BOUNDING THE NUMBER OF SELF-AVOIDING WALKS: HAMMERSLEY-WELSH WITH POLYGON INSERTION

HUGO DUMINIL-COPIN, SHIRSHENDU GANGULY, ALAN HAMMOND, AND IOAN MANOLESCU

ABSTRACT. Let $c_n = c_n(d)$ denote the number of self-avoiding walks of length n starting at the origin in the Euclidean nearest-neighbour lattice $\mathbb{Z}^d$. Let $\mu = \lim_n c_n^{1/n}$ denote the connective constant of $\mathbb{Z}^d$. In 1962, Hammersley and Welsh [HW62] proved that, for each $d \geq 2$, there exists a constant $C > 0$ such that $c_n \leq \exp(Cn^{1/2})\mu^n$ for all $n \in \mathbb{N}$. While it is anticipated that $c_n\mu^{-n}$ has a power-law growth in n, the best known upper bound in dimension two has remained of the form $n^{1/2}$ inside the exponential.

We consider two planar lattices and prove that $c_n \leq \exp(Cn^{1/2-\varepsilon})\mu^n$ for an explicit constant $\varepsilon > 0$ (where here μ denotes the connective constant for the lattice in question). The result is conditional on a lower bound on the number of self-avoiding polygons of length n, which is proved to hold on the hexagonal lattice $\mathbb{H}$ for all n, and subsequentially in n for $\mathbb{Z}^2$. A power-law upper bound on $c_n\mu^{-n}$ for $\mathbb{H}$ is also proved, contingent on a non-quantitative assertion concerning this lattice's connective constant.

Contents

1. INTRODUCTION

1.1. **Definitions and results.** We will denote by $\mathbb{N}$ the set of *positive* integers. For $u \in \mathbb{R}^2$, let $\|u\|$ denote the Euclidean norm of u. The two-dimensional nearest-neighbour Euclidean lattice $\mathbb{Z}^2 = (\mathbb{Z}^2, E(\mathbb{Z}^2))$ has origin $O = (0,0)$.

A *walk* of length $n \in \mathbb{N} \cup \{0\}$ is a map $\gamma : \{0, \ldots, n\} \to \mathbb{Z}^2$ such that $\|\gamma_{i+1} - \gamma_i\| = 1$ for each $i \in \{0, \ldots, n-1\}$. An injective walk is called *self-avoiding*. Write SAW_n for the set of all self-avoiding walks of length n with $\gamma_0 = O$. The book [MS13] and lecture notes [BDCGS12] offer introductions to the topic of self-avoiding walk.

It follows from a simple sub-multiplicativity result that there exists a constant μ, called the *connective constant*, such that

$$\mu^n \leq |\mathsf{SAW}_n| \leq \exp(o(n))\, \mu^n.$$

The question that we address here concerns the upper bound. Until recently, the best known upper bound on $|\mathsf{SAW}_n|$ was given by the celebrated work of Hammersley and Welsh [HW62].

Theorem 1.1 (Hammersley-Welsh bound). *For $d \geq 2$, there exists $C > 0$ such that, for all $n \in \mathbb{N}$,*

$$|\mathsf{SAW}_n| \leq \exp(Cn^{1/2})\, \mu^n. \tag{1.1}$$

Our definition of SAW_n is made for $d = 2$ because this article's results concern this dimension, but the model's definition in higher dimensions is no different. Theorem 1.1's proof depends on unfolding self-avoiding walks so that certain special self-avoiding walks known as bridges result. Kesten substantially improved Theorem 1.1 for dimensions $d \geq 3$ in [Kes64]. Recently, [Hut18] has observed that, in any dimension at least two, the sub-ballisticity of self-avoiding walk [DCH13] entails that the constant $C > 0$ in Theorem 1.1 may be chosen arbitrarily. In sufficiently high dimension d, Slade [Sla89] has proved the far stronger inference that $|\mathsf{SAW}_n| \sim c_d\, \mu^n$ for some $c_d > 0$; when $d \geq 5$, this result is due to Hara and Slade [HS92].

Our aim is to improve the exponent $1/2$ in the power of n in the exponential in Theorem 1.1 in the two-dimensional case. We mention first that it is expected that

$$|\mathsf{SAW}_n| = n^{11/32+o(1)} \mu^n,$$

where the exponent $11/32$ is predicted [Nie82] for any planar lattice, in contrast to the lattice-dependent value of μ. See [LSW04] for discussions of this prediction and the conjectural $\mathrm{SLE}_{8/3}$ scaling limit of planar self-avoiding walk.

The improved bound that we present requires a power-law lower bound on the correction to exponential growth for self-avoiding polygons of a given length, where such polygons are in essence self-avoiding walks that return to their starting points. This bound is available only subsequentially for $\mathbb{Z}^2$. Here then is the first of our main results.

Theorem 1.2 (Improved Hammersley-Welsh on $\mathbb{Z}^2$). *For any $\varepsilon < \frac{1}{466}$, there exist infinitely many values of $n \in \mathbb{N}$ such that*

$$|\mathsf{SAW}_n| \leq \exp(n^{1/2-\varepsilon})\, \mu^n. \tag{1.2}$$

Indeed, there are infinitely many $j \in \mathbb{N}$ for which all $n \in \mathbb{N} \cap [j, j^{4/3}]$ verify (1.2).

The technique of proof of Theorem 1.2 is rather robust, and it is plausible that a similar result may be obtained for many planar lattices. For a particular lattice, special structure permits the derivation of a stronger result. This lattice is the hexagonal lattice $\mathbb{H}$, which is dual to the triangular lattice $\mathbb{T} = \mathbb{Z} + e^{i\pi/3}\mathbb{Z}$: see the later Figure 8 for a depiction. On $\mathbb{H}$, an analysis of a discretely holomorphic observable, which was exploited in [DCS12] and is discussed in Section 6, leads to a bound of the form (1.2) for all n. Let $\mathsf{SAW}_n(\mathbb{H})$ denote the set of self-avoiding walks of length n starting from a given vertex in $\mathbb{H}$. The hexagonal lattice's connective constant $\mu(\mathbb{H})$ was proved to equal $\sqrt{2 + \sqrt{2}}$ in [DCS12].
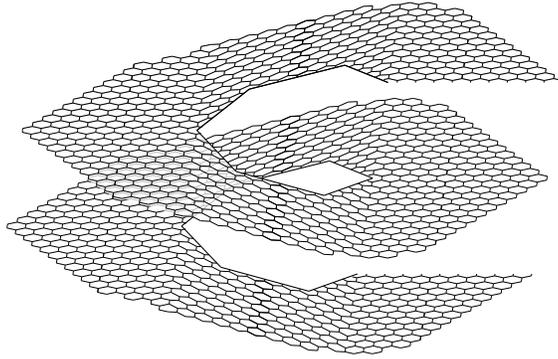
FIGURE 1. A finite section of some $\mathbb{U}_k^\infty$ is embedded in $\mathbb{R}^3$.

Theorem 1.3 (Improved Hammersley-Welsh on $\mathbb{H}$). *Let $\varepsilon \in (0, \frac{1}{42})$. Then, for any $n \in \mathbb{N}$ high enough,*

$$|\mathsf{SAW}_n(\mathbb{H})| \leq \exp(n^{1/2-\varepsilon})\,\mu(\mathbb{H})^n\,. \tag{1.3}$$

Beyond these two bounds, we present a third theorem, in which a polynomial upper bound on the hexagonal lattice's normalized walk count $\mu(\mathbb{H})^{-n}|\mathsf{SAW}_n(\mathbb{H})|$ is obtained subject to a qualitative conjecture concerning numbers of self-avoiding walks. In order to pose the conjecture, consider the universal cover $\mathbb{U}^\infty$ of $\mathbb{H}$ with singularity at the origin; this informal description will be made precise in Section 7.1. Let π_∞ be the canonical projection of $\mathbb{U}^\infty$ to $\mathbb{H}$. For $k \geq 0$, let Λ_k be the ball of radius k for the graph distance on $\mathbb{T}$, centred at the origin; we will view Λ_k as a set of faces of $\mathbb{H}$; note that Λ_0 contains one face, whose centre is the origin. Let $\mathbb{U}_k^\infty := \mathbb{U}^\infty \setminus \pi_\infty^{-1}(\Lambda_k)$, this being the universal cover of $\mathbb{H} \setminus \Lambda_k$; note that $\mathbb{U}^\infty \subset \mathbb{U}_0^\infty$.

Let $k \in \mathbb{N}$. For $v \in \mathbb{U}_k^\infty$, set $\mathsf{SAW}_n(\mathbb{U}_k^\infty, v)$ equal to the set of self-avoiding walks in $\mathbb{U}_k^\infty$ of length n that start at v; absence of translation invariance entails the specification of v in this notation since it is the cardinality of these sets that will concern us. When a self-avoiding walk is split into two, the pieces are also self-avoiding. Thus, the sequence $\left\{\sup_{v \in \mathbb{U}_k^\infty} |\mathsf{SAW}_n(\mathbb{U}_k^\infty, v)| : n \in \mathbb{N}\right\}$ is submultiplicative, so that Fekete's lemma permits us to define

$$\mu(\mathbb{U}_k^\infty) := \lim_{n \to \infty} \sup_{v \in \mathbb{U}_k^\infty} \left|\mathsf{SAW}_n(\mathbb{U}_k^\infty, v)\right|^{1/n}\,. \tag{1.4}$$

Conjecture 1. *There exists $k \in \mathbb{N} \cup \{0\}$ for which $\mu(\mathbb{U}_k^\infty)$ is equal to $\mu(\mathbb{H})$.*

The conjecture is in our view likely to be valid even when $k = 0$. In principle, however, a typical self-avoiding walk on $\mathbb{U}_k^\infty$ may wind continually around the axis of the universal cover, resulting in $\mu(\mathbb{U}_k^\infty) > \mu(\mathbb{H})$.

Here is our third main result.

Theorem 1.4. *Assume Conjecture 1. There exist positive constants C_0 and C such that, for $n \in \mathbb{N}$,*

$$\left|\mathsf{SAW}_n(\mathbb{H})\right| \leq C_0\,n^C \mu(\mathbb{H})^n\,. \tag{1.5}$$

Although this theorem is a conditional result, it is plausible that it considerably simplifies the task of proving (1.5). Indeed, the task of deriving a power-law upper bound for $\mu(\mathbb{H})^{-n}c_n$ has been reduced to proving that the number of self-avoiding walks on $\mathbb{U}_k$ is bounded above by $\mu(\mathbb{H})^{n+o(n)}$ with a non-quantitative $o(n)$ bound.

The introduction continues by presenting some fundamental notation and concepts. It ends by explaining the paper's structure.

1.2. Notation.
For $a < b$, write $[\![a, b]\!] := [a, b] \cap \mathbb{Z}$.

Let (e_1, e_2) denote the standard unit vectors that generate $\mathbb{Z}^2$. For a point $z \in \mathbb{Z}^2$, we write $x(z)$ and $y(z)$ for its first and second coordinates. For a finite set of points $A \subset \mathbb{Z}^2$, write $y_{\max}(A) = \max\{y(a) : a \in A\}$ and $y_{\min}(A) = \min\{y(a) : a \in A\}$; the same applies to the x-coordinate. For walks γ, $y_{\max}(\gamma)$ and $y_{\min}(\gamma)$ refer to the definitions above, with γ being the set of points visited by γ. Then $\mathsf{h}(\gamma) = y_{\max}(\gamma) - y_{\min}(\gamma)$ is called the *height* of γ. (The word 'height' will also refer to the y-coordinate of a point in $\mathbb{Z}^2$.)

For a walk $\gamma \in \mathsf{SAW}_n$ and $0 \le a \le b \le n$, $\gamma_{[a,b]}$ is the walk $(\gamma_a, \gamma_{a+1}, \ldots, \gamma_b)$. We define the *concatenation* $\gamma \circ \gamma'$ of two walks γ and γ' of respective lengths n and m by the formula

$$(\gamma \circ \gamma')_k = \begin{cases} \gamma_k & \text{for } k \le n \\ \gamma_n + \gamma'_{k-n} & \text{for } k \ge n. \end{cases}$$

The length of a self-avoiding walk γ will be denoted by $|\gamma|$.

1.3. Multi-valued maps.
Our main arguments will be presented using a simple combinatorial inequality.

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

Lemma 1.5. *Let $\Phi : A \to \mathcal{P}(B)$. Set m_0 to be the minimum over $a \in A$ of the number of arrows outgoing from a, and M_0 to be the maximum over $b \in B$ of the number of arrows incoming to b. Then $|B| \ge m_0 M_0^{-1} |A|$.*

Proof. The quantities $M_0|B|$ and $m_0|A|$ are upper and lower bounds on the total number of arrows. $\qquad\square$

When there is merely one outgoing arrow from each element of A, we call the multi-valued map Φ *degenerate*; in this case, we may instead view Φ as a function from A to B.

1.4. Self-avoiding bridges and polygons.
A self-avoiding walk in $\mathbb{Z}^2$ of length n that starts at O and verifies $0 < y(\gamma_k) \le y(\gamma_n)$ for $1 \le k \le n$ is called a bridge. Let SAB_n denote the set of self-avoiding bridges of length n.

When bridges are concatenated, the outcome is also a bridge. The resulting supermultiplicativity leads to

$$|\mathsf{SAB}_n| \leq \mu_b^n,$$

where $\mu_b = \lim_n |\mathsf{SAB}_n|^{1/n} \leq \mu$.

Self-avoiding polygons play an essential role in the proofs of the strengthened Hammersley-Welsh bounds, Theorems 1.2 and 1.3.

A self-avoiding walk $\gamma : \{0, \ldots, n\} \to \mathbb{Z}^2$ is called *closing* if $\|\gamma_n - \gamma_0\| = 1$. A closing walk has odd length. A self-avoiding polygon is formed from such a closing walk γ by taking the union of the set of nearest-neighbour edges that interpolate consecutive endpoints of γ with the further edge $\{\gamma_n, \gamma_0\}$. This polygon has length $n + 1$, which is even. A self-avoiding polygon of length four is called a plaquette; it consists of the four edges that border a certain unit square.

Let SAP_n denote the set of equivalence classes of self-avoiding polygons of length $n \in 2\mathbb{N}$, where two such polygons are identified if there exists a vector in $\mathbb{Z}^2$ that translates one to the other.

An element of SAP_n corresponds to $2n$ closing walks of length $n - 1$ that start at the origin: any one of n edges may be removed from a length n polygon, and a choice of two orientations then made for the resulting closing walk.

The next lemma follows from an argument of Kesten [Kes63].

Lemma 1.6. *For any $n \in \mathbb{N}$,*

$$\left|\mathsf{SAP}_{2n+2}\right| \geq \frac{\left|\mathsf{SAB}_n\right|^2}{4(2n + 1)n(n + 1)^3}.$$

Proof. The final inequality of the proof of [BDCGS12, Theorem 2.9] asserts, when $d = 2$, that the number of closing walks in $\mathbb{Z}^2$ of length $2n + 1$ that begin at the origin is at least $n^{-1}(n + 1)^{-2}(2n + 1)^{-1}|\mathsf{SAB}_n|^2$. The number of elements of SAP_{2n+2} is thus seen to be at least the product of $2^{-1}(2n + 2)^{-1}$ and the latter quantity. $\qquad\square$

Two remarks: 1. We study self-avoiding walks, but walks in which the constraint of self-avoidance is violated play a vital role in the proofs of Theorems 1.2 and 1.3. Walks, bridges and polygons are understood to be self-avoiding: the violation of this constraint will be noted with emphasis.

2. We will make such assertions as 'a walk and a polygon are disjoint' and 'a polygon intersects a given set of vertices'. These assertions are abuses of notation, but their meaning is, we hope, clear.

1.5. **Structure of the paper.** The celebrated Hammersley-Welsh unfolding argument proves Theorem 1.1. Our first two principal results depend on developing this argument, and we begin Section 2 by reviewing it. We then explain roughly how we will modify the argument with the use of polygon insertion.

This heuristic discussion will make clear that certain basic tools are needed: polygon abundance, and Madras' join technique for polygons in $\mathbb{Z}^2$. These tools are reviewed or developed in Section 3. By polygon abundance, we mean a power-law lower bound on the normalized number of certain useful polygons of given length n. This bound is obtained subsequentially in n for $\mathbb{Z}^2$. A stronger bound valid for the hexagonal lattice and for *all* lengths has a proof that involves the parafermionic observable. Section 6 contains this derivation.

Sections 4 and 5 offer the rigorous counterpart to the heuristics of Section 2: assertions of polygon abundance will be used to prove the square lattice Theorem 1.2. Section 4 begins by explaining how the task will be carried out over the two sections. Section 5 ends with some brief comments concerning the prospect for obtaining improvements to Theorem 1.2 by making changes to this theorem's proof. Certain perturbations of the argument for Theorem 1.2 are needed alongside the stronger polygon abundance estimate available for the hexagonal lattice to yield Theorem 1.3. The construction and changes are described in Appendix A.

The main part of the paper ends with two sections treating the hexagonal lattice, the first of which we have mentioned. The final Section 7 contains the proof of the strong but conditional Theorem 1.4.

2. The Hammersley-Welsh argument and its prospective modification

In two subsections, we recount this argument and, in a third, we lay out a plan to modify it.

2.1. **Bridge lists.** We define the bridge-product space $\text{B}\Pi_{\ell,j}$, for $\ell, j \in \mathbb{N}$, to be the set of ordered lists $(b_1, \ldots, b_j)$ of length j whose elements are bridges whose respective lengths ℓ_k satisfy $\sum_{k=1}^{j} \ell_k = \ell$ and whose respective heights h_k form a strictly decreasing sequence.

Lemma 2.1. *We have the following two properties.*

(1) If $j \leq (2\ell)^{1/2}$, then $\left|\text{B}\Pi_{\ell,j}\right| \leq \ell^j \mu^\ell$, and if $j > 2\ell^{1/2}$, then $\text{B}\Pi_{\ell,j} = \emptyset$.

(2) There exists $C > 0$ such that, for $\ell \in \mathbb{N}$,

$$\left| \bigcup_{j=1}^{\ell} \text{B}\Pi_{\ell,j} \right| \leq \exp(C\ell^{1/2})\mu^\ell.$$

Remark. The proofs demonstrate the statements with the change $\mu \to \mu_b$. However, the results under review leading to Theorem 1.1 demonstate that $\mu_b = \mu$. We will thus make no further reference to μ_b.

Proof of Lemma 2.1. (1). Fix j and ℓ. If $j > (2\ell)^{1/2}$, the sum of heights of any element $(b_1, \ldots, b_j) \in \mathsf{B\Pi}_{\ell,j}$ is greater than $1 + 2 + \cdots + j > \ell$, which proves that no such bridge-list exists. Assume now that $j \leq (2\ell)^{1/2}$.

The concatenation map $\mathcal{C}$ that sends $(b_1, \ldots, b_j)$ to $b_1 \circ \cdots \circ b_j$ maps $\mathsf{B\Pi}_{\ell,j}$ to SAB_ℓ. Treating this operation as a multi-valued map, though it is a degenerate one, we apply Lemma 1.5 with $A = \mathsf{B\Pi}_{\ell,j}$ and $B = \mathsf{SAB}_\ell$. Clearly, $m_0 = 1$, while the fact noted early in Section 1.4 implies that $|B| \leq \mu^\ell$. The map $\mathcal{C}$ is not injective, because of the ambiguity in how several bridges are concatenated to form a given output bridge. We gauge distance from injectivity by asserting that the quantity M_0 in Lemma 1.5 is at most ℓ^j. The assertion is valid because, for given $b \in \mathsf{SAB}_\ell$, the identity of an element in $\mathcal{C}^{-1}(b)$ may be inferred from the j concatenation points of consecutive bridges along b, and each of these points lies in a set of cardinality ℓ.

(2). Fix ℓ. When the concatenation map is extended to act on $\bigcup_{j=1}^{\ell} \mathsf{B\Pi}_{\ell,j}$, its range again lies in SAB_ℓ. Applying Lemma 1.5 once more, it is enough to argue that $M_0 \leq \exp(C\ell^{1/2})$ for a suitable positive constant C. To bound M_0, note that for given $b \in \mathsf{SAB}_\ell$, it suffices to determine the strictly decreasing sequence $\{h_i : i \geq 1\}$ of heights of $b_1, \ldots, b_j$ in order to determine $b_1, \ldots, b_j$. This sequence is a decreasing partition of the height of b, which is at most ℓ.

In view of a much finer result of Hardy and Ramanujan [HR17], there exists a positive constant C such that the number of decreasing partitions of a given integer $k \geq 1$ is bounded above by $\exp(Ck^{1/2})$. Taking k to equal the height of b, we verify that Lemma 1.5's M_0 is at most $\exp(C\ell^{1/2})$. $\qquad\square$

2.2. The Hammersley-Welsh argument.

Although the proof of Theorem 1.1 that we are reviewing is valid in any dimension $d \geq 2$, our development of this argument requires that $d = 2$. We thus take $d = 2$ throughout the article; even if in this section it is purely as a matter of notational convenience.

A *half-space self-avoiding walk* (which, in accordance with our convention, we will call simply a half-space walk) of length n is a self-avoiding walk γ of length n starting at O with $0 < y(\gamma_k)$ for all $0 < k \leq n$. Write HSW_n for the set of such walks.

Decomposing any self-avoiding walk of length n at its final lowest point, we find that

$$|\mathsf{SAW}_n| \leq \sum_{k=0}^{n} |\mathsf{HSW}_{k+1}| \, |\mathsf{HSW}_{n-k}| \, . \tag{2.1}$$

Lemma 2.2. *There exists $C > 0$ such that, for $n \in \mathbb{N}$,*

$$|\mathsf{HSW}_n| \leq \exp(Cn^{1/2}) \, \mu^n \, .$$

Proof of Theorem 1.1. This is immediate from the two preceding bounds. $\qquad\square$

Proof of Lemma 2.2. Let $n \in \mathbb{N}$. In light of Lemma 2.1(2), it is enough to construct an injective map $\Psi : \mathsf{HSW}_n \to \bigcup_{j=1}^{n} \mathsf{B\Pi}_{n,j}$ from length-n half space walks to bridge-lists of sum length n.

Let $\gamma \in \mathsf{HSW}_n$. Define the *record points* of γ recursively as follows. Let $a_0 = 0$. For $k \geq 1$ odd, let a_k be the last time that $\gamma_{[a_{k-1},n]}$ reaches its highest y-coordinate. For $k \geq 2$ even, let a_k be the last time that $\gamma_{[a_{k-1},n]}$ reaches its lowest y-coordinate. Formally,

$$a_k := \begin{cases} \max\{t > a_k : y(\gamma_t) = y_{\max}(\gamma_{[a_{k-1},n]})\} & \text{if } k \text{ is odd,} \\ \max\{t > a_k : y(\gamma_t) = y_{\min}(\gamma_{[a_{k-1},n]})\} & \text{if } k \text{ is even.} \end{cases} \tag{2.2}$$

Stop this procedure on the first occasion that $a_k = n$.

We call the walks $\gamma_{[a_k,a_{k+1}]}$, $0 \leq k < r$, the *branches* of γ. In this way, we associate to γ its *branch decomposition*

$$\left(\gamma_{[0,a_1]}, \gamma_{[a_1,a_2]}, \ldots, \gamma_{[a_{r-1},a_r]}\right), \tag{2.3}$$

with $r = r(\gamma)$ the number of branches of γ. The branch heights form a strictly decreasing sequence.

When k is even, $\gamma_{[a_k,a_{k+1}]}$ is a bridge after a suitable translation. When this index is odd, it is the vertical reflection of $\gamma_{[a_k,a_{k+1}]}$ that is the translate of a bridge. More formally, write τ for the reflection with respect to the horizontal axis. The *bridge decomposition* of γ is formed by applying τ to every second component of γ's branch decomposition and suitably translating each resulting component so that it is a bridge: that is, the new decomposition is

$$\left(\gamma_{[0,a_1]}, \tau(\gamma_{[a_1,a_2]} - \gamma_{a_2}), \gamma_{[a_2,a_3]} - \gamma_{a_2}, \ldots\right). \tag{2.4}$$

We define $\Psi(\gamma)$ to be the bridge decomposition of γ. Since the transformation that specifies this map is invertible, Ψ is injective, and Lemma 2.2 is proved. $\qquad\square$

2.3. **A rough guide to how polygon insertion will modify the Hammersley-Welsh argument.** The injective map Ψ constructed in Lemma 2.2 sends a half-space walk to its bridge decomposition. Since this map is so vital to the proof of Theorem 1.1, we call Ψ the *Hammersley-Welsh* map. It will be by perturbing the construction of Ψ that we prove Theorems 1.2 and 1.3, and thus we emphasise its role. Note, however, that the proof of Theorem 1.1 amounts to considering the composition $\mathcal{C} \circ \Psi : \mathsf{HSW}_n \to \mathsf{SAB}_n$, where $\mathcal{C}$ is the concatenation map from Lemma 2.1's proof, and noting that the number of arrows incoming to any point in the range of this map is bounded above by $\exp(Cn^{1/2})$. (We have called Hammersley-Welsh's proof an unfolding argument because $\mathcal{C} \circ \Psi$ unfolds walks into bridges.)

Here, we outline the improvement of the Hammersley-Welsh bound. Fix $\varepsilon > 0$ to be the desired reduction of the $1/2$ exponent of (1.2); we will see below that the improvement is limited, as ε needs to be smaller than some threshold.

A simple argument will permit us to restrict the domain of Ψ to a special class of half-space walks. This set will be denoted (for a reason that we explain shortly) by MHSW_n. In essence, a member $\gamma \in \mathsf{MHSW}_n$ is an element of HSW_n that

(1) has *many branches*: the branch, and thus also the bridge, decomposition of γ has at least an order of $n^{1/2-\varepsilon}$ terms;

(2) and is *horizontally confined*: γ is contained in a vertical strip of width of order $n^{1/2+\varepsilon}$.

