The Annals of Probability
2016, Vol. 44, No. 5, 3474–3545
DOI: 10.1214/15-AOP1055

CONFORMAL WELDINGS OF RANDOM SURFACES: SLE AND THE QUANTUM GRAVITY ZIPPER[1]

BY SCOTT SHEFFIELD

Massachusetts Institute of Technology

We construct a conformal welding of two Liouville quantum gravity random surfaces and show that the interface between them is a random fractal curve called the Schramm–Loewner evolution (SLE), thereby resolving a variant of a conjecture of Peter Jones. We also demonstrate some surprising symmetries of this construction, which are consistent with the belief that (path-decorated) random planar maps have (SLE-decorated) Liouville quantum gravity as a scaling limit. We present several precise conjectures and open questions.

CONTENTS

Received October 2014; revised August 2015.

[1]Supported in part by a Grant from the Simons Foundation while at the Newton Institute, and also by DMS-12-09044, a fellowship from the Simons Foundation and EPSRC Grants EP/L018896/1 and EP/I03372X/1.

MSC2010 subject classifications. 60K35, 60D05.

Key words and phrases. Conformal welding, Gaussian free field, SLE, imaginary geometry.

1. Introduction.

1.1. *Overview.* Liouville quantum gravity and the Schramm–Loewner evolution (SLE) rank among the great mathematical physics discoveries of the last few decades. Liouville quantum gravity, introduced in the physics literature by Polyakov in 1981 in the context of string theory, is a canonical model of a random two-dimensional Riemannian manifold [74, 75]. The Schramm–Loewner evolution, introduced by Schramm in 1999, is a canonical model of a random path in the plane that does not cross itself [82, 85]. Each of these models is the subject of a large and active literature spanning physics and mathematics.

Our goal here is to connect these two objects to each other in the simplest possible way. Roughly speaking, we will show that if one glues together two independent Liouville quantum gravity random surfaces along boundary segments (in a boundary-length-preserving way)—and then conformally maps the resulting surface to a planar domain—then the interface between the two surfaces is an SLE.

Peter Jones conjectured several years ago that SLE could be obtained in a similar way—specifically, by gluing (what in our language amounts to) one Liouville quantum gravity random surface and one deterministic Euclidean disc. Astala, Jones, Kupiainen and Saksman showed that the construction Jones proposed produces a well-defined curve [5, 6], but Binder and Smirnov recently announced a proof (involving multifractal exponents) that this curve is not a form of SLE, and hence the original Jones conjecture is false [96] (see Section 1.5). Our construction shows that a simple variant of the Jones conjecture is in fact true.

Beyond this, we discover some surprising symmetries. For example, it turns out that there is one particularly natural random simply connected surface (called a γ*-quantum wedge*) that has an infinite-length boundary isometric to $\mathbb{R}$ (almost surely) which contains a distinguished "origin." Although this surface is simply connected, it is almost surely highly nonsmooth and it has a random fractal structure. We will explain precisely how it is defined in Section 1.6. The origin divides the boundary into two infinite-length boundary arcs. Suppose we glue (in a boundary-length preserving way) the right arc of one such surface to the left arc of an independent random surface with the same law, then conformally map the combined surface to the complex upper half plane $\mathbb{H}$ (sending the origin to the origin and ∞ to ∞—see figure below), and then *erase* the boundary interface. The

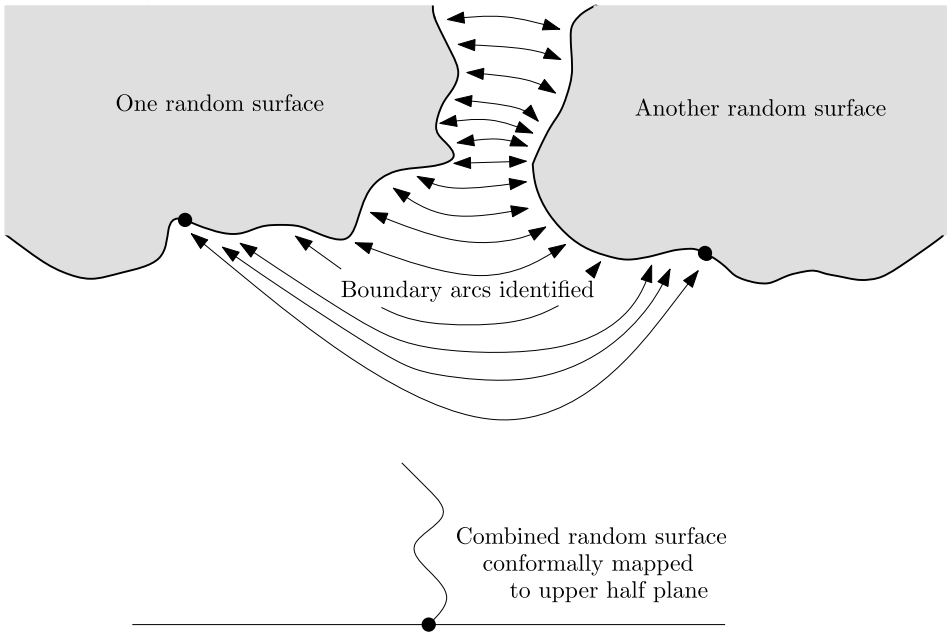

FIG. 1. *Welding surfaces.*

geometric structure of the combined surface can be pushed forward to give geometric structure (including an area measure) on $\mathbb{H}$. One wonders if one can guess, from this geometric structure on $\mathbb{H}$, where the now-erased interface used to be, see Figure 1.

We will show that the geometric structure yields no information at all. That is, the conditional law of the interface is that of an SLE in $\mathbb{H}$ *independently* of the underlying geometry (a fact formally stated as part of Theorem 1.8). Another way to put this is that conditioned on the combined surface, all of the information about the interface is contained in the *conformal structure* of the combined surface, which determines the embedding in $\mathbb{H}$ (up to rescaling $\mathbb{H}$ via multiplication by a positive constant, which does not affect the law of the path, since the law of SLE is scale-invariant).

This apparent coincidence is actually quite natural from one point of view. We recall that one reason (among many) for studying SLE is that it arises as the fine mesh "scaling limit" of random simple paths on lattices. Liouville quantum gravity is similarly believed (though not proved) to be the scaling limit of random discretized surfaces and random planar maps. The independence mentioned above turns out to be consistent with (indeed, at least heuristically, a *consequence of*) certain scaling limit conjectures (and a related conformal invariance Ansatz) that we will formulate precisely (in Section 2.2) for the first time here.

Polyakov initially proposed Liouville quantum gravity as a model for the intrinsic Riemannian manifold parameterizing the space–time trajectory of a string [74]. From this point of view, the welding/subdivision of such surfaces is analogous to the concatenation/subdivision of one-dimensional time intervals (which parameterize point-particle trajectories). It seems natural to try to understand complicated string trajectories by decomposing them into simpler pieces (and/or gluing pieces together), which should involve subdividing and/or welding the corresponding Liouville quantum gravity surfaces. The purpose of this paper is to study these weldings and subdivisions mathematically. We will not further explore the physical implications here.

In a recent memoir [76], Polyakov writes that he first became convinced of the connection between the discrete models and Liouville quantum gravity in the 1980s after jointly deriving, with Knizhnik and Zamolodchikov, the so-called *KPZ formula* for certain Liouville quantum gravity scaling dimensions and comparing them with known combinatorial results for the discrete models [46]. With Duplantier, the present author recently formulated and proved the KPZ formula in a mathematical way [28, 29] (see also [9, 79]). This paper is in some sense a sequel to [29], and we refer the reader there for references and history.

We will find it instructive to develop Liouville quantum gravity along with a closely related construction called the *AC geometry* or *imaginary geometry*. Both Liouville quantum gravity and the imaginary geometry are based on a simple object called the Gaussian free field.

1.2. *Random geometries from the Gaussian free field.* The two-dimensional Gaussian free field (GFF) is a natural higher dimensional analog of Brownian motion that plays a prominent role in mathematics and physics. See the survey [91] and the introductions of [86, 87] for a detailed account. On a planar domain D, one can define both a zero boundary GFF and a free boundary GFF (the latter being defined only modulo an additive constant, which we will sometimes fix arbitrarily). In both cases, an instance of the GFF is a random sum

$$h = \sum_i \alpha_i f_i,$$

where the α_i are i.i.d. mean-zero unit-variance normal random variables, and the f_i are an orthonormal basis for a Hilbert space of real-valued functions on D (or in the free boundary case, functions modulo additive constants) endowed with the Dirichlet inner product

$$(f_1, f_2)_\nabla := (2\pi)^{-1} \int_D \nabla f_1(z) \cdot \nabla f_2(z)\, dz.$$

The Hilbert space is the completion of either the space of smooth compactly supported functions $f : D \to \mathbb{R}$ (zero boundary) or the space of all smooth functions $f : D \to \mathbb{R}$ modulo additive constants with $(f, f)_\nabla < \infty$ (free boundary). In each

case, h is understood not as a random function on D but as a random distribution or generalized function on D. (Mean values of h on certain sets are also defined, but the value of h at a particular point is not defined.) One can fix the additive constant for the free boundary GFF in various ways, for example, by requiring the mean value of h on some set to be zero. We will review these definitions in Section 3.

There are two natural ways to produce a "random geometry" from the Gaussian free field. The first construction is (critical) *Liouville quantum gravity*. Here, one replaces the usual Lebesgue measure dz on a smooth domain D with a random measure $\mu_h = e^{\gamma h(z)}\,dz$, where $\gamma \in [0, 2)$ is a fixed constant and h is an instance of (for now) the free boundary GFF on D (with an additive constant somehow fixed—we will actually consider various ways of fixing the additive constant later in the paper; one way is to require the mean value of h on some fixed set to be 0). Since h is not defined as a function on D, one has to use a regularization procedure to be precise:

$$\mu = \mu_h := \lim_{\varepsilon \to 0} \varepsilon^{\gamma^2/2} e^{\gamma h_\varepsilon(z)}\,dz, \tag{1.1}$$

where dz is Lebesgue measure on D, $h_\varepsilon(z)$ is the mean value of h on the circle $\partial B_\varepsilon(z)$ and the limit represents weak convergence (on compact subsets) in the space of measures on D. (The limit exists almost surely, at least if ε is restricted to powers of two [29].) We interpret μ_h as the area measure of a random surface conformally parameterized by D. When $x \in \partial D$, we let $h_\varepsilon(x)$ be the mean value of h on $D \cap \partial B_\varepsilon(x)$. On a linear segment of ∂D, we may define a boundary length measure by

$$\nu = \nu_h := \lim_{\varepsilon \to 0} \varepsilon^{\gamma^2/4} e^{\gamma h_\varepsilon(x)/2}\,dx, \tag{1.2}$$

where dx is Lebesgue measure on ∂D. (For details, see [29], which also relates the above random measures to the curvature-based action used to define Liouville quantum gravity in the physics literature.)

We could also parameterize the same surface with a different domain $\widetilde{D}$, and our regularization procedure implies a simple rule for changing coordinates. Suppose that ψ is a conformal map from a domain $\widetilde{D}$ to D and write $\tilde{h}$ for the distribution on $\widetilde{D}$ given by $h \circ \psi + Q \log|\psi'|$ where

$$Q := \frac{2}{\gamma} + \frac{\gamma}{2},$$

as in Figure 2.[2] Then μ_h is almost surely the image under ψ of the measure $\mu_{\tilde{h}}$. That is, $\mu_{\tilde{h}}(A) = \mu_h(\psi(A))$ for $A \subset \widetilde{D}$. Similarly, ν_h is almost surely the image

[2] We use the same distribution composition notation as [29]: that is, if ϕ is a conformal map from D to a domain $\widetilde{D}$ and h is a distribution on D, then we define the pullback $h \circ \phi^{-1}$ of h to be a distribution on $\widetilde{D}$ defined by $(h \circ \phi^{-1}, \tilde{\rho}) = (h, \rho)$ whenever $\rho \in H_s(D)$ and $\tilde{\rho} = |\phi'|^{-2}\rho \circ \phi^{-1}$. (Here, ϕ' is the complex derivative of ϕ, and (h, ρ) is the value of the distribution h integrated against ρ.)

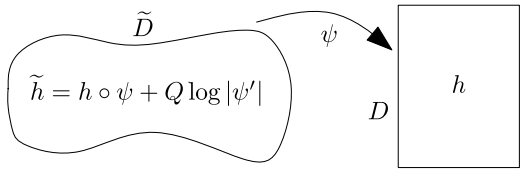

FIG. 2. *A quantum surface coordinate change.*

under ψ of the measure $\nu_{\tilde{h}}$ [29]. In fact, [29] formally defines a *quantum surface* to be an equivalence class of pairs (D, h) under the equivalence transformations (see Figure 2)

$$(D, h) \to \psi^{-1}(D, h) := \big(\psi^{-1}(D), h \circ \psi + Q \log|\psi'|\big) = (\widetilde{D}, \tilde{h}), \tag{1.3}$$

noting that both area and boundary length are well defined for such surfaces. The invariance of ν_h under (1.3) actually yields a definition of the quantum boundary length measure ν_h when the boundary of D is not piecewise linear, that is, in this case, one simply maps to the upper half plane (or any other domain with a piecewise linear boundary) and computes the length there.[3]

The second construction involves "flow lines" of the unit vector field $e^{ih/\chi}$ where $\chi \neq 0$ is a fixed constant (see Figure 9), or alternatively flow lines of $e^{i(h/\chi + c)}$ for a constant $c \in [0, 2\pi)$. The author has proposed calling this collection of flow lines the *AC geometry*[4] of h, but a recent series of joint works with Jason Miller uses the term *imaginary geometry* [64–67]. Makarov once proposed the term "magnetic gravity" in a lecture, suggesting that in some sense the AC geometry is to Liouville quantum gravity as electromagnetism is to electrostatics. We will discuss additional interpretations in Section 2 and the Appendix.

Although h is a distribution and not a function, one can make sense of flow lines using the couplings between the Schramm–Loewner evolution (SLE) and the GFF

Note that if h is a continuous function [viewed as a distribution via the map $\rho \to \int_D \rho(z)h(z)\,dz$], then the distribution $h \circ \phi^{-1}$ thus defined is the ordinary composition of h and ϕ^{-1} (viewed as a distribution).

[3]It remains an open question whether the *interior* of a quantum surface is canonically a metric space. A pair (D, h) is a metric space parameterized by D when, for distinct $x, y \in D$ and $\delta > 0$, one defines the distance $d_\delta(x, y)$ to be the smallest number of Euclidean balls in D of μ_h mass δ required to cover *some* continuous path from x to y in D. We conjecture but cannot prove that for some constant β the limiting metric

$$\lim_{\delta \to 0} \delta^\beta d_\delta$$

exists a.s. and is invariant under the transformations described by (1.3).

[4]AC stands for "altimeter-compass." If the graph of h is viewed as a mountainous terrain, then a hiker holding an analog altimeter—with a needle indicating altitude modulo $2\pi\chi$—in one hand and a compass in the other can trace an AC ray by walking at constant speed (continuously changing direction as necessary) in such a way that the two needles always point in the same direction.

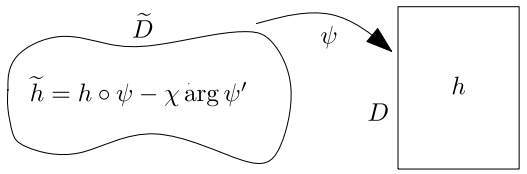

FIG. 3. *An AC surface coordinate change.*

in [87, 90], which were further developed in [20] and more recently in [38, 39, 59]. The paths in these couplings are generalizations of the GFF contour lines of [87].

We define an *AC surface* to be an equivalence class of pairs under the following variant of (1.3):

$$(D, h) \to \left(\psi^{-1}(D), h \circ \psi - \chi \arg \psi'\right) = (\widetilde{D}, \tilde{h}), \tag{1.4}$$

as in Figure 3. The reader may observe that (at least when h is smooth) the flow lines of the LHS of (1.4) are the ψ images of the flow lines of the RHS. To check this, first consider the simplest case: if ψ^{-1} is a rotation (i.e., multiplication by a modulus-one complex number), then (1.4) ensures that the unit flow vectors $e^{ih/\chi}$ (as in Figure 9) are rotated by the same amount that D is rotated. The general claim follows from this, since every conformal map looks locally like the composition of a dilation and a rotation (see Section 2.1).

Recalling the conformal invariance of the GFF, if the h on the left-hand side of (1.3) and (1.4) is a centered (expectation zero) Gaussian free field on D then the distribution on the right-hand side is a centered (expectation zero) GFF on $\widetilde{D}$ *plus* a deterministic function. In other words, *changing the domain of definition is equivalent to recentering the GFF*. The deterministic function is harmonic if D is a planar domain, but it can also be defined (as a nonharmonic function) when D is a surface with curvature (see [29]). In what follows, we will often find it convenient to define quantum and AC surfaces on the complex half-plane $\mathbb{H}$ using a (free or zero boundary) GFF on $\mathbb{H}$, sometimes recentered by the addition of a deterministic function that we will call $\mathfrak{h}_0$. We will state our main results in the introduction for fairly specific choices of $\mathfrak{h}_0$. We will extend these results to more general underlying geometries in Section 4 and Section 5.

1.3. *Theorem statements*: *SLE/GFF couplings*. We will give explicit relationships between the Gaussian free field and both "forward" and "reverse" forms of SLE in Theorems 1.1 and 1.2 below. We will subsequently interpret these theorems as statements about AC geometry and Liouville quantum gravity, respectively. We will prove Theorem 1.1 in Section 4.1 using a series of calculations. These calculations are not really new to this paper, although the precise form of the argument we give has not been published elsewhere.[5] Our main reason for proving Theorem 1.1

[5]The argument presented in Section 4.1, together with the relevant calculations, appeared in lecture slides some time ago [90], and is by now reasonably well known. Dubédat presented another short

in Section 4.1 is that we wish to simultaneously prove Theorem 1.2. Theorem 1.2 is completely new to this paper (and essential to the other results obtained in this paper), but it is very closely related to Theorem 1.1. Proving the two results side by side allows us to highlight the similarities and differences.

We will not give a detailed introduction to SLE here, but there are many excellent surveys on SLE; see, for example, the introductory references [48, 104] for basic SLE background. To set notation, we recall that when η is an instance of chordal SLE_κ in $\mathbb{H}$ from 0 to ∞, the conformal maps $g_t : \mathbb{H} \setminus \eta([0,t]) \to \mathbb{H}$, normalized so that $\lim_{z\to\infty} |g_t(z) - z| = 0$, satisfy

$$dg_t(z) = \frac{2}{g_t(z) - W_t}\, dt, \tag{1.5}$$

with $W_t = \sqrt{\kappa} B_t = g_t(\eta(t))$, where B_t is a standard Brownian motion. In fact, this can be taken as the definition of SLE_κ. Rohde and Schramm proved in [82] that for each κ and instance of B_t, there is almost surely a unique continuous curve η in $\mathbb{H}$ from 0 to ∞, parameterized by $[0,\infty)$, for which (1.5) holds for all t. When η is parameterized so that (1.5) holds, the quantity t is called the (half-plane) *capacity* of $\gamma([0,t])$. The curve η is almost surely a simple curve when $\kappa \in [0,4]$, a self-intersecting but nonspace-filling curve when $\kappa \in (4,8)$, and a space-filling curve (ultimately hitting every point in $\mathbb{H}$) when $\kappa \geq 8$ [82].

The maps

$$f_t(z) := g_t(z) - W_t$$

satisfy

$$df_t(z) = \frac{2}{f_t(z)}\, dt - \sqrt{\kappa}\, dB_t,$$

and $f_t(\eta(t)) = 0$. Throughout this paper, we will use f_t rather than g_t to describe the Loewner flow. If $\eta_T = \eta([0,T])$ is a segment of an SLE trace, denote by K_T the complement of the unbounded component of $\mathbb{H} \setminus \eta_T$. In the statements of Theorem 1.1 and Theorem 1.2 below and throughout the paper, we will discuss several kinds of random distributions on $\mathbb{H}$. To show that these objects are well defined as distributions on $\mathbb{H}$, we will make implicit use of some basic facts about distributions:

1. If h is a distribution on a domain D then its restriction to a subdomain is a distribution on that subdomain. (This follows by simply restricting the class of test functions to those supported on the subdomain.)

---

derivation of this statement within a long foundational paper [20]. More recent variants appear in [38, 39], and in the series of papers [64–67], which studies the couplings in further detail. Prior to these works, Kenyon and Schramm derived (but never published) a calculation relating SLE to the GFF in the case $\kappa = 8$. One could also have inferred the existence of such a relationship from the fact—due to Lawler, Schramm and Werner—that SLE_8 is a continuum scaling limit of uniform spanning tree boundaries [50], and the fact—due to Kenyon—that the winding number "height functions" of uniform spanning trees have the GFF as a scaling limit [43–45].

2. If h a distribution on a domain D and ϕ is a conformal map from D to a domain $\widetilde{D}$ then $h \circ \phi^{-1}$ is a distribution on $\widetilde{D}$. (Recall footnote 2.)

3. An instance of the zero boundary GFF on a subdomain of D is also well defined as a distribution on all of D. (See Section 2.1 of [87].)

4. If h is an L^1 function on D, then h can be understood as a distribution on D defined by $(h, \rho) = \int_D \rho(z)h(z)\,dz$.

In the proof of Theorem 1.1 in Section 4.1, we will show that even though the function $\arg f_t'$ that appears in the theorem statement is a.s. unbounded, it can also a.s. be understood as a distribution on $\mathbb{H}$ (see the discussion after the theorem statement below).

THEOREM 1.1. *Fix $\kappa \in (0, 4]$ and let η_T be the segment of SLE_κ generated by the Loewner flow*

$$df_t(z) = \frac{2}{f_t(z)}\,dt - \sqrt{\kappa}\,dB_t, \qquad f_0(z) = z \tag{1.6}$$

up to a fixed time $T > 0$. Write

$$\mathfrak{h}_0(z) := \frac{-2}{\sqrt{\kappa}} \arg z, \qquad \chi := \frac{2}{\sqrt{\kappa}} - \frac{\sqrt{\kappa}}{2},$$
$$\mathfrak{h}_t(z) := \mathfrak{h}_0\big(f_t(z)\big) - \chi \arg f_t'(z).$$

Here, $\arg(f_t(z))$ (which is a priori defined only up to an additive multiple of 2π) is chosen to belong $(0, \pi)$ when $f_t(z) \in \mathbb{H}$; we similarly define $\arg f_t'(z)$ by requiring that (when t is fixed) it is continuous on $\mathbb{H} \setminus \eta_T$ and tends to 0 at ∞. Let $\tilde{h}$ be an instance of the zero boundary GFF on $\mathbb{H}$, independent of B_t. Then the following two random distributions on $\mathbb{H}$ agree in law:[6]

$$h := \mathfrak{h}_0 + \tilde{h},$$
$$h \circ f_T - \chi \arg f_T' = \mathfrak{h}_T + \tilde{h} \circ f_T.$$

The two distributions above also agree in law when $\kappa \in (4, 8)$ if we replace $\tilde{h} \circ f_T$ with a GFF on $\mathbb{H} \setminus \eta([0, t])$ (which in this case means the sum of an independent zero boundary GFF on each component of $\mathbb{H} \setminus \eta([0, t])$) and take $\mathfrak{h}_t(z) := \lim_{s \to \tau(z)_-} \mathfrak{h}_s(z)$ if z is absorbed at time $\tau(z) \le t$.

Alternative statement of Theorem 1.1: Using our coordinate change and AC surface definitions, we may state the theorem when $\kappa < 4$ somewhat more elegantly as follows: the law of the AC surface $(\mathbb{H}, h)$ is invariant under the operation of independently sampling f_T using a Brownian motion and (1.6), transforming

[6]Note that f_T maps $\mathbb{H} \setminus K_T$ to $\mathbb{H}$, so $(\mathbb{H}, h)$ and $(\mathbb{H} \setminus K_T, h \circ f_T - \chi \arg f_T')$ describe equivalent AC surfaces by (1.4).

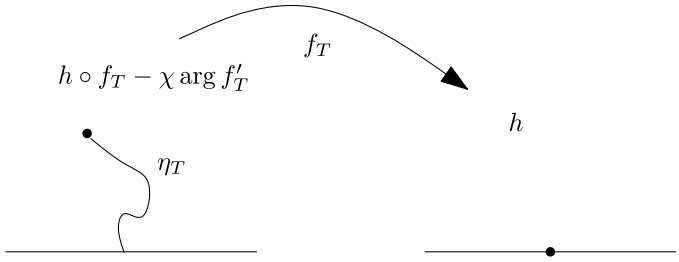

FIG. 4. *Forward coupling.*

the AC surface via the coordinate change f_T^{-1} (going from right to left in Figure 4[7]—see also Figure 11) in the manner of (1.4), and erasing the path η_T (to obtain an AC surface parameterized by $\mathbb{H}$ instead of $\mathbb{H} \setminus \eta_T$). We discuss the geometric intuition behind the alternative statement in Section 2.1.

Note that, as a function, $\mathfrak{h}_T$ is not defined on η_T itself. However, we will see in Section 4.1 that $\mathfrak{h}_T$ is a.s. well defined as a distribution, independently of how we define it as a function on η_T itself. This will follow from the fact that, when $\kappa = 4$, this $\mathfrak{h}_T$ is almost surely a bounded function off of η_T, and when $\kappa \neq 4$, the restriction of $\mathfrak{h}_T$ to any compact subset of $\mathbb{H}$ is almost surely in L^p for each $p < \infty$ (see Section 4). The fact that $\tilde{h} \circ f_T$ is well defined as a distribution on $\mathbb{H}$ (not just as a distribution on $\mathbb{H} \setminus \eta_T$) follows from conformal invariance of the GFF, and the fact (mentioned above, proved in [87]) that a zero boundary GFF instance on a subdomain can be understood as a distribution on the larger domain.

Another standard approach for generating a segment η_T of an SLE curve is via the reverse Loewner flow, whose definition is recalled in the statement of the following theorem. (Note that if T is a fixed constant, then the law of the η_T generated by reverse Loewner evolution is the same as that generated by forward Loewner evolution; see Figures 4 and 5.)

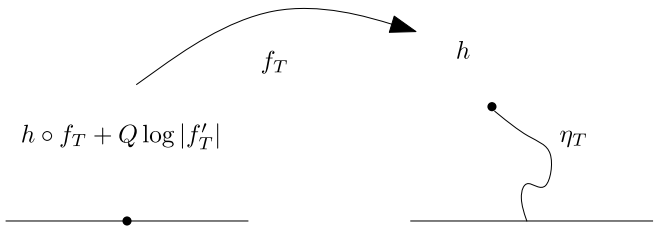

FIG. 5. *Reverse coupling.*

[7]All figures in this paper are sketches, not representative simulations.

THEOREM 1.2. *Fix $\kappa > 0$ and let η_T be the segment of SLE_κ generated by a reverse Loewner flow*

$$df_t(z) = \frac{-2}{f_t(z)}\,dt - \sqrt{\kappa}\,dB_t, \qquad f_0(z) = z \tag{1.7}$$

up to a fixed time $T > 0$. Write

$$\mathfrak{h}_0(z) := \frac{2}{\sqrt{\kappa}} \log|z|, \qquad Q := \frac{2}{\sqrt{\kappa}} + \frac{\sqrt{\kappa}}{2},$$
$$\mathfrak{h}_t(z) := \mathfrak{h}_0\big(f_t(z)\big) + Q \log\big|f_t'(z)\big|,$$

and let $\tilde{h}$ be an instance of the free boundary GFF on $\mathbb{H}$, independent of B_t. Then the following two random distributions (modulo additive constants) on $\mathbb{H}$ agree in law:[8]

$$h := \mathfrak{h}_0 + \tilde{h},$$
$$h \circ f_T + Q \log\big|f_T'\big| = \mathfrak{h}_T + \tilde{h} \circ f_T.$$

Alternative statement of Theorem 1.2: A more elegant way to state the theorem is that the law of $(\mathbb{H}, h)$ is invariant under the operation of independently sampling f_T, cutting out K_T (equivalent to η_T when $\kappa \le 4$), and transforming via the coordinate change f_T^{-1} (going from right to left in Figure 5) in the manner of (1.3).

