On self-avoiding polygons and walks:
the snake method via polygon joining[*]

Alan Hammond[†]

Abstract

For $d \geq 2$ and $n \in \mathbb{N}$, let W_n denote the uniform law on self-avoiding walks beginning at the origin in the integer lattice $\mathbb{Z}^d$, and write Γ for a W_n-distributed walk. We show that the *closing probability* $\mathsf{W}_n\big(||\Gamma_n|| = 1\big)$ that Γ's endpoint neighbours the origin is at most $n^{-4/7+o(1)}$ for a positive density set of odd n in dimension $d = 2$. This result is proved using the snake method, a general technique for proving closing probability upper bounds, which originated in [3] and was made explicit in [8]. Our conclusion is reached by applying the snake method in unison with a polygon joining technique whose use was initiated by Madras in [13].

Contents

[†]University of California at Berkeley. E-mail: alanmh@math.berkeley.edu

1 Introduction

Self-avoiding walk was introduced in the 1940s by Flory and Orr [6, 17] as a model of a long polymer chain in a system of such chains at very low concentration. It is well known among the basic models of discrete statistical mechanics for posing problems that are simple to state but difficult to solve. Two recent surveys are the lecture notes [1] and [11, Section 3].

1.1 The model

We will denote by $\mathbb{N}$ the set of non-negative integers. Let $d \geq 2$. For $u \in \mathbb{R}^d$, let $||u||$ denote the Euclidean norm of u. A *walk* of length $n \in \mathbb{N}$ with $n > 0$ is a map $\gamma : \{0, \cdots, n\} \to \mathbb{Z}^d$ such that $||\gamma(i+1) - \gamma(i)|| = 1$ for each $i \in \{0, \cdots, n-1\}$. An injective walk is called *self-avoiding*. A self-avoiding walk γ of length $n \geq 2$ is said to close (and to be closing) if $||\gamma(n)|| = 1$. When the *missing edge* connecting $\gamma(n)$ and $\gamma(0)$ is added, a polygon results.

Definition 1.1. *For $n \geq 4$ an even integer, let $\gamma : \{0, \ldots, n-1\} \to \mathbb{Z}^d$ be a closing self-avoiding walk. For $1 \leq i \leq n-1$, let u_i denote the unordered nearest neighbour edge in $\mathbb{Z}^d$ with endpoints $\gamma(i-1)$ and $\gamma(i)$. Let u_n denote γ's missing edge, with endpoints $\gamma(n-1)$ and $\gamma(0)$. (Note that we have excluded the case $n = 2$ so that u_n is indeed not among the other u_i.) We call the collection of edges $\{u_i : 1 \leq i \leq n\}$ the polygon of γ. A self-avoiding polygon in $\mathbb{Z}^d$ is defined to be any polygon of a closing self-avoiding walk in $\mathbb{Z}^d$. The polygon's length is its cardinality.*

We will usually omit the adjective self-avoiding in referring to walks and polygons. Recursive and algebraic structure has been used to analyse polygons in such domains as strips, as [2] describes.

Note that the polygon of a closing walk has length that exceeds the walk's by one. Polygons have even length and closing walks, odd.

Let SAW_n denote the set of self-avoiding walks γ of length n that start at 0, i.e., with $\gamma(0) = 0$. We denote by W_n the uniform law on SAW_n. The walk under the law W_n will be denoted by Γ. The *closing probability* is $\mathsf{W}_n(\Gamma \text{ closes})$.

1.2 A combinatorial view of the closing probability

Let the walk number c_n equal the cardinality of SAW_n. By equation (1.2.10) of [14], the limit $\lim_{n \in \mathbb{N}} c_n^{1/n}$ exists and is positive and finite; it is called the connective constant and denoted by μ, and we have $c_n \geq \mu^n$.

Define the polygon number p_n to be the number of length n polygons up to translation. By (3.2.5) and (3.2.9) of [14], $\lim_{n \in 2\mathbb{N}} p_n^{1/n} \in (0, \infty)$ exists and coincides with μ, and $p_n \leq (d-1)\mu^n$.

The closing probability may be written in terms of the polygon and walk numbers. There are $2n$ closing walks whose polygon is a given polygon of length n, since there are n choices of missing edge and two of orientation. Thus,

$$\mathsf{W}_n(\Gamma \text{ closes}) = \frac{2(n+1)p_{n+1}}{c_n}, \tag{1.1}$$

for any $n \in \mathbb{N}$ (but non-trivially only for odd values of n).

1.3 Corrections to exponential growth

We define the real-valued *polygon number deficit* and *walk number excess* exponents θ_n and ξ_n according to the formulas

$$p_n = n^{-\theta_n} \cdot \mu^n \text{ for } n \in 2\mathbb{N} \tag{1.2}$$

and

$$c_n = n^{\xi_n} \cdot \mu^n \text{ for } n \in \mathbb{N}. \tag{1.3}$$

Since $c_n \geq \mu^n$ and $p_n \leq (d-1)\mu^n$, $\liminf_{n \in 2\mathbb{N}} \theta_n$ and $\inf_{n \in \mathbb{N}} \xi_n$ are non-negative in any dimension $d \geq 2$.

1.4 Closing probability upper bounds: a review

An interesting first task regarding closing probability upper bounds is to show that $W_n\big(\Gamma_n \text{ closes}\big) \to 0$ as $n \to \infty$ for any $d \geq 2$; this question was posed by Itai Benjamini to the authors of [3] who showed an upper bound of $n^{-1/4+o(1)}$ whenever $d \geq 2$. One way of approaching the question – not the method of [3] – is to improve the above stated $\liminf_{n \in 2\mathbb{N}} \big(\xi_{n-1} + \theta_n\big) \geq 0$ so that the right-hand side takes the form $1 + \chi$, for some $\chi > 0$, in which case, a closing probability upper bound of the form $n^{-\chi+o(1)}$ will result from (1.1).

Hara and Slade [10, Theorem 1.1] used the lace expansion to prove that $c_n \sim C\mu^n$ when $d \geq 5$, but the problem of showing that $c_n \mu^{-n}$ tends to infinity as $n \to \infty$ appears to be open when d equals two or three (in which cases, it is anticipated to be true). As such, this combinatorial approach to closing probability upper bounds has depended on advances on the polygon rather than the walk side. Madras and Slade [14, Theorem 6.1.3] have proved that $\theta_n \geq d/2 + 1$ when $d \geq 5$ for *spread-out* models, in which the vertices of $\mathbb{Z}^d$ are connected by edges below some bounded distance. Thus, the conclusion that the closing probability decays as fast as $n^{-d/2}$ has been reached for such models when $d \geq 5$. (In passing: this conclusion may be expected to be sharp, but the opposing lower bound is not known to the best of the author's knowledge.) When the model is nearest neighbour, the upper bound is in essence known for $d \geq 5$ in an averaged sense: [10, Theorem 1.3] states that $\sum_n n^a W_n(\Gamma \text{ closes}) < \infty$ for any $a < d/2 - 1$; in fact, the counterpart to this closing probability bound for a walk ending at any given displacement from the origin is also proved, uniformly in this displacement.

It is interesting to note that, while the problem of showing that $\limsup_{n \in 2\mathbb{N}} \theta_n$ is finite is open, it is known from classical results that, for any dimension $d \geq 2$, $\liminf_{n \in 2\mathbb{N}} \theta_n < \infty$. Indeed, a technique [14, Theorem 3.2.4] that rearranges two self-avoiding bridges into a closing walk (and, by the addition of the missing edge, into a polygon) combines with the subsequential availability of bridges of length n (meaning $b_n \mu^{-n} \geq n^{1-o(1)}$ for infinitely many n), an availability which is due to the divergence of the bridge generating function at its critical point [14, Corollary 3.1.8], to yield the inference that $\liminf_{n \in 2\mathbb{N}} \theta_n \leq d + 5$.

In dimension $d = 2$, Madras [13] showed that $\liminf \theta_n \geq 1/2$ in 1995. To show this, he introduced a polygon joining technique of which we will also make use in the present article (and which is reviewed in Section 4). His technique has recently been exploited to prove the next result, [9, Theorem 1.3].

Definition 1.2. *The limit supremum density of a set A of even, or odd, integers is*

$$\limsup_n \frac{\big|A \cap [0,n]\big|}{\big|2\mathbb{N} \cap [0,n]\big|} = \limsup_n n^{-1}\big|A \cap [0,2n]\big|.$$

When the corresponding limit infimum density equals the limit supremum density, we naturally call it the limiting density.

Theorem 1.3. *Let $d = 2$. For any $\delta > 0$, the limiting density of the set of $n \in 2\mathbb{N}$ for which $\theta_n \geq 3/2 - \delta$ is equal to one.*

As such, the combinatorial perspective offered by (1.1) alongside polygon joining via Theorem 1.3 has yielded the inference that $W_n\big(\Gamma \text{ closes}\big) \leq n^{-1/2+o(1)}$ on a full density set of odd n when $d = 2$.

This bound complements the $n^{-1/4+o(1)}$ general dimensional upper bound from [3]. In fact, the method of [3] was reworked in [8] to achieve the following $n^{-1/2+o(1)}$ bound.

Theorem 1.4. *Let $d \geq 2$. For any $\varepsilon > 0$ and $n \in 2\mathbb{N} + 1$ sufficiently high,*

$$\mathsf{W}_n\big(\Gamma \text{ closes}\big) \leq n^{-1/2+\varepsilon}.$$

The method of proof of this result was not a combinatorial analysis via (1.1) but rather a probabilistic approach called the snake method (about which more shortly).

We may summarise this review of existing closing probability upper bounds as follows. An upper bound of the form $n^{-1/2+o(1)}$ in dimension $d = 2$ has been obtained in two very different ways. Via combinatorics and polygon joining, Theorem 1.3 asserts this bound for typical odd n. Via the probabilistic snake method, Theorem 1.4 asserts the bound for each $d \geq 2$ and for all high n.

1.5 The main result

This article's principal conclusion strengthens this upper bound when $d = 2$ beyond the threshold of one-half in the two existing methods.

Theorem 1.5. *Let $d = 2$.*

1. *For any $\varepsilon > 0$, the bound*

$$\mathsf{W}_n\big(\Gamma \text{ closes}\big) \leq n^{-4/7+\varepsilon}$$

   holds on a set of $n \in 2\mathbb{N} + 1$ of limit supremum density at least $1/1250$.

2. *Suppose that the limits $\theta := \lim_{n\in 2\mathbb{N}} \theta_n$ and $\xi := \lim_{n\in\mathbb{N}} \xi_n$ exist in $[0,\infty]$. Then $\theta + \xi \geq 5/3$. Since*

$$\mathsf{W}_n\big(\Gamma \text{ closes}\big) = n^{-\theta-\xi+1+o(1)} \tag{1.4}$$

   as $n \to \infty$ through odd values of n by (1.1), the closing probability is seen to be bounded above by $n^{-2/3+o(1)}$.

(When $\theta + \xi = \infty$, (1.4) should be interpreted as asserting a superpolynomial decay in n for the left-hand side.)

However far from rigorous the hypotheses of Theorem 1.5(2) may be, their validity is uncontroversial. For example, the limiting value θ is predicted to exist and to satisfy a relation with the Flory exponent ν for mean-squared radius of gyration. The latter exponent is specified by the putative formula $\mathbb{E}_{\mathsf{W}_n} ||\Gamma(n)||^2 = n^{2\nu+o(1)}$, where $\mathbb{E}_{\mathsf{W}_n}$ denotes the expectation associated with W_n (and where note that $\Gamma(n)$ is the non-origin endpoint of Γ); in essence, $||\Gamma(n)||$ is supposed to be typically of order n^ν. The hyperscaling relation that is expected to hold between θ and ν is $\theta = 1 + d\nu$ where the dimension $d \geq 2$ is arbitrary. In $d = 2$, $\nu = 3/4$ and thus $\theta = 5/2$ is expected. That $\nu = 3/4$ was predicted by the Coulomb gas formalism [15, 16] and then by conformal field theory [4, 5]. We mention also that $\xi = 11/32$ is expected when $d = 2$; in light of the $\theta = 5/2$ prediction and (1.1), $\mathsf{W}_n\big(\Gamma \text{ closes}\big) = n^{-\psi+o(1)}$ with $\psi = 59/32$ is expected. The $11/32$ value was predicted by Nienhuis in [15] and can also be deduced from calculations concerning $\mathrm{SLE}_{8/3}$: see [12, Prediction 5]. We will refer to ψ as the *closing* exponent.

1.6 The snake method and polygon joining in unison

We have mentioned that the snake method is a probabilistic tool for proving closing probability upper bounds. It is introduced in [8]; (it originated, though was not made explicit, in [3]). The method states a condition that must be verified in order to achieve some such upper bound. In [8], this condition was verified with the use of a technique we may call Gaussian pattern fluctuation, in order to obtain Theorem 1.4.

This article is devoted to proving Theorem 1.5. We will obtain the result by a second use of the snake method. In order to verify the necessary condition, we will replace the use of pattern fluctuation with the technique of polygon joining seen in [13, 9].

That is, Theorem 1.5 pushes above the threshold of one-half for the exponent in the closing probability upper bound by combining the combinatorial and probabilistic perspectives that had been used in the two proofs that led to the lower border of this threshold. The first part of the theorem is probably the more significant conclusion. Theorem 1.5(2) is only a conditional result, but it serves a valuable expository purpose: its proof is that of the theorem's first part with certain technicalities absent.

Structure of the paper. It will by now be apparent that some groundwork is needed before the proof of Theorem 1.5 may be given. The two elements that must be reviewed (and analysed) are the polygon joining technique and the snake method. After some general notation in Section 2, the respective reviews are contained in Sections 3, 4 and 5 and in Section 6. The proof of the main result then appears in the remaining eight sections, of which the first, Section 7, offers an exposition. See the beginning of Section 3 for a summary of the content of the joining material in the ensuing three sections and the beginning of Section 8 for a summary of the structure of the sections giving the proof of Theorem 1.5.

A few words on possible reading of related works. This article is intended to be read on its own. It is nonetheless the case that, since this article makes use of a fairly subtle combination of existing techniques, the reader may benefit from reading some of the related material. For example, the application of the snake method made here is rather more technical than the one in [8], and it is imaginable that reading parts of that article might bring the reader a certain familiarity with the method that would be a useful (though not necessary) preparation for reading the present one. Our polygon joining method refines that in [9] and again some review of that work is not necessary but possibly helpful. We also mention that the online article [7] gathers together the content of the three articles [9], [8] and the present work, while also providing some further exposition and highlighting some other connections between the concepts in the articles; the reader who wishes to study all the various theorems collectively is encouraged to consult this work.

Alternatively, optional further reading may be selected by noting the outside inputs used by this paper. Theorem 1.3 and Corollary 3.3 are proved in [9]: they are Theorem 1.3 and Corollary 4.6 of that article. Lemma 6.1 and Theorem 6.3 appear in [8]: see Lemma 2.2 and Theorem 3.2 in that work. The latter two results concern the application and setup of the snake method. It is worth pointing out that their proofs are respectively almost trivial and fairly short (three to four pages in the latter case). It is thus suggested to the reader who wishes to make a moderate attempt at understanding the paper's inputs to read the proofs of these results from Sections 2.2 and 3.2 of [8] when their statements are reached.

2 Some general notation and tools

2.0.1 Cardinality of a finite set A

This is denoted by either $\#A$ or $|A|$.

2.0.2 Multi-valued maps

For a finite set B, let $\mathcal{P}(B)$ denote its power set. Let A be another finite set. A *multi-valued map* from A to B is a function $\Psi : A \to \mathcal{P}(B)$. An arrow is a pair $(a, b) \in A \times B$ for which $b \in \Psi(a)$; such an arrow is said to be outgoing from a and incoming to b. We consider multi-valued maps in order to find lower bounds on $|B|$, and for this, we need upper (and lower) bounds on the number of incoming (and outgoing) arrows. This next lemma is an example of such a lower bound.

Lemma 2.1. *Let $\Psi : A \to \mathcal{P}(B)$. Set m to be the minimum over $a \in A$ of the number of arrows outgoing from a, and M to be the maximum over $b \in B$ of the number of arrows incoming to b. Then $|B| \geq mM^{-1}|A|$.*

Proof. The quantities $M|B|$ and $m|A|$ are upper and lower bounds on the total number of arrows. $\qquad\square$

2.0.3 Denoting walk vertices and subpaths

For $i, j \in \mathbb{N}$ with $i \leq j$, we write $[i, j]$ for $\{k \in \mathbb{N} : i \leq k \leq j\}$. For a walk $\gamma : [0, n] \to \mathbb{Z}^d$ and $j \in [0, n]$, we write γ_j in place of $\gamma(j)$. For $0 \leq i \leq j \leq n$, $\gamma_{[i,j]}$ denotes the subpath $\gamma_{[i,j]} : [i, j] \to \mathbb{Z}^d$ given by restricting γ.

2.0.4 Notation for certain corners of polygons

Definition 2.2. *The Cartesian unit vectors are denoted by e_1 and e_2 and the coordinates of $u \in \mathbb{Z}^2$ by $x(u)$ and $y(u)$. For a finite set of vertices $V \subseteq \mathbb{Z}^2$, we define the northeast vertex $\mathrm{NE}(V)$ in V to be that element of V of maximal e_2-coordinate; should there be several such elements, we take $\mathrm{NE}(V)$ to be the one of maximal e_1-coordinate. That is, $\mathrm{NE}(V)$ is the uppermost element of V, and the rightmost among such uppermost elements if there are more than one. (Rightmost means of maximal e_1-coordinate.) Using the four compass directions, we may similarly define eight elements of V, including the lexicographically minimal and maximal elements of V, $\mathsf{WS}(V)$ and $\mathsf{EN}(V)$. We extend the notation to any self-avoiding walk or polygon γ, writing for example $\mathrm{NE}(\gamma)$ for $\mathrm{NE}(V)$, where V is the vertex set of γ. For a polygon or walk γ, set $y_{\max}(\gamma) = y\big(\mathrm{NE}(\gamma)\big)$, $y_{\min}(\gamma) = y\big(\mathrm{SE}(\gamma)\big)$, $x_{\max}(\gamma) = x\big(\mathrm{EN}(\gamma)\big)$ and $x_{\min}(\gamma) = x\big(\mathrm{WN}(\gamma)\big)$. The height $h(\gamma)$ of γ is $y_{\max}(\gamma) - y_{\min}(\gamma)$ and its width $w(\gamma)$ is $x_{\max}(\gamma) - x_{\min}(\gamma)$.*

2.0.5 Polygons with northeast vertex at the origin

For $n \in 2\mathbb{N}$, let SAP_n denote the set of length n polygons ϕ such that $\mathrm{NE}(\phi) = 0$. The set SAP_n is in bijection with equivalence classes of length n polygons where polygons are identified if one is a translate of the other. Thus, $p_n = |\mathrm{SAP}_n|$.

We write P_n for the uniform law on SAP_n. A polygon sampled with law P_n will be denoted by Γ, as a walk with law W_n is.

There are $2n$ ways of tracing the vertex set of a polygon ϕ of length n: n choices of starting point and two of orientation. We now select one of these ways. Abusing notation, we may write ϕ as a map from $[0, n]$ to $\mathbb{Z}^2$, setting $\phi_0 = \mathrm{NE}(\phi)$, $\phi_1 = \mathrm{NE}(\phi) - e_1$, and successively defining ϕ_j to be the previously unselected vertex for which ϕ_{j-1} and ϕ_j

form the vertices incident to an edge in ϕ, with the final choice $\phi_n = \mathrm{NE}(\phi)$ being made. Note that $\phi_{n-1} = \mathrm{NE}(\phi) - e_2$.

2.0.6 Plaquettes

The shortest non-empty polygons contain four edges. Certain such polygons play an important role in several arguments and we introduce notation for them now.

Definition 2.3. *A plaquette is a polygon with four edges. Let ϕ be a polygon. A plaquette P is called a join plaquette of ϕ if ϕ and P intersect at precisely the two horizontal edges of P. Note that when P is a join plaquette of ϕ, the operation of removing the two horizontal edges in P from ϕ and then adding in the two vertical edges in P to ϕ results in two disjoint polygons whose lengths sum to the length of ϕ. We use symmetric difference notation and denote the output of this operation by $\phi \triangle P$.*

The operation may also be applied in reverse: for two disjoint polygons ϕ^1 and ϕ^2, each of which contains one vertical edge of a plaquette P, the outcome $\left(\phi^1 \cup \phi^2\right) \triangle P$ of removing P's vertical edges and adding in its horizontal ones is a polygon whose length is the sum of the lengths of ϕ^1 and ϕ^2.

3 A very brief overview of polygon joining

The next three sections define and explain our method of polygon joining. In the present section, we set the scene with a brief review of pertinent concepts. Section 4 reviews the local details of Madras' joining procedure. Section 5 introduces and analyses the detailed mechanism for joining that we will use to prove Theorem 1.5.

Recall that the two-dimensional Theorem 1.3 is proved in [9] via polygon joining. Since it may be conceptually useful, we now recall some elements of [9] and give a very brief overview of the method of proof of its conclusion that $\theta_n \geq 3/2 - o(1)$ for typical n. Section 3 of [9] is a heuristical discussion of polygon joining which expands on the next few paragraphs. We review three joining arguments, the third one leading to the above conclusion.

Consider first a pair of polygons in $d = 2$ of say equal length n. The second may be pushed to a location where it is disjoint from, and to the right of, the first, so that there exists a plaquette whose left vertical edge lies in the first polygon and whose right vertical edge lies in the second. The reverse operation in the latter paragraph of Definition 2.3 may be applied to this plaquette and a single polygon of length $2n$ results. A moment's thought shows that, since the original pair of polygons had equal length, the location of the concerned plaquette is identifiable from the image polygon. Thus, polygon joining in its simplest guise yields the polygon superadditivity bound $p_{2n} \geq p_n^2$.

