CERTIFIED FEASIBILITY FOR INEXACT RIEMANNIAN FRANK–WOLFE ON STRONGLY CONVEX SETS

Anonymous authors
Paper under double-blind review

ABSTRACT

We study constrained optimization on Riemannian manifolds where every iterate must remain feasible despite numerical approximation of the updates. Riemannian Frank–Wolfe preserves feasibility by moving along geodesics within a geodesically convex constraint set. Numerical approximation and finite-precision storage, however, can destroy this guarantee even when the output remains a valid manifold point. For numerical maps, we derive an acceptance rule combining a geometric tolerance from strong convexity of the feasible set with a quadratic descent tolerance; given feasible oracle targets, valid error bounds, and the stated oracle-accuracy and step-schedule conditions, accepted iterates remain feasible and retain the standard sublinear bound for smooth geodesically convex objectives. We then derive an intrinsic tangent-cone margin and, under smooth logarithms, an exact-oracle scaling inequality with a multiplicative conditioning factor. We use this inequality to transfer known faster rates when the gradient norm is bounded away from zero. We derive sharp rotation-ball formulas that make the geometric tolerances explicit. Finally, for a fixed-sub-step Euler integrator in a rotation chart, we prove that mantissa precision logarithmic in the inverse target accuracy suffices, given an exact feasible oracle, the stated primitive-error model, and sufficient available precision. Rotation experiments show that the acceptance rule maintains feasibility and recovers reference accuracy under low-precision storage through accurate fallback.

1 INTRODUCTION

We study optimization over a geodesically convex subset of a Riemannian manifold when every computed iterate must satisfy the constraint. One concrete case is a constrained form of rotation averaging, a basic problem in computer vision and robotics (Hartley et al., 2013), in which the orientation estimated from noisy rotations must remain within a prescribed angle of a reference attitude throughout optimization. In exact arithmetic, Riemannian Frank–Wolfe preserves feasibility through its update geometry. It chooses a feasible target by minimizing a linear model over the feasible set and moves toward it along the connecting geodesic, so geodesic convexity keeps the update feasible without a projection (Weber & Sra, 2023).

This exact argument, however, does not automatically survive numerical realization. Retractions replace the exponential map (Absil et al., 2008; Boumal, 2023), numerical integrators approximate the geodesic, and finite-precision storage can perturb the output. Even an output represented or restored as a valid manifold point can violate the additional constraint. Feasibility alone is also insufficient for the optimization analysis, because a feasible perturbation can erase the predicted descent. The relevant quantity is therefore the total error of the final represented endpoint, and its tolerance must reflect both the constraint geometry and the progress that the analysis requires.

Existing Frank–Wolfe analyses admit several forms of inexact direction information. Approximate linear minimization is standard in Euclidean Frank–Wolfe (Jaggi, 2013) and in its Riemannian extension (Weber & Sra, 2023), and stochastic Riemannian variants account for gradient uncertainty (Weber & Sra, 2022). Strong convexity of the feasible set and the associated scaling inequalities yield faster rates (Scieur et al., 2026; Chen, 2026). An exact-update feasibility argument, however, does not by itself control a perturbed endpoint, even when the chosen target remains feasible.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Other approaches control the update map or test the trial point directly. Li et al. (2024) use a lifted feasible construction in tangent space or a line search that checks both membership and an objective condition for each retracted trial point, and Chen & Fukuda (2026) analyse conditional-gradient steps under feasibility assumptions on the retraction curves. A membership test certifies feasibility of the points it accepts, but by itself it bounds neither their deviation from the intended update nor the resulting loss of progress. Analyses of inexact manifold primitives quantify approximation errors (Díaz et al., 2026), but membership in the manifold alone does not certify an additional constraint within it. Finally, an error indicator suffices only if it is a valid bound covering every operation that affects the represented output (Rump, 2010).

This paper therefore addresses the following question.

How accurately must a Riemannian Frank–Wolfe update be computed so that every iterate remains feasible and the method retains its convergence guarantees?

Contributions. First, we derive sufficient conditions on the endpoint error, that is, the distance from the computed update to the exact geodesic update. Section 3 combines a geometric budget that guarantees feasibility with a progress tolerance that controls the objective perturbation; a numerical update is accepted only when a valid endpoint-error bound satisfies both. Section 4 proves that accepted iterates retain the standard sublinear bound under the stated oracle and schedule assumptions and shows how exact-oracle scaling transfers faster rates.

Second, we derive the geometry behind these guarantees (Section 5). Intrinsic strong convexity yields a tangent-cone margin without boundary regularity. Under smooth logarithms, exact-oracle optimality turns this margin into a scaling inequality with a multiplicative conditioning factor and no additive residual, whereas the general intrinsic route of Scieur et al. (2026, Proposition 3.3) leaves an additive residual. A gradient norm bounded away from zero makes the scaling uniform. For rotation balls, we prove the sharp set modulus and the exact worst-case finite-step clearance.

Third, we analyse and test numerical realizations on rotations (Section 6 and Appendix D). Under an exact feasible oracle, explicit primitive-error assumptions, and sufficient available precision, a fixed-sub-step Euler policy in a rotation chart needs at most a mantissa precision logarithmic in the inverse target objective accuracy. The experiments show infeasible storage proposals, feasible low-precision iterates that lose accuracy, and the recovery of that accuracy through frequent accurate fallback under a conservative rule.

The framework concerns strongly convex feasible sets with nonempty interior in the chosen manifold, so equality constraints require an appropriate manifold formulation (Remark C.11). It assumes feasible represented targets and valid endpoint-error bounds; it does not certify a general numerical linear oracle. Proofs are in Appendix A; numerical analysis and experiments are in Appendix D.

2 Setting and geometry

Throughout, $(\mathcal{M}, g)$ denotes a complete, connected, finite-dimensional Riemannian manifold with tangent spaces $T_x\mathcal{M}$, inner products $\langle\cdot,\cdot\rangle_x$, norms $\|\cdot\|_x$, distance d, exponential map Exp_x, and, where defined, logarithm Log_x; $\operatorname{dist}(x, S) := \inf_{y\in S} d(x, y)$ for $S \subseteq \mathcal{M}$. The problem is

$$\min_{x\in\mathcal{C}} F(x), \tag{1}$$

where $\mathcal{C} \subset \mathcal{M}$ is the feasible set and F the objective, C^1 near $\mathcal{C}$ and geodesically convex and smooth on $\mathcal{C}$ (Definition 2.2). The objective need only be geodesically convex; strong convexity is imposed on the feasible set to supply the clearance that tolerates endpoint error.

Assumption 2.1 (set regularity)**.** $\mathcal{C} \subset \mathcal{M}$ is nonempty, compact, and uniquely geodesically convex, that is, any $x, y \in \mathcal{C}$ are joined by a unique minimizing geodesic of $\mathcal{M}$ (unique among minimizing segments of $\mathcal{M}$, not merely among those contained in $\mathcal{C}$), and it lies in $\mathcal{C}$. The restricted logarithm $\operatorname{Log}_x\colon \mathcal{C} \to T_x\mathcal{M}$ returns the initial velocity of that geodesic, so $\|\operatorname{Log}_x y\|_x = d(x, y)$ and $\operatorname{Exp}_x \operatorname{Log}_x y = y$.

The assumption fixes a single feasible chord between two feasible points, hence a single exact endpoint against which a numerical output is compared.

Definition 2.2 (function classes)**.** Let $U \subseteq \mathcal{M}$ be uniquely geodesically convex and let φ be differentiable on an open neighbourhood of U. For constants $0 \le \mu \le L$, φ is geodesically μ-strongly convex and L-smooth on U if, for all $x, y \in U$,

$$\tfrac{\mu}{2} d^2(x,y) \;\le\; \varphi(y) - \varphi(x) - \langle \operatorname{grad} \varphi(x), \operatorname{Log}_x y \rangle_x \;\le\; \tfrac{L}{2} d^2(x,y).$$

The case $\mu = 0$ is geodesic convexity.

F is in this class with $\mu = 0$ and constant L_F; a positive μ enters only the appendices.

We study Riemannian Frank–Wolfe (Algorithm 1), whose exact geodesic updates remain feasible. For smooth geodesically convex F, $\eta_k = 2/(k+2)$ gives an $O(1/k)$ objective-gap bound (Weber & Sra, 2023). Our focus is preserving these guarantees under inexact updates.

Algorithm 1 Exact Riemannian Frank–Wolfe (Weber & Sra, 2023)

Require: $x_0 \in \mathcal{C}$; steps $\eta_k \in (0,1]$; iteration count K
1: **for** $k = 0, \dots, K-1$ **do**
2: $\quad q_k \in \arg\min_{q \in \mathcal{C}} \langle \operatorname{grad} F(x_k), \operatorname{Log}_{x_k} q \rangle_{x_k}$ {exact linear minimization oracle}
3: $\quad x_{k+1} \leftarrow \operatorname{Exp}_{x_k}(\eta_k \operatorname{Log}_{x_k} q_k)$ {exact geodesic step}
4: **end for**

The following definition (Scieur et al., 2026) is central to our analysis.

Definition 2.3 (geodesic strong convexity of a set)**.** Let $\mathcal{C}$ satisfy Assumption 2.1 and fix $\alpha > 0$. $\mathcal{C}$ is *geodesically α-strongly convex* if, for every $x, y \in \mathcal{C}$ and $t \in [0,1]$,

$$\{z \in \mathcal{M} : d(\gamma_{xy}(t), z) \le \alpha t(1-t) d^2(x,y)\} \subseteq \mathcal{C}, \tag{2}$$

where γ_{xy} is the constant-speed minimizing geodesic from x to y.

Each non-endpoint of a geodesic chord between distinct points of $\mathcal{C}$ has a ball of positive radius contained in $\mathcal{C}$. The radius in (2) lower-bounds its *clearance*, the distance to $\mathcal{M} \setminus \mathcal{C}$ (Figure 2).

The following condition is used only where the logarithm is differentiated (Theorem 5.1(ii)) and in the sublevel-set theorem (Theorem A.20); closed balls of radius below $\pi/2$ in $\mathrm{SO}(3)$ satisfy it under the metric of Section 5, with a slightly larger open ball as the set V below.

Assumption 2.4 (normal neighbourhood for a set $S \subseteq \mathcal{M}$)**.** There is a totally normal, geodesically convex open neighbourhood $V \supseteq S$ on which $(x,y) \mapsto \operatorname{Log}_x y$ is smooth, and every geodesic segment contained in V is the unique minimizing segment between its endpoints.

3 Certified inexact updates

Let $\mathcal{C}$ satisfy Assumption 2.1 and be geodesically α-strongly convex with $D := \operatorname{diam} \mathcal{C} > 0$. For $x \in \mathcal{C}$, a feasible target $q \in \mathcal{C}$, and a step $\eta \in [0,1]$, the exact geodesic update of Algorithm 1 is

$$z = \operatorname{Exp}_x(\eta \operatorname{Log}_x q) = \gamma_{xq}(\eta) \in \mathcal{C}, \tag{3}$$

with γ_{xq} the unique constant-speed minimizing geodesic from x to q (Assumption 2.1); z is feasible by geodesic convexity alone. An implementation, however, returns only stored data.

Definition 3.1 (represented endpoint)**.** The *represented endpoint* $\tilde{z} \in \mathcal{M}$ is the point a declared decoding map assigns to the data returned by an implementation of (3); the *endpoint error* is $d(\tilde{z}, z)$.

For example, on $\mathrm{SO}(3)$ the implementation returns a floating-point matrix with positive determinant, and $\tilde{z}$ is its polar factor. Every error bound below controls the endpoint error, including that of any chart remap or polar restoration. An accepted iterate is a represented endpoint, $x_{k+1} = \tilde{z}$.

Corollary 3.2 (endpoint rule)**.** *Let $x, q \in \mathcal{C}$, $\eta \in [0,1]$, z as in* (3)*, and define the* geometric budget $\beta(x,q,\eta) := \alpha\, \eta(1-\eta)\, d^2(x,q)$*. Then every $\tilde{z} \in \mathcal{M}$ with $d(\tilde{z}, z) \le \beta(x,q,\eta)$ lies in $\mathcal{C}$. The budget is a lower bound on* $\operatorname{dist}(z, \mathcal{M} \setminus \mathcal{C})$ *and vanishes at $\eta \in \{0,1\}$ and quadratically as $d(x,q) \to 0$.*

Corollary 3.2 follows by evaluating Definition 2.3 at the interpolated point (Figure 2). Its budget depends only on the chord data and the modulus, so it is available before z is computed. The

acceptance test, called the endpoint guard below, compares this budget with a computable upper bound on $d(\tilde{z}, z)$ supplied by the numerical map, an *enclosure* of the endpoint error; Appendix D constructs such a bound for chart Euler on rotations, that is, explicit Euler integration of the geodesic in the logarithmic chart at the ball centre. Rejection does not prove infeasibility, and the rule requires the represented target itself to be feasible.

Progress tolerance. Feasibility alone does not preserve the descent estimate. Once z and $\tilde{z}$ both lie in $\mathcal{C}$, the gradient bound $G := \max_{\mathcal{C}} \|\operatorname{grad} F\|$ gives $F(\tilde{z}) \le F(z) + G\, d(\tilde{z}, z)$, while the smoothness remainder along the exact chord is of order $\eta^2 d^2(x, q)$ (Section 4). A second tolerance, the *progress tolerance* $c\,\eta^2 d^2(x, q)$ with $c \in (0, \alpha]$, keeps the induced objective error $G\, d(\tilde{z}, z)$ at the same quadratic order and changes only the remainder's coefficient. The *accepted budget* is their minimum,

$$\beta_k := \min\{\alpha\eta_k(1 - \eta_k),\ c\,\eta_k^2\}\, d_k^2, \qquad d_k := d(x_k, q_k) \text{ with } q_k \text{ from (5)}. \tag{4}$$

For $\eta_k \le \frac{1}{2}$ the progress tolerance is no larger than the geometric budget because $c \le \alpha$, so a feasible candidate whose enclosure exceeds it is rejected even when the true clearance is much larger; both tolerances are sufficient, not necessary, conditions.

Inexact oracle. Fix $\delta \ge 0$. At iteration k the linear oracle returns a feasible $q_k \in \mathcal{C}$ with

$$\langle \operatorname{grad} F(x_k), \operatorname{Log}_{x_k} q_k \rangle_{x_k} \le \min_{q \in \mathcal{C}} \langle \operatorname{grad} F(x_k), \operatorname{Log}_{x_k} q \rangle_{x_k} + \delta_k, \qquad \delta_k \le \delta\, \underline{\eta}_k. \tag{5}$$

The step may be reduced before acceptance, so we distinguish the proposed step η_k^0, the accepted step η_k used in (4), and a lower bound $\underline{\eta}_k \le \eta_k$ known before the oracle call. The numerical realization (S) below never reduces the step, so $\underline{\eta}_k = \eta_k^0 = \eta_k$; the retraction realization (R) halves it, and $\underline{\eta}_k = \min\{\eta_k^0, \bar{\eta}_r/2\}$ with $\bar{\eta}_r$ the uniform step cap defined after (7). Lemma 3.3 ensures $\delta_k \le \delta\eta_k$ in either case.

Numerical realization (S). A numerical map is queried at refinement levels $j = 0, 1, \ldots$ and returns either a status (domain or primitive failure) or a candidate $\tilde{z}_j$ with a computable truncation term $E_j^{\mathrm{tr}} \ge 0$ and floor term $E_j^{\mathrm{fl}} \ge 0$ such that

$$d(\tilde{z}_j, z) \le E_j := E_j^{\mathrm{tr}} + E_j^{\mathrm{fl}}, \tag{6}$$

where $E_j^{\mathrm{tr}} \to 0$ if every trial returns a candidate, whereas E_j^{fl} bounds the remaining error (arithmetic, storage, restoration, and any nonvanishing analytic discrepancy) and does not necessarily converge to zero at fixed precision. The realization keeps the proposed step $\eta_k = \eta_k^0$ and refines the numerical map, increasing j, until the total-error test $E_j \le \beta_k$ passes. If the test fails while $E_j^{\mathrm{fl}} \ge (1 - \theta_{\mathrm{stop}})\beta_k$ for a fixed reserve fraction $\theta_{\mathrm{stop}} \in (0, 1)$, the loop stops refining (the *reserve exit*) and calls a fallback, an alternative evaluation of the same step with its own enclosure E^{ref}, accepted only if $E^{\mathrm{ref}} \le \beta_k$; otherwise the step ends with a precision request or the status *cannot certify*. While $E_j^{\mathrm{fl}} < (1 - \theta_{\mathrm{stop}})\beta_k$, at least $\theta_{\mathrm{stop}}\beta_k > 0$ of the budget remains for the vanishing component, which eventually fits (Lemma 3.3). The two branches that copy an existing point ($x_{k+1} = x_k$ when $d_k = 0$, $x_{k+1} = q_k$ when $\eta_k = 1$) have zero endpoint error despite a zero budget. Algorithm 2 implements this numerical realization.

Exactly evaluated retractions (R). Let $\mathcal{R}_x$ be a retraction evaluated in exact arithmetic, with $d(\mathcal{R}_x(v), \operatorname{Exp}_x v) \le K\|v\|_x^r$ for an exponent $r \ge 2$, all $x \in \mathcal{C}$, and $\|v\|_x \le r_0$, where $r_0 \in (0, D]$ and the bound must hold with one constant K on this whole domain (a sampled maximum only calibrates K). Then $\mathcal{R}_x(\eta \operatorname{Log}_x q) \in \mathcal{C}$ whenever

$$\eta\, d(x, q) \le r_0 \qquad \text{and} \qquad K\, \eta^{\,r-1}\, d(x, q)^{\,r-2} \le \alpha(1 - \eta) \tag{7}$$

(Corollary A.9). Realization (R) halves the proposed step until (7) holds and returns $\mathcal{R}_{x_k}(\eta_k v_k)$, where $v_k := \operatorname{Log}_{x_k} q_k$.

For $r = 2$ and $r_0 = D$, condition (7) is the distance-independent cap $\eta \le \alpha/(\alpha + K)$, worst-case sharp in the plane (Proposition C.1), and for any known $\widehat{D} \ge D$ both conditions hold for

Algorithm 2 Certified inexact Riemannian Frank–Wolfe, numerical realization (S). Algorithm 3 (Appendix D) also gives the separate retraction realization (R), the complete input contract (Table 4), and all return statuses.

Require: $x_0 \in \mathcal{C}$; $\alpha > 0$; $c \in (0, \alpha]$; steps $\eta_k \in (0, 1]$; an oracle satisfying (5); APPROXEXP(x, v, j) returning a candidate $(\tilde{z}_j, E_j^{\mathrm{tr}}, E_j^{\mathrm{fl}})$ satisfying (6) or a status; a reserve $\theta_{\mathrm{stop}} \in (0, 1)$; a FALLBACK returning $(\text{ACCEPTED}, \tilde{z}, E^{\mathrm{ref}})$ with $d(\tilde{z}, \mathrm{Exp}_{x_k}(\eta_k v_k)) \le E^{\mathrm{ref}}$, or a status

for $k = 0, 1, 2, \dots$ **do**
  $q_k \leftarrow$ oracle at x_k; $v_k \leftarrow \mathrm{Log}_{x_k} q_k$; $d_k \leftarrow \|v_k\|$
  if $d_k = 0$ **then** $x_{k+1} \leftarrow x_k$; **continue** $\{x_k, q_k$ coincide as represented points; Algorithm 3 requests precision if the chord cannot be resolved$\}$
  if $\eta_k = 1$ **then** $x_{k+1} \leftarrow q_k$; **continue** $\{$a full step returns the feasible target$\}$
  $\beta_k \leftarrow \min\{\alpha\eta_k(1 - \eta_k),\ c\,\eta_k^2\}\, d_k^2$
  for $j = 0, 1, 2, \dots$ **do**
    $r_j \leftarrow$ APPROXEXP$(x_k, \eta_k v_k, j)$; **if** r_j is a status **then return** (r_j, k) $\{$no new iterate$\}$
    $(\tilde{z}_j, E_j^{\mathrm{tr}}, E_j^{\mathrm{fl}}) \leftarrow r_j$; $E_j \leftarrow E_j^{\mathrm{tr}} + E_j^{\mathrm{fl}}$
    if $E_j \le \beta_k$ **then**
      $x_{k+1} \leftarrow \tilde{z}_j$; **break** $\{$accepted on the bound $E_j\}$
    else if $E_j^{\mathrm{fl}} \ge (1 - \theta_{\mathrm{stop}})\beta_k$ **then**
      $(s, \tilde{z}, E^{\mathrm{ref}}) \leftarrow$ FALLBACK$(x_k, q_k, \eta_k, \beta_k)$; **if** $s =$ ACCEPTED **and** $E^{\mathrm{ref}} \le \beta_k$ **then** $x_{k+1} \leftarrow \tilde{z}$; **break else return** (s', k) $\{s' = s$, or *cannot certify* if $s =$ ACCEPTED$\}$
    end if
  end for
end for

all q once $\eta \le \bar{\eta}_r := \sup\{\eta \in (0, 1] : K\eta^{r-1}\widehat{D}^{r-2} \le \alpha(1 - \eta),\ \eta\widehat{D} \le r_0\} > 0$. Under (7), $d(\mathcal{R}_x(\eta \mathrm{Log}_x q), z) \le K\eta^r d(x, q)^r \le KD^{r-2}\eta^2 d(x, q)^2$, so the endpoint error of (R) is a quadratic allowance with coefficient KD^{r-2}, not necessarily at most α; (R) applies no progress-tolerance test. A floating-point evaluation of a retraction instead belongs to (S), with its analytic discrepancy counted in the nonvanishing component E_j^{fl} and its arithmetic errors in the total-error enclosure.

Lemma 3.3 (finite backtracking and finite return; constants in Lemma A.7)**.** *In realization (R) with $d_k > 0$, the loop halves η_k finitely often, the accepted step satisfies $\eta_k \ge \underline{\eta}_k = \min\{\eta_k^0, \bar{\eta}_r/2\}$, and for $\eta_k^0 = 2/(k + 2)$ no halving occurs for $k \ge m_r - 2$, where $m_r := \min\{m \in \mathbb{N} : m \ge 2,\ 2/m \le \bar{\eta}_r\}$. In realization (S) with $d_k > 0$ and $0 < \eta_k < 1$, $\beta_k > 0$ and the loop of Algorithm 2 returns after finitely many trials, by acceptance, by an explicit status, or by a reserve exit. In both realizations every produced iterate lies in $\mathcal{C}$; a run that returns a status at step k has produced exactly the feasible iterates $x_0, \dots, x_k$, and a precision request resumes the same k.*

A status return preserves the guarantees of the iterates already produced but does not ensure that a new iterate is accepted. For chart Euler on $\mathrm{SO}(3)$, a fixed-sub-step policy certifies every step before the first ε-accurate iterate with at most $\max\{b_0, b_\varepsilon\}$ mantissa bits, where b_0 is the initial precision and $b_\varepsilon = O(\log(1/\varepsilon))$, under an exact feasible oracle, the uniform primitive-error model, the theorem's schedule, and a sufficient precision cap (Theorem D.5).

4 CONVERGENCE WITH CERTIFICATION ACTIVE

Let F be C^1 near $\mathcal{C}$ and geodesically convex and L_F-smooth on $\mathcal{C}$, let $x^\star \in \mathcal{C}$ be a minimizer, and write $F^\star = F(x^\star)$, $h_k := F(x_k) - F^\star$, $G = \max_{\mathcal{C}} \|\operatorname{grad} F\|$, $D = \operatorname{diam}\mathcal{C}$, and $g_k := -\langle \operatorname{grad} F(x_k), \mathrm{Log}_{x_k} q_k\rangle$ for the oracle's directional gap. Convexity and (5) give $g_k \ge h_k - \delta_k$; for an exact oracle g_k is the usual Frank–Wolfe gap, and for an inexact oracle it can be negative. The analysis separates two errors. The oracle suboptimality δ_k measures how well the feasible target minimizes the linear model, and the endpoint error $e_k := d(x_{k+1}, z_{k+1})$, where $z_{k+1} := \mathrm{Exp}_{x_k}(\eta_k v_k)$, measures how accurately the geodesic step is realized. Every accepted step of either realization obeys the one-step inequality (Appendix A)

$$h_{k+1} \le h_k - \eta_k g_k + \tfrac{L_F}{2}\eta_k^2 d_k^2 + G\, e_k, \qquad e_k \le \kappa\, \eta_k^2 d_k^2, \tag{8}$$

with $\kappa = c$ in (S) and $\kappa = KD^{r-2}$ in (R). The estimate separates the exact step from its realization. Smoothness along the feasible chord bounds $F(z_{k+1})$, the oracle inequality $g_k \geq h_k - \delta_k$ bounds the linear decrease $\eta_k g_k$ from below, and feasibility of both z_{k+1} and x_{k+1} lets the gradient bound G transfer the estimate between them along a geodesic in $\mathcal{C}$. The accepted endpoint error is quadratic in $\eta_k d_k$, so it only enlarges the quadratic coefficient from $\frac{L_F}{2}$ to $\frac{L_F}{2} + G\kappa$; the inequality bounds the objective suboptimality along the accepted sequence and does not require $F(x_{k+1}) \leq F(x_k)$ at every step.

Theorem 4.1 (certified descent and convergence)**.** *Let $\mathcal{C}$ satisfy Assumption 2.1 and be geodesically α-strongly convex, let F be as above, and run Algorithm 2 or its retraction realization (R) from $x_0 \in \mathcal{C}$ with a feasible oracle satisfying* (5)*. The statements concern the iterates produced before a status return and after each successful resumption; k counts accepted outer updates only. Put $Q_\kappa := (\frac{L_F}{2} + G\kappa)D^2$ with κ as in* (8)*.*

(i) Feasibility and descent. *Every produced iterate lies in $\mathcal{C}$, in both realizations, and every accepted step satisfies* (8)*, hence $h_{k+1} \leq (1-\eta_k)h_k + \eta_k\delta_k + (\frac{L_F}{2} + G\kappa)\eta_k^2 d_k^2$.*

(ii) Rate. *In (S) let $\eta_k = 2/(k+m)$ for an integer $m \geq 2$; in (R) let $\eta_k^0 = 2/(k+m_r)$ and $m = m_r$, so that no halving occurs. Then $h_k \leq \max\{m\, h_0,\, 4(Q_\kappa + \delta)\}/(k+m)$ for every produced iterate with index $k \geq 0$, and the backtracked schedule $\eta_k^0 = 2/(k+2)$ retains this order after the finite shift $m_r - 2$ (Proposition A.13).*

The proof turns (8) into the recurrence of (i) and applies a shifted Frank–Wolfe induction (Lemma A.11, Remark A.15). The rate concerns the objective gap h_k. A constrained minimizer with nonzero gradient is unique, and on an infinite accepted run satisfying the stated rate, $x_k \to x^\star$ by Lemma A.18. For the sublinear rate alone, an absolute allowance of order $\eta_k^2 D^2$ would suffice (Remark A.16); the chord-relative tolerance of (4) is retained because faster rates use more information. When the exact-oracle gap controls d_k^2, the error term $\kappa\eta_k^2 d_k^2$ can be absorbed into the gap itself.

Corollary 4.2 (fast open-loop rates are preserved under a scaling inequality)**.** *Under the hypotheses of Theorem 4.1 with an exact oracle ($\delta = 0$), suppose the oracle directions satisfy $g_k \geq \sigma d_k^2$ with $\sigma > 0$ for all k, and put $B := \frac{L_F}{2} + G\kappa > 0$. Then every accepted step with $\eta_k \leq \sigma/B$ satisfies $h_{k+1} \leq \big(1 - \eta_k + \frac{B}{\sigma}\eta_k^2\big)h_k$, and two rates follow. (i) For the proposed schedule $\eta_k^0 = a/(k+a)$ with a fixed $a \geq 2$, $h_k = O(k^{-a})$ on any run with infinitely many accepted steps, after a finite transient that covers the initial backtracking of (R) and the steps with $\eta_k > \sigma/B$; Proposition A.14 gives the transient index and the constant. (ii) For the constant proposed step $\eta_* := \min\{\frac{1}{2}, \sigma/(2B), \bar{\eta}_r\}$ ($\bar{\eta}_r$ omitted in (S)), $h_k \leq (1 - \eta_*/2)^k h_0$ for every produced iterate.*

The rate exponent is inherited from the exact-map analysis of Chen (2026), building on Euclidean open-loop analyses (Wirth et al., 2023; 2025); Euclidean linear-rate precedents are Levitin & Polyak (1966), Garber & Hazan (2015, Section 4.1), and Kerdreux et al. (2021). The new conclusion is that the rate survives a certified endpoint error of order $\eta_k^2 d_k^2$, which replaces $\frac{L_F}{2}$ by B in this recurrence. Theorem 5.1(ii) supplies this scaling inequality for exact oracles under Assumption 2.4 when the gradient norm is bounded away from zero. These guarantees need progressively stronger assumptions. Feasible targets and valid enclosures give feasibility, oracle accuracy and the schedule give the objective bound, and the precision result of Appendix D additionally assumes an exact feasible oracle and a primitive model to ensure that enough updates are accepted.

5 Geometry of strongly convex sets and faster rates

Corollary 4.2 requires a lower bound on the exact-oracle gap in terms of $d^2(x, q)$. **Our main geometric result**, Theorem 5.1, derives it from the strong convexity of $\mathcal{C}$ through a first-order margin, formulated in the *contingent cone*

$$T_{\mathcal{C}}(x) := \{v \in T_x\mathcal{M} : \liminf_{\tau \to 0^+} \operatorname{dist}(\operatorname{Exp}_x(\tau v), \mathcal{C})/\tau = 0\} \tag{9}$$

of first-order feasible directions at $x \in \mathcal{C}$ ($T_{\mathcal{C}}(x) = T_x\mathcal{M}$ at interior points). Its polar $T_{\mathcal{C}}(x)^\circ := \{n : \langle n, v\rangle_x \leq 0\ \forall v \in T_{\mathcal{C}}(x)\}$ is the set of outward normals.

Theorem 5.1 (intrinsic margins and exact-oracle scaling)**.** *Let $\mathcal{C}$ satisfy Assumption 2.1 and be geodesically α-strongly convex.*

(i) Intrinsic margin. *For all $p, q \in \mathcal{C}$,*

$$\big\{\mathrm{Log}_p\, q + \alpha d^2(p,q)\, w :\ w \in T_p\mathcal{M},\ \|w\|_p \le 1\big\} \subseteq T_{\mathcal{C}}(p), \tag{10}$$

and consequently, for every $n \in T_{\mathcal{C}}(p)^\circ$,

$$\langle n, \mathrm{Log}_p\, q\rangle_p \le -\alpha\, d^2(p,q)\, \|n\|_p. \tag{11}$$

(ii) Exact-oracle scaling. *Assume additionally Assumption 2.4 with $S = \mathcal{C}$, let F be C^1 near $\mathcal{C}$, and let $m_{\log} := \min_{x,q\in\mathcal{C}} \sigma_{\min}(D_q \mathrm{Log}_x)$, where $D_q \mathrm{Log}_x$ is the differential of $y \mapsto \mathrm{Log}_x\, y$ at q and $\sigma_{\min}$ denotes its smallest singular value (positive, Appendix A). For $x \in \mathcal{C}$ and an exact minimizer q of $\langle \mathrm{grad}\, F(x), \mathrm{Log}_x \cdot\rangle_x$ over $\mathcal{C}$,*

$$-\langle \mathrm{grad}\, F(x), \mathrm{Log}_x\, q\rangle_x \ \ge\ \alpha\, m_{\log}\, \|\,\mathrm{grad}\, F(x)\|_x\, d^2(x,q).$$

If $\nu := \min_{\mathcal{C}} \|\,\mathrm{grad}\, F\| > 0$, the scaling inequality of Corollary 4.2 holds with $\sigma = \alpha m_{\log} \nu$.

Part (i) is a first-order version of Definition 2.3. At any $p \in \mathcal{C}$, the direction toward another feasible point q points strictly into $\mathcal{C}$, with a margin that grows as $d^2(p,q)$; no curvature bound or boundary regularity is needed, and (11) is the intrinsic counterpart of the normal-cone characterization of Euclidean strong convexity (Vial, 1983; Goncharov & Ivanov, 2017; Remark C.3). Part (ii) applies this margin at the oracle target, where the negative gradient of the linear model is an outward normal; unlike the general intrinsic bound of Scieur et al. (2026, Proposition 3.3), the resulting scaling inequality has no additive residual. On the rotation balls of Proposition 5.2, $m_{\log} = 1$, and $\nu > 0$ holds when F is geodesically convex and its constrained minimizer has a nonzero gradient (Lemma A.18).

Geodesic balls on $\mathrm{SO}(3)$. Consider rotations within angle ρ of a reference R_c. We equip $\mathrm{SO}(3)$ with the bi-invariant metric $\langle X, Y\rangle = \frac{1}{2}\,\mathrm{tr}(X^\top Y)$, for which d is the rotation angle, the sectional curvature is $\frac{1}{4}$, and the universal cover is the sphere $S^3(2)$ of radius 2. For $\rho < \pi/2$, the closed ball $\mathcal{C}_\rho := \overline{B}(R_c, \rho)$ has diameter below π, and its lift to $S^3(2)$ is isometric to $\mathcal{C}_\rho$; the ball is uniquely geodesically convex and satisfies Assumption 2.4 with $m_{\log} = 1$ (Appendix A).

Proposition 5.2 (sharp modulus and exact clearance on $\mathrm{SO}(3)$)**.** *Let $C_\rho := \overline{B}(R_c, \rho) \subset \mathrm{SO}(3)$ under the metric of Section 5, with $0 < \rho < \pi/2$, and set $\alpha_\rho := \frac{1}{4}\cot(\rho/2)$.*

(i) $\mathcal{C}_\rho$ is geodesically α_ρ-strongly convex, and no $\alpha > \alpha_\rho$ satisfies (2) for $\mathcal{C}_\rho$.

(ii) For every geodesic segment γ_{xy} of $\mathcal{C}_\rho$ with $d(x,y) = \ell$ and every $\eta \in [0,1]$,

$$\rho - d(\gamma_{xy}(\eta), R_c) \ge \Delta_\rho(\eta, \ell) := \rho - 2\arccos\Big[\frac{\cos(\rho/2)\cos((1-2\eta)\ell/4)}{\cos(\ell/4)}\Big] \ge \alpha_\rho\, \eta(1-\eta)\, \ell^2,$$

with equality in the first inequality, for $0 < \eta < 1$, if and only if $x, y \in \partial\mathcal{C}_\rho$; hence no larger tolerance depending only on (ρ, η, ℓ) is valid.

The proof in Appendix A lifts each chord to the spherical cover. Its inner product with the lifted centre satisfies a second-order linear equation, which yields the worst-case clearance. Boundary-to-boundary chords attain this clearance, and short boundary chords establish sharpness of the modulus.

On C_ρ, the sharper clearance replaces the geometric budget in Corollary 3.2, giving $\min\{\Delta_\rho(\eta_k, d_k),\ c\eta_k^2 d_k^2\}$ in (4). Acceptance is unchanged when the progress tolerance binds (Section 6). Chen (2026, Proposition 5.5) obtains the rescaled counterpart of α_ρ as an exact-oracle scaling coefficient on spherical balls; here it is also the sharp intrinsic set modulus. Proposition A.19 gives the boundary-normal margin for every boundary point and target in the ball, together with the small-radius expansion showing that the Euclidean value $1/(2\rho)$ is not a valid modulus. For sublevel-set constraints, Theorem A.20 establishes strong convexity under ambient-domain and stationary-minimizer assumptions; Proposition A.22 shows why restricted-domain assumptions alone are insufficient.

6 Experiments

The experiments test agreement with the proved geometry, native feasibility, and the effect of the endpoint guard on accuracy and work; Table 1 in Appendix D lists the studies with their sample

units, horizons, and evidence status. A *native proposal* is the output of a numerical kernel before any guard or fallback. A trajectory is reported as violating when an accepted iterate lies more than 10^{-10} rad outside the ball, and native proposals are classified by the sign of their clearance recomputed in 400-bit arithmetic; these values are consistency checks, not certificates. The primitive bounds of Assumption D.1 remain hypotheses, and, except in the exact-oracle instance, oracle solutions and reference optima are numerical approximations. Timings are indicative and are not used to rank methods; no runtime advantage is claimed.

The optimization studies estimate a constrained Fréchet mean of rotations. Given N_s attitudes R_i, they minimize $F(R) = (2N_s)^{-1} \sum_{i=1}^{N_s} d(R, R_i)^2$ over a ball, usually of radius $60°$. The reported gap $F(R_k) - F_{\mathrm{ref}}$ uses a numerical reference optimum, except in Experiment F, where the optimum is known analytically. The nonradial family, so called because its iterates do not stay on one geodesic through the ball centre (Appendix D), draws 50 attitudes from three noisy clusters of sizes 25, 15, and 10 at two separations, giving boundary-active and interior reference minimizers. Ten seeds, two separations, and two starts give 40 trajectories per treatment, each run for 300 iterations; the starts are at the centre and near the boundary, and confidence intervals resample the 10 seeds, the four trajectories of a seed forming one bundle. The kernels are chart Euler, fp16 or fp32 storage of the float64 endpoint with polar restoration, and QR.

Feasibility does not ensure accuracy (H). The unguarded and membership-guarded runs report no trajectory violation ($0/40$ per kernel). Unguarded fp16 storage nevertheless increased the final $\log_{10}$ gap relative to the float64 reference run by a paired mean of 1.74 in the boundary-active regime, with a 95% seed-bundle bootstrap interval of $[1.67, 1.81]$; the interior estimate is -0.086, with interval $[-0.204, 0.003]$, showing no clear degradation. In-ball membership guards never intervened on their own trajectories and left the fp16 loss unchanged, whereas the endpoint-guarded fp16 method recovered the reference run's accuracy (Figure 1(a), Table 13).

What recovers the accuracy (H). Across both regimes, the endpoint-guarded fp16 method reused the already computed float64 endpoint, each time under its own enclosure, on 97.3% of its steps, and the endpoint-guarded fp32 method accepted 6.45% of its boundary-active proposals natively. These two methods and unguarded fp32 storage stayed close to the reference run, with paired mean differences in final $\log_{10}$ gap below $5 \cdot 10^{-3}$ in magnitude in both regimes (largest magnitude $4.5 \cdot 10^{-3}$; Table 13). This establishes neither low-precision savings nor an advantage over a fixed fp32 policy. The oracle dominates the measured online time (Table 16), and the timings were collected under concurrent load.

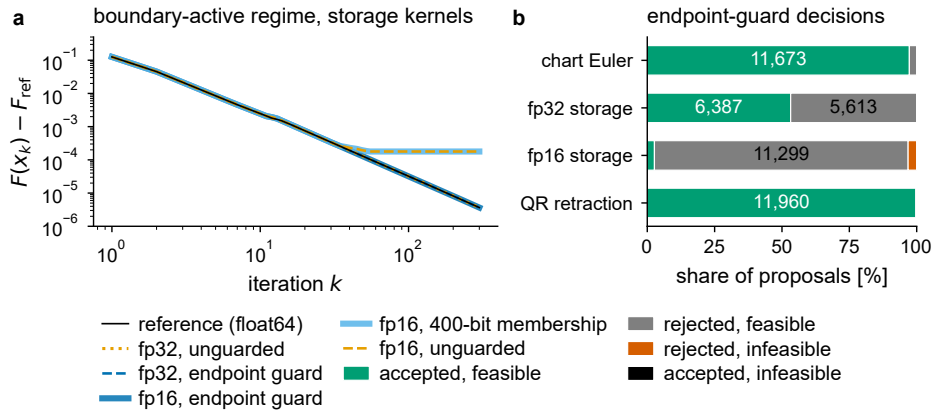

Figure 1: Accuracy and guard decisions in Experiment H. (a) Median gap per iteration over the 20 boundary-active trajectories for the storage kernels. The fp16 kernel stalls, with or without a membership guard, unless endpoint-guarded (accurate fallback), whereas fp32 tracks the float64 reference with or without a guard. (b) Endpoint-guard decisions, across both regimes, on the $12{,}000$ native proposals recorded along each kernel's endpoint-guarded trajectories, classified by the sign of the 400-bit clearance (same candidates); segments above 14% are labelled (all counts in Table 14).

Decisions on the same proposals (H). Because the guards generate different trajectories, we also compare both rules on the same native proposals from endpoint-guarded runs. Among the 12,000 fp16 proposals, the 400-bit membership guard and the endpoint guard both reject all 382 infeasible proposals. The endpoint guard additionally rejects 11,299 feasible fp16 proposals, and 327 Euler, 5,613 fp32, and 40 QR proposals, all feasible (Figure 1(b), Table 14); these additional rejections reflect its total-error acceptance requirement and show its substantial conservatism.

Where the conservatism comes from (H). With $c = \alpha$ and the ideal schedule $\eta_k = 2/(k+3)$, the geometric budget divided by the progress tolerance is $(1-\eta_k)/\eta_k = (k+1)/2$; the implemented ratio uses the represented step and differs only by schedule rounding. The progress tolerance is the smaller term from $k = 2$ on, that is, on 298 of the 300 iterations, with a median ratio of about 75 (Table 15). The endpoint clearance of Remark C.2, evaluated offline at the float64 reference endpoint (Appendix D), exceeds the generic geometric budget by median factors of about 10^2 (interior) to 10^6 (boundary-active). Replacing that budget by this clearance, with the progress tolerance retained, leaves the Euler and storage acceptance counts unchanged and raises QR acceptances only from 11,960 to 11,980. A second diagnostic replaces the enclosure by the 400-bit endpoint error against the same reference; it accepts every recorded Euler proposal but raises boundary-active fp16 acceptances only from 73 to 148 of the 6,000 proposals. These diagnostics motivate evaluating the sublinear-only allowance of Remark A.16 and less restrictive conditions that also preserve the faster rates.

Native feasibility near the boundary (G). On a declared grid of 216 one-step configurations whose starting points have clearance between 10^{-2} and 10^{-4} rad, fp16 storage followed by polar restoration produced 10 infeasible outputs, all at the smallest clearance, with endpoint error up to 4.3 times the clearance of the reference endpoint (Appendix D). The endpoint guard, using the storage-error enclosure of Appendix D, rejected all 216 fp16 proposals, including all 206 feasible outputs. The fp32, one-step Euler, and QR outputs stayed feasible (Table 8).

Agreement with the geometric formulas (A, B). For six radii up to 89° the minimum sampled values of the boundary-normal ratio $-\langle n, \operatorname{Log}_p q\rangle_p / d(p,q)^2$ in (11), for unit n, lie within $3 \cdot 10^{-6}$ of the sharp modulus α_ρ, and no sampled chord violates Proposition 5.2 (Tables 6–7; Figure 3(d) illustrates Proposition A.22).

Precision continuation and cost (D, E). On the 1,500-iteration continuation study the float64 (53-bit) routine returns *cannot certify* at $k = 448$, whereas the economical adaptive policy (Proposition D.4) completes the horizon with at most 67 mantissa bits and 2.0 right-hand-side evaluations per iteration; the unguarded float64 controls remain feasible and reach the same final gap as the adaptive policy, so the demonstrated benefit is continued conditional certification, not rescued feasibility. The fixed-N policy of Theorem D.5 ($N = 256$) costs about 259 evaluations per iteration in the cost study, whose numerical oracle and schedule $\eta_k = 2/(k+3)$ lie outside the theorem's hypotheses.

The theorem-aligned instance (F). The axis-invariant instance has an exact represented oracle, satisfies the theorem's schedule conditions, and has zero Euler truncation. The run reached $\varepsilon = 10^{-3}$, 10^{-5}, 10^{-7}, and 10^{-9} at $k = 31$, 329, 3,310, and 33,113, using at most 53, 54, 74, and 94 mantissa bits against the sufficient thresholds $b_\varepsilon = 66$, 92, 119, and 146, and every decoded iterate stayed feasible (Table 25, Figure 8). Over each of the intervals from 10^{-5} to 10^{-7} and from 10^{-7} to 10^{-9}, the observed maximum precision increased by 20 bits, a finite-range observation consistent with logarithmic growth.

7 Conclusion and limitations

Strong convexity of the feasible set turns an endpoint-error enclosure into a feasibility guarantee, and a quadratic endpoint-error allowance preserves the Frank–Wolfe convergence bounds along accepted updates, under feasible-oracle and valid-enclosure contracts. In the nonradial experiments, unguarded fp16 iterates stayed feasible yet lost objective accuracy near the boundary, which the endpoint rule recovered; the progress tolerance binds on almost every iteration, so the rule rejects many feasible proposals and relies on an accurate fallback, without a demonstrated runtime advantage. Next steps are tighter validated enclosures, certified oracles, and a less restrictive progress tolerance.

