

A APPENDIX / SUPPLEMENTARY: FULL PROOFS AND ALGORITHM

S.0 NOTATION AND PRELIMINARIES

For a fixed layer, let $z : \mathbb{R}^d \rightarrow \mathbb{R}^p$ be the representation map and $J_z(x) \in \mathbb{R}^{p \times d}$ its Jacobian. For a sample pool $\mathcal{N}$ we write $Z = \{z(x)\}_{x \in \mathcal{N}}$, empirical mean μ and covariance Σ . Global whitening is $\tilde{z}(x) = \Sigma^{-1/2}(z(x) - \mu)$ (any fixed symmetric $\Sigma^{-1/2}$ suffices). A loop $\gamma = (x_0, \dots, x_{L-1}, x_L = x_0)$ has edges $e_i = (x_i, x_{i+1})$. For an edge e_i we select a *shared* index set $\mathcal{I}_i \subset \{1, \dots, |\mathcal{N}|\}$ of size k by k -NN around the midpoint $\frac{1}{2}(\tilde{z}(x_i) + \tilde{z}(x_{i+1}))$ in whitened feature space. Let $\tilde{Z}_{\mathcal{I}_i} \in \mathbb{R}^{k \times p}$ be the whitened feature matrix restricted to those rows. Write $m_i := \frac{1}{2}(\tilde{z}(x_i) + \tilde{z}(x_{i+1}))$. Define weights on the same rows by $w_j^{(i)} \propto \exp(-\|\tilde{Z}_{\mathcal{I}_i} - m_i\|/\sigma_i)$ (normalized on $\mathcal{I}_i$ to sum to 1), and the shared midpoint center $\bar{\mu}_i := \sum_{j \in \mathcal{I}_i} w_j^{(i)} \tilde{Z}_{\mathcal{I}_i}$. Set the centered clouds $\tilde{Z}_i^{\text{src}} = \tilde{Z}_{\mathcal{I}_i} - \bar{\mu}_i$, $\tilde{Z}_i^{\text{tgt}} = \tilde{Z}_{\mathcal{I}_i} - \bar{\mu}_i$. Let $B_i \in \mathbb{R}^{p \times q}$ be the top- q right singular vectors of $\begin{bmatrix} \tilde{Z}_i^{\text{src}} \\ \tilde{Z}_i^{\text{tgt}} \end{bmatrix}$. In the q -subspace, the orthogonal Procrustes solution is $U_i \Sigma_i V_i^\top = \text{SVD}((X_i B_i)^\top W_i (Y_i B_i))$, $R_i^{(q)} = U_i V_i^\top$, where $W_i = \text{diag}(w_j^{(i)})_{j \in \mathcal{I}_i}$. $R_i^{(q)} = U_i V_i^\top \in \text{SO}(q)$ (enforce $\det = +1$ if necessary). We embed to $\mathbb{R}^p$ by

$$\hat{R}_i = B_i R_i^{(q)} B_i^\top + (I - B_i B_i^\top) \in \text{SO}(p).$$

The empirical holonomy is $\hat{H}(\gamma) = \hat{R}_{L-1} \cdots \hat{R}_0$ and $\hat{h}_{\text{norm}} = \|\hat{H} - I\|_F / (2\sqrt{p})$. We denote spectral and Frobenius norms by $\|\cdot\|_2$ and $\|\cdot\|_F$, and principal angle matrices by $\sin \Theta(\cdot, \cdot)$.

Matrix perturbation tools. We use (i) Davis–Kahan/Wedin: for symmetric A, E , if $A = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}$ in the eigenbasis and $\text{gap} = \min_{\lambda \in \sigma(A_{11}), \mu \in \sigma(A_{22})} |\lambda - \mu| > 0$, then $\|\sin \Theta(\hat{U}, U)\|_2 \leq \|E\|_2 / \text{gap}$ for the top- q subspace. For rectangular SVD subspaces, Wedin’s theorem yields the same bound for left/right singular subspaces. (ii) For $Q \in \text{O}(p)$, $\|Q - I\|_F^2 = 2 \sum_{j=1}^p (1 - \cos \theta_j) \leq 4p$.

S.1 FULL PROOFS OF INVARIANCES, NULLS, AND NORMALIZATION

Proposition 1 (Gauge invariance under orthogonal reparameterizations; full proof). *Let $U \in \text{O}(p)$ and $\tilde{z}'(x) = U \tilde{z}(x)$. The shared index sets $\mathcal{I}_i$ are unchanged (same midpoint up to left multiplication by U), and for every edge i , $\hat{R}'_i = U \hat{R}_i U^\top$. Hence $\hat{H}' = U \hat{H} U^\top$, $\|\hat{H}' - I\|_F = \|\hat{H} - I\|_F$, and the eigen-angle multisets coincide.*

