Relative-Entropy Endpoint–Path Gaps as a Diagnostic of RG Coarse-Graining Failure

Zihe Zhao*
Independent Researcher
China

21 December 2025

Abstract

RG-based coarse graining is intrinsically *viewpoint-conditioned*: the observable interface, protocol, regulator/scheme choices, model class, reconstruction map, divergence, and aggregation rule jointly determine what is regarded as "the same" across scales. Fix a viewpoint $\mathsf{V} = (\mathcal{O}, \text{protocol}, R, S, \text{threshold rule}, \mathcal{M}, \pi, D, \text{Agg})$. We then construct two complementary cross-scale objects. *Route I* defines an endpoint divergence Δ_{I} as a relative-entropy-type discrepancy between endpoint predictive objects. *Route II* induces an information metric G from a local second-order expansion and defines the information speed $\Psi = \sqrt{\beta^\top G\, \beta}$, yielding the trajectory length $\Delta_{\mathrm{II}} = \int \Psi \, \mathrm{d}\gamma$.

At the strict level, we establish an endpoint–path bridge for relative entropy along a scale chain under V. By DPI, the endpoint-pair marginal divergence satisfies $\Delta_{\mathrm{I}}^{(\mathrm{pair})} \le \Delta_{\mathrm{path}}$. By the chain-rule decomposition, the strict path object obeys $\Delta_{\mathrm{path}} = \Delta_{\mathrm{I}}^{(\mathrm{pair})} + \Gamma_{\mathrm{path}}$, where $\Gamma_{\mathrm{path}} \ge 0$ and $\Gamma_{\mathrm{path}} = 0$ is equivalent to endpoint sufficiency for distinguishing the compared scale-chain processes under V. At the computable level, we define the geometric proxy gap $\Gamma_{\mathrm{geo}} := \Delta_{\mathrm{II}} - d_I$, where $d_I := \mathrm{dist}_G\big(g(\gamma_0), g(\gamma_1)\big)$ is the endpoint distance under the same induced structure.

Main results. (i) Under Condition G_{geo}, Γ_{geo} is invariant under smooth reparameterizations of coupling coordinates and monotone reparameterizations of the scale parameter, provided G is transformed covariantly (Theorem 5.1). (ii) Under the same Condition G_{geo}, $\Gamma_{\mathrm{geo}} = 0$ holds if and only if the RG trajectory segment is a minimizing geodesic between the endpoints under G (Theorem 5.2). (iii) For a minimal two-dimensional EFT truncation $\mathcal{M}_2 = \{\theta, u\}$ with $\dot\theta = b$ and $\dot u = -pu$, we obtain closed forms for Δ_{II} and show $\Gamma_{\mathrm{geo}}(T) > 0$ for $u_0 \neq 0,\ p > 0,\ T > 0$, together with a small-slope quadratic approximation controlled by the variance of a scale-dependent slope ratio (Theorem 7.1 and Proposition 7.1). We calibrate the construction on a one-loop, threshold-free QED window where the one-dimensional truncation yields $\Gamma_{\mathrm{geo}} = 0$.

Keywords: viewpoint-relative; relative entropy; information geometry; renormalization group; coarse-graining; QED; EFT; truncation gap.

Contents

*ybing2093@gmail.com

1 Introduction

In the standard practice of the renormalization group (RG) and effective field theory (EFT), scale dependence is often accessed through *endpoint objects*[1–4]: one compares renormalized couplings, effective parameters, or predictive/observable outputs at a few fixed momentum points between two energy scales μ_0 and μ_1, and then performs matching, calibration, and constraints accordingly. Endpoint comparison is operationally efficient, but it carries an implicit premise: under the declared interfaces and coarse-graining rules, endpoint information is sufficient to represent the distinguishable structural content of scale evolution within the window.

Main message. Within a fixed, explicitly declared viewpoint V, endpoint-only *comparison* is sufficient for distinguishing the two scale-chain processes being compared if and only if the strict endpoint–path residual Γ_{path} vanishes (equivalently: the endpoints are sufficient statistics for that comparison, i.e. for discriminating $\mathbb{P}_Z$ from $\mathbb{Q}_Z$ as defined in Sec. 5). The geometric proxy Γ_{geo} inherits a sharp zero criterion only on the declared domain of Condition G_{geo} (Sec. 8.2). A nonzero residual identifies intermediate-scale structure that remains distinguishable under the declared interface and cannot be compressed into endpoints at the declared resolution; it therefore provides an operational diagnostic of coarse-graining/truncation breakdown along RG flows.