The next polygon joining argument is due to Madras [13]. The first polygon may be oriented to have height at least $n^{1/2}$. Vertically displacing the second before attempting joining should yield order $n^{1/2}$ locations for joining and thus $p_{2n} \geq n^{1/2} p_n^2$ and in essence $\theta_n \geq 1/2$.

The third joining argument is performed in [9]: working with $n/2$ in place of n, we vary the length pair for the concerned polygons from $(n/2, n/2)$ to $(n/2 - j, n/2 + j)$ where j varies over choices such that $|j| \leq n/4$, and we seem to obtain

$$p_n \geq n^{1/2 - o(1)} \sum_{j=-n/4}^{n/4} p_{n/2-j} p_{n/2+j} \tag{3.1}$$

whence $\liminf_{m \in 2\mathbb{N}} \theta_m \geq 3/2 - o(1)$. In fact, some care is needed to derive this bound. To obtain it, it must be the case that a typical length-n polygon formed by such joinings is

such that the plaquette used to join is identifiable among at most $n^{o(1)}$ possible locations (we can call this circumstance *efficient plaquette identification*). The join plaquette in question is 'global' in the rough sense that it borders two macroscopic chambers. Here is the precise definition from [9], of which we will again make use.

Definition 3.1. *For $n \in 2\mathbb{N}$, let $\phi \in \mathrm{SAP}_n$. A join plaquette P of ϕ is called* global *if the two polygons comprising $\phi \triangle P$ may be labelled ϕ^ℓ and ϕ^r in such a way that*

- *every rightmost vertex in ϕ is a vertex of ϕ^r;*
- *and $\mathrm{NE}(\phi)$ is a vertex of ϕ^ℓ.*

Write GJ_ϕ for the set of global join plaquettes of the polygon ϕ.

For example, the four edges that form the boundary of the unit square shaded red in each of the polygons ϕ in the lower part of Figure 1 are a plaquette belonging to GJ_ϕ.

In [9], efficient plaquette identification is proved to work only in the case that the image polygon length lies in a set HPN of indices of *high polygon number*. Next we state Definition 4.1 and Corollary 4.6 of [9].

Definition 3.2. *For $\zeta > 0$, the set $\mathrm{HPN}_\zeta \subseteq 2\mathbb{N}$ of ζ-high polygon number indices is given by*

$$\mathrm{HPN}_\zeta = \left\{ n \in 2\mathbb{N} : p_n \geq n^{-\zeta} \mu^n \right\}.$$

Corollary 3.3. *For all $\zeta > 0$ and $C_2 > 0$, there exist $C_3, C_4 > 0$ such that, for $n \in \mathrm{HPN}_\zeta$,*

$$\mathsf{P}_n \left(\left| \mathrm{GJ}_\Gamma \right| \geq C_3 \log n \right) \leq C_4 n^{-C_2} .$$

Thus, when $n \in \mathrm{HPN}_\zeta$, a form of (3.1) is obtained, in [9, Proposition 4.2].

Many of these ideas will also be needed in the present article, which is why we have just informally recounted them.

4 Madras' polygon joining procedure

When a pair of polygons is close, there may not be a plaquette whose vertical edges are divided between the two elements of the pair in the manner discussed above. Local details of the structure of the two polygons in the locale of near contact must be broken into cases and a suitable local modification made in each one, in order to join the pair of polygons. Madras [13] defined such a joining technique which works in a general way. For the most part, the precise details of this mechanism need not be the focus of the reader's attention; it is only in the proof of one technical result, Lemma 5.11, that they must be analysed, (and even then the result is plausible without this analysis). In this section, we explain in rough terms the principal relevant features. The reader who would like to see a detailed overview is invited to read [9, Section 4.1] where Madras' polygon joining technique is reviewed in detail and the specifics of the notation that we use in relation to it are given; a figure depicting the various cases also appears there.

In outline, then, here is Madras' technique: consider two polygons τ and σ of lengths n and m for which the intervals

$$\big[y_{\min}(\tau) - 1, y_{\max}(\tau) + 1 \big] \text{ and } \big[y_{\min}(\sigma) - 1, y_{\max}(\sigma) + 1 \big] \text{ intersect} .$$

Madras' procedure joins τ and σ to form a new polygon of length $n + m + 16$ in the following manner. The polygon σ is first pushed far to the right and then back to the left until it is a few units from touching τ. A local modification is made to both polygons around the contact point, with a net gain in length to both polygons of eight. The modification causes the polygons to reach closer to one another horizontally by four, five
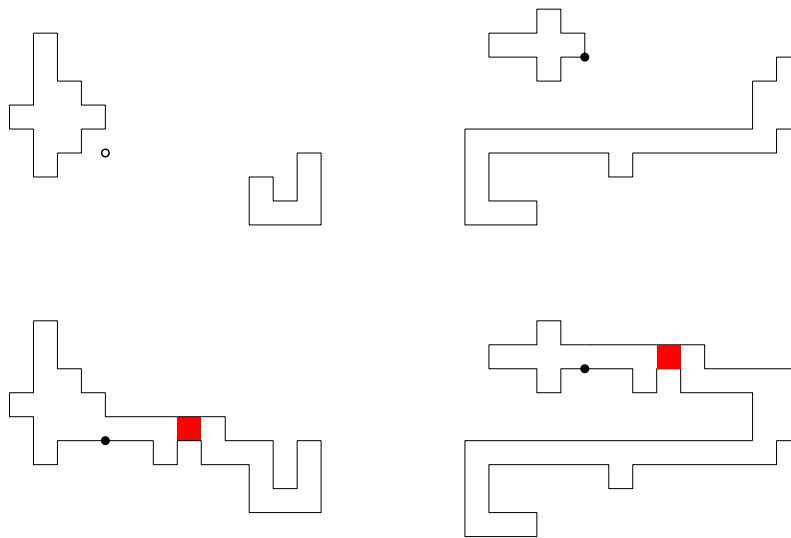
Figure 1: Two pairs of Madras joinable polygons and the Madras join polygon of each pair. The junction plaquette is shaded red (or, a touch pedantically, the interior of the union of its constituent edges is), while the circle or disk indicates an aspect of the detailed mechanism of Madras joining omitted in our overview.

or six units and a *horizontal translational adjustment* of the right polygon in one or other direction may be needed in order to obtain the desired outcome, which is a polygon pair for which there exists a plaquette in the locale of modification with left vertical edge in the left polygon and right vertical edge in the right polygon. The reverse operation in Definition 2.3 applied to this plaquette then yields a joined polygon of length $n + 16$. This polygon is defined to be the *Madras join polygon* $J(\tau, \sigma)$. The plaquette used at the final step of the construction of this polygon is called the *junction* plaquette.

Sometimes a pair of polygons are located in such a relative position (with the right polygon a few units away from the left near the point of contact) that in the above procedure, no horizontal adjustment is needed. In this case, the pair is called *Madras joinable*.

5 Polygon joining in preparation for the proof of Theorem 1.5

Our purpose in this section is to specify the polygon pairs that we will join to prove Theorem 5 and discuss some properties of the joining.

It may be apparent from the review in Section 3 that, in the proof of Theorem 1.3 that appears in [9], certain *left* polygons are Madras joined to other *right* polygons. It serves our purpose of specifying the polygon joining details needed for Theorem 1.5's proof to recall the definitions and some properties of these polygons, and we do so in the first of this section's six subsections. In the second subsection, we refine the definitions of these polygon types, and also of the joining mechanism for them. The polygons that are formed under the new, stricter, mechanism will be called *regulation global join* polygons: in the third subsection, these polygons are defined and their three key properties stated. The remaining three subsections prove these properties in turn.

5.1 Left and right polygons

Let ϕ be a polygon. Recall from Definition 2.2 the notation $y_{\max}(\phi)$ and $y_{\min}(\phi)$, as well as the height $h(\phi)$ and width $w(\phi)$.

Definition 5.1. *For $n \in 2\mathbb{N}$, let $\mathrm{SAP}_n^{\mathfrak{l}}$ denote the set of* left *polygons $\phi \in \mathrm{SAP}_n$ such that*

- *$h(\phi) \geq w(\phi)$ (and thus, by a trivial argument, $h(\phi) \geq n^{1/2}$),*
- *and $y\big(\mathrm{ES}(\phi)\big) \leq \frac{1}{2}\big(y_{\min}(\phi) + y_{\max}(\phi)\big)$.*

Let $\mathrm{SAP}_n^{\mathfrak{r}}$ denote the set of right *polygons $\phi \in \mathrm{SAP}_n$ such that*

- *$h(\phi) \geq w(\phi)$.*

Lemma 5.2. *For $n \in 2\mathbb{N}$,*

$$\big|\mathrm{SAP}_n^{\mathfrak{l}}\big| \geq \tfrac{1}{4} \cdot \big|\mathrm{SAP}_n\big| \quad and \quad \big|\mathrm{SAP}_n^{\mathfrak{r}}\big| \geq \tfrac{1}{2} \cdot \big|\mathrm{SAP}_n\big|.$$

Proof. An element $\phi \in \mathrm{SAP}_n$ not in $\mathrm{SAP}_n^{\mathfrak{r}}$ is brought into this set by right-angled rotation. If, after the possible rotation, it is not in $\mathrm{SAP}_n^{\mathfrak{l}}$, it may brought there by reflection in the x-axis. $\qquad\square$

Definition 5.3. *A Madras joinable polygon pair (ϕ^1, ϕ^2) is called globally Madras joinable if the junction plaquette of the join polygon $J(\phi^1, \phi^2)$ is a global join plaquette of $J(\phi^1, \phi^2)$.*

Both polygon pairs in the upper part of Figure 1 are globally Madras joinable.
The next result is [9, Lemma 4.11].

Lemma 5.4. *Let $n, m \in 2\mathbb{N}$ and let $\phi^1 \in \mathrm{SAP}_n^{\mathfrak{l}}$ and $\phi^2 \in \mathrm{SAP}_m^{\mathfrak{r}}$.*
Every value

$$k \in \left[y\big(\mathrm{ES}(\phi^1)\big), \, y\big(\mathrm{ES}(\phi^1)\big) + \min\left\{ n^{1/2}/2, m^{1/2} \right\} - 1 \right] \tag{5.1}$$

is such that ϕ^1 and some horizontal shift of $\phi^2 + ke_2$ is globally Madras joinable.

Write $\mathsf{GlobalMJ}_{(\phi^1, \phi^2)}$ for the set of $\vec{u} \in \mathbb{Z}^2$ such that the pair ϕ^1 and $\phi^2 + \vec{u}$ is globally Madras joinable. Then

$$\left| \mathsf{GlobalMJ}_{(\phi^1, \phi^2)} \right| \geq \min\left\{ n^{1/2}/2, m^{1/2} \right\}.$$

5.2 Left and right polygons revisited

We will refine these notions now, specifying new left and right polygon sets $\mathrm{SAP}_n^{\mathrm{left}}$ and $\mathrm{SAP}_n^{\mathrm{right}}$. Each is a subset of its earlier counterpart.

In order to specify $\mathrm{SAP}_n^{\mathrm{left}}$, we now define a *left-long* polygon. Let $\phi \in \mathrm{SAP}_n$. We employ the notationally abusive parametrization $\phi : [0, n] \to \mathbb{Z}^2$ such that $\phi_0 = \phi_n = \mathrm{NE}(\phi)$ and $\phi_1 = \mathrm{NE}(\phi) - e_1$. If $j \in [0, n]$ is such that $\phi_j = \mathrm{SE}(\phi)$, note that ϕ may be partitioned into two paths $\phi_{[0,j]}$ and $\phi_{[j,n]}$. The two paths are edge-disjoint, and the first lies to the left of the second: indeed, writing H for the horizontal strip $\big\{(x,y) \in \mathbb{R}^2 : y\big(\mathrm{SE}(\phi)\big) \leq y \leq y\big(\mathrm{NE}(\phi)\big)\big\}$, the set $H \setminus \cup_{i=0}^{j-1}[\phi_i, \phi_{i+1}]$ has two unbounded connected components; one of these, the right component, contains $H \cap \big\{(x,y) \in \mathbb{R}^2 : x \geq C\big\}$ for large enough C; and the union of the edges $[\phi_i, \phi_{i+1}]$, $j \leq i \leq n-1$, excepting the points $\phi_j = \mathrm{SE}(\phi)$ and $\phi_n = \mathrm{NE}(\phi)$, lies in the right component. It is thus natural to call $\phi_{[0,j]}$ the left path, and $\phi_{[j,n]}$, the right path. We call ϕ *left-long* if $j \geq n/2$ and *right-long* if $j \leq n/2$.

For $n \in 2\mathbb{N}$, let $\mathrm{SAP}_n^{\mathrm{left}}$ denote the set of left-long elements of $\mathrm{SAP}_n^{\mathfrak{l}}$. We simply set $\mathrm{SAP}_n^{\mathrm{right}}$ equal to $\mathrm{SAP}_n^{\mathfrak{r}}$. Note that ITA in Figure 2 is an element of $\mathrm{SAP}_n^{\mathrm{left}}$.
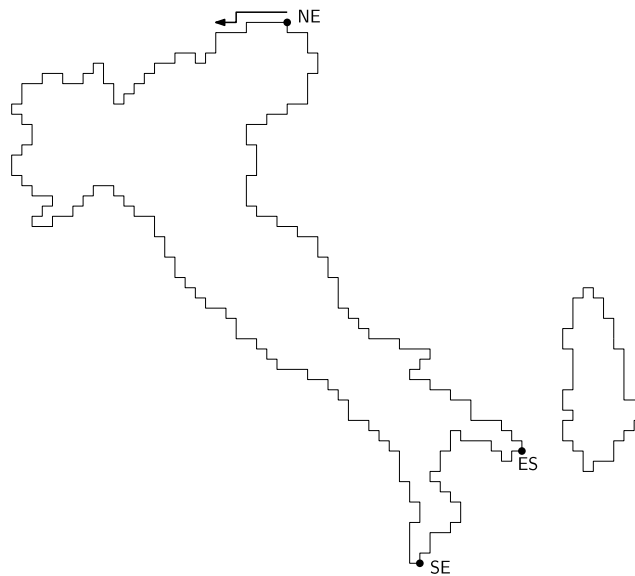
Figure 2: Illustrating Lemma 5.4 with two polygons, ITA and ALB. Denoting the polygons' lengths by n and m, note that the translates of ITA and ALB with $\mathrm{NE} = 0$ belong to $\mathrm{SAP}_n^{\mathfrak{l}}$ and $\mathrm{SAP}_m^{\mathfrak{r}}$. The start of the counterclockwise tour of ITA from $\mathrm{NE}(\mathrm{ITA})$ dictated by convention is marked by an arrow. The depicted polygons are globally Madras joinable after a horizontal shift of ALB, with the necessary surgery to ITA occurring in a vicinity of $\mathrm{ES}(\mathrm{ITA})$. Vertical shifts of ALB that leave this polygon in an easterly line of sight from $\mathrm{ES}(\mathrm{ITA})$ will maintain this state of affairs.

Lemma 5.5. *For* $n \in 2\mathbb{N}$,

$$\left|\mathrm{SAP}_n^{\mathrm{left}}\right| \geq \tfrac{1}{8} \cdot \left|\mathrm{SAP}_n\right| \ \ \text{and} \ \ \left|\mathrm{SAP}_n^{\mathrm{right}}\right| \geq \tfrac{1}{2} \cdot \left|\mathrm{SAP}_n\right|.$$

Proof. The second assertion has been proved in Lemma 5.2. In regard to the first, write the three requirements for an element $\phi \in \mathrm{SAP}_n$ to satisfy $\phi \in \mathrm{SAP}_n^{\mathrm{left}}$ in the order: $h(\phi) \geq w(\phi)$; ϕ is left-long; and $y\big(\mathrm{ES}(\phi)\big) \leq \frac{1}{2}\big(y_{\min}(\phi) + y_{\max}(\phi)\big)$. These conditions may be satisfied as follows.

- The first property may be ensured by a right-angled counterclockwise rotation if it does not already hold.
- It is easy to verify that, when a polygon is reflected in a vertical line, the right path is mapped to become a sub-path of the left path of the image polygon. Thus, a right-long polygon maps to a left-long polygon under such a reflection. (We note incidentally the asymmetry in the definition of left- and right-long: this last statement is not always true *vice versa*.) A polygon's height and width are unchanged by either horizontal or vertical reflection, so the first property is maintained if a reflection is undertaken at this second step.
- The third property may be ensured if necessary by reflection in a horizontal line. The first property's occurrence is not disrupted for the reason just noted. Could the reflection disrupt the second property? When a polygon is reflected in a horizontal line, NE and SE in the domain map to SE and NE in the image. The left path maps to the reversal of the left path, (and similarly for the right path). Thus, any left-long polygon remains left-long when it is reflected in such a way. We see that the second property is stable under the reflection in this third step.

Each of the three operations has an inverse, and thus $\#\text{SAP}_n^{\text{left}} \geq \frac{1}{8}\#\text{SAP}_n$. $\qquad\square$

5.3 Regulation global join polygons: three key properties

Definition 5.6. *Let* $k, \ell \in 2\mathbb{N}$ *satisfy* $k/2 \leq \ell \leq 35k$. *Let* $\text{RGJ}_{k,\ell}$ *denote the set of regulation global join* polygons (with length pair (k, ℓ))*, whose elements are formed by Madras joining the polygon pair* $(\phi^1, \phi^2 + \vec{u})$, *where*

1. $\phi^1 \in \text{SAP}_k^{\text{left}}$;

2. $\phi^2 \in \text{SAP}_\ell^{\text{right}}$;

3. $y_{\max}(\phi^2 + \vec{u}) \in \left[y(\text{ES}(\phi^1)), y(\text{ES}(\phi^1)) + \lfloor k^{1/2}/10 \rfloor - 1 \right]$;

4. *and* $(\phi^1, \phi^2 + \vec{u})$ *is Madras joinable.*

Let RGJ *denote the union of the sets* $\text{RGJ}_{k,\ell}$ *over all such choices of* $(k, \ell) \in 2\mathbb{N} \times 2\mathbb{N}$.

Note that Lemma 5.4 implies that whenever a polygon pair is joined to form an element of RGJ, this pair is globally Madras joinable.

The three key properties of regulation polygons are now stated as propositions.

5.3.1 An exact formula for the size of $\text{RGJ}_{k,\ell}$

Proposition 5.7. *Let* $k, \ell \in 2\mathbb{N}$ *satisfy* $k/2 \leq \ell \leq 35k$. *For any* $\phi \in \text{RGJ}_{k,\ell}$, *there is a unique choice of* $\phi^1 \in \text{SAP}_k^{\text{left}}$, $\phi^2 \in \text{SAP}_\ell^{\text{right}}$ *and* $\vec{u} \in \mathbb{Z}^2$ *for which* $\phi = J(\phi^1, \phi^2 + \vec{u})$. *We also have that*

$$\left| \text{RGJ}_{k,\ell} \right| = \lfloor k^{1/2}/10 \rfloor \cdot \left| \text{SAP}_k^{\text{left}} \right| \left| \text{SAP}_\ell^{\text{right}} \right|.$$

5.3.2 Law of the initial path of a given length in a random regulation polygon

We show that initial subpaths in regulation polygons are distributed independently of the length of the right polygon.

Definition 5.8. *Write* $\mathsf{P}_n^{\text{left}}$ *for the uniform law on* $\text{SAP}_n^{\text{left}}$.

Proposition 5.9. *Let* $k, \ell \in 2\mathbb{N}$ *satisfy* $k/2 \leq \ell \leq 35k$. *Let* $j \in \mathbb{N}$ *satisfy* $j \leq k/2 - 1$. *For any* $\phi \in \text{SAP}_k^{\text{left}}$,

$$\mathsf{P}_{k+\ell+16}\Big(\Gamma_{[0,j]} = \phi_{[0,j]} \,\Big|\, \Gamma \in \text{RGJ}_{k,\ell} \Big) = \mathsf{P}_k^{\text{left}}\Big(\Gamma_{[0,j]} = \phi_{[0,j]} \Big).$$

5.3.3 Regulation polygon joining is almost injective

We reviewed in Section 3 how a form of (3.1) is obtained in [9] by establishing efficient plaquette identification. We need something similar in our present endeavour. This next result is a counterpart to [9, Proposition 4.2].

Proposition 5.10. *For any* $\Theta > 0$, *there exist* $c > 0$ *and* $n_0 \in \mathbb{N}$ *such that the following holds. Let* $i \geq 4$ *be an integer and let* $n \in 2\mathbb{N} \cap \left[2^{i+2}, 2^{i+3} \right]$ *satisfy* $n \geq n_0$. *Suppose that* R *is a subset of* $2\mathbb{N} \cap [2^i, 2^{i+1}]$ *that contains an element* k^* *such that* $\max\left\{ \theta_{k^*}, \theta_{n-16-k^*} \right\} \leq \Theta$. *Then*

$$\left| \bigcup_{j \in \mathsf{R}} \text{RGJ}_{j,n-16-j} \right| \geq c \frac{n^{1/2}}{\log n} \sum_{j \in \mathsf{R}} p_j \, p_{n-16-j} \, .$$

Section 5 concludes with three subsections consecutively devoted to the proofs of these propositions.

5.4 Proof of Proposition 5.7

We need only prove the first assertion, the latter following directly. To do so, it is enough to determine the junction plaquette associated to the Madras join that forms ϕ. By Lemma 5.4, any such polygon pair $(\phi^1, \phi^2 + \vec{u})$ is globally Madras joinable. Recalling Definition 5.3, the associated junction plaquette is thus a global join plaquette of ϕ.