AI USE STATEMENT

We used ChatGPT (GPT 5.6-Sol and GPT 6-Astra) to assist with language polishing, proof checking, implementation of the numerical experiments, and reference formatting. The authors take responsibility for the final content of this work, including text, claims, code, and artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

The appendices provide proofs of the theoretical results and describe the numerical contracts, experimental protocols, parameter choices, random-number streams and seeds, statistical procedures, and reported results. Main-text proofs are collected in Appendix A; the numerical analysis and experimental details, including the represented inputs of the exact-oracle instance, are given in Appendix D. The implementation guarantees remain conditional on the primitive-error assumptions stated there. Except in Experiment F, the linear-oracle solutions and reference optima are high-accuracy numerical approximations, not certified global optima; Experiment F establishes an exact feasible oracle and optimum for its represented problem. The reported 400-bit checks provide numerical consistency evidence and are distinct from outward-rounded interval certification. After completion of the review process, we will release the manuscript on arXiv and make the experiment code, dependency versions, configurations, seeds, raw results, verification scripts and tests, and the table- and figure-generation scripts publicly available through GitHub.

REFERENCES

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

Aaron D. Ames, Xiangru Xu, Jessy W. Grizzle, and Paulo Tabuada. Control barrier function based quadratic programs for safety critical systems. *IEEE Transactions on Automatic Control*, 62(8): 3861–3876, 2017.

Aaron D. Ames, Samuel Coogan, Magnus Egerstedt, Gennaro Notomista, Koushil Sreenath, and Paulo Tabuada. Control barrier functions: theory and applications. In *European Control Conference (ECC)*, pp. 3420–3431, 2019.

Jean-Pierre Aubin and Arrigo Cellina. *Differential Inclusions: Set-Valued Maps and Viability Theory*. Springer, 1984.

Jean-Pierre Aubin and Hélène Frankowska. *Set-Valued Analysis*. Birkhäuser, 1990.

Franco Blanchini. Set invariance in control. *Automatica*, 35(11):1747–1767, 1999.

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

Kangming Chen. Open-loop Riemannian Frank–Wolfe: fast rates under error bounds and scaling inequalities, 2026. arXiv:2608.21598.

Kangming Chen and Ellen H. Fukuda. Riemannian conditional gradient methods for composite optimization problems. *Computational and Applied Mathematics*, 45:151, 2026. doi: 10.1007/s40314-025-03496-1. arXiv:2412.19427.

Francis H. Clarke, Yuri S. Ledyaev, Ronald J. Stern, and Peter R. Wolenski. *Nonsmooth Analysis and Control Theory*. Springer, 1998.

Christopher Criscitiello and Nicolas Boumal. Negative curvature obstructs acceleration for strongly geodesically convex optimization, even with exact first-order oracles. In *Proceedings of the 35th Conference on Learning Theory (COLT)*, volume 178 of *Proceedings of Machine Learning Research*, pp. 496–542, 2022.

Damek Davis, Dmitriy Drusvyatskiy, and Zhan Shi. Stochastic optimization over proximally smooth sets. *SIAM Journal on Optimization*, 35(1):157–179, 2025.

Massimiliano de Sa, Pio Ong, and Aaron D. Ames. From bundles to backstepping: geometric control barrier functions for safety-critical control on manifolds. In *2026 American Control Conference (ACC)*, pp. 1928–1935, 2026. arXiv:2510.20202.

Mateo Díaz, Benjamin Grimmer, and Ian McPherson. Nonsmooth Riemannian optimization with inexact manifold primitives via bundle methods, 2026. arXiv:2604.27078.

Hélène Frankowska, Sławomir Plaskacz, and Tadeusz Rzeżuchowski. Measurable viability theorems and the Hamilton–Jacobi–Bellman equation. *Journal of Differential Equations*, 116(2):265–305, 1995. doi: 10.1006/jdeq.1995.1036.

Dan Garber and Elad Hazan. Faster rates for the Frank–Wolfe method over strongly-convex sets. In *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, volume 37 of *Proceedings of Machine Learning Research*, pp. 541–549, 2015.

Vladimir V. Goncharov and Grigorii E. Ivanov. Strong and weak convexity of closed sets in a Hilbert space. In *Operations Research, Engineering, and Cyber Security: Trends in Applied Mathematics and Technology*, volume 113 of *Springer Optimization and Its Applications*, pp. 259–297. Springer, 2017. doi: 10.1007/978-3-319-51500-7_12.

Richard Hartley, Jochen Trumpf, Yuchao Dai, and Hongdong Li. Rotation averaging. *International Journal of Computer Vision*, 103(3):267–305, 2013.

Adrian Hauswirth, Saverio Bolognani, and Florian Dörfler. Projected dynamical systems on irregular, non-Euclidean domains for nonlinear optimization. *SIAM Journal on Control and Optimization*, 59(1):635–668, 2021.

Martin Jaggi. Revisiting Frank–Wolfe: projection-free sparse convex optimization. In *Proceedings of the 30th International Conference on Machine Learning (ICML)*, volume 28 of *Proceedings of Machine Learning Research*, pp. 427–435, 2013.

Michel Journée, Yurii Nesterov, Peter Richtárik, and Rodolphe Sepulchre. Generalized power method for sparse principal component analysis. *Journal of Machine Learning Research*, 11(15):517–553, 2010.

Thomas Kerdreux, Alexandre d'Aspremont, and Sebastian Pokutta. Projection-free optimization on uniformly convex sets. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 130 of *Proceedings of Machine Learning Research*, pp. 19–27, 2021.

Evgeny S. Levitin and Boris T. Polyak. Constrained minimization methods. *USSR Computational Mathematics and Mathematical Physics*, 6(5):1–50, 1966. doi: 10.1016/0041-5553(66)90114-5.

Yuchen Li, Laura Balzano, Deanna Needell, and Hanbaek Lyu. Convergence and complexity guarantee for inexact first-order Riemannian optimization algorithms. In *Proceedings of the 41st International Conference on Machine Learning (ICML)*, volume 235 of *Proceedings of Machine Learning Research*, pp. 27376–27398, 2024.

Aaron Lou, Derek Lim, Isay Katsman, Leo Huang, Qingxuan Jiang, Ser-Nam Lim, and Christopher De Sa. Neural manifold ordinary differential equations. In *Advances in Neural Information Processing Systems 33 (NeurIPS)*, pp. 17548–17558, 2020.

David Martínez-Rubio and Sebastian Pokutta. Accelerated Riemannian optimization: handling constraints with a prox to bound geometric penalties. In *Proceedings of the 36th Conference on Learning Theory (COLT)*, volume 195 of *Proceedings of Machine Learning Research*, pp. 359–393, 2023.

Emile Mathieu and Maximilian Nickel. Riemannian continuous normalizing flows. In *Advances in Neural Information Processing Systems 33 (NeurIPS)*, pp. 2503–2515, 2020.

Mitio Nagumo. Über die Lage der Integralkurven gewöhnlicher Differentialgleichungen. *Proceedings of the Physico-Mathematical Society of Japan, 3rd Series*, 24:551–559, 1942. doi: 10.11429/ppmsj1919.24.0_551.

Siegfried M. Rump. Verification methods: rigorous results using floating-point arithmetic. *Acta Numerica*, 19:287–449, 2010.

Damien Scieur, David Martínez-Rubio, Thomas Kerdreux, Alexandre d'Aspremont, and Sebastian Pokutta. Strong convexity of sets in Riemannian manifolds, 2024. arXiv:2312.03583, version 3, 13 December 2024, https://arxiv.org/abs/2312.03583v3.

Damien Scieur, David Martínez-Rubio, Thomas Kerdreux, Alexandre d'Aspremont, and Sebastian Pokutta. Strongly convex sets in Riemannian manifolds. In *The Fourteenth International Conference on Learning Representations (ICLR)*, pp. 23433–23468, 2026. URL https://openreview.net/forum?id=4sDszSYKP6.

Jean-Philippe Vial. Strong and weak convexity of sets and functions. *Mathematics of Operations Research*, 8(2):231–259, 1983.

Dongming Wang and Wei Ren. Geodesic strong convexity does not imply forward invariance under gradient flow on SO(3): a certified counterexample, 2026. arXiv:2608.28976.

Melanie Weber and Suvrit Sra. Projection-free nonconvex stochastic optimization on Riemannian manifolds. *IMA Journal of Numerical Analysis*, 42(4):3241–3271, 2022.

Melanie Weber and Suvrit Sra. Riemannian optimization via Frank–Wolfe methods. *Mathematical Programming*, 199:525–556, 2023.

Elias Wirth, Thomas Kerdreux, and Sebastian Pokutta. Acceleration of Frank–Wolfe algorithms with open-loop step-sizes. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 206 of *Proceedings of Machine Learning Research*, pp. 77–100, 2023.

Elias Wirth, Javier Peña, and Sebastian Pokutta. Accelerated affine-invariant convergence rates of the Frank–Wolfe algorithm with open-loop step-sizes. *Mathematical Programming*, 214:201–245, 2025. doi: 10.1007/s10107-024-02180-2.

Wei Xiao, Tsun-Hsuan Wang, Ramin Hasani, Mathias Lechner, Yutong Ban, Chuang Gan, and Daniela Rus. On the forward invariance of neural ODEs. In *Proceedings of the 40th International Conference on Machine Learning (ICML)*, volume 202 of *Proceedings of Machine Learning Research*, pp. 38100–38124, 2023.

Hongyi Zhang and Suvrit Sra. First-order methods for geodesically convex optimization. In *Proceedings of the 29th Conference on Learning Theory (COLT)*, volume 49 of *Proceedings of Machine Learning Research*, pp. 1617–1638, 2016.

CONTENTS OF THE APPENDICES

- Appendix A: proofs of the main-text results, with the supporting lemmas on logarithmic images and tangent cones, the explicit fast-rate constants, the rotation-ball margin and small-radius expansion, the ambient-tube theorem, and the restricted-domain counterexample.
- Appendix B (optional): the continuous-time forward-invariance analogue of the margin; no main-text result depends on it.
- Appendix C: notation and terminology, the certificate schematic (Figure 2), the sharpness of the retraction cap, exact clearances on rotation balls, comparisons of certification routes, and remarks on retractions, charts, projections, and equality constraints.
- Appendix D: represented quantities and the primitive-error model, the chart-Euler enclosure and storage ledger, the acceptance machinery (contract table and operational routine), the precision policies and the sufficient-precision theorem, the shared experimental protocol, and Experiments A–H.

A PROOFS

This appendix proves the main-text results and states the auxiliary lemmas used in their proofs. Section A.1 compares the forms of set strong convexity and establishes the logarithmic-image and tangent-cone lemmas; Section A.2 proves the intrinsic margins; Sections A.3–A.4 analyse certified updates and convergence, including the explicit fast-rate constants and a sublinear-only relaxation of the progress tolerance; Sections A.5–A.8 give the rotation-ball, ambient-tube, restricted-domain, and oracle arguments. Main-text statements keep their numbers.

A.1 VARIANTS OF SET STRONG CONVEXITY, LOGARITHMIC IMAGES, AND TANGENT CONES

Throughout the appendices, $A_p := \mathrm{Log}_p(\mathcal{C}) \subset T_p\mathcal{M}$ denotes the *logarithmic image* of $\mathcal{C}$ at $p \in \mathcal{C}$, with Log_p the restricted logarithm of Assumption 2.1.

Definition A.1 (tangent-ball and logarithmic-image forms; Scieur et al., 2026)**.** Let $\mathcal{C}$ satisfy Assumption 2.1, let $\alpha > 0$, and for $x, y \in \mathcal{C}$ let $\gamma\colon [0,1] \to \mathcal{C}$ be the minimizing geodesic from x to y, parametrized with constant speed. Each item quantifies over all $x, y \in \mathcal{C}$ and $t \in [0,1]$; (a) denotes Definition 2.3.

(b) $\mathcal{C}$ is *double geodesically α-strongly convex* if $\mathrm{Exp}_{\gamma(t)}(e) \in \mathcal{C}$ for all $e \in T_{\gamma(t)}\mathcal{M}$ with $\|e\|_{\gamma(t)} \le \alpha t(1-t)d^2(x,y)$.

(c) $\mathcal{C}$ is *Riemannian α-strongly convex* if, for every $p \in \mathcal{C}$, the logarithmic image A_p is α-strongly convex in $T_p\mathcal{M}$. Explicitly, for all $u, w \in A_p$, $t \in [0,1]$, and $\zeta \in T_p\mathcal{M}$ with $\|\zeta\|_p \le 1$, we require $tu + (1-t)w + \alpha t(1-t)\|u-w\|_p^2\zeta \in A_p$.

In (a) the perturbation is a metric ball about the chord point, in (b) a tangent ball at the chord point, and in (c) a tangent ball at the base point p, with the chord length measured by $\|u-w\|_p$ instead of $d(x,y)$ (a comparison distance with constants $\ell_{\mathcal{M}} \le L_{\mathcal{M}}$ is allowed by Scieur et al., 2026, and is set to 1 here; Remark C.5).

Proposition A.2 (hierarchy of the notions)**.** *Under Assumption 2.1, (a) and (b) are equivalent with the same modulus, and (c) implies (b), hence (a), with the same modulus, on every complete manifold. Whether (a) implies (c) in general is not asserted here.*

This agrees with Scieur et al. (2026) at $\ell_{\mathcal{M}} = L_{\mathcal{M}} = 1$: their Proposition 2.7 (Proposition 4.3 of Scieur et al., 2024) is the equivalence of (a) and (b), and their Proposition 2.6 (Proposition 4.2 of the preprint) is the implication from (c); we give a direct proof that does not use curvature comparison.

Proof. $\mathcal{C}$ satisfies Assumption 2.1; $\mathcal{M}$ is complete.

(a)⇔(b). If (a) holds and $\|e\|_{\gamma(t)} \le r := \alpha t(1-t)d^2(x,y)$, then $d(\gamma(t), \mathrm{Exp}_{\gamma(t)} e) \le \|e\| \le r$ (the curve $\tau \mapsto \mathrm{Exp}_{\gamma(t)}(\tau e)$ has length $\|e\|$), so $\mathrm{Exp}_{\gamma(t)} e \in \mathcal{C}$, which proves (b). If (b) holds and $d(\gamma(t), \tilde{z}) \le r$, then by Hopf–Rinow there is a minimizing geodesic from $\gamma(t)$ to $\tilde{z}$ with initial

velocity e, $\|e\| = d(\gamma(t), \tilde{z}) \le r$, and $\tilde{z} = \operatorname{Exp}_{\gamma(t)} e \in \mathcal{C}$, which proves (a). Neither direction uses injectivity of $\operatorname{Exp}_{\gamma(t)}$; the modulus is unchanged.

(c)⇒(b). Let $x, y \in \mathcal{C}$, $t \in [0,1]$, $p := \gamma(t)$, $V := \dot\gamma(t)$, so $\|V\|_p = d(x,y)$. The restrictions $\gamma|_{[t,1]}$ and $\gamma|_{[0,t]}$ are minimizing geodesics from p to y and to x (subsegments of minimizing geodesics are minimizing), and by Assumption 2.1 they are the unique ones; hence $u := \operatorname{Log}_p y = (1-t)V$ and $w := \operatorname{Log}_p x = -tV$, both in A_p. Then $tu + (1-t)w = 0$ and $\|u - w\|_p = \|V\|_p = d(x,y)$. Definition A.1(c) at p with these u, w, t gives $\alpha t(1-t)d^2(x,y)\,\zeta \in A_p$ for every $\|\zeta\|_p \le 1$, i.e., for every $e \in T_p\mathcal{M}$ with $\|e\|_p \le \alpha t(1-t)d^2(x,y)$ there is $y \in \mathcal{C}$ with $\operatorname{Log}_p y = e$, and then $\operatorname{Exp}_p(e) = y \in \mathcal{C}$. This is (b) with the same modulus. No curvature comparison enters. □

The following statements, used in Proposition A.6, in the proofs of Theorem 5.1(ii), Lemma A.18, and Corollary A.5, and in Appendix B, relate the logarithmic image to the contingent cone under Assumption 2.1 alone.

Lemma A.3 (logarithmic image and cones)**.** *Let $\mathcal{C}$ satisfy Assumption 2.1 and fix $p \in \mathcal{C}$. Then (i) the restricted logarithm $\operatorname{Log}_p\colon \mathcal{C} \to T_p\mathcal{M}$ is continuous and $A_p = \operatorname{Log}_p(\mathcal{C})$ is compact; (ii) $T_{\mathcal{C}}(p) = T_{A_p}(0)$, the contingent cone of $A_p \subset T_p\mathcal{M}$ at the origin (defined by the liminf condition of* (9) *in $T_p\mathcal{M}$, equivalently as $\{\lim a_j/\tau_j : a_j \in A_p,\ \tau_j \to 0^+\}$ because A_p is closed); (iii) $p \in \partial\mathcal{C}$ iff $0 \in \partial A_p$; (iv) if A_p is convex, then $T_{A_p}(0)$ is the closed conic hull of A_p, $T_{\mathcal{C}}(p)^\circ = N_{A_p}(0) =: N_{\mathcal{C}}(p)$, and $T_{\mathcal{C}}(p) = \{v : \langle n, v\rangle_p \le 0\ \forall n \in N_{\mathcal{C}}(p)\}$; (v) the map $(x,y) \mapsto \operatorname{Log}_x y$ is continuous on $\mathcal{C} \times \mathcal{C}$.*

Proof. (i) If $y_n \to y$ in $\mathcal{C}$, the initial velocities $\operatorname{Log}_p y_n$ have norms $d(p, y_n)$, hence are bounded; any limit point is the initial velocity of a minimizing geodesic from p to y, which is unique, so $\operatorname{Log}_p y_n \to \operatorname{Log}_p y$. Compactness follows. (ii)–(iii) For ε smaller than the injectivity radius at p, the minimizing geodesic from p to any $y \in B(p,\varepsilon)$ is radial, so $\{v \in A_p : \|v\| < \varepsilon\} = \operatorname{Log}_p(\mathcal{C} \cap B(p,\varepsilon)) = (\operatorname{Exp}_p|_{B(0,\varepsilon)})^{-1}(\mathcal{C} \cap B(p,\varepsilon))$. Thus, near p, the sets $\mathcal{C}$ and A_p correspond under a diffeomorphism that fixes the base point and has identity differential. This correspondence preserves their local contingent cones and boundaries. Part (iv) follows from standard convex analysis. For a closed convex set, the contingent cone at a point is the closed conic hull of the set translated by that point. Its polar is the normal cone, and the bipolar theorem gives the last description. (v) If $(x_j, y_j) \to (x, y)$ in $\mathcal{C} \times \mathcal{C}$, the initial velocities $\operatorname{Log}_{x_j} y_j$ have norms $d(x_j, y_j)$, hence are bounded in $T\mathcal{M}$; any limit point (x, v) is the initial velocity of a curve $\tau \mapsto \operatorname{Exp}_x(\tau v)$ of length $\|v\| = d(x,y)$ from x to y (continuity of Exp on $T\mathcal{M}$ and of d), i.e., of a minimizing geodesic, which is unique by Assumption 2.1; so $\operatorname{Log}_{x_j} y_j \to \operatorname{Log}_x y$. □

Lemma A.4 (convexity and derivability of the tangent cone)**.** *Let $\mathcal{C}$ satisfy Assumption 2.1 and $p \in \mathcal{C}$. Then $A_p = \operatorname{Log}_p(\mathcal{C})$ is star-shaped about the origin, $T_{\mathcal{C}}(p) = \overline{\operatorname{cone} A_p}$ with $\operatorname{cone} A_p := \{\lambda a : \lambda \ge 0,\ a \in A_p\}$, and $T_{\mathcal{C}}(p)$ is a closed convex cone. Moreover $\lim_{t\to 0^+} \operatorname{dist}(\operatorname{Exp}_p(tv), \mathcal{C})/t = 0$ for every $v \in T_{\mathcal{C}}(p)$ ($\mathcal{C}$ is* derivable *at p), so $T_{\mathcal{C}}(p) = (T_{\mathcal{C}}(p)^\circ)^\circ$ and, on the product manifold $\mathcal{M}^N$ with the product metric, $T_{\mathcal{C}^N}(x_1, \dots, x_N) = \prod_i T_{\mathcal{C}}(x_i)$.*

Proof. Geodesic convexity gives $t \operatorname{Log}_p q = \operatorname{Log}_p(\gamma_{pq}(t)) \in A_p$ for $0 \le t \le 1$ (subsegments of the unique minimizing geodesic are the unique minimizing geodesics between their endpoints), so A_p is star-shaped. For a closed star-shaped set $S \ni 0$ in a Euclidean space, the contingent cone at the origin is $\overline{\operatorname{cone} S}$. Indeed, each term a_j/τ_j of a tangent sequence lies in $\operatorname{cone} S$. Conversely, each $\lambda a \in \operatorname{cone} S$ is tangent because $t\lambda a \in S$ for all sufficiently small $t > 0$. More strongly, for $v \in \overline{\operatorname{cone} S}$ and $\varepsilon > 0$ pick λa with $\|v - \lambda a\| < \varepsilon$; then $\operatorname{dist}(tv, S)/t \le \|v - \lambda a\| < \varepsilon$ for all small t, so the quotient tends to 0. Lemma A.3(ii) transfers the first statement to $\mathcal{C}$ at p; the second transfers because Exp_p is bi-Lipschitz on $B(0,\varepsilon)$ with constants tending to 1 as $\varepsilon \to 0$, and the nearest points to tv in A_p lie within $2t\|v\|$ of 0. To prove convexity, fix $v, w \in T_{\mathcal{C}}(p)$ and $\theta \in [0,1]$. The limit established above supplies $a_t, b_t \in \mathcal{C}$ whose normal coordinates at p are $tv + o(t)$ and $tw + o(t)$; for small t they lie in a common convex normal neighbourhood, their geodesic interpolation lies in $\mathcal{C}$, and smooth dependence of geodesic interpolation on its endpoints gives it the normal coordinate $t[(1-\theta)v + \theta w] + o(t)$; hence $(1-\theta)v + \theta w \in T_{\mathcal{C}}(p)$. The set defined by (9) is closed (if $v_n \to v$ then $\liminf_\tau \operatorname{dist}(\operatorname{Exp}_p(\tau v), \mathcal{C})/\tau \le \Lambda\|v - v_n\|$ for every n, with Λ a local Lipschitz constant of

Exp_p), and the bipolar identity follows. For the product, the squared distance to $\mathcal{C}^N$ is the sum of the squared component distances, so the ordinary limits give $\prod_i T_{\mathcal{C}}(x_i) \subseteq T_{\mathcal{C}^N}(x)$, while projection onto each factor gives the reverse inclusion. □

Corollary A.5 (margin in tangent form)**.** *Let $\mathcal{C}$ satisfy Assumption 2.1 and be geodesically α-strongly convex. For $p, q \in \mathcal{C}$ and $e \in T_p\mathcal{M}$, the direction $\mathrm{Log}_p q + e$ lies in $T_{\mathcal{C}}(p)$ whenever $\|e\|_p \le \alpha d^2(p,q)$. By Lemma A.4, it also suffices that $\langle n, e\rangle_p \le \alpha d^2(p,q)$ for all unit $n \in T_{\mathcal{C}}(p)^\circ$. This condition controls only the positive pairings of the error with outward polar normals. The proof appears in Section A.3, and Appendix B uses this form for the continuous-time certificates.*

Proposition A.6 (chart form and membership test)**.** *Let $\mathcal{C}$ satisfy Assumption 2.1 and be Riemannian α-strongly convex (Definition A.1(c)). Then for every $p \in \partial\mathcal{C}$, unit $n \in N_{\mathcal{C}}(p)$, and $q \in \mathcal{C}$, $\langle n, \mathrm{Log}_p q\rangle_p \le -\alpha d^2(p,q)$, and $T_{\mathcal{C}}(p) = \{v : \langle n, v\rangle_p \le 0 \ \forall n \in N_{\mathcal{C}}(p)\}$.*

Proof. By Lemma A.3(i),(iii), A_p is compact with $0 \in \partial A_p$, and by Definition A.1(c) it is α-strongly convex, hence convex; the Euclidean normal-cone characterization of strong convexity (Vial, 1983; Goncharov & Ivanov, 2017), $\langle w, u - x\rangle \ge \alpha\|w\|\|u - x\|^2$ for $u \in \partial A$, $w \in N_A(u)$, $x \in A$, applied with $u = 0$, $x = \mathrm{Log}_p q$, $w = n$, gives the inequality; the cone description is Lemma A.3(iv). □

Proposition A.6 is a chart-based special case of Theorem 5.1 (Proposition A.2). Assumption 2.1 alone makes the contingent cone closed and convex (Lemma A.4), so the intrinsic polar characterization and the multi-anchor certificates of Appendix B do not require convexity of the full logarithmic image. The paper distinguishes three normal-cone constructions. The chart normal cone $N_{\mathcal{C}}(p)$ is defined in $T_p\mathcal{M}$ and used only when A_p is convex. The intrinsic polar cone $T_{\mathcal{C}}(p)^\circ$ of Theorem 5.1 does not require convexity of A_p, and it equals $N_{\mathcal{C}}(p)$ when A_p is convex. The proximal normal cone belongs to the theory of prox-regular sets discussed in Remark C.9.

A.2 Proof of Theorem 5.1 (intrinsic margins and exact-oracle scaling)

Proof of part (i). $\mathcal{C}$ satisfies Assumption 2.1 and is geodesically α-strongly convex; $\mathcal{M}$ is complete, so Exp_p is defined and smooth on all of $T_p\mathcal{M}$. The case $q = p$ is trivial (the set in (10) is $\{0\} \subseteq T_{\mathcal{C}}(p)$). Let $\ell := d(p,q) > 0$, let γ be the minimizing geodesic from p to q, fix w with $\|w\|_p \le 1$, and set $v := \mathrm{Log}_p q + \alpha\ell^2 w$. For $t \in (0,1]$ define

$$\Lambda_t := \max\big\{1,\ \sup\{\|D\,\mathrm{Exp}_p(u)\| : \|u\|_p \le t(\ell + \alpha\ell^2)\}\big\}, \qquad v_t := \mathrm{Log}_p q + (1-t)\,\frac{\alpha\ell^2}{\Lambda_t}\,w,$$

where $\|D\,\mathrm{Exp}_p(u)\|$ denotes the operator norm from $(T_p\mathcal{M}, \|\cdot\|_p)$ to $(T_{\mathrm{Exp}_p u}\mathcal{M}, \|\cdot\|)$. We use only the scalar function $u \mapsto \|D\,\mathrm{Exp}_p(u)\|$, which is continuous and therefore bounded on the compact ball. The points $t\,\mathrm{Log}_p q$, tv_t, and tv have norm at most $t(\ell + \alpha\ell^2)$, so the segments joining them lie in that ball, and the mean-value inequality (the length of the image of a segment under Exp_p is at most Λ_t times its length) gives

$$d\big(\mathrm{Exp}_p(tv_t), \gamma(t)\big) = d\big(\mathrm{Exp}_p(tv_t), \mathrm{Exp}_p(t\,\mathrm{Log}_p q)\big) \le \Lambda_t\,\|tv_t - t\,\mathrm{Log}_p q\|_p \le \alpha t(1-t)\ell^2.$$

By Definition 2.3, $\mathrm{Exp}_p(tv_t) \in \mathcal{C}$. Since Exp_p is smooth with $D\,\mathrm{Exp}_p(0) = \mathrm{id}$, $\Lambda_t \to 1$ as $t \to 0^+$, hence $\Lambda_t - (1-t) = (\Lambda_t - 1) + t \to 0$ and

$$\mathrm{dist}\big(\mathrm{Exp}_p(tv), \mathcal{C}\big) \le d\big(\mathrm{Exp}_p(tv), \mathrm{Exp}_p(tv_t)\big) \le \Lambda_t\,t\,\|v - v_t\|_p \le t\,\alpha\ell^2\big(\Lambda_t - (1-t)\big) = o(t),$$

which is the $\liminf$ condition in (9); thus $v \in T_{\mathcal{C}}(p)$. For $n \in T_{\mathcal{C}}(p)^\circ \setminus \{0\}$, choosing $w = n/\|n\|_p$ gives $\langle n, \mathrm{Log}_p q\rangle + \alpha\ell^2\|n\| \le 0$, which is (11). For $n = 0$ the inequality holds trivially. □

Proof of part (ii). Under Assumption 2.4, $(x,y) \mapsto \mathrm{Log}_x y$ is smooth on a neighbourhood of $\mathcal{C} \times \mathcal{C}$ and $D_q\,\mathrm{Log}_x = (D\,\mathrm{Exp}_x|_{\mathrm{Log}_x q})^{-1}$ is invertible, so $\sigma_{\min}(D_q\,\mathrm{Log}_x)$ is continuous and positive on the compact set $\mathcal{C} \times \mathcal{C}$ (including $x = q$, where it equals 1); hence $m_{\log} > 0$. Fix $x \in \mathcal{C}$, put $\nabla := \mathrm{grad}\,F(x)$, $g(x,q) := -\langle \nabla, \mathrm{Log}_x q\rangle_x$, and $\phi(y) := \langle \nabla, \mathrm{Log}_x y\rangle_x$, a C^1 function near $\mathcal{C}$ with $\mathrm{grad}\,\phi(q) = (D_q\,\mathrm{Log}_x)^*\nabla =: w$. We prove the pointwise bound

$$g(x,q) \ \ge\ \alpha\,d(x,q)^2\,\|w\|_q \ \ge\ \alpha\,m_{\log}\,\|\nabla\|_x\,d(x,q)^2,$$

whose two inequalities combine to the bound stated in the theorem; the first is the sharper pointwise adjoint form. A minimizer q of ϕ over $\mathcal{C}$ exists because ϕ is continuous (Lemma A.3(v)) and $\mathcal{C}$ is compact. If $\nabla = 0$, then $g(x,q) = 0$ and both right-hand sides vanish; if $q = x$, then $d(x,q) = g(x,q) = 0$; the claims are trivial in either case, so assume $\nabla \neq 0$ and $q \neq x$. Since q minimizes ϕ over the closed set $\mathcal{C}$, every $v \in T_{\mathcal{C}}(q)$ satisfies $\langle w, v\rangle_q \geq 0$. Indeed, (9) supplies $\tau_j \to 0$ and $y_j \in \mathcal{C}$ with $d(\mathrm{Exp}_q(\tau_j v), y_j) = o(\tau_j)$, and consequently $0 \leq \phi(y_j) - \phi(q) = \tau_j \langle w, v\rangle_q + o(\tau_j)$. Thus $-w \in T_{\mathcal{C}}(q)^\circ$ and part (i) at $p = q$ with anchor x gives $\langle -w, \mathrm{Log}_q x\rangle_q \leq -\alpha d(x,q)^2 \|w\|_q$, i.e., $\langle w, \mathrm{Log}_q x\rangle_q \geq \alpha d(x,q)^2 \|w\|_q$. Along the reversed minimizing geodesic $\gamma(t) = \mathrm{Exp}_q(t \,\mathrm{Log}_q x)$ one has $\mathrm{Log}_x \gamma(t) = (1-t)\,\mathrm{Log}_x q$ (uniqueness of minimizing geodesics, Assumption 2.1), so $D_q \,\mathrm{Log}_x[\mathrm{Log}_q x] = -\,\mathrm{Log}_x q$ and $\langle w, \mathrm{Log}_q x\rangle_q = \langle \nabla, D_q \,\mathrm{Log}_x[\mathrm{Log}_q x]\rangle_x = -\langle \nabla, \mathrm{Log}_x q\rangle_x = g(x,q)$. Finally $\|w\|_q \geq \sigma_{\min}(D_q \,\mathrm{Log}_x)\|\nabla\|_x \geq m_{\log}\|\nabla\|_x$. On $\mathrm{SO}(3)$ with the metric of Section 5, $D\,\mathrm{Exp}_x$ at v with $\|v\| = \ell < \pi$ has singular values 1 (radial) and $\sin(\ell/2)/(\ell/2) \leq 1$ (transverse) (cf. Remark C.8; Jacobi fields as in Boumal, 2023, Section 10.8), so its inverse has smallest singular value 1 and $m_{\log} = 1$ on $\mathcal{C}_\rho$, whose chords have length below π. If $\nu > 0$, the bound just proved gives $g(x,q) \geq \alpha m_{\log} \nu \, d(x,q)^2$ at every exact target q, which is the scaling inequality of Corollary 4.2 with $\sigma = \alpha m_{\log} \nu$. For the remark following the theorem, if $x^\star$ is a constrained minimizer with $\mathrm{grad}\, F(x^\star) \neq 0$, Lemma A.18(ii) makes $x^\star$ the unique minimizer of F over $\mathcal{C}$; a point $y \in \mathcal{C}$ with $\mathrm{grad}\, F(y) = 0$ would be a global minimizer of the geodesically convex F over $\mathcal{C}$, forcing $y = x^\star$, a contradiction; hence $\|\,\mathrm{grad}\, F\| > 0$ on the compact $\mathcal{C}$ and $\nu > 0$. The argument concerns an exact minimizer q and does not give a scaling inequality for arbitrary or inexact targets. □

A.3 Proofs of Corollaries 3.2, A.9, A.8, and A.5, and of Lemmas 3.3 and A.7

Lemma A.7 (finite return, quantitative form)**.** *Under the hypotheses of Lemma 3.3, in realization (R) with $d_k > 0$ the loop halves η_k at most $\lceil \log_2(\eta_k^0/\bar{\eta}_r)\rceil^+$ times, where $\lceil u\rceil^+ := \max\{\lceil u\rceil, 0\}$, and if $r = 2$ and $r_0 = \widehat{D} = D$, then m_2 ($:= m_r$ for $r = 2$) $= \lceil 2(\alpha + K)/\alpha\rceil$. In realization (S), fix $\theta \in (\theta_{\mathrm{stop}}, 1]$ and let $j^*(\theta)$ be the first index with $E_{j^*}^{\mathrm{tr}} \leq \theta\beta_k$; if every trial through j^* returns a candidate and $E_j^{\mathrm{fl}} \leq (1-\theta)\beta_k$ for every $j \leq j^*$, the loop accepts at or before j^*.*

Corollary A.8 (approximate maps with feasible inexact targets)**.** *If $\tilde{z}$ comes with a verifiable bound $d(\tilde{z}, z) \leq \varepsilon$ (a series remainder, an a priori integrator enclosure, a rounding bound), then $\tilde{z} \in \mathcal{C}$ as soon as $\varepsilon \leq \beta(x,q,\eta)$; for random error this holds pathwise on the event $\{d(\tilde{z}, z) \leq \beta\}$, with no trajectory-level guarantee asserted. Any feasible $q \in \mathcal{C}$ is admissible. Its suboptimality contributes no separate feasibility error, although the choice of q changes the chord, the tolerance, and the work; oracle suboptimality enters the convergence analysis through Theorem 4.1, and a gradient evaluated with error at most ξ contributes oracle error at most $2\xi D$ (Lemma A.10). A represented target outside $\mathcal{C}$ is not covered.*

Corollary A.9 (retractions with an analytic map-error bound)**.** *Let $\mathcal{R}_x \colon T_x\mathcal{M} \to \mathcal{M}$ be a retraction with $d(\mathcal{R}_x(v), \mathrm{Exp}_x(v)) \leq K\|v\|_x^r$ for all $x \in \mathcal{C}$ and $\|v\|_x \leq r_0$, where $K \geq 0$, the* endpoint-error exponent *$r \geq 2$, and the certified radius $r_0 \in (0, D]$ are fixed. Then $\mathcal{R}_x(\eta\,\mathrm{Log}_x q) \in \mathcal{C}$ whenever $\eta\, d(x,q) \leq r_0$ and $K\,\eta^{r-1}\, d(x,q)^{r-2} \leq \alpha(1-\eta)$, that is, whenever* (7) *holds. For $r = 2$ and $r_0 = D$ this is the distance-independent cap $\eta \leq \alpha/(\alpha + K)$; for any known $\widehat{D} \geq D$, both conditions hold for all q once $\eta \leq \bar{\eta}_r := \sup\{\eta \in (0,1] : K\eta^{r-1}\widehat{D}^{r-2} \leq \alpha(1-\eta),\ \eta\widehat{D} \leq r_0\} > 0$; the admissible set is the interval $(0, \bar{\eta}_r]$, so m_r of Lemma 3.3 is the smallest integer $m \geq 2$ with $K(2/m)^{r-1}\widehat{D}^{r-2} \leq \alpha(1-2/m)$ and $2\widehat{D}/m \leq r_0$.*

Proof. $\mathcal{C}$ satisfies Assumption 2.1 and is geodesically α-strongly convex. *Corollary 3.2.* $z = \gamma_{xq}(\eta)$ with γ_{xq} the minimizing geodesic from x to q, so Definition 2.3 at $t = \eta$ says exactly that the closed ball of radius $\beta(x,q,\eta)$ about z lies in $\mathcal{C}$, hence $\mathrm{dist}(z, \mathcal{M} \setminus \mathcal{C}) \geq \beta(x,q,\eta)$; the tangent form is Definition A.1(b), equivalent by Proposition A.2. The stated vanishing limits follow directly from the formula for β.

Corollary A.9. If $\eta = 0$ then $\mathcal{R}_x(0) = x \in \mathcal{C}$; assume $\eta > 0$. With $v = \eta\,\mathrm{Log}_x q$, $\|v\| = \eta d(x,q) \leq r_0$ by the step condition, so $d(\mathcal{R}_x(v), z) \leq K\eta^r d(x,q)^r$. If $d(x,q) = 0$ then $v = 0$ and $\mathcal{R}_x(0) = x \in \mathcal{C}$. Otherwise $K\eta^r d^r \leq \alpha\eta(1-\eta)d^2$ is equivalent to (7), and Corollary 3.2 applies. For $r = 2$, (7) reads $K\eta \leq \alpha(1-\eta)$, i.e., $\eta \leq \alpha/(\alpha + K)$. For general r, $d(x,q) \leq D$ shows that $K\eta^{r-1}D^{r-2} \leq \alpha(1-\eta)$ implies (7) and $\eta D \leq r_0$ implies $\eta d(x,q) \leq r_0$; the left sides are

nondecreasing in η and tend to 0 as $\eta \to 0$ while $\alpha(1-\eta) \to \alpha > 0$ and $r_0 > 0$, so $\bar{\eta}_r > 0$ and both conditions hold for all $\eta \le \bar{\eta}_r$.

Corollary A.8 is Corollary 3.2 with $\varepsilon \le \beta$, applied pathwise on the stated event in the random-error case; the statement about feasible targets is geodesic convexity.

Corollary A.5. For the norm condition, if $p = q$ then $e = 0$ and $\mathrm{Log}_p q + e = 0 \in T_{\mathcal{C}}(p)$; if $p \neq q$, apply (10) with $w = e/(\alpha d^2(p,q))$. For the normal-component condition, Lemma A.4 makes $T_{\mathcal{C}}(p)$ closed and convex, so by the bipolar theorem it suffices that $\langle n, \mathrm{Log}_p q + e\rangle_p \le 0$ for every unit $n \in T_{\mathcal{C}}(p)^\circ$; by (11) and the assumed bound, $\langle n, \mathrm{Log}_p q + e\rangle_p \le -\alpha d^2(p,q) + \langle n, e\rangle_p \le 0$. If the polar cone is $\{0\}$ the condition is vacuous and $T_{\mathcal{C}}(p) = T_p\mathcal{M}$, so the conclusion also holds.

Lemmas 3.3 and A.7. Fix k with $d_k > 0$. Both acceptance conditions, $\eta d_k \le r_0$ and $K\eta^{r-1} d_k^{r-2} \le \alpha(1-\eta)$, have nondecreasing left sides and nonincreasing right sides in η, and both hold for every $\eta \le \bar{\eta}_r$ because $d_k \le D \le \widehat{D}$ and the global caps $K\eta^{r-1}\widehat{D}^{r-2} \le \alpha(1-\eta)$, $\eta\widehat{D} \le r_0$ hold there. Starting from $\eta_k^0 \le 1$, after j halvings the trial step is $2^{-j}\eta_k^0 \le \bar{\eta}_r$ as soon as $j \ge \log_2(\eta_k^0/\bar{\eta}_r)$, so the loop performs at most $\lceil \log_2(\eta_k^0/\bar{\eta}_r)\rceil^+$ halvings and one final test, and none when $\eta_k^0 \le \bar{\eta}_r$. For the open-loop schedule, m_r is finite because the conditions $K(2/m)^{r-1}\widehat{D}^{r-2} \le \alpha(1-2/m)$ and $2\widehat{D}/m \le r_0$ hold for all large m, and for $k \ge m_r - 2$ we have $2/(k+2) \le 2/m_r$, at which the condition holds by definition of m_r and hence, by monotonicity, at $2/(k+2)$ too. For $r = 2$ the condition reads $2(K+\alpha)/m \le \alpha$, i.e., $m \ge 2(\alpha+K)/\alpha$. If halving occurs, the last rejected trial $2\eta_k$ exceeds $\bar{\eta}_r$ (the test holds for every $\eta \le \bar{\eta}_r$), so $\eta_k > \bar{\eta}_r/2$. Otherwise $\eta_k = \eta_k^0$. In both cases $\eta_k \ge \underline{\eta}_k$, and $\delta_k \le \delta\underline{\eta}_k \le \delta\eta_k$. In realization (S), $d_k > 0$ and $0 < \eta_k < 1$ give $\beta_k > 0$. If a trial returns a domain or primitive-failure status, the algorithm returns at that finite trial. Otherwise every trial under consideration returns a candidate. If neither acceptance nor a reserve exit occurs, then $E_j^{\mathrm{fl}} < (1-\theta_{\mathrm{stop}})\beta_k$ for every j, while the candidate-interface hypothesis gives $E_j^{\mathrm{tr}} \to 0$; some finite j then has $E_j^{\mathrm{tr}} \le \theta_{\mathrm{stop}}\beta_k$, hence $E_j \le \beta_k$ and the loop accepts (acceptance is tested first, so equality cases are accepted), a contradiction. Thus the loop returns finitely. This argument uses only the declared interface. A positive reserve is necessary for this finite-return guarantee. To see why, suppose $\theta_{\mathrm{stop}} = 0$, a value excluded by Algorithm 2, and consider the valid conservative enclosure $E_j^{\mathrm{fl}} = \beta_k(1-2^{-j-1})$, $E_j^{\mathrm{tr}} = 2^{-j}\beta_k$. Then $E_j = \beta_k(1+2^{-j-1}) > \beta_k$ and $E_j^{\mathrm{fl}} < \beta_k$ for every j, so the loop neither accepts nor exits. If $E_j^{\mathrm{fl}} \le (1-\theta)\beta_k$ with $\theta > \theta_{\mathrm{stop}}$ for all j up to the first index j^* with $E_{j^*}^{\mathrm{tr}} \le \theta\beta_k$, the reserve exit cannot trigger before j^* and the loop accepts at or before j^*. An accepted point satisfies $d(\tilde{z}_j, \mathrm{Exp}_{x_k}(\eta_k v_k)) \le E_j \le \beta_k$ by the bound, hence lies in $\mathcal{C}$ by Corollary 3.2; a fallback point is accepted only with its own enclosure $E^{\mathrm{ref}} \le \beta_k$, and a run that cannot certify stops before producing an uncertified iterate. The acceptance test uses the error enclosure and does not require evaluation of $d(\tilde{z}_j, \mathrm{Exp}_{x_k}(\eta_k v_k))$. □

A.4 Proof of Theorem 4.1 (convergence) and the fast-rate constants

Lemma A.10 (inexact gradients as oracle error)**.** *Let $g := \mathrm{grad}\, F(x)$, $\hat{g} \in T_x\mathcal{M}$ with $\|\hat{g} - g\| \le \xi$, and let $\hat{q}$ minimize $\langle \hat{g}, \mathrm{Log}_x q\rangle$ over $q \in \mathcal{C}$. Then $\langle g, \mathrm{Log}_x \hat{q}\rangle \le \min_{q\in\mathcal{C}}\langle g, \mathrm{Log}_x q\rangle + 2\xi D$.*

Proof. For any $q \in \mathcal{C}$, $|\langle \hat{g} - g, \mathrm{Log}_x q\rangle| \le \xi d(x,q) \le \xi D$. Hence $\langle g, \mathrm{Log}_x \hat{q}\rangle \le \langle \hat{g}, \mathrm{Log}_x \hat{q}\rangle + \xi D \le \langle \hat{g}, \mathrm{Log}_x q\rangle + \xi D \le \langle g, \mathrm{Log}_x q\rangle + 2\xi D$ for every q. □

Lemma A.11 (shifted Frank–Wolfe recursion)**.** *Let $h_k \ge 0$ satisfy $h_{k+1} \le (1-\eta_k)h_k + \eta_k^2\Gamma$ with $\Gamma \ge 0$ and $\eta_k = 2/(k+m)$, $m \ge 2$. Then $h_k \le A/(k+m)$ for all $k \ge 0$ with $A := \max\{m\, h_0, 4\Gamma\}$. If $m = 2$ then in addition $h_k \le 4\Gamma/(k+2)$ for all $k \ge 1$.*

Proof. We argue by induction. The definition of A gives $h_0 \le A/m$. If $h_k \le A/(k+m)$, then the recurrence gives

$$h_{k+1} \le \Big(1 - \frac{2}{k+m}\Big)\frac{A}{k+m} + \frac{4\Gamma}{(k+m)^2} \le \frac{A(k+m-2)+A}{(k+m)^2} = \frac{A(k+m-1)}{(k+m)^2} \le \frac{A}{k+m+1},$$

since $(k+m-1)(k+m+1) = (k+m)^2 - 1$. For $m = 2$, $\eta_0 = 1$ gives $h_1 \le \Gamma \le 4\Gamma/3$, and the same induction from $k = 1$ with $A = 4\Gamma$ gives $h_k \le 4\Gamma/(k+2)$ for $k \ge 1$. □

Lemma A.12 (rounded schedule)**.** *Let $h_k \geq 0$ satisfy $h_{k+1} \leq (1 - T_k)h_k + CT_k^2$ with $C \geq 0$ and represented steps $T_k = \frac{2}{k+m}(1+\vartheta_k)$, $|\vartheta_k| \leq u_s < \frac{1}{2}$, $m \geq 3$. Then $h_k \leq A_s/(k+m)$ for all $k \geq 0$ with $A_s := \max\{m\, h_0,\, 4C(1+u_s)^2/(1-2u_s)\}$.*

Proof. Write $a_- := 2(1-u_s)$ and $a_+ := 2(1+u_s)$, so that $h_{k+1} \leq (1 - \frac{a_-}{k+m})h_k + \frac{Ca_+^2}{(k+m)^2}$ (the factor $1 - T_k \geq 0$ because $T_k \leq a_+/m < 1$ for $m \geq 3$). We again argue by induction. The definition of A_s gives $h_0 \leq A_s/m$. If $h_k \leq A_s/(k+m)$, the recurrence yields $h_{k+1} \leq \frac{A_s(k+m-a_-)+Ca_+^2}{(k+m)^2} \leq \frac{A_s(k+m-1)}{(k+m)^2} \leq \frac{A_s}{k+m+1}$, using $A_s(a_- - 1) \geq Ca_+^2$, which is the definition of A_s since $a_- - 1 = 1 - 2u_s$. □

The lemma is applied in Appendix D to the binary steps actually used; with an oracle satisfying its tolerance relative to the accepted T_k, $C = Q_\kappa + \delta$ with $\kappa = c$.

