CAP$_{\mathcal{M}}$: Curvature-Aware Pulling on Riemannian Manifolds

Anonymous Authors

Abstract

Learning signed distance functions (SDFs) on curved spaces is common in geometry processing and non-Euclidean machine learning, yet most "pull/projection" updates are ambient and ignore curvature. We study an *intrinsic* pull operator that advances along geodesics with an explicit normalization to tame gradient-magnitude errors. Our contributions are threefold: (i) a *curvature-explicit* contraction guarantee in a tubular neighborhood around the target surface, including a necessary–sufficient step-size window and a closed-form optimal step; (ii) a *practical path-safety rule* and an off-band Lyapunov "funneling" effect that make the update usable from realistic initializations; and (iii) a *discretization-aware* refinement that absorbs mesh, distance, transport and retraction errors into measurable constants. The result is a small, verifiable recipe: an intrinsic step, and theory that predicts what one should observe in practice.

Keywords: geometry, topology, signed distance functions, Riemannian optimization

1. Motivation and Overview

In Euclidean settings, SDF-based "pull" steps toward a surface are straightforward. On a manifold $(\mathcal{M}, g)$, however, taking ambient steps misrepresents distances and directions. **In brief:** we design a *geodesic* pull step that (1) behaves like the Euclidean update when curvature is small, (2) quantifies how curvature penalizes overly large steps, and (3) comes with knobs and readouts practitioners can *measure* to stay in the safe regime.

2. Problem Setup and Notation

Let $(\mathcal{M}, g)$ be a complete C^2 Riemannian manifold with sectional curvature $|K| \le \kappa$ on the region of interest and injectivity radius $\mathrm{inj}(\mathcal{M}) \ge \iota > 0$. Let the target be a closed, embedded, oriented C^2 hypersurface $\mathcal{S} \subset \mathcal{M}$ with reach $\rho > 0$. For $0 < r < \min\{\rho, \iota\}$ define the tube $\mathcal{B}_r := \{x : \mathrm{dist}_g(x, \mathcal{S}) < r\}$. Let d be the signed geodesic distance (orientation fixes the sign), with unit normal $n = \nabla_g d$, $\|n\|_g = 1$. Assume curvature control in $\mathcal{B}_r$: $\left\|\nabla_g^2 d\right\| \le C_d = C_d(\kappa, \|\mathrm{II}\|, r)$.
We learn $f : \mathcal{M} \to \mathbb{R}$ to approximate d in $\mathcal{B}_r$.

3. Related Work

Implicit fields and Euclidean SDF learning. Neural implicit representations have become a standard tool for geometry, with models that learn continuous signed distance functions or occupancy fields from raw data (Park et al., 2019; Mescheder et al., 2019; Atzmon and Lipman, 2020). Architectural choices such as periodic activations further improve fidelity and signal propagation (Sitzmann et al., 2020), while geometric regularizers promote distance-like behavior (Gropp et al., 2020). These approaches typically update points or parameters in *ambient* space; on curved domains this can distort directions and step lengths, which is precisely what our intrinsic pull addresses.

Geodesic distances and manifold numerics. A large body of work studies distances and PDEs on surfaces. Fast marching and related methods approximate geodesics by propagating fronts (Sethian, 1996; Kimmel and Sethian, 1998), while the heat method provides robust, mesh-friendly distance estimates (Crane et al., 2013). For computations on embedded surfaces, closest-point retractions and embedding techniques allow one to evolve intrinsic flows using standard solvers (Ruuth and Merriman, 2008). Our analysis assumes a standard tubular neighborhood where nearest-point projection is unique; this is governed by classical reach and curvature arguments (Federer, 1959; Lee, 2018; do Carmo, 1992) and enters our bounds via explicit curvature/radius constants.

Optimization on manifolds. Riemannian optimization provides principled updates (retractions, parallel transport, trust regions) and convergence theory beyond Euclidean space (Absil et al., 2008; Boumal, 2023). Stochastic and variance-reduced methods further adapt these ideas to large-scale learning (Bonnabel, 2013; Zhang and Sra, 2016). While this literature offers general-purpose tools, it usually optimizes abstract objectives; in contrast, we target a *distance-to-surface* objective and quantify how curvature and discretization affect a simple, normal-direction pull—yielding a concrete contraction factor with a necessary–sufficient step-size window and a closed-form optimum.