By planarity, the complement of the union of the edges that comprise ϕ is a bounded region in $\mathbb{R}^2$. Also by planarity, the removal from the closure of this region of the closed unit square associated to any of ϕ's global join plaquettes will disconnect $\mathrm{NE}(\phi)$ and $\mathrm{ES}(\phi)$. From this, we see that

- each global join plaquette of ϕ has one horizontal edge traversed in the outward journey along ϕ from $\mathrm{NE}(\phi)$ to $\mathrm{ES}(\phi)$, and one traversed on the return journey;
- and the set of ϕ's global join plaquettes is totally ordered under a relation in which the upper element is both reached earlier on the outward journey and later on the return.

We thus infer that the map that sends a global join plaquette P of ϕ to the length of the polygon in $\phi \, \Delta \, P$ that contains $\mathrm{NE}(\phi) = 0$ is injective. Since this length must equal $k + 8$ for any admissible choice of $\phi^1 \in \mathrm{SAP}_k^{\mathrm{left}}$, $\phi^2 \in \mathrm{SAP}_\ell^{\mathrm{right}}$ and $\vec{u} \in \mathbb{Z}^2$, this choice is unique, and we are done. $\qquad\square$

Note further that the set of join locations stipulated by the third condition in Definition 5.6 is restricted to an interval of length of order $k^{1/2}$. The restriction causes the formula in Proposition 5.7 to hold. It may be that many elements of $\mathrm{SAP}_k^{\mathrm{left}}$ and $\mathrm{SAP}_\ell^{\mathrm{right}}$ have heights much exceeding $k^{1/2}$, so that, for pairs of such polygons, there are many more than an order of $k^{1/2}$ choices of translate for the second element that result in a globally Madras joinable polygon pair. The term *regulation* has been attached to indicate that a specific rule has been used to produce elements of RGJ and to emphasise that such polygons do not exhaust the set of polygons that we may conceive as being globally joined.

5.5 Deriving Proposition 5.9

The next result is needed for this derivation.

Lemma 5.11. *Consider a globally Madras joinable polygon pair (ϕ^1, ϕ^2). Let $j \in \mathbb{N}$ be such that ϕ_j^1 equals $\mathrm{SE}(\phi^1)$. The join polygon $J(\phi^1, \phi^2)$ has the property that the initial subpaths $\phi_{[0,j-1]}^1$ and $J(\phi^1, \phi^2)_{[0,j-1]}$ coincide.*

Proof. The detailed structure of Madras joining from [9, Section 4.1] will be needed. Since (ϕ^1, ϕ^2) is globally Madras joinable, the northeast vertex of the join polygon coincides with $\mathrm{NE}(\phi^1)$. As such, it suffices to confirm that, absent its final edge, ϕ^1's left path $\phi_{[0,j-1]}^1$ forms part of $J(\phi^1, \phi^2)$.

Suppose that a vertex v in ϕ^1 is such that the horizontal line segment extending to the right from v intersects the vertex set of ϕ^1 only at v. Any such v must equal ϕ_i^1 for some $i \geq j$ (i.e., must lie in the right path of ϕ^1), because any vertex of the left path of ϕ^1, except its endpoint $\mathrm{SE}(\phi^1)$, is directly to the left of some other vertex of ϕ^1.

Review [9, Figure 1]. Applied to the present context, each sketch depicts a locale of ϕ^1 in which surgery (and joining) will occur. A special vertex Y is marked by a circle or a disc in each sketch. Recall from the Madras join construction that the horizontal line segment emanating rightwards from Y has no intersection with the vertex set of ϕ^1, except possibly at its leftmost point; moreover, this is also true when Y is replaced by either $Y - e_2$ or $Y + e_2$.

Thus, we see that, whichever case for [9, Figure 1] obtains, any vertex of ϕ^1 lying in $\{Y - e_2, Y, Y + e_2\}$ has the form ϕ_i^1 for $i \geq j$. Inspection of the figure shows that the
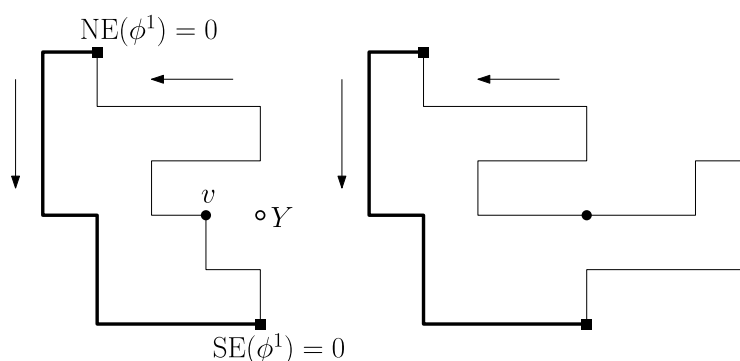
Figure 3: A case in the proof of Lemma 5.11. The polygon ϕ^1 depicted on the left lies in SAP_n for some $n \in 2\mathbb{N}$. The polygon is indexed counterclockwise from $\mathrm{NE}(\phi^1)$ as the arrows indicate. Recalling [9, Figure 1], the circle in the left sketch, and the disk in the right, marks the point Y used in the formation of the right sketch τ_{mod} with $\tau = \phi^1$, (during the process of Madras joining with an undepicted ϕ^2). Case IIcii obtains for ϕ^1. The left sketch disk marks a vertex v as discussed in the proof. Emboldened is the left path of ϕ^1, which passes undisturbed by the surgery from the left sketch to the right to form an initial subpath of the postsurgical polygon.

surgery conducted to form the Madras join polygon leaves $\phi^1_{[0,j-1]}$ unaltered: in case IIb, this inference depends on further noting that both $Y - e_1$ and $Y - e_1 + e_2$ take the form ϕ^1_i for $i \geq j$; in case IIci, it depends on noting that $Y - e_2 - e_1$ and $Y - e_1$ take such a form. (Case IIcii is illustrated in Figure 3.) Moreover, this assertion of constancy under surgery includes cases IIIa, IIIb, IIIci and IIIcii which are not depicted but are discussed in the caption of [9, Figure 1].

In fact, in all cases other than case IIa, even $\phi^1_{[0,j]}$ suffers no alteration in surgery. It is case IIa and the possibility that $Y + e_2 = \mathrm{SE}(\phi_1)$ that leads us to claim merely that $\phi^1_{[0,j-1]} = J(\phi^1, \phi^2)_{[0,j-1]}$ in Lemma 5.11. $\qquad\square$

Proof of Proposition 5.9. By the first assertion in Proposition 5.7, we find that the conditional distribution of $\mathsf{P}_{k+\ell+16}$ given that $\Gamma \in \mathrm{RGJ}_{k,\ell}$ has the law of the output in this procedure:

- select a polygon according to the law $\mathsf{P}^{\mathrm{left}}_k$;
- independently select another uniformly among elements of $\mathrm{SAP}^{\mathrm{right}}_\ell$;
- choose one of the $\lfloor \frac{1}{10} k^{1/2} \rfloor$ vertical displacements specified in Definition 5.6 uniformly at random, and shift the latter polygon by this displacement;
- translating the displaced polygon by a suitable horizontal shift, Madras join the first and the translated second polygon to obtain the output.

Denote by ϕ^1 the element of $\mathrm{SAP}^{\mathrm{left}}_k$ selected in the first step. This polygon is left-long; let $m \geq k/2$ satisfy $\phi^1_m = \mathrm{SE}(\phi^1)$. By Lemma 5.11, the output polygon has an initial subpath that coincides with the left path $\phi^1_{[0,m-1]}$ of ϕ^1 (absent the final edge). Since $j \leq k/2 - 1 \leq m - 1$, we see that the initial length j subpath of the output polygon coincides with $\phi^1_{[0,j]}$. Thus, its law is that of $\Gamma_{[0,j]}$ where the polygon Γ is distributed according to $\mathsf{P}^{\mathrm{left}}_k$. $\qquad\square$

5.6 Proof of Proposition 5.10

This argument is based on the proof of [9, Lemma 4.12], (which is the principal component of Proposition 4.2 in that paper). We mentioned in Section 3 that efficient

plaquette identification is proved to work in [9] only when the length of the joined polygon lies in the index set HPN_ζ (for a suitable choice of $\zeta > 0$). It is Corollary 3.3 that demonstrates that identification is efficient. We use this approach again now.

Let $\Theta > 0$, $i \geq 4$ and $n \in 2\mathbb{N} \cap [2^{i+2}, 2^{i+3}]$. Suppose that a set R contains an element k^* in the way that Proposition 5.10 describes. We begin by pointing out that for any choice of ζ exceeding 2Θ, the assertion that $n \in \text{HPN}_\zeta$ holds provided that n is at least the value $n_0(\zeta) = \mu^{16/(\zeta-2\Theta)}$. That this is true follows from the straightforward $p_n \geq p_{n-16}$, polygon superadditivity $p_{n-16} \geq p_{k^*} p_{n-16-k^*}$ (see [14, (3.2.3)]) and the hypothesised $\max\{\theta_{k^*}, \theta_{n-16-k^*}\} \leq \Theta$.

We now consider a multi-valued map $\Psi : A \to \mathcal{P}(B)$. We take $A = \bigcup_{j \in \text{R}} \text{SAP}_j^{\text{left}} \times \text{SAP}_{n-16-j}^{\text{right}}$ and $B = \cup_{j \in \text{R}} \text{RGJ}_{j,n-16-j}$ (which is a subset of $\text{SAP}_n \cap \text{RGJ}$). Note the indices concerned in specifying A are non-negative because $\max \text{R} \leq 2^{i+1}$, $n \geq 2^{i+2}$ and $i \geq 3$ ensure that $n-16-j \geq 0$ whenever $j \in \text{R}$. (Moreover, the choice $(k, \ell) = (j, n-16-j)$ in Definition 5.6 does satisfy the bounds stated there. We use $i \geq 4$ to verify that $k/2 \leq \ell$.) We specify Ψ so that the generic domain point $(\phi^1, \phi^2) \in \text{SAP}_j^{\text{left}} \times \text{SAP}_{n-16-j}^{\text{right}}$, $j \in \text{R}$, is mapped to the collection of length-n polygons formed by Madras joining ϕ^1 and $\phi^2 + \vec{u}$ as $\vec{u}$ ranges over the subset of $\text{GlobalMJ}_{(\phi^1,\phi^2)}$ specified for the given pair (ϕ^1, ϕ^2) by conditions (3) and (4) in Definition 5.6. (That this set is indeed a subset of $\text{GlobalMJ}_{(\phi^1,\phi^2)}$ is noted after Definition 5.6.)

We have that $\left|\Psi\big((\phi^1, \phi^2)\big)\right| = \lfloor j^{1/2}/10 \rfloor$. Since $j \geq 2^i$ and $n \leq 2^{i+3}$, we have that $j \geq 2^{-3}n$, so that $\left|\Psi\big((\phi^1, \phi^2)\big)\right| \geq \lfloor \frac{1}{10} 2^{-3/2} n^{1/2} \rfloor$.

Applying Lemma 5.5, we learn that the number of arrows in Ψ is at least

$$2^{-4} \cdot \lfloor \tfrac{1}{10} 2^{-3/2} n^{1/2} \rfloor \sum_{j \in \text{R}} p_j p_{n-16-j} \, . \tag{5.2}$$

For a constant $C_6 > 0$, denote by

$$H_n = H_n(C_6) = \left\{ \phi \in B : \left|\text{GJ}_\phi\right| \geq C_6 \log n \right\}$$

the set of length-n polygons with a *high* number of global join plaquettes.

We now set the value of $C_6 > 0$. Specify a parameter $\zeta = 2\Theta + 1$, and recall that $n \in \text{HPN}_\zeta$ is known provided that $n \geq n_0(\zeta)$. We set the parameter C_6 equal to the value C_3 furnished by Corollary 3.3 applied with C_2 taking the value $2\Theta + 1$ and for the given value of ζ (which happens to be the same value). This use of the corollary also provides a value of C_4, and, since $\text{P}_n\big(|\text{GJ}_\Gamma| \geq C_6 \log n\big) = p_n^{-1} |H_n|$, we learn that $|H_n| \leq C_4 n^{-(2\Theta+1)} p_n$.

There are two cases to analyse, the first of which is specified by at least one-half of the arrows in Ψ pointing to H_n. In the first case, we have

$$n \times C_4 n^{-(2\Theta+1)} p_n \geq \max\left\{ |\Psi^{-1}(\phi)| : \phi \in H_n \right\} \times \left|H_n\right| \geq \tfrac{1}{2} \#\left\{ \text{ arrows in } \Psi \right\}$$
$$\geq \quad 2^{-5} \lfloor \tfrac{1}{10} 2^{-3/2} n^{1/2} \rfloor \sum_{j \in \text{R}} p_j p_{n-16-j} \geq \alpha\, n^{1/2} p_{k^*} p_{n-16-k^*} \geq \alpha\, n^{1/2} n^{-2\Theta} \mu^{n-16} \, .$$

where the value of α may be chosen to equal $\frac{1}{10} 2^{-7}$ since n is supposed to be least n_0, and the latter quantity may be increased if need be. The consecutive inequalities are due to: $\left|\Psi^{-1}(\phi)\right| \leq n$ for $\phi \in \text{SAP}_n$; the first case obtaining; the lower bound (5.2); the membership of R by the hypothesised index k^*; and the hypothesised properties of k^*. We have found that $p_n \geq c\, n^{1/2} \mu^n$ with $c = \mu^{-16} \alpha C_4^{-1}$. Since $p_n \leq \mu^n$ (as noted in the introduction), the first case ends in contradiction if we stipulate that the hypothesised lower bound n_0 on n is at least $c^{-2} = 100 \cdot 2^{14} \mu^{32} C_4$ (where this bound has a Θ-dependence that is transmitted via C_4).

The set of preimages under Ψ of a given $\phi \in B \subset \mathrm{SAP}_n$ may be indexed by the junction plaquette associated to the Madras join polygon $J(\phi^1, \phi^2)$ that equals ϕ. Recalling Definition 5.3, this plaquette is a global join plaquette of ϕ. Thus,

$$|\Psi^{-1}(\phi)| \leq |\mathsf{GJ}_\phi| \text{ for any } \phi \in B. \tag{5.3}$$

We find then that, when the second case obtains,

$$|B| \geq |B \setminus H_n| \geq C_6^{-1} (\log n)^{-1} \times 2^{-5} \lfloor \tfrac{1}{10} 2^{-3/2} n^{1/2} \rfloor \sum_{j \in \mathsf{R}} p_j p_{n-16-j}.$$

Recalling the definition of the set B completes the proof of Proposition 5.10. $\qquad\square$

6 The snake method: general elements

In this section, we recall from [8] the snake method. The method is used to prove upper bounds on the closing probability, and assumes to the contrary that to some degree this probability has slow decay. For the technique to be used, two ingredients are needed.

1. A charming snake is a walk or polygon γ many of whose subpaths beginning at $\mathrm{NE}(\gamma)$ have high conditional closing probability, when extended by some common length. It must be shown that charming snakes are not too atypical.

2. A general procedure is then specified in which a charming snake is used to manufacture huge numbers of alternative self-avoiding walks. These alternatives overwhelm the polygons in number and show that the closing probability is very small, contradicting the assumption.

The first step certainly makes use of the assumption of slow decay on the closing probability. It is not however a simple consequence of this assumption. After this section, we will carry out the first step via the polygon joining apparatus we have constructed to prove Theorem 1.5; in [8], the first step was implemented via a technique of Gaussian pattern fluctuation to prove the general dimensional Theorem 1.4.

This section is devoted to explaining the second step, which is formulated as a general tool, valid in any dimension $d \geq 2$.

6.1 Some necessary notation and tools

To set up the apparatus of the method, we present some preliminaries.

6.1.1 The two-part decomposition

In the snake method, we represent any given walk γ in a *two-part* decomposition. This consists of an ordered pair of walks (γ^1, γ^2) that emanate from a certain common vertex and that are disjoint except at that vertex. The two walks are called the *first part* and the *second part*. To define the decomposition, consider any walk γ of length n. We first mention that the common vertex is chosen to be the northeast vertex $\mathrm{NE}(\gamma)$. Choosing $j \in [0, n]$ so that $\gamma_j = \mathrm{NE}(\gamma)$, the walk γ begins at γ_0 and approaches $\mathrm{NE}(\gamma)$ along the subwalk $\gamma_{[0,j]}$, and then continues to its endpoint γ_n along the subwalk $\gamma_{[j,n]}$. The reversal $\overleftarrow{\gamma}_{[0,j]}$ of the first walk, and the second walk $\gamma_{[j,n]}$, form a pair of walks that emanate from $\mathrm{NE}(\gamma)$. (When j equals zero or n, one of the walks is the length zero walk that only visits $\mathrm{NE}(\gamma)$; in the other cases, each walk has positive length.) The two walks will be the two elements in the two-part decomposition; all that remains is to order them, to decide which is the first part. If one of the walks visits $\mathrm{NE}(\gamma) - e_1$ at its first step, then it
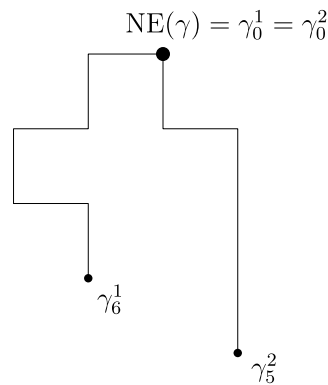
Figure 4: The two-part decomposition of a walk of length eleven.

is set equal to γ^1; of course, γ^2 is then the other walk. If neither visits $\mathrm{NE}(\gamma) - e_1$, then one of the walks has length zero. This walk is γ^1 in this case.

We use square brackets to indicate the two-part decomposition, writing $\gamma = [\gamma^1, \gamma^2]$.

As a small aid to visualization, it is useful to note that if the first γ^1 part of a two-dimensional walk γ for which $\mathrm{NE}(\gamma) = 0$ has length j, then $\gamma^1 : [0, j] \to \mathbb{Z}^2$ satisfies

- $\gamma_0^1 = 0$ and $\gamma_1^1 = -e_1$;
- $y(\gamma_i^1) \leq 0$ for all $i \in [0, j]$;
- $\gamma_i^1 \notin \mathbb{N} \times \{0\}$ for any $i \in [1, j]$.

6.1.2 Notation for walks not beginning at the origin

Let $n \in \mathbb{N}$. We write SAW_n^* for the set of self-avoiding walks γ of length n (without stipulating the location γ_0). We further write SAW_n^0 for the subset of SAW_n^* whose elements γ verify $\mathrm{NE}(\gamma) = 0$. Naturally, an element $\gamma \in \mathrm{SAW}_n^0$ is said to close (and be closing) if $||\gamma_n - \gamma_0|| = 1$. The uniform law on SAW_n^0 will be denoted by W_n^0. The sets SAW_n^0 and SAW_n are in bijection via a clear translation; we will use this bijection implicitly.

6.1.3 First parts and closing probabilities

First part lengths with low closing probability are rare, as the following result, [8, Lemma 2.2], attests.

Lemma 6.1. *Let $n \in 2\mathbb{N} + 1$ be such that, for some $\alpha' > 0$, $\mathrm{W}_n\big(\Gamma \text{ closes}\big) \geq n^{-\alpha'}$. For any $\delta' > 0$, the set of $i \in [0, n]$ for which*

$$\#\Big\{\gamma \in \mathrm{SAW}_n^0 : |\gamma^1| = i\Big\} \geq n^{\alpha' + \delta'} \cdot \#\Big\{\gamma \in \mathrm{SAW}_n^0 : |\gamma^1| = i, \gamma \text{ closes}\Big\}$$

has cardinality at most $2n^{1-\delta'}$.

6.1.4 Possible first parts and their conditional closing probabilities

For $n \in \mathbb{N}$, let $\mathrm{First}_n \subseteq \mathrm{SAW}_n$ denote the set of walks $\gamma : [0, n] \to \mathbb{Z}^2$ whose northeast vertex is $\gamma_0 = 0$. We wish to view First_n as the set of possible first parts of walks $\phi \in \mathrm{SAW}_m^0$ of some length m that is at least n. (We could in fact be more restrictive in specifying First_n, stipulating if we wish that any element γ satisfies $\gamma_1 = -e_1$. As it
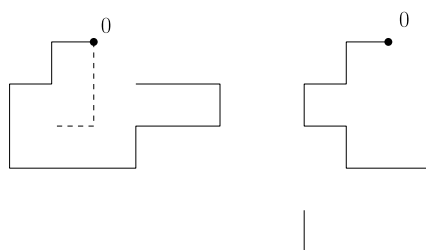
Figure 5: *Left*: the bold $\phi \in \mathrm{SAW}_{14}$ and dashed $\gamma \in \mathrm{SAW}_3$ are such that $[\phi, \gamma]$ is a two-part decomposition. Note that $\phi \in \mathrm{First}_{14,14+3} \cap \mathrm{First}^c_{14,14+4}$. *Right*: An element of $\cap^\infty_{m=1} \mathrm{First}_{14,14+m}$.

stands, First_n actually contains all possible second parts. What matters, however, is only that First_n contains all possible first parts.)

Note that, as Figure 5 illustrates, for given $m > n$, only some elements of First_n appear as such first parts, and we now record notation for the set of such elements. Write $\mathrm{First}_{n,m} \subseteq \mathrm{First}_n$ for the set of $\gamma \in \mathrm{First}_n$ for which there exists an element $\phi \in \mathrm{SAW}_{m-n}$ (necessarily with $\mathrm{NE}(\phi) = 0$) such that $[\gamma, \phi]$ is the two-part decomposition of some element $\chi \in \mathrm{SAW}_m$ with $\mathrm{NE}(\chi) = 0$.