Positioning. The DPI/chain-rule identity used to derive the strict decomposition is not the contribution by itself. The contribution is to (i) treat the viewpoint V as part of the mathematical specification, so cross-scale statements become reproducible and comparable across implementations, and (ii) construct an invariant, computable proxy gap Γ_{geo} with a sharp zero criterion (minimizing geodesicity) and closed-form calibration examples that quantify when endpoint compression fails.

A segment of scale evolution is, however, itself a coarse-graining process [2, 5]. Even if endpoint objects appear consistent under a given interface, intermediate scales may carry additional distinguishable structure. When such structure cannot be absorbed by endpoint compression at the declared resolution, the "endpoint-only" description and the "full-trajectory" description cease to be equivalent. This paper formalizes the resulting incompressible discrepancy as a *viewpoint-conditioned information gap* (denoted by Γ)[1] and adopts the following operational criterion: under a fixed viewpoint V and a declared resolution clause, if the gap transitions from negligible to non-negligible, then the window is identified as a window where endpoint compression fails.

To make this criterion reproducible, we elevate the viewpoint to a first-order structural object. A viewpoint is specified by

$$\mathsf{V} = (\mathcal{O}, \text{protocol}, R, S, \text{threshold rule}, \mathcal{M}, \pi, D, \text{Agg}),$$

including the observable interface and protocol ($\mathcal{O}$), coarse-graining procedure (R), renormalization scheme (S), threshold rules, model class ($\mathcal{M}$) and reconstruction map (π), divergence functional (D), and aggregation rule (Agg). All reported quantities—the endpoint divergence Δ_{I}[2], the induced-geometry length Δ_{II}[3], and the gap terms $\Gamma_{\text{path}}, \Gamma_{\text{geo}}$—are treated as V-conditioned objects. Accordingly, changes in $\mathcal{O}, R, S, \mathcal{M}, \pi, D$, or Agg represent *changes of specification* (and hence changes of the diagnostic task), rather than alternative expressions of a single viewpoint-independent quantity (Sec. 2).

Notation in this section is anticipatory; formal definitions are given in Secs. 3–5.

1.1 Contributions and main results

The technical ingredients used in this paper (relative entropy, DPI/chain rule, and Fisher information geometry) are standard. The contribution is to assemble them into a single V-locked diagnostic that separates *endpoint-compressible* windows from windows where intermediate-scale structure remains distinguishable at the declared resolution. Concretely, under a fixed viewpoint V, the paper provides:

- **Viewpoint as part of the mathematical specification.** All divergences and lengths are defined as V-conditioned objects; altering $\mathcal{O}, R, S, \mathcal{M}, \pi, D$, or Agg changes the diagnostic specification (Sec. 2).
- **Strict endpoint–path bridge with a nonnegative residual.** Introducing path-process objects and applying DPI and the chain rule yields the exact decomposition $\Delta_{\text{path}} = \Delta_{\mathrm{I}}^{(\text{pair})} + \Gamma_{\text{path}}$ with $\Gamma_{\text{path}} \geq 0$, and an explicit equality condition equivalent to endpoint sufficiency for distinguishing the compared scale flows (Sec. 5 and Appendix A).
- **Computable geometric proxy with invariant meaning and a sharp zero criterion.** On the domain where Route II yields a genuine endpoint distance under a positive-definite metric (Condition G_{geo}), defining $\Gamma_{\text{geo}} := \Delta_{\mathrm{II}} - d_I$ yields a quantity that is invariant under smooth coordinate reparameterizations and monotone reparameterizations of the scale parameter, and satisfies $\Gamma_{\text{geo}} = 0$ if and only if the RG trajectory segment is a minimizing geodesic between the endpoints (Theorems 5.1–5.2).

[1]Here Γ denotes the residual history contribution in the endpoint–path decomposition; its strict definition is Γ_{path} in Sec. 5 (via the chain rule) and Appendix A, and its computable proxy is Γ_{geo} in Sec. 5.

[2]A relative-entropy-type discrepancy between endpoint predictive objects; see Sec. 3.

[3]Obtained by integrating the information speed along an RG trajectory; see Sec. 4.

