Brief Report

Reference-Measure Geometry in Quantum Parameter Estimation: When Coordinate Surrogates Optimize the Wrong Objective

Christopher P. Fulton and Lawrence V. Fulton *

Posted Date: 6 May 2026

doi: 10.20944/preprints202605.0359.v1

Keywords: reference measure; Haar measure; information geometry; surrogate optimization; quantum parameter estimation; gate-set tomography; Lie group; Fisher information

 doi:10.20944/preprints202605.0359.v1

Brief Report

Reference-Measure Geometry in Quantum Parameter Estimation: When Coordinate Surrogates Optimize the Wrong Objective

Christopher P. Fulton [1] and Lawrence V. Fulton [2,*]

[1] United States Air Force Test Pilot School, Edwards Air Force Base, CA 93524, USA
[2] Applied Analytics, Boston College Woods College of Advancing Studies, 140 Commonwealth Ave., Chestnut Hill, MA 02467, USA
* Correspondence: fultonl@bc.edu

Abstract

Quantum gate estimation and tomography pipelines routinely combine intrinsically defined likelihoods with priors or regularization terms specified in local Euclidean coordinates. This practice implicitly replaces the Haar reference measure on $SU(2)$ with Lebesgue measure, specifying a different statistical model rather than a reparametrization of the intended one. We show that omitting the associated chart-volume factor alters the optimization objective itself, modifying its gradient field and stationary-point structure. The mismatch persists arbitrarily close to the identity, so that flat-coordinate surrogate objectives can converge to points that are non-stationary for the corresponding Haar-consistent objective even in regimes where local Gaussian approximations are assumed valid. We prove a formal non-equivalence proposition and validate a leading-order Fisher-information correction analytically and numerically. Large-scale multi-start optimization experiments ($N = 11{,}900$ runs) demonstrate that the discrepancy is regime-dependent and most pronounced under moderate-to-strong regularization or limited data. The fix requires a single-line modification to any gradient-based optimizer. These results identify reference-measure selection as an explicit modeling decision with direct consequences for optimization and inference in gate-set tomography, randomized benchmarking, and Bayesian gate estimation on curved parameter manifolds.

Keywords: reference measure; Haar measure; information geometry; surrogate optimization; quantum parameter estimation; gate-set tomography; Lie group; Fisher information

1. Introduction

Modern quantum estimation and control workflows routinely combine two distinct geometric structures. Forward models evolve quantum states or gates on curved manifolds such as $SU(2)$ using intrinsically defined physical laws and likelihoods [1]. Optimization, however, is typically performed in local coordinate charts $\theta = (\theta_1, \ldots, \theta_d) \in \mathbb{R}^d$ using Euclidean objectives built from priors, noise models, or regularization terms specified directly in those coordinates [2,13,14]. This practice implicitly adopts Lebesgue measure on $\mathbb{R}^d$ as the reference measure for inference and optimization.

When probability models are defined intrinsically with respect to Haar (bi-invariant Riemannian) volume on $SU(2)$, their evaluation in local coordinates necessarily includes a chart-volume factor equal to the square root of the metric determinant. This is a modeling choice, not a methodological error: flat-coordinate formulations remain appropriate when priors or noise are intentionally Euclidean, whereas Haar-consistent formulations apply when isotropy is intended with respect to intrinsic group geometry. Omitting this factor specifies a *different* statistical model—not a reparametrization of the same one.

This distinction is operationally important. Flat-coordinate surrogate objectives can converge to points that satisfy standard optimizer tolerances while being non-stationary for the corresponding Haar-

 doi:10.20944/preprints202605.0359.v1

consistent objective (Proposition 1). Coordinate-gradient ascent directions can therefore correspond to descent directions under the intrinsic geometry, and coordinate-space optima can lie on ridges of the intrinsic loss landscape.

Our contribution is methodological rather than geometric. While the differential geometry is classical [3,11,12], we show that reference-measure inconsistency systematically distorts optimization in modern gate-estimation pipelines [2,4,5,16,17], producing stationary-point misalignment, ridge-side convergence, and boundary-amplified failures under realistic regularization and data regimes. We make this mechanism explicit in three steps: (i) a global illustration via stereographic projection (Appendix B); (ii) local analysis in exponential (Lie-algebra) coordinates (Section 2); (iii) large-scale multi-start gate-estimation experiments (Section 3).