In this light, we now define the conditional closing probability

$$q_{n,m} : \mathrm{First}_{n,m} \to [0,1], \quad q_{n,m}(\gamma) = \mathsf{W}^0_m \Big(\Gamma \text{ closes} \,\Big|\, \Gamma^1 = \gamma \Big),$$

where here $m, n \in \mathbb{N}$ satisfy $m > n$; note that since $\gamma \in \mathrm{First}_{n,m}$, the event in the conditioning on the right-hand side occurs for some elements of SAW_m, so that the right-hand side is well-defined.

We also identity a set of first parts with *high* conditional closing probability: for $\alpha > 0$, we write

$$\mathrm{HighFirst}^\alpha_{n,m} = \Big\{ \gamma \in \mathrm{First}_{n,m} : q_{n,m}(\gamma) > m^{-\alpha} \Big\}.$$

6.2 The snake method: recall of its general apparatus

6.2.1 Parameters

The snake method has three exponent parameters:

- the inverse charm $\alpha > 0$;
- the snake length $\beta \in (0,1]$;
- and the charm deficit $\eta \in (0, \beta)$.

It has two index parameters:

- $n \in 2\mathbb{N} + 1$ and $\ell \in \mathbb{N}$, with $\ell \leq n$.

6.2.2 Charming snakes

Here we define these creatures.

Definition 6.2. *Let $\alpha > 0$, $n \in 2\mathbb{N}+1$, $\ell \in [0, n]$, $\gamma \in \mathrm{First}_{\ell,n}$, and $k \in [0, \ell]$ with $\ell - k \in 2\mathbb{N}$. We say that γ is (α, n, ℓ)-charming at (or for) index k if*

$$\mathsf{W}^0_{k+n-\ell} \Big(\Gamma \text{ closes} \,\Big|\, |\Gamma^1| = k,\, \Gamma^1 = \gamma_{[0,k]} \Big) > n^{-\alpha}. \tag{6.1}$$

(The event that $|\Gamma^1| = k$ in the conditioning is redundant and is recorded for emphasis.) Note that an element $\gamma \in \mathrm{First}_{\ell,n}$ is (α, n, ℓ)-charming at index k if in selecting a length $n - \ell$ walk beginning at 0 uniformly among those choices that form the second part of a walk whose first part is $\gamma_{[0,k]}$, the resulting $(k + n - \ell)$-length walk closes with probability exceeding $n^{-\alpha}$. (Since we insist that n is odd and that ℓ and k share their parity, the length $k + n - \ell$ is odd; the condition displayed above could not possibly be satisfied if this were not the case.) Note that, for any $\ell \in [0, n]$, $\gamma \in \mathrm{First}_{\ell,n}$ is (α, n, ℓ)-charming at the special choice of index $k = \ell$ precisely when $\gamma \in \mathrm{HighFirst}_{\ell,n}^{\alpha}$. When $k < \ell$ with $k + n - \ell$ of order n, the condition that $\gamma \in \mathrm{First}_{\ell,n}$ is (α, n, ℓ)-charming at index k is *almost* the same as $\gamma_{[0,k]} \in \mathrm{HighFirst}_{k,k+n-\ell}^{\alpha}$; (the latter condition would involve replacing n by $k + n - \ell$ in the right-hand side of (6.1)).

For $n \in 2\mathbb{N} + 1$, $\ell \in [0, n]$, $\alpha, \beta > 0$ and $\eta \in (0, \beta)$, define the charming snake set

$$\mathrm{CS}_{\beta,\eta}^{\alpha,\ell,n} \;=\; \Big\{ \gamma \in \mathrm{First}_{\ell,n} : \gamma \text{ is } (\alpha, n, \ell)\text{-charming}$$

$$\text{for at least } n^{\beta - \eta}/4 \text{ indices belonging to the interval } \big[\ell - n^{\beta}, \ell\big] \Big\} .$$

A charming snake is an element of $\mathrm{CS}_{\beta,\eta}^{\alpha,\ell,n}$ and, as such, it is a walk. To a charming snake γ, however, is associated the sequence $\big(\gamma_{[0,\ell - n^{\beta}]}, \gamma_{[0,\ell - n^{\beta}+1]}, \cdots, \gamma_{[0,\ell]}\big)$ of $n^{\beta} + 1$ terms. This sequence may be depicted as evolving as an extending snake does; since $\gamma \in \mathrm{CS}_{\beta,\eta}^{\alpha,\ell,n}$, the sequence has many (α, n, ℓ)-charming elements, for each of which there is a high conditional closing probability for extensions of a *shared* continuation length $n - \ell$.

6.2.3   A general tool for the method's second step

For the snake method to produce results, we must work with a choice of parameters for which $\beta - \eta - \alpha > 0$. (The method could be reworked to handle equality in some cases.) Here we present the general tool for carrying out the method's second step. The tool asserts that, if $\beta - \eta - \alpha > 0$ and even a very modest proportion of snakes are charming, then the closing probability drops precipitously.

Theorem 6.3. *Let $d \geq 2$. Set $c = 2^{\frac{1}{5(4d+1)}} > 1$ and set $K = 20(4d + 1)\frac{\log(4d)}{\log 2}$. Suppose that the exponent parameters satisfy $\delta = \beta - \eta - \alpha > 0$. If the index parameter pair (n, ℓ) satisfies $n \geq K^{1/\delta}$ and*

$$\mathsf{P}_{n+1}\Big(\Gamma_{[0,\ell]} \in \mathrm{CS}_{\beta,\eta}^{\alpha,\ell,n}\Big) \geq c^{-n^{\delta}/2}, \tag{6.2}$$

then

$$\mathsf{W}_n\Big(\Gamma \text{ closes}\Big) \leq 2(n + 1)\, c^{-n^{\delta}/2}.$$

The proof appears in [8, Section 3.2].

Note that since the closing probability is predicted to have polynomial decay, the hypothesis (6.2) is never satisfied in practice. For this reason, the snake method will always involve argument by contradiction, with (6.2) being established under a hypothesis that the closing probability is, to some degree, decaying slowly.

7   An impression of the ideas for the main proof

We have assembled the principal inputs needed to prove Theorem 1.5 and will begin proving this result in the next section. The proof is rather technical and delicate, and, before we begin it, we attempt to set the reader's bearings by giving a necessarily impressionistic account of the ideas that drive the derivation.

Remember that our basic aim in this paper to prove that, when $d = 2$, the closing probability has a rate of decay more rapid than $n^{-1/2-\varepsilon}$, for some $\varepsilon > 0$. Since we

will proceed by proof by contradiction, it is helpful to reflect on the circumstance that the opposite case holds, which we may summarise by the informal assertion that $\mathsf{W}_n(\Gamma \text{ closes}) \geq n^{-1/2-o(1)}$. In this section, we will refer to this bound, assumed for the present purpose of exposition to hold for all high odd n, as the *standing assumption*. Recalling (1.1) and $c_n \geq \mu^n$, our assumption forces $p_n \geq n^{-3/2-o(1)}\mu^n$ for even n. (The opposing bound $p_n \leq n^{-3/2+o(1)}\mu^n$ holds for typical such n by Theorem 1.3. The standing assumption thus yields the perhaps provocatively implausible bound $c_n \leq n^{o(1)}\mu^n$ for typical n. This remark is however an aside, because we will make no use of this bound.)

The lower bound on polygon number forces all sufficiently high even numbers to be high polygon number indices in the sense of Definition 3.2, so that the hypothesis $n \in \mathsf{HPN}_\zeta$ of Corollary 3.3 (for any given $\zeta > 3/2$) is verified for high even n. This means that the circumstance of efficient plaquette identification obtains, so that a form of (3.1) holds for large even n.

Now, the right-hand side of (3.1) is an expression for the number of ways of joining together pairs of polygons each of size of order n, provided that a set of order $n^{1/2}$ possibilities is used for each pair of polygons. More joinings may be possible, but we do not know this for sure: we will call the Madras join polygons arising from a specific set of order $n^{1/2}$ join locations 'regulation global join polygons' in the upcoming proof, though here we prefer to call them simply 'global join polygons', to emphasise their key geometric feature of being comprised of two chambers of the same order of length.

Now a very distinctive geometric feature obtains under the standing assumption: a non-negligible fraction (of size at least $n^{-o(1)}$) of polygons of length a given high even value n must be global join polygons. We may call this feature 'ample supply of global join polygons'. The reason that such polygons are in ample supply is that, if these special polygons represented only a negligible fraction of polygons of a given high length, then (3.1) could be improved by the presence of an addition factor of n^ε on its right-hand side, and the inference that $p_n\mu^{-n} \leq n^{-3/2-\varepsilon}$ would result, contrary to the bound $p_n \geq n^{-3/2-o(1)}\mu^n$ that has been seen to follow from the standing assumption.

The ample supply of global join polygons will play a critical role in validating the snake method hypothesis (6.2). Under the standing assumption, we will endeavour to verify this hypothesis with the parameter choice $\beta = 1$; this choice reflects the ambition to implement the method in its strongest form, in which a charming snake whose length is macroscopic (of order n) must be shown to be not atypical. We further set, for reasons that will shortly be discussed, $\alpha = 1/2 + o(1)$ and $\eta = o(1)$. (The bound that $\delta = \beta - \eta - \alpha$ be positive is satisfied; when we turn later to rigorous argument, the consideration that $\delta > 0$ dictates the degree to which we succeed in showing that the closing probability decays more quickly than $n^{-1/2}$.)

To explain the basic meaning of the method hypothesis (6.2), first recall that we have been treating n in the preceding paragraphs as a general parameter and thus permitting it to be even. In the snake method, however, it is fixed at given high odd value, (reflecting its role as the length of a closing walk). So fixing n, we also make a choice of the parameter $\ell \leq n$ of the same order as n. Think of a journey travelled counterclockwise along a typical polygon under P_{n+1}. To each moment of the journey – to each vertex of the polygon – corresponds the first part of the walk formed by the removal from the polygon of the edge over which the traveller will next pass. Since we take $\beta = 1$, the journey includes order n steps, so that the traveller passes over a macroscopic part of the polygon. To each first part, we consider a second part of fixed length $n - \ell$, chosen uniformly among the choices of second part that are admissible given the first part. The hypothesis (6.2) is satisfied when it is with at least modest probability (for simplicity, say at least $n^{-o(1)}$), that a non-negligible fraction of these first parts (specifically, at least $n^{1-o(1)}$ of them) are charming in the sense that this random

selection of second part has a probability at least $n^{-1/2-o(1)}$ of closing the first part, i.e., of ending next to the non-origin endpoint of the given first part.

Roughly speaking, this choice of cutoff $n^{-1/2-o(1)}$, dictated by setting $\alpha = 1/2 + o(1)$, is compatible with the standing assumption that the closing probability decays at rate very close to $n^{-1/2}$. Indeed, if we were to vary the form of the hypothesis (6.2) a little (calling the new version the 'simple' form of the hypothesis), we could readily deduce that many first parts have the requisite property. In order to specify the simple form of hypothesis (6.2), we adopt the definition that, for a polygon of length $n + 1$, a first part, of a length k of order $n/2$, is 'simply' charming if, when a second part extension of length $n - k$ is sampled, it closes the first part with probability at least $n^{-1/2-o(1)}$.

Because the closing walks so formed correspond to polygons of length n, it is now a direct consequence of our standing assumption that, for a typical polygon sampled from P_{n+1}, most first parts are simply charming. Indeed, the number of first parts that fail to be simply charming is bounded above by Lemma 6.1. (The proof of this result appears in [8], but it is very straightforward.) That is, simply charming first parts are typical, with $n^{1-o(1)}$ exceptions permitted among order n lengths of first part.

Our discussion has led us to a clear expression of the challenge we face in implementing the snake method. The hypothesis (6.2) is validated in its 'simple' form, that is, when we are prepared to decrease the second part length as the first part length rises (so that the sum of lengths equals n). In its actual form, of course, we must accept that the second part length remains fixed (at the value $n - \ell$) even as the first part length k increases over a range of order n values. In fact, the simple version of the hypothesis has been comfortably validated. From the form of (6.2), we see that actually charming first parts must be shown to represent a non-negligible, $n^{-o(1)}$, fraction of all first parts for a non-negligible fraction of all length $n + 1$ polygons; but first parts have been seen to be simply charming apart from a small, $n^{1-o(1)}$-sized, set of violations, with probability at least $1 - n^{-o(1)}$.

It is the ample supply of global join polygons that allows us to bridge the gap between the simple and actual forms of the snake method hypothesis. Building this bridge is a matter of understanding a certain similarity of measure between the polygon laws P_{n+1} and P_{m+1}, where m is odd and of order n. We are trying to gain an understanding that there is at least probability $n^{-o(1)}$ that a first part of a given length k arising from the law P_{n+1} will have the property that a length $n - \ell$ second part will close with probability at least $n^{-1/2-o(1)}$. By considering the choice $m = n - \ell + k$, we have this information for typical such k in the case that the polygon law is P_{m+1} in place of P_{n+1}: this is the deduction made from knowing that the snake method hypothesis is valid in its simple form.

To build the bridge, recall that a non-negligible, $m^{-o(1)}$, fraction of length-$(m + 1)$ polygons are global join polygons. There is an attractive random means of mapping such polygons into their length-$(n + 1)$ counterparts. Any such polygon is comprised of a left polygon and a right polygon. Pull out the right polygon, replace it by a polygon, uniformly chosen among the space of polygons whose length exceeds the removed polygon's by $n - m$ (whatever the sign of this difference), and rejoin the new polygon in one of the order $n^{1/2}$ spots that we regard as admissible for the formation of globally joined polygons.

The resulting polygon has length $n + 1$; (clearly, we are neglecting the details of the Madras join construction here). Moreover, this random function, mapping the space of length-$(m+1)$ globally joined polygons to the space of their length-$(n+1)$ counterparts, is roughly measure-preserving when projected to first parts that do not reach the location where surgery occurs (and thus are undisturbed by surgery).

The probability that a first part of given length k fails the simple charm test, when
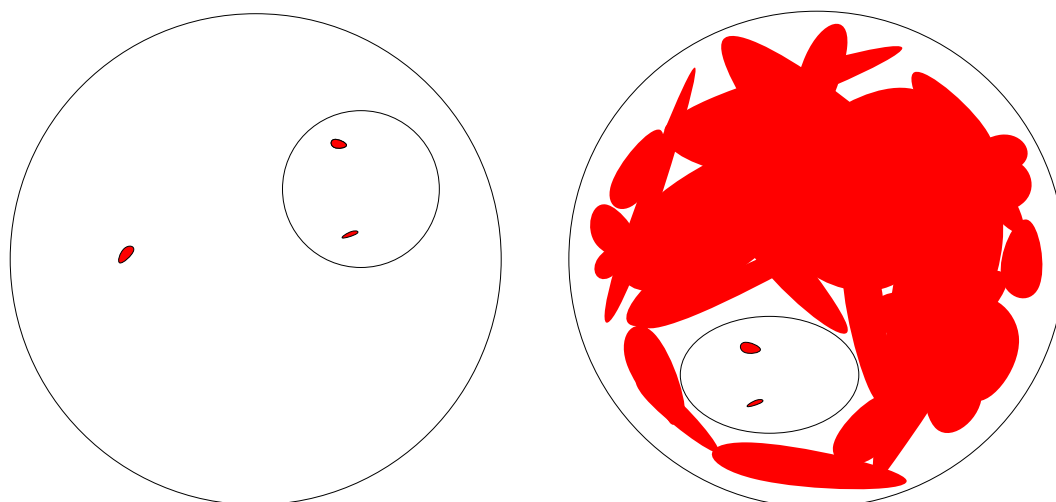
Figure 6: For k given of order n, the two large balls respectively represent the laws P_{m+1} and P_{n+1}, where $m = k + n - \ell$. Imagine that each ball is partitioned into tiny disjoint cells that represent the polygons of the length in question. Inside the two balls, the cells associated to polygons whose length-k first part fails to be charming for a second part extension of length $n - \ell$ are depicted in red. The smaller ball on the left represents globally joined polygons of length $m + 1$. These are transported in a roughly area preserving way to the interior of the ellipse on the right, which represents globally joined polygons of length $n + 1$. Under the standing assumption, the red region on the left has smaller area than does the ball, so that much of the ellipse interior on the right is seen to be not red. Moreover, the ellipse interior on the right represents a non-negligible area of the right-hand ball. Thus, even though the right-hand ball may be predominately red outside the ellipse, a non-negligible region on the right is not red, so that length k first parts are not atypically charming under P_{n+1}.

polygon length equals $m + 1 = n - \ell + k + 1$, is low enough that it is much smaller than the proportion of length-$(m + 1)$ polygons that are global join polygons. (Ample supply of global join polygons means that the latter proportion, while not of unit order, is non-negligible. When we reprise these arguments rigorously, parameter choices may be made, in correspondance to different $o(1)$ terms, so that the failure probability for the simple charm test for a given length first part will indeed be much smaller than the global join polygon proportion.)

Consider now the random mapping of globally joined polygons of length $m + 1$ into their length $n + 1$ counterparts (and consult Figure 6 for an illustration of the argument). This map roughly preserves measure, so that the charm evinced by a length-k first part (namely, that a length-$(n - \ell)$ second part closes with probability at least $n^{-1/2-o(1)}$) is seen to be typical not only under a uniformly chosen globally joined polygon of length $m + 1$ but also under the counterpart measure for length $n + 1$. By the ample supply of globally joined polygons, a non-negligible fraction of length-$(n + 1)$ polygons are globally joined. Therefore, a non-negligible, $n^{-o(1)}$, fraction of such polygons are charming at index k.

By varying k over order n values, we find that the snake method hypothesis (6.2) is verified under the standing assumption, because it must be the case that, with probability at least $n^{-o(1)}$, the P_{n+1}-sampled polygon has at least $n^{1-o(1)}$ indices lengths at which it is charming. Naturally, Theorem 6.3 then applies to show that the closing probability

has a rapid decay, in contradiction to the standing assumption.

In this way, the standing assumption is seen to be false, and the aim of showing that the closing probability decays as quickly as $n^{-1/2-\varepsilon}$ is achieved subsequentially (for some positive ε). The argument demonstrates both the beauty and horror of the author's experience of self-avoiding walk: it is rather delicate and intricate, and it achieves a specific inference, regarding closing probability, by examining in detail interesting geometric questions about walks and polygons, concerning say the proportion of large polygons that are globally joined; however, it leaves these questions unsettled.

8 Proving Theorem 1.5(2)

The paper has seven further sections, and these are devoted to proving Theorem 1.5. We will gather the two set of preliminaries we have assembled and prove the theorem using the snake method implemented via polygon joining.

To summarise the structure of the rest of the paper, we mention first that we will begin with Theorem 1.5(2) because the proof of the first part of the theorem shares the framework of the second but is slightly more technical. We first reduce (in the rest of this section) Theorem 1.5(2) to Proposition 8.1; in Section 9, we state the conclusions of the three main steps that lead to the proof of the proposition; and in the three ensuing sections, we prove the results associated to the respective steps and thereby complete the proof of Theorem 1.5(2). Section 13 presents a counterpart to Proposition 8.1 for Theorem 1.5(1) and collects some useful statements for the proof of this part of the main theorem, which proof is presented with the same three step structure in the final Section 14.

Focussing now on the proof of Theorem 1.5(2), note that it is enough to suppose the existence of θ and ξ and to argue that $\theta + \xi \geq 5/3$, since the latter assertion of Theorem 1.5(2) was justified as it was stated. In order to prove this inequality by finding a contradiction, we suppose that $\theta + \xi < 5/3$. The sought contradiction is exhibited in the next result.

Proposition 8.1. *Let $d = 2$. Assume that the limits $\theta := \lim_{n \in 2\mathbb{N}} \theta_n$ and $\xi = \lim_{n \in \mathbb{N}} \xi_n$ exist. Suppose further that $\theta + \xi < 5/3$. Then, for some $c > 1$ and $\delta > 0$, the set of $n \in 2\mathbb{N} + 1$ for which*

$$\mathsf{W}_n\big(\Gamma \text{ closes}\big) \leq c^{-n^\delta}$$

intersects the shifted dyadic scale $\big[2^i - 1, 2^{i+1} - 1\big]$ for all but finitely many $i \in \mathbb{N}$.

Proof of Theorem 1.5(2). The non-negative limits θ and ξ are hypothesised to exist. If their sum is finite, then the formula (1.4) exhibits a polynomial decay. When the sum is less than $5/3$, this contradicts Proposition 8.1. $\qquad\square$

It remains of course to derive the proposition. Assume throughout the derivation that the proposition's hypotheses are satisfied. Define an exponent χ so that $\theta + \xi = 3/2 + \chi$. Since $\xi \geq 0$ and $\theta \geq 3/2$ (classically and by Theorem 1.3), χ is non-negative. Note that $\mathsf{W}_n\big(\Gamma \text{ closes}\big)$ equals $n^{-1/2-\chi+o(1)}$, and also that we are supposing that $\chi < 1/6$.

The proof of Proposition 8.1 is an application of the snake method. We now set the snake method exponent parameters for this application, taking

- the snake length β equal to one;
- the inverse charm α equal to $1/2 + 2\chi + 4\varepsilon$;
- and the deficit η equal to $\chi + 8\varepsilon$;

where henceforth in proving Proposition 8.1, $\varepsilon > 0$ denotes a given but arbitrarily small quantity. The quantity $\delta = \beta - \eta - \alpha = 1/2 - 3\chi - 12\varepsilon$ must be positive if the snake method is to work, and thus we choose $\varepsilon < (1/6 - \chi)/4$.

