arXiv: 0710.0379

CONSISTENT ESTIMATES OF DEFORMED ISOTROPIC GAUSSIAN RANDOM FIELDS ON THE PLANE

By Ethan Anderes* and Sourav Chatterjee†

University of California at Berkeley

This paper proves fixed domain asymptotic results for estimating a smooth invertible transformation $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ when observing the deformed random field $Z \circ f$ on a dense grid in a bounded simply connected domain Ω where Z is assumed to be an isotropic Gaussian random field on $\mathbb{R}^2$. The estimate, $\hat{f}$, is constructed on a simply connected domain U such that $\overline{U} \subset \Omega$ and is defined using kernel smoothed quadratic variations, Bergman projections and results from quasiconformal theory. We show under mild assumptions on the random field Z and the deformation f that $\hat{f} \to R_\theta f + c$ uniformly on compact subsets of U with probability one as the grid spacing goes to zero, where R_θ is an unidentifiable rotation and c is an unidentifiable translation.

1. Introduction. The use of deformations to model nonstationary processes was first introduced to the spatial statistics literature by Sampson and Guttorp [39]. Their work, as well as that of subsequent authors (see for example, [40], [14], [36], [25])) consider estimating the deformation f when observing a deformed random field $Z \circ f$ at sparse observation locations with independent replicates of the random field.

Two recent papers [11], [6] study a different problem: estimating a deformation f from dense observations of a single realization of a deformed isotropic random field $Z \circ f$ in two dimensions. These deformed isotropic random fields provide a flexible semi-parametric model of nonstationarity for random fields. In addition, this observation scenario is becoming increasingly important with the abundance of high resolution digital imagery and remote sensing. One of the more recent motivations for the deformation model under the one-realization observation scenario is gravitational lensing of the cosmic microwave background (CMB). The gravitational effect from intervening matter distort the CMB images to produce deformed random field observations. In the hope of improving estimates of cosmological parameters and the mass distribution in the universe (including dark matter) there is considerable interest in detecting and measuring the lensing of the CMB ([24], [42]).

*Supported in part by an NSF Postdoctoral Fellowship DMS-0503227.

†Supported in part by a Sloan Research Fellowship in Mathematics and NSF grant DMS-0707054.

AMS 2000 subject classifications: Primary 60G60, 62M30, 62M40; secondary 62G05

Keywords and phrases: Deformation, Quasiconformal Maps, Nonstationary Random Fields, Bergman Space

In this paper we establish a strong consistency result for the estimation of the deformation f when observing $Z \circ f$ on a dense grid in a bounded simply connected domain in $\mathbb{R}^2$, as the grid spacing goes to zero. We first construct estimates of the complex dilatation and log-scale of the map f (see Section 4) which converge uniformly on compact subsets of the observation region with probability one. Then we construct a deformation estimate $\hat{f}$ on a subset of the observation region which converges uniformly on compact subsets with probability one. We show this result under mild assumptions on the map f and the two dimensional isotropic random field Z.

Most attempts at recovering the deformation f from a single realization of $Z \circ f$ rely on estimating local properties of f, usually related to the Jacobian of f, from the local behavior of the random field $Z \circ f$. Intuitively, the random field $Z \circ f$ is locally stretched and sheared by f, as determined by the Jacobian. One can clearly see the visual consequences of this shear as seen in the left plot from Figure 1. When the random field Z is isotropic the identification of all four parameters of the Jacobian becomes difficult from the local behavior of $Z \circ f$. In particular, by decomposing the Jacobian matrix as $U\Lambda V^T$ (using singular value decomposition so that U, V are orthogonal matrices and Λ is diagonal) the rotation matrix U becomes particularly hard, if not impossible, to estimate when observing $Z \circ f$ in a small neighborhood. An important object for us is the complex dilatation and log-scale of f, determined by the Jacobian (and defined in Section 4), which are invariant under left multiplication of rotation matrices to the Jacobian. It is this invariance which allows us to estimate these parameters under the isotropy assumption for Z.

Guyon and Perrin [22] tackle the problem of developing consistent estimates of deformations in two dimensions and succeed in proving consistency within a subclass of deformations when observing random fields that are stationary but not isotropic. The subclass of deformations, however, is restrictive. In particular, it is not closed under post composition with rotations which, in some sense, removes some complications for estimating the Jacobian generated by general deformations that can locally twist as well as stretch. On the other hand, Anderes and Stein [6] consider a large, nonparametric class of deformations. However, their results are methodological in nature and do not provide proofs of consistency. This paper contributes to bridging the gap between these two papers by considering the same flexible class of nonparametric deformations as in [6] while proving consistency for the estimated deformation as in [22]. For further references on densely observed deformed random fields see, [10], [35], [31], [34], [20], [37].

In this paper we use kernel smoothed directional quadratic variations to estimate local properties of f which are used to construct an estimate $\hat{f}$ of f. We establish sufficient conditions on the rate of bandwidth decay, in relation to the grid spacing, for strong uniform convergence of the kernel smoothed quadratic variations. There is a significant amount of literature studying the convergence of quadratic variations (see for example [32], [7], [19], [16], [27], [9], [43],[1], [8], [21], [12], [26]). One of the crucial inequalities used in many of the recent convergence results is the Hanson and Wright bound for quadratic forms [23].

Indeed, we also depend heavily on this bound, and use it to establish Claim 2 in Section 3.2, which in turn gives uniform convergence on compact subsets for the estimated complex dilatation and log-scale and ultimately the convergence of the estimated deformation.

The kernel smoothed quadratic variations used in this paper are based on second order increments of the deformed process. Second order increments—rather than first order—are used in equation (4) to obtain sufficient spatial de-correlation for uniform convergence. Using higher order increments for quadratic variations is not new. They have been used in [26] and [8] for identification of a local fractional index and in [13] to identify the singularity function of a fractional process. The heuristic is that by increasing the order of increment, one can increase the rate of decay of the variance of the quadratic variation. However, this rate improvement holds only to a point after which additional increments no longer improves the situation. The interaction between the number of increments, the fractional index of the random field, and the dimension of the domain of the random field is investigated in Chapter 3 of [4].

One of the main theoretical tools we use in this paper is the theory of quasiconformal maps. We believe this paper demonstrates how quasiconformal theory can provide a flexible theoretical framework for estimating smooth invertible transformations whereby making these objects available to statisticians for modeling a diverse range of physical phenomena. In two dimensions, an important object in the theory of quasiconformal maps is the complex dilatation $\mu\colon \Omega \to \mathbb{D}$ (here Ω is the observation region for $Z \circ f$, and $\mathbb{D}$ is the unit disk in the complex plane). A more detailed discussion of the complex dilatation is presented in Section 4. Besides characterizing quasiconformal maps up to post composition with conformal maps, the complex dilatation, μ, has two other useful properties. First, μ can be interpreted as measuring the ellipticity and inclination of the local ellipse which gets mapped to a local circle under the quasiconformal map which it characterizes. This is important to us for developing estimates of μ locally from $Z \circ f$. Second, the only requirement on μ is measurability and $\|\mu\|_\infty < 1$. In other words, it suffices to measurably assign eccentricity and inclinations of local infinitesimal ellipses and by keeping the eccentricity bounded there is a quasiconformal map which send these infinitesimal ellipses to circles (unique up to post composition with conformal maps). This property allows us to find a smooth invertible transformation that corresponds to the estimated complex dilatation.

The other object we use for estimating f is the log-scale $\tau := \log |\partial f|$. We will discuss both μ and τ in detail later, however it is worth while to notice that μ and τ provide enough information to uniquely specify the map f up to an overall rotation and translation. As we will see, one of the difficulties with τ, as compared to μ, is that τ lies in a complicated subspace of functions mapping $\mathbb{R}^2$ to $\mathbb{R}$. This is where we employ the Bergman projection as a tool to overcome the restrictive nature of the log-scale parameter τ.

Figure 2 illustrates the estimates $\hat{\mu}$ and $\widehat{|\partial f|}$ from the simulation shown in Figure 1. These are obtained by convolving the second order quadratic increments with a Gaussian smoothing kernel and transforming these smoothed increments as discussed in Section 4.

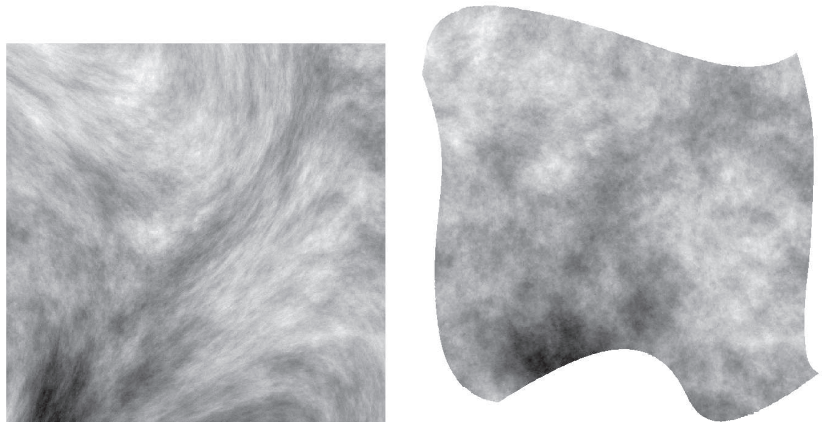

FIG 1. *Left: One realization of a deformed isotropic random field. Right: The recovered isotropic random field using the estimated deformation from Figure* 3.

An estimated deformation $\hat{f}$ corresponding to $\hat{\mu}$ and $|\widehat{\partial f}|$ is show in Figure 3 (*left*) along with the true deformation (*right*). The image shown in Figure 1 (*right*) shows $Z \circ f \circ \hat{f}^{-1}$, which "unwinds" the deformed process in an attempt at recovering the isotropic process Z. Note: the deformation $\hat{f}$ in Figure 3 is constructed using methods from both this paper and from [6]. To be explicit, all the estimation methods from this paper are used for $\hat{f}$ with the exception of the Bergman projection where the methods from [6] were used. The reason is that the computational techniques are not yet developed, to the authors' knowledge, for accurate approximation of the Bergman projection used in constructing $\hat{f}$. However, since the Bergman space is a reproducing kernel Hilbert space, there is potential for accurate approximation with using spline methodology.

There are three main parts to this paper. Section 2 discusses the assumptions on Z and the smooth invertible map f. In Section 3 we show that kernel smoothed directional quadratic variations converge uniformly on compact subsets with probability one. These directional quadratic variations are then used, in Section 4, to get estimates $\hat{\mu}$ and $\hat{\tau} := \log|\widehat{\partial f}|$. Finally, in Section 5 we show how to convert $\hat{\mu}$ and $\hat{\tau}$ to an estimated map $\hat{f}$ on simply connected subsets U such that $\overline{U} \subset \Omega$, and show that $\hat{f}$ converges to f uniformly on compact subsets of U with probability one.

2. The random field $\boldsymbol{Z}$ and the map $\boldsymbol{f}$. In this section we list our assumptions on the isotropic random field Z and the smooth invertible transformation f. This section starts with a brief discussion of our assumptions on the autocovariance function of Z. Then a detailed discussion follows on our assumptions for the smooth invertible transformation $f\colon \mathbb{R}^2 \to \mathbb{R}^2$.

We require the following three conditions on the isotropic random field Z:

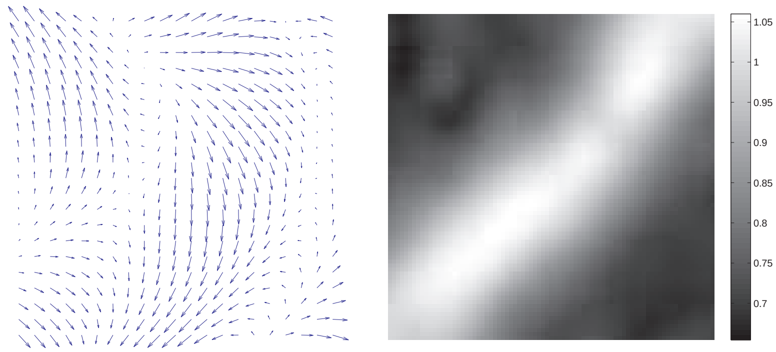

FIG 2. *The estimated dilatation $\hat{\mu}$ (left) and scale $\widehat{|\partial f|}$ (right) using kernel smoothed second order directional increments and the results of Section* 4.

R1. Z is a constant mean Gaussian process on $\mathbb{R}^2$ with autocovariance $R(|\boldsymbol{t} - \boldsymbol{s}|) = \mathrm{cov}(Z(\boldsymbol{t}), Z(\boldsymbol{s}))$.

R2. $R(|t|) = R(0) - |t|^\alpha + o(|t|^{\alpha+\gamma})$, as $|t| \to 0$ for some $0 < \alpha < 2$, $\gamma > 0$.

R3. R is C^4 away from the origin and there exists a $c > 0$ such that $|R^{(4)}(t)| \le c\, t^{\alpha-4}$ for all sufficiently small $t > 0$.

The assumption *R2* establishes the local quadratic variation behavior of the process Z to be similar to that of a fractional Brownian sheet with Hurst index $\alpha/2$. Informally, the assumption *R3* ensures that the second order increments of Z have spatial de-correlation like that of a fractional Brownian sheet. *Remark 1:* Most of the following results can be extended, with some additional technical assumptions, to a larger class of autocovariance functions by replacing the principle term $|t|^\alpha$ in *R2* with $L(|t|)|t|^\alpha$, where L is a slowly varying function at 0. The main difference is that the quadratic variation process defined in equation (4) below will need to be normalized by $n^{-\alpha}L(1/n)$ instead of $n^{-\alpha}$. *Remark 2:* The class of autocovariances satisfying *R2*-*R3* encompasses a broad range of random fields that are continuous but not differentiable. Examples include the Mátern autocovariance function with smoothness parameter $0 < \nu < 1$ (see [41], [33], [45]) and the exponential family $\exp(-c|t|^\alpha)$ where $\alpha \in (0, 2)$. One way to extend our results to random fields with higher order differentiability is to use quadratic variations of higher order increments of the deformed random field to obtain sufficient spatial de-correlation. In the interest of space we only prove results for the non-differentiable case.

Our basic assumption on the smooth invertible map f is that there exists a local affine approximation. In particular,

$$f(\boldsymbol{t} + \boldsymbol{h}) = f(\boldsymbol{t}) + J_f^{\boldsymbol{t}} \boldsymbol{h} + o(|\boldsymbol{h}|) \tag{1}$$

where $J_f^{\boldsymbol{t}} := \left(\frac{\partial f_i}{\partial t_j}(\boldsymbol{t})\right)_{i,j}$ is the Jacobian of the map f at $\boldsymbol{t}$. This local linear behavior is

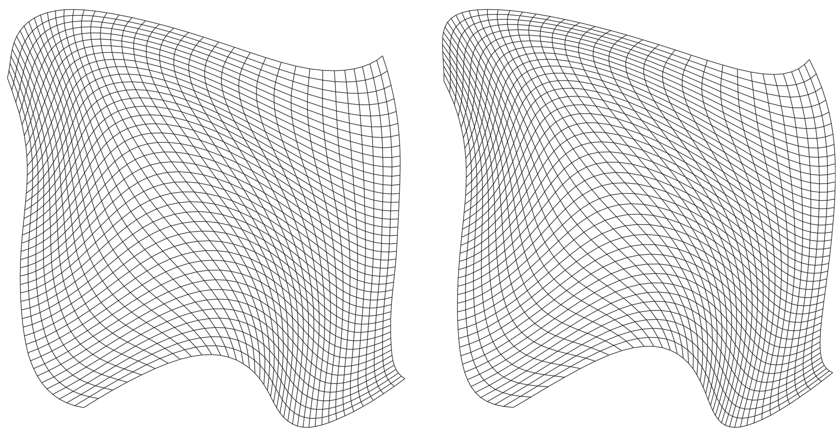

Fig 3. *Left: The estimated deformation recovered from the estimated dilatation and scale shown in Figure 2. The deformation was constructed from the estimated dilatation and scale using methods outlined in [6]. Right: The true deformation f.*

important since we don't have replicates of the deformed random field $Z \circ f$ and therefore most of the statistical information is contained in the local variation of the process $Z \circ f$. When f behaves locally like the Jacobian matrix transformation, the distribution of the random field $Z \circ f(\boldsymbol{t} + \boldsymbol{h})$, as $\boldsymbol{h}$ varies in a small neighborhood of the origin for a fixed $\boldsymbol{t}$, behaves similar to that of $Z(J_f^{\boldsymbol{t}} \boldsymbol{h})$. Therefore one can hope to estimate parameters of the Jacobian $J_f^{\boldsymbol{t}}$ using the local quadratic variation of one realization of the process $Z \circ f$ near $\boldsymbol{t}$. Of course, higher order terms in a Taylor expansion may also be estimated; however, presumably these require smaller neighborhoods for accurate estimation.

In addition to the local affine behavior of f we will need extra smoothness conditions. We discuss three notions of differentiability for a map $f \colon U \to V$ between planer open subsets U, V: Fréchet, Gâteaux, and $C^1(U)$. For a point $\boldsymbol{x}_0 \in U$, f is said to be Gâteaux differentiable at $\boldsymbol{x}_0$ in the direction $\boldsymbol{h}$ if the limit

$$\lim_{\epsilon \to 0} \frac{f(\boldsymbol{x}_0 + \epsilon \boldsymbol{h}) - f(\boldsymbol{x}_0)}{\epsilon}$$

exists. A stronger notion of differentiability is Fréchet differentiable. The map f is said to be Fréchet differentiable at $\boldsymbol{x}_0$ if there exits a continuous linear map, $T \colon \mathbb{R}^2 \to \mathbb{R}^2$, such that

$$f(\boldsymbol{x}_0 + \boldsymbol{h}) - f(\boldsymbol{x}_0) = T(\boldsymbol{h}) + o(|\boldsymbol{h}|).$$

If such a T exists and f has continuous partial derivatives then T is the map corresponding to left multiplication by the Jacobian matrix $J_f^{\boldsymbol{x}_0}$. Clearly, if a function f is Fréchet differentiable at $\boldsymbol{x}_0$, then it's Gâteaux differentiable at $\boldsymbol{x}_0$ and the Gâteaux derivative in the direction $\boldsymbol{h}$ is equal to $J_f^{\boldsymbol{x}_0} \boldsymbol{h}$.

The third notion of differentiability, $C^1(U)$, is satisfied if the partials $\frac{\partial f_i}{\partial x_j}$ exist and are continuous on U. This notion of differentiability is different in that it is not defined pointwise. The reason for this is that there is not much one can say about the local behavior near $\boldsymbol{x}_0$ of a function where the partials exist. In fact, it may be neither Fréchet nor Gâteaux differentiable. If, however, we require the partials $\frac{\partial f_i}{\partial x_j}$ to exits and be continuous on U, this is enough to imply Fréchet differentiability at all points in U. Even more is true: f is $C^1(U)$ if and only if $\boldsymbol{x} \mapsto J_f^{\boldsymbol{x}}$ is continuous as a mapping from U into the space of continuous linear functions on $\mathbb{R}^2$ (8.9.1 from [15]).

Define the class of $C^1(U)$ diffeomorphisms to be the set of all continuous invertible maps $f\colon U \to \mathbb{R}^2$ such that f is $C^1(U)$ and f^{-1} is $C^1(V)$ where $V = f(U)$ is the range of f. By the Inverse Function Theorem, necessary and sufficient conditions are that f be invertible, $C^1(U)$ and $\det J_f \neq 0$. We write C^1 as short for $C^1(\mathbb{R}^2)$. Now every C^1 diffeomorphism is Fréchet differentiable and so there is a Jacobian matrix $J_f^{\boldsymbol{t}}$ such that

$$f(\boldsymbol{t}+\boldsymbol{h}) = f(\boldsymbol{t}) + J_f^{\boldsymbol{t}}\boldsymbol{h} + o(|\boldsymbol{h}|).$$

Moreover, the directional derivative in the direction θ, denoted $\partial_\theta f$, is $J_f^{\boldsymbol{t}}\boldsymbol{u}_\theta$ where $\boldsymbol{u}_\theta = (\cos\theta, \sin\theta)$.

In the following paper we will restrict the definition of C^1 diffeomorphisms to have $\det J_f > 0$ on $\mathbb{R}^2$, which characterizes the diffeomorphism to be orientation preserving. In some sense this is a trivial restriction since when f is a C^1 diffeomorphism, either $\det J_f > 0$ everywhere or $\det J_f < 0$ everywhere. These are referred to as orientation-preserving and orientation-reversing respectively (see page 10 of [30]). Finally define $C^r(U)$ diffeomorphisms to be the $C^1(U)$ diffeomorphisms with order r continuous mixed partials.

We list some consequences of a C^2 diffeomorphic assumption on f which we use in the following proofs.

- *D1.* f is a quasiconformal map on bounded simply connected domains (see the appendix of [5] for a definition of quasiconformal maps).
- *D2.* $\sup_{\boldsymbol{t}\in\Theta}\left|\frac{f(\boldsymbol{t}+\epsilon\boldsymbol{h})-f(\boldsymbol{t})}{\epsilon} - J_f^{\boldsymbol{t}}\boldsymbol{h}\right| \to 0$ as $\epsilon \to 0$ for every compact set Θ.
- *D3.* For any vector $\boldsymbol{h} \neq 0$ and compact set Θ there exists a constant c such that $|\partial_{\boldsymbol{h}}^{(2,2)} R(|f(\boldsymbol{s}) - f(\boldsymbol{t})|)| \leq c|\boldsymbol{s}-\boldsymbol{t}|^{\alpha-4}$ for all $\boldsymbol{s}, \boldsymbol{t} \in \Theta$ such that $\boldsymbol{s} \neq \boldsymbol{t}$. Note: R and α are defined in assumptions *R1-R3* and $\partial_{\boldsymbol{h}}^{(2,2)}$ is defined in the next section.
- *D4.* For every compact subset Θ there exists constants $c_1, c_2 > 0$ such that $c_1|\boldsymbol{h}| \leq |J_f^{\boldsymbol{x}}\boldsymbol{h}| \leq c_2|\boldsymbol{h}|$ for all $\boldsymbol{h}$ and all $\boldsymbol{x} \in \Theta$.
- *D5.* $|J_f^{\boldsymbol{t}}\boldsymbol{h}|^\alpha$ is Hölder continuous in $\boldsymbol{t} \in \Theta$ for any $\boldsymbol{h}$ and compact set Θ.

Note: D3, D4 and *D5* are the only statements that depend on the extra C^2 assumption rather than the C^1. The proofs of *D1-D4* for C^2 diffeomorphisms are included in the Appendix of [5]. Notice that *D5* follows from *D4* and the fact that $J_f^{\boldsymbol{t}}$ has C^1 components.

3. Kernel smoothed squared increments. In this section we study the convergence of the kernel smoothed squared increments of the deformed process $Y(\boldsymbol{x}) := Z \circ f(\boldsymbol{x})$ observed on some dense grid in a bounded simply connected open subset $\Omega \subset \mathbb{R}^2$. The asymptotic regime we consider is as the grid spacing goes to zero and the region Ω stays fixed, sometimes called infill asymptotics.

For a fixed nonzero vector $\boldsymbol{h} \in \mathbb{R}^2$ let

$$\Delta_{\boldsymbol{h}} Y(\boldsymbol{t}) := Y(\boldsymbol{t}+\boldsymbol{h}) - Y(\boldsymbol{t})$$
$$\Delta^m_{\boldsymbol{h}} Y(\boldsymbol{t}) := \Delta_{\boldsymbol{h}}\Delta^{m-1}_{\boldsymbol{h}} Y(\boldsymbol{t}).$$

If $\boldsymbol{t}$ is near $\partial\Omega$, computing $\Delta^m_{\boldsymbol{h}}$ may require observing Y outside of Ω. Therefore we will suppose that we observe Y on Ω plus some points within a small distance from the boundary $\partial\Omega$. Now, for a function of two variables $F(\boldsymbol{s},\boldsymbol{t})$ let $\Delta^{(m,n)}_{\boldsymbol{h}} F(\boldsymbol{s},\boldsymbol{t}) := \Delta^m_{\boldsymbol{h}}\Delta^n_{\boldsymbol{h}} F(\boldsymbol{s},\boldsymbol{t})$ where $\Delta^m_{\boldsymbol{h}}$ acts on the variable $\boldsymbol{s}$ and $\Delta^n_{\boldsymbol{h}}$ acts on the variable $\boldsymbol{t}$. Define $\partial_{\boldsymbol{h}} := \boldsymbol{h}\cdot\nabla$ to be the directional derivative in the direction $\boldsymbol{h}$ and $\partial^{(m,n)}_{\boldsymbol{h}} F(\boldsymbol{s},\boldsymbol{t}) := \partial^m_{\boldsymbol{h}}\partial^n_{\boldsymbol{h}} F(\boldsymbol{s},\boldsymbol{t})$ where $\partial^m_{\boldsymbol{h}}$ acts on the variable $\boldsymbol{s}$ and $\partial^n_{\boldsymbol{h}}$ acts on $\boldsymbol{t}$. The following notation will be used throughout this paper

$$g(\boldsymbol{t}) := (8-2^{\alpha+1})|J_f^{\boldsymbol{t}}\boldsymbol{h}|^\alpha \tag{2}$$
$$\Omega_n := \Omega \cap \{\mathbb{Z}^2/n\} \tag{3}$$

i.e. Ω_n is the grid of spacing $1/n$ in Ω.

Here is a summary of the results of this section. First we show Lemma 1 which establishes that $\mathsf{E}(\Delta^2_{\boldsymbol{h}/n}Y(\boldsymbol{t}))^2 \approx n^{-\alpha}g(\boldsymbol{t})$. Motivated by this lemma we estimate $g(\boldsymbol{t})$ by locally averaging the squared increments $(\Delta^2_{\boldsymbol{h}/n}Y(\boldsymbol{w}))^2/n^{-\alpha}$ for the points $\boldsymbol{w}$ near $\boldsymbol{t}$ then show that there is enough spatial decorrelation for convergence. To this end, define $B_{n,b}(\boldsymbol{t}) : \Omega \to \mathbb{R}$ as follows

$$B_{n,b}(\boldsymbol{t}) := \frac{1}{n^2b^2}\sum_{\boldsymbol{w}\in\Omega_n} K\left(\frac{\boldsymbol{w}-\boldsymbol{t}}{b}\right)\frac{(\Delta^2_{\boldsymbol{h}/n}Y(\boldsymbol{w}))^2}{n^{-\alpha}}. \tag{4}$$

Here, K is a convolution kernel satisfying certain conditions stated below. This section then culminates with Theorems 4 and 5 concerning the uniform convergence of $B_{n,b}$ and $\partial_{\boldsymbol{u}}B_{n,b}$ for some $\boldsymbol{u} \neq 0$. In particular, Theorem 4 shows that under appropriate conditions,

$$\sup_{\boldsymbol{t}\in\Theta}\big|B_{n,b}(\boldsymbol{t}) - g(\boldsymbol{t})\big| \longrightarrow 0, \quad w.p.1$$

as $n\to\infty$, $b\to 0$, and $n^{-1}b^{-3}\to 0$ for all compact sets $\Theta\subset\Omega$. Theorem 5 shows that with some additional smoothness assumptions and $n^{-1}b^{-4}\to 0$ the directional derivatives $\partial_{\boldsymbol{u}}B_{n,b}$ converge uniformly on compact subsets w.p.1 to $\partial_{\boldsymbol{u}}g$.

3.1. *Assumptions on K and Ω.* The assumptions on K are as follows:

K1. K has bounded, continuous first and second order mixed partial derivatives.
K2. $\iint K(\boldsymbol{x})d\boldsymbol{x} = 1$ and $\iint |\boldsymbol{x}|K(\boldsymbol{x})d\boldsymbol{x} < \infty$.

Notice these assumptions imply that K is Riemann integrable, K and the first partials of K are Hölder continuous. Finally we assume:

O1. Ω is a bounded simply connected domain of $\mathbb{R}^2$.

Notice this assumption ensures that the number of points in Ω_n is of order n^2.

3.2. *Strong convergence of $B_{n,b}$ and $\partial_{\boldsymbol{u}} B_{n,b}$.* For the remainder of this section let $Y := Z \circ f$, $B_{n,b}$ be defined as in (4), g defined as in (2), Ω_n defined as in (3) and set $X_{\boldsymbol{t}} := B_{n,b}(\boldsymbol{t}) - \mathsf{E}B_{n,b}(\boldsymbol{t})$. In the following, 'universal constant' means any constant which does not depend on n, b, Θ, or the process $(X_{\boldsymbol{t}})_{\boldsymbol{t}\in\Theta}$.

LEMMA 1. *Suppose R1-R2, O1 and f is a C^2 diffeomorphism. Then*

$$\sup_{\boldsymbol{t}\in\overline{\Omega}} \left| \frac{\mathsf{E}\big(\Delta^2_{\epsilon\boldsymbol{h}} Y(\boldsymbol{t})\big)^2}{\epsilon^\alpha} - g(\boldsymbol{t}) \right| \to 0$$

as $\epsilon \downarrow \infty$.

PROOF. By assumption *R2* we can write $R(|t|) = R(0) - |t|^\alpha + r(|t|)$ where $r(|t|) = o(|t|^{\alpha+\gamma})$ as $|t| \to 0$. Write $\mathsf{E}\big(\Delta^2_{\epsilon\boldsymbol{h}} Y(\boldsymbol{t}_0)\big)^2$ for $\boldsymbol{t}_0 \in \overline{\Omega}$ as a sum of two terms

$$\mathsf{E}\big(\Delta^2_{\epsilon\boldsymbol{h}} Y(\boldsymbol{t}_0)\big)^2 = \Delta^{(2,2)}_{\epsilon\boldsymbol{h}} \Big\{ \mathrm{cov}(Y(\boldsymbol{s}), Y(\boldsymbol{t})) \Big\}\Big|_{\boldsymbol{s},\boldsymbol{t}=\boldsymbol{t}_0} = \mathcal{I}_1 + \mathcal{I}_2$$

where

$$\mathcal{I}_1 = \Delta^{(2,2)}_{\epsilon\boldsymbol{h}} \Big\{ -\big|f(\boldsymbol{s}) - f(\boldsymbol{t})\big|^\alpha \Big\}\Big|_{\boldsymbol{s},\boldsymbol{t}=\boldsymbol{t}_0} \tag{5}$$

$$\mathcal{I}_2 = \Delta^{(2,2)}_{\epsilon\boldsymbol{h}} \Big\{ r(|f(\boldsymbol{s}) - f(\boldsymbol{t})|) \Big\}\Big|_{\boldsymbol{s},\boldsymbol{t}=\boldsymbol{t}_0}. \tag{6}$$

Write the increment operator $\Delta^2_{\epsilon\boldsymbol{h}}$ as a linear filter so that its action on a function $Q\colon \mathbb{R}^2 \to \mathbb{R}$ can be expressed as $\Delta^2_{\epsilon\boldsymbol{h}} Q(\boldsymbol{t}) = \sum_{j=0}^2 d_j Q(\boldsymbol{t} + \epsilon \boldsymbol{s}_j)$ where $d_j = (-1)^{2-j}\binom{2}{j}$ and $\boldsymbol{s}_j = j\boldsymbol{h}$. Now the first term can be computed as follows

$$\mathcal{I}_1/\epsilon^\alpha = -\sum_{i,j=0}^{2} d_i d_j \left| f(\boldsymbol{t}_0 + \epsilon\boldsymbol{s}_i) - f(\boldsymbol{t}_0 + \epsilon\boldsymbol{s}_j) \right|^\alpha / \epsilon^\alpha \tag{7}$$

$$= -\sum_{i,j=0}^{2} d_i d_j \left| J_f^{\boldsymbol{t}_0}(\boldsymbol{s}_i - \boldsymbol{s}_j) \right|^\alpha + o(1) \tag{8}$$

where $o(1) \to 0$ uniformly over $\boldsymbol{t}_0 \in \overline{\Omega}$ as $\epsilon \to 0$ by *D2* and *D4*.

Similarly the second term can be computed as

$$\mathcal{J}_2/\epsilon^\alpha = \sum_{i,j=0}^{2} d_i d_j \frac{r(|f(\boldsymbol{t}_0 + \epsilon \boldsymbol{s}_i) - f(\boldsymbol{t}_0 + \epsilon \boldsymbol{s}_j)|)}{\epsilon^\alpha} \tag{9}$$

where $\sup_{\boldsymbol{t}_0 \in \overline{\Omega}} |r(|f(\boldsymbol{t}_0 + \epsilon \boldsymbol{s}_i) - f(\boldsymbol{t}_0 + \epsilon \boldsymbol{s}_j)|)/\epsilon^\alpha|$ converges to zero by *R2*, *D2* and *D4*. Combining terms $\mathcal{J}_1$ and $\mathcal{J}_2$ we get

$$E(\Delta_{\epsilon \boldsymbol{h}} Y(\boldsymbol{t}_0))^2/\epsilon^\alpha \to -\sum_{i,j=0}^{2} d_i d_j |J_f^{\boldsymbol{t}_0}(\boldsymbol{s}_i - \boldsymbol{s}_j)|^\alpha$$

uniformly for $\boldsymbol{t}_0 \in \overline{\Omega}$. This completes the proof since $-\sum_{i,j=0}^{2} d_i d_j |J_f^{\boldsymbol{t}_0}(\boldsymbol{s}_i - \boldsymbol{s}_j)|^\alpha = g(\boldsymbol{t}_0)$. □

LEMMA 2. *Suppose R1-R3, O1 and f is a C^2 diffeomorphism. Then there exists a constant $c > 0$ such that*

$$\left|\mathsf{E}\Delta^2_{\boldsymbol{h}/n} Y(\boldsymbol{t}) \Delta^2_{\boldsymbol{h}/n} Y(\boldsymbol{s})\right| \le c\, n^{-4} \left|\boldsymbol{t} - \boldsymbol{s}\right|^{\alpha-4}$$

for all $\boldsymbol{s}, \boldsymbol{t} \in \Omega$ such that $|\boldsymbol{s} - \boldsymbol{t}| > |3\boldsymbol{h}/n|$.

PROOF. The idea is that

$$\begin{aligned}
|\mathsf{E}\Delta^2_{\boldsymbol{h}/n} Y(\boldsymbol{t}) \Delta^2_{\boldsymbol{h}/n} Y(\boldsymbol{s})| &= |\Delta^{(2,2)}_{\boldsymbol{h}/n} R(|f(\boldsymbol{s}) - f(\boldsymbol{t})|)| \\
&\approx |n^{-4} \partial^{(2,2)}_{\boldsymbol{h}} R(|f(\boldsymbol{s}) - f(\boldsymbol{t})|)| \\
&\le c n^{-4} |\boldsymbol{s} - \boldsymbol{t}|^{\alpha-4}.
\end{aligned}$$

To make this precise let $F(\boldsymbol{s}, \boldsymbol{t}) := \operatorname{cov}(Y(\boldsymbol{s}), Y(\boldsymbol{t})) = R(|f(\boldsymbol{s}) - f(\boldsymbol{t})|)$ and H be the 2 by 2 matrix with each column $\boldsymbol{h}/n$. Then

$$\begin{aligned}
|\mathsf{E}\Delta^2_{\boldsymbol{h}/n} Y(\boldsymbol{t}) \Delta^2_{\boldsymbol{h}/n} Y(\boldsymbol{s})| &= |\Delta^{(2,2)}_{\boldsymbol{h}/n} F(\boldsymbol{s}, \boldsymbol{t})| \\
&= n^{-4} \left| \int_{[0,1]^2} \int_{[0,1]^2} \partial^{(2,2)}_{\boldsymbol{h}} F(\boldsymbol{s} + H\boldsymbol{\xi}, \boldsymbol{t} + H\boldsymbol{\eta}) d\boldsymbol{\xi} d\boldsymbol{\eta} \right| \\
&\le c_1 n^{-4} \int_{[0,1]^2} \int_{[0,1]^2} \left|\boldsymbol{s} - \boldsymbol{t} + H(\boldsymbol{\xi} - \boldsymbol{\eta})\right|^{\alpha-4} d\boldsymbol{\xi} d\boldsymbol{\eta} \\
&\le c_2 n^{-4} \sup_{\boldsymbol{\eta}, \boldsymbol{\xi} \in [0,1]^2} \left|\boldsymbol{s} - \boldsymbol{t} + \boldsymbol{h}(\xi_1 + \xi_2 - \eta_1 - \eta_2)/n\right|^{\alpha-4} \\
&= c_2 n^{-4} \sup_{-1 \le \tau \le 1} \left|\boldsymbol{s} - \boldsymbol{t} + 2\boldsymbol{h}\tau/n\right|^{\alpha-4} \\
&\le c_3 n^{-4} |\boldsymbol{s} - \boldsymbol{t}|^{\alpha-4} \quad \text{when } |\boldsymbol{s} - \boldsymbol{t}| > |3\boldsymbol{h}/n|.
\end{aligned}$$

Notice that the above proof requires that $|\partial_{\boldsymbol{h}}^{(2,2)} F(\boldsymbol{s},\boldsymbol{t})| \leq c_1 |\boldsymbol{s}-\boldsymbol{t}|^{\alpha-4}$ which is why we need *D3*. $\square$

CLAIM 1. *Suppose R1-R3, K1-K2, O1 and f is a C^2 diffeomorphism. Then*

$$\sup_{\boldsymbol{t}\in\Theta} \big|\mathsf{E}(B_{n,b}(\boldsymbol{t})) - g(\boldsymbol{t})\big| \to 0$$

for all compact subsets $\Theta \subset \Omega$ as $n\to\infty$, $b \to 0$ and $n^{-1}b^{-3} = o(1)$.

PROOF. First let $g_n(\boldsymbol{w}) := n^\alpha \mathsf{E}(\Delta^2_{\boldsymbol{h}/n} Y(\boldsymbol{w}))^2$ for $\boldsymbol{w}\in\Omega$. We show

$$\begin{aligned}
\mathsf{E}(B_{n,b}(\boldsymbol{t})) &= \frac{1}{n^2b^2}\sum_{\boldsymbol{w}\in\Omega_n} K\left(\frac{\boldsymbol{w}-\boldsymbol{t}}{b}\right) g_n(\boldsymbol{w})\\
&= \frac{1}{n^2b^2}\sum_{\boldsymbol{w}\in\Omega_n} K\left(\frac{\boldsymbol{w}-\boldsymbol{t}}{b}\right) g(\boldsymbol{w}) + e_I\\
&= \frac{1}{b^2}\iint_\Omega K\left(\frac{\boldsymbol{x}-\boldsymbol{t}}{b}\right) g(\boldsymbol{x})d\boldsymbol{x} + e_{II} + e_I\\
&= g(\boldsymbol{t}) + e_{III} + e_{II} + e_I.
\end{aligned}$$

where $e_I = o(1)$, $e_{II} = O(n^{-1}b^{-3})$ and $e_{III} = o(1)$ uniformly on compact subsets of Ω as $n\to\infty$, $b\to 0$, and $n^{-1}b^{-3}\to 0$.

To show the results about e_I and e_{II} we need to control the error when approximating Riemann integrals of Hölder continuous functions on Ω by Riemann sums on Ω_n. This error is bounded by the difference between the upper and lower Riemann sums which is bounded by $a_\Omega\, a\, n^{-1}$ where a is the Hölder constant of the function and a_Ω is a constant only depending on the region Ω. To show $e_I = o(1)$ and $e_{II} = O(n^{-1}b^{-3})$ uniformly for $\boldsymbol{t}\in\Theta$ we will use the fact that $b^{-2}K((\cdot - \boldsymbol{t})/b)$ is Hölder continuous with Hölder constant $c\, b^{-3}$ for some constant c.

To show $e_I = o(1)$ fix some compact subset $\Theta\subset\Omega$ and notice that

$$\begin{aligned}
|e_I| &\leq \frac{1}{n^2b^2}\sum_{\boldsymbol{w}\in\Omega_n} |K|\left(\frac{\boldsymbol{w}-\boldsymbol{t}}{b}\right)|g_n(\boldsymbol{w}) - g(\boldsymbol{w})|\\
&\leq \Big(\sup_{\boldsymbol{w}\in\overline{\Omega}} |g_n(\boldsymbol{w}) - g(\boldsymbol{w})|\Big)\left(\frac{1}{n^2b^2}\sum_{\boldsymbol{w}\in\Omega_n}|K|\left(\frac{\boldsymbol{w}-\boldsymbol{t}}{b}\right)\right).
\end{aligned}$$

The term $\sup_{\boldsymbol{w}\in\overline{\Omega}}|g_n(\boldsymbol{w}) - g(\boldsymbol{w})| \to 0$ by Lemma 1. Now by the comments in the previous paragraph the Riemann sum $\frac{1}{n^2b^2}\sum_{\boldsymbol{w}\in\Omega_n}|K|\left(\frac{\boldsymbol{w}-\boldsymbol{t}}{b}\right)$ is approximately $b^{-2}\iint |K|((\boldsymbol{x}-\boldsymbol{t})/b)d\boldsymbol{x}$ (which is bounded) with error $O(n^{-1}b^{-3})$. Therefore $e_I\to 0$ as $n\to\infty$, $b\to 0$ and $n^{-1}b^{-3}\to 0$.

Similarly, to show $e_{II} = O(n^{-1}b^{-3})$ we notice that the Hölder continuity of K and g are sufficient for the Riemann sums of $b^{-2}K((\cdot - \boldsymbol{t})/b)g(\cdot)$ to converge to the Riemann integral with an error $e_{II} = O(n^{-1}b^{-3})$ uniformly in $\boldsymbol{t} \in \Theta$.

Finally to show $e_{III} = o(1)$ we need that

$$\frac{1}{b^2}\iint_{\Omega} K\left(\frac{\boldsymbol{x}-\boldsymbol{t}}{b}\right) g(\boldsymbol{x})d\boldsymbol{x} = \iint_{(\Omega-\boldsymbol{t})/b} K(\boldsymbol{w})\, g(b\boldsymbol{w}+\boldsymbol{t})d\boldsymbol{w} \longrightarrow g(\boldsymbol{t})$$

as $b \to 0$ uniformly in $\boldsymbol{t} \in \Theta$. Here we use the Hölder continuity of g and assumption *K2*. Notice that the error term e_{III} does not converge to zero uniformly in $\boldsymbol{t} \in \overline{\Omega}$. This is why we can only show the result uniformly on compacts instead of uniformly on $\overline{\Omega}$. $\square$

In what follows we will not only show convergence results about $B_{n,b}(\boldsymbol{t})$ but also $\partial_{\boldsymbol{u}} B_{n,b}(\boldsymbol{t})$, and $B_{n,b}(\boldsymbol{t}) - B_{n,b}(\boldsymbol{s})$ all of which have the form

$$Q_n = n^{\alpha-2} \sum_{\boldsymbol{i}\in\Omega_n} (\Delta^2_{\boldsymbol{h}/n} Y(\boldsymbol{i}))^2 \mathcal{K}_{\boldsymbol{i}} \tag{10}$$

where $\mathcal{K}_{\boldsymbol{i}}$ may depend on $\boldsymbol{t}, \boldsymbol{s} \in \Omega$, and the bandwidth parameter b.

CLAIM 2. *Let Q_n be defined as in* (10). *Suppose R1-R3, O1 and f is a C^2 diffeomorphism. In addition suppose there exists a function $G(\boldsymbol{t}, \boldsymbol{s}, b)$ such that $|\mathcal{K}_{\boldsymbol{i}}| \le G(\boldsymbol{t}, \boldsymbol{s}, b)$ for all $\boldsymbol{s}, \boldsymbol{t} \in \Omega$ and b in a neighborhood of the origin. Then for all $\epsilon > 0$, b sufficiently small and n sufficiently large*

$$\mathsf{P}\Big[|Q_n - \mathsf{E}Q_n| \ge \epsilon\Big] \le c_1 \exp\left(-\frac{c_2 \epsilon n^2}{G(\boldsymbol{t},\boldsymbol{s},b)} \wedge \frac{c_3\epsilon^2 n^2}{G(\boldsymbol{t},\boldsymbol{s},b)^2}\right)$$

where c_1, c_2, c_3 are universal constants.

PROOF. Writing $\mathcal{K}_{\boldsymbol{i}} = \mathcal{K}^+_{\boldsymbol{i}} - \mathcal{K}^-_{\boldsymbol{i}}$, we get the decomposition $Q_n = Q^1_n - Q^2_n$ where

$$Q^1_n := n^{\alpha-2} \sum_{\boldsymbol{i}\in\Omega_n} (\Delta^2_{\boldsymbol{h}/n} Y(\boldsymbol{i}))^2 \mathcal{K}^+_{\boldsymbol{i}}$$
$$Q^2_n := n^{\alpha-2} \sum_{\boldsymbol{i}\in\Omega_n} (\Delta^2_{\boldsymbol{h}/n} Y(\boldsymbol{i}))^2 \mathcal{K}^-_{\boldsymbol{i}}$$

Therefore

$$\mathsf{P}(|Q_n - \mathsf{E}Q_n| \ge \epsilon) \le \mathsf{P}\left(|Q^1_n - \mathsf{E}Q^1_n| \ge \frac{\epsilon}{2}\right) + \mathsf{P}\left(|Q^2_n - \mathsf{E}Q^2_n| \ge \frac{\epsilon}{2}\right).$$

First we find a bound for $\mathsf{P}\left(|Q^1_n - \mathsf{E}Q^1_n| \ge \frac{\epsilon}{2}\right)$. Let ΔY be the column vector with elements $n^{\alpha/2-1}\Delta^2_{\boldsymbol{h}/n} Y(\boldsymbol{i})\mathcal{K}^{+/2}_{\boldsymbol{i}}$ for $\boldsymbol{i} \in \Omega_n$, where $\mathcal{K}^{+/2}_{\boldsymbol{i}} := (\mathcal{K}^+_{\boldsymbol{i}})^{1/2}$. Let $\Sigma = \mathsf{E}\Delta Y \Delta Y^T$ be the covariance matrix for ΔY so that $Q_n^1 = \Delta Y^T \Delta Y \overset{\mathcal{D}}{=} W^T \Sigma W$ where W is a vector of *iid* standard Gaussian random variables. Also let $\Sigma(\boldsymbol{i},\boldsymbol{j})$ denote the matrix entries of Σ for $\boldsymbol{i},\boldsymbol{j}\in\Omega_n$.

Using the bound on quadratic forms for Gaussian random variables found in Hanson and Wright [23] we now get

$$\mathsf{P}\Big[|W^T\Sigma W - \mathsf{E} W^T \Sigma W| \geq \epsilon\Big] \leq 2\exp\left(-\frac{c_1\epsilon}{\|\Sigma_{\text{abs}}\|_2} \wedge \frac{c_2\epsilon^2}{\|\Sigma_{\text{abs}}\|_F^2}\right)$$

where $\|\cdot\|_2$ and $\|\cdot\|_F$ are the spectral and Frobenius matrix norms respectively and Σ_{abs} is the matrix with elements $|\Sigma(\boldsymbol{i},\boldsymbol{j})|$ for $\boldsymbol{i},\boldsymbol{j}\in\Omega_n$. Now

$$\begin{aligned}|\Sigma(\boldsymbol{i},\boldsymbol{j})| &= \frac{\mathcal{K}_{\boldsymbol{i}}^{+/2}\mathcal{K}_{\boldsymbol{j}}^{+/2}}{n^{2-\alpha}}\big|\mathsf{E}\Delta^2_{\boldsymbol{h}/n}Y(\boldsymbol{i})\Delta^2_{\boldsymbol{h}/n}Y(\boldsymbol{j})\big| \\ &\leq c_3\frac{\mathcal{K}_{\boldsymbol{i}}^{+/2}\mathcal{K}_{\boldsymbol{j}}^{+/2}}{n^{2-\alpha}} n^{-4}|\boldsymbol{i}-\boldsymbol{j}|^{\alpha-4}\end{aligned}$$

for all $|\boldsymbol{i}-\boldsymbol{j}| > |3\boldsymbol{h}/n|$ by Lemma 2. By assumption $|\mathcal{K}_{\boldsymbol{i}}^{+/2}\mathcal{K}_{\boldsymbol{j}}^{+/2}| \leq G(\boldsymbol{t},\boldsymbol{s},b)$. Therefore for all $\boldsymbol{i},\boldsymbol{j}\in\Omega_n$ such that $|\boldsymbol{i}-\boldsymbol{j}| > |3\boldsymbol{h}/n|$

$$|\Sigma(\boldsymbol{i},\boldsymbol{j})| \leq c_3\, G(\boldsymbol{t},\boldsymbol{s},b)\, n^{\alpha-6}|\boldsymbol{i}-\boldsymbol{j}|^{\alpha-4}. \tag{11}$$

To finish the proof we show $\|\Sigma_{abs}\|_F^2 = O(|G(\boldsymbol{t},\boldsymbol{s},b)|^2 n^{-2})$ and $\|\Sigma_{abs}\|_2 = O(|G(\boldsymbol{t},\boldsymbol{s},b)|n^{-2})$ uniformly for all $\boldsymbol{t},\boldsymbol{s}\in\Omega$, b sufficiently small, and n sufficiently large.

To show the bound for $\|\Sigma_{abs}\|_F^2$ notice

$$\begin{aligned}\sum_{\boldsymbol{i},\boldsymbol{j}\in\Omega_n}\Sigma(\boldsymbol{i},\boldsymbol{j})^2 &\leq \sum_{|\boldsymbol{i}-\boldsymbol{j}|\leq|3\boldsymbol{h}/n|}\Sigma(\boldsymbol{i},\boldsymbol{j})^2 + c_3^2\, G^2 n^{2\alpha-8}\sum_{|\boldsymbol{i}-\boldsymbol{j}|>|3\boldsymbol{h}/n|} n^{-4}\,|\boldsymbol{i}-\boldsymbol{j}|^{2\alpha-8} \\ &=^* \sum_{|\boldsymbol{i}-\boldsymbol{j}|\leq|3\boldsymbol{h}/n|}\Sigma(\boldsymbol{i},\boldsymbol{j})^2 + c_3^2\, G^2 n^{2\alpha-8} O(n^{-2\alpha+6}) \\ &= O(n^{-2}G^2) + O(n^{-2}G^2).\end{aligned}$$

where the last equality is because $|\Sigma(\boldsymbol{i},\boldsymbol{j})| \leq \max_{\boldsymbol{i}}\Sigma(\boldsymbol{i},\boldsymbol{i}) = O(n^{-2}G(\boldsymbol{t},\boldsymbol{s},b))$, by Lemma 1. To get $=^*$ there are some technical difficulties but the heuristic is when $0<\alpha<2$

$$\begin{aligned}\sum_{|\boldsymbol{i}-\boldsymbol{j}|>|3\boldsymbol{h}/n|} n^{-4}\,|\boldsymbol{i}-\boldsymbol{j}|^{2\alpha-8} &\asymp c_4\int_{\{1/n<|\boldsymbol{x}|<1\}}|\boldsymbol{x}|^{2\alpha-8}d\boldsymbol{x} \\ &= c_5\int_{1/n}^1 r^{2\alpha-7}dr \\ &= O(n^{-2\alpha+6}).\end{aligned}$$

For the full details see [4], Lemma 3, page 41. Finally, to show the bound for $\|\Sigma_{abs}\|_2$ notice that

$$\begin{aligned}
\|\Sigma_{abs}\|_2 &\leq \max_{\boldsymbol{i}\in\Omega_n}\sum_{\boldsymbol{j}\in\Omega_n}|\Sigma(\boldsymbol{i},\boldsymbol{j})| \\
&\leq \max_{\boldsymbol{i}\in\Omega_n}\sum_{|\boldsymbol{i}-\boldsymbol{j}|\leq|3\boldsymbol{h}/n|}|\Sigma(\boldsymbol{i},\boldsymbol{j})| + c_3\,G\max_{\boldsymbol{i}\in\Omega_n}\sum_{|\boldsymbol{i}-\boldsymbol{j}|>|3\boldsymbol{h}/n|}\frac{|\boldsymbol{i}-\boldsymbol{j}|^{\alpha-4}}{n^{6-\alpha}} \quad \text{by (11)} \\
&= O(n^{-2}G) + c_3\,Gn^{\alpha-4}\max_{\boldsymbol{i}\in\Omega_n}\sum_{|\boldsymbol{i}-\boldsymbol{j}|>|3\boldsymbol{h}/n|} n^{-2}\,|\boldsymbol{i}-\boldsymbol{j}|^{\alpha-4} \\
&= O(n^{-2}G) + O(n^{-2}G)
\end{aligned}$$

where the last equality uses the fact that $0<\alpha<2$.

This establishes the desired bound for $\mathsf{P}\left(\left|Q_n^1-\mathsf{E}Q_n^1\right|\geq\frac{\epsilon}{2}\right)$. The result for $\mathsf{P}\left(\left|Q_n^2-\mathsf{E}Q_n^2\right|\geq\frac{\epsilon}{2}\right)$ is exactly similar. This completes the proof. $\square$

Corollary 1. *Fix a point $\boldsymbol{t}_0\in\Omega$, suppose R1-R3, O1 and let f be a C^2 diffeomorphism. If K is bounded then*

$$\left|B_{n,b}(\boldsymbol{t}_0)-\mathsf{E}B_{n,b}(\boldsymbol{t}_0)\right|\xrightarrow{a.s.}0$$

as $n\to\infty$, $b\to0$ and $n^{-1}b^{-2}=O(n^{-\beta})$ for some $\beta>0$.

Proof. This follows by Claim 2 and Borel-Cantelli using $\mathcal{K}_{\boldsymbol{i}}=\frac{1}{b^2}K\left(\frac{\boldsymbol{i}-\boldsymbol{t}_0}{b}\right)$ and $G(\boldsymbol{t},\boldsymbol{s},b)=b^{-2}\|K\|_\infty$. $\square$

Corollary 2. *Fix a point $\boldsymbol{t}_0\in\Omega$, suppose R1-R3, O1 and let f be a C^2 diffeomorphism. If K has continuous partial derivatives and $\boldsymbol{u}\neq0$ then*

$$\left|\partial_{\boldsymbol{u}}B_{n,b}(\boldsymbol{t}_0)-\mathsf{E}\partial_{\boldsymbol{u}}B_{n,b}(\boldsymbol{t}_0)\right|\xrightarrow{a.s.}0$$

as $n\to\infty$, $b\to0$ and $n^{-1}b^{-3}=O(n^{-\beta})$ for some $\beta>0$.

Proof. This follows by Claim 2 and Borel-Cantelli using $\mathcal{K}_{\boldsymbol{i}}=\frac{1}{b^3}(\partial_{\boldsymbol{u}}K)\left(\frac{\boldsymbol{i}-\boldsymbol{t}_0}{b}\right)$ and $G(\boldsymbol{t},\boldsymbol{s},b)=b^{-3}\|\partial_{\boldsymbol{u}}K\|_\infty$. $\square$

The following corollary will be used for the uniform convergence of B to g in the next subsection. Remember that $X_{\boldsymbol{t}}$ is defined as $B_{n,b}(\boldsymbol{t})-\mathsf{E}B_{n,b}(\boldsymbol{t})$.

COROLLARY 3. *Fix a point $\boldsymbol{t}_0 \in \Omega$, suppose R1-R3, O1 and let f be a C^2 diffeomorphism. If K is Hölder continuous then*

$$\mathsf{P}(|X_{\boldsymbol{t}} - X_{\boldsymbol{s}}| \geq \epsilon) \leq c_1 \exp\left(-\frac{c_2 \epsilon n^2 b^3}{|\boldsymbol{t}-\boldsymbol{s}|} \wedge \frac{c_3 \epsilon^2 n^2 b^6}{|\boldsymbol{t}-\boldsymbol{s}|^2}\right) \tag{12}$$

were c_1, c_2, c_3 are universal constants.

PROOF. First write $X_{\boldsymbol{t}} - X_{\boldsymbol{s}}$ in the form $Q_n - \mathsf{E}Q_n$ where $Q_n := B_{n,b}(\boldsymbol{t}) - B_{n,b}(\boldsymbol{s})$. Then the corollary follows by Claim 2 using

$$\mathcal{K}_{\boldsymbol{i}} = \frac{1}{b^2} K\left(\frac{\boldsymbol{t}-\boldsymbol{i}}{b}\right) - \frac{1}{b^2} K\left(\frac{\boldsymbol{s}-\boldsymbol{i}}{b}\right)$$

so that $|\mathcal{K}_{\boldsymbol{i}}| \leq c\, b^{-3}|\boldsymbol{s}-\boldsymbol{t}|$ for some Hölder constant $c > 0$. □

3.3. *Uniform Convergence of $B_{n,b}$ and $\partial_{\boldsymbol{u}} B_{n,b}$.* In this subsection we use the results from the previous section to establish the uniform convergence of $B_{n,b}$ and the directional derivative $\partial_{\boldsymbol{u}} B_{n,b}$ on compact subsets of the observation region Ω. These results are then used to establish consistent estimators of the complex dilatation μ and log-scale τ of the diffeomorphism f in Section 4.

LEMMA 3. *For any $b > 0$ and $a \geq e$, we have*

$$\int_0^\infty (ae^{-bt^2} \wedge 1)dt \leq 2\sqrt{\frac{\log a}{b}} \quad \textit{and} \quad \int_0^\infty (ae^{-bt} \wedge 1)dt \leq 2\frac{\log a}{b}.$$

PROOF. Given $b > 0$ and $a \geq e$, let $\gamma = \sqrt{b^{-1}\log a}$. Then

$$\begin{aligned}\int_0^\infty (ae^{-bt^2} \wedge 1)dt &= \gamma + \int_\gamma^\infty ae^{-bt^2} dt \\ &\leq \gamma + \int_\gamma^\infty \frac{t}{\gamma} ae^{-bt^2} dt \\ &= \gamma + \frac{1}{2b\gamma} \\ &\leq 2\gamma \ \ (\text{since } a \geq e \implies \gamma \geq 1/2b\gamma).\end{aligned}$$

Similarly, putting $\nu = b^{-1}\log a$,

$$\int_0^\infty (ae^{-bt} \wedge 1)dt = \nu + \int_\nu^\infty ae^{-bt} dt = \nu + \frac{1}{b} \leq 2\nu.$$

This completes the proof of the lemma. □

LEMMA 4. *Let Θ be a compact subset of Ω. Suppose the assumptions in Corollary 3. Let $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_m$ and $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_m$ be arbitrary points in Θ, where $m \geq 2$. Let $M = \max_{1\leq i\leq m} |X_{\boldsymbol{t}_i} - X_{\boldsymbol{s}_i}|$ and $\delta = \max_{1\leq i\leq m} |\boldsymbol{t}_i - \boldsymbol{s}_i|$. Then for any $r > 0$,*

$$\mathsf{P}(M \geq r) \leq c_1 m \exp\biggl(-\frac{c_2 r n^2 b^3}{\delta}\biggr) + c_1 m \exp\biggl(-\frac{c_3 r^2 n^2 b^6}{\delta^2}\biggr),$$

where c_1, c_2, and c_3 are universal constants.

PROOF. By Corollary 3, for each $r > 0$ and $\boldsymbol{t}, \boldsymbol{s} \in \Theta$,

$$\begin{aligned}\mathsf{P}(|X_{\boldsymbol{t}} - X_{\boldsymbol{s}}| \geq r) &\leq c_1 \exp\biggl(-\frac{c_2 r n^2 b^3}{|\boldsymbol{t}-\boldsymbol{s}|} \wedge \frac{c_3 r^2 n^2 b^6}{|\boldsymbol{t}-\boldsymbol{s}|^2}\biggr)\\ &\leq c_1 \exp\biggl(-\frac{c_2 r n^2 b^3}{|\boldsymbol{t}-\boldsymbol{s}|}\biggr) + c_1 \exp\biggl(-\frac{c_3 r^2 n^2 b^6}{|\boldsymbol{t}-\boldsymbol{s}|^2}\biggr).\end{aligned}$$

From the above bound, we see that

$$\mathsf{P}(M \geq r) \leq c_1 m \exp\biggl(-\frac{c_2 r n^2 b^3}{\delta}\biggr) + c_1 m \exp\biggl(-\frac{c_3 r^2 n^2 b^6}{\delta^2}\biggr).$$

This completes the proof. □

CLAIM 3. *Fix a compact set $\Theta \subset \Omega$ and a point $\boldsymbol{t}_0 \in \Theta$. Let $M = \max_{\boldsymbol{t}\in\Theta} |X_{\boldsymbol{t}} - X_{\boldsymbol{t}_0}|$ and suppose the assumptions in Corollary 3. Then there exists universal constants L_1 and L_2 such that for all $R > 0$, we have*

$$\mathsf{P}\biggl(M \geq \frac{R}{nb^3}\biggr) \leq L_1 \exp(-L_2 \min\{Rn, R^2\}).$$

PROOF. Suppose we have a sequence of finite sets $A_0, A_1, A_2, \ldots \subseteq \Theta$ and constants $c < 1 < B$ and D satisfying the following properties.

(i) $A_0 = \{\boldsymbol{t}_0\}$.
(ii) $A_k \subseteq A_{k+1}$ for all k.
(iii) $|A_k| \leq B^k$ for all k, where $|A_k|$ denotes the cardinality of A_k.
(iv) For each $k \geq 1$ and each $\boldsymbol{t} \in A_k$ there exists a 'parent' $\boldsymbol{t}_p \in A_{k-1}$ such that $|\boldsymbol{t} - \boldsymbol{t}_p| \leq Dc^k$. Note that $\boldsymbol{t} = \boldsymbol{t}_p$ if $x \in A_{k-1}$.
(v) The sequence has a 'limiting denseness property' in the sense that for any nonnegative continuous function f on Θ, we have $\max_{\boldsymbol{t}\in\Theta} f(\boldsymbol{t}) = \lim_{k\to\infty} \max_{\boldsymbol{t}\in A_k} f(\boldsymbol{t})$.

It is easy to see how the constants c, B, and D can be chosen, with $D \propto \mathrm{diam}(\Theta)$, and the sets $\{A_k\}_{k\geq 1}$ constructed by successive dyadic partitioning. For each $k \geq 1$, let

$$M_k = \max_{\boldsymbol{t}\in A_k} |X_{\boldsymbol{t}} - X_{\boldsymbol{t}_p}|.$$

Applying the 'limiting denseness property' we see that $M \le \sum_{k=1}^{\infty} M_k$. For each $k \ge 1$, let $r_k = \frac{kc^k}{\sum_{j=1}^{\infty} jc^j}$. Then $r_k > 0$ and $\sum_k r_k = 1$. Thus for any $R > 0$,

$$\begin{aligned} \mathsf{P}\Big(M \ge \frac{R}{nb^3}\Big) &\le \mathsf{P}\Big(M_k \ge \frac{Rr_k}{nb^3} \text{ for some } k \ge 1\Big) \\ &\le \sum_{k=1}^{\infty} \mathsf{P}\Big(M_k \ge \frac{Rr_k}{nb^3}\Big). \end{aligned} \tag{13}$$

By the tail bound in Lemma 4, we have

$$\begin{aligned} \mathsf{P}\Big(M_k \ge \frac{Rr_k}{nb^3}\Big) &\le c_1 B^k \exp\Big(-\frac{c_2 R r_k n}{Dc^k}\Big) + c_1 B^k \exp\Big(-\frac{c_3 R^2 r_k^2}{D^2 c^{2k}}\Big) \\ &\le c_1 B^k (\exp(-c_5 Rkn) + \exp(-c_6 R^2 k^2)), \end{aligned}$$

where c_5 and c_6 are universal constants. Thus, if $R > c_7 \log B$ for some suitably large constant c_7 (that need not depend on n), then from the above bound and (13) we can conclude that

$$\mathsf{P}\Big(M \ge \frac{R}{nb^3}\Big) \le c_8 \exp(-c_9 Rn) + c_8 \exp(-c_{10} R^2).$$

The condition $R > c_7 \log B$ can be dropped by choosing c_8 large enough, because B, c_9 and c_{10} do not vary with n. This completes the proof. □

Now we obtain the main results of this section: Theorems 4 and 5. These establish uniform convergence results for the kernel smoothed squared increments.

THEOREM 4. *Suppose R1-R3, K1-K2, O1 hold and f is a C^2 diffeomorphism. Then for all compact sets $\Theta \subset \Omega$*

$$\sup_{\boldsymbol{t}\in\Theta} \big| B_{n,b}(\boldsymbol{t}) - g(\boldsymbol{t}) \big| \longrightarrow 0, \quad w.p.1$$

as $n \to \infty$, $b \to 0$, and $n^{-1}b^{-3} = O(n^{-\beta})$ for some $\beta > 0$, where $g(\boldsymbol{t}) := (8 - 2^{\alpha+1})|J_f^{\boldsymbol{t}} \boldsymbol{h}|^{\alpha}$.

PROOF. Let Θ be a compact subset of Ω and let $\boldsymbol{t}_0 \in \Theta$. Let $b_n \to 0$ as $n \to \infty$ so that $n^{-1} b_n^{-3} \le c\, n^{-\beta}$. Set $M_n = \sup_{\boldsymbol{t}\in\Theta} |X_{\boldsymbol{t}} - X_{\boldsymbol{t}_0}|$ where $X_{\boldsymbol{t}} := B_{n,b_n}(\boldsymbol{t}) - \mathsf{E} B_{n,b_n}(\boldsymbol{t})$. Then

$$\begin{aligned} \sup_{\boldsymbol{t}\in\Theta} \big| B_{n,b_n}(\boldsymbol{t}) - g(\boldsymbol{t}) \big| &\le \sup_{\boldsymbol{t}\in\Theta} \big| B_{n,b_n}(\boldsymbol{t}) - \mathsf{E} B_{n,b_n}(\boldsymbol{t}) \big| + \sup_{\boldsymbol{t}\in\Theta} \big| \mathsf{E} B_{n,b_n}(\boldsymbol{t}) - g(\boldsymbol{t}) \big| \\ &\le M_n + \big| B_{n,b_n}(\boldsymbol{t}_0) - \mathsf{E} B_{n,b_n}(\boldsymbol{t}_0) \big| + o(1), \quad \text{by Claim 1} \end{aligned}$$

Now by the bound in Claim 3 we have

$$\sum_{n=1}^{\infty} \mathsf{P}\big(M_n \geq n^{-\beta/2}\big) \leq \sum_{n=1}^{\infty} L_2 \exp(-L_3 \min\{n^{1+\beta/2}, n^{\beta}\}) < \infty.$$

Therefore $M_n \to 0$ with probability one. Finally, $n^{-1}b_n^{-3} = O(n^{-\beta})$ for some $\beta > 0$ is a sufficient condition for $\big|B_{n,b_n}(\boldsymbol{t}_0) - \mathsf{E}B_{n,b_n}(\boldsymbol{t}_0)\big| \xrightarrow{a.e.} 0$ by Corollary 1. □

THEOREM 5. *Suppose the assumptions of Theorem 4 hold along with the addition assumption that K is a compactly supported kernel and f is a C^3 diffeomorphism. Then for all compact sets $\Theta \subset \Omega$*

$$\sup_{\boldsymbol{t}\in\Theta}\big|\partial_{\boldsymbol{u}} B_{n,b}(\boldsymbol{t}) - \partial_{\boldsymbol{u}} g(\boldsymbol{t})\big| \longrightarrow 0, \quad w.p.1$$

as $n \to \infty$, $b \to 0$, $n^{-1}\,b^{-4} = O(n^{-\beta})$ for some $\beta > 0$ such that $\beta > 1 - 4\gamma$. Note: $\gamma > 0$ is the number appearing in assumption R2 and $g(\boldsymbol{t}) := (8 - 2^{\alpha+1})|J_f^{\boldsymbol{t}}\boldsymbol{h}|^{\alpha}$.

PROOF. Let Θ be a compact subset of Ω and let $\boldsymbol{t}_0 \in \Theta$. Let $b_n \to 0$ as $n \to \infty$ so that $n^{-1}b_n^{-4} \leq c\, n^{-\beta}$. Set $M_n = \sup_{\boldsymbol{t}\in\Theta}\big|\tilde{X}_{\boldsymbol{t}} - \tilde{X}_{\boldsymbol{t}_0}\big|$ where $\tilde{X}_{\boldsymbol{t}} := \partial_{\boldsymbol{u}} B_{n,b_n}(\boldsymbol{t}) - \mathsf{E}\partial_{\boldsymbol{u}} B_{n,b_n}(\boldsymbol{t})$. Then

$$\begin{aligned}\sup_{\boldsymbol{t}\in\Theta}\big|\partial_{\boldsymbol{u}} B_{n,b_n}(\boldsymbol{t}) - \partial_{\boldsymbol{u}} g(\boldsymbol{t})\big| &\leq \sup_{\boldsymbol{t}\in\Theta}\big|\partial_{\boldsymbol{u}} B_{n,b_n}(\boldsymbol{t}) - \mathsf{E}\partial_{\boldsymbol{u}} B_{n,b_n}(\boldsymbol{t})\big| \\ &\qquad + \sup_{\boldsymbol{t}\in\Theta}\big|\mathsf{E}\partial_{\boldsymbol{u}} B_{n,b_n}(\boldsymbol{t}) - \partial_{\boldsymbol{u}} g(\boldsymbol{t})\big| \\ &\leq M_n + \big|\partial_{\boldsymbol{u}} B_{n,b_n}(\boldsymbol{t}_0) - \mathsf{E}\partial_{\boldsymbol{u}} B_{n,b_n}(\boldsymbol{t}_0)\big| + o(1)\end{aligned}$$

where $o(1)$ is established by similar methods to that of Claim 1 using the fact that K is a C^2 compactly supported kernel.

Now Corollary 2 establishes that $\big|\partial_{\boldsymbol{h}} B_{n,b_n}(\boldsymbol{t}_0) - \mathsf{E}\partial_{\boldsymbol{h}} B_{n,b_n}(\boldsymbol{t}_0)\big| \xrightarrow{a.s.} 0$. The proof that $M_n \xrightarrow{a.s.} 0$ is similar to Theorem 4 with the exception that now instead of the bound from Claim 3 one can show

$$\mathsf{P}\Big(M_n \geq \frac{R}{nb^4}\Big) \leq L_1 \exp(-L_2 \min\{Rn, R^2\})$$

which is sufficient to prove the claim. □

4. Consistent estimates of $(\boldsymbol{\mu}, \boldsymbol{\tau})$. At this point it becomes advantageous to switch to complex variable notation so that points in the plane $(x, y) \in \mathbb{R}^2$ correspond to points in the complex plane $x + iy \in \mathbb{C}$. Under this correspondence, C^1 diffeomorphisms $f(x, y) = (u(x, y), v(x, y))$ of $\mathbb{R}^2$ can be considered as C^1 diffeomorphisms of $\mathbb{C}$ by writing $f(x+iy) =$

$u(x, y) + iv(x, y)$. For the remainder of the paper we use $\mathrm{Re}f$ and $\mathrm{Im}f$ to denote the real and imaginary parts of the complex representation of the map f.

The main utility of switching to complex notation is that we can directly use the results and techniques of conformal and quasiconformal theory to establish consistent estimates of f. This section starts by defining the complex dilatation μ and log-scale τ of a C^1 diffeomorphism. We then conclude by showing how the results of the previous section can be used to construct consistent estimates of μ and τ. In the next section we show how the estimates of μ and τ can be used to establish consistent estimates of f.

For a function $f \in C^1(U)$ define the complex derivatives

$$\partial f := \frac{1}{2}\left(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}\right), \qquad \overline{\partial} f := \frac{1}{2}\left(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\right).$$

The complex dilatation and the log-scale are then defined as

$$\mu := \overline{\partial} f / \partial f \tag{14}$$

$$\tau := \log |\partial f|. \tag{15}$$

The complex dilatation, μ, characterizes the infinitesimal ellipse with inclination $\arg(-\mu/2)$ and eccentricity $\frac{1+|\mu|}{1-|\mu|}$ that gets mapped to an infinitesimal circle under the image of f. In addition, μ uniquely determines f up to post composition with conformal maps. The log-scale τ is then used to recover the conformal post composition so that, together, μ and τ uniquely determine f up to a rotation and translation. For a short introduction to quasiconformal theory see the appendices of [6] or [5]. For more a complete treatment see [3], [30], [29], [28].

In the previous section we constructed a sequence of functions $B_{n,b_n}(\boldsymbol{t})$ which converge uniformly on compacts as $n \to \infty$ to $(8 - 2^{\alpha+1})\big|J_f^{\boldsymbol{t}}\boldsymbol{h}\big|^{\alpha}$ where $\boldsymbol{h} := (h_1, h_2)$ is a vector of our choice. Using complex variable notation we can write $\big|J_f^{\boldsymbol{t}}\boldsymbol{h}\big| = |h\partial f + \overline{h}\,\overline{\partial} f|$ where $h = h_1 + ih_2$. By factoring out ∂f we get $|h\partial f + \overline{h}\,\overline{\partial} f|^{\alpha} = |\partial f|^{\alpha}|h + \overline{h}\mu|^{\alpha}$. Now by choosing $h = 1, i, 1+i$ (for increments in the north-south, east-west, and diagonal directions) we can construct three functions $W_{1,n}$, $W_{2,n}$ and $W_{3,n}$ which converge to $|\partial f||1 + \mu|$, $|\partial f||1 - \mu|$, and $|\partial f||1 + i + \mu(1 - i)|$ respectively. In particular

$$W_{1,n}(\boldsymbol{t}) \to \left|J_f^{\boldsymbol{t}}\begin{pmatrix}1\\0\end{pmatrix}\right|, \quad W_{2,n}(\boldsymbol{t}) \to \left|J_f^{\boldsymbol{t}}\begin{pmatrix}0\\1\end{pmatrix}\right|, \quad W_{3,n}(\boldsymbol{t}) \to \left|J_f^{\boldsymbol{t}}\begin{pmatrix}1\\1\end{pmatrix}\right|$$

where the convergence is uniform in $\boldsymbol{t}$ on compacts subsets of Ω as $n \to \infty$. Notice that $W_{1,n}$, $W_{2,n}$, and $W_{3,n}$ are the factors of stretching the affine transformation $J_f^{\boldsymbol{t}}$ applies to the lines in the three different directions: horizontal, vertical and diagonal. Therefore, the points, $(W_{1,n}^{-1}, 0)$, $(0, W_{2,n}^{-1})$ and $(W_{3,n}^{-1}, W_{3,n}^{-1})$ asymptotically lie on the ellipse that gets mapped to a circle with unit radius under the affine transformation induced by the matrix $J_f^{\boldsymbol{t}}$. Since the general equation for an ellipse is $ax^2 + bxy + cy^2 = 1$ we have

$$\begin{aligned} a_n &:= W_{1,n}^2 \to a \\ c_n &:= W_{2,n}^2 \to c \\ b_n &:= W_{3,n}^2 - W_{1,n}^2 - W_{2,n}^2 \to b. \end{aligned}$$

The area of the ellipse specified by $ax^2 + bxy + cy^2 = 1$ is $2\pi/\sqrt{4ac - b^2}$. Since $J_f^{\boldsymbol{t}}$ sends the ellipse $ax^2 + bxy + cy^2 = 1$ to the unit circle, we have that $\pi = \det(J_f^{\boldsymbol{t}})\frac{2\pi}{\sqrt{4ac-b^2}}$, which gives

$$\sqrt{4a_nc_n - b_n^2} \to 2\det(J_f^{\boldsymbol{t}}) = 2(|\partial f|^2 - |\overline{\partial} f|^2).$$

Now to solve for (μ, τ), first notice that $a_n + c_n \to |\partial f|^2(|1+\mu|^2 + |1-\mu|^2) = 2|\partial f|^2 + 2|\overline{\partial} f|^2$. Therefore $\sqrt{4a_nc_n - b_n^2} + a_n + c_n \to 4|\partial f|^2$. Similarly

$$4\,\mathrm{Re}(\mu) = |1+\mu|^2 - |1-\mu|^2 \tag{16}$$

$$4\,\mathrm{Im}(\mu) = |1+i+\mu(1-i)|^2 - |1+\mu|^2 - |1-\mu|^2 \tag{17}$$

which gives

$$\frac{a_n - c_n}{\sqrt{4a_nc_n - b_n^2} + a_n + c_n} \longrightarrow \mathrm{Re}(\mu) \tag{18}$$

$$\frac{b_n}{\sqrt{4a_nc_n - b_n^2} + a_n + c_n} \longrightarrow \mathrm{Im}(\mu) \tag{19}$$

$$\log(\sqrt{4a_nc_n - b_n^2} + a_n + c_n) \longrightarrow 2\tau + \log 4. \tag{20}$$

Therefore under the conditions of Theorem 4 we can construct estimates $\hat{\mu}$ and $\hat{\tau}$ that converge to μ and τ, respectively, where the convergence is uniform on compact subsets of Ω with probability one as $n \to \infty$. Moreover, under the extra conditions of Theorem 5, $\partial\hat{\mu} \to \partial\mu$ uniformly on compact subsets of Ω with probability one.

5. Estimating $\boldsymbol{f}$. In this section we show how to construct an estimate $\hat{f}$ on a simply connected domain U such that $\overline{U} \subset \Omega$. Then we show that $\hat{f}$ converges to f uniformly on compact subsets of U. The construction of $\hat{f}$ is on U instead of Ω because we need the uniform convergence of $\hat{\mu}$, $\partial\hat{\mu}$ and $\hat{\tau}$ to establish the convergence of $\hat{f}$. It is open as to whether one can construct $\hat{f}$ on the full observation region which converges uniformly on compact subsets.

We start, in Subsection 5.1, with a discussion on how f can be recovered, uniquely up to a rotation and translation, from the true μ and τ. This will indicate how we recover $\hat{f}$ from the estimated $\hat{\mu}$ and $\hat{\tau}$. Finally we show $\hat{f} \to f$ uniformly on compact subsets of U in Section 5.3.

5.1. *Recovering f from (μ, τ).* First let U be a simply connected domain such that $\overline{U} \subset \Omega$. The C^1 diffeomorphism f now satisfies $f = g \circ f_\mu$ on U where f_μ is the unique normalized quasiconformal map with dilatation μ which maps U to the unit disk, $\mathbb{D}$ (see the appendix in [5]). Since f and f_μ have the same complex dilatation, $g = f \circ f_\mu^{-1}$ is a conformal map defined on $\mathbb{D}$. It turns out this decomposition of f is useful in that the complex dilatation μ determines f_μ and τ is used to recover the conformal map g.

To see how to recover g from τ notice that $\partial f = \partial(g \circ f_\mu) = (g' \circ f_\mu)\partial f_\mu$. Therefore

$$\log|g'| = \log|\partial f| \circ f_\mu^{-1} - \log|\partial f_\mu| \circ f_\mu^{-1}. \tag{21}$$

Since g is conformal on $\mathbb{D}$, $\log g'$ is holomorphic on $\mathbb{D}$. Moreover $\log g' = \log|g'| + i\arg(g')$. Therefore, using (21), $\operatorname{Re}\log g' = \tau \circ f_\mu^{-1} - \log|\partial f_\mu| \circ f_\mu^{-1}$, which can be recovered from (μ, τ). Since the real and imaginary parts of holomorphic maps are harmonic conjugates, which are unique up to a constant, we can recover $\operatorname{Im}\log g' + \theta$ where $\theta \in \mathbb{R}$ is an unknown factor. Now by exponentiating we can recover $e^{i\theta} g'$. Then

$$e^{i\theta} g(z) + c = \int_{z_0}^{z} e^{i\theta} g'(w)dw$$

where the integral is taken over a line connecting z_0 to z. Therefore μ and τ are sufficient to recover $f = g \circ f_\mu$ on U up to a rotation and translation.

5.2. *Constructing $\hat{f}$ from $(\hat{\mu}, \hat{\tau})$.* The technique for recovering $\hat{f}$ in Section 5.1 would work if we knew that there existed a quasiconformal map $\hat{f}$ with complex dilatation $\hat{\mu}$ such that $\hat{\tau} = \log|\partial\hat{f}|$. Unfortunately, there is no simple condition on $\hat{\tau}$ for the existence of such an $\hat{f}$. The main problem is that we don't precisely measure $\log|g'|$, which is required to be harmonic. Instead we only have an estimate of $\log|g'|$. The estimate, which is motivated by (21), is defined by

$$\widehat{\log|g'|} := \hat{\tau} \circ f_{\hat{\mu}}^{-1} - \log|\partial f_{\hat{\mu}}| \circ f_{\hat{\mu}}^{-1}. \tag{22}$$

Since $\widehat{\log|g'|}$ is not guaranteed to be harmonic it may not always be possible to find the harmonic conjugate used to recover $\log g'$. In what follows we notice that $\log g'$ is in the Bergman space of holomorphic functions with finite L_2 integrals. We then use the Bergman projection to find a holomorphic function whose real part approximates $\widehat{\log|g'|}$.

We define our estimate $\hat{f}$ of f in the region U as

$$\hat{f} = \hat{g} \circ f_{\hat{\mu}} \tag{23}$$

where $f_{\hat{\mu}}$ is the unique normalized quasiconformal map sending U to $\mathbb{D}$ and the function $\hat{g}$ is the holomorphic map, unique up to translations, defined on the unit disk $\mathbb{D}$, satisfying

$$\hat{g}' = \exp\big(\mathcal{P}\widehat{\log|g'|}\big)$$

where the operator $\mathcal{P}$ is defined by

$$\mathcal{P}h(w) := \frac{2}{\pi}\int_{\mathbb{D}} \frac{h(z)}{(1-\overline{z}w)^2}dxdy - \mathrm{Re}h(0)$$

where $z = x+iy$. The integral transform in the above definition is the Bergman Projection (see [17] for an introduction to Bergman spaces).

To motivate our choice of operator $\mathcal{P}$, we first mention that the true conformal map g satisfies

$$\begin{aligned}(\mathcal{P}\log|g'|)(w) &= \mathcal{P}\left(\frac{\log g' + \overline{\log g'}}{2}\right)(w)\\ &= \frac{1}{\pi}\int_{\mathbb{D}}\frac{\log g'(z)}{(1-\overline{z}w)^2}dxdy + \frac{1}{\pi}\int_{\mathbb{D}}\frac{\overline{\log g'(z)}}{(1-\overline{z}w)^2}dxdy - \log|g'(0)|\\ &= \log g'(w) + \overline{\log g'(0)} - \log|g'(0)|\\ &= \log g'(w) + i\theta\end{aligned}$$

where $\theta = -\mathrm{Im}\log g'(0)$. In the above computation we used the fact that for any conformal map g defined on $\mathbb{D}$, the holomorphic function $\log g'$ is in the Bergman space $A^2(\mathbb{D})$ (this is true by Theorem 9.4 of [38] along with the fact that the Bloch space is a subset of the Bergman space). Therefore the projection $\mathcal{P}$ can be used to recover the harmonic conjugate of $\log|g'|$ up to an unknown constant θ. In what follows we show $\widehat{\log|g'|} \to \log|g'|$ and $\mathcal{P}\widehat{\log|g'|} \to \mathcal{P}\log|g'|$ uniformly on compact subsets of $\mathbb{D}$ as $n\to\infty$.

5.3. *$\hat{f}$ converges to f.* We show that under appropriate conditions $\hat{f}$ converges uniformly on compact subsets of U, with probability 1. First we establish the following lemma.

LEMMA 5. *Suppose $\mu_n, \mu \in C^2(U)$ are complex dilatations such that $\mu_n \xrightarrow{L_\infty(U)} \mu$ and $\partial\mu_n \xrightarrow{L_\infty(U)} \partial\mu$ on a bounded simply connected domain U. Suppose, in addition, one can extend μ_n, μ to functions $\mu_n^*, \mu^* \in C^2(W)$ on a simply connected domain W containing $\overline{U}$ such that $\mu_n \xrightarrow{L_\infty(W)} \mu$ and $\partial\mu_n \xrightarrow{L_\infty(W)} \partial\mu$. Then,*

$$\log|\partial f_{\mu_n}|\circ f_{\mu_n}^{-1} \to \log|\partial f_\mu|\circ f_\mu^{-1} \tag{24}$$

$$\mathcal{P}\log|\partial f_{\mu_n}|\circ f_{\mu_n}^{-1} \to \mathcal{P}\log|\partial f_\mu|\circ f_\mu^{-1}, \tag{25}$$

uniformly on compact subsets of $\mathbb{D}$ as $n\to\infty$.

PROOF. First decompose f_{μ_n} and f_μ so that,

$$f_{\mu_n} = h_n\circ\tilde{f}_n, \qquad f_\mu = h\circ\tilde{f},$$

where $\tilde{f}_n$ and $\tilde{f}$ are normalized quasiconformal maps on the whole plane with complex dilatations obtained by extending μ_n and μ to the whole plane by smoothly truncating to zero away from U. Here h_n and h are conformal maps sending $\tilde{f}_n(U)$ and $\tilde{f}(U)$ to the unit disk $\mathbb{D}$, respectively. Note: the truncation must be done in a way so that $\tilde{\mu}_n$ has uniformly bounded compact support, $\tilde{\mu}_n$ is as smooth as μ_n, $\tilde{\mu}_n \xrightarrow{L_\infty} \tilde{\mu}$ and $\partial\tilde{\mu}_n \xrightarrow{L_\infty} \partial\tilde{\mu}$ (existence guaranteed by the existence of the extensions μ^*, μ_n^* on W). Now, the quasiconformal maps $\tilde{f}_n$ converge uniformly to $\tilde{f}$ on U since $\tilde{\mu}_n \xrightarrow{L_\infty} \tilde{\mu}$ with uniformly bounded support (see Lemma 1 on page 55 of [3]).

Notice $\partial f_{\mu_n} = (h_n' \circ \tilde{f}_n)\partial \tilde{f}_n$ which gives

$$(\log \partial f_{\mu_n}) \circ f_{\mu_n}^{-1} = \log h_n' \circ h_n^{-1} + \log \partial \tilde{f}_n \circ f_{\mu_n}^{-1}. \tag{26}$$

See the appendix in [5] for a discussion on how to define a continuous version of $\log \partial f_{\mu_n}$ and $\log \partial \tilde{f}_n$. To establish (24) and (25) we show that both terms in (26) converge uniformly on compact subsets of $\mathbb{D}$, as well as the result of applying the operator $\mathcal{P}$ to both terms.

For the first term, $\log h_n' \circ h_n^{-1}$, in (26), notice that $h_n^{-1} = \tilde{f}_n \circ f_{\mu_n}^{-1}$ and $h^{-1} = \tilde{f} \circ f_\mu^{-1}$. Since $\tilde{f}_n \to \tilde{f}$ uniformly on U, $\tilde{f}$ is Hölder continuous on U and $f_{\mu_n}^{-1} \to f_\mu^{-1}$ uniformly on compact subsets of $\mathbb{D}$, $h_n^{-1} \to h^{-1}$ uniformly on compact subsets of $\mathbb{D}$ (this follows by Corollary 9, and Lemmas 10 and 13 of [5]). Since the functions $\xi_n := h_n^{-1}$ and $\xi := h^{-1}$ are conformal maps of $\mathbb{D}$, $\log \xi_n'$ and $\log \xi'$ are both holomorphic and $\log \xi_n' \to \log \xi'$ uniformly on compact subsets. Noticing that $\log \xi_n' = -\log h_n' \circ h_n^{-1}$ and $\log \xi' = -\log h' \circ h^{-1}$ gives

$$\log h_n' \circ h_n^{-1} \to \log h' \circ h^{-1} \tag{27}$$

uniformly on compact subsets of $\mathbb{D}$ as $n \to \infty$. In addition, $\log \xi_n'$ and $\log \xi'$ are both in the Bergman space $A_2(\mathbb{D})$ (and are therefore unaffected by the Bergman projection), which establishes that

$$\mathcal{P} \log h_n' \circ h_n^{-1} \to \mathcal{P} \log h' \circ h^{-1} \tag{28}$$

uniformly on compact subsets of $\mathbb{D}$ as $n \to \infty$.

For the second term, $\log \partial \tilde{f}_n \circ f_{\mu_n}^{-1}$, in (26), notice that the results in the appendix of [5] establish that $\log \partial \tilde{f}$ is Hölder continuous on U and $\log \partial \tilde{f}_n \xrightarrow{L_\infty(U)} \log \partial \tilde{f}$. Therefore, $\log \partial \tilde{f}_n \circ f_{\mu_n}^{-1}$ converges to $\log \partial \tilde{f} \circ f_\mu^{-1}$ uniformly on compact subsets. Moreover, since the continuity of $\log \partial \tilde{f}$ on $\mathbb{C}$ implies it is bounded on U, $\log \partial \tilde{f}_n \circ f_{\mu_n}^{-1}$ also convergences in $L_2(\mathbb{D})$. Therefore

$$\log \partial \tilde{f}_n \circ f_{\mu_n}^{-1} \to \log \partial \tilde{f} \circ f_\mu^{-1}, \tag{29}$$

$$\mathcal{P} \log \partial \tilde{f}_n \circ f_{\mu_n}^{-1} \to \mathcal{P} \log \partial \tilde{f} \circ f_\mu^{-1} \tag{30}$$

uniformly on compact subsets of $\mathbb{D}$ as $n \to \infty$. The last convergence is due to the fact that the Bergman projection is a bounded operator on $L_2(\mathbb{D})$ and that convergence in the Bergman space implies convergence on compacts. Finally, (27), (28), (29) and (30) establishes the Lemma. $\square$

THEOREM 6. *Suppose R1-R3, O1, K1-K2, K is a compactly supported kernel, f is a C^3 diffeomorphism, $n \to \infty$, $b \to 0$, and $n^{-1}b^{-4} = O(n^{-\beta})$ for some $\beta > 0$ such that $\beta > 1 - 4\gamma$. Let U be a simply connected open subset of the observation region Ω such that $\overline{U} \subset \Omega$. Then the estimated map $\hat{f}$, defined on U by (23), converges to $e^{i\theta} f + c$ uniformly on compact subsets of U with probability one, where θ is an unidentifiable rotation angle and c is an unidentifiable translation.*

PROOF. The results of Theorem 4 and Theorem 5, along with the comments made in Section 4, establish that $\hat{\mu} \to \mu$, $\partial\hat{\mu} \to \partial\mu$ and $\hat{\tau} \to \tau$ uniformly on U with probability one as $n \to \infty$. Therefore, $f_{\hat{\mu}} \to f_\mu$ uniformly on compact subsets of U (by Corollary 9 of [5]). Now it is sufficient to show that $\hat{g}$ converges uniformly on compact subsets of $\mathbb{D}$ to g (where sufficiency is by Lemma 11 of [5]).

We first show

$$\widehat{\log|g'|} \to \log|g'|, \tag{31}$$

$$\mathcal{P}\widehat{\log|g'|} \to \mathcal{P}\log|g'| \tag{32}$$

uniformly on compact subsets of $\mathbb{D}$ as $n \to \infty$. Remember, $\widehat{\log|g'|}$ is defined by

$$\widehat{\log|g'|} := \hat{\tau} \circ f_{\hat{\mu}}^{-1} - \log|\partial f_{\hat{\mu}}| \circ f_{\hat{\mu}}^{-1}. \tag{33}$$

Lemma 5 immediately establishes the required convergence for the second term, $\log|\partial f_{\hat{\mu}}| \circ f_{\hat{\mu}}^{-1}$. The first term, $\hat{\tau} \circ f_{\hat{\mu}}^{-1}$, converges to $\tau \circ f_\mu^{-1}$ both uniformly on compacts of $\mathbb{D}$ and in $L_2(\mathbb{D})$. This follows by Lemma 13 of [5] since $\hat{\tau} \xrightarrow{L_\infty(U)} \tau$ and that τ is Hölder continuous and bounded on U (since f is assumed to be a C^3 diffeomorphism). Since the Bergman projection is a bounded operator from $L_2(\mathbb{D})$ to $A_2(\mathbb{D})$, we also have that $\mathcal{P}\,\hat{\tau} \circ f_{\hat{\mu}}^{-1} \xrightarrow{A_2(\mathbb{D})} \mathcal{P}\,\tau \circ f_\mu^{-1}$. The convergence is also uniformly on compact subsets by Lemma 12 of [5]. This establishes (31) and (32).

Now by the comments made in Section 5.2, $\mathcal{P}\log|g'| = \log g' + i\theta$. Therefore, by (31) and (32), $\exp(\mathcal{P}\widehat{\log|g'|})$ converges uniformly on compacts to $e^{i\theta}g'$. Since U is simply connected, $\exp(\mathcal{P}\widehat{\log|g'|})$ has an antiderivative $\hat{g}$ such that $\hat{g}' = \exp(\mathcal{P}\widehat{\log|g'|})$ and $\hat{g} \to e^{i\theta}g + c$ uniformly on compact subsets of $\mathbb{D}$. Therefore $\hat{f}$ converges to $e^{i\theta}f + c$ uniformly on compact subsets of U.

$\square$

APPENDIX A: QUASICONFORMAL MAPS

For a short introduction to quasiconformal theory see [6]. For more a complete treatment see [3], [30], [29], [28]. Here we will collect some the standard results from quasiconformal theory. We also include a section where we use the non-homogeneous Beltrami equation to establish the convergence of $\log \partial f_{\mu_n}$. This is not necessarily new, and most likely a somewhat trivial result in mathematics, however the authors could not find it actually written down to reference. None the less, it is necessary for our results, therefore we include it for completeness.

A.1. Quasiconformal maps and the Beltrami equation. For a function $f \in C^1(U)$ define

$$\partial f = \frac{1}{2}\left(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}\right), \qquad \overline{\partial} f = \frac{1}{2}\left(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}\right).$$

If $f \in C^1(U)$ is an orientation-preserving homeomorphism, then f has the following expansion

$$f(z_0 + z) = f(z_0) + \partial f(z_0)(z) + \overline{\partial} f(z_0)(\overline{z}) + o(|z|).$$

This is just equation (1) using complex variable notation. By factoring out ∂f (which is non-zero by the assumption that $\det J_f > 0$) one gets

$$f(z_0 + z) - f(z_0) = \partial f(z_0)(z + \mu(z_0)\overline{z}) + o(|z|)$$

where $\mu(z_0) = \overline{\partial} f(z_0)/\partial f(z_0)$. Now define the complex dilatation and the log-scale (respectively) as

$$\mu := \overline{\partial} f/\partial f \tag{34}$$

$$\tau := \log|\partial f|. \tag{35}$$

Then by the above expansion, μ characterizes the infinitesimal ellipse with inclination $\arg(-\mu/2)$ and eccentricity $\frac{1+|\mu|}{1-|\mu|}$ that gets mapped to an infinitesimal circle under the image of f. We will see that μ uniquely determines f up to post composition with conformal maps.

More generally, if f is locally integrable in a region U, ∂f, $\overline{\partial} f$ are said to be generalized derivatives of f if they are locally integrable and satisfy

$$\iint_U (\partial f)\varphi dxdy = -\iint_U f\partial\varphi dxdy$$
$$\iint_U (\overline{\partial} f)\varphi dxdy = -\iint_U f\overline{\partial}\varphi dxdy$$

for all $\varphi \in C_0^\infty(U)$. Any derivatives written in the following are assumed to be generalized derivatives unless stated otherwise.

In general, let $W_p^k(U)$ be the Sobolev space of all functions with k generalized derivatives with finite L_p norm. For $k \geq 1$ and $0 < \alpha \leq 1$ let $C^{k,\alpha}(U)$ denote the Hölder space of functions with k continuous derivatives which are uniformly Hölder continuous of order α. We will also need local versions of the Sobolev spaces and Hölder spaces. Let $W^k_{p,\,loc}(U)$ consists of the functions which are in $W_p^k(U')$ for all U' such that $\overline{U'} \subset U$. Similarly let $C^{k,\alpha}_{loc}(U)$ consists of the functions which are in $C^{k,\alpha}(U')$ for all U' such that $\overline{U'} \subset U$.

The class of quasiconformal maps are defined as the class of orientation preserving homeomorphisms f such that $f \in W^1_{2,loc}(U)$ and

$$\max_\alpha |\partial_\alpha f| \leq K \min_\alpha |\partial_\alpha f|$$

for some $K > 1$, where ∂_α is the directional derivative in the direction α. The above inequality need to only hold a.e..

At first glance, the $L_{2,loc}$ requirement on the generalized derivative seems to be a bit confusing. Of course, we need the derivative to exist almost everywhere to define the complex dilatation. So why does one need the extra requirement on the regularity for the definition of the quasiconformal maps? The reason is that there exists badly behaved homeomorphism with almost everywhere derivatives (see [30], page 167).

Notice that not much is lost going from continuous derivatives to weak derivatives since we still have the following expansion

$$f(z) = f(z_0) + \partial f(z_0)(z - z_0) + \overline{\partial} f(z_0)(\overline{z} - \overline{z_0}) + o(|z - z_0|)$$

for almost all $z_0 \in U$ (see [30] page 128). However, we gain compactness of the class of quasiconformal maps with $\|\mu\|_\infty \leq k < 1$.

The following is the mapping theorem and, as stated, can be found in [30].

Theorem 7. *Let U and V be conformally equivalent simply connected domains and μ a measurable function on U with $\|\mu\|_\infty < 1$. Then there exists a quasiconformal mapping $f\colon U \to V$ whose complex dilatation coincides with μ almost everywhere. This mapping is uniquely determined up to a conformal mapping of V onto itself.*

Notice also that for a given dilatation, there exist non-homeomorphic solutions to the Beltrami equation. These are called quasiregular functions and are found by post composing a quasiconformal map with a holomorphic function, rather than a conformal one (i.e. holomorphic and homeomorphic). Still we have the nice state of affairs that all of the homeomorphic solutions are conformal maps of one another.

In this paper, for a given complex dilatation μ defined on a simply connected domain U, we define a unique representative quasiconformal map f_μ with dilatation μ by the following appropriate normalization. If μ is defined on $\mathbb{C}$ then set f_μ to be the unique map fixing

$f(0) = 0$, $f(1) = 1$ and $f(z) \to \infty$ as $z \to \infty$. If, however, μ is defined on U which is conformally equivalent to the unit disk we use a different normalization. Fix two points $z_0^U, z_1^U \in U$. For any $\mu\colon U \to \mathbb{D}$ such that $\|\mu\|_\infty \le k < 1$, define $f_\mu\colon U \to \mathbb{D}$ to be the unique quasiconformal map with dilatation μ, normalized by $f(z_0^U) = 0$, $\mathrm{Re} f(z_1^U) > 0$ and $\mathrm{Im} f(z_1^U) = 0$.

Now we have the following theorem (after renormalizing and using Lemma 1 on page 55 of [3]).

THEOREM 8. *Let μ_n and μ be defined on $\mathbb{C}$ such that $\|\mu_n\|_\infty, \|\mu\|_\infty \le k < 1$ with uniformly bounded supports. Then if $\mu_n \to \mu$ a.e. then*

$$f_{\mu_n} \to f_\mu$$

uniformly on compact subsets of $\mathbb{C}$.

After a simple application of the Carathéodory Convergence Theorem (see Theorem 3.1 of [18] for example) one gets the following corollary.

COROLLARY 9. *If $\mu_n \to \mu$ a.e. on a bounded simply connected region U then*

$$f_{\mu_n} \to f_\mu$$

uniformly on compact subsets of U .

The following two theorems are obtained from Theorems 2.4 and 2.5 (pages 87-89) of [44].

THEOREM 10. *Let $\mu\colon U \to \mathbb{D}$ be a complex dilatation such that $\|\mu\|_\infty \le k < 1$. If $\mu \in C^{m,\alpha}_{loc}(U)$ for some $0 < \alpha < 1$ and $m \ge 1$ then $f_\mu \in C^{m+1,\alpha}_{loc}(U)$.*

THEOREM 11. *Let $\mu\colon U \to \mathbb{D}$ be a complex dilatation such that $\|\mu\|_\infty \le k < 1$. If $\mu \in W^m_{p,loc}(U)$ for some $p > 2$ and $m \ge 1$ then $f_\mu \in W^{m+1}_{p,loc}(U)$.*

A.2. The non-homogeneous Beltrami equation. The main purpose of this section is to show that if μ, $\mu_n \in C^2$ with uniformly bounded support such that $\|\mu\|_\infty, \|\mu_n\|_\infty \le k < 1$ and $\partial\mu, \partial\mu_n \in L_p$ (here $p > 2$ satisfies $kC_p < 1$ where C_p is defined in Lemma 3 of [2]), then for any compact subset $\Xi \subset \mathbb{C}$

$$\log|\partial f_{\mu_n}(z)| \stackrel{L_\infty(\Xi)}{\longrightarrow} \log|\partial f_\mu(z)|$$

as $n \to \infty$, if $\|\mu - \mu_n\|_\infty \to 0$ and $\|\partial\mu - \partial\mu_n\|_p \to 0$. In addition we show $\log|\partial f_{\mu_n}|$ and $\log|\partial f_\mu|$ are Hölder continuous of order $1 - 2/p$.

In this section we work on the whole plane and assume every dilatation μ is subject to the constraint $\|\mu\|_\infty \leq k < 1$ and p is such that $p > 2$ and $kC_p < 1$.

We first introduce the Banach space B_p of functions w defined on the whole plane, which satisfy a global Hölder condition of order $1-2/p$, vanish at the origin and have generalized derivatives w_z, $w_{\overline{z}}$ that belong to $L_p(\mathbb{C})$. Define a norm on B_p as follows

$$\|w\|_{B_p} = \sup_{z_1, z_2 \in \mathbb{C}} \frac{|w(z_1) - w(z_2)|}{|z_1 - z_2|^{1-2/p}} + \|w_z\|_p + \|w_{\overline{z}}\|_p.$$

The following theorem is proved in [2]

THEOREM 12. *If $\sigma \in L_p(\mathbb{C})$ the equation*

$$w_{\overline{z}} = \mu w_z + \sigma \tag{36}$$

has a unique solution $w^{\mu,\sigma} \in B_p$. This is the only solution with $w(0) = 0$ and $w_z \in L_p(\mathbb{C})$.

Moreover, if $\|\mu_1\|_\infty, \|\mu_2\|_\infty \leq k < 1$, and $\sigma_1, \sigma_2 \in L_p(\mathbb{C})$, (where $p > 2$ such that $kC_p < 1$), then

$$\|w^{\mu_1,\sigma_1} - w^{\mu_2,\sigma_2}\|_{B_p} \leq c\big(\|\mu_1 - \mu_2\|_\infty \|\sigma_2\|_p + \|\sigma_1 - \sigma_2\|_p\big). \tag{37}$$

The constant c depends only on k and p.

In what follows we will show that under suitable conditions, $\log\left(\frac{\partial f_\mu(z)}{\partial f_\mu(0)}\right) = w^{\mu,\partial\mu}$. However we must first add some conditions on μ to ensure we can construct a differentiable function $h = \log \partial f_\mu(z)$ such that $e^h = \partial f_\mu(z)$.

To this end, we suppose $\mu, \mu_n \in C_0^2(\mathbb{C})$ are complex dilatations with uniformly bounded supports. Notice that $\mu_n, \mu \in C_0^2(\mathbb{C})$ is a sufficient condition for $\mu \in C^{1,\alpha}_{loc}(\mathbb{C}) \cap W^2_{1,loc}(\mathbb{C})$ and $\partial\mu, \partial\mu_n \in L_p(\mathbb{C})$. We will need the following lemma found in [3].

LEMMA 6. *If p and q are continuous and have generalized derivatives that satisfy $\overline{\partial} p = \partial q$ on a simply connected domain Ω, then there exists a function $f \in C^1(\Omega)$ which satisfies $\partial f = p$ and $\overline{\partial} f = q$.*

Since $|\partial f_\mu| > 0$ we can set

$$p := \frac{\partial^2 f_\mu}{\partial f_\mu}, \quad q := \frac{\overline{\partial}\partial f_\mu}{\partial f_\mu}.$$

Notice that since $\mu \in C^{1,\alpha}_{loc}$, p and q are continuous on $\mathbb{C}$ (by Theorem 10) and since $\mu \in W^2_{1,loc}(\mathbb{C})$, p and q have generalized derivatives such that $\overline{\partial} p = \partial q$ (by Theorem 11).

Therefore there exists a function $h \in C^1$ such that $\partial h = p$ and $\overline{\partial} h = q$. Now we notice that such an h suffices as a definition for $\log \partial f_\mu$ by checking that

$$\partial\left(e^{-h}\partial f_\mu\right) = \overline{\partial}\left(e^{-h}\partial f_\mu\right) = 0$$

which implies, by adding a constant to h if necessary, $e^h = \partial f_\mu$. Also notice $\mathrm{Re}(\log \partial f_\mu) = \log|\partial f_\mu|$.

COROLLARY 13. *Suppose μ, $\mu_n \in C^2$ with uniformly bounded support such that $\|\mu\|_\infty, \|\mu_n\|_\infty \leq k < 1$ and $\partial\mu, \partial\mu_n \in L_p$ (here $p > 2$ satisfies $kC_p < 1$ where C_p is defined in Lemma 3 of [2]), then for any compact subset $\Xi \subset \mathbb{C}$*

$$\log|\partial f_{\mu_n}(z)| \xrightarrow{L_\infty(\Xi)} \log|\partial f_\mu(z)|$$

as $n \to \infty$, if $\|\mu - \mu_n\|_\infty \to 0$ and $\|\partial\mu - \partial\mu_n\|_p \to 0$. Moreover, $\log|\partial f_{\mu_n}|$ and $\log|\partial f_\mu|$ are Hölder continuous of order $1 - 2/p$.

PROOF. Switching to a different normalization let $\tilde{f}_\mu = z + w^{\mu,\mu}$ and $\tilde{f}_{\mu_n} = z + w^{\mu_n,\mu_n}$ so that $f_\mu = \tilde{f}_\mu(z)/\tilde{f}_\mu(1)$ and $f_{\mu_n} = \tilde{f}_{\mu_n}(z)/\tilde{f}_{\mu_n}(1)$. In what follows we show the result for $\tilde{f}_\mu$ and $\tilde{f}_{\mu_n}$ then renormalize.

We first show $\log\left(\frac{\partial\tilde{f}_\mu(z)}{\partial\tilde{f}_\mu(0)}\right) = w^{\mu,\partial\mu}(z)$. Since clearly $\log\partial\tilde{f}_\mu$ satisfies (36) it will be sufficient to show that $\frac{\partial^2\tilde{f}_\mu}{\partial\tilde{f}_\mu} \in L_p(\mathbb{C})$. The support of μ is bounded and therefore $\tilde{f}_\mu$ is conformal on the complement of a ball $\mathcal{B}_r(0)$ of radius r about the origin. Since $\partial\tilde{f}_\mu - 1 \in L_p(\mathbb{C})$, $\partial\tilde{f}_\mu$ has a removable singularity at ∞ and therefore $\partial\tilde{f}_\mu(z) = 1 + O(|z|^{-1})$ and $\partial^2\tilde{f}_\mu(z) = O(|z|^{-2})$. Thus $\frac{\partial^2\tilde{f}_\mu}{\partial\tilde{f}_\mu} \in L_p(\mathbb{C})$.

Using the bound (37), $\log\left(\frac{\partial\tilde{f}_{\mu_n}(z)}{\partial\tilde{f}_\mu(0)}\right)$ converges uniformly on compact subsets of $\mathbb{C}$. Indeed, this implies that for any $z_0 \in \mathbb{C}$

$$\log\left(\frac{\partial\tilde{f}_{\mu_n}(z)}{\partial\tilde{f}_{\mu_n}(z_0)}\right) \to \log\left(\frac{\partial\tilde{f}_\mu(z)}{\partial\tilde{f}_\mu(z_0)}\right). \tag{38}$$

uniformly on compact subset $\Xi \subset \mathbb{C}$.

Now, the bound (37) implies $\partial w^{\mu_n,\mu_n} \xrightarrow{L_p} \partial w^{\mu,\mu}$. Since μ_n and μ have uniformly bounded support, $\partial w^{\mu_n,\mu_n}$ and $\partial w^{\mu,\mu}$ are conformal on the complement of the ball $B_r(0)$ for some sufficiently large $r > 0$. Therefore $\partial w^{\mu_n,\mu_n} \to \partial w^{\mu,\mu}$ uniformly on compact subsets of $\mathcal{B}_r^c(0)$ (this is a variation of Lemma 12). Since $\partial\tilde{f}_\mu = 1 + \partial w^{\mu,\mu}$, we then have $\log\partial\tilde{f}_{\mu_n}(z) \to \log\partial\tilde{f}_\mu(z)$ uniformly on compact subsets of $\mathcal{B}_r^c(0)$. Therefore there exists a $z_0 \in \mathcal{B}_r^c(0)$ such that $\log\partial\tilde{f}_{\mu_n}(z_0) \to \log\partial\tilde{f}_\mu(z_0)$. In combination with (38)

$$\log\partial\tilde{f}_{\mu_n}(z) \xrightarrow{L_\infty(\Xi)} \log\partial\tilde{f}_\mu(z)$$

for all compact subsets $\Xi \subset \mathbb{C}$.

To returning back to the original normalization notice that $\log|\tilde{f}_{\mu_n}(1)| \to \log|\tilde{f}_\mu(1)|$ as $n \to \infty$ and therefore

$$\begin{aligned}\log|\partial f_{\mu_n}(z)| &= \log|\partial \tilde{f}_{\mu_n}(z)| - \log|\tilde{f}_{\mu_n}(1)| \\ &\xrightarrow{L_\infty(\Xi)} \log|\partial \tilde{f}_\mu(z)| - \log|\tilde{f}_\mu(1)| = \log|\partial f_\mu(z)|\end{aligned}$$

for all compact subsets $\Xi \subset \mathbb{C}$. □

A.3. Lemmas. We now show some results that are mentioned in the main body of the paper.

Lemma 7. *If f is a C^1 diffeomorphism then for every compact subset Θ and every fixed $\boldsymbol{h} \neq 0$*

$$\sup_{z\in\Theta}\left|\frac{f(\boldsymbol{t}+\epsilon\boldsymbol{h}) - f(\boldsymbol{t})}{\epsilon} - J_f^{\boldsymbol{t}}\boldsymbol{h}\right| \xrightarrow{\epsilon\to 0} 0. \tag{39}$$

Proof. This is proved using the mean value theorem (see [15]) which can be stated as follows. If f is a mapping of a plane region which contains the line segment S connecting $\boldsymbol{x}_0$ to $\boldsymbol{x}_0 + \boldsymbol{h}$ and f is differentiable on S then

$$|f(\boldsymbol{x}_0+\boldsymbol{h}) - f(\boldsymbol{x}_0)| \le |\boldsymbol{h}| \cdot \sup_{0\le\xi\le 1}\left\|J_f^{\boldsymbol{x}_0+\xi\boldsymbol{h}}\right\|$$

where $\|L\| = \sup_{|\boldsymbol{x}|\le 1}|L(u)|$ for $L \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$. If we apply this theorem to the mapping $\boldsymbol{x} \mapsto f(\boldsymbol{x}) - J_f^{\boldsymbol{x}_0}\boldsymbol{x}$, we get

$$|f(\boldsymbol{x}_0+\boldsymbol{h}) - f(\boldsymbol{x}_0) - J_f^{\boldsymbol{x}_0}\boldsymbol{h}| \le |\boldsymbol{h}| \cdot \sup_{0\le\xi\le 1}\left\|J_f^{\boldsymbol{x}_0+\xi\boldsymbol{h}} - J_f^{\boldsymbol{x}_0}\right\|.$$

Now if $\boldsymbol{x} \mapsto J_f^{\boldsymbol{x}}$ is a continuous map sending $\mathbb{R}^2$ to $\mathcal{L}(\mathbb{R}^2; \mathbb{R}^2)$, then it is uniformly continuous on every compact subset and thus implies (39). Therefore to finish the proof of (39) we notice that $f \in C^1(\mathbb{R}^2)$ implies $\boldsymbol{x} \mapsto J_f^{\boldsymbol{x}}$ is uniformly continuous on any compact set Θ. □

Lemma 8. *Suppose $f \in C^1(\overline{U})$, $\det J_f^{\boldsymbol{x}} \ge k > 0$ on $U \subset \mathbb{R}^2$, then there exists constants $a, b > 0$ such that $a|\boldsymbol{h}| \le |J_f^{\boldsymbol{x}}\boldsymbol{h}| \le b|\boldsymbol{h}|$ for all $\boldsymbol{h}$ and all $\boldsymbol{x} \in U$.*

Proof. Switching to complex notation write $\boldsymbol{h}/|\boldsymbol{h}| = e^{i\theta}$ so that $|J_f^{\boldsymbol{x}}\boldsymbol{h}|/|\boldsymbol{h}| = |\partial f e^{i\theta} + \overline{\partial} f e^{-i\theta}|$. Therefore $\max_{\boldsymbol{h}\neq 0}|J_f^{\boldsymbol{x}}\boldsymbol{h}|/|\boldsymbol{h}| = |\partial f| + |\overline{\partial} f|$ and $\min_{\boldsymbol{h}\neq 0}|J_f^{\boldsymbol{x}}\boldsymbol{h}|/|\boldsymbol{h}| = |\partial f| - |\overline{\partial} f|$. Since $f \in C^1(\overline{U})$, there exits a $b > 0$ such that $\max_{\boldsymbol{h}\neq 0}|J_f^{\boldsymbol{x}}\boldsymbol{h}|/|\boldsymbol{h}|| \le b$ on U. Since $\det J_f^{\boldsymbol{x}} = |\partial f|^2 - |\overline{\partial} f|^2 \ge k > 0$, there exits an $a > 0$ such that $a \le \min_{\boldsymbol{h}\neq 0}|J_f^{\boldsymbol{x}}\boldsymbol{h}|/|\boldsymbol{h}|$ on U. This completes the proof. □

LEMMA 9. *Suppose f is a C^2 diffeomorphism such that $\det J_f > 0$. Also suppose $R\colon \mathbb{R}^+ \to \mathbb{R}$ is an autocovariance function such that $R \in C^4(\mathbb{R}^+ \setminus \{0\})$ and $|R^{(4)}(t)| \leq C_1 t^{\alpha-4}$ for all $t > 0$ sufficiently small, where $\alpha \in (0,2)$. Then for any bounded open set $U \subset \mathbb{R}^2$ and fixed vector $\boldsymbol{h} \neq 0$ there exists a constant C such that*

$$|\partial_{\boldsymbol{h}}^{(2,2)}[R(|f(\boldsymbol{s}) - f(\boldsymbol{t})|)]| \leq C|\boldsymbol{s} - \boldsymbol{t}|^{\alpha-4} \tag{40}$$

for all $\boldsymbol{s}, \boldsymbol{t} \in U$ such that $\boldsymbol{s} \neq \boldsymbol{t}$.

PROOF. First notice that since $f \in C^2$ and $f(\boldsymbol{s}) = f(\boldsymbol{t})$ only when $\boldsymbol{s} = \boldsymbol{t}$, $R(|f(\boldsymbol{s}) - f(\boldsymbol{t})|)$ is twice differentiable in $\boldsymbol{s}$ and $\boldsymbol{t}$ when $\boldsymbol{s} \neq \boldsymbol{t}$. Define $R_{sq}(t) := R(\sqrt{t})$ and $F(\boldsymbol{s},\boldsymbol{t}) := |f(\boldsymbol{s}) - f(\boldsymbol{t})|^2$ so that

$$\partial_{\boldsymbol{h}}^{(2,2)}[R(|f(\boldsymbol{s}) - f(\boldsymbol{t})|)] = \partial_{\boldsymbol{h}}^{(2,2)} R_{sq}(F(\boldsymbol{s},\boldsymbol{t})) \tag{41}$$

$$= \sum_{k=1}^{4} \sum_{\substack{0 \leq j \leq k \\ j+k \leq 4}} R_{sq}^{(k)}(F(\boldsymbol{s},\boldsymbol{t}))(\partial_{\boldsymbol{h}}^{*} F(\boldsymbol{s},\boldsymbol{t}))^{k-j} B_{kj} \tag{42}$$

where the functions B_{kj} are bounded for all $\boldsymbol{s}, \boldsymbol{t} \in U$ such that $\boldsymbol{s} \neq \boldsymbol{t}$ and $\partial_{\boldsymbol{h}}^{*}$ denotes a generic directional derivative on either the variable $\boldsymbol{s}$ or $\boldsymbol{t}$ (therefore $(\partial_{\boldsymbol{h}}^{*} F)^2$ could denote $(\partial_{\boldsymbol{h}}^{(1,0)} F)(\partial_{\boldsymbol{h}}^{(0,1)} F)$, for example). First notice that $|\partial_{\boldsymbol{h}}^{*} F(\boldsymbol{s},\boldsymbol{t})|^{k-j} \leq c_1 |f(\boldsymbol{s}) - f(\boldsymbol{t})|^{k-j}$ for some constant $c_1 > 0$. Secondly notice that our assumptions on R guarantees the existence of a constant c_2 such that $|R_{sq}^{(k)}(t)| \leq c_2 t^{\alpha/2-k}$ for all $t \in (0, \operatorname{diam}(T))$ and $k = 1,2,3,4$. Now

$$\begin{aligned}
|R_{sq}^{(k)}(F(\boldsymbol{s},\boldsymbol{t}))(\partial_{\boldsymbol{h}}^{*} F(\boldsymbol{s},\boldsymbol{t}))^{k-j} B_{kj}| &\leq c_3 |F(\boldsymbol{s},\boldsymbol{t})|^{\alpha/2-k} |f(\boldsymbol{s}) - f(\boldsymbol{t})|^{k-j} \\
&\leq c_4 |f(\boldsymbol{s}) - f(\boldsymbol{t})|^{\alpha-(k+j)} \\
&\leq c_5 |f(\boldsymbol{s}) - f(\boldsymbol{t})|^{\alpha-4}
\end{aligned}$$

where the last inequality follows since $k + j \leq 4$. Therefore there exists a $c_5 > 0$ such that

$$\partial_{\boldsymbol{h}}^{(2,2)} R(|f(\boldsymbol{s}) - f(\boldsymbol{t})|) \leq c_5 |f(\boldsymbol{s}) - f(\boldsymbol{t})|^{\alpha-4} \tag{43}$$

for all $\boldsymbol{s}, \boldsymbol{t} \in U$ such that $\boldsymbol{s} \neq \boldsymbol{t}$.

To finish the proof we show $|f(\boldsymbol{s}) - f(\boldsymbol{t})| \geq c_6 |\boldsymbol{s} - \boldsymbol{t}|$ for $\boldsymbol{s}, \boldsymbol{t} \in U$ and some $c_6 > 0$. Since f is a C^2 diffeomorphism with $\det J_f > 0$, the inverse f^{-1} is differentiable and $J_{f^{-1}}^{\boldsymbol{t}} = \big(J_f^{f^{-1}(\boldsymbol{t})}\big)^{-1}$ for all $\boldsymbol{t} \in \mathbb{R}^2$. Therefore by the Mean Value Theorem

$$|f^{-1}(\boldsymbol{a}) - f^{-1}(\boldsymbol{b})| \leq |\boldsymbol{a} - \boldsymbol{b}| \sup_{\zeta \in f^{-1}(S)} \|(J_f^{\zeta})^{-1}\|$$

where S is the line segment from $\boldsymbol{a}$ to $\boldsymbol{b}$. Now

$$\sup_{\zeta\in f^{-1}(S)} \|(J_f^\zeta)^{-1}\| = \sup_{\zeta\in f^{-1}(S)} \frac{1}{\inf_{|\boldsymbol{u}|=1} |J_f^\zeta \boldsymbol{u}|} = \sup_{\zeta\in f^{-1}(S)} \frac{1}{(|\partial f(\zeta)| - |\overline{\partial} f(\zeta)|)}.$$

Since $\det J_f = |\partial f|^2 - |\overline{\partial} f|^2 > 0$ on $\mathbb{R}^2$, the function $(|\partial f(\zeta)| - |\overline{\partial} f(\zeta)|)^{-1}$ is bounded on compact subsets of $\mathbb{R}^2$. Therefore a constant c_7 can be chosen so that $|f^{-1}(\boldsymbol{a}) - f^{-1}(\boldsymbol{b})| \le c_7|\boldsymbol{a}-\boldsymbol{b}|$ for all $\boldsymbol{a},\boldsymbol{b}\in f(U)$. Therefore

$$|\boldsymbol{s}-\boldsymbol{t}| \le c_7|f(\boldsymbol{t}) - f(\boldsymbol{s})|$$

for all $\boldsymbol{s},\boldsymbol{t}\in U$. □

LEMMA 10. *Suppose f, f_n are continuous one-to-one maps sending a domain Θ onto $\mathbb{D}$ such that f_n converses to f uniformly on compact subsets of Θ. Then f_n^{-1} converges to f^{-1} uniformly on compact subsets of $\mathbb{D}$.*

PROOF. Let F be a closed subset of $\mathbb{D}$ and D be an open set such that $F\subset D$ and $\overline{D}\subset\mathbb{D}$. For notational purposes let $F' := f^{-1}(F)$ and $D' := f^{-1}(D)$. Notice that F' is a closed set and D' is open set such that $F'\subset D'$ and $\overline{D'}\subset\Theta$. Since f_n converges uniformly on compacts to f, $f(F')\subset f_n(D')$ for sufficiently large n. Therefore $f_n^{-1}(F)\subset f^{-1}(D)$ for sufficiently large n.

Now we show f_n^{-1} converges to f^{-1} uniformly on F by contradiction. So suppose there exists a sequence $z_n\in F$ and a $b>0$ such that

$$\left|f_n^{-1}(z_n) - f^{-1}(z_n)\right| > b > 0 \tag{44}$$

for all n. Let n_k be a subsequence such that $f_{n_k}^{-1}(z_{n_k})\to w'$ and $z_{n_k}\to z_0$ as $k\to\infty$. Since f is continuous

$$f(w') = \lim_{k\to\infty} f(f_{n_k}^{-1}(z_{n_k})) = \lim_{k\to\infty} z_{n_k} + \varepsilon_{n_k}(f_{n_k}^{-1}(z_{n_k}))$$

where ε_n converges to zero uniformly on compact subsets of Θ as $n\to\infty$. Since $\{f_{n_k}^{-1}(z_{n_k})\}\subset \overline{f^{-1}(D)}\subset\Theta$, $f(w')=z_0$. Therefore $w' = f^{-1}(z_0) = \lim_{k\to\infty} f^{-1}(z_{n_k})$. This contradicts (44) and therefore f_n^{-1} converges to f^{-1} uniformly on F. □

LEMMA 11. *Suppose $\hat{g}$ and g are holomorphic maps of $\mathbb{D}$ such that $\hat{g}\to g$ uniformly on compact subsets of $\mathbb{D}$. Suppose $f_{\hat\mu}, f_\mu$ are quasiconformal maps defined on Ω and mapping surjectively onto $\mathbb{D}$ such that $f_{\hat\mu}\to f_\mu$ uniformly on compact subsets of Ω. Then $\hat{g}\circ f_{\hat\mu}\to g\circ f_\mu$ uniformly on compact subsets of Ω.*

Proof. Let Θ be a compact subset of Ω.

$$\begin{aligned}\sup_\Theta |\hat g\circ f_{\hat\mu} - g\circ f_\mu| &\leq \sup_\Theta |\hat g\circ f_{\hat\mu} - \hat g\circ f_\mu| + \sup_\Theta |\hat g\circ f_\mu - g\circ f_\mu| \\ &= \sup_\Theta |\hat g\circ f_{\hat\mu} - \hat g\circ f_\mu| + \sup_{f_\mu(\Theta)} |\hat g - g|\end{aligned}$$

Since f_μ is continuous $f_\mu(\Theta)$ is a compact subset and therefore the second term $\sup_{f_\mu(\Theta)}|\hat g - g| \rightarrow 0$. For the first term, notice that we can find a compact subset K of $\mathbb{D}$ such that $\cup f_{\hat\mu}(\Theta)\subset K$ (this is because $\sup_\Theta |f_{\hat\mu} - f_\mu|\rightarrow 0$). Therefore

$$\sup_\Theta |\hat g\circ f_{\hat\mu} - \hat g\circ f_\mu| \leq \sup_K |\hat g'| \sup_\Theta |f_{\hat\mu} - f_\mu|.$$

Since $\hat g$ converges uniformly on compacts and are holomorphic functions, $\hat g'$ converges uniformly on compacts and therefore $\sup_K |\hat g'|$ is bounded. $\square$

Lemma 12. *Let h_n, h be holomorphic maps of $\mathbb{D}$ such that $h_n \xrightarrow{L^2(\mathbb{D})} h$ then $h_n\rightarrow h$ where the convergence is uniform on compact subsets of $\mathbb{D}$.*

Proof. Since $h_n - h$ is holomorphic, mean value theorem says that for any disk $D_r(w)\subset \mathbb{D}$ centered at $w\in\mathbb{D}$ with radius r

$$h_n(w) - h(w) = \frac{1}{\pi r^2}\int_{D_r(w)} h_n(z) - h(z) dxdy.$$

Therefore

$$\begin{aligned}|h_n(w) - h(w)| &\leq \left|\frac{1}{\pi r^2}\int_{D_r(w)} h_n(z) - h(z)dxdy\right| \\ &\leq \frac{1}{\sqrt{\pi} r}\left(\int_{D_r(w)} |h_n(z) - h(z)|^2 dxdy\right)^{1/2} \quad \text{by Hölder} \\ &\leq \frac{1}{\sqrt{\pi} r}\|h_n - h\|_{L^2(\mathbb{D})}.\end{aligned}$$

Now, for any compact subset $\Gamma\in\mathbb{D}$, we can find a $0<r_0<1$ so that $\overline{D_{r_0}(w)}\subset\mathbb{D}$ for all $w\in\Gamma$. Therefore $\sup_{w\in\Gamma}|h_n(w) - h(w)| \leq \frac{1}{\sqrt{\pi} r_0}\|h_n - h\|_{L^2(\mathbb{D})}$ which proves the lemma. $\square$

Lemma 13. *Let U be a simply connected domain, $g_n, g\colon U\rightarrow\mathbb{C}$ and $f_n, f\colon\mathbb{D}\rightarrow U$. If $g_n \xrightarrow{L_\infty(U)} g$, g is Hölder continuous on U, and $f_n\rightarrow f$ uniformly on compact subsets of $\mathbb{D}$ then $g_n\circ f_n \rightarrow g\circ f$ uniformly on compact subsets of $\mathbb{D}$. Moreover, if g is bounded, then the convergence is $L_q(\mathbb{D})$ for all $q\geq 1$.*

PROOF. Let Γ be a compact subset of $\mathbb{D}$ and let C be the Hölder constant for g. Then

$$\begin{aligned}\sup_{\Gamma}|g_n \circ f_n - g\circ f| &\leq \sup_{\Gamma}|g_n\circ f_n - g\circ f_n| + \sup_{\Gamma}|g\circ f_n - g\circ f| \\ &\leq \|g_n - g\|_\infty + C\sup_{\Gamma}|f_n - f|\end{aligned}$$

where $\|g_n - g\|_\infty \to 0$ and $\sup_\Gamma |f_n - f| \to 0$ by assumption.

In addition, if g is bounded on U, then there exists $B > 0$ such that $|g_n \circ f_n| \leq B$ on $\mathbb{D}$ for all sufficiently large n. Therefore by the Dominated Convergence Theorem

$$g_n \circ f_n \xrightarrow{L_q(\mathbb{D})} g\circ f$$

for any $q \geq 1$. □

LEMMA 14. *Let h_n, h be holomorphic maps of $\mathbb{D}$ and suppose $h_n' \to h'$ uniformly on compact subsets of $\mathbb{D}$. Then $h_n - h_n(0) \to h - h(0)$ uniformly on compact subsets of $\mathbb{D}$.*

PROOF. For any $w \in \mathbb{D}$, $h_n(w) - h_n(0) = \int_0^w h_n'(z)dz$ and $h(w) - h(0) = \int_0^w h'(z)dz$, where the integration is taken over the line connecting w and 0. Let $\Gamma \subset \mathbb{D}$ be compact and let D_r be the disk about 0 with radius $r < 1$ such that $\Gamma \subset D_r$. Then,

$$\begin{aligned}\sup_{w\in\Gamma}\big|h_n(w) - h_n(0) - h(w) + h(0)\big| &= \sup_{w\in\Gamma}\Big|\int_0^w (h_n'(z) - h'(z))dz\Big| \\ &\leq \sup_{w\in\overline{D_r}}\Big(|w|\sup_{t\in[0,1]}\big|h_n'(tw) - h'(tw)\big|\Big) \\ &\leq r\sup_{w\in\overline{D_r}}\big|h_n'(w) - h'(w)\big| \xrightarrow{n\to\infty} 0.\end{aligned}$$

□

REFERENCES

[1] R. J. Adler and R. Pyke. Uniform quadratic variation for Gaussian processes. *Stoch. Proc. Appl.*, 48:191–209, 1993.

[2] L. Ahlfors and L. Bers. Riemann's mapping theorem for variable metrics. *Ann. of Math.*, 72:385–404, 1960.

[3] L. V. Ahlfors. *Lectures on Quasiconformal Mappings (with additional chapters by C. J. Earle, I. Kra, M. Shishikura, J.H. Hubbard)*, volume 38 of *University Lecture Series*. Amer. Math. Soc., Providence, RI., 2006.

[4] E. B. Anderes. *Estimating Deformations of Isotropic Gaussian Random Fields*. PhD thesis, University of Chicago, 2005.

[5] E. B. Anderes and S. Chatterjee. Consistent estimates of deformed isotropic gaussian random fields on the plane. Technical Report 739, Statistics Department. University of California at Berkeley, `http://www.stat.berkeley.edu/tech-reports/739.pdf`, 2008.

[6] E. B. Anderes and M. L. Stein. Estimating deformations of isotropic gaussian random fields on the plane. *Ann. Stat.*, 36:719–741, 2008.

[7] G. Baxter. A strong limit theorem for Gaussian processes. *Proc. Amer. Math. Soc.*, 7:522–527, 1956.

[8] A. Benassi, S. Cohen, J. Istas, and S. Jaffard. Identification of filtered white noises. *Stoch. Proc. Appl.*, 75:31–49, 1998.

[9] S. M. Berman. A version of the Lévy-Baxter theorem for the increments of Brownian motion of several parameters. *Proc. Amer. Math. Soc.*, 18:1051–1055, 1967.

[10] M. Clerc and S. Mallat. The texture gradient equation for recovering shape from texture. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24:536–549, 2002.

[11] M. Clerc and S. Mallat. Estimating deformations of stationary processes. *Ann. Stat.*, 31:1772–1821, 2003.

[12] S. Cohen, X. Guyon, O. Perrin, and M. Pontier. Identification of an isometric transformation of the standard brownian sheet. *J. of Stat. Planning and Inference.*, 136:1317–1330, 2006.

[13] S. Cohen, X. Guyon, O. Perrin, and M. Pontier. Singularity functions for fractional processes: application to the fractional brownian sheet. *Ann. Inst. Henri Poincaré*, PR 42:187–205, 2006.

[14] D. Damian, P. Sampson, and P. Guttorp. Bayesian estimation of semi-parametric non-stationary spatial covariance structures. *Environmetrics*, 12:161–178, 2001.

[15] J. Dieudonné. *Foundations of Modern Analysis.* Academic Press, New York, NY, 1960.

[16] R. M. Dudley. Sample functions of the Gaussian process. *Ann. Probability*, 1:66–103, 1973.

[17] P. Duren and A. Schuster. *Bergman Spaces*, volume 100. Mathematical Surveys and Monographs, New York, NY, 2004.

[18] P. L. Duren. *Univalent Functions.* Springer-Verlag, New York, NY, 1983.

[19] E. G. Gladyshev. A new limit theorem for stochastic processes with Gaussian increments. *Theor. Probability Appl.*, 6:52–61, 1961.

[20] J. Gårding. Shape from texture for smooth curved surfaces in perspective projection. *J. Math. Imaging Vision*, 2:327–350, 1992.

[21] X. Guyon and G. Leon. Convergence en loi des h-variations d'un processus gaussien stationnaire. *Ann. Inst. Henri Poincaré*, 25:265–282, 1989.

[22] X. Guyon and O. Perrin. Identification of space deformation using linear and superficial quadratic variations. *Stat. and Prob. Letters*, 47:307–316, 2000.

[23] D.L. Hanson and F.T. Wright. A bound on tail probabilities for quadratic form in independent random variables. *Ann. Math. Stat.*, 42:1079–1083, 1971.

[24] W. Hu. Dark synergy: Gravitational lensing and the cmb. *Phys. Rev. D.*, 65, 2001.

[25] S. Iovleff and O. Perrin. Estimating a nonstationary spatial structure using simulated annealing. *J. of Comput. and Graph. Stat.*, 13:90–105, 2004.

[26] J. Istas and G. Lang. Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann. Inst. Henri Poincaré*, 33:407–436, 1997.

[27] R. Klein and E. Gine. On quadratic variation of processes with Gaussian increments. *Ann. Probability*, 3:716–721, 1975.

[28] S. L. Krushkal'. *Quasiconformal Mappings and Riemann Surfaces.* V. H. Winston & Sons, 1979.

[29] J. Ławrynowicz. *Quasiconformal Mappings in the Plane: Parametrical Methods.* Springer-Verlag, 1983.

[30] O. Lehto and K.I Virtanen. *Quasiconformal Mappings in the Plane.* Springer-Verlag, 1965.

[31] J. Leon and J. Ortega. Weak convergence of different types of variation for biparametric Gaussian processes. *Colloquia Math. Soc. J. Bolyal. 57. Limit Theorems in Probability and Statistics*, 1998.

[32] P. Lévy. Le mouvement brownien plan. *Amer. J. Math.*, 62:487–550, 1940.

[33] W. Loh. Fixed-domain asymptotics for a subclass of matern-type gaussian random fields. *Ann. Stat.*, 33:2344–2394, 2005.

[34] J. Malik and R. Rosenholtz. Computing local surface orientation and shape from texture for curved surfaces. *Int. J. Comp. Vision*, 23(2):149–168, 1997.

[35] O. Perrin. Functional convergence in distribution of quadratic variations for a large class of Gaussian processes: application to a time deformation model. Technical report, Unit of Biometrics at Avignon, 1998.
[36] O. Perrin and W. Meiring. Identifiability for non-stationary spatial structure. *J. of Appl. Prob.*, 36:1244–1250, 1999.
[37] O. Perrin and R. Senoussi. Reducing non-stationary random fields to stationarity and isotropy using a space deformation. *Stat. and Prob. Letters*, 48:23–32, 2000.
[38] C. Pommerenke. *Univalent Functions*. Vandenhoeck and Ruprecht, Göttingen, Germany, 1975.
[39] P. Sampson and P. Guttorp. Nonparametric estimation of nonstationary spatial covariance structure. *J. Amer. Stat. Assoc.*, 87:108–119, 1992.
[40] A. Schmidt and A. O'Hagan. Bayesian inference for nonstationary spatial covariance structure via spatial deformations. *J. Roy. Stat. Soc. B*, 65:745–758, 2003.
[41] M. L. Stein. *Interpolation of Spatial Data: Some Theory for Kriging*. Springer, New York, 1999.
[42] R. Stompor and G. Efstathiou. Gravitational lensing of the cosmic microwave background anisotropies and cosmological parameter estimation. *Monthly Notices of the Royal Astronomical Society*, 302(4):735–747, 1999.
[43] P. T. Strait. On Berman's version of the Lévy-Baxter theorem. *Proc. Amer. Math. Soc.*, 23:91–93, 1969.
[44] I. N. Vekua. *Generalized Analytic Functions*. Pergamon Press, 1962.
[45] H. Zhang. Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *J. Amer. Stat. Assoc.*, 99:250–261, 2004.

Department of Statistics
University of California at Berkeley
367 Evans Hall
Berkeley, CA 94720-3860