Positioning and gap. Compared to Euclidean pull/projection steps used with neural SDFs (Park et al., 2019; Atzmon and Lipman, 2020; Gropp et al., 2020), our operator is *intrinsic* (geodesic) and *normalized* to mitigate gradient-scale errors, with theory that predicts what one should observe in practice under measurable band conditions. Relative to general Riemannian methods (Absil et al., 2008; Boumal, 2023), we trade quadratic rates (which would require accurate Hessians/SDFs) for a lightweight step whose constants expose the role of curvature, band radius, and discretization. Finally, when discussing band-to-uniform guarantees for the diagnostic losses we rely on standard uniform-convergence tools (Bartlett and Mendelson, 2002; Shalev-Shwartz and Ben-David, 2014), which justify reporting simple, interpretable metrics (eikonal error, angle alignment, value fidelity) alongside the predicted contraction curve.

4. Method: A Geodesic Pull with Normalization

We take a geodesic step with a scale that normalizes the gradient magnitude:

$$T_\gamma(x) \;=\; \exp_x\Big(-\gamma\,\frac{f(x)}{\mu(x)}\,\nabla_g f(x)\Big), \qquad \mu(x) = \max\{\|\nabla_g f(x)\|_g\,, \tau\}, \tag{1}$$

with step-size $\gamma \in (0, 1]$ and clamp $\tau > 0$.

Rationale: Moving along geodesics respects the geometry; scaling by μ damps spurious overshoot when $\|\nabla_g f\| \neq 1$. We measure and train inside the band $\Omega_r := \{x : |f(x)| \leq r\}$.

5. Assumptions

In $\mathcal{B}_r$, suppose:

- **(A1) Near-eikonal.** $\|\nabla_g f\|$ is close to 1 (distance-like scaling).
- **(A2) Angle control.** $\nabla_g f$ points mostly along the true normal n (directional alignment).
- **(A3) Value fidelity.** f is a small relative error proxy for d (signed distance accuracy).
- **(A4) Clamp inactive in-band.** Choose $\tau \leq 1 - \eta$ so normalization is intrinsic where we prove contraction.

These are exactly the measurable properties an SDF should have near a surface; and they translate into a clean distance-decay bound.

6. Theoretical Claims

We study the intrinsic pull (1) on a tubular neighborhood $\mathcal{B}_r$ of the target hypersurface $\mathcal{S} \subset (\mathcal{M}, g)$, where d denotes signed geodesic distance and $\nabla_g^2 d$ is bounded by C_d on $\mathcal{B}_r$. Inside $\mathcal{B}_r$, assume the standard in-band conditions:

$$\text{(A1) } 1-\eta \leq \|\nabla_g f\|_g \leq 1+\eta, \quad \text{(A2) } \left\langle \widehat{\nabla_g f},\, \nabla_g d \right\rangle_g \geq 1-\eta, \quad \text{(A3) } |f-d| \leq \varepsilon |d|, \quad \text{(A4) } \tau \leq 1-\eta.$$

C1 — Curvature-aware one-step contraction and optimal step. Let $y = T_\gamma(x)$ with $x \in \mathcal{B}_r$. A geodesic Taylor bound gives

$$|d(y)| \;\leq\; \left(1 - a\gamma + b\gamma^2\right) |d(x)|, \qquad a := (1-\eta)(1-\varepsilon), \quad b := \tfrac{1}{2}\, C_d (1+\varepsilon)^2 r. \tag{2}$$

Hence $q(\gamma) := 1 - a\gamma + b\gamma^2 < 1$ *iff* $0 < \gamma < \frac{a}{b}$ (necessary & sufficient window), and the minimizer is

$$\gamma^\star = \arg\min_{\gamma > 0} q(\gamma) = \frac{a}{2b} = \frac{(1-\eta)(1-\varepsilon)}{C_d (1+\varepsilon)^2 r}. \tag{3}$$

Rationale: The linear term $(-a\gamma)$ is the desired normal pull; the quadratic penalty $(+b\gamma^2)$ comes from curvature and band radius. Small positive γ guarantees contraction; $\gamma^\star$ trades these effects optimally.