9 Three steps to the proof of Proposition 8.1

The argument leading to Proposition 8.1 has three principal steps. We explain these now in outline, stating the conclusion of each one. The proofs follow, in the three sections that follow this one.

9.1 Step one: finding many first parts with high closing probability

The outcome of the first step may be expressed as follows, using the high conditional closing probability set notation $\mathrm{HighFirst}$ introduced in Subsection 6.1.4.

Proposition 9.1. *For $i \in \mathbb{N}$ sufficiently high, there exists $m' \in 2\mathbb{N} \cap \left[2^{i+3}, 2^{i+4}\right]$ such that, writing K' for the set of values $k \in \mathbb{N}$, $2^i \leq k \leq 2^i + 2^{i-2}$, that satisfy*

$$\mathsf{P}_{m'+k}\Big(\Gamma_{[0,k]} \in \mathrm{HighFirst}_{k,k+m'-1}^{\alpha}\Big) \geq 1 - 2^{-i\chi}\,,$$

we have that $|K'| \geq 2^{-8}m'$.

This first step is a soft, Fubini, argument: such index pairs $(k, m'+k)$ are characteristic given our assumption on closing probability decay, and the proposition will follow from Lemma 6.1.

9.2 Step two: individual snake terms are often charming

To apply the snake method, we must verify its fundamental hypothesis (6.2) that charming snakes are not rare under the uniform polygon law. Proposition 9.1 is not adequate for verifying this hypothesis, because the polygon law index $m' + k$ is a variable that changes with k. We want a version of the proposition where this index is fixed. The next assertion, which is the conclusion of our second step, is such a result. Note from the differing forms of the right-hand sides in the two results that a behaviour determined to be highly typical in Proposition 9.1 has a counterpart in Proposition 9.2 which is merely shown to be not unusual.

Proposition 9.2. *For each $i \in \mathbb{N}$ sufficiently high, there exist $m \in 2\mathbb{N} \cap \left[2^{i+4}, 2^{i+5}\right]$ and $m' \in \left[2^{i+3}, 2^{i+4}\right]$ such that, writing K for the set of values $k \in \mathbb{N}$, $1 \leq k \leq m - m'$, that satisfy*

$$\mathsf{P}_m\Big(\Gamma_{[0,k]} \in \mathrm{HighFirst}_{k,k+m'-1}^{\alpha}\Big) \geq m^{-\eta+\varepsilon}\,,$$

we have that $|K| \geq 2^{-9}m$.

The second step, leading to the proof of this proposition, is quite subtle. Expressed in its simplest terms, the idea of the proof is that the subscript index in P can be changed from the k-dependent $m' + k$ in Proposition 9.1 to the k-independent m in the new result by using the polygon joining technique. This index change involves proving a *similarity of measure* between polygon laws with distinct indices on a given dyadic scale. The set of regulation global join polygons introduced in Section 5 provides a convenient reservoir of macroscopically joined polygons whose left polygons are shared between polygon laws of differing index, so that this set provides a means by which this similarity may be proved. Figure 7 offers some further outline of how we will exploit the ample supply of regulation polygons to prove similarity of measure.

9.3 Step three: applying the snake method Theorem 6.3

In this step, it is our goal to use Proposition 9.2 in order to verify the snake method hypothesis (6.2). In the step, we will fix the method's two index parameters: for $i \in \mathbb{N}$ given, we will use Proposition 9.2 to take n equal to $m - 1$, and ℓ equal to $m - m'$.
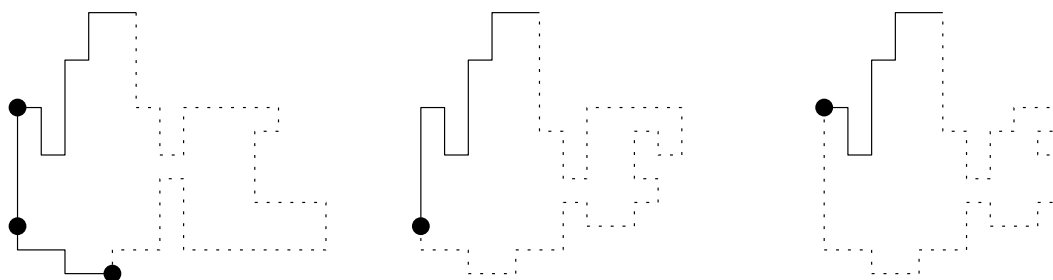
Figure 7: The snake method via polygon joining. Black dots mark vertices of polygons γ: fixing a typical $m' \in 2\mathbb{N}$, assign a black dot at each vertex γ_j, $m'/4 \leq j \leq 3m'/4$, of a polygon γ of length of order m', that has the property that $\mathrm{SAW}^0_{m'+j}\big(\Gamma \text{ closes} \,\big|\, |\Gamma^1| = j, \Gamma^1 = \gamma_{[0,j]}\big) \geq (m')^{-1/2-\chi-\varepsilon}$. Our assumptions readily imply that most polygons γ with length n' of order say $2m'$ will have a black dot at $\gamma_{n'-m'}$. What about at other locations γ_j where j is also of order m'? A black dot is likely to appear at Γ_j when Γ is $\mathsf{P}_{j+m'}$-distributed. If we can make a *comparison of measure*, showing that the laws $\mathsf{P}_{j+m'}$ and $\mathsf{P}_{n'}$ are to some degree similar, then the black dot will be known to also appear at γ_j in a typical sample of $\mathsf{P}_{n'}$. Applying this for many such j, black dots will appear along the course of a typical length n' polygon. These black dots index charming snake terms and permit the use of the snake method. Comparison of measure will be undertaken in Proposition 11.2 by the use of polygon joining. Lemma 11.1 shows that a non-negligible proportion of polygons are regulation global join polygons; thus, such polygons themselves typically have black dots m' steps from their end. The law of the left polygon under the uniform law on regulation polygons of a given length is largely unchanged as that total length is varied on a given dyadic scale, because the length discrepancy can be absorbed by altering the length of the right polygon. Indeed, in the above sketches, we deplete the length of the right polygon in a possible extension of the depicted first part, as the length of this first part shortens; in this way, we show that first parts typically arising in regulation polygons at one length are also characteristic in such polygons with lengths on the same scale.

We may now emphasise why the proposition is useful for this verification goal. Note that $\max K \leq \ell$. Since $n - \ell = m' - 1$, we have that, for $k \in [1, \ell]$, the walk $\gamma \in \mathrm{First}_{\ell,n}$ is (α, n, ℓ)-charming at index k if and only if

$$\mathsf{W}^0_{k+m'-1}\Big(\Gamma \text{ closes} \,\Big|\, |\Gamma^1| = k\,,\, \Gamma^1_{[0,k]} = \gamma_{[0,k]}\Big) > n^{-\alpha}\,;$$

note the displayed condition is almost the same as $\gamma_{[0,k]} \in \mathrm{HighFirst}^\alpha_{k,m'+k-1}$, the latter occurring when $n^{-\alpha}$ is replaced by $(m' + k - 1)^{-\alpha}$, a change which involves only a bounded factor.

Thus, Proposition 9.2 shows that individual snake terms are charming with probability at least $m^{-\chi-o(1)}$. It will be a simple Fubini argument that will allow us to verify (6.2) by confirming that it is not rare that such terms gather together to form charming snakes. Theorem 6.3 may be invoked to prove Proposition 8.1 immediately. The third step will provide this Fubini argument.

The next three sections are devoted to implementing these three steps. In referring to the dyadic scale index $i \in \mathbb{N}$ in results stated in these sections, we will sometimes neglect to record that this index is assumed to be sufficiently high.

10 Step one: the proof of Proposition 9.1

Before proving this result, we review its meaning in light of Figure 7. Proposition 9.1 is a precisely stated counterpart to the assertion in the figure's caption that a black dot typically appears m' steps from the end of a polygon drawn from $\mathsf{P}_{n'}$ where n' has order $2m'$ (or say 2^{i+3}). Note the discrepancy that we have chosen α to be marginally above $1/2 + 2\chi$ while in the caption a choice just above $1/2 + \chi$ was made. The extra margin permits $1 - 2^{-i\chi}$ rather than $1 - o(1)$ in the conclusion of the proposition. We certainly need the extra margin; we will see how at the end of step two.

Consider two parameters $\epsilon_2 > \epsilon_1 > 0$. (During the proof of Proposition 9.2, we will ultimately set $\epsilon_2 = 4\epsilon_1$ and $\epsilon_1 = \varepsilon$, where $\varepsilon > 0$ is the parameter that has been used to specify the three snake method exponents.)

Definition 10.1. *Let $i \in \mathbb{N}$. An index pair $(k, j) \in \mathbb{N} \times 2\mathbb{N}$ will be called* closing probability typical *on dyadic scale i if*

- $k \in \left[2^i, 2^i + 2^{i-2}\right]$ *and* $j \in \left[2^{i+4}, 2^{i+5}\right]$;
- *and, for any* $a \geq 0$,

$$\mathsf{P}_j\left(\Gamma_{[0,k]} \notin \mathrm{HighFirst}_{k,j-1}^{1/2+(a+1)\chi+\epsilon_2}\right) \leq (j-1)^{-(a\chi+\epsilon_2-\epsilon_1)}. \tag{10.1}$$

Proposition 10.2. *For all $i \in \mathbb{N}$ high enough, there exists $k \in 2\mathbb{N} \cap \left[2^{i+4}, 2^{i+4} + 2^i\right]$ for which there are at least 2^{i-4} values of $j \in 2\mathbb{N} \cap [0, 2^{i-2}]$ with the index pair $\left(2^i + j, k + j\right)$ closing probability typical on dyadic scale i.*

When we prove Proposition 9.2 in step two, we will in fact do so using a direct consequence of Proposition 10.2, rather than Proposition 9.1. The latter result, a byproduct of the proof of Proposition 10.2, has been stated merely because it permitted us to make a direct comparison between steps one and two in the preceding outline. As such, our goal in step one is now to prove Proposition 10.2.

Definition 10.3. *Define E to be the set of pairs $(k, j) \in \mathbb{N} \times 2\mathbb{N}$,*

- *where $k \in \left[2^i, 2^i + 2^{i-2}\right]$ and $j \in \left[2^{i+4}, 2^{i+5}\right]$;*
- *and the pair (k, j) is such that*

$$\#\left\{\gamma \in \mathrm{SAW}_{j-1}^0 : |\gamma^1| = k\right\} \tag{10.2}$$

$$< \quad (j-1)^{1/2+\chi+\epsilon_1} \cdot \#\left\{\gamma \in \mathrm{SAW}_{j-1}^0 : |\gamma^1| = k, \gamma \text{ closes}\right\}.$$

We will now apply Lemma 6.1 to show that membership of E is typical. The application is possible because by hypothesis there is a lower bound on the decay of the closing probability. Indeed, we know that $\mathsf{W}_n\left(\Gamma \text{ closes}\right) \geq n^{-1/2-\chi-o(1)}$. Thus, the lemma with $\alpha' = 1/2 + \chi + \epsilon_1/2$ and $\delta' = \epsilon_1/2$ proves the following.

Lemma 10.4. *Provided the index $i \in \mathbb{N}$ is sufficiently high, for each $j \in 2\mathbb{N} \cap \left[2^{i+4}, 2^{i+5}\right]$, the set of $k \in \left[2^i, 2^i + 2^{i-2}\right]$ such that $(k, j) \notin E$ has cardinality at most $2j^{1-\epsilon_1/2} \leq 2\left(2^{i+5}\right)^{1-\epsilon_1/2}$.*

Lemma 10.5. *Any pair $(k, j) \in E$ is closing probability typical on dyadic scale i.*

Proof. We must verify that (k, j) verifies (10.1) for any $a \geq 0$. Recalling the notation W_{j-1}^0 from Subsection 6.1.2, note that the assertion

$$\mathsf{W}_{j-1}^0\left(\Gamma^1 \notin \mathrm{HighFirst}_{k,j-1}^{1/2+(a+1)\chi+\epsilon_2} \;\middle|\; \Gamma \text{ closes}, |\Gamma^1| = k\right) \leq (j-1)^{-(a\chi+\epsilon_2-\epsilon_1)} \tag{10.3}$$

is a reformulation of this condition. Note that

$$\#\left\{\gamma \in \mathrm{SAW}^0_{j-1} : |\gamma^1| = k \,,\, \gamma^1 \notin \mathrm{HighFirst}^{1/2+(a+1)\chi+\epsilon_2}_{k,j-1}\right\} \tag{10.4}$$

is at least the product of $(j-1)^{1/2+(a+1)\chi+\epsilon_2}$ and

$$\#\left\{\gamma \in \mathrm{SAW}^0_{j-1} : |\gamma^1| = k \,,\, \gamma^1 \notin \mathrm{HighFirst}^{1/2+(a+1)\chi+\epsilon_2}_{k,j-1} \,,\, \gamma \text{ closes}\right\}.$$

The quantity (10.4) is bounded above by (10.2); thus is (10.3) obtained. □

Proof of Proposition 10.2. By Lemma 10.5, it is enough to argue that there exists $k \in 2\mathbb{N} \cap \left[2^{i+4}, 2^{i+4} + 2^i\right]$ for which there are at least 2^{i-4} values of $j \in 2\mathbb{N} \cap [0, 2^{i-2}]$ with $\left(2^i + j, k + j\right) \in E$.

By Lemma 10.4,

$$\sum_{j \in 2\mathbb{N} \cap [2^{i+4}+2^{i-2}, 2^{i+4}+2^i]} \sum_{\ell=2^i}^{2^i+2^{i-2}} \mathbb{1}_{(\ell,j) \in E} \tag{10.5}$$

$$\geq\ \left(2^{i-2} + 1 - 2\left(2^{i+5}\right)^{1-\epsilon_1/2}\right) \cdot \tfrac{1}{2}\left(2^i - 2^{i-2}\right).$$

As Figure 8 illustrates, the left-hand side here is bounded above by

$$\sum_{k \in 2\mathbb{N} \cap [2^{i+4}, 2^{i+4}+2^i]} \sum_{\ell=2^i}^{2^i+2^{i-2}} \mathbb{1}_{(\ell, k+\ell-2^i) \in E}\,. \tag{10.6}$$

There being $2^{i-1} + 1 \leq 2^i$ indices $k \in 2\mathbb{N} \cap \left[2^{i+4}, 2^{i+4} + 2^i\right]$, one such k satisfies

$$\sum_{\ell=2^i}^{2^i+2^{i-2}} \mathbb{1}_{(\ell, k+\ell-2^i) \in E} \geq \tfrac{3}{8}\left(2^{i-2} + 1 - 2\left(2^{i+5}\right)^{1-\epsilon_1/2}\right). \tag{10.7}$$

Noting that $\tfrac{3}{4}\left(2^{i+5}\right)^{1-\epsilon_1/2} \leq \tfrac{1}{8}2^{i-2}$, we obtain the sought statement and so conclude the proof. □

Proof of Proposition 9.1. Proposition 10.2 furnishes the existence of a minimal $Q \in 2\mathbb{N} \cap \left[2^{i+4}, 2^{i+4} + 2^i\right]$ such that there are at least 2^{i-4} values of $k \in 2\mathbb{N} \cap [0, 2^{i-2}]$ with $\left(2^i + k, Q + k\right)$ closing probability typical on dyadic scale i.

Let $\left(k_1, \ldots, k_{2^{i-4}}\right)$ be an increasing sequence of such values of k associated to the value Q. Set $m' = Q - 2^i$ and note that $m' \in \left[2^{i+3}, 2^{i+4}\right]$. Take $a = 1$ in (10.1) and specify that $\epsilon_2 = 4\varepsilon$ alongside $\epsilon_2 > \epsilon_1$ to find that, for all $j \in [1, 2^{i-4}]$, $2^i + k_j \in K'$. Thus, $|K'| \geq 2^{i-4} \geq 2^{-8}m'$.

This completes the proof of the proposition. □

We end this section by recording the explicit artefact that will be used in step two.

Definition 10.6. *Set $L = 2^{i-4}$. We specify a constant difference sequence of index pairs (r_j, s_j), $1 \leq j \leq L$. Its elements are closing probability typical for dyadic scale i, the difference $s_j - r_j$ is independent of j and takes a value in $[2^{i+4} - 2^i, 2^{i+4}]$, and $\{r_j : 1 \leq j \leq L\}$ is increasing. Moreover, $r_j \in 2\mathbb{N} \cap [2^i, 2^i + 2^{i-2}]$ and $s_j \in 2\mathbb{N} \cap [2^{i+4}, 2^{i+4}+2^i+2^{i-2}]$ for each $j \in [1, L]$.*

Indeed, we may work with the sequence constructed in the preceding proof, setting $r_j = 2^i + k_j$ and $s_j = Q + k_j$ for $j \in [1, L]$.

We are *en route* to Proposition 9.2, and take the opportunity to mention how its two parameters m and m' will be specified. The latter will be chosen, as it was in the proof of Proposition 9.1, to be the value of the constant difference in the sequence in Definition 10.6, while the former will be taken equal to s_L in this definition.
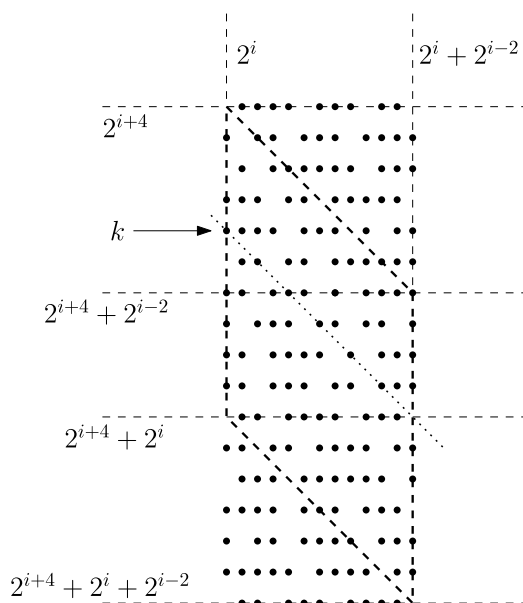
Figure 8: Illustrating the proof of Proposition 10.2. The dots depict elements of E. The quantity (10.5) is the number of dots in the middle rectangle; (in these terms, ℓ is the x-coordinate in the sketch and j is the negative y-coordinate). Similarly, the quantity (10.6) is the number of dots in the parallelogram bounded by dashed lines. In (10.7), we find a diagonal, such as the one depicted with a pointed line, on which many dots lie.

11 Step two: deriving Proposition 9.2

The key tool for this proof is similarity of measure between polygon laws with distinct indices, with the reservoir of regulation global join polygons providing the mechanism for establishing this similarity. Recall that these polygons were introduced and discussed in Section 5; the reader may wish to review this material now.

This step two section has three subsections. In the first, we provide a piece of apparatus, which augments the regulation polygon properties discussed in Section 5, that we will need in order to state our similarity of measure result. This result, Proposition 11.2, is stated and proved in the next subsection. The final subsection gives the proof of Proposition 9.2.

11.1 Relating polygon laws P_n for distinct n via global join polygons

The next lemma shows that there is an ample supply of regulation polygons in the sense that they are not rare among all polygons. In contrast to the regulation polygon tools in Section 5, Lemma 11.1 is contingent on our standing assumption that the hypotheses of Proposition 8.1 are satisfied. (In fact, it is the existence and finiteness of θ and ξ that are needed, rather than the condition $\theta + \xi < 5/3$.)

Define $\mathsf{R}_i := 2\mathbb{N} \cap \left[2^i, 2^{i+1}\right]$.

Lemma 11.1. *Let $\varphi > 0$. For any $i \in \mathbb{N}$ sufficiently high, and $n \in 2\mathbb{N} \cap \left[2^{i+4}, 2^{i+5}\right]$,*

$$\mathsf{P}_n\left(\Gamma \in \bigcup_{k \in \mathsf{R}_{i+2}} \mathrm{RGJ}_{k,n-16-k}\right) \geq n^{-\chi-\varphi}.$$

The set R_i may be thought as a set of *regular* indices $j \in 2\mathbb{N} \cap \left[2^i, 2^{i+1}\right]$, with θ_j for example being neither atypically high nor low. As we prove Theorem 1.5(2), all indices

are regarded as regular, so that we make the trivial specification $\mathsf{R}_i = 2\mathbb{N} \cap \left[2^i, 2^{i+1}\right]$; the set will be respecified non-trivially in the proof of Theorem 1.5(1).

(As we explained in the three-step outline, the snake method parameters, including n, will be set for the proof of Proposition 8.1 at the start of step three. Until then, we are treating n as a free variable.)

Proof of Lemma 11.1. Note that the probability in question is equal to the ratio of the cardinality of the union $\bigcup_{j \in \mathsf{R}_{i+2}} \mathrm{RGJ}_{j,n-16-j}$ and the polygon number p_n.