Our goal will be to prove that $|\mathsf{W\!HSW}_n|\mu^{-n} \leq \exp(Cn^{1/2-\varepsilon})$ for some constant C. The bound $|\mathsf{HSW}_n \setminus \mathsf{W\!HSW}_n|\mu^{-n} \leq \exp(Cn^{1/2-\varepsilon})$ is straightforward: see Lemmas 4.2 and 4.3. These two bounds will then imply Theorem 1.2.

Fix $r \in \mathbb{N}$ to be a given value whose order is at least $n^{1/2-\varepsilon}$. Set $\mathsf{W\!HSW}_n(r) = \{\gamma \in \mathsf{W\!HSW}_n : \gamma$ has r branches$\}$. We will explain how to bound above $\big|\mathsf{W\!HSW}_n(r)\big|\mu^{-n}$. The desired bound on $|\mathsf{W\!HSW}_n|\mu^{-n}$ will then be obtained by summing this bound over such r.

Consider the restriction of Ψ to $\mathsf{W\!HSW}_n(r)$; its image is $\mathrm{B}\Pi_{n,r}$. How may we seek to improve Ψ for our purpose, when this map is already injective? The answer lies in modifying Ψ into a multi-valued map with a large number of outgoing arrows from each point of $\mathsf{W\!HSW}_n(r)$, without significantly compromising the map's injectivity.

To survey options for introducing outgoing arrows from a given domain point, denote such a point by $\gamma \in \mathsf{W\!HSW}_n(r)$. Due to the high number of branches of γ, it is in essence true that at least one-half of them have height exceeding $n^{1/2-\varepsilon}$. These will be called the *tall branches*. Since the vertical intervals occupied by consecutive branches form a decreasing sequence under containment, we see that γ's tall branches all cross a given vertical interval of length $n^{1/2-\varepsilon}$. Moreover, as γ is horizontally confined, every tall branch crosses between its upper and its lower side some given translate R of the rectangle $[-n^{1/2+\varepsilon}, n^{1/2+\varepsilon}] \times [-n^{1/2-\varepsilon}, n^{1/2-\varepsilon}]$, all the while remaining in R.

Thus, γ is necessarily a rather dense object: the notation W is intended to evoke γ's repeated up-down movement at close quarters.

In particular, the tall branches of γ often pass close to each other inside R. Say that a point of γ in R is a *near self-touch* if it is the rightmost point of a tall branch at a given height among the y-coordinates assumed by R, and if the horizontal interval of length $n^{2\varepsilon}$ to its right intersects another tall branch of γ. Each of the at least $r/2$ tall branches of γ contains one rightmost point for each of the $n^{1/2-\varepsilon}$ heights of R. Due to the limited volume of R, most of these rightmost points may be shown to be near self-touches. In conclusion, γ has of order $r\, n^{1/2-\varepsilon}$ near self-touches.

We will alter the definition of Ψ in a way that seeks to exploit the abundance of near self-touches among elements in its domain. Our alteration of Ψ, which we will call Φ, entails the insertion of certain self-avoiding polygons around near self-touches. The polygons to be inserted will be drawn from a certain class WSAP_m^u of *wide* polygons; here, u and m are parameters. The set WSAP_m^u is a collection of self-avoiding polygons of length m each of which has, roughly expressed, width at least u. A choice of u will be made of order n^δ, where $\delta > 0$ is a parameter chosen so that the resulting choice equals or slightly exceeds the horizontal near self-touch distance $n^{2\varepsilon}$; while m will be a given value that is at most $n^{2\delta}$. The value δ should be at least 2ε and, in this overview, we in fact set $\delta = 2\varepsilon$.

In a first attempt to modify the definition of Ψ, suppose that we attempt to insert a single polygon on to a half-space walk in the domain of Ψ, and then take the bridge decomposition of this modified walk. That is, suppose that we define Φ to be a multi-valued map whose domain is the product space $\mathsf{W\!HSW}_n(r) \times \mathsf{WSAP}_m^u$.
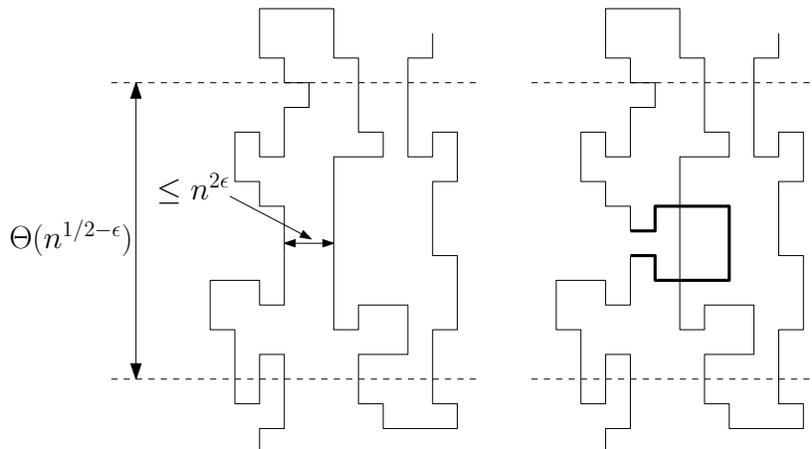
FIGURE 2. By means of a local deformation on its left, a square-shaped poly-
gon P is inserted into a branch walk γ at a height where this branch nearly
touches the branch to its right. The implanted copy of P intersects the branch
on the right. The edges introduced into the walk are shown in bold in the
right sketch.

Fix a domain point $(\gamma, P) \in \mathsf{MHSW}_n(r) \times \mathsf{WSAP}_m^u$ and consider the decomposition $b_1, \ldots, b_r$
of γ into branches. In this paragraph, we will abusively treat each b_i as if it were bridge, even
though alternate vertical reflection is in fact needed to ensure this. The set $\Psi(\gamma, P)$ of bridge-
lists associated with (γ, P) will be obtained by inserting P onto a given tall branch of γ. Let
b_j denote this branch; we choose the index j so that the cardinality of the collection of near
self-touches of γ along b_j has the typical order of $n^{1/2-\varepsilon}$. Let γ_t denote a generic element in
this collection of near self-touches. Define $\tilde{b}_j$ to be the bridge obtained by gluing the polygon
P to the right of b_j, at the level of the point γ_t. The procedure used to glue a polygon to a
bridge is called the *Madras join* and will be reviewed in the next section. By definition of the
Madras join, the result $\tilde{b}_j$ has length $|b_j| + m + 16$ and, excepting a few marginal examples, is
a bridge with the same height as b_j. Consider now the list $(b_1, \ldots, \tilde{b}_j, \ldots, b_r)$ which is given
by $\Psi(\gamma)$ except that b_j is replaced by $\tilde{b}_j$; this is a bridge-list of total length $n + m + 16$. Then
$\Phi(\gamma, P)$ is the set of all bridge-lists obtained in this manner, as γ_t ranges over points of near
self-touch for γ that lie in b_j. Thus, Φ is a multi-valued map from $\mathsf{MHSW}_n(r) \times \mathsf{WSAP}_m^u$ to
$\mathsf{B\Pi}_{n+m+16,r}$.

We now evaluate how successful this modification Φ of the Hammersley-Welsh map Ψ is in
finding an improved upper bound on $|\mathsf{MHSW}_n(r)|$. Lemma 1.5 and Lemma 2.1(1) imply that

$$|\mathsf{MHSW}_n(r)| \cdot |\mathsf{WSAP}_m^u| \leq \frac{M_0}{m_0} |\mathsf{B\Pi}_{n+m,r}| \leq \frac{M_0}{m_0} \cdot (n + m)^r \mu^{n+m}, \qquad (2.5)$$

where M_0 is the maximum number of arrows incoming to a given range point; and m_0, the
minimum number of arrows outgoing from any domain point, is, in light of the preceding
paragraph, at least $n^{1/2-\varepsilon}$ up to a constant factor which we here neglect.

For Φ to produce a meaningful result, the constant M_0 associated with it needs to be bounded from above. We next explain how this is done.

Let $\bar{b} = (b_1, \ldots, b_r) \in \mathsf{B\Pi}_{n+m+16,r}$ be a point in the image set of Φ. Consider the walk resulting from the concatenation of the elements $b_1, \ldots, b_r$ after vertical reflection of every other bridge in this sequence. Vitally, this result is not self-avoiding: the choice of P and of the point γ_t of near self-touch have been made so that the branch to which P was added intersects its right-adjacent counterpart. We may thus infer the branch to which P was added (though in the present case, we know it: it is b_j); and, in view of the fixed length of P, we further in essence infer the location of the joining of P to that branch. Only a finite number C of pre-images of $\bar{b}$ are then possible, where the value C accounts for possible microscopic variations in the surgery used to join P to the branch of γ. Thus, (2.5) becomes

$$|\mathsf{WHSW}_n(r)| \cdot \mu^{-n} \leq C n^{-(1/2-\varepsilon)} \cdot (n+m)^r \cdot \left(|\mathsf{WSAP}_m^u| \cdot \mu^{-m}\right)^{-1}. \tag{2.6}$$

This should be compared to the naive bound involving the original map Ψ, which yields

$$|\mathsf{WHSW}_n(r)| \cdot \mu^{-n} \leq n^r. \tag{2.7}$$

The difference between the right-hand factors of the last two inequalities represents the change in the outcome due to the insertion of a single polygon. The order $n^{-(1/2-\varepsilon)} \ll 1$ benefit due to polygon placement entropy is possibly compromised by the polygon scarcity cost $(|\mathsf{WSAP}_m^u|\mu^m)^{-1}$ (it will transpire that the loss due to n^r being replaced by $(n+m)^r$ is insignificant). To understand the role of the term $(|\mathsf{WSAP}_m^u|\mu^m)^{-1}$, note that, since WSAP_m^u is a collection of polygons of length m, we have $|\mathsf{WSAP}_m^u| \leq |\mathsf{SAP}_m| \leq \mu^m$: see [MS93, (3.2.5)]. Insofar as this inequality on $|\mathsf{WSAP}_m^u|$ fails to be sharp, the resulting polygon scarcity constitutes an opposing force that undoes some of the benefit achieved by modifying Ψ to Φ. Thus, it is an upcoming challenge to demonstrate a lower bound on the number of wide polygons. If we posit that, for some constant $\alpha > 0$, there exist u of order $n^{2\varepsilon}$ and $m \leq u^2$ for which $|\mathsf{WSAP}_m^u|\mu^{-m} \geq u^{-\alpha}$, then (2.6) would become

$$|\mathsf{WHSW}_n(r)| \cdot \mu^{-n} \leq C n^{-(1/2-\varepsilon)+2\alpha\varepsilon} \cdot (n+m)^r, \tag{2.8}$$

where the influence of the modified argument is transmitted again through the first term on the right-hand side, namely $C n^{\varepsilon(1+2\alpha)-1/2}$. Given the information that suitable wide polygons are plentiful in the sense of the parameter α, the modification is seen to return a benefit when $\varepsilon(2\alpha+1) < 1/2$. Even when this condition is met, the improvement of $n^{\varepsilon(1+2\alpha)-1/2}$ of (2.8) compared to (2.7) is only very modest: after all, r has order at least $n^{1/2-\varepsilon}$.

To obtain a greater improvement, we have little choice but to iterate our procedure: instead of adding one polygon, we will add as many as needed to reduce the right-hand side of (2.6) below $\exp(n^{1/2-\varepsilon})$. Suppose that we add $K = \kappa r$ polygons independently and uniformly into admissible near-touch slots; note that $\kappa > 0$ is a parameter indicating the order of the number of polygons inserted per tall branch. Adding a polygon to a near-touch slot may block nearby near self-touches from receiving polygons. However, only at most order $n^{2\varepsilon}$ other near self-touches are thus affected. Indeed, since there is at least an order of $n^{1-2\varepsilon}$ near self-touches in total, the assumption that $K = o(n^{1-4\varepsilon})$, which in practice certainly permits κ to be chosen to be of unit order, entails that there are of order $n^{K(1-2\varepsilon)}$ sets of near self-touches for the K polygons to be added.

In this light, we reconsider Φ to be a multi-valued map from $\mathsf{MHSW}_n(r) \times (\mathsf{WSAP}_m^u)^K$ to $\mathrm{B\Pi}_{n+K(m+16),r}$ obtained by adding K polygons drawn from WSAP_m^u to the tall branches of $\gamma \in \mathsf{MHSW}_n(r)$. The reasoning offered for the case of the insertion of a single polygon now shows that $m_0 \geq n^{K(1-2\varepsilon)}$ and $M_0 \leq C^K$. Assuming again that $|\mathsf{WSAP}_m^u|\mu^{-m} \geq u^{-\alpha}$, we find that

$$|\mathsf{MHSW}_n(r)| \cdot \mu^{-n} \leq \left[Cn^{-(1/2-\varepsilon)+2\alpha\varepsilon}\right]^K \cdot (n+Km)^r$$
$$\leq \exp\left[\left(-\kappa\big(1/2 - \varepsilon(1+2\alpha)\big) + 1\right)r\log n + r\log(2C^\kappa)\right],$$

where the latter bound depends on the practically harmless replacement of $(n+Km)^r$ by n^r. The choice of $\varepsilon > 0$ so that $\varepsilon(1+2\alpha) < 1/2$, and the further choice of κ to be a unit-order quantity verifying $\kappa > \big(1/2 - \varepsilon(1+2\alpha)\big)^{-1}$, yield the inference that $|\mathsf{MHSW}_n(r)| \cdot \mu^{-n} \to 0$; in particular, for a large enough $C' > 0$, we learn that

$$|\mathsf{MHSW}_n(r)| \cdot \mu^{-n} \leq \exp\left(C'n^{1-\varepsilon}\right).$$

These heuristic considerations offer a road map for the rigorous argument for Theorem 1.2 which will be presented in Sections 4 and 5 (with the variation needed for Theorem 1.3 explained in Appendix A). The proof hinges on the abundance of polygons, which is to say, on a power-law lower bound on $|\mathsf{WSAP}_m^u|\mu^{-m}$. Such bounds will be studied in Section 3; they are proved to hold subsequentially for the square lattice and for all u (and well-chosen values of m) for the hexagonal lattice.

3. Self-avoiding polygons: counting and joining

In this section, we present two tools which are needed to rigorously implement the plan that has just been sketched. In the preceding discussion, we invoked an assertion of polygon abundance, namely $|\mathsf{WSAP}_m^u| \geq u^{-\alpha}$. In Section 3.1, we will present a precise definition of the wide polygon set WSAP_m^u. Propositions 3.1 and 3.2 are our rigorous assertions of polygon plenitude. We also review in Section 3.2 Madras' joining technique which allows polygons to be attached to the branches of γ.

3.1. **Wide polygons.** For a polygon $p \in \mathsf{SAP}_m$, the *width* of p is

$$\mathsf{w}(p) := x_{\max}(P) - x_{\min}(P),$$

where P is any representative of the equivalence class p. The *line-width* of p is

$$\mathsf{Lw}(p) := \max\{\, x_{\max}(P \cap \ell_y) - x_{\min}(P \cap \ell_y) : y \in \mathbb{Z} \text{ for which } P \cap \ell_y \neq \emptyset \,\},$$

where ℓ_y is the horizontal line $\mathbb{Z} \times \{y\}$. Note that $\mathsf{w}(p) \geq \mathsf{Lw}(p)$; the former quantity may much exceed the latter.

Let $m \in 2\mathbb{N}$ and $u \in \mathbb{N}$. We define the *wide polygon* set

$$\mathsf{WSAP}_m^u = \{\, p \in \mathsf{SAP}_m : \mathsf{Lw}(p) \geq u \,, \mathsf{h}(p) \leq 16u \,\},$$

where the height notation h from Section 1.2 has been extended to polygons.

For practical purposes, we will choose one particular rooting for wide polygons. Indeed, the walk P corresponding to any $p \in \mathsf{WSAP}_m^u$ is such that

- the line-width of p is realized by P at height 0, and this is the lowest level where it is so realized:

$$\mathsf{Lw}(p) = x_{\max}(P \cap \ell_0) - x_{\min}(P \cap \ell_0) > \max\{x_{\max}(P \cap \ell_y) - x_{\min}(P \cap \ell_y) : y < 0\}$$

- and the origin is the leftmost point of P at height 0, i.e. $0 = x_{\min}(P \cap \ell_0)$.

By convention, the walk that represents p begins and ends at the origin, and traverses the edges that comprise P in a counterclockwise manner, so that the interior of P is on the left during the walk. The correspondence between polygon and walk offered by this convention is a bijection, because it specifies a representative of the equivalence class in question, which amounts to distinguishing a vertex on the polygon; and it then selects one of the two possible orientations of the resulting walk.

The abundance of wide polygons is important for our application. We verify this abundance by two different means, one for $\mathbb{Z}^2$ and the other for $\mathbb{H}$. The latter inference is stronger. It is in deriving this inference that the parafermionic observable available for $\mathbb{H}$ is used during the proof of Theorem 1.3.

Proposition 3.1 (Subsequential abundance of polygons for $\mathbb{Z}^2$). *For any $\alpha > 29$, there exists an infinite set $A \subset \mathbb{N}$ such that, for all $u \in A$ there exists $m \in [\![u/5, u^2/4]\!]$ for which*

$$\left|\mathsf{WSAP}_m^u\right| \geq u^{-\alpha}\mu^m\,.$$

Write $\mathsf{WSAP}_m^u(\mathbb{H})$ for the set of self-avoiding polygons on $\mathbb{H}$, with the same properties as those defining wide polygons on $\mathbb{Z}^2$.

Proposition 3.2 (Abundance of polygons for $\mathbb{H}$). *There exist positive constants C and c_0 such that, for all $u \in \mathbb{N}$, we may find $m \leq c_0 u^2$ for which*

$$|\mathsf{WSAP}_m^u(\mathbb{H})| \geq C\,u^{-10}\mu(\mathbb{H})^m\,.$$

Proposition 3.2 is a principal element of the variation of the proof of Theorem 1.2 that is needed to obtain Theorem 1.3. Its proof appears after the derivation of Theorem 1.2 is completed, in Section 6.

The rest of Section 3.1 is dedicated to proving Proposition 3.1. We start with a lemma that bounds the number of wide polygons in terms of the number of bridges.

Lemma 3.3. *For any $m \in \mathbb{N}$, there exists $u \in [m^{1/2}, m]$ such that*

$$\left|\mathsf{WSAP}_{4m+4}^u\right| \geq (\log 2)^2 3^{-1} 2^{-20} \left(\log u\right)^{-2} u^{-21} \left|\mathsf{SAB}_m\right|^4\,.$$

Proof. Let $m \in \mathbb{N}$. Specifying the diameter $\mathsf{diam}(p)$ of a polygon $p \in \mathsf{SAP}_{2m+2}$ to be $\max\{\mathsf{h}(p), \mathsf{w}(p)\}$, it is easily seen that $m^{1/2} \leq \mathsf{diam}(p) \leq m$. Let $u \in [m^{1/2}, m]$ be a number of the form $2^{-j}m$ such that the set

$$\{\, p \in \mathsf{SAP}_{2m+2} : u/2 \leq \mathsf{diam}(p) \leq u\,\}$$
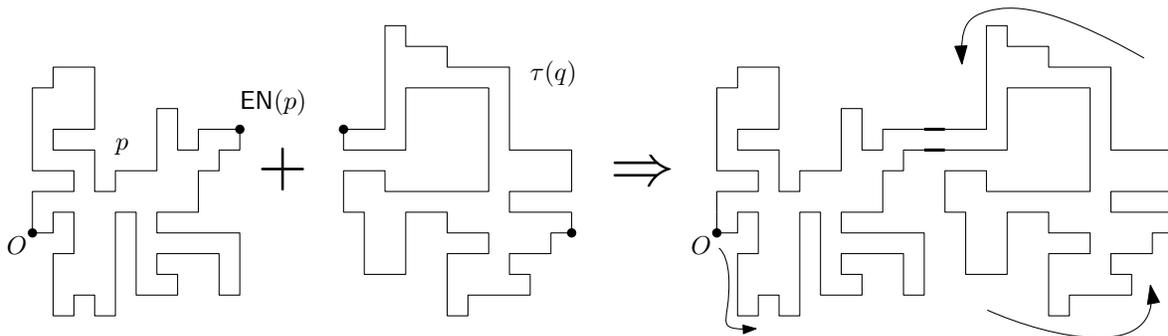
FIGURE 3. Two elements p, q in $\overline{\mathsf{WSAP}}^u_{2m+2}$ with $y\big(\mathsf{EN}(p)\big) = y\big(\mathsf{EN}(q)\big)$ are joined to form $J\big(p, \tau(q)\big) \in \mathsf{WSAP}^u_{4m+4}$. The dot at O and the curved arrows show the root and orientation of the joined polygon that is dictated by our convention.

has maximal cardinality among such sets. At least half of the polygons in the displayed set are at least as wide as they are high; (indeed, any polygon that does not satisfy this condition is the right-angled rotation of one that does). Writing $\overline{\mathsf{WSAP}}^u_{2m+2}$ for the set of polygons $p \in \mathsf{SAP}_{2m+2}$ with $\mathsf{w}(p) \geq u/2$ and $\mathsf{h}(p) \leq u$, we thus see that

$$\big|\overline{\mathsf{WSAP}}^u_{2m+2}\big| \geq \frac{1}{2 \log_2 m} \big|\mathsf{SAP}_{2m+2}\big|. \tag{3.1}$$

When $p \in \overline{\mathsf{WSAP}}^u_{2m+2}$ is depicted as a walk, we may specify the west-south vertex of p, $\mathsf{WS}(p)$, (and its east-north vertex $\mathsf{EN}(p)$) to be the lowest among the leftmost (and the highest among the rightmost) elements of $\mathbb{Z}^2$ visited by the walk p. Clearly each $p \in \overline{\mathsf{WSAP}}^u_{2m+2}$ has a unique representation such that $\mathsf{WS}(p) = 0$. Then $x\big(\mathsf{EN}(p)\big) = \mathsf{w}(p)$.

Let $p, q \in \overline{\mathsf{WSAP}}^u_{2m+2}$ be represented by walks such that $\mathsf{WS}(p) = \mathsf{WS}(q) = 0$ and be such that $y\big(\mathsf{EN}(p)\big) = y\big(\mathsf{EN}(q)\big)$. We will join the polygon p to a reflection $\tau(q)$ of q to form a new polygon $J\big(p, \tau(q)\big)$ about which we will claim that $J(p, q) \in \mathsf{WSAP}^u_{4m+4}$. The joining operation J does the same job as the Madras join which we will review in Section 3.2, but the special geometry arising from the assumption that $y\big(\mathsf{EN}(p)\big) = y\big(\mathsf{EN}(q)\big)$ permits the use of a simpler technique in the present case. Figure 3 illustrates the operation J.

The polygon q may be reflected in the vertical axis, via the reflection τ, and then horizontally translated to such a position that its vertex in correspondence with $\mathsf{EN}(q)$ is to be found one unit directly to the right of $\mathsf{EN}(p)$. The union of the edges that comprise p and this translation of $\tau(q)$ intersects the four edges of the plaquette whose north-west corner is $\mathsf{EN}(p)$ along the west and east sides of the plaquette. By replacing these two edges in the union by the plaquette's north and south sides, a polygon of length $2(2m+2)$ results. This polygon is $J\big(p, \tau(q)\big)$.

We show now that $J\big(p, \tau(q)\big)$ belongs to WSAP_{4m+4}. The height of this polygon is at most $\mathsf{h}(p) + \mathsf{h}(q) \leq 2u$. Moreover, $\mathsf{WS}(J\big(p, \tau(q)\big)) = \mathsf{WS}(p) = 0$ and $J\big(p, \tau(q)\big)$ realizes its line-width $\mathsf{Lw}(J\big(p, \tau(q)\big)) = \mathsf{w}(p) + \mathsf{w}(q) + 1 > u$ at height zero. Finally, O is the leftmost point

of $J\big(p, \tau(q)\big)$ at height 0; hence $J\big(p, \tau(q)\big)$ is rooted according to our convention for wide polygons. The claim that $J\big(p, \tau(q)\big) \in \mathsf{WSAP}^u_{4m+4}$ has been confirmed.