Both theorems give us an *alternate* way of sampling a distribution with the law of h, that is, by first sampling the B_t process (which determines η_T), then sampling a (fixed or free boundary) GFF $\tilde{h}$ and taking

$$h = \mathfrak{h}_T + \tilde{h} \circ f_T.$$

This two part sampling procedure produces a coupling of η_T with h. In the forward SLE setting of Theorem 1.1, it was shown in [20] that in any such coupling, η_T is almost surely equal to a particular path-valued *function* of h. (This was also done in [87] in the case $\kappa = 4$.) In other words, in such a coupling, h determines η_T almost surely. This is important for our geometric interpretations. Even though h is not defined pointwise as a function, we would like to geometrically interpret η as a level set of h (when $\kappa = 4$) or a flow line of $e^{ih/\chi}$ (when $\kappa < 4$), as we stated above and will explain in more detail in Section 2.1. It is thus conceptually natural that such curves are uniquely determined by h (as they would be if h were a smooth function, see Section 2.1).

As mentioned earlier, this paper introduces and proves Theorem 1.2 while highlighting its similarity to Theorem 1.1. Indeed, it won't take us much more work to

[8]Note that f_T maps $\mathbb{H}$ to $\mathbb{H} \setminus K_T$, so $(\mathbb{H}, h \circ f_T + Q \log|f_T'|)$ and $(\mathbb{H} \setminus K_T, h)$ describe equivalent quantum surfaces by (1.3). Indeed, $(\mathbb{H}, h \circ f_T + Q \log|f_T'|) = f_T^{-1}(\mathbb{H} \setminus K_T, h)$.

prove Theorems 1.1 and 1.2 together than it would take to prove one of the two theorems alone. It turns out that in both Figure 4 (which illustrates Theorem 1.1) and Figure 5 (which illustrates Theorem 1.2), the field illustrated on the left side of the figure (which agrees with h in law) actually determines η_T and the map f_T, at least when $\kappa < 4$. In the former context (Figure 4), this a major result due to Dubédat [20] (see also the exposition on this point in [64]). It says that a certain "flow line" is a.s. uniquely determined by h. The statement in the latter context is a major result obtained in this paper, stated in Theorems 1.3 and 1.4. With some hard work, we will be able to show that the map f_T describes a conformal welding in which boundary arcs of equal quantum boundary length are "welded together." Once we have this, the fact that the boundary measure uniquely characterizes f_T will be obtained by applying a general "removability" result of Jones and Smirnov, as we will explain in Section 1.4.

1.4. *Theorem statements*: *Conformal weldings*. We will now try to better understand Theorem 1.2 in the special case $\kappa < 4$. Note that a priori the h in Theorem 1.2 is defined only up to additive constant. We can either choose the constant arbitrarily (e.g., by requiring that the mean value of h on some set be zero) or avoid specifying the additive constant and consider the measures μ_h and ν_h to be defined only up to a global multiplicative constant. The choice does not affect the theorem statement below.

THEOREM 1.3. *Suppose that $\kappa < 4$ and that h and η_T are coupled in the way described at the end of the previous section, that is, h is generated by first sampling the B_t process up to time T in order to generate f_T via a reverse Loewner flow, and then choosing $\tilde{h}$ independently and writing $h = \mathfrak{h}_T + \tilde{h} \circ f_T$, and $\eta_T((0,T]) = \mathbb{H} \setminus f_T\mathbb{H}$.*[9] *Given a point z along the path η_T, let $z_- < 0 < z_+$ denote the two points in $\mathbb{R}$ that f_T (continuously extended to $\mathbb{R}$) maps to z. Then almost surely*

$$\nu_h([z_-, 0]) = \nu_h([0, z_+])$$

for all z on η_T.

Theorem 1.3 is a relatively difficult theorem, and it will be the last thing we prove. We next define $R = R_h : (-\infty, 0] \to [0, \infty)$ so that $\nu_h([x, 0]) = \nu_h([0, R(x)])$ for all x (recall that ν is a.s. atom free [29]). This R gives a homeomorphism from $[0_-, 0]$ to $[0, 0_+]$ that we call a *conformal welding* of these two intervals. We stress that the values 0_- and 0_+ depend on T, but the overall homeomorphism R between $(-\infty, 0]$ and $[0, \infty)$ is determined by the boundary measure ν_h, whose law does not depend on T (although the coupling between h, $\tilde{h}$ and

[9] It is not known whether an analog of Theorem 1.3 can obtained in the case $\kappa = 4$. The standard procedure for constructing the boundary measure ν_h breaks down when $\kappa = 4, \gamma = 2$, but a scheme was introduced [25, 26] to create a nontrivial boundary measure ν_h. The open problems listed in Section 6 also address a related question in the $\kappa > 4$ setting.

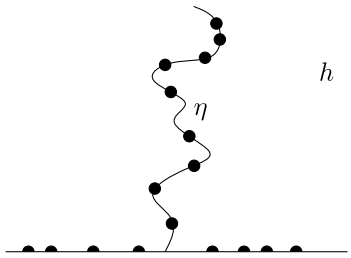

FIG. 6. *Sketch of η with marks spaced at intervals of the same v_h length along ∂D_1 and ∂D_2. Here, $(-\infty, 0]$ and $[0, \infty)$ are the two open strands of the "zipper" while η is the closed (zipped up) strand. Semicircular dots on $\mathbb{R}$ are "zipped together" by f_t^h. Circular dots on η are "pulled apart" by f_t^η. (Recall that under the reverse Loewner flow f_t^h, the center of a semicircle on the negative real axis will reach the origin at the same time as the center of the corresponding semicircle on the positive real axis.) The law of $((D_1, h^{D_1}), (D_2, h^{D_2}))$ is invariant under "zipping up" by t capacity units or "zipping down" by t capacity units.*

η_T described in the theorem statement clearly depends on T). Since η_T is simple, it clearly determines the restriction of R to $[0_-, 0]$. (See Figure 6.) It turns out that R also determines η_T.

THEOREM 1.4. *For $\kappa < 4$, in the setting of Theorem 1.3, the homeomorphism R from $[0_-, 0]$ to $[0, 0_+]$ uniquely determines the curve η_T. In other words, it is almost surely the case that if $\tilde{\eta}_{\tilde{T}}$ is any other simple curve in $\mathbb{H}$ such that the homeomorphism induced by its reverse Loewner flow is the same as R on $[0_-, 0]$, then $\tilde{\eta}_{\tilde{T}} = \eta_T$. In particular, h determines η_T almost surely.*

PROOF. The author learned from Smirnov that Theorem 1.4 follows almost immediately from Theorem 1.3 together with known results in the literature. If there were a distinct candidate $\tilde{\eta}_T$ with a corresponding $\tilde{f}_T$, then $\phi = \tilde{f}_T \circ f_T^{-1}$—extended from $\mathbb{H}$ to $\mathbb{R}$ by continuity, and to all of $\mathbb{C}$ by Schwarz reflection—would be a nontrivial homeomorphism of $\mathbb{C}$ [with $\lim_{z\to\infty} \phi(z) - z = 0$] which was conformal on $\mathbb{C} \setminus (\eta_T \cup \bar{\eta}_T)$, where $\bar{\eta}_T$ denotes the complex conjugate of η_T. Thus, to prove Theorem 1.4, it suffices to show that no such map exists. In complex analysis terminology, this is equivalent by definition to showing that the curve $\eta_T \cup \bar{\eta}_T$ is *removable*. Rohde and Schramm showed that the complement of $\eta([0, T])$ is a.s. a Hölder domain for $\kappa < 4$ (see Theorem 5.2 of [82]) and that η is a.s. a simple curve in this setting. In particular, $\eta_T \cup \bar{\eta}_T$ is almost surely the boundary of its complement, and this complement is a Hölder domain. (More about Hölder continuity appears in work of Beliaev and Smirnov [7] and Kang [42] and Lind [57].) Jones and Smirnov showed generally that boundaries of Hölder domains are removable (Corollary 2 of [41]). The same observations are used in [6]. □

We remark that the above arguments also show that $\eta \cup \bar{\eta}$ is removable when η is the entire SLE path. In the coming sections, we will often interpret the left and

right components of $\mathbb{H} \setminus \eta$ as distinct quantum surfaces, where the right boundary arc of one surface is welded (along η) to the left boundary arc of another surface in a quantum-boundary-length-preserving way. When the law of η is given by SLE_κ with $\kappa < 4$, removability implies that η is almost surely determined (up to a constant rescaling of $\mathbb{H}$) by the way that these boundary arcs are identified. In other words, aside from constant rescalings, there is no homeomorphism of $\mathbb{H}$, fixing 0 and ∞, whose restriction to $\mathbb{H} \setminus \eta$ is conformal.

1.5. *Corollary*: *Capacity stationary quantum zipper.* This subsection contains some discussion and interpretation of some simple consequences of Theorems 1.3 and 1.4, in particular Corollary 1.5 below. We first observe that for $\kappa < 4$, Theorem 1.4 implies that R determines η_T almost surely for any given $T > 0$. In particular, this means that R determines an entire reverse Loewner evolution $f_t = f_t^h$ for all $t \geq 0$, and that this f_t^h is (in law) a *reverse* SLE_κ flow. Similarly, given a chordal curve η from 0 to ∞ in $\mathbb{H}$, we denote by f_t^η the *forward* Loewner flow corresponding to η. The following is now an immediate corollary of the domain Markov property for SLE and Theorems 1.2, 1.3 and 1.4. As usual, transformations $f(D, h)$ are defined using (1.3).

COROLLARY 1.5. *Fix $\kappa \in (0, 4)$. Let $h = \mathfrak{h}_0 + \tilde{h}$ be as in Theorem 1.2 and let η be an SLE_κ on $\mathbb{H}$ chosen independently of h. Let D_1 be the left component of $\mathbb{H} \setminus \eta$ and h^{D_1} the restriction of h to D_1. Let D_2 be the right component of $\mathbb{H} \setminus \eta$ and h^{D_2} the restriction of h to D_2. For $t \geq 0$, write*

$$\mathcal{Z}_t^{\mathrm{CAP}}\big((D_1, h^{D_1}), (D_2, h^{D_2})\big) = \big(f_t^h(D_1, h^{D_1}), f_t^h(D_2, h^{D_2})\big),$$
$$\mathcal{Z}_{-t}^{\mathrm{CAP}}\big((D_1, h^{D_1}), (D_2, h^{D_2})\big) = \big(f_t^\eta(D_1, h^{D_1}), f_t^\eta(D_2, h^{D_2})\big).$$

Note that both h and η are determined by the pair $((D_1, h^{D_1}), (D_2, h^{D_2}))$, and that f_t^h and f_t^η are also a.s. determined by this pair, so that the maps $\mathcal{Z}_t^{\mathrm{CAP}}$ and $\mathcal{Z}_{-t}^{\mathrm{CAP}}$ are well defined for almost all pairs $((D_1, h^{D_1}), (D_2, h^{D_2}))$ chosen in the manner described above. Then the law of $((D_1, h^{D_1}), (D_2, h^{D_2}))$ is invariant under $\mathcal{Z}_t^{\mathrm{CAP}}$ for all t. Also, for all s and t,

$$\mathcal{Z}_{s+t}^{\mathrm{CAP}} = \mathcal{Z}_s^{\mathrm{CAP}} \mathcal{Z}_t^{\mathrm{CAP}}$$

almost surely.

Because the forward and reverse Loewner evolutions are parameterized according to half plane capacity, we refer to the group of transformations $\mathcal{Z}_t^{\mathrm{CAP}}$ as the *capacity quantum zipper*; see Figure 6. (The term "zipper" in the Loewner evolution context has been used before; see the "zipper algorithm" for numerically computing conformal mappings in [61] and the references therein.) When $t > 0$, applying $\mathcal{Z}_t^{\mathrm{CAP}}$ is called "zipping up" the pair of quantum surfaces by t capacity

units and applying $\mathcal{Z}_{-t}^{\mathrm{CAP}}$ is called "zipping down" or "unzipping" by t capacity units.

To begin to put this construction in context, we recall that the general *conformal welding* problem is usually formulated in terms of identifying unit discs $\mathbb{D}_1$ and $\mathbb{D}_2$ along their boundaries via a given homeomorphism ϕ from $\partial\mathbb{D}_1$ to $\partial\mathbb{D}_2$ to create a sphere with a conformal structure. Precisely, one wants a simple loop η in the complex sphere, dividing the sphere into two pieces such that if conformal maps ψ_i from the $\mathbb{D}_i$ to the two pieces are extended continuously to their boundaries, then $\psi_1 \circ \psi_2^{-1}$ is ϕ. In general, not every homeomorphism ϕ between disc boundaries is a conformal welding in this way, and when it is, it does not always come from an η that is (modulo conformal automorphisms of the sphere) unique; in fact, arbitrarily small changes to ϕ can lead to large changes in η and some fairly exotic behavior (see, e.g., [13]).

The theorems of this paper can also be formulated in terms of a sphere obtained by gluing two discs along their boundaries: in particular, one can zip up the quantum surfaces of Corollary 1.5 "all the way" (see Figure 16 and Section 5.2), which could be viewed as welding two Liouville quantum surfaces (each of which is topologically homeomorphic to a disc) to obtain an SLE loop in the sphere, together with an instance of the free boundary GFF on the sphere.

Note that in the construction described above, the quantum surfaces are defined only modulo an additive constant for the GFF, and we construct the two surfaces together in a particular way. In Section 1.6 (Theorem 1.8), we will describe a related construction in which one takes two *independent* quantum surfaces (each with its additive constant well defined) and welds them together to obtain SLE.

As mentioned in Section 1.1, Peter Jones conjectured several years ago that an SLE loop could be obtained by (what in our language amounts to) welding a quantum surface to a deterministic Euclidean disc. (The author first learned of this conjecture during a private conversation with Jones in early 2007 [40].) Astala, Jones, Kupiainen and Saksman recently showed that such a welding exists and determines a unique loop (up to conformal automorphism of the sphere) [5, 6]. Binder and Smirnov recently announced (to the author, in private communication [96]) that they have obtained a proof that the original conjecture of Jones is false. By computing a multifractal spectrum, they showed that the loop constructed in [5, 6] does not look locally like SLE. However, our construction, together with Theorem 1.8 below, shows that a natural variant of the Jones conjecture—involving two independent quantum surfaces instead of one quantum surface and one Euclidean disc—is in fact true.

We also remark that the "natural" d-dimensional measure on (or parameterization of) an SLE curve of Hausdorff dimension d was only constructed fairly recently [49, 51, 52], and it was shown to be uniquely characterized by certain symmetries, in particular the requirement that it transforms like a d-dimensional measure under the maps f_t (i.e., if the map locally stretches space by a factor of r, then it locally increases the measure by a factor of r^d). Our construction here can

be viewed as describing, for $\kappa < 4$, a natural "quantum" parameterization of SLE_κ, which is similarly characterized by transformation laws, in particular the requirement that adding C to h—which scales area by a factor of $e^{\gamma C}$—scales length by a factor of $e^{\gamma C/2}$. These ideas are discussed further in [30].

The relationship between Euclidean and quantum natural fractal measures and their evolution under capacity invariant quantum zipping is developed in [30] in a way that makes use of the KPZ formula [29, 46].

1.6. *Quantum wedges and quantum length stationarity.* This subsection contains ideas and definitions that are important for the proofs of Theorem 1.3 and 1.4, as well as the statement of another of this paper's main results, Theorem 1.8, which we will actually prove before Theorem 1.3. The reader who prefers to first see proofs of Theorems 1.1 and 1.2 and some discussion of the consequences may read Sections 3 and 4, as well as much of Section 5, independently of this subsection.

Theorem 1.8 includes a variant of Corollary 1.5 in which one parameterizes time by "amount of quantum length zipped up" instead of by capacity. The "stationary" picture will be described as a particular random quantum surface $\mathcal{S}$ with two marked boundary points and a chordal SLE η connecting the two marked points. The theorem will state that this η divides $\mathcal{S}$ into two quantum surfaces $\mathcal{S}_1$ and $\mathcal{S}_2$ that are *independent* of each other. (One can also reverse the procedure and first choose the $\mathcal{S}_i$—these are the so-called γ-quantum wedges mentioned earlier—and then weld them together to produce $\mathcal{S}$ and the interface η.) As we have already mentioned, this independence appears at first glance to be a rather bizarre coincidence. However, as we will see in Section 2.2, this kind of result is to be expected if SLE-decorated Liouville quantum gravity is (as conjectured) the scaling limit of path-decorated random planar maps.

Before we state Theorem 1.8 formally, we will need to spend a few paragraphs constructing a particular kind of scale invariant random quantum surface that we will call an "α quantum wedge." The reader who has never encountered quantum wedges before may wish to first read Section 1.4 of [23], which contains a more recent and better illustrated discussion of the quantum wedge construction.

We begin this construction by making a few general remarks. Recall that given any quantum surface represented by $(\widetilde{D}, \tilde{h})$—with two distinguished boundary points—we can change coordinates via (1.3) and represent it as the pair $(\mathbb{H}, h)$ for some h, where $\mathbb{H}$ is the upper half plane, and the two marked points are taken to be 0 and ∞. We will represent the "quantum wedges" we construct in this way, and we will focus on constructions in which there is almost surely a finite amount of μ_h mass and ν_h mass in each bounded neighborhood of 0 and an infinite amount in each neighborhood of ∞. In this case, the corresponding quantum surface is *half-plane-like* in the sense that it has one distinguished boundary point "at infinity" and one distinguished "origin"—and each neighborhood of "infinity" includes infinite

area and an infinite length portion of the surface boundary, while the complement of such a neighborhood contains only finite area and a finite-length portion of the surface boundary. We will let $\mathcal{S}_h$ denote the doubly marked quantum surface described by h in this way.

The h describing $\mathcal{S}_h$ is canonical except that we still have one free parameter corresponding to constant rescalings of $\mathbb{H}$ by (1.3). For each $a > 0$, such a rescaling is given by

$$(\mathbb{H}, h) \to \big(\mathbb{H}, h(a\cdot) + Q \log|a|\big). \tag{1.8}$$

We can fix this parameter by requiring that $\mu_h(B_1(0) \cap \mathbb{H}) = 1$. We will let μ_h be zero on the negative half plane so that we write this slightly more compactly as $\mu_h(B_1(0)) = 1$. (Alternatively, one could normalize so that $\nu_h([-1, 1]) = 1$.) We call the h for which this holds the *canonical description* of the doubly marked quantum surface.

Now to construct a "quantum wedge" it will suffice to give the law of the corresponding h. To this end, we first recall that one can decompose the Hilbert space for the free boundary GFF into an orthogonal sum of the space of functions which are radially symmetric about zero and the space of functions with zero mean about all circles centered at zero [29]. Consequently, we can write $h(\cdot) = h_{|\cdot|}(0) + h^\dagger(\cdot)$, where $h^\dagger_\varepsilon(0) = 0$ for all ε, and $h_{|z|}(0)$ is (of course) a continuous and radially symmetric function of z. This is a decomposition of the GFF h into its projection onto two $(\cdot, \cdot)_\nabla$ orthogonal subspaces, so $h_{|\cdot|}(0)$ and $h^\dagger(\cdot)$ are independent of each other [91]; the latter is a scale invariant random distribution and defined without an additive constant (since its mean is set to be zero on all circles centered at the origin). Now we define three types of quantum surfaces (the first two being defined only up an additive constant for h, which corresponds to a constant-factor rescaling of the surface itself). The third may seem unmotivated; however, the reader may note that it is similar in the spirit to the second, except that the third h is actually a well-defined random distribution (as opposed to a random distribution modulo additive constant), so that $(\mathbb{H}, h)$ is a well-defined quantum surface.

1. *Definition—unscaled quantum wedge on* $\mathbb{H}$: The quantum surface $(\mathbb{H}, h)$ where h is an instance of the free boundary GFF (which is defined up to additive constant, so that the quantum surface is defined only up to rescaling). In this case, $h_{|\cdot|}$ agrees in law with $B_{-\log|\cdot|}$ when $B_t, t \in \mathbb{R}$ is $\sqrt{2}$ times a standard Brownian motion defined up to a global additive constant. We think of B_t as a Brownian motion with diffusive rate 2, which will be understood throughout the discussion below. We can write

$$h = h^\dagger(\cdot) + B_{-\log|\cdot|},$$

where $h^\dagger(\cdot)$ and $B_{-\log|\cdot|}$ are independent.

2. *Definition—α-log-singular free quantum surface on* $\mathbb{H}$: The quantum surface $(\mathbb{H}, h)$ where

$$h = h^{\dagger}(\cdot) + \alpha\big(-\log|\cdot|\big) + B_{-\log|\cdot|}, \tag{1.9}$$

with $h^{\dagger}$ and B as above (and h also defined only up to additive constant).

3. *Definition—α-quantum wedge*: for $\alpha < 0$, the quantum surface $(\mathbb{H}, h)$ where

$$h = h^{\dagger}(\cdot) + Q\big(-\log|\cdot|\big) + A_{-\log|\cdot|}, \tag{1.10}$$

and the process $A_t, t \in \mathbb{R}$ is defined in a particular way: namely, for $t \geq 0$, A_t is a Brownian motion with drift $\alpha - Q$, that is, $A_t = B_t + (\alpha - Q)t$. Also, for $t \geq 0$, the negative-time process A_{-t} is chosen independently as a Brownian motion with drift $-(\alpha - Q)$ *conditioned* not to revisit zero. This involves conditioning on a probability zero event, so let us state this another way to be clear. Note that $\widetilde{B}_t = B_t - (\alpha - Q)t$ has positive drift, and hence a.s. $s_0 = \sup\{s : \widetilde{B}_s = 0\} < \infty$. Then the law of A_{-t} (for $t \geq 0$) is the law of $\widetilde{B}_{t+s_0}$, for $t \geq 0$.

To begin to motivate the definition above, note that applying the coordinate transformation (1.8) to the α-quantum wedge defined by (1.10), where the coordinate change map is a rescaling by a factor of a, amounts to replacing (1.10) with

$$\begin{aligned} &h^{\dagger}(a\cdot) + Q\big(-\log|a\cdot|\big) + A_{-\log|a\cdot|} + Q\log|a| \\ &\quad = h^{\dagger}(a\cdot) + Q\big(-\log|\cdot|\big) + A_{\log a - \log|\cdot|}. \end{aligned}$$

Since the law of $h^{\dagger}$ is scaling invariant, we find that the coordinate change described amounts to a horizontal translation of A by $-\log a$. That is, the quantum surface obtained by sampling A and then sampling $h^{\dagger}$ independently agrees in law with the quantum surface obtained by sampling A, translating the graph of A horizontally by some (possibly random) amount, and then sampling $h^{\dagger}$ independently.

We think of A_t as a Brownian process that drifts steadily as a Brownian motion with drift $(\alpha - Q)$ from $-\infty$, reaches zero at some point, and then subsequently evolves as a regular Brownian motion with the same drift. Since translating the graph of A_t horizontally does not affect the law of the quantum surface obtained, we choose (for concreteness) the translation for which $\inf\{t : A_t = 0\} = 0$. (We remark that the process A_t can also be interpreted as the log of a Bessel process, reparameterized by quadratic variation, noting that the graph of such a reparameterization is *a priori* only defined up to a horizontal translation; this point of view is explained and used extensively in [23].)

Now we make another simple claim: the α-quantum wedge is a doubly marked quantum surface whose law is invariant under the multiplication of its area by a constant. To explain what this means, let us observe that when $C \in \mathbb{R}$, we can "multiply the surface area by the constant e^C" by replacing h with $h + C/\gamma$, or equivalently, by replacing A with $A + C/\gamma$. Let $t_0 = \inf\{t : \widetilde{A}_t = 0\}$ and write $\widetilde{A}_t = A_{t_0+t} + C/\gamma$. By the definition of t_0, we find that $\widetilde{A}_t$ (like A_t) is a process that drifts

up from $-\infty$, reaches zero for the first time when $t = 0$, and then subsequently evolves as a Brownian motion with drift. Indeed, it is not hard to see that $\widetilde{A}_t$ has the same law as A_t. To deduce the claim, we then observe that the distribution of $h^\dagger$ is fixed; and since the radial parts $h_{|\cdot|}(0)$ of the GFF are continuous and independent of $\mu_{h^\dagger}$ and converge to a limit in law, we may conclude that $e^{\gamma h_{|\cdot|}(0)}\, d\mu_{h^\dagger}$ converges in law.

For future reference, we mention that one has a natural notion of "convergence" for quantum surfaces of this type: if $h^1, h^2, \ldots$ are the canonical descriptions of a sequence of doubly marked quantum surfaces and h is the canonical description of $\mathcal{S}_h$, then we say that the sequence $\mathcal{S}_{h^i}$ *converges* to $\mathcal{S}_h$ if the corresponding measures μ_{h^i} converge weakly to μ_h on all bounded subsets of $\mathbb{H}$.

One motivation for the definition of a quantum wedge is the following, which can be deduced from the description of quantum typical points given in Section 6 of [29]. It says (in a certain special setting; for a stronger result, see Proposition 5.5) that if one zooms in near a "quantum-boundary-measure-typical" point, one finds that the quantum surface looks like a γ-quantum wedge near that point.

PROPOSITION 1.6. *Fix $\gamma \in [0, 2)$ and let D be a bounded subdomain of $\mathbb{H}$ for which $\partial D \cap \mathbb{R}$ is a segment of positive length. Let $\tilde{h}$ be an instance of the GFF with zero boundary conditions on $\partial D \setminus \mathbb{R}$ and free boundary conditions on $\partial D \cap \mathbb{R}$. Let $[a, b]$ be any subinterval of $\partial D \cap \mathbb{R}$ and let $\mathfrak{h}_0$ be a continuous function on D that extends continuously to the interval (a, b). Let dh be the law of $\mathfrak{h}_0 + \tilde{h}$, and let $\nu_h[a, b]\, dh$ denote the measure whose Radon–Nikodym derivative w.r.t. dh is $\nu_h[a, b]$. (Assume that this is a finite measure, i.e., the dh expectation of $\nu_h[a, b]$ is finite.) Now suppose we*:

1. *sample h from $\nu_h[a, b]\, dh$ (normalized to be a probability measure),*
2. *then sample x uniformly from ν_h restricted to $[a, b]$ (see Figure 7) (normalized to be a probability measure),*
3. *and then let h^* be h translated by $-x$ units horizontally (i.e., recentered so that x becomes the origin).*

Then as $C \to \infty$ the random quantum surfaces $\mathcal{S}_{h^*+C/\gamma}$ converge in law (w.r.t. the topology of convergence of doubly-marked quantum surfaces) to a γ-quantum wedge.

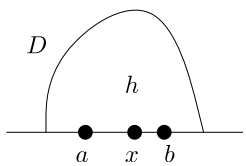

FIG. 7. *Point x sampled from ν_h (restricted to $[a, b]$).*

PROOF. We first recall that in this setting the description of quantum typical points in Section 6 of [29] implies a very explicit description of the joint law of the pair x and h sampled in Proposition 1.6. The marginal law of x is absolutely continuous with respect to Lebesgue measure, and conditioned on x the law of h is that of its original law *plus* a deterministic function that has the form $-\gamma \log|x - \cdot|$ plus a deterministic smooth function. In a small neighborhood of x, this deterministic smooth function is approximately constant, which means that h^* looks like (up to additive constant) the h used to define an α-log-singular free quantum surface in (1.9), with $\alpha = \gamma$. If we write $A'_t = B_t + (\alpha - Q)t$, then we find that h^* looks like the h used to define a γ-quantum wedge in (1.10), except with A replaced by A'.

Now replacing h^* by $h^* + C/\gamma$ corresponds to adding C/γ to the process B from (1.9), and hence also corresponds to adding C/γ to the process A', which translates the graph of A' vertically. Recall from above that translating the graph of A' horizontally corresponds to a coordinate change; so we can translate A' so that it hits zero for the first time at the origin. It is not hard to see that as $C \to \infty$, the law of A' thus translated converges to the law of A. Since the law of $h^\dagger$ is scale invariant and can be chosen independently, this implies the proposition statement. □

We will later show (see Proposition 5.5) that the conclusion of the proposition still holds if (when generating x and h) we condition on particular values for $\nu_h[a, x]$ and $\nu_h[x, b]$.

The following is an immediate consequence of Proposition 1.6. It tells us that the γ-quantum wedge is stationary with respect to shifting the origin by a given amount of quantum length. (When $\gamma = 0$, the proposition simply states that $\mathbb{H}$ itself is invariant under horizontal translations. Proposition 1.7 is the general quantum analog of this invariance.)