A compact notation summary is provided in Appendix A. We distinguish Euclidean (flat-coordinate) quantities with subscript E from geometric (Haar-consistent) quantities with subscript G.

2. Reference-Measure Geometry and Lie-Algebra Analysis

2.1. Reference-Measure Geometry

Let M be a smooth n-dimensional Riemannian manifold with metric tensor g and volume measure $d\mu_g$. If $\psi : V \subset \mathbb{R}^n \to M$ is a smooth coordinate chart, the pullback of the volume measure satisfies

$$\psi^*(d\mu_g) = \sqrt{\det g(y)}\, dy, \tag{1}$$

where dy denotes Lebesgue measure on $\mathbb{R}^n$ (see [11]; cf. Appendix C).

In quantum estimation workflows, likelihood functions are defined intrinsically on $SU(2)$ or $SU(2^n)$, while optimization is carried out in local coordinates using Euclidean objectives [2,13,15,16]. Priors, noise models, and regularization terms are specified directly in coordinate space, implicitly adopting Lebesgue measure as the reference measure. When isotropy or noninformativeness is intended with respect to the intrinsic group geometry, specifying priors in flat coordinates without the chart-volume factor encodes a *different* statistical model. The resulting objective is not a reparametrization of an intrinsic likelihood, but a surrogate defined with respect to a different reference measure—with distinct gradient field and stationary-point structure.

2.2. Exponential Coordinates and Haar Volume on $SU(2)$

In the Pauli basis, any element of $SU(2)$ sufficiently close to the identity admits a unique representation $U(v) = \exp(iv \cdot \sigma/2)$, $v \in \mathbb{R}^3$, $\|v\| < 2\pi$ [1,3]. In this chart, the bi-invariant Riemannian volume (Haar measure) pulls back as

$$d\mu_{\text{Haar}}(U(v)) = J_{\exp}(v)\, dv, \qquad J_{\exp}(v) = \left(\frac{\sin(\|v\|/2)}{\|v\|/2}\right)^2, \tag{2}$$

where $J_{\exp}(0) = 1$ by continuity [3,12].

A Lebesgue-coordinate density $f_E(v)$ (e.g., a zero-mean isotropic Gaussian) induces a probability measure on $SU(2)$ via the exponential map that is *not* Haar unless modified by the chart-volume factor. The Haar-consistent density is

$$f_G(v) = C\, f_E(v)\, J_{\exp}(v), \tag{3}$$

where C is a normalization constant. Although f_E and f_G are both Lebesgue densities on $\mathbb{R}^3$, they induce different measures on $SU(2)$ [3,11].

2.3. Local Expansion and Curvature Consequences

Expanding $\log J_{\exp}$ about $v = 0$ via the Maclaurin series $\sin x = x - x^3/6 + O(x^5)$:

$$-\log J_{\exp}(v) = \frac{1}{12}\|v\|^2 + O(\|v\|^4). \tag{4}$$

 doi:10.20944/preprints202605.0359.v1

Thus $\log f_G(v) = -\frac{1}{2\tau^2}\|v\|^2 - \frac{1}{12}\|v\|^2 + O(\|v\|^4) + \text{const}$, showing that even arbitrarily close to the identity, omitting the chart-volume factor introduces a second-order correction to the log-density. Because the Jacobian contributes at second order, curvature-based quantities—including Fisher information [6,18]—computed from f_E and f_G disagree at $O(\tau^{-2})$ even near the identity.

2.4. Implications for Optimization Objectives

A typical regularized Euclidean objective is

$$L_E(v) = -\log \mathbb{P}(D \mid U(v)) + \tfrac{\lambda}{2}\|v\|^2, \tag{5}$$

corresponding to a Gaussian penalty in flat parameter space. The Haar-consistent objective is

$$L_G(v) = L_E(v) - \log J_{\exp}(v). \tag{6}$$

Differentiating, $\nabla L_G(v) = \nabla L_E(v) - \nabla \log J_{\exp}(v)$. The additional term contributes a nonzero gradient that modifies the stationary-point equations even at $v \approx 0$, since $\log J_{\exp}$ departs from zero at quadratic order and therefore produces a linear-order correction in its gradient. Solutions satisfying $\nabla L_E(v^*) = 0$ therefore need not satisfy $\nabla L_G(v^*) = 0$.

