Large Deviations in First Passage Percolation

RIDDHIPRATIM BASU
*International Centre for Theoretical Sciences*

SHIRSHENDU GANGULY
*University of California, Berkeley*

AND

ALLAN SLY
*Princeton University*

**Abstract**

For first passage percolation on $\mathbb{Z}^2$ with i.i.d. bounded edge weights, we consider the upper tail large deviation event, i.e., the rare situation where the first passage time between two points at distance $n$ is macroscopically larger than typical. It was shown by Kesten [24] that the probability of this event decays as $\exp(-\Theta(n^2))$. However, the question of existence of the rate function, i.e., whether the log-probability normalized by $n^2$ tends to a limit, remains open. We show that under some additional mild regularity assumption on the passage time distribution, the rate function for upper tail large deviation indeed exists. The key intuition behind the proof is that a limiting metric structure that is atypical causes the upper tail large deviation event. The formal argument then relies on an approximate version of the above which allows us to use independent copies of the large deviation environment at a given scale to form an environment at a larger scale satisfying the large deviation event. Using this, we compare the upper tail probabilities for various values of $n$. © 2021 Wiley Periodicals LLC.

**Contents**



Communications on Pure and Applied Mathematics, Vol. LXXIV, 1577–1640 (2021)

# 1 Introduction and Main Result

First passage percolation is a popular model of fluid flow through inhomogeneous random media, where one puts random weights on the edges of a graph and considers the first passage time between two vertices, which is obtained by minimizing the total weight among all paths between the two vertices. First passage percolation on Euclidean lattices was introduced by Hammersley and Welsh [19] in 1965 and has been studied extensively both in statistical physics and probability literature ever since. This model served as one of the motivations of developing the theory of subadditive stochastic processes, and the early progress using subadditivity was made by Hammersley-Richardson-Kingman [19, 27, 32] and culminated in the proof of the celebrated Cox-Durrett shape theorem [12] establishing the first-order law of large numbers behaviour for passage times between faraway points. Further progress was made in the 80s and 90s through efforts of Kesten [24–26] and Talagrand [36] establishing concentration inequalities for passage times, and Newman and others [30] on more geometric aspects of the model. Much progress has been made since [8,20] including a flurry of results in the last five years [1,9,15,16]. Despite this impressive progress, most of the fundamental questions still remain major mathematical challenges; see the survey [2] for a comprehensive history as well as an extensive list of the major open problems in this field.

One other reason planar first passage percolation came into prominence is that this model is believed to be in the KPZ universality class that was introduced by Kardar, Parisi, and Zhang [23] in 1986. Using nonrigorous renormalization group techniques, KPZ predicted universal scaling exponents for many (1+1)-dimensional growth models including first and last passage percolation under very general conditions on the passage time distribution (precise definitions later). An explosion of rigorous results in the last two decades starting with the seminal work of Baik, Deift, and Johansson [3] has now verified the KPZ prediction for a handful of models including last passage percolation with exponential, geometric, or Bernoulli passage times. However, this progress has been mostly restricted to the so-called exactly solvable (or, integrable) models where exact formulae are available using deep connections to algebraic combinatorics, representation theory, and random matrix theory; extremely detailed information has been obtained about such models by analyzing those formulae. Although the same results are qualitatively expected to hold for a much larger class of models, these methods rely very crucially on the exact formulae, and moving beyond the exactly solvable models remains a major challenge.

Our focus in this paper is such a problem in the nonintegrable setting of first passage percolation in the large deviation regime. The question first arose in the work of Kesten [24], who considered the probability of large deviation events in first passage percolation. Postponing the precise definitions momentarily, let us first describe informally the setup. Consider the passage time $\mathbf{T}_n$ from $(0,0)$ to $(n,0)$. The already mentioned shape theorem dictates that under some regularity conditions $\frac{\mathbf{T}_n}{n} \to \mu$ almost surely for some $\mu \in (0, \infty)$. The study of large deviations is concerned with the unlikely events $\{\mathbf{T}_n \geq (\mu + \varepsilon)n\}$ (upper tail) and $\{\mathbf{T}_n \leq (\mu - \varepsilon)n\}$ (lower tail). In the classical theory of large deviations, the log of such probabilities suitably scaled (by the so-called speed of large deviations) converges to a function of $\varepsilon$, known as the rate function. For first passage percolation, Kesten [24] showed the large deviation speed of $n$ and existence of the rate function for the lower tail using a subadditive argument. For the upper tail, Kesten showed a large deviation speed of $n^2$ for bounded edge weight distribution; however, the existence of the rate function remained open (see Open Question 18 in [2]). Our main result in this paper (see Theorem 1 below) answers this question establishing the existence of the rate function for the upper tail, which to our knowledge is the first such result beyond the exactly solvable models.

### 1.1 Model definitions and statement of result

We start with formal definitions of standard first passage percolation on $\mathbb{Z}^d$, $d \geq 2$. Let $E(\mathbb{Z}^d)$ denote the set of all nearest-neighbour edges in $\mathbb{Z}^d$. Let $\nu$ be a probability measure supported on the nonnegative real line. Let $\Pi = \{X_e : e \in E(\mathbb{Z}^d)\}$ denote a field of i.i.d. random variables where each $X_e$ (called the passage time of the edge $e$) has distribution $\nu$. For a sequence $\gamma = e_1 e_2 \cdots e_k$ of neighbouring edges (called a path), the passage time of the path, denoted by $\ell(\gamma)$, is defined as

$$\ell(\gamma) = \sum_{i=1}^{k} X_{e_i}.$$

For any two vertices $u$ and $v$, the first passage time between $u$ and $v$, denoted $\mathbf{PT}(u, v)$, is defined as the infimum of $\ell(\gamma)$ where $\gamma$ varies over all paths starting at $u$ and ending at $v$. Let $\mathbf{0}$ denote the origin. Under very mild conditions on $\nu$, it is a fundamental fact that for all $v \in \mathbb{Z}^d$, there exists $\mu(d, \nu, v) \geq 0$ such that

$$\lim_{n \to \infty} \frac{\mathbf{PT}(\mathbf{0}, nv)}{n} = \mu(d, \nu, v)$$

almost surely. For the special case when $v = (1, 0, \ldots, 0)$ is the unit vector along the first coordinate, we denote the limiting constant by just $\mu$, also known as the time constant in the literature. For the rest of this paper we shall focus on the planar case ($d = 2$) of the above model, and hence shall be in the setting of standard first passage percolation on $\mathbb{Z}^2$ unless otherwise mentioned. Let $\mathbf{n} := (n, 0)$ and let us denote the passage time $\mathbf{PT}(\mathbf{0}, \mathbf{n})$ by $\mathbf{T}_n$. As mentioned above we are concerned with the probability of the upper tail large deviation event

$$\mathscr{U}_\zeta(n) := \{\mathbf{T}_n \geq (\mu + \zeta)n\} \tag{1.1}$$

for some $\zeta > 0$. Throughout the paper we will work with the assumption that the passage time distribution has a continuous density with support $[0, b]$. Although we expect our method to extend beyond this condition, throughout the paper this will be our standing assumption, which is general enough, and yet, makes some of the proofs cleaner. For future reference we record this assumption below.

DEFINITION 1.1. For $b > 0$, let $\mathcal{P}(b)$ denote the set of all probability measures with support $[0, b]$ and a continuous density.

It is well-known that if $\nu \in \mathcal{P}(b)$ for any $b > 0$, then we have $0 < \mu < b$ (e.g., see [19]). Also observe that for $\nu \in \mathcal{P}(b)$, we have deterministically that $\mathbf{T}_n \leq bn$. So while considering the large deviation event $\mathscr{U}_\zeta$ in the above scenario, it suffices to consider $\zeta \in (0, b - \mu)$. Our main theorem shows that the large deviation rate function exists in the above setting.

THEOREM 1. *Consider standard first passage percolation on $\mathbb{Z}^2$ with passage time distribution $\nu \in \mathcal{P}(b)$ for some $b > 0$. Then for $\zeta \in (0, b - \mu)$ there exists $r = r(\nu, \zeta) \in (0, \infty)$ such that*

$$\lim_{n \to \infty} -\frac{\log \mathbb{P}(\mathscr{U}_\zeta(n))}{n^2} = r.$$

A couple of remarks are in order. First, there is nothing special about the direction $(1, 0)$; the same result holds for any unit vector $v$ with different rate function $r$, with minor adjustments in the proof. Also, a variant of this result is expected to hold in higher dimensions as well where the speed of the large deviation is $n^d$ rather than $n^2$ (see, e.g., (1.4)), and we expect that the same argument proving Theorem 1 may be used to prove the higher-dimensional analogue. However, in this paper we shall only focus on proving Theorem 1. We would also like to point out that our proof implies that $r(\nu, \cdot)$ is continuous on $(0, b - \zeta)$. Indeed, the proof crucially uses as an ingredient a result, Proposition 1.4, which implies that $r(\nu, \cdot)$ must be continuous on $(0, b - \zeta)$ if it exists. We are unable to establish convexity at this point and further smoothness properties seem out of reach of our current methods.

Observe that the condition in Theorem 1 is not optimal, and we have not made an attempt to make it the weakest possible. Together with the standard assumptions that the mass at 0 is less than the critical bond percolation probability on $\mathbb{Z}^2$ and that the edge distribution is not degenerate at a single point, Kesten assumed boundedness.[1] It is important to observe that one cannot completely remove this additional hypothesis; some condition is needed to ensure even the $n^2$ speed of the large deviation. For example, if the passage times are exponentially distributed, just increasing all the passage times around the origin by $(\mu + \zeta)n$ would force the large deviation event, while its probability being only exponentially small in $n$ (see the recent work [29], where indeed a sharp large deviations rate function has been established in this particular case). One can however prove Kesten's result for passage times with sufficiently fast decaying tails (see [13]), and one believes that

[1] Our additional assumption, of continuous density and full support in Definition 1.1 will only be used in the proof of Proposition 1.4, and we believe can be relaxed.

the rate function will exist in such a case too, possibly under some additional hypotheses. However, in this paper we have not pursued those directions, and instead focused on proving the result in the simplest possible case that is still sufficiently general to be of interest.

### 1.2 Background and related works

First passage percolation can be thought of as putting a random metric on $\mathbb{Z}^d$, where the distance between two vertices is given by the first passage time between them. As alluded to above, the most fundamental result about first passage percolation says that under suitable rescaling these metrics converge almost surely to a deterministic metric on $\mathbb{R}^d$ in a pointed Gromov-Hausdorff sense. More precisely, we have the following. Suppose $\nu \in \mathcal{P}(b)$ for some $b \in (0, \infty)$ (actually one only needs some moment condition and that the aforementioned standard assumption that mass of any atom at 0 is sufficiently small), and let $\widetilde{B}(t)$ denote the set of all vertices that are within distance $t$ of $\mathbf{0}$ in the FPP metric, and let $B(t) = \widetilde{B}(t) + [-\frac{1}{2}, \frac{1}{2}]^d$. Then there exists a nonrandom compact convex set $\mathcal{B} = \mathcal{B}_\nu$ with obvious symmetries such that for each $\varepsilon > 0$

$$\mathbb{P}\left((1-\varepsilon)\mathcal{B}_\nu \subset \frac{B(t)}{t} \subset (1+\varepsilon)\mathcal{B}_\nu \text{ for all large } t\right) = 1. \tag{1.2}$$

The set $\mathcal{B}$ is called the limit shape for this model. See, e.g., [12] for a proof of this. Recall the limiting constant $\mu(d, \nu, v)$ in direction $v$. It is not hard to see that $\mu(d, \nu, \cdot)$ can be extended to a norm in $\mathbb{R}^d$ and $\mathcal{B}$ is the unit ball corresponding to this norm. The shape theorem implies that at large scales, the distance function in the FPP metric in a fixed direction grows approximately linearly with the Euclidean distance, and the convexity of the limit shape is then just a consequence of the triangle inequality.

The shape theorem is a law of large number result, and the natural next question of obtaining fluctuations has been extensively investigated. The moderate deviation estimates are interesting, particularly in $d = 2$, where KPZ scaling predicts a fluctuation exponent of $\frac{1}{3}$. However, the best-known fluctuation and concentration bounds (for $\mathbf{T}_n$) have so far been proved at $n^{1/2+o(1)}$ scale [8, 26, 36]. In this paper, we are looking at the large deviation regime, i.e., where we consider a linear deviation of $\mathbf{T}_n$ from its long term value. Although we recall standard results only for $\mathbf{T}_n$, qualitatively the same results hold in all directions. Also, we are assuming throughout that the passage time distribution is in $\mathcal{P}(b)$ for some $b$, although many of these results hold under weaker assumptions.

Kesten [24] considered both upper and lower tail large deviations for first passage percolation. Let $\mathscr{L}_\zeta(n) := \{\mathbf{T}_n \le (\mu - \zeta)n\}$ (throughout this section for brevity we will use $\mathbf{T}_n$ to denote the passage time between $(0, 0, \ldots, 0)$ and $(n, 0, \ldots, 0)$ in $\mathbb{Z}^d$ although it was initially defined only for $\mathbb{Z}^2$) denote the lower tail large deviation event. Using a subadditive argument, Kesten showed that for

$\zeta \in (0, \mu)$,

$$\lim_{n\to\infty} -\frac{\log \mathbb{P}(\mathscr{L}_\zeta(n))}{n} = r_\ell(\zeta) \in (0,\infty). \tag{1.3}$$

While the existence of $r_\ell(\zeta)$ follows from subadditivity, showing positivity of the same requires more work and relies on percolation arguments.

For the upper tail large deviations, Kesten showed that

$$0 < \liminf_{n\to\infty} -\frac{\log \mathbb{P}(\mathscr{U}_\zeta(n))}{n^d} \le \limsup_{n\to\infty} -\frac{\log \mathbb{P}(\mathscr{U}_\zeta(n))}{n^d} < \infty. \tag{1.4}$$

The existence of the limit was left open, and this open question was reiterated in [2] (see Question 18 there), which we answer in our Theorem 1.

Observe that the speed of large deviations is different in the upper and lower tails. This is not unexpected and can be intuitively explained as follows: For $\mathbf{T}_n$ to be much smaller than $\mu n$, one needs only one path that is atypically small; however it is much more unlikely for $\mathbf{T}_n$ to be atypically large, since typically one can find $n^{d-1}$ many 'parallel' short paths between the origin and $(n,0,0,\dots,0)$ that are disjoint except at the beginning and the end. Thus to attain the upper tail event, all such paths need to be large, each of which costs $e^{-\Theta(n)}$ and hence the total cost is at least $(e^{-\Theta(n)})^{n^{d-1}}$. Indeed, this feature is quite common in many growth models, e.g., last passage percolation, the parabolic Anderson model, and deviation of the spectrum of GUE (see [14] and the references therein).

As a matter of fact, among the only cases of growth models where the existence of rate function is known for both tails are the so-called exactly solvable models of last passage percolation. As an illustration, we only describe the result for the case of exponential directed last passage percolation in $\mathbb{Z}^2$ [22]; however, the same qualitative result is known in the case of Poissonian directed last passage percolation in $\mathbb{R}^2$ [17] and last passage percolation on $\mathbb{Z}^2$ with geometric edge weights [22]. Consider the following last passage percolation model on $\mathbb{Z}^2$ where each vertex is equipped with an i.i.d. sample of Exp(1) random variable. As before, the weight of any path is the sum of weights on it. The difference from the first passage percolation model is that we only consider up/right directed paths, and the last passage time between two vertices is calculated by maximizing the weight over all such paths between the two vertices. This is one of the first exactly solvable models rigorously shown to be in the KPZ universality class by Johansson [22] using exact determinantal formulae. Let $L_n$ denote the last passage time from $(0,0)$ and $(n,n)$. It is well-known [33] that $\frac{L_n}{n} \to 4$ almost surely as $n \to \infty$. Johansson proved large and moderate deviation estimates for $L_n$. In particular, he proved that

$$\lim_{n\to\infty} \frac{\log \mathbb{P}(L_n \ge (4+\zeta)n)}{n} = -I_u(\zeta), \quad \zeta > 0,$$

and

$$\lim_{n\to\infty} \frac{\log \mathbb{P}(L_n \le (4-\zeta)n)}{n^2} = -I_\ell(\zeta), \quad \zeta \in (0,4).$$

The functions $I_\ell$ and $I_u$ could in principle be explicitly evaluated there. Observe that for last passage percolation, as expected, the roles of the upper tail and the lower tail are reversed, but qualitatively there is no other difference from the FPP case. Johansson [22] also proved a similar result for last passage percolation with geometric passage times. Even prior to [22], the $n$-speed upper tail rate function for exponential LPP was obtained by Seppäläinen in [34]. The analogous result in the context of Poissonian LPP was proved by Deuschel and Zeitouni in [17] and Seppäläinen in [35].

However, the above results concerning LDP at speed $n^2$ use some form of integrability, and the proofs rely heavily on the specific passage time distributions that are intimately connected to the integrable features in these models.

Although as far as we are aware, our result is the first one proving the existence of a large deviation rate function for the $n^2$-speed tail for point-to-point passage times in a nonintegrable setting. One variant of such a result was proved by Chow and Zhang [10] in the case of line-to-line first passage time in standard first passage percolation. Formally Chow and Zhang considered the minimum passage time over all paths with one endpoint in $A = \{(0,i) : i \in \{0,1,\ldots,n\}\}$ and the other endpoint in $B = \{(n,i) : i \in \{0,1,\ldots,n\}\}$; moreover, they consider the geodesic restricted to lie in the square $[0,n]^2$. Let us denote the passage time by $\mathbf{T}_n^*$. It is a standard result [24] that $\frac{\mathbf{T}_n^*}{n} \to \mu$ almost surely as $n \to \infty$. In [10], Chow and Zhang showed that for $\zeta > 0$

$$\lim_{n\to\infty} -\frac{\log \mathbb{P}(\mathscr{U}_\zeta(n))}{n^2}$$

exists and is nontrivial. The appropriate variant of their result holds in all dimensions. Even though the specific geometric setting considered in [10] causes significant simplification, and in particular rules out backtracks of the geodesic, it is worth mentioning that the argument in [10] is an approximate subadditive argument, which bears resemblance to our approach at least at a high level (see Section 1.3 for more details). The open question addressed by Theorem 1 was also mentioned in [10].

We end this section with a brief discussion about a related line of work concerning geometric consequences of large deviation events in first/last passage percolation. Formally one considers the measure obtained by conditioning on the large deviation events, and investigates how the geometry of the random field of weights changes. These questions were considered in the setting of exactly solvable Poissonian last passage percolation for the upper tail (i.e., the tail with large deviation speed $n$) by Deuschel and Zeitouni who, in [17], showed that under the upper tail large deviation event, the maximizing paths between two faraway points is with high probability localized around the straight line segment joining the two endpoints. This was refined recently in the case of exponential LPP in [5], which established the precise exponent governing the localization. For the harder lower tail case, in a recent paper [6] we showed that forcing the large deviation event makes the path delocalized with high probability. Although the basic framework of the latter was last passage percolation, the argument does not rely on integrable probability (see remarks in [6] for more details).

### 1.3 An approximate monotonicity and the proof of Theorem 1

The argument proving Theorem 1 is quite involved and has many pieces going into the proof. The purpose of this section is to provide a broad overview of the key steps. At a very high level, our argument intuitively is predicated on the existence of a limiting metric structure as in (1.2) even in the upper tail large deviation regime, which roughly implies that conditional on the large deviation event, the distances in a fixed direction grow linearly at large scales, and as the direction is varied, the gradient changes in a reasonably regular way. The reason to expect this is intimately tied to the reason behind the $n^2$ speed of large deviation, which causes the edge distributions of $\Theta(n^2)$ many edges to change.

Although we believe the above statement to be true, for the purposes of the proof it suffices to have subsequential limits. In fact, the exact statement that we prove is much less refined (see Proposition 2.4).

For the remainder of the paper, let $b > 0$ and $\nu \in \mathcal{P}(b)$ be fixed. Recall that $\mu$ denotes the time constant in the $x$-direction for the standard first passage percolation on $\mathbb{Z}^2$ with $\nu$-distributed edge weights. Let $\zeta \in (0, b - \mu)$ be fixed. For $n \in \mathbb{N}$, let $a_n = a_n(\zeta)$ be defined by

$$a_n = \log \mathbb{P}(\mathscr{U}_\zeta(n)).$$

Theorem 1 will follow easily from the following approximate monotonicity result.

PROPOSITION 1.2. *For each $\varepsilon > 0$, there exists $N_0 > 0$ such that the following holds. For all $n \in \mathbb{N}$ with $n > N_0$ there exists $M_0 = M_0(n)$ such that for all $m > M_0$ we have*

$$\frac{a_m}{m^2} \geq \frac{a_n}{n^2} - \varepsilon.$$

Most of this paper is devoted to proving Proposition 1.2, but before we outline its proof let us quickly finish the proof of Theorem 1 assuming the above.

PROOF OF THEOREM 1. Let

$$\mathfrak{a} = \limsup_{n\to\infty} \frac{a_n}{n^2}, \quad \mathfrak{a}' = \liminf_{n\to\infty} \frac{a_n}{n^2}.$$

By Kesten's result (1.4) we know that $-\infty < \mathfrak{a}' \leq \mathfrak{a} < 0$, and hence it suffices to prove that for all $\varepsilon > 0$, we have $\mathfrak{a}' \geq \mathfrak{a} - 2\varepsilon$. Fix $\varepsilon > 0$, and let $N_0$ be such that the conclusion of Proposition 1.2 holds. Pick $N_1 > N_0$ such that $a_{N_1}/N_1^2 \geq \mathfrak{a} - \frac{\varepsilon}{2}$, and pick $N_2 > M_0(N_1)$ as in Proposition 1.2 such that $a_{N_2}/N_2^2 \leq \mathfrak{a}' + \frac{\varepsilon}{2}$. Proposition 1.2 now implies that $\mathfrak{a}' \geq \mathfrak{a} - 2\varepsilon$, as required. This completes the proof of the theorem. □

### 1.4 Auxiliary propositions, key ideas, and the proof of Proposition 1.2

The rest of this paper proves Proposition 1.2. In this section we will state Propositions 1.3 and 1.4. The former is the key result that lower-bounds the probability of the large deviation at a larger scale $m$ (with a slightly decreased excess) in terms of the large deviation probability at a smaller scale $n$. The second proposition proves continuity of the normalized log-probability of the large deviation event. We will complete the proof of Proposition 1.2 modulo the two just-mentioned propositions and discuss the key ideas behind the proofs of the latter. For the sake of exposition, we will not be very precise in our outline of the ideas. In particular, the discussion will involve terms such as $o(n^2)$ or $o(1)$, which should be interpreted as terms that are an arbitrarily small constant times $n^2$ or 1, where the arbitrarily small constant depends on other parameters without necessarily being a sequence of constants that go to 0 with $n$.

Observe that to prove Proposition 1.2, we need to obtain a lower bound to $\mathbb{P}(\mathscr{U}_\zeta(m))$ in terms of $\mathbb{P}(\mathscr{U}_\zeta(n))$ for $m \gg n \gg 1$. The first (and the most important) step is to construct an event with probability at least $\mathbb{P}(\mathscr{U}_\zeta(n))^{m^2/n^2}$ (up to an error of $e^{-o(m^2)}$) on which we shall have $\{\mathbf{T}_m \geq (\mu + \zeta')m\}$ for $\zeta'$ smaller but arbitrarily close to $\zeta$.

Formally, via this construction, we will prove the following proposition.

PROPOSITION 1.3. *For each $\varepsilon' \in (0, \zeta)$ and $\varepsilon > 0$, there exist $N_0$ and $H_0$ such that for all $n > N_0$ and $m > nH_0$ we have*

$$\log \mathbb{P}(\mathscr{U}_{\zeta-\varepsilon'}(m)) \geq \frac{m^2}{n^2} \log \mathbb{P}(\mathscr{U}_\zeta(n)) - \varepsilon m^2.$$

In fact, it would be convenient to note that proving Proposition 1.3 for the case when $m$ is divisible by $n$ implies the same for all $m$. To see this, write $m = nk + r$ where $r < n$ and note that applying Proposition 1.3 to $nk$ and $n$, with $\frac{\varepsilon'}{2}$ and $\frac{\varepsilon}{2}$ in place of $\varepsilon'$ and $\varepsilon$, we get

$$\begin{aligned} \log \mathbb{P}(\mathscr{U}_{\zeta-\varepsilon'}(m)) \geq \log \mathbb{P}(\mathscr{U}_{\zeta-\frac{\varepsilon'}{2}}(nk)) &\geq k^2 \log \mathbb{P}(\mathscr{U}_\zeta(n)) - \frac{\varepsilon}{2} n^2 k^2 \\ &\geq \frac{m^2}{n^2} \log \mathbb{P}(\mathscr{U}_\zeta(n)) - \varepsilon m^2, \end{aligned} \tag{1.5}$$

where the first inequality follows by observing that

$$\begin{aligned} \mathbf{T}_{nk} \geq \left(\mu + \zeta - \frac{\varepsilon'}{2}\right) nk \implies \mathbf{T}_m &\geq (\mu + \zeta - \frac{\varepsilon'}{2}) nk - rb \\ &\geq (\mu + \zeta - \varepsilon')m \end{aligned}$$

for all large enough $k$. The second inequality in (1.5) is an application of the above proposition and the final inequality follows since $m \geq nk$ and $\log(\mathbb{P}(\mathscr{U}_\zeta(n)) \leq 0$.

Once we have Proposition 1.3 at our disposal, all we need to prove Proposition 1.2 is a way to compare $\mathbb{P}(\mathscr{U}_\zeta(n))$ and $\mathbb{P}(\mathscr{U}_{\zeta'}(n))$ when $\zeta$ and $\zeta'$ are close. To this end we have the following proposition, which essentially says that if the rate function exists it must be continuous in $\zeta$.

PROPOSITION 1.4. *For each $\varepsilon > 0$, there exists $\varepsilon' > 0$ such that for all $n$ sufficiently large we have*

$$\frac{\log \mathbb{P}(\mathscr{U}_{\zeta-\varepsilon'}(n))}{n^2} \le \frac{\log \mathbb{P}(\mathscr{U}_{\zeta}(n))}{n^2} + \varepsilon.$$

Our assumption of the edge distribution possessing a continuous density (see Definition 1.1) is essentially only used in the proof of the above. Although we expect that this result might be proven more generally, we have not made such an attempt in this paper. It is easy to complete the proof of Proposition 1.2 using Propositions 1.3 and 1.4.

PROOF OF PROPOSITION 1.2. The proof follows immediately by noticing that

$$\frac{a_m}{m^2} \ge \frac{\log \mathbb{P}(\mathscr{U}_{\zeta-\varepsilon'}(m))}{m^2} - \varepsilon \ge \frac{a_n}{n^2} - 2\varepsilon,$$

where the first inequality is the content of Proposition 1.4 and the second inequality is the content of Proposition 1.3. □

The rest of this paper deals with proving Propositions 1.3 and 1.4. Proof of Proposition 1.4 is easier. Essentially one shows that to increase the passage time $\mathbf{T}_n$ by $\varepsilon' n$, it suffices to increase the passage times of all the edges inside a box of size $O(n)$ by $O(\varepsilon')$. The cost of such a change can be made as small as possible in the exponential scale by choosing $\varepsilon'$ small enough and using the continuity of the density of $\nu$. The only subtle point is that since the variables are supported on $[0, b]$, one cannot increase the values of the edges that already have values close to $b$. However, by choosing the parameters carefully we ensure that there are not too many edges of the latter kind and that the geodesic necessarily passes through many edges whose values are away from $b$, in which case the perturbation strategy works. The formal proof appears in Section 5. The remainder of this section presents an outline of the proof of Proposition 1.3, which is really the heart of this paper.

For the purpose of illustration, we shall only outline the proof in the special case $m = 2n$. Also, we shall pretend, for the time being, that the event $\{\mathbf{T}_n \ge (\mu+\zeta)n\}$ only depends on the edge weights in the box $B = [\![0, n]\!] \times [\![-\frac{n}{2}, \frac{n}{2}]\!]$, where $[\![a, b]\!] := [a, b] \cap \mathbb{Z}$. Observe that while this is not deterministically true because the paths are allowed to backtrack, a version of this holds with high probability if one replaces $B$ by a box of side length being a large ($\nu$ dependent) constant times $n$ and centered at the origin. This is what we will do throughout the rest of the paper (see, e.g., the discussion around (2.2)).

Let $\varepsilon$ be an arbitrary small positive number, and suppose $\mathbb{P}(\mathbf{T}_n \ge (\mu + \zeta)n) = p$. So our task is to create an environment on $B_1 = [\![0, 2n]\!] \times [\![-n, n]\!]$ with probability at least $p^4 e^{-o(n^2)}$, on which we shall have $\{\mathbf{T}_{2n} \ge (\mu + \zeta - \varepsilon)2n\}$. The basic idea of such a construction is as follows. Consider the large deviation event $\{\mathbf{T}_n \geq (\mu + \zeta)n\}$. We shall show that $B$ can be tiled by subboxes of size $k \times k$ (which we will call 'tiles'; see Figure 1.1) such that there exists a subevent of $\{\mathbf{T}_n \geq (\mu + \zeta)n\}$ with probability at least $pe^{-o(n^2)}$ such that, for any environment $\Pi$ in this subevent, most of the tiles in $B$ are *stable*. We shall choose $k = \frac{n}{2^J}$ for some large $J$ that remains bounded independent of $n$.

The notion of stability of a tile is defined precisely later (see Definition 2.8), but here we give a simplified (and somewhat vague) description for the purpose of exposition. Roughly, a $k \times k$ tile being stable means that for all points $\mathbf{z}$ in the tile, the passage time starting from $\mathbf{z}$ in each fixed direction grows approximately linearly with the Euclidean distance at scale $k$; i.e., the passage time from $\mathbf{z}$ to the point in direction $\theta$ at distance $2k$ is approximately twice the passage time from $\mathbf{z}$ in the direction $\theta$ at distance $k$ and so on). For example, for $\mathbf{z}$ in the tile and any $\theta \in \mathbb{S}^1$, let $\mathbf{z}_1$ and $\mathbf{z}_2$ be points such that $\mathbf{z}, \mathbf{z}_1, \mathbf{z}_2$ lie in a straight line making angle $\theta$ with the $x$-axis and $\|\mathbf{z} - \mathbf{z}_1\| = \|\mathbf{z}_1 - \mathbf{z}_2\| = k$ where $\|\cdot\|$ denotes the Euclidean norm.[2] The box is then said to stable if

$$\mathbf{PT}(\mathbf{z}, \mathbf{z}_1) = (1 + o(1))\mathbf{PT}(\mathbf{z}_1, \mathbf{z}_2) = (1 + o(1))\frac{\mathbf{PT}(\mathbf{z}, \mathbf{z}_2)}{2} \tag{1.6}$$

for each such $\mathbf{z}$ and each direction $\theta$. The actual definition of stability will ask for something stronger (e.g., a similar condition for a sequence of larger number of equally separated points $\mathbf{z}, \mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_\ell$ on a line and a larger number of different scales of separations instead of the separation being only $k$; see Section 2.1 for precise definitions).

