

# Uniform fluctuation and wandering bounds in first passage percolation

Kenneth S. Alexander\*

## Abstract

We consider first passage percolation on certain isotropic random graphs in  $\mathbb{R}^d$ . We assume exponential concentration of passage times  $T(x, y)$ , on some scale  $\sigma(r)$  whenever  $|y - x|$  is of order  $r$ , with  $\sigma(r)$  “growing like  $r^\chi$ ” for some  $0 < \chi < 1$ . Heuristically this means transverse wandering of geodesics should be at most of order  $\Delta(r) = (r\sigma(r))^{1/2}$ . We show that in fact uniform versions of exponential concentration and wandering bounds hold: except with probability exponentially small in  $t$ , there are no  $x, y$  in a natural cylinder of length  $r$  and radius  $K\Delta(r)$  for which either (i)  $|T(x, y) - ET(x, y)| \geq t\sigma(r)$ , or (ii) the geodesic from  $x$  to  $y$  wanders more than distance  $\sqrt{t}\Delta(r)$  from the cylinder axis. We also establish that for the time constant  $\mu = \lim_n ET(0, ne_1)/n$ , the “nonrandom error”  $ET(0, x) - \mu|x|$  is at most a constant multiple of  $\sigma(|x|)$ .

**Keywords:** first passage percolation; geodesic; exponential concentration; multiscale; coarse graining.

**MSC2020 subject classifications:** Primary 60K35, Secondary 82B43.

Submitted to EJP on November 22, 2021, final version accepted on October 6, 2023.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Context of our results . . . . .	5
1.2	Definitions and assumptions . . . . .	6
1.3	Statement of results . . . . .	9
<b>2</b>	<b>Existence of acceptable random graphs</b>	<b>11</b>
<b>3</b>	<b>Straightness of geodesics and and regularity of means</b>	<b>14</b>

---

\*University of Southern California, United States of America. E-mail: alexandr@usc.edu

<b>4 Proof of Theorem 1.3—downward deviations</b>	<b>23</b>
4.1 Proof sketch for (1.14)	23
4.2 Step 0. Preliminaries and easy cases	25
4.3 Step 1. Setting up the coarse-graining	29
4.4 Step 2. Performing the $j_1$ th-scale (first) iteration of coarse-graining	38
4.5 Step 3. Bounding the tracking-error event for the $j_1$ th-scale-iteration	42
4.6 Step 4. Further iterations of coarse-graining: preparation	44
4.7 Step 5. First stage of the $(j_1 - 1)$ th-scale (second) iteration of coarse-graining: shifting to the $(j_1 - 1)$ th-scale grid	52
4.8 Step 6. Second stage of the $(j_1 - 1)$ th-scale (second) coarse-graining iteration: removing marked points	54
4.9 Step 7. Further iterations of coarse-graining	68
4.10 Step 8. Final marked CG paths	73
<b>5 Proof of Theorem 1.6 for fixed <math>(x, y)</math></b>	<b>75</b>
<b>6 Proof of Theorem 1.5</b>	<b>78</b>
<b>7 Proof of Theorem 1.3—upward deviations</b>	<b>82</b>
<b>8 Proofs of Theorem 1.6 and Lemma 1.2</b>	<b>84</b>
<b>References</b>	<b>85</b>

## 1 Introduction

In i.i.d. first passage percolation (FPP) on a graph  $G = (V, E)$ , i.i.d. (edge) passage times  $t_e$  are attached to the edges  $e \in E$ , and for a path  $\Gamma$  in  $G$ , the (path) passage time  $T(\Gamma)$  is the sum of the times  $t_e$  over  $e \in \Gamma$ . For  $x, y \in V$ , the passage time from  $x$  to  $y$  is

$$T(x, y) = \inf\{T(\Gamma) : \Gamma \text{ is a path from } x \text{ to } y \text{ in } G\}. \tag{1.1}$$

The geodesic from  $x$  to  $y$  is the path, denoted  $\Gamma_{xy}$ , which minimizes the path passage time; when  $t_e$  is a continuous random variable (as we always assume), a unique geodesic exists a.s. [22].

There are two exponents of primary interest in the study of FPP. First, the fluctuations (i.e. standard deviation) of passage times  $T(x, y)$  for  $|y - x|$  of scale  $r$  in  $\mathbb{Z}^d$  are believed to be of order  $r^\chi$  for some  $\chi = \chi_d < 1/2$ , with  $\chi_2 = 1/3$ . Second, the typical transverse wandering of a geodesic, meaning the maximum distance from any point on  $\Gamma_{xy}$  to the straight line (denoted  $\Pi_{xy}$ ) from  $x$  to  $y$ , is believed to be of order  $r^\xi$  for some  $\xi = \xi_d$ . For  $|y - x|$  of order  $r$ , if  $\Gamma_{xy}$  contains a vertex  $z$  at distance of order  $r^\xi$  from  $\Pi_{xy}$  (not too near  $x$  or  $y$ ), then the associated extra distance  $|z - x| + |y - z| - |y - x|$  traveled by the geodesic in order to pass through  $z$  is of order  $r^{2\xi - 1}$ . For such wandering to have non-negligible probability, the passage time fluctuations  $r^\chi$  should be at least as large as the extra distance; heuristically this leads to the relation  $\chi = 2\xi - 1$ . There are various ways to formally define the exponents  $\chi, \xi$ ; these must allow for the fact that the true scales of fluctuations and wandering are not known to be pure powers of  $r$ . Chatterjee [14] gave a rigorous version of the relationship  $\chi = 2\xi - 1$ , under the assumption that multiple possible definitions of each exponent actually agree.

Looking more finely than just at the level of exponents, the heuristic for  $\chi = 2\xi - 1$  says that if the fluctuation scale is  $\sigma(r)$  for  $|y - x|$  of order  $r$ , then the scale of transverse wandering should be

$$\Delta(r) = (r\sigma(r))^{1/2}.$$

In [2] and (for  $d = 2$ ) in [19] it was shown that under natural assumptions, the transverse wandering with high probability does not exceed  $(r\sigma(r) \log r)^{1/2}$ . One of our main results here is an upper bound on wandering: under somewhat weaker assumptions, for all  $d \geq 2$ , the probability of wandering greater than  $s\Delta(r)$  decays as  $e^{-cs^2}$  or faster, for  $s \leq r/\Delta(r)$ . Previously such a result has only been known for integrable cases of last passage percolation (LPP) in  $d = 2$ , from [13] (with  $e^{-cs}$  in place of  $e^{-cs^2}$ ) and [12]. The bound is optimal in the sense of being on the scale  $\Delta(r)$ , though the second power of  $s$  in the exponent may not be optimal, as suggested by LPP results in  $d = 2$  in [20].

In fact we have this bound uniformly over many geodesics simultaneously, in the following sense: Consider a cylinder of length  $r$  and radius  $K\Delta(r)$ , and let  $\epsilon > 0$  and  $s > 2K$ . Then under the assumptions we will make, the probability that there exists any geodesic  $\Gamma_{xy}$ , with  $x, y$  in the cylinder and  $|y - x| \geq \epsilon r$ , which wanders farther than  $s\Delta(r)$  from the cylinder axis decays as  $e^{-cs^2}$  or faster, for  $s \leq r/\Delta(r)$ .

By comparison, in [2] it was shown roughly that if there is exponential concentration on some scale  $\sigma(r)$  which “grows like a power of  $r$ ,” uniformly for passage times over distance  $r$ , then the probability of a transverse fluctuation of size  $t(\log r)^{1/2}\Delta(r)$  for a single geodesic  $\Gamma_{xy}$  is bounded by  $C_1 e^{-C_2 t^2 \log t}$  for all  $t > 0$ . This tells us nothing, though, about transverse fluctuations of size  $t\Delta(r)$  with  $1 \ll t \ll (\log r)^{1/2}$ , which should also be subject to exponential concentration, as in our present result.

It should be emphasized that in our transverse wandering (and other) results,  $\sigma(r)$  is not necessarily the actual scale of the standard deviation—it need only be an upper bound in the sense that exponential concentration holds for passage times  $T(x, y)$  on scale  $\sigma(|y - x|)$ . Then the corresponding value  $\Delta(r)$  is the scale that appears in the upper bound for transverse wandering.

Our uniform wandering bound will be a byproduct of another uniform-bound result for passage times; to describe it we first discuss exponential bounds. For the lattice  $\mathbb{Z}^d$ , Kesten [23] proved that, assuming

$$Ee^{\lambda t_e} < \infty \text{ for some } \lambda > 0, \quad \text{and} \quad P(t_e = 0) < p_c(\mathbb{Z}^d) \tag{1.2}$$

(where  $p_c(\mathbb{Z}^d)$  is the bond percolation threshold for  $\mathbb{Z}^d$ ), there is exponential concentration of  $T(x, y)$  on scale  $r^{1/2}$  for  $|x - y| \leq r$ :

$$P(|T(x, y) - ET(x, y)| \geq tr^{1/2}) \leq C_3 e^{-C_4 t} \quad \text{for all } t \leq C_5 r.$$

Talagrand [33] improved this: assuming just an exponential moment for  $t_e$ ,

$$P(|T(x, y) - ET(x, y)| \geq tr^{1/2}) \leq C_6 e^{-C_7 \min(t^2, tr^{1/2})} \quad \text{for all } t > 0.$$

Damron, Hanson, and Sosoe [16] improved the bound to a subgaussian scale: under (1.2),

$$P\left(|T(x, y) - ET(x, y)| \geq t \left(\frac{r}{\log r}\right)^{1/2}\right) \leq C_8 e^{-C_9 t} \quad \text{for all } t > 0. \tag{1.3}$$

None of these bounds are near-optimal, though—an optimal bound would be on the scale of the standard deviation of  $T(x, y)$ . What we will prove here is roughly as follows. Suppose passage times satisfy exponential concentration on a scale  $\sigma(\cdot)$ , uniformly:

$$P\left(|T(x, y) - ET(x, y)| \geq t\sigma(|y - x|)\right) \leq C_{10} e^{-C_{11} t} \quad \text{for all } x, y, \tag{1.4}$$

for some  $\sigma(r)$  which “grows like  $r^\chi$ ” for some  $\chi \in (0, 1)$ , in a sense we will make precise. Then for  $G_r(K)$  a cylinder of length  $r$  and radius  $K\Delta(r)$  for some fixed  $K$ , we have concentration on the same scale, uniformly over  $x, y \in G_r(K)$ :

$$P\left(|T(x, y) - ET(x, y)| \geq t\sigma(r) \text{ for some } x, y \in G_r(K) \text{ with } |y - x| \geq \epsilon r\right) \leq C_{12} e^{-C_{13} t} \tag{1.5}$$

for all  $r$  large and  $t \geq cK^2$ . This has previously been proved for integrable models of LPP, but only for  $d = 2$  ([13, 11]), and even the non-integrable part of the proof there does not carry over to FPP—see Remark 1.8.

For  $d \geq 3$  there is no generally-agreed-upon value of  $\chi$  in the physics literature. Heuristics and simulations suggest that  $\chi$  should decrease with dimension; simulations in [32] for a model believed to be in the same (KPZ) universality class as FPP show a decrease from  $\chi = .33$  to  $\chi = .054$  as  $d$  increases from 2 to 7. Some have predicted the existence of a finite upper critical dimension, possibly as low as 3.5, above which  $\chi = 0$  ([18], [25]); others predict that  $\chi$  is positive for all  $d$  ([4], [29]), with simulations in [24] showing  $\chi > 0$  all the way to  $d = 12$ , decaying approximately as  $1/(d + 1)$ . Our results here require  $\chi > 0$  so they only have content below the upper critical dimension, should it be finite. In fact, results in [3] strongly suggest that Theorems 1.3 and 1.5 below should not be valid above the upper critical dimension, should such exist.

In the preceding and throughout the paper,  $c_1, c_2, \dots$  and  $C_1, C_2, \dots$ , and  $\epsilon_0, \epsilon_1, \dots$  represent unspecified constants which depend only on the graph  $\mathbb{G}$  (or its distribution, if it is random) and the distribution of the passage times  $t_e$  (or speeds  $\eta_e$ , to be given below.) We use  $C_i$  for constants which occur outside of proofs and may be referenced later; any given  $C_i$  has the same value at all occurrences. We use  $c_i$  for those which do not recur and are only needed inside one proof. For the  $c_i$ 's we restart the numbering with  $c_0$  in each proof, and the values are different in different proofs.

As is standard, since passage times  $T(x, y)$  are subadditive, assumptions much weaker than (1.2) guarantee the a.s. existence (positive and finite for  $x \neq 0$ ) of the limit

$$g(x) = \lim_n \frac{T(0, nx)}{n} = \lim_n \frac{ET(0, nx)}{n} = \inf_n \frac{ET(0, nx)}{n} \quad \text{a.s. and in } L^1 \quad (1.6)$$

for  $x \in \mathbb{Z}^d$ ;  $g$  extends to  $x$  with rational coordinates by considering only  $n$  with  $nx \in \mathbb{Z}^d$ , and then to a norm on  $\mathbb{R}^d$  by uniform continuity. We write  $\mathfrak{B}_g$  for the unit ball of this norm. To obtain the optimal uniform results for wandering and (1.5) for fluctuations, we need to understand both parts of the discrepancy

$$T(0, x) - g(x) = \left( T(0, x) - ET(0, x) \right) + \left( ET(0, x) - g(x) \right). \quad (1.7)$$

Here in the parentheses on the right are the *random part* and *nonrandom part* of the discrepancy. In [1] it was shown that under (1.2),

$$g(x) \leq ET(0, x) \leq g(x) + C_{14}|x|^{1/2} \log |x| \quad \text{for all } |x| > 1. \quad (1.8)$$

In [34] the error term was improved to  $C_{14}(|x| \log |x|)^{1/2}$ , and in [19] to  $c_\eta |x|^{1/2} (\log |x|)^\eta$  for all  $\eta > 0$ . For the Euclidean first passage percolation of [21], the analog of (1.8) was proved in [17] with an error term of  $C_{14} \Psi(|x|) \log^{(k)} |x|$  for arbitrary  $k \geq 1$ , where  $\Psi(|x|)$  is a scale on which an exponential bound is known (analogous to  $\sigma(|x|)$  in (1.4)) and  $\log^{(k)} |x|$  is the  $k$ -times-iterated logarithm;  $\Psi$  is subject to regularity conditions which are incompatible with  $\chi = 0$ . Here we will obtain an essentially optimal bound for the non-random part: if  $\sigma(\cdot)$  satisfies certain regularity (also incompatible with  $\chi = 0$ ) and (1.4) holds, then a log factor as in the earlier bounds is unnecessary in our context: we have

$$g(x) \leq ET(0, x) \leq g(x) + C_{15} \sigma(|x|). \quad (1.9)$$

There is a strong interdependence among this result, our uniform wandering bounds, and (1.5), as discussed in Remark 1.7.

Analogous to (1.9), of (1.5), and of our uniform wandering result, with optimal scale  $\sigma(|x|) = \text{var}(T(0, x))^{1/2} \asymp |x|^{1/3}$ , are known for certain integrable models of directed last passage percolation (LPP) for  $d = 2$ . We note that an exponential bound like (1.4),

but with centering at the analog of  $g(x)$  instead of at  $ET(0, x)$ , shows that (i) (1.9) must hold, and (ii) an exponential bound like (1.4) must also hold with centering at  $g(y - x)$ . Such a recentered bound appears in [13] (extracted from [5]) and in [26] for LPP on  $\mathbb{Z}^2$  with exponential passage times, in [27, 28] for LPP based on a Poisson process in the unit square, and in [15] for LPP on  $\mathbb{Z}^2$  with geometric passage times. An analog of our transverse wandering bound for LPP on  $\mathbb{Z}^2$  with exponential passage times appears in [12]. All of these require integral probability methods, which are not available for FPP.

### 1.1 Context of our results

For  $d = 2$  it is useful to discuss further the relation of our work to recent LPP results. For integrable models of LPP on  $\mathbb{Z}^2$ , from a probabilistic perspective, many important results are “black boxes” in the sense that the proofs are based on ideas from algebraic combinatorics and asymptotic analysis which frequently have no probabilistic interpretation. A number of authors have recently employed the approach of taking certain of these black-box results as inputs to proofs, which are otherwise purely probabilistic, of a rich range of results. Some primary black-box inputs used this way (stated loosely) are the following, with  $u$  denoting a unit vector:

- (i) The standard deviation  $\sigma(T(0, ru))$  is asymptotically  $\approx Cr^{1/3}$ .
- (ii) Exponential bound for passage times over distance  $r$ :

$$P\left(|T(0, ru) - ET(0, ru)| \geq \lambda r^{1/3}\right) \leq c_1 e^{-c_2 \lambda} \quad \text{for (nearly) all } r, \lambda, u.$$

The rate  $r^{1/3}$  is optimal, due to (i).

- (iii) The asymptotic growth rate  $g(x)$  from (1.6) satisfies

$$g(ru) - ET(0, ru) \approx Cr^{1/3}.$$

- (iv)  $g(ru)$  varies nicely both transversally as a function of  $u$  (e.g. the boundary of “ $g \leq \text{constant}$ ” has nontrivial curvature) and longitudinally as a function of  $r$ ; therefore by (iii),  $ET(0, ru)$  also varies nicely.
- (v) The normalization  $(T(0, ru) - ET(0, ru))/Cr^{1/3}$  converges to a nondegenerate distribution (GUE Tracy–Widom.)
- (vi) LPP paths are by definition directed (simplifying a number of proofs.)

The use of these inputs in this manner occurs for example in [7, 8, 9, 10, 11, 12, 20, 30, 36]. In particular, it has become clear in recent years that coalescence properties of geodesics are central to the understanding of LPP and FPP, particularly in the context of limiting objects like the directed landscape; several of these papers answer such questions for LPP, using some of the inputs (i)–(vi). For FPP, for this and other questions, it would certainly be desirable to derive a correspondingly rich set of consequences from (i)–(vi), but no analog of any of these inputs is known. The closest would be (1.3) from [16], which is like (ii) but with a far-from-optimal  $(r/\log r)^{1/2}$  in place of  $r^{1/3}$ . Therefore ever since the ‘90’s, some authors have taken the path of *assuming* variants of some of (i)–(vi), and seeing what follows from them. For example, some (e.g. [2, 31], and for non-integrable LPP [20]) have assumed a form of (iv): the boundary of  $\{x : g(x) \leq 1\}$  has nontrivial curvature, either globally or in some direction. In [14] the full strength of the main result requires the assumption that, in a certain sense, the powers of  $r$  appearing in (i) and (ii) are “the same to within an arbitrarily small  $\epsilon$ .”

It should be emphasized that results established for LPP will not transfer readily to FPP even if one were to assume or prove versions of *all* of (i)–(vi), as they are not the only black-box inputs, just perhaps the primary ones. For example, [7, 11, 36] all

explicitly exploit the fact that the limiting GUE Tracy–Widom distribution in (v) has a negative mean. New approaches will be necessary, beyond what works for FPP, even if (i)–(v) are proved or assumed. This paper is itself an example of the need for new approaches—see Remark 1.8.

To access a richer understanding of FPP, it certainly seems undesirable to assume *all* of the unproven inputs (i)–(vi), so what to do? Can the necessary assumptions at least be minimized? Input (iv) can be obtained by working on a graph which is translation-invariant and isotropic, instead of  $\mathbb{Z}^2$ ; a random graph is most natural for this, and then  $g(x)$  is necessarily a constant multiple of  $|x|$ . One of our main results says that in this isotropic context, a version of (iii) (with the two terms on the left interchanged, as appropriate for FPP) is a consequence of mainly just a version of (ii): if (ii) holds (along with certain milder assumptions) with  $r^{1/3}$  replaced by a function which “grows like  $r^\chi$ ” in a nice way for some  $\chi \in (0, 1)$ , then the FPP version of (iii) also holds for that function, at least as a one-sided bound. Thus inputs (ii)–(iv) all become available with only (ii) assumed—and this is true in all dimensions where  $\chi > 0$ . One can hope that this widens the door a bit for future results about the geometry of FPP. More broadly, the situation for LPP points to the centrality of (i)–(v) as questions for FPP.

Fortunately, working on isotropic random graphs doesn’t increase technicalities relative to the usual  $\mathbb{Z}^d$  in our context. In fact it decreases them substantially, as isotropy is a precise and powerful tool. Technicalities are also reduced because we don’t try to make the random graph as general as possible, in fact quite the opposite—we just need to know that *some* random graph exists with certain properties convenient for FPP, such as a finite radius of dependence (see A1 below), so that we can use input (iv). That finite radius of dependence is not present in previously considered isotropic models of FPP, for example in [21].

## 1.2 Definitions and assumptions

As mentioned we will consider isotropic models, built on a random graph  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  embedded in  $\mathbb{R}^d$ , to ensure a strong form of (iv) in Section 1.1. We want this random graph to be “as lattice-like as possible,” as quantified in A1 below, despite being isotropic. To that end, the *dilation* of such an embedded graph is the least  $C$  such that, for every  $x, y \in \mathbb{V}$  there is a path from  $x$  to  $y$  in  $\mathbb{G}$  for which the total (Euclidean) length of the edges is at most  $C|y - x|$ . We say that such a graph has *bounded dilation* if there exists  $C$  such that with probability one the dilation of  $\mathbb{G}$  is at most  $C$ . For  $A \subset \mathbb{R}^d$ , the *restriction of  $\mathbb{G}$  to  $A$*  is the graph with vertex set  $\mathbb{V}_A = \{x \in \mathbb{V} : \langle x, y \rangle \in \mathbb{E} \text{ for some } y \in \mathbb{V} \cap A\}$  and edge set  $\mathbb{E}_A = \{\langle x, y \rangle \in \mathbb{E} : x \in A\}$ .

We require that the graph  $\mathbb{G}$  satisfy the assumptions A1 below, which are somewhat stringent and include bounded dilation, but we will construct an example that works. (We see no need to make the graph as general as possible; we simply need one to know we can work with one that has certain properties that make it lattice-like.) That example is built roughly as follows. We first construct a point process  $\mathbb{V}$  to serve as vertices, with  $\mathbb{V}$  satisfying those parts of A1 which involve only the vertices. To make a graph from  $\mathbb{V}$  we use the Voronoi diagram, which divides  $\mathbb{R}^d$  into closed polyhedrons  $\{Q_x : x \in \mathbb{V}\}$  (called *Voronoi cells*), the interior of the polygon  $Q_x$  consisting of those points which are strictly closer to  $x$  than to any other point of  $\mathbb{V}$ . We define  $\varphi$  by  $\varphi(y) = x$  for  $y \in Q_x$ . To produce the *Delaunay graph* (or *Delaunay triangulation* in  $d = 2$ ) one places an edge between each pair  $x, y \in \mathbb{V}$  for which  $Q_x$  and  $Q_y$  have a face of positive  $(d - 1)$ -volume in common. For  $d = 2$ , it is known that the dilation of the Delaunay graph of any locally finite subset of  $\mathbb{R}^2$  is at most 1.998 [35], ensuring A1 is fully satisfied, but such bounded dilation for  $d \geq 3$  is not known. We therefore modify the Delaunay graph by adding certain non-nearest-neighbor edges of uniformly bounded length, by a deterministic

local rule, and show that bounded dilation then holds. Here and in what follows, by the *length* of an edge  $e = \langle x, y \rangle$  we mean the Euclidean distance  $|y - x|$ , which we denote  $|e|$ .

Our FPP proofs for isotropic random graphs should adapt to  $\mathbb{Z}^d$ , but (we expect) only by assuming two unproven properties of FPP on  $\mathbb{Z}^d$ : first, uniform curvature of  $\partial\mathcal{B}_g$ , and second, a kind of smoothness of the mean as the direction changes:

$$\begin{aligned} & \sup \left\{ |ET(0, x) - ET(0, y)| : \max(|g(x) - r|, |g(y) - r|) \leq C_{16}, |y - x| \leq C_{17}\Delta(r) \right\} \\ & = O(\sigma(r)) \quad \text{as } r \rightarrow \infty. \end{aligned}$$

Results in [1] show that for lattice FPP, the left side is  $O(\sigma(r) \log r)$ . This is (implicitly) improved to  $O(\sigma(r)(\log r)^{1/2})$  in [34] and to  $O(\sigma(r)(\log r)^\kappa)$  for all  $\kappa > 0$ , in [19].

Here then are the assumptions  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  must satisfy.

### A1. Acceptability of random graphs

- (i)  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  is isotropic, stationary, and ergodic;
- (ii) Bounded hole size: every open ball in  $\mathbb{R}^d$  of radius 1 contains at least one vertex of  $\mathbb{V}$ ;
- (iii) Finite range of dependence: there exists  $\beta$  such that if  $A, B$  are Lebesgue-measurable subsets of  $\mathbb{R}^d$  separated by distance  $d(A, B) \geq \beta$ , then the restrictions of  $\mathbb{G}$  to  $A$  and to  $B$  are independent;
- (iv) Bounded dilation: the dilation of  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  is bounded a.s. (and hence equal to some nonrandom  $C_{18}$  a.s., by (i));
- (v) Exponential bound for the local density: given  $r_0 > 0$  there exist  $C_{19}, C_{20}$  such that for all  $r > r_0$  and  $a \geq 1$ ,  $P(|\mathbb{V} \cap B_r(0)| \geq ar^d) \leq C_{19}e^{-C_{20}a}$ ;

We say a random graph with these properties is *acceptable*. We will show that acceptable random graphs exist. By rescaling, we may replace radius 1 in (ii) by any other positive value. Condition (v) can be weakened from exponential to stretched exponential; we use exponential to simplify the exposition.

Conditionally on  $\mathbb{G}$  we define a collection of i.i.d. nonnegative continuous random variables  $\eta = \{\eta_e, e \in \mathbb{E}\}$ . Formally the pair  $\omega = (\mathbb{V}, \eta)$  is defined on a probability space  $(\Omega, \mathcal{F}, P)$ , with  $\mathbb{G}$  determined by  $\mathbb{V}$ .

In contrast to the usual FPP on a true lattice, here we view  $\eta_e$  not as a time but as a speed. We thus define the passage time of an edge  $e$  to be  $\eta_e|e|$ , and proceed “as usual”: for  $x, y \in \mathbb{V}$ , a *path*  $\Gamma$  from  $x$  to  $y$  is a finite sequence of alternating vertices and edges of  $\mathbb{G}$ , of the form  $\Gamma = (x = x_0, \langle x_0, x_1 \rangle, x_1, \dots, x_{n-1}, \langle x_{n-1}, x_n \rangle, x_n = y)$ . We may designate a path by specifying only the vertices. The (*path*) *passage time* of  $\Gamma$  is

$$T(\Gamma) := \sum_{e \in \Gamma} \eta_e |e|,$$

and the *passage time* from  $x$  to  $y$  is

$$T(x, y) := \inf \{T(\Gamma) : \Gamma \text{ is a path from } x \text{ to } y \text{ in } \mathbb{G}\}. \tag{1.10}$$

More generally, for  $x, y \in \mathbb{R}^d$  we define

$$T(x, y) := T(\varphi(x), \varphi(y)).$$

For technical convenience we do not require that paths be self-avoiding, but for the moment this is irrelevant because geodesics are always self-avoiding. For general  $x, y$  not necessarily in  $\mathbb{V}$ , we take “a geodesic from  $x$  to  $y$ ” to mean a geodesic from  $\varphi(x)$  to  $\varphi(y)$ . Let  $\zeta(\lambda) = Ee^{\lambda\eta_e}$ . We assume the following.



**A2.  $\eta_e$  properties**

- (i)  $\eta_e$  is a continuous random variable.
- (ii) There exists  $\lambda > 0$  such that  $\zeta(\lambda) < \infty$ .

Here (i) guarantees that there is at most one geodesic from  $x$  to  $y$  a.s., for each  $x, y$ .

We do not assume that  $\sigma(r)$  in (1.4) is a power of  $r$ , but we do require that it “grows like  $r^\chi$ ” for some  $0 < \chi < 1$  in a sense similar to [2], as follows. We call a nonnegative function  $\{\sigma(r), r > 0\}$  *powerlike (with exponent  $\chi$ )* if there exist  $0 < \chi_1 < \chi < \chi_2$  and constants  $C_i$  such that

$$\lim_{r \rightarrow \infty} \frac{\log \sigma(r)}{\log r} = \chi \quad \text{and for all } s \geq r \geq C_{21}, \quad C_{22} \left(\frac{s}{r}\right)^{\chi_1} \leq \frac{\sigma(s)}{\sigma(r)} \leq C_{23} \left(\frac{s}{r}\right)^{\chi_2}. \quad (1.11)$$

If (1.11) holds with  $\chi_2 < 1$  we say  $\sigma(\cdot)$  is *sublinearly powerlike*.

The requirement that  $\sigma(\cdot)$  be sublinearly powerlike is perhaps not as stringent as it first appears, due to Lemma 1.2 below. It implies that if one is only interested in fluctuations at the level of exponents  $\chi, \xi$ , and one knows uniform the exponential tightness (1.4) but not necessarily the powerlike property for  $\sigma(\cdot)$ , then there is a sublinearly powerlike function with the same  $\chi$  for which (1.4) holds.

**Remark 1.1.** If  $\sigma(\cdot)$  is powerlike, then so is the increasing function  $\hat{\sigma}(r) = \sup_{s \leq r} \sigma(s)$ ; by further increasing  $\hat{\sigma}$  (though by at most a constant factor) we may make it strictly increasing and continuous while preserving the powerlike property. Therefore we may and do without loss of generality always assume  $\sigma(\cdot)$  is strictly increasing and continuous. The inverse function  $\Delta^{-1}$  is well-defined, and for  $\xi = (1 + \chi)/2 \in (\frac{1}{2}, 1)$  we have  $\Delta(r) \asymp r^\xi$  and  $\Delta^{-1}(a) \asymp a^{1/\xi}$  in the sense that

$$\lim_{r \rightarrow \infty} \frac{\log \Delta(r)}{\log r} = \xi, \quad \lim_{a \rightarrow \infty} \frac{\log \Delta^{-1}(a)}{\log a} = \frac{1}{\xi}.$$

The following is proved in section 8.

**Lemma 1.2.** Let  $\rho : (1, \infty) \rightarrow (0, \infty)$  satisfy

$$\lim_{r \rightarrow \infty} \frac{\log \rho_r}{\log r} = \chi \in (0, \infty).$$

Then there exist  $\rho^- \leq \rho \leq \rho^+$  which are powerlike with this same exponent  $\chi$  and strictly increasing, and

- (i) given  $\epsilon > 0$  we may take  $\rho^\pm$  to satisfy (1.11) with  $|\chi_i - \chi| < \epsilon$ ,  $i = 1, 2$ ,
- (ii) if  $\chi < 1$  then given  $0 \leq \delta < 1 - \chi$  we may take  $\rho^\pm$  such that  $\rho^\pm(r)/r^{1-\delta}$  is decreasing.

For general  $x, y \in \mathbb{R}^d$  not necessarily in  $\mathbb{V}$ , we write  $\Gamma_{xy}$  for  $\Gamma_{\varphi(x), \varphi(y)}$ . In general we view  $\Gamma_{xy}$  as an undirected path, but at times we will refer to, for example, the first point of  $\Gamma_{xy}$  with some property. Hence when appropriate, and clear from the context, we view  $\Gamma_{xy}$  as a path from  $\varphi(x)$  to  $\varphi(y)$ .

Our final standard assumption is the following.

**A3. Uniform exponential tightness** For some  $\sigma(\cdot)$  which is sublinearly powerlike with exponent  $\chi \in (0, 1)$ ,

$$P\left(|T(x, y) - ET(x, y)| \geq t\sigma(|y - x|)\right) \leq C_{24}e^{-C_{25}t} \quad \text{for all } x, y \in \mathbb{R}^d. \quad (1.12)$$



The isotropic property assumed for  $\mathbb{V}$  in A1 means that  $g(x) = \mu|x|$  for all  $x$ , where  $\mu = g(e_1)$ , and  $ET(0, x)$  depends only on  $|x|$ , so we define

$$h(r) = ET(0, re_1).$$

Let  $\mathfrak{B}_{d-1}$  be the Euclidean unit ball of  $\mathbb{R}^{d-1}$  and define the cylinders

$$G_r(K) = [0, r] \times K\Delta(r)\mathfrak{B}_{d-1}.$$

### 1.3 Statement of results

Here is our first main result. We write  $x_i$  for the  $i$ th coordinate of  $x \in \mathbb{R}^d$ .

**Theorem 1.3.** *Suppose  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  and  $\{\eta_e, e \in \mathbb{E}\}$  satisfy A1, A2, and A3. Then given  $\epsilon > 0$  there exist constants  $C_i = C_i(\epsilon)$ ,  $26 \leq i \leq 30$ , such that for all  $r \geq 1, K \geq 1, t \geq C_{26}K^2$ ,*

$$P\left(\left|T(x, y) - ET(x, y)\right| \geq t\sigma(r) \text{ for some } x, y \in G_r(K) \text{ with } |y - x| \geq \epsilon r\right) \leq C_{27}e^{-C_{28}t}, \tag{1.13}$$

and

$$P\left(T(x, y) \leq h(|(y - x)_1|) - t\sigma(r) \text{ for some } x, y \in G_r(K)\right) \leq C_{29}e^{-C_{30}t}. \tag{1.14}$$

Equation (1.14) is weaker than (1.13) in the sense that  $h(|(y - x)_1|)$  is “typically” smaller than  $ET(x, y) = h(|y - x|)$  (and always bounded by a constant multiple of it, see Lemma 3.6), but stronger in that it isn’t limited to  $|y - x| \geq \epsilon r$ .

We can split (1.13) into upward and downward deviations:

$$T(x, y) \geq ET(x, y) + t\sigma(r) \quad \text{and} \quad T(x, y) \leq ET(x, y) - t\sigma(r).$$

Then the downward-deviations part of (1.13) is a consequence of (1.14) and Proposition 3.3 below, because

$$x, y \in G_r(K), |y - x| \geq \epsilon r \implies \left| |y - x| - |(y - x)_1| \right| \leq C_{31}K^2\sigma(r) \tag{1.15}$$

for some  $C_{31}(\epsilon)$ . In general, if we think of  $x$  to  $y$  as an increment a path might make within  $G_r(K)$  in going from the one end to the other in the  $e_1$  direction, then  $(y - x)_1$  measures progress made by that increment in the  $e_1$  direction, so it is a natural normalization of  $T(x, y)$  in the context of such paths. It is also sufficient for application to the next two theorems. We restrict to  $|y - x| \geq \epsilon r$  in (1.13) so that we can use (1.15) as above, and so that, even for the shorter geodesic segments considered in our multiscale argument, the direction of the increment of the segment is always close to  $e_1$ ; see the discussion around (4.25) and (7.2). But the restriction to  $|y - x| \geq \epsilon r$  can be eliminated, as we now describe.

**Remark 1.4.** For (1.13) in Theorem 1.3 we can replace the conditions  $K \geq 1, t \geq C_{26}K^2$  with  $K \geq C_{32}, t \geq \epsilon K^2$ . This is because there exists  $m$  such that  $G_r(K)$  is contained in  $C_{32}\epsilon^{-m}$  “thin cylinders” (not necessarily oriented parallel to  $e_1$ ) of length between  $\epsilon r$  and  $r$  and radius  $K_\epsilon\Delta(r) = \sqrt{\epsilon/C_{26}}K\Delta(r)$  such that every pair  $x, y$  as in (1.13) is contained in one of these cylinders. Here  $C_{32}$  does not depend on  $K$  or  $r$ . We can apply the theorem to each thin cylinder, since  $t \geq C_{26}K_\epsilon^2$ , then sum over the thin cylinders.

This same idea of summing over small cylinders can be used to eliminate the condition  $|y - x| \geq \epsilon r$  in (1.13); to deal with pairs  $x, y$  with  $2^{-(k+1)}r < |y - x| \leq 2^{-k}r$  one can use cylinders of length  $2^{-(k-1)}r$  and apply (1.13), with  $2^{-(k-1)}r$  in place of  $r$  and  $t\sigma(r)/\sigma(2^{-(k-1)}r)$  in place of  $t$ , to each cylinder. We omit the details, though, as it seems not particularly natural to use  $\sigma(r)$  in (1.13) when  $|y - x|$  is much smaller than  $r$ .

We will use Theorem 1.3 together with a coarse-graining scheme in establishing the following.

**Theorem 1.5.** *Suppose  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  and  $\{\eta_e, e \in \mathbb{E}\}$  satisfy A1, A2, and A3. There exists  $C_{33}$  such that*

$$\mu|x| \leq h(|x|) \leq \mu|x| + C_{33}\sigma(|x|) \quad \text{for all } x \in \mathbb{R}^d. \tag{1.16}$$

The regularity required on  $\sigma(\cdot)$  in A3 is not compatible with a fluctuation exponent  $\chi = 0$ , which is as expected since (as noted above) results in [3] suggest that in that case (1.16) should not hold with  $\sigma$  of the order of the standard deviation.

As we have noted, for  $|y - x|$  of order  $r$ , if  $\Gamma_{xy}$  contains a vertex  $z$  at distance of order  $\gg \Delta(r)$  from  $\Pi_{xy}$  (not too near  $x$  or  $y$ ), then the associated extra distance  $g(z - x) + g(y - z) - g(y - x)$  traveled by the geodesic is of order  $\gg \Delta(r)^2/r = \sigma(r)$ , and by Theorem 1.5 the same is true for  $h$  in place of  $g$ . Since the corresponding passage times satisfy  $T(x, z) + T(z, y) - T(x, y) = 0$ , this means that either

$$T(x, y) - ET(x, y) \gg \sigma(r), \quad T(x, z) - ET(x, z) \ll -\sigma(r), \quad \text{or} \quad T(z, y) - ET(z, y) \ll -\sigma(r).$$

The assumption (1.12) says the first of these is unlikely, and Theorems 1.3 and 1.5 can be used to show it is unlikely that there exists a  $z$  for which the second or third occurs. (Not without complications, though, as we cannot assume  $z \in G_r$ .) This is the idea behind the following. For  $r, s > 0$  define intervals enlarging  $[0, r]$ :

$$\mathcal{I}_{r,s} = \begin{cases} [-s^2\sigma(r) \log r, r + s^2\sigma(r) \log r] & \text{if } s \leq (C_{34} \log r)^{1/2}; \\ [-s^2\sigma(r), r + s^2\sigma(r)] & \text{if } (C_{34} \log r)^{1/2} < s \leq r/\Delta(r); \\ [-s\Delta(r), r + s\Delta(r)] & \text{if } s > r/\Delta(r), \end{cases} \tag{1.17}$$

where  $C_{34}$  “sufficiently large” will be specified later, and

$$G_{r,s} = \mathcal{I}_{r,s} \times s\Delta(r)\mathfrak{B}_{d-1}.$$

For  $s > K$  we have  $G_r(K) \subset G_{r,s}$ ; in this case we may view  $G_{r,s}$  as being the cylinder  $G_r(K)$  fattened transversally to width  $s\Delta(r)$ , and lengthened by an amount which varies with the size of  $s$  relative to  $r$ , chosen to be “enough to make wandering of geodesics out of the cylinder end at least as unlikely as out of the sides.”

For  $x \in \mathbb{R}^d$  we write  $x^*$  for  $(x_2, \dots, x_d)$  so  $x = (x_1, x^*)$ .

**Theorem 1.6.** *Suppose  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  and  $\{\eta_e, e \in \mathbb{E}\}$  satisfy A1, A2, and A3. There exist  $C_{35}$ – $C_{38}$  such that for all  $K \geq C_{35}$ ,*

$$P(\Gamma_{xy} \not\subset G_{r,s} \text{ for some } x, y \in G_r(K) \text{ with } |(y - x)^*| \leq (y - x)_1) \leq \begin{cases} C_{36}e^{-C_{37}s^2} & \text{for all } C_{38}K \leq s \leq r/\Delta(r), \\ C_{36}e^{-C_{37}s\Delta(r)/\sigma(s\Delta(r))} & \text{for all } s > r/\Delta(r). \end{cases} \tag{1.18}$$

We include the condition  $|(y - x)^*| \leq (y - x)_1$  because we are primarily interested in transverse fluctuations of geodesics out the side of  $G_{r,s}$ , so we wish to avoid  $y - x$  oriented in a direction too far from  $e_1$ .

**Remark 1.7.** The strategy for proving Theorems 1.3–1.6 is as follows:

- (1) prove Theorem 1.3 for downward deviations—this is the most difficult part;
- (2) use Theorem 1.3 for downward deviations to prove Theorem 1.6 restricted to a fixed  $(x, y)$ ;
- (3) use Theorem 1.3 for downward deviations and the restricted Theorem 1.6 to prove Theorem 1.5;
- (4) use Theorem 1.5 to prove Theorem 1.3 for upward deviations;
- (5) use the full Theorem 1.3 to prove the unrestricted Theorem 1.6.

Step (1) occupies section 4; after some definitions and preliminary steps there we give a proof sketch for (1) in Remark 4.4, with further elaboration in Remark 4.8.

**Remark 1.8.** In [13] and [11], an alternate strategy was used to prove LPP analogs (in integrable cases) of Theorems 1.3 and 1.6 in  $d = 2$ . Theorem 1.5 was already known for that context—see the comments following (1.9). Essentially the strategy for the Theorem 1.3 analog in [13] is this, when translated to FPP: first the easier upward-deviations half of Theorem 1.3 is proved. For downward deviations, consider the points  $0$  and  $3re_1$ , and a cylinder  $\mathcal{G}_r$  of radius  $\Delta(r)$  with axis from  $re_1$  to  $2re_1$ . Suppose there are (random) vertices  $u, v \in \mathcal{G}_r$  with  $u_1 < v_1$  for which the passage time is fast:  $T(u, v) \leq h(|v - u|) - 3t\sigma(r)$  for some large  $t$ . From the upward-deviations half of Theorem 1.3, with high probability we have also  $T(x, u) \leq h(|u - x|) + t\sigma(r)$  and  $T(v, y) \leq h(|y - v|) + t\sigma(r)$ . From this, using that  $\sigma(r)$  is proportional to  $r^{1/3}$  and  $h(r) = \mu r + O(\sigma(r))$ , assuming  $t$  is large enough,

$$T(x, y) \leq h(|u - x|) + h(|v - u|) + h(|y - v|) - t\sigma(r) \leq h(|y - x|) - \frac{t}{2}\sigma(r). \quad (1.19)$$

This has probability exponentially small in  $t$ , by (1.12), hence so does the probability of such  $u, v$  existing. This strategy does not work for FPP, however, as it requires one to already *know* Theorem 1.5 to obtain the second inequality in (1.19); we instead need to establish the uniformity in Theorem 1.3 to *obtain* Theorem 1.5.

## 2 Existence of acceptable random graphs

We construct a random graph  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  satisfying A1. We begin by constructing the point process  $\mathbb{V}$  of vertices. We start with a “space-time” Poisson process  $\mathbb{V}_0$  (which we view as a random countable set) of density 1 with respect to Lebesgue measure in  $\mathbb{R}^d \times (0, \infty)$ . We say a point  $v \in \mathbb{R}^d$  appears at time  $t$  in  $\mathbb{V}_0$  if  $(v, t) \in \mathbb{V}_0$ , and we say  $A \subset \mathbb{R}^d$  is empty at time  $t-$  if no point in  $A$  appears in  $\mathbb{V}_0$  during  $(0, t)$ . We then define

$$\mathbb{V} = \left\{ v \in \mathbb{R}^d : \text{for some } t > 0 \text{ and } x \in \mathbb{R}^d \text{ with } v \in B_1(x), v \text{ appears at time } t \text{ in } \mathbb{V}_0 \right. \\ \left. \text{and } B_1(x) \text{ is empty at time } t- \right\}.$$

In other words, we keep in  $\mathbb{V}$  the  $\mathbb{R}^d$  coordinate  $v$  of a point of  $\mathbb{V}_0$  if  $v$  is the first point to appear in some ball of radius 1. Then almost surely, for each  $v \in \mathbb{V}$  there is a unique point  $(v, t_v^*) \in \mathbb{V}_0$ , and we view  $t_v^*$  as the time at which  $v$  appeared in  $\mathbb{V}_0$ . With probability one, every unit ball in  $\mathbb{R}^d$  contains a point of  $\mathbb{V}$ . We call  $\mathbb{V}$  the *available-space point process*.

For the set  $\mathbb{V}$  recall that  $\{Q_v, v \in \mathbb{V}\}$  denotes the corresponding Voronoi cells. More generally we write  $Q_x$  for the Voronoi cell containing any  $x$  (with some arbitrary convention if  $x$  is on the boundary of multiple cells), and  $\varphi(x)$  for the unique point of  $\mathbb{V}$  in  $Q_x$ . When convenient we view  $e = \langle x, y \rangle$  as the line segment joining  $x$  and  $y$ . Let  $B_r(x)$  denote the open Euclidean ball of radius  $r$  about  $x$ . For  $d = 2$ , the Delaunay graph of  $\mathbb{V}$  is our graph  $\mathbb{G}$ . For  $d \geq 3$  we fix  $0 < \delta_{\mathbb{G}} < 1$  and use

$$\mathbb{E} = \{ \langle x, y \rangle : d(Q_x, Q_y) \leq \delta_{\mathbb{G}} \}.$$

We call  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  the *augmented Delaunay graph* of  $\mathbb{V}$ . The edges in  $\mathbb{E}$  which are not in the Delaunay graph are called *augmentation edges*. We write  $x \sim y$  to denote that  $x, y$  are adjacent vertices in  $\mathbb{G}$ , and  $x \sim_{\text{Del}} y$  to denote adjacency in the Delaunay graph of  $\mathbb{V}$ . If  $y \sim_{\text{Del}} z$ , then for all  $u \in Q_y \cap Q_z, B_{|y-z|/2}(u) \cap \mathbb{V} = \emptyset$ . Hence by A1(ii),

$$|z - y| < 2 \text{ whenever } y \sim_{\text{Del}} z. \quad (2.1)$$

Similarly if  $z \in Q_y$  then  $B_{|z-y|}(z)$  contains no point of  $\mathbb{V}$ , so  $|z - y| < 1$ ; thus

$$Q_y \subset B_1(y). \tag{2.2}$$

**Remark 2.1.** The purpose of augmentation is roughly the following. Consider the line segment  $\Pi_{xy}$  between  $x, y \in \mathbb{V}$ . It passes through a sequence of Voronoi cells  $Q_x = Q_{x_0}, Q_{x_1}, \dots, Q_{x_m} = Q_y$ , and there is a corresponding path  $x = x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_m = y$  in the Delaunay graph. If too many of the cells  $Q_{x_j}$  are “thin,” then the Delaunay path length  $\sum_{j=1}^m |x_j - x_{j-1}|$  may be much greater than  $|y - x|$ , making it difficult to bound the dilation. The augmentation effectively allows paths in  $\mathbb{G}$  that skip over such problematic sequences of cells, at least for a small distance, enabling us to prove bounded dilation while preserving other properties of the Delaunay triangulation. We can reduce the occurrence of augmentation to involve an arbitrarily small proportion of vertices by using a small enough  $\delta > 0$ , but we will not give details of that fact here.

For  $A \subset \mathbb{R}^d$  and  $r > 0$  let  $A^r = \{x : d(x, A) < r\}$ .

**Proposition 2.2.** *The augmented Delaunay graph of the available-space point process satisfies A1.*

*Proof.* A1(i) for  $\mathbb{V}$  follows from the same properties for  $\mathbb{V}_0$ ; since  $\mathbb{G}$  is constructed from  $\mathbb{V}$  via isotropic and translation-invariant local rules, A1(i) also holds for  $\mathbb{G}$ . A1(ii) follows from the fact that the first point of  $\mathbb{V}_0$  to appear in any radius-1 ball is always a point of  $\mathbb{V}$ .

To prove A1(iii), observe first that by (2.1), given  $x \in \mathbb{V}$  the cell  $Q_x$  and all Delaunay edges  $\langle x, y \rangle$  are determined by  $\mathbb{V} \cap B_2(x)$ . Therefore the collection of Voronoi cells intersecting  $B_3(x)$  is determined by  $\mathbb{V} \cap B_5(x)$ , and hence so are all augmentation edges  $\langle x, y \rangle$ .  $\mathbb{V} \cap B_5(x)$ , in turn, is determined by  $\mathbb{V}_0 \cap (B_7(x) \times (0, \infty))$ , so for  $A \subset \mathbb{R}^d$ , the restriction  $(\mathbb{V}_A, \mathbb{E}_A)$  is determined by  $\mathbb{V}_0 \cap (A^7 \times (0, \infty))$ . Since  $\mathbb{V}_0$  is independent in disjoint regions, A1(iii) follows with  $\beta = 14$ .

Turning to A1(v), let  $q = 1/(1 + \lfloor 3\sqrt{d} \rfloor)$  and  $r > r_0 > 0$ , and let  $r_1 = r_0 \wedge 1$ . For  $x \in qr_1\mathbb{Z}^d$  let  $J_x$  denote the cube  $\prod_{i=1}^d [x_i, x_i + qr_1)$  and  $\mathcal{J}_x = \{J_y : y \in qr_1\mathbb{Z}^d, |y - x|_\infty = 2qr_1\}$ . This means the  $5^d - 3^d$  cubes in  $\mathcal{J}_x$  form a shell around  $J_x$ , with a smaller shell of cubes in between, and the diameter of  $\mathcal{J}_x$  is less than  $2r_1$ . Then any radius-1 ball intersecting  $J_x$  must contain a cube in  $\mathcal{J}_x$ . Letting  $t^*(J_x) = \max\{t_v^* : v \in \mathbb{V} \cap J_x\}$  it follows that at time  $t^*(J_x)$ , for some  $J_y \in \mathcal{J}_x$ , at least  $|\mathbb{V} \cap J_x|$  points of  $\mathbb{V}_0$  have appeared in  $J_x$  but none in  $J_y$ . Letting  $N_{xy}$  be the number of points of  $\mathbb{V}_0$  appearing in  $J_x$  before the first point of  $\mathbb{V}_0$  appears in  $J_y$ , this says that  $N_x := \max_{J_y \in \mathcal{J}_x} N_{xy} \geq |\mathbb{V} \cap J_x|$ . Since  $|\mathcal{J}_x| \leq 5^d$  it follows that

$$P(|\mathbb{V} \cap J_x| \geq n) \leq P(N_x \geq n) \leq \sum_{J_y \in \mathcal{J}_x} P(N_{xy} \geq n) \leq 5^d 2^{-n}, \tag{2.3}$$

and hence for  $\lambda < \log 2$ ,

$$Ee^{\lambda N_x} \leq \frac{5^d}{1 - \frac{e^\lambda}{2}}. \tag{2.4}$$

Let  $\mathcal{I}_0 = \{J_x : J_x \cap [-r, r)^d \neq \emptyset\}$ , so  $|\mathcal{I}_0| \leq (2(1 + \frac{r}{qr_1}))^d$  and  $B_r(0) \subset \cup_{J_x \in \mathcal{I}_0} J_x$ . Now  $\mathcal{I}_0$  is a lattice of cubes, and we divide it into  $5^d$  sublattices, each consisting of a cube  $J_x$  together with all its translates by vectors in  $5qr_1\mathbb{Z}^d$  which intersect  $[-r, r)^d$ . We label these sublattices of cubes as  $\mathcal{I}_{0,1}, \dots, \mathcal{I}_{0,5^d}$ , and the cardinality satisfies

$$|\mathcal{I}_{0,j}| \leq \left(2 \left(1 + \frac{r}{5qr_1}\right)\right)^d \tag{2.5}$$

for each  $j$ . We denote the corresponding union of cubes as  $I_{0,j} = \cup_{J_x \in \mathcal{I}_{0,j}} J_x$ . The spacing of the cubes in  $\mathcal{I}_{0,j}$  means that the shells  $\{J_x : J_x \in \mathcal{I}_{0,j}\}$  are disjoint, so the variables  $\{N_x : J_x \in \mathcal{I}_{0,j}\}$  are i.i.d.

