The Annals of Probability
2020, Vol. 48, No. 1, 1–52
https://doi.org/10.1214/18-AOP1326

DIMERS AND IMAGINARY GEOMETRY

BY NATHANAËL BERESTYCKI[1], BENOÎT LASLIER[2] AND GOURAB RAY[3]

[1] *Fakultät für Mathematik, Universität Wien, nathanael.berestycki@univie.ac.at*
[2] *Université Paris Diderot, UFR de Mathématiques, laslier@math.univ-paris-diderot.fr*
[3] *Department of Mathematics and Statistics, University of Victoria, gourabray@uvic.ca*

We show that the winding of the branches in a uniform spanning tree on a planar graph converge in the limit of fine mesh size to a Gaussian free field. The result holds assuming only convergence of simple random walk to Brownian motion and a Russo–Seymour–Welsh type crossing estimate, thereby establishing a strong form of universality. As an application, we prove universality of the fluctuations of the height function associated to the dimer model, in several situations.

The proof relies on a connection to imaginary geometry, where the scaling limit of a uniform spanning tree is viewed as a set of flow lines associated to a Gaussian free field. In particular, we obtain an explicit construction of the a.s. unique Gaussian free field coupled to a continuum uniform spanning tree in this way, which is of independent interest.

CONTENTS

Received November 2016; revised August 2018.
MSC2010 subject classifications. 60B05, 60B99.
Key words and phrases. Dimer model, imaginary geometry, uniform spanning tree, SLE, Gaussian free field.

1. Introduction.

1.1. *Main results*. Let G be a finite bipartite planar graph. A dimer covering of G is a set of edges such that each vertex is incident to exactly one edge; in other words it is a perfect edge-matching of its vertices. The *dimer model* on G is simply a uniformly chosen dimer covering of G. It is a classical model of statistical physics, going back to work of Kasteleyn [17] and Temperley–Fisher [41] who computed its partition function. It is the subject of an extensive physical and mathematical literature; we refer the reader to [22] for a relatively recent discussion of some of the most important progress. A key feature of this model is its "exact solvability" which comes from its determinantal structure [17] and brings in tools from subjects such as discrete complex analysis, algebraic combinatorics and algebraic geometry. This is one reason the study of this model has been so successful.

An important tool for the dimer model is a notion of height function introduced by Thurston [42] which turns a dimer configuration into a discrete random surface in $\mathbb{R}^3$ (i.e., a random function indexed by the faces of G with values in $\mathbb{R}$). Therefore a key question concerns the large-scale behaviour of this height function. It is widely believed that in the planar case and under very general assumptions, the fluctuations of the height function are described by (a variant of) the *Gaussian free field* (see Figure 1).

In this paper, we present a robust approach to proving such results. We now state an example of application of this technique. Consider lozenge tiling of the plane by lozenges with angles $\pi/3$ and $2\pi/3$ and sidelength δ, which can be seen equivalently as dimer configurations on the hexagonal lattice of mesh size δ, or stack of cubes in $\mathbb{R}^3$ of size δ. As usual,

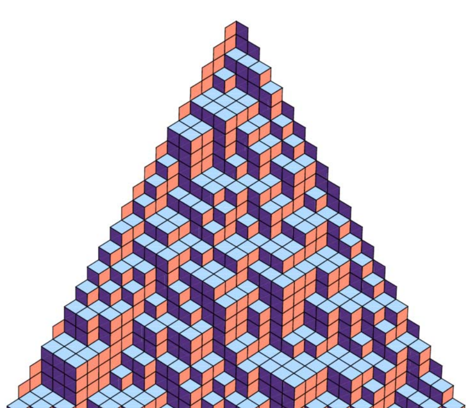

FIG. 1. *Height function of a dimer model* (*or lozenge tiling*) *with planar boundary conditions on a triangle. Picture by R. Kenyon.*

we describe a tiling by its height function $h^{\#\delta}$ which we can take to be the z coordinate in the stack of cubes at each point of the tiling. Given a bounded domain that can be tiled with lozenges, we define the *boundary height* to be the curve in $\mathbb{R}^3$ obtained by considering the height function along the outermost lozenges (which does not depend on the tiling configuration). Set $\chi = 1/\sqrt{2}$ (this is the parameter in imaginary geometry associated with $\kappa = 2$; see [33, 34]).

THEOREM 1.1. *Let P be a plane in $\mathbb{R}^3$ whose normal vector has positive coordinates, and let $D \subset \mathbb{R}^2$ be a simply connected bounded domain with locally connected boundary. Then there exists a sequence of domains $U^{\#\delta} \subset \mathbb{R}^2$, which can be tiled by lozenges of size δ, with the following properties. The boundary height of $\partial U^{\#\delta}$ stays at distance $o(1)$ of P, $\partial U^{\#\delta}$ converges to ∂D in Hausdorff sense, and*

$$\frac{h^{\#\delta} - \mathbb{E}(h^{\#\delta})}{\delta} \xrightarrow[\delta \to 0]{} \frac{1}{2\pi\chi} h^0_{\mathrm{GFF}} \circ \ell,$$

in distribution where ℓ is an explicit linear map determined by P and h^0_{GFF} is a Gaussian free field with Dirichlet boundary conditions in $\ell(D)$.

Note that convergence holds in distribution on the Sobolev space $H^{-1-\eta}(D)$ for all $\eta > 0$, once $h^{\#\delta}$ has been extended to a continuous function on D (essentially by interpolation); see Section 5.2 for more details. We emphasise here that in Theorem 1.1 above we only prove the existence of a sequence of domains $U^{\#\delta}$ such that the result holds. See Section 4.2 of [3] for details of the construction of $U^{\#\delta}$.

Theorem 1.1 is the consequence of a more general theorem (Theorem 1.2) which will be the focus of this article. The connection between these two theorems is explained in [3] and exploits a relation between the dimer model and the uniform spanning tree model on a modified graph called the T-graph introduced in [27]. More precisely, this connection is a generalisation of Temperley's celebrated bijection, which equates the height function of a dimer configuration to the winding of branches in an associated uniform spanning tree.

We now state the general theorem which concerns the winding of branches in a uniform spanning tree (see Figure 2). Let $G^{\#\delta}$ be a sequence of planar (possibly directed) graphs

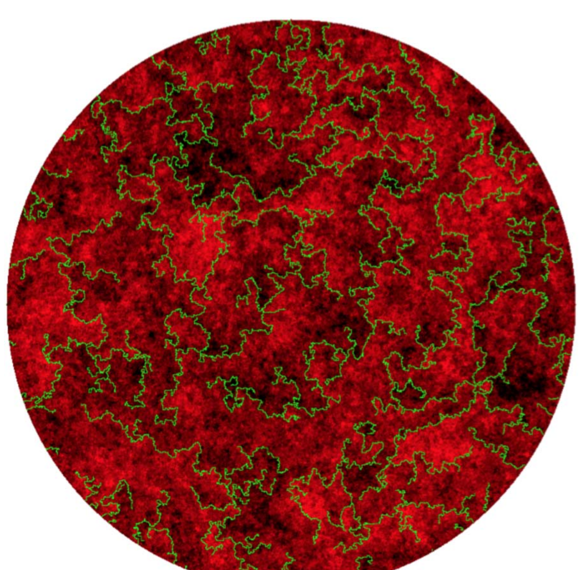

FIG. 2. *A Uniform spanning tree and its winding field.*

properly embedded in the plane. We assume that $G^{\#\delta}$ satisfies some natural conditions (stated precisely in Section 4.1). In particular, the two main assumptions are: (i) simple random walk on $G^{\#\delta}$ converges to a Brownian motion as $\delta \to 0$, and (ii) a Russo–Seymour–Welsh type crossing condition, namely, simple random walk can cross any rectangle of fixed aspect ratio and size at least δ, with a probability uniformly positive over the position, orientation and scale of the rectangle.

Let $D \subset \mathbb{C}$ be a bounded domain with locally connected boundary. Let $D^{\#\delta}$ be the graph induced by the vertices of $G^{\#\delta}$ in D with boundary $\partial D^{\#\delta}$ (the precise description is in Section 5.1). Recall also that a wired uniform spanning tree is simply the uniform spanning tree on the graph obtained from $D^{\#\delta}$ by identifying all the boundary vertices of $D^{\#\delta}$. For more details on this topic, see Section 1.3 in the Supplementary Material [4] file as well as [32] for (much) more background.

THEOREM 1.2. *Let $\mathcal{T}^{\#\delta}$ be a wired uniform spanning tree on $D^{\#\delta}$, and for any $v \in D^{\#\delta}$ let $h^{\#\delta}(v)$ denote the winding of the branch of $\mathcal{T}^{\#\delta}$ connecting v and $\partial D^{\#\delta}$. Then*

$$h^{\#\delta} - \mathbb{E}(h^{\#\delta}) \xrightarrow[\delta \to 0]{} \frac{1}{\chi} h^0_{\mathrm{GFF}}$$

in the sense of distributions, where h^0_{GFF} is a Gaussian free field with Dirichlet boundary conditions in D.

By *winding* in Theorem 1.2, we mean the intrinsic winding, that is, the sum of the turning angles along the path. See equation (2.3) for precise definition. Note that the scaling is somewhat different from Theorem 1.1 (there is no renormalisation here) because in that theorem we measure the height defined by lozenges of diameter $O(\delta)$ whereas here we measure the winding (unnormalised) along paths in the tree.

A more precise form of Theorem 1.2 is stated later on in Theorem 5.1. Furthermore, in Theorem 6.1 we prove a stronger version of this theorem: we obtain the joint convergence of the winding function *and* spanning tree to a pair (GFF, continuum spanning tree) which are coupled together according to the imaginary geometry coupling. The connection to the theory of *imaginary geometry*, initiated in [10] and further developed in a sequence of papers of which [33] and [34] will be the most relevant here, will in fact play a crucial role in this work. Very informally, imaginary geometry provides a coupling between a Gaussian free field and an SLE curve so that the "pointwise values" of the field along the curve are given by the "intrinsic winding" of the SLE curve. Hence this coupling can be viewed as a continuum analogue of Temperley's bijection, an observation already alluded to in [10]. In particular, our approach provides an explicit construction of the a.s. unique Gaussian free field associated to a continuum uniform spanning tree which may be of independent interest; see Theorem 3.1 for a statement and the discussion immediately below.

Theorem 1.2 may be applied to various other dimer models to show Gaussian free field fluctuations. We give a brief overview of such examples:

- generalised Temperleyan domains, as described in [26], on graphs which satisfy the assumptions of Section 4.1.
- dimers on double isoradial graphs with uniformly elliptic angles. This recovers and in fact significantly strengthens a result of Li [31] as her work requires the discrete boundary of the domain to contain a macroscopic straight line. (Note that the assumptions in Theorem 1.2 are satisfied in this this case by results of Chelkak and Smirnov [7]: for instance, the crossing assumption is an easy consequence of Theorem 3.10 in [7].)

- dimers in random environment: for example, on $\mathbb{Z}^2$ with random i.i.d. weights on the even edges of $\mathbb{Z}^2$ (in which case the law of the dimers is simply proportional to the product of the weights). We restrict the randomness of the weights to the even sublattice in order to apply the Temperley bijection, and assume for instance the weights to be balanced and uniformly elliptic.
- dimers with a defect line: suppose the weight of all the edges in a horizontal line of $\mathbb{Z}^2$ is changed from 1 to $z > 0$.

1.2. *Discussion of the results.*

Mean height in dimer models and spanning trees. Theorem 1.1 describes the limiting distribution of $h^{\#\delta} - \mathbb{E}(h^{\#\delta})$ and the reader might be interested to know what can be said about the mean itself, $\mathbb{E}(h^{\#\delta})$. First, we point out that on the law of large number scale, the mean height of the lozenge tiling is known by a result of Cohn, Kenyon and Propp [8] to converge to a deterministic function which here is simply an affine function (due to our assumptions about the boundary values of the height function).

Our approach yields further information about $\mathbb{E}(h^{\#\delta})$. In the spanning tree setting, if $h^{\#\delta}$ is the winding of branches in a uniform spanning tree (as in the setup of Theorem 1.2) from a fixed marked point x on the boundary then we obtain

$$\mathbb{E}(h^{\#\delta}) = m^{\#\delta} + u_{D,x} + \frac{\pi}{2} + o(1),$$

where $u_{D,x}$ is the harmonic extension of the anticlockwise winding from x (see equation (2.7) for a precise definition) and $m^{\#\delta}$ depends only on the graph and the vertex v at which we are computing the winding (but interestingly not the domain in which the spanning tree/dimer configuration is being sampled). Note that a consequence of the above mentioned result of Cohn–Kenyon–Propp [8] is that $m^{\#\delta} = o(1/\delta)$ uniformly over the graph; in fact much better bounds can be derived.

For many "reasonable" graphs we suspect that $m^{\#\delta}$ actually converges to 0, as it is essentially the expected winding of a path converging to a full-plane SLE_2. Nevertheless some assumptions are clearly needed, as the fact that random walk converges to Brownian motion alone is *not* enough to give control on the mean winding in a UST. For an example, take the usual square grid and add a spiral path at every vertex. This example shows that it is only the *fluctuations* which may be hoped to be universal, while the mean itself will usually depend on the microscopic details of the graph.

Relation to earlier results on fluctuations of dimer models. The study of fluctuations in dimer models has a long and distinguished history, which is not the purpose of this paper to recall; see [22] for references. However, we mention a few highlights. In [18, 19], Kenyon showed that the height function on the square lattice for Temperleyan domains (for which the boundary conditions are planar of slope 0) converge to a multiple of the Gaussian free field with Dirichlet boundary conditions. The study of dimers on graphs more general than the square or hexagonal lattices was initiated in [25] where they consider tilings on arbitrary periodic bipartite planar graphs. The nonperiodic case was first mentioned in [20], also in the whole plane setting. Convergence to the full plane Gaussian free field on isoradial periodic bipartite graphs (including ergodic lozenge tilings of arbitrary slope), as well as on Temperleyan superpositions of isoradial (not necessarily periodic) graphs, is a consequence of a remarkable work by De Tilière [9].

The interest in the role of boundary conditions was sparked by the observation of the arctic circle phenomenon: for some domains, in the limit the dimer configuration outside of some region (the liquid or temperate region) is deterministic (also called frozen). This was first identified in the case of the aztec diamond by Jokusch, Propp and Shor [16] (see also the more recent paper [38] by Romik for a different approach and fascinating connections to alternating sign matrices). The case of general boundary conditions for the hexagonal lattice was solved later by Cohn, Kenyon and Propp [8] who obtained a variational problem determining the law of large numbers behaviour for the height function. This variational principle was studied by Kenyon and Okounkov in [24] who discovered that in polygonal domains the boundary between the frozen and liquid regions are always explicit algebraic curves. In this direction we also point out the recent paper by Petrov [36] and by Bufetov–Gorin [6] who obtained convergence of the height function fluctuations to the GFF in liquid regions for some polygonal domains.

A paper by Kenyon [21] discusses the question of fluctuations, with the goal of proving convergence of the centered height function to a (deformation of) the Gaussian free field in the liquid region. Unfortunately, the crucial argument in his proof, Lemma 3.6, is incomplete and at this point it is unclear how to fix it.[1] The issue is the following. The central limit theorem proved in [28] provides an information about convergence of discrete harmonic functions to continuous harmonic functions. However what is needed in [21] is an estimate on the discrete derivative of such functions (i.e., the entries of the inverse Kasteleyn matrix) as well as a control on the speed of convergence so that the errors can be summed when integrating along paths. (There is a more general question here, which is to better understand the links between discrete and continuous harmonic functions on quasi-periodic graphs.) Our work can be seen as a way to get around these issues but more importantly provides a unified and robust approach to the convergence of fluctuations.

Finally, let us mention that all the above works on fluctuations rely on writing an exact determinantal formula for the correlations between dimers. The main body of work is then to find the asymptotic of the entries of these determinants using either exact combinatorics or discrete complex analytic techniques. Our approach is completely orthogonal, relying on properties of the limiting objects in the continuum rather than exact computations at the microscopic level. This is one reason why the results we obtain are valid under less restrictive conditions on the regularity of the boundary (while such assumptions are typically needed for the tools of discrete complex analysis). In particular, we do not assume the domain to be Jordan or smooth, only to have a locally connected boundary. This is the condition required so that the conformal map from the unit disc to the domain extends to the boundary (Theorem 2.1 in [37]). It is plausible that even this mild condition can be relaxed by appealing to a suitable notion of conformal boundary (e.g., prime ends; see Section 2.4 in [37]) but we did not pursue this here in an attempt to keep the paper at a reasonable length.

1.3. *A conjecture.* Theorem 3.1 provides a continuum analogue of Theorem 1.2, in the sense that the continuum field is regularised by truncating the SLE branches rather than discretisation. As already mentioned, this is of independent interest since it gives an explicit construction of the GFF coupled to a uniform spanning tree according to imaginary geometry. We strongly believe that the same result holds for other values of κ. Our proof of Theorem 3.1 is written in a way that is mostly independent of the value of κ except for a few lemmas, gathered in Section 2.3. These lemmas concern fairly basic properties of flow lines which seem very plausible for arbitrary values of κ. However, we did not try to establish them, preferring to focus on the case $\kappa = 2$ only since we also need the analogous discrete statements later on in the paper.

The above discussion suggests a number of results concerning interacting dimers recently introduced by Giuliani, Mastropietro and Toninelli [15]. We conjecture that if one applies

[1] We thank Fabio Toninelli and Rick Kenyon for helpful discussions regarding this lemma.

Temperley's bijection to a configuration of interacting dimers as in [15], the Peano curve of the resulting tree converges to certain space-filling $\mathrm{SLE}_{\kappa'}$ defined by Miller and Sheffield [34] in these cases and that by adjusting the interaction parameter one can at least obtain any $\kappa' \in (8-\varepsilon, 8+\varepsilon)$. However, it is quite speculative at the moment as we lack tools (like Wilson's algorithm) to study interacting dimers or corresponding Temperleyan spanning trees. See [14] (which appeared after a draft of our paper was first put on arxiv) for additional support for our conjecture, and see [23] for a related question.

1.4. *Overview of the proof.* For the convenience of the reader, we summarise briefly the main steps of the proof of Theorem 1.2.

Step 1. We first formulate in Theorem 3.1 a continuous analogue of this theorem, where we study the winding of truncated branches in a continuum wired Uniform Spanning Tree. Branches of this tree are SLE_2 curves, and therefore a *key idea* is to introduce a suitable notion of (intrinsic) winding. To do so we rely on a simple deterministic observation (see Lemma 2.1), which shows that the intrinsic winding of a smooth simple curve is equal to the sum of its topological winding with respect to either endpoints. After that, we prove by hand a version of the change of coordinate formula in imaginary geometry:

$$\tilde{h} \circ \varphi - \chi \arg \varphi' = h,$$

where $\varphi : D \to \tilde{D}$ is a conformal mapping, $\chi = \frac{2}{\sqrt{\kappa}} - \frac{\sqrt{\kappa}}{2}$ is the constant of imaginary geometry (note that $\chi = 1/\sqrt{2}$ for $\kappa = 2$), and h, $\tilde{h}$ are GFF with appropriate boundary conditions in the domains D, $\tilde{D}$. This equation is taken as the starting point of the theory of imaginary geometry (see, e.g., [33, 40]) but here it must be derived from the model and our definition of winding. Together with the domain Markov property of the GFF and of the continuum UST (inherited from the domain Markov property of SLE), this implies that the winding of a continuum UST is a Gaussian free field with appropriate boundary conditions.

Step 2. After Theorem 3.1 is proved, we return to the discrete UST, and we write

$$h^{\#\delta} = h_t^{\#\delta} + \epsilon^{\#\delta}, \tag{1.1}$$

where $h^{\#\delta}$ is the winding of the branches of the discrete tree, $h_t^{\#\delta}$ is the winding of the branches truncated at capacity t, and $\epsilon^{\#\delta}$ is the difference. When t is fixed and $\delta \to 0$ there is no problem in showing that $h_t^{\#\delta}$ converges to the regularised winding of the continuum UST (this follows from results of Yadin and Yehudayoff [43] and results about winding in Step 1). By Theorem 3.1 mentioned above, we also know that as $t \to \infty$, h_t converges to a GFF.

Step 3. It remains to deal with the error term $\epsilon^{\#\delta}$. The main idea for this is to construct a multiscale coupling (Theorem 4.21) with independent full plane USTs, which relies on a modification of a lemma of Schramm [39]. This allows us to show that the terms $\epsilon^{\#\delta}$ from point to point have a fixed mean and are independent of each other, even if the points come close to each other. This is enough to show that when we integrate against a test function, the contribution of these terms will vanish.

Step 4. In order to do so, we need to evaluate the moments of $h^{\#\delta}$ integrated against a test function; however this requires precise a priori bounds on the moments of the discrete winding to deal with bad events when the coupling fails. We therefore first derive a priori tail estimates on the winding of loop-erased random walks (Proposition 4.12). This is where we make use of our RSW crossing assumptions.

1.5. *Organisation of the paper.* The paper is organised as follows. In Section 2, some background and definitions are provided. In Section 3 we formulate and prove the continuum analogue of Theorem 1.2, Theorem 3.1. In Section 4, we derive the required a priori estimates on winding and describe the multiscale coupling. We put all those ingredients together in Section 5, which completes the proof of Theorem 1.2.

The paper includes fairly technical proofs related to several different areas. This makes a full account quite long. For the sake of brevity and readability, we will defer the proofs of some technical statements to a separate file containing the Supplementary Material [4].

Throughout the paper, c, C, c', C' etc. will denote constants whose numerical value may change from line to line. Arg will denote the principal branch of argument with branch cut $(-\infty, 0]$. Also all our domains are bounded unless explicitly stated.

Throughout this paper, universal constants mean constants which do not depend upon anything else in consideration. This should not be confused with our results of "universality" which is the main topic of this article.

2. Background. For background on SLE and Gaussian free field we refer the reader to Section A of the Supplementary Material [4]. Our normalisation of the Gaussian free field is such that the two point function blows up like $-\log|z-w|$ as $w \to z$.

NOTATION. For $z \in D$, we denote by $R(z, D)$ the conformal radius of z in the domain D. That is, if g is any conformal map sending D to the unit disc $\mathbb{D}$ and z to 0, then $R(z, D) = |g'(z)|^{-1}$.

2.1. *Winding of curves.* In this section, we recall simple facts about the winding of smooth curves, which we think are important motivations for the definitions we will use later. Let $\gamma : [0, 1] \to \mathbb{C}$ be a (continuous) curve. For $0 \le s < t \le 1$, we will write $\gamma[s, t]$ for the curve $\gamma|_{[s,t]}$.

Topological winding. The *topological* winding of a curve around a point $p \notin \gamma[0, 1]$ is defined as follows. We can write

$$\gamma(t) - p = r(t)e^{i\theta(t)}, \tag{2.1}$$

where the function $\theta(t) : [0, \infty) \mapsto [0, \infty)$ is taken to be continuous. We define the winding of γ around p, denoted $W(\gamma, p)$, to be $\theta(1) - \theta(0)$. We extend this definition to $p = \gamma(0)$ or $p = \gamma(1)$ by the following formulas when they make sense (i.e., the limits exist):

$$W(\gamma, \gamma(1)) = \lim_{t\to 1} W(\gamma[0, t], \gamma(1)); \qquad W(\gamma, \gamma(0)) = \lim_{s\to 0} W(\gamma[s, 1], \gamma(0)).$$

Intrinsic winding. The *intrinsic* winding of a (smooth) curve is defined as follows. Suppose that γ is continuously differentiable and $\forall t$, $\gamma'(t) \neq 0$, and write

$$\gamma'(t) = r_{\text{int}}(t)e^{i\theta_{\text{int}}(t)}, \tag{2.2}$$

where again $\theta_{\text{int}}(t) : [0, \infty) \mapsto [0, \infty)$ is taken to be continuous. We define the intrinsic winding of γ to be

$$W_{\text{int}}(\gamma) := \theta_{\text{int}}(1) - \theta_{\text{int}}(0). \tag{2.3}$$

The definition can be extended to piecewise smooth paths by summing the intrinsic winding of each smooth piece together with the jumps in between these pieces. In general, these two definitions are very different; think, e.g., of an "8" curve whose intrinsic winding is 0 while its topological winding is either -1, 0, or 1 depending on the point. For simple curves however they are related by the following topological lemma which in a sense says that the only amount of nontrivial winding that a simple curve can accumulate is near its endpoints – anything else has to be unwinded (cancelled out). Its proof can be found in the Supplementary Material [4] (Lemma B.1).

LEMMA 2.1. *Let $\gamma : [0,1] \mapsto \mathbb{C}$ be a smooth simple curve with $\gamma'(s) \neq 0$ for all s. We have*

$$W_{\text{int}}(\gamma) = W(\gamma, \gamma(1)) + W(\gamma, \gamma(0)). \tag{2.4}$$

A further important fact is that the topological winding of any path around a boundary point can only arise due to winding of the domain itself. To state this precisely, we recall the following notion of argument $\arg_{D;x}$ in D with respect to a boundary point x. This is defined so that

$$\arg_{D;x}(b) - \arg_{D;x}(a) = \Im\left(\int_p \frac{dz}{z-x}\right)$$

over any smooth path $p \subset D$ going from a to b. In other words, x is taken to be the origin and the argument is determined in a continuous way in the simply connected domain D. A priori this is defined only up to a global additive constant, whose choice for now can be made in an arbitrary way. Note that if the boundary is locally smooth at x, and if γ is a smooth path in $\bar{D}$ such that $\gamma(0) = x$ and $\gamma(0,1] \subset D$ then we can define with an abuse of notation $\arg_{D;x}(\gamma'(0))$ as $\lim_{\varepsilon\to 0} \arg_{D;x}(\gamma(\varepsilon))$, up to the same global additive constant. With these definitions, we have the following obvious lemma.

LEMMA 2.2. *Let D be a simply connected domain and let x be a fixed boundary point. Let γ be a smooth curve with $\gamma(0) = x$ and $\gamma(0,1] \in D$. We have*

$$W(\gamma, x) = \arg_{D;x}(\gamma(1)) - \arg_{D;x}(\gamma'(0)).$$

In particular, if γ is in addition simple:

$$W_{\text{int}}(\gamma) = W(\gamma, \gamma(1)) + \arg_{D;x}(\gamma(1)) - \arg_{D;x}(\gamma'(0)).$$

REMARK 2.3. We will be interested in branches of the uniform spanning trees which are rough self avoiding curves between the boundary of a fixed domain and an inside point. Furthermore in the discrete, the natural relation is between the intrinsic winding of branches (it is easily extended to piecewise smooth curves) and the height function so we want to make sense of the intrinsic winding of an SLE curve. Lemma 2.1 will be crucial because it motivates the definition of intrinsic winding for a simple curve using only regularity at the endpoint. Actually as long as we work in a fixed domain the second formula in Lemma 2.2 will allow us to think that $W_{\text{int}}(\gamma) = W(\gamma, \gamma(1))$ losing only unimportant deterministic correction terms.

We now state a lemma showing how the intrinsic winding behaves under conformal maps. This is one of the key deterministic statements used in this paper: it states that the change in winding under an application of conformal map ψ is roughly $\arg \psi'$. See Remark 2.5 below for a clean corresponding statement, which however is only valid for smooth curves.