Note that $J\big(p, \tau(q)\big)$ determines $(p, q) \in \mathsf{WSAP}^u_{2m+2}$. Indeed, there is a unique vertical line whose coordinate has the form $k + 1/2$ for some $k \in \mathbb{Z}$ that cuts $J\big(p, \tau(q)\big)$ through a consecutive pair of horizontal edges, and for which the replacement in $J\big(p, \tau(q)\big)$ of the horizontal with the vertical sides of the associated plaquette results in two polygons, of which the one whose vertex set contains the origin has $2m + 2$ edges (and, as it happens, so does the other). The plaquette in question was used in the surgical formation of $J\big(p, \tau(q)\big)$, and the operation just described undoes this surgery so that p and $\tau(q)$ are recovered. This injectivity on the part of J implies that

$$\big|\mathsf{WSAP}^u_{4m+4}\big| \geq \sum_{j=-\lceil u \rceil}^{\lfloor u \rfloor} \big|\big\{ p \in \overline{\mathsf{WSAP}}^u_{2m+2} : y(\mathsf{ne}(p)) = j \big\}\big|^2,$$

where it is since u may not be an integer that use is made of rounding. This right-hand side is seen to be at least $(2u+1)^{-1}\big|\overline{\mathsf{WSAP}}^u_{2m+2}\big|^2$ by means of the Cauchy-Schwarz inequality. Applying (3.1) and Lemma 1.6 with $n = m$, we see that

$$\big|\mathsf{WSAP}^u_{4m+4}\big| \geq (2u+1)^{-1}\left(\frac{\big|\mathsf{SAB}_m\big|^2}{8(2m+1)m(m+1)^3 \log_2 m}\right)^2.$$

Since $u^2 \geq m \geq 1$, a short algebraic manipulation leads to

$$\big|\mathsf{WSAP}^u_{4m+4}\big| \geq (\log 2)^2 3^{-1} 2^{-20} \big(\log u\big)^{-2} u^{-21} \big|\mathsf{SAB}_m\big|^4.$$

This completes the proof of Lemma 3.3. $\qquad\square$

Proof of Proposition 3.1. As [MS13, Corollary 3.1.8] reviews, a variation of the Hammersley-Welsh argument for Theorem 1.1 establishes that the partition function for bridges diverges at its radius of convergence μ^{-1}:

$$\sum_{m=1}^{\infty} \mu^{-m} |\mathsf{SAB}_m| = \infty.$$

Let $\varepsilon > 0$. There are thus infinitely many values of $m \in \mathbb{N}$ such that $|\mathsf{SAB}_m| \geq m^{-1-\varepsilon}\mu^m$. Inserting this bound into Lemma 3.3, we obtain infinitely many pairs (m, u) with $u \in [m^{1/2}, m]$ such that

$$\big|\mathsf{WSAP}^u_{4m+4}\big| \geq c_0\, u^{-21}(\log u)^{-2} m^{-4-4\varepsilon}\, \mu^{4m} \geq c_0\, u^{-29-8\varepsilon}(\log u)^{-2}\, \mu^{4m},$$

where $c_0 = (\log 2)^2 3^{-1} 2^{-20}$. Since $\varepsilon > 0$ is arbitrary, the proposition follows by relabelling m. $\qquad\square$

3.2. **Madras joining.** In [Mad91], Madras introduced a technique by which a pair of polygons may be joined to form a new polygon. We will be using his technique in order to attach polygons to the side of bridges or their vertical reflections. We begin by briefly reviewing his procedure. Then we state and prove Lemma 3.4. This result gathers together all properties of the Madras join which we will use. This section's hexagonal lattice counterpart is postponed to Appendix A.

The eight parts of Lemma 3.4 are all needed in this article. However, the reader may lose little by perusing the lemma briefly and omitting its proof; the core of the argument may even be followed by merely interpreting the Madras join as a procedure that inserts a polygon at a given height on the right of a walk while effecting a bounded change of length near a distinguished plaquette, called the junction plaquette, about which the join is made.

The reader who wishes to follow details precisely is encouraged, in reading the lemma and its proof, to consult either [Mad91] or the article [Ham18] which has also made use of Madras' technique. Section 4.1 of the latter article contains a detailed review of the technique, including a depiction in Figure 1 of the various cases of local deformation used in the operation.

3.2.1. *An overview of Madras' joining technique.* For any given polygon P, the integer interval $[\![y_{\min}(P), y_{\max}(P)]\!]$ comprises the vertical coordinates of vertices in $\mathbb{Z}^2$ visited by the polygon. Suppose given a pair of polygons P and Q such that

$$[\![y_{\min}(P) - 1, y_{\max}(P) + 1]\!] \cap [\![y_{\min}(Q) - 1, y_{\max}(Q) + 1]\!] \neq \emptyset. \qquad (3.2)$$

Let $n, m \in 2\mathbb{N}$ denote the respective lengths of P and Q. The *Madras join* of P and Q is a polygon, which we will denote by $M(P, Q)$, formed by attaching Q after horizontal displacement to the right side of P. The procedure for joining begins by shifting Q far to the right of P, and bringing it back, step by step, until there is a vertex in P and a vertex in the translate of Q that are either equal or whose displacement is a vector of vertical orientation and of Euclidean norm at most two. We then define the vertex Y to be the vertex z of maximal y-coordinate among those for which $\{z - e_2, z, z + e_2\}$ intersects P and the translate of Q. A local surgery is performed on the two polygons in the neighborhood of Y. The local geometry of P determines the form of the surgery on P according to a division into several cases; and similarly for the translate of Q. The two local surgeries are performed in such a way that the modified translate of Q may be translated a certain few steps to right into such a position that there is a plaquette whose intersection with the locally modified copy of P consists of its left vertical edge, and whose intersection with the locally modified translate of Q consists of its right vertical edge.

With Q indeed relocated in this way, these two vertical edges are replaced in the collection of edges in P and the new translate of Q by the two horizontal edges in the plaquette. The Madras join polygon $M(P, Q)$ is defined to be the resulting polygon. The plaquette involved in the final step of this procedure will be called the junction plaquette associated with the joining of P and Q.

The local surgeries performed are such that the length of $M(P, Q)$ equals $n + m + 16$, so there is always a net gain of sixteen in the number of edges involved as a result of the joining operation. The collection of edges in the symmetric difference of the edge collections of $M(P, Q)$ and of the union of P with the surgically determined translate of Q may be viewed as an error set arising in surgery. This error set contains a bounded number of edges (there are at most twenty of them), all these in the locale of the junction plaquette associated with P and Q.

3.2.2. *Our application of Madras joining, with bridges and polygons.* In the present paper, we will not be joining pairs of polygons, but rather pairs (γ, P) where γ is a bridge, or the

vertical reflection of one, and P is a polygon. In a natural extension of notation, the walk resulting from the join, in which a horizontal translate of P is in essence incorporated into the range of γ, will be denoted by $M(\gamma, P)$.

In our applications, this outcome $M(\gamma, P)$ will be a bridge, or the reflection of one: indeed, the first part of the next result ensures this property. Note that the condition (3.3) on vertical alignment of γ and P is slightly stronger than that (3.2) which ensures that Madras joining for a pair of polygons may occur.

Lemma 3.4. *Let γ be a bridge and let P be a polygon. Suppose that*

$$[\![y_{\min}(P) - 3, y_{\max}(P) + 3]\!] \subseteq [\![y_{\min}(\gamma), y_{\max}(\gamma)]\!] . \tag{3.3}$$

(1) *The Madras join $M(\gamma, P)$ is well-defined and is a bridge whose endpoints are shared with γ.*

(2) *Write $n \in \mathbb{N}$ for the length of γ and $m \in 2\mathbb{N}$ for the length of P. The endpoints of the right vertical edge in the junction plaquette of $M(\gamma, P)$ belong to $M(\gamma, P)$; their chemical distance along $M(\gamma, P)$ equals $m + 7$.*

(3) *Denote the junction plaquette of $M(\gamma, P)$ by J, and by $M(\gamma, P)\Delta J$ the modification formed from $M(\gamma, P)$ by the removal of the horizontal edges in J and the insertion of the vertical ones. Then $M(\gamma, P)\Delta J$ comprises a walk γ' and a polygon Q. No right translation of Q, including Q itself, intersects γ'.*

(4) *Suppose that P' is a polygon that also verifies (3.3) and for which*

$$[\![y_{\min}(P) - 2, y_{\max}(P) + 2]\!] \cap [\![y_{\min}(P') - 2, y_{\max}(P') + 2]\!] = \emptyset .$$

Then $M\big(M(\gamma, P), P'\big) = M\big(M(\gamma, P'), P\big)$.

(5) *Suppose given the output $M(\gamma, P)$ and the junction plaquette employed in Madras joining a bridge γ of length $n \in \mathbb{N}$ and a polygon P of length $2m$. The number of consistent values of (γ, P) is at most four.*

(6) *Let $P + (i, 0)$ denote the horizontal translate of P which is locally modified to form $M(\gamma, P)$. Let SE (or NE) denote the rightmost among the lowest (or the highest) vertices in $P + (i, 0)$. The right-side of $P + (i, 0)$ is the collection of vertices in $P + (i, 0)$ encountered on a counterclockwise journey between SE and NE, including the journey's endpoints. Every element of the right-side of $P + (i, 0)$ lies in $M(\gamma, P)$.*

(7) *All but at most two edges in γ lie in $M(\gamma, P)$.*

(8) *The walk $M(\gamma, P)$ is the concatenation of three subwalks. The first is an initial subwalk of γ; the second has length either $m + 17$ or $m + 18$; and the third is a final subwalk of γ.*

If γ is instead supposed to be the vertical reflection of a bridge, then the same statements hold, except that $M(\gamma, P)$ is a vertically reflected bridge.

Proof: (1). The y-coordinate of the vertex Y in Madras' construction belongs to $[\![y_{\min}(P) - 1, y_{\max}(P) + 1]\!]$. Any endpoint of an edge that is added or removed during surgery has y-coordinate equal to $y(Y) - 1$, $y(Y)$ or $y(Y) + 1$. Thus the y-coordinates of affected endpoints

are seen to lie in the interval $[\![y_{\min}(P) - 2, y_{\max}(P) + 2]\!]$. By hypothesis, such coordinates lie in $[\![y_{\min}(\gamma) + 1, y_{\max}(\gamma) - 1]\!]$. Since modifications are not made at y-coordinate values that are extremal for γ, the status of γ as a bridge or a vertical reflection of one is unaffected by the change $\gamma \to M(\gamma, P)$.

(2). Because [Ham18, Figure 1] presents the left, rather than the right, of the two local deformations made during surgery, it is helpful, in regard to the first assertion, to argue the claim in which γ is a polygon of length n and P is a walk (the choice of notation is incongruous but it is adopted only temporarily); and where it is instead claimed that the chemical distance along $M(\gamma, P)$ between the two endpoints of the left vertical edge in the junction plaquette equals $n + 7$. In each of the seven cases depicted in [Ham18, Figure 1], consider the two endpoints of the union of the dotted line segments. The vertices so denoted are elements of γ whose longer chemical distance around γ equals $n - r$ where $r \in \{1, 2\}$ is the number of dotted line segments in the case in question. In each case, the sketch in the third column has two consecutive vertical solid edges in the rightmost coordinate. It is the higher of these two edges that is the left vertical edge in the junction plaquette. The combined length of the journeys that do not use this edge between the endpoints of this edge and the endpoints of the union of the dotted line segments is seen by inspection to equal $7 + r$. Thus the post-surgical chemical distance around γ between these two vertices of the junction plaquette is $n + 7$.

We want to establish a similar assertion concerning the journey around P. The Madras join operation involves the half-circle rotation of the right element P about Y; the identical surgery on the resulting polygon; and then a further half-circle rotation about Y of the post-surgical polygon. Since the form of surgery is the same, the preceding analysis is applicable, and we obtain Lemma 3.4(2).

(3). The polygonal component Q of $M(\gamma, P)\Delta J$ lies in the union $C_1 \cup C_2$, where C_1 is the part of the strip $\mathbb{R} \times [y(Y) - 1, y(Y) + 1]$ on or to the right of the right vertical edge in the junction plaquette J, and C_2 is the translate of P used in the Madras join. As such, no translation of Q to the right may intersect γ. The walk component γ' of $M(\gamma, P)\Delta J$ lies in the union of γ and certain edges in $\mathbb{R} \times [y(Y) - 1, y(Y) + 1]$ on or to the left of the left vertical edge in the junction plaquette. Thus, nor may a translation of Q to the right intersect γ'.

(4). Any deformation to γ under the surgery $\gamma \to M(\gamma, P)$ occurs in the strip $\mathbb{R} \times [y_{\min}(P) - 2, y_{\max}(P) + 2]$. Thus the joining of P and P' to γ is a commutative operation.

(5). It is perhaps helpful to begin with an example of how the presurgical data is not uniquely specified given the output including the location of the junction plaquette. Suppose that γ in a certain locale makes a sequence of down moves until a vertex z; then makes one left move; and then continues with down moves. As [Ham18, Figure 1] indicates, right-attachment of a polygon with $Y = z$ in Case *Ib*, or with $Y = z - e_2$ in Case *IIb*, leads to outcomes in the third column of the figure that are indistinguishable up to affine shift. In this example, then, two presurgical left-inputs are compatible with the output including the junction plaquette. Inspection of cases shows that there are never more than two inputs on the left compatible with this data. The same considerations apply on the right, so that Lemma 3.4(5) is obtained.

(6). Similarly to the derivation of Lemma 3.4(2), [Ham18, Figure 1] makes it convenient to argue for a symmetric statement. Suppose instead that the polygon P is on the left, and

another polygon is being joined to its right. Introduce the vertices $\mathsf{SW}(P)$ and $\mathsf{NW}(P)$ in accordance with Lemma 3.4(6)'s notation, and specify the left-side of P as the collection of vertices in P encountered on a clockwise journey between this pair of vertices. It is enough to demonstrate that the left-side of P enters the joined output without modification by surgery, because Lemma 3.4(6) is an assertion symmetrical to this. Consult [Ham18, Figure 1]. A moment's thought shows that no black edge depicted in the second column may abut a vertex in P's left side. Since the edges removed in surgery, which are the dotted edges in the third column, are a subset of such edges, the sought assertion is demonstrated.

(7). The edges removed from γ in the formation of $M(\gamma, P)$ are the dotted edges in the third column of [Ham18, Figure 1]. There are at most two such edges in each case.

(8). The walk $M(\gamma, P)$ is formed by the removal of either one edge, or two consecutive edges from γ, and the insertion of a new walk that connects the two endpoints that arise from this removal. The inserted piece has length either $m + 17$ or $m + 18$ because $M(\gamma, P)$ has length $n + m + 16$. $\qquad\square$

3.2.3. *Notation for the Madras join in applications.* Let $\gamma \in \mathsf{SAW}_n$ and $p \in \mathsf{SAP}_m$. Recall from Section 3.1 that the polygon p is depicted as containing the origin as the leftmost point whose vertical coordinate attains p's line-width. In our application, we will wire p into γ after translating p vertically; after all, it is the entropic benefit among those translations that we seek to exploit. We now set up the notation to be used for the Madras join in this application.

Fix $j \in \mathbb{Z}$ with
$$y_{\min}(\gamma) + \mathsf{h}(p) + 3 \leq j \leq y_{\max}(\gamma) - \mathsf{h}(p) - 3 \,.$$
We write $\mathsf{MJ}_j(\gamma, p) \in \mathsf{SAW}_{n+m+16}$ for the Madras join $M(\gamma, p + je_2)$ of γ with the copy of p vertically translated by j units. The condition on j ensures that the hypothesis (3.3) is met.

4. Hammersley-Welsh with polygon insertion: reducing to a key estimate

We begin to implement the plan laid out in Section 2.3 to prove Theorem 1.2. The present section states but does not prove Proposition 4.4, an upper bound on the number of horizontally confined half-space walks with many branches. This estimate is at the heart of the proposed plan. The section contains the proof of Theorem 1.2 that invokes this proposition. In the next Section 5, we give the proof of this key estimate.

We start the rigorous implementation by recalling from the overview the cast of characters $(\varepsilon, \delta, \alpha, u, m)$, the roles of this quintet's members and the conditions that it is necessary to impose on them.

- The sought improvement in the exponent of the Hammersley-Welsh bound is ε, so that $n^{1/2-\varepsilon}$ will replace $n^{1/2}$;

- the line-width of inserted polygons will be at least u, a quantity of order n^δ;

- a lower bound on the number of wide polygons, suitable for insertion, is expressed for $\mathbb{Z}^2$ in terms of the exponent α in Proposition 3.1;

- and all inserted polygons have a common length, namely m.

Of these five positive parameters, the first three are exponents on which we impose the conditions that

$$2\varepsilon \le \delta < 1/2 - \varepsilon \qquad \text{and} \qquad (4.1)$$

$$\varepsilon + \alpha\delta < 1/2. \qquad (4.2)$$

The polygon line-width n^δ is chosen, in the lower bound of (4.1), to be at least the horizontal near self-touch distance $n^{2\varepsilon}$ between consecutive crossings. There is no use in strict inequality, and so we choose $\delta = 2\epsilon$, as we did in the plan in Section 2.3. The upper bound of (4.1) is a non-degeneracy condition, which is not restrictive. The vital condition (4.2) coincides with the upper bound on ε, noted after (2.8): it ensures that a given polygon insertion is helpful, with the entropic gain in the resulting outgoing arrows being larger than the cost of purchasing the polygon.

By imposing (4.1) and (4.2), we are developing rather directly the suggestion prompted by the heuristic discussion of Section 2.3. In fact, such a choice for $(\varepsilon, \delta, \alpha)$ as these conditions impose will permit the derivation of an upper bound on the number $|\mathsf{HSW}_n|$ of length n half-space walks. In order to use this information to obtain Theorem 1.2, namely a similar bound on $|\mathsf{SAW}_n|$, we will invoke (2.1). To be able to do so, we will need control on $|\mathsf{HSW}_n|$ not merely for infinitely many n, but for infinitely many long ranges of consecutive n. To achieve such a bound, we will need to strengthen the conditions (4.1) and (4.2) on the triple $(\varepsilon, \delta, \alpha)$.

Reflecting these considerations, the next proposition has two parts. The first part articulates the notion that polygon abundance, which limits the cost of polygon purchase to which we alluded a few moments ago, secures an improvement in the Hammersley-Welsh bound for half-space walks. The second part shows that such an improvement is also secured for all self-avoiding walks. As such, it is in this second part that a strengthening of the conditions (4.1) and (4.2) is hypothesised.

We encourage the reader to focus on the elaboration of the heuristic guide and thus on a choice of parameters satisfying (4.1) and (4.2). In this regard, note that, when $n \in \mathbb{N}$ verifies (4.3), we formally specify the final two members (u, m) of the parameter quintet given the first three $(\varepsilon, \delta, \alpha)$.

Proposition 4.1. (1) *Fix a triple* $(\varepsilon, \delta, \alpha)$ *of positive exponents satisfying* (4.1) *and* (4.2). *There exists a constant* $C = C(\varepsilon) > 0$ *such that, whenever* $n \in \mathbb{N}$ *satisfies the condition*

$$\text{there exist } u \in [n^\delta, 2n^\delta] \text{ and } m \le 4^{-1}u^2, \text{ for which } |\mathsf{WSAP}^u_m| \ge n^{-\alpha\delta}\mu^m, \qquad (4.3)$$

we have that

$$|\mathsf{HSW}_n| \le C\exp(21n^{1/2-\varepsilon}\log n) \cdot \mu^n. \qquad (4.4)$$

(2) *Fix a triplet* $(\varepsilon, \delta, \alpha)$ *of positive exponents satisfying the stronger conditions*

$$4\varepsilon \le \delta < 1/2 - \varepsilon \quad \text{and} \quad \varepsilon + 2\alpha\delta < 1/2. \qquad (4.5)$$

Then there exists a constant $C = C(\varepsilon) > 0$ *such that, whenever* $n \in \mathbb{N}$ *verifies* (4.3), *any* $\ell \in [\![n^{3/2}, n^2]\!]$ *satisfies*

$$|\mathsf{SAW}_\ell| \le \exp(C\ell^{1/2-\varepsilon}\log \ell) \cdot \mu^\ell.$$

Proposition 4.1(1) only has value if α, δ and ε are such that there are infinitely many values n satisfying (4.3). Since $\mathsf{WSAP}_m^u \subset \mathsf{SAP}_m$, and it is expected that $|\mathsf{SAP}_m| = m^{-5/2+o(1)}\mu^m$ for m even (see [Nie82]), the best upper bound on $|\mathsf{HSW}_n|$ that the present implementation of our method may achieve takes $\varepsilon = 1/12 + o(1)$. This is the value of ε in (4.2) consistent with $\alpha = 5/2$. That said, aspects of the method other than the α-value could be varied: see Section 5.5.

Our main result for $\mathbb{Z}^2$ follows readily from Proposition 4.1(2) and the demonstration of polygon abundance in Section 3.

Proof of Theorem 1.2. Fix $\varepsilon_0 < \varepsilon < \frac{1}{466}$. We will prove the theorem for ε_0. With a view to choosing the triple $(\varepsilon, \delta, \alpha)$ to satisfy the hypotheses of Proposition 4.1(2), we set $\delta = 4\varepsilon$. A value of α may then be selected so that the hypothesis $\alpha > 29$ of Proposition 3.1 holds, while the bound $\varepsilon(1 + 8\alpha) < 1/2$ of (4.5) is also met.

Proposition 3.1 ensures that there are infinitely many values of (u, m) satisfying $\left|\mathsf{WSAP}_m^u\right| \geq u^{-\alpha}\mu^m$. For any such value of u, choose n such that $u/2 \leq n^\delta \leq u$. Then Proposition 4.1(2) implies that $|\mathsf{SAW}_\ell| \leq \exp(C\ell^{1/2-\varepsilon}\log\ell) \cdot \mu^\ell$ whenever $\ell \in [\![n^{3/2}, n^2]\!]$. This in turn implies that $|\mathsf{SAW}_\ell| \leq \exp(\ell^{1/2-\varepsilon_0}) \cdot \mu^\ell$ for such ℓ when n is supposed to be high enough. Taking $j = n^{3/2}$, we obtain the latter, stronger, assertion of Theorem 1.2. $\qquad\square$

We now prepare to give the proof of Proposition 4.1, which is to say, to rigorously implement the plan from Section 2.3. To that end,

fix a triple $(\varepsilon, \delta, \alpha)$ of positive exponents satisfying (4.1) **and** (4.2)**.**

The plan began by roughly explaining why the walks in the domain of the Hammersley-Welsh map Ψ might be chosen to be a certain subset of HSW_n. This subset was called MHSW_n. As we turn to precise specification, we mention that we will be working with a more general object, MHSW_k^n.

Here, we are fixing parameters $k, n \in \mathbb{N}$ with $k \leq n$, and specify MHSW_k^n to be a subset of the space HSW_k of length-k half-space walks. Elements of MHSW_k^n will adhere nonetheless to the informal description offered at the start of Section 2.3 of members of MHSW_n: they are horizontally confined walks with many branches.

By horizontally confined walks, we mean the elements of a set $\mathsf{HorConfHSW}_k^n$ defined to be the subset of HSW_k whose elements are contained in the vertical strip $[-n^{1/2+\varepsilon}, n^{1/2+\varepsilon}] \times \mathbb{Z}$.

The set $\mathsf{FewBranchHSW}_k^n$ of length-k half-space walks with few branches consists of those members of HSW_k with fewer than $7n^{1/2-\varepsilon}$ elements in their branch decomposition. We then define MHSW_k^n to be $\mathsf{HorConfHSW}_k^n \setminus \mathsf{FewBranchHSW}_k^n$.

The implementation of our plan comprises three tasks: the first short, the second substantial, and the third again short. The first task ascertains that, for the purpose of bounding $|\mathsf{HSW}_n|$, attention may indeed by focussed on the sets MHSW_k^n; the second provides an upper bound on the cardinality of these sets, with which the proof of Proposition 4.1(1) may easily be presented; and the third proves Proposition 4.1(2).

The next two lemmas carry out the first task. The following Proposition 4.4 states the outcome of the second. After these statements, we prove Proposition 4.1(1) by invoking the three results. Proofs of the two lemmas then follow. The second task, namely the proof of Proposition 4.4, appears in Section 5; naturally, it is there that that the central ideas in our plan are enacted. The present Section 4 ends with the third task, namely the proof of Proposition 4.1(2).

Lemma 4.2 (Restricting to horizontally confined walks)**.** *For $n \geq 4$, we have that*

$$|\mathsf{HSW}_n| \cdot \mu^{-n} \leq 3\mu \exp(2n^{1/2-\varepsilon} \log n) \Big(\sum_{k=0}^{n} |\mathsf{HorConfHSW}_k^n| \cdot \mu^{-k} \Big)^2 .$$

Lemma 4.3 (Despatching walks with few branches)**.** *For n high enough and for $k \leq n$,*

$$\left| \mathsf{FewBranchHSW}_k^n \right| \leq \exp(8n^{\frac{1}{2}-\varepsilon} \log n) \, \mu^k . \tag{4.6}$$

Proposition 4.4 (The key bound)**.** *For some positive constant C, for $n \in \mathbb{N}$ verifying* (4.3) *and any $k \leq n$,*

$$\left| \mathsf{MHSW}_k^n \right| \leq C\mu^k . \tag{4.7}$$

Proof of Proposition 4.1(1). Since $\mathsf{HorConfHSW}_k^n \subseteq \mathsf{FewBranchHSW}_k^n \cup \mathsf{MHSW}_k^n$, Lemma 4.3 and Proposition 4.4 imply that, when $n \in \mathbb{N}$ verifying (4.3) is high enough, and $k \leq n$,

$$|\mathsf{HorConfHSW}_k^n| \leq |\mathsf{FewBranchHSW}_k^n| + |\mathsf{MHSW}_k^n| \leq \exp(9n^{\frac{1}{2}-\varepsilon} \log n) \, \mu^k .$$

Using Lemma 4.2, we deduce that, for such n,

$$|\mathsf{HSW}_n| \, \mu^{-n} \leq \exp(21n^{\frac{1}{2}-\varepsilon} \log n) .$$

Multiplying this right-hand side by a suitable constant $C(\varepsilon) > 0$ permits us to discard the condition that $n \in \mathbb{N}$ verifying (4.3) is supposed to be high enough. We have obtained Proposition 4.1(1). $\qquad\square$

Proof of Lemma 4.2. Let $\gamma \in \mathsf{HSW}_n$. Let ℓ denote the number of elements in the branch decomposition of γ whose heights exceed $n^{1/2+\varepsilon}$, and note that $\ell \leq n^{1/2-\varepsilon}$. We may consider the injective map Ξ defined on HSW_n and specified by

$$\Xi : \gamma \mapsto \big(\gamma_{[0,a_1]}, \dots, \gamma_{[a_{\ell-1},a_\ell]}, \gamma_{[a_\ell,n]} \big) ,$$

where $\gamma_{[0,a_1]}, \dots, \gamma_{[a_{\ell-1},a_\ell]}$ are the first ℓ branches of γ. See Figure 4 for an illustration of the decomposition of γ under Ξ.