For  $a \geq 1$ , taking  $\lambda = 1/5$  in (2.4) we obtain  $Ee^{\lambda N_x} \leq 3 \cdot 5^d$  and hence using (2.5), provided  $a$  is large (depending on  $r_1$ ),

$$\begin{aligned} P(|\mathbb{V} \cap B_r(0)| \geq ar^d) &\leq \sum_{j=1}^{5^d} P\left(|\mathbb{V} \cap I_{0,j}| \geq \frac{ar^d}{5^d}\right) \\ &\leq \sum_{j=1}^{5^d} P\left(\sum_{J_x \in \mathcal{I}_{0,j}} N_x \geq \frac{ar^d}{5^d}\right) \\ &\leq 5^d (3 \cdot 5^d)^{|\mathcal{I}_{0,j}|} e^{-ar^d/5^{d+1}} \\ &\leq e^{-ar^d/2 \cdot 5^{d+1}}, \end{aligned} \tag{2.6}$$

proving A1(v).

Finally we consider A1(iv). Let  $x, y \in \mathbb{V}$ . As in Remark 2.1,  $\Pi_{xy}$  passes through a sequence of Voronoi cells  $Q_x = Q_{x_0}, Q_{x_1}, \dots, Q_{x_m} = Q_y$ , and there is a corresponding path  $x = x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_m = y$  in the Delaunay graph. (There is probability 0 that  $\Pi_{xy}$  intersects some  $Q_u$  in just a single point, so we will ignore this possibility, meaning that “passes through” here is unambiguous.) For  $j < m$  let  $a_j$  be the first point of  $Q_{x_j}$  in  $\Pi_{xy}$  and let  $a_m = y$ , so by convexity of cells,  $Q_{x_j} \cap \Pi_{xy} = [a_j, a_{j+1}]$  for all  $0 \leq j < m$ . We select indices  $0 = j(0) < j(1) < \dots < j(\ell) = m$  iteratively, taking  $j(k+1)$  as the least index  $j > j(k)$  for which either  $|a_{j+1} - a_{j(k)+1}| > \delta_{\mathbb{G}}$  or  $j = m$ . Then  $\langle x_{j(k)}, x_{j(k+1)} \rangle$  is always a Delaunay or augmentation edge, so we consider the path  $x = x_{j(0)} \rightarrow x_{j(1)} \rightarrow \dots \rightarrow x_{j(\ell)} = y$  in  $\mathbb{G}$ ; in particular we want to bound  $|x_{j(k+1)} - x_{j(k)}|$  relative to  $|a_{j(k+1)+1} - a_{j(k)+1}|$  for  $0 \leq k \leq \ell - 2$ .

For  $k \leq \ell - 1$  we have using (2.2)

$$|x_{j(k+1)} - x_{j(k)}| \leq |x_{j(k+1)} - a_{j(k+1)}| + |a_{j(k+1)} - a_{j(k)+1}| + |a_{j(k)+1} - x_{j(k)}| \leq 2 + \delta_{\mathbb{G}}$$

so for  $k \leq \ell - 2$ ,

$$|a_{j(k+1)+1} - a_{j(k)+1}| > \delta_{\mathbb{G}} \geq \frac{\delta_{\mathbb{G}}}{2 + \delta_{\mathbb{G}}} |x_{j(k+1)} - x_{j(k)}|.$$

Therefore if  $\ell \geq 2$ ,

$$\begin{aligned} \sum_{k=0}^{\ell-2} |x_{j(k+1)} - x_{j(k)}| &\leq \frac{2 + \delta_{\mathbb{G}}}{\delta_{\mathbb{G}}} \sum_{k=0}^{\ell-2} |a_{j(k+1)+1} - a_{j(k)+1}| \\ &= \frac{2 + \delta_{\mathbb{G}}}{\delta_{\mathbb{G}}} |a_{j(\ell-1)+1} - a_1| \leq \frac{2 + \delta_{\mathbb{G}}}{\delta_{\mathbb{G}}} |y - x|. \end{aligned} \tag{2.7}$$

Having  $\ell \geq 2$  also ensures  $|y - x| \geq |a_{j(1)} - a_{j(0)}| > \delta_{\mathbb{G}}$  so

$$\begin{aligned} |x_{j(\ell)} - x_{j(\ell-1)}| &= |y - x_{j(\ell-1)}| \leq |y - a_{j(\ell-1)+1}| + |a_{j(\ell-1)+1} - x_{j(\ell-1)}| \\ &\leq |y - x| + 1 \leq \frac{1 + \delta_{\mathbb{G}}}{\delta_{\mathbb{G}}} |y - x| \end{aligned}$$

which with (2.7) yields

$$\sum_{k=0}^{\ell-1} |x_{j(k+1)} - x_{j(k)}| \leq \frac{3 + 2\delta_{\mathbb{G}}}{\delta_{\mathbb{G}}} |y - x|. \tag{2.8}$$

On the other hand, if  $\ell = 1$  then

$$\sum_{k=0}^{\ell-1} |x_{j(k+1)} - x_{j(k)}| = |y - x|$$

so (2.8) still holds. This proves bounded dilation. □

### 3 Straightness of geodesics and and regularity of means

For  $q > 0$  and  $x \in \mathbb{R}^d$ , let

$$\begin{aligned} \psi_q(x) &= \text{the point of } q\mathbb{Z}^d \text{ closest to } x \text{ (with ties broken arbitrarily),} \\ F_y &= \psi_q^{-1}(y), y \in q\mathbb{Z}^d, \end{aligned} \tag{3.1}$$

the latter being a cube (ignoring the boundary.) More generally for  $u \in \mathbb{R}^d$  we define  $F_u$  to be the cube  $F_y$  containing  $u$ , with some arbitrary rule for cube-boundary points.

The bound (1.4) applies to deterministic  $x, y$ ; we cannot for example take  $x, y \in \mathbb{V}$ . Instead for random  $x, y$  we can apply (1.4) to nearby points of  $q\mathbb{Z}^d$  for some  $q$ , and use the following. It is the only place the assumption A2(iv) of bounded dilation is used.

**Lemma 3.1.** *Suppose  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  and  $\{\eta_e, e \in \mathbb{E}\}$  satisfy A1, A2, and A3. There exist constants  $C_{39}$ – $C_{42}$  as follows.*

(i) *Let  $r \geq 2$  and  $t \geq C_{39} \log r$ . Then*

$$P(\text{there exist } x, y \in B_r(0) \cap \mathbb{V} \text{ with } T(x, y) \geq tr) \leq C_{40}e^{-C_{41}t}.$$

(ii) *For all  $x, y \in \mathbb{R}^d$ ,*

$$ET(x, y) \leq C_{42}(|y - x| \vee 1).$$

*Proof.* To prove (i) we condition on  $\mathbb{V}$ . Given  $x, y \in B_r(0) \cap \mathbb{V}$ , by A1(iv) there exists a path  $x = x_0, x_1, \dots, x_m = y$  in  $\mathbb{G}$  with

$$\sum_{j=1}^m |x_j - x_{j-1}| \leq C_{18}|y - x| \leq 2C_{18}r.$$

Writing  $\eta_j$  for  $\eta_{\langle x_{j-1}x_j \rangle}$ , for  $\lambda > 0$  we have

$$\begin{aligned} P\left(T(x, y) \geq t|y - x| \mid \mathbb{V}\right) &\leq P\left(\sum_{j=1}^m |x_j - x_{j-1}| \eta_j \geq t|y - x| \mid \mathbb{V}\right) \\ &\leq e^{-\lambda t|y-x|} \prod_{j=1}^m \zeta(\lambda|x_j - x_{j-1}|) \\ &\leq e^{-\lambda t|y-x|} \zeta(C_{18}\lambda|y - x|), \end{aligned} \tag{3.2}$$

so using A2(ii) it follows that for some  $c_i$ ,

$$P\left(T(x, y) \geq t|y - x| \mid \mathbb{V}\right) \leq e^{-I_\eta(t/C_{18})} \leq c_0e^{-c_1t}, \tag{3.3}$$

where  $I_\eta(t) = \sup_{\gamma > 0} (\gamma t - \log \zeta(\gamma))$  is the large-deviations rate function of the variables  $\eta_e$ . From this, using A1(v),

$$P(\text{there exist } x, y \in B_r(0) \cap \mathbb{V} \text{ with } T(x, y) \geq tr) \leq \sum_{n=2}^{\infty} P(|B_r(0) \cap \mathbb{V}| = n) \binom{n}{2} c_1e^{-c_1t/2}$$

$$\begin{aligned} &\leq c_1 E|B_r(0) \cap \mathbb{V}|^2 e^{-c_1 t/2} \\ &\leq c_2 r^4 e^{-c_2 t/2} \\ &\leq c_2 e^{-c_3 t}, \end{aligned} \tag{3.4}$$

proving (i).

To prove (ii) we apply (3.3) to  $\varphi(x)$  and  $\varphi(y)$ , and use the fact that  $|\varphi(x) - x| \leq 1$  for all  $x$ . □

Building on Lemma 3.1 we obtain the following.

**Lemma 3.2.** *Suppose  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  and  $\{\eta_e, e \in \mathbb{E}\}$  satisfy A1, A2, and A3. There exist constants  $C_{43}$ – $C_{45}$  as follows. For all  $r \geq 2, u, v \in \mathbb{R}^d$  and  $t > 0$  with  $\sigma(|u - v|) \geq r$  and  $t\sigma(|u - v|) \geq C_{43}r \log r$ ,*

$$\begin{aligned} &P\left(\text{there exist } x \in B_r(u) \cap \mathbb{V}, y \in B_r(v) \cap \mathbb{V} \text{ with } |T(x, y) - ET(u, v)| \geq t\sigma(|u - v|)\right) \\ &\leq C_{44}e^{-C_{45}t}. \end{aligned} \tag{3.5}$$

*Proof.* Using Lemma 3.1(i) and (1.4) we have

$$\begin{aligned} &P\left(\text{there exist } x \in B_r(u) \cap \mathbb{V}, y \in B_r(v) \cap \mathbb{V} \text{ with } |T(x, y) - ET(u, v)| \geq t\sigma(|u - v|)\right) \\ &\leq P\left(\text{there exists } x \in B_r(u) \cap \mathbb{V} \text{ with } T(x, u) \geq t\sigma(|u - v|)/3\right) \\ &\quad + P\left(\text{there exists } y \in B_r(v) \cap \mathbb{V} \text{ with } T(y, v) \geq t\sigma(|u - v|)/3\right) \\ &\quad + P\left(|T(u, v) - ET(u, v)| \geq t\sigma(|u - v|)/3\right) \\ &\leq 2C_{40}e^{-C_{41}t\sigma(|u - v|)/3r} + C_{24}e^{-C_{25}t/3} \\ &\leq c_0e^{-c_1t}. \end{aligned} \tag{3.6}$$

□

We write

$$x \rightarrow y_1 \rightarrow \cdots \rightarrow y_k \rightarrow y \text{ in } \Gamma_{xy}$$

to denote that the vertices  $y_i \in \mathbb{V}$  appear in the order  $y_1, \dots, y_k$  in traversing  $\Gamma_{xy}$  from  $x$  to  $y$ . For  $a$  preceding  $b$  in  $\Gamma_{xy}$  we write  $\Gamma_{xy}[a, b]$  for the segment of  $\Gamma_{xy}$  from  $a$  to  $b$ . (Here we do not require  $a, b \in \mathbb{V}$ .) For  $v$  in a geodesic  $\Gamma_{xy}$  and  $0 < s < |v - x|$ , let  $u$  be the first vertex in  $\Gamma_{xy} \cap \mathbb{V}$  before  $v$  satisfying  $\Gamma_{uv} \subset B_s(v)$ . We then call  $\Gamma_{uv}$  the *trailing  $s$ -segment* of  $v$  in  $\Gamma_{xy}$ . Note that by (2.1),  $s \geq |u - v| \geq s - 2$ . Define the hyperplanes, slabs, and halfspaces

$$\begin{aligned} H_s &= \{(x_1, x^*) \in \mathbb{R}^d : x_1 = s\}, \quad H_{[r,s]} = \{(x_1, x^*) \in \mathbb{R}^d : r \leq x_1 \leq s\}, \\ H_s^+ &= \{(x_1, x^*) \in \mathbb{R}^d : x_1 \geq s\}, \quad H_s^- = \{(x_1, x^*) \in \mathbb{R}^d : x_1 \leq s\}. \end{aligned}$$

We turn next to a weaker version of Theorem 1.5 similar to (1.8); we will need it on the way to the proof of Theorem 1.5. The proof of (1.8) in [1] does not carry over immediately to the present situation, as it uses translation invariance of the lattice. But we can use radial symmetry here to give a distinctly shorter proof than that of (1.8) in [1].

Recall  $h(r) - \mu r$  is nonnegative by subadditivity of  $h$ . For technical convenience we assume  $\beta > 4$  in A1(ii).

**Proposition 3.3.** *Suppose  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  and  $\{\eta_e, e \in \mathbb{E}\}$  satisfy A1, A2, and A3. There exists  $C_{46}$  such that for all  $r \geq 2$ ,*

$$\mu r \leq h(r) \leq \mu r + C_{46}\sigma(r) \log r. \tag{3.7}$$



*Proof.* It is sufficient to prove the bound for all sufficiently large  $r$ , so we will tacitly assume  $r$  is large, as needed.

We consider the geodesic  $\Gamma_{0,nre_1}$  for a fixed large  $n$ . Let  $\beta$  be as in A1(iii). For  $0 \leq j \leq n - 1$  let  $v_j$  be the first vertex  $v$  in  $\Gamma_{0,nre_1} \cap \mathbb{V}$  with the property that the trailing  $(r - 2\beta)$ -segment of  $v$  in  $\Gamma_{0,nre_1}$  is contained in  $H_{jr}^+$ , and let  $u_j$  be the starting point of this trailing segment. From (2.1) it is easy to see that we must have  $jr \leq (u_j)_1 \leq jr + 2$ . It follows that  $\Gamma_{0,nre_1}[u_j, v_j] \subset B_{r-2\beta}(v_j) \cap H_{jr}^+ \cap H_{(j+1)r-2\beta+2}^-$ ; we denote this last region by  $W_{v_j}$ . Note that any two sets  $W_{v_j}$  are separated by distance at least  $2\beta - 2$ .

We need to control the entropy of the collection of pairs  $\{(u_j, v_j) : 0 \leq j < n\}$ . To do this, we enlarge this collection in such a way that we can put a natural tree structure on it. To that end, let  $v_n$  be the first vertex in  $\Gamma_{0,nre_1} \cap \mathbb{V}$ , if one exists, which is at distance at least  $r$  from  $\cup_{0 \leq j < n} W_{v_j}$ . Then let  $u_n \in \mathbb{V}$  be such that  $\Gamma_{0,nre_1}[u_n, v_n]$  is the trailing  $(r - 2\beta)$ -segment of  $v_n$  in  $\Gamma_{0,nre_1}$ , and let  $W_{v_n} = B_{r-2\beta}(v_n)$ , which contains  $\Gamma_{0,nre_1}[u_n, v_n]$ . We repeat this to obtain  $(u_{n+1}, v_{n+1}), \dots, (u_N, v_N)$ , stopping when no  $v_{N+1}$  exists. Note this preserves the property of separation by distance at least  $2\beta - 2$ . Also, each new  $v_j$  is within distance  $2r - 2\beta + 2$  of some already-existing  $v_i$ .

We define the discrete approximations  $\hat{u}_j = \psi_{1/\sqrt{d}}(u_j)$ ,  $\hat{v}_j = \psi_{1/\sqrt{d}}(v_j)$ , and then let

$$\hat{W}_j = \begin{cases} B_{r-2\beta+1}(\hat{v}_j) \cap H_{jr-1}^+ \cap H_{(j+1)r-2\beta+3}^- & \text{if } 0 \leq j < n, \\ B_{r-2\beta+1}(\hat{v}_j) & \text{if } n \leq j \leq N. \end{cases}$$

which contains  $W_{v_j}$ ; the sets  $\hat{W}_j$ ,  $0 \leq j \leq N$  are separated from each other by distance at least  $2\beta - 6$ . We call  $\{(\hat{u}_j, \hat{v}_j) : 0 \leq j < n\}$  **primary pairs**, and  $\{(\hat{u}_j, \hat{v}_j) : n \leq j \leq N\}$  **secondary pairs**. Since  $r - 2\beta \geq |u_j - v_j| \geq r - 2\beta - 2$ , we have

$$r > r - 2\beta + 2 \geq |\hat{u}_j - \hat{v}_j| \geq r - 2\beta - 4 > r - 3\beta. \tag{3.8}$$

We now make a graph with vertices  $\{(\hat{u}_j, \hat{v}_j) : 0 \leq j \leq N\}$  by placing an edge between the  $i$ th and  $j$ th pairs if  $|\hat{v}_j - \hat{v}_i| \leq 4r$ . The construction ensures that the resulting graph is connected, and it is easy to see that the disjointness of the sets  $\hat{W}_j$  means the number of neighbors of any pair is bounded by some  $c_0$ . We label  $(u_0, v_0)$  as the root, and by some arbitrary algorithm, we take a spanning tree of the graph, which we denote  $\mathcal{T}(\Gamma_{0,nre_1})$ . For counting purposes, we view two such trees as the same if they have the same pairs  $\{(\hat{u}_j, \hat{v}_j) : 0 \leq j \leq N\}$ , and the same set of primary pairs. We define *parents* and *offspring* in this rooted tree in the usual way: for a given pair  $(\hat{u}_j, \hat{v}_j)$ , its parent is the first pair after  $(\hat{u}_j, \hat{v}_j)$  in the unique path from  $(\hat{u}_j, \hat{v}_j)$  to the root, and its offspring are those pairs having  $(\hat{u}_j, \hat{v}_j)$  as parent.

The tree  $\mathcal{T}(\Gamma_{0,nre_1})$  determines what we will call an abstract tree, in which all that is specified is the number of offspring of the root, then the number of offspring of each of these offspring, etc. The number of possible abstract trees here with  $N + 1$  vertices is at most  $c_0^N$ , and, provided  $r$  is large, for each such abstract tree the number of corresponding actual trees  $\mathcal{T}(\Gamma_{0,nre_1})$  is at most  $((8r)^d)^{2N+2}$ , so the number of trees  $\mathcal{T}(\Gamma_{0,nre_1})$  consisting of  $N + 1$  pairs is at most  $(c_1 r)^{2d(N+1)}$ .

We now consider all possible trees  $\mathcal{T}_0$ , with vertices  $\{(\hat{u}_j, \hat{v}_j) : 0 \leq j \leq N\}$ , and with  $\{(\hat{u}_j, \hat{v}_j) : 0 \leq j < n\}$  primary. Let  $v(\mathcal{T}_0)$  denote the number of vertices in the tree  $\mathcal{T}_0$ . We have

$$\begin{aligned} P\left(T(0, nre_1) \leq n(\mu r + 1)\right) &= \sum_{N \geq n} \sum_{\mathcal{T}_0: v(\mathcal{T}_0)=N} P\left(T(0, nre_1) \leq n(\mu r + 1), \mathcal{T}(\Gamma_{0,nre_1}) = \mathcal{T}_0\right) \\ &\leq \sum_{N \geq n} (c_1 r)^{2d(N+1)} \max_{\mathcal{T}_0: v(\mathcal{T}_0)=N} P\left(T(0, nre_1) \leq n(\mu r + 1), \mathcal{T}(\Gamma_{0,nre_1}) = \mathcal{T}_0\right) \end{aligned} \tag{3.9}$$

We proceed by contradiction: we want to show that for some  $C_{46}$ , if  $h(r) > \mu r + C_{46}\sigma(r) \log r$  then the right side of (3.9) approaches 0 as  $n \rightarrow \infty$ , which means that  $\limsup_n T(0, nre_1)/n \geq \mu r + 1$  a.s., contradicting the definition of  $\mu$ .

Thus suppose

$$h(r) > \mu r + 2C_{46}\sigma(r) \log r, \tag{3.10}$$

with  $C_{46}$  to be specified, and fix  $\mathcal{T}_0$  with vertices  $\{(\hat{u}_j, \hat{v}_j) : 0 \leq j \leq N\}$  with  $N \geq n$ . For  $u, v \in \mathbb{V} \cap \hat{W}_j$  let

$$T_{\hat{W}_j}(u, v) = \inf\{T(\Gamma) : \Gamma \text{ is a path from } u \text{ to } v \text{ in } \mathbb{G}, \Gamma \subset \hat{W}_j\},$$

taking the value  $\infty$  when there is no such path.  $T_{\hat{W}_j}(u, v)$  is determined by the configuration in

$$\hat{W}_j^1 = \{u \in \mathbb{R}^2 : d(u, \hat{W}_j) < 1\}.$$

There must exist  $u_j, v_j \in \hat{W}_j \cap \mathbb{V}$  satisfying  $|u_j - \hat{u}_j| \leq 1, |v_j - \hat{v}_j| \leq 1, \Gamma_{u_j, v_j} \subset \Gamma_{0, nre_1} \cap \hat{W}_j$  and

$$\sum_{j=0}^N T_{\hat{W}_j}(u_j, v_j) \leq T(0, nre_1).$$

It follows that, letting

$$\hat{T}_j = \inf\{T_{\hat{W}_j}(u_j, v_j) : u_j, v_j \in \mathbb{V}, |u_j - \hat{u}_j| \leq 1, |v_j - \hat{v}_j| \leq 1\},$$

the  $\hat{T}_j$ 's are independent (since the sets  $\hat{W}_j^1$  are separated from each other by distance more than  $2\beta - 8 > \beta$ ) and satisfy

$$\sum_{j=0}^N \hat{T}_j \leq T(0, nre_1). \tag{3.11}$$

Each  $\hat{T}_j$  is stochastically larger than

$$T_j = \inf\{T(u_j, v_j) : u_j, v_j \in \mathbb{V}, |u_j - \hat{u}_j| \leq 1, |v_j - \hat{v}_j| \leq 1\}$$

From (1.11) and (3.8) we have for some  $c_2$  that  $\sigma(|\hat{u}_j - \hat{v}_j|) \leq c_2\sigma(r)$ . From (3.8), (3.10), and subadditivity we also have

$$ET(\hat{u}_j, \hat{v}_j) \geq h(r) - h(r - |\hat{u}_j - \hat{v}_j|) \geq h(r) - c_3 \geq \mu r + 2C_{46}\sigma(r) \log r - c_3 \geq \mu r + C_{46}\sigma(r) \log r.$$

Hence from Lemma 3.2 we obtain that provided  $r$  is large, for all  $t \geq 1$ ,

$$\begin{aligned} P(T_j \leq \mu r + C_{46}\sigma(r) \log r - t\sigma(r)) &\leq P(T_j \leq ET(\hat{u}_j, \hat{v}_j) - t\sigma(r)) \\ &\leq P(T_j \leq ET(\hat{u}_j, \hat{v}_j) - c_2^{-1}t\sigma(|\hat{u}_j - \hat{v}_j|)) \\ &\leq C_{44}e^{-C_{45}t/c_2}. \end{aligned} \tag{3.12}$$

By increasing  $C_{44}$  we may make this valid for all  $t > 0$ . We then have for  $\lambda = C_{45}/2c_2\sigma(r)$  and  $M = \mu r + C_{46}\sigma(r) \log r$ ,

$$\begin{aligned} Ee^{-\lambda \hat{T}_j} &\leq Ee^{-\lambda T_j} \\ &= \int_0^\infty P(T_j \leq x) \lambda e^{-\lambda x} dx \\ &= e^{-\lambda M} \int_{-\infty}^{M/\sigma(r)} P(T_j \leq M - t\sigma(r)) \lambda \sigma(r) e^{\lambda \sigma(r)t} dt \end{aligned}$$

$$\begin{aligned} &\leq e^{-\lambda M} \left[ \int_{-\infty}^0 \lambda \sigma(r) e^{\lambda \sigma(r)t} dt + \int_0^{M/\sigma(r)} C_{44} e^{-C_{45}t/c_2} \lambda \sigma(r) e^{\lambda \sigma(r)t} dt \right] \\ &\leq (1 + C_{44}) e^{-\lambda M}. \end{aligned} \tag{3.13}$$

Recalling (3.9) and (3.11) we then have

$$\begin{aligned} P\left(T(0, nre_1) \leq n(\mu r + 1), \mathcal{T}(\Gamma_{0, nre_1}) = \mathcal{T}_0\right) &\leq P\left(\sum_{j=0}^N \hat{T}_j \leq n(\mu r + 1)\right) \\ &\leq e^{\lambda(\mu r + 1)n} (1 + C_{44})^{N+1} e^{-\lambda M(N+1)}, \end{aligned} \tag{3.14}$$

so by (3.9), provided  $r$  (and hence  $\lambda M$ ) and  $C_{46}$  (in (3.10)) are large,

$$\begin{aligned} P\left(T(0, nre_1) \leq n(\mu r + 1)\right) &\leq \sum_{N \geq n} (c_1 r)^{2d(N+1)} e^{\lambda(\mu r + 1)n} (1 + C_{44})^{N+1} e^{-\lambda M(N+1)} \\ &\leq 2(c_1 r)^{2d(n+1)} e^{\lambda(\mu r + 1)n} (1 + C_{44})^{n+1} e^{-\lambda M(n+1)} \\ &\leq e^{-n \log r}. \end{aligned} \tag{3.15}$$

As we have noted, since this approaches 0 as  $n \rightarrow \infty$ , it contradicts the fact that  $T(0, nre_1)/n \rightarrow \mu r$  a.s. Thus (3.10) must be false.  $\square$

We need to use a result from [2] to the effect that “geodesics are very straight.” It is proved there for FPP on a lattice, but the proof goes through essentially unchanged for the present context. The heuristics are as follows: suppose the geodesic  $\Gamma_{0, re_1}$  passes through a vertex  $u = (u_1, u^*)$  at distance  $s$  from  $\Pi_{0, re_1}$ ; by symmetry we may suppose  $u \in H_{r/2}^-$ . The geodesic then travels a corresponding extra distance  $g(u) + g(re_1 - u) - g(re_1)$ . If the angle between  $u$  and  $re_1$  is small, this extra distance is of order  $|u^*|^2/u_1$ , and from (1.12), the cost of this (meaning log of the probability) is of order  $|u^*|^2/u_1 \sigma(u_1)$ . If instead the angle between  $u$  and  $re_1$  is not small, the extra distance is of order  $|u|$  and the cost is of order  $|u|/\sigma(|u|)$ . We can combine these into a single statement by saying the cost for general  $u$  should be whichever of these two costs is smaller, at least for  $u \in H_{[0, r]}$ .

The exact formulation of the straightness result contains extra log factors relative to the preceding heuristic, due to the need to bound the probability for all  $u$  simultaneously. It is as follows, using the constants  $C_i, \chi_i$  of (1.11). Define  $\sigma^*(s)$  and  $\Phi(s)$  by

$$\Phi(s) = \frac{s}{C_{23} \sigma^*(s) \log(2 + s)} = \frac{s^{\chi_2}}{C_{23}} \sup_{t \leq s} \frac{t^{1-\chi_2}}{\sigma(t) \log(2 + t)}.$$

Here factoring out a power of  $s$  on the right, and the use of the sup, ensure that  $\Phi$  is strictly increasing. Note that by (1.11) we have

$$C_{23}^{-1} \sigma(s) \leq \sigma^*(s) \leq \sigma(s). \tag{3.16}$$

Then define

$$\Xi(s) = (s \sigma(s) \log(2 + s))^{1/2}, \quad D(u) = \begin{cases} \min\left(\frac{|u^*|^2}{\Xi(s)}, \Phi(\max(|u_1|, |u^*|))\right) & \text{if } u_1 \geq 0, \\ \Phi(\max(|u_1|, |u^*|)) & \text{if } u_1 < 0. \end{cases} \tag{3.17}$$

Note that by (1.11) and (3.16), for large  $s$ ,

$$\frac{1}{C_{23} \Phi(s)} \leq \left(\frac{\Xi(s)}{s}\right)^2 \leq \frac{1}{\Phi(s)}. \tag{3.18}$$

The “min” in the definition of  $D$  is in accordance with our heuristic: from [2], we have for large  $|u|$  that for  $C_{23}$  from (1.11),

$$D(u) = \begin{cases} \Phi(\max(|u_1|, |u^*|)) & \text{if } |u^*| \geq u_1, \\ \frac{|u^*|^2}{\Xi(u_1)^2} & \text{if } |u^*| \leq C_{23}^{-1/2}u_1. \end{cases} \tag{3.19}$$

Finally, define the symmetric version of  $D$ :

$$D_r(u) = \begin{cases} D(u) & \text{if } u_1 \leq \frac{r}{2} \\ D(re_1 - u) & \text{if } u_1 > \frac{r}{2}. \end{cases} \tag{3.20}$$

This makes the right half of the region  $\{u : D_r(u) \leq c\}$  the mirror image of the left half; this region is a “tube” (narrower near the ends) surrounding the line from 0 to  $re_1$  bounded by the shell  $\{u : |u^*| = c^{1/2}\Xi(u_1)\}$ , augmented by a cylinder of radius  $\Phi^{-1}(c)$  and length  $2\Phi^{-1}(c)$  around each endpoint, so we will call it a *tube-and-cylinders region*.

We will also consider tube-and-cylinders regions around general pairs  $u, v$  in place of  $0, re_1$ . To that end, let  $\Theta_{uv} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  be translation by  $-u$  followed by some unitary transformation which takes  $v - u$  to the positive horizontal axis, so that  $\Theta_{uv}(u) = 0, \Theta_{uv}(v) = |v - u|e_1$ . (The particular choice of unitary transformation does not matter.) Then  $\Theta_{uv}^{-1}(\{w : D_{|v-u|}(w) \leq c\})$  is a tube-and-cylinders region containing  $\Pi_{uv}$ .

The proof of the acceptable-random-graphs version of the straightness bound is little changed from the lattice-FPP version in [2]; we can readily use Lemma 3.1(i) to change the result from “point-to-point” (say, 0 to  $re_1$ ) to “ball-to-ball” ( $B_1(0)$  to  $B_1(re_1)$ ), with a sup over  $x$  and  $y$ . The proof in [2] uses only A3, not the additional assumption there essentially that  $\chi$  from A3 is the same as the value of  $\lim_{|x| \rightarrow \infty} \log(\text{var}(T(0, x))^{1/2}) / \log |x|$ , and not the fact that the FPP there is on a lattice. We omit the details of the direct adaptation.

**Proposition 3.4.** *Suppose  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  and  $\{\eta_e, e \in \mathbb{E}\}$  satisfy A1, A2, and A3. There exist constants  $C_{47}, C_{48}$  as follows. For all  $r, t > 0$ ,*

$$P\left(\sup_{x \in B_1(0) \cap \mathbb{V}} \sup_{y \in B_1(re_1) \cap \mathbb{V}} \sup_{u \in \Gamma_{xy}} D_r(u) \geq t\right) \leq C_{47}e^{-C_{48}t \log t}. \tag{3.21}$$

We next bound transverse increments of passage times from 0, that is, increments which are (approximately) along the boundary of a ball of large radius  $r$ , over distances  $\ll \Delta(r)$ . The following lattice-FPP result from [2] carries over straightforwardly to the present context with the help of Lemma 3.1(i), and again we omit details. Recalling the cubes  $F_u$  of side  $q$  define

$$\hat{T}(u, v) = \min\{T(y, z) : y \in F_u, z \in F_v\}, \quad u, v \in \mathbb{R}^d.$$

**Proposition 3.5.** *Suppose  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  and  $\{\eta_e, e \in \mathbb{E}\}$  satisfy A1, A2, and A3. There exist constants  $C_{49} - C_{54}$  as follows. For all  $u, v \in q\mathbb{Z}^d$  with*

$$|u| \geq C_{49}, \quad |g(u) - g(v)| \leq C_{50}, \quad \text{and} \quad 3 \leq |u - v| \leq C_{51}\Delta(|u|), \tag{3.22}$$

and all  $\lambda \geq C_{52}$ , we have

$$P\left(\hat{T}(v, 0) - \hat{T}(u, 0) \geq \lambda\sigma(\Delta^{-1}(|u - v|)) \log |u - v|\right) \leq C_{53}e^{-C_{54}\lambda \log |u - v|}. \tag{3.23}$$

We next prove a seemingly obvious fact:  $h(r) = ET(0, re_1)$  is approximately increasing in  $r$ .

**Lemma 3.6.** Suppose  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  and  $\{\eta_e, e \in \mathbb{E}\}$  satisfy A1, A2, and A3. Given  $0 < \epsilon < 1$  there exist constants  $C_{55}, C_{56}$  as follows.

(i) Let  $q > (1 + \chi)/(1 - \chi)$ . Let  $r, s > 0$  satisfying  $\min(r, s) \geq C_{55}\epsilon^{-q}$ . Then

$$h(r + s) \geq h(r) + (1 - \epsilon)h(s). \tag{3.24}$$

Here  $C_{55} = C_{55}(q)$ .

(ii) For all  $r, s \geq 0$ ,

$$h(r + s) \geq h(r) + (1 - \epsilon)h(s) - C_{56}. \tag{3.25}$$

**Remark 3.7.** We only need Lemma 3.6(ii); we include (i) because it is the main step in proving (ii) and may be useful elsewhere. Part (i) can be reformulated as follows: given  $\gamma > \chi/\xi$  there exist  $C_{57}(\gamma), C_{58}(\gamma)$  such that

$$h(r + s) \geq h(r) + h(s) - C_{57}s^\gamma \quad \text{for all } r \geq s \geq C_{58}.$$

Also, from Proposition 3.3 we have for some  $C_{59}$

$$h(r + s) \geq \mu(r + s) \geq h(r) + h(s) - 2C_{46}\sigma(r) \log r \quad \text{for all } r \geq s \geq 2,$$

but when  $s \leq \sigma(r) \log r$  this does not yield (3.24).

*Proof of Lemma 3.6.* We prove (i), then obtain (ii) as a straightforward consequence. The idea is to show that  $\Gamma_{0,(r+s)e_1}$  must with high probability approach  $(r + s)e_1$  approximately horizontally, which forces  $T(0, (r + s)e_1) - T(0, re_1)$  to be near  $h(s)$  with high probability; a non-horizontal approach would force the sup in (3.21) to be large.

Suppose (3.24) holds under the added condition  $s \leq r/4$ . Then for  $s > r/4$  we can take  $n$  with  $s/n \leq r/4 < s/(n - 1)$  and see that the hypotheses are satisfied with  $s/n$  in place of  $s$ . (This may require increasing  $C_{55}$ , but without dependence on  $n$ .) Applying (3.24)  $n$  times then yields

$$h(r + s) = h\left(r + n\frac{s}{n}\right) \geq h(r) + n(1 - \epsilon)h\left(\frac{s}{n}\right) \geq h(r) + (1 - \epsilon)h(s).$$

Therefore it is sufficient to prove the lemma for  $s \leq r/4$ . It is also sufficient to consider  $0 < \epsilon < \epsilon_0$  for any fixed  $\epsilon_0 = \epsilon_0(q) > 0$ .

With  $c_0$  to be specified, define

$$m = \inf \left\{ u > 0 : \sigma(u) \log u > \frac{\epsilon\mu s}{16c_0} \right\} \wedge \frac{r}{2}, \quad t = \epsilon\Phi(m)^{1/2}, \tag{3.26}$$

$$S_1 = \{w \in \mathbb{R}^d : r + s - m \leq w_1 \leq r + s - m + 2, D_{r+s}(w) \leq t\},$$

$$S_2 = \{w \in \mathbb{R}^d : r + s - m \leq w_1 \leq r + s - m + 2, |w^*| \leq \epsilon m\}.$$

Note that provided  $C_{55}$  (and hence  $s$  and  $m$ ) is large, which we henceforth tacitly assume, and provided  $\epsilon_0$  is small, we have  $\Phi^{-1}(t) < m/2$  and  $m \geq 2s$ . Suppose  $w \in S_1$ . We have  $|w_1 - (r + s)| \geq m - 2 > \Phi^{-1}(t)$  (meaning  $w$  lies to the left of the cylinder around  $(r + s)e_1$  in the tube-and-cylinders region  $\{u : D_{r+s}(u) \leq t\}$ ) and so by (3.18),

$$\frac{|w^*|}{m} \leq \frac{t^{1/2}\Xi(m)}{m} \leq \frac{t}{\Phi(m)^{1/2}} = \epsilon. \tag{3.27}$$

Thus  $S_1 \subset S_2$ . We let  $\hat{w} = (r + s - m, w^*)$ , so  $|w - \hat{w}| \leq 2$  for  $w \in S_2$ . See Figure 1. Then for such  $w$ , recalling  $m \geq 2s$ , we have

$$|\hat{w} - (r + s)e_1| - |\hat{w} - re_1| = (m^2 + |w^*|^2)^{1/2} - ((m - s)^2 + |w^*|^2)^{1/2} \geq \frac{sm}{(m^2 + |w^*|^2)^{1/2}} \geq (1 - \epsilon^2)s,$$

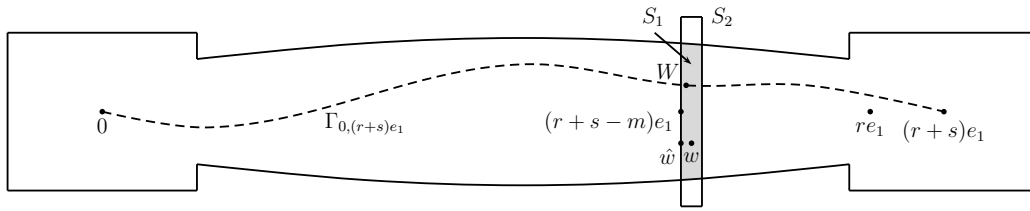


Figure 1: Diagram for the proof of Lemma 3.6.  $S_1$  is shaded.

so assuming  $C_{55}$  is large and  $\epsilon_0$  is small,

$$|w - (r + s)e_1| - |w - re_1| \geq (1 - \epsilon^2)s - 2 \geq \left(1 - \frac{\epsilon}{4}\right)s. \quad (3.28)$$

Also for  $w \in S_2$ ,  $m/2 \leq m - s \leq m - 2 \leq |w - re_1| \leq 2m$  and hence by (1.11),  $\sigma(|w - re_1|) \leq 3C_{23}\sigma(m - 2)$ . It follows that provided we choose  $c_0$  large enough in (3.26),

$$C_{46}\sigma(|w - re_1|) \log |w - re_1| \leq 4C_{23}C_{46}\sigma(m - 2) \log(m - 2) \leq \frac{\epsilon}{12}\mu s.$$

With (3.28) and Proposition 3.3, this yields that provided  $C_{55}$  is large,

$$\begin{aligned} h(|w - (r + s)e_1|) - h(|w - re_1|) &\geq \mu|w - (r + s)e_1| - \mu|w - re_1| - C_{46}\sigma(|w - re_1|) \log |w - re_1| \\ &\geq \left(1 - \frac{\epsilon}{3}\right)\mu s \\ &\geq \left(1 - \frac{\epsilon}{2}\right)h(s). \end{aligned} \quad (3.29)$$

We need a lower bound for  $t$ . Since  $q > \hat{q} = (1 + \chi)/(1 - \chi)$ , we can choose  $b > 0$  small enough so  $(q - 1)(1 - b)^2 > (1 + b)(\hat{q} - 1)$ . Provided  $C_{55}$  is large enough (in the lemma statement, and depending on  $b, q$ ), since  $s \geq C_{55}\epsilon^{-q}$  we have

$$m \geq (\epsilon s)^{(1-b)/\chi}, \quad \Phi(m) \geq m^{(1-b)(1-\chi)}, \quad t = \epsilon\Phi(m)^{1/2} \geq c_1\epsilon^{1-(1-b)^2(q-1)(1-\chi)/2\chi} \geq c_1\epsilon^{-b}. \quad (3.30)$$

Observe that for every vertex  $w \in \Gamma_{0,(r+s)e_1} \cap \mathbb{V}$  we have

$$T(0, (r + s)e_1) - T(0, re_1) = T(0, w) + T(w, (r + s)e_1) - T(0, re_1) \geq T(w, (r + s)e_1) - T(w, re_1). \quad (3.31)$$

Let

$$T^* = \min\{T(w, (r + s)e_1) - T(w, re_1) : w \in S_2 \cap V\}$$

and let  $W$  be the first vertex in  $\Gamma_{0,(r+s)e_1} \cap \mathbb{V}$  satisfying  $r + s - m \leq W_1 \leq r + s - m + 2$ . If  $W \in S_2$  then by (3.31) and subadditivity we have

$$\begin{aligned} T(0, (r + s)e_1) - T(0, re_1) &\geq T^*, \quad |T(0, (r + s)e_1) - T(0, re_1)| \leq T(re_1, (r + s)e_1), \\ |T^*| &\leq T(re_1, (r + s)e_1). \end{aligned}$$

Hence

$$\begin{aligned} h(r + s) - h(r) &= E[T(0, (r + s)e_1) - T(0, re_1)] \\ &\geq E(T^* \mathbf{1}_{\{W \in S_2\}}) - E(T(re_1, (r + s)e_1) \mathbf{1}_{\{W \notin S_2\}}) \\ &\geq E(T^*) - 2E(T(re_1, (r + s)e_1) \mathbf{1}_{\{W \notin S_2\}}). \end{aligned} \quad (3.32)$$

It follows from (1.4) that for some  $c_2$

$$E(T(re_1, (r+s)e_1)^2)^{1/2} \leq c_2 s.$$

Assuming  $\epsilon_0$  is small, by Proposition 3.4 and (3.30) we have

$$P(W \notin S_2) \leq P(W \notin S_1) \leq C_{47} e^{-C_{48} t \log t} \leq C_{47} e^{-c_3 \epsilon^{-b}} \leq \left(\frac{\epsilon \mu}{8c_2}\right)^2. \quad (3.33)$$

Combining these yields

$$E(T(re_1, (r+s)e_1) 1_{\{W \notin S_2\}}) \leq E(T(re_1, (r+s)e_1)^2)^{1/2} P(W \notin S_2)^{1/2} \leq \frac{\epsilon \mu s}{8}. \quad (3.34)$$

To use (3.32) we also need a lower bound for  $E(T^*)$ . For  $y > 0$ , using (3.29),

$$\begin{aligned} P\left(T^* \leq \left(1 - \frac{\epsilon}{2}\right) h(s) - 2y\sigma(m) \log m\right) \\ \leq P\left(\text{for some } w \in S_2 \cap \mathbb{V}, T(w, (r+s)e_1) - T(w, re_1) \leq \right. \\ \left. h(|w - (r+s)e_1|) - h(|w - re_1|) - 2y\sigma(m) \log m\right) \\ \leq P\left(\text{for some } w \in S_2 \cap \mathbb{V}, T(w, (r+s)e_1) - h(|w - (r+s)e_1|) \leq -y\sigma(m) \log m\right) \\ + P\left(\text{for some } w \in S_2 \cap \mathbb{V}, T(w, re_1) - h(|w - re_1|) \geq y\sigma(m) \log m\right). \end{aligned} \quad (3.35)$$

We consider the first probability on the right in (3.35); the second probability is similar. For  $w \in S_2$  we have using (3.27) that

$$m - 2 \leq |w - (r+s)e_1| \leq (1 + \epsilon^2) m,$$

and then from (1.11),

$$\frac{\sigma(m)}{\sigma(|w - (r+s)e_1|)} \geq \frac{C_{22}}{2}.$$

From these and Lemma 3.2 (see  $C_{45}$  there) we obtain that if  $c_0$  is large enough, then for  $y \geq c_0$ ,

$$\begin{aligned} P\left(\text{for some } w \in S_2 \cap \mathbb{V}, T(w, (r+s)e_1) - h(|w - (r+s)e_1|) \leq -y\sigma(m) \log m\right) \\ \leq \sum_{\hat{w} \in S_2 \cap d^{-1/2} \mathbb{Z}^d} P\left(\text{for some } w \in B_1(\hat{w}) \cap \mathbb{V}, \right. \\ \left. T(w, (r+s)e_1) - h(|\hat{w} - (r+s)e_1|) \leq h(|w - \hat{w}|) - y\sigma(m) \log m\right) \\ \leq \sum_{\hat{w} \in S_2 \cap d^{-1/2} \mathbb{Z}^d} P\left(\text{for some } w \in B_1(\hat{w}) \cap \mathbb{V}, \right. \\ \left. T(w, (r+s)e_1) - h(|\hat{w} - (r+s)e_1|) \leq -\frac{y}{2}\sigma(m) \log m\right) \\ \leq c_4 m^{d-1} \exp\left(-\frac{C_{45} C_{22} y}{4} \log m\right) \\ \leq \frac{1}{2} e^{-c_5 y \log m}. \end{aligned} \quad (3.36)$$



The same bound holds for the second probability on the right in (3.35). With (3.35) and the definition of  $m$  in (3.26) this shows that

$$ET^* \geq \left(1 - \frac{\epsilon}{2}\right) h(s) - 4c_0\sigma(m) \log m \geq \left(1 - \frac{3\epsilon}{4}\right) h(s).$$

Combining this with (3.32) and (3.34) yields (3.24).

We now prove (ii). From (i), there exists  $c_6$  such that  $h(r + s) \geq h(r) + (1 - \epsilon)h(s)$  whenever  $r, s \geq c_6$ ; there then exists  $c_7$  such that  $h(s) \leq c_7$  whenever  $0 \leq s \leq c_6$ . We therefore have

$$h(r + s) \geq \begin{cases} h(r) + (1 - \epsilon)h(s) & \text{if } r, s \geq c_6 \\ h(r) - h(s) \geq h(r) + (1 - \epsilon)h(s) - 2c_7 & \text{if } s < c_6 \\ 0 \geq h(r) + (1 - \epsilon)h(s) - 2c_7 & \text{if } r < c_6 \end{cases}$$

which proves (3.25). □

## 4 Proof of Theorem 1.3—downward deviations

### 4.1 Proof sketch for (1.14)

The proof operates by making and comparing coarse-grain (CG) approximations on multiple scales. Given a geodesic  $\Gamma_{xy}$ , which we take here to be approximately in the  $e_1$  direction, we form a collection  $\mathcal{H}_{xy}$  of hyperplanes perpendicular to  $e_1$ , between  $x$  and  $y$ , and we make a CG approximation in each hyperplane near the point where  $\Gamma_{xy}$  crosses it. The nature of  $\mathcal{H}_{xy}$  is illustrated in Figure 3—there are a small number of black hyperplanes there, called “ $j$ th scale hyperplanes” and spaced apart by a multiple of  $\delta^j r$  for some  $j$ , where  $\delta$  (small) is a parameter we choose. Then there are medium-gray hyperplanes sandwiching the black ones; these are spaced apart by a multiple of  $\delta^{j+1} r$ . Finally there are light gray hyperplanes sandwiching the medium-gray ones, spaced apart by a multiple of  $\delta^{j+2} r$ . There are additional hyperplanes in  $\mathcal{H}_{xy}$ , not shown, called joining hyperplanes. There are also in general many more length scales  $j_2 \leq j \leq j_1$  and corresponding sets of hyperplanes in  $\mathcal{H}_{xy}$ , beyond the three scales shown.

To coarse-grain on the  $j$ th scale, we approximate hyperplane-crossing points of  $\Gamma_{xy}$  to within a multiple of  $\beta^j \Delta(r)$  (“ $j$ th-scale approximation”), where  $\beta$  is another small parameter. Initially, the CG is on the  $j_1$ th (finest) scale in all hyperplanes. Our initial goal—which occupies most of the proof—is to create a modified CG approximation  $\Gamma_{xy}^{CG} : u^0 \rightarrow \dots \rightarrow u^n$  using a small subset of  $\mathcal{H}_{xy}$ , containing just two hyperplanes for each scale  $j$ , without overly disrupting the probability (1.14) that we wish to bound. By this notation for  $\Gamma_{xy}^{CG}$  we mean that  $T(\Gamma_{xy}^{CG})$  is defined to be  $\sum_{i=1}^n T(u^{i-1}, u^i)$ . Such a modified approximation  $\Gamma_{xy}^{CG}$  is illustrated by the path with vertices  $u^i$  in Figure 3—as we move from the endpoints  $u^0, u^7$  towards the center of the path, the approximation becomes coarser, being  $(j + 2)$ th-scale in the gray hyperplanes at  $u^1, u^6$ , then  $(j + 1)$ th-scale in the medium-gray hyperplanes at  $u^2, u^5$ , and so on. To obtain the modified CG approximation we iterate a process of coarsening to a  $j$ th-scale approximation and then deleting unwanted  $(j + 1)$ th-scale hyperplanes, one iteration each for  $j = j_1, j_1 - 1, \dots, j_2$ . Each iteration (except the initial  $j = j_1 - 1$ , which has no deletions) has its two stages illustrated in Figure 6. In the first stage of coarsening, called “shifting to the grid,” in each hyperplane not being deleted, we shift the point from a  $j$ th-scale approximation to  $(j - 1)$ th-scale; in Figure 6 this changes the black CG path to the dashed one, in the section of path shown. In the second “point-deletion” stage, we remove the CG approximation points in all hyperplanes to be deleted; in Figure 6 this changes the CG path to the gray one.