LEMMA 2.4. *Let D, D' be bounded domains with locally connected boundary and let ψ be conformal map sending D to D'. Let $\gamma : [0,1] \mapsto \bar{D}$ be a curve in $\bar{D}$. Assume further that $\arg(\psi')$ extends continuously to $\gamma(0)$ and $\gamma(1)$. Let z be a point in $D \setminus \gamma[0,1]$ and let $R = R(z, D)$ be its conformal radius and assume that $|z - \gamma(1)| \leq R/8$. Then, letting $x = \gamma(0)$ and $x' = \psi(x)$,*

$$\begin{aligned} &W(\psi(\gamma), \psi(z)) - W(\gamma, z) \\ &\quad = \arg_{\psi'(D)}(\psi'(z)) + \arg_{D;x}(z) - \arg_{D';x'}(\psi(z)) + O(|z - \gamma(1)|/R), \end{aligned} \tag{2.5}$$

where the implicit constant in the $O(|z-\gamma(1)|/R)$ *is universal and we* choose *the global constants defining the arguments so that the chain rule holds at* $x=\gamma(0)$, *that is*,

$$\arg_{D';x'}\big((\psi\circ\gamma)'(0)\big)=\arg_{D,x}\big(\gamma'(0)\big)+\arg_{\psi'(D)}\big(\psi'(x)\big). \tag{2.6}$$

Furthermore if $\arg(\psi')$ *does not extend to* x, *the formula still holds up to a global constant in* $\mathbb{R}$ *depending on the choice of the constants for the arguments and not on* γ.

The proof of this lemma can be found in the Supplementary Material [4] (Corollary B.7). See also Lemma B.8 in the Supplementary Material [4] for a simple geometric condition guaranteeing that $\arg\psi'$ extends continuously near some fixed boundary point x: essentially all that is required, beyond local connectedness, is a bit of smoothness for ∂D locally around x. It is this condition which explains why without smoothness, the height function is only defined up to a global additive constant (see (3.2)).

REMARK 2.5. By letting $z\to\gamma(1)$, for a smooth curve γ in D, we deduce the following somewhat cleaner statement:

$$W_{\text{int}}\big(\psi(\gamma)\big)=W_{\text{int}}(\gamma)+\arg_{\psi'(D)}\big(\psi'(\gamma(1))\big)-\arg_{\psi'(D)}\big(\psi'(\gamma(0))\big),$$

where $\arg_{\psi'(D)}$ here is any determination of the argument on the image of ψ'. This is significant for the following reasons. The SLE/GFF coupling results developed by Dubédat, Miller and Sheffield [10, 34] (referred to as imaginary geometry) was defined using a change in coordinate formula under conformal map using $\arg\psi'$. Lemma 2.4 shows that this definition is consistent with the idea that along a branch, the field takes values equal to the intrinsic winding of the branch. In that setting, a key insight is that while the intrinsic winding itself does not make sense, its harmonic extension does and this is the only information needed for the GFF.

2.2. *Continuum uniform spanning tree and coupling with GFF.* The breakthrough papers of Schramm [39] followed by the paper of Lawler, Schramm and Werner [30] established, among other things, the existence and a precise description of the scaling limit of a uniform spanning tree of a domain on a square lattice. We call this limit the *continuum uniform spanning tree*. The following lemma is a consequence of their work which relies on the major result in [30] that loop erased random walk when rescaled converges to a SLE_2 curve and Wilson's algorithm (see the Supplementary Material [4], Section A.3 for background on Wilson's algorithm). For now, we state the following proposition which is a simple consequence of their work.

PROPOSITION 2.6 (Wilson's algorithm in the continuum). *Let* D *be a simply connected domain and* $z_1,\dots,z_k\in D$. *We can sample the* (*a.s. unique*) *branches of the continuum wired UST in a domain* D *from* $z_1,\dots,z_k$ *as follows. Given the branches* γ_i *from* z_i *for* $1\le i<j$, *we inductively sample the branch from* z_j *as follows. We pick a point* p *from the boundary of* $D':=D\setminus\bigcup_{1\le i<j}\gamma_i$ *according to harmonic measure from* z_j *and draw a radial* SLE_2 *curve in* D' *from* p *to* z_j. *The joint law of the branches does not depend on the order in which we sample the branches.*

Readers interested in a more precise exposition are referred to Section A.4 of the Supplementary Material [4].

Coupling with a GFF. Let D be a simply connected domain whose boundary β is a smooth closed curve and let x be a marked point in the boundary of the domain. Let us parametrise the boundary β of D in an anticlockwise direction (meaning that D lies to left of the curve) and such that $\beta(0) = x$. We define *intrinsic winding boundary condition* on (D, x) to be a function u defined on the boundary by $u_{(D,x)}(\beta(t)) := W_{\text{int}}(\beta[0, t])$. We call $u_{(D,x)}$ the intrinsic winding boundary function and extend it harmonically to D.

We extend this definition to any simply connected domain D smooth in a neighbourhood of a marked point x (but not necessarily smooth elsewhere on the boundary and possibly unbounded) as follows. Let $\varphi : \mathbb{D} \to D$ be a conformal map which maps x to 1. Let $u_{(\mathbb{D},1)}$ be the intrinsic winding boundary function on $(\mathbb{D}, 1)$. Define $u_{(D,x)}$ on D by

$$u_{(\mathbb{D},1)} = u_{(D,x)} \circ \varphi - \arg_{\varphi'(\mathbb{D})} \varphi', \tag{2.7}$$

where we define $\arg_{\varphi'(\mathbb{D})}$ as the argument defined continuously in $\varphi'(\mathbb{D})$ (note $\varphi'(\mathbb{D})$ do not contain 0 since φ is conformal) with the global constant chosen such that $u_{(D,x)}$ jumps from 2π to 0 at x. One can check that this choice is such that $\arg_{\varphi'(\mathbb{D})} \varphi'$ verifies the chain rule at x as in (2.6). It is elementary but tedious to check that this definition is unambiguous in the sense that it does not in fact depend on the choice of the conformal map φ: indeed, if one applies a Möbius transform of the disc, winding boundary conditions are changed into winding boundary conditions.

REMARK 2.7. We can still define $u_{(D,x)}$ up to a global constant for domains with general boundary.

THEOREM 2.8 (Imaginary geometry coupling). *Let D be a simply connected domain with a marked point x on the boundary and let $\chi = \frac{1}{\sqrt{2}}$. Let $h = \chi u_{(D,x)} + h^0_D$ where h^0_D is a GFF with Dirichlet boundary conditions in D and $u_{(D,x)}$ is defined as in* (2.7) *and Remark* 2.7. *There exists a coupling between the continuum wired UST on D and h such that the following is true. Let $\{\gamma_i\}_{1\le i\le k}$ be the branches of the continuum wired UST from points $\{z_i\}_{1\le i\le k}$ in D and let $D' = D \setminus \bigcup_{1\le i\le k} \gamma_i$. Then the conditional law of h given $\{\gamma_i\}_{1\le i\le k}$ is the same as $\chi u_{(D',x)} + h^0_{D'}$ where $h^0_{D'}$ is a GFF with Dirichlet boundary condition in D'. Furthermore, h is completely determined by the UST and vice-versa.*

The proof of Theorem 2.8 is implicit in [10, 30, 33, 34]. We provide a detailed proof in Section C (Theorem C.2) of the Supplementary Material [4].

2.3. SLE_2 *estimates.* In this section, we gather some estimates purely about SLE which are needed for Section 3. We note that these estimates are the only place in Section 3 where we need to restrict ourself to $\kappa = 2$. These lemmas are no doubt true for SLE_κ curves with $\kappa \in (0, 8)$ and seem fairly well known in the folklore; however we could not find precise references. Since in any case we will need the corresponding discrete statement for loop-erased random walk, we prefer to provide discrete proofs and deduce the continuum statements below from the known convergence of loop-erased random walk to SLE_2 [30]. Since these are the only estimates specific to the case $\kappa = 2$, Theorem 3.1 extends immediately to other values of κ if the statements below are generalised to the corresponding flow lines. (Note however that when $\kappa \neq 2$ flow lines are not simply SLE_κ curves but rather specific types of $\text{SLE}_\kappa(\rho)$ with marked points.)

The first estimate controls the probability that an SLE targeted to a point w comes close to another point z in a uniform way and follows from Proposition 4.11, Lemma 4.17 and [30].

LEMMA 2.9. *Let D be a domain in $\mathbb{C}$. There exists a universal constant $c_0 > 0$ such that the following holds. Let $z, w \in D$ and let γ_w be a radial* SLE_2 *started from a point on the boundary picked according to harmonic measure from w and targeted at w. Let $r = |z - w| \wedge \mathrm{dist}(z, \partial D) \wedge \mathrm{dist}(w, \partial D)$. Then for all $0 < \varepsilon < 1/4$,*

$$\mathbb{P}(|\gamma_w - z| < \varepsilon r) < \varepsilon^{c_0}.$$

There also exists absolute constants c, c' such that if $r' := \mathrm{dist}(w, \partial D) < \mathrm{Diam}(D)/10$, then for all $R > r'$

$$\mathbb{P}(\gamma_w \subset B(w, R)) \geq 1 - c\left(\frac{r'}{R}\right)^{c'}.$$

When we work in general domains with possibly rough boundaries, we also need *a priori* bounds on moments of the winding, which follow directly from Proposition 4.12 and [30].

LEMMA 2.10. *Let D be a simply connected domain, and let $z \in D$. Let γ_z be radial* SLE_2 *towards z, started from a point chosen according to harmonic measure on ∂D viewed from z. There exist constants $C, c > 0$ such that the following holds. For all $t \geq 0$ and $n \geq 1$,*

$$\mathbb{P}\Big(\sup_{t \leq t_1, t_2 \leq t+1} \big|W(\gamma_z[0, t_1], z) - W(\gamma_z[0, t_2], z)\big| > n\Big) < Ce^{-cn}.$$

In a reference domain such as the unit disc, the winding of a single SLE branch has been studied extensively starting with the original paper of Schramm [39] itself. In particular, Schramm obtained the following result, which will be used to say that arbitrary moments of the winding at a fixed point blow up at most logarithmically.

THEOREM 2.11 ([39], Theorem 7.2). *Suppose D is the unit disc and let γ_0 be a radial* SLE_2 *to 0 started from a point chosen according to the harmonic measure (which is just the uniform measure in this case). We have the following equalities in law*

$$W(\gamma_0[0, t], 0) = B(2t) + y_t, \qquad \gamma_0(0) = e^{i\Theta},$$

where $B(\cdot)$ is a standard Brownian motion started from 0, y_t is a random variable having uniform exponential tail and $\Theta \sim \mathrm{Unif}[0, 2\pi)$. In fact, $e^{i(B(2t)+\Theta)}$ is the driving function of γ_0. Also there exists constants C, c such that for all $s > 0$,

$$\mathbb{P}(|\gamma_0(t)| > e^{-t+s}) \leq Ce^{-cs}. \tag{2.8}$$

As a side note, we remark that it is precisely this observation which led Schramm to conjecture that loop-erased random walk converges to SLE_2, by combining this result together with Kenyon's work on the dimer model and his computation of the asymptotic pointwise variance of the height function.

3. Continuum windings and GFF. The goal of this section is to show that the winding of the branches in a continuum UST gives a Gaussian free field. By analogy with the discrete, we wish to show that the *intrinsic winding* (in the sense of earlier definitions) of the branches of the continuum UST up to the end points is the Gaussian free field. However, there are two obstacles if we want to deal with this. First, the branches are rough and hence intrinsic winding does not make sense. Second, the winding up to the end point blows up because the branches wind infinitely often in every neighbourhood of their endpoints (indeed this should be the case since the GFF is not defined pointwise).

To tackle the first problem, we note that the topological winding is well defined even for rough curves. We will therefore study the topological winding and add the correction term from Lemma 2.2 by hand (see Remarks 2.3 and 2.5 for additional details).

We address the second problem by regularising the winding to obtain a well defined function. The regularisation we use is simply to truncate the UST branches at some point. We will therefore have to show that this regularised winding field converges to a GFF as the regularisation is removed.

3.1. *Winding in the continuum and statement of the result.* Let D be a bounded simply connected domain with a locally connected boundary and a marked point x on its boundary. Let $\mathcal{T}$ be a continuum wired uniform spanning tree in D. Recall that viewed as a random variable in Schramm's space, a.s. for Lebesgue-almost every $z \in D$ there is a unique branch connecting z to ∂D and for a fixed z this has the law of a radial SLE_2. For $z \in \mathbb{D}$, let γ_z be the UST branch starting from z to the boundary (let z^* be the point where it hits ∂D), continued by going clockwise along ∂D from z^* to x. Note that since ∂U is locally connected, we can think of ∂U as a curve with some parametrisation [37] and hence this description indeed makes sense. Recall that for any point z, z^* has the distribution of a sample from the harmonic measure on the boundary seen from z, which we denote by $\mathrm{Harm}_D(z, \cdot)$. Also given z^*, the part of the curve from z^* to z is a radial SLE_2 curve in D from z^* to z in law. We parametrise the part of γ_z which lies in ∂D by $[-1, 0]$ so that $\gamma_z(-1) = x$ and $\gamma_z(0) = z^*$. We parametrise the rest of the curve by capacity, that is, for all $t \geq 0$, $-\log(R(z, D \setminus \gamma_z[-1, t]) = t - \log(R(z, D))$, where note that the term $-\log(R(z, D))$ is necessary for continuity.

If the boundary of D is smooth in a neighbourhood of x, then γ is smooth near -1 and we can define

$$h_t^D(z) = W(\gamma_z[-1, t], z) + \arg_{D;x}(z) - \arg_{D;x}(\gamma'(-1)), \tag{3.1}$$

where $\arg_{D;x}$ is defined as in Lemma 2.2. The intuition behind adding these extra terms is to work with (an emulation of) the intrinsic winding rather than the topological one; see Lemma 2.2. Note that h_t^D is defined almost surely as an almost everywhere function and hence in particular can be viewed (a.s.) as a random distribution.

For a domain D with general (not necessarily smooth) boundary, the additive constant $\arg_{D;x}(\gamma'(-1))$ might becomes ambiguous. We can nevertheless define h_t^D (and write simply h_t when there is no chance of confusion) as follows:

$$h_t^D(z) = W(\gamma_z[-1, t], z) + \arg_{D;x}(z) \qquad \text{up to a global constant in } \mathbb{R}. \tag{3.2}$$

For a.e. z, we get a branch γ_z which is an SLE_2 and to which Theorem 2.11 naturally applies. In particular, for a.e. z we get a driving Brownian motion $B_z(2t)$, which forms a Gaussian stochastic process when indexed by $\mathbb{D}$. Informally, the next result, which is the main result of this section, says that this Gaussian process converges to the Gaussian free field as $t \to \infty$. (In fact, the result below even deals with the error term y_t.) Recall that $u_{(D,x)}$ is the function which gives the intrinsic winding of the boundary curve ∂D, harmonically extended to D [see (2.7)].

THEOREM 3.1. *Let D be a bounded simply connected domain with locally connected boundary and a marked point $x \in \partial D$. As $t \to \infty$, we have the following convergence in probability*:

$$h_t \xrightarrow[t \to \infty]{} h_{\mathrm{GFF}}.$$

The convergence is in the Sobolev space $H^{-1-\eta}$ for all $\eta > 0$, and holds almost surely along the set of integers, that is, if we only take a limit with $t \in \mathbb{Z}$. Moreover, $\mathbb{E}(\|h_t -$

$h_{\mathrm{GFF}}\|^k_{H^{-1-\eta}}) \to 0$ *for any* $k \geq 1$. *The limit* h_{GFF} *is a Gaussian free field with variance* 2 *and winding boundary conditions*: *that is*, *we have*

$$h_{\mathrm{GFF}} = (1/\chi)h^0_{\mathrm{GFF}} + \pi/2 + u_{(D,x)},$$

where h^0_{GFF} *is a GFF with Dirichlet boundary conditions on* D *and* $u_{(D,x)}$ *is defined as in equation* (2.7) *and Remark* 2.7. *When the boundary is rough everywhere*, *the above convergence should be viewed up to a global constant in* $\mathbb{R}$.

REMARK 3.2. The coupling defined above between $\mathcal{T}$ and h_{GFF} is in fact the imaginary geometry coupling of Theorem 2.8. In particular, this result recovers the fact that h_{GFF} is measurable with respect to $\mathcal{T}$, furthermore providing a fairly explicit construction. It was already proved in [10] that actually both $\mathcal{T}$ and h_{GFF} are measurable with respect to each other and a little known fact is that Section 8.1 in that paper already sketches an explicit construction of the field as a function of the tree which is however different from our own. Note also that the construction in [10] was also conceived as an analogue to Temperley's bijection.

The rest of this section is dedicated to the proof of Theorem 3.1. The general strategy is to first study the k-point functions $\mathbb{E}[\prod h_t(x_i)]$ and to only integrate them at the last step to obtain moments of the integral of h_t against test functions. The advantage of working with the k-point function is that it only depends on k branches of the tree, which we know how to sample using Proposition 2.6. The existence of $\lim_{t\to\infty} \mathbb{E}[\prod h_t(x_i)]$ will follow from relatively simple distortion arguments and is proved in Lemma 3.7. This essentially shows that $\lim h_t$ exists in the sense of moments (in particular, this does not rely on Imaginary Geometry yet).

To identify the limit, we show that the conditional expectation of $\lim h_t$ given some tree branches agrees with the imaginary geometry definition (Sections 3.3 and 3.4). The uniqueness in imaginary geometry concludes. Finally, Section 3.5 covers the extension from the disc to general smooth domains and Section 3.6 upgrades the convergence from finite dimensional marginals to $H^{-1-\eta}$ using the moment bounds derived earlier.

3.2. *Convergence in the unit disc*: *One point function*. We first prove Theorem 3.1 in the case $D = \mathbb{D}$ of the unit disc, with the marked point 1. The extension of the results to general domains is discussed in Section 3.5. *Until that section, we henceforth assume* $D = \mathbb{D}$.

Recall from (3.1) that the definition of h_t for this case is given by

$$h_t(z) = W(\gamma_z[-1, t], z) + \arg_{\mathbb{D};1}(z) - \pi/2, \tag{3.3}$$

where, as in Lemma 2.2, $\arg_{\mathbb{D};1}$ is chosen so that $\arg_{\mathbb{D};1}(0) = \pi$.

LEMMA 3.3. *Let* $a_1, \dots, a_k \in \mathbb{D}$ *be distinct and let* $K = \gamma_{a_1}[0, t_1] \cup \dots \cup \gamma_{a_k}[0, t_k]$ *where* $0 \leq t_i \leq \infty$. *Fix* $z \in \mathbb{D}$ *distinct from any of the* a_i, *and* $T > 0$. *Let* $D' = \mathbb{D} \setminus K$ *and assume that* 1 *is a smooth point of* $\partial D'$. *Let* $g : D' \to \mathbb{D}$ *be a conformal map such that* $g(1) = 1$ (*note such a map is not unique*). *If* $\gamma_z(T) \in B(z, \varepsilon)$ *with* $\varepsilon \leq R(z, D')/8$. *Then*

$$\begin{aligned} &W\big(g(\gamma_z[-1, T]), g(z)\big) + \arg_{\mathbb{D};1}\big(g(z)\big) - \arg_{g'(D')}\big(g'(z)\big) + \epsilon(T) \\ &\quad = W\big(\gamma_z[-1, T], z\big) + \arg_{\mathbb{D};1}(z), \end{aligned} \tag{3.4}$$

where $|\epsilon(T)| = O(\varepsilon/R(z, D'))$ *with the implied constant being universal*, *and as before* $\arg_{g'(D')}$ *is chosen so that* $\arg_{g'(D')}(g'(1)) = 0$.

Furthermore, *assume that* $\mathrm{dist}(z, K \cup \partial\mathbb{D}) > \mathrm{dist}(z, \gamma_z[0, T])$. *Then*

$$\frac{e^{-T}}{4} \leq \frac{R(g(z), \mathbb{D} \setminus g(\gamma_z[0, T]))}{R(g(z), \mathbb{D})} \leq e^{-T} \,\mathrm{dist}(z, K \cup \partial\mathbb{D})^{-1}. \tag{3.5}$$

PROOF. (3.4) is just an application of Lemma 2.4. We only have to check the choice of the constant in the arguments. First, note that $D' \subset \mathbb{D}$ so we can choose $\arg_{D';1}$ to be a restriction of $\arg_{\mathbb{D};1}$. Thus, it remains to check that the chain rule (2.6) implies that $\arg_{g'(D')}(g'(1)) = 0$, which however is easy to check thanks to the fact that $g'(1) > 0$.

Moreover, (3.5) follows easily from conformal invariance and domain monotonicity of the conformal radius as well as Koebe's 1/4 theorem (see Theorem 3.17 in [29]). □

Theorem 2.11 deals with SLE curves towards 0. We now provide an extension of this result for SLE curves towards an arbitrary point in the unit disc.

LEMMA 3.4. *Let $z \in \mathbb{D}$ and let $\psi : \mathbb{D} \mapsto \mathbb{D}$ be the Möbius transformation mapping z to* 0 *and* 1 *to* 1. *If $\gamma_z(t) \in B(z, \varepsilon)$ where $\varepsilon \leq R(z, \mathbb{D})/8$, then we have*:

$$W\big(\psi(\gamma_z[-1,t]), 0\big) = W\big(\gamma_z[-1,t], z\big) + \pi - \arg_{\mathbb{D};1}(z) + \epsilon(t), \tag{3.6}$$

where the error term $|\epsilon(t)| \leq C\varepsilon/R(z, \mathbb{D})$ for some universal constant $C > 0$ and $\arg_{\mathbb{D};1}$ is chosen so that $\arg_{\mathbb{D};1}(0) = \pi$. Also for all s, t

$$\mathbb{P}\big(\big|\gamma_z(t) - z\big| > e^{-t+s} R(z, \mathbb{D})\big) \leq ce^{-c's}, \tag{3.7}$$

where c, c' are independent of z.

The proof can be found in the Supplementary Material [4], Lemma B.9.

We now want to regularise h_t a bit further by restricting it to an event where the tip is not too far away from the endpoint. This is something we often need to do in the following, so we will define for $t \geq 0$ and $z \in \mathbb{D}$,

$$\mathcal{A}(t, z) := \big\{\big|\gamma_z(t) - z\big| < e^{-t/2} R(z, \mathbb{D})\big\}; \qquad \hat{h}_t(z) := h_t(z)\mathbb{1}_{\mathcal{A}(t,z)}. \tag{3.8}$$

The event $\mathcal{A}(t, z)$ and corresponding field $\hat{h}_t$ will be used throughout our proof of Theorem 3.1. By Lemma 3.4, $\mathcal{A}(t, z)$ is a very likely event:

$$\mathbb{P}\big(\mathcal{A}(t, z)\big) > 1 - ce^{-c't} \tag{3.9}$$

for some universal constants $c, c' > 0$.

LEMMA 3.5. *We have for every $z \in \mathbb{D}$,*

$$\lim_{t\to\infty} \mathbb{E}\big(\hat{h}_t(z)\big) = 2\arg_{\mathbb{D};1}(z) - \frac{\pi}{2}. \tag{3.10}$$

Also, *we have the following bounds on the moments*:

$$\mathbb{E}\big(\big|\hat{h}_t(z)\big|^k\big) \leq c(k)\big(1 + t^{k/2}\big), \qquad \mathbb{E}\big(\big|h_t(z)\big|^k\big) \leq c(k)\big(1 + t^{k/2}\big). \tag{3.11}$$

PROOF. We first check equation (3.10) and equation (3.11) for $z = 0$. Then $\gamma_z(0)$ is uniformly distributed on $\partial\mathbb{D}$, contributing an expected topological winding of π (using the fact that the Loewner equation is invariant under $z \mapsto \bar{z}$). Adding the term $\arg_{\mathbb{D};1}(0) - \pi/2$ in the definition of h_t in equation (3.3) shows that $\mathbb{E}(h_t) = 3\pi/2$. Furthermore, by Theorem 2.11, we have $\mathbb{E}(h_t(0)^2) \leq 2t + o(t)$ and since $\mathbb{P}(\mathcal{A}(t, z)) \geq 1 - e^{-ct}$ we deduce from Cauchy–Schwarz that $\lim_{t\to\infty} \mathbb{E}(\hat{h}_t(0)) = 3\pi/2$. The moment bound for $h_t(0)$ (and then for $\hat{h}_t(0)$ using Cauchy–Schwarz) follows again from Theorem 2.11 and the inequality $|a + b|^k \leq 2^{k-1}(|a|^k + |b|^k)$.

For any other $z \in \mathbb{D}$, we start by proving the moment bound equation (3.11). We join $\gamma_z(t)$ and z by a hyperbolic geodesic in $\mathbb{D}$, call the resulting union γ', and apply ψ to it.

Then the image becomes a concatenation of an SLE_2 curve targeted towards 0 and another hyperbolic geodesic. Using equation (3.6) (which is deterministic) with $\varepsilon = 0$, $W(\psi(\gamma'), 0) - W(\gamma', z) = \pi - \arg_{\mathbb{D};1}(z)$. Since the winding of the hyperbolic geodesics are bounded by at most π, and the winding of $\psi(\gamma)$ possesses the required moment bounds, this proves equation (3.11) in $\mathbb{D}$.

Now using (3.6):

$$\begin{aligned}
&\mathbb{E}(W(\psi(\gamma_z[-1,t]),0)\mathbb{1}_{\mathcal{A}(t,z)}) \\
&\quad = \mathbb{E}(W(\gamma_z[-1,t],z)\mathbb{1}_{\mathcal{A}(t,z)}) \\
&\qquad + (\pi - \arg_{\mathbb{D};1}(z))\mathbb{P}(\mathcal{A}(t,z)) + \mathbb{E}(\epsilon(t)\mathbb{1}_{\mathcal{A}(t,z)}).
\end{aligned} \tag{3.12}$$

Since $\psi(\gamma_z[-1,t))$ is an SLE_2 curve towards 0 the left hand side of (3.12) converges to π. Also from Lemma 3.4, the error term $|\epsilon(t)| < ce^{-c't} \to 0$ on $\mathcal{A}(t,z)$ and hence converges to 0 as $t \to \infty$. Recall also the terms added in the definition of h_t in (3.1). Combining all these together with equation (3.11), we have our result. $\square$

3.3. *Conformal covariance of k-point function.* In the next lemma, we prove the existence of the limit of the k-point function of the regularised winding field of the continuum UST. However, we do not identify the limit at this point as this requires a separate argument. For this separate argument, we will also need a convergence result of the k-point function given several branches of $\mathcal{T}$, the continuum UST.