The final component, $\gamma_{[a_\ell,n]}$, is a half-space walk, contained in a horizontal strip of height $n^{1/2+\varepsilon}$. Let s denote the highest index of a point of maximal x-coordinate of $\gamma_{[a_\ell,n]}$. After right-angled rotation and translation, $\gamma_{[s,n]}$ becomes an element of $\mathsf{HorConfHSW}_{n-s}^n$, while a suitable such operation on $\gamma_{[a_\ell,s]}$ may be followed by the prefixing of a vertically oriented edge so that an element of $\mathsf{HorConfHSW}_{s-a_\ell+1}^n$ results. The range of Ξ is thus seen to be a subset of a set in bijection with

$$\bigcup_{\ell,a,s} \mathsf{B\Pi}_{a,\ell} \times \mathsf{HorConfHSW}_{s-a+1}^n \times \mathsf{HorConfHSW}_{n-s}^n ,$$
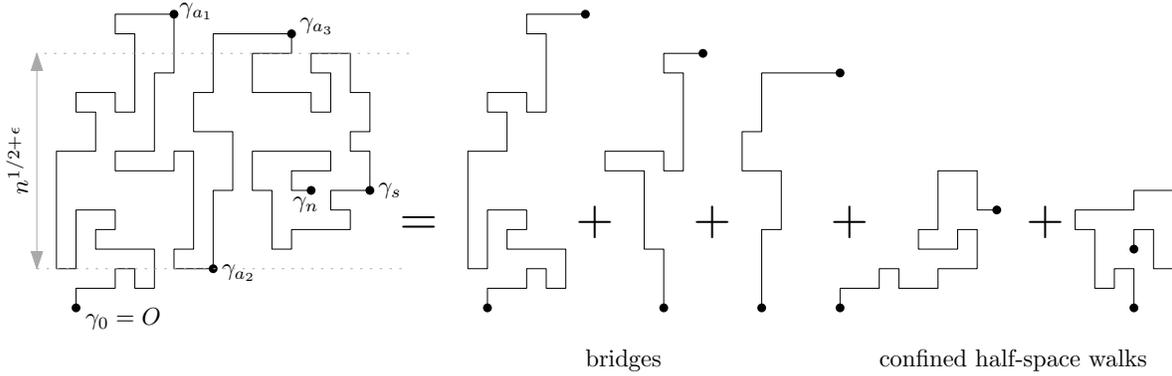
FIGURE 4. A half-space walk is decomposed into three bridges of height larger than $n^{1/2+\varepsilon}$ and two confined half-space walks.

where the union is over $0 \leq \ell \leq n^{1/2-\varepsilon}$ and $0 \leq a \leq s \leq n$. By Lemma 2.1(1), the displayed set has cardinality at most

$$\sum_{\ell=0}^{\lfloor n^{1/2-\varepsilon} \rfloor} \sum_{a=0}^{n} \sum_{s=a}^{n} \mu^a a^\ell \cdot \left| \mathsf{HorConfHSW}_{s-a+1}^n \right| \cdot \left| \mathsf{HorConfHSW}_{n-s}^n \right|$$

$$\leq \left(n^{1/2-\varepsilon} + 1 \right) n^{n^{1/2-\varepsilon}} \mu^{n+1}$$

$$\times \sum_{a=0}^{n} \sum_{s=a}^{n} \mu^{-(s-a+1)} \left| \mathsf{HorConfHSW}_{s-a+1}^n \right| \cdot \mu^{-(n-s)} \left| \mathsf{HorConfHSW}_{n-s}^n \right|$$

$$\leq 3 n^{2n^{1/2-\varepsilon}} \mu^{n+1} \left(\sum_{k=0}^{n} \mu^{-k} \left| \mathsf{HorConfHSW}_k^n \right| \right)^2 .$$

The final inequality is due to $\left| \mathsf{HorConfHSW}_{k+1}^n \right| \leq 3 \left| \mathsf{HorConfHSW}_k^n \right|$, a bound seen by the removal of the last edge from each element of $\mathsf{HorConfHSW}_{k+1}^n$; to $x+1 \leq e^x$ with $x = n^{1/2-\varepsilon}$; and to $n \geq 4$ in the guise $e^x \leq n^x$. Since Ξ is injective, we obtain Lemma 4.2. $\square$

Proof of Lemma 4.3. The Hammersley-Welsh map Ψ from Lemma 2.2 injectively maps $\mathsf{FewBranchHSW}_k^n$ to $\bigcup_{j=1}^{\lfloor 7n^{1/2-\varepsilon} \rfloor} \mathsf{B\Pi}_{k,j}$, the set of bridge-lists with at most $7n^{1/2-\varepsilon}$ terms, which terms have combined length k. By Lemma 2.1(1), the latter set has cardinality at most $\mu^k \sum_{j=1}^{\lfloor 7n^{1/2-\varepsilon} \rfloor} k^j \leq \mu^k \cdot 7n^{1/2-\varepsilon} \cdot n^{7n^{1/2-\varepsilon}}$. By supposing $n \in \mathbb{N}$ to be high enough, we obtain the bound in Lemma 4.3. $\square$

Next we prove Proposition 4.1(2) For $\eta, \beta > 0$, $n \in \mathbb{N}$ is said to satisfy Assertion $A(\eta, \beta)$ when

there exist $u \in [n^\eta, 2n^\eta]$ and $m \in \mathbb{N}$ with $m \leq 4^{-1}u^2$ for which $|\mathsf{WSAP}_m^u| \geq n^{-\beta\eta} \mu^m$.

Note that $n \in \mathbb{N}$ satisfies Assertion $A(\delta, \alpha)$ precisely when n verifies (4.3).

Lemma 4.5 (Stability of polygon abundance under changes in length). *Suppose that $n \in \mathbb{N}$ satisfies Assertion $A(\delta, \alpha)$. Then any $\ell \in [\![n, n^2]\!]$ satisfies Assertion $A(\delta', 2\alpha)$ for some $\delta' = \delta'(\ell) \in [\delta/2, \delta]$.*

Proof. Pick $\ell \in [\![n, n^2]\!]$. Assertion $A(\delta, \alpha)$ for n entails the existence of u and $m \leq 4^{-1}u^2$ such that $u \in [n^\delta, 2n^\delta]$ and $|\mathsf{WSAP}_m^u| \geq n^{-\alpha\delta}\mu^m$. The first condition can be rewritten as $u \in [\ell^{\delta'}, 2\ell^{\delta'}]$ with $\delta' = \delta\frac{\log n}{\log \ell} \in [\delta/2, \delta]$. Since $\ell \geq n$ and $2\alpha\delta' \geq \alpha\delta$, the second implies $|\mathsf{WSAP}_m^u| \geq \ell^{-\alpha 2\delta'}\mu^m$. Overall, ℓ satisfies Assertion $A(\delta', 2\alpha)$.

$\square$

Proof of Proposition 4.1(2). Assume the fixed values $\varepsilon, \delta, \alpha$ satisfy the stronger condition (4.5) rather than merely (4.1) and (4.2). Fix $n \in \mathbb{N}$ that satisfies (4.3) and consider $n^{3/2} \leq \ell < n^2$; the case where $\ell = n^2$ follows from $|\mathsf{SAW}_{n^2}| \leq 4|\mathsf{SAW}_{n^2-1}|$.

Applying (2.1) and then the Cauchy-Schwarz inequality gives

$$|\mathsf{SAW}_\ell| \leq \sum_{k=0}^{\ell} |\mathsf{HSW}_{k+1}|\,|\mathsf{HSW}_{\ell-k}| \leq \sum_{k=0}^{\ell+1} |\mathsf{HSW}_k|^2 \qquad (4.8)$$

The classical Hammersley-Welsh bound Theorem 1.1 implies that, for all $k \leq n$,

$$|\mathsf{HSW}_k| \leq \exp(Ck^{1/2})\mu^k, \qquad (4.9)$$

for some constant C independent of k and n. Furthermore, by Lemma 4.5, any $k \in [\![n, n^2]\!]$ satisfies Assertion $A(\delta', 2\alpha)$ for some $\delta' \in [\delta/2, \delta]$. The strengthened hypothesis (4.5) on $(\varepsilon, \delta, \alpha)$ ensures that the triple $(\varepsilon, \delta', 2\alpha)$ verifies the conditions (4.1) and (4.2). Thus, Proposition 4.1(1) gives

$$|\mathsf{HSW}_k| \leq \exp(Ck^{1/2-\varepsilon}\log k)\mu^k, \qquad (4.10)$$

for all $k \in [\![n, n^2]\!]$, where $C = C(\varepsilon) > 0$ is some constant (which we can assume to be larger than the constant in (4.9)).

Applying (4.9) for $k \leq n$ and (4.10) for $n < k \leq \ell + 1$ to (4.8), we obtain

$$|\mathsf{SAW}_\ell| \leq (\ell + 2)\exp(2C\ell^{1/2-\varepsilon}\log \ell)\mu^\ell$$

since $n^{1/2} \leq \ell^{1/3} \leq \ell^{1/2-\varepsilon}$, where $\varepsilon < 1/6$ is imposed by (4.5). The desired bound is thus obtained for $\ell \in [\![n^{3/2}, n^2 - 1]\!]$ by a suitable adjustment of the constant C. $\square$

5. Proving the key bound concerning confined walks with many branches

Here we prove Proposition 4.4, implementing the essential aspects of the polygon insertion plan. For the whole section, the triple $(\varepsilon, \delta, \alpha)$ satisfying (4.1) and (4.2) is fixed.

Recall from the conclusion of the heuristically presented plan that we will endeavour to alter a horizontally confined walk γ that has a high number $r = r(\gamma)$ of branches by the surgical insertion of κr polygons. Here, the parameter $\kappa > 0$ will be of unit order, chosen so that $\kappa r \in \mathbb{N}$. (This membership of $\mathbb{N}$ will be straightforward to arrange: recall that $r \geq 7n^{1/2-\varepsilon}$ when $\gamma \in \mathsf{MHSW}_k^n$.) The value of $\kappa > 0$ will be fixed when needed; the consideration that determines its value has been explained at the end of the guide in Section 2.3, though the precise condition will be slightly modified, for a rather inconsequential reason.

5.1. **Locations for polygon insertion.** Let $\gamma \in \mathsf{MHSW}_k^n$. In this subsection, we specify a counterpart Rect of the rectangle R through which, as we argued in Section 2.3, so many vertical crossings would be made by such walks as γ; we specify a set $\mathsf{Join}(\gamma)$ of indices for join locations at the side of these vertical crossings at which a polygon may be inserted so as to overlap the adjacent crossing; and we record the set $\mathsf{JoinLocLists}(\gamma)$ of length-κr lists of such sites. Section 5.1 ends with a lower bound, Lemma 5.2, on the prospective forward arrow number, $\big| \mathsf{JoinLocLists}(\gamma) \big|$, of the multi-valued map Φ that we aim to construct.

Let (n, u, m) verify (4.3). We denote by $\big(\gamma_{[0,a_1]}, \gamma_{[a_1,a_2]}, \ldots, \gamma_{[a_{r-1},a_r]} \big)$ the branch decomposition of γ. Let $\gamma_{[a_{t-1},a_t]}$ be the last branch in the decomposition of γ whose height exceeds $n^{1/2-\varepsilon}$. Since the branch heights form a strictly decreasing sequence, we have $t \geq r - n^{1/2-\varepsilon}$; and, since $r \geq 7n^{1/2-\varepsilon}$, we see that t has the same order as r.

Write $j_{\max} = y_{\max}(\gamma_{[a_{t-1},a_t]})$ and $j_{\min} = y_{\min}(\gamma_{[a_{t-1},a_t]})$, so that $j_{\max} - j_{\min} > n^{1/2-\varepsilon}$. The branches $\gamma_{[0,a_1]}, \ldots, \gamma_{[a_{t-1},a_t]}$ traverse vertically the rectangle

$$\mathsf{Rect} := [-n^{1/2+\varepsilon}, n^{1/2+\varepsilon}] \times [j_{\min}, j_{\max}],$$

which permits us to order them from left to right. More precisely, a given such branch may traverse this rectangle several times, and these traversals may be ordered from left to right. The leftmost traversals of the t branches are themselves ordered from left to right, and it is in this order that we record the branches $\phi^{(1)}, \ldots, \phi^{(t)}$. It is these objects that we call the *tall branches* of γ.

For $l \in [\![1, t]\!]$, denote by $z_{l,j}$ the rightmost among the vertices of $\phi^{(l)}$ whose y-coordinate equals j. Now let $(\ell, j) \in [\![1, t-1]\!] \times \mathbb{Z}$. The pair (ℓ, j) is called a *viable join index* if

- $j_{\min} + 16u + 3 \leq j \leq j_{\max} - 16u - 3$,

- and $x(z_{\ell+1,j}) - x(z_{\ell,j}) \leq u$.

Let $\mathsf{Join}(\gamma)$ be the set of viable join indices of γ. The next lemma indicates why we call such pairs (ℓ, j) viable: when a wide polygon is inserted on the right of $\phi^{(\ell)}$ at height j, it will overlap the right-adjacent tall branch $\phi^{(\ell+1)}$ of γ. Recall from Subsection 3.2.3 that the subscript in the notation MJ_j indicates the vertical displacement to which the right polygon is subjected before Madras joining is undertaken.

Lemma 5.1. *Suppose that $n \geq 9^{1/(1-\varepsilon)}$. Let (ℓ, j) be a viable join location for γ. Then, for any $p \in \mathsf{WSAP}_m^u$, there exists a vertex that belongs to $\mathsf{MJ}_j\big(\phi^{(\ell)}, p\big)$ but not to $\phi^{(\ell)}$ and that also belongs to $\phi^{(\ell+1)}$.*

Proof. By assumption on (p, j), the polygon $P := p + (0, j)$ is contained in the region $\mathsf{Strip} := \mathbb{R} \times [j_{\min} + 3, j_{\max} - 3]$. The pair $\big(\phi^{(\ell)}, P \big)$ is thus seen to verify (3.3), so Lemma 3.4(1) permits the construction of $\mathsf{MJ}_j\big(\phi^{(\ell)}, p\big) = M\big(\phi^{(\ell)}, P\big)$.

In the Madras join of γ and P, let $(i, 0)$ denote the vector of horizontal translation which relocates P before surgery takes place. This translate of P occupies vertices of y-coordinate j at locations, including (i, j), strictly to the right of such locations, including $z_{\ell,j}$, for vertices in $\phi^{(\ell)}$; thus, $i > x(z_{\ell,j})$. The polygon p is wide, and thus $P + (i, 0)$ attains its line-width of at least u at height j. Denoting by w the rightmost point at height j in this polygon, we thus

see that w lies on the semi-infinite line $z_{\ell,j} + [u, \infty) \times \{0\}$. Since $z_{\ell+1,j}$ lies at a distance of at most u directly to the right of $z_{\ell,j}$, the point w lies directly to the right of $z_{\ell+1,j}$.

The branch $\phi^{(\ell+1)}$ separates Strip into several regions, of which two are infinite. Call these two the left and right regions of Strip. By definition, the half line $z_{\ell+1,j} + (0, \infty) \times \{0\}$ is contained in the right region of Strip. The point w lies on the right-side of $P + (i, 0)$ in the sense of Lemma 3.4(6); this result thus implies that $\mathsf{MJ}_j(\phi^{(\ell)}, p)$ contains w, which is a point in the right region of Strip. However, $\phi^{(\ell)} \cap$ Strip is contained in the left region of Strip. Note that $\mathsf{MJ}_j(\phi^{(\ell)}, p)$ has non-empty intersection with $\phi^{(\ell)} \cap$ Strip, since, by Lemma 3.4(7), $\mathsf{MJ}_j(\phi^{(\ell)}, p)$ contains all but at most two edges in $\phi^{(\ell)}$, while $\phi^{(\ell)}$ contains at least $j_{\max} - j_{\min} - 6 \geq n^{1-\varepsilon} - 6 \geq 3$ edges in view of the assumption that $n \geq 9^{1/(1-\varepsilon)}$. We see then that $\mathsf{MJ}_j(\phi^{(\ell)}, p)$ intersects $\phi^{(\ell+1)}$ inside Strip. Any vertex in that intersection verifies the assertion made by Lemma 5.1, because such a vertex cannot lie in $\phi^{(\ell)}$ in view of the disjointness of $\phi^{(\ell)}$ and $\phi^{(\ell+1)}$. $\qquad\square$

A *join location list* for γ is a set list $\subset$ Join of cardinality κr with the property that if $(\ell_1, j) \in$ list and $(\ell_2, j') \in$ list are two viable join indices whose first components satisfy $|\ell_1 - \ell_2| \leq 1$, then $|j - j'| \geq 32u + 5$. (The term 'list' is used because we prefer to think of inserting polygons one at a time; note, however, that a join location list is an unordered set. Note also that the term 'join location list' is a misnomer: the elements of list merely *index* physical locations in $\mathbb{Z}^2$.) Let $\mathsf{JoinLocLists}(\gamma)$ be the set of join location lists for γ.

Set $\mathsf{MHSW}_k^n(r) = \{\gamma \in \mathsf{MHSW}_k^n : \gamma \text{ has } r \text{ branches}\}$.

Lemma 5.2. *If n is large enough, then $\gamma \in \mathsf{MHSW}_k^n(r)$ verifies*

$$|\mathsf{Join}(\gamma)| \geq \tfrac{1}{4} \, r \, n^{1/2-\varepsilon} \, . \tag{5.1}$$

By further increasing n if need be, we have that

$$|\mathsf{JoinLocLists}(\gamma)| \geq \exp\left(\kappa(1/2 - \varepsilon) r \log n - \kappa \log(8\kappa) r \right). \tag{5.2}$$

Proof. We start by proving (5.1). Recall that t denotes the number of γ's branches whose height exceeds $n^{1/2-\varepsilon}$. There are $(t-1)(j_{\max} - j_{\min} - 32u - 5)$ pairs (ℓ, j) with $1 \leq \ell < t$ and $j \in \mathbb{Z} \cap [j_{\min} + 16u + 3, j_{\max} - 16u - 3]$. If such a pair fails to be a viable join index, it is because the second condition in the definition of such an index is not met.

Fix such a pair (ℓ, j), so that indeed the second condition fails. Let L denote the elements of $\mathbb{Z}^2$ that are encountered strictly and directly to the right of $z_{\ell,j}$, up to and including $z_{\ell,j} + (u, 0)$. No vertex of the form $z_{\ell',j}$ lies in L. For this reason, we say that (ℓ, j) *blocks* the u elements of L. Observe that if $(\ell, j), (\ell', j') \in [1, t-1] \times [j_{\min} + 16u + 3, j_{\max} - 16u - 3]$ are two distinct pairs that are not viable join indices, then the sets of points that they block are disjoint. Moreover, all blocked points lie in $[-n^{1/2+\varepsilon}, n^{1/2+\varepsilon} + u] \times [j_{\min} + 16u + 3, j_{\max} - 16u - 3]$.

It follows that there are at most $u^{-1}(2n^{1/2+\varepsilon} + u + 1)(j_{\max} - j_{\min} - 32u - 5)$ pairs that are not viable join indices. Thus, we may bound from below the number of viable join indices:

$$
\begin{aligned}
\left|\mathsf{Join}(\gamma)\right| &\geq \left(t - 1 - 2u^{-1}n^{1/2+\varepsilon} - 1 - u^{-1}\right)(j_{\max} - j_{\min} - 32u - 5) \\
&\geq \left(r - 3n^{1/2-\varepsilon} - 3\right)(j_{\max} - j_{\min} - 32u - 5),
\end{aligned}
$$

where the latter inequality is due to $u \geq n^{2\varepsilon} \geq 1$ and $r - t \leq n^{1/2-\varepsilon}$. Since $r \geq 7n^{1/2-\varepsilon}$, the first term in the latter product is bounded below by $r/2$ when n is high enough. Since $u \leq 2n^\delta$, while $\delta < 1/2 - \varepsilon$ by (4.1), the second term in this product is bounded below by $2^{-1}n^{1/2-\varepsilon}$, provided that n is supposed high enough. This proves (5.1).

Suppose now that n is indeed large enough for (5.1) to hold. In order to count join location lists $\mathsf{list} \subset \mathsf{Join}(\gamma)$, we record such an element list, one item at a time. In view of Lemma 3.4(4), each choice of a viable join index disallows the prospect of later inclusion in the growing list for at most $3(32u + 5)$ other elements of $\mathsf{Join}(\gamma)$, where the factor of three is contributed because the relevant disjointness condition concerns pairs of indices whose first components differ by at most one from each other. Thus, we find that

$$
\begin{aligned}
|\mathsf{JoinLocLists}(\gamma)| \;\; &\geq \;\; \frac{1}{(\kappa r)!} \prod_{i=0}^{\kappa r - 1} \big(|\mathsf{Join}(\gamma)| - 3(32u + 5)i\big) \\
&\geq \;\; \frac{1}{(\kappa r)!} \big(8^{-1} r\, n^{1/2-\varepsilon}\big)^{\kappa r} \geq (2e)^{-1} e^{\kappa r} \big(8^{-1}\kappa^{-1} n^{1/2-\varepsilon}\big)^{\kappa r} \\
&\geq \;\; (2e)^{-1} \exp\big((1/2 - \varepsilon)\kappa r \log n - Cr\big),
\end{aligned}
$$

where $C = \kappa\big(\log(8\kappa) - 1\big)$. The second inequality is due to (5.1), and, since $u \leq 2n^\delta$, holds when n is high enough that $3(64n^\delta + 5)\kappa \leq 8^{-1}n^{1/2-\varepsilon}$. The third makes use of $h! \leq 2e\, h^h e^{-h}$ with $h = \kappa r$. Replacing C by $\kappa \log(8\kappa)$ and invoking again that n is high enough, we obtain Lemma 5.2. $\qquad\square$

5.2. Key properties of the multi-valued map Φ.
Henceforth, we let $r \in \mathbb{N}$ denote any *given* value satisfying $r \geq 7n^{1/2-\varepsilon}$. In order to bound above the number of elements of MHSW_k^n with r branches, we construct a multi-valued map

$$
\Phi : \mathsf{MHSW}_k^n(r) \times \big(\mathsf{WSAP}_m^u\big)^{\kappa r} \to \mathsf{B\Pi}_{k+\kappa r(m+16),r}\,,
$$

where the bridge-list set that is the range was specified in Section 2.1.

An element in the domain of Φ will be recorded in the form $(\gamma;\, p_1, \ldots, p_{\kappa r})$. The arrows under Φ outgoing from this element will be indexed by the set $\mathsf{JoinLocLists}(\gamma)$ of length-κr join location lists.

Next we state a key property that Φ will be constructed to verify. We then use this property to close out the proof of Proposition 4.4. In two further subsections, Φ is constructed, and the key property is proved.

Lemma 5.3 (The multi-valued map Φ is close to injective). *When $m \geq 17$,*

- *the minimum number m_0 of arrows of Φ outgoing from any element of its domain is at least $\exp\big(\kappa r\big[(1/2 - \varepsilon) \log n - \log(8\kappa)\big]\big)$;*

- *and the maximum number M_0 of arrows incoming to any element in its range is at most $L^{\kappa r}$, where $L = 12$.*

Proof of Proposition 4.4. It suffices to prove the statement for n large enough, since smaller values may be incorporated by adjusting the value of the constant C. Fix n satisfying

(4.3). By the multi-valued map Lemma 1.5, (4.3) and Lemma 5.3,

$$\big|\mathsf{WHSW}_k^n(r)\big| \;\leq\; \Big(\frac{L}{n^{-\delta\alpha}\mu^m}\exp\big(-(\tfrac{1}{2}-\varepsilon)\log n + \log(8\kappa)\big)\Big)^{\kappa r}\,\big|\mathsf{B\Pi}_{k+\kappa r(m+16),r}\big|\,.$$

(The hypothesis in Lemma 5.3 that $m \geq 17$ is satisfied because m, being the length of a polygon whose line-width is at least u, must satisfy $m \geq u$; while $u \geq n^{2\varepsilon}$; and $n \geq 17^{1/(2\varepsilon)}$, since we are permitted to suppose that n is sufficiently high.) Using Lemma 2.1(1) to bound the right-hand side and the crude bound $k + \kappa r(m+16) \leq n^{1+4\varepsilon}$, we find that

$$\big|\mathsf{WHSW}_k^n(r)\big|\,\mu^{-k} \;\leq\; \exp\Big(-r\big[\kappa(\tfrac{1}{2}-\varepsilon-\delta\alpha)-(1+4\varepsilon)\big]\log n + rC'\Big), \qquad (5.3)$$

where $C' = \kappa\log(8\kappa) + \kappa\log L + 16\kappa\log\mu = \kappa\log\big(96\kappa\mu^{16}\big)$, since $L = 12$ for $\mathbb{Z}^2$. By (4.2), the coefficient above for κ is strictly positive. Set $\kappa = \lceil\frac{1+4\varepsilon}{1/2-\varepsilon-\delta\alpha}+1\rceil$; the rounding up is a device that ensures that $\kappa r \in \mathbb{N}$ in view of $r \in \mathbb{N}$. Then (5.3) reads

$$\big|\mathsf{WHSW}_k^n(r)\big|\,\mu^{-k} \;\leq\; \exp(-r\log n + rC') \leq n^{-1}\,,$$

the latter by assuming that n is large enough. Then

$$\big|\mathsf{WHSW}_k^n\big|\,\mu^{-k} \;=\; \sum_{r=\lceil 7n^{1/2-\varepsilon}\rceil}^{n}\big|\mathsf{WHSW}_k^n(r)\big|\,\mu^{-k} \leq 1\,,$$

and Proposition 4.4 is proved. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

5.3. **Construction of Φ.** An element in the domain of Φ may be recorded in the form $(\gamma\,;\,p_1,\ldots,p_{\kappa r})$, and an arrow under Φ outgoing from this element may be recorded as $(\gamma\,;\,p_1,\ldots,p_{\kappa r}\,;\,\mathsf{list})$, where $\mathsf{list} \in \mathsf{JoinLocLists}(\gamma)$. In a notational device that is intended to draw attention to the proposed surgical locations of the polygons, we instead denote the respective polygons $p_1,\ldots,p_{\kappa r}$ in the form $\big(p_{(\ell,j)} : (\ell,j) \in \mathsf{list}\big)$, where the elements of list are ordered by the value of ℓ, and by the value of j when the value of ℓ is shared between elements.