Of course we need to analyze how the passage time of the path changes through the various iterations and stages; the very purpose of the iterative process is to have tight control over these changes, which would be not be possible if we jumped directly from original to fully CG path. We want to move from the need to bound

$$P\left(T(x, y) - h(|(y - x)_1|) \leq -t\sigma(r) \text{ for some } x, y \in G_r(K)\right), \tag{4.1}$$

as in (1.14), to just the need to bound something like

$$P\left(T(\Gamma_{xy}^{CG}) - h(|(y - x)_1|) \leq -\frac{1}{2}t\sigma(r) + f(\Gamma_{xy}^{CG}) \text{ for some } (x, y)\right), \tag{4.2}$$

where the  $f(\Gamma_{xy}^{CG})$  (not too large) is there to allow the bound to depend on  $\Gamma_{xy}^{CG}$ . Let us illustrate how the point-removal stage of an iteration moves toward this goal, with a simplified example. For simplicity here we approximate  $h(p, q)$  by  $\mu|q - p|$  for various  $p, q$ . Consider just a single path  $\Gamma_{wz}^1 : w = v^0 \rightarrow v^1 \rightarrow v^2 \rightarrow v^3 = z$ , with  $w_1 < z_1$ ; it should be emphasized that our actual computations are done uniformly over  $(x, y)$  as in (4.1). Suppose that the probability we want to bound is  $P(T(\Gamma_{wz}^1) \leq h(|(z - w)_1|) - t\sigma(r) - a)$  for some  $a$  (so (4.1) has  $a = 0$ , but this changes as iterations occur.) We define an “intermediate path” in which (in some cases) only some of the scheduled points get deleted; suppose this is  $\Gamma_{wz}^{int} : v^0 \rightarrow v^2 \rightarrow v^3$ . Let

$$U_h(\Gamma_{wz}^1) = \sum_{i=1}^3 h(|v^i - v^{i-1}|)$$

and analogously for  $U_h(\Gamma_{wz}^{int})$ . Let  $c, \delta > 0$  and define

$$\tilde{b} = U_h(\Gamma_{wz}^1) - U_h(\Gamma_{wz}^{int}) - 3c, \quad b = \delta (U_h(\Gamma_{wz}^1) - h(|(z - w)_1|)) - 3c,$$

and similarly for  $U_h(\Gamma_{wz}^{int})$ . Then  $(b + 3c)/\delta$  represents the extra  $h$ -length in  $\Gamma_{wz}^1$  versus a single-step path  $w$  to  $z$ . This extra length is related to what we call the bowedness of the path. We always choose the intermediate path so that at least  $\delta$  fraction of this extra length is present when we compare  $\Gamma_{wz}^1$  to  $\Gamma_{wz}^{int}$  instead of to the single-step path, that is, we choose so that  $\tilde{b} \geq b$ ; we call this *deterministic tracking*. We then have

$$\begin{aligned} &P\left(T(\Gamma_{wz}^1) - h(|(z - w)_1|) \leq -t\sigma(r) - a\right) \\ &\leq P\left(T(\Gamma_{wz}^{int}) - h(|(z - w)_1|) \leq -t\sigma(r) - a - \tilde{b}\right) + P\left(T(\Gamma_{wz}^1) - T(\Gamma_{wz}^{int}) \leq \tilde{b}\right) \\ &\leq P\left(T(w, z) - h(|(z - w)_1|) \leq -t\sigma(r) - a - b\right) \\ &\quad + P\left(T(\Gamma_{wz}^1) - T(\Gamma_{wz}^{int}) \leq U_h(\Gamma_{wz}^1) - U_h(\Gamma_{wz}^{int}) - 3c\right) \\ &\leq P\left(T(w, z) - h(|(z - w)_1|) \leq -t\sigma(r) - a - b\right) \\ &\quad + P\left(|T(u^0, u^1) - h(|u^1 - u^0|)| \geq c \text{ or } |T(u^1, u^2) - h(|u^2 - u^1|)| \geq c \right. \\ &\quad \left. \text{or } |T(u^0, u^2) - h(|u^2 - u^0|)| \geq c\right) \\ &= (I) + (II). \end{aligned} \tag{4.3}$$

It should be noted that the use of the intermediate path means that the link  $u^2 \rightarrow u^3$  does not appear in (II), since it is common to both  $\Gamma_{wz}^1$  and  $\Gamma_{wz}^{int}$ . This is useful in a case where  $u^2 \rightarrow u^3$  is much longer than the three links appearing in (II); excluding  $u^2 \rightarrow u^3$  therefore allows for a much smaller value of  $c$  necessary to make (II) small. We refer to  $c$

as an *allocation*; it is like a “tax” subtracted from  $\delta(U_h(\Gamma_{wz}^1) - h(|z - w|))$  in defining  $b$ , and allocated to one of the links  $u^i \rightarrow u^u$ , so that (II) can be small.

We refer to the event on the fourth line of (4.3) as a *tracking failure*; it says that when we remove a point from  $\Gamma_{wz}^1$  to create  $\Gamma_{wz}^{int}$ , the reduction in passage time  $T(\cdot)$  is significantly less than the reduction in  $h$ -length  $U_h(\cdot)$ . Provided we can make the tracking failure probability (II) small while keeping  $c$  small enough,  $b$  will be positive and we have then achieved an improvement from the left side of (4.3) to (I) in the sense that we have replaced  $-t\sigma(r) - a$  with  $-t\sigma(r) - a - b$ . As we go through the iterations, starting from (4.1), these improvements accumulate and we obtain something like (see (4.166))

$$\begin{aligned} P(T(x, y) - h(|(y - x)_1|) \leq -t\sigma(r) \text{ for some } x, y \in G_r(K)) \\ \leq P\left(T(\Gamma_{xy}^{CG}) - h(|(y - x)_1|) \leq -\frac{1}{2}t\sigma(r) - \delta\mu(U_{Euc}(\Gamma_{xy}^{(j_1)}) - U_{Euc}(\Gamma_{xy}^{CG}))\right) \\ + (\text{sum of allocations) for some } x, y) \\ + (\text{sum of probabilities of tracking failures}), \end{aligned} \tag{4.4}$$

where  $\Gamma_{xy}^{(j_1)}$  denotes the initial finely-coarse-grained path. The sum of allocations here contains a term which we split:

$$\delta\mu(U_{Euc}(\Gamma_{xy}^{(j_1)}) - |(y - x)_1|) = \delta\mu(U_{Euc}(\Gamma_{xy}^{(j_1)}) - U_{Euc}(\Gamma_{xy}^{CG})) + \delta\mu(U_{Euc}(\Gamma_{xy}^{CG}) - |(y - x)_1|),$$

with the first term on the right representing the decrease in path length due to iterations and the second representing the remaining excess path length in the CG path, after the iterations are complete. The first of these is cancelled by the similar term on the right in (4.4), which came from the accumulated improvements, and the problem can then be reduced (see (4.173)) to bounding a probability purely involving the final CG approximation:

$$P\left(T(\Gamma_{xy}^{CG}) - U_h(\Gamma_{xy}^{CG}) \leq -\frac{1}{3}t\sigma(r) - \frac{\mu}{2}(U_{Euc}(\Gamma_{xy}^{CG}) - |(y - x)_1|) \text{ for some } x, y \in G_r(K)\right). \tag{4.5}$$

If we compare (4.5) to (4.1), we see that besides replacing the geodesic with its course-grained approximation, we have improved the right side of the inequality to a factor proportional to the excess length in the CG path. This improvement makes it possible to sum where needed over all possible CG paths. A key tool in the bounding of (4.5), as well as the various tracking failure probabilities, is Lemma 4.6, which says that with high probability, there are no CG links  $(u, v)$  anywhere in  $G_r^+$  for which  $T(u, v)$  deviates from its mean by more than a specified allocation, which shrinks as the coarse-graining scale becomes finer.

Most of the preceding is oversimplified from what we actually do; we give further proof-outlining information in Remarks 4.4 and 4.8.

#### 4.2 Step 0. Preliminaries and easy cases

As we have noted, it is sufficient to prove (1.14), thanks to (1.15) and Lemma 3.3. We use a multiscale argument which is related to chaining. First we dispense with simpler cases that only require Lemma 3.2 and Proposition 3.4; these allow us to take  $|y - x|$  at least of order  $r/(\log r)^{1/\chi_1}$ . The first such case is pairs  $x, y$  which are close together. For technical convenience later, we prove the following also for geodesics with endpoints in the set  $G_r^+$ , satisfying  $G_r^+ \supset G_r(K)$  for  $K$  fixed and  $r$  large, given by

$$G_r^+ = [-C_{60}(\log r)^{2/(1-\chi)}, r + C_{60}(\log r)^{2/(1-\chi)}] \times C_{60}\Delta(r)(\log r)\mathfrak{B}_{d-1}, \tag{4.6}$$

with  $C_{60}$  to be specified.

**Lemma 4.1.** *Suppose  $G = (V, E)$  and  $\{\eta_e, e \in E\}$  satisfy A1, A2, and A3. There exist constants  $C_{61}$ – $C_{64}$  such that for all  $r \geq 2, K \geq 1$ , and  $t > C_{61}K^2$ ,*

$$P\left(T(x, y) \leq ET(x, y) - t\sigma(r) \text{ for some } x, y \in G_r(K) \cup G_r^+ \text{ with } |y - x| \leq \frac{C_{62}r}{(\log r)^{1/\chi_1}}\right) \leq C_{63}e^{-C_{64}t \log r}. \tag{4.7}$$

*Proof.* We may assume  $t$  is large. We first discretize: by Lemma 3.1(ii), for  $x \in B_1(\hat{x}), y \in B_1(\hat{y})$ , we have  $|ET(x, y) - ET(\hat{x}, \hat{y})| \leq 2C_{42} \leq t\sigma(r)/2$ , so for  $c_0 > 0$ ,

$$P\left(T(x, y) \leq ET(x, y) - t\sigma(r) \text{ for some } x, y \in G_r^+ \text{ with } |y - x| \leq \frac{c_0r}{(\log r)^{1/\chi_1}}\right) \leq \sum_{\substack{\hat{x}, \hat{y} \in d^{-1/2}\mathbb{Z}^d \cap (G_r(K) \cup G_r^+) \\ |\hat{x} - \hat{y}| \leq 2c_0r/(\log r)^{1/\chi_1}}} P\left(T(x, y) \leq ET(\hat{x}, \hat{y}) - \frac{t\sigma(r)}{2} \text{ for some } x \in B_1(\hat{x}), y \in B_1(\hat{y})\right). \tag{4.8}$$

From (1.11),

$$|\hat{x} - \hat{y}| \leq \frac{2c_0r}{(\log r)^{1/\chi_1}} \implies \frac{\sigma(r)}{\sigma(|\hat{x} - \hat{y}|)} \geq \frac{c_1}{(2c_0)^{1/\chi_1}} \log r,$$

so if we take  $c_0$  small enough then by Lemma 3.2,

$$\begin{aligned} &\sum_{\substack{\hat{x}, \hat{y} \in d^{-1/2}\mathbb{Z}^d \cap (G_r(K) \cup G_r^+) \\ |\hat{x} - \hat{y}| \leq c_0r/2(\log r)^{1/\chi_1}}} P\left(T(x, y) \leq ET(\hat{x}, \hat{y}) - \frac{t\sigma(r)}{2} \text{ for some } x \in B_1(\hat{x}), y \in B_1(\hat{y})\right) \\ &\leq c_2r(K\Delta(r) \log r)^{d-1} \max\left\{e^{-C_{45}t\sigma(r)/2\sigma(|\hat{x} - \hat{y}|)} : |\hat{x} - \hat{y}| \leq c_0r/2(\log r)^{1/\chi_1}\right\} \\ &\leq e^{-C_{45}t \log r}, \end{aligned} \tag{4.9}$$

which with (4.8) proves (4.7). □

Lemma 4.1 means we need only consider  $x, y \in G_r(K)$  satisfying

$$|y - x| > \frac{C_{62}r}{(\log r)^{1/\chi_1}}. \tag{4.10}$$

Writing  $\alpha_{uv}$  for the angle between nonzero vectors  $u, v \in \mathbb{R}^d$ , this means that for large  $r$ ,

$$\alpha_{y-x, e_1} \leq \beta_r := c_0 \frac{\Delta(r)(\log r)^{1/\chi_1}}{r} \leq r^{-(1-\chi)/4}. \tag{4.11}$$

A second simple case is small  $r$ . For fixed  $r_0$  and  $1 \leq r \leq r_0$ , from Lemma 3.1 for all  $x, y \in G_r$  we have  $ET(x, y) \leq c_1r$ , so for  $t \geq c_2r_0$  we have  $ET(x, y) - t\sigma(r) \leq c_1r - t\sigma(r) < 0$ , and hence the probability in (1.14) is 0. Therefore there exist  $C_{29}, C_{30}$  such that (1.14) is valid for all  $1 \leq r \leq r_0$  and  $t > 0$ .

A third simple case is  $t \geq c_3 \log(Kr)$ , with  $c_3$  large enough. As in (4.8) and (4.9), for  $C_{45}$  from Lemma 3.2 we then have

$$\begin{aligned} &P\left(T(x, y) \leq ET(x, y) - t\sigma(r) \text{ for some } x, y \in G_r(K)\right) \\ &\leq c_4r(K\Delta(r))^{d-1} \max_{u, v \in G_r(K)} e^{-C_{45}t\sigma(r)/2\sigma(|u-v|)} \\ &\leq c_4r(K\Delta(r))^{d-1} e^{-c_5t} \\ &\leq c_6e^{-c_5t/2}. \end{aligned} \tag{4.12}$$

It follows that, for  $C_{26}$  from Theorem 1.3, we need only consider  $C_{26}K^2 \leq t < c_3 \log(Kr)$ , and therefore also  $K \leq c_7(\log r)^{1/2}$ , which means  $G_r(2K) \subset G_r^+$ .

A fourth simple case is pairs  $x, y$  for which  $\Gamma_{xy}$  goes well outside  $G_r(K)$ , when  $r$  is large and  $t < c_3 \log(Kr)$ . For that we have the following.

**Lemma 4.2.** *Suppose  $G = (V, E)$  and  $\{\eta_e, e \in E\}$  satisfy A1, A2, and A3. Then given  $C > 0$  there exist constants  $C_i = C_i(C)$  such that for  $c_7$  as above, for all  $r \geq C_{65}, K \leq c_7(\log r)^{1/2}$  we have*

$$P(\Gamma_{xy} \not\subset G_r^+ \text{ for some } x, y \in G_r(K) \text{ satisfying (4.10)}) \leq C_{68}e^{-C_{69}(\log \log r) \log r}. \quad (4.13)$$

For cases of Theorem 1.3 not covered by the first three simple cases, since  $r \geq r_0$  and (as noted above)  $C_{26}K^2 \leq t < c_3 \log Kr$ , provided we take  $r_0$  and  $C_{26}$  large enough, the exponent in (4.13) is greater than  $t$ , so (4.13) is in keeping with (1.14).

*Proof of Lemma 4.2.* Note the assumptions guarantee  $G_r(K) \subset G_r^+$ . Let  $\hat{x}, \hat{y} \in d^{-1/2}\mathbb{Z}^d \cap G_r(K)$  with  $\hat{x}_1 < \hat{y}_1$ , and let  $x, y \in G_r$  satisfy (4.10), with  $|x - \hat{x}| < 1, |y - \hat{y}| < 1$ . Recall the transformation  $\Theta_{\hat{x}\hat{y}}$  which takes  $\Pi_{\hat{x}\hat{y}}$  to  $\Pi_{0,|\hat{y}-\hat{x}|e_1}$ . Suppose  $w \in \Gamma_{\hat{x}\hat{y}} \setminus G_r^+$  and let  $\tilde{w} = \Theta_{\hat{x}\hat{y}}w$ ; we may take  $w$  with  $d(w, G_r^+) \leq 2$ . We claim that, for  $D_r$  from (3.20), we have  $D_r(\tilde{w}) \geq C_{60} \log r$ , with  $C_{60}$  from the definition of  $G_r^+$ . We consider several cases.

**Case 1.**  $-C_{60}(\log r)^{2/(1-\chi)} \leq w_1 \leq r + C_{60}(\log r)^{2/(1-\chi)}$ , that is,  $w$  is to the side of  $G_r^+$ . See Figure 2. From symmetry we may assume  $-C_{60}(\log r)^{2/(1-\chi)} \leq w_1 \leq r/2$ . Note that in the definition (3.17) of  $D(u)$ , the case  $u_1 \geq 0$  is the relevant one here if and only if  $\tilde{w}_1 \in [0, |\hat{y} - \hat{x}|]$ . We have

$$d(\tilde{w}, \Pi_{0,|\hat{y}-\hat{x}|e_1}) = d(w, \Pi_{\hat{x}\hat{y}}) \geq d(w, G_r(K)) \geq \frac{C_{60}}{2}\Delta(r) \log r, \quad |w^*| \geq C_{60}\Delta(r) \log r, \quad (4.14)$$

and  $\tilde{w}_1 \in [0, |\hat{y} - \hat{x}|] \implies d(\tilde{w}, \Pi_{0,|\hat{y}-\hat{x}|e_1}) = |\tilde{w}^*|$ , so

$$\tilde{w}_1 \in \left[0, \frac{|\hat{y} - \hat{x}|}{2}\right] \implies \frac{|\tilde{w}^*|^2}{\Xi(\tilde{w}_1)^2} \geq \frac{d(\tilde{w}, \Pi_{0,|\hat{y}-\hat{x}|e_1})^2}{\Xi\left(\frac{1}{2}|\hat{y} - \hat{x}|\right)^2} \geq \frac{C_{60}^2\Delta(r)^2(\log r)^2}{4\Xi(r)^2} \geq \frac{C_{60}^2}{5} \log r.$$

Also as in (4.14),

$$\max(|\tilde{w}_1|, |\tilde{w}^*|) \geq \frac{1}{\sqrt{2}}|\tilde{w}| = \frac{1}{\sqrt{2}}|w - \hat{x}| \geq \frac{1}{\sqrt{2}}d(w, \Pi_{\hat{x}\hat{y}}) \geq \frac{C_{60}}{3}\Delta(r) \log r$$

so

$$\Phi(\max(|\tilde{w}_1|, |\tilde{w}^*|)) \geq C_{60} \log r.$$

Therefore under Case 1,

$$-\infty < \tilde{w}_1 \leq \frac{|\hat{y} - \hat{x}|}{2} \implies D_{|\hat{y}-\hat{x}|}(\tilde{w}) \geq C_{60} \log r, \quad (4.15)$$

and symmetrically the same holds for  $\tilde{w}_1 > |\hat{y} - \hat{x}|/2$ .

**Case 2.**  $w_1 < -C_{60}(\log r)^{2/(1-\chi)}$  with  $\tilde{w}_1 > 0$ , or  $w_1 > r + C_{60}(\log r)^{2/(1-\chi)}$  with  $\tilde{w}_1 < |\hat{y} - \hat{x}|$ , that is,  $w$  is past the end of  $G_r^+$  but  $\tilde{w}$  is to the side of  $\Pi_{0,|\hat{y}-\hat{x}|e_1}$ . See Figure 2. The two ends are symmetric so we need only consider  $w_1 < -C_{60}(\log r)^{2/(1-\chi)}$  with  $\tilde{w}_1 > 0$ . Let  $H_{(\hat{x})}$  be the hyperplane through  $\hat{x}$  perpendicular to  $\hat{y} - \hat{x}$ , and let  $\bar{w}$  be the orthogonal projection of  $w$  into  $H_{(\hat{x})}$ , so  $|\bar{w} - \hat{x}| = |\tilde{w}^*|$  and  $|w - \bar{w}| = |\tilde{w}_1|$ . Since  $K \leq c_7(\log r)^{1/2}$ , by (4.11) the angle  $\alpha_{\hat{y}-\hat{x},e_1}$  is small enough that we have  $\bar{w}_1 < w_1 < 0$  and hence, for  $\beta_r$  from (4.11),

$$C_{60}(\log r)^{2/(1-\chi)} \leq \hat{x}_1 - \bar{w}_1 \leq |\bar{w} - \hat{x}| \sin \alpha_{\hat{y}-\hat{x},e_1} \leq |\tilde{w}^*| \beta_r; \quad (4.16)$$

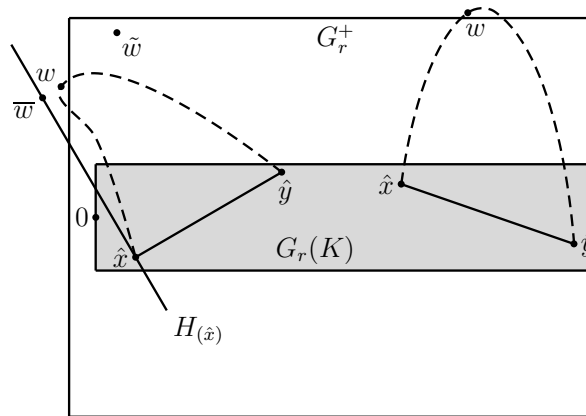


Figure 2: Diagram for Case 1 (right) and Case 2 (left) in the proof of Lemma 4.2. On the left, if we translate and rotate the picture so that  $\hat{x}$  becomes 0 and  $\hat{y} - \hat{x}$  becomes horizontal, then  $w$  becomes  $\tilde{w}$ . Case 3 is like Case 2 except that  $\tilde{w}$  is to the left of 0.

therefore

$$\Phi(\max(|\tilde{w}_1|, |\tilde{w}^*|)) \geq \Phi\left(\frac{C_{60}(\log r)^{2/(1-\chi)}}{\beta_r}\right) \geq C_{60} \log r.$$

Further, similarly to (4.16), since the small angle  $\alpha_{\hat{y}-\hat{x}, e_1}$  ensures  $\frac{1}{2}\tilde{w}_1 = \frac{1}{2}|w - \bar{w}| \leq (w - \bar{w})_1$ , we have

$$\frac{1}{2}\tilde{w}_1 \leq (w - \bar{w})_1 \leq (\hat{x} - \bar{w})_1 = |\bar{w} - \hat{x}| \sin \alpha_{\hat{y}-\hat{x}, e_1} \leq |\tilde{w}^*| \beta_r \tag{4.17}$$

and therefore using (1.11),

$$\frac{|\tilde{w}^*|^2}{\Xi(\tilde{w}_1)^2} \geq \frac{|\tilde{w}^*|^2}{\Xi(2\beta_r|\tilde{w}^*|)^2} \geq \frac{C_{22}|\tilde{w}^*|}{\beta_r^{\chi_1} \sigma(|\tilde{w}^*|) \log(2 + |\tilde{w}^*|)} \geq \frac{1}{\beta_r^{\chi_1}} \geq C_{60} \log r.$$

This proves the right side of (4.15) under Case 2.

Case 3.  $w_1 < -C_{60}(\log r)^{2/(1-\chi)}$  with  $\tilde{w}_1 \leq 0$ , or  $w_1 > r + C_{60}(\log r)^{2/(1-\chi)}$  with  $\tilde{w}_1 \geq |\hat{y} - \hat{x}|$ , that is,  $w$  is past the end of  $G_r^+$  and  $\tilde{w}$  is past the end of  $\Pi_{0, |\hat{y}-\hat{x}| e_1}$ . The two ends are again symmetric so we need only consider  $w_1 < -C_{60}(\log r)^{2/(1-\chi)}$  with  $\tilde{w}_1 \leq 0$ . Then  $|\tilde{w}| = |w - \hat{x}| \geq \hat{x}_1 - w_1 \geq C_{60}(\log r)^{2/(1-\chi)}$  so

$$D_r(\tilde{w}) = \Phi(\max(|\tilde{w}_1|, |\tilde{w}^*|)) \geq \Phi\left(\frac{|\tilde{w}|}{2}\right) \geq C_{60} \log r.$$

Thus in all cases we have  $D_r(\tilde{w}) \geq C_{60} \log r$ , so by Proposition 3.4,

$$\begin{aligned} & P\left(\Gamma_{xy} \not\subset G_r^+ \text{ for some } x, y \in G_r(K) \text{ satisfying (4.10)}\right) \\ & \leq \sum_{\hat{x}, \hat{y} \in d^{-1/2}\mathbb{Z}^d \cap G_r(K)} P\left(\Gamma_{xy} \not\subset G_r^+ \text{ for some } x, y \in G_r(K) \text{ satisfying (4.10)}\right) \\ & \qquad \qquad \qquad \text{with } |x - \hat{x}| < 1, |y - \hat{y}| < 1 \\ & \leq \sum_{\hat{x}, \hat{y} \in d^{-1/2}\mathbb{Z}^d \cap G_r(K)} P\left(\sup_{x \in B_1(\hat{x}) \cap \mathbb{V}, y \in B_1(\hat{y}) \cap \mathbb{V}} \sup_{w \in \Gamma_{xy}} D_{|\hat{y}-\hat{x}|}(\Theta_{\hat{x}\hat{y}} w) \geq C_{60} \log r\right) \\ & \leq c_0 r(K\Delta(r))^{d-1} e^{-C_{47}(\log \log r) \log r} \\ & \leq c_1 e^{-c_2(\log \log r) \log r}. \end{aligned}$$

□

In summary, having handled the four simple cases we may restrict to the following situation (still with  $x, y \in G_r(K)$ ):

$$r \geq r_0, \quad K \leq c_7(\log r)^{1/2}, \quad c_8 \leq t \leq c_3 \log r, \quad |y - x| > \frac{C_{62}r}{(\log r)^{1/\chi_1}}, \quad \Gamma_{xy} \subset G_r^+, \quad (4.18)$$

with  $C_{62}$  from Lemma 4.1.

### 4.3 Step 1. Setting up the coarse-graining

For purposes of coarse-graining and multiscale analysis of paths, we build grids inside  $G_r^+$  on various scales, using small parameters  $\lambda, \delta, \beta$  satisfying

$$1 \gg \lambda \gg \delta^{\chi_1} \gg \delta^{(1+\chi_1)/2} \gg \beta \quad (4.19)$$

in the sense that the ratio of each term to the one following must be taken sufficiently large, in a manner to be specified. We choose these so  $1/\delta$  and  $1/\beta$  are integers. There is also a fourth parameter  $\rho > 1$ , and we further require

$$\frac{\beta}{\lambda\rho\delta^{(1+\chi_2)/2}} \ll 1, \quad \frac{\rho\beta}{\lambda\delta^{(1-\chi_1)/2}} \ll 1, \quad \frac{\beta^2}{\lambda^2\delta} \ll 1, \quad \frac{\beta^{2\chi_1/(1+\chi_1)}}{\lambda\delta^3} \ll 1, \quad \rho^2\lambda \gg 1; \quad (4.20)$$

all of (4.20) can be satisfied by taking  $\beta$  small enough after choosing  $\rho, \lambda, \delta$ . For  $j \geq 1$ , a  $j$ th-scale hyperplane is one of form  $H_{k\delta^j r}, k \in \mathbb{Z}$ , and the  $j$ th-scale grid in  $G_r^+$  is

$$\mathbb{L}_j = \mathbb{L}_j(r) = \{u \in G_r^+ : u_1 \in \delta^j r \mathbb{Z}, u^* \in K_0 \beta^j \Delta(r) \mathbb{Z}^{d-1}\},$$

where  $K_0 \in [1, 2]$  is to be specified. Note that larger  $j$  values correspond to smaller scales, and since  $1/\delta$  is an integer, a  $j$ th-scale hyperplane for some  $j$  is also a  $k$ th-scale hyperplane for  $k > j$ . A hyperplane is *maximally  $j$ th-scale* if it is  $j$ th-scale but not  $(j - 1)$ th-scale.

The  $j$ th-scale grid divides a  $j$ th-scale hyperplane into cubes which we call  *$j$ th-scale blocks*. For concreteness we take these blocks to be products of left-open-right-closed intervals. Each point  $u$  of the hyperplane then lies in a unique such block. For a point  $u$  in a  $j$ th-scale hyperplane, the  *$j$ th-scale coarse-grain approximation* of  $u$  is the point  $V_j(u)$  which is the unique corner point in the block containing  $u$ . We abbreviate coarse-grain as CG. The definition ensures that two points with the same  $j$ th-scale CG approximation also have the same  $k$ th-scale CG approximation for all larger scales  $k < j$ .

A *transverse step* in the  $j$ th-scale grid is a step from some  $u \in \mathbb{L}_j$  to some  $v \in \mathbb{L}_j$  satisfying  $v_1 = u_1, |v^* - u^*| = K_0 \beta^j \Delta(r)$ ; a *longitudinal step* is from  $u$  to  $v$  satisfying  $v_1 = u_1 + \delta^j r, v^* = u^*$ . Define

$$\ell_1 = \ell_1(j) = \frac{\Delta(\delta^j r)}{\beta^j \Delta(r)}, \quad \ell_2 = \ell_2(j) = \rho^j \ell_1(j),$$

so from (1.11),

$$\ell_1 \geq \frac{\delta^{(1+\chi_2)j/2}}{C_{23}\beta^j}, \quad \frac{\delta^{(1+\chi_2)j/2}\rho^j}{C_{23}\beta^j} \leq \ell_2 \leq \frac{\delta^{(1+\chi_1)j/2}\rho^j}{C_{22}\beta^j}. \quad (4.21)$$

Here  $\ell_1(j)$  is chosen so that the typical transverse fluctuation  $\Delta(\delta^j r)$  for a geodesic making one longitudinal step is of order  $\ell_1(j)$  transverse steps.

On short enough length scales, coarse-graining is unnecessary because we can use Proposition 3.5 and Lemma 4.1. More precisely, we will need only consider  $j \leq \hat{j}_1 = j_1(r)$  where  $j_1$  is the least  $j$  for which

$$\left(\frac{\lambda}{\delta^{\chi_1}}\right)^{j_1} \geq (\log r)^2, \quad \text{so } j_1(r) \asymp \log \log r. \quad (4.22)$$



Provided  $r$  is large, this means the spacings  $\delta^{j_1}r$  and  $K_0\beta^{j_1}\Delta(r)$  of the  $j_1$ th-scale grid are large, and therefore we can choose  $q \in [4, 5]$  so that  $\delta^{j_1}r$  is an integer multiple of  $q$ , and then choose  $K_0 \in [1, 2]$  such that  $K_0\beta^{j_1}\Delta(r)$  is an integer multiple of  $q$ . Then (since  $1/\delta$  and  $1/\beta$  are integers) for all  $j \leq j_1$ , the  $j$ th-scale grid in every  $j$ th-scale hyperplane is contained in  $q\mathbb{Z}^d$ , which we call the *basic grid*.

We say an interval in  $\mathbb{R}$  has  $k$ th-scale length if its length is between  $10\delta^{k+1}r$  and  $10\delta^k r$ .

Given  $x, y \in G_r(K)$  with  $x_1 < y_1$ , satisfying (4.18), we define a *hyperplane collection*  $\mathcal{H}_{xy}$ , which depends on the geodesic  $\Gamma_{xy}$ , constructed inductively as follows. All hyperplanes  $H_s \in \mathcal{H}_{xy}$  have  $s \in [x_1, y_1]$ , and we view the hyperplanes as ordered by their indices. Let  $j_2(x, y)$  be the least  $j$  such that there are at least 4  $j$ th-scale hyperplanes between  $x$  and  $y$ . Subject always to the constraint  $s \in [x_1, y_1]$ , at scale  $j_2$  we put in  $\mathcal{H}_{xy}$  the  $j_2$ th scale hyperplanes  $H_s$  second closest to  $x$  and to  $y$ ; we call these  $j_2$ -terminal hyperplanes. The gap between the  $j_2$ -terminal hyperplanes is at least  $\delta^{j_2}r$  and at most  $4\delta^{j_2-1}r$ . In general, when we have chosen the  $j$ th-scale hyperplanes in  $\mathcal{H}_{xy}$  for some  $j \geq j_2$ , each gap between consecutive ones is called a  $j$ th-scale interval, and the hyperplanes bounding it are the *endpoint hyperplanes* of the interval. The number of  $j$ th-scale intervals in  $\mathcal{H}_{xy}$  is at most  $|\hat{x} - \hat{y}|/\delta^j \leq \delta^{-j}$ . For any interval  $I$  we also call  $\{H_s : s \in I\}$  an interval; which meaning should be clear from the context. A  $j$ th-scale interval is *short* if it has  $j$ th-scale length, and *long* otherwise. We then add  $(j + 1)$ th-scale hyperplanes to  $\mathcal{H}_{xy}$ , of 3 types.

- (i) The first type consists of two hyperplanes, which are the  $(j + 1)$ th-scale hyperplanes second closest to  $x$  and to  $y$ , which we call  $(j + 1)$ -terminal hyperplanes. A  $(j + 1)$ -terminal hyperplane may also be a  $k$ th-scale hyperplane on some larger scale  $k < j + 1$ , in which case we call it an *incidental  $k$ th-scale hyperplane*.
- (ii) As a second type, for each non-incidental  $j$ th-scale hyperplane  $H_{k\delta^j r} \in \mathcal{H}_{xy}$  we put in  $\mathcal{H}_{xy}$  the closest  $(j + 1)$ th-scale hyperplanes on either side of  $H_{k\delta^j r}$ , that is,  $H_{(k\delta^j - \delta^{j+1})r}$  and  $H_{(k\delta^j + \delta^{j+1})r}$ , which we call *sandwiching hyperplanes*.
- (iii) The third type is  $(j + 1)$ th-scale *joining hyperplanes*; we place between 1 and 4 of these in each long  $j$ th-scale interval, depending on the behavior of  $\Gamma_{xy}$  in the interval in a manner to be specified below. Joining hyperplanes are always placed in the “extremal 10ths” of the long interval; more precisely, if the  $j$ th-scale interval has  $k$ th-scale length then they are placed at distance  $\delta^\ell r$  from one of the endpoints for some  $k + 1 \leq \ell \leq j$ , with at most 2 such hyperplanes at either end.

We use superscripts  $-$  and  $+$  for quantities associated with left-end and right-end joining hyperplanes, respectively. In (iii), if we add 2 joining hyperplanes at one end of the interval, then their distances from the endpoint have form  $\delta^\ell r$  and  $\delta^{\ell+1}r$  for some  $\ell$ ; this  $\ell$  corresponds to the scale where bowedness is maximal, in a sense we will describe below.

We continue adding hyperplanes through all scales from  $j_2$  to  $j_1$ ; after adding  $j_1$ th-scale hyperplanes,  $\mathcal{H}_{xy}$  is complete and we stop.

A *terminal  $j$ th-scale interval* in  $[x_1, y_1]$  is an interval between a terminal  $(j + 1)$ th-scale hyperplane and the terminal  $j$ th-scale hyperplane closest to it; the length of such an interval is necessarily between  $\delta^j r/2$  and  $2\delta^j r$ , so it is short. In Figure 3, the interval between the hyperplanes containing  $u^2$  and  $u^3$  is a terminal  $j$ th-scale interval.

At most 6  $(j + 1)$ th-scale hyperplanes are added inside each  $j$ th-scale interval, and only 1 if the interval is terminal. Therefore if  $\mathcal{H}_{xy}$  contains  $n$   $j$ th-scale hyperplanes, then the number of  $(j + 1)$ th-scale hyperplanes is at most  $7(n - 1) + 5$ . It follows that

$$\left| \{H_s : H_s \text{ is a } j\text{th-scale hyperplane in } \mathcal{H}_{xy}\} \right| \leq 7^j - 1. \tag{4.23}$$

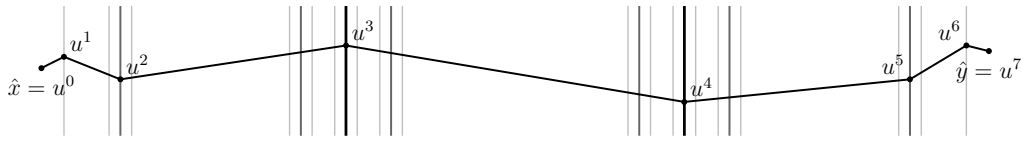


Figure 3: Diagram showing 3 scales of hyperplanes in  $\mathcal{H}_{xy}$ :  $j$ th-scale (black),  $(j + 1)$ th-scale (medium gray),  $(j + 2)$ th-scale (light gray.) Points  $u^i, 1 \leq i \leq 6$ , are in the terminal hyperplanes; if these are all the scales (i.e.  $j_1 = j_2 + 2$ ) then the path is an example of a possible final CG path  $\Gamma_{xy}^{CG}$ , or a path  $\Omega_{xy}$  in Section 7. The other hyperplanes are sandwiching ones. Joining hyperplanes are not shown.

In keeping with (iii) above, we will designate up to four random values  $\mu_{xy}^{-,1}(I) < \mu_{xy}^{-,2}(I) < \mu_{xy}^{+,2}(I) < \mu_{xy}^{+,1}(I)$  in  $I$  for each long  $j$ th-scale interval  $I = [a, b]$ , for each  $j$ , depending on  $\Gamma_{xy}$ . These will satisfy

$$|I| \in [\delta^k r - 4\delta^k r, \delta^k r] \implies \mu_{xy}^{-,1}(I) = a + \delta^{\ell+1} r, \mu_{xy}^{-,2}(I) = a + \delta^\ell r \text{ for some } k < \ell \leq j, \\ \mu_{xy}^{+,1}(I) = b - \delta^{\ell'+1} r, \mu_{xy}^{+,2}(I) = b - \delta^{\ell'} r \text{ for some } k < \ell' \leq j.$$

We call the values  $\mu_{xy}^{\pm,1}(I)$  *outer joining points*, and  $\mu_{xy}^{\pm,2}(I)$  *inner joining points*; the inner ones will represent locations where a certain other path traversing  $I$  can be guided to coalesce with  $\Gamma_{xy}$ , and we call  $H_{\mu_{xy}^{\pm,\epsilon}(I)}$  the (potential) *joining hyperplanes of the interval*  $I$ . We say “potential” because not all are necessarily actually included in  $\mathcal{H}_{xy}$ ; which are included, and with what values of  $\ell, \ell'$ , depend on rules to be described.

Recall that  $\psi_q(u)$  denotes the closest point to  $u$  in the basic grid, and  $F_y = \psi_q^{-1}(y), y \in q\mathbb{Z}^d$ . We need only consider the case in which  $x, y$  each share a Voronoi cell with a basic grid point, that is,

$$x = \varphi(\hat{x}), \quad y = \varphi(\hat{y}) \quad \text{for some } \hat{x}, \hat{y} \in q\mathbb{Z}^d; \tag{4.24}$$

we readily obtain the general case from this via Lemma 3.1, since for every  $x_0 \in \mathbb{V}$  the point  $x = \varphi(\psi_q(x_0))$  satisfies (4.24) and  $|x_0 - x| \leq q\sqrt{d}$ . Since  $q \geq 2$ , (4.24) ensures that  $\psi_q(x) = \hat{x}, \psi_q(y) = \hat{y}$ . For  $0 \leq s \leq r$  and  $\gamma$  a path in  $\mathbb{G}$  from  $B_1(0)$  to  $B_1(re_1)$ , for  $v, w$  vertices in  $\gamma$ , write  $\gamma_{[v,w]}$  for the segment of  $\gamma$  from  $v$  to  $w$  and let  $u_s(\gamma)$  denote the *entry point* of  $\gamma$  into  $H_s^+$ , that is, the first vertex of  $\gamma$  in  $H_s^+$ , necessarily next to  $H_s$  (in the sense that its Voronoi cell intersects  $H_s$ .) Our aim is to approximate a general geodesic  $\Gamma_{xy}$  (subject to (4.18)) by a CG one via certain *marked (basic) grid points* which we will designate, lying in hyperplanes  $H_s \in \mathcal{H}_{xy}$ . In the geodesic  $\Gamma_{xy}$ , the first and last marked grid points are  $\hat{x} = \psi_q(x)$  and  $\hat{y} = \psi_q(y)$  (see (4.24)). Initially, the ones in between are the discrete approximations  $\psi_q(u_s(\gamma))$  of the entry points  $u_s(\gamma)$  corresponding to each of the hyperplanes  $H_s \in \mathcal{H}_{xy}$ . Here since  $q \geq 4$  and  $d(u_s(\gamma), H_s) \leq 2$ , we always have  $\psi_q(u_s(\gamma)) \in H_s$ . Later we will remove some of these initial marked grid points, and replace others with their  $j$ th-scale CG approximations for various  $j$ . This means not every  $H_s \in \mathcal{H}_{xy}$  necessarily contains a marked grid point, in all our CG approximations. When a path  $\Gamma$  has a marked grid point in some  $H_s$ , we denote that marked grid point as  $\mathbf{m}_s(\Gamma)$ . For all our CG approximations, the first and last marked grid points are  $\psi_q(x)$  and  $\psi_q(y)$ , and for some  $j$  the ones in between each lie in the  $j$ th-scale grid in some  $j$ th-scale hyperplane  $H_{k\delta^j r} \in \mathcal{H}_{xy}$ . If such a CG path has marked grid points  $v, w$  in consecutive  $j$ th-scale hyperplanes of  $\mathcal{H}_{xy}$ , we say  $\gamma$  makes a  *$j$ th-scale transition* from  $v$  to  $w$ . Such a transition involves one longitudinal step and some number  $m_i$  of transverse steps in direction  $i$  for each  $2 \leq i \leq d$ , going from  $v$  to  $w$ . We say a  $j$ th-scale transition is *normal* if  $m_i \in [-\ell_2, \ell_2]$  for all  $2 \leq i \leq d$ , and *sidestepping* otherwise. Even on the smallest scale  $j_1$ , every  $j$ th-scale transition with  $v, w \in G_r^+$  is nearly in the  $e_1$  direction, in that its angle

satisfies (similarly to (4.11))

$$\alpha_{e_1, w-v} \leq \frac{4C_{60}\sqrt{d-1}\Delta(r)\log r}{\delta^{j_1}r} \leq r^{-(1-\chi_2)/2}, \tag{4.25}$$

provided  $r$  is large, with  $C_{60}$  from the definition of  $G_r^+$ .

We refer to a path  $\gamma$  from  $x$  to  $y$ , together with its marked grid points, as a *marked path*; the path alone, without the marked grid points, is called the *underlying path*. We may write a marked path as  $\psi_q(x) = v^0 \rightarrow v^1 \rightarrow \dots \rightarrow v^m \rightarrow v^{m+1} = \psi_q(y)$ ; here the  $v^i$  are grid points. The pairs  $(v^{i-1}, v^i)$  are called *links* of the path. If  $\gamma$  is the concatenation of the geodesics  $\Gamma_{v^{i-1}, v^i}$ , we call it a *marked piecewise-geodesic path*; we abbreviate piecewise-geodesic as **PG**. (Recall that when  $v, w \notin \mathbb{G}$ ,  $\Gamma_{vw}$  denotes  $\Gamma_{\varphi(v), \varphi(w)}$ .) Unless otherwise specified, when we give a marked path by writing its marked points in this way, we assume the path is the (unique) marked PG path given by those marked points.

Recall that

$$\hat{T}(u, v) = \min\{T(y, z) : y \in F_u, z \in F_v\},$$

with  $F_u, F_v$  being cubes of side  $q \in [4, 5]$ . Suppose we have a marked PG path

$$\Gamma^{CG} : v^0 \rightarrow v^1 \rightarrow \dots \rightarrow v^m \rightarrow v^{m+1},$$

contained in  $G_r^+$ . We associate four quantities to this path:

$$\Upsilon_{Euc}(\Gamma^{CG}) = \Upsilon_{Euc}(v^0, \dots, v^{m+1}) = \sum_{i=1}^{m+1} |v^i - v^{i-1}|, \quad \Upsilon_h(\Gamma^{CG}) = \sum_{i=1}^{m+1} h(|v^i - v^{i-1}|),$$

$$\Upsilon_{\hat{T}}(\Gamma^{CG}) = \sum_{i=1}^{m+1} \hat{T}(v^{i-1}, v^i), \quad \Psi(\Gamma^{CG}) = \sum_{i=1}^{m+1} \frac{|(v^i - v^{i-1})^*|^2}{|(v^i - v^{i-1})_1|}.$$

Using the standard fact that for some  $c_9 \leq 1$ ,

$$\frac{3}{8} \frac{|w^*|^2}{|w_1|} \leq |w| - |w_1| \leq \frac{|w^*|^2}{2|w_1|} \quad \text{whenever} \quad \frac{|w^*|}{|w_1|} \leq c_9, \tag{4.26}$$

we see that provided  $r$  is large and all  $v^i \in G_r^+$ ,  $\Psi(\Gamma^{CG})$  represents added length in  $\Gamma^{CG}$  relative to the lower bound  $|(v^{m+1} - v^0)_1|$ , in that

$$\Upsilon_{Euc}(\Gamma^{CG}) \geq \sum_{i=1}^{m+1} |(v^i - v^{i-1})_1| + \frac{3}{8}\Psi(\Gamma^{CG}) = |(v^{m+1} - v^0)_1| + \frac{3}{8}\Psi(\Gamma^{CG}). \tag{4.27}$$

Informally we refer to  $\Upsilon_{Euc}(v^0, \dots, v^{m+1}) - |(v^{m+1} - v^0)_1|$  as the *extra length* of the path  $\Gamma^{CG}$ . From (4.26), subadditivity of  $h$ , and Lemma 3.6(ii) (applied with any  $\epsilon < \frac{3}{8} - \frac{1}{3}$ ) we obtain, after reducing  $c_9$  if necessary,

$$\frac{\mu}{3} \frac{|w^*|^2}{|w_1|} - c_{10} \leq h(|w|) - h(|w_1|) \leq \frac{2\mu}{3} \frac{|w^*|^2}{|w_1|} + c_{10} \quad \text{whenever} \quad \frac{|w^*|}{|w_1|} \leq c_9. \tag{4.28}$$

In our applications of (4.28), the last condition will always be satisfied due to (4.25). In general, provided  $r$  is large and all  $v^i$  lie in  $G_r^+$ , with  $(v^i - v^{i-1})_1$  much larger than the width  $2C_{60}\Delta(r)\log r$  of  $G_r^+$ , we have from (4.28) that

$$\begin{aligned} \Upsilon_h(\Gamma^{CG}) - h((v^{m+1} - v^0)_1) &\geq \sum_{i=1}^{m+1} \left[ h(|v^i - v^{i-1}|) - h((v^i - v^{i-1})_1) \right] \\ &\geq \frac{\mu}{3}\Psi(\Gamma^{CG}) - (m+1)c_{10}. \end{aligned} \tag{4.29}$$

We can now define the joining points  $\mu_{xy}^{\pm, \epsilon}(I)$  in a long  $j$ th-scale interval  $I = [a, b]$  of  $k$ th-scale length, where  $j_2 \leq k < j \leq j_1 - 1$ . We begin with  $\mu_{xy}^{-, \epsilon}(I)$ ,  $\epsilon = 1, 2$ , which lie in the left part of  $I$ . We first define modified values of  $t$  which will appear in Lemmas 4.5 and 4.6:

$$t^*(v) = \frac{1}{3}t + \frac{\delta\mu}{18} \left( \frac{|v^*|}{\Delta(r)} \right)^2 \quad \text{and} \quad t^*(v, w) = \frac{1}{3}t + \frac{\delta\mu}{18} \left[ \left( \frac{|v^*|}{\Delta(r)} \right)^2 + \left( \frac{|w^*|}{\Delta(r)} \right)^2 \right]. \quad (4.30)$$

We consider a marked PG path with marked points in the  $(j + 1)$ th-scale grid: let

$$v^0 = \mathbf{m}_a(\Gamma_{xy}) = V_{j+1}(u_a(\Gamma_{xy})), \quad v^{fin} = \mathbf{m}_b(\Gamma_{xy}) = V_{j+1}(u_b(\Gamma_{xy})),$$

and for each  $k + 1 \leq \ell \leq j + 1$  let

$$v^\ell = \mathbf{m}_{a+\delta^\ell r}(\Gamma_{xy}) = V_{j+1}(u_{a+\delta^\ell r}(\Gamma_{xy})), \quad w^\ell = \Pi_{v^0 v^{fin}} \cap H_{a+\delta^\ell r},$$

and let

$$\alpha(\ell) = \frac{|v^\ell - w^\ell|^2}{\delta^\ell r}, \quad \kappa(\ell) = \frac{|v^\ell - w^\ell|^2}{\delta^\ell r \sigma(\delta^\ell r)}. \quad (4.31)$$

We may view  $\alpha(\ell)/2$  as an approximation of the “extra distance at scale  $\delta^\ell r$ ,” that is, of  $|v^\ell - v^0| + |v^{fin} - v^\ell| - |v^{fin} - v^0|$ ;  $\kappa(\ell)/2$  is this extra distance normalized by the fluctuation size. For  $k + 1 \leq \ell < j + 1$  let  $g^{\ell+1} = \Pi_{v^0 v^\ell} \cap H_{a+\delta^{\ell+1} r}$ , and let

$$\theta(\ell + 1) = \frac{|g^{\ell+1} - w^{\ell+1}|}{|v^{\ell+1} - w^{\ell+1}|} = \delta \frac{|v^\ell - w^\ell|}{|v^{\ell+1} - w^{\ell+1}|}; \quad (4.32)$$

see Figure 4. Then

$$\frac{|v^{\ell+1} - g^{\ell+1}|^2}{\delta^{\ell+1} r} = (1 - \theta(\ell + 1))^2 \alpha(\ell + 1) \quad (4.33)$$

so in view of (4.25), there is an “extra distance”

$$\Upsilon_{Euc}(v^0, v^{\ell+1}, v^\ell) - |v^\ell - v^0| \geq \frac{1}{3} \frac{|v^{\ell+1} - g^{\ell+1}|^2}{\delta^{\ell+1} r} = \frac{(1 - \theta(\ell + 1))^2}{3} \alpha(\ell + 1). \quad (4.34)$$

Suppose that for some  $\ell \in [k + 1, j]$  we have

$$2^{j-\ell} \kappa(j + 1) \leq \kappa(\ell + 1), \quad 2\kappa(\ell + 1) \geq \kappa(\ell), \quad (4.35)$$

as happens if  $2^m \kappa(m)$  is maximized at  $m = \ell$ . From (1.11), the second inequality ensures that

$$\theta(\ell + 1)^2 \leq 2\delta \frac{\sigma(\delta^\ell r)}{\sigma(\delta^{\ell+1} r)} \leq \frac{1}{4} \quad \text{and} \quad 2\alpha(\ell + 1) \geq \frac{\sigma(\delta^{\ell+1} r)}{\sigma(\delta^\ell r)} \alpha(\ell) \geq C_{23}^{-1} \delta^{\chi_2} \alpha(\ell), \quad (4.36)$$

and the first inequality in (4.35) tells us that

$$\alpha(\ell + 1) \geq 2^{j-\ell} \frac{\sigma(\delta^{\ell+1} r)}{\sigma(\delta^{j+1} r)} \alpha(j + 1) \geq C_{22} \left( \frac{2}{\delta^{\chi_1}} \right)^{j-\ell} \alpha(j + 1). \quad (4.37)$$

It then follows from (4.34), (4.36), and (4.37) that the path from  $v^0$  to  $v^\ell$  is bowed in the sense that

$$\Upsilon_{Euc}(v^0, v^{\ell+1}, v^\ell) - |v^\ell - v^0| \geq \frac{C_{22}}{12} \left( \frac{2}{\delta^{\chi_1}} \right)^{j-\ell} \alpha(j + 1). \quad (4.38)$$

Motivated by this we let

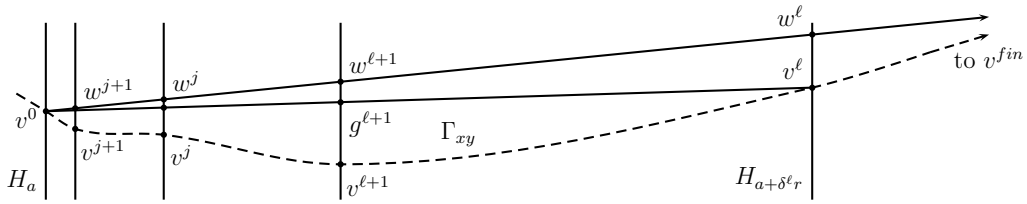


Figure 4: The left end of a long  $j$ th-scale interval in which  $L^-(I) = \ell$ . Starting from the left, there is an endpoint hyperplane, then one of its sandwiching hyperplanes; the rightmost two are the joining hyperplanes, containing  $v^{\ell+1}$  and  $v^\ell$ . In the bowed case,  $v^{\ell+1}$  (at distance  $\delta^{\ell+1}r$  from  $H_a$ ) is “sufficiently far” from  $w^{\ell+1}$ , creating more bowedness than exists on other length scales  $\delta^i r$ .

$$L^-(I) = \begin{cases} j & \text{if } \max(\alpha(j+1), \delta^{-1}\alpha(j)) \leq \frac{1}{16\mu} \left(\frac{\lambda}{7}\right)^{j+1} t^*(v^0)\sigma(r), \\ \arg \max_{\ell \in [k+1, j]} 2^\ell \kappa(\ell) - 1 & \text{if } \max(\alpha(j+1), \delta^{-1}\alpha(j)) > \frac{1}{16\mu} \left(\frac{\lambda}{7}\right)^{j+1} t^*(v^0)\sigma(r) \\ & \text{and } \arg \max_{\ell \in [k+1, j]} 2^\ell \kappa(\ell) > k+1, \\ k+1 & \text{otherwise.} \end{cases} \tag{4.39}$$

We refer to the 3 options in (4.39) as the *forward*, *bowed*, and *totally unbowed* cases, respectively. They may be interpreted as follows. In the forward case the initial steps  $v^0 \rightarrow v^1 \rightarrow v^2$  have little sidestepping, and we will see that this eliminates the need to exploit bowedness; this should be viewed as the “baseline” or “most likely” case. Otherwise we look for a scale  $\delta^\ell r$ , with  $k+1 \leq \ell \leq j$ , on which  $\Gamma_{xy}$  is bowed as in (4.38), by seeking a scale (the arg max) satisfying (4.35). In the bowed case such a scale exists; see Figure 4. In the totally unbowed case there is no such scale, meaning  $2^\ell \kappa(\cdot)$  is maximized for essentially the full length scale of the interval  $I$ . By (4.37) this forces the extra distance  $\alpha(k+1)$  to be very large. We define the inner and outer joining points as

$$\mu_{xy}^{-,2}(I) = a + \delta^{L^-(I)} r; \quad \mu_{xy}^{-,1}(I) = a + \delta^{L^-(I)+1} r.$$

We define  $\mu_{xy}^{+,\epsilon}(I)$  in a mirror image manner to  $\mu_{xy}^{-,\epsilon}(I)$  in  $[a, b]$ , going backwards from  $b$  to  $a$  instead of forward from  $a$  to  $b$ . That is, we use the points  $\hat{v}^\ell = \mathfrak{m}_{b-\delta^\ell r}(\Gamma_{xy})$  in place of the  $v^\ell$ 's and  $t^*(v^{fin})$  in place of  $t^*(v^0)$ ; otherwise the definition is the same, and the analogs of (4.34), (4.36), and (4.37) are valid for the analog  $L^+(I)$  of  $L^-(I)$ .