PROPOSITION 3.6. *Let $\{z_1, \dots, z_k, w_1, \dots, w_{k'}\}$ be a set of points in $\mathbb{D}$ all of which are distinct. Then the following are true.*

- *Both $\lim_{t\to\infty} \mathbb{E}(\prod_{i=1}^k \hat{h}_t(z_i))$ and $\lim_{t\to\infty} \mathbb{E}(\prod_{i=1}^k h_t(z_i))$ exist and are equal.*
- *Let $A = \{\gamma_{w_1}, \dots, \gamma_{w_{k'}}\}$ be a set of branches of $\mathcal{T}$. Let $\mathbb{E}^A$ denote the conditional expectation given A. Let $g_A : \mathbb{D} \setminus A \mapsto \mathbb{D}$ be some conformal map which fixes 1. Let $\tilde{h}_t$ be an independent copy of h_t in $\mathbb{D}$. Then*

$$\begin{aligned}
\lim_{t\to\infty} \mathbb{E}^A\left(\prod_{i=1}^k h_t(z_i)\right) &= \lim_{t\to\infty} \mathbb{E}^A\left(\prod_{i=1}^k \hat{h}_t(z_i)\right) \\
&= \lim_{t\to\infty} \mathbb{E}\left[\prod_{i=1}^k (\tilde{h}_t(g_A(z_i)) - \arg_{g_A'(\mathbb{D}\setminus A)}(g_A'(z_i)))\right] \qquad a.s.
\end{aligned}$$

We call the function defined by the first point of the proposition the k-point function and we write it H:

$$H(z_1, \dots, z_k) := \lim_{t\to\infty} \mathbb{E}\left(\prod_{i=1}^k h_t(z_i)\right). \tag{3.13}$$

The technical part of the proof of Proposition 3.6 is accomplished in the following lemma.

LEMMA 3.7. *Let $\{z_1, z_2, \dots, z_k\}$ be a set of points in $\mathbb{D}$ all of which are distinct. Let $r = \min_i \operatorname{dist}(z_i, \partial\mathbb{D}) \wedge \min_{i\neq j} |z_i - z_j|$. Let $t_1 \geq t_2 > \cdots \geq t_k > t \geq -10\log r + 1$ such that $t_1 < 10t_k$. Then there are constants c, c' depending only on k such that*

$$\left|\mathbb{E}\left(\prod_{i=1}^k \hat{h}_{t_i}(z_i)\right) - \mathbb{E}\left(\prod_{i=1}^k \hat{h}_t(z_i)\right)\right| < ct^k e^{-c't}.$$

The same inequality holds with h instead of $\hat{h}$.

Let us comment on why we need to go through the trouble of considering multiple times in Lemma 3.7. A first issue is that when we apply a conformal map, the conformal radius changes differently depending on the point. To get any control both before and after applying the map we therefore need to allow for different t_i's (this is, e.g., the case in equation (3.29)).

PROOF. We first claim that it is enough to prove that for t_i's as above,

$$\left|\mathbb{E}\Big(\prod_i \hat{h}_{t_i}(z_i)\Big) - \mathbb{E}\Big(\prod_i \hat{h}_t(z_i)\Big)\right| \le c t_1^{k/2} e^{-c't}. \tag{3.14}$$

This clearly completes the proof since we can break up the interval $[t, t_k]$ into $\bigcup_{i=1}^{J}[t2^{i-1}, t2^i]$ where $t2^{J-1} \le t_k < t2^J$. Using the bound (3.14) for each such interval and using $t_1 < 10t_k$,

$$\left|\mathbb{E}\Big(\prod_i \hat{h}_{t_i}(z_i)\Big) - \mathbb{E}\Big(\prod_i \hat{h}_t(z_i)\Big)\right| \le c\sum_{i=1}^{\infty}(t2^i)^{k/2}e^{-c't2^{i-1}} \le ct^{k/2}e^{-c't} \tag{3.15}$$

from which Lemma 3.7 follows. The bound for the term involving h_t follows from that of $\hat{h}_t$ using equation (3.11), Hölder's inequality (generalised for k terms) and the exponential bound on the probability of events $\mathcal{A}(t_i, z_i)$.

To prove (3.14), the idea is to consider several cases depending on how close γ_{z_i} gets to the other points. If it gets very close, the distortion of the conformal map becomes more pronounced and the estimate in Lemma 3.3 carries large errors. But γ_{z_i} getting close to z_j for some $j \neq i$ is unlikely by Lemma 2.9 and comes at a price. So there is a tradeoff between these two situations. Let $d_i = \inf_{j\neq i} \operatorname{dist}(z_i, \gamma_{z_j}[0,\infty)) \wedge r$ and $d_{\min} := \min_i(d_i)$.

Case 1: $-\log(d_{\min}) > t/4$. By Lemma 2.9 and a union bound, $\mathbb{P}(-\log d_{\min} > t/4) \le ck(\frac{e^{-t/4}}{r})^{c_0}$ for universal constants c, c_0. Using the fact that $t > -10\log r + 1$, we see that $\mathbb{P}(-\log d_{\min} > t/4) \le cke^{-c't}$ for some $c' > 0$. Using the one-point moment bounds (3.11) and Hölder's (generalised) inequality,

$$\left|\mathbb{E}\Big(\prod_i \hat{h}_{t_i}(z_i)\Big) - \mathbb{E}\Big(\prod_i \hat{h}_t(z_i)\Big)\mathbb{1}_{-\log(d_{\min})>t/4)}\right| < ct_1^{k/2}e^{-c't} \tag{3.16}$$

for some positive universal constants c, c' since $t > 1$.

Case 2: $-\log(d_{\min}) \le t/4$. Let $A_i := \{\gamma_{z_j}[0,\infty) : j \neq i\}$. First we observe that it is enough to prove

$$\left|\mathbb{E}\big(h_{t_i}(z_i) - h_t(z_i)|A_i\big)\mathbb{1}_{-\log(d_{\min})<t/4}\right| \le ct_ie^{-c't} \tag{3.17}$$

since we can use the decomposition

$$\begin{aligned}
&\left|\mathbb{E}\Big(\prod_i \hat{h}_{t_i}(z_i)\Big) - \mathbb{E}\Big(\prod_i \hat{h}_t(z_i)\Big)\right| \\
&\quad \le \sum_{i=1}^{k}\big|\mathbb{E}\big(\mathbb{E}\big(\hat{h}_{t_i}(z_i) - \hat{h}_t(z_i)|A_i\big)\hat{h}_t(z_1)\cdots \\
&\qquad \hat{h}_t(z_{i-1})\hat{h}_{t_{i+1}}(z_{i+1})\cdots\hat{h}_{t_n}(z_k)\big)\big|
\end{aligned} \tag{3.18}$$

and then use Hölder's inequality, (3.17) and the one-point moment bounds (3.11) to obtain the required bound.

We now concentrate on the proof of (3.17). We wish to use Lemma 3.3 and map out A_i by a conformal map φ mapping z_i to 0 and 1 to 1 and record the change in winding of γ_{z_i}. By (3.5),

$$\frac{e^{-t_i}}{4} \leq R(\varphi(z_i), \mathbb{D} \setminus \varphi(\gamma_{z_i}[0, t_i])) \leq e^{-(t_i - t/4)}$$

since $-\log(d_{\min}) \leq t/4$. Therefore, using Lemma 3.3 for an independent copy $\tilde{h}$ of h (note that the $\arg_{\varphi'(\mathbb{D}\setminus A_i)}$ term cancels), we have

$$\begin{aligned} &\mathbb{E}\big((\hat{h}_{t_i}(z_i) - \hat{h}_t(z_i))\mathbb{1}_{\mathcal{A}(t,z_i)} | A_{z_i}\big) \\ &\quad = \mathbb{E}\big((\tilde{h}_{t_i'}(z_i) - \tilde{h}_{t'}(z_i) + \epsilon(t_i) - \epsilon(t))\mathbb{1}_{\mathcal{A}(t,z_i)}\big), \end{aligned} \tag{3.19}$$

where $|t_i' - t_i| < t_i/2$ and $|t' - t| < t/2$ and $|\epsilon(t_i)| \vee |\epsilon(t)| \leq e^{-ct}$ on $\mathcal{A}(t, z_i)$. Now notice that by symmetry, $\mathbb{E}((\tilde{h}_{t_i'}(z_i) - \tilde{h}_{t'}(z_i))) = 0$. We conclude using Cauchy–Schwarz, the moment bound (3.11) and the bound on the probability on $\mathcal{A}(t, z_i)^c$. $\square$

We also need the following estimate which says that the k-point function blows up at most like a power of $\log(r)$ as the points come close.

LEMMA 3.8 (Logarithmic divergence). *For any $k \geq 1$ and any k distinct points $z_1, z_2, \ldots, z_k \in \mathbb{D}$ and $t > 0$,*

$$\left|\mathbb{E}\left(\prod_{i=1}^k h_t(z_i)\right)\right| \leq c\big(1 + \log^k(1/r)\big); \qquad \left|\mathbb{E}\left(\prod_{i=1}^k \hat{h}_t(z_i)\right)\right| \leq c\big(1 + \log^k(1/r)\big),$$

where $r = \min_i \operatorname{dist}(z_i, \partial\mathbb{D}) \wedge \min_{i \neq j} |z_i - z_j|$ and $c = c(k) > 0$ is a constant.

PROOF. We only check the first of these inequalities as the proof of the other is identical. Let $t = -10 \log r + 1$. By Lemma 3.5, for $t' \leq t$, we obtain $\mathbb{E}(\prod_{i=1}^k |h_{t'}(z_i)|) \leq C(1 + (t')^{k/2}) \leq C(1 + t^{k/2})$ which is what we wanted for $t' \leq t$. On the other hand if $t' \geq t$, by Lemma 3.7, $|\mathbb{E}(\prod_{i=1}^k h_{t'}(z_i) - \prod_{i=1}^k h_t(z_i))| < ct^k e^{-ct}$. Combining with the result for $t' = t$ we obtained the desired bound also for $t' \geq t$. $\square$

PROOF OF PROPOSITION 3.6. Notice that Lemma 3.7 implies that the quantity $\mathbb{E}(\prod_{i=1}^k h_t(z_i))$ (resp. $\mathbb{E}(\prod_{i=1}^k \hat{h}_t(z_i))$) is a Cauchy sequence and hence converges. Moreover,

$$\begin{aligned} \left|\mathbb{E}\left(\prod_{i=1}^k h_t(z_i) - \prod_{i=1}^k \hat{h}_t(z_i)\right)\right| &\leq \mathbb{E}\left(\prod_{i=1}^k |h_t(z_i)| \mathbb{1}_{\bigcup_i \mathcal{A}(t,z_i)^c}\right) \\ &\leq c\big(1 + t^k\big)e^{-c't} \to 0. \end{aligned} \tag{3.20}$$

Hence the limits are the same, proving the first point.

To simplify notation, we write g in place of g_A and arg in place of $\arg_{g_A'(\mathbb{D}\setminus A)}$. Let $r = \min_i \operatorname{dist}(z_i, A \cup \partial\mathbb{D})$ and take $t > -11 \log r + 1$. From Lemma 3.3, we see (using the obvious domain Markov property and conformal invariance of the UST) that given A, we have the equality in distribution

$$h_t(z) = \tilde{h}_{t_i}\big(g(z)\big) - \arg g'(z_i) + \epsilon_i(t), \tag{3.21}$$

where

$$t_i = -\log R\big(g(z_i), \mathbb{D} \setminus g\big(\gamma_{z_i}[0, t]\big)\big) + \log R\big(g(z_i), \mathbb{D}\big).$$

Hence,

$$\mathbb{E}^A\left(\prod_{i=1}^k \hat{h}_t(z_i)\right) = \mathbb{E}\left(\prod_{i=1}^k (\tilde{h}_{t_i}(g(z_i)) - \arg(g'(z_i)) + \epsilon_i(t))\mathbb{1}_{\mathcal{A}(t,z_i)}\right). \tag{3.22}$$

By equation (3.5), $|t_i - t| \le \log 4 + \log(1/r)$. Therefore almost surely, $9t/10 \le t_i \le 11t/10$ for all i from the choice of t. Thus $t_i \to \infty$ as $t \to \infty$. Further $|\epsilon_i(t)| = O(e^{-t/2}/r) = O(e^{-t/2+t/10}) \to 0$ for all i on the event $\mathcal{A}(t, z_i)$ from Lemma 3.3. Using all this information, Cauchy–Schwarz, Lemmas 3.7 and 3.8, we obtain

$$\begin{aligned}\Bigg|\mathbb{E}\left(\prod_{i=1}^k (\tilde{h}_{t_i}(g(z_i)) - \arg(g'(z_i)) + \epsilon_i(t))\mathbb{1}_{\mathcal{A}(t,z_i)}\right)\\ - \mathbb{E}(\prod_{i=1}^k (\tilde{h}_{9t/10}(g(z_i)) - \arg(g'(z_i)))\Bigg| \le ct^k e^{-c't}\end{aligned}$$

almost surely given A. The second item of the proposition now follows from the first item. $\square$

To prepare for the proof of convergence in the Sobolev space $H^{-1-\eta}$ for all $\eta > 0$ we need the following convergence of h_t integrated against test functions.

LEMMA 3.9. *Let $\{f_i\}_{1\le i\le n}$ be smooth compactly supported functions in $\mathbb{D}$. Then for any sequence of integers $k_1, \dots, k_n$,*

$$\lim_{t\to\infty} \mathbb{E}\prod_{i=1}^n \left(\int_{\mathbb{D}} h_t(z) f_i(z)\, dz\right)^{k_i} = \int_{\mathbb{D}^{\sum_{i=1}^n k_i}} H(z_{11}, \dots, z_{nk_n}) \prod_{i=1}^n \prod_{j=1}^{k_i} f_i(z_{ij})\, dz_{ij},$$

where H is as in Proposition 3.6.

PROOF. Straightforward expansion and Fubini's theorem yield

$$\begin{aligned}&\mathbb{E}\prod_{i=1}^n \left(\int_{\mathbb{D}} h_t(z) f_i(z)\, dz\right)^{k_i}\\ &\quad = \int_{\mathbb{D}^{\sum_{i=1}^n k_i}} \mathbb{E}\left(\prod_{i=1}^n \prod_{j=1}^{k_i} h_t(z_{ij})\right) \prod_{i=1}^n \prod_{j=1}^{k_i} f_i(z_{ij})\, dz_{ij}.\end{aligned} \tag{3.23}$$

We can apply Fubini because the term inside the integral is integrable from the moment bounds in Lemma 3.5. We want to take the limit as $t \to \infty$ on both sides of (3.23) and apply dominated convergence theorem and Proposition 3.6 to complete the proof. To justify the application of dominated convergence theorem note that by Lemma 3.8, $\mathbb{E}|\prod_{i=1}^n \prod_{j=1}^{k_i} h_t(z_{ij}))| \le \log^{\sum_{i=1}^n k_i/2}(r)$ where $r = \min_{(i,j),(i',j')} |z_{ij} - z_{i'j'}| \wedge \min_{ij} |z_{ij} - \partial\mathbb{D}|$, which is integrable. Further the functions f_i's are uniformly bounded. $\square$

3.4. *Identifying winding as the GFF*: *Imaginary geometry.* At this stage, we have proven that h_t converges as t goes to infinity, even if we are yet to state a precise meaning for this convergence. However, we do not have any information about the limit law and in particular the k-point function H is unknown. In this section, we will identify the limit with the GFF, using the imaginary geometry coupling.

Recall from Section 2.2 that imaginary geometry provides a coupling between a UST and h_{GFF}, such that conditionally on some branches γ_i of the UST, h_{GFF} is a GFF in $\mathbb{D} \setminus \bigcup \gamma_i$, plus an argument term. Note that this argument term is exactly the same as the one for the conditional law of the regularised winding h_t (see Proposition 3.6). The key idea will be to say that if we take a large but finite number of branches, then in $\mathbb{D} \setminus \bigcup \gamma_i$ all points are close to the boundary and therefore both h_{GFF} and h_t have a small conditional variance. The means are essentially the argument terms so they match up to small errors. This will show that h_{GFF} and h_t are close in L^2, hence identifying the limit.

Note that the only nontrivial fact about imaginary geometry that we need to use is the existence of a field h_{GFF} with such a conditional law.

We first need the fact that the centred two-point function $G(x, y)$ (defined below) is small when one of the points, say x, is near the boundary. For this, we start by a deterministic lemma about the argument of conformal maps that remove a small set.

LEMMA 3.10 (Distortion of argument). *Let K be a closed subset of $\bar{\mathbb{D}}$ such that $H = \mathbb{D} \setminus K$ is simply connected (i.e., K is a hull). Further assume that the diameter of K is smaller than some $\delta < 1/2$ and $1 \notin \bigcup_{x \in K} B(x, \delta^{1/2})$. Let $\tilde{g}$ denote the conformal map sending H to $\mathbb{D}$ with $\tilde{g}(0) = 0$ and $\tilde{g}(1) = 1$. Then*

$$\left|\arg_{\tilde{g}'(H)}\bigl(\tilde{g}'(0)\bigr) - \arg_{\tilde{g}'(H)}\bigl(\tilde{g}'(1)\bigr)\right| < C\delta^{1/2},$$

where C is a universal constant. Here, $\arg_{\tilde{g}'(H)}(\cdot)$ is the argument in $\tilde{g}'(H)$ (which does not contain 0), defined up to a global unimportant additive constant.

The proof is given in the Supplementary Material ([4], Lemma B.3).

We define

$$G(z_1, z_2, \dots, z_k) = \lim_{t \to \infty} \mathbb{E}\left(\prod_{i=1}^{k} \bigl(h_t(z_i) - \mathbb{E}\bigl(h_t(z_i)\bigr)\bigr)\right) \tag{3.24}$$

to be the k-point covariance function which exists by Proposition 3.6. Using Lemma 3.10, we can show that the two-point function is small close to the boundary (i.e., the field has Dirichlet boundary conditions):

LEMMA 3.11. *For all $|z| \geq 3/4$ and $t_2 \geq t_1 > -10 \log \operatorname{dist}(z, \partial\mathbb{D}) + 1$ such that $t_2 < 10 t_1$,*

$$\left|\mathbb{E}\bigl(h_{t_1}(z) h_{t_2}(0) - \mathbb{E}\bigl(h_{t_1}(z)\bigr)\mathbb{E}\bigl(h_{t_2}(0)\bigr)\bigr)\right| \leq c \operatorname{dist}(z, \partial\mathbb{D})^{c'}.$$

In particular, as $z \to \partial\mathbb{D}$, $G(0, z) \to 0$.

PROOF. Let $r = \operatorname{dist}(z, \partial\mathbb{D}) = 1 - |z|$. Set $t = -10 \log r + 1$. By Lemma 3.7,

$$\left|\mathbb{E}\bigl(h_{t_1}(z) h_{t_2}(0) - h_t(z) h_t(0)\bigr)\right| \leq c t e^{-c' t},$$

$$\left|\mathbb{E} h_{t_1}(z) \mathbb{E} h_{t_2}(0) - \mathbb{E} h_t(z) \mathbb{E} h_t(0)\right| \leq c t e^{-c' t},$$

and observe that $cte^{-c't} \leq c r^{c'}$. Let us define the event $\mathcal{G} := \mathcal{A}(t, z) \cap \{\gamma_z \subset B(z, \sqrt{r})\}$ where here and in the rest of the proof by γ_z we mean $\gamma_z[0, \infty)$. From the exponential bound on the probability of $\mathcal{A}(t, z)$, $\mathcal{A}(t, 0)$ and Lemma 2.9, we have

$$\mathbb{P}\bigl(\mathcal{G} \cap \mathcal{A}(t, 0)\bigr) \geq 1 - c e^{-c' t}. \tag{3.25}$$

By Lemma 3.5 and Cauchy–Schwarz, we see that

$$\left|\mathbb{E}\bigl(h_t(z) h_t(0)\bigr) - \mathbb{E}\bigl(h_t(z) h_t(0) \mathbb{1}_{\mathcal{G}, \mathcal{A}(t,0)}\bigr)\right| \leq c t e^{-c' t}. \tag{3.26}$$

Let $g : \mathbb{D} \setminus \gamma_z \mapsto \mathbb{D}$ be a conformal map fixing 0 and 1. Then from Lemma 3.3, we have for some independent copy $\tilde{h}_t(0)$ of $h_t(0)$,

$$\begin{aligned}
&\mathbb{E}(h_t(0)\mathbb{1}_{\mathcal{A}(t,0)}|\gamma_z)\mathbb{1}_{\mathcal{G}} \\
&\quad = \mathbb{E}\big((\tilde{h}_s(0) - \arg_{g'(\mathbb{D}\setminus\gamma_z)}(g'(0)) + \epsilon(t))\mathbb{1}_{\mathcal{A}(t,0)}\big)\mathbb{1}_{\mathcal{G}},
\end{aligned}\tag{3.27}$$

where $|\epsilon(t)| < ce^{-c't}$ on $\mathcal{A}(t,0)$ and $t + \log(|z| - \sqrt{r}) \le s \le t + \log 4$ (as in equation (3.5)), and where $\arg_{g'(\mathbb{D}\setminus\gamma_z)}$ is chosen as in Lemma 3.3, that is, $\arg_{g'(\mathbb{D}\setminus\gamma_z)}(g'(1)) = 0$. Note also that $|z| - \sqrt{r} \ge 1 - 2\sqrt{r} > 0$. Thus, we obtain using equation (3.26),

$$\begin{aligned}
&\mathbb{E}\big(h_t(z)h_t(0)\big) = O\big(e^{-ct}\big) \\
&\quad + \mathbb{E}\big(h_t(z)\mathbb{E}\big((\tilde{h}_s(0) - \arg_{g'(\mathbb{D}\setminus\gamma_z)}(g'(0)) + \epsilon(t))\mathbb{1}_{\mathcal{A}(t,0)}|\gamma_z\big)\mathbb{1}_{\mathcal{G}}\big).
\end{aligned}\tag{3.28}$$

We now expand the terms in the right-hand side and treat each of them separately. Observe that while $\tilde{h}$ is independent of h, s is still measurable with respect to γ_z. Hence by symmetry, $\mathbb{E}(\tilde{h}_s(0)|\gamma_z) = H(0)$ a.s. and hence

$$\mathbb{E}\big(h_t(z)\mathbb{E}\big(\tilde{h}_s(0)|\gamma_z\big)\big) = H(0)\mathbb{E}\big(h_t(z)\big) \to H(0)H(z)$$

as $t \to \infty$. (In fact a symmetry argument holds here as well and there is no need to let $t \to \infty$.)

Regarding the second term, we claim that $\mathbb{E}(h_t(z)\arg_{g'(\mathbb{D}\setminus\gamma_z)}(g'(0))\mathbb{1}_{\mathcal{G}})$ converges to 0. Indeed, this follows from the distortion estimate on the argument we did in Lemma 3.10 and the fact that on $\mathcal{G}$, $\mathrm{Diam}(\gamma_z(t)) < \sqrt{r}$. Hence by Cauchy–Schwarz, we conclude

$$\mathbb{E}\big|h_t(z)\arg_{g'(\mathbb{D}\setminus\gamma_z)}(g'(0))\mathbb{1}_{\mathcal{G}}\big| < cte^{-c't}.$$

Finally, for the third term, since $|\epsilon(t)| \le e^{-c't}$ on $\mathcal{A}(t,0)$, we deduce that $\mathbb{E}(h_t(0)\epsilon(t) \times \mathbb{1}_{\mathcal{A}(t,0),\mathcal{G}}) \le te^{-ct}$ by Cauchy–Schwarz and the moment bound. Consequently, we have proved

$$\big|\mathbb{E}\big(h_t(z)h_t(0)\big) - H(z)H(0)\big| \le ce^{-c't}.$$

Using Lemma 3.7, we deduce that $|\mathbb{E}(h_t(z)h_t(0)) - \mathbb{E}(h_t(z))\mathbb{E}(h_t(0))| \le ce^{-c't}$. This proves the lemma. □

LEMMA 3.12. *Let $\{w_1, w_2, \dots, w_k\}$ be a set of points in $\mathbb{D}$ all of which are distinct. Let $A = \{\gamma_{w_1}, \dots, \gamma_{w_k}\}$ be the corresponding set of branches of $\mathcal{T}$ in $\mathbb{D}$. Let $g : \mathbb{D} \setminus A \to \mathbb{D}$ be a conformal map fixing 1. Let $g_z : \mathbb{D} \setminus A \to \mathbb{D}$ be a conformal map which maps z to 0 and 1 to 1. Then for any test function f in $C^\infty(\bar{\mathbb{D}})$,*

$$\begin{aligned}
&\lim_{t\to\infty} \mathbb{E}\bigg[\int_{\mathbb{D}} \hat{h}_t(z)f(z)\mathbb{1}_{\mathrm{dist}(z,A)>e^{-t/10}}\,dz\big|A\bigg] \\
&\quad = \int_{\mathbb{D}} \bigg(2\arg_{\mathbb{D};1}\big(g(z)\big) - \frac{\pi}{2} - \arg_{g'(\mathbb{D}\setminus A)}\big(g'(z)\big)\bigg) f(z)\,dz,
\end{aligned}$$

where $\arg_{g'(\mathbb{D}\setminus A)}$ is chosen so that $\arg_{g'(\mathbb{D}\setminus A)}(g'(1)) = 0$.

$$\begin{aligned}
&\lim_{t\to\infty} \mathrm{Var}\bigg[\int_{\mathbb{D}} \hat{h}_t(z)f(z)\mathbb{1}_{\mathrm{dist}(z,A)>e^{-t/10}}\,dz\big|A\bigg] \\
&\quad = \int_{\mathbb{D}\times\mathbb{D}} G\big(0, g_z(w)\big)f(z)f(w)\,dz\,dw
\end{aligned}$$

almost surely, where $G(\cdot,\cdot)$ is the two-point covariance function defined in equation (3.24).

PROOF. This proof is an application of dominated convergence theorem. For the first item, note that for a fixed t we can take the expectation inside by Fubini and the moment bounds of $\hat{h}_t$. Again observe that from (3.22), we have for an independent copy $\tilde{h}_t$ of h_t in $\mathbb{D}$, and if we write $\mathcal{D}(z,A) = \{\mathrm{dist}(z,A) > e^{-t/10}\}$,

$$\begin{aligned}
&\int_{\mathbb{D}} \mathbb{E}\big[\hat{h}_t(z) f(z) \mathbb{1}_{\mathcal{D}(z,A)}\, dz | A\big] \\
&\quad = \int_{\mathbb{D}} \mathbb{1}_{\mathrm{dist}(z,A) > e^{-t/10}} \mathbb{E}\big(\mathbb{1}_{\mathcal{A}(t,z)}\big(\tilde{h}_{t'}(g(z)) - \arg_{g'(\mathbb{D}\setminus A)} g'(z)\big) | A\big) f(z)\, dz \\
&\qquad + O(e^{-ct}),
\end{aligned}$$

where $9t/10 < t' = t'(z) < 11t/10$ almost surely by Lemma 3.3. Therefore, the first item follows by taking limit on both sides and using dominated convergence theorem (whose application is justified by, say, Lemma 3.8).