PROPOSITION 1.7. *Fix a constant $L > 0$. Suppose that $(\mathbb{H}, h)$ is a γ-quantum wedge. Then choose $y > 0$ so that $\nu_h[0, y] = L$, and let h^* be h translated by $-y$ units horizontally (i.e., recentered so that y becomes the origin). Then $(\mathbb{H}, h^*)$ is a γ-quantum wedge.*

PROOF. Suppose that x is the point chosen uniformly from the quantum boundary measure in Proposition 1.6, and x' is the point translated δL quantum length units to the right from x, so that $\nu_h[x, x'] = \delta L$. Note that such an x' exists with a probability that tends to 1 as $\delta \to 0$, and that the law of x' converges (in total variation sense) to the law of x as $\delta \to 0$. In the rescaled surfaces in Proposition 1.6, boundary lengths are scaled by $e^{C/2}$, so if we set $\delta = e^{-C/2}$, then the distance between x and x' is L after the rescaling. Since this δ tends to zero as $C \to \infty$ we conclude that the limiting surface law is (as desired) invariant under the operation that translates the origin by L units of quantum boundary length. □

The following will be proved in Section 5.

THEOREM 1.8. Wedge decomposition: *Fix $\gamma \in (0,2)$, and let $\mathcal{S}$ be a $(\gamma - 2/\gamma)$-quantum wedge with canonical description h. Let η be a chordal SLE_κ in $\mathbb{H}$ from 0 to ∞, with $\kappa = \gamma^2$, chosen independently of h. Let D_1 and D_2 be the left and right components of $\mathbb{H} \setminus \eta$, and let h^{D_1} and h^{D_2} be the restrictions of h to these domains. Then the quantum surfaces represented by (D_1, h^{D_1}) and (D_2, h^{D_2}) are* independent *γ-quantum wedges (marked at 0 and ∞), and their quantum boundary lengths along η agree.*

Zipper stationarity: *Moreover, suppose we define*

$$\mathcal{Z}_{-t}^{\mathrm{LEN}}\big((D_1, h^{D_1}), (D_2, h^{D_2})\big)$$

as follows. First, find z on η for which the quantum boundary lengths along D_1 and D_2 (which are well defined by unzipping) along η between 0 and z are both equal to t. Let t' be the time that η hits z (when η is parameterized by capacity) and define

$$\mathcal{Z}_{-t}^{\mathrm{LEN}}\big((D_1, h^{D_1}), (D_2, h^{D_2})\big) = \text{rescaling of } \big(f_{t'}^{\eta}(D_1, h^{D_1}), f_{t'}^{\eta}(D_2, h^{D_2})\big),$$

where the rescaling is done via (1.8) with the parameter a chosen so that $B_1(0)$ has area one in the transformed quantum measure. Then the following hold:

1. *The inverse $\mathcal{Z}_t^{\mathrm{LEN}}$ of the operation $\mathcal{Z}_{-t}^{\mathrm{LEN}}$ is a.s. uniquely defined (via conformal welding).*
2. *$\mathcal{Z}_{s+t}^{\mathrm{LEN}} = \mathcal{Z}_s^{\mathrm{LEN}} \mathcal{Z}_t^{\mathrm{LEN}}$ almost surely for $s, t \in \mathbb{R}$.*
3. *The law of the pair $((D_1, h^{D_1}), (D_2, h^{D_2}))$ is invariant under $\mathcal{Z}_t^{\mathrm{LEN}}$ for all $t \in \mathbb{R}$.*

It also follows from Theorem 1.4 and the subsequent discussion that the two independent γ-quantum wedges uniquely determine h and η almost surely. We refer to the group of transformations $\mathcal{Z}_t^{\mathrm{LEN}}$ as the *length quantum zipper*. When $t > 0$, applying $\mathcal{Z}_t^{\mathrm{LEN}}$ is called "zipping up" the pair of quantum surfaces by t quantum length units and applying $\mathcal{Z}_{-t}^{\mathrm{LEN}}$ is called "zipping down" or "unzipping" by t quantum length units. When we defined the operations $\mathcal{Z}_t^{\mathrm{CAP}}$, h was defined only up to additive constant, and the zipping maps f_t were independent of that constant. By contrast, $\mathcal{Z}_t^{\mathrm{LEN}}$ represents zipping by an actual quantity of quantum length, and hence cannot be defined without the additive constant being fixed.

We will give a detailed proof in Section 5, which is in some sense the heart of the paper. But for now, let us give a brief overview of the proof and the relationship to our other results. We will start with the scenario described in Figure 6, with h normalized to have mean zero on $\partial B_1(0)$, except that the measure dh on h is replaced by the probability measure whose Radon–Nikodym derivative w.r.t. dh is $\nu_h(-\delta, 0)$ for some fixed δ (see Figure 8).

Then we will sample x from ν_h restricted to $(-\delta, 0)$ (normalized to be a probability measure) and "zip up" until x hits the origin (to obtain a "quantum-length-typical" configuration). We then zoom in near the origin (multiplying the area by

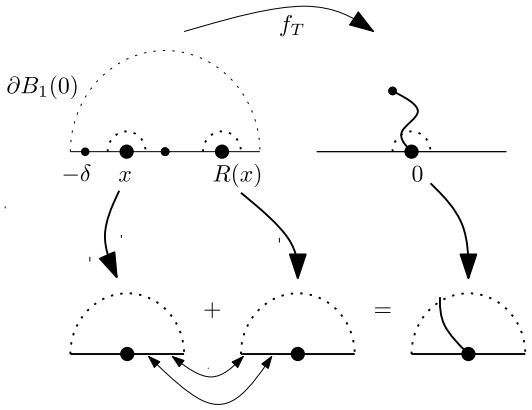

FIG. 8. *Choose h as in Theorem 1.2 [normalized by $h_1(0) = 0$] except with the law of h* weighted *by $\nu_h([-\delta, 0])$ for some fixed $\delta \in (0, 1)$. Then conditioned on h, sample x from ν_h restricted to $[-\delta, 0]$ (normalized to be a probability measure). Take T so that f_T is the map zipping up $[x, 0]$ with $[0, R(x)]$. Consider the three random surfaces obtained by choosing a semi-disc of quantum area $\tilde{\varepsilon}$ centered at each of x and $R(x)$ (on the left-hand side) and 0 (on the right side), and multiplying areas by $1/\tilde{\varepsilon}$ (zooming in) so that all three balls have quantum area 1. In the $\tilde{\varepsilon} \to 0$ limit, the left two quantum surfaces become independent γ-quantum wedges, and the right is the conformal welding of these two.*

$\tilde{\varepsilon}^{-1}$—and hence the boundary length by $\tilde{\varepsilon}^{-1/2}$—say). We then use a variant of Proposition 1.6 (namely, Proposition 5.5) to show that (in the $\tilde{\varepsilon} \to 0$ limit) the lower two rescaled surfaces on the lower left of Figure 8 become independent γ-quantum wedges.

The fact that the curve on the right in Figure 8 is (in the $\tilde{\varepsilon} \to 0$ limit) an SLE_κ independent of the canonical description h on the right will be shown in Section 5 by directly calculating the law of the process that "zips up" $[x, 0]$ with $[0, R(x)]$. It could also be seen by showing that we can construct an equivalent pair of glued surfaces by beginning with Figure 6 [with h normalized to have mean zero on $\partial B_1(0)$] and then zipping *down* by a random amount (chosen uniformly from an interval) of quantum length, then zooming in by multiplying lengths by $1/\tilde{\varepsilon}$, and then taking the $\tilde{\varepsilon} \to 0$ limiting law. (In this case, the domain Markov property of the original SLE, and its independence from the original GFF, would imply that the conditional law of the still-zipped portion of the curve is an SLE_κ, independent of h.)

Similar arguments to those in [29] will show that the procedure in Figure 8 produces a configuration related to the one in Figure 5 except that it is in some sense *weighted* by the amount of quantum mass near zero. It will turn out that this weighting effectively adds $-\gamma \log|\cdot|$ to the $\mathfrak{h}_0$ of Theorem 1.2 and Corollary 1.5. This is why Theorem 1.8 involves a $(\gamma - 2/\gamma)$-quantum wedge, instead of a $(-2/\gamma)$-quantum wedge, as one might initially guess based on Theorem 1.2. Once we have all of this structure in place, the really crucial step will be showing

that parameterizing time by the amount of "left boundary quantum length" zipped up yields the same stationary picture as parameterizing time the amount of "right boundary quantum length" zipped up. Given this, we will then use the ergodic theorem to show that over the long term, the amount of left bounday quantum length zipped up approximately agrees with the amount of right boundary length zipped up. Using scale invariance symmetries, we will then deduce that this agreement almost surely holds exactly on all scales.

1.7. *Organization*. Section 2 provides heuristic justification and motivation for the main results about AC geometry and Liouville quantum gravity. (An interpretation of AC geometry in terms of "imaginary curvature" appears in the Appendix.) Section 3 gives a brief overview of the Gaussian free field. Section 4 proves Theorems 1.1 and 1.2, along with a generalization to other underlying geometries. Section 5 proves Theorems 1.8 and 1.3 (in that order), along with some additional results about zipping processes and time changes. (Recall that we have already proved that Theorem 1.4 is a consequence of Theorem 1.3.) Section 6, finally, presents a list of open problems and conjectures.

2. Geometric interpretation. This section summarizes some of the conjectures and intuition behind our main results, including some discrete-model-based reasons that one would expect the coupling and welding theorems to be true. This section may be skipped by the reader who prefers to proceed directly to the proofs.

2.1. *Forward coupling*: *Flow lines of* $e^{ih/\chi}$. Fix a planar domain D, viewed as a subset of $\mathbb{C}$, a function $h : D \to \mathbb{R}$, and a constant $\chi > 0$. An *AC ray* of h is a flow line of the complex vector field $e^{ih/\chi}$ beginning at a point $x \in \overline{D}$, that is, a path $\eta : [0, \infty) \to \mathbb{C}$ that is a solution to the ODE:

$$\eta'(t) := \frac{\partial}{\partial t}\eta(t) = e^{ih(\eta(t))/\chi} \qquad \text{when } t > 0, \eta(0) = x, \tag{2.1}$$

until time $T = \inf\{t > 0 : \eta(t) \notin D\}$. When h is Lipschitz, the standard Picard–Lindelöf theorem implies that if $x \in D$, then (2.1) has a unique solution up until time T (and T is itself uniquely determined). The reader can visually follow the flow lines in Figure 9.

If h is continuous, then the time derivative $\eta'(t)$ moves continuously around the unit circle, and $h(\eta(t)) - h(\eta(0))$ describes the net amount of *winding* of η' around the circle between times 0 and t.

A major problem (addressed in depth in an imaginary geometry series [64–67]) is to make sense of these flow lines when h is a multiple of the Gaussian free field. We will give here just a short overview of the way these objects are constructed. Suppose that η is a smooth simple path in $\mathbb{H}$ beginning at the origin, with (forward) Loewner map $f_t = f_t^{\eta}$. We may assume that η starts out in the vertical direction, so that the winding number is $\pi/2$ for small times. Then when η and h are both

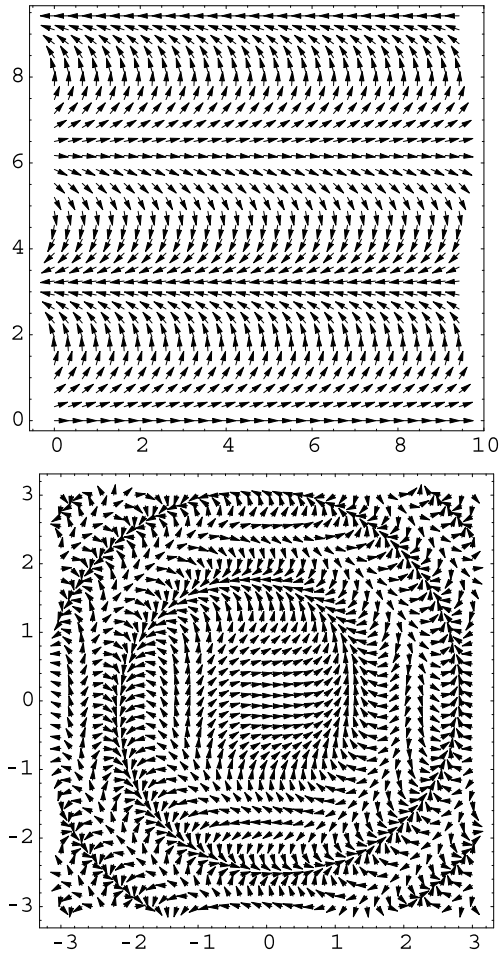

FIG. 9. *The complex vector flow* e^{ih}: $h(x, y) = y$, $h(x, y) = x^2 + y^2$.

smooth, the statement that η is a flow line of $e^{ih/\chi}$ is equivalent to the statement that for each x on $\eta((0, t))$ we have

$$\chi \arg f_t'(z) \to -h(x) \tag{2.2}$$

as z approaches x from the left side of η and

$$\chi \arg f_t'(z) \to -h(x) + \chi\pi \tag{2.3}$$

as z approaches x from the right side of η (as Figure 10 illustrates). Recall that $\arg f_t'(z)$—a priori determined only up to a multiple of 2π—is chosen to be continuous on $\mathbb{H} \setminus \eta([0, t])$ and 0 on $\mathbb{R}$. If $\chi = 0$, then (2.2) and (2.3) hold if and only if h is identically zero along the path η, that is, η is a zero-height *contour line* of h.

In [86, 87], it is shown that when one takes certain approximations h^ε of an instance h of the GFF that are piecewise linear on an ε-edge-length triangular

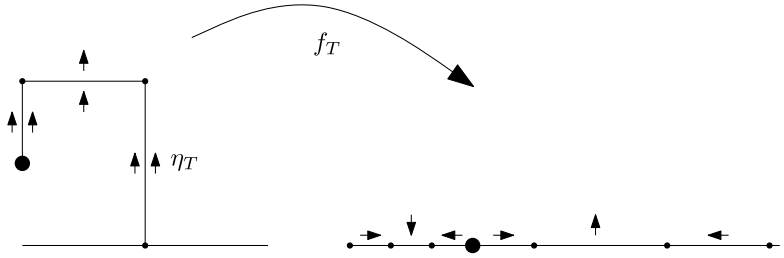

FIG. 10. *Winding number along η_T determines* $\arg f_T'$, *which is the amount a small arrow near η_T is rotated by f_T.*

mesh, then conditioned on a zero chordal contour line of h^ε there is in some $\varepsilon \to 0$ limiting sense a constant "height gap" between the expected heights immediately to one side of the contour line and those heights on the other. We might similarly expect that if one looked at the expectation of h^ε, given a chordal flow line η^ε of $e^{ih^\varepsilon/\chi}$, there would be a constant order limiting height gap between the two sides; see Figure 11.

This suggests the form of $\mathfrak{h}_t$ given in Theorem 1.1, which comes from taking (2.2) and (2.3) and modifying the height gap between the two sides by adding a multiple of $\arg f_t$. (As in [87], the size of the height gap—and hence the coefficient of $\arg f_t$ in the definition of $\mathfrak{h}_t$—is determined by the requirement that $\mathfrak{h}_t(z)$ be a martingale in t; see Section 4.) Interestingly, the fact that winding may be ill-defined at a particular point on a fractal curve turns out to be immaterial. It is the harmonic extension of the boundary winding values (the $\arg f_t'$) that is needed to define $\mathfrak{h}_t$, and this is defined even for nonsmooth curves.

The time-reversal of a flow line of $e^{ih^\varepsilon/\chi}$ is a flow line of $e^{i(h^\varepsilon/\chi+\pi)}$, which at first glance appears to imply that there should not be a height gap between the two sides (since if the left side were consistently smaller for the forward path, then the right side would be consistently smaller for the reverse path). To counter this intuition, observe that, in the left diagram in Figure 9, the left-going infinite

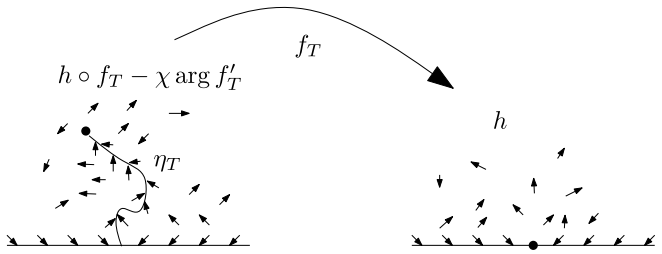

FIG. 11. *Forward coupling with arrows in $e^{ih/\chi}$ direction* (*sketch*), *illustrating the constant angle gap between the two sides of the curve η, constant angles along the positive and negative real axes, and random angles* (*not actually point-wise defined if h is the GFF*) *in $\mathbb{H} \setminus \eta$.*

horizontal flow lines (at vertical heights of $k\pi$, k odd) are "stable" in that the flow line beginning at a generic point slightly off one of these lines will quickly converge to the line. The right-going horizontal flow lines (at heights $k\pi$, k even) are unstable. In a stable flow line, h appears to generally be larger to the right side of the flow line and smaller to the left side. It is reasonable to expect that a flow line of $e^{ih^\varepsilon/\chi}$ started from a generic point would be approximately stable in that direction—and in particular would look qualitatively different from the time reversal of a flow line of $e^{i(h^\varepsilon/\chi+\pi)}$ started from a generic point.

2.2. *Reverse coupling*: *Planar maps and scaling limits*. In this section, we conjecture a connection between path-decorated planar maps and SLE-decorated Liouville quantum gravity (in particular, the quantum-length-invariant decorated quantum wedge of Theorem 1.8). We will explain the details in just one example based on the uniform spanning tree. (Variants based on Ising and $O(n)$ and FK models on random planar maps—or on random planar maps without additional decoration besides the chordal paths—are also possible. Many rigorous results for percolation and the Ising model have been obtained for deterministic graphs in [14, 16, 97–100] (and in many other papers we will not survey here), and one could hope to extend these results to random graphs. One could also consider discrete random surfaces decorated by loops and in the continuum replace SLE decorations with CLE decorations [92, 94].) As mentioned earlier, we will see that the more surprising elements of Theorem 1.8 are actually quite natural from the discrete random surface point of view.

Let G be a planar map with exactly n edges (except that each edge on the outer face is counted as *half* an edge) and let T be a subgraph consisting of a single boundary cycle, a chordal path from one boundary vertex a to another boundary vertex b that otherwise does not hit the boundary cycle, and a spanning forest rooted at this "Figure 8" structure. (See Figure 12.) Here, T is like the wired spanning tree (in which the entire boundary is considered to be one vertex), except that

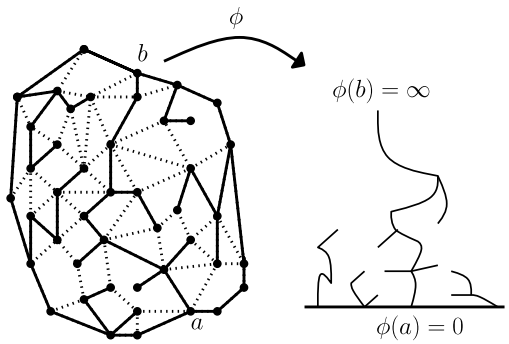

FIG. 12. *Planar map with a distinguished outer-boundary-plus-one-chord-rooted spanning tree* (*solid black edges*), *with chord joining marked boundary points* a *and* b, *plus image of tree under conformally uniformizing map* ϕ *to* $\mathbb{H}$ (*sketch*).

there is also one chord connecting a pair of boundary vertices. What happens if we consider the uniform measure on all pairs (G, T) of this type? This model is fairly well understood combinatorially (tree-rooted maps on the sphere are in bijective correspondence with certain walks in $\mathbb{Z}^2$ — see, e.g., [73] as well as [12] and the references therein—and our model is a simple variant of this) and in particular, it follows from these bijections that the length of the boundary of the outer face of this map will be of order $\sqrt{n}$ with high probability when n is large. Now, can we understand the scaling limit of the random pair (G, T) as $n \to \infty$?

There are various ways to pose this problem. For example, one could consider G as a metric space and aim for convergence in law w.r.t. the Gromov–Hausdorff metric on metric spaces. The reader is probably aware that there is a sizable literature on the realization of a random metric space called the Brownian map as a Gromov–Hausdorff scaling limit of random planar maps of various types. However, since this paper is concerned with the conformal structure of random geometry, we will try to phrase the problem in a way that keeps track of that structure.

First, we would like to understand how to conformally map the planar map to the half plane, as in Figure 12. We may consider G as embedded in a two-dimensional manifold with boundary in various ways, one of which we sketch here: first add an interior vertex to each face of G and an edge joining it to each vertex of that face (as in Figure 13). Each interior edge of G is now part of a quadrilateral (containing one vertex for each interior face of G and one for each vertex of G) and we will endow that quadrilateral with the metric of a unit square $[0, 1] \times [0, 1]$. Similarly, the triangle containing an exterior edge of G is endowed with the metric of half a unit square (split on its diagonal, with the exterior edge as the hypotenuse). When two squares or half squares share an edge, the points along that edge are identified with one another in a length preserving way. We may view the collection of (whole and half) unit squares, glued together along boundaries, as a manifold (with isolated conical singularities at vertices whose number of incident squares is not four) with a uniquely defined conformal structure (note that it is trivial to define a Brownian motion on the manifold, since it a.s. never hits the singularities). We may choose a conformal map ϕ from this manifold to $\mathbb{H}$, sending a to 0 and b to ∞, as sketched in Figure 12.

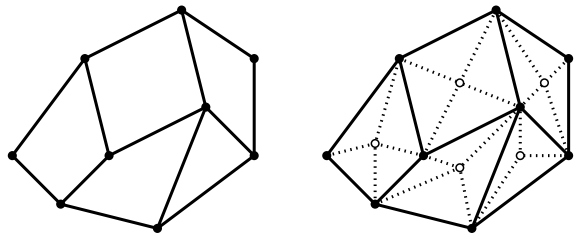

FIG. 13. *An arbitrary planar map can be used to construct a collection of stitched-together unit squares and half unit squares. The result is viewed as a two-dimensional manifold with boundary.*

This ϕ is determined only up to scaling, but we can fix the scaling in many ways. We will do so by considering a number $k < n$ and requiring that the area of $\phi^{-1}(B_1(0))$ be equal to k. Then ϕ determines a random measure on $\mathbb{H}$ (the image of the area measure on the manifold) in which the measure of $B_1(0)$ is deterministically equal to k; let $\mu_{n,k}$ denote this random measure divided by k, so that $\mu_{n,k}(B_1(0)) = 1$. We expect that if one lets n and k tend to ∞ in such a way that n/k tends to ∞, then the random measures $\mu_{n,k}$ will converge in law with respect to the metric of weak convergence on bounded subsets of $\mathbb{H}$ to the $\mu = \mu_h$ corresponding to the canonical description h of the $(\gamma - 2/\gamma)$-quantum wedge of Theorem 1.8. [By compactness, the laws of the $\mu_{n,k}$ restricted to the closure of $B_1(0)$ have at least a subsequential limit.] We similarly conjecture that $\nu_{n,k}$—defined to be $1/\sqrt{k}$ times the image of the manifold's boundary measure—will converge in law to the corresponding ν_h. (We remark that one could alternatively formulate the conjecture by taking an infinite volume limit first, that is, letting n go to infinity while keeping k constant to define a limiting measure $\mu_{\infty,k} := \lim_{n\to\infty} \mu_{n,k}$. This kind of infinite volume limit of random planar maps was constructed in [3]. One can subsequently take $k \to \infty$ and conjecture that the limit is μ_h. A similar conjecture in [29] was formulated in terms of infinite volume limits.)

We are currently unable to prove these conjectures, but related questions about Brownian motion on random surfaces have been explored in [32], where it was shown that certain infinite random triangulations and quadrangulations (*without* boundaries) are parabolic (as opposed to hyperbolic) Riemann surfaces [32]. (This is equivalent to showing that a Brownian motion visits each face infinitely often almost surely; see analogous discrete results in [3].)

Now let us make some more observations. If we take k, n, and n/k to be large and condition on G, a, and b, then what is the conditional law of $\phi(T)$, as depicted in Figure 12? The conditional law of T itself is uniform among all valid 8-rooted spanning forest configurations. The physics literature frequently invokes a kind of "conformal invariance Ansatz" which suggests that this random path (and many other random sets in critical two-dimensional statistical physics) should be a conformally invariant object.

In this case, we claim that the law of the chordal path should be approximately that of a chordal SLE_2 *even after we have conditioned on G, a, and b, which determine the measure $\mu_{n,k}$*. The reason for our claim is that a related SLE_2 convergence result is obtained in [50] in the case that G is a deterministic lattice graph, and this was generalized substantially in [105] where it was shown that if a graph can be embedded in the plane in such a way that simple random walk approximates Brownian motion, then the uniform spanning tree paths approximate a form of SLE_2. We do not know whether the hypotheses of [105] hold in our setting. Brownian motion is conformally invariant, but it is not clear whether simple random walk on our random G approximates Brownian motion on the corresponding quadrangulated manifold with high probability. However, it seems very natural to conjecture that the hypotheses hold. In any case, we stress the following: if our scaling limit

conjecture holds, then the asymptotic independence of the chordal path from $\mu_{n,k}$ would be consistent with the independence of η and h in Theorem 1.8.

Next let D_1 and D_2 be the wired-spanning-tree decorated manifolds to the left and right of the chordal path. Note that once we condition on the length of the chordal path in (G, T) and the number of edges on each side of it, the laws of D_1 and D_2 are independent of one another. We might guess that the local behavior of D_1 and D_2 near a would be approximately independent of these global numbers. We expect a similar property to hold in the scaling limit, which would be consistent with the independence of the left and right quantum surfaces described in Theorem 1.8. (The idea of gluing together independent discrete surfaces in this manner has been explored in many works by Duplantier and others, beginning perhaps in [22]. The idea of gluing a whole series of discrete surfaces was used in [21] to heuristically derive certain "cascade relations" via the KPZ formula.)

Finally, if we condition on the point b and on D_1 and D_2, then the length of the path along which D_1 and D_2 are glued to each other is uniform among all possibilities (which range between 1 and the minimum M of the boundary lengths of the two D_i's minus 1). In other words, once D_1 and D_2 and b are all fixed, we can randomly decide how far to "zip up" or "unzip" these two surfaces (moving the vertex a accordingly). If r is the random number of steps we zip, then r and $r + m$ have approximately the same law (as long as m/M is small). We expect a similar property to hold in the scaling limit, which would be consistent with the quantum-length-zipper invariance described in Theorem 1.8.

3. Gaussian free field overview. We refer the reader to [91] for a survey of the Gaussian free field (GFF) and several additional references. For completeness, we include a short overview, closely following [87, 91]. For the reader who is already familiar with the zero and free boundary GFF, it may be sufficient (to set notation) to read only the numbered equations in this section and the statement of Proposition 3.1.

3.1. *GFF definitions.*

3.1.1. *Dirichlet inner product.* Fix a simply connected planar domain $D \subset \mathbb{C}$ (with $D \neq \mathbb{C}$). Let $H_s(D)$ be the space of smooth, compactly supported functions on D, and let $H(D)$ [sometimes denoted by $\mathbb{H}_0^1(D)$ or $W^{1,2}(D)$] be its Hilbert space closure under the Dirichlet inner product

$$(f_1, f_2)_\nabla := (2\pi)^{-1} \int_D \nabla f_1(z) \cdot \nabla f_2(z)\, dz.$$

Let ψ be a conformal map from another domain $\widetilde{D}$ to D. Then an elementary change of variables calculation shows that

$$\int_{\widetilde{D}} \nabla(f_1 \circ \psi) \cdot \nabla(f_2 \circ \psi)\, dx = \int_D (\nabla f_1 \cdot \nabla f_2)\, dx.$$

In other words, the Dirichlet inner product is invariant under conformal transformations.

We write $(f_1, f_2) = \int_D f_1(x) f_2(x)\, dx$ for the L^2 inner product on D. We write $\|f\| := (f, f)^{1/2}$ and $\|f\|_\nabla := (f, f)_\nabla^{1/2}$. If $f_1, f_2 \in H_s(D)$, then integration by parts gives

$$(f_1, f_2)_\nabla = \frac{1}{2\pi}(f_1, -\Delta f_2). \tag{3.1}$$

3.1.2. *Distributions and the Laplacian.* It is conventional to use $H_s(D)$ as a space of test functions. This space is a topological vector space in which the topology is defined so that $\phi_k \to 0$ in $H_s(D)$ if and only if there is a compact set on which all of the ϕ_k are supported and the mth derivative of ϕ_k converges uniformly to zero for each integer $m \geq 1$.