In order to bound below the numerator in this ratio, we aim to apply Proposition 5.10 with the role of $i \in \mathbb{N}$ played by $i + 2$ and with $\mathsf{R} = \mathsf{R}_{i+2}$. In order to do so, set $\overline{\theta} = \sup_{n \in 2\mathbb{N}} \theta_n$, a quantity that is finite by hypothesis. Set the proposition's Θ equal to $\overline{\theta}$. The index k^* in the proposition may then be any element of R_{i+2}. Applying the proposition, we find that there exists $c = c(\Theta) > 0$ such that

$$\left| \bigcup_{j \in \mathsf{R}_{i+2}} \mathrm{RGJ}_{j,n-16-j} \right| \geq c \, \frac{n^{1/2}}{\log n} \sum_{j \in \mathsf{R}_{i+2}} p_j p_{n-16-j} \, .$$

The existence of θ implies that $p_m \mu^{-m} \in \left[m^{-\theta-\varphi}, m^{-\theta+\varphi}\right]$ whenever $m \geq m_0(\varphi)$. Since $i \in \mathbb{N}$ is supposed sufficiently high, we thus find that

$$\mathsf{P}_n\!\left(\Gamma \in \bigcup_{j \in \mathsf{R}_{i+2}} \mathrm{RGJ}_{j,n-16-j}\right)$$
$$\geq \; \mu^{-16} c \, \frac{n^{1/2}}{\log n} \sum_{j \in \mathsf{R}_{i+2}} j^{-(\theta+\varphi)}(n-16-j)^{-(\theta+\varphi)} \cdot n^{\theta-\varphi} \geq 2^{-4} \mu^{-16} c \, \frac{n^{3/2-\theta-3\varphi}}{\log n} \, ,$$

where we used $\#\mathsf{R}_{i+2} \geq 2^{i+1} \geq 2^{-4}n$. Noting that $\theta \leq 3/2 + \chi$ and relabelling $\varphi > 0$ completes the proof. $\qquad\square$

11.2 Similarity of measure

We may now state our similarity of measure Proposition 11.2. Recall from Definition 10.6 the constant difference sequence (r_j, s_j), $1 \leq j \leq L$, and also the regular index set R_{i+2} in the preceding subsection.

Proposition 11.2. *For each $j \in [1, L]$, there is a subset $D_j \subseteq \mathrm{First}_{r_j}$ satisfying*

$$\mathsf{P}_{s_j}\!\left(\Gamma_{[0,r_j]} \notin D_j \, , \, \Gamma \in \mathrm{RGJ}_{k,s_j-16-k} \text{ for some } k \in \mathsf{R}_{i+2}\right) \leq 4\mu^{16} s_j^{-10} \tag{11.1}$$

such that, for each $\phi \in D_j$,

$$\mathsf{P}_{s_L}\!\left(\Gamma_{[0,r_j]} = \phi\right)$$
$$\geq \; \tfrac{1}{40} C_8^{-1}\left(\log s_L\right)^{-1}\left(\tfrac{3}{16}\right)^{\theta} s_L^{-4\varepsilon}$$
$$\times \; \mathsf{P}_{s_j}\!\left(\Gamma_{[0,r_j]} = \phi \, , \, \Gamma \in \mathrm{RGJ}_{k,s_j-16-k} \text{ for some } k \in \mathsf{R}_{i+2}\right).$$

Preparing for the proof, we define, for each $\gamma \in \mathrm{SAP}_{s_j}$, $1 \leq j \leq L$, the set of γ's regulation global join indices

$$\tau_\gamma = \left\{ k \in \mathsf{R}_{i+2} : \gamma \in \mathrm{RGJ}_{k,s_j-16-k} \right\}.$$

For $\gamma \in \mathrm{SAP}_{s_j}$, the map sending $k \in \tau_\gamma$ to the junction plaquette associated to the join of elements of $\mathrm{SAP}_k^{\mathrm{left}}$ and $\mathrm{SAP}_{s_j-16-k}^{\mathrm{right}}$ is an injective map into GJ_γ: see the first

assertion of Proposition 5.7. Thus, $|\tau_\gamma| \leq |\mathsf{GJ}_\gamma|$. Corollary 3.3 implies that there exists a constant $C_8 > 0$ such that

$$\mathsf{P}_{s_L}\left(|\tau_\Gamma| \geq C_8 \log s_L\right) \leq C_8 \, s_L^{-4\overline{\theta}-22} \,, \tag{11.2}$$

where here we have reintroduced $\overline{\theta} = \sup_{n \in 2\mathbb{N}} \theta_n$, a quantity which we recall is finite by hypothesis; (this finiteness permits the application of the corollary).

We make use of this constant to specify for each $j \in [1, L]$ a set $D_j \subseteq \mathrm{First}_{r_j}$ of length r_j first parts that, when extended to form a length s_L polygon, do not typically produce far-above-average numbers of regulation joinings. We define

$$D_j = \left\{\phi \in \mathrm{First}_{r_j} : \mathsf{P}_{s_L}\left(|\tau_\Gamma| \leq C_8 \log s_L \,\Big|\, \Gamma_{[0,r_j]} = \phi\right) \geq 1 - C_8 \, s_L^{-2\overline{\theta}-11}\right\},$$

The use of the next lemma lies in its final part, which establishes the P_{s_j}-rarity assertion (11.1). The first two parts of the lemma are steps towards the third. The rarity under P_{s_L} of the length r_j initial path failing to be in D_j is a direct consequence of definitions, as the first part of the lemma shows, but a little work is needed to make the change of index $L \to j$.

Lemma 11.3. *Let $j \in [1, L]$.*

1. *We have that*
$$\mathsf{P}_{s_L}\left(\Gamma_{[0,r_j]} \in D_j^c\right) \leq s_L^{-2\overline{\theta}-11} \,.$$

2. *Let $k \in \mathsf{R}_{i+2}$. Then*
$$\mathsf{P}_k^{\mathrm{left}}\left(\Gamma_{[0,r_j]} \notin D_j\right) \leq 16\mu^{16} s_L^{-11} \,.$$

3. *We have that*
$$\mathsf{P}_{s_j}\left(\Gamma_{[0,r_j]} \notin D_j \,,\, \Gamma \in \mathrm{RGJ}_{k,s_j-16-k} \ \textit{for some } k \in \mathsf{R}_{i+2}\right) \leq 4\mu^{16} s_j^{-10} \,.$$

Proof (1). Note that

$$\begin{aligned}
\mathsf{P}_{s_L}\left(|\tau_\Gamma| \geq C_8 \log s_L\right) &\geq \mathsf{P}_{s_L}\left(|\tau_\Gamma| \geq C_8 \log s_L \,,\, \Gamma_{[0,r_j]} \in D_j^c\right) \\
&= \mathsf{P}_{s_L}\left(\Gamma_{[0,r_j]} \in D_j^c\right) \mathsf{P}_{s_L}\left(|\tau_\Gamma| \geq C_8 \log s_L \,\Big|\, \Gamma_{[0,r_j]} \in D_j^c\right) \\
&\geq C_8 s_L^{-2\overline{\theta}-11} \, \mathsf{P}_{s_L}\left(\Gamma_{[0,r_j]} \in D_j^c\right) .
\end{aligned}$$

Lemma 11.3(1) thus follows from (11.2).

(2). In light of the preceding part, it suffices to prove

$$\mathsf{P}_k^{\mathrm{left}}\left(\Gamma_{[0,r_j]} \notin D_j\right) \leq 16\mu^{16} s_L^{2\overline{\theta}} \, \mathsf{P}_{s_L}\left(\Gamma_{[0,r_j]} \notin D_j\right). \tag{11.3}$$

We will derive this bound with a simple joining argument. Note that $s_L \geq k + 16$ since $s_L \geq 2^{i+4}$, $k \leq 2^{i+3}$ and $i \geq 1$. Consider the map

$$\Psi : \left\{\phi \in \mathrm{SAP}_k^{\mathrm{left}} : \phi_{[0,r_j]} \in D_j^c\right\} \times \mathrm{SAP}_{s_L-k-16}^{\mathrm{right}} \to \left\{\phi \in \mathrm{SAP}_{s_L} : \phi_{[0,r_j]} \in D_j^c\right\}$$

defined by setting $\Psi(\phi_1, \phi_2)$ equal to the join polygon $J(\phi_1, \phi_2 + \vec{u})$, where the second argument is chosen so that the pair is globally Madras joinable. (A choice of $\vec{u}$ ensuring that the pair is *globally* joinable may be made by placing the second polygon strictly below the x-axis and then displacing it horizontally to the rightmost location at which it is Madras joinable to the first polygon.)

We now wish to apply Lemma 5.11 to learn that $\Psi(\phi, \phi')_{[0,r_j]} = \phi_{[0,r_j]}$, so that Ψ's image lies in the stated set. In order to do so, it is enough to verify that $r_j \leq k/2 - 1$, since $\phi \in \mathrm{SAP}_k$ is left-long. This bound follows from $k \geq 2^{i+2}$, $r_j \leq 2^i + 2^{i-2}$ (as noted in Definition 10.6), and $i \geq 1$.

The decomposition uniqueness in Proposition 5.7 shows that Ψ is injective. Thus, by considering Ψ, we find that

$$\#\big\{\phi \in \mathrm{SAP}_{s_L} : \phi_{[0,r_j]} \in D_j^c\big\} \geq \#\big\{\phi \in \mathrm{SAP}_k^{\mathrm{left}} : \phi_{[0,r_j]} \in D_j^c\big\} \#\big\{\phi \in \mathrm{SAP}_{s_L-k-16}^{\mathrm{right}}\big\}.$$

By Lemma 5.5,

$$\mathsf{P}_{s_L}\big(\Gamma_{[0,r_j]} \in D_j^c\big) \geq \tfrac{1}{16}\mathsf{P}_k^{\mathrm{left}}\big(\Gamma_{[0,r_j]} \in D_j^c\big) p_{s_L}^{-1} p_k p_{s_L-k-16}.$$

Note that $p_{s_L}^{-1} p_k p_{s_L-k-16} \geq \mu^{-16} s_L^{-2\overline{\theta}}$, where recall that $\overline{\theta} = \sup_{n \in 2\mathbb{N}} \theta_n$. In this way, we obtain (11.3).

(3). We first make the key observation that, for $k \in \mathsf{R}_{i+2}$ and $j' \in [1, L]$,

$$\mathsf{P}_{s_{j'}}\Big(\Gamma_{[0,r_j]} = \phi \,\Big|\, \Gamma \in \mathrm{RGJ}_{k,s_{j'}-16-k}\Big) = \mathsf{P}_k^{\mathrm{left}}\Big(\Gamma_{[0,r_j]} = \phi\Big). \tag{11.4}$$

Proposition 5.9 implies this, provided that its hypotheses that $r_j \leq k/2 - 1$ and $k/2 \leq s_{j'} - 16 - k \leq 35k$ are valid. That the first inequality holds has been noted in the proof of the lemma's second part. That the second holds follows from $2^{i+3} \geq k \geq 2^{i+2}$, $2^{i+5} \geq s_{j'} \geq 2^{i+4}$ and $i \geq 2$.

The quantity on the left-hand side of the inequality in Lemma 11.3(3) is at most

$$\sum_{k \in \mathsf{R}_{i+2}} \mathsf{P}_{s_j}\Big(\Gamma_{[0,r_j]} \notin D_j \,,\, \Gamma \in \mathrm{RGJ}_{k,s_j-16-k}\Big)$$

and thus by (11.4) with $j' = j$, at most $|\mathsf{R}_{i+2}| \max_{k \in \mathsf{R}_{i+2}} \mathsf{P}_k^{\mathrm{left}}\big(\Gamma_{[0,r_j]} \notin D_j\big)$. We may now apply the second part of the lemma as well as the bounds $|\mathsf{R}_{i+2}| \leq 4^{-1} s_L$ and $s_L \geq s_j$ to obtain Lemma 11.3(3). (This upper bound on $|\mathsf{R}_{i+2}|$ is due to $|\mathsf{R}_{i+2}| = 2^{i+1} + 1 \leq 2^{i+2}$ and $s_L \geq 2^{i+4}$.) $\qquad\square$

Lemma 11.4. *Recall that $L = 2^{i-4}$. Let $j \in [1, L]$.*

1. *For $\phi \in D_j$,*

$$\mathsf{P}_{s_L}\Big(\Gamma_{[0,r_j]} = \phi\Big) \geq \tfrac{1}{400} C_8^{-1}\big(\log s_L\big)^{-1} s_L^{-2\varepsilon} \mu^{-16} \sum_{k \in \mathsf{R}_{i+2}} \mathsf{P}_k^{\mathrm{left}}\Big(\Gamma_{[0,r_j]} = \phi\Big) k^{1/2} p_k \mu^{-k}.$$

2. *For $\phi \in \mathrm{First}_{r_j}$,*

$$\mathsf{P}_{s_j}\Big(\Gamma_{[0,r_j]} = \phi \,,\, \Gamma \in \mathrm{RGJ}_{k,s_j-16-k} \text{ for some } k \in \mathsf{R}_{i+2}\Big)$$
$$\leq \quad \tfrac{1}{10}\big(\tfrac{16}{3}\big)^{\theta} \mu^{-16} \cdot s_j^{2\varepsilon} \sum_{k \in \mathsf{R}_{i+2}} \mathsf{P}_k^{\mathrm{left}}\Big(\Gamma_{[0,r_j]} = \phi\Big) k^{1/2} p_k \mu^{-k}.$$

Proof: *(1).* Note that $\mathsf{P}_{s_L}\big(\Gamma_{[0,r_j]} = \phi\big)$ is at least

$$\mathsf{P}_{s_L}\Big(\Gamma_{[0,r_j]} = \phi \,,\, \Gamma \in \bigcup_{k \in \mathsf{R}_{i+2}} \mathrm{RGJ}_{k,s_L-16-k} \,,\, \big|\tau_\Gamma\big| \leq C_8 \log s_L\Big)$$
$$\geq \quad \big(C_8 \log s_L\big)^{-1} \sum_{k \in \mathsf{R}_{i+2}} \mathsf{P}_{s_L}\Big(\Gamma_{[0,r_j]} = \phi \,,\, \Gamma \in \mathrm{RGJ}_{k,s_L-16-k} \,,\, \big|\tau_\Gamma\big| \leq C_8 \log s_L\Big).$$

The summand in the last line may be written

$$P_{s_L}\left(\Gamma_{[0,r_j]} = \phi\,,\, \Gamma \in \mathrm{RGJ}_{k,s_L-16-k}\right)$$

$$-\quad P_{s_L}\left(\Gamma_{[0,r_j]} = \phi\,,\, \Gamma \in \mathrm{RGJ}_{k,s_L-16-k}\,,\, \big|\tau_\Gamma\big| > C_8 \log s_L\right),$$

with the subtracted term here being at most

$$P_{s_L}\left(\Gamma_{[0,r_j]} = \phi\,,\, \big|\tau_\Gamma\big| > C_8 \log s_L\right) \le C_8 s_L^{-2\overline{\theta}-11} P_{s_L}\left(\Gamma_{[0,r_j]} = \phi\right)$$

since $\phi \in D_j$. We noted at the end of the preceding proof that $|\mathsf{R}_{i+2}| \le 4^{-1}s_L$. Thus, $P_{s_L}\left(\Gamma_{[0,r_j]} = \phi\right)$ is at least

$$\left(C_8 \log s_L\right)^{-1} \sum_{k \in \mathsf{R}_{i+2}} P_{s_L}\left(\Gamma_{[0,r_j]} = \phi\,,\, \Gamma \in \mathrm{RGJ}_{k,s_L-16-k}\right) \tag{11.5}$$

$$-\quad \left(C_8 \log s_L\right)^{-1} 4^{-1}s_L \cdot C_8 s_L^{-2\overline{\theta}-11} P_{s_L}\left(\Gamma_{[0,r_j]} = \phi\right).$$

Taking $j' = L$ in (11.4) and using Proposition 5.7, we learn that the summand in the first line satisfies

$$P_{s_L}\left(\Gamma_{[0,r_j]} = \phi\,,\, \Gamma \in \mathrm{RGJ}_{k,s_L-16-k}\right) \tag{11.6}$$

$$=\quad P_k^{\mathsf{left}}\left(\Gamma_{[0,r_j]} = \phi\right) P_{s_L}\left(\Gamma \in \mathrm{RGJ}_{k,s_L-16-k}\right)$$

$$=\quad P_k^{\mathsf{left}}\left(\Gamma_{[0,r_j]} = \phi\right) \cdot p_{s_L}^{-1} \cdot \lfloor \tfrac{1}{10} k^{1/2} \rfloor \cdot \big|\mathrm{SAP}_k^{\mathsf{left}}\big| \big|\mathrm{SAP}_{s_L-16-k}^{\mathsf{right}}\big|,$$

and thus by Lemma 5.5 (as well as $k \ge 2^{i+2}$ and $i \ge 7$) is at least

$$P_k^{\mathsf{left}}\left(\Gamma_{[0,r_j]} = \phi\right) \cdot p_{s_L}^{-1} \cdot \tfrac{1}{20} k^{1/2} \cdot \tfrac{1}{8} p_k \cdot \tfrac{1}{2} p_{s_L-16-k}\,.$$

Returning to (11.5), we find that $P_{s_L}\left(\Gamma_{[0,r_j]} = \phi\right)$ is at least

$$\left(1 + 4^{-1}\left(\log s_L\right)^{-1} s_L^{-2\overline{\theta}-10}\right)^{-1} \left(C_8 \log s_L\right)^{-1} \tfrac{1}{320} p_{s_L}^{-1}$$

$$\times \sum_{k \in \mathsf{R}_{i+2}} P_k^{\mathsf{left}}\left(\Gamma_{[0,r_j]} = \phi\right) k^{1/2} p_k p_{s_L-16-k} \tag{11.7}$$

$$\ge\quad \left(C_8 \log s_L\right)^{-1} \tfrac{1}{400} p_{s_L}^{-1} \mu^{s_L-16}$$

$$\times \sum_{k \in \mathsf{R}_{i+2}} P_k^{\mathsf{left}}\left(\Gamma_{[0,r_j]} = \phi\right) k^{1/2} p_k \mu^{-k} (s_L - 16 - k)^{-\theta-\varepsilon}\,.$$

Using $p_{s_L} \le \mu^{s_L} s_L^{-\theta+\varepsilon}$, we obtain Lemma 11.4(1).

(2). For $\phi \in \mathrm{First}_{r_j}$,

$$P_{s_j}\left(\Gamma_{[0,r_j]} = \phi\,,\, \Gamma \in \mathrm{RGJ}_{k,s_j-16-k} \text{ for some } k \in \mathsf{R}_{i+2}\right)$$

$$\le\quad \sum_{k \in \mathsf{R}_{i+2}} P_{s_j}\left(\Gamma_{[0,r_j]} = \phi\,,\, \Gamma \in \mathrm{RGJ}_{k,s_j-16-k}\right).$$

By taking $j' = j$ in (11.4), we see that the two equalities in the display (11.6) hold with s_j in place of s_L. Since $\big|\mathrm{SAP}_k^{\mathsf{left}}\big| \le p_k$ and $\big|\mathrm{SAP}_{s_j-16-k}^{\mathsf{right}}\big| \le p_{s_j-16-k}$, the last sum is seen

to be at most

$$
\tfrac{1}{10}\, p_{s_j}^{-1} \sum_{k\in\mathsf{R}_{i+2}} \mathsf{P}_k^{\mathrm{left}}\!\Big(\Gamma_{[0,r_j]}=\phi\Big) k^{1/2} p_k p_{s_j-16-k} \tag{11.8}
$$

$$
\leq\quad \tfrac{1}{10} s_j^{\theta+\varepsilon}\mu^{-16} \sum_{k\in\mathsf{R}_{i+2}} \mathsf{P}_k^{\mathrm{left}}\!\Big(\Gamma_{[0,r_j]}=\phi\Big) k^{1/2} p_k \mu^{-k}(s_j-16-k)^{-\theta+\varepsilon}
$$

$$
\leq\quad \tfrac{1}{10}\big(\tfrac{16}{3}\big)^{\theta}\mu^{-16}\cdot s_j^{2\varepsilon} \sum_{k\in\mathsf{R}_{i+2}} \mathsf{P}_k^{\mathrm{left}}\!\Big(\Gamma_{[0,r_j]}=\phi\Big) k^{1/2} p_k \mu^{-k} ,
$$

where we used $2^{i+5}\geq s_j\geq 2^{i+4}$, $k\leq 2^{i+3}$ and $i\geq 3$ in the form $s_j-16-k\geq 3s_j/16$. $\square$

Proof of Proposition 11.2. A consequence of Lemmas 11.3(3) and 11.4. $\square$

11.3 Obtaining Proposition 9.2 via similarity of measure

This subsection is devoted to the next proof.

Proof of Proposition 9.2. Note that the right-hand side of the bound

$$
\mathsf{P}_{s_L}\!\Big(\Gamma_{[0,r_j]}\in\mathrm{HighFirst}_{r_j,s_j-1}^{1/2+2\chi+\epsilon_2}\Big) \geq \mathsf{P}_{s_L}\!\Big(\Gamma_{[0,r_j]}\in\mathrm{HighFirst}_{r_j,s_j-1}^{1/2+2\chi+\epsilon_2}\,,\ \Gamma_{[0,r_j]}\in D_j\Big)
$$

is, by the similarity of measure Proposition 11.2, at least the product of the quantity $c\big(\log s_L\big)^{-1} s_L^{-4\varepsilon}$ and the probability

$$
\mathsf{P}_{s_j}\!\Big(\Gamma_{[0,r_j]}\in\mathrm{HighFirst}_{r_j,s_j-1}^{1/2+2\chi+\epsilon_2}\,,\ \Gamma\in\mathrm{RGJ}_{k,s_j-16-k}\text{ for some }k\in\mathsf{R}_{i+2}\,,\ \Gamma_{[0,r_j]}\in D_j\Big) ,
$$

where the constant c equals $\tfrac{1}{40}C_8^{-1}\big(\tfrac{3}{16}\big)^{\theta}$. Note further that

- $\mathsf{P}_{s_j}\!\Big(\Gamma_{[0,r_j]}\in\mathrm{HighFirst}_{r_j,s_j-1}^{1/2+2\chi+\epsilon_2}\Big)\geq 1-(s_j-1)^{-(\chi+\epsilon_2-\epsilon_1)}$, by (10.1) with $a=1$; and

- $\mathsf{P}_{s_j}\!\Big(\Gamma\notin\mathrm{RGJ}_{k,s_j-16-k}\text{ for every }k\in\mathsf{R}_{i+2}\Big)\leq 1-s_j^{-(\chi+\epsilon_2/2)}$, by Lemma 11.1 with $\varphi=\epsilon_2/2$.