- **Closed, reproducible calibration examples in standard QED/EFT windows.** In a one-loop, threshold-free QED window, the one-dimensional truncation yields $\Gamma_{\rm geo} = 0$ as a baseline calibration (Sec. 6); extending to the minimal two-dimensional truncation $\mathcal{M}_2 = \{\theta, u\}$ yields $\Gamma_{\rm geo}(T) > 0$ generically for $u_0 \neq 0,\ p > 0,\ T > 0$, with closed forms and a variance-controlled small-slope approximation (Sec. 7 and Appendix B).

Under a fixed V, this paper provides two complementary construction routes. Route I (the endpoint route) rewrites the difference between endpoint-induced predictive objects under the interface $\mathcal{O}$ as a relative-entropy-type distance[6, 7], yielding the endpoint divergence $\Delta_{\rm I}$ (Sec. 3). Route II (the full-trajectory route) localizes the relative entropy in Route I at neighboring scales, induces an information metric G via a second-order expansion[8–11], and combines it with the RG vector field β to define the information speed

$$\Psi = \sqrt{\beta^\top G\, \beta},$$

which is then integrated along the scale variable γ to obtain the path length

$$\Delta_{\rm II} = \int \Psi \, {\rm d}\gamma$$

(Sec. 4).

We then introduce path-process objects at the rigorous level and use the data-processing inequality (DPI) together with the chain-rule decomposition of relative entropy to establish the endpoint–path bridge

$$\Delta_{\rm path} = \Delta_{\rm I}^{(\rm pair)} + \Gamma_{\rm path}, \qquad \Gamma_{\rm path} \geq 0,$$

where the equality condition corresponds to the endpoints being sufficient statistics under the viewpoint (Sec. 5 and Appendix A). At the computable level, we propose the geometric gap proxy

$$\Gamma_{\rm geo} = \Delta_{\rm II} - d_I, \qquad d_I(\mu_0 \to \mu_1) := {\rm dist}_G\big(g(\gamma_0), g(\gamma_1)\big),$$

where d_I is the endpoint geodesic distance under the induced information-geometric structure; within the validity window of the local second-order expansion, the operational approximation $d_I \approx \sqrt{2\Delta_{\rm I}}$ is available, thereby yielding a diagnostic that compares "endpoint compression" against "full-trajectory accumulation" (Sec. 5).

To demonstrate operability, we perform two reproducible checks in one-loop QED within a threshold-free window [12–15]. First, under a one-dimensional truncation and a stable interface, we obtain the baseline case $\Gamma_{\rm geo} = 0$ (Sec. 6); second, within the same window, after extending to a two-dimensional truncation that includes an EFT-irrelevant direction u[1, 3, 4, 16, 17], a minimal nontrivial gap $\Gamma_{\rm geo}(T) > 0$ generically appears, and we provide closed forms and controlled approximations (Sec. 7). We also provide a reproducible protocol: a required checklist of viewpoint elements to report, pseudocode for the operational procedure, and threshold/window segmentation rules (Sec. 8 and Appendices D–E).

The goal of this paper is to provide an operational, viewpoint-conditioned diagnostic for cross-scale endpoint–path gaps together with invariant meaning (reparameterization invariance and a sharp geodesic zero criterion) and closed-form calibration examples. All statements are made conditional on a locked viewpoint V and a declared resolution clause; under that specification, non-negligibility of the gap term provides a structural criterion for endpoint-compression failure and a reproducible method for locating where such gaps emerge.

From the next section onward, all quantities are defined and compared conditional on the locked viewpoint V*; statements outside that specification are outside the scope of the paper.*

The paper is organized as follows: Sec. 2 fixes the normative specification of the viewpoint V; Secs. 3–4 define the Route I endpoint divergence and the Route II information-geometric length, respectively; Sec. 5 establishes the endpoint–path consistency bridge and introduces the gap term; Secs. 6–7 present the baseline and nontrivial one-loop QED examples; Sec. 8 discusses the physical meaning of the gap term, constraint methods, and falsifiable statements; Sec. 9 concludes. Derivation details and closed-form examples are given in Appendices A–F.

Terminology and Structural-Quantity Reference

We adopt a terminology system defined relative to a fixed viewpoint V. The main structural quantities used throughout the paper are summarized in section 1.1.