C2 — Variable-step convergence (diminishing/adaptive). For any sequence $\gamma_k \in (0, a/b)$,

$$|d(x_{k+1})| \;\leq\; \left(1 - a\gamma_k + b\gamma_k^2\right) |d(x_k)| \;\Rightarrow\; |d(x_K)| \;\leq\; |d(x_0)|\, \exp\!\Big(-a \sum_{k<K} \gamma_k + b \sum_{k<K} \gamma_k^2\Big). \tag{4}$$

If $\sum_k \gamma_k = \infty$ and $\sum_k \gamma_k^2 < \infty$, then $d(x_k) \to 0$ (linear-to-tolerance decay).

Rationale: The exponential envelope exposes how accumulated step length drives progress, while curvature/discretization only taxes the quadratic part.

C3 — Path-safety cap and forward invariance. Contraction controls endpoints; to keep the entire geodesic inside valid charts (reach ρ, injectivity ι), enforce

$$\gamma \;\le\; \frac{r_{\text{safety}}}{(1+\varepsilon)r}, \qquad r_{\text{safety}} := \tfrac{1}{2}\min\{\rho - r,\ \iota - r\}. \tag{5}$$

If additionally $q(\gamma) < 1$ and $x \in \mathcal{B}_r$, then $y = T_\gamma(x) \in \mathcal{B}_r$ and $|d(y)| < |d(x)|$ (forward invariance).

Rationale: A conservative path-length bound keeps the step inside the trusted tube where the Taylor bound is valid.

C4 — Discretization-aware version. With mesh scale h, set $\tilde{\eta} = \eta + c_1 O(h)$, $\tilde{\varepsilon} = \varepsilon + c_2 O(h)$ and

$$b_{\text{disc}} := \tfrac{1}{2}\, C_{\text{ret}}(h)\,(1+\tilde{\varepsilon})^2 r, \qquad C_{\text{ret}}(h) = O(h). \tag{6}$$

Then the bound persists as

$$|d(T_\gamma(x))| \;\le\; \big(1 - \tilde{a}\,\gamma + (\tilde{b} + b_{\text{disc}})\gamma^2\big)\,|d(x)|, \quad \tilde{a} = (1-\tilde{\eta})(1-\tilde{\varepsilon}), \quad \tilde{b} = \tfrac{1}{2}\,\tilde{C}_d(1+\tilde{\varepsilon})^2 r,$$

and the admissible window shrinks smoothly with h.

Rationale: First-order errors nudge the linear term; retraction/transport mismatches enter quadratically (via step length), preserving the same contraction structure.

7. Limitations and Scope

Local tube assumptions; linear (not Newton) rates; measured constants depend on mesh quality and shape regularity; the Lyapunov funneling requires a small off-band step relative to a gradient Lipschitz constant.

References

P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2008.

M. Atzmon and Y. Lipman. Sal: Sign agnostic learning of shapes from raw data. In *CVPR*, 2020.

P. L. Bartlett and S. Mendelson. Rademacher and gaussian complexities: Risk bounds and structural results. *Journal of Machine Learning Research*, 3:463–482, 2002.

S. Bonnabel. Stochastic gradient descent on riemannian manifolds. *IEEE Transactions on Automatic Control*, 58(9):2217–2229, 2013.

N. Boumal. *An Introduction to Optimization on Smooth Manifolds*. Cambridge University Press, 2023.

K. Crane, C. Weischedel, and M. Wardetzky. Geodesics in heat: A new approach to computing distance based on heat flow. *ACM Transactions on Graphics*, 32(5):152:1–152:11, 2013.

M. P. do Carmo. *Riemannian Geometry*. Birkhäuser, 1992.

H. Federer. Curvature measures. *Transactions of the American Mathematical Society*, 93 (3):418–491, 1959.

A. Gropp, L. Yariv, N. Haim, M. Atzmon, and Y. Lipman. Implicit geometric regularization for learning shapes. In *NeurIPS*, 2020.

R. Kimmel and J. A. Sethian. Computing geodesic paths on manifolds. *Proceedings of the National Academy of Sciences*, 95(15):8431–8435, 1998.

J. M. Lee. *Introduction to Riemannian Manifolds*. Springer, 2 edition, 2018.

L. Mescheder, M. Oechsle, M. Niemeyer, S. Nowozin, and A. Geiger. Occupancy networks: Learning 3d reconstruction in function space. In *CVPR*, 2019.

J. J. Park, P. Florence, J. Straub, R. A. Newcombe, and S. Lovegrove. Deepsdf: Learning continuous signed distance functions for shape representation. In *CVPR*, 2019.

