PAPER • OPEN ACCESS

Stable reconstruction of simple Riemannian manifolds from unknown interior sources

To cite this article: Maarten V de Hoop *et al* 2023 *Inverse Problems* **39** 095002

View the article online for updates and enhancements.

Stable reconstruction of simple Riemannian manifolds from unknown interior sources

Maarten V de Hoop[1]**, Joonas Ilmavirta**[2] ⓘ**, Matti Lassas**[3]
and Teemu Saksala[4,*] ⓘ

[1] Computational Applied Mathematics & Operations Research, Rice University, 6100 Main MS-134, Houston, TX 77005-1892, United States of America
[2] Department of Mathematics and Statistics, University of Jyväskylä, PO Box 35 (MaD), FI-40014 Jyväskylä, Finland
[3] Department of Mathematics and Statistics, University of Helsinki, PO Box 68 (Gustaf Hällströmin katu 2B), FI-00014 Helsinki, Finland
[4] Department of Mathematics, North Carolina State University, 2311 Stinson Drive, Raleigh, NC 27695-8205, United States of America

E-mail: tssaksal@ncsu.edu

Abstract

Consider the geometric inverse problem: there is a set of delta-sources in space-time that emit waves travelling at unit speed. If we know all the arrival times at the boundary cylinder of the spacetime, can we reconstruct the space, a Riemannian manifold with boundary? With a finite set of sources we can only hope to get an approximate reconstruction, and we indeed provide a discrete metric approximation to the manifold with explicit data-driven error bounds when the manifold is simple. This is the geometrization of a seismological inverse problem where we measure the arrival times on the surface of waves from an unknown number of unknown interior microseismic events at unknown times. The closeness of two metric spaces with a marked boundary is measured by a labeled Gromov–Hausdorff distance. If measurements are done for

infinite time and spatially dense sources, our construction produces the true Riemannian manifold and the finite-time approximations converge to it in the metric sense

(Some figures may appear in colour only in the online journal)

1. Introduction

We study a geometric inverse problem arising from global seismology. Multiple sources go off at unknown times and we measure the arrival times, not knowing which arrivals belong to the same event. The aim is to extract as much information as possible about the planet, modeled here by a simple Riemannian manifold M of dimension $n = \dim(M) \geqslant 2$, and the sources from this information.

Mathematically, the problem boils down to this: Suppose we know the boundary of a Riemannian manifold and for some unknown collection of interior points we know the union of their boundary distance graphs, each shifted with an unknown offset. What can we say about the manifold and these special points from this data? A more detailed description of the data can be found in section 2 below.

The distances between the source points turn out to be determined exactly by the data (theorem 2), but there is no hope of reconstructing the manifold if the source set is finite. However, the source set can be regarded as a discrete metric space, and this space approximates the Riemannian manifold. It is not only close as a metric space, but we can also assign approximate boundary points with good accuracy (theorem 9). To describe the similarity of two metric spaces 'with the same boundary', we define a labeled Gromov–Hausdorff distance. This extends the classical Gromov–Hausdorff distance and compares both the similarity of the metric spaces and the sameness of the boundaries—with a fixed model space for the boundary. It is in the sense of this distance that our approximation is good, and the quality of the approximation can be estimated explicitly directly from data.

The sources can be point sources in space time produced by a suitable stochastic process (proposition 12), for example. We present a method to look at the data on the boundary without knowing the manifold and reconstruct an approximate manifold—and get an explicit error bound on our reconstruction. This increases the applicability of geometric inverse problems to the messy cases with real data, but we do not attempt to optimize the bounds. When it comes to explicit geometric estimates, this paper should be seen as a proof of concept.

We assume the source events to be discrete in the spacetime, but their projections to the space can be dense. In this case the approximate reconstruction is actually perfect, and even the smooth and Riemannian structure is determined (theorem 1). If measurements are made for finite but increasing time, the approximate reconstructions converge to the true manifold in the labeled Gromov–Hausdorff distance if the sources are spatially dense (theorem 11).

Estimation of the quality of the approximation from data alone is lossy. To prove that it is not hopelessly so, we prove that by making the source points dense enough the observed density can be made arbitrarily small. There is an explicit estimate between the true and observed density in both directions (theorems 1 and 10).

We solve a problem stemming from physics by developing geometry rather than by applying ad hoc tricks. These developments include the introduction of the labeled Gromov–Hausdorff distance and various quantitative descriptions of simplicity of a Riemannian manifold.

From a more applied point of view, the key advances are that we introduce of a new and more physically relevant problem and provide quantitative and fully data-driven error estimates.

The precise definitions and statements of results are given in section 2. We will discuss the results and their context in section 3. The overall strategy of proofs is given in section 4 and the actual proofs follow in subsequent sections.

2. Definitions and theorems

2.1. Arrival time data

Let M be a Riemannian manifold with boundary. We denote by $d\colon M \times M \to \mathbb{R}$ its Riemannian distance. The spacetime $M \times \mathbb{R}$ comes with two natural projections, π onto M and τ onto $\mathbb{R}$.

Let $S \subset M \times \mathbb{R}$ be a set of sources. The source $s \in S$ goes off at the point $\pi(s)$ at the time $\tau(s)$. The data measured is the arrival times of signals at the boundary, given as the set

$$Q(S) = \{(x, \tau(s) + d(x, \pi(s))); \, x \in \partial M, s \in S\}.$$

We emphasize that we know the set $Q(S)$ but not the set S. We only observe the data as this point set without any labels to tell which arrivals correspond to the same sources.

The arrival time function of a source $s \in S$ is the function $a_s\colon \partial M \to \mathbb{R}$ given by $a_s(x) = \tau(s) + d(\pi(s), x)$. Let us denote the graph of the function a_s by $\mathcal{G}(a_s) \subset \partial M \times \mathbb{R}$. The data set can be rewritten as $Q(S) = \bigcup_{s \in S} \mathcal{G}(a_s)$.

Let $\phi\colon \partial M_1 \to \partial M_2$ be a diffeomorphism between two boundaries. If $A \subset \partial M_2 \times \mathbb{R}$, we define

$$\phi^* A = \{(\phi^{-1}(x), t); \, (x, t) \in A\} \subset \partial M_1 \times \mathbb{R}.$$

The set A would typically be the data set $Q(S_2)$ or a subset thereof. When there are two manifolds, we decorate all the objects with the subscripts 1 and 2 when clarity requires so.

A subset of the spacetime $M \times \mathbb{R}$ is said to be discrete when it has no accumulation points. If M is compact, it follows that any time slice $M \times [a, b]$ contains only finitely many points of a discrete set.

Our results apply to so-called simple manifolds. A Riemannian manifold is called simple if it is compact with strictly convex boundary (in the sense of definiteness of the second fundamental form) and the exponential map at every point is a diffeomorphism on its maximal domain of definition. Consequently there are no conjugate points and any two points on the manifold, including the boundary, are joined by a unique geodesic depending smoothly on the endpoints.

A key concept in our analysis is density. We say that a subset $A \subset X$ of a metric space X is ε-dense if the balls $B(a, \varepsilon)$ with $a \in A$ cover X. Equivalently, every point in X has a point of A at a distance below ε.

We set out to study whether and how a Riemannian manifold M can be determined by the set $Q(S)$ and prior knowledge of the boundary ∂M.

2.2. Precise determination

We begin by studying what can be determined from the data with no error.

The source sets $S_i \subset M_i \times \mathbb{R}$ can be finite or infinite. However, the discreteness assumption implies that every $S_i \cap (M_i \times [-T, T])$ is finite and so the source sets are at most countably infinite. Although the sources are not dense in the spacetime, the source points in space can be dense. In that case the same data determines the whole manifold up a diffeomorphism.

Theorem 1. *Let M_1 and M_2 be simple Riemannian manifolds so that the diameter of both manifolds is at most $\mathscr{C}_{\mathrm{diam}} > 0$ and the sectional curvature is bounded from above by $\mathscr{C}_{\mathrm{sec}+} > 0$ so that*

$$\mathscr{C}_{\mathrm{diam}} \sqrt{\mathscr{C}_{\mathrm{sec}+}} < \pi. \tag{1}$$

Let $\phi\colon \partial M_1 \to \partial M_2$ be a Riemannian isometry. Suppose the two sets $S_i \subset \mathrm{int}(M_i) \times \mathbb{R}$ are discrete in $M_i \times \mathbb{R}$. Let $Q(S_i) \subset \partial M_i \times \mathbb{R}$ be defined as above.

If the set $\pi_1(S_1) \subset M_1$ is dense and $Q(S_1) = \phi^* Q(S_2)$, then

(1) there is a smooth Riemannian isometry $\Phi\colon M_1 \to M_2$ so that $\phi = \Phi|_{\partial M_1}$ and
(2) there is a bijection $\xi\colon S_1 \to S_2$ between the sources so that

$$\xi(p,t) = (\Phi(p), t) \tag{2}$$

for all $(p,t) \in S_1$.

The condition (1) in most of our theorems always holds when the manifold is *a priori* known to have non-positive sectional curvature. See section 3.2.4 for details.

The theorem immediately implies that the source times coincide: $\tau_1(s) = \tau_2(\xi(s))$ for all $s \in S_1$.

The main novelty of theorem 1 is in allowing the data set $Q(S)$ to be a union rather than an indexed collection of graphs of the arrival time functions. The rest of our results do not have a similar precedent, and our method of proof for theorem 1 is new and constructive. See the discussion in section 3 below for details.

The next result concerns a situation with a finite amount of sources. Informally, the next theorem states that knowledge of ∂M as a smooth manifold and the set $Q(S)$ determines the distances between the sources, the time differences, and two kinds of distance differences at the boundary. More formally:

Theorem 2. *Let M_1 and M_2 be simple Riemannian manifolds and $\phi\colon \partial M_1 \to \partial M_2$ a diffeomorphism. Suppose the two sets $S_i \subset \mathrm{int}(M_i) \times \mathbb{R}$ are discrete in $M_i \times \mathbb{R}$ and $\#\pi(S_1) \geqslant 2$. Let $Q(S_i) \subset \partial M_i \times \mathbb{R}$ be defined as above.*

If $Q(S_1) = \phi^* Q(S_2)$, then

(1) there is a bijection $\xi\colon S_1 \to S_2$,
(2) $d_1(\pi_1(s), \pi_1(r)) = d_2(\pi_2(\xi(s)), \pi_2(\xi(r)))$ for all $s, r \in S_1$,
(3) $\tau_1(s) - \tau_1(r) = \tau_2(\xi(s)) - \tau_2(\xi(r))$ for all $s, r \in S_1$,
(4) $d_1(\pi_1(s), x) - d_1(\pi_1(r), x) = d_2(\pi_2(\xi(s)), \phi(x)) - d_2(\pi_2(\xi(r)), \phi(x))$ for all $x \in \partial M_1$ and $s, r \in S_1$, and
(5) $d_1(\pi_1(s), x) - d_1(\pi_1(s), y) = d_2(\pi_2(\xi(s)), \phi(x)) - d_2(\pi_2(\xi(s)), \phi(y))$ for all $x, y \in \partial M_1$ and $s \in S_1$.

While measuring for infinite time gives full uniqueness, it is not realistic. Therefore we consider measurements for a finite but increasing time and show that there is an approximate reconstruction and in the limit of infinite time it tends to the correct one in a suitable sense. To be able to state the results, we need to set up a way to compare the true manifold to a discrete approximation. Before pursuing this direction, we record a result that helps clean up the data by disentangling the union of graphs into separate graphs. The arrival time functions a_s are easily verified to be smooth when $\pi(s) \notin \partial M$ and the manifold is simple.

Proposition 3. *Let M be a simple Riemannian manifold. Suppose $S \subset \mathrm{int}(M) \times \mathbb{R}$ is discrete in $M \times \mathbb{R}$. Then the set Q(S) and the Riemannian structure of ∂M determine the set*

$$\{a_s; s \in S\},$$

of arrival time functions. Furthermore, if $\Omega \subset \partial M \times \mathbb{R}$ is any open set, then Q(S) determines the set of connected components of the sets $\mathcal{G}(a_s) \cap \Omega$, $s \in S$.

2.3. Labeled Gromov–Hausdorff distance

This subsection is devoted to the metric geometry we will use to compare our approximate reconstructions to the true manifold.

Definition 4. Let *X* and *Y* be compact metric spaces and *L* any set (understood as a set of labels). Let $\alpha\colon L \to X$ and $\beta\colon L \to Y$ be any functions. We define the labeled Gromov–Hausdorff distance between (X, α) and (Y, β) to be

$$d_{GH}^L(X, \alpha; Y, \beta) = \inf\{d_H^Z(f(X), g(Y)) + \sup_{\ell \in L} d_Z(f(\alpha(\ell)), g(\beta(\ell)));$$

$$Z \text{ is a compact metric space,}$$

$$f\colon X \to Z \text{and } g\colon Y \to Z \text{are isometric embeddings}\}.$$

Here d_H^Z is the Hausdorff distance on the metric space *Z*. If the set of labels is empty, then this distance reduces to the usual Gromov–Hausdorff distance; in this context we understand the supremum of an empty set to be zero.

We will be comparing our discrete reconstruction to the manifold *M* with labels given by the known set $L = \partial M$, labeled by the inclusion $\iota\colon \partial M \to M$. To make it meaningful to state that the labeled Gromov–Hausdorff distance is small, we must ensure that this is a well-behaved concept of distance.

The point of this definition is that in addition to getting a good metric approximation of the manifold in the sense of Gromov–Hausdorff distance, we want to get knowledge of how well the known boundary ∂M sits inside the approximation. This is what the concept is designed to do.

Proposition 5. *The labeled Gromov–Hausdorff distance is symmetric and satisfies the triangle inequality. Moreover,*

$$d_{GH}^L(X, \alpha; Y, \beta) = 0$$

if and only if there is an isometry $h\colon X \to Y$ so that $h \circ \alpha = \beta$.

2.4. Quantitative simplicity

We will make approximate reconstructions of the whole manifold, and we can quantify the error of the reconstruction precisely once we have *a priori* bounds on curvature, diameter, and similar geometric properties.

This set of properties might not be minimal in any way, but the following are the estimates we shall need. By proposition 7 all simple manifolds do indeed have bounded geometry in the sense of the next definition.

Definition 6. We say that a manifold *M* with boundary has bounded geometry with the constants $\mathscr{C}_{\mathrm{JF}}$, $\mathscr{C}_{\mathrm{SFF}}$, $\mathscr{C}_{\mathrm{diam}}$, $\mathscr{C}_{\mathrm{sec}-}$, $\mathscr{C}_{\mathrm{sec}+}$, $\mathscr{C}_{\mathrm{exp}}$, $\mathscr{C}_{\mathrm{dist}}$, $\mathscr{C}_{\mathrm{H1}}$, and $\mathscr{C}_{\mathrm{H2}}$ (each of them > 0) if the following properties hold:

(1) The diameter of M is at most $\mathscr{C}_{\text{diam}}$.
(2) The sectional curvature is always in the range $[-\mathscr{C}_{\text{sec}-}, \mathscr{C}_{\text{sec}+}]$.
(3) For any $x \in M$ and $\eta_1, \eta_2 \in T_x M$ we have

$$|\eta_1 - \eta_2| \leqslant \mathscr{C}_{\exp} d(\exp_x \eta_1, \exp_x \eta_2)$$

as long as the exponentials are defined.
(4) Every Jacobi field J with $J(0) = 0$ along any unit speed geodesic satisfies

$$\frac{t}{2} \partial_t |J(t)|^2 \leqslant \mathscr{C}_{\text{JF}} |J(t)|^2$$

for all $t > 0$ for which the geodesic is defined.
(5) If h_1 and h_2 are, respectively, the first and second fundamental forms of the boundary ∂M, then $h_2 \leqslant \mathscr{C}_{\text{SFF}} h_1$.
(6) The distances on M and along ∂M satisfy

$$d_M(x,y) \leqslant d_{\partial M}(x,y) \leqslant \mathscr{C}_{\text{dist}} d_M(x,y)$$

for all $x, y \in \partial M$.
(7) Take any unit speed geodesic γ on M and let ρ be the distance function $\rho(x) = d(x, \gamma(0))$. For any $t > 0$, let $w \in T_{\gamma(t)} M$ be any vector orthogonal to $\dot{\gamma}(t)$. Then the Hessian $\mathcal{H}_\rho$ of ρ satisfies

$$\langle w, \mathcal{H}_\rho w \rangle \geqslant |w|^2 (\mathscr{C}_{\text{H1}} t^{-1} - \mathscr{C}_{\text{H2}}).$$

We have to point out, however, that bounded geometry does not imply simplicity. This is because the estimates below do not ensure strictly convex boundary, and indeed any smooth domain in a simple manifold satisfies all the estimates. Strict convexity of the boundary is important for our results, but it is used in a qualitative rather than a quantitative way. See section 3.2.2 and proposition 3 for the use of convexity. If item (5) is improved from $h_2 \lesssim h_1$ to $h_1 \lesssim h_2 \lesssim h_1$, then strict convexity does indeed follow.

We will derive a number of other estimates from those of definition 6 with constants depending on the ones listed here. See section 9 for details. The constants appearing in the derived estimates are all given in appendix B. We will refer to the appendix whenever a new constant is introduced in a claim. All claims are given with simple constants, so there will be a number of relations between the various constants as is clear from the appendix.

It should be noted in part (7) of the definition that w is a vector and $\mathcal{H}_\rho w$ is a covector. The inner product is their duality pairing, which we find most convenient to write as $\langle w, \mathcal{H}_\rho w \rangle$ instead of $(\mathcal{H}_\rho w)(w)$ or $\mathcal{H}_\rho(w, w)$.

Having these estimates and constants is not an added assumptions, but merely a quantification of the simplicity of the manifold M. Our next result justifies that the constants of definition 6 can well be called constants of quantitative simplicity.

Proposition 7. *Every simple manifold has bounded geometry in the sense of definition 6 for some constants.*

2.5. Approximate determination

Theorems 1 and 2 state what can be determined exactly from our data. We now turn to studying how well the whole manifold can be reconstructed with finite data.

The source set is as before. The measurements at the boundary start at time $t = 0$ and continue up to time $t = T$. From this finite amount of data we can then draw an approximate conclusion about the geometry of our manifold. The error can be quantified very concretely in terms of the geometric bounds.

The only points we know on the manifold are the finite number of sources that have produced a boundary signal on the time interval $[0, T]$ and the *a priori* known boundary itself. Therefore what we have is a discrete approximation of the smooth Riemannian manifold. In addition to getting a metric approximation of the manifold itself, we need to approximately embed the known boundary ∂M into the approximation, and so we will use the Gromov–Hausdorff distance labeled by ∂M.

Let us denote the set of spatial source points by $P = \pi(S)$. To each $p \in P$ is associated a boundary distance function $r_p \colon \partial M \to \mathbb{R}$ given by $r_p(x) = d(x, p)$. Let $c(p) \subset \partial M$ be the set of critical points of this function. By compactness $\#c(p) \geqslant 2$.

Consider a simple manifold of bounded geometry in the sense of definition 6. At $x \in c(p)$ let $\lambda(p, x)$ be the smallest eigenvalue of the Hessian of r_p at x. This requires that ∂M is known as a Riemannian manifold, not just as a smooth one. For $y \in c(p)$ we define

$$E(p, y) = \begin{cases} \frac{\mathscr{C}_{\mathrm{JF}}}{\lambda(p, y) - \mathscr{C}_{\mathrm{SFF}}}, & \text{when } \lambda(p, y) > \mathscr{C}_{\mathrm{SFF}} \\ \infty, & \text{otherwise} \end{cases} \tag{3}$$

and for all $x \in \partial M$

$$E(x) = \inf_{p \in P, y \in c(p)} \left(E(p, y) + d_{\partial M}(x, y) \right). \tag{4}$$

Finally, we let

$$E = \sup_{x \in \partial M} E(x). \tag{5}$$

These estimate the following geometric quantities from above:

- $E(p, y)$ estimates the distance $d(p, y)$.
- $E(x)$ estimates the distance from x to the nearest source point.
- E estimates how far at worst a boundary point can be from a source point.

These three types of Es will be all determined by the data and the constants of bounded geometry. Observe that if the Hessians of the distance functions are never large enough, all these Es may well be infinite.

For any $p, q \in P$ and $r, s > 0$ we define the lentil

$$L_{r,s}^{p,q} = B(p, r) \cap B(q, s).$$

The lentil is said to have thickness

$$\delta_{r,s}^{p,q} = r + s - d(p, q).$$

One can check that on a simple manifold the thickness is the length of the segment of the geodesic $\gamma_{p,q}$ between the two points that lies within the lentil, provided that $r, s < d(p, q)$. To see this, observe that a point $\gamma_{p,q}(t)$ is in the lentil if and only if $t < r$ and $d(p, q) - t < s$.

Our inversion procedure starts with preprocessing the data. Proposition 3 gives one aspect of it, and the next one says that the derived quantities introduced above are determined by the data.

Proposition 8. *Let M_1 and M_2 be simple Riemannian manifolds. Let $\phi \colon \partial M_1 \to \partial M_2$ be a smooth isometry. Suppose that the two sets $S_i \subset \mathrm{int}(M_i) \times \mathbb{R}$ are discrete in $M_i \times \mathbb{R}$ and*

$\#\pi(S_1) \geqslant 2$. Let $Q(S_i) \subset \partial M_i \times \mathbb{R}$ be defined as above. Let $\xi\colon S_1 \to S_2$ be the bijection of theorem 2 and let $\Xi\colon P_1 \to P_2$ be the map for which $\Xi \circ \pi_1 = \pi_2 \circ \xi$. Then:

(1) The critical point sets correspond via $c(\Xi(p)) = \phi(c(p))$ for all $p \in P_1$. The boundary distance quantities correspond via $E_1(p,y) = E_2(\Xi(p), \phi(y))$ for all $p \in P$ and $y \in c(p)$. Moreover and $E_1(x) = E_2(\phi(x))$ for all $x \in \partial M_1$ and $E_1 = E_2$.
(2) If we fix any $\varepsilon_1 > 0$ and define

$$\Gamma_i = \{p \in P_i;\ E_i(p,y) < \varepsilon_1 \text{for some } y \in c(p)\}, \tag{6}$$

then $\Gamma_2 = \Xi(\Gamma_1)$.
(3) Fix any $\varepsilon_1 > 0$ and $\delta > 0$. Define Γ_i for $i = 1,2$ as above. Take any $x_1, y_1 \in \Gamma_1$ so that $d(x_1, y_1) > \delta$, and pick any $r \in (\delta, d_1(x_1, y_1))$. Set $s = d_1(x_1, y_1) - r + \delta$ and let

$$L_1 = L_{r,s}^{x_1,y_1} \subset M_1 \quad and \quad L_2 = L_{r,s}^{x_2,y_2} \subset M_2,$$

where $x_2 = \Xi(x_1)$ and $y_2 = \Xi(y_1)$. Then

$$\Xi(L_1 \cap P_1) = L_2 \cap P_2$$

and so

$$L_1 \cap P_1 \neq \emptyset \quad if\ and\ only\ if \quad L_2 \cap P_2 \neq \emptyset.$$

Now that we know the data to determine the relevant auxiliary quantities uniquely, we move to working on a single manifold. The next theorem states that under known bounds on the geometry and metrically known boundary one can estimate directly, from the data, how dense the source points are. For our purposes this is to assign approximate boundary points, and estimate the metric distance of an approximate reconstruction from the true manifold.

In practice, the set P in the next theorem stands for the spatial source set $\pi(S)$ as above, but the claim is valid for any subset of $\mathrm{int}(M)$.

Theorem 9. *Let M be simple a Riemannian manifold with the constants of definition 6 (by proposition 7) satisfying the condition (1), and let $P \subset \mathrm{int}(M)$ be any set. This theorem uses the constants given in equations (40), (41), (42) and (43).*

Fix any $\varepsilon_1 > 0$ and define the set $\Gamma \subset P$ as in (6). Set $\varepsilon_2 = \varepsilon_1 + E$ and $\delta = C_1\varepsilon_2$. Suppose that for every $x,y \in \Gamma$ with $d(x,y) > \delta$ and $r \in (\delta, d(x,y))$ we have

$$L_{r,d(x,y)-r+\delta}^{x,y} \cap P \neq \emptyset. \tag{7}$$

Then:

(1) The set P is ε-dense in M with

$$\varepsilon = C_2\varepsilon_2 + C_3\sqrt{\varepsilon_2}. \tag{8}$$

(2) A function $\alpha\colon \partial M \to P$ satisfying

$$E(x) + \varepsilon_1 \geqslant \inf_{y \in c(\alpha(x))} \left(E(\alpha(x),y) + d_{\partial M}(x,y) \right) \tag{9}$$

for all $x \in \partial M$ can be constructed from the data. With any such function the pair (P, α) is close to (M, ι), with the inclusion $\iota \colon \partial M \to M$, in the sense of the labeled Gromov–Hausdorff distance:

$$d_{GH}^{\partial M}(P, \alpha; M, \iota) \leqslant C_4 \varepsilon_2 + C_3 \sqrt{\varepsilon_2}. \tag{10}$$

The first part of this theorem says that under an additional curvature assumption the quantities defined by the data (due to proposition 8) give an explicit estimate for the density of the source points. The second part is an estimate of the labeled Gromov–Hausdorff distance between the approximating discrete space P with the approximate boundary described by α and the true manifold M with boundary ∂M. The key assumption is that all lentils of thickness δ, which is proportional to ε_2, contain a source point, and by proposition 8 this is determined uniquely by the data.

The quality of the estimates given in claim (1) of theorem 9 depend on the choice of ε_1. When faced with a practical reconstruction task, one ought to vary ε_1 and see which choice gives the best error bound ε_2 while still having non-empty lentils. The only parameter we choose is ε_1; the second parameter ε_2 which plays a more central role is based on the observed quantity E of (5). Heuristically, ε_1 describes how close the chosen source points are to the boundary and ε_2 how far from any given boundary point these near-boundary sources are. These constants are then used to establish estimates for interior density of the sources.

The estimates of claim (1) of theorem 9 are vacuous if $\varepsilon_2 = \infty$. For the theorem to be meaningful, we need to ensure that for some ε_1 we indeed have $\varepsilon_2 < \infty$. To be able to get an arbitrarily good estimate for density and metric distance, we also need to ensure that ε_2 will become as small as we like if the source set $P \subset M$ is dense enough. This is a reverse density estimate: If the source set is very dense, it will appear dense by our estimates.

Theorem 10. *Let M be a simple manifold of bounded geometry with the constants of definition 6 (by proposition 7) satisfying the condition (1) and $S \subset \mathrm{int}(M) \times \mathbb{R}$ a discrete source set. In this theorem we use the constants of (44), (45), (46), (47), (43), and (42).*

Take any $\varepsilon > 0$. If $P = \pi(S) \subset M$ is $\hat{\varepsilon}$-dense with

$$\hat{\varepsilon} = \min \left(C_6, C_7 \varepsilon, C_8 \varepsilon^2 \right), \tag{11}$$

and $\varepsilon_1 = C_5 \hat{\varepsilon}$, then:

(1) The quantity $\varepsilon_2 = \varepsilon_1 + E$ as used in theorem 9 satisfies

$$C_4 \varepsilon_2 + C_3 \sqrt{\varepsilon_2} < \varepsilon. \tag{12}$$

(2) Denote $\delta = C_1 \varepsilon_2$. For any $x, y \in P$ that are more than distance δ apart and any $r \in \left(\delta, d(x, y) \right)$ the lentil $L_{r, d(x,y)-r+\delta}^{x,y}$ contains an element of P.

(3) The set P satisfies the assumptions of theorem 9 for the given choice of ε_1; especially the lentil intersection property (7) holds. Therefore one can conclude from the data set $Q(S)$ and the geometric constants alone that P is ε-dense and the discrete approximation is ε-good.

Choosing ε_1 correctly is important for theorem 10. If ε_1 is too large, the resulting estimates are not tight enough. If it is too small, then too few boundary sources are identified and not all lentils contain source points.

2.6. Convergence of approximations

The next theorem treats the case of gradually acquired and interpreted data. Assuming the sources are eventually dense, a finite amount of data determines an approximate manifold. The data itself gives an estimate for how good the approximation is and the approximations converge to the true manifold with boundary in the labeled Gromov–Hausdorff sense.

Theorem 11. *Let M be a simple Riemannian manifold with geometry bounded in the sense of definition 6 satisfying the condition (1). Let $S \subset \text{int}(M) \times \mathbb{R}$ be a countable discrete set of sources so that the projection $\pi(S \cap (M \times [0, \infty))) \subset M$ is dense.*

Then for each $T > 0$ the boundary ∂M as a Riemannian manifold and the set

$$Q(S, T) = Q(S) \cap (\partial M \times [0, T]) \tag{13}$$

determine a finite metric space M_T and a map $\alpha_T \colon \partial M \to M_T$ so that

$$\lim_{T \to \infty} d_{GH}^{\partial M}(M_T, \alpha_T; M, \iota) = 0.$$

Thus $(M_T, \alpha_T) \to (M, \iota)$ as $T \to \infty$ in the labeled Gromov–Hausdorff sense.

Finally, we point out that if the geometric assumptions of theorem 11 are satisfied, then the assumptions on the sources are almost certainly satisfied in a simple stochastic model. That is, for almost any source pattern a finite amount of data determines an approximation and the approximation converges to the true solution with increasing data in a manner quantifiable from data itself.

We say that two measures are uniformly comparable when they are absolutely continuous with respect to each other and the Radon–Nikodym derivatives both ways are bounded. Let ν be the product of the volume measure of M and the Lebesgue measure of $\mathbb{R}$, and let μ be a measure on $M \times \mathbb{R}$ uniformly comparable with ν. A homogeneous Poisson point process on the spacetime $M \times \mathbb{R}$ with intensity $\lambda > 0$ is such a point process that the number of points in the set $A \subset M \times \mathbb{R}$ follows the Poisson distribution with parameter $\lambda \bar{\mu}(A)$.

Proposition 12. *Let M be a simple Riemannian manifold with bounded geometry satisfying (1). Let μ be a measure on $M \times \mathbb{R}$ which is uniformly comparable with the natural product measure ν. Suppose the set $S \subset M \times \mathbb{R}$ is given by a homogeneous Poisson point process with intensity $\lambda > 0$ and the measure μ.*

Then almost surely:

(1) S is countable and discrete,
(2) the projection $\pi(S \cap (M \times [0, \infty)))$ is contained in $\text{int}(M)$ and dense in M, and therefore
(3) the conclusions of theorem 11 hold.

That is, if the events $S \subset M \times \mathbb{R}$ are given by a Poisson process, then almost surely the reconstruction from data for finite but increasing time converges to the true solution as time increases.

According to the widely used epidemic-type aftershock model [21] earthquakes and their aftershocks do indeed follow a Poisson-type stochastic process.

3. Discussion

3.1. Geometric inverse problems and seismology

An elastic body—e.g. a planet—can be modeled as a manifold, where distance is measured in travel time: the distance between two points is the shortest time it takes for a wave to go from one point to the other. If the material is isotropic or elliptically anisotropic, then this elastic geometry is Riemannian. However, this sets a very stringent assumption on the stiffness tensor describing the elastic system, and Riemannian geometry is therefore insufficient to describe the propagation of seismic waves in the Earth. If no structural assumptions on the stiffness tensor apart from the physically necessary symmetry and positivity properties are made, this leads necessarily modeling the planet by a Finsler manifold as was explained in [13].

An isotropically elastic medium carries pressure (*P*) and shear (*S*) wave speeds that are conformally Euclidean metrics. Of these two the *P*-waves are faster [8]. In order to be true to the isotropic elasticity we should measure both *P*- and *S*-wave arrivals. As we have many sources going off in different locations and we do not know which arrival is related to which source we cannot *a priori* catalog arrivals of different wave types. We simplify this aspect of the problem by disregarding polarizations and considering only one type of wave.

In the two following subsections we review some seismologically relevant geometric inverse problems on Riemannian and Finsler manifolds. We also draw relations between the mathematical data, considered in this paper, and some real world seismic measurements. In the final subsection we discuss our contributions to field of inverse problems and possible future research directions suggested by the results of this paper.

3.1.1. Geometric inverse problems on Riemannian manifolds.   The problem of determination of the isometry type of a compact Riemannian manifold from its *boundary distance data*

$$\{r_p \colon \partial M \to (0, \infty); \, p \in \text{int}(M)\}$$

was introduced for the first time in [31]. The reconstruction of the smooth atlas on the manifold and the metric tensor in these coordinates was originally considered in [28]. In contrast to the paper at hand these earlier results do not need any extra assumption for the geometry, but have a complete data in the sense that the distance function r_p to the boundary is known for any interior source point $p \in \text{int}(M)$.

The problem of boundary distance data is related to many other geometric inverse problems. For instance, it is a crucial step in proving uniqueness for Gel'fand's inverse boundary spectral problem [28]. Gel'fand's problem concerns the question whether the data

$$\left(\partial M, (\lambda_j, \partial_\nu \phi_j|_{\partial M})_{j=1}^\infty\right)$$

determine (M, g) up to isometry, when (λ_j, ϕ_j) are the Dirichlet eigenvalues and the corresponding L^2-orthonormal eigenfunctions of the Laplace–Beltrami operator. Belishev and Kurylev provide an affirmative answer to this problem in [4].

In [29] the authors studied a question of approximating a Riemannian manifold under the assumption: On the boundary one can measure the travel times $\{r_p(z); z \in R\}$ for a collection of points $p \in \text{int}(M)$ with small errors. This problem corresponds to measuring wave fronts produced by point sources that go off at known initial times and the wave fronts produced by these sources can be separately measured for each interior source, whereas in the present paper we measure wave fronts produced by several point sources, that go off at unknown times, and moreover, the signals from the different point sources cannot be *a priori* separated but are

mixed together, that is, we observe the union of the wave fronts produced by several point sources.

We recall that for a source $s \in S$ we do not know the initial time $\tau(s)$, but due to proposition 3 we can recover the arrival times $a_s(z) = d(z, \pi(s)) + \tau(s)$ for each $z \in \partial M$. Taking the difference of the arrival times one obtains a boundary distance difference function

$$D_{\pi(s)}(z_1, z_2) := d(\pi(s), z_1) - d(\pi(s), z_2)$$

for all $z_1, z_2 \in \partial M$, and this is independent of the initial time. In [37] it is shown that if $U \subset N$ is a compact subset of a closed Riemannian manifold (N, g) and $\mathrm{int}(U) \neq \emptyset$, then *distance difference data*

$$((U, g|_U), \{D_x \colon U \times U \to \mathbb{R} \mid x \in N\})$$

determine (N, g) up to an isometry. This result was recently generalized to complete Riemannian manifolds [26] and for Riemannian manifolds with boundary [15, 25]. These results require sources at all interior points, unlike the ones in the present paper.

If the sign in the definition of the distance difference functions is changed, we arrive at the distance sum functions,

$$D_x^+(z_1, z_2) = d(z_1, x) + d(z_2, x)$$

for all $x \in M$ and $z_1, z_2 \in \partial M$. These functions give the lengths of the broken geodesics, that is, the union of the shortest geodesics connecting z_1 to x and the shortest geodesics connecting x to z_2. Also, the gradients of $D_x^+(z_1, z_2)$ with respect to z_1 and z_2 give the velocity vectors of these geodesics. The inverse problem of determining the manifold (M, g) from the *broken geodesic data*, consisting of the initial and the final points and directions, and the total length, of the broken geodesics, has been considered in [32]. In [32] the authors show that broken geodesic data determine the boundary distance data and use then the results of [28, 31] to prove that the broken geodesic data determine the Riemannian manifold up to an isometry. A different variant of broken geodesic data was recently considered in [41].

The Riemannian wave operator is a globally hyperbolic linear partial differential operator of real principal type. Therefore, the Riemannian distance function and the propagation of a singularity initiated by a point source in space time are related to one another. We let u be the solution of the Riemannian wave equation with a point source $s \in S$. In [17, 19] it is shown that the image, Λ, of the wave front set of u, under the musical isomorphism $T^*M \ni (x, \xi) \mapsto (x, g^{ij}(x)\xi_i) \in TM$, coincides with the image of the unit sphere $S_{\pi(s)}M$ at $\pi(s)$ under the geodesic flow of g. Thus $\Lambda \cap \partial(SM)$, where SM is the unit sphere bundle of (M, g), coincides with the exit directions of geodesics emitted from $\pi(s)$. In [38] the authors show that if (M, g) is a compact smooth non-trapping Riemannian manifold with smooth strictly convex boundary, then generically the *scattering data of point sources* $(\partial M, R_{\partial M}(M))$ determine (M, g) up to isometry. Here, $R_{\partial M}(x) \in R_{\partial M}(M)$ for $x \in M$ stands for the collection of tangential components to boundary of exit directions of geodesics from x to ∂M.

If M is an open subset of a complete Riemannian manifold (N, g), then one more important data set, given by the wave front set of u, is related to a *generalized spheres* of radius $r > 0$, that is given by the formula

$$S(p, r) := \{\exp_p(v) \in N; \, v \in T_pM, \, \|v\|_g = r, \quad \text{and } \exp_p \text{ is not singular at } v\}.$$

In [11] the authors show that the *spherical surface data*

$$\{S(q, t) \cap N \setminus M; \, q \in M \text{ and } t > 0\}$$

determine the universal cover space of *N*. If a generalized sphere $S(p,r)$ is given the authors show that there exists a specific coordinate structure in a neighborhood of any maximal normal geodesic to $S(p,r)$ such that in these coordinates metric tensor *g* can be determined. However this does not determine *g* globally. The authors provide an example of two different metric tensors which produce the same spherical surface data.

A classical geometric inverse problem, that is closely related to the distance functions, asks: does the Dirichlet-to-Neumann mapping of a Riemannian wave operator determine a Riemannian manifold up to isometry? For the full boundary data this problem was solved originally in [4] using the Boundary control method. Partial boundary data questions have been studied for instance in [36, 44]. Recently [34] extended these results for connection Laplacians. Lately also inverse problems related to non-linear hyperbolic equations have been studied extensively [33, 39, 58]. For a review of inverse boundary value problems for partial differential equations see [35, 56].

Maybe the most studied geometric inverse problem formulated with the distance functions is the Boundary rigidity problem. This problem asks: Does the *boundary distance function*, that gives a distance between any two boundary points, determine (M,g) up to an isometry? In an affirmative case (M,g) is said to be boundary rigid. For a general Riemannian manifold the problem is false: Suppose the manifold contains a domain with very slow wave speed, such that all the geodesics starting and ending at the boundary avoid this domain. Then in this domain one can perturb the metric in such a way that the boundary distance function does not change. It was conjectured in [43] that for all compact simple Riemannian manifolds the answer is affirmative. In two dimensions it was solved in [48]. For higher dimensional case the problem is still open, but different variations of it has been considered for instance in [7, 9, 53, 54].

A general feature of geometric inverse problems related to seismology is that each point source is analyzed in isolation and there is a source in every point on the manifold. We drop both of these assumptions to step towards a physically more accurate model. The source set can be finite and the data does not *a priori* make a distinction between different events. This distinction is achieved by proposition 3. Another novelty is that our discrete approximation is quantitatively stable in the sense of the labeled Gromov–Hausdorff distance. All our estimates only depend on the manifold *M* through the data and the constants of quantitative simplicity.

3.1.2. Related geometric inverse problems on Finsler manifolds. In [13] we studied the recovery of a compact Finsler manifold from its boundary distance data. In contrast to earlier Riemannian results [28, 31] the data only determines the topological and smooth structures, but not the global geometry. The Finsler function $F\colon TM \to [0,\infty)$ can be however recovered in a closure of the set $G(M,F) \subset TM$, that consists of points $(p,v) \in TM$ such that the corresponding geodesic $\gamma_{p,v}$ is distance minimizing to the terminal boundary point. We also showed that if the set $TM \setminus G(M,F)$ is non-empty then any small perturbation of *F* in this set leads to a Finsler metric whose boundary distance data agrees with the one of *F*. If $G(M,F) = TM$, then the boundary distance data determines (M,F) up to a Finsler isometry. For instance the isometry class of any simple Finsler manifold is determined by this data. The same is not true if only the boundary distance function is known [24]. Thus a simple Finsler manifold is never boundary rigid. In [14] we utilized the main result of [13] and generalized the result of [32], about the broken geodesic data, on reversible Finsler manifolds, satisfying a convex foliation condition.

Although simple Finsler manifolds are not boundary rigid there are results considering their rigidity questions for some special Finsler metrics. For instance it was shown in [45]

that Randers metrics $F_i = G_i + \beta_i$ indexed with $i \in \{1, 2\}$ with simple and boundary rigid Riemannian norm $G_i(x, v) = \sqrt{g_{ij}(x)v^iv^j}$ and closed one-form β_i, have the same boundary distance function if and only if $G_1 = \Psi^*G_2$ for some boundary fixing diffeomorphism $\Psi: M \to M$ and $\beta_1 - \beta_2 = \mathrm{d}\phi$ for some smooth function ϕ vanishing on ∂M. It is worth of mentioning that analogous results have been presented earlier on a Riemannian manifold in the presence of a magnetic field [2, 10].

The sphere data described in connection to the Riemannian results of [11] above has also been studied on Finsler manifolds [12]. Knowledge of the spheres uniquely determine the fundamental tensor and the curvature operator along any geodesic passing through the known domain, but in contrast to Riemannian geometry this information is insufficient for a full reconstruction of the universal cover.

3.1.3. Geophysical literature.

In fact, the inverse problem considered here can be directly related to seismology. The sources correspond with microseismic events with unknown locations and origin times while the metric corresponds with the wave speed. Here, we consider only one wave speed associated with elastic *P*-waves, but incorporating a second wave speed associated with elastic *S*-waves is quite straightforward. In the past decade there has been extensive research on the joint recovery of wave speed and event locations, with mining, geothermal and hydraulic-fracturing applications of induced seismicity but also in studies of the crustal structure. Here, we present a comprehensive analysis providing fundamental insight in the feasibility of succeeding in this, while focusing on recovering the wave speed. We use geometric data rather than wave fields; in applications, the extraction of these have been routine, for example, with template matching [18].

Initial empirical studies [5, 27] assumed a simple wave speed model varying in one coordinate only and one or two strings of receivers (in boreholes) penetrating the manifold. Strategies have been broadly based on optimization, such as variations of gradient descent [22, 50], with intertwining updating [42] and a neighborhood algorithm [55] employing different optimization criteria [52, 59] or via intermediate (approximate) interior wave-field recovery [51]. A basic statistical, Bayesian framework has been developed in parallel [61].

3.2. Discussion of technical assumptions

In this subsection we present a discussion of many technical assumptions we made in our theorems. Our focus here is on the key phenomenology of having multiple sources with unknown times and we have chosen to simplify other aspects of the problem.

3.2.1. Density of source points.

We assume in all our results that the sources are discrete in the spacetime $M \times \mathbb{R}$. This is physically reasonable and also practical. If the sources were dense or even accumulated somewhere, proposition 3 would become far more complicated.

However, there is little hope of full uniqueness with finitely many source points. The sources can be discrete in spacetime but dense in space, and this is indeed the setting of theorems 1 and 11 in which the Riemannian manifold is determined uniquely either from full time data or asymptotically from increasing time data.

3.2.2. Full data and convexity.

All our results concern full data in the sense that arrivals are recorded on the whole boundary ∂M. The full boundary is crucial for proposition 3 which disentangles the data into a collection of graphs, and we make so heavy use of differential tools
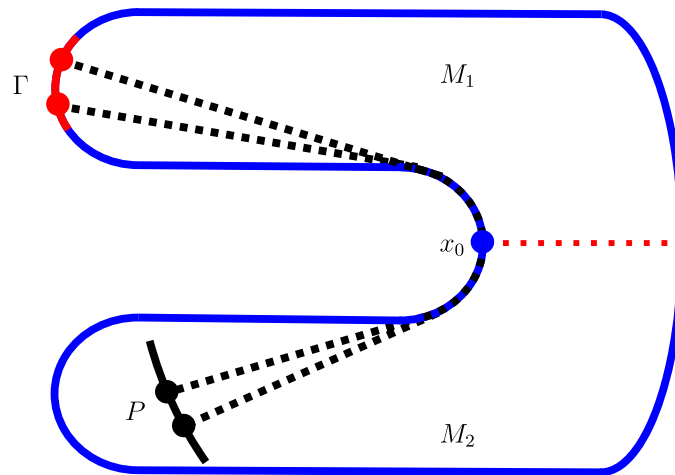
Figure 1. A domain where partial data is insufficient. We split the domain *M* into two
pieces M_1 and M_2 with respect to the line (red dotted line) that is normal to ∂M at $x_0 \in \partial M$ (blue dot). Then we choose a domain $\Gamma \subset \partial M_1$ (red arch) so that any minimizing
curve joining a point on Γ and a point in M_2 touches the boundary near x_0. The curve
$P \subset M$ is any involute of the boundary, meaning that the distance from all points on *P*
to x_0 is the same.

on the boundary throughout the paper that a discrete subset of the boundary is beyond current
reach.

Let us construct an explicit example of a surface *M* and a subset $\Gamma \subset \partial M$ so that our results
fail with data recorded only on Γ. Every pair of points on a smooth compact Riemannian man-
ifold with boundary is connected by a C^1-smooth distance minimizing curve [1]. We choose
our a manifold to be the horseshoe-shaped domain of figure 1.

Because *P* is an involute of the boundary as in the figure, $d(x,p) = d(x,q)$ for any $x \in \Gamma$ and
$p,q \in P$. Therefore from the point of view of our data, the set *P* appears to collapse to a point.

Worse, if Γ happens to be an involute as well, then all the distance functions to points
$p \in M_2$ are constants in Γ. As the unknown origin times produce unknown constant offsets
to the boundary distances, all points in M_2 look alike when seen from an involutive Γ in this
sense.

The problems of figure 1 also illustrate the problems lack of convexity of the boundary may
cause.

Similar problems may arise in higher dimensions if data is recorded on too small sets. Let
the manifold *M* be the closed unit ball in $\mathbb{R}^n$ and consider partial data on $\Gamma = \partial M \cap H$, where
$H \subset \mathbb{R}^n$ is a hyperplane through the origin. Let p_1 and p_2 be two points in $\text{int}(M) \setminus H$ situated
symmetrically about *H*. Then the boundary distance functions of p_1 and p_2 agree on Γ. The
sets *M* and $\Gamma \subset \partial M$ have a reflection symmetry which leaves the data invariant, making it
impossible to distinguish the two sides.

3.2.3. Infinitely many lentils.   The crucial condition for the estimates in theorem 9 was that
all of the lentils meet the spatial source set *P*. There is an infinite family of lentils, so we have
a large number of conditions to check from our data. The lentils are open and cover a certain
compact subset of *M* (excluding a layer near the boundary), so in fact using a finite number

will suffice. However, it is not easy to identify a covering collection of lentils—for which one would then easily check whether they contain source points—from data.

3.2.4. Conjugate points and boundary sources. If there is a source point on the boundary, the graph of the distance function is singular at that point. The separation of the set $Q(S)$ into graphs in proposition 3 relies on smoothness. If several corners happen to coincide, especially if $n = 2$, it may be difficult to choose how to continue the graphs correctly.

This problem becomes far worse if there are conjugate points. They also cause the boundary distance function to be non-smooth but can do so for several different points. This makes both disentangling the data into graphs and the analysis of those graphs substantially more complicated.

The condition (1) on the constants of definition 6 is assumed in theorems 1, 9–11. This condition is used in lemma 24 and proposition 30 when the manifold is compared to model manifold with constant sectional curvature. For the comparison to work, we need the model manifold of the same diameter to not have conjugate points, and this is exactly what the condition ensures.

The condition (1) always holds in negative curvature. If the manifold is known to have non-positive sectional curvature and an explicit bound on the diameter, one can simply choose the sectional curvature upper bound $\mathscr{C}_{\mathrm{sec}+} > 0$ to be small enough to satisfy (1). The condition also holds on all simple manifolds of constant sectional curvature.

The constants C_{a} of (37), C_{b} of (38), and C_{c} of (39), as well as all constants derived from them, become worse when $\mathscr{C}_{\mathrm{diam}}\sqrt{\mathscr{C}_{\mathrm{sec}+}}$ gets close to π.

3.2.5. Constants of quantitative simplicity. The constants of quantitative simplicity have to be the same for both manifolds M_i in theorem 9. The set Γ_i and the approximate boundary inclusion map α depend on these constants. The distances between the source points are determined irrespective of the constants by theorem 2, but the constants have an effect on the approximate boundary structure and the estimates on the quality of the reconstruction.

3.2.6. Scaling of small quantities. Consider the case when all the various epsilons are very small. In the setting of theorem 9 we have $\delta \approx \varepsilon_2 \approx \varepsilon^2$ and in theorem 10 we have $\delta \approx \varepsilon_2 \lesssim \varepsilon^2 \approx \hat{\varepsilon}$. Ideally, we would expect to see no second powers so that all concepts of density—the size of lentils, the true density, the observed quantity ε_2, the density estimated from data—to be bi-Lipschitz equivalent to each other. The difference in scaling is all due to the lentils being shaped so that

$$\text{outer radius} \approx \sqrt{\text{inner radius.}}$$

This scaling is easy to verify in Euclidean geometry, and the relevant Riemannian aspects are covered in lemmas 24 and 25, and proposition 30.

3.2.7. Generalization to simple Berwald manifolds. As discussed above, our methods depend on the smoothness of the distance function. Thus these techniques are not applicable for general Riemannian metrics and relaxing the simplicity assumption would likely require different methods. There is however hope to apply these techniques to study analogous questions in the case of certain simple Finsler metrics.

Any constant speed geodesic γ of a smooth connected and complete Finsler manifold (N, F) is given by the solution of the geodesic equations, in local coordinates, $\ddot{\gamma}^k(t) + \Gamma_{ij}^k(\gamma, \dot{\gamma})\dot{\gamma}^i(t)\dot{\gamma}^j(t) = 0$. Here $\Gamma_{ij}^k(x, v)$ are the coefficients of the *Chern connection* on the tangent bundle [3]. Due to the dependence of the directional variable $v \in T_x N$ a reverse curve of a Finsler geodesic does not need to be a geodesic. If the Chern connection coefficients are directionally independent the Finsler metric F is called *Berwald*. Thus Berwald manifolds have a well defined canonical covariant differential operator for tensor fields of any order. Therefore Berwald metrics carry a Gauss type of formula to connect boundary and interior Hessians for the distance function. This is not true for a general Finsler metrics and the lack of Gauss formula would cause issues in several parts of our proofs. It is worth of mentioning that Berwald metrics are not just a theoretical curiosity but have a connection to linear elasticity. It has been observed that elastic Finsler metrics, arising from transverse isotropic medium in weak anisotropy, are actually Berwaldian [60].

4. Plan of proofs

Our data is given in the form of a set, a union of graphs, and we begin by disentangling it into the separate graphs. We prove proposition 3 to this effect in section 5.

In section 6 we turn to finding distances between source points and other information that the data gives exactly. With the help of proposition 3, we may start with the knowledge of the arrival time functions a_s. Each graph corresponds to a unique source point, giving us the first claim. Computing suitable differences of arrival times, especially along the unique geodesic between two given source points, we cancel out unwanted contributions and find the distance between the two points. Once we have these first two claims, the rest of theorem 2 follows straightforwardly.

In section 7 we depart from differential geometry to the geometry of metric spaces and study the labeled Gromov–Hausdorff distance in more detail, including the proof of proposition 5.

The most substantial part of our proofs is in estimating the density of the sources, and we will do this in section 9 using the explicit bounds on geometric quantities found in section 8 — including the proof of proposition 7.

The first task is to estimate the density of sources near the boundary, which practically amounts to estimating how near the boundary a given source point is. This is done in terms of the curvature of the graph of the arrival time function; when the point is very close to the boundary, the Hessian of the said function blows up. We will prove that $d(p, x) \leqslant E(p, x)$ for any source point $p \in P$ and a critical point $x \in \partial M$ of its arrival time function. This gives an estimate in terms of data for how close to any given boundary point must there be a source point. The initial data is first processed into the form of $E(p, x)$ and quantities derived from it, and we verify in proposition 8 that the necessary auxiliary quantities are indeed determined by our data.

To get a density estimate deep inside the manifold, a different set of tools must be employed. A key tool is the lentil, an intersection of two metric balls. Whether a source point belongs in a lentil defined by two other source points is entirely decided by the distances between the three points, and by theorem 2 this information is indeed determined by the data. We will show that with a suitable choice of $\delta > 0$ (as given in theorem 9) the lentils cover the deep interior of the manifold but are small. This requires estimates in two directions for the lentils so that they are large enough to cover the relevant subset of M without gaps but are small enough so that they provide a useful density estimate. If every point is contained in a lentil and that lentil contains

a source point, then the density of sources is bounded by the diameters of the lentils. The more complicated side of lentils is to ensure that they cover enough of the manifold, and this relies on a number of estimates based on bounded geometry. These density estimates constitute a proof of theorem 9.

Density estimates turn out to be simpler in the reverse direction, showing that the methods used in the proof of theorem 9 are fine enough to give a decent estimate of density. These reverse estimates prove theorem 10.

In section 10 we turn our attention to matters of convergence. The proof of theorem 11 is mostly based on theorems 9 and 10; as T increases, the source points become denser and denser both in reality and in terms of data-driven estimates, and thus the quality of the discrete approximation improves.

Theorem 1 follows a similar path and can be regarded essentially as a corollary of theorem 11, and thus it has a relatively simple proof now that all the tools are available. The metric convergence results show that the two manifolds of theorem 1 are isometric as metric spaces, and by the Myers–Steenrod theorem the isometry is in fact smooth. Proposition 12 gets its proof, using similar ideas and basic properties of Poisson point processes, at the very end.

5. Separation of graphs on a bundle

Let us denote the scalar second fundamental form by h_2 and the boundary metric (the first fundamental form) by h_1. If $p \in \text{int}(M)$, $\rho \colon M \to \mathbb{R}$ is the distance function $\rho(x) = d(p,x)$ at p then the corresponding boundary distance function $r \colon \partial M \to \mathbb{R}$, is just the restriction of ρ on the boundary ∂M. By the Gauss formula for hypersurfaces (see e.g. [40, theorem 8.13 a]) and the definition of Hessians we have that

$$\mathcal{H}_r(x) = \mathcal{H}_\rho(x) + h_2(x). \tag{14}$$

We will make use of this throughout the paper.

Our method of proving proposition 3 is by means of lifting the data from the boundary $\partial M \times \mathbb{R}$ of the spacetime to a suitable bundle where the graphs do not intersect. We start by showing that when $s \neq s'$, then the graphs of a_s and $a_{s'}$ can only have up to first order tangency when they intersect. Therefore 2-jets will separate them.

Lemma 13. *Let M be a simple Riemannian manifold and $s, s' \in \text{int}(M) \times \mathbb{R}$ be two distinct points in the spacetime. Then if $a_s(x) = a_{s'}(x)$ at some point $x \in \partial M$, then the gradient or the Hessian of the two functions will differ at x.*

Proof. Take any two $s, s' \in S$ so that a_s and $a_{s'}$, their differentials, and their Hessians agree at a point $x \in \partial M$. We aim to show that $s = s'$.

First, observe that the differential $da_s(x)$ is the covector corresponding to the tangential component of the velocity of the unit speed geodesic from $\pi(s)$ to x. Therefore $da_s(x) = da_{s'}(x)$ implies that $p := \pi(s)$ and $p' := \pi(s')$ are on the same geodesic starting from x.

Let γ be this unit speed geodesic starting at x. We have $\gamma(t_0) = p$ and $\gamma(t_0') = p'$ for some $t_0, t_0' > 0$.

Consider the Hessian $\mathcal{H} := \mathcal{H}_{\rho_p}$ of the interior distance function $\rho_p \colon M \to \mathbb{R}$ defined by $\rho_s(x) = d(x,p)$ and its counterpart $\mathcal{H}'$ corresponding to p'.

If a normal Jacobi field J along γ satisfies $\mathcal{H}J(t) = -D_t J(t)$ for some $t < t_0$, then $J(t_0) = 0$, (see for instance [40, proposition 11.2]). Similarly, $\mathcal{H}'J(t) = -D_t J(t)$ for any $t < t_0'$ implies $J(t_0') = 0$.

Both Hessians satisfy $\mathcal{H}^{(\prime)}\dot{\gamma}(t) = -\dot{\gamma}(t)$ for $t < \min(t_0, t_0')$. As the geodesic meets ∂M transversally at x and the boundary Hessians of a_s and $a_{s'}$ agree there, we have in fact the equality $\mathcal{H}(x) = \mathcal{H}'(x)$ for the Hessian operators on $T_x M$. This follows from (14).

Now take any nonzero $w \in T_x M$ normal to $\dot{\gamma}(0)$ and let J be the Jacobi field along γ with $J(0) = w$ and $D_t J(0) = \mathcal{H}w = \mathcal{H}'w$. As observed above, this implies $J(t_0) = 0$ and $J(t_0') = 0$. Due to the lack of conjugate points $p = \gamma(t_0) = \gamma(t_0') = p'$.

As the two source points are equal distance $t_0 = t_0'$ from x and the two arrivals are at the same time, we also have $\tau(s) = \tau(s')$. This concludes the proof of $s = s'$. $\qquad\square$

Proof of proposition 3. Let us abbreviate $Q := Q(S) \subset \partial M \times \mathbb{R}$. As $S \subset M \times \mathbb{R}$ is discrete and M is compact, there are only finitely many points of S in any interval $M \times [a, b]$. By simplicity there is no geodesic longer than the diameter of the manifold. Thus Q written as

$$Q = \bigcup_{s \in S} \mathcal{G}(a_s)$$

is a locally finite union on $\partial M \times \mathbb{R}$.

Let $\sigma(Q)$ be the set of 'smooth points of Q', where Q is given locally as a single graph. By definition $\sigma(Q) \subset Q$ is open. It is also dense, for otherwise there would be two graphs that coincide in an open set, contradicting lemma 13.

Let E be the bundle over $\partial M \times \mathbb{R}$ with fibers

$$E_{(x,t)} = T_x \partial M \times T_x^{\otimes 2} \partial M.$$

This E is the product of the bundles $T_x \partial M$ and $T_x \partial M \otimes T_x \partial M$ pulled to $\partial M \times \mathbb{R}$ over the projection $\partial M \times \mathbb{R} \to \partial M$.

For any smooth function $h \colon U \to \mathbb{R}$ defined in an open set $U \subset \partial M$ we call the lift of its graph the subset $L\mathcal{G}(h) \subset E$ so that

$$L_{x,h(x)}\mathcal{G}(h) = (\nabla h(x), \nabla^2 h(x)) \in T_x \partial M \times T_x^{\otimes 2} \partial M.$$

This defines a smooth submanifold of E.

In this way we lift all of $\sigma(Q)$ into a submanifold $L\sigma(Q)$ of E. Since $\sigma(Q) \subset Q$ is dense, we have

$$\overline{L\sigma(Q)} = \bigcup_{s \in S} L\mathcal{G}(a_s). \tag{15}$$

There are no second order intersections by lemma 13, so the smooth submanifolds $L(\mathcal{G}(a_s)) \subset E$ are pairwise disjoint.

Therefore Q determines the set

$$\{L\mathcal{G}(a_s) \subset E; \, s \in S\}.$$

Projecting from E down to $\partial M \times \mathbb{R}$ gives the graphs $\mathcal{G}(a_s)$ and thus also the functions a_s, proving the first part of the claim.

For the second claim we simply localize (15) to be over $\Omega \subset \partial M \times \mathbb{R}$, and we get a disjoint union of the sets $L[\mathcal{G}(a_s) \cap \Omega]$. A single graph $\mathcal{G}(a_s)$ may be cut into several pieces by this procedure and we do not necessarily know which pieces correspond to the same source. Therefore what we obtain is the collection of connected components of graphs restricted to Ω. $\qquad\square$

It was proven in [31] that the boundary distance function $r_x \colon \partial M \to \mathbb{R}$ of $x \in \mathrm{int}(M)$ determines the point x uniquely. The next lemma improves this slightly: the boundary distance function modulo constants is enough for this uniqueness.

Lemma 14. *Let M be a simple Riemannian manifold and $x, y \in \mathrm{int}(M)$ any two distinct points. Denote by r_x and r_y the boundary distance functions $\partial M \to \mathbb{R}$ from these points. The difference function $r_x - r_y$ cannot be a constant function.*

Proof. Suppose that the difference function $r_x - r_y$ takes the constant value $c \in \mathbb{R}$ despite $x \neq y$. If $c = 0$, the boundary distance functions coincide and thus by [31] $x = y$, yields a contradiction. Therefore we may suppose that $c \neq 0$.

Let γ be the maximal geodesic through the points x and y with endpoints $x', y' \in \partial M$ ordered so that x' is closer to x and y' to y. By assumption we have

$$d(x', x) = r_x(x') = r_y(x') + c = d(x', y) + c.$$

Due to simplicity and the order of the points, we have $d(x', y) = d(x', x) + d(x, y)$. Therefore $0 = d(x, y) + c$. The same calculation with the roles of x and y reversed shows that $0 = d(x, y) - c$. These two together imply that $d(x, y) = c = 0$, which is a contradiction.　　□

6. Exact observables

Proof of theorem 2. Due to proposition 3 the data determines the arrival time functions $a_s(x) = d(\pi(s), x) + \tau(s)$ at all points $x \in \partial M$ for every $s \in S$. It also follows from the same proposition that distinct sources have distinct arrival time functions. Therefore now that the two manifolds M_1 and M_2 and their source sets S_1 and S_2 have the same data modulo identification by ϕ, we have a bijection between the source sets given by identifying the corresponding arrival time functions which coincide. This proves claim (1).

We will prove all subsequent claims by describing how the arrival time functions determine the quantities in question. As these functions coincide on the two manifolds, the reconstructed quantities coincide. We therefore drop the subscripts and work with a single manifold. We can think of the source points being indexed by their arrival time graphs. To simplify notation, we denote the source points as $p_s = \pi(s)$ and source times as $t_s = \tau(s)$.

The arrival time functions now have the form $a_s(x) = d(p_s, x) + t_s$. Two source points p_s and p_r coincide if and only if $a_s - a_r$ is a constant function—this is a straightforward corollary of lemmas 13 and 14. We pass to a subset of S so that each source point p_s is only present once; the claims extend easily to the duplicated source points.

For any $s, r \in S$ we define the functions $f_{rs} \colon \partial M \times \partial M \to \mathbb{R}$ by

$$f_{rs}(x, y) = a_r(x) - a_s(y).$$

These functions are determined by the data. Taking $s = r$ proves claim (5).

From now on we suppose that $s \neq r$. The differentials $\mathrm{d}a_s$ and $\mathrm{d}a_r$ agree at $x \in \partial M$ if and only if the geodesics joining x to p_s and p_r start in the same direction at x. Therefore the differentials agree at exactly two points, the endpoints of the unique maximal geodesic through the points p_s and p_r. Let these boundary points be x and y.

Depending on how the four points (p_s, p_r, x, y) are ordered on the geodesic, we have

$$\begin{aligned}
f_{rr}(x, y) - f_{ss}(x, y) &= d(p_r, x) - d(p_s, x) + d(p_s, y) - d(p_r, y) \\
&= \pm 2 d(p_r, p_s).
\end{aligned} \tag{16}$$

Thus claim (2) is given by

$$d(p_r, p_s) = \frac{1}{2} \left| f_{rr}(x, y) - f_{ss}(x, y) \right|.$$

This determination of distances between source points will play a central role in the proofs of the other theorems.

By switching the points x and y if needed (depending on the sign in (16)), we can assume to have $d(p_r, y) = d(p_r, p_s) + d(p_s, y)$. Then for any $z \in \partial M$, we have

$$f_{rr}(z, y) - f_{ss}(z, y) + d(p_r, p_s) = d(p_r, z) - d(p_s, z).$$

As everything on the left-hand side is determined by the data, so is then the right-hand side, and we have obtained the function $f^{rs} \colon \partial M \to \mathbb{R}$ given by

$$f^{rs}(z) = d(p_r, z) - d(p_s, z)$$

for all $s, r \in S$. This proves claim (4).

Finally, we note that for any $z \in \partial M$ we have

$$t_r - t_s = f_{rs}(z, z) - f^{rs}(z)$$

and claim (3) follows. $\qquad\square$

As mentioned, the proof of theorem 1 will be postponed until we have approximation tools available.

7. The labeled Gromov–Hausdorff distance

Our proof of the basic properties of the labeled Gromov–Hausdorff distance is similar to well known proofs of the usual Gromov–Hausdorff distance (see for instance [20, chapter 3] or [6, section 7.3]) but with the labels interwoven into it. For the sake of completeness we record our proof in its entirety, although the aspects unrelated to labels are indeed well known.

Proof of proposition 5. The symmetry of the labeled Gromov–Hausdorff distance is evident. The triangle inequality can be verified as in [6, proposition 7.3.16]. Also, if there is h as described, then choosing $Z = Y$, $f = h$, and $g = \mathrm{id}$ shows that the distance is zero. The only nontrivial claim is that if the distance is zero, then such an h exists. Let us prove that.

Take any compact metric space Z and isometric embeddings $f \colon X \to Z$ and $g \colon Y \to Z$. We can define a semimetric (all properties of a metric but $d(x, y) = 0$ need not imply $x = y$) d on the disjoint union $W = X \sqcup Y$ by letting

$$d(a, b) = \begin{cases} d_Z(f(a), f(b)), & \text{when } a \in X \text{ and } b \in X, \\ d_Z(f(a), g(b)), & \text{when } a \in X \text{ and } b \in Y, \\ d_Z(g(a), f(b)), & \text{when } a \in Y \text{ and } b \in X, \\ d_Z(g(a), g(b)), & \text{when } a \in Y \text{ and } b \in Y. \end{cases}$$

The embeddings are isometric, so $d_Z(f(a), f(b)) = d_X(a, b)$ and similarly for Y and g.

Take any $k \in \mathbb{N}$. By assumption we have a metric space Z_k and thus a semimetric d_k on W so that $d_H^{Z_k}(X, Y) < \frac{1}{k}$ and $d_k(\alpha(\ell), \beta(\ell)) < \frac{1}{k}$ for all $\ell \in L$. These semimetrics agree with the metrics on X and Y when restricted to either set.

By compactness there is a finite set $A_k' \subset X$ so that the semiballs

$$\left\{ p \in X; \, d_k(x, p) < \frac{1}{k} \right\}_{x \in A_k'}$$

cover X. Similarly, there is a finite $A_k'' \subset X$ so that the semiballs

$$\left\{ p \in Y; d_k(x,p) < \frac{1}{k} \right\}_{x \in A_k''}$$

cover Y. Similarly, there are finite sets B_k' and B_k'' in Y so that the semiballs of radius $\frac{1}{k}$ cover Y and X. We define $A_k = A_k' \cup A_k''$ and $B_k = B_k' \cup B_k''$. We can make all of these choices so that $A_k \subset A_{k+1}$ and $B_k \subset B_{k+1}$.

For any $k \in \mathbb{N}$ on the finite set $A_k \sqcup B_k$ the sequence of semimetrics $(d_j)_{j=k}^{\infty}$ is bounded pointwise, since $d_j|_{A_k \times A_k} = d_{j'}|_{A_k \times A_k}$ and $d_j|_{B_k \times B_k} = d_{j'}|_{B_k \times B_k}$ for $j, j' \geqslant k$. For $x \in A_k$ and $y \in B_k$ the assumption $d_H^{Z_k}(X,Y) < \frac{1}{k}$ yields

$$d_j(x,y) \leqslant \max(\operatorname{diam}(X), \operatorname{diam}(Y)) + \frac{2}{k}$$

for all $j \geqslant k$. Thus $(d_j)_{j=k}^{\infty}$ has a converging subsequence and the pointwise limit of semimetrics is a semimetric on $A_k \sqcup B_k$. Constructing a diagonal sequence gives us a subsequence of (d_j) which converges pointwise on $W' = \bigcup_{k \in \mathbb{N}} (A_k \sqcup B_k)$ to a semimetric δ. Because each d_k agrees with the original metrics on X and Y, the only thing to inspect are the 'cross-distances' between X and Y.

We can extend δ to all of W as follows. When x and y are both in X or both in Y, we use the metrics on these spaces. When $x \in X$ and $y \in Y$, we pick for each $k \in \mathbb{N}$ points $x_k \in B^{d_X}(x, \frac{1}{k}) \cap A_k \subset X$ and $y_k \in B^{d_Y}(y, \frac{1}{k}) \cap B_k \subset Y$, where we have indicated the metric defining the balls as a superscript. We then let

$$\delta(x,y) = \lim_{k \to \infty} \delta(x_k, y_k).$$

A simple argument shows that this limit is independent of the choice of the approximating sequences. It is also straightforward to check that the extended δ is indeed a semimetric on W.

We want to show that for every $x \in X$ there is a unique $y \in Y$ so that $\delta(x,y) = 0$. The triangle inequality and δ agreeing with the metric on Y shows that the point is unique. For existence, there is a point in $x_k \in A_k$ that is $\frac{1}{k}$-close to x, and there is $y_k \in B_k$ that is $\frac{1}{k}$-close to x_k. We can then set $y = \lim_{k \to \infty} y_k$. Similarly, each $y \in Y$ has a unique $x \in X$ so that $\delta(x,y) = 0$.

This gives rise to a bijection $h: X \to Y$ that satisfies $\delta(x, h(x)) = 0$. This can be checked to be an isometry.

Take any $\ell \in L$ and $k \in \mathbb{N}$. We have $d_k(\alpha(\ell), \beta(\ell)) < \frac{1}{k}$ by construction of d_k, so in the limit of the subsequence that we got we find $\delta(\alpha(\ell), \beta(\ell)) = 0$. This means that $h(\alpha(\ell)) = \beta(\ell)$. Therefore $h \circ \alpha = \beta$. $\qquad\square$

We record two additional propositions. The proofs are immediate and we omit them.

Proposition 15. *Let X and Y be compact metric spaces and L any set. Let $\alpha: L \to X$ and $\beta: L \to Y$ be any two functions. If $L' \subset L$, then*

$$d_{GH}^{L'}(X, \alpha|_{L'}; Y, \beta|_{L'}) \leqslant d_{GH}^{L}(X, \alpha; Y, \beta).$$

As $d_{GH}^{\emptyset}(X, \emptyset; Y, \emptyset)$ is the usual Gromov–Hausdorff distance between X and Y, the labeled kind of convergence implies the usual kind of convergence.

Proposition 16. *Let X be a compact metric space and $Y \subset X$ a subset. Let $\alpha: L \to X$ and $\beta: L \to Y$ be any two functions on a set L. If $Y \subset X$ is ε_1-dense and $\sup_{\ell \in L} |d(\alpha(\ell), \beta(\ell))| \leqslant \varepsilon_2$, then*

$$d_{GH}^{L}(X, \alpha; Y, \beta) \leqslant \varepsilon_1 + \varepsilon_2.$$

8. Bounded geometry

We begin our study of bounded geometry by proving proposition 7 with the help of two lemmas, and then we move on to finding further estimates based on these basic bounds.

Lemma 17. *Let M be a simple manifold, let $U_x \subset T_xM$ be the maximal domain of definition of the exponential map, and let $U \subset TM$ be the subset with fibers U_x. The map $\exp \times \pi \colon U \to M \times M$ that maps*

$$U \ni (x,v) \mapsto (\exp_x(v), x) \in M \times M$$

is a diffeomorphism and so has a smooth inverse $\theta \colon M \times M \to U \subset TM$.

Proof. As each $\exp_x \colon U_x \to M$ is a diffeomorphism on a simple manifold, the map $\exp \times \pi$ is clearly smooth and bijective. What remains to check is the invertibility of the differential at every point $(x,v) \in U$. The differential has a convenient block structure due to $\pi(x,v) = x$ being independent of v, and so $\mathrm{d}(\exp \times \pi)$ is invertible at (x,v) if and only if $\mathrm{d}\exp_x$ is invertible at v. The invertibility of $\mathrm{d}\exp_x$ is true by assumption. □

Lemma 18. *Let $W^1, \ldots, W^n$ be vector fields on a simple manifold M constituting a global orthonormal frame. Let $\Gamma^i{}_{jk}(x,y)$ be the Christoffel symbol at $x \in M$ written in the normal coordinates centered at $y \in M$. Let $W^1(y), \ldots, W^n(y)$ be the coordinate basis for T_yM. Then*

$$\sup_{x,y \in M} \left| \Gamma^i{}_{jk}(x,y) \right| < \infty$$

for all indices i,j,k.

There is a global orthonormal frame on every simple manifold given by vector fields W_i as stated. Such a frame can be produced by the Gram–Schmidt method from a general frame coming from the trivializability of the tangent bundle.

Proof of lemma 18. The vectors $W_1(y), \ldots, W_n(y)$ constitute an orthonormal basis for T_yM. With the help of lemma 17 we see that the basis vectors for T_xM in the normal coordinates about y are given by

$$w_i(x,y) := \mathrm{d}^{[x]} \exp_y(\theta(x,y)) W_i(y).$$

Here and later in this proof we indicate the variable of differentiation by superscript in [square brackets] when needed for clarity.

The invariant coordinate map for the normal coordinates of y is $\theta_y \colon M \to T_yM$ given by $\theta_y(x) = \theta(x,y)$. Using the basis given by the frame, we get a proper coordinate map $\overline{\theta_y} \colon M \to \mathbb{R}^n$ given by $\overline{\theta_y}^i = W_i(y)^\flat \circ \theta_y$, which means

$$\overline{\theta_y}(x) = (\langle W_1(y), \theta_y(x) \rangle, \ldots, \langle W_n(y), \theta_y(x) \rangle).$$

Now we have a concrete description of the normal coordinates in terms of the frame.

Let us denote $w^i(x,y) = \mathrm{d}^{[x]} \overline{\theta_y}^i(x) \in T_x^*M$. This is the differential of a coordinate map and thus a basis covector for T_x^*M induced by the normal coordinates about y. The corresponding basis vectors on T_xM are $w_i(x,y)$. In terms of the coordinate map $z = \overline{\theta_y} \colon M \to \mathbb{R}^n$ the Christoffel symbols are given by

$$\begin{aligned}
\Gamma^i{}_{jk}(x,y) &= \mathrm{d}z^i \left(\nabla_{\partial/\partial z^j} \frac{\partial}{\partial z^k} \right) \\
&= w^i(x,y) (\nabla^{[x]}_{w_j(x,y)} w_k(x,y)).
\end{aligned}$$

All the items in this formula depend smoothly on x and y, up to the boundary. The claim then follows from compactness of $M \times M$.

$\square$

Proof of proposition 7. Parts (1), (2), (5), and (6) of definition 6 follow easily from compactness. We will prove the other parts more carefully.

Part (3): let us denote $y_i = \exp_x(\eta_i) \in M$. In terms of lemma 18 we have $\theta(y_i, x) = \eta_i$. Let us equip TM with the Sasaki metric and $M \times M$ with the product metric. An important property of the Sasaki metric on TM is that it respects the natural distance on the fibers: the Sasaki distance between $v, w \in T_x M$ is simply $|v - w|$ because the fibers are totally geodesic [49]. Because θ is a diffeomorphism between compact manifolds with boundary, it is Lipschitz-continuous. The Lipschitz constant is denoted by $\mathscr{C}_{\exp}$, that is the maximum value of the norm of the Jacobian of θ.

Part (4): let us decompose an arbitrary Jacobi field J starting with $J(0) = 0$ into parallel and normal components relative to $\dot{\gamma}$:

$$J(t) = J^{\parallel}(t) + J^{\perp}(t).$$

As $|J(t)|^2 = \left|J^{\parallel}(t)\right|^2 + \left|J^{\perp}(t)\right|^2$ and both components are Jacobi fields, it suffices to prove the estimate separately for parallel and normal Jacobi fields.

If J is tangential to γ then $J(t) = tc\dot{\gamma}(t)$ for some $c \in \mathbb{R}$ and $|J(t)|^2 = c^2 t^2$. Thus for the tangential Jacobi fields the claim holds for any constant $\mathscr{C}_{\mathrm{JF}} \geqslant 2$.

Now we assume that J is normal to $\dot{\gamma}$. Let $\rho(x) = d(x, \gamma(0))$. Then

$$\partial_t |J(t)|^2 = 2 \langle D_t J, J \rangle \leqslant 2 |D_t J| |J| = 2 |\mathcal{H}_\rho J| |J| \leqslant 2 |\lambda_{\max}(t)| |J|^2,$$

where $\lambda_{\max}(t)$ is the eigenvalue of the Hessian $\mathcal{H}_\rho$ of ρ at $\gamma(t)$ with the largest absolute value.

Let us choose normal coordinates at $\gamma(0)$ as in lemma 18. In these coordinates we have $\rho(x) = |x|_e$ for all $x \in M$, where the subscript e stands for the Euclidean norm. Geodesics with initial unit speed $v \in S_{\gamma(0)} M$ are given by $\gamma(t) = tv$.

Therefore the Hessian of the distance function ρ is given in these coordinates by

$$\mathcal{H}_\rho(x) = \frac{1}{|x|_e} A_{ij} \mathrm{d}x^i \otimes \mathrm{d}x^j, \tag{17}$$

where

$$A_{ij}(x, \gamma(0)) = \delta_{ij} - \frac{x_i x_j}{|x|_e^2} - \Gamma^k_{ij}(x, \gamma(0)) x_k. \tag{18}$$

The numbers A_{ij} are all bounded uniformly for all $\gamma(0) \in M$ and $x \in M \setminus \{\gamma(0)\}$ due to lemma 18. Thus the Hessians $\mathcal{H}_\rho$, as quadratic forms, are uniformly bounded by some constant $\frac{1}{2}\mathscr{C}_{\mathrm{JF}}$, and $|\lambda_{\max}(t)| \leqslant \frac{1}{2}\mathscr{C}_{\mathrm{JF}}$. The claim follows.

Part (7): this is similar to part (4) and we continue with the same proof setup. The point where everything is evaluated is $x = tv$, so by equations (17) and (18) we have for all $w \in \dot{\gamma}(t)^{\perp}$

$$\langle w, \mathcal{H}_\rho w \rangle = t^{-1} [\delta_{ij} w^i w^j - \Gamma^k_{ij}(tv) t v_k w^i w^j].$$

The bound on the Christoffel symbols implies that $\left|\Gamma^k_{ij}(tv)\right| \leqslant C$ for some absolute constant $C \geqslant 0$ independent of t, v, and $\gamma(0)$. In the Euclidean norm $\left|v^{\flat}\right|_e = 1$, so $|v_k| \leqslant 1$. Thus we find

$$\langle w, \mathcal{H}_\rho w \rangle \geqslant t^{-1} |w|_e^2 - C |w|_e^2$$

for all $w \in \dot{\gamma}(t)^\perp$. By the diffeomorphic nature of $\exp \times \pi$ of the proof of lemma 18 the Euclidean norm of w is bi-Lipschitz equivalent with the Riemannian norm of w, with bi-Lipschitz constants independent of x, and $\gamma(0)$. This ensures that the desired two constants exist. $\qquad\square$

We will next establish a number of estimates based on the results of proposition 7. Most of them we will rewrite in section 9.1 in the language of lentils. Before getting started, we recall three standard comparison results, stated in a form suitable for our use.

We will only use lemma 19 to compare to constant sectional curvature, and we state the next two lemmas only in this setting. We denote by M_κ the complete Riemannian manifold with constant sectional curvature $\kappa \in \mathbb{R}$, either the hyperbolic space, the Euclidean space, or the sphere, depending on the sign of κ.

Lemma 19 (Rauch's comparison theorem [16, chapter 10, theorem 2.3]). *Let M be a simple Riemannian manifold with sectional curvature $\geqslant \kappa$. Let J and $\tilde{J}$ be Jacobi fields on unit speed geodesics γ and $\tilde{\gamma}$ on the manifolds M and M_κ, with $J(0) = 0 = \tilde{J}(0)$, $\langle D_t J(0), \dot{\gamma}(0) \rangle = \langle D_t \tilde{J}(0), \dot{\tilde{\gamma}}(0) \rangle$ and $|D_t J(0)| = |D_t \tilde{J}(0)|$. Then*

$$|J(t)| \leqslant |\tilde{J}(t)|$$

for all $t > 0$ for which both geodesics are defined.

Consider a simple Riemannian manifold M with sectional curvature bounded from above by $\mathscr{C}_{\mathrm{sec}+}$. For any three points $x, y, z \in M$ we want to pick three points $\tilde{x}, \tilde{y}, \tilde{z} \in M_{\mathscr{C}_{\mathrm{sec}+}}$ so that

$$d(x,y) = d(\tilde{x}, \tilde{y}), \quad d(x,z) = d(\tilde{x}, \tilde{z}), \quad d(y,z) = d(\tilde{y}, \tilde{z})$$

or

$$d(x,y) = d(\tilde{x}, \tilde{y}), \quad d(x,z) = d(\tilde{x}, \tilde{z}), \quad \angle yxz = \angle \tilde{y}\tilde{x}\tilde{z}.$$

We denote distances on both manifolds by d, as there should be no confusion as to which space each point belongs. As long as M satisfies the estimate (1), such points on $M_{\mathscr{C}_{\mathrm{sec}+}}$ exist.

The following two lemmas compare triangles on M with corresponding triangles on M_κ. They both follow from Rauch's comparison theorem, although not exactly the same version as lemma 19 above. We recall from definition 6 that the sectional curvature is bounded from above by the constant $\mathscr{C}_{\mathrm{sec}+} > 0$.

Lemma 20 ([16, chapter 10, proposition 2.5]). *Let M be a simple manifold whose constants (of definition 6 and proposition 7) satisfy the constraint (1). Take any three points $x, y, z \in M$. Then there are points $\tilde{x}, \tilde{y}, \tilde{z} \in M_{\mathscr{C}_{\mathrm{sec}+}}$ so that $d(x,y) = d(\tilde{x}, \tilde{y})$, $d(x,z) = d(\tilde{x}, \tilde{z})$, and the angle is the same at x and $\tilde{x}$. The remaining distances satisfy $d(y,z) \leqslant d(\tilde{y}, \tilde{z})$.*

Finally we give a comparison result for the angles of geodesic triangles.

Lemma 21 ([30, theorem 2.7.6]). *Let M be a simple manifold whose constants (of definition 6 and proposition 7) satisfy the constraint (1). Take any three points $x, y, z \in M$. Then there are points $\tilde{x}, \tilde{y}, \tilde{z} \in M_{\mathscr{C}_{\mathrm{sec}+}}$ so that $d(x,y) = d(\tilde{x}, \tilde{y})$, $d(x,z) = d(\tilde{x}, \tilde{z})$, and $d(y,z) = d(\tilde{y}, \tilde{z})$, and for the respective geodesic triangles the angle at x is at most that at $\tilde{x}$.*

We will then proceed to estimates specific to our setting.

Lemma 22. *Let M be a simple Riemannian manifold of bounded geometry. Let J be a Jacobi field along any constant (not necessarily unit) speed geodesic with $J(0) = 0$. Then*

$$|J(t)| \leqslant C_{\mathrm{b}} |D_t J(0)| t$$

for all t > 0, where the constant is given by (38) in appendix B.

Proof. Let us first prove the claim for unit speed geodesics.

In constant sectional curvature $k = -\mathscr{C}_{\mathrm{sec}-} < 0$ every Jacobi field J with the initial conditions $J(0) = 0$ and $|D_t J(0)| = 1$ is of the form $J(t) = A(-k)^{-1/2} \sinh(t\sqrt{-k})$ with a parallel unit vector A; see e.g. [40, proposition 10.12].

If M has sectional curvature $\geqslant k$, then by lemma 19 every Jacobi field J with the initial conditions $J(0) = 0$ and $|D_t J(0)| = 1$ satisfies $|J(t)| \leqslant (-k)^{-1/2} \sinh(t\sqrt{-k})$ before a conjugate point. As M is simple, there are no conjugate points.

The function $x \mapsto x^{-1} \sinh(x)$ is increasing for $x \in [0, \infty)$ and $\mathrm{diam}(M) \leqslant \mathscr{C}_{\mathrm{diam}}$. Thus we get the desired estimate with the constant

$$\sup_{t \in (0, \mathscr{C}_{\mathrm{diam}})} t^{-1}(-k)^{-1/2} \sinh(t\sqrt{-k}) \leqslant \frac{\sinh(\mathscr{C}_{\mathrm{diam}}\sqrt{\mathscr{C}_{\mathrm{sec}-}})}{\mathscr{C}_{\mathrm{diam}}\sqrt{\mathscr{C}_{\mathrm{sec}-}}} = C_{\mathrm{b}}$$

as claimed.

Consider then a geodesic $\hat{\gamma}$ with non-unit constant speed $\lambda = \left|\dot{\hat{\gamma}}(0)\right| > 0$. It can be written as $\hat{\gamma}(t) = \gamma(\lambda t)$ for a unit speed geodesic γ.

Let $\Gamma(t,s)$ be a family of geodesics so that $\Gamma(t,0) = \gamma(t)$ and $J(t) = \partial_s \Gamma(t,s)|_{s=0}$ for a Jacobi field J along γ satisfying the assumptions of the claim. Consider then the new Jacobi field $\hat{J}(t') = \partial_{s'} \Gamma(\lambda t', \lambda^{-1} s')|_{s'=0}$ along the geodesic $\hat{\gamma}$. The scaling of the two parameters was chosen so that $\hat{J}$ is a vector field along $\hat{\gamma}$ and $D_{t'}\hat{J}(0) = D_t J(0)$, so that both covariant derivatives have unit norm at $\gamma(0) = \hat{\gamma}(0)$. As we have $\hat{J}(t') = \lambda^{-1} J(\lambda t')$, the already proven claim for J along γ gives

$$\begin{aligned}
\left|\hat{J}(t')\right| &= \lambda^{-1} \left|\hat{J}(\lambda t')\right| \\
&\leqslant \lambda^{-1} C_{\mathrm{b}} \lambda t' \\
&= C_{\mathrm{b}} t'.
\end{aligned}$$

This is the claimed estimate for $\hat{J}$. $\qquad\square$

Lemma 23. *Let $\gamma_1, \gamma_2 \colon [0,1] \to M$ be any two constant (not necessarily unit) speed geodesics with $\gamma_1(0) = \gamma_2(0)$. Then*

$$d(\gamma_1(t), \gamma_2(t)) \leqslant C_{\mathrm{d}} t \, d(\gamma_1(1), \gamma_2(1))$$

for all $t \in (0,1)$, where the constant is given by (48) in appendix B.

Proof. To do so, we write each geodesic as $\gamma_i(t) = \exp_x(t\eta_i)$. Let η_s with $s \in [1,2]$ interpolate linearly between η_1 and η_2.

Then we may estimate the distance between the two required points by taking the *s*-curve with *t* constant. First, the Jacobi field growth estimate of lemma 22 gives

$$|\partial_s \exp_x(t\eta_s)| \leqslant C_{\mathrm{b}} t |\eta_2 - \eta_1|.$$

Therefore

$$d(\gamma_1(t), \gamma_2(t)) \leqslant \int_1^2 |\partial_s \exp_x(t\eta_s)| \, \mathrm{d}s \leqslant t C_{\mathrm{b}} |\eta_2 - \eta_1|. \tag{19}$$

By the uniform invertibility of the exponential maps due to bounded geometry we have

$$|\eta_2 - \eta_1| \leqslant \mathscr{C}_{\exp} d(\exp_x \eta_1, \exp_x \eta_2)$$

and thus (19) turns into

$$d(\gamma_1(t), \gamma_2(t)) \leqslant t C_{\mathrm{b}} \mathscr{C}_{\exp} d(\gamma_1(1), \gamma_2(1)),$$

as claimed. □

9. Density estimates in bounded geometry

9.1. Lentils

We remind the reader that a lentil is defined simply as

$$L_{r,s}^{x,y} = B(x,r) \cap B(y,s).$$

We will always assume that $r, s < d(x,y)$.

We say that the thickness $\delta_{r,s}^{x,y}$ of this lentil is the length of the geodesic segment $\gamma_{x,y} \cap L_{r,s}^{x,y}$, where $\gamma_{x,y}$ is the unit speed geodesic from x to y. Under the assumptions $r, s < d(x,y)$ this leads to

$$\delta_{r,s}^{x,y} = r + s - d(x,y).$$

The assumptions $r, s < d(x,y)$ imply that $\delta_{r,s}^{x,y} < r, s$.

The midpoint of the lentil is

$$m_{r,s}^{x,y} = \gamma_{x,y}(r - \tfrac{1}{2}\delta).$$

The transversal radius of the lentil $L_{r,s}^{x,y}$ with midpoint $m = m_{r,s}^{x,y}$ is

$$R_{r,s}^{x,y} = \sup\{R > 0; \exp_m(\eta) \in L \text{ whenever } \eta \in T_m M, \eta \perp \dot{\gamma} \text{ and } |\eta| < R\}.$$

That is, (the closure of) the lentil contains a geodesic disc—a ball of dimension $n-1$ — of this radius normal to $\gamma_{x,y}$.

Lemma 24. *Consider a simple manifold of bounded geometry with the constants satisfying the condition (1). If a lentil $L_{r,s}^{x,y}$ has thickness $\delta > 0$, then its diameter satisfies*

$$\mathrm{diam}(L_{r,s}^{x,y}) \leqslant \delta + C_{\mathrm{e}}\sqrt{\delta},$$

where the constant is given by (49) in appendix B.

The condition (1) ensures that $\sin(\mathscr{C}_{\mathrm{diam}}\sqrt{\mathscr{C}_{\mathrm{sec}+}}) > 0$ and thus C_{e} is a well defined and positive constant.

Proof of lemma 24. The proof is composed of two parts. In the first part we show that if δ is small, then the opening angle of the lentil as seen from x is small. In the second part we use this to show that the lentil itself is small.

Let us denote $L := L_{r,s}^{x,y}$. Let $\gamma = \gamma_{x,y}$ be the geodesic between x and y and let $\ell := d(x,y)$. Take any point $z \in L$.

The three points x, y, z form a triangle with sides ℓ, $d(x,z) =: a$, and $d(y,z) =: b$. Let us call the angle at x by the name $\beta \in [0, \pi]$.

We will estimate this angle by comparing to a constant curvature reference manifold. As earlier, let $M_{\mathscr{C}_{\mathrm{sec}+}}$ be the sphere of the same dimension as M with constant sectional curvature

$\mathscr{C}_{\text{sec}+} > 0$. In the corresponding geodesic triangle with the same side lengths ℓ, a, b the corresponding angle is $\tilde{\beta}$. By lemma 21 we have $\beta \leqslant \tilde{\beta}$.

By the spherical law of cosines we have

$$\cos(\tilde{\beta}) = \frac{\cos(\tilde{b}) - \cos(\tilde{\ell})\cos(\tilde{a})}{\sin(\tilde{\ell})\sin(\tilde{a})},$$

where we denoted $\tilde{\ell} = \ell\sqrt{\mathscr{C}_{\text{sec}+}}$, $\tilde{a} = a\sqrt{\mathscr{C}_{\text{sec}+}}$, and $\tilde{b} = b\sqrt{\mathscr{C}_{\text{sec}+}}$. We will also use scaled versions of the two radii and the thickness: $\tilde{r} = r\sqrt{\mathscr{C}_{\text{sec}+}}$, $\tilde{s} = s\sqrt{\mathscr{C}_{\text{sec}+}}$, and $\tilde{\delta} = \delta\sqrt{\mathscr{C}_{\text{sec}+}}$.

By the assumption that $z \in L$ we have $a < r$ and $b < s$, and the reverse triangle inequality gives $a > \ell - s$. These same inequalities hold when scaled with $\sqrt{\mathscr{C}_{\text{sec}+}}$, where

$$\cos(\tilde{\beta}) \geqslant \frac{\cos(\tilde{s}) - \cos(\tilde{\ell})\cos(\tilde{\ell} - \tilde{s})}{\sin(\tilde{\ell})\sin(\tilde{r})}.$$

Using

$$\cos(\tilde{s}) = \cos(\tilde{\ell})\cos(\tilde{\ell} - \tilde{s}) + \sin(\tilde{\ell})\sin(\tilde{\ell} - \tilde{s})$$

and $\ell - s = r - \delta$, the estimate simplifies to

$$\cos(\tilde{\beta}) \geqslant \frac{\sin(\tilde{r} - \tilde{\delta})}{\sin(\tilde{r})}.$$

We will simplify this further using $\sin(\tilde{r} - \tilde{\delta}) \geqslant \sin(\tilde{r}) - \tilde{\delta}$.

Finally, we make use of the fact that $\sin(t) \geqslant T^{-1}\sin(T)t$ for all $t \in [0, T]$ when $T \in (0, \pi)$. All the distances on the manifold are bounded by the diameter bound $\mathscr{C}_{\text{diam}}$, so in our setting $T = \mathscr{C}_{\text{diam}}\sqrt{\mathscr{C}_{\text{sec}+}}$, which was indeed assumed to be below π. Combining these, we find

$$\cos(\tilde{\beta}) \geqslant 1 - C_{\text{f}}\frac{\delta}{r}$$

with the constant given by (50). (In the limit $\mathscr{C}_{\text{sec}+} \to 0$ we have $C_{\text{f}} = 1$, and one can indeed verify the estimate with this constant using Euclidean comparison geometry when the sectional curvature is non-positive.)

Using $\cos(\beta) \leqslant 1 - \frac{1}{5}\beta^2$, which is valid for all $\beta \in [0, \pi]$, we get

$$1 - \tfrac{1}{5}\beta^2 \geqslant \cos(\beta) \geqslant 1 - C_{\text{f}}\frac{\delta}{r},$$

which leads to

$$\beta^2 \leqslant 5C_{\text{f}}\frac{\delta}{r}. \tag{20}$$

This explicit estimate for β in terms of δ and r concludes the first part of the proof.

Now consider any two points $z_1, z_2 \in L$. Estimating their distance will amount to estimating the diameter of the lentil. Let β_i be the angle at x between $\gamma_{x,y}$ and γ_{x,z_i}.

Let $\hat{z}_1$ be the unique point on γ with $d(x, z_1) = d(x, \hat{z}_1)$. The point $\hat{z}_1$ is the closest one to y on the metric sphere $S(x, d(x, z_1))$ because γ meets the sphere orthogonally. Thus $d(y, \hat{z}_1) \leqslant d(y, z_1) < s$. As also $d(x, \hat{z}_1) = d(x, z_1) < r$, we have $\hat{z}_1 \in L$. Let $\eta, \hat{\eta} \in T_x M$ be the vectors for which $\exp_x(\eta) = z_1$ and $\exp_x(\hat{\eta}) = \hat{z}_1$. Let $\tilde{\sigma} : [0, \beta_1] \to T_x M$ be the unit speed circular arc with constant norm and endpoints $\tilde{\sigma}(0) = \hat{\eta}$ and $\tilde{\sigma}(\beta_1) = \hat{\eta}_1$. Then $\sigma = \exp_x \circ \tilde{\sigma}$ is a curve joining $\hat{z}_1$ to z_1.

If sectional curvature and length of geodesics is bounded uniformly, then

$$|\dot{\sigma}(t)| \leqslant C_b d(x,z_1) < C_b r$$

by lemma 22. Estimating the distance between the two points by the length of a curve joining them, we thus have $d(z_1, \hat{z}_1) \leqslant C_b r \beta_1$.

By (20) we have $\beta_i \leqslant \sqrt{5C_f \delta / r}$, and clearly $d(\hat{z}_1, \hat{z}_2) < \delta$ when $\hat{z}_2$ is defined analogously to $\hat{z}_1$ above. The triangle inequality gives

$$d(z_1, z_2) \leqslant C_b r \beta_1 + \delta + C_b r \beta_2 \leqslant \delta + 2\sqrt{5C_f} C_b \sqrt{r\delta}.$$

Thus the diameter of the lentil can be estimated by

$$\mathrm{diam}(L) \leqslant \delta + 2\sqrt{5C_f} C_b \sqrt{\mathrm{diam}(M)} \sqrt{\delta}$$

as claimed. □

Lemma 25. *Consider a simple manifold M of bounded geometry with the constants satisfying the condition (1). Take any $x, y \in M$ and consider the lentil $L_{r,s}^{x,y}$ with $r, s \in (0, d(x,y))$. If m is the midpoint of $L_{r,s}^{x,y}$, then $B(m, \frac{1}{2}\delta) \subset L_{r,s}^{x,y}$ and thus $R_{r,s}^{x,y} \geqslant \frac{1}{2}\delta$.*

Proof. Recall that $L_{r,s}^{x,y} = B(x,r) \cap B(y,s)$. As the distances from m to x and y are $r - \frac{1}{2}\delta$ and $s - \frac{1}{2}\delta$, respectively, it follows from the triangle inequality that $B(m, \frac{1}{2}\delta)$ is contained in both balls defining the lentil. The estimate for the transversal radius follows immediately. □

Lemma 25 is sufficient for our needs, but we point out that the exponent of δ can be improved. We provide this observation as proposition 30 in appendix A.

The endpoints x and y of the relevant geodesics will be source points in $\Gamma \subset B(\partial M, \varepsilon_1)$. As the source set is countable in practice and can well be finite, the geodesics will not cover the deep interior of the manifold. But if the geodesics are fattened suitably, with the suitable level given by lemma 27, they do cover the desired set without gaps. Such a thickening is most naturally realized in our setting as a union of lentils with x, y, and $\delta > 0$ fixed. Lemma 25 and proposition 30 gives a lower bound on the fatness of this 'sausage'.

9.2. Boundary density estimates

We will now describe how the quantities defined in equations (3), (4), and (5) estimate the distance of source points to the boundary.

Consider a source $s \in S$ located at the point $p = \pi(s)$. The boundary distance function $r_p \colon \partial M \to \mathbb{R}$ is given by $r_p(x) = d(p,x)$ and the arrival time function a_s is this function shifted by a constant. Proximity of a source point $p \in \mathrm{int}(M)$ to the boundary causes the Hessians of both these functions to blow up. However, some of the curvature of the graph can be caused by the curvature of the boundary itself rather than the proximity of the source point. Bounded geometry gives a concrete bound on this effect and allows us to ensure with concrete estimates that the source point really is close to the boundary.

Recall that $c(p) \subset \partial M$ is the set of critical points of the boundary distance function r_p.

Lemma 26. *Let M be a simple Riemannian manifold with bounded geometry and $\emptyset \neq P \subset \mathrm{int}(M)$. Let $\varepsilon_1 > 0$. The quantities defined in equations (3), (4), and (5) satisfy the following estimates:*

(1) Take any $p \in P$. If $x \in c(p)$, then $d(p,x) \leqslant E(p,x)$.
(2) For any $x \in \partial M$ there is $p \in P$ with $d(p,x) \leqslant E(x) + \varepsilon_1$.
(3) For any $x \in \partial M$ there is $p \in P$ so that $d(p,x) \leqslant E + \varepsilon_1$.

Proof. Part (1): take any $p \in \text{int}(M)$. Let $\rho \colon M \to \mathbb{R}$ be the distance function $\rho(x) = d(p,x)$ and $r = \rho|_{\partial M}$. The Hessians $\mathcal{H}_\rho$ and $\mathcal{H}_r$ of these functions are symmetric quadratic forms on TM and $T\partial M$, respectively.

Let γ be any unit speed geodesic with $\gamma(0) = p$ and J a Jacobi field along γ so that $J(0) = 0$. Suppose γ meets the boundary orthogonally, which is equivalent with the exit point being a critical point of r. Then the Hessian of the distance function has the property (cf proof of lemma 13)

$$D_t J(t) = \mathcal{H}_\rho J(t)$$

for all $t > 0$ and thus by part (4) of definition 6 we also have

$$\langle J(t), \mathcal{H}_\rho J(t) \rangle \leqslant \mathscr{C}_{\text{JF}} t^{-1} |J(t)|^2.$$

As there are no conjugate points, this amounts to

$$\mathcal{H}_\rho \leqslant \mathscr{C}_{\text{JF}} t^{-1} g$$

in the sense of quadratic forms on $\dot{\gamma}^\perp \subset T_{\gamma(t)} M$.

If $\lambda(x)$ denotes the smallest eigenvalue of $\mathcal{H}_r(x)$ at a boundary point $x \in \partial M$, then (14) yields

$$\lambda(x) h_1(x) \leqslant \mathcal{H}_r(x) = \mathcal{H}_\rho(x) + h_2(x) \leqslant [\mathscr{C}_{\text{JF}} r(x)^{-1} + \mathscr{C}_{\text{SFF}}] h_1(x)$$

as quadratic forms on the boundary. If $\lambda(x) > \mathscr{C}_{\text{SFF}}$, this estimate gives

$$r(x) \leqslant \frac{\mathscr{C}_{\text{JF}}}{\lambda(x) - \mathscr{C}_{\text{SFF}}} =: E(p,x).$$

If $\lambda(x) \leqslant \mathscr{C}_{\text{SFF}}$, then $E(p,x) = \infty$. The claimed estimate thus holds in both cases.

Part (2): let $x \in \partial M$. By the triangle inequality, the estimate $d(y,x) \leqslant d_{\partial M}(y,x)$ and Part (1) we have

$$d(p,x) \leqslant d(p,y) + d_{\partial M}(y,x) \leqslant E(p,y) + d_{\partial M}(y,x)$$

for all $p \in P$ and $y \in c(p)$. Thus by the definition of $E(x)$ in (4) there is a point $p \in P$ with a distance to x less than $E(x) + \varepsilon_1$ to x.

Part (3): follows immediately from the previous one. $\square$

With the aid of lemma 26 we can now prove proposition 8.

Proof of proposition 8. Proposition 3 implies that the data determines the arrival time functions a_s for all $s \in S$, although we do not have a description of the index set S yet. As these functions differ from the boundary distance functions $r_s \colon \partial M \to \mathbb{R}$, $r_s(x) = d(x, \pi(s))$, only by a constant, we thus know the differential and the Hessian of each r_s on all of ∂M. Therefore the data determines the critical points of these functions and of the function $E(p,y)$, for $p = \pi(s)$ and $y \in c(p)$, defined in (3). From this one can easily compute $E(x)$ and E from (4) and (5).

The set Γ is defined in terms of these quantities, so it is uniquely determined by the data as well.

Theorem 2 indicates that the data determines the pointwise spatial distances of the source points, and the last part of the claim follows. $\square$

9.3. Interior density estimates

For the density of sources deep in the manifold *M*, we use two estimates. The first one concerns the density of the set of geodesics connecting near-boundary source points in Γ. The second one ensures that the lentils cover enough of the manifold.

Lemma 27. *Let M be a simple manifold of bounded geometry and let $\varepsilon_2 > 0$. Let $\Gamma \subset M$ be such that $\partial M \subset B(\Gamma, \varepsilon_2)$. Then for every $z \in \mathrm{int}(M)$ there are points $x, y \in \Gamma$ so that*

$$d(z, \gamma_{x,y}([0, d(x, y)])) \leqslant C_d \varepsilon_2.$$

Proof. Fix any $x \in \Gamma$ and take any $z \in \mathrm{int}(M)$. Let $\hat{\gamma} = \gamma_{x,z}$ be a constant speed geodesic with $\hat{\gamma}(0) = x$ and $\hat{\gamma}(t) = z$ for some $t > 0$. We extend this geodesic beyond *z* so that it meets ∂M at some time $t' > t$. We scale the constant speed so that $t' = 1$ and we denote $\hat{y} := \hat{\gamma}(1) \in \partial M$.

There is $y \in \Gamma$ so that $d(y, \hat{y}) < \varepsilon_2$. Let $\gamma \colon [0, 1] \to M$ be the constant speed geodesic for which $\gamma(0) = x$ and $\gamma(1) = y$.

By lemma 23 we have

$$d(\hat{\gamma}(t), \gamma(t)) \leqslant C_d t d(\hat{y}, y).$$

As $t < 1$, we have thus

$$d(z, \gamma([0, 1])) \leqslant d(\hat{\gamma}(t), \gamma(t)) < C_d \varepsilon_2$$

as claimed. $\qquad\square$

Lemma 28. *Take any $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$, and let M be a simple manifold with bounded geometry and $\Gamma \subset B(\partial M, \varepsilon_1)$ so that $\partial M \subset B(\Gamma, \varepsilon_2)$. Let $z \in M$. Whenever $d(z, \partial M) \geqslant \varepsilon_1 + C_g \varepsilon_2$, then there is a lentil $L_{r,s}^{x,y}$ with $x, y \in \Gamma$ and thickness $\delta = C_1 \varepsilon_2 < \min(r, s)$ containing z. The constants are given by (40) and (51) in appendix B.*

Proof. We showed in lemma 27 above that for any $z \in \mathrm{int}(M)$ there are $x, y \in \Gamma$ so that the distance from *z* to the trace of the geodesic $\gamma_{x,y}$ is at most $C_d \varepsilon_2$. We will choose $h > 0$ later (see (24) for an explicit expression) and require that $d(x, z) \geqslant h$ and $d(y, z) \geqslant h$. These additional requirements may not hold for the pair of points $x, y \in \Gamma$ without further assumptions for *z*. As both *x* and *y* are ε_1-close to the boundary, the additional requirements $d(x, z) \geqslant h$ and $d(y, z) \geqslant h$ are certainly satisfied if $d(z, \partial M) \geqslant \varepsilon_1 + h$. This is why we placed a boundary distance assumption on *z* in the statement of this lemma.

Let $z' \in \gamma_{x,y}$ be the nearest point on this geodesic to the point *z*. After choosing *h* large enough we obtain by the reverse triangle inequality

$$d(x, z'), d(y, z') \geqslant h - C_d \varepsilon_2 > 0. \tag{21}$$

This implies that z' is an interior point of $\gamma_{x,y}$, and the geodesics $\gamma_{x,y}$ and $\gamma_{z,z'}$ meet orthogonally at z'. To make sure that *z* is contained in a lentil $L_{r,s}^{x,y}$, with thickness δ and midpoint z', we require that the transversal radius of that lentil is more than $d(z, z')$.

By lemma 25 the aforementioned condition of the transversal radius is satisfied when

$$C_d \varepsilon_2 \leqslant \tfrac{1}{2}\delta. \tag{22}$$

The condition for z' to be a midpoint of a lentil of thickness δ is that $d(x, z') > \tfrac{1}{2}\delta$ and $d(y, z') > \tfrac{1}{2}\delta$, which by equation (21) are satisfied when

$$h - C_d \varepsilon_2 > \tfrac{1}{2}\delta. \tag{23}$$

Combining the two conditions (22) and (23) gives

$$C_d \varepsilon_2 < h - C_d \varepsilon_2$$

and therefore

$$h > 2 C_d \varepsilon_2.$$

We choose

$$h = 3 C_d \varepsilon_2. \tag{24}$$

The conditions (22) and (23) then become

$$2 C_d \varepsilon_2 \leqslant \delta < 4 C_d \varepsilon_2,$$

so we may choose

$$\delta = 2 C_d \varepsilon_2. \tag{25}$$

This is why we chose C_1 in (40) as we did.

Finally we note that by the choice of $z' \in \gamma_{x,y}$ and due to (21), (24), and (25) we have $d(x,y) > \delta$. If we set $r = d(x,z') + \frac{1}{2}\delta$ and $s = \delta + d(x,y) - r = d(z',y) + \frac{1}{2}\delta$, then the lentil $L_{r,s}^{x,y}$ is of width δ and has z' as a midpoint. Moreover, by the previous argument this lentil contains the point z, if $d(z,\partial M) \geqslant \varepsilon_1 + C_g \varepsilon_2$, as required in the claim of this lemma. □

The choices of h and δ in the proof above are optimal up to a constant and the explicit choice simplifies our estimates.

9.4. Global density estimates

With all the estimates we have collected, we are ready to prove both parts of theorem 9. Claim (1) states that if all lentils of thickness δ meet a source point (which is verifiable from data), then the source set $P_1 \subset M_1$ has a concrete density estimate. The thickness δ is chosen carefully depending on ε_2. Claim (2) states that under the assumptions of the previous claim the measurements define a discrete metric space which is a quantitatively good approximation for the true space.

Proof of theorem 9. For claim (1), take any $z \in M$. We wish to show that there is $p \in P$ so that $d(z,p) < \varepsilon$ with ε chosen as in (8). We split the proof in two cases: near the boundary and deep within the manifold, and use different tools in either of these cases to verify the validity of the density claim.

The boundary case: if $d(z,\partial M) < \varepsilon_1 + C_g \varepsilon_2$, then there is $x \in \partial M$ so that $d(z,x) < \varepsilon_1 + C_g \varepsilon_2$. By claim (3) of lemma 26, there is $p \in P$ so that $d(x,p) < E + \varepsilon_1$. Recall that $\varepsilon_2 = E + \varepsilon_1 > \varepsilon_1$. By the triangle inequality we have

$$d(z,P) \leqslant d(z,p) \leqslant d(z,x) + d(x,p) < (C_g + 2)\varepsilon_2. \tag{26}$$

This is a sufficient estimate near the boundary.

The deep interior case: by definition of the set Γ in (6) and lemma 26 we have $\Gamma \subset B(\partial M, \varepsilon_1)$ and $\partial M \subset B(\Gamma, \varepsilon_2)$. If $d(z,\partial M) \geqslant \varepsilon_1 + C_g \varepsilon_2$, then lemma 28 shows that there is a lentil $L = L_{r,s}^{x,y}$ of the correct thickness $\delta = C_1 \varepsilon_2 < r,s$ so that $z \in L$. By the crucial assumption of the theorem, there is $p \in P \cap L$. Lemma 24 then gives us

$$d(z,P) \leqslant d(z,p) \leqslant \delta + C_e\sqrt{\delta} = C_1\varepsilon_2 + C_e\sqrt{C_1\varepsilon_2}. \tag{27}$$

This is a sufficient estimate for deep interior points.

Now it remains to combine the estimates for points in the deep interior and near the boundary. Combining (26) and (27) gives

$$d(z,P) \leqslant \max\left((C_g + 2)\varepsilon_2, C_1\varepsilon_2 + C_e\sqrt{C_1\varepsilon_2}\right)$$
$$\leqslant (C_g + 2 + C_1)\varepsilon_2 + C_e\sqrt{C_1\varepsilon_2}$$

for all $z \in M$. This proves that $P \subset M$ is ε-dense with the choice of (8) and concludes the proof of claim (1).

The labeled Gromov–Hausdorff distance is now straightforward to estimate for claim (2). By the definition of $E(x)$ in (4) we can choose a function $\alpha \colon \partial M \to P$ that satisfies (9); heuristically, the point $\alpha(x)$ is an ε_1-approximate minimizer of the right-hand side of (4). This function can be actually identified from the data $Q(S)$ as we know the first fundamental form of ∂M and the value of $E(p,y)$ for all $p \in \pi(S)$ and $y \in c(p)$, given by (3), is obtained by the computing the gradient and the Hessian of the corresponding arrival time function. Then by lemma 26 it follows that for each $x \in \partial M$ there exists $y \in c(\alpha(x))$ that satisfies

$$\varepsilon_2 \geqslant E(x) + \varepsilon_1 > E(\alpha(x),y) + d_{\partial M}(x,y)$$
$$\geqslant d(\alpha(x),y) + d(x,y)$$
$$\geqslant d(\alpha(x),x).$$

Since $P \subset M$ is ε-dense, proposition 16 finally gives

$$d_{GH}^{\partial M}(P,\alpha;M,\iota) \leqslant \varepsilon + \varepsilon_2.$$

By the choice of ε in (8) this proves estimate (10) and completes the proof. $\qquad\square$

9.5. Reverse density estimates

We do not need any more preparations before the next proof.

Proof of theorem 10. Consider a unit speed geodesic γ starting at a source point $p = \gamma(0) \in P$. Take any $t > 0$ and any $w \in T_{\gamma(t)}M$ orthogonal to $\dot\gamma(t)$. By bounded geometry we have

$$\langle w, \mathcal{H}_\rho w \rangle \geqslant |w|^2 (\mathscr{C}_{H1}t^{-1} - \mathscr{C}_{H2}).$$

Now suppose that $y = \gamma(t)$ is on ∂M and $\dot\gamma(t)$ is normal to the boundary.

We recall the earlier notation that $\rho \colon M \to \mathbb{R}$ is the distance function to the point p and $r = \rho|_{\partial M}$. The second fundamental form h_2 is positive definite by simplicity and by (14) we deduce that the Hessians on ∂M and M evaluated at y are related by

$$\mathcal{H}_r(y) = \mathcal{H}_\rho(y) + h_2(y) \geqslant (\mathscr{C}_{H1}t^{-1} - \mathscr{C}_{H2})h_1(y),$$

where h_1 is the first fundamental form. Therefore the smallest eigenvalue $\lambda(p,y)$ of $\mathcal{H}_r(y)$ satisfies

$$\lambda(p,y) \geqslant \mathscr{C}_{H1}d(p,y)^{-1} - \mathscr{C}_{H2}. \tag{28}$$

This enables us to estimate $E(p,y)$.

To utilize the assumption that $P \subset M$ is $\hat{\varepsilon}$-dense with $\hat{\varepsilon}$ chosen by (11), suppose that $d(\partial M, p) < \hat{\varepsilon}$ and let $z_p \in \partial M$ be the nearest boundary point to it. Estimate (28) then implies that

$$E(p, z_p) \leqslant \frac{\mathscr{C}_{\mathrm{JF}}}{\mathscr{C}_{\mathrm{H1}}\hat{\varepsilon}^{-1} - \mathscr{C}_{\mathrm{H2}} - \mathscr{C}_{\mathrm{SFF}}}$$

when $\hat{\varepsilon} < \mathscr{C}_{\mathrm{H1}}/(\mathscr{C}_{\mathrm{H2}} + \mathscr{C}_{\mathrm{SFF}})$. Due to our choice of $\hat{\varepsilon}$, we have

$$\hat{\varepsilon} \leqslant \frac{\mathscr{C}_{\mathrm{H1}}}{2(\mathscr{C}_{\mathrm{H2}} + \mathscr{C}_{\mathrm{SFF}})}, \tag{29}$$

and so

$$E(p, z_p) \leqslant 2\mathscr{C}_{\mathrm{JF}}\mathscr{C}_{\mathrm{H1}}^{-1}\hat{\varepsilon}. \tag{30}$$

Consider then an arbitrary point $x \in \partial M$. There is a source point $q \in P$ with $d(x, q) < \hat{\varepsilon}$. Let $z_q \in \partial M$ be the closest boundary point to q. We have $d(z_q, q) < \hat{\varepsilon}$, and by bounded geometry (part (6) of definition 6)

$$\begin{aligned} d_{\partial M}(x, z_q) &\leqslant \mathscr{C}_{\mathrm{dist}} d_M(x, z_q) \\ &\leqslant \mathscr{C}_{\mathrm{dist}}(d_M(x, q) + d_M(z_q, q)) \\ &< 2\mathscr{C}_{\mathrm{dist}}\hat{\varepsilon}. \end{aligned}$$

By (30) we get

$$E(x) \leqslant E(q, z_q) + d_{\partial M}(x, z_q) \leqslant 2\mathscr{C}_{\mathrm{JF}}\mathscr{C}_{\mathrm{H1}}^{-1}\hat{\varepsilon} + 2\mathscr{C}_{\mathrm{dist}}\hat{\varepsilon}.$$

As this bound it is independent of $x \in \partial M$, we have the same bound for E from (5):

$$E \leqslant 2(\mathscr{C}_{\mathrm{JF}}\mathscr{C}_{\mathrm{H1}}^{-1} + \mathscr{C}_{\mathrm{dist}})\hat{\varepsilon}, \tag{31}$$

under the assumption (29).

Combining our choices $\varepsilon_1 = C_5\hat{\varepsilon}$ and $\varepsilon_2 = \varepsilon_1 + E$ with (31), we get

$$\varepsilon_2 = \varepsilon_1 + E \leqslant (C_5 + 2\mathscr{C}_{\mathrm{JF}}\mathscr{C}_{\mathrm{H1}}^{-1} + 2\mathscr{C}_{\mathrm{dist}})\hat{\varepsilon}. \tag{32}$$

All other estimates for ε_2 follow from this one.

To satisfy (12), we require that $C_4\varepsilon_2 < \frac{1}{2}\varepsilon$ and $C_3\sqrt{\varepsilon_2} < \frac{1}{2}\varepsilon$. By (32) both follow from requiring

$$C_4(C_5 + 2\mathscr{C}_{\mathrm{JF}}\mathscr{C}_{\mathrm{H1}}^{-1} + 2\mathscr{C}_{\mathrm{dist}})\hat{\varepsilon} < \frac{1}{2}\varepsilon$$

and

$$C_3\sqrt{(C_5 + 2\mathscr{C}_{\mathrm{JF}}\mathscr{C}_{\mathrm{H1}}^{-1} + 2\mathscr{C}_{\mathrm{dist}})\hat{\varepsilon}} < \frac{1}{2}\varepsilon,$$

or equivalently

$$\hat{\varepsilon} < \min\left(\frac{\varepsilon}{2C_4(C_5 + 2\mathscr{C}_{\mathrm{JF}}\mathscr{C}_{\mathrm{H1}}^{-1} + 2\mathscr{C}_{\mathrm{dist}})}, \frac{\varepsilon^2}{4C_3^2(C_5 + 2\mathscr{C}_{\mathrm{JF}}\mathscr{C}_{\mathrm{H1}}^{-1} + 2\mathscr{C}_{\mathrm{dist}})}\right). \tag{33}$$

Thus (29) and (33) are satisfies due to (11), and by the previous remark the estimate (12) follows.

It remains to show that each lentil $L = L_{r, d(x,y) - r + \delta}^{x,y}$ with $x, y \in P$ contain source points as claimed. The thickness of the lentil L is $\delta = C_1\varepsilon_2$, so by lemma 25 $B(m, \frac{1}{2}\delta) \subset L$, where m is

the midpoint of the lentil L. Therefore each lentil of thickness δ contains a ball of radius $\frac{1}{2}\delta$. It follows from $\hat{\varepsilon}$-density of $P \subset M$ that that each such lentil contains a source point because

$$\hat{\varepsilon} = C_5^{-1}\varepsilon_1 = \tfrac{1}{2}C_1\varepsilon_1 \leqslant \tfrac{1}{2}C_1\varepsilon_2 = \tfrac{1}{2}\delta.$$

This is why we chose C_5 so that $C_5^{-1} = \frac{1}{2}C_1$.

$\square$

10. Convergence of discrete approximations

10.1. Deterministic convergence

With proposition 8, theorems 9 and 10 the proof of theorem 11 is straightforward.

Proof of theorem 11. We begin with proposition 3 with $\Omega = \partial M \times (0,T)$. Out of the parts of graphs we only choose the ones that are full graphs of a function $a_s\colon \partial M \to \mathbb{R}$. It follows from the diameter bound that

$$\tau(s) \leqslant a_s(x) \leqslant \tau(s) + \mathscr{C}_{\text{diam}}$$

for all $x \in \partial M$. Therefore the set $Q(S,T)$ of (13) contains the full graph of a_s for all the sources $s \in S$ with $0 < \tau(s) < T - \mathscr{C}_{\text{diam}}$. Let us denote the set of source points with their complete graphs contained in $\partial M \times (0,T)$ by $P_T \subset \pi(S)$.

If T is too small, we may have $P_T = \emptyset$. In this case we choose the metric approximation M_T to be a set of one point and $\alpha_T\colon \partial M \to M_T$ the constant map. When there are sources, we set $M_T = P_T$.

Take any $\varepsilon > 0$ and let $\hat{\varepsilon}$ be given by (11). As $\bigcup_{T>0} P_T$ is dense in M by assumption, by compactness there is $T(\varepsilon) > 0$ so that P_T is $\hat{\varepsilon}$-dense for all $T \geqslant T(\varepsilon)$. With the choices $\varepsilon_1 = C_5\hat{\varepsilon}$ and $\varepsilon_2 = \varepsilon_1 + E(T)$, where $E(T)$ is defined as in (5), and $\delta = C_1\varepsilon_2$ we have

$$C_4\varepsilon_2 + C_3\sqrt{\varepsilon_2} < \varepsilon, \tag{34}$$

by theorem 10. By the last claim of theorem 10 the assumption of theorem 9 is satisfied. Thus the set P_T is also ε-dense as required in claim ((2)) of theorem 9. We choose a map $\alpha_T\colon \partial M \to P_T$ so that (9) is satisfied when $\alpha = \alpha_T$ and $P = P_T$. Finally the estimate (10) with (34) yields

$$d_{GH}^{\partial M}(P_T, \alpha_T; M, \iota) < \varepsilon.$$

It should be noted that due to proposition 8, the auxiliary quantity $E(T)$ and the map α_T used above are defined by the data.

$\square$

10.2. Dense sources

Now we are finally able to prove theorem 1. We will make use of a version of the Myers–Steenrod Theorem [46, 47] which states that a metric isometry between smooth Riemannian manifolds is necessarily smooth. In the case of simple manifolds it is straightforward to prove and the boundary causes no technical trouble. We record the proof here for the sake of completeness.

Lemma 29 (Myers–Steenrod Theorem). *If $\Phi\colon M_1 \to M_2$ is a metric isometry between simple Riemannian manifolds, it is smooth up to the boundary and an isometry also in the sense that $\Phi^* g_2 = g_1$.*

Proof. It suffices to show that Φ is smooth. Then preserving distances quickly implies $\Phi^* g_2 = g_1$ via differentiating distance functions.

Take any $x_1 \in \mathrm{int}(M_1)$ and let $x_2 = \Phi(x_1)$. If γ is any constant speed geodesic with $\gamma(0) = x_1$, then $\Phi \circ \gamma$ is a constant speed geodesic starting at x_2 with the same speed. Let $\Phi_* : T_{x_1} M_1 \to T_{x_2} M_2$ be the map that maps the velocity vectors of these geodesics to each other. That is, if $\partial_t \gamma(t)|_{t=0} = v$, then $\partial_t \Phi(\gamma(t))|_{t=0} = \Phi_*(v)$. We have $\Phi \circ \exp_{x_1} = \exp_{x_2} \circ \Phi_*$.

The map Φ_* is clearly homogeneous in positive scalings, maps unit vectors to unit vectors, and is bijective. By considering reversed geodesics we find $\Phi_*(-v) = -\Phi_*(v)$ for all $v \in T_{x_1} M_1$.

For any two $v, w \in T_{x_1} M_1$ we have [28, equation (2.15)]

$$d(\exp_{x_1}(tv), \exp_{x_1}(tw))^2 = t^2(|v|^2 + |w|^2 - 2\langle v, w \rangle) + \mathcal{O}(t^4)$$

for small $t > 0$. Due to Φ_* preserving the norm, the intertwining property $\Phi \circ \exp_{x_1} = \exp_{x_2} \circ \Phi_*$, and the isometric nature of Φ, this implies that

$$\langle v, w \rangle = \langle \Phi_* v, \Phi_* w \rangle$$

for all $v, w \in T_{x_1} M_1$.

The map Φ_* preserves norms and inner products, and therefore it preserves the (squared) distance between any pair of vectors in $T_{x_1} M_1$. Thus Φ_* is an isometry between finite-dimensional inner product spaces, and as a Euclidean isometry fixing the origin it is linear and thus smooth. As the exponential maps are diffeomorphisms on their maximal domain of definition, the map $\Phi = \exp_{x_2} \circ \Phi_* \circ \exp_{x_1}^{-1}$ is smooth. $\qquad \square$

Proof of theorem 1. Let $\iota_1 : \partial M_1 \to M_1$ be the usual inclusion map and let $\iota_2 = \phi : \partial M_1 \to M_2$ (with extended codomain here for convenience) so that we have boundary inclusions $\iota_i : \partial M_1 \to M_i$ with the same domain.

For any $T > 0$ consider data on the bounded set $\partial M \times (-T, T)$. This time the time interval extends into both the past and the future. As in the proof of theorem 11 above, we get a finite metric space M_T and a map $\alpha_T : \partial M_1 \to M_T$ so that

$$d_{GH}^{\partial M_1}(M_T, \alpha_T; M_i, \iota_i) \to 0$$

as $T \to \infty$ for both $i = 1, 2$. By proposition 8 the approximating metric space M_T constructed from data is the same for the two manifolds M_1 and M_2.

We then apply proposition 5. By the triangle inequality

$$d_{GH}^{\partial M_1}(M_1, \iota_1; M_2, \iota_2) \to 0$$

as $T \to \infty$ and so $d_{GH}^{\partial M_1}(M_1, \iota_1; M_2, \iota_2) = 0$. Now proposition 5 gives an isometry $\Phi : M_1 \to M_2$ so that $\Phi \circ \iota_1 = \iota_2$. The condition $\Phi \circ \iota_1 = \iota_2$ simply means that $\Phi|_{\partial M_1} = \phi$. By lemma 29 the isometry Φ is actually smooth.

It remains to check that the sources $s = (p, t) \in S_1 \subset \mathrm{int}(M_1) \times \mathbb{R}$ are mapped correctly. By theorem 2 we know that there is a bijection, but we have to verify that it corresponds to the isometry Φ as claimed. In light of proposition 3, the data can be seen as a collection of graphs. Let us denote $A_i = \{a_s; s \in S_i\}$. The two manifolds having equivalent data means that the map $\phi^* : A_2 \to A_1$ that takes $a \mapsto \phi^* a = a \circ \phi$ is bijective. We will show that the maps $b_i : S_i \to A_i$ with $b_i(s) = a_s$ are bijective, and therefore the natural bijection between the sources $S_1 \to S_2$ is $\xi = b_2^{-1} \circ (\phi^*)^{-1} \circ b_1$. We need this map to satisfy (2), which now amounts to

$$a_{(p,t)} = a_{(\Phi(p),t)} \circ \phi$$

as functions $\partial M_1 \to \mathbb{R}$ for all $(p,t) \in S_1$. This follows straightforwardly from the definitions and Φ being an isometry.

Let us then show that b_i is bijective for both i. To this end, take any two distinct sources $s, \hat{s} \in S_i$. The two arrival time functions are $a_s = r_s + \tau(s)$ and $a_{\hat{s}} = r_{\hat{s}} + \tau(\hat{s})$.

If the spatial source points are different, $\pi(s) \neq \pi(\hat{s})$, then $r_s - r_{\hat{s}}$ cannot be a constant function on ∂M due to lemma 14. Therefore a_s and $a_{\hat{s}}$ do not coincide. If the spatial points are the same, then $r_s - r_{\hat{s}}$ is the constant function 0. As $s \neq \hat{s}$ but $\pi(s) = \pi(\hat{s})$, we must have $\tau(s) \neq \tau(\hat{s})$ and thus $a_s - a_{\hat{s}}$ is a non-zero constant function.

We have thus proven that $s \neq \hat{s}$ implies $b_i(s) \neq b_i(\hat{s})$. This concludes the proof that b_i is injective and thus the theorem is proven. $\qquad\square$

10.3. Stochastic convergence

The proof of the stochastic result only requires checking that the point process almost surely produces the correct kind of source set.

Proof of proposition 12. Countability of the set $S \subset M \times \mathbb{R}$ given by the homogeneous Poisson point process is certain. We need to prove that the following properties hold almost surely:

(1) $\pi(S) \cap \partial M = \emptyset$.
(2) $S \subset M \times \mathbb{R}$ is discrete.
(3) $\pi(S \cap (M \times [0, \infty))) \subset M$ is dense.

The probability that a measurable set $A \subset M \times \mathbb{R}$ contains $k \in \mathbb{N}$ points of the source set S is

$$P[\#(A \cap S) = k] = e^{-\lambda \mu(A)} \frac{(\lambda \mu(A))^k}{k!}. \tag{35}$$

For this and other basic properties of homogeneous Poisson point processes, see e.g. [23, 57].

The first property follows simply from $\mu(\partial M \times \mathbb{R}) = 0$. The second property follows from $A \cap S$ being almost certainly finite when $\mu(A) < \infty$.

For the third property, let $(x_i)_{i=1}^{\infty}$ be a dense sequence in M and define for any $i, j \geqslant 1$ the set

$$A_{i,j} = B_M(x_i, j^{-1}) \times [0, \infty) \subset M \times \mathbb{R}.$$

Comparability with the natural product measure gives $\mu(A_{i,j}) = \infty$. Therefore it follows from (35) that $P[\#(A_{i,j} \cap S) = 0] = 0$.

The projection $\pi(S \cap (M \times [0, \infty)))$ is dense in M if $\#(A_{i,j} \cap S) > 0$ for all i and j. As this event for individual indices has probability 1, the event for density is a countable intersection of events of full probability. Therefore density is almost certain.

The conditions of theorem 11 are thus met almost surely. $\qquad\square$

Data availability statement

No new data were created or analysed in this study.

Acknowledgments

Appendix A. An improved estimate on transversal radius

As mentioned in connection to lemma 25, the estimate can be improved although it is not necessary for the proof of our theorems. The estimate $R \gtrsim \delta$ of lemma 25 can be improved to $R \gtrsim \sqrt{\delta}$ as follows.

Proposition 30. *Consider a simple manifold of bounded geometry with the constants satisfying the condition (1). Take any $x,y \in M$ and consider the lentil $L_{r,s}^{x,y}$ with $r,s \in (0,d(x,y))$. Suppose $\delta_{r,s}^{x,y} = r + s - d(x,y)$ satisfies $\delta_{r,s}^{x,y} \leqslant \min(r,s)$. The transversal radius of any lentil satisfies*

$$R_{r,s}^{x,y} > \min\left(C_\mathrm{h} \sqrt{\delta_{r,s}^{x,y} \min(r,s)}, \tfrac{1}{2}\mathscr{C}_\mathrm{diam} \right),$$

where the constant is given by (52) in appendix B.

Most of our constants are used for estimates from above, but C_h is used for estimating from below. Therefore, unlike most of our constants, it can be made smaller but not larger if needed. It is also natural that the transversal radius cannot exceed half of the diameter. Lemma 25 is more convenient to use and we do not benefit significantly from the improved exponent, so we will not employ proposition 30 in the proofs of our main results.

Proof of proposition 30. Consider a lentil $L = L_{r,s}^{x,y}$ with thickness $\delta = \delta_{r,s}^{x,y}$ and any point $z \in M$ for which the closest point on $\gamma_{x,y}$ is the midpoint $m = m_{r,s}^{x,y}$. We want to find conditions on the distance $w := d(z,m)$ which ensure that $z \in L$. We will do so by making explicit estimates on a comparison manifold of constant sectional curvature and then translating the resulting estimate to the actual manifold.

We assume that $w \leqslant \tfrac{1}{2}\mathscr{C}_\mathrm{diam}$. Due to (1) this implies $w^2\mathscr{C}_\mathrm{sec+} < \pi^2/4$, and so $1 - \tfrac{1}{2}w^2\mathscr{C}_\mathrm{sec+} \in (-1,1]$. This will ensure that the argument of $\arccos\colon [-1,1] \to [0,\pi]$ stays within the domain in the following treatment. We may freely assume $w > 0$ when convenient, as the case $w = 0$ can be given a trivial separate treatment.

Let us denote $R := d(x,m)$ and $d := d(x,z)$. The points x, z, and m form a triangle with a right angle at m. Consider the corresponding triangle on the $M_{\mathscr{C}_\mathrm{sec+}}$ with constant sectional curvature $\mathscr{C}_\mathrm{sec+}$ with with the right angle at $\tilde{m}$ with the same lengths R and w of the catheti. The length of the hypotenuse is

$$\tilde{d} = \mathscr{C}_\mathrm{sec+}^{-1/2} \arccos[\cos(R\sqrt{\mathscr{C}_\mathrm{sec+}})\cos(w\sqrt{\mathscr{C}_\mathrm{sec+}})].$$

The cosine satisfies $\cos(x) \geqslant 1 - \frac{1}{2}x^2$, so

$$\tilde{d} \leqslant \mathscr{C}_{\text{sec}+}^{-1/2} \arccos[\cos(R\sqrt{\mathscr{C}_{\text{sec}+}})[1 - \tfrac{1}{2}w^2\mathscr{C}_{\text{sec}+}]].$$

The function $t \mapsto \arccos(1 - t^2)$ is Lipschitz-continuous on $[0, T]$ with the Lipschitz constant

$$\frac{2}{\sqrt{2 - T^2}}$$

whenever $T < \sqrt{2}$. In our setting $T = \sqrt{C_{\text{c}}}$ (with the constant of (39)) and the two values of t where we compare $\arccos(1 - t^2)$ are

$$\sqrt{1 - \cos(R\sqrt{\mathscr{C}_{\text{sec}+}})[1 - \tfrac{1}{2}w^2\mathscr{C}_{\text{sec}+}]}$$

and

$$\sqrt{1 - \cos(R\sqrt{\mathscr{C}_{\text{sec}+}})}.$$

The condition (1) implies that indeed $T < \sqrt{2}$, so the Lipschitz constant is C_{i} as given in (53). This Lipschitz-continuity now gives

$$\tilde{d} \leqslant R + \mathscr{C}_{\text{sec}+}^{-1/2} C_{\text{i}} \Bigg[\sqrt{1 - \cos(R\sqrt{\mathscr{C}_{\text{sec}+}})[1 - \tfrac{1}{2}w^2\mathscr{C}_{\text{sec}+}]}$$
$$- \sqrt{1 - \cos(R\sqrt{\mathscr{C}_{\text{sec}+}})} \Bigg].$$

Using the estimate $\sqrt{1 - a + ab} \leqslant \sqrt{1 - a} + \frac{1}{2}|ab|\sqrt{1 - a}$ (which is valid for $a < 1$ and $b \in \mathbb{R}$) leads to

$$\tilde{d} \leqslant R + \mathscr{C}_{\text{sec}+}^{-1/2} C_{\text{i}} \frac{\mathscr{C}_{\text{sec}+}\left|\cos(R\sqrt{\mathscr{C}_{\text{sec}+}})\right|}{4\sqrt{1 - \cos(R\sqrt{\mathscr{C}_{\text{sec}+}})}} w^2.$$

A simple calculation gives $|\cos(x)|[1 - \cos(x)]^{-1/2} \leqslant 4/x$ for all $x \in (0, \pi)$. Thus we find

$$\tilde{d} \leqslant R + C_{\text{i}}R^{-1}w^2,$$

completing our estimate of the comparison length $\tilde{d}$.

By lemma 20 we have $d \leqslant \tilde{d}$, and so

$$d \leqslant R + C_{\text{i}}R^{-1}w^2. \tag{36}$$

Let us use this to see when $z \in L$.

To ensure $z \in L$, we need $d < r$. As $r = R + \frac{1}{2}\delta$, estimate (36) says that $d < r$ whenever $C_{\text{i}}R^{-1}w^2 < \frac{1}{2}\delta$. Thus we get the condition $w^2 < (2C_{\text{i}})^{-1}\delta(r - \frac{1}{2}\delta)$ and the rougher bound $\frac{1}{2}\mathscr{C}_{\text{diam}}$. A similar treatment of the condition $d(y, z) < s$ leads us to $w^2 < (2C_{\text{i}})^{-1}\delta(s - \frac{1}{2}\delta)$. We have thus shown that

$$R_{r,s}^{x,y} > \min\left((2C_{\text{i}})^{-1/2}\sqrt{\delta(r - \tfrac{1}{2}\delta)}, (2C_{\text{i}})^{-1/2}\sqrt{\delta(s - \tfrac{1}{2}\delta)}, \tfrac{1}{2}\mathscr{C}_{\text{diam}}\right).$$

As $\delta < r, s$, this implies the claimed inequality. $\qquad\square$

Appendix B. Constants

The constants in our results depend explicitly on the constants describing the simplicity of the manifold in definition 6. The same applies to a number of auxiliary constants appearing in our lemmas and their proofs. To express these constants neatly, we make use of the three constants

$$C_{\mathrm{a}} = \mathrm{sinc}(\mathscr{C}_{\mathrm{diam}}\sqrt{\mathscr{C}_{\mathrm{sec}+}}), \tag{37}$$

$$C_{\mathrm{b}} = \mathrm{sinhc}(\mathscr{C}_{\mathrm{diam}}\sqrt{\mathscr{C}_{\mathrm{sec}-}}), \tag{38}$$

$$C_{\mathrm{c}} = \mathrm{vercos}(\mathscr{C}_{\mathrm{diam}}\sqrt{\mathscr{C}_{\mathrm{sec}+}}). \tag{39}$$

defined in terms of three less common trigonometric functions: $\mathrm{sinc}(x) = \sin(x)/x$, its hyperbolic counterpart $\mathrm{sinhc}(x) = \sinh(x)/x$, and the vercosine $\mathrm{vercos}(x) = 1 + \cos(x)$.

We will express our constants below in two forms. The first one is neatly expressed in terms of the fundamental constants and those of (37)–(39). The second one is the one that arises naturally from the proof. We recommend using the first ones for reading the claims and the second ones for reading the proofs. Our main results use these constants:

$$C_1 = 2\mathscr{C}_{\mathrm{exp}}C_{\mathrm{b}} \tag{40}$$

$$= 2C_{\mathrm{d}},$$

$$C_2 = 2 + 5\mathscr{C}_{\mathrm{exp}}C_{\mathrm{b}} \tag{41}$$

$$= 2 + C_1 + C_{\mathrm{g}},$$

$$C_3 = 2\sqrt{10}\mathscr{C}_{\mathrm{diam}}^{1/2}\mathscr{C}_{\mathrm{exp}}^{1/2}C_{\mathrm{a}}^{1/2}C_{\mathrm{b}}^{3/2} \tag{42}$$

$$= C_{\mathrm{e}}\sqrt{C_1},$$

$$C_4 = 3 + 5\mathscr{C}_{\mathrm{exp}}C_{\mathrm{b}} \tag{43}$$

$$= C_2 + 1,$$

$$C_5 = \mathscr{C}_{\mathrm{exp}}^{-1}C_{\mathrm{b}}^{-1} \tag{44}$$

$$= 2C_1^{-1},$$

$$C_6 = \frac{\mathscr{C}_{\mathrm{H1}}}{2(\mathscr{C}_{\mathrm{H2}} + \mathscr{C}_{\mathrm{SFF}})}, \tag{45}$$

$$C_7 = \frac{1}{2(3 + 5\mathscr{C}_{\mathrm{exp}}C_{\mathrm{b}})(\mathscr{C}_{\mathrm{exp}}^{-1}C_{\mathrm{b}}^{-1} + 2\mathscr{C}_{\mathrm{JF}}\mathscr{C}_{\mathrm{H1}}^{-1} + 2\mathscr{C}_{\mathrm{dist}})} \tag{46}$$

$$= \frac{1}{2C_4(C_5 + 2\mathscr{C}_{\mathrm{JF}}\mathscr{C}_{\mathrm{H1}}^{-1} + 2\mathscr{C}_{\mathrm{dist}})},$$

$$C_8 = \frac{1}{160\mathscr{C}_{\mathrm{diam}}\mathscr{C}_{\mathrm{exp}}\mathscr{C}_a C_{\mathrm{b}}^3 (\mathscr{C}_{\mathrm{exp}}^{-1}C_{\mathrm{b}}^{-1} + 2\mathscr{C}_{\mathrm{JF}}\mathscr{C}_{\mathrm{H1}}^{-1} + 2\mathscr{C}_{\mathrm{dist}})} \tag{47}$$

$$= \frac{1}{4C_3^2(C_5 + 2\mathscr{C}_{\mathrm{JF}}\mathscr{C}_{\mathrm{H1}}^{-1} + 2\mathscr{C}_{\mathrm{dist}})}.$$

Our auxiliary results make use of these constants:

$$C_{\mathrm{d}} = \mathscr{C}_{\mathrm{exp}}C_{\mathrm{b}}, \tag{48}$$

$$C_{\mathrm{e}} = 2\sqrt{5}C_{\mathrm{a}}^{1/2}C_{\mathrm{b}}\mathscr{C}_{\mathrm{diam}}^{1/2}, \tag{49}$$

$$C_{\mathrm{f}} = C_{\mathrm{a}}^{-1}, \tag{50}$$

$$C_{\mathrm{g}} = 3 C_{\mathrm{d}}. \tag{51}$$

Finally, these constants appear in proposition 30 in appendix A:

$$C_{\mathrm{h}} = 2^{-3/2} C_{\mathrm{c}}^{1/4}, \tag{52}$$

$$C_{\mathrm{i}} = 2 C_{\mathrm{c}}^{-1/2}. \tag{53}$$

ORCID iDs

Joonas Ilmavirta ![ORCID] https://orcid.org/0000-0002-2399-0911
Teemu Saksala ![ORCID] https://orcid.org/0000-0002-3785-9623

References

[1] Alexander R and Alexander S 1981 Geodesics in Riemannian manifolds-with-boundary *Indiana Univ. Math. J.* **30** 481–8
[2] Assylbekov Y M and Zhou H 2015 Boundary and scattering rigidity problems in the presence of a magnetic field and a potential *Inverse Problems Imaging* **9** 935–50
[3] Bao D, Chern S-S and Shen Z 2012 *An Introduction to Riemann-Finsler Geometry* vol 200 (Springer Science & Business Media)
[4] Belishev M I and Kurylev Y V 1992 To the reconstruction of a Riemannian manifold via its spectral data (Bc–Method) *Commun. PDE* **17** 767–804
[5] Blias E and Grechka V 2013 Analytic solutions to the joint estimation of microseismic event locations and effective velocity model *Geophysics* **78** KS51–61
[6] Burago D, Burago Y and Ivanov S 2001 *A Course in Metric Geometry* vol 33 (American Mathematical Society Providence)
[7] Burago D and Ivanov S 2010 Boundary rigidity and filling volume minimality of metrics close to a flat one *Ann. Math.* **171** 1183–211
[8] Cerveny V 2005 *Seismic Ray Theory* (Cambridge University Press)
[9] Croke C B *et al* 1991 Rigidity and the distance between boundary points *J. Differ. Geom.* **33** 445–64
[10] 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
[11] de Hoop M V, Holman S F, Iversen E, Lassas M and Ursin B 2015 Recovering the isometry type of a Riemannian manifold from local boundary diffraction travel times *J. Math. Pures Appl.* **103** 830–48
[12] de Hoop M V, Ilmavirta J and Lassas M 2021 Reconstruction along a geodesic from sphere data in Finsler geometry and anisotropic elasticity (arXiv:2102.10383)
[13] de Hoop M V, Ilmavirta J, Lassas M and Saksala T 2019 Determination of a compact Finsler manifold from its boundary distance map and an inverse problem in elasticity (arXiv:1901.03902)
[14] de Hoop M V, Ilmavirta J, Lassas M and Saksala T 2020 A foliated and reversible Finsler manifold is determined by its broken scattering relation (arXiv:2003.12657)
[15] de Hoop M V and Saksala T 2019 Inverse problem of travel time difference functions on a compact riemannian manifold with boundary *J. Geom. Anal.* **29** 3308–27
[16] do Carmo M 1992 *Riemannian Geometry* (Birkhäuser)
[17] Duistermaat J J and Hormander L 1996 *Fourier Integral Operators* vol 2 (Springer)
[18] Feenstra J, Thurber C, Townend J, Roecker S, Bannister S and Boese C 2016 Microseismicity and P-wave tomography of the central Alpine Fault, New Zealand *N. Z. J. Geol. Geophys.* **59** 483–95
[19] Greenleaf A and Uhlmann G 1993 Recovering singularities of a potential from singularities of scattering data *Commun. Math. Phys.* **157** 549–72
[20] Gromov M 2007 *Metric Structures for Riemannian and non-Riemannian Spaces* (Springer Science & Business Media)
[21] Helmstetter A and Sornette D 2002 Diffusion of epicenters of earthquake aftershocks, omori's law and generalized continuous-time random walk models *Phys. Rev.* E **66** 061104

[22] Igonin N and Innanen K 2018 Analysis of simultaneous velocity and source parameter updates in microseismic FWI *SEG Technical Program Expanded Abstracts 2018* (Society of Exploration Geophysicists) p 5520

[23] Isham V 1981 An introduction to spatial point processes and Markov random fields *Int. Stat. Rev.* **49** 21–43

[24] Ivanov S 2013 Local monotonicity of Riemannian and Finsler volume with respect to boundary distances *Geom. Dedicata* **164** 83–96

[25] Ivanov S 2020 Distance difference functions on non-convex boundaries of Riemannian manifolds (arXiv:2008.13153)

[26] Ivanov S 2020 Distance difference representations of Riemannian manifolds *Geom. Dedicata* **207** 167–92

[27] Jansky J, Plicka V and Eisner L 2010 Feasibility of joint 1D velocity model and event location inversion by the Neighbourhood algorithm *Geophys. Prospect.* **58** 229–34

[28] Katchalov A, Kurylev Y and Lassas M 2001 *Inverse Boundary Spectral Problems* (*Monographs and Surveys in Pure and Applied Mathematics* vol 123) (Chapman & Hall/CRC)

[29] Katsuda A, Kurylev Y and Lassas M 2007 Stability of boundary distance representation and reconstruction of Riemannian manifolds *Inverse Problems Imaging* **1** 135

[30] Klingenberg W P 1995 *Riemannian Geometry* vol 1 (Walter de Gruyter)

[31] Kurylev Y 1997 Multidimensional Gel'fand inverse problem and boundary distance map *Inverse Problems Related With Geometry* ed H Soga (Ibaraki University Press) pp 1–15

[32] Kurylev Y, Lassas M and Uhlmann G 2010 Rigidity of broken geodesic flow and inverse problems *Am. J. Math.* **132** 529–62

[33] Kurylev Y, Lassas M, and Uhlmann G 2014 Inverse problems for Lorentzian manifolds and non-linear hyperbolic equations *Inventiones Mathematicae* (to appear)

[34] Kurylev Y, Oksanen L and Paternain G P 2018 Inverse problems for the connection Laplacian *J. Differ. Geom.* **110** 457–94

[35] Lassas M 2018 Inverse problems for linear and non-linear hyperbolic equations *Proc. Int. Congress of Mathematicians (2018 Rio de Janeiro)* vol 3 pp 3739–60

[36] Lassas M and Oksanen L 2014 Inverse problem for the Riemannian wave equation with Dirichlet data and Neumann data on disjoint sets *Duke Math. J.* **163** 1071–103

[37] Lassas M and Saksala T 2019 Determination of a Riemannian manifold from the distance difference functions *Asian J. Math.* **23** 173–200

[38] Lassas M, Saksala T and Zhou H 2018 Reconstruction of a compact manifold from the scattering data of internal sources *Inverse Problems Imaging* **12** 993–1031

[39] Lassas M, Uhlmann G and Wang Y 2018 Inverse problems for semilinear wave equations on lorentzian manifolds *Commun. Math. Phys.* **360** 1–55

[40] Lee J M 2018 *Introduction to Riemannian Manifolds* vol 2 (Springer)

[41] Meyerson R 2020 Stitching data: recovering a manifold's geometry from geodesic intersections (arXiv:2008.07642)

[42] Michel O and Tsvankin I 2017 Waveform inversion for microseismic velocity analysis and event location in VTI media *Geophysics* **82** WA95–103

[43] Michel R 1981 Sur la rigidité imposée par la longueur des géodésiques *Invent. Math.* **65** 71–83

[44] Milne T and Mansouri A-R 2019 Codomain rigidity of the Dirichlet to Neumann operator for the Riemannian wave equation *Trans. Am. Math. Soc.* **371** 8781–810

[45] Mönkkönen K 2020 Boundary rigidity for Randers metrics (arXiv:2010.11484)

[46] Myers S B and Steenrod N E 1939 The group of isometries of a Riemannian manifold *Ann. Math.* **40** 400–16

[47] Palais R S 1957 On the differentiability of isometries *Proc. Am. Math. Soc.* **8** 400–16

[48] Pestov L and Uhlmann G 2005 Two dimensional compact simple Riemannian manifolds are boundary distance rigid *Ann. Math.* **161** 1093–110

[49] Sasaki S 1958 On the differential geometry of tangent bundles of Riemannian manifolds *Tohoku Math. J.* **10** 338–54

[50] Shekar B and Nath A 2015 Gradient computation for simultaneous microseismic event location and velocity inversion by eikonal tomography *SEG Technical Program Expanded Abstracts 2015* (Society of Exploration Geophysicists) p 5634

[51] Song C and Alkhalifah T 2019 Microseismic event estimation based on an efficient wavefield inversion *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **12** 4664–71

[52] Song C, Alkhalifah T and Wu Z 2019 Microseismic event estimation and velocity analysis based on a source-focusing function *Geophysics* **84** KS85–94

[53] Stefanov P, Uhlmann G and Vasy A 2016 Boundary rigidity with partial data *J. Am. Math. Soc.* **29** 299–332

[54] Stefanov P, Uhlmann G and Vasy A 2017 Local and global boundary rigidity and the geodesic x-ray transform in the normal gauge (arXiv:1702.03638)

[55] Tan Y, He C and Mao Z 2018 Microseismic velocity model inversion and source location: the use of neighborhood algorithm and master station method *Geophysics* **83** KS49–63

[56] Uhlmann G 1998 Inverse boundary value problems for partial differential equations *Proc. Int. Congress of Mathematicians* pp 77–86

[57] van Lieshout M-C 2010 *Spatial Point Process Theory* (Chapman & Hall/CRC) pp 263–82

[58] Wang Y and Zhou T 2016 Inverse problems for quadratic derivative nonlinear wave equations *Communications in Partial differential equations* (to appear)

[59] Witten B and Jeffrey Shragge J 2017 Image-domain velocity inversion and event location for microseismic monitoring *Geophysics* **82** KS71–83

[60] Yajima T, Yamasaki K and Nagahama H 2011 Finsler metric and elastic constants for weak anisotropic media *Nonlinear Anal. Real World Appl.* **12** 3177–84

[61] Zhang Z, Rector J and Nava M 2017 Simultaneous inversion of multiple microseismic data for event locations and velocity model with Bayesian inference *Geophysics* **82** KS27–39