Recall from Section 5.1 that γ's tall branches $\phi^{(1)},\ldots,\phi^{(t)}$ are ordered according to the left-right order of the leftmost crossings of Rect that they make.

For $\ell \in \{1,\ldots,t-1\}$, define $\tilde{\phi}^{(\ell)}$ to be the Madras join of $\phi^{(\ell)}$ with all polygons of the form $p_{\ell,j}$ with $(\ell,j) \in \mathsf{list}$ at the corresponding heights j. Formally, if the pairs $(\ell,j) \in \mathsf{list}$ are those with values $j_1,\cdots,j_J$, where $J = J(\ell)$, then

$$\tilde{\phi}^{(\ell)} = \mathsf{MJ}_{j_J}\Big(\cdots\mathsf{MJ}_{j_2}\big(\mathsf{MJ}_{j_1}(\phi^{(\ell)},p_{(\ell,j_1)}),p_{(\ell,j_2)}\big)\cdots\Big),p_{(\ell,j_J)}\Big).$$

Although this specification of the iterated Madras joins is perhaps notationally cumbersome, the joins are commutative in view of the condition imposed on list by its membership of $\mathsf{JoinLocLists}$. Indeed, since distinct j-values for given ℓ differ by at least $32u + 5$, while the intervals of y-coordinates adopted by the translations used in surgery of the polygons $\{p_{(\ell,j_i)} : 1 \leq i \leq J\}$ each have length at most $16u$, Lemma 3.4(3) ensures this commutativity. Moreover, all of these j-values are contained in $[j_{\min}+3, j_{\max}-3]$, so that (3.3) holds for each proposed join. Hence, Lemma 3.4(1), or its counterpart concerning vertical reflection, implies that $\tilde{\phi}^{(\ell)}$ inherits from $\phi^{(\ell)}$ the status of bridge or vertical reflection of one. The length of $\tilde{\phi}^{(\ell)}$ is the sum of that of $\phi^{(\ell)}$ and the quantity $J(m+16)$, since each insertion contributes $m+16$
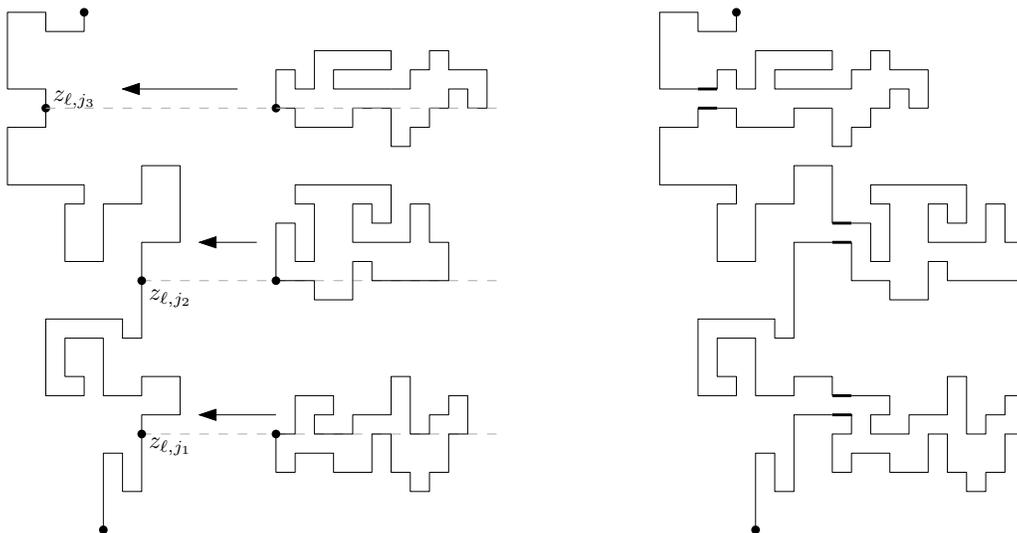
FIGURE 5. A toy depiction of the Madras join of a branch $\phi^{(\ell)}$ and three polygons $p_{(\ell,j_1)}$, $p_{(\ell,j_2)}$ and $p_{(\ell,j_3)}$ at respective heights j_1, j_2 and j_3. The edges introduced in surgery have been represented merely by two thick edges in each case.

edges, with m due to the polygon's length and a net gain of sixteen arising from surgery. Figure 5 illustrates the several surgical attachments that a given tall branch may endure.

The arrow under Φ that we index by $(\gamma\,;\,p_1,\dots,p_{\kappa r}\,;\,\mathsf{list})$ is incoming to a certain range point in $\mathrm{B\Pi}_{k+(m+16)\kappa r}$. We now specify this point precisely, so that Φ is indeed constructed. The half-space walk γ has a branch decomposition (2.3) in which appear each of the tall branches $\phi^{(\ell)}$, $1 \le \ell \le t$, (though they may not do so in increasing order). Replace each of these terms by its post-surgical counterpart $\tilde{\phi}^{(\ell)}$. Then apply the operation, mapping (2.3) to (2.4), that replaces any given branch decomposition by its bridge counterpart. The outcome reported by this operation is an element of $\mathrm{B\Pi}_{k+(m+16)\kappa r}$. The arrow under Φ that we consider points to this element. We have constructed Φ.

5.4. **The near injectivity of Φ: Proof of Lemma 5.3.** That
$$m_0 \ge \exp\big(\kappa r\big[(1/2 - \varepsilon)\log n - \log(8\kappa)\big]\big)$$
for Φ follows directly from its construction and (5.2). We turn to the upper bound on M_0.

Let $(b_1,\dots,b_r)$ be an element of $\mathrm{B\Pi}_{k+\kappa r(m+16),r}$ to which an arrow under Φ is incoming. We wish to show that the domain point $(\gamma\,;\,p_1,\dots,p_{\kappa r})$ from which the arrow is outgoing may be determined from $(b_1,\dots,b_r)$ provided that κr choices, each from among at most twelve options, are made successively. We will establish the stronger assertion in which the data list is also accurately determined by these successful choices.

Let χ be the concatenation of suitable translations of $b_1, \tau(b_2), b_3, \dots$. Then χ is a post-surgical walk, whose branches are the translations of the just listed terms. Incorporated in these branches are local deformations of the polygons $p_1,\dots,p_{\kappa r}$ which have been attached

by the Madras join operation. Violations of self-avoidance on the part of χ are the residues of surgery and they will permit a very accurate inference to be made by an observer of χ as to the form of the κr surgeries by which this walk has been formed.

Recall from the start of Section 5.1 that the tall branches $\phi^{(1)}, \ldots, \phi^{(t)}$ of γ are those among γ's branches whose heights exceed $n^{1/2-\varepsilon}$. It is only on to these branches that polygons are grafted in the formation of χ, and any insertion of a polygon on to a branch leaves the height of the branch unchanged. Thus it is that a record of the branches of χ whose heights exceed $n^{1/2-\varepsilon}$ is given by the first t terms in the sequence $b_1, \tau(b_2), b_3, \ldots$.

We have recalled that the tall branches $\phi^{(1)}, \ldots, \phi^{(t)}$ of γ are recorded in the left-to-right order of their leftmost crossings of Rect. The insertion of polygons into these branches does not produce new crossings, and nor does it affect the intersections of crossings with the horizontal sides of Rect, at heights $j_{\min}$ and $j_{\max}$. The leftmost crossing made by a given branch of either γ or χ occupies the leftmost coordinates at height $j_{\min}$ (and also at height $j_{\max}$) among all crossings made by this branch. For this reason, inspection of the first t terms of the sequence $b_1, \tau(b_2), b_3, \ldots$ permits a permutation to be recorded in the form $\psi^{(1)}, \ldots, \psi^{(t)}$ so that the terms correspond to the list $\phi^{(1)}, \ldots, \phi^{(t)}$ of the tall branches of γ.

The branch decompositions of γ and χ coincide beyond the first t terms. This is also true for the tth term, because no polygon insertion is attempted in this, the rightmost, case. Regarding the earlier terms, let $\ell \in \{1, \ldots, t-1\}$. The branch $\psi^{(\ell)}$ is the Madras join $\phi^{(\ell)}$ with a certain number of polygons which we label $p_{(\ell,j_1)}, \ldots, p_{(\ell,j_J)}$, where $j_1 < \cdots < j_J$ are the height parameters j in the Madras join operations.

We aim to retrieve each $\phi^{(\ell)}$ and the polygons $p_{(\ell,j_1)}, \ldots, p_{(\ell,j_J)}$ by interpreting intersections between $\psi^{(\ell)}$ and $\psi^{(\ell+1)}$ as signposts that mark where these polygons have been inserted into $\phi^{(\ell)}$.

First we set $\ell = 1$. Several polygon insertions modify $\phi^{(1)}$ into $\psi^{(1)}$. Any such insertion alters the present walk by removing or adding edges whose endpoints have heights that occupy an integer interval. These intervals are disjoint for distinct polygon insertions, and so they may be ordered, from bottom to top. The first polygon insertion that we seek to undo is the lowest, in which $p_{(1,j_1)}$ is inserted into $\phi^{(1)}$. Lemma 5.1 indicates that a vertex in $\phi^{(2)}$ exists that lies in the outcome $\mathsf{MJ}_{j_1}(\phi^{(1)}, p_{(1,j_1)})$ of this insertion but does not lie in the input walk $\phi^{(1)}$. For definiteness, we select the leftmost of the lowest among such vertices, and call it v: this vertex is supposed to offer a clue as to the location of the junction plaquette used in this surgery. Naturally, the observer of Φ's range point $(b_1, \cdots, b_r)$ should be able to find v. This observer computes $\psi^{(1)}$ and $\psi^{(2)}$, and identifies v as the lowest point of intersection between these walks, choosing the leftmost among such points if need be. The vertex v is correctly identified by this means because, in view of the disjointness conditions placed on join location lists in Section 5.1, the form of $\psi^{(1)}$ within the horizontal strip whose the vertical coordinate interval is that in which perturbations occur in the surgery $\phi^{(1)} \to \mathsf{MJ}_{j_1}(\phi^{(1)}, p_{(1,j_1)})$ coincides with the form of $\mathsf{MJ}_{j_1}(\phi^{(1)}, p_{(1,j_1)})$ within this strip; and likewise the form of $\psi^{(2)}$ within this strip coincides with the form of $\phi^{(2)}$ therein. That is, the observer is able to infer the location of v because the essential attributes of the first polygon insertion are not disrupted by later insertions.

The next definition and lemma are needed alongside Lemma 3.4(2) and (3) to locate the junction plaquette given the value of v.

Definition 5.4. *Let $\ell \in \mathbb{N}$ and $\rho \in \mathsf{SAW}_\ell$. An index pair (j, k), $0 \leq j < k \leq \ell$, is called a right-detachable ρ-adjacency of gap $k - j$ if*

- *$\{\rho_j, \rho_k\}$ is the endpoint pair of a vertically oriented unit edge;*

- *the horizontal edges in the plaquette P whose right border is this edge belong to ρ and the vertical edges do not;*

- *the modification $\rho \, \Delta \, P$ formed from ρ by the removal of the horizontal edges in P and the insertion of the vertical ones is a disjoint union of a walk ρ' and a polygon Q;*

- *and any translation of Q directly to the right is disjoint from ρ'.*

Lemma 5.5. *Let $\ell \in \mathbb{N}$ and let $\rho \in \mathsf{SAW}_\ell$ attains its minimal height at ρ_0 and its maximal height at ρ_ℓ. For $h \in \mathbb{N}$, two right-detachable ρ-adjacencies (j_1, k_1) and (j_2, k_2) of gap h satisfy $|k_2 - k_1| \geq h$.*

Lemma 5.5 depends on planarity, so this tool has a certain importance. However, we defer its slightly irksome proof until the derivation of Lemma 5.3 is finished.

Set $\rho = \mathsf{MJ}_{j_1}\big(\phi^{(1)}, p_{(1,j_1)}\big)$. To find the junction plaquette used in the surgery $\phi^{(1)} \to \rho$, let $i \in \mathbb{N}$ satisfy $v = \rho_i$. Since v does not belong to $\phi^{(1)}$, Lemma 3.4(8) implies that v lies in ρ's surgically implanted middle section, which has length at most $m + 18$. The upper-right vertex of the junction plaquette also lies in this section, and thus this vertex's index in $M(\gamma, P)$ lies in the interval $I := [i - m - 18, i + m + 18]$. By Lemma 3.4(2) and (3), the indices in $M(\gamma, P)$ of the lower-right and upper-right corners of the junction plaquette are the elements of a right-detachable ρ-adjacency of gap $m + 7$. By Lemma 5.5, the distance between the higher indices in a pair of such adjacencies is at least $m + 7$. Since the condition $m \geq 17$ ensures that $2m + 37 < 3(m + 7)$, the interval I may contain the higher index of at most three such adjacencies.

In this way, the junction plaquette in χ associated with the lowest polygon insertion among those that construct $\psi^{(1)}$ may be detected by correctly choosing an element from a set of size at most three. (With further analysis of the Madras join, the junction plaquette may perhaps be detected uniquely, but we will not attempt this detection.) This choice made, the walk and polygon used in this surgery may be determined via Lemma 3.4(5). There are at most four forms for the presurgical walk-polygon pair given the surgical outcome alongside the identity of the junction plaquette.

Thus, the data $\big(p_{(1,j_1)}, j_1\big)$ and the form of the deformation to $\phi^{(1)}$ as a result of the lowest insertion on to this branch may be recovered by correctly choosing an element in a set of size at most twelve.

The counterpart recovery should be accomplished for the second lowest polygon insertion for the leftmost tall branch; and then the third; and then all higher such; and then we may increase the value of ℓ so that it equals two; and recover all polygons attached to this branch, from the lowest to the highest; and likewise for all tall branches; and so we remove all inserted
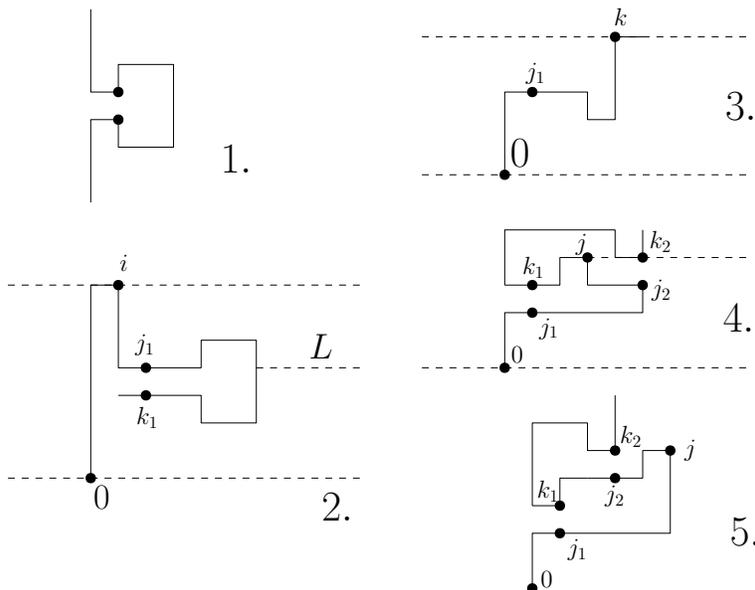
FIGURE 6. Claims and cases in the proof of Lemma 5.5. Each sketch depicts a walk ρ. Each label refers to the chemical distance i along ρ of the vertex ρ_i indicated by the nearby dot.

polygons. Each of the κr removals introduces a factor of at most twelve to our estimate of incoming arrow number to the given range point of Φ. At the end, we have identified $\left(\phi^{(1)}, \ldots, \phi^{(t-1)}\right)$, the form of each polygon, and the order of the polygons. Since the higher indexed $\phi^{(i)}$ are already determined, γ is recovered as the concatenation $\phi^{(1)} \circ \cdots \circ \phi^{(r)}$. That is, $(\gamma; p_1, \ldots, p_{\kappa r})$ is known; and so is list, though this is incidental. This completes the proof of Lemma 5.3. □

Proof of Lemma 5.5. Let P denote the plaquette whose right border has endpoints ρ_{j_1} and ρ_{k_1}, and let Q denote the polygonal component of $\rho \Delta P$. Note that the vertices of Q are those of $\rho_{[j_1,k_1]}$. The rightmost vertex of maximal height in Q, $\mathsf{NE}(Q)$, will be denoted by ρ_j, where note that the index j lies in $[\![j_1, k_1]\!]$. Let ρ' denote the walk component of $\rho \Delta P$. Figure 6(1) depicts the vertex pair, such as $\{\rho_{j_1}, \rho_{k_1}\}$, corresponding to a right-detachable ρ-adjacency as a pair of dots along the walk ρ.

First we show two claims.

Claim 1. The vertex ρ_{k_1} lies one unit upwards from ρ_{j_1}.

Proof. In the opposing case, ρ_{k_1} is located one unit downwards from ρ_{j_1}. Let ρ_i denote a vertex of maximal height on $\rho_{[0,k_1]}$: see Figure 6(2). Let L denote a semi-infinite horizontal line segment whose left endpoint is a vertex of maximal x-coordinate on $\rho_{[j_1,k_1]}$; and note that this vertex lies in Q. By our assumption on ρ_{k_1} and planarity, ρ_{k_1} lies in a component of the upper half-plane after the removal of the union of $\rho_{[0,k_1]} \setminus \{\rho_{k_1}\}$ and L which contains no point of height greater than or equal to that of ρ_i. Since ρ terminates at such a height, the path along ρ from ρ_{k_1} must cross L. Since the left endpoint of L lies in Q, we see that some

right translation of Q intersects $\rho_{[k_1+1,\ell]}$ and thus also intersects ρ'. However, this is contrary to the hypothesis that (j_1, k_1) is a right-detachable ρ-adjacency. □

Claim 2. The journey along $\rho_{[j_1,k_1]}$ from ρ_{j_1} to ρ_{k_1} first encounters a vertex of maximal height at $\rho_j = \mathsf{NE}(Q)$.

Proof. After this journey first reaches this height, at ρ_k say, a rightward turn on the part of the walk would, as Figure 6(3) depicts, leave the walk in a different component of the strip $\mathbb{R} \times [y(\rho_0), y(\rho_k)]$ than the point that is one unit above ρ_{j_1}. By Claim 1, the latter location is the endpoint of the journey, ρ_{k_1}. Thus such a rightward turn is impossible, and Claim 2 is proved. □

Recall that (j_2, k_2) is also a right-detachable ρ-adjacency, and that we seek to show that $k_2 - k_1$ is at least the shared value $k_1 - j_1 = k_2 - j_2$. It is thus enough to argue that the intervals $[j_1, k_1]$ and $[j_2, k_2]$ are disjoint, except possibly for some coincidence between endpoints. To establish this, suppose that the contrary holds; and, without loss of generality, that $j_2 \geq j_1$. Note then that $j_2 < k_1$.

We will analyse three cases: $j_2 < j$; $j_2 = j$; and $j > j_2$.

Case 1. Suppose that $j_2 < j$, and see Figure 6(4). Claim 2 implies that, in this case, when the journey along ρ from ρ_{j_1} to ρ_j passes through ρ_{j_2}, it does so at a strictly lower height than is achieved at its ending point. That is, $y(\rho_{j_2}) < y(\rho_j)$. Recall that ρ_{k_2} is supposed to be adjacent to ρ_{j_2}. The path $\rho_{[k_1,k_2]}$ is not permitted to visit heights less than that of ρ_0. In order to reach its endpoint ρ_{k_2}, this path must, by planarity, intersect, at a point other than ρ_{k_1}, the union of $\rho_{[0,j]} \cup \rho_{[j_2,k_1]} = \rho_{[0,k_1]}$ and the semi-infinite planar line whose left endpoint is $\mathsf{NE}(Q) = \rho_j$. Since self-avoidance prevents such crossing of $\rho_{[0,k_1]}$, we see that such crossing occurs on the semi-infinite line. Thus, some translation of Q to the right intersects $\rho_{[k_1+1,k_2]}$. But the latter is a subpath of ρ', the walk component of $\rho \Delta P$, so this inference is contrary to assumption.

Case 2. Suppose that $j_2 = j$. The vertex ρ_{j_2-1} lies directly to the left of ρ_{j_2}, because j_2 is an element in a right-detachable ρ-adjacency. However, ρ_{j-1} lies directly downwards from ρ_j, by Claim 2. Case 2 is thus impossible.

Case 3. Suppose instead that $j < j_2 < k_1$, and see Figure 6(5). Write $\overline{P}$ for the plaquette whose right border has endpoints ρ_{j_2} and ρ_{k_2}, and write $\bar{\rho}'$ and $\overline{Q}$ for the walk and polygon of which $\rho \Delta \overline{P}$ is comprised.

By Claim 2, the walk ρ turns left after reaching ρ_j. It does not reach a higher y-coordinate until at least index k_1, because $\rho_{[j,k_1]}$ describes part of the polygon Q. Thus, $\rho_{[0,j]}$ disconnects every point in $\rho_{[j+1,k_1]}$ from points arbitrarily far to the right within the horizontal strip of boundary heights $y(\rho_0)$ and $y(\rho_j)$. In Case 3, $\rho_{[j_2,k_1]}$ is a subwalk of $\rho_{[j+1,k_1]}$. Since $\rho_{[j_2,k_1]}$ is a subwalk of $\overline{Q}$, we see that, if $\overline{Q}$ is moved out to the right, it will encounter $\rho_{[0,j]}$ at some point. The walk $\bar{\rho}'$ contains $\rho_{[0,j_2-1]}$ and thus also $\rho_{[0,j]}$. Thus, the rightward movement of $\overline{Q}$ will encounter $\bar{\rho}'$. This is contrary to the assumption that (j_2, k_2) is a right-detachable ρ-adjacency. □

5.5. **Prospects for polygon insertion.** As we remarked after Proposition 4.1, it seems unlikely that the present method may improve Theorem 1.2 beyond $\varepsilon = 1/12 + o(1)$, since the parameter α in (4.3) is predicted to be at least $5/2$. Indeed, our method of polygon insertion has been presented in an effort to communicate how an improvement of some $\varepsilon > 0$ may be achieved, and we have not been concerned with the explicit value of ε so obtained. Here we make two comments about how further progress could be made by varying the method's implementation.

Obtaining suitably wide polygons more efficiently. Elements of WSAP^u_{4m+4} were produced in the proof of Lemma 3.3 by joining, under a suitable circumstance, pairs of elements of $\overline{\mathsf{WSAP}}^u_{2m+2}$. Since the latter elements were formed of two bridges in Kesten's proof of Lemma 1.6, four bridges are needed to build an element of WSAP^u_{4m+4}. However, we might replace the use of the set WSAP^u_m in the proof of Theorem 1.2 by a set such as $\{p \in \overline{\mathsf{WSAP}}^u_m : y(\mathsf{ne}(p)) = j\}$, for any given value of j. The new polygons, each formed of merely two bridges, may not have sufficient line-width, but they may have significant width when oriented suitably. A potential improvement in Proposition 4.1(1) and Theorem 1.2 may arise from this approach, due to an improvement on the polygon abundance lower bound $\alpha > 29$ in Proposition 3.1.

Making inserted polygons of variable length. Every polygon that is inserted in the proof of the key estimate, Proposition 4.4, is drawn from WSAP^u_m. As such, each of these polygons has given length m. One could attempt polygon insertion by permitting this length to vary, say between m and $2m$. There would seem to be a gain in outgoing arrow entropy in the construction of the multi-valued map Φ which leads to an effective drop of one in the value of the parameter α. In principle, then, the limit of this variant of the method takes α equal to $3/2$, rather than $5/2$. The challenge for this variant is that incoming arrow number may increase because, in the counterpart to the proof of Lemma 5.3, the value of polygon length on the interval $[m, 2m]$ must be surmised as each polygon insertion is undone. The post-surgical walk crosses the lower side of the junction plaquette used in surgery, journeys over the inserted polygon, and travels back across this plaquette's upper side. As such the plaquette's location can be roughly surmised, unless the inserted polygon has many locations at which chemically distant vertices are neighbours. The rarity of such locations may be gauged by [Ham18, Proposition 4.5].

6. Abundance of polygons for the hexagonal lattice

Two remaining sections treat self-avoiding walks on the hexagonal lattice. In the present section, we prove the polygon abundance estimate for this lattice, Proposition 3.2, which in its exponent and its validity for all lengths improves on Proposition 3.1. Theorem 1.4 will be proved in Section 7. Both results depend vitally on a very specific integrability enjoyed by self-avoiding walk on $\mathbb{H}$, expressed in terms of the *parafermionic observable*. Use of the observable entails some changes in model definition that we present in a first subsection. In a second, we offer a quick introduction to the observable. In a third, we prove Proposition 3.2.

6.1. **Definitions for the hexagonal lattice.** A union Ω of faces in $\mathbb{H}$ is called a *domain* if there exists a self-avoiding polygon P on $\mathbb{T} = \mathbb{Z} + e^{i\pi/3}\mathbb{Z}$, the graph which is dual to $\mathbb{H}$, such

that Ω is the set of faces contained in the finite connected component of $\mathbb{R}^2 \setminus P$. Let $\partial\Omega$ be the intersection of P with the edges of $\mathbb{H}$. Note that the elements of $\partial\Omega$ are the midpoints of those edges of $\mathbb{H}$ that have exactly one endpoint in the interior of P. A triangular domain is depicted in the upcoming Figure 8.