We now describe the rules for which of the four  $(j+1)$ th-scale joining hyperplanes  $H_s$  with  $s = \mu_{xy}^{\pm,\epsilon}(I)$ , in a long  $j$ th-scale interval  $I$ , are included in  $\mathcal{H}_{xy}$ . We note again that the endpoint and sandwiching hyperplanes at both ends of  $I$  are always included; in some instances the sandwiching hyperplanes coincide with outer joining hyperplanes, so these criteria never rule out the inclusion of such hyperplanes.

- (i) If both  $\pm$  ends of  $I$  have the forward case, then we include the inner joining hyperplanes in  $\mathcal{H}_{xy}$ ; these are at distance  $\delta^j r$  from the interval ends. The outer ones coincide with the sandwiching hyperplanes at distance  $\delta^{j+1} r$  from the interval ends, so they are included as well.
- (ii) If both  $\pm$  ends have the bowed case, then we include both the inner and outer joining hyperplanes in  $\mathcal{H}_{xy}$ . We note that when either of  $L^\pm(I) = j$ , the corresponding outer joining hyperplane coincides with the sandwiching hyperplane as in (i), so it is already in  $\mathcal{H}_{xy}$  on that basis.
- (iii) If both  $\pm$  ends have the totally unbowed case, then we include only the inner joining hyperplanes.

- (iv) If the two ends have different cases, we determine which end of the interval is *dominant* according to the criterion described next. We then include the joining hyperplane(s) only at the dominant end, 1 or 2 hyperplanes in accordance with (i)–(iii) above. We call this the *mixed case*.

To determine the dominant end of a long  $j$ th-scale interval  $I = [a, b]$ , necessarily having  $k$ th-scale length for some  $k < j$ , in the mixed case, we first select those non-sandwiching hyperplanes which are candidates for inclusion in  $\mathcal{H}_{xy}$ , in accordance with (i)–(iii) above. (For example, if the path has the bowed case with  $L^-(I) \neq j$  at the left end and the forward case at the right, we select the inner and outer joining hyperplanes on the left, and only the inner on the right.) For these candidate hyperplanes, along with the 4 endpoint and sandwiching hyperplanes in  $I$ , we consider the corresponding “tentative” marked PG path (part of  $\Gamma_{xy}$ ) with a marked point in each of the hyperplanes:  $u^0 \rightarrow \dots \rightarrow u^n$  with  $n = 5$  or  $6$ , where  $u^0 = V_j(\mathbf{m}_{(u^0)_1}(\Gamma_{xy}))$  and  $u^i = V_{j+1}(\mathbf{m}_{(u^i)_1}(\Gamma_{xy}))$ ,  $1 \leq i \leq n$ , are  $j$ th and  $(j+1)$ th-scale CG approximations. The inner joining hyperplanes contain  $u^\ell, u^{\ell+1}$ , with  $\ell = 2$  or  $3$ ; we call this  $\ell$  the *central index*; the gap from  $u^\ell$  to  $u^{\ell+1}$  is the longest in  $I$ , at least  $8\delta^{k+1}r$ . Let  $w_\perp^i$  be the orthogonal projection of  $u^i$  into the line  $\Pi_{u^0u^n}$ ; see Figure 5. We use the fact that the excess length  $\mathcal{E}(u^0, \dots, u^n) = \Upsilon_{Euc}(u^0, \dots, u^n) - |u^n - u^0|$  can be approximately split into components associated with the two ends, as follows. When the central index is  $\ell$  we have

$$\begin{aligned} \mathcal{E}(u^0, \dots, u^n) &= \left[ \Upsilon_{Euc}(u^0, \dots, u^\ell) - |w_\perp^\ell - u^0| \right] + \left[ |u^{\ell+1} - u^\ell| - |w_\perp^{\ell+1} - w_\perp^\ell| \right] \\ &\quad + \left[ \Upsilon_{Euc}(u^{\ell+1}, \dots, u^n) - |w_\perp^{\ell+1} - u^n| \right]. \end{aligned} \tag{4.40}$$

For the first difference on the right we have from (4.26)

$$\Upsilon_{Euc}(u^0, \dots, u^\ell) - |w_\perp^\ell - u^0| \geq |u^\ell - u^0| - |w_\perp^\ell - u^0| \geq \frac{|u^\ell - w_\perp^\ell|^2}{3\delta^{k+1}r},$$

and similarly for the third difference in (4.40), while for the middle one, by (4.26),

$$\begin{aligned} |u^{\ell+1} - u^\ell| - |w_\perp^{\ell+1} - w_\perp^\ell| &\leq \frac{(|u^\ell - w_\perp^\ell| + |u^{\ell+1} - w_\perp^{\ell+1}|)^2}{2 \cdot 8\delta^{k+1}r} \\ &\leq \frac{1}{8} \left( \frac{|u^\ell - w_\perp^\ell|^2}{\delta^{k+1}r} + \frac{|u^{\ell+1} - w_\perp^{\ell+1}|^2}{\delta^{k+1}r} \right) \end{aligned} \tag{4.41}$$

so the middle difference in (4.40) is only a limited fraction of the whole:

$$|u^{\ell+1} - u^\ell| - |w_\perp^{\ell+1} - w_\perp^\ell| \leq \frac{3}{8} \mathcal{E}(u^0, \dots, u^n). \tag{4.42}$$

We designate the left end of  $I$  as *dominant* if the first of the 3 differences on the right in (4.40) is larger than the third difference, and the right end in the reverse case. As given in (iv) above, we include in  $\mathcal{H}_{xy}$  only the candidate joining hyperplanes from the dominant end; the non-dominant end has its endpoint and sandwiching hyperplanes but no joining ones.

We note that if (for illustration) the left end is dominant, then, after excluding the right-end joining hyperplanes, we are left with the marked PG path  $u^0 \rightarrow \dots \rightarrow u^\ell \rightarrow u^{n-1} \rightarrow u^n$ , for which the contribution of the right end to the extra length can be bounded: we have

$$|u^n - u^{n-1}| - |u^n - w_\perp^{n-1}| \leq \Upsilon_{Euc}(u^{\ell+1}, \dots, u^n) - |w_\perp^{\ell+1} - u^n| \leq \Upsilon_{Euc}(u^0, \dots, u^\ell) - |w_\perp^\ell - u^0| \tag{4.43}$$

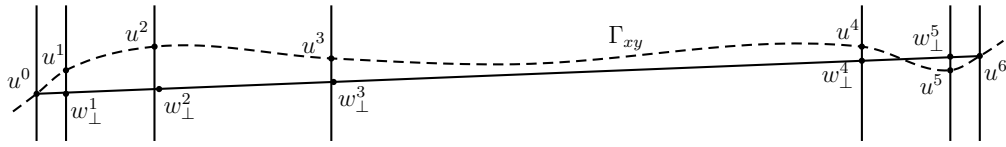


Figure 5: The mixed case with the bowed case at the (dominant) left end of the interval, and the forward case at the right end, showing the candidate hyperplanes (the 3 middle ones) and tentative marked PG path. The central index is 3. Since the right end is not dominant, the outer joining hyperplane there, containing  $v^4$ , is not included in  $\mathcal{H}_{xy}$ . The full length of the interval is between  $10\delta^{k+1}r$  and  $10\delta^k r$ , and all hyperplanes lie in the leftmost or rightmost 1/10 of the interval.

where the last inequality follows from dominance of the left end, so similarly to (4.42),

$$\frac{|u^{n-1} - w_{\perp}^{n-1}|^2}{3\delta^{j+1}r} \leq |u^n - u^{n-1}| - |u^n - w_{\perp}^{n-1}| \leq \frac{1}{2}\mathcal{E}(u^0, \dots, u^{\ell}, u^{n-1}, u^n). \quad (4.44)$$

The same bound with  $|u^1 - w_{\perp}^1|$  in place of  $|u^{n-1} - w_{\perp}^{n-1}|$  holds symmetrically when the right end is dominant.

**Remark 4.3.** In the bowed case at the left end of a  $j$ th-scale interval  $I = [a, b]$ , with  $L^-(I) = \ell$ , define

$$z^0 = v^0, \quad z^m = \Pi_{v^0, v^{\ell}} \cap H_{a+\delta^m r}, \quad \ell \leq m \leq j+1,$$

and symmetrically at the right end. See Figure 4;  $g^{\ell+1}$  there is  $z^{\ell+1}$  here. We have from (4.37) and (4.39)

$$\begin{aligned} \frac{|v^{\ell+1} - w^{\ell+1}|^2}{\delta^{\ell+1}r} \frac{\sigma(\delta^{j+1}r)}{\sigma(\delta^{\ell+1}r)} &\geq \max \left( 2^{j-\ell} \frac{|v^{j+1} - w^{j+1}|^2}{\delta^{j+1}r}, 2^{j-\ell-1} \frac{\sigma(\delta^{j+1}r)}{\sigma(\delta^j r)} \frac{|v^j - w^j|^2}{\delta^j r} \right) \\ &\geq 2^{j-\ell-1} \frac{\delta}{16\mu} \frac{\sigma(\delta^{j+1}r)}{\sigma(\delta^j r)} \left( \frac{\lambda}{7} \right)^{j+1} t^*(v^0)\sigma(r) \end{aligned}$$

and

$$\frac{|v^{\ell} - w^{\ell}|^2}{\delta^{\ell}r} \leq \frac{2\sigma(\delta^{\ell}r)}{\sigma(\delta^{\ell+1}r)} \frac{|v^{\ell+1} - w^{\ell+1}|^2}{\delta^{\ell+1}r}, \quad (4.45)$$

and as in (4.36) it follows from these that  $|v^{\ell+1} - z^{\ell+1}| \geq \frac{1}{2}|v^{\ell+1} - w^{\ell+1}|$  and then that

$$\frac{|v^{\ell+1} - z^{\ell+1}|^2}{\delta^{\ell+1}r} \geq 2^{j-\ell} \frac{\delta}{128\mu} \left( \frac{\lambda}{7} \right)^{j+1} \frac{\sigma(\delta^{\ell+1}r)}{\sigma(\delta^j r)} t^*(v^0)\sigma(r). \quad (4.46)$$

The advantage of (4.46) is that it depends only on  $(v^0, v^1, v^2, v^3)$  and not on  $v^{fin}$ , whereas in (4.45) the points  $w^i$  do depend on  $v^{fin}$ . What we have shown is that if there exists  $v^{fin} \in G_r^+$  with  $(v^{fin} - v^0)_1 \geq \delta^{\ell}r$  for which (4.45) holds, then (4.46) holds, not involving  $v^{fin}$ . We further have  $w^{j+1} - z^{j+1} = \delta^{j-\ell}(w^{\ell+1} - z^{\ell+1})$ , while by the first half of (4.36) we have  $|v^{\ell+1} - w^{\ell+1}| \leq 2|v^{\ell+1} - z^{\ell+1}|$  and  $|w^{\ell+1} - z^{\ell+1}| \leq |v^{\ell+1} - z^{\ell+1}|$ , so

$$\begin{aligned} \frac{|v^{j+1} - z^{j+1}|^2}{\delta^{j+1}r} &\leq 2 \frac{|v^{j+1} - w^{j+1}|^2}{\delta^{j+1}r} + 2 \frac{|w^{j+1} - z^{j+1}|^2}{\delta^{j+1}r} \\ &= 2 \frac{|v^{j+1} - w^{j+1}|^2}{\delta^{j+1}r} + 2\delta^{2(j-\ell)} \frac{|w^{\ell+1} - z^{\ell+1}|^2}{\delta^{j+1}r} \\ &\leq 2^{-(j-\ell-3)} \frac{\sigma(\delta^{j+1}r)}{\sigma(\delta^{\ell+1}r)} \frac{|v^{\ell+1} - z^{\ell+1}|^2}{\delta^{\ell+1}r} + 2\delta^{j-\ell} \frac{|v^{\ell+1} - z^{\ell+1}|^2}{\delta^{\ell+1}r} \\ &\leq 2^{-(j-\ell-4)} \frac{\sigma(\delta^{j+1}r)}{\sigma(\delta^{\ell+1}r)} \frac{|v^{\ell+1} - z^{\ell+1}|^2}{\delta^{\ell+1}r}. \end{aligned} \quad (4.47)$$



Again the left and right expressions in (4.47) depend only on  $(v^0, v^{j+1}, v^{\ell+1}, v^\ell)$ , not on  $v^{fin}$ .

In (1.14) we may interpret  $t\sigma(r)$  as a reduction in the time allotted to go from  $x$  to  $y$ , relative to  $h(|(y-x)_1|)$ . In place of the reduction  $t\sigma(r)$  relative to  $h(|(y-x)_1|)$ , we can consider a modified reduction, call it  $R_0$ , which is relative to  $\Upsilon_h(\Gamma^{CG})$ :

$$h(|(y-x)_1|) - t\sigma(r) = \Upsilon_h(\Gamma^{CG}) - R_0.$$

The modified reduction is larger: using (4.29) we see that

$$R_0 \geq t\sigma(r) + \frac{\mu}{3}\Psi(\Gamma^{CG}).$$

We will need to (roughly speaking) allocate pieces of  $R_0$  to the various transitions made by  $\Gamma^{CG}$  and certain related paths. Motivated by this, we define the  $j$ th-scale allocation  $A_j^0(v, w)$  of a transition  $v \rightarrow w$  to be

$$A_j^0(v, w) = \lambda^j \left( \frac{t\sigma(r)}{7^j} + \delta\mu \frac{|(w-v)^*|^2}{|(w-v)_1|} \right). \tag{4.48}$$

In (4.48) the factor  $7^j$  is used due to (4.23). For a marked PG path  $\Gamma^{CG} : v^0 \rightarrow v^1 \rightarrow \dots \rightarrow v^{m+1}$  as above, with some number  $m \leq 7^j - 1$  of marks (in the  $j$ th-scale grid, in cases of interest), we have from (4.27)

$$\begin{aligned} \sum_{i=1}^{m+1} A_j^0(v^{i-1}, v^i) &\leq \lambda^j [t\sigma(r) + \delta\mu\Psi(\Gamma^{CG})] \\ &\leq \lambda^j \left[ t\sigma(r) + 3\delta\mu \left( \Upsilon_{Euc}(\Gamma^{CG}) - (v^{m+1} - v^0)_1 \right) \right]. \end{aligned} \tag{4.49}$$

Also, since  $j \leq j_1 = O(\log \log r)$ , all  $A_j^0(v, w)$  are large provided  $r$  is large and  $t \geq 1$ .

**Remark 4.4.** We now add some detail to the proof sketch given in Section 4.1. Each geodesic  $\Gamma_{xy}$  can be viewed as a marked PG path with marks in the basic grid in each of the hyperplanes of  $\mathcal{H}_{xy}$ . The goal is to gradually coarsen this approximation on successively larger length scales until we obtain a final path  $\Gamma_{xy}^{CG}$ . The number of possible final paths (outside of a collection of “bad” paths having negligible probability) is small enough so that a version of (1.14) can be proved for final paths.

To perform the coarsening we iterate a two-stage process, with the exception that the first iteration has only one stage. The first iteration is on the  $j_1$ th scale, the second on the larger  $(j_1 - 1)$ th scale, and so on. For the  $j_1$ th-scale iteration, we perform a set of operations on the original marked PG path (essentially  $\Gamma_{xy}$ ) called *shifting to the  $j_1$ th-scale grid*, replacing each marked grid point (located in the basic grid) in each hyperplane in  $\mathcal{H}_{xy}$  with a nearby point in the  $j_1$ th-scale grid. Each further iteration has two stages. For the  $(j_1 - 1)$ th scale (second) iteration, in the first stage we shift those marked points lying in  $(j_1 - 1)$ th-scale hyperplanes in  $\mathcal{H}_{xy}$  to the  $(j_1 - 1)$ th-scale grid. In the second stage of the iteration, we remove from the marked PG path those marked points not lying in  $(j_1 - 1)$ th-scale hyperplanes, with exceptions for points in terminal hyperplanes. See Figure 6. In general, for the  $j$ th-scale iteration ( $j \leq j_1$ ), at the start of the iteration all the marked points in non-terminal hyperplanes are  $(j + 1)$ th-scale grid points in  $(j + 1)$ th-scale hyperplanes; in the first stage we shift the ones in  $j$ th-scale hyperplanes to the  $j$ th-scale grid, and in the second stage we remove the ones not in  $j$ th-scale hyperplanes, again with exceptions in terminal hyperplanes.

The difficulty is that as we alter the marked PG path, the length  $\Upsilon_{Euc}(\cdot)$  and corresponding  $h$ -sum  $\Upsilon_h(\cdot) \approx \mu\Upsilon_{Euc}(\cdot)$  change, with marked-point removals always reducing



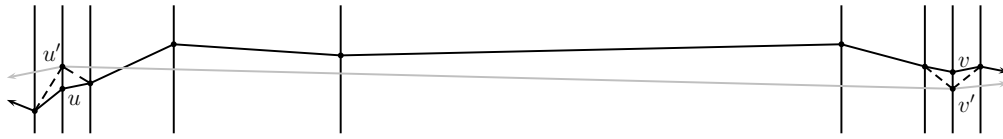


Figure 6: Illustration of the two stages of the  $j$ th-scale iteration. The black path is the current one at the start of the iteration. The points  $u, u', v, v'$  lie in the endpoint hyperplanes of a long  $j$ th-scale interval; the adjacent hyperplanes close on either side of these are the sandwiching ones, and the other three are joining hyperplanes. In the first stage, we shift to the  $j$ th-scale grid in the endpoint hyperplanes, replacing  $u, v$  in the marked PG path with  $u', v'$ , thus updating to the dashed path. In the second stage, the marked points between the endpoint hyperplanes are removed, updating to the gray path with marked points  $u', v'$ .

these sums, and we need to ensure that, with high probability, the corresponding sum of passage times  $\Upsilon_{\hat{T}}(\cdot)$  “tracks” these changes at least partly, to within a certain allocated error related to the above-mentioned  $A_j^0(v, w)$ . For shifting to a grid the tracking is not too difficult to achieve, as the allowed error turns out to be larger than the change in  $h$ -sum being tracked. But for removal of marked points the tracking requires multiple different strategies, depending on the options in (4.39) for the marked grid point locations in the gap between each two successive  $j$ th-scale hyperplanes in  $\mathcal{H}_{xy}$ . The particular tracking needed is that, with high probability to within the allocated errors, when marked points are removed from a gap, the decrease in total passage time  $\Upsilon_{\hat{T}}(\cdot)$  is at least a positive fraction  $\delta$  of the decrease in  $h$ -sum. The primary difficulty in achieving this is that if the gap has a large length  $L$ , then the relevant passage time fluctuation size  $\sigma(L)$  may overwhelm both the reduction in  $h$ -sum and the allocated errors; here the remedy involves the “joining hyperplanes.” We also make use of what we call *intermediate paths*, which (in most cases) have total passage time and  $h$ -sum in between the values that exist before and after the marked-point removal, and are chosen so that they are relatively easy to compare to the pre-removal path.

In (1.14) the passage time reduction for the full path is  $t\sigma(r)$ , which *a priori* suggests that the total of the allocated errors associated to a path should not exceed this amount. There is no natural way to work with such a small total error yet achieve bounds uniformly over all  $\Gamma_{xy}$ . However, as described in Section 4.1, the tracking enables us to increase the total of the allocated errors by an amount proportional to the extra Euclidean length  $\Upsilon_{Euc}(\cdot)$  of the original marked PG path relative to the “horizontal” distance  $(y - x)_1$ , which gives the second term in parentheses in the formula (4.48). With this the necessary uniformity can be achieved, both for the tracking and for the fluctuation bounds on passage times of final paths  $\Gamma_{xy}^{CG}$ .

#### 4.4 Step 2. Performing the $j_1$ th-scale (first) iteration of coarse-graining

As described in Remark 4.4, for the  $j_1$ th-scale iteration we perform a sequence of operations on marked PG paths called shifting to the  $j$ th-scale grid, in the hyperplanes of  $\mathcal{H}_{xy}$ . The  $j_1$ th-scale iteration is different from those that follow, in that there is no second stage of removing marked points, and no need for the “tracking” of Remark 4.4. In general we will refer to the marked PG path existing before a shift or removal operation as the *current (marked) path*, and the modified one resulting from the operation as the *updated (marked) path*. The allocations  $A_j^0(\cdot, \cdot)$  of (4.48) are used only for the  $j_1$ th-scale iteration; we will define other allocations later.

For the  $j_1$ th scale, every  $H_s \in \mathcal{H}_{xy}$  is a  $j_1$ th-scale hyperplane. Suppose  $\mathcal{H}_{xy} =$

$\{H_{s_i}, 1 \leq i \leq m\}$  with  $x_1 < s_1 < \dots < s_m < y_1$ . Let  $x^i = \mathbf{m}_{s_i}(\Gamma_{xy}) = \psi_q(u_{s_i}(\Gamma_{xy}))$ ,  $1 \leq i \leq m$ , and define  $x^0, p^0, x^{m+1}, p^{m+1}$  by  $x^0 = p^0 = \hat{x} = \psi_q(x)$ ,  $x^{m+1} = p^{m+1} = \hat{y} = \psi_q(y)$ . At the start, the current path  $\Gamma_{xy}^{j_1,0}$  is  $\Gamma_{xy}$  with marks at the grid points  $x^i$ :

$$\Gamma_{xy}^{j_1,0} : x^0 \rightarrow x^1 \rightarrow \dots \rightarrow x^{m+1}.$$

Note that  $\Gamma_{xy}^{j_1,0}$  is a marked path, but not necessarily a marked PG path, since  $\Gamma_{xy}$  need only pass near  $\psi_q(x^i)$ , not necessarily through it. By near we mean that both  $u_{s_i}(\Gamma_{xy})$  and  $\varphi(x^i)$  lie in the same cube  $F_{x^i}$  of the basic grid.

Recall that the blocks of  $\mathbb{L}_{j_1}$  have side  $K_0\beta^{j_1}\Delta(r)$ . The first shift to the  $j_1$ th-scale grid happens in  $H_{s_1}$ , replacing  $x^1$  with  $p^1 = V_{j_1}(x^1)$  to produce the updated marked path

$$\Gamma_{xy}^{j_1,1} : p^0 \rightarrow p^1 \rightarrow x^2 \rightarrow \dots \rightarrow x^{m+1},$$

with the underlying path being the concatenation of geodesics  $\Gamma_{xp^1}, \Gamma_{p^1, u_{s_2}(\Gamma_{xy})}, \Gamma_{u_{s_2}(\Gamma_{xy}), y}$ . Next we repeat this in  $H_{s_2}$ , replacing  $x^2$  with  $p^2 = V_{j_1}(x^2)$ . We continue this way performing shifts to the  $j_1$ th-scale grid in  $H_{s_3}, \dots, H_{s_m}$ , producing the updated path

$$\Gamma_{xy}^{j_1,m} : p^0 \rightarrow p^1 \rightarrow \dots \rightarrow p^m \rightarrow p^{m+1},$$

with the underlying path now (in view of (4.24)) being the marked PG path given by these points.

Let us analyze the effect of these shifts on  $\Upsilon_h(\Gamma_{xy}^{j_1,\cdot})$ . Consider the  $i$ th shift, replacing  $x^i$  with  $p^i$ . From (4.11) and basic geometry we have

$$|p^i - p^{i-1}| \geq |x^i - p^{i-1}| - c_{11} \frac{|(p^i - x^i)^*| |(p^i - p^{i-1})^*| + |(p^i - x^i)^*|^2}{(p^i - p^{i-1})_1}. \tag{4.50}$$

Consider first the ‘‘sidestepping’’ case:  $|(p^i - p^{i-1})^*| \geq \ell_2(j_1)\beta^{j_1}\Delta(r)$ . Since  $|(p^i - x^i)^*| \leq \sqrt{d-1}K_0\beta^{j_1}\Delta(r)$ , we have from (4.20), (4.21), and (4.50) that provided  $r$  is large, for  $1 \leq i \leq m$ ,

$$\begin{aligned} |p^i - p^{i-1}| &\geq |x^i - p^{i-1}| - \frac{2c_{11}K_0\sqrt{d-1}}{\ell_2(j_1)} \frac{|(p^i - p^{i-1})^*|^2}{(p^i - p^{i-1})_1} \\ &\geq |x^i - p^{i-1}| - 2c_{11}K_0\sqrt{d-1} \left( \frac{\beta}{\rho\delta^{(1+\chi_2)/2}} \right)^{j_1} \frac{|(p^i - p^{i-1})^*|^2}{(p^i - p^{i-1})_1} \\ &\geq |x^i - p^{i-1}| - \frac{1}{32\mu} A_{j_1}^0(p^{i-1}, p^i). \end{aligned} \tag{4.51}$$

Now consider the ‘‘normal’’ case:  $|(p^i - p^{i-1})^*| < \ell_2(j_1)\beta^{j_1}\Delta(r)$ . From (4.20), (4.21), and (4.50) we have

$$\begin{aligned} |p^i - p^{i-1}| &\geq |x^i - p^{i-1}| - 2c_{11}K_0\ell_2(j_1)\sqrt{d-1} \frac{(\beta^{j_1}\Delta(r))^2}{\delta^{j_1 r}} \\ &\geq |x^i - p^{i-1}| - c_{12} \left( \frac{\rho\beta}{\delta^{(1-\chi_1)/2}} \right)^{j_1} \sigma(r) \\ &\geq |x^i - p^{i-1}| - \frac{1}{32\mu} A_{j_1}^0(p^{i-1}, p^i). \end{aligned} \tag{4.52}$$

We can interchange the roles of  $p^i$  and  $x^i$  and/or replace  $p^{i-1}$  with  $x^{i+1}$ , so it follows from (4.51) and (4.52) that replacing  $x_i$  with  $p_i$  does not change the path length much—we have

$$\left| |p^i - p^{i-1}| - |x^i - p^{i-1}| \right| \leq \frac{1}{32\mu} \left( A_{j_1}^0(p^{i-1}, p^i) + A_{j_1}^0(p^{i-1}, x^i) \right) \tag{4.53}$$

and

$$\left| |p^i - x^{i+1}| - |x^i - x^{i+1}| \right| \leq \frac{1}{32\mu} \left( A_{j_1}^0(p^i, x^{i+1}) + A_{j_1}^0(x^i, x^{i+1}) \right) \quad (4.54)$$

and then also, for the  $h$ -length,

$$\left| h(|p^i - p^{i-1}|) - h(|x^i - p^{i-1}|) \right| \leq \frac{1}{16} \left( A_{j_1}^0(p^{i-1}, p^i) + A_{j_1}^0(p^{i-1}, x^i) \right) \quad (4.55)$$

and

$$\left| h(|p^i - x^{i+1}|) - h(|x^i - x^{i+1}|) \right| \leq \frac{1}{16} \left( A_{j_1}^0(p^i, x^{i+1}) + A_{j_1}^0(x^i, x^{i+1}) \right). \quad (4.56)$$

We claim that

$$A_{j_1}^0(p^{i-1}, p^i) \leq \left( 1 + \frac{1}{4} \lambda^{j_1} \right) A_{j_1}^0(x^{i-1}, x^i). \quad (4.57)$$

It is enough to show

$$|(p^i - p^{i-1})^*|^2 - |(x^i - x^{i-1})^*|^2 \leq \frac{1}{4\delta\mu} \lambda^{j_1} \frac{t\sigma(r)}{7^{j_1}} (x^i - x^{i-1})_1 + \frac{1}{4} \lambda^{j_1} |(x^i - x^{i-1})^*|^2. \quad (4.58)$$

To that end, we have from the definition of  $p^i$  that

$$\begin{aligned} & |(p^i - p^{i-1})^*| \\ & \leq |(x^i - x^{i-1})^*| + |(x^i - p^i)^*| + |(p^{i-1} - x^{i-1})^*| \leq |(x^i - x^{i-1})^*| + 2K_0\sqrt{d-1}\beta^{j_1}\Delta(r), \end{aligned}$$

so

$$|(p^i - p^{i-1})^*|^2 - |(x^i - x^{i-1})^*|^2 \leq 4K_0\sqrt{d-1}\beta^{j_1}\Delta(r)|(x^i - x^{i-1})^*| + 4K_0^2(d-1)\beta^{2j_1}r\sigma(r). \quad (4.59)$$

From (4.20), the last term satisfies

$$4K_0^2(d-1)\beta^{2j_1}r\sigma(r) \leq \frac{1}{8} \left( \frac{\lambda\delta}{7} \right)^{j_1} t\sigma(r) \leq \frac{1}{8} \left( \frac{\lambda}{7} \right)^{j_1} t\sigma(r)(x^i - x^{i-1})_1. \quad (4.60)$$

If the  $i$ th transition has very small sidestep, that is,

$$|(x^i - x^{i-1})^*| \leq 32K_0\sqrt{d-1} \left( \frac{\beta}{\lambda} \right)^{j_1} \Delta(r), \quad (4.61)$$

then provided  $r$  is large, since  $(x^i - x^{i-1})_1 \geq \delta^{j_1}r$ , using (4.20) the first term on the right in (4.59) satisfies

$$\begin{aligned} & 4K_0\sqrt{d-1}\beta^{j_1}\Delta(r)|(x^i - x^{i-1})^*| \\ & \leq 128K_0^2(d-1) \left( \frac{\beta^2}{\lambda} \right)^{j_1} r\sigma(r) \\ & \leq 128K_0^2(d-1) \left( \frac{\beta^2}{\lambda\delta} \right)^{j_1} \sigma(r)(x^i - x^{i-1})_1 \leq \frac{1}{8\delta\mu} \left( \frac{\lambda}{7} \right)^{j_1} t\sigma(r)(x^i - x^{i-1})_1. \end{aligned} \quad (4.62)$$

If instead the  $i$ th transition has larger sidestep, meaning

$$|(x^i - x^{i-1})^*| > 32K_0\sqrt{d-1} \left( \frac{\beta}{\lambda} \right)^{j_1} \Delta(r), \quad (4.63)$$

then the first term on the right in (4.59) satisfies

$$4K_0\sqrt{d-1}\beta^{j_1}\Delta(r)|(x^i - x^{i-1})^*| \leq \frac{1}{8} \lambda^{j_1} |(x^i - x^{i-1})^*|^2. \quad (4.64)$$

Together, (4.59)–(4.64) prove (4.58), and thus also (4.57). This same proof shows that

$$A_{j_1}^0(p^{i-1}, p^i), A_{j_1}^0(p^{i-1}, x^i), A_{j_1}^0(x^{i-1}, x^i) \text{ are all within a factor of } 1 + \frac{1}{4}\lambda^{j_1} \leq \frac{4}{3}, \quad (4.65)$$

and similarly for  $A_{j_1}^0(x^i, x^{i+1}), A_{j_1}^0(p^i, x^{i+1}), A_{j_1}^0(p^i, p^{i+1})$ .

From (4.49), (4.53), (4.54), and (4.65) we bound the total change in length from all shifts to the  $j_1$ th-scale grid:

$$\begin{aligned} |\Upsilon_{Euc}(\Gamma_{xy}^{j_1, m}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_1, 0})| &\leq \frac{1}{6\mu} \sum_{i=1}^{m+1} A_{j_1}^0(p^{i-1}, p^i) \\ &\leq \frac{\lambda^{j_1}}{6\mu} \left[ t\sigma(r) + 3\delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1, m}) - (\hat{y} - \hat{x})_1 \right) \right] \end{aligned} \quad (4.66)$$

and similarly, using (4.55)–(4.56) instead of (4.53)–(4.54),

$$|\Upsilon_h(\Gamma_{xy}^{j_1, m}) - \Upsilon_h(\Gamma_{xy}^{j_1, 0})| \leq \frac{\lambda^{j_1}}{3} \left[ t\sigma(r) + 3\delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1, m}) - (\hat{y} - \hat{x})_1 \right) \right]. \quad (4.67)$$

In view of (4.65), the derivation of (4.66) and (4.67) is also valid if we replace  $A_{j_1}^0(p^{i-1}, p^i)$  with  $A_{j_1}^0(x^{i-1}, x^i)$ , which gives the alternate bound

$$|\Upsilon_h(\Gamma_{xy}^{j_1, m}) - \Upsilon_h(\Gamma_{xy}^{j_1, 0})| \leq \frac{\lambda^{j_1}}{3} \left[ t\sigma(r) + 3\delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1, 0}) - (\hat{y} - \hat{x})_1 \right) \right], \quad (4.68)$$

and we may similarly replace  $m$  with  $0$  on the right in (4.66).

For basic grid points  $u$  and the cubes  $F_u$  of (3.1), define

$$M(u) = \max\{T(y, z) : y, z \in F_u\}.$$

We have

$$T(x, y) \geq \sum_{i=1}^{m+1} \hat{T}(x^{i-1}, x^i) = \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1, 0}) \quad (4.69)$$

and a form of approximate subadditivity holds: for basic grid points  $u, v, w$ ,

$$\hat{T}(u, v) \leq \hat{T}(u, w) + \hat{T}(w, v) + M(w). \quad (4.70)$$

From Lemma 3.1 we have for sufficiently large  $s$  that the event

$$J^{(0)}(s) : M(u) \geq s \log r \text{ for some } u \in q\mathbb{Z}^d \cap G_r^+$$

satisfies

$$P\left(J^{(0)}(s)\right) \leq c_{13}|G_r^+|e^{-c_{14}s \log r} \leq c_{13}e^{-c_{15}s \log r}. \quad (4.71)$$

Before proceeding we stress that  $m, p^i$ , and  $x^i$  should always be viewed at functions of  $(x, y, \omega)$ . From (4.49) we have

$$\sum_{i=1}^{m+1} A_{j_1}^0(x^{i-1}, x^i) \leq \lambda^{j_1} \left( t\sigma(r) + 3\delta\mu \left[ \Upsilon_{Euc}(\Gamma_{xy}^{j_1, 0}) - (\hat{y} - \hat{x})_1 \right] \right). \quad (4.72)$$

Therefore using (4.66),

$$\begin{aligned} &\left\{ T(x, y) \leq h((\hat{y} - \hat{x})_1) - t\sigma(r) \right\} \\ &\subset \left\{ \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1, 0}) - h((\hat{y} - \hat{x})_1) \leq -(1 - 2\lambda^{j_1}) t\sigma(r) + 4\delta\mu\lambda^{j_1} \left[ \Upsilon_{Euc}(\Gamma_{xy}^{j_1, m}) - (\hat{y} - \hat{x})_1 \right] \right\} \end{aligned}$$

$$\begin{aligned} & - \sum_{i=1}^{m+1} A_{j_1}^0(x^{i-1}, x^i) \Big\} \\ \subset & \left\{ \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1, m}) - h((\hat{y} - \hat{x})_1) \leq - (1 - 2\lambda^{j_1}) t\sigma(r) + 4\delta\mu\lambda^{j_1} [\Upsilon_{Euc}(\Gamma_{xy}^{j_1, m}) - (\hat{y} - \hat{x})_1] \right\} \\ & \cup \left\{ \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1, m}) - \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1, 0}) > \sum_{i=1}^{m+1} A_{j_1}^0(x^{i-1}, x^i) \right\}. \end{aligned} \tag{4.73}$$

The key inclusion here is the second one, as it takes us from an event involving the passage time of the original path  $\Gamma_{xy}^{j_1, 0}$  to an event involving the  $j_1$ th-scale CG approximation  $\Gamma_{xy}^{j_1, m}$ , up to the “tracking error event” given in the last line. The name is only partly suitable here—bounding the last probability in (4.73) is related to the tracking of Remark 4.4, in that we are ensuring that changing the path from  $\Gamma_{xy}^{j_1, 0}$  to  $\Gamma_{xy}^{j_1, m}$  doesn’t change  $\Upsilon_{\hat{T}}(\cdot)$  too much, but the change in  $h$ -sum here is too small to require being tracked.

### 4.5 Step 3. Bounding the tracking-error event for the $j_1$ th-scale-iteration

Consider next the tracking-error event

$$J_{xy}^{(1)} : \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1, m}) - \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1, 0}) > \sum_{i=1}^{m+1} A_{j_1}^0(x^{i-1}, x^i)$$

from the right side of (4.73). Recalling Remark 4.4, this reflects the failure of the passage time to track well when the path changes from  $\Gamma_{xy}^{j_1, 0}$  to its  $j_1$ th-scale CG approximation  $\Gamma_{xy}^{j_1, m}$ . We have

$$\Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1, m}) - \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1, 0}) \leq \sum_{i=1}^m \left[ \left( \hat{T}(p^{i-1}, p^i) - \hat{T}(p^{i-1}, x^i) \right) + \left( \hat{T}(p^i, x^{i+1}) - \hat{T}(x^i, x^{i+1}) \right) \right] \tag{4.74}$$

and therefore using (4.65),

$$\begin{aligned} J_{xy}^{(1)} \subset & \bigcup_{i=1}^m \left\{ \hat{T}(p^{i-1}, p^i) - \hat{T}(p^{i-1}, x^i) \geq \frac{3}{8} A_{j_1}^0(p^{i-1}, x^i) \right\} \\ & \cup \bigcup_{i=1}^m \left\{ \hat{T}(p^i, x^{i+1}) - \hat{T}(x^i, x^{i+1}) \geq \frac{3}{8} A_{j_1}^0(x^i, x^{i+1}) \right\}. \end{aligned} \tag{4.75}$$

Let us consider any one of the events in the first union on the right in (4.75); we prepare to apply Proposition 3.5. We assume  $|p^i - p^{i-1}| \geq |x^i - p^{i-1}|$  as the opposite case is similar. Define  $\epsilon$  by

$$(1 - \epsilon)|p^i - p^{i-1}| = |x^i - p^{i-1}|$$

and let

$$\bar{p}^i = p^{i-1} + (1 - \epsilon)(p^i - p^{i-1}), \quad \tilde{p}^i = \psi_q(\varphi(\bar{p}^i)).$$

From (4.51) and (4.52) we have

$$|\bar{p}^i - p^{i-1}| = |x^i - p^{i-1}|, \quad |p^i - \bar{p}^i| = \epsilon|p^i - p^{i-1}| = |p^i - p^{i-1}| - |x^i - p^{i-1}| \leq \frac{1}{32\mu} A_{j_1}^0(x^{i-1}, p^i). \tag{4.76}$$

We split  $\hat{T}(p^{i-1}, p^i)$  into two corresponding increments, using (4.70):

$$\left\{ \hat{T}(p^{i-1}, p^i) - \hat{T}(p^{i-1}, x^i) \geq \frac{3}{8} A_{j_1}^0(p^{i-1}, x^i) \right\}$$

$$\subset \left\{ \hat{T}(p^{i-1}, \tilde{p}^i) - \hat{T}(p^{i-1}, x^i) \geq \frac{3}{16} A_{j_1}^0(p^{i-1}, x^i) \right\} \cup \left\{ \hat{T}(\tilde{p}^i, p^i) + M(\tilde{p}^i) \geq \frac{3}{16} A_{j_1}^0(p^{i-1}, x^i) \right\} \tag{4.77}$$

and define the corresponding unions

$$J_{xy}^{(1a)} = \bigcup_{i=1}^m \left\{ \hat{T}(p^{i-1}, \tilde{p}^i) - \hat{T}(p^{i-1}, x^i) \geq \frac{3}{16} A_{j_1}^0(p^{i-1}, x^i) \right\},$$

$$J_{xy}^{(1b)} = \bigcup_{i=1}^m \left\{ \hat{T}(\tilde{p}^i, p^i) + M(\tilde{p}^i) \geq \frac{3}{16} A_{j_1}^0(p^{i-1}, x^i) \right\},$$

so that (4.77) says the first union in (4.75) is contained in  $J_{xy}^{(1a)} \cup J_{xy}^{(1b)}$ . Define the set of  $(x, y)$  corresponding to (4.18):

$$X_r = \left\{ (x, y) \in \mathbb{V} \times \mathbb{V} : x, y \in G_r(K), |y - x| > \frac{C_{62}r}{(\log r)^{1/\chi_1}}, (4.24) \text{ holds} \right\},$$

with  $C_{62}$  from (4.18), and define the events

$$J^{(1a)} = \cup_{(x,y) \in X_r} J_{xy}^{(1a)}, \quad J^{(1b)} = \cup_{(x,y) \in X_r} J_{xy}^{(1b)}, \quad J^{(1c)} = \left\{ \Gamma_{xy} \not\subset G_r^+ \text{ for some } (x, y) \in X_r \right\}.$$

For configurations  $\omega \notin J^{(1c)}$ , all  $p^{i-1}, p^i, x^i, \tilde{p}^i$  lie in  $G_r^+$ , so the number of possible tuples  $(p^{i-1}, p^i, x^i, \tilde{p}^i)$  arising from some  $(x, y) \in X_r$  is at most  $c_{16}|G_r^+|^4$ . Define  $s, \alpha$  by

$$\Delta(s) = 2K_0\sqrt{d-1}\beta^{j_1}\Delta(r), \quad \alpha = \frac{3\lambda^{j_1}t\sigma(r)}{16\sigma(s)\log(2K_0\sqrt{d-1}\beta^{j_1}\Delta(r))},$$

so that  $|\tilde{p}^i - x^i| \leq |\bar{p}^i - x^i| + |\tilde{p}^i - \bar{p}^i| \leq 2|p^i - x^i| + q\sqrt{d-1} \leq \Delta(s)$ . By (1.11),

$$\frac{r\sigma(r)}{s\sigma(s)} = \frac{\Delta(r)^2}{\Delta(s)^2} \geq \frac{c_{17}}{\beta^{2j_1}} \quad \text{and hence} \quad \frac{\sigma(r)}{\sigma(s)} \geq \frac{c_{18}}{\beta^{2\chi_1 j_1/(1+\chi_1)}}.$$

From this, (4.20), and (4.22),

$$\alpha \geq c_{19} \left( \frac{\lambda}{\beta^{2\chi_1/(1+\chi_1)}} \right)^{j_1} \frac{t}{\log r} \geq c_{20}t(\log r)^2. \tag{4.78}$$

Now

$$\frac{3}{16} A_{j_1}^0(p^{i-1}, x^i) \geq \frac{3\lambda^{j_1}t\sigma(r)}{16} = \alpha\sigma(s)\log(2K_0\sqrt{d-1}\beta^{j_1}\Delta(r)) = \alpha\sigma(s)\log\Delta(s)$$

so for  $\omega \in J_{xy}^{(1a)}$ , using  $\Delta(s) \geq |\tilde{p}^i - x^i|$  we have for some  $i$  that

$$\hat{T}(p^{i-1}, \tilde{p}^i) - \hat{T}(p^{i-1}, x^i) \geq \alpha\sigma(s)\log\Delta(s) \geq \alpha\sigma(s\log\Delta(s)) \geq \alpha\sigma\left(\Delta^{-1}(|\tilde{p}^i - x^i|)\log|\tilde{p}^i - x^i|\right).$$

In view of (4.76) (first equality) and (4.78) we then get from Proposition 3.5 and Lemma 4.2 (noting the comments following it) that

$$P\left(J^{(1a)} \cup J^{(1c)}\right) \leq P\left(J^{(1a)} \setminus J^{(1c)}\right) + P\left(J^{(1c)}\right) \leq c_{21}|G_r^+|^4 e^{-C_{54}\alpha} + C_{68}e^{-t} \leq e^{-c_{22}t}. \tag{4.79}$$

Next, recalling (4.76), we observe that  $|\tilde{p}^i - \bar{p}^i| \leq q\sqrt{d-1}$  and provided  $r$  is large,

$$|u - v| \leq \frac{1}{32\mu} A_{j_1}^0(x^{i-1}, p^i) + q\sqrt{d-1} \implies h(|u - v|) \leq \frac{1}{16} A_{j_1}^0(x^{i-1}, p^i),$$

so, applying this to  $(u, v) = (p^i, \tilde{p}^i)$ , using (4.76) and the bound on  $t$  in (4.18) we have from Lemmas 3.1 and 3.2 that

$$\begin{aligned}
 & P\left(J^{(1b)} \setminus J^{(1c)}\right) \\
 & \leq P\left(\text{for some } u, v \in q\mathbb{Z}^d \cap G_r^+ \text{ and } A \geq \left(\frac{\lambda}{7}\right)^{j_1} t\sigma(r) \text{ we have} \right. \\
 & \quad \left. h(|u - v|) \leq \frac{A}{16} \text{ and } \hat{T}(u, v) \geq \frac{A}{8}\right) \\
 & \quad + P\left(\text{for some } u \in q\mathbb{Z}^d \cap G_r^+ \text{ we have } M(u) \geq \frac{1}{16} \left(\frac{\lambda}{7}\right)^{j_1} t\sigma(r)\right) \\
 & \leq c_{23}|G_r^+|^2 \exp\left(-c_{24} \left[\left(\frac{\lambda}{7}\right)^{j_1} t\sigma(r)\right]^{1-\chi_2}\right) + c_{25}|G_r^+| \exp\left(-c_{26} \left(\frac{\lambda}{7}\right)^{j_1} t\sigma(r)\right) \\
 & \leq c_{27}e^{-c_{28}t}. \tag{4.80}
 \end{aligned}$$

From (4.77), (4.79), and (4.80) the first union in (4.75) combined over  $(x, y)$ , together with  $J^{(1c)}$ , has probability bounded as

$$P\left(J^{(1a)} \cup J^{(1b)} \cup J^{(1c)}\right) \leq e^{-c_{22}t} + c_{27}e^{-c_{28}t},$$

and the same bound applies to the second union in (4.75), while from (4.18) and (4.71), for large  $c_{29}$  we have

$$P\left(J^{(0)}(c_{29})\right) \leq c_{30}e^{-c_{31}t}.$$

Hence from (4.73) we obtain

$$\begin{aligned}
 & P\left(T(x, y) \leq h((y - x)_1) - t\sigma(r) \text{ for some } (x, y) \in X_r\right) \\
 & \leq P\left(\Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1, m}) - h((y - x)_1) \leq -(1 - 2\lambda^{j_1}) t\sigma(r) + 4\delta\mu\lambda^{j_1} \left(\Upsilon_{Euc}(\Gamma_{xy}^{j_1, m}) - (y - x)_1\right) \right. \\
 & \quad \left. \text{for some } (x, y) \in X_r; \omega \notin J^{(0)} \cup J^{(1c)}\right) + c_{32}e^{-c_{33}t}. \tag{4.81}
 \end{aligned}$$

This completes the  $j_1$ th-scale (first) iteration.

It should be noted that, due to (4.73), the terms with coefficients 2 and 4 in (4.81) represent a reduction taken from the original bound  $t\sigma(r)$  in (1.14), used to bound errors created in the  $j_1$ th-scale iteration.

#### 4.6 Step 4. Further iterations of coarse-graining: preparation

For the  $(j_1 - 1)$ th-scale and later iterations of coarse-graining we use allocations  $A_j^1(\Gamma_{xy}^{cur}, u^i)$  associated not to a particular transition, but to the shifting of the marked grid point  $u^i$  in some  $(j + 1)$ th-scale current marked PG path  $\Gamma_{xy}^{cur}$  to the  $j$ th-scale grid. Specifically, in such a shift the marked grid point  $u^i$  in the  $(j + 1)$ th-scale grid is replaced by the  $j$ th-scale CG approximation  $V_j(u^i)$ , and the other marked grid points are left unchanged. Consider an initial marked PG path when an iteration of shifts to the  $j$ th-scale grid begins:

$$\Gamma^0 : u^0 \rightarrow u^1 \rightarrow \dots \rightarrow u^{n+1},$$

with  $n \leq 7^{j+1} - 1$ . Suppose that for some  $I \subset \{1, \dots, n\}$  not containing two consecutive integers, the marked grid points  $(u^i, i \in I)$  are the ones shifted, one at a time, in the iteration, updating the path from  $\Gamma^0$  via a sequence of intermediate paths  $\Gamma^1, \dots, \Gamma^{|I|-1}$

to a final  $\Gamma^I$ . Note that since  $I$  does not contain two consecutive integers, the shifts of different  $u^i$ 's do not "interact," as shifting  $u^i$  only affects the path between  $\varphi(u^{i-1})$  and  $\varphi(u^{i+1})$  and only affects  $u^i$  among the marks; we will refer to this aspect as *noninteraction of shifts*. The need for such noninteraction is a primary reason we must create the final CG path gradually over many iterations of coarsening.

Recall  $t^*(\cdot, \cdot)$  from (4.30). The allocation we use most widely, in both stages of each iteration, is  $A_j^1(\cdot, \cdot)$  which appears in the next lemma.  $A_j^2(\cdot, \cdot)$  is a variant of  $A_j^1(\cdot, \cdot)$ .