A *distribution* on D is a continuous linear functional on $H_s(D)$. Since $H_s(D) \subset L^2(D)$, we may view every $h \in L^2(D)$ as a distribution $\rho \mapsto (h, \rho)$. A *modulo-additive-constant* distribution on D is a continuous linear functional on the subspace of $H_s(D)$ consisting of ρ for which $\int_D \rho(z)\, dz = 0$. We will frequently abuse notation and use h—or more precisely the map denoted by $\rho \to (h, \rho)$—to represent a general distribution (which is a functional of ρ), even though h may not correspond to an element of $L^2(D)$. [Later, we will further abuse notation and use ρ to represent a nonsmooth function or a measure; in the latter case (h, ρ), when defined, will represent the integral of h against that measure.]

We define partial derivatives and integrals of distributions in the usual way (via integration by parts), that is, for $\rho \in H_s(D)$,

$$\left(\frac{\partial}{\partial x} h, \rho\right) := -\left(h, \frac{\partial}{\partial x}\rho\right).$$

In particular, if h is a distribution then Δh is a distribution defined by $(\Delta h, \rho) := (h, \Delta\rho)$. When h is a distribution and $\rho \in H_s(D)$, we also write

$$(h, \rho)_\nabla := \frac{1}{2\pi}(-\Delta h, \rho) = \frac{1}{2\pi}(h, -\Delta\rho).$$

When $x \in D$ is fixed, we let $\widetilde{G}_x(y)$ be the harmonic extension to $y \in D$ of the function of y on ∂D given by $-\log|y - x|$. Then *Green's function in the domain* D is defined by

$$G(x, y) = -\log|y - x| - \widetilde{G}_x(y).$$

When $x \in D$ is fixed, Green's function may be viewed as a distributional solution of $\Delta G(x, \cdot) = -2\pi\delta_x(\cdot)$ with zero boundary conditions [91]. It is nonnegative for all $x, y \in D$ and $G(x, y) = G(y, x)$.

For any $\rho \in H_s(D)$, we write

$$-\Delta^{-1}\rho := \frac{1}{2\pi}\int_D G(\cdot, y)\rho(y)\, dy.$$

This is a C^∞ function in D whose Laplacian is $-\rho$. Indeed, a similar definition can be made if ρ is any signed measure (with finite positive and finite negative mass) rather than a smooth function. Recalling (3.1), if $f_1 = -2\pi\Delta^{-1}\rho_1$ then $(h, f_1)_\nabla = (h, \rho_1)$, and similarly if $f_2 = -2\pi\Delta^{-1}\rho_2$. Now $(f_1, f_2)_\nabla = (\rho_1, -2\pi\Delta^{-1}\rho_2)$ describes a covariance that can (by the definition of $-\Delta^{-1}\rho_2$ above) be rewritten as

$$\operatorname{Cov}\big((h, \rho_1), (h, \rho_2)\big) = \int_{D\times D} \rho_1(x) G(x, y) \rho_2(y)\, dx\, dy. \tag{3.2}$$

If $\rho \in H_s(D)$, may define the map $(h, \cdot)$ by $(h, \rho) := (h, -2\pi\Delta^{-1}\rho)_\nabla$, and this definition describes a distribution [91]. [It is not hard to see that $-2\pi\Delta^{-1}\rho \in H(D)$, since its Dirichlet energy is given explicitly by (3.2).]

3.1.3. *Zero boundary GFF.* An instance of the GFF with zero boundary conditions on D is a random sum of the form $h = \sum_{j=1}^\infty \alpha_j f_j$ where the α_j are i.i.d. one-dimensional standard (unit variance, zero mean) real Gaussians and the f_j are an orthonormal basis for $H(D)$. This sum almost surely does not converge within $H(D)$ (since $\sum_{j=1}^\infty |\alpha_j|^2$ is a.s. infinite). However, it does converge almost surely within the space of distributions—that is, the limit $(\sum_{j=1}^\infty \alpha_j f_j, \rho)$ almost surely exists for all $\rho \in H_s(D)$, and the limiting value as a function of ρ is almost surely a continuous functional on $H_s(D)$ [91]. We may view h as a sample from the measure space $(\Omega, \mathcal{F})$ where $\Omega = \Omega_D$ is the set of distributions on D and $\mathcal{F}$ is the smallest σ-algebra that makes (h, ρ) measurable for each $\rho \in H_s(D)$, and we sometimes denote by dh the probability measure which is the law of h. If f_j are chosen in $H_s(D)$, then the values α_j are clearly $\mathcal{F}$-measurable. In fact, for any $f \in H(D)$ with $f = \sum_j \beta_j f_j$ the sum $(h, f)_\nabla := \sum_j \alpha_j \beta_j$ is a.s. well defined and is a Gaussian random variable with mean zero and variance $(f, f)_\nabla$.

3.2. *Green's functions on $\mathbb{C}$ and $\mathbb{H}$: Free boundary GFF.* The GFF with free boundary conditions is defined the same way as the GFF with zero boundary conditions except that we replace $H_s(D)$ with the space of all smooth functions with gradients in $L^2(D)$ (i.e., we remove the requirement that the functions be compactly supported). However, to make the correspondingly defined $H(D)$ a Hilbert space, we have to consider functions only modulo additive constants (since all constant functions have norm zero). On the whole plane $\mathbb{C}$, we may define the Dirichlet inner product on the Hilbert space closure $H(\mathbb{C})$ of the space of such functions defined modulo additive constants.

Generally, given a compactly supported ρ (or more generally, a signed measure), we can write

$$-\Delta^{-1}\rho(\cdot) := \frac{1}{2\pi}\int_{\mathbb{C}} G(\cdot, y)\rho(y)\, dy, \tag{3.3}$$

with $G(x, y) = -\log|x - y|$.

As before, for compactly supported f and g, we have $(f,g)_\nabla = \frac{1}{2\pi}(f,-\Delta g)$ by integration by parts, and moreover $(f,-\Delta^{-1}\rho)_\nabla = \frac{1}{2\pi}(\rho,f)$. The same holds for bounded and not necessarily compactly supported smooth functions f and g if the gradient of $-\Delta^{-1}\rho$ tends to zero at infinity, which in turn happens if and only if $\int_{\mathbb{C}} \rho(z)\,dz = 0$.

If $\int_{\mathbb{C}} \rho(z)\,dz \neq 0$, then the Dirichlet energy of $-\Delta^{-1}\rho$ will be infinite and moreover (h,ρ) will not be independent of the additive constant chosen for h. [If we view $\mathbb{C}$ as a Riemann sphere, then $\int_{\mathbb{C}} \rho(z)\,dz \neq 0$ can also be interpreted as the statement that the Laplacian of $-\Delta^{-1}\rho$ has a point mass at ∞.] When h is the free boundary GFF on $\mathbb{C}$, we will thus define the random variables (h,ρ) only if the integral of ρ over $\mathbb{C}$ is zero. If ρ_1 and ρ_2 each have total integral zero, we may write

$$\operatorname{Cov}\big((h,\rho_1),(h,\rho_2)\big) = \int_{\mathbb{C}\times\mathbb{C}} \rho_1(x)G(x,y)\rho_2(y)\,dx\,dy. \tag{3.4}$$

Using $z \to \bar{z}$ to denote complex conjugation, we define, for smooth functions $h \in H(\mathbb{C})$, the pair of projections

$$h^{\mathrm{O}}(z) := \frac{1}{\sqrt{2}}\big(h(z) - h(\bar{z})\big),$$

$$h^{\mathrm{E}}(z) := \frac{1}{\sqrt{2}}\big(h(z) + h(\bar{z})\big).$$

If h is an instance of the free boundary GFF on $\mathbb{C}$, we may still define h^{O} and h^{E} as projections of h onto complementary orthogonal subspaces. Their restrictions to $\mathbb{H}$ are instances of the zero boundary GFF and free boundary GFF, respectively, on $\mathbb{H}$. For ρ supported on $\mathbb{H}$ we write (for $z \in \mathbb{C}$) $\rho^*(z) := \rho(\bar{z})$. Then we have by definition

$$\big(h^{\mathrm{O}},\rho\big) = \frac{1}{\sqrt{2}}\big(h,\rho-\rho^*\big),$$

$$\big(h^{\mathrm{E}},\rho\big) = \frac{1}{\sqrt{2}}\big(h,\rho+\rho^*\big).$$

Note that (h^{E},ρ) is only defined if the total integral of ρ is zero, while (h^{O},ρ) is defined without that restriction (since in any case the total integral of $\rho-\rho^*$ will be zero).

For ρ_1 and ρ_2 supported on $\mathbb{H}$ we now compute the following (first integral taken over $\mathbb{C}\times\mathbb{C}$, second over $\mathbb{H}\times\mathbb{H}$):

$$\begin{aligned}
&\operatorname{Cov}\big(\big(h^{\mathrm{O}},\rho_1\big),\big(h^{\mathrm{O}},\rho_2\big)\big)\\
&\quad= \frac{1}{2}\int \big(\rho_1(x)-\rho_1^*(x)\big)\log|x-y|\big(\rho_2(y)-\rho_2^*(y)\big)\,dx\,dy\\
&\quad= \int \rho_1(x)G^{\mathbb{H}_0}(x,y)\rho_2(y)\,dx\,dy,
\end{aligned} \tag{3.5}$$

where $G^{\mathbb{H}_0}(x, y) := \log|x - \bar{y}| - \log|x - y|$. Similarly (first integral over $\mathbb{C} \times \mathbb{C}$, second over $\mathbb{H} \times \mathbb{H}$),

$$\begin{aligned} &\operatorname{Cov}\big((h^{\mathrm{E}}, \rho_1), (h^{\mathrm{E}}, \rho_2)\big) \\ &\quad = \frac{1}{2}\int \big(\rho_1(x) + \rho_1^*(x)\big)\log|x - y|\big(\rho_2(y) + \rho_2^*(y)\big)\,dx\,dy \\ &\quad = \int \rho_1(x) G^{\mathbb{H}_F}(x, y)\rho_2(y)\,dx\,dy, \end{aligned} \tag{3.6}$$

where $G^{\mathbb{H}_F}(x, y) := -\log|x - \bar{y}| - \log|x - y|$.

3.3. *GFF as a continuous functional.* Note that we could have used (3.5) and (3.6) to give an alternate and more direct definition of the zero and free boundary Gaussian free fields on $\mathbb{H}$. Here, (3.5) and (3.6) define inner products on the space of functions ρ on $\mathbb{H}$. They are well defined when ρ_1 and ρ_2 are smooth and compactly supported functions on $\mathbb{H}$ [each with total integral zero in the case of (3.6)]. By taking the Hilbert space closure of functions of this type, we get a larger space of ρ, which correspond to Laplacians of elements of $H(\mathbb{H})$, and which cannot all be interpreted as functions on $\mathbb{H}$. For example, the ρ for which (h, ρ) is $h_\varepsilon(z)$, the mean value of h on $\partial B_\varepsilon(z)$, is not a function, though it can be interpreted as a measure—a uniform measure on $\partial B_\varepsilon(z)$—and the inner products (3.5) and (3.6) still make sense when $\rho_1(z)\,dz$ and $\rho_2(z)\,dz$ are replaced with more general measures, as do the definitions of $-\Delta^{-1}\rho_1$ and $-\Delta^{-1}\rho_2$.

The (h, ρ) are centered jointly Gaussian random variables, defined for each ρ in this Hilbert space, with covariances given by the inner products (3.5) and (3.6) (which can be defined on the entire Hilbert space). For each particular ρ in this Hilbert space, (h, ρ) is a.s. well defined and finite; however, $\rho \to (h, \rho)$ is almost surely not a continuous linear functional defined on the entire Hilbert space, since a.s. $h \notin H(\mathbb{H})$.

In addition to the description of h as a distribution above, there are various ways to construct a space of ρ values—a subset of the complete Hilbert space—endowed with a topology that makes $\rho \to (h, \rho)$ almost surely continuous. For example, the map $h \to h_\varepsilon(z)$ is an a.s. Hölder continuous function of ε and z [29]. Also, the zero boundary GFF can be defined as a random element of $(-\Delta)^{-\varepsilon}L^2(D)$ for any $\varepsilon > 0$, and is hence a continuous linear function on $(-\Delta)^{\varepsilon}L^2(D)$, if D is bounded. (See [91] for definitions and further discussion of fractional powers of the Laplacian in this context.) Also, as mentioned earlier, both the free and zero boundary GFFs can be understood as random distributions [91].

The issues that come up when defining $\rho \to (h, \rho)$ as a continuous function on some topological space of ρ values are the same ones that come up when rigorously constructing a Brownian motion B_t: one can give the joint law of B_t for any finite set of t values explicitly by specifying covariances, and this determines the law for any fixed countable set of t values, but one needs to overcome some (mild)

technicalities in order to say "B_t is almost surely a continuous function." Indeed, if one uses the smallest σ-algebra in which B_t is measurable for each fixed t, then the event that B_t is continuous is not even in the σ-algebra.

On the other hand, if we are given a construction that produces a random continuous function with the right finite dimensional marginals, then it must be a Brownian motion. A standard fact (proved using characteristic functions and Fourier transforms) states that a random variable on a finite dimensional space is a centered Gaussian with a given covariance structure if and only if all of its one-dimensional projections are centered Gaussians with the appropriate variance. Thus, to establish that B_t is a Brownian motion, it is enough to show that each finite linear combination of B_t values is a (one-dimensional) centered Gaussian with the right variance. The following proposition formalizes the analogous notion in the GFF context. It is a standard and straightforward result about Gaussian processes (see [91] for a proof in the zero boundary case; the free boundary case is identical).

PROPOSITION 3.1. *The zero boundary GFF on $\mathbb{H}$ is the only random distribution h on $\mathbb{H}$ with the property that for each $\rho \in H_s(\mathbb{H})$ the random variable (h, ρ) is a mean-zero Gaussian with variance given by (3.5) (with $\rho_1 = \rho_2 = \rho$). Similarly, the free boundary GFF is the only random modulo-additive-constant distribution on $\mathbb{H}$ with the property that for each $\rho \in H_s(\mathbb{H})$ with $\int_{\mathbb{H}} \rho(z)\, dz = 0$ the random variable (h, ρ) is a mean-zero Gaussian with variance given by (3.6).*

In our proofs of Theorem 1.1 and Theorem 1.2 in Section 4, we will first construct a random distribution in the manner prescribed by the theorem statement and then check the laws of the one dimensional projections (which determine the laws of the finite and countably infinite dimensional projections) to conclude by Proposition 3.1 that it must be the GFF.

We remark that knowing h as a distribution determines the values of α_j in a basis expansion $h = \sum_j \alpha_j f_j$, as long as the $-\Delta f_j$ are sufficiently smooth. This in turn determines the value of $h_\varepsilon(z)$ almost surely for a countable dense set of ε and z values, which determines the values for all ε and z by the almost sure continuity of $h_\varepsilon(z)$ [29]. This is convenient because it means that h, as a distribution, a.s. determines $(z, \varepsilon) \to h_\varepsilon(z)$ as a function, which in turn determines μ_h and ν_h. [We could alternatively have defined $h_\varepsilon(z)$—and hence μ_h and ν_h—using weighted averages of h defined by integrating against smooth bump functions on $B_\varepsilon(z)$ instead of averages on $\partial B_\varepsilon(z)$. Though we will not do this here, one can easily construct measures this way that are almost surely equivalent to μ_h and ν_h.]

4. Coupling the GFF with forward and reverse SLE.

4.1. *Proofs of coupling theorems.* This section will simultaneously prove Theorem 1.1 and Theorem 1.2. It is instructive to prove them together, and we will put

TABLE 1

Forward flow SLE	**Reverse flow SLE**
$df_t(z) = \frac{2}{f_t(z)}\,dt - \sqrt{\kappa}\,dB_t$	$df_t(z) = \frac{-2}{f_t(z)}\,dt - \sqrt{\kappa}\,dB_t$
$d\log f_t(z) = \frac{(4-\kappa)}{2f_t(z)^2}\,dt - \frac{\sqrt{\kappa}}{f_t(z)}\,dB_t$	$d\log f_t(z) = \frac{-(4+\kappa)}{2f_t(z)^2}\,dt - \frac{\sqrt{\kappa}}{f_t(z)}\,dB_t$
$df_t'(z) = \frac{-2f_t'(z)}{f_t(z)^2}\,dt$	$df_t'(z) = \frac{2f_t'(z)}{f_t(z)^2}\,dt$
$d\log f_t'(z) = \frac{-2}{f_t(z)^2}\,dt$	$d\log f_t'(z) = \frac{2}{f_t(z)^2}\,dt$

the relevant calculations in tables, with those for the forward SLE coupling of Theorem 1.1 on the left side and those for the reverse SLE coupling of Theorem 1.2 on the right side.

Now, using the language of stochastic differential equations and applying Itô's formula in the case $W_t = \sqrt{\kappa}B_t$, we compute the time derivatives of the four processes $f_t(z)$, $\log f_t(z)$, $f_t'(z)$, and $\log f_t'(z)$ in both forward and reverse SLE settings. Here, $f_t'(z)$ denotes the spatial derivative $\frac{\partial}{\partial z} f_t$. (Similar calculations appear in [87] in the case $\kappa = 4$.)

We next define the martingales $\mathfrak{h}_t$ in both settings and compute their stochastic derivatives. The purpose of the stochastic calculus below is to show that the quantities $(\mathfrak{h}_t, \rho)$ are continuous local martingales (the fact that they are martingales will become apparent later) and to explicitly computing their quadratic variations, so that they can be understood as Brownian motions subject to an explicit time change. Ultimately, we will use the properties of these Brownian motions to establish couplings between SLE and the Gaussian free field.

Note that while the two columns have differed only in signs until now, the definitions of $\mathfrak{h}_t$ below will diverge in that one involves the imaginary and one the real part of $\mathfrak{h}_t^*$. We will write $\gamma := \sqrt{\min(\kappa, 16/\kappa)} \in [0, 2]$.

Before continuing with the calculation, we make several remarks.

TABLE 2

Forward flow SLE	**Reverse flow SLE**
$\chi := \frac{2}{\sqrt{\kappa}} - \frac{\sqrt{\kappa}}{2}$	$Q := \frac{2}{\sqrt{\kappa}} + \frac{\sqrt{\kappa}}{2} = \frac{2}{\gamma} + \frac{\gamma}{2}$
$\mathfrak{h}_t^*(z) := \frac{-2}{\sqrt{\kappa}}\log f_t(z) - \chi \log f_t'(z)$	$\mathfrak{h}_t^*(z) := \frac{2}{\sqrt{\kappa}}\log f_t(z) + Q\log f_t'(z)$
$d\mathfrak{h}_t^*(z) = \frac{2}{f_t(z)}\,dB_t$	$d\mathfrak{h}_t^*(z) = \frac{-2}{f_t(z)}\,dB_t$
$\mathfrak{h}_t(z) := \operatorname{Im}\mathfrak{h}_t^*(z)$	$\mathfrak{h}_t(z) := \operatorname{Re}\mathfrak{h}_t^*(z)$
$d\mathfrak{h}_t(z) = \operatorname{Im}\frac{2}{f_t(z)}\,dB_t$	$d\mathfrak{h}_t(z) = \operatorname{Re}\frac{-2}{f_t(z)}\,dB_t$

REMARK 4.1. The form of $d\mathfrak{h}_t(z)$ in the forward case is significant. At time $t = 0$, the function $-2\,\mathrm{Im}(f_t(z)^{-1})$ is simply $-2\,\mathrm{Im}(z^{-1})$. This is a positive harmonic function whose level sets are circles in $\mathbb{H}$ that are tangent to $\mathbb{R}$ at the origin. It is a multiple of the so-called Poisson kernel, and it is a derivative of the Green's function $G(y,z) = G^{\mathbb{H}_0}(y,z) = \log|\frac{y-\bar{z}}{y-z}|$ in the following sense:

$$\left[\frac{\partial}{\partial s}G(is,z)\right]_{s=0} = \frac{\partial}{\partial s}\left|\frac{z+is}{z-is}\right|_{s=0} = \mathrm{Re}\,\frac{2iz}{|z^2|} = 2\,\mathrm{Im}(z^{-1}).$$

Intuitively, the value of $-2\,\mathrm{Im}(f_t(z)^{-1})$ represents the harmonic measure of the tip of $\eta_t := \eta([0,t])$ as seen from the point z. Roughly speaking, as one makes observations of the GFF at points near the tip of η_t, the conditional expectation of h goes up or down by multiples of this function.

REMARK 4.2. Also, in the forward case, $\mathfrak{h}_0$ is the harmonic function on $\mathbb{H}$ with boundary conditions $-2\pi/\sqrt{\kappa}$ on the negative real axis and 0 on the positive real axis. We could have (for sake of symmetry) added a constant to $\mathfrak{h}_0$ (and general $\mathfrak{h}_t$) so that $\mathfrak{h}_0$ is equal to $-\lambda$ on the negative real axis and λ on the positive real axis, where $\lambda := \pi/\sqrt{\kappa}$. Observe that when $\kappa = 4$, we have $\chi = 0$, and hence each $\mathfrak{h}_t$ would be the harmonic function on $\mathbb{H}\setminus\eta_t$ with boundary conditions $-\lambda$ on the left side of the tip of η_t and λ on the right side. In this case, the $\lambda = \pi/2$ is the same (up to a $\sqrt{2\pi}$ factor stemming from a different choice of normalization for the GFF) as the value $\lambda = \sqrt{\pi/8}$ obtained in [87].

REMARK 4.3. In the reverse case, the expression for $d\mathfrak{h}_t$ has $\mathrm{Re}\,\frac{-2}{f_t(z)}$ in place of $\mathrm{Im}\,\frac{2}{f_t(z)}$. Intuitively, at time zero, when one observes what f_t looks like for small t, one learns something about the *difference* between h just to the left of 0 and h just to the right of 0. (It is this difference that determines the ratio of the ν_h densities to the left and to the right of zero, which is what determines how the zipping-up should behave in the short term.) The conditional expectation of h thus changes by a small multiple of $\mathrm{Re}\,\frac{2}{f_t(z)}$, which is negative on one side of the imaginary axis and positive on the other side. Unlike $\mathrm{Im}\,\frac{2}{f_t(z)}$, the function $\mathrm{Re}\,\frac{2}{f_t(z)}$ is nonzero on $\mathbb{R}$.

We use $\langle X_t, Y_t\rangle$ to denote cross variation between processes X_t and Y_t up to time t, so that $\langle X_t, X_t\rangle$ represents the quadratic variation of the process X_t up to time t. (The cross variation $\langle X_t, Y_t\rangle$ is also often written as $\langle X, Y\rangle_t$.) In both forward and reverse flow settings, $\mathfrak{h}_t(z)$ is a continuous local martingale for each fixed z and is thus a Brownian motion under the quadratic variation parameterization, which we can give explicitly.

If z is a point in a simply connected domain D, and ϕ conformally maps the unit disc to D, with $\phi(0) = z$, then we refer to the quantity $|\phi'(0)|$ as the *conformal radius* of D viewed from z. If, in the above definition of conformal radius, we replaced the unit disc with $\mathbb{H}$ and 0 with i, this would only change the definition by

TABLE 3

Forward flow SLE	**Reverse flow SLE**
$C_t(z) := \log \operatorname{Im} f_t(z) - \operatorname{Re} \log f_t'(z)$	$C_t(z) := -\log \operatorname{Im} f_t(z) - \operatorname{Re} \log f_t'(z)$
$d\langle \mathfrak{h}_t(z), \mathfrak{h}_t(z)\rangle = -dC_t(z)$	$d\langle \mathfrak{h}_t(z), \mathfrak{h}_t(z)\rangle = -dC_t(z)$

an additive constant. Thus, in the forward flow case, $C_t(z)$ is (up to an additive constant) the log of the conformal radius of $\mathbb{H} \setminus \eta([0,t])$ viewed from z. In both cases, $\mathfrak{h}_t(z)$ is a Brownian motion when parameterized by the time parameter $-C_t(z)$ (which is increasing as a function of t). The fact that $d\langle \mathfrak{h}_t(z), \mathfrak{h}_t(z)\rangle = -dC_t(z)$ may be computed directly via Itô's formula but it is also easy to see by taking $y \to z$ in the formulas for $\langle \mathfrak{h}_t(y), \mathfrak{h}_t(z)\rangle$ and $-dG_t(y,z)$ that we will give below.

We will now show that weighted averages of $\mathfrak{h}_t$ over multiple points in $\mathbb{H}$ are also continuous local martingales (and hence Brownian motions when properly parameterized). The calculation will make use of the function $G(y,z)$, which we take to be the zero boundary Green's function $G^{\mathbb{H}_0}(y,z)$ on $\mathbb{H}$ in the forward case and the free boundary Green's function $G^{\mathbb{H}_F}(y,z)$ in the reverse case.

Now write $G_t(y,z) = G(f_t(y), f_t(z))$ in the reverse case. In the forward case, write $G_t(y,z) = G(f_t(y), f_t(z))$ when y and z are both in the infinite component of $\mathbb{H} \setminus \eta_t$—otherwise, let $G_t(y,z)$ be the limiting value of $G_s(y,z)$ as s approaches the first time at which one of y or z ceases to be in this infinite component. The reader may check that for fixed y and z, this limit exists almost surely when $4 < \kappa < 8$: it is equal to zero when y and z are in different connected components of $\mathbb{H} \setminus \eta_t$, and when y and z lie in the same component, it is simply the Green's function of y and z on this bounded domain. Now we let ρ be a smooth compactly supported function on $\mathbb{H}$ (which we will assume has mean zero in the reverse case) and do some more calculations.

Each of the equations above comes from a straightforward Itô calculation. To explain their derivation, we begin by expanding the dG_t computation in the forward case (the reverse case is similar):

$$\begin{aligned}
dG_t(x,y) &= -d \operatorname{Re} \log\bigl[f_t(x) - f_t(y)\bigr] + d \operatorname{Re} \log\bigl[f_t(x) - \overline{f_t(y)}\bigr] \\
&= -2 \operatorname{Re} \frac{f_t(x)^{-1} - f_t(y)^{-1}}{f_t(x) - f_t(y)}\, dt \\
&\quad + 2 \operatorname{Re} \frac{f_t(x)^{-1} - \overline{f_t(y)^{-1}}}{f_t(x) - \overline{f_t(y)}}\, dt \\
&= 2 \operatorname{Re}\bigl(f_t(x)^{-1} f_t(y)^{-1}\bigr)\, dt - 2 \operatorname{Re}\bigl(f_t(x)^{-1} \bigl(\overline{f_t(y)}\bigr)^{-1}\bigr)\, dt \\
&= 2 \operatorname{Re}\bigl(i f_t(x)^{-1} \operatorname{Im}\bigl[2 f_t(y)^{-1}\bigr]\bigr)\, dt \\
&= -\operatorname{Im} \frac{2}{f_t(x)} \operatorname{Im} \frac{2}{f_t(y)}\, dt.
\end{aligned}$$

TABLE 4

Forward flow SLE	**Reverse flow SLE**
$G(y,z) := \log\lvert y-\bar{z}\rvert - \log\lvert y-z\rvert$	$G(y,z) := -\log\lvert y-z\rvert - \log\lvert y-\bar{z}\rvert$
$G_t(y,z) := G(f_t(y), f_t(z))$	$G_t(y,z) := G(f_t(y), f_t(z))$
$dG_t(y,z) = -\operatorname{Im}\frac{2}{f_t(y)}\operatorname{Im}\frac{2}{f_t(z)}\,dt$	$dG_t(y,z) = -\operatorname{Re}\frac{2}{f_t(y)}\operatorname{Re}\frac{2}{f_t(z)}\,dt$
$d\langle \mathfrak{h}_t(y), \mathfrak{h}_t(z)\rangle = -dG_t(y,z)$	$d\langle \mathfrak{h}_t(y), \mathfrak{h}_t(z)\rangle = -dG_t(y,z)$
$E_t(\rho) := \int_{\mathbb{H}} \rho(y)G_t(y,z)\rho(z)\,dy\,dz$	$E_t(\rho) := \int_{\mathbb{H}} \rho(y)G_t(y,z)\rho(z)\,dy\,dz$
$d\langle(\mathfrak{h}_t,\rho),(\mathfrak{h}_t,\rho)\rangle = -dE_t(\rho)$	$d\langle(\mathfrak{h}_t,\rho),(\mathfrak{h}_t,\rho)\rangle = -dE_t(\rho)$

The fact that $d\langle \mathfrak{h}_t(y), \mathfrak{h}_t(z)\rangle = -dG_t(y,z)$ is then immediate from our calculation of $d\mathfrak{h}_t$.