An internal edge in Ω is an edge in $\mathbb{H}$ that is contained in Ω. The midpoint of an edge in Ω refers either to the midpoint of an internal edge or to an element of $\partial\Omega$. A vertex in Ω is a vertex in $\mathbb{H}$ that lies in Ω.

We now specify notation for self-avoiding walks and polygons on $\mathbb{H}$ and on graphs useful for their analysis. The definitions are at variance with our earlier usage.

Let Ω be a domain. Two midpoints of edges in Ω are said to be adjacent if the edges share an endpoint. A self-avoiding walk on Ω is a sequence of adjacent edge midpoints in Ω in which there are no repetitions among the list of vertices in Ω that are the shared endpoints of consecutive members of the sequence. This formulation permits the possibility that the first and final elements in a self-avoiding walk are equal; in this case, the walk is called a self-avoiding polygon. As such, polygons are rooted, and have orientation. When the root and the orientation are forgotten, we recover a counterpart to the definition of polygon used for $\mathbb{Z}^2$; here, we will call this object a *polygon trace*.

The length of a self-avoiding walk is one less than the number of sequence elements; namely, the number of vertices of Ω visited by the walk. A length zero walk thus refers to one that begins and ends at the same edge midpoint without passing through any domain vertex.

Let $\mathsf{SAW}_n(\Omega)$ denote the set of self-avoiding walks of length n on Ω; the usage is at variance with that of Section 1.4, because the starting point of the walks is not fixed. When the subscript n is absent, the set of such walks of arbitrary length is designated. The notation SAP refers to polygons in place of walks. Omission of Ω is understood to imply $\mathbb{H}$.

Henceforth x denotes $\mu(\mathbb{H})^{-1}$. Let z_0 and z be midpoints of edges of Ω. We specify the normalized walk count, or *partition function*,

$$G_\Omega(z) = G_\Omega(z_0, z) = \sum_{\substack{\gamma \in \mathsf{SAW}(\Omega) \\ \gamma : z_0 \to z}} x^{|\gamma|}, \tag{6.1}$$

where $\gamma : z_0 \to z$ indicates that the starting and ending points of γ are z_0 and z, and where G's first argument z_0 may be omitted when this value is understood. When $z = z_0$, a normalized count of polygons is being made.

6.2. A quick introduction to the parafermionic observable.

That $\mu = \mu(\mathbb{H}) = \sqrt{2 + \sqrt{2}}$ is a consequence [DCS12] of analysis of the observable. Applications also include [BBMdG$^+$14].

Given $\sigma \in \mathbb{R}$, a domain Ω and $z_0 \in \partial\Omega$, define the parafermionic observable at any midpoint z of an edge in Ω to be

$$F(z) = F_\Omega(z_0, z) := \sum_{\substack{\gamma \in \mathsf{SAW}(\Omega) \\ \gamma : z_0 \to z}} e^{-i\sigma \mathrm{wind}(\gamma)} x^{|\gamma|}, \tag{6.2}$$

where $\mathrm{wind}(\gamma)$ is the total rotation of γ from z_0 to z.

It is by now classical (see [Smi10, Lemma 4]) that F satisfies the following relations when σ is set equal to $5/8$ in addition to the setting $x = \mu(\mathbb{H})^{-1}$. For the midpoints p, q, r of the three edges incident to a vertex $v \in \Omega$,

$$(p - v)F(p) + (q - v)F(q) + (r - v)F(r) = 0, \tag{6.3}$$

where $p - v$, $q - v$ or $r - v$ are interpreted as complex numbers.

A simple and important observation is that, for $z \in \partial\Omega$ and $x > 0$, the observable can be related to the generating function $G(z) = G_\Omega(z_0, z)$ of walks from z_0 to z staying in Ω. Indeed, since z lies on the boundary of Ω, the winding of all paths going from z_0 to z is the same, so that

$$F(z) = e^{-i\sigma \operatorname{wind}(z)} G(z), \tag{6.4}$$

where $\operatorname{wind}(z)$ is a deterministic constant. Here it is crucial that Ω be simply connected, as is the case here by Ω's definition.

We set $\sigma = 5/8$ henceforth.

6.3. **An improved lower bound for** WSAP: **deriving Proposition 3.2.** It is proved in [DCS12, Remark 2] that there exists $c > 0$ such that, for $u \geq 1$,

$$B_u := \sum_{n \geq 1} \left| \{ \gamma \in \mathsf{SAB}_n : \mathsf{h}(\gamma) = \tfrac{\sqrt{3}}{2} u \} \right| \cdot \mu^{-n} \geq c\, u^{-1}. \tag{6.5}$$

This bound does not directly yield a lower bound on $b_n \mu^{-n}$, but Lemma 3.3 may perhaps be adapted to express a lower bound on $|\mathsf{WSAP}_m^u|$ in terms of B_u, so that some form of Proposition 3.2 would follow from (6.5). Here, we apply (6.5) in a different fashion in order to obtain this proposition. A similar argument has appeared in [GM17].

Recall from the introduction that, for $k \in \mathbb{N}$, Λ_k is the set of faces in $\mathbb{H}$ whose distance from the face containing the origin is at most k, in the sense of distance on the triangular lattice; and note that Λ_k is a domain of $\mathbb{H}$. Let $\mathsf{RSAP}^{(k)}$ be the set of *round* self-avoiding polygons: those polygons that are included in Λ_{8k}; begin and end at $(k + 1/2, 0)$; surround, but do not intersect, the planar line segment that interpolates $(0, 0)$ and $(k, 0)$; and which intersect the positive x-axis only at $(k + 1/2, 0)$. There is no demand that the polygons' length be given.

Round polygons are readily available:

Proposition 6.1. *There exists a constant $c > 0$ such that, for $k \in \mathbb{N}$,*

$$\sum_{\gamma \in \mathsf{RSAP}^{(k)}} x^{|\gamma|} \geq c\, k^{-1} B_{8k}^6.$$

This proposition will be used the proof of Theorem 1.4. We now use it to prove Proposition 3.2 and then give its proof.

Proof of Proposition 3.2. This result offers a lower bound on the cardinality of WSAP_m^u. We mention that, in our present notation, this is a set of polygon traces, rather than polygons.
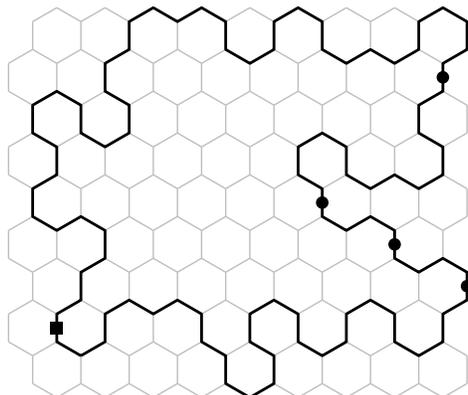
FIGURE 7. A polygon trace. Dots mark certain starting points for which the polygon is in $\mathsf{RSAP}^{(u)}$. The square marks the starting point used in the parameterization of elements of WSAP^u_m.

Due to Proposition 6.1,

$$\sum_{\gamma \in \mathsf{RSAP}^{(u)}} x^{|\gamma|} = \sum_{m=u}^{|\Lambda_{8u}|} \left|\{\gamma \in \mathsf{RSAP}^{(u)} : |\gamma| = m\}\right| \mu^{-m} \geq c\, u^{-1} B^6_{8u}.$$

The equality is due to each $\gamma \in \mathsf{RSAP}^{(u)}$ having length between u and $|\Lambda_{8u}|$. The number of terms in the sum is bounded by $c_0 u^2$ for some constant c_0. Thus, there exists at least one value for m such that

$$\left|\{\gamma \in \mathsf{RSAP}^{(u)} : |\gamma| = m\}\right| \mu^{-m} \geq c_1\, u^{-3} B^6_{8u}.$$

for some universal $c_1 > 0$. Finally, notice that, for any γ in the set above, the trace of γ is an element of WSAP^u_m. Indeed, the line-width of γ is at least u, since γ surrounds the segment $[0, u] \times \{0\}$; and its height is bounded by $16u$ because it is contained in Λ_{8u}.

The passage $\{\gamma \in \mathsf{RSAP}^{(u)} : |\gamma| = m\} \to \mathsf{WSAP}^u_m$ from polygon to trace is not injective. Any $\gamma \in \mathsf{RSAP}^{(u)}$ is rooted at the rightmost point of its intersection with the horizontal line containing its root: see Figure 7. Since γ visits at most $16u$ different heights, any given trace corresponds to at most $16u$ polygons of $\mathsf{RSAP}^{(u)}$. As such, using (6.5), we find that

$$\left|\mathsf{WSAP}^u_m\right| \mu^{-m} \geq \frac{c_1}{16u^4} B_{8u}(x_c)^6 \geq c_2\, u^{-10},$$

for some constant $c_2 > 0$. □

In the rest of this subsection, Proposition 6.1 is proved. We start with a lemma.

Fix $k \in \mathbb{N}$. Let the strip $\mathcal{S}_k$ of height k and the equilateral triangle $\mathcal{T}_k$ of side length $2k+1$ be defined as the domains whose internal edges are precisely those edges of $\mathbb{H}$ that are respectively contained in

$$\left\{(x, y) \text{ such that } 0 < y < \tfrac{\sqrt{3}}{2}k\right\} \subset \mathbb{R}^2 \qquad \text{and}$$
$$\left\{(x, y) \text{ such that } 0 < y < \sqrt{3}(k + \tfrac{1}{2} - |x - k - \tfrac{1}{2}|)\right\} \subset \mathbb{R}^2\,;$$
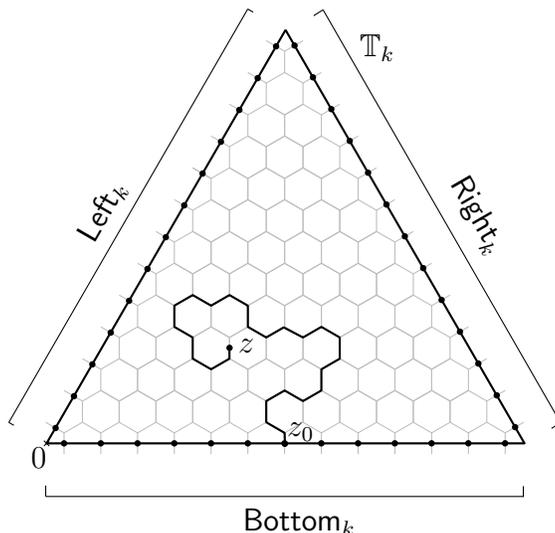
FIGURE 8. The domain $\Omega = \mathcal{T}_k$: the polygon P around it is depicted in bold; the edges of $\mathcal{T}_k$ are all those with at least one endpoint in P; and the boundary edges are those with exactly one endpoint inside P. The centres of the boundary edges, marked with dots, form $\partial \mathcal{T}_k$. The three sides of $\partial \mathcal{T}_k$ are denoted by Bottom_k, Left_k, Right_k; z_0 is the midpoint of the edge in $\mathcal{T}_k$ emanating from a. The depicted path starts at z_0 and ends at the midpoint z of an edge inside $\mathcal{T}_k$, as in the definition of the parafermionic observable. This path has winding 2π at z.

see Figure 8. In an evident notation, let Bottom and Top_k partition $\partial \mathcal{S}_k$; and let Bottom_k, Left_k and Right_k do so for $\partial \mathcal{T}_k$. Note that B_k is the partition function of walks contained in $\mathcal{S}_k$ that start at some fixed point on Bottom and end on Top_k.

Lemma 6.2. *There exists $c > 0$ such that, for even $k \in \mathbb{N}$,*

$$\sum_{\substack{\gamma \in \mathsf{SAW}(\mathcal{T}_k) \\ \gamma : z_0 \to \mathsf{Left}_k}} x^{|\gamma|} \geq x \, B_{2k} \,,$$

where $z_0 = \left(k + 1/2, 0\right)$ is in the middle of Bottom_k.

Proof. Consider the observable $F = F_{\mathcal{T}_k}$ defined in (6.2). Summing (6.3) over all vertices $v \in \mathcal{T}_k$, we find that the contributions of each internal edge to the relations around its endpoints cancel each other out; whence

$$e^{\pi i/3} \sum_{z \in \mathsf{Left}_k} F(z) + e^{-\pi i/3} \sum_{z \in \mathsf{Right}_k} F(z) - \sum_{z \in \mathsf{Bottom}_k \backslash \{z_0\}} F(z) = F(z_0) \,. \tag{6.6}$$

Now, $F(z_0) = 1$, since only the walk of length zero contributes. Note also that $\mathrm{wind}(z)$, as defined in (6.4), is equal to $\pi/3$ on Left_k; to $-\pi/3$ on Right_k; and to $\pm \pi$ on Bottom_k, the choice of sign depending on whether z is on the left or on the right of z_0. In particular, (6.4)

enables us to transform (6.6) into

$$2\cos\left[\tfrac{\pi}{3}(1-\sigma)\right]\sum_{z\in\mathsf{Left}_k} G_{\mathcal{T}_k}(z) + \cos[\pi(1-\sigma)]\sum_{z\in\mathsf{Bottom}_k\setminus\{z_0\}} G_{\mathcal{T}_k}(z) = 1\,, \qquad (6.7)$$

where we used the symmetry of $\mathcal{T}_k$ with respect to the vertical line running through z_0.

Proceeding for $\mathcal{S}_{2k}$ as we did for $\mathcal{T}_k$, we obtain

$$\sum_{z\in\mathsf{Top}_{2k}} G_{\mathcal{S}_{2k}}(z) + \cos[\pi(1-\sigma)]\sum_{z\in\mathsf{Bottom}\setminus\{z_0\}} G_{\mathcal{S}_{2k}}(z) = 1. \qquad (6.8)$$

Here, it may seem problematic that $\mathcal{S}_{2k}$ is infinite. However, an argument, whose details we omit because it is arguably rather standard, in which a multi-valued map is constructed between $\mathsf{SAW}_n(\mathcal{S}_{2k})$ and a space of rather longer walks under which height k polygons are grafted on to a domain point at a small constant fraction of horizontal coordinates, demonstrates that $\mu(\mathcal{S}_{2k}) < \mu(\mathbb{H}) = x^{-1}$. Thus is the summation in (6.8) permitted.

Finally, observe that, since $\mathcal{T}_k \subset \mathcal{S}_{2k}$, for any $z\in\mathsf{Bottom}_k$, $G_{\mathcal{T}_k}(z) \le G_{\mathcal{S}_{2k}}(z)$. Moreover, the coefficients $\cos[\pi(1-\sigma)]$ and $2\cos\left[\tfrac{\pi}{3}(1-\sigma)\right] = 1/x$ are both positive. Thus, by subtracting (6.8) from (6.7), we obtain that, for every $k\in\mathbb{N}$,

$$\sum_{z\in\mathsf{Left}_k} G_{\mathcal{T}_k}(z) \ge \frac{1}{2\cos\left[\tfrac{\pi}{3}(1-\sigma)\right]}\sum_{z\in\mathsf{Top}_{2k}} G_{\mathcal{S}_{2k}}(z) = x\,B_{2k}\,. \qquad (6.9)$$

$\square$

Remark. Equation (6.8) implies that $B_k = \sum_{z\in\mathsf{Top}_k} G_{\mathcal{S}_k}(z)$ is decreasing in k.

Aside. The same equation shows that

$$B_k - B_{k+1} = \cos[\pi(1-\sigma)]\sum_{z\in\mathsf{Bottom}\setminus\{z_0\}} [G_{\mathcal{S}_{k+1}}(z) - G_{\mathcal{S}_k}(z)] \le \frac{\cos[\pi(1-\sigma)]}{x}B_{k+1}^2\,,$$

since any walk contributing to $G_{\mathcal{S}_{k+1}}(z) - G_{\mathcal{S}_k}(z)$ is what may be called an arc from z_0 to z: a walk that visits $\mathcal{S}_{k+1}\setminus\mathcal{S}_k$; and hence such a walk decomposes into two bridges of height $k+1$ (after a small alteration). The above was used in [DCS12] to prove (6.5).

We now present the proof of Proposition 6.1. A similar strategy was used in [DCGPS17] for loop $O(n)$ models.

Proof of Proposition 6.1. For every $k', k'' \ge 0$, define the quantity

$$G(k', k'') := G_{\mathcal{T}_k}\left((\tfrac{1}{2}+k', 0)\,, (\tfrac{1}{2}+k'', 0)\mathrm{e}^{\mathrm{i}\pi/3}\right),$$

and note that $G(k', k'') = 0$ if $k'' > 2k'$. Let Λ_{8k}^+ denote the intersection of Λ_{8k} and the upper half-plane. Concatenating suitable rotations of three walks that contribute to $G(k, k_1)$, $G(k_1, k_2)$ and $G(k_2, k')$, we obtain a walk in Λ_{8k}^+ from $(\tfrac{1}{2}+k, 0)$ to $(\tfrac{1}{2}-k', 0)$: see Figure 9.

Thus,

$$\sum_{k'=0}^{8k} G_{\Lambda_{8k}^+}\left(z_0, (\tfrac{1}{2}-k', 0)\right) \ge \sum_{k_1\le 2k}\sum_{k_2\le 2k_1}\sum_{k'\le 2k_2} G(k, k_1)G(k_1, k_2)G(k_2, k')\,, \qquad (6.10)$$
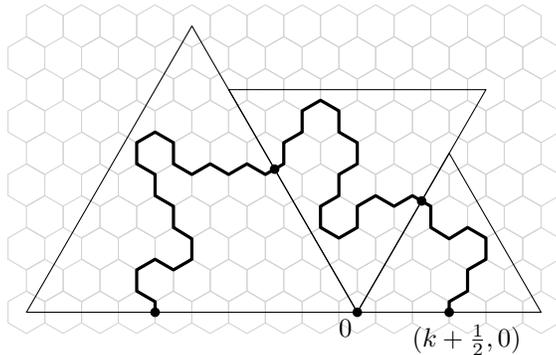
FIGURE 9. The concatenation of three walks rotations of which contribute to $G(k, k_1)$, $G(k_1, k_2)$ and $G(k_2, k')$ produces a walk contained in the upper half of Λ_{8k}, from $(\frac{1}{2} + k, 0)$ to $(\frac{1}{2} - k', 0)$.

whose right-hand side is at most $x^3 B_{8k}^3$ in view of (6.9) and the monotonicity of B on which we remarked before the proof.

Consider two walks that contribute to the same summand in the preceding left-hand side. The concatenation of one with the vertical reflection of the other is a polygon of $\mathsf{RSAP}^{(k)}$. Thus, and by the Cauchy-Schwarz inequality,

$$\sum_{\gamma \in \mathsf{RSAP}^{(k)}} x^k \geq \frac{x^6}{8k} B_{8k}^6 .$$

$\square$

7. The proof of Theorem 1.4

The three arenas in which the proof of Theorem 1.4 will unfurl are the lattice $\mathbb{H}$; its universal cover $\mathbb{U}^\infty$; and its eight-fold cover, which we will call $\mathbb{U}^8$. Our description of $\mathbb{U}^\infty$ in the introduction was informal. We first precisely specify $\mathbb{U}^\infty$ and $\mathbb{U}^8$ and some pertinent subgraphs.

7.1. **The universal and eight-fold covers of the hexagonal lattice.** The positive x-axis intersects those edges in $\mathbb{H}$ that are vertical, symmetric about this axis and whose x-coordinates have the form $k + 1/2$ with $k \in \mathbb{N} \cup \{0\}$. When this semi-infinite line is removed from the plane, each of these edges is cut into two half-edges at its midpoint. The endpoints of the resulting half-edges that lie on the x-axis may be labelled with the terms 'lower' and 'upper', according to whether the argument values of these points, when specified via continuity, are equal to zero or to 2π. The branch-cut hexagonal lattice is the graph so formed: its vertices are the midpoints of edges in $\mathbb{H}$, with the exception of those on the positive x-axis, for which there are two copies, labelled 'upper' and 'lower'.

The universal cover $\mathbb{U}^\infty$ of $\mathbb{H}$ is formed from a $\mathbb{Z}$-indexed collection of branch-cut hexagonal lattices; any upper vertex on the positive x-axis in one lattice is identified with its lower counterpart in the lattice whose index is one greater.

Let $j \in \mathbb{N}$. The j-fold cover $\mathbb{U}^j$ of $\mathbb{H}$ is formed from a collection of j such lattices. The collection is cyclically ordered, and the cover results when the same identifications are made using this ordering of the lattice copies.

The canonical projection from $\mathbb{U}^\infty$ to $\mathbb{H}$ has been denoted by π_∞. The counterpart map from $\mathbb{U}^8$ to $\mathbb{H}$ will be called π_8.

At the centre of each face of $\mathbb{H}$ is an element of the triangular lattice $\mathbb{T}$. The distance between a pair of these faces is equal to the distance in $\mathbb{T}$ between the corresponding vertices. Let $k \in \mathbb{N} \cup \{0\}$. Recall Λ_k, the set of faces in $\mathbb{H}$ whose distance from the face containing the origin is at most k. The union of the six planar line segments that interpolate the consecutive rotations of the point $(k+1,0)$ about the origin by the sixth roots of unity is a planar polygon P_k. When P_k is removed from the plane, the lattice $\mathbb{H}$ fragments into the bounded domain Λ_k and an unbounded piece $\mathbb{U}_k$. (Reuse of the symbol $\mathbb{U}$ is intentional: note that $\pi_\infty^{-1}(\mathbb{U}_k)$ equals $\mathbb{U}_k^\infty$ from Conjecture 1. In fact, the specification of $\mathbb{U}_k^\infty$ in the introduction was rather informal, and we may define it to be $\pi_\infty^{-1}(\mathbb{U}_k)$.) The two pieces Λ_k and $\mathbb{U}_k$ share a common set $\partial\Lambda_k$ of boundary edge-midpoints, whose elements lie in the intersection of P_k with the edges in $\mathbb{H}$ and are midpoints of such edges.

7.2. **The road map for proving Theorem 1.4.** The expression (6.1) is a weighted count of walks. Upper bounds on such counts are a close surrogate for bounds on the cardinality of the elements of $\mathsf{SAW}_n(\mathbb{H})$ with a given starting point. Indeed, we begin deriving Theorem 1.4 by reducing it to the next result.

Proposition 7.1. *There exists $C > 0$ such that, for $n, k \in \mathbb{N}$ with $n \geq k$, $z_0 \in \partial\Lambda_n$ and z the midpoint of an internal edge in Λ_{n-k},*

$$G_{\Lambda_n}(z, z_0) \leq C(k)\, n^{C(6k+7)+103},$$

where $\{C(k) : k \in \mathbb{N}\}$ is a sequence that verifies $\sup_k k^{-1}\big(\log C(k)\big)^{1/3} < \infty$.

Remark. The constant $C > 0$ in the exponent in this result arises from the upcoming Lemma 7.5. We have not expanded effort on optimizing exponents in our bounds. We mention, however, that $C = 32$ is obtained in the proof of Lemma 7.5, with corresponding explicit values for exponents resulting in Proposition 7.1 and in Theorem 1.4, the latter shown in the proof which we now give.

Proof of Theorem 1.4. First, we note that the definition of self-avoiding walk implied in the introduction, in which such walks begin and end at vertices in $\mathbb{H}$, is at variance with our actual usage, in which it is the midpoints of edges in $\mathbb{H}$ which accommodate these endpoints. It is a trivial matter, however, to interpolate between the two forms of Theorem 1.4.

Let o denote the midpoint of the vertical edge in $\mathbb{H}$ that borders the right side of the face containing the origin.

Let γ be a walk in $\mathbb{H}$ of length n that starts at o. Write $m = \min\{r \in \mathbb{N} : \gamma \subseteq \Lambda_r\}$. Let γ_ℓ be the first point of γ adjacent to $\partial\Lambda_m$, and let $z_0 \in \partial\Lambda_m$ be adjacent to γ_ℓ. Then $\gamma_{[0,\ell]}$ may be extended by one step to form a walk that contributes to $G_{\Lambda_m}(0, z_0)$, while $\gamma_{[\ell,n]}$ may be reversed and then so extended to form a walk that contributes to $G_{\Lambda_m}(\gamma_n, z_0)$. Suppose for now that the edge midpoint γ_n lies in Λ_{m-k}. We may apply Proposition 7.1 with $n = m$ twice, and sum over all possible values of $m \leq n$, $z_0 \in \partial\Lambda_m$ and $\gamma_n \in \Lambda_m$; we thus find that the number $c_n(\mathbb{H})$ of walks of length n starting at a given edge midpoint satisfies

$$\mu^{-n} c_n(\mathbb{H}) \leq C(k)^2 n^{2C(6k+7)+210}. \tag{7.1}$$

It remains to treat the case where γ_n lies in $\Lambda_m \setminus \Lambda_{m-k}$. The reversal of $\gamma_{[\ell,n]}$ may be extended from its endpoint γ_ℓ by k steps in such a fashion that γ_n is at least k steps closer to the origin than is the new opposing endpoint of this walk. At the expense of the presence of a further k-dependent constant on its right-hand side, we obtain (7.1) in this case also, because Proposition 7.1 has become applicable to the reversed and extended walk. □

When $z = o$ is taken in Proposition 7.1, the proof of the result coincides with the general case's except for the elimination of some mildly distracting notation. There are four principal steps in the proof of the proposition.

In the next subsection, we state four results, attached to these steps, in the special case when $z = o$; and give the proof of Proposition 7.1 for $z = o$. Four ensuing subsections present the proofs of the four upcoming results. In a final subsection, the notational changes needed to obtain Proposition 7.1 in its general form are explained. Our proof of Theorem 1.4 in fact succeeds under a slightly weaker hypothesis than Conjecture 1. The present Section 7 ends with a remark in this regard.