Proof of Theorem 4.1. Throughout the proof, $\mathcal{C}$ satisfies Assumption 2.1 and is geodesically α-strongly convex. The function F is C^1 near $\mathcal{C}$, and is geodesically convex and L_F-smooth on $\mathcal{C}$ in the sense of Definition 2.2. Compactness gives $G = \max_{\mathcal{C}} \|\operatorname{grad} F\| < \infty$, and we write $D = \operatorname{diam}\mathcal{C}$. The nonnegative oracle tolerances δ_k, δ satisfy $\delta_k \leq \delta\eta_k \leq \delta\eta_k$ at every accepted step by Lemma 3.3.

(i) If $d_k = 0$ then $x_{k+1} = x_k \in \mathcal{C}$. Otherwise $z_{k+1} = \gamma_{x_k q_k}(\eta_k) \in \mathcal{C}$ by Assumption 2.1. In realization (R), the accepted η_k satisfies the test of Corollary A.9, which gives $x_{k+1} \in \mathcal{C}$. In realization (S), the accepted point satisfies $d(x_{k+1}, z_{k+1}) \leq E_j \leq \beta_k \leq \alpha\eta_k(1-\eta_k)d_k^2$, or $d(x_{k+1}, z_{k+1}) \leq E^{\mathrm{ref}} \leq \beta_k$ for a fallback point, so Corollary 3.2 gives $x_{k+1} \in \mathcal{C}$. When $\eta_k = 1$, the update is $x_{k+1} = q_k \in \mathcal{C}$.

Descent along the geodesic. Write $\nabla_k := \operatorname{grad} F(x_k)$ and $g_k = -\langle \nabla_k, \operatorname{Log}_{x_k} q_k\rangle$. If $d_k = 0$ then $q_k = x_k$, $g_k = 0$, the oracle guarantee gives $0 \leq \min_q \langle \nabla_k, \operatorname{Log}_{x_k} q\rangle + \delta_k$, and convexity toward $x^\star$ gives $h_k \leq -\langle \nabla_k, \operatorname{Log}_{x_k} x^\star\rangle \leq \delta_k$, so $h_{k+1} = h_k = (1-\eta_k)h_k + \eta_k h_k \leq (1-\eta_k)h_k + \eta_k\delta_k$ (using $h_k \leq \delta_k$, with η_k the scheduled step, which Algorithm 3 assigns as $\eta_k \leftarrow \eta_k^0$ before this branch), which is the recursion (14) below without the error term, and (8) holds trivially with $e_k = 0$. Otherwise, since $z_{k+1} \in \mathcal{C}$ and $\operatorname{Log}_{x_k} z_{k+1} = \eta_k \operatorname{Log}_{x_k} q_k$ (unique minimizing geodesic), L_F-smoothness on $\mathcal{C}$ gives

$$F(z_{k+1}) \leq F(x_k) - \eta_k g_k + \tfrac{L_F}{2}\eta_k^2 d_k^2. \tag{12}$$

By convexity at x_k toward $x^\star$ and the oracle guarantee, $F^\star \geq F(x_k) + \langle \nabla_k, \operatorname{Log}_{x_k} x^\star\rangle \geq F(x_k) - g_k - \delta_k$, so $g_k \geq h_k - \delta_k$ and

$$F(z_{k+1}) - F^\star \leq (1-\eta_k)h_k + \eta_k\delta_k + \tfrac{L_F}{2}\eta_k^2 d_k^2 \leq (1-\eta_k)h_k + \eta_k\delta_k + \tfrac{L_F}{2}\eta_k^2 D^2. \tag{13}$$

From z_{k+1} to x_{k+1}. Both points lie in $\mathcal{C}$, so the minimizing geodesic γ_k between them lies in $\mathcal{C}$, where $\|\operatorname{grad} F\| \leq G$; hence $F(x_{k+1}) - F(z_{k+1}) = \int_0^1 \langle \operatorname{grad} F(\gamma_k(t)), \dot\gamma_k(t)\rangle\, dt \leq G\, e_k$, $e_k = d(x_{k+1}, z_{k+1})$. Adding this to (12) gives the one-step inequality (8); in realization (S), $e_k \leq \beta_k \leq c\eta_k^2 d_k^2$, and in realization (R), $e_k \leq K\eta_k^r d_k^r \leq KD^{r-2}\eta_k^2 d_k^2$ (using $\eta_k \leq 1$, $d_k \leq D$), which are the two values of κ. Thus, whenever $e_k \leq \kappa\,\eta_k^2 d_k^2 \leq \kappa\,\eta_k^2 D^2$,

$$h_{k+1} \leq (1-\eta_k)h_k + \eta_k^2\Big(\big(\tfrac{L_F}{2} + G\kappa\big)D^2 + \delta\Big), \tag{14}$$

valid for every $\eta_k \in (0, 1]$. The recurrence displayed in (i) is the intermediate form $h_{k+1} \leq (1-\eta_k)h_k + \eta_k\delta_k + (\frac{L_F}{2} + G\kappa)\eta_k^2 d_k^2$ before $\delta_k \leq \delta\eta_k$ and $d_k \leq D$ are applied.

(ii) In realization (S), (14) holds with $\kappa = c$; in realization (R) with $\eta_k^0 = 2/(k+m_r)$, no halving occurs by Lemma 3.3 and (14) holds with $\kappa = KD^{r-2}$. In both cases Lemma A.11 with $\Gamma = Q_\kappa + \delta$ and the shift m (respectively $m = m_r$) gives the bound. □

Proposition A.13 (retraction schedules)**.** *Under the hypotheses of Theorem 4.1 in realization (R), write $Q_K := Q_\kappa$ with $\kappa = KD^{r-2}$. The following bounds hold. (i) If $\eta_k^0 = 2/(k+m_r)$, no halving occurs and $h_k \leq \max\{m_r h_0, 4(Q_K+\delta)\}/(k+m_r)$ for all $k \geq 0$. (ii) If the open-loop proposal is $\eta_k^0 = 2/(k+2)$, halving occurs only for $k < m_r - 2$, and $h_k \leq \max\{m_r \max\{h_0, Q_K+\delta\}, 4(Q_K+\delta)\}/(k+2)$ for $k \geq m_r - 2$.*

Proof. (i) is Theorem 4.1(ii) in realization (R). For (ii), every accepted step satisfies (14) with $\kappa = KD^{r-2}$ and some $\eta_k \in (0,1]$, hence $h_{k+1} \le (1-\eta_k)h_k + \eta_k\Gamma \le \max\{h_k, \Gamma\}$ with $\Gamma = Q_K + \delta$, and $h_k \le \max\{h_0, \Gamma\}$ for all k. By Lemma 3.3, for $k \ge k_0 := m_r - 2$ the accepted step is $\eta_k = 2/(k+2) = 2/((k-k_0)+m_r)$. Applying Lemma A.11 to $h'_j := h_{k_0+j}$ with shift m_r gives $h_k \le \max\{m_r h_{k_0}, 4\Gamma\}/(k+2) \le \max\{m_r \max\{h_0, \Gamma\}, 4\Gamma\}/(k+2)$ for $k \ge k_0$. □

Proposition A.14 (fast rates with explicit transients)**.** *Under the hypotheses of Corollary 4.2, let $\eta_k^0 = a/(k+a)$ with $a \ge 2$, let $k_R := \lceil a(\bar\eta_r^{-1} - 1)\rceil^+$ in realization (R) and $k_R := 0$ in realization (S) (no halving occurs for $k \ge k_R$), and $K_1 := \max\{k_R, \lceil aB/\sigma\rceil\}$. Then for every produced iterate with $k \ge K_1$,*

$$h_k \le h_{K_1}\Big(\tfrac{K_1+a}{k+a}\Big)^a \exp\Big(\tfrac{a^2 B}{\sigma(K_1+a-1)}\Big), \qquad h_{K_1} \le h_0 + BD^2 \sum_{j<K_1} \eta_j^2,$$

so $h_k = O(k^{-a})$ on any run with infinitely many accepted steps.

Proof of Proposition A.14 and Corollary 4.2. With $\delta = 0$ the oracle is exact, so $g_k = \max_{q\in\mathcal{C}}\langle -\nabla_k, \mathrm{Log}_{x_k} q\rangle \ge \langle -\nabla_k, \mathrm{Log}_{x_k} x^\star\rangle \ge h_k$ by convexity. From (8), $h_{k+1} \le h_k - \eta_k g_k + B\eta_k^2 d_k^2 \le h_k - \eta_k g_k(1 - B\eta_k/\sigma)$ using $d_k^2 \le g_k/\sigma$; when $B\eta_k/\sigma \le 1$ the coefficient of g_k is nonpositive and $g_k \ge h_k$ gives $h_{k+1} \le (1 - \eta_k + \frac{B}{\sigma}\eta_k^2)h_k$. For every accepted step, $h_{k+1} \le h_k + B\eta_k^2 D^2$ (drop $-\eta_k g_k \le 0$), which bounds h_{K_1} using the accepted steps $\eta_j \le \eta_j^0$. (i) In realization (R), the accepted step equals the proposed $a/(k+a)$ whenever $a/(k+a) \le \bar\eta_r$ (Lemma 3.3; the test uses the current chord and may also pass above the global cap), in particular for $k \ge a(\bar\eta_r^{-1} - 1)$, hence for $k \ge k_R$; in realization (S) the accepted step is always the proposed one. For $j \ge K_1$ the accepted step is $\eta_j = a/(j+a) \le \sigma/B$ (as $j \ge \lceil aB/\sigma\rceil$ gives $a/(j+a) \le a/(aB/\sigma + a) \le \sigma/B$), so $\prod_{j=K_1}^{k-1}(1 - \eta_j + \frac{B}{\sigma}\eta_j^2) \le \exp\big(-\sum_{j=K_1}^{k-1}\eta_j + \frac{B}{\sigma}\sum_{j\ge K_1}\eta_j^2\big)$, with $\sum_{j=K_1}^{k-1}\frac{a}{j+a} \ge a\log\frac{k+a}{K_1+a}$ and $\sum_{j\ge K_1}\frac{a^2}{(j+a)^2} \le \frac{a^2}{K_1+a-1}$. (ii) The proposed constant step $\eta_* \le \bar\eta_r$ passes the test of realization (R) without halving (Lemma 3.3), $\eta_* \le \frac12 < 1$ keeps $\beta_k > 0$ in realization (S), and $\eta_* \le \sigma/(2B)$ gives the factor $1 - \eta_* + \frac{B}{\sigma}\eta_*^2 \le 1 - \eta_*/2$. Corollary 4.2(i) is the proposition with the transient index K_1 and the constant left implicit; the steps with $k < K_1$ are the finite transient named there. □

Remark A.15 (what certification changes in the outer recurrence). The shift $m_r - 2$ delays the schedule until the step cap is met (about $2K/\alpha$ for $r = 2$ when $r_0 = D$ and the error bound dominates; the domain constraint contributes a shift of order $2\widehat{D}/r_0$ otherwise), and GKD^r (realization (R)) or GcD^2 (realization (S)) in Q_κ is the coefficient the accepted endpoint error adds to the η_k^2 term.

Remark A.16 (a sublinear-only allowance). The chord-relative progress tolerance $c\eta_k^2 d_k^2$ in (4) is sufficient, not necessary, for the sublinear bound. In realization (S), replace the nondegenerate acceptance budget by $\beta_k^{\mathrm{abs}} := \min\{\alpha\eta_k(1-\eta_k)d_k^2,\, c\eta_k^2\widehat{D}^2\}$ with $\widehat{D} \ge D$, keeping the zero-direction and full-step branches and the positive reserve. Every produced iterate is still feasible, because the geometric term is unchanged; for $d_k > 0$ and $0 < \eta_k < 1$ the budget is positive, so the finite-return argument of Lemma 3.3 applies verbatim; and an accepted step has $e_k \le c\eta_k^2\widehat{D}^2$, so (13) and the transfer step $F(x_{k+1}) - F(z_{k+1}) \le Ge_k$ give $h_{k+1} \le (1-\eta_k)h_k + \eta_k^2\big[\frac{L_F}{2}D^2 + Gc\widehat{D}^2 + \delta\big]$, and Lemma A.11 yields $h_k \le \max\{mh_0, 4(\frac{L_F}{2}D^2 + Gc\widehat{D}^2 + \delta)\}/(k+m)$ for $\eta_k = 2/(k+m)$, $m \ge 2$. The chord-relative bound $e_k \le c\eta_k^2 d_k^2$ in (8) and the fast-rate argument of Corollary 4.2, which absorbs e_k into the gap through $d_k^2 \le g_k/\sigma$, are not retained under this relaxation, and no experiment in Appendix D evaluates it.

Proposition A.17 (refinement count under an error-decay model)**.** *In realization (S) fix an iteration with $d_k > 0$ and $0 < \eta_k < 1$, suppose $E_j^{\mathrm{tr}} \le C_E\, 2^{-pj}(\eta_k d_k)^{p+1}$ with $p \ge 1$, $C_E > 0$, and that trial j costs $W_0 2^j$ primitive evaluations. For $\theta \in (\theta_{\mathrm{stop}}, 1]$ put $c_k := \min\{c, \alpha(1-\eta_k)/\eta_k\}$ (= c when $\eta_k \le \frac12$) and define the deterministic index $J_k := \big\lceil \frac1p \log_2\big(\frac{C_E}{\theta c_k}(\eta_k d_k)^{p-1}\big)\big\rceil^+$. If every trial $0 \le j \le J_k$ passes its primitive and domain checks and returns a candidate, and $E_j^{\mathrm{fl}} \le (1-\theta)\beta_k$ for every such j, the loop accepts at an index $j_k^{\mathrm{acc}} \le J_k$ (earlier acceptance is possible); when $\eta_k \le \frac12$, $J_k \le J := \big\lceil \frac1p \log_2(C_E D^{p-1}/(\theta c))\big\rceil^+$, and the map work through acceptance, counting rejected trials, is at most $W_0(2^{j_k^{\mathrm{acc}}+1} - 1) < W_0 2^{J_k+1}$. In realization (R) each nonzero-direction iteration evaluates the retraction once after at most $\lceil \log_2(\eta_k^0/\bar\eta_r)\rceil^+$ rejected scalar tests, and a*

zero-direction iteration copies the current point. Hence, provided the contract (including the bound $E_j^{\mathrm{fl}} \le (1-\theta)\beta_k$ for $j \le J_k$) holds through the horizon, reaching $h_k \le \varepsilon$ in Theorem 4.1(ii) takes at most $K_\varepsilon := \lceil A/\varepsilon \rceil$ oracle calls, A the numerator of the bound. The map work is at most $W_0 2^{J+1} K_\varepsilon + W_{\mathrm{init}}$ in (S), W_{init} the finite map work of the iterations with $\eta_k > \frac{1}{2}$, and (R) makes at most K_ε retraction evaluations. Oracle calls, distance evaluation, enclosure construction, scalar tests, and precision changes are excluded from this work count; which implementation satisfies the truncation model, on which horizon, is a separate statement (Proposition D.2 for chart Euler on $\mathrm{SO}(3)$).

Proof of Proposition A.17. Since $\beta_k = \min\{\alpha\eta_k(1-\eta_k), c\eta_k^2\} d_k^2 = c_k \eta_k^2 d_k^2$ with $c_k = \min\{c, \alpha(1-\eta_k)/\eta_k\}$, and $c_k = c$ when $\eta_k \le \frac{1}{2}$ (because $c \le \alpha$), the test $E_j = E_j^{\mathrm{tr}} + E_j^{\mathrm{fl}} \le \beta_k$ is implied by $E_j^{\mathrm{tr}} \le \theta\beta_k$ together with $E_j^{\mathrm{fl}} \le (1-\theta)\beta_k$, and the former by $C_E 2^{-pj} (\eta_k d_k)^{p+1} \le \theta c_k \eta_k^2 d_k^2$, i.e., by $2^{pj} \ge (C_E/(\theta c_k))(\eta_k d_k)^{p-1}$, which holds at $j = J_k$ by definition of J_k. The bound on E_j^{fl}, together with $\theta > \theta_{\mathrm{stop}}$, prevents a reserve exit for $j \le J_k$. The total-error test holds at J_k, so the first accepted index satisfies $j_k^{\mathrm{acc}} \le J_k$. Acceptance may occur earlier because the returned bounds can be smaller than the sufficient upper bound used to define J_k. Since $\eta_k d_k \le D$, $J_k \le J$ when $\eta_k \le \frac{1}{2}$. The work through acceptance, including rejected trials, is $\sum_{i=0}^{j_k^{\mathrm{acc}}} W_0 2^i = W_0(2^{j_k^{\mathrm{acc}}+1} - 1) < W_0 2^{J_k+1}$. In realization (R), Lemma 3.3 bounds the rejected tests and one retraction is evaluated. From Lemma A.11, $h_k \le A/(k+m) \le \varepsilon$ as soon as $k + m \ge A/\varepsilon$, which needs at most $K_\varepsilon = \lceil A/\varepsilon \rceil$ iterations, each with one oracle call; summing $W_0 2^{J+1}$ over the iterations with $\eta_k \le \frac{1}{2}$ and the finite work of the others gives the stated map work. For $\eta_k = 2/(k+m)$ and $m \ge 4$, $\eta_k \le \frac{1}{2}$ for all $k \ge 0$. □

Lemma A.18 (quadratic growth and target collapse at a boundary-active minimizer)**.** *Let $\mathcal{C}$ satisfy Assumption 2.1 and be geodesically α-strongly convex, let F be C^1 on a neighbourhood of $\mathcal{C}$ with a minimizer $x^\star$ of F over $\mathcal{C}$ at which $g^\star := \operatorname{grad} F(x^\star) \ne 0$, and put $a_\star := \alpha\|g^\star\| > 0$. Then (i) $\langle g^\star, \operatorname{Log}_{x^\star} q\rangle \ge a_\star\, d^2(x^\star, q)$ for all $q \in \mathcal{C}$; (ii) if F is geodesically convex on $\mathcal{C}$, then $F(x) - F^\star \ge a_\star d^2(x^\star, x)$ for all $x \in \mathcal{C}$, so $x^\star$ is the unique minimizer and $h_k \to 0$ implies $x_k \to x^\star$, without strong convexity of F; (iii) if $x_k \to x^\star$ and $q_k \in \mathcal{C}$ satisfies $\langle \operatorname{grad} F(x_k), \operatorname{Log}_{x_k} q_k\rangle \le \min_{q\in\mathcal{C}} \langle \operatorname{grad} F(x_k), \operatorname{Log}_{x_k} q\rangle + \delta_k$ with $\delta_k \to 0$, then $q_k \to x^\star$ and $d(x_k, q_k) \to 0$.*

Proof. For $q \in \mathcal{C}$ the geodesic γ from $x^\star$ to q lies in $\mathcal{C}$, so $F(\gamma(t)) \ge F(x^\star)$ for $t \in [0,1]$ and $\langle g^\star, \operatorname{Log}_{x^\star} q\rangle = \frac{d}{dt} F(\gamma(t))|_{t=0} \ge 0$. Put $n := -g^\star$; then $\langle n, a\rangle \le 0$ for all $a \in A_{x^\star}$. By Lemma A.3(ii), every $v \in T_{\mathcal{C}}(x^\star)$ is a limit of a_j/τ_j with $a_j \in A_{x^\star}$, $\tau_j \to 0$, so $\langle n, v\rangle \le 0$ for every $v \in T_{\mathcal{C}}(x^\star)$. Hence $n \in T_{\mathcal{C}}(x^\star)^\circ$. Inequality (11) gives $\langle n, \operatorname{Log}_{x^\star} q\rangle \le -\alpha d^2(x^\star, q)\|n\|$, which is (i). For (ii), geodesic convexity gives $F(x) - F^\star \ge \langle g^\star, \operatorname{Log}_{x^\star} x\rangle \ge a_\star d^2(x^\star, x)$ by (i). For (iii), the map $\phi(x, q) := \langle \operatorname{grad} F(x), \operatorname{Log}_x q\rangle$ is continuous on $\mathcal{C} \times \mathcal{C}$ by continuity of $\operatorname{grad} F$ and Lemma A.3(v). The oracle guarantee with the competitor $q = x_k$ gives $\phi(x_k, q_k) \le \phi(x_k, x_k) + \delta_k = \delta_k$. If $d(x^\star, q_k) \ge \varepsilon > 0$ along a subsequence, compactness yields a further subsequence with $q_k \to \bar{q}$, $d(x^\star, \bar{q}) \ge \varepsilon$, and continuity gives $\phi(x_k, q_k) \to \phi(x^\star, \bar{q}) \ge \alpha\|g^\star\|\varepsilon^2 > 0$, contradicting $\phi(x_k, q_k) \le \delta_k \to 0$. Hence $q_k \to x^\star$ and $d(x_k, q_k) \le d(x_k, x^\star) + d(x^\star, q_k) \to 0$. □

A.5 Proof of Proposition 5.2 and the boundary-normal margin on SO(3) balls

Proposition A.19 (boundary-normal margin and small-radius expansion on $\mathrm{SO}(3)$)**.** *Let $0 < \rho < \pi/2$ and $\alpha_\rho = \frac{1}{4}\cot(\rho/2)$ as in Proposition 5.2.*

- *(i) For $R \in \partial\mathcal{C}_\rho$, the outward unit normal $n = -\operatorname{Log}_R R_c/\rho$ generates $T_{\mathcal{C}}(R)^\circ = \mathbb{R}_+ n$, and every $Q \in \mathcal{C}_\rho$ with $\ell := d(R, Q) > 0$ satisfies $\langle n, \operatorname{Log}_R Q\rangle \le -\ell \cot(\frac{\rho}{2}) \tan(\frac{\ell}{4}) \le -\alpha_\rho \ell^2$, with equality in the first inequality if and only if $Q \in \partial\mathcal{C}_\rho$.*
- *(ii) Along boundary anchors $Q_k \to R$ the exact margin in (i) is $\alpha_\rho \ell_k^2 (1 + o(1))$, and the ratio of the exact margin to $\alpha_\rho \ell^2$ is strictly increasing in ℓ; hence no $\alpha > \alpha_\rho$ satisfies (11) for $\mathcal{C}_\rho$, which gives the sharpness claim of Proposition 5.2(i).*
- *(iii) As $\rho \to 0$, $\alpha_\rho = \frac{1}{2\rho} - \frac{\rho}{24} + O(\rho^3)$, but $1/(2\rho)$ is not a valid modulus for any fixed $\rho > 0$.*

Proof of Propositions 5.2 and A.19. Let $\varpi\colon S^3(2) \to \mathrm{SO}(3)$ be the universal cover, a local isometry whose deck transformation is the antipodal map. For $a, b \in S^3(2)$ one has $d_{\mathrm{SO}(3)}(\varpi a, \varpi b) = \min\{d_S(a,b),\, 2\pi - d_S(a,b)\}$, so on any set of spherical diameter at most π the cover is an isometry onto its image. Since $\rho < \pi/2$, the lift $\hat{\mathcal{C}}$ of $\mathcal{C}_\rho$ about a lift $\hat{c}$ of R_c has diameter at most $2\rho < \pi$; distances between points of this lift, their minimizing segments, and angles are preserved, and the clearance balls used below correspond under the covering map. We work on the unit sphere $S^3 \subset \mathbb{R}^4$, where lengths are half their values on $S^3(2)$; accordingly, we write $\hat\rho = \rho/2 < \pi/4$ and $\hat\ell = \ell/2$; every hatted length below is such a unit-sphere quantity, half of its $\mathrm{SO}(3)$ value. Distances on S^3 are $\hat d(a,b) = \arccos\langle a, b\rangle$.

Proposition 5.2: modulus and exact clearance. Let $x, y \in \hat{\mathcal{C}}$ and $\hat\ell := \hat d(x,y) \le 2\hat\rho < \pi/2$. If $x = y$ the claims are trivial ($\gamma \equiv x \in \hat{\mathcal{C}}$, and $\Delta_\rho(\eta, 0) = 0$ by the definition of Δ_ρ in Proposition 5.2(ii)). Otherwise let $\gamma(\theta) = \big(\sin((1-\theta)\hat\ell)\, x + \sin(\theta\hat\ell)\, y\big)/\sin\hat\ell$ be the minimizing geodesic (unique since $\hat\ell < \pi$). Put $\varphi(\theta) := \langle \gamma(\theta), \hat c\rangle = \cos\hat\varrho(\theta)$ with $\hat\varrho(\theta) := \hat d(\gamma(\theta), \hat c)$. Then

$$\varphi(\theta) = \frac{\sin((1-\theta)\hat\ell)\,\varphi(0) + \sin(\theta\hat\ell)\,\varphi(1)}{\sin\hat\ell}, \qquad \varphi(0), \varphi(1) \ge \cos\hat\rho, \tag{15}$$

with nonnegative coefficients because $(1-\theta)\hat\ell, \theta\hat\ell \in [0, \pi/2]$. Hence

$$\cos\hat\varrho(\theta) \;\ge\; \cos\hat\rho\; S(\theta, \hat\ell), \qquad S(\theta,\hat\ell) := \frac{\sin((1-\theta)\hat\ell) + \sin(\theta\hat\ell)}{\sin\hat\ell}, \tag{16}$$

with equality, for $0 < \theta < 1$, if and only if $\varphi(0) = \varphi(1) = \cos\hat\rho$, i.e., $x, y \in \partial\hat{\mathcal{C}}$ (at $\theta = 0$, resp. $\theta = 1$, only x, resp. y, needs to be on the boundary). By the sum-to-product formulas, $S(\theta,\hat\ell) = \cos((1-2\theta)\hat\ell/2)/\cos(\hat\ell/2)$, the form displayed in Proposition 5.2(ii), which is continuous at $\hat\ell = 0$ with value 1. For $\hat\ell \le 2\hat\rho$ boundary-to-boundary segments of length $\hat\ell$ exist, and for them (16) is an equality with left side at most 1; so $\cos\hat\rho\, S(\theta,\hat\ell) \le 1$ for all admissible $(\theta, \hat\ell)$, and since $\hat\varrho(\theta) \in [0, \hat\rho] \subset [0, \pi/2]$ where cos is decreasing, $\hat\varrho(\theta) \le \arccos(\cos\hat\rho\, S(\theta,\hat\ell))$ with equality when both chord endpoints lie on the boundary. Writing $\hat\Delta(\theta,\hat\ell) := \hat\rho - \arccos\big(\cos\hat\rho\, S(\theta,\hat\ell)\big)$ and doubling lengths, $\rho - d(\gamma(\eta), R_c) \ge 2\hat\Delta(\eta, \ell/2) = \Delta_\rho(\eta,\ell)$ with equality when both chord endpoints lie on the boundary. This proves the clearance bound of Proposition 5.2(ii) and its equality case.

For the modulus we bound S from below. Differentiating (15), $\varphi'' = -\hat\ell^2\varphi$; the ball of radius $\hat\rho < \pi/2$ is geodesically convex, so $\hat\varrho \le \hat\rho$, $\varphi \ge \cos\hat\rho > 0$, and $\varphi'' \le -\hat\ell^2\cos\hat\rho$. The function $\psi(\theta) := \varphi(\theta) - (1-\theta)\varphi(0) - \theta\varphi(1) - \frac12\hat\ell^2\cos\hat\rho\,\theta(1-\theta)$ is concave with $\psi(0) = \psi(1) = 0$, so $\psi \ge 0$ on $[0,1]$; with $\varphi(0), \varphi(1) \ge \cos\hat\rho$ this gives

$$\cos\hat\varrho(\theta) \;\ge\; \cos\hat\rho\,\big(1 + \tfrac{\varkappa}{2}\big), \qquad \varkappa := \theta(1-\theta)\hat\ell^2. \tag{17}$$

Let $\hat k := \frac12\cot\hat\rho$ and $\hat r := \hat k\,\varkappa$. Since $\varkappa \le \hat\ell^2/4 \le \hat\rho^2$ and $\hat\rho\cot\hat\rho \le 1$, we have $\hat r \le \hat\rho/2$, so $\hat\rho - \hat r \in [\hat\rho/2, \hat\rho] \subset [0, \pi/2]$. Using $\cos\hat r \le 1$ and $\sin\hat r \le \hat r$,

$$\cos(\hat\rho - \hat r) = \cos\hat\rho\cos\hat r + \sin\hat\rho\sin\hat r \le \cos\hat\rho + \sin\hat\rho\cdot\tfrac12\cot\hat\rho\,\varkappa = \cos\hat\rho\big(1 + \tfrac{\varkappa}{2}\big) \le \cos\hat\varrho(\theta)$$

by (17). Since $\hat\varrho(\theta)$ and $\hat\rho - \hat r$ both lie in $[0, \pi/2]$, this inequality implies $\hat\varrho(\theta) \le \hat\rho - \hat r$. The triangle inequality therefore places the closed ball of radius $\hat r = \frac12\cot\hat\rho\,\theta(1-\theta)\hat\ell^2$ about $\gamma(\theta)$ inside $\hat{\mathcal{C}}$. Transporting this containment to $\mathrm{SO}(3)$ doubles every length, so the radius becomes $2\hat r = \cot(\rho/2)\,\theta(1-\theta)\,(\ell/2)^2 = \alpha_\rho\,\theta(1-\theta)\,\ell^2$ with $\ell = d(x,y)$. Since the modulus multiplies a squared length, doubling all lengths halves it: $\alpha_\rho = \frac12\hat k = \frac14\cot(\rho/2)$. This is Definition 2.3 with modulus α_ρ; Definition A.1(b) follows from Proposition A.2. Applied to a boundary-to-boundary segment of length ℓ, for which the clearance equals $\Delta_\rho(\eta,\ell)$, it gives $\Delta_\rho(\eta,\ell) \ge \alpha_\rho\eta(1-\eta)\ell^2$. A small-chord expansion of Δ_ρ also follows from $S(\theta,\hat\ell) = 1 + \frac12\theta(1-\theta)\hat\ell^2 + O(\hat\ell^4)$, uniformly in $\theta \in [0,1]$, because the Taylor coefficients of $\cos((1-2\theta)\hat\ell/2)/\cos(\hat\ell/2)$ are bounded in θ. The defining identity $\cos(\hat\rho - \hat\Delta) = \cos\hat\rho\, S(\theta,\hat\ell) = \cos\hat\rho(1 + \frac{\varkappa}{2}) + O(\hat\ell^4)$ then gives $\hat\Delta = \frac12\cot\hat\rho\,\varkappa + O(\hat\ell^4)$, whose leading term is the quadratic radius $\hat r$; doubling lengths yields $\Delta_\rho = \alpha_\rho\eta(1-\eta)\ell^2 + O(\ell^4)$. Finally, the clearance $\rho - d(z, R_c)$ of a point $z \in \mathcal{C}_\rho$ equals $\mathrm{dist}(z, \mathcal{M}\setminus\mathcal{C}_\rho)$. The triangle inequality gives

the lower bound. If $z \neq R_c$, the radial geodesic from R_c through z is minimizing up to length $\pi > \rho$, and its points just beyond the boundary sphere are exterior points at distance arbitrarily close to $\rho - d(z, R_c)$ from z; if $z = R_c$, any radial direction gives the same conclusion. Hence a budget depending only on (ρ, η, ℓ) that exceeds $\Delta_\rho(\eta, \ell)$ is violated by a segment of length ℓ whose endpoints both lie on the boundary, at which the clearance equals $\Delta_\rho(\eta, \ell)$.

Proposition A.19(i). $\partial\mathcal{C}_\rho$ is the level set $\{d(\cdot, R_c) = \rho\}$ of a function that is smooth with nonvanishing gradient $-\operatorname{Log}_R R_c/\rho$ away from R_c and its cut locus ($\rho < \pi$), so $T_{\mathcal{C}}(R) = \{v : \langle n, v\rangle \leq 0\}$ and $T_{\mathcal{C}}(R)^\circ = \mathbb{R}_+ n$. Let ϑ_R be the angle at R between $\operatorname{Log}_R R_c$ and $\operatorname{Log}_R Q$, so $\langle n, \operatorname{Log}_R Q\rangle = -\ell\cos\vartheta_R$; angles are preserved by the isometry. In the spherical triangle $(\hat{c}, R, Q)$ on S^3 with sides $\hat{d}(R, \hat{c}) = \hat{\rho}$, $\hat{a} := \hat{d}(\hat{c}, Q) \leq \hat{\rho}$, $\hat{d}(R, Q) = \hat{\ell} \in (0, \pi/2)$, the law of cosines gives

$$\cos\vartheta_R = \frac{\cos\hat{a} - \cos\hat{\rho}\cos\hat{\ell}}{\sin\hat{\rho}\sin\hat{\ell}} \;\geq\; \frac{\cos\hat{\rho}\,(1 - \cos\hat{\ell})}{\sin\hat{\rho}\sin\hat{\ell}} = \cot\hat{\rho}\,\tan\!\big(\tfrac{\hat{\ell}}{2}\big) = \cot\!\big(\tfrac{\rho}{2}\big)\tan\!\big(\tfrac{\ell}{4}\big),$$

with equality iff $\cos\hat{a} = \cos\hat{\rho}$, i.e., $Q \in \partial\mathcal{C}_\rho$. Since $\ell \leq 2\rho < \pi$, $\ell/4 < \pi/4$ and $\tan(\ell/4) \geq \ell/4$, giving the quadratic bound.

Sharpness, Proposition 5.2(i), and Proposition A.19(ii). Suppose Definition 2.3 held for $\mathcal{C}_\rho$ with some $\alpha > \alpha_\rho$. By Theorem 5.1, $\langle n, \operatorname{Log}_R Q\rangle \leq -\alpha\ell^2$ for all $Q \in \mathcal{C}_\rho$. For boundary anchors $Q_k \to R$ with $\ell_k \to 0$, equality in Proposition A.19(i) gives $-\langle n, \operatorname{Log}_R Q_k\rangle/\ell_k^2 = \cot(\rho/2)\tan(\ell_k/4)/\ell_k \to \alpha_\rho < \alpha$, a contradiction. The same computation shows that the ratio of the exact margin to $\alpha_\rho\ell_k^2$ is $4\tan(\ell_k/4)/\ell_k \to 1$; for finite ℓ_k the exact margin is strictly larger than the quadratic bound.

Proposition A.19(iii). $\cot u = u^{-1} - u/3 + O(u^3)$ gives the expansion. For the last claim, the ratio $4\tan(\ell/4)/\ell$ of the exact margin to $\alpha_\rho\ell^2$ is strictly increasing in ℓ (as $\tan(u)/u$ is), which is the monotonicity asserted in (ii). Thus the infimum of $-\langle n, \operatorname{Log}_R Q\rangle_R/d(R, Q)^2$ over boundary anchors $Q \neq R$ is α_ρ. Since $\alpha_\rho < 1/(2\rho)$, the Euclidean coefficient $1/(2\rho)$ is not valid in (i) or Definition 2.3. □

A.6 The ambient-tube theorem and its proof

Under Assumption 2.4 for a set S, L-smoothness of φ on S (Definition 2.2) implies the *exponential form*

$$\varphi(\operatorname{Exp}_x v) \leq \varphi(x) + \langle \operatorname{grad}\varphi(x), v\rangle_x + \tfrac{L}{2}\|v\|_x^2 \qquad \text{whenever } \operatorname{Exp}_x([0,1]v) \subseteq S, \tag{18}$$

because that segment lies in V and is therefore the unique minimizing geodesic from x to $\operatorname{Exp}_x v$, so $\operatorname{Log}_x \operatorname{Exp}_x v = v$ and $d(x, \operatorname{Exp}_x v) = \|v\|_x$ in the two-point inequality of Definition 2.2.

Theorem A.20 (strongly convex sublevel sets, tube form)**.** *Let $\mathcal{M}$ be complete, let $U \subseteq \mathcal{M}$ be open, uniquely geodesically convex, and satisfy Assumption 2.4; let $f\colon U \to \mathbb{R}$ be differentiable, geodesically L-smooth and μ-strongly convex on U with $0 < \mu \leq L$, with a stationary point $x^\star \in U$, $\operatorname{grad} f(x^\star) = 0$, $f^\star := f(x^\star)$; let $s > 0$ and $Q_s := \{x \in U : f(x) - f^\star \leq s\}$. Set $\alpha_0 := \mu/(2\sqrt{2sL})$, $\tau := \sqrt{2s/L}$, and $T_\tau := \{z \in \mathcal{M} : \operatorname{dist}(z, Q_s) \leq \tau\}$, and assume $T_\tau \subseteq U$. Then Q_s is compact, uniquely geodesically convex, satisfies $0 \leq f(q) - f^\star \leq s$ and $d(q, x^\star) \leq \sqrt{2s/\mu}$ for all $q \in Q_s$, and is geodesically α_0-strongly convex.*

Lemma A.21 (gradient budget from smoothness)**.** *Let U be uniquely geodesically convex and satisfy Assumption 2.4, let $L > 0$ and $s > 0$, and let f be geodesically convex and L-smooth on U, with $\operatorname{grad} f(x^\star) = 0$ for some $x^\star \in U$ and $f^\star := f(x^\star)$. If $x \in U$ satisfies $f(x) - f^\star \leq s$ and $\overline{B}(x, \sqrt{2s/L}) \subseteq U$, then $\|\operatorname{grad} f(x)\|_x \leq \sqrt{2L(f(x) - f^\star)} \leq \sqrt{2Ls}$.*

Proof. First-order convexity at $x^\star$ gives $f \geq f^\star$ on U. Write $g = \|\operatorname{grad} f(x)\|_x$, $\varrho = \sqrt{2s/L}$, and assume $g > 0$. If $g > L\varrho$, the trial geodesic segment $t \mapsto \operatorname{Exp}_x(-t\operatorname{grad} f(x)/g)$, $0 \leq t \leq \varrho$, lies in $\overline{B}(x, \varrho) \subseteq U$, so the exponential form (18) gives, at $x_\varrho := \operatorname{Exp}_x(-\varrho\operatorname{grad} f(x)/g)$, $f(x_\varrho) \leq f(x) - \varrho g + \frac{L}{2}\varrho^2 < f(x) - \frac{L}{2}\varrho^2 = f(x) - s \leq f^\star$, contradicting $f \geq f^\star$. Hence $g \leq L\varrho$, the trial point $\operatorname{Exp}_x(-\operatorname{grad} f(x)/L)$ lies within distance $g/L \leq \varrho$ of x, hence in U together with its segment, and smoothness gives $f^\star \leq f(\operatorname{Exp}_x(-\operatorname{grad} f(x)/L)) \leq f(x) - g^2/(2L)$. □

Proof of Theorem A.20. The tube assumption keeps the descent probes used to bound the gradient within U and ensures that the candidate clearance balls are evaluated where the function inequalities

hold. The set U is open, uniquely geodesically convex, and satisfies Assumption 2.4. The function f is differentiable, L-smooth, and μ-strongly convex on U, with $0 < \mu \le L$ and $\operatorname{grad} f(x^\star) = 0$. The tube inclusion $T_\tau \subseteq U$ holds for $\tau = \sqrt{2s/L}$.

(1) Boundedness and range. First-order convexity at $x^\star$ gives $f \ge f^\star$ on U, hence $0 \le f(q) - f^\star \le s$ on Q_s, the range statement. For $q \in Q_s$, strong convexity at $x^\star$ gives $s \ge f(q) - f^\star \ge \frac{\mu}{2} d^2(q, x^\star)$, so $d(q, x^\star) \le \sqrt{2s/\mu}$ and $D := \operatorname{diam} Q_s \le 2\sqrt{2s/\mu}$.

(2) Closedness and compactness. Let $q_n \in Q_s$, $q_n \to q \in \mathcal{M}$. Then $\operatorname{dist}(q, Q_s) = 0 \le \tau$, so $q \in T_\tau \subseteq U$, and continuity of f on U gives $f(q) - f^\star \le s$. Hence $q \in Q_s$, so Q_s is closed; being also bounded in the complete manifold $\mathcal{M}$, it is compact (Hopf–Rinow).

(3) Geodesic convexity and the chord inequality. For $x, y \in Q_s \subseteq U$ let γ be the unique minimizing geodesic, contained in U. Along γ the subsegments are the unique minimizing geodesics between their endpoints, so $\operatorname{Log}_{\gamma(t)} x = -t\dot\gamma(t)$ and $\operatorname{Log}_{\gamma(t)} y = (1-t)\dot\gamma(t)$. Applying the strong-convexity inequality at $\gamma(t)$ toward x and toward y, multiplying by $(1-t)$ and t, and adding, the first-order terms cancel and

$$f(\gamma(t)) \le (1-t)f(x) + tf(y) - \tfrac{\mu}{2}\, t(1-t)\, d^2(x,y) \le f^\star + s - \tfrac{\mu}{2} t(1-t) d^2(x,y).$$

Hence $\gamma \subset Q_s$; Q_s is uniquely geodesically convex and satisfies Assumption 2.1 (nonempty, as $x^\star \in Q_s$).

(4) Gradient budget. For $p \in Q_s$, $\overline{B}(p, \tau) \subseteq T_\tau \subseteq U$, so Lemma A.21 gives $\|\operatorname{grad} f(p)\| \le \sqrt{2L(f(p) - f^\star)}$.