The fact that $d\langle(\mathfrak{h}_t,\rho),(\mathfrak{h}_t,\rho)\rangle = -dE_t(\rho)$ is essentially a Fubini calculation but it requires some justification. First, we claim that the $(\mathfrak{h}_t,\rho)$ are continuous martingales. We begin by considering $\mathfrak{h}_t(z)$ for a fixed z in the support of ρ. We have shown above that the quantity $\mathfrak{h}_t(z)$ is a Brownian motion under a certain parameterization. In the reverse case, the Loewner evolution gives that $|\frac{\partial}{\partial t}C_t(z)|$ is uniformly bounded above for z in the support of ρ and for all times t. [Note that $\operatorname{Im} f_t(z)$ is strictly increasing in t.] This immediately implies that $\mathfrak{h}_t(z)$ is a martingale (not merely a local martingale) because for each z and t, $\mathfrak{h}_t(z)$ represents the value of a Brownian motion stopped at a random time that is strictly less than a constant times t. The fact that the expectation of $\mathfrak{h}_t(z)$—given the filtration up to time $s<t$—is $\mathfrak{h}_s(z)$ is then immediate from the optional stopping theorem.

In the forward case, one obtains something similar by noting that the law of the conformal radius r of z in $\mathbb{H}\setminus\eta([0,t])$ has a power law decay as $r\to 0$—that is, the probability that $-C_t(z) > c$ decays exponentially in c, and is in fact bounded by an exponentially decaying function that is independent of z, for z in the support of ρ. (A precise description of the law of the conformal radius at time infinity appears as the main construction in [88].) This implies that $\mathfrak{h}_t(z)$ is a Brownian motion stopped at a time whose law decays exponentially (uniformly over z in the support of ρ) which is again enough to apply the optional stopping theorem and conclude that $\mathfrak{h}_t(z)$ is martingale. In both cases, we obtain that for any t, the probability distribution function for $|\mathfrak{h}_t(z)|$ decays exponentially fast, uniformly for z in the support of ρ. In both cases, we also see that (for any fixed t), $\mathfrak{h}_t(z)$ is an L^1 function of z and the probability space, which allows us to use Fubini's theorem and conclude that the $(\mathfrak{h}_t,\rho)$ are martingales.

Let L^p_{loc} denote the set of ψ for which the integral of $|\psi|^p$ over every compact subset of $\mathbb{H}$ is finite. The exponential decay above implies that $\mathfrak{h}_t$ is almost surely in L^1_{loc}, since the expected integral of $|\mathfrak{h}_t|$ over any compact set is finite. (Note that we can define $\mathfrak{h}_t$ arbitrarily on the measure zero set $\eta([0,t])$ without affecting

the definition of $\mathfrak{h}_t$ as an element of $L^1_{\mathrm{loc}}(\mathbb{H})$.) In fact, since $\mathbb{E}|\mathfrak{h}_t(z)|^p$ is bounded uniformly for z in a compact set, it follows that $\mathfrak{h}_t$ is almost surely in $L^p_{\mathrm{loc}}(\mathbb{H})$ for any $p \in (1, \infty)$. The fact that $\mathfrak{h}_t$ is almost surely in L^1_{loc} also implies that it can be understood as a random distribution on $\mathbb{H}$.

Moreover,

$$\sup_{s\in[0,t]} \left|\mathfrak{h}_s(z)\right| \tag{4.1}$$

also has, by Doob's inequality, a law that decays exponentially, uniformly in z. Thus, (4.1) also belongs a.s. to $L^p_{\mathrm{loc}}(\mathbb{H})$ for any $p \in (1, \infty)$. From this and the a.s. continuity of SLE, it follows that $(\mathfrak{h}_t, \rho)$ is a.s. continuous in t. [This continuity is obvious in the reverse case; in the forward case, it is also obvious if one replaces ρ by ρ_ε, which we define to be zero on an ε neighborhood of η and ρ elsewhere. The fact that (4.1) belongs to $L^p_{\mathrm{loc}}(\mathbb{H})$ implies that the $(\mathfrak{h}_t, \rho_\varepsilon)$ converge to $(\mathfrak{h}_t, \rho)$ uniformly, for almost all η, and a uniform limit of continuous functions is continuous.]

Now we can show $d\langle(\mathfrak{h}_t, \rho), (\mathfrak{h}_t, \rho)\rangle = -dE_t(\rho)$, as noted in [87], either via a stochastic Fubini's theorem (see, e.g., [77], Section IV.4) or by using the following simpler approach proposed in private communication by Jason Miller.

First, note that $\langle(\mathfrak{h}_t, \rho_1), (\mathfrak{h}_t, \rho_2)\rangle$ is characterized by the fact that

$$(\mathfrak{h}_t, \rho_1)(\mathfrak{h}_t, \rho_2) - \big\langle(\mathfrak{h}_t, \rho_1), (\mathfrak{h}_t, \rho_2)\big\rangle$$

is a local martingale. Thus, it suffices for us to show that

$$(\mathfrak{h}_t, \rho_1)(\mathfrak{h}_t, \rho_2) + \int \rho_1(x)\rho_2(y)G_t(x, y)\,dx\,dy \tag{4.2}$$

is a martingale. We know from the above calculations that

$$\mathfrak{h}_t(x)\mathfrak{h}_t(y) + G_t(x, y)$$

is a martingale for fixed x and y in $\mathbb{H}$. Since $G_t(x, y)$ is nonincreasing and the $\mathfrak{h}_t(z)$ have laws that decay exponentially, uniformly in z, we can use Fubini's theorem to conclude that (4.2) is a martingale. Thus, we have that $(\mathfrak{h}_t, \rho)$ is a Brownian motion when parameterized by time $-E_t(\rho)$. To complete the proofs of Theorems 1.1 and 1.2, recall that in the theorem statements $\tilde{h}$ denotes an instance of the free boundary GFF on $\mathbb{H}$, and note that since each $(\mathfrak{h}_T + \tilde{h} \circ f_T, \rho)$ is a sum of a standard Brownian motion stopped at time $E_0(\rho) - E_T(\rho)$ and a conditionally independent Gaussian of variance $E_T(\rho)$, it has the same law as a Gaussian of variance $E_0(\rho)$ and mean $(\mathfrak{h}_0, \rho)$. (See Figure 14.) For future reference, we note that in the reverse flow case one may integrate the expression for $d\mathfrak{h}_t(z)$ above to find (using the stochastic Fubini's theorem) that

$$d(\mathfrak{h}_t, \rho) = \big(-2\,\mathrm{Re}(f_t)^{-1}, \rho\big)\,dB_t. \tag{4.3}$$

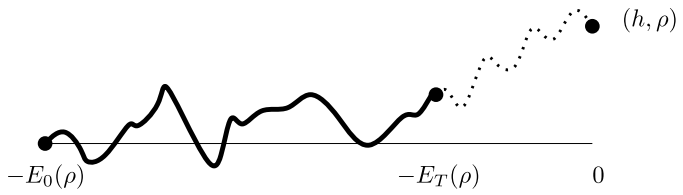

FIG. 14. *The pair* $(-E_t(\rho), (\mathfrak{h}_t, \rho))$ *traces the graph of a Brownian motion* (*solid curve*) *as* t *ranges from* 0 *to* T. *Conditioned on this, the difference between* (h, ρ) *and* $(\mathfrak{h}_T, \rho)$ *is a centered Gaussian of variance* $E_T(\rho)$. *Choosing* (h, ρ) *to be* $(\mathfrak{h}_T, \rho)$ *plus a Gaussian of this variance is equivalent to continuing the Brownian motion parameterized by* $-E_t(\rho)$ *time* (*solid curve*) *all the way to time zero* (*dotted curve*) *and letting* (h, ρ) *be its value at time zero.*

REMARK 4.4. The statement of Theorem 1.1 excluded the case $\kappa \geq 8$, since SLE(κ) is space-filling in that case and $\mathfrak{h}_t$ cannot be defined as a function almost everywhere. Nonetheless, we may still define $(\mathfrak{h}_t, \rho)$ to be the solution to the stochastic differential equation $d(\mathfrak{h}_t, \rho) = (-2\,\mathrm{Im}(f_t)^{-1}, \rho)\, dB_t$. In this case, the calculations above again yield that $d\langle(\mathfrak{h}_t, \rho), (\mathfrak{h}_t, \rho)\rangle = -dE_t(\rho)$, which as before implies that $(\mathfrak{h}_T + \tilde{h} \circ f_T, \rho)$ and $(\mathfrak{h}_0 + \tilde{h}, \rho)$ agree in law for each ρ, just as in the $\kappa < 8$ case, which yields a $\kappa \geq 8$ analog of Theorem 1.1. (Figure 14 still makes sense then $\kappa \geq 8$.)

It will be useful for later purposes to note that (at least in the reverse SLE case) the graph in Figure 14 actually uniquely determines (and is uniquely determined by) the process $W_t = \sqrt{\kappa} B_t$ almost surely; see Figure 15. This is a special case of a much more general theorem about stochastic processes (see Chapter IX, Theorem 2.1 of [78]—it suffices that $(\mathfrak{h}_t, \rho)$ satisfies an SDE in W_t with a diffusive coefficient that remains strictly bounded away from zero and infinity, at least as long as we stop at any time strictly before $t = \infty$). This means that the evolution of η can be described by the Brownian motion in Figure 14, as well as by the Brownian motion B_t. Context will determine which description is more convenient to work with.

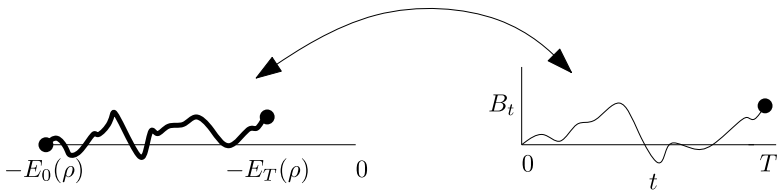

FIG. 15. *The graph traced by* $(-E_t(\rho), (\mathfrak{h}_t, \rho))$ *as* t *ranges from* 0 *to* T (*left*) *and the graph traced by* (t, B_t) (*right*), *where* $W_t = \sqrt{\kappa} B_t$. *The left graph uniquely determines the right graph, and vice versa, almost surely. Each has the law of a standard Brownian motion* (*up to a stopping time*).

4.2. *Alternative underlying geometries and SLE$_{\kappa,\rho}$.* Both Theorems 1.1 and 1.2 can be generalized to other values of $\mathfrak{h}_0$ using the so-called SLE$_{\kappa,\rho}$ processes. (As discussed at the end of Section 1.2, changing $\mathfrak{h}_0$ can be interpreted as changing the underlying geometry on which Liouville quantum gravity is defined.) We generalize the latter here (see [20] for the former). We take $G = G^{\mathbb{H}_F}$ in the following. Slightly abusing notation, we will consider situations where $\rho(y)\,dy$ represents a general signed measure (instead of requiring that ρ be a smooth test function). In this case, $(F, \rho) = \int F(y)\rho(y)\,dy$ represents integration of F w.r.t. this measure.

THEOREM 4.5. *Fix $\kappa > 0$ and a signed measure $\rho(y)\,dy$ on $\overline{\mathbb{H}}$ with finite positive and finite negative mass supported on some closed $\mathcal{C} \subset \mathbb{H}$. Write*

$$\hat{\mathfrak{h}}_t(z) = \mathfrak{h}_t(z) + \frac{1}{2\sqrt{\kappa}} \int G_t(y, z)\rho(y)\,dy, \tag{4.4}$$

where $\mathfrak{h}_t(z)$ is as in Theorem 1.2, and let $\tilde{h}$ be an instance of the free boundary GFF on $\mathbb{H}$, independent of B_t. Let η_T be the segment generated by the reverse Loewner flow

$$df_t(z) = \frac{-2}{f_t(z)}\,dt - dW_t, \tag{4.5}$$

where

$$\begin{aligned} dW_t &= \left(\int \operatorname{Re} \frac{-1}{f_t(y)} \rho(y)\,dy \right) dt + \sqrt{\kappa}\,dB_t \\ &= \left(-\operatorname{Re}(f_t)^{-1}, \rho\right) dt + \sqrt{\kappa}\,dB_t, \end{aligned} \tag{4.6}$$

up to any stopping time $T \geq 0$ at or before the smallest t for which $0 \in f_t(\mathcal{C})$ (here f_t is extended continuously from $\mathbb{H}$ to $\overline{\mathbb{H}}$). Then the following two random distributions (modulo additive constants) on $\mathbb{H}$ agree in law: $h = \hat{\mathfrak{h}}_0 + \tilde{h}$ and $\hat{\mathfrak{h}}_T + \tilde{h} \circ f_T$.

We will make several observations before we prove Theorem 4.5. First, if ρ is supported on a set of n points $y_1, \ldots, y_n$ in $\mathbb{H}$, with masses given by real numbers $\rho_1, \rho_2, \ldots, \rho_n$, then the process defined by (4.5) and (4.6) is (the reverse form of) what is commonly called an SLE$_{\kappa,\rho}$ process in the literature: in this case, (4.6) takes the form

$$dW_t = \sum_{i=1}^{n} \operatorname{Re} \frac{-\rho_i}{f_t(y_i)}\,dt + \sqrt{\kappa}\,dB_t, \tag{4.7}$$

which is the same expression one finds in the usual definition of the forward-flow SLE$_{\kappa,\rho}$ process, as in [89]. (Note that the notation here differs from [89], since

here we use ρ to denote the measure, not the vector of mass values ρ_i.) In the special case that ρ is supported at a single point $x \in \mathbb{R}$, with mass ρ_1 we find that

$$df_t(x) = \frac{-2}{f_t(x)}\,dt - dW_t = \frac{-2+\rho_1}{f_t(x)}\,dt - \sqrt{\kappa}\,dB_t, \tag{4.8}$$

so that $f_t(x)/\sqrt{\kappa}$ is a Bessel process of dimension δ satisfying $(\delta-1)/2 = (\rho_1 - 2)/\kappa$, that is,

$$\delta = 1 + \frac{2(\rho_1 - 2)}{\kappa}. \tag{4.9}$$

For this and future discussion, it will be useful to recall a few standard facts:

1. The Bessel process X_t of dimension δ by definition satisfies $dX_t = dB_t + \frac{\delta-1}{2}X_t^{-1}\,dt$. Hence $d\log X_t = \frac{1}{X_t}\,dB_t + \frac{\delta-1}{2X_t^2}\,dt - \frac{1}{2X_t^2}\,dt$. The process $\log X_t$, when parameterized by its quadratic variation, is a Brownian motion with a constant drift of magnitude $\frac{\delta-2}{2}$.

2. If X_t is a Bessel process of dimension δ started at $X_0 = x$ and run until the first time T that it reaches zero, then the time reversal X_{T-t} has the law of a Bessel process of dimension δ', started at zero and run until the last time that it hits x, where δ' is the dimension one gets by changing the sign of the drift in $\log X_t$. That is, $\frac{\delta-2}{2} = -\frac{\delta'-2}{2}$, so that

$$\delta = 4 - \delta'. \tag{4.10}$$

3. In the usual forward flow definition of $\mathrm{SLE}_{\kappa,\rho_1'}$ the function $f_t(x)$ is a Bessel process of dimension

$$\delta' = 1 + \frac{2(\rho_1' + 2)}{\kappa}. \tag{4.11}$$

The reason for the difference from (4.9) can be seen by considering the case $\rho_1 = \rho_1' = 0$. In the reverse process, the Loewner drift is pulling $f_t(x)$ toward the origin, while in the forward process the Loewner drift is pushing $f_t(x)$ away from the origin. In both cases, ρ_1 (or ρ_1') indicates a quantity of additional force pushing $f_t(x)$ away from the origin.

4. Combining (4.9), (4.10) and (4.11) gives a relationship between ρ_1 and ρ_1'. Namely, $1 + \frac{2(\rho_1'+2)}{\kappa} = 4 - (1 + \frac{2(\rho_1-2)}{\kappa})$, so that

$$\rho_1' = \kappa - \rho_1. \tag{4.12}$$

This means that if we run a reverse $\mathrm{SLE}_{\kappa,\rho_1'}$ until the time T at which $f_t(x)$ hits zero, then f_T maps $\mathbb{H}$ to $\mathbb{H} \setminus \eta_T$ where η_T has the law of an initial segment of a forward $\mathrm{SLE}_{\kappa,\rho_1}$. In particular, if $\rho_1' = \kappa$, then η_T has the law of an ordinary SLE stopped at a time T (which corresponds to the last time that a Bessel process hits a certain value). This will be important later.

Recall from Section 1.2 that changing $\mathfrak{h}_0$ to $\hat{\mathfrak{h}}_0$ can be interpreted as changing the underlying geometry on which Liouville quantum gravity is defined. Moreover, ρ is proportional to $-\Delta(\hat{\mathfrak{h}}_0 - \mathfrak{h}_0)$, and $-\Delta\hat{\mathfrak{h}}_0$ is proportional to the overall Gaussian curvature density (see the Appendix).

We will give a formal proof of Theorem 4.5 below using Itô calculus, but first let us offer an informal explanation of why the result is true. The idea behind Theorem 4.5 is to interpret (4.4) as the expectation of h in a certain weighted measure and (4.6) as the description of the law of W_t in that measure. This is easiest to understand when we first switch coordinates using the correspondence shown in Figure 15. Suppose first that ρ is such that (3.6) is finite with $\rho_1 = \rho_2 = \rho$ and that the total integral of ρ is zero. If dh is the law of a (centered or not centered) GFF, then the standard Gaussian complete-the-square argument shows that $e^{(h,\rho)}\,dh = e^{(h,-2\pi\Delta^{-1}\rho)\nabla}\,dh$ (normalized to be a probability measure) is the law of the standard GFF plus $-2\pi\Delta^{-1}\rho$. When we weight the law of the Brownian motion in Figure 14 by $e^{\alpha(h,\rho)}$ for some constant α [note that (h, ρ) is the terminal value that the Brownian motion in that figure reaches at time zero] this is equivalent to adding a constant drift term to the Brownian motion [parameterized by $-E_t(\rho)$] in Figure 14.

We take $\alpha = \frac{1}{2\sqrt{\kappa}}$ and weight by

$$e^{(h,(1/(2\sqrt{\kappa}))\rho)}, \tag{4.13}$$

which, as explained above, modifies the law in a way that amounts to adding the drift term of $\frac{1}{2\sqrt{\kappa}}(E_0(\rho) - E_t(\rho))$ to the Brownian motion in Figure 14. Recalling the correspondence shown in Figure 15, the fact that the left figure is a Brownian motion with this constant drift (up to a stopping time) completely determines the law of W_t up to that stopping time. Indeed, recalling (4.3), we find that the law of W_t is necessarily the one described by (4.6).

We have now related the weighted measure to (4.6), but what does this have to do with (4.4)? Observe that

$$(\hat{\mathfrak{h}}_t, \rho) = (\mathfrak{h}_t, \rho) + \frac{1}{2\sqrt{\kappa}}E_t(\rho)$$

represents the conditional expectation (in the weighted measure) of (h, ρ) given $B_{\cdot}$ up to time t. In fact, by the standard complete-the-square argument, the function $\hat{\mathfrak{h}}_t$ in Theorem 4.5 represents the conditional expectation (in the weighted measure) of h, given $B_{\cdot}$ up to time t, and is thus a martingale in t. The above construction (and a bit of thought) actually constitutes a proof of Theorem 4.5 when (3.6) is finite and the total integral of ρ is zero.

The argument above can be adapted to more general ρ. If the total integral of ρ is not zero, we may modify ρ by adding some mass very far from the origin, so that the total integral becomes zero but the drift in (4.6) does not change very much. If (3.6) is infinite, we may be able to modify it to make it finite: for example,

TABLE 5

Reverse flow SLE
$df_t(z) = \frac{-2}{f_t(z)}\,dt + (\int \operatorname{Re}\frac{1}{f_t(y)}\rho(y)\,dy)\,dt - \sqrt{\kappa}\,dB_t$
$d\log f_t(z) = \frac{-(4+\kappa)}{2f_t(z)^2}\,dt + f_t(z)^{-1}(\int \operatorname{Re}\frac{1}{f_t(y)}\rho(y)\,dy)\,dt - \frac{\sqrt{\kappa}}{f_t(z)}\,dB_t$
$df_t'(z) = \frac{2f_t'(z)}{f_t(z)^2}\,dt$
$d\log f_t'(z) = \frac{2}{f_t(z)^2}\,dt$

if ρ is a point mass, then we may replace it with a uniform measure on a tiny ball centered at that point mass, and the harmonicity of $G_t(\cdot, z)$ in (4.4) and of $\operatorname{Re}(f_t)^{-1}$ in (4.6) show that (outside of this small ball) neither (4.4) nor (4.6) is affected by this replacement. We will present a more direct Itô calculation below, which also applies when (3.6) is infinite.

PROOF OF THEOREM 4.5. We will follow the calculations of Theorem 1.2 and check where differences appear. First, we find Table 5.

Also, as before, we compute

$$dG_t(y,z) = -\operatorname{Re}\frac{2}{f_t(y)}\operatorname{Re}\frac{2}{f_t(z)}\,dt,$$

and recall that

$$\hat{\mathfrak{h}}_t(z) := \frac{2}{\sqrt{\kappa}}\log|f_t(z)| + Q\log|f_t'(z)| + \frac{1}{2\sqrt{\kappa}}\int G_t(y,z)\rho(y)\,dy.$$

We then find that when computing $d\hat{\mathfrak{h}}_t(z)$ the extra term in $d\frac{2}{\sqrt{\kappa}}\operatorname{Re}\log f_t(z)$ cancels the term $d\frac{1}{2\sqrt{\kappa}}\int G_t(y,z)\rho(y)\,dy$ so that $d\hat{\mathfrak{h}}_t(z) = \operatorname{Re}\frac{-2}{f_t(z)}\,dB_t$, just as in the proof of Theorem 1.2. The remaining calculations are the same as in the proof of Theorem 1.2. □

We next remark, in the context of Theorem 4.5, that if $(\mathbb{H}, \hat{\mathfrak{h}}_T + \tilde{h}\circ f_T)$ is a quantum surface, then

$$\begin{aligned} &f_T(\mathbb{H}, \hat{\mathfrak{h}}_T + \tilde{h}\circ f_T) \\ &\quad = f_T\Big(\mathbb{H}, \mathfrak{h}_T + \frac{1}{2\sqrt{\kappa}}\int G_T(y,\cdot)\rho(y)\,dy + \tilde{h}\circ f_T\Big), \end{aligned} \tag{4.14}$$

and this can be written

$$\Big(\mathbb{H}\setminus K_T, \mathfrak{h}_0 + \frac{1}{2\sqrt{\kappa}}\int G\big(f_T(y),\cdot\big)\rho(y)\,dy + \tilde{h}\Big). \tag{4.15}$$

When $\kappa < 4$, this suggests the following interpretation of Theorem 4.5. We start with $\hat{\mathfrak{h}}_0 + \tilde{h}$, which is actually only defined up to additive constant, so it determines a quantum surface up to a multiplicative constant. We zip up this (modulo multiplicative constant) quantum surface until a stopping time T, and condition on the zipper map f_T. Then the conditional law of the new zipped-up quantum surface (which is also defined only up to a multiplicative constant) is the same as the original law except that ρ is replaced by the f_T image of ρ. We have not fully established that this interpretation is correct, because we have not yet shown that $\hat{\mathfrak{h}}_0 + \tilde{h}$ uniquely determines f_T. As in Theorem 1.2, we have only shown that sampling h from $\hat{\mathfrak{h}}_0 + \tilde{h}$ is equivalent to first zipping up according to a given law, then sampling the field from a putative conditional law in the zipped up picture, and then unzipping.

5. Quantum zippers and conformal welding.

5.1. *Overview of zipper proofs.* The goal of this section is to prove the statements in Section 1 that we have not yet proved: namely, Theorem 1.3 (which, by the a.s. removability of the SLE paths, implies Theorem 1.4 and Corollary 1.5, as explained in Section 1) and Theorem 1.8. Both proofs will be completed in Section 5.4.

Given what we know now, Theorem 1.3 may not seem surprising. In light of Theorem 1.2, we know that, given the independent pair (h, η) described there, the quantum boundary lengths along the left and right sides of η are both a.s. well defined. Indeed, recall that to measure the ν_h length along the left side of $\eta([0, t])$, we may "unzip" via f_t^{η}—and the transformation rule (1.3)—so that $\eta([0, t])$ maps to an interval of $\mathbb{R}$, and then measure the quantum length of that interval. We only need to show that the quantum boundary measure of η measured from the left agrees with the quantum boundary measure of η measured from the right.

We already have some information about how these measures depend on h. For example, it is immediate from the definition of ν_h that if we change h—by adding a smooth function to h that is equal to a constant C in a region $A \subset \mathbb{H}$—then this has the effect of multiplying the length of $\eta([0, t]) \cap A$ (as measured from both left and right sides) by $e^{C\gamma/2}$. More generally, if ν_h is the measure from one of the two sides (viewed as a measure on $\overline{\mathbb{H}}$, which happens to be supported along $\eta \cup \mathbb{R}$) and $\phi : \overline{\mathbb{H}} \to \mathbb{R}$ is smooth, then

$$\nu_{h+\phi} = e^{(\gamma/2)\phi} \nu_h, \tag{5.1}$$

(i.e., the measure whose Radon–Nikodym derivative w.r.t. ν_h is $e^{(\gamma/2)\phi}$). This implies that the left and right measures depend on h in a similar way, but it does not show that they agree.

One possible way to prove agreement might be to suppose otherwise for contradiction. Then if we cover η with a lot of small balls of comparable quantum size,

we will find that in some of the balls the quantum length of η is greater measured from the left side and in some greater from the right side. Still, we would expect some long range near-independence and a law of large numbers to show that all of these differences average out when the balls are small enough; taking limits as the balls get smaller should show that the lengths on the two sides are in fact equal. (A related argument based was used in [86] to prove the so-called height gap lemma for discrete Gaussian free fields.)

An alternative approach to proving Theorem 1.3 would be to try to show that both the left and right boundary measures give "quantum length" measures of the curve that correspond to the quantum analog (as in [30]) of the natural time parameterization constructed via the Doob–Meyer decomposition in [51, 52]. The uniqueness arguments introduced in [51] could then be used to show that these two measures must agree.

While both of the above approaches appear viable, we will actually prove Theorem 1.3 with a third approach, which we feel is instructive. Namely, we will first construct the invariant measure described by Theorem 1.8, in the manner outlined in Figure 8, and then use symmetries and the ergodic theorem to deduce that the quantum measures on the two sides of the path are almost surely equal. Before doing this, we establish some results of independent interest. Section 5.2 describes a $T = \infty$ analog of Theorem 1.2 and Section 5.3 describes an interesting space–time symmetry: by changing the underlying geometry of the space used to define Liouville quantum gravity, one may construct a quantum zipper that is invariant with respect to a modified notion of capacity time.

5.2. *Zipping up "all the way" and capacity stationarity.* Let Γ^0 be the law of the pair (h, η) in which h is $\frac{2}{\sqrt{\kappa}} \log|z| + \tilde{h}$ (with $\tilde{h}$ a free boundary GFF on $\mathbb{H}$, defined modulo additive constant) and η is an independent SLE_κ. This is the measure that appears in Theorem 1.2 and Corollary 1.5. Note that the space of $((D_1, h^{D_1}), (D_2, h^{D_2}))$ configurations is in one to one correspondence with the space of pairs (h, η) on a full Γ^0 measure set, so we can view the $\mathcal{Z}_t^{\mathrm{CAP}}$ of Corollary 1.5 as acting on the pair (h, η) and we will denote the thus transformed pair by $\mathcal{Z}_t^{\mathrm{CAP}}(h, \eta)$. By Theorem 1.2, the $\mathcal{Z}_t^{\mathrm{CAP}}$ of Corollary 1.5 with $t < 0$ (the "unzipping" direction) are measure preserving transformations of Γ^0.