**Lemma 4.5.** *Let  $K \geq 1$  and  $1 \leq j \leq j_1$ . Consider a marked PG path*

$$\Gamma : u^0 \rightarrow u^1 \rightarrow \dots \rightarrow u^{n+1}$$

*in  $G_r^+$  with  $u^0, u^{n+1} \in G_r(K)$ ,  $n \leq 7^j - 1$ , all  $u^i \in \mathbb{L}_j$ , and  $(u^0)_1 < \dots < (u^{n+1})_1$ , and let  $I \subset \{1, \dots, n\}$ , not containing two consecutive integers. Define*

$$A_j^1(\Gamma, i) = \lambda^j \left[ \frac{1}{7^j} t^*(u^{i-1}, u^i) \sigma(r) + \frac{\delta\mu}{9} \left( \frac{|(u^i - u^{i-1})^*|^2}{|(u^i - u^{i-1})_1|} + \frac{|(u^{i+1} - u^i)^*|^2}{|(u^{i+1} - u^i)_1|} \right) \right],$$

*and for  $v, w \in \mathbb{L}_j$  with  $v_1 < w_1$ , let*

$$A_j^2(v, w) = \frac{1}{4} \lambda^j \left[ \frac{1}{7^j} t^*(v, w) \sigma(r) + \frac{\delta\mu}{9} \frac{|(w - v)^*|^2}{|(w - v)_1|} \right]. \tag{4.82}$$

*Then provided  $t/K^2$  is sufficiently large,*

$$\sum_{i \in I} A_j^1(\Gamma, i) \leq \frac{2}{3} \lambda^j \left[ t \sigma(r) + \delta\mu \left( \Upsilon_{Euc}(\Gamma) - (u^{n+1} - u^0)_1 \right) \right] \tag{4.83}$$

*and*

$$\sum_{i=1}^{n+1} A_j^2(u^{i-1}, u^i) \leq \frac{1}{4} \lambda^j \left[ t \sigma(r) + \delta\mu \left( \Upsilon_{Euc}(\Gamma) - (u^{n+1} - u^0)_1 \right) \right]. \tag{4.84}$$

The bound (4.83) would be trivial if we used  $t$  instead of  $t^*(u^{i-1}, u^i)$  the definition of  $A_j^1(\Gamma, i)$ . The point of Lemma 4.5 is that when the transitions affected by the shifting start or end at points  $u^i$  which are at distance a large multiple of  $\Delta(r)$  from  $G_r(K)$ , it increases the usable value of  $t$  by a multiple of  $(|(u^i)^*|/\Delta(r))^2$ , manifested in the definition of  $t^*(u^{i-1}, u^i)$ . Here by "usable" we mean "not so big that (4.83) or (4.84) fails."

*Proof of Lemma 4.5.* First note that, similarly to (4.25), every increment  $u^m - u^i$  with  $i < m$  is nearly horizontal; in particular the "whenever" in (4.26) applies to each increment.

If  $u^k \notin G_r(2K)$  for some  $k$  then  $d(u^k, \Pi_{u^0 u^{n+1}}) \geq |(u^k)^*|/2$  so we have from (4.26)

$$\begin{aligned} |u^{n+1} - u^0| + \frac{|(u^k)^*|^2}{3r} &\leq |u^0 - u^k| + |u^k - u^{n+1}| \\ &\leq \sum_{i=1}^{n+1} |u^i - u^{i-1}| \\ &\leq (u^{n+1} - u^0)_1 + \sum_{i=1}^{n+1} \frac{|(u^i - u^{i-1})^*|^2}{2|(u^i - u^{i-1})_1|}, \end{aligned} \tag{4.85}$$

so

$$\frac{\delta\mu |(u^k)^*|^2}{18r} \leq \sum_{i=1}^{n+1} \frac{\delta\mu |(u^i - u^{i-1})^*|^2}{12 |(u^i - u^{i-1})_1|}, \tag{4.86}$$



while if  $u^k \in G_r(2K)$ , since  $t/K^2$  is large,

$$\frac{\delta\mu|(u^k)^*|^2}{18r} \leq \frac{2\delta\mu K^2\sigma(r)}{9} \leq \frac{\delta t\sigma(r)}{3}, \tag{4.87}$$

so using (4.26),

$$\sum_{k=0}^n \frac{\delta\mu|(u^k)^*|^2}{18r7^j} \leq \frac{\delta t\sigma(r)}{3} + \sum_{i=1}^{n+1} \frac{\delta\mu|(u^i - u^{i-1})^*|^2}{12|(u^i - u^{i-1})_1|} \leq \frac{1}{3} \left[ \delta t\sigma(r) + \delta\mu \left( \Upsilon_{Euc}(\Gamma) - (u^{n+1} - u^0)_1 \right) \right]. \tag{4.88}$$

From (4.26) we also get

$$\frac{1}{3} \sum_{i \in I} \left( \frac{|(u^i - u^{i-1})^*|^2}{|(u^i - u^{i-1})_1|} + \frac{|(u^{i+1} - u^i)^*|^2}{|(u^{i+1} - u^i)_1|} \right) \leq \Upsilon_{Euc}(\Gamma) - (u^{n+1} - u^0)_1$$

which with (4.88) yields (4.83). The proof of (4.84) is similar, the only (inconsequential) difference being that when we sum the terms  $t^*(u^{i-1}, u^i)$  over all  $i \leq n$ , most terms  $\delta\mu|(u^k)^*|^2/18r7^j$  get counted twice, for  $k = i$  and  $k = i + 1$ , whereas each was counted at most once when we summed over  $i \in I$  in (4.83).  $\square$

**Lemma 4.6.** *There exist constants  $C_{70}, C_{71}$  such that, provided (4.20) holds and  $t$  is sufficiently large, we have for all  $j \leq j_1$*

$$\begin{aligned} &P\left(|\hat{T}(v, w) - h(|w - v|)| \geq A_j^2(v, w) \text{ for some } v, w \in G_r^+ \cap \mathbb{L}_j \right. \\ &\quad \left. \text{with } \delta^j r \leq (w - v)_1 \leq 10\delta^{j-1}r\right) \\ &\leq C_{70} \exp\left(-C_{71} \left(\frac{\lambda}{7\delta\chi_1}\right)^j t\right). \end{aligned} \tag{4.89}$$

*Proof.* Let

$$\begin{aligned} \nu_0 &= \nu_0(r) = \min\{\nu : (2^\nu t)^{1/2} \geq C_{60} \log r\}, \\ \ell_0 &= \ell_0(r, j) = \min\{\ell : 2^\ell \mathbb{L}_j K_0 \beta^j \geq 2C_{60} \log r\} \end{aligned}$$

for  $C_{60}$  from the definition (4.6) of  $G_r^+$ . We decompose the pairs  $(v, w)$  appearing in (4.89) into subclasses according to the size of  $|v^*|$  and the degree of sidestepping  $|(w - v)^*|$ , as follows. Fix  $1 \leq j \leq j_1$ . We use a variable  $\nu$  to index the transverse scale of the starting point  $v$  of the increment, and  $\ell$  to index the transverse scale of the increment  $w - v$ . Their maximum values  $\nu_0$  and  $\ell_0$  are the largest ones compatible with  $v, w \in G_r^+$ . There are only  $10/\delta$  possible values of the longitudinal increment  $(w - v)_1$  in (4.89) so we do not need to index scales for those increments. The value  $\ell = 0$  corresponds to transverse increments  $|(w - v)^*|$  not more than an order- $\rho^j$  factor larger than the typical scale- $j$  transverse increment  $\Delta(\delta^j r)$ . In detail, let

$$\begin{aligned} R_{r,j}^{\nu,\ell} &= \left\{ (v, w) \in (G_r^+ \cap \mathbb{L}_j)^2 : \delta^j r \leq (w - v)_1 \leq 10\delta^{j-1}r, (2^{\nu-1}t)^{1/2}\Delta(r) < |v^*| \leq (2^\nu t)^{1/2}\Delta(r), \right. \\ &\quad \left. 2^{\ell-1}\ell_2(j)K_0\beta^j\Delta(r) < |(w - v)^*| \leq 2^\ell\ell_2(j)K_0\beta^j\Delta(r) \right\}, \\ &\quad \text{defined for } 1 \leq \nu \leq \nu_0, 1 \leq \ell \leq \ell_0(r, j + 1), \\ R_{r,j}^{\nu,0} &= \left\{ (v, w) \in (G_r^+ \cap \mathbb{L}_j)^2 : \delta^j r \leq (w - v)_1 \leq 10\delta^{j-1}r, (2^{\nu-1}t)^{1/2}\Delta(r) < |v^*| \leq (2^\nu t)^{1/2}\Delta(r), \right. \\ &\quad \left. |(w - v)^*| \leq \ell_2(j)K_0\beta^j\Delta(r) \right\}, \quad \text{defined for } 1 \leq \nu \leq \nu_0, \\ R_{r,j}^{0,\ell} &= \left\{ (v, w) \in (G_r^+ \cap \mathbb{L}_j)^2 : \delta^j r \leq (w - v)_1 \leq 10\delta^{j-1}r, |v^*| \leq t^{1/2}\Delta(r), \right. \end{aligned}$$

$$2^{\ell-1} \ell_2(j) K_0 \beta^j \Delta(r) < |(w-v)^*| \leq 2^\ell \ell_2(j) K_0 \beta^j \Delta(r) \}, \quad 1 \leq \ell \leq \ell_0(r, j+1),$$

$$R_{r,j}^{0,0} = \left\{ (v, w) \in (G_r^+ \cap \mathbb{L}_j)^2 : \delta^j r \leq (w-v)_1 \leq 10\delta^{j-1}r, |v^*| \leq t^{1/2} \Delta(r), \right. \\ \left. |(w-v)^*| \leq \ell_2(j) K_0 \beta^j \Delta(r) \right\}.$$

First, for  $1 \leq \nu \leq \nu_0, 1 \leq \ell \leq \ell_0(r, j+1)$ , and  $(v, w) \in R_{r,j}^{\nu,\ell}$ , from the definition of  $\ell_2(j)$  we have

$$\frac{A_j^2(v, w)}{\sigma(10\delta^{j-1}r)} \geq \frac{1}{4} \lambda^j \left[ \frac{1}{7^j} \frac{\delta\mu}{36} 2^\nu t \frac{\sigma(r)}{\sigma(10\delta^{j-1}r)} + \frac{\delta\mu(2^\ell \ell_2(j) K_0 \beta^j \Delta(r))^2}{360\delta^{j-1}r\sigma(10\delta^{j-1}r)} \right] \\ \geq \lambda^j \left[ \frac{c_0}{7^j} 2^\nu t \frac{1}{\delta^{\chi_1 j}} + c_1 2^{2\ell} \rho^{2j} \right], \quad (v, w) \in R_{r,j}^{\nu,\ell}$$

and from (4.21)

$$|R_{r,j}^{\nu,\ell}| \leq \frac{c_2}{\delta^j} \left( \frac{(2^\nu t)^{1/2}}{\beta^j} \right)^{d-1} (2^\ell \ell_2(j))^{d-1} \leq c_3 (2^\nu t)^{(d-1)/2} 2^{(d-1)\ell} \frac{1}{\delta^j} \left( \frac{\rho\delta^{(1+\chi_1)/2}}{\beta^2} \right)^{(d-1)j}.$$

Hence by Lemma 3.2, provided  $t$  is large,

$$\sum_{\nu=1}^{\nu_0} \sum_{\ell=1}^{\ell_0} P\left(|\hat{T}(v, w) - h(|w-v|)| \geq A_j^2(v, w) \text{ for some } (v, w) \in R_{r,j}^{\nu,\ell}\right) \\ \leq c_3 \frac{1}{\delta^j} \left( \frac{\rho\delta^{(1+\chi_1)/2}}{\beta^2} \right)^{(d-1)j} \\ \cdot \sum_{\nu=1}^{\nu_0} (2^\nu t)^{(d-1)/2} \sum_{\ell=1}^{\ell_0} 2^{(d-1)\ell} C_{44} \exp\left(-C_{45} \lambda^j \left[ \frac{c_0}{7^j} 2^\nu t \frac{1}{\delta^{\chi_1 j}} + c_1 2^{2\ell} \rho^{2j} \right]\right) \\ \leq c_4 \frac{1}{\delta^j} \left( \frac{\rho\delta^{(1+\chi_1)/2}}{\beta^2} \right)^{(d-1)j} \sum_{\nu=1}^{\nu_0} (2^\nu t)^{(d-1)/2} \exp\left(-c_5 \left( \frac{\lambda}{7\delta^{\chi_1}} \right)^j 2^\nu t\right) \\ \leq c_4 \frac{1}{\delta^j} \left( \frac{\rho\delta^{(1+\chi_1)/2}}{\beta^2} \right)^{(d-1)j} \exp\left(-c_5 \left( \frac{\lambda}{7\delta^{\chi_1}} \right)^j t\right). \tag{4.90}$$

Second, for  $1 \leq \nu \leq \nu_0, \ell = 0$  we can drop the second term in brackets in (4.82) and the first term  $t/3$  in the definition (4.30) of  $t^*(v, w)$ :

$$\frac{A_j^2(v, w)}{\sigma(10\delta^{j-1}r)} \geq \frac{1}{4} \left( \frac{\lambda}{7} \right)^j \frac{\delta\mu}{36} 2^\nu t \frac{\sigma(r)}{\sigma(10\delta^{j-1}r)} \geq c_6 \left( \frac{\lambda}{7\delta^{\chi_1}} \right)^j 2^\nu t, \quad (v, w) \in R_{r,j}^{\nu,0}$$

and from (4.21)

$$|R_{r,j}^{\nu,0}| \leq \frac{c_7}{\delta^j} \left( \frac{(2^\nu t)^{1/2} \ell_2(j)}{\beta^j} \right)^{d-1} \leq c_8 (2^\nu t)^{(d-1)/2} \frac{1}{\delta^j} \left( \frac{\rho\delta^{(1+\chi_1)/2}}{\beta^2} \right)^{(d-1)j}.$$

Hence again using Lemma 3.2, provided  $t$  is large,

$$\sum_{\nu=1}^{\nu_0} P\left(|\hat{T}(v, w) - h(|w-v|)| \geq A_j^2(v, w) \text{ for some } (v, w) \in R_{r,j}^{\nu,0}\right) \\ \leq c_8 \frac{1}{\delta^j} \left( \frac{\rho\delta^{(1+\chi_1)/2}}{\beta^2} \right)^{(d-1)j} \sum_{\nu=1}^{\nu_0} (2^\nu t)^{(d-1)/2} \exp\left(-c_9 \left( \frac{\lambda}{7\delta^{\chi_1}} \right)^j 2^\nu t\right)$$

$$\leq c_8 \frac{1}{\delta^j} \left( \frac{\rho \delta^{(1+\chi_1)/2}}{\beta^2} \right)^{(d-1)j} \exp \left( -c_9 \left( \frac{\lambda}{7\delta^{\chi_1}} \right)^j t \right). \tag{4.91}$$

Third, for  $\nu = 0, 1 \leq \ell \leq \ell_0(r, j + 1)$ , using  $t^*(v, w) \geq t/3$  we have

$$\begin{aligned} \frac{A_j^2(v, w)}{\sigma(10\delta^{j-1}r)} &\geq \frac{1}{4} \lambda^j \left[ \frac{1}{7^j} \frac{t}{3} \frac{\sigma(r)}{\sigma(10\delta^{j-1}r)} + \frac{\delta \mu(2^\ell \ell_2(j) K_0 \beta^j \Delta(r))^2}{360 \delta^{j-1} r \sigma(10\delta^{j-1}r)} \right] \\ &\geq c_{10} \left( \frac{\lambda}{7\delta^{\chi_1}} \right)^j t + c_{11} 2^{2\ell} (\lambda \rho^2)^j, \quad (v, w) \in R_{r,j}^{0,\ell} \end{aligned}$$

and from (4.21)

$$\left| R_{r,j}^{0,\ell} \right| \leq \frac{c_{12}}{\delta^j} \left( \frac{t^{1/2}}{\beta^j} \right)^{d-1} (2^\ell \ell_2(j))^{d-1} \leq c_{12} t^{(d-1)/2} 2^{(d-1)\ell} \frac{1}{\delta^j} \left( \frac{\rho \delta^{(1+\chi_1)/2}}{\beta^2} \right)^{(d-1)j}.$$

so similarly to (4.91)

$$\begin{aligned} &\sum_{\ell=1}^{\ell_0} P \left( \left| \hat{T}(v, w) - h(|w - v|) \right| \geq A_j^2(v, w) \text{ for some } (v, w) \in R_{r,j}^{0,\ell} \right) \\ &\leq c_{13} t^{(d-1)/2} \frac{1}{\delta^j} \left( \frac{\rho \delta^{(1+\chi_1)/2}}{\beta^2} \right)^{(d-1)j} \sum_{\ell=1}^{\ell_0} 2^{(d-1)\ell} \exp \left( -c_{14} \left[ \left( \frac{\lambda}{7\delta^{\chi_1}} \right)^j t + 2^{2\ell} (\lambda \rho^2)^j \right] \right) \\ &\leq c_{13} \frac{1}{\delta^j} \left( \frac{\rho \delta^{(1+\chi_1)/2}}{\beta^2} \right)^{(d-1)j} \exp \left( -c_{14} \left( \frac{\lambda}{7\delta^{\chi_1}} \right)^j t \right). \end{aligned} \tag{4.92}$$

Finally, for  $k = \ell = 0$  we can drop the second term in brackets in (4.82) and use  $t^*(v, w) \geq t/3$ , and we have analogously

$$\frac{A_j^2(v, w)}{\sigma(10\delta^{j-1}r)} \geq \frac{1}{4} \left( \frac{\lambda}{7} \right)^j t \frac{\sigma(r)}{\sigma(10\delta^{j-1}r)} \geq c_{15} \left( \frac{\lambda}{7\delta^{\chi_1}} \right)^j t, \quad (v, w) \in R_{r,j}^{0,0},$$

and

$$\left| R_{r,j}^{0,0} \right| \leq \frac{c_{16}}{\delta^j} \left( \frac{t^{1/2} \ell_2(j)}{\beta^j} \right)^{d-1} \leq c_{17} t^{(d-1)/2} \frac{1}{\delta^j} \left( \frac{\rho \delta^{(1+\chi_1)/2}}{\beta^2} \right)^{(d-1)j}.$$

so

$$\begin{aligned} &P \left( \left| \hat{T}(v, w) - h(|w - v|) \right| \geq A_j^2(v, w) \text{ for some } (v, w) \in R_{r,j}^{0,0} \right) \\ &\leq c_{18} t^{(d-1)/2} \frac{1}{\delta^j} \left( \frac{\rho \delta^{(1+\chi_1)/2}}{\beta^2} \right)^{(d-1)j} \exp \left( -c_{19} \left( \frac{\lambda}{7\delta^{\chi_1}} \right)^j t \right) \\ &\leq c_{20} \frac{1}{\delta^j} \left( \frac{\rho \delta^{(1+\chi_1)/2}}{\beta^2} \right)^{(d-1)j} \exp \left( -c_{21} \left( \frac{\lambda}{7\delta^{\chi_1}} \right)^j t \right). \end{aligned} \tag{4.93}$$

Since  $t$  is large, (4.89) follows from (4.20) and (4.90)–(4.93). □

For  $j_2 \leq j < j_1$  and  $j_2 < \ell \leq j - 1$ , a  $(j, \ell)$ th-scale joining 4-path is a marked PG path  $\Gamma : v^0 \rightarrow v^1 \rightarrow v^2 \rightarrow v^3$  with  $v^i \in \mathbb{L}_{j+1} \cap G_r^+$ , such that for some  $j$ th-scale hyperplane  $H_a$ ,

$$v^0 \in H_a, \quad v^1 \in H_{a+\delta^{j+1}r}, \quad v^2 \in H_{a+\delta^\ell r}, \quad v^3 \in H_{a+\delta^\ell r}. \tag{4.94}$$

In a  $j$ th-scale interval with left endpoint  $a$ , these four hyperplanes represent the hyperplane at  $a$ , a sandwiching  $(j+1)$ th-scale hyperplane, and two possible joining hyperplanes.

The pairs  $(v^0, v^1), (v^1, v^2), (v^2, v^3)$  are the *links* of the joining 4-path. Let  $z^i$  be the point where  $\Pi_{v^0, v^3}$  intersects the hyperplane  $H$ , containing  $v^i$ . The *intermediate path corresponding to  $\Gamma$*  is  $\Gamma^{int} : z^0 \rightarrow z^1 \rightarrow z^2 \rightarrow z^3$ ; note these points are collinear. Figure 4 contains a  $(j, \ell)$ th-scale joining 4-path  $v^0 \rightarrow v^{j+1} \rightarrow v^{\ell+1} \rightarrow v^\ell$  and intermediate path  $v^0 \rightarrow g^{j+1} \rightarrow g^{\ell+1} \rightarrow v^\ell$ . Recalling Remark 4.3, we say the  $(j, \ell)$ th-scale joining 4-path is *internally bowed* if

$$\frac{|v^2 - z^2|^2}{\delta^{\ell+1}r} \geq 2^{j-\ell} \frac{\delta}{128\mu} \left(\frac{\lambda}{7}\right)^{j+1} \frac{\sigma(\delta^{\ell+1}r)}{\sigma(\delta^j r)} t^*(v^0)\sigma(r) \tag{4.95}$$

(compare to (4.46)) and

$$\frac{|v^1 - z^1|^2}{\delta^{j+1}r} \leq 2^{-(j-\ell-4)} \frac{\sigma(\delta^{j+1}r)}{\sigma(\delta^{\ell+1}r)} \frac{|v^2 - z^2|^2}{\delta^{\ell+1}r} \tag{4.96}$$

(compare to (4.47)) or equivalently

$$\left(\frac{|v^1 - z^1|}{\Delta(\delta^{j+1}r)}\right)^2 \leq 2^{-(j-\ell-4)} \left(\frac{|v^2 - z^2|}{\Delta(\delta^{\ell+1}r)}\right)^2. \tag{4.97}$$

As an example, in Figure 4 the joining 4-path is internally bowed if  $v_{\ell+1}$  is sufficiently far from  $g^{\ell+1}$ , and  $v^{j+1}$  is not too far from  $g^{j+1}$ , with “far” being expressed in terms of the extra distance of a path  $v^0 \rightarrow x \rightarrow v^\ell$  relative to  $v^0 \rightarrow v^\ell$ .

We define special allocations to deal with internally bowed paths. For each  $(j, \ell)$ th-scale joining 4-path  $\Gamma : v^0 \rightarrow v^1 \rightarrow v^2 \rightarrow v^3$  and intermediate path  $\Gamma^{int} : z^0 \rightarrow z^1 \rightarrow z^2 \rightarrow z^3$  let

$$A_j^3(\Gamma) = \frac{\delta\mu}{324} \left( \frac{|v^2 - z^2|^2}{\delta^{\ell+1}r} + \lambda^j \frac{|(v^3 - v^0)^*|^2}{\delta^\ell r} \right).$$

Note that by (4.94),  $\ell$  here is a function of  $\Gamma$ . We will need an analog of Lemma 4.6 which enables us to deal with joining 4-paths as single units. Normally for  $\ell$ th-scale links (length of order  $\delta^\ell r$ ), where the fluctuations of the passage time  $\hat{T}(v^{i-1}, v^i)$  are of order  $\sigma(\delta^\ell r)$ , we need  $\ell$ th-scale allocations (proportional to  $\lambda^\ell$ ) in the role of  $t\sigma(|u - v|)$  in Lemma 3.2 to get a good bound, but in the next lemma we are able to use  $j$ th-scale allocations  $A_j^3(\Gamma)$  in that role even for much longer  $\ell$ th-scale lengths, by taking advantage of bowedness.

**Lemma 4.7.** *There exist constants  $C_i$  as follows. Let  $j_2 \leq j < j_1$  and  $j_2 < \ell \leq j - 1$ .*

(i) *For every  $(j, \ell)$ th-scale joining 4-path  $\Gamma : v^0 \rightarrow v^1 \rightarrow v^2 \rightarrow v^3$  in  $G_r^+$ , and  $C_{56}$  from (3.25),*

$$6A_{j+1}^3(\Gamma) \leq \sum_{i=1}^3 A_{j+1}^2(v^{i-1}, v^i) + \frac{\delta\mu}{18} \left[ \Upsilon_{Euc}(\Gamma) - |v^3 - v^0| \right] - 6C_{56}. \tag{4.98}$$

(ii)

$$\begin{aligned} &P\left(\text{there exists an internally bowed } (j, \ell)\text{th-scale joining 4-path } \Gamma : v^0 \rightarrow v^1 \rightarrow v^2 \rightarrow v^3 \right. \\ &\quad \left. \text{in } G_r^+, \text{ and } 1 \leq i \leq 3, \text{ for which } |\hat{T}(v^{i-1}, v^i) - h(|v^i - v^{i-1}|)| \geq A_{j+1}^3(\Gamma)\right) \\ &\leq C_{72} \left(\frac{2}{\delta\chi_2+1}\right)^{j-\ell} \left(\frac{\delta}{\beta^2}\right)^j \exp\left(-C_{73} \left(\frac{\lambda}{\delta\chi_1}\right)^{j+1} 2^{j-\ell}t\right). \end{aligned} \tag{4.99}$$

(iii)

$$P\left(\text{there exists an internally bowed } (j, \ell)\text{th-scale joining 4-path } \Gamma \text{ in } G_r^+, \text{ with}\right.$$

corresponding intermediate path  $\Gamma^{int} : u^0 \rightarrow u^1 \rightarrow u^2 \rightarrow u^3$ ,  
 and  $1 \leq i \leq 3$ , for which  $|\hat{T}(u^{i-1}, u^i) - h(|u^i - u^{i-1}|)| \geq A_{j+1}^3(\Gamma^{int})$ )

$$\leq C_{72} \left( \frac{2}{\delta^{\chi_2+1}} \right)^{j-\ell} \left( \frac{\delta}{\beta^2} \right)^j \exp \left( -C_{73} \left( \frac{\lambda}{\delta^{\chi_1}} \right)^{j+1} 2^{j-\ell} t \right). \quad (4.100)$$

*Proof.* (i). Using (4.26) and  $t^*(\cdot, \cdot) \geq t/3$  we get

$$\begin{aligned} 6A_{j+1}^3(\Gamma) &\leq \frac{\delta\mu}{54} \left( 3 \left[ \Upsilon_{Euc}(v^0, v^2, v^3) - |v^3 - v^0| \right] + 3\lambda^j \left[ |v^3 - v^0| - (v^3 - v^0)_1 \right] \right) \\ &\leq \frac{\delta\mu}{18} \left( \left[ \Upsilon_{Euc}(\Gamma) - |v^3 - v^0| \right] + \frac{1}{2} \lambda^j \Psi(\Gamma) \right) \\ &\leq \frac{\delta\mu}{18} \left[ \Psi_{Euc}(\Gamma) - |v^3 - v^0| \right] + \sum_{i=1}^3 A_{j+1}^2(v^{i-1}, v^i) - \frac{1}{4} \left( \frac{\lambda}{7} \right)^j t\sigma(r) \\ &\leq \frac{\delta\mu}{18} \left[ \Psi_{Euc}(\Gamma) - |v^3 - v^0| \right] + \sum_{i=1}^3 A_{j+1}^2(v^{i-1}, v^i) - 6C_{56}, \end{aligned} \quad (4.101)$$

where the last inequality follows from  $j \leq j_1 = O(\log \log r)$ .

(ii) and (iii). The proof of (iii) is a slightly simplified version of the proof of (ii), so we only prove (ii). We proceed as in Lemma 4.6. We decompose the set of joining 4-paths according to the sizes of  $|(v^0)^*|$  and  $|(v^3 - v^0)^*|$ , and the degree of bowedness, measured by the left side of (4.95). From Remark 4.3, every internally bowed  $(j, \ell)$ th-scale joining 4-path  $\Gamma : v^0 \rightarrow v^1 \rightarrow v^2 \rightarrow v^3$  in  $G_r^+$  satisfies

$$\frac{|v^2 - z^2|^2}{\delta^{\ell+1}r} \geq \frac{\delta}{128\mu} \left( \frac{\lambda}{7} \right)^{j+1} t^*(v^0)\sigma(r) \frac{\sigma(\delta^{\ell+1}r)}{\sigma(\delta^j r)}. \quad (4.102)$$

Thus define for  $\nu, m_2, m_3 \geq 1$

$R_{r,j,\ell}(\nu, m_2, m_3)$

$$\begin{aligned} &= \left\{ \Gamma : v^0 \rightarrow v^1 \rightarrow v^2 \rightarrow v^3 \mid \Gamma \text{ is an internally bowed } (j, \ell)\text{th-scale joining 4-path in } G_r^+ \right. \\ &\quad \text{such that } (2^{\nu-1}t)^{1/2}\Delta(r) \leq |(v^0)^*| \leq (2^\nu t)^{1/2}\Delta(r), \\ &\quad 2^{m_2-1}2^{j-\ell} \frac{\delta}{128\mu} \left( \frac{\lambda}{7} \right)^{j+1} t^*(v^0)\sigma(r) \frac{\sigma(\delta^{\ell+1}r)}{\sigma(\delta^j r)} < \frac{|v^2 - z^2|^2}{\delta^{\ell+1}r} \\ &\quad \leq 2^{m_2}2^{j-\ell} \frac{\delta}{128\mu} \left( \frac{\lambda}{7} \right)^{j+1} t^*(v^0)\sigma(r) \frac{\sigma(\delta^{\ell+1}r)}{\sigma(\delta^j r)}, \text{ and} \\ &\quad \left. 2^{m_3-1}2^{j-\ell} t^*(v^0)\sigma(r) \frac{\sigma(\delta^\ell r)}{\sigma(\delta^{j+1}r)} < \frac{|(v^3 - v^0)^*|^2}{\delta^\ell r} \leq 2^{m_3}2^{j-\ell} t^*(v^0)\sigma(r) \frac{\sigma(\delta^\ell r)}{\sigma(\delta^{j+1}r)} \right\}. \end{aligned} \quad (4.103)$$

$R_{r,j,\ell}(0, m_2, m_3)$  is defined similarly with the first condition in (4.103) replaced by  $|(v^0)^*| \leq t^{1/2}\Delta(r)$ .  $R_{r,j,\ell}(\nu, m_2, 0)$  is defined similarly with the last condition in (4.103) replaced by

$$\frac{|(v^3 - v^0)^*|^2}{\delta^\ell r} \leq 2^{j-\ell} t^*(v^0)\sigma(r) \frac{\sigma(\delta^\ell r)}{\sigma(\delta^{j+1}r)}.$$

Every internally bowed joining 4-path in  $G_r^+$  is in one of these classes, since by (4.102) we don't need classes with  $m_2 = 0$ .

Suppose  $\Gamma : v^0 \rightarrow v^1 \rightarrow v^2 \rightarrow v^3$  is in  $R_{r,j,\ell}(\nu, m_2, m_3)$  for some  $\nu \geq 0, m_3 \geq 0$  and  $m_2 \geq 1$ . Using  $t^*(v^0) \geq c_0(|v_0^*|/\Delta(r))^2$  we obtain

$$\begin{aligned} \frac{A_{j+1}^3(\Gamma)}{\sigma(\delta^\ell r)} &\geq \frac{\delta\mu}{324} 2^{j-\ell} \frac{t^*(v^0)\sigma(r)}{\sigma(\delta^j r)} \left( \frac{\delta}{128\mu} \left(\frac{\lambda}{7}\right)^{j+1} 2^{m_2-1} \frac{\sigma(\delta^{\ell+1} r)}{\sigma(\delta^\ell r)} + \lambda^{j+1} 2^{m_3-1} \right) \\ &\geq c_1 \left(\frac{\lambda}{7\delta\chi_1}\right)^{j+1} 2^{j-\ell} (2^{m_2} + 2^{m_3}) 2^\nu t. \end{aligned} \tag{4.104}$$

Hence for each  $1 \leq i \leq 3$ , by Lemma 3.2,

$$\begin{aligned} P\left(|\hat{T}(v^{i-1}, v^i) - h(|v^i - v^{i-1}|)| \geq A_{j+1}^3(\Gamma)\right) &\leq C_{44} \exp\left(-C_{45} \frac{A_{j+1}^3(\Gamma)}{2\sigma(\delta^\ell r)}\right) \\ &\leq C_{44} \exp\left(-c_2 \left(\frac{\lambda}{7\delta\chi_1}\right)^{j+1} 2^{j-\ell} (2^{m_2} + 2^{m_3}) 2^\nu t\right). \end{aligned} \tag{4.105}$$

(Strictly speaking, when  $m_3 = 0$  we need to remove the terms involving  $2^{m_3}$  in (4.104) and (4.105), but this doesn't change the validity of the bound on the right in each case, possibly with different  $c_0, c_1$ .) Regarding the number of paths in  $R_{r,j,\ell}(\nu, m_2, m_3)$ , the number of possible  $v^0$  (necessarily in  $\mathbb{L}_{j+1}$ ) for such a path in a given  $j$ th-scale hyperplane is at most

$$2 \left(\frac{2(2^\nu t)^{1/2}}{K_0\beta^{j+1}}\right)^{d-1}.$$

The upper bound for  $|(v^0)^*|$  in (4.103) gives  $t^*(v^0) \leq c_3 2^\nu t$ , which with the last upper bound in (4.103) yields

$$\left(\frac{|(v^3 - v^0)^*|}{\Delta(r)}\right)^2 \leq c_4 2^{m_3} \left(\frac{2}{\delta\chi_2}\right)^{j-\ell} \delta^\ell 2^\nu t,$$

so for a given  $v^0$  the number of possible  $v^3$  is at most

$$2 \left(c_5 2^{m_3} \left(\frac{2}{\delta\chi_2}\right)^{j-\ell} \frac{1}{\beta^{2(j+1)}} \delta^\ell 2^\nu t\right)^{(d-1)/2} = 2 \left(\frac{c_5}{\beta^2} 2^{m_3} \left(\frac{2}{\delta\chi_2+1}\right)^{j-\ell} \left(\frac{\delta}{\beta^2}\right)^j 2^\nu t\right)^{(d-1)/2}.$$

The upper bound for  $|v^2 - z^2|/\delta^{\ell+1}r$  in (4.103) implies

$$\left(\frac{|(v^2 - z^2)|}{\Delta(r)}\right)^2 \leq c_6 2^{m_2} \left(\frac{2}{\delta\chi_2}\right)^{j-\ell} \delta^\ell 2^\nu t$$

and for given  $v^0, v^3$ , the point  $z^2$  is determined, and then the number of possible  $v^2$  is at most

$$2 \left(\frac{c_7}{\beta^2} 2^{m_2} \left(\frac{2}{\delta\chi_2+1}\right)^{j-\ell} \left(\frac{\delta}{\beta^2}\right)^j 2^\nu t\right)^{(d-1)/2}.$$

Combining these, we see that

$$|R_{r,j,\ell}(\nu, m_2, m_3)| \leq \frac{c_8}{\delta^j} \left(\frac{(2^\nu t)^{1/2}}{\beta^{j+1}}\right)^{d-1} \left(2^{m_2/2} 2^{m_3/2} \left(\frac{2}{\delta\chi_2+1}\right)^{j-\ell} \left(\frac{\delta}{\beta^2}\right)^j 2^\nu t\right)^{d-1}.$$

Multiplying this by (4.105) and summing over  $\nu, m_2, m_3$  gives (4.99). □

**4.7 Step 5. First stage of the  $(j_1 - 1)$ th-scale (second) iteration of coarse-graining: shifting to the  $(j_1 - 1)$ th-scale grid**

The current marked PG path at the start of the  $(j_1 - 1)$ th-scale iteration step is  $\Gamma_{xy}^{j_1, m}$ . We rename it now as

$$\Gamma_{xy}^{j_1-1, 0} : p^0 \rightarrow p^1 \rightarrow \dots \rightarrow p^m \rightarrow p^{m+1}.$$

We shift certain points to the  $(j_1 - 1)$ th-scale grid; the procedure is somewhat different from the  $j_1$ th-scale iteration step, as we are starting from a  $j_1$ th-scale marked PG path  $\Gamma_{xy}^{j_1-1, 0}$ . Fix  $(x, y) \in X_r$  and, recalling  $\mathcal{H}_{xy} = \{H_{s_i}, 1 \leq i \leq m\}$ , let

$$I_{xy} = \left\{ i \in \{1, \dots, m\} : H_{s_i} \text{ is a non-incident } (j_1 - 1)\text{-scale hyperplane in } \mathcal{H}_{xy} \right\},$$

then relabel  $\{s_i : i \in I_{xy} \cup \{1, m\}\}$  as  $a_1 < \dots < a_n$ , so  $H_{a_1} = H_{s_1}$  and  $H_{a_n} = H_{s_m}$  are the terminal  $j_1$ th-scale hyperplanes, and define indices  $\gamma(N)$  by  $a_N = s_{\gamma(N)}$ , so the  $p^{\gamma(N)}, 2 \leq N \leq n - 1$ , are the marked points in non-incident  $(j_1 - 1)$ th-scale hyperplanes. Observe that every  $(j_1 - 1)$ th-scale interval, including terminal ones, has at least one  $j_1$ th-scale hyperplane in its interior; hence  $\gamma(N + 1) \geq \gamma(N) + 2$  for all  $1 \leq N < n$ .

For each non-incident  $(j_1 - 1)$ th-scale hyperplane in  $\mathcal{H}_{xy}$  let  $b^N = V_{j_1-1}(p^{\gamma(N)})$ . Also let  $b^0 = p^0 = \hat{x}, b^1 = p^1, b^n = p^m$ , and  $b^{n+1} = p^{m+1} = \hat{y}$ . We shift to the  $(j_1 - 1)$ th scale grid in each non-incident  $(j_1 - 1)$ th-scale hyperplane  $H_{s_i}, i \in I_{xy}$ , replacing  $p^{\gamma(N)}$  with  $b^N$  for  $2 \leq N \leq n - 1$ , to create the updated path which we denote  $\Gamma_{xy}^{j_1-1, 1}$ . Letting  $\hat{p}^i = m_{s_i}(\Gamma_{xy}^{j_1-1, 1})$  (so  $\hat{p}^i = p^i$  if  $i \notin I_{xy}, \hat{p}^i = b^{\gamma^{-1}(i)}$  if  $i \in I_{xy}$ ), we may equivalently write this as

$$\Gamma_{xy}^{j_1-1, 1} : \hat{p}^0 \rightarrow \hat{p}^1 \rightarrow \hat{p}^2 \rightarrow \dots \rightarrow \hat{p}^{m+1}.$$

The black path from  $u'$  to  $v'$  in Figure 6 illustrates a segment of such a path, with  $u' = b^{N-1} = \hat{p}^{\gamma(N-1)}, v' = b^N = \hat{p}^{\gamma(N)}$  for some  $N$  there; the vertices between  $u'$  and  $v'$  are points  $p^i = \hat{p}^i$  with  $\gamma(N - 1) < i < \gamma(N)$ . Viewing  $\Gamma_{xy}^{j_1-1, 0}$  as a CG approximation, on the  $j_1$ th scale, of the original  $\Gamma_{xy}$ , we may view  $\Gamma_{xy}^{j_1-1, 1}$  as a coarsening of the approximation to the  $(j_1 - 1)$ th scale, but only coarsened in the non-incident  $(j_1 - 1)$ th-scale hyperplanes. In the second stage of the iteration we will remove marked points in those non-terminal  $j_1$ th-scale hyperplanes where coarsening did not occur, to create the path

$$\Gamma_{xy}^{j_1-1, 2} : b^0 \rightarrow b^1 \rightarrow \dots \rightarrow b^n \rightarrow b^{n+1}$$

which is entirely a CG approximation on the  $(j_1 - 1)$ th scale.

By Lemma 4.5 we have

$$\sum_{i \in I_{xy}} A_{j_1}^1(\Gamma_{xy}^{j_1-1, 0}, i) \leq \frac{2}{3} \lambda^{j_1} \left[ t\sigma(r) + \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1, 0}) - (\hat{y} - \hat{x})_1 \right) \right], \quad (4.106)$$

while similarly to (4.66) (noting the comment after (4.68)),

$$\left| \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1, 1}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1, 0}) \right| \leq \frac{\lambda^{j_1}}{2\mu} \left[ \frac{t\sigma(r)}{3} + \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1, 0}) - (\hat{y} - \hat{x})_1 \right) \right] \quad (4.107)$$

and

$$\left| \Upsilon_h(\Gamma_{xy}^{j_1-1, 1}) - \Upsilon_h(\Gamma_{xy}^{j_1-1, 0}) \right| \leq \lambda^{j_1} \left[ \frac{t\sigma(r)}{3} + \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1, 0}) - (\hat{y} - \hat{x})_1 \right) \right]. \quad (4.108)$$

(The only notable change from the derivation of (4.66) and (4.68) is that now, since we are shifting to the  $(j_1 - 1)$ th scale, we have the bound

$$|(p^i - \hat{p}^i)^*| \leq \sqrt{d-1} K_0 \beta^{j_1-1} \Delta(r),$$

larger only by a constant  $\beta^{-1}$  compared to the analogous bound in (4.51)–(4.52). Then (4.106) and (4.108) give

$$\sum_{i \in I_{xy}} A_{j_1}^1(\Gamma_{xy}^{j_1-1,0}, i) + \Upsilon_h(\Gamma_{xy}^{j_1-1,1}) - \Upsilon_h(\Gamma_{xy}^{j_1-1,0}) \leq \lambda^{j_1} \left[ t\sigma(r) + \frac{5}{3} \delta \mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y} - \hat{x})_1 \right) \right]. \quad (4.109)$$

From (4.107) and (4.109), analogously to (4.73), we have for the event on the right in (4.81) that

$$\begin{aligned} & \left\{ \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,0}) - h((y-x)_1) \leq -(1-2\lambda^{j_1})t\sigma(r) + 4\delta\mu\lambda^{j_1} \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y} - \hat{x})_1 \right) \right\} \\ & \subset \left\{ \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,1}) - h((y-x)_1) \leq -(1-4\lambda^{j_1})t\sigma(r) + 6\delta\mu\lambda^{j_1} \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y} - \hat{x})_1 \right) \right. \\ & \quad \left. - \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,1}) \right) \right\} \\ & \cup \left\{ \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,0}) - \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,1}) \leq \Upsilon_h(\Gamma_{xy}^{j_1-1,0}) - \Upsilon_h(\Gamma_{xy}^{j_1-1,1}) - \sum_{i \in I_{xy}} A_{j_1}^1(\Gamma_{xy}^{j_1-1,0}, i) \right\}. \end{aligned} \quad (4.110)$$

**Remark 4.8.** To see the meaning of (4.110), let us momentarily rename the events there as  $M_1, M_2, M_3$  (left to right) and rename the right side of the inequality in each event as  $-\tau\sigma(r) + z_1, -\tau\sigma(r) + z_1 - z_2$ , and  $z_3$ , respectively. We have  $z_1 > 0$ ;  $z_2, z_3$  may be positive or negative. The validity of  $M_1 \subset M_2 \cup M_3$  requires that  $z_2 \leq z_3$ , which follows from (4.109). The completed shifting to the grid may increase or reduce the  $h$ -length  $\Upsilon_h(\cdot)$ , its approximation  $\mu\Upsilon_{Euc}(\cdot)$ , and/or the passage time  $\Upsilon_{\hat{T}}(\cdot)$  of the path. Suppose that the shifting reduces the  $h$ -length. The event  $M_3$  says that the passage time fails to *track* (i.e. keep up with) this reduction, to within an allowed error given by the sum of allocations in  $M_3$ . If we can choose those allocations so that  $M_3$  is unlikely, we can have  $z_2 > 0$  (or at least nearly so, as reflected in the increase in coefficients from 2, 4 to 4, 6 going from  $M_1$  to  $M_2$ ), which means the right side in  $M_2$  improves on (i.e. is smaller than) the right side in  $M_1$ , by an amount involving a multiple  $\delta$  of the path length reduction.

If we can successfully do this for all iterations, for both the shifting-to-the-grid and point-removal stages, then in the left side of (4.81) we have effectively replaced  $\Gamma_{xy}$  with the final CG path  $\Gamma_{xy}^{CG}$  (i.e. replaced  $T(x, y)$  with  $\Upsilon_{\hat{T}}(\Gamma_{xy}^{CG})$ ) and improved  $-\tau\sigma(r)$  to a significantly more negative value, incorporating a positive fraction of the total path-length reduction from all iterations. It is this improvement of  $-\tau\sigma(r)$  that will enable us to sum the probability of an unusually fast  $\Upsilon_{\hat{T}}(\Gamma_{xy}^{CG})$  over all possible  $\Gamma_{xy}^{CG}$ .

Informally we refer to  $M_3$ , and similar events, as *tracking failure*, and say we *establish tracking* when we show the probability of tracking failure is small. *Tracking error* means the difference between the increment of  $\Upsilon_{\hat{T}}$  and the increment of  $\Upsilon_h$ , as in  $M_3$ , for either a path or a single link.

In the shifting-to-the-grid stage, (4.108) says that the change in  $h$ -length is small, which makes it relatively easy to establish tracking—the primary part of  $z_3$  is the sum of allocations. But for the point-removal stage, the (always positive) reduction in  $h$ -length may be large, and establishing tracking becomes harder. Further, in (4.81) and in  $M_1$ ,  $2\lambda^{j_1}t\sigma(r) + 4\delta\mu\lambda^{j_1}(\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y} - \hat{x})_1)$  represents roughly the accumulated error allocations used in the completed  $j_1$ th-scale iteration; the increase in coefficients from 2, 4 to 4, 6 going from  $M_1$  to  $M_2$  represents additional accumulated error allocations used



in the shifting-to-the-grid stage of the  $(j_1 - 1)$ th-scale iteration. If these allocations are too large, then the improvement  $z_2$  may be too small, or even negative.

We slightly change (4.110) in the point-removal stage (see (4.113) and (4.116))—in  $M_3$  we use  $\mu\Upsilon_{Euc}$  in place of  $\Upsilon_h$  and a different term in place of the sum of allocations, but the principle is the same.

We note that failure to track is a one-sided phenomenon—we only care about failure of  $\Upsilon_{\hat{T}}(\cdot)$  to track decreases in  $\Upsilon_h(\cdot)$  created by shifting to the grid, not increases.

The primary tool to establish tracking is Lemma 4.6, which tells us that with high probability, there are *no* relevant links anywhere in  $G_r^+$  for which tracking error exceeds a certain allocation, as in our next step.

Using Lemma 4.6 we have for the last event in (4.110)

$$\begin{aligned}
 &P\left(\Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,0}) - \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,1}) \leq \Upsilon_h(\Gamma_{xy}^{j_1-1,0}) - \Upsilon_h(\Gamma_{xy}^{j_1-1,1}) - \sum_{i \in I_{xy}} A_{j_1}^1(\Gamma_{xy}^{j_1-1,0}, i) \right. \\
 &\quad \left. \text{for some } (x, y) \in X_r; \omega \notin J^{(0)} \cup J^{(1c)}\right) \\
 &\leq P\left(\max\left(|\hat{T}(\hat{p}^{i-1}, \hat{p}^i) - h(|\hat{p}^i - \hat{p}^{i-1}|)|, |\hat{T}(p^{i-1}, p^i) - h(|p^i - p^{i-1}|)|\right) \right. \\
 &\quad \left. \geq \frac{1}{4} A_{j_1}^1(\Gamma_{xy}^{j_1-1,0}, i) \text{ for some } 1 \leq i \leq m+1 \text{ and } (x, y) \in X_r \right. \\
 &\quad \left. \text{with } \{i-1, i\} \cap I_{xy} \neq \emptyset; \omega \notin J^{(0)} \cup J^{(1c)}\right) \\
 &\leq P\left(|\hat{T}(v, w) - h(|w - v|)| \geq A_{j_1}^2(v, w) \right. \\
 &\quad \left. \text{for some } j_1\text{-th-scale transition } v \rightarrow w \text{ with } v, w \in G_r^+ \cap \mathbb{L}_{j_1}\right) \\
 &\leq C_{70} \exp\left(-C_{71} \left(\frac{\lambda}{7\delta\chi_1}\right)^{j_1} t\right), \tag{4.111}
 \end{aligned}$$

establishing tracking as desired. Combining this with (4.81) and (4.110) yields

$$\begin{aligned}
 &P\left(T(x, y) \leq h((y - x)_1) - t\sigma(r) \text{ for some } (x, y) \in X_r\right) \\
 &\leq P\left(\Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,1}) - h((y - x)_1) \leq -(1 - 4\lambda^{j_1})t\sigma(r) + 6\delta\mu\lambda^{j_1} \left(\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y} - \hat{x})_1\right) \right. \\
 &\quad \left. - \delta\mu \left(\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,1})\right) \text{ for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)}\right) \\
 &\quad + c_{32}e^{-c_{33}t} + C_{70} \exp\left(-C_{71} \left(\frac{\lambda}{7\delta\chi_1}\right)^{j_1} t\right). \tag{4.112}
 \end{aligned}$$

**4.8 Step 6. Second stage of the  $(j_1 - 1)$ th-scale (second) coarse-graining iteration: removing marked points**

For the next update of the current marked PG path  $\Gamma_{xy}^{j_1-1,1}$ , we remove all the marked points in non-terminal  $j_1$ th-scale hyperplanes to create the updated marked PG path

$$\Gamma_{xy}^{j_1-1,2} : b^0 \rightarrow b^1 \rightarrow \dots \rightarrow b^n \rightarrow b^{n+1}.$$

Here we recall that  $b^0 = \hat{x}$ ,  $b^{n+1} = \hat{y}$ ,  $b^1$  and  $b^n$  lie in terminal  $j_1$ th-scale hyperplanes, and  $b^2, \dots, b^{n-1}$  lie in  $(j_1 - 1)$ th-scale hyperplanes. As with  $p^i, \hat{p}^i$ , etc.,  $n$  and  $b^i$  should be viewed as functions of  $x, y, \omega$ .

In this second stage, the removal of marked points always reduces the Euclidean length of the path (see Figure 5), meaning  $\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,1}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,2}) \geq 0$ , and on average the reduction in passage time,  $\Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,1}) - \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,2})$ , should be about  $\mu$  times the reduction in length. Here a tracking failure means the actual reduction in passage time is at most  $\delta\mu$  times the reduction in length, to within an allocated error; see the last probability in (4.113).