Proof. For the same rows, $(\tilde{Z}_i^{\text{src}})' = U \tilde{Z}_i^{\text{src}}$ and similarly for $\tilde{Z}_i^{\text{tgt}}$ because soft centers transform as $\tilde{\mu}'_i = U \tilde{\mu}_i$. Let B_i be an orthonormal basis for the span of the stacked clouds; then $B'_i = U B_i$ is an orthonormal basis for the transformed span. The cross-covariance in the subspace transforms as

$$(\tilde{Z}_i^{\text{src}} B_i)^\top (\tilde{Z}_i^{\text{tgt}} B_i) \mapsto (B_i^\top U^\top \tilde{Z}_i^{\text{src}}) (U \tilde{Z}_i^{\text{tgt}} B_i) = B_i^\top \tilde{Z}_i^{\text{src}} \tilde{Z}_i^{\text{tgt}} B_i.$$

Hence $U_i \Sigma_i V_i^\top$ is unchanged; $R_i^{(q)}$ is identical. Embedding gives $\hat{R}'_i = B'_i R_i^{(q)} B_i^\top + (I - B'_i B_i^\top) = U(B_i R_i^{(q)} B_i^\top + I - B_i B_i^\top) U^\top = U \hat{R}_i U^\top$. Composition and Frobenius invariance under conjugation conclude the proof. $\square$

Proposition 2 (Affine invariance after global whitening; full proof). *Let raw features be $z'(x) = Az(x) + b$ with $A \in \text{GL}(p)$. Let Σ' and μ' be the pool covariance and mean of z' . Then there exists $Q \in \text{O}(p)$ such that $\tilde{z}'(x) = Q \tilde{z}(x)$ for all x . Consequently Proposition above applies.*

Proof. We have $\mu' = A\mu + b$ and $\Sigma' = A\Sigma A^\top$. Choose symmetric square roots. Then

$$\tilde{z}'(x) = \Sigma'^{-1/2}(Az(x) + b - \mu') = \Sigma'^{-1/2}A(z(x) - \mu).$$

Write the polar decomposition of $\Sigma'^{-1/2}A\Sigma^{1/2}$ as QP with $Q \in \text{O}(p)$, P symmetric positive definite. Then

$$\Sigma'^{-1/2}A = QP\Sigma^{-1/2} \Rightarrow \tilde{z}'(x) = Q(P\Sigma^{-1/2}(z(x) - \mu)).$$

But $P = I$ because

$$P^2 = \Sigma^{\frac{1}{2}} A^\top \Sigma'^{-1} A \Sigma^{\frac{1}{2}} = \Sigma^{\frac{1}{2}} A^\top (A^{-\top} \Sigma^{-1} A^{-1}) A \Sigma^{\frac{1}{2}} = \Sigma^{\frac{1}{2}} \Sigma^{-1} \Sigma^{\frac{1}{2}} = I.$$

Thus $\tilde{z}'(x) = Q\tilde{z}(x)$.

Equivalently, set $M := \Sigma'^{-1/2} A \Sigma^{1/2}$. Then $M^\top M = I$, so $M \in O(p)$ and $\tilde{z}'(x) = M\tilde{z}(x)$. $\square$

Proposition 3 (Linear null; full proof). *If $z(x) = Bx + c$ (affine) and the same index set $\mathcal{I}_i$ is used for both directions of each edge, then $\hat{R}_i = I$ and $\hat{H}(\gamma) = I$ for any loop γ .*

Proof. With shared neighbors and the shared midpoint soft center $\bar{\mu}_i$, for any affine $z(x) = Bx + c$ we have on rows I_i that $\tilde{Z}_i^{\text{src}} = \tilde{Z}_i^{\text{tgt}} = \tilde{Z}_{I_i} - \bar{\mu}_i$. Hence the $\text{SO}(q)$ Procrustes optimum is $R_i^{(q)} = I$ and the embedded map is $\hat{R}_i = I$, so $\hat{H}(\gamma) = I$ for any loop γ . $\square$

Proposition 4 (Orientation, reparametrization, and normalization; full proof). *Reversing edge order in γ inverts the orthogonal product so $\hat{H}(\gamma^{-1}) = \hat{H}(\gamma)^{-1}$ and $\|\hat{H}(\gamma^{-1}) - I\|_F = \|\hat{H}(\gamma) - I\|_F$. Cyclic reparameterizations do not change the product. Moreover, $\hat{h}_{\text{norm}} \in [0, 1]$ with equality 1 iff all eigen-angles are π . (The upper bound is attained by $H = -I$; within $\text{SO}(p)$ this is attainable only when p is even.)*

Proof. All $\hat{R}_i \in \text{SO}(p)$; the first two claims follow from group identities. For normalization, for $H \in O(p)$ with eigenvalues $e^{i\theta_j}$, $\|H - I\|_F^2 = \text{tr}((H - I)^\top (H - I)) = 2 \sum_j (1 - \cos \theta_j) \leq 4p$. $\square$

S.2 SMALL-RADIUS LIMIT; FULL PROOF

Assumption 1 (Regularity and neighbor stability). *(i) z is C^2 with L -Lipschitz Jacobian on a neighborhood of the loop. (ii) The loop γ_r lies on a C^2 2D manifold in input space with total length $O(r)$. (iii) Shared-midpoint k -NN selection has overlap probability $1 - O(r)$ as $r \rightarrow 0$. (iv) The subspace dimension q contains the rank of the local feature covariance of the shared rows.*