Also recall (11.1). Thus,

$$
\mathsf{P}_{s_L}\!\Big(\Gamma_{[0,r_j]}\in\mathrm{HighFirst}_{r_j,s_j-1}^{1/2+2\chi+\epsilon_2}\Big)
$$

$$
\geq\quad c\big(\log s_L\big)^{-1} s_L^{-4\varepsilon}\left(s_j^{-(\chi+\epsilon_2/2)} - (s_j-1)^{-(\chi+\epsilon_2-\epsilon_1)} - 4\mu^{16} s_j^{-10}\right) .
$$

Recall that $s_j\geq 2^{i+4}$. Provided that the index $i\in\mathbb{N}$ is supposed to be sufficiently high, the displayed right-hand side is thus at least $\tfrac{1}{2}c\big(\log s_L\big)^{-1} s_L^{-\chi-\epsilon_2/2-4\varepsilon}$ if we insist that $\epsilon_2>2\epsilon_1$ as well as $\chi+\epsilon_2<10$ (the latter following from $\chi<1/6$ and the harmlessly assumed $\epsilon_2<1$).

We are now ready to obtain Proposition 9.2. We must set the values of the quantities m and m' in the proposition's statement. We take $m=s_L$ and m' equal to the constant difference s_j-r_j between the terms in any pair in the sequence constructed in Definition 10.6. Recalling the bound on this difference stated in the definition, we see that $2^{i+3}\leq m'\leq 2^{i+4}\leq m\leq 2^{i+5}$. Set $\epsilon_2=4\epsilon_1$ and $\epsilon_1=\varepsilon$, so that $\chi+\epsilon_2/2+4\varepsilon=\chi+6\varepsilon<\eta-\varepsilon$ and $1/2+2\chi+\epsilon_2=\alpha$. Note that r_j belongs to the set K specified in Proposition 9.2 whenever $j\in[1,L]$, because $r_j=s_j-m'\leq m-m'$. Thus, $\#K\geq 2^{i-4}\geq 2^{-9}m$, and Proposition 9.2 is proved. $\square$

12 Step three: completing the proof of Proposition 8.1

As we explained in this step's outline, we set the snake method index parameters at the start of step three. For a given sufficiently high choice of $i \in \mathbb{N}$, Proposition 9.2 specifies m and m'. We then set the two index parameters, with n equal to $m-1$ and ℓ equal to $m - m'$.

For $\gamma \in \mathrm{SAP}_{n+1}$, set

$$X_\gamma = \sum_{k=0}^{\ell} \mathbb{1}_{\gamma_{[0,k]} \in \mathrm{HighFirst}_{k,m'+k-1}^{\alpha}} \, .$$

Recall that $\gamma_{[0,\ell]} \in \mathrm{First}_{\ell,n}$ is (α, n, ℓ)-charming at an index $k \in [1, \ell]$ if

$$\mathsf{W}_{k+n-\ell}^0 \Big(\Gamma \text{ closes } \Big| \, \big|\Gamma^1\big| = k \, , \, \Gamma^1 = \gamma_{[0,k]} \Big) > n^{-\alpha} \, ;$$

on the other hand, for given $k \in [1, \ell]$, the property that such a γ satisfies $\gamma_{[0,k]} \in \mathrm{HighFirst}_{k,n-\ell+k}^{\alpha}$ takes the same form with $n^{-\alpha}$ replaced by the slightly larger quantity $(n - \ell + k)^{-\alpha}$. Since $n - \ell = m' - 1$ and $\beta = 1$, we find that, for any $\gamma \in \mathrm{SAP}_{n+1}$,

$$X_\gamma \geq n^{\beta - \eta}/4 \text{ implies that } \gamma_{[0,\ell]} \in \mathrm{CS}_{\beta,\eta}^{\alpha,\ell,n} \, . \tag{12.1}$$

We now show that

$$\mathsf{P}_{n+1}\Big(X_\Gamma \geq n^{1-\eta+\varepsilon/2} \Big) \geq n^{-\eta+\varepsilon/2} \, . \tag{12.2}$$

To derive this, consider the expression

$$S = \sum_{\gamma \in \mathrm{SAP}_{n+1}} \sum_{k=0}^{\ell} \mathbb{1}_{\gamma_{[0,k]} \in \mathrm{HighFirst}_{k,m'+k-1}^{\alpha}} \, .$$

Recall that p_{n+1} denotes $\#\mathrm{SAP}_{n+1}$; using Proposition 9.2 in light of $n + 1 = m$,

$$\begin{aligned} S &= p_{n+1} \sum_{k=0}^{\ell} \mathsf{P}_{n+1}\Big(\Gamma_{[0,k]} \in \mathrm{HighFirst}_{k,m'+k-1}^{\alpha} \Big) \\ &\geq p_{n+1} \cdot \#K \cdot (n+1)^{-\eta+\varepsilon} \geq p_{n+1} \cdot 2^{-9} n \cdot 2^{-1} n^{-\eta+\varepsilon} \, . \end{aligned}$$

Let q denote the left-hand side of (12.2). Note that

$$S \leq p_{n+1} \cdot \Big(q\big(\ell+1\big) + (1-q)n^{1-\eta+\varepsilon/2} \Big) \, .$$

From the lower bound on S, and $n \geq \ell$,

$$q\big(n+1\big) + n^{1-\eta+\varepsilon/2} \geq 2^{-10}\, n^{1-\eta+\varepsilon} \, ,$$

which implies for sufficiently high n that $q \geq n^{-\eta+\varepsilon/2}$; in this way, we obtain (12.2).

Trivially, $n^{1-\eta+\varepsilon/2} \geq n^{1-\eta}/4$ (for all $n \in \mathbb{N}$, including our choice of n). Since the snake length exponent β is set to one, we learn from (12.1) and (12.2) that

$$\mathsf{P}_{n+1}\Big(\Gamma_{[0,\ell]} \in \mathrm{CS}_{\beta,\eta}^{\alpha,\ell,n} \Big) \geq \mathsf{P}_{n+1}\Big(X_\Gamma \geq n^{1-\eta+\varepsilon/2} \Big) \geq n^{-\eta+\varepsilon/2} \, .$$

Thus, if the dyadic scale parameter $i \in \mathbb{N}$ is chosen so that $n \geq 2^{i+4}$ is sufficiently high, the charming snake presence hypothesis (6.2) is satisfied. By Theorem 6.3, $\mathsf{W}_n\big(\Gamma \text{ closes}\big) \leq 2(n+1)c^{-n^\delta/2}$. This deduction has been made for some value of $n \in (2\mathbb{N}+1) \cap \big[2^{i+4}-1, 2^{i+5}-1\big]$, where here $i \in \mathbb{N}$ is arbitrary to the right of a finite interval. Relabelling $c > 1$ to be any value in $(1, c^{1/2})$ completes the proof of Proposition 8.1. $\qquad\square$

13 Preparing for the proof of Theorem 1.5(1)

This argument certainly fits the template offered by the proof of Theorem 1.5(2). We will present the proof by explaining how to modify the statements and arguments used in the preceding derivation.

Recall first of all that the hypotheses of Theorem 1.5(2) implied the existence of the closing exponent, and that we chose to label $\chi \in \mathbb{R}$ such that $\mathsf{P}_n\big(\Gamma \text{ closes}\big) = n^{-1/2-\chi+o(1)}$ for odd n; moreover, χ could be supposed to be non-negative. Specifically, it was the lower bound $\mathsf{P}_n\big(\Gamma \text{ closes}\big) \geq n^{-1/2-\chi-o(1)}$ that was invoked, in part one of the preceding proof.

In the present proof, the closing exponent is not hypothesised to exist, and so we must abandon this usage of χ. However, this parameter will be used again, with the basic role of $1/2 + \chi$ as the exponent in such a closing probability lower bound being maintained.

Definition 13.1. *For $\zeta > 0$, define the set of indices* $\mathsf{HCP}_\zeta \subseteq 2\mathbb{N}$ *of ζ-high closing probability,*

$$\mathsf{HCP}_\zeta = \left\{ n \in 2\mathbb{N} : \mathsf{W}_{n-1}\big(\Gamma \text{ closes}\big) \geq n^{-\zeta} \right\}.$$

For $\chi > 0$ arbitrary, we introduce the *closing probability*

Hypothesis CP_χ**.** The set $2\mathbb{N} \setminus \mathsf{HCP}_{1/2+\chi}$ has limit supremum density in $2\mathbb{N}$ less than $1/1250$.

We begin the proof by the same type of reduction as was used in the earlier derivation. Here is the counterpart to Proposition 8.1.

Proposition 13.2. *Let $d = 2$. Let $\chi \in (0, 1/14)$, and assume Hypothesis CP_χ. For some $c > 1$ and $\delta > 0$, the set of odd integers n satisfying $n + 1 \in \mathsf{HCP}_{1/2+\chi}$ and*

$$\mathsf{W}_n\big(\Gamma \text{ closes}\big) \leq c^{-n^\delta}$$

intersects the dyadic scale $\big[2^i, 2^{i+1}\big]$ for all but finitely many $i \in \mathbb{N}$.

Proof of Theorem 1.5(1). The two properties of the index n asserted by the proposition's conclusion are evidently in contradiction for n sufficiently high. Thus Hypothesis CP_χ is false whenever $\chi \in (0, 1/14)$. This implies the result. $\qquad\square$

The rest of the article is dedicated to proving Proposition 13.2. The three step plan of attack for deriving Proposition 8.1 will also be adopted for the new proof. Of course some changes are needed. We must cope with a deterioration in the known regularity of the θ-sequence. We begin by overviewing what we know in the present case and how the new information will cause changes in the three step plan.

In the proof of Theorem 1.5(2) via Proposition 8.1, we had the luxury of assuming the existence of the closing exponent. Throughout the proof of Proposition 13.2 (and thus henceforth), we instead *fix $\chi \in (0, 1/14)$, and suppose that Hypothesis CP_χ holds*. Thus, our given information is that the closing probability is at least $n^{-1/2-\chi}$ for a uniformly positive proportion of indices n in any sufficiently long initial interval of positive integers.

We may state and prove right away the θ-sequence regularity offered by this information.

Proposition 13.3. *Let $\varepsilon \in (0, 1)$. For all but finitely many values of $i \in \mathbb{N}$,*

$$\#\left\{ j \in 2\mathbb{N} \cap \big[2^i, 2^{i+1}\big] : 3/2 - \varepsilon \leq \theta_j \leq 3/2 + \chi + \varepsilon \right\} \geq 2^{i-1}\big(1 - \tfrac{1}{300}\big).$$

Proof. Theorem 1.3 implies that, for all but finitely many values of $i \in \mathbb{N}$,

$$\#\left\{ j \in 2\mathbb{N} \cap \big[2^i, 2^{i+1}\big] : \theta_j \leq 3/2 - \varepsilon \right\} \leq \tfrac{1}{600} \, 2^{i-1}. \tag{13.1}$$

Thus, it is enough to show that, also for all but finitely many i,

$$\#\left\{j \in 2\mathbb{N} \cap \left[2^i, 2^{i+1}\right] : \theta_j > 3/2 + \chi + \varepsilon\right\} \le \tfrac{1}{600}\, 2^{i-1}\,. \tag{13.2}$$

We will invoke Hypothesis CP_χ as we verify (13.2). First, we make a

Claim. For any $\varphi > 0$, the set $\mathsf{HCP}_{1/2+\chi} \cap \left\{n \in 2\mathbb{N} : \theta_n > 3/2 + \chi + \varphi\right\}$ is finite.

To see this, note that $2np_n/c_{n-1} \ge n^{-1/2-\chi}$ if $n \in \mathsf{HCP}_{1/2+\chi}$. From $c_{n-1} \ge \mu^{n-1}$ and $p_n = \mu^n n^{-\theta_n}$ follows $n^{-\theta_n} \ge \frac{1}{2\mu} n^{-3/2-\chi}$ for such n, and thus the claim.

The claim implies that, under Hypothesis CP_χ, the set $\left\{n \in 2\mathbb{N} : \theta_n > 3/2 + \chi + \varepsilon\right\}$ has limit supremum density at most $1/1250$. From this, (13.2) follows directly. □

That is, the weaker regularity information forces a positive proportion of the polygon number deficit exponents θ_n to lie in any given open set containing the interval $[3/2, 3/2 + \chi]$. In the proof of Theorem 1.5(1), we knew this for the one-point set $\{3/2 + \chi\}$, for all high n.

In the three step plan, we will adjust step one so that the constant difference sequence constructed there incorporates the θ-regularity that Proposition 13.3 shows to be typical, as well as a closing probability lower bound permitted by Hypothesis CP_χ.

In step two, the weaker regularity will lead to a counterpart Lemma 14.6 to the regulation polygon ample supply Lemma 11.1. Where before we found a lower bound on the probability that a polygon is regulation global join of the form $n^{-\chi-o(1)}$, now we will find only a bound $n^{-2\chi-o(1)}$.

Continuing this step, the mechanism of measure comparison for initial subpaths of polygons drawn from the laws P_n for differing lengths n is no longer made via all regulation global join polygons but rather via such polygons whose length index lies in a certain regular set. The new similarity of measure result counterpart to Proposition 11.2 will be Proposition 14.7. It will be applied at the end of step two to prove a slight variant of Proposition 9.2, namely Proposition 14.4.

This last result will yield Proposition 13.2 in step three by a verbatim argument to that by which Proposition 8.1 followed from Proposition 9.2.

14 The three steps for Proposition 13.2's proof in detail

In a first subsection of this section, we reset the snake method's parameters to handle the weaker information available. In the following three subsections, we state and prove the assertions associated to each of the three steps.

14.1 The snake method exponent parameters

Since $\chi < 1/14$, we may fix a parameter $\varepsilon \in \left(0, (1/2 - 7\chi)/14\right)$, and do so henceforth. The three exponent parameters are then set so that

- $\beta = 1$;
- $\alpha = 1/2 + 3\chi + 5\varepsilon$;
- and $\eta = 4\chi + 9\varepsilon$.

Note that the quantity $\delta = \beta - \eta - \alpha$, which must be positive if the method to work, is equal to $1/2 - 7\chi - 14\varepsilon$. The constraint imposed on ε ensures this positivity.

14.2 Step one

The essential conclusion of the first step one was the construction of the constant difference sequence in Definition 10.6. Now our aim is similar.

Fixing as previously two parameters $\epsilon_2 > \epsilon_1 > 0$ (which in the present case we will ultimately set $\epsilon_2 = 5\epsilon_1$ and $\epsilon_1 = \varepsilon$ in terms of our fixed parameter ε), we adopt the Definition 10.1 of a closing probability typical index pair on dyadic scale i, where of course the parameter χ is fixed in the way we have explained. An index pair (k, j) that satisfies this definition is called *regularity typical* for dyadic scale i if moreover $j \in \mathsf{HCP}_{1/2+\chi}$ and

$$3/2 - \varepsilon \le \theta_j \le 3/2 + \chi + \varepsilon. \tag{14.1}$$

We then replace Definition 10.6 with the following.

Definition 14.1. *Set $L = 2^{i-4}$. We specify a* constant difference *sequence of index pairs (r_j, s_j), $1 \le j \le L$, whose elements are regularity typical for dyadic scale i, with $s_j - r_j$ independent of j and valued in $[2^{i+4} - 2^i, 2^{i+4}]$, and $\{r_j : 1 \le j \le L\}$ increasing. Moreover, $s_j \in 2\mathbb{N} \cap [2^{i+4}, 2^{i+4} + 2^i + 2^{i-2}]$ for each $j \in [1, L]$.*

Our job in step one is to construct such a sequence. First, recall the set E of index pairs (k, j) specified in Definition 10.3. Let E' be the subset of E consisting of such pairs for which $j \in \mathsf{HCP}_{1/2+\chi}$ and (14.1) holds. Counterpart to Lemma 10.4, we have:

Lemma 14.2. *Provided that the index $i \in \mathbb{N}$ is sufficiently high, except for at most $\frac{1}{100}2^{i+3}$ values of $j \in 2\mathbb{N} \cap [2^{i+4}, 2^{i+5}]$, the set of $k \in [2^i, 2^i + 2^{i-2}]$ such that $(k, j) \notin E'$ has cardinality at most $2(2^{i+5})^{1-\epsilon_1}$.*

Proof. Proposition 13.3 implies that the set of $j \in 2\mathbb{N} \cap [2^{i+4}, 2^{i+5}]$ satisfying (14.1) – a set containing all second coordinates of pairs in E' – has cardinality at least $(1 - \frac{1}{300})2^{i+3}$. Invoking Hypothesis CP_χ, we see that, of these values of j, at most $(2^{i+4} + 1)\frac{1}{1250} \le \frac{1}{600}2^{i+3}$ fail the test of membership of $\mathsf{HCP}_{1/2+\chi}$. We claim that the set of remaining values of j satisfies the statement in the lemma. Indeed, all these values belong to $\mathsf{HCP}_{1/2+\chi}$. As such, we are able to invoke Lemma 6.1, as we did in proving Lemma 10.4, but this time with $\alpha' = 1/2 + \chi$ and $\delta' = \epsilon_1$, to find that at most $2(2^{i+5})^{1-\epsilon_1}$ values of $k \in [2^i, 2^i + 2^{i-2}]$ violate the condition (10.2). $\qquad\square$

Analogously to Proposition 10.2,

Proposition 14.3. *There exists $k \in 2\mathbb{N} \cap [2^{i+4}, 2^{i+4} + 2^i]$ for which there are at least 2^{i-4} values of $j \in 2\mathbb{N} \cap [0, 2^{i-2}]$ with the index pair $(2^i + j, k + j)$ regularity typical on dyadic scale i.*

Proof. Using Lemma 14.2, we find, analogously to (10.5),

$$\sum_{j \in 2\mathbb{N} \cap [2^{i+4}+2^{i-2}, 2^{i+4}+2^i]} \sum_{\ell=2^i}^{2^i+2^{i-2}} \mathbb{1}_{(\ell, j) \in E'}$$
$$\ge \left(2^{i-2} + 1 - 2(2^{i+5})^{1-\epsilon_1}\right) \cdot \left(\tfrac{1}{2}(2^i - 2^{i-2}) - \tfrac{1}{100}2^{i+3}\right).$$

The argument for Proposition 10.2 now yields the result. $\qquad\square$

We may now conclude step one, because the proof of Proposition 9.1 is valid with the notion of regularity typical replacing closing probability typical; thus, we have specified a constant difference sequence in the sense of Definition 14.1. (Incidentally, the unused Proposition 9.1 is seen to be valid with right-hand side strengthened to $1 - 2^{-2i\chi}$ if we choose $\epsilon_2 = 5\varepsilon$ and $\epsilon_2 > \epsilon_1$ by considering $a = 2$ in (10.1).)

14.3 Step two

The new version of Proposition 9.2 differs from the original only in asserting that the constructed m belongs to $\mathsf{HCP}_{1/2+\chi}$.

Proposition 14.4. *For each $i \in \mathbb{N}$ sufficiently high, there exist $m \in 2\mathbb{N} \cap \left[2^{i+4}, 2^{i+5}\right] \cap$ $\mathrm{HCP}_{1/2+\chi}$ and $m' \in \left[2^{i+3}, 2^{i+4}\right]$ such that, writing K for the set of values $k \in \mathbb{N}$, $1 \leq k \leq$ $m - m'$, that satisfy*

$$\mathsf{P}_m\left(\Gamma_{[0,k]} \in \mathrm{HighFirst}_{k,k+m'-1}^{\alpha}\right) \geq m^{-\eta+\varepsilon} \,,$$

we have that $|K| \geq 2^{-9}m$.

In this step, we provide a suitable similarity of measure result, Proposition 14.7, and use it to derive Proposition 14.4.

Before doing so, we must adapt our regulation polygon tools to cope with the weaker regularity that we are hypothesising. The ample regulation polygon supply Lemma 11.1 will be replaced by Lemma 14.6.

In preparing to state the new lemma, recall that Lemma 11.1 was stated in terms of a regular index set R_{i+2} that was specified in a trivial way. Recalling that $L = 2^{i-4}$, we now specify a non-trivial counterpart set $\mathsf{R}_{n,i+2}^{\chi+\varepsilon}$ when n is an element in the sequence $\left\{s_j : 1 \leq j \leq L\right\}$ from step one.

Definition 14.5. *For $j \in \left[1, L\right]$, define $\mathsf{R}_{s_j,i+2}^{\chi+\varepsilon}$ to be the set of $k \in 2\mathbb{N} \cap \left[2^{i+2}, 2^{i+3}\right]$ such that*

$$\max\left\{\theta_k, \theta_{s_j-16-k}, \theta_{s_L-16-k}\right\} \leq 3/2 + \chi + \varepsilon \ \text{and} \ \theta_{s_j-16-k} \geq 3/2 - \varepsilon \,.$$

Here is the new ample supply lemma.

Lemma 14.6. *Let $j \in [1, L]$. Then*

$$\mathsf{P}_{s_j}\left(\Gamma \in \mathrm{RGJ}_{k,s_j-16-k} \ \text{for some } k \in \mathsf{R}_{s_j,i+2}^{\chi+\varepsilon}\right) \geq \frac{c}{\log s_j}\, s_j^{-(2\chi+3\varepsilon)} \,,$$

where $c > 0$ is a universal constant.

The similarity of measure Proposition 11.2 is replaced by the next result.