Expanding near the identity: $\nabla L_G(v) = \nabla L_E(v) + \frac{1}{6}v + O(\|v\|^3)$. If $\nabla L_E(v_E^*) = 0$ then $\nabla L_G(v_E^*) = \frac{1}{6}v_E^* + O(\|v_E^*\|^3)$, which is generically nonzero. This local mismatch underlies the global stationary-point misalignment demonstrated in Section 3.

Furthermore, a gradient step $\Delta v = -\eta \nabla L_E(v)$ that decreases L_E can *increase* L_G whenever $v^\top \nabla L_E(v) < -6\|\nabla L_E(v)\|^2$, demonstrating that convergence on L_E does not certify stationarity or even descent for L_G.

Proposition 1 (Stationary-point non-equivalence). *Let $L_E(v) = L_D(v) + R(v)$ be a smooth coordinate-space objective and $L_G(v) = L_E(v) - \log J_{\exp}(v)$. If v^* satisfies $\nabla L_E(v^*) = 0$ and $\nabla \log J_{\exp}(v^*) \neq 0$, then:*

(i) *v^* is not a stationary point of L_G, and $\nabla L_G(v^*) = -\nabla \log J_{\exp}(v^*)$;*
(ii) *$-\nabla L_G(v^*)$ is a strict descent direction for L_G;*
(iii) *$\|\nabla L_G(v^*)\| = \|\nabla \log J_{\exp}(v^*)\|$ is generically $O(\|v^*\|)$ and need not be small even when $\|\nabla L_E(v^*)\|$ meets optimizer tolerances.*

Proof. Since $\nabla L_G(v) = \nabla L_E(v) - \nabla \log J_{\exp}(v)$ and $\nabla L_E(v^*) = 0$, part (i) is immediate. For (ii), the directional derivative of L_G along $-\nabla L_G(v^*)$ equals $-\|\nabla L_G(v^*)\|^2 < 0$. Part (iii) follows from (i). □

3. Materials and Methods: Multi-Start Gate Estimation

We demonstrate that the analytical mismatch produces measurable optimization artifacts in a large-scale single-qubit gate-estimation study mirroring standard practice in tomography and calibration [4,5,13–15,17].

3.1. Experimental Design

Gates are parametrized in axis-angle (exponential) coordinates $v \in \mathbb{R}^3$ via $U(v) = \exp(iv \cdot \sigma/2)$. For a true gate $U_{\text{true}} = U(v_{\text{true}})$, data are generated by sampling random pure input states uniformly on the Bloch sphere, evolving under U_{true}, measuring in the computational basis, and applying independent bit-flip error with probability $p_{\text{flip}} = 0.01$. This mirrors Born-rule likelihood evaluation used in gate-set tomography [2,13] and randomized benchmarking [4], with Gaussian coordinate-space regularization as a stabilizing prior [5].

The negative log-likelihood is $L_D(v) = -\sum_\ell \log \mathbb{P}(m_\ell | \psi_\ell, U(v))$, which is intrinsically defined on $SU(2)$ and coordinate invariant. We compare the two MAP objectives in Equations (5) and (6), with $J_{\exp}(v) = [\sin(\|v\|/2)/(\|v\|/2)]^2$.

 doi:10.20944/preprints202605.0359.v1

We vary the measurement budget N and regularization strength λ across configurations spanning likelihood-dominated, balanced, and prior-sensitive conditions, as well as stress cases with large parameter norms. Each configuration uses 340 random initializations drawn from $v_0 \sim \mathcal{N}(0, 0.09\, I_3)$ with box constraints $v_i \in [-3, 3]$. In total, 11,900 converged multi-start L-BFGS-B runs were analyzed.

3.2. Diagnostics

For each flat-coordinate solution v_E^* satisfying $\|\nabla L_E(v_E^*)\| < \epsilon_{\text{opt}} = 10^{-5}$, we assess: (1) geometric gradient magnitude $\|\nabla L_G(v_E^*)\|$; (2) descent certificate: a small step $\Delta v = -\eta \nabla L_G(v_E^*)$ with $\eta = 10^{-3}$ verified to strictly decrease L_G; (3) intrinsic loss gap $L_G(v_E^*) - L_G(v_G^*)$; (4) gate infidelity $1 - |\langle v_{\text{true}}|U^\dagger(v_E^*)U(v_G^*)|v_{\text{true}}\rangle|^2$.