Theorem 1 (Small-radius limit). *As $r \rightarrow 0$, $\|\hat{R}_i - I\|_F = O(r)$ for each edge and $\|\hat{H}(\gamma_r) - I\|_F = O(r)$. Hence $\hat{h}_{\text{norm}}(\gamma_r) = O(r)$.*

Proof. Let $\delta_i = x_{i+1} - x_i$ with $\|\delta_i\| = O(r)$ and $x(t)$ be a C^2 parameterization. A second-order expansion gives $z(x + \delta) = z(x) + J_z(x)\delta + \frac{1}{2} \text{Hc}_z(x)[\delta, \delta] + O(\|\delta\|^3)$. Soft centering on the *same* rows cancels translations, leaving two clouds whose covariance difference is $O(\|J_z(x + \delta) - J_z(x)\|) = O(\|\delta\|)$ by Lipschitzness. Orthogonal Procrustes between two centered clouds that differ by an $O(\|\delta\|)$ linear term has solution $I + O(\|\delta\|)$ (Procrustes perturbation Lemma S.4 below). Therefore $\|\hat{R}_i - I\|_F = O(\|\delta_i\|) = O(r)$. Since $L = O(1)$ and products of $I + E_i$ with $E_i = O(r)$ deviate from I by $O(\sum_i \|E_i\|) = O(r)$, the holonomy claim follows. $\square$

S.3 PROCRUSTES PERTURBATION (FULL STATEMENT AND PROOF)

We quantify how the orthogonal Procrustes optimum $UV^\top$ changes under perturbations of the cross-covariance. This will be invoked per edge on the q -dimensional subspace.

Lemma 1 (Procrustes perturbation via singular subspaces). *Let $M \in \mathbb{R}^{q \times q}$ with SVD $U\Sigma V^\top$ and orthogonal Procrustes minimizer $R^* = UV^\top$. Let $\hat{M} = M + E$, with SVD $\hat{U}\hat{\Sigma}\hat{V}^\top$ and minimizer $\hat{R} = \hat{U}\hat{V}^\top$ (take the $\text{SO}(q)$ correction by flipping the last column of $\hat{U}$ if needed so $\det \hat{R} = +1$). If the smallest singular value gap of M satisfies $\text{gap} = \min\{\sigma_j - \sigma_{j+1} : 1 \leq j < q\} > 0$, then*

$$\|\hat{R} - R^*\|_F \leq 2(\|\sin \Theta(\hat{U}, U)\|_F + \|\sin \Theta(\hat{V}, V)\|_F) \leq \frac{4\|E\|_2}{\text{gap}} \sqrt{q}.$$

Moreover $\|\hat{R} - R^*\|_2 \leq 2\|E\|_2/\text{gap}$.

Proof. Write $\hat{R} - R^* = \hat{U}\hat{V}^\top - UV^\top = (\hat{U} - UQ_U)\hat{V}^\top + UQ_U(\hat{V} - VQ_V)^\top + U(Q_UQ_V^\top - I)V^\top$, for orthogonal Q_U, Q_V chosen to realize the principal-angle alignments between the subspaces spanned by columns of U and $\hat{U}$, and of V and $\hat{V}$ (CS decomposition). The third term is bounded by $\|Q_UQ_V^\top - I\|_F \leq \|Q_U - I\|_F + \|Q_V - I\|_F \leq 2(\|\sin \Theta(\hat{U}, U)\|_F + \|\sin \Theta(\hat{V}, V)\|_F)$. The first two terms are each bounded by the same sine-angle norms. Summing yields the first bound. For the second inequality, Wedin’s theorem gives $\|\sin \Theta(\hat{U}, U)\|_2 \leq \|E\|_2/\text{gap}$ and similarly for V ; Frobenius then adds a $\sqrt{q}$ factor, while the spectral bound is direct. $\square$

Remark. The $\text{SO}(q)$ correction (flip the last singular vector if $\det < 0$) changes R by at most 2 in Frobenius norm and is absorbed by the same bound when $\|E\|_2$ is small relative to the gap; empirically it eliminates spurious π flips.

S.4 SUBSPACE TRUNCATION AND DAVIS–KAHAN/WEDIN

We justify the q -dimensional embedding error.

Lemma 2 (Subspace truncation bound). *Let $S = \begin{pmatrix} \tilde{Z}_i^{\text{src}} \\ \tilde{Z}_i^{\text{tgt}} \end{pmatrix}$, $\Sigma_S = \frac{1}{k}S^\top S$, and let B be the top- q right singular vectors of S . Let $\Pi = BB^\top$ and $\Pi_\perp = I - \Pi$. Suppose the singular value gap $\Delta = \sigma_q(S) - \sigma_{q+1}(S) > 0$. Then for any two centered clouds X, Y formed from the same rows, the Procrustes minimizers satisfy*

$$\|(BR^{(q)}B^\top + \Pi_\perp) - R^*\|_F \leq C \frac{\|X^\top X - Y^\top Y\|_2}{\Delta} + \|\Pi_\perp\|_F,$$

where R^* is the (untruncated) Procrustes optimum on the full span and C depends on local condition numbers of $X^\top X, Y^\top Y$.