Using Theorem 1.2 alone, we do not yet have a definition of $\mathcal{Z}_t^{\mathrm{CAP}}$ for $t > 0$ (the "zipping up" direction). However, we can construct a stationary process (h^t, η^t), defined for all t, such that whenever $t < 0$ and $s \in \mathbb{R}$ we have

$$(h^{s+t}, \eta^{s+t}) = \mathcal{Z}_t^{\mathrm{CAP}}(h^s, \eta^s). \tag{5.2}$$

To construct this process, note that once we choose (h^T, η^T) for some large fixed constant T from Γ^0, the evolution rule (5.2) determines (h^t, η^t) for all $t < T$. [The fact that the unzipping maps preserve Γ^0 implies that for each fixed $t < T$, the pair (h^t, η^t) also has the law of Γ^0.] Taking a limit as $T \to +\infty$ (using Kolmogorov's

consistency theorem), we obtain a process defined for all $t \in \mathbb{R}$. Denote by Γ the law of this process. We frequently write $h = h^0$ and $\eta = \eta^0$.

For the remainder of Section 5, we will use f_t to denote "zipping down" maps when $t < 0$ and "zipping up" maps when $t > 0$. Once Corollary 1.5 is established, this will amount to defining f_t for all $t \in \mathbb{R}$ (using the notation of Corollary 1.5) as

$$f_t = \begin{cases} f_t^h, & t \geq 0, \\ f_{-t}^\eta, & t \leq 0. \end{cases} \tag{5.3}$$

(We stress that this contrasts with earlier notation, where we defined f_t only for $t > 0$ and—to highlight similarities—used f_t to describe both forward Loewner evolution in the AC-geometry context and reverse Loewner evolution in the Liouville quantum gravity context.) However, using Theorem 1.2 alone, it is not a priori clear that h determines the map f_t^h in (5.3). Even so, it *is* clear that these maps are determined by the process (h^t, η^t) with the law Γ constructed above, so for now we will define f_t to be the maps determined by process (h^t, η^t).

Now compare (1.6) and (1.7) and note that the Brownian motion corresponding to the forward Loewner evolution of an SLE segment η_T (the restriction of a Brownian motion to an interval of time) is (up to additive constant) the time reversal of the Brownian motion corresponding to the reverse Loewner evolution f_t for the same curve. Indeed, if we write $\widetilde{B}_t = B_{T-t}$ and $\tilde{f}_t = f_{T-t}$, then (1.7) holds for a general pair B_t and f_t (indexed by $t \in [0, T]$) precisely when (1.6) holds for $\widetilde{B}_t$ and $\tilde{f}_t$. The f_t map is driven by a Brownian motion in the $t < 0$ "unzipping" direction (since the curve that it unzips is an SLE curve) and from the definition of f_t for general times (which involved starting at stationarity for some large T and unzipping from there to get f_t for $t < T$, taking the limit as $T \to \infty$) we see that we may define a standard Brownian motion B_t for all times t so that $B_0 = 0$ and f_t satisfies (1.7) for all time. [This is also to be expected from (5.3), since once we show that f_t^h is determined by h, we will expect that the forward flow f_t^η and reverse flow f_t^h are driven by Brownian motions, for $t \geq 0$, that are independent of one another.]

We have shown above how to construct the f_t corresponding to Γ from a single Brownian motion (indexed by all of $\mathbb{R}$). We will now give an alternate description of the law Γ by coupling this f_t with h using Theorem 1.2 applied with $T = +\infty$. Informally, this amounts to "zipping up all the way" and then sampling a GFF on the fully zipped up surface. (This description comprises the remainder of Section 5.2. It will not be used in subsequent sections and may be skipped on a first read. See [23] for an analysis of what happens when two sides of a general quantum wedge are zipped together.) To explain what this means, recall that we showed in the proof of Theorem 1.2 that for each fixed ρ with total integral zero, the process $(\mathfrak{h}_t, \rho)$ is a Brownian motion parameterized by $-E_t(\rho)$; in particular, this implies that $\lim_{t \to +\infty}(\mathfrak{h}_t, \rho)$ exists almost surely and has variance at most $E_0(\rho)$.

We will be interested in $\mathfrak{h}_t$ modulo additive constant. Let us focus on the function $\mathfrak{h}_t(z) - \mathfrak{h}_t(y)$ for some fixed point $y \in \mathbb{H}$. By harmonicity, $\mathfrak{h}_t(z)$ is equivalent to its mean value on a ball centered at z, and hence $\mathfrak{h}_t(z) - \mathfrak{h}_t(y)$ can be written as $(\mathfrak{h}_t, \rho)$ for a smooth, compactly supported ρ with mean zero. Thus, $\mathfrak{h}_t(z) - \mathfrak{h}_t(y)$ is a Brownian motion when parameterized by $-E_t(\rho)$, and the total amount of quadratic variation time elapsed as $t \to +\infty$ is at most $E_0(\rho)$ almost surely. When z is restricted to a bounded domain D with closure in $\mathbb{H}$, we can define such a ρ for each z so that the quantity $E_0(\rho)$ is bounded uniformly in z. By Doob's inequality, this implies that $\sup_t |\mathfrak{h}_t(z) - \mathfrak{h}_t(y)|$ has a finite second moment, uniformly bounded for z in D, and hence

$$\mathbb{E} \int_D \sup_t |\mathfrak{h}_t(z) - \mathfrak{h}_t(y)|^2 \, dz < \infty.$$

In fact, since $\mathfrak{h}_t(z) - \mathfrak{h}_t(y)$ converges almost surely for each z, the above shows that $\mathfrak{h}_t(\cdot) - \mathfrak{h}_t(y)$ restricted to D is almost surely Cauchy in the space $L^2(D)$; since the $\mathfrak{h}_t$ are harmonic on D, this implies that $\mathfrak{h}_t$ almost surely converges uniformly (on each compact subset of $\mathbb{H}$, modulo additive constant) to a harmonic limit $\mathfrak{h}_\infty$ as $t \to +\infty$.

On the other hand, the maps f_t clearly do not converge to a finite limit; the Loewner evolution shows that when $z \in \mathbb{H}$ is fixed, $f_t(z)$ cannot converge to any value except for ∞ as $t \to +\infty$. However, let us define $\hat{f}_t$ to be $a_t f_t + b_t$ where complex constants a_t and b_t are chosen for each t so that $\hat{f}_t(i-1) = (i-1)$ and $\hat{f}_t(i+1) = i+1$. If f_T is as in Figure 5, then $\hat{f}_T$ would be as in Figure 16.

PROPOSITION 5.1. *The limit*

$$\hat{f}_\infty(z) := \lim_{t \to \infty} \hat{f}_t(z) \tag{5.4}$$

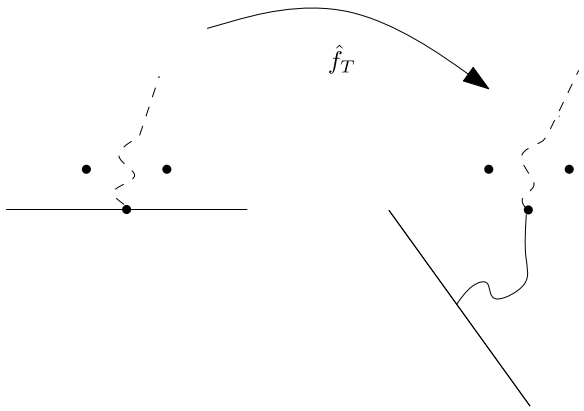

FIG. 16. *The map $\hat{f}_T$ is first normalized so that $\hat{f}_T(i-1) = (i-1)$ and $\hat{f}_T(i+1) = i+1$. (Here, $i \pm 1$ are shown as dots; a curve η and its image under $\hat{f}_T$ is also shown.) The limit $\hat{f}_\infty(z) := \lim_{t \to \infty} \hat{f}_t(z)$ a.s. exists as a conformal map from $\mathbb{H}$ to the complement in $\mathbb{C}$ of a semi-infinite path from some starting point to ∞.*

exists almost surely as an analytic map from $\mathbb{H}$ *to* $\mathbb{C}$. *It conformally maps* $\mathbb{H}$ *to the complement, in* $\mathbb{C}$, *of a certain random path from some finite starting point in* $\mathbb{C}$ *to* ∞. (*One may then translate the surface so that this starting point is the origin.*)

PROOF. To establish the existence of this limit, we argue that each of the two terms on the RHS of

$$\mathfrak{h}_t(z) = \mathfrak{h}_0\big(f_t(z)\big) + Q\log\big|f_t'(z)\big| \tag{5.5}$$

converges almost surely (modulo additive constant) to a limit. By (5.4) the LHS has a limit, so it will suffice to show that $\mathfrak{h}_0(f_t(z)) - \mathfrak{h}_0(f_t(y))$ a.s. converges uniformly to zero (for $y, z \in D$). For contradiction, suppose otherwise. Then, for some fixed C, let T_k be the first time after time 2^k for which there exist y and z in D such that

$$\big|\log|f_T(z)| - \log|f_T(y)|\big| \geq C. \tag{5.6}$$

By assumption, we may choose C so that with positive probability each of the T_k is finite. In each case, we may assume $|f_T(z)| \geq |f_T(y)|$ so that (5.6) implies $|f_T(z)|/|f_T(y)| > 1 + C$.

Now, conditioned on f_t up to such a time $T = T_k$, there is at least a constant C_1 probability that $\mathfrak{h}_t(z) - \mathfrak{h}_t(y)$ will change by at least some constant C_2 during the next $\min\{|f_T(y)|^2, |f_T(z)|^2\}$ units of capacity time after time T. By scaling, it is enough to observe that this is true when $|f_T(y)| < |f_T(z)|$ and $|f_T(y)| = 1$, which follows from the claim in the caption to Figure 17.

To prove this claim, recall (5.5) (and the definition of $\mathfrak{h}_0$ in Theorem 1.2):

$$\begin{aligned}&\mathfrak{h}_t(z) - \mathfrak{h}_t(y)\\ &\quad= \frac{2}{\sqrt{\kappa}}\big(\log|f_t(z)| - \log|f_t(y)|\big) + Q\big(\log|f_t'(z)| - \log|f_t'(y)|\big).\end{aligned} \tag{5.7}$$

To deal with the first difference on the RHS of (5.7), note that

$$\log|f_T(z) - a| - \log|f_T(y) - a|$$

cannot, as a function of a, be constant on the interval $[-1/2, 0]$ or constant on the interval $[0, 1/2]$. Indeed (if $|f_T(z)| > 1 + C$) the amount by which this function

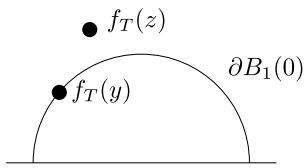

FIG. 17. *Assume that* $|f_T(y)| = 1$ *and* $|f_T(z)| > 1 + C$ *and consider the reverse Loewner flow driven by* $W_t = \sqrt{\kappa}B_t$). *Claim: for some* $C_1, C_2 > 0$ (*depending only on* C) *there is at least a* C_1 *probability that* $\mathfrak{h}_t(z) - \mathfrak{h}_t(y)$ *changes by at least* C_2 *after time* T.

varies over each interval is at least some $c > 0$. Thus, if W_t goes up or down by $1/2$ during t_Δ units of time after time T (and t_Δ is small enough), then the Loewner flow (1.7) shows that $f_t(z) - f_T(z)$ is approximately $W_T - W_t$ during this time, so that the first difference on the LHS of (5.7) changes by at least $\frac{c}{\sqrt{\kappa}}$. If we choose t_Δ small enough, we can also ensure that the second difference on the RHS of (5.7) changes by less than $\frac{c}{2\sqrt{\kappa}}$ [recall $d \log f_t' = (2/f_t^2)\, dt$]. Thus, (5.7) changes by at least $C_2 := \frac{c}{2\sqrt{\kappa}}$. The probability of this depends only on t_Δ, which depends only on C.

It follows from the above that on the event that the T_k are all finite, the quantity (5.7) a.s. changes by C_2 after T_k for infinitely many choices of k, contradicting the fact (which we have already shown) that $\mathfrak{h}_t(z) - \mathfrak{h}_t(y)$ a.s. converges uniformly to a limit for $y, z \in D$. We conclude that there is almost surely a last time for which $\log|f_T(z)| - \log|f_T(y)| \geq C$ for some y and z in D, and since this holds for any C, the RHS of (5.5) a.s. converges uniformly to zero for $y, z \in D$. Since $\log|f_t'(z)|$ [the second term on the RHS of (5.5)] then converges almost surely uniformly (modulo additive constant) to a limit, it follows (adding i times the harmonic conjugate) that $\log f_t'(z)$ converges almost surely uniformly on D (modulo additive constant), which implies the uniform convergence of f_t' (modulo multiplicative constant) and hence the convergence of $\hat{f}_t$ to a limit. Since the limiting map is analytic, it follows that the path (as in Figure 16) converges to a limiting continuous path as well. □

Theorem 1.2 can now be stated for $T = +\infty$. The statement is the same except that we replace

$$h \circ f_T + Q \log|f_T'| = \mathfrak{h}_0 \circ f_T + \tilde{h} \circ f_T + Q \log|f_T'|$$

with

$$\tilde{h} \circ \hat{f}_\infty + Q \log|\hat{f}_\infty'|,$$

considered modulo additive constant. [Recall that the $\mathfrak{h}_0(f_t(z))$ term on the RHS of (5.5) a.s. tends to zero modulo additive constant as $t \to +\infty$.]

Note that in this case, after "zipping up all the way to $+\infty$" the image $\hat{f}_\infty(\mathbb{H})$ is a (dense) subset of $\mathbb{C}$, rather than $\mathbb{H}$, and thus $\tilde{h}$ is defined as a free boundary GFF on $\mathbb{C}$ rather than just on $\mathbb{H}$. Given this change, there is nothing in the proof of Theorem 1.2 in Section 4, where the law of (h, ρ) was checked one test function ρ at a time, that fails to hold if we take $T = +\infty$.

To summarize, we have the following.

PROPOSITION 5.2. *We may produce a sample from* Γ *explicitly as follows*: *first sample a standard Brownian motion* B_t *for* $t \in \mathbb{R}$, *which determines* f_t (*and hence* η^t) *for all time* (*positive and negative*) *by solving* (1.7). *Then define* (*modulo additive constant*) $\mathfrak{h}_\infty := \lim_{t \to +\infty} \mathfrak{h}_t$ *and* $\hat{f}_\infty := \lim_{t \to +\infty} \hat{f}_t$ *and write*

$$h = h^0 = \tilde{h} \circ \hat{f}_\infty + \mathfrak{h}_\infty,$$

where $\tilde{h}$ is a free boundary GFF on $\mathbb{C}$. Define h^t for all other finite t as the quantum surface transformation of h^0 under f_t. That is, $h^t = h^0 \circ f_t^{-1} + Q \log |(f_t^{-1})'|$.

The entire process (h^t, η^t) is thus determined by $\tilde{h}$ and the process B_t. Note that for each t, the curve η^t is a random path. The forward Loewner evolution of this path—parameterized by its capacity time s—is given by f_{t-s}, with s ranging from 0 to $+\infty$. Its Loewner driving function is (as a function of s) $\sqrt{\kappa} B_{t-s}$. Note that $\mathcal{Z}_t^{\mathrm{CAP}}$ has the effect of translating the process $B_\cdot$ by t units to the left (and adding a constant to maintain $B_0 = 0$).

5.3. *Re-parameterizing time*: *General observations*. Now we make a general observation about the GFF. Suppose we write dh for the law of a (not necessarily centered) free boundary GFF h, and that f belongs to the Hilbert space on which h is defined, so that $\mathrm{Var}(h, f)_\nabla = (f, f)_\nabla$. As mentioned in Section 4.2, the standard Gaussian complete-the-square calculation shows that $e^{(h,f)_\nabla}\, dh$ (multiplied by a constant to make it a probability measure) is the law of the original GFF plus f. In other words, weighting the law of h by $e^{(h,f)_\nabla}$ is equivalent to deterministically adding f to h. Thus, weighting the law of h by $e^{(h,\rho)}$ (we assume that ρ has total integral zero) is equivalent to adding $-2\pi\Delta^{-1}\rho$ to h. Here Δ^{-1} is defined in the Neumann sense, that is, by taking by (3.3) with $G(x, y) = G^{\mathbb{H}_F}(x, y)$. Note also the following.

PROPOSITION 5.3. *Let (h^t, η^t) be the process with law Γ (as in Section 5.2), let ρ be a signed measure with total integral zero* [*for which (h^0, ρ) is a.s. finite*], *and write*

$$s_\rho(t) := \int_0^t e^{(h^s,\rho)}\, ds.$$

Then the weighted measure $e^{(h^0,\rho)}\Gamma$ (normalized to be a probability measure) is stationary w.r.t. s_ρ time.

PROOF. First, observe that the constant $\mathbb{E}_\Gamma e^{(h^0,\rho)}$ is finite, and by Fubini's theorem $\mathbb{E}_\Gamma \int_0^T e^{(h^t,\rho)}\, dt$ is T times that constant. The ergodic theorem implies that as $T \to \infty$ along integers the random variables $s_\rho(T)/T$ converge to a (possibly random) value c almost surely and in L^1. To see that c is a.s. constant, we would need to check that the pair $(h^\cdot, \eta^\cdot)$ is ergodic. We do not really need ergodicity for the proof of this proposition but we mention as an aside that it is not hard to prove [the Brownian motion generating η_t is ergodic; and if t_1 and t_2 are distinct times, one can use to GFF properties to show that the restrictions of h to tiny neighborhoods of $\eta(t_1)$ and $\eta(t_2)$ are nearly independent; letting the neighborhoods get small, and using scale invariance, one can establish a long-range mixing property that implies ergodicity]. The reason that ergodicity is not necessary for the proof of the proposition is that if $(h^\cdot, \eta^\cdot)$ had multiple ergodic components but each component were

invariant under the operation of translation by a fixed amount of s_ρ time, then the entire process would be invariant under this operation as well; so it is enough to focus on a single ergodic component.

Now, suppose that π_T denotes uniform measure on $[0, T]$ and we sample s from π_T. Then consider the measure $\Gamma \times \pi_T$. Once we have a sample from this measure, we may use it to generate a pair $(\bar{h}^t, \bar{\eta}^t) = (h^{t+s}, \eta^{t+s})$. That is, $(\bar{h}^\cdot, \bar{\eta}^\cdot)$ is the original process $(h^\cdot, \eta^\cdot)$ translated by the random quantity s, which was chosen uniformly from $[0, T]$. The stationarity of Γ implies that $(\bar{h}^\cdot, \bar{\eta}^\cdot)$ also has the law Γ. Now suppose that we instead sample from the weighted measure $\widetilde{\Gamma} = e^{(h^s, \rho)}\Gamma \times \pi_T$ (normalized to be a probability measure). Since $e^{(h^s, \rho)} = e^{(\bar{h}^0, \rho)}$, the induced measure on $(\bar{h}^\cdot, \bar{\eta}^\cdot)$ has the law $e^{(h^0, \rho)}\Gamma$.

However, sampling from $\widetilde{\Gamma}$ can also be done by first choosing the process $(h^\cdot, \eta^\cdot)$ from its marginal law, which is the measure Γ weighted by $\int_0^T e^{(h^t, \rho)}\, dt$ (which converges in law to $c\Gamma$—normalized to be a probability measure—in the total variation sense as $T \to \infty$, by the L^1 ergodic theorem) and then sampling s from $e^{(h^t, \rho)}\pi_T$ (normalized to be a probability measure), which amounts to choosing a uniform s_ρ time from the long interval. Since, as noted above, the process centered at this random time has the law $e^{(h^0, \rho)}\Gamma$, we find, by letting T tend to infinity, that $e^{(h^0, \rho)}\Gamma$ must be stationary with respect to s_ρ time. □

Recall from Section 1.2 that adding $-2\pi\Delta^{-1}\rho$ to h is equivalent to changing the underlying geometry w.r.t. which Liouville quantum gravity is defined. Thus, Proposition 5.3 says that a Liouville quantum gravity measure on an alternative underlying geometry (h still taken modulo additive constant) is stationary under zipping/unzipping in a modified version of capacity time.

This symmetry between space and time is intriguing for its own sake. It is related to but should not be confused with Theorem 4.5. Both theorems involve modifying the function $\mathfrak{h}_0$. However, Theorem 4.5 does not describe a new stationary process; instead, it gives a modified law for the zipping up map f_t and does not involve changing the time parameter. On the other hand, one can use Theorem 4.5 to describe how the process of Proposition 5.3 evolves in capacity time. The following is an immediate consequence of Theorem 4.5.

PROPOSITION 5.4. *Given a signed measure ρ on $\mathbb{H}$ with $(-\Delta^{-1}\rho, \rho) < \infty$, let Γ_ρ denote the measure whose Radon–Nikodym derivative with respect to Γ is proportional to $e^{(h^0, \rho)}$ (normalized to be a probability measure). Then the Γ_ρ law of the zipping up map f_t, for $t \geq 0$, is the law of the modified SLE process in Theorem 4.5. Moreover, given f_t for some time $t \geq 0$, the Γ_ρ conditional law of h^0 is that of $\tilde{h} \circ f_t + \hat{\mathfrak{h}}_t$, as defined in the statement of Theorem 4.5.*

5.4. *Conclusion of zipper proofs.*

PROPOSITION 5.5. *The conclusion of Proposition* 1.6 (*see Figure* 7) *remains true even if, when we choose* h *and* x, *we condition on a particular pair of values* $L_1 = \nu_h[a, x]$ *and* $L_2 = \nu_h[x, b]$.

PROOF. We use the setup and notation of Proposition 1.6. Recall the explicit form of the weighted measure $\nu_h[a, b]\, dh$ given in [29], as described earlier. Once we condition on x, the conditional law of h is simply that of a zero boundary GFF (except with free boundary conditions along $\mathbb{R}$) *plus* a random harmonic function *plus* $-\gamma \log|x - \cdot|$. Roughly speaking, the proposition follows from the fact that the restriction of h to an extremely small neighborhood of x (which tells us what the quantum surface looks like when we zoom in near x) is almost independent of the pair (L_1, L_2).

To express this point more carefully, consider smooth functions ϕ_1 and ϕ_2 on ∂D that are supported on disjoint neighborhoods $U_1 \subset D$ and $U_2 \subset D$ with $\overline{U}_1 \cap \mathbb{R} \subset (a, x)$ and $\overline{U}_2 \cap \mathbb{R} \subset (x, b)$. We assume (for $i \in \{1, 2\}$) that each ϕ_i is equal to one on some interval of $\overline{U}_i \cap \mathbb{R}$ and $\phi_i(z) \in [0, 1]$ for all $z \in D$, and also that $x \notin \overline{U}_1 \cup \overline{U}_2$.

Now, by the definition of the GFF we can write $h = \alpha_1\phi_1 + \alpha_2\phi_2 + h_0$ where α_1 and α_2 are centered Gaussian random variables and h_0 is projection of h onto the orthogonal complement of the span of ϕ_1 and ϕ_2, so that $(h_0, \phi_1)_\nabla = (h_0, \phi_2)_\nabla = 0$ almost surely. In this construction, α_1, α_2 and h_0 are independent of each other. Recalling (5.1), we have, for $i \in \{1, 2\}$, that the restriction of ν_h to ∂U_i is given by

$$\nu_h = e^{(\gamma/2)\alpha_i\phi_i}\, \nu_{h_0}.$$

In particular, this implies that once we condition on h_0, each $\nu_h(\partial U_i \cap \mathbb{R})$ is a.s. given by an increasing smooth function of α_i. Here, the smoothness can be verified by differentiating with respect to α and noting that no matter how many times one differentiates one obtains a compactly supported test function integrated against ν_{h_0}. This in particular implies that, once we condition on h_0, the quantity $\nu_h(\partial U_i \cap \mathbb{R})$ has a law which is absolutely continuous with respect to Lebesgue measure on (a, ∞) (for some a that depends on h_0) and has a smooth density function. Let ψ_1 and ψ_2 be the density functions for these laws. If we fix h_0 outside of some small $B_{\bar{\varepsilon}}(x)$ disjoint from $U_1 \cup U_2$ but rerandomize the restriction h_0' of h_0 to $B_{\bar{\varepsilon}}(x)$—conditioned on (L_1, L_2)—then this is the same as rerandomizing h_0' without conditioning on (L_1, L_2), except that we weight the law by (a quantity proportional to) the product below (viewed as a function of h_0')

$$\psi_1\big(L_1 - \nu_h([a, x] \setminus \partial U_1)\big)\psi_2\big(L_2 - \nu_h([x, b] \setminus \partial U_2)\big).$$

As $\bar{\varepsilon}$ tends to zero, the amount by which resampling h_0' changes either $\nu_h[a, x]$ or $\nu_h[x, b]$ is a quantity that tends to zero in probability. We conclude that weighting the law of h_0' by the big expression above (which is a smooth bounded function of $\nu_h[a, x]$ and $\nu_h[b, x]$) affects this law by an amount that (in total variation sense) tends to zero as $\varepsilon \to 0$. □

In order to prove Theorem 1.8, we will follow the argument sketched in Figure 8 in Section 1.6. We begin with a lemma.

LEMMA 5.6. *Suppose we first sample* (h^t, η^t) *from* $\nu_h[-\delta, 0]\Gamma$ (*normalized to be a probability measure*; *here, we write* $h = h^0$) *and then sample* x *from* $\nu_h[-\delta, 0]$. *Consider the joint law of the* x *and the* (h^t, η^t) *process chosen in this way. Then given* x, *the conditional law of the zipping up process* f_t *is that of the modified SLE process in Theorem 4.5, with the* ρ *measure given by* γ^2 *times a Dirac distribution at* x *minus* γ^2 *times a unit of uniformly distributed mass on the unit circle. Moreover* (*as in Proposition 5.4*), *given* f_t *for some time* $t \geq 0$, *the conditional law of* h^0 *is that of* $\tilde{h} \circ f_t + \hat{\mathfrak{h}}_t$, *as defined in the statement of Theorem 4.5.*

PROOF. Recall (1.2) and define

$$\nu_h^\varepsilon := \varepsilon^{\gamma^2/4} e^{\gamma h_\varepsilon(x)/2}\, dx. \tag{5.8}$$

Suppose that for some positive $\delta' < \delta$ and $\varepsilon < \delta'$ we weight by $\nu_h^\varepsilon[-\delta, -\delta']$ instead of $\nu_h[-\delta, 0]$—and then, given h, we choose x from ν_h^ε restricted to $[-\delta, -\delta']$ and normalized to be a probability measure. Then we know that this is the same as sampling from

$$e^{\gamma[h_\varepsilon(x) - h_1(0)]/2}\, dx\, dh \tag{5.9}$$

(times a normalizing constant) where x is Lebesgue on $[-\delta, \delta']$ and dh is the Γ law of h. [We have $h_\varepsilon(x) - h_1(0)$ instead of $h_\varepsilon(x)$ in the exponent because the former is defined independently of the additive constant for h, and is what we obtain if we choose the additive constant so that $h_1(0) = 0$.] Thus, given x, the conditional law of h is $e^{\gamma(h_\varepsilon(x) - h_1(0))/2}\, dh$, and Proposition 5.4 exactly determines the law of f_t [where the relevant ρ measure is a uniform measure on $\partial B_\varepsilon(x)$ minus a uniform measure on $\partial B_1(0)$].

The above gives an explicit way to construct a sample from $\nu_h^\varepsilon[-\delta, -\delta']$ and the corresponding f_t. To establish the lemma, we need to argue that an analogous result holds when ν_h^ε is replaced by ν_h. One way to deduce the ν_h result from the ν_h^ε result is as follows. Let U be the ε neighborhood of $[-\delta, -\delta']$ in $\mathbb{H}$, as in Figure 18. Decompose the free boundary GFF Hilbert space into the space Supp_U (the closure of the space of smooth functions supported on U that vanish on $\partial U \setminus \mathbb{R}$, with free boundary conditions on $\partial U \cap \mathbb{R}$) and the space Harm_U (the orthogonal complement of Supp_U, which consists of functions that are harmonic on U with

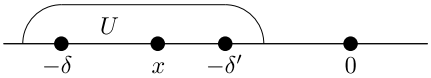

FIG. 18. $U \subset \mathbb{H}$ *is an* ε *neighborhood of* $[-\delta, -\delta']$.

Neumann boundary conditions along $\mathbb{R}$ [87]). Let h^{HARM} be the projection of h onto Harm_U and h^{SUPP} the projection of h onto Supp_U, so that

$$h = h^{\mathrm{HARM}} + h^{\mathrm{SUPP}}. \tag{5.10}$$

Roughly speaking, h^{HARM} represents the conditional expectation of h given all of the values of h outside of U. Note that h^{SUPP}, which is independent of h^{HARM}, has the law of a GFF on U with zero boundary conditions on $\partial U \setminus \mathbb{R}$ and free boundary conditions on $\partial U \cap \mathbb{R}$.