Before proceeding further, we discuss our choice of scales. As indicated above, we shall show that for some $J \in \mathbb{N}$ (depending in a somewhat complicated fashion on parameters governing stability, but remaining bounded as $n \to \infty$), there exists an event with probability at least $pe^{-o(n^2)}$ (called the **Base-event**) contained in the large deviation event $\{\mathbf{T}_n \geq (\mu + \zeta)n\}$ with the following property. On this subevent, we can tile the environment with tiles of size $\frac{n}{2^J}$ such that except at most $O(\varepsilon)$ fraction of the tiles, all other tiles are stable. Note that, in the definition of stability, the gradient of the linear function at a given point and in a given direction can a priori depend on the environment $\Pi$. However, by a picking a fine enough mesh and rounding to the nearest mesh point, we shall restrict ourselves to a further subevent, still of probability at least $pe^{-o(n^2)}$, such that each environment in the subevent yields the same mesh point after rounding (in particular, this implies that the ratios of $\mathbf{PT}(\mathbf{z}, \mathbf{z}_1)$ computed on any two different environments in this subevent agree up to $1 + o(1)$ multiplicative factors). See Proposition 2.4 and Lemma 3.11 for precise statements of the above results.

[2] Throughout the paper we shall use $\|\cdot\|$ to denote the Euclidean norm for points in $\mathbb{R}^2$. Occasionally we shall also need to use the $\ell_1$ norm for points and vectors in $\mathbb{R}^2$, which will be denoted by $\|\cdot\|_1$.

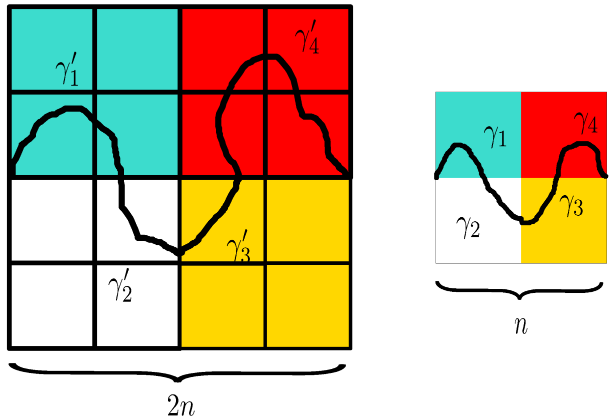


FIGURE 1.1. Figure illustrating the proof sketch below where for every path $\gamma'$ in the dilated environment $\Pi$ there exists a path $\gamma$ obtained by scaling down the endpoints of the excursions. However, note that a priori $\gamma_i$ need not be excursions even though $\gamma_i'$ are by definition. The former are just taken to be the shortest path in the environment $\Pi$ between the endpoints of $\gamma_i'$ divided by 2. Here $k = \frac{n}{2}$.

For convenience of exposition, let us ignore the unstable tiles, and explain how to use the **Base-event** (a tiling of $B$ into stable $k \times k$ tiles) with probability $pe^{-o(n^2)}$ to construct a subset of configurations on the larger box $B_1$ with the desired properties. We construct an environment on $B_1$ by independently sampling environments $\Pi_1, \Pi_2, \Pi_3, \Pi_4$ on $B$ with the same law as $\Pi$ (conditioned on **Base-event**). Equipped with this, we now tile $B_1$ using tiles of size $2k \times 2k$ where each such tile is formed from four tiles of size $k \times k$ (one from each $\Pi_i$) as illustrated in Figure 1.1. Let us call the constructed environment $\Pi'$. This procedure will be referred to as *dilating the environment* in what follows. Given such a construction, we will be done once we establish the following two properties:

(1) The constructed event has probability at least $p^4 e^{-o(n^2)}$.
(2) Any path $\gamma'$ in $\Pi'$ between $\mathbf{0}$ and $\mathbf{2n}$ has passage time at least $(\mu+\zeta-\varepsilon)2n$. (In the formal treatment, we will actually choose our boxes to be bigger so that any path that exits the box before reaching $\mathbf{2n}$ can automatically be guaranteed to have passage time at least $(\mu + \zeta - \varepsilon)2n$).

The lower bound on the probability of **Base-event** easily implies the first item, and our construction of four independent environments conditioned on **Base-event**. The second item is more involved, and this argument (in full generality) takes up Section 4. The basic idea is the following: We decompose $\gamma'$ into excursions $\gamma_1', \gamma_2', \ldots,$ where each $\gamma_i'$ resides in a tile of size $2k \times 2k$ and $\gamma_i'$ and $\gamma_{i+1}'$ reside in separate tiles. (Observe that such a decomposition into excursions can a priori be quite wild involving a lot of backtracks, but we shall add some thin corridors between the tiles with high passage times so that it would suffice to consider a nicer class of paths, with well-behaved excursions; see Section 4.1 for details.) Thus $\ell(\gamma') = \sum_i \ell(\gamma_i')$. Now the key is to observe that for such a $\gamma'$ one can create a path $\gamma$ between $\mathbf{0}$ and $\mathbf{n}$ such that $\gamma$ is a concatenation of paths $\gamma_1, \gamma_2, \ldots$. This is done by just taking $\gamma_i$ to be the shortest path between points that are the endpoints of $\gamma_i'$ scaled down by a factor 2 (see Figure 1.1). However, note that $\gamma_i$ need not be excursions even though $\gamma_i'$s are by definition. The former are just taken to be the shortest path in the environment $\Pi_1$ (say) between the points obtained by dividing the endpoints of $\gamma_i'$ by 2.

The stability of the tiles now imply that $\ell(\gamma_i') \geq (1+o(1))2\ell(\gamma_i)$, where the left-hand side is computed on $\Pi'$ and the right-hand side on $\Pi_1$ say. Thus it follows that

$$\begin{aligned}\ell(\gamma') &= \sum_i \ell(\gamma_i') \geq 2(1+o(1)) \sum_i \ell(\gamma_i) \\ &= 2(1+o(1))\ell(\gamma) \geq 2(1+o(1))(\mu+\zeta)n,\end{aligned} \tag{1.7}$$

where the last inequality follows by definition, as $\gamma$ is a path formed by concatenating $(\gamma_i)_i$ between $\mathbf{0}$ and $\mathbf{n}$ in the environment $\Pi_1$, which is in $\mathscr{U}_\zeta(n)$.

As indicated before, we have only attempted to present the high level ideas involved in the arguments without much discussion on the quantifiers involved in the precise statements. We end with a brief discussion on some of the technical aspects.

(1) The most important step is to prove that $B$ can be divided into such stable tiles. In fact, we prove that there exists a tiling of $B$ where most tiles are stable, i.e., the total number of points in unstable tiles is $o(n^2)$. This essentially only uses the fact that with high probability, the FPP metric is bi-Lipschitz with respect to the Euclidean metric at all large enough scales, which in turn is a consequence of the shape theorem in (1.2). Under the conditioned large deviation event, this continues to persist due to the FKG inequality (we record this observation in Lemma 2.2). The formal stability result is Proposition 2.4 in this paper, and the proof is provided in Section 7, where a detailed outline of the proof and an elaborate explanation of the key ideas can be found.

Intuitively the result says that any subsequential limiting metric structure due to its bi-Lipschitz nature should have a reasonably smooth gradient function (see (2.10) and Definition 2.11). Thus the size of the tiles capture the scale at which an approximate smoothness is witnessed. However, formally we show (see Proposition 2.4) that all but at most a small fraction of tiles are stable, and the unstable tiles can be handled by replacing all the edge values in those by values close to $b$ (recall that $\nu$ has support $[0, b]$). This operation only can increase the passage time and hence makes the upper tail event more likely and on the other hand it only costs $e^{-o(n^2)}$ in probability and hence does not change any of the conclusions.

(2) Finally, we describe briefly another point among many which we have swept under the carpet so far. All the discussion above describes how to construct

a $2n \times 2n$ environment out of an $n \times n$ environment preserving (up to an error) the upper tail large deviation event. However, observe that in order to prove Proposition 1.3, we need to be able to dilate the original environment by factor $h = \frac{m}{n}$, which could be arbitrarily large. To ensure that the error term $(1 + o(1))$ in (1.7) does not blow up, we will in fact modify the notion of stable tiles which allows dilation by an arbitrary factor $h$. (As mentioned earlier, we shall choose the tile size $k = \frac{n}{2^J}$ for some large scale $J$ that depends on certain parameters including $n$ and $m$, but remains uniformly bounded by an absolute constant $J_2$. Now such a $J$ would a priori be random and dependent on the environment. However, on account of the uniform boundedness, a simple application of the pigeonhole principle implies there exists a $J$ such that certain desired properties such as stability hold at the scale determined by $J$ with probability at least $\frac{1}{J_2}$. This is the scale we will choose. We will in fact show the existence of such a $J$ by an application of the probabilistic method by introducing a certain additional artificial randomness which will be explicitly discussed in Section 7). To ensure this, we prove that stable tiles have a couple of additional properties:

- First of all, we need to ensure stability at most locations at many consecutive length scales (this was already alluded to before) rather than just two as in (1.6).
- More importantly, we show that as the direction vector is varied at a given location, the gradient field has approximate convexity properties. This result should be thought of as a weak analogue of the convexity of the limiting shape in (1.2) in the upper tail large deviation regime, and this will enable us to compare the distance function between the $k \times k$ box and the $kh \times kh$ box. The formal convexity statement is stated as Proposition 3.4 and the proof is presented in Section 6.

## 1.5 Related future directions

We end this section by briefly pointing out that the general technique developed in this paper is expected to be applicable to a wide array of problems, a few of which are discussed below. We expect our methods to be adaptable to the case of lower tail ($n^2$-speed) large deviations in directed last passage percolation in $\mathbb{Z}^2$.

Another related object of study is the entire space-time evolution profile of the last passage time or polymer energy, i.e.,

$$\{L_v : v = (v_1, v_2), nt_1 \le v_1 + v_2 \le nt_2, nu_1 \le |v_1 - v_2| \le nu_2\},$$

where $L_v$ denotes the last passage time from $(0,0)$ to $v$. For the case of exponential LPP, using the correspondence to TASEP, this question is equivalent to understanding the height function of the so-called corner growth process; large deviations for the $n$-speed tail in this case was obtained in [21,38] while for the $n^2$-speed tail, only upper and lower bounds have recently been obtained in [31] starting from general initial data. We believe that our methods could possibly be sufficiently robust to handle the lower tail large deviation for the space-time evolution for a large class of passage time distributions going beyond the integrable case of exponential LPP.

Another promising direction of possible applications lie in the realm of positive temperature polymer models in $(1+1)$-dimension. These are variants of last passage percolation models where instead of choosing the maximal weight path, one puts a probability measure on the space of all directed paths (from $(0,0)$ to $(n,n)$, say) which assigns a probability proportional $\exp(\beta H(\gamma))$ to a path $\gamma$ where $\beta$ is the inverse temperature and $H(\gamma)$ is the sum of weights along $\gamma$. The quantity of interest here is the log of the partition function $Z_{n,\beta} := \sum_\gamma \exp(\beta H(\gamma))$. In [7], results of [13] were generalized to establish the $n^2$ speed of the lower tail large deviations for $\log Z$ under certain tail conditions on the weight distribution. A precise upper tail large deviation rate function was established for the $n$-speed upper tail for the exactly solvable log-gamma polymer [18]. We believe that our methods could be useful to prove the existence of rate functions for lower tail large deviations of $\log Z$ under certain conditions that guarantee $n^2$ speed of the large deviations. In this context, it is worth mentioning the recent progress [11, 37] on the related problem of the lower tail large deviation for the KPZ equation, using techniques depending crucially on the exactly solvable nature of the problem. In a forthcoming project, with Manan Bhatia [4], the first two authors address the lower tail large deviations of a Poissonized positive temperature model, which is not known to exhibit any integrable properties.

Finally, to deduce interesting geometric consequences of large deviations, several key steps often have to be established beyond proving the existence of a rate function. In many natural cases the rate function turns out to have nice analytic properties like convexity, which we also expect in our case. Moreover, as the previous discussion on the key idea of the proof in Section 1.3 suggests, we expect a shape theorem (a limiting metric space) even in the large deviation regime analogous to the typical behaviour mentioned in (1.2). The above and other related directions form a general program of systematically studying large deviations in nonintegrable settings to be pursued in future research.

## 1.6 Organization

We finish this introduction by describing the organization of this paper. In Section 2 we set up the notation and make a precise statement of the stability result Proposition 2.4. We also state precisely the regularity results of the gradient field. The proofs of these results are postponed until later. In Section 3, we state Proposition 3.4, a key approximate convexity result for the distance function. This section also contains the definition of the key event (see Lemma 3.5) that acts as our building block in going from a lower to a higher scale. In Section 4 we use these results to prove Proposition 1.3. In Sections 5 and 6 we provide the proofs of the continuity of rate function (Proposition 1.4) and the approximate convexity result Proposition 3.4, respectively. Finally, in Section 7 we prove the stability result Proposition 2.4 to complete the argument.

## 2 Formal Definitions and Notations

Throughout the remainder of this paper we shall fix a passage time distribution $\nu$ that satisfies the hypothesis of Theorem 1; i.e., it is supported on $[0, b]$ with a continuous density function. This in particular implies that passage times are not concentrated on one point and there is no mass at 0, which in turn implies that the shape theorem (1.2) holds. For this passage time distribution and a direction vector $\mathbf{v} \in \mathbb{S}^1$, we shall denote by $\mu_{\mathbf{v}}$ the time constant in direction $\mathbf{v}$ (as in the previous section, for $\mathbf{v} = (1, 0)$ we shall drop the subscript). Under these conditions one can prove the following basic concentration estimate (see, e.g., [24]): for each $\varepsilon > 0$, $\mathbf{v} \in \mathbb{S}^1$, some $c > 0$, and all $n$ sufficiently large we have ($\lfloor n\mathbf{v} \rfloor$ is the vertex in $\mathbb{Z}^2$ obtained by taking coordinatewise integer parts of $n\mathbf{v}$):

$$\mathbb{P}(|\mathbf{PT}(\mathbf{0}, \lfloor n\mathbf{v} \rfloor) - \mu_{\mathbf{v}} n| \geq \varepsilon n) \leq e^{-cn}. \tag{2.1}$$

We shall rely on the above often, sometimes implicitly without referring to it. Notice that we are concerned with the large deviation regime, whereas (2.1) is for typical environments. To use it in the large deviation regime, we need a tool to compare the environment in the large deviation regime with the typical environment. This is provided by the FKG inequality. For a fixed $\zeta \in (0, b - \mu)$, let $\Pi = (X_e : e \in E(\mathbb{Z}^2))$ and $\Pi^{\mathscr{U}} = (X_e^{\mathscr{U}} : e \in E(\mathbb{Z}^2))$ be the typical and conditional (on $\mathscr{U}_\zeta(n)$) edge weight environments, respectively.

The following lemma is a well-known consequence of the FKG inequality and Strassen's theorem (see [28]).

LEMMA 2.1. *There exists a coupling* $(\Pi, \Pi^{\mathscr{U}})$ *such that almost surely, for each edge $e$ we have* $X_e \leq X_e^{\mathscr{U}}$.

There are two main consequences of Lemma 2.1 that will be useful for us. First, this will provide lower bounds on the FPP metric conditional on $\mathscr{U}_\zeta(n)$; second, it will enable us to restrict our attention to finite boxes. Before we proceed with the relevant statements, we extend the function $\mathbf{PT}$ from $\mathbb{Z}^2 \times \mathbb{Z}^2$ to $\mathbb{R}^2 \times \mathbb{R}^2$; this will reduce notational complexities significantly. There is not one canonical way to do this; we choose the following extension for concreteness. For every $x, y \in \mathbb{R}^2$ define $\mathbf{PT}(x, y) := \mathbf{PT}(\hat{x}, \hat{y})$ where $\hat{x}$ and $\hat{y}$ are the nearest lattice points to $x, y$, respectively (in case of a tie, we choose the one that is smallest in the usual lexicographic order on $\mathbb{Z}^2$.). We introduce some more useful notations. Throughout we will use $\mathbf{Box}(r)$ to denote the (continuous) $r \times r$ box $[-\frac{r}{2}, \frac{r}{2}]^2 \subseteq \mathbb{R}^2$.

We now proceed to show that geodesics do not wander too much even in the large deviation regime. Let $\mu_{\min} = \min_{\mathbf{v} \in \mathbb{S}^1} \mu_{\mathbf{v}}$. As a consequence of (1.2), $\mu_{\min} > 0$. Let us fix

$$\mathscr{C} = \frac{16b}{\mu_{\min}}. \tag{2.2}$$

This $\mathscr{C}$ will be important for us and will be fixed throughout the paper.

The next lemma shows that in the environment $\Pi^{\mathscr{U}}$, with high probability the FPP metric within $\mathbf{Box}(2\mathscr{C}n)$ is lower-bounded by a constant multiple of the Euclidean metric. To this end let us define the following events. Let

$$\begin{aligned}\mathscr{E}_{n,1} &:= \big\{\forall \mathbf{z}, \forall \mathbf{w} \in \mathbf{Box}(2\mathscr{C}n) \text{ such that } \|\mathbf{z}-\mathbf{w}\| \geq \sqrt{n}, \\ &\qquad \mathbf{PT}(\mathbf{z},\mathbf{w}) \geq \widetilde{\mu}\|\mathbf{z}-\mathbf{w}\|\big\}, \\ \mathscr{E}_{n,2} &:= \big\{\forall \mathbf{z} \in \mathbf{Box}(2\mathscr{C}n), \mathbf{PT}(\mathbf{z}, \mathbb{Z}^2 \setminus \mathbf{Box}(4\mathscr{C}^2 n)) \geq 2\widetilde{\mu}\mathscr{C}^2 n\big\},\end{aligned} \tag{2.3}$$

where $\widetilde{\mu} = \frac{\mu_{\min}}{2}$ and $\mathbf{PT}(\mathbf{z}, \mathbb{Z}^2 \setminus \mathbf{Box}(4\mathscr{C}^2 n))$ denotes the minimum passage time from $\mathbf{z}$ to a point in $\mathbb{Z}^2 \setminus \mathbf{Box}(4\mathscr{C}^2 n)$ (clearly this is attained at some point on the boundary of $\mathbf{Box}(4\mathscr{C}^2 n)$).

Let us define $\mathscr{E}_n := \mathscr{E}_{n,1} \cap \mathscr{E}_{n,2}$. Observe that since any path attaining

$$\mathbf{PT}(\mathbf{z}, \mathbb{Z}^2 \setminus \mathbf{Box}(4\mathscr{C}^2 n))$$

must necessarily be contained in $\mathbf{Box}(4\mathscr{C}^2 n)$, the event $\mathscr{E}_{n,2}$ is measurable with respect to the passage times in $\mathbf{Box}(4\mathscr{C}^2 n)$. Also since $\mathbf{PT}(\mathbf{z},\mathbf{w}) \leq 2b\|\mathbf{z}-\mathbf{w}\|$ for $\|\mathbf{z}-\mathbf{w}\| \geq \sqrt{n}$, by our choice of $\mathscr{C}$, on $\mathscr{E}_{n,2}$ whether or not $\mathscr{E}_{n,1}$ holds is also a function of the edge weights on $\mathbf{Box}(4\mathscr{C}^2 n)$. A similar reasoning implies that on $\mathscr{E}_{n,1} \cap \mathscr{E}_{n,2}$, $\mathbf{PT}(\mathbf{0},\mathbf{n})$ is a deterministic function of the edge weights on $\mathbf{Box}(\mathscr{C}n)$. Summarising, $\mathscr{E}_n$ is an event measurable with respect to the passage times in $\mathbf{Box}(4\mathscr{C}^2 n)$ on which the following hold:

$$\begin{gathered}\mathbf{PT}(\mathbf{0}, \mathbb{Z}^2 \setminus \mathbf{Box}(\mathscr{C}n)) \geq 4bn, \mathbf{Geo}(\mathbf{0},\mathbf{n}) \subset \mathbf{Box}(\mathscr{C}n) \\ \mathbf{Geo}(\mathbf{z},\mathbf{w}) \subset \mathbf{Box}(4\mathscr{C}^2 n) \quad \forall \mathbf{z}, \forall \mathbf{w} \in \mathbb{Z}^2 \cap \mathbf{Box}(\mathscr{C}n).\end{gathered}$$

where $\mathbf{Geo}(\mathbf{z},\mathbf{w})$ denotes the almost surely unique geodesic between the points $\mathbf{z}$ and $\mathbf{w}$.

We now show that $\mathscr{E}_n$ holds with high probability conditional on $\mathscr{U}_\zeta(n)$.

LEMMA 2.2. *There exists $c > 0$ such that for all sufficiently large $n$, for all $\zeta \in (0, b-\mu)$, with conditional (on $\mathscr{U}_\zeta(n)$) probability at least $1 - e^{-c\sqrt{n}}$, $\mathscr{E}_n$ holds.*

PROOF. By taking a union bound over all pairs of lattice points in $\mathbf{Box}(2\mathscr{C}n)$ with mutual distance at least $\sqrt{n} - 3$ and using (2.1), it follows that $\mathbb{P}(\mathscr{E}_{n,1}^c) \leq e^{-c\sqrt{n}}$. By taking a union bound over all pairs of lattice points, one of which is in $\mathbf{Box}(2\mathscr{C}_n)$ and the other is on the boundary of $\mathbf{Box}(4\mathscr{C}^2 n)$, it follows that $\mathbb{P}(\mathscr{E}_{n,2}^c) \leq e^{-cn}$. These together imply $\mathbb{P}(\mathscr{E}_n) \geq 1 - e^{-c\sqrt{n}}$ for all $n$ sufficiently large, and the proof is completed by invoking Lemma 2.1. $\square$

From now on we will restrict ourselves to $\mathbf{Box}(4\mathscr{C}^2 n)$ by defining the event

$$\mathscr{U}_\zeta^*(n) = \mathscr{U}_\zeta(n) \cap \mathscr{E}_n,$$

which by the above discussion satisfies

$$(1 - e^{-c\sqrt{n}})\mathbb{P}(\mathscr{U}_\zeta(n)) \leq \mathbb{P}(\mathscr{U}_\zeta^*(n)). \tag{2.4}$$

This allows us to work with $\mathscr{U}_\zeta^*(n)$ instead of $\mathscr{U}_\zeta(n)$, which we will often do throughout this article.

### 2.1 Gradients and stability

To precisely state the stabilization that we have alluded to, we need to develop some more notation. For our purposes, we shall be comparing distance functions for fixed directions, so we introduce the following notation. For $\mathbf{z} \in \mathbb{R}^2$ and $\theta \in \mathbb{S}^1$ (the unit circle), let $\mathbb{L}_{\theta,\mathbf{z}} = \{\mathbf{z} + \lambda\theta : \lambda > 0\}$; i.e., in the standard parametrization of $\mathbb{S}^1$, $\mathbb{L}_{\theta,\mathbf{z}}$ denotes the ray starting from $\mathbf{z}$ in the direction $\theta$. Throughout this article, we will use $\theta$ interchangeably to denote an angle or a unit vector making the corresponding angle with the $x$-axis. The usage will be clear from context and we expect it to not create confusion.Also as will be clear from context, we shall make use of polar coordinates and denote by $(\theta, \ell)$ the point on $\mathbb{L}_{\theta,\mathbf{0}}$ at distance $\ell$ from $\mathbf{0}$. We shall consider a sequence of equally spaced points along $\mathbb{L}_{\theta,\mathbf{z}}$ defined as follows. For $\mathbf{z} \in \mathbb{R}^2$, $\theta \in \mathbb{S}^1$, $k \in \mathbb{N}$, and $\ell > 0$, let us define the discrete segment

$$\mathscr{U}(\mathbf{z}, \theta, \ell, k) = [\mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_k] \tag{2.5}$$

where $\mathbf{z}_0 = \mathbf{z}$ and $\mathbf{z}_{i+1} = \mathbf{z}_i + \ell\theta$; see Figure 2.1.

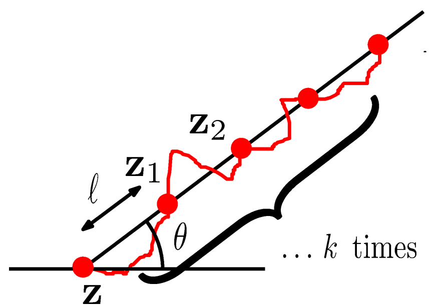


FIGURE 2.1. $k$ points spaced at distance $\ell$ along a line making angle $\theta$ with the $x$-axis forming $\mathscr{U}(\mathbf{z}, \theta, \ell, k)$.

We define the passage time for the segment $\mathscr{U}$ by

$$\mathbf{PT}(\mathbf{z}, \theta, \ell, k) := \sum_{i=0}^{k-1} \mathbf{PT}(\mathbf{z}_i, \mathbf{z}_{i+1}). \tag{2.6}$$

Note that the starting point and ending points of $\mathscr{U}(\mathbf{z}, \theta, \frac{\ell}{2}, 2k)$ and $\mathscr{U}(\mathbf{z}, \theta, \ell, k)$ are the same, and the former is obtained from the latter by subdividing subintervals of the latter intp two equal halves.

As a consequence of the triangle inequality, we have the following straightforward lemma.

LEMMA 2.3. $\mathbf{PT}(\mathbf{z}, \theta, \frac{\ell}{2}, 2k) \geq \mathbf{PT}(\mathbf{z}, \theta, \ell, k)$.

The main arguments in this paper rely on a notion of stability of the passage times from a point $\mathbf{z}$. Fix a tolerance parameter $\delta > 0$. For $k \in \mathbb{N}$, $\ell > 0$, and $\theta \in \mathbb{S}^1$, we say that $\mathbf{z} \in \mathbb{R}^2$ is $(\delta, \theta, \ell, k)$-**Stable** (with respect to any edge weight configuration $\Pi$) if for $1 \leq k' \leq k$,

$$\frac{k'\mathbf{PT}(\mathbf{z}, \theta, \ell, 1)}{(1+\delta)} \leq \mathbf{PT}(\mathbf{z}, \theta, \ell k', 1) \leq (1+\delta)k'\mathbf{PT}(\mathbf{z}, \theta, \ell, 1). \tag{2.7}$$

In words, $\mathbf{z} \in \mathbb{R}^2$ is $(\delta, \theta, \ell, k)$-**Stable** if the passage time from $z$ to $z + (\theta, \ell k')$ can be approximated up to a $(1+\delta)$ multiplicative error by $k'$ times the passage time from $z$ to $z + (\theta, \ell)$ for all $1 \leq k' \leq k$. This captures the approximately linear growth of the distance function.

In the following, for convenience, we would work with a discretized version of $\mathbb{S}^1$. For any $\eta > 0$, let

$$\mathbb{S}^1(\eta) = \{0, \eta, 2\eta, \ldots 2\pi - \eta\} \tag{2.8}$$

($2\pi/\eta$ is assumed to be an integer to avoid rounding issues). We will say that $\mathbf{z}$ is $(\delta, \mathbb{S}^1(\eta), \ell, k)$-**Stable** if $\mathbf{z}$ is $(\delta, \theta, \ell, k)$-**Stable** for each $\theta \in \mathbb{S}^1(\eta)$ and similarly we will say that $\mathbf{z}$ is $(\delta, \ell, k)$-**Stable** if $\mathbf{z}$ is $(\delta, \theta, \ell, k)$-**Stable** for each $\theta \in \mathbb{S}^1$.

With this preparation, we can now state an initial version of our stabilization result.

PROPOSITION 2.4. *Fix $\delta, \varepsilon, \eta > 0$, and $k \in \mathbb{N}$ and $J_1 \in \mathbb{N}$. There exists $J_2 \in \mathbb{N}$ such that for all large enough $n$, conditioned on $\mathscr{U}_\zeta^*(n)$ the following holds: there exists $J_1 \leq j \leq J_2$ (random depending on $\Pi \in \mathscr{U}_\zeta^*(n)$) such that*

$$\#\left\{\mathbf{z} \in \mathbb{Z}^2 \cap \mathbf{Box}(\mathscr{C}n) : \mathbf{z} \textit{ is not } \left(\delta, \mathbb{S}^1(\eta), \frac{\mathscr{C}n}{2^j}, k\right)\text{-}\mathbf{Stable}\right\} \leq \varepsilon n^2.$$

The proof of Proposition 2.4 is rather technical and is postponed until Section 7.

This is one of the three main ingredients of our proofs. Although we have stated the result in terms of the lattice points in $\mathbf{Box}(\mathscr{C}n)$, to avoid having to address rounding issues, it will be convenient to work with a version of stability for *all* points in $\mathbf{Box}(\mathscr{C}n)$. This follows from the fact that stability of a point implies stability of nearby points with possibly slightly worse parameters. The next lemma makes this precise.

LEMMA 2.5. *For $k > C > m > 0$ and a fixed $j$, the following holds on $\mathscr{E}_n$ for all $n$ sufficiently large and $\ell > \frac{n}{2^j}$: for any $\mathbf{z} \in \mathbf{Box}(\mathscr{C}n)$ that is $(\delta, \theta, \ell, k)$-**Stable** and any $\mathbf{z}'$ such that $\|\mathbf{z} - \mathbf{z}'\| \leq \ell m$, we have $\mathbf{z}'$ is $(\delta', \theta, C\ell, \frac{k}{C})$-**Stable**, where $\delta' = \delta + O(\frac{m}{C})$.*

PROOF. The proof follows by an application of the triangle inequality where we observe the following: for any $\ell'$,

$$\begin{aligned} \mathbf{PT}(\mathbf{z}, \mathbf{z} + (\theta, \ell')) &\leq \mathbf{PT}(\mathbf{z}', \mathbf{z}' + (\theta, \ell')) + O(\ell m), \\ \mathbf{PT}(\mathbf{z}', \mathbf{z}' + (\theta, \ell')) &\leq \mathbf{PT}(\mathbf{z}, \mathbf{z} + (\theta, \ell')) + O(\ell m). \end{aligned} \tag{2.9}$$

Hence for any $\ell' \geq C\ell$ as in the proof of the previous lemma, using the event $\mathscr{E}_n$, it follows that $\mathbf{PT}(\mathbf{z}, \mathbf{z} + (\theta, \ell')) = (1 + O(\frac{m}{C}))\mathbf{PT}(\mathbf{z}', \mathbf{z}' + (\theta, \ell'))$ by using the lower bounds on $\mathbf{PT}(\cdot, \cdot)$ on $\mathscr{E}_n$ (see Figure 2.2 for an illustration). □

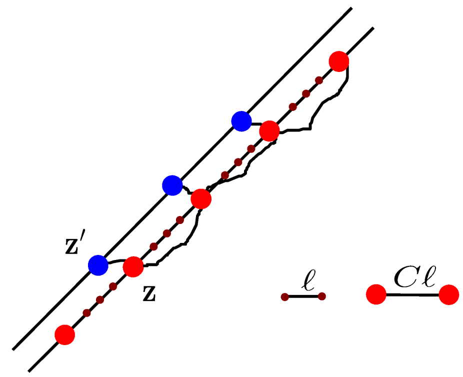


FIGURE 2.2. Stability for the discrete segment formed by the red points implies the stability for the nearby segment formed by the blue points.

