

Stability estimate for the broken non-abelian x-ray transform in Minkowski space

Simon St-Amant* 

Department of Pure Mathematics and Mathematical Statistics,
University of Cambridge, Cambridge CB3 0WB, United Kingdom

E-mail: sas242@cam.ac.uk

Received 7 February 2022, revised 28 June 2022

Accepted for publication 11 August 2022

Published 13 September 2022



CrossMark

Abstract

We study the broken non-abelian x-ray transform in Minkowski space. This transform acts on the space of Hermitian connections on a causal diamond and is known to be injective up to an infinite-dimensional gauge. We show a stability estimate that takes the gauge into account, leading to a new proof of the transform's injectivity. Our proof leads us to consider a special type of connections that we call light-sink connections. We then show that we can consistently recover a light-sink connection from noisy measurement of its x-ray transform data through Bayesian inversion.

Keywords: non-abelian x-ray transform, stability estimate, inverse problem for Yang–Mills equations, broken ray transform, gauge, light-sink connection, Bayesian inverse problem

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

*Author to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

1. Introduction and main results

We start by defining the broken non-abelian x-ray transform and provide the motivation for its study. We then state our main results. Sections 2 and 3 contain the proofs of those results, while section 4 concerns some of their statistical applications.

1.1. The broken non-abelian x-ray transform

Consider the causal diamond in Minkowski space $(\mathbb{R}^{1+3}, -dt^2 + dx_1^2 + dx_2^2 + dx_3^2)$ given by

$$\mathbb{D} := \{(t, x) \in \mathbb{R}^{1+3} : |x| \leq t + 1, |x| \leq 1 - t\}.$$

The origin's world line is $\mathcal{O} = (-1, 1) \times \{(0, 0, 0)\} \subset \mathbb{D}$. For $0 < \varepsilon \leq \varepsilon_0 < 1/2$, consider the ε -neighbourhood of \mathcal{O}

$$\mathcal{U}_\varepsilon := \{(t, x) \in \text{int } \mathbb{D} : |x| < \varepsilon\}.$$

We implicitly write \mathcal{U} for $\mathcal{U}_{\varepsilon_0}$ and write \mathcal{U}_ε whenever we want to emphasise the dependence on ε . Given $x, y \in \mathbb{D}$, we write $x < y$ if there is a future-pointing causal curve from x to y . We also write $(x, y) \in \mathbb{L}$ if $x < y$ and there is a lightlike geodesic from x to y . See figure 1 for a visual representation of \mathbb{D} , \mathcal{U} and \mathcal{O} in \mathbb{R}^{1+2} .

Recall that a line segment $\gamma : [0, T] \rightarrow \mathbb{R}^{1+3}$, $\gamma(s) = x + sv$ for $x \in \mathbb{R}^{1+3}$, $v = (v_0, v_1, v_2, v_3)$ is a lightlike geodesic if

$$v_0^2 = v_1^2 + v_2^2 + v_3^2$$

and that it is future-pointing if $v_0 > 0$ and past-pointing if $v_0 < 0$. We say that γ is parametrised by arc length if $|v|_{\mathbb{R}^4} = 1$. The set of points $y \in \mathbb{R}^{1+3}$ such that there is a future-pointing (past-pointing) lightlike geodesic from x to y is called the future (past) light cone at x . Hence, $(x, y) \in \mathbb{L}$ if and only if y is in the future light cone of x , or equivalently, x is in the past light cone of y .

We will work with Hermitian connections on the trivial bundle $\mathbb{D} \times \mathbb{C}^n$. Such a connection A is a $u(n)$ -valued one-form on \mathbb{D} and we can write it as

$$A = A_0 dt + A_1 dx_1 + A_2 dx_2 + A_3 dx_3$$

for some matrix fields $A_i \in C^\infty(\mathbb{D}, u(n))$. We denote the set of Hermitian connections on \mathbb{D} by \mathcal{U} . A connection induces a covariant derivative on functions $f : \mathbb{D} \rightarrow \mathbb{C}^n$ given by $d_A f = df + Af$. Given a smooth curve $\gamma : [0, T] \rightarrow \mathbb{D}$, the parallel transport isomorphism $P_\gamma^A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ is given by the solution of the matrix ODE

$$\begin{cases} \dot{U}(t) + A(\dot{\gamma}(t))U(t) = 0, \\ U(0) = \text{Id}, \end{cases}$$

at time T . Hence, the parallel transport of a vector $v \in \mathbb{C}^n$ along γ is $P_\gamma^A v := U(T)v$. One can check that P_γ^A does not depend on the parametrisation of γ and that it takes values in $U(n)$ since A is Hermitian. Given $x, y \in \mathbb{D}$, we denote by $P_{y \leftarrow x}^A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ the parallel transport from x to y along the straight line between the two points. The direction of the arrow in the notation is chosen as to behave nicely with compositions.

We can now define the broken non-abelian x-ray transform. In [CLOP21a] and [CLOP21b], they define it as follows. Consider the set

$$\mathbb{S}^+(\mathcal{U}) := \{(x, y, z) \in \mathbb{D}^3 : (x, y), (y, z) \in \mathbb{L}, x < y < z \text{ with } x, z \in \mathcal{U}, y \notin \mathcal{U}\}.$$

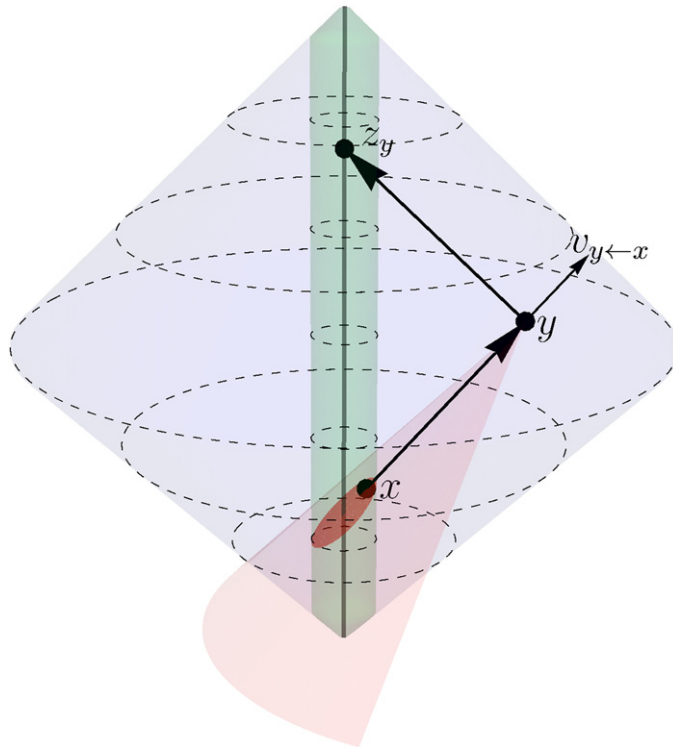


Figure 1. Setting for the broken non-abelian x-ray transform in \mathbb{R}^{1+2} for a future-determined path. The point y lies inside the causal diamond \mathbb{D} (in blue), but outside the set \mathcal{U} (in green). The point x can take values in the fibre \mathcal{F}_y^x which is given by the intersection of \mathcal{U} and the past light cone at y (in red). The point z_y is always taken on the origin's world line and is uniquely determined by y . The vector $v_{y \leftarrow x}$ is based at y and points in the direction coming from x .

This set is comprised of light rays starting from $x \in \mathcal{U}$ that exit \mathcal{U} and break at $y \notin \mathcal{U}$ before returning to \mathcal{U} at z . One such ray is shown in figure 1. We denote by

$$\mathcal{U}^X := \bigcup_{(x,y,z) \in \mathbb{S}^+(\mathcal{U})} \{x\} \quad \text{and} \quad \mathcal{U}^Z := \bigcup_{(x,y,z) \in \mathbb{S}^+(\mathcal{U})} \{z\}$$

the sets of values that x and z can take in \mathcal{U} , respectively. It is important to note that neither \mathcal{U}^X or \mathcal{U}^Z cover \mathcal{U} , but that $\mathcal{U} = \mathcal{U}^X \cup \mathcal{U}^Z$. Given a Hermitian connection A as above, its broken non-abelian x-ray transform is the unitary map

$$S^A_{z \leftarrow y \leftarrow x} := P^A_{z \leftarrow y} P^A_{y \leftarrow x} : \mathbb{C}^n \rightarrow \mathbb{C}^n \quad (x, y, z) \in \mathbb{S}^+(\mathcal{U}).$$

We are interested in recovering the connection A from its scattering data S^A . However, the map $A \mapsto S^A$ is not injective as it has a gauge given by the following right group action. For $\varphi \in C^\infty(\mathbb{D}, U(n))$, we denote

$$A \triangleleft \varphi := \varphi^{-1} d\varphi + \varphi^{-1} A \varphi.$$

The next proposition, whose proof is straightforward, states that the action of φ on the connection amounts to a conjugation of the parallel transports.

Proposition 1.1. *Let A be a connection on \mathbb{D} and let $\varphi \in C^\infty(\mathbb{D}, U(n))$. Then, for any smooth curve $\gamma : [0, T] \rightarrow \mathbb{D}$,*

$$P_\gamma^{A \triangleleft \varphi} = \varphi(\gamma(T))^{-1} P_\gamma^A \varphi(\gamma(0)).$$

In particular, if $\varphi|_{\mathcal{U}} = \text{Id}$, then

$$S_{z \leftarrow y \leftarrow x}^{A \triangleleft \varphi} = S_{z \leftarrow y \leftarrow x}^A$$

for all $(x, y, z) \in \mathbb{S}^+(\mathcal{U})$.

Therefore, the scattering data of $A \triangleleft \varphi$ coincides with that of A whenever φ is in the gauge group

$$\mathcal{G} := \{\varphi \in C^\infty(\mathbb{D}, U(n)) : \varphi|_{\mathcal{U}} = \text{Id}\}.$$

This natural obstruction to recovering A from S^A turns out to be the only one. Indeed, it is shown in [CLOP21a, theorem 5] that Hermitian connections A and B share the same scattering data if and only if they are in the same gauge orbit, that is, there exists $\varphi \in \mathcal{G}$ such that $B = A \triangleleft \varphi$.

Our goal is to find a stability estimate relating the scattering data of two connections A and B with some measure of distance between them in a gauge invariant way. In other words, we want to show that A and B must be relatively similar whenever S^A and S^B are close.

1.2. The non-abelian x-ray transform and broken Radon transform

The usual non-abelian x-ray transform assigns to a matrix field $A \in C^\infty(\mathbb{R}^d \times \mathbb{S}^{d-1}, \mathbb{C}^{n \times n})$ the scattering data map

$$(x, \theta) \in \mathbb{R}^d \times \mathbb{S}^{d-1} \mapsto \lim_{s \rightarrow \infty} \psi^+(x + s\theta, \theta) \in \mathbb{C}^{n \times n}$$

where ψ^+ is the unique solution of the transport equation

$$\sum_{i=1}^d \theta_i \partial_{x_i} \psi + A(x, \theta) \psi = 0, \quad x \in \mathbb{R}^d, \quad \theta \in \mathbb{S}^{d-1},$$

such that

$$\lim_{s \rightarrow -\infty} \psi^+(x + s\theta, \theta) = \text{Id}.$$

Given that A decays sufficiently fast as $|x| \rightarrow \infty$, the transform is well-defined and one can ask whether it is possible to recover A from the scattering data. The non-abelian x-ray transform has been studied extensively in the last 20 years and has applications in many different types of tomographies, such as single-photon emission computed tomography or neutron polarisation tomography. See [Nov19] for a recent survey on the non-abelian x-ray transform and its applications.

The non-abelian x-ray transform has also been studied on simple surfaces [PS20, MNP21] and compact manifolds with strictly convex boundary [Boh21] where the transport equation is now solved along unit-speed geodesics with endpoints on the boundary of the manifold. For more details and background on the two-dimensional problem, see [PSU21].

When $n = 1$, the broken non-abelian x-ray transform is also called the broken-ray Radon transform. In [FMS11], they consider the broken-ray Radon transform with rays breaking at a fixed angle within a slab and provide an inversion formula. The broken-ray Radon transform has applications in optical tomography, see [AS09] for a survey. The V-line Radon transform [Amb12, ALJ19] is another example of an inverse problem making use of broken rays and has applications in imaging.

1.3. Physical motivation

The broken non-abelian x-ray transform has been introduced in [CLOP21a] where they began to analyse inverse problems for the Yang–Mills–Higgs equations. They show that one can recover a Hermitian connection A from the source-to-solution map L_A taking a source $f \in C_c^4(\mathcal{U}, \mathbb{C}^n)$ to

$$L_A f = \phi|_{\mathcal{U}}$$

where ϕ solves

$$\begin{cases} \square_A \phi + |\phi|^2 \phi = f & \text{in } (-1, 2) \times \mathbb{R}^3, \\ \phi|_{t < -1} = 0. \end{cases} \quad (1)$$

Here \square_A is the connection wave operator given by $\square_A = d_A^* d_A$. Note that when $A = 0$, we recover the usual wave operator $\square = \partial_t^2 - \Delta$. The map L_A is well-defined as long as f is sufficiently small. They show that the maps L_A and L_B agree if and only if A and B are gauge equivalent. To do so, they first show that L_A determines the broken non-abelian x-ray transform $S_{z \leftarrow y \leftarrow x}^A$ for all $(x, y, z) \in \mathbb{S}^+(\mathcal{U})$. Injectivity up to gauge of L_A then follows from that of the broken x-ray transform.

To determine $S_{z \leftarrow y \leftarrow x}^A$ from the source-to-solution map L_A , they construct a source of the form

$$f = \epsilon_1 f_1 + \epsilon_2 f_2 + \epsilon_3 f_3$$

where each f_j is a conormal distribution supported near $x \in \mathcal{U}$. Let ϕ be the solution of (1) corresponding to such an f . The functions $\partial_{\epsilon_j} \phi|_{\epsilon_j=0}$ satisfy a wave equation and, when the sources are chosen carefully, can produce an artificial source at y which emits a singular wave front that reaches z . This interaction is encoded in the operator $f \mapsto \partial_{\epsilon_1} \partial_{\epsilon_2} \partial_{\epsilon_3} \phi|_{\epsilon=0}$, whose principal symbol determines $S_{z \leftarrow y \leftarrow x}^A$. The creation of an artificial source is only possible thanks to the nonlinearity in (1) and shows how one can exploit nonlinearities in an advantageous way, similar to what is shown in [KLU18].

1.4. Statistical motivation

The second motivation for considering the broken non-abelian x-ray transform is to use it as an example for dealing with injectivity issues that arise in the study of Bayesian inverse problems. We give a summary to the Bayesian approach to solving inverse problems in section 4, as introduced in [Stu10].

Let $\mathcal{G} : \Theta \rightarrow Y$ be a mapping between Banach spaces that we refer to as the forward map. We say that Θ is the parameter space while Y is the sample space. For fixed $y \in Y$, the inverse problem related to \mathcal{G} consists of finding $\theta_* \in \Theta$ such that

$$y = \mathcal{G}(\theta_*).$$

The Bayesian approach to inverse problems aims to find a suitable candidate $\hat{\theta}$ for θ_* through Bayesian statistics on Θ corresponding to noisy measurements of y . See section 4 for more details.

In our setting, the set Θ could be the set of Hermitian connections A on \mathbb{D} , Y the set of matrix fields on $\mathbb{S}^+(\mathcal{U})$ and $\theta \mapsto \mathcal{G}_\theta$ the mapping that sends a connection A to its scattering data S^A .

However, in the case of the broken non-abelian x-ray transform, the map $\mathcal{G} : A \mapsto S^A$ is not injective and so the true underlying parameter θ_* is not uniquely identifiable. Indeed, all connections in the same \mathcal{G} -orbit yield the same scattering data. Moreover, these orbits are all infinite dimensional as well as non-linear. Can we still find a way to get a meaningful candidate for A from samples of S^A through the framework of Bayesian inverse problems?

The first approach one could use to deal with injectivity issues is as follows. Let us assume, as is our case, that a group G acts on Θ and that \mathcal{G} is injective up to the action of G . This means that for every $g \in G$ and $\theta \in \Theta$, we have $\mathcal{G}(\theta \triangleleft g) = \mathcal{G}(\theta)$ and that $\mathcal{G}(\theta_1) = \mathcal{G}(\theta_2)$ if and only if there is $g' \in G$ with $\theta_1 = \theta_2 \triangleleft g'$. Then \mathcal{G} naturally induces an injective map on the quotient space

$$\tilde{\mathcal{G}} : \Theta/G \rightarrow Y.$$

One could try to prove statistical guarantees for this map. However, as settings such as the present one where \mathcal{G} is non-linear, the quotient space Θ/G is intractable as it is unclear how one would parametrise the equivalence classes. What one needs is a choice of representative for each class in the quotient, that is, a continuous map $s : \Theta/G \rightarrow \Theta$ such that the following diagram commutes.

The existence of s is nontrivial and it is often the case that such a lift simply does not exist, see [Sin78] for examples where topological obstructions prevent its existence. And even if s exists, it might only be theoretical and not correspond to an explicit choice (not constructive or numerically computable). Hence, we need a new approach that is adapted to the problem we want to consider.

What we will end up doing is finding another group H of which G is a proper subgroup and for which we can find an explicit section $s_H : \Theta/H \rightarrow \Theta$. Although the forward map will not be invariant under the action of H , our stability estimates will. Those same estimates will guarantee that the forward map is injective when restricted to the image of s_H . We will then show that we can use Bayesian inversion to solve this restricted problem. Finally, through some choice of extension operator, we will show that, from the solution to the restricted problem, we can recover an element that is G -equivalent to the true solution θ_* . See the discussion after proposition 1.9 for more details.

1.5. Definitions and notation

Before presenting the main results, we use this section to gather some notation and additional definitions that will be used throughout.

Unlike in [CLOP21a], we will not consider all paths in $\mathbb{S}^+(\mathcal{U})$. We will mostly consider two types of paths that we refer to as **past-determined** and **future-determined** paths. A past-determined path is a path of the form $z \leftarrow y \leftarrow x_y$ for $(x_y, y, z) \in \mathbb{S}^+(\mathcal{U})$ where x_y is the unique point such that $(x_y, y) \in \mathbb{L}$ and $x_y \in \mathcal{O}$, that is, $x_y = (t, 0, 0, 0)$ for some $t \in [-1, 1]$. Similarly, a future-determined path is a path of the form $z_y \leftarrow y \leftarrow x$ for $(x, y, z_y) \in \mathbb{S}^+(\mathcal{U})$ where now z_y is the unique point on \mathcal{O} such that $(y, z_y) \in \mathbb{L}$. We denote the corresponding scattering data as

$$S^A_{z \leftarrow y \leftarrow x_y} \quad \text{and} \quad S^A_{z_y \leftarrow y \leftarrow x}.$$

Hence, the wiggle room in \mathcal{U} will only be used to move x in \mathcal{U}^X or z in \mathcal{U}^Z , but not both.

Remark 1.2. It is not sufficient to only consider paths that are both past-determined and future-determined, that is, paths of the form $z_y \leftarrow y \leftarrow x_y$. Indeed, in polar coordinates (t, r, ϑ, ϕ) , the tangent vector along the path $z_y \leftarrow y$ is $\frac{1}{\sqrt{2}}(\frac{\partial}{\partial t} - \frac{\partial}{\partial r})$ while the tangent vector along the path $y \leftarrow x_y$ is $\frac{1}{\sqrt{2}}(\frac{\partial}{\partial t} + \frac{\partial}{\partial r})$. Hence, the angular components of the connection play no role in the forward problem for such paths.

Any future-determined path can be identified by its break point $y \in \mathbb{D} \setminus \mathcal{U}$ and its first end-point x which lies in the intersection between \mathcal{U}^X and the past light cone of y . We can represent the admissible future-determined paths as a B_3 -bundle $\pi : \mathcal{F}^X \rightarrow \mathbb{D} \setminus \mathcal{U}$, where B_3 stands for the unit ball in \mathbb{R}^3 . For every $y \in \mathbb{D} \setminus \mathcal{U}$, the fibre is given by

$$\mathcal{F}_y^X = \{x \in \mathcal{U}^X : (x, y) \in \mathbb{L}\} \cong B_3.$$

See figure 1 where the fibre \mathcal{F}_y^X is represented in dark red. Similarly, the set of admissible past-determined paths can be represented through the B_3 -bundle $\pi : \mathcal{F}^Z \rightarrow \mathbb{D} \setminus \mathcal{U}$ with fibre

$$\mathcal{F}_y^Z = \{z \in \mathcal{U}^Z : (y, z) \in \mathbb{L}\} \cong B_3.$$

We write $\mathcal{F}_\varepsilon^X$ or $\mathcal{F}_\varepsilon^Z$ whenever we want to emphasise the dependence of the bundles on ε through \mathcal{U}_ε .

For two different points x and y in \mathbb{D} , let $\gamma_{y \leftarrow x} : [0, T] \rightarrow \mathbb{D}$ be the straight line from x to y parametrised by its (Euclidean) arc length. We denote

$$v_{y \leftarrow x} := \dot{\gamma}_{y \leftarrow x}(T) = \frac{y - x}{|y - x|_e},$$

that is, $v_{y \leftarrow x}$ is the unit length vector pointing from x to y , but based at y (see figure 1). For a function $\Phi : \mathbb{D}^2 \setminus \Gamma \rightarrow \mathbb{C}^n$ where Γ is the diagonal of \mathbb{D}^2 , we define the differential operator

$$(\partial\Phi)(x, y) = \partial_{y \leftarrow x} \Phi(x, y) := \left. \frac{d}{dt} \Phi(x, y + tv_{y \leftarrow x}) \right|_{t=0}. \tag{2}$$

This is simply the directional derivative of Φ at y in the direction $v_{y \leftarrow x}$. Note that if $(x, y) \in \mathcal{F}^X$ and the domain of Φ is \mathcal{F}^X , the operators $\partial_{y \leftarrow x}$ and $\partial_{x \leftarrow y}$ are both well-defined since $x \in \mathcal{F}_{y+tv_{y \leftarrow x}}^X$ and $x + tv_{x \leftarrow y} \in \mathcal{F}_y^X$ whenever $x \in \mathcal{F}_y^X$ and t is sufficiently small. One can see $\partial_{y \leftarrow x}$ and $\partial_{x \leftarrow y}$ as horizontal and vertical vector fields on \mathcal{F}^X , respectively. Similarly, $\partial_{y \leftarrow z}$ and $\partial_{z \leftarrow y}$ are well defined operators if $(y, z) \in \mathcal{F}^Z$ and the domain of Φ is \mathcal{F}^Z .

We define the L^2 -norm of a function $\Phi : \mathcal{F}^X \rightarrow \mathbb{C}^n$ as

$$\|\Phi\|_{L^2(\mathcal{F}^X)} = \left(\int_{\mathbb{D} \setminus \mathcal{U}} \int_{\mathcal{F}_y^X} |\Phi(x, y)|^2 dx dy \right)^{1/2}$$

where dx is the natural measure on \mathcal{F}_y^X induced by Euclidean space. Note that this norm scales down as ε goes to 0 at a rate of ε^3 .

Given a linear map \mathcal{T} from \mathbb{R}^m to \mathbb{C}^n , we will denote its operator norm as

$$\|\mathcal{T}\| := \sup_{v \in \mathbb{R}^m \setminus \{0\}} \frac{|\mathcal{T}(v)|}{|v|}$$

where $|\cdot|$ denotes the usual norm on \mathbb{R}^m or \mathbb{C}^n . This induces a pointwise norm on $\mathbb{C}^{n \times n}$ -valued one-forms ω on \mathbb{D} at any given point $y \in \mathbb{D}$ by seeing ω_y as a mapping from \mathbb{R}^4 to \mathbb{C}^{n^2} . Note

that we then have $|\omega_y(v)| = \text{Tr}([\omega_y, \omega_y^*](v))$ and so $\|\omega_y\|$ is invariant under the action of $U(n)$. This also induces an L^2 -norm on the space of one-forms by

$$\|\omega\|_{L^2(\mathbb{D})} = \left(\int_{\mathbb{D}} \|\omega_y\|^2 dy \right)^{1/2}.$$

Given two connections A and B and a matrix field $Q \in C^\infty(\mathbb{D}, \mathbb{C}^{n \times n})$, we define $E(A, B) \in C^\infty(\mathbb{D}, \text{End}(\mathbb{C}^{n \times n}))$ by $E(A, B)Q = AQ - QB$. If A and B are Hermitian, then so is $E(A, B)$, in the sense that $E(A, B)^* = -E(A, B)$.

1.6. Main results

We state our results only for future-determined paths, but equivalent statements hold for past-determined paths by seeing $[S_{z \leftarrow y \leftarrow x}^A]^{-1}$ as a future-determined path. We will first show the stability estimates below for the values of a connection inside and outside \mathcal{U} .

Theorem 1.3. *Let A and B be Hermitian connections on \mathbb{D} . There is a constant $C > 0$ independent of ε such that*

$$\|A - B\|_{L^2(\mathcal{U}_\varepsilon^X)} \leq C \|\partial_{x \leftarrow y} (S_{z \leftarrow y \leftarrow x}^A [S_{z \leftarrow y \leftarrow x}^B]^{-1})\|_{L^2(\mathcal{F}_\varepsilon^X)}.$$

Theorem 1.4. *Let A and B be Hermitian connections. There exists a smooth function $p \in C^\infty(\mathbb{D}, \mathbb{C}^{n \times n})$ vanishing on \mathcal{O} and $C > 0$ such that for all $0 < \varepsilon < \varepsilon_0$,*

$$\|A - B - d_{E(A,B)}p\|_{L^2(\mathbb{D} \setminus \mathcal{U}_\varepsilon)} \leq \frac{C}{\varepsilon^4} \|\partial_{y \leftarrow x} ([S_{z \leftarrow y \leftarrow x}^A]^{-1} S_{z \leftarrow y \leftarrow x}^B)\|_{L^2(\mathcal{F}_\varepsilon^X)}. \quad (3)$$

By combining both theorems we can get a new proof of the injectivity (up to the gauge \mathcal{G}) of the broken non-abelian x-ray transform.

Corollary 1.5. *Let A and B be Hermitian connections. Then S^A and S^B agree for all past-determined and future-determined paths if and only if A and B are gauge equivalent.*

Proof. Since the scattering data of A and B agree for all future-determined paths, theorem 1.3 implies that A and B must agree on \mathcal{U}^X . By the same estimate for past-determined paths, the connections must also agree on \mathcal{U}^Z and hence they agree on \mathcal{U} . Theorem 1.4 yields $p \in C^\infty(\mathbb{D}, \mathbb{C}^{n \times n})$ such that

$$B = A - d_{E(A,B)}p = A - dp - Ap + pB \quad (4)$$

on $\mathbb{D} \setminus \mathcal{U}$. Let $\varphi = \text{Id} - p$. As the proof of theorem 2.2 will reveal, φ takes values in $U(n)$ since actually $\varphi = P_{y \leftarrow z, y}^A P_{z \leftarrow y, y}^B$. We can rewrite (4) as

$$\varphi B = A\varphi + d\varphi$$

and so $B = A \triangleleft \varphi$. It follows that A and B are gauge equivalent since they agree on \mathcal{U} and so $\varphi|_{\mathcal{U}} = \text{Id}$. The converse implication is the statement of proposition 1.1. \square

This can be seen as a partial data result improving on theorem 5 in [CLOP21a] as we only considered past-determined and future-determined paths. It also suggests that always taking such paths might be a more efficient problem to study.

Both stability estimates are invariant under the action of \mathcal{G} , but they are also invariant under the action of the bigger group

$$\mathcal{H} := \{\varphi \in C^\infty(\mathbb{D}, U(n)) : \varphi|_{\mathcal{O}} = \text{Id}\}.$$

In fact, we can rewrite the left-hand side of (3) in a way that highlights this.

Theorem 1.6. *Let p be as in theorem 1.4. Then for all $y \in \mathbb{D} \setminus \mathcal{U}$,*

$$\|(A - B - d_{E(A,B)}p)_y\| = \|(A \triangleleft P_{y \leftarrow z_y}^A)_y - (B \triangleleft P_{y \leftarrow z_y}^B)_y\|.$$

Hence, this defines a distance between the connections A and B that is invariant under the action of \mathcal{H} , and so gauge independent as $\mathcal{G} \subset \mathcal{H}$. With a little bit of work, we can combine this expression with theorem 1.4 to get the following H^1 estimate.

Corollary 1.7. *Let A and B be Hermitian connections. There exists a constant $C > 0$ such that*

$$\|(A \triangleleft P_{y \leftarrow z_y}^A) - (B \triangleleft P_{y \leftarrow z_y}^B)\|_{L^2(\mathbb{D} \setminus \mathcal{U}_\varepsilon)} \leq \frac{C\Psi(A, B)}{\varepsilon^4} \|S^A - S^B\|_{H^1(\mathcal{F}_\varepsilon^X)}$$

where

$$\Psi(A, B) = 1 + \min\{\|F_A\|_{L^\infty(\mathbb{D})} + \|A(\partial_t)\|_{L^\infty(\mathcal{O})}, \|F_B\|_{L^\infty(\mathbb{D})} + \|B(\partial_t)\|_{L^\infty(\mathcal{O})}\}$$

and F_A is the curvature two-form of A .

Theorem 1.6 also suggests we should naturally try to fix the gauge by considering connections such that $A \triangleleft P_{y \leftarrow z_y}^A = A$. We call them **light-sink** connections. They form a linear space \mathcal{L} and we can characterise them, see proposition 3.2.

Proposition 1.8. *Every connection A is \mathcal{H} -equivalent to a unique light-sink connection and the map $\rho: \mathcal{U} / \mathcal{H} \rightarrow \mathcal{L}$,*

$$\rho([A]) := A \triangleleft P_{y \leftarrow z_y}^A$$

is well-defined.

The map ρ is almost a fixing of the gauge. Contrary to that of \mathcal{G} , the action of \mathcal{H} on \mathcal{U} does not preserve the scattering data. Therefore, the map ρ does not define a lift as we defined it in section 1.4. Nonetheless, if a light-sink connection A is \mathcal{H} -equivalent to another connection B , we can use their scattering data and the map ρ to make them gauge equivalent.

Proposition 1.9. *Let A be a light-sink connection and let B be a Hermitian connection such that*

$$A = B \triangleleft P_{y \leftarrow z_y}^B. \quad (5)$$

From the past-determined and future-determined scattering data of A and B , we can find a map $\Phi \in \mathcal{H}$ such that $A \triangleleft \Phi$ and B are gauge equivalent (with respect to \mathcal{G}).

The map Φ is defined up to an extension operator

$$\mathcal{E}: C^\infty(\mathcal{U}, U(n)) \hookrightarrow C^\infty(\mathbb{D}, U(n)).$$

We have reduced the choice of a gauge to the choice of an extension operator \mathcal{E} . Note that such an operator can be constructed by first extending with values in $GL(n, \mathbb{C})$ and then projecting onto $U(n)$ through a strong deformation retract (a continuous map $F: [0, 1] \times GL(n, \mathbb{C}) \rightarrow GL(n, \mathbb{C})$ such that $F(0, x) = x$ and $F(1, x) \in U(n)$ for all $x \in GL(n, \mathbb{C})$, and $F(t, \cdot)|_{U(n)} = \text{Id}$ for all $t \in [0, 1]$).

In practice, say that we observe the scattering data S^B on past-determined and future-determined paths for some connection B and that we have complete knowledge of the forward

map $A \mapsto S^A$. We wish to find the gauge equivalence class of B from S^B , which amounts to finding a connection A such that $A = B \triangleleft \varphi$ for some $\varphi \in \mathcal{G}$. Our results give the following strategy to do so.

- (a) By taking y on the boundary of \mathcal{U} , use theorem 1.3 to determine B inside \mathcal{U} from the scattering data of B along past-determined and future-determined paths.
- (b) Minimise the mapping

$$A \mapsto \|S_{z_y \leftarrow y \leftarrow x}^A - S_{z_y \leftarrow y \leftarrow x}^B P_{x \leftarrow z_x}^B\|_{H^1(\mathcal{F}_\varepsilon^x)}$$

over all light-sink connections A . Note that we can compute $P_{x \leftarrow z_x}^B$ from the first step since we know B inside \mathcal{U} .

- (c) By corollary 1.7 and the definition of ρ , the unique minimiser of this problem is $A = \rho([B])$.
- (d) Use proposition 1.9 to get a connection $A \triangleleft \Phi$ that is gauge-equivalent to B .

Note that in step (b), $S_{z_y \leftarrow y \leftarrow x}^B P_{x \leftarrow z_x}^B$ is precisely the scattering data of $\rho([B])$, which explains why corollary 1.7 implies that $A = \rho([B])$ is the unique minimiser of the problem.

One can implement this algorithm with the use of Bayesian inversion. Step (b) is equivalent to recovering a light-sink connection from its scattering data and we will show in section 4 that we can consistently do so through Bayesian inversion, see theorem 4.2. Using similar arguments, one could also provide guarantees for recovering B on \mathcal{U} in step (a) using Bayesian inversion. As steps (c) and (d) are only simple direct computations, the above algorithm fits within the framework of Bayesian inverse problems. Therefore, by following these steps, one should be able to compute a connection that is close to being gauge-equivalent to B from noisy measurements of its scattering data.

2. Stability estimate

The goal of this section is to prove the following two pointwise estimates from which theorems 1.3 and 1.4 will follow.

Theorem 2.1. *Let A and B be Hermitian connections on \mathbb{D} . Then, there is a constant $C > 0$ such that for all $x \in \mathcal{U}^x$,*

$$\|(A - B)_x\| \leq C \sup_{\substack{y \in \mathbb{D} \setminus \mathcal{U} \\ (x,y) \in \mathbb{L}}} |\partial_{x \leftarrow y} (S_{z_y \leftarrow y \leftarrow x}^A [S_{z_y \leftarrow y \leftarrow x}^B]^{-1})|.$$

Theorem 2.2. *Let A and B be Hermitian connections on \mathbb{D} . There exists a smooth function $p \in C^\infty(\mathbb{D}, \mathbb{C}^{n \times n})$ vanishing on \mathcal{O} and $C > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ and $y \in \mathbb{D} \setminus \mathcal{U}_\varepsilon$, it holds that*

$$\|(A - B - d_{E(A,B)} p)_y\| \leq \frac{C}{\varepsilon^4} \int_{(\mathcal{F}_\varepsilon^x)_y} |\partial_{y \leftarrow x} ([S_{z_y \leftarrow y \leftarrow x}^A]^{-1} S_{z_y \leftarrow y \leftarrow x}^B)| \, dx. \tag{6}$$

To do so, we introduce the attenuated x-ray transform, as well as a pseudolinearisation identity. We also show how to reformulate the theorems in the form of an H^1 estimate.

2.1. The attenuated x-ray transform

Let $\gamma : [0, T] \rightarrow \mathbb{D}$ be a smooth curve and let $\omega \in \Omega^1(\mathbb{D}, \mathbb{C}^n)$, that is, ω is a one-form on \mathbb{D} with values in \mathbb{C}^n (we will actually use $\mathbb{C}^{n \times n}$ in the proof of theorem 2.2, but everything will be

defined analogously through the isomorphism with \mathbb{C}^{n^2}). Fix a Hermitian connection A on \mathbb{D} as above. The attenuated x-ray transform of ω along γ with respect to A is given by

$$I_\gamma^A(\omega) := \int_0^T P_{\gamma(0)\leftarrow\gamma(t)}^A \omega(\dot{\gamma}(t)) dt. \quad (7)$$

Similar to the parallel transport, we can express $I_\gamma^A(\omega)$ as the solution of a matrix ODE.

Lemma 2.3. *Let u be the unique solution along $\gamma : [0, T] \rightarrow \mathbb{D}$ of the matrix ODE*

$$\begin{cases} \dot{u} + A(\dot{\gamma}(t))u = -\omega(\dot{\gamma}(t)), \\ u(0) = u_0. \end{cases}$$

Then $u(T) = P_\gamma^A(u_0 - I_\gamma^A(\omega))$.

Proof. Let U solve

$$\begin{cases} \dot{U} + A(\dot{\gamma}(t))U = 0, \\ U(0) = \text{Id}. \end{cases}$$

A quick computation shows that $(U^{-1})^\cdot = U^{-1}A$. Therefore, along γ we have

$$(U^{-1}\dot{u}) = U^{-1}Au + U^{-1}\dot{u} = U^{-1}(Au + \dot{u}) = -U^{-1}\omega.$$

Integrating both sides from 0 to T yields

$$U^{-1}(T)u(T) - U^{-1}(0)u(0) = \int_0^T (U^{-1}\dot{u})(t) dt = - \int_0^T U^{-1}(t)\omega(\dot{\gamma}(t)) dt.$$

By definition of the parallel transport, $U(t) = P_{\gamma(0)\leftarrow\gamma(t)}^A$. Isolating $u(T)$ in the previous equation and replacing U by the parallel transport yields the result. \square

If A vanishes identically, the attenuated x-ray is simply the integral of the one-form ω along γ , and so if ω is potential ($\omega = df$ for some $f \in C^\infty(\mathbb{D}, \mathbb{C}^n)$), then the attenuated x-ray of ω is the difference between the values of f at both endpoints of γ by the fundamental theorem of calculus. This is not exactly true when A does not vanish as we have to account for the parallel transport in the definition of $I_\gamma^A(\omega)$. Instead of potential forms with respect to d , we actually have to consider potential forms with respect to $d_A = d + A$ to get an analog of the fundamental theorem of calculus.

Lemma 2.4. *Let $f : \mathbb{D} \rightarrow \mathbb{C}^n$ be a smooth function on \mathbb{D} . Then*

$$I_\gamma^A(d_A f) = (P_\gamma^A)^{-1} f(\gamma(T)) - f(\gamma(0))$$

where $d_A f = df + Af$.

Recall the definition of $\partial_{y \leftarrow x}$ as in (2). We can apply $\partial_{y \leftarrow x}$ to the attenuated x-ray to evaluate the values of a one-form from the tangent space at y .

Lemma 2.5. *Let ω be a one-form on \mathbb{D} . For $x \neq y$,*

$$\partial_{y \leftarrow x}(I_{y \leftarrow x}^A(\omega)) = P_{x \leftarrow y}^A(\omega_y(v_{y \leftarrow x})).$$

Proof. Let $\gamma : [0, T]$ be the line segment from x to y parametrised by arclength. By extending γ , we see that $\gamma(s) + tv_{y \leftarrow x} = \gamma(s + t)$. Hence, we get

$$\begin{aligned} \partial_{y \leftarrow x} I_{y \leftarrow x}^A(\omega) &= \frac{d}{dt} (I_{y+tv \leftarrow x}^A(\omega)) \Big|_{t=0} \\ &= \frac{d}{dt} \left(\int_0^{T+t} P_{x \leftarrow \gamma(s)}^A \omega_{\gamma(s)} (\dot{\gamma}(s)) ds \right) \Big|_{t=0} \\ &= P_{x \leftarrow y}^A \omega_y (\dot{\gamma}(T)) \end{aligned}$$

since $\gamma(T) = y$. The result follows since $v_{y \leftarrow x} = \dot{\gamma}(T)$. □

2.2. *The broken attenuated x-ray transform*

We will actually be interested in a broken version of the attenuated x-ray transform. One could define naively the broken attenuated x-ray transform $I_{z \leftarrow y \leftarrow x}^A(\omega)$ as $I_{y \leftarrow x}(\omega) + I_{y \leftarrow z}(\omega)$. However, this is not compatible with lemma 2.4 as we would want

$$I_{z \leftarrow y \leftarrow x}^A(d_A f) = P_{x \leftarrow y}^A P_{y \leftarrow z}^A f(z) - f(x) \tag{8}$$

to hold in general. It also does not coincide with the usual attenuated x-ray transform $I_{z \leftarrow x}^A(\omega)$ if x, y and z lie on the same line in order. Instead, we need to define the broken attenuated x-ray transform as

$$I_{z \leftarrow y \leftarrow x}^A(\omega) := I_{y \leftarrow x}^A(\omega) + P_{x \leftarrow y}^A I_{z \leftarrow y}^A(\omega). \tag{9}$$

One can check that (8) holds under this definition and $I_{z \leftarrow y \leftarrow x}^A(\omega) = I_{z \leftarrow x}^A(\omega)$ whenever the curve $z \leftarrow y \leftarrow x$ is smooth.

2.3. *Pseudolinearisation identity*

The key tool in the proofs of theorems 2.1 and 2.2 is the following pseudolinearisation identity. It relates parallel transports along a curve with respect to two different connections with an attenuated x-ray of their difference. See [PSU21, chapter 13.2] for more details on the pseudolinearisation identity. We shall adapt their proof to our setting.

Lemma 2.6. *For any smooth curve $\gamma : [0, T] \rightarrow \mathbb{D}$ and connections A and B ,*

$$[P_\gamma^A]^{-1} P_\gamma^B - \text{Id} = I_\gamma^{E(A,B)}(A - B) \tag{10}$$

where $E(A, B) \in \text{End}(\mathbb{C}^{n \times n})$ is given by $E(A, B)Q = AQ - QB$ for $Q \in \mathbb{C}^{n \times n}$.

The right-hand side of (10) is the attenuated x-ray of $A - B$ with respect to $E(A, B)$. This is slightly different to how we introduced the attenuated x-ray earlier. However, we can see $A - B$ as a one-form taking values in $\mathbb{C}^{n^2} \cong \mathbb{C}^{n \times n}$ and $E(A, B)$ as a connection on the trivial bundle $\mathbb{D} \times \mathbb{C}^{n^2}$. Before proving lemma 2.6, we state another useful lemma.

Lemma 2.7. *Let $\gamma : [0, T] \rightarrow \mathbb{D}$ be a smooth curve and let A and B be connections on \mathbb{D} . For any $Q \in \mathbb{C}^{n \times n}$,*

$$P_\gamma^{E(A,B)} Q = P_\gamma^A Q [P_\gamma^B]^{-1}. \tag{11}$$

Proof. Let u and v solve

$$\begin{cases} \dot{u} + E(A, B)(\dot{\gamma}(t))u = 0, \\ u(0) = Q, \end{cases} \quad \text{and} \quad \begin{cases} \dot{v} + B(\dot{\gamma}(t))v = 0, \\ v(0) = \text{Id}, \end{cases}$$

respectively. On one hand, by the definition of parallel transport,

$$u(T)v(T) = (P_{\gamma}^{E(A,B)}Q)P_{\gamma}^B.$$

On the other hand,

$$(\dot{u}v) = [-A(\dot{\gamma}(t))u + uB(\dot{\gamma}(t))]v - uB(\dot{\gamma}(t))v = -A(\dot{\gamma}(t))uv.$$

Hence, uv satisfies the parallel transport equation for A along γ with $u(0)v(0) = Q$, and so $u(T)v(T) = P_{\gamma}^A Q$. Combining the two expressions for $u(T)v(T)$ yields (11). \square

Proof of lemma 2.6. Let u_A solve

$$\begin{cases} \dot{u}_A + A(\dot{\gamma}(t))u_A = 0, \\ u_A(0) = \text{Id}, \end{cases}$$

and let u_B be the solution of the same equation with the connection A replaced by B . Consider the function $q := u_A u_B^{-1} - \text{Id}$. From the definition of parallel transport, evaluating q at T yields

$$q(T) = P_{\gamma}^A [P_{\gamma}^B]^{-1} - \text{Id}.$$

Moreover, one can check that q solves

$$\begin{cases} \dot{q} + E(A, B)(\dot{\gamma}(t))q = -(A - B)(\dot{\gamma}(t)), \\ q(0) = 0. \end{cases}$$

Lemma 2.3 then yields $q(T) = -P_{\gamma}^{E(A,B)} I_{\gamma}^{E(A,B)} (A - B)$. By combining the expressions for $q(T)$ and applying lemma 2.7, we get

$$P_{\gamma}^A [P_{\gamma}^B]^{-1} - \text{Id} = -P_{\gamma}^A I_{\gamma}^{E(A,B)} (A - B) [P_{\gamma}^B]^{-1}.$$

Rearranging the last equation yields (10). \square

Importantly, the pseudolinearisation identity is also valid in the broken case, where the parallel transports are replaced by the scattering data.

Lemma 2.8. Let A and B be connections on \mathbb{D} and let $x, y, z \in \mathbb{D}$. Then

$$[S_{z \leftarrow y \leftarrow x}^A]^{-1} S_{z \leftarrow y \leftarrow x}^B - \text{Id} = I_{z \leftarrow y \leftarrow x}^{E(A,B)} (A - B).$$

Proof. By expanding $I_{z \leftarrow y \leftarrow x}^{E(A,B)} (A - B)$, we get

$$I_{z \leftarrow y \leftarrow x}^{E(A,B)} (A - B) = I_{y \leftarrow x}^{E(A,B)} (A - B) + P_{x \leftarrow y}^{E(A,B)} I_{z \leftarrow y}^{E(A,B)} (A - B).$$

We can use lemma 2.6 on both attenuated x-ray transforms and lemma 2.7 on the parallel transport to get

$$\begin{aligned}
 I_{z \leftarrow y \leftarrow x}^{E(A,B)}(A - B) &= P_{x \leftarrow y}^A P_{y \leftarrow x}^B - \text{Id} + P_{x \leftarrow y}^A (P_{y \leftarrow z}^A P_{z \leftarrow y}^B - \text{Id}) P_{y \leftarrow x}^B \\
 &= P_{x \leftarrow y}^A P_{y \leftarrow x}^B - \text{Id} + P_{x \leftarrow y}^A P_{y \leftarrow z}^A P_{z \leftarrow y}^B P_{y \leftarrow x}^B - P_{x \leftarrow y}^A P_{y \leftarrow x}^B \\
 &= [S_{z \leftarrow y \leftarrow x}^A]^{-1} S_{z \leftarrow y \leftarrow x}^B - \text{Id}
 \end{aligned}$$

as claimed. □

The pseudolinearisation identity and lemma 2.5 are enough to prove theorem 2.1.

Proof of theorem 2.1. By interchanging the role of x and z , we see that the pseudolinearisation identity can also be written as

$$S_{z \leftarrow y \leftarrow x}^A [S_{z \leftarrow y \leftarrow x}^B]^{-1} - \text{Id} = I_{x \leftarrow y \leftarrow z}^{E(A,B)}(A - B).$$

Hence, by definition of the broken x-ray, we have

$$\begin{aligned}
 \partial_{x \leftarrow y} (S_{z \leftarrow y \leftarrow x}^A [S_{z \leftarrow y \leftarrow x}^B]^{-1}) &= \partial_{x \leftarrow y} (I_{y \leftarrow z}^{E(A,B)}(A - B) \\
 &\quad + P_{z \leftarrow y}^{E(A,B)} I_{x \leftarrow y}^{E(A,B)}(A - B)).
 \end{aligned}$$

The operator $\partial_{x \leftarrow y}$ is essentially a derivative with respect to x , and so the first term in the definition of the broken x-ray vanishes. Moreover, $P_{z \leftarrow y}^{E(A,B)}$ is unaffected. It follows from lemmas 2.5 and 2.7 that

$$\begin{aligned}
 \partial_{x \leftarrow y} (S_{z \leftarrow y \leftarrow x}^A [S_{z \leftarrow y \leftarrow x}^B]^{-1}) &= P_{z \leftarrow y}^{E(A,B)} \partial_{x \leftarrow y} (I_{x \leftarrow y}^{E(A,B)}(A - B)) \\
 &= P_{z \leftarrow y}^{E(A,B)} P_{y \leftarrow x}^{E(A,B)} (A - B)_x(v_{x \leftarrow y}) \\
 &= S_{z \leftarrow y \leftarrow x}^A (A - B)_x(v_{x \leftarrow y}) [S_{z \leftarrow y \leftarrow x}^B]^{-1}.
 \end{aligned}$$

Taking norms, the scattering data terms vanish since they belong in $U(n)$ and we get

$$|(A - B)_x(v_{x \leftarrow y})| = |\partial_{x \leftarrow y} (S_{z \leftarrow y \leftarrow x}^A [S_{z \leftarrow y \leftarrow x}^B]^{-1})|. \tag{12}$$

The choice of z on the right-hand side is irrelevant, and we take $z = z_y$. Since vectors of the form $v_{x \leftarrow y}$ form a basis of the tangent plane at x without degenerating when ε goes to 0, we can find a constant $C > 0$ such that

$$\begin{aligned}
 \sup_{\substack{v \in T_x \mathbb{D} \\ |v|=1}} |(A - B)_x(v)| &\leq C \sup_{\substack{y \in \mathbb{D} \setminus \cup \\ (x,y) \in \mathbb{L}}} |(A - B)_x(v_{x \leftarrow y})| \\
 &= C \sup_{\substack{y \in \mathbb{D} \setminus \cup \\ (x,y) \in \mathbb{L}}} |\partial_{x \leftarrow y} (S_{z \leftarrow y \leftarrow x}^A [S_{z \leftarrow y \leftarrow x}^B]^{-1})|
 \end{aligned}$$

and the theorem follows. Note that C can be chosen independently of $x \in \cup^X$ by symmetry. □

To prove theorem 1.3, it only remains to integrate over \cup^X to get a global estimate.

Proof of theorem 1.3. Note that the map

$$\omega_x \mapsto \int_{y: x \in \mathcal{F}_y^X} |\omega_x(v_{x \leftarrow y})| dy$$

is a norm on $T_x^*\mathbb{R}^{1+3}$ since the vectors $v_{x\leftarrow y}$ span $T_x\mathbb{R}^{1+3}$. By equivalence of norms, we can find $C > 0$ independent of ε , $x \in \mathcal{U}^X$ (by symmetry) and the connections A and B such that

$$\|(A - B)_x\| \leq C \int_{y : x \in \mathcal{F}_y^X} |(A - B)_x(v_{x\leftarrow y})| dy.$$

After changing the integrand through equation (12) with $z = z_y$, integrating over $x \in \mathcal{U}^X$ and using Cauchy–Schwarz yields the desired estimate. \square

The proof of theorem 2.1 crucially relies on the fact that x is always an endpoint of the path and is not the breaking point, since then the operator $\partial_{x\leftarrow y}$ only hits $I_{x\leftarrow y}^{E(A,B)}$ in the expression for the broken attenuated x-ray. This allows us to evaluate $A - B$ inside \mathcal{U} , but such an approach does not immediately work for evaluating $A - B$ outside \mathcal{U} . This is where we need to take the gauge into account.

2.4. Dealing with the gauge through a potential form

In order to use similar techniques as in the proof of theorem 2.1 to estimate the connection outside \mathcal{U} , we aim to make the second term in (9) vanish. To do so, we will modify the argument of the attenuated x-ray by a potential form.

For a connection A and a one-form ω , we define the function

$$p(y) := p_\omega^A(y) = P_{y\leftarrow z_y}^A I_{y\leftarrow z_y}^A(\omega). \tag{13}$$

This function will serve as an approximate potential for ω . We chose p in this way so that $I_{z_y\leftarrow y}^A(\omega - d_A p)$ vanishes for all $y \in \mathbb{D} \setminus \mathcal{U}$, as the next lemma shows.

Lemma 2.9. *Let $\gamma : [0, T] \rightarrow \mathbb{D}$ be the unit-speed lightlike geodesic from z_y to y . Then, with p defined as above, we have*

$$\omega(\dot{\gamma}(t)) = d_A p(\dot{\gamma}(t))$$

for all $t \in (0, T)$. In particular, $\omega(v_{y\leftarrow z_y}) = d_A p(v_{y\leftarrow z_y})$.

Proof. Consider the unique solution u of

$$\begin{cases} \dot{u} + A(\dot{\gamma}(t))u = -\omega(\dot{\gamma}(t)), \\ u(0) = 0, \end{cases}$$

along γ . By lemma 2.3, it holds that $u(t) = -P_{\gamma(t)\leftarrow z_y}^A I_{\gamma(t)\leftarrow z_y}^A(\omega) = -p(\gamma(t))$. Hence, we have

$$-d_A p(\dot{\gamma}(t)) = \dot{u}(t) + A(\dot{\gamma}(t))u(t) = -\omega(\dot{\gamma}(t))$$

and the result follows. \square

We can deduce from lemma 2.9 and (7) that $I_{z_y\leftarrow y}^A(\omega - d_A p) = 0$ for all y and so, on the one hand,

$$I_{z_y\leftarrow y\leftarrow x}^A(\omega - d_A p) = I_{y\leftarrow x}^A(\omega - d_A p).$$

On the other hand, by (8), we have

$$\begin{aligned} I_{z_y \leftarrow y \leftarrow x}^A(\omega - d_A p) &= I_{z_y \leftarrow y \leftarrow x}^A(\omega) - P_{x \leftarrow y}^A P_{y \leftarrow z_y}^A p(z_y) + p(x) \\ &= I_{z_y \leftarrow y \leftarrow x}^A(\omega) + p(x) \end{aligned}$$

since $p(z_y) = 0$ and so by combining both expressions, we get

$$I_{z_y \leftarrow y \leftarrow x}^A(\omega) = I_{y \leftarrow x}^A(\omega - d_A p) - p(x). \quad (14)$$

By applying $\partial_{y \leftarrow x}$ to both sides of the last expression, lemma 2.5 yields

$$\partial_{y \leftarrow x} \left(I_{z_y \leftarrow y \leftarrow x}^A(\omega) \right) = P_{x \leftarrow y}^A \left([\omega - d_A p]_y(v_{y \leftarrow x}) \right) \quad (15)$$

since $\partial_{y \leftarrow x}$ is essentially a derivative in y , and so $\partial_{y \leftarrow x} p(x) = 0$. To prove theorem 2.2, we will replace A by $E(A, B)$ and ω by $A - B$ in (15) in order to use the pseudolinarisation identity.

2.5. Evaluating from the tangent space at y

As shown in [CLOP21a, lemma 1], the set of vectors $v_{y \leftarrow x}$ for $x \in (\mathcal{F}_\varepsilon^X)_y$ form a basis of the tangent space at y , but this basis degenerates when ε goes to 0. We therefore need estimates to quantify how well we can estimate $\omega - d_A p$ at $y \in \mathbb{D} \setminus \mathcal{U}$ from moving x around in the intersection of \mathcal{U} and the past light cone of y .

Lemma 2.10. *Let $0 < \varepsilon < \varepsilon_0$ and let $y \in \mathbb{D} \setminus \mathcal{U}_\varepsilon$. Then*

$$\|(\omega - d_A p)_y\| \leq \frac{8}{\varepsilon} \sup_{x \in (\mathcal{F}_\varepsilon^X)_y} |(\omega - d_A p)(v_{y \leftarrow x})|.$$

The key to proving lemma 2.10 is this small linear algebra lemma whose proof is straightforward.

Lemma 2.11. *Let b_1, \dots, b_m be a basis of \mathbb{R}^m with $|b_i|_{\mathbb{R}^m} = 1$ and let $\mathcal{T} : \mathbb{R}^m \rightarrow \mathbb{C}^n$ be a linear map. Then*

$$\|\mathcal{T}\| \leq \sqrt{m} \|B^{-1}\| \max_{i=1, \dots, m} |\mathcal{T}(b_i)|$$

where B is the matrix whose columns are the b_i 's and $\|B^{-1}\|$ is the operator norm of its inverse.

Proof of lemma 2.10. As stated earlier, lemma 1 in [CLOP21a] guarantees that the set of vectors $v_{y \leftarrow x}$ generate $T_y \mathbb{R}^{1+3}$. Hence, we wish to apply lemma 2.11 by evaluating from $T_y \mathbb{R}^{1+3}$ using different light rays γ_x from x to y for different $x \in \mathcal{U}_\varepsilon$ with $(x, y) \in \mathbb{L}$, that is, x in the fibre of $\mathcal{F}_\varepsilon^X$ at y .

We first claim that it suffices to compute the case where $y = (0, 1, 0, 0)$. Through a rotation in space and a translation in time, we can identify the sets $\{v_{y \leftarrow x}\}_{x \in \mathcal{U}_\varepsilon}$ and $\{v_{y' \leftarrow x}\}_{x \in \mathcal{U}_\varepsilon}$ whenever y and y' share the same spatial norm. By symmetry, this does not intervene in norm estimates. Therefore, without loss of generality, we can choose $y = y_r = (0, r, 0, 0)$. Moreover, whenever $r_1 < r_2$, we can see that $\{v_{y_{r_2} \leftarrow x}\}_{x \in \mathcal{U}_\varepsilon} \subset \{v_{y_{r_1} \leftarrow x}\}_{x \in \mathcal{U}_\varepsilon}$ and so any stability estimate for y_{r_2} is also valid for y_{r_1} since we are taking the supremum over a larger set. Hence, it suffices to show the case $r = 1$, as claimed.

To apply lemma 2.11, we need a basis of $T_y \mathbb{R}^{1+3}$. Let $b_i := v_i / |v_i|$ where

$$\begin{aligned}v_1 &= (1, 1, 0, 0), \\v_2 &= (\sqrt{1 + \varepsilon^2}, 1, \varepsilon, 0), \\v_3 &= (\sqrt{1 + \varepsilon^2}, 1, 0, \varepsilon), \\v_4 &= (-1, 1, 0, 0).\end{aligned}$$

It is obvious that the b_i 's are linearly independent and hence form a basis of the tangent space at y . The vector b_4 is $v_{y \leftarrow z_y}$, while the other vectors b_i correspond to $v_{y \leftarrow x_i}$ with $x_1 = (-1, 0, 0, 0)$, $x_2 = (-\sqrt{1 + \varepsilon^2}, 0, -\varepsilon, 0)$ and $x_3 = (-\sqrt{1 + \varepsilon^2}, 0, 0, -\varepsilon)$. Notice that $(x_i, y) \in \mathbb{L}$ and that $x_i \in \overline{U_\varepsilon}$ for $i = 1, 2, 3$. A somewhat tedious computation verified with Mathematica yields

$$\|B^{-1}\|_F = \frac{\sqrt{8\varepsilon^2 + 8}}{\varepsilon} \leq \frac{4}{\varepsilon}$$

for $0 < \varepsilon < 1$. The operator norm and the Frobenius norm are equivalent with $\|B^{-1}\| \leq \|B^{-1}\|_F$ and so by lemma 2.11,

$$\begin{aligned}\|(\omega - d_A p)_y\| &= \sup_{v \in T_y \mathbb{R}^{1+3}} \frac{|(\omega - d_A p)(v)|}{|v|} \\&\leq 2 \|B^{-1}\| \max_{i=1,2,3,4} |(\omega - d_A p)(b_i)| \\&\leq \frac{8}{\varepsilon} \sup_{\substack{x \in \overline{U_\varepsilon} \\ (x,y) \in \mathbb{L}}} |(\omega - d_A p)(v_{y \leftarrow x})|.\end{aligned}$$

The last inequality follows from the fact that $(\omega - d_A p)(b_4) = 0$ by lemma 2.9 and since $\{b_1, b_2, b_3\}$ is in the closure of $\{v_{y \leftarrow x}\}_{x \in \overline{U_\varepsilon}}$. \square

2.6. Proof of theorems 1.4 and 2.2

We finally have everything we need to prove theorem 2.2. The main idea is to use the pseudo-linearisation identity to relate the scattering data with an attenuated x-ray transform of $A - B$, and then use the operator $\partial_{y \leftarrow x}$ to evaluate $A - B - d_{E(A,B)} p$ from $T_y \mathbb{R}^{1+3}$. Theorem 1.4 then immediately follows by integrating over $\mathbb{D} \setminus \mathcal{U}$.

Proof of theorem 2.2. By lemma 2.8, we have

$$[S_{z_y \leftarrow y \leftarrow x}^A]^{-1} S_{z_y \leftarrow y \leftarrow x}^B - \text{Id} = I_{z_y \leftarrow y \leftarrow x}^{E(A,B)}(A - B).$$

We can see A and B as one-forms taking values in \mathbb{C}^{n^2} and $E(A, B)$ a Hermitian connection taking values in $\mathfrak{u}(n^2)$. Hence, if we let

$$p = p_{A-B}^{E(A,B)} = P_{y \leftarrow z_y}^{E(A,B)} I_{y \leftarrow z_y}^{E(A,B)}(A - B)$$

then (14) yields

$$I_{z_y \leftarrow y \leftarrow x}^{E(A,B)}(A - B) = I_{y \leftarrow x}^{E(A,B)}(A - B - d_{E(A,B)} p) - p(x).$$

Since $\partial_{y \leftarrow x} \text{Id} = \partial_{y \leftarrow x} p(x) = 0$, it now follows from lemma 2.5 that

$$\begin{aligned}\partial_{y \leftarrow x} \left([S_{z_y \leftarrow y \leftarrow x}^A]^{-1} S_{z_y \leftarrow y \leftarrow x}^B \right) &= P_{y \leftarrow x}^{E(A,B)} (A - B - d_{E(A,B)} p)(v_{y \leftarrow x}) \\ &= P_{y \leftarrow x}^A (A - B - d_{E(A,B)} p)(v_{y \leftarrow x}) P_{x \leftarrow y}^B.\end{aligned}$$

The parallel transports are in $U(n)$ and so

$$|\partial_{y \leftarrow x} \left([S_{z_y \leftarrow y \leftarrow x}^A]^{-1} S_{z_y \leftarrow y \leftarrow x}^B \right)| = |(A - B - d_{E(A,B)} p)_y(v_{y \leftarrow x})|. \quad (16)$$

We can finally apply lemma 2.10 to get

$$\|(A - B - d_{E(A,B)} p)_y\| \leq \frac{8}{\varepsilon} \sup_{x \in (\mathcal{F}_\varepsilon^X)_y} |(A - B - d_{E(A,B)} p)_y(v_{y \leftarrow x})|.$$

Finally, note that

$$\sup_{x \in (\mathcal{F}_\varepsilon^X)_y} |\omega_y(v_{y \leftarrow x})| \sim \frac{1}{\text{vol}((\mathcal{F}_\varepsilon^X)_y)} \int_{(\mathcal{F}_\varepsilon^X)_y} |\omega_y(v_{y \leftarrow x})| \, dx$$

for any one-form ω as ε goes to 0, where \sim means that their quotient tends to one. Moreover, both sides of the last equation can be considered as norms on $T_y^* \mathbb{R}^{1+3}$. Therefore, by equivalence of norms and the fact that $\text{vol}((\mathcal{F}_\varepsilon^X)_y)$ is proportional to ε^3 , we can find a constant $C > 0$ independent of ε as well as A and B such that

$$\|(A - B - d_{E(A,B)} p)_y\| \leq \frac{C}{\varepsilon^4} \int_{(\mathcal{F}_\varepsilon^X)_y} |\partial_{y \leftarrow x} \left([S_{z_y \leftarrow y \leftarrow x}^A]^{-1} S_{z_y \leftarrow y \leftarrow x}^B \right)| \, dx.$$

The constant C degenerates the further y is from \mathcal{O} and so it can be chosen uniformly by compactness of \mathbb{D} . \square

Remark 2.12. Note that even though there is a supremum in the right-hand side of (6), one does not need to know $\partial_{y \leftarrow x} \left([S_{z_y \leftarrow y \leftarrow x}^A]^{-1} S_{z_y \leftarrow y \leftarrow x}^B \right)$ for all $x \in \mathcal{U}_\varepsilon$ in the past light cone of y to get an estimate. Indeed, the important equation is (16) as it reveals the linear structure behind the estimate. In practice, one only needs to evaluate $A - B - d_{E(A,B)} p$ at three different linearly independent vectors $v_{y \leftarrow x}$ since we already know it vanishes when evaluated at $v_{y \leftarrow z_y}$. Lemma 2.11 then yields an estimate for those vectors.

2.7 H^1 estimate

It remains to show corollary 1.7, which relates S^A and S^B in a linear fashion rather than through the group multiplication in $U(n)$. To do so, we follow the argument in [MNP21, corollary 2.3].

Lemma 2.13. *There is a constant $C > 0$ such that*

$$\|\partial([S^A]^{-1} S^B)\|_{L^2(\mathcal{F}^X)} \leq C \left(1 + \|A \triangleleft P_{y \leftarrow z_y}^A\|_{L^\infty(\mathbb{D} \setminus \mathcal{U})} \right) \|S^A - S^B\|_{H^1(\mathcal{F}^X)}.$$

Proof. To simplify notation, we omit the paths in what follows and write S^A for $S_{z_y \leftarrow y \leftarrow x}^A$. We can expand

$$\begin{aligned}
 |\partial_{y \leftarrow x}([S^A]^{-1}S^B)| &= |(\partial_{y \leftarrow x}[S^A]^{-1})S^B + [S^A]^{-1}\partial_{y \leftarrow x}S^B| \\
 &= |[S^A]^{-1}\partial_{y \leftarrow x}S^B - [S^A]^{-1}(\partial_{y \leftarrow x}S^A)[S^A]^{-1}S^B| \\
 &= |\partial_{y \leftarrow x}(S^B - S^A) + (\partial_{y \leftarrow x}S^A)(\text{Id} - [S^A]^{-1}S^B)| \\
 &\leq |\partial_{y \leftarrow x}(S^B - S^A)| + |\partial_{y \leftarrow x}S^A||S^A - S^B|.
 \end{aligned}$$

The third equality follows from the second by using that $S^A \in U(n)$ as well as adding and subtracting $\partial_{y \leftarrow x}S^A$. Taking the supremum over the fibres \mathcal{F}_y^X , squaring, integrating and using that $(a + b)^2 \leq 2(a^2 + b^2)$ yields

$$\begin{aligned}
 \|\partial([S^A]^{-1}S^B)\|_{L^2(\mathcal{F}^X)}^2 &\leq 2\left(\|\partial(S^A - S^B)\|_{L^2(\mathcal{F}^X)}^2 \right. \\
 &\quad \left. + \|\partial S^A\|_{L^\infty(\mathcal{F}^X)}^2 \|S^A - S^B\|_{L^2(\mathcal{F}^X)}^2\right).
 \end{aligned}$$

It remains to estimate $\|\partial S^A\|_{L^\infty(\mathcal{F}^X)}$. We did not show it yet, but the proof of theorem 1.6 reveals that

$$|\partial_{y \leftarrow x}S_{z_y \leftarrow y \leftarrow x}^A| = |(A \triangleleft P_{y \leftarrow z_y}^A)(v_{y \leftarrow x})|$$

and so $\|\partial S^A\|_{L^\infty(\mathcal{F}^X)} \leq \|A \triangleleft P_{y \leftarrow z_y}^A\|_{L^\infty(\mathbb{D} \setminus \zeta)}$. The estimate follows by taking square roots and using that $\sqrt{1 + x^2}/(1 + x)$ is bounded. \square

The last estimate is again invariant under \mathcal{G} and involves the L^∞ -norm of the light-sink connection $A \triangleleft P_{y \leftarrow z_y}^A$. We can get an estimate on that norm involving the curvature of A and the value of A along \mathcal{O} .

Lemma 2.14. *There is a constant C such that*

$$\|A \triangleleft P_{y \leftarrow z_y}^A\|_{L^\infty(\mathbb{D})} \leq C(\|F_A\|_{L^\infty(\mathbb{D})} + \|A(\partial_t)\|_{L^\infty(\mathcal{O})})$$

where $F_A = dA + A \wedge A$ is the curvature two-form of A and

$$\|F_A\|_{L^\infty(\mathcal{D})} = \sup_{\substack{y \in \mathbb{D} \\ |u|=|v|=1}} |(F_A)_y(u, v)|.$$

Corollary 1.7 will then directly follow from theorems 1.4, 1.6, lemmas 2.13 and 2.14. However, we need another lemma before proving lemma 2.14.

Lemma 2.15. *Let $\gamma : [0, T] \rightarrow \mathbb{D}$ be a smooth curve and let $\gamma_s : [0, T] \rightarrow \mathbb{D}$ be a smooth variation of γ where $s \in I = [-\delta, \delta]$ for some $\delta > 0$. Then*

$$\begin{aligned}
 \left. \frac{d}{ds} P_{\gamma_s}^A \right|_{s=0} &= P_{\gamma}^A A_{\gamma(0)}(\partial_s \gamma_s(0)) - A_{\gamma(T)}(\partial_s \gamma_s(T)) P_{\gamma}^A \\
 &\quad + \int_0^T P_{\gamma[t, T]}^A F_A(\dot{\gamma}(t), \partial_s \gamma_s(t)) P_{\gamma[0, t]}^A dt
 \end{aligned} \tag{17}$$

where $\partial_s \gamma_s(t) = \left. \frac{d}{ds} \gamma_s(t) \right|_{s=0}$ and $P_{\gamma[0, t]}^A$ is the parallel transport along the segment of γ restricted to the interval $[0, t]$.

Proof. Let $U(s, t) = P_{\gamma_s[0,t]}^A$. Then U solves

$$\begin{cases} \partial_t U(s, t) + A(\dot{\gamma}_s(t))U(s, t) = 0, & (s, t) \in I \times [0, T]; \\ U(s, 0) = \text{Id}. \end{cases}$$

By differentiating with respect to s , we get

$$\begin{cases} \partial_s \partial_t U(s, t) + \partial_s [A(\dot{\gamma}_s(t))]U(s, t) + A(\dot{\gamma}_s(t))\partial_s U(s, t) = 0, & (s, t) \in I \times [0, T]; \\ \partial_s U(s, 0) = 0. \end{cases}$$

Let $v(s, t) = \partial_s U(s, t)$. We are interested in computing $v(0, T)$. From the previous equation, we see that v satisfies the inhomogeneous differential equation

$$\begin{cases} \partial_t v + A(\dot{\gamma}_s(t))v = -\partial_s [A(\dot{\gamma}_s(t))]P_{\gamma_s[0,t]}^A, & (s, t) \in I \times [0, T]; \\ v(s, 0) = 0. \end{cases}$$

By Duhamel's principle, the solution of this differential equation is given by

$$v(s, t) = \int_0^t u^r(s, t) \, dr$$

where u^r solves

$$\begin{cases} \partial_t u^r + A(\dot{\gamma}_s(t))u^r = 0, & (s, t) \in I \times (r, T); \\ u^r(s, r) = -\partial_s [A(\dot{\gamma}_s(r))]P_{\gamma_s[0,r]}^A. \end{cases}$$

The equation defining u^r is simply that of a parallel transport and so

$$u^r(s, t) = P_{\gamma_s[r,t]}^A (-\partial_s [A(\dot{\gamma}_s(r))])P_{\gamma_s[0,r]}^A.$$

Hence, we get

$$v(0, T) = \int_0^T P_{\gamma[0,T]}^A (-\partial_s [A(\dot{\gamma}_s(r))]|_{s=0})P_{\gamma[0,r]}^A \, dr.$$

Expanding the curvature two-form $F_A = dA + A \wedge A$ yields

$$\begin{aligned} F_A(\dot{\gamma}(r), \partial_s \gamma_s(r)) &= \partial_r [A(\partial_s \gamma_s(r))] - \partial_s [A(\dot{\gamma}_s(r))] - A([\dot{\gamma}(r), \partial_s \gamma_s(r)]) \\ &\quad + A(\dot{\gamma}(r))A(\partial_s \gamma_s(r)) - A(\partial_s \gamma_s(r))A(\dot{\gamma}(r)). \end{aligned}$$

The vectors $\dot{\gamma}(r)$ and $\partial_s \gamma_s(r)$ commute so the term with the commutator vanishes. We can isolate $-\partial_s [A(\dot{\gamma}_s(r))]$ in that expression to get

$$\begin{aligned} v(0, T) &= \int_0^T P_{\gamma[0,T]}^A [F_A(\dot{\gamma}(r), \partial_s \gamma_s(r)) - \partial_r [A(\partial_s \gamma_s(r))] \\ &\quad - A(\dot{\gamma}(r))A(\partial_s \gamma_s(r)) + A(\partial_s \gamma_s(r))A(\dot{\gamma}(r))]P_{\gamma[0,r]}^A \, dr. \end{aligned} \quad (18)$$

We can integrate by parts the last term using that $A(\dot{\gamma}(r))P_{\gamma[0,r]}^A = -\partial_r P_{\gamma[0,r]}^A$. This yields

$$\begin{aligned} & \int_0^T P_{\gamma[r,T]}^A A(\partial_s \gamma_s(r)) A(\dot{\gamma}(r)) P_{\gamma[0,r]}^A dr \\ &= [-P_{\gamma[r,T]}^A A(\partial_s \gamma_s(r)) P_{\gamma[0,r]}^A]_0^T \\ & \quad + \int_0^T \partial_r [P_{\gamma[r,T]}^A A(\partial_s \gamma_s(r))] P_{\gamma[0,r]}^A dr. \end{aligned}$$

The boundary term corresponds to the first two terms in (17) and we can expand the integrand of the second term to get

$$\begin{aligned} \partial_r [P_{\gamma[r,T]}^A A(\partial_s \gamma_s(r))] P_{\gamma[0,r]}^A &= P_{\gamma[r,T]}^A [A(\dot{\gamma}(r)) A(\partial_s \gamma_s(r)) \\ & \quad + \partial_r [A(\partial_s \gamma_s(r))]] P_{\gamma[0,r]}^A \end{aligned}$$

since $\partial_r P_{\gamma[r,T]}^A = P_{\gamma[r,T]}^A A(\dot{\gamma}(r))$. These terms cancel with the second and third terms in (18) to simplify $v(0, T)$ to (17). \square

Proof of lemma 2.14. For $y \in \mathbb{D}$ and unit $v \in T_y \mathbb{D}$, we have

$$(A \triangleleft P_{y \leftarrow z_y}^A)_y(v) = P_{z_y \leftarrow y}^A dP_{y \leftarrow z_y}^A(v) + P_{z_y \leftarrow y}^A A_y(v) P_{y \leftarrow z_y}^A.$$

We can compute $dP_{y \leftarrow z_y}^A(v)$ by using lemma 2.15 with the variation γ_s given by the lightlike geodesic from z_{y+sv} to $y + sv$. This yields

$$\begin{aligned} dP_{y \leftarrow z_y}^A(v) &= P_{y \leftarrow z_y}^A A_{z_y}(c \partial_t) - A_y(v) P_{y \leftarrow z_y}^A \\ & \quad + \int_0^T P_{y \leftarrow \gamma(t)}^A F_A(\dot{\gamma}(t), \partial_s \gamma_s(t)) P_{\gamma(t) \leftarrow z_y}^A dt \end{aligned}$$

for some $-1 \leq c \leq 1$ and hence

$$(A \triangleleft P_{y \leftarrow z_y}^A)(v) = c A_{z_y}(\partial_t) + \int_0^T P_{y \leftarrow \gamma(t)}^A F_A(\dot{\gamma}(t), \partial_s \gamma_s(t)) P_{\gamma(t) \leftarrow z_y}^A dt.$$

The result follows since $z_y \in \mathcal{O}$ for all y and

$$\left| \int_0^T P_{y \leftarrow \gamma(t)}^A F_A(\dot{\gamma}(t), \partial_s \gamma_s(t)) P_{\gamma(t) \leftarrow z_y}^A dt \right| \leq T \|F_A\|_{L^\infty(\mathbb{D})}.$$

\square

2.8. Forward estimates

We finish the section by collecting forward estimates that will be useful for the statistical applications in section 4.

Lemma 2.16. *Let A and B be Hermitian connections. Then*

$$\|S_{z \leftarrow y \leftarrow x}^A - S_{z \leftarrow y \leftarrow x}^B\|_{L^\infty(\mathbb{S}^+(\mathbb{U}))} \leq C \|A - B\|_{L^\infty(\mathbb{D})}.$$

Proof. Since $S_{z \leftarrow y \leftarrow x}^A$ and the parallel transports lie in $U(n)$, lemma 2.8 yields the pointwise estimate

$$\begin{aligned}
 |S_{z \leftarrow y \leftarrow x}^A - S_{z \leftarrow y \leftarrow x}^B| &= |I_{z \leftarrow y \leftarrow x}^{E(A,B)}(A - B)| \\
 &\leq |I_{y \leftarrow x}^{E(A,B)}(A - B)| + |I_{z \leftarrow y}^{E(A,B)}(A - B)| \\
 &\leq \int_0^{T_1} |(A - B)(\dot{\gamma}_{y \leftarrow x}(t))| dt \\
 &\quad + \int_0^{T_2} |(A - B)(\dot{\gamma}_{z \leftarrow y}(t))| dt \\
 &\leq |x - y|_e \sup_{t \in [0, T_1]} |(A - B)(\dot{\gamma}_{y \leftarrow x}(t))| \\
 &\quad + |y - z|_e \sup_{t \in [0, T_2]} |(A - B)(\dot{\gamma}_{z \leftarrow y}(t))|
 \end{aligned}$$

where $\gamma_{y \leftarrow x} : [0, T_1] \rightarrow \mathbb{D}$ and $\gamma_{z \leftarrow y} : [0, T_2] \rightarrow \mathbb{D}$ are the unit-speed geodesics from x to y and y to z , respectively, and $|x - y|_e$ is the Euclidean distance between x and y . In particular, by taking the supremum over all values of $(x, y, z) \in \mathbb{S}^+(\mathcal{U})$ and using that the distance between x and y is bounded, we get

$$\begin{aligned}
 \|S_{z \leftarrow y \leftarrow x}^A - S_{z \leftarrow y \leftarrow x}^B\|_{L^\infty(\mathbb{S}^+(\mathcal{U}))} &\leq C \sup_{\substack{x, z \in \mathcal{U}, y \in \mathbb{D} \\ (x, y), (y, z) \in \mathbb{L}}} \{|(A - B)(v_{y \leftarrow x})|, |(A - B)(v_{y \leftarrow z})|\} \\
 &\leq C \|A - B\|_{L^\infty(\mathbb{D})}
 \end{aligned}$$

since the second supremum is taken over a larger set. Note that we did not restrict y outside \mathcal{U} in the first supremum since the curves $\gamma_{y \leftarrow x}$ and $\gamma_{y \leftarrow z}$ cross \mathcal{U} . \square

Lemma 2.17. *Let A be a Hermitian connection and ω a one-form on \mathbb{D} with values in \mathbb{C}^n . For all $k \geq 0$, there is a constant $C > 0$ such that*

$$\|I^A(\omega)_{y \leftarrow x}\|_{H^k(\mathcal{F}^X)} \leq C \|P_{y \leftarrow x}^A\|_{C^k(\mathbb{D} \times \mathbb{D})} \|\omega\|_{H^k(\mathbb{D})}.$$

Proof. We will show that I can be extended to a continuous map from $H^k(\mathbb{D})$ to $H^k(\mathcal{F}^X)$ by following the approach of [Sha94, theorem 4.2.1]. The attenuated x-ray can be written as

$$I_{y \leftarrow x}^A(\omega) = \int_0^1 P_{x \leftarrow x + t(y-x)}^A \omega_{x+t(y-x)}(y-x) dt.$$

Firstly, for $k = 0$, we have

$$\|I^A(\omega)\|_{L^2(\mathcal{F}^X)}^2 \leq C \int_{\mathcal{F}^X} \int_0^1 |\omega_{x+t(y-x)}(y-x)|^2 dt dx dy.$$

Denoting the future light cone at z by $J^+(z)$, we claim that for every $z \in \mathbb{D}$, there is a nonnegative, bounded and compactly supported function $g_z : J^+(z) \rightarrow \mathbb{R}$ such that

$$\int_{\mathcal{F}^X} \int_0^1 |\omega_{x+t(y-x)}(y-x)|^2 dt dx dy = \int_{z \in \mathbb{D}} \int_{v \in J^+(z)} |\omega_z(v)|^2 g_z(v) dv dy.$$

Indeed, identifying z and v from the integral on the right with $x + t(y - x)$ and $y - x$ respectively in the integral on the left, the function g_z corresponds to the amount of times the term

$\omega_z(v)$ appears in the left integral. The only way to vary x and y while keeping $y - x$ unchanged is to perturb both x and y by the same amount, say w . But then, $x + w + t_1(y - x) = x + t_2(y - x)$ if and only if $w = (t_2 - t_1)(y - x)$ and so w must be parallel to $y - x$. The function $g_z(v)$ is hence given by the length of the maximal interval \mathcal{I} such that

$$(y, x) = (z + (1 - t)v, z - tv) \in \mathcal{F}^X$$

for all $t \in \mathcal{I}$. We can bound g_z and its support uniformly in z . It follows that there are constants $C, M > 0$ such that

$$\|I_{y \leftarrow x}^A(\omega)\|_{L^2(\mathcal{F}^X)}^2 \leq C \int_{z \in \mathbb{D}} \int_{v \in J^+(z)} |\omega_z(v)|^2 1_{\{|v|_e < M\}} \, dv \, dy \leq C \|\omega\|_{L^2(\mathbb{D})}^2$$

with the last inequality following from equivalence of norms and symmetry.

Now suppose that ω is smooth. Let $\{\partial^1, \dots, \partial^\gamma\}$ be a global commuting frame on \mathcal{F}^X . Such a frame exists since we can choose global coordinates on a past light cone when we exclude its apex and we can see \mathcal{F}^X as an open submanifold of the product between \mathbb{D} and a past light cone. Then for every multi-index α with $|\alpha| = k$,

$$\partial^\alpha I_{y \leftarrow x}^A(\omega) = \sum_{\beta + \gamma = \alpha} C_{\beta\gamma}^\alpha \int_0^1 \partial^\beta [P_{x \leftarrow x + t(y-x)}^A] \partial^\gamma [\omega_{x+t(y-x)}(y-x)] \, dt$$

and so

$$\begin{aligned} & \|\partial^\alpha I_{y \leftarrow x}^A(\omega)\|_{L^2(\mathcal{F}^X)}^2 \\ & \leq C \|P_{y \leftarrow x}^A\|_{C^k(\mathbb{D} \times \mathbb{D})}^2 \sum_{\gamma \leq \alpha} \int_{\mathcal{F}^X} \int_0^1 |\partial^\gamma [\omega_{x+t(y-x)}(y-x)]|^2 \, dt \, dx \, dy \\ & \leq C \|P_{y \leftarrow x}^A\|_{C^k(\mathbb{D} \times \mathbb{D})}^2 \sum_{\gamma \leq \alpha} \|\partial^\gamma \omega\|_{L^2(\mathbb{D})}^2 \end{aligned}$$

since the map $(y, x) \mapsto x + t(y - x)$ is smooth with bounded derivatives and we used the same argument as above to get the bound with $\|\partial^\gamma \omega\|_{L^2(\mathbb{D})}$. It follows that

$$\|I_{y \leftarrow x}^A(\omega)\|_{H^k(\mathcal{F}^X)} \leq C \|P_{y \leftarrow x}^A\|_{C^k(\mathbb{D} \times \mathbb{D})} \|\omega\|_{H^k(\mathbb{D})}. \tag{19}$$

Now if $\omega \in H^k(\mathbb{D})$, we can choose a smooth sequence ω_j converging to ω in $H^k(\mathbb{D})$. On one hand, the sequence $I^A(\omega_j)$ converges to $I(\omega)$ in $L^2(\mathcal{F}^X)$. On the other, by the estimate (19), the sequence $I^A(\omega_j)$ is a Cauchy sequence in $H^k(\mathcal{F}^X)$, and so $I^A(\omega) \in H^k(\mathcal{F}^X)$. Moreover, the estimate holds for all $\omega \in H^k(\mathbb{D})$ and not only for smooth ω . \square

Lemma 2.18. *Let A and B be Hermitian connections. There is a constant $C > 0$ such that*

$$\|S_{z_y \leftarrow y \leftarrow x}^A - S_{z_y \leftarrow y \leftarrow x}^B\|_{L^2(\mathcal{F}^X)} \leq C \|A - B\|_{L^2(\mathbb{D})}.$$

Proof. By the pseudolinearisation identity and the definition of the broken attenuated x-ray, we have

$$\begin{aligned} \|S_{z_y \leftarrow y \leftarrow x}^A - S_{z_y \leftarrow y \leftarrow x}^B\|_{L^2(\mathcal{F}^X)} & \leq \|I_{y \leftarrow x}^{E(A,B)}(A - B)\|_{L^2(\mathcal{F}^X)} \\ & \quad + \|I_{z_y \leftarrow y}^{E(A,B)}(A - B)\|_{L^2(\mathcal{F}^X)}. \end{aligned}$$

The first term can be bounded by a multiple of $\|A - B\|_{L^2(\mathbb{D})}$ by lemma 2.17 with $k = 0$. For the second term, we have

$$\|I_{z_y \leftarrow y}^{E(A,B)}(A - B)\|_{L^2(\mathcal{F}^X)}^2 \leq C \int_{y \in \mathbb{D} \setminus \mathcal{U}} \int_0^T |(A - B)(\dot{\gamma}_{z_y \leftarrow y}(t))|^2 dt dy \tag{20}$$

where T is the length of the segment $\gamma_{z_y \leftarrow y}$.

We claim that the integral in the previous display can be rewritten as

$$\int_{y \in \mathbb{D} \setminus \mathcal{O}} |(A - B)_y(v_{y \leftarrow z_y})|^2 g(y) dy + \int_{z \in \mathcal{O} \cap \mathcal{U}^X} \int_{v \in J^+(z)} |(A - B)_z(v)|^2 g(z) dv dz$$

for some bounded nonnegative function $g : \mathbb{D} \rightarrow \mathbb{R}$ and where $J^+(z)$ is the set of unit length vectors in the future light cone of z with dv its standard measure induced by the Lebesgue measure. For $y \in \mathbb{D} \setminus \mathcal{U}$, we let $g(y)$ be the exit time (from \mathbb{D}) of the lightlike geodesic starting at y with velocity $v_{y \leftarrow z_y}$. For $y \in \mathcal{U}$, we let $g(y)$ be the difference between that same exit time and the exit time from \mathcal{U} . Note that by symmetry, g is well-defined on \mathcal{O} . The integrand of (20) is either of the form $|(A - B)_y(-v_{y \leftarrow z_y})|^2$ for some $y \in \mathbb{D} \setminus \mathcal{O}$ and each such term appears with frequency $g(y)$ or if $y \in \mathcal{O}$, then $A - B$ takes for input all the vectors in $J^+(y)$ again with frequency $g(y)$. This justifies the claim.

We can bound g uniformly and we have

$$|\omega_y(v_{y \leftarrow z_y})|^2 \leq c_1 \int_{v \in S_y \mathbb{D}} |\omega_y(v)|^2 dv,$$

$$\int_{v \in J^+(z)} |\omega_z(v)|^2 dv \leq c_2 \int_{v \in S_z \mathbb{D}} |\omega_z(v)|^2 dv$$

for all one forms ω where $S\mathbb{D} = \{v \in T\mathbb{D} : |v| = 1\}$ is the unit sphere bundle. The first inequality follows from choosing a basis v_0, v_1, v_2, v_3 for $T_y\mathbb{D}$ with $v_0 = v_{y \leftarrow z_y}$. Then $(\sum_{i=0}^3 |\omega_y(v_i)|^2)^{1/2}$ defines a norm on $T_y^*\mathbb{D}$ and the inequality follows from equivalence of norms. The second inequality also follows from equivalence of norms. Note that by symmetry, both c_1 and c_2 can be chosen uniformly. Therefore,

$$\|I_{z_y \leftarrow y}^{E(A,B)}(A - B)\|_{L^2(\mathcal{F}^X)}^2 \leq C \int_{y \in \mathbb{D}} \int_{v \in S_y \mathbb{D}} |(A - B)_y(v)|^2 dv dy$$

$$= C \|A - B\|_{L^2(\mathbb{D})}^2$$

and the result follows. □

Lemma 2.19. *Let A be a Hermitian connection. For every $k \geq 0$, there is a constant c_k such that*

$$\|P^A\|_{C^k(\mathbb{D} \times \mathbb{D})} \leq c_k(1 + \|A\|_{C^k(\mathbb{D})})^k.$$

Proof. Before proving the claim, we justify why P^A is indeed smooth whenever A is itself smooth. Let ∂ be the vector field on $\mathbb{D} \times \mathbb{D}$ given by

$$(\partial\Phi)(x, y) = \left. \frac{d}{dt} \right|_{t=0} \Phi(x, y + t(y - x)).$$

Note that the vector field $\partial_{y \leftarrow x}$ we defined earlier is equivalent to $\partial/|\partial|$. Then $P_{y \leftarrow x}^A$ can be characterised as the unique solution to

$$(\partial + A)U_A := \partial U_A(x, y) + A_y(y - x)U_A(x, y) = 0, \quad (x, y) \in \mathbb{D} \times \mathbb{D}$$

such that $U_A|_\Gamma = \text{Id}$ where $\Gamma = \{(x, x) : x \in \mathbb{D}\}$. To see U_A is smooth, we can consider the function

$$\tilde{U}_{x,y}(t) = (P_{x+t(y-x)\leftarrow x}^A, x, y) = (\tilde{U}_1, \tilde{U}_2, \tilde{U}_3).$$

It follows from the characterisation of parallel transport that $\tilde{U}_{x,y}$ solves the ODE

$$\begin{cases} \frac{d\tilde{U}_{x,y}}{dt}(t) = (-A_{\tilde{U}_2+t(\tilde{U}_3-\tilde{U}_2)}(\tilde{U}_3 - \tilde{U}_2), 0, 0) \\ \tilde{U}_{x,y}(0) = (\text{Id}, x, y) \end{cases}$$

and we can see x and y as initial conditions. By the standard theory of nonautonomous ODEs [Har82], the smoothness of A guarantees that $\tilde{U}_{x,y}$ depends smoothly on the initial conditions (x, y) and so

$$U_A(x, y) = \pi_1(\tilde{U}_{x,y}(1))$$

is smooth on $\mathbb{D} \times \mathbb{D}$.

Now to prove the claim, we will proceed inductively by using lemma 2.15. Let $\alpha, \beta \in \mathbb{Z}_{\geq 0}^4$ be multi-indices with $|\alpha| + |\beta| = k$. We can compute $\partial_x^\alpha \partial_y^\beta P_{y \leftarrow x}^A$ through a k -fold variation of the curve joining x and y . For $x, y \in \mathbb{D}$, we denote by $\gamma[x, y] : [0, 1] \rightarrow \mathbb{D}$ the curve

$$\gamma[x, y](t) := x + t(y - x).$$

Now consider the curve

$$\gamma_{\alpha,\beta}(t) = \gamma \left[x + \sum_{i=1}^4 \sum_{j=1}^{\alpha_i} a_j^i e_i, y + \sum_{i=1}^4 \sum_{j=1}^{\beta_i} b_j^i e_i \right] (t)$$

where $(e_i)_{i=1}^4$ is the canonical basis of \mathbb{R}^4 and the a_j^i, b_j^i are the k -fold variation's coefficients. Then,

$$\partial_x^\alpha \partial_y^\beta P_{y \leftarrow x}^A = \left. \frac{d^{|\alpha|}}{da^{|\alpha|}} \frac{d^{|\beta|}}{db^{|\beta|}} P_{\gamma_{\alpha,\beta}}^A \right|_{a=b=0}$$

where the right-hand side means that we differentiate $P_{\gamma_{\alpha,\beta}}^A$ once with respect to each coefficient a_j^i or b_j^i and then evaluate them all at 0.

We will denote by γ_k any k -fold variation corresponding to some $\partial_x^\alpha \partial_y^\beta P_{y \leftarrow x}^A$ and relabel the coefficients by s_1, \dots, s_k . We claim that for every $k \geq 0$, there is a constant c_k such that

$$\left| \frac{d}{ds_1} \cdots \frac{d}{ds_k} P_{\gamma_k}^A \right|_{s_1=\dots=s_k=0} \leq c_k (1 + \|A\|_{C^k(\mathbb{D})})^k$$

where c_k is independent of the variation γ_k . The case $k = 0$ corresponds to no variation at all and holds trivially. Let us now suppose that the claim holds for some $k - 1 \geq 0$ and all lower

order variations. Let γ_k be a k -fold variation as above and denote by γ_{k-1} the $(k - 1)$ -fold variation obtained by setting $s_k = 0$. Then, according to lemma 2.15, we have

$$\begin{aligned} & \frac{d}{ds_1} \cdots \frac{d}{ds_k} P_{\gamma_k}^A \Big|_{s=0} \\ &= \frac{d}{ds_1} \cdots \frac{d}{ds_{k-1}} \left(P_{\gamma_{k-1}}^A A_{\gamma_{k-1}(0)} (\partial_{s_k} \gamma_k(0)) \right. \\ & \quad - A_{\gamma_{k-1}(1)} (\partial_{s_k} \gamma_k(1)) P_{\gamma_{k-1}}^A \\ & \quad \left. + \int_0^1 P_{\gamma_{k-1}[t,T]}^A F_A(\dot{\gamma}_{k-1}(t), \partial_{s_k} \gamma_k(t)) P_{\gamma_{k-1}[0,t]}^A dt \right) \Big|_{s=0}. \end{aligned}$$

Taking norms on both sides, we can use the induction hypothesis and Leibniz’s rule to bound the first two terms by a multiple of $(1 + \|A\|_{C^{k-1}(\mathbb{D})})^{k-1} \|A\|_{C^{k-1}(\mathbb{D})}$. For the integral, the curvature F_A is obtained through differentiating A and so its C^{k-1} norm is bounded by the C^k norm of A . Hence, again through Leibniz’s rule, we can bound the integral by a multiple of $(1 + \|A\|_{C^{k-1}(\mathbb{D})})^{k-1} \|A\|_{C^k(\mathbb{D})}$. The claim readily follows.

The lemma then follows by taking a supremum over all the admissible variations γ_k since the previous claim then yields

$$\begin{aligned} \|P_{y \leftarrow x}^A\|_{C^k(\mathbb{D} \times \mathbb{D})} &= \sum_{j=0}^k \sup_{|\alpha|+|\beta|=j} |\partial_x^\alpha \partial_y^\beta P_{y \leftarrow x}^A| \\ &\leq \sum_{j=0}^k c_j (1 + \|A\|_{C^j(\mathbb{D})})^j \leq c_k (1 + \|A\|_{C^k(\mathbb{D})})^k. \end{aligned}$$

□

3. Gauge invariance

3.1. Gauge invariance

We now study the quantity $\|A - B - d_{E(A,B)} p\|$ in (6) in order to prove theorem 1.6. Recall the definition of p_ω^A as in (13). The p used in theorem 2.2 is actually $p_{A-B}^{E(A,B)}$, which we will denote as $p_{(A,B)}$ to emphasise the dependence on A and B more concisely. Let us denote

$$\Delta(A, B) := A - B - d_{E(A,B)} p_{(A,B)}.$$

Recall that $\varphi \in C^\infty(\mathbb{D}, U(n))$ is in \mathcal{H} if $\varphi|_O = \text{Id}$. The following lemma shows how $\Delta(A, B)$ changes under the action of \mathcal{H} .

Lemma 3.1. *Let $\Delta(A, B)$ be as above and let $\varphi, \psi \in \mathcal{H}$. Then*

- (a) $\Delta(A, B)^* = \Delta(B, A)$,
- (b) $\Delta(A \triangleleft \varphi, B) = \varphi^{-1} \Delta(A, B)$,
- (c) $\Delta(A \triangleleft \varphi, B \triangleleft \psi) = \varphi^{-1} \Delta(A, B) \psi$.

Proof. We start by proving (a). Let u be the solution along the segment $\gamma : [0, T] \rightarrow \mathbb{D}$ from z_y to y of

$$\begin{cases} \dot{u} + E(A, B)(\dot{\gamma}(t))u = -(A - B)(\dot{\gamma}(t)), \\ u(0) = 0. \end{cases}$$

We have

$$(E(A, B)u)^* = (Au - uB)^* = u^*A^* - B^*u^* = -E(B^*, A^*)u^* = E(B, A)u^*$$

where the last equality follows from the fact that both A and B are Hermitian. Hence, by taking the conjugate transpose on both sides of the equation defining u , we see that u^* solves

$$\begin{cases} (\dot{u}^*) + E(B, A)(\dot{\gamma}(t))u^* = -(B - A)(\dot{\gamma}(t)), \\ u^*(0) = 0. \end{cases}$$

It then follows from lemma 2.3 that

$$p_{(A,B)}^* = -u(T)^* = p_{(B,A)}.$$

Therefore, we can compute that

$$\begin{aligned} (d_{E(A,B)}p_{(A,B)})^* &= (dp_{(A,B)} + Ap_{(A,B)} - p_{(A,B)}B)^* \\ &= dp_{(A,B)}^* + p_{(A,B)}^*A^* - B^*p_{(A,B)}^* \\ &= dp_{(B,A)} + Bp_{(B,A)} - p_{(B,A)}A \\ &= d_{E(B,A)}p_{(B,A)} \end{aligned}$$

and so

$$\Delta(A, B)^* = (A - B - d_{E(A,B)}p_{(A,B)})^* = B - A - d_{E(B,A)}p_{(B,A)} = \Delta(B, A)$$

as claimed.

Let us now prove (b). Let u be as in the proof of (a) above and consider $\varphi^{-1}u$. We first claim that $\varphi^{-1}u$ solves

$$\begin{cases} (\varphi^{-1}\dot{u}) + E(A \triangleleft \varphi, B)(\dot{\gamma}(t))\varphi^{-1}u = -\varphi^{-1}(A - B), \\ [\varphi^{-1}u](0) = 0. \end{cases}$$

Indeed,

$$\begin{aligned} (\varphi^{-1}\dot{u}) + E(A \triangleleft \varphi, B)\varphi^{-1}u &= (\varphi^{-1}\dot{u}) + \varphi^{-1}\dot{u} + (\varphi^{-1}\dot{\varphi})\varphi^{-1}u \\ &\quad + (\varphi^{-1}A\varphi)\varphi^{-1}u - \varphi^{-1}uB \\ &= \varphi^{-1}\dot{u} + \varphi^{-1}Au - \varphi^{-1}uB \\ &= -\varphi^{-1}(A - B + E(A, B)u) + \varphi^{-1}E(A, B)u \\ &= -\varphi^{-1}(A - B). \end{aligned}$$

Now notice that

$$d_{E(A \triangleleft \varphi, B)} \varphi^{-1} = d(\varphi^{-1}) + \varphi^{-1}(d\varphi)\varphi^{-1} + \varphi^{-1}A\varphi\varphi^{-1} - \varphi^{-1}B = \varphi^{-1}(A - B)$$

since $d(\varphi^{-1}) = -\varphi^{-1}(d\varphi)\varphi^{-1}$. Therefore, we can rewrite

$$\varphi^{-1}(A - B) = d_{E(A \triangleleft \varphi, B)}(\varphi^{-1} - \text{Id} + \text{Id}) = A \triangleleft \varphi - B + d_{E(A \triangleleft \varphi, B)}(\varphi^{-1} - \text{Id}).$$

To make notation less cumbersome, we write p for $p_{(A, B)}$ and q for $p_{(A \triangleleft \varphi, B)}$. We can now apply lemma 2.3 again to get

$$\begin{aligned} \varphi^{-1}p &= -\varphi^{-1}u(T) \\ &= P_{y^* \leftarrow z_y}^{E(A \triangleleft \varphi, B)} I_{y^* \leftarrow z_y}^{E(A \triangleleft \varphi, B)} (A \triangleleft \varphi - B + d_{E(A \triangleleft \varphi, B)}(\varphi^{-1} - \text{Id})) \\ &= q + P_{y^* \leftarrow z_y}^{E(A \triangleleft \varphi, B)} I_{y^* \leftarrow z_y}^{E(A \triangleleft \varphi, B)} (d_{E(A \triangleleft \varphi, B)}(\varphi^{-1} - \text{Id})). \end{aligned}$$

Notice that $\varphi^{-1} - \text{Id}$ vanishes at z_y since $\varphi|_{\mathcal{O}} = \text{Id}$, and so lemma 2.4 yields

$$P_{y^* \leftarrow z_y}^{E(A \triangleleft \varphi, B)} I_{y^* \leftarrow z_y}^{E(A \triangleleft \varphi, B)} (d_{E(A \triangleleft \varphi, B)}(\varphi^{-1} - \text{Id})) = \varphi^{-1}(y) - \text{Id},$$

that is, $q = \varphi^{-1}p - \varphi^{-1} + \text{Id}$. We can now compute

$$\begin{aligned} d_{E(A \triangleleft \varphi, B)} q &= dq + (A \triangleleft \varphi)q - qB \\ &= d(\varphi^{-1}p) + (A \triangleleft \varphi)\varphi^{-1}p - \varphi^{-1}pB - d\varphi^{-1} - (A \triangleleft \varphi)\varphi^{-1} \\ &\quad + \varphi^{-1}B + A \triangleleft \varphi - B \\ &= (d\varphi^{-1})p + \varphi^{-1}dp + \varphi^{-1}(d\varphi)\varphi^{-1}p + \varphi^{-1}Ap - \varphi^{-1}pB \\ &\quad - d\varphi^{-1} - \varphi^{-1}(d\varphi)\varphi^{-1} - \varphi^{-1}A + \varphi^{-1}B + A \triangleleft \varphi - B \\ &= \varphi^{-1}dp + \varphi^{-1}Ap - \varphi^{-1}pB - \varphi^{-1}A + \varphi^{-1}B + A \triangleleft \varphi - B \\ &= \varphi^{-1}d_{E(A, B)}p - \varphi^{-1}(A - B) + A \triangleleft \varphi - B \end{aligned}$$

and therefore

$$\Delta(A \triangleleft \varphi, B) = A \triangleleft \varphi - B - d_{E(A \triangleleft \varphi, B)} q = \varphi^{-1}(A - B - d_{E(A, B)} p) = \varphi^{-1} \Delta(A, B)$$

as claimed.

Finally, to prove (c), it suffices to combine (a) and (b) to get

$$\Delta(A \triangleleft \varphi, B \triangleleft \psi) = \varphi^{-1} \Delta(B \triangleleft \psi, A)^* = \varphi^{-1} (\psi^{-1} \Delta(B, A))^* = \varphi^{-1} \Delta(A, B) \psi$$

where we used that ψ takes values in $U(n)$ and therefore $\psi^{-1} = \psi^*$. \square

3.2. Proof of theorem 1.6

We are now ready to prove theorem 1.6. The proof mostly amounts to using lemma 3.1 with the right choice of matrix fields in \mathcal{H} .

Proof of theorem 1.6. Recall that the gauge takes values in $U(n)$ and therefore if $\varphi, \psi \in \mathcal{G}$, lemma 3.1 yields

$$\|\Delta(A \triangleleft \varphi, B \triangleleft \psi)\| = \|\varphi^{-1} \Delta(A, B) \psi\| = \|\Delta(A, B)\|.$$

This shows that the estimate in theorem 2.2 is gauge invariant. Therefore, we can actually choose a gauge to compute $\|\Delta(A, B)\|$. We take $\varphi = P^A_{y \leftarrow z_y}$ and $\psi = P^B_{y \leftarrow z_y}$.

Using that $d_{E(A,B)} \text{Id} = A - B$, we can rewrite $p_{(A,B)}$ as

$$p_{(A,B)} = P^{E(A,B)}_{y \leftarrow z_y} I^{E(A,B)}_{y \leftarrow z_y} (d_{E(A,B)} \text{Id}) = \text{Id} - P^{E(A,B)}_{y \leftarrow z_y} \text{Id} = \text{Id} - P^A_{y \leftarrow z_y} P^B_{z_y \leftarrow y}$$

where we used lemma 2.4 for the second equality. Hence,

$$\begin{aligned} \Delta(A, B) &= A - B - d_{E(A,B)} p_{(A,B)} \\ &= A - B - d_{E(A,B)} (\text{Id} - P^{E(A,B)}_{y \leftarrow z_y} \text{Id}) \\ &= A - B - (A - B) + d_{E(A,B)} P^{E(A,B)}_{y \leftarrow z_y} \text{Id} \\ &= d_{E(A,B)} P^{E(A,B)}_{y \leftarrow z_y} \text{Id}. \end{aligned}$$

We can expand this last expression with the definition of $d_{E(A,B)}$ and lemma 2.7. This yields

$$\Delta(A, B) = d(P^A_{y \leftarrow z_y} P^B_{z_y \leftarrow y}) + AP^A_{y \leftarrow z_y} P^B_{z_y \leftarrow y} - P^A_{y \leftarrow z_y} P^B_{z_y \leftarrow y} B. \tag{21}$$

We have $P^{A \triangleleft \varphi}_{y \leftarrow z_y} = \varphi^{-1}(y) P^A_{y \leftarrow z_y} \varphi(z_y)$ from proposition 1.1 and since $\varphi|_{\mathcal{O}} = \text{Id}$, $\varphi(z_y) = \text{Id}$. We chose $\varphi = P^A_{y \leftarrow z_y}$, and so $P^{A \triangleleft \varphi}_{y \leftarrow z_y} = \text{Id}$. Similarly, we have $P^{B \triangleleft \psi}_{z_y \leftarrow y} = \text{Id}$. Therefore, plugging $A \triangleleft \varphi$ for A and $B \triangleleft \psi$ for B in (21) gives

$$\Delta(A \triangleleft \varphi, B \triangleleft \psi) = A \triangleleft \varphi - B \triangleleft \psi.$$

The result readily follows. □

3.3. Light-sink connections

Recall that we called a connection light-sink if $A \triangleleft P^A_{y \leftarrow z_y} = A$. We can characterise such connections.

Proposition 3.2. *Let A be a Hermitian connection on \mathbb{D} . Then $A \triangleleft P^A_{y \leftarrow z_y} = A$ if and only if*

$$A_y(\partial_t) = A_y(\partial_r)$$

for all $y \in \mathbb{D}$.

Here, r is the outward radial component in space, that is, $r^2 = x_1^2 + x_2^2 + x_3^2$. Changing to polar coordinates in the space variables, we can therefore write any light-sink connection as

$$A = A_0(dt + dr) + A_\vartheta d\vartheta + A_\phi d\phi$$

for some matrix fields A_0, A_ϑ, A_ϕ with values in $u(n)$.

Proof of proposition 3.2. Let $\gamma : [0, T] \rightarrow \mathbb{D}$ be the unit-speed lightlike geodesic from z_y to y . Then if $A = A \triangleleft P^A_{y \leftarrow z_y}$,

$$\begin{aligned} A(\dot{\gamma}(t)) &= (A \triangleleft P^A_{y \leftarrow z_y})(\dot{\gamma}(t)) \\ &= P^A_{z_y \leftarrow y} [dP^A_{y \leftarrow z_y}](\dot{\gamma}(t)) + P^A_{z_y \leftarrow y} A(\dot{\gamma}(t)) P^A_{y \leftarrow z_y} \\ &= 0 \end{aligned}$$

since $[dP_{y \leftarrow z_y}^A](\dot{\gamma}(t)) = -A(\dot{\gamma}(t))P_{y \leftarrow z_y}^A$ by definition of the parallel transport. In particular, by taking $t = T$, we get $A(v_{y \leftarrow z_y}) = 0$ and so

$$A_y(\partial_t) = A_y(\partial_r) \quad (22)$$

since $v_{y \leftarrow z_y} = \frac{1}{\sqrt{2}}(\partial_r - \partial_t)$. On the other hand, if (22) holds for all $y \in \mathbb{D}$, then $P_{z_y \leftarrow y}^A = \text{Id}$, and so $A \triangleleft P_{z_y \leftarrow y}^A = A$. \square

Since \mathcal{G} is a proper subgroup of \mathcal{H} , each \mathcal{G} -orbit does not necessarily contain a light-sink connection. However, proposition 1.8 guarantees that they are \mathcal{H} -equivalent to a unique light-sink connection.

Proof of proposition 1.8. Suppose that $B = A \triangleleft \varphi$ is a light-sink connection for some $\varphi \in \mathcal{H}$. Then the previous proof guarantees that

$$P_{y \leftarrow z_y}^{A \triangleleft \varphi} = \text{Id}$$

for all $y \in \mathbb{D}$ and so $\varphi(y) = P_{y \leftarrow z_y}^A$ by proposition 1.1 and the fact that $\varphi(z_y) = \text{Id}$ because $\varphi \in \mathcal{H}$. It also follows that

$$(A \triangleleft \varphi) \triangleleft P_{y \leftarrow z_y}^{A \triangleleft \varphi} = A \triangleleft P_{y \leftarrow z_y}^A$$

for all $\varphi \in \mathcal{H}$ and so the map $[A] \mapsto A \triangleleft P_{y \leftarrow z_y}^A$ is well-defined. \square

We now show how one can retrieve a connection from its scattering data and its unique \mathcal{H} -equivalent light-sink connection.

Proof of proposition 1.9. Since $A = B \triangleleft P_{y \leftarrow z_y}^B$, we have by proposition 1.1

$$S_{z_y \leftarrow y \leftarrow x}^A = S_{z_y \leftarrow y \leftarrow x}^{B \triangleleft P^B} = S_{z_y \leftarrow y \leftarrow x}^B P_{x \leftarrow z_x}^B$$

for all $(x, y) \in \mathcal{F}^X(\mathcal{U})$, and so we can determine $P_{y \leftarrow z_y}^B$ inside \mathcal{U}^X from the past-determined scattering data $S_{z_y \leftarrow y \leftarrow x}^A$ and $S_{z_y \leftarrow y \leftarrow x}^B$. The same applies on \mathcal{U}^Z with the future-determined scattering data $S_{z \leftarrow y \leftarrow x_y}^A$ and $S_{z \leftarrow y \leftarrow x_y}^B$. Hence, we can recover $P_{y \leftarrow z_y}^B$ on the whole of \mathcal{U} . Let $\Phi \in C^\infty(\mathbb{D}, U(n))$ be any smooth extension of $P_{z_y \leftarrow y}^B|_{\mathcal{U}}$ to the whole of \mathbb{D} . Then, by (5),

$$A \triangleleft \Phi = (B \triangleleft P_{y \leftarrow z_y}^B) \triangleleft \Phi = B \triangleleft [P_{y \leftarrow z_y}^B \Phi]$$

and so $A \triangleleft \Phi$ is gauge equivalent to B since $P_{y \leftarrow z_y}^B \Phi|_{\mathcal{U}} = \text{Id}$. \square

4. Statistical application

We show that, when restricting ourselves to light-sink connections, one can use Bayesian inversion to consistently recover a connection from its scattering data. To do so, we follow the approach first laid out in [MNP21]. Specifically, we will use theorem 5.1 in [BN21] as it only requires checking a nice set of conditions.

We will only consider light-sink connections as the problem then becomes injective. It would be natural to then only consider future-determined paths for the scattering data, but in doing so, the endpoints of our paths would never lie in $\mathcal{U} \setminus (\mathcal{U}^X \cup \mathcal{O})$. Therefore, we also need to consider past-determined paths.

To simplify the statistics slightly, we will only consider connections with values in $\mathfrak{so}(n)$ instead of $\mathfrak{u}(n)$. We do not lose any generality in doing so, but no longer have to deal with complex noise.

4.1. Bayesian approach to inverse problems

We start with a short summary of the Bayesian approach to solving inverse problems, as introduced in [Stu10]. Recall that for some forward map $\mathcal{G} : \Theta \rightarrow Y$ between Banach spaces, and $y \in Y$, we wish to find $\theta \in \Theta$ such that

$$y = \mathcal{G}(\theta).$$

Let us take $Y = L^2_\lambda(\mathcal{X}, \mathcal{V})$, the set of square-integrable functions on a probability space $(\mathcal{X}, \mathcal{F}, \lambda)$ with values in a finite-dimensional normed space \mathcal{V} . Rather than working with the whole infinite-dimensional L^2 space, we discretise it by considering the following regression model which mimics the setting of an experiment. Let $(X_i)_{i=1}^N$ be i.i.d. random variables on \mathcal{X} with distribution λ . These random variables correspond to experimental measurement of $\mathcal{G}_\theta = \mathcal{G}(\theta)$ with input X_i . Such measurements come with experimental noise that we model through the random variables

$$V_i = \mathcal{G}_\theta(X_i) + \mathcal{E}_i, \quad i = 1, \dots, N,$$

where the \mathcal{E}_i are i.i.d. standard Gaussian variables on \mathcal{V} , independent of the X_i .

Let $D_N = \{(V_i, X_i) : i = 1, \dots, N\} \subset (\mathcal{V} \times \mathcal{X})^N$ be the full data vector and let P_θ^N be its law. By making a choice of prior Π on the parameter space Θ , Bayes' rule yields a posterior distribution on Θ given the data D_N . For a Borel set $O \subset \Theta$, it is given by

$$\Pi^N(\theta \in O | D_N) = \frac{\int_O e^{\ell_N(\theta)} \Pi(d\theta)}{\int_\Theta e^{\ell_N(\theta)} \Pi(d\theta)}$$

where the log-likelihood is, up to additive constants,

$$\ell_N(\theta) = \ell_N(\theta | D_N) = -\frac{1}{2} \sum_{i=1}^N |V_i - \mathcal{G}_\theta(X_i)|_{\mathcal{V}}^2.$$

One can study how the posterior distribution Π^N behaves when N gets large. If we suppose there exists a unique underlying parameter $\theta_* \in \Theta$ from which the observations are made, we would want the posterior distribution to concentrate around θ_* (see [GN16, chapter 7.3] or [GvdV17]), that is, we would want that

$$\Pi^N(\|\theta - \theta_*\|_{L^2(\Theta)} > \delta_N | D_N) = o_{P_{\theta_*}^N}(1) \quad (23)$$

as $N \rightarrow \infty$ for some sequence $\delta_N \rightarrow 0$ that dictates the rate of convergence. Following substantial developments in the field, one should then get a good estimator $\hat{\theta}$ for θ_* by computing the expectation of the posterior distribution Π^N through MCMC sampling. Depending on the inverse problem, can we get estimates such as (23) and can we guarantee that the posterior mean indeed converges to θ_* , legitimising Bayesian inversion? This question has been studied for a range of different inverse problems and is an active area of research, see [MNP21] as well as [AN19, Boh21, GN20] for some examples.

4.2. Our setting

We consider the following experimental setup similar to the one we just described. Let λ be the uniform distribution on $\mathbb{S}^+(\mathbb{U})$ induced by the Lebesgue measure on \mathbb{D} and consider the random variables

$$(X_i, Y_i, Z_i)_{i=1}^N \sim \text{i.i.d. } \lambda$$

corresponding to random draws from $\mathbb{S}^+(\mathbb{U})$. For a light-sink Hermitian connection $A \in \Omega^1(\mathbb{D}, \mathfrak{so}(n))$, we denote its future-determined scattering data $S_{z \leftarrow y \leftarrow x}^A$ by $S_+^A(x, y)$ and its past-determined scattering data $S_{z \leftarrow y \leftarrow x}^A$ by $S_-^A(y, z)$. Suppose that we observe noisy versions of the scattering data corresponding to both types of paths according to our random draws, that is, we observe

$$S_i = (S_+^A(X_i, Y_i) + \mathcal{E}_i^+, S_-^A(Y_i, Z_i) + \mathcal{E}_i^-) \quad i = 1, \dots, N. \quad (24)$$

The matrices \mathcal{E}_i^\pm correspond to independent Gaussian noise in the sense that $\mathcal{E}_i^\pm = (\varepsilon_{i,j,k}^\pm)_{1 \leq j,k \leq n}$ and all the $\varepsilon_{i,j,k}^\pm$'s are i.i.d. $N(0, 1)$ that are independent from the other random variables. We denote by P_A^N the joint law of the random variables $(S_i, (X_i, Y_i, Z_i))_{i=1}^N$.

In order to estimate the connection A from D_N , we need to choose a prior Π on the space of $\mathfrak{so}(n)$ -valued light-sink connections. Any such connection can be represented by three skew-symmetric matrix fields since

$$A = A_0(dt + dr) + A_\vartheta d\vartheta + A_\phi d\phi,$$

and therefore by $d_n := 3 \dim \mathfrak{so}(n) = 3n(n-1)/2$ continuous functions on \mathbb{D} . Following [BN21], we choose the prior Π by prescribing an orthonormal basis on $L^2(\mathbb{D}, \mathbb{R})$ as well as a sequence of positive scalars. For conciseness, we choose as basis the normalised eigenfunctions $(e_j)_{j \in \mathbb{N}}$ of the Laplacian with Neumann boundary conditions and choose their eigenvalues $(\lambda_j)_{j \in \mathbb{N}}$ as scalars. It follows from classical L^∞ estimates for eigenfunctions from [Hör68] and Weyl's law [Hör09] that we can choose $\tau = 3/4$ and $d = 4$ in condition 3.1 of [BN21]. This choice gives rise to Sobolev-type spaces

$$H^s(\mathbb{D}, \mathbb{R}) = \left\{ f \in L^2(\mathbb{D}, \mathbb{R}) : \sum_{j \in \mathbb{N}} \lambda_j^s \langle f, e_j \rangle_{L^2(\mathbb{D}, \mathbb{R})}^2 < \infty \right\}$$

which, in this specific case, agree with the usual Sobolev spaces. These Sobolev spaces naturally induce Sobolev spaces on the space of light-sink connections

$$H^s = H^s(\mathbb{D}, \mathfrak{so}(n)^3) \cong H^s(\mathbb{D}, \mathbb{R}^{d_n}) = \times_{i=1}^{d_n} H^s(\mathbb{D}, \mathbb{R})$$

which plays the role of our parameter space Θ . The eigenfunctions (e_j) naturally induce the basis $\{e_{j,i} : 1 \leq i \leq d_n, j \in \mathbb{N}\}$ on $H^s(\mathbb{D}, \mathfrak{so}(n)^3)$ where

$$e_{j,i} = (\delta_{i,1} e_j, \dots, \delta_{i,d_n} e_j), \quad \delta_{i,j} = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases}$$

For D , an integer multiple of d_n , let E_D be the span of the first D vectors of the basis, that is,

$$E_D := \{e_{j,i} : 1 \leq i \leq d_n, 1 \leq j \leq D/d_n\}.$$

For $\alpha > 0$, we take as prior on E_D

$$A = N^{-1/(\alpha+2)} \sum_{i \leq d_n} \sum_{j \leq D/d_n} \lambda_n^{-\alpha/2} g_{j,i} e_{j,i}, \quad g_{j,i} \sim \text{i.i.d. } \mathcal{N}(0, 1). \tag{25}$$

Theorem 4.2 can also be proved for $D \rightarrow \infty$, giving rise to a commonly used Matérn prior of order α for the Laplacian, see e.g. [GvdV17, chapter 11]. For simplicity, we take a truncated prior as it reflects what happens in practice. We denote the law of A by Π and its density by π . Through Bayes' rule, the choice of prior gives rise to the posterior distribution

$$\Pi(A \in O | (S_i, (X_i, Y_i, Z_i))_{i=1}^N) = \frac{\int_O e^{\ell_N(A)} d\Pi(A)}{\int_{E_D} e^{\ell_N(A)} d\Pi(A)}, \quad O \subseteq E_D \text{ Borel}$$

with log-likelihood given by

$$\ell_N(A) = -\frac{1}{2} \sum_{i=1}^N |S_i - S(A)(X_i, Y_i, Z_i)|_{(\mathbb{R}^n \times \mathbb{R}^n)^2} - \frac{1}{2} N \delta_N^2 \|A\|_{H^\alpha}^2$$

up to some additive constant, where $\delta_N = N^{-\alpha/(2\alpha+4)}$. See [BN21] for more details.

4.3. Statistical guarantees for light-sink connections

To apply [BN21, theorem 5.1], it remains to show that the map $S : A \mapsto (S_+^A, S_-^A)$ satisfies their condition 3.2 which contains three parts. In our setting, condition 3.2 amounts to the following three statements.

(a) (Uniform boundedness). There exists a constant $C \geq 1$ such that

$$\sup_{A \in \mathcal{L}, (x,y,z) \in \mathbb{S}^+(\mathcal{U})} |S(A)(x, y, z)| \leq C.$$

(b) (Global Lipschitz). There exists a constant $C > 0$ such that

$$\begin{aligned} \|S(A) - S(B)\|_{L^\infty(\mathbb{S}^+(\mathcal{U}))} &\leq C \|A - B\|_{L^\infty(\mathbb{D})}, \\ \|S(A) - S(B)\|_{L^2(\mathbb{S}^+(\mathcal{U}))} &\leq C \|A - B\|_{L^2(\mathbb{D})}. \end{aligned}$$

(c) (Inverse continuity modulus). For every M there exists a constant L' and $0 < \gamma \leq 1$ such that for all $\delta > 0$ small enough and the given $\alpha > 0$,

$$\sup \{ \|A - B\|_{L^2(\mathbb{D})} : \|A\|_{H^\alpha} + \|B\|_{H^\alpha} \leq M, \|S(A) - S(B)\|_{L^2(\mathbb{S}^+(\mathcal{U}))} \leq \delta \} \leq L' \delta^\gamma.$$

The uniform boundedness condition is immediately satisfied since S^A takes values in $U(n)$. The global Lipschitz condition follows immediately from lemma 2.16 in the L^∞ case and from lemma 2.18 in the L^2 case. Hence, it only remains to show that the inverse continuity modulus condition holds. This is the content of the next lemma.

Lemma 4.1. *The inverse continuity modulus condition as described above holds.*

Proof. First, note that we can find $C_\varepsilon > 0$ that depends on ε such that

$$\|S(A) - S(B)\|_{L^2(\mathbb{S}^+(\mathcal{U}))}^2 \geq C_\varepsilon \left(\|S_+^A - S_+^B\|_{L^2(\mathcal{F}X)}^2 + \|S_-^A - S_-^B\|_{L^2(\mathcal{F}Z)}^2 \right).$$

By the interpolation inequality for Sobolev spaces (see [Boh21, lemma 7.4]), for every $s > 0$,

$$\|\cdot\|_{H^1(\mathcal{F}^X)} \lesssim_s \|\cdot\|_{L^2(\mathcal{F}^X)}^{1-1/s} \|\cdot\|_{H^s(\mathcal{F}^X)}^{1/s}.$$

Here, \lesssim_s means that the left-hand side is bounded by a constant multiple of the right-hand side where the constant is allowed to depend on s . By applying this inequality to $[S_+^A]^{-1}S_+^B - \text{Id}$ and using the pseudolinearisation identity, we get

$$\|[S_+^A]^{-1}S_+^B - \text{Id}\|_{H^1(\mathcal{F}^X)} \lesssim_s \|S_+^A - S_+^B\|_{L^2(\mathcal{F}^X)}^{1-1/s} \|I_{y^x \leftarrow y^x - y}^{E(A,B)}(A - B)\|_{H^s(\mathcal{F}^X)}^{1/s}.$$

Since A and B are light-sink connections, the broken attenuated x-ray is equal to the simple attenuated x-ray $I_{y^x \leftarrow x}^{E(A,B)}(A - B)$. For $s = k$ an integer, lemmas 2.17 and 2.19 yield

$$\|I_{y^x \leftarrow x}^{E(A,B)}(A - B)\|_{H^k(\mathcal{F}^X)} \lesssim_k \|A - B\|_{H^k(\mathcal{F}^X)} (1 + \|E(A, B)\|_{C^k(\mathbb{D})})^k.$$

Taking $\alpha > k$ sufficiently large such that $\|A\|_{H^\alpha} + \|B\|_{H^\alpha} \leq M$, we can bound the C^k norms of A and B via Sobolev embedding inequalities. Note that $\alpha = k + 3$ suffices. This in turn allows us to bound the C^k -norm of $E(A, B)$, and so

$$\|[S_+^A]^{-1}S_+^B - \text{Id}\|_{H^1(\mathcal{F}^X)} \lesssim_{k,M} \|S_+^A - S_+^B\|_{L^2(\mathcal{F}^X)}^{1-1/k}.$$

It follows from theorem 1.4 for light-sink connections ($p = 0$ by theorem 1.6) that

$$\|A - B\|_{L^2(\mathbb{D} \setminus \cup)} \lesssim_{\varepsilon,k,M} \|S_+^A - S_+^B\|_{L^2(\mathcal{F}^X)}^{1-1/k}.$$

Similar inequalities also hold on \cup^X and \cup^Z by theorem 1.3. Hence, we can choose $\gamma = \frac{\alpha-4}{\alpha-3}$ for $\alpha \geq 5$ by taking $\alpha = k + 3$. □

We can finally apply theorem 5.1 in [BN21] to get the following estimate regarding the concentration of the posterior distribution around the real parameter A_* obtained through noisy samples of S^{A_*} as the number of samples goes to infinity.

Theorem 4.2. *Let the posterior distribution $\Pi(\cdot | (S_i, (X_i, Y_i, Z_i))_{i=1}^N)$ arise from the prior (25) with $\alpha \geq 5$ and data $(S_i, (X_i, Y_i, Z_i)) \sim P_{A_*}^N$ as in (24). Suppose that $A_* \in H^\alpha$ and $D \simeq N^{2/(\alpha+2)}$. Let $\gamma = \frac{\alpha-4}{\alpha-3}$. Then, there is $M > 0$ such that*

$$\Pi(\|A - A_*\|_{L^2(\mathbb{D})} > M\delta_N^\gamma | (S_i, (X_i, Y_i, Z_i))_{i=1}^N) = o_{P_{A_*}^N}(1)$$

as $N \rightarrow \infty$, where $\delta_N = N^{-\alpha/(2\alpha+4)}$.

In short, the posterior distribution converges to a delta distribution about A_* in $P_{A_*}^N$ -probability at a rate that depends on the smoothness of the prior and of A_* . The smoother A_* is, the smoother we can choose the prior, and the faster the posterior distribution concentrates. Moreover, by the same arguments used at the end of [MNP21] to complete their proof of their theorem 3.2, one can expect the rate of theorem 4.2 to carry over to the posterior mean, that is,

$$\|E^\Pi[A | (S_i, (X_i, Y_i, Z_i))_{i=1}^N] - A_*\|_{L^2(\mathbb{D})} = O_{P_{A_*}^N}(\delta_N^\gamma)$$

as $N \rightarrow \infty$.

Note that we have convergence to A_* in theorem 4.2 and not only its projection $A_{*,D}$ on E_D as in the statement of theorem 5.1 in [BN21]. This is due to the fact that the estimate in lemma 4.1 holds for all A, B in the whole parameter space $H^\alpha(\mathbb{D}, \mathfrak{so}(n)^3)$, and not just E_D . Indeed, remark 5.2 in [BN21] guarantees that we can then replace $A_{*,D}$ by A_* .

Acknowledgments

I would like to thank Gabriel Paternain for suggesting this project and for his guidance. I would also like to thank Richard Nickl, Lauri Oksanen and Jan Bohr for their helpful comments. Furthermore, I thank both anonymous referees for their valuable feedback. The research of the author is supported by the Cambridge Trust, NSERC's PGS D scholarship and the CCIMI.

Data availability statement

No new data were created or analysed in this study.

ORCID iDs

Simon St-Amant  <https://orcid.org/0000-0003-2031-2709>

References

- [ALJ19] Ambartsoumian G and Latifi Jebelli M J 2019 The V-line transform with some generalizations and cone differentiation *Inverse Problems* **35** 034003
- [Amb12] Ambartsoumian G 2012 Inversion of the V-line Radon transform in a disc and its applications in imaging *Comput. Math. Appl.* **64** 260–5
- [AN19] Abraham K and Nickl R 2019 On statistical Calderón problems *Math. Stat. Learn.* **2** 165–216
- [AS09] Arridge S R and Schotland J C 2009 Optical tomography: forward and inverse problems *Inverse Problems* **25** 123010
- [BN21] Bohr J and Nickl R 2021 On log-concave approximations of high-dimensional posterior measures and stability properties in non-linear inverse problems (arXiv:2105.07835)
- [Boh21] Bohr J 2021 Stability of the non-abelian x-ray transform in dimension ≥ 3 *J. Geom. Anal.* **31** 11226–69
- [CLOP21a] Chen X, Lassas M, Oksanen L and Paternain G P 2022 Detection of Hermitian connections in wave equations with cubic non-linearity *J. Eur. Math. Soc.* **24** 2191–232
- [CLOP21b] Chen X, Lassas M, Oksanen L and Paternain G P 2021 Inverse problem for the Yang–Mills equations *Commun. Math. Phys.* **384** 1187–225
- [FMS11] Florescu L, Markel V A and Schotland J C 2011 Inversion formulas for the broken-ray Radon transform *Inverse Problems* **27** 025002
- [GN16] Giné E and Nickl R 2016 *Mathematical Foundations of Infinite-Dimensional Statistical Models* (Cambridge Series in Statistical and Probabilistic Mathematics vol 40) (Cambridge: Cambridge University Press)
- [GN20] Giordano M and Nickl R 2020 Consistency of Bayesian inference with Gaussian process priors in an elliptic inverse problem *Inverse Problems* **36** 085001
- [GvdV17] Ghosal S and van der Vaart A 2017 *Fundamentals of Nonparametric Bayesian Inference* (Cambridge Series in Statistical and Probabilistic Mathematics vol 44) (Cambridge: Cambridge University Press)
- [Har82] Hartman P 1982 *Ordinary Differential Equations* 2nd edn (Boston, MA: Birkhäuser)
- [Hör68] Hörmander L 1968 The spectral function of an elliptic operator *Acta Math.* **121** 193–218
- [Hör09] Hörmander L 2009 *The Analysis of Linear Partial Differential Operators: IV. Fourier Integral Operators* (Classics in Mathematics) (Berlin: Springer) Reprint of the 1994 edition
- [KLU18] Kurylev Y, Lassas M and Uhlmann G 2018 Inverse problems for Lorentzian manifolds and non-linear hyperbolic equations *Invent. Math.* **212** 781–857
- [MNP21] Monard F, Nickl R and Paternain G P 2021 Consistent inversion of noisy non-abelian x-ray transforms *Commun. Pure Appl. Math.* **74** 1045–99
- [Nov19] Novikov R G 2019 5. Non-abelian Radon transform and its applications *The Radon Transform: The First 100 Years and Beyond* (Berlin: de Gruyter & Co) pp 115–28

- [PS20] Paternain G P and Salo M 2020 The non-abelian x-ray transform on surfaces (arXiv:2006.02257)
- [PSU21] Paternain G P, Salo M and Uhlmann G 2021 Geometric inverse problems with emphasis on two dimensions https://dpmms.cam.ac.uk/~7Egpp24/GIP2D_driver.pdf
- [Sha94] Sharafutdinov V A 1994 *Integral Geometry of Tensor Fields (Inverse and Ill-posed Problems Series)* (Utrecht: VSP)
- [Sin78] Singer I M 1978 Some remarks on the Gribov ambiguity *Commun. Math. Phys.* **60** 7–12
- [Stu10] Stuart A M 2010 Inverse problems: a Bayesian perspective *Acta Numer.* **19** 451–559