3.3. Fisher-Information Validation

The flat and geometric formulations differ only in the induced measure on $SU(2)$. With the geometry-consistent local density $f_G(v) \propto \exp(-\|v\|^2/2\tau^2) J_{\exp}(v)$, rejection sampling from the Gaussian envelope f_E (using $J_{\exp}(v) \leq 1$) yields Monte Carlo Fisher estimates via the score outer product

$$\widehat{I} = \frac{1}{N} \sum_{\ell=1}^{N} s(\xi_\ell) s(\xi_\ell)^\top, \quad s(v) = \nabla_v \log f(v). \tag{7}$$

The analytic small-noise prediction is $I_E = \tau^{-2} I_3$, $I_G = (\tau^{-2} + \frac{1}{6}) I_3$, $(I_G - I_E)/I_E = \tau^2/6$. Table 1 confirms this leading-order correction for $\tau \lesssim 0.2$; deviations at larger τ reflect higher-order curvature contributions [6,18].

Table 1. Score-based Fisher-information validation ($N = 2 \times 10^4$ per τ). Both score functions evaluated on samples from f_G. $\widehat{I}_E$ uses the Euclidean score $s_E = -v/\tau^2$; $\widehat{I}_G$ uses the geometric score $s_G = -v/\tau^2 + \nabla \log J_{\exp}(v)$. Predicted ratio $(I_G/I_E)_{\text{pred}} = 1 + \tau^2/6$.

τ	$\widehat{I}_E$	$\widehat{I}_G$	$\widehat{I}_G/\widehat{I}_E$	Predicted
0.05	399.50	399.83	1.0008	1.0004
0.10	100.73	101.06	1.0033	1.0017
0.20	24.73	25.06	1.0134	1.0067
0.50	3.82	4.16	1.0869	1.0417
1.00	0.85	1.19	1.3934	1.1667

4. Results

4.1. Fisher-Information Validation

Table 1 presents score-based Fisher-information validation across concentration parameters spanning two orders of magnitude. Close agreement with the analytic Laplace prediction for $\tau \lesssim 0.2$ validates the correction as a genuine local geometric effect; deviations at $\tau \gtrsim 0.5$ reflect higher-order curvature contributions outside the quadratic regime.

4.2. Multi-Start Optimization Results

Of the 11,900 converged runs, 11,734 (98.6%) were strictly interior to the box constraints; 166 (1.4%) terminated at the boundary.

4.2.1. Interior Ridge-Valley Mismatch

Interior flat-coordinate optima frequently exhibited substantial $\|\nabla L_G(v_E^*)\|$, positive intrinsic loss gaps, and verified descent directions under L_G, confirming they are not local optima of the intrinsic objective. Table 2 reports the ten most extreme interior cases. The most pronounced improvement (config B1 small, $\lambda = 10^{-6}$) reduces infidelity from 0.872 to 0.024 after geometric re-optimization. This is characteristic of stress configurations: large-norm axis-angle parameters correspond to "long-way-round" representations that are geometrically disfavored under a Haar-consistent objective; the

 doi:10.20944/preprints202605.0359.v1

Jacobian term $-\log J_{\text{exp}}$ counteracts the Euclidean regularizer's spurious radial bias. Cases where re-optimization marginally increases infidelity (negative Diff) show differences of at most 0.063, consistent with noise-level variation near a shared basin.

Table 2. Top 10 interior extremes ranked by $\|\nabla L_G(v_E^*)\|$. All cases strictly interior (no box constraints active). Config labels: B = boundary-stress truth (large-norm parameters); C = baseline interior truth; R = randomized interior controls. Regime: *small* = moderate data with stress truth; *lo* = low data; *hi* = high data. Diff = flat infidelity − geometric infidelity.