If we take any $\varepsilon' < \varepsilon$ then the measure $e^{\gamma h_{\varepsilon'}(x)/2}\,dh$ (normalized to be a probability measure) induces a law for h^{HARM} that does not depend on ε'. We claim that the Γ law of f_t [up until the first time T_U that a point in $f_t(U)$ reaches the origin] is independent of the projection of $h = h^0$ onto Supp_U. This follows from Theorem 1.2 and the way we constructed Γ: we may sample h and f_{T_U} by first sampling f_{T_U}, then sampling h_{T_U} from the law of $\tilde{h} + \mathfrak{h}_0$, and then using a coordinate transformation via $f_{T_U}^{-1}$ to recover h^0. The decomposition (5.10) can be made using $f_{T_U}(U)$ instead of U before we apply this coordinate transformation. It then follows from the conformal invariance of the zero boundary GFF (with free boundary conditions along $\mathbb{R}$) that the $h = h^0$ we obtain after the coordinate transformation has a projection to Supp_U that is independent of f_{T_U}.

Thus, the conditional law of f_{T_U} given h depends only on the projection of h onto Harm_U. Since weighting Γ by $\nu_h^{\varepsilon'}[-\delta, -\delta']$ and by $\nu_h[-\delta, -\delta']$ affects the law of the projection of h to Harm_U in an identical way, the law of this projection, and hence the law of f_t (up until T_U), is the same in each weighted measure. Since this holds for any interval $[-\delta, -\delta']$ we conclude that the joint law of x and f_t (up until T_U) does not change if we replace ν_h^{ε} with ν_h. To be explicit now about what we get from Proposition 5.4, recall from (4.6) and the subsequent discussion that we have $df_t(z) = \frac{-2}{f_t(z)}\,dt - dW_t$, where

$$dW_t = \bigl(-\mathrm{Re}(f_t)^{-1}, \rho\bigr)\,dt + \sqrt{\kappa}\,dB_t,$$

in the case that we weight by $e^{(h,(1/(2\sqrt{\kappa}))\rho)}$, the weight from (4.13). In our case, we take ρ to be such that

$$\left(h, \frac{1}{2\sqrt{\kappa}}\rho\right) = \frac{\gamma}{2}\bigl(h_\varepsilon(x) - h_1(0)\bigr),$$

which implies

$$(h, \rho) = \gamma^2\bigl(h_\varepsilon(x) - h_1(0)\bigr),$$

and the lemma follows by taking ε to zero. □

PROOF OF THEOREM 1.8. We begin by analyzing the construction described in Lemma 5.6 in more detail. By Lemma 5.6, we find that if h and x are taken from

the measure (5.9), then given x, the law of the zipping up procedure in Figure 8 [up until $f_t(x+\varepsilon)$ reaches zero] is determined by

$$dW_t = \left(-\frac{\gamma^2}{f_t(x)} + \operatorname{Re}\int_0^\pi \frac{\gamma^2}{f_t(e^{i\theta})}\frac{d\theta}{\pi}\right)dt + \sqrt{\kappa}\,dB_t. \tag{5.11}$$

As explained in the proof of Lemma 5.6, this law is independent of ε and also holds when ν_h^ε is replaced by ν_h in the construction of x and h above.

Now if the second term in the expression for the dt piece of (5.11) were not there, then (5.11) would correspond to the one-force-point $\mathrm{SLE}_{\kappa,\rho_1}$ with $\rho_1 = \gamma^2 = \kappa$, and the corresponding $f_t(x)$ would evolve precisely as a Bessel process of [recall (4.9)] dimension

$$1 + \frac{2(\gamma^2-2)}{\kappa} = 3 - \frac{4}{\gamma^2} < 2, \tag{5.12}$$

which would imply that $f_t(x)$ eventually reaches zero almost surely [78]. Moreover, by (4.12), the zipped up curve obtained, at the time x hit zero, would look like ordinary SLE up to some time with a size of order δ^2. For the next point in the proof, it is (slightly) easier if we replace the mean value on $\partial B_1(0)\cap\mathbb{H}$ [in Figure 8 and in (5.11)] with the mean value on an arc of $\partial B_1(0)\cap\mathbb{H}$ that is bounded away from $\mathbb{R}$. [The statement and proof of Lemma 5.6 do not change if we replace $\partial B_1(0)\cap\mathbb{H}$ with such an arc.] In this case, since $\operatorname{Re} f_t(z)$ is increasing in t for each $z\in\mathbb{H}$, the magnitude of the second term in the dt piece of (5.11) is a.s. bounded by a constant for all t. Thus, when we zoom in (as in Figure 8), the effective drift from this term in the zoomed-in process tends uniformly to zero.

We now claim that the limit sketched on the right in Figure 8 is a $(\gamma - 2/\gamma)$-quantum wedge with an independent SLE_κ curve. Lemma 5.6 implies that the limiting curve is an SLE_κ independent of the field. To see that the field corresponds to a quantum wedge, note that before rescaling it looks like a GFF with a $(\gamma - 2/\gamma)(-\log|\cdot|)$ singularity at the origin. (Recall that the -2γ corresponds to the log singularity present in the capacity invariant model, and the γ comes from the extra force point that has just collided with the origin.) The fact that the limit is then a quantum wedge follows from the argument in Proposition 1.6 (which shows generally that if h looks like a free boundary GFF plus a log singularity at some point, then one obtains a quantum wedge when one zooms in near that point). We also claim that in this limiting object [the $(\gamma - 2/\gamma)$-quantum wedge decorated by an independent SLE_κ curve] the two quantum surfaces divided by the curve each have the law of a γ-quantum wedge. This follows for the left side from Proposition 1.6 and for the right side by symmetry. [Note that a $(\gamma - 2\gamma)$-quantum wedge decorated by an independent SLE_κ is an object with left-right symmetry.]

At this point, we need to show that the quantum length measures along the two sides of η almost surely agree. Since x was sampled uniformly from quantum measure (before we zipped up and zoomed in), the SLE_κ decorated $(\gamma - 2/\gamma)$

quantum wedge must be invariant under the operation of unzipping by a fixed quantity of quantum boundary length as measured along the left of the two γ quantum wedges. (This is similar to the argument used to show Proposition 1.7.) Thus, the pair $((D_1, h_1), (D_2, h_2))$ is invariant under zipping and unzipping by the boundary length measure.

Let $F(s)$ denote the quantum length (as measured from the right side) of the curve segment $\eta([0,t])$, where t is chosen so that the quantum length (as measured from the left side) of $\eta([0,t])$ is s. The ergodic theorem implies that $\lim_{s\to\infty} F(s)/s$ exists almost surely; denote this (possibly random) limit by c. Since the law of the $(\gamma - 2/\gamma)$-quantum wedge is scale invariant, the fact that $F(s) \approx cs$ holds for large scales (by the ergodic theorem) implies that the same holds on all scales, and indeed we must have identically $F(s) = cs$. Since c is determined by the restriction of η and h to arbitrarily small balls centered at zero. In other words, if $\mathcal{F}_\delta$ denotes the σ algebra generated by the restriction of h to $B_\delta(0)$ and the curve η stopped when it first exits $B_\delta(0)$, then c is measurable w.r.t. the intersection $\bigcap_{\delta>0} \mathcal{F}_\delta$. However, it is easy to see that events in $\bigcap_{\delta>0} \mathcal{F}_\delta$ have probability zero or one, so that c must be a.s. constant. (This is proved explicitly in Lemma 8.2 of [23] in a setting that includes the field but not the path η. The statement including the path η then follows from the well-known fact that if $\mathcal{G}_\delta$ is generated by the restriction of Brownian motion to $[0,\delta]$ then events in $\bigcap_{\delta>0} \mathcal{G}_\delta$ have probability zero or one.) By the left-right symmetry of the law of SLE_κ and the quantum wedge, we have $c = 1$ almost surely.

Now that we know that the length measures on the two sides agree, we next claim that each of the two sides in the limit sketched on the right in Figure 8 are *independent* γ-quantum wedges. We will deduce this by applying Proposition 5.5 to Figure 8. More precisely, we consider the setting of the upper image in Figure 8 and we condition on the restriction of the GFF to the complement of $B^1 := B_{\bar\varepsilon}(x)$ and $B^2 := B_{\bar\varepsilon}(R(x))$, and we also condition on the quantum lengths $\nu_h(B^1 \cap \mathbb{R})$ and $\nu_h(B^2 \cap \mathbb{R})$. We may choose δ small enough so that with high probability $R(x) \in B_1(0)$, and we will focus our attention on the event that this is the case. If $\bar\varepsilon$ are small, then with high probability these balls do not contain the origin and are contained in $B_1(0)$.

If we resample the restriction of h to $B^1 \cap \mathbb{H}$—conditioned on $\nu_h(B^1 \cap [x,\infty))$ and $\nu_h(B_1 \cap (-\infty, x])$ and the restriction of h to the complement of B^1—then Proposition 5.5 implies that even with this conditioning it is a.s. the case that the zoomed-in figures (as in lower left in Figure 8) converges in law to a γ-quantum wedge as $\bar\varepsilon \to 0$. The conditional law of h (just given its values outside of B^1) is that of a random harmonic function plus a GFF on $B^1 \cap \mathbb{H}$ with free boundary conditions on $\mathbb{R}$ and zero boundary conditions on $\partial B^1 \cap \mathbb{H}$ [91].

The same applies when we zoom in near $R(x)$. Now if we condition on the GFF values on $\partial B_1(0)$ together with the values on the line $i\mathbb{R}$, together with the lengths of $[-1, x]$ and $[x, 0]$ and $[0, R(x)]$ and $[R(x), 1]$, then the conditional law of the restrictions of h to the two halfs of $B_1(0) \cap \mathbb{H}$ are independent by the standard

GFF Markov property. A similar analysis to the above shows that even with this conditioning, one still obtains the laws of quantum wedges, near each of x and $R(x)$, upon zooming in near those points. We conclude that in the $\bar{\varepsilon} \to 0$ limit the γ-quantum wedges are independent of each other. This completes the proof of the items in the first paragraph of the theorem statement. Along the way, we have also established the invariance of the law of the pair $((h_1, D_1), (h_2 D_2))$ under zipping up or down by a unit of quantum length zipper stationarity, and the $\mathcal{Z}_t^{\mathrm{LEN}}$ properties are now immediate from this. □

Deriving Theorem 1.3 as a consequence of Theorem 1.8 is now fairly straightforward. Essentially, one uses absolute continuity of the corresponding fields (at least when restricted to certain compact sets, away from the origin) to say that if the left and right boundary lengths a.s. agree in the setting of Theorem 1.8 then they almost surely agree in the setting of Theorem 1.3 as well.

PROOF OF THEOREM 1.3. The above proof shows that if one draws an independent SLE_κ on a $(\gamma - 2/\gamma)$-quantum wedge, the quantum lengths measured along the left and right sides of the curve agree almost surely. The setting of Theorem 1.3 is different because the law of h is different. On the other hand, we would like to argue that it is not *so* different, that is, that the laws of the two random fields (restricted to a compact set, bounded away from the boundary) are absolutely continuous w.r.t. each other, so that a statement that is a.s. true for one is a.s. for the other.

That is, we would like to say that the h of Theorem 1.3 "looks like" the canonical description h of a $(\gamma - 2/\gamma)$-quantum wedge, at least in the sense of absolute continuity of the restriction to compact subsets of $\mathbb{H}$. This is most immediate for compact sets that lie outside of $B_1(0)$. If we consider a canonical description h of a quantum wedge, then we know that by definition $\mu_h(B_1(0)) = 1$. However, once we condition on the restriction of h to $B_1(0)$, the conditional law of h restricted to $\mathbb{H} \setminus \overline{B_1(0)}$ is that of a GFF (with zero boundary conditions on $\partial B_1(0)$, free elsewhere) plus a certain random smooth function on $\mathbb{H} \setminus \overline{B_1(0)}$. [Note that if h is a canonical description, it will remain a canonical description even after we resample its values outside of $B_1(0)$ in this way.] In particular, the restriction of h to some compact $D \subset \mathbb{H} \setminus \overline{B_1(0)}$ is absolutely continuous with respect to the restriction to D of the h of Theorem 1.3, defined up to additive constant. (This is immediate from the absolutely continuity results in Section 3.1 of [87].) We thus find that in the setting of Theorem 1.3 the left and right quantum lengths along η are almost surely equal, at least outside of $B_1(0)$. The object in Theorem 1.3 is scale invariant; hence, if the left and right quantum boundary lengths a.s. agree outside of a unit ball of radius 1, then they a.s. agree outside of a unit ball of any radius, which means that they a.s. agree everywhere along the SLE_κ curve. □

REMARK 5.7. The above argument also holds if we add any smooth function to h—which affects the restriction of h to each compact subset of $\mathbb{H}$ in an absolutely continuous way (see [87, 91])—and then independently draw any path whose law is absolutely continuous w.r.t. that of SLE. That is, even in this setting one can use absolute continuity results about the GFF to prove that the quantum lengths along the two sides of the path agree almost surely. In particular, Theorems 1.3 and 1.4 still apply in the setting of Theorem 4.5.

6. Twenty questions. *Update*: *Since the first version of this paper was posted to the arXiv in* 2010, *there has been progress on several of the questions listed below. The current version leaves the questions as they were in* 2010 *but includes brief updates on work completed since.*

Many of the most fundamental questions about quantum gravity are open. Before formulating some of these questions, we present some definitions that will appear in the questions.

We have already seen that quantum wedges are natural random quantum surfaces of infinite area and infinite boundary length. We now describe some natural random quantum surfaces of unit area or unit boundary length, in terms of limits. We will say more about the sense in which the limits exist below, and more detailed constructions appear in [23].

1. Fix a smooth bounded domain D. Let h be a GFF with free boundary conditions on a linear segment L of ∂D and zero boundary conditions on $\partial D \setminus L$. Fix $C > 0$ and condition on $\mu_h(D) = C$. Let $\hat{h} = h - (\log C)/\gamma$, so that $\mu_{\hat{h}}(D) = 1$. We claim that as $C \to \infty$, the law of $(D, \hat{h})$ (viewed as a quantum surface) tends to a limit that does not depend on D or L. Call this the *unit area quantum disc*.

2. The *unit boundary length quantum disc* is defined the same way except that we condition on $\nu_h(L) = \sqrt{C}$ and the normalization gives $\nu_{\hat{h}}(L) = 1$.

3. The *unit area quantum sphere* is defined the same way as the unit area quantum disc except that we take D to have zero boundary conditions on all of D.

In the case of the discs (the first and second objects described above), one way to formulate the convergence is to fix a point $x \in D \setminus L$ and then to choose x_1 and x_2 on L at random from $\nu_{\hat{h}}$. We change coordinates via (1.3) to $\mathbb{H}$ so that x_1, x_2 and x map respectively to 0, 1 and ∞ (see the left side of Figure 19). We then obtain a random measure on $\mathbb{H}$, and these unit random measures should converge in law with respect to the topology of weak convergence of measures on $\mathbb{H}$.

In the case of the sphere, one can choose two points x_1 and x_2 from $\mu_{\hat{h}}$ and conformally map D to an origin-centered disc in such a way that one point goes to 0 and one to 1. (This determines the radius of the disc; see the right side of Figure 19.) We then obtain a random measure on $\mathbb{C}$ (which is incidentally supported on a disc of finite radius; the radius of this supporting disc tends to ∞ as $C \to \infty$), and these random measures should converge in law with respect to the topology of weak convergence of measures on $\mathbb{C}$.

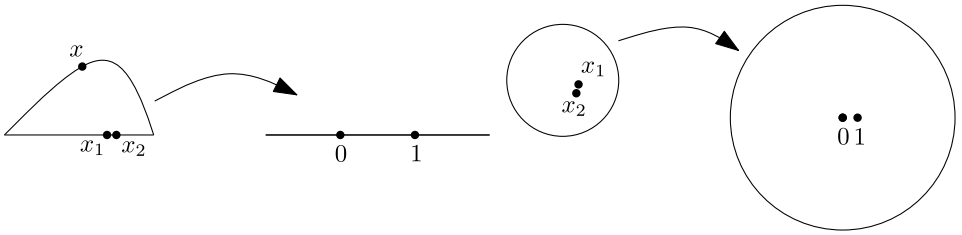

FIG. 19. *Constructing unit area/length quantum disc (left) or unit area quantum sphere (right). When C is large, we expect x_1 and x_2 to be close with high probability, and for most quantum area to be near those points.*

We now present some conjectures and questions. Some are very concrete and specific, and some are more speculative and open-ended. Some of the physics questions have of course been extensively discussed in the physics literature already, but this literature is far too broad for us to survey here (though we can at least mention Polyakov's reflections, with references, on the history of the subject in [76], as well as the extensive reference list in [29]).

Conjectures relevant to discrete model scaling limits:

1. Discrete random planar map models have Liouville quantum gravity as a scaling limit in the metric of weak convergence of area measures. To be more specific, in addition to the conjectures in Section 2 and in [29], we conjecture the following:

(a) The uniform quadrangulation of a sphere with n quadrilaterals scales to the unit area quantum sphere with $\gamma = \sqrt{8/3}$. One way to formulate this is that if we choose three points uniformly on the quadrangulation and conformally map the quadrangulated surface to $\mathbb{C}$ with these three points going to 0, 1 and ∞, then the image of the area measure in $\mathbb{C}$ (normalized to the total area is one) converges in law (w.r.t. the topology of the local weak convergence) to the unit area quantum sphere measure described above.

(b) The random quadrangulation of a disc with boundary length n, where the probability of a quadrangulation is proportional to a (critical) constant to the number of quadrilaterals, scales to the unit boundary length quantum disc with $\gamma = \sqrt{8/3}$.

(c) The random quadrangulation of a disc with n quadrilaterals, where the probability of a quadrangulation is proportional to a (critical) constant to the boundary length, scales to the unit area quantum disc with $\gamma = \sqrt{8/3}$.

(d) Similar statements hold for random quadrangulations weighted by the partition functions of Ising configurations, $O(n)$ configurations, uniform spanning trees, etc., with $\gamma = \sqrt{\kappa}$ for the appropriate κ. In each case, the cluster boundary loops (of the statistical model on the random surface) scale to a CLE_κ independent of the measure. (Question: What happens if, for some d, we make

the probability of a quadrangulation proportional to the partition function of the d-dimensional discrete Gaussian-free field on the corresponding graph?)

Update: *The problem as stated remains open. However, convergence of FK-decorated random planar maps to CLE-decorated LQG has been established in the so-called peanosphere topology and in the so-called loop structure topology, in both infinite and finite volume settings* [24, 34–37, 70, 93]; *see also related results in* [11, 17, 60, 101]. *The construction of finite volume LQG spheres and disks* (*briefly considered above*) *is discussed in more detail in* [4, 19, 24, 70]. *Partial results in the direction of understanding the conformal structure of the discrete models appear in* [18]. *Roman Boikii and Stanislav Smirnov have also described progress in private communication*.

2. A quantum surface is a coordinate-change-invariant random metric space in the sense of the footnote in Section 1.2. In the special case $\gamma = \sqrt{8/3}$, this is equivalent to the *Brownian map* (see, e.g., [53, 56, 62]), which is a term given to the random metric space scaling limit of uniform quadrangulations on the sphere. *Update*: *The metric space structure of LQG has now been constructed in the pure quantum gravity case* $\gamma = \sqrt{8/3}$ *in recent work* (*some in progress*) *showing that* $\sqrt{8/3}$*-LQG is equivalent to the Brownian map* [68, 69, 71, 72]. *No such construction is currently available for other values of* $\gamma \in (0, 2)$.

3. Assuming that 2 holds, the scaling limits in 1 also hold in the topology of Gromov–Hausdorff convergence of metric spaces. *Update: In light of the equivalence of* $\sqrt{8/3}$*-LQG and the Brownian map* [68, 69, 71, 72]*, and the Gromov–Hausdorff convergence of discrete random planar maps to the Brownian map* [54, 55, 63]*, this conjecture has now been established in the pure quantum gravity case* $\gamma = \sqrt{8/3}$. *It remains open for all other* $\gamma \in (0, 2)$.

Related but more modest requests:

4. Show that simple random walk on a random quadrangulated surface is a good approximation for Brownian motion on that surface. *Update*: *This problem remains open. However, works by Gill and Rohde, by Gurel-Gurevich and Nachmias and by Benjamini and Curien have derived some fundamental properties for random walks and Brownian motion on infinite volume versions of these surfaces* [8, 32, 33].

5. Prove any SLE convergence result at all for random planar maps. *Update*: *As mentioned above, there are now a number of convergence results involving the peanosphere and loop structure topologies.*

6. Show that the following toy model has a scaling limit which is a metric space. Begin with a unit square S. After an exponentially random amount of time, divide S into new four squares of equal size. Each time a new square is created, give it a new exponential clock (independent of the others) and so that it too will divide into four new squares after an exponentially random amount of time. [Each time a clock at an existing box rings, the total number of boxes increases by 3.

Thus, if $N(t)$ is the number of boxes at time t, then $e^{-3t}N(t)$ is a martingale.] After time t, consider two squares to be adjacent if they have part of an edge in common. Show that for some $\beta > 0$, the graph metric on the set of squares times $e^{-\beta t}$ converges almost surely to a random metric space parameterized by S. *Update: This toy problem remains completely unsolved. However, we remark that another interesting model for producing random squarings appears in* [1].

Structural questions about Liouville quantum gravity and AC geometry:

7. If a quantum surface is well defined as a metric space, what does the boundary of a unit ball look like? Is it a form of SLE (some collection of SLE loops) or something completely different? How about a shortest (geodesic) path between two points z_1 and z_2? How about the boundary of the set of points closer to z_1 than to z_2? What is the continuum analog of the breadth first search tree (the tree that appears in the Schaeffer bijection [84])? Can one even formulate a conjecture? *Update*: *As mentioned above*, *much has been now been established in the case* $\gamma = \sqrt{8/3}$, *where the ball boundaries are constructed using the quantum Loewner evolution* [68, 69, 71, 72], *and a canonical embedding of the Brownian map in the sphere has been constructed. The problem remains completely open for other* $\gamma \in (0, 2)$.

8. If $\kappa \in (4, 8)$, then the complement of SLE_κ, run to time T, is not a single simply connected domain; it consists of infinitely many components. The set of components that lie to the left side of η comes with a tree-like hierarchical structure (where a component A_1 is "above" a component A_2 if η traces ∂A_1 after it has started tracing, but before it has finished tracing, ∂A_2). The set of components on the right side comes with a similar structure. This suggest that when we "unzip" along η (as in Theorem 1.2) we obtain not merely a quantum surface parameterized by $\mathbb{H}$ but a quantum surface parameterized by $\mathbb{H}$ together with a tree-like structure of quantum-surface "beads" hanging off of its boundary. Does an analog of Theorem 1.3 hold in this setting if one properly defines the boundary length of the tree-like structure? How about an analog of Theorem 1.4? Proving the latter would likely require first showing that SLE_κ is removable when $\kappa \in (4, 8)$. Is this the case? Can a $\kappa \in (4, 8)$ version of the quantum zipper be used to prove the time reversal symmetry of SLE for $\kappa \in (4, 8)$? *Update*: *The time-reversal symmetry results were established in the imaginary geometry papers* [64–67], *and the questions about the* $\kappa \in (4, 8)$ *analogs of Theorems* 1.3 *and* 1.4 *were answered affirmatively in* [24]. *Interestingly*, *the zipping up*" *results in* [24] *were established without a proof that* SLE_κ *is removable when* $\kappa \in (4, 8)$. *It remains an open question whether these* SLE_κ *processes are removable. A new result on the inversion symmetry of the quantum zipper welding* (*for the* $\kappa < 4$ *case*) *appears in* [83].

9. What is the Hausdorff dimension of a quantum surface (assuming it is defined as metric space) for general γ? Can one at least handle the special case $\gamma = \sqrt{8/3}$ (where the answer should be 4, by analogy to discrete random triangulations; see e.g. [3])? *Update*: *This question is settled in the case* $\gamma = \sqrt{8/3}$, *in*

light of the above-mentioned equivalence between the $\sqrt{8/3}$*-LQG sphere and the Brownian map. In that case, the Hausdorff dimension is indeed* 4. *For general* γ, *there is a Hausdorff dimension conjecture due to Watabiki* [103], *which has found some support in recent simulations* [2]. *This conjecture is further explained and discussed in* [68].

10. Let h^ε be projections of h onto the space of functions that are piecewise linear on an ε-edge-length triangular lattice (or alternatively, mollifications of h obtained by convolving with a bump function supported on a disc of radius ε). Do the flow lines of $e^{ih^\varepsilon/\chi}$ converge in law to SLE curves (AC geometry flow lines) as $\varepsilon \to 0$? If so, this would extend the main result of [86] to $\kappa \neq 4$. Do the other results of [86, 87] for $\kappa = 4$ extend to $\kappa \neq 4$?

11. If h^ε are smooth mollifications or piecewise linear projections (as above) of the *free* boundary GFF of Theorem 1.2, do the curves obtained by zipping up part of the boundary via conformal welding (using the identification map R as in Section 1.4, except that we use the approximate measures $e^{\gamma h^\varepsilon(x)/2}\,dx$ instead of ν_h to define R) converge in law to SLE?

12. What happens when $\gamma = 2$ (so $\kappa = 4$, $Q = 2$)? The definition of a quantum surface as an equivalence class—as given in Section 1.2—makes sense for $\gamma = 2$, and as mentioned in the Introduction. Is the "zipping up" map in the coupling of Theorem 1.2 determined by h, as in the $\gamma < 2$ case? Is it the limit of the curves one obtains by using the same GFF $\tilde{h}$ but with different values of γ, and letting γ approach 2? Is SLE_4 almost surely removable, like SLE_κ for $\kappa < 4$? *Update*: *This question remains open. However, recent results enable one to construct a quantum length measure* ν_h (*as well as a quantum area measure* μ_h) *in the* $\gamma = 2$ *case* [25, 26]. *It is natural to conjecture that zipping up gives a welding that respects this measure, and also that* SLE_4 *is removable. We note that a closely related result has been established by Tecu in* [102], *which builds on the work of* [6].

13. Can the continuum limit of simple random walk on a random quadrangulation (or of Brownian motion on the corresponding Riemannian surface) be somehow understood directly, in terms of the Schaeffer or Mullin bijections, as a random process on an identified pair of continuum trees? *Update*: *The results in* [24] *and* [68, 69, 71, 72] *allow one to define a conformal structure* (*and hence a Brownian motion, with a time change defined using the* "*Liouville Brownian motion*" *theory developed in* [10, 31]) *on these mated pairs of continuum trees. The reader may decide whether these* (*SLE/LQG/QLE-based*) *constructions ought to be considered* "*direct*."

14. What is the most natural formulation of a higher genus quantum surface, where the conformal modulus is allowed to be random? What is the right conjectural scaling limit of a random quadrangulation of a torus (or higher genus surface) with n faces, as $n \to \infty$? How about a surface with n holes (weighting by the combined length of the hole boundaries in a critical way)? One way to construct a quantum surface with holes is to put an independent conformal loop ensemble on top of a quantum surface and then cut out the regions surrounded by some of

the loops. (Here, one would have to give a rule for specifying which loops to cut out, for example, one might try to cut out the k loops with the longest quantum boundary lengths.) Is it possible to weld together surfaces constructed this way and produce higher genus surfaces that answer the questions above? *Update*: *See* [81] *for some recent work on LQG tori by David, Rhodes and Vargas, which cites many earlier works from the physics literature on higher genus LQG constructions. It remains an open problem to implement the ideas mentioned above* (*i.e., to construct canonical higher-genus LQG random surfaces via weldings*) *and to show that these approaches are equivalent to those discussed in* [81].

15. One way to combine Liouville quantum gravity and AC geometry is to let h be a complex valued free field. The real part of h encodes Liouville quantum gravity and the imaginary part an AC geometry, which could be coupled with a conformal loop ensemble on the quantum surface [92]. In this context, what new meaning, if any, comes out of the extra symmetries of the complex Gaussian free field (e.g., its invariance w.r.t. replacing h with a modulus-one complex constant times h)?