From this point onwards, whenever we talk about the stability of a point, it will refer to a point in $\mathbb{R}^2$ unless explicitly mentioned otherwise.

We next define the gradient function for **Stable** points naturally in the following way: For $\theta \in \mathbb{S}^1$ and $\ell \in \mathbb{R}$, let

$$\nabla(\mathbf{z}, \theta, \ell) = \frac{\mathbf{PT}(\mathbf{z}, \mathbf{z} + (\theta, \ell))}{\ell}. \tag{2.10}$$

An easy consequence of the notion of stability is that the gradient function stays almost constant over a range of values of $\ell$.

LEMMA 2.6. *Fix $j \in \mathbb{N}$ and $\eta > 0$. On the event $\mathscr{E}_n$ from Lemma 2.2 for all sufficiently large $n$, for any $\ell \geq \frac{n}{2^j}$, and for any $(\delta, \mathbb{S}^1(\eta), \ell, k)$-**Stable** point $\mathbf{z} \in$ **Box**$(\mathscr{C}n)$, for any $\frac{k}{4}\ell \leq \ell', \ell'' \leq k\ell$ and for any $\theta \in \mathbb{S}^1$*

$$\nabla(\mathbf{z}, \theta, \ell') = \left(1 + O\left(\eta + \delta + \frac{1}{k}\right)\right)\nabla(\mathbf{z}, \theta, \ell'').$$

PROOF. The above lemma without the $O(\eta)$ term in the multiplicative factor follows immediately from the definition of stability for all $\theta$ in $\mathbb{S}^1(\eta)$. However, we need to extend this to all $\theta \in \mathbb{S}^1$, and a further approximation is necessary. For any $\theta \in \mathbb{S}^1$, let $\widehat{\theta}$ be the closest point in $\mathbb{S}^1(\eta)$. Then by the triangle inequality for any $\ell$, it follows that

$$\begin{aligned} \mathbf{PT}(\mathbf{z}, \mathbf{z} + (\theta, \ell)) &\leq \mathbf{PT}(\mathbf{z}, \mathbf{z} + (\widehat{\theta}, \ell)) + O(\eta\ell), \\ \mathbf{PT}(\mathbf{z}, \mathbf{z} + (\widehat{\theta}, \ell)) &\leq \mathbf{PT}(\mathbf{z}, \mathbf{z} + (\theta, \ell)) + O(\eta\ell), \end{aligned} \tag{2.11}$$

since the passage times on the edges are bounded (by $b$). Now, since we are on the event $\mathscr{E}_n$, the lower bounds on the terms $\mathbf{PT}(\mathbf{z}, \mathbf{z} + (\widehat{\theta}, \ell))$ and $\mathbf{PT}(\mathbf{z}, \mathbf{z} + (\theta, \ell))$ complete the proof of the lemma with the addition of the $O(\eta)$ term in the multiplicative error. $\square$

An immediate but important corollary of Lemma 2.5 is the following smoothness of the gradient field, which we state without proof.

COROLLARY 2.7. *Given $\delta$ and $\delta'$, for all large $k$ and $C$ as in Lemma 2.5, and for all large $n$ and any $\mathbf{z}, \mathbf{z}'$ satisfying the hypothesis of that lemma, and for all $\theta \in \mathbb{S}^1$,*

$$\frac{1}{1+\delta'} \leq \frac{\nabla(\mathbf{z}, \theta, \ell)}{\nabla(\mathbf{z}', \theta, C\ell)} < 1 + \delta'.$$

## 2.2 Stability of Tiles

In this subsection we introduce the notion of stability of tiles parallel to the notion of stability for points, which will be convenient for the proofs. The section contains a few lemmas which, even though quite similar to the ones already stated, have various associated quantifiers that could make it a little hard to read and the reader can choose to skip the straightforward proofs in this section. This will not affect the readability of the future sections.

Given the square $\mathbf{Box}(n)$ ($\subseteq \mathbb{R}^2$), we will often think of it as being made up of boxes of a particular scale $j$; i.e., think of the box as being naturally tiled using boxes of size $n/2^j$. Using the natural bijection between the set of tiles and the set $[\![1, 2^j]\!]^2$, we will denote the tile corresponding to $v \in [\![1, 2^j]\!]^2$ by $\mathbf{Tile}_n(j, v)$ (see Figure 2.3).

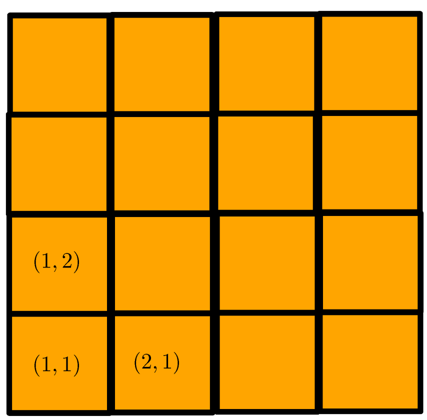


FIGURE 2.3. The figure illustrates the tiling of an $n \times n$ box into tiles of size $\frac{n}{4}$. Thus the set of tiles has a natural bijection with $[\![1, 4]\!]^2$.

DEFINITION 2.8. For any $v \in [\![1, 2^j]\!]^2$, we call a tile $\mathbf{Tile}_n(j, v)$ $(\delta, \mathbb{S}^1(\eta), \ell, k, \varepsilon)$-**Stable** if at least $1-\varepsilon$ fraction of the lattice points in $\mathbf{Tile}_n(j, v)$ are $(\delta, \mathbb{S}^1(\eta), \ell, k)$-**Stable**.

In what follows, while dealing with tiling of $\mathbf{Box}(n)$, we will choose $\ell = \frac{n}{2^{j+m}}$ and $k = 2^{2m}$, where the choice of $j$ and $m$ will vary through the paper and will depend on some other parameters relevant for specific applications. Nonetheless, importantly, they will not depend on $n$, and the reader should think of them as fixed constants and $n$ as a much larger number.

Using Lemma 2.5, we now prove that if at least $(1-\varepsilon)$ fraction of the lattice points in a $\mathbf{Tile}_n(j,v)$ are stable for some values of the parameters, then *all the points* are stable for a slightly different range of parameters.

LEMMA 2.9. *Let $j,m \in \mathbb{N}$, and $\ell = \frac{n}{2^{j+m}}, k = 2^{2m}$. Fix $\delta, \eta > 0$. There exists $C > 0$ sufficiently large such that for all sufficiently small $\varepsilon > 0$ on $\mathscr{E}_n$, the following holds for all sufficiently large $n$: If $\mathbf{Tile}_n(j,v)$ is $(\delta, \mathbb{S}^1(\eta), \ell, k, \varepsilon)$-**Stable**, then $\mathbf{Tile}_n(j,v)$ is $(2\delta, \mathbb{S}^1(\eta), \ell', k', 0)$-**Stable** where $\ell' = \max(\frac{n}{2^j} C\sqrt{\varepsilon}, \ell)$ and $k' = k\ell/\ell'$.*

PROOF. Observe that for every $\mathbf{Tile}_n(j,v)$ that is $(\delta, \mathbb{S}^1(\eta), \ell, k, \varepsilon)$-**Stable**, and any $\mathbf{z} \in \mathbf{Tile}_n(j,v)$, there exists $\mathbf{w} \in \mathbf{Tile}_n(j,v)$ with $\|\mathbf{z}-\mathbf{w}\| \le 8\sqrt{\varepsilon}\frac{n}{2^j}$ and $\mathbf{w}$ is $(\delta, \mathbb{S}^1(\eta), \ell, k)$-**Stable**. This is because the existence of a $\mathbf{z}$ for which there is no such $\mathbf{w}$ contradicts the hypothesis that $\mathbf{Tile}_n(j,v)$ is $(\delta, \mathbb{S}^1(\eta), \ell, k, \varepsilon)$-**Stable**. The proof now follows from Lemma 2.5 for $C$ sufficiently large (and $\varepsilon$ sufficiently small). $\square$

Observe that even though Lemma 2.9 refers only to the stability of all lattice points in $\mathbf{Tile}_n(j,v)$, the proof actually shows that all points in $\mathbf{Tile}_n(j,v)$ are $(\delta, \mathbb{S}^1(\eta), \ell, k)$-**Stable**. From now on we will call such a tile a $(\delta, \mathbb{S}^1(\eta), \ell, k)$-**Stable** tile. We now show that the above in fact implies stability for all angles $\theta \in \mathbb{S}^1$.

LEMMA 2.10. *Let $j,m$ be as in the previous lemma. Then on $\mathscr{E}_n$ the following holds for all sufficiently large $n$: for a $(\delta, \mathbb{S}^1(\eta), \ell, k)$-**Stable** $\mathbf{Tile}_n(j,v)$ for $\mathbf{z}, \mathbf{z}' \in \mathbf{Tile}_n(j,v)$, we have for all $\theta \in \mathbb{S}^1$,*

$$\frac{1}{1+\delta'} \le \frac{\nabla(\mathbf{z}, \theta, k_1\ell)}{\nabla(\mathbf{z}', \theta, k_2\ell)} < 1+\delta', \tag{2.12}$$

*with $\delta' = O(\delta + \eta + \frac{1}{2^m})$ and $1 \le k_1, k_2 \le k$.*

PROOF. The proof is quite similar to that of Lemma 2.9. Recalling (2.11), if for any $\theta \in \mathbb{S}^1$, $\widehat{\theta}$ is the closest point in $\mathbb{S}^1(\eta)$, then for any $k_1 \le k$,

$$|\mathbf{PT}(\mathbf{z}, \mathbf{z} + (\theta, k_1\ell)) - \mathbf{PT}(\mathbf{z}, \mathbf{z} + (\widehat{\theta}, k_1\ell))| \le O(\eta k_1 \ell),$$

which along with the hypothesis that $\mathbf{z}$ is $(\delta, \mathbb{S}^1(\eta), \ell, k)$-**Stable** implies that

$$\frac{1}{1+O(\delta+\eta)} \le \frac{\nabla(\mathbf{z}, \theta, k_1\ell)}{\nabla(\mathbf{z}, \theta, k_2\ell)} < 1 + O(\delta+\eta)).$$

Now another application of the triangle inequality as in (2.9) shows that for any $\mathbf{z}, \mathbf{z}'$ as in the statement of the lemma,

$$\mathbf{PT}(\mathbf{z}, \mathbf{z} + (\theta, k\ell)) \leq \mathbf{PT}(\mathbf{z}', \mathbf{z}' + (\theta, k\ell)) + O\left(\frac{n}{2^j}\right).$$

Hence using the fact that $k\ell = \frac{2^m n}{2^j}$, it follows that

$$\frac{1}{1 + O(\delta + \eta + \frac{1}{2^m})} \leq \frac{\nabla(\mathbf{z}, \theta, k\ell)}{\nabla(\mathbf{z}', \theta, k\ell)} < 1 + O\left(\delta + \eta + \frac{1}{2^m}\right). \qquad \square$$

Thus from now on we shall refer to a tile as $(\delta, \ell, k)$-**Stable** if (2.12) is satisfied with $\delta$ in place of $\delta'$. Now for a $(\delta, \ell, k)$-**Stable** $\mathbf{Tile}_n(j, v)$ as above, (2.12) allows us to define a gradient function not for every individual point $\mathbf{z}$ but for the whole tile itself.

DEFINITION 2.11. For a $(\delta, \ell, k)$-**Stable** $\mathbf{Tile}_n(j, v)$, define for any $\theta \in \mathbb{S}^1$,

$$\nabla_n((j, v), \theta) = \nabla_n((j, v), \theta, \ell) := \nabla(\mathbf{z}, \theta, \ell)$$

for the center point $\mathbf{z}$ of $\mathbf{Tile}_n(j, v)$.

Observe that even though this definition implicitly depends on $\ell$, we shall drop it from our notation as the length scale $\ell$ will always be clear from the context. The reason for calling $\nabla_n((j, v), \cdot)$ the gradient function for $\mathbf{Tile}_n(j, v)$ is the following: even if we replace the centre of $\mathbf{Tile}_n(j, v)$ by any arbitrary $\mathbf{z} \in \mathbf{Tile}_n(j, v)$, the value of the gradient changes only by a multiplicative factor of $(1 + \delta)$, and in all our applications, by a proper choice of parameters, $\delta$ will be made arbitrarily close to 0.

Note that on the event in Lemma 2.2, the following straightforward bounds hold:

$$\left(1 - O\left(\frac{1}{\ell}\right)\right)\widetilde{\mu} \leq \nabla_n((j, v), \theta) \leq \left(1 + O\left(\frac{1}{\ell}\right)\right) b\|(\theta, 1)\|_1, \tag{2.13}$$

where, as earlier, $(\theta, 1)$ is the unit vector in direction $\theta$. While the lower bound is essentially the content of Lemma 2.2, the upper bound follows from the fact that the edge weights are bounded by $b$. The errors arise from rounding, since passage times between real points are defined to be passage times between nearest lattice points.

With the above preparation we shall now go back to the setting of Proposition 2.4 and show that there exists a scale $j$ such that, conditional on $\mathscr{U}_\zeta^*(n)$, with probability bounded below, most of the scale $j$ tiles in $\mathbf{Box}(\mathscr{C}n)$ are stable. Recall that $\mathscr{U}_\zeta^*(n)$ was an event on $\mathbf{Box}(4\mathscr{C}^2 n)$.

LEMMA 2.12. *Condition on* $\mathscr{U}_\zeta^*(n)$. *Then given* $\eta, m, \delta, \varepsilon_1, J_1$ *such that* $\frac{1}{2^m} \geq \sqrt{\varepsilon_1}$, *there exists a constant* $J_2$ *such that for all large enough* $n$, *there exists a scale* $J_1 \leq j \leq J_2$ (*depending on* $n$) *such that with probability at least* $\frac{1}{J_2}$, *for all but* $\varepsilon_1$ *fraction of* $v \in [\![1, 2^j]\!]^2$, $\mathbf{Tile}_{\mathscr{C}n}(j, v)$ *is* $(\delta, \mathbb{S}^1(\eta), \ell, k, \varepsilon_1)$-**Stable** (*see Definition* 2.8) *where* $\ell = \frac{\mathscr{C}n}{2^{j+m}}$ *and* $k = 2^{2m}$.

PROOF. Note that from the statement of Proposition 2.4, choosing $k = 2^{4m}$ and $\varepsilon/\mathscr{C}^2 = \varepsilon_1^2$, it follows that there exists a scale $j$ such that with probability at least $\frac{1}{J_2}$ ($J_2$ appearing in the statement of Proposition 2.4) the fraction of lattice points $\mathbf{z}$ in $\mathbf{Box}(\mathscr{C}n)$ that are not $(\delta, \mathbb{S}^1(\eta), \ell, k)$-**Stable** is at most $\varepsilon n^2$ where $\ell = \frac{\mathscr{C}n}{2^{j+m}}$ and $k = 2^{2m}$. Thus the total fraction of $v \in [\![1, 2^j]\!]^2$ such that $\mathbf{Tile}_{\mathscr{C}n}(j, v)$ is not $(\delta, \mathbb{S}^1(\eta), \ell, k, \varepsilon_1)$-**Stable** is at most $\varepsilon_1$ since otherwise the total fraction of lattice points $\mathbf{z} \in \mathbf{Box}(\mathscr{C}n)$ that are not $(\delta, \mathbb{S}^1(\eta), \ell, k)$-**Stable** will be more than $\varepsilon_1^2 = \varepsilon/\mathscr{C}^2$, contradicting the conclusion of Proposition 2.4. □

The above result now implies that most of the tiles are stable (in the sense of Lemma 2.9) for an appropriate choice of parameters.

LEMMA 2.13. *Given small enough $\delta_1, \varepsilon_1 > 0$ and a positive integer $m_1$ such that $\frac{1}{2^{m_1}} \geq \varepsilon_1^{1/4}$ and $J_1 \in \mathbb{N}$, there exists $J_2$ such that for all large enough $n$, conditioned on $\mathscr{U}_\zeta^*(n)$, there exists $J_1 \leq j_1 < J_2$ (depending on $n$) such that with probability at least $\frac{1}{J_2}$, the fraction of $v \in [\![1, 2^{j_1}]\!]^2$ such that* $\mathbf{Tile}_{\mathscr{C}n}(j_1, v)$ *is not* $(\delta_1, \ell_1, k_1)$-**Stable** *is at most $\varepsilon_1$ where $\ell_1 = \frac{\mathscr{C}n}{2^{j_1+m_1}}$ and $k_1 = 2^{2m_1}$.*

PROOF. For $\eta, \delta$ sufficiently small and $m$, $J_1$ (to be chosen appropriately later depending on $\delta_1, \varepsilon_1, m_1$ as in the statement of the lemma), Lemma 2.12 implies the existence of $J_2 \in \mathbb{N}$ and $J_1 \leq j_1 \leq J_2$ such that with probability at least $\frac{1}{J_2}$ and with $\ell = \frac{\mathscr{C}n}{2^{j_1+m}}$ and $k = 2^{2m}$ for all but an $\varepsilon_1$ fraction of $v \in [\![1, 2^{j_1}]\!]^2$, $\mathbf{Tile}_{\mathscr{C}n}(j_1, v)$ are $(\delta, \mathbb{S}^1(\eta), \ell, k, \varepsilon_1)$-**Stable** (see Definition 2.8). We shall show that all of these stable tiles are also $(\delta_1, \ell_1, k_1)$-**Stable**. We now fix $m$ such that

$$\ell_1 = \frac{\mathscr{C}n}{2^{j_1+m_1}} = \max\left(\frac{\mathscr{C}n}{2^{j_1}} C\sqrt{\varepsilon}, \ell\right)$$

(where $C$ is as in Lemma 2.9) and $k_1 = 2^{2m_1} = \frac{k\ell}{\ell_1}$ (notice that such a choice is possible because we have assumed $\frac{1}{2^{m_1}} \geq \varepsilon_1^{1/4}$). Notice first that Lemma 2.9 implies that all the stable tiles above are also $(2\delta, \mathbb{S}^1(\eta), \ell', k', 0)$-**Stable**. Now applying Lemma 2.10 we conclude that each such tile is in fact $(\delta_1, \ell_1, k')$-**Stable** by choosing $\delta, \eta$ sufficiently small so that $\delta_1 = O(\delta + \eta + \frac{1}{2^{m_1}})$ as in Lemma 2.10. □

Throughout this article Lemma 2.13 will govern our choices of parameters.

## 3 Technical Preliminaries

As mentioned in our proof strategy, we shall take a configuration from the large deviation regime at some length scale $n$ and replicate/dilate the same configuration to obtain a configuration at a larger length scale. The obvious problem one notices is that for continuous passage time distributions, each configuration has probability 0. Hence to carry out our proof strategy, we will not be able to work with the edge weight configurations directly. We will project it to a discrete set of $e^{o(n^2)}$ many elements and pick the most likely one among them (still in the large deviation regime). We describe below the discretization we shall employ.

Since by the upper bound on passage times on edges, deterministically,

$$\mathbf{PT}(\mathbf{x},\mathbf{y}) \le b\|\mathbf{x}-\mathbf{y}\|_1 + 3b \tag{3.1}$$

for any $\mathbf{x},\mathbf{y} \in \mathbb{R}^2$ (the additive constant is needed to take care of rounding errors as $\mathbf{x},\mathbf{y} \in \mathbb{R}^2$ and are not points in $\mathbb{Z}^2$). For a discretization parameter $\eta_1$, we will round the normalized distances (passage time divided by Euclidean distance) to be in the set $\{0,\eta_1,2\eta_1,\ldots,3b\}$ (without loss of generality assuming $3b/\eta_1 \in \mathbb{N}$), and project the distance functions $\mathbf{PT}(\cdot,\cdot)$ onto a discrete space accordingly. We shall use this only for $\mathbf{x},\mathbf{y}$ sufficiently far apart so that the normalized passage times will always be upper-bounded by $3b$ using (3.1).

To define things formally, first let the set of all points in $\mathbf{Box}(n)\cap\frac{n}{2^j}\mathbb{Z}^2$ be called $\mathbf{Grid}_n(j)$. We will also need the following variant: Let $\ell = \frac{n}{2^{j+m}}$ for some some $m \in \mathbb{N}$. By $\mathbf{Grid}_n(\ell; j)$, we shall denote the set of all points in $\mathbf{Grid}_n(j+m)$ that lie on the line segments joining the nearest neighbours in $\mathbf{Grid}_n(j)$ as elements of $\frac{n}{2^j}\mathbb{Z}^2$ (see Figure 3.1). Observe that the number of vertices in $\mathbf{Grid}_n(j)$ depends only on $j$, whereas the number of vertices in $\mathbf{Grid}_n(\ell; j)$ depends on $j$ and $m$; in particular, as the parameters $j$ and $m$ will never depend on $n$, these numbers will not depend on $n$ either.

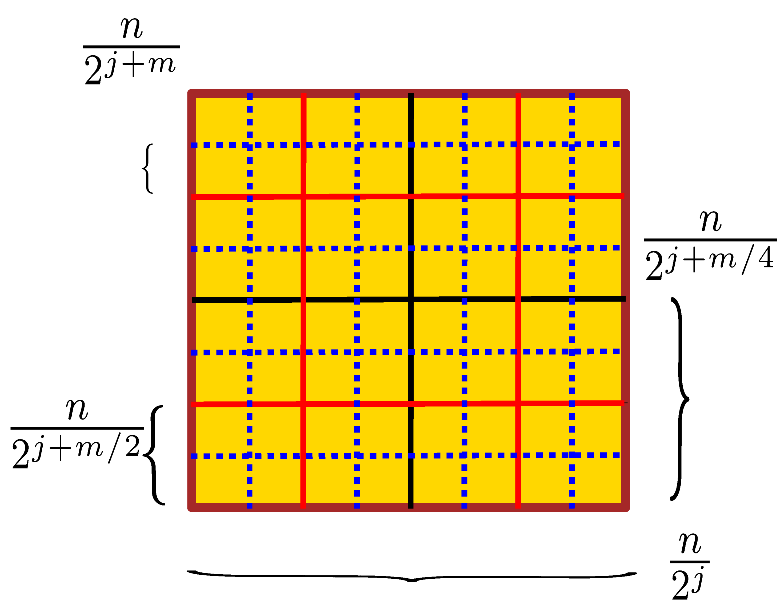


FIGURE 3.1. Figure illustrating the various grid points, with spacing between the grid lines indicated as well.

Now given $\eta_1, m_1, j_1$, let the projection map

$$\overset{\eta_1, j_1+m_1}{\mathbf{Proj}} : \mathbf{Grid}_{\mathscr{C}n}(j_1+m_1) \times \mathbf{Grid}_{\mathscr{C}n}(j_1+m_1) \to \mathbb{R}_+$$

be defined as follows: for any $\mathbf{z}, \mathbf{w} \in \mathbf{Grid}_{\mathscr{C}n}(j_1+m_1)$, with

$$\overset{\eta_1, j_1+m_1}{\mathbf{Proj}}(\mathbf{z}, \mathbf{w}) = \eta_1 \left\lfloor \frac{\mathbf{PT}(\mathbf{z}, \mathbf{w})}{\eta_1 \|\mathbf{z}-\mathbf{w}\|} \right\rfloor \|\mathbf{z}-\mathbf{w}\|. \tag{3.2}$$

when $\mathbf{z} \neq \mathbf{w}$; otherwise we define it to be 0.

Observe that the function $\overset{\eta_1, j_1+m_1}{\mathbf{Proj}}$ is random, but we choose to suppress the dependence on the underlying noise for brevity. Since these will be the only parameters we will use, we will also drop the dependence on $\eta_1$, $j_1$, and $m_1$ in the notation. Note that $\mathbf{Proj}$ induces a weighted graph with vertex set $\mathbf{Grid}_{\mathscr{C}n}(j_1+m_1)$, with the weight on any edge $(\mathbf{z}, \mathbf{w})$ being $\mathbf{Proj}(\mathbf{z}, \mathbf{w})$. Let $\mathscr{P}\mathscr{V}_{\eta_1, j_1+m_1}$ denote the set of all possible such graphs induced by $\mathbf{Proj}$ (as the weight configuration varies). Observe that a very basic counting argument yields that the cardinality of $\mathscr{P}\mathscr{V}_{\eta_1, j_1+m_1}$ satisfies

$$|\mathscr{P}\mathscr{V}_{\eta_1, j_1+m_1}| \leq e^{O(2^{2(j_1+m_1)}) \log \frac{1}{\eta_1}} \tag{3.3}$$

and in particular is independent of $n$. It will also be useful to extend the definition of $\mathbf{Proj}$ to a larger set of pairs. For all pairs of points $\mathbf{z}, \mathbf{w} \in \mathbf{Box}(\mathscr{C}n)$ we will extend the definition by letting $\mathbf{Proj}(\mathbf{z}, \mathbf{w}) = \mathbf{Proj}(\widetilde{\mathbf{z}}, \widetilde{\mathbf{w}})$, where $\widetilde{\mathbf{z}}, \widetilde{\mathbf{w}}$ are the nearest points to $\mathbf{z}, \mathbf{w}$, respectively, in $\mathbf{Grid}_{\mathscr{C}n}(j_1+m_1)$ (as before, breaking ties by picking the smallest in the lexicographic order). Note that if $\mathbf{z}$ and $\mathbf{w}$ get rounded to the same point, then $\mathbf{Proj}(\mathbf{z}, \mathbf{w})$ is 0, which is not a realistic definition. However, we will only be interested in pairs $\mathbf{z}$ and $\mathbf{w}$ that are reasonably far apart, so the above issue will not arise and hence we will not bother with this aspect of the definition.

The first thing we show now is that the error introduced by using $\mathbf{Proj}(\cdot, \cdot)$ instead of $\mathbf{PT}(\cdot, \cdot)$ can be neglected at sufficiently large length scales. For reasons that will become clear, we shall work with **Stable** tiles, although the approximation is valid independent of that. Fix $\delta_1, \varepsilon_1, m_1, J_1$ as in Lemma 2.13, which then guarantees that there exists $j_1$ with probability bounded away from 0 such that for all but $\varepsilon_1$ fraction of $v \in [\![1, 2^{j_1}]\!]^2$, $\mathbf{Tile}_{\mathscr{C}n}(j_1, v)$ is $(\delta_1, \ell_1, k_1)$-**Stable** where $\ell_1$ and $k_1$ are $\frac{\mathscr{C}n}{2^{j_1+m_1}}$ and $2^{2m_1}$, respectively. Let $m_1$ also be even. For later reference, let us call $v \in [\![1, 2^{j_1}]\!]^2$ to be $(\delta_1, \ell_1, k_1)$-**Stable** or $(\delta_1, \ell_1, k_1)$-**Unstable** depending on whether $\mathbf{Tile}_{\mathscr{C}n}(j_1, v)$ is $(\delta_1, \ell_1, k_1)$-**Stable** or not, respectively.

LEMMA 3.1. *Fix $\delta_1, \eta_1, j_1, m_1$ and accordingly $\ell_1$ and $k_1$ as above. Then conditioned on $\mathscr{U}^*_\zeta(n)$, consider $v \in [\![1, 2^{j_1}]\!]^2$ such that* $\mathbf{Tile}_{\mathscr{C}n}(j_1, v)$ *is $(\delta_1, \ell_1, k_1)$-***Stable**. *Then for any $\mathbf{z} \neq \mathbf{w} \in \mathbf{Grid}_{\mathscr{C}n}(j_1+m_1/2)$ such that $\mathbf{z}, \mathbf{w} \in \mathbf{Tile}_{\mathscr{C}n}(j_1, v)$, we have the following*:

$$1 \leq \frac{\mathbf{PT}(\mathbf{z}, \mathbf{w})}{\mathbf{Proj}(\mathbf{z}, \mathbf{w})} \leq 1 + O(\eta_1).$$

Note that $\mathbf{Grid}_{\mathscr{C}n}(j_1 + m_1/2) \subset \mathbf{Grid}_{\mathscr{C}n}(j_1 + m_1)$ when $m_1$ is even (in fact, to avoid dealing with such divisibility issues further, we will henceforth assume without loss of generality that $m_1$ is divisible by 16).

PROOF. Observe that by definition for all $\mathbf{z}, \mathbf{w}$,

$$\mathbf{Proj}(\mathbf{z}, \mathbf{w}) \leq \mathbf{PT}(\mathbf{z}, \mathbf{w}) \leq \mathbf{Proj}(\mathbf{z}, \mathbf{w}) + O(\eta_1)\|\mathbf{z} - \mathbf{w}\|. \tag{3.4}$$

The proof now follows immediately by noticing that since $\mathbf{z}$ and $\mathbf{w}$ are at distance apart of at least $\frac{\mathscr{C}n}{2^{j_1+m_1/2}}$ and, on $\mathscr{U}_\zeta^*(n)$, by definition

$$\mathbf{PT}(\mathbf{z}, \mathbf{w}) \geq \widetilde{\mu}\|\mathbf{z} - \mathbf{w}\|. \tag{3.5}$$

□

We now define a gradient function corresponding to the projected distances analogous to (2.10). As in the above setting, let $\mathbf{z}, \mathbf{w} \in \mathbf{Grid}_{\mathscr{C}n}(j_1 + m_1/2)$, and let $\theta$ and $d > 0$ be such that $\mathbf{w} = \mathbf{z} + (\theta, d)$. Then let

$$\nabla_{\mathbf{Proj}}(\mathbf{z}, \theta, d) = \frac{\mathbf{Proj}(\mathbf{z}, \mathbf{w})}{\|\mathbf{z} - \mathbf{w}\|}. \tag{3.6}$$

Defining the projected gradients only for pairs of points in $\mathbf{Grid}_{\mathscr{C}n}(j_1 + m_1/2)$, we then define projected gradients in all directions at a slightly coarser scale, i.e., for all points in $\mathbf{Grid}_{\mathscr{C}n}(j_1 + m_1/4)$. For any $\mathbf{z} \in \mathbf{Tile}_{\mathscr{C}n}(j_1, v) \cap \mathbf{Grid}_{\mathscr{C}n}(j_1 + m_1/4)$ and for any $\theta \in \mathbb{S}^1$ and

$$\frac{\mathscr{C}n}{2^{j_1+m_1/4}} < d < \frac{\mathscr{C}n}{2^{j_1}} 2^{m_1/4}, \tag{3.7}$$

let

$$\nabla_{\mathbf{Proj}}(\mathbf{z}, \theta, d) = \frac{\mathbf{Proj}(\mathbf{z}, \mathbf{w})}{d}, \tag{3.8}$$

where $\mathbf{w}$ is the closest point to $\mathbf{z} + (\theta, d)$ in $\mathbf{Grid}_{\mathscr{C}n}(j_1 + m_1/2)$. Note that $\mathbf{Grid}_{\mathscr{C}n}(j_1 + m_1/4) \subset \mathbf{Grid}_{\mathscr{C}n}(j_1 + m_1/2)$.