For the variance computation, recall that we write $\mathbb{E}^A$ for the conditional expectation given A. Then, applying the conformal map g_z and using (3.21),

$$\begin{aligned}
&\mathbb{E}^A\big[\big(\hat{h}_t(z) - \mathbb{E}^A \hat{h}_t(z)\big)\mathbb{1}_{\mathcal{D}(z,A)}\big(\hat{h}_t(w) - \mathbb{E}^A \hat{h}_t(w)\big)\mathbb{1}_{\mathcal{D}(w,A)}\big] \\
&\quad = \mathbb{E}^A\Big[\prod_{y\in\{z,w\}} \big(\tilde{h}_{t_y}\big(g_z(y)\big) + \epsilon_y(t) - \mathbb{E}\tilde{h}_{t_x}\big(g_z(y)\big)\big)\mathbb{1}_{\mathcal{A}(t,y);\mathcal{D}(y,A)}\Big]
\end{aligned} \tag{3.29}$$

because once we condition on A, the term $\arg_{g'_z(\mathbb{D}\setminus A)}(g'_z(y))$ is nonrandom and hence cancels out in $\hat{h}_t(y) - \mathbb{E}^A \hat{h}_t(y)$. Note that since z, w are at least at a distance $e^{-t/10}$ away from A, we have $9t/10 \le t_y \le 11t/10$ for $y \in \{z,w\}$, and that $|\epsilon_y(t)| \le ce^{-c't}$ on $\mathcal{A}(t,y)$. By Cauchy–Schwarz and Lemmas 3.5 and 3.7, note that in the right-hand side we can replace t_z, t_w by t provided that we add an error term bounded by $ce^{-c't}$, uniformly in z and w.

Hence, by Fubini and equation (3.29),

$$\begin{aligned}
&\mathrm{Var}\Big[\int_{\mathbb{D}} \hat{h}_t(z) f(z) \mathbb{1}_{\mathcal{D}(z,A)}\, dz) | A\Big] \\
&\quad = \int_{\mathbb{D}^2} \mathbb{E}\Big[\prod_{y\in\{z,w\}} \big(\tilde{h}_t\big(g_z(y)\big) + \epsilon_y(t) - \mathbb{E}\tilde{h}_t\big(g_z(y)\big)\big)\mathbb{1}_{\mathcal{A}(t,y);\mathcal{D}(y,A)}\Big] \\
&\qquad \times f(z) f(w)\, dz\, dw \\
&\qquad + \mathsf{error}(t) \\
&\quad = \int_{\mathbb{D}^2} \mathrm{Cov}\big(\tilde{h}_t(0), \tilde{h}_t\big(g_z(w)\big)\big) f(z) f(w)\, dz\, dw + \mathsf{error}(t) \\
&\quad = \int_{z\in\mathbb{D}} f(z) \int_{y\in\mathbb{D}} \mathrm{Cov}\big(\tilde{h}_t(0)\tilde{h}_t(y)\big)\big|\big(g_z^{-1}\big)'(y)\big|^2 f\big(g_z^{-1}(y)\big)\, dy\, dz + \mathsf{error}(t),
\end{aligned}$$

where the error term satisfies $|\mathsf{error}(t)| \le cte^{-c't}$.

Pointwise convergence of the integrand comes from the definition of the two-point correlation function G. To conclude, we check that we can apply the dominated convergence theorem. By Lemma 3.11, one can find a δ such that for all $y \in \mathbb{D}$ with $|y| > 1-\delta$ and all $t > -12\log\delta + 1$, $|\mathrm{Cov}(\tilde{h}_t(0)\tilde{h}_t(y))| < 1$ almost surely. Therefore, on the set $\{|y| \ge 1-\delta\}$, the integrand in the last equality is bounded by $a(y,z) := |(g_z^{-1})'(y)|^2 \|f\|_\infty^2$. On the other hand, by Lemma 3.8 the integrand is bounded by $b(y,z) := \log(|y| \wedge (1-|y|))|(g_z^{-1})'(y)|^2 \|f\|_\infty^2$ when $|y| \le 1-\delta$.

Note that

$$\int_{\{|y|>1-\delta\}} \big|\big(g_z^{-1}\big)'(y)\big|^2\, dy = \mathrm{Leb}\big(\{w : \big|g_z(w) > 1-\delta\big|\}\big) \le \pi,$$

therefore $a(y,z)$ is integrable on $\{|y|>1-\delta\}$. Note also that if $|y|\le 1-\delta$, $|(g_z^{-1})'(y)| = R(g_z^{-1}(y), g_z^{-1}(\mathbb{D}))/R(y,\mathbb{D}) < c\delta^{-1}$, so b too is integrable on $\{|y|\le 1-\delta\}$. Thus, we can take limit inside the integral. Finally, one can use Proposition 3.6, item 1 to conclude. $\square$

We are now going to use the imaginary geometry coupling of Theorem 2.8 to prove the following consequence.

THEOREM 3.13. *Let f be any test function in $C^\infty(\bar{\mathbb{D}})$. Let $h = h^0_{\mathrm{GFF}} + \chi u_{\mathbb{D},1}$ be the Gaussian free field coupled with the UST according to Theorem 2.8, and let $h_{\mathrm{GFF}} = (1/\chi)h + \pi/2$. Then we have*

$$\lim_{t\to\infty} \mathbb{E}\big((h_t, f) - (h_{\mathrm{GFF}}, f)\big)^2 = 0.$$

In particular, (h_t, f) converges to (h_{GFF}, f) in $L^2(\mathbb{P})$ and in probability as $t\to\infty$.

PROOF. Fix $\varepsilon>0$. Using Lemma 3.11, pick δ such that we have $G(0,y)+2\log(1/|y|) < \varepsilon$ if $|y|\in(1-\delta,1)$. Fix η to be chosen suitably later (in a way which is allowed to depend on ε and δ). Let A be the set of branches of $\mathcal{T}$ from a "dense" set of points, $\frac{\eta}{4}\mathbb{Z}^2\cap\mathbb{D}$, to 1. Note that that $R(z,\mathbb{D}\setminus A)<\eta$ for any $z\in\mathbb{D}$ by Koebe's 1/4 theorem. Let $D' = \mathbb{D}\setminus A$. Define $\bar{h}_t(z) = \hat{h}_t \mathbb{1}_{\mathrm{dist}(z,A)>e^{-t/10}}$. First of all notice that

$$\begin{aligned}
&\mathbb{E}\big((h_t, f) - (h_{\mathrm{GFF}}, f)\big)^2\\
&\quad\le 2\mathbb{E}\big((\bar{h}_t, f) - (h_{\mathrm{GFF}}, f)\big)^2 + 2\mathbb{E}\big((\bar{h}_t, f) - (h_t, f)\big)^2.
\end{aligned}\tag{3.30}$$

By adding and removing $\hat{h}_t$ the last expression on the right hand side of (3.30) converges to 0 by (3.20) and the following fact: using Cauchy–Schwarz and the moment bounds on $\hat{h}_t(z)$,

$$\begin{aligned}
&\mathbb{E}\big((\bar{h}_t, f) - (\hat{h}_t, f)\big)^2\\
&\quad= \mathbb{E}\bigg(\int_{\mathbb{D}} \hat{h}_t(z) f(z)\mathbb{1}_{\mathrm{dist}(z,A)\le e^{-t/10}}\bigg)^2\\
&\quad\le \int_{\mathbb{D}^2} \mathbb{E}\big|\hat{h}_t(z)\hat{h}_t(w)\big| \|f\|_\infty^2 \mathbb{1}_{\mathrm{dist}(z,A)\vee\mathrm{dist}(w,A)\le e^{-t/10}}\,dz\,dw\\
&\quad\le c\|f\|_\infty^2(1+t)\int_{\mathbb{D}^2} \mathbb{P}\big(\mathrm{dist}(z,A)\vee\mathrm{dist}(w,A)\le e^{-t/10}\big)^{1/2}\,dz\,dw\\
&\quad\le c\|f\|_\infty^2(1+t)\mathbb{P}\big(\mathrm{dist}(U_1,A)\vee\mathrm{dist}(U_2,A)\le e^{-t/10}\big)^{1/2},
\end{aligned}$$

where $U_i\sim\mathrm{Unif}(\mathbb{D})$ and are independent of everything else and each other. Now if z is a point in $\frac{\eta}{4}\mathbb{Z}^2\cap\mathbb{D}$, then

$$\mathbb{P}\big(\mathrm{dist}(U,\gamma_z)\le e^{-t/10}\big)\le c_\eta e^{-c't}$$

by Lemma 2.9. Hence summing up over $O(1/\eta^2)$ points and using a union bound, we see that (for every fixed η), $\mathbb{P}(\mathrm{dist}(U_1,A)\vee\mathrm{dist}(U_2,A)\le e^{-t/10})^{1/2}\to 0$ exponentially fast and thus the second term on the right-hand side of equation (3.30) tends to 0.

Let $\mathbb{E}^A$ and Var^A denote the conditional expectation and variance given A. It is easy to see

$$\begin{aligned}
&\mathbb{E}^A\big((\bar{h}_t, f) - (h_{\mathrm{GFF}}, f)\big)^2\\
&\quad\le 3\,\mathrm{Var}^A(\bar{h}_t, f) + 3\,\mathrm{Var}^A(h_{\mathrm{GFF}}, f) + 3\big(\mathbb{E}^A(\bar{h}_t, f) - \mathbb{E}^A(h_{\mathrm{GFF}}, f)\big)^2.
\end{aligned}\tag{3.31}$$

Note that it is enough to show that as $t\to\infty$ the left-hand side of (3.31) can be made smaller than ε (in expectation) by choosing η suitably since this implies that the first term in the right-hand side of (3.30) is smaller than ε plus a term converging to zero, which completes the proof.

The last term of (3.31) converges to 0 for every η from the convergence of expectations in Lemma 3.12 and the fact that h_{GFF} satisfies the correct boundary conditions given A (which is a consequence of the imaginary geometry coupling).

For the other terms, recall that conditionally on A, h_{GFF} is just a free field in D' with variance $1/\chi^2$ and with Dirichlet boundary condition plus a harmonic function. Recall that the variance of a GFF integrated against a test function is given by an integral of the Green's function in the domain. Also recall that the Green's function is conformally invariant. In particular if g_z is the conformal map from D' to $\mathbb{D}$ sending z to 0 and 1 to 1, using the explicit formula for the Green's function in the unit disc, we have

$$\mathrm{Var}^A(h_{\mathrm{GFF}}) = -\int_{\mathbb{D}\times\mathbb{D}} \frac{1}{\chi^2} \log\left|g_z(w)\right| f(z) f(w)\, dz\, dw.$$

Plugging in the variance formula derived in Lemma 3.12 and since $\chi = 1/\sqrt{2}$,

$$\begin{aligned} &\mathbb{E}\left(\mathrm{Var}^A\left((h_{\mathrm{GFF}}, f)\right)\right) + \mathbb{E}\Big(\lim_{t\to\infty} \mathrm{Var}^A(\bar{h}_t, f)\Big)\Big) \\ &\quad = \mathbb{E}\left(\int_{\mathbb{D}\times\mathbb{D}} \left(G\left(0, g_z(w)\right) - 2\log\left|g_z(w)\right|\right) f(z) f(w)\, dz\, dw\right). \end{aligned} \tag{3.32}$$

By a change of variable $y = g_z(w)$,

$$\begin{aligned} &\int_{\mathbb{D}} \left(G\left(0, g_z(w)\right) - 2\log\left|g_z(w)\right|\right) f(w)\, dw \\ &\quad \le \|f\|_\infty \int_{\mathbb{D}} \left|\left(G(0, y) - 2\log|y|\right)\right| \left|(g_z^{-1})'(y)\right|^2 dy. \end{aligned} \tag{3.33}$$

As in Lemma 3.12, we are going to estimate the integral on two domains, $B_\delta = \{|y| < 1-\delta\}$ and $B'_\delta := \mathbb{D} \setminus B_\delta$. Recall that by the choice of δ, $G(0, y) + 2\log 1/|y| < \varepsilon$ if $y \in B'_\delta$. Hence,

$$\int_{B'_\delta} \left|\left(G(0, y) - 2\log|y|\right)\right| \left|(g_z^{-1})'(y)\right|^2 dy < \varepsilon\, \mathrm{Leb}\left(g_z^{-1}(B'_\delta)\right) < \pi\varepsilon.$$

To estimate the integral in B_δ, notice that $|(g_z^{-1})'(y)| = R(g_z^{-1}(y), D')/R(y, \mathbb{D}) < \eta/\delta$ if $y \in B_\delta$, by Koebe's 1/4 theorem. Hence,

$$\begin{aligned} &\int_{B_\delta} \left|\left(G(0, y) - 2\log|y|\right)\right| \left|(g_z^{-1})'(y)\right|^2 dy \\ &\quad < \frac{\eta^2}{\delta^2} \int_{B_\delta} \left|G(0, y) - 2\log|y|\right| dy. \end{aligned} \tag{3.34}$$

The integral on the right-hand side is finite via the bound Lemma 3.8. After bounding the integral over w, it remains to bound the integral over z in (3.32) by $\|f\|_\infty$ times the area. This completes the proof since we can choose η such that $\eta/\delta < \varepsilon$ where ε is arbitrary. $\square$

COROLLARY 3.14. *In the same setup as Theorem 3.13, for any $p > 0$, and any sequence of test functions $(f_1, \dots, f_n) \in C^\infty(\bar{\mathbb{D}})$ and integers $k_1, \dots, k_n$, as $t \to \infty$,*

$$\mathbb{E}\left[\left|\prod_{i=1}^n (h_t, f_i)^{k_i} - \prod_{i=1}^n (h_{\mathrm{GFF}}, f_i)^{k_i}\right|^p\right] \to 0. \tag{3.35}$$

PROOF. It is enough to prove this fact when p is an even integer. For $n = 1$, this follows from the fact that (h_t, f) converges in $L^2(\mathbb{P})$ towards (h_{GFF}, f), and (h_t, f) is bounded in L^p for any $p > 1$ by Lemma 3.8.

For general $n \geq 1$, we proceed by induction, and note that by the triangle inequality in L^p (i.e., Minkowski's inequality), if $X_n \to X$ and $Y_n \to Y$ in L^p for every $p > 1$ then $X_n Y_n \to XY$ in L^p for every $p > 1$. □

3.5. *General domains*. In this section, we state our result when D is a bounded domain with a locally connected boundary and we defer the proof to Theorem F.1 of the Supplementary Material [4]. Recall that our definition of $u_{(D,x)}$ in (2.7) only makes sense when the boundary is smooth in a neighbourhood of a marked point $x \in \partial D$ (while otherwise it is only defined up to a global additive constant; see Remark 2.7). The general idea is to show that in the limit one has

$$h^D \circ \varphi(z) = h^{\mathbb{D}}(z) + \arg_{\varphi'(\mathbb{D})}(\varphi'(z)), \qquad z \in \mathbb{D}$$

which is the imaginary geometry change of coordinates (see [33, 34]).

THEOREM 3.15. *Let D be as above. Let f be any bounded Borel test function defined on $\bar{D}$. Let $h = h^0_{\text{GFF}} + \chi u_{(D,x)}$ be the GFF coupled to the UST according to the imaginary geometry coupling of Theorem 2.8 and $u_{(D,x)}$ is as in (2.7). Then (h^D_t, f) converges to (h^D_{GFF}, f) in $L^2(\mathbb{P})$ and in probability as $t \to \infty$, where $h^D_{\text{GFF}} = \chi^{-1} h + \pi/2$.*

3.6. *Convergence in $H^{-1-\eta}(D)$*. Let D be a domain with locally connected boundary and now assume also that D is bounded. Let $(e_j)_{j\geq 1}$ denote the orthonormal basis of $L^2(D)$ given by the eigenfunctions of $-\Delta$ in D. Let h_{GFF} denote the process defined in Theorem 3.13, which is $(1/\chi)$ times a GFF with winding boundary conditions multiplied by χ. We now strengthen the convergence from a convergence in probability or $L^2(\mathbb{P})$ for finite-dimensional marginals to a convergence in the Sobolev space $H^{-1-\eta}$.

PROPOSITION 3.16. *For every $\eta > 0$, the field h_t converges to h_{GFF} in $H^{-1-\eta}$ in probability as $t \to \infty$. Further, $\{h_n\}_{n\in\mathbb{N}}$ converges almost surely to h_{GFF} as $n \to \infty$ along positive integers. Also for all $1 \leq k < \infty$, $\mathbb{E}[\|h_u - h_\infty\|^k_{H^{-1-\eta}}] \to 0$.*

PROOF. The basic idea is to show that h_t is a Cauchy sequence in $H^{-1-\eta}$. Let $u \geq t$. We start by getting bounds on $\mathbb{E}[(h_u - h_t, e_j)^2]$. By Fubini's theorem and Cauchy–Schwarz,

$$\begin{aligned}
&(\mathbb{E}[(h_u - h_t, e_j)^2])^2 \\
&\quad = \left(\int_{D^2} \mathbb{E}[(h_u(z) - h_t(z))(h_u(w) - h_t(w))] e_j(z) e_j(w)\, dz\, dw\right)^2 \\
&\quad \leq \int_{D^2} (\mathbb{E}[(h_u(z) - h_t(z))(h_u(w) - h_t(w))])^2\, dz\, dw \\
&\qquad \times \int_{D^2} e_j^2(z) e_j^2(w)\, dz\, dw \\
&\quad = \int_{D^2} (\mathbb{E}[(h_u(z) - h_t(z))(h_u(w) - h_t(w))])^2\, dz\, dw \qquad (3.36)
\end{aligned}$$

since e_j forms an orthonormal basis of $L^2(D)$. Let $r(z, w) = |z - w| \wedge \text{dist}(z, \partial D) \wedge \text{dist}(w, \partial D)$. We are going to break up the integral in (3.36) into two cases, either $t \leq$

$-10\log(r(z,w))+1$ (i.e., $r^{10} \le e^{1-t}$) or otherwise. In the first case Cauchy–Schwarz and the bound on moment of order two yield

$$\begin{aligned} &\int_{D^2} \big(\mathbb{E}\big[\big(h_u(z)-h_t(z)\big)\big(h_u(w)-h_t(w)\big)\big]\big)^2 \mathbb{1}_{t\le -10\log(r(z,w))+1}\,dz\,dw \\ &\quad \le c\big(1+u^2\big)\int_{D^2} \mathbb{1}_{r(z,w)^{10}\le e^{1-t}}\,dz\,dw \le c\big(1+u^2\big)e^{-c't}. \end{aligned} \tag{3.37}$$

On the other hand,

$$\begin{aligned} &\int_{D^2} \big(\mathbb{E}\big[\big(h_u(z)-h_t(z)\big)\big(h_u(w)-h_t(w)\big)\big]\big)^2 \mathbb{1}_{t> -10\log(r(z,w))+1}\,dz\,dw \\ &\quad \le cu^2 e^{-ct}\int_{D^2} \mathbb{1}_{t> -10\log(r(z,w))+1}\,dz\,dw \le c\big(1+u^2\big)e^{-c't}, \end{aligned} \tag{3.38}$$

where the second inequality above follows from (F.1) of the Supplementary Material [4] to formulate the correlation in the unit disc, and Lemma 3.7 to control this correlation by $r(z,w)$ (note that the bound of Lemma 3.7 holds also if $\hat{h}$ is replaced by h, because of the control on moments of h_t in Lemma 3.5 and the exponential bound on the probability of $\mathcal{A}(t,z)^c$). Combining (3.37) and (3.38), we obtain

$$\mathbb{E}\big[(h_u-h_t,e_j)^2\big] \le c\big(1+u^2\big)e^{-ct}. \tag{3.39}$$

Now let $t=n$ and $n\le u\le n+1$. Then, by Jensen's inequality,

$$\begin{aligned} \big[\mathbb{E}\|h_u-h_n\|_{H^{-1-\eta}}\big]^2 &\le \mathbb{E}\big[\|h_u-h_n\|^2_{H^{-1-\eta}}\big] \\ &= \sum_{j\ge 1}\mathbb{E}\big[(h_u-h_n,e_j)^2\big]\lambda_j^{-1-\eta} \\ &\le c\big(1+n^2\big)e^{-c'n}\sum_{j=1}^{\infty}\lambda_j^{-1-\eta} \end{aligned} \tag{3.40}$$

and since $\sum_{j=1}^{\infty}\lambda_j^{-1-\eta}<\infty$ by Weyl's law we deduce (by applying Markov's inequality and the Borel–Cantelli lemma) that h_n is almost surely a Cauchy sequence in $H^{-1-\eta}$ along the integers, and hence converge to a limit h_∞ almost surely in $H^{-1-\eta}$. Furthermore by the triangle inequality and (3.40) we get $\mathbb{E}(\|h_\infty-h_n\|_{H^{-1-\eta}})\le Ce^{-c'n}$. Then using (3.40) again, we deduce

$$\mathbb{E}\big[\|h_u-h_\infty\|_{H^{-1-\eta}}\big] \le C\big(1+u^2\big)e^{-c'u}$$

and hence h_u converges in probability in $H^{-1-\eta}$ to h_∞. To get convergence of $\mathbb{E}[\|h_u-h_\infty\|^k_{H^{-1-\eta}}$ for any $k\ge 1$ a similar argument would work: one needs to consider $\mathbb{E}(h_t-h_u,e_j)^{2k}$ and hence there would be k terms inside the integrals around (3.36). We skip this here because an exact similar argument with minor modifications is done in the proof of Theorem 5.1 later. Furthermore we have that $h_\infty=h_{\mathrm{GFF}}$ by considering the action on test functions and Theorem 3.13. This finishes the proof of Proposition 3.16 and hence also of Theorem 3.1. □

4. Discrete estimates on uniform spanning trees. The goal of this section is to gather the lemmas needed in Section 5 for the proof of the main result, Theorem 1.2. To make the purpose of the results in this section more clear, it will be useful for the reader to recall the general overview of the proof in Section 1.4. Recall that one additional difficulty comes from the fact that we need to deal with convergence of moments and not just convergence in law. Therefore we also need a priori estimates on the tails of our variables to use Cauchy–Schwarz bounds and dominated convergence theorems. In particular a bound on the tail of the winding of loop-erased random walk is derived in Section 4.3.

4.1. *Assumptions on the graph.* Let $\{G^{\#\delta}\}_{\delta>0}$ be a sequence of planar infinite (directed, weighted) graphs embedded properly in the plane. This means that for any $\delta > 0$ the embedding is such that no two edges cross each other. (The reader may think of δ usefully as a "mesh size" or microscopic scale). Vertices of the graph are identified with some points in $\mathbb{C}$ given by the embedding. We allow $G^{\#\delta}$ to have oriented edges with weights. A continuous time simple random walk $\{X_t\}_{t\geq 0}$ on such a graph $G^{\#\delta}$ is defined in the usual way: the walker jumps from u to v at rate $w(u,v)$ where $w(u,v)$ denotes the weight of the oriented edge (u,v). Given a vertex u in $G^{\#\delta}$, let $\mathbb{P}_u$ denote the law of continuous time simple random walk on $G^{\#\delta}$ started from u. For $A \subset \mathbb{C}$, we denote by $A^{\#\delta}$ the set of vertices of $G^{\#\delta}$ in A.

In this section, $B(a,r)$ will denote the set $\{z : |z-a| < r\}$. For $A \subset \mathbb{C}$, denote by $A + z := \{z + x : x \in A\}$ to be the translation of A by z. We assume $G^{\#\delta}$ has the following properties.

(i) (*Bounded density*) There exists C such that for any $x \in \mathbb{C}$, the number of vertices of $G^{\#\delta}$ in the square $x + [0,\delta]^2$ is smaller than C.

(ii) (*Good embedding*) The edges of the graph are embedded in such a way that they are piecewise smooth, do not cross each other and have uniformly bounded winding. Also, 0 is a vertex.

(iii) (*Irreducible*) The continuous time random walk on $G^{\#\delta}$ is irreducible in the sense that for any two vertices u and v in $G^{\#\delta}$, $\mathbb{P}_u(X_1 = v) > 0$.

(iv) (*Invariance principle*) The continuous time random walk $\{X_t\}_{t\geq 0}$ on $G^{\#\delta}$ started from 0 satisfies

$$(X_{t/\delta^2})_{t\geq 0} \xrightarrow[\delta\to 0]{(d)} (B_{\phi(t)})_{t\geq 0},$$

where $(B_t, t \geq 0)$ is a two dimensional standard Brownian motion in $\mathbb{C}$ started from 0, and ϕ is a nondecreasing, continuous, possibly random function satisfying $\phi(0) = 0$ and $\phi(\infty) = \infty$. The above convergence holds in law in Skorokhod topology.

(v) (*Uniform crossing estimate*) Let $\mathcal{R}$ be the horizontal rectangle $[0,3] \times [0,1]$ and $\mathcal{R}'$ be the vertical rectangle with same dimensions, and let $B_1 := B((1/2,1/2),1/4)$ be the *starting ball* and $B_2 := B((5/2,1/2),1/4)$ be the *target ball* (see Figure 3). There exist constants $\delta_0 > 0$ and $\alpha_0 > 0$ such that for all $z \in \mathbb{C}$, $\delta > 0$, $\ell \geq 1/\delta_0$, $v \in \ell\delta B_1$ such that $v + z \in G^{\#\delta}$,

$$\mathbb{P}_{v+z}\big(X \text{ hits } (\ell\delta B_2 + z) \text{ before exiting } (\ell\delta\mathcal{R} + z)\big) > \alpha_0. \tag{4.1}$$

The same statement as above holds for crossing from right to left, that is, for any $v \in \ell\delta B_2$, (4.1) holds if we replace B_2 by B_1. Also, the corresponding statements hold for the vertical rectangle $\mathcal{R}'$.

We point out that the invariance principle starting from zero, together with the crossing estimate, imply an invariance principle starting from arbitrary vertices $v^{\#\delta}$ in G^δ converging to a point $z \in \mathbb{C}$ as $\delta \to 0$. Briefly, by the invariance principle a walk starting from zero has a positive probability of making a small circuit around z; so that a walk starting from z can be coupled to the walk starting from zero the first time it hits this circuit (note that this time is a.s. finite by the crossing assumption).

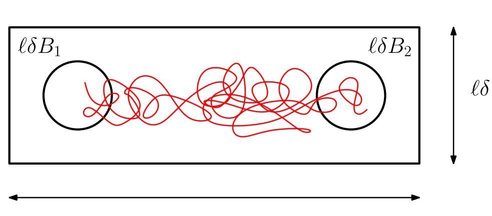

FIG. 3. *An illustration of the crossing condition.*

However, we also point out that that the crossing estimate would not follow from the invariance principle even if we assume it for all starting points; this would require a uniformity in the rate of convergence which does not necessarily hold in interesting examples for applications (in particular in the case of T-graphs which motivates our work).

REMARK 4.1. In this paper, we make crucial use of a result of Yadin and Yehudayoff [43] showing convergence of loop-erased random walk to SLE_2. This holds under an assumption of invariance principle for simple random walk. However, we need to spare a few words on their paper since they do not state their main theorems with quite the level of generality that we need here. Here are the points to note in order to check that their proofs extend to our setting.

1. They considered *backward* loop erased random walk, whereas we consider forward LERW (where loops are erased in chronological order). However, these have the same law, even when the graph is directed as is the case here (see [29]).
2. They consider scaled versions of a single infinite graphs instead of an arbitrary family of graphs with a scale parameter δ. However this is just for ease of notation as the proofs never use the relation between the graphs at different scales and all estimates are uniform over the underlying graphs.
3. The result of [43] is stated with $D = \mathbb{D}$ and $z = 0$, but this does not play a role in the proof. Notice in particular that the key estimate on the Poisson kernel ([43], Lemma 1.2) is stated with the generality we require, namely on an arbitrary domain and an arbitrary target point. More precisely, the target point is 0 but the domain is an arbitrary domain which contains 0: this, of course, amounts to the same thing as fixing and choosing an arbitrary fixed point inside a given domain. They also require that the inner radius of the domain (with respect to the target point 0) is greater than $1/2$. Up to a change of scale, this amounts to requiring that the point z is at positive distance from the boundary. The convergence in Lemma 1.2 of [43] is hence uniform in the domain if we assume that the distance from the boundary is bounded below. In our case, we will only use the result of [43] at a finite number of points which sit in the support of a compactly supported function f on D, so this assumption is certainly verified.