Proof. Decompose both clouds into in-span and out-of-span components; Wedin/Davis–Kahan ensures $\|\sin \Theta(\text{span}(B), \text{span}(S))\|_2 \leq \|E\|_2/\Delta$ for the empirical perturbation E of the covariance. The cross-covariance restricted to $\text{span}(B)$ deviates from the full one by $O(\|E\|_2/\Delta)$. Apply Lemma 1 inside the subspace and add the residual $\|\Pi_\perp\|_F$ from identity on the complement. $\square$

S.5 PER-EDGE AND HOLONOMY ERROR BOUNDS; FULL DERIVATION

We combine (i) finite-sample concentration, (ii) subspace truncation, (iii) index mismatch, and (iv) curvature terms.

Assumption 2 (Sampling and gaps). *Neighbors are i.i.d. from a distribution with covariance Σ_i whose top- q eigenspace is separated by a gap $\Delta_i > 0$. Empirical covariances concentrate: $\|\hat{\Sigma}_i - \Sigma_i\|_2 \leq c\sigma\sqrt{\frac{\log(1/\delta)}{k}}$ with prob. $\geq 1 - \delta$.*

Theorem 2 (Per-edge transport error). *With probability $\geq 1 - \delta$,*

$$\|\hat{R}_i - R_i^*\|_F \leq C_1 \frac{\sigma}{\sqrt{k}} + C_2 \frac{\|\Pi_\perp^i \Sigma_i^{1/2}\|_F}{\lambda_q(\Sigma_i)^{1/2}} + C_3 \text{TV}(\mathcal{I}_i, \mathcal{I}_i^*) + C_4 \|J_z(x_{i+1}) - J_z(x_i)\|_2,$$

where R_i^* is the population Procrustes minimizer on the true shared rows and full span, $\Pi_\perp^i$ projects onto the discarded right singular directions, and $\text{TV}(\cdot, \cdot)$ is the (normalized) total-variation distance between empirical and population index sets. Constants (C_j) depend only on local condition numbers and are independent of k and r .

Proof. (1) *Finite sample:* concentration of empirical cross-covariances (sub-Gaussian or bounded support) yields $\|E\|_2 \lesssim \sigma/\sqrt{k}$. Lemma 1 gives the first term.

(2) *Truncation:* Lemma 2 yields the second term with $\|\Pi_\perp^i \Sigma_i^{1/2}\|_F/\lambda_q(\Sigma_i)^{1/2}$ measuring residual energy outside the top- q .

(3) *Index mismatch:* if empirical indices differ from population $\mathcal{I}_i^*$ by fraction τ , then centered clouds differ by $O(\tau)$ in Frobenius norm; propagate through Procrustes continuity to obtain $C_3 \tau$. Set $\tau = \text{TV}(\mathcal{I}_i, \mathcal{I}_i^*)$.

(4) *Curvature:* with shared rows, the first-order difference in centered clouds is controlled by $\|J_z(x_{i+1}) - J_z(x_i)\|_2 = O(\|x_{i+1} - x_i\|) = O(r)$ by Lipschitzness; this yields the last term. $\square$

Corollary 1 (Holonomy error accumulation). *For $L = O(1)$ edges,*

$$\|\hat{H} - H^*\|_F \leq \sum_{i=0}^{L-1} \|\hat{R}_i - R_i^*\|_F + O(r),$$

where $H^* = R_{L-1}^* \cdots R_0^*$ is the population holonomy.

Proof. Write $\hat{H} - H^* = \sum_{i=0}^{L-1} (\hat{R}_{L-1} \cdots \hat{R}_{i+1})(\hat{R}_i - R_i^*)(R_{i-1}^* \cdots R_0^*)$ and use submultiplicativity with $\|\hat{R}_j\|_2 = \|R_j^*\|_2 = 1$. $\square$

S.6 ALGORITHM (MIRRORS THE IMPLEMENTATION)

We include a compact algorithm in `algorithmic` style. The steps and symbols match the code.

Algorithm 1 GAUGE-INVARIANT REPRESENTATION HOLONOMY (SO(p) subspace Procrustes)