S. J. Ruuth and B. Merriman. A simple embedding method for solving partial differential equations on surfaces. *Journal of Computational Physics*, 227(3):1943–1961, 2008.

J. A. Sethian. A fast marching level set method for monotonically advancing fronts. *Proceedings of the National Academy of Sciences*, 93(4):1591–1595, 1996.

S. Shalev-Shwartz and S. Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014.

V. Sitzmann, J. N. P. Martel, A. W. Bergman, D. B. Lindell, and G. Wetzstein. Implicit neural representations with periodic activation functions. In *NeurIPS*, 2020.

H. Zhang and S. Sra. Riemannian svrg: Fast stochastic optimization on riemannian manifolds. In *NeurIPS*, 2016.

Appendix:

Notation: $\langle\cdot,\cdot\rangle_g$ and $\|\cdot\|_g$ denote the metric and norm; ∇_g and ∇_g^2 are gradient and covariant Hessian; $\exp_x$ is the exponential map at x.

Lemma 1 (Riemannian Taylor bound) *Let $u:\mathcal{M}\to\mathbb{R}$ be C^2 on a geodesically convex set $\mathcal{U}$. For any $x\in\mathcal{U}$ and $v\in T_x\mathcal{M}$ such that $\gamma(t)=\exp_x(tv)\in\mathcal{U}$ for $t\in[0,1]$, there exists $\xi=\gamma(\theta)$ with $\theta\in(0,1)$ s.t.*

$$u(\exp_x v) \;=\; u(x)+\langle\nabla_g u(x),v\rangle_g+\tfrac{1}{2}\,\nabla_g^2 u(\xi)[v,v].$$

In particular, $|u(\exp_x v)-u(x)-\langle\nabla_g u(x),v\rangle_g|\le\frac{1}{2}\|\nabla_g^2 u\|_{L^\infty(\mathcal{U})}\,\|v\|_g^2$.

Proof sketch: Apply the 1D Taylor theorem to $t\mapsto u(\gamma(t))$ and use $\dot\gamma(0)=v$ and $\frac{d^2}{dt^2}u(\gamma(t))=\nabla_g^2 u(\gamma(t))[\dot\gamma(t),\dot\gamma(t)]$.

Lemma 2 (Distance field structure) *Let d be the signed distance to a C^2 oriented hypersurface $\mathcal{S}$ on a tube $\mathcal{B}_r$ where nearest-point projection is unique. Then $\nabla_g d=n$ with $\|n\|_g=1$, and $\|\nabla_g^2 d\|\le C_d(\kappa,\|\mathrm{II}\|,r)$ on $\mathcal{B}_r$.*

Proof sketch: Standard properties of the distance function on a tubular neighborhood; see normal coordinates and reach/injectivity results.

Theorem 3 (Contraction and optimal step) *Under (A1)–(A4) on $\mathcal{B}_r$, let $y=T_\gamma(x)=\exp_x(v)$ with $v=-\gamma\frac{f}{\mu}\nabla_g f$ and $x\in\mathcal{B}_r$. Then*

$$|d(y)|\;\le\;\big(1-a\gamma+b\gamma^2\big)\,|d(x)|,\qquad a=(1-\eta)(1-\varepsilon),\quad b=\tfrac{1}{2}\,C_d(1+\varepsilon)^2 r,$$

so $q(\gamma)=1-a\gamma+b\gamma^2<1$ iff $0<\gamma<\frac{a}{b}$, with minimizer $\gamma^\star=\frac{a}{2b}$.

Proof sketch: Apply Lemma 1 to $u=d$: $d(y)=d(x)+\langle\nabla_g d(x),v\rangle_g+R_2$ with $|R_2|\le\frac{1}{2}C_d\|v\|_g^2$. By (A4), $\mu=\|\nabla_g f\|_g$ on $\mathcal{B}_r$, so $\langle\nabla_g d,v\rangle_g=-\gamma\,\frac{f}{\mu}\langle\nabla_g d,\nabla_g f\rangle_g=-\gamma f\,\Big\langle\widehat{\nabla_g f},\,\nabla_g d\Big\rangle_g\le-\gamma(1-\eta)f$. With (A3), f shares the sign of d and $|f|\ge(1-\varepsilon)|d|$, giving the linear term $-a\gamma|d|$. Also $\|v\|_g=\gamma|f|\le\gamma(1+\varepsilon)|d|\le\gamma(1+\varepsilon)r$, so $|R_2|\le\frac{1}{2}C_d\gamma^2(1+\varepsilon)^2 r\,|d|$. Combine and take absolute values to obtain the factor $1-a\gamma+b\gamma^2$. Convexity of $q(\gamma)$ yields the window and $\gamma^\star$.

Proposition 4 (Variable-step convergence) *For $x_{k+1}=T_{\gamma_k}(x_k)$ with $\gamma_k\in(0,a/b)$ and $x_0\in\mathcal{B}_r$,*

$$|d(x_K)|\;\le\;|d(x_0)|\prod_{k=0}^{K-1}\big(1-a\gamma_k+b\gamma_k^2\big)\;\le\;|d(x_0)|\,\exp\Big(-a\sum_{k<K}\gamma_k+b\sum_{k<K}\gamma_k^2\Big).$$

If $\sum_k\gamma_k=\infty$ and $\sum_k\gamma_k^2<\infty$, then $d(x_K)\to 0$.

Proof sketch: Iterate Theorem 3 and use $1 + u \le e^u$ with $u_k = -a\gamma_k + b\gamma_k^2$.

Proposition 5 (Path-safety cap) *Let* $r_{\text{safety}} := \frac{1}{2}\min\{\rho - r,\ \iota - r\}$. *If* $\|v\|_g \le r_{\text{safety}}$ *and* $x \in \mathcal{B}_r$*, the whole geodesic* $\exp_x(tv)$ *stays within the normal/injective neighborhood for* $t \in [0, 1]$*. In particular, it suffices to enforce*

$$\gamma \ \le\ \frac{r_{\text{safety}}}{(1+\varepsilon)r}.$$

Proof sketch: By reach and injectivity assumptions, balls of radius $\rho - r$ and $\iota - r$ around x lie within the valid tube and normal neighborhood. Any geodesic of length $\le r_{\text{safety}}$ remains inside; since $\|v\|_g \le \gamma(1+\varepsilon)r$, the stated bound guarantees this.

Proposition 6 (Off-band Lyapunov funneling) *Let* $F(x) = \frac{1}{2}f(x)^2$*. Assume* $\nabla_g f$ *is* L*-Lipschitz off-band and* $\mu \ge \tau > 0$*. For a* C^2 *retraction and* $v = -\gamma\frac{f}{\mu}\nabla_g f$*,*

$$F(\mathrm{Retr}_x(v)) \ \le\ F(x) - \gamma\,\frac{\|\nabla_g f(x)\|_g^2}{\mu(x)}\,F(x) + c\,\gamma^2\|f(x)\nabla_g f(x)\|_g^2.$$

Thus choosing $\gamma \le \tau/(2cL)$ *yields* F *strictly decreasing off-band, funneling iterates into* Ω_r*.*

Proof sketch: Use the Riemannian descent lemma on F along the retraction; $\nabla_g F = f\nabla_g f$ gives the linear decrease term, and Lipschitzness controls the quadratic remainder.

Theorem 7 (Discretization-aware contraction) *With mesh scale* h *and first-order errors in gradient and distance, plus* $O(h)\|v\|_g^2$ *retraction/transport errors, set* $\tilde{\eta} = \eta + c_1 O(h)$*,* $\tilde{\varepsilon} = \varepsilon + c_2 O(h)$*,* $b_{\text{disc}} = \frac{1}{2}C_{\text{ret}}(h)(1+\tilde{\varepsilon})^2 r$ *with* $C_{\text{ret}}(h) = O(h)$*. Then*

$$|d(T_\gamma(x))| \ \le\ \big(1 - \tilde{a}\,\gamma + (\tilde{b} + b_{\text{disc}})\gamma^2\big)\,|d(x)|, \quad \tilde{a} = (1-\tilde{\eta})(1-\tilde{\varepsilon}), \quad \tilde{b} = \tfrac{1}{2}\,\tilde{C}_d(1+\tilde{\varepsilon})^2 r,$$

and the continuous bound is recovered as $h \to 0$*.*

Proof sketch: Repeat Theorem 3 tracking perturbations: the linear term is modified by $\tilde{\eta}, \tilde{\varepsilon}$; remainder collects $C_d\|v\|_g^2$ and $O(h)\|v\|_g^2$ from retraction/transport, grouped into $\tilde{b} + b_{\text{disc}}$.