Config	$\|\nabla L_G\|$	Flat Infid	Geom Infid	Diff	Gap	$\|v_E - v_G\|$	Dist_G
B1 small, $\lambda = 10^{-6}$	34.633	0.872	0.024	0.848	24.434	3.524	0.420
C2, $\lambda = 10^{-6}$	1.280	0.049	0.112	−0.063	2.001	5.643	0.743
C2, $\lambda = 0$	1.280	0.049	0.112	−0.063	2.001	5.643	0.743
C2 lo, $\lambda = 10^{-6}$	1.126	0.075	0.131	−0.056	1.658	5.672	0.823
C2 lo, $\lambda = 10^{-6}$	1.112	0.087	0.131	−0.044	1.632	5.626	0.823
C2, $\lambda = 10^{-6}$	1.007	0.300	0.112	0.188	1.504	4.896	0.743
C2, $\lambda = 0$	1.007	0.300	0.112	0.188	1.504	4.896	0.743
C2 lo, $\lambda = 10^{-6}$	0.926	0.931	0.131	0.799	1.271	4.943	0.823
B3 small, $\lambda = 10^{-6}$	0.925	0.997	0.900	0.097	0.153	1.045	3.706
B3 small, $\lambda = 10^{-6}$	0.920	0.995	0.900	0.095	0.143	0.969	3.706

4.2.2. Boundary Saturation

Boundary-terminated runs exhibit a distinct behavior: both objectives converge to the same feasible solution ($v_E^* = v_G^*$) with zero intrinsic loss gap, because steps that would reduce L_G lie outside the feasible region. These runs confirm that large interior discrepancies are not artifacts of the box constraint.

4.2.3. Configuration-Level Structure

Table 3 summarizes regime dependence. Stress and low-data configurations produce the largest interior violations; high-data settings exhibit small interior gradients. These statistics confirm that the mismatch is structural and regime-dependent, becoming most pronounced when regularization meaningfully shapes the objective relative to the likelihood—precisely the setting described in [5,15,16].

Table 3. Per-configuration mismatch summary (340 multi-start runs each). Columns: boundary-terminated runs and maximum intrinsic gradient magnitude for interior and boundary solutions.

Config	Bdy Runs	max $\|\nabla L_G\|$ (Int.)	max $\|\nabla L_G\|$ (Bdy)
B1 small	3	34.633	0.792
B2 small	14	0.809	17.095
B3 small	18	0.925	13.330
C1 lo, 10^{-6}	8	0.427	22.454
C2 lo, 10^{-6}	6	1.126	3.640
C2, $\lambda = 0$	3	1.280	114.882
C3 hi	3	0.225	461.229
C4 hi	7	0.187	654.953
C5 hi	0	0.330	—

5. Discussion

5.1. Geometry Versus Physical Noise

Flat coordinate-space priors are not intrinsically incorrect. When the scientific intent is to model additive noise in control parameters—uncorrelated Gaussian perturbations introduced by hardware imperfections in axis-angle coordinates—a coordinate-Gaussian prior is physically appropriate [5,16].

 doi:10.20944/preprints202605.0359.v1

Our analysis applies specifically to workflows combining intrinsically defined likelihoods (e.g., Born-rule models on $SU(2)$) with coordinate-space regularization *while intending* isotropy with respect to group geometry. In small-noise, high-data regimes, Euclidean surrogates may provide accurate local approximations near the identity. The significance of reference-measure mismatch is structural rather than pointwise: the Jacobian term modifies the objective, its gradient field, and its stationary-point structure, as demonstrated by strictly interior solutions with moderate parameter norms in Tables 2 and 3.

5.2. Gradient Corruption Versus Curvature Bias

The $1/6$ coefficient in Equation (4) is sometimes interpreted as a small numerical correction to objective values. This understates the impact. The more consequential effect occurs in the gradient field: omitting $-\log J_{\exp}(v)$ removes a term whose gradient $-\nabla \log J_{\exp}(v)$ is generally nonzero away from the identity, radial, and direction-dependent. The stationary-point conditions for L_E and L_G are not equivalent even arbitrarily close to the identity. Flat-coordinate surrogates can therefore converge to parameter values far from stationary under the intrinsic objective, with $\|\nabla L_G(v_E^*)\|$ remaining substantial relative to convergence tolerances.

5.3. Relationship to Natural-Gradient Methods

Natural-gradient methods [7–9] respect the information geometry of the *specified* statistical model. If that model is defined using a coordinate-space reference measure, the same reference-measure mismatch is inherited by the natural gradient. The choice of reference measure (which model is specified) is therefore logically prior to the choice of optimizer metric (how it is optimized). Correcting the reference measure—not changing the optimizer—resolves the mismatch.