Symbol	Defined in	Brief meaning
Δ_{I}	Sec. 3	The maximal endpoint discrepancy of predictive objects on the interface set $\mathcal{O}$, defined as the maximum relative entropy (KL divergence) over $\mathcal{O}$; interpreted as the endpoint structural divergence.
Δ_{II}	Sec. 4	The path length accumulated along the RG flow by integrating the information speed Ψ; it measures distinguishability along the entire trajectory.
Ψ	Sec. 4	Information speed, defined from the RG vector field β and the local Fisher metric G by $\Psi = \sqrt{\beta^{\top} G \beta}$.
d_I	Sec. 5	The information-geometric geodesic distance induced by G; it is the shortest-path length between endpoints under the metric G.
Δ_{path}	Sec. 5	The joint KL divergence of path objects (scale chains), serving as the strict definition of the "full-trajectory" divergence.
Γ_{path}	Sec. 5	The strict gap term, representing the incompressible residual information in Δ_{path} beyond the endpoint-pair marginal divergence $\Delta_{\mathrm{I}}^{(\mathrm{pair})}$: $\Delta_{\mathrm{path}} = \Delta_{\mathrm{I}}^{(\mathrm{pair})} + \Gamma_{\mathrm{path}}$.
Γ_{geo}	Sec. 5	A computable geometric proxy gap, defined by $\Gamma_{\mathrm{geo}} = \Delta_{\mathrm{II}} - d_I$; under a second-order expansion regime, it approximates Γ_{path}.
V	Sec. 2	The viewpoint object, comprising interfaces, protocols, model classes, coarse-graining rules, distance functionals, etc.; it provides the reference frame relative to which all structural quantities are defined.

Table: Core structural quantities and their semantics. All quantities are defined relative to a fixed viewpoint V.

2 Normative Setup: Scale Axis, Viewpoint Object, Observable Interfaces, Coarse-Graining, and Model Classes

This section fixes the unique comparability convention (normative setup) used throughout the paper. All divergence quantities, length quantities, inequalities, and gap terms appearing later are *defined and compared* only under the viewpoint configuration locked in this section; if any component is changed (observable interface, coarse-graining, scheme, threshold handling, model class, reconstruction rule, or divergence functional), we treat this as a *change of viewpoint*, which yields another set of objects rather than a different writing of the same object.

2.1 Viewpoint object and relativity statement

We treat "viewpoint" as a first-order structural object. A viewpoint V is defined as the following ordered tuple of data:

$$\mathsf{V} := \big(\mathcal{O},\ \text{observation protocol},\ R,\ S,\ \text{threshold rule},\ \mathcal{M},\ \pi,\ D,\ \mathrm{Agg}\big). \tag{2.1}$$

Here: $\mathcal{O}$ and the observation protocol specify the interfaces for reading/comparing information; R and S specify the theoretical presentation, coarse-graining procedure, and renormalization scheme; the threshold rule specifies the convention for active degrees of freedom as the scale varies; $\mathcal{M}$ and π specify the admissible truncation class and the reconstruction (projection) rule; the divergence functional D specifies how distinguishable information differences are compared *within* a fixed interface; and Agg specifies the fixed cross-interface aggregation rule used when $\mathcal{O}$ is treated as a family of admissible interfaces (e.g. $\sup_{\mathcal{O}}$, an average over $\mathcal{O}$, or another declared aggregation). In the degenerate single-interface case, Agg reduces to the identity and may be absorbed into the choice of D; we nonetheless record Agg explicitly whenever cross-interface comparability is intended.

Remark 2.1 (Relativity statement). The endpoint divergence Δ_{I}, the path divergence/length Δ_{II}, the information speed Ψ, and the gap term (denoted by Γ, with concrete forms $\Gamma_{\text{path}}, \Gamma_{\text{geo}}$) are all viewpoint-relative objects in this paper, and are strictly written as $\Delta_{\mathrm{I}}^{(\mathsf{V})}$, $\Delta_{\mathrm{II}}^{(\mathsf{V})}$, $\Psi^{(\mathsf{V})}$, $\Gamma^{(\mathsf{V})}$. Unless otherwise stated, we work under a fixed viewpoint V and omit the superscripts.

For convenience, we may introduce the family of *admissible viewpoints* $\mathfrak{V}$, which denotes the set of viewpoint configurations satisfying the normative clauses (N1–N6) of this section. The structural conclusions of this paper (the endpoint–path inequality and the gap decomposition) hold in form for any $\mathsf{V} \in \mathfrak{V}$; but the numerical values $\Delta_{\mathrm{I}}^{(\mathsf{V})}, \Delta_{\mathrm{II}}^{(\mathsf{V})}, \Gamma^{(\mathsf{V})}$ vary with the viewpoint.