From (4.107) we have

$$\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,2}) - (\hat{y} - \hat{x})_1 \leq \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,1}) - (\hat{y} - \hat{x})_1 \leq \frac{1}{2}t\sigma(r) + \frac{3}{2}(\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y} - \hat{x})_1)$$

and therefore, assuming  $\delta$  is small,

$$\begin{aligned} & \frac{1}{3} \left[ 2t\sigma(r) + \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,1}) - (\hat{y} - \hat{x})_1 + \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,2}) - (\hat{y} - \hat{x})_1 \right) \right] \\ & \leq t\sigma(r) + \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y} - \hat{x})_1 \right). \end{aligned}$$

From this and (4.112) we obtain the setup to establish tracking:

$$\begin{aligned} & P\left(T(x, y) \leq h((y-x)_1) - t\sigma(r) \text{ for some } (x, y) \in X_r\right) \\ & \leq P\left(\Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,1}) - h((y-x)_1) \leq -(1-4\lambda^{j_1})t\sigma(r) + 6\delta\mu\lambda^{j_1} \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y} - \hat{x})_1 \right) \right. \\ & \quad \left. - \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,1}) \right) \text{ for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)}\right) \\ & \quad + c_{32}e^{-c_{33}t} + C_{70} \exp\left(-C_{71} \left(\frac{\lambda}{7\delta\chi_1}\right)^{j_1} t\right) \\ & \leq P\left(\Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,2}) - h((y-x)_1) \leq -(1-5\lambda^{j_1})t\sigma(r) + 7\delta\mu\lambda^{j_1} \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y} - \hat{x})_1 \right) \right. \\ & \quad \left. - \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,2}) \right) \text{ for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)}\right) \\ & \quad + P\left(\Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,1}) - \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,2}) \right. \\ & \quad \left. \leq -\frac{1}{3}\lambda^{j_1} \left[ 2t\sigma(r) + \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,1}) - (\hat{y} - \hat{x})_1 + \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,2}) - (\hat{y} - \hat{x})_1 \right) \right] \right. \\ & \quad \left. + \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,1}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,2}) \right) \right. \\ & \quad \left. \text{for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)}\right) \\ & \quad + c_{32}e^{-c_{33}t} + C_{70} \exp\left(-C_{71} \left(\frac{\lambda}{7\delta\chi_1}\right)^{j_1} t\right). \end{aligned} \tag{4.113}$$

Let us consider the contribution to the difference of sums  $\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,1}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,2})$  (which is nonnegative), appearing in the last probability in (4.113), from a single  $(j_1 - 1)$ th-scale interval  $I_N = [b^N, b^{N+1}] = [\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N+1)}]$ . Removing the marked points from the two hyperplanes in the interior of  $I_N$  changes the marked PG path from the *full path*

$$\Gamma^{N,full} : \hat{p}^{\gamma(N)} \rightarrow \hat{p}^{\gamma(N)+1} \rightarrow \hat{p}^{\gamma(N)+2} \rightarrow \hat{p}^{\gamma(N+1)}$$

to the *direct path*  $\hat{p}^{\gamma(N)} \rightarrow \hat{p}^{\gamma(N+1)}$  (that is,  $b^N \rightarrow b^{N+1}$ .) We then have

$$\Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,1}) - \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,2}) = \sum_{N=1}^{n+1} \left[ \Upsilon_{\hat{T}}(\Gamma^{N,full}) - \hat{T}(\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N+1)}) \right]. \tag{4.114}$$

Given  $n \geq 1$  and  $\Gamma : u_1 \rightarrow \dots \rightarrow u_n$  with (for some  $j$ ) all  $u_i$  in the grid  $\mathbb{L}_j$ , we define

$$S_j(u_0, \dots, u_n) = S_j(\Gamma) = \sum_{i=1}^n A_j^2(u_{i-1}, u_i).$$

From Lemma 4.5 we have

$$\sum_{N=0}^n S_{j_1}(\Gamma^{N,full}) = \sum_{i=1}^{m+1} A_{j_1}^2(\hat{p}^{i-1}, \hat{p}^i) \leq \frac{1}{4} \lambda^j \left[ t\sigma(r) + \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,1}) - (\hat{y} - \hat{x})_1 \right) \right], \tag{4.115}$$

so the last probability in (4.113) is bounded above by

$$\begin{aligned} P \left( \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,1}) - \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,2}) \leq \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,1}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,2}) \right) \right. \\ \left. - \sum_{N=0}^n S_{j_1}(\Gamma^{N,full}) - \frac{1}{3} \lambda^{j_1} \left[ t\sigma(r) + \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,2}) - (\hat{y} - \hat{x})_1 \right) \right] \right. \\ \left. - \frac{1}{12} \lambda^{j_1} t\sigma(r) \text{ for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \right). \end{aligned} \tag{4.116}$$

This is the tracking-failure event (see Remark 4.4) for the marked-point-removal stage of the iteration, and our main task is to bound its probability. The last of the 3 terms on the right inside the probability can be viewed as part of the allocation of allowed errors. As noted in Remark 4.8, the quantity

$$4\lambda^{j_1} t\sigma(r) + 6\delta\mu \lambda^{j_1} \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y} - \hat{x})_1 \right)$$

in the second probability in (4.113) represents the accumulated error allocations used in the 1.5 iterations completed so far; the allocations for the present stage increase the 4 and 6 to 5 and 7 in the third probability in (4.113). Our ability to bound the tracking-failure event is what allows us to replace the quantity

$$\delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,1}) \right) \quad \text{with} \quad \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,2}) \right)$$

in that second probability in (4.113) to obtain the third probability. We need each iteration to involve similar such replacement, so that when the iterations are complete this term becomes

$$\delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - \Upsilon_{Euc}(\Gamma_{xy}^{CG}) \right)$$

which is typically positive and can in part cancel the accumulated error allocations (see (4.167)).

Let  $\hat{w}^i$  be the point  $\Pi_{\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N+1)}} \cap H_{(\hat{p}^i)_1}$  (see Figure 7); note  $\hat{w}^i$  does not necessarily lie in any  $j$ th-scale grid. Let  $w_\perp^i$  be the orthogonal projection of  $\hat{p}^i$  into  $\Pi_{\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N+1)}}$ , noting that by (4.25),  $|w_\perp^i - \hat{w}^i|$  is much smaller than  $|\hat{p}^i - \hat{w}^i|$ . (Note the indexing of marked points differs here from that used in Step 1 in defining  $L^-(I)$ . Our  $\hat{w}^i$  here has index  $i$  matching that of the point  $\hat{p}^i$  in the hyperplane, whereas  $w^\ell$  in Step 1 has index corresponding to the distance  $\delta^\ell r$  from the left end of the interval.)

We continue considering the contribution to the difference of sums  $\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,1}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,2})$  in (4.116) from a single  $(j_1 - 1)$ th-scale interval  $I_N = [b^N, b^{N+1}] = [\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N+1)}]$ . As noted after (4.22), the number (call it  $m$ ) of  $(j_1 - 1)$ th-scale intervals in  $\mathcal{H}_{xy}$  satisfies  $m \leq \delta^{-j_1}$ . We split into cases according to the type of interval  $I_N$ .

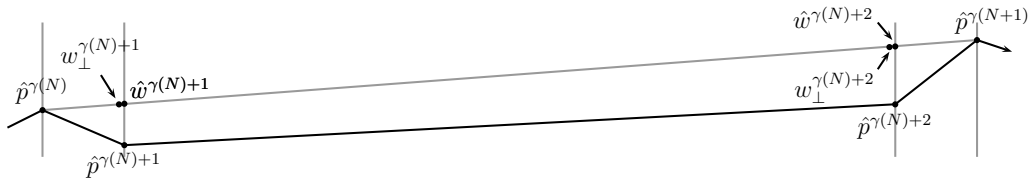


Figure 7: Diagram for Case 1 showing  $(j_1 - 1)$ th-scale endpoint hyperplanes and  $j_1$ th-scale sandwiching hyperplanes in a short  $(j_1 - 1)$ th-scale interval. The black path is  $\Gamma^{N,full}$  and the gray is  $\Gamma^{N,int}$ .

Case 1.  $I_N$  is a short non-terminal  $(j_1 - 1)$ th-scale interval. Here  $\mathcal{H}_{xy}$  includes no joining hyperplanes in the interval, so it includes exactly two maximally  $j_1$ th-scale (sandwiching) hyperplanes there, at distance  $\delta^{j_1}r$  from each end; see Figure 7. We introduce the *intermediate path*

$$\Gamma^{N,int} : w_{\perp}^{\gamma(N)} \rightarrow w_{\perp}^{\gamma(N)+1} \rightarrow w_{\perp}^{\gamma(N)+2} \rightarrow w_{\perp}^{\gamma(N+1)}$$

which has the same endpoints and satisfies  $\Upsilon_{\hat{T}}(\Gamma^{N,int}) \geq \hat{T}(\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N+1)})$ . To bound (4.116) we use the expression on the right in (4.114). Applying Lemma 3.6(ii) with  $\epsilon = 1 - \delta$  yields

$$h(|\hat{p}^i - \hat{p}^{i-1}|) - h(|w_{\perp}^i - w_{\perp}^{i-1}|) \geq \delta\mu (|\hat{p}^i - \hat{p}^{i-1}| - |w_{\perp}^i - w_{\perp}^{i-1}|) - C_{56}$$

and therefore

$$\Upsilon_h(\Gamma^{N,full}) - \Upsilon_h(\Gamma^{N,int}) \geq \delta\mu \left( \Upsilon_{Euc}(\Gamma^{N,full}) - |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| \right) - 3C_{56}. \quad (4.117)$$

This is the essential property of the intermediate path: when we look at the bowedness of the full path relative to the intermediate path (represented by the left side of (4.117)), and also relative to the direct path (right side of (4.117), without  $\delta$ ), the first is at least  $\delta$  fraction of the second, to within a constant. By (4.70), for  $\omega \notin J^{(0)}(c_{29})$  the intermediate path also satisfies

$$\hat{T}(\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N+1)}) \leq \Upsilon_{\hat{T}}(\Gamma^{N,int}) + (\gamma(N+1) - \gamma(N) - 1)c_{29} \log r. \quad (4.118)$$

Similarly to (4.115), since in Case 1  $\Upsilon_{Euc}(\Gamma^{N,int}) = |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| = |b^{N+1} - b^N|$ , we have

$$\sum_{N=0}^n S_{j_1}(\Gamma^{N,int}) = \sum_{i=1}^{m+1} A_{j_1}^2(w_{\perp}^{i-1}, w_{\perp}^i) \leq \frac{1}{4} \lambda^{j_1} \left[ t\sigma(r) + \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,2}) - (\hat{y} - \hat{x})_1 \right) \right]. \quad (4.119)$$

Using (4.22) we have

$$\frac{\lambda^{j_1} t\sigma(r)}{6m} \geq (\lambda\delta)^{j_1} \sigma(r) \geq 2c_{29} \log r + 2C_{56}. \quad (4.120)$$

From this and (4.114)–(4.119) it follows that the contribution to the tracking-failure probability (4.116) from short non-terminal intervals is bounded by

$$P \left( \hat{T}(\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N+1)}) - \Upsilon_{\hat{T}}(\Gamma^{N,full}) \geq S_{j_1}(\Gamma^{N,full}) + S_{j_1}(\Gamma^{N,int}) + \frac{\lambda^{j_1} t\sigma(r)}{6m} + \delta\mu \left( |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| - \Upsilon_{Euc}(\Gamma^{N,full}) \right) \text{ for some } N \text{ with } I_N \text{ short non-terminal } (j_1 - 1)\text{th-scale, and some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1e)} \right)$$

$$\begin{aligned}
 &\leq P\left(\Upsilon_{\hat{T}}(\Gamma^{N,int}) - \Upsilon_{\hat{T}}(\Gamma^{N,full}) \geq S_{j_1}(\Gamma^{N,full}) + S_{j_1}(\Gamma^{N,int}) + \frac{\lambda^{j_1} t \sigma(r)}{6m} - 2c_{29} \log r \right. \\
 &\quad \left. + \Upsilon_h(\Gamma^{N,int}) - \Upsilon_h(\Gamma^{N,full}) - 2C_{56} \text{ for some } N \text{ with } I_N \text{ short non-terminal} \right. \\
 &\quad \left. (j_1 - 1)\text{th-scale, and some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)}\right) \\
 &\leq P\left(|\Upsilon_{\hat{T}}(\Gamma^{N,full}) - \Upsilon_h(\Gamma^{N,full})| \geq S_{j_1}(\Gamma^{N,full}) \text{ or } |\Upsilon_{\hat{T}}(\Gamma^{N,int}) - \Upsilon_h(\Gamma^{N,int})| \right. \\
 &\quad \left. \geq S_{j_1}(\Gamma^{N,int}) \text{ for some } N \text{ with } I_N \text{ short non-terminal} \right. \\
 &\quad \left. (j_1 - 1)\text{th-scale, and some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)}\right) \\
 &\leq P\left(\left|\hat{T}(\hat{p}^{i-1}, \hat{p}^i) - h(|\hat{p}^i - \hat{p}^{i-1}|)\right| \geq A_{j_1}^2(\hat{p}^{i-1}, \hat{p}^i) \text{ or } \left|\hat{T}(w_{\perp}^{i-1}, w_{\perp}^i) - h(|w_{\perp}^i - w_{\perp}^{i-1}|)\right| \right. \\
 &\quad \left. \geq A_{j_1}^2(w_{\perp}^{i-1}, w_{\perp}^i) \text{ for some } \gamma(N) + 1 \leq i \leq \gamma(N + 1), \text{ for some } N \text{ with } I_N \text{ short} \right. \\
 &\quad \left. \text{non-terminal } (j_1 - 1)\text{th-scale, and some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)}\right).
 \end{aligned} \tag{4.121}$$

Since the corresponding intervals are short, all of the increments  $\hat{T}(u, v)$  in the last event have  $\delta^{j_1} r \leq (v - u)_1 \leq 10\delta^{j_1-1} r$ . We can therefore bound the last probability similarly to Lemma 4.6, with the main difference being that in place of pairs  $(u, v)$  in the definitions of  $R_{r,j}^{*,*}$  we need to consider 4-tuples  $(u, v, w, z)$  corresponding to values of  $(\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N)+1}, \hat{p}^{\gamma(N)+2}, \hat{p}^{\gamma(N+1)}) \in (\mathbb{L}_{j_1} \cap G_r^+)^4$ . This means the exponents  $d - 1$  in the bounds on  $|R_{r,j}^{*,*}|$  become  $3(d - 1)$ . The values  $w_{\perp}^i$  are determined by  $(\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N)+1}, \hat{p}^{\gamma(N)+2}, \hat{p}^{\gamma(N+1)})$  so their presence does not increase the necessary size of  $R_{r,j}^{*,*}$  in the lemma. As with the sets  $R_{r,j}^{\nu,\ell}$  in the lemma proof, we can decompose the possible 4-tuples  $(\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N)+1}, \hat{p}^{\gamma(N)+2}, \hat{p}^{\gamma(N+1)})$  according to the size of  $|\hat{p}^{\gamma(N)+1} - \hat{p}^{\gamma(N)}|$ ,  $|\hat{p}^{\gamma(N)+1} - w_{\perp}^{\gamma(N)+1}|$ , and  $|\hat{p}^{\gamma(N)+2} - w_{\perp}^{\gamma(N)+2}|$  and sum over the possible size ranges. Otherwise the proof remains the same, and we get that the last probability in (4.121) is bounded by

$$c_{34} \exp\left(-c_{35} \left(\frac{\lambda}{7\delta^{\chi_1}}\right)^{j_1} t\right). \tag{4.122}$$

**Case 2.**  $I_N$  is a terminal  $(j_1 - 1)$ th-scale interval (meaning either  $[\hat{p}^{\gamma(1)}, \hat{p}^{\gamma(2)}]$  or  $[\hat{p}^{\gamma(n-1)}, \hat{p}^{\gamma(n)}]$ ). The proof is similar to Case 1, except that the interval includes only one maximally  $j_1$ th-scale (sandwiching) hyperplane between the two terminal hyperplanes that are at the ends of the interval, so the full path in the interval has form  $\hat{p}^{\gamma(N)} \rightarrow \hat{p}^{\gamma(N)+1} \rightarrow \hat{p}^{\gamma(N+1)}$ . We obtain for the terminal-interval contribution to the tracking-failure probability (4.116) the bound

$$\begin{aligned}
 &P\left(\hat{T}(\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N+1)}) - \Upsilon_{\hat{T}}(\Gamma^{N,full}) \geq S_{j_1}(\Gamma^{N,full}) + S_{j_1}(\Gamma^{N,int}) \right. \\
 &\quad \left. + \delta\mu\left(|\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| - \Upsilon_{Euc}(\Gamma^{N,full})\right) \text{ for some } N \text{ with } I_N \text{ terminal} \right. \\
 &\quad \left. (j_1 - 1)\text{th-scale, and some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)}\right)
 \end{aligned}$$

$$\leq c_{34} \exp \left( -c_{35} \left( \frac{\lambda}{7\delta\chi_1} \right)^{j_1} t \right). \tag{4.123}$$

Case 3.  $I_N$  is a long non-terminal  $(j_1 - 1)$ th-scale interval. Such an interval, and thus also the middle increment  $\hat{p}^{\gamma(N+1)-1} - \hat{p}^{\gamma(N)+1}$  of the 3 comprising  $\Gamma^{N,full}$  in Case 1, may be much longer than  $\delta^{j_1-1}r$ . Therefore the quantity  $A_{j_1}^2(\hat{p}^{\gamma(N)+1}, \hat{p}^{\gamma(N+1)-1})$  used on the right side of (4.121) for that increment is no longer large enough to give a useful bound on the probability. To avoid this problem we will sometimes use a different intermediate path in  $I_N$  which coincides with the full path between the inner joining hyperplanes, so differs from the full path only near the ends of  $I_N$ , while preserving the property (4.117) (this preservation being the purpose of our choice of  $L^\pm(I)$ .) Equation (4.117) represents what we may informally call *deterministic tracking*, a nonrandom analog of the tracking of Remark 4.4 which facilitates our desired (random) form of tracking.

Fix a long non-terminal  $(j_1 - 1)$ th-scale interval  $I_N$ , and suppose it has  $k$ th-scale length for some  $k < j_1 - 1$ . The hyperplanes of  $\mathcal{H}_{xy}$  in  $I_N$  are the  $(j_1 - 1)$ th-scale ones at each endpoint, two sandwiching  $j_1$ th-scale hyperplanes at distance  $\delta^{j_1}r$  from each end, and between 2 and 4  $j_1$ th-scale joining hyperplanes between these. These hyperplanes are at the joining points  $(\hat{p}^{\gamma(N)})_1 + \delta^{L^-(I)+1}r$ ,  $(\hat{p}^{\gamma(N)})_1 + \delta^{L^-(I)}r$ ,  $(\hat{p}^{\gamma(N+1)})_1 - \delta^{L^+(I)}r$ , and  $(\hat{p}^{\gamma(N+1)})_1 - \delta^{L^+(I)+1}r$ , with two exceptions. First, if  $L^-(I_N) = j_1 - 1$  then the first of these 4 joining hyperplanes coincides with the left-end sandwiching one, and similarly for  $L^+(I_N)$ , which reduces the number of  $j_1$ th-scale joining hyperplanes to fewer than 4, as discussed in criterion (ii) after (4.39). In Figure 8,  $L^-(I_N) = j_1 - 1$  means the middle two hyperplanes coincide.) Second, in the totally unbowed case (third option in (4.39)) at either end of  $I_N$ , there is no outer joining hyperplane at that end, as in criterion (iii). We define  $4 \leq V(N) \leq 6$  to be the number of  $j_1$ th-scale hyperplanes in the interior of  $I_N$ .

Case 3a. The bowed case (second option in (4.39)) for both  $L^\pm(I)$ , with  $V(N) = 6$ . Since  $V(N) = 6$  there are two joining hyperplanes at each end of  $I_N$ , and we must have  $L^-(I_N) < j_1 - 1$  and  $L^+(I_N) < j_1 - 1$ . Here the full path is

$$\Gamma^{N,full} : \hat{p}^{\gamma(N)} \rightarrow \hat{p}^{\gamma(N)+1} \rightarrow \dots \rightarrow \hat{p}^{\gamma(N)+6} \rightarrow \hat{p}^{\gamma(N+1)}$$

and the direct path again is  $\hat{p}^{\gamma(N)} \rightarrow \hat{p}^{\gamma(N+1)}$ . Define

$$\hat{z}^i = \begin{cases} \text{the point } \Pi_{\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N)+3}} \cap H_{(\hat{p}^i)_1} & \text{if } i = \gamma(N) + 1, \gamma(N) + 2, \\ \text{the point } \Pi_{\hat{p}^{\gamma(N)+4}, \hat{p}^{\gamma(N+1)}} \cap H_{(\hat{p}^i)_1} & \text{if } i = \gamma(N) + 5, \gamma(N) + 6, \\ \hat{p}^i & \text{if } i = \gamma(N), \gamma(N) + 3, \gamma(N) + 4, \gamma(N) + 1; \end{cases}$$

see Figure 8. This time the intermediate path is defined as

$$\Gamma^{N,int} : \hat{z}^{\gamma(N)} \rightarrow \hat{z}^{\gamma(N)+1} \rightarrow \dots \rightarrow \hat{z}^{\gamma(N)+6} \rightarrow \hat{z}^{\gamma(N+1)},$$

which coincides with  $\Gamma^{N,full}$  between the inner joining hyperplanes (i.e. from  $\hat{p}^{\gamma(N)+3}$  to  $\hat{p}^{\gamma(N)+4}$ ), which is “most” of  $I_N$ . We denote the parts of the paths outside the inner joining hyperplanes by

$$\begin{aligned} \Gamma^{N,full,-} &: \hat{p}^{\gamma(N)} \rightarrow \hat{p}^{\gamma(N)+1} \rightarrow \hat{p}^{\gamma(N)+2} \rightarrow \hat{p}^{\gamma(N)+3}, \\ \Gamma^{N,full,+} &: \hat{p}^{\gamma(N)+4} \rightarrow \hat{p}^{\gamma(N)+5} \rightarrow \hat{p}^{\gamma(N)+6} \rightarrow \hat{p}^{\gamma(N+1)}, \end{aligned}$$

with  $\Gamma^{N,int,\pm}$  defined similarly with  $\hat{z}^i$  in place of  $\hat{p}^i$ , so that

$$\begin{aligned} \Upsilon_{\hat{T}}(\Gamma^{N,full}) - \Upsilon_{\hat{T}}(\Gamma^{N,int}) &= \left[ \Upsilon_{\hat{T}}(\Gamma^{N,full,-}) - \Upsilon_{\hat{T}}(\Gamma^{N,int,-}) \right] \\ &\quad + \left[ \Upsilon_{\hat{T}}(\Gamma^{N,full,+}) - \Upsilon_{\hat{T}}(\Gamma^{N,int,+}) \right]. \end{aligned} \tag{4.124}$$

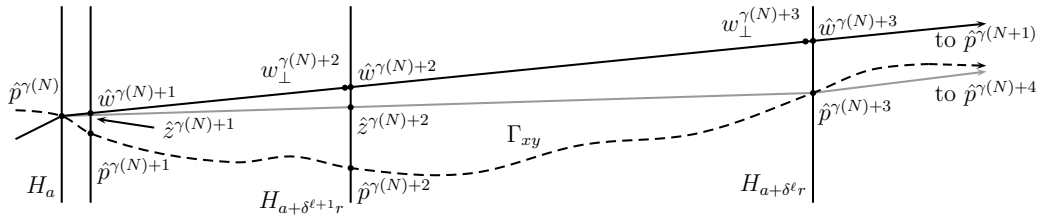


Figure 8: The left end of a long  $(j_1 - 1)$ th-scale interval  $I_N = [a, b]$  with the bowed case, with  $L^-(I_n) = \ell$ , showing the direct path (black) and intermediate path (gray.) The full path has marked points at the  $\hat{p}^{\gamma(N)+i}$ .

For the corresponding quantities  $\Upsilon_h(\cdot)$ , similarly to (4.117), but using  $\epsilon = 1/2$  in Lemma 3.6(ii), we have

$$\begin{aligned} & \Upsilon_h(\Gamma^{N,full}) - \Upsilon_h(\Gamma^{N,int}) \\ &= \left[ \Upsilon_h(\Gamma^{N,full,-}) - \Upsilon_h(\Gamma^{N,int,-}) \right] + \left[ \Upsilon_h(\Gamma^{N,full,+}) - \Upsilon_h(\Gamma^{N,int,+}) \right] \\ &\geq \frac{\mu}{2} \left[ \Upsilon_{Euc}(\Gamma^{N,full,-}) - |\hat{p}^{\gamma(N)+3} - \hat{p}^{\gamma(N)}| \right] \\ &\quad + \frac{\mu}{2} \left[ \Upsilon_{Euc}(\Gamma^{N,full,+}) - |\hat{p}^{\gamma(N)+1} - \hat{p}^{\gamma(N)+4}| \right] - 6C_{56} \\ &= \frac{\mu}{2} \left[ \Upsilon_{Euc}(\Gamma^{N,full}) - \Upsilon_{Euc}(\Gamma^{N,int}) \right] - 6C_{56}. \end{aligned} \tag{4.125}$$

We claim that deterministic tracking holds in the sense that

$$\Upsilon_{Euc}(\Gamma^{N,full}) - \Upsilon_{Euc}(\Gamma^{N,int}) \geq 3\delta \left[ \Upsilon_{Euc}(\Gamma^{N,full}) - |\hat{p}^{\gamma(N)+1} - \hat{p}^{\gamma(N)}| \right]. \tag{4.126}$$

This says that at least  $3\delta$  fraction of the excess length in  $\Gamma^{N,full}$  is not present in  $\Gamma^{N,int}$ ; it comes instead from the bowedness of  $\Gamma^{N,full,-}$  and/or  $\Gamma^{N,full,+}$ . (4.125) and (4.126) give the full analog of (4.117) for Case 3a. To prove the claim, suppose  $L^-(I_N) = \ell^-$  for some  $k + 1 \leq \ell^- < j_1 - 1$ , the first inequality being a property of the bowed case, from (4.39). Recall  $\alpha(\cdot)$  and  $\theta(\cdot)$  from (4.31) and (4.32), noting that in our present notation,  $v, w, z$  there have become  $\hat{p}, \hat{w}, \hat{z}$ , with different superscript labeling. Using (4.34)–(4.36) with  $j = j_1 - 1$ ,

$$\alpha(\ell^- + 1) = \frac{|\hat{p}^{\gamma(N)+2} - \hat{w}^{\gamma(N)+2}|^2}{\delta^{\ell^-+1}r}, \quad \theta(\ell^- + 1) = \frac{|\hat{z}^{\gamma(N)+2} - \hat{w}^{\gamma(N)+2}|}{|\hat{p}^{\gamma(N)+2} - \hat{w}^{\gamma(N)+2}|} \leq \frac{1}{2}$$

and hence

$$\begin{aligned} \Upsilon_{Euc}(\Gamma^{N,full,-}) - |\hat{p}^{\gamma(N)+3} - \hat{p}^{\gamma(N)}| &\geq \Upsilon_{Euc}(\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N)+2}, \hat{p}^{\gamma(N)+3}) - |\hat{p}^{\gamma(N)+3} - \hat{p}^{\gamma(N)}| \\ &\geq \frac{|\hat{p}^{\gamma(N)+2} - \hat{z}^{\gamma(N)+2}|^2}{3\delta^{\ell^-+1}r} \\ &\geq \frac{\alpha(\ell^- + 1)}{12}. \end{aligned} \tag{4.127}$$

Here the second inequality uses (4.26) and the fact that  $\hat{p}^{\gamma(N)+2} - \hat{z}^{\gamma(N)+2}$  is nearly perpendicular to  $\Pi_{\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N)+3}}$ , by (4.25). In the other direction, recalling  $I_N$  has  $k$ th-scale length and supposing  $L^+(I_N) = \ell^+$ , since  $k + 1 \leq \ell^+$  and  $k + 1 \leq \ell^-$  we have  $(\hat{p}^{\gamma(N)+3} - \hat{p}^{\gamma(N)})_1 \leq \delta^{k+1}r$  and  $(\hat{p}^{\gamma(N)+1} - \hat{p}^{\gamma(N)+4})_1 \leq \delta^{k+1}r$ , so  $(\hat{p}^{\gamma(N)+4} - \hat{p}^{\gamma(N)+3})_1 \geq 8\delta^{k+1}r$ . Hence similarly to (4.41) and (4.42)

$$\Upsilon_{Euc}(\Gamma^{N,int}) - |\hat{p}^{\gamma(N)+1} - \hat{p}^{\gamma(N)}| \leq \frac{|\hat{p}^{\gamma(N)+3} - w^{\gamma(N)+3}|^2}{2\delta^{\ell^-}r} + \frac{|\hat{p}^{\gamma(N)+4} - w^{\gamma(N)+4}|^2}{2\delta^{\ell^+}r}$$

$$\begin{aligned}
 & + \frac{|(\hat{p}^{\gamma(N)+3} - w_{\perp}^{\gamma(N)+3}) - (\hat{p}^{\gamma(N)+4} - w_{\perp}^{\gamma(N)+4})|^2}{16\delta^{k+1}r} \\
 & \leq \frac{|\hat{p}^{\gamma(N)+3} - \hat{w}^{\gamma(N)+3}|^2}{\delta^{\ell^-}r} + \frac{|\hat{p}^{\gamma(N)+4} - \hat{w}^{\gamma(N)+4}|^2}{\delta^{\ell^+}r}.
 \end{aligned}
 \tag{4.128}$$

Now the first ratio on the right is  $\alpha(\ell^-)$ , so it follows from (4.36) (last two inequalities) and (4.127) that

$$\frac{|\hat{p}^{\gamma(N)+3} - \hat{w}^{\gamma(N)+3}|^2}{\delta^{\ell^-}r} \leq 24C_{23}\delta^{-\chi_2} \left[ \Upsilon_{Euc}(\Gamma^{N,full,-}) - |\hat{p}^{\gamma(N)+3} - \hat{p}^{\gamma(N)}| \right],$$

and similarly

$$\frac{|\hat{p}^{\gamma(N)+4} - \hat{w}^{\gamma(N)+4}|^2}{\delta^{\ell^-}r} \leq 24C_{23}\delta^{-\chi_2} \left[ \Upsilon_{Euc}(\Gamma^{N,full,+}) - |\hat{p}^{\gamma(N)+4} - \hat{p}^{\gamma(N)}| \right].$$

Hence from (4.128) and the second equality in (4.26),

$$\Upsilon_{Euc}(\Gamma^{N,int}) - |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| \leq 24C_{23}\delta^{-\chi_2} \left[ \Upsilon_{Euc}(\Gamma^{N,full}) - \Upsilon_{Euc}(\Gamma^{N,int}) \right] \tag{4.129}$$

which implies

$$\Upsilon_{Euc}(\Gamma^{N,full}) - |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| \leq (1 + 24C_{23}\delta^{-\chi_2}) \left[ \Upsilon_{Euc}(\Gamma^{N,full}) - \Upsilon_{Euc}(\Gamma^{N,int}) \right], \tag{4.130}$$

proving the claim (4.126) when we take  $\delta$  small.

From Lemma 4.7(i) we have

$$6A_{j_1}^3(\Gamma^{N,full,-}) \leq S_{j_1}(\Gamma^{N,full,-}) + \frac{\delta\mu}{18} \left[ \Upsilon_{Euc}(\Gamma^{N,full,-}) - |\hat{p}^{\gamma(N)+3} - \hat{p}^{\gamma(N)}| \right] - 6C_{56}, \tag{4.131}$$

and the equivalent “mirror image” of Lemma 4.7 covers  $\Gamma^{N,full,+}$  symmetrically, incorporating a symmetric definition of joining 4-path.

Inequality (4.118) remains valid, and we now have the ingredients for the analog of (4.121), bounding the bowed-case contribution to the tracking-failure probability (4.116)—see (4.133) below:

$$\begin{aligned}
 & P \left( \Upsilon_{\hat{T}}(\Gamma^{N,full}) - \hat{T}(\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N+1)}) \leq -S_{j_1}(\Gamma^{N,full,-}) - S_{j_1}(\Gamma^{N,full,+}) - \frac{\lambda^{j_1}t\sigma(r)}{3m} \right. \\
 & \quad \left. + \delta\mu \left( \Upsilon_{Euc}(\Gamma^{N,full}) - |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| \right) \text{ for some } N \text{ with } V(N) = 6 \right. \\
 & \quad \left. \text{and } I_N \text{ long non-terminal } (j_1 - 1)\text{-th-scale, in the bowed case of (4.39)} \right. \\
 & \quad \left. \text{for both of } L^{\pm}(I_N), \text{ for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \right) \\
 & \leq P \left( \left[ \Upsilon_{\hat{T}}(\Gamma^{N,full,-}) - \Upsilon_{\hat{T}}(\Gamma^{N,int,-}) \right] + \left[ \Upsilon_{\hat{T}}(\Gamma^{N,full,+}) - \Upsilon_{\hat{T}}(\Gamma^{N,int,+}) \right] \right. \\
 & \quad \leq -6A_{j_1}^3(\Gamma^{N,full,-}) - 6A_{j_1}^3(\Gamma^{N,full,+}) + \Upsilon_h(\Gamma^{N,full}) - \Upsilon_h(\Gamma^{N,int}) \\
 & \quad \left. + 6c_{29} \log r - \frac{\lambda^{j_1}t\sigma(r)}{3m} \text{ for some } N \text{ with } V(N) = 6 \text{ and } I_N \text{ long non-terminal} \right. \\
 & \quad \left. (j_1 - 1)\text{-th-scale, in the bowed case of (4.39) for both of } L^{\pm}(I_N), \right. \\
 & \quad \left. \text{and for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \right)
 \end{aligned}$$



$$\leq \sum_{\ell^- = k+1}^{j_1-2} \sum_{\ell^+ = k+1}^{j_1-2} P \left( \left| \hat{T}(\hat{p}^{i-1}, \hat{p}^i) - h(|\hat{p}^i - \hat{p}^{i-1}|) \right| \geq A_{j_1}^3(\Gamma^{N,full,\pm}) \text{ or} \right. \\ \left. \left| \hat{T}(\hat{z}^{i-1}, \hat{z}^i) - h(|\hat{z}^i - \hat{z}^{i-1}|) \right| \geq A_{j_1}^3(\Gamma^{N,full,\pm}) \text{ for some } \gamma(N) + 1 \leq i \leq \gamma(N + 1) \right. \\ \text{with } i \neq \gamma(N) + 4, \text{ for some } N \text{ with } V(N) = 6 \text{ and } I_N \text{ long non-terminal} \\ \text{(} j_1 - 1 \text{)th-scale with } L^-(I_N) = \ell^- \text{ and } L^+(I_N) = \ell^+, \text{ in the bowed case of} \\ \left. (4.39) \text{ for both of } L^\pm(I_N), \text{ and some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \right). \quad (4.132)$$

Here the second inequality uses a slight modification of (4.120), and the first inequality uses (4.118), (4.124), and

$$\hat{T}(\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N+1)}) \leq \Upsilon_{\hat{T}}(\Gamma^{N,int}) + 6c_{29} \log r,$$

and that from (4.131), (4.126), and then (4.125),

$$\begin{aligned} & -S_{j_1}(\Gamma^{N,full,-}) - S_{j_1}(\Gamma^{N,full,+}) + \delta\mu \left( \Upsilon_{Euc}(\Gamma^{N,full}) - |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| \right) \\ & \leq -6A_{j_1}^3(\Gamma^{N,full,-}) - 6A_{j_1}^3(\Gamma^{N,full,+}) \\ & \quad + \frac{10}{9}\delta\mu \left( \Upsilon_{Euc}(\Gamma^{N,full}) - |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| \right) - 12C_{56} \\ & \leq -6A_{j_1}^3(\Gamma^{N,full,-}) - 6A_{j_1}^3(\Gamma^{N,full,+}) + \frac{10}{27}\mu \left[ \Upsilon_{Euc}(\Gamma^{N,full}) - \Upsilon_{Euc}(\Gamma^{N,int}) \right] - 12C_{56} \\ & \leq -6A_{j_1}^3(\Gamma^{N,full,-}) - 6A_{j_1}^3(\Gamma^{N,full,+}) + \Upsilon_h(\Gamma^{N,full}) - \Upsilon_h(\Gamma^{N,int}). \end{aligned} \quad (4.133)$$

For the  $\pm$  symbol in the last event in (4.132), the  $-$  applies to  $\gamma(N) + 1 \leq i \leq \gamma(N) + 3$  and the  $+$  applies to  $\gamma(N) + 5 \leq i \leq \gamma(N + 1)$ . An application of Lemma 4.7(ii) and (iii) with  $j = j_1 - 1$ , followed by summing over  $\ell^\pm$ , bounds the last (tracking-failure) probability in (4.132) by

$$\begin{aligned} & \left[ \sum_{\ell=1}^{j_1-2} C_{72} \left( \frac{2}{\delta\chi_2+1} \right)^{j_1-1-\ell} \left( \frac{\delta}{\beta^2} \right)^{j_1-1} \exp \left( -C_{73} \left( \frac{\lambda}{\delta\chi_1} \right)^{j_1-1} 2^{j_1-1-\ell} t \right) \right]^2 \\ & \leq c_{36} \left( \frac{\delta}{\beta^2} \right)^{2(j_1-1)} \exp \left( -C_{73} \left( \frac{\lambda}{\delta\chi_1} \right)^{j_1-1} t \right). \end{aligned} \quad (4.134)$$

**Case 3b.** The bowed case (second option in (4.39)) for both  $L^\pm(I)$ , with  $V(N) < 6$ . This means we have at least one of  $L^\pm(I_N) = j_1 - 1$ . If for example  $L^-(I_N) = j_1 - 1$ , then what were in Case 3a the two points  $\hat{p}^{\gamma(N)+1}, \hat{p}^{\gamma(N)+2}$  are now the same point, so effectively there is no separate point  $\hat{p}^{\gamma(N)+1}$ . Instead we have only the equivalent of  $\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N)+2}, \hat{p}^{\gamma(N)+3}$ , so joining 4-paths become joining 3-paths. This has no significant effect on the arguments, including Lemma 4.7, other than some simplifications, and the bound in (4.134) still applies for the corresponding contribution to the tracking-failure probability (4.116):

$$\begin{aligned} & P \left( \hat{T}(\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N+1)}) - \Upsilon_{\hat{T}}(\Gamma^{N,full}) \geq S_{j_1}(\Gamma^{N,full,-}) + S_{j_1}(\Gamma^{N,full,+}) \right. \\ & \quad \left. + \delta\mu \left( |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| - \Upsilon_{Euc}(\Gamma^{N,full}) \right) \text{ for some } N \text{ with } V(N) < 6 \right. \\ & \quad \left. \text{and } I_N \text{ long non-terminal (} j_1 - 1 \text{)th-scale, in the bowed case of (4.39)} \right. \\ & \quad \left. \text{for both of } L^\pm(I_N), \text{ and for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \right) \end{aligned}$$

$$\leq c_{36} \left(\frac{\delta}{\beta^2}\right)^{2(j_1-1)} \exp\left(-C_{73} \left(\frac{\lambda}{\delta x_1}\right)^{j_1-1} t\right). \tag{4.135}$$

Case 3c. The forward case (first option in (4.39)) for both  $L^\pm(I_N)$ . Here we have  $V(N) = 4$  as there are only inner joining points, at distance  $\delta^{j_1}r$  from each end of  $I_N$ . As before the direct path is  $\hat{p}^{\gamma(N)} \rightarrow \hat{p}^{\gamma(N+1)}$ , and the full path is

$$\Gamma^{N,full} : \hat{p}^{\gamma(N)} \rightarrow \hat{p}^{\gamma(N)+1} \rightarrow \dots \rightarrow \hat{p}^{\gamma(N)+4} \rightarrow \hat{p}^{\gamma(N+1)};$$

by (4.70), for  $\omega \notin J^{(0)}(c_{29})$  we have

$$\Upsilon_{\hat{T}}(\Gamma^{N,full}) - \hat{T}(\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N+1)}) \geq -4c_{29} \log r. \tag{4.136}$$

We claim that in the forward case we have

$$S_{j_1}(\Gamma^{N,full}) \geq \delta\mu \left(\Upsilon_{Euc}(\Gamma^{N,full}) - |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}|\right) + 4c_{29} \log r. \tag{4.137}$$

By (4.136), this means that in the forward case there is no tracking failure:

$$\begin{aligned} P\left(\Upsilon_{\hat{T}}(\Gamma^{N,full}) - \hat{T}(\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N+1)}) \leq -S_{j_1}(\Gamma^{N,full}) \right. \\ \left. + \delta\mu \left(\Upsilon_{Euc}(\Gamma^{N,full}) - |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}|\right) \text{ for some } N \text{ with } \right. \\ \left. I_N \text{ long non-terminal } (j_1 - 1)\text{th-scale, in the forward case of (4.39)} \right. \\ \left. \text{for both of } L^\pm(I_N), \text{ and for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)}\right) = 0. \tag{4.138} \end{aligned}$$

To prove (4.137), we first observe that from the definitions (4.30) and (4.82),

$$S_{j_1}(\Gamma^{N,full}) \geq \frac{1}{4} \left(\frac{\lambda}{7}\right)^{j_1} \left(t^*(\hat{p}^{\gamma(N)}) + t^*(\hat{p}^{\gamma(N+1)})\right) \sigma(r). \tag{4.139}$$

From (4.26) and (4.39), since  $I_N$  long entails  $k < j_1 - 1$ ,

$$\begin{aligned} &\Upsilon_{Euc}(\Gamma^{N,full}) - |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| \\ &\leq \frac{1}{2} \sum_{i=\gamma(N)+1}^{\gamma(N+1)} \frac{|(\hat{p}^i - w_\perp^i) - (\hat{p}^{i-1} - w_\perp^{i-1})|^2}{|w_\perp^i - w_\perp^{i-1}|} \\ &\leq \frac{1}{2} \left[ \frac{|\hat{p}^{\gamma(N)+1} - w_\perp^{\gamma(N)+1}|^2}{\delta^{j_1}r} + \frac{(|\hat{p}^{\gamma(N)+1} - w_\perp^{\gamma(N)+1}| + |\hat{p}^{\gamma(N)+2} - w_\perp^{\gamma(N)+2}|)^2}{(1-\delta)\delta^{j_1-1}r} \right. \\ &\quad \left. + \frac{(|\hat{p}^{\gamma(N)+2} - w_\perp^{\gamma(N)+2}| + |\hat{p}^{\gamma(N)+3} - w_\perp^{\gamma(N)+3}|)^2}{10\delta^{k+1}r - 2\delta^{j_1-1}r} \right. \\ &\quad \left. + \frac{(|\hat{p}^{\gamma(N)+3} - w_\perp^{\gamma(N)+3}| + |\hat{p}^{\gamma(N)+4} - w_\perp^{\gamma(N)+4}|)^2}{(1-\delta)\delta^{j_1-1}r} + \frac{|\hat{p}^{\gamma(N)+4} - w_\perp^{\gamma(N)+4}|^2}{\delta^{j_1}r} \right] \\ &\leq \frac{|\hat{p}^{\gamma(N)+1} - w_\perp^{\gamma(N)+1}|^2}{\delta^{j_1}r} + \frac{|\hat{p}^{\gamma(N)+4} - w_\perp^{\gamma(N)+4}|^2}{\delta^{j_1}r} + \frac{9}{\delta^{j_1-1}r} \max_{1 \leq i \leq 4} |\hat{p}^{\gamma(N)+i} - w_\perp^{\gamma(N)+i}|^2 \\ &\leq \frac{1+9\delta}{16\mu} \left(\frac{\lambda}{7}\right)^{j_1} \left(t^*(\hat{p}^{\gamma(N)}) + t^*(\hat{p}^{\gamma(N+1)})\right) \sigma(r). \tag{4.140} \end{aligned}$$

Since (4.22) ensures the right side of (4.139) is much larger than  $\log r$ , the claim (4.137) follows from (4.139) and (4.140).

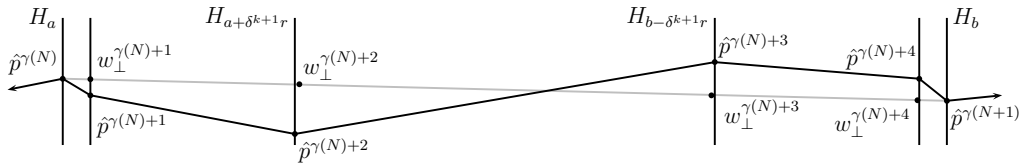


Figure 9: A long  $(j_1 - 1)$ th-scale interval  $I_N = [a, b]$  of  $k$ th-scale length with the totally unbowed case, with  $L^-(I_n) = k + 1$ , showing the full path (black) and intermediate path (gray.) The full path has marked points at the  $\hat{p}^{\gamma(N)+i}$ .

Case 3d. The totally unbowed case (third option in (4.39)) for both  $L^\pm(I_N)$ . Here we have  $V(N) = 4$  as there are the sandwiching hyperplanes at distance  $\delta^{j_1}r$  from each end of  $I_N$ , and inner joining hyperplanes at distance  $\delta^{k+1}r$  from each end, with  $k + 1 \leq j_1 - 1$  (since  $I_N$  is long) such that  $I_N$  is of  $k$ th-scale length; there are no outer joining hyperplanes. See Figure 9. From (4.37) (valid now for  $k + 1$  in place of  $\ell + 1$ ) and (4.39), this means that

$$\frac{|\hat{p}^{\gamma(N)+2} - w_\perp^{\gamma(N)+2}|^2}{\delta^{k+1}r} \geq \frac{1}{2} \max \left( 2^{j_1-k-1} \frac{\sigma(\delta^{k+1}r)}{\sigma(\delta^{j_1}r)} \frac{|\hat{p}^{\gamma(N)+1} - w_\perp^{\gamma(N)+1}|^2}{\delta^{j_1}r}, \right. \\ \left. 2^{j_1-k-2} \frac{\sigma(\delta^{k+1}r)}{\sigma(\delta^{j_1-1}r)} \frac{\delta}{16\mu} \left(\frac{\lambda}{7}\right)^{j_1} t^*(\hat{p}^{\gamma(N)})\sigma(r) \right). \quad (4.141)$$

Here the second term in the max comes from the fact that to come under the third (totally unbowed) option in (4.39), we must have  $2^{k+1}\kappa(k+1) \geq 2^{j_1-1}\kappa(j_1-1)$ . We note that the definition of the totally unbowed case gives (4.141) for  $\hat{w}^i$  in place of  $w_\perp^i$ , but by (4.25) this only changes each of the 4 terms in (4.141) by a factor  $1 + o(1)$  as  $r \rightarrow \infty$ , so we have accounted for this via the factor  $1/2$  in front of the max. The bound (4.141) is for the left end of  $I_N$ ; a symmetric bound is valid for the right end. As before, the direct path is  $\hat{p}^{\gamma(N)} \rightarrow \hat{p}^{\gamma(N+1)}$ , and the full path is

$$\Gamma^{N,full} : \hat{p}^{\gamma(N)} \rightarrow \hat{p}^{\gamma(N)+1} \rightarrow \dots \rightarrow \hat{p}^{\gamma(N)+4} \rightarrow \hat{p}^{\gamma(N+1)},$$

but this time the intermediate path is

$$\Gamma^{N,int} : w_\perp^{\gamma(N)} \rightarrow w_\perp^{\gamma(N)+1} \rightarrow \dots \rightarrow w_\perp^{\gamma(N)+4} \rightarrow w_\perp^{\gamma(N+1)}$$

These points are collinear so  $\Upsilon_{Euc}(\Gamma^{N,int}) = |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}|$ .