5.4. Practical Fix and Scope

The required modification is a one-line change to any gradient-based optimizer: replace $\nabla L_E(y)$ with

$$\nabla L_G(y) = \nabla L_E(y) - \nabla \log J_{\text{chart}}(y), \tag{8}$$

where $J_{\text{chart}}(y) = \sqrt{\det g(y)}$ is the chart-volume factor. This correction is model-based rather than algorithm-specific and applies equally to SGD, Adam, RMSProp, and quasi-Newton methods. If the intrinsic measure is not available in closed form, $\nabla \log J$ can be computed via automatic differentiation of the model map $F : \mathbb{R}^d \to M$.

The mismatch arises whenever priors, regularizers, or surrogate objectives are formulated directly in parameter space on a curved manifold [8,9], including regularized gate-set tomography [2,14], Bayesian gate estimation with Lie-algebra or Euler-angle priors [5,15], and learning-based control objectives penalizing deviations in axis-angle coordinates [16]. Standard Born-rule tomography without regularization is unaffected.

5.5. Limitations

This work is restricted to single-qubit systems. Extending to $SU(2^n)$ requires explicit treatment of the intrinsic metric and volume form, which depend on entanglement structure. For independently parametrized qubits the correction factorizes as $J_{\exp}(v_1, \ldots, v_n) = \prod_j J_{\exp}(v_j)$ and scales linearly with system size. Global gates and entangled states do not admit such factorization and require separate analysis. Our validation also focuses on axis-angle parametrization and quadratic regularization; alternative coordinates or priors may change the magnitude but not the mechanism of the mismatch.

6. Conclusions

Reference-measure selection is an explicit modeling decision with direct consequences for optimization and inference on curved parameter manifolds. When intrinsically defined likelihoods are combined with flat-coordinate regularization while intending Haar isotropy, omitting the chart-volume

 doi:10.20944/preprints202605.0359.v1

factor specifies a different statistical model with a different gradient field and distinct stationary-point structure. We have demonstrated this analytically via Proposition 1 and empirically via 11,900 multi-start gate-estimation experiments. The fix is minimal: add $-\log J_{\text{chart}}(y)$ to the loss and replace ∇L_E with the Jacobian-corrected gradient (Equation (8)). Coordinate-space optima that satisfy flat stationarity conditions can remain non-stationary under the intrinsic objective, reflecting the mismatch between the two gradient fields.

Author Contributions: Conceptualization, C.P.F. and L.V.F.; methodology, C.P.F. and L.V.F.; software, C.P.F. and L.V.F.; validation, C.P.F. and L.V.F.; formal analysis, C.P.F. and L.V.F.; investigation, C.P.F. and L.V.F.; writing—original draft preparation, C.P.F. and L.V.F.; writing—review and editing, C.P.F. and L.V.F.; visualization, C.P.F. and L.V.F. All authors have read and agreed to the published version of the manuscript. All authors contributed equally to the conceptualization, methodology, software development, analysis, visualization, and writing of this manuscript.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Informed Consent Statement: This study did not involve human participants, human subjects data, or animal experimentation and therefore did not require institutional ethical approval or informed consent.

Data Availability Statement: All data are simulated. Code is available here: https://github.com/dustoff06/SU2.

Use of Artificial Intelligence: During the preparation of this work, the authors used OpenAI's ChatGPT to assist with drafting and editing of prose for clarity and readability, and Anthropic's Claude Sonnet 4.6 to assist with programming tasks, including code generation, refinement, and debugging related to data generation, analysis, and visualization. After using these tools, the authors reviewed and edited all content as needed and take full responsibility for the content of the published article.

Conflicts of Interest: The authors declare that they have no known financial or non-financial competing interests that could have influenced the work reported in this paper.

Appendix A. Notation Summary

Table A1. Key notation used throughout.

Symbol	Definition
μ_{Haar}	Haar (bi-invariant) measure on $SU(2)$
μ_{Leb}	Lebesgue measure on $\mathbb{R}^3$
$v \in \mathbb{R}^3$	Exponential (axis-angle) coordinates, $\|v\| < 2\pi$
$U(v) = \exp(iv \cdot \sigma/2)$	Exponential map $\mathbb{R}^3 \to SU(2)$
$J_{\exp}(v)$	Exponential chart-volume factor (Equation (2))
$f_E(v)$	Euclidean (flat-coordinate) density on $\mathbb{R}^3$
$f_G(v)$	Haar-consistent density (Equation (3))
$L_E(v)$	Euclidean objective (Equation (5))
$L_G(v)$	Haar-consistent objective (Equation (6))
∇	Euclidean gradient in coordinates
Subscript E/G	Euclidean/flat vs. Geometric/Haar