(5) Clearance balls. Let $x, y \in Q_s$, $t \in [0,1]$, $p := \gamma(t)$, and $r := \alpha_0 t(1-t) d^2(x,y)$. If $x = y$ or $t \in \{0, 1\}$ then $r = 0$ and the claim is $p \in Q_s$, shown in step (3). Otherwise let $z \in \mathcal{M}$ with $d(p, z) \le r$. By step (1), $r \le \alpha_0 D^2/4 \le \tau/2$, using $D^2 \le 8s/\mu$ from step (1) and $\alpha_0 \cdot 2s/\mu = \frac{1}{2}\sqrt{2s/L} = \tau/2$, so every point of a minimizing geodesic from p to z is within $\tau/2$ of $p \in Q_s$ and lies in $T_\tau \subseteq U$. By Assumption 2.4 that segment is $\operatorname{Exp}_p([0,1]e)$ with $\|e\|_p = d(p, z)$, and the exponential form (18) together with step (4) gives

$$f(z) - f^\star \le (f(p) - f^\star) + \|\operatorname{grad} f(p)\|\, d(p,z) + \tfrac{L}{2} d(p,z)^2 \le \left(\sqrt{f(p) - f^\star} + \sqrt{\tfrac{L}{2}}\, d(p,z)\right)^2.$$

(6) Conclusion. By step (3), $f(p) - f^\star \le s - \omega$ with $\omega := \frac{\mu}{2} t(1-t) d^2(x,y) \in [0, s]$, and concavity of $\sqrt{\cdot}$ gives $\sqrt{s-\omega} \le \sqrt{s} - \omega/(2\sqrt{s})$. Hence $f(z) - f^\star \le s$ as soon as $\sqrt{L/2}\, d(p,z) \le \omega/(2\sqrt{s}) = \mu t(1-t) d^2(x,y)/(4\sqrt{s})$, i.e., $d(p,z) \le \frac{\mu}{2\sqrt{2sL}} t(1-t) d^2(x,y) = r$, which holds. As $z \in U$, $z \in Q_s$. This is Definition 2.3 with modulus α_0. The modulus is sharp on Euclidean balls. To see this, take $f = \frac{1}{2}\|\cdot\|^2$ on $\mathbb{R}^n$ with $n \ge 2$ and $\mu = L = 1$. Then Q_s is the ball of radius $R := \sqrt{2s}$, the classical strongly convex set of radius R (Vial, 1983), and $\alpha_0 = 1/(2R)$ is its modulus in Definition 2.3, the constant of the normal-cone inequality (Goncharov & Ivanov, 2017, Theorem 2.1(g)). A chord of length ℓ with both endpoints on the boundary has midpoint clearance $R - \sqrt{R^2 - \ell^2/4} = (\ell^2/4)/(R + \sqrt{R^2 - \ell^2/4})$, which is at least $\ell^2/(8R) = \alpha_0 \cdot \frac{1}{4}\ell^2$ and equals $\ell^2/(8R) + O(\ell^4)$, so no modulus larger than α_0 is valid. In dimension one, $Q_s = [-R, R]$. For the chord joining its endpoints, the ratio of the clearance $2Rt$ to $t(1-t)(2R)^2$ tends to $1/(2R)$ as $t \to 0$, giving the same conclusion. □

A.7 Restricted-domain smoothness is insufficient: the counterexample

Proposition A.22 (restricted-domain smoothness is insufficient)**.** *In $\mathbb{R}^2$ let $\mathcal{C} = \{(x,y) : x^4 + y^4 \le 1\}$, let $f\colon \mathcal{C} \to \mathbb{R}$ be the restriction of $\frac{1}{2}\|\cdot\|^2$ with the derivatives of the ambient function, and let $\frac{1}{2} < s < 2^{-1/2}$. Then f is 1-smooth and 1-strongly convex on the compact convex $\mathcal{C}$ with interior minimizer 0, and the restricted sublevel set $Q_s := \{p \in \mathcal{C} : f(p) \le s\} = \mathcal{C} \cap \{\|\cdot\|^2 \le 2s\}$ is a proper, compact, strictly convex subset, but it is not geodesically α-strongly convex for any $\alpha > 0$. Indeed, for sufficiently small $b > 0$, the chord with endpoints $p_b, q_b := ((1-b^4)^{1/4}, \pm b)$ has midpoint clearance $1 - (1-b^4)^{1/4} \le b^4$, whereas a modulus α would require clearance at least αb^2.*

The example of Proposition A.22 satisfies the stated restricted-domain hypotheses of Theorem 4.2 of Scieur et al. (2026), previously Theorem 5.2 of the 2024 preprint (Scieur et al., 2024), including

strict convexity of Q_s, but violates its conclusion. The proof there applies the smoothness bound to nearby trial points without first establishing their membership in the function domain. Theorem A.20 restores the conclusion under the explicit ambient-tube assumption, which this example violates because $(1,0) \in Q_s \cap \partial\mathcal{C}$; Section 5 states the consequence for sublevel-set constraints.

Proof. $\mathcal{C}$ is compact and convex, hence uniquely geodesically convex in $\mathbb{R}^2$; f is the restriction of the smooth ambient function $\frac{1}{2}\|\cdot\|^2$, whose gradient is used, so it is 1-smooth and 1-strongly convex on $\mathcal{C}$ with $\operatorname{grad} f(0) = 0$, $0 \in \operatorname{int}\mathcal{C}$, $f^\star = 0$. Since $\max_{\mathcal{C}} f = f(2^{-1/4}, 2^{-1/4}) = 2^{-1/2} > s$, $Q_s = \mathcal{C} \cap \{\|\cdot\|^2 \le 2s\}$ is a proper subset of $\mathcal{C}$; it is compact and strictly convex as the intersection of two strictly convex compact sets. For $0 < b < 1$ the points p_b, q_b lie on $\partial\mathcal{C}$ and $\|p_b\|^2 = (1-b^4)^{1/2} + b^2 \to 1 < 2s$, so $p_b, q_b \in Q_s$ for small b, with $d(p_b, q_b) = 2b$ and midpoint $m_b = ((1-b^4)^{1/4}, 0)$. Any point $(m_{b,1} + \varrho, 0)$ with $\varrho > 1 - (1-b^4)^{1/4}$ has first coordinate exceeding 1, hence lies outside $\mathcal{C} \supseteq Q_s$. Conversely, with $a_b := (1-b^4)^{1/4}$ and $\xi_b := 1 - a_b$, every $h \in \mathbb{R}^2$ with $\|h\|_2 \le \xi_b$ satisfies $\|m_b + h\|_4 \le \|m_b\|_4 + \|h\|_4 \le a_b + \|h\|_2 \le 1$ and $\|m_b + h\|_2 \le a_b + \xi_b = 1 < \sqrt{2s}$, so the Euclidean ball of radius ξ_b about m_b lies in Q_s. Together with the outward displacement argument, this proves that the midpoint clearance is exactly $\xi_b = 1 - (1-b^4)^{1/4}$. If Q_s were geodesically α-strongly convex, Definition 2.3 at $t = \frac{1}{2}$ would place the ball of radius $\alpha \cdot \frac{1}{4} \cdot (2b)^2 = \alpha b^2$ about m_b inside Q_s, forcing $\alpha b^2 \le 1 - (1-b^4)^{1/4} \le b^4$ (using $(1-u)^{1/4} \ge 1 - u$ on $[0,1]$), which implies $\alpha \le b^2$ for every sufficiently small $b > 0$. Letting $b \to 0$ contradicts $\alpha > 0$. Finally, no tube of positive width about Q_s lies in $\mathcal{C}$, because $(1,0) \in Q_s \cap \partial\mathcal{C}$, so the hypothesis of Theorem A.20 fails, consistent with its conclusion. □

A.8 Linear minimization over a ball is attained on the boundary

Lemma A.23 (linear minimization over a geodesic ball)**.** *Let $\mathcal{C} = \overline{B}(o, \rho) \subset \mathcal{M}$ be a closed geodesic ball with $\rho < \pi/2$ on* $\mathrm{SO}(3)$ *(metric of Section 5), $R \in \mathcal{C}$, and $0 \neq g \in T_R\mathcal{M}$. Then every minimizer of $Q \mapsto \langle g, \operatorname{Log}_R Q\rangle$ over $\mathcal{C}$ lies on $\partial\mathcal{C}$.*

Proof. $\varphi_g(Q) := \langle g, \operatorname{Log}_R Q\rangle$ is continuous on the compact set $\mathcal{C}$, so a minimizer exists. Since $d(R, Q) \le 2\rho < \pi$ for $Q \in \mathcal{C}$, $D_Q \operatorname{Log}_R$ is invertible on a neighbourhood of $\mathcal{C}$ (the singular values of $D \operatorname{Exp}_R$ are 1 and $\sin(\ell/2)/(\ell/2) > 0$ for $\ell < \pi$, Remark C.8), so $\operatorname{grad}\varphi_g(Q) = (D_Q \operatorname{Log}_R)^* g \neq 0$ for every $Q \in \mathcal{C}$. An interior minimizer of the smooth function φ_g would have zero gradient, a contradiction; hence every minimizer lies on $\partial\mathcal{C}$. □

B Continuous-time forward invariance from the same geometric margin

This appendix records the continuous-time use of the margin of Theorem 5.1(i), with complete hypotheses. Nothing in the main text depends on it; the statements are analogues of the discrete certificate and do not transfer to Algorithm 2 without a separate theorem. Throughout, $\mathcal{C}$ is *forward invariant* for $\dot{x} = X(x)$ if every maximal solution with $x(0) \in \mathcal{C}$ satisfies $x(t) \in \mathcal{C}$ for every $t \ge 0$ in its interval of existence. Under the regularity assumptions below, compactness and forward invariance of $\mathcal{C}$ imply that these solutions are forward complete. We use the Nagumo tangency condition (Nagumo, 1942; Aubin & Cellina, 1984; Blanchini, 1999). If $\mathcal{C} \subset \mathcal{M}$ is closed and X is locally Lipschitz on an open neighbourhood of $\mathcal{C}$, then $\mathcal{C}$ is forward invariant if and only if $X(p) \in T_{\mathcal{C}}(p)$ for every $p \in \partial\mathcal{C}$. This local statement reduces to the Euclidean theorem in a chart, and local Lipschitz continuity upgrades viability of some solution to invariance of the unique one. A merely continuous field gives viability of one solution, not invariance of all; a time-dependent field needs Carathéodory bounds (measurable in t, with locally integrable bound and Lipschitz modulus in x; Aubin & Cellina, 1984, Chapter 2) and the measurable viability theorems of Frankowska et al. (1995); and a differential inclusion needs a set-valued tangency condition (Clarke et al., 1998). Barrier-function and sublevel-set versions of the certificate, with examples of what fails when the closedness, uniqueness, or boundary-nondegeneracy hypotheses are omitted, are recorded in a companion note to be released with the code; nothing in this appendix depends on that note.

Further related work. Riemannian proximal and projected methods study optimization complexity and the control of the iterate region under their own oracle models, including inexact proximal subproblems in Martínez-Rubio & Pokutta (2023) and exact projected steps in Zhang & Sra (2016). Criscitiello & Boumal (2022) establish curvature-dependent lower bounds for exact first-order oracles. These results address different questions from certification of the final represented endpoint. Discrete methods on proximally smooth sets (Davis et al., 2025) and projected dynamics on irregular domains (Hauswirth et al., 2021) use exterior-ball or prox-regularity hypotheses (Remark C.9). Manifold-valued continuous normalizing flows (Lou et al., 2020; Mathieu & Nickel, 2020) keep trajectories on the manifold, which is a different requirement from invariance of a subset; forward invariance of neural ODEs (Xiao et al., 2023) and control barrier functions (Ames et al., 2017; 2019; de Sa et al., 2026) certify invariance of a subset for continuous-time dynamics, the setting of this appendix. Strongly convex sublevel sets in Euclidean space appear in Journée et al. (2010).

Anchor-directed flows with disturbances. Let $\mathcal{C}$ satisfy Assumption 2.1 and be geodesically α-strongly convex; its contingent cones are closed and convex by Lemma A.4 (for regular sublevel sets they are half-spaces). Consider

$$X(p) = \sum_{j=1}^{N} \lambda_j(p) \operatorname{Log}_p q_j(p) + e(p), \qquad \lambda_j(p) \geq 0,\ q_j(p) \in \mathcal{C}, \tag{19}$$

and define the *geometric margin* and the *outward disturbance*

$$m_{\mathrm{geo}}(p) := \alpha \sum_{j=1}^{N} \lambda_j(p)\, d^2\big(p, q_j(p)\big), \qquad E_\perp(p) := \sup_{n \in T_{\mathcal{C}}(p)^\circ,\ \|n\|_p \leq 1} \langle n, e(p) \rangle_p \ \geq 0.$$

Corollary B.1 (anchor-directed fields)**.** *If $E_\perp(p) \leq m_{\mathrm{geo}}(p)$ for every $p \in \partial\mathcal{C}$, then $X(p) \in T_{\mathcal{C}}(p)$ on $\partial\mathcal{C}$; if moreover the weights, anchors, disturbance, and logarithms extend so that X is locally Lipschitz on an open neighbourhood of $\mathcal{C}$, then $\mathcal{C}$ is forward invariant. Pointwise, $m_{\mathrm{geo}}(p) - E_\perp(p) > 0$ gives $\langle n, X(p) \rangle_p \leq -(m_{\mathrm{geo}}(p) - E_\perp(p)) < 0$ for every unit $n \in T_{\mathcal{C}}(p)^\circ$. If $\sigma := \inf_{\partial\mathcal{C}}(m_{\mathrm{geo}} - E_\perp) > 0$, then every additional disturbance e' with $\|e'(p)\| \leq \sigma$ on $\partial\mathcal{C}$ preserves the tangency condition, and $\mathcal{C}$ remains forward invariant when the perturbed field satisfies the same Lipschitz hypothesis; this uniform condition is sufficient for robustness and is not necessary, because $m_{\mathrm{geo}} - E_\perp$ is a conservative lower bound on the inward margin. The margin $m_{\mathrm{geo}}(p)$ vanishes when all positively weighted anchors coincide with p. The norm condition $\|e(p)\| \leq m_{\mathrm{geo}}(p)$ is sufficient because $E_\perp(p) \leq \|e(p)\|$. It can be conservative, since the sharper condition depends only on $E_\perp(p)$, the positive pairings of $e(p)$ with outward polar normals. If $q_0 \in \operatorname{int}\mathcal{C}$ with $\varrho := \operatorname{dist}(q_0, \partial\mathcal{C}) > 0$ and $X(p) = \lambda(p) \operatorname{Log}_p q_0 + e(p)$ with $\lambda \geq \underline{\lambda} > 0$, the uniform bound $\|e(p)\| \leq \alpha\underline{\lambda}\varrho^2$ on $\partial\mathcal{C}$ suffices.*

Proof. By (11), for unit $n \in T_{\mathcal{C}}(p)^\circ$, $\langle n, X(p) \rangle \leq -m_{\mathrm{geo}}(p) + \langle n, e(p) \rangle \leq -m_{\mathrm{geo}}(p) + E_\perp(p) \leq 0$; by the bipolar theorem for the closed convex cone $T_{\mathcal{C}}(p)$, $X(p) \in T_{\mathcal{C}}(p)$. The Nagumo condition gives invariance. For an additional disturbance e', the supremum defining $E_\perp$ is subadditive in the disturbance and the polar normals have norm at most one, so the outward disturbance of $e + e'$ is at most $E_\perp(p) + \|e'(p)\| \leq E_\perp(p) + \sigma \leq m_{\mathrm{geo}}(p)$, and the first part applies to the perturbed field. For the fixed interior anchor, $d(p, q_0) \geq \varrho$ on $\partial\mathcal{C}$ gives $m_{\mathrm{geo}}(p) \geq \alpha\underline{\lambda}\varrho^2 \geq \|e(p)\| \geq E_\perp(p)$. □

Remark B.2 (networks on product manifolds; product geometry). Equip $\mathcal{M}^N$ with the product metric, let $\mathcal{C}^N$ be the common constraint with $\mathcal{C}$ geodesically α-strongly convex, and for $i = 1, \ldots, N$ consider $\dot{x}_i = \sum_{j=1}^{N} a_{ij}(t) \operatorname{Log}_{x_i} x_j + e_i(t, x)$, defined on an open normal neighbourhood of $\mathcal{C}^N$ on which all logarithms are single-valued and smooth, with measurable, locally essentially bounded weights $a_{ij}(t) \geq 0$ and disturbances e_i satisfying the Carathéodory bounds above; assume Assumption 2.4 for $\mathcal{C}$. Lemma A.4 gives derivability of $\mathcal{C}$ and the product formula $T_{\mathcal{C}^N}(x) = \prod_i T_{\mathcal{C}}(x_i)$ (for products of sets that are derivable at the respective points the contingent cone of the product is the product of the contingent cones, Aubin & Frankowska, 1990, Chapter 4), so the tangency condition decouples agent by agent; in particular the corollary applies to the $\mathrm{SO}(3)$ balls of Proposition 5.2 without convexity of their logarithmic images. Hence, if for a.e. t, every i, and every $x \in \mathcal{C}^N$ with $x_i \in \partial\mathcal{C}$, $\sup_{n \in T_{\mathcal{C}}(x_i)^\circ, \|n\| \leq 1} \langle n, e_i(t, x) \rangle \leq \alpha \sum_{j=1}^{N} a_{ij}(t)\, d^2(x_i, x_j)$, then $\mathcal{C}^N$ is forward invariant for Carathéodory solutions. The certificate remains valid under any measurable, locally essentially bounded switching of the weights; its instantaneous magnitude adapts

to the selected neighbours and may vanish near consensus. State-triggered switching makes the field discontinuous in x and is not covered. When $N \geq 2$ and $\mathcal{C}$ contains more than one point, the product set $\mathcal{C}^N$ has no positive isotropic strong-convexity modulus (Remark C.11). The isotropic certificate of Definition 2.3 therefore cannot be applied directly to the whole product. The continuous-time tangency test above applies factorwise, whereas a discrete product method needs its own endpoint-error allocation and convergence analysis (Remark C.11).

Remark B.3 (cross-field caution). If the feasible set is a sublevel set $Q_s = \{V \leq s\}$ of V with $\operatorname{grad} V \neq 0$ on its boundary, while the flow is $-\operatorname{grad} F$ with $F \neq V$, the boundary requirement is $\langle \operatorname{grad} V, \operatorname{grad} F \rangle \geq 0$ on ∂Q_s. The geometry of Q_s alone does not imply this condition. At a boundary-active constrained minimizer $x^\star$ of F with $\operatorname{grad} F(x^\star) + \lambda \operatorname{grad} V(x^\star) = 0$ and $\lambda > 0$, the derivative of V along $-\operatorname{grad} F$ equals $\lambda \| \operatorname{grad} V(x^\star) \|^2 > 0$. Thus $-\operatorname{grad} F$ points strictly outward at $x^\star$, which prevents Q_s from being forward invariant under this flow. The nonvanishing boundary gradient is essential to this outward-pointing conclusion. For example, take $V(x) = (x^2 - 1)^3$, $Q_0 = [-1, 1]$, and $F(x) = (x - 1)^2$. At $x^\star = 1$, every $\lambda > 0$ satisfies the stationarity condition because $V'(1) = F'(1) = 0$. Nevertheless, the flow $\dot{x} = -2(x - 1)$ preserves $[-1, 1]$ and converges to $x^\star$. This is why the discrete certificates of Section 3 are stated for anchor-directed (Frank–Wolfe) updates rather than gradient steps. Wang & Ren (2026) give an explicit chart-quadratic example on $\mathrm{SO}(3)$ whose cost is geodesically strongly convex on a prescribed ball and whose unique minimizer over that ball is interior, while the negative-gradient vector field points strictly outward at a boundary point; the corresponding trajectory therefore exits the ball immediately. This establishes failure of forward invariance; the later return toward the minimizer reported there is a numerical observation and is not used here. Objective strong convexity alone does not certify forward invariance of an independently specified feasible set.

C Supporting geometry, examples, and numerical interpretations

Notation and terminology. We write η_k^0 for the proposed step, η_k for the accepted step, $\underline{\eta}_k$ for its guaranteed lower bound, and $\bar{\eta}_r$ for the uniform retraction cap. The geometric budget β is a lower bound on clearance, while the accepted budget β_k of (4) additionally imposes the progress tolerance $c\eta_k^2 d_k^2$ with coefficient c; Δ_ρ is the exact worst-case clearance on the $\mathrm{SO}(3)$ balls of Section 5, and $\zeta(\theta)$ is the chart coefficient of Proposition D.2. The *exact geodesic endpoint* $z_{k+1} = \operatorname{Exp}_{x_k}(\eta_k \operatorname{Log}_{x_k} q_k)$ is a mathematical point; the *represented endpoint* $\tilde{z}$ is the point that the declared decoding map assigns to stored data (Definition 3.1), and the *reference exponential* is the float64 evaluation of the exact endpoint, used as a baseline. A *feasible oracle* returns a point of $\mathcal{C}$ satisfying (5), and an *exact oracle* attains the minimum ($\delta_k = 0$). We distinguish analytic bounds proved in exact arithmetic from implementation bounds that are conditional on the rounding model of Assumption D.1; a *calibrated constant* is a value taken from samples, and an *error indicator*, such as a step-doubling discrepancy, has no established enclosing guarantee. Figure 2 is a schematic of the certificate.

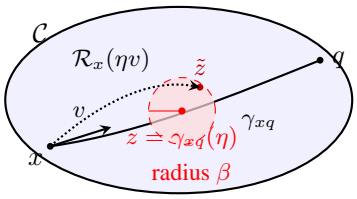

Figure 2: The geometric certificate for an inexact update (schematic). The exact interpolation z toward a feasible target q lies on the geodesic chord γ_{xq} with initial velocity $v = \operatorname{Log}_x q$; strong convexity places the ball of radius $\beta = \alpha\eta(1 - \eta)d^2(x, q)$ about z inside $\mathcal{C}$ (Corollary 3.2), so a represented endpoint $\tilde{z}$, produced by a retraction $\mathcal{R}_x(\eta v)$ or a numerical map that leaves the segment, is feasible whenever a valid enclosure of $d(\tilde{z}, z)$ is at most β. For numerical realization (S), Algorithm 2 additionally imposes the progress tolerance $c\eta^2 d^2(x, q)$, and the smaller tolerance decides acceptance. The tolerance vanishes at $\eta \in \{0, 1\}$ and as $d(x, q) \to 0$; the drawing does not depict a curved manifold isometrically.

Proposition C.1 (the retraction cap is worst-case sharp in the plane)**.** *Let $\mathcal{C}$ be the closed unit disk in $\mathbb{R}^2$ (modulus $\alpha = \frac{1}{2}$), $K > 0$, and $\mathcal{R}_x(v) := x + v + K\|v\|^2 x$, a smooth retraction with $d(\mathcal{R}_x(v), \operatorname{Exp}_x v) = K\|v\|^2\|x\| \le K\|v\|^2$ on $\mathcal{C}$ ($r = 2$, $r_0 = D$). Corollary A.9 certifies every step with $\eta \le \alpha/(\alpha + K) = 1/(1 + 2K)$, and no larger uniform cap is valid for this retraction: for every $\eta > 1/(1 + 2K)$ there are boundary chords whose retracted endpoint is infeasible.*

Proof. Set $x = (1, 0)$, $q = (\cos t, \sin t)$, $s := \|q - x\|^2 = 2 - 2\cos t$, and $v = \eta(q - x)$, so that $\langle x, v\rangle = \eta(\cos t - 1) = -\eta s/2$ and $\|v\|^2 = \eta^2 s$. Expanding the squared norm, $\|x + v + K\|v\|^2 x\|^2 = 1 + \|v\|^2 + K^2\|v\|^4 + 2\langle x, v\rangle + 2K\|v\|^2 + 2K\|v\|^2\langle x, v\rangle$, which reduces to

$$\|\mathcal{R}_x(\eta(q - x))\|^2 - 1 = s\big[-\eta(1 - \eta) + 2K\eta^2\big] + s^2\big[-K\eta^3 + K^2\eta^4\big].$$

For every $\eta > 1/(1 + 2K)$ the leading coefficient $\eta(\eta(1 + 2K) - 1)$ is positive, so sufficiently short boundary chords produce an infeasible endpoint. $\mathcal{R}_x(0) = x$ and $D\mathcal{R}_x(0) = \mathrm{id}$, so $\mathcal{R}$ is a retraction, and $\operatorname{Exp}_x v = x + v$ in the plane. (The identity is also checked numerically, Appendix D.) □

Remark C.2 (clearance from the chord length and anchor radii on $\mathrm{SO}(3)$ balls). Let $x, q \in \mathcal{C}_\rho$ with $r_x := d(R_c, x)$, $r_q := d(R_c, q)$, $\ell := d(x, q) > 0$, and $\eta \in [0, 1]$. Then, in the notation of the proof of Proposition 5.2,

$$\rho - d(\gamma_{xq}(\eta), R_c) = \Delta_\rho(\eta, \ell; r_x, r_q) := \rho - 2\arccos\left[\frac{\sin((1-\eta)\ell/2)\cos(r_x/2) + \sin(\eta\ell/2)\cos(r_q/2)}{\sin(\ell/2)}\right],$$

extended by continuity to the coincident case $x = q$ ($\ell = 0$, $r_q = r_x$), where it equals $\rho - r_x$; this is (15) with $\varphi(0) = \cos(r_x/2)$, $\varphi(1) = \cos(r_q/2)$ and doubled lengths, reduces to $\Delta_\rho(\eta, \ell)$ when $r_x = r_q = \rho$, and is the "clearance of the reference endpoint" curve of Figure 6(a), evaluated there at the float64 reference endpoint (checked numerically, Appendix D). It is a valid tolerance for the segment at hand and is at least $\Delta_\rho(\eta, \ell)$; for $0 < \eta < 1$ the inequality is strict when at least one endpoint lies in the interior, while at $\eta \in \{0, 1\}$ the clearance depends only on the selected endpoint. This explains why the chord-based budget can be many orders of magnitude below the actual clearance; it does not lift the progress tolerance $c\eta^2 d^2$, which is a separate restriction. The worst-case formula Δ_ρ subtracts nearly equal angles for short chords, and the arbitrary-segment formula does so when the exact endpoint is near the boundary; the implemented form is cancellation-free. Writing $r := \rho/2$ and $\delta := \Delta_\rho/2$, the identities $\cos((1 - 2\eta)\ell/4)/\cos(\ell/4) - 1 = \varsigma$ with $\varsigma := 2\sin((1-\eta)\ell/4)\sin(\eta\ell/4)/\cos(\ell/4)$, $\cos(r - \delta) = \cos(r)(1 + \varsigma) =: A$, and the half-angle identity $\tan(\delta/2) = (\cos(r - \delta) - \cos r)/(\sin r + \sin(r - \delta))$ give $\Delta_\rho(\eta, \ell) = 4\arctan\big[\cos(\frac{\rho}{2})\,\varsigma/(\sin(\frac{\rho}{2}) + \sqrt{1 - A^2})\big]$, which evaluates the numerator directly at order ℓ^2. A certified lower bound additionally requires directed rounding of this expression.

Remark C.3 (why the normal-cone form, not the scaling inequality). The Riemannian scaling inequality used for Frank–Wolfe rates (Scieur et al., 2026; Chen, 2026) is imposed at maximizers of $z \mapsto \langle w, \operatorname{Log}_x z\rangle_x$ over $\mathcal{C}$. At a boundary point p with outward polar normal n and $w = n$, the functional $z \mapsto \langle n, \operatorname{Log}_p z\rangle$ is nonpositive on $\mathcal{C}$ by (11) and vanishes at $z = p$, so its maximizer is p itself and the inequality collapses to $0 \ge 0$. A feasibility certificate for an arbitrary target q needs the inequality for every q, which is (11). In Hilbert space the corresponding formulations are equivalent. On a curved manifold the oracle-based scaling inequality does not directly give this boundary test for an arbitrary anchor; Theorem 5.1 supplies the required normal-cone inequality directly from geodesic strong convexity, without assuming convexity of the logarithmic images.

Remark C.4 (formal comparison of certification routes on $\mathrm{SO}(3)$). On a manifold of bounded curvature, geodesic balls of sufficiently small radius are Riemannian strongly convex by the bounded-curvature certificate of Scieur et al. (2026) (Theorem D.4 of the published version); its radius bound involves the sectional-curvature bound κ_{sec}, a bound on $\|\nabla \mathrm{Rm}\|$, and the strong convexity and smoothness constants of $\frac{1}{2}d^2$ on the ball. Specializing those constants to $\mathrm{SO}(3)$ ($\kappa_{\mathrm{sec}} = \frac{1}{4}$, $\nabla \mathrm{Rm} = 0$, smoothness constant 1, strong-convexity constant $\frac{r}{2}\cot\frac{r}{2}$ of $\frac{1}{2}d^2$ on the ball of radius r), a formal substitution of these constants yields the condition $\tan(r/2) \le \frac{1}{4}$, i.e., $r \le 2\arctan\frac{1}{4} \approx 28.07°$. Their theorem is stated for a globally uniquely geodesic ambient manifold, which the compact, positively curved $\mathrm{SO}(3)$ is not, and the localization lemma that would turn this substitution into a theorem on $\mathrm{SO}(3)$ is not proved here. The value therefore establishes no geometric or safety limit, and Proposition 5.2 covers every $\rho < \pi/2$ with the sharp modulus. For the normalized coupling $\operatorname{Log}_R Q/\ell$ the exact margin $\cot(\rho/2)\tan(\ell/4)$ applies; componentwise sign couplings are *not* covered, as they can rotate an inward logarithmic direction into an outward one.

Remark C.5 (choice of distance). All statements use the Riemannian distance d. If strong convexity and smoothness of f are measured with respect to an equivalent distance $\tilde{d}$ with $\ell_{\mathcal{M}} d \le \tilde{d} \le L_{\mathcal{M}} d$, then f is $\mu\ell_{\mathcal{M}}^2$-strongly convex and $LL_{\mathcal{M}}^2$-smooth with respect to d, and every result applies with these constants; e.g., the modulus of Theorem A.20 becomes $\mu\ell_{\mathcal{M}}^2/(2L_{\mathcal{M}}\sqrt{2sL})$, with tube width $\sqrt{2s/L}/L_{\mathcal{M}}$. This is how the constants $\ell_{\mathcal{M}}, L_{\mathcal{M}}$ of Scieur et al. (2026) enter.

Remark C.6 (what vanishes near convergence). The accepted budget $\beta_k = \min\{\alpha\eta_k(1-\eta_k), c\eta_k^2\}d_k^2$ can decrease as the step size η_k decreases or as the chord length d_k tends to zero. When $\delta_k \to 0$, Lemma A.18(iii) shows that $d_k \to 0$ if the iterates converge to a constrained minimizer $x^\star$ with $\operatorname{grad} F(x^\star) \neq 0$. At such a minimizer, (11) gives the quadratic growth $\langle \operatorname{grad} F(x^\star), \operatorname{Log}_{x^\star} q\rangle \ge \alpha\|\operatorname{grad} F(x^\star)\| d^2(x^\star, q)$, so $x^\star$ uniquely minimizes its own linearization over $\mathcal{C}$. The retraction error bound and the truncation model of Proposition A.17 scale with the step length $\eta_k d_k$; their certification guarantees also require the stated step-size, domain, and rounding-error conditions. A map with a fixed absolute error enclosure ε no longer satisfies the acceptance rule once $\beta_k < \varepsilon$. This failure of certification does not by itself establish infeasibility. Actual feasibility depends on the error direction and on the true clearance of z_{k+1}, which may exceed β_k, because the budget uses the chord-based geometric lower bound together with the progress tolerance rather than the full clearance of the interpolated point (Section 6). Appendix B derives analogous bounds for continuous-time vector-field disturbances from the same margin, where a uniform positive geometric margin yields a uniform disturbance tolerance under the stated regularity assumptions. These continuous-time statements do not transfer to the discrete algorithm without a separate theorem.

Remark C.7 (axis-preserving and off-axis retractions on $\mathrm{SO}(3)$). Geodesic convexity says nothing about retraction curves in general, and Corollary A.9 is the mechanism that controls their endpoint error. In exact arithmetic on $\mathrm{SO}(3)$ with the bi-invariant metric, however, any retraction of the form $R\,\phi(\Omega)$ with ϕ a matrix function of the single skew matrix Ω (Cayley, polar, truncated exponential followed by polar projection) returns a rotation about the axis of Ω, so the result lies *on the geodesic* $\tau \mapsto R\exp(\tau\Omega)$. For the Cayley map and $0 \le \theta := \|\Omega\| < \pi$, the discrepancy is $d(R\,\mathrm{cay}(\Omega), R\exp(\Omega)) = \theta - 2\arctan(\theta/2) \le \theta^3/12$ (unit tests, Appendix D; for $\theta \ge \pi$ the geodesic distance wraps and only the inequality survives), and every anchor-directed update $\Omega = \eta \operatorname{Log}_x q$ with $\eta \in [0,1]$ satisfies this restriction since $\theta \le d(x,q) \le 2\rho < \pi$; for such an update the Cayley-retracted point lies on the segment from the current point to the exact endpoint on the chord to q, hence in $\mathcal{C}$ by geodesic convexity alone (an arbitrary Cayley step that overshoots a chord is not covered). The QR retraction $R \mapsto \mathrm{qf}(R(I+\Omega))$ is not a matrix function of Ω and is off-axis; its error is second order in $\|\Omega\|$, $d(\mathrm{qf}(I+\Omega), \exp\Omega) = O(\|\Omega\|^2)$. We do not have a proof of a uniform constant K on $\{\|\Omega\| \le D\}$; the maximum of the ratio $d(\mathrm{qf}(I+\Omega), \exp\Omega)/\|\Omega\|^2$ over 20,000 random samples is $\hat{K} \approx 0.342$ both on $\|\Omega\| \le 1.1$ (the step domain $\eta_0 \cdot 2\rho \approx 1.05$ of the shifted QR schedule $\eta_k = 2/(k+4)$ of Experiment C) and on $\|\Omega\| \le 2.1$ (an upper bound on the diameter $2\pi/3$ of the 60° ball) (Appendix D), and the experiments use $K = 0.40$. A sampled maximum is not a uniform bound, so the resulting schedule is labelled "calibrated", not certified, throughout; a validated bound (interval arithmetic with a Taylor argument near $\Omega = 0$) would upgrade it. On Stiefel and Grassmann manifolds the QR and polar retractions are off-axis in general and the corollary is the relevant tool once such a bound is available.

Remark C.8 (an approximate exponential map on $\mathrm{SO}(3)$ with off-geodesic endpoint errors). In exact arithmetic, any approximation of $\exp\Omega$ by a matrix function of Ω (truncated series, Padé, scaling and squaring, Runge–Kutta applied to $\dot{R} = R\Omega$ with constant Ω, quaternion propagation) is on-axis by the argument of Remark C.7 once it is mapped back to $\mathrm{SO}(3)$ by an axis-preserving projection (polar projection, quaternion renormalization; a raw truncated series or Runge–Kutta step is not orthogonal), and it is feasible whenever the returned point lies on the minimizing segment from the current iterate to the exact endpoint; floating-point evaluation can break the exact commutation on which this on-axis property rests, which is the subject of Appendix D. Off-geodesic errors arise when the update is computed in a chart. The rotation-vector kinematics $\dot{w} = J_r(w)^{-1}\Omega$, with J_r the right Jacobian of $\mathrm{SO}(3)$ (nonsingular for $\|w\| < 2\pi$) and $w = \operatorname{Log}_{R_c} R$ the exponential coordinate about a fixed base R_c on the principal chart $\|w\| < \pi$, describes the geodesic $R_c \exp(\widehat{w(t)})$ with constant body rate Ω. These kinematics are standard in attitude propagation. When the chart coordinate and the body rate are not collinear, discretization can produce endpoint errors transverse to the exact geodesic, as illustrated by the Euler and Runge–Kutta updates in Appendix D. Chart differences bound geodesic distances, $d(R_c\exp\widehat{w_1}, R_c\exp\widehat{w_2}) \le \|w_1 - w_2\|$, for all $w_1, w_2 \in \mathbb{R}^3$, because the singular values of $J_r(w)$ are 1 and $2|\sin(\|w\|/2)|/\|w\| \le 1$ (continuously extended at $w = 0$); this converts a

step-doubling estimate in the chart into an estimate of $d(\tilde{z}, z)$ (an indicator; a proved enclosure is needed for a certificate, Proposition D.2), and a rounding bound $\|w - \mathrm{fl}(w)\| \le u\|w\| + \sqrt{3}\, q_{\min}/2$ (componentwise round-to-nearest with unit roundoff u and least positive subnormal spacing $q_{\min}$, no overflow or flush-to-zero; remapped by $R_c \exp \widehat{\mathrm{fl}(w)}$ in float64, whose own rounding error is not included) into a coordinate-storage bound. Under an exact decoding map, this estimate bounds the induced geodesic storage perturbation. Experiment C uses it in a coordinate-storage ablation whose float64 remapping error is not included; Experiment H instead uses matrix storage with the full storage and restoration ledger of Lemma D.3 (Appendix D). Certification of the represented endpoint requires a total error bound, including the errors of all remapping and restoration operations actually performed, to be at most β_k.

Remark C.9 (projected dynamics and prox-regularity). The projected vector field $x \mapsto \Pi_{T_{\mathcal{C}}(x)}(-\operatorname{grad} F(x))$ can be discontinuous at $\partial\mathcal{C}$, so its dynamics require a specified solution notion; existence, invariance, uniqueness under Lipschitz $\operatorname{grad} F$, and the relations between Carathéodory, Krasovskii, and normal-cone-inclusion solutions are established under Clarke-regularity and prox-regularity hypotheses by Hauswirth et al. (2021); see also Davis et al. (2025). Geodesic strong convexity supplies a quadratic interior clearance along feasible chords (the ball radii vanish at the chord endpoints, so it is not a uniform interior rolling-ball condition; intersections of two equal-radius balls have corners and are strongly convex). Prox-regularity instead controls local nearest-point geometry and provides a framework for many nonconvex feasible sets; such sets lie outside the present scope.

Example C.10 (attitude keep-in set on $\mathrm{SO}(3)$ as a sublevel set)**.** Let $0 < \rho < \pi/4$, $f(R) = \frac{1}{2}d^2(R, R_0)$, $s = \frac{1}{2}\rho^2$, and let U be the open ball of radius $r' \in (2\rho, \pi/2)$ about R_0; then $Q_s = \{x \in U : f(x) \le s\}$ is the closed ball of radius ρ about R_0. On U, $\operatorname{grad} f(R) = -\operatorname{Log}_R R_0$ and f is geodesically μ-strongly convex and L-smooth with $\mu = \frac{r'}{2}\cot\frac{r'}{2}$ and $L = 1$ (constant curvature $\frac{1}{4}$); the tube of width $\tau = \sqrt{2s} = \rho$ about Q_s lies in U, so Theorem A.20 applies for every such r', and the supremum over $r' \in (2\rho, \pi/2)$ of the resulting moduli is $\frac{1}{2}\cot\rho$, the value compared with α_ρ in Table 6 and Figure 3(b); explicitly $\frac{1}{2}\cot\rho = \alpha_\rho - \frac{1}{4}\tan(\rho/2) < \alpha_\rho$, so this route is valid but not sharp. Since the ball is closed, the ball inclusions of Definition 2.3 pass to this limit, so $\frac{1}{2}\cot\rho$ is itself a valid modulus by this route. For the kinematics $\dot{R} = \operatorname{Log}_R R_0 + e(R)$ with $e(R) \in T_R\mathrm{SO}(3)$, Corollary B.1 with the fixed interior anchor R_0 gives the uniform disturbance budget $\|e(R)\|_R \le \alpha_\rho\rho^2$ on the boundary (a time-dependent disturbance needs the Carathéodory bounds stated at the start of Appendix B); in body coordinates, $\dot{R} = R\widehat{\Omega}$ with $\Omega = (R^\top \operatorname{Log}_R R_0)^\vee + e_b$ and $e = R\widehat{e_b}$, and the half-trace metric makes this trivialization an isometry, so $\|e\|_R = \|e_b\|_2$.

Remark C.11 (equality constraints and full-dimensional scope). For $\mathcal{C} = \{h_a \le 0,\ g_b = 0\}$ under a constraint qualification, $T_{\mathcal{C}}(p) = \{v : \langle \operatorname{grad} h_a(p), v\rangle \le 0\ (a \text{ active}),\ \langle \operatorname{grad} g_b(p), v\rangle = 0\}$; a set with empty interior is all boundary, and tangency is required at every point. A feasible set with empty interior and at least two distinct points has no positive strong-convexity modulus in the sense of Definition 2.3 in the ambient manifold (a singleton satisfies the definition vacuously); if $\{g_b = 0\}$ is itself a suitable manifold, the problem must be reformulated on it, with the inequality constraints defining a full-dimensional subset there. Products of at least two nontrivial rotation balls also lack a positive isotropic modulus. Under the product metric, a segment that holds one factor at a boundary point while moving another factor remains on the boundary of the product. The isotropic certificate of Definition 2.3 therefore cannot be applied directly to the whole product. A discrete product extension requires a separate analysis, using factorwise endpoint-error bounds or a certified lower bound on the product clearance together with a convergence-compatible allocation of the total error; the continuous-time tangency test of Appendix B decouples factorwise (Remark B.2).

D Numerical contracts, precision analysis, and experimental details

Roadmap. This appendix has five parts. Section D.1 defines the represented quantities, the evidence classes, and the primitive-error model (Assumption D.1). Section D.2 derives the chart-Euler enclosure (Proposition D.2) and the storage ledger (Lemma D.3) and describes the acceptance machinery: the reserve policy, the comparison arithmetic, the complete contract (Table 4), the operational routine (Algorithm 3), and the independent checker. Section D.3 analyses the economical

policy (Proposition D.4) and the fixed-N sufficient policy with its precision theorem (Theorem D.5). Section D.4 states the shared experimental protocol. The studies then follow in the order of their subsections: the geometric checks A and B, the native stress grid G, the nonradial comparisons H, the baseline diagnostics C, the precision studies D and E, and the represented exact-oracle instance F. Each study distinguishes its mathematical inputs, its numerical observations, and its remaining limitations. Section 6 names each study by its question and gives its letter in parentheses; this appendix uses the letters. Table 1 associates each experiment with the question it answers, the result it bears on, its instances, sample unit, and horizon, and its oracle and evidence status. Each study either audits the implementation of a proved statement or investigates a question about the implementation that the theory leaves open.

Exp.	Question it answers	Result it bears on	Instances, sample unit, and horizon	Oracle and evidence status
A	Do independent evaluations agree with the rotation-ball identities, and how closely do sampled and limiting configurations approach the proved sharp bounds?	Prop. 5.2	six radii, 20,000 configurations and 20,000 chords each; static	identities proved; agreement checked
B	Does the quartic sublevel set have fourth-order midpoint clearance?	Prop. A.22	one quartic set; static	counterexample proved; illustrated in Figure 3(d)
C	How does Algorithm 2 behave under native and injected endpoint error?	Alg. 2, Prop. D.2	50 seeds, boundary-active and interior; 300 iterations	numerical oracle; nearly radial; baseline diagnostics reported in this appendix
D	What happens past the float64 enclosure floor, and what does escalation cost?	Prop. D.4	10 boundary-active seeds; 1,500 iterations	economical policy, shift 3, numerical oracle; outside Theorem D.5
E	What does the fixed-N sufficient policy cost per iteration?	Thm. D.5	5 boundary-active seeds; 300 iterations	fixed-N policy, shift 3, numerical oracle; cost diagnostic outside the theorem
F	Does the sufficient policy meet its precision bound on a theorem-aligned instance?	Thm. D.5	one represented axis-invariant problem, four targets; to the first ε-accurate iterate	exact oracle, shift 4; conditional primitives; zero Euler truncation
G	Does native error produce infeasible represented points, and what does the ledger decide?	Cor. 3.2, Lemma D.3	216 declared near-boundary cases per kernel; one step	conditional ledgers; calibrated QR bound
H	Feasibility, accuracy, and cost against membership guards, projection, and precision controls	Alg. 2	40 trajectories per treatment (10 seed bundles, 2 separations, 2 starts); 300 iterations	numerical reference oracle; shared-proposal comparison kept separate

Table 1: The studies, with the question each answers, the result it bears on, its instances, sample unit, and horizon, and its oracle and evidence status. Section 6 reports A, B, and D–H; the intervals of H resample the 10 seed bundles.

D.1 Represented quantities, evidence classes, and the primitive model

Coordinates. Throughout this appendix a tangent vector $R\widehat{w} \in T_R\mathrm{SO}(3)$ is identified with its body-coordinate vector $w \in \mathbb{R}^3$; under the half-trace metric this identification is an isometry, $\|R\widehat{w}\|_R = \|w\|$. In numerical formulas $\mathrm{Log}_R Q$ denotes the coordinate vector $(\log R^\top Q)^\vee$ and $\mathrm{Exp}_R(w) = R\exp\widehat{w}$; the coordinate of R in the exponential chart about the ball centre R_c is $w = \mathrm{Log}_{R_c} R$ (Section D.4 lists the remaining conventions).