Thus, with $\delta_1, \ell_1, k_1$ as above, if $\mathbf{Tile}_{\mathscr{C}n}(j_1, v)$ is $(\delta_1, \ell_1, k_1)$-**Stable**, then the following result about smoothness of the projected gradient field follows as in (2.12) and (2.11); we omit the proof.

LEMMA 3.2. *Fix $\eta_1 > 0$ as in Lemma 3.1. For a $(\delta_1, \ell_1, k_1)$-**Stable** $\mathbf{Tile}_{\mathscr{C}n}(j_1, v)$ and for any $\mathbf{z}, \mathbf{z}' \in \mathbf{Tile}_{\mathscr{C}n}(j_1, v)$ and $\theta_1, \theta_2 \in \mathbb{S}^1$ such that $\|\theta_1 - \theta_2\| \leq \eta_1$ and $d_1, d_2$ satisfy (3.7), with $\nabla_{\mathbf{Proj}}(\mathbf{z}, \theta_1, d_1)$ and $\nabla_{\mathbf{Proj}}(\mathbf{z}', \theta_2, d_2)$ defined via (3.8), we have*

$$\frac{1}{1 + O(\delta_1 + \eta_1 + 2^{-m_1/4})} \leq \frac{\nabla_{\mathbf{Proj}}(\mathbf{z}, \theta_1, d_1)}{\nabla_{\mathbf{Proj}}(\mathbf{z}', \theta_2, d_2)} < 1 + O(\delta_1 + \eta_1 + 2^{-m_1/4}).$$

Note that above we choose $\|\theta_1 - \theta_2\| \leq \eta_1$ where the latter appeared in the definition of **Proj**. This is done deliberately to avoid introducing new notation since for us any small enough value of $\eta_1$ serves both the purposes.

This allows us to define a projected gradient for the entire tile as we did in Definition 2.11.

DEFINITION 3.3. If $\mathbf{Tile}_{\mathscr{C}n}(j_1, v)$ is $(\delta_1, \ell_1, k_1)$-**Stable**, then let

$$\nabla_{\mathbf{Proj}}((j_1, v), \theta) := \nabla_{\mathbf{Proj}}(\mathbf{z}, \theta, d)$$

where $\mathbf{z}$ is the center point of $\mathbf{Tile}_n(j, v)$ and $d = \frac{n}{2^{j_1+m_1/8}}$.

Observe that if in the above definition we had chosen some arbitrary $\mathbf{z}$ contained in $\mathbf{Tile}_n(j_1, v) \cap \mathbf{Grid}_{\mathscr{C}n}(j_1 + m_1/2)$ and $d$ such that the RHS is defined via (3.8), then the definition would change only by a multiplicative factor of $(1 + O(\delta_1 + \frac{1}{2^{m_1/4}}))$. In our applications, the multiplicative error will be made suitably close to 1 by choosing the parameters appropriately.

We now move towards our second main technical ingredient. Note that the convexity of the limit shape $\mathcal{B}$ in (1.2) is essentially due to **PT** satisfying the triangle inequality (by definition). We shall establish an analogous approximate convexity statement corresponding to **Proj**. To formally state things, it would be convenient to consider the following function on all of $\mathbb{R}^2$: for any $\mathbf{w} = (\theta, r)$,

$$\|\mathbf{w}\|_{(j,v)} \equiv \|(\theta, r)\|_{(j,v)} := r\nabla_{\mathbf{Proj}}((j, v), \theta).$$

Note that as in Definition 3.3, this definition implicitly depends on the choice of $\mathbf{z}$ and $d$. We record the following consequence of the above definitions. If $\mathbf{Tile}_{\mathscr{C}n}(j_1, v)$ is $(\delta_1, \ell_1, k_1)$-**Stable**, then for any $\mathbf{w}_1, \mathbf{w}_2 \in \mathbf{Tile}_{\mathscr{C}n}(j_1, v)$, with $\|\mathbf{w}_1 - \mathbf{w}_2\| \ge \frac{\mathscr{C}n}{2^{j_1+m_1/4}}$,

$$\big(1 - O(\delta_1 + 2^{-\frac{m_1}{4}})\big) \le \frac{\mathbf{PT}(\mathbf{w}_1, \mathbf{w}_2)}{\|\mathbf{w}_1 - \mathbf{w}_2\|_{j,v}} \le \big(1 + O(\delta_1 + 2^{-\frac{m_1}{4}})\big). \tag{3.9}$$

The following fact analogous to (2.13) will be useful as well.

$$\begin{aligned}\left(1 - O\left(\delta_1 + \frac{1}{2^{m_1/4}}\right)\right)\widetilde{\mu} &\le \nabla_{\mathbf{Proj}}((j, v), \theta) \\ &\le \left(1 + O\left(\delta_1 + \frac{1}{2^{m_1/4}}\right)b\right)\|(\theta, 1)\|_1.\end{aligned} \tag{3.10}$$

The next lemma shows the approximate convexity of the above-defined function that allows us to think of the above as roughly a norm.

PROPOSITION 3.4. *If* $\mathbf{Tile}_{\mathscr{C}n}(j_1, v)$ *is* $(\delta_1, \ell_1, k_1)$-**Stable***, then for any set of vectors* $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_t$*, if* $\mathbf{w} = \sum_{i=1}^t \mathbf{w}_i$*, then*

$$\|\mathbf{w}\|_{(j_1,v)} \le \big(1 + O\big(\delta_1 + 2^{-\frac{m_1}{16}}\big)\big)\left(\sum_{i=1}^t \|\mathbf{w}_i\|_{(j_1,v)}\right).$$

The proof relying on an approximate triangle inequality is technical and is postponed to Section 6.

The next and final result in this section will show the existence of an event that approximates the large deviation event in the log scale and will be used as the building block of our dilation construction in the next section. Given $\delta_1$, $\varepsilon_1$, $m_1$, and $J_1$ satisfying the hypothesis of Lemma 2.13, let $j_1$ be the scale obtained from that lemma and recall the definitions of $\ell_1$ and $k_1$ from the statement of the same. Recall

$$\eta_1 \quad \text{and} \quad \mathbf{Proj} = \overset{\eta_1, j_1+m_1}{\mathbf{Proj}}$$

from Lemma 3.1. Recalling $\mathscr{PV}_{\eta_1, j_1+m_1}$, the set of weighted graphs induced on $\mathbf{Grid}_{\mathscr{C}n}(j_1+m_1)$, let $\mathscr{PV}_{\eta_1, j_1+m_1/2}$ be the induced graphs on the vertex set $\mathbf{Grid}_{\mathscr{C}n}(j_1+m_1/2) \subset \mathbf{Grid}_{\mathscr{C}n}(j_1+m_1)$. By an abuse of notation, we shall denote by $\mathbf{Proj}^{-1}(\Im)$, for $\Im \in \mathscr{PV}_{\eta_1, j_1+m_1/2}$, the set of all weight configurations for which $\Im$ is the weighted graph induced by $\mathbf{Proj}$. Now for $\Im \in \mathscr{PV}_{\eta_1, j_1+m_1/2}$ and $A \subset [\![1, 2^{j_1}]\!]^2$, let us define the key event

$$\text{(3.11)} \quad \mathbf{Base\text{-}event}(\eta_1, \delta_1, j_1, m_1, \varepsilon_1, A, \Im) :=$$
$$\mathscr{U}^*_\zeta(n) \cap \mathbf{Proj}^{-1}(\Im) \cap \big\{\{v \in [\![1, 2^{j_1}]\!]^2 : v \text{ is } (\delta_1, \ell_1, k_1) - \mathbf{Unstable}\} \subset A\big\}.$$

The following simple lemma based on the pigeonhole principle lower-bounds the probability of the above event.

LEMMA 3.5. *Given $\varepsilon_4 > 0$ and the parameters as above, i.e., $\delta_1$, $\varepsilon_1$, $m_1$, and $j_1$, there exists $\Im \in \mathscr{PV}_{\eta_1, j_1+m_1/2}$, and $A \subset [\![1, 2^{j_1}]\!]^2$ with $|A| = \varepsilon_1 2^{2j_1}$ such that*

$$\frac{\log\big(\mathbb{P}(\mathbf{Base\text{-}event}(\eta_1, \delta_1, j_1, m_1, \varepsilon_1, A, \Im))\big)}{n^2} \geq \kappa - \varepsilon_4$$

*for all large enough $n$, where $\kappa = \kappa_n := \frac{\log \mathbb{P}(\mathscr{U}^*_\zeta(n))}{n^2}$.*

Note that $\varepsilon_4$ is independent of the remaining parameters.

PROOF. By our choice of parameters, it follows from Lemma 2.13 that for any $\varepsilon_4 > 0$, for all $n$ sufficiently large,

$$\frac{\log\big(\mathbb{P}(\mathscr{U}^*_\zeta(n) \cap \#\big\{\{v \in [\![1, 2^{j_1}]\!]^2 : v \text{ is } (\delta_1, \ell_1, k_1) - \mathbf{Unstable}\} \leq \varepsilon_1 2^{2j_1}\big\})\big)}{n^2}$$
$$\geq \kappa - \frac{\varepsilon_4}{2}.$$

Recall now the trivial bound mentioned in (3.3),

$$|\mathscr{PV}_{\eta_1, j_1+m_1/2}| = e^{O(2^{2(j_1+m_1)} \log \frac{1}{\eta_1})} = e^{O(1)}.$$

Moreover, the number of possible subsets $A$ of $[\![1, 2^{j_1}]\!]^2$ of size at most $\varepsilon_1 2^{2j_1}$ is at most $e^{O(H(\varepsilon_1))2^{2j_1}}$, where $H(\cdot)$ is the entropy functional. Thus by the pigeonhole principle the result follows. □

Observe that **Base-event** is measurable with respect to the edges in $\mathbf{Box}(4\mathscr{C}^2 n)$ and, as already mentioned, this will be the building block in our constructions in the next section. We end this section by pointing out that although the definition of **Base-event** depends on the set $A$ as well as the element $\Im \in \mathscr{PV}_{\eta_1, j_1+m_1/2}$, in what follows we shall often refer to **Base-event** given the parameters $\eta_1$, $\delta_1$, $j_1$, $m_1$, $\varepsilon_1$, and $A$ and $\Im$ would then be understood as given by Lemma 3.5.

# 4 Constructing a Large Deviation Event at a Higher Scale

In this section we prove Proposition 1.3. With the definitions and results from the previous section at our disposal, we now follow the strategy outlined in Section 1.3. Given arbitrary small positive constants $\varepsilon, \varepsilon'$, and $n$, for any $n_1$ large enough, we will create the favourable event $\mathbf{Fav} := \mathbf{Fav}(n_1)$, which will imply $\mathscr{U}_{\zeta'}(n_1)$ where $\zeta' \geq \zeta - \varepsilon'$, and

$$\frac{\log \mathbb{P}(\mathbf{Fav}(n_1))}{n_1^2} \geq \frac{\log \mathbb{P}(\mathscr{U}_\zeta(n))}{n^2} - \varepsilon.$$

Note that (1.5) allows us to assume that $n_1$ is divisible by $n$, which is what we will assume throughout this section.

We start by defining certain key ingredients: Fixing $\varepsilon_6 > 0$ (to be chosen appropriately later) and recalling the constant $\mathscr{C}$ from (2.2), for brevity we adopt the following abbreviations

$$\begin{gathered} \mathfrak{n}_0 := \mathscr{C} n_1(1+2\varepsilon_6), \quad \mathfrak{n}_1 := \mathscr{C} n_1(1+\varepsilon_6), \quad \mathfrak{n}_2 := \mathscr{C} n_1, \\ \mathfrak{n}_3 := \mathscr{C} n(1+\varepsilon_6), \quad \mathfrak{n}_4 := \mathscr{C} n. \end{gathered} \tag{4.1}$$

Moreover, in what follows we will denote $\mathbf{Box}(\mathfrak{n}_i)$ as $\mathfrak{B}_i$. We will often identify each such box with the set of all the lattice edges contained in its closure, i.e., including the edges along the boundary thought of as a subset of $\mathbb{R}^2$. Consequently, throughout the discussion we will say a path is contained in a box (or more generally in a subset of edges) if all of its constituent edges are.

## 4.1 Construction of Fav

**Fav** will be measurable with respect to the edge weights in $\mathfrak{B}_0$, with the property that on the event **Fav**,

$$\mathbf{PT}_{\mathfrak{B}_0}(\mathbf{0}, \mathfrak{B}_0^c) \geq 2bn_1 \quad \text{and} \quad \mathbf{PT}_{\mathfrak{B}_0}(\mathbf{0}, \mathbf{n_1}) \geq (\mu + \zeta - \varepsilon')n_1 \tag{4.2}$$

for some small $\varepsilon'$ where $\mathbf{PT}_{\mathfrak{B}_0}(\cdot,\cdot)$ denotes the passage time between points restricted to $\mathfrak{B}_0$; i.e., one only considers paths that do not exit $\mathfrak{B}_0$. Clearly this implies that $\mathbf{Fav} \subset \mathscr{U}_{\zeta'}(n_1)$ for $\zeta' = \zeta - \varepsilon'$.

Throughout this section we will work with parameters as in Lemma 3.5, i.e., $\delta_1, \varepsilon_1, \varepsilon_4, j_1, m_1$, and $\ell_1 = \frac{n}{2^{j_1+m_1}}$ and $k_1 = 2^{2m_1}$. Further, we have already introduced a parameter $\varepsilon_6$ in (4.1). We will introduce another parameter $\varepsilon_7$ in the definition of **Fav**.

The basic objects we will be working with are the following. Tile the box $\mathfrak{B}_0$ by $\mathbf{Tile}_{\mathfrak{n}_0}(j_1, v)$ for $v \in [\![1, 2^{j_1}]\!]^2$. Now each such tile is a square of size $\frac{\mathfrak{n}_0}{2^{j_1}}$. For $v \in [\![1, 2^{j_1}]\!]^2$, consider the square with the same centre as $\mathbf{Tile}_{\mathfrak{n}_0}(j_1, v)$ and side length $\frac{\mathfrak{n}_1}{2^{j_1}}$. Call this square (closed) $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$; see Figure 4.1. It follows that neighbouring $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$'s are separated by vertical and horizontal strips of width $\varepsilon_6 \frac{\mathfrak{n}_1}{2^{j_1}}$. We will call the set of all edges in $\mathfrak{B}_0$ that do not belong to any $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$

as a $\mathbf{Corridor}^{\mathrm{ext}}(j_1, \mathfrak{n}_0)$ (“ext” stands for exterior; we will also consider corridors inside $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$, which will be defined shortly).

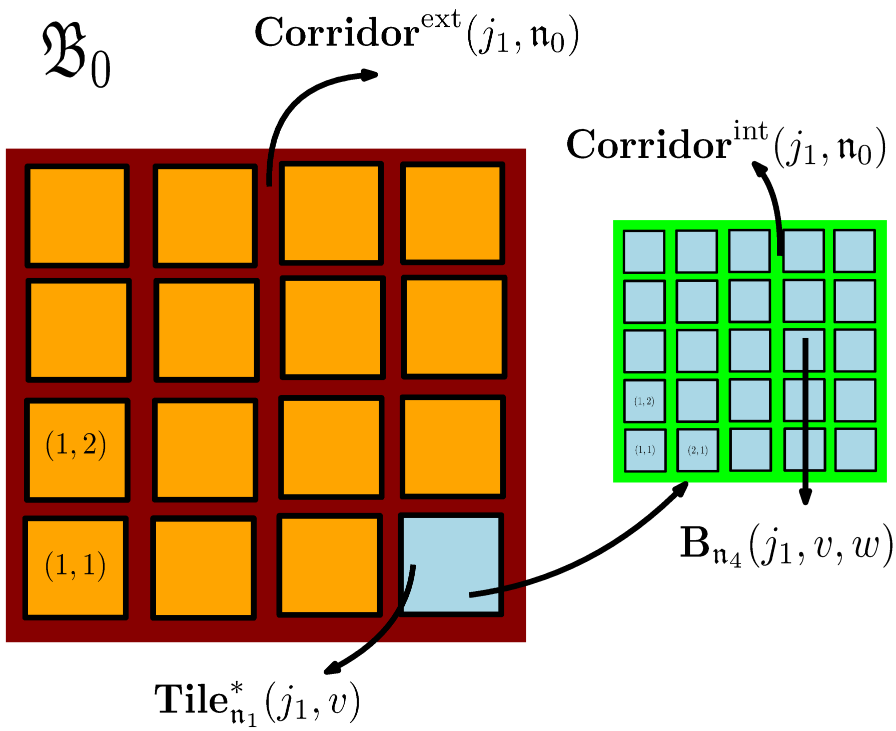


FIGURE 4.1. The figure illustrates the basic structural definitions inside $\mathfrak{B}_0$. On the LHS the figure shows the $\mathbf{Corridor}^{\mathrm{ext}}(j_1, \mathfrak{n}_0)$ (red region) and the tiling of the remaining area by $\mathbf{Tile}^*_{\mathfrak{n}_1}(j, v)$ for $v \in [\![1, 2^{j_1}]\!]^2$, (the associated $v$ to a tile has been indicated). The RHS zooms into one particular $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$, namely $v = (4, 1)$. $\mathbf{B}_{\mathfrak{n}_4}(j_1, v, w)$ and the surrounding $\mathbf{C}_{\mathfrak{n}_4}(j_1, v, w)$, which form a part of $\mathbf{Corridor}^{\mathrm{int}}(j_1, \mathfrak{n}_0)$, are shown with the corresponding $w$'s indicated.

Our construction of **Fav** will have two steps:

(i) specifying the environment inside $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$ for $v \in [\![1, 2^{j_1}]\!]^2$, and
(ii) specifying the environment in $\mathbf{Corridor}^{\mathrm{ext}}(j_1, \mathfrak{n}_0)$.

Part (i) involves a large deviation environment in the smaller scale $n$, whereas for the second part we make all the edge weights close to $b$. We shall formalize part (i) later, but for now let us make part (ii) precise as follows. Let $\mathbf{Barrier}^{\mathrm{ext}}(\mathfrak{n}_0, j_1)$ denote the event that the passage time on each edge in $\mathbf{Corridor}^{\mathrm{ext}}(j_1, \mathfrak{n}_0)$ is in $[b - \varepsilon_7, b]$ for some small but fixed $\varepsilon_7$. Observe that by our assumption on the edge weight distribution $\nu$ from Definition 1.1, $\nu([b - \varepsilon_7, b]) > 0$ for all $\varepsilon_7 > 0$.

As the total number of edges in $\mathbf{Corridor}^{\mathrm{ext}}(j_1, \mathfrak{n}_0)$ is $O(\varepsilon_6 \mathfrak{n}_0^2)$, it follows that

$$-\log \mathbb{P}(\mathbf{Barrier}^{\mathrm{ext}}(\mathfrak{n}_0, j_1)) = O_{\varepsilon_7}(\varepsilon_6 n_1^2) \tag{4.3}$$

(the constant in the $O_{\varepsilon_7}(\cdot)$ notation depends on $\varepsilon_7$, and $\varepsilon_6$ will be chosen to be much smaller than $\varepsilon_7$ depending on the edge weight distribution $\nu$).

Having constructed $\mathbf{Barrier}^{\mathrm{ext}}(\mathfrak{n}_0, j_1)$, we are left to do two more things:

(1) specifying the environments inside $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$ using the **Base-event** defined in Lemma 3.5, and
(2) verifying the two properties listed in (4.2): recall that it involves showing that any path $\gamma$ between $\mathbf{0}$ and $\mathbf{n_1}$ contained in $\mathfrak{B}_0$ and any path from $\mathbf{0}$ to $\mathfrak{B}_0^c$ has lengths at least $(\mu + \zeta - \varepsilon')n_1$ and $2bn_1$, respectively.

For $v \in [\![1, 2^{j_1}]\!]^2$, to specify the environments inside $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$, it will be convenient to think of each $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$ as naturally made up of $(\frac{n_1}{n})^2$ copies of $\mathbf{Tile}_{\mathfrak{n}_3}(j_1, v)$, which we will denote as $\mathbf{A}_{\mathfrak{n}_3}(j_1, v, w)$ for $w \in [\![1, \frac{n_1}{n}]\!]^2$. As before, each $\mathbf{A}_{\mathfrak{n}_3}(j_1, v, w)$ can be thought of as a copy of $\mathbf{Tile}_{\mathfrak{n}_4}(j_1, v)$ (to be called $\mathbf{B}_{\mathfrak{n}_4}(j_1, v, w)$) surrounded by an annulus $\mathbf{C}_{\mathfrak{n}_4}(j_1, v, w)$ of width $\frac{\varepsilon_6}{2}\frac{\mathfrak{n}_4}{2^{j_1}}$ (see Figure 4.1). We denote the union of edges in $\mathbf{C}_{\mathfrak{n}_4}(j_1, v, w)$ (union over $v \in [\![1, 2^{j_1}]\!]^2$ and $w \in [\![1, \frac{n_1}{n}]\!]^2$) as $\mathbf{Corridor}^{\mathrm{int}}(j_1, \mathfrak{n}_0)$. As before, only the edges that are not contained in the closure of any $\mathbf{B}_{\mathfrak{n}_4}(j_1, v, w)$ will be counted in $\mathbf{Corridor}^{\mathrm{int}}(j_1, \mathfrak{n}_0)$. Similarly to $\mathbf{Barrier}^{\mathrm{ext}}(\mathfrak{n}_0, j_1)$, let $\mathbf{Barrier}^{\mathrm{int}}(\mathfrak{n}_0, j_1)$ denote the event that the passage time on each edge in $\mathbf{Corridor}^{\mathrm{int}}(j_1, \mathfrak{n}_0)$ is in $[b - \varepsilon_7, b]$, and as in (4.3), we have

$$-\log \mathbb{P}(\mathbf{Barrier}^{\mathrm{int}}(\mathfrak{n}_0, j_1)) = O_{\varepsilon_7}(\varepsilon_6 n_1^2). \tag{4.4}$$

We will use **Barrier** to denote the intersection of the events $\mathbf{Barrier}^{\mathrm{ext}}(\mathfrak{n}_0, j_1)$ and $\mathbf{Barrier}^{\mathrm{int}}(\mathfrak{n}_0, j_1)$.

We are now left with the task of prescribing the environment inside $\mathbf{B}_{\mathfrak{n}_4}(j_1, v, w)$. However, before formally doing that, we address a rounding issue. Note that there is a natural identification between the continuous boxes $\mathbf{B}_{\mathfrak{n}_0}(j_1, v)$ and $\mathbf{Tile}_{\mathfrak{n}_0}(j_1, v)$ and similarly between $\mathbf{B}_{\mathfrak{n}_4}(j_1, v, w)$ and $\mathbf{Tile}_{\mathfrak{n}_4}(j_1, v)$, since each member of the pair is a translate of the other. However, there might be microscopic discrepancies in their intersections with $\mathbb{Z}^2$. For example, $\mathbb{Z}^2 \cap \mathbf{B}_{\mathfrak{n}_4}(j_1, v, w)$ might not be identifiable with $\mathbb{Z}^2 \cap \mathbf{Tile}_{\mathfrak{n}_4}(j_1, v)$. This discrepancy is very minor and can be handled in a number of ways. One would be to translate $\mathbf{B}_{\mathfrak{n}_4}(j_1, v, w)$'s by a distance $\leq 2$ (this will lead to local changes in the width of the corridors by $O(1)$ and will not affect any of our arguments) to allow an exact identification to hold even at the lattice level. For the sake of exposition and to avoid introducing new notation to handle this trivial issue completely precisely, we will be ignoring this and similar rounding issues throughout our discussion, and assume that for each $v$, $\mathbf{Tile}_{\mathfrak{n}_4}(j_1, v)$, and additionally for each $w$, $\mathbf{B}_{\mathfrak{n}_4}(j_1, v, w)$, is a closed lattice box, i.e., of the form $[\![a, b]\!] \times [\![c, d]\!]$ for some integers $a, b, c, d$.

Towards prescribing the environment inside $\mathbf{B}_{\mathrm{n}_4}(j_1, v, w)$, recall the set $A$ of size $\varepsilon_1 2^{2j_1}$ in the statement of Lemma 3.5. For $v \in A$ and $v \in [\![1, 2^{j_1}]\!]^2 \setminus A$, we shall use two different events **Boosting** and **Dilation**, respectively, to define the environments in $\mathbf{B}_{\mathrm{n}_4}(j_1, v, w)$ where $w \in [\![1, \frac{n_1}{n}]\!]^2$. First we set **Boosting** as the event that the passage time on all the edges in $\bigcup_{v \in A} \bigcup_w \mathbf{B}_{\mathrm{n}_4}(j_1, v, w)$ is in $[b - \varepsilon_7, b]$. Clearly, by the bound on $|A|$ we have

$$-\log \mathbb{P}(\mathbf{Boosting}) = O_{\varepsilon_7}(\varepsilon_1 n_1^2).$$

Finally, we define the event **Dilation**. Recall **Base-event** from Lemma 3.5. Let $\mathbf{Base\text{-}event}|_{A^c}$ denote the projection of **Base-event** to the edges in

$$\bigcup_{v \in [\![1, 2^{j_1}]\!]^2 \setminus A} \mathbf{Tile}_{\mathrm{n}_4}(j_1, v).$$

Up to a simple relabeling of the edges, **Dilation** is the intersection of $(\frac{n_1}{n})^2$ i.i.d. copies of $\mathbf{Base\text{-}event}|_{A^c}$. The relabeling will be clear from the following constructive description of how to generate a configuration of edge weights belonging to **Dilation** (see Figure 4.2): Sample $(n_1/n)^2$ many independent realizations of **Base-event**, which yield environments on $\mathbf{Box}(4\mathscr{C}^2 n)$ and let

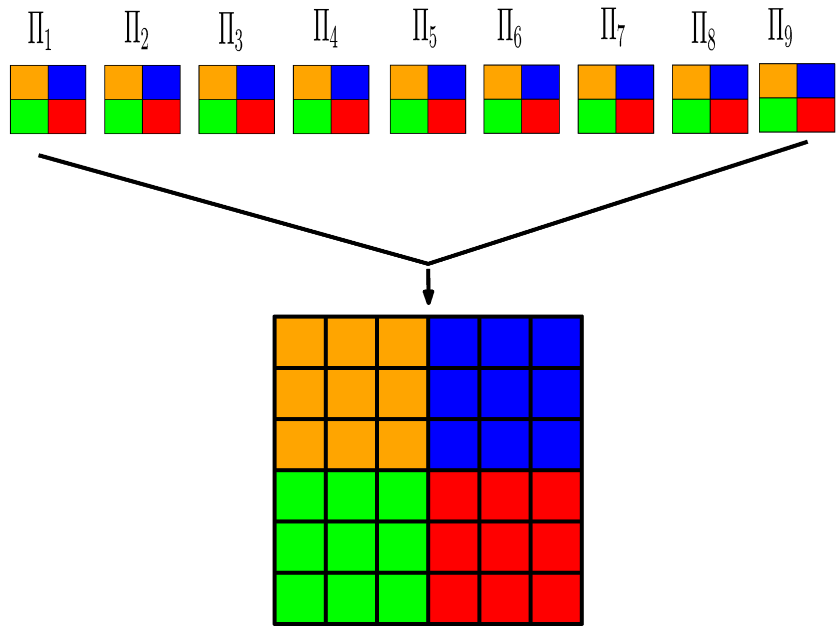


FIGURE 4.2. The figure illustrates the event **Dilation**. In the figure $\frac{n_1}{n} = 3$, $j_1 = 2$, and $A = \varnothing$. Corresponding tiles in the environments $\Pi_1, \ldots \Pi_9$ are denoted by the same color, all of which are clubbed together to form the corresponding tile in the enlarged environment which has nine times the volume.

$$\Pi_1, \Pi_2 \ldots, \Pi_{(n_1/n)^2} \tag{4.5}$$

be their restrictions on $\mathbf{Box}(\mathscr{C}n)$. For each $v \in [\![1, 2^{j_1}]\!]^2 \setminus A$ and $w \in [\![1, \frac{n_1}{n}]\!]^2$, let the edge weights on the edges in $\mathbf{B}_{\mathfrak{n}_4}(j_1, v, w)$ be the same as the edge weights of $\Pi_w$ in $\mathbf{Tile}_{\mathfrak{n}_4}(j_1, v)$, where we use the natural identification between $\mathbf{B}_{\mathfrak{n}_4}(j_1, v, w)$ and $\mathbf{Tile}_{\mathfrak{n}_4}(j_1, v)$ and the lexicographic bijection between $w \in [\![1, \frac{n_1}{n}]\!]^2$ and $\{1, 2, \ldots, (\frac{n_1}{n})^2\}$.

This bijection is used above to interchangeably use the notations $\Pi_i$ and $\Pi_w$.

Note that the choice of the term **Dilation** is natural, as by using $(\frac{n_1}{n})^2$ copies of **Base-event**, we ensure that for any $v \in [\![1, 2^{j_1}]\!]^2 \setminus A$, the environments in $\mathbf{B}_{\mathfrak{n}_4}(j_1, v, w)$ for different values of $w$ are the same in some coarse sense. Ultimately, we define

$$\mathbf{Fav} := \mathbf{Dilation} \cap \mathbf{Barrier} \cap \mathbf{Boosting}.$$

Note that the event **Barrier** along with **Dilation** describe the projection of the event **Fav** on all the edges except the edges in $\bigcup_{v \in A, w \in [\![1, \frac{n_1}{n}]\!]^2} \mathbf{B}_{\mathfrak{n}_4}(j_1, v, w)$, whereas **Boosting** defines the projections on the latter, and the three events are independent. Hence

$$\begin{aligned} \mathbb{P}(\mathbf{Fav}) &= [\mathbb{P}(\mathbf{Base\text{-}event}|_{A^c})]^{\frac{n_1^2}{n^2}} \nu([b - \varepsilon_7, b])^{O((\varepsilon_1 + \varepsilon_6) n_1^2)} \\ &\geq [\mathbb{P}(\mathbf{Base\text{-}event})]^{\frac{n_1^2}{n^2}} \nu([b - \varepsilon_7, b])^{O((\varepsilon_1 + \varepsilon_6) n_1^2)} \end{aligned} \tag{4.6}$$

where $\nu$ is the passage time distribution satisfying the hypothesis in Theorem 1. As mentioned before, by our assumption on the edge weight distribution $\nu$ from Definition 1.1, $\nu([b - \varepsilon_7, b]) > 0$ for all $\varepsilon_7 > 0$.

The proof of Proposition 1.3 will now be complete from the following lemma.