Appendix B. Stereographic Projection and Reference-Measure Distortion

This appendix provides a closed-form global illustration of reference-measure distortion under stereographic projection on $S^3 \cong SU(2)$, intended purely as an intuition aid. Stereographic coordinates are not used in practical quantum workflows; all quantitative results in the main text use exponential (axis-angle) coordinates.

 doi:10.20944/preprints202605.0359.v1

$SU(2)$ is diffeomorphic to the 3-sphere $S^3 \subset \mathbb{R}^4$ via $U = \left(\begin{smallmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{smallmatrix}\right)$, $|\alpha|^2 + |\beta|^2 = 1$ [10]. Stereographic projection $\pi : S^3 \setminus \{N\} \to \mathbb{R}^3$ from $N = (0,0,0,1)$ is $\pi(\omega) = (\omega_1, \omega_2, \omega_3)/(1-\omega_4)$. The conformal factor is $\lambda(y) = 2/(1+\|y\|^2)$ [11], giving Jacobian $J_\pi(y) = \lambda(y)^3$ and

$$d\mu_{\text{Haar}} = \left(\frac{2}{1+\|y\|^2}\right)^3 dy. \tag{A1}$$

For Haar-uniform f_{Haar}, the induced Lebesgue density satisfies $f_{Y,\text{Leb}}(y) \propto \|y\|^{-6}$ as $\|y\| \to \infty$—a heavy-tailed coordinate artifact; the intrinsic distribution on S^3 remains uniform. Figure A1 illustrates this mechanism.

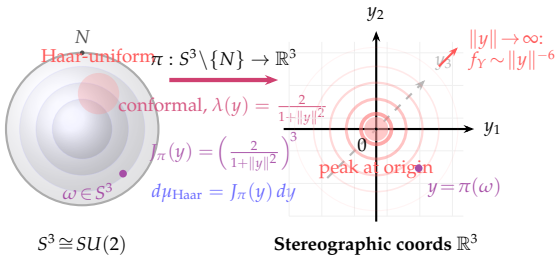

Figure A1. Stereographic projection $\pi : S^3 \setminus \{N\} \to \mathbb{R}^3$ and the induced reference-measure distortion. A Haar-uniform intrinsic distribution becomes nonuniform in coordinates solely through the chart-volume factor $J_\pi(y) = \left(\frac{2}{1+\|y\|^2}\right)^3$. The heavy tail $f_Y \sim \|y\|^{-6}$ is a coordinate artifact; the intrinsic distribution on S^3 remains uniform.

Appendix C. Reference-Measure Pullback

Let (M, g) be an n-dimensional Riemannian manifold with volume measure μ_g, and let $\psi : V \subset \mathbb{R}^n \to M$ be a smooth coordinate map. The pullback satisfies $\psi^*(d\mu_g) = \sqrt{\det g(y)}\, dy$. If X on M has density $p_X(\omega)$ with respect to μ_g, then $Y = \psi^{-1}(X)$ has Lebesgue density $f_Y(y) = p_X(\psi(y))\sqrt{\det g(y)}$ [11]. For a compact Lie group with bi-invariant metric, the Riemannian volume form coincides (up to normalization) with Haar measure [3,12]. Omitting the chart-volume factor replaces Haar volume by Lebesgue measure, specifying a different statistical model.