7.3. **The four steps to Proposition 7.1.** Lemma 7.2 is the first step to Proposition 7.1 with $z = o$. It is this step which dictates the use of the eight-fold cover. Set $\Lambda_n^8 = \pi_8^{-1}(\Lambda_n)$, $\mathbb{U}_n^8 = \pi_8^{-1}(\mathbb{U}_n)$, and note that $\partial\Lambda_n^8 = \pi_8^{-1}(\partial\Lambda_n)$. A walk in the eight-fold cover $\mathbb{U}^8$ that begins and ends in $\partial\Lambda_n^8$ is called 'inner' if every vertex in $\mathbb{U}^8$ that it visits lies in Λ_n^8; and 'outer' if every such vertex lies in $\mathbb{U}_n^8$. Such definitions may equally be made for the universal cover, with the change $8 \to \infty$.

Walks in $\mathbb{H}$ from o to $\partial\Lambda_n$, as measured by $G_{\Lambda_n}(o, z_0)$, are shown to be no more numerous than those walks in the eight-fold cover that begin and end in $\partial\Lambda_n^8$ and that are inner.

Lemma 7.2. *Let $n \in \mathbb{N}$. For $z_0 \in \partial\Lambda_n$, let $\overline{z}_0 \in \partial\Lambda_n^8$ verify $\pi_8(\overline{z}_0) = z_0$. Then*

$$G_{\Lambda_n}(o, z_0) \leq \frac{8}{\cos(3\pi/16)} \sum_{z \in \partial\Lambda_n^8} G_{\Lambda_n^8}(\overline{z}_0, z). \tag{7.2}$$

The second principal step provides a lower bound on the number of walks in the universal cover $\mathbb{U}^\infty$ between given suitable elements in $\pi_\infty^{-1}(\partial\Lambda_n)$ that are outer and have lengths of order roughly Cn^2.

We extend the notation (6.1), writing $G_{\Omega,\ell}(z_0, z)$ when the sum is indexed by walks whose length is at most ℓ; we also use it in cases such as $\Omega = \mathbb{U}_n^\infty$ which are not domains, as specified in Section 6.1, but for which associated definitions have clear counterparts.

Lemma 7.3. *There exist $c, C > 0$ such that, for $n \in \mathbb{N}$,*

$$G_{\mathbb{U}^\infty, Cn^2}(z_0, z) \geq c\, n^{-103}, \tag{7.3}$$

whenever $z_0, z \in \pi_\infty^{-1}(\partial\Lambda_n)$ are at distance in $\mathbb{U}^\infty$ of at most $48n$.

Any eight-fold cover inner walk that contributes to the sum on the right-hand side of (7.2) may be lifted to the universal cover $\mathbb{U}^\infty$. An outer walk that shares the endpoints of this lift is furnished by Lemma 7.3. The concatenation at both endpoints of inner and outer walk is a polygon in the universal cover. Thus it is that, for roughly the price of a walk in $\mathbb{H}$ from the origin to distance n, we obtain a polygon in $\mathbb{U}^\infty$ of length of order at most n^2 that runs through a given vertex in $\partial\Lambda_n^\infty$. Our purpose will be served, however, only if we succeed in obtaining a polygon in $\mathbb{U}_k^\infty$, rather than one merely in the superset $\mathbb{U}^\infty$. To this end, the polygon in $\mathbb{U}^\infty$ will be broken into several polygons so that each avoids Λ_k^∞. Our third step thus asserts that, for the given price, we may obtain possibly several such polygons in $\mathbb{U}_k^\infty$.

Corollary 7.4. *For $n, k \in \mathbb{N}$ with $n \geq k$, and $z_0 \in \partial\Lambda_n$,*

$$G_{\Lambda_n}(o, z_0) \leq C(k)\, n^{103} \left(\max_{z \in \mathbb{U}_k^\infty} G_{\mathbb{U}_k^\infty, 72n^2}(z, z) \right)^{6k+7}, \tag{7.4}$$

where $\{C(k) : k \in \mathbb{N}\}$ is a sequence that verifies $\sup_k k^{-1}\big(\log C(k)\big)^{1/3} < \infty$.

Our goal is the walk rarity bound $G_{\Lambda_n}(o, z_0) \leq n^C$. The first three steps have shown that polygons in $\mathbb{U}_k^\infty$ of length at most Cn^2 are as readily available as are walks contributing to this G-count. To reach our goal of a polynomial bound on $G_{\Lambda_n}(o, z_0)$, our fourth step demonstrates that such polygons are rare in a sense suitable for this purpose. It is at this moment in the proof of Theorem 1.4 that Conjecture 1 is needed.

Lemma 7.5. *There exists positive constants C_0 and C such that, whenever $k \in \mathbb{N}$ verifies $\mu(\mathbb{U}_k^\infty) = \mu(\mathbb{H})$, $n \in \mathbb{N}$ satisfies $n \geq k$, and $z \in \mathbb{U}_k^\infty$,*

$$G_{\mathbb{U}_k^\infty, n}(z, z) \leq C_0\, n^C.$$

Proof of Proposition 7.1 with $z = o$. By Corollary 7.4 and Lemma 7.5,

$$G_{\Lambda_n}(o, z_0) \leq C_0^{6k+7} C(k)\, n^{C(6k+7)+103}.$$

The relabelling of $C(k)$ so that it denotes $C_0^{6k+7}C(k)$ does not affect the growth rate of this sequence as the corollary describes it. Thus is the special case of Proposition 7.1 obtained. $\square$

7.4. **Proof of Lemma 7.2.**

Recall that $n \in \mathbb{N}$; that $z_0 \in \partial\Lambda_n$; that $\overline{z}_0 \in \partial\Lambda_n^8$ projects to z_0 under π_8; and that o is the midpoint of a particular edge that borders the face in $\mathbb{H}$ containing the origin. Eight points in Λ_n^8 project to o under π_8. Let v be one among these. The edge in $\mathbb{U}^8$ of which v is the midpoint will be denoted by e. A domain to be called $\Lambda_n^8(v)$ is formed from Λ_n^8 by cutting e in two half-edges. These two half-edges may be labelled 'plus' and 'minus' in an arbitrary fashion.

The domain $\Lambda_n^8(v)$ contains, in place of v, two vertices v^+ and v^-, which are endpoints of the new half-edges, and where v^+ is the endpoint of the plus half-edge. Note then that $\partial\Lambda_n^8(v) = \partial\Lambda_n^8 \cup \{v^-, v^+\}$.

In $\Lambda_n^8(v)$, self-avoiding walks may end at v^+, their journey ending via the plus half-edge generated by the cut at v; and they may end at v^-, if they finish by moving towards this vertex via the minus half-edge; they may not, however, traverse the edge e.

We define the observable in $\Lambda_n^8(v)$ as we did before, in (6.2):

$$F(z) = F_{\Lambda^8(v)}(\overline{z}_0, z) := \sum_{\substack{\gamma \in \mathsf{SAW}(\Lambda_n^8(v)) \\ \gamma : \overline{z}_0 \to z}} e^{-i\sigma \, \mathrm{wind}(\gamma)} x^{|\gamma|} \,. \tag{7.5}$$

Since $\overline{z}_0 \in \partial\Lambda_n^8$, while o lies on the edge of the singular face in $\mathbb{H}$, the difference of value for $\mathrm{wind}(\gamma)$ between two walks $\gamma : \overline{z}_0 \to o$ in Λ_n^8 is equal to an integer multiple of 16π. Since $\sigma = 5/8$, the σ-multiple of this difference is divisible by 2π, so the phase factor in (7.5) of all contributing walks is constant; (it is this consideration that determines our use of the eight-fold cover). Thus, it is that the local relation (6.3) (with the presently adopted values of x and σ) is obtained in the same way as it is in the planar case. Using the equivalent for $\Lambda_n^8(v)$ of (6.6), we deduce that

$$\left| F(v_+) - F(v_-) \right| \leq \sum_{z \in \partial\Lambda_n^8(v) \setminus \{v_+, v_-\}} G_{\Lambda_n^8(v)}(\overline{z}_0, z) \leq \sum_{z \in \partial\Lambda_n^8} G_{\Lambda_n^8}(\overline{z}_0, z) \,. \tag{7.6}$$

For a certain constant $\mathrm{wind}(v_+)$, the winding of a walk $z_0 \to v_+$ lies in $\mathrm{wind}(v_+) + 16\pi\mathbb{Z}$; while the winding of a walk $z_0 \to v_-$ lies in $\mathrm{wind}(v_+) - \pi + 16\pi\mathbb{Z}$. Since $-e^{-5i\pi/8} = e^{3i/8}$, we see that

$$\left| F(v_+) - F(v_-) \right| = \left| G_{\Lambda_n^8(v)}(\overline{z}_0, v_+) e^{i3\pi/16} + G_{\Lambda_n^8(v)}(\overline{z}_0, v_-) e^{-i3\pi/16} \right| \,.$$

For $a \in \mathbb{C}$, $|a| \geq (a + \bar{a})/2$; thus is the last right-hand side at least

$$\cos(\tfrac{3}{16}\pi) \left(G_{\Lambda_n^8(v)}(\overline{z}_0, v_+) + G_{\Lambda_n^8(v)}(\overline{z}_0, v_-) \right) = \cos(\tfrac{3}{16}\pi) \, G_{\Lambda_n^8}(\overline{z}_0, v) \,.$$

Any self-avoiding walk in Λ_n is the projection of at least one in $\overline{\Lambda}_n$; whence

$$G_{\Lambda_n}(o, z_0) \leq \sum_{v : \pi_8(v) = o} G_{\Lambda_n^8}(\overline{z}_0, v) \,.$$

Since there are eight such v,

$$G_{\Lambda_n}(o, z_0) \leq \frac{8}{\cos(3\pi/16)} \sum_{z \in \partial\Lambda_n^8} G_{\Lambda_n^8}(\overline{z}_0, z)$$

by (7.6) and the bound that is derived after it. This completes the proof of Lemma 7.2. □

7.5. **Deriving Lemma 7.3.** Two estimates for partition functions in the half-plane hexagonal lattice $\mathbb{H}^+ = (\mathbb{R} \times [0, +\infty)) \cap \mathbb{H}$ will be needed in this derivation. Recall the triangle $\mathcal{T}_k$ of Figure 8; translate it so that o is the center of Bottom_k. Due to (6.9) and (6.5), there exists a constant $c > 0$ such that, for $k \in \mathbb{N}$,

$$\sum_{z \in \mathsf{Left}_k} G_{\mathcal{T}_k}(o, z) \geq c \, k^{-1} \,. \tag{7.7}$$

Our second estimate refers to walks from o to arbitrary edge midpoints on the positive x-axis. It states that there exists $c > 0$ such that, for $k \in \mathbb{N}$,

$$G_{\Lambda_{4k} \cap \mathbb{H}^+}\big(o, (k+1/2, 0)\big) \geq c\, k^{-7}\,. \tag{7.8}$$

We now prove (7.8). Consider the rectangle $\mathcal{R}_k$ illustrated in Figure 10. Its base, Base_k, is centred at o; it has width $2k+1$; and height $4\sqrt{3}\, k$. Write Left_k', Top_{8k}' and Right_k' for its left, top and right sides. Writing (6.7) for $\mathcal{R}_k$ and subtracting (6.9) for $\mathcal{S}_{8k}$, we find that

$$\sum_{z \in \mathsf{Left}_k'} 2\cos\big(\tfrac{(3\pm1)\pi}{16}\big)\, G_{\mathcal{R}_k}(o, z) \;=\; \cos(\tfrac{3\pi}{8}) \sum_{z \in \mathsf{Base}\backslash\mathsf{Base}_k} G_{\mathcal{S}_{8k}}(o, z) \;+\; \sum_{z \in \mathsf{Top}_{8k}\backslash\mathsf{Top}_{8k}'} G_{\mathcal{S}_{8k}}(o, z)\,.$$

The angle in the left sum depends on the orientation of z; it alternates between $\pi/8 = (1-\sigma)\tfrac{\pi}{3}$ and $\pi/4 = (1-\sigma)\tfrac{2\pi}{3}$. These factors are strictly positive; thus, there exist constants $c_1, c_2 > 0$ such that

$$\sum_{z \in \mathsf{Left}_k'} G_{\mathcal{R}_k}(o, z) \;\geq\; c_1 \sum_{z \in \mathsf{Base}\backslash\mathsf{Base}_k} G_{\mathcal{S}_{8k}}(o, z) \;\geq\; c_2\, k^{-3}\,.$$

The last inequality is due to (6.10) and (6.5).

Next, observe that, for any $z \in \mathsf{Left}_k'$, two walks contributing to $G_{\mathcal{R}_k}(o, z)$ may be combined, after one of them is reflected horizontally and translated, to form a walk from o to $(-2k - 1/2, 0)$; see again Figure 10. Moreover, this walk is contained in the upper half-plane $\mathbb{H}^+$ and in Λ_{8k}. Thus, there exists $c_3 > 0$ such that

$$G_{\Lambda_{8k} \cap \mathbb{H}^+}\big(o, (-2k - \tfrac{1}{2}, 0)\big) \;\geq\; \sum_{z \in \mathsf{Left}_k'} x G_{\mathcal{R}_k}(o, z)^2 \;\geq\; c_3\, k^{-7}\,.$$

(The factor x is due to the red piece in Figure 10.) A composition of the walks with the addition of a suitable pair of edges yields the same lower bound for $G_{\Lambda_{8k} \cap \mathbb{H}^+}\big(o, (-2k - \tfrac{3}{2}, 0)\big)$, with a modified value for c_3. By relabelling $2k \to k$, monotonicity of G in the domain, and symmetry, (7.8) is proved.

Proof of Lemma 7.3. Recall that $n \in \mathbb{N}$ and that the distance in $\mathbb{U}^\infty$ between $z_0, z \in \pi_\infty^{-1}(\partial\Lambda_n)$ is at most $48n$. The elements of $\pi_\infty^{-1}(\partial\Lambda_n)$ may be labelled by $\mathbb{Z}$, with increasing index corresponding to counterclockwise movement along the copy of $\partial\Lambda_n$ obtained by projection via π_∞. The index set $\mathbb{Z}$ is partitioned into countably many intervals, each of length n, according to the segment in the planar polygon P_n on which the π_∞-image of the indexed element lies. Indeed, an unending counterclockwise journey along this polygon may be lifted via π_∞ so that the elements of $\pi_\infty^{-1}(\partial\Lambda_n)$ are visited in order; in fact, in order for this to make sense, π_∞ should be extended, compatibly with the present definition, to be a projection from the universal cover of $\mathbb{C}$ without the origin to this base space. Any point on this journey whose projection under π_∞ is the endpoint of one of the segments comprising the polygon P_n is called a corner. Corners naturally carry half-integral indices, and each partition interval is bookended by the indices of a pair of corners.

It is without loss of generality that we may suppose that z is higher than z_0 in this ordering of $\pi_\infty^{-1}(\partial\Lambda_n)$. Let $x_1, x_2, \ldots$ denote the corners encountered on the journey from z_0 to z along $\pi_\infty^{-1}(\partial\Lambda_n)$ dictated by the ordering. This sequence may be empty, and may have at
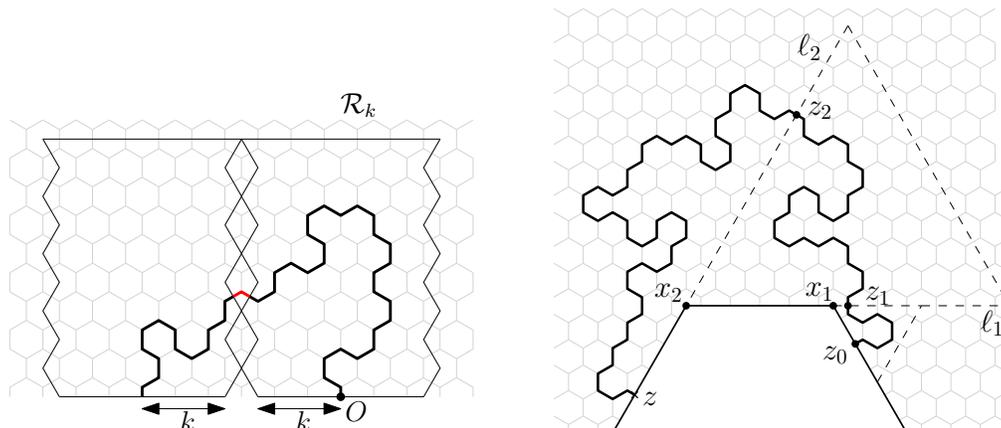
FIGURE 10. *Left:* The domain $\mathcal{R}_k$ and a translate of it. Two walks contributing to $G_{\mathcal{R}_k}(o, z)$ for some $z \in \mathsf{Left}'_k$ (one is reflected and translated) may be used to create an arc of between o and $(-2k - \frac{1}{2}, 0)$. The red edges are added in the concatenation *Right:* To connect z_0 to z we use (7.7) to arrive to some point z_i on the same line as z, then use (7.8) to connect z_i to z.

most forty-eight terms. Write ℓ_i for the semi-infinite planar line segment that emanates from $\pi_\infty(x_i)$ in the direction opposite to $\pi_\infty(x_{i+1}) - \pi_\infty(x_i)$. See Figure 10 for an illustration.

Let $c > 0$ be such that (7.7) and (7.8) are valid. If there is no corner between z and z_0, then (7.8) shows that

$$G_{\mathbb{U}_n^\infty, 3Cn^2}(z_0, z) \geq c\, n^{-7};$$

since the number of edge midpoints in $\Lambda_{4n} \cap \mathbb{H}$ is at most $3 \cdot 3/2 \cdot (4n)^2 = 72n^2$, this bound holds provided that $C \geq 24$.

Suppose now that the sequence $x_1, x_2, \ldots$ has $i \in \mathbb{N}$ terms. Thus, $\pi_\infty(z)$ lies on the planar line segment that contains $\pi_\infty(x_i)$ and $\pi_\infty(x_{i+1})$. Applying (7.7) with $k = n$ after a suitable translation, and using $|\mathsf{Left}_n| = n$, we find that there exists a point z_1 on ℓ_1 such that $G_{\mathcal{T}(z_0)}(z_0, z_1) \geq c\, n^{-2}$, where $\mathcal{T}(z_0)$ is the equilateral triangle the centre of whose base is z_0 and whose left corner is x_1. (More accurately, the image under π_∞ of $\mathcal{T}(z_0)$ is an equilateral triangle with corresponding data; we abuse notation in a moment by disregarding a similar nicety.) The distance between $\pi_\infty(x_2)$ and $\pi_\infty(z_1)$ is at most $3n$. Applying (7.7) again, there exists a point z_2 on ℓ_2 such that $G_{\mathcal{T}(z_1)}(z_1, z_2) \geq c\, (3n)^{-2}$, where $\mathcal{T}(z_1)$ is the equilateral triangle whose base is centred at z_1 and whose left corner is x_2. This process may be repeated until a point z_i on ℓ_i is reached. There are at most forty-eight steps and the size of each triangle is at most three times the preceding one's. When walks that respectively contribute to the terms $G_{\mathcal{T}(z_0)}(z_0, z_1)$, $G_{\mathcal{T}(z_1)}(z_1, z_2), \ldots$ are concatenated, an outer walk between z_0 and z_i is obtained. Thus, we find that

$$G_{\mathbb{U}_n^\infty, Cn^2}(z_0, z_i) \geq c'\, n^{-96},$$

where $C = 3^{2 \cdot 47} \cdot 9/8$ and $c' = c \cdot 3^{-2(1+2+\cdots+47)} \geq c\, 3^{-(47)^2}$. Finally, we apply (7.8) to connect z_i to z. By relabelling C and c, we thus obtain Lemma 7.3. $\square$

7.6. **Proof of Corollary 7.4.** Recall that $n \in \mathbb{N}$ and $z_0 \in \partial\Lambda_n$. In view of Lemma 7.2, we wish to bound above $\sum_{z \in \partial\Lambda_n^8} G_{\Lambda_n^8}(\overline{z}_0, z)$, where $\overline{z}_0 \in \partial\Lambda_n^8$ satisfies $\pi_8(\overline{z}_0) = z_0$.

Let γ be a walk contributing to this sum; γ is an inner walk in Λ_n^8, moving from $\partial\Lambda_n^8$ to itself. Suppose that γ visits Λ_k^8. We may decompose γ's lifetime between its first and last visits to Λ_{k+1}^8 into outer excursions γ_s from $\partial\Lambda_{k+1}^8$ to itself; and intervening journeys ψ_s within Λ_{k+1}^8. That is, for some $S = S(\gamma) \leq |\partial\Lambda_{k+1}^8|$, we may write $\gamma = \phi_{\text{start}} \circ \psi_1 \circ \gamma_1 \circ \psi_1 \circ \cdots \circ \gamma_{S-1} \circ \phi_S \circ \phi_{\text{end}}$, where $\phi_{\text{start}} : \overline{z}_0 \to \partial\Lambda_{k+1}^8$ and $\phi_{\text{end}} : \partial\Lambda_{k+1}^8 \to z$ lie in $\Lambda_n^8 \setminus \Lambda_{k+1}^8$; for $1 \leq i \leq S$, $\psi_i : \partial\Lambda_{k+1}^8 \to \partial\Lambda_{k+1}^8$ lies in Λ_{k+1}^8; and, for $1 \leq i \leq S-1$, $\gamma_i : \partial\Lambda_{k+1}^8 \to \partial\Lambda_{k+1}^8$ lies in $\Lambda_n^8 \setminus \Lambda_{k+1}^8$. In this description, and later in this proof, 'ϕ lies in Ω' means 'excepting ϕ's endpoints, every ϕ_i is the midpoint of an internal edge in Ω'.

The endpoints of ϕ_{start} and ϕ_{end} on $\partial\Lambda_{k+1}^8$ may be joined by a walk γ' that lies in $\Lambda_n^8 \setminus \Lambda_{k+1}^8$. For $1 \leq i \leq S-1$, the endpoints of γ_s may be joined by a walk that lies in $\Lambda_{k+1}^8 \setminus \Lambda_k^8$ so that a polygon p_s in $\mathbb{U}_k^8$ results. The ordering of the endpoints of γ_s used to produce the walk may be selected so that p_s does not disconnect the origin from infinity in $\mathbb{U}^8$. Intersections between the walk γ' and the polygons so produced are possible and do not concern us.

The topological property of the produced polygons permits that they be lifted to the universal cover, where they are contained in $\mathbb{U}_k^\infty$; we redesignate p_i to denote the polygon so lifted.

When γ does not visit Λ_k^8, set $\gamma' = \gamma$ and $S = 1$. Any point in the range of the application $\gamma \to (\gamma', p_1, \ldots, p_{S-1})$ has at most C_0 preimages, where C_0 is a certain constant that depends on k. Indeed, γ may be inferred from the image point by the reconstruction of the walks ψ_i, of which there are at most $|\partial\Lambda_{k+1}^8| \leq 6(k+1)$; since each ψ_i lies in Λ_{k+1}^8, we may choose $C_0 = 2^{54k^2(k+1)}$ since $\left(2^{|\Lambda_{k+1}^8|}\right)^{6(k+1)} \leq 2^{9k^2 \cdot 6(k+1)}$; here, and later in this proof, $|\Lambda_\ell^8|$ abusively denotes the number of internal edges in Λ_ℓ^8. Each polygon p_i has length at most $|\Lambda_n^8| \leq 8 \cdot 9n^2$. Setting $C_1 = 72$, we find that, since $S - 1 \leq 6(k+1)$,

$$G_{\Lambda_n^8}(\overline{z}_0, z) \leq C_0 \left(\max_{z_1 \in \partial\Lambda_{k+1}^\infty} G_{\mathbb{U}_k^\infty, C_1 n^2}(z_1, z_1) \right)^{6(k+1)} G_{\Lambda_n^8 \setminus \Lambda_k^8}(\overline{z}_0, z). \tag{7.9}$$

Next, we bound $G_{\Lambda_n^8 \setminus \Lambda_k^8}(\overline{z}_0, z)$. Every walk contributing to this sum is an inner walk in Λ_n^8, moving from $\partial\Lambda_n^8$ to itself. The lattice Λ_n^8 is a fusion of eight copies of branch cuts of Λ_n; by applying the map $x \to x^{1/8}$ before fusing these copies, Λ_n^8 may be embedded in the plane. This embedding easily demonstrates that the lift $\gamma : \overline{z}_0 \to z$ to $\mathbb{U}^\infty$ of any contributing walk has $|\text{wind}(\gamma)| \leq 16\pi$. (By an abuse of notation, we continue to denote the endpoints of γ by $\overline{z}_0$ and z.) Any path γ' contributing to $G_{\mathbb{U}_n^\infty, Cn^2}(\overline{z}_0, z)$ may be concatenated at both its endpoints to the lift γ, so that a polygon $p : \overline{z}_0 \to \overline{z}_0$ in $\mathbb{U}_k^\infty$ results. This polygon has length at most $|\Lambda_n^8| + 1 + Cn^2 \leq C_1 n^2$ for some constant C_1. The application $(\gamma, \gamma') \to p$ is injective, because p crosses $\pi_\infty^{-1}(\partial\Lambda_n)$ exactly twice, at $\overline{z}_0$ and z. What we have learned is that

$$G_{\mathbb{U}_k^\infty, C_1 n^2}(\overline{z}_0, \overline{z}_0) \geq \sum_{z \in \partial\Lambda_n^8} G_{\Lambda_n^8 \setminus \Lambda_k^8}(\overline{z}_0, z) G_{\mathbb{U}_n^\infty, Cn^2}(\overline{z}_0, z),$$

where our notational abuse permits $\overline{z}_0$ and z to refer both to certain points in the eight-fold cover and to their counterparts in the universal cover. Applying (7.9), we find that

$$C_0 \left(\max_{z_1 \in \partial \Lambda_{k+1}^\infty} G_{\mathbb{U}_k^\infty, C_1 n^2}(z_1, z_1) \right)^{6(k+1)} G_{\mathbb{U}_k^\infty, C_1 n^2}(\overline{z}_0, \overline{z}_0)$$

$$\geq \sum_{z \in \partial \Lambda_n^8} G_{\Lambda_n^8}(\overline{z}_0, z) G_{\mathbb{U}_n^\infty, Cn^2}(\overline{z}_0, z).$$

By Lemmas 7.2 and 7.3, this right-hand side is at least $c\, n^{-103} \cdot 8^{-1} \cos\left(3\pi/16\right) G_{\Lambda_n}(z, z_0)$. Since $\partial \Lambda_{k+1}^\infty \cup \{\overline{z}_0\} \subseteq \mathbb{U}_k^\infty$, we obtain Corollary 7.4 with $C(k) = c^{-1} C_0(k) \cdot 8 \cos\left(3\pi/16\right)^{-1}$. $\square$

7.7. **Proof of Lemma 7.5.** Our presentation will be fairly detailed but inexplicit regarding certain more minor aspects.