Objects being certified. A stored float64 (or b-bit) matrix A_k represents the rotation $x_k = \Pi(A_k)$, its polar factor in exact arithmetic. (On every run of Experiment C, the only study that records them, the orthogonality residuals $\|A^\top A - I\|_F$ of the stored matrices stay below $3 \cdot 10^{-14}$.) The stored oracle output Q_k represents $q_k = \Pi(Q_k)$; $\Omega_k := \mathrm{Log}_{x_k} q_k$ is the exact body rate of the represented feasible chord and $\tilde{\Omega}_{k,b}$ the rate computed at precision b. T_k is the binary step actually used. We model its rounding by $T_k = \frac{2}{k+m}(1+\vartheta_k)$ with $|\vartheta_k| \le u_{53}$ ($u_b = 2^{-b}$ and the constants $c_{\mathrm{in}}, c_\Omega, \dots$ below are the rounding-model quantities of Assumption D.1). The counter $k+m$ is an exact integer, and T_k is the correctly rounded binary value of the exact rational $2/(k+m)$, whether or not $k+m$ is representable in 53 bits. Experiment F rounds the exact rational directly; on the horizons of Experiments C–E, G, and H, $k+m < 2^{53}$, so one correctly rounded division by the exactly represented denominator gives the same value, as a unit test confirms. The rounded-schedule extension of Theorem 4.1(ii) (Lemma A.12) gives the same $O(1/k)$ rate with the constant

$A_s = \max\{m\, h_0, 4(Q_\kappa + \delta)(1 + u_{53})^2/(1 - 2u_{53})\}$, $\kappa = c$; feasibility budgets and map enclosures use T_k itself. As in Section 4, $z_{k+1} := x_k \exp(T_k \widehat{\Omega}_k)$ is the exact endpoint that Corollary 3.2 needs, $\bar{z}_{k+1,b} := x_k \exp(T_k \widehat{\Omega}_{k,b})$ the exact endpoint of the computed-rate geodesic, A_{k+1} the stored output matrix (the computed remap, or the restored matrix), and $\tilde{z}_{k+1} = \Pi(A_{k+1})$ the represented output, which becomes x_{k+1} on acceptance. The certificate bounds $d(\tilde{z}_{k+1}, z_{k+1})$ under the stated numerical model; the endpoint errors reported in the experiments are numerical estimates of this distance. In Experiments D–F they are 400-bit distances to references that recompute the exact endpoint of the represented chord (Sections D.9 and D.10), and in Experiment C the independent checker of Section D.2 recomputes this endpoint in 400-bit arithmetic for every accepted step of the enclosure-certified method. The per-iteration error curves of Experiment C (Figure 6) and the errors of Experiments G and H are instead distances to the represented float64 reference endpoint of the same step, evaluated in float64 in Experiment C and in 400-bit arithmetic in Experiments G and H, and the clearances recorded in Experiments G and H are those of that reference endpoint; the reference endpoint's own distance to the exact endpoint is at most $(c_{\mathrm{out}} + T_k c_\Omega) u_{53} < 10^{-14}$ under Assumption D.1, so each recorded error and clearance is within this allowance of its exact counterpart. This proxy error is invisible at the reported magnitudes, except for the chart-Euler proposals of the boundary-active regime of Experiment H, more than half of whose recorded errors lie below this bound; the affected ratios of Table 15 and Figure 4 are resolution-limited and are marked there. The numerical reference oracle used outside Experiment F returns targets on the sphere of radius $\rho(1 - 2^{-40})$ about R_c, so that under the model below the represented target is feasible (its clearance is at least $\rho 2^{-40} - (\rho + c_{\mathrm{out}}) u_{53} > 0$, the extra term covering the rounding of the scaled direction); the change in the oracle value is below 10^{-11} and is absorbed in $\widehat{\delta}_k$. Experiment F instead returns the stored boundary target that defines its represented ball, as specified in its construction.

Roles and status classes. Experiments A and B are identity and sharpness checks of closed-form statements, Experiments C–E, G, and H are diagnostics of the acceptance rules, and Experiment F tests the sufficient policy on a theorem-aligned instance. We distinguish four classes of evidence for numerical components, with labels chosen so that they cannot be confused with the experiment letters or the appendix letters. Class **[AN]** (analytic) comprises analytic inequalities and exact-arithmetic bounds, including the geometric budget, the truncation term of Proposition D.2, and the storage perturbation of Lemma D.3. Class **[CO]** (conditional) comprises implementation enclosures conditional on the rounding model of Assumption D.1, including the rounding term of Proposition D.2, the nominal float64 endpoint, and the restoration allowance. Class **[DI]** (diagnostic) comprises empirical numerical checks without a proved rounding enclosure, such as a float64 membership test or a 400-bit recomputation. Class **[CA]** (calibrated) comprises calibrated or heuristic estimates, including the QR constant and a step-doubling indicator. A method carries several labels when its guarantee combines components of different status; a guarantee combining analytic and assumed components remains conditional on the assumed components, so an [AN+CO] method has a conditional total guarantee. The auxiliary labels **[ung.]** (unguarded) and **[stress]** (an injected error of prescribed size) describe the treatment, not the evidence. The float64 *reference exponential* approximates the exact geodesic update and is itself covered only by the rounding model [CO]. The Cayley retraction is feasible by a different route, namely on-geodesic under-rotation in exact arithmetic (Remark C.7) [AN], which does not exercise the endpoint-error budget. The chart-Euler map with the a priori enclosure of Proposition D.2 has a proved truncation term [AN] and a conditional rounding term [CO], and every accepted step is checked independently offline in 400-bit floating-point arithmetic (the independent checker of Section D.2) [DI]. Reduced-precision chart coordinates carry a proved storage bound that excludes the float64 remap, so "guarded coordinate rounding" is a storage-only ablation [AN]. Reduced-precision *matrix* storage with polar restoration is guarded by the full representation ledger of Lemma D.3 (storage perturbation [AN]; nominal endpoint, residual evaluation, and restoration [CO]). The step-doubling guard of the adaptive integrators is an error indicator [CA]. Membership backtracking uses a direct float64 membership test [DI] rather than an a priori endpoint-error certificate (a validated membership test would be a legitimate feasible baseline; Experiment H compares float64 and high-precision floating membership tests). The QR constant is calibrated [CA]. Outside Experiment F, the linear oracle and the reference optimum F_{ref} are high-accuracy numerical references rather than certified global minimizers, so conclusions requiring their exactness are diagnostics; Experiment F establishes oracle optimality and the optimum analytically for its represented instance, and the primitive model remains conditional in that experiment as well.

Assumption D.1 (rounding model at precision b)**.** We use binary floating-point arithmetic with $b \geq 53$ mantissa bits, unit roundoff $u_b = 2^{-b}$, and round-to-nearest, and assume that no overflow or underflow occurs. The implementation satisfies the following conditions.

(i) The computed chart entry $\tilde{w}_0$ of the represented point satisfies $\|\tilde{w}_0 - \mathrm{Log}_{R_c} x\| \leq c_{\mathrm{in}} u_b$.

(ii) The computed body rate satisfies $\|\tilde{\Omega} - \mathrm{Log}_x q\| \leq c_\Omega u_b$.

(iii) One floating evaluation $\hat{f}(w)$ of $f(w) = J_r(w)^{-1}\tilde{\Omega}$ satisfies the normwise relative bound $\|\hat{f}(w) - f(w)\| \leq c_f u_b \|f(w)\|$.

(iv) For an exactly represented step h, the componentwise multiply-then-add update satisfies the absolute bound $|\mathrm{fl}(w + h\hat{f}) - (w + h\hat{f})| \leq u_b|w| + (2u_b + u_b^2)h|\hat{f}|$ (this follows for round-to-nearest from $\mathrm{fl}(w + \mathrm{fl}(h\hat{f})) = (w + h\hat{f}(1+\epsilon_1))(1+\epsilon_2)$ with $|\epsilon_i| \leq u_b$; a relative bound on the final update fails under cancellation, e.g. $w = \frac{1}{2}$, $h = \mathrm{fl}(\frac{1}{3})$, $\hat{f} = -\frac{3}{2}$ has exact value 2^{-55} but computed value 0, a case included in the unit tests).

(v) Three output-construction routines are covered by the following bounds: (v-a) the represented point of the computed remap $R_c \exp \widehat{w}$ is within $c_{\mathrm{out}} u_b$ of $R_c \exp \widehat{w}$; (v-b) the complete reference-endpoint routine, which takes the stored matrix A of a represented point $x = \Pi(A)$, a binary step $T \in (0, 1]$, and a represented body rate $\tilde{\Omega}$ of a feasible chord, and returns the stored matrix A_{ref} of the computed product of A with the exponential of the rounded scaled rate $\mathrm{fl}(T\tilde{\Omega})$, satisfies $d\big(\Pi(A_{\mathrm{ref}}), x \exp(T\widehat{\widehat{\Omega}})\big) \leq c_{\mathrm{out}} u_b$, where the target uses the exact real product $T\tilde{\Omega}$ and the bound covers the stored base, the rounding of the scaled rate, the exponential, the matrix product, and the storage of the output; and (v-c) the represented output of the float64 polar restoration Π of a matrix B with $\det B > 0$ and $\|B^\top B - I\|_F \leq 3 \cdot 10^{-3}$ is within $c_{\mathrm{pol}} u_{53}$ of $\Pi(B)$, the domain being established for every matrix that is restored.

(vi) The floating evaluations of the budget and of every enclosure component that is compared (E_n, E_n^{tr}, E_n^{fl}, E^{ref}, E_{total}) have relative error below $\tau := 2^{-40}$, and the final comparison $(1+\tau)\widehat{E} \leq (1-\tau)\widehat{B}$ is performed exactly, in rational arithmetic on the represented binary values (nonfinite inputs are rejected).

(vii) Domain tests are performed separately from the final comparison and use the represented values. A chart node is used only after an exact rational comparison confirms $\sum_i w_i^2 \leq \rho'^2$, with both its binary coordinates and the binary chart radius interpreted exactly. The resolved-chord predicate is likewise decided exactly from the represented binary rate. It succeeds if and only if $\sum_{i=1}^3 \tilde{\Omega}_{b,i}^2 \geq (3c_\Omega u_b)^2$, with both sides compared in rational arithmetic; the threshold $3c_\Omega u_b = 192u_b$ is exactly representable. This predicate is equivalent to $\|\tilde{\Omega}_b\| \geq 3c_\Omega u_b$, and its decision does not use a rounded norm value. The reserve $\rho + Tc_\Omega u_b < \rho'$ is a fixed inequality between represented constants, and polar restoration is invoked only after a sound domain predicate establishes, by exact rational arithmetic on the represented entries of the stored matrix B, that $\det B > 0$ and $\sum_{i,j}[(B^\top B - I)_{ij}]^2 \leq (3/1000)^2$ (failure returns a numerical-domain status without invoking the primitive, and the a priori bound $2\delta_A + \delta_A^2$ of the ledger below is recorded as an analytic upper bound but does not decide the domain); for the float64 residual $s_A = \|A^\top A - I\|_F$ of a raw matrix, the computed value satisfies $|\hat{s}_A - s_A| \leq c_{\mathrm{res}} u_{53}$ on the stated input domain, and the exact real sum $s_A^+ := \hat{s}_A + c_{\mathrm{res}} u_{53} \geq s_A$ is the quantity used in every ledger; the floating evaluation of the ledger is covered by the relative-error contract in (vi).

Constants and calibration. The constants used are $c_{\mathrm{in}} = 16$, $c_\Omega = 64$, $c_f = 32$, $c_{\mathrm{out}} = 8$, $c_{\mathrm{pol}} = 128$, $c_{\mathrm{res}} = 64$. These are modelling constants. A calibration against 400-bit reference evaluations used 300 random configurations for the float64 primitives and 100 per precision for the b-bit kernels with $b \in \{53, 64, 96, 128\}$, with chords up to 2ρ. In the test-script run (seed 0; the log is shipped with the code), the largest absolute errors observed over these calibrations for the chart entry, the body rate, and the remap were 4.8, 27.6, and 1.7 units of u_b; for the vector-field evaluation, the largest observed normwise relative error $\|\hat{f} - f\|/\|f\|$ was 3.3 units of u_b (3.4 in a separate 64-bit calibration log), matching the form of (iii). Sampling the residual domain of the float64 polar-restoration primitive

on 300 perturbed rotations gave a largest observed error of 47.2 units of u_{53}. The sampled errors lie below the chosen constants (margin at least 2.3); this agreement does not establish a uniform bound and is not a floating-point proof of (i)–(iii) and (v). The remap samples draw the stored base matrix at random and pass an unscaled vector, so they exercise the reference-endpoint routine of (v-b) except for the scaling of the rate, whose componentwise rounding moves the endpoint by at most $T\|\tilde{\Omega}\|u_b$ (at most about $2.1u_b$ for the chords of the 60° ball). The complete routine was not sampled separately, and its bound with the shared constant c_{out} is a hypothesis.

What is proved and what is assumed. Item (iv) is proved. The final comparison in (vi) and the chart-node and restoration-domain predicates in (vii) use exact rational arithmetic on represented binary inputs. The uniform relative-error bound for the enclosure and budget evaluations in (vi), the residual-evaluation bound in (vii), and the primitive bounds (i)–(iii) and (v) remain hypotheses of the conditional analysis (Table 2). The comparison test in (vi) is safe because it is equivalent to $\widehat{E}/(1-\tau) \le \widehat{B}/(1+\tau)$, whose sides enclose E from above and B from below; the individual factors are not one-sided enclosures. The uniform relative error in (vi) presumes stable evaluation of $\phi(z) = (e^z - 1)/z$ and of the small-angle series of ζ; the constants are asserted on chords shorter than $2\rho < \pi$, away from the cut locus of the logarithm.

contract component	status
multiply-then-add absolute bound (iv)	proved
final comparison $(1+\tau)\widehat{E} \le (1-\tau)\widehat{B}$ (vi)	exact rational implementation
chart-node predicate $\sum_i w_i^2 \le \rho'^2$ (vii)	exact rational implementation
resolved-chord predicate $\sum_i \tilde{\Omega}_{b,i}^2 \ge (3c_\Omega u_b)^2$ (vii)	exact rational implementation
domain-reserve test $\rho + T_k c_\Omega u_b < \rho'$ (hypothesis of Proposition D.2)	exact rational implementation
restoration-domain predicate $\det B > 0$, $\|B^\top B - I\|_F^2 \le (3/1000)^2$ (vii)	exact rational implementation
relative accuracy of the enclosure and budget evaluations (vi)	assumed uniform bound
chart entry, body rate, vector field, remap, restoration (i)–(iii), (v-a), (v-c)	assumed uniform bounds, calibrated by samples
complete reference-endpoint routine (v-b)	assumed uniform bound with shared c_{out}; complete routine not sampled separately
residual evaluation $\lvert\hat{s}_A - s_A\rvert \le c_{\mathrm{res}} u_{53}$ (vii)	assumed absolute bound
400-bit recomputation of accepted steps	empirical high-precision consistency check

Table 2: Status of the components of the numerical contract. The table distinguishes proved bounds, exact decisions on represented inputs, and hypotheses of the conditional analysis.

D.2 Endpoint-error enclosures and the acceptance machinery

Notation specific to this appendix: θ is the truncation share of a budget, ϕ the function $(e^z-1)/z$ in Proposition D.2 and the cluster separation in Experiment H, B the analytical budget, s the chart reserve $\rho' - \rho$, and γ the comparison factor $((1-\tau)/(1+\tau))^2$; these letters have other meanings in the main text.

Proposition D.2 (a priori enclosure for chart Euler on SO(3))**.** *Let $0 < \rho < \pi/2$, $\rho < \rho' < \pi$, $R_c \in \mathrm{SO}(3)$, $x, q \in \mathcal{C}_\rho$, $\Omega := \mathrm{Log}_x q$ (so the exact geodesic $x\exp(t\widehat{\Omega})$, $0 \le t \le 1$, is the chord to q and stays in $\mathcal{C}_\rho$), let $T \in (0,1]$ be the positive binary step actually used and n a power of two, so that $h = T/n$ is exactly represented. In the exponential chart about R_c the geodesic is $w(t) = \mathrm{Log}_{R_c}(x\exp(t\widehat{\Omega}))$ with $\dot{w} = f(w) := J_r(w)^{-1}\Omega$, $J_r(w)^{-1} = I + \frac{1}{2}\widehat{w} + \zeta(\|w\|)\widehat{w}^2$, and $\zeta(\theta) = \theta^{-2} - (1+\cos\theta)/(2\theta\sin\theta) = \sum_{k\ge1} |B_{2k}|\theta^{2k-2}/(2k)!$, $\zeta(0) = \frac{1}{12}$ (the series is used near 0). Let $\tilde{\Omega}$ be the computed body rate, let $\tilde{w}_0, \dots, \tilde{w}_n$ be the explicit Euler iterates for $\dot{w} = J_r(w)^{-1}\tilde{\Omega}$ with step $h = T/n$ computed in b-bit floating point from the computed entry $\tilde{w}_0$, and assume the domain reserve $\rho + Tc_\Omega u_b < \rho'$ and $\|\tilde{w}_i\| \le \rho'$ for all i (checked online; on failure the routine returns a status and evaluates no enclosure). With $\kappa_1 := (\rho'/2)/\sin(\rho'/2)$, $\Lambda := \frac{1}{2} + 2\rho'\zeta(\rho') + \rho'^2\zeta'(\rho')$, and $L := \Lambda\|\tilde{\Omega}\|$, the represented returned point $\tilde{z} = \Pi(A_{\mathrm{out}})$, A_{out} the stored remap matrix with*

$d(\Pi(A_{\mathrm{out}}), R_c \exp \widehat{\tilde{w}_n}) \le c_{\mathrm{out}} u_b$, *satisfies, under Assumption D.1,*

$$d\big(\tilde{z}, \mathrm{Exp}_x(T\Omega)\big) \le E_n := \underbrace{\tfrac{h}{2}\,\kappa_1 \|\tilde{\Omega}\|\,(e^{LT}-1)}_{E_n^{\mathrm{tr}},\ \textit{proved}} + \underbrace{u_b\, K_{\mathrm{fl}}(n)}_{E_n^{\mathrm{fl}},\ \textit{conditional}},$$

$$K_{\mathrm{fl}}(n) := e^{LT}\big(c_{\mathrm{in}} + n(2\rho' + (3+c_f)h\kappa_1\|\tilde{\Omega}\|)\big) + c_{\mathrm{out}} + Tc_\Omega.$$

$E_n^{\mathrm{tr}} \to 0$ as $n \to \infty$ and E_n^{fl} is nondecreasing in n. On trials passing the domain checks, the routine with $n = 2^j$ supplies the candidate and error decomposition required by Algorithm 2; a failed check returns a numerical-domain status without an enclosure.

Uniform envelope and intermediate nodes. *With the envelope $\overline{D} := D + c_\Omega u_{53} \ge \|\tilde{\Omega}\|$ and $\overline{C}_E := \frac{1}{2}\kappa_1 \Lambda\, \phi(\Lambda \overline{D})$, $\phi(z) := (e^z - 1)/z$, $\phi(0) = 1$, one has $E_n^{\mathrm{tr}} \le \overline{C}_E\,(T\|\tilde{\Omega}\|)^2/n$, and every intermediate node satisfies the same bound $\|\tilde{w}_i - \bar{w}(t_i)\| \le E_n^{\mathrm{tr}} + u_b K_{\mathrm{fl}}(n) - c_{\mathrm{out}} u_b - Tc_\Omega u_b$.*

Proof. Constants. On the plane orthogonal to w, $J_r(w)^{-1}$ has eigenvalues $(\theta/2)\cot(\theta/2) \pm i\theta/2$, $\theta = \|w\|$, of modulus $(\theta/2)/\sin(\theta/2)$, and eigenvalue 1 along w; hence $\|J_r(w)^{-1}\| \le \kappa_1$ on $\|w\| \le \rho'$ and $\|f\| \le \kappa_1\|\tilde{\Omega}\|$. Since $(\theta/2)\cot(\theta/2) = 1 - \sum_{k\ge1} |B_{2k}|\theta^{2k}/(2k)!$, ζ has positive coefficients, so ζ and ζ' are nonnegative and increasing on $[0,\pi)$. From $f(w) = \tilde{\Omega} + \frac{1}{2} w \times \tilde{\Omega} + \zeta(\theta)\, w \times (w \times \tilde{\Omega})$, $Df(w)[v] = \frac{1}{2} v \times \tilde{\Omega} + \zeta(\theta)\big(v \times (w \times \tilde{\Omega}) + w \times (v \times \tilde{\Omega})\big) + \zeta'(\theta)\frac{\langle w, v\rangle}{\theta}\, w \times (w \times \tilde{\Omega})$, so $\|Df(w)\| \le \|\tilde{\Omega}\|(\frac{1}{2} + 2\theta\zeta(\theta) + \theta^2\zeta'(\theta)) \le L$ on the convex ball $\|w\| \le \rho'$; f is L-Lipschitz there, and $\|\ddot{w}\| = \|Df(w)f(w)\| \le L\kappa_1\|\tilde{\Omega}\|$ along any solution in the ball.

Domain. The exact chord $x\exp(t\widehat{\Omega})$, $0 \le t \le T$, lies in $\mathcal{C}_\rho$. The exponential map of $\mathrm{SO}(3)$ is globally nonexpansive for the adopted metric. Its differential at a tangent vector of norm θ has singular values 1 and $|\sin(\theta/2)|/(\theta/2) \le 1$, the latter extended continuously at $\theta = 0$ (Remark C.8). Integrating this differential bound along the straight segment from $t\Omega$ to $t\tilde{\Omega}$ in $T_x\mathrm{SO}(3)$ gives $d(x\exp(t\widehat{\tilde{\Omega}}), x\exp(t\widehat{\Omega})) \le t\|\tilde{\Omega} - \Omega\| \le Tc_\Omega u_b$ by (ii). Consequently $d(R_c, x\exp(t\widehat{\tilde{\Omega}})) \le \rho + Tc_\Omega u_b < \rho' < \pi$ by the domain reserve, so $\bar{w}(t) := \mathrm{Log}_{R_c}(x\exp(t\widehat{\tilde{\Omega}}))$ is well defined on $[0,T]$, solves $\dot{\bar{w}} = J_r(\bar{w})^{-1}\tilde{\Omega}$ with $\bar{w}(0) = w(0)$, and satisfies $\|\bar{w}(t)\| < \rho'$; the computed iterates lie in the same ball by hypothesis, so the Lipschitz bound applies to every pair $(\tilde{w}_i, \bar{w}(t_i))$.

Recurrence for the computed iterates. Let $e_i := \|\tilde{w}_i - \bar{w}(t_i)\|$. One Euler step in exact arithmetic from $\tilde{w}_i$ differs from $\bar{w}(t_{i+1})$ by at most $(1+hL)e_i + \frac{h^2}{2}L\kappa_1\|\tilde{\Omega}\|$ (Lipschitz propagation plus the local truncation error $\frac{h^2}{2}\sup\|\ddot{\bar{w}}\|$). By (iii) the computed right-hand side is $\hat{f} = f + \delta$ with $\|\delta\| \le c_f u_b \|f\|$, so $h\|\delta\| \le c_f u_b h\|f\|$ and $\|\hat{f}\| \le (1 + c_f u_b)\|f\|$; by the absolute bound (iv) the floating update adds at most $u_b\|\tilde{w}_i\| + (2u_b + u_b^2)h\|\hat{f}\|$ (the componentwise bound implies the Euclidean one). Hence the local arithmetic perturbation is at most $\varepsilon_i := u_b\big(\|\tilde{w}_i\| + [c_f + (2+u_b)(1+c_f u_b)]h\|f\|\big) \le u_b\big(2\rho' + (3+c_f)h\kappa_1\|\tilde{\Omega}\|\big)$, using $c_f + (2+u_b)(1+c_f u_b) \le c_f + 3$ for $c_f = 32$ and $b \ge 53$ and the bound $\|\tilde{w}_i\| \le 2\rho'$ in the arithmetic-error estimate (the factor 2 is spare room; the chart-node check gives $\|\tilde{w}_i\| \le \rho'$). Thus $e_{i+1} \le (1+hL)e_i + \frac{h^2}{2}L\kappa_1\|\tilde{\Omega}\| + \varepsilon_i$, and by induction, for every $i \le n$, $e_i \le (1+hL)^i e_0 + \frac{(1+hL)^i - 1}{hL}\big(\frac{h^2}{2}L\kappa_1\|\tilde{\Omega}\| + \varepsilon\big) \le e^{LT}c_{\mathrm{in}}u_b + \frac{h}{2}\kappa_1\|\tilde{\Omega}\|(e^{LT} - 1) + ne^{LT}\varepsilon$, using $(1+hL)^i \le e^{LT}$, $((1+hL)^i - 1)/(hL) \le ne^{LT}$, and $e_0 \le c_{\mathrm{in}}u_b$ by (i) (when $L = 0$ the quotient $((1+hL)^i - 1)/(hL)$ is read as its limit i, which is the direct sum of the zero-field recurrence); this is the intermediate-node bound.

Back to $\mathrm{SO}(3)$. $d(R_c\exp\widehat{w_1}, R_c\exp\widehat{w_2}) \le \|w_1 - w_2\|$ (Remark C.8) converts e_n into a geodesic distance between $R_c\exp\widehat{\tilde{w}_n}$ and $x\exp(T\widehat{\tilde{\Omega}})$; (v-a) adds $c_{\mathrm{out}}u_b$ for the represented remap, and the body-rate discrepancy adds $Tc_\Omega u_b$ (domain step). Summing gives E_n. Finally, with $h = T/n$ and $LT = \Lambda T\|\tilde{\Omega}\|$, $E_n^{\mathrm{tr}} = \frac{1}{2n}\kappa_1\Lambda\,\phi(\Lambda T\|\tilde{\Omega}\|)\,(T\|\tilde{\Omega}\|)^2$, and ϕ is increasing with $T\|\tilde{\Omega}\| \le \|\tilde{\Omega}\| \le \|\Omega\| + c_\Omega u_b \le \overline{D}$ by (ii), which gives the decay model with $\overline{C}_E$. □

Lemma D.3 (storage followed by exact polar decoding)**.** *Let $R \in \mathrm{SO}(3)$ be stored entrywise in a binary floating-point format with unit roundoff u and least positive subnormal spacing $q_{\min}$ (round-to-nearest, no overflow or flush-to-zero), let $\delta_F := \sqrt{3}\,u + \frac{3}{2}q_{\min} < 1$, and let Π denote the*

exact polar projection onto $\mathrm{SO}(3)$. *Then* $d(\Pi(\mathrm{fl}(R)), R) \le \frac{\pi}{\sqrt{2}}\delta_F$*; for half precision (*$u = 2^{-11}$, $q_{\min} = 2^{-24}$*) the bound is* $\approx 1.88 \cdot 10^{-3}$ *rad, for single precision* $2.29 \cdot 10^{-7}$ *rad. The bound also has a raw-matrix form. Let* $A \in \mathbb{R}^{3\times 3}$ *have* $\det A > 0$ *and represented rotation* $R_A := \Pi(A)$*, let* $r_A \ge \|A - R_A\|_F$*, and put* $\delta_A := r_A + u(\sqrt{3} + r_A) + \frac{3}{2}q_{\min}$*. If* $\delta_A < 1$ *then, for* $B = \mathrm{fl}(A)$*,* $\|B - R_A\|_F \le \delta_A$*,* $\|B^\top B - I\|_F \le 2\delta_A + \delta_A^2$*, and* $d(\Pi(B), R_A) \le \frac{\pi}{\sqrt{2}}\delta_A$*; a computable choice is* $r_A = \|A^\top A - I\|_F/(1 + \sqrt{1 - \|A^\top A - I\|_F})$ *when* $\|A^\top A - I\|_F < 1$*. For chart coordinates,* $d(R_c \exp\widehat{\mathrm{fl}(w)}, R_c \exp\widehat{w}) \le u\|w\| + \frac{\sqrt{3}}{2}q_{\min}$.

Proof. $\|\mathrm{fl}(R) - R\|_F \le u\|R\|_F + 3q_{\min}/2 = \delta_F < 1$, so $\mathrm{fl}(R) = R(I + R^\top(\mathrm{fl}(R) - R))$ is nonsingular with positive determinant, and $\Pi(\mathrm{fl}(R))$ is the nearest rotation to $\mathrm{fl}(R)$ in the Frobenius norm; hence $\|\Pi(\mathrm{fl}(R)) - R\|_F \le \|\Pi(\mathrm{fl}(R)) - \mathrm{fl}(R)\|_F + \delta_F \le 2\delta_F$. For rotations, $\|Q - R\|_F = 2\sqrt{2}\sin(d(Q,R)/2)$, so $d(Q,R) = 2\arcsin(\|Q - R\|_F/(2\sqrt{2})) \le \frac{\pi}{2\sqrt{2}}\|Q - R\|_F$. For the raw-matrix form, the triangle inequality gives $\|B - R_A\|_F \le \|B - A\|_F + \|A - R_A\|_F \le u\|A\|_F + \frac{3}{2}q_{\min} + r_A \le \delta_A$, since $\|A\|_F \le \sqrt{3} + r_A$; the same argument about R_A gives the distance bound, and $B^\top B - I = R_A^\top E + E^\top R_A + E^\top E$ with $E = B - R_A$ gives the residual bound. To obtain a computable r_A, let σ_i be the singular values of A. Then $\|A - \Pi(A)\|_F^2 = \sum_i(\sigma_i - 1)^2 \le \sum_i(\sigma_i^2 - 1)^2/(1 + \sigma_{\min})^2$ and $\sigma_{\min} \ge \sqrt{1 - \|A^\top A - I\|_2} \ge \sqrt{1 - \|A^\top A - I\|_F}$. The coordinate bound is Remark C.8 with the componentwise rounding error $u|w_i| + q_{\min}/2$. (The bounds are checked by unit tests on 20,000 sampled rotations and 3000 raw matrices; the observed maxima for exact rotations are $3.3 \cdot 10^{-4}$ rad (half) and $4.3 \cdot 10^{-8}$ rad (single) against the bounds $1.88 \cdot 10^{-3}$ and $2.29 \cdot 10^{-7}$.) □

The lemma is a perturbation bound for low-precision storage followed by exact polar decoding. The implemented restoration workflow stores the *raw* float64 reference endpoint matrix A (whose represented rotation is $R_A = \Pi(A)$) and restores it by a float64 SVD, so its representation ledger, relative to the exact endpoint z_{k+1} of the represented chord, is

$$E_{\mathrm{total}} \le (c_{\mathrm{out}} + T_k c_\Omega)u_{53} + \tfrac{\pi}{\sqrt{2}}\delta_A + c_{\mathrm{pol}}u_{53},$$

where the three terms are the nominal-endpoint error, the storage perturbation, and the restoration error. The first satisfies $d(R_A, z_{k+1}) \le (c_{\mathrm{out}} + T_k c_\Omega)u_{53}$ under the rate bound (ii) and the reference-endpoint bound (v-b) of Assumption D.1, by the chain displayed for the reference fallback in the reserve-policy paragraph below (the computed-rate discrepancy enters here); the second uses δ_A from the raw-matrix form with $r_A = s_A^+/(1 + \sqrt{1 - s_A^+})$ constructed only after $0 \le s_A^+ < 1$ and $\det A > 0$ have been established for the padded residual $s_A^+ = \hat{s}_A + c_{\mathrm{res}}u_{53}$ (otherwise the ledger returns a numerical-domain status before the square root is evaluated); the third is used only after the exact rational predicate of Assumption D.1(vii), applied to the stored matrix B, establishes $\det B > 0$ and $\|B^\top B - I\|_F \le 3 \cdot 10^{-3}$; the analytic upper bound $2\delta_A + \delta_A^2$ is recorded in the ledger but does not decide this domain. For fp16 storage of a near-rotation, $\delta_A \approx 8.5 \cdot 10^{-4}$, giving a conservative residual bound of about $1.7 \cdot 10^{-3}$; the restoration contract applies on the domain with residual at most $3 \cdot 10^{-3}$, and the exact rational predicate checks membership in that domain before restoration. The guarded storage method accepts a low-precision branch only when the safeguarded comparison $(1+\tau)\widehat{E}_{\mathrm{total}} \le \underline{\beta}_k$ passes ($\underline{\beta}_k$ is the safe budget defined in the refinement-count paragraph below); its fp64 branch, which skips storage and restoration, uses $(c_{\mathrm{out}} + T_k c_\Omega)u_{53}$. The storage perturbation is proved; the logarithm, nominal endpoint, residual evaluation, and restoration terms are conditional on the primitive contracts. Coordinate rounding excludes the float64 remap and remains a storage-only ablation.

Refinement count for chart Euler. Let $\ell_b := (\|\tilde{\Omega}\| - c_\Omega u_b)_+$ be the lower chord bound and $r_\Omega := \|\tilde{\Omega}\|/\ell_b$ the computed-rate/lower-chord ratio. Then $r_\Omega \le \frac{3}{2}$ once $\|\tilde{\Omega}\| \ge 3c_\Omega u_b$. In Experiments C–E and H the represented chords stay above $9 \cdot 10^{-7}$, and in Experiment F the chord reaches $5.7 \cdot 10^{-9}$ only at 94 bits, so $r_\Omega \le 1 + 10^{-8}$ throughout. Write $B_{k,b} := c_k T_k^2 \ell_b^2 = \min\{\alpha T_k(1 - T_k), cT_k^2\}\ell_b^2$ for the analytical local budget before comparison margins, with $c_k = \min\{c, \alpha(1 - T_k)/T_k\}$, $\widehat{B}_{k,b}$ for its floating evaluation, $\underline{\beta}_k := (1 - \tau)\widehat{B}_{k,b}$ for the safe budget that the implementation compares

against, and $\gamma := ((1-\tau)/(1+\tau))^2$. Under Assumption D.1(vi), the truncation-share comparison of the operational implementation is guaranteed whenever $E_n^{\mathrm{tr}} \le \theta\gamma B_{k,b}$ (this is the same sufficient condition as in the proof of Theorem D.5); since $E_n^{\mathrm{tr}} \le \overline{C}_E r_\Omega^2 T_k^2 \ell_b^2/n$, the deterministic sufficient sub-step count is $N_k = 2^{J_k}$ with $J_k = \max\{0, \lceil \log_2(\overline{C}_E r_\Omega^2/(\theta\gamma c_k))\rceil\}$, and if $E_n^{\mathrm{fl}} \le (1-\theta)\gamma B_{k,b}$ for every trial through N_k, with all chart-domain checks passing, then $E_n \le \gamma B_{k,b}$ at the sufficient trial and the safeguarded comparison accepts. Since $\theta > \theta_{\mathrm{stop}}$, no earlier reserve exit occurs, so the routine accepts by N_k (Proposition A.17 with $p = 1$; both error components receive compatible shares of the same comparison-safe budget). For the configuration used in the experiments, these constants give the following sufficient sub-step counts. For $\rho = 60°$, $\rho' = \rho + 0.1$, $D = 2\rho$, and $c = \alpha_\rho$, the constants are approximately $\kappa_1 = 1.057$, $\Lambda = 0.700$, $\phi(\Lambda D) = 2.27$, and $\overline{C}_E = 0.841$. With $\theta = \frac{1}{2}$ and $T_k \le \frac{1}{2}$, one has $c_k = \alpha_\rho = 0.433$, $r_\Omega \approx 1$, and $\gamma \approx 1 - 4\tau$, so $\overline{C}_E/(\theta\gamma c_k) \approx 3.9$, giving at most 4 sub-steps; at $k = 0$, where $T_0 = \frac{2}{3}$ and $c_0 = \alpha_\rho/2$, the bound is 8 sub-steps. The factor γ does not change these powers of two. The observed counts are reported separately (Experiment D). Using the truncation share $\theta = 0.68$ left by the largest observed rounding-floor share in Experiment C (0.32 of the safe budget) gives $\overline{C}_E/(0.68\,c) = 2.9$, so the sufficient power-of-two count remains four. This share is a sample statistic, not a uniform bound.

Reserve policy and comparison arithmetic. Throughout this appendix E and B denote the analytical enclosure and budget, $\widehat{E}$ and $\widehat{B}$ their floating evaluations, $\underline{\beta}_k = (1-\tau)\widehat{B}$ the safe budget as above, and the accepted test $(1+\tau)\widehat{E} \le \underline{\beta}_k$ is equivalent to $\widehat{E}/(1-\tau) \le \widehat{B}/(1+\tau)$, a comparison of the valid upper enclosure of E with the valid lower enclosure of B; the decision ratio $R_{\mathrm{acc}} = (1+\tau)\widehat{E}/\underline{\beta}_k$ is at most one exactly when this test passes. The implementation follows Algorithm 2 with $\theta_{\mathrm{stop}} = \frac{1}{4}$. It first establishes the domain reserve $\rho + Tc_\Omega u_b < \rho'$ of Proposition D.2 by an exact rational test on the represented values; a failed reserve test is a precision request, and no enclosure is evaluated. The routine then tests $(1+\tau)\widehat{E}_n \le \underline{\beta}_k$ by exact rational comparison of the represented values and accepts if the test passes. If it fails, the routine checks whether $(1+\tau)\widehat{E}_n^{\mathrm{fl}} \ge (1-\theta_{\mathrm{stop}})\underline{\beta}_k$; when this reserve threshold is reached, the routine exits with a request, and otherwise it doubles n (Lemma 3.3; the limiting-floor sequence of its proof and a constant-floor variant are regression tests). The finite-return argument of Lemma 3.3 transfers to these evaluated comparisons: if no reserve exit occurs, then $E_n^{\mathrm{fl}} \le \widehat{E}_n^{\mathrm{fl}}/(1-\tau) < (1-\theta_{\mathrm{stop}})\widehat{B}/(1+\tau)$, so the acceptance test holds as soon as $(1+\tau)^2 E_n^{\mathrm{tr}} \le [\theta_{\mathrm{stop}}(1+\tau) - 2\tau]\widehat{B}$, a positive reserve because $\theta_{\mathrm{stop}} = \frac{1}{4} > 2\tau/(1+\tau)$. A reference-exponential fallback is accepted only when its own conditional enclosure $E^{\mathrm{ref}} = (c_{\mathrm{out}} + Tc_\Omega)u_b$ passes the safeguarded comparison $(1+\tau)\widehat{E}^{\mathrm{ref}} \le \underline{\beta}_k$. This enclosure follows from the reference-endpoint bound of Assumption D.1(v-b) and the rate bound (ii). With A_{ref} the stored output, $z = x\exp(T\widehat{\Omega})$, and $\bar{z} = x\exp(T\widehat{\tilde{\Omega}})$ the exact endpoint of the computed-rate geodesic (the $\bar{z}_{k+1,b}$ of Section D.1), the nonexpansiveness used in the proof of Proposition D.2 gives $d(\Pi(A_{\mathrm{ref}}), z) \le c_{\mathrm{out}}u_b + d(\bar{z}, z) \le c_{\mathrm{out}}u_b + T\|\tilde{\Omega} - \Omega\| \le (c_{\mathrm{out}} + Tc_\Omega)u_b$. The adaptive kernel serves a request by a precision increase (Proposition D.4). If the applicable recovery policy cannot supply an accepted point, the run returns the status *cannot certify at the current precision* without producing a new iterate. The diagnostics distinguish four quantities. The violation *reporting threshold* is 10^{-10}, and the *resolution* of the float64 distance evaluation is about 10^{-15}. The conditional *enclosure* of the implemented map is E_n, or the applicable storage bound. The *reserve threshold* for the comparison quantity $(1+\tau)\widehat{E}_n^{\mathrm{fl}}$ is $(1-\theta_{\mathrm{stop}})\underline{\beta}_k$. On the boundary-active instance (50 seeds, 300 iterations) the checker reports the minimum safe budget over all accepted steps ($1.0 \cdot 10^{-14}$), the maximum accepted-step floor share $\max E^{\mathrm{fl}}/\underline{\beta}_k$ (0.32), and the first reserve request; these are generated from one checker ledger, used to report the work and consistency checks of Experiment C. Table 3 collects the comparison quantities. Table 4 lists the inputs of Algorithm 2, what the theorems assume about each, and where a valid value comes from; Algorithm 3 is the operational form of Algorithm 2 with both realizations, its complete input contract, and every status.

Independent checker. An independent script reconstructs, for every accepted step of every seed, 400-bit approximations to the represented current point, target, and output from the stored matrices (Π by a 400-bit polar iteration), and evaluates the represented-chord logarithm $\Omega_k = \mathrm{Log}_{x_k} q_k$ and the exact endpoint z_{k+1} in 400-bit arithmetic. The contract implies the chain $d(\tilde{z}_{k+1}, z_{k+1}) \le E_k \le$

$\widehat{E}_k/(1-\tau) \le \underline{\beta}_k/(1-\tau^2) \le \beta_k$, since $\underline{\beta}_k/(1-\tau^2) = \widehat{B}_{k,b}/(1+\tau) \le B_{k,b}$, with $\widehat{E}_k$ the evaluated enclosure that the run compared and β_k evaluated from the represented chord rather than from the stored approximation to its rate. The stronger inequality $d(\tilde{z}_{k+1}, z_{k+1}) \le \widehat{E}_k \le \underline{\beta}_k \le \beta_k$ also held numerically on every accepted step and is reported as an observation, not as a consequence of the contract. The script separately records the stored-rate endpoint error, the rate discrepancy $\|\tilde{\Omega}_k - \Omega_k\|$ against $c_\Omega u$, the feasibility of the represented target, the accepted-sub-step histogram, the total number of right-hand-side evaluations including rejected trials, and the minimum margin of every inequality. These are high-precision floating-point consistency checks, not outward-rounded interval certificates; the statement below that every accepted enclosure passed the 400-bit check refers to this independent consistency check.

symbol	meaning
E, $\widehat{E}$	analytical endpoint-error enclosure (exact arithmetic) and its floating-point evaluation
B, $\widehat{B}$	analytical local budget $\min\{\alpha T_k(1-T_k), cT_k^2\}\ell_b^2$ from the represented rate and its floating-point evaluation
$\underline{\beta}_k = (1-\tau)\widehat{B}$	safe budget, the right-hand side of the implemented comparison
$(1+\tau)\widehat{E} \le \underline{\beta}_k$	the acceptance decision, made exactly in rational arithmetic on the represented values
β_{geo}, β_{conv}	geometric budget $\alpha\eta(1-\eta)d^2$ and progress tolerance $c\eta^2 d^2$ of the exact chord (Experiment H)

Table 3: The comparison quantities of the implementation. The analytical enclosure and budget are never compared directly; their floating evaluations are tested using $(1+\tau)\widehat{E} \le (1-\tau)\widehat{B}$, the safeguarded comparison of Assumption D.1(vi).