See in particular [43], Proposition 6.4 for a statement about the convergence of the driving function to Brownian motion in the general setup we require. Note that planarity of the graph plays a crucial role to prove this estimate. Also, the proof of tightness in the sense of Lemma 6.17 in [43] follows through in our situation with no significant modification.

REMARK 4.2. Let us briefly discuss the role of these assumptions. The invariance principle should be essentially a minimal assumption for the convergence. Indeed the Gaussian free field depends on the Euclidean structure of the plane and it is difficult to imagine any graph converging in a sense to the Euclidean plane without satisfying an invariance principle. In practice the invariance principle and irreducibility, together with the fact that there is no accumulation point, are exactly the assumptions needed for the convergence of the loop-erased random walk to SLE_2 from [43].

Our main additional assumption is the uniform crossing estimate. It is used extensively to derive various *a priori* estimates on the behaviour of the random walk, the uniformity over starting points and scale being a key factor for different multi-scale arguments. We believe however that there should be some room in our proofs to weaken this assumption.

The bounded density assumption is actually only needed for a union bound in the proof of Lemma 4.18. It is clear from that proof that it would not be needed if the uniform crossing assumption was allowed to “scale” with the local density of the graph.

For future reference, we note that the uniform crossing assumption can be rephrased equivalently by saying that there exist δ_0, α_0 such that for any $r > 0$, for any $\delta \le r\delta_0$, the probability to cross $r\mathcal{R}$ from rB_1 to rB_2 (left to right) and from rB_2 to rB_1 (right to left) is at least α_0, and likewise for the vertical rectangle.

4.2. *Russo–Seymour–Welsh type estimates*. Let D be a domain with locally connected boundary. To define the wired UST in the discrete domain, we perform the following surgery. For every oriented edge (xy) which intersect ∂D, we add an extra auxiliary vertex at the first intersection point (when following the embedded edge (xy)). We then replace (xy) by an oriented edge from x to this auxiliary vertex, keeping the same weight. The *wired graph* is the graph induced by all the vertices in $D^{\#\delta}$ along with all the auxiliary vertices and then wiring (or gluing) together all the auxiliary vertices. We denote by $\partial D^{\#\delta}$ all the edges with one endpoint being an auxiliary vertex and another endpoint inside D. The wired UST $\mathcal{T}^{\#\delta}$ is defined to be a uniform spanning tree on the wired graph. It is useful to think of the wired tree being sampled by Wilson's algorithm with the wired vertex being the initial root vertex. All the results in this section hold without the assumption of CLT (just assumptions (i) and (v) from Section 4.1 are needed).

We denote by $A(x, r, R)$ the annulus $\{z \in \mathbb{C} : r < |z - x| < R\}$. Let $v \in A(x, r, R)^{\#\delta}$. The random walk trajectory from a vertex v is the union of the edges it crosses (viewed as embedded in $\mathbb{C}$). We say random walk from v does a *full turn* in $A(x, r, R)$ if the random walk trajectory intersects every curve in the plane starting from $\{|z| = r\}$ and ending in $\{|z| = R\}$. We will write $X[a, b]$ for the random walk trajectory between times a and b. We will allow ourselves to see $X[a, b]$ and the loop-erased walk Y either as sequences of vertices, continuous paths in $\mathbb{C}$, or as sets depending on the place but this should not lead to any confusion. For any continuous curve $\lambda \in \mathbb{C}$, with a slight abuse of terminology we will freely say that "$(X_t, t \ge 0)$ *crosses (or hits)* λ at time $t > 0$" to mean that $X_t \ne X_{t^-}$ and the half-open edge $(X_{t^-}, X_t]$ intersects the range of λ.

In this section and the next, we will always assume that the loop-erased walk is generated by erasing loops chronologically from a simple random walk. We will allow ourselves to refer to the simple random walk associated to a loop-erased walk without further mention of this.

LEMMA 4.3. *Fix* $0 < r < R$, $\Delta = R - r$. *There exists constants* $c > 0$ *and* $\alpha > 0$ *depending only on* R/r *such that the following holds. For all* $\delta \le cr\delta_0$ *where* δ_0 *is as in item (v), for all* $x \in \mathbb{C}$ *and* $v \in A(x, r + \Delta/3, R - \Delta/3)^{\#\delta}$, *the probability that the random walk starting at* v *does a full turn before exiting* $A(x, r, R)$ *is at least* α.

PROOF. We use the uniform crossing assumption here and use the notation and terminology as described in Section 4.1. It is easy to see that we can find a sequence of rectangles $\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_k \subset A(x, r + \Delta/4, R - \Delta/4)^{\#\delta}$ where each such rectangle is a rectangle of the form $\varepsilon\mathcal{R} + z$ or $\varepsilon\mathcal{R}' + z$ (i.e., a scaling and translation of $\mathcal{R}$ or $\mathcal{R}'$) such that the starting ball of $\mathcal{R}_i$ coincides with the target ball of $\mathcal{R}_{i-1}$, v is in the starting ball of $\mathcal{R}_1$ and the following holds. If the simple random walk iteratively moves from the starting ball to the target ball of $\mathcal{R}_i$ for each $i = 1, \dots, k$ such that the starting vertex of $\mathcal{R}_{i+1}$ is the vertex where the walk enters the target ball of $\mathcal{R}_i$, then the walker accomplishes a full turn in $A(x, r, R)$. Here we can choose the scaling ε as a function of the ratio r, R and the number k to be bounded above by a constant $k_0(R/r)$. Applying the uniform crossing estimate and the Markov property of the walk, we see that this probability is bounded below by $\alpha_0^{k_0}$, thus completing the proof. $\square$

Actually, we will need estimates such as Lemma 4.3 to hold even when we condition on the exit point of the annulus which we will prove now. The first step is to prove a conditional version of the uniform crossing estimate.

LEMMA 4.4. *Fix $0 < r < R$ and $\epsilon < (R-r)/3$. There exist constants $c = c(R/r, \epsilon/r) > 0$, $\alpha = \alpha(R/r, \epsilon/r) > 0$ such that if $\delta \le cr\delta_0$ and $x, y \in \mathbb{C}$, the following holds. Let τ be the stopping time when the random walk exits $A(x, r, R)^{\#\delta}$. Let $\mathcal{R}$ be a rectangle of the form $y + [0, 3\eta] \times [0, \eta]$ such that $\mathcal{R} \subset A(x, r+\epsilon, R-\epsilon)$ and $\eta \ge \epsilon$. Let B_1 and B_2 be balls defined as in the uniform crossing estimate, i.e $B_1 = y + B((\frac{\eta}{2}, \frac{\eta}{2}), \frac{\eta}{4})$ and $B_2 = y + B((\frac{5\eta}{2}, \frac{\eta}{2}), \frac{\eta}{4})$. For all $v \in B_1^{\#\delta}$ and $u \in \partial A(x, r, R)^{\#\delta}$ such that $\mathbb{P}_v[X_\tau = u] > 0$,*

$$\mathbb{P}_v[X \textit{ hits } B_2 \textit{ before exiting } \mathcal{R} | X_\tau = u] > \alpha.$$

PROOF. The following argument is inspired by [43]. Let $h(v) = \mathbb{P}_v[X_\tau = u]$. We start by giving a rough bound on h restricted to $A(x, r+\epsilon, R-\epsilon)^{\#\delta}$. Let us fix $v, v' \in A(x, r+\epsilon, R-\epsilon)^{\#\delta}$. Since h is harmonic, there exists a path $\gamma = \{v, v_1, \dots\}$ from v to $\partial(A(x, r, R)^{\#\delta})$ along which h is nondecreasing. Also since h is harmonic and bounded, if τ_γ denotes the hitting time of $\gamma \cup \partial A(x, r, R)$ by a simple random walk, we have

$$h(v') = \mathbb{E}_{v'}\big[h(X_{\tau_\gamma})\big] \ge h(v)\mathbb{P}_{v'}[X_{\tau_\gamma} \in \gamma].$$

Using the crossing estimate a bounded number of times as in the proof of Lemma 4.3, it is clear that there exists a constant $\beta = \beta(R/r, \epsilon/r)$ independent of δ and v' such that

$$\begin{aligned}\mathbb{P}_{v'}\big[&X \text{ does a full turn in } A(x, r, r+\epsilon) \text{ and} \\ &\text{in } A(x, R-\epsilon, R) \text{ before exiting } A(x, r, R)\big] \ge \beta.\end{aligned}$$

We see that on the above event we have $X_{\tau_\gamma} \in \gamma$ so we have proved the Harnack inequality

$$\forall v, v' \in A(x, r+\epsilon, R-\epsilon)^{\#\delta}, \qquad \beta h(v) \le h(v') \le (1/\beta)h(v).$$

Now, together with the Markov property and the uniform crossing estimate, this gives

$$\begin{aligned}&\mathbb{P}_v[X \text{ hits } B_2 \text{ before exiting } \mathcal{R} \text{ and } X_\tau = u] \\ &\quad \ge \mathbb{P}_v(X \text{ hits } B_2 \text{ before exiting } \mathcal{R}) \inf_{v' \in B_2^{\#\delta}} h(v') \ge \alpha\beta h(v).\end{aligned} \tag{4.2}$$

Dividing by $h(v)$, the proof is complete. □

Using a bounded number of rectangles to surround the center x in $A(x, r, R)$ as in Lemma 4.3, we get the following corollaries.

COROLLARY 4.5. *Suppose we are in the setup of Lemma 4.4. Let τ be the stopping time when the random walk exits $A(x, r, R)^{\#\delta}$. Let $v \in A(x, r+\varepsilon, R-\varepsilon)$ and $u \in \partial A(x, r, R)^{\#\delta}$ such that $\mathbb{P}_v(X_\tau = u) > 0$. Then if $\delta < cr\delta_0$,*

$$\mathbb{P}_v\big(X \textit{ does a full turn in } A(x, r, R) | X_\tau = u\big) \ge \alpha = \alpha(R/r) > 0.$$

COROLLARY 4.6. *Fix $0 < r < R$. There exists a constant $c = c(R/r) > 0$ and $\alpha(R/r) > 0$ such that if $\delta \le cr\delta_0$ and $x \in \mathbb{C}$, $v \in A(x, r + \frac{R-r}{3}, R - \frac{R-r}{3})^{\#\delta}$ and if u is such that $\mathbb{P}_v(X_\tau = u) > 0$, where τ is the exit time of $B(x, R)$,*

$$\mathbb{P}_v\big[X \textit{ enters } B(x, r) \textit{ before exiting } B(x, R) | X_\tau = u\big] \ge \alpha.$$

The next lemma establishes an exponential tail for the winding of the simple random walk in an annulus conditioned to exit at a vertex, a key estimate to get an exponential tail on the winding of loop-erased random walk (Proposition 4.12). Recall that we write $X[t, t']$ for the random walk path between times t and t' and $W(\gamma, x)$ for the topological winding of a path γ around x.

LEMMA 4.7. *Fix $0 < r < R$. There exists $\alpha = \alpha(R/r) \in (0, 1)$ and c such that for all $x \in \mathbb{C}$, $\delta \in (0, cr\delta_0)$, $v \in A(x, r + \frac{R-r}{3}, R - \frac{R-r}{3})^{\#\delta}$, and for all u such that $\mathbb{P}_v(X_\tau = u) > 0$ where τ is the exit time of $A(x, r, R)$,*

$$\forall n \geq 1, \qquad \mathbb{P}_v\Big(\sup_{\mathcal{Y} \subset X[0,\tau]} \big|W(\mathcal{Y}, x)\big| \geq n | X_\tau = u\Big) \leq C(1-\alpha)^n,$$

where the supremum is over all continuous paths $\mathcal{Y}$ obtained by erasing portions from $X[0, \tau]$.

PROOF. The proof is technical and can be found in Lemma D.1 of the Supplementary Material [4]. The main idea is similar to the Harnack inequality in Lemma 4.4: we show that there is a curve cutting the annulus such that every time we wind around x in the annulus we hit this curve and there is then a positive probability to escape the annulus without increasing the winding number. □

Lemma 4.7 will allow us to control the random walk at all scales except the biggest one (the annulus around the center which intersects the complement of D) since in reality we stop the walk when it exits D; the following lemma allows us to control this largest scale.

LEMMA 4.8. *Fix $0 < r < R$. There exists $\alpha = \alpha(R/r) > 0$ and c such that for all $x \in \mathbb{C}$, $\delta \in (0, cr\delta_0)$, $v \in A(x, r + \frac{R-r}{3}, R - \frac{R-r}{3})^{\#\delta}$, writing τ for the exit time of $D \setminus B(x, r)$,*

$$\forall n \geq 1, \qquad \mathbb{P}_v\Big(\sup_{\mathcal{Y} \subset X[0,\tau]} \big|W(\mathcal{Y}, x)\big| \geq n | X_\tau \in \partial D^{\#\delta}\Big) \leq C(1-\alpha)^n,$$

and for all $u \in B(x, r)$ such that $\mathbb{P}(X_\tau = u) > 0$,

$$\forall n \geq 1, \qquad \mathbb{P}_v\Big(\sup_{\mathcal{Y} \subset X[0,\tau]} \big|W(\mathcal{Y}, x)\big| \geq n | X_\tau = u\Big) \leq C(1-\alpha)^n.$$

In both cases, the supremum is over all continuous paths $\mathcal{Y}$ obtained by erasing portions from $X[0, \tau]$.

Note that when we condition on exiting the domain, it is essential that we do not condition on the precise exit point. Indeed the stronger statement where we condition on this exit point does not necessarily hold if the domain itself winds many times around x.

PROOF. Details are similar to Lemma 4.7 and can be found in Lemma D.2 in the Supplementary Material [4]. □

Finally, recalling the definition of γ_z from (3.1), we need to bound the winding along the boundary of a domain $D^{\#\delta}$ from an arbitrary marked point to X_τ where τ is the hitting time of $\partial D^{\#\delta}$. Once again, as in Lemma 4.8, it is important here to note that we can only hope for a statement that is not conditional on X_τ since a priori the boundary winding can be arbitrarily large in some places—but these will have small harmonic measure.

LEMMA 4.9. *Let $\gamma_v^{\#\delta}$ denote the branch of the tree between v and the marked boundary vertex $x^{\#\delta}$. There exists $\alpha > 0$ and $c, C > 0$ depending only on the constants in the uniform crossing conditions item (v) such that the following holds. For all $\delta < c\delta_0 R$,*

$$\forall n \geq 1, \qquad \mathbb{P}_v\big(\big|W\big(\gamma_v^{\#\delta}[-1,0], v\big) - \mathbb{E}W\big(\gamma_v^{\#\delta}[-1,0], v\big)\big| \geq n\big) \leq C(1-\alpha)^n.$$

PROOF. The idea is to use the replica method, that is, to consider two independent samples of $\gamma_v^{\#\delta}$ and compare their boundary winding. Roughly, we can form a simple loop by considering the two paths after their last intersection and the portion of the boundary joining them. Since any simple loop has winding bounded by 2π, the winding of the portion of the boundary between the two paths is controlled by the winding of these two paths after their last intersection, whose tails are given by Lemma 4.8. The details are in Lemma D.3 of the Supplementary Material [4]. □

REMARK 4.10. We emphasise that in general the result of Lemma 4.9 does not hold if we remove the centering, in the same way that the result cannot hold conditionally on $\gamma_v(0)$.

Recall that we aim to decompose the winding into $h^{\#\delta} = h_t^{\#\delta} + \varepsilon^{\#\delta}$ (see equation (1.1)). We have already dealt with the limit of $h_t^{\#\delta}$ as $\delta \to 0$ in the continuum part. It now remains to say that $\varepsilon^{\#\delta}$ does not contribute because when $x \neq x'$ and t is large, $\varepsilon^{\#\delta}(x)$ and $\varepsilon^{\#\delta}(x')$ are nearly independent. This is proved in Section 4.4 by constructing a coupling of the sub tree around x and x' with independent variables. This coupling is built by sampling tree branches in the right order using Wilson algorithm and analysing carefully which part of the graph the random walks visits while performing the algorithm. In particular, a crucial step is to control the probability that the loop-erased walk from x comes close to x' which we now prove.

PROPOSITION 4.11. *There exists constants $c > 0$ and $\alpha > 0$ depending only on the constants α_0, δ_0 in the uniform crossing assumption item (v) such that the following holds. Let $D \subset \mathbb{C}$ be a domain and let $u, v \in D$. Let $r = |u - v| \wedge \operatorname{dist}(v, \partial D) \wedge \operatorname{dist}(u, \partial D)$. Let $v^{\#\delta}$ be the closest vertex to v in $G^{\#\delta}$. Let γ be a loop erased walk starting from $v^{\#\delta}$ until it exits $D^{\#\delta}$. For all $\delta \in (0, c\delta_0]$, for all $n \leq \log_4(cr\delta_0/\delta) - 1$ in $\mathbb{N}$,*

$$\mathbb{P}\big(\operatorname{dist}(u, \gamma) < 4^{-n}r\big) < (1-\alpha)^n.$$

PROOF. We assume $|u - v| = r$ for otherwise we can wait until the simple random walk comes closer to u. The idea for the proof is the following. If the loop erased walk comes within distance $4^{-k}r$ to u, then after the *last* time the random walk was within distance $4^{-k}r$ to u, it crossed k annuli without performing a full turn. The probability of this event is exponentially small in k via Corollary 4.5. Some care is needed since the above time is obviously not a stopping time.

Now we write the details. Let $i_{\max}(\delta) = \lfloor \log_4(c\delta_0 r/\delta)\rfloor$. Let $\{C_i\}_{0\leq i\leq i_{\max}}$ denote the circle of radius $4^{-i}r$ around u and define $C_{-1} = \partial D$. We inductively define a sequence of times $\{\tau_k\}_{k\geq 0}$ as follows. We have $\tau_0 = 0$. Having defined τ_k to be a time when the random walk crosses (or hits) some circle $C_{i(k)}$, we define τ_{k+1} to be the smallest time when leaving the annulus defined by $C_{i(k)-1}$ and $C_{i(k)+1}$, and define $i(k+1)$ to be the index of the circle by which the random walk leaves the annulus. If $i(k) = i_{\max}$, we define τ_{k+1} to be smallest time after τ_k when the walk leaves the ball defined by $C_{i_{\max}-1}$. We stop if we leave D and let N be the largest index of τ after which we stop.

Let $\mathcal{S} := (X_{\tau_k})_{0\leq k\leq N}$ denote the sequence of crossing positions of the C_i. Notice that conditioned on any realisation of $\mathcal{S}$, if $k < N$, the simple random walk between X_{τ_k} and $X_{\tau_{k+1}}$ is a simple random walk in the annulus $A(u, C_{i(k)-1}, C_{i(k)+1})$ conditioned to exit at

$X_{\tau_{k+1}}$. Furthermore by the Markov property of the walk, conditioned on $\mathcal{S}$, the portions of random walk $(X[\tau_k, \tau_{k+1}])_{0\le k\le N-1}$ are independent.

On the event $\mathrm{dist}(u,\gamma) < 4^{-n}r$, the sequence of positions $\mathcal{S} := \{X_{\tau_k}\}_{0\le k\le N}$ contains an index when the random walk crosses C_n since the loop-erased walk is obviously a subset of the walk. Let κ be the index of the last crossing of C_n by the walk and let γ' be the path obtained by erasing loops from $X[0,\tau_\kappa]$, that is, the "current loop-erased path" at time τ_κ. On the event $\mathrm{dist}(u,\gamma) < 4^{-n}r$, necessarily the random walk did not hit γ' after τ_κ, otherwise the part of the path closer to u would have been erased. In particular, the random walk did not do any full turn after time τ_κ. Also by construction we have $N-\kappa \ge n$.

Therefore, we see that conditioned on the sequence of positions $\mathcal{S}$, the event $\mathrm{dist}(u,\gamma) < 4^{-n}r$ is included in the event that there was no full turn in the last n intervals $[\tau_k,\tau_{k+1}]$. Choosing the constant c to be the constant from Corollary 4.5 associated with annuli of aspect ratio $R/r = 16$, we can apply that result for each annulus, since we condition on $\mathcal{S}$ and for each $i \le i_{\max}$, if r_i is the radius of C_i, we have $\delta \le c r_i \delta_0$ by our choice of $i_{\max}(\delta)$. Hence using the independence noted above, the conditional probability of this is at most $(1-\alpha)^n$ for some $\alpha > 0$, which concludes the proof. □

4.3. *Tail estimate for winding of loop-erased random walk.* Let $\gamma_v^{\#\delta}[0,\infty)$ denote the branch of the wired UST in $D^{\#\delta}$ from v to the wired vertex. As in the continuum we parametrise it by its capacity plus $\log R(v,D)$, so that time 0 corresponds to being on ∂D, and time ∞ to hitting v. We prove that $W(\gamma_v[t,t+1]^{\#\delta}, v)$ has exponential tail uniformly in δ and t. (Note that here we do not consider the contribution coming from the winding of the boundary between the marked boundary vertex $x^{\#\delta}$ and $\gamma_v^{\#\delta}(0)$.)

PROPOSITION 4.12. *There exist constants $C, c > 0$ depending only on the constant in the uniform crossing assumption item (v) such that the following holds. For all $v \in D^{\#\delta}$, for all $t \ge 0$, for all $\delta < ce^{-t} d(v, \partial D^{\#\delta})\delta_0$ and $n \ge 1$,*

$$\mathbb{P}\Big(\sup_{t\le t_1,t_2\le t+1} \big|W(\gamma_v[0,t_1]^{\#\delta}, v) - W(\gamma_v[0,t_2]^{\#\delta}, v)\big| > n\Big) < Ce^{-cn}.$$

We first set up some notation; these are analogous to the ones in the proof of Proposition 4.11 except that here we are considering circles of growing size.

Let $r_i = (4e)^{i-1}e^{-t}R(v,D)$ for $i \ge -1$ and $r_{-2} = 0$. Let C_i be the circle of radius r_i centered at v as long as $C_i \subset D$. As soon as C_i is not a subset of D, define $C_i = \partial D$ (call the maximum index $i_{\max}$ and allow us to make small abuse of notation such as calling $C_{i_{\max}}$ a circle or writing $D \setminus B(v, r_{i_{\max}-1}) = A(r_{i_{\max}-1}, r_{i_{\max}})$). Let X be a random walk from v run until it leaves the domain D. Let $Y(t)$ be the loop-erasure of X, reparametrised by capacity seen from v plus $\log R(v,D)$ (so in particular $Y(\infty) = v$). We emphasise that we are indexing circles starting from $i = -2$.

We inductively define a sequence of times $\{\tau_k\}_{k\ge 0}$ as follows. We have $\tau_0 = 0$ and $i(0) = -2$, τ_1 is the time the random walk crosses C_{-1} and $i(1) = -1$. Having defined τ_k to be the smallest time when the random walk crosses (or hits) some circle $C_{i(k)}$, we define τ_{k+1} to be the smallest time when it hits $C_{i(k)-1}$ or $C_{i(k)+1}$ and define $i(k+1)$ to be the index of the circle it crosses. When $i(k) = -1$ we define τ_{k+1} only as the next crossing of C_0.

Let k_{exit} be the first index such that $i(k) = i_{\max}$, that is, the index corresponding to the exit from $D^{\#\delta}$ and we let $\mathcal{S} = (X_{\tau_k})_{0\le k<k_{\mathrm{exit}}}$ the sequence of crossing positions, not including the exit position. For any j, we also define $\mathcal{V}_j$ to be the sequence of crossings of the circle C_j, that is, $\mathcal{V}_j = \{k | i(k) = j\}$. We will call sets of the form $X[\tau_k,\tau_{k+1}]$ *elementary piece of random walk* and we note that conditionally on $\mathcal{S}$ they are all independent. In this proof, we call *crossing* a time of the form τ_k.

Actually we prove something stronger than the statement in the theorem. Let B_0 be the disc with boundary C_0. We look at the portion of Y *outside* B_0 counted from the first time it enters C_1 until the last time it enters C_0 and bound the maximal winding of any sub-portion of this part of the path. Note that from Koebe's 1/4 theorem for any $t \geq 0$, the circle C_0 does not intersect ∂D and the times t and $t+1$ happen in the interval above.

LEMMA 4.13. *There exist constants $C, c > 0$ depending only on the constant in the uniform crossing assumption item (v) such that the following holds. For all $v \in D^{\#\delta}$, for all $t \geq 0$, for all $\delta < Ce^{-t} d(v, \partial D^{\#\delta})\delta_0$, let $\mathcal{Y}$ denote the portion of Y from the first time it enters C_1 until the last time it enters C_0, then*

$$\forall n \geq 1, \qquad \mathbb{P}\Big(\sup_{\tilde{\mathcal{Y}} \subset \mathcal{Y} \cap B_0^c} |W(\tilde{\mathcal{Y}}, v)| > n \Big) < Ce^{-cn}.$$

We now prove Lemma 4.13 which immediately implies Proposition 4.12. The main idea of the proof of Lemma 4.13 will be to show that this portion of the loop erased walk can be generated by erasing loops from a small number of elementary pieces of random walk. Then Lemma 4.7 will show that each piece does not contribute too much to the winding.

To control the number of pieces, we will use two elements. First, conditionally on $\mathcal{S}$ we will argue that we only need to look at the last few visits of the simple random walk to any circle because everything else was erased by a loop. Second, we will show that the sequence $\mathcal{S}$ is not too badly behaved even when we are looking close to the last visit to C_0. Note that this is non trivial because the last visit to C_0 is very far from being a stopping time. We now proceed to the actual proof, writing each of these steps as lemmas.

We first note some deterministic facts about the loop erasure (Lemma 4.14), therefore until further notice we work on a given realisation of the random walk. Let $k_{\max} - 1$ be the index of the last crossing of C_1 by the walk (or k_{exit} if $C_1 = \partial D$). Only indices less than $k_{\max}$ will be of interest to us, as only those can contribute to the loop-erasure in the range we are considering.

Let κ_{-1} be the last k such that $i(k) = -1$ and in the interval $[\tau_k, \tau_{k+1}]$ the random walk did a full turn in $A(v, r_{-1}, r_0)$. If there was no such full turn, set $\kappa_{-1} = 0$. Now we define κ_i inductively as follows.

- The time κ_i is the last time after κ_{i-1} (but still before $k_{\max}$) where a full turn occurs in the annulus $A(v, r_i, r_{i+1})$.
- In case no such full turn occurs, κ_i is the index of the first crossing of C_i after κ_{i-1} (but before $k_{\max}$).
- Finally, if there is no crossing of C_i between κ_{i-1} and $k_{\max}$ define $\kappa_i = +\infty$.

Finally, we define $I = \max\{i : \kappa_i < \infty\}$. The idea is that a full turn erases a lot of the past. After a full turn, we may still visit larger scales without doing a full turn in those scales and these might not get erased, so we will need to consider them. This is the role of the random variable I which gives a bound on the largest such scale.