Require: Model f , layer $z(\cdot)$, neighbor loader $\mathcal{L}$, loop points $\{x_i\}_{i=0}^{L-1}$, k neighbors, subspace q
Ensure: Holonomy matrix $\hat{H} \in \text{SO}(p)$, normalized score $\hat{h}_{\text{norm}}$, eigen-angles $\{\theta_j\}$
1: **Pool features:** collect $Z = \{z(x)\}_{x \in \mathcal{N}}$ from $\mathcal{L}$; compute mean μ and covariance Σ .
2: **Global whitening:** $\tilde{z}(x) = \Sigma^{-1/2}(z(x) - \mu)$; build whitened pool $\tilde{Z}$.
3: Initialize $\hat{H} \leftarrow I_p$.
4: **for** $i = 0, \dots, L-1$ **do** $\triangleright$ edges $x_i \rightarrow x_{i+1}$ with $x_L = x_0$
5: $\tilde{z}_i \leftarrow \tilde{z}(x_i)$, $\tilde{z}_{i+1} \leftarrow \tilde{z}(x_{i+1})$, midpoint $m_i = \frac{1}{2}(\tilde{z}_i + \tilde{z}_{i+1})$
6: **Shared neighbors:** $\mathcal{I}_i \leftarrow k$ -NN indices of m_i in $\tilde{Z}$; $S \leftarrow \tilde{Z}[\mathcal{I}_i, :]$
7: **Shared soft center:** compute $\bar{\mu}_i$ on rows $\mathcal{I}_i$ using weights at m_i ; set $X = Y = S - \bar{\mu}_i$.
8: **Shared subspace:** $B \in \mathbb{R}^{p \times q}$ = top- q right singular vectors of $\begin{bmatrix} X \\ Y \end{bmatrix}$
9: **Procrustes in $\mathbb{R}^q$:** $USV^\top = \text{SVD}((XB)^\top(YB))$; $R^{(q)} = UV^\top$; enforce $\det R^{(q)} = +1$
10: **Embed:** $\hat{R}_i \leftarrow BR^{(q)}B^\top + (I - BB^\top)$ $\triangleright \in \text{SO}(p)$
11: **Compose:** $\hat{H} \leftarrow \hat{R}_i \cdot \hat{H}$
12: **end for**
13: **Return:** $\hat{H}$, $\hat{h}_{\text{norm}} = \|\hat{H} - I\|_F / (2\sqrt{p})$, eigen-angles $\{\theta_j\}$ of $\hat{H}$

S.7 COMPLEXITY (EXPANDED)

Let k be neighbors, q the subspace, p the feature dimension, L edges, and N_{pool} pooled samples.

- One-time pool: $O(N_{\text{pool}} \cdot \text{forward}(z))$ to extract features and $O(N_{\text{pool}} p^2)$ to form Σ (streaming computation eliminates storing Z ; memory is $O(p^2)$ for Σ and $O(p)$ for μ).
- Per edge: thin SVD of a $(2k) \times p$ matrix to get B : $O(kpq)$; Procrustes SVD in $\mathbb{R}^q$: $O(q^3)$; embedding $O(pq)$.
- Total loop: $O(L(kpq + q^3 + pq))$, typically dominated by kpq with $q \ll p$.

S.8 PRACTICAL IMPLICATIONS OF THE BOUNDS

The per-edge error bound (Thm. 2) recommends: (i) choose $k \gg q$ (e.g., $k \in [128, 192]$ with $q \in [64, 96]$ in vision), (ii) keep radii small enough to ensure large neighbor overlap, (iii) use global whitening and $\text{SO}(p)$ projection to avoid stepwise gauge drift and reflection flips (empirically, self-loop bias $< 10^{-6}$).

S.9 EXTENDED TRAINING DYNAMICS

To examine how holonomy evolves during optimization, we tracked h_{norm} across epochs. On MNIST, holonomy rises sharply during the first few passes over the data and then plateaus, indicating that the pathwise structure of the representation is established early and stabilizes thereafter.

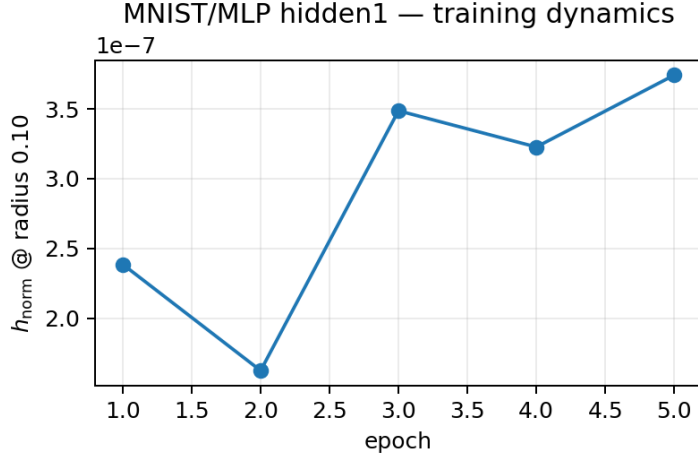


Figure 6: **MNIST training dynamics.** Mean $\pm 95\%$ CI of h_{norm} across epochs.

S.10 EIGEN-ANGLE SPECTRA

Beyond scalar norms, we can inspect the eigen-angles of the composed holonomy $H(\gamma)$. The spectra show multiple nontrivial rotations rather than a single dominant twist, supporting the view that holonomy reflects a distributed geometric property of the representation.

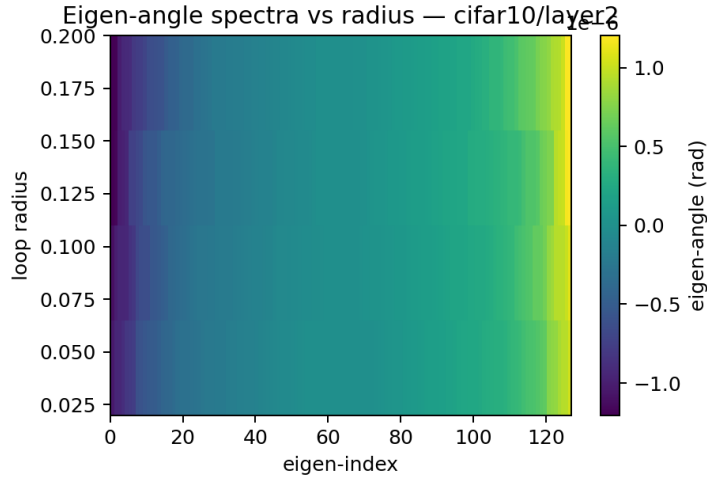


Figure 7: **Eigen-angle spectra (CIFAR-10, layer2).** Distribution of loop holonomy eigen-angles.

S.11 EFFICIENCY AND COMPRESSION

We benchmark the estimator’s runtime and memory with and without dimension compression. Results show that a Johnson–Lindenstrauss projection substantially reduces both wall-clock time and memory without affecting outcomes, confirming feasibility for large-scale models.

Table 4: Wall-clock and memory requirements of the estimator.

	compressed	False	True
k	q		
96	32	27747.000000	892.000000
	64	28986.000000	828.000000
	96	28403.000000	908.000000
128	32	55952.000000	936.000000
	64	57022.000000	925.000000
	96	55239.000000	958.000000
192	32	69469.000000	1011.000000
	64	70521.000000	943.000000
	96	72477.000000	1065.000000

S.12 FURTHER ABLATIONS

We varied estimator hyperparameters to test robustness. Table 5 shows that holonomy values are stable across a wide (k, q) grid. Table 6 highlights the importance of the guardrails: local whitening or separate k -NNs produce inflated or unstable estimates, whereas the shared-midpoint + $\text{SO}(p)$ choice yields consistent results.

Table 5: MNIST Hidden1: mean h_{norm} at $r = 0.10$ across (k, q) .

k	32	64	96
96	8.31e-07	6.33e-07	6.74e-07
128	6.44e-07	5.21e-07	5.68e-07
192	4.14e-07	4.54e-07	4.7e-07

Table 6: Ablations on MNIST Hidden1 at $r = 0.10$. Δ is relative to the best (smallest) h_{norm} configuration.

whitening	neighbors	group	h_{norm}	delta
global	shared	SO	6.42e-07	4.2e-07
global	shared	O	6.42e-07	4.2e-07
local	shared	SO	2.22e-07	0
global	separate	SO	0.222	0.222

S.13 STABILITY TO POOL SIZE

We also varied the standardization pool size N_{pool} . Table 7 shows that increasing from 10^3 to 8×10^3 samples produces only minor changes, indicating practical insensitivity to this parameter.

Table 7: Effect of standardization pool size on h_{norm} (MNIST Hidden1, $r = 0.10$).

pool size	mean	std
1024	2.92e-07	nan
2048	3.81e-07	nan
4096	5.83e-07	nan
8192	5.26e-07	nan

S.14 GAUGE, ORIENTATION, AND BIAS FLOOR CONTROLS

To confirm validity, we tested invariances and null cases. Random orthogonal reparameterization leaves holonomy unchanged (Table 8); near-zero self-loops yield $\mathcal{O}(10^{-8})$ bias floors (Table 9); and replacing nonlinearities with identity (linear null) collapses holonomy to noise level (Table 10).

Table 8: Gauge invariance: change in h_{norm} after random orthogonal reparameterization.

setting	$mean \Delta h $	$max \Delta h $
MNIST Hidden1	nan	nan
CIFAR-10 layer2	nan	nan

Table 9: Self-loop bias floor on MNIST Hidden1.

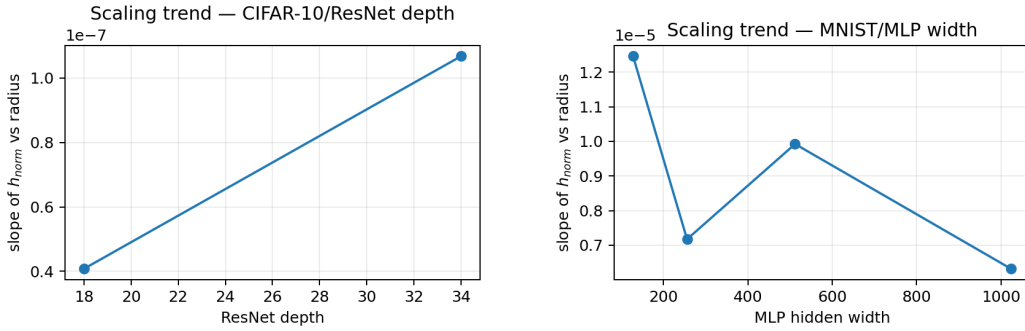
mean bias	std bias	max bias	n
1.99e-07	4.82e-08	2.79e-07	60

Table 10: Linear-null control: replacing nonlinearities with identity drives h_{norm} to noise.

mean h norm	std	n
3.32e-07	4.29e-08	5

S.15 DEPTH/WIDTH SCALING

Holonomy scales systematically with network size. On CIFAR-10, deeper ResNets exhibit steeper slopes of h_{norm} vs. radius (Table 11), while on MNIST, wider MLPs show consistent though saturating growth (Table 12).

Figure 8: **Scaling.** Left: CIFAR-10 depth slice; Right: MNIST width slice.Table 11: CIFAR-10: slope of h_{norm} vs. radius by network depth.

depth	slope h per r
18	4.08e-08
34	1.07e-07

Table 12: MNIST: slope of h_{norm} vs. radius by hidden width.

width	slope h per r
128	1.25e-05
256	7.17e-06
512	9.92e-06
1024	6.32e-06

S.16 CIFAR-10 RESULTS

On CIFAR-10, both `layer1` and `layer2` exhibit positive holonomy at $r = 0.10$, with magnitudes similar to CIFAR-10 (Table 13). This indicates that holonomy generalizes across dataset complexity.

Table 13: CIFAR-10: h_{norm} at $r = 0.10$ by layer.

layer	mean	std
layer1	6.01e-07	7.92e-09
layer2	4.45e-07	4.73e-08

S.17 ALTERNATIVE PATHWISE METRICS

We compared holonomy against other proposed pathwise statistics. Scatter plots versus Lipschitz constants and path curvature (Figure 9) show that while correlated, these alternatives do not subsume holonomy, supporting its distinctiveness as a geometric descriptor.

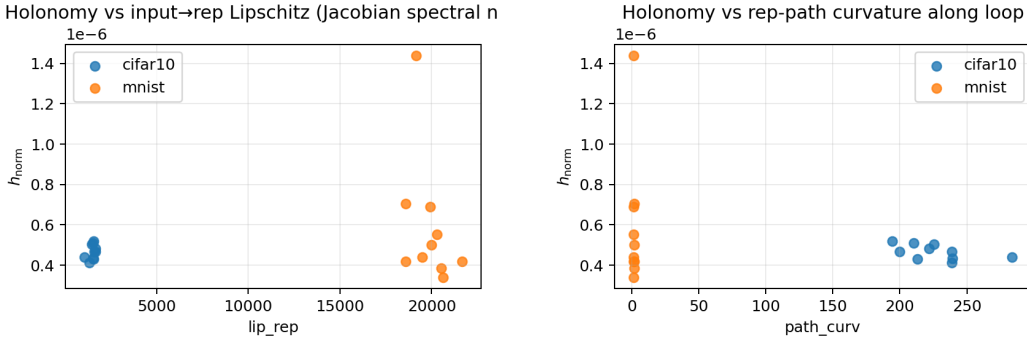


Figure 9: **Comparison to other pathwise statistics.** Holonomy vs. representation Lipschitz (left) and path curvature (right).

B ADDITIONAL DIAGNOSTIC EXPERIMENTS

In this section we report three additional diagnostic experiments that probe (i) behaviour in simple “ground truth” settings, (ii) sensitivity to loop discretisation, and (iii) the small-radius / self-loop regime of the estimator.

B.1 EXPERIMENT A: TOY EQUIVARIANCE VS. ALIASING ON MNIST

We construct a simple convolutional toy setting on MNIST with two small networks that are intentionally different from a geometric point of view:

1. **EquivConvNetMNIST**: a CNN with only stride-1 convolutions, circular padding, and no pooling, which is approximately translation-equivariant on the grid;

2. **AliasConvNetMNIST**: a CNN with zero padding and max-pooling, which introduces strong aliasing and boundary artefacts.

Both models are trained on MNIST with the same optimisation hyperparameters. We then measure representation holonomy at the second convolutional layer along a short loop of integer translations of a single test image, using the discrete path

$$(0, 0) \rightarrow (0, 1) \rightarrow (1, 1) \rightarrow (1, 0) \rightarrow (0, 0)$$

in pixel space (implemented via circular shifts of the image).

Table 14 reports the normalized holonomy h_{norm} , the Frobenius norm $\|H - I\|_{\text{Fro}}$ of the holonomy operator, and the maximum eigen-angle (in radians) of H .

Table 14: **Experiment A (MNIST, toy equivariance vs. aliasing)**. Holonomy at the second convolutional layer along a discrete translation loop. The nearly equivariant network exhibits holonomy close to zero, whereas the aliased network shows substantially larger holonomy.

Model	h_{norm}	$\ H - I\ _{\text{Fro}}$	max eigen-angle
EquivConvNetMNIST	8.41×10^{-7}	1.3×10^{-5}	6×10^{-6}
AliasConvNetMNIST	2.99×10^{-4}	4.78×10^{-3}	3×10^{-3}

Under the same translation loop, the aliased network exhibits roughly three orders of magnitude larger holonomy than the approximately equivariant one, providing a clean “ground truth” sanity check: when translation symmetry is respected, holonomy is essentially zero, and when it is broken by padding and pooling artefacts, holonomy is large.