Proposition 14.7. *For each $j \in [1, L]$, there is a subset $D_j \subseteq \mathrm{First}_{r_j}$ satisfying*

$$\mathsf{P}_{s_j}\left(\Gamma_{[0,r_j]} \notin D_j \,, \Gamma \in \mathrm{RGJ}_{k,s_j-16-k} \ \text{for some } k \in \mathsf{R}_{s_j,i+2}^{\chi+\varepsilon}\right) \leq 4\mu^{16} s_j^{-10} \tag{14.2}$$

such that, for each $\phi \in D_j$,

$$\begin{aligned}
\mathsf{P}_{s_L}\left(\Gamma_{[0,r_j]} = \phi\right) \ &\geq \ \tfrac{1}{40}\, C_8^{-1}\left(\tfrac{3}{16}\right)^{3/2}\left(\log s_L\right)^{-1} s_L^{-2\chi-4\varepsilon} \\
&\quad \times \ \mathsf{P}_{s_j}\left(\Gamma_{[0,r_j]} = \phi \,, \Gamma \in \mathrm{RGJ}_{k,s_j-16-k} \ \text{for some } k \in \mathsf{R}_{s_j,i+2}^{\chi+\varepsilon}\right).
\end{aligned}$$

In the rest of this subsection of step two proofs, we prove in turn Lemma 14.6 and Propositions 14.7 and 14.4.

14.3.1 Deriving Lemma 14.6

We will use the next result.

Lemma 14.8. *For $i \in \mathbb{N}$ with $i \geq 6$, and $n \in 2\mathbb{N} \cap \left[2^{i+4}, 2^{i+5}\right]$,*

$$\left|\mathsf{R}_{s_j,i+2}^{\chi+\varepsilon}\right| \geq \left(\tfrac{9}{10} - \tfrac{1}{32}\right)2^{i+1} \,.$$

Proof. It is enough to prove two claims.

For $\varphi > 0$, set $\mathsf{K}_{s_j,i+2}^{\varphi} = \left\{k \in 2\mathbb{N} \cap \left[2^{i+2}, 2^{i+3}\right] : \max\left\{\theta_k, \theta_{s_j-16-k}\right\} \leq 3/2 + \varphi\right\}$.
Claim 1. The set $\mathsf{K}_{s_j,i+2}^{\chi+\varepsilon} \setminus \mathsf{R}_{s_j,i+2}^{\chi+\varepsilon}$ has cardinality at most 2^{i-4}.
Claim 2. $\left|\mathsf{K}_{s_j,i+2}^{\chi+\varepsilon}\right| \geq \tfrac{9}{10}\, 2^{i+1}$.

Proof of Claim 1. Note that $k \in \mathsf{K}^{\chi+\varepsilon}_{s_j,i+2} \setminus \mathsf{R}^{\chi+\varepsilon}_{s_j,i+2}$ implies that either $s_j - 16 - k$ belongs to the union up to index $i + 4$ (in the role of i) of the sets in (13.1), or $s_L - 16 - k$ belongs to the comparable union of the sets in (13.2). (The reason that the index here is $i + 4$ is because s_j and s_L are at most $2^{i+4} + 2^i + 2^{i-2} \leq 2^{i+5}$ in view of Definition 14.1.) Thus, the cardinality in question is at most $\frac{4}{600} 2^{i+3}$, where the factor of four accomodates the two alternatives as well as the sum over indices of dyadic scales up to index $i + 4$. We obtain Claim 1.

Proof of Claim 2. For $\varphi > 0$, define the set of φ-high θ values on dyadic scale i,

$$\mathsf{High\Theta}^{\varphi}_i = \left\{ k \in 2\mathbb{N} \cap \left[2^i, 2^{i+1}\right] : \theta_k \geq 3/2 + \varphi \right\},$$

and also set $\mathsf{S}^{\varphi}_{s_j,i+2} = 2\mathbb{N} \cap \left[2^{i+2}, 2^{i+3}\right] \setminus \mathsf{K}^{\varphi}_{s_j,i+2}$. We have then that, for any $\varphi > 0$,

$$\mathsf{S}^{\varphi}_{s_j,i+2} \subseteq \mathsf{High\Theta}^{\varphi}_{i+2} \cup \left\{ k \in 2\mathbb{N} \cap \left[2^{i+2}, 2^{i+3}\right] : \theta_{s_j-16-k} \geq 3/2 + \varphi \right\};$$

the latter set in the right-hand union is a subset of

$$2\mathbb{N} \cap \left[2^{i+3} - 2^4, 2^{i+3}\right] \cup \left\{ k \in 2\mathbb{N} \cap \left[2^{i+2}, 2^{i+3} - 2^4\right] : s_j - 16 - k \in \mathsf{High\Theta}^{\varphi}_{i+3} \cup \mathsf{High\Theta}^{\varphi}_{i+4} \right\},$$

because $2^{i+3} \leq s_j - 16 - k \leq 2^{i+5}$ for $k \in \left[2^{i+2}, 2^{i+3} - 2^4\right]$. Thus,

$$\#\mathsf{S}^{\chi+\varepsilon}_{s_j,i+2} \leq \#\mathsf{High\Theta}^{\chi+\varepsilon}_{i+2} + \#\mathsf{High\Theta}^{\chi+\varepsilon}_{i+3} + \#\mathsf{High\Theta}^{\chi+\varepsilon}_{i+4} + 9 \,,$$

so that (13.2) implies $\left| \mathsf{S}^{\chi+\varepsilon}_{s_j,i+2} \right| \leq \frac{7}{600} 2^{i+1} + 9 \leq \frac{1}{10} 2^{i+1}$ (for $i \geq 6$). Thus, we obtain Claim 2. $\qquad \square$

Proof of Lemma 14.6. The proof is similar to that of Lemma 11.1. Now, the probability in question equals the ratio of $\left| \bigcup_{l \in \mathsf{R}^{\chi+\varepsilon}_{s_j,i+2}} \mathrm{RGJ}_{l,s_j-16-l} \right|$ and p_{s_j}.

To bound the numerator below, we again seek to apply Proposition 5.10 with the role of $i \in \mathbb{N}$ played by $i + 2$, but now with $n = s_j$ and $\mathsf{R} = \mathsf{R}^{\chi+\varepsilon}_{s_j,i+2}$. Set Θ in the proposition to be $3/2 + \chi + \varepsilon$. Any element of $\mathsf{R}^{\chi+\varepsilon}_{s_j,i+2}$ may act as $k^* \in \mathsf{R}$ in the way that the proposition demands, by the definition of $\mathsf{R}^{\chi+\varepsilon}_{s_j,i+2}$; (and Lemma 14.8 shows that $\mathsf{R}^{\chi+\varepsilon}_{s_j,i+2}$ is non-empty if $i \geq 6$). Applying the proposition and then using Lemma 14.8, $2^{i+1} \geq 2^{-4} s_j$, and the definition of $\mathsf{R}^{\chi+\varepsilon}_{s_j,i+2}$, we find that

$$\left| \bigcup_{l \in \mathsf{R}^{\chi+\varepsilon}_{s_j,i+2}} \mathrm{RGJ}_{l,s_j-16-l} \right| \geq c \frac{s_j^{1/2}}{\log s_j} \cdot \left(\tfrac{9}{10} - \tfrac{1}{32} \right) 2^{-4} s_j \cdot \mu^{s_j - 16} s_j^{-(3+2\chi+2\varepsilon)} \,,$$

which is at least a small constant multiple of $\mu^{s_j} \left(\log s_j \right)^{-1} s_j^{-3/2-2\chi-2\varepsilon}$.

The ratio's denominator is bounded above with the bound $p_{s_j} \leq \mu^{s_j} s_j^{-3/2+\varepsilon}$ which is due to the lower bound in (14.1) with s_j playing the role of j. The lemma follows by relabelling $c > 0$. $\qquad \square$

14.3.2 Deriving Proposition 14.7.

We begin by respecifying the set τ_γ. For each $\gamma \in \mathrm{SAP}_{s_j}$, $1 \leq j \leq L$, we now set

$$\tau_\gamma = \left\{ k \in \mathsf{R}^{\chi+\varepsilon}_{s_j,i+2} : \gamma \in \mathrm{RGJ}_{k,s_j-16-k} \right\}.$$

Note that $\mathsf{R}^{\chi+\varepsilon}_{s_j,i+2}$ has replaced R_{i+2} here. As we comment on the needed changes, it is understood that this replacement is always made.

We reinterpret the quantity $\overline{\theta}$ to be $3/2 + \chi + \varepsilon$. The constant $C_8 > 0$ may again be chosen in order that (11.2) be satisfied. Indeed, since (r_L, s_L) is by construction regularity typical on dyadic scale i, we have that s_L is an element of $\mathsf{HCP}_{1/2+\chi}$ and thus also of $\mathsf{HPN}_{3/2+\chi+o(1)}$. Corollary 3.3 is thus applicable with $n = s_L$, so that (11.2) results.

The set D_j for $j \in [1, L]$ is specified as before, and Lemma 11.3(1) is also obtained, by the verbatim argument.

The definition of $\overline{\theta}$ has been altered in order to maintain the validity of Lemma 11.3(3). The moment in the proof of this result where the new definition is needed is in deriving the bound

$$p_{s_L}^{-1} p_k p_{s_L - 16 - k} \geq \mu^{-16} s_L^{-2\overline{\theta}} \tag{14.3}$$

for $k \in \mathsf{R}_{s_j, i+2}^{\chi+\varepsilon}$ in the proof of Lemma 11.3(2) (which is needed to prove the third part). That $k \in \mathsf{R}_{s_j, i+2}^{\chi+\varepsilon}$ ensures that both θ_k and $\theta_{s_L - 16 - k}$ are at most $3/2 + \chi + \varepsilon = \overline{\theta}$, from which follows (14.3), and thus Lemma 11.3(2) and (3).

The weaker form of the inference that we may make in view of the weaker available regularity information is evident in the counterpart to Lemma 11.4. Note the presence of the new factors of $s_L^{-\chi}$ and s_j^{χ} in the two right-hand sides.

Lemma 14.9. *Recall that $L = 2^{i-4}$. Let $j \in [1, L]$.*

1. *For $\phi \in D_j$,*

$$\mathsf{P}_{s_L}\Big(\Gamma_{[0, r_j]} = \phi\Big) \geq \frac{1}{400} C_8^{-1} \big(\log s_L\big)^{-1} s_L^{-\chi - 2\varepsilon} \mu^{-16}$$
$$\times \sum_{k \in \mathsf{R}_{s_j, i+2}^{\chi+\varepsilon}} \mathsf{P}_k^{\mathrm{left}}\Big(\Gamma_{[0, r_j]} = \phi\Big) k^{1/2} p_k \mu^{-k} .$$

2. *For $\phi \in \mathrm{First}_{r_j}$,*

$$\mathsf{P}_{s_j}\Big(\Gamma_{[0, r_j]} = \phi, \, \Gamma \in \mathrm{RGJ}_{k, s_j - 16 - k} \text{ for some } k \in \mathsf{R}_{s_j, i+2}^{\chi+\varepsilon}\Big)$$
$$\leq \frac{1}{10}\Big(\frac{16}{3}\Big)^{3/2} \mu^{-16} \cdot s_j^{\chi+2\varepsilon} \sum_{k \in \mathsf{R}_{s_j, i+2}^{\chi+\varepsilon}} \mathsf{P}_k^{\mathrm{left}}\Big(\Gamma_{[0, r_j]} = \phi\Big) k^{1/2} p_k \mu^{-k} .$$

Proof (1). The earlier proof runs its course undisturbed until $\mathsf{P}_{s_L}\big(\Gamma_{[0, r_j]} = \phi\big)$ is found to be at least the expression in the double line ending at (11.7). The proof of this part is then completed by noting the bound

$$p_{s_L}^{-1} p_{s_L - 16 - k} \geq \mu^{-16-k} s_L^{-\chi - 2\varepsilon}$$

for $k \in \mathsf{R}_{s_j, i+2}^{\chi+\varepsilon}$. This bound is due to $\theta_{s_L - 16 - k} \leq 3/2 + \chi + \varepsilon$ and $\theta_{s_L} \geq 3/2 - \varepsilon$, the latter being the lower bound in (14.1) with $j = s_L$.

(2). The left-hand side of the claimed inequality is bounded above by (11.8). The proof is then completed by invoking the bound

$$p_{s_j}^{-1} p_{s_j - 16 - k} \leq \mu^{-16-k} \Big(\frac{16}{3}\Big)^{3/2} s_j^{\chi+2\varepsilon} \tag{14.4}$$

for $k \in \mathsf{R}_{s_j, i+2}^{\chi+\varepsilon}$. To derive this bound, note that

$$p_{s_j - 16 - k} \leq \mu^{s_j - 16 - k} \big(s_j - 16 - k\big)^{-3/2+\varepsilon} \leq \mu^{s_j - 16 - k} \Big(\frac{16}{3}\Big)^{3/2} s_j^{-3/2+\varepsilon},$$

where the first inequality follows from $\theta_{s_j - 16 - k} \geq 3/2 - \varepsilon$, which is due to $k \in \mathsf{R}_{s_j, i+2}^{\chi+\varepsilon}$; and where the second depends on $s_j - 16 - k \geq 3s_j/16$, which we saw at the corresponding moment in the derivation that is being adapted, as well as on $0 \leq \varepsilon \leq 3/2$.

We also have
$$p_{s_j} \geq \mu^{s_j} s_j^{-3/2-\chi-\varepsilon}$$

due to the upper bound in (14.1) with the role of j played by s_j.

The last two displayed inequalities combine to yield (14.4). □

Proof of Proposition 14.7. A consequence of Lemmas 11.3(3) and 14.9. □

14.3.3 Proof of Proposition 14.4.

Note that the right-hand side of the bound
$$\mathsf{P}_{s_L}\left(\Gamma_{[0,r_j]} \in \text{HighFirst}_{r_j,s_j-1}^{1/2+3\chi+\epsilon_2}\right) \geq \mathsf{P}_{s_L}\left(\Gamma_{[0,r_j]} \in \text{HighFirst}_{r_j,s_j-1}^{1/2+3\chi+\epsilon_2}, \Gamma_{[0,r_j]} \in D_j\right)$$

is, by Proposition 14.7, at least the product of $c_1\left(\log s_L\right)^{-1} s_L^{-2\chi-4\varepsilon}$ and
$$\mathsf{P}_{s_j}\left(\Gamma_{[0,r_j]} \in \text{HighFirst}_{r_j,s_j-1}^{1/2+3\chi+\epsilon_2}, \Gamma \in \text{RGJ}_{k,s_j-16-k} \text{ for some } k \in \mathsf{R}_{s_j,i+2}^{\chi+\varepsilon}, \Gamma_{[0,r_j]} \in D_j\right)$$
(14.5)
where $c_1 = \frac{1}{40} C_8^{-1}\left(\frac{3}{16}\right)^{3/2}$.

Recall that the index pair (r_j, s_j) is regularity, and thus closing probability, typical on dyadic scale i. We may thus apply (10.1) with s_j playing the role of j and with $a = 2$ to find that

- $\mathsf{P}_{s_j}\left(\Gamma_{[0,r_j]} \in \text{HighFirst}_{r_j,s_j-1}^{1/2+3\chi+\epsilon_2}\right) \geq 1 - (s_j-1)^{-(2\chi+\epsilon_2-\epsilon_1)}$.

Also recall

- that, by Lemma 14.6, the quantity $\mathsf{P}_{s_j}\left(\Gamma \notin \text{RGJ}_{k,s_j-16-k} \text{ for every } k \in \mathsf{R}_{s_j,i+2}^{\chi+\varepsilon}\right)$ is at most $1 - c\left(\log s_j\right)^{-1} s_j^{-(2\chi+3\varepsilon)}$; and
- the inequality (14.2).

The three inequalities stated or recalled in these bullet points may be used to find a lower bound on the expression (14.5). We see then that
$$\begin{aligned}
&\mathsf{P}_{s_L}\left(\Gamma_{[0,r_j]} \in \text{HighFirst}_{r_j,s_j-1}^{1/2+3\chi+\epsilon_2}\right) \\
\geq \quad & c_1\left(\log s_L\right)^{-1} s_L^{-2\chi-4\varepsilon}\left(\frac{c}{\log s_j} s_j^{-(2\chi+3\varepsilon)} - (s_j-1)^{-(2\chi+\epsilon_2-\epsilon_1)} - 4\mu^{16} s_j^{-10}\right).
\end{aligned}$$

Provided that the index $i \in \mathbb{N}$ is supposed to be sufficiently high, the right-hand side is at least $\frac{1}{2} c_1 c\left(\log s_L\right)^{-2} s_L^{-4\chi-7\varepsilon} \geq s_L^{-4\chi-8\varepsilon} = s_L^{-\eta+\varepsilon}$ if we insist that $\epsilon_2 - \epsilon_1 > 3\varepsilon$ as well as $2\chi + \epsilon_2 - \epsilon_1 < 10$. We may set $\epsilon_2 = 5\epsilon_1$ and $\epsilon_1 = \varepsilon$, because we imposed at the outset that $\chi \in (0, 1/14)$ and in Section 14.1 the condition $\varepsilon \in \left(0, (1/2 - 7\chi)/14\right)$, and these choices do ensure the stated bounds. Note that $\alpha = 1/2 + 3\chi + \epsilon_2$ results from this choice.

We may now obtain Proposition 14.4. We set the values of m and m' as we did in proving Proposition 9.2, taking $m = s_L$ and m' equal to the constant difference $s_j - r_j$ associated to any element in the sequence (r_j, s_j) from Definition 14.1 (in place of course of Definition 10.6). The bound $\#K \geq 2^{i-4} \geq 2^{-9}m$ holds as in the earlier proof. Note however that we are also claiming that $m \in \text{HCP}_{1/2+\chi}$. This holds because $m = s_L$ is the second coordinate of a regularity typical index pair for dyadic scale i. Proposition 14.4 is proved. □

14.4 Step three

At the start of the earlier step three, the index parameters n and ℓ were respectively set equal to $m - 1$ and $m - m'$, where Proposition 9.2 provided m and m'. We now make the same choice, with Proposition 14.4 providing the latter two quantities.

Proof of Proposition 13.2. The completion of the proof of Proposition 8.1 in Section 12 applies verbatim after recalling that the snake method's parameter $\delta = \beta - \eta - \alpha$ is positive, and noting that the method's index parameter n satisfies $n + 1 \in \mathsf{HCP}_{1/2+\chi}$ because it has been set equal to $m - 1$ where m is specified in Proposition 14.4. $\qquad\square$

References

[1] R. Bauerschmidt, H. Duminil-Copin, J. Goodman, and G. Slade. Lectures on self-avoiding walks. In D. Ellwood, C. Newman, V. Sidoravicius, and W. Werner, editors, *Lecture notes, in Probability and Statistical Physics in Two and More Dimensions*. CMI/AMS – Clay Mathematics Institute Proceedings, 2011.

[2] Mireille Bousquet-Mélou and Richard Brak. Exactly solved models. In *Polygons, polyominoes and polycubes*, volume 775 of *Lecture Notes in Phys.*, pages 43–78. Springer, Dordrecht, 2009.

[3] Hugo Duminil-Copin, Alexander Glazman, Alan Hammond, and Ioan Manolescu. On the probability that self-avoiding walk ends at a given point. *Ann. Probab.*, 44(2):955–983, 2016.

[4] B. Duplantier. Fractals in two dimensions and conformal invariance. *Phys. D*, 38(1-3):71–87, 1989. Fractals in physics (Vence, 1989).

[5] B. Duplantier. Renormalization and conformal invariance for polymers. In *Fundamental problems in statistical mechanics VII (Altenberg, 1989)*, pages 171–223. North-Holland, Amsterdam, 1990.

[6] P. Flory. *Principles of Polymer Chemistry*. Cornell University Press, 1953.

[7] Alan Hammond. On self-avoiding polygons and walks: counting, joining and closing. arXiv:1504.05286, 2017.

[8] Alan Hammond. On self-avoiding polygons and walks: the snake method via pattern fluctuation. *Trans. Amer. Math. Soc., to appear*, 2017. https://math.berkeley.edu/~alanmh/papers/snakemethodpattern.pdf.

[9] Alan Hammond. An upper bound on the number of self-avoiding polygons via joining. *Ann. Probab.*, 46(1):175–206, 2018.

[10] T. Hara and G. Slade. Self-avoiding walk in five or more dimensions. I. The critical behaviour. *Comm. Math. Phys.*, 147(1):101–136, 1992.

[11] G. Lawler. Random walk problems motivated by statistical physics. http://www.math.uchicago.edu/~lawler/russia.pdf, 2013.

[12] G. F. Lawler, O. Schramm, and W. Werner. On the scaling limit of planar self-avoiding walk. In *Fractal geometry and applications: a jubilee of Benoît Mandelbrot, Part 2*, volume 72 of *Proc. Sympos. Pure Math.*, pages 339–364. Amer. Math. Soc., Providence, RI, 2004.

[13] N. Madras. A rigorous bound on the critical exponent for the number of lattice trees, animals, and polygons. *Journal of Statistical Physics*, 78(3-4):681–699, 1995.

[14] N. Madras and G. Slade. *The self-avoiding walk*. Probability and its Applications. Birkhäuser Boston Inc., Boston, MA, 1993.

[15] B. Nienhuis. Exact critical point and critical exponents of $O(n)$ models in two dimensions. *Phys. Rev. Lett.*, 49:1062–1065, 1982.

[16] B. Nienhuis. Coulomb gas description of 2D critical behaviour. *J. Statist. Phys.*, 34:731–761, 1984.

[17] W.J.C. Orr. Statistical treatment of polymer solutions at infinite dilution. *Transactions of the Faraday Society*, 43:12–27, 1947.

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm