

EFFICIENT NONPARAMETRIC BAYESIAN INFERENCE FOR X-RAY TRANSFORMS

BY FRANÇOIS MONARD^{*,1}, RICHARD NICKL^{†,2} AND
GABRIEL P. PATERNAIN^{†,3}

University of California, Santa Cruz^{} and University of Cambridge[†]*

We consider the statistical inverse problem of recovering a function $f : M \rightarrow \mathbb{R}$, where M is a smooth compact Riemannian manifold with boundary, from measurements of general X-ray transforms $I_a(f)$ of f , corrupted by additive Gaussian noise. For M equal to the unit disk with “flat” geometry and $a = 0$ this reduces to the standard Radon transform, but our general setting allows for anisotropic media M and can further model local “attenuation” effects—both highly relevant in practical imaging problems such as SPECT tomography. We study a nonparametric Bayesian inference method based on standard Gaussian process priors for f . The posterior reconstruction of f corresponds to a Tikhonov regulariser with a reproducing kernel Hilbert space norm penalty that does not require the calculation of the singular value decomposition of the forward operator I_a . We prove Bernstein–von Mises theorems for a large family of one-dimensional linear functionals of f , and they entail that posterior-based inferences such as credible sets are valid and optimal from a frequentist point of view. In particular we derive the asymptotic distribution of smooth linear functionals of the Tikhonov regulariser, which attains the semiparametric information lower bound. The proofs rely on an invertibility result for the “Fisher information” operator $I_a^* I_a$ between suitable function spaces, a result of independent interest that relies on techniques from microlocal analysis. We illustrate the performance of the proposed method via simulations in various settings.

1. Introduction. The Radon transform and its variants play a key role in image reconstruction problems, with important applications in physics, engineering and other areas of scientific imaging. The classical case is where a function f in \mathbb{R}^2 is reconstructed from integrals over straight lines:

$$Rf(s, \omega) = \int_{-\infty}^{\infty} f(s\omega + t\omega^\perp) dt, \quad s \in \mathbb{R}, \omega \in S^1,$$

where ω^\perp is the rotation of ω by 90 degrees counterclockwise. Often it is natural to confine the function f to a bounded subset M of Euclidean space such as the unit

Received August 2017; revised February 2018.

¹Supported in part by NSF Grant DMS-17-12790.

²Supported in part by ERC Grant UQMSI/647812.

³Supported in part by by EPSRC Grant EP/M023842/1.

MSC2010 subject classifications. Primary 62G20; secondary 58J40, 65R10, 62F15.

Key words and phrases. Inverse problem, Bernstein–von Mises theorem, MAP estimate, Tikhonov regulariser, Gaussian prior, Radon transform, semiparametric efficiency.

disk, where integrals are now taken along lines connecting boundary points of M . Such transforms constitute the basis for imaging methods such as *computerised tomography* (CT) and *positron emission tomography* (PET), and their mathematical properties are well studied [28, 40].

Two generalisations of the standard Radon transform are important in applications: (a) to model an attenuation or absorption effect within M , for example regions of different levels of biological activity in the physical medium M , and (b) to model anisotropy or physical heterogeneity of M , for instance when “shortest travel times” of waves through the earth follow geodesics of a non-Euclidean metric. The methods used for (a) form the basis for SPECT imaging techniques (see, for instance, [5, 34]) and (b) occurs naturally in seismology, helioseismology and acoustic tomography problems, to mention a few [14, 19, 39, 52]. Both effects can be tackled by the general notion of *attenuated geodesic X-ray transforms* that are given by the formula

$$(1.1) \quad I_a f(x, v) = \int_0^{\tau(x,v)} f(\gamma_{(x,v)}(t)) e^{\int_0^t a(\gamma_{(x,v)}(s)) ds} dt,$$

where (x, v) parametrises the set of geodesics $\gamma_{(x,v)}$ through M , and where $\tau(x, v)$ is the “exit time” of the geodesic started at a point x at the boundary ∂M in the direction of v —see Section 2.1 for precise definitions. The case $a = 0$ corresponds to the case when no attenuation is present, and the ‘geometry’ of M is naturally encoded in the set of geodesics.

The mathematical inverse problem here is to recover f from the line integral values $I_a(f)$ along all geodesics. Explicit reconstruction formulas are available in some specific settings: in the case of the flat disk in \mathbb{R}^2 and when $a = 0$ this was proved in Radon’s celebrated 1917 paper [47], and it has been shown in the last 2 decades that explicit inversion formulas hold also in a variety of other more involved settings, namely, “simple” geometries; see [43], [45] and the paper [37] on numerical implementation. It is, however, generally not clear how the inversion step should be done in case of observations corrupted by statistical noise. The general approach to noisy inverse problems that can be found in the statistical literature is typically based on obtaining a singular value decomposition (SVD) of the forward operator I_a and to then construct a procedure based on spectral regularisation; see, for example, the papers [11, 12, 24, 31–33, 48], just to mention a few. For the standard Radon transform such methods have been suggested in the seminal paper by Johnstone and Silverman [31] where the SVD basis is given by Zernike polynomials. Another approach consists in “rebinning” fan-beam data into parallel data, for which regularisation methods are well understood thanks to the Fourier-slice theorem; see [40]. However, neither approach adapts well to more general X-ray transforms: with attenuation and/or general geodesics, the SVD can rarely be computed analytically; nor is rebinning an option, as the space of geodesics is not homogeneous in general, and this removes the possibility of regularisation methods based on parallel geometry.

In the present paper we follow the Bayesian approach to inverse problems [17, 54] and study a basic nonparametric inference method built around a standard Gaussian prior for the unknown function f which does not require the identification of the SVD basis of I_a . We show how this method can be implemented in a standard way and the resulting maximum a posteriori (MAP) point estimates correspond to a Tikhonov regulariser with a common Sobolev norm penalty, where the Sobolev norm is defined in a classical way (and not implicitly via the SVD of I_a). We prove a Bernstein–von Mises theorem that entails asymptotic normality of various “semiparametric aspects” of the posterior distribution. From it we deduce in particular asymptotic normality and statistical efficiency of the plug-in Tikhonov regulariser for linear integral functionals $\langle f, \psi \rangle_{L^2}$, where ψ is any smooth test function on M . In other words we establish that the semiparametric information bound in this problem is attained by a standard regularisation method that does not require the calculation of the SVD basis. The proof is based on a combination of ideas from Bayesian nonparametric statistics [6, 8, 9] with an inversion result for the “Fisher information” operator $I_a^* I_a$ between suitable function spaces (here I_a^* is a natural adjoint operator defined below).

Heuristically (by analogy to the finite-dimensional linear model) the semiparametric information lower bound for inference on $\langle f, \psi \rangle_{L^2}$ should be $\mathcal{I}_\psi = \|I_a(I_a^* I_a)^{-1} \psi\|_{L^2}^2$, but in our infinite-dimensional setting it has to be clarified for which ψ this quantity is well defined. In Section 4 we invert the “Fisher information” operator by solving the homogeneous Dirichlet boundary value problem for the pseudo-differential operator $I_a^* I_a$, using techniques from micro-local analysis. The mapping properties we deduce imply in particular Theorem 2.2(c) below, which rigorously establishes that \mathcal{I}_ψ exists for all smooth ψ (and equals the information lower bound). In our inversion result for $I_a^* I_a$, a key analytical difficulty, explained in more detail at the outset of Section 4, arises at the boundary ∂M of M : for example, when applied to smooth (say constant) functions, $I_a^* I_a$ can generate singularities at ∂M . And even if one assumes that the unknown f , and thus relevant test functions ψ , are supported *strictly within* M , an application of $(I_a^* I_a)^{-1}$ to such ψ will produce a function that is *fully* supported in M (in view of the nonlocality of the inverse operator). Dealing with boundary issues can therefore not be dispensed with. These nonlocality effects can also be seen in numerical simulations (Example 3).

The connection to partial differential equation (PDE) models just mentioned deserves a final remark: For M a bounded domain in \mathbb{R}^d with smooth boundary ∂M , consider the transport equation

$$(1.2) \quad v \cdot \nabla_x u(x, v) + a(x)u(x, v) = -f(x), \quad x \in M, v \in S^{d-1},$$

subject to the boundary condition $u(x, v) = 0$ for $x \in \partial M, v \cdot \nu(x) \geq 0$, where $\nu(x)$ is the outer normal at x . Along each straight line the transport equation (1.2) becomes an ordinary differential equation that is easily solved to find that the influx

trace of u is precisely the function $I_a(f)$. Our results can thus be cast into the setting of Bayesian inference for parameters of partial differential equations (here f) from noisy observations of their solutions [here $I_a(f)$], studied by A. Stuart and others in the inverse problems literature; see [17, 54] for an overview and [41, 42] for recent related theoretical contributions for parabolic and elliptic PDEs.

This article is organised as follows: In Section 2.1 we introduce general X -ray transforms and state the invertibility theorem for the information operator. In Section 2.2 we propose a Bayesian nonparametric method for inference from noisy X -ray transform data, and in Section 2.3 we give the theoretical results about the performance of the Bayes method and the associated Tikhonov regulariser. All proofs can be found in subsequent sections.

2. Main results.

2.1. *Geodesic X -ray transforms and an inversion result for the information operator.* In this section we introduce the geodesic X -ray transform I of a compact Riemannian manifold with boundary as well as the attenuated version I_a . Our main objective is to establish mapping properties for the normal (information) operator $I_a^* I_a$.

The geodesic X -ray transform acts on functions defined on a compact oriented d -dimensional Riemannian manifold (M, g) with boundary ∂M ($d \geq 2$). In essence, it integrates the function along all possible geodesics running between boundary points. To define the transform with precision we need to introduce some language that conveniently describes the geodesics on a manifold. Geodesics in a Riemannian manifold can be defined in many ways, but for our purposes it suffices to say that they are curves that locally minimize the distance between two points. It turns out that they obey a second order ordinary differential equation on M and thus a geodesic is uniquely determined by its initial position and velocity (i.e., a point in phase space). Geodesics travel at constant speed, so we might as well from now on fix the speed to be one. It is hence convenient to pack positions and velocities together in what we call the *unit sphere bundle* SM . This consists of pairs (x, v) , where $x \in M$ and v is a tangent vector at x with norm $|v|_g = 1$, where g is the inner product in the tangent space at x (i.e., the Riemannian metric).

Unit tangent vectors at the boundary of M constitute the boundary ∂SM of SM and will play a special role. Specifically

$$\partial SM := \{(x, v) \in SM : x \in \partial M\}.$$

We will need to distinguish those tangent vectors pointing inside (“influx boundary”) and those pointing outside (“outflux boundary”), so we define two subsets of ∂SM

$$\partial_{\pm} SM := \{(x, v) \in \partial SM : \pm \langle v, \nu(x) \rangle_g \leq 0\},$$

where $\nu(x)$ is the outward unit normal vector on ∂M at x .

Given $(x, v) \in SM$, we denote by $\gamma_{x,v} : \mathbb{R} \rightarrow M$ the unique geodesic with $\gamma_{x,v}(0) = x$ and $\frac{d\gamma_{x,v}}{dt}(0) = v$ and let $\tau(x, v)$ be the first time when the geodesic $\gamma_{x,v}$ exits M .

We say that (M, g) is *nontrapping* if $\tau(x, v) < \infty$ for all $(x, v) \in SM$. In this paper we will work exclusively with nontrapping manifolds and this is all we need to define the geodesic X -ray transform. Let $C^\infty(W)$ denote the set of infinitely differentiable functions on a manifold W .

DEFINITION 2.1. The *geodesic X -ray transform* of a function $f \in C^\infty(M)$ is the function $I f : \partial_+ SM \rightarrow \mathbb{R}$ given by

$$I f(x, v) = \int_0^{\tau(x,v)} f(\gamma_{x,v}(t)) dt, \quad (x, v) \in \partial_+ SM.$$

In order to obtain good mapping properties for I , we need additional conditions on M . The second condition that we will impose is that M has *strictly convex boundary*, that is, the second fundamental form $\Pi_x(v, v) := \langle \nabla_v v, v \rangle_g$, for v any tangent vector at x , is positive definite for all $x \in \partial M$. This ensures that $I : C^\infty(M) \rightarrow C^\infty(\partial_+ SM)$ since strict convexity of the boundary implies $\tau \in C^\infty(\partial_+ SM)$ ([51], Lemma 4.1.1).

Effectively, the influx boundary $\partial_+ SM$ parametrizes all geodesics going through M . The space of geodesics carries a natural measure (or volume form) which in turn equips $\partial_+ SM$ with the measure

$$d\mu(x, v) := |\langle v(x), v \rangle_g| dx dv$$

and we shall denote $L^2_\mu(\partial_+ SM)$ the space of functions on $\partial_+ SM$ with inner product

$$\langle u, w \rangle_{L^2_\mu(\partial_+ SM)} = \int_{\partial_+ SM} u w d\mu.$$

The measure $d\mu$ is natural in the following sense. If we consider the canonical map

$$\Phi : \{(x, v, t) : (x, v) \in \partial_+(SM); t \in [0, \tau(x, v)]\} \rightarrow SM$$

given by $\Phi(x, v, t) = (\gamma_{(x,v)}(t), \dot{\gamma}_{(x,v)}(t))$ (the geodesic flow) then a calculation shows that

$$(2.1) \quad \Phi^*(\Theta) = |\langle v(x), v \rangle_g| dx dv dt,$$

where Θ is the canonical volume form of SM (also called Liouville form in classical mechanics) and $\Phi^*(\Theta)$ is a new volume form obtained by pulling back Θ via Φ .

It is not hard to prove that I extends as a bounded linear map [51], Theorem 4.2.1,

$$I : L^2(M) \rightarrow L^2_\mu(\partial_+ SM)$$

and hence we have a well defined adjoint $I^* : L^2_\mu(\partial_+ SM) \rightarrow L^2(M)$ that can be easily computed using (2.1). Explicitly

$$I^*w(x) = \int_{S_x M} w^\sharp(x, v) dv,$$

where $w^\sharp(x, v) := w(\gamma_{x,v}(-\tau(x, -v)), \dot{\gamma}_{x,v}(-\tau(x, -v)))$ and $S_x M$ denotes the set of unit tangent vectors at x . In the literature that discusses the standard Radon transform, this operator is usually referred to as *back-projection* operator and appears prominently in the celebrated *filtered back-projection formula* [40, 47] (see [34] for an excellent recent presentation of the classical Radon transform). We can now define the “information operator” $I^*I : L^2(M) \rightarrow L^2(M)$.

The third and final condition that we will impose on M is that it is *free of conjugate points*. Intuitively, this means that beams of geodesics emanating from a point do not focus on or converge to another point (as it would happen for the geodesics on the sphere connecting south and north poles). Equivalently, two points in M are joined by a unique geodesic (note that M nontrapping and with strictly convex boundary implies that M is contractible [56]). This property is fundamental for us since it implies that the information operator is an elliptic pseudo-differential operator of order -1 . Manifolds satisfying the three conditions—nontrapping, strict convexity of the boundary and absence of conjugate points—are called *simple*.

The theory of the X -ray transform is well developed in the case of simple manifolds. If one considers only nontrapping manifolds with strictly convex boundary but allows for conjugate points, the operator I^*I loses its pseudo-differential character. Strict convexity of the boundary is seen as less essential, but dropping it causes technical complications mostly arising from the noncontinuity of the exit time τ .

The discussion above extends without difficulties to the attenuated case. The *attenuated geodesic X -ray transform* $I_a f$ of a function $f \in C^\infty(M)$ with attenuation coefficient $a \in C^\infty(M)$ can be defined as the integral:

$$I_a f(x, v) := \int_0^{\tau(x,v)} f(\gamma_{(x,v)}(t)) \exp\left[\int_0^t a(\gamma_{(x,v)}(s)) ds\right] dt, \quad (x, v) \in \partial_+ SM.$$

The transform I_a extends as a bounded operator $I_a : L^2(M) \rightarrow L^2_\mu(\partial_+ SM)$ with adjoint $I_a^* : L^2_\mu(\partial_+ SM) \rightarrow L^2(M)$. In the case of simple manifolds, the information operator $I_a^* I_a$ displays the same features as $I^* I$.

We will consider noisy observations Y of the X -ray transform $I_a f$ of an unknown function f . If \mathbb{W} is a standard Gaussian white noise in the Hilbert space $L^2_\mu(\partial_+ SM)$ and $\varepsilon > 0$ a noise level, our data is

$$(2.2) \quad Y = I_a f + \varepsilon \mathbb{W}.$$

Up to a discretisation step described in the next section and the usual “Gaussianisation” of Poisson count data, this is a realistic approximate noise model for physical X -ray transform measurements. Assuming this model the following properties of the information operator $I_a^* I_a$ and its inverse will be crucial for the theory

that follows: They imply that the inverse Fisher information exists for a variety of semiparametric inference problems. Their proofs using techniques from microlocal analysis are given in Section 4 below.

THEOREM 2.2. *Let M be a simple manifold and suppose*

$$I_a : C^\infty(M) \rightarrow C^\infty(\partial_+ SM)$$

is injective. Let d_M be any C^∞ function that equals (the Riemannian) $\text{dist}(\cdot, \partial M)$ near the boundary and is positive on the interior of M .

(a) *The information operator $I_a^* I_a$ defines a bijection between $\{d_M^{-1/2} g : g \in C^\infty(M)\}$ and $C^\infty(M)$ and hence has a well defined inverse*

$$(I_a^* I_a)^{-1} : C^\infty(M) \rightarrow \{d_M^{-1/2} g : g \in C^\infty(M)\}$$

such that $I_a^ I_a (I_a^* I_a)^{-1} \psi = \psi$ for all $\psi \in C^\infty(M)$.*

(b) *We have for some constant $c > 0$ that depends only on d, M*

$$\|I_a(d_M^{-1/2} h)\|_{L^2_\mu(\partial_+ SM)} \leq c \|h\|_\infty$$

for every $h \in C(M)$.

(c) *For any $\psi \in C^\infty(M)$ we have that $I_a(I_a^* I_a)^{-1} \psi \in L^2_\mu(\partial_+ SM)$ and*

$$(2.3) \quad \|I_a(I_a^* I_a)^{-1} \psi\|_{L^2_\mu(\partial_+(SM))}^2 < \infty$$

is the Cramér–Rao lower bound (inverse Fisher information) for estimation of the parameter $\langle f, \psi \rangle_{L^2(M)}$ in the model (2.2).

Injectivity of I_a for simple manifolds is known in virtually all cases, so assuming it in the theorem is not a serious restriction. When $a = 0$, injectivity of I is a classical landmark result due to Mukhometov [38]. In two dimensions, injectivity of I_a is known in general [49] and in dimensions ≥ 3 , I_a is injective as long as M admits a strictly convex function [44, 58].

Given parts (a) and (b) the proof of the first assertion in Theorem 2.2(c) is straightforward. The second assertion in Part c) then follows from standard semiparametric theory (Chapter 25 in [59]): An application of Lemma 3.1 below implies that the model (2.2) is locally (asymptotically) normal (LAN) with LAN-norm $\|\cdot\|_{\text{LAN}} = \|I_a(\cdot)\|_{L^2_\mu(\partial_+ SM)}$, and since we have for all $h \in L^2(M)$ that

$$(2.4) \quad \langle I_a h, I_a \tilde{\psi} \rangle_{L^2_\mu(\partial_+ SM)} = \langle h, \psi \rangle_{L^2(M)}, \quad \tilde{\psi} = (I_a^* I_a)^{-1} \psi,$$

we can argue as in Section 7.5 in [41] to deduce the information lower bound $\|\tilde{\psi}\|_{\text{LAN}}^2$ from (2.3). This identifies in particular the (lower bound for the) asymptotic minimax constant

$$(2.5) \quad \liminf_{\varepsilon \rightarrow 0} \inf_{\hat{\psi}} \sup_f \varepsilon^{-2} E_f (\hat{\psi} - \langle f, \psi \rangle)^2 \geq \|I_a(I_a^* I_a)^{-1} \psi\|_{L^2_\mu(\partial_+(SM))}^2,$$

where the infimum is taken over all estimators $\hat{\psi} = \hat{\psi}(Y)$ of $\langle f, \psi \rangle$ based on observations in the model (2.2), and where the supremum in f extends over arbitrary L^2 -neighbourhoods of f_0 of diameter ε .

2.2. *Bayesian inference with Gaussian priors.* We now address the statistical problem of recovering f from a noisy observation of the X-ray transform $I_a f$, and propose numerical illustrations of the feasibility of the approach to general geometries. In what follows, we will take $M = \{(x_1, x_2) \in \mathbb{R}^2, x_1^2 + x_2^2 \leq 1\}$, endowed with either the Euclidean metric $g_e = dx_1^2 + dx_2^2$ (generating the classical Radon transform), or the metric

$$(2.6) \quad \begin{aligned} g(x_1, x_2) &= e^{2\lambda(x_1, x_2)}(dx_1^2 + dx_2^2), \\ \lambda(x_1, x_2) &:= 0.45(e^{-8((x_1-0.3)^2+(x_2-0.3)^2)} - e^{-8((x_1+0.3)^2+(x_2+0.3)^2)}); \end{aligned}$$

see Figure 1. We will concentrate on the unattenuated case $a = 0$ for conciseness. We parameterise $\partial_+ SM$ using *fan-beam coordinates*, defined for $(\beta, \alpha) \in [0, 2\pi) \times (-\pi/2, \pi/2)$ by

$$(\beta, \alpha) \mapsto \left(x = \begin{pmatrix} \cos \beta \\ \sin \beta \end{pmatrix}, v = e^{-\lambda(x)} \begin{pmatrix} \cos(\beta + \pi + \alpha) \\ \sin(\beta + \pi + \alpha) \end{pmatrix} \right) \in \partial_+ SM,$$

with area element $d\mu(x, v) = \cos \alpha \, d\alpha \, d\beta$.

Discretisation. We assume in practice that we are given noisy data at geodesics $\{\gamma_i\}_{i=1}^n$ emanating from a fan-beam sample $\{(\beta_i, \alpha_i)\}_{i=1}^n$, and that the unknown function is modelled as a finite sum $f = \sum_{j=1}^m f_j \phi_j$. Specifically, the domain is a triangular mesh with m nodes x_1, \dots, x_m (see Figure 1, left), so that f_j represents $f(x_j)$ and ϕ_j is a piecewise linear function on the mesh, uniquely defined by the relation $\phi_j(x_k) = \delta_{jk}$. We then seek to reconstruct $X = (f_1, \dots, f_m)^T$ from the observation

$$(2.7) \quad Y = AX + \varepsilon W^{(n)}, \quad Y = (y_1, \dots, y_n)^T,$$

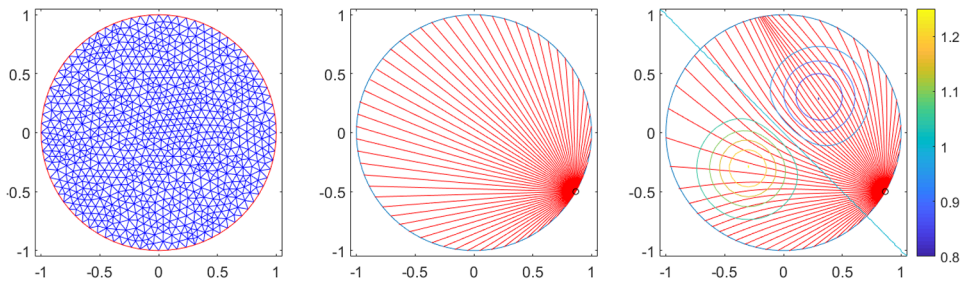


FIG. 1. Left to right: example of a mesh with 886 nodes; geodesics for the Euclidean geometry; geodesics for the metric given in (2.6), superimposed to a contour plot of the “sound speed” $e^{-\lambda}$.

where the discretised forward operator A is an $n \times m$ matrix with entries $A_{ij} = I\phi_j(\beta_i, \alpha_i)$, and $W^{(n)}$ is Gaussian white noise on an *ad hoc* n -dimensional subspace of $L^2_\mu(\partial_+SM)$. By *ad hoc* we mean that for this problem to be a faithful discretisation of the continuous one (2.2), one must endow the domain and range of $A : \mathbb{R}^m \rightarrow \mathbb{R}^n$ with inner products (described by matrices m and n , respectively) which mimick the continuous inner products on $L^2(M)$ and $L^2(\partial_+SM)$: More precisely, if $f = \sum_j f_j\phi_j$ and $f' = \sum_j f'_j\phi_j$, then

$$\int_M f(x)f'(x) dx = \sum_{i,j=1}^m m_{ij} f_i f'_j, \quad m_{ij} := \int_M \phi_i(x)\phi_j(x) dx.$$

Similarly, assuming here that the data comes from a uniform cartesian discretisation of ∂_+SM of size $n = n_\beta n_\alpha$, the n -dimensional subspace of $L^2_\mu(\partial_+SM)$ on which (2.7) is posed has an orthogonal basis $\{e_i\}_{i=1}^n$, where e_i equals 1 on a pixel of dimensions $\frac{2\pi}{n_\beta} \times \frac{\pi}{n_\alpha}$ centered at (β_i, α_i) and 0 elsewhere. A data sample $Y = (y_1, \dots, y_n)^T$ can then be viewed as a function $y = \sum_i y_i e_i \in L^2_\mu(\partial_+SM)$, so that

$$\int_{\partial_+SM} y(\beta, \alpha)y'(\beta, \alpha) \cos \alpha d\alpha d\beta = \sum_{i,j=1}^n n_{ij} y_i y'_j,$$

where

$$n_{ij} = \int_{\partial_+SM} e_i(\beta, \alpha)e_j(\beta, \alpha) \cos \alpha d\alpha d\beta \approx \delta_{ij} \frac{2\pi^2}{n_\beta n_\alpha} \cos \alpha_i.$$

In particular, $\{e_i/\sqrt{n_{ii}}\}_{i=1}^n$ is an orthonormal family, and a realisation of a standard Gaussian white noise on this basis takes the form $\sum_{i=1}^n w_i(e_i/\sqrt{n_{ii}})$ with $w_i \sim \mathcal{N}(0, 1)$. In short, i.i.d. standard Gaussian noise in the “ad hoc” space used for discretisation equals $W^{(n)} \sim \mathcal{N}(0, n^{-1})$ in usual coordinates. The computation of the forward matrix A is now done by solving geodesics by ODEs first, then computing the integrals via Riemann sums, as in [37]. The main difference is that here f is defined on an unstructured triangular mesh generated via the package [62], and interpolating these values at any point is done using barycentric coordinates.

Approach. We take a basic Bayesian approach to this inverse problem: given a Gaussian prior Π on the function f (or its coefficient vector X), we assume the $Y_i|f$ ’s are generated from model (2.7) conditional on f , and obtain the posterior distribution on $f|Y$ by an application of Bayes’ rule. For inference one needs to be able to calculate the posterior distribution, at least approximately. To this end, with the notation above, since $W^{(n)} \sim \mathcal{N}(0, n^{-1})$, then $Y|X \sim \mathcal{N}(AX, \varepsilon^2 n^{-1})$. Assuming the prior distribution is of the form $X \sim \mathcal{N}(0, \sigma^{-1}\Gamma)$ where the prior covariance matrix Γ and the precision parameter σ are known, a standard calculation gives the posterior distribution $X|Y \sim \mathcal{N}(X_c, \Gamma_c)$, where

$$(2.8) \quad \Gamma_c := (\varepsilon^{-2}A^T nA + \sigma \Gamma^{-1})^{-1}, \quad X_c := \varepsilon^{-2}\Gamma_c A^T nY,$$

where A^T denotes the standard matrix transpose. As the posterior distribution is Gaussian, the posterior mean equals the posterior mode X_c (or MAP-estimate) and thus the Tikhonov-regulariser; see [16] or [22]. As a consequence the centre of mass of the posterior distribution is an approximation to the solution of the optimisation problem

$$\min_{f \in V_\Pi} [\varepsilon^{-2} \|Y - I_a f\|_{L^2_\mu(\partial_+ SM)}^2 + \sigma \|f\|_{V_\Pi}^2],$$

discretised into

$$\min_{X \in \mathbb{R}^m} [\varepsilon^{-2} (Y - AX)^T \mathfrak{n}(Y - AX) + \sigma X^T \Gamma^{-1} X],$$

where $V_\Pi \subset L^2(M)$ is the reproducing kernel Hilbert space (RKHS) of Π . (See [22, 23] for standard properties of Gaussian processes and their RKHS.) Natural choices for V_Π are those coming from kernel-type Gaussian process whose covariance is prescribed by a fixed positive definite function $K(\cdot, \cdot)$; see also Remark 2.6 below. In particular, we choose here the Matérn kernel $K(x_i, x_j) = k_{\nu, \ell}(|x_i - x_j|)$, where

$$k_{\nu, \ell}(r) := \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu r}}{\ell}\right)^\nu K_\nu\left(\frac{\sqrt{2\nu r}}{\ell}\right),$$

and where K_ν denotes the modified Bessel function of the second kind. In the examples below, the four parameters $(\varepsilon, \sigma, \nu, \ell)$ are assumed to be known. To address uncertainty on these parameters, hierarchical models can be considered and efficient methods can be derived to compute features of the posterior distribution; see, for example, the recent article [4].

Experiments. The phantoms used are given Figure 2, f_1 is the so-called ‘‘modified Shepp–Logan’’ phantom (compactly supported) and $f_2 = h_2/\sqrt{d_M}$ with $h_2 \in C^\infty(\overline{M})$ and $d_M(x, y) := \frac{1}{2}(1 - x^2 - y^2)$ (as discussed in the next section, the scaling by $\sqrt{d_M}$ is natural in this inverse problem). In all examples, the mesh has $m = 6027$ nodes and we use $n = 14,450$ geodesics. The other parameters are given by $\varepsilon = 10^{-3}$, $\sigma = 1$, $\nu = 1.5$ and $\ell = 0.2$. Sampling the posterior distribution is done by drawing $X = X_c + GZ$, where $Z \sim \mathcal{N}(0, I_{m \times m})$ and G is a matrix satisfying $GG^T = \Gamma_c$ [defined in (2.8)], obtained for instance by Cholesky decomposition (here one may notice that this step is a much cheaper option than computing the SVD of the information operator). To compute forward data, we use the code [37] which allows to produce ‘‘true’’ data with higher accuracy, thereby avoiding the *inverse crime* of using a forward and inverse solver on the same computational grid.

EXAMPLE 1. Euclidean geometry, reconstruction of f_1 from its Euclidean ray transform. We compute the posterior distribution and visualise the mean and sample draws. Results are visualised Figure 3. As the Shepp–Logan phantom has

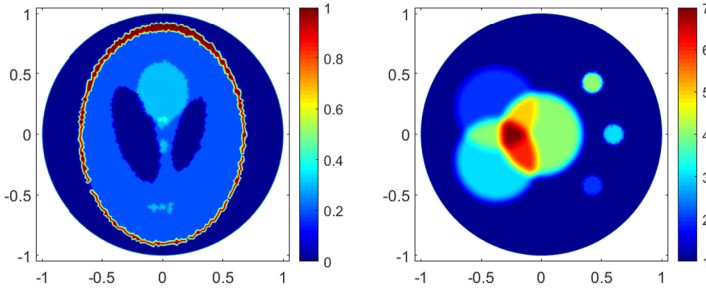


FIG. 2. Left: f_1 , the Shepp–Logan phantom (compactly supported). Right: the function h_2 such that $f_2 := h_2/\sqrt{d_M}$ blows up at the boundary.

spatial variations which may be too sharp to be captured by the prior, we expect over-smoothed reconstructions near sharp edges. This can be seen, for example, on Figure 3 (bottom-right).

EXAMPLE 2. Same as Example 1, except that the geometry is the non-Euclidean one characterized by the metric in (2.6) and geodesics displayed Figure 1 (right). Results are displayed Figure 4, illustrating the applicability of the approach to nonstandard geometries.

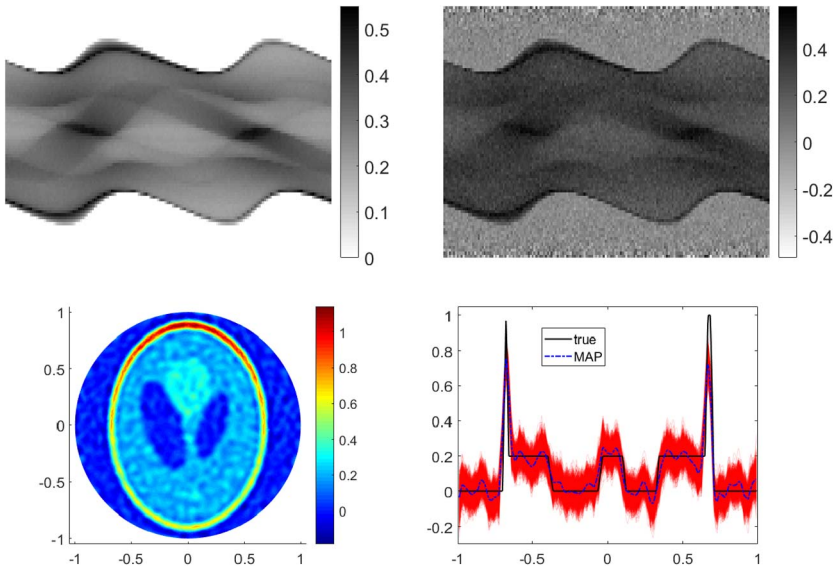


FIG. 3. Example 1. Left to right. Top row: $I f_1$; $I f_1$ noisy (with β on the horizontal axis and α on the vertical axis). Bottom row: posterior mean; cross-section on $\{x_2 = 0\}$ of 2000 posterior samples.

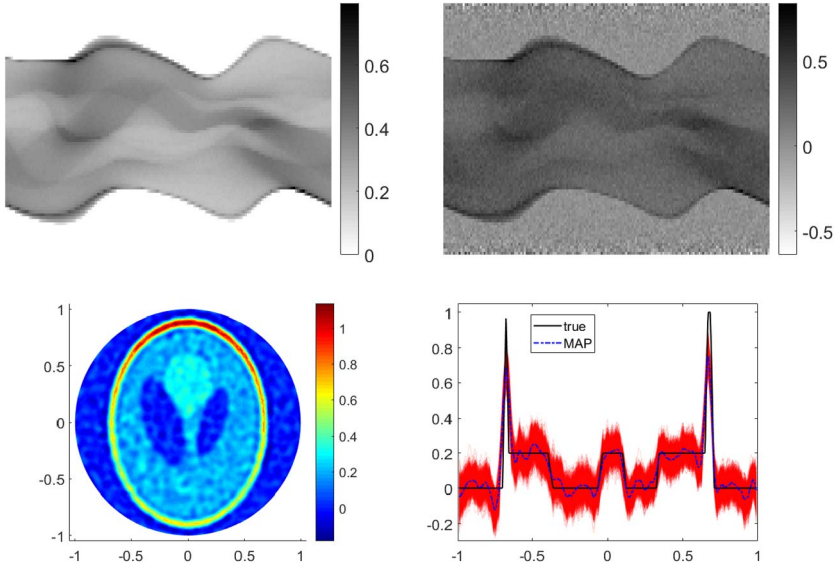


FIG. 4. Example 2. Left to right. Top row: f_1 ; f_1 noisy (with β on the horizontal axis and α on the vertical axis). Bottom row: posterior mean; cross-section on $\{x_2 = 0\}$ of 2000 posterior samples.

EXAMPLE 3. Reconstruction of f_2 in Euclidean geometry, with noise level $\varepsilon = 10^{-2}$ (all other parameters unchanged). As explained in the theory that follows in the next section, an appropriate prior for f_2 should be of the form $d_M^{-1/2}h_2$ where h_2 is drawn from a “standard” Gaussian prior modelling a regular function. For numerical purposes, it should be more stable to work with h_2 , and try to reconstruct h_2 from the transform $I_d h_2 := I(d_M^{-1/2}h_2)$, as the transform I_d naturally compensates for the blowup by integrating. In the implementation, the only change is to work with the discretised version of I_d rather than I (call the corresponding matrix A_d), everything else being kept equal. As may be observed on the middle row of Figure 5, the reconstruction of h_2 is quite robust, especially at the boundary despite the blowup of f_2 there. For comparison, the bottom row of Figure 5 gives the outcome of just inverting for f_2 using A with the usual prior on f_2 instead of h_2 (as in Example 1). As expected, the latter approach is manifestly more unstable near the boundary, and this instability is propagated to the reconstruction in the interior of M , as Figure 5 illustrates.

2.3. Asymptotic normality of the posterior distribution and of the MAP estimator. The statistical methodology laid out in the previous section did not rely on any problem-specific regularisation (it just was based on a standard Gaussian process and the penalty norm of its RKHS), particularly no computation of the SVD of the forward operator I_a was required. One may wonder whether the Bayes solution of this inverse problem can be objectively trusted in the sense that it solves

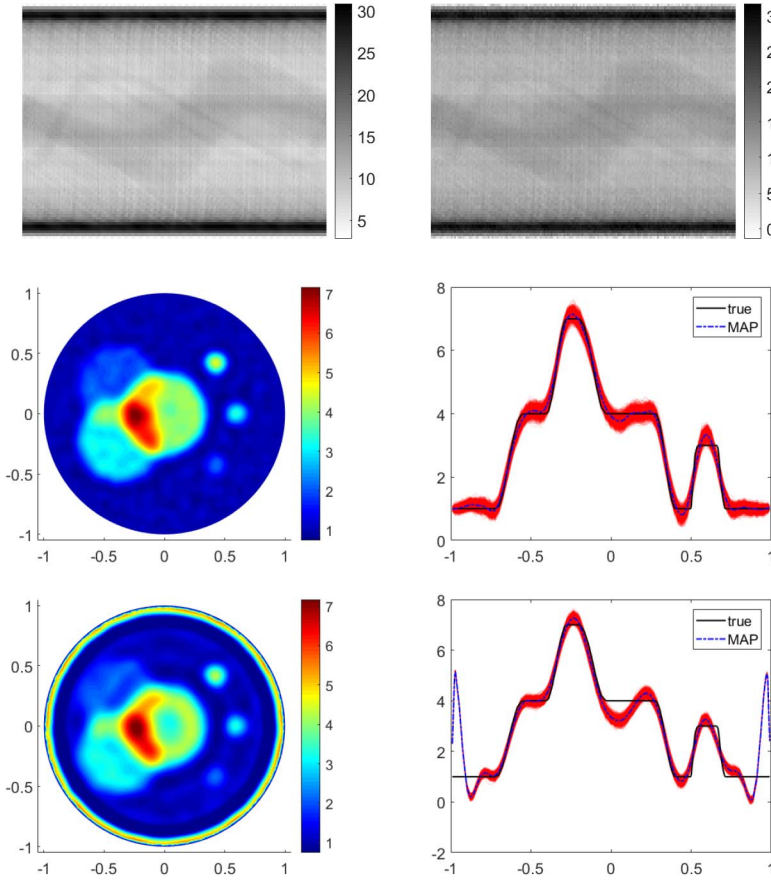


FIG. 5. Example 3. Left to right. Top row: f_2 ; f_2 noisy. Middle row: posterior mean and 2000 cross-sectioned posterior samples for h_2 . Bottom row: posterior mean and 2000 cross-sectioned posterior samples for f_2 , divided by $d_M^{-1/2}$ for comparison with middle row.

the inverse problem in a prior-independent way. We will now show that for the recovery of arbitrary C^∞ -aspects of f , posterior based inference is not only valid but actually optimal from an information theoretic point of view. The theory will be given in the “continuous” Gaussian white noise model (2.2).

We start with a Gaussian Borel probability measure $\tilde{\Pi}$ supported in the space $C(M)$ of bounded continuous functions on M . If $h \sim \tilde{\Pi}$ we let the prior Π for f be the law of the random function

$$(2.9) \quad f(x) = D(h)(x) := h(x)/\sqrt{d_M(x)}, \quad x \in M,$$

where d_M is any function as in Theorem 2.2. By standard arguments (Exercise 2.6.5 in [23] or Lemma I.16 in [22]), if $V_{\tilde{\Pi}}$ is the RKHS of the initial Gaussian measure $\tilde{\Pi}$ then the RKHS V_Π of the induced prior has norm $\|\cdot\|_{V_\Pi} = \|\sqrt{d_M}(\cdot)\|_{V_{\tilde{\Pi}}}$.

The linear mapping D transforms a standard Gaussian prior into one that allows for singularities of functions at the boundary ∂M of a form suggested by Theorem 2.2.

We will now give some precise asymptotic ($\varepsilon \rightarrow 0$) results about the statistical behaviour of the posterior distribution arising from such a prior, under the frequentist assumption that a fixed f_0 generates the observations in (2.2). We will require a mild condition on the prior and on f_0 expressed through the concentration function of the initial probability measure $\tilde{\Pi}$

$$(2.10) \quad \phi_{\tilde{\Pi}, f_0}(\delta) = \inf_{v \in V_{\tilde{\Pi}}, \|v - \sqrt{d_M} f_0\|_\infty \leq \delta} \left[\frac{\|v\|_{V_{\tilde{\Pi}}}^2}{2} - \log \tilde{\Pi}(h : \|h\|_\infty \leq \delta) \right],$$

which characterises the asymptotics of the small ball probabilities $\tilde{\Pi}(h : \|h - \sqrt{d_M} f_0\|_\infty \leq \delta)$ of $\tilde{\Pi}$ as $\delta \rightarrow 0$. The concentration function of Gaussian priors is well studied (see [60] or also Chapter 2.6 in [23] and [22]), and the condition that follows is mild—it can be shown to be satisfied for all sufficiently rich Gaussian processes arising from positive definite kernels K , as soon as $\sqrt{d_M} f_0$ satisfies standard smoothness conditions; see Remark 2.6.

CONDITION 2.3. *Let $f_0 : M \rightarrow \mathbb{R}$ such that $\sqrt{d_M} f_0 \in C(M)$. Let $\tilde{\Pi}$ be a Gaussian Borel probability measure on $C(M)$ whose RKHS $V_{\tilde{\Pi}}$ contains $C^\infty(M)$ and whose concentration function satisfies, for c equal the constant from Theorem 2.2(b) and for some sequence $\delta_\varepsilon \rightarrow 0$ such that $\delta_\varepsilon/\varepsilon \rightarrow \infty$,*

$$(2.11) \quad \phi_{\tilde{\Pi}, f_0}(\delta_\varepsilon/2c) \leq (\delta_\varepsilon/\varepsilon)^2.$$

CONDITION 2.4. *Let $P_{f_0}^Y$ be the law generating the equation $Y = I_a f_0 + \varepsilon \mathbb{W}$, where $I_a : L^2(M) \rightarrow L^2_\mu(\partial_+(SM))$ is the X -ray transform under the conditions of Theorem 2.2, \mathbb{W} is a white noise in $L^2_\mu(\partial_+(SM))$, and $\varepsilon > 0$ is a noise level. Let $\Pi(\cdot|Y)$ be the posterior distribution arising from observing (2.2) under prior $\Pi = \mathcal{L}(f)$, where $\mathcal{L}(f) = \mathcal{L}(d_M^{-1/2} h)$, $h \sim \tilde{\Pi}$, with $\tilde{\Pi}$ satisfying Condition 2.3 for the given f_0 .*

Our main statistical result is the following Bernstein–von Mises theorem for posterior inference on $\langle f, \psi \rangle_{L^2(M)}$ for arbitrary test functions $\psi \in C^\infty(M)$. The idea of its proof is partly inspired by [6–10], where however priors have to be used that are diagonal in the inner product induced by the information operator. This is not the case in the inverse problem setting we consider here, but the invertibility result in Theorem 2.2 combined with an adaptation of ideas in [8, 9] allow to overcome this difficulty. We give the result for smooth ψ but our techniques can be used to obtain results for less regular ψ as well in principle; see Remark 2.10.

We employ the usual notion of weak convergence of laws $\mathcal{L}(X_n) \rightarrow \mathcal{L}(X)$ of real random variables X_n, X that converge in distribution, $X_n \rightarrow^d X$. In (2.12) below we claim convergence of random laws $\mu_n \rightarrow^{\mathcal{L}} \mu$ in probability, which means

that for β any metric for weak convergence of laws (11.3 in [18]), the real random variables $\beta(\mu_n, \mu)$ converge to zero in probability.