LEMMA 4.1. *Given $\varepsilon_8 > 0$ and $\varepsilon_9 > 0$, there exists $\varepsilon_4 > 0$ and the choice of parameters in the definition of* **Base-event** *in Lemma* 3.5*, and $\varepsilon_6, \varepsilon_7$ in the definition of* **Fav***, such that*

$$\frac{\log(\mathbb{P}(\mathbf{Fav}))}{n_1^2} \geq \kappa - \varepsilon_8 \quad \text{and} \quad \mathbf{Fav} \subset \mathscr{U}_{\zeta'}(n_1),$$

*where $\zeta' = \zeta - \varepsilon_9$.*

Note that, given any $\varepsilon_7 > 0$, the lower bound on the probability of **Fav** is a straightforward consequence of (4.6) and the lower bound on $\mathbb{P}(\mathbf{Base\text{-}event})$ from Lemma 3.5, by choosing $\varepsilon_1, \varepsilon_4, \varepsilon_6$ small enough. The rest of this section is devoted to the proof of the inclusion $\mathbf{Fav} \subset \mathscr{U}_{\zeta'}(n_1)$, which will follow from a series of lemmas. Before stating the lemmas we roughly describe our strategy. The proof involves showing that on the event **Fav** two things occur:

$$\mathbf{PT}_{\mathfrak{B}_0}(\mathbf{0}, \mathbf{n_1}) \geq (\mu + \zeta') n_1, \tag{4.7}$$

$$\mathbf{PT}_{\mathfrak{B}_0}(\mathbf{0}, \mathfrak{B}_0^c) \geq 2 b n_1. \tag{4.8}$$

The proof of both of the above bounds is obtained by the same strategy. Consider the two random fields given by **Fav** and **Base-event** on $\mathfrak{B}_0$ and $\mathfrak{B}_4$, respectively. Recall that the former is a 'dilation' of the latter by a factor of $\frac{n_1}{n}$, with some additional changes involving the setting up of the barriers and the boosting on the unstable tiles.

As outlined in Section 1.3, given the above, the strategy for showing (4.7) is to show that for any path $\gamma$ (joining $\mathbf{0}$ and $\mathbf{n_1}$), which we can assume to be self-avoiding, in $\mathfrak{B}_0$ there exists a path $\gamma_{\mathbf{S}}$ (joining, approximately, $\mathbf{0}$ and $\mathbf{n}$) in $\mathfrak{B}_4$ such that

$$\ell(\gamma) \geq \frac{n_1}{n}(1 - o(1))\ell(\gamma_{\mathbf{S}}), \tag{4.9}$$

where the above inequality holds pointwise for any sample point in **Fav** on the LHS and any sample point from **Base-event** on the RHS. Informally, $\gamma$ can be thought of as a path obtained by dilating the path $\gamma_{\mathbf{S}}$ by a factor $\frac{n_1}{n}$ (a similar strategy will be employed for a path joining $\mathbf{0}$ and $\mathfrak{B}_0^c$ to prove (4.8)). By definition, on **Base-event** we have $\ell(\gamma_{\mathbf{S}}) \geq (\mu + \zeta)n$, and this yields the sought lower bound of $\ell(\gamma)$. To make (4.9) rigorous, we need some regularity properties of the path $\gamma$, which will be obtained by a preprocessing. This is done in the following subsection.

## 4.2 Preprocessing of Paths

We shall see later that it suffices to consider only self-avoiding paths contained in $\mathfrak{B}_0$ that start and end on the boundary of some $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$ and $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v')$, respectively. Observe that any such path (thought of as a sequence of edges) admits a unique decomposition as a concatenation of a number of paths, i.e., $\gamma = \alpha_0\chi_0\alpha_1\chi_1\alpha_2\chi_2\dots\alpha_L\chi_L\alpha_{L+1}$ with the following properties:

(i) Each $\alpha_i$ is contained in some $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v_i)$ for some $v_i \in [\![1, 2^{j_1}]\!]^2$ (recall that we identify any box with all the lattice edges contained in its closure); $\alpha_0$ and $\alpha_{L+1}$ could be empty.

(ii) Each $\chi_i$ is contained in $\mathbf{Corridor}^{\text{ext}}(j_1, \mathfrak{n}_0)$ (again thought of as a union of edges).

Given $\varepsilon_6$ as in (4.1), let us call the paths $\alpha_i$ for $i \in \{1, 2, \dots, L\}$ as *excursions* of $\gamma$, and let us call the above decomposition of $\gamma$ its decomposition into excursions. Let $\mathbf{x}_i$ (resp., $\mathbf{y}_i$) denote the starting (resp., ending) vertex of $\alpha_i$. Let us call the excursion $\alpha_i$ *large* if there exists a vertex $\mathbf{z}_i$ on $\alpha_i$ such that

$$\min\{\|\mathbf{x}_i - \mathbf{z}_i\|, \|\mathbf{y}_i - \mathbf{z}_i\|\} \geq \varepsilon_6^2 \frac{\mathfrak{n}_0}{2^{j_1}}.$$

Observe that $\alpha_i$ is large if $\|\mathbf{x}_i - \mathbf{y}_i\| \geq 2\varepsilon_6^2 \frac{\mathfrak{n}_0}{2^{j_1}}$.

We shall need to define one more property of a path. Consider a path $\gamma$ with the decomposition into excursions as above. Observe that each $\chi_i$ must start at a boundary vertex of $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$ and end at a boundary vertex of $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v')$ for some $v = v(\chi_i), v' = v'(\chi_i) \in [\![1, 2^{j_1}]\!]^2$. Note that $v$ can in fact be equal to $v'$. We call the path $\gamma$ *regular* if for each $\chi_i$ we have $\|v(\chi_i) - v'(\chi_i)\|_1 = 1$ (i.e., they

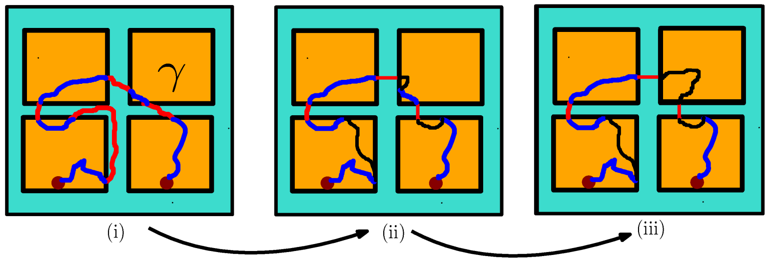


FIGURE 4.3. A schematic diagram describing the preprocessing. (i) illustrates the decomposition of the path $\gamma$ into $\alpha_i$ (blue segments) and $\chi_i$ (red segments). (ii) describes the content of Lemma 4.3, where each red segment is replaced by a regular path. (iii) describes the content of Lemma 4.4, where if an excursion is small (the part in the northeast tile), then we replace it by a larger excursion without changing the length too much. Note that the path starts and ends on the boundary of tiles, which we show can be assumed and would be convenient.

are neighbouring vertices) and $\chi_i$ is a vertical or horizontal segment depending on the relative positions of $v$ and $v'$.

The next lemma allows us to consider regular paths only. Recall the parameters $\varepsilon_6$ and $\varepsilon_7$ in the definition of the event **Fav**.

LEMMA 4.2. *For any path $\gamma$ and contained in $\mathfrak{B}_0$ whose endpoints are at a distance larger than $\frac{n_1}{2}$ and are located on boundaries of some tiles $\textbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$ and $\textbf{Tile}^*_{\mathfrak{n}_1}(j_1, v')$, respectively, there exists a regular path $\mathcal{P}$ with the same endpoints as $\gamma$ such that*

(i) *All the excursions of $\mathcal{P}$ are large.*
(ii) *On $\textbf{Barrier}^{\text{ext}}(\mathfrak{n}_0, j_1)$, we have $\ell(\gamma) \geq (1 - O(\varepsilon_7 + \varepsilon_6))\ell(\mathcal{P})$.*

The proof of the above lemma is done in two steps (see Figure 4.3 for an illustration). Let $\gamma$ be fixed as in the lemma. Consider its decomposition into excursions: $\gamma = \alpha_0\chi_0\alpha_1\chi_1\alpha_2\chi_2\cdots$. Observe that if we can replace each $\chi_i$ by an $\mathsf{L}$-shaped path with the same endpoints, the resulting path will be regular. The following lemma shows that this can be done without increasing the length of the path by more than a factor of $(1 - O(\varepsilon_7))^{-1}$.

LEMMA 4.3. *Consider a path $\chi$ completely contained in $\textbf{Corridor}^{\text{ext}}(j_1, \mathfrak{n}_0)$ whose starting and ending vertices are located on the boundaries of $\textbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$ and $\textbf{Tile}^*_{\mathfrak{n}_1}(j_1, v')$, respectively. Then there exists a regular path $\mathcal{P}_\chi$ with the same starting and ending points such that on $\textbf{Barrier}^{\text{ext}}(\mathfrak{n}_0, j_1)$, we have $\ell(\chi) \geq (1 - O(\varepsilon_7))\ell(\mathcal{P}_\chi)$.*

PROOF. Let $\mathbf{x}$ and $\mathbf{y}$ be the starting and ending point of $\chi$ respectively. Consider the 1−norm minimizing path from $\mathbf{x}$ to $\mathbf{y}$ that consists of a horizontal path followed by a vertical path (this choice is arbitrary): i.e., for $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$

consider the piecewise $\mathbb{L}$ shaped path obtained by concatenating the straight line segment obtained by joining $\mathbf{x}$ to $(y_1, x_2)$ followed by the straight line segment obtained by joining $(y_1, x_2)$ to $\mathbf{y}$. Call this path $\mathbb{L}$ and consider $\mathbb{L}$ as a path on the nearest neighbour graph of $\mathbb{Z}^2$.

Observe that there exists points $u_0 = \mathbf{x}, u_1, \dots, u_\ell = \mathbf{y}$ on $\mathbb{L}$, all on boundaries of $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, u)$'s such that $\mathbb{L}$ restricted between $u_i$ and $u_{i+1}$ (called $\mathbb{L}_i$) is either (a) contained in $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, u)$ for some $u$ (type $A$, say) or (b) is entirely contained in $\mathbf{Corridor}^{\mathrm{ext}}(j_1, \mathfrak{n}_0)$, and further $u_i \in \mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, u), u_{i+1} \in \mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, u')$ for some $u, u'$ that have $\ell_1$ distance 1 (type $B$). Observe again that such a decomposition is unique.

If $\mathbb{L}_i$ is type $A$, let us set $\mathcal{P}_i$ to be the shortest path between $u_i$ and $u_{i+1}$ contained in $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, u)$, and if $\mathbb{L}_i$ is type $B$ we set $\mathcal{P}_i = \mathbb{L}_i$. Consider the path $\mathcal{P}_\chi = \mathcal{P}_0\mathcal{P}_1 \cdots \mathcal{P}_{\ell-1}$ obtained by concatenating the $\mathcal{P}_i$. It is clear that the path $\mathcal{P}_\chi$ obtained as above is regular (see Figure 4.3 for an illustration), and hence it only remains to show that on $\mathbf{Barrier}^{\mathrm{ext}}(\mathfrak{n}_0, j_1)$, we have $\ell(\chi) \geq (1 - O(\varepsilon_7))\ell(\mathcal{P}_\chi)$. Observe that on $\mathbf{Barrier}^{\mathrm{ext}}(\mathfrak{n}_0, j_1)$ we have $\ell(\chi) \geq (b - \varepsilon_7)\|\mathbf{x} - \mathbf{y}\|_1$. It also follows from the definitions that $\ell(\mathcal{P}_i) \leq b\|u_i - u_{i+1}\|_1$. The proof is completed by observing $\sum_{i=0}^{\ell-1} \|u_i - u_{i+1}\|_1 = \|\mathbf{x} - \mathbf{y}\|_1$. □

Lemma 4.3 tells us that for any $\gamma$ as in the statement of Lemma 4.2, one can replace the paths $\chi_i$ in its decomposition by the paths $\mathcal{P}_{\chi_i}$ as constructed in Lemma 4.3 to end up with a regular path $\mathcal{P}_*$ with the same endpoints such that $\ell(\gamma) \geq (1 - O(\varepsilon_7))\ell(\mathcal{P}_*)$. The following lemma ensures the largeness of the excursions and therefore suffices to complete the proof of Lemma 4.2.

LEMMA 4.4. *For any regular path $\gamma$ contained in $\mathfrak{B}_0$ whose endpoints are at a distance larger than $n_1/2$ and are located on boundaries of some tiles $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$ and $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v')$, respectively, there exists a regular path $\mathcal{P}$ with the same endpoints such that*

(i) *Each excursion of $\mathcal{P}$ is large.*
(ii) *On $\mathbf{Barrier}^{\mathrm{ext}}(\mathfrak{n}_0, j_1)$, we have $\ell(\gamma) \geq (1 - O(\varepsilon_6))\ell(\mathcal{P})$.*

PROOF. Let $\gamma$ be as in the statement of the lemma. Consider its decomposition into excursions $\gamma = \alpha_0\chi_0\alpha_1\chi_1\alpha_2\chi_2 \cdots \alpha_L\chi_L\alpha_{L+1}$. The proof again will be a step-by-step procedure. We shall inspect the short excursions one by one, and remove them by modifying the path locally without increasing the lengths too much. Let $\alpha_i$ be contained in $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v_i)$. For each nonempty $\alpha_i$, we shall replace it, if necessary, by a path $\alpha_i'$ contained in $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v_i)$ with the same starting and ending point as $\alpha_i$. If $\alpha_i$ is a large excursion or is empty (recall that $\alpha_0$ and $\alpha_{L+1}$ can be empty), we set $\alpha_i' = \alpha_i$. Consider any excursion $\alpha_i$ that is not large. Let $\mathbf{x}_i$ and $\mathbf{y}_i$ be its starting and ending points, respectively. Fix a vertex $\mathbf{z}_i$ in $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v_i)$ such that

$$\|\mathbf{x}_i - \mathbf{z}_i\|, \|\mathbf{y}_i - \mathbf{z}_i\| \in \left(\varepsilon_6^2 \frac{\mathfrak{n}_0}{2^{j_1}}, 2\varepsilon_6^2 \frac{\mathfrak{n}_0}{2^{j_1}}\right);$$

clearly such a vertex exists. Now consider the path $\alpha'_i = \alpha_i^{(1)}\alpha_i^{(2)}$ where $\alpha_i^{(1)}$ (resp., $\alpha_i^{(2)}$) is the shortest path between $\mathbf{x}_i$ and $\mathbf{z}_i$ (resp., $\mathbf{z}_i$ and $\mathbf{y}_i$) contained in $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v_i)$. Clearly $\alpha'_i$ is a large excursion. Consider the path

$$\mathcal{P} = \alpha'_0\chi_0\alpha'_1\chi_1\cdots\alpha'_L\chi_L\alpha'_{L+1}.$$

Notice that $\mathcal{P}$ satisfies the first conclusion of the lemma by definition. We shall show that it also satisfies the required upper bound on $\ell(\mathcal{P})$.

We shall show that, on $\mathbf{Barrier}^{\mathrm{ext}}(\mathfrak{n}_0, j_1)$, for each $i = 0, 1, \ldots, L-1$ we have $\ell(\alpha_i\chi_i) \geq (1 - O(\varepsilon_6))\ell(\alpha'_i\chi_i)$, and further we have

$$\ell(\alpha_L\chi_L\alpha_{L+1}) \geq (1 - O(\varepsilon_6))\ell(\alpha'_L\chi_i\alpha'_{L+1}).$$

Clearly this suffices. Consider any $i \leq L-1$ such that $\alpha_i \neq \alpha'_i$ (there is nothing to prove otherwise). To get an upper bound on $\ell(\alpha'_i\chi_i)$, observe that $\ell(\alpha'_i) \leq 8b\varepsilon_6^2\frac{\mathfrak{n}_0}{2^{j_1}}$ and on $\mathbf{Barrier}^{\mathrm{ext}}(\mathfrak{n}_0, j_1)$, by taking $\varepsilon_6, \varepsilon_7$ sufficiently small, we have $\ell(\chi_i) \geq (b - \varepsilon_7)\varepsilon_6\frac{\mathfrak{n}_0}{2^{j_1}}$ and $\ell(\alpha_i\chi_i) \geq (1 - O(\varepsilon_6))\ell(\alpha'_i\chi_i)$. The same argument gives $\ell(\alpha_L\chi_L\alpha_{L+1}) \geq (1 - O(\varepsilon_6))\ell(\alpha'_L\chi_i\alpha'_{L+1})$ and completes the proof of the lemma. □

Given the regular path $\mathcal{P} = \alpha_0\chi_0\alpha_1\chi_1\alpha_2\chi_2\cdots\alpha_L\chi_L\alpha_{L+1}$ from Lemma 4.2, we use essentially the same arguments on each of the nonempty excursions $\alpha_i$ as in the proof of Lemma 4.2 to obtain a further decomposition. For a path $\alpha$ contained in some $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$ with endpoints on the boundary of $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$, consider the unique decomposition $\alpha = \chi'_0\beta_1\chi'_1\cdots\beta_{L'}\chi'_{L'}$ such that each $\beta_i$ is contained in $\mathbf{B}(\mathfrak{n}_4, v_i, w)$ for some $w \in [\![1, \frac{n_1}{n}]\!]^2$ and each $\chi'_i$ contained in $\mathbf{Corridor}^{\mathrm{int}}(j_1, \mathfrak{n}_0)$ (observe that since the endpoints of $\alpha$ are on the boundary of $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$, $\chi'_0$ and $\chi'_{L'}$ are both nonempty but it is possible that $L' = 0$ and $\alpha$ is completely contained in $\mathbf{Corridor}^{\mathrm{int}}(j_1, \mathfrak{n}_0)$).

We call such a path $\alpha$ *strongly regular* if each $\chi'_i$ is an $\mathbb{L}$-shaped path joining its endpoints and if the nonempty excursions $\beta_i$ are *large* where the definition of excursion is large is as before except $\mathfrak{n}_0$ is replaced by $\mathfrak{n}_4$. Note that in the above definition all the $\chi'_i$s must in fact be straight lines except possibly $\chi'_0$ and $\chi'_{L'}$.

The same arguments as in the proof of the two previous lemmas yield the following result whose proof we omit.

LEMMA 4.5. *For a path $\alpha$ contained in some $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$ with endpoints on the boundary of $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$, there exists a strongly regular path $\mathcal{P}_\alpha$ with the same endpoints such that, on $\mathbf{Barrier}^{\mathrm{int}}(\mathfrak{n}_0, j_1)$, we have $\ell(\alpha) \geq (1 - O(\varepsilon_7 + \varepsilon_6))\ell(\mathcal{P}_\alpha)$.*

### 4.3 Rescaling Paths

Equipped with the above results, we are now ready to prove (4.7) (a very similar argument will take care of (4.8)). Following the strategy indicated in (4.9), for a given path joining $\mathbf{0}$ and $\mathbf{n_1}$ contained in $\mathfrak{B}_0$, we want to create its scaled version. As we want to apply Lemma 4.2, we shall work with paths joining with endpoints

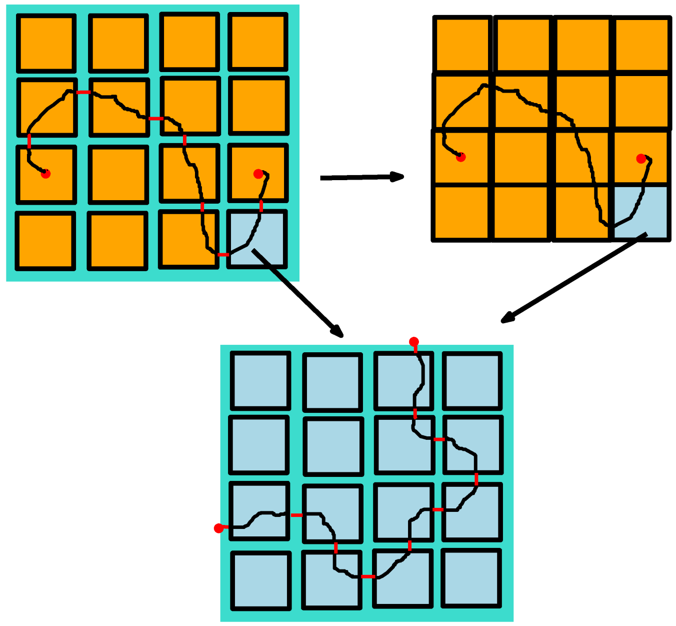

FIGURE 4.4. The left-to-right arrow in the top illustrates a natural identification $\mathfrak{B}_0 \setminus \mathbf{Corridor}^{\text{ext}}(j_1, \mathfrak{n}_0) \to \mathbf{Box}(\mathfrak{n}_1 + 2^{j_1})$ by collapsing the corridors $\mathbf{Corridor}^{\text{ext}}(j_1, \mathfrak{n}_0)$. The bottom figure zooms in on the southeast-most tile and describes a strongly regular path.

close to $\mathbf{0}$ and $\mathbf{n_1}$ instead. Let $\widetilde{\mathbf{0}}$ (resp., $\widetilde{\mathbf{n_1}}$) denote the closest lattice point to $\mathbf{0}$ (resp., $\mathbf{n_1}$) in any $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$. Clearly,

$$\left|\mathbf{PT}_{\mathfrak{B}_0}(\mathbf{0}, \mathbf{n_1}) - \mathbf{PT}_{\mathfrak{B}_0}(\widetilde{\mathbf{0}}, \widetilde{\mathbf{n_1}})\right| = O\left(\frac{n_1}{2^{j_1}}\right), \tag{4.10}$$

and by choosing $j_1$ sufficiently large depending on $\varepsilon_9$ (which appears in the statement of Lemma 4.1), obtaining a suitable lower bound for $\mathbf{PT}_{\mathfrak{B}_0}(\widetilde{\mathbf{0}}, \widetilde{\mathbf{n_1}})$ would be sufficient.

Let us now fix a self-avoiding path $\gamma$ contained in $\mathfrak{B}_0$ between $\widetilde{\mathbf{0}}$ and $\widetilde{\mathbf{n_1}}$. Lemma 4.2 yields a regular path $\gamma_{\text{reg}} = \alpha_0 \chi_0 \alpha_1 \chi_1 \alpha_2 \chi_2 \cdots \alpha_L$, in which $\alpha_i$ is contained in $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v_i)$, which, on applying Lemma 4.5 to the $\alpha_i$s, subsequently yields the path $\gamma_{\text{s-reg}} = \beta_0 \chi_0 \beta_1 \chi_1 \beta_2 \chi_2 \ldots \beta_L$ with each $\beta_i$ strongly regular. Observe that

$$\ell(\gamma) \geq (1 - O(\varepsilon_6 + \varepsilon_7))\ell(\gamma_{\text{s-reg}}), \tag{4.11}$$

and hence it suffices to get a lower bound for $\ell(\gamma_{\text{s-reg}})$. In the next few lemmas we create a scaled version $\gamma^{\mathbf{S}}$ of this path $\gamma_{\text{s-reg}}$. What we do is rather simple and natural.

Using the definition of a regular path, $\gamma_{\text{s-reg}}$ corresponds to a natural path one can form in $\mathfrak{B}_3$, by collapsing all the external corridors. For the technical convenience of not having to identify the boundaries of neighbouring tiles $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v)$,

$\mathbf{Tile}^*_{\mathfrak{n}_2 1}(j_1, v')$ for some $v, v'$, we will collapse the external corridors to width 1 corridors. Formally, we will use the natural identification

$$\mathfrak{B}_0 \setminus \mathbf{Corridor}^{\mathrm{ext}}(j_1, \mathfrak{n}_0) \to \mathbf{Box}(\mathfrak{n}_1 + 2^{j_1}), \tag{4.12}$$

which is easily seen by replacing $\mathbf{Corridor}^{\mathrm{ext}}(j_1, \mathfrak{n}_0)$ by corridors of width 1. For brevity, we will denote $\mathbf{Box}(\mathfrak{n}_1 + 2^{j_1})$ by $\widetilde{\mathbf{Box}}$. This allows us to identify the path $\gamma_{\text{s-reg}}$ with a path $\widetilde{\gamma}$ in $\widetilde{\mathbf{Box}}$ formed by replacing the bridges $\chi_i$ by a length 1 path (notice that it is possible because $\gamma_{\text{s-reg}}$ is regular).

Under the above operation, $\widetilde{\gamma}$ admits a decomposition $\widetilde{\gamma} = \widetilde{\gamma}_0\widetilde{\gamma}_1 \cdots \widetilde{\gamma}_L$, where the ending point of $\widetilde{\gamma}_i$ is adjacent to the beginning point of $\widetilde{\gamma}_{i+1}$ (the connecting edge being obtained by collapsing the external corridors to width 1 corridors) and $\widetilde{\gamma}_i$ belongs to $\widetilde{\mathbf{Tile}}(j_1, v_i)$, the box that $\mathbf{Tile}^*_{\mathfrak{n}_1}(j_1, v_i)$ maps to under the above operation.

Let the starting and ending points of $\widetilde{\gamma}_i$ be $\widetilde{x}_i$ and $\widetilde{y}_i$. We will now scale all these points by a factor $\frac{n}{n_1}(1 + O(\varepsilon_6))$ to obtain the sequence of points

$$\widehat{x}_i = \frac{\mathfrak{n}_4}{\mathfrak{n}_1}\widetilde{x}_i, \quad \widehat{y}_i = \frac{\mathfrak{n}_4}{\mathfrak{n}_1}\widetilde{y}_i, \tag{4.13}$$

for $i = 0, 1, \ldots, L$. (Note that the notations $\widehat{x}$ and $\widetilde{x}$ have been used before to denote various approximations of a point $x$; see the paragraph after Lemma 2.1 and the paragraph following (3.3). However, the usage of the above notations in this section is rather minimal and local, and we allow ourselves this abuse of notation to avoid introducing new symbols.) Of course, one cannot expect these to be lattice points. Further, the fact that $\widetilde{\gamma}$ was in $\widetilde{B}$ whose size is slightly bigger than $\mathfrak{n}_1$ means scaling by $\frac{\mathfrak{n}_4}{\mathfrak{n}_1}$ does not quite map $\widetilde{\mathbf{Box}}$ to $\mathfrak{B}_4 = \mathbf{Box}(\mathscr{C}n)$, nor does $\widetilde{\mathbf{Tile}}(j_1, v_i)$ map to $\mathbf{Tile}_{\mathscr{C}n}(j_1, v_i)$. Nonetheless, as one would expect, the discrepancies are minor and only introduce the need for some bookkeeping to maintain precision and do not affect our arguments.

We round off the points to the nearest points on a grid. To this end, recall $m_1$, $j_1$, from the definition of **Base-event** and the notation $\mathbf{Grid}_{\mathfrak{n}_4}(\ell_2; j_1)$ where

$$\ell_2 := \frac{\mathscr{C}n}{2^{j_1 + \frac{m_1}{2}}}.$$

We want to define $x_i^{\mathbf{S}}$ to be the closest point in $\mathbf{Tile}_{\mathfrak{n}_4}(j_1, v_i) \cap \mathbf{Grid}_{\mathfrak{n}_4}(\ell_2; j_1)$ to $\widehat{x}_i$. However, these might not be points in $\mathbb{Z}^2$, and hence we define them to be the closest point in $\mathbb{Z}^2$ to the closest point in $\mathbf{Tile}_{\mathfrak{n}_4}(j_1, v_i) \cap \mathbf{Grid}_{\mathfrak{n}_4}(\ell_2; j_1)$ instead. Similarly, define $y_i^{\mathbf{S}}$. However since $y_i^{\mathbf{S}}$ and $x_{i+1}^{\mathbf{S}}$ are rather close, we will ignore the former and consider the path $\gamma^{\mathbf{S}}$ to be the shortest path between the sequence of points

$$x_0^{\mathbf{S}}, x_1^{\mathbf{S}}, \ldots x_L^{\mathbf{S}}, x_{L+1}^{\mathbf{S}} \tag{4.14}$$

where $x_{L+1}^{\mathbf{S}} = y_L^{\mathbf{S}}$. Let $\gamma_i^{\mathbf{S}}$ be the segment of $\gamma^{\mathbf{S}}$ between $x_i^{\mathbf{S}}$ and $x_{i+1}^{\mathbf{S}}$. Observe that by our choice of the starting and ending points of $\gamma$ and by an appropriate choice of parameters (i.e., making $j_1$ large) we get that

$$\|\widetilde{x}_{\mathbf{0}}\| = O(\varepsilon_6 n_1), \quad \|\mathbf{n_1} - \widetilde{y}_L\| = O(\varepsilon_6 n_1),$$

and so

$$\|x_0^{\mathbf{S}}\| = O(\varepsilon_6 n), \quad \|\mathbf{n} - x_{L+1}^{\mathbf{S}}\| = O(\varepsilon_6 n); \tag{4.15}$$

hence $\gamma^{\mathbf{S}}$ can thought of as a path between $\mathbf{0}$ and $\mathbf{n}$.

As a consequence of the approximate convexity statement in Proposition 3.4, we have the following key result.

LEMMA 4.6. *Given any $\varepsilon_{11} > 0$, there exists a choice of the parameters in the definition of* **Base-event** *and* **Fav**, *namely $\varepsilon_6$ and $\varepsilon_7$ small enough, and then $\delta_1$ small enough followed by $j_1$ and $m_1$ large enough, such that, deterministically for any $i$,*

$$\max\left(\ell(\widetilde{\gamma}_i), (b-\varepsilon_7)\frac{\varepsilon_6 \mathscr{C} n_1}{2^{j_1}}\right) \geq (1-\varepsilon_{11})\frac{n_1}{n}\ell\big(\gamma_i^{\mathbf{S}}\big), \tag{4.16}$$

*where, as mentioned in* (4.9)*, the LHS is computed on any sample point on the event* **Fav** *and the RHS is computed on any sample point in* **Base-event**.

The scaling in (4.13) in fact implies a slightly stronger bound where the factor $\frac{n_1}{n}$ is replaced by $\frac{\mathrm{n}_1}{\mathrm{n}_4} = (1+\varepsilon_6)\frac{n_1}{n}$. However, for our purposes the above weaker bound suffices.

Before proving the above lemma we finish the proof of Lemma 4.1 using the above results.

PROOF OF LEMMA 4.1. The proof will clearly follow by showing (4.7) and (4.8). We will show only the former; the latter has a similar proof. Recall the definition of $\widetilde{\mathbf{0}}$ and $\widetilde{\mathbf{n_1}}$ and fix any path $\gamma$ from $\mathbf{0}$ to $\mathbf{n_1}$ in $\mathfrak{B}_0$. Now applying Lemmas 4.2 and 4.5, we obtain the path $\gamma_{\text{s-reg}}$ as described above and the collapsing operation above leads to the path $\widetilde{\gamma}$. Finally, consider the scaled path $\gamma^{\mathbf{S}}$. We now have the following string of inequalities where the first two inequalities are immediate consequences of our above constructions and the third inequality is the lemma above:

$$\begin{aligned}
\ell(\gamma) &\overset{(4.11)}{\geq} (1 - O(\varepsilon_7 + \varepsilon_6))\ell(\gamma_{\text{s-reg}}) \\
&\geq (1 - O(\varepsilon_7 + \varepsilon_6)) \sum_{i=0}^{L-1}\left[\ell(\widetilde{\gamma}_i) + (b-\varepsilon_7)\frac{\varepsilon_6 \mathscr{C} n_1}{2^{j_1}}\right] + \ell(\widetilde{\gamma}_L) \\
&\geq (1 - O(\varepsilon_7 + \varepsilon_6))(1-\varepsilon_{11})\frac{n_1}{n}\sum_{i=0}^{L}\ell(\gamma_i^{\mathbf{S}}) \\
&= (1 - O(\varepsilon_7 + \varepsilon_6))(1-\varepsilon_{11})\frac{n_1}{n}\ell(\gamma^{\mathbf{S}}),
\end{aligned} \tag{4.17}$$

where $\ell(\gamma), \ell(\gamma_{\text{s-reg}})$, and $\ell(\widetilde{\gamma}_i)$ are computed on any sample point on the event **Fav**, and $\ell(\gamma_i^{\mathbf{S}}), \ell(\gamma^{\mathbf{S}})$ are computed on any sample point in **Base-event**.