References

1. Sakurai, J.J.; Napolitano, J. *Modern Quantum Mechanics*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2017. DOI: 10.1017/9781108499996.
2. Nielsen, E.; Blume-Kohout, R.; Gamble, J.K.; Rudinger, K.M. Gate set tomography. *Quantum Sci. Technol.* **2021**, *6*, 044006. DOI: 10.1088/2058-9565/abf6a9.
3. Hall, B.C. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, 2nd ed.; Springer: New York, NY, USA, 2015. DOI: 10.1007/978-3-319-13467-3.
4. Magesan, E.; Gambetta, J.M.; Emerson, J. Efficient measurement of quantum gate error by interleaved randomized benchmarking. *Phys. Rev. Lett.* **2012**, *109*, 080505. DOI: 10.1103/PhysRevLett.109.080505.
5. Hincks, I.; Wallman, J.J.; Ferrie, C.; Granade, C.; Cory, D.G. Bayesian inference for randomized benchmarking protocols. *arXiv* **2018**, arXiv:1802.00401. URL: https://arxiv.org/abs/1802.00401.
6. Amari, S.; Nagaoka, H. *Methods of Information Geometry*; Translations of Mathematical Monographs, Vol. 191; American Mathematical Society: Providence, RI, USA, 2000. DOI: 10.1090/mmono/191.
7. Amari, S. Natural gradient works efficiently in learning. *Neural Comput.* **1998**, *10*, 251–276. DOI: 10.1162/089976698300017746.
8. Girolami, M.; Calderhead, B. Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *J. R. Stat. Soc. B* **2011**, *73*, 123–214. DOI: 10.1111/j.1467-9868.2010.00765.x.
9. Byrne, S.; Girolami, M. Geodesic Monte Carlo on embedded manifolds. *Scand. J. Stat.* **2013**, *40*, 825–845. DOI: 10.1111/sjos.12036.
10. Naber, G.L. *Topology, Geometry and Gauge Fields: Foundations*, 2nd ed.; Springer: New York, NY, USA, 2011. DOI: 10.1007/978-1-4419-7254-5.

 doi:10.20944/preprints202605.0359.v1

11. Lee, J.M. *Introduction to Smooth Manifolds*, 2nd ed.; Graduate Texts in Mathematics, Vol. 218; Springer: New York, NY, USA, 2013. DOI: 10.1007/978-1-4419-9982-5.
12. Helgason, S. *Differential Geometry, Lie Groups, and Symmetric Spaces*; Graduate Studies in Mathematics, Vol. 34; American Mathematical Society: Providence, RI, USA, 2001. DOI: 10.1090/gsm/034.
13. Blume-Kohout, R.; Gamble, J.K.; Nielsen, E.; Rudinger, K.; Mizrahi, J.; Fortier, K.; Maunz, P. Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography. *Nat. Commun.* **2017**, *8*, 14485. DOI: 10.1038/ncomms14485.
14. Nielsen, E.; Rudinger, K.; Proctor, T.; Russo, A.; Young, K.; Blume-Kohout, R. Probing quantum processor performance with pyGSTi. *Quantum Sci. Technol.* **2020**, *5*, 044002. DOI: 10.1088/2058-9565/ab8aa4.
15. Evans, T.J.; Huang, W.; Yoneda, J.; Harper, R.; Tanttu, T.; Chan, K.W.; Hudson, F.E.; Itoh, K.M.; Saraiva, A.; Yang, C.H.; Dzurak, A.S.; Bartlett, S.D. Fast Bayesian tomography of a two-qubit gate set in silicon. *Phys. Rev. Appl.* **2022**, *17*, 024068. DOI: 10.1103/PhysRevApplied.17.024068.
16. Granade, C.E.; Ferrie, C.; Wiebe, N.; Cory, D.G. Robust online Hamiltonian learning. *New J. Phys.* **2012**, *14*, 103013. DOI: 10.1088/1367-2630/14/10/103013.
17. Ferrie, C. Self-guided quantum tomography. *Phys. Rev. Lett.* **2014**, *113*, 190404. DOI: 10.1103/PhysRevLett.113.190404.
18. Paris, M.G.A. Quantum estimation for quantum technology. *Int. J. Quantum Inf.* **2009**, *7*, 125–137. DOI: 10.1142/S0219749909004839.
19. Dehollain, J.P.; Muhonen, J.T.; Blume-Kohout, R.; Rudinger, K.M.; Gamble, J.K.; Nielsen, E.; Laucht, A.; Simmons, S.; Kalra, R.; Dzurak, A.S.; Morello, A. Optimization of a solid-state electron spin qubit using gate set tomography. *New J. Phys.* **2016**, *18*, 103018. DOI: 10.1088/1367-2630/18/10/103018.
20. Wiebe, N.; Granade, C. Efficient Bayesian phase estimation. *Phys. Rev. Lett.* **2016**, *117*, 010503. DOI: 10.1103/PhysRevLett.117.010503.
21. Stace, T.M.; Chen, J.; Li, L.; Perunicic, V.S.; Carvalho, A.R.R.; Hush, M.; Valahu, C.H.; Tan, T.R.; Biercuk, M.J. Optimized Bayesian system identification in quantum devices. *Phys. Rev. Appl.* **2024**, *21*, 014012. DOI: 10.1103/PhysRevApplied.21.014012.