2.2 Scale parameter and scale stepping

Let $\mu > 0$ denote the renormalization scale (also interpretable as a resolution or a matching scale of an effective theory), and fix a reference scale μ_0. Define the dimensionless scale coordinate

$$\gamma := \ln(\mu/\mu_0). \tag{2.2}$$

To establish the endpoint–path relation, we use both a discrete scale grid and a continuous limit:

- **Discrete grid:** take $\mu_0 < \mu_1 < \cdots < \mu_N$, and write $\gamma_k = \ln(\mu_k/\mu_0)$.
- **Continuous limit:** let $\Delta\gamma_k = \gamma_{k+1} - \gamma_k \to 0$, yielding a continuous parameter $\gamma \in [\gamma_0, \gamma_1]$.

We interpret a single scale step as a Wilsonian scale-advancement operation (realizable as coarse-graining or its inverse, depending on the chosen direction and implementation details): from an effective description at scale μ_k to an effective description at scale μ_{k+1}. To avoid introducing process-mechanism drift in comparisons, when comparing two scale flows we always require that they are generated under the same viewpoint clauses by the same class of scale-stepping maps (see the normative clause N3 of this section). Here "the same class" means: adopting the same set of RG transformations / the same set of integration-elimination and rescaling rules; otherwise, differences will be contaminated by mechanism differences and will no longer correspond to "distinguishable information differences under the same process."

2.3 Observable interface $\mathcal{O}$ and observation protocol

To rewrite "theoretical differences" as "distinguishable information differences," one must fix the observable interface set $\mathcal{O}$ together with its observation protocol. We adopt the following abstract setup:

- $\mathcal{O}$ is a set of observables (or a family of observable functionals). Each $O \in \mathcal{O}$ is bound to a fixed experimental/measurement protocol, including the sampling method, definition of statistics, choice of momentum points/energy-scale points, normalization convention, etc.
- For any theoretical description T_μ at scale μ and any $O \in \mathcal{O}$, "theory + protocol" jointly induce a predictive distribution (or a quantum state), denoted by

$$P^{(O)}_{T_\mu}. \tag{2.3}$$

  This induced mapping is fixed throughout the paper; it must not be replaced in different parts by a "more convenient definition," in order to avoid convention drift.

Interface selection principle (QED context) The calibration target of this paper is a one-loop, threshold-free QED window, and hence $\mathcal{O}$ is taken by default to be gauge-invariant observable interfaces; for example, with defining readings of the "effective charge/effective coupling" as the core (specified by a normalization condition on a two-point function, or by a prescription using some class of scattering cross sections at fixed momentum points). We do not enforce a unique choice of O at the normative level, but require that once a concrete interface implementation is chosen in the main text, it is kept consistent throughout all derivations and numerical examples under that implementation.

2.4 Regulator, renormalization scheme, and threshold rule

Relative entropy / relative-entropy-type divergences in field theory typically come with volume and UV-regulator dependence, so we fix the regulator and scheme explicitly[15]:

- **Regulator R:** one may take a finite volume V and a UV cutoff Λ, or a lattice spacing a, etc. Any regulator implementation is allowed, but it must be kept unchanged within the same derivation chain.
- **Scheme S:** fix a renormalization scheme (e.g. $\overline{\mathrm{MS}}$), and define parameter coordinates under it (e.g. $\alpha(\mu)$). Once fixed, the scheme is not switched within the paper.
- **Threshold rule:** to avoid non-structural non-uniqueness caused by piecewise changes of degrees of freedom when crossing mass thresholds, the main calibration (one-loop QED) part assumes by default a scale window that does not cross thresholds; if threshold-splicing

across windows is discussed, we will explicitly specify consistent rules for the entrance/exit of active degrees of freedom in the relevant subsections.

To make relative-entropy-type quantities comparable, we require that one of the following standardization options be fixed in a concrete implementation:

- **Densitization:** for quantities that grow proportionally to V, take $\frac{1}{V}$ as a "relative-entropy density";
- **Differencing:** compare differences or ratios under the same regulator so that regulator dependence cancels explicitly;
- **Explicit retention of regulator dependence:** report (V, Λ) (or equivalent regulator parameters) as part of the output, avoiding an implicit limit.