Now, by definition, $\gamma^{\mathbf{S}}$ is a path joining $x_0^{\mathbf{S}}$ and $x_{L+1}^{\mathbf{S}}$ in $\mathfrak{B}_4$ and hence on **Base-event**, using (4.15) we have $\ell(\gamma^{\mathbf{S}}) \geq (\mu + \zeta)n(1 - O(\varepsilon_6))$. It therefore follows that

$$\mathbf{PT}_{\mathfrak{B}_0}(\widetilde{\mathbf{0}}, \widetilde{\mathbf{n}_1}) \geq (1 - O(\varepsilon_7 + \varepsilon_6))(1 - \varepsilon_{11})(\mu + \zeta)n_1.$$

Choosing $\varepsilon_6$, $\varepsilon_7$, and $\varepsilon_{11}$ small enough depending on $\varepsilon_9$, and using (4.10) and choosing $j_1$ sufficiently large, we establish (4.7). An identical argument involving paths starting from $\mathbf{0}$ to $\mathfrak{B}_0^c$ now implies (4.8) and completes the proof. $\square$

We now prove Lemma 4.6 using Lemmas 4.2 and 4.5 and Proposition 3.4.

PROOF OF LEMMA 4.6. Recall the set of unstable tiles $A$ in the definition of **Base-event**. For the proof let us consider an environment $\Pi_1 \in$ **Base-event**. Recall that the latter is measurable with respect to the edges in $\mathbf{Box}(4\mathscr{C}^2 n)$, with the property that the weight of any path from $\mathbf{0}$ to $\mathbb{Z}^2 \setminus \mathbf{Box}(\mathscr{C}n)$ exceeds $4bn$.

Fix any $i$, and let $\widetilde{\gamma}_i = \chi'_{i,0}\beta_{i,1}\chi'_{i,1}\beta_{i,2}\cdots\beta_{i,L'}\chi'_{i,L'}$ such that each $\beta_{i,j}$ is large and contained in $\mathbf{B}(\mathfrak{n}_4, v_i, w)$ for some $w \in [\![1, \frac{n_1}{n}]\!]^2$ and each $\chi'_i$ is contained in $\mathbf{Corridor}^{\text{int}}(j_1, \mathfrak{n}_0)$. (This follows from the discussion preceding Lemma 4.5).

Since all the $\beta_{i,j}$ are large, for each $j$ there is a point (say $a_{i,j}^2$) that is at least $\varepsilon_6^2 \frac{\mathscr{C}n}{2^{j_1}}$ distance away from the endpoints $a_{i,j}^0$ and $a_{i,j}^1$. Let

$$\overrightarrow{w}_{i,j}^{1} \quad \text{and} \quad \overrightarrow{w}_{i,j}^{2}$$

be the vectors obtained by taking the difference of $a_{i,j}^2 - a_{i,j}^0$ and $a_{i,j}^1 - a_{i,j}^2$, respectively. Also, let $\beta_{i,j}^1$ and $\beta_{i,j}^2$ be the two subpaths of $\beta_{i,j}$ split at $a_2$. Further, by construction, the segments $\chi'_{i,j}$ are also large, in the sense that their endpoints are also at least $\varepsilon_6^2 \frac{\mathscr{C}n}{2^{j_1}}$ away. Let the vector joining their endpoints be

$$\overrightarrow{w}_{i,j}^{3}.$$

Thus for all $i, j$,

$$\min\left(\|\overrightarrow{w}_{i,j}^{1}\|, \|\overrightarrow{w}_{i,j}^{2}\|, \|\overrightarrow{w}_{i,j}^{3}\|\right) \geq \varepsilon_6^2 \frac{\mathscr{C}n}{2^{j_1}}. \tag{4.18}$$

Now recall that for any $\Pi \in$ **Base-event**, and for all $v \in [\![1, 2^{j_1}]\!]^2 \setminus A$, $\mathbf{Tile}_{\mathfrak{n}_4}(j_1, v)$ is $(\delta_1, \ell_1, k_1)$-**Stable**, and moreover recall the approximate norm $\|\cdot\|_{(j_1, v_i)}$ from the statement in Proposition 3.4.

Now the following argument is split into two cases depending on whether $v_i \in A$ or not in $A$ where $v_i$ is the index of the tile such that $\widetilde{\gamma}_i \in \widetilde{\mathbf{Tile}}(j_1, v_i)$. In the subsequent arguments the parameter $\varepsilon''$ will be used to denote a small number whose value will change from line to line. We will make explicit the dependence $\varepsilon$ on the various other parameters at the end to finally achieve the desired smallness.

*Case* 1. $v_i \notin A$. In this case, the following string of inequalities holds (we provide an explanation for the inequalities following the statement):

$$
\begin{aligned}
\ell(\widetilde{\gamma}_i) &\geq \left[\sum_{j=0}^{L'-1}\left[\ell(\chi'_{i,j}) + \ell(\beta_{i,j+1})\right] + \ell(\chi'_{i,L'})\right] \\
&\geq (1-\varepsilon'')\left[\sum_{j=0}^{L'-1}\left(\|\vec{w}^{3}_{i,j}\|_{(j_1,v_i)} + \|\vec{w}^{1}_{i,j}\|_{(j_1,v_i)} + \|\vec{w}^{2}_{i,j}\|_{(j_1,v_i)}\right)\right. \\
&\qquad\qquad \left. + \|\vec{w}^{3}_{i,L'}\|_{(j_1,v_i)}\right] \\
&\geq (1-\varepsilon'')\left\|\sum_{j=0}^{L'-1}\left(\vec{w}^{1}_{i,j} + \vec{w}^{2}_{i,j} + \vec{w}^{3}_{i,j}\right) + \vec{w}^{3}_{i,L'}\right\|_{(j_1,v_i)} \\
&\geq (1-\varepsilon'')\|\vec{\widetilde{\gamma}}_i\|_{(j_1,v_i)} \geq (1-\varepsilon'')\frac{n_1}{n}\left\|\vec{\gamma}^{\mathbf{S}}_i + \vec{O}(\ell_2)\right\|_{(j_1,v_i)}
\end{aligned} \tag{4.19}
$$

where $\vec{\widetilde{\gamma}}_i$ (resp., $\vec{\gamma^{\mathbf{S}}_i}$) is the vector obtained by taking the difference of the end points of the path $\widetilde{\gamma}_i$ (resp., $\gamma_i^{\mathbf{S}}$) (where $\vec{O}(x)$ denotes a vector with Euclidean norm bounded by $O(x)$).

The first equality is straightforward since $\chi'_{i,j}$ and $\beta_{i,j}$ are disjoint subpaths of $\widetilde{\gamma}_i$. For the inequality in the second line, we use that $\beta_{i,j} \subset \mathbf{B}(\mathfrak{n}_4, v_i, w)$ for some $w \in [\![1, \frac{n_1}{n}]\!]^2$. Recall that we are on the event **Fav**, and given $v_i \notin A$, for each $w$, the environment in $\mathbf{B}(\mathfrak{n}_4, v_i, w)$ is a restriction of $\Pi_w$ (see (4.5)) to $\mathbf{Tile}_{\mathfrak{n}_4}(j_1, v_i)$. Thus on the event **Fav**, the passage times on $\mathbf{B}(\mathfrak{n}_4, v_i, w)$ between points whose Euclidean separation is at least what is prescribed by (3.7), induce the same approximate norm $\|\cdot\|_{(j_1,v_i)}$, which satisfies the conclusion of Proposition 3.4. The second inequality now follows from the definition of the approximate norm $\|\cdot\|_{(j_1,v_i)}$, using the fact that $\mathbf{Tile}_{\mathfrak{n}_4}(j_1, v_i)$ is stable, together with the lower bound on the Euclidean norms of the vectors in (4.18), which implies

$$
\ell(\beta_{i,j}) = \ell\big(\beta^1_{i,j}\big) + \ell\big(\beta^2_{i,j}\big) \geq (1-\varepsilon'')\Big[\|\vec{w}^{1}_{i,j}\|_{(j_1,v_i)} + \|\vec{w}^{2}_{i,j}\|_{(j_1,v_i)}\Big].
$$

Note that by (3.9), one can make $\varepsilon''$ above arbitrarily small by choosing $\delta_1$ small and $m_1$ large. Note that (4.18) is needed crucially since (3.9) holds for pairs of points which are at distance at least $\frac{\mathscr{C}n}{2^{j_1+m_1/4}}$.

However, the segments $\chi'_{i,j}$ are subsets of $\mathbf{Corridor}^{\text{int}}(j_1, \mathfrak{n}_0)$ and not $\mathbf{B}(\mathfrak{n}_4, v_i, w)$, and hence one cannot, in general, relate $\ell(\chi'_{i,j})$ to $\|\overrightarrow{w}^3_{i,j}\|_{(j_1,v_i)}$. Nonetheless, since the edge weights are bounded by $b$, it follows that for any vector $\overrightarrow{w}$,

$$\|\overrightarrow{w}\|_{(j_1,v_i)} \leq (1+\varepsilon'')b\|\overrightarrow{w}\|_1,$$

where similarly as above, by (3.10), $\varepsilon''$ can be made small enough by choosing $\delta_1$ small and $m_1$ large. Now, since on **Barrier** we deterministically have the bound $\ell(\chi'_{i,j}) \geq (1-O(\varepsilon_7))b\|\overrightarrow{w}^3_{i,j}\|_1$, it follows that

$$\ell(\chi'_{i,j}) \geq (1-\varepsilon'')\|\overrightarrow{w}^3_{i,j}\|_{(j_1,v_i)}$$

The third inequality is the approximate triangle inequality, which is the content of Proposition 3.4, where again $\varepsilon''$ can be made small enough by choosing $\delta_1$ small and $m_1$ large. The fourth inequality is straightforward since by definition,

$$\sum_{j=0}^{L'-1}\left(\overrightarrow{w}^1_{i,j} + \overrightarrow{w}^2_{i,j} + \overrightarrow{w}^3_{i,j}\right) + \overrightarrow{w}^3_{i,L'} = \overrightarrow{\widetilde{\gamma}}_i.$$

For the final inequality, note that due to rounding to points in $\mathbf{Grid}_{\mathfrak{n}_4}(\ell_2; j_1)$ to obtain the points $x^{\mathbf{S}}_0, x^{\mathbf{S}}_1, \dots x^{\mathbf{S}}_L, x^{\mathbf{S}}_{L+1}$, it follows that

$$\overrightarrow{\widetilde{\gamma}}_i = \frac{\mathfrak{n}_1}{\mathfrak{n}_4}(\overrightarrow{\gamma^{\mathbf{S}}_i} + \overrightarrow{O}(\ell_2)). \tag{4.20}$$

We have thus verified the string of inequalities ending with (4.19).

To finish the proof of (4.16), note that at this point one of two cases can occur.

- $\|\overrightarrow{\gamma^{\mathbf{S}}_i}\|_1 \leq \frac{\varepsilon_6 \mathscr{C} n}{2^{j_1}}$, which implies $\ell(\gamma^{\mathbf{S}}_i) \leq b\frac{\varepsilon_6 \mathscr{C} n}{2^{j_1}}$ and hence (4.16) is trivially true.

Note that, by construction all the excursions $\widetilde{\gamma}_i$ are large, even though their endpoints might be close. Thus instead of working with the latter, if instead we split up the excursions into subpaths with guaranteed separation between the endpoints, only the second case analysis would have sufficed.

- $\|\overrightarrow{\gamma^{\mathbf{S}}_i}\|_1 \geq \frac{\varepsilon_6 \mathscr{C} n}{2^{j_1}}$, which in particular implies $\|\overrightarrow{\gamma^{\mathbf{S}}_i}\| \geq \frac{1}{\sqrt{2}}\frac{\varepsilon_6 \mathscr{C} n}{2^{j_1}}$. This along with (3.9) implies that

$$\left\|\overrightarrow{\gamma}^{\mathbf{S}}_i\right\|_{(j_1,v_i)} \geq (1-\varepsilon'')\ell(\gamma^{\mathbf{S}}_i).$$

Moreover, since $\|\overrightarrow{\gamma^{\mathbf{S}}_i}\| \geq \frac{1}{\sqrt{2}}\frac{\varepsilon_6 \mathscr{C} n}{2^{j_1}}$, for $m_1$ large enough in the definition of $\ell_2$, the $\overrightarrow{O}(\ell_2)$ term can be made relatively arbitrarily small. Formally, one can again use the triangle inequality property of the norm (Proposition 3.4) to argue

$$\left\|\overrightarrow{\gamma^{\mathbf{S}}_i} + \overrightarrow{O}(\ell_2)\right\|_{(j_1,v_i)} \geq (1-\varepsilon'')\left\|\overrightarrow{\gamma^{\mathbf{S}}_i}\right\|_{(j_1,v_i)}.$$

The above two displays along with (4.19) imply

$$\ell(\widetilde{\gamma}_i) \geq (1-\varepsilon'')\frac{n_1}{n}\ell\big(\gamma_i^{\mathbf{S}}\big),$$

which completes the proof of (4.16) when $v_i \notin A$, by choosing $\varepsilon_6, \varepsilon_7$ small enough compared to $\varepsilon_{11}$ and further choosing $\delta_1$ small enough and $m_1$ large enough.

*Case* 2. $v_i \in A$. In case $\|\overrightarrow{\gamma_i^{\mathbf{S}}}\|_1 \leq \frac{\varepsilon_6 \mathscr{C} n}{2^{j_1}}$, the proof can be completed as above, so we assume otherwise. In this case, since we are on **Fav** and hence on **Boosting**, the conclusion follows from the following inequalities.

$$\ell(\widetilde{\gamma}_i) \geq (b-\varepsilon_7)\|\overrightarrow{\widetilde{\gamma}_i}\|_1, \quad b\|\overrightarrow{\gamma_i^{\mathbf{S}}}\|_1 \geq (1-O(\varepsilon_7))\ell\big(\gamma_i^{\mathbf{S}}\big)$$

and (4.20), by choosing $m_1$ large enough. $\square$

The next three sections prove the three key technical results, Propositions 2.4, 3.4, and 1.4 regarding stability, approximate convexity of the distance function as well as continuity of the rate function. We start with the continuity result.

## 5 Continuity of the Rate Function

In this section we prove Proposition 1.4. Recall that the statement says that for each $\varepsilon > 0$, there exists $\varepsilon' > 0$ such that for all $n$ sufficiently large we have

$$\frac{\log \mathbb{P}(\mathscr{U}_{\zeta-\varepsilon'}(n))}{n^2} \leq \frac{\log \mathbb{P}(\mathscr{U}_{\zeta}(n))}{n^2} + \varepsilon.$$

This is where the assumption of continuous density of the edge distribution will simplify the proof significantly. Moreover, to avoid introducing new notation, we will use several letters in this section that were used earlier to denote different quantities. However, this section will be completely self-contained and hence we expect that this should not create any confusion or conflict.

The basic approach is simply to start with an environment $\Pi \in \mathscr{U}^*_{\zeta-\varepsilon'}(n)$ and then increase the weight of 'all' the edges slightly to construct an environment $\Pi' \in \mathscr{U}_{\zeta}(n)$. However, a technical issue arises since we have assumed that the variables are bounded by a constant $b > 0$. Hence the variables in $\Pi$ that are very close to $b$ cannot be increased. Thus the first step is to localize the set of such really high-valued edges. In fact, we will also localize the set of edges that takes values where the density $f_\nu$ is close to 0.

Let us now formally carry out this strategy. Let $\varepsilon > 0$ be fixed. We shall define a series of parameters $\varepsilon_1$ through $\varepsilon_7$ depending on $\varepsilon$ and $\nu$, and $\varepsilon' > 0$ satisfying the conclusion of Proposition 1.4 will be chosen sufficiently small based on these parameters. Let us first describe how the parameters are chosen; the reason behind these choices will become clear shortly during the course of the proof.

For $\varepsilon_4 > 0$ sufficiently small, let us define

$$\varepsilon_5 = -[\varepsilon_4 \log(\varepsilon_4) + (1-\varepsilon_4)\log(1-\varepsilon_4)],$$

and we shall choose $\varepsilon_4$ sufficiently small so that $\varepsilon_5$ sufficiently close to 0 (depending on $\varepsilon$). This is possible since $\varepsilon_5 \to 0$ as $\varepsilon_4 \to 0$. We shall then choose $\varepsilon_1$ sufficiently small depending on $\varepsilon_4$ and choose $0 < \varepsilon_3 < \varepsilon_2$ sufficiently small depending on $\varepsilon_1$ and $\zeta$ such that the following conditions are satisfied. First, we ask $\mathbb{P}(X_e \in [b-\varepsilon_2, b]) \le \varepsilon_1$ and $c := \frac{\mu+\zeta}{b-\varepsilon_2} < 1$. We also require

$$\inf\{f_\nu(x) : x \in [b-\varepsilon_2, b-\varepsilon_2+\varepsilon_3]\} \ge \varepsilon_3 \tag{5.1}$$

and that $\varepsilon_3$ be sufficiently small compared to $\varepsilon_4$. Observe that existence of such $\varepsilon_2$ and $\varepsilon_3$ are guaranteed by our assumptions on $\nu$. Let

$$\mathbf{B} = \left\{x \in [0, b-\varepsilon_2] : f_\nu(x) \le \frac{\varepsilon_3^3}{b}\right\}. \tag{5.2}$$

Thus by definition $\nu(\mathbf{B}) \le \varepsilon_3^3$. Next, set $\mathbf{D} = [0,b] \setminus \{\mathbf{B} \cup [b-\varepsilon_2, b]\}$ and observe that as a consequence of continuity of $f_\nu$ on $[0,b]$, $f_\nu$ is uniformly away from 0 (at least $\varepsilon_3^3/b$) on $\overline{\mathbf{D}}$.

We shall also choose $\varepsilon_6$ sufficiently small depending on $\varepsilon$ and set $\varepsilon_7 > 0$ such that for any $x \in \mathbf{D}$

$$\frac{1}{1+\varepsilon_6} \le \frac{f_\nu(x+\varepsilon_7)}{f_\nu(x)}. \tag{5.3}$$

The existence of such an $\varepsilon_7$ is guaranteed by the fact that $\overline{\mathbf{D}}$ is compact and hence $f_\nu$ is uniformly continuous on the same. Finally, we set

$$\varepsilon' = \frac{\min(\frac{\varepsilon_3}{2}, \varepsilon_7)(1-c)}{2}.$$

We shall now show that $\varepsilon'$ as above satisfies the conclusion of Proposition 1.4, if $\varepsilon_4$ and $\varepsilon_6$ are chosen sufficiently small depending on $\varepsilon$ and other parameters are chosen as above.

For any $n$, recall the notation $\mathfrak{B}_4 = \mathbf{Box}(\mathscr{C}n)$ from (4.1). We will work with the event $\mathscr{U}^*_{\zeta-\varepsilon'}(n)$ (recall the definition from (2.4)), which is measurable with respect to the edges in $\mathbf{Box}(4\mathscr{C}^2 n)$. However, recall that on $\mathscr{U}^*_{\zeta-\varepsilon'}(n)$, any path from $\mathbf{0}$ that exited $\mathfrak{B}_4$ has length bigger than $bn$; thus it would suffice to increase the value of the edges only inside $\mathfrak{B}_4$.

Let $\mathbf{H}_1 = \{e \in \mathfrak{B}_4 : X_e \in [b-\varepsilon_2, b]\}$. Now by a straightforward union bound over all possible choices of $\mathbf{H}_1$ (at most $2^{O(n^2)}$), for any fixed $\varepsilon_4 > 0$ we have

$$\begin{aligned}\mathbb{P}(|\mathbf{H}_1| \ge \varepsilon_4 n^2) \le 2^{O(n^2)} \varepsilon_1^{\varepsilon_4 n^2} &= e^{O(n^2) - \varepsilon_4 \log(\frac{1}{\varepsilon_1}) n^2} \text{ and hence} \\ &= o\big(\mathbb{P}(\mathscr{U}^*_\zeta(n))\big) \text{ for all small enough } \varepsilon_1 \\ &= o\big(\mathbb{P}(\mathscr{U}^*_{\zeta-\varepsilon'}(n))\big) \text{ for all } \varepsilon' > 0.\end{aligned} \tag{5.4}$$

Similarly, letting $\mathbf{H}_2 = \{e \in \mathfrak{B}_4 : X_e \in \mathbf{B}\}$, we get

$$\begin{aligned}\mathbb{P}(|\mathbf{H}_2| \geq \varepsilon_4 n^2) \leq 2^{O(n^2)} \varepsilon_3^{3\varepsilon_4 n^2} &= e^{O(n^2) - 3\varepsilon_4 \log(\frac{1}{\varepsilon_3}) n^2} \text{ and hence} \\ &= o\big(\mathbb{P}(\mathscr{U}_\zeta^*(n))\big) \text{ for all small enough } \varepsilon_3 \\ &= o\big(\mathbb{P}(\mathscr{U}_{\zeta-\varepsilon'}^*(n))\big) \text{ for all } \varepsilon' > 0.\end{aligned} \tag{5.5}$$

The above allows us to localize $\mathbf{H}_1$ and $\mathbf{H}_2$ without paying too much in the probability. As we have chosen $\varepsilon_1, \varepsilon_3$ sufficiently small depending on $\varepsilon_4$, it follows that

$$\mathbb{P}(\{|\mathbf{H}_1| \leq \varepsilon_4 n^2\} \cap \{|\mathbf{H}_2| \leq \varepsilon_4 n^2\} \cap \mathscr{U}_{\zeta-\varepsilon'}^*(n)) \geq \mathbb{P}(\mathscr{U}_{\zeta-\varepsilon'}(n))(1 - o(1)).$$

Since the total number of subsets of $\mathfrak{B}_4$ of size at most $\varepsilon_4 n^2$ is at most $2^{O(\varepsilon_5) n^2}$, by the pigeonhole principle it follows that there exists subsets $A_1, A_2$ of $\mathfrak{B}_4$, each of size at most $\varepsilon_4 n^2$, such that

$$\mathbb{P}(\{\mathbf{H}_1 = A_1\} \cap \{\mathbf{H}_2 = A_2\} \cap \mathscr{U}_{\zeta-\varepsilon'}^*(n)) \geq \mathbb{P}(\mathscr{U}_{\zeta-\varepsilon'}(n)) e^{-O(\varepsilon_5 n^2)}. \tag{5.6}$$

For easy referencing, let us call the event $\{\mathbf{H}_1 = A_1\} \cap \{\mathbf{H}_2 = A_2\} \cap \mathscr{U}_{\zeta-\varepsilon'}^*(n)$ as $\mathbf{C}$. Now let us modify the event $\mathbf{C}$ to get an event $\mathbf{C}_1$ that will possess the property that

$$\log(\mathbb{P}(\mathbf{C})) - \log(\mathbb{P}(\mathbf{C}_1)) \ll \varepsilon n^2 \quad \text{(we will quantify } \ll \text{ shortly)},$$

and most importantly, $\mathbf{C}_1 \subset \mathscr{U}_\zeta$. Note that by definition $A_1$ and $A_2$ are disjoint. For any weight configuration $\Pi \in \mathbf{C}$, let us set

$$\begin{aligned}\mathbf{C}_1(\Pi) = \{\Pi' : \Pi'(e) &= \Pi(e) \quad \forall e \in A_1, \\ \Pi'(e) &\in \left[b - \varepsilon_2 + \frac{\varepsilon_3}{2}, b - \varepsilon_2 + \varepsilon_3\right] \quad \forall e \in A_2, \\ \Pi'(e) &= \Pi(e) + \varepsilon_7 \quad \forall e \in \mathfrak{B}_4 \setminus A_1 \cup A_2\}.\end{aligned}$$

Let $\mathbf{C}_1 = \bigcup_{\Pi \in \mathbf{C}} \mathbf{C}_1(\Pi)$. We now compute $\mathbb{P}(\mathbf{C}_1)$. For any $\Pi$, and a subset $B$ of edges in $\mathfrak{B}_4$, it would be convenient to let $\Pi|_B$ be the restriction of $\Pi$ on the edges in $B$; for any event $\mathbf{E}$, let

$$\mathbf{E}(B) = \{\Pi|_B : \Pi \in \mathbf{E}\} \quad \text{and} \quad f_\nu(\Pi|_B) := \prod_{e \in B} f_\nu(\Pi(e)).$$

Thus

$$\mathbb{P}(\mathbf{C}) = \int_{\mathbf{C}} f_\nu(\Pi) \mathrm{d}\Pi \leq \varepsilon_3^{3|A_2|} \int_{\mathbf{C}(\mathfrak{B}_4 \setminus A_2)} f_\nu(\Pi|_{\mathfrak{B}_4 \setminus A_2}) \mathrm{d}\Pi|_{\mathfrak{B}_4 \setminus A_2}, \tag{5.7}$$

where the second inequality follows from the definition of $A_2$ and from the fact that our choice of $\varepsilon_3$ guarantees $\nu(\mathbf{B}) \leq \varepsilon_3^3$. Note that by definition for any $\widetilde{\Pi}|_{\mathfrak{B}_4 \setminus A_2} \in \mathbf{C}_1(\mathfrak{B}_4 \setminus A_2)$ there exists a $\Pi|_{\mathfrak{B}_4 \setminus A_2} \in \mathbf{C}(\mathfrak{B}_4 \setminus A_2)$ such that

$$\widetilde{\Pi}|_{\mathfrak{B}_4 \setminus A_2} \geq \Pi|_{\mathfrak{B}_4 \setminus A_2} + \mathbf{v}$$

where

$$\mathbf{v}(e) = \begin{cases} 0 & \text{if } e \in A_1, \\ \varepsilon_7 & \text{if } e \in \mathfrak{B}_4 \setminus \{A_1 \cup A_2\}. \end{cases}$$

Using the above it follows that

$$\begin{aligned} \mathbb{P}(\mathbf{C}_1) &\geq \left(\frac{\varepsilon_3}{2}\right)^{2|A_2|} \int_{\mathbf{C}_1(\mathfrak{B}_4 \setminus A_2)} f_\nu(\Pi|_{\mathfrak{B}_4 \setminus A_2}), \\ &\geq \left(\frac{\varepsilon_3}{2}\right)^{2|A_2|} \left(\frac{1}{1+\varepsilon_6}\right)^{|\mathfrak{B}_4|} \int_{\mathbf{C}(\mathfrak{B}_4 \setminus A_2)} f_\nu(\Pi|_{\mathfrak{B}_4 \setminus A_2}) \\ &\geq e^{-O(\varepsilon_6 + \varepsilon_4)n^2} \mathbb{P}(\mathbf{C}), \end{aligned} \tag{5.8}$$

where the first inequality follows from the definition of $\mathbf{C}_1$ and (5.1), the second inequality is by (5.3), and the final equality is by (5.7) and our choice of parameters. Now, by (5.6) and (5.8) we can choose $\varepsilon_4$ and $\varepsilon_6$ sufficiently small depending on $\varepsilon$ such that $\log \mathbb{P}(\mathscr{U}^*_{\zeta-\varepsilon'}(n)) \leq \log \mathbb{P}(\mathbf{C}_1) + \frac{\varepsilon}{2}n^2$. Using the fact that

$$|\log \mathbb{P}(\mathscr{U}_{\zeta-\varepsilon'}(n)) - \log \mathbb{P}(\mathscr{U}^*_{\zeta-\varepsilon'}(n))| = o(n^2),$$

the proof will now be complete once we show that $\mathbf{C}_1 \subset \mathscr{U}_\zeta(n)$. To do this note that for any $\Pi' \in \mathbf{C}_1$ there exists $\Pi \in \mathbf{C}$ such that $\Pi'(e) \geq \Pi(e) + \min(\frac{\varepsilon_3}{2}, \varepsilon_7)$ for all $\varepsilon \in \mathfrak{B}_4 \setminus A_1$ and $\Pi'(e) = \Pi(e)$ for $e \in A_1$.

Note that since $\Pi \in \mathscr{U}^*_{\zeta-\varepsilon'}(n)$, any path $\mathscr{P}$ starting from the origin, which exits $\mathfrak{B}_4$ has weight at least $bn$ in $\Pi$ and hence by the above discussion also in $\Pi'$. Thus to prove the lemma we only consider the path $\mathscr{P}$, which is the shortest path between $\mathbf{0}$ and $\mathbf{n}$ lying inside $\mathfrak{B}_4$, in the environment $\Pi'$. We want to show $\ell_{\Pi'}(\mathscr{P}) \geq (\mu + \zeta)n$, where $\ell_\Pi(\mathscr{P}), \ell_{\Pi'}(\mathscr{P})$ denote the weights of $\mathscr{P}$ in the environments $\Pi$ and $\Pi'$, respectively.

Now since by construction, $\ell_{\Pi'}(\mathscr{P}) \geq \ell_\Pi(\mathscr{P})$, there is nothing to show if $\ell_\Pi(\mathscr{P}) > (\mu + \zeta)n$. Assuming otherwise, it follows that $|\mathscr{P} \cap A_1| \leq cn$ where $c = \frac{(\mu+\zeta)n}{b-\varepsilon_2} < 1$ was defined above (by $\mathscr{P} \cap A_1$ we denote the set of edges in $A_1$ that $\mathscr{P}$ passes through). Indeed, this is true since each edge in $A_1$ has weight at least $b - \varepsilon_2$. However, note that since $\mathscr{P}$ connects $\mathbf{0}$ and $\mathbf{n}$, trivially $\mathscr{P}$ passes through at least $n$ edges. Thus $|\mathscr{P} \cap A_1^c| \geq (1-c)n$ and hence $\ell_{\Pi'}(\mathscr{P}) - \ell_\Pi(\mathscr{P}) \geq (1-c)n \min(\frac{\varepsilon_3}{2}, \varepsilon_7)$. By definition $\ell_\Pi(\mathscr{P}) \geq (\mu + \zeta - \varepsilon')n$, and hence our choice of $\varepsilon'$ implies the sought bound $\ell_{\Pi'}(\mathscr{P}) \geq (\mu + \zeta)n$.