We use a special allocation for the totally unbowed case, the same for all 10 links of  $\Gamma^{N,full}$  and  $\Gamma^{N,int}$ , defined (when  $I_N$  is of  $k$ th-scale length) as

$$A_j^4(\Gamma^{N,full}) \\ = \frac{\delta\mu}{540} \left( \frac{|\hat{p}^{\gamma(N)+2} - w_\perp^{\gamma(N)+2}|^2}{\delta^{k+1}r} + \frac{|\hat{p}^{\gamma(N)+3} - w_\perp^{\gamma(N)+3}|^2}{\delta^{k+1}r} + \lambda^j \frac{|\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}|^2}{3\delta^k r} \right).$$

Now  $\Upsilon_{Euc}(\Gamma^{N,int}) = |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}|$  so by Lemma 3.6(ii), applied with  $\epsilon = 1/4$ ,

$$\Upsilon_h(\Gamma^{N,full}) - \Upsilon_h(\Gamma^{N,int}) \geq \frac{\mu}{2} \left[ \Upsilon_{Euc}(\Gamma^{N,full}) - |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| \right] - 4C_{56}.$$

while from (4.26),

$$S_{j_1}(\Gamma^{N,full}) \geq \lambda^{j_1} \frac{\delta\mu}{18} \left[ \Upsilon_{Euc}(\Gamma^{N,full}) - |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| \right] + \frac{1}{3} \left(\frac{\lambda}{7}\right)^{j_1} t\sigma(r). \quad (4.142)$$

From these we get deterministic tracking:

$$\begin{aligned}
 & 10A_{j_1}^4(\Gamma^{N,full}) \\
 & \leq \frac{\delta\mu}{54} \left( 3 \left[ \Upsilon_{Euc}(\Gamma^{N,full}) - |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| \right] \right. \\
 & \quad \left. + 3\lambda^{j_1} \left[ |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| - (\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)})_1 \right] \right) \\
 & \leq \left( \frac{1}{2} - \delta \right) \mu \left[ \Upsilon_{Euc}(\Gamma^{N,full}) - |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| \right] \\
 & \quad + \lambda^{j_1} \frac{\delta\mu}{18} \left[ \Upsilon_{Euc}(\Gamma^{N,full}) - (\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)})_1 \right] \\
 & \leq \Upsilon_h(\Gamma^{N,full}) - \Upsilon_h(\Gamma^{N,int}) - \delta\mu \left[ \Upsilon_{Euc}(\Gamma^{N,full}) - |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| \right] \\
 & \quad + S_{j_1}(\Gamma^{N,full}) - \frac{1}{3} \left( \frac{\lambda}{7} \right)^{j_1} t\sigma(r) + 4C_{56}. \tag{4.143}
 \end{aligned}$$

We now can establish the analog of (4.121) and (4.132) for the totally-unbowed-case contribution to the tracking-failure probability (4.116), using (4.70) and (4.143):

$$\begin{aligned}
 & P \left( \Upsilon_{\hat{T}}(\Gamma^{N,full}) - \hat{T}(\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N+1)}) \leq -S_{j_1}(\Gamma^{N,full}) \right. \\
 & \quad \left. + \delta\mu \left( \Upsilon_{Euc}(\Gamma^{N,full}) - |\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| \right) \text{ for some } N \text{ with } I_N \text{ long non-terminal} \right. \\
 & \quad \left. (j_1 - 1)\text{th-scale, in the totally unbowed case of (4.39) for both of } L^\pm(I_N), \right. \\
 & \quad \left. \text{for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \right) \\
 & \leq P \left( \Upsilon_{\hat{T}}(\Gamma^{N,full}) - \Upsilon_{\hat{T}}(\Gamma^{N,int}) \leq \Upsilon_h(\Gamma^{N,full}) - \Upsilon_h(\Gamma^{N,int}) - 10A_{j_1}^4(\Gamma^{N,full}) \right. \\
 & \quad \left. - \frac{1}{3} \left( \frac{\lambda}{7} \right)^{j_1} t\sigma(r) + 4c_{29} \log r + 4C_{56} \text{ for some } N \text{ with } I_N \text{ long} \right. \\
 & \quad \left. \text{non-terminal } (j_1 - 1)\text{th-scale, in the totally unbowed case of (4.39)} \right. \\
 & \quad \left. \text{for both of } L^\pm(I_N), \text{ for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \right) \\
 & \leq P \left( \left| \hat{T}(\hat{p}^{\gamma(N)+i-1}, \hat{p}^{\gamma(N)+i}) - h(|\hat{p}^{\gamma(N)+i} - \hat{p}^{\gamma(N)+i-1}|) \right| \geq A_{j_1}^4(\Gamma^{N,full}) \text{ or} \right. \\
 & \quad \left| \hat{T}(w_\perp^{\gamma(N)+i-1}, w_\perp^{\gamma(N)+i}) - h(|w_\perp^{\gamma(N)+i} - w_\perp^{\gamma(N)+i-1}|) \right| \geq A_{j_1}^4(\Gamma^{N,full}) \\
 & \quad \text{for some } 1 \leq i \leq 5 \text{ and some } N \text{ with } I_N \text{ long non-terminal } (j_1 - 1)\text{th-scale,} \\
 & \quad \text{in the totally unbowed case of (4.39) for both of } L^\pm(I_N), \\
 & \quad \left. \text{for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \right). \tag{4.144}
 \end{aligned}$$

Here we have again used  $j_1 = O(\log \log r)$ , from (4.22), to ensure  $(\lambda/7)^{j_1} t\sigma(r)/6 \geq 4c_{29} \log r + 4C_{56}$ . With the second term of the max in (4.141) in mind, suppose that for some  $N$  and some  $\nu, m_0, m^* \geq 1$  we have

$$(2^{\nu-1}t)^{1/2} \Delta(r) < |(\hat{p}^{\gamma(N)})^*| \leq (2^\nu t)^{1/2} \Delta(r), \tag{4.145}$$

$$\begin{aligned}
 & 2^{m_0-1} 2^{j_1-k-2} \frac{\sigma(\delta^{k+1}r)}{\sigma(\delta^{j_1-1}r)} \frac{\delta}{16\mu} \left(\frac{\lambda}{7}\right)^{j_1} t^*(\hat{p}^{\gamma(N)})\sigma(r) \\
 & \leq \frac{|\hat{p}^{\gamma(N)+2} - w_{\perp}^{\gamma(N)+2}|^2}{\delta^{k+1}r} + \frac{|\hat{p}^{\gamma(N)+3} - w_{\perp}^{\gamma(N)+3}|^2}{\delta^{k+1}r} \\
 & \leq 2^{m_0} 2^{j_1-k-2} \frac{\sigma(\delta^{k+1}r)}{\sigma(\delta^{j_1-1}r)} \frac{\delta}{16\mu} \left(\frac{\lambda}{7}\right)^{j_1} t^*(\hat{p}^{\gamma(N)})\sigma(r), \tag{4.146}
 \end{aligned}$$

and

$$\begin{aligned}
 & 2^{m^*-1} 2^{j_1-k-2} \frac{\sigma(\delta^{k+1}r)}{\sigma(\delta^{j_1-1}r)} \frac{\delta}{16\mu} \left(\frac{\lambda}{7}\right)^{j_1} t^*(\hat{p}^{\gamma(N)})\sigma(r) \\
 & \leq \lambda^{j_1-1} \frac{|(\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)})^*|^2}{\delta^k r} \\
 & \leq 2^{m^*} 2^{j_1-k-2} \frac{\sigma(\delta^{k+1}r)}{\sigma(\delta^{j_1-1}r)} \frac{\delta}{16\mu} \left(\frac{\lambda}{7}\right)^{j_1} t^*(\hat{p}^{\gamma(N)})\sigma(r). \tag{4.147}
 \end{aligned}$$

Then using (1.11),

$$\frac{A_{j_1}^4(\Gamma^{N,full})}{\sigma(\delta^k r)} \geq c_{37}(2^{m_0} + 2^{m^*}) 2^{j_1-k} \left(\frac{\lambda}{7\delta^{\chi_1}}\right)^{j_1} 2^\nu t, \tag{4.148}$$

and by Lemma 3.2, for every possible value  $(u^1, \dots, u^6)$  of  $(\hat{p}^{\gamma(N)}, \dots, \hat{p}^{\gamma(N+1)})$  and every link  $(u^{i-1}, u^i)$ , we have

$$P\left(\left|\hat{T}(u^i - u^{i-1}) - h(|u^{i-1} - u^i|)\right| \geq A_{j_1}^4(u^1, \dots, u^6)\right) \leq C_{44} \exp\left(-C_{45} \frac{A_{j_1}^4(u^1, \dots, u^6)}{\sigma(\delta^k r)}\right). \tag{4.149}$$

It is important here that the lower bound (4.148) not depend on the length scale  $k$  of  $I_N$ , except through the factor  $2^{j_1-k}$  which is always at least 1. Similarly to the entropy bounds in the proofs of Lemmas 4.6 and 4.7, and in Case 3c, we see that the number of possible choices of  $\Gamma^{N,full}$  satisfying (4.145)–(4.147) is at most

$$\begin{aligned}
 & \frac{c_{38}}{\delta^{2j_1}} \cdot \left(\frac{2^\nu t}{\beta^{2j_1}}\right)^{(d-1)/2} \cdot \left(2^{m^*} 2^\nu \left(\frac{2}{\delta^{1+\chi_2}}\right)^{j_1-k} \left(\frac{\delta}{7\beta^2}\right)^{j_1} t\right)^{(d-1)/2} \\
 & \cdot \left(2^{m_0} 2^\nu \left(\frac{2}{\delta^{1+\chi_2}}\right)^{j_1-k} \left(\frac{\lambda\delta}{7\beta^2}\right)^{j_1} t\right)^{2(d-1)}. \tag{4.150}
 \end{aligned}$$

Here the dots separate bounds for the number of choices (up to a constant) first of endpoint hyperplanes  $(\hat{p}^{\gamma(N)})_1, (\hat{p}^{\gamma(N+1)})_1$ , and then of  $\hat{p}^{\gamma(N)}$ , of  $\hat{p}^{\gamma(N+1)}$ , and finally of  $(\hat{p}^{\gamma(N+1)}, \dots, \hat{p}^{\gamma(N+4)})$ . For  $|(\hat{p}^{\gamma(N)})^*| \leq t^{1/2} \Delta(r)$ , (4.148) and (4.150) remain valid for  $m_0, m^* \geq 1$  with  $2^\nu$  replaced by 1. For  $\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N+1)}$  with

$$\lambda^{j_1-1} \frac{|(\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)})^*|^2}{\delta^k r} \leq 2^{j_1-k-2} \frac{\sigma(\delta^{k+1}r)}{\sigma(\delta^{j_1-1}r)} \frac{\delta}{16\mu} \left(\frac{\lambda}{7}\right)^{j_1} t^*(\hat{p}^{\gamma(N)})\sigma(r)$$

(i.e. the left side too small to satisfy (4.147) for any  $m^* \geq 1$ ) they remain valid with  $2^{m^*}$  replaced by 1. (Note that by (4.141),  $m_0 \geq 1$  covers all cases.) Hence as in previous cases, inserting the bound (4.148) into (4.149), multiplying by the entropy factor (4.150), and summing over  $\nu \geq 0, m^* \geq 0, m_0 \geq 1$ , and  $k \leq j_1 - 2$  we get that the right side of (4.144) is bounded by

$$c_{39} \left(\frac{2\lambda}{7^5 \beta^{12} \delta^5 \chi_2}\right)^{(d-1)j_1/2} t^{3(d-1)} \exp\left(-c_{40} \left(\frac{\lambda}{7\delta^{\chi_1}}\right)^{j_1} t\right). \tag{4.151}$$

We have thus bounded the totally-unbowed-case contribution to the tracking-failure probability (4.116).

Case 3e. Mixed cases, which we subdivide as mixed forward case, mixed bowed case, and mixed totally unbowed case, according to the condition at the dominant end (as defined after (4.42)) of  $I_N$ . In Cases 3a–3d we have assumed that the same option in (4.39) occurs at both ends of the interval  $I_N$ , but this is strictly for clarity of exposition. As explained in Step 1, in mixed cases we have joining hyperplanes only at the dominant end of  $I_N$ . The situation is symmetric so let us assume the left end is dominant. In mixed cases the full path in  $I_N$  is always  $\hat{p}^{\gamma(N)} \rightarrow \hat{p}^{\gamma(N)+1} \rightarrow \dots \rightarrow \hat{p}^{\gamma(N+1)}$ , with  $\gamma(N+1) - \gamma(N) = 4$  or  $5$ . Suppose  $I_N$  is a  $(j_1 - 1)$ th-scale interval of  $k$ th-scale length. The intermediate paths are analogous to those in Cases 3a–3d, as follows. If the dominant left end has the forward case then there is one joining hyperplane in  $I_N$ , containing  $\hat{p}^{\gamma(N)+2}$ , at the left end at distance  $\delta^{j_1-1}r$  from  $\hat{p}^{\gamma(N)}$ , and the intermediate path is

$$\Gamma^{N,int} : \hat{p}^{\gamma(N)} \rightarrow w_{\perp}^{\gamma(N)+1} \rightarrow \hat{p}^{\gamma(N)+2} \rightarrow \hat{p}^{\gamma(N)+3} \rightarrow \hat{p}^{\gamma(N+1)}.$$

If the (dominant) left end has the totally unbowed case then there is one (inner) joining hyperplane in  $I_N$ , containing  $\hat{p}^{\gamma(N)+2}$ , at the left end at distance  $\delta^{k+1}r$  from  $\hat{p}^{\gamma(N)}$ , and the intermediate path follows the line from  $\hat{p}^{\gamma(N)}$  to  $\hat{p}^{\gamma(N+1)}$ :

$$\Gamma^{N,int} : \hat{p}^{\gamma(N)} \rightarrow w_{\perp}^{\gamma(N)+1} \rightarrow \dots \rightarrow w_{\perp}^{\gamma(N)+3} \rightarrow \hat{p}^{\gamma(N+1)},$$

similar to Figure 9 but without the third hyperplane from the right. If the left end has the bowed case with  $L^-(N) = \ell < j_1 - 1$  then there are two joining hyperplanes in  $I_N$ , containing  $\hat{p}^{\gamma(N)+2}$  and  $\hat{p}^{\gamma(N)+3}$ , at the left end at distances  $\delta^{\ell+1}r$  and  $\delta^{\ell}r$  from  $\hat{p}^{\gamma(N)}$ , and the intermediate path is

$$\Gamma^{N,int} : \hat{p}^{\gamma(N)} \rightarrow \hat{z}^{\gamma(N)+1} \rightarrow \hat{z}^{\gamma(N)+2} \rightarrow \hat{p}^{\gamma(N)+3} \rightarrow \hat{p}^{\gamma(N)+4} \rightarrow \hat{p}^{\gamma(N+1)};$$

see Figure 8. If the left end has the bowed case with  $L^-(N) = j_1 - 1$  then there is one (inner) joining hyperplane in  $I_N$ , containing  $\hat{p}^{\gamma(N)+2}$ , at the left end at distance  $\delta^{j_1-1}r$  from  $\hat{p}^{\gamma(N)}$ , and the intermediate path is

$$\Gamma^{N,int} : \hat{p}^{\gamma(N)} \rightarrow \hat{z}^{\gamma(N)+1} \rightarrow \hat{p}^{\gamma(N)+2} \rightarrow \hat{p}^{\gamma(N)+3} \rightarrow \hat{p}^{\gamma(N+1)},$$

similar to Figure 8 but with the middle two hyperplanes coinciding. In each case, we define the subpaths  $\Gamma^{N,full,-}$  and  $\Gamma^{N,int,-}$  between the left end of the interval and the left inner joining hyperplane. In all cases the arguments are essentially the same as in the analogous non-mixed case, with the addition that for the final transition ending at  $\hat{p}^{\gamma(N+1)}$ , one uses (4.44) to bound  $|\hat{p}^{\gamma(N+1)-1} - w_{\perp}^{\gamma(N+1)-1}|$ . We will not reiterate the arguments here, but simply state that the result is again the analog of (4.121) and (4.132), bounding the mixed-case contribution to the tracking-failure probability (4.116):

$$\begin{aligned} &P\left(\hat{T}(\hat{p}^{\gamma(N)}, \hat{p}^{\gamma(N+1)}) - \Upsilon_{\hat{T}}(\Gamma^{N,full}) \geq S_{j_1}(\Gamma^{N,full}) \right. \\ &\quad + \delta\mu\left(|\hat{p}^{\gamma(N+1)} - \hat{p}^{\gamma(N)}| - \Upsilon_{Euc}(\Gamma^{N,full})\right) + \frac{1}{3}\lambda^{j_1}\left[t\sigma(r) + \delta\mu\left(\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,2}) - (\hat{y} - \hat{x})_1\right)\right] \\ &\quad + \frac{\lambda^{j_1}t\sigma(r)}{6m} \text{ for some } N \text{ with } I_N \text{ non-terminal long } (j_1 - 1)\text{th-scale, in the} \\ &\quad \text{mixed case of (4.39) for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)}\Big) \\ &\leq P\left(\left|\hat{T}(u, v) - h(|v - u|)\right| \geq A \text{ for some link } (u, v) \text{ of } \Gamma^{N,full,-} \text{ or } \Gamma^{N,int,-}\right) \end{aligned}$$

(or  $\Gamma^{N,full}, \Gamma^{N,int}$  in the mixed totally unbowed case) for some  $N$  with  $I_N$  long non-terminal  $(j_1 - 1)$ th-scale, in the mixed case of (4.39)

$$\text{for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \Big), \tag{4.152}$$

where

$$A = \begin{cases} \infty & \text{in the mixed forward case for } I_N, \\ A_{j_1}^3(\Gamma^{N,full,-}) & \text{in the mixed bowed case for } I_N, \\ A_{j_1}^4(\Gamma^{N,full}) & \text{in the mixed totally unbowed case for } I_N. \end{cases}$$

As in Cases 3a–3d, the right side of (4.152), and thus the mixed-case contribution to the tracking-failure event (4.116), is bounded above by

$$f(\lambda, \delta, \beta)^{j_1} \exp \left( -c_{41} \left( \frac{\lambda}{7\delta\chi_1} \right)^{j_1} t \right) \tag{4.153}$$

for some  $f(\lambda, \delta, \beta)$ .

We have now bounded all contributions to (4.116), by the sum of (4.122), (4.123), (4.134), (4.135), (4.138), (4.151), and (4.153). From this and (4.113) we have the completion of the  $(j_1 - 1)$ th-scale iteration:

$$\begin{aligned} &P\left(T(x, y) \leq h((y - x)_1) - t\sigma(r) \text{ for some } (x, y) \in X_r\right) \\ &\leq P\left(\Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-1,2}) - h((y - x)_1) \leq - (1 - 5\lambda^{j_1}) t\sigma(r) + 7\delta\mu\lambda^{j_1} \left(\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y} - \hat{x})_1\right) \right. \\ &\quad \left. + \delta\mu \left(\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,2}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0})\right) \text{ for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)}\right) \\ &\quad + c_{32}e^{-c_{33}t} + c_{42} \exp \left( -c_{43} \left( \frac{\lambda}{7\delta\chi_1} \right)^{j_1} t \right). \end{aligned} \tag{4.154}$$

As in the comments after (4.81) and (4.112), the increase of the coefficients 4, 6 in (4.81) to be 5, 7 in (4.112), together with an increment to the “length-change” term with coefficient  $\delta\mu$ , represent a further reduction taken from the original bound  $t\sigma(r)$  in (1.14), and allocated to bound errors in the second stage of the  $(j_1 - 1)$ th-scale iteration, just completed. In this second stage, however, the increment of the length-change term, due to replacing  $\Gamma_{xy}^{j_1-1,1}$  with  $\Gamma_{xy}^{j_1-1,2}$ , is negative (removing points reduces path length) and may compensate at least in part for the increases from 4, 6 to 5, 7. This compensation, made possible by control of tracking errors, is what allows the total allocation through all iterations to potentially exceed the entire original bound  $t\sigma(r)$ —see (4.167).

#### 4.9 Step 7. Further iterations of coarse-graining

The further iterations proceed quite similarly to the  $(j_1 - 1)$ th-scale iteration; for the most part, to do the  $j$ th-scale iteration we simply replace  $j_1 - 1, j_1$  throughout by  $j, j + 1$ . We will sketch the  $(j_1 - 2)$ th-scale (third) iteration to make the pattern clear.

For the first stage of the  $(j_1 - 2)$ th-scale iteration, the current marked PG path at the start is  $\Gamma_{xy}^{j_1-1,2} : b^0 \rightarrow b^1 \rightarrow \dots \rightarrow b^{n+1}$ , which we rename  $\Gamma_{xy}^{j_1-2,0}$ . Letting

$$I_{xy}^2 = \{i \in \{3, \dots, n - 2\} : b^i \text{ lies in a non-incident } (j_1 - 2)\text{th-scale hyperplane in } \mathcal{H}_{xy}\}$$

we shift each  $b^i$ ,  $i \in I_{xy}^2$ , to the  $(j_1 - 2)$ th-scale grid, creating the updated marked PG path

$$\Gamma_{xy}^{j_1-2,1} : \hat{b}^0 \rightarrow \hat{b}^1 \rightarrow \dots \rightarrow \hat{b}^n \rightarrow \hat{b}^{n+1}.$$

Removing those  $b^i$  which lie in non-terminal  $(j_1 - 1)$ th-scale hyperplanes then creates the marked PG path

$$\Gamma_{xy}^{j_1-2,2} : \zeta^0 \rightarrow \zeta^1 \rightarrow \dots \rightarrow \zeta^{n_2} \rightarrow \zeta^{n_2+1},$$

in which (recalling (4.24))  $\zeta^0 = \hat{x}$ ,  $\zeta^{n_2+1} = \hat{y}$ ,  $\zeta^1$  and  $\zeta^{n_2}$  lie in terminal  $j_1$ th-scale hyperplanes,  $\zeta^2, \zeta^{n_2-1}$  lie in terminal  $(j_1 - 1)$ th-scale hyperplanes, and  $\zeta^3, \dots, \zeta^{n_2-2}$  lie in  $(j_1 - 2)$ th-scale hyperplanes. Noninteraction of shifts again applies (see Step 4), since if  $b^i$  and  $b^k$  both lie in  $(j_1 - 2)$ th-scale hyperplanes then there are sandwiching (and possibly also joining) hyperplanes in between, ensuring  $|i - k| \geq 2$ . For  $N \leq n_2$  we write  $\Gamma^{2,N,full}$  for the portion of  $\Gamma_{xy}^{j_1-2,1}$  in the interval (denoted  $I_N^2$ ) from  $(\zeta^N)_1$  to  $(\zeta^{N+1})_1$ , and define  $\gamma(2, N)$  by  $\zeta^N = \hat{b}^{\gamma(2,N)}$ ,  $N \leq n_2 + 1$ , analogous to  $\gamma(N)$  used in the previous iteration.

Similarly to (4.106)–(4.108) we have

$$\sum_{i \in I_{xy}} A_{j_1-1}^1(\Gamma_{xy}^{j_1-2,0}, i) \leq \frac{2}{3} \lambda^{j_1-1} \left[ t\sigma(r) + \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-2,0}) - (\hat{y} - \hat{x})_1 \right) \right], \quad (4.155)$$

$$\delta\mu \left| \Upsilon_{Euc}(\Gamma_{xy}^{j_1-2,1}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_1-2,0}) \right| \leq \frac{\delta\lambda^{j_1-1}}{2} \left[ \frac{t\sigma(r)}{3} + \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-2,0}) - (\hat{y} - \hat{x})_1 \right) \right], \quad (4.156)$$

and

$$\left| \Upsilon_h(\Gamma_{xy}^{j_1-2,1}) - \Upsilon_h(\Gamma_{xy}^{j_1-2,0}) \right| \leq \lambda^{j_1-1} \left[ \frac{t\sigma(r)}{3} + \delta\mu \left( \Upsilon_{Euc}(\Gamma_{xy}^{j_1-2,0}) - (\hat{y} - \hat{x})_1 \right) \right]. \quad (4.157)$$

These lead to a tracking bound like (4.111) for the first stage (shifting to the  $(j_1 - 2)$ th-scale grid) using Lemma 4.6:

$$\begin{aligned} & P \left( \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-2,0}) - \Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-2,1}) < - \sum_{i \in I_{xy}^2} A_{j_1-1}^1(\Gamma_{xy}^{j_1-2,0}, i) + \Upsilon_h(\Gamma_{xy}^{j_1-2,0}) - \Upsilon_h(\Gamma_{xy}^{j_1-2,1}) \right. \\ & \quad \left. \text{for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \right) \\ & \leq P \left( \max \left( |\hat{T}(\hat{b}^{i-1}, \hat{b}^i) - h(|\hat{b}^i - \hat{b}^{i-1}|)|, |\hat{T}(b^{i-1}, b^i) - h(|b^i - b^{i-1}|)| \right) \right. \\ & \quad \left. > \frac{1}{4} A_{j_1-1}^1(\Gamma_{xy}^{j_1-2,0}, i) \text{ for some } 1 \leq i \leq n+1 \text{ and } (x, y) \in X_r \right. \\ & \quad \left. \text{with } \{i-1, i\} \cap I_{xy}^2 \neq \emptyset; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \right) \\ & \leq P \left( |\hat{T}(v, w) - h(|w - v|)| \geq A_{j_1-1}^2(v, w) \right. \\ & \quad \left. \text{for some } (j_1 - 1)\text{th-scale transition } v \rightarrow w \text{ with } v, w \in G_r^+ \cap \mathbb{L}_{j_1-1} \right) \\ & \leq C_{70} \exp \left( -C_{71} \left( \frac{\lambda}{7\delta x_1} \right)^{j_1-1} t \right). \quad (4.158) \end{aligned}$$

As with (4.81) and (4.112), with (4.154) this yields

$$P \left( T(x, y) \leq h((y - x)_1) - t\sigma(r) \text{ for some } (x, y) \in X_r \right)$$



$$\begin{aligned} &\leq P\left(\Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-2,1}) - h((y-x)_1) \leq -(1-2\lambda^{j_1})t\sigma(r) + 4\delta\mu\lambda^{j_1}\left(\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y}-\hat{x})_1\right)\right. \\ &\quad + 3\lambda^{j_1}\left(t\sigma(r) + \delta\mu\left[\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y}-\hat{x})_1\right]\right) \\ &\quad + 2\lambda^{j_1-1}\left(t\sigma(r) + \delta\mu\left[\Upsilon_{Euc}(\Gamma_{xy}^{j_1-2,0}) - (\hat{y}-\hat{x})_1\right]\right) \\ &\quad - \delta\mu\left(\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_1-2,1})\right) \text{ for some } (x,y) \in X_r; \\ &\quad \left.\omega \notin J^{(0)}(c_{29}) \cup J^{(1c)}\right) \\ &\quad + c_{32}e^{-c_{33}t} + c_{42}\exp\left(-c_{43}\left(\frac{\lambda}{7\delta\lambda_1}\right)^{j_1}t\right) + C_{70}\exp\left(-C_{71}\left(\frac{\lambda}{7\delta\lambda_1}\right)^{j_1-1}t\right). \end{aligned} \quad (4.159)$$

**Remark 4.9.** The second probability here contains three quantities of the form

$$\lambda^{j+1}\left(t\sigma(r) + \delta\mu\left[\Upsilon_{Euc}(\Gamma_{xy}^{j,0}) - (\hat{y}-\hat{x})_1\right]\right),$$

with certain additional integer coefficients: 2 and 4 in the first quantity, 3 in the second quantity, and 2 in the third. As noted in Remark 4.8 and after (4.112), (4.116), and (4.154), these represent the accumulated error allocations from the  $j_1$ th,  $(j_1 - 1)$ th, and (first-stage)  $(j_1 - 2)$ th-scale iterations, respectively. We have split these out here for clarity; in (4.113) and (4.154) the first two of the three quantities are combined, giving the terms with coefficients 5 and 7. In general after the first stage of the  $j$ th-scale iteration (which is the source of the third quantity), the second quantity would represent allocations from all stages  $j_1 - 1$  through  $j + 1$ . With this in mind we define the accumulated-allocations upper bounds

$$\begin{aligned} \mathcal{A}_j^1(\Gamma_{xy}) &= 2\lambda^{j_1}\left(t\sigma(r) + 2\delta\mu\left[\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y}-\hat{x})_1\right]\right) \\ &\quad + \sum_{k=j+1}^{j_1-1} 3\lambda^{k+1}\left(t\sigma(r) + \delta\mu\left[\Upsilon_{Euc}(\Gamma_{xy}^{k,0}) - (\hat{y}-\hat{x})_1\right]\right) \\ &\quad + 2\lambda^{j+1}\left(t\sigma(r) + \delta\mu\left[\Upsilon_{Euc}(\Gamma_{xy}^{j,0}) - (\hat{y}-\hat{x})_1\right]\right) \end{aligned} \quad (4.160)$$

and

$$\begin{aligned} \mathcal{A}_j^2(\Gamma_{xy}) &= 2\lambda^{j_1}\left(t\sigma(r) + 2\delta\mu\left[\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y}-\hat{x})_1\right]\right) \\ &\quad + \sum_{k=j}^{j_1-1} 3\lambda^{k+1}\left(t\sigma(r) + \delta\mu\left[\Upsilon_{Euc}(\Gamma_{xy}^{k,0}) - (\hat{y}-\hat{x})_1\right]\right), \end{aligned} \quad (4.161)$$

with  $\mathcal{A}_j^1(\Gamma_{xy})$  a valid upper bound for accumulated allocations after the first stage of the  $j$ th-scale iteration, and  $\mathcal{A}_j^2(\Gamma_{xy})$  valid after the second stage. Comparing (4.160) to (4.112), in this case we have  $j = j_1 - 1$ , the sum in (4.160) has no terms, and the other two terms on the right in (4.112) merge into one. Comparing (4.161) to (4.154), in this instance also  $j = j_1 - 1$ , and the sum in (4.161) has one term which in (4.154) is merged with the other term on the right. Equation (4.161) may also be compared to (4.81), where  $j = j_1$  and the sum in (4.161) has no terms. We can rewrite the second probability in (4.159) as

$$P\left(\Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-2,1}) - h((y-x)_1) \leq -t\sigma(r) + \mathcal{A}_{j_1-2}^1(\Gamma_{xy})\right)$$

$$- \delta\mu \left( \Upsilon_{Euc} (\Gamma_{xy}^{j_1-1,0}) - \Upsilon_{Euc} (\Gamma_{xy}^{j_1-2,1}) \right) \text{ for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \Big). \tag{4.162}$$

Moving on to the second stage of the  $(j_1 - 2)$ th-scale iteration, from (4.159), using the form (4.162), we have similarly to (4.113) and (4.116)

$$\begin{aligned} & P \left( T(x, y) \leq h((y - x)_1) - t\sigma(r) \text{ for some } (x, y) \in X_r \right) \\ & \leq P \left( \Upsilon_{\hat{T}} (\Gamma_{xy}^{j_1-2,2}) - h((y - x)_1) \leq -t\sigma(r) + \mathcal{A}_{j_1-2}^2(\Gamma_{xy}) \right. \\ & \quad \left. - \delta\mu \left( \Upsilon_{Euc} (\Gamma_{xy}^{j_1-1,0}) - \Upsilon_{Euc} (\Gamma_{xy}^{j_1-2,2}) \right) \text{ for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \right) \\ & \quad + P \left( \Upsilon_{\hat{T}} (\Gamma_{xy}^{j_1-2,1}) - \Upsilon_{\hat{T}} (\Gamma_{xy}^{j_1-2,2}) \leq - \sum_{N=0}^{n_2} S_{j_1-1}(\Gamma^{2,N,full}) \right. \\ & \quad \quad \left. - \frac{1}{3} \lambda^{j_1-1} \left( t\sigma(r) + \delta\mu \left[ \Upsilon_{Euc} (\Gamma_{xy}^{j_1-2,2}) - (\hat{y} - \hat{x})_1 \right] \right) \right. \\ & \quad \quad \left. + \delta\mu \left( \Upsilon_{Euc} (\Gamma_{xy}^{j_1-2,1}) - \Upsilon_{Euc} (\Gamma_{xy}^{j_1-2,2}) \right) \text{ for some } (x, y) \in X_r; \right. \\ & \quad \quad \left. \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \right) \\ & \quad + c_{32} e^{-c_{33}t} + c_{42} \exp \left( -c_{43} \left( \frac{\lambda}{7\delta\chi_1} \right)^{j_1} t \right) + C_{70} \exp \left( -C_{71} \left( \frac{\lambda}{7\delta\chi_1} \right)^{j_1-1} t \right), \tag{4.163} \end{aligned}$$

the last probability being the tracking-failure probability. We now divide into Cases 1–5 as in the  $(j_1 - 1)$ th-scale iteration. In each case (excluding 3c) there is an intermediate path  $\Gamma^{2,N,int}$  in each  $(j_1 - 2)$ th-scale interval  $I_N^2$ , which is slower than the direct path (up to a multiple of  $\log r$ ), satisfying

$$\Upsilon_{\hat{T}}(\Gamma^{2,N,int}) \geq \hat{T}(\hat{b}^{\gamma(2,N)}, \hat{b}^{\gamma(2,N+1)}) - (\gamma(2, N + 1) - \gamma(2, N) - 1)c_{29} \log r$$

assuming  $\omega \notin J^{(0)}(c_{29})$ , and hence

$$\sum_{N=0}^{n_2} \Upsilon_{\hat{T}}(\Gamma^{2,N,int}) \geq \Upsilon_{\hat{T}} (\Gamma_{xy}^{j_1-2,2}) - (n - n_2)c_{29} \log r,$$

and which satisfies deterministic tracking in a form like (4.117), sometimes with extra terms in the form of error allocations, as in (4.143); this deterministic tracking says that the intermediate path is sufficiently shorter relative to the full path, to within the error given by the allocations. This enables us to bound the last probability in (4.163) by a tracking-failure probability of the form

$$\begin{aligned} & P \left( \Upsilon_{\hat{T}} (\Gamma^{2,N,full}) - \Upsilon_{\hat{T}} (\Gamma^{2,N,int}) \leq \Upsilon_h (\Gamma^{2,N,full}) - \Upsilon_h (\Gamma^{2,N,int}) - (\text{allocations}) \right. \\ & \quad \left. \text{for some } N \text{ for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \right), \tag{4.164} \end{aligned}$$

differently for each of the 5 cases; see for example (4.144) for the totally-unbowed case. This is then bounded by summing probabilities of form

$$P \left( \left| \hat{T}(v, w) - h(|w - v|) \right| \geq A \right)$$

over all possible links  $(v, w)$  of paths  $\Gamma^{2,N,full}, \Gamma^{2,N,int}$  for all  $N$ , using Lemmas 4.6 and 4.7. The end result of the  $(j_1 - 2)$ th-scale iteration is that

$$\begin{aligned}
 & P\left(T(x, y) \leq h((y - x)_1) - t\sigma(r) \text{ for some } (x, y) \in X_r\right) \\
 & \leq P\left(\Upsilon_{\hat{T}}(\Gamma_{xy}^{j_1-2,2}) - h((y - x)_1) \leq -t\sigma(r) + \mathcal{A}_{j_1-2}^2(\Gamma_{xy}) \right. \\
 & \quad \left. - \delta\mu\left(\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_1-2,2})\right) \text{ for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)}\right) \\
 & \quad + c_{32}e^{-c_{33}t} + \sum_{j=j_1-2}^{j_1-1} c_{42} \exp\left(-c_{43} \left(\frac{\lambda}{7\delta\lambda_1}\right)^{j+1} t\right). \tag{4.165}
 \end{aligned}$$

After all iterations are completed through the  $j_2$ th scale, this becomes

$$\begin{aligned}
 & P\left(T(x, y) \leq h((y - x)_1) - t\sigma(r) \text{ for some } (x, y) \in X_r\right) \\
 & \leq P\left(\Upsilon_{\hat{T}}(\Gamma_{xy}^{j_2,2}) - h((y - x)_1) \leq -t\sigma(r) + \mathcal{A}_{j_2}^2(\Gamma_{xy}) - \delta\mu\left(\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_2,2})\right) \right. \\
 & \quad \left. \text{for some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)}\right) \\
 & \quad + c_{32}e^{-c_{33}t} + \sum_{j=j_2}^{j_1-1} c_{42} \exp\left(-c_{43} \left(\frac{\lambda}{7\delta\lambda_1}\right)^{j+1} t\right). \tag{4.166}
 \end{aligned}$$

We claim that

$$\mathcal{A}_{j_2}^2(\Gamma_{xy}) - \delta\mu\left(\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - \Upsilon_{Euc}(\Gamma_{xy}^{j_2,2})\right) \leq \frac{1}{2}t\sigma(r) + \delta\mu\left(\Upsilon_{Euc}(\Gamma_{xy}^{j_2,2}) - (\hat{y} - \hat{x})_1\right). \tag{4.167}$$

This says that the second term on the left, representing part of the cumulative effects of tracking, cancels enough of the cumulative allocations  $\mathcal{A}_{j_2}^2(\Gamma_{xy})$ ; see the comments after (4.116) and (4.154). To prove (4.167) we make the subclaim

$$D := \max_{j_2 \leq j \leq j_1-1} \Upsilon_{Euc}(\Gamma_{xy}^{j,0}) - (\hat{y} - \hat{x})_1 \leq 2\left(\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y} - \hat{x})_1\right) + t\sigma(r). \tag{4.168}$$

Assuming (4.168) and assuming  $\lambda$  satisfies  $7\lambda/(1 - \lambda) < 1/4$  we have from the definition (4.161)

$$\begin{aligned}
 \mathcal{A}_{j_2}^2(\Gamma_{xy}) & \leq \frac{7\lambda}{1 - \lambda} \left( t\sigma(r) + \delta\mu \left[ \max_{j_2 \leq j \leq j_1-1} \Upsilon_{Euc}(\Gamma_{xy}^{j,0}) - (\hat{y} - \hat{x})_1 \right] \right) \\
 & \leq \frac{1 + \delta\mu}{4} t\sigma(r) + \delta\mu\left(\Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y} - \hat{x})_1\right) \tag{4.169}
 \end{aligned}$$

and (4.167) follows. To prove (4.168) we use (4.107) (which generalizes to all  $j$  in place of  $j_1 - 1$ ) and the fact that, since removing points reduces length, we have  $\Upsilon_{Euc}(\Gamma_{xy}^{j,2}) \leq \Upsilon_{Euc}(\Gamma_{xy}^{j,1})$  for all  $j$ . This yields that for all  $j \geq 1$ ,

$$\begin{aligned}
 \Upsilon_{Euc}(\Gamma_{xy}^{j,0}) & \leq \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) + \sum_{k=j+1}^{j_1-1} \left| \Upsilon_{Euc}(\Gamma_{xy}^{k,1}) - \Upsilon_{Euc}(\Gamma_{xy}^{k,0}) \right| \\
 & \leq \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) + \sum_{k=j+1}^{j_1-1} \frac{\lambda^{k+1}}{2\mu} \left[ \frac{t\sigma(r)}{3} + \delta\mu D \right] \tag{4.170}
 \end{aligned}$$

and therefore

$$D \leq \Upsilon_{Euc}(\Gamma_{xy}^{j_1-1,0}) - (\hat{y} - \hat{x})_1 + \frac{\lambda^3}{6\mu(1-\lambda)}t\sigma(r) + \frac{\delta\lambda^3}{2(1-\lambda)}D, \tag{4.171}$$

from which (4.168) follows, and hence also (4.167). The right side of (4.166) is therefore bounded above by

$$P\left(\Upsilon_{\hat{T}}(\Gamma_{xy}^{j_2,2}) - h((y-x)_1) \leq -\frac{1}{2}t\sigma(r) + \delta\mu\left(\Upsilon_{Euc}(\Gamma_{xy}^{j_2,2}) - (\hat{y} - \hat{x})_1\right) \right. \\ \left. \text{for some } (x,y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)}\right) + c_{44} \exp\left(-c_{45} \left(\frac{\lambda}{7\delta x_1}\right)^{j_2+1} t\right). \tag{4.172}$$

**4.10 Step 8. Final marked CG paths**

In  $\Gamma_{xy}^{j_2,2}$ , only terminal hyperplanes contain marked points (excluding  $\hat{x}, \hat{y}$ ), one near each end of  $[\hat{x}_1, \hat{y}_1]$  for each scale  $j_2 \leq j \leq j_1$ . The marked point in the  $j$ -terminal hyperplane is in the grid  $L_j$  for all  $j$ . As noted in Step 1, the gap between the  $j_2$ -terminal hyperplanes is at least  $4\delta^{j_2}r$  and at most  $5\delta^{j_2-1}r$ . Now  $\Gamma_{xy}^{j_2,2}$  is our final CG path, so we rename it

$$\Gamma_{xy}^{CG} : \hat{x} = u^0 \rightarrow \dots \rightarrow u^{R+1} = \hat{y},$$

where  $R = 2(j_1 - j_2 + 1)$ . We also define a projected path with collinear marked points

$$\Gamma_{xy}^{CG,pr} : v^0 \rightarrow \dots \rightarrow v^{R+1},$$

where  $v^i = (u^i)_1 e_1$  is the projection onto the  $e_1$  axis. For  $j_1 - 1 \leq j \leq j_2$ , the links  $(u^{j_1-j}, u^{j_1-j+1})$  and  $(u^{R-j_1+j}, u^{R-j_1+j+1})$  each have one end in a  $j$ -terminal hyperplane and the other in a  $(j + 1)$ -terminal hyperplane; these will be called the *jth-scale links* of  $\Gamma_{xy}^{CG}$ . The links  $(u^0, u^1)$  and  $(u^R, u^{R+1})$  are called *final links*, and the link  $(u^{R/2}, u^{R/2+1})$  between  $j_2$ -terminal hyperplanes is called a *macroscopic link*. See Figure 3 in Section 4.3, where  $u^0, \dots, u^7$  are the marked points in  $\Gamma_{xy}^{CG}$  and  $(u^3, u^4)$  is the macroscopic link, which (by definition of  $j_2$ ) always covers at least  $1/6$  the length of  $[\hat{x}_1, \hat{y}_1]$ .

We have  $\Upsilon_{Euc}(\Gamma_{xy}^{CG,pr}) = (\hat{y} - \hat{x})_1$  so by Lemma 3.6(ii) (applied with  $\epsilon = 1/4$  and assuming, as we may, that  $\delta < 1/4$ ), since  $R \leq 2j_1$ ,

$$\Upsilon_h(\Gamma_{xy}^{CG}) - \Upsilon_h(\Gamma_{xy}^{CG,pr}) \geq \left(\frac{1}{2} + \delta\right)\mu\left(\Upsilon_{Euc}(\Gamma_{xy}^{CG}) - (\hat{y} - \hat{x})_1\right) - 2j_1 C_{56}.$$

Therefore, using (4.70), the probability in (4.172) is bounded above by

$$P\left(\Upsilon_{\hat{T}}(\Gamma_{xy}^{CG}) - \Upsilon_h(\Gamma_{xy}^{CG}) \leq -\frac{1}{3}t\sigma(r) + \delta\mu\left(\Upsilon_{Euc}(\Gamma_{xy}^{CG}) - (\hat{y} - \hat{x})_1\right) \right. \\ \left. - \left[\Upsilon_h(\Gamma_{xy}^{CG}) - \Upsilon_h(\Gamma_{xy}^{CG,pr})\right] \text{ for some } (x,y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)}\right) \\ \leq P\left(\Upsilon_{\hat{T}}(\Gamma_{xy}^{CG}) - \Upsilon_h(\Gamma_{xy}^{CG}) \leq -\frac{1}{3}t\sigma(r) - \frac{\mu}{2}\left(\Upsilon_{Euc}(\Gamma_{xy}^{CG}) - (\hat{y} - \hat{x})_1\right) \right. \\ \left. \text{for some } (x,y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)}\right). \tag{4.173}$$

To bound this we use allocations

$$A_j^{CG}(v, w) = \frac{1}{4} \lambda^j \left[ t^*(v, w) \sigma(r) + \frac{\delta \mu}{18} \frac{|(w-v)^*|^2}{|(w-v)_1|} \right]$$

for  $j$ th-scale links,  $A_{j_1}^{CG}(v, w)$  for final links, and  $A_{j_2}^{CG}(v, w)$  for macroscopic links. From (4.86) and (4.87) we have

$$\frac{\delta \mu}{18} \max_{k \leq R+1} \frac{|(u^k)^*|^2}{r \sigma(r)} \leq \frac{\delta t}{3} + \frac{\delta \mu}{12 \sigma(r)} \sum_{i=1}^{R+1} \frac{|(u^{i-1} - u^i)^*|^2}{|(u^{i-1} - u^i)_1|}$$

Therefore using also (4.27) the total of these for all links in a path  $\Gamma_{xy}^{CG}$  satisfies

$$\begin{aligned} & A_{j_2}^{CG}(u^{R/2}, u^{R/2+1}) + A_{j_1}^{CG}(u^0, u^1) + A_{j_1}^{CG}(u^R, u^{R+1}) \\ & + \sum_{j=j_2}^{j_1-1} \left[ A_{j+1}^{CG}(u^{j_1-j}, u^{j_1-j+1}) + A_{j+1}^{CG}(u^{R-j_1+j}, u^{R-j_1+j+1}) \right] \\ & \leq 4 \sum_{j=j_2-1}^{j_1-1} \frac{\lambda^{j+1}}{4} \left[ \left( \frac{t}{3} + 2 \frac{\delta \mu}{18} \max_{k \leq R+1} \frac{|(u^k)^*|^2}{r \sigma(r)} \right) \sigma(r) + \frac{\delta \mu}{18} \sum_{i=1}^{R+1} \frac{|(u^{i-1} - u^i)^*|^2}{|(u^{i-1} - u^i)_1|} \right] \\ & \leq \sum_{j=j_2}^{j_1} \frac{1}{3} \lambda^j (1 + 2\delta) t \sigma(r) + \sum_{j=j_2}^{j_1} \lambda^j \left( \frac{\delta \mu}{6} + \frac{\delta \mu}{18} \right) \sum_{i=1}^{R+1} \frac{|(u^{i-1} - u^i)^*|^2}{|(u^{i-1} - u^i)_1|} \\ & \leq \frac{\lambda}{1-\lambda} t \sigma(r) + \frac{16 \delta \mu \lambda}{27(1-\lambda)} \left( \Upsilon_{Euc}(\Gamma_{xy}^{CG}) - (\hat{y} - \hat{x})_1 \right) \\ & \leq \frac{1}{3} t \sigma(r) + \frac{\mu}{2} \left( \Upsilon_{Euc}(\Gamma_{xy}^{CG}) - (\hat{y} - \hat{x})_1 \right). \end{aligned} \tag{4.174}$$

Therefore the last probability in (4.173) is bounded above by

$$\begin{aligned} & \sum_{j=1}^{j_1} P \left( \left| \hat{T}(u^{i-1}, u^i) - h(|u^{i-1} - u^i|) \right| \geq A_{j+1}^{CG}(u^{i-1}, u^i) \text{ for some } j\text{th-scale link} \right. \\ & \quad \left. (u^{i-1}, u^i) \text{ of } \Gamma_{xy}^{CG} \text{ and some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \right) \\ & + P \left( \left| \hat{T}(u^{i-1}, u^i) - h(|u^{i-1} - u^i|) \right| \geq A_{j_2}^{CG}(u^{i-1}, u^i) \text{ for some macroscopic link} \right. \\ & \quad \left. (u^{i-1}, u^i) \text{ of } \Gamma_{xy}^{CG} \text{ and some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \right) \\ & + P \left( \left| \hat{T}(u^{i-1}, u^i) - h(|u^{i-1} - u^i|) \right| \geq A_{j_1}^{CG}(u^{i-1}, u^i) \text{ for some final link} \right. \\ & \quad \left. (u^{i-1}, u^i) \text{ of } \Gamma_{xy}^{CG} \text{ and some } (x, y) \in X_r; \omega \notin J^{(0)}(c_{29}) \cup J^{(1c)} \right). \end{aligned} \tag{4.175}$$

Considering the probability for final links  $(v, w)$ , the number of possible such links in  $G_r^+$  is at most  $c_{46} r^{2d}$  and for each such link we have from Lemma 3.2 and (4.22), using  $A_{j_1}^{CG}(v, w) \geq \lambda^{j_1} t \sigma(r) / 12$ ,

$$\begin{aligned} & P \left( \left| \hat{T}(v, w) - h(|w-v|) \right| \geq A_{j_1}^{CG}(v, w) \right) \leq C_{44} \exp \left( -C_{45} \frac{A_{j_1}^{CG}(v, w)}{\sigma(2\delta^{j_1} r)} \right) \\ & \leq c_{47} \exp \left( -\frac{c_{48}}{j_1} \left( \frac{\lambda}{\delta \lambda^1} \right)^{j_1} t \right) \leq c_{47} \exp(-c_{49} t \log r). \end{aligned} \tag{4.176}$$

The probabilities for  $j$ th-scale and macroscopic links can be bounded using minor modifications of Lemma 4.6, showing that (4.175) (and hence also the probability on the right in (4.166)) is bounded above by

$$\sum_{j=1}^{j_1} c_{50} \exp\left(-c_{51} \left(\frac{\lambda}{7\delta\chi_1}\right)^j t\right) + c_{50} \exp\left(-c_{51} \frac{\lambda}{7\delta\chi_1} t\right) + c_{46}c_{47}r^{2d} \exp(-c_{49}t \log r) \leq c_{52}e^{-c_{53}t}. \tag{4.177}$$

With (4.166) this completes the proof of (1.14). As noted after Theorem (1.3), the downward-deviations part of (1.13) is a consequence, so the proof of the downward-deviations part of Theorem (1.3) is complete.

### 5 Proof of Theorem 1.6 for fixed $(x, y)$

We move on to step (2) of the strategy in Remark 1.7. As with the proof of Theorem 1.3 (downward deviations), there are simpler cases which can be proved from Lemma 3.2 and do not require Theorem 1.3. These we can handle uniformly over  $(x, y)$ ; after dealing with these, in Lemma 5.1 we will handle the main case, for the moment only for fixed  $(x, y)$ .

Let

$$r_0 = \frac{c_0 r}{(\log r)^{1/(1+\chi_1)}}, \tag{5.1}$$

with  $c_0$  to be specified; we will consider separately short geodesics (meaning  $|\hat{y} - \hat{x}| \leq r_0$ ) and longer ones. When  $\Gamma_{xy} \not\subset G_{r,s}$  for some  $x, y \in \mathbb{V} \cap G_r(K)$  with  $|(y-x)^*| \leq (y-x)_1$ , taking  $u$  to be the first vertex in  $\Gamma_{xy}$  outside  $G_{r,s}$ , we see that one of the following cases must hold.

#### Transverse wandering cases (see Figure 10):

- (i) (*non-transverse wandering*) there exist  $\hat{x}, \hat{y}, \hat{u} \in q\mathbb{Z}^d$  with  $\hat{x}, \hat{y} \in G_r(K), \hat{u} \notin G_{r,s}$  with  $d(\hat{u}, G_{r,s}) \leq 5$ , and  $x, y, u \in \mathbb{V}$  with

$$\max(|x - \hat{x}|, |y - \hat{y}|, |u - \hat{u}|) \leq 5\sqrt{d-1}, \quad u \in \Gamma_{xy}, \quad \hat{u}_1 \notin [\hat{x}_1, \hat{y}_1],$$

- (ii) (*wandering by short geodesics*) there exist  $\hat{x}, \hat{y}, \hat{u} \in q\mathbb{Z}^d$  with  $\hat{x}, \hat{y} \in G_r(K), \hat{u} \notin G_{r,s}$  with  $d(\hat{u}, G_{r,s}) \leq 5$ , and  $x, y, u \in \mathbb{V}$  with

$$\max(|x - \hat{x}|, |y - \hat{y}|, |u - \hat{u}|) \leq 5\sqrt{d-1}, \quad u \in \Gamma_{xy}, \quad \hat{u}_1 \in [\hat{x}_1, \hat{y}_1], \quad |\hat{y} - \hat{x}| \leq r_0,$$

- (iii) (*large wandering*)  $s > c_1(\log r)^{1/2}$  and there exist  $\hat{x}, \hat{y}, \hat{u} \in q\mathbb{Z}^d$  with  $\hat{x}, \hat{y} \in G_r(K), \hat{u} \notin G_{r,s}$  with  $d(\hat{u}, G_{r,s}) \leq 5$ , and  $x, y, u \in \mathbb{V}$  with

$$\max(|x - \hat{x}|, |y - \hat{y}|, |u - \hat{u}|) \leq 5\sqrt{d-1}, \quad |y - x| > r_0, \quad u \in \Gamma_{xy}, \quad u_1 \in [x_1, y_1], \tag{5.2}$$

- (iv) (*moderate wandering*)  $s \leq c_1(\log r)^{1/2}$  and there exist  $x, y, u \in \mathbb{V}$  with

$$x, y \in G_r(K), \quad u \notin G_{r,s}, \quad d(u, G_{r,s}) \leq 2, \quad |y - x| > r_0, \quad u \in \Gamma_{xy}, \quad u_1 \in [x_1, y_1], \tag{5.3}$$

where  $c_1$  is to be specified. In all cases (i)–(iv) we assume  $|(y-x)^*| \leq (y-x)_1$ .

Cases (i)–(iii) are the ones which do not require Theorem 1.3 and can be proved from Lemma 3.2. This is because the exponent obtained from that lemma is at least of order  $\log r$ , meaning it dominates the entropy from the number of possible choices of  $\hat{x}, \hat{y}, \hat{u}$ .

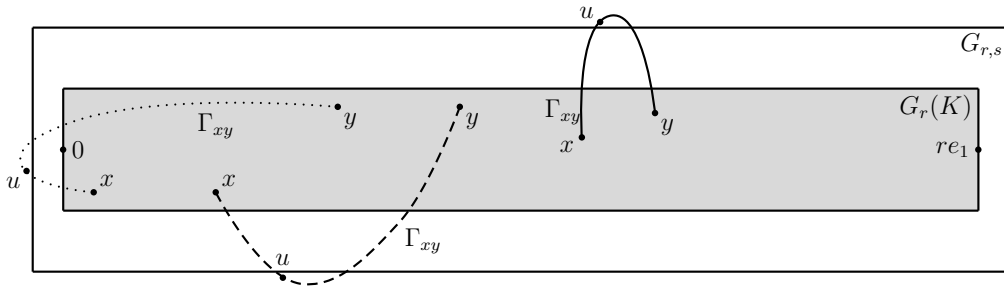


Figure 10: The inner box  $G_r(K)$  has height  $2K\Delta(r)$ , and the outer box  $G_{r,s}$  has height  $2s\Delta(r)$ . The dotted line shows  $\Gamma_{xy}$  in transverse wandering case (i), the solid line shows case (ii), and the dashed line shows cases (iii) and (iv), for large and small  $s$ , respectively.