For every $i \leq I$, we let $\mathcal{G}_i$ be the set of visits to C_i after κ_{i-1} but before $k_{\max}$, i.e., $\mathcal{G}_i = \{k \in \mathcal{V}_i | \kappa_{i-1} \leq k \leq k_{\max}\}$. Observe that $\mathcal{G}_i$ are "good" indices which matter for the loop-erasure.

LEMMA 4.14. *Recall $\mathcal{Y}$, B_0 from Lemma 4.13. Then*

$$\mathcal{Y} \cap B_0^c \subset \bigcup_{0 \leq i \leq I} \bigcup_{k \in \mathcal{G}_i} X[\tau_k, \tau_{k+1}].$$

Furthermore, one can write $\mathcal{Y} \cap B_0^c = \bigcup_{i \leq I} \bigcup_{k \in \cup \mathcal{G}_i} \mathcal{Y}_k$ where $\mathcal{Y}_k$ are disjoint intervals of the loop erased random walk of the form $\mathcal{Y}_k = (Y_{j_k}, Y_{j_k+1}, \dots, Y_{j_k+i_k})$ and $\mathcal{Y}_k \subset X[\tau_k, \tau_{k+1}]$.

PROOF. The proof is by inspection; see Lemma D.4 of the Supplementary Material [4]. □

The next step is to control the law of the size of the sets $\mathcal{G}_j$, therefore we go back to considering X as random.

LEMMA 4.15. *There exist constants* $C, c, c' > 0$, *such that for all* $\delta \le ce^{-t}\,\mathrm{dist}(v, \partial D^{\#\delta})\delta_0$, *for all* $n > 0$,

$$\mathbb{P}\Big(\sum_{0\le i\le I} |\mathcal{G}_i| \ge n\Big) \le C\exp(-c'n).$$

PROOF. This is the most delicate part of the proof of Proposition 4.12 and is included in Lemma D.5 of the Supplementary Material [4]. To explain briefly:

– It is easy to check that I itself has geometric tail (conditionally on $\mathcal{S}$, each i such that $\kappa_i < \infty$ requires not making a full turn immediately after the last visit to C_i).

– It is also immediate to see with a similar argument that the number of crossings of C_i after τ_{κ_i} has geometric tail.

– Therefore, to get an exponential tail on $|\mathcal{G}_i|$ it remains to exclude the possibility that the walk oscillates many times between C_i and C_{i+1} before the next visit to C_{i-1}. Of course, the idea is to exploit the fact that every time the walk visits C_i there is a positive chance to hit C_{i-1} first rather than C_{i+1}. However this is technically tedious to implement since these visits are not stopping times and we can not directly condition on $\mathcal{S}$ this time. Instead we choose to discover $\mathcal{S}$ step by step by revealing only the portion of $\mathcal{S}$ that is outside of C_i and using Corollary 4.6. □

Now it is easy to complete the proof of Lemma 4.13 using Lemma 4.7. By Lemma 4.14, we can write $\mathcal{Y} \cap B_0^c = \bigcup_k \mathcal{Y}_k$ with for all k, $\mathcal{Y}_k \subset X[\tau_k, \tau_{k+1}]$. Therefore, the winding around v of any $\mathcal{Y}_k$ is bounded by the maximal winding difference between two times in $[\tau_k, \tau_{k+1}]$ which has uniform exponential tail by Lemma 4.7 or Lemma 4.8 for the pieces in $\mathcal{G}_{i_{\max}}$ and are independent since the walks are independent conditionally on $\mathcal{S}$. Note that crucially this independence holds even when we do not condition on $\gamma_v^{\#\delta}(0) = X(\tau_{k_{\text{exit}}})$, making Lemma 4.8 applicable. By Lemma 4.15 the number of terms in the union has exponential tail, so the proposition follows.

We now put together a consequence of the above estimates in a single lemma for ease of reference later on. Pick $v \in D^{\#\delta}$ and $t \ge 0$. Recall the definition of $\gamma_v^{\#\delta}[0, t]$ from Proposition 4.12. Now add to it a portion of the boundary $\partial D^{\#\delta}$ from the marked boundary point on $\partial D^{\#\delta}$ to $\gamma_v^{\#\delta}(0)$ (i.e., the point where the branch hits the wired boundary). Parametrise the resulting curve as $\gamma_v^{\#\delta}[-1, t]$.

LEMMA 4.16. *For any* $k \ge 1$, *there exists a constant* $A = A_k > 0$ *depending only on the constants in the crossing assumptions* (*item* (v)) *and* k *such that for all* $\delta < ce^{-t}\,\mathrm{dist}(v, \partial D^{\#\delta})\delta_0$, *we have*

$$\mathbb{E}\big(\big|W(\gamma_v^{\#\delta}[-1, t), v) - \mathbb{E}[W(\gamma_v^{\#\delta}[-1, t), v)]\big|^k\big) \le A(t+1)^k.$$

PROOF. Recall the definition of the circles C_i used in the proof of Lemma 4.13. Note that Lemma 4.13 provides exponential tail for the winding of $\gamma_v^{\#\delta}$ from the first entry into C_{i+1} until the first entry into C_i for $i \ge i_{\max} - 2$ (note that this is a continuous subportion of the loop erasure seen until last entry into C_i and this portion is completely outside C_i), which immediately implies the same bound for the same centered random variable. Likewise, the first item of Lemma 4.8 provides exponential tail on the (centered) winding from $\gamma_v^{\#\delta}$ until the first entry into $C_{i_{\max}-1}$. Finally, Lemma 4.9 provides exponential tail of the centered winding of $\gamma_v^{\#\delta}[-1, 0]$. Putting all of this together, we see that we need to bound the kth moment of a sum of $O(t+1)$ random variables with uniform exponential tails, which is elementary. $\square$

4.4. *Local coupling of spanning trees.* Let $z_1, \dots, z_k \in D^{\#\delta}$. The goal of this section is to establish a coupling between a wired uniform spanning tree $\mathcal{T}^{\#\delta}$ in $D^{\#\delta}$ and k independent copies of full plane spanning tree measure $\mathcal{T}_i$ such that for all i, there is a neighbourhood N_i around z_i on which $\mathcal{T}^{\#\delta}$ matches with $\mathcal{T}_i^{\#\delta}$. The diameter of the neighbourhoods N_i are going to be random but nevertheless we will have a good bound on the probability of the diameter being very small; typically the neighbourhoods will be of the optimal scale (i.e., not much smaller than the distance to the nearest vertex v_j with $j \neq i$ or to the boundary). Note that *a priori* it is not even clear that the full plane local weak limit of a wired spanning tree exists. For undirected graphs, the existence of this limit follows from the theory of electrical networks [35]. However our setting includes directed graphs where the electrical network theory no longer applies. The existence of this limit will actually come out of our coupling procedure.

The overall strategy will be to sample the spanning tree $\mathcal{T}^{\#\delta}$ in $D^{\#\delta}$ and $(\mathcal{T}_i^{\#\delta})_{1 \le i \le k}$ using Wilson's algorithm. The coupling will mostly be achieved by using the same random walks for $\mathcal{T}^{\#\delta}$ and $(\mathcal{T}_i^{\#\delta})_{1 \le i \le k}$. To achieve independence and obtain the tail estimate for the diameters of the neighbourhoods N_i, we will choose carefully the points from which we sample loop-erased walks and keep track of the distances from $\{z_i\}$ to the sub-tree discovered at any step.

We start with a simple lemma regarding the hitting probability of random walk. This is a reformulation of Lemma 2.1 from Schramm [39] in our setting.

LEMMA 4.17. *There exist constants $C, c, c' > 0$ such that for all connected set $K \subset \mathbb{C}$ such that the diameter* (*in the metric inherited from the Euclidean plane*) *of K is at least R, for all $\delta \in (0, C \operatorname{dist}(v, K)\delta_0)$,*

$$\mathbb{P}_v\big(X \text{ exits } B(v, R)^{\#\delta} \text{ before hitting } K^{\#\delta}\big) \le c\left(\frac{\operatorname{dist}(v, K)}{R}\right)^{c'}.$$

PROOF. Let C_i denote the circle of radius 2^{-i} around v for $i \in \mathbb{Z}$. Consider a sequence of stopping times $\{T_k\}_{k \ge 0}$ defined as in Proposition 4.11: if T_k is the time when the walk crosses C_i then T_{k+1} is the smallest time after T_k when the simple random walk crosses C_{i+1} or C_{i-1}. The number of circles which intersect K is at least $c \log_2(\frac{R}{\operatorname{dist}(v,K)})$ for some $c > 0$. The choice of δ is small enough for Lemma 4.3 to apply for the annuli bounded by these circles. Whenever the walk at T_k is in a circle C_i such that both C_{i-1} and C_{i+1} are subsets of D and intersect K, then the walk has probability at least $\alpha > 0$ of performing a full turn in $A(v, 2^{-i-1}, 2^{-i+1})$ via Lemma 4.3. But doing such a full turn implies the walk must hit K. Hence, the probability of the walk exiting $D^{\#\delta}$ without hitting K has probability at most $(1-\alpha)^{c' \log_2(\frac{R}{\operatorname{dist}(v,K)})}$ for some $c' > 0$ which concludes the proof. $\square$

Let D be a fixed bounded domain, let $v \in D^{\#\delta}$, and let r be such that $B(v, r) \subset D$. Using Wilson's algorithm, we now prescribe a way to sample the portion of the wired uniform spanning tree $\mathcal{T}^{\#\delta}$ of $D^{\#\delta}$ which contains all the branches emanating from vertices in $B(v, r/2)^{\#\delta}$. Consider $\{\frac{r}{2} 6^{-j}\mathbb{Z}^2\}_{j \ge 0}$, a sequence of scalings of the square lattice $\mathbb{Z}^2$ which divides the plane into square cells. At step j, pick a vertex from each cell f of $\frac{r}{2}6^{-j}\mathbb{Z}^2$ which is farthest from v in $B(v, \frac{r}{2}(1 + 2^{-j}))^{\#\delta} \cap f$ (break ties arbitrarily) and is not chosen in any previous step. Call $\mathcal{Q}_j$ the set of vertices picked in step j. Now sample branches of $\mathcal{T}^{\#\delta}$ from each of these vertices in any prescribed order via Wilson's algorithm, resulting in a partial tree $\mathcal{T}_j^{\#\delta}$. We continue until we exhaust all the vertices in $B(v, r/2)^{\#\delta}$. We call this algorithm the *good algorithm* $\mathrm{GA}_D^{\#\delta}(r, v)$ to sample the portion of $\mathcal{T}^{\#\delta}$ containing all branches emanating from vertices in $B(v, r/2)^{\#\delta}$ (and in particular, the restriction of $\mathcal{T}^{\#\delta}$ to $B(v, r/2)^{\#\delta}$). Note in particular that $\mathrm{GA}_D^{\#\delta}(r, v)$ is sure to terminate after step $j = \log_6(Cr/\delta)$, where C depends only on the constant appearing in the bounded density assumption (assumption (i)).

The next lemma is similar to Schramm's finiteness theorem from [39]. Roughly, this says that for all $\varepsilon > 0$, if we fix a ρ sufficiently small depending only on ε, and reveal the branches of the spanning tree at a finite number of points with density approximately $1/\rho$, then with high probability none of the remaining branches would have diameter greater than ε. Schramm's finiteness theorem is originally stated for the diameter of the remaining branches of the spanning tree (which are loop-erased paths) but in fact the result holds for the underlying random walks themselves. Also the original theorem is interested in sampling the whole tree while we only want to sample $\mathcal{T}^{\#\delta} \cap B(v, r)$ for some r. Since we will need these properties later on, we write it for completeness, but the proof is exactly the same as in [39].

LEMMA 4.18 (Schramm's finiteness theorem). *Fix $\varepsilon > 0$ and let D, v, r be as above. Then there exists a $j_0 = j_0(\varepsilon)$ depending solely on ε such that for all $j \geq j_0$ and all $\delta \leq 6^{-j_0}\delta_0 r$, where δ_0 is as in item (v), the following holds with probability at least $1 - \varepsilon$:*

- *The random walks emanating from all vertices in $\mathcal{Q}_j$ for $j > j_0$ stay in $B(v, r)$.*
- *All the branches of $\mathcal{T}^{\#\delta}$ sampled from vertices in $\mathcal{Q}_j \cap B(v, r/2)$ for $j > j_0$ until they hit $\mathcal{T}_{j_0}^{\#\delta} \cup \partial D^{\#\delta}$ have Euclidean diameter at most εr. More precisely, the connected components of $\mathcal{T}^{\#\delta} \setminus \mathcal{T}_{j_0}^{\#\delta}$ within $B(v, r/2)$ have Euclidean diameter at most εr.*

PROOF. For $j \geq 1$, the number of vertices in $\mathcal{Q}_j$ is at most $c6^j$ where c is a universal constant. Let $j_{\max} := \lfloor \log_6(\frac{C\delta_0 r}{\delta}) \rfloor$. The choice of $j_{\max}$ is such that for $j \leq j_{\max}$ our uniform crossing assumption holds and in particular we can apply Lemma 4.17. Notice each vertex in $D^{\#\delta}$ is within Euclidean distance $4 \cdot 6^{-j} r$ from a vertex in $\mathcal{Q}_{j-1}$. By Lemma 4.17 and the choice of δ, for $j \leq j_{\max}$, there exists a C_0 such that the probability that the simple random walk from a vertex in $\mathcal{Q}_j$ reaches Euclidean distance $C_0 6^{-j} r$ from its starting point without hitting $\mathcal{T}_{j-1}^{\#\delta}$ is at most $1/2$. Notice that $j^2 6^{-j} < 2^{-j}$ for all $j \in \mathbb{N}$ and hence using the Markov property, we can iteratively apply the same bound for the walk j^2/C_0 times (this is the reason why in the good algorithm we sample from balls of decreasing size at each step). This shows that the probability that the random walk emanating from a vertex w in $\mathcal{Q}_j$ has diameter greater than $j^2 6^{-j} r$ (call this event $\mathcal{D}(w, j)$) is at most $(1/2)^{j^2/C_0}$.

Recall that the bounds above hold for $j \leq j_{\max}$. When $j > j_{\max}$ and $w \in \mathcal{Q}_j$ we define $\mathcal{D}(w, j)$ to be the event that the random walk emanating from w reaches distance $j_{\max}^2 6^{-j_{\max}} r$ without hitting $\mathcal{T}_j^{\#\delta}$ (and in particular $\mathcal{T}_{j_{\max}}^{\#\delta}$). Then observe that we still have $\mathbb{P}(\mathcal{D}(w, j)) \leq (1/2)^{j_{\max}^2/C_0}$ in this case. Notice that the number of lattice points in $\bigcup_{j > j_{\max}} \mathcal{Q}_j$ is at most $B6^{j_{\max}}$ for some constant B depending only on δ_0 in the uniform crossing estimate assumption and bounded density assumptions (items (i) and (v)).

Notice that on the complement of

$$\mathcal{D} := \bigcup_{j_0 \leq j} \bigcup_{w \in \mathcal{Q}_j} \mathcal{D}(w, j)$$

no random walk emanating from a vertex $w \in \mathcal{Q}_j$ can reach distance $j^2 6^{-j} r$ from its starting point and hence stays in $B(v, r)$. Furthermore, by a union bound,

$$\mathbb{P}(\mathcal{D}) \leq \sum_{j \geq j_0} c 6^j (1/2)^{j^2/C_0} < \varepsilon \tag{4.3}$$

for large enough choice of $j_0 = j_0(\varepsilon)$ which shows the first property of the lemma.

Also, let $\mathcal{E}(w, j)$ be the event that a connection from w to $\mathcal{T}_{j-1}^{\#\delta}$ has diameter at least $j^2 6^{-j} r$. Observe that conditionally on $\mathcal{T}_{j-1}^{\#\delta}$, the probability of the event $\mathcal{E}(w, j)$ does not depend on the order of the points in $\mathcal{Q}_j$ and hence we may assume that w is the first point in $\mathcal{Q}_j$ when we compute this probability. In that case, the probability in question is at most the one we computed above, and we deduce that $\mathbb{P}(\mathcal{E}(w, j) | \mathcal{T}_{j-1}^{\#\delta}) \leq (1/2)^{(j \wedge j_{\max})^2/C_0}$. On the complement of

$$\mathcal{E} := \bigcup_{j_0 \leq j} \bigcup_{w \in \mathcal{Q}_j} \mathcal{E}(w, j) \tag{4.4}$$

each point $w \in \mathcal{Q}_j$ is connected to a point in $\mathcal{T}_{j_0}^{\#\delta}$ by a path of diameter at most $\sum_{j > j_0} j^2 6^{-j} r \leq \varepsilon r$ provided that j_0 is large enough. As $\mathbb{P}(\mathcal{E}) \leq \varepsilon$ if j_0 is large enough, the proof is complete. $\square$

Now we shall describe the coupling between a wired spanning tree in $D^{\#\delta}$ and a full plane spanning tree around a single point, which we call *base coupling*. (The final coupling will be nothing more but an iteration of this procedure with an extra step initially which we call cutset exploration.) Recall that a priori it is not even clear that the full plane local weak limit of a wired spanning tree exists (the existence of this limit will actually come out of our coupling procedure).

Basically, the idea is the following. We wish to couple a uniform spanning in tree in $D^{\#\delta}$ to a uniform spanning tree in $\tilde{D}^{\#\delta}$ within a neighbourhood of some fixed vertex v. To do this, we first make sure that the branches from a finite number of vertices coincide for UST's in $D^{\#\delta}$ and $\tilde{D}^{\#\delta}$ in some neighbourhood, and then we apply Schramm's finiteness theorem. We now explain this in detail.

Base coupling. The base coupling we describe now takes the following parameters as input: two domains $D^{\#\delta}$, $\tilde{D}^{\#\delta}$ and a neighbourhood $B(v, 10r)$ of a vertex v such that $B(v, 10r)^{\#\delta} \subset D^{\#\delta} \cap \tilde{D}^{\#\delta}$. Let $\mathcal{T}$ and $\tilde{\mathcal{T}}$ denote a sample of uniform spanning tree in $D^{\#\delta}$ and $\tilde{D}^{\#\delta}$, respectively. For any vertex u in $D^{\#\delta}$ (resp. $\tilde{D}^{\#\delta}$), let $\gamma(u)$ (resp. $\tilde{\gamma}(u)$) denote the branch of $\mathcal{T}$ (resp. $\tilde{\mathcal{T}}$) from u to the boundary of $D^{\#\delta}$ (resp. $\tilde{D}^{\#\delta}$).

(i) Pick a point u_1 in $A^{\#\delta}(v, 8r, 9r)$ and sample $\gamma(u_1)$, $\tilde{\gamma}(u_1)$ independently (any joint law could work but we take them independent for concreteness). Let $\mathcal{E}_1$ be the event that both $\gamma(u_1)$ and $\tilde{\gamma}(u_1)$ stay outside $B(v, 7r)$ and suppose $\mathcal{E}_1$ holds.

(ii) Let $u_2 \in A^{\#\delta}(v, 2r, 3r)$. We will use the same underlying random walk to couple $\gamma(u_2)$ and $\tilde{\gamma}(u_2)$. More precisely, start a simple random walk from u_2 until it is in one of $\gamma(u_1) \cup \partial D^{\#\delta}$ or $\tilde{\gamma}(u_1) \cup \partial \tilde{D}^{\#\delta}$ at time t_1. Suppose without loss of generality that the walk hits $\gamma(u_1) \cup \partial D^{\#\delta}$ at t_1. Then we continue the walk from that point until it hits the other path or the boundary at time t_2. We then define $\gamma(u_2)$ to be the loop erased path up to time t_1 and $\tilde{\gamma}(u_2)$ to be the loop erased path from time 0 to t_2. Let $\mathcal{E}_2$ be the event that $\gamma(u_2)$ and $\tilde{\gamma}(u_2)$ agree in $B(v, 6r)$, and suppose $\mathcal{E}_2$ holds.

(iii) Fix a $j_0 = j_0(1/2)$ as defined in Lemma 4.18. Let $\mathcal{Q}_j$ be a set of points in $B(v, (r/2)(1 + 2^{-j})^{\#\delta}$, one in each cell of $6^{-j}(r/2)\mathbb{Z}^2$ chosen that it is furthest away from v within that cell, as described in the good algorithm above. Let $\mathcal{E}_3$ be the event that the branches from all the vertices in $\bigcup_{j\le j_0} \mathcal{Q}_j$ of $\mathcal{T}$, $\tilde{\mathcal{T}}$ agree in $B(v,5r)$, and suppose that $\mathcal{E}_3$ holds.

(iv) Finally, let $\mathcal{E}_4$ be the event that all the branches from vertices in $\bigcup_i \mathcal{Q}_i$ of $\mathcal{T}$, $\tilde{\mathcal{T}}$ agree in $B(v,r/2)$.

In the steps above if $\bigcap_i \mathcal{E}_i$ does not occur, we say that the *base coupling has failed*. We think of the above process as sampling branches one by one from the prescribed vertices. We stop this process of sampling branches at the first time a sampled branch causes the base coupling to fail.

LEMMA 4.19. *There exist constants* $0 < p_0 < p_0' < 1$ *and* $c > 0$ *such that for all* $\delta \le c\delta_0 r$,

$$p_0 < \mathbb{P}(\textit{base coupling succeeds in } D^{\#\delta}) < p_0'.$$

PROOF. The proof essentially follows from Proposition 4.11, which says that loop erased random walk does not come too close to a particular vertex, and Lemma 4.3 which says that random walk makes a full turn in a given annulus with positive probability. To start with, it follows from Proposition 4.11 (possibly replacing 4 there by some other number) that $\mathcal{E}_1$ has a positive probability p_1. Also using the crossing estimate it is easy to see that $\mathbb{P}(\mathcal{E}_1) < p_0' < 1$ which completes the proof of the upper bound. Moreover, independently of $\mathcal{E}_1$, the walk started from u_2 after exiting $B(v,7r)$ has probability at least p_2 to make a full turn in $A(v,9r,10r)$ without first hitting $B(v,6r)$ by Lemma 4.3. In particular, this implies $\mathcal{E}_2$ has probability at least p_2, conditionally on $\mathcal{E}_1$.

Now assume $\mathcal{E}_1 \cap \mathcal{E}_2$ holds. Let $w \in \mathcal{Q}_j$ for $j \le j_0$, and assume that revealing previous branches did not make the coupling fail. Then the walk started from w has a positive probability p_3 to do a full turn in $A(v,3r,5r)$ before leaving $B(v,5r)$. If this occurs then the corresponding branches $\gamma(w)$ and $\tilde{\gamma}(w)$ will agree in $B(v,5r)$ (since the walk is then certain to hit at least both $\gamma(u_2)$ and $\tilde{\gamma}(u_2)$ in that ball). Iterating this bound over a bounded number of points (of order 6^{j_0}) shows that, conditionally on $\mathcal{E}_1 \cap \mathcal{E}_2$, the probability of $\mathcal{E}_3$ is uniformly bounded below.

Finally, conditionally on $\bigcap_{1\le i\le 3} \mathcal{E}_i$, $\mathcal{E}_4$ has probability at least $1/2$ by Schramm's finiteness theorem (Lemma 4.18), which finishes the proof. □

The general idea for the full coupling around one point v will be that when the base coupling fails there is a not too small neighbourhood around v which was not intersected by any of the paths we sampled so far. Therefore, we will be able to retry the coupling in a new smaller neighbourhood. To implement this strategy, we first show that if the base coupling fails, v remains reasonably isolated at the point when we stop the process with high probability. We say a vertex u (possibly different from the vertex v around which we make the coupling) has *isolation radius* 6^{-k} at scale r at any step in the above base coupling (centered around v) if

$$B(u, 6^{-k} r) \qquad \text{does not contain any vertex from a sampled branch.}$$

We then set I_u to be the minimal such $k \ge 1$ at the time when the base coupling fails.

LEMMA 4.20. *Let* I_u *be as above and suppose that either* $u = v$ *or* $|u - v| \ge 10r$. *Then there exist constants* $\delta', c, c' > 0$ *such that for all* $\delta \in (0, \delta' r)$ *and for all* $i \in (0, \log_6(\delta' r/\delta) - 1)$,

$$\mathbb{P}(I_u \ge i \mid \textit{coupling fails}) \le c e^{-c' i}.$$

PROOF. From Proposition 4.11, if one of $\mathcal{E}_k$ fails for $k = 1, 2, 3$, the probability that the isolation radius is at least i is at most $C(1-\alpha)^i$ for some $C > 0$ and $\alpha > 0$ (since it is the maximum of a finite number of variables each with exponential tail). It remains to consider is the branches drawn while doing the good algorithm in item (iv) of the base coupling.

Notice that the number of vertices in $\mathcal{Q}_j$ for $j \geq j_0$ is at most $C_0 6^j$ and each of them is at a distance at least $r6^{-j-1}$ from u (note that this holds both when $u = v$ or $|u - v| \geq 10r$). Let $\mathcal{A}(i, j)$ be the event that coupling fails in step $j \geq j_0$ and I_u after this step is at least i. For $i \in (j^2, \log_6(\delta_0 r/\delta))$, the probability of $\mathcal{A}(i, j)$ is at most the probability that the branch $\gamma(w)$ from one of the vertices $w \in \mathcal{Q}_j$ comes within distance $6^{-i}r$ of u. By Proposition 4.11 and a union bound, this is bounded by $C_0 6^j (1-\alpha)^{i-j} \leq C_0 6^{\sqrt{i}} (1-\alpha)^{i-\sqrt{i}}$, which decays exponentially even when we sum over j such that $j \leq \sqrt{i}$.

Finally, if $i < j^2$, we bound the probability of $\mathcal{A}(i, j)$ by using the explicit error bound equation (4.3) which we obtained in the proof of Schramm's finiteness theorem: indeed, we showed that the probability one of the branches emanating from a vertex $w \in \mathcal{Q}_j$ leaves the ball of radius $j^2 6^{-j} r$ around w is less than $(1/2)^{j^2/C_0} = (1-\alpha)^{j^2}$. In particular, this is also a bound on the probability that one of these branches leaves $B(v, r)$. Hence, we conclude that $\mathbb{P}(\mathcal{A}(i, j)) \leq C_0 6^j (1-\alpha)^{j^2} \leq C(1-\alpha)'^{j^2}$ for some $\alpha' > 0$. Summing over $j \geq \sqrt{i}$, we get $\mathbb{P}(I_u \geq i) \leq C(1-\alpha')^i$.

It remains to condition on the event that the coupling fails. But since the coupling fails with probability bounded below by Lemma 4.19, the result follows. $\square$

Iteration of base coupling around a single point. We now describe how to iterate the base coupling at different scales which is the key step to construct the full coupling. We start with a domain $D \subset \mathbb{C}$ and suppose $v \in D^{\#\delta}$. Suppose $r < 1$ is small enough such that $B(v, 10r)$ is disjoint and contained in $D \cap \tilde{D}$. Fix a small constant c so that Lemma 4.19 holds and assume that $\delta \leq c\delta_0 r$.

(i) Perform a base coupling with $D^{\#\delta}$, $\tilde{D}^{\#\delta}$ and $B(v, r)$. If the coupling succeeds, we are done.