THEOREM 2.5. *Assume Condition 2.4. If $f \sim \Pi(\cdot|Y)$, then for every $\psi \in C^\infty(M)$ we have as $\varepsilon \rightarrow 0$ that*

$$(2.12) \quad \mathcal{L}(\varepsilon^{-1}(\langle f, \psi \rangle_{L^2(M)} - \hat{\Psi})|Y) \rightarrow \mathcal{L} \mathcal{N}(0, \|I_a(I_a^* I_a)^{-1} \psi\|_{L^2_\mu(\partial_+(SM))}^2)$$

in $P_{f_0}^Y$ -probability, where

$$\hat{\Psi} = \langle f_0, \psi \rangle_{L^2(M)} - \varepsilon \langle I_a(I_a^* I_a)^{-1} \psi, \mathbb{W} \rangle_{L^2_\mu(\partial_+(SM))}.$$

REMARK 2.6 (Examples of Gaussian priors and of f_0 's). Regarding M as a subset of \mathbb{R}^d , most Gaussian processes that model regular functions in $C(\mathbb{R}^d)$, when restricted to $C(M)$, will satisfy Condition 2.3, if we assume that $\sqrt{d_M} f_0$ is sufficiently regular. For example let $K : \mathbb{R}^d \rightarrow \mathbb{R}$ be a positive definite kernel function whose Fourier transform FK satisfies, for all $\|u\|$ large enough,

$$(2.13) \quad c_0 \|u\|^{-2s} \leq FK(u) \leq c_1 \|u\|^{-2s}, \quad c_0 < c_1, s > d/2,$$

a concrete example being provided by the Matérn kernel (see page 313 in [22] and also Section 9.6 of [1]). For such K we can define a unique centred stationary Gaussian process $(G(x) : x \in \mathbb{R}^d)$ with covariance $EG(x)G(y) = K(x - y)$, $x, y \in \mathbb{R}^d$. The Gaussian process $(G(x) : x \in M)$ obtained by restriction to $M \subset \mathbb{R}^d$ defines a tight Gaussian measure γ_K on $C(M)$, and its RKHS coincides with the standard Sobolev space $H^s(M)$ obtained from restricting elements of $H^s(\mathbb{R}^d)$ to M . Moreover if $f_0 = d_M^{-1/2} \phi_0$ for some $\phi_0 \in H^s(M)$, $s > d/2$, then $\tilde{\Pi} = \gamma_K$ satisfies Condition 2.3 with $\delta_\varepsilon \approx \varepsilon^{2s/(2s+d)}$. Likewise, if ϕ_0 is α -Hölder continuous on M for some $\alpha > 0$ [including the case of arbitrary $f_0 \in C^\infty(M)$], it can be approximated from elements in $H^s(M)$ in $\|\cdot\|_\infty$ -norm and a sequence $\delta_\varepsilon \rightarrow 0$ for which Condition 2.3 holds can still be found. These facts can be proved just as in [22], page 330f.

From the previous theorem we can deduce the asymptotic distribution of the posterior mean $E^\Pi[f|Y]$, which, since the posterior distribution is also a Gaussian measure, equals the posterior mode (MAP estimate). From Corollary 3.10 in [16] (see also Section 11.7 in [22]), MAP estimates can further be seen to equal the Tikhonov-regularisers with RKHS norm as penalty function. Note that in our infinite-dimensional setting the Tikhonov regulariser is defined as the maximiser in f of the Onsager–Machlup functional

$$(2.14) \quad Q(f) = \frac{1}{\varepsilon^2} \langle I_a f, Y \rangle_{L^2_\mu(\partial_+ SM)} - \frac{1}{2\varepsilon^2} \|I_a f\|_{L^2_\mu(\partial_+ SM)}^2 - \frac{1}{2} \|f\|_{V_\Pi}^2.$$

In the discrete setting from Section 2.2 this is equivalent to minimising $Q(f) = \frac{1}{\varepsilon^2} \|Y - I_a f\|^2 + \|f\|_{V_\Pi}^2$ as usual, but in our setting $Y \notin L^2_\mu(\partial_+ SM)$, so the preceding formulation is the appropriate one.

THEOREM 2.7. *Let $\bar{f} = \bar{f}(Y) = E^\Pi[f|Y] \in C(M)$ be the mean of the posterior distribution in Theorem 2.5. Then for every $\psi \in C^\infty(M)$ we have $\langle \bar{f}(Y), \psi \rangle_{L^2(M)} - \hat{\Psi} = o_{P_{f_0}^Y}(\varepsilon)$ as $\varepsilon \rightarrow 0$ and thus also, under $P_{f_0}^Y$,*

$$(2.15) \quad \frac{1}{\varepsilon} \langle \bar{f} - f_0, \psi \rangle_{L^2(M)} \rightarrow^d Z \sim \mathcal{N}(0, \|I_a(I_a^* I_a)^{-1} \psi\|_{L^2_\mu(\partial_+(SM))}^2).$$

In particular in (2.12) in Theorem 2.5 we may replace the centring $\hat{\Psi}$ by $\langle \bar{f}, \psi \rangle_{L^2(M)}$.

REMARK 2.8 (Exact asymptotic minimaxity). The proof of the last theorem implies that convergence of all moments in (2.15) occurs, and hence \bar{f} attains the lower bound constant from (2.5) in the small noise limit. Thus $\langle \bar{f}, \psi \rangle_{L^2(M)}$ is an asymptotically exact minimax estimator of $\langle f_0, \psi \rangle_{L^2(M)}$.

REMARK 2.9 (Confident credible sets). Theorems 2.5 and 2.7 justify the following construction of a confidence set for the Tikhonov regulariser: Consider a credible interval

$$C_\varepsilon = \{x \in \mathbb{R} : |\langle \bar{f}, \psi \rangle - x| \leq R_\varepsilon\}, \quad R_\varepsilon \text{ s.t. } \Pi(C_\varepsilon|Y) = 1 - \alpha,$$

for some given significance level $0 < \alpha < 1$. The frequentist coverage probability of C_ε will satisfy (arguing as in the proof of Theorem 7.3.23 in [23])

$$P_{f_0}^Y((f_0, \psi) \in C_\varepsilon) \rightarrow 1 - \alpha$$

and

$$\varepsilon^{-1} R_\varepsilon \rightarrow P_{f_0}^Y \Phi^{-1}(1 - \alpha)$$

as $\varepsilon \rightarrow 0$. Here Φ^{-1} is the continuous inverse of $\Phi = \Pr(|Z| \leq \cdot)$ with Z as in (2.15). To implement this confidence set we use the posterior sampling method from Section 2.2 to numerically approximate the quantile constants R_ε —computation of $\text{Var}(Z)$, which could be intricate, is not required.

REMARK 2.10 (Extensions). The above theorem shows that semiparametrically efficient recovery of C^∞ aspects of f is possible. Following the program laid out in the papers [6–9, 41] one could in principle proceed to use the estimates in the proof of Theorem 2.5 to derive a result for posterior reconstruction of the entire parameter f in suitable norms via bounding $\varepsilon^{-1} \langle f - f_0, \psi \rangle|Y$ uniformly in collections of functions ψ of bounded Sobolev norms. The approximation theoretic arguments required to do that in the present setting involve delicate boundary issues, with standard Sobolev spaces as approximation scales having to be replaced by the Hörmander spaces introduced in Section 4 below. The execution of these arguments is possible but quite technical and beyond the scope of this paper.

3. Proofs for Section 2.3.

3.1. *Proof of Theorem 2.5.* Let $H_i, i = 1, 2$ be separable Hilbert spaces and consider the equation

$$Y = G(f) + \varepsilon\mathbb{W}, \quad \varepsilon > 0,$$

where $G : H_1 \rightarrow H_2$ is a Borel measurable mapping and \mathbb{W} is a centred Gaussian white noise process ($\mathbb{W}(h) : h \in H_2$) with covariance $E\mathbb{W}(h)\mathbb{W}(g) = \langle h, g \rangle_{H_2}$. Observing Y then means that we observe a realisation of the Gaussian process ($Y(h) = \langle Y, h \rangle_{H_2} : h \in H_2$). We sometimes write $\langle \mathbb{W}, h \rangle_{H_2}$ for the random variable $\mathbb{W}(h)$. Arguing as in Section 7.4 in [41] the posterior distribution of $f|Y$ exists and equals

$$(3.1) \quad \Pi(B|Y) = \frac{\int_B p_f(Y) d\Pi(f)}{\int_{\mathcal{F}} p_f(Y) d\Pi(f)}, \quad B \in \mathcal{B}_{H_1} \quad \text{a Borel set in } H_1,$$

where $p_f(Y)$ is a likelihood function with respect to a suitable dominating measure. The following result is a standard application of the Cameron-Martin theorem (see equation (116) in [41]).

LEMMA 3.1. *Let $\ell(f) = \log p_f(Y)$ and assume $Y = G(f_0) + \varepsilon\mathbb{W}$ for some fixed $f_0 \in H_1$. Then if G is also linear, we have for any $f, g \in H_1$,*

$$\ell(f) - \ell(g) = -\frac{1}{2\varepsilon^2} (\|G(f - f_0)\|_{H_2}^2 - \|G(g - f_0)\|_{H_2}^2) + \frac{1}{\varepsilon} \langle G(f - g), \mathbb{W} \rangle_{H_2}.$$

We now prove Theorem 2.5, and will use the above lemma with $H_1 = L^2(M)$, $H_2 = L^2_\mu(\partial_+ SM)$, $G = I_a$. In what follows the total variation norm between finite measures μ, ν is defined to equal the supremum $\|\mu - \nu\|_{TV} := \sup_B |\mu(B) - \nu(B)|$ over all Borel sets B .

LEMMA 3.2. *Let $\tilde{\Pi}$ be a Gaussian Borel probability measure on $C(M)$, and for fixed f_0 assume its concentration function $\phi_{\tilde{\Pi}, f_0}$ satisfies (2.11) for some $\delta = \delta_\varepsilon \rightarrow 0$. Let Π be the prior for f corresponding to the law of $f = D(h), h \sim \tilde{\Pi}$ with D as in (2.9), and let $\Pi(\cdot|Y)$ be the resulting posterior distribution arising from observing $Y = I_a f + \varepsilon\mathbb{W}$, where I_a is the X-ray transform from Theorem 2.2. Then for any Borel set $D_\varepsilon \subset L^2(M)$ for which*

$$(3.2) \quad \Pi(D_\varepsilon^c) \leq e^{-D_0(\delta_\varepsilon/\varepsilon)^2} \quad \text{for some } D_0 > 3$$

and all $\varepsilon > 0$ small enough, we have

$$(3.3) \quad \Pi(D_\varepsilon^c|Y) \rightarrow 0 \quad \text{and} \quad \|\Pi^{D_\varepsilon}(\cdot|Y) - \Pi(\cdot|Y)\|_{TV} \rightarrow 0$$

as $\varepsilon \rightarrow 0$ in $P_{f_0}^Y$ -probability. Here $\Pi^{D_\varepsilon}(\cdot|Y)$ is the posterior distribution arising from the prior $\tilde{\Pi}(\cdot \cap D_\varepsilon) / \Pi(D_\varepsilon)$ restricted to D_ε and renormalised.

PROOF. It suffices to prove the first limit in (3.3), the second then follows from the basic inequality $\|\Pi^{D_\varepsilon}(\cdot|Y) - \Pi(\cdot|Y)\|_{TV} \leq 2\Pi(D_\varepsilon^c|Y)$.

We have from (3.1) that

$$\Pi(B|Y) = \frac{\int_B e^{\ell(f)-\ell(f_0)} d\Pi(f)}{\int_{\mathcal{F}} e^{\ell(f)-\ell(f_0)} d\Pi(f)}, \quad B \in \mathcal{B}_{L^2(M)},$$

and under $P_{f_0}^Y$ can use Lemma 3.1 to see

$$\ell(f) - \ell(f_0) = -\frac{1}{2\varepsilon^2} \|I_a(f - f_0)\|_{H_2}^2 + \frac{1}{\varepsilon} \langle I_a(f - f_0), \mathbb{W} \rangle_{H_2}.$$

Let ν be any probability measure on the set $B = \{f : \|I_a(f - f_0)\|_{H_2}^2 \leq \delta^2\}$. For any $C > 0$ we have from Jensen's inequality

$$\begin{aligned} P_{f_0}^Y \left(\int_B e^{\ell(f)-\ell(f_0)} d\nu(f) \leq e^{-(1+C)(\delta_\varepsilon/\varepsilon)^2} \right) &\leq \Pr \left(\int_B \left(-\frac{1}{2\varepsilon^2} \|I_a(f - f_0)\|_{H_2}^2 \right. \right. \\ &\quad \left. \left. + \frac{1}{\varepsilon} \langle I_a(f - f_0), \mathbb{W} \rangle_{H_2} \right) d\nu(f) \leq -(1+C) \frac{\delta_\varepsilon^2}{\varepsilon^2} \right) \\ &\leq \Pr \left(\left| \int_B \frac{1}{\varepsilon} \langle I_a(f - f_0), \mathbb{W} \rangle_{H_2} d\nu(f) \right| \geq C \frac{\delta_\varepsilon^2}{\varepsilon^2} \right) \leq e^{-C^2(\delta_\varepsilon/\varepsilon)^2/2} \end{aligned}$$

since the standard Gaussian tail bound $P(|Z - EZ| > u) \leq e^{-u^2/2\text{Var}(Z)}$ applies to the random variable $Z = \int_B \varepsilon^{-1} \langle I_a(f - f_0), \mathbb{W} \rangle_{H_2} d\nu(f)$ which has a centred normal distribution with variance bounded, again using Jensen's inequality, by

$$E \left[\int_B \varepsilon^{-1} \langle I_a(f - f_0), \mathbb{W} \rangle_{H_2} d\nu(f) \right]^2 \leq \varepsilon^{-2} \int_B E \langle I_a(f - f_0), \mathbb{W} \rangle_{H_2}^2 d\nu(f) \leq \frac{\delta^2}{\varepsilon^2},$$

recalling that \mathbb{W} is a centred Gaussian white noise in H_2 . Now we choose $\nu = \Pi(\cdot \cap B)/\Pi(B)$ and let

$$A_\varepsilon = \left\{ \int_B e^{\ell(f)-\ell(f_0)} d\nu(f) \leq e^{-2(\delta_\varepsilon/\varepsilon)^2} \right\},$$

for which $P_{f_0}^Y(A_\varepsilon) \leq e^{-(\delta_\varepsilon/\varepsilon)^2/2} \rightarrow 0$ by what precedes (with $C = 1$). For $E_{f_0}^Y$ the expectation operator corresponding to $P_{f_0}^Y$ and by Markov's inequality, it suffices to prove convergence to zero of

$$E_{f_0}^Y \Pi(D_\varepsilon^c|Y) = E_{f_0}^Y \Pi(D_\varepsilon^c|Y)1_{A_\varepsilon} + E_{f_0}^Y \Pi(D_\varepsilon^c|Y)1_{A_\varepsilon^c}.$$

Since $\Pi(\cdot|Y) \leq 1$ the first quantity is less than $P_{f_0}^Y(A_\varepsilon)$ and hence converges to zero. For the second term we have

$$\begin{aligned}
 & E_{f_0}^Y \Pi(D_\varepsilon^c|Y) 1_{A_\varepsilon^c} \\
 (3.4) \quad & \leq \frac{e^{2(\delta_\varepsilon/\varepsilon)^2}}{\Pi(f : \|I_a(f - f_0)\|_{H_2}^2 \leq \delta_\varepsilon^2)} \int_{D_\varepsilon^c} E_{f_0}^Y [e^{\ell(f) - \ell(f_0)}] d\Pi(f) \\
 & \leq e^{2(\delta_\varepsilon/\varepsilon)^2} e^{\phi_{\Pi, f_0}(\delta_\varepsilon/2)} \Pi(D_\varepsilon^c)
 \end{aligned}$$

noting that $E_{f_0}^Y [e^{\ell(f) - \ell(f_0)}] = 1$ and where

$$(3.5) \quad \phi_{\Pi, f_0}(\delta) = \inf_{w \in V_\Pi, \|I_a(w - f_0)\|_{L_\mu^2(\partial_+ SM)} \leq \delta} \left[\frac{\|w\|_{V_\Pi}^2}{2} - \log \Pi(f : \|I_a f\|_{L_\mu^2(\partial_+ SM)} \leq \delta) \right],$$

using Proposition 2.6.19 and Exercise 2.6.5 in [23], with RKHS $I_a(V_\Pi)$ of $I_a f$ isometric to V_Π since I_a is linear and injective. Now we have for all $\delta > 0$ that $\phi_{\Pi, f_0}(\delta) \leq \phi_{\tilde{\Pi}, f_0}(\delta/c)$ since Theorem 2.2 implies

$$\|I_a f\|_{L_\mu^2(\partial_+ SM)} = \|I_a(D(h))\|_{L_\mu^2(\partial_+ SM)} \leq c\|h\|_\infty$$

so that

$$-\log \Pi(f : \|I_a f\|_{L_\mu^2(\partial_+ SM)} \leq \delta) \leq -\log \tilde{\Pi}(h : \|h\|_\infty \leq \delta/c)$$

as well as

$$\|I_a(w - f_0)\|_{L_\mu^2(\partial_+ SM)} = \|I_a(D(v - D^{-1}f_0))\|_{L_\mu^2(\partial_+ SM)} \leq c\|v - D^{-1}f_0\|_\infty,$$

where $v = \sqrt{d_M}w \in V_{\tilde{\Pi}}$ corresponds to $w = Dv = d_M^{-1/2}v \in V_\Pi$. Thus by (2.11) the right-hand side of (3.4) is bounded above by

$$e^{2(\delta_\varepsilon/\varepsilon)^2} e^{\phi_{\tilde{\Pi}, f_0}(\delta_\varepsilon/2c)} \Pi(D_\varepsilon^c) \leq e^{(3-D_0)(\delta_\varepsilon/\varepsilon)^2} \rightarrow 0$$

for $D_0 > 3$, completing the proof. \square

For $\psi \in C^\infty(M)$ define now $\tilde{\psi} = -(I_a^* I_a)^{-1} \psi$. We have from Theorem 2.2 that $\tilde{\psi}$ can be written as $\tilde{\psi} = d_M^{-1/2} \bar{\psi}$ for some $\bar{\psi} \in C^\infty(M)$. Therefore, since the RKHS $V_{\tilde{\Pi}}$ of $\tilde{\Pi}$ contains $C^\infty(M)$ we have

$$(3.6) \quad \|\tilde{\psi}\|_{V_\Pi}^2 = \|\sqrt{d_M} \tilde{\psi}\|_{V_{\tilde{\Pi}}}^2 = \|\bar{\psi}\|_{V_{\tilde{\Pi}}}^2 \leq C.$$

Next, the random variable $\langle \tilde{\psi}, f \rangle_{V_\Pi}$, $f \sim \Pi$, is $\mathcal{N}(0, \|\tilde{\psi}\|_{V_\Pi}^2)$ and the standard Gaussian tail inequality guarantees for all $u, \delta \geq 0$ that

$$\Pi\left(f : \frac{|\langle \tilde{\psi}, f \rangle_{V_\Pi}|}{\|\tilde{\psi}\|_{V_\Pi}} > u \frac{\delta}{\varepsilon}\right) \leq e^{-u^2(\delta/\varepsilon)^2/2}$$

hence Lemma 3.2 applies to the set

$$D_\varepsilon = \left\{ f : \frac{|\langle \tilde{\psi}, f \rangle_{V_\Pi}|}{\|\tilde{\psi}\|_{V_\Pi}} \leq K \frac{\delta_\varepsilon}{\varepsilon} \right\},$$

whenever $K > \sqrt{6}$, and in deriving the asymptotic distribution of the posterior measure we can restrict to the posterior distribution $\Pi^{D_\varepsilon}(\cdot|Y)$ arising from the prior $\Pi^{D_\varepsilon} = \Pi(\cdot \cap D_\varepsilon)/\Pi(D_\varepsilon)$.

PROPOSITION 3.3. *Assume Condition 2.4. For $\psi \in C^\infty(M)$, define the random variables*

$$(3.7) \quad \hat{\Psi} = \langle f_0, \psi \rangle_{H_1} - \varepsilon \langle I_a(I_a^* I_a)^{-1} \psi, \mathbb{W} \rangle_{H_2}.$$

Then for all $\tau \in \mathbb{R}$ and as $\varepsilon \rightarrow 0$ we have

$$(3.8) \quad E^{\Pi_{D_\varepsilon}} \left[e^{\frac{\tau}{\varepsilon} (\langle f, \psi \rangle_{H_1} - \hat{\Psi})} | Y \right] = e^{\frac{\tau^2}{2} \|I_a(I_a^* I_a)^{-1} \psi\|_{H_2}^2} \times (1 + o_{P_{f_0}^Y}(1)).$$

PROOF. The left-hand side of (3.8) equals, for $f_\tau = f + \tau \varepsilon \tilde{\psi}$,

$$(3.9) \quad \begin{aligned} & E^{\Pi_{D_\varepsilon}} \left[e^{\frac{\tau}{\varepsilon} \langle f - f_0, \psi \rangle_{H_1} + \tau \langle I_a(I_a^* I_a)^{-1} \psi, \mathbb{W} \rangle_{H_2}} | Y \right] \\ &= e^{\tau \langle I_a(I_a^* I_a)^{-1} \psi, \mathbb{W} \rangle_{H_2}} \frac{\int_{\mathcal{F}} e^{\frac{\tau}{\varepsilon} \langle f - f_0, \psi \rangle_{H_1} + \ell(f) - \ell(f_\tau) + \ell(f_\tau)} d\Pi^{D_\varepsilon}(f)}{\int_{\mathcal{F}} e^{\ell(f)} d\Pi^{D_\varepsilon}(f)} \\ &= e^{\frac{\tau^2}{2} \|I_a \tilde{\psi}\|_{H_2}^2} \frac{\int_{D_\varepsilon} e^{\ell(f_\tau)} d\Pi(f)}{\int_{D_\varepsilon} e^{\ell(f)} d\Pi(f)} \end{aligned}$$

since by Lemma 3.1

$$\begin{aligned} & \ell(f) - \ell(f_\tau) \\ &= -\frac{1}{2\varepsilon^2} (\|I_a(f - f_0)\|_{H_2}^2 - \|I_a f - I_a f_0 + \tau \varepsilon I_a \tilde{\psi}\|_{H_2}^2) + \tau \langle I_a \tilde{\psi}, \mathbb{W} \rangle_{H_2} \\ &= -\tau \langle I_a(I_a^* I_a)^{-1} \psi, \mathbb{W} \rangle_{H_2} + \frac{\tau^2}{2} \|I_a \tilde{\psi}\|_{H_2}^2 + \frac{\tau}{\varepsilon} \langle I_a(f - f_0), I_a \tilde{\psi} \rangle_{H_2} \end{aligned}$$

and since by Theorem 2.2

$$(3.10) \quad \langle I_a(f - f_0), I_a(I_a^* I_a)^{-1} \psi \rangle_{H_2} = \langle f - f_0, \psi \rangle_{H_1}.$$

By the Cameron–Martin theorem ([23], Theorem 2.6.13) the last ratio in (3.9) equals, for Π_τ the shifted law of f_τ , $f \sim \Pi$,

$$\frac{\int_{D_{\varepsilon,\tau}} e^{\ell(g)} \frac{d\Pi_\tau}{d\Pi}(g) d\Pi(g)}{\int_{D_\varepsilon} e^{\ell(g)} d\Pi(g)} = \frac{\int_{D_{\varepsilon,\tau}} e^{\ell(g)} e^{\tau \varepsilon \langle \tilde{\psi}, g \rangle_{V_\Pi} - (\tau \varepsilon)^2 \|\tilde{\psi}\|_{V_\Pi}^2 / 2} d\Pi(g)}{\int_{D_\varepsilon} e^{\ell(g)} d\Pi(g)},$$

where $D_{\varepsilon, \tau} = \{g = f_\tau : f \in D_\varepsilon\}$. In view of (3.6) we have $\varepsilon \|\tilde{\psi}\|_{V_\Pi} \rightarrow 0$ as $\varepsilon \rightarrow 0$, and by definition of D_ε and the Cauchy–Schwarz inequality we have convergence to zero of

$$\varepsilon \sup_{g \in D_{\varepsilon, \tau}} |\langle \tilde{\psi}, g \rangle_{V_\Pi}| = \varepsilon \sup_{f \in D_\varepsilon} |\langle \tilde{\psi}, f + \tau \varepsilon \tilde{\psi} \rangle_{V_\Pi}| \leq K \varepsilon \frac{\delta_\varepsilon}{\varepsilon} \|\tilde{\psi}\|_{V_\Pi} + |\tau| \varepsilon^2 \|\tilde{\psi}\|_{V_\Pi}^2$$

since $\delta_\varepsilon \rightarrow 0$. Conclude that the last ratio is, for every $\tau \in \mathbb{R}$,

$$(1 + o(1)) \frac{\int_{D_{\varepsilon, \tau}} e^{\ell(g)} d\Pi(g)}{\int_{D_\varepsilon} e^{\ell(g)} d\Pi(g)} = (1 + o(1)) \frac{\Pi(D_{\varepsilon, \tau}|Y)}{\Pi(D_\varepsilon|Y)}$$

as $\varepsilon \rightarrow 0$, and the proof is completed by showing that both the numerator and the denominator of the last ratio converge to one in probability: The denominator $\Pi(D_\varepsilon|Y)$ converges to one in $P_{f_0}^Y$ -probability by Lemma 3.2. The same is true for the numerator by applying Lemma 3.2 once more, since the Gaussian tail inequality guarantees for $\sqrt{6} < k < K$ and every $\tau \in \mathbb{R}$ that for ε small enough that

$$\begin{aligned} \Pi(D_{\varepsilon, \tau}^c) &= \Pi\left(v : \frac{|\langle \tilde{\psi}, v - \tau \varepsilon \tilde{\psi} \rangle_{V_\Pi}|}{\|\tilde{\psi}\|_{V_\Pi}} > K \delta_\varepsilon / \varepsilon\right) \\ &\leq \Pi\left(v : \frac{|\langle \tilde{\psi}, v \rangle_{V_\Pi}|}{\|\tilde{\psi}\|_{V_\Pi}} > K \frac{\delta_\varepsilon}{\varepsilon} - |\tau| \|\tilde{\psi}\|_{V_\Pi} \varepsilon\right) \leq e^{-k^2(\delta_\varepsilon/\varepsilon)^2/2}. \end{aligned}$$

□

Theorem 2.5 now follows from the fact that convergence in total variation distance implies convergence in any metric for weak convergence, so that in view of Lemma 3.2 it suffices to prove the theorem with $\Pi^{D_\varepsilon}(\cdot|Y)$ replacing $\Pi(\cdot|Y)$, and using the previous proposition plus the fact that pointwise convergence of Laplace transforms (in probability) implies weak convergence (in probability); see, for example, Proposition 28 in [41].

3.2. *Proof of Theorem 2.7.* Let $(\Omega, \mathcal{S}, \text{Pr})$ be the probability space supporting the random variable Y from (2.2) with law $P_{f_0}^Y$ (this space is implicitly constructed before (3.1) via the results from Section 7.4 in [41]). We show that $\varepsilon^{-1} E^\Pi[\langle f, \psi \rangle_{H_1} - \hat{\Psi}|Y]$ converges to 0 in $P_{f_0}^Y$ -probability which implies the result since then by definition of $\hat{\Psi}$ we then have

$$\varepsilon^{-1} \langle \bar{f}(Y) - f_0, \psi \rangle_{H_1} = -\langle I_a(I_a^* I_a)^{-1} \psi, \mathbb{W} \rangle_{H_2} + o_{P_{f_0}^Y}(1).$$

We argue by contradiction: Let ε_m be any sequence such that $\varepsilon_m \rightarrow 0$ but assume $\varepsilon_m^{-1} E^\Pi[\langle f, \psi \rangle_{H_1} - \hat{\Psi}|Y]$ does not converge to 0 in probability. Then there exists an event $\Omega' \in \mathcal{S}$ of positive probability $\text{Pr}(\Omega') > 0$ and $\xi > 0$ such that along a subsequence of m ,

$$(3.11) \quad |\varepsilon_m^{-1} E^\Pi[\langle f, \psi \rangle_{H_1} - \hat{\Psi}|Y(\omega)]| \geq \xi \quad \forall \omega \in \Omega'.$$

Since convergence in probability implies convergence almost surely along a subsequence, we can extract a further subsequence, still denoted by ε_m , for which we deduce from Theorem 2.5 that

$$\beta(\mathcal{L}(\varepsilon_m^{-1}(\langle f, \psi \rangle_{H_1} - \hat{\Psi})|Y), \mathcal{L}(Z)) \rightarrow 0$$

almost surely for Pr, and where Z is a $\mathcal{N}(0, \|I_a(I_a^* I_a)^{-1} \psi\|_{L^2_\mu(\partial_+ SM)}^2)$ random variable. Fix the event $\Omega_0 \subset \Omega$ of probability one where the last limit holds: then for every fixed $\omega \in \Omega_0$ we have the convergence in distribution

$$\Psi_m(\omega) \equiv \varepsilon_m^{-1}(\langle f, \psi \rangle_{H_1} - \hat{\Psi})|Y(\omega) \rightarrow^d Z.$$

By Skorohod’s theorem on almost surely convergent realisations of weakly convergent random variables (Theorem 11.7.2 in [18]) we can find, for every fixed $\omega \in \Omega_0$, a probability space on which we can define random variables $\tilde{\Psi}_m(\omega)$, \tilde{Z} such that $\mathcal{L}(\Psi_m(\omega)) = \mathcal{L}(\tilde{\Psi}_m(\omega))$, $\mathcal{L}(Z) = \mathcal{L}(\tilde{Z})$ and

$$\tilde{\Psi}_m(\omega) - \tilde{Z} \rightarrow 0$$

almost surely as $m \rightarrow \infty$. By standard conjugacy arguments the law of $h = I_a f|Y$ is a Gaussian measure on $L^2_\mu(\partial_+ SM)$. By Theorem 2.2, when integrating against $\psi \in C^\infty(M)$ we see

$$\langle f, \psi \rangle_{L^2(M)} = \langle h, I_a(I_a^* I_a)^{-1} \psi \rangle_{L^2_\mu(\partial_+ SM)}$$

which is a well defined normal distribution on the real line since the mapping $h \mapsto \langle h, I_a(I_a^* I_a)^{-1} \psi \rangle_{L^2_\mu(\partial_+ SM)}$ from $L^2_\mu(\partial_+ SM) \rightarrow \mathbb{R}$ is linear and continuous in view of $I_a(I_a^* I_a)^{-1} \psi \in L^2_\mu(\partial_+ SM)$, using Theorem 2.2 once more. Thus for every $\omega \in \Omega_0$ the variables $\tilde{\Psi}_m(\omega) - \tilde{Z}$, $m \in \mathbb{N}$, are all Gaussian and by the usual Paley–Zygmund argument (e.g., Exercise 2.1.4 in [23]) almost sure convergence implies convergence of all moments, in particular $E|\tilde{\Psi}_m(\omega) - \tilde{Z}| \rightarrow 0$ as $m \rightarrow \infty$. From this we deduce, for all $\omega \in \Omega_0$, $\Pr(\Omega_0) = 1$, that

$$\varepsilon_m^{-1} E^\Pi[\langle f, \psi \rangle_{H_1} - \hat{\Psi}|Y(\omega)] = E\Psi_m(\omega) = E[\tilde{\Psi}_m(\omega)] \rightarrow E\tilde{Z} = EZ = 0$$

as $m \rightarrow \infty$, a contradiction to (3.11) with $\Pr(\Omega') > 0$, completing the proof.

4. Proofs for Section 2.1. In this section we prove Theorem 2.2 and we will do so by putting the theory into the framework of the transmission condition as developed in [26, 29]. We will give full details for the case of the geodesic X -ray transform I and indicate the (minor) modifications necessary for the proof to work also for the attenuated X -ray transform I_a at the end. We note that previously known results only give that $I^* I$ is injective on $L^2(M)$ and surjectivity properties were only obtained after enlarging M (as in [46]). These results are not sufficient to obtain the theorems in Section 2.3, nor do they expose the precise boundary behaviour as we do here.

4.1. *Setting up the scene and main ideas.* We shall denote by N the normal, or “information” operator $I^*I : L^2(M) \rightarrow L^2(M)$ introduced in Section 2.1. An integral formula for N can be derived directly from the expressions for I and I^* :

$$(4.1) \quad Nf(x) = 2 \int_{S_x M} dv \int_0^{\tau(x,v)} f(\gamma_{x,v}(t)) dt.$$

A property of fundamental importance is that whenever (M, g) has no conjugate points, then, in the interior of M , the operator N is an elliptic pseudo-differential operator (Ψ DO) of order -1 with principal symbol $c_d|\xi|^{-1}$; cf. [27], Section 6.3, [53] or Lemma 3.1 in [46]. (The reference [27] states this property under the so called *Bolker condition*, which is seen to be equivalent in our case to the absence of conjugate points.) We refer to [57] for a treatment of Ψ DOs. In particular recall that for P a classical Ψ DO of order $m \in \mathbb{C}$, a full symbol in local coordinates is denoted by $p(x, \xi) \sim \sum_{j=0}^\infty p_j(x, \xi)$ where $p_j(x, t\xi) = t^{m-j} p_j(x, \xi)$, and where p_0 is the *principal symbol*. The operator P is *elliptic* if $p_0(x, \xi) \neq 0$ for all (x, ξ) in the cotangent bundle, $\xi \neq 0$.

Recall that (M, g) is called *simple* if it is nontrapping, has strictly convex boundary and no conjugate points. Simple manifolds are simply connected; in fact they are diffeomorphic to balls in Euclidean space. From now on we shall assume that (M, g) is simple. It will be convenient for what follows to consider (M, g) isometrically embedded into a closed manifold (S, g) . Since M is simple, there is an open neighborhood U_1 of M in S , such that its closure $M_1 := \overline{U_1}$ is a compact simple manifold; see Figure 6. Let I_1 denote the geodesic ray transform associated to (M_1, g) and let $N_1 = I_1^* I_1$.

Following [46] we may cover (S, g) with finitely many simple open sets U_k with $M \subset U_1$, $M \cap \overline{U_j} = \emptyset$ for $j \geq 2$, and consider a partition of unity $\{\varphi_k\}$ subordinate to $\{U_k\}$ so that $\varphi_k \geq 0$, $\text{supp } \varphi_k \subset U_k$ and $\sum \varphi_k^2 = 1$. We pick φ_1 such that $\varphi_1 \equiv 1$ on

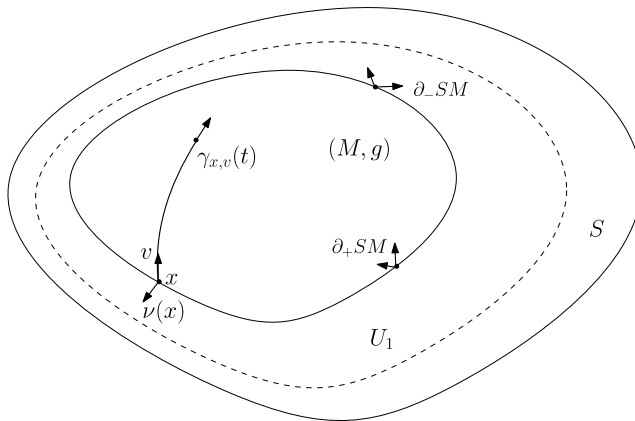


FIG. 6. *Setting.*

a neighborhood of M compactly supported in U_1 . Hence, for I_k the ray transform associated to (U_k, g) , we can define

$$(4.2) \quad Pf := \sum_k \varphi_k(I_k^* I_k)(\varphi_k f), \quad f \in C^\infty(S).$$

Each operator $I_k^* I_k : C_c^\infty(U_k) \rightarrow C^\infty(U_k)$ is an elliptic Ψ DO of order -1 and principal symbol $c_d|\xi|^{-1}$, and hence so is P . Having P defined on a closed manifold is convenient, since one can use standard mapping properties for Ψ DOs. For instance for P defined by (4.2) we have

$$P : H^s(S) \rightarrow H^{s+1}(S) \quad \text{for all } s \in \mathbb{R},$$

where $H^s(S)$ denotes the standard L^2 Sobolev space of the closed manifold S (when s is a nonnegative integer, $H^s(S)$ can be identified with the set of $u \in L^2(S)$ such that $Du \in L^2(S)$ for all differential operators D of order $\leq s$ with coefficients in $C^\infty(S)$; see [55] for the definition for arbitrary $s \in \mathbb{R}$).

Let $r_M : L^2(S) \rightarrow L^2(M)$ denote restriction to M^{int} , the interior of M , and $e_M : L^2(M) \rightarrow L^2(S)$ extension by zero. (We could consider restriction to M as well, but this makes no difference since the boundary of M has measure zero.) Both operators are bounded and dual to each other. Since $\varphi_1 = 1$ near M , given $f \in C_c^\infty(M^{\text{int}})$ (smooth functions with compact support contained in M^{int}) we have

$$r_M P e_M f = r_M N_1 \varphi_1 e_M f = r_M N_1 e_M f.$$

Equation (4.1) shows that $r_M N_1 e_M f = Nf$ and thus by density of $C_c^\infty(M^{\text{int}})$ in $L^2(M)$, we have that P and N are related by the following truncation process:

$$(4.3) \quad N = r_M P e_M \quad \text{in } L^2(M).$$

Since $P : L^2(S) \rightarrow H^1(S)$, this gives immediately the mapping property $N : L^2(M) \rightarrow H^1(M)$ when the spaces $H^s(M)$ are defined by restriction (4.7). Since the embedding $H^1(M) \hookrightarrow L^2(M)$ is compact, obviously $N : L^2(M) \rightarrow L^2(M)$ is compact (and hence I).

However, without further analysis not much more can be said about the mapping properties of N , especially if we are interested in functions supported all the way to the boundary of M . From (4.3), we see that e_M could produce singularities for higher order Sobolev spaces, preventing good mapping properties in ‘smooth topologies’. A key input of Boutet de Monvel [2, 3] (see also [13]) was to show that a necessary and sufficient condition for P , a Ψ DO of order m defined on S , to satisfy $r_M P e_M(C^\infty(M)) \subset C^\infty(M)$, is that P satisfies the *transmission condition* with respect to ∂M in the sense that

$$(4.4) \quad \partial_x^\beta \partial_\xi^\alpha p_j(x, \nu(x)) = e^{\pi i(m-j-|\alpha|)} \partial_x^\beta \partial_\xi^\alpha p_j(x, -\nu(x)),$$

for all j, α, β and $x \in \partial M$.

Unfortunately such a condition does not hold in the case of N_1 (or P) defined above, as the following example shows: let M be the unit disk in \mathbb{R}^2 . An elementary calculation gives that $N(1) = 4E(r)/\pi$ where E is the complete elliptic integral of the second kind and r is the radial coordinate. As r approaches 1, $E'(r)$ blows up and hence $N(1) \notin C^\infty(M)$, therefore $N = r_M N_1$ cannot satisfy the transmission condition.

Furthermore for purposes of inversion, even if $h \in C^\infty(M)$, we cannot expect the solutions f to $N(f) = h$ to be in $C^\infty(M)$ either: in the previous example, it is not hard to check that $I((1 - r^2)^{-1/2})$ (r denoting distance to the origin) is a constant function [35], Corollary 3.3, and therefore so is

$$(4.5) \quad N((1 - r^2)^{-1/2}) = c.$$

While N_1 does not satisfies condition (4.4), we show that it satisfies a modified transmission condition as introduced by Hörmander in [29] and recently expanded and enhanced by Grubb in [26]. Namely, given $\mu \in \mathbb{C}$ with real part $\Re \mu > -1$, we say that P , a ΨDO of order m defined on S , satisfies a *transmission condition of type μ* with respect to ∂M if its symbol satisfies

$$(4.6) \quad \partial_x^\beta \partial_\xi^\alpha p_j(x, \nu(x)) = e^{\pi i(m-2\mu-j-|\alpha|)} \partial_x^\beta \partial_\xi^\alpha p_j(x, -\nu(x)),$$

for all j, α, β and $x \in \partial M$, generalizing the case $\mu = 0$ given by (4.4). To tie this condition with mapping properties, using μ as above, let us define

$$\mathcal{E}_\mu(M) := \{e_M d_M(x)^\mu \varphi, \varphi \in C^\infty(M)\},$$

where $d_M(x)$ is a $C^\infty(M)$ -function equal to $\text{dist}(x, \partial M)$ near ∂M and positive on the interior of M . Then, as a generalization to Boutet de Monvel’s result above, the following theorem appears in [30], Theorem 18.2.18:

THEOREM 4.1. *A necessary and sufficient condition in order that $r_M P u \in C^\infty(M)$ for all $u \in \mathcal{E}_\mu(M)$ is that P satisfies the μ -transmission condition (4.6).*

To make use of the theorem above, we first prove in Section 4.3 that

LEMMA 4.2. *The operator P defined in (4.2) satisfies the transmission condition of type $\mu = -1/2$ with respect to ∂M .*

In particular, Lemma 4.2 and Theorem 4.1 imply that

$$N : d_M^{-1/2} C^\infty(M) \rightarrow C^\infty(M)$$

is well defined. Notice that the domain allows for functions which blow up near the boundary like $\text{dist}(x, \partial M)^{-\frac{1}{2}}$, explaining (4.5).

After constructing in Section 4.2 appropriate Hilbert-scale versions of $\mathcal{E}_{-1/2}(M)$, namely, the Hörmander spaces $H^{-1/2(s)}(M)$, the first basic result [26], Theorem 4.2, applied to $r_M P$ gives further mapping properties:

THEOREM 4.3. $r_M P$ maps $H^{-1/2(s)}(M)$ continuously into $H^{s+1}(M)$, where P is defined in (4.2) and $s > -1$.

While all the results above only discuss forward mapping properties, using ellipticity will show that such an operator is in fact Fredholm in the functional settings mentioned above. Then proving that its kernel and co-kernel are trivial will ensure that it will be invertible in these settings as well. In particular, the main result we prove below provides a full solution to the homogeneous Dirichlet problem for P on the domain M .

THEOREM 4.4. Let P be the elliptic Ψ DO of order -1 given by (4.2). For $s > -1$ the map $r_M P : H^{-1/2(s)}(M) \rightarrow H^{s+1}(M)$ is a homeomorphism. Moreover, $N : d_M^{-1/2} C^\infty(M) \rightarrow C^\infty(M)$ is a bijection.

The outline of the remainder is as follows. Section 4.2 contains details on Sobolev and Hörmander spaces. Section 4.3 will be devoted to the proof of Lemma 4.2, and Section 4.4 to the proof of Theorem 4.4, requiring a few technical lemmas, followed by the proof of Theorem 2.2.

4.2. Sobolev spaces and Hörmander spaces. In this section we summarize the main functional setting that we will be using. Here we shall be concerned only with L^2 -Sobolev and Hörmander spaces. The Sobolev spaces are standard but the Hörmander spaces are less so. For the latter we will follow [26] and for the former [36, 55] (with minor departures in notation). As before we let (M, g) be a compact Riemannian manifold with boundary which we think isometrically embedded into a closed manifold (S, g) . We write $H^s(S)$ for the standard L^2 based Sobolev space of the closed manifold S . We denote

$$H^s(M) = r_M H^s(S) = \{u|_{M^{\text{int}}} : u \in H^s(S)\}$$

equipped with the quotient norm

$$(4.7) \quad \|u\|_{H^s(M)} := \inf\{\|w\|_{H^s(S)} : w \in H^s(S), r_M w = u\}.$$

We denote

$$H_M^s(S) := \{u \in H^s(S) : \text{supp}(u) \subset M\}.$$

REMARK 4.5. The space $H_M^s(M)$ can also be seen as the closure of $C_c^\infty(M^{\text{int}})$ in $H^s(S)$. Finally we can also define $H_0^s(M)$ as the closure of $C_c^\infty(M^{\text{int}})$ in $H^s(M)$. When $s \notin \mathbb{Z} + \frac{1}{2}$, there is a natural identification with $H_M^s(S)$. When s is a nonnegative integer, $H^s(M)$ can be identified with the set of $u \in L^2(M)$ such that $Du \in L^2(M)$ for all differential operators D of order $\leq s$ with coefficients in $C^\infty(M)$.

One of the main inputs of [26] is the introduction of particularly efficient *order reducing operators*; cf. [26], Theorem 1.3. These are classical elliptic Ψ DOs on S of order μ (denoted $\Lambda_+^{(\mu)}$) preserving support in M and defining homeomorphisms

$$(4.8) \quad \Lambda_+^{(\mu)} : H_M^s(S) \rightarrow H_M^{s-\Re\mu}(S),$$

where $\Re\mu$ denotes the real part of μ . These operators are used to define the Hörmander spaces (also known as μ -transmission spaces)

$$H^{\mu(s)}(M) := \Lambda_+^{(-\mu)} e_M H^{s-\Re\mu}(M), \quad s > \Re\mu - 1/2.$$

For $s > \Re\mu - 1/2$, the maps $r_M \Lambda_+^{(\mu)} : H^{\mu(s)}(M) \rightarrow H^{s-\Re\mu}(M)$ are homeomorphisms with inverse $\Lambda_+^{(-\mu)} e_M$ [26], Proposition 1.7. We have a natural embedding $H_M^s(S) \subset H^{\mu(s)}(M)$.

As explained in Section 4.1, these spaces are specifically adapted to the μ -transmission condition (4.6) and will provide natural spaces of solutions to the equation $Nf = h$ where $\mu = -1/2$.

The order reducing operators $\Lambda_+^{(\mu)}$ are used in conjunction with their adjoints $\Lambda_-^{(\bar{\mu})}$ by considering a new operator $Q = \Lambda_-^{(\mu-m)} P \Lambda_+^{(-\mu)}$, where m is the order of P . The point is that if P satisfies the μ -transmission condition (4.6), then Q satisfies the transmission condition (4.4) with $\mu = 0$ and fits the Boutet de Monvel calculus. This is the main idea in [26].

Let $d_M(x)$ be a $C^\infty(M)$ -function equal to $\text{dist}(x, \partial M)$ near ∂M and positive on the interior of M . For $\mu \in \mathbb{C}$ with $\Re\mu > -1$, let $\mathcal{E}_\mu(M)$ denote the space of functions u such that $u = e_M d_M(x)^\mu \varphi$ with $\varphi \in C^\infty(M)$. One can show that ([26], Proposition 4.1)

$$\mathcal{E}_\mu(M) = \bigcap_s H^{\mu(s)}(M).$$

The spaces $H^{\mu(s)}(M)$ were introduced in [29] as the completion of $\mathcal{E}_\mu(M)$ in the topology defined by the seminorms $u \mapsto \|r_M P u\|_{H^{s-\Re m}(M)}$, where P runs through the operators satisfying the μ -transmission condition (see below) and any order $m \in \mathbb{C}$. Hörmander’s starting point was the work of Vishik and Eskin [20, 61].

4.3. Proof of Lemma 4.2.

DEFINITION 4.6. We shall say that P has *even symbol* if $p_j(x, -\xi) = (-1)^j p_j(x, \xi)$ for all $j \geq 0$. It is easy to check that this condition is independent of the coordinates chosen. (Recall that the full symbol is not defined intrinsically.)

LEMMA 4.7. *The symbol of N (or N_1) is even.*

PROOF. There are (at least) three possible proofs of this lemma. As explained in [50] the full *geometric* symbol of N coincides with its principal symbol $c_d|\xi|^{-1}$. In [52] a relation is established between the full geometric symbol and the ordinary full symbol in local coordinates. See, for example, equations (1.6) and (1.7) in [52]. An inspection of those formulas shows that the symbol is even starting from the fact that $c_d|\xi|^{-1}$ is even. Another more natural proof was suggested to us by Gunther Uhlmann and is based on the calculation of the full symbol in [53]. Equation (17) in [53] gives an explicit formula for the amplitude $M(x, y, \xi)$ of the Ψ DO for the case of 2-tensors. In the case of functions the formula is

$$M(x, y, \xi) = \int e^{-i\xi \cdot z} (G^{(1)} z \cdot z)^{\frac{-n+1}{2}} \frac{|\det G^{(3)}|}{\sqrt{\det g}} dz,$$

where $G^{(1)}(x, y)$ and $G^{(3)}(x, y)$ are defined in [53], Lemma 3, but we do not need to know what they are. The terms $p_j(x, \xi)$ may be derived from the amplitude by

$$p_j(x, \xi) = \sum_{|\alpha|=j} \frac{1}{\alpha!} \partial_\xi^\alpha D_y^\alpha M(x, y, \xi)|_{y=x}.$$

Since $M(x, y, \xi)$ is even in ξ , we see that once we start taking derivatives in ξ , the parity of p_j in ξ changes according to $(-1)^j = (-1)^{|\alpha|}$.

The quickest way is perhaps to use [15], Lemma B.1, which covers a broad range of operators for the form

$$Af(x) = \int_{S_x U_1} \int_{\mathbb{R}} A(x, r, w) f(x + rw) dr dS_x(w).$$

Our operator N is certainly of this form (after some change of variables). The lemma proves that A is a classical Ψ DO of order -1 and computes explicitly the full symbol deriving a formula

$$p_k(x, \xi) = 2\pi \frac{i^k}{k!} \int_{S_x U_1} \partial_r^k A(x, 0, w) \delta^{(k)}(w \cdot \xi) dS_x(w).$$

From this formula we see right away that $p_k(x, -\xi) = (-1)^k p(x, \xi)$ since the delta function δ is even. \square

PROOF OF LEMMA 4.2. In a tubular neighbourhood of ∂M , the full symbol of P coincides with that of N_1 . The result is then a direct consequence of Lemma 4.7 and the fact that $m = -1$. \square

4.4. *Proofs of Theorems 4.4 and 2.2.* To prove Theorem 4.4, a first step is to prove that for P defined in (4.2), $r_M P$ is a Fredholm operator in the functional settings $H^{-1/2(s)}(M) \rightarrow H^{s+1}(M)$ for $s > -1$, and $\mathcal{E}_{-1/2}(M) \rightarrow C^\infty(M)$. This is mainly due to the ellipticity of P , and one additional concept from [26], the *factorization index* μ_0 . This is defined for elliptic operators of order m as

$$\mu_0 := m/2 + (a_+ - a_-)/2\pi i,$$

where $\exp a_{\pm} = p_0(x, \mp \nu(x))$. For the case of N_1 , we have $\mu_0 = -1/2$ since the principal symbol of N_1 is (up to a constant) $|\xi|^{-1}$; cf. [26], Example 3.2. In general, for elliptic operators of order m with even symbol, $\mu = \mu_0 = m/2$. Then as a direct consequence of [26], Theorem 4.4, we obtain the following theorem.

THEOREM 4.8. *Assume $s > -1$. Suppose $u \in H_M^\sigma(S)$ for some $\sigma > -1$ and let P be the elliptic Ψ DO of order -1 given by (4.2). If $r_M P u \in H^{s+1}(M)$, then $u \in H^{-1/2(s)}(M)$. Moreover, the mapping $r_M P : H^{-1/2(s)}(M) \rightarrow H^{s+1}(M)$ is Fredholm. In particular, if $r_M P u \in C^\infty(M)$, then $u \in \mathcal{E}_{-1/2}(M)$. The mapping $r_M P : \mathcal{E}_{-1/2}(M) \rightarrow C^\infty(M)$ is also Fredholm.*

Before proving Theorem 4.4, we state and prove a few preliminary lemmas. We begin with:

LEMMA 4.9. *If (M, g) is a nontrapping manifold with strictly convex boundary, there exists a constant $C_0 > 0$ such that*

$$d_M(\gamma_{x,v}(t), \partial M) \geq C_0 t(\tau(x, v) - t) \quad \text{for all } (x, v) \in \partial_+ SM, t \in [0, \tau(x, v)].$$

PROOF. In what follows, denote $\alpha_A : \partial_+ SM \rightarrow \partial_+ SM$ the antipodal scattering relation, that is, the map $(x, v) \mapsto (\gamma_{x,v}(\tau(x, v)), -\dot{\gamma}_{x,v}(\tau(x, v)))$. α_A satisfies $\alpha_A^2 = Id$ and if $g(\nu_x, v) = 0$, then $\alpha_A(x, v) = (x, -v)$.

It is enough to show that the nonnegative function $F(x, v, t) := \frac{d_M(\gamma_{x,v}(t), \partial M)}{t(\tau(x, v) - t)}$ is uniformly bounded away from zero on the set

$$G = \{(x, v) \in \partial_+ SM, t \in (0, \tau(x, v))\}.$$

In what follows, we will also use [51], Lemma 4.1.2, page 113, stating that there exists $C_2 > 0$ such that

$$(4.9) \quad \tau(x, v) \leq C_2 |\langle \nu(x), v \rangle_g|, \quad (x, v) \in \partial_+ SM, \quad \langle \nu(x), v \rangle_g \neq 0,$$

where $\nu(x)$ is the outer unit normal at $x \in \partial M$. This is essentially a consequence of the strict convexity of the boundary of M . Using that $d_M(\gamma_{x,v}(t), \partial M)|_{t=0} = 0$ and $\frac{d}{dt} d_M(\gamma_{x,v}(t), \partial M)|_{t=0} = |\langle \nu(x), v \rangle_g|$ (e.g., by using normal geodesic coordinates), l'Hôpital's rule implies

$$\lim_{t \rightarrow 0^+} F(x, v, t) = \frac{|\langle \nu(x), v \rangle_g|}{\tau(x, v)} = \frac{\mu(x, v)}{\tau(x, v)} \geq \frac{1}{C_2}.$$

Moreover, since $\gamma_{x,v}(t) = \gamma_{\alpha_A(x,v)}(\tau(x, v) - t)$, we have the symmetry property $F(x, v, t) = F(\alpha_A(x, v), \tau(x, v) - t)$, and this allows to deduce the limit

$$\lim_{t \rightarrow \tau(x,v)^-} F(x, v, t) = \frac{\mu(\alpha_A(x, v))}{\tau(\alpha_A(x, v))} \geq \frac{1}{C_2}.$$

By compactness, the result follows since F is uniformly bounded away from zero outside any neighborhood of $\{t = 0\} \cup \{t = \tau(x, v)\}$ in G . \square

An important ingredient in what follows is the consideration of the following weighted space $L^2(M, d_M^{1/2})$ where the measure is $d_M^{1/2} dx$. Recall that I^* denotes the usual backprojection, that is, the adjoint of $I : L^2(M) \rightarrow L^2_\mu(\partial_+ SM)$.

LEMMA 4.10. *The following hold:*

- (i) *The map $I : L^2(M, d_M^{1/2}) \rightarrow L^2_\mu(\partial_+ SM)$ is bounded with adjoint $I_w^* = d_M^{-1/2} I^*$.*
- (ii) *The map $I^* : L^2_\mu(\partial_+ SM) \rightarrow L^2(M, d_M^{-1/2})$ is bounded.*

PROOF. Let $f \in L^2(M, d_M^{1/2})$ and write $f = d_M^{-1/4} g$ for some $g \in L^2(M)$. We write

$$If(x, v) = \int_0^{\tau(x, v)} g(\gamma_{x, v}(t)) \frac{dt}{d_M(\gamma_{x, v}(t), \partial M)^{1/4}},$$

$$|If(x, v)|^2 \leq \int_0^{\tau(x, v)} |g(\gamma_{x, v}(t))|^2 dt \int_0^{\tau(x, v)} \frac{dt}{d_M(\gamma_{x, v}(t), \partial M)^{1/2}},$$

using the Cauchy–Schwarz inequality. Using Lemma 4.9, we have, for any $(x, v) \in \partial_+ SM$,

$$\begin{aligned} & \int_0^{\tau(x, v)} \frac{dt}{d_M(\gamma_{x, v}(t), \partial M)^{1/2}} \\ & \leq \frac{1}{C_0^{1/2}} \int_0^{\tau(x, v)} \frac{dt}{(t(\tau - t))^{\frac{1}{2}}} \stackrel{t=\tau u}{=} \frac{1}{C_0^{1/2}} \int_0^1 \frac{du}{(u(1 - u))^{\frac{1}{2}}} \\ & = \frac{\pi}{C_0^{1/2}}. \end{aligned}$$

Integrating over $\partial_+ SM$, we then obtain

$$\begin{aligned} \int_{\partial_+ SM} |If(x, v)|^2 d\mu & \leq \frac{\pi}{C_0^{1/2}} \int_{\partial_+ SM} \int_0^{\tau(x, v)} |g(\gamma_{x, v}(t))|^2 dt d\mu \\ & = \frac{\pi}{C_0^{1/2}} \int_{SM} |g(x)|^2 dx dv \quad (\text{by Santaló's formula [51]}) \\ & = \frac{\pi \text{Vol}(S^{d-1})}{C_0^{1/2}} \int_M |g(x)|^2 dx \\ & = \frac{\pi \text{Vol}(S^{d-1})}{C_0^{1/2}} \|f\|_{L^2(M, d_M^{1/2})}^2, \end{aligned}$$

hence (i) holds. Then (ii) is a direct consequence of the factorization

$$I^* : L^2_\mu(\partial_+ SM) \xrightarrow{I_w^*} L^2(M, d_M^{1/2}) \xrightarrow{d_M^{1/2}} L^2(M, d_M^{-1/2}),$$

where I_w^* is continuous by (i) and the second operator (multiplication by $d_M^{1/2}$) is an isometry in the setting above. The proof of Lemma 4.10 is complete. \square

LEMMA 4.11. *Given $\varphi \in C^\infty(M)$, $d_M^{-1/2}\varphi \in L^2(M, d_M^{1/2}) \cap L^1(M)$.*

PROOF. Obviously φ is bounded in M . Since $(d_M^{-1/2}\varphi)^2 d_M^{1/2} = d_M^{-1/2}\varphi^2$, we just need to prove that $d_M^{-1/2}$ is in $L^1(M)$. By taking local geodesic normal coordinates where x_n denotes distance to the boundary, the lemma follows from the elementary observation

$$\int_0^\epsilon x_n^{-1/2} dx_n < \infty$$

since locally $d_M = x_n$. \square

REMARK 4.12. The same proof shows that $d_M^{-1/2}H^s(M) \subset L^2(M, d_M^{1/2})$ as long as $s > \dim M/2$. The latter condition ensures that elements in $H^s(M)$ are continuous and hence bounded.

We are now ready to give the proof of Theorem 4.4.

PROOF OF THEOREM 4.4. By Theorem 4.8, the map $r_M P$ is Fredholm with finite dimensional kernel and co-kernel independent of s ; in fact elements in the kernel must be in $\mathcal{E}_{-1/2}(M)$; cf. [25], Theorem 3.5. Hence, it suffices to check that these kernel and co-kernel are trivial. We begin by proving that the kernel is trivial.

Suppose there is $u \in \mathcal{E}_{-1/2}(M)$ such that $r_M P u = 0$. Writing $u = e_M d_M^{-1/2}\varphi$ with $\varphi \in C^\infty(M)$, we see that $Nf = 0$ where $f = d_M^{-1/2}\varphi \in L^2(M, d_M^{1/2})$ by Lemma 4.11. But $I_w^* I = d_M^{-1/2} N$, hence $I_w^* I f = 0$. This implies $(I_w^* I f, f)_{L^2(M, d_M^{1/2})} = 0$ and hence $I f = 0$. To show that f must in fact be smooth, extend f by zero to U_1 and call the extension f_1 . By Lemma 4.11, $f_1 \in L^1(U_1)$ so that, using Santaló’s formula, it is easy to see that $I_1 f_1$ makes sense in $L^1(\partial_+ S U_1)$ and also that $N_1 f_1 = I_1^* I_1 f$ makes sense in $L^1(U_1)$. Then $I_1 f_1 = 0$ and thus $N_1 f_1 = 0$. Since N_1 is elliptic, the function f_1 must be smooth in U_1 and hence f is smooth in M . Now we use the standard injectivity result for I acting on smooth functions on a simple manifold [38] to conclude that $f = 0$.

Let us now check that the co-kernel of $r_M P$ is trivial. Consider the injection

$$\iota : H_M^s(S) \hookrightarrow H^{-1/2(s)}(M),$$

where $H_M^s(S)$ consists of elements in $H^s(S)$ with support in M (cf. Section 4.2). Let us compute $(r_M P \iota)^*$. The point of using ι is to end up with standard dualities not involving the Hörmander spaces. Note

$$(r_M P \iota)^* : (H^{s+1}(M))^* \rightarrow (H_M^s(S))^*$$

where

$$(H^{s+1}(M))^* = H_M^{-s-1}(S),$$

$$(H_M^s(S))^* = H^{-s}(M)$$

are the standard dualities. Take $u \in H_M^{-s-1}(S)$ and $f \in H_M^s(S)$ and observe

$$(4.10) \quad (r_M P f, u)_M = (P f, u)_S = (f, P_{-s-1} u)_S = (f, r_M P u)_M.$$

Thus $\iota^*(r_M P)^* = (r_M P \iota)^* = r_M P$. Hence if $u \in H_M^{-s-1}(M)$ is such that $(r_M P)^* u = 0$ we see that $r_M P u = 0$. By ellipticity $u \in \mathcal{E}_{-1/2}(M)$ and since we have already proved injectivity of $r_M P$ on this space we deduce that the co-kernel of $r_M P$ is trivial as well. Thus $r_M P : \mathcal{E}_{-1/2}(M) \rightarrow C^\infty(M)$ is a bijection and since $N = r_M P e_M$, from the definition of $\mathcal{E}_{-1/2}(M)$ we conclude that $N : d_M^{-1/2} C^\infty(M) \rightarrow C^\infty(M)$ is also a bijection. \square

PROOF OF THEOREM 2.2. For the case of I , part (a) in Theorem 2.2 follows immediately from Theorem 4.4. Part (b) in Theorem 2.2 is a direct consequence of Lemma 4.10 and the fact that $d_M^{-1/2} h \in L^2(M, d_M^{1/2})$ if $h \in C(M)$.