Algorithm 3 Operational routine of Algorithm 2 with its complete contract and return statuses

Require: $x_0 \in \mathcal{C}$; a lower bound α on the modulus; $\widehat{D} \geq \operatorname{diam}\mathcal{C}$; $c \in (0, \alpha]$; proposed steps $\eta_k^0 \in (0, 1]$; a linear oracle returning a *feasible* $q \in \mathcal{C}$ with $\langle \operatorname{grad} F(x), \operatorname{Log}_x q\rangle \leq \min_{q' \in \mathcal{C}} \langle \operatorname{grad} F(x), \operatorname{Log}_x q'\rangle + \delta_k$, $\delta_k \leq \delta\, \underline{\eta}_k$ ($\underline{\eta}_k$ as in Lemma 3.3, computable before the call); and *either* (R) a retraction $\mathcal{R}$ with a proven bound $d(\mathcal{R}_x(v), \operatorname{Exp}_x v) \leq K\|v\|^r$ on $\{\|v\| \leq r_0\}$, $r \geq 2$, *or* (S) a routine APPROXEXP(x, v, j) returning either a candidate $(\text{CANDIDATE}, \tilde{z}_j, E_j^{\text{tr}}, E_j^{\text{fl}})$ with computable bounds $E_j^{\text{tr}}, E_j^{\text{fl}} \geq 0$ and $d(\tilde{z}_j, \operatorname{Exp}_x v) \leq E_j := E_j^{\text{tr}} + E_j^{\text{fl}}$, or an explicit numerical-domain or primitive-failure status (if every trial returns a candidate, $E_j^{\text{tr}} \to 0$ as $j \to \infty$); a reserve $\theta_{\text{stop}} \in (0, 1)$; and a FALLBACK returning either $(\text{ACCEPTED}, \tilde{z}, E^{\text{ref}})$ with $d(\tilde{z}, \operatorname{Exp}_{x_k}(\eta_k v_k)) \leq E^{\text{ref}} \leq \beta_k$ or a precision-request or *cannot certify* status

for $k = 0, 1, 2, \ldots$ **do**
  $q_k \leftarrow$ oracle at x_k; $v_k \leftarrow \operatorname{Log}_{x_k} q_k$; $d_k \leftarrow \|v_k\|$; $\eta_k \leftarrow \eta_k^0$
  if $d_k = 0$ **then**
    $x_{k+1} \leftarrow x_k$ {represented endpoints coincide; an unresolved chord is a precision request}
  else if realization (R) **then**
    while $\eta_k d_k > r_0$ **or** $K\eta_k^{r-1} d_k^{r-2} > \alpha(1 - \eta_k)$ **do**
      $\eta_k \leftarrow \eta_k/2$ {rejected trial; no map evaluation}
    end while
    $x_{k+1} \leftarrow \mathcal{R}_{x_k}(\eta_k v_k)$
  else if $\eta_k = 1$ **then**
    $x_{k+1} \leftarrow q_k$ {realization (S) returns the feasible target itself for a full step}
  else
    $\beta_k \leftarrow \min\{\alpha\eta_k(1 - \eta_k),\ c\,\eta_k^2\}\, d_k^2$
    for $j = 0, 1, 2, \ldots$ **do**
      $r_j \leftarrow$ APPROXEXP$(x_k, \eta_k v_k, j)$
      if r_j is not a candidate **then return** $(r_j.\text{status}, k, x_k, q_k, \eta_k)$ {failed check; no new iterate}
      $(\tilde{z}_j, E_j^{\text{tr}}, E_j^{\text{fl}}) \leftarrow r_j.\text{data}$; $E_j \leftarrow E_j^{\text{tr}} + E_j^{\text{fl}}$
      if $E_j \leq \beta_k$ **then**
        $x_{k+1} \leftarrow \tilde{z}_j$; **break** {accepted on the bound E_j}
      else if $E_j^{\text{fl}} \geq (1 - \theta_{\text{stop}})\beta_k$ **then**
        $(s, \tilde{z}, E^{\text{ref}}) \leftarrow$ FALLBACK$(x_k, q_k, \eta_k, \beta_k)$ {reserve exhausted}
        if $s =$ ACCEPTED **and** $E^{\text{ref}} \leq \beta_k$ **then** $x_{k+1} \leftarrow \tilde{z}$; **break** {defensive re-test}
        else return $(s', k, x_k, q_k, \eta_k)$ with $s' = s$, or $s' =$ *cannot certify* if $s =$ ACCEPTED {precision request or *cannot certify*; no new iterate}
      end if
    end for
  end if
end for

input	what must be known, and where it comes from
modulus lower bound α, $c \in (0, \alpha]$, $\widehat{D} \geq \operatorname{diam} \mathcal{C}$	a valid α for $\mathcal{C}$ is available in closed form for balls on $\mathrm{SO}(3)$ (Proposition 5.2) and spheres, and follows from (μ, L, s) for sublevel sets satisfying the ambient-tube assumptions of Theorem A.20
feasible target $q_k \in \mathcal{C}$ with $\langle \operatorname{grad} F(x_k), \operatorname{Log}_{x_k} q_k \rangle \leq \min_{\mathcal{C}} \langle \operatorname{grad} F(x_k), \operatorname{Log}_{x_k} \cdot \rangle + \delta_k$	represented feasibility of q_k is assumed, not tested; $\delta_k \leq \delta \underline{\eta}_k$ with $\underline{\eta}_k$ computable before the call (Lemma 3.3); the condition $\delta_k \leq \delta \eta_k$ is standard (Jaggi, 2013; Weber & Sra, 2023)
chord $d_k = \| \operatorname{Log}_{x_k} q_k \|$, proposed step $\eta_k^0 \in (0, 1]$	computed by the update anyway; a lower enclosure of β_k is used (this appendix)
(R) retraction $\mathcal{R}$ with $d(\mathcal{R}_x(v), \operatorname{Exp}_x v) \leq K \|v\|^r$ on $\{\|v\| \leq r_0\}$, $r \geq 2$	a uniform K proved for the retraction at hand; a sampled maximum only calibrates K (Remark C.7)
(S) routine ApproxExp(x, v, j) returning a candidate $(\tilde{z}_j, E_j^{\mathrm{tr}}, E_j^{\mathrm{fl}})$ with $d(\tilde{z}_j, \operatorname{Exp}_x v) \leq E_j := E_j^{\mathrm{tr}} + E_j^{\mathrm{fl}}$, or a domain or primitive-failure status	an a priori enclosure (a rounding bound, a series remainder, a Grönwall enclosure such as Proposition D.2), not a step-doubling indicator; $E_j^{\mathrm{tr}} \to 0$ if every trial returns a candidate; E_j^{fl} need not vanish at fixed precision
reserve $\theta_{\mathrm{stop}} \in (0, 1)$; Fallback returning $(\text{ACCEPTED}, \tilde{z}, E^{\mathrm{ref}})$ with $d(\tilde{z}, \operatorname{Exp}_{x_k}(\eta_k v_k)) \leq E^{\mathrm{ref}} \leq \beta_k$, or a precision request, or *cannot certify*	the fallback carries its own enclosure; a status returns without a new iterate

Table 4: The computational contract of Algorithm 2. The table lists each input, what the theorems assume about it, and where a valid value comes from. Under an error-decay model, Proposition A.17 bounds the refinement index, the map work, and the number of oracle calls needed to reach $h_k \leq \varepsilon$.

D.3 Precision policies and sufficient precision

Precision-adaptive certification. The rounding component of the enclosure is $E_n^{\mathrm{fl}} = u_b K_{\mathrm{fl}}(n)$. For fixed n and fixed problem constants, the stated model gives a uniform upper bound on $K_{\mathrm{fl}}(n)$, so executing the same kernel in b-bit arithmetic with a larger b makes this component arbitrarily small. The body rate, the chord bound, and the budget also depend on b, so the implemented policy recomputes every quantity at the precision finally used; domain and acceptance conditions are checked at that precision.

Proposition D.4 (operational precision policy)**.** *Fix $\theta \in (\theta_{\mathrm{stop}}, 1)$ and start from $b = \max\{53, b_{k-1}\}$. At each trial, the policy decodes the represented current point and target, then recomputes at precision b the body rate $\tilde{\Omega}_b$, the chord bound $\ell_b = (\|\tilde{\Omega}_b\| - c_\Omega u_b)_+$, and the safe budget $\underline{\beta}_{k,b}$. If the exact resolved-chord predicate of Assumption D.1(vii) fails (unresolved chord), the policy raises b and restarts. Otherwise, it chooses the smallest power of two n whose evaluated truncation enclosure passes $(1+\tau)\widehat{E}_n^{\mathrm{tr}} \leq \theta \underline{\beta}_{k,b}$ and the smallest $b' \geq b$ whose rounding enclosure meets the complementary budget, $(1+\tau) u_{b'} K_{\mathrm{fl}}(n) \leq (1-\theta)\underline{\beta}_{k,b}$, in the evaluation convention of Assumption D.1(vi). The domain reserve $\rho + T_k c_\Omega u_b < \rho'$ of Proposition D.2 is tested before this selection, by an exact rational test on the represented values; the scalar truncation formula only selects the trial count, and the enclosure is issued only after the computed chart nodes pass the domain checks. If $b' > b$, it sets $b = b'$ and restarts. Otherwise, it runs the kernel at precision b, checks every chart node, and accepts and stores the iterate at precision b if the safeguarded comparison $(1+\tau)\widehat{E}_n \leq \underline{\beta}_{k,b}$ passes. A failed trial raises b (by eight bits in the implementation) and restarts, and reaching a finite resource cap returns* cannot certify. *Precision never decreases, so the stored binary matrix entries of an iterate remain exactly representable at every later precision, and Assumption D.1(i) bounds the decoding of their represented rotation. Every accepted step satisfies Proposition D.2 at the precision used and is feasible by Corollary 3.2; with the declared finite resource cap, the policy either accepts or returns a status after finitely many restarts. The zero-direction branch of Algorithm 2 is taken before this search only when the stored matrices of x_k and q_k coincide; if the chord remains unresolved at the resource cap the policy returns the status* cannot certify *at the retry cap and does not infer optimality. This is the* economical *policy used in Experiment D; its outcomes there are conditional numerical observations. The theorem below concerns the separate sufficient policy that follows.*

Proof. An accepted step passes the safeguarded comparison with the enclosure of Proposition D.2 evaluated at the precision used, so under Assumption D.1 its represented endpoint satisfies $d(\tilde{z}_{k+1}, z_{k+1}) \le \beta_k$, and Corollary 3.2 gives feasibility. Every restart raises b by at least one bit, so the finite resource cap bounds the number of restarts, after which the policy has accepted or returned a status. □

Sufficient policy. The fixed-N policy below is designed to guarantee acceptance before the target accuracy is reached. Fix the problem constants, the chart reserve $s := \rho' - \rho$, and $\theta \in (\theta_{\mathrm{stop}}, 1)$, and choose N as in Theorem D.5. At iteration k, start from $b = \max\{53, b_{k-1}\}$; recompute the represented input decoding, body rate, chord bound, enclosure, and budget at precision b; attempt the chart-Euler update with exactly N sub-steps; accept iff the domain reserve $\rho + T_k c_\Omega u_b < \rho'$ of Proposition D.2 holds (an exact rational test on the represented radius bound, the binary step, and u_b, made before any enclosure is evaluated), the represented chord is resolved (the exact predicate of Assumption D.1(vii), equivalent to $\|\tilde{\Omega}_b\| \ge 3c_\Omega u_b$), every chart-node check succeeds, and the safeguarded comparison $(1+\tau)\widehat{E}_N \le \underline{\beta}_{k,b}$ succeeds; otherwise replace b by $b+1$ and repeat; a trial at the declared precision cap b_{cap} is permitted, and after a failed trial the policy returns *cannot certify* if the next precision would exceed b_{cap}. The accepted output is stored at precision b; a unit test checks that the implemented routine follows this policy.

Theorem D.5 (sufficient precision before the accuracy threshold)**.** *Work in the* $\mathrm{SO}(3)$ *ball and chart setting of Proposition D.2, with its constants* $\overline{D}$, $\overline{C}_E$, κ_1, Λ. *Assume the hypotheses of Theorem 4.1(ii) with an exact feasible oracle,* $G > 0$, $c \in (0, \alpha]$, *and an integer* $m \ge 4$. *Let* T_k *be the correctly rounded binary value of* $2/(k+m)$ *at the schedule precision (*53 *bits, no larger than any working precision). More generally,* T_k *may be any represented step with* $T_k = \frac{2}{k+m}(1+\vartheta_k)$, $|\vartheta_k| \le u_s < \frac{1}{2}$, *and* $0 < T_k \le \frac{1}{2}$, *represented at a precision no larger than every working precision used. For the rounded schedule and* $m \ge 4$, *monotonicity of rounding and exact representability of* $\frac{1}{2}$ *give* $T_k \le \frac{1}{2}$, *so the accepted-budget coefficient is c. Let* A_s *be the numerator of Lemma A.12, so that every sequence of accepted steps satisfies* $h_k \le A_s/(k+m)$; *assume Assumption D.1 at every precision used and the chart reserve* $s := \rho' - \rho > 0$; *let* $\gamma := ((1-\tau)/(1+\tau))^2$ *be the comparison-quality factor of Assumption D.1(vi) and fix* $\theta \in (\theta_{\mathrm{stop}}, 1)$. *Let* $\varepsilon > 0$, $H_\varepsilon := \lceil A_s/\varepsilon \rceil$, *and*

$$B_\varepsilon := \frac{c\,\gamma\,(1-u_s)^2\,\varepsilon^4}{G^2\,[A_s + (m+1)\varepsilon]^2}, \quad N := \textit{the smallest power of two } \ge \max\Big\{1,\ \frac{9\,\overline{C}_E}{4\theta c\gamma},\ \frac{4\overline{C}_E\overline{D}^2}{s}\Big\},$$

$$\overline{K} := e^{\Lambda\overline{D}}\big[c_{\mathrm{in}} + 2N\rho' + (3+c_f)\kappa_1\overline{D}\big] + c_{\mathrm{out}} + c_\Omega,$$

$$b_\varepsilon := \max\big\{53,\ \lceil \log_2(4Gc_\Omega/\varepsilon)\rceil,\ \lceil \log_2(\overline{K}/((1-\theta)B_\varepsilon))\rceil,\ \lceil \log_2(4\overline{K}/s)\rceil,\ \lceil \log_2(4c_\Omega/s)\rceil\big\}.$$

Then at every iteration $k \le H_\varepsilon$ *with* $h_k > \varepsilon$, *the chart-Euler step with* N *sub-steps in* b*-bit arithmetic, for any* $b \ge b_\varepsilon$, *resolves the chord, keeps every chart node in the ball of radius* ρ', *satisfies* $E_N \le \gamma\, cT_k^2\ell_b^2$ *so that the safeguarded comparison* $(1+\tau)\widehat{E}_N \le \underline{\beta}_{k,b}$ *of Assumption D.1(vi) accepts, and has a feasible represented endpoint. Consequently the fixed-*N *sufficient policy certifies every step before the first* ε*-accurate iterate, provided its declared precision cap satisfies* $b_{\mathrm{cap}} \ge \max\{b_0, b_\varepsilon\}$ *and no other resource limit interrupts execution (otherwise it returns* cannot certify *without a new iterate), using* N *sub-steps per attempted kernel call and at most* $b_{\max} \le \max\{b_0, b_\varepsilon\}$ *mantissa bits; its precision increases at most* $(b_\varepsilon - b_0)_+$ *times over the horizon, so it makes at most* $H_\varepsilon + (b_\varepsilon - b_0)_+$ *kernel attempts before producing an* ε*-accurate iterate, and Lemma A.12 gives* $h_k \le \varepsilon$ *by* $k = H_\varepsilon$. *As* $\varepsilon \to 0$ *with the problem constants fixed,* $B_\varepsilon = \Theta(\varepsilon^4)$, N *and* $\overline{K}$ *do not depend on* ε, *and* $b_\varepsilon = O(\log(1/\varepsilon))$. *These bounds do not claim that the economical policy of Proposition D.4 attains the same guarantee. Every attempted Euler step uses the binary* T_k, *so* $h = T_k/N$ *is exact and Proposition D.2 applies literally.*

Proof. Resolved chord. With an exact oracle, $Gd_k \ge g_k \ge h_k > \varepsilon$, so $d_k > \varepsilon/G$. By (ii), $\|\tilde{\Omega}_b\| \ge d_k - c_\Omega u_b$ and $\ell_b \ge d_k - 2c_\Omega u_b$ (the allowance appears twice because the primitive model bounds the error of the computed rate and the definition of ℓ_b subtracts a further $c_\Omega u_b$). The second requirement in b_ε gives $c_\Omega u_b \le \varepsilon/(4G)$, hence $\ell_b \ge \varepsilon/(2G) > 0$ and $\|\tilde{\Omega}_b\|/\ell_b = 1 + c_\Omega u_b/\ell_b \le \frac{3}{2}$; in particular $\|\tilde{\Omega}_b\| \ge 3c_\Omega u_b$, which is exactly the predicate the policy decides (Assumption D.1(vii)), so the resolution branch passes.

A budget independent of b. Since $T_k \le \frac{1}{2}$ and $c \le \alpha$, the progress tolerance is no larger than the geometric cap, so the local budget before comparison margins is $B_{k,b} = cT_k^2\ell_b^2$, and the comparison of Assumption D.1(vi) succeeds whenever $E_N \le \gamma B_{k,b}$ (the test $(1+\tau)\widehat{E} \le (1-\tau)\widehat{B}$ holds if $(1+\tau)^2 E \le (1-\tau)^2 B$). On the horizon, $T_k \ge 2(1-u_s)\varepsilon/(A_s + (m+1)\varepsilon)$ for $k \le H_\varepsilon$, which gives $B_\varepsilon \le \gamma B_{k,b}$. Therefore B_ε is a uniform lower bound on the guarded local budget, independent of b. We use it to choose the rounding precision and compare the truncation term directly with the local budget.

Truncation. $E_N^{\mathrm{tr}} \le \overline{C}_E(T_k\|\tilde{\Omega}_b\|)^2/N \le \overline{C}_E\,\frac{9}{4}\,T_k^2\ell_b^2/N \le \theta\gamma B_{k,b}$ by the first requirement on N, and $E_N^{\mathrm{tr}} \le \overline{C}_E\overline{D}^2/N \le s/4$ by the second.

Rounding and bits. $K_{\mathrm{fl}}(N) \le \overline{K}$ because $nh\kappa_1\|\tilde{\Omega}\|(3+c_f) \le (3+c_f)\kappa_1\overline{D}$, $T_k \le 1$, and $e^{LT} \le e^{\Lambda\overline{D}}$. The third requirement in b_ε gives $u_b\overline{K} \le (1-\theta)B_\varepsilon \le (1-\theta)\gamma B_{k,b}$, and hence $E_N = E_N^{\mathrm{tr}} + u_b K_{\mathrm{fl}}(N) \le \gamma B_{k,b}$, so the fixed-$N$ candidate passes the safeguarded comparison once its domain validity is established below. The last two requirements give $u_b\overline{K} \le s/4$ and $c_\Omega u_b \le s/4$.

Chart domain (first exit). The computed-rate path lies within radius $\rho + T_k c_\Omega u_b \le \rho + s/4$. The initial node satisfies $\|\tilde{w}_0 - w_0\| \le c_{\mathrm{in}}u_b \le \overline{K}u_b \le s/4$, so it lies in the ball. Suppose some later node left the ball of radius ρ' and let i be the first such index; all earlier nodes lie in the ball, so the recurrence of Proposition D.2 applies up to i and gives $\|\tilde{w}_i - \bar{w}(t_i)\| \le E_N^{\mathrm{tr}} + u_b K_{\mathrm{fl}}(N) \le s/4 + s/4$, whence $\|\tilde{w}_i\| \le \rho + 3s/4 < \rho'$, a contradiction. Thus every exact node test of Assumption D.1(vii) passes and Proposition D.2 applies to the whole step; Corollary 3.2 gives feasibility.

The sufficient policy. At an attempt with $b < b_\varepsilon$ the policy either accepts (an accepted step at any precision has passed the exact node-domain checks and the safeguarded comparison, so Proposition D.2 and Corollary 3.2 make it feasible; the threshold estimates above only ensure eventual success) or raises b by one; at any $b \ge b_\varepsilon$ the resolved-chord, first-exit, and comparison arguments guarantee acceptance, and the exact domain-reserve test passes because $c_\Omega u_b \le s/4$ gives $\rho + T_k c_\Omega u_b < \rho'$. Since each failed attempt increases b by exactly one bit, the policy does not exceed the sufficient integer threshold, so $b_{\max} \le \max\{b_0, b_\varepsilon\}$. Because precision never decreases, there are at most $(b_\varepsilon - b_0)_+$ failed attempts over the horizon and at most $H_\varepsilon + (b_\varepsilon - b_0)_+$ kernel attempts in total. Lemma A.12 applied to the accepted steps gives $h_k \le \varepsilon$ by $k = H_\varepsilon$. □

The theorem proves constructively that a sufficient pair (N, b) exists for every step before the first ε-accurate iterate and that the fixed-N policy reaches it. The fixed-N policy controls both the truncation error and the chart-domain reserve, whereas the economical policy of Proposition D.4 chooses its sub-step count from the truncation budget alone, so increasing the precision cannot rule out an exact-arithmetic discretization excursion beyond the chart; Theorem D.5 therefore does not cover that policy. Experiment D tests it empirically, and Experiment F tests the sufficient policy under the theorem's oracle and schedule assumptions. For the experimental configuration ($\rho = 60^\circ$, $\rho' = \rho + 0.1$, $\theta = \frac{1}{2}$) the theorem gives $N = 256$, determined by the domain-reserve term. The statement $N = O(1)$ expresses independence from ε and does not imply a small numerical constant. The theorem does not provide a computable certificate that the current iterate is ε-optimal, an exact oracle implementation, a proof of the primitive model, or a full bit-complexity bound. Let $\mathsf{C}_{\mathrm{att}}(N, b)$ bound the map work of one attempted update at precision b, including represented-input decoding, the body rate, N vector-field evaluations and updates, the chart-node tests, and the output representation. The map work before the first ε-accurate iterate is then $O((H_\varepsilon + (b_\varepsilon - b_0)_+)\,\mathsf{C}_{\mathrm{att}}(N, b_{\max}))$. This bound excludes the oracle, the enclosure and budget evaluation, and a stopping certificate; it counts mantissa bits only and assumes a sufficient exponent range (no underflow). The theorem therefore bounds the outer iteration count, the total attempt count, and the mantissa precision. A sharper use of the intermediate chord clearance (the truncation error at node t_i is at most $(t_i/T)E_n^{\mathrm{tr}}$, which the clearance $\alpha t_i(1-t_i)d^2$ of the intermediate chord point absorbs) might extend the guarantee to the economical policy, for which no failing instance was found; we do not claim it here.

D.4 Shared experimental protocol: environment, conventions, seeds, and unit tests

Except for Experiment F, the optimization studies use the constrained Fréchet objective of Section 6, the numerical reference oracle and objective reference defined below, and the reporting conventions

of Experiment C (Section D.8); schedule changes are specified in each study. This subsection also collects the computing environment, the conventions on $\mathrm{SO}(3)$, the random streams and seeds, and the unit tests that every study uses.

Reference oracle and objective reference. Outside Experiment F, the linear minimization oracle evaluates $\langle g, \mathrm{Log}_R Q\rangle$ on a Fibonacci grid of 1000 boundary points (Lemma A.23) and refines by Nelder–Mead in local coordinates (coordinate and function-value tolerances 10^{-10} and 10^{-14}). This is a high-accuracy numerical reference, not a certified global minimizer; measured oracle discrepancies are denoted $\widehat{\delta}_k$ because they are relative to this reference oracle, not to the true linear minimum. The objective reference F_{ref} is computed by projected Riemannian gradient descent (step $1/L_F$, 3000 iterations; projection onto the ball moves along the geodesic to R_c), whose reference-oracle Frank–Wolfe gap is at most $2.4 \cdot 10^{-13}$ in absolute value on every Experiment C instance and at most $1.2 \cdot 10^{-13}$ on every Experiment H instance (the reference oracle uses the sphere shrunk by the factor $1 - 2^{-40}$, which can make this value slightly negative). It is a numerical reference, independent of all compared methods, not a certified optimum; all reported gaps are $F(x_k) - F_{\mathrm{ref}}$, and gaps below about 10^{-13} are not resolved.

Environment. Python 3.12.7, NumPy 1.26.4 (MKL BLAS), SciPy 1.13.1, Matplotlib 3.9.2, mpmath 1.3.0, Windows 11 (build 26200), one 13th-generation Intel Core i7-13650HX laptop CPU (14 cores). The baseline methods of Experiment C and the fixed-precision kernel use IEEE binary64 arithmetic with a 53-bit significand. The adaptive kernels of Experiments D–F use b-bit binary floating-point arithmetic in mpmath with $b \geq 53$. The checker, the calibration tests, and the recomputed slacks and errors of Experiments D–H use 400-bit mpmath arithmetic. Python-level work is assigned to separate processes (up to 17 concurrently in the paper runs); BLAS thread counts were not controlled, and there was no warm-up or randomized method order, so wall-clock measurements are indicative single-machine timings and differences between similarly timed methods are not interpreted. The code, configurations, and results will be released after completion of the review process, as described in the reproducibility statement.

Randomization and seeds. Every random draw uses NumPy's default generator (PCG64). Table 5 specifies the streams and seeds of Experiments A–H; these seeds are the only source of randomness in those studies, Experiments D and E regenerate the indicated instances of Experiment C, and the deterministic studies draw nothing. The unit tests and the calibrations draw from one stream with seed 0, consumed in the order of the test script, except for the multiply-then-add, exact-comparison, and resolved-chord tests, which use separate streams with seeds 7, 11, and 3.

Conventions on $\mathrm{SO}(3)$. With the coordinate identification of Section D.1, $\mathrm{Log}_R Q = (\log R^\top Q)^\vee$ is evaluated by the Rodrigues formulas, with the branch at angle π handled by the symmetric part, and $d(R, Q)$ is the rotation angle. The Cayley retraction is $R\,\mathrm{cay}(\widehat{w}) = R(I - \frac{1}{2}\widehat{w})^{-1}(I + \frac{1}{2}\widehat{w})$ and the QR retraction is $\mathrm{qf}(R(I + \widehat{w}))$ with $\mathrm{qf}(I) = I$. The exponential chart about the ball centre is $w = \mathrm{Log}_{R_c} R \in \mathbb{R}^3$; in it the geodesic with body rate Ω solves $\dot{w} = J_r(w)^{-1}\Omega$ (Remark C.8).

Unit tests. The numerical checks use 2000–20,000 random configurations each, with the exceptions stated below, and we report the worst residuals. We check the Exp / Log round trip, whose largest residual is $1.6 \cdot 10^{-12}$, and the identity $\langle n, \mathrm{Log}_R Q\rangle = -\ell \cot(\rho/2)\tan(\ell/4)$ for boundary anchors, whose largest residual is $4 \cdot 10^{-14}$. We also check both inequalities of Proposition A.19(i) for interior anchors. For the set-level inequality $\rho - d(\gamma(t), R_c) \geq \alpha_\rho t(1-t)d^2(x, y)$, the largest value of budget minus clearance is $-2.0 \cdot 10^{-10}$, so the inequality holds on every sample. We check the exact-clearance identity of Proposition 5.2(ii) for chords whose two endpoints lie on the boundary, obtaining a largest residual of $5 \cdot 10^{-14}$. The remaining geometric checks cover the inequality $\rho - d(\gamma(t), R_c) \geq \Delta_\rho(t, \ell)$ for chords with at least one interior endpoint, the inequality $\Delta_\rho \geq \alpha_\rho t(1-t)\ell^2$, and the anchor-dependent clearance of Remark C.2, whose largest residual is $3 \cdot 10^{-13}$. The radial law of the volume-uniform sampler is checked on $2 \cdot 10^5$ samples, with a Kolmogorov–Smirnov distance of $2 \cdot 10^{-3}$. The two identity checks draw 2000 configurations with radii uniform on $[0.05, \pi/2 - 0.02]$ rad, so their largest residuals refer to this pool and are smaller than those of the 20,000-sample sweep of Experiment A at $89°$ (Table 7).

The numerical-map checks include the Cayley identity $d(\operatorname{cay}\widehat{w}, \exp\widehat{w}) = \theta - 2\arctan(\theta/2) \le \theta^3/12$. When the integration step is halved, the median global-error ratios over 200 configurations are 2.00 for Euler and 16.0 for RK4. We also check the inequality $d(R_c \exp\widehat{w_1}, R_c \exp\widehat{w_2}) \le \|w_1 - w_2\|$, the storage bounds of Lemma D.3, and the sharp-cap identity of Proposition C.1. The constants of Proposition D.2 are checked by comparing ζ' with a central difference and by evaluating $\|J_r^{-1}\| \le \kappa_1$ and the Lipschitz bound on random pairs. The enclosure itself is checked on 400 random feasible steps, with a maximum ratio of evaluated error to E_n of 0.22.

The checks also include the high-precision calibration of Assumption D.1. The same script reports the *empirical* QR ratio $\max d(\mathrm{qf}(I + \widehat{w}), \exp\widehat{w})/\|w\|^2$ over 20,000 samples in each range, namely 0.335, 0.342, and 0.342 on $\|w\| \le 0.5$, 1.1, and 2.1, respectively. The experiments use $K = 0.40$. These sampled values calibrate that choice but do not prove a uniform bound (Remark C.7).

study	streams and what they draw
A	One stream with seed 0 is consumed radius by radius in the order $10^\circ, 25^\circ, 45^\circ, 70^\circ, 85^\circ, 89^\circ$. At each radius it draws 20,000 configurations, each containing a centre, a boundary point, a boundary anchor, and an interior anchor with a volume-uniform radius, and then 20,000 chords, with each endpoint on the boundary with probability $\frac{1}{2}$ and otherwise volume-uniform. No draw is shared across radii; within a radius, the chord checks use the eligible subsets of the same 20,000-chord pool, and the anchor checks use the configurations of the same pass. The asymptotic sharpness study uses seed 3 with three orientations per chord length. Panel (a) of Figure 3 uses seed 1 with 3000 samples at 45°.
B	This study is deterministic. It uses 14 logarithmically spaced values of b from $10^{-2.5}$ to 0.35, fits slopes on the eight smallest values, and samples 400,001 points on each boundary piece.
C	For instance $j \in \{0, \dots, 49\}$, the stream seeded j draws the centre, the nominal direction, and the sample noise (with the 105° rejection), in that order. The random-direction error methods draw their directions from the paired stream $10^4 + j$, the shrinking-perturbation oracle draws its tangential direction from a stream seeded 777, and the seed bootstrap uses the stream seeded 123 (4000 resamples).
D, E	The boundary-active instances $j \in \{0, \dots, 9\}$ (D) and $j \in \{0, \dots, 4\}$ (E) of Experiment C are regenerated from the same streams. The precision policies draw nothing.
F	This study is deterministic (the represented instance of Table 26).
G	One stream with seed 2027 is consumed in grid order, varying clearance first, then length ratio, then step fraction, with eight repetitions per cell. Every case draws a fresh centre and two unit vectors, so no orientation is shared across cells.
H	For instance (j, ϕ), the stream seeded $5000 + j$ draws the centre, the two reference directions, the cluster noise (with the 105° rejection), and the near-boundary start direction, in that order. The stream is re-seeded for each separation, so the two separations of a seed share the centre and reference directions. The stratum bootstrap intervals use a stream seeded 1 and the pooled controls a stream seeded 2 (4000 resamples each).

Table 5: The NumPy PCG64 streams and seeds used in each study.

D.5 Geometric checks and the sublevel counterexample (Experiments A and B)

Experiment A. At each radius, we generate 20,000 random configurations, each consisting of a centre, a boundary point, a boundary anchor, and an interior anchor. Interior anchor radii follow the volume-uniform law of $\mathrm{SO}(3)$, $\Pr(R \le r) = (r - \sin r)/(\rho - \sin\rho)$. Its radial density is proportional to $\sin^2(r/2)$, so the Euclidean law $\rho u^{1/3}$ is not volume-uniform on $\mathrm{SO}(3)$. For the set-level and exact-clearance checks, we generate a single pool of 20,000 random chords. Each endpoint is independently placed on the boundary or volume-uniform inside with probability $\frac{1}{2}$ each, and t is uniform on $[0, 1]$. The set-level check uses the full pool, the exact-clearance identity its boundary-to-boundary subset, and the Δ_ρ ratio the complementary subset; the realized subset sizes are given in the caption of Table 7. Table 6 reports the moduli and the minimum sampled boundary-normal ratios; Table 7 reports the residual of the margin identity, the minimum clearance-to-budget ratio, and the residual of the exact-clearance identity.

Experiment B (counterexample). The clearance $\mathrm{dist}(m_b, \partial Q_s)$ is computed both in closed form and by dense sampling ($400{,}001$ points on each boundary piece); the two agree to machine precision. Slopes are least-squares fits on the eight smallest b. This experiment only illustrates Proposition A.22.

ρ	α_ρ (sharp)	α_0 (tube)	$1/(2\rho)$	min sampled ratio, boundary	sampling gap to α_ρ	min sampled ratio, interior
10°	2.857513	2.8356	2.8648	2.857513	9.9e-08	2.8594
25°	1.127677	1.0723	1.1459	1.127678	4.1e-07	1.1290
45°	0.603553	–	0.6366	0.603554	9.2e-07	0.6052
70°	0.357037	–	0.4093	0.357039	2.2e-06	0.3590
85°	0.272827	–	0.3370	0.272827	2.2e-07	0.2765
89°	0.254402	–	0.3219	0.254404	2.4e-06	0.2581

Table 6: Experiment A compares the sharp modulus α_ρ, the tube-theorem modulus α_0 available for $\rho < 45^\circ$, and the Euclidean value $1/(2\rho)$ on $\mathrm{SO}(3)$. The sampled columns use $20{,}000$ configurations per radius and report the minimum sampled ratio $-\langle n, \mathrm{Log}_R Q\rangle/\ell^2$ over boundary anchors, its gap to the infimum α_ρ, and the minimum sampled ratio over interior anchors. Moduli and ratios are in rad^{-1}. A finite sample with positive chord length does not attain the infimum, which is approached as $\ell \to 0$.

ρ	margin-identity residual	min clearance/quadratic budget	exact-clearance identity residual	min clearance/Δ_ρ (interior chords)
10°	3.3e-16	1.0000	1.4e-15	1.0000
25°	6.7e-16	1.0000	1.4e-15	1.0000
45°	1.3e-15	1.0001	5.2e-15	1.0001
70°	5.8e-15	1.0000	8.9e-15	1.0000
85°	8.2e-14	1.0001	8.2e-13	1.0000
89°	1.7e-12	1.0001	3.5e-12	1.0001

Table 7: Experiment A reports numerical residuals. The margin-identity residual uses the $20{,}000$ boundary-anchor configurations of each radius; the chord columns use one pool of $20{,}000$ chords per radius and its eligible subsets. The columns give the residual of the closed-form margin identity, the minimum ratio of chord clearance to the quadratic budget over the full pool, the residual of the exact-clearance identity of Proposition 5.2(ii) over the chords with both endpoints on the boundary (4957, 5020, 5003, 4984, 4990, and 5055 chords at the six radii), and the minimum ratio of clearance to Δ_ρ over the remaining chords with at least one interior endpoint ($15{,}043$, $14{,}980$, $14{,}997$, $15{,}016$, $15{,}010$, and $14{,}945$ chords).

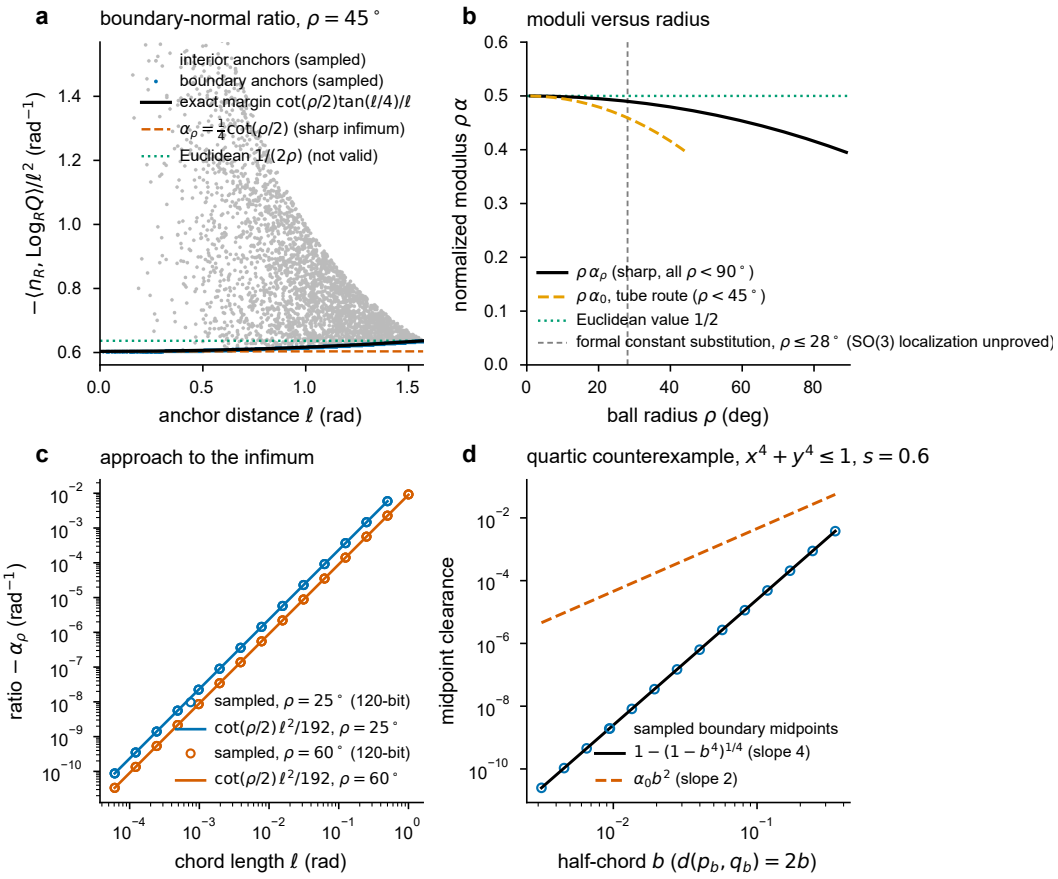

Figure 3: Experiments A and B. (a) Boundary-normal ratio versus anchor distance on a $45°$ ball; the exact margin (solid) is attained by boundary anchors and α_ρ (dashed) is its infimum. (b) Normalized moduli $\rho\,\alpha$ versus radius; the tube route is available for $\rho < 45°$ and the vertical line marks the range $\rho \le 28°$ that a formal substitution into the bounded-curvature certificate produces (Remark C.4), neither being an intrinsic limit. (c) Asymptotic sharpness study for $\rho = 25°$ and $60°$, using boundary anchors at deterministic chord lengths $\ell = 2^{-j}$, with three orientations per ℓ and 120-bit arithmetic. The evaluated excess of the ratio over α_ρ follows the leading term $\cot(\rho/2)\ell^2/192$ of $\cot(\rho/2)(\tan(\ell/4)/\ell - \frac{1}{4})$. (d) Quartic counterexample, slope 4 against the slope 2 that a positive modulus would require.

D.6 The native near-boundary stress grid (Experiment G)

Experiment G (a declared, admissible near-boundary grid for native kernels). To measure native endpoint error relative to the available clearance without an injected error, we take chords of $\mathcal{C}_{60°}$ with the anchor x at radius $r_x = \rho - \delta_x$, $\delta_x \in \{10^{-2}, 10^{-3}, 10^{-4}\}$ rad, the target q on the oracle sphere of radius $r_q = \rho(1 - 2^{-40})$, chord length $\ell = d(x, q)$ with $\ell/\delta_x \in \{1.1, 3, 10\}$ (every cell satisfies the reverse triangle inequality $\ell \ge r_q - r_x$), step fractions $\eta \in \{0.5, 0.1, 0.02\}$, and 8 random orientations per cell (216 admissible cases). The chord is constructed from the spherical law of cosines on the cover. Before any kernel is applied, the realized distances $d(R_c, x)$, $d(R_c, q)$, $d(x, q)$ are recomputed in 400-bit arithmetic and checked against the declared values with tolerance 10^{-12}; the intended and realized geometry are stored for every case.

Four native kernels approximate the same exact chord endpoint. The fp16 and fp32 matrix-storage kernels quantize a float64 reference endpoint and restore it by the float64 polar projection; the other two kernels are one-step chart Euler about the centre and the QR retraction. For every case we record in 400-bit arithmetic the clearance of the float64 reference endpoint (the proxy of Section D.1), the kernel's endpoint error relative to it, whether the represented output is feasible, and the decision of

the stated bound (the storage ledger, the enclosure of Proposition D.2 with $n = 1$, or the calibrated QR bound) against the safe budget. All cases are reported (Table 8).

fp16 storage with restoration produced 10 infeasible outputs in 216 cases, all with $\delta_x = 10^{-4}$, where its error ($\leq 2.8 \cdot 10^{-4}$ rad) exceeds the clearance by up to $4.3\times$. Its ledger bound exceeds the acceptance budget in every case, so the guard rejects the entire fp16 grid, a conservative precision veto that also rejects the 206 feasible outputs rather than a selective prediction of the 10 failures. fp32 storage stayed feasible and was accepted in 80 cases; one-step Euler and QR stayed feasible, were accepted in all cases, and reached 3.7% and 11% of the clearance. These ratios quantify native endpoint error relative to the available clearance, not merely the certificate's conservatism. Because the cases form a declared diagnostic grid rather than a search for failures, the observed fractions are not estimates of failure rates in deployment.

kernel	infeasible	accepted	infeasible, accepted	feasible, rejected
fp16 storage + restoration	10	0	0	206
fp32 storage + restoration	0	80	0	136
chart Euler, 1 step	0	216	0	0
QR retraction	0	216	0	0

kernel	max endpoint err [rad]	max err/clearance(z)
fp16 storage + restoration	2.8e-04	4.34e+00
fp32 storage + restoration	3.1e-08	4.76e-04
chart Euler, 1 step	2.2e-04	3.68e-02
QR retraction	6.6e-04	1.09e-01

Table 8: Experiment G evaluates native kernels on the declared admissible near-boundary grid (216 cases per kernel, no injected error). The upper panel reports infeasible represented outputs, cases accepted under the stated bound and model, and counts classified jointly by that decision and the 400-bit feasibility check. The lower panel reports the maximum endpoint error and the maximum ratio of endpoint error to the clearance of the reference endpoint; the smallest clearance in the grid is $5.0 \cdot 10^{-5}$ rad.

D.7 Nonradial comparisons and precision controls (Experiment H)

Instances and replication. Experiment H compares native kernels, direct alternatives, and precision controls on a prespecified nonradial family. The instances of Experiment C are nearly radial and have inexpensive membership tests and projections. The family used here, fixed before these runs, moves off the radial line and compares the same native error source under five treatments, including the direct alternatives, together with two matched precision controls. Each instance is indexed by (j, ϕ) with $j \in \{0, \dots, 9\}$ and has a random centre R_c. It contains three nominal attitudes $R_c \exp(r_i \widehat{v_i})$. The unit directions are v_1 and its rotations by $\pm\phi$ within a randomly oriented plane in the body-coordinate space $\mathbb{R}^3$. The radii are $r = (\rho + 10^\circ, \rho + 5^\circ, \rho + 15^\circ)$, the cluster sizes are $(25, 15, 10)$, with isotropic 10° noise and the same 105° rejection rule as Experiment C. The separation $\phi \in \{30^\circ, 60^\circ\}$ controls the directional spread of the nominal attitudes. As shown in Table 9, $\phi = 30^\circ$ gives boundary-active reference minimizers with gradient norms between 0.069 and 0.12, whereas $\phi = 60^\circ$ gives interior reference minimizers. On every instance the reference-oracle Frank–Wolfe gap of the reference minimizer is at most $1.2 \cdot 10^{-13}$ in absolute value, far below the smallest final gaps reported ($2.5 \cdot 10^{-6}$ and $7.2 \cdot 10^{-6}$ medians, Tables 11–12); this numerical diagnostic is not a certified global error bound. Each instance uses two starts, $x_0 = R_c$ and the near-boundary point $x_0 = R_c \exp(0.9\rho\, \widehat{u})$ with a random unit u. Every run uses the schedule $\eta_k = 2/(k + 3)$, 300 iterations, and the reference oracle of Section D.4. The ten instance seeds are the independent randomization units. The two separations and two starts of a seed form one *seed bundle*, so each treatment has 40 trajectories in 10 bundles. The native kernels are one-step chart Euler about R_c, fp16 matrix storage with float64 polar restoration, fp32 matrix storage with float64 polar restoration (the always-fp32 control), and the QR retraction. The storage kernels first compute the float64 reference endpoint and then degrade its representation, and that construction is charged to the native proposal.

Native kernels and decision rules. We compare five treatments of the native output $\tilde{z}$. The unguarded treatment accepts the native output without an additional check. The float64 membership guard tests $d(\tilde{z}, R_c) \le \rho$ in float64 with tolerance 0 [DI]. The 400-bit membership guard uses the sign of the computed 400-bit clearance of the represented output [DI]; this is a high-precision floating-point test, not an interval one, and its computation is charged to the online guard. The endpoint-error guard applies the safeguarded comparison $(1+\tau)\widehat{E} \le \underline{\beta}_k$ of the stated bound against the safe budget. It uses the enclosure of Proposition D.2 with $n = 1$ for chart Euler [AN+CO], the storage ledger with the restoration domain established by the exact predicate of Assumption D.1(vii) for the storage kernels [AN+CO], and the calibrated bound for QR [CA]. The fifth treatment applies radial metric projection of $\tilde{z}$ onto the ball in float64. The fp32 kernel under the endpoint guard is the prespecified fp32-first guarded control.

Fallback and cost accounting. After a rejection, each of the three guards attempts the same float64 reference exponential, and, as in Algorithm 2, the fallback is accepted only on its own rule. Under the endpoint guard it carries the conditional enclosure $E^{\mathrm{ref}} = (c_{\mathrm{out}} + T_k c_\Omega)u_{53}$, $u_{53} = 2^{-53}$, with T_k the binary step actually used (the float64 branch of Lemma D.3, which rests on the reference-endpoint bound of Assumption D.1(v-b) and the rate bound (ii)). Writing $\widehat{E}^{\mathrm{ref}}$ for its floating-point evaluation, acceptance requires the safeguarded exact comparison $(1+\tau)\widehat{E}^{\mathrm{ref}} \le \underline{\beta}_k$ against the current safe budget. Each membership guard instead repeats its membership test on the fallback. A failed re-test returns the status *cannot certify* and ends the trajectory without a new iterate. No re-test failed in the reported runs. Over the 17,661 endpoint-guard fallbacks of the four kernels, the largest ratio $(1+\tau)\widehat{E}^{\mathrm{ref}}/\underline{\beta}_k$ was 0.092 (Table 16). For the chart-Euler kernel the endpoint guard first establishes the domain reserve of Proposition D.2 by the exact test. For the storage kernels the fallback reuses the endpoint that the native workflow has already computed, whereas for chart Euler and QR it is a new reference evaluation. The two cases are counted separately, and the re-test is charged to output construction.