16. Can one fully construct all rays in the AC geometry of an instance of the GFF (starting from all points in D) and determine which rays intersect each other and themselves, etc.? Does the list of properties shown to hold for smooth h in the Appendix (Proposition A.1) hold in the GFF setting as well? *Update*: *This question has been affirmatively answered in* [64–67].

Returning to the original physics motivation...

17. Can one construct a variant of Liouville quantum gravity with some of the additional complexities of a physical string theory (and figure out what the following things mean in this context mathematically)?

(a) Minkowski space time.
(b) Complex weights (in place of a probability measure).
(c) Supersymmetry.
(d) Gauge theories.
(e) Embeddings in Calabi–Yau manifolds.
(f) Quantum surfaces of nondeterministic genus.

18. When introducing Liouville quantum gravity, Polyakov wrote in 1981 that it was necessary to develop a theory of random surfaces "because today gauge invariance plays the central role in physics. Elementary excitations in gauge theories are formed by the flux lines (closed in the absence of charges) and the time development of these lines forms the world surfaces. All transition amplitude(s) are given by the sums over all possible surfaces with fixed boundary." [74] Can one formulate any version of this gauge/string duality as a theorem or conjecture relating gauge theory flux lines to Liouville quantum gravity? *Update*: *We mention one recent rigorous formulation of Yang–Mills gauge-string duality on a lattice due to Chatterjee* (*not obviously related to Liouville quantum gravity*) [15].

19. Can any Liouville quantum gravity insights be used to study higher dimensional random metrics? What are the most natural models? *Update*: *We remark that certain higher dimensional analogs are easy to describe. If h is an instance of the log-correlated Gaussian field (LGF) as defined, for example, in the surveys* [27, 58], *then $e^{\gamma h(z)}\,dz$ is easy to define as a random measure when γ is in the right range*; *see, for example, the survey* [80]. *Furthermore, the restriction of a higher dimensional LGF to a (two-dimensional) plane yields a GFF on the plane, so that the higher dimensional LGF can be interpreted as a coupling of planar GFFs, one for each planar subspace*; *it remains unclear whether the imaginary geometry or level set structures corresponding to the individual slices can be unified in a coherent way* [58].

20. Are there any natural three- or four-dimensional versions of AC geometry? For example, is it possible to interpret paths in the three-dimensional uniform spanning tree scaling limit [47] as geodesics of a random affine connection? *Update*: *This remains open. We mention one attempt to describe a three-dimensional analog of SLE_6 (at least on the discrete level) via so-called tricolor percolation* [95], *which also references relevant field theoretic constructions from the physics literature*.

APPENDIX: AC GEOMETRY: REAL VS. IMAGINARY GAUSSIAN CURVATURE

This section gives a formal definition/interpretation/explanation of the AC geometry (and its relationship to Liouville quantum gravity) when $h : D \to \mathbb{R}$ is smooth. Denote by $\mathbb{F}(h, D)$ the collection of curves that are flow lines of $e^{i(h/\chi+c)}$ (beginning at a point in $\overline{D}$), for some $c \in [0, 2\pi)$. For each $x \in D$, there is a one-parameter family of "rays" in $\mathbb{F}(h, D)$ (indexed by c) that begin at x and ultimately hit the boundary of D. We refer to c as the *angle* of the ray. Two rays of D are *parallel* if they have the same angle. When h is constant, these are the rays of Euclidean geometry. In general, the collection of rays is quite interesting and satisfies some of the axioms describing rays in Euclidean geometry. For example, the reader may verify the following.

PROPOSITION A.1. *If h is a Lipschitz function on a simply connected domain $D \subset \mathbb{C}$, then the following hold*:

1. *Each ray beginning at $x \in D$ is a differentiable simple path (i.e., it does not intersect itself).*
2. *Two parallel rays in $\mathbb{F}(h, D)$ are disjoint unless one is a subset of the other.*
3. *Two nonparallel rays in $\mathbb{F}(h, D)$ intersect at most once in D.*
4. *If $x \in D$ lies on a ray beginning at $y \in D$, then y lies on a ray of opposite direction beginning at x.*
5. *The sum of the angles of a "triangle"—whose edges are segments of rays in $\mathbb{F}(h, D)$—is always π.*

When h is the GFF, the "flow line" described by the coupling in Theorem 1.1 begins at a special location on the boundary of D, but in principle one would like to make sense of the entire set $\mathbb{F}(h, D)$ when h is an instance of the Gaussian free field and χ is a positive constant. Even when h is smooth, the AC geometry is not a Euclidean or non-Euclidean geometry in the usual sense; in particular, it does not come with a notion of length or distance. We will interpret the AC rays of a smooth h as the geodesics of a random torsion-free (defined below) affine connection that is naturally dual to the Levi–Civita connection of a Liouville quantum surface.

Recall that an *affine connection* determines, for any smooth path segment P in D, an orthogonal map on the two-dimensional tangent space, which we may represent as multiplication by a complex number $\alpha(P)$. Intuitively, it describes the way a small object transforms under parallel transport ("sliding without rotating") along the path P. The Levi–Civita connection for a Riemannian metric is (by definition) the unique metric-preserving and torsion-free affine connection. For a two-dimensional metric conformally parameterized by a subset of D with area measure $e^{\gamma h(z)}\,dz$, metric preservation implies that

$$\log|\alpha(P)| = \frac{\gamma}{2}h(z_1) - \frac{\gamma}{2}h(z_2), \tag{A.1}$$

where z_1 and z_2 are the first and last endpoints of P. In particular, this quantity is independent of the trajectory P takes between z_1 and z_2.

We recall the following standard fact [29] (see also keywords "isothermal coordinates" and "Gaussian curvature" in any text on Riemannian surfaces): write $\lambda = \gamma h$ and note that given a measurable subset A of D, the integral $\int_A e^{\lambda(z)}\,dz$ (where dz denotes Lebesgue measure on D) is the area of the portion of $\mathcal{M}$ parameterized by A. The function $K = -e^{-\lambda}\Delta\lambda$ (where $\Delta\lambda := \lambda_{xx} + \lambda_{yy}$ is the Laplacian operator) is called the *Gaussian curvature* of $\mathcal{M}$. If A is a measurable subset of the (x, y) parameter space, then the integral of the Gaussian curvature with respect to the portion of $\mathcal{M}$ parameterized by A can be written $\int_A e^{\lambda(z)}K(z)\,dz = \int_A -\Delta\lambda(z)\,dz$ where dz denotes Lebesgue measure on D. In other words, $-\Delta\lambda$ gives the density of Gaussian curvature in the isothermal coordinate space. In particular, $\mathcal{M}$ is flat if and only if h is harmonic.

We may define $\mathcal{A}(z, v)$ for each $z \in D$ and $v \in \mathbb{C}$ so that if P is parameterized by $t \in [0, 1]$ we have

$$\log\alpha(P) = \int_0^1 \mathcal{A}\left(P(t), \frac{\partial}{\partial t}P(t)\right)dt.$$

We may identify $\mathbb{C}$ with $\mathbb{R}^2$ so that ∇h is a vector-valued function of $\mathbb{R}^2$; given $v \in \mathbb{C}$, let $[\nabla h \cdot v]$ denote the dot product of this vector with the vector v (viewed as an element of $\mathbb{R}^2$), which is a real number.

In the case of the Levi–Civita connection, we have $\operatorname{Re}\mathcal{A}(z, v) = \frac{-\gamma}{2}[\nabla h \cdot v]$. Given this, $\operatorname{Im}\mathcal{A}(z, v)$ is determined by the requirement that the connection is

torsion-free. Using this notation, the statement that the connection is torsion-free means that $\mathcal{A}(z, iv) = -i\mathcal{A}(z, v)$ for all z and v, which implies

$$\mathcal{A}(z, v) = \frac{-\gamma}{2}[\nabla h \cdot v] + i\frac{-\gamma}{2}[\nabla h \cdot iv].$$

If P is the boundary of a smooth region $R \subset D$, then Green's theorem implies that

$$\log \alpha(P) = i \int_R -\frac{-\gamma}{2}\Delta h(z)\, dz. \tag{A.2}$$

The RHS of (A.2) is (up to a constant imaginary multiplicative factor) the integral of the Gaussian curvature over the region of the surface parameterized by R. In fact, one may even *define* Gaussian curvature to be the function for which this the case.

The AC rays are the geodesics of a connection satisfying an analog of (A.1):

$$\arg \alpha(P) = h(z_2)/\chi - h(z_1)/\chi, \tag{A.3}$$

and if we require that this connection also be torsion-free we obtain

$$\mathcal{A}(z, v) = i[\nabla h \cdot v]/\chi - [\nabla h \cdot iv]/\chi,$$

and hence, if P is the boundary of a smooth region R, then

$$\log \alpha(P) = \int_R (-\Delta h(z)/\chi)\, dz \in \mathbb{R}.$$

Intuitively, this means that a small object sliding around the path P will undergo no net rotation, but will change in size (the opposite of what happens in a Levi–Civita connection of a Riemannian surface). The definition of Gaussian curvature density suggested above, in terms of (A.2), suggests an imaginary Gaussian curvature proportional to the real curvature obtained in Liouville quantum gravity.

We remark that the Laplacian of the GFF has a natural interpretation as the charge density of a two-dimensional Coulomb gas (see the discussion and references in the introduction to [86]). Thus, Liouville quantum gravity can be interpreted as imposing a Gaussian curvature density equal to a real constant times this charge density. AC geometry is the same, except that the constant is required to be imaginary.

Acknowledgments. For valuable insights, helpful conversations, and comments on early drafts, the author thanks Julien Dubédat, Bertrand Duplantier, Richard Kenyon, Greg Lawler, Jason Miller, Curtis McMullen, Steffen Rohde, Oded Schramm, Stanislav Smirnov, Fredrik Johansson Viklund, Wendelin Werner and David Wilson. The author also thanks the organizers and participants of the PIMS 2010 Summer School in Probability (Seattle) and the Clay Mathematics Institute 2010 Summer School (Búzios, Brazil), where drafts of this paper were used as lecture notes for courses. We also thank Nathanaël Berestycki, Ewain Gwynne, Xin Sun and a reading group at the Newton Institute in Cambridge for helpful comments.

REFERENCES

[1] ADDARIO-BERRY, L. (2014). Random infinite squarings of rectangles. Preprint. Available at arXiv:1405.2870.

[2] AMBJØRN, J. and BUDD, T. G. (2014). Geodesic distances in Liouville quantum gravity. *Nuclear Phys. B* **889** 676–691. MR3280382

[3] ANGEL, O. and SCHRAMM, O. (2003). Uniform infinite planar triangulations. *Comm. Math. Phys.* **241** 191–213. MR2013797

[4] ARU, J., HAUNG, Y. and SUN, X. (2015). Two perspectives of the unit area quantum sphere and their equivalence. In preparation.

[5] ASTALA, K., JONES, P., KUPIAINEN, A. and SAKSMAN, E. (2010). Random curves by conformal welding. *C. R. Math. Acad. Sci. Paris* **348** 257–262. MR2600118

[6] ASTALA, K., JONES, P., KUPIAINEN, A. and SAKSMAN, E. (2011). Random conformal weldings. *Acta Math.* **207** 203–254. MR2892610

[7] BELIAEV, D. and SMIRNOV, S. (2009). Harmonic measure and SLE. *Comm. Math. Phys.* **290** 577–595. MR2525631

[8] BENJAMINI, I. and CURIEN, N. (2013). Simple random walk on the uniform infinite planar quadrangulation: Subdiffusivity via pioneer points. *Geom. Funct. Anal.* **23** 501–531. MR3053754

[9] BENJAMINI, I. and SCHRAMM, O. (2009). KPZ in one dimensional random geometry of multiplicative cascades. *Comm. Math. Phys.* **289** 653–662. MR2506765

[10] BERESTYCKI, N. (2015). Diffusion in planar Liouville quantum gravity. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** 947–964. MR3365969

[11] BERESTYCKI, N., LASLIER, B. and RAY, G. (2015). Critical exponents on Fortuin–Kastelyn weighted planar maps. Preprint. Available at arXiv:1502.00450.

[12] BERNARDI, O. (2007). Bijective counting of tree-rooted maps and shuffles of parenthesis systems. *Electron. J. Combin.* **14** 36 pp. (electronic). MR2285813

[13] BISHOP, C. J. (2007). Conformal welding and Koebe's theorem. *Ann. of Math.* (2) **166** 613–656. MR2373370

[14] CAMIA, F. and NEWMAN, C. M. (2008). SLE_6 and CLE_6 from critical percolation. In *Probability, Geometry and Integrable Systems. Math. Sci. Res. Inst. Publ.* **55** 103–130. Cambridge Univ. Press, Cambridge. MR2407594

[15] CHATTERJEE, S. (2015). Rigorous solution of strongly coupled $SO(N)$ lattice gauge theory in the large N limit. Preprint. Available at arXiv:1502.07719.

[16] CHELKAK, D. and SMIRNOV, S. (2012). Universality in the 2D Ising model and conformal invariance of fermionic observables. *Invent. Math.* **189** 515–580. MR2957303

[17] CHEN, L. (2015). Basic properties of the infinite critical-FK random map. Preprint. Available at arXiv:1502.01013.

[18] CURIEN, N. (2015). A glimpse of the conformal structure of random planar maps. *Comm. Math. Phys.* **333** 1417–1463. MR3302638

[19] DAVID, F., KUPIAINEN, A., RHODES, R. and VARGAS, V. (2014). Liouville quantum gravity on the Riemann sphere. Preprint. Available at arXiv:1410.7318.

[20] DUBÉDAT, J. (2009). SLE and the free field: Partition functions and couplings. *J. Amer. Math. Soc.* **22** 995–1054. MR2525778

[21] DUPLANTIER, B. (1998). Random walks and quantum gravity in two dimensions. *Phys. Rev. Lett.* **81** 5489–5492. MR1666816

[22] DUPLANTIER, B. and KOSTOV, I. (1988). Conformal spectra of polymers on a random surface. *Phys. Rev. Lett.* **61** 1433–1437. MR0960093

[23] DUPLANTIER, B., MILLER, J. and SHEFFIELD, S. (2014). Liouville quantum gravity as a mating of trees. Preprint. Available at arXiv:1409.7055.

[24] DUPLANTIER, B., MILLER, J. and SHEFFIELD, S. (2014). Liouville quantum gravity as a mating of trees. Preprint. Available at arXiv:1409.7055.
[25] DUPLANTIER, B., RHODES, R., SHEFFIELD, S. and VARGAS, V. (2012). Renormalization of critical Gaussian multiplicative chaos and kpz formula. Preprint. Available at arXiv:1212.0529.
[26] DUPLANTIER, B., RHODES, R., SHEFFIELD, S. and VARGAS, V. (2014). Critical Gaussian multiplicative chaos: Convergence of the derivative martingale. *Ann. Probab.* **42** 1769–1808. MR3262492
[27] DUPLANTIER, B., RHODES, R., SHEFFIELD, S. and VARGAS, V. (2014). Log-correlated Gaussian fields: An overview. Preprint. Available at arXiv:1407.5605.
[28] DUPLANTIER, B. and SHEFFIELD, S. (2009). Duality and the Knizhnik–Polyakov–Zamolodchikov relation in Liouville quantum gravity. *Phys. Rev. Lett.* **102** 150603, 4. MR2501276
[29] DUPLANTIER, B. and SHEFFIELD, S. (2011). Liouville quantum gravity and KPZ. *Invent. Math.* **185** 333–393. MR2819163
[30] DUPLANTIER, B. and SHEFFIELD, S. (2011). Schramm–Loewner evolution and Liouville quantum gravity. *Phys. Rev. Lett.* **107** 131305.
[31] GARBAN, C., RHODES, R. and VARGAS, V. (2014). On the heat kernel and the Dirichlet form of Liouville Brownian motion. *Electron. J. Probab.* **19** no. 96, 25. MR3272329
[32] GILL, J. T. and ROHDE, S. (2013). On the Riemann surface type of random planar maps. *Rev. Mat. Iberoam.* **29** 1071–1090. MR3090146
[33] GUREL-GUREVICH, O. and NACHMIAS, A. (2013). Recurrence of planar graph limits. *Ann. of Math.* (2) **177** 761–781. MR3010812
[34] GWYNNE, E., MAO, C. and SUN, X. (2015). Scaling limits for the critical Fortuin–Kasteleyn model on a random planar map I: Cone times. Preprint. Available at arXiv:1502.00546.
[35] GWYNNE, E. and MILLER, J. (2015). Convergence of the topology of critical Fortuin–Kasteleyn planar maps to that of CLE_κ on a Liouville quantum surface. In preparation.
[36] GWYNNE, E. and SUN, X. (2015). Scaling limits for the critical Fortuin–Kastelyn model on a random planar map II: Local estimates and empty reduced word exponent. Preprint. Available at arXiv:1505.03375.
[37] GWYNNE, E. and SUN, X. (2015). Scaling limits for the critical Fortuin–Kastelyn model on a random planar map III: Finite volume case. In preparation.
[38] HAGENDORF, C., BERNARD, D. and BAUER, M. (2010). The Gaussian free field and SLE_4 on doubly connected domains. *J. Stat. Phys.* **140** 1–26. MR2651436
[39] IZYUROV, K. and KYTÖLÄ, K. (2013). Hadamard's formula and couplings of SLEs with free field. *Probab. Theory Related Fields* **155** 35–69. MR3010393
[40] JONES, P. Private communciation.
[41] JONES, P. W. and SMIRNOV, S. K. (2000). Removability theorems for Sobolev functions and quasiconformal maps. *Ark. Mat.* **38** 263–279. MR1785402
[42] KANG, N.-G. (2007). Boundary behavior of SLE. *J. Amer. Math. Soc.* **20** 185–210. MR2257400
[43] KENYON, R. (2000). Conformal invariance of domino tiling. *Ann. Probab.* **28** 759–795. MR1782431
[44] KENYON, R. (2001). Dominos and the Gaussian free field. *Ann. Probab.* **29** 1128–1137. MR1872739
[45] KENYON, R. (2008). Height fluctuations in the honeycomb dimer model. *Comm. Math. Phys.* **281** 675–709. MR2415464
[46] KNIZHNIK, V. G., POLYAKOV, A. M. and ZAMOLODCHIKOV, A. B. (1988). Fractal structure of 2D-quantum gravity. *Modern Phys. Lett. A* **3** 819–826. MR0947880
[47] KOZMA, G. (2007). The scaling limit of loop-erased random walk in three dimensions. *Acta Math.* **199** 29–152. MR2350070

[48] LAWLER, G. F. (2009). Conformal invariance and 2D statistical physics. *Bull. Amer. Math. Soc.* (*N.S.*) **46** 35–54. MR2457071

[49] LAWLER, G. F. and REZAEI, M. A. (2015). Minkowski content and natural parameterization for the Schramm–Loewner evolution. *Ann. Probab.* **43** 1082–1120. MR3342659

[50] 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

[51] LAWLER, G. F. and SHEFFIELD, S. (2011). A natural parametrization for the Schramm–Loewner evolution. *Ann. Probab.* **39** 1896–1937. MR2884877

[52] LAWLER, G. F. and ZHOU, W. (2013). *SLE* curves and natural parametrization. *Ann. Probab.* **41** 1556–1584. MR3098684

[53] LE GALL, J.-F. (2007). The topological structure of scaling limits of large planar maps. *Invent. Math.* **169** 621–670. MR2336042

[54] LE GALL, J.-F. (2013). Uniqueness and universality of the Brownian map. *Ann. Probab.* **41** 2880–2960. MR3112934

[55] LE GALL, J.-F. (2014). The Brownian map: A universal limit for random planar maps. In *XVIIth International Congress on Mathematical Physics* 420–428. World Sci. Publ., Hackensack, NJ. MR3204495

[56] LE GALL, J.-F. and PAULIN, F. (2008). Scaling limits of bipartite planar maps are homeomorphic to the 2-sphere. *Geom. Funct. Anal.* **18** 893–918. MR2438999

[57] LIND, J. R. (2008). Hölder regularity of the SLE trace. *Trans. Amer. Math. Soc.* **360** 3557–3578. MR2386236

[58] LODHIA, A., SHEFFIELD, S., SUN, X. and WATSON, S. S. (2014). Fractional Gaussian fields: A survey. Preprint. Available at arXiv:1407.5598.

[59] MAKAROV, N. and SMIRNOV, S. (2010). Off-critical lattice models and massive SLEs. In *XVIth International Congress on Mathematical Physics* 362–371. World Sci. Publ., Hackensack, NJ. MR2730811

[60] MAO, C., SUN, X. and ZHOU, T. (2014). Inventory accumulation with k products. Available at http://math.mit.edu/research/undergraduate/urop-plus/documents/Zhou.pdf.

[61] MARSHALL, D. E. and ROHDE, S. (2007). Convergence of a variant of the zipper algorithm for conformal mapping. *SIAM J. Numer. Anal.* **45** 2577–2609 (electronic). MR2361903

[62] MIERMONT, G. (2008). On the sphericity of scaling limits of random planar quadrangulations. *Electron. Commun. Probab.* **13** 248–257. MR2399286

[63] MIERMONT, G. (2013). The Brownian map is the scaling limit of uniform random plane quadrangulations. *Acta Math.* **210** 319–401. MR3070569

[64] MILLER, J. and SHEFFIELD, S. (2016). Imaginary Geometry I: Interacting SLEs *Probab. Theory Related Fields* **164** 553–705. MR3477777

[65] MILLER, J. and SHEFFIELD, S. (2016). Imaginary geometry II: Reversibility of $\mathrm{SLE}_\kappa(\rho_1;\rho_2)$ for $\kappa \in (0,4)$. *Ann. Probab.* **44** 1647–1722. MR3502592

[66] MILLER, J. and SHEFFIELD, S. (2012). Imaginary geometry III: Reversibility of SLE_κ for $\kappa \in (4,8)$. *Ann. Math.* To appear.

[67] MILLER, J. and SHEFFIELD, S. (2013). Imaginary geometry IV: Interior rays, whole-plane reversibility, and space-filling trees. Preprint. Available at arXiv:1302.4738.

[68] MILLER, J. and SHEFFIELD, S. (2013). Quantum Loewner Evolution. Preprint. Available at arXiv:1312.5745.

[69] MILLER, J. and SHEFFIELD, S. (2015). Liouville quantum gravity and the Brownian map I: The QLE$(8/3,0)$ metric. Preprint. Available at arXiv:1507.00719.

[70] MILLER, J. and SHEFFIELD, S. (2015). Liouville quantum gravity spheres as matings of finite-diameter trees. Preprint. Available at arXiv:1506.03804.

[71] MILLER, J. and SHEFFIELD, S. (2015). Liouville quantum gravity and the Brownian map II: Geodesics and continuity of the embedding. In preparation.

[72] MILLER, J. and SHEFFIELD, S. (2015). Liouville quantum gravity and the Brownian map III: The conformal structure is determined. In preparation.
[73] MULLIN, R. C. (1967). On the enumeration of tree-rooted maps. *Canad. J. Math.* **19** 174–183. MR0205882
[74] POLYAKOV, A. M. (1981). Quantum geometry of bosonic strings. *Phys. Lett. B* **103** 207–210. MR0623209
[75] POLYAKOV, A. M. (1981). Quantum geometry of fermionic strings. *Phys. Lett. B* **103** 211–213. MR0623210
[76] POLYAKOV, A. M. (2008). From Quarks to Strings. Preprint. Available at arXiv:0812.0183.
[77] PROTTER, P. (1990). *Stochastic Integration and Differential Equations. Applications of Mathematics* (*New York*) **21**. Springer, Berlin. MR1037262
[78] REVUZ, D. and YOR, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **293**. Springer, Berlin. MR1725357
[79] RHODES, R. and VARGAS, V. (2011). KPZ formula for log-infinitely divisible multifractal random measures. *ESAIM Probab. Stat.* **15** 358–371. MR2870520
[80] RHODES, R. and VARGAS, V. (2014). Gaussian multiplicative chaos and applications: A review. *Probab. Surv.* **11** 315–392. MR3274356
[81] RHODES, R. and VARGAS, V. (2015). Liouville Brownian motion at criticality. *Potential Anal.* **43** 149–197. MR3374108
[82] ROHDE, S. and SCHRAMM, O. (2005). Basic properties of SLE. *Ann. of Math.* (2) **161** 883–924. MR2153402
[83] ROHDE, S. and ZHAN, D. (2013). Backward SLE and the symmetry of the welding. *Probab. Theory Related Fields* 1–49.
[84] SCHAEFFER, G. (1997). Bijective census and random generation of Eulerian planar maps with prescribed vertex degrees. *Electron. J. Combin.* **4** 14 pp. (electronic). MR1465581
[85] SCHRAMM, O. (2000). Scaling limits of loop-erased random walks and uniform spanning trees. *Israel J. Math.* **118** 221–288. MR1776084
[86] SCHRAMM, O. and SHEFFIELD, S. (2009). Contour lines of the two-dimensional discrete Gaussian free field. *Acta Math.* **202** 21–137. MR2486487
[87] SCHRAMM, O. and SHEFFIELD, S. (2013). A contour line of the continuum Gaussian free field. *Probab. Theory Related Fields* **157** 47–80. MR3101840
[88] SCHRAMM, O., SHEFFIELD, S. and WILSON, D. B. (2009). Conformal radii for conformal loop ensembles. *Comm. Math. Phys.* **288** 43–53. MR2491617
[89] SCHRAMM, O. and WILSON, D. B. (2005). SLE coordinate changes. *New York J. Math.* **11** 659–669 (electronic). MR2188260
[90] SHEFFIELD, S. (2005). Local sets of the Gaussian free field: Slides and audio. Available at www.fields.utoronto.ca/audio/05-06/percolation_SLE/sheffield1, www.fields.utoronto.ca/audio/05-06/percolation_SLE/sheffield2, www.fields.utoronto.ca/audio/05-06/percolation_SLE/sheffield3.
[91] SHEFFIELD, S. (2007). Gaussian free fields for mathematicians. *Probab. Theory Related Fields* **139** 521–541. MR2322706
[92] SHEFFIELD, S. (2009). Exploration trees and conformal loop ensembles. *Duke Math. J.* **147** 79–129. MR2494457
[93] SHEFFIELD, S. (2011). Quantum gravity and inventory accumulation. Preprint. Available at arXiv:1108.2241.
[94] SHEFFIELD, S. and WERNER, W. (2012). Conformal loop ensembles: The Markovian characterization and the loop-soup construction. *Ann. of Math.* (2) **176** 1827–1917. MR2979861
[95] SHEFFIELD, S. and YADIN, A. (2014). Tricolor percolation and random paths in 3D. *Electron. J. Probab.* **19** no. 4, 23. MR3164757

[96] SMIRNOV, S. Private communciation.
[97] SMIRNOV, S. (2001). Critical percolation in the plane: Conformal invariance, Cardy's formula, scaling limits. *C. R. Acad. Sci. Paris Sér. I Math.* **333** 239–244. MR1851632
[98] SMIRNOV, S. (2005). Critical percolation and conformal invariance. In *XIVth International Congress on Mathematical Physics* 99–112. World Sci. Publ., Hackensack, NJ. MR2227824
[99] SMIRNOV, S. (2006). Towards conformal invariance of 2D lattice models. In *International Congress of Mathematicians. Vol. II* 1421–1451. Eur. Math. Soc., Zürich. MR2275653
[100] SMIRNOV, S. (2010). Conformal invariance in random cluster models. I. Holomorphic fermions in the Ising model. *Ann. of Math.* **2** 1435–1467.
[101] SUN, X. and WILSON, D. B. (2015). Sandpiles and unicycles on random planar maps. Preprint. Available at arXiv:1506.08881.
[102] TECU, N. (2012). Random conformal weldings at criticality. Preprint. Available at arXiv:1205.3189.
[103] WATABIKI, Y. (1993). Analytic study of fractal structure of quantized surface in two-dimensional quantum gravity. *Progr. Theoret. Phys. Suppl.* **114** 1–17. MR1268911
[104] WERNER, W. (2004). Random planar curves and Schramm–Loewner evolutions. In *Lectures on Probability Theory and Statistics. Lecture Notes in Math.* **1840** 107–195. Springer, Berlin. MR2079672
[105] YADIN, A. and YEHUDAYOFF, A. (2011). Loop-erased random walk and Poisson kernel on planar graphs. *Ann. Probab.* **39** 1243–1285. MR2857239

DEPARTMENT OF MATHEMATICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS 02139
USA
E-MAIL: sheffield@math.mit.edu