B.2 EXPERIMENT C: SENSITIVITY TO LOOP DISCRETISATION

To probe sensitivity to the number of loop points, we fix a trained MNIST MLP and measure holonomy at Hidden 1 for a fixed radius r while varying only the loop discretisation. Loops are constructed as regular polygons with $n_{\text{points}} \in \{6, 8, 12, 16, 24\}$ in a local two-dimensional PCA plane around each base point; all other estimator settings (neighbourhood size, subspace dimension, whitening pool, etc.) are kept fixed.

Table 15 reports the resulting h_{norm} values.

Table 15: **Experiment C (MNIST MLP, loop discretisation)**. Holonomy at Hidden 1 for different numbers of loop samples n_{points} at fixed radius. Values vary smoothly with n_{points} , with no indication of instability.

n_{points}	h_{norm}
6	3.51×10^{-7}
8	3.87×10^{-7}
12	3.95×10^{-7}
16	4.15×10^{-7}
24	4.70×10^{-7}

Holonomy varies smoothly and monotonically with n_{points} (see Figure 10). Increasing the number of loop points by a factor of four changes h_{norm} by only $\sim 1.2 \times 10^{-7}$ (about a 30% relative change, but of the same order of magnitude), suggesting that the estimator is numerically stable with respect to reasonable changes in loop discretisation. In the main text we therefore adopt $n_{\text{points}} = 12$ as a compute-accuracy compromise.

B.3 EXPERIMENT D: SMALL-RADIUS AND SELF-LOOP REGIME

Finally, we study the very small-radius regime and the numerical floor of the estimator. On a (separately) trained MNIST MLP at Hidden 1 we construct:

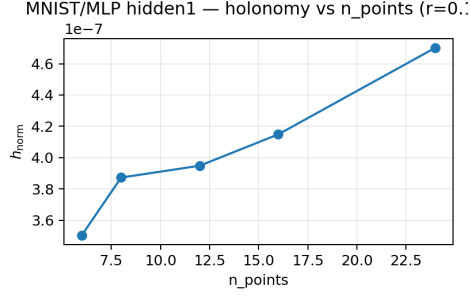


Figure 10: **MNIST/MLP Hidden 1: holonomy vs. loop discretisation.** Normalised holonomy h_{norm} at radius $r = 0.10$ as a function of the number of loop samples $n_{\text{points}} \in \{6, 8, 12, 16, 24\}$. The curve is smooth and monotone, with all values in the range $3.5\text{--}4.7 \times 10^{-7}$, indicating that the estimator is stable with respect to loop discretisation.

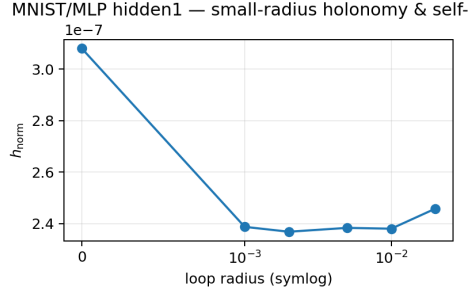


Figure 11: **MNIST/MLP Hidden 1: small-radius holonomy and self-loop.** Normalised holonomy h_{norm} for an exact self-loop ($r = 0$) and PCA loops with very small radii $r \in \{10^{-3}, 2 \times 10^{-3}, 5 \times 10^{-3}, 10^{-2}, 2 \times 10^{-2}\}$. All non-zero radii lie in the narrow band $h_{\text{norm}} \in [2.37, 2.46] \times 10^{-7}$, i.e. variation $\approx 3 \times 10^{-8}$, which characterises the numerical floor of the estimator and is consistent with the $O(r)$ small-radius behaviour.

- a *self-loop* in which the same image is repeated at all loop points (radius $r = 0$); and
- PCA-based loops with radii $r \in \{10^{-3}, 2 \times 10^{-3}, 5 \times 10^{-3}, 10^{-2}, 2 \times 10^{-2}\}$.

Table 16 shows h_{norm} as a function of radius.

Table 16: **Experiment D (MNIST MLP, small-radius and self-loop).** Holonomy at Hidden 1 for an exact self-loop ($r = 0$) and for very small PCA loops. All values lie in a narrow band, characterising the numerical floor of the estimator in this setting.

Radius	h_{norm}
0	3.08×10^{-7}
0.001	2.39×10^{-7}
0.002	2.37×10^{-7}
0.005	2.38×10^{-7}
0.010	2.38×10^{-7}
0.020	2.45×10^{-7}

For $r \leq 0.02$, holonomy remains essentially flat in the narrow band $h_{\text{norm}} \in [2.37, 2.46] \times 10^{-7}$ (variation $\approx 3 \times 10^{-8}$). Together with the self-loop value at $r = 0$, these numbers characterise the numerical floor of our estimator and are consistent with the $O(r)$ small-radius behaviour established in Theorem 1: the increases with radius reported in the main figures only become visible once we leave this floor and move to radii where perturbations have a semantic effect, see Figure 11.