(ii) If the coupling fails, let $6^{-I_{v,1}}r$ be the isolation radius of v at scale r at the step the coupling has failed. If $I_{v,1} \geq \log_6(c\delta_0 r/\delta)$, we *abort* the whole process and we say that the full coupling failed.

(iii) If the base coupling has failed but we haven't aborted, we move to a smaller scale. Let $\mathcal{T}_1$ (resp. $\tilde{\mathcal{T}}_1$) be the portion of the uniform spanning tree in $D^{\#\delta}$ (resp. $\tilde{D}^{\#\delta}$) sampled up to this point. We perform the base coupling in the domains $D^{\#\delta} \setminus \mathcal{T}_1$, $\tilde{D}^{\#\delta} \setminus \tilde{\mathcal{T}}_1^{\#\delta}$ in the ball $B(v, 6^{-I_{v,1}}r)$ around v.

(iv) If the coupling fails, let $6^{-I_{v,1}-I_{v,2}}r$ be the isolation radius at scale r around v at the step the coupling has failed. Let $\mathcal{T}_2 \supset \mathcal{T}_1$ (resp. $\tilde{\mathcal{T}}_2 \supset \tilde{\mathcal{T}}_1$) be the uniform spanning tree of $D^{\#\delta}$ (resp. $\tilde{D}^{\#\delta}$) sampled up to this point. If $(I_{v,1} + I_{v,2}) \geq \log_6(c\delta_0 r/\delta)$, we abort the whole process. Otherwise we perform the base coupling with $D^{\#\delta} \setminus \mathcal{T}_2$, $\tilde{D}^{\#\delta} \setminus \tilde{\mathcal{T}}_2$ and $B(v_1, 6^{-(I_{v,1}+I_{v,2})}r)$.

(v) We continue in this way until we either abort or the base coupling succeeds at the Nth iteration. If we haven't aborted the process along the way, we have obtained a partial tree $\mathcal{T}_N$ which is coupled with a uniform spanning tree $\tilde{\mathcal{T}}_N$ in $\tilde{D}^{\#\delta}$ so that they are the same in $B(v, 6^{-(I_{v,1}+I_{v,2}+\cdots+I_{v,N-1})}r)$.

(vi) If we abort the process at step m, define $N = m$ by convention.

We call $I = \sum_{\ell=1}^{N-1} I_{v,\ell}$, so that $6^{-I}r$ is the isolation radius at scale r when we have succeeded in coupling the trees around v.

Full coupling. We now describe how to perform the full coupling around a fixed number of given points. For this, we introduce a new idea. We first sample all the branches from a cutset separating each of the vertices from the rest. Conditioned on these sampled branches, the neighbourhood of the vertices are now independent. We then show that the "unexplored" neighbourhoods around the points are still big with high probability and apply our one-point iterated base coupling for each such neighbourhood.

Let us start with some notation. Let $v_1, v_2, \dots, v_k$ be k distinct points and let r be chosen so that so that $B(v_i, 10r)$ are disjoint and are all contained in $D \cap \tilde{D}$. Let H_i be a set of vertices in $A(v_i, 9r/2, 5r)$ which disconnect v_i from ∂D and $\partial \tilde{D}$, and let $H = \bigcup_i H_i$. We simply reveal the branches emanating from H_i, $1 \le i \le k$ in some arbitrary order by Wilson's algorithm, resulting in a subgraph $\mathcal{T}_H^{\#\delta}$. We call this step a *cutset exploration*. Let J_{v_i} be the minimum k such that $B(v_i, 6^{-k}r) \cap \mathcal{T}_H^{\#\delta} = \varnothing$. Let $J = \max_i J_{v_i}$, and let $D_i^{\#\delta}$ be the remaining unexplored domain around v_i, that is, the connected component containing v_i in $D^{\#\delta} \setminus \mathcal{T}_H^{\#\delta}$. Say that we abort if $c6^{-J} r\delta_0 \le \delta$, where c is as in Lemma 4.19.

Conditionally on $\mathcal{T}_H^{\#\delta}$, on the event that we haven't aborted, the component of $\mathcal{T}$ containing v_i is distributed as a wired uniform spanning tree in $D_i^{\#\delta}$. We perform the iterated base coupling of this wired spanning tree with a uniform spanning tree $\mathcal{T}^{\#\delta}(i)$ of $\tilde{\mathcal{D}}^{\#\delta}$ with a base neighbourhood $B(v_i, 6^{-J} r)$. We also do this coupling around each point to obtain conditionally independent subtrees $(\mathcal{T}^{\#\delta}(i))_{1 \le i \le k}$ given the cutset exploration $\mathcal{T}_H^{\#\delta}$. We say that we abort the full coupling either if we aborted at the cutset exploration step or if we abort in any of the iterated base couplings. Let I_i be the isolation radius at scale r around v_i after performing the iterated base coupling around v_i.

THEOREM 4.21. *On the event $\mathcal{A}$ that we do not abort the full coupling, we obtain a coupling between $\mathcal{T}^{\#\delta}$ and independent copies of uniform spanning trees $\tilde{\mathcal{T}}^{\#\delta}(i)$ in $\tilde{D}^{\#\delta}$ for $1 \le j \le k$ such that*

$$\mathcal{T}^{\#\delta} \cap B(v_i, 6^{-I_i} r) = \tilde{\mathcal{T}}^{\#\delta}(i) \cap B(v_i, 6^{-I_i} r).$$

Furthermore, there exists a universal constant $c > 0$ and $C > 0$ such that for all $\delta \le c\delta_0 r$ and $1 \le i \le k$,

$$\mathbb{P}(I_i \ge n; \mathcal{A}) \le C e^{-cn}. \tag{4.5}$$

PROOF. Observe that I_i is a sum of the form $\sum_{\ell=1}^{N_i - 1} I_{v_i,\ell} + J_{v_i}$. By exactly the same proof as Lemma 4.20, J_{v_i} has an exponential tail so we concentrate on the sum. Observe that by Lemma 4.19, N_i has geometric tail (since the base coupling has probability uniformly bounded below to succeed at every step, independently of the past). Moreover, each $I_{v_j,\ell}$ has uniform exponential tail conditionally on all previous steps by Lemma 4.20. Hence, the situation is similar to Lemma D.5 in the Supplementary Material [4]: by Markov's inequality

$$\begin{aligned}\mathbb{P}\left(\sum_{\ell=1}^{N_i-1} I_{v_i,\ell} \ge n\right) &\le \mathbb{P}(N_i \ge \varepsilon n) + P\left(\sum_{\ell=1}^{\varepsilon n} I_{v_i,\ell} \ge n\right) \\ &\le e^{-c\varepsilon n} + \mathbb{E}\left(e^{c' \sum_{\ell=1}^{\varepsilon n} I_{v_i,\ell}}\right) e^{-c'n},\end{aligned}$$

where c' is as in Lemma 4.20. Now from that lemma we see that, even when we condition on $I_{v_i,1}, \dots, I_{v_i,\ell_1}$, $\mathbb{E}(e^{c' I_{v_i,\ell}} | I_{v_i,1}, \dots, I_{v_i,\ell-1}) \le C_1$ for some C_1. Hence the right-hand side above is less than

$$e^{-c\varepsilon n} + (C_1)^{\varepsilon n} e^{-c'n}$$

so by choosing ε sufficiently small this decays exponentially, as desired. $\square$

The following two consequences are immediate.

COROLLARY 4.22. *There exist constants $C, c > 0$ such that the probability of aborting the above process is at most $C(\delta/r\delta_0)^c$.*

PROOF. The event $\mathcal{A}^c$ occurs precisely when one of the variables I in step (vii) exceeds $\log_6(cr\delta_0/\delta)$. As above, this has exponential tail. $\square$

COROLLARY 4.23. *The wired uniform spanning tree measures in $D^{\#\delta}$ has a local limit when $\tilde{D} \to \mathbb{C}$ and this limit is independent of the exhaustion taken. We call this measure the* whole plane spanning tree. *In Theorem 4.21 and all the above statements, the spanning tree measure on $\tilde{D}^{\#\delta}$ can be replaced by a whole plane spanning tree.*

PROOF. For the first sentence, consider an exhaustion D_n that is, an increasing sequence such that $\bigcup_n D_n = \mathbb{C}$. Using Theorem 4.21 along with the control of the abortion probability from Corollary 4.22, we conclude that the spanning tree measures form a Cauchy sequence in total variation, therefore it converges. For the second one, it then follows immediately from the fact that all the statements are uniform on domains $\tilde{D}$. $\square$

REMARK 4.24. Suppose that $(\gamma, \tilde{\gamma})$ are branches emanating from z_0 in $\mathcal{T}$ and $\tilde{\mathcal{T}}$ coupled by a global coupling as above. For $i \in \mathbb{Z}$, let T_i be the smallest time t such that $\gamma(t) \in B(z_0, r6^{-i})$, and let $T = T_I$ (here we view γ as parametrised towards z_0). Then given $I = i$ the law of $(\gamma_t, t \geq T)$ is absolutely continuous with respect to the unconditional law of $(\gamma_t, t \geq T_i)$ with Radon–Nikodym derivative bounded above by C for some universal constant $C > 0$ as we vary δ (indeed, it is just the law of a loop-erased random walk conditioned on some event of uniformly positive probability, where these events are described in the base coupling).

Cutting coupled paths. At this point, we have proved that each branch γ of the UST has a portion where it can be coupled with a path $\tilde{\gamma}$ in the whole plane independent UST. Also the part far away from the starting point can be approximated by an SLE because of the convergence of loop erased random walk to SLE. It will therefore be natural to cut the branches into two parts and to use a different approximation for each piece.

A subtle issue arises here because of the choice of approximation. Indeed for the portion approximated by SLE we want to cut the path at a fixed capacity to be in the setup of Section 3, while for the discrete part we want to cut γ and $\tilde{\gamma}$ exactly at the same point so that their (diverging) windings cancel exactly. This is a problem because the capacity of a curve depends on the whole curve and therefore will never agree exactly between γ and $\tilde{\gamma}$. Our solution to this issue is to parametrise by capacity but cut at a randomised time in both γ and $\tilde{\gamma}$ in such a way that the corresponding positions match exactly. The key to doing this will be to observe that not only are the capacities of the two paths close to one another, but also their derivatives.

Recall the full coupling $(\gamma, \tilde{\gamma})$ in Theorem 4.21, where γ is a loop-erased random walk in $D^{\#\delta}$ and $\tilde{\gamma}$ is a loop-erased random walk in $\tilde{D}^{\#\delta}$ starting from a vertex v where $\tilde{D}$ is arbitrary (we can think of $\tilde{D}$ as the full plane). Suppose they are parametrised by capacity plus log of the conformal radius seen from v in their respective domains.

LEMMA 4.25. *There exist constants $C, c > 0$ such that the following holds. For any $t > 0$ there exists $\delta = \delta(t)$ such that for any $\delta \in (0, \delta(t))$, we can find a pair of random variables $(X, \tilde{X})$ such that individually, X and $\tilde{X}$ are each independent of $(\gamma, \tilde{\gamma})$ and*

$$\mathbb{P}[\gamma(t + X) = \tilde{\gamma}(t + \tilde{X})] \geq 1 - Ce^{-ct}.$$

Furthermore, both X and $\tilde{X}$ are random variables which are bounded (by $1/20$).

The proof of this lemma is quite technical so we only give a short description of the ideas; see the Supplementary Material [4] for details.

The idea behind the coupling is to see $\gamma(t+X)$ where X is an exponential variable as the first point in a Poisson point process on γ with an intensity given by a multiple of the derivative of the capacity over $[t, t+1/20]$ (or more formally a multiple of the push forward of the Lebesgue measure by $s \mapsto \gamma(s)$ over that interval), where the multiple is itself exponentially large in t. This derivative measure will be almost the same in γ and $\tilde{\gamma}$ with exponentially small error in $t - I$ (this is the technical part of the proof). Therefore the point processes can be coupled so that their first points are the same with high probability and this preserves some independence because of the independence inherent to a Poisson process.

5. Convergence of the winding of uniform spanning tree to GFF.

5.1. *Discrete winding, definitions and notation.* We define the discrete winding fields in a finite domain properly here. This is completely analogous to the definition in the continuum from Section 3. Let us fix a bounded domain $D \subset \mathbb{C}$ with a locally connected boundary and a marked point $x \in \partial D$. Using [37] Theorem 2.1, ∂D is a curve. Let $\delta > 0$ and let $\mathcal{T}^{\#\delta}$ be a wired spanning tree of $D^{\#\delta}$. Let $\gamma_v^{\#\delta}$ be the branch connecting v to the wired boundary in $\mathcal{T}^{\#\delta}$. As in the continuum definition, we add to each $\gamma_v^{\#\delta}$ a path following ∂D clockwise to x. More precisely, one endpoint of $\gamma_v^{\#\delta}$ is some auxiliary vertex (a point on the continuum boundary ∂D; see Section 4.2). We add to $\gamma_v^{\#\delta}$ the *continuum* curve joining this vertex and the marked boundary point x in the clockwise direction. For simplicity of notation, we still call this path $\gamma_v^{\#\delta}$.

We parametrise the part of the curve in ∂D by $[-1, 0]$ and the rest by capacity in D plus $\log R(z, D)$. Observe that this definition is analogous to the continuum definition in Section 3, so that $t = 0$ will always correspond to hitting the boundary. Motivated by the formula connecting intrinsic and topological winding (cf. Lemma 2.2), we define, in the smooth case first

$$h^{\#\delta}(v) := W_{\text{int}}(\gamma_v^{\#\delta}[-1, \infty)) \tag{5.1}$$

and

$$h_t^{\#\delta}(v) := W(\gamma_v^{\#\delta}[-1, t], z) - \operatorname{Arg}(-(\gamma_v^{\#\delta})'(-1)) + \arg_{D;x}(v). \tag{5.2}$$

If D is not smooth near x, then we can still define $h^{\#\delta}(v)$ via (5.1) up to a global constant,

$$h_t^{\#\delta}(v) := W(\gamma_v^{\#\delta}[-1, t], z) + \arg_{D;x}(v), \tag{5.3}$$

where $\arg_{D;x}(v)$ is defined up to a global constant.

Consider a full plane discrete UST on $G^{\#\delta}$, $\tilde{\mathcal{T}} = (\tilde{\gamma}_v)_{v \in G^{\#\delta}}$. (We write $\tilde{\gamma}_v$ instead of $\tilde{\gamma}_v^{\#\delta}$ for simplicity.) We parametrise the paths $\tilde{\gamma}_v$ by full plane capacity plus $\log R(v, D)$, going from $-\infty$ far away to $+\infty$ at v (we need to add $\log R(v, D)$ to match its parametrisation with the parametrisation of γ_v as much as possible). We extend the definition of the regularised winding to that setting by defining $\tilde{h}_T(v) - \tilde{h}_S(v) := W(\tilde{\gamma}_v(S, T), v)$ and $\tilde{h}(v) - \tilde{h}_T(v) := W(\tilde{\gamma}_v(T, \infty), v)$. Note that the definition of the increments (in T) of $\tilde{h}$ make sense even though we cannot define $\tilde{h}$ pointwise, so this definition is a slight abuse of notation. Finally,

$$m^{\#\delta}(v) := \mathbb{E}[\tilde{h}^{\#\delta}(v) - \tilde{h}^{\#\delta}_{\log R(v,D)}(v)] \tag{5.4}$$

and note that random variable

$$\tilde{h}^{\#\delta}(v) - \tilde{h}^{\#\delta}_{\log R(v,D)}(v) = W(\tilde{\gamma}_v(\log(R(v, D)), \infty); v)$$

is well defined (and its expectation is well defined too) and that it does not depend on D since $\tilde{\gamma}(\log R(v, D))$ is just the point of full plane capacity 0. We point out that as $\delta \to 0$, the path $\tilde{\gamma}_v$ converges (uniformly on compacts) towards a full plane SLE_2; this follows by a combination of results from Lawler, Schramm and Werner; Yadin and Yehudayoff [30, 43] as well as Field and Lawler [11]. As a consequence, the arbitrary choice of truncating at capacity 0 is irrelevant: this is because the asymptotics of $m^{\#\delta}(v)$ as $\delta \to 0$ is independent of the choice of truncation (indeed, for a full plane SLE path, the expected winding between capacity 0 and 1 is zero by symmetry). Readers who are uncomfortable with the notion of full plane SLE can replace the full plane by a disc of some large radius in the definition of $m^{\#\delta}$; in which case the above remark relies just on the convergence result of Lawler, Schramm and Werner [30] and Yadin and Yehudayoff [43].

To help with the intuition, recall from the Introduction that we need to remove by hand microscopic contributions to the expected winding. This is the purpose of $m^{\#\delta}$. Subtracting $m^{\#\delta}$ also allows us to take into account the possible contribution to winding coming from intermediate (mesoscopic) scales.

5.2. *Statement of the main result.* We now state the main theorem in this section, which is a stronger version of Theorem 1.2 in the Introduction. Since we are going to integrate the discrete winding field against test functions, we need to make precise what we mean by this. There are two natural choices to do this integral: one which takes into account the geometry of the underlying graph and another one which accounts only for the ambient Euclidean space in which the graph is embedded. The latter turns out to be slightly more natural since the limit in that case is a (conformally invariant) Gaussian free field, that is, does not depend on the limiting density of vertices.

We proceed as follows. Given $h^{\#\delta}$ defined on the vertices of the graph, we can extend $h^{\#\delta}$ to a function on the whole domain D using various forms of interpolation. One way to do this is to linearly extend the value of $h^{\#\delta}$ to the edges and then take a harmonic extension on the faces (this includes the outer face, and then we restrict this extension to D). However for concreteness, we will look at the following extension: consider the Voronoi tesselation of D defined by the vertices of the graph. We then define the extension $h^{\#\delta}_{\mathrm{ext}}$ to be constant on each Voronoi cell, equal to $h^{\#\delta}(v)$ where v is the centre of the cell. This allows us to use the regular L^2 product to integrate $h^{\#\delta}$ against test functions.

$$(h^{\#\delta}, f) := \int_D h^{\#\delta}_{\mathrm{ext}}(z) f(z)\, dz.$$

This extension procedure can also be applied to $m^{\#\delta}$, leading to a function defined on all of D. We then have the following theorem.

THEOREM 5.1. *Let G be a graph satisfying the assumptions of Section 4.1, let $D \subset \mathbb{C}$ be a simply connected domain with a locally connected boundary and a marked point $x \in \partial D$. Let $\mathcal{T}$ be a uniform spanning tree of $D^{\#\delta}$ and let $h^{\#\delta}(v)$ denote the intrinsic winding as above, and let $m^{\#\delta}$ be defined as above. Then*

$$h^{\#\delta} - m^{\#\delta} \xrightarrow[\delta \to 0]{} h_{\mathrm{GFF}}.$$

The convergence is in law in the Sobolev space $H^{-1-\eta}(D)$ for any $\eta > 0$. The limit h_{GFF} is a free field with intrinsic winding boundary conditions, that is, $h_{\mathrm{GFF}} = (1/\chi)h + \pi/2$ where $h = h^0_{\mathrm{GFF}} + \chi u_{(D,x)}$. Here h^0_{GFF} is a standard GFF with Dirichlet boundary conditions and $u_{(D,x)}$ is defined as in equation (2.7) and Remark 2.7.

Moreover, for all $n \geq 1$, for all $f_1, \ldots, f_n \in H^{1+\eta}$, and for all positive reals $k_1, \ldots, k_n$ we have

$$\mathbb{E}\prod_i (h^{\#\delta} - m^{\#\delta}, f_i)^{k_i} \xrightarrow[\delta \to 0]{} \mathbb{E}\prod_i (h_{\mathrm{GFF}}, f_i)^{k_i}$$

and for all $k \geq 1$

$$\mathbb{E}(\|h^{\#\delta} - m^{\#\delta}\|^k_{H^{-1-\eta}}) \xrightarrow[\delta \to 0]{} \mathbb{E}(\|h_{\mathrm{GFF}}\|^k_{H^{-1-\eta}}).$$

REMARK 5.2. We emphasise that the function $m^{\#\delta}$ is a deterministic function which depends only on the point in the graph, and in particular does not depend on the domain D. Note also that it follows from this result that $\mathbb{E}(h^{\#\delta} - m^{\#\delta}) \to \mathbb{E}(h_{\mathrm{GFF}})$ and hence we deduce

$$h^{\#\delta} - \mathbb{E}h^{\#\delta} \to \frac{1}{\chi} h^0_{\mathrm{GFF}}$$

in the same sense as above, where h^0_{GFF} is a Gaussian free field with Dirichlet boundary conditions.

5.3. *Other notions of integration.* We now comment briefly on other possible definitions of integration against test functions. Another definition which is a priori natural is to consider

$$(h^{\#\delta} - m^{\#\delta}, f)_{\#\delta} := \frac{1}{\mu^{\#\delta}(D)} \sum_v (h^{\#\delta}(v) - m^{\#\delta}) f(v).$$

In that case, $h^{\#\delta} - m^{\#\delta}$ is viewed as a random measure which is a sum of weighted Dirac masses. We can first ask about convergence of this object as a stochastic process indexed by test functions (see, e.g., [2], Definition 1.10). It can be checked that if the uniform distribution on the vertices of the graph converges to a measure μ in $\mathbb{C}$, we have that

$$(h^{\#\delta} - m^{\#\delta}, f)_{\#\delta} \xrightarrow[\delta \to 0]{} h^{\mu}_{\mathrm{GFF}},$$

where now h^{μ}_{GFF} is a Gaussian stochastic process indexed by test functions such that $(h^{\mu}_{\mathrm{GFF}}, \phi) = (h_{\mathrm{GFF}}, \phi \frac{d\mu}{d\,\mathrm{Leb}})$ where Leb denotes the Lebesgue measure.

Note that in most cases, for example, in any periodic lattice or isoradial graphs or T-graphs where our results apply, μ is just the Lebesgue measure. In that case, note that $h^{\#\delta}$ lies a priori within $H^{-1-\varepsilon}$ for any $\varepsilon > 0$ (this is the Sobolev regularity of any Dirac mass) and it is easy to check that our proof implies convergence of $h^{\#\delta} - m^{\#\delta}$ towards h_{GFF} in the stronger sense of Sobolev spaces $H^{-1-\eta}(D)$ for any η, as in Theorem 5.1.

However there are also interesting examples of graphs where the convergence of random walk to a (time-changed) Brownian motion holds but μ is different from Lebesgue measure. An exotic example of such a situation is a conformally embedded random planar map where such a convergence is expected to hold and the measure μ is a variant of Gaussian multiplicative chaos (see [12] and [1, 13] for an introduction to this topic).

5.4. *Proof of the main result.* Now we collect the results of the two previous sections to prove Theorem 5.1.

LEMMA 5.3. *Fix a domain $D \subset \mathbb{C}$ with locally connected boundary and let $x_1, \ldots, x_k \in D$. For all $v_1^{\delta}, \ldots, v_k^{\delta} \in D^{\#\delta}$ converging to $x_1, \ldots, x_k$, for all $T_1, \ldots, T_k$,*

$$\mathbb{E}\Big[\prod_i h^{\#\delta}_{T_i}(v_i)\Big] \xrightarrow[\delta \to 0]{} \mathbb{E}\Big(\prod_i h_{T_i}(x_i)\Big),$$

where h_T is the regularised winding field of a continuum UST in D as defined in equation (2.7) and Remark 2.7.

PROOF. By our assumptions (see Remark 4.1) and by Wilson's algorithm, the paths $(\gamma^{\#\delta}_{v_i})_{i=1}^k$ converge to $(\gamma_{x_i})_{i=1}^k$ where the γ_{x_i} are the paths connecting x_i to ∂D in a continuous UST. Furthermore, observe that the function $h_T(v)$ is a continuous function of $\gamma^{\#\delta}_v$ (this is because the topological winding up to capacity $t + \log(R(v, D))$ is continuous in the curve). Hence, $\prod h^{\#\delta}_T(v_i)$ converges in distribution to $\prod h_T(x_i)$. Using Lemma 4.16, we see that it is also uniformly integrable and hence the expectation converges. □

Combining the above lemma with Theorem 3.1 and Lemma 3.7, we can find a sequence $T(\delta)$ (depending on the v_i's) going to infinity slowly enough such that whenever $T(\delta) \le T_i(\delta) \le T(\delta) + 1/20$,

$$\mathbb{E}\Big[\prod h^{\#\delta}_{T_i(\delta)}(v_i)\Big] \xrightarrow[\delta\to 0]{} \mathbb{E}\Big(\prod h_{\mathrm{GFF}}(x_i)\Big).$$

Using the previous lemma with the full coupling of Theorem 4.21, we can control the k-point function for the winding up to the endpoint, which is the key step in the proof of Theorem 5.1.

PROPOSITION 5.4. *For all k, for all bounded domains D with locally connected boundary, for all $v_1^\delta, \dots, v_k^\delta \in D$ converging to $x_1, \dots, x_k$,*

$$\mathbb{E}\Big[\prod_i \big(h^{\#\delta}(v_i^\delta) - m^{\#\delta}(v_i^\delta)\big)\Big] \xrightarrow[\delta\to 0]{} \mathbb{E}\Big(\prod_i h_{\mathrm{GFF}}(x_i)\Big).$$

Recall that the right hand side is a notation for the k-point function of a GFF with winding boundary condition, as in Theorem 5.1.

PROOF. By definition of our extension of $h^{\#\delta}$ to D, we may assume without loss of generality that $v_i^\delta \in D^{\#\delta}$ (this is one of the advantages of working with the Voronoi extension of $h^{\#\delta}$ to D). We write γ_i for $\gamma^{\#\delta}_{v_i}$. If δ is small enough, one can apply the coupling of Theorem 4.21. Focusing only on the paths from the v_i, we obtain random variables $I_1, \dots, I_k$, all with exponential tails, and independent full plane loop-erased paths $\tilde\gamma_1, \dots, \tilde\gamma_k$ such that

$$\forall j, \qquad \gamma_j \cap B\big(v_j, 6^{-I_j} r\big) = \tilde\gamma_j \cap B\big(v_j, 6^{-I_j} r\big)$$

on the event that the coupling succeeds, where $r = (1/10)(\min_{i\neq j} |v_i - v_j| \wedge \min_i d(v_i, \partial D) \wedge 1)$. Here r is a constant and $\delta \to 0$ so we will not worry about r or the offset in the parametrisation (which recall is capacity plus $\log R(v_i, D)$). Let $\tilde h^{\#\delta}$ be the associated field defined as in equation (5.4) (recall that only its increments in T are defined). Let $T(\delta)$ be some sequence such that for any T_j such that $T(\delta) \le T_j \le T(\delta) + 1/20$ we have

$$\mathbb{E}\Big[\prod h^{\#\delta}_{T_j}(v_j)\Big] \xrightarrow[\delta\to 0]{} \mathbb{E}\prod h_{\mathrm{GFF}}(x_j), \tag{5.5}$$

$$\mathbb{E}\big[\tilde h^{\#\delta}_{T(\delta)}(v_j) - \tilde h^{\#\delta}_{\log(R(v,D))}(v_j)\big] \xrightarrow[\delta\to 0]{} 0. \tag{5.6}$$

Recall that Lemma 4.25 holds for $t = T(\delta)$. We can further modify our choice of $T(\delta)$ if needed so that

$$\mathbb{P}\big(\gamma_i(T(\delta), \infty) \not\subset B\big(v_i, e^{-T(\delta)/2}\big)\big) \le c e^{-cT(\delta)}$$

and the same statement holds with $\tilde\gamma_i$ instead of γ_i. This is possible because $T(\delta)$ can be chosen to grow to infinity arbitrarily slowly and these estimates hold in the continuum. Indeed, γ_j converges to radial SLE_2 for which we can apply (2.8) and $\tilde\gamma_j$ converges to whole plane SLE_2 which is symmetric with respect to conjugation hence has zero expected winding.