We separate the time of each step into oracle evaluation, native proposal construction, online guard evaluation, and output construction. The guard cost includes the 400-bit clearance when the 400-bit membership guard is used. Output construction is a new reference evaluation, a reused endpoint, or the projection, as applicable. Offline diagnostics are recorded separately because they are computed only for reporting and never enter a decision (Table 16). Measurements were collected under concurrent load without warm-up or controlled BLAS threads and do not establish comparative runtime performance. For every step we record in 400-bit arithmetic the clearance Δ of the endpoint z of the represented chord and the native endpoint error $e = d(\tilde{z}, z)$, both with z represented by the float64 reference endpoint (Section D.1), the sign of the computed clearance of the native output and of the accepted iterate, the stated bound $\widehat{E}$, the safe budget $\underline{\beta}_k$, the chord length, and the decision ratio $R_{\mathrm{acc}} = (1+\tau)\widehat{E}/\underline{\beta}_k$; feasibility counts use that sign. This Δ approximates the clearance of the particular endpoint (Remark C.2) to within the reference-endpoint allowance of Section D.1; it is not the worst-case function Δ_ρ of Proposition 5.2, which the exact clearance dominates and which every recorded value exceeds by at least $2.3 \cdot 10^{-5}$ rad. A seed bundle has a reported violation if any accepted iterate in its four trajectories has computed 400-bit clearance $\rho - d(x_k, R_c) < -10^{-10}$, and the Clopper–Pearson interval concerns that bundle event (Table 10); the sign classification of native proposals is recorded separately from this reporting threshold.

Paired comparisons. Tables 11 and 12 report the strata by separation and start. Table 13 compares the storage family within the boundary-active and interior regimes. For each regime, the paired effect is the equal-weight mean, over its 20 trajectories (both starts of the 10 seeds), of $\log_{10}(F(x_{300}) - F_{\mathrm{ref}})$ minus the same quantity for the reference-exponential run with the same seed, start, and separation. The 95% percentile interval uses 4000 cluster-bootstrap resamples of the 10 seed bundles, with all trajectories of a resampled bundle entering together. The effect is omitted, never clipped, if any included gap is at or below the resolution floor 10^{-13} or if any trajectory of the group violates. Table 14 gives the offline common-candidate comparison, in which the native proposals recorded on the endpoint-guard trajectory are classified by the 400-bit membership and endpoint-error rules without changing the points. These proposals are generated at the endpoint-guarded states, so their infeasibility count is not the count on the standalone unguarded or membership-guarded trajectories, which report no violation.

Table 15 decomposes the conservatism on the same proposals. For the represented chord, define $\beta_{\mathrm{geo}} := \alpha\eta(1-\eta)d^2$, $\beta_{\mathrm{conv}} := c\eta^2 d^2$, and $B := \min\{\beta_{\mathrm{geo}}, \beta_{\mathrm{conv}}\}$, and let $\underline{\beta}$ denote the safe budget. For proposals with positive recorded error e, clearance Δ (both at the reference endpoint, as recorded above), and chord length, the decomposition is $\widehat{E}/\underline{\beta} = (e/\Delta)(\widehat{E}/e)(\Delta/\beta_{\mathrm{geo}})(\beta_{\mathrm{geo}}/B)(B/\underline{\beta})$. Its factors measure the endpoint error relative to the available clearance, the stated bound relative to the evaluated error, the conservatism of the generic geometric budget, the additional restriction imposed by the progress tolerance, and the combined reduction introduced by the lower bound on the chord length and the factor $1-\tau$ of the safe budget. With $c = \alpha$ and the ideal schedule $\eta_k = 2/(k+3)$, $\beta_{\mathrm{geo}}/\beta_{\mathrm{conv}} = (1-\eta_k)/\eta_k = (k+1)/2$, so the schedule fixes the fourth factor at $\max\{1, (k+1)/2\}$; the implemented factor uses the represented step T_k and differs from this value by at most $3 \cdot 10^{-14}$, and the float64 values tabulated below differ from it by less than 10^{-13}. The progress tolerance is therefore the smaller term at every $k \geq 2$ (298 of the 300 iterations; the two terms coincide at $k = 1$) whatever the geometry, and the median of the fourth factor over the 300 iterations rounds to 75.25 for both schedules. The marginal medians of the five factors are descriptive and do not multiply to the median of their product. Consequently, replacing the generic geometric budget by the reference clearance Δ while retaining the cap changes the acceptance decisions only marginally: the twenty additional QR acceptances are the $k = 0$ steps of the centre-start trajectories, where the geometric term is the smaller one and the clearance is large, whereas at the near-boundary starts the QR bound at $k = 0$ exceeds even the recorded clearance. The table also counts the proposals that would be accepted under the current budget, under the reference clearance Δ capped by the progress tolerance $c\eta^2 d^2$ of the computed chord (the stated bound compared without the comparison margins), under the reference error e evaluated in 400-bit arithmetic against the current safe budget (an offline diagnostic, not an implementable guard), and by membership. The findings are stated in Section 6; the protocol, including the two controls, was fixed before these runs and no parameter was tuned to the outcome.

separation ϕ	seeds	reference slack $\rho - d(x_{\mathrm{ref}}, R_c)$ [rad] (min/max)	$\|\operatorname{grad} F(x_{\mathrm{ref}})\|$ (min/max)	max $d(R_c, R_i)$ [deg]
30°	10	-2.2e-16 / 4.4e-16	6.9e-02 / 1.2e-01	97.9
60°	10	6.5e-02 / 1.3e-01	5.2e-18 / 6.3e-17	100.6

Table 9: The Experiment H instances summarized by separation ϕ. For each separation, the table gives the minimum and maximum slack and gradient norm of the reference minimizer over the 10 seeds, together with the largest sample distance from the centre.

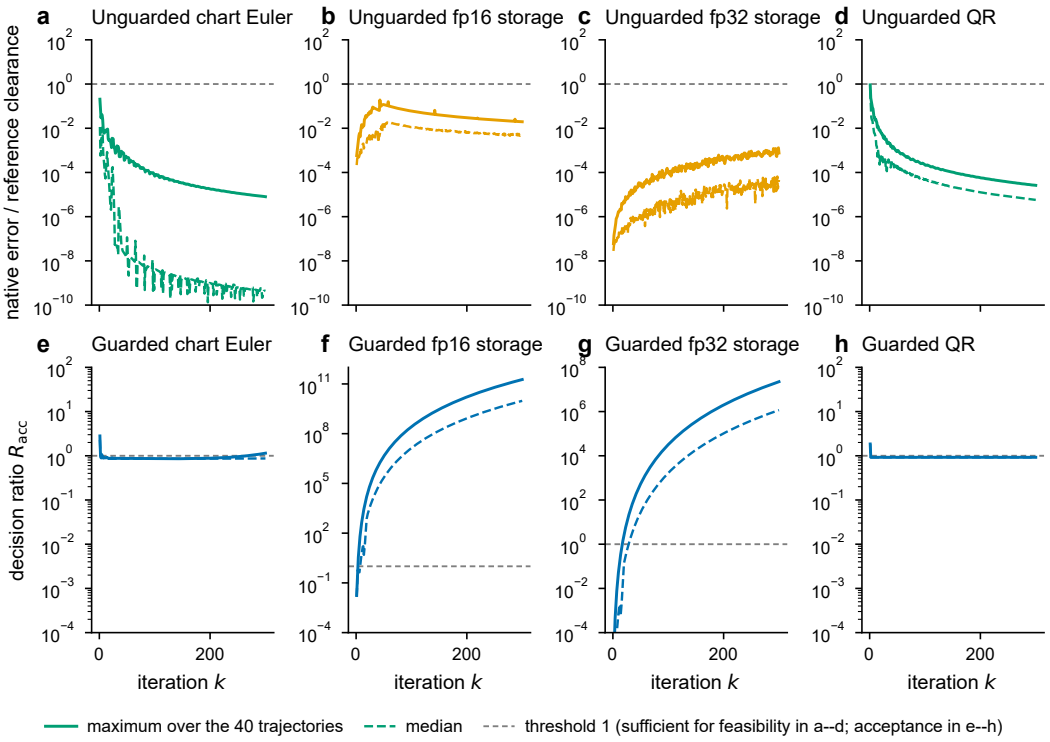

Figure 4: Experiment H. Panels (a–d) show, for each unguarded native kernel, the native endpoint error divided by the endpoint clearance, both evaluated at the represented float64 reference endpoint as the proxy for the exact endpoint (Section D.1), with the maximum and median over the 40 trajectories at each iteration; the plotted 400-bit values are diagnostics, and in the chart-Euler panel more than half of the recorded errors of the boundary-active regime lie below the proxy bound of Section D.1, so the smallest ratios there are resolution-limited. For exactly evaluated distances, a ratio at most one is sufficient for feasibility of that proposal (triangle inequality about the reference endpoint), and a larger ratio is inconclusive. Panels (e–h) show the endpoint-guard decision ratio $R_{\mathrm{acc}} = (1+\tau)\widehat{E}/((1-\tau)\widehat{B})$ on the native proposals of the endpoint-guarded trajectories, including the rejected ones; the guard accepts exactly when this ratio is at most one after the domain and input checks succeed, its denominator is the safe budget, which is generally far smaller than the clearance in (a–d), and the numerator $(1+\tau)\widehat{E}$ is a comparison device rather than a separate upper enclosure. The two rows use different trajectories and do not decompose the same proposals; Table 15 isolates the sources of conservatism on unchanged recorded candidates.

Kernel	Treatment	Proposals	Native inf.	Inf. acc.	Feas. rej.	Fallbacks new / reused	Traj. inf.	Bundles inf. [95% CI]
reference Exp (float64)	–	12,000	–	–	–	–	0/40	0/10 [0.00,0.31]
chart Euler, 1 step	no guard	12,000	0	–	–	–	0/40	0/10 [0.00,0.31]
	float64 membership	12,000	0	0	0	0 / 0	0/40	0/10 [0.00,0.31]
	400-bit membership	12,000	0	0	0	0 / 0	0/40	0/10 [0.00,0.31]
	endpoint-error guard	12,000	0	0	327	327 / 0	0/40	0/10 [0.00,0.31]
	radial projection	12,000	0	–	–	–	0/40	0/10 [0.00,0.31]
fp16 storage	no guard	12,000	0	–	–	–	0/40	0/10 [0.00,0.31]
	float64 membership	12,000	0	0	0	0 / 0	0/40	0/10 [0.00,0.31]
	400-bit membership	12,000	0	0	0	0 / 0	0/40	0/10 [0.00,0.31]
	endpoint-error guard	12,000	382	0	11299	0 / 11681	0/40	0/10 [0.00,0.31]
	radial projection	12,000	0	–	–	–	0/40	0/10 [0.00,0.31]
fp32 storage	no guard	12,000	0	–	–	–	0/40	0/10 [0.00,0.31]
	float64 membership	12,000	0	0	0	0 / 0	0/40	0/10 [0.00,0.31]
	400-bit membership	12,000	0	0	0	0 / 0	0/40	0/10 [0.00,0.31]
	endpoint-error guard	12,000	0	0	5613	0 / 5613	0/40	0/10 [0.00,0.31]
	radial projection	12,000	0	–	–	–	0/40	0/10 [0.00,0.31]
QR retraction	no guard	12,000	0	–	–	–	0/40	0/10 [0.00,0.31]
	float64 membership	12,000	0	0	0	0 / 0	0/40	0/10 [0.00,0.31]
	400-bit membership	12,000	0	0	0	0 / 0	0/40	0/10 [0.00,0.31]
	endpoint-error guard	12,000	0	0	40	40 / 0	0/40	0/10 [0.00,0.31]
	radial projection	12,000	0	–	–	–	0/40	0/10 [0.00,0.31]

Table 10: Experiment H uses $\rho = 60^\circ$ and 10 seeds, each with 2 separations and 2 starts. Each row therefore summarizes 40 trajectories of 300 iterations, giving 12,000 proposals. The columns report native infeasible outputs according to the sign of the computed 400-bit clearance; infeasible proposals accepted and feasible proposals rejected by the guard decision relative to that classification; and fallbacks to the reference exponential, separated into new reference evaluations and reused endpoints of the storage workflow. They also report the number of trajectories containing an accepted iterate whose computed 400-bit clearance $\rho - d(x_k, R_c)$ is below -10^{-10}, followed by the number of seed bundles containing any such violation ("Native inf." is the number of natively infeasible proposals, "Inf. acc." the infeasible proposals accepted, and "Feas. rej." the feasible proposals rejected). The trajectory count is a raw count over the 40 dependent trajectories. The bundle count uses the 10 independent seeds and carries a 95% Clopper–Pearson interval for the bundle event.

kernel	treatment	native infeasible	feasible & rejected	fallback share	max err/clearance; max R_{acc}
centre start					
reference Exp	–	–	–	–	–
chart Euler, 1 step	no guard	0	–	–	5.44e-03; 2.20e+00
	endpoint-error guard	0	286	9.5%	5.44e-03; 2.20e+00
fp16 storage + restoration	no guard	0	–	–	1.90e-01; 4.61e+07
	endpoint-error guard	301	2669	99.0%	7.96e+00; 1.84e+11
fp32 storage + restoration	no guard	0	–	–	1.31e-03; 2.24e+07
	endpoint-error guard	0	2829	94.3%	1.22e-03; 2.24e+07
QR retraction	no guard	0	–	–	4.01e-01; 1.85e+00
	endpoint-error guard	0	10	0.3%	4.01e-01; 1.85e+00
near-boundary start					
reference Exp	–	–	–	–	–
chart Euler, 1 step	no guard	0	–	–	2.06e-01; 2.81e+00
	endpoint-error guard	0	12	0.4%	2.06e-01; 2.81e+00
fp16 storage + restoration	no guard	0	–	–	1.59e-01; 3.37e+07
	endpoint-error guard	81	2876	98.6%	4.28e+00; 8.46e+10
fp32 storage + restoration	no guard	0	–	–	6.93e-04; 1.03e+07
	endpoint-error guard	0	2784	92.8%	6.67e-04; 1.03e+07
QR retraction	no guard	0	–	–	9.18e-01; 1.85e+00
	endpoint-error guard	0	10	0.3%	9.18e-01; 1.85e+00

kernel	treatment	gap k=300 (median)	$\Delta \log_{10}$ gap [CI]
centre start			
reference Exp	–	2.49e-06	–
chart Euler, 1 step	no guard	2.49e-06	-0.0028 [-0.0066,-0.0001]
	endpoint-error guard	2.49e-06	-0.0028 [-0.0066,-0.0001]
fp16 storage + restoration	no guard	1.60e-04	$+1.835$ [$+1.757$,$+1.910$]
	endpoint-error guard	2.49e-06	-0.001 [-0.015,$+0.011$]
fp32 storage + restoration	no guard	2.50e-06	$+4.3\mathrm{e}{-04}$ [$+1.7\mathrm{e}{-04}$,$+7.9\mathrm{e}{-04}$]
	endpoint-error guard	2.49e-06	$-9.7\mathrm{e}{-07}$ [$-1.8\mathrm{e}{-06}$,$-2.3\mathrm{e}{-07}$]
QR retraction	no guard	4.24e-06	$+0.231$ [$+0.166$,$+0.298$]
	endpoint-error guard	2.96e-06	$+0.129$ [$+0.048$,$+0.218$]
near-boundary start			
reference Exp	–	4.95e-06	–
chart Euler, 1 step	no guard	5.12e-06	$+0.002$ [-0.013,$+0.016$]
	endpoint-error guard	4.93e-06	-0.006 [-0.010,-0.001]
fp16 storage + restoration	no guard	1.99e-04	$+1.646$ [$+1.575$,$+1.718$]
	endpoint-error guard	4.95e-06	$+7.5\mathrm{e}{-06}$ [$-1.1\mathrm{e}{-04}$,$+1.2\mathrm{e}{-04}$]
fp32 storage + restoration	no guard	4.95e-06	$+3.8\mathrm{e}{-04}$ [$+1.5\mathrm{e}{-04}$,$+6.4\mathrm{e}{-04}$]
	endpoint-error guard	4.95e-06	$-1.4\mathrm{e}{-07}$ [$-5.9\mathrm{e}{-07}$,$+3.0\mathrm{e}{-07}$]
QR retraction	no guard	4.66e-06	$+0.002$ [-0.054,$+0.058$]
	endpoint-error guard	4.47e-06	-0.017 [-0.033,-0.001]

Table 11: The boundary-active stratum of Experiment H uses $\phi = 30^\circ$ and 10 trajectories per start. The two panels share the kernel and treatment columns. The upper panel compares the unguarded and endpoint-guard treatments through native infeasible proposals, feasible proposals rejected by the guard, fallback share (fallbacks per 3,000 proposals of the stratum), and the maxima of two different quantities, native error over the endpoint clearance (both at the float64 reference endpoint, Section D.1) and the decision ratio R_{acc}; these maxima need not occur at the same step. The lower panel reports the median final gap to F_{ref} and the paired $\Delta \log_{10}$ effect versus the reference exponential, with a seed-paired bootstrap interval computed only when every gap is resolved.

kernel	treatment	native infeasible	feasible & rejected	fallback share	max err/clearance; max R_{acc}
centre start					
reference Exp	–	–	–	–	–
chart Euler, 1 step	no guard	0	–	–	4.16e-02; 2.20e+00
	endpoint-error guard	0	13	0.4%	4.16e-02; 2.20e+00
fp16 storage + restoration	no guard	0	–	–	4.18e-03; 1.59e+03
	endpoint-error guard	0	2891	96.4%	4.69e-03; 6.16e+03
fp32 storage + restoration	no guard	0	–	–	4.82e-07; 7.52e-01
	endpoint-error guard	0	0	0.0%	4.82e-07; 7.52e-01
QR retraction	no guard	0	–	–	4.13e-01; 1.85e+00
	endpoint-error guard	0	10	0.3%	4.13e-01; 1.85e+00
near-boundary start					
reference Exp	–	–	–	–	–
chart Euler, 1 step	no guard	0	–	–	1.69e-01; 2.81e+00
	endpoint-error guard	0	16	0.5%	1.69e-01; 2.81e+00
fp16 storage + restoration	no guard	0	–	–	4.26e-03; 3.25e+03
	endpoint-error guard	0	2863	95.4%	4.77e-03; 6.66e+02
fp32 storage + restoration	no guard	0	–	–	5.36e-07; 8.12e-02
	endpoint-error guard	0	0	0.0%	5.36e-07; 8.12e-02
QR retraction	no guard	0	–	–	9.01e-01; 1.85e+00
	endpoint-error guard	0	10	0.3%	9.01e-01; 1.85e+00

kernel	treatment	gap k=300 (median)	$\Delta \log_{10}$ gap [CI]
centre start			
reference Exp	–	7.22e-06	–
chart Euler, 1 step	no guard	7.22e-06	-0.0380 [-0.1113,-0.0002]
	endpoint-error guard	7.22e-06	-0.0404 [-0.1190,-0.0002]
fp16 storage + restoration	no guard	7.23e-06	-0.162 [-0.389,$+0.004$]
	endpoint-error guard	7.22e-06	-0.009085 [-0.025551,$+0.000001$]
fp32 storage + restoration	no guard	7.22e-06	-0.0087219 [-0.0259464,-0.0000004]
	endpoint-error guard	7.22e-06	-0.0087219 [-0.0259464,-0.0000004]
QR retraction	no guard	7.18e-06	-0.1603 [-0.3854,-0.0004]
	endpoint-error guard	7.18e-06	-0.160 [-0.386,-0.001]
near-boundary start			
reference Exp	–	7.18e-06	–
chart Euler, 1 step	no guard	7.18e-06	$-2.1\mathrm{e}{-05}$ [$-3.5\mathrm{e}{-05}$,$-2.7\mathrm{e}{-06}$]
	endpoint-error guard	7.18e-06	$-1.2\mathrm{e}{-05}$ [$-1.7\mathrm{e}{-05}$,$-8.3\mathrm{e}{-06}$]
fp16 storage + restoration	no guard	7.19e-06	-0.010 [-0.035,$+0.012$]
	endpoint-error guard	7.18e-06	$+2.1\mathrm{e}{-06}$ [$-1.0\mathrm{e}{-06}$,$+7.5\mathrm{e}{-06}$]
fp32 storage + restoration	no guard	7.18e-06	$+1.4\mathrm{e}{-06}$ [$-3.5\mathrm{e}{-06}$,$+6.9\mathrm{e}{-06}$]
	endpoint-error guard	7.18e-06	$+1.4\mathrm{e}{-06}$ [$-3.5\mathrm{e}{-06}$,$+6.9\mathrm{e}{-06}$]
QR retraction	no guard	7.17e-06	$+6.0\mathrm{e}{-05}$ [$-5.8\mathrm{e}{-05}$,$+1.6\mathrm{e}{-04}$]
	endpoint-error guard	7.18e-06	$-2.5\mathrm{e}{-05}$ [$-6.1\mathrm{e}{-05}$,$+1.1\mathrm{e}{-05}$]

Table 12: Experiment H, interior stratum ($\phi = 60°$); panels and columns as in Table 11.

Treatment	$\Delta \log_{10}$ gap [95% CI]		infeasible; natively accepted	
	boundary-active	interior	boundary-active	interior
reference Exp (float64)	0 (reference)	0 (reference)	0; –	0; –
fp32 storage, no guard	$+4.1e{-}04$ [$+2.2e{-}04$,$+6.3e{-}04$]	$-4.4e{-}03$ [$-1.3e{-}02$,$+1.4e{-}06$]	0; 6,000	0; 6,000
fp16 storage, no guard	$+1.74$ [$+1.67$,$+1.81$]	-0.086 [-0.204,$+0.003$]	0; 6,000	0; 6,000
fp16 storage, 400-bit membership	$+1.74$ [$+1.67$,$+1.81$]	-0.086 [-0.201,$+0.002$]	0; 6,000	0; 6,000
fp16 storage, radial projection	$+1.74$ [$+1.67$,$+1.81$]	-0.086 [-0.203,$+0.002$]	0; 6,000	0; 6,000
fp16 storage, endpoint-error guard	$-7.4e{-}04$ [$-7.5e{-}03$,$+5.3e{-}03$]	$-4.5e{-}03$ [$-1.2e{-}02$,$+1.2e{-}06$]	382; 73	0; 246
fp32 storage, endpoint-error guard	$-5.5e{-}07$ [$-1.0e{-}06$,$-1.0e{-}07$]	$-4.4e{-}03$ [$-1.3e{-}02$,$+1.3e{-}06$]	0; 387	0; 6,000

Treatment	Reference endpoint calls / reuses per step		Online time [s] (oracle time [s])	
	boundary-active	interior	boundary-active	interior
reference Exp (float64)	1 / –	1 / –	2.87 (2.86)	2.55 (2.54)
fp32 storage, no guard	1 (workflow) / 0.00	1 (workflow) / 0.00	2.93 (2.90)	2.91 (2.88)
fp16 storage, no guard	1 (workflow) / 0.00	1 (workflow) / 0.00	3.00 (2.97)	2.80 (2.77)
fp16 storage, 400-bit membership	1 (workflow) / 0.00	1 (workflow) / 0.00	3.62 (2.92)	3.51 (2.82)
fp16 storage, radial projection	1 (workflow) / 0.00	1 (workflow) / 0.00	3.01 (2.97)	2.91 (2.87)
fp16 storage, endpoint-error guard	1 (workflow) / 0.99	1 (workflow) / 0.96	3.04 (2.94)	3.06 (2.96)
fp32 storage, endpoint-error guard	1 (workflow) / 0.94	1 (workflow) / 0.00	2.88 (2.77)	2.99 (2.89)

Table 13: The matched precision controls of Experiment H compare the storage family within each regime, using 20 trajectories per regime. The upper panel reports the paired $\Delta \log_{10}$ final-gap effect versus the reference exponential, with a seed-cluster bootstrap interval over the 10 seeds, and, as two separate counts, infeasible proposals and natively accepted proposals, each out of $6{,}000$. The lower panel reports reference-endpoint calls and reuses per step, together with the online time per 300-iteration trajectory and the oracle time in parentheses. The storage workflow computes one reference endpoint per step and reuses it on fallback. The timings are indicative measurements collected under concurrent load.

Kernel	Proposals	Infeasible	Membership rejects	Endpoint rejects
chart Euler, 1 step	12,000	0	0	327
fp16 storage	12,000	382	382	11,681
fp32 storage	12,000	0	0	5,613
QR retraction	12,000	0	0	40

Kernel	Both reject	Endpoint only	Accepted feasible
chart Euler, 1 step	0	327	11,673
fp16 storage	382	11,299	319
fp32 storage	0	5,613	6,387
QR retraction	0	40	11,960

Table 14: Experiment H, common-candidate comparison on the $12{,}000$ native proposals recorded for each kernel's endpoint-guard trajectories. Infeasibility and membership rejection use the sign of the 400-bit clearance; on the unchanged points, the endpoint rule accepts exactly when $R_{\mathrm{acc}} \leq 1$ after the domain and input checks pass, and all other proposals are rejected. "Endpoint only" counts feasible proposals rejected only by that rule; all endpoint-accepted proposals are feasible under the 400-bit check. Storage includes the stated polar-restoration workflow.

Kernel	Regime	e/Δ	$\widehat{E}/e$	$\Delta/\beta_{\mathrm{geo}}$	β_{geo}/B	$B/\underline{\beta}$
chart Euler, 1 step	boundary-active	6.6e-12 (1.4e-09; 1.7e-03)	1.7e+03 (5.6e+03; 2.1e+04)	1.2e+06 (7.6e+06; 1.4e+07)	7.6e+01 (1.4e+02; 1.5e+02)	1.0e+00 (1.0e+00; 1.0e+00)
	interior	2.5e-05 (3.6e-04; 1.3e-02)	4.6e+00 (7.0e+01; 6.0e+03)	9.3e+01 (1.6e+02; 2.4e+03)	7.5e+01 (1.4e+02; 1.5e+02)	1.0e+00 (1.0e+00; 1.0e+00)
fp16 storage	boundary-active	4.8e-01 (2.8e+00; 5.4e+00)	2.1e+01 (5.3e+01; 1.5e+02)	1.2e+06 (7.6e+06; 1.4e+07)	7.5e+01 (1.4e+02; 1.5e+02)	1.0e+00 (1.0e+00; 1.0e+00)
	interior	8.3e-04 (1.9e-03; 3.3e-03)	2.2e+01 (5.5e+01; 1.4e+02)	9.4e+01 (1.6e+02; 2.4e+03)	7.5e+01 (1.4e+02; 1.5e+02)	1.0e+00 (1.0e+00; 1.0e+00)
fp32 storage	boundary-active	6.3e-05 (2.9e-04; 6.5e-04)	2.2e+01 (5.4e+01; 1.3e+02)	1.2e+06 (7.5e+06; 1.4e+07)	7.5e+01 (1.4e+02; 1.5e+02)	1.0e+00 (1.0e+00; 1.0e+00)
	interior	1.0e-07 (2.2e-07; 3.7e-07)	2.2e+01 (5.3e+01; 1.3e+02)	9.3e+01 (1.6e+02; 2.3e+03)	7.5e+01 (1.4e+02; 1.5e+02)	1.0e+00 (1.0e+00; 1.0e+00)
QR retraction	boundary-active	5.0e-09 (2.8e-06; 1.8e-02)	1.8e+00 (4.2e+00; 7.9e+00)	1.1e+06 (6.7e+06; 1.3e+07)	7.5e+01 (1.4e+02; 1.5e+02)	1.0e+00 (1.0e+00; 1.0e+00)
	interior	7.8e-05 (1.3e-03; 5.6e-02)	1.8e+00 (3.1e+00; 3.7e+00)	8.5e+01 (1.5e+02; 1.6e+02)	7.5e+01 (1.4e+02; 1.5e+02)	1.0e+00 (1.0e+00; 1.0e+00)

Kernel	Regime	Tolerance active	Positive-error proposals	Current	Reference clearance	Reference error	Membership
chart Euler, 1 step	boundary-active	5,960	5,990	5,702	5,702	6,000	6,000
	interior	5,960	5,990	5,971	5,971	6,000	6,000
fp16 storage	boundary-active	5,960	6,000	73	73	148	5,618
	interior	5,960	6,000	246	246	1,257	6,000
fp32 storage	boundary-active	5,960	6,000	387	387	691	6,000
	interior	5,960	6,000	6,000	6,000	6,000	6,000
QR retraction	boundary-active	5,960	6,000	5,980	5,990	5,983	6,000
	interior	5,960	6,000	5,980	5,990	5,983	6,000

Table 15: Experiment H decomposes the conservatism on the recorded endpoint-guard proposals. Each cell of the upper panel gives the median and, in parentheses on the following lines, the 90th and 99th percentiles of the five factors of $\widehat{E}/\underline{\beta}$ defined in the text. These summaries use the proposals with positive evaluated error, clearance, and chord length; the ten excluded Euler proposals per regime are the first steps of the ten centre-start trajectories, on which the single chart-Euler step from the zero chart coordinate evaluates the same float64 expression as the reference endpoint, so their recorded native error is exactly zero; after this exclusion both Euler medians of β_{geo}/B are 75.5, printed as 7.5e+01 or 7.6e+01 by rounding, whereas the other rows have 75.25. Of the remaining chart-Euler proposals of the boundary-active regime, $3{,}439$ of $5{,}990$ have recorded errors below the proxy bound $(c_{\mathrm{out}} + T_k c_\Omega)u_{53}$ of Section D.1, so the e/Δ and $\widehat{E}/e$ entries of that row are resolution-limited. Among the positive-error proposals, the lower panel counts those on which the progress tolerance $c\eta^2 d^2$ is the smaller budget term. Using all $6{,}000$ proposals of the regime, it then reports the number that would be accepted under the current safe budget, under the reference clearance Δ (the endpoint clearance of Remark C.2 evaluated at the float64 reference endpoint) capped by the progress tolerance $c\eta^2 d^2$ (without the comparison margins), under the reference error e evaluated in 400-bit arithmetic against the current safe budget (an offline diagnostic), and by 400-bit membership.

Kernel	Treatment	Oracle [s]	Proposal [s]	Guard [s]	Output [ms]	Fallbacks new / reused	Re-test failed / max R^{ref}	Offline [s]
reference Exp (float64)	–	2.70	0.010	–	–	–	–	1.31
chart Euler, 1 step	no guard	2.73	0.022	0.000	0.12	–	–	1.30
	float64 membership	2.68	0.022	0.005	0.12	0 / 0	0 / n/a	1.29
	400-bit membership	2.75	0.022	0.645	0.15	0 / 0	0 / n/a	1.25
	endpoint-error guard	2.83	0.023	0.020	0.43	327 / 0	0 / 0.092	1.35
	radial projection	2.93	0.024	0.000	5.46	–	–	1.99
fp16 storage	no guard	2.87	0.032	0.000	0.13	–	–	1.40
	float64 membership	2.84	0.031	0.007	0.13	0 / 0	0 / n/a	1.39
	400-bit membership	2.87	0.031	0.661	0.15	0 / 0	0 / n/a	1.34
	endpoint-error guard	2.95	0.032	0.064	3.38	0 / 11681	0 / 0.092	2.00
	radial projection	2.92	0.032	0.000	6.60	–	–	2.05
fp32 storage	no guard	2.89	0.031	0.000	0.12	–	–	1.41
	float64 membership	2.81	0.030	0.006	0.13	0 / 0	0 / n/a	1.39
	400-bit membership	2.82	0.030	0.653	0.14	0 / 0	0 / n/a	1.32
	endpoint-error guard	2.83	0.031	0.068	1.62	0 / 5613	0 / 0.092	1.62
	radial projection	2.67	0.030	0.000	6.05	–	–	1.94
QR retraction	no guard	2.61	0.027	0.000	0.12	–	–	1.27
	float64 membership	2.67	0.028	0.006	0.13	0 / 0	0 / n/a	1.29
	400-bit membership	2.79	0.029	0.648	0.14	0 / 0	0 / n/a	1.25
	endpoint-error guard	2.82	0.029	0.010	0.17	40 / 0	0 / 5.3e-14	1.33
	radial projection	2.80	0.029	0.000	6.13	–	–	1.94

Table 16: Experiment H, indicative component costs per 300-iteration trajectory. Timing columns are means over the 40 trajectories of each treatment (output construction in milliseconds, the other columns in seconds; a guard time printed as 0.000 is below 0.5 ms); the fallback counts are totals over those 40 trajectories. Online policy cost is the sum of the oracle, native proposal, online guard, and output-construction columns, after converting the output-construction milliseconds to seconds, and includes every computation needed to produce the proposal, evaluate the selected guard (the 400-bit clearance is charged to the 400-bit membership guard), and construct the accepted output; a reference endpoint is reused without a second charge only when the storage workflow itself computed it, and its construction is included in the native-proposal column. The re-test column gives, per treatment, the number of fallbacks that failed their own re-test (Algorithm 2's fallback contract; all zero) and, for the endpoint guard, the largest ratio $R^{\mathrm{ref}} := (1+\tau)\widehat{E}^{\mathrm{ref}}/\underline{\beta}_k$ over its fallbacks. The offline column contains only computations unnecessary for that policy's decisions. Measurements were collected under concurrent load without warm-up or controlled BLAS threads and do not establish comparative runtime performance.

D.8 Baseline diagnostics and tolerance sensitivity (Experiment C)

Experiment C (inexact RFW diagnostics). Experiment C examines inexact Riemannian Frank–Wolfe methods on two families of instances; Table 17 and Figure 5 summarize the boundary-active instance for a curated subset of the methods. For each $j \in \{0, \dots, 49\}$, seed j generates a random centre R_c, a random unit vector v, and a nominal attitude $\bar{R} = \mathrm{Exp}_{R_c}(r_{\mathrm{n}} v)$. We use $r_{\mathrm{n}} = \rho + 10^\circ$ for the boundary-active instance and $r_{\mathrm{n}} = \rho/2$ for the interior instance. Each instance contains $N_{\mathrm{s}} = 50$ samples $R_i = \bar{R}\exp\widehat{\xi_i}$ with $\xi_i \sim \mathcal{N}(0, (10^\circ)^2 I)$. Samples with $d(R_c, R_i) \geq 105^\circ$ are

rejected, so that every distance $d(R, R_i)$ with $R \in \mathcal{C}_\rho$ stays below $165^\circ < \pi$, which keeps $\frac{1}{2}d^2(\cdot, R_i)$ smooth and geodesically convex on $\mathcal{C}_\rho$. The objective $F(R) = \frac{1}{2N_s}\sum_i d^2(R, R_i)$ has $L_F = 1$ and $\operatorname{grad} F(R) = -\frac{1}{N_s}\sum_i \operatorname{Log}_R R_i$; the constraint is $\mathcal{C}_\rho$ with $\rho = 60^\circ$, $\alpha_\rho = 0.433$, $D = 2\rho = 2.094$. The metric is bi-invariant, so randomizing the centre does not change the intrinsic geometry. The data are concentrated about one attitude with isotropic noise and the run starts at the chart centre, so the chart dynamics are nearly radial, and these instances provide limited evidence about integration error away from radial trajectories. They serve as a geometry and certification testbed, not as a demanding integration benchmark.

Experiment C oracles and schedule. Experiment C uses the numerical reference oracle and the objective reference of Section D.4. Because that oracle is not a certified global minimizer, a replicate with a 2000-point grid and halved tolerances is also reported, and oracle-dependent statements are diagnostics.

Perturbed oracles. The fixed coarse oracle uses 60 grid points and no refinement; the shrinking-perturbation oracle rotates the fine reference output along the boundary sphere by the angle $\psi_k = 0.35\sqrt{\eta_k}$. At an exact stationary boundary minimizer such a perturbation has quadratic value growth and hence an $O(\eta_k)$ error under a uniform second-derivative bound. Our reference oracle is numerical, so in both cases we measure $\widehat{\delta}_k$ against the fine reference oracle and record only the maximum of $\widehat{\delta}_k/\eta_k$ over the run, which on the reported runs is 0.082 for the shrinking-perturbation oracle and 1.65 for the fixed coarse oracle. Such a finite-run value is not a certificate of the error relative to the true linear minimum.

Schedule. All methods start at $x_0 = R_c$, use $\eta_k = 2/(k+3)$ unless stated (so $\eta_0 = 2/3 < 1$ and every tolerance is positive), run 300 iterations, and share the paired noise directions of seed $10^4 + j$.

Experiment C methods. The methods vary the endpoint approximation, the error direction, and the acceptance rule. The *fixed absolute error* method applies a perturbation of requested magnitude $\varepsilon_{\mathrm{abs}} = 10^{-2}$ rad at the reference geodesic point z, either in the radially outward direction $-\operatorname{Log}_z R_c/\|\cdot\|$, which is the most adverse direction for ball feasibility at a fixed small displacement, or in a random direction; we also test it with the constant step $\eta \equiv 0.05$. The *stress test* requests a perturbation of magnitude $\beta_k = \alpha\eta_k \min\{1-\eta_k, \eta_k\} d_k^2$ (Algorithm 2 with $c = \alpha$), outward or random. The *projection* baseline applies the outward absolute error and then projects metrically onto the ball. This combination is favourable in these nearly radial instances, so this baseline's small gap is an artefact of the injected direction and not a neutral advantage. We test the *QR retraction* with the open-loop schedule and with the shifted schedule $\eta_k = 2/(k+m)$, where $m = \lceil 2(\alpha+K)/\alpha \rceil = 4$ for the calibrated value $K = 0.40$. We also test the *Cayley retraction*, whose error is along the geodesic axis (Remark C.7). The *chart integrators* apply explicit Euler or classical RK4 to $\dot{w} = J_r(w)^{-1}\Omega$ over $[0, \eta_k]$ with n_{step} equal sub-steps and $\Omega = \operatorname{Log}_{x_k} q_k$. The variant "1 step" uses $n_{\mathrm{step}} = 1$, a first-order chart retraction. The variant "adaptive (step-doubling estimate)" doubles n_{step} until the step-doubling estimate is at most β_k and returns the $2n$ solution; the estimate is $2\|w_{2n} - w_n\|$ for Euler and $\|w_{2n} - w_n\|$ for RK4 ($n = n_{\mathrm{step}}$; safety factors 2 and 15 relative to the Richardson extrapolation), and it is an error indicator, not a certified bound, checked only a posteriori against the reference endpoint below. The variant "membership backtracking" doubles n_{step} until $d(x_{k+1}, R_c) \le \rho$ in float64. The variant "chart Euler [D.2: truncation AN, rounding CO]" (the labels of Table 18) is the method of Proposition D.2 with the reserve policy above. Work is counted in right-hand-side evaluations per iteration, *including rejected trials* (an accepted resolution of $n = 4$ after trials at 1 and 2 costs 7); the accepted-sub-step histogram is reported separately. The *coordinate-storage* variants round the chart coordinate $w = \operatorname{Log}_{R_c} z$ of the reference endpoint to IEEE half or single precision and map it back using $\tilde{z} = R_c \exp \widehat{\mathrm{fl}(w)}$; the coordinate-storage error is bounded by the proved storage bound of Lemma D.3. The variant "guarded coordinate rounding" uses half precision when its bound is at most β_k, single precision when the single-precision bound is, and double precision otherwise. The *matrix-storage* variants store the reference endpoint matrix entrywise in half precision and restore it to $\mathrm{SO}(3)$ by the polar projection, following the storage-and-restoration workflow of Lemma D.3. We test an unguarded variant ("fp16 matrix storage + restoration") and a guarded variant that chooses the storage precision by the full representation ledger following Lemma D.3 against the safe budget $\underline{\beta}_k$ ("guarded storage precision + restoration"). The fp64 fallback of the guarded storage method skips reduced-precision storage and restoration and

is accepted only when its full conditional enclosure $(c_{\text{out}} + T_k c_\Omega)u_{53}$ passes the same safeguarded comparison against the safe budget; the reduced-precision coordinate branches remain storage-only ablations because their float64 remap error is not included.

method [status]	gap k=300	$\Delta \log_{10}$ gap [CI]	runs viol. [CI]	map work
reference Exp (float64) [CO]	4.0e-06	–	0/50 [0.00,0.07]	–
fixed abs. error, outward [ung.]	n.c. (infeasible)	–	50/50 [0.93,1.00]	–
full budget outward [stress]	2.6e-06	−0.19020 [−0.19021,−0.19019]	0/50 [0.00,0.07]	–
full budget, shrinking-perturbation LMO [stress]	1.4e-04	+1.55 [+1.54,+1.55]	0/50 [0.00,0.07]	–
QR, shifted schedule [CA]	1.3e-05	+0.510 [+0.506,+0.513]	0/50 [0.00,0.07]	–
chart Euler, membership backtracking [DI]	4.0e-06	−6.3e−08 [−9.6e−08,−4.0e−08]	0/50 [0.00,0.07]	1.0 RHS/it.
chart Euler, enclosure [D.2: AN+CO]	4.0e-06	−6.3e−08 [−9.7e−08,−4.0e−08]	0/50 [0.00,0.07]	1.2 RHS/it.
fp16 storage + restoration [ung.]	2.8e-04	+1.85 [+1.81,+1.90]	0/50 [0.00,0.07]	fp64 in 0% of it.
guarded storage precision [AN+CO]	4.0e-06	−4.2e−05 [−1.2e−04,+3.0e−05]	0/50 [0.00,0.07]	fp64 in 95% of it.

Table 17: This curated subset of Experiment C uses $\rho = 60°$, the boundary-active instance, 50 seeds, and 300 iterations. The table reports the median final gap to F_{ref}, the paired $\Delta \log_{10}$ gap versus the reference exponential, the number of runs violating, and map work. The paired effect and the violation frequency carry 95% CIs. The algebraic difference $F(x) - F_{\text{ref}}$ can be evaluated at an infeasible point, but it is not a feasible suboptimality measure; a method with an infeasible final iterate on any seed is therefore marked "n.c.". Map work is measured in right-hand-side evaluations per iteration, including rejected trials, or in the fp64 share. Tables 18 and 19 report all 23 methods with slopes, maximum violations, and timings. The shrinking-perturbation LMO has measured $\max_k \widehat{\delta}_k/\eta_k \le 0.082$ over the reported runs, relative to the numerical reference oracle; this is not an asymptotic or certified true-oracle bound. Status classes as in Section D.1.

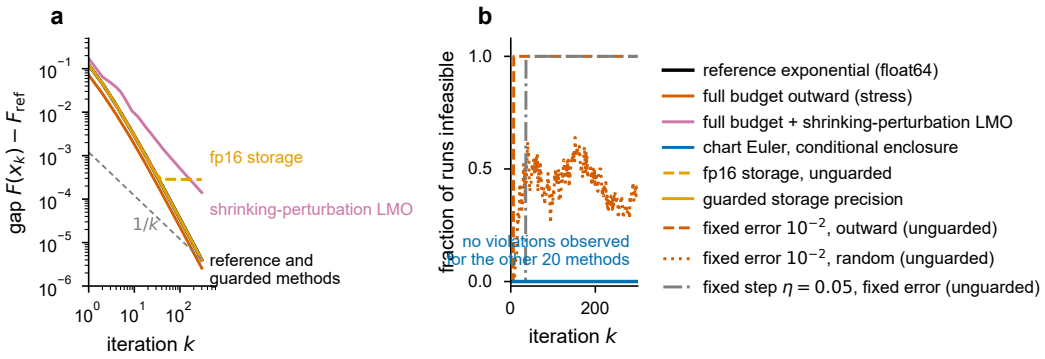

Figure 5: Experiment C, boundary-active instance. (a) Median gap to F_{ref} ($1/k$ guide dotted). (b) Fraction of runs whose current iterate is infeasible at iteration k (reporting threshold 10^{-10}); the twenty methods other than the three unguarded fixed-error variants show no observed violation, and this group contains unguarded native kernels as well as guarded ones, so the observation is not itself an arithmetic certificate. Figure 6 reports the budgets, enclosures, clearances, and ratios.

Experiment C statistics and numerical floor. The statistics distinguish observed feasibility, resolved objective gaps, and the numerical resolution floor. A violation is counted when $d(x_k, R_c) - \rho > 10^{-10}$ (float64 resolution about 10^{-15}; raw maxima are reported). "Max violation" is the maximum of $d(x_k, R_c) - \rho$ over all iterations and seeds (clipped at 0); "runs violating" counts seeds with at least one violation and carries a 95% Clopper–Pearson interval, which quantifies the chosen random experiment and is not a certificate; a first-violation iteration is a median over the violating runs only; "no observed violation" means that no violation larger than 10^{-10} was observed in the 50 runs and is not a universal statement; reported maxima are positive-part maxima; bootstrap intervals are descriptive and not multiplicity-adjusted.

