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.

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

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

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

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

We implicitly write $\mathcal{U}$ for $\mathcal{U}_{\varepsilon_0}$ and write $\mathcal{U}_\varepsilon$ whenever we want to emphasise the dependence on $\varepsilon$. Given $x, y \in D$, we write $x < y$ if there is a future-pointing causal curve from $x$ to $y$. We also write $(x, y) \in \mathcal{L}$ if $x < y$ and there is a lightlike geodesic from $x$ to $y$.

Recall that a line segment $\gamma : [0, T] \to \mathbb{R}^{1+3}$, $\gamma(s) = x + sv$ for $x \in \mathbb{R}^{1+3}, v = (v_0, v_1, v_2, v_3) = (v_0, v_1, v_2, v_3)$ 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 $\gamma$ 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, $(x, y) \in \mathcal{L}$ if and only if $y$ is in the future light cone of $x$, or equivalently, $x$ is in the past light cone of $y$.

Date: January 21, 2022.
We will work with Hermitian connections on the trivial bundle $\mathbb{D} \times \mathbb{C}^n$. Such a connection $A$ is a $\mathfrak{u}(n)$-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 $A_i \in C^\infty(\mathbb{D}, \mathfrak{u}(n))$. We denote the set of Hermitian connections on $\mathbb{D}$ by $\mathcal{H}$. A connection induces a covariant derivative on functions $f : \mathbb{D} \to \mathbb{C}^n$ given by $d_A f = df + Af$. Given a smooth curve $\gamma : [0, T] \to \mathbb{D}$, the parallel transport isomorphism $P_\gamma^A : \mathbb{C}^n \to \mathbb{C}^n$ 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 $\gamma$ is $P_\gamma^A v := U(T)v$.

One can check that $P_\gamma^A$ does not depend on the parametrisation of $\gamma$ and that it takes values in $U(n)$ since $A$ is Hermitian. Given $x, y \in \mathbb{D}$, we denote by $P_\gamma^A_{y \leftarrow x}$ the parallel transport from $x$ to $y$ along the straight line between the two points. 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

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

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 $z$. We denote by

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

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

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

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 $\varphi$ 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 $\gamma : [0, T] \to \mathbb{D}$,

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

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

$$S^A_{z \leftarrow y \leftarrow x} = S^A_{z \leftarrow y \leftarrow x}$$

for all $(x, y, z) \in \mathcal{S}^+(\mathcal{U})$. 
Therefore, the scattering data of $A \triangleleft \varphi$ coincides with that of $A$ whenever $\varphi$ is in the gauge group

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

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 $A \in C^\infty(\mathbb{R}^d \times S^{d-1}, \mathbb{C}^{n \times n})$ the scattering data map

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

where $\psi^+$ is the unique solution of the transport equation

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

such that

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

Given that $A$ decays sufficiently fast as $|x| \to \infty$, 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 $f \in C^2_{c}(\mathcal{U}, \mathbb{C}^n)$ to

$$L_Af = \phi|_{t=1}$$

where $\phi$ solves

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

Here $\Box_A$ is the connection wave operator given by $\Box_A = d^*A \otimes A$. Note that when $A = 0$, we recover the usual wave operator $\Box = \partial_t^2 - \Delta$. 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 $S^{A}_{y \to y \to x}$ for all $(x, y, z) \in S^+(\mathcal{U})$. Injectivity up to gauge of $L_A$ then follows from that of the broken X-ray transform.

To determine $S^{A}_{y \to y \to x}$ 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 $\phi$ 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 $S^{A}_{y \to y \to x}$. 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 short summary to the Bayesian approach to solving inverse problems, as introduced in [Stu10].

For some mapping $\mathcal{G} : \Theta \to Y$ between Banach spaces, and $y \in Y$, we wish to find $\theta \in \Theta$ such that

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

Let us take $Y = L^2_{\lambda}(\mathcal{X}, \mathcal{Y})$, the set of square-integrable functions on a probability space $(\mathcal{X}, \lambda)$ with values in a finite-dimensional normed space $\mathcal{Y}$. 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 $\lambda$. 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

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

where the $\mathcal{E}_i$ are i.i.d. standard Gaussian variables on $\mathcal{Y}$, independent of the $X_i$. 
In our setting, the set $\Theta$ could be the set of Hermitian connections $A$ on $\mathbb{D}$, $Y$ the set of matrix fields on $S^+(\mathbb{D})$ and $\theta \mapsto G_\theta$ the mapping that sends a connection $A$ to its scattering data $S^A$. Each $X_i$ then amounts to a random choice of path $z \leftarrow y \leftarrow x$ in $S^+(\mathbb{D})$ and $V_i$ a noisy version of $S^A_{z \leftarrow y \leftarrow x}$.

Let $D_N = \{(V_i, X_i) : i = 1, \ldots, N\} \subset (V \times X)^N$ be the full data vector and let $P^N_\theta$ be its law. By making a choice of prior $\Pi$ on the parameter space $\Theta$, Bayes’ rule yields a posterior distribution on $\Theta$ 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,

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

One can study how the posterior distribution $\Pi^N$ 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 $\theta_*$ (see [GN16, Chapter 7.3] or [GvdV17]), that is, we would want that

$$
\Pi^N(\|\theta - \theta_*\|_{L^2(\theta)} > \delta_N | D_N) = o_{P^N_{\theta_*}}(1)
$$
as $N \to \infty$ for some sequence $\delta_N \to 0$ that dictates the rate of convergence. Following substantial developments in the field, one should then get a good estimator $\hat{\theta}$ for $\theta_*$ by computing the expectation of the posterior distribution $\Pi^N$ through MCMC sampling. Depending on the inverse problem, can we get estimates such as (2) and can we guarantee that the posterior mean indeed converges to $\theta_*$, 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.

However, in the case of the broken non-abelian X-ray transform, the map $G : A \mapsto S^A$ is not injective and so the true underlying parameter is not uniquely identifiable. Indeed, all connections in the same $G$-orbit yield the same scattering data. Moreover, these orbits are all infinite dimensional. 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 $\Theta$ and that $G$ is injective up to the action of $G$. This means that for every $g \in G$ and $\theta \in \Theta$, we have $G(\theta \circ g) = G(\theta)$ and that $G(\theta_1) = G(\theta_2)$ if and only if there is $g' \in G$ with $\theta_1 = \theta_2 \circ g'$. Then $G$ naturally induces an injective map on the quotient space

$$
\tilde{G} : \Theta/G \to Y.
$$

One could try to prove statistical guarantees for this map. However, as settings such as the present one where $G$ is non-linear, the quotient space $\Theta/G$ 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 $s : \Theta/G \to \Theta$ 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 \( s_H : \Theta/H \to \Theta \). 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 \( \theta_* \). 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 \( S^+(\mathcal{U}) \). 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 \( z \leftarrow y \leftarrow x \) where \( x \) is the unique point such that \( (x, y) \in \mathcal{L} \) and \( x \in \mathcal{O} \), that is, \( x = (t, 0, 0, 0) \) for some \( t \in [-1, 1] \). Similarly, a future-determined path is a path of the form \( z \leftarrow y \leftarrow x \) for \( (x, y, z) \in \mathcal{S}^+(\mathcal{U}) \) where now \( z \) is the unique point on \( \mathcal{O} \) such that \( (y, z) \in \mathcal{L} \). We denote the corresponding scattering data as

\[
S^A_{z \leftarrow y \leftarrow x} \quad \text{and} \quad S^A_{z \leftarrow y \leftarrow x}.
\]

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 \( z \leftarrow y \leftarrow x \). Indeed, in polar coordinates \((t, r, \vartheta, \phi)\), the tangent vector along the path \( z \leftarrow y \) is \( \frac{1}{\sqrt{2}} \left( \frac{\partial}{\partial t} - \frac{\partial}{\partial r} \right) \) while the tangent vector along the path \( y \leftarrow x \) is \( \frac{1}{\sqrt{2}} \left( \frac{\partial}{\partial t} + \frac{\partial}{\partial r} \right) \). 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 endpoint \( x \) 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 \( \pi : \mathcal{F}^X \to \mathbb{D} \setminus \mathcal{U} \), 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

\[
\mathcal{F}^X_y = \{ x \in \mathcal{U}^X : (x, y) \in \mathcal{L} \} \cong B_3.
\]
Similarly, the set of admissible past-determined paths can be represented through the $B_3$-bundle $\pi : F^Z \rightarrow \mathbb{D} \setminus \mathcal{U}$ with fibre

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

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

For two 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

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

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

$$\begin{equation}
(\partial \Phi)(x,y) = \partial_{y\leftarrow x}\Phi(x, y) := \frac{d}{dt}\Phi(x, y + tv_{y\leftarrow x}) \bigg|_{t=0}.
\end{equation}$$

Note that if $(x, y) \in F^X_y$ and the domain of $\Phi$ is $F^X_y$, the operators $\partial_{y\leftarrow x}$ and $\partial_{x\leftarrow y}$ are both well-defined since $x \in F^X_{y+tv_{y\leftarrow x}}$ and $x + tv_{x\leftarrow y} \in F^X_y$ whenever $x \in F^X_y$ and $t$ is sufficiently small. One can see $\partial_{y\leftarrow x}$ and $\partial_{x\leftarrow y}$ as horizontal and vertical vector fields on $F^X_y$, respectively. Similarly, $\partial_{y\leftarrow z}$ and $\partial_{z\leftarrow y}$ are well defined operators if $(y, z) \in F^Z_y$ and the domain of $\Phi$ is $F^Z_y$.

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

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

where $dx$ is the natural measure on $F^X_y$ induced by Euclidean space. Note that this norm scales down as $\varepsilon$ goes to 0 at a rate of $\varepsilon^3$.

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 $\omega$ on $\mathbb{D}$ at any given point $y \in \mathbb{D}$ by seeing $\omega_y$ 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 $E(A, B)^* = -E(A, B)$.
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 (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 $x$.

1.6. Main results. We state our results only for future-determined paths, but equivalent statements hold for past-determined paths by seeing $[S_{z_y\leftarrow y\leftarrow x}]^{-1}$ 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 $\varepsilon$ such that

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

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

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

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 $\overline{\mathcal{O}}^X$. By the same estimate for past-determined paths, the connections must also agree on $\overline{\mathcal{O}}^Z$ and hence they agree on $\overline{\mathcal{O}}$. Theorem 1.4 yields $p \in C^\infty(\mathbb{D}, \mathbb{C}^{n \times n})$ such that

\[ B = A - d_{E(A,B)}p = A - dp - Ap + pB \]

on $\mathbb{D} \setminus \overline{\mathcal{O}}$. Let $\varphi = \text{Id} - p$. As the proof of Theorem 2.2 will reveal, $\varphi$ takes values in $U(n)$ since actually $\varphi = P_{y \mathcal{L} y}^A P_{z \mathcal{L} z}^B$. We can rewrite (5) as

\[ \varphi B = A \varphi + d \varphi \]

and so $B = A \bowtie \varphi$. It follows that $A$ and $B$ are gauge equivalent since they agree on $\overline{\mathcal{O}}$ and so $\varphi|_X = \text{Id}$. The converse implication is the statement of Proposition 1.1. $\Box$

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 (4) in a way that highlights this.

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

\[ \|(A - B - d_{E(A,B)}p)y\| = \|(A \bowtie P_{y \mathcal{L} y}^A - B \bowtie P_{y \mathcal{L} y}^B)y\| \]

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

\[ \|(A \bowtie P_{y \mathcal{L} y}^A - B \bowtie P_{y \mathcal{L} y}^B)\|_{L^2(\mathbb{D} \setminus \overline{\mathcal{O}})} \leq C \Psi(A, B) \|S^A - S^B\|_{H^1(F^X_\mathcal{O})} \]

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 $A \bowtie P_{y \mathcal{L} y}^A = A$. We call them light-sink connections. They form a linear space 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 $\rho : \mathcal{Y} / \mathcal{H} \to \mathcal{Y}$,

$$\rho([A]) := A \circ P_{y^{-z}}$$

is well-defined.

The map $\rho$ is almost a fixing of the gauge. Contrary to that of $\mathcal{G}$, the action of $\mathcal{H}$ on $\mathcal{Y}$ does not preserve the scattering data. Therefore, the map $\rho$ 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 $\rho$ to make them gauge equivalent.

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

$$A = B \circ P_{y^{-z}}.$$ 

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

The map $\Phi$ is defined up to an extension operator

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

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 $F : [0, 1] \times GL(n, \mathbb{C}) \to GL(n, \mathbb{C})$ such that $F(0, x) = x$ and $F(1, x) \in U(n)$ for all $x \in GL(n, \mathbb{C})$, and $F(t, \cdot)|_{U(n)} = Id$ 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 $A = B \circ \varphi$ for some $\varphi \in \mathcal{G}$.

Our results give the following strategy to do so.

(1) 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.

(2) Minimise the mapping

$$A \mapsto \|S_{y^{-x}}^A - S_{y^{-x}}^B P_{x^{-z}}^B\|_{H^1(F^X)}$$

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

(3) By Corollary 1.7 and the definition of $\rho$, the unique minimiser of this problem is $A = \rho([B])$.

(4) Use Proposition 1.9 to get a connection $A \circ \Phi$ that is gauge-equivalent to $B$.

Note that in step (2), $S_{y^{-x}}^B P_{x^{-z}}^B$ 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 (2) 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 (1) using Bayesian inversion. As steps (3) and (4) 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.

Acknowledgements. 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. This research was supported by the Cambridge Trust, NSERC’s PGS D scholarship and the CCIMI.

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 $x \in \mathcal{U}$,

$$\| (A - B)_x \| \leq C \sup_{y \in \mathcal{D} \setminus \mathcal{U}} | \partial_{x \leftarrow y} \left( S_{2y \leftarrow y \leftarrow x}^A [S_{2y \leftarrow y \leftarrow x}^B]^{-1} \right) |.$$

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

$$\| (A - B - dE_{(A,B)p})_y \| \leq C \frac{\varepsilon^4}{\varepsilon^4} \int_{(y \leftarrow x)_y} | \partial_{y \leftarrow x} \left( [S_{2y \leftarrow y \leftarrow x}^A]^{-1} S_{2y \leftarrow y \leftarrow x}^B \right) | dx.$$

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, $\omega$ 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 $\omega$ along $\gamma$ with respect to $A$ is given by

$$I^A_\gamma(\omega) := \int_0^T P^A_{\gamma(0) \leftarrow \gamma(t)} \omega(\dot{\gamma}(t)) dt.$$

Similar to the parallel transport, we can express $I^A_\gamma(\omega)$ 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

$$(\dot{u} + A(\dot{\gamma}(t)) u = -\omega(\dot{\gamma}(t)),
\quad u(0) = u_0.)$$
Then $u(T) = P^A_\gamma(u_0 - I^A_\gamma(\omega))$.

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 $(U^{-1}) = U^{-1}A$. Therefore, along $\gamma$ we have

$$(U^{-1}u) = U^{-1}Au + U^{-1}\dot{u} = U^{-1}(Au + \dot{u}) = -U^{-1}\omega.$$ Integrating both sides from 0 to $T$ yields

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

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

If $A$ vanishes identically, the attenuated X-ray is simply the integral of the one-form $\omega$ along $\gamma$, and so if $\omega$ is potential ($\omega = df$ for some $f \in C^\infty(D, \mathbb{C}^n)$), then the attenuated X-ray of $\omega$ is the difference between the values of $f$ at both endpoints of $\gamma$ 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 $I^A_\gamma(\omega)$. 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 $f : D \to \mathbb{C}^n$ be a smooth function on $D$. Then

$$I^A_\gamma(d_A f) = (P^A_\gamma)^{-1}f(\gamma(T)) - f(\gamma(0))$$

where $d_A f = df + Af$.

Recall the definition of $\partial_{y \leftarrow x}$ as in (3). 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 $\omega$ be a one-form on $D$. For $x \neq y$,

$$
\partial_{y \leftarrow x}(I^A_{y \leftarrow x}(\omega)) = P^A_{x\leftarrow y}(\omega_y(v_{y \leftarrow x})).
$$

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

$$
\begin{align*}
\partial_{y \leftarrow x}I^A_{y \leftarrow x}(\omega) &= \frac{d}{dt}(I^A_{y + tv_{y \leftarrow x}}(\omega)) \bigg|_{t=0} \\
&= \frac{d}{dt} \left( \int_0^{T+t} P^A_{x\leftarrow \gamma(s)\omega_{\gamma(s)}(\dot{\gamma}(s))ds} \right) \bigg|_{t=0} \\
&= P^A_{x\leftarrow y}\omega_y(\dot{\gamma}(T))
\end{align*}
$$

since $\gamma(T) = y$. The result follows since $v_{y \leftarrow x} = \dot{\gamma}(T)$. \qed
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 $I_{y\leftarrow x}^A(\omega) + I_{y\leftarrow z}^A(\omega)$. However, this is not compatible with Lemma 2.4 as we would want

$$I_{z\leftarrow y\leftarrow x}^A(dAf) = P_{x\leftarrow y}^AP_{y\leftarrow z}^Af(z) - f(x)$$

(9)

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

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

(10)

One can check that (9) 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 $B$,

$$[P^A_\gamma]^{-1}P^B_\gamma - \text{Id} = I^{E(A,B)}_\gamma(A-B)$$

(11)

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 (11) 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 $\mathbb{C}^{n^2} \simeq \mathbb{C}^{n\times n}$ 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 $Q \in \mathbb{C}^{n\times n}$,

$$P^{E(A,B)}_\gamma Q = P^A_\gamma Q[P^B_\gamma]^{-1}.$$

(12)

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, $u(T)v(T) = (P^{E(A,B)}_\gamma Q)P^B_\gamma$. On the other hand,

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

Hence, $uv$ satisfies the parallel transport equation for $A$ along $\gamma$ with $u(0)v(0) = Q$, and so $u(T)v(T) = P^A_\gamma Q$. Combining the two expressions for $u(T)v(T)$ yields (12). □
Lemma 2.3 then yields

\[
\frac{d}{dt} A(t) + A(t) \frac{d}{dt} \gamma(t) u_A = 0,
\]

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 (11).

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

\[
[S^A_{z \leftarrow y \leftarrow x}]^{-1} S^B_{z \leftarrow y \leftarrow x} - \text{Id} = I^E_{z \leftarrow y \leftarrow x}(A - B).
\]

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

\[
I^E_{z \leftarrow y \leftarrow x}(A - B) = I^E_{y \leftarrow x}(A - B) + P^E_{x \leftarrow y} I^E_{z \leftarrow y}(A - B).
\]

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

\[
I^E_{z \leftarrow y \leftarrow x}(A - B) = P^A_{x \leftarrow y} P^B_{y \leftarrow x} - \text{Id} + P^A_{x \leftarrow y} (P^A_{y \leftarrow z} P^B_{z \leftarrow y} - \text{Id}) P^B_{y \leftarrow x}
\]

\[
= P^A_{x \leftarrow y} P^B_{y \leftarrow x} - \text{Id} + P^A_{x \leftarrow y} P^A_{y \leftarrow z} P^B_{z \leftarrow y} P^B_{y \leftarrow x} - P^A_{x \leftarrow y} P^B_{y \leftarrow x}
\]

\[
= [S^A_{z \leftarrow y \leftarrow x}]^{-1} S^B_{z \leftarrow y \leftarrow x} - \text{Id}
\]

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

\[
S^A_{z \leftarrow y \leftarrow x} [S^B_{z \leftarrow y \leftarrow x}]^{-1} - \text{Id} = I^E_{z \leftarrow y \leftarrow z}(A - B).
\]

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

\[
\partial_{x \leftarrow y} \left( S^A_{z \leftarrow y \leftarrow x} [S^B_{z \leftarrow y \leftarrow x}]^{-1} \right) = \partial_{x \leftarrow y} \left( I^E_{y \leftarrow z}(A - B) + P^E_{z \leftarrow y} I^E_{x \leftarrow y}(A - B) \right).
\]
The operator $\partial_{x \leftrightarrow y}$ 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

$$\partial_{x \leftrightarrow y} \left( S^A_{z \leftarrow y \leftarrow x} [S^B_{z \leftarrow y \leftarrow x}]^{-1} \right) = P_{z \leftarrow y}^{E(A,B)} \partial_{x \leftrightarrow y} \left( I_{z \leftarrow y}^{E(A,B)} (A - B) \right)$$

$$= P_{z \leftarrow y}^{E(A,B)} P_{y \leftarrow x}^{E(A,B)} (A - B)_{x}(v_{x \leftarrow y})$$

$$= S^A_{z \leftarrow y \leftarrow x} [(A - B)_{x}(v_{x \leftarrow y})] S^B_{z \leftarrow y \leftarrow x}.$$

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

$$\sup_{v \in T_{x \leftarrow y} \frac{D}{|v|=1}} |(A - B)_{x}(v_{x \leftarrow y})| = \sup_{y \in \frac{D}{(x,y) \in \mathbb{L}}} |(A - B)_{x}(v_{x \leftarrow y})| = \sup_{y \in \frac{D}{(x,y) \in \mathbb{L}}} |\partial_{x \leftrightarrow y} \left( S^A_{z \leftarrow y \leftarrow x} [S^B_{z \leftarrow y \leftarrow x}]^{-1} \right) |.$$

and the theorem follows. $\square$

To prove Theorem 1.3, it only remains to integrate over $\mathcal{U}^X$ to get a global estimate.

**Proof of Theorem 1.3.** By equivalence of norms, we can find $C > 0$ independent of both $\varepsilon$ and $x \in \mathcal{U}^X$ such that

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

After changing the integrand through equation (13) 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 (10) 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 $\omega$, we define the function

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

This function will serve as an approximate potential for $\omega$. 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 \( \gamma : [0, T] \to \mathbb{D} \) be the unit-speed lightlike geodesic from \( y \) to \( z_y \). 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, \( \omega(v_{y-z_y}) = d_A p(v_{y-z_y}) \).

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 \( \gamma \). By Lemma 2.3, it holds that \( u(t) = -P_{\gamma(t)\leftarrow z_y}^{A} I_{\gamma(t)\leftarrow z_y}^{A} = -p(\gamma(t)) \). 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 (8) that \( I_{z_y \leftarrow y}^{A}(\omega - d_A p) = 0 \) for all \( y \) and so, on the one hand,

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

On the other hand, by (9), we have

\[I_{z_y \leftarrow y}^{A}(\omega - d_A p) = I_{z_y \leftarrow y}^{A}(\omega) - P_{z_y \leftarrow y}^{A} P_{y \leftarrow z_y}^{A} p(z_y) + p(x) = I_{z_y \leftarrow y}^{A}(\omega) + p(x)\]

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

\[(15) \quad I_{z_y \leftarrow y}^{A}(\omega) = I_{y \leftarrow x}^{A}(\omega - d_A p) - p(x).\]

By applying \( \partial_{y \leftarrow x} \) to both sides of the last expression, Lemma 2.5 yields

\[(16) \quad \partial_{y \leftarrow x} \left( I_{z_y \leftarrow y}^{A}(\omega) \right) = P_{x \leftarrow y}^{A} ([\omega - d_A p]_y(v_{y \leftarrow x}))\]

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 \( \omega \) by \( A - B \) in (16) in order to use the pseudolinearisation 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 \( x \in (F_\varepsilon^X)_y \) form a basis of the tangent space at \( y \), but this basis degenerates when \( \varepsilon \) goes to 0. We therefore need estimates to quantify how well we can estimate \( \omega - d_A p \) at \( y \in \mathbb{D} \setminus \tilde{U} \) from moving \( x \) around in the intersection of \( \tilde{U} \) and the past lightcone of \( y \).

Lemma 2.10. Let \( 0 < \varepsilon < \varepsilon_0 \) and let \( y \in \mathbb{D} \setminus \tilde{U}_\varepsilon \). Then

\[\|(\omega - d_A p)_y\| \leq \frac{8}{\varepsilon} \sup_{x \in (F_\varepsilon^X)_y} |(\omega - d_A p)(v_{y \leftarrow x})|\]

The key to proving Lemma 2.10 is this small linear algebra lemma whose proof is straightforward.
Lemma 2.11. Let $b_1, \ldots, b_m$ be a basis of $\mathbb{R}^m$ with $|b_i|_{\mathbb{R}^m} = 1$ and let $T : \mathbb{R}^m \to \mathbb{C}^n$ be a linear map. Then

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

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 $v_{y-x}$ 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 $\gamma_x$ from $x$ to $y$ for different $x \in \overline{U}_{\varepsilon}$ 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 $\{v_{y-x}\}_{x \in \overline{U}_{\varepsilon}}$ and $\{v'_{y-x}\}_{x \in \overline{U}_{\varepsilon}}$ 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 $\{v_{y_{r_1-x}}\}_{x \in \overline{U}_{\varepsilon}} \subset \{v_{y_{r_2-x}}\}_{x \in \overline{U}_{\varepsilon}}$ and so any stability estimate for $y_{r_2}$ is also valid for $y_{r_1}$ since we're 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

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

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 $v_{y-z_4}$, while the other vectors $b_i$ correspond to $v_{y-x_i}$ 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 quick computation 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,

$$
\|((\omega - dAP)y)\| = \sup_{v \in T_y \mathbb{R}^{1+3}} \frac{|(\omega - dAP)(v)|}{|v|} \leq 2\|B^{-1}\| \max_{i=1, 2, 3, 4} |(\omega - dAP)(b_i)|
$$

$$
\leq \frac{8}{\varepsilon} \sup_{x \in \overline{U}_{\varepsilon}} |(\omega - dAP)(v_{y-x})|.
$$

The last inequality follows from the fact that $(\omega - dAP)(b_4) = 0$ by Lemma 2.9 and since $\{b_1, b_2, b_3\}$ is in the closure of $\{v_{y-x}\}_{x \in \overline{U}_{\varepsilon}}$. \qed
2.6. Proof of Theorems 1.4 and 2.2. We finally have everything to prove Theorem 2.2. The main idea is to use the pseudolinearisation identity to relate the scattering data with an attenuated X-ray transform of \( A - B \), and then use the operator \( \partial_{y^{\nu-x}} \) to evaluate \( A - B - d_{E(A,B)}p \) from \( T_{y}^{1}\mathbb{R}^{1+3} \). Theorem 1.4 then immediately follows by integrating over \( \mathbb{D} \setminus \mathbb{U} \).

**Proof of Theorem 2.2.** By Lemma 2.8, we have

\[
[S^{A}_{z_{y^{\nu-x}}} - 1]S^{B}_{z_{y^{\nu-x}}} - \text{Id} = I_{z_{y^{\nu-x}}}^{E(A,B)}(A - B).
\]

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 \( U(n^{2}) \). Hence, if we let

\[
p = p_{A-B}^{E(A,B)} = p_{y^{\nu-x}}^{E(A,B)}I_{y^{\nu-x}}^{E(A,B)}(A - B)
\]

then (15) yields

\[
I_{z_{y^{\nu-x}}}^{E(A,B)}(A - B) = I_{y^{\nu-x}}^{E(A,B)}(A - B - d_{E(A,B)}p) - p(x).
\]

Since \( \partial_{y^{\nu-x}} \text{Id} = \partial_{y^{\nu-x}} p(x) = 0 \), it now follows from Lemma 2.5 that

\[
\partial_{y^{\nu-x}} \left( [S^{A}_{z_{y^{\nu-x}}} - 1]S^{B}_{z_{y^{\nu-x}}} \right) = p_{y}^{E(A,B)}(A - B - d_{E(A,B)}p)(v_{y^{\nu-x}})
\]

\[
= p_{y}^{A}(A - B - d_{E(A,B)}p)(v_{y^{\nu-x}}) P_{x^{\nu-y}}^{B}.
\]

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

\[
|\partial_{y^{\nu-x}} \left( [S^{A}_{z_{y^{\nu-x}}} - 1]S^{B}_{z_{y^{\nu-x}}} \right)| = |(A - B - d_{E(A,B)}p)(v_{y^{\nu-x}})|.
\]

We can finally apply Lemma 2.10 to get

\[
\|(A - B - d_{E(A,B)}p)\|_{y} \leq \frac{8}{\varepsilon} \sup_{x \in (F_{\varepsilon}^{A})_{y}} |(A - B - d_{E(A,B)}p)(v_{y^{\nu-x}})|.
\]

Finally, note that

\[
\sup_{x \in (F_{\varepsilon}^{A})_{y}} |\omega(v_{y^{\nu-x}})| \sim \frac{1}{\text{vol}((F_{\varepsilon}^{A})_{y})} \int_{(F_{\varepsilon}^{A})_{y}} |\omega(v_{y^{\nu-x}})| \, dx
\]

for any one-form \( \omega \) as \( \varepsilon \) goes to 0. Combining this with the fact that \( \text{vol}((F_{\varepsilon}^{A})_{y}) \) is proportional to \( \varepsilon^{3} \), we can find a constant \( C > 0 \) independent of \( \varepsilon \) such that

\[
\|(A - B - d_{E(A,B)}p)\|_{y} \leq C \int_{(F_{\varepsilon}^{A})_{y}} |\partial_{y^{\nu-x}} \left( [S^{A}_{z_{y^{\nu-x}}} - 1]S^{B}_{z_{y^{\nu-x}}} \right)| \, dx.
\]

\[\square\]

**Remark 2.12.** Note that even though there is a supremum in the right-hand side of (7), one does not need to know \( \partial_{y^{\nu-x}} \left( [S^{A}_{z_{y^{\nu-x}}} - 1]S^{B}_{z_{y^{\nu-x}}} \right) \) for all \( x \in \mathbb{U}_{\varepsilon} \) in the past light cone of \( y \) to get an estimate. Indeed, the important equation is (17) 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 \( v_{y^{\nu-x}} \), since we already know it vanishes when evaluated at \( v_{y^{\nu-x}} \). 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

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

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

$$
\| \partial([S^A]^{-1}S^B) \|_{L^2(F^X)} = |(\partial_{y \rightarrow x} [S^A]^{-1})S^B + [S^A]^{-1}\partial_{y \rightarrow x}S^B|
$$

$$
= |[S^A]^{-1}\partial_{y \rightarrow x}S^B - [S^A]^{-1}(\partial_{y \rightarrow x}S^A)[S^A]^{-1}S^B|
$$

$$
= |\partial_{y \rightarrow x}(S^B - S^A) + (\partial_{y \rightarrow x}S^A)(\text{Id} - [S^A]^{-1}S^B)|
$$

$$
\leq |\partial_{y \rightarrow x}(S^B - S^A)| + |\partial_{y \rightarrow x}S^A||S^A - S^B|.
$$

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

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

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

$$
|\partial_{y \rightarrow x}S^A_{z_y \rightarrow y \rightarrow x}| = |(A \triangleleft P_{y \rightarrow z_y}^A)(v_{y \rightarrow x})|
$$

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

The last estimate is again invariant under $\mathcal{G}$ and involves the $L^\infty$-norm of the lightsink connection $A \triangleleft P_{y \rightarrow z_y}^A$. 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

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

where $F_A = dA + A \wedge A$ is the curvature 2-form of $A$ and

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

Corollary 1.7 will then directly follow from Theorem 1.4, Theorem 1.6, Lemma 2.13 and Lemma 2.14. However, we need another lemma before proving Lemma 2.14.
Lemma 2.15. Let \( \gamma : [0, T] \to \mathbb{D} \) be a smooth curve and let \( \gamma_s : [0, T] \to \mathbb{D} \) be a smooth variation of \( \gamma \) where \( s \in I = [\delta, \delta] \) for some \( \delta > 0 \). Then

\[
\frac{d}{ds} P^A_{\gamma_s} \bigg|_{s=0} = P^A_{\gamma} A_{\gamma(0)}(\partial_s \gamma_s(0)) - A_{\gamma(T)}(\partial_s \gamma_s(T)) P^A_{\gamma} + \int_0^T P^A_{\gamma[t, T]} F_A(\dot{\gamma}(t), \partial_s \gamma_s(t)) P^A_{\gamma[0, t]} dt
\]

where \( \partial_s \gamma_s(t) = \frac{d}{ds} \gamma_s(t) \bigg|_{s=0} \) and \( P^A_{\gamma[0, t]} \) is the parallel transport along the segment of \( \gamma \) restricted to the interval \([0, t]\).

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

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

By differentiating with respect to \( s \), we get

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

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

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

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

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

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

\[
u^r(s, t) = P^A_{\gamma_s[r, t]} \left(-\partial_s \left[A(\dot{\gamma}_s(r))\right]\right) P^A_{\gamma_s[0, r]}.
\]

Hence, we get

\[
v(0, T) = \int_0^T P^A_{\gamma[r, T]} \left(-\partial_s \left[A(\dot{\gamma}_s(r))\right]\right) \bigg|_{s=0} P^A_{\gamma[0, r]} dr.
\]

Expanding the curvature 2-form \( F_A = dA + A \wedge A \) yields

\[
F_A(\dot{\gamma}(r), \partial_s \gamma_s(r)) = \partial_r \left[A(\partial_s \gamma_s(r)) - \partial_s \left[A(\dot{\gamma}_s(r))\right]\right] - A(\dot{\gamma}(r), \partial_s \gamma_s(r)) + A(\dot{\gamma}(r), A(\partial_s \gamma_s(r)) - A(\partial_s \gamma_s(r)) A(\dot{\gamma}(r)).
\]
The vectors $\dot{y}(r)$ 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

$$ v(0, T) = \int_0^T P_{\gamma[0,T]}^A \left[ F_A(\dot{\gamma}(r), \partial_s \gamma_s(r)) - \partial_r [A(\partial_s \gamma_s(r))] - A(\dot{\gamma}(r))A(\partial_s \gamma_s(r)) + A(\partial_s \gamma_s(r))A(\dot{\gamma}(r)) \right] P_{\gamma[0,r]}^A \, dr. $$

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

$$ \int_0^T P_{\gamma[r,T]}^A A(\partial_s \gamma_s(r))A(\dot{\gamma}(r))P_{\gamma[0,r]}^A \, dr = \left[ -P_{\gamma[r,T]}^A A(\partial_s \gamma_s(r))P_{\gamma[0,r]}^A \right]_0^T + \int_0^T \partial_r [P_{\gamma[r,T]}^A A(\partial_s \gamma_s(r))] P_{\gamma[0,r]}^A \, dr. $$

The boundary term corresponds to the first two terms in (18) and we can expand the integrand of the second term to get

$$ \partial_r [P_{\gamma[r,T]}^A A(\partial_s \gamma_s(r))] P_{\gamma[0,r]}^A = P_{\gamma[r,T]}^A \left[ A(\dot{\gamma}(r))A(\partial_s \gamma_s(r)) + \partial_r [A(\partial_s \gamma_s(r))] \right] P_{\gamma[0,r]}^A $$

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 (19) to simplify $v(0, T)$ to (18).

**Proof of Lemma 2.14.** For $y \in D$ and unit $v \in T_y D$, we have

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

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

$$ dP_{y \rightarrow z_y}^A(v) = P_{y \rightarrow z_y}^A A_{z_y}(c\dot{t}) - A_y(v) P_{y \rightarrow z_y}^A $$

$$ + \int_0^T P_{y \rightarrow \gamma(t)}^A F_A(\dot{\gamma}(t), \partial_s \gamma_s(t)) P_{\gamma(t) \rightarrow z_y}^A \, dt. $$

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

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

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

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

□
2.8. Forward estimates. We finish the section by collecting forward estimates that will be useful for Section 4.

Lemma 2.16. Let $A$ and $B$ be Hermitian connections. Then
\[
\|S^A_{z \rightarrow y - x} - S^B_{z \rightarrow y - x}\|_{L^\infty(S^+(\mathcal{U}))} \leq 2\sqrt{2}\|A - B\|_{L^\infty(S\mathbb{D})}
\]
where $S\mathbb{D} = \{(x, v) \in T\mathbb{D} : |v|_e = 1\}$ is the (Euclidean) sphere bundle on $\mathbb{D}$.

Proof. Since $S^A_{z \rightarrow y - x}$ and the parallel transports lie in $U(n)$, Lemma 2.8 yields the pointwise estimate
\[
|S^A_{z \rightarrow y - x} - S^B_{z \rightarrow y - x}| = |I^{E(A, B)}_{y \rightarrow x}(A - B)|
\leq |I^{E(A, B)}_{y \rightarrow x}(A - B)| + |I^{E(A, B)}_{z \rightarrow y}(A - B)|
\leq \int_0^{T_1} |(A - B)(\gamma_{y \rightarrow x}(t))| \, dt + \int_0^{T_2} |(A - B)(\gamma_{z \rightarrow y}(t))| \, dt
\leq |x - y|_e \sup_{t \in [0, T_1]} |(A - B)(\gamma_{y \rightarrow x}(t))| +
\leq |y - z|_e \sup_{t \in [0, T_2]} |(A - B)(\gamma_{z \rightarrow y}(t))|
\]
where $\gamma_{y \rightarrow x} : [0, T_1] \rightarrow \mathbb{D}$ and $\gamma_{z \rightarrow y} : [0, T_2] \rightarrow \mathbb{D}$ 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 $(x, y, z) \in S^+(\mathcal{U})$ and using that the distance between $x$ and $y$ is at most $\sqrt{2}$, we get
\[
\|S^A_{z \rightarrow y - x} - S^B_{z \rightarrow y - x}\|_{L^\infty(S^+(\mathcal{U}))} \leq 2\sqrt{2} \sup_{x, z \in \mathbb{D}, y \in \mathbb{D}} \{(A - B)(v_{y \rightarrow x})|, |(A - B)(v_{y \rightarrow z})|\}
\leq 2\sqrt{2}\|A - B\|_{L^\infty(S\mathbb{D})}
\]
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 $\gamma_{y \rightarrow x}$ and $\gamma_{y \rightarrow z}$ cross $\mathcal{U}$. $\square$

Lemma 2.17. Let $A$ be a Hermitian connection and $\omega$ a one-form on $\mathbb{D}$. There is a constant $C > 0$ such that
\[
\|I^A_{y \rightarrow x}(\omega)\|_{H^k(\mathbb{D} \times \mathbb{D})} \leq C\|P^A_{y \rightarrow x}\|_{C^k(\mathbb{D} \times \mathbb{D})}\|\omega\|_{H^k(S\mathbb{D})}
\]
for all $k \geq 0$.

Proof. Let $I : C^\infty(\mathbb{D} \times S\mathbb{D}) \rightarrow C^\infty(\mathbb{D} \times \mathbb{D})$ be the usual ray transform on $\mathbb{D} \times S\mathbb{D}$ given by
\[
I(x, y)(\beta) = \int_0^{T(x, y)} \beta(x, \gamma_{y \rightarrow x}(t), \gamma_{y \rightarrow x}(t)) \, dt.
\]
for $\beta \in C^\infty(\mathbb{D} \times S\mathbb{D})$. It is well-known that $I$ is continuous from $H^k(\mathbb{D} \times S\mathbb{D})$ to $H^k(\mathbb{D} \times \mathbb{D})$ for all $k \geq 0$ [Sha94, Theorem 4.2.1]. Consider the natural projections in the following diagram.
\[
S\mathbb{D} \xleftarrow{\pi_1} \mathbb{D} \times S\mathbb{D} \xrightarrow{\pi_2} \mathbb{D} \times \mathbb{D}.
\]
These projections induce pullbacks on functions.

\[ C^\infty(S^\Delta) \xrightarrow{\pi_1^*} C^\infty(\Delta \times S^\Delta) \leftarrow C^\infty(\Delta \times \Delta). \]

We can view \( P^A_{x \leftrightarrow \gamma(t)} \) as a \( U(n) \) valued function on \( \Delta \times \Delta \) and \( \omega \) as a \( \mathbb{C}^{n \times n} \) valued function on \( S^\Delta \). Therefore, we can rewrite \( I^A_{x \leftrightarrow y}(\omega) \) as

\[ I^A_{y \leftrightarrow x}(\omega) = I(x, y) \left( (\pi_1^* P^A)(\pi_1^* \omega) \right). \]

The result follows by continuity of \( I \) and the pullbacks.

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

\[ \|S^A_{y \leftrightarrow x} - S^B_{y \leftrightarrow x}\|_{L^2(F^X)} \leq C\|A - B\|_{L^2(S^\Delta)}. \]

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

\[ \|S^A_{y \leftrightarrow x} - S^B_{y \leftrightarrow x}\|_{L^2(F^X)} \leq \|I^E_{y \leftrightarrow x}(A - B)\|_{L^2(F^X)} + \|I^E_{y \leftrightarrow x}(A - B)\|_{L^2(F^X)}. \]

The first term can be bounded by the same \( L^2 \)-norm on \( \Delta \times \Delta \) and is hence bounded by a multiple of \( \|A - B\|_{L^2(S^\Delta)} \) by Lemma 2.17 with \( k = 0 \). For the second term, we have

\[ \|I^E_{y \leftrightarrow x}(A - B)\|_{L^2(F^X)}^2 \leq C \int_{y \in \Delta \setminus \Delta} \int_0^T |(A - B)(\dot{\gamma}_{y \leftrightarrow x}(t))|^2 \, dt \, dy \]

where \( T \) is the length of the segment \( \gamma_{y \leftrightarrow x} \). We can rewrite the integral in the previous display as

\[ \int_{y \in \Delta} \int_{v \in S_y \Delta} |(A - B)_y(v)|^2 g(y, v) \, dv \, dy \]

for some bounded nonnegative function \( g : S^\Delta \to \mathbb{R} \). To see this, pick a point \( y^* \in \Delta \) and a direction \( v \in S_{y^* \Delta} \). Then, given \( y \in \Delta \setminus \Delta \), \( \dot{\gamma}_{y \leftrightarrow x}(t) = v \) for some unique \( t \) precisely when \( y^*, y \) and \( z_y \) lie on the line spanned by \( v \) based at \( y^* \) and \( y < y^* < z_y \).

Since \( y \) and \( y^* \) must lie on a bounded line and \( \Delta \) is bounded, it follows that \( g \) is also bounded. The result readily follows by bounding \( g \). \[ \square \]

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

\[ \|P^A_{y \leftrightarrow x}\|_{C^k(\Delta \times \Delta)} \leq c_k(1 + \|A\|_{C^k(S^\Delta)})^k. \]

**Proof.** We follow the inductive approach laid out in [Boh21]. Let \( \Gamma \) be the diagonal of \( \Delta \times \Delta \), that is, \( \Gamma = \{(x, x) : x \in \Delta\} \). Let \( \partial = \partial_{y \leftrightarrow x} \) be the vector field on \( \Delta \times \Delta \setminus \Gamma \) defined via (3). Then \( P^A_{y \leftrightarrow x} \) can be characterised as the unique smooth function \( U_A \) on \( \Delta \times \Delta \) such that

\[ (\partial + A)U_A := \partial_{y \leftrightarrow x}U_A(x, y) + A_y(v_{y \leftrightarrow x})U_A(x, y) = 0, \quad (x, y) \in \Delta \times \Delta \setminus \Gamma, \]

and \( U_A|_{\Gamma} = \text{Id} \).
Note that if \( G : \mathbb{D} \times \mathbb{D} \to \mathbb{C}^{n \times n} \) solves
\[
(\partial + A)G = -F, \quad (x, y) \in \mathbb{D} \times \mathbb{D} \setminus \Gamma
\]
for some \( F : \mathbb{D} \times \mathbb{D} \to \mathbb{C}^{n \times n} \) with \( G|\Gamma = G_0 \), then it follows from Lemma 2.3 that
\[
G = P^A_{y \leftarrow x} \left( G_0(x) - \int_0^T P^A_{x \leftarrow \gamma(t)} F(x, \gamma(t)) \, dt \right)
\]
where \( \gamma : [0, T] \to \mathbb{D} \) is the line segment from \( x \) to \( y \) parametrised by arc length. Hence, it holds that
\[
\|G\|_{L^\infty(\mathbb{D} \times \mathbb{D})} \leq C \left( \|G_0\|_{L^\infty(\Gamma)} + \|F\|_{L^\infty(\mathbb{D} \times \mathbb{D} \setminus \Gamma)} \right).
\]

By continuity, since \( P^A_{y \leftarrow x} \) is smooth, its \( C^k(\mathbb{D} \times \mathbb{D}) \)-norm agrees with its \( C^k(\mathbb{D} \times \mathbb{D} \setminus \Gamma) \)-norm. Let \( \{\partial, L_1, \ldots, L_7\} \) be a global commuting frame on \( \mathbb{D} \times \mathbb{D} \setminus \Gamma \) and denote \( L^\alpha = L_{1}^\alpha \cdots L_{7}^\alpha \) for \( \alpha \in \mathbb{Z}^7 \). Such a frame exists since we can choose global coordinates on \( \mathbb{D} \times \mathbb{D} \setminus \Gamma \) by first prescribing the usual coordinates for \( x \) in \( \mathbb{R}^4 \) and then choosing polar coordinates based at \( x \) to describe \( y \). The vector \( \partial \) then corresponds to the coordinate vector field related to the radial coordinate of \( y \) with respect to \( x \). We claim that
\[
\|P^A_{y \leftarrow x}\|_k := \sup_{j+|\alpha|=k} \|\partial^j L^\alpha P^A_{y \leftarrow x}\|_{L^\infty(\mathbb{D} \times \mathbb{D} \setminus \Gamma)} \lesssim_k \|A\|_{C^k(\mathbb{S}^D)}^k
\]
for all \( k \geq 0 \). For \( k = 0 \), this holds trivially as \( P^A_{y \leftarrow x} \) takes values in \( U(n) \). Suppose now that (21) holds for some \( k - 1 \geq 0 \). Take \( j \) and \( \alpha \) such that \( j + |\alpha| = k \). Then, if we let \( G := \partial^j L^\alpha P^A_{y \leftarrow x} \), we see that \( G \) solves
\[
(\partial + A)G = [A, \partial^j L^\alpha] P^A_{y \leftarrow x}, \quad (x, y) \in \mathbb{D} \times \mathbb{D} \setminus \Gamma,
\]
with \( G|\Gamma = \partial^j L^\alpha P^A_{y \leftarrow x}|_{y=x} \). Any differential operator \( V \) on \( \mathbb{D} \times \mathbb{D} \) can be decomposed as \( V = V_1 + V_2 \) where \( V_1 \) acts on the first coordinate and \( V_2 \) on the second with corresponding vectors \( v_1 \) and \( v_2 \) in \( T\mathbb{D} \) based at \( x \) and \( y \) respectively. Lemma 2.15 gives
\[
V P^A_{y \leftarrow x}|_{y=x} = (P^A_{y \leftarrow x} A(x_1) - A(y_2) P^A_{y \leftarrow x})|_{y=x} = A(x_1 - v_1 - v_2)
\]
and so we see that \( \|G|\Gamma\|_{L^\infty(\Gamma)} \leq \|A\|_{C^{k-1}(\mathbb{S}^D)} \). Equation (20) therefore gives
\[
\|\partial^j L^\alpha P^A_{y \leftarrow x}\|_{L^\infty(\mathbb{D} \times \mathbb{D} \setminus \Gamma)} \leq C \left( \|A\|_{C^{k-1}(\mathbb{S}^D)} + \|[A, \partial^j L^\alpha] P^A_{y \leftarrow x}\|_{L^\infty(\mathbb{D} \times \mathbb{D} \setminus \Gamma)} \right).
\]
The bracket \([A, \partial^j L^\alpha]\) is a differential operator of order \( k - 1 \) whose coefficients are derivatives of \( A \) and can be bounded by \( \|A\|_{C^k(\mathbb{S}^D)} \). Therefore, by the induction hypothesis, we have
\[
\|[A, \partial^j L^\alpha] P^A_{y \leftarrow x}\|_{L^\infty(\mathbb{D} \times \mathbb{D} \setminus \Gamma)} \lesssim_k \|A\|_{C^k(\mathbb{S}^D)} \|P^A_{y \leftarrow x}\|_{L^\infty(\mathbb{D} \times \mathbb{D} \setminus \Gamma)} \lesssim_k \|A\|_{C^k(\mathbb{S}^D)}^k\]
Absorbing the term for \( G|\Gamma \) in (22) and taking the supremum over all \( j \) and \( \alpha \) such that \( j + |\alpha| = k \), we see that (21) holds. Hence,
\[
\|P^A_{y \leftarrow x}\|_{C^k(\mathbb{D} \times \mathbb{D})} \lesssim \sum_{j=0}^k \|P^A_{y \leftarrow x}\|_j \lesssim \sum_{j=0}^k \|A\|_{C^k(\mathbb{S}^D)}^k \lesssim (1 + \|A\|_{C^k(\mathbb{S}^D)})^k.
\]
\[\square\]
3. Gauge invariance

3.1. Gauge invariance. We now study the quantity \( \| A - B - d_{E(A,B)}p \| \) in (7) in order to prove Theorem 1.6. Recall the definition of \( p^4_u \) as in (14). The \( p \) used in Theorem 2.2 is actually \( p^{E(A,B)}_A \), 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 \( \varphi \in C^\infty(D, U(n)) \) is in \( \mathcal{H} \) if \( \varphi|_0 = \text{Id} \). 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

(i) \( \Delta(A,B)^* = \Delta(B,A) \),

(ii) \( \Delta(A \triangleleft \varphi, B) = \varphi^{-1} \Delta(A,B) \),

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

**Proof.** We start by proving (i). 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

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

It then follows from Lemma 2.3 that

\[ p^*_u(A,B) = -u(T)^* = p_u(B,A). \]

Therefore, we can compute that

\[
(d_{E(A,B)}p(A,B))^* = (dp_{l}(A,B) + Ap_{l}(A,B) - p_{l}(A,B)B)^* \\
= dp^*_l(A,B) + p^*_l(A,B)A^* - B^* p^*_l(A,B) \\
= dp_l(B,A) + Bp_l(B,A) - p_l(B,A)A \\
= d_{E(B,A)}p(B,A)
\]

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 (iii). Let \( u \) be as in the proof of (i) above and consider \( \varphi^{-1}u \). We first claim that \( \varphi^{-1}u \) solves
\[
\begin{aligned}
\begin{cases}
(\varphi^{-1}u) + E(A \varphi, B)(\dot{\gamma}(t))\varphi^{-1}u = -\varphi^{-1}(A - B), \\
|\varphi^{-1}u|'(0) = 0.
\end{cases}
\end{aligned}
\]
Indeed,
\[
(\varphi^{-1}u) + E(A \varphi, B)\varphi^{-1}u = (\varphi^{-1})u + \varphi^{-1}\dot{u} + (\varphi^{-1})\varphi^{-1}u + (\varphi^{-1}A\varphi)\varphi^{-1}u - \varphi^{-1}uB
= \varphi^{-1}u + \varphi^{-1}Au - \varphi^{-1}uB
= -\varphi^{-1}(A - B + E(A, B)u) + \varphi^{-1}E(A, B)u
= -\varphi^{-1}(A - B).
\]

Now notice that
\[
d_{E(A\varphi, B)}\varphi^{-1} = d(\varphi^{-1}) + \varphi^{-1}(d\varphi)\varphi^{-1} + \varphi^{-1}A\varphi^{-1} - \varphi^{-1}B = \varphi^{-1}(A - B)
\]
since \( d(\varphi^{-1}) = -\varphi^{-1}(d\varphi)\varphi^{-1} \). Therefore, we can rewrite
\[
\varphi^{-1}(A - B) = d_{E(A\varphi, B)}(\varphi^{-1} - \text{Id} + \text{Id}) = A \varphi - B + d_{E(A\varphi, B)}(\varphi^{-1} - \text{Id}).
\]
To make notation less cumbersome, we write \( p \) for \( p_{(A, B)} \) and \( q \) for \( p_{(A\varphi, B)} \). We can now apply Lemma 2.3 again to get
\[
\varphi^{-1}p = -\varphi^{-1}u(T)
= P_{y^t z_y}^{E(A\varphi, B)} I_{y^t z_y}^{E(A\varphi, B)}(A \varphi - B + d_{E(A\varphi, B)}(\varphi^{-1} - \text{Id}))
= q + P_{y^t z_y}^{E(A\varphi, B)} I_{y^t z_y}^{E(A\varphi, B)}(d_{E(A\varphi, B)}(\varphi^{-1} - \text{Id})).
\]
Notice that \( \varphi^{-1} - \text{Id} \) vanishes at \( z_y \) since \( \varphi|_\circ = \text{Id} \), and so Lemma 2.4 yields
\[
P_{y^t z_y}^{E(A\varphi, B)} I_{y^t z_y}^{E(A\varphi, B)}(d_{E(A\varphi, B)}(\varphi^{-1} - \text{Id})) = \varphi^{-1}(y) - \text{Id},
\]
that is, \( q = \varphi^{-1}p - \varphi^{-1} + \text{Id} \). We can now compute
\[
d_{E(A\varphi, B)}q = dq + (A \varphi)q - qB
= d(\varphi^{-1}p) + (A \varphi)\varphi^{-1}p - \varphi^{-1}pB - d\varphi^{-1} - (A \varphi)\varphi^{-1} + \varphi^{-1}B
\]
\[
\quad + A \varphi - B
= (d\varphi^{-1})p + \varphi^{-1}dp + \varphi^{-1}(d\varphi)\varphi^{-1}p + \varphi^{-1}Ap - \varphi^{-1}pB - d\varphi^{-1}
\quad - \varphi^{-1}(d\varphi)\varphi^{-1} - \varphi^{-1}A + \varphi^{-1}B + A \varphi - B
= \varphi^{-1}dp + \varphi^{-1}Ap - \varphi^{-1}pB - \varphi^{-1}A + \varphi^{-1}B + A \varphi - B
= \varphi^{-1}d_{E(A, B)}p - \varphi^{-1}(A - B) + A \varphi - B
\]
and therefore
\[
\Delta(A \varphi, B) = A \varphi - B - d_{E(A\varphi, B)}q = \varphi^{-1}(A - B - d_{E(A, B)}p) = \varphi^{-1}\Delta(A, B)
\]
as claimed.
Finally, to prove (iii), it suffices to combine (i) and (ii) to get
\[
\Delta(A \varphi, B \psi) = \varphi^{-1}\Delta(B \psi, A)^* = \varphi^{-1}(\psi^{-1}\Delta(B, A))^* = \varphi^{-1}\Delta(A, B)\psi
\]
where we used that $\psi$ takes values in $U(n)$ and therefore $\psi^{-1} = \psi^*$. \hfill \Box

### 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

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

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

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

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

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

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

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

$$= d_{E(A,B)} P^{E(A,B)}_{y \leftarrow z_y} \text{Id}.$$ 

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

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

We have $P^A_{y \leftarrow z_y} = \varphi^{-1}(y) P^A_{y \leftarrow z_y} \varphi(z_y)$ from Proposition 1.1 and since $\varphi|_\varphi = \text{Id}$, $\varphi(z_y) = \text{Id}$. We chose $\varphi = P^A_{y \leftarrow z_y}$, and so $P^A_{y \leftarrow z_y} = \text{Id}$. Similarly, we have $P^{B\psi}_{z_y \leftarrow y} = \text{Id}$. Therefore, plugging $A \varphi$ for $A$ and $B \psi$ for $B$ in (23) gives

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

The result readily follows. \hfill \Box

### 3.3. Light-sink connections.

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

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

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

for some matrix fields $A_0, A_\theta, A_\phi$ with values in $u(n)$. 


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

$$A(\gamma(t)) = (A \circ P^A_{y \leftarrow z_y})(\gamma(t))$$

$$= P^A_{z_y \leftarrow y}[dP^A_{y \leftarrow z_y}](\gamma(t)) + P^A_{z_y \leftarrow y}A(\dot{\gamma}(t))P^A_{z_y \leftarrow y}$$

$$= 0$$

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

$$(24) \quad A_y(\partial_t) = A_y(\partial_r)$$

since $v_{y \leftarrow z_y} = \frac{1}{\sqrt{2}}(\partial_r - \partial_t)$. On the other hand, if (24) holds for all $y \in \mathbb{D}$, then $P^A_{z_y \leftarrow y} = \text{Id}$, and so $A \circ P^A_{z_y \leftarrow y} = A$. □

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 $B = A \circ \varphi$ is a light-sink connection for some $\varphi \in \mathcal{H}$. Then the previous proof guarantees that

$$P^{A \circ \varphi}_{y \leftarrow z_y} = \text{Id}$$

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

$$(A \circ \varphi) \circ P^{A \circ \varphi}_{y \leftarrow z_y} = A \circ P^A_{y \leftarrow z_y}$$

for all $\varphi \in \mathcal{H}$ and so the map $[A] \mapsto A \circ P^A_{y \leftarrow z_y}$ is well-defined. □

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 $A = B \circ P^B_{y \leftarrow z_y}$, we have

$$S^A_{z_y \leftarrow y \leftarrow x} = S^{B \circ P^B}_{z_y \leftarrow y \leftarrow x} = S^B_{z_y \leftarrow y \leftarrow x} P^B_{x \leftarrow z_x}$$

for all $(x,y) \in \mathcal{F}^X(\mathcal{U})$, and so, we can determine $P^B_{y \leftarrow z_y}$ inside $\mathcal{U}^X$ from the past-determined scattering data $S^A_{z_y \leftarrow y \leftarrow x}$ and $S^B_{z_y \leftarrow y \leftarrow x}$. The same applies on $\mathcal{U}^Z$ with the future-determined scattering data $S^A_{z_y \leftarrow y \leftarrow x}$ and $S^B_{z_y \leftarrow y \leftarrow x}$. Hence, we can recover $P^B_{y \leftarrow z_y}$ on the whole of $\mathcal{U}$. Let $\Phi \in C^\infty(\mathbb{D}, U(n))$ be any smooth extension of $P^B_{z_y \leftarrow y}$ to the whole of $\mathbb{D}$. Then, by (6),

$$A \circ \Phi = \Phi \circ A \circ P^B_{y \leftarrow z_y} \circ \Phi = B \circ [P^B_{y \leftarrow z_y} \Phi]$$

and so $A \circ \Phi$ is gauge equivalent to $B$ since $P^B_{y \leftarrow z_y} \Phi|_{\mathcal{U}} = \text{Id}$. □
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 $\mathcal{O} \setminus (\mathcal{U}^X \cup \mathcal{O})$. 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. Setting. We consider the following experimental setup similar to the one we described in Section 1.4. Let $\lambda$ be the uniform distribution on $\mathbb{S}^+(\mathcal{O})$ 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 $\mathbb{S}^+(\mathcal{O})$. For a light-sink Hermitian connection $A \in \Omega^1(\mathbb{D}, \mathfrak{so}(n))$, we denote its future-determined scattering data $S^A_{x \leftarrow y \leftarrow x}$ by $S^A_+(x, y)$ and its past-determined scattering data $S^A_{y \leftarrow x \leftarrow y}$ 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

$$S_i = (S^A_+(X_i, Y_i) + \mathcal{E}^+_i, S^A_-(Y_i, Z_i) + \mathcal{E}^-_i) \quad i = 1, \ldots, N.$$  \hfill (25)

The matrices $\mathcal{E}^\pm_i$ correspond to independent Gaussian noise in the sense that $\mathcal{E}^\pm_i = (\epsilon^\pm_{i,j,k})_{1 \leq j,k \leq n}$ and all the $\epsilon^\pm_{i,j,k}$’s are i.i.d. $N(0, 1)$ that are independent from the other random variables. We denote by $P^N_A$ 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 $\Pi$ on the space of $\mathfrak{so}(n)$-valued light-sink connections. Any such connection can be represented by three skew-symmetric matrix fields since

$$A = A_0(dt + dr) + A_d d\theta + A_{\phi} d\phi,$$

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 $\Pi$ 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 $L^\infty$ 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

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

\[ H^s(\mathbb{D}, \mathfrak{so}(n)^3) \simeq H^s(\mathbb{D}, \mathbb{R}^{d_n}) = \bigotimes_{i=1}^{d_n} H^s(\mathbb{D}, \mathbb{R}). \]

which plays the role of our parameter space \( \Theta \). The eigenfunctions \((e_i)\) 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

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

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 \(E_D\)

\[ A = N^{-1/(\alpha+2)} \sum_{i \leq d_n} \sum_{j \leq D/d_n} \lambda_i^{-\alpha/2} g_{j,i} e_{j,i}, \quad g_{j,i} \sim i.i.d. \mathcal{N}(0,1). \tag{26} \]

Theorem 4.2 can also be proved for \(D \to \infty\), giving rise to a commonly used Matérn prior of order \(\alpha\) 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 \(\Pi\) and its density by \(\pi\). 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

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

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

### 4.2. Statistical guarantees for light-sink connections.

To apply [BN21, Theorem 5.1], it remains to show that the map \(S : A \mapsto (S_1^A, S_2^A)\) satisfies their Condition 3.2 that contains three parts. The first part, uniform boundedness, is immediately satisfied since \(S^A\) takes values in \(U(n)\). The second part consists of global Lipschitz estimates for the forward map and follows from Lemma 2.16 in the \(L^\infty\) case and from Lemma 2.18 in the \(L^2\) case. Hence, it only remains to show that the last part of their Condition holds, which they have called inverse continuity modulus. That is the content of the next lemma.

**Lemma 4.1.** For every \(M\) there exists a constant \(L'\) and \(0 < \gamma \leq 1\) such that for all \(\delta > 0\) small enough and the given \(\alpha > 0\),

\[ \sup \{ \|A - B\|_{L^2(\mathbb{D})} : \|A\|_{H^\alpha(\mathbb{D})} + \|B\|_{H^\alpha(\mathbb{D})} \leq M, \|S(A) - S(B)\|_{L^2(\mathbb{S}^+(\mathbb{D}))} \leq \delta \} \leq L' \delta^\gamma. \]
Proof. First, note that we can find $C_\varepsilon > 0$ that depends on $\varepsilon$ such that
\[
\|S(A) - S(B)\|_{L^2(S^+(\mathbb{H}))}^2 \geq C_\varepsilon \left( \|S_A^\star - S_B^\star\|_{L^2(\mathbb{H})}^2 + \|S_A - S_B\|_{L^2(\mathbb{H})}^2 \right).
\]
By the interpolation inequality for Sobolev spaces (see [Boh21, Lemma 7.4]), for every $s > 0$,
\[
\| \cdot \|_{H^s(\mathbb{H})} \lesssim_s \| \cdot \|_{L^2(\mathbb{H})}^{1-1/s} \| \cdot \|_{H^s(\mathbb{H})}^{1/s}.
\]
By applying this inequality to $[S_A^\star]^{-1}S_B^\star - \text{Id}$ and using the pseudolinarisation identity, we get
\[
\| [S_A^\star]^{-1}S_B^\star - \text{Id} \|_{H^s(\mathbb{H})} \lesssim \|S_A^\star - S_B^\star\|_{L^2(\mathbb{H})}^{1-1/s} \|I^{E(A,B)}_{\varepsilon_1}\|_{H^{1/2}(\mathbb{H})}^{1/s}.
\]
Since $A$ and $B$ are light-sink connections, the broken attenuated X-ray is equal to the simple attenuated X-ray $I^{E(A,B)}_{\varepsilon_1}(A - B)$. For $s = k$ an integer, Lemmas 2.17 and 2.19 yield
\[
\|I^{E(A,B)}_{\varepsilon_1}(A - B)\|_{H^k(\mathbb{H})} \lesssim_k \|A - B\|_{H^k(\mathbb{H})}(1 + \|E(A,B)\|_{C^k(\mathbb{H})})^k.
\]
Taking $\alpha > k$ sufficiently large such that $\|A\|_{H^\alpha(\mathbb{H})} + \|B\|_{H^\alpha(\mathbb{H})} \leq M$, 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
\[
\| [S_A^\star]^{-1}S_B^\star - \text{Id} \|_{H^s(\mathbb{H})} \lesssim_{k,M} \|S_A^\star - S_B^\star\|_{L^2(\mathbb{H})}^{1-1/k}.
\]
It follows from Theorem 1.4 for light-sink connections ($p = 0$ by Theorem 1.6) that
\[
\|A - B\|_{L^2(\mathbb{H})} \lesssim_{k,M} \|S_A - S_B\|_{L^2(\mathbb{H})}^{1-1/k}.
\]
Similar inequalities also hold on $\Omega^X$ and $\Omega^Z$ by Theorem 1.3. Hence, we can choose $\gamma = \frac{\alpha - 4}{\alpha - 3}$ for $\alpha > 5$ by taking $\alpha = k + 3$. \qed

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 $A^\star$ obtained through noisy samples of $S_A^\star$ 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 (26) with $\alpha \geq 5$ and data $(S_i, (X_i, Y_i, Z_i)) \sim P^N_A$ as in (25). Suppose that $A^\star \in H^\alpha$ and $D \simeq N^{2/(\alpha + 2)}$. Let $\gamma = \frac{\alpha - 4}{\alpha - 3}$. Then, there is $M > 0$ such that
\[
\Pi\left( \|A - A^\star\|_{L^2(\mathbb{H})} > M\delta \|S_i, (X_i, Y_i, Z_i)\|_{i=1}^N \right) = o_{P^N_A}(1)
\]
as $N \to \infty$, where $\delta = N^{-\alpha/(2\alpha + 4)}$.

In short, the posterior distribution converges to a delta distribution around $A^\star$ in $P^N_A$-probability at a rate that depends on the smoothness of the prior and of $A^\star$. The smoother $A^\star$ 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,
\[
\|E\Pi[A | (S_i, (X_i, Y_i, Z_i))_{i=1}^N] - A^\star\|_{L^2(\mathbb{H})} = O_{P^N_A}(\delta_N^{\gamma})
\]
as $N \to \infty$.

Note that we have convergence to $A_\star$ in Theorem 4.2 and not only its projection $A_{\star,D}$ 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 $H^\alpha(D, \mathfrak{so}(n)^3)$, and not just $E_D$. Indeed, Remark 5.2 in [BN21] guarantees that we can then replace $A_{\star,D}$ by $A_\star$.

REFERENCES

[ALJ19] G. Ambartsoumian and M. J. Latifi Jebelli. The V-line transform with some generalizations and cone differentiation. Inverse Problems, 35(3):034003, 29, 2019.

[Amb12] Gaik Ambartsoumian. Inversion of the V-line Radon transform in a disc and its applications in imaging. Comput. Math. Appl., 64(3):260–265, 2012.

[AN19] Kweku Abraham and Richard Nickl. On statistical Calderón problems. Math. Stat. Learn., 2(2):165–216, 2019.

[AS09] Simon R. Arridge and John C. Schotland. Optical tomography: forward and inverse problems. Inverse Problems, 25(12):123010, 59, 2009.

[BN21] Jan Bohr and Richard Nickl. On log-concave approximations of high-dimensional posterior measures and stability properties in non-linear inverse problems. arXiv preprint arXiv:2105.07835, 2021.

[Boh21] Jan Bohr. Stability of the non-abelian X-ray transform in dimension $\geq 3$. J. Geom. Anal., 31(11):11226–11269, 2021.

[CLOP21a] Xi Chen, Matti Lassas, Lauri Oksanen, and Gabriel P Paternain. Detection of Hermitian connections in wave equations with cubic non-linearity. J. Eur. Math. Soc., 2021. To appear.

[CLOP21b] Xi Chen, Matti Lassas, Lauri Oksanen, and Gabriel P. Paternain. Inverse problem for the Yang-Mills equations. Comm. Math. Phys., 384(2):1187–1225, 2021.

[FMS11] Lucia Florescu, Vadim A. Markel, and John C. Schotland. Inversion formulas for the broken-ray Radon transform. Inverse Problems, 27(2):025002, 13, 2011.

[GN16] Evarist Giné and Richard Nickl. Mathematical foundations of infinite-dimensional statistical models. Cambridge Series in Statistical and Probabilistic Mathematics, [40]. Cambridge University Press, New York, 2016.

[GN20] Matteo Giordano and Richard Nickl. Consistency of Bayesian inference with Gaussian process priors in an elliptic inverse problem. Inverse Problems, 36(8):085001, 35, 2020.

[GvdV17] Subhashis Ghosal and Aad van der Vaart. Fundamentals of nonparametric Bayesian inference, volume 44 of Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2017.

[Hör68] Lars Hörmander. The spectral function of an elliptic operator. Acta Math., 121:193–218, 1968.

[Hör09] Lars Hörmander. The analysis of linear partial differential operators. IV. Classics in Mathematics. Springer-Verlag, Berlin, 2009. Fourth integral operators, Reprint of the 1994 edition.

[KLU18] Yaroslav Kurylev, Matti Lassas, and Gunther Uhlmann. Inverse problems for Lorentzian manifolds and non-linear hyperbolic equations. Invent. Math., 212(3):781–857, 2018.

[MNP21] François Monard, Richard Nickl, and Gabriel P. Paternain. Consistent inversion of noisy non-Abelian X-ray transforms. Comm. Pure Appl. Math., 74(5):1045–1099, 2021.

[Nov19] Roman Novikov. Non-abelian x-ray transform and its applications. In The Radon Transform: The First 100 Years and Beyond, pages 115–128. De Gruyter, 2019.

[PS20] Gabriel P Paternain and Mikko Salo. The non-abelian x-ray transform on surfaces. arXiv preprint arXiv:2006.02257, 2020.
[PSU21] Gabriel P. Paternain, Mikko Salo, and Gunther Uhlmann. *Geometric inverse problems with emphasis on two dimensions*. 2021. Preprint available on https://www.dpmms.cam.ac.uk/~g7egpp24/GIP2D_driver.pdf.

[Sha94] V. A. Sharafutdinov. *Integral geometry of tensor fields*. Inverse and Ill-posed Problems Series. VSP, Utrecht, 1994.

[Sin78] I. M. Singer. Some remarks on the Gribov ambiguity. *Comm. Math. Phys.*, 60(1):7–12, 1978.

[Stu10] A. M. Stuart. Inverse problems: a Bayesian perspective. *Acta Numer.*, 19:451–559, 2010.

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

Email address: sas242@cam.ac.uk