## 6 Approximate Convexity Properties

In this section we will prove Proposition 3.4, i.e., given $\delta_1$, $m_1$, and $j_1$, for any $\mathbf{Tile}_{\mathscr{C}n}(j_1, v)$ that is $(\delta_1, \ell_1, k_1)$-**Stable** where $\ell_1 = \frac{\mathscr{C}n}{2^{j_1+m_1}}$ and $k = 2^{2m_1}$, and any set of vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_t$, if $\mathbf{w} = \sum_{i=1}^{t} \mathbf{w}_i$, then

$$\|\mathbf{w}\|_{(j_1,v)} \le (1+\delta) \sum_{i=1}^{t} \|\mathbf{w}_i\|_{(j_1,v)} \tag{6.1}$$

where $\delta = O(\delta_1 + 2^{-m_1/16})$. Note that it suffices to only consider $m_1$ sufficiently large, since for smaller values, the above inequality following using (3.9), the standard triangle inequality for the Euclidean norm, and the lower bound and upper bound on the passage times as in (2.3) and the discussion following it. The proof essentially follows by noticing that any set of vectors as above can be scaled down to get a sum of vectors inside $\mathbf{Tile}_{\mathscr{C}n}(j_1, v)$ followed by an application of stability and the triangle inequality. To formalize this, we need some notation: For every $\phi \in \mathbb{S}^1(\eta_1)$ (value of $\eta_1$ will be specified later and be sufficiently small), let

$$\mathcal{C}(\phi) = \{\mathbf{b} \in \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_t\} : \arg(\mathbf{b}) \in [\phi, \phi + \eta_1)\}$$

where $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_t$ are as in the statement of the proposition; i.e., $\mathcal{C}(\phi)$ denotes the collection of vectors among $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_t\}$ whose angle with the $x$-axis falls in the interval $[\phi, \phi + \eta_1)$. Let

$$\mathbf{w}_\phi = \sum_{i=1}^{t} \mathbf{w}_i \mathbf{1}(\mathbf{w}_i \in \mathcal{C}(\phi))$$

be the sum of the vectors in $\mathcal{C}(\phi)$. Thus by definition

$$\sum_{\phi \in \mathbb{S}^1(\eta_1)} \mathbf{w}_\phi = \mathbf{w}.$$

Also, for every $\mathbf{w}_i \in \mathcal{C}(\phi)$, by Lemma 3.2 (and the definitions of $\|\mathbf{w}_i\|_{(j_1,v)}$ and $\nabla_{\mathbf{Proj}}((j_1, v), \phi)$, we have

$$\|\mathbf{w}_i\|_{(j_1,v)} = \|\mathbf{w}_i\| \big(1 + O(\delta_1 + \eta_1 + 2^{-m_1/4})\big) \nabla_{\mathbf{Proj}}((j_1, v), \phi), \tag{6.2}$$

$$\sum_{\mathbf{w}_i \in \mathcal{C}(\phi)} \|\mathbf{w}_i\|_{(j_1,v)} = A_\phi \big(1 + O(\delta_1 + \eta_1 + 2^{-m_1/4})\big) \nabla_{\mathbf{Proj}}((j_1, v), \phi) \tag{6.3}$$

where $A_\phi = \sum_{i=1}^{t} \|\mathbf{w}_i\| \mathbf{1}(\mathbf{w}_i \in \mathcal{C}(\phi))$.

In what follows, for brevity, we shall use the notation $\widetilde{\nabla}_{\mathbf{Proj}}((j_1, v), \phi)$ to denote terms of the form

$$\big(1 + O(\delta_1 + \eta_1 + 2^{-m_1/4})\big) \nabla_{\mathbf{Proj}}((j_1, v), \phi) \tag{6.4}$$

in (6.3) by $\widetilde{\nabla}_{\mathbf{Proj}}((j_1, v), \phi)$. Thus different instances of such usages might denote different quantities, all within a $(1 + O(\delta_1 + \eta_1 + 2^{-m_1/4}))$ multiplicative factor of each other.

Notice now that it follows from the definition of $A_\phi$ that

$$(1 - \eta_1^2) A_\phi < \|\mathbf{w}_\phi\| \le A_\phi. \tag{6.5}$$

Indeed, the upper bound is just the triangle inequality. For the lower bound, notice that if $\theta_i$ is the angle between $\mathbf{w}_i$ and $\mathbf{w}_\phi$, then the projection $\widetilde{\mathbf{w}_i}$ of $\mathbf{w}_i$ in the direction of $\mathbf{w}_\phi$ satisfies $\|\widetilde{\mathbf{w}}_i\| = \cos\theta_i \|\mathbf{w}_i\|$. The lower bound follows by observing that $|\theta_i| \le \eta_1$, choosing $\eta_1$ sufficiently small to ensure $\cos\eta_1 \ge 1 - \eta^2$, and summing over $\mathbf{w}_i \in \mathcal{C}(\phi)$.

Now, (6.5), together with the definition of $\|\mathbf{w}_\phi\|_{(j_1,v)}$, implies

$$(1-\eta_1^2) A_\phi \widetilde{\nabla}_{\mathbf{Proj}}((j_1,v),\phi) \le \|\mathbf{w}_\phi\|_{(j_1,v)} \le A_\phi \widetilde{\nabla}_{\mathbf{Proj}}((j_1,v),\phi) \tag{6.6}$$

where the term $\widetilde{\nabla}_{\mathbf{Proj}}((j_1,v),\phi)$ is as indicated above.

For each $\phi$, let us consider the value $b_\phi = \frac{A_\phi}{\|\mathbf{w}\|}$. We first claim that without loss of generality we can assume that there exists a universal constant $C$ sufficiently large such that

$$b_\phi \le C \tag{6.7}$$

for all $\phi$. Otherwise, the fact that $\nabla_{\mathbf{Proj}}((j_1,v),\phi)$ is bounded away from 0 and infinity for any $\phi$ (see (3.10)), together with (6.3) and the definition of $\|\mathbf{w}\|_{(j_1,v)}$, will imply (6.1).

Since our proof strategy relies on using the connection to passage times that breaks down for vectors having very small Euclidean norms, we will ignore the vectors in $\mathcal{C}(\phi)$ with small $b_\phi$. Towards formalizing this, we now define the set $\mathcal{B} = \{\phi \in \mathbb{S}^1(\eta_1) : b_\phi \le 2^{-m_1/5}\}$ and hence by setting $\eta_1 = 2^{-m_1/8}$ we get

$$\Big\|\sum_{\phi\in\mathcal{B}} \mathbf{w}_\phi\Big\| \le \frac{2\pi}{\eta_1} \times 2^{-m_1/5}\|\mathbf{w}\| = O(2^{-m_1/16})\|\mathbf{w}\|. \tag{6.8}$$

Now let

$$\mathbf{c}_\phi = \frac{\mathbf{w}_\phi}{\|\mathbf{w}\|}\frac{\mathscr{C}n}{C\times 100\times 2^{j_1}}\frac{\widetilde{\mu}}{b} \quad\text{and}\quad \mathbf{c} = \frac{\mathbf{w}}{\|\mathbf{w}\|}\frac{\mathscr{C}n}{C\times 100\times 2^{j_1}}\frac{\widetilde{\mu}}{b}. \tag{6.9}$$

Thus we have rescaled $\mathbf{w}$ to get a vector $\mathbf{c}$ of length $\frac{\mathscr{C}n}{C\times 100\times 2^{j_1}}\frac{\widetilde{\mu}}{b}$ and scaled all the $\mathbf{w}_\phi$'s by the same factor to obtain the $\mathbf{c}_\phi$'s. The constant $\widetilde{\mu}$ here is the one appearing in (3.5), $C$ appears in (6.7), and $b$ is the upper bound on the individual vertex weights.

For convenience, let us denote $\mathbb{S}^1(\eta_1) \setminus \mathcal{B} = \{\phi_1, \phi_2, \dots, \}$. We now consider the sequence of points $\mathbf{v}_0, \mathbf{v}_1, \dots$ such that $\mathbf{v}_i - \mathbf{v}_{i-1} = \mathbf{c}_{\phi_i}$ and let, for concreteness, $\mathbf{v}_0$ be the center point of $\mathbf{Tile}_{\mathscr{C}n}(j_1, v)$. Notice that for $m_1$ sufficiently large, we have $\|\mathbf{c}_{\phi_i}\| \ge \frac{\mathscr{C}n}{2^{j_1+m/4}}$. Now consider the path $\gamma$ obtained by concatenation of paths $\gamma_1, \gamma_2, \dots$ where $\gamma_i$ is the shortest path between $\mathbf{v}_{i-1}$ and $\mathbf{v}_i$. There are two cases to be considered.

*Case* 1. Each $\mathbf{v}_i$ is a point in $\mathbf{Tile}_{\mathscr{C}n}(j_1, v)$. In this case, since $\mathbf{Tile}_{\mathscr{C}n}(j_1, v)$ is $(\delta_1, \ell_1, k_1)$-**Stable** as in the hypothesis of the proposition, it follows, using the lower bound on $\|\mathbf{c}_{\phi_i}\|$, (3.4), (3.7), and Lemma 3.2 that

$$\ell(\gamma_i) \le \|\mathbf{c}_{\phi_i}\| \widetilde{\nabla}_{\mathbf{Proj}}((j_1,v),\phi_i).$$

Hence it follows that

$$\ell(\gamma) = \sum_i \ell(\gamma_i) \le \big(1 + O(\delta_1 + \eta_1 + 2^{-\frac{m_1}{4}})\big) \sum_i \|\mathbf{c}_{\phi_i}\|_{(j_1,v)}. \tag{6.10}$$

Observe now that $\gamma$ is a path (not necessarily the shortest) joining $\mathbf{v}_0$ and $\mathbf{v}_0 + \mathbf{c}_*$, where $\mathbf{c}_* = \sum_i \mathbf{c}_{\phi_i}$. Observe also that from (6.9) and (6.8) we have $\|\mathbf{c} - \mathbf{c}_*\| \le O(2^{-m_1/16})\|\mathbf{c}\|$, and hence we also have $\|\mathbf{c}_*\| \ge \frac{\mathscr{C}n}{2^{j_1+m/4}}$. Using the stability of $\mathbf{Tile}_{\mathscr{C}n}(j_1, v)$, this implies

$$\ell(\gamma) \ge \big(1 - O(\delta_1 + \eta_1 + 2^{-\frac{m_1}{4}})\big)\|\mathbf{c}_*\|_{(j_1,v)}. \tag{6.11}$$

Now on account of the closeness between $\mathbf{c}_*$ and $\mathbf{c}$, by Lemma 3.2 we also have

$$\|\mathbf{c}_*\|_{(j_1,v)} \ge \big(1 - O(\delta_1 + 2^{-m_1/16} + 2^{-\frac{m_1}{4}})\big)\|\mathbf{c}\|_{(j_1,v)}. \tag{6.12}$$

Putting the above together (letting $1+B = 1 + O(\delta_1 + \eta_1 + 2^{-m_1/4})$ appearing in (6.4)) it follows that

$$\begin{aligned}
\sum_{i=1}^{t} \|\mathbf{w}_i\|_{(j_1,v)} &\overset{(6.2)(6.3)(6.6)}{\ge} \big(1-\eta_1^2\big)(1-B) \sum_{\phi\in\mathbb{S}^1(\eta_1)} \|\mathbf{w}_\phi\|_{(j_1,v)} \\
&\ge (1-\eta_1^2)(1-B) \sum_{\phi\in\mathbb{S}^1(\eta_1)\setminus\mathcal{B}} \|\mathbf{w}_\phi\|_{(j_1,v)}, \\
&= (1-\eta_1^2)(1-B) \sum_{\phi\in\mathbb{S}^1(\eta_1)\setminus\mathcal{B}} \|\mathbf{c}_\phi\|_{(j_1,v)} \frac{C \times 100 \times 2^{j_1} b\|\mathbf{w}\|}{\widetilde{\mu}\mathscr{C}n}, \\
&\overset{(6.10),(6.11),(6.12)}{\ge} (1-\eta_1^2)(1-B)^2\big(1 - B - O\big(2^{-\frac{m_1}{16}}\big)\big) \\
&\qquad \cdot \|\mathbf{c}\|_{(j_1,v)} \frac{C \times 100 \times 2^{j_1} b\|\mathbf{w}\|}{\widetilde{\mu}\mathscr{C}n} \\
&\ge (1 - O(\delta_1 + 2^{-\frac{m_1}{16}}))\|\mathbf{w}\|_{(j_1,v)}
\end{aligned}$$

where the final inequality follows by recalling that $\eta_1 = 2^{-m_1/8}$.

*Case* 2. Let $j$ be the first index such that $\mathbf{v}_j \notin \mathbf{Tile}_{\mathscr{C}n}(j_1, v)$. In this case, using the same reasoning as in (6.10) for the path $\widetilde{\gamma} = \gamma_1\gamma_2\cdots\gamma_{j-1}$ yields

$$\ell(\widetilde{\gamma}) = \sum_{i=1}^{j-1} \ell(\gamma_i) \le \big(1 + O(\delta_1 + \eta_1 + 2^{-\frac{m_1}{4}})\big) \sum_i \|\mathbf{c}_{\phi_i}\|_{(j_1,v)}. \tag{6.13}$$

On the other hand, observe that $\mathbf{v}_j \notin \mathbf{Tile}_{\mathscr{C}n}(j_1, v)$ and (6.7) implies $\|\mathbf{v}_j - \mathbf{v}_{j-1}\| \le C\|\mathbf{c}\|$, and hence it follows that the Euclidean distance between the endpoints of $\widetilde{\gamma}$, $\mathbf{v}_0$, and $\mathbf{v}_{j-1}$ is at least

$$\left(1 - \frac{\widetilde{\mu}}{100 \times b}\right)\frac{\mathscr{C}n}{2^{j_1}}.$$

We now get from (3.5) that

$$\ell(\widetilde{\gamma}) \geq \widetilde{\mu}\left(1 - \frac{\widetilde{\mu}}{100 \times b}\right)\frac{\mathscr{C}n}{2^{j_1}}.$$

Finally, by (3.10) and the above, we get $\|\mathbf{c}\|_{(j_1,v)} \leq \sqrt{2}b\frac{\mathscr{C}n}{C\times 100\times 2^{j_1}}\frac{\widetilde{\mu}}{b} \leq \ell(\widetilde{\gamma})$, which, together with (6.13), completes the proof in this case. □

## 7 Stability of the Gradient

This section is devoted to proving Proposition 2.4. It turns out that this property has little to do with the specific details of the first passage percolation metric; rather it is a property of general distance functions on $\mathbb{R}^2$ that are comparable to the Euclidean metric; i.e., it satisfies the triangle inequality and that for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ such that $\|\mathbf{x}-\mathbf{y}\|$ is large enough (possibly $n$ dependent when considering boxes of size $O(n)$)

$$\widetilde{\mu}\|\mathbf{x}-\mathbf{y}\| \leq \mathbf{PT}(\mathbf{x},\mathbf{y}) \leq 3b\|\mathbf{x}-\mathbf{y}\|, \tag{7.1}$$

which in our case is a consequence of (2.3) and (3.1). Although, as explained in the introduction, we believe that results of similar flavour (and with essentially similar proofs) would be useful for studying a larger class of models, e.g., last passage percolation and positive temperature polymer models, we did not find a model where this result would be directly quotable, and hence we have decided to not introduce extra notations to write the result in its most general form.

For the ease of reading, we recall the statement of the proposition. Recall our terminology that $\mathbf{z} \in \mathbb{R}^2$ is $(\delta, \mathbb{S}^1(\eta), \ell, k)$-**Stable** if $\mathbf{z}$ is $(\delta, \theta, \ell, k)$-**Stable** for each $\theta \in \mathbb{S}^1(\eta)$.

PROPOSITION 7.1. *Fix $\delta, \varepsilon, \eta > 0$, $k \in \mathbb{N}$, and $J_1 \in \mathbb{N}$ and suppose that* (7.1) *holds for all $\mathbf{x}, \mathbf{y} \in \mathbf{Box}(10n)$ such that $\|\mathbf{x}-\mathbf{y}\| \geq \sqrt{n}$. Then there exists $J_2 \in \mathbb{N}$ such that for all large enough n, there exists $J_1 \leq j \leq J_2$ such that*

$$\#\left\{\mathbf{z} \in \mathbb{Z}^2 \cap \mathbf{Box}(n) : \mathbf{z} \textit{ is not } \left(\delta, \mathbb{S}^1(\eta), \frac{n}{2^j}, k\right)\text{-}\mathbf{Stable}\right\} \leq \varepsilon n^2.$$

Above we have replaced $\mathscr{C}n$ in the statement of Proposition 2.4 by $n$, and similarly the conditioning on $\mathscr{U}_{\zeta}^*(n)$ therein by simply assuming that (7.1) holds for well separated points in $\mathbf{Box}(10n)$ to reduce notational overhead; the reader will notice the arguments imply Proposition 2.4 by simply changing certain constants.

### 7.1 A roadmap of the proof

Since we are not seeking optimal bounds, the proofs will often rely on several crude averaging arguments and applications of the pigeonhole principle along with the bi-Lipschitz nature of the FPP metric. However, there are many technical steps involved, and for the sake of exposition we give a brief overview of the argument at this point.

The argument is divided into two parts with similar proofs. The first part shows that for a fixed direction $\theta \in \mathbb{S}^1$ and a fixed point $\mathbf{z} \in \mathbf{Box}(n)$, and for $J_1 \in \mathbb{N}$ and for $J_2$ sufficiently large, there exists $\mathfrak{m} \in \mathbb{N}$ and $J_1 \le j \le J_2$, and such that most points $\mathbf{z}_i$ in the discrete segment $\mathscr{U}(\mathbf{z}, \theta, \frac{n}{2^{j\mathfrak{m}}}, 2^{j\mathfrak{m}})$ are stable with parameters depending on $j$ and $\mathfrak{m}$. This is carried out in Section 7.2. (See Lemma 7.3 for a precise statement.) The second part of the argument strengthens this result to Proposition 7.1 by showing that a common $j$ can be found for all directions and all points. The two parts of the arguments are similar; we describe here a roadmap of the proof of Lemma 7.3, and the second part of the argument is provided in Section 7.3; see the beginning of that subsection for a discussion of the extra ingredients needed.

Our argument relies on the following observations.

(1) Fix $\mathbf{z} \in \mathbf{Box}(n)$ and $\theta \in \mathbb{S}^1(\eta)$. Observe that for all $J_2 > J_1$,

$$\begin{aligned}&\mathbf{PT}\Big(\mathbf{z}, \theta, \frac{n}{2^{J_2}}, 2^{J_2}\Big) - \mathbf{PT}\Big(\mathbf{z}, \theta, \frac{n}{2^{J_1}}, 2^{J_1}\Big)\\ &\quad = \sum_{j=J_1}^{J_2-1} \Big[\mathbf{PT}(\mathbf{z}, \theta, \frac{n}{2^{j+1}}, 2^{j+1}) - \mathbf{PT}\Big(\mathbf{z}, \theta, \frac{n}{2^j}, 2^j\Big)\Big].\end{aligned} \tag{7.2}$$

The LHS in (7.2) is trivially bounded by $3bn$, and all the terms in the RHS are nonnegative by the triangle inequality (as in Lemma 2.3).

(2) Thus, if $J_2 - J_1 \ge \frac{1}{\varepsilon}$, by the pigeonhole principle there must exist one $J_1 \le j \le J_2$ such that $[\mathbf{PT}(\mathbf{z}, \theta, \frac{n}{2^{j+1}}, 2^{j+1}) - \mathbf{PT}(\mathbf{z}, \theta, \frac{n}{2^j}, 2^j)] \le O(\varepsilon)n$. As a matter of fact, if further $J_2 - J_1 \gg \frac{1}{\varepsilon}$, we should be able to find many consecutive integers $j$ satisfying the above property.

(3) Now for $j$ as in (2), consider the discrete segments

$$\begin{aligned}\mathscr{U}\left(\mathbf{z}, \theta, \frac{n}{2^j}, 2^j\right) &= [\mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_{2^j}]\\ \mathscr{U}\left(\mathbf{z}, \theta, \frac{n}{2^{j+1}}, 2^{j+1}\right) &= [\mathbf{z}_0, \mathbf{z}_{0,1}, \mathbf{z}_1, \mathbf{z}_{1,2}, \mathbf{z}_2, \dots, \mathbf{z}_{2^j}],\end{aligned}$$

where $\mathbf{z}_{i,i+1}$ is the midpoint of the line segment joining $\mathbf{z}_i$ and $\mathbf{z}_{i+1}$. Thus the above observation together with the lower bound in (7.1) suggests that for most $i$,

$$\mathbf{PT}(\mathbf{z}_i, \mathbf{z}_{i,i+1}) + \mathbf{PT}(\mathbf{z}_{i,i+1}, \mathbf{z}_{i+1}) \le (1 + O(\varepsilon))\mathbf{PT}(\mathbf{z}_i, \mathbf{z}_{i+1}).$$

However, this is not quite enough to establish stability, and in fact we need something along the lines of the following stronger fact (see Lemma 7.2): for most $i$,

$$\mathbf{PT}(\mathbf{z}_i, \mathbf{z}_{i,i+1}) \approx \mathbf{PT}(\mathbf{z}_{i,i+1}, \mathbf{z}_{i+1}) \approx \frac{1}{2}\mathbf{PT}(\mathbf{z}_i, \mathbf{z}_{i+1}).$$

(4) Suppose the contrary to the expression directly above. Without loss of generality assume that

$$\mathbf{PT}(\mathbf{z}_i, \mathbf{z}_{i,i+1}) \geq \left(\frac{1}{2} + \delta\right)\mathbf{PT}(\mathbf{z}_i, \mathbf{z}_{i+1}).$$

The contradiction will come from the fact that the above cannot be true for many consecutive scales. Indeed, if it were true for $j'$ many consecutive scales, then recursively picking one half of an interval at each scale in which the above inequality holds leads to an interval $[\mathbf{w}_1, \mathbf{w}_2]$ such that $\|\mathbf{w}_1 - \mathbf{w}_2\| = \frac{\|\mathbf{z}_i - \mathbf{z}_{i+1}\|}{2^{j'}}$ but

$$\mathbf{PT}(\mathbf{w}_1, \mathbf{w}_2) \geq (1 + 2\delta)^{j'} \frac{\mathbf{PT}(\mathbf{z}_i, \mathbf{z}_{i+1})}{2^{j'}}.$$

Clearly, using the lower bound in (7.1), for $j'$ large enough (depending on $\delta$, $\widetilde{\mu}$, and $b$), this contradicts the upper bound in (7.1). We now make the argument above formal.

Recalling the notion of stability from (2.7), the following crude lemma shows how the above observations are useful to show stability.

LEMMA 7.2. *Let $\delta > 0$, $\theta \in S^1$, and $\ell, k \in \mathbb{N}$ be fixed. Recalling that*

$$\mathscr{U}(\mathbf{z}, \theta, \ell, k) = [\mathbf{z} = \mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_k],$$

*suppose*

$$\sup_{0 \leq i,j \leq k-1} \frac{\mathbf{PT}(\mathbf{z}_i, \mathbf{z}_{i+1})}{\mathbf{PT}(\mathbf{z}_j, \mathbf{z}_{j+1})} \leq 1 + \delta \quad \textit{and}$$

$$\frac{k\mathbf{PT}(\mathbf{z}_0, \mathbf{z}_1)}{1 + \delta} \leq \mathbf{PT}(\mathbf{z}_0, \mathbf{z}_k) \leq k\mathbf{PT}(\mathbf{z}_0, \mathbf{z}_1)(1 + \delta).$$

*Then, for each $i \leq k$ and $k' \leq k - i$, $\mathbf{z}_i$ is $(\delta', \theta, \ell, k')$-***Stable** *where $\delta' = O(\delta k)$.*

PROOF. Using the hypotheses, for any $i \leq k$ and $k' \leq k - i$, we have

$$\mathbf{PT}(\mathbf{z}_0, \mathbf{z}_k) \leq \mathbf{PT}(\mathbf{z}_i, \mathbf{z}_{i+k'}) + (k - k')(1 + \delta)\mathbf{PT}(\mathbf{z}_0, \mathbf{z}_1).$$

Thus it follows that

$$\begin{aligned}\mathbf{PT}(\mathbf{z}_i, \mathbf{z}_{i+k'}) &\geq \frac{1}{(1 + \delta)} k\mathbf{PT}(\mathbf{z}_0, \mathbf{z}_1) - (k - k')(1 + \delta)\mathbf{PT}(\mathbf{z}_0, \mathbf{z}_1)\\ &\geq k'\mathbf{PT}(\mathbf{z}_0, \mathbf{z}_1)(1 - O(\delta k)),\\ &\geq k'\mathbf{PT}(\mathbf{z}_i, \mathbf{z}_{i+1})(1 - O(\delta k)).\end{aligned}$$

Moreover, note that by the triangle inequality and the hypothesis, $\mathbf{PT}(\mathbf{z}_i, \mathbf{z}_{i+k'}) \leq (1 + \delta)k'\mathbf{PT}(\mathbf{z}_i, \mathbf{z}_{i+1})$. This completes the proof. □

Thus, in what follows, to prove the stability of a point we will only prove that the hypothesis of Lemma 7.2 is satisfied following the strategy outlined above. Going back to the proof of Proposition 7.1, as explained in the roadmap, we shall first deduce the existence of many stable points along a fixed line in a given direction. We shall first state and prove this weaker version.

### 7.2 Stability on a fixed line

Let us consider the discrete segment $\mathscr{U}(\mathbf{z}, \theta, \frac{n}{2^j}, 2^j)$ for some $\mathbf{z} \in \mathbf{Box}(n)$ and $\theta \in \mathbb{S}^1(\eta)$. We shall show most points on this segment are stable for $k$ consecutive intervals.

LEMMA 7.3. *Let $\mathbf{z} \in \mathbf{Box}(n)$, $\theta \in \mathbb{S}^1(\eta)$, $k \in \mathbb{N}$, and $J_1 \in \mathbb{N}$ be fixed. There exists an absolute constant $C_0 > 0$ such that for each $\delta_2 > 0$, there exists $\mathfrak{m}$, depending on $\delta_2$ and $k$ only, for which the following holds: for all small enough $\delta_3$ (depending on $\mathfrak{m}$ and $\delta_2$) and for all $n$ large enough depending on all the other parameters, there exists $j \in \mathbb{N}$ with $J_1 \leq j \leq (J_1 + 1/\delta_3^2)$ for which all but a $C_0\delta_2$ fraction of the points $\mathbf{z}_i$ in the discrete segment $\mathscr{U}(\mathbf{z}, \theta, n/2^{j\mathfrak{m}}, 2^{j\mathfrak{m}})$ are $(\delta_2, \theta, n/2^{j\mathfrak{m}}, k)$-**Stable**.*

The quantification in the above statement might be a little hard to parse, but it will create some simplification in the notational choices later. For the moment, let us fix a value of $\mathfrak{m}$ to be specified later. It will be convenient to associate trees to the intervals in $\mathscr{U}(\mathbf{z}, \theta, n/2^{\mathfrak{m}J_1}, 2^{\mathfrak{m}J_1}) = [\mathbf{z} = \mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_{2^{\mathfrak{m}J_1}}]$. Let

$$T_1, T_2, \ldots, T_{2^{\mathfrak{m}J_1}} \tag{7.3}$$

be complete $2^{\mathfrak{m}}$-ary trees of depth $J_2 - J_1$, where the value of $J_2$ will be specified to be a large enough number later (for convenience we shall index the levels of these trees by $j = J_1, J_1 + 1, \ldots, J_2$). Let $L_j^{(i)}$ denote the vertices at the $j^{\text{th}}$ level of $T_i$, and let $L_j = \bigcup_i L_j^{(i)}$ denote the union of the vertices at the $j^{\text{th}}$ level. We will identify $T_i$ with the interval $[\mathbf{z}_{i-1}, \mathbf{z}_i]$. Now for any $J_1 \leq j \leq J_2$, consider the discrete segment

$$\mathscr{U}\left(\mathbf{z}, \theta, \frac{n}{2^{j\mathfrak{m}}}, 2^{j\mathfrak{m}}\right) = \left[\mathbf{z}_0^{*,j}, \ldots \mathbf{z}_{2^{(j-J_1)\mathfrak{m}}}^{*,j}, \mathbf{z}_{2^{(j-J_1)\mathfrak{m}}+1}^{*,j}, \ldots, \mathbf{z}_{2^{(j+1-J_1)\mathfrak{m}}}^{*,j}, \ldots, \mathbf{z}_{2^{(j-1)\mathfrak{m}}}^{*,j} \ldots \mathbf{z}_{2^{j\mathfrak{m}}}^{*,j}\right].$$

Naturally,

$$\left[\mathbf{z}_0^{*,j}, \ldots, \mathbf{z}_{2^{(j-J_1)\mathfrak{m}}}^{*,j}\right]$$

is a discretization of the interval $[\mathbf{z}_0, \mathbf{z}_1]$, and hence can be associated to $L_j^{(1)}$, where each vertex in $L_j^{(1)}$ corresponds to $[\mathbf{z}_h^{*,j}, \mathbf{z}_{h+1}^{*,j}]$ in the natural order (e.g., the root of $T_1$ corresponds to the interval $[\mathbf{z}_0^{*,J_1}, \mathbf{z}_1^{*,J_1}]$, which is nothing but the interval $[\mathbf{z}_0, \mathbf{z}_1]$). The same correspondence holds for the other intervals and trees. See Figure 7.1 for an illustration. In what follows, the level of the trees under consideration would be clear, and hence to avoid cluttering the notation, we will suppress the $j$ superscript.

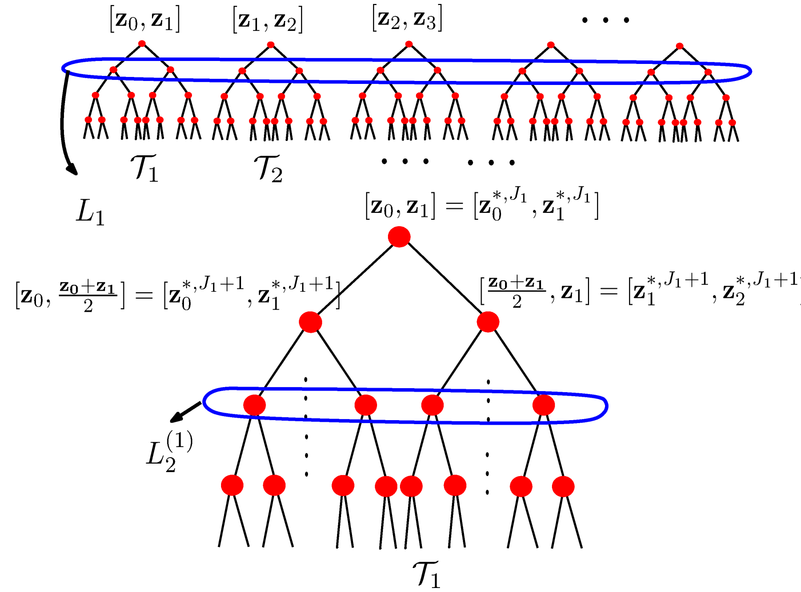


FIGURE 7.1. This figure illustrates the various definitions introduced in this section related to the trees in (7.3) in the case $\mathfrak{m} = 1$ where the trees are binary trees.