Gaps, paired effects, and slopes. We compare each method with the reference exponential on the same seed by averaging over seeds the paired difference $\log_{10}(F(x_{300}) - F_{\text{ref}})$ minus the same quantity for the reference method, with a 95% bootstrap interval (4000 resamples of seeds). An effect of 0 corresponds to a geometric-mean ratio of paired final gaps equal to 1, and an effect of +1

corresponds to a geometric-mean ratio of 10. Log–log slopes of the median gap over $k \in [30, 300]$ (interval fixed in advance; least squares on the median curve) are reported as empirical descriptors, not as rate theorems. We report a method's aggregate constrained objective gap only when the final iterate of every seed is feasible under the stated reporting rule. If any final iterate in a method's seed group is infeasible, the aggregate final gap, the paired log-gap effect, and the slope are omitted and the row is marked "not comparable (infeasible)", since an infeasible point can have a lower objective value than the constrained optimum; the runs-violating count instead records whether a run violated at any iteration. For a feasible iterate, a difference of magnitude at most 10^{-13} is unresolved at the reference precision, and a difference below -10^{-13} is flagged as a reference inconsistency requiring investigation. Paired logarithmic effects are computed only when both gaps exceed 10^{-13} on every included seed; no clipping or pseudocount is used, and feasibility status is the first branch of the formatting rule shared by every generated table.

Display convention. One display convention is shared by every generated table: an effect below half a unit of the displayed precision is printed, together with its interval, in scientific notation; otherwise the effect and its interval endpoints receive the decimals needed for both endpoints to print nonzero and distinct. These display choices do not increase the numerical accuracy of the underlying evaluations, and no value is clipped. Paired effects far below the resolution of the objective evaluation, such as the RK4 effects of order 10^{-11} in $\log_{10}$ units, indicate agreement with the reference to working precision and support no superiority claim.

Threshold crossings and requested perturbations. The saved budget curves locate the threshold crossings. The median accepted budget of the stress test falls below the reporting threshold 10^{-10} at $k = 55$ and below 10^{-13} at $k = 176$, reaching $4.3 \cdot 10^{-15}$ at $k = 300$. For the reference-exponential trajectory, the corresponding crossings occur at $k = 64$ and $k = 204$, and the final budget is $1.0 \cdot 10^{-14}$. From about $k = 176$ on, the stress test requests a perturbation that approaches the resolution of the float64 construction and distance evaluator, so we distinguish the requested perturbation from its realization and do not resolve its feasibility excursions below the reporting threshold. The guarantee applies to an implementation satisfying the stated endpoint-error enclosure; for the conditionally certified map the relevant quantities are the safe budget $\underline{\beta}_k$, the floor E_n^{fl} at the accepted trial, and the reserve threshold, reported from the checker's ledger (on the reported 300-iteration runs no reserve request occurred; the first occurs at $k = 341$ in Experiment D). For the rounding modes we also record the ratio of the numerically evaluated endpoint error relative to the reference endpoint to the stated bound, over iterations where the bound exceeds 10^{-13}. Figure 6 shows these evaluated endpoint errors relative to the reference endpoint, not estimator values. The step-doubling variants use error indicators rather than proved bounds, whereas the enclosure-based chart-Euler variant uses Proposition D.2 under Assumption D.1 and its accepted steps are checked independently in 400-bit arithmetic. The "map work" column counts right-hand-side evaluations for integrators and the fraction of fp64 iterations for storage modes; the appendix table adds mean wall-clock time per run with the oracle and map shares (indicative timings, see the environment paragraph; the oracle dominates every method, and the chart integrators cost about as much as the reference exponential, so no speed-up is claimed).

Experiment C tolerance-halving replicate and interior instance. We also re-run the boundary-active instance with 2000 grid points and halved Nelder–Mead tolerances (Table 21), and report the interior-optimum instance ($r = \rho/2$) in Table 20. On the interior instance none of the 13 methods run has an observed violation, including the unguarded absolute-error methods. The clearance of the reference endpoint stays near 0.5 rad, providing a substantially larger feasibility margin than in the boundary-active instance, even as the acceptance budget decreases.

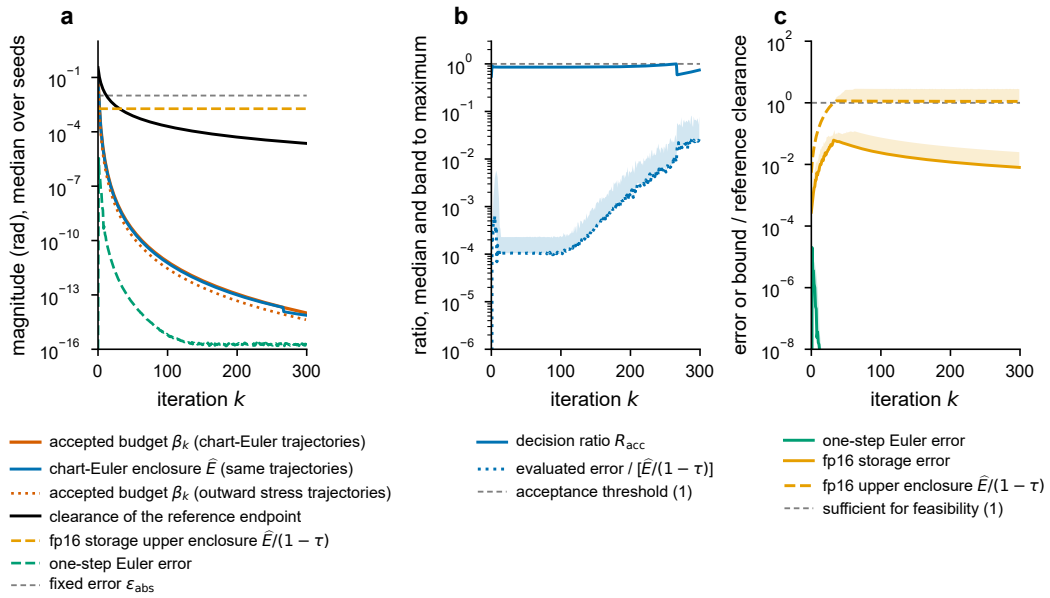

Figure 6: Experiment C, boundary-active instance. Only the chart-Euler enclosure and budget curves are paired quantities that enter an acceptance decision; every other curve is a diagnostic. Each curve shows the median over seeds, with a band extending to the maximum. (a) The accepted budget and the evaluated enclosure $\widehat{E}$ of the same accepted chart-Euler trajectories (paired records; the per-seed ratio of enclosure to budget never exceeds one), the accepted budget of the outward stress-test trajectories, drawn separately for scale comparison only, the clearance of the reference endpoint on the reference trajectories, the fp16 storage upper enclosure $\widehat{E}/(1-\tau)$ on the unguarded fp16 trajectories, and the fixed and native error levels. (b) For chart Euler under the stated conditional enclosure, the decision ratio $R_{\mathrm{acc}} = (1+\tau)\widehat{E}/((1-\tau)\widehat{B})$, which the guard accepts precisely when it is at most one after the domain and input checks succeed, together with the endpoint error, evaluated in float64 relative to the reference endpoint, divided by the conditional upper enclosure $\widehat{E}/(1-\tau)$, a numerical consistency check of the enclosure on the sampled steps (the 400-bit check of Section D.2 covers every accepted step). (c) The native endpoint errors of unguarded one-step Euler and fp16 storage, together with the conditional storage enclosure $\widehat{E}/(1-\tau)$, each divided by the clearance of the reference endpoint. For the corresponding mathematical error or valid bound, a ratio at most one is sufficient for feasibility of the corresponding proposal; a larger ratio is inconclusive. The plotted float64 evaluations are diagnostic, and this threshold is not the guard's acceptance test.

method [status; ung. = unguarded]	gap k=300 (n.c. = not comparable)	$\Delta \log_{10}$ gap [CI]	slope $[30, 300]$
reference Exp (float64) [CO, nominal]	4.01e-06	–	-1.97
fixed abs. err. 10^{-2}, outward [ung.]	n.c. (infeasible)	–	–
fixed abs. err. 10^{-2}, random [ung.]	n.c. (infeasible)	–	–
full budget, outward [stress]	2.59e-06	-0.19020 [-0.19021,-0.19019]	-1.97
full budget, random [stress]	4.68e-06	$+0.07$ [$+0.05$,$+0.10$]	-1.97
full budget + fixed coarse LMO [stress]	5.14e-03	$+3.09$ [$+3.04$,$+3.14$]	-0.04
full budget + shrinking-perturbation LMO [stress]	1.42e-04	$+1.55$ [$+1.54$,$+1.55$]	-1.19
fixed step $\eta = 0.05$, abs. err. [ung.]	n.c. (infeasible)	–	–
abs. err. + metric projection [ung.]	unresolved ($\leq 10^{-13}$)	– (unres.)	–
QR retraction, open-loop [ung.]	5.20e-06	$+0.12$ [$+0.11$,$+0.12$]	-1.97
QR retraction, shifted [CA: calibrated K]	1.30e-05	$+0.510$ [$+0.506$,$+0.513$]	-1.96
Cayley retraction [AN: on-axis]	4.33e-06	$+0.0331597$ [$+0.0331596$,$+0.0331598$]	-1.97
chart Euler, 1 step [ung.]	4.01e-06	$-6.3\text{e}{-08}$ [$-9.9\text{e}{-08}$,$-4.0\text{e}{-08}$]	-1.97
chart Euler, adaptive [CA: step-doubling est.]	4.01e-06	$-3.2\text{e}{-08}$ [$-5.0\text{e}{-08}$,$-2.1\text{e}{-08}$]	-1.97
chart Euler, membership backtracking [DI: float64 test]	4.01e-06	$-6.3\text{e}{-08}$ [$-9.6\text{e}{-08}$,$-4.0\text{e}{-08}$]	-1.97
chart Euler [D.2: truncation AN, rounding CO]	4.01e-06	$-6.3\text{e}{-08}$ [$-9.7\text{e}{-08}$,$-4.0\text{e}{-08}$]	-1.97
chart RK4, 1 step [ung.]	4.01e-06	$-7.1\text{e}{-11}$ [$-1.6\text{e}{-10}$,$+1.9\text{e}{-11}$]	-1.97
chart RK4, adaptive [CA: step-doubling est.]	4.01e-06	$-3.7\text{e}{-11}$ [$-1.2\text{e}{-10}$,$+5.2\text{e}{-11}$]	-1.97
fp16 chart coordinates [ung.]	5.68e-04	$+2.15$ [$+2.14$,$+2.16$]	0.00
fp32 chart coordinates [ung.]	4.09e-06	$+0.006$ [$+0.005$,$+0.008$]	-1.97
guarded coordinate rounding [AN: storage only, remap excluded]	4.01e-06	$+1.6\text{e}{-04}$ [$-1.7\text{e}{-04}$,$+4.7\text{e}{-04}$]	-1.97
fp16 matrix storage + restoration [ung.]	2.82e-04	$+1.85$ [$+1.81$,$+1.90$]	-0.02
guarded storage precision + restoration [ledger: storage AN, nominal/restoration CO]	4.01e-06	$-4.2\text{e}{-05}$ [$-1.2\text{e}{-04}$,$+3.0\text{e}{-05}$]	-1.97

Table 18: Experiment C (diagnostic) compares all 23 methods on the boundary-active instance with $\rho = 60°$, 50 seeds, and 300 iterations. The default schedule is $\eta_k = 2/(k+3)$; the fixed-step method uses $\eta = 0.05$ and the shifted QR method uses $\eta_k = 2/(k+4)$. This table reports the median final gap to F_{ref}, the paired $\Delta \log_{10}$ gap versus the reference exponential, and the fitted log–log slope of the median gap over $k \in [30, 300]$; Table 19 continues with the positive-part maximum violation, the number of runs violating, map work, and the mean wall-clock seconds per run with the oracle and map shares; these timings are indicative. Gaps at or below 10^{-13} are marked unresolved. Paired effects carry 95% bootstrap CIs and are omitted when any included gap is unresolved. Violation frequencies also carry 95% CIs. The [AN] label of the Cayley row refers to the exact-arithmetic on-axis argument of Remark C.7; its float64 evaluation is unguarded. The status classes [AN]/[CO]/[DI]/[CA]/[ung.]/[stress] are defined in Section D.1.

method [status; ung. = unguarded]	max viol. [rad]	runs viol. [CI]	map work	time [s] (oracle/map)
reference Exp (float64) [CO, nominal]	0.0e+00	0/50 [0.00,0.07]	–	3.5 (3.4/0.03)
fixed abs. err. 10^{-2}, outward [ung.]	2.5e-01	50/50 [0.93,1.00]	–	3.5 (3.3/0.03)
fixed abs. err. 10^{-2}, random [ung.]	8.5e-02	50/50 [0.93,1.00]	–	3.5 (3.4/0.03)
full budget, outward [stress]	0.0e+00	0/50 [0.00,0.07]	–	3.5 (3.4/0.03)
full budget, random [stress]	0.0e+00	0/50 [0.00,0.07]	–	3.5 (3.4/0.03)
full budget + fixed coarse LMO [stress]	0.0e+00	0/50 [0.00,0.07]	–	3.5 (3.4/0.03)
full budget + shrinking-perturbation LMO [stress]	0.0e+00	0/50 [0.00,0.07]	–	6.9 (6.8/0.03)
fixed step $\eta = 0.05$, abs. err. [ung.]	2.0e-01	50/50 [0.93,1.00]	–	3.5 (3.4/0.03)
abs. err. + metric projection [ung.]	8.9e-16	0/50 [0.00,0.07]	–	3.5 (3.3/0.04)
QR retraction, open-loop [ung.]	0.0e+00	0/50 [0.00,0.07]	–	3.6 (3.5/0.03)
QR retraction, shifted [CA: calibrated K]	0.0e+00	0/50 [0.00,0.07]	–	3.7 (3.5/0.03)
Cayley retraction [AN: on-axis]	0.0e+00	0/50 [0.00,0.07]	–	3.6 (3.5/0.01)
chart Euler, 1 step [ung.]	0.0e+00	0/50 [0.00,0.07]	1.0 RHS/it.	3.6 (3.5/0.04)
chart Euler, adaptive [CA: step-doubling est.]	0.0e+00	0/50 [0.00,0.07]	3.0 RHS/it.	3.6 (3.5/0.06)
chart Euler, membership backtracking [DI: float64 test]	0.0e+00	0/50 [0.00,0.07]	1.0 RHS/it.	3.6 (3.5/0.05)
chart Euler [D.2: truncation AN, rounding CO]	0.0e+00	0/50 [0.00,0.07]	1.2 RHS/it.	3.9 (3.7/0.08)
chart RK4, 1 step [ung.]	0.0e+00	0/50 [0.00,0.07]	4.0 RHS/it.	3.6 (3.5/0.07)
chart RK4, adaptive [CA: step-doubling est.]	0.0e+00	0/50 [0.00,0.07]	12.0 RHS/it.	3.7 (3.5/0.12)
fp16 chart coordinates [ung.]	0.0e+00	0/50 [0.00,0.07]	fp64 in 0% of it.	3.8 (3.7/0.04)
fp32 chart coordinates [ung.]	0.0e+00	0/50 [0.00,0.07]	fp64 in 0% of it.	4.0 (3.8/0.04)
guarded coordinate rounding [AN: storage only, remap excluded]	0.0e+00	0/50 [0.00,0.07]	fp64 in 93% of it.	4.0 (3.8/0.05)
fp16 matrix storage + restoration [ung.]	0.0e+00	0/50 [0.00,0.07]	fp64 in 0% of it.	4.0 (3.8/0.07)
guarded storage precision + restoration [ledger: storage AN, nominal/restoration CO]	0.0e+00	0/50 [0.00,0.07]	fp64 in 95% of it.	4.0 (3.8/0.07)

Table 19: Experiment C (diagnostic), boundary-active instance: violation, work, and timing columns for the methods of Table 18, with the conventions stated there.

method [status]	gap k=300	$\Delta \log_{10}$ gap [CI]	max viol. [rad]	runs viol. [CI]
reference Exp (float64) [CO, nominal]	8.71e-06	–	0.0e+00	0/50 [0.00,0.07]
fixed abs. err. 10^{-2}, outward [ung.]	6.70e-06	$+0.06$ [-0.14,$+0.26$]	0.0e+00	0/50 [0.00,0.07]
fixed abs. err. 10^{-2}, random [ung.]	6.94e-05	$+1.01$ [$+0.77$,$+1.27$]	0.0e+00	0/50 [0.00,0.07]
full budget, outward [stress]	4.90e-06	-0.08 [-0.42,$+0.27$]	0.0e+00	0/50 [0.00,0.07]
fp16 chart coordinates [ung.]	5.82e-06	-0.37 [-0.72,-0.04]	0.0e+00	0/50 [0.00,0.07]
guarded coordinate rounding [AN: storage only, remap excluded]	7.83e-06	-0.05 [-0.17,$+0.07$]	0.0e+00	0/50 [0.00,0.07]
fp16 matrix storage + restoration [ung.]	6.10e-06	-0.15 [-0.46,$+0.16$]	0.0e+00	0/50 [0.00,0.07]
guarded storage precision + restoration [ledger: storage AN, nominal/restoration CO]	8.54e-06	-0.04 [-0.14,$+0.05$]	0.0e+00	0/50 [0.00,0.07]
chart Euler, 1 step [ung.]	8.71e-06	$-2.1\mathrm{e}{-03}$ [$-4.7\mathrm{e}{-02}$,$+3.9\mathrm{e}{-02}$]	0.0e+00	0/50 [0.00,0.07]
chart Euler, adaptive [CA: step-doubling est.]	8.71e-06	-0.0165 [-0.0448,$+0.0005$]	0.0e+00	0/50 [0.00,0.07]
chart Euler [D.2: truncation AN, rounding CO]	8.71e-06	-0.04 [-0.16,$+0.03$]	0.0e+00	0/50 [0.00,0.07]
QR retraction, open-loop [ung.]	6.20e-06	-0.23 [-0.56,$+0.09$]	0.0e+00	0/50 [0.00,0.07]
abs. err. + metric projection [ung.]	6.70e-06	$+0.06$ [-0.14,$+0.27$]	0.0e+00	0/50 [0.00,0.07]

Table 20: The interior-optimum instance of Experiment C uses $\rho = 60^\circ$, 50 seeds, 300 iterations, and $\eta_k = 2/(k+3)$. This run comprises 13 of the 23 methods. The table reports the final gap, paired effect, maximum violation, and runs violating, with the conventions of Table 18.

method [status]	gap k=300	$\Delta \log_{10}$ gap [CI]	max viol. [rad]	runs viol. [CI]
reference Exp (float64) [CO, nominal]	4.01e-06	–	0.0e+00	0/50 [0.00,0.07]
fixed abs. err. 10^{-2}, outward [ung.]	n.c. (infeasible)	–	2.5e-01	50/50 [0.93,1.00]
full budget, outward [stress]	2.59e-06	-0.19020 [-0.19021,-0.19019]	0.0e+00	0/50 [0.00,0.07]
full budget + shrinking-perturbation LMO [stress]	1.42e-04	$+1.55$ [$+1.54$,$+1.55$]	0.0e+00	0/50 [0.00,0.07]
fp16 chart coordinates [ung.]	5.68e-04	$+2.15$ [$+2.14$,$+2.16$]	0.0e+00	0/50 [0.00,0.07]
guarded coordinate rounding [AN: storage only, remap excluded]	4.01e-06	$+1.6\text{e}{-04}$ [$-1.7\text{e}{-04}$,$+4.9\text{e}{-04}$]	0.0e+00	0/50 [0.00,0.07]
fp16 matrix storage + restoration [ung.]	2.82e-04	$+1.85$ [$+1.81$,$+1.90$]	0.0e+00	0/50 [0.00,0.07]
guarded storage precision + restoration [ledger: storage AN, nominal/restoration CO]	4.01e-06	$-4.2\text{e}{-05}$ [$-1.1\text{e}{-04}$,$+3.2\text{e}{-05}$]	0.0e+00	0/50 [0.00,0.07]
chart Euler, adaptive [CA: step-doubling est.]	4.01e-06	$-3.2\text{e}{-08}$ [$-5.0\text{e}{-08}$,$-2.1\text{e}{-08}$]	0.0e+00	0/50 [0.00,0.07]
chart Euler [D.2: truncation AN, rounding CO]	4.01e-06	$-6.3\text{e}{-08}$ [$-9.8\text{e}{-08}$,$-4.0\text{e}{-08}$]	0.0e+00	0/50 [0.00,0.07]
QR retraction, open-loop [ung.]	5.20e-06	$+0.12$ [$+0.11$,$+0.12$]	0.0e+00	0/50 [0.00,0.07]
QR retraction, shifted [CA: calibrated K]	1.30e-05	$+0.510$ [$+0.506$,$+0.513$]	0.0e+00	0/50 [0.00,0.07]

Table 21: Experiment C, boundary-active instance, tolerance-halving replicate (2000-point oracle grid, halved refinement tolerances). This run comprises 12 of the 23 methods, with the schedule $\eta_k = 2/(k+3)$ except for the shifted QR method, which uses $\eta_k = 2/(k+4)$; columns as in Table 20.

D.9 Precision continuation and the fixed-N policy (Experiments D and E)

Experiment D (precision escalation beyond the acceptance limit of the stated float64 enclosure). This experiment uses the same boundary-active instances, oracle, and schedule as Experiment C, with 10 seeds and 1,500 iterations. We compare four methods. The first is the float64 kernel with the reserve policy and the reference-exponential fallback. The second is the precision-adaptive kernel of Proposition D.4, using $\theta = \frac{1}{2}$, $\theta_{\mathrm{stop}} = \frac{1}{4}$, and mpmath binary floating point with b mantissa bits; its precision never decreases, each iterate is stored in b bits, and the chart centre and oracle targets are decoded as polar factors. The two unguarded controls are one-step float64 chart Euler and the float64 reference exponential, both continued through all iterations. At every step, a 400-bit computation evaluates the endpoint error against the exact endpoint of the represented chord (logarithm recomputed in 400-bit arithmetic), the stored-rate endpoint error, the rate discrepancy, and the slack $\rho - d(x_k, R_c)$ of the represented iterate; these values are 400-bit estimates, not interval enclosures.

The run records show that the float64 kernel first exits the refinement loop with a reserve request at $k = 341$; Table 22 and Figure 7 show that it returns *cannot certify* at $k = 448$ on every seed, when the fallback enclosure $E^{\mathrm{ref}} = 9.2 \cdot 10^{-16}$ exceeds the safe budget. Across the ten seeds, the fixed-precision routine completed 4480 outer updates before stopping; the work ledger in Table 23 records 6820 right-hand-side evaluations, including rejected chart trials, that is, 1.52 evaluations per completed outer step, and 1070 separately counted reference-exponential fallbacks, with the accepted-resolution histogram; the terminal failed attempts are excluded from these totals. Its 400-bit error-to-enclosure ratio rises to 0.15 near the acceptance limit of the stated float64 enclosure.

The adaptive policy first raises the precision at $k = 324$ and reaches $k = 1{,}500$ on every seed with at most 67 bits (57, 63, and 67 at $k = 500$, 1,000, 1,500, following $\log_2(K_{\mathrm{fl}}(n)/((1-\theta)\underline{\beta}_k))$), at most 8 sub-steps (8 at $k = 0$ and 2 thereafter; the explicit half-budget allocation uses two sub-steps where the fixed-precision routine often accepts one), and 2.0 right-hand-side evaluations per iteration. Every accepted enclosure passed the 400-bit consistency check (maximum ratio 0.082; rate discrepancy at most $1.4\,u_b$), and the final gap is $1.6 \cdot 10^{-7}$ (median) against $1.8 \cdot 10^{-6}$ for float64 at its stop. A b-bit step costs about 4.6 ms against 0.3 ms in float64 (indicative timings under concurrent load), a robustness cost, not a speed claim.

The unguarded controls, continued to $k = 1{,}500$, remain feasible with a minimum 400-bit slack of about $9.3 \cdot 10^{-7}$ rad and reach the same final gap. The adaptive method therefore continues to satisfy the conditional acceptance test after the stated float64 enclosure becomes too large for the budget, and the benefit shown by this experiment is continued conditional certification. Failure of that enclosure to meet the acceptance budget is not evidence that ordinary float64 computation cannot solve the problem safely at this horizon. The oracle is the uncertified numerical reference of Experiment C and the schedule uses $m = 3$, including an initial step $\eta_0 = \frac{2}{3}$ in the finite phase excluded from Theorem D.5. The experiment therefore illustrates finite-precision continuation under the implementation's model with the economical policy, without instantiating Theorem D.5.

method	runs stopped (iteration)	max bits	final gap (median)
fixed float64 (53 bits), reserve policy	10/10 ($k = 448$)	53	1.81e-06
precision-adaptive (Prop. D.4, $\theta = \frac{1}{2}$)	0/10	67	1.62e-07
one-step chart Euler, unguarded (control)	0/10	53	1.62e-07
reference exponential, unguarded (control)	0/10	53	1.62e-07

method	max err$/E$	min slack [rad]	max endpoint err [rad]
fixed float64 (53 bits), reserve policy	0.148	1.0e-05	7.2e-06
precision-adaptive (Prop. D.4, $\theta = \frac{1}{2}$)	0.082	9.3e-07	3.6e-06
one-step chart Euler, unguarded (control)	–	9.3e-07	7.2e-06
reference exponential, unguarded (control)	–	9.3e-07	3.2e-16

Table 22: Experiment D compares fixed float64 and precision-adaptive certification over $1{,}500$ iterations (10 seeds, boundary-active instance, $\rho = 60^\circ$), with two unguarded controls. The upper panel reports stopping status, maximal precision, and median final gap. The lower panel reports error ratios, slacks, and endpoint errors, all recomputed in 400-bit arithmetic against the exact endpoint of the represented chord; the ratio column divides the 400-bit endpoint error by the evaluated enclosure $\widehat{E}$.

metric (fixed float64 kernel, reserve policy, ten seeds)	value
completed outer steps	4480
accepted chart-Euler outputs	3410
accepted with $n = 1$ sub-step	2660
accepted with $n = 2$ sub-steps	640
accepted with $n = 4$ sub-steps	110
reference-exponential fallbacks	1070
right-hand-side evaluations on accepted-Euler steps (including rejected trials)	5350
right-hand-side evaluations on the failed trials preceding a fallback	1470
total counted right-hand-side evaluations	6820
right-hand-side evaluations per completed outer step	1.52

Table 23: Work ledger of the fixed float64 kernel over the ten seeds of Experiment D. The table reports completed outer steps, accepted chart-Euler outputs with their accepted-resolution histogram, reference-exponential fallbacks, and right-hand-side evaluations on accepted-Euler steps (including rejected trials) and on the failed trials preceding a fallback. The 1070 reference-exponential evaluations and the terminal *cannot certify* attempt are not included in the right-hand-side counts.

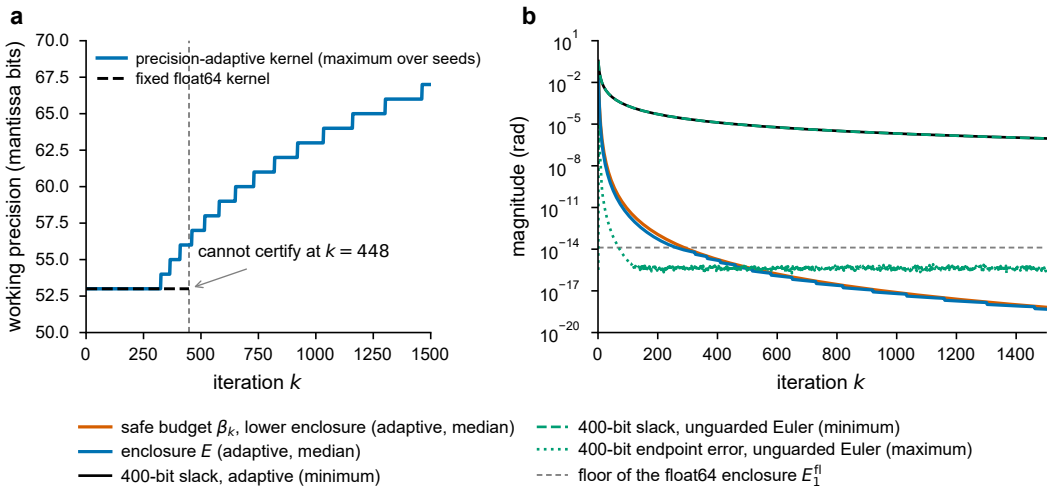

Figure 7: Experiment D. (a) Working precision of the adaptive policy (maximum over seeds); the float64 kernel ends with *cannot certify*. (b) Safe budget and enclosure of the adaptive kernel (medians over seeds), the 400-bit slack of the adaptive and unguarded iterates (minima), and the unguarded one-step Euler endpoint error (maximum), against the rounding component E_1^{fl} of the stated float64 enclosure; the curves are different aggregations and not pointwise inequalities between the quantities of one trajectory.

Experiment E (the cost of the fixed-N sufficient policy). Experiment E measures the cost of the fixed-N, one-bit-increment policy analysed in Theorem D.5. It uses $N = 256$ (the theorem's value in this configuration, $\theta = \frac{1}{2}$, $s = 0.1$) on five boundary-active instances for 300 iterations, with the same uncertified reference oracle and $m = 3$; these oracle and schedule choices lie outside the theorem's hypotheses, so the experiment checks the policy's implementation and cost. All 1,500 outer iterations completed with an accepted update (Table 24). The precision rose from 53 to at most 56 bits, with 15 failed attempts followed by one-bit increases across these iterations. The kernel used about 259 right-hand-side evaluations per iteration against 2.0 for the economical policy, every accepted enclosure passed the 400-bit consistency check (maximum ratio 0.40), and a step took about 92 ms. These results measure the cost of the implemented fixed-N policy; Experiment F next tests a theorem-aligned represented instance.

fixed-N sufficient policy	value
seeds × iterations	5 × 300
sub-steps per attempt N	256
maximal precision [bits]	56
failed attempts (each followed by a one-bit increase)	15
right-hand-side evaluations per iteration	259
largest 400-bit error$/E$	4.0e-01
smallest 400-bit slack [rad]	2.3e-05
step time [ms], indicative	92

Table 24: Experiment E applies the fixed-N, one-bit-increment sufficient policy of Theorem D.5 to the boundary-active instances (5 seeds, 300 iterations, $N = 256$). The table reports maximum precision, failed attempts (each followed by a one-bit increase), right-hand-side evaluations per iteration, the 400-bit error-to-enclosure ratio, the minimum 400-bit slack, and the indicative step time.

D.10 The represented exact-oracle instance (Experiment F)

Represented problem. Experiment F is a represented instance that meets the oracle and schedule hypotheses of Theorem D.5, conditional on Assumption D.1. It uses a same-axis rotation ball with an exact represented oracle. The centre is $R_c = I$ and the axis $u = e_3$. A stored binary block

matrix $A_q = \left[\begin{smallmatrix} 1/2 & -s_q & 0 \\ s_q & 1/2 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right]$ with $s_q = \mathrm{fl}(\sqrt{3}/2)$ defines the target $q^* = \Pi(A_q)$, the rotation about e_3 by the exact angle $\rho := \mathrm{atan2}(s_q, \frac{1}{2})$; the block is a scaled rotation, which identifies its polar factor analytically. The feasible set is the closed ball of radius ρ about I, on whose boundary q^* lies, and A_q satisfies the input residual contract.

Exact oracle. The objective is $F(R) = \frac{1}{2}d(R, A)^2$ with $A = \mathrm{Exp}_I(ae_3)$ and $a = \rho + \pi/18$. On the ball, $d(R, A) \le a + \rho < \pi$, so F is smooth and geodesically convex with $L_F = 1$, $G \le a + \rho$, and $\|\operatorname{grad} F\| \ge a - \rho > 0$. For every decoded iterate $x = \mathrm{Exp}_I(re_3)$, $0 \le r \le \rho$, geodesic convexity of $r_c := d(I, \cdot)$ gives $\langle \operatorname{grad} r_c(x), \mathrm{Log}_x\, q\rangle \le \rho - r$ for all q in the ball, with equality at q^*, and $\operatorname{grad} F(x) = -(a - r)\operatorname{grad} r_c(x)$, so q^* is a global exact minimizer of the linear oracle at every iterate (at $x = I$ by the Cauchy–Schwarz inequality). The stored entries of A_q are returned at every iteration, and the oracle is exact and feasible for the represented problem. The exact gap of the decoded angle r is $h(r) = \frac{1}{2}[(a - r)^2 - (a - \rho)^2]$; we evaluate this expression in 400-bit arithmetic.

Axis preservation. The iterates remain exactly on the axis because block-structured matrices are closed under the b-bit products, logarithms, Rodrigues exponentials, and the Newton polar iteration of the kernel (zeros arise as $0 \cdot x + 0 \cdot y$ and the $(3, 3)$ entry as $0 + 0 + 1 \cdot 1$); the stored zeros and unit entry are checked after every step, and a violation aborts the run. Starting from $r_0 = 0$, the accepted-step enclosure also keeps every decoded angle on the positive half-axis. To see this, suppose $d_k = \rho - r_k > 0$. The exact next angle is $r_k + T_k d_k$, while the accepted angular error is at most $cT_k^2 d_k^2 < T_k d_k$ because $T_k \le \frac{1}{2}$ and $c\rho/2 < 1$, and feasibility supplies $r_{k+1} \le \rho$ (the axis angle difference is the geodesic distance inside the ball, since $2\rho < \pi$). The exact-oracle argument therefore applies inductively to every accepted represented iterate.

Constants and policy. The run starts at $x_0 = I$ with initial precision $b_0 = 53$ and uses the fixed-N one-bit-increment sufficient policy with $N = 256$, $m = 4$, $T_k = \mathrm{fl}(2/(k + 4))$ correctly rounded from the exact rational, $\theta = \frac{1}{2}$, $c = \underline{\alpha}$, and a represented binary chart radius $\rho'_{\mathrm{bin}} = \mathrm{fl}(\overline{\rho} + 0.1) > \rho$, which is the radius compared in the exact node predicate and in every derived constant, together with the certified lower bound $s = \rho'_{\mathrm{bin}} - \overline{\rho} \le \rho'_{\mathrm{bin}} - \rho$ on the reserve. The feasible set is exactly $\mathcal{C}_\rho$. Its radius $\rho = \mathrm{atan2}(s_q, \frac{1}{2})$ is transcendental ($\tan\rho = 2s_q$ is a nonzero rational) and is never represented; it enters only through a rigorous enclosure $\underline{\rho} \le \rho \le \overline{\rho}$ by adjacent doubles, obtained from 400-bit interval arithmetic on the exact binary inputs (interval width 2^{-399}) followed by outward rounding, with the margins between the interval and the enclosing doubles (between 2^{-56} and 2^{-52}) checked and recorded. The same route encloses $\alpha_\rho = \frac{1}{4}\cot(\rho/2) = (\sqrt{1/4 + s_q^2} + 1/2)/(4s_q)$ and gives $\underline{\alpha}$. Each interval endpoint is used only in the direction required by the bound. The upper endpoint $\overline{\rho}$ enters the diameter bound $D = 2\overline{\rho}$, the chart radius ρ'_{bin}, the reserve bound (hence N), and the domain-reserve test $\overline{\rho} + T_k c_\Omega u_b < \rho'_{\mathrm{bin}}$ of Proposition D.2, made exactly before every attempt. The lower endpoint $\underline{\alpha}$ is used for c and for the modulus. The lower endpoint $\underline{\rho}$ is recorded but needed by no computation, since the feasibility of every accepted iterate is certified by the endpoint test and the exact oracle returns the stored target on the boundary of $\mathcal{C}_\rho$. All constants (h_0, G, L_F, A_s, H_ε, B_ε, $\overline{K}$, b_ε) are computed from these declared bounds before the run. Table 26 lists the stored target entries, the enclosure, the rounding conventions, the precision cap, and the extension wall-clock cap from which the table thresholds are generated. The precision cap is $b_{\mathrm{cap}} = 4096$ bits and the extension cap is 60 min of wall-clock time; the objective parameter is the exact real number $a = \rho + \pi/18$, and the stored values of h_0 and G and the recorded a^+ are upper bounds, each rounded up from a 400-bit value, on the exact initial gap, on $a + \rho$, and on a; and $A_s \approx 26.01$ and $\overline{K} \approx 3022$ are rounded only for presentation, with Table 26 holding the binary values used and the complete execution records retained for the post-review release.

Observed thresholds. Targets $\varepsilon \in \{10^{-3}, 10^{-5}, 10^{-7}\}$ were declared, with 10^{-9} as a resource extension under a wall-clock cap; one trajectory serves every ε because the policy does not depend on ε, and for each ε we record the first k with $h_k \le \varepsilon$, the maximal precision and the number of failed attempts before it, and the theorem's b_ε, $(b_\varepsilon - b_0)_+$, and H_ε (Table 25, Figure 8). Every attempted step was either accepted or followed by a one-bit increase, and no step returned a domain, primitive, or off-axis status. Every accepted enclosure passed the 400-bit consistency check against the exact endpoint of the represented chord, and the 400-bit checks found $r_k \le \rho$ for every decoded iterate. The three declared targets were reached at $k = 31$, 329, and 3310 with maximal precisions 53, 54, and 74 bits against $b_\varepsilon = 66$, 92, and 119 and with 0, 1, and 21 failed attempts against $(b_\varepsilon - b_0)_+ = 13$, 39,

and 66, far inside the horizons H_ε, because the instance converges much faster than the worst-case recurrence that defines A_s. The declared 10^{-9} extension completed within its cap at $k = 33{,}113$ with 94 bits against $b_\varepsilon = 146$ and 41 failed attempts against 93.

Interpretation. Once the 53-bit floor ceases to dominate, the observed maximum precision increases by 20 bits for each two-decade reduction of the target (54, 74, 94 bits at 10^{-5}, 10^{-7}, 10^{-9}). This is about 10 bits per decade on the two finer intervals, against the asymptotic slope $4\log_2 10 \approx 13.3$ of the theorem's threshold. The finite-range behaviour is consistent with logarithmic growth; it is neither a fitted asymptotic rate nor evidence that the sufficient threshold is tight. (On the exact axis recurrence the gap is of order k^{-2} and the quadratic budget of order k^{-6}, which suggests a local sufficient precision of order $3\log_2(1/\varepsilon)$ on this special trajectory.) The chart field is constant on the axis, so Euler truncation vanishes in exact arithmetic. This run therefore tests the precision policy, the represented target, and the finite-rounding contract in a special case satisfying the theorem's oracle and schedule assumptions; it does not test chart integration error, higher-dimensional behaviour, or practical speed, and the primitive bounds of Assumption D.1 remain conditional.

target accuracy ε	10^{-3}	10^{-5}	10^{-7}	10^{-9}
first k with $h_k \le \varepsilon$	31	329	3,310	33,113
horizon H_ε of Theorem D.5	26,012	2,601,145	260,114,461	26,011,446,019
maximal precision used before it [bits]	53	54	74	94
sufficient threshold b_ε [bits]	66	92	119	146
failed attempts before it	0	1	21	41
bound $(b_\varepsilon - b_0)_+$ on failed attempts	13	39	66	93
largest 400-bit error$/E$	7.9e-05	4.1e-01	4.2e-01	4.2e-01
smallest 400-bit slack [rad]	5.6e-03	5.7e-05	5.7e-07	5.7e-09

Table 25: Experiment F applies the fixed-N sufficient policy to the represented exact-oracle instance (N and $m = 4$ as in Theorem D.5). Each column corresponds to one of the three declared targets or to the resource-extension target. The table compares the first ε-accurate iterate with the theorem's horizon H_ε, the maximal precision used before it with the sufficient threshold b_ε, and the failed attempts before it with $(b_\varepsilon - b_0)_+$; it also reports the largest 400-bit error-to-enclosure ratio and the smallest 400-bit slack of an accepted iterate.

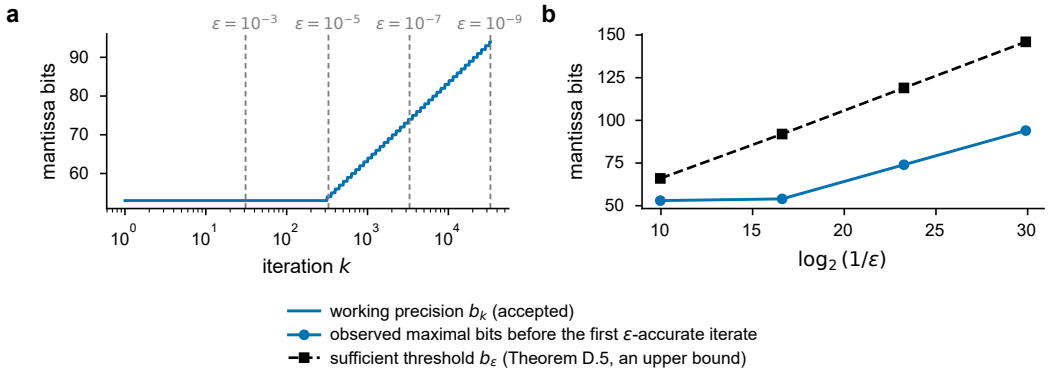

Figure 8: Experiment F. (a) Working precision of the accepted step against the iteration, with the first ε-accurate iterates marked. (b) Maximal precision before the first ε-accurate iterate against $\log_2(1/\varepsilon)$, with the sufficient threshold b_ε of Theorem D.5; b_ε is a sufficient upper bound, not a fitted rate, and all four targets are read from one trajectory.

quantity	role	binary value (hexadecimal float)	decimal
s_q	stored target entry $\mathrm{fl}(\sqrt{3}/2)$	`0x1.bb67ae8584caap-1`	0.8660254037844386
$\underline{\rho}$	lower endpoint of the enclosure of ρ (recorded, unused)	`0x1.0c152382d7365p+0`	1.0471975511965976
$\overline{\rho}$	upper endpoint; enters D, ρ'_{bin}, s, N, the domain-reserve test	`0x1.0c152382d7366p+0`	1.0471975511965979
$\underline{\alpha} = c$	lower endpoint of the enclosure of α_ρ	`0x1.bb67ae8584caap-2`	0.4330127018922193
ρ'_{bin}	chart radius $\mathrm{fl}(\overline{\rho} + 0.1)$	`0x1.25aebd1c70d00p+0`	1.147197551196598
s	reserve lower bound $\rho'_{\mathrm{bin}} - \overline{\rho}$ (exact difference)	`0x1.99999999999a0p-4`	0.10000000000000009
D	diameter bound $2\overline{\rho}$	`0x1.0c152382d7366p+1`	2.0943951023931957
$\overline{D}$	rate envelope $D + c_\Omega u_{53}$	`0x1.0c152382d7376p+1`	2.094395102393203
a^+	upper bound on the exact attractor angle $a = \rho + \pi/18$ (recorded; the objective uses the exact a)	`0x1.38c35418a5bf7p+0`	1.2217304763960308
h_0	upper bound on the exact initial gap	`0x1.76505acbb952ep-1`	0.7310818074881007
G	upper bound on the exact maximum gradient $a + \rho$ (not the sum of the stored bounds)	`0x1.226c3bcdbe7aep+1`	2.2689280275926285
γ	comparison-quality factor $((1-\tau)/(1+\tau))^2$	`0x1.fffffffff8000p-1`	0.999999999996362
$\overline{C}_E$	truncation envelope of Proposition D.2	`0x1.ae8558302e1cfp-1`	0.8408610876859316
A_s	numerator of Lemma A.12, float64 evaluation (see caption)	`0x1.a02ee20553468p+4`	26.011446018983776
$\overline{K}$	rounding envelope of Theorem D.5	`0x1.79b25048036a3p+11`	3021.572299963636

Table 26: The represented inputs and the derived binary constants of Experiment F, as stored and used by the run. The exact parameters are $L_F = 1$, $m = 4$, $u_s = 2^{-53}$, $\theta = \frac{1}{2}$, $\tau = 2^{-40}$, $N = 256$, $b_0 = 53$, and $b_{\mathrm{cap}} = 4096$; the declared targets are $10^{-3}, 10^{-5}, 10^{-7}$, with an extension target of 10^{-9} under a 60-minute wall-clock cap. The radius $\rho = \mathrm{atan2}(s_q, \frac{1}{2})$ is transcendental and never represented; the computations use an enclosing pair of doubles obtained from a 400-bit interval evaluation followed by outward rounding. The 400-bit interval enclosures of ρ and α_ρ have widths 2^{-399} and 2^{-400}, and the margins between each interval and its enclosing doubles are $2^{-53.3}$, $2^{-52.7}$ (ρ) and $2^{-54.6}$, $2^{-55.6}$ (α_ρ). The stored A_s is the float64 evaluation, with each operation rounded to nearest, of $\max\{mh_0, 4(\frac{L_F}{2} + Gc)D^2(1+u_s)^2/(1-2u_s)\}$ with the declared bounds; the exact value of that expression is $26.0114460189837808\ldots$, whose upward rounding is $A_s^+ = \texttt{0x1.a02ee2055346ap+4} = 26.011446018983783$, and the same expression with the exact diameter 2ρ is $26.0114460189837742\ldots < A_s$, so the stored value is a valid numerator for the represented problem; A_s^+ gives the same horizons H_ε and thresholds b_ε.