It follows from routine geometry that under (i), for  $C_{34}$  from (1.17) we have

$$|\hat{u} - \hat{x}| + |\hat{y} - \hat{u}| - |\hat{y} - \hat{x}| \geq d(G_r(K), G_{r,s}^c) \geq \begin{cases} s^2\sigma(r) \log r & \text{if } s \leq (C_{34} \log r)^{1/2}, \\ s^2\sigma(r) & \text{if } (C_{34} \log r)^{1/2} < s \leq r/\Delta(r), \\ s\Delta(r) & \text{if } s > r/\Delta(r) \end{cases}$$

so also, using Lemma 3.6(ii),

$$h(|\hat{u} - \hat{x}|) + h(|\hat{y} - \hat{u}|) - h(|\hat{y} - \hat{x}|) \geq \begin{cases} \frac{\mu}{2} s^2\sigma(r) \log r & \text{if } s \leq (C_{34} \log r)^{1/2}, \\ \frac{\mu}{2} s^2\sigma(r) & \text{if } (C_{34} \log r)^{1/2} < s \leq r/\Delta(r), \\ \frac{\mu}{2} s\Delta(r) & \text{if } s > r/\Delta(r), \end{cases} \quad (5.4)$$

while

$$\left| \hat{T}(\hat{x}, \hat{u}) + \hat{T}(\hat{u}, \hat{y}) - \hat{T}(\hat{x}, \hat{y}) \right| \leq M(\hat{x}) + M(\hat{y}) + M(\hat{u}).$$

Therefore one of the following 6 quantities

$$\left| \hat{T}(\hat{x}, \hat{u}) - h(|\hat{u} - \hat{x}|) \right|, \quad \left| \hat{T}(\hat{u}, \hat{y}) - h(|\hat{y} - \hat{u}|) \right|, \quad \left| \hat{T}(\hat{x}, \hat{y}) - h(|\hat{y} - \hat{x}|) \right|, \quad M(\hat{x}), \quad M(\hat{y}), \quad M(\hat{u})$$

exceeds  $1/6$  of the right side of (5.4). There are at most  $c_2 s^{d_r 2d}$  possible values of  $(\hat{x}, \hat{y}, \hat{u})$  under (i), so from Lemmas 3.1 and 3.2, provided  $C_{34}$  is sufficiently large we have

$$P\left( (i) \text{ holds} \right) \leq \begin{cases} c_2 s^{d_r 2d} \cdot C_{44} \exp\left(-c_3 \frac{s^2\sigma(r) \log r}{\sigma(2r)}\right) \leq c_4 e^{-c_5 s^2} & \text{if } s \leq (C_{34} \log r)^{1/2} \\ c_2 s^{d_r 2d} \cdot C_{44} \exp\left(-c_3 \frac{s^2\sigma(r)}{\sigma(2r)}\right) \leq c_4 e^{-c_5 s^2} & \text{if } (C_{34} \log r)^{1/2} < s \leq r/\Delta(r), \\ c_2 s^{d_r 2d} \cdot C_{44} \exp\left(-c_3 \frac{s\Delta(r)}{\sigma(s\Delta(r))}\right) \leq c_4 e^{-c_5 s\Delta(r)/\sigma(s\Delta(r))} & \text{if } s > r/\Delta(r). \end{cases} \quad (5.5)$$

Under (ii) we have for some  $c_6$

$$|\hat{u} - \hat{x}| + |\hat{y} - \hat{u}| - |\hat{y} - \hat{x}| \geq \begin{cases} c_6 s^2 \Delta(r)^2 / r_0 & \text{if } s \leq r_0 / \Delta(r), \\ c_6 s \Delta(r) & \text{if } s > r_0 / \Delta(r). \end{cases}$$

Further, for  $s \leq r/\Delta(r)$  we have  $1/s\Delta(r)\sigma(s\Delta(r)) \leq c_7/\Delta(r)^2$  or equivalently  $s\Delta(r)/\sigma(s\Delta(r)) \geq c_7 s^2$ . Hence similarly to (i), provided we take  $c_0$  large in (5.1),

$$P\left( (ii) \text{ holds} \right)$$

$$\leq \begin{cases} c_8 s^d r^{2d} \cdot C_{44} \exp\left(-c_9 \frac{s^2 r \sigma(r)}{r_0 \sigma(r_0)}\right) \leq c_{10} e^{-c_{11} s^2 \log r} & \text{if } s \leq r_0/\Delta(r) \\ c_8 s^d r^{2d} \cdot C_{44} \exp\left(-c_9 \frac{s \Delta(r)}{\sigma(s \Delta(r))}\right) \leq c_{10} e^{-c_{11} s^2} & \text{if } r_0/\Delta(r) < s \leq r/\Delta(r), \\ c_8 s^d r^{2d} \cdot C_{44} \exp\left(-c_9 \frac{s \Delta(r)}{\sigma(s \Delta(r))}\right) \leq c_{10} e^{-c_{11} s \Delta(r)/\sigma(s \Delta(r))} & \text{if } s > r/\Delta(r). \end{cases} \tag{5.6}$$

Under (iii) we have

$$|\hat{u} - \hat{x}| + |\hat{y} - \hat{u}| - |\hat{y} - \hat{x}| \geq \begin{cases} c_{12} s^2 \sigma(r) & \text{if } c_1 (\log r)^{1/2} \leq s \leq r/\Delta(r), \\ c_{12} s \Delta(r) & \text{if } s > r/\Delta(r). \end{cases}$$

Hence as in (i) and (ii) we get

$$\begin{aligned} &P\left(\text{(iii) holds}\right) \\ &\leq \begin{cases} c_{13} s^d r^{2d} \cdot C_{44} \exp\left(-c_{14} \frac{s^2 \sigma(r)}{\sigma(2r)}\right) \leq c_{15} e^{-c_{16} s^2} & \text{if } c_1 (\log r)^{1/2} \leq s \leq r/\Delta(r) \\ c_{13} s^d r^{2d} \cdot C_{44} \exp\left(-c_{14} \frac{s \Delta(r)}{\sigma(s \Delta(r))}\right) \leq c_{14} e^{-c_{15} s \Delta(r)/\sigma(s \Delta(r))} & \text{if } s > r/\Delta(r). \end{cases} \end{aligned} \tag{5.7}$$

To deal with (iv) and complete the proof for fixed  $(x, y)$ , we have the following lemma. We note the difference from Lemma 4.2, which covered many  $x, y$  simultaneously but considered only larger transverse wandering, of order  $\Delta(r) \log r$  or more. Note that so far we have not had to impose any conditions on  $c_1$  appearing in (iii) and (iv), so we can take  $c_1$  to be  $C_{78}$  from the lemma.

**Lemma 5.1.** *Suppose  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  and  $\{\eta_e, e \in \mathbb{E}\}$  satisfy A1, A2, and A3. There exist  $C_i$  such that for all  $K \geq C_{74}$  and all  $x, y \in G_r(K)$  with  $|(y - x)^*| \leq (y - x)_1$ ,*

$$P\left(\Gamma_{xy} \not\subset G_{r,s}\right) \leq C_{75} e^{-C_{76} s^2} \quad \text{for all } C_{77} K \leq s \leq C_{78} (\log r)^{1/2}. \tag{5.8}$$

*Proof.* Fix  $x, y \in G_r(K)$  and let  $r_1 = (y - x)_1$ . We consider first the case of  $r_1 < r/2$ . To facilitate application of (1.14) in Theorem 1.3, we need to replace  $G_{r,s}$  with a cylinder of length only  $2r_1$ . To that end, let  $m \geq 0$  satisfy  $mr_1 \leq x_1 < y_1 \leq (m + 2)r_1$ , and for  $C_{22}$  from (1.11) let

$$s_1 = 2C_{22} s \left(\frac{r}{2r_1}\right)^{(1+\chi_1)/2}, \quad G_{r,2s}^{(m)} = [mr_1, (m + 2)r_1] \times 2s \Delta(r) \mathfrak{B}_{d-1},$$

the latter being a slice of  $G_{r,2s}$ . Define  $s_2$  by  $2s \Delta(r) = s_2 \Delta(2r_1)$ . Then  $s_2 \geq s$ , and from (1.11) we have  $s_2 \geq s_1$ , and  $G_{r,2s}^{(m)}$  is a translate of  $G_{2r_1}(s_2)$ .

In view of (5.5) and (5.6), we need only consider  $(x, y, u)$  as in (5.3); in particular  $r_1 \geq r_0/2$ , with  $r_0$  from (5.1). Here we have for some  $c_0$

$$\begin{aligned} &|u - x| + |y - u| - |y - x| \geq c_0 s_2^2 \sigma(2r_1) \\ &\text{and hence } h(|u - x|) + h(|y - u|) - h(|y - x|) \geq \frac{c_0 \mu}{2} s_2^2 \sigma(2r_1), \end{aligned}$$

so as in transverse wandering cases (i)–(iii), since  $u \in \Gamma_{xy}$  one of the quantities

$$h(|u - x|) - T(x, u), \quad h(|y - u|) - T(u, y), \quad T(x, y) - h(|y - x|)$$

must exceed  $c_0 \mu s_2^2 \sigma(2r_1)/6$ . It follows from (1.12) and the downward-deviations part of Theorem 1.3 (with  $\epsilon, r, K, t$  there taken as  $1/2, 2r_1, K s_2/2s, c_0 \mu s_2^2/6$  here) that for  $x_0 = x - mr_1 e_1, y_0 = y - mr_1 e_1$  we have

$P(\Gamma_{xy}$  contains a vertex  $u$  as in (5.3))



$$\begin{aligned}
 &\leq P\left(\max\left(h(|u-x|) - T(x,u), h(|y-u|) - T(u,y)\right) \geq \frac{c_0\mu}{6}s_2^2\sigma(2r_1) \text{ for some } u \in G_{r,2s}^{(m)}\right) \\
 &\quad + P\left(T(x,y) - h(|y-x|) \geq \frac{c_0\mu}{6}s_2^2\sigma(2r_1)\right) \\
 &\leq P\left(\max\left(h(|\tilde{u}-x_0|) - T(x_0,\tilde{u}), h(|y_0-\tilde{u}|) - T(\tilde{u},y_0)\right) \right. \\
 &\quad \left. \geq \frac{c_0\mu}{6}s_2^2\sigma(2r_1) \text{ for some } \tilde{u} \in G_{2r_1}(s_2)\right) \\
 &\quad + P\left(T(x,y) - h(|y-x|) \geq \frac{c_0\mu}{6}s_2^2\sigma(2r_1)\right) \\
 &\leq C_{27}e^{-c_1s_2^2} \\
 &\leq C_{27}e^{-c_1s^2}. \tag{5.9}
 \end{aligned}$$

The proof in the case  $r_1 \geq r/2$  is the same but without the need to shorten  $G_{r,s}$ .  $\square$

This completes the proof of Theorem 1.6 for fixed  $(x, y)$ , and thus of (2) in Remark 1.7.

### 6 Proof of Theorem 1.5

We move to (3) of Remark 1.7. We begin with an extension of a consequence of Theorem 1.3, removing the requirement that  $y$  lie in the same tube  $G_r(K)$  as  $x$ , to create a bound on the probability of a fast “hyperplane-block to hyperplane” passage time. Define

$$\Lambda_r = [0, 2] \times [-\Delta(r), \Delta(r)]^{d-1}.$$

**Lemma 6.1.** *Suppose  $G = (V, E)$  and  $\{\eta_e, e \in E\}$  satisfy A1, A2, and A3. There exist constants  $C_i$  such that for all  $r \geq C_{79}$  and  $t \geq C_{80}$ ,*

$$P\left(T(x, y) \leq h(r) - t\sigma(r) \text{ for some } x \in \Lambda_r \cap V \text{ and } y \in H_r^+\right) \leq C_{81}e^{-C_{82}t}. \tag{6.1}$$

*Proof.* It is enough to consider  $y \in H_{[r,r+2]}$  and  $t \leq h(r)/\sigma(r)$ . We first deal with large  $|y^*|$ . With  $c_0$  to be specified we have

$$\begin{aligned}
 &P\left(T(x, y) \leq h(r) - t\sigma(r) \text{ for some } x \in \Lambda_r \cap V \text{ and } y \in H_{[r,r+2]} \text{ with } |y^*| > c_0(\log r)^{1/2}\Delta(r)\right) \\
 &\leq \sum_{k=1}^{\infty} P\left(T(x, y) \leq h(r) - t\sigma(r) \text{ for some } x \in \Lambda_r \cap V \text{ and } y \in H_{[r,r+2]} \right. \\
 &\quad \left. \text{with } 2^{k-1}c_0(\log r)^{1/2}\Delta(r) < |y^*| \leq 2^k c_0(\log r)^{1/2}\Delta(r)\right) \\
 &\leq \sum_{k=1}^{\infty} P\left(\hat{T}(\hat{x}, \hat{y}) \leq h(r) - t\sigma(r) \text{ for some } \hat{x} \in \Lambda_r \cap q\mathbb{Z}^d \text{ and } \hat{y} \in H_r \cap q\mathbb{Z}^d \right. \\
 &\quad \left. \text{with } 2^{k-1}c_0(\log r)^{1/2}\Delta(r) < |\hat{y}^*| \leq 2^k c_0(\log r)^{1/2}\Delta(r)\right). \tag{6.2}
 \end{aligned}$$

For  $c_9$  from (4.26), let  $k_0 = \max\{k : 2^k c_0(\log r)^{1/2}\Delta(r) \leq c_9 r\}$ . For  $k \leq k_0$ , and for  $\hat{x}, \hat{y}$  of the last event, we have from (4.26) and Lemma 3.6(ii) that

$$|\hat{y} - \hat{x}| \geq r + \frac{|\hat{y}^*|^2}{3r} \geq r + \frac{2^{2k}c_0^2\sigma(r)\log r}{12} \quad \text{so} \quad h(|\hat{y} - \hat{x}|) \geq h(r) + \frac{\mu C_0^2}{24} 2^{2k}\sigma(r)\log r, \tag{6.3}$$

and the number of allowed  $\hat{x}, \hat{y}$  is at most  $c_1(2^k(\log r)^{1/2})^{d-1}\Delta(r)^{2(d-1)} \leq 2^{kd}r^{2d}$ . Therefore by Lemma 3.2, provided  $c_0$  is large, the sum up to  $k_0$  on the right side of (6.2) is bounded above by

$$\begin{aligned} & \sum_{k=1}^{k_0} P\left(\hat{T}(\hat{x}, \hat{y}) \leq h(|\hat{y} - \hat{x}|) - t\sigma(r) - \frac{\mu c_0^2}{24} 2^{2k} \sigma(r) \log r \text{ for some } \hat{x} \in \Lambda_r \cap q\mathbb{Z}^d \right. \\ & \quad \left. \text{and } \hat{y} \in H_r \cap q\mathbb{Z}^d \text{ with } 2^{k-1}c_0(\log r)^{1/2}\Delta(r) < |\hat{y}^*| \leq 2^k c_0(\log r)^{1/2}\Delta(r)\right) \\ & \leq \sum_{k=1}^{k_0} 2^{kd}r^{2d}C_{44} \exp\left(-C_{45} \frac{\sigma(r)}{\sigma(2r)} \left(t + 2^{2k} \frac{\mu c_0^2}{24} \log r\right)\right) \\ & \leq c_2 e^{-c_3 t}. \end{aligned} \tag{6.4}$$

For  $k > k_0$ , in place of (6.3) we have

$$|\hat{y} - \hat{x}| \geq r + c_4 |\hat{y}^*| \geq r + c_4 2^{k-1} c_0 (\log r)^{1/2} \Delta(r) \quad \text{so} \quad h(|\hat{y} - \hat{x}|) \geq h(r) + \frac{c_4 c_0 \mu}{4} 2^k (\log r)^{1/2} \Delta(r),$$

and we obtain similarly that the sum from  $k_0 + 1$  to  $\infty$  on the right side of (6.2) is bounded above by

$$\begin{aligned} & \sum_{k=k_0+1}^{\infty} P\left(\hat{T}(\hat{x}, \hat{y}) \leq h(|\hat{y} - \hat{x}|) - t\sigma(r) - \frac{c_4 c_0 \mu}{4} 2^k (\log r)^{1/2} \Delta(r) \text{ for some } \hat{x} \in \Lambda_r \cap q\mathbb{Z}^d \text{ and} \right. \\ & \quad \left. \hat{y} \in H_r \cap q\mathbb{Z}^d \text{ with } 2^{k-1}c_0(\log r)^{1/2}\Delta(r) < |\hat{y}^*| \leq 2^k c_0(\log r)^{1/2}\Delta(r)\right) \\ & \leq \sum_{k=k_0+1}^{\infty} 2^{kd}r^{2d}C_{44} \exp\left(-\frac{C_{45}}{\sigma(r) + 2^k c_0 (\log r)^{1/2} \Delta(r)} \left(t\sigma(r) + \frac{c_4 c_0 \mu}{4} 2^k (\log r)^{1/2} \Delta(r)\right)\right) \\ & \leq c_5 \exp\left(-\frac{c_6(t\sigma(r) + r)}{\sigma(2r)}\right) \\ & \leq c_2 e^{-c_3 t}. \end{aligned} \tag{6.5}$$

The last inequality uses  $t \leq h(r)/\sigma(r)$ , which we may assume, as noted above.

We now deal with the remaining case  $|y^*| \leq c_0(\log r)^{1/2}\Delta(r)$ . For  $C_{26}$  from Theorem 1.3, we take  $K = \sqrt{t/C_{26}(d-1)}$  and subdivide  $H_{[r, r+2]}$  into blocks of side  $K\Delta(r)$ :

$$Y_{\mathbf{m}} = [r, r+2] \times \prod_{i=1}^{d-1} [(m_i - 1)K\Delta(r), m_i K\Delta(r)], \quad \mathbf{m} \in \mathbb{Z}^{d-1}.$$

Let  $z_{\mathbf{m}}$  denote the center point of  $Y_{\mathbf{m}}$  and let

$$\mathfrak{M} = \left\{ \mathbf{m} \in \mathbb{Z}^{d-1} : 0 \notin Y_{\mathbf{m}}, \exists y \in Y_{\mathbf{m}} \text{ with } |y^*| \leq c_0(\log r)^{1/2}\Delta(r) \right\}.$$

Given  $\mathbf{m} \in \mathfrak{M}$ , there exists a cylinder  $\mathcal{C}_{\mathbf{m}}$ , with axis containing  $\Pi_{0, z_{\mathbf{m}}}$ , radius  $2\sqrt{d-1}K\Delta(r)$ , and length  $2r$ , which contains  $\Lambda_r$  and  $Y_{\mathbf{m}}$ . See Figure 11. Further,  $d_{\infty}(Y_{\mathbf{m}}, [r, r+2] \times \{0\})$  is a positive integer multiple of  $K\Delta(r)$ , and for  $x \in \Lambda_r$  and  $y \in Y_{\mathbf{m}}$  satisfying  $d_{\infty}(Y_{\mathbf{m}}, [r, r+2] \times \{0\}) = jK\Delta(r)$  for some  $j$ , we have using (4.26) and Lemma 3.6(ii) that

$$|y - x| \geq r + \frac{(jK\Delta(r))^2}{3r} = r + \frac{(jK)^2}{3}\sigma(r) \quad \text{and hence} \quad h(|y - x|) \geq h(r) + \frac{\mu(jK)^2}{6}\sigma(r).$$

The definition of  $K$  says  $t = C_{26}(d-1)K^2$ , so by rotational invariance we can now apply

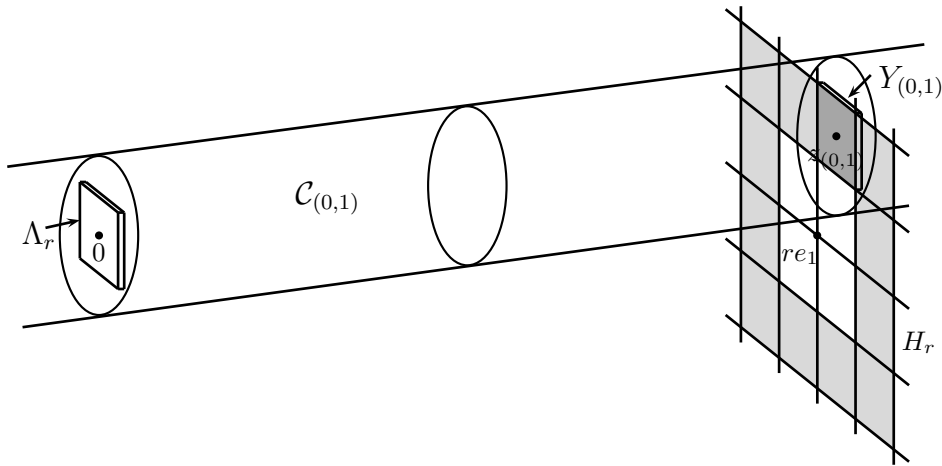


Figure 11: The block  $\Lambda_r$  in  $H_0$  (“fattened” slightly to thickness 2), the similar but smaller block  $Y_{\mathbf{m}}$  of  $H_r$ , and the cylinder  $\mathcal{C}_{\mathbf{m}}$  containing both, for  $\mathbf{m} = (0, 1)$ . The gray blocks in  $H_r$  are those with  $j = 1$ .

Theorem 1.3 (for downward deviations) to the cylinder  $\mathcal{C}_{\mathbf{m}}$  and obtain

$$\begin{aligned}
 &P\left(T(x, y) \leq h(r) - t\sigma(r) \text{ for some } x \in \Lambda_r \cap \mathbb{V} \text{ and } y \in Y_{\mathbf{m}}\right) \\
 &\leq P\left(T(x, y) \leq h(|y - x|) - \left(t + \frac{\mu(jK)^2}{6}\right)\sigma(r) \text{ for some } x \in \Lambda_r \cap \mathbb{V} \text{ and } y \in Y_{\mathbf{m}}\right) \\
 &\leq C_{27} \exp\left(-C_{28} \left(t + \frac{\mu(jK)^2}{6}\right)\right). \tag{6.6}
 \end{aligned}$$

Summing over  $j \leq j_0 = \max\{j : jK\Delta(r) \leq c_0(\log r)^{1/2}\Delta(r)\}$  gives

$$\begin{aligned}
 &P\left(T(x, y) \leq h(r) - t\sigma(r) \text{ for some } x \in \Lambda_r \cap \mathbb{V}, y \in H_{[r, r+2]} \text{ with } |y^*| \leq c_0(\log r)^{1/2}\Delta(r)\right) \\
 &\leq P\left(T(x, y) \leq h(r) - t\sigma(r) \text{ for some } x \in \Lambda_r \cap \mathbb{V}, y \in Y_{\mathbf{m}}, \text{ and } \mathbf{m} \in \mathfrak{M}\right. \\
 &\quad \left. \text{with } d_\infty(Y_{\mathbf{m}}, [r, r+2] \times \{0\}) \leq j_0K\Delta(r)\right) \\
 &\leq \sum_{j=0}^{j_0} P\left(T(x, y) \leq h(r) - t\sigma(r) \text{ for some } x \in \Lambda_r \cap \mathbb{V}, y \in Y_{\mathbf{m}}, \text{ and } \mathbf{m} \in \mathfrak{M}\right. \\
 &\quad \left. \text{with } d_\infty(Y_{\mathbf{m}}, [r, r+2] \times \{0\}) = jK\Delta(r)\right) \\
 &\leq \sum_{j=0}^{j_0} (2j+2)^{d-1} C_{27} \exp\left(-C_{28} \left(t + \frac{\mu(jK)^2}{6}\right)\right) \\
 &\leq c_7 e^{-c_8 t}. \tag{6.7}
 \end{aligned}$$

With (6.2), (6.4), and (6.5) this completes the proof of Lemma 6.1. □

Define the “small box”

$$\mathcal{G}_r = [0, r] \times [-\Delta(r), \Delta(r)]^{d-1}.$$

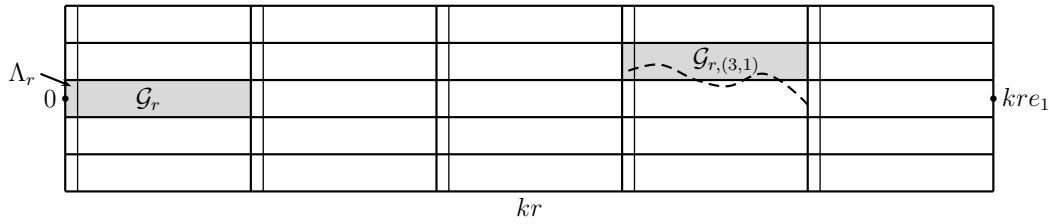


Figure 12: The big box  $\mathcal{G}^+$ , of size  $kr \times (4j+2)\Delta(r) \approx kr \times K\Delta(kr)$ , here with  $k = 5, j = 2$ . Each small box  $\mathcal{G}_{r,m}$  has size  $r \times 2\Delta(r)$ .  $\Lambda_{r,(3,1)}$  is a fast source if some path like the dashed one is “sufficiently fast” relative to  $h(r)$ .

We now tile a region of  $\mathbb{R}^d$  with translates of  $\mathcal{G}_r$ . Fix  $k, K$  to be specified and let  $j = j(r, k, K)$  be the least  $j$  for which  $(2j + 1)\Delta(r) \geq K\Delta(kr)$ . Define the “big box”

$$\mathcal{G}^+ = [0, kr] \times [-(2j + 1)\Delta(r), (2j + 1)\Delta(r)]^{d-1},$$

so the width of  $\mathcal{G}^+$  is near  $2K\Delta(kr)$ . For  $\mathbf{m} \in \mathbb{Z}^d$  let  $\mathcal{G}_{r,m}, \Lambda_{r,m}$  be  $\mathcal{G}_r, \Lambda_r$  translated by  $(rm_1, 2\Delta(r)\mathbf{m}^*)$ , and let  $\mathcal{M} = \{\mathbf{m} \in \mathbb{Z}^d : \mathcal{G}_{r,m} \subset \mathcal{G}^+\}$ ; see Figure 12. Then the number of small boxes comprising  $\mathcal{G}^+$  is

$$|\mathcal{M}| = k(2j + 1)^{d-1} \leq 2k \left( \frac{K\Delta(kr)}{\Delta(r)} \right)^{d-1} \leq c_0 K^{d-1} k^{1+(d-1)\chi_2},$$

and we have  $\mathcal{G}^+ \supset G_{kr,K} \cap H_{[0,kr]}$ . Let  $\mathcal{E}_{\mathbf{m}}(t)$  be the event in (6.1) translated to  $\mathcal{G}_{r,m}$ , that is, we replace  $\Lambda_r$  with  $\Lambda_{r,m}$  and  $H_r^+$  with  $H_{(m_1+1)r}^+$ . Fix  $C$  to be specified; when  $\omega \in \mathcal{E}_{\mathbf{m}}(C \log k)$  we say  $\Lambda_{r,m}$  is a fast source. Let  $\mathcal{L}(r) = (h(r) - r\mu)/\sigma(r)$ , so

$$0 \leq \mathcal{L}(r) \leq C_{46} \log r \tag{6.8}$$

by Proposition 3.3.

We now consider the passage time between the ends of the big box. If  $\Gamma_{xy} \subset G_{kr,K}$  then  $\Gamma_{0,kre_1}$  must intersect at least  $k$  of the regions  $\Lambda_{r,m}$ , one for each  $0 \leq m_1 < k$ . If there are no fast sources, this means  $T(0, kre_1) > kr\mu + k\mathcal{L}(r)\sigma(r) - C\sigma(r)k \log k$ . Provided  $K$  is large, it then follows from Lemmas 5.1 and 6.1 that

$$\begin{aligned} P\left(T(0, kre_1) \leq kr\mu + k\mathcal{L}(r)\sigma(r) - C\sigma(r)k \log k\right) & \\ \leq P\left(\Gamma_{0,kre_1} \not\subset G_{kr,K}\right) + P\left(\cup_{\mathbf{m} \in \mathcal{M}} \mathcal{E}_{\mathbf{m}}(C \log k)\right) & \\ \leq C_{75}e^{-C_{76}K^2} + C_{81}|\mathcal{M}|e^{-C_{82}C \log k} & \\ \leq C_{75}e^{-C_{76}K^2} + c_1K^{d-1}k^{1+(d-1)\chi_2}e^{-C_{82}C \log k}. & \end{aligned} \tag{6.9}$$

By taking  $K$  large and then  $C$  large, we can make the right side of (6.9) less than 1/2 for all  $k \geq 3$ . On the other hand, from (1.12) for large  $c > 1$  we have

$$P\left(T(0, kre_1) \leq kr\mu + \mathcal{L}(kr)\sigma(kr) + c\sigma(kr)\right) \geq 1 - C_{24}e^{-C_{25}c} > \frac{1}{2}.$$

From this and (6.9) it follows that

$$(\mathcal{L}(kr) + c)\sigma(kr) > k(\mathcal{L}(r) - C \log k)\sigma(r) \quad \text{for all } r \text{ large and } k \geq 3. \tag{6.10}$$

Fix  $k \geq 3$  large enough so  $k^{1-\chi_2} \geq 3C_{23}c$ , with  $C_{23}$  from (1.11), and, to get a contradiction, suppose there exists  $r$  arbitrarily large with  $\mathcal{L}(r) \geq 2C \log k$ . By (1.11) and (6.10) we then have

$$\mathcal{L}(kr) \geq \frac{\sigma(r)}{\sigma(kr)} \frac{k\mathcal{L}(r)}{2} - c \geq \frac{k^{1-\chi_2}}{3C_{23}} \mathcal{L}(r) \geq c\mathcal{L}(r).$$

Then iteration and (6.8) give that for all  $n \geq 1$ ,  $C_{46}(n \log k + \log r) \geq \mathcal{L}(k^n r) \geq c^n \mathcal{L}(r)$ . Since this is false for large  $n$ , we must have  $\mathcal{L}(r) < 2C \log k$  for all large  $r$ , so  $\mathcal{L}(\cdot)$  is bounded, which completes the proof of Theorem 1.5, and thus (3) of Remark 1.7.

### 7 Proof of Theorem 1.3—upward deviations

We move on to (4) of Remark 1.7. The proof is similar to the LPP proof for  $d = 2$  in ([13] Proposition 10.1). As with downward deviations, we need only consider  $x, y$  as in (4.24). Let  $\mathcal{H}_{xy}$  be as in the downward-deviations proof, and

$$\mathcal{H}_{xy}^{ter} = \{\text{all terminal hyperplanes in } \mathcal{H}_{xy}\},$$

so  $\mathcal{H}_{xy}^{ter} = \{H_{s_1}, \dots, H_{s_n}\}$  for some  $n, s_1, \dots, s_n$  depending on  $(x, y)$ . Note that  $\mathcal{H}_{xy}^{ter}$  does not depend on  $\Gamma_{xy}$ ; the  $j$ -terminal hyperplane at each end of  $[\hat{x}_1, \hat{y}_1]$  is just the second-closest  $j$ th-scale hyperplane to the endpoint. We form a marked PG path

$$\Omega_{xy} : \hat{x} = u^0 \rightarrow \dots \rightarrow u^{n+1} = \hat{y}$$

by taking  $u^i$  as the closest point to  $w^i = \Pi_{\hat{x}\hat{y}} \cap H_{s_i}$  in the  $j$ th-scale grid, when  $H_{s_i}$  is a  $j$ -terminal hyperplane; this means  $|u^i - w^i| \leq \sqrt{d-1} \beta^j \Delta(r)$  and  $n \leq 2j_1$  (from (4.22)). Thus the CG approximation gets coarser as we move away from the endpoints of the interval. See Figure 3.

We proceed generally as in Step 8 of the proof for downward deviations (Section 4.10). Let  $w_\perp^i$  be the orthogonal projection of  $u^i$  into  $\Pi_{\hat{x}\hat{y}}$ . We use the terminology  $j$ th-scale link, final link, and macroscopic link from there. For  $j \leq j_1$ , for a  $j$ th-scale link  $(u^{i-1}, u^i)$  we have  $(u^i - u^{i-1})_1 \geq (1 - \delta) \delta^j r$  and

$$|u^i - u^{i-1}| \leq |w_\perp^i - w_\perp^{i-1}| + \frac{(|u^{i-1} - w_\perp^{i-1}| + |u^i - w_\perp^i|)^2}{\delta^j r} \leq |w_\perp^i - w_\perp^{i-1}| + \frac{2(d-1)\beta^{2j}\Delta(r)^2}{\delta^j r}. \tag{7.1}$$

Further, the angle  $\alpha$  between  $e_1$  and  $\hat{y} - \hat{x}$  satisfies  $\tan \alpha \leq 2K\Delta(r)/\epsilon r$  and therefore

$$|w^i - w_\perp^i| \leq c_0 \frac{K\Delta(r)}{\epsilon r} |u^i - w^i| \leq c_1 \frac{K\beta^j}{\epsilon} \sigma(r), \tag{7.2}$$

and similarly for  $i-1$  in place of  $i$ . We have also  $\delta^j r/2 \leq (u^i - u^{i-1})_1 \leq 2\delta^j r$  so from (4.26) and the bound on  $\tan \alpha$  we also get

$$|w^i - w^{i-1}| - (w^i - w^{i-1})_1 = \frac{(w^i - w^{i-1})_1}{(\hat{y} - \hat{x})_1} \left( |\hat{y} - \hat{x}| - (\hat{y} - \hat{x})_1 \right) \leq \frac{2\delta^j}{\epsilon} \frac{2K^2}{\epsilon} \sigma(r).$$

Combining this with (4.19), (7.1), and (7.2) and using  $(u^i - u^{i-1})_1 = (w^i - w^{i-1})_1$  we see that

$$\begin{aligned} & |u^i - u^{i-1}| \\ & \leq (u^i - u^{i-1})_1 + \left( |u^i - u^{i-1}| - |w_\perp^i - w_\perp^{i-1}| \right) + |w^i - w_\perp^i| + |w^{i-1} - w_\perp^{i-1}| \\ & \quad + \left( |w^i - w^{i-1}| - (w^i - w^{i-1})_1 \right) \\ & \leq (u^i - u^{i-1})_1 + \left( \frac{2(d-1)\beta^{2j}}{\delta^j} + \frac{2c_1 K \beta^j}{\epsilon} + \frac{4\delta^j K^2}{\epsilon^2} \right) \sigma(r) \leq (u^i - u^{i-1})_1 + \frac{8\delta^j K^2}{\epsilon^2} \sigma(r) \end{aligned}$$

and therefore by Theorem 1.5, using (4.19) and the assumption  $t \geq C_{26} K^2$ ,

$$\begin{aligned}
 & h(|u^i - u^{i-1}|) \\
 & \leq h((u^i - u^{i-1})_1) + \frac{16\mu\delta^j K^2}{\epsilon^2} \sigma(r) \leq \mu(u^i - u^{i-1})_1 + 2C_{33}\sigma(2\delta^j r) + \frac{16\mu\delta^j K^2}{\epsilon^2} \sigma(r) \\
 & \leq \mu(u^i - u^{i-1})_1 + \frac{c_2\delta^{j\lambda_1} K^2}{\epsilon^2} \sigma(r) \leq \mu(u^i - u^{i-1})_1 + \frac{\lambda^j}{2} t\sigma(r). \tag{7.3}
 \end{aligned}$$

It follows using Lemma 3.2 that

$$\begin{aligned}
 P\left(\hat{T}(u^{i-1}, u^i) - \mu(u^i - u^{i-1})_1 \geq \lambda^j t\sigma(r)\right) & \leq P\left(\hat{T}(u^{i-1}, u^i) - h(|u^i - u^{i-1}|) \geq \frac{\lambda^j}{2} t\sigma(r)\right) \\
 & \leq C_{44} \exp\left(-C_{45} \frac{\lambda^j t\sigma(r)}{\sigma(2\delta^j r)}\right) \\
 & \leq C_{44} \exp\left(-c_3 \left(\frac{\lambda}{\delta\lambda_1}\right)^j t\right). \tag{7.4}
 \end{aligned}$$

The same is valid for a final link, with  $j = j_1$ . For macroscopic links it is valid with  $\sigma(2\delta^j r)$  replaced by  $\sigma(r)$ , giving

$$P\left(\hat{T}(u^{i-1}, u^i) - \mu(u^i - u^{i-1})_1 \geq \lambda t\sigma(r)\right) \leq C_{44} e^{-c_4 \lambda t}. \tag{7.5}$$

The numbers of possible links inside  $G_r(K)$  are as follows:

- (i)  $j$ th-scale links: at most  $c_5 K^{2(d-1)} / \delta^{2j} \beta^{2j(d-1)}$ ,
- (ii) final links: at most  $c_6 r^{2d}$ ,
- (iii) macroscopic links: at most  $c_7 K^{2(d-1)}$ .

Since  $t \geq C_{26} K^2$ , provided  $C_{26}$  is taken large enough it follows from these bounds and (4.20), (4.22), (7.4), (7.5) that

$$\begin{aligned}
 & P\left(\text{for some } \hat{x}, \hat{y} \in G_r(K) \text{ and } j \leq j_1, \text{ there is a } j\text{th-scale link } (v, w) \text{ in } \Omega_{xy} \right. \\
 & \quad \left. \text{satisfying } \hat{T}(v, w) - \mu(w - v)_1 \geq \lambda^j t\sigma(r)\right) \leq c_8 \exp\left(-c_{10} \left(\frac{\lambda}{\delta\lambda_1}\right)^j t\right), \\
 & P\left(\text{for some } \hat{x}, \hat{y} \in G_r(K) \text{ there is a final link } (v, w) \text{ in } \Omega_{xy} \right. \\
 & \quad \left. \text{satisfying } \hat{T}(v, w) - \mu(w - v)_1 \geq \lambda^{j_1} t\sigma(r)\right) \leq c_8 \exp\left(-c_{10} \left(\frac{\lambda}{\delta\lambda_1}\right)^{j_1} t\right), \\
 & P\left(\text{for some } \hat{x}, \hat{y} \in G_r(K) \text{ the macroscopic link } (v, w) \text{ in } \Omega_{xy} \right. \\
 & \quad \left. \text{satisfies } \hat{T}(v, w) - \mu(w - v)_1 \geq \lambda t\sigma(r)\right) \leq c_8 e^{-c_{10} \lambda t}. \tag{7.6}
 \end{aligned}$$

If  $\omega$  is not in any of the events in (7.6), and  $\omega \notin J^{(0)}(c_{29})$ , then for all  $x, y$  as in (1.13),

$$\begin{aligned}
 T(x, y) = T(\hat{x}, \hat{y}) & \leq \sum_{i=1}^{n+1} \hat{T}(u^{i-1}, u^i) + c_{29} n \log r \\
 & \leq \mu(y - x)_1 + c_{29} j_1 \log r + (\lambda + \lambda^{j_1}) t\sigma(r) + \sum_{j=1}^{j_1} \lambda^j t\sigma(r) \\
 & \leq h(|y - x|) + \frac{2\lambda}{1 - \lambda} t\sigma(r).
 \end{aligned}$$

Taking  $\lambda < 1/3$  it follows that

$$P\left(T(x, y) - ET(x, y) \geq t\sigma(r) \text{ for some } x, y \in G_r(K) \text{ with } |y - x| \geq \epsilon r\right) \leq c_{10} e^{-c_{11} t},$$

which completes the proof of Theorem 1.3 for downward deviations.

### 8 Proofs of Theorem 1.6 and Lemma 1.2

We finish with (5) of Remark 1.7 by proving Theorem 1.6. In the proof in section 5 for fixed  $x, y$ , transverse wandering cases (i)–(iii) were dealt with uniformly over  $x, y$ , so we need only consider the case there:

(iv)  $s \leq c_0(\log r)^{1/2}$  and there exist  $x, y, u \in \mathbb{V}$  with

$$x, y \in G_r(K), \quad u \notin G_{r,s}, \quad d(u, G_{r,s}) \leq 2, \quad |y - x| > r_0, \quad u \in \Gamma_{xy}, \quad u_1 \in [x_1, y_1]. \quad (8.1)$$

See the dashed line in Figure 10. As in section 5, this means that, for  $C_{56}$  from Lemma 3.6,

$$|u - x| + |y - u| - |y - x| \geq s^2\sigma(r), \quad \text{so} \quad h(|u - x|) + h(|y - u|) - h(|y - x|) \geq \frac{\mu s^2}{2}\sigma(r) - C_{56}$$

while

$$T(x, u) + T(u, y) - T(x, y) = 0$$

so by Theorem 1.3 and Remark 1.4,

$$\begin{aligned} P\left((iv) \text{ holds}\right) &\leq P\left(\text{there exist } v, w \in G_r(2s) \text{ with } |T(v, w) - h(|w - v|)| \geq \frac{\mu s^2}{7}\sigma(r)\right) \\ &\leq C_{36}e^{-c_1s^2}. \end{aligned} \quad (8.2)$$

Together with cases (i)–(iii) in section 5, this completes the proof of Theorem 1.6.

*Proof of Lemma 1.2.* We consider  $\rho^+$  first. Fix  $M$  (large) and  $\epsilon > 0$ , and define

$$f(r) = \log \rho(e^r) - \chi r, \quad r \geq 0,$$

$$\beta_0 = \sup\{f(r) : 0 < r \leq M\}, \quad \beta_k = \sup\{f(r) : 2^{k-1}M < r \leq 2^kM\}, \quad k \geq 1,$$

so  $f(r) = o(r)$  and hence  $\beta_k = o(2^k)$ . Letting

$$a_k = \frac{(2^{k-1} + 2^k)M}{2}, \quad k \geq 0,$$

we define the piecewise linear function  $f^+$  as follows. First let  $f^+ \equiv \beta_1$  on  $[0, a_0]$ . Then for those  $k \geq 0$  with  $\beta_k \geq \beta_{k+1}$ , define  $f^+$  on  $(a_k, a_{k+1}]$  by

$$f^+(t) = \begin{cases} \beta_k & \text{if } t \in (a_k, 2^kM] \\ \beta_{k+1} & \text{if } t = a_{k+1} \\ \text{linear on } [2^kM, a_{k+1}]. \end{cases}$$

For  $k \geq 0$  with  $\beta_k < \beta_{k+1}$ , define  $f^+$  on  $(a_k, a_{k+1}]$  by

$$f^+(t) = \begin{cases} \beta_{k+1} & \text{if } t \in [2^kM, a_{k+1}] \\ \beta_k & \text{if } t = a_k \\ \text{linear on } [a_k, 2^kM]. \end{cases}$$

Since  $\beta_k = o(2^k)$  it is easily seen that if we take  $M$  sufficiently large, then  $|(f^+)'| < \epsilon$  everywhere  $(f^+)'$  exists, and  $(f^+)'(r) \rightarrow 0$  as  $r \rightarrow \infty$ . Defining  $\rho^+$  on  $[1, \infty)$  by

$$\log \rho^+(e^r) = f^+(r) + \chi r$$

it follows that for all  $s \geq r > 1$ ,

$$\frac{\log \rho^+(s) - \log \rho^+(r)}{\log s - \log r} = \frac{f^+(\log s) - f^+(\log r)}{\log s - \log r} + \chi \in [\chi - \epsilon, \chi + \epsilon],$$

and the powerlike property follows with  $|\chi_i - \chi| \leq \epsilon, i = 1, 2$ . From the definition,  $\rho^+$  is increasing provided  $|(f^+)'| \leq \chi$  everywhere, and  $\rho^+(r)/r^{1-\delta}$  is decreasing provided  $|(f^+)'| \leq 1 - \chi - \delta$  everywhere. Both of these are valid provided  $M$  is large enough.

The proof for  $\rho^-$  is similar—we simply replace  $f$  with  $-f$  starting with the definition of the  $\beta_k$ 's. □

## References

- [1] Alexander, K. S. Approximation of subadditive functions and rates of convergence in limiting shape results. *Ann. Probab.* **24** (1997), 30–55. MR1428498
- [2] Alexander, K. S. (2020). Geodesics, bigeodesics, and coalescence in first passage percolation in general dimension. To appear in *Electron. J. Probab.* arXiv:math.PR/2001.08736. MR4673327
- [3] Alexander, K. S. (2023). Properties of first passage percolation above the (hypothetical) critical dimension. arXiv:math.PR/2304.08669
- [4] Alves, S. G., Oliveira, T. J., and Ferreira, S. C. Universality of fluctuations in the Kardar-Parisi-Zhang class in high dimensions and its upper critical dimension. *Phys. Rev. E* **90** (2018), 020103. arXiv:cond-mat.stat-mech/1405.0974. MR4536678
- [5] Baik, J., Ferrari, P. L., and Pécché, S. Convergence of the two-point function of the stationary TASEP. *Singular phenomena and scaling in mathematical models*, 91–110, (2014), Springer, Cham. arXiv:math-ph/1209.0116. MR3205038
- [6] Balázs, M., Busani, O., and Seppäläinen, T. Nonexistence of bi-infinite geodesics in the exponential corner growth model. *Forum Math. Sigma* **8** (2020), Paper No. e46, 34 pp. arXiv:math.PR/1909.06883. MR4176750
- [7] Basu, R. and Bhatia, M. (2021). Small deviation estimates and small ball probabilities for geodesics in last passage percolation. arXiv:math.PR/2101.01717
- [8] Basu, R., Ganguly, S., Hammond, A., and Hegde, M. Interlacing and scaling exponents for the geodesic watermelon in last passage percolation. *Comm. Math. Phys.* **393** (2020), no. 3, 1241–1309. arXiv:math.PR/2006.11448. MR4453234
- [9] Basu, R. and Ganguly, S. Time correlation exponents in last passage percolation. *In and out of equilibrium 3. Celebrating Vladas Sidoravicius*, 101–123, *Progr. in Probab.* **77** (2021), Birkhäuser/Springer, Cham. arXiv:math.PR/1807.09260. MR4237265
- [10] Basu, R., Ganguly, S., and Zhang, L. Temporal correlation in last passage percolation with flat initial condition via Brownian comparison. *Comm. Math. Phys.* **383** (2021), no. 3, 1805–1888. arXiv:1912.04891. MR4244262
- [11] Basu, R., Hoffman, C., and Sly, A. Nonexistence of bigeodesics in planar exponential last passage percolation. *Comm. Math. Phys.* **389** (2022), no. 1, 1–30. arXiv:math.PR/1811.04908. MR4365136
- [12] Basu, R., Sarkar, S., and Sly, A. Coalescence of geodesics in exactly solvable models of last passage percolation. *J. Math. Phys.* **60** (2019), 093301, 22 pp. arXiv:math.PR/1704.05219. MR4002528
- [13] Basu, R., Sidoravicius, V., and Sly, A. (2016). Last passage percolation with a defect line and the solution of the slow bond problem. arXiv:math.PR/1408.3464
- [14] Chatterjee, S. The universal relation between scaling exponents in first-passage percolation. *Ann. of Math. (2)* **127** (2013), no. 2, 663–697. arXiv:math.PR/1105.4566. MR3010809
- [15] Corwin, I., Liu, Z., and Wang, D. Fluctuations of TASEP and LPP with general initial data. *Ann. Appl. Probab.* **26** (2016), 2030–2082. arXiv:math.PR/1412.5087. MR3543889
- [16] Damron, M., Hanson, J., and Sosoë, P. Subdiffusive concentration in first-passage percolation. *Electron. J. Probab.* **19** (2014), Paper No. 109, 27 pp. arXiv:math.PR/1401.0917. MR3286463
- [17] Damron, M. and Wang, X. Entropy reduction in Euclidean first-passage percolation. *Electron. J. Probab.* **21** (2016), Paper No. 65, 23 pp. arXiv:math.PR/1605.06665. MR3580031
- [18] Fogedby, H. C. Kardar-Parisi-Zhang equation in the weak noise limit: Pattern formation and upper critical dimension. *Phys. Rev. E* **73** (2006), 031104. arXiv:cond-mat.stat-mech/0510268. MR2231344
- [19] Gangopadhyay, U. Fluctuations of transverse increments in two-dimensional first passage percolation. *Electron. J. Probab.* **27** (2022), 1–61. arXiv:math.PR/2011.14686. MR4417201
- [20] Ganguly, S. and Hegde, M. Optimal tail exponents in general last passage percolation via bootstrapping & geodesic geometry. *Probab. Theory Related Fields* **186** (2023), no. 1–2, 221–284. arXiv:math.PR/2007.03594. MR4586220



- [21] Howard, C. D. and Newman, C. M. Euclidean models of first-passage percolation. *Probab. Th. Rel. Fields* **108** (1997), 153–170. MR1452554
- [22] Kesten, H. (1986). H. Kesten. Aspects of first passage percolation. In *Ecole d'Eté de probabilité de Saint-Flour XIV – 1984, Lecture Notes in Math.* **1180**, Springer, Berlin, 125–264. MR0876084
- [23] Kesten, H. On the speed of convergence in first-passage percolation. *Ann. Appl. Probab.* **3** (1993), 296–338. MR1221154
- [24] Kim, S.-W. and Kim, J. M. (2014). A restricted solid-on-solid model in higher dimensions. *J. Stat. Mech.* **2014** (2014), P07005. MR3249152
- [25] Le Doussal, P. and Wiese, K. J. Two-loop functional renormalization for elastic manifolds pinned by disorder in  $N$  dimensions. *Phys. Rev. E* **72** (2005), 035101. arXiv:cond-mat.dis-nn/0501315
- [26] Ledoux, M. and Rider, B. Small deviations for beta ensembles. *Electron. J. Probab.* **15** (2010), 1319–1343. arXiv:math.PR/0912.5040. MR2678393
- [27] Loŵe, M. and Merkl, F. Moderate deviations for longest increasing subsequences: The upper tail. *Comm. Pure Appl. Math.* **54** (2001), 1488–1519. MR1852980
- [28] Loŵe, M., Merkl, F., and Rolles, S. Moderate deviations for longest increasing subsequences: The lower tail. *J. Theor. Probab.* **15** (2002), 1031–1047. MR1937784
- [29] Marinari, E., Pagnani, A., Parisi, G., Rácz, Z. Width distributions and the upper critical dimension of Kardar-Parisi-Zhang interfaces. *Phys. Rev. E* **65** (2002), 026136. arXiv:cond-mat.stat-mech/0105158
- [30] Martin, J. B., Sly, A., and Zhang, L. (2021). Convergence of the environment seen from geodesics in exponential last-passage percolation. arXiv:math.PR/2106.05242
- [31] Newman, C. M., A surface view of first passage percolation. *Proceedings of the International Congress of Mathematicians*, Vol. 1, 2 (Zürich, 1994), 1047–1023, Birkhäuser, Basel, 1995. MR1404001
- [32] Rodrigues, E. A., Oliveira, F. A., and Mello, B. A. On the existence of an upper critical dimension for systems within the KPZ universality class. *Acta. Phys. Pol. B* **46** (2015), 1231–1234. arXiv:cond-mat.stat-mech/1502.06121. MR3369006
- [33] Talagrand, M. Concentration of measure and isoperimetric inequalities in product spaces. *Publications Mathématiques de l'Institut des hautes Etudes Scientifiques* **81**(1) (1995), 73–205. MR1361756
- [34] Tessera, R. Speed of convergence in first passage percolation and geodesicity of the average distance. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** (2018), 569–586. arXiv:math.PR/1410.1701. MR3795059
- [35] Xia, G. The stretch factor of the Delaunay triangulation is less than 1.998. *SIAM J. Comput.* **42** (2013), 1620–1659. arXiv:cs.CG/1103.4361. MR3082502
- [36] Zhang, L. Optimal exponent for coalescence of finite geodesics in exponential last passage percolation. *Electron. Commun. Probab.* **25** (2020), Paper no. 74, 14 pp. arXiv:math.PR/1912.07733. MR4167385