It is straightforward that $\mu(\mathbb{U}_l^\infty)$, defined in (1.4), decreases in l to a limit that is at least $\mu(\mathbb{H})$. We suppose that $k \in \mathbb{N} \cup \{0\}$ satisfies $\mu(\mathbb{U}_k^\infty) = \mu(\mathbb{H})$; and we may thus harmlessly suppose that k is even.

The positive y-axis meets edge midpoints in $\mathbb{H}$ at y-coordinates of the form $2^{-1}\sqrt{3}(1 + 2l)$, where $l \in \mathbb{N} \cup \{0\}$. Note that $\left(0, 2^{-1}\sqrt{3}(1 + k)\right)$ is the only element in this set of points that lies in $\partial \Lambda_k$.

Set $z_0 = \left(0, 2^{-1}\sqrt{3}(1 + k), 0\right)$. We embed $\mathbb{U}_k^\infty$ in $\mathbb{R}^3$ so that the vertical coordinate of any point $z \in \mathbb{U}_k^\infty$ is the winding around the singularity of any path $\gamma : z_0 \to z$ in $\mathbb{U}_k^\infty$.

The cover $\mathbb{U}_k^\infty$ lacks the translation invariance of $\mathbb{H}$, so that two self-avoiding polygons in $\mathbb{U}_k^\infty$ cannot in general be concatenated; thus, the standard method which permits an upper bound on the number of polygons on $\mathbb{H}$ is not directly available. However, $\mathbb{U}_k^\infty$ has a rotational $\mathbb{Z}$-invariance, and thus pairs among a certain class of polygons can be joined. We are about to make a definition in this regard, but first we mention that, in the present proof, we abuse the notation of Section 6 and interpret polygons as unrooted objects: a polygon here can be said to be a polygon trace as understood elsewhere in Sections 6 and 7.

A polygon in $\mathbb{U}_k^\infty$ is (i, j)-*good* if it has a unique lowest edge, and this edge is centred on $2^{-1}\sqrt{3}(0, i, 0)$; and it has a unique highest edge, which is the rotation of the edge centred at $2^{-1}\sqrt{3}(0, j, 0)$ by an angle in $\pi/3 \cdot \mathbb{N}$. Here, and later, highest and lowest refer to the third coordinate in $\mathbb{R}^3$; note that i and j must be odd for such a polygon to exist. Let $g_n(i, j) = g_n(i, j; k)$ be the number of (i, j)-good polygons of length n. Then, for every i, j, n, l, we have that $g_{n+l-2}(i, i) \geq g_n(i, j) g_l(j, i)$. It is a simple consequence of the $\mathbb{Z}$-invariance of $\mathbb{U}_k^\infty$ that $g_n(i, j) = g_n(j, i)$. By Fekete's lemma applied to the sequence $g_n'(i, i) = g_{n+2}(i, i)$, we deduce that, for every odd i and j,

$$g_{n+2}(i, j)^{1/n} \leq \limsup_N g_N(i, i)^{1/N}. \tag{7.10}$$

Recall the specification of $\mu(\mathbb{U}_k^\infty)$ in (1.4). Each polygon is a walk and thus $g_N(i, i) \leq \sup_{v \in \mathbb{U}_k^\infty} \left| \mathsf{SAW}_N(\mathbb{U}_k^\infty, v) \right|$. Thus is the right-hand side of (7.10) seen to be at most $\mu(\mathbb{U}_k^\infty)$.

Since $\mu(\mathbb{U}_k^\infty) = \mu(\mathbb{H})$, we obtain that, for $n \in \mathbb{N}$, and $i, j \in \mathbb{N}$ odd,

$$g_{n+2}(i,j)^{1/n} \leq \mu(\mathbb{H}) . \tag{7.11}$$

Now that we have bounded the number of good polygons, we wish to bound $G_{\mathbb{U}_k^\infty, n}(z, z)$ for every $z \in \mathbb{U}_k^\infty$. The idea will be to transform a polygon containing z into a good polygon by concatenating polygons at its ends.

The trace of a walk $\gamma : z \to z$ that contributes to the sum $G_{\mathbb{U}_k^\infty, n}(z, z)$ is a polygon p containing z of length at most n. We may index contributions by such polygons, there being two walks γ in correspondence to each polygon p. Those p that contain a point at a distance greater than $n/2$ from the centre of the singular hexagonal face are planar polygons that contain the given point z (because they are too short to wind around the singularity); the contribution that the corresponding γ make to $G_{\mathbb{U}_k^\infty, n}(z, z)$ is thus at most $2n$.

Thus, we may suppose that the contributing $p : z \to z$ is contained in $\Lambda_{n+3k/2}^\infty$. We now rotate p so that its lowest point has height between zero and $\pi/3$. Write ℓ for the line $\{(0, x, 0) : x \geq 0\}$ and write ℓ' for the line at height $\pi/3$ given by a counterclockwise rotation of ℓ by $\pi/3$ about the centre of the singular hexagon. Let $u = (u_1, u_2, \pi/3)$ denote the point on $\ell' \cap p$ whose distance from this centre is the greatest. Note that $|u_1|$ is the planar distance, measured using the first two coordinates, between u and the line ℓ.

Suppose for now that $|u_1| \geq 3k/2$. A polygon in $\mathbb{H}$ is *useful* if its width is at least $2^{-1}\sqrt{3}\, n$; and, writing e for the most northerly among the rightmost (and thus vertical) edges in the polygon, there exists a vertex in the polygon whose x-coordinate is $|u_1| - 3k/2$ less than e's and whose y-coordinate is at most that of the lower of e's endpoints. A convenient class of useful polygons is offered by the set $\mathsf{RSAP}^{(n)}$ from Section 6.3: any given element of this set is either useful or its vertical reflection is.

Let q denote a useful polygon, formed, by vertical reflection if need be, from an element of $\mathsf{RSAP}^{(n)}$. We now join to the rightmost edge e of q the blue arrow-shaped polygon that is depicted in Figure 11. The result will be called q'. When q' is embedded in $\mathbb{U}_k^\infty$ and translated so that its rightmost edge lies on the line ℓ, the underside of the added blue polygon fits the contour that runs from $\ell \cap \partial\Lambda_k$ along $\partial\Lambda_k$ until $\ell' \cap \partial\Lambda_k$ and further along ℓ' until the contour recovers its original y-coordinate.

Since $|u_1| \geq 3k/2$, u lies on ℓ' beyond the end of the contour we have just described. The translate of q' is next pushed away from the singularity in the direction of the line ℓ, so that it is more advanced in this direction than is p. We then return this translate back along this line until it has distance two from p. This condition will be achieved before q' enters Λ_k: indeed, because q is useful, $|u_1| \geq 3k/2$ and the inserted blue polygon has width $3k/2$, the translate will meet the point u before it enters Λ_k.

In fact, the same circumstance will arise when $|u_1| < 3k/2$, whatever the definition of a useful polygon that is adopted in this case, because the blue underside of the polygon added on the right will contact u before entering Λ_k.

When p and the translate of q' have such a relative displacement, a hexagon h may be found that borders both of these polygons, with contact with a given one of the polygons occurring
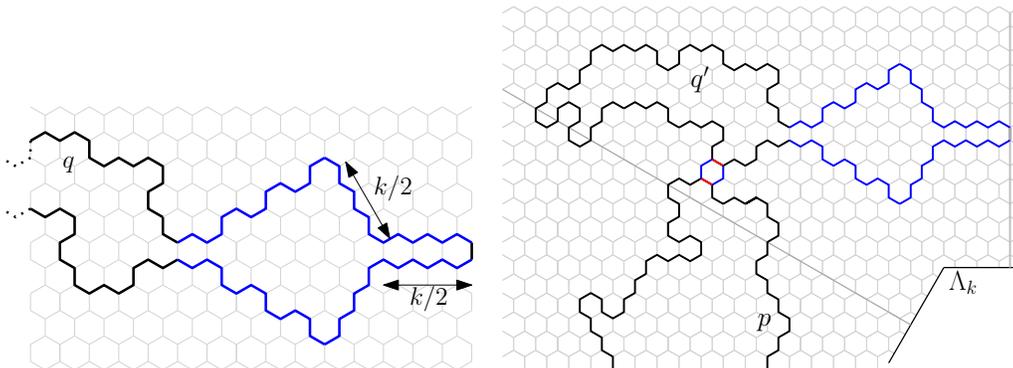
FIGURE 11. *Left:* Construct q' by joining to the rightmost edge of q the blue polygon shaped as an arrow, and possibly reflecting vertically. *Right:* The polygon q' is manoeuvred along ℓ into a position whereby it may be joined to p via the indicated hexagon. In this instance, the join occurs by the removal of the two red edges and the addition of the four blue ones.

along consecutive edges in h. Using the exclusive-or operation on h, we join p and the translate of q'.

The resulting polygon $\overline{p}$ has a length that differs from the sum of the constituents' by -2, 0 or 2. This polygon enjoys one of the two defining features of a good polygon, in that its least height is attained by a unique edge at height zero. In order to produce a good polygon, a similar circumstance should be effected at the greatest height. This is ensured by a symmetric construction, where ℓ and ℓ' are now specified so that the sector between these line segments includes the highest point in $\overline{p}$, and with the useful polygon and its enhancement called $\bar{q}$ and $\bar{q}'$.

The polygon that results from this double surgery has length at most $|p|+|q'|+|\bar{q}'|+8 \leq C'n^2$, where C' may be chosen independently of n and k (provided that $n \geq k$). Given the result of such a procedure, the values of p, q and $\bar{q}$ may be retrieved given the join locations and the starting points of these three polygons. Thus, the map sending $(\gamma, q, \bar{q})$ to the resulting good polygon is $C_0 n^{12}$-to-one for some constant C_0. We deduce that

$$\tfrac{1}{2}\,\overline{G}_{\mathbb{U}_k^\infty,n}(z_0,z_0) \cdot \left(\tfrac{1}{2} \sum_{\gamma \in \mathsf{SAP}^{(n)}} x^{|\gamma|} \right)^2 \leq C_0\, n^{12} \sum_{i,j,m \leq C'n^2} g_m(i,j)\, x^m,$$

where recall that $x = \mu(\mathbb{H})^{-1}$, and where the bar over G on the left-hand side is used to indicate that contributing p are contained in $\Lambda_{n+3k/2}^\infty$.

Applying (6.5) and Proposition 6.1, we obtain

$$\overline{G}_{\mathbb{U}_k^\infty,n}(z_0,z_0) \leq C_0\, n^{26} \sum_{i,j,m \leq C'n^2} g_m(i,j)\, x^m$$

after a suitable k-independent increase in the value of C_0. By (7.11), we see that, with a further such increase, $\overline{G}_{\mathbb{U}_k^\infty,n}(z_0,z_0) \leq C_0 n^{32}$. Recalling that we noted earlier that $G_{\mathbb{U}_k^\infty,n}(z_0,z_0) - \overline{G}_{\mathbb{U}_k^\infty,n}(z_0,z_0) \leq 2n$, we obtain Lemma 7.5. $\square$

7.8. **Proof of Proposition 7.1 in the general case.** The defining role of the edge midpoint o in the proof of Proposition 7.1 with $z = o$ is that this point lies on the edge of the singular hexagon, containing the origin, that specifies the covers $\mathbb{U}^8$ and $\mathbb{U}^\infty$. This property is used in the proof of Lemma 7.2. The general case, where z denotes the midpoint of any internal edge in Λ_{n-k}, may lack this feature. A device which permits our statements to remain largely unchanged respecifies the covers so that the singular hexagon has z at its boundary. Now, $\partial\Lambda_n$ is off-centre relative to the singular hexagon; but Λ_n continues to contain this hexagon, so that the statement and proof of Lemma 7.2 are unaffected. A change is needed in Corollary 7.4: in this statement, $\mathbb{U}_k^\infty$ must be interpreted as the infinite component of $\mathbb{U}^\infty$ after the removal of the polygons at a distance at most k from the singular hexagon, rather than from the hexagon containing the origin. Since z is an internal edge midpoint in Λ_{n-k}, the boundary $\partial\mathbb{U}_k^\infty$ remains in the set of internal-edge midpoints of Λ_n^∞. This permits the proof of the corollary in its new guise to proceed. $\qquad\square$

Remark. The proof of Lemma 7.5 succeeds under the hypothesis that, for some $k \in \mathbb{N}$, the number of (i,j)-good polygons verifies $\limsup_N g_N(i,i;k)^{1/N} \le \mu(\mathbb{H})$ for every odd i. Thus, Theorem 1.4 is valid when Conjecture 1, which concerns walks, is replaced by the last condition, which merely concerns polygons.

Appendix A. Polygon insertion for the hexagonal lattice

Our goal is to prove Theorem 1.3 by means of Proposition 3.2. To do this, we will introduce a join procedure on $\mathbb{H}$; record its pertinent properties; and explain perturbations needed to the proof in Sections 4 and 5 of Theorem 1.2. Three subsections perform these respective tasks.

A.1. **A counterpart for $\mathbb{H}$ of Madras' join procedure.** Here, we present an $\mathbb{H}$-version of the procedure reviewed in Section 3.2.

We continue to adopt the notation $y_{\min}$ and $y_{\max}$. Let γ be a self-avoiding bridge and P be a polygon on $\mathbb{H}$ such that

$$\llbracket y_{\min}(P) - 3, y_{\max}(P) + 3 \rrbracket \subseteq \llbracket y_{\min}(\gamma), y_{\max}(\gamma) \rrbracket. \tag{A.1}$$

Let $F(\gamma)$ and $F(P)$ are the faces of $\mathbb{H}$ adjacent to at least one edge in γ and P respectively. Call a face of $F(\gamma)$ good if there are at most two edges of γ adjacent to it and, in the case where there are two such edges, these two are also adjacent to each other. The same applies to $F(P)$.

Slide P horizontally to its rightmost position such that $F(\gamma)$ and $F(P)$ at a distance four: that is, the shortest path of adjacent faces from $F(\gamma)$ to $F(P)$ contains five faces. This may be done because moving P one unit horizontally changes the distance between $F(\gamma)$ and $F(P)$ by at most one.

Let $f_0, \ldots, f_4$ be a path of adjacent faces with $f_0 \in F(\gamma)$ and $f_4 \in F(P)$. Notice that f_0 and f_4 may be chosen to be good. Indeed, given f_1, f_2, f_3, there is always a good neighbour of f_1 in $F(\gamma)$; and similarly for f_3. Moreover, by the minimality of the distance, all faces neighbouring f_2 are adjacent to neither γ nor P.
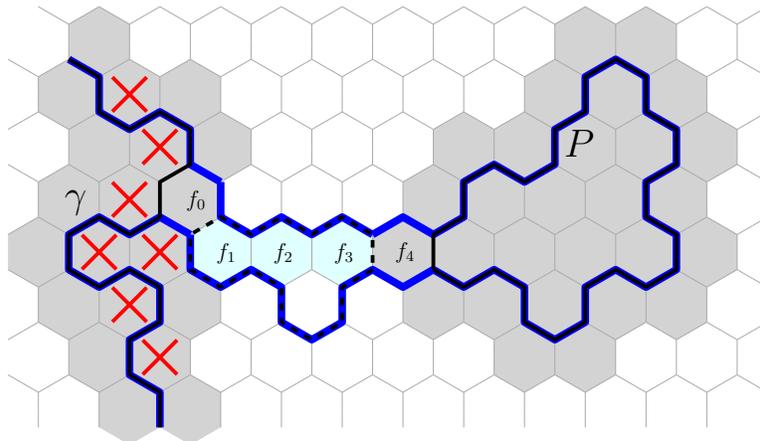
FIGURE 12. The join $M(\gamma, P)$ is emboldened. The grey faces are part of $F(\gamma)$ and $F(P)$; the bad faces of γ are marked. The choice of f_0 and f_4 may be such that both are good. The polygon Q is dotted.

Let k_γ be the number of edges of γ adjacent to f_0, and let k_P be the number of edges of P adjacent to f_4. Set $K = k_\gamma + k_P$. The join of γ and P is then defined by using the exclusive-or operation xor:

$$M(\gamma, P) := \gamma \, \mathsf{xor} \, f_0 \, \mathsf{xor} \, Q \, \mathsf{xor} \, f_4 \, \mathsf{xor} \, P \,,$$

where, in an abuse of notation, f_0 and f_4 denote the polygons of length six that respectively surround the faces actually denoted by f_0 and f_4; and where Q is a polygon whose interior contains f_2 and whose form depends on the position of f_0 and of f_4 relative to f_2, as well as on K, in a manner that is indicated in Figure 13. An example appears in Figure 12.

A.2. **Pertinent properties of the join on $\mathbb{H}$.** Lemma 3.4 records all required features of the Madras join. We need a counterpart for $\mathbb{H}$, but it would be unduly repetitive to explicitly state the analogous lemma.

Rather, we focus on essentials. First note that, in view of the choice of $f_0, \ldots, f_4$ and of Q, the join $M(\gamma, P)$ is a bridge of length

$$|M(\gamma, P)| = |\gamma| + |P| + |Q| + 12 - 2K - 4 = |\gamma| + |P| + 18 \,;$$

so the change $16 \to 18$ is made from $\mathbb{Z}^2$ to $\mathbb{H}$. The role of the junction plaquette is now assumed by a *junction hexagon*, which we may set equal to f_3. The two adjacent vertices of f_3 whose intervening edge borders the grey hexagon in P indicated in Figure 13 have chemical distance along P (excluding this edge) equal to $|P| - 1$.

A.3. **Perturbations to proofs.** Assisted by fairly clear definitional changes, the $\mathbb{H}$-course of the proof of Theorem 1.3 offered by Sections 4 and 5 runs rather unruffled in its $\mathbb{Z}^2$-tracks. The change $29 \to 10$ occurs in the lower bound on α due to the use of Proposition 3.2 in place of Proposition 3.1. Since Proposition 3.2 holds for all $u \in \mathbb{N}$, every $n \in \mathbb{N}$ verifies (4.3); or rather the variant of (4.3) obtained by the change $4^{-1} \to c_0$. The $\mathbb{H}$-counterpart to Proposition 4.1(1) thus offers an upper bound on $|\mathsf{HSW}_n|$ for all n. This permits the counterpart to

Position of f_0 and f_4 with respect to f_2	$K = 2$	$K = 3$	$K = 4$
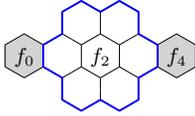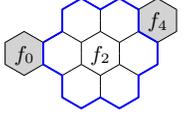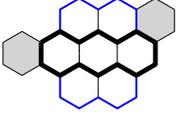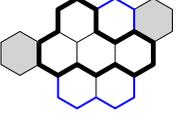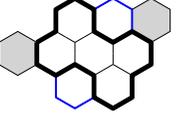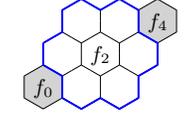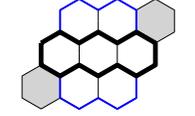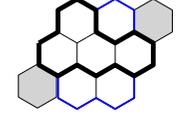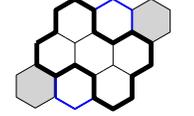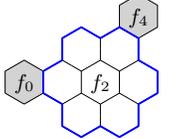
FIGURE 13. The choice of Q, which is depicted in bold, is determined by K and by the position of f_0 and f_4 relative to f_2.

Proposition 4.1(2) in which conditions (4.1) and (4.2) are used, rather than the strengthened conditions needed in the $\mathbb{Z}^2$ case. The $\mathbb{H}$-Proposition 4.1(2) follows directly from (2.1). The bound $\varepsilon < (42)^{-1}$ in Theorem 1.3 arises by choosing $\alpha = \alpha(\varepsilon) > 10$ small enough that (4.2) with $\delta = 2\varepsilon$ holds.

In regard to the near injectivity of Φ, the value $L = L(\mathbb{H})$ in Lemma 5.3 equals three, whereas $L(\mathbb{Z}^2) = 12$. Indeed, it may be argued by examining the cases that, when $M(\gamma, P)$ and the junction hexagon are known, there is only one compatible choice of (γ, P), whereas before there were possibly as many as four.

REFERENCES

[BBMdG+14] Nicholas R. Beaton, Mireille Bousquet-Mélou, Jan de Gier, Hugo Duminil-Copin, and Anthony J. Guttmann. The critical fugacity for surface adsorption of self-avoiding walks on the honeycomb lattice is $1 + \sqrt{2}$. *Comm. Math. Phys.*, 326(3):727–754, 2014.

[BDCGS12] Roland Bauerschmidt, Hugo Duminil-Copin, Jesse Goodman, and Gordon Slade. Lectures on self-avoiding walks. In *Probability and statistical physics in two and more dimensions*, volume 15 of *Clay Math. Proc.*, pages 395–467. Amer. Math. Soc., Providence, RI, 2012.

[DCGPS17] H. Duminil-Copin, A. Glazman, R. Peled, and Y. Spinka. Macroscopic loops in the loop $O(n)$ model at Nienhuis' critical point. *arXiv:1707.09335*, 2017.

[DCH13] H. Duminil-Copin and A. Hammond. Self-avoiding walk is sub-ballistic. *Commun. Math. Phys.*, 324(2):401–423, 2013.

[DCS12] Hugo Duminil-Copin and Stanislav Smirnov. The connective constant of the honeycomb lattice equals $\sqrt{2 + \sqrt{2}}$. *Ann. Math. (2)*, 175(3):1653–1665, 2012.

[GM17]   Alexander Glazman and Ioan Manolescu. Self-avoiding walk on $\mathbf{Z}^2$ with Yang-Baxter weights: universality of critical fugacity and 2-point function. *arXiv:1708.00395*, 2017.

[Ham18]   Alan Hammond. An upper bound on the number of self-avoiding polygons via joining. *Ann. Probab.*, 46(1):175–206, 01 2018.

[HR17]   Godfrey Harold Hardy and Srinivasa Ramanujan. Asymptotic formul for the distribution of integers of various types. *Proc. London Math. Soc.*, 2(1):112–132, 1917.

[HS92]   Takashi Hara and Gordon Slade. Self-avoiding walk in five or more dimensions. I. The critical behaviour. *Comm. Math. Phys.*, 147(1):101–136, 1992.

[Hut18]   Tom Hutchcroft. The Hammersley-Welsh bound for self-avoiding walk revisited. *Electron. Commun. Probab.*, 23:8 pp., 2018.

[HW62]   J. M. Hammersley and D. J. A. Welsh. Further results on the rate of convergence to the connective constant of the hypercubical lattice. *Quart. J. Math. Oxford Ser. (2)*, 13:108–110, 1962.

[Kes63]   Harry Kesten. On the number of self-avoiding walks. *J. Math. Phys.*, 4:960–969, 1963.

[Kes64]   Harry Kesten. On the number of self-avoiding walks. II. *J. Math. Phys.*, 5:1128–1137, 1964.

[LSW04]   Gregory F. Lawler, Oded Schramm, and Wendelin 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.

[Mad91]   Neal Madras. Bounds on the critical exponent of self-avoiding polygons. In *Random walks, Brownian motion, and interacting particle systems*, volume 28 of *Progr. Probab.*, pages 359–371. Birkhäuser Boston, Boston, MA, 1991.

[MS93]   Neal Madras and Gordon Slade. *The self-avoiding walk*. Probability and its Applications. Birkhäuser Boston Inc., Boston, MA, 1993.

[MS13]   Neal Madras and Gordon Slade. *The self-avoiding walk*. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2013. Reprint of the 1993 original.

[Nie82]   Bernard Nienhuis. Exact Critical Point and Critical Exponents of O(n) Models in Two Dimensions. *Physical Review Letters*, 49(15):1062–1065, 1982.

[Sla89]   Gordon Slade. The scaling limit of self-avoiding random walk in high dimensions. *Ann. Probab.*, 17(1):91–107, 1989.

[Smi10]   Stanislav Smirnov. Discrete complex analysis and probability. In *Proceedings of the International Congress of Mathematicians. Volume I*, pages 595–621, New Delhi, 2010. Hindustan Book Agency.

Institut des Hautes Études Scientifiques and University of Geneva, Le Bois-Marie 35 route de Chartres, 91440 Bures-sur-Yvette, France. Member of SwissMAP.

E-mail address: duminil@ihes.ch, hugo.duminil@unige.ch

Department of Statistics, U.C. Berkeley, Berkeley, CA, 94720-3840, U.S.A.

E-mail address: sganguly@berkeley.edu

Departments of Mathematics and Statistics, U.C. Berkeley, Berkeley, CA, 94720-3840, U.S.A.

E-mail address: alanmh@stat.berkeley.edu

Département de mathématiques, Université de Fribourg, Chemin du Musée 23, CH-1700 Fribourg, Switzerland. Member of SwissMAP.

E-mail address: ioan.manolescu@unifr.ch