For each $1 \le j \le k$ we use Lemma 4.25 and we write $T_j = T(\delta) + X_j$ and $\tilde T_j = T(\delta) + \tilde X_j$ the resulting times. Let $\mathcal{C}$ denote the σ-algebra generated by the cutset exploration in

Theorem 4.21 as well as $\tilde{T}_j$, $1 \le j \le k$. Let G be the good event that $\tilde{\gamma}_i(\tilde{T}_i, \infty) = \gamma_i(T_i, \infty) \subset B(v_i, re^{-I_i})$ for all $1 \le i \le k$. In other words, G is the event that $\gamma_i(T_i) = \tilde{\gamma}_i(\tilde{T}_i)$ and $T_i \ge \Lambda_i$ where Λ_i is the last time at which γ_i enters $B(v_i, re^{-I_i})$. Unfortunately G is not measurable with respect to $\mathcal{C}$, but fortunately we will see that its complement has exponential small probability in $T(\delta)$.

For clarity, we remove the superscripts $\cdot^{\#\delta}$ from notation, hence, for instance, $h(v_i)$ means $h^{\#\delta}(v_i)$. Note that on the good event G, since $T_i \ge \Lambda_i$ and $\tilde{T}_i \ge \tilde{\Lambda}_i$ (where $\tilde{\Lambda}_i$ is defined in the obvious analogue way to Λ_i), and since γ and $\tilde{\gamma}$ also agree on $B(v_i, re^{-I_i})$, we can write

$$h(v_i) = h_{T_i}(v_i) + \Delta\tilde{h}(v_i), \tag{5.7}$$

where $\Delta\tilde{h}(v_i) = \tilde{h}(v_i) - \tilde{h}_{\tilde{T}_i}(v_i)$. The first term is the "main" term (well described by SLE); see (5.5). The second term is a mesoscopic term: when we subtract $m(v_i)$ we will get independent terms with mean approximately zero.

We wish to take the conditional expectation given $\mathcal{C}$, but since G is not measurable with respect to $\mathcal{C}$, some care is needed. We introduce the following notation: if $s \le t$ we write $h(s,t)$ for the winding around v_i of γ_i during $[s,t]$; that is, $h(s,t) = W(\gamma_i([s,t]), v_i)$. By an abuse of notation, we write $h(v_i) = h(-1,\infty)$. When $s \ge t$ we put $h(s,t) = -h(t,s)$. We then write

$$h(-1,\infty) = h(-1,T_i) + h(T_i,\Lambda_i) + h(\Lambda_i,\infty).$$

As before, the main term is $h(-1,T_i) = h_{T_i}(v_i)$ which is thus governed by (5.5). Note that we also have $h(\Lambda_i,\infty) = \tilde{h}(\tilde{\Lambda}_i,\infty)$ (this is also trivially true even when the coupling fails).

Moreover, the middle term $h(T_i,\Lambda_i)$ may be rewritten by adding and taking away $\tilde{h}(\tilde{T}_i,\tilde{\Lambda}_i)$ as

$$h(T_i,\Lambda_i) = \tilde{h}(\tilde{T}_i,\tilde{\Lambda}_i) + \xi_i,$$

where

$$\xi_i = h(T_i,\Lambda_i) - \tilde{h}(\tilde{T}_i,\tilde{\Lambda}_i)$$

is an error that is typically zero except on an event of very small probability.

Consequently, we obtain the decomposition

$$h(v_i) - m_i = \underbrace{h_{T_i}(v_i)}_{\text{type 1}} + \underbrace{[\tilde{h}(\tilde{T}_i,\infty) - m_i]}_{\text{type 2}} + \underbrace{\xi_i}_{\text{type 3}}.$$

We now explain how we will finish the proof. We need to compute the expected value of the product over i of the expression in the left hand side above. We expand the product in the right hand side above to get a finite sum of products of terms involving one of the three types of terms above for each $1 \le i \le k$. We take the conditional expectation given $\mathcal{C}$ and then the total expectation. If only the first type of terms occur in the sum, we can simply use (5.5) as already mentioned. To finish the proof, we simply make the following observations:

- Terms of type two, $B_i = [\tilde{h}(\tilde{T}_i,\infty) - m_i]$, satisfy $\mathbb{E}(B_i) = o(1)$ and B_i is independent of $\mathcal{C}$ and of any of the terms in the product given $\mathcal{C}$.
- Since $\xi_i = 0$ except if G does not hold, we have $\mathbb{P}(\xi_i \ne 0) \le Ce^{-cT}$. Indeed,

$$\begin{aligned}\mathbb{P}(G^c) &\le \mathbb{P}\big(\gamma(T_i) \ne \tilde{\gamma}_i(\tilde{T}_i)\big) + \mathbb{P}\big(I_i \ge T(\delta)/2 - \log r + \log R(v_i,D)\big)\\ &\quad + \mathbb{P}\big(I_i \le T(\delta)/2 - \log r + \log R(v_i,D) \text{ and}\\ &\quad \gamma_i(T(\delta),\infty) \not\subset B(v_i, e^{-T(\delta)/2})\big)\\ &\le Ce^{-cT(\delta)},\end{aligned}$$

where we have used Lemma 4.25 for the first term, Theorem 4.21 for the second (and the fact that r is fixed), and (2.8) and the choice of $T(\delta)$ for the third. Hence, using Lemma 4.16,

$$\mathbb{E}(\xi_i^k) \le ce^{-cT}\mathbb{E}(T_i^k + \tilde{T}_i^k) \le Ce^{-cT}.$$

- Moreover, $\mathbb{E}(h_{T_i}^k) \le C(1+T^k)$ by Lemma 4.16.

When we take the conditional expectation, all the terms of the type 2 contribute $\mathbb{E}(B_i|\mathcal{C}) = \mathbb{E}(B_i) = o(1)$ to the product since they are independent of any other term in the product. Hence if the product contains only terms of types 1 and 2, then this contributes $o(1)\mathbb{E}(\prod h_{T_i}(v_i)) = o(1)$ by (5.5).

Otherwise, if the product contains any term involving ξ_i (i.e., type 3), using Hölder's inequality and (5.5), this contributes at most $O(e^{-cT}) = o(1)$. Hence the result is proved. $\square$

The above proposition gives a pointwise convergence of the k-point function. We now need some a priori bounds to allow us to integrate these moments against test functions via the dominated convergence theorem.

LEMMA 5.5. *For all $k \ge 2$, for all bounded domains D with locally connected boundary, there exist constants $C = C_k$, $c > 0$ such that for all $\delta < c\delta_0$, for all $v_1^{\#\delta}, \dots, v_k^{\#\delta} \in D^{\#\delta}$,*

$$\left|\mathbb{E}\left[\prod\left(h^{\#\delta}(v_i^{\#\delta}) - m^{\#\delta}(v_i^{\#\delta})\right)\right]\right| \le C(1 + \log^{2k} r),$$

where $r = (1/10)(\min_{i\neq j}|v_i^{\#\delta} - v_j^{\#\delta}| \wedge \min_j d(v_j^{\#\delta}, \partial D) \wedge 1)$. The same holds even if we replace $v_i^{\#\delta}$ by any point in its Voronoi cell as in our extended definition of $h_{\text{ext}}^{\#\delta}(z)$.

PROOF. Let us assume $r \ge \delta$ for now and consider the full coupling of Theorem 4.21. We use the notation from the proof of Proposition 5.4. We exploit the following decomposition which is analogous to the decomposition in Proposition 5.4 except we do away with T_i (indeed, if the points v_i are really close to each other, e^{-T_i} might be much greater that this distance):

$$\begin{aligned} h(v_i) - m_i &= h(-1, \Lambda_i) + \tilde{h}(\tilde{\Lambda}_i, \log R(v, D)) \\ &\quad + \left[\tilde{h}(\log R(v, D), \infty) - m_i\right]. \end{aligned} \tag{5.8}$$

Let $\mathcal{C}$ be the sigma algebra generated by the cutset exploration. Note that conditionally on $\mathcal{C}$, the third term is independent of any of the above terms involving $j \neq i$ and has 0 mean. Thus, we can ignore the terms in the expansion of the product which has at least one term of the third type.

We now provide an argument on how to bound the first term and the same argument can be used to bound the second term by the same quantity. Let Φ_i be the time of first entry of γ_i into $B(v_i, re^{-I_i})$ (in contrast with Λ_i which is the last entry into $B(v_i, re^{-I_i})$). Note that

$$h(-1, \Lambda_i) = h(-1, \Phi_i) + h(\Phi_i, \Lambda_i)$$

and the second term is deterministically bounded above in absolute value by 2π for elementary topological reasons (essentially, the winding number of a Jordan curve is either 0 or $\pm 2\pi$). Furthermore, note that

$$\left|h(-1, \Phi_i)\right| \le \sum_{j=-1}^{\lfloor \Phi_i \rfloor - 1} \left|h(j, j+1)\right| + \left|h\left(\lfloor \Phi_i \rfloor, \Phi_i\right)\right|.$$

Now each of these terms have exponential tail by Proposition 4.12. Moreover, $\Phi_i - \log R(v, D) \le -\log r + I_i$ by monotonicity of conformal radius, and I_i has exponential tails by equation (4.5) in Theorem 4.21, hence by convexity of the function $x \mapsto x^k$,

$$\mathbb{E}(|h(-1, \Phi_i)|^k) \le \mathbb{E}\Bigg((-\log r + I_i + \log R(v_i, D))^{k-1} \times \sum_{j=-1}^{-\log r + I_i + \log R(v_i, D)} |h(j, j+1)|^k\Bigg) + C$$

$$\le C(-\log r + 1)^{2k} + C$$

by Cauchy–Schwarz, as desired (the above bound is not optimal, but this is unimportant).

Finally, if $r < \delta$, we can use Hölder to bound the moment of $h^{\#\delta} - m^{\#\delta}$ by $C(1 + \log^{2k} \delta)$ as above which is at most the required bound. □

We can now prove our main theorem.

PROOF OF THEOREM 5.1. Fix $f_1, \dots, f_n$ to be smooth functions in $\bar{D}$ and $k_1, \dots, k_n \ge 1$. Combining Proposition 5.4 and Lemma 5.5, we see that we can apply the dominated convergence theorem to $\mathbb{E}[\prod_{i=1}^n (h^{\#\delta} - m^{\#\delta}, f_i)^{k_i}]$ and therefore

$$\mathbb{E}\Bigg[\prod_{i=1}^n (h^{\#\delta} - m^{\#\delta}, f_i)^{k_i}\Bigg] \to \mathbb{E}\Bigg[\prod_{i=1}^n (h_{\text{GFF}}, f_i)^{k_i}\Bigg].$$

Since the right-hand side is Gaussian (and therefore moments characterise the distribution), $(h^{\#\delta} - m^{\#\delta}, f_i)_{i=1}^n$ converges in distribution to $(h_{\text{GFF}}, f_i)_{i=1}^n$. In other words, at this point we already know $h^{\#\delta} - m^{\#\delta}$ converges to h_{GFF} in the sense of finite dimensional marginals (when viewed as a stochastic process indexed by smooth functions with compact support, say). We now check tightness in the Sobolev space $H^{-1-\eta}$, from which convergence in $H^{-1-\eta}$ will follow.

Note that by the Rellich–Kondrachov embedding theorem, to get tightness in $H^{-1-\eta}$ it suffices to prove that $\mathbb{E}(\|h^{\#\delta} - m^{\#\delta}\|^2_{H^{-1-\eta'}}) < C$ for some constant C, for any $\eta' < \eta$. More generally we will check that $\mathbb{E}(\|h^{\#\delta} - m^{\#\delta}\|^{2k}_{H^{-1-\eta}}) < C_k$ for any $k \ge 1$ and any $\eta > 0$.

Let (e_j) be an orthonormal basis of eigenfunctions of $-\Delta$ in $L^2(D)$, corresponding to eigenvalues $\lambda_j > 0$. Then writing $h = h^{\#\delta} - m^{\#\delta}$ for convenience,

$$\mathbb{E}(\|h\|^{2k}_{H^{-1-\eta}}) = \mathbb{E}\Bigg(\sum_{j=1}^\infty (h, e_j)^2_{L^2} \lambda_j^{-1-\eta}\Bigg)^k \le C \sum_{j=1}^\infty \mathbb{E}((h, e_j)^{2k}_{L^2}) \lambda_j^{-1-\eta}$$

by Fubini's theorem and Jensen's inequality (since by Weyl's law, $\sum_j \lambda_j^{-1-\eta} < \infty$ is summable).

Furthermore,

$$\mathbb{E}((h, e_j)^{2k}_{L^2}) = \int_{D^{2k}} \mathbb{E}(h(z_1) \cdots h(z_{2k})) e_j(z_1) \cdots e_j(z_{2k})\, dz_1 \cdots dz_{2k}$$

and note that by Lemma 5.5, $\mathbb{E}(h(z_1) \cdots h(z_{2k})) \le C(1 + \log^{4k} r)$ where $r = r(z_1, \dots, z_{2k})$ is as in that lemma. Note also that $(\log r)^a$ is integrable for any $a > 0$ and D is bounded hence using Cauchy–Schwarz, we deduce that $\mathbb{E}((h, e_j)^{2k}_{L^2}) \le C(\int_{D^{2k}} e_j(z_1)^2 \cdots e_j(z_{2k})^2\, dz_1 \cdots dz_{2k})^{1/2} = C$ since e_j is orthonormal in L^2. Consequently,

$$\mathbb{E}(\|h\|^{2k}_{H^{-1-\eta}}) \le \sum_{j=1}^\infty C \lambda_j^{-1-\eta} < \infty$$

by Weyl's law. This finishes the proof of Theorem 5.1 and hence Theorem 1.2. Let us remind the reader here that the proofs of moment bounds in $H^{-1-\eta}$ follows through in exactly the same way in the continuum proof in Proposition 3.16. □

6. Joint convergence. In this section, we prove the joint convergence of (dimer height function, wired UST) to (GFF, continuum wired UST) where the latter is coupled together through the imaginary geometry coupling in Theorem 2.8.

Let us first introduce the setup. The topology on the height function is the Sobolev space $H^{-1-\eta}$ (recall this is a complete, separable Hilbert space). The topology on the tree is the Schramm topology Ω_1 introduced in [39] (and described in the Supplementary Material [4]). As usual we have a bounded domain (D, x) with a marked point $x \in \partial D$ and locally connected boundary. We work with the space $\Omega := \Omega_1 \times H^{-1-\eta}(D)$ equipped with the product topology. We also view Ω also as a metric space with metric defined by $d_1 + d_2$ where d_1 and d_2 are the metrics in each coordinate. Let $\mathcal{T}^d$, $h^{\#\delta}$ be as in Section 5.2. Let $(\mathcal{T}, h_{\mathrm{GFF}}(\mathcal{T}))$ denote the continuum wired UST in D with $h_{\mathrm{GFF}}(\mathcal{T})$ denoting the GFF which is coupled with $\mathcal{T}$ using the imaginary geometry coupling (cf. Theorem 2.8). The point here is again that the height function is not continuous as a function of the discrete tree, hence we have to use the results about the truncated winding we proved in this paper.

THEOREM 6.1. *In the above setup, we have the following joint convergence in law in the product topology described above*:

$$(\mathcal{T}^{\#\delta}, h^{\#\delta} - m^{\#\delta}) \xrightarrow[\delta \to 0]{(d)} (\mathcal{T}, h_{\mathrm{GFF}}(\mathcal{T})).$$

PROOF. To simplify notation, we write $h^{\#\delta}$ for $h^{\#\delta} - m^{\#\delta}$ admitting a slight abuse of notation. Let h_t denote the continuum winding truncated at capacity t plus log conformal radius seen from the point as before. Notice that from Theorem 3.13, $(\mathcal{T}, h_t) \xrightarrow[t \to \infty]{P} (\mathcal{T}, h_{\mathrm{GFF}}(\mathcal{T}))$, where the convergence is in probability in the metric space defined above. Note that for any fixed t, $h_t^{\#\delta}$ and h_t are obtained by applying the same continuous function to respectively $\mathcal{T}^{\#\delta}$ and $\mathcal{T}$. Hence, we have

$$(\mathcal{T}^{\#\delta}, h_t^{\#\delta}) \xrightarrow[\delta \to 0]{(d)} (\mathcal{T}, h_t).$$

Thus there exists a sequence $t(\delta)$ growing slow enough such that

$$(\mathcal{T}^{\#\delta}, h_{t(\delta)}^{\#\delta}) \xrightarrow[\delta \to 0]{(d)} (\mathcal{T}, h_{\mathrm{GFF}}(\mathcal{T})).$$

Now recall that we proved in Proposition 5.4 and Lemma 5.5 that $\mathbb{E}(\|h_{t(\delta)}^{\#\delta} - h^{\#\delta}\|_{H^{-1-\eta}}) \to 0$ which implies that $h_{t(\delta)}^{\#\delta} - h^{\#\delta}$ converges to 0 in probability. Hence, the result follows from Slutsky's lemma (e.g., Theorem 3.1 in [5]). □

Acknowledgements. We are grateful to a number of people for useful and stimulating discussions, including: Dima Chelkak, Julien Dubédat, Laure Dumaz, Alexander Glazman, Rick Kenyon, Jason Miller, Marianna Russikh, Xin Sun, Vincent Tassion and Fabio Toninelli. Special thanks to Vincent Beffara for useful discussions on the winding of SLE towards the beginning of this project. This work started while visiting the *Random Geometry* programme at the Isaac Newton Institute. We wish to express our gratitude for the hospitality and the stimulating atmosphere.

The first author was supported in part by EPSRC grants EP/L018896/1 and EP/I03372X/1. On leave from the University of Cambridge.

The second author was supported in part by EPSRC grant EP/I03372X/1.

The third author was supported in part by EPSRC grant EP/I03372X/1, NSERC 50311-57400 and University of Victoria start-up 10000-27458.

SUPPLEMENTARY MATERIAL

Supplementary materials to dimers and Imaginary geometry (DOI: 10.1214/18-AOP1326SUPP; .pdf). In the supplementary material we provide all the details of the intermediary results which have been deferred.

REFERENCES

[1] BERESTYCKI, N. (2015). Diffusion in planar Liouville quantum gravity. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** 947–964. MR3365969 https://doi.org/10.1214/14-AIHP605

[2] BERESTYCKI, N. (2015). Introduction to the Gaussian free field and Liouville quantum gravity. Lecture notes, available on author's webpage.

[3] BERESTYCKI, N., LASLIER, B. and RAY, G. (2016). A note on dimers and T-graphs. Available at https://arxiv.org/abs/1610.07994.

[4] BERESTYCKI, N., LASLIER, B. and RAY, G. (2020). Supplement to "Dimers and imaginary geometry." https://doi.org/10.1214/18-AOP1326SUPP.

[5] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York. MR0233396

[6] BUFETOV, A. and GORIN, V. (2018). Fluctuations of particle systems determined by Schur generating functions. *Adv. Math.* **338** 702–781. MR3861715 https://doi.org/10.1016/j.aim.2018.07.009

[7] CHELKAK, D. and SMIRNOV, S. (2011). Discrete complex analysis on isoradial graphs. *Adv. Math.* **228** 1590–1630. MR2824564 https://doi.org/10.1016/j.aim.2011.06.025

[8] COHN, H., KENYON, R. and PROPP, J. (2001). A variational principle for domino tilings. *J. Amer. Math. Soc.* **14** 297–346. MR1815214 https://doi.org/10.1090/S0894-0347-00-00355-6

[9] DE TILIÈRE, B. (2007). Scaling limit of isoradial dimer models and the case of triangular quadri-tilings. *Ann. Inst. Henri Poincaré Probab. Stat.* **43** 729–750. MR3252428

[10] DUBÉDAT, J. (2009). SLE and the free field: Partition functions and couplings. *J. Amer. Math. Soc.* **22** 995–1054. MR2525778 https://doi.org/10.1090/S0894-0347-09-00636-5

[11] FIELD, L. S. and LAWLER, G. F. (2013). Reversed radial SLE and the Brownian loop measure. *J. Stat. Phys.* **150** 1030–1062. MR3038676 https://doi.org/10.1007/s10955-013-0729-5

[12] GARBAN, C. (2013). Quantum gravity and the KPZ formula [after Duplantier–Sheffield]. *Astérisque* **352** 315–354. MR3087350

[13] GARBAN, C., RHODES, R. and VARGAS, V. (2016). Liouville Brownian motion. *Ann. Probab.* **44** 3076–3110. MR3531686 https://doi.org/10.1214/15-AOP1042

[14] GIULIANI, A., MASTROPIETRO, V. and TONINELLI, F. L. (2017). Haldane relation for interacting dimers. *J. Stat. Mech. Theory Exp.* **3** 034002. MR3637384 https://doi.org/10.1088/1742-5468/aa5d1f

[15] GIULIANI, A., MASTROPIETRO, V. and TONINELLI, F. L. (2017). Height fluctuations in interacting dimers. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 98–168. MR3606736 https://doi.org/10.1214/15-AIHP710

[16] JOCKUSCH, W., PROPP, J. and SHOR, P. (1998). Random domino tilings and the arctic circle theorem. arXiv preprint. Available at arXiv:math/9801068.

[17] KASTELEYN, P. W. (1961). The statistics of dimers on a lattice I. The number of dimer arrangements on a quadratic lattice. *Physica* **27** 1209–1225.

[18] KENYON, R. (2000). Conformal invariance of domino tiling. *Ann. Probab.* **28** 759–795. MR1782431 https://doi.org/10.1214/aop/1019160260

[19] KENYON, R. (2001). Dominos and the Gaussian free field. *Ann. Probab.* **29** 1128–1137. MR1872739 https://doi.org/10.1214/aop/1015345599

[20] KENYON, R. (2002). The Laplacian and Dirac operators on critical planar graphs. *Invent. Math.* **150** 409–439. MR1933589 https://doi.org/10.1007/s00222-002-0249-4

[21] KENYON, R. (2008). Height fluctuations in the honeycomb dimer model. *Comm. Math. Phys.* **281** 675–709. MR2415464 https://doi.org/10.1007/s00220-008-0511-8

[22] KENYON, R. (2009). Lectures on dimers. In *Statistical Mechanics*. *IAS/Park City Math. Ser.* **16** 191–230. Amer. Math. Soc., Providence, RI. MR2523460

[23] KENYON, R., MILLER, J., SHEFFIELD, S. and WILSON, D. B. (2017). Six-vertex model and Schramm–Loewner evolution. *Phys. Rev. E* **95** 052146.

[24] KENYON, R. and OKOUNKOV, A. (2006). Planar dimers and Harnack curves. *Duke Math. J.* **131** 499–524. MR2219249 https://doi.org/10.1215/S0012-7094-06-13134-4

[25] KENYON, R., OKOUNKOV, A. and SHEFFIELD, S. (2006). Dimers and amoebae. *Ann. of Math.* (2) **163** 1019–1056. MR2215138 https://doi.org/10.4007/annals.2006.163.1019

[26] KENYON, R. W., PROPP, J. G. and WILSON, D. B. (2000). Trees and matchings. *Electron. J. Combin.* **7** 25. MR1756162

[27] KENYON, R. W. and SHEFFIELD, S. (2004). Dimers, tilings and trees. *J. Combin. Theory Ser. B* **92** 295–317. MR2099145 https://doi.org/10.1016/j.jctb.2004.07.001

[28] LASLIER, B. (2013). Central limit theorem for T-graphs. arXiv preprint. Available at arXiv:1312.3177.

[29] LAWLER, G. F. (2005). *Conformally Invariant Processes in the Plane*. *Mathematical Surveys and Monographs* **114**. Amer. Math. Soc., Providence, RI. MR2129588

[30] LAWLER, G. F., SCHRAMM, O. and WERNER, W. (2004). Conformal invariance of planar loop-erased random walks and uniform spanning trees. *Ann. Probab.* **32** 939–995. MR2044671 https://doi.org/10.1214/aop/1079021469

[31] LI, Z. (2017). Conformal invariance of dimer heights on isoradial double graphs. *Ann. Inst. Henri Poincaré D* **4** 273–307. MR3713018 https://doi.org/10.4171/AIHPD/41

[32] LYONS, R. and PERES, Y. (2016). *Probability on Trees and Networks*. *Cambridge Series in Statistical and Probabilistic Mathematics* **42**. Cambridge Univ. Press, New York. MR3616205 https://doi.org/10.1017/9781316672815

[33] MILLER, J. and SHEFFIELD, S. (2016). Imaginary geometry I: Interacting SLEs. *Probab. Theory Related Fields* **164** 553–705. MR3477777 https://doi.org/10.1007/s00440-016-0698-0

[34] MILLER, J. and SHEFFIELD, S. (2017). Imaginary geometry IV: Interior rays, whole-plane reversibility, and space-filling trees. *Probab. Theory Related Fields* **169** 729–869. MR3719057 https://doi.org/10.1007/s00440-017-0780-2

[35] PEMANTLE, R. (1991). Choosing a spanning tree for the integer lattice uniformly. *Ann. Probab.* **19** 1559–1574. MR1127715

[36] PETROV, L. (2015). Asymptotics of uniformly random lozenge tilings of polygons. Gaussian free field. *Ann. Probab.* **43** 1–43. MR3298467 https://doi.org/10.1214/12-AOP823

[37] POMMERENKE, C. (1992). *Boundary Behaviour of Conformal Maps*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **299**. Springer, Berlin. MR1217706 https://doi.org/10.1007/978-3-662-02770-7

[38] ROMIK, D. (2012). Arctic circles, domino tilings and square Young tableaux. *Ann. Probab.* **40** 611–647. MR2952086 https://doi.org/10.1214/10-AOP628

[39] SCHRAMM, O. (2000). Scaling limits of loop-erased random walks and uniform spanning trees. *Israel J. Math.* **118** 221–288. MR1776084 https://doi.org/10.1007/BF02803524

[40] SHEFFIELD, S. (2016). Conformal weldings of random surfaces: SLE and the quantum gravity zipper. *Ann. Probab.* **44** 3474–3545. MR3551203 https://doi.org/10.1214/15-AOP1055

[41] TEMPERLEY, H. N. V. and FISHER, M. E. (1961). Dimer problem in statistical mechanics—an exact result. *Philos. Mag.* **6** 1061–1063. MR0136398

[42] THURSTON, W. P. (1990). Conway's tiling groups. *Amer. Math. Monthly* **97** 757–773. MR1072815 https://doi.org/10.2307/2324578

[43] YADIN, A. and YEHUDAYOFF, A. (2011). Loop-erased random walk and Poisson kernel on planar graphs. *Ann. Probab.* **39** 1243–1285. MR2857239 https://doi.org/10.1214/10-AOP579