To complete the proof of Theorem 2.2 we just need to explain why the same proof works for the attenuated geodesic X -ray transform I_a . The microlocal properties of $N_a = I_a^* I_a$ are studied in detail in [21]. With this in hand, it is straightforward to check that N_a will fit the theory developed above. For this we need to extend $a \in C^\infty(M)$ smoothly to S and observe that the third proof of Lemma 4.7 applies to N_a since these operators are covered by [15], Lemma B.1; see [21], Section 4, for a proof. Hence the full symbol of N_a is also even, $N_{a,1}$ satisfies the transmission condition with $\mu = -1/2$ with respect to ∂M , and we can derive all the required mapping properties. The proof of Lemma 4.10 works for I_a and Theorem 4.4 holds as well for I_a as long as we know that I_a is injective on smooth functions. \square

Acknowledgments. The authors are grateful to Gerd Grubb, Plamen Stefanov and Gunther Uhlmann for helpful conversations and comments regarding the microlocal aspects of this paper. The authors also thank Sarah Vallélian for sharing insights and code pertaining to [4], and Matteo Giordano and Hanne Kekkonen for proofreading parts of the manuscript.

REFERENCES

- [1] ABRAMOWITZ, M. and STEGUN, I. A., eds. (1966). *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York. [MR0208797](#)
- [2] BOUTET DE MONVEL, L. (1966). Comportement d'un opérateur pseudo-différentiel sur une variété à bord. II. Pseudo-noyaux de Poisson. *J. Anal. Math.* **17** 255–304. [MR0239255](#)
- [3] BOUTET DE MONVEL, L. (1971). Boundary problems for pseudo-differential operators. *Acta Math.* **126** 11–51. [MR0407904](#)
- [4] BROWN, D. A., SAIBABA, A. and VALLÉLIAN, S. (2017). Low rank independence samplers in bayesian inverse problems. ArXiv preprint. Available at [arXiv:1609.07180v2](#).
- [5] BUDINGER, T. F., GULLBERG, G. T. and HUESMAN, R. H. (1979). Emission computed tomography. In *Image Reconstruction from Projections* (G. T. Herman, ed.). *Topics in Applied Physics* **32** 147–246. Springer, Berlin.
- [6] CASTILLO, I. (2014). On Bayesian supremum norm contraction rates. *Ann. Statist.* **42** 2058–2091. [MR3262477](#)
- [7] CASTILLO, I. (2017). Pólya tree posterior distributions on densities. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 2074–2102. [MR3729648](#)
- [8] CASTILLO, I. and NICKL, R. (2013). Nonparametric Bernstein-von Mises theorems in Gaussian white noise. *Ann. Statist.* **41** 1999–2028. [MR3127856](#)
- [9] CASTILLO, I. and NICKL, R. (2014). On the Bernstein-von Mises phenomenon for nonparametric Bayes procedures. *Ann. Statist.* **42** 1941–1969. [MR3262473](#)
- [10] CASTILLO, I. and ROUSSEAU, J. (2015). A Bernstein–von Mises theorem for smooth functionals in semiparametric models. *Ann. Statist.* **43** 2353–2383. [MR3405597](#)
- [11] CAVALIER, L. (2008). Nonparametric statistical inverse problems. *Inverse Probl.* **24** Article ID 034004. [MR2421941](#)
- [12] CAVALIER, L., GOLUBEV, G. K., PICARD, D. and TSYBAKOV, A. B. (2002). Oracle inequalities for inverse problems. *Ann. Statist.* **30** 843–874. [MR1922543](#)
- [13] CHAZARAIN, J. and PIRIOU, A. (1982). *Introduction to the Theory of Linear Partial Differential Equations. Studies in Mathematics and Its Applications* **14**. North-Holland, Amsterdam. Translated from the French. [MR0678605](#)
- [14] CREAGER, K. C. (1992). Anisotropy of the inner core from differential travel times of the phases PKP and PKIPK. *Nature* **356** 309–314.
- [15] DAIRBEKOV, N. S., PATERNAIN, G. P., STEFANOV, P. and UHLMANN, G. (2007). The boundary rigidity problem in the presence of a magnetic field. *Adv. Math.* **216** 535–609. [MR2351370](#)
- [16] DASHTI, M., LAW, K. J. H., STUART, A. M. and VOSS, J. (2013). MAP estimators and their consistency in Bayesian nonparametric inverse problems. *Inverse Probl.* **29** Article ID 095017. [MR3104933](#)
- [17] DASHTI, M. and STUART, A. (2016). The Bayesian approach to inverse problems. In *Handbook of Uncertainty Quantification* (R. Ghanem, D. Higdon and H. Owhadi, eds.) Springer, Berlin.
- [18] DUDLEY, R. M. (2002). *Real Analysis and Probability. Cambridge Studies in Advanced Mathematics* **74**. Cambridge Univ. Press, Cambridge. Revised reprint of the 1989 original. [MR1932358](#)
- [19] DUVAL, T. L. JR, JEFFERIES, S. M., HARVEY, J. W. and POMERANTZ, M. A. (1993). Time–distance helioseismology. *Nature* **362** 430–432.
- [20] ESKIN, G. I. (1981). *Boundary Value Problems for Elliptic Pseudodifferential Equations. Translations of Mathematical Monographs* **52**. Amer. Math. Soc., Providence, RI. Translated from the Russian by S. Smith. [MR0623608](#)

- [21] FRIGYIK, B., STEFANOV, P. and UHLMANN, G. (2008). The X-ray transform for a generic family of curves and weights. *J. Geom. Anal.* **18** 89–108. [MR2365669](#)
- [22] GHOSAL, S. and VAN DER VAART, A. (2017). *Fundamentals of Nonparametric Bayesian Inference. Cambridge Series in Statistical and Probabilistic Mathematics* **44**. Cambridge Univ. Press, Cambridge. [MR3587782](#)
- [23] GINÉ, E. and NICKL, R. (2016). *Mathematical Foundations of Infinite-Dimensional Statistical Models*. Cambridge Univ. Press, New York. [MR3588285](#)
- [24] GOLDENSHLUGER, A. and PEREVERZEV, S. V. (2000). Adaptive estimation of linear functionals in Hilbert scales from indirect white noise observations. *Probab. Theory Related Fields* **118** 169–186. [MR1790080](#)
- [25] GRUBB, G. (2014). Local and nonlocal boundary conditions for μ -transmission and fractional elliptic pseudodifferential operators. *Anal. PDE* **7** 1649–1682. [MR3293447](#)
- [26] GRUBB, G. (2015). Fractional Laplacians on domains, a development of Hörmander’s theory of μ -transmission pseudodifferential operators. *Adv. Math.* **268** 478–528. [MR3276603](#)
- [27] GUILLEMIN, V. and STERNBERG, S. (1977). *Geometric Asymptotics. Mathematical Surveys* **14**. Amer. Math. Soc., Providence, RI. [MR0516965](#)
- [28] HELGASON, S. (1999). *The Radon Transform*, 2nd ed. *Progress in Mathematics* **5**. Birkhäuser, Boston, MA. [MR1723736](#)
- [29] HÖRMANDER, L. (1965). *Ch II, Boundary Problems for “Classical” Pseudo-Differential Operators. Lectures Notes at IAS*, Princeton. Available at <http://www.math.ku.dk/~grubb/LH65.pdf>.
- [30] HÖRMANDER, L. (1985). *The Analysis of Linear Partial Differential Operators. III. Pseudodifferential operators. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **274**. Springer, Berlin. [MR0781536](#)
- [31] JOHNSTONE, I. M. and SILVERMAN, B. W. (1990). Speed of estimation in positron emission tomography and related inverse problems. *Ann. Statist.* **18** 251–280. [MR1041393](#)
- [32] KERKYACHARIAN, G., KYRIAZIS, G., LE PENNEC, E., PETRUSHEV, P. and PICARD, D. (2010). Inversion of noisy Radon transform by SVD based needlets. *Appl. Comput. Harmon. Anal.* **28** 24–45. [MR2563258](#)
- [33] KNAPIK, B. T., VAN DER VAART, A. W. and VAN ZANTEN, J. H. (2011). Bayesian inverse problems with Gaussian priors. *Ann. Statist.* **39** 2626–2657. [MR2906881](#)
- [34] KUCHMENT, P. (2014). *The Radon Transform and Medical Imaging. CBMS-NSF Regional Conference Series in Applied Mathematics* **85**. SIAM, Philadelphia, PA. [MR3137435](#)
- [35] LOUIS, A. K. (1984). Orthogonal function series expansions and the null space of the Radon transform. *SIAM J. Math. Anal.* **15** 621–633. [MR0740700](#)
- [36] MCLEAN, W. (2000). *Strongly Elliptic Systems and Boundary Integral Equations*. Cambridge Univ. Press, Cambridge. [MR1742312](#)
- [37] MONARD, F. (2014). Numerical implementation of geodesic X-ray transforms and their inversion. *SIAM J. Imaging Sci.* **7** 1335–1357. [MR3217215](#)
- [38] MUHOMETOV, R. G. (1977). The reconstruction problem of a two-dimensional Riemannian metric, and integral geometry. *Dokl. Akad. Nauk SSSR* **232** 32–35. [MR0431074](#)
- [39] MUNK, W. and WUNSCH, C. (1979). Ocean acoustic tomography: A scheme for large scale monitoring. *Deep-Sea Res.* **26A** 123–161.
- [40] NATTERER, F. (1986). *The Mathematics of Computerized Tomography*. Wiley, Chichester. [MR0856916](#)
- [41] NICKL, R. (2017). Bernstein–von Mises theorems for statistical inverse problems I: Schrödinger equation. ArXiv preprint. Available at [arxiv:1707.01764](https://arxiv.org/abs/1707.01764).
- [42] NICKL, R. and SÖHL, J. (2017). Nonparametric Bayesian posterior contraction rates for discretely observed scalar diffusions. *Ann. Statist.* **45** 1664–1693. [MR3670192](#)
- [43] NOVIKOV, R. G. (2002). An inversion formula for the attenuated X-ray transformation. *Ark. Mat.* **40** 145–167. [MR1948891](#)

- [44] PATERNAIN, G. P., SALO, M., UHLMANN, G. and ZHOU, H. (2018). The geodesic X -ray transform with matrix weights. *Am. J. Math.* To appear. Available at arXiv:1605.07894.
- [45] PESTOV, L. and UHLMANN, G. (2004). On characterization of the range and inversion formulas for the geodesic X -ray transform. *Int. Math. Res. Not.* **80** 4331–4347. MR2126628
- [46] PESTOV, L. and UHLMANN, G. (2005). Two dimensional compact simple Riemannian manifolds are boundary distance rigid. *Ann. of Math. (2)* **161** 1093–1110. MR2153407
- [47] RADON, J. Über die Bestimmung Von Funktionen Durch Ihre Integralwerte Längs Gewisser Mannigfaltigkeiten. *Berichte über die Verhandlungen der Königlich-Sächsischen Akademie der Wissenschaften zu Leipzig. Mathematisch-Physische Klasse* **69** 262–277.
- [48] RAY, K. (2013). Bayesian inverse problems with non-conjugate priors. *Electron. J. Stat.* **7** 2516–2549. MR3117105
- [49] SALO, M. and UHLMANN, G. (2011). The attenuated ray transform on simple surfaces. *J. Differential Geom.* **88** 161–187. MR2819758
- [50] SHARAFUTDINOV, V., SKOKAN, M. and UHLMANN, G. (2005). Regularity of ghosts in tensor tomography. *J. Geom. Anal.* **15** 499–542. MR2190243
- [51] SHARAFUTDINOV, V. A. (1994). *Integral Geometry of Tensor Fields*. VSP, Utrecht. MR1374572
- [52] SHARAFUTDINOV, V. A. (2005). Geometric symbol calculus for pseudodifferential operators. I. [Translation of *Mat. Tr.* **7**(2) (2004) 159–206; MR2124544]. *Siberian Adv. Math.* **15** 81–125. MR2191866
- [53] STEFANOV, P. and UHLMANN, G. (2004). Stability estimates for the X -ray transform of tensor fields and boundary rigidity. *Duke Math. J.* **123** 445–467. MR2068966
- [54] STUART, A. M. (2010). Inverse problems: A Bayesian perspective. *Acta Numer.* **19** 451–559. MR2652785
- [55] TAYLOR, M. E. (1996). *Partial Differential Equations. I. Basic Theory. Applied Mathematical Sciences* **115**. Springer, New York. MR1395148
- [56] THORBERGSSON, G. (1978). Closed geodesics on non-compact Riemannian manifolds. *Math. Z.* **159** 249–258. MR0493872
- [57] TRÈVES, F. (1980). *Introduction to Pseudodifferential and Fourier Integral Operators. Vol. 1. Pseudodifferential Operators*. Plenum Press, New York. MR0597144
- [58] UHLMANN, G. and VASY, A. (2016). The inverse problem for the local geodesic ray transform. *Invent. Math.* **205** 83–120. MR3514959
- [59] VAN DER VAART, A. W. (1998). *Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge Univ. Press, Cambridge. MR1652247
- [60] VAN DER VAART, A. W. and VAN ZANTEN, J. H. (2008). Rates of contraction of posterior distributions based on Gaussian process priors. *Ann. Statist.* **36** 1435–1463. MR2418663
- [61] VIŠIK, M. I. and ÈSKIN, G. I. (1965). Convolution equations in a bounded region. *Uspekhi Mat. Nauk* **20** 89–152. MR0185273
- [62] Seppacher’s FEM_beta package. Available at <http://math.univ-lyon1.fr/homes-www/seppacher/codes.php>.

F. MONARD
 DEPARTMENT OF MATHEMATICS
 UNIVERSITY OF CALIFORNIA, SANTA CRUZ
 SANTA CRUZ, CALIFORNIA 95064
 USA
 E-MAIL: fmonard@ucsc.edu

R. NICKL
 G. P. PATERNAIN
 DEPARTMENT OF PURE MATHEMATICS
 AND MATHEMATICAL STATISTICS
 UNIVERSITY OF CAMBRIDGE
 WILBERFORCE ROAD, CAMBRIDGE CB3 0WB
 UNITED KINGDOM
 E-MAIL: r.nickl@statslab.cam.ac.uk
 g.p.paternain@dpmmms.cam.ac.uk