Now for any vertex $v$ in any of the trees, let $Y_v := \mathbf{PT}(\mathbf{z}', \mathbf{z}'')$ where $[\mathbf{z}', \mathbf{z}'']$ is the discrete segment associated to the vertex $v$. We need some further notation: let

$$U_{i,j} = \sum_{v \in L_j^{(i)}} Y_v \quad \text{and} \quad U_j = \sum_i U_{i,j}.$$

It will be convenient to frame our arguments using the pigeonhole principle as applications of 'the probabilistic method', and hence we define a set of random variables. For any $j$, pick uniformly any edge $e_{j+1} = (v, w)$ at the $(j+1)^{\text{th}}$ level across all the trees, i.e., connecting $L_j$ and $L_{j+1}$, where $v$ is closer to the root, and let

$$X_{e_{j+1}} := \frac{2^{\mathfrak{m}} Y_w}{Y_v}. \tag{7.4}$$

For brevity we will identify the set of such edges with the set $L_{j+1}$ using the natural correspondence. Now by the triangle inequality (again, as in Lemma 2.3)

$$\mathbb{E}(X_{e_{j+1}} \mid v, Y_v) \geq 1. \tag{7.5}$$

However, the distributions of $X_{e_j}$ across various $j$ will not be independent, and the joint distribution can be defined in the following way: pick uniformly a vertex among all the leaf vertices across all the trees (note that it is naturally and uniquely associated with a uniformly chosen simple path from the root to the leaf in a uniformly chosen tree) and label the edges on the path as $(e_{J_1+1}, e_{J_1+2}, \dots, e_{J_2})$ where $e_i$ denotes the intersection of the path with the $i^{\text{th}}$ level. It is clear that $e_j$ is uniformly distributed among all edges connecting $L_{j-1}$ and $L_j$. Notice that

$$\prod_{j=J_1+1}^{J_2} X_{e_j} \stackrel{d}{=} \frac{2^{(J_2-J_1)\mathfrak{m}} Y_w}{Y_v}$$

where $Y_v$ and $Y_w$ are the variables attached to the root of a randomly chosen tree $T_i$ and a randomly chosen leaf of $L_{J_2}^{(i)}$, respectively. We now bound the expectation of $X_{e_j}$ for all $J_1 < j \le J_2$. To do this consider the ratio $U_{j+1}/U_j$. By definition, we have the following:

$$\frac{U_{j+1}}{U_j} = \frac{\sum_{w \in L_{j+1}} Y_w}{\sum_{v \in L_j} Y_v} = \frac{\sum_{v \in L_j} Y_v \mathbb{E}(X_{e_{j+1}} \mid v, Y_v)}{\sum_{v \in L_j} Y_v} \ge 1 \tag{7.6}$$

where the second equality follows from (7.4) and the fact that the trees $T_i$ are $2^{\mathfrak{m}}$-ary, and the final inequality follows from (7.5). The following lemma completes the proof of Lemma 7.3 under the further assumption that on a sufficiently large interval contained in $[\![J_1, J_1 + 1/\delta_3^2]\!]$, $U_{j+1}/U_j$ is also upper-bounded by $1 + \delta_3$ for some small enough $\delta_3$.

LEMMA 7.4. *Fix $c > 0$. In the setting of Lemma 7.3, suppose there exists an interval $I \subseteq [\![J_1, J_1 + 1/\delta_3^2]\!]$, with $|I| \ge c/\delta_3$ such that for all $j \in I$, for some small enough $\delta_3$, depending on $\mathfrak{m}$ and $\delta_2$,*

$$0 \le \frac{U_{j+1}}{U_j} - 1 \le \delta_3. \tag{7.7}$$

*Then the conclusion of Lemma 7.3 holds.*

PROOF. Without loss of generality for this proof we shall write $I = [\![J_1, J_2]\!]$, where $I$ is given by the hypothesis. It is a consequence of (7.1) that for $j \le k$, $v \in L_j$ and $w \in L_k$,

$$\frac{1}{C} \le \frac{Y_v}{2^{(k-j)\mathfrak{m}} Y_w} \le C \tag{7.8}$$

for some universal constant $C = C(b, \widetilde{\mu}) > 1$. Define now the probability measure $\mu_j$ on the $j^{\text{th}}$-level vertices given by $\mu_j(v) = Y_v/(\sum_{v \in L_j} Y_v)$. In particular, for $k = j$, (7.8) implies that the Radon-Nikodym derivative of $\mu_j$ with respect to the uniform measure $\mathfrak{u}_j$ on $L_j$ is bounded above and below by $C^2$ and $C^{-2}$, respectively. Now, (7.7), along with (7.6), implies $\mathbb{E}_{\mu_j}(\mathbb{E}(X_{e_{j+1}} \mid v, Y_v) - 1) \le \delta_3$. This, together with the above observation and (7.5) (the fact that $\mathbb{E}(X_{e_{j+1}} \mid v, Y_v) - 1 \ge 0$) implies that

$$\mathbb{E}_{\mathfrak{u}_j}(\mathbb{E}(X_{e_{j+1}} \mid v, Y_v) - 1) \le C^2 \delta_3. \tag{7.9}$$

By (7.1), $C$ in (7.8) can be chosen such that deterministically $\frac{1}{C} \le X_{e_j} \le C$ and moreover,

$$\frac{1}{C} \le \prod_{j=J_1+1}^{J_2} X_{e_j} \le C, \text{ which implies } \left| \sum_{j=J_1+1}^{J_2} \mathbb{E}_{\mathfrak{u}}(\log X_{e_j}) \right| \le \log C. \tag{7.10}$$

Thus, it follows that there exists $J_1 + 1 \le j \le J_2$ such that

$$\mathbb{E}_{\mathfrak{u}_j}(\log X_{e_j}) \ge -c^{-1}(\log C)\delta_3.$$

Hence we have found a $J_1 + 1 \le j \le J_2$ with the following two properties:

$$1 \le \mathbb{E}_{\mathfrak{u}_j}(X_{e_j}) \overset{(7.9)}{\le} 1 + C^2\delta_3 \quad \text{and} \quad \mathbb{E}_{\mathfrak{u}_j}(\log(X_{e_j})) \ge -c^{-1}(\log C)\delta_3.$$

Now for any edge $e$, denoting $X_e - 1 = y_e$, the above can be restated as

$$0 \le \frac{1}{2^{j\mathfrak{m}}} \sum_{e \in L_j} y_e \le C^2\delta_3 \quad \text{and} \quad \frac{1}{2^{j\mathfrak{m}}} \sum_{e \in L_j} \log(1 + y_e) \ge -c^{-1}(\log C)\delta_3.$$

Now note that by (7.8), $-1 \le y_e \le C$, and hence using Taylor expansion, $\log(1 + y_e) \le y_e - C' y_e^2$ for some universal constant $C'$. Using the above inequalities it follows that

$$\mathbb{E}_{\mathfrak{u}_j}(X_{e_j} - 1)^2 = \frac{1}{2^{j\mathfrak{m}}} \sum_{e \in L_j} y_e^2 \le \frac{1}{C'}\left[\frac{1}{2^{j\mathfrak{m}}} \sum_{e \in L_j} \big(y_e - \log(1 + y_e)\big)\right] = O(\delta_3).$$

Thus by Chebyshev's inequality, for at least a $1 - O(\sqrt{\delta_3})$ fraction of $e \in L_j$, we have $|X_e - 1| \le \delta_3^{1/4}$. Let us call such an edge $e$, a good edge. Now let us consider all $v \in L_{j-1}$ such that all the children of $v$ are good (let us call such $v$ good). A naive bound shows that the fraction of good $v$ is at least $1 - O(2^{\mathfrak{m}}\sqrt{\delta_3})$. Now for any good $v$ corresponding to an interval $[\mathbf{w}_1, \mathbf{w}_2]$, say, if the discrete segment $[\mathbf{w}_1 = \mathbf{w}_0^*, \mathbf{w}_1^*, \ldots, \mathbf{w}_{2^{\mathfrak{m}}}^* = \mathbf{w}_2]$ corresponds to the $2^{\mathfrak{m}}$ children, then Lemma 7.2 implies the following: each $\mathbf{w}_i^*$ for $i \in [\![0, 2^{\mathfrak{m}} - k]\!]$ is $(\delta', \theta, n/2^{j\mathfrak{m}}, k)$-**Stable**, where $\delta' = O(2^{\mathfrak{m}}\delta_3^{1/4})$ (note that $k$ here is the same as in Lemma 7.3, and not that in the statement of Lemma 7.2; the latter is applied by setting $k = 2^{\mathfrak{m}}$). Thus the total fraction of points on $\mathscr{U}[\mathbf{z}, \theta, n/2^{j\mathfrak{m}}, 2^{j\mathfrak{m}}]$ that are not $(\delta', \theta, n/2^{j\mathfrak{m}}, k)$-**Stable** is at most $O(k/2^{\mathfrak{m}} + 2^{\mathfrak{m}}\sqrt{\delta_3})$. Now choose $\mathfrak{m}$ large enough and then $\delta_3$ small enough such that $\max(k/2^{\mathfrak{m}} + 2^{\mathfrak{m}}\sqrt{\delta_3}, \delta') \le \delta_2$. $\square$

It remains to prove that (7.7) holds for a large number of consecutive scales. This is ensured by the following lemma using another pigeonhole argument.

LEMMA 7.5. *In the setting of Lemma* 7.4, *there exists* $c > 0$ *and* $I \subseteq [\![J_1, J_1 + 1/\delta_3^2]\!]$ *with* $|I| \ge \frac{c}{\delta_3}$ *such that for all* $j \in I$

$$0 \le \frac{U_{j+1}}{U_j} - 1 \le \delta_3.$$

PROOF. For $c > 0$ to be specified later, we divide the $1/\delta_3^2$ many scales into consecutive blocks of $c/\delta_3$ many scales each. For $i \in [\![1, 1/c\delta_3]\!]$, let $a_i = U_{J_1+ic/\delta_3} - U_{J_1+(i-1)c/\delta_3}$. By the triangle inequality, $a_i \geq 0$ for all $i$, and by (7.8), there exists a universal constant $C$ such that $U_{J_1+1/\delta_3^2} \leq CU_{J_1}$. As a consequence, $\sum_{i\in[\![1,1/c\delta_3]\!]} a_i \leq CU_{J_1}$, and by choosing $c$ sufficiently small, it follows there exists some $i \in [\![1, 1/c\delta_3]\!]$ such that $a_i \leq U_{J_1}\delta_3$. Now this implies that for any $J_1 + (i-1)c/\delta_3 \leq j \leq J_1 + ic/\delta_3$ we have

$$\frac{U_{j+1} - U_j}{U_j} \leq \frac{a_i}{U_{J_1}} \leq \delta_3,$$

where the final inequalities are consequences of $U_{j+1} - U_j \leq a_i$ and $U_j \geq U_{J_1}$, both of which, in turn, follow from (7.6). This completes the proof. $\square$

### 7.3 Strengthening Lemma 7.3 to Proposition 7.1

We now provide the extra ingredients needed to extend the argument of the previous subsection to establish the stronger statement of Proposition 7.1. To avoid repetition, instead of providing the full formal proof, we shall describe the main ideas and present an elaborate sketch. Observe that to establish Proposition 7.1, one needs to extend Lemma 7.3 in the following two directions:

(a) At the same scale $j$, get the stability simultaneously in all directions in $\mathbb{S}^1(\eta)$ . We will do this by first ensuring that there is a single scale $j$ such that for all $\theta \in \mathbb{S}^1(\eta)$ there exists a 'dense' set of stable points, which we don't quantify yet.

(b) Deduce stability of most lattice points from the stability of a nearby point in the above-mentioned dense set.

We describe below how to take care of these two items. To address the issue in (a), note that one cannot naively apply the above argument separately for all $\theta \in \mathbb{S}^1(\eta)$ since a priori one might not end up with the same scale $j$ for all $\theta \in \mathbb{S}^1(\eta)$.

Instead we do the following: for each $\theta \in \mathbb{S}^1(\eta)$, consider the set of parallel lines

$$\mathfrak{L}_\theta := \{\mathbb{L}^\theta_{-K}, \dots, \mathbb{L}^\theta_{-1}, \mathbb{L}^\theta_0, \mathbb{L}^\theta_1, \dots, \mathbb{L}^\theta_K\}$$

where for any $i \in [\![-K, K]\!]$, $\mathbb{L}^\theta_i$ is a line segment of length $4n$, making angle $\theta$ with the $x$-axis; $\mathbb{L}^\theta_0$ is centered at the origin; and $\mathbb{L}^\theta_i$ is obtained by translating $\mathbb{L}^\theta_0$ in the orthogonal direction by $\frac{in}{2^{J_3\mathfrak{m}}}$ where $J_3 = J_1 + 1/\delta_3^4$ and $K = 3 \times 2^{J_3\mathfrak{m}}$ (see Figure 7.2). For each

$\theta \in \mathbb{S}^1(\eta)$ and each $i \in [\![-K, K]\!]$, let $\mathscr{U}_{i,\theta}$ be the discrete line segment formed by the points on $\mathbb{L}^\theta_i$ at spacing $\frac{n}{2^{J_1\mathfrak{m}}}$ (without loss of generality we assume that the starting and ending points of $\mathbb{L}^\theta_i$ and $\mathscr{U}_{i,\theta}$ are the same to avoid rounding ). Thus $\mathscr{U}_{i,\theta} = [\mathbf{z}^{i,\theta}_0, \mathbf{z}^{i,\theta}_1, \dots, \mathbf{z}^{i,\theta}_M]$ where $M = 4 \times 2^{J_1\mathfrak{m}}$. We now create a tree $T_{i,\theta,\ell}$ for each $i \in [\![-K, K]\!]$, $\theta \in \mathbb{S}^1(\eta)$, and $\ell \in [\![0, M-1]\!]$ corresponding to the interval

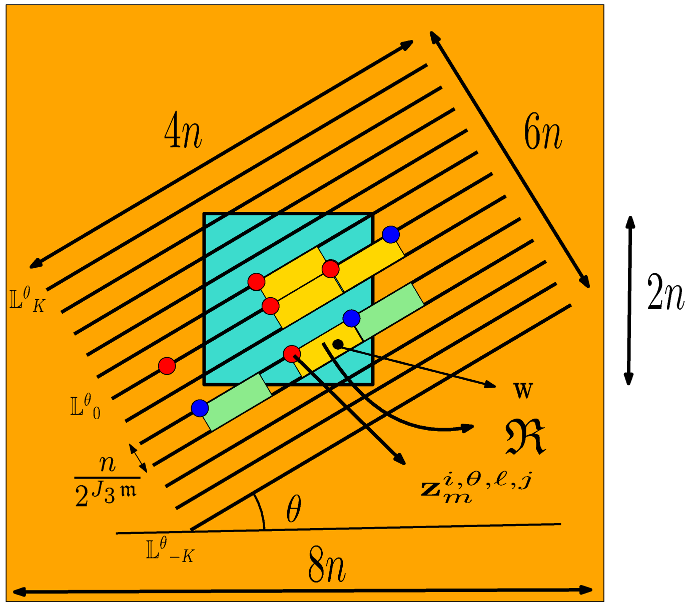


FIGURE 7.2. This figure illustrates the set of parallel lines $\mathfrak{L}_\theta := \{\mathbb{L}^\theta_{-K}, \ldots, \mathbb{L}^\theta_{-1}, \mathbb{L}^\theta_0, \mathbb{L}^\theta_1, \ldots, \mathbb{L}^\theta_K\}$. The red and blue dots denote the points $\mathbf{z}_h^{i,\theta,\ell,j}$ for $i \in [\![-K, K]\!]$, $\theta \in \mathbb{S}^1(\eta)$, $\ell \in [\![0, M-1]\!]$, and $h \in [\![1, 2^{(j-J_1)\mathrm{m}}]\!]$. The red and blue colors denote whether the point is $(\delta_2, \theta, \frac{n}{2^{j\mathrm{m}}}, k_1)$-**Stable** or not, respectively. For each such point we associate a rectangular box with one of the sides parallel to $\mathbb{L}^\theta_0$, where the point is at the northwest corner of the associated rectangle. A particular example of a point $\mathbf{z}_h^{i,\theta,\ell,j}$ and the associated rectangle $\mathfrak{R}$ and a lattice point $\mathbf{w}$ inside $\mathfrak{R}$ are marked in the figure. The green boxes are associated to the blue points and the yellow boxes are associated to the red points.

$[\mathbf{z}_\ell^{i,\theta}, \mathbf{z}_{\ell+1}^{i,\theta}]$ as in (7.3). As before, for any $j \geq J_1$, let $L_j^{i,\theta,\ell}$ denote the $j^{\text{th}}$ level of the tree $T_{i,\theta,\ell}$ and $L_j = \bigcup_{i,\theta,\ell} L_j^{i,\theta,\ell}$.

Using the same argument as before with these trees in place of the ones in (7.3) now gives us $J_1 \leq j \leq J_1 + \frac{1}{\delta_3^2}$ with the following property: If

$$\mathscr{U}_{i,\theta,\ell,j} = [\mathbf{z}_{\mathbf{0}}^{i,\theta,\ell,j}, \mathbf{z}_{\mathbf{1}}^{i,\theta,\ell,j}, \ldots, \mathbf{z}_{2^{(j-J_1)\mathrm{m}}}^{i,\theta,\ell,j}]$$

denotes the discrete segment corresponding to $L_j^{i,\theta,\ell}$, i.e., the $2^{(j-J_1)\mathrm{m}}$ vertices in the latter correspond to the intervals $[\mathbf{z}_h^{i,\theta,\ell,j}, \mathbf{z}_{h+1}^{i,\theta,\ell,j}]$ for $i \in [\![-K, K]\!]$, $\theta \in \mathbb{S}^1(\eta)$, $\ell \in [\![0, M-1]\!]$, and $h \in [\![1, 2^{(j-J_1)\mathrm{m}} - 1]\!]$, then for any $k_1$ (to be specified below and small enough compared to $J_1$), for any $\theta \in \mathbb{S}^1(\eta)$, except for at most an $O(\frac{k_1}{2^{\mathfrak{m}}}+\frac{2^{\mathfrak{m}}\sqrt{\delta_3}}{\eta})$ fraction, all the remaining $\mathbf{z}_h^{i,\theta,\ell,j}$, are $(\delta',\theta,\frac{n}{2^{j\mathfrak{m}}},k_1)$-**Stable** where $\delta' = O(2^{\mathfrak{m}}\delta_3^{1/4})$. Note that the $\frac{1}{\eta}$ term appears in the fraction of unstable points to ensure uniformity across $\theta \in \mathbb{S}^1(\eta)$.

Thus by choosing $\mathfrak{m}$ large enough depending on $k_1$ and $\delta_2$, followed by choosing $\delta_3$ small enough, provides for any $\theta \in \mathbb{S}^1(\eta)$ a dense set of points at spacing $\frac{n}{2^{j\mathfrak{m}}}$ (of fraction $1-O(\frac{k_1}{2^{\mathfrak{m}}}+\frac{2^{\mathfrak{m}}\sqrt{\delta_3}}{\eta})$), which are $(\delta_2,\theta,\frac{n}{2^{j\mathfrak{m}}},k_1)$-**Stable** and hence addresses the issue in (a).

To address the issue in (b) we will use the above along with Lemma 2.5 to imply stability for most points in $\mathbb{Z}^2 \cap \mathbf{Box}(n)$ with slightly worse parameters. Fixing $\theta \in \mathbb{S}^1(\eta)$, for any $(\delta_2,\theta,\frac{n}{2^{j\mathfrak{m}}},k_1)$-**Stable** point $\mathbf{z}_h^{i,\theta,\ell,j}$, consider any lattice point $\mathbf{w}$ in the associated rectangular box $\mathfrak{R}$ as illustrated in Figure 7.2. Thus $\|\mathbf{w}-\mathbf{z}_h^{i,\theta,\ell,j}\| \le 2\frac{n}{2^{j\mathfrak{m}}}$. Hence, applying Lemma 2.5 (by taking $\ell = \frac{n}{2^{j\mathfrak{m}}}$, $m=2$, $k=k_1$ and $C=\sqrt{k_1}$) implies that $\mathbf{w}$ is $(\delta',\theta,\frac{n\sqrt{k_1}}{2^{j\mathfrak{m}}},\sqrt{k_1})$-**Stable** where $\delta' = \delta_2 + O(\frac{1}{\sqrt{k_1}})$. Thus it follows that

$$\begin{aligned}&\#\{\mathbf{z}\in\mathbb{Z}^2\cap\mathbf{Box}(n) : \mathbf{z} \text{ is not } (\delta',\theta,\frac{n\sqrt{k_1}}{2^{j\mathfrak{m}}},\sqrt{k_1})\text{-}\mathbf{Stable}\}\\ &\quad\le O(\frac{k_1}{2^{\mathfrak{m}}}+\frac{2^{\mathfrak{m}}\sqrt{\delta_3}}{\eta})n^2,\end{aligned}$$

By a simple union bound over $\theta \in \mathbb{S}^1(\eta)$, it follows that

$$\begin{aligned}&\#\left\{\mathbf{z}\in\mathbb{Z}^2\cap\mathbf{Box}(n) : \mathbf{z} \text{ is not } \left(\delta',\mathbb{S}^1(\eta),\frac{n\sqrt{k_1}}{2^{j\mathfrak{m}}},\sqrt{k_1}\right)\text{-}\mathbf{Stable}\right\}\\ &\quad\le O\left(\frac{1}{\eta}\left(\frac{k_1}{2^{\mathfrak{m}}}+\frac{2^{\mathfrak{m}}\sqrt{\delta_3}}{\eta}\right)\right)n^2.\end{aligned} \tag{7.11}$$

The statement of Proposition 7.1 now follows from choosing $\sqrt{k_1} \gtrsim \max(\frac{1}{\delta},k)$, followed by $\delta_2$ small enough to ensure $\delta' \le \delta$, and then $\mathfrak{m}$ large enough followed by $\delta_3$ small enough to ensure that the

$$O\left(\frac{1}{\eta}\left(\frac{k_1}{2^{\mathfrak{m}}}+\frac{2^{\mathfrak{m}}\sqrt{\delta_3}}{\eta}\right)\right) \text{ term}$$

is less than $\varepsilon$. Moreover, we take the value of $J_2$ to be $\mathfrak{m}(J_1+1/\delta_3^2)$. Note that the value of $j$ in Proposition 7.1 can be taken to be $j\mathfrak{m}-\frac{\log k_1}{2}$, where the latter $j$ appears in (7.11).

**Acknowledgment.** The authors thank an anonymous referee for a careful reading of the manuscript and many useful comments and suggestions that helped improve the quality of the paper. RB is partially supported by an ICTS-Simons Junior Faculty Fellowship, a Ramanujan Fellowship (SB/S2/RJN-097/2017) from the Science and Engineering Research Board, and by ICTS via project no. 12-R&D-TFR-5.10-1100 from DAE, Government of India. SG was partially supported by a Miller Research Fellowship. AS is supported by NSF Grant DMS-1855527, a Simons Investigator grant, and a MacArthur Fellowship.

## Bibliography

[1] Auffinger, A.; Damron, M. A simplified proof of the relation between scaling exponents in first-passage percolation. *Ann. Probab.* **42** (2014), no. 3, 1197–1211. doi:10.1214/13-AOP854

[2] Auffinger, A.; Damron, M.; Hanson, J. 50 years of first passage percolation. Preprint, 2015. arXiv:1511.03262 [math.PR]

[3] Baik, J.; Deift, P.; Johansson, K. On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc* **12** (1999), no. 4, 1119–1178. doi:10.1090/S0894-0347-99-00307-0

[4] Basu, R.; Bhatia, M.; Ganguly, S. Lower tail large deviations in Poissonian polymers. In preparation.

[5] Basu, R.; Ganguly, S. Connecting eigenvalue rigidity with polymer geometry: Diffusive transversal fluctuations under large deviation. Preprint, 2019. arXiv:1902.09510 [math.PR]

[6] Basu, R.; Ganguly, S.; Sly, A. Delocalization of polymers in lower tail large deviation. *Comm. Math. Phys.* **370** (2019), no. 3, 781–806. doi:10.1007/s00220-019-03526-0

[7] Ben-Ari, I. Large deviations for partition functions of directed polymers in an IID field. *Ann. Inst. H. Poincaré Probab. Statist.* **45** (2009), no. 3, 770–792. doi:10.1214/08-AIHP185

[8] Benjamini, I.; Kalai, G.; Schramm, O. First passage percolation has sublinear distance variance. *Ann. Probab.* **31** (2003), no. 4, 1970–1978. doi:10.1214/aop/1068646373

[9] Chatterjee, S. The universal relation between scaling exponents in first-passage percolation. *Ann. Math. (2)* **177** (2013), no. 2, 663–697. doi:10.4007/annals.2013.177.2.7

[10] Chow, Y.; Zhang, Y. Large deviations in first-passage percolation. *Ann. Appl. Probab.* **13** (2003), no. 4, 1601–1614. doi:10.1214/aoap/1069786513

[11] Corwin, I.; Ghosal, P. Lower tail of the KPZ equation. *Duke Math. J.* **169** (2020), no. 7, 1329–1395. doi:10.1215/00127094-2019-0079

[12] Cox, J. T.; Durrett, R. Some limit theorems for percolation processes with necessary and sufficient conditions. *Ann. Probab.* **9** (1981), no. 4, 583–603.

[13] Cranston, M.; Gauthier, D.; Mountford, T. On large deviation regimes for random media models. *Ann. Appl. Probab.* **19** (2009), no. 2, 826–862. doi:10.1214/08-AAP535

[14] Cranston, M.; Gauthier, D.; Mountford, T. On large deviations for the parabolic Anderson model. *Probab. Theory Related Fields* **147** (2010), no. 1-2, 349–378. doi:10.1007/s00440-009-0249-z

[15] Damron, M.; Hanson, J. Busemann functions and infinite geodesics in two-dimensional first-passage percolation. *Comm. Math. Phys.* **325** (2014), no. 3, 917–963. doi:10.1007/s00220-013-1875-y

[16] Damron, M.; Hanson, J. Bigeodesics in first-passage percolation. *Comm. Math. Phys.* **349** (2017), no. 2, 753–776. doi:10.1007/s00220-016-2743-3

[17] Deuschel, J.-D.; Zeitouni, O. On increasing subsequences of iid samples. *Combin. Probab. Comput.* **8** (1999), no. 3, 247–263. doi:10.1017/S0963548399003776

[18] Georgiou, N.; Seppäläinen, T. Large deviation rate functions for the partition function in a log-gamma distributed random potential. *Ann. Probab.* **41** (2013), no. 6, 4248–4286. doi:10.1214/12-AOP768

[19] Hammersley, J. M.; Welsh, D. J. A. First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. *Proc. Internat. Res. Semin., Statist. Lab., Univ. California, Berkeley, Calif.*, 61-–110. Springer, New York, 1965.

[20] Hoffman, C. Geodesics in first passage percolation. *Ann. Appl. Probab.* **18** (2008), no. 5, 1944–1969. doi:10.1214/07-AAP510

[21] Jensen, L. The asymmetric exclusion process in one dimension. Ph.D. thesis, New York University, 2000.

[22] Johansson, K. Shape fluctuations and random matrices. *Comm. Math. Phys.* **209** (2000), no. 2, 437–476. doi:10.1007/s002200050027

[23] Kardar, M.; Parisi, G.; Zhang, Y.-C. Dynamic scaling of growing interfaces. *Phys. Rev. Lett.* **56** (1986), 889–892. doi:10.1103/PhysRevLett.56.889

[24] Kesten, H. Aspects of first passage percolation. *École d'Été de Probabilités de Saint Flour XIV - 1984*, 125–264. Lecture Notes in Math., 1180. Springer, Berlin, 1986. doi:10.1007/BFb0074919

[25] Kesten, H. Percolation theory and first-passage percolation. *Ann. Probab.* **15** (1987), no. 4, 1231–1271.

[26] Kesten, H. On the speed of convergence in first-passage percolation. *Ann. Appl. Probab.* **3** (1993), no. 2, 296–338.

[27] Kingman, J. F. C. Subadditive ergodic theory. *Ann. Probability* **1** (1973), 883–909. doi:10.1214/aop/1176996798

[28] Lindvall, T. On Strassen's theorem on stochastic domination. *Electron. Comm. Probab.* **4** (1999), 51–59. doi:10.1214/ECP.v4-1005

[29] Nakajima, S. Upper tail large deviations for a class of distributions in first-passage percolation. Preprint, 2019. arXiv:1912.13212 [math.PR]

[30] Newman, C. M. A surface view of first-passage percolation. *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)*, 1017-–1023. Birkhäuser, Basel, 1995.

[31] Olla, S.; Tsai, L.-C. Exceedingly large deviations of the totally asymmetric exclusion process. *Electron. J. Probab.* **24** (2019), Paper No. 16, 71 pp. doi:10.1214/19-EJP278

[32] Richardson, D. Random growth in a tessellation. *Proc. Cambridge Philos. Soc.* **74** (1973), 515–528. doi:10.1017/s0305004100077288

[33] Rost, H. Nonequilibrium behaviour of a many particle process: Density profile and local equilibria. *Z. Wahrsch. Verw. Gebiete* **58** (1981), no. 1, 41–53. doi:10.1007/BF00536194

[34] Seppäläinen, T. Coupling the totally asymmetric simple exclusion process with a moving interface. *Markov Processes and Related Fields* **4** (1998), no. 4, 593–628.

[35] Seppäläinen, T. Large deviations for increasing sequences on the plane. *Probab. Theory Related Fields* **112** (1998), no. 2, 221–244. doi:10.1007/s004400050188

[36] Talagrand, M. Concentration of measure and isoperimetric inequalities in product spaces. *Inst. Hautes Études Sci. Publ. Math.* **81** (1995), no. 1, 73–205.

[37] Tsai, L.-C. Exact lower tail large deviations of the KPZ equation. Preprint, 2018. arXiv:1809.03410 [math.PR]

[38] Varadhan, S. R. S. Large deviations for the asymmetric simple exclusion process. *Stochastic analysis on large scale interacting systems*, 1–27. Adv. Stud. Pure Math., 39. Math. Soc. Japan, Tokyo, 2004. doi:10.2969/aspm/03910001

RIDDHIPRATIM BASU
International Centre
for Theoretical Sciences
Tata Institute of Fundamental Research
Bangalore 560089
INDIA
E-mail: rbasu@icts.res.in

SHIRSHENDU GANGULY
Department of Statistics
UC Berkeley
Berkeley, CA 94720
USA
E-mail: sganguly@berkeley.edu

ALLAN SLY
Department of Mathematics
Princeton University
Princeton, NJ 08540
USA
E-mail: allansly@princeton.edu

Received November 2018.
Revised December 2020.