2.5 Model class $\mathcal{M}$ and reconstruction map π

Wilsonian scale advancement generally generates an infinite-dimensional tower of operators. To turn "cross-scale differences" into computable objects[1, 2, 16], we introduce a truncated model class and fix its reconstruction rule:

- **Full description:** denote the effective description at scale μ by T_μ (viewable as a Wilsonian effective action or an equivalent object).
- **Model class:** choose a family of finite-dimensional model classes $\mathcal{M}$, whose elements are uniquely parameterized by a finite-dimensional parameter $g \in \mathbb{R}^d$, denoted by $M(g)$. For example:
  - Baseline calibration: $\mathcal{M}_1 = \{\alpha(\mu)\}$ (one-dimensional parameter);
  - EFT extension: $\mathcal{M}_2 = \{\alpha(\mu), u(\mu)\}$ (two-dimensional parameter, where u denotes a dimensionless coefficient along an irrelevant-operator direction).
- **Reconstruction/projection map:** fix a reconstruction map from the full description to the model class,

$$\pi: \; T_\mu \;\mapsto\; \widehat{T}_\mu = \pi(T_\mu) \in \mathcal{M}, \tag{2.4}$$

  and require that π be uniquely specified by a single criterion (numerical approximations are allowed in implementation, but the criterion must not be changed across different parts).

(P) Minimal-information-discrepancy reconstruction (normalized choice) Under fixed $\mathcal{O}$ and a fixed divergence functional D, let $\widehat{T}_\mu = M(\widehat{g}(\mu))$, where

$$\widehat{g}(\mu) := \arg\min_{g \in \mathbb{R}^d} \sup_{O \in \mathcal{O}} D\Big(P^{(O)}_{T_\mu} \,\|\, P^{(O)}_{M(g)}\Big)\,. \tag{2.5}$$

If an aggregated interface is used instead of sup, then sup is replaced by the same aggregation rule, consistently throughout the paper. This definition emphasizes that the model class is not an arbitrarily scale-adjusted "fitting tool," but rather a comparability coordinate system bound to the interface $\mathcal{O}$; changing π is equivalent to changing the coordinate system and the comparison problem itself.

Remark 2.2 (Existence and implementation). The existence/uniqueness of the above $\arg\min$ depends on the concrete implementations of $\mathcal{M}$, $\mathcal{O}$, and D. We fix it here as a normative clause: in practical computations, approximate optimization or computable proxies may be used for implementation, but the criterion must not be switched across different scale segments in pursuit of more "ideal" numerical behavior.

2.6 Divergence functional D and normative clauses (viewpoint stability)

We use relative-entropy-type divergences to characterize distinguishable information differences. The divergence functional $D(\cdot\|\cdot)$ is fixed throughout the paper. Typical choices include the KL divergence (for classical distributions) or its quantum counterpart (for quantum states). In what follows we only use its structural properties to derive the endpoint–path inequality and the gap decomposition, for example:

- Non-negativity: $D(X\|Y) \geq 0$;
- Uniqueness of the zero point: if two objects are indistinguishable under the given interface, then $D = 0$;
- Monotonicity under coarse-graining (data processing inequality, DPI): for an appropriate coarse-graining map Φ (a stochastic map in the classical case, a CPTP map in the quantum case), one has $D(X\|Y) \geq D(\Phi(X)\|\Phi(Y))$.

To ensure that subsequent conclusions do not fail due to viewpoint drift, we organize the viewpoint-stability requirements into six normative clauses:

- **N1 (Interface fixed):** once $\mathcal{O}$ and the observation protocol are chosen, they are not changed throughout the paper; any change is treated as a change of viewpoint.
- **N2 (Regulator and scheme fixed):** the regulator R and scheme S are fixed; the threshold rule is fixed; the chosen handling method among densitization/differencing/explicit retention of regulator dependence is fixed.
- **N3 (Same scale-advancement mechanism):** when comparing two scale flows, they must share the same class of scale-stepping maps (the same scale-advancement mechanism); otherwise, differences are contaminated by mechanism differences and no longer correspond to distinguishable information differences under the same process.
- **N4 (Model class and reconstruction fixed):** the model class $\mathcal{M}$ and reconstruction map π are fixed; one must not switch truncation classes or switch optimization criteria across different scale segments to obtain more "ideal" numerical behavior.
- **N5 (Viewpoint equivalence):** if two viewpoints V, V' are equivalent on the predictive objects defined by the interface $\mathcal{O}$, i.e. for the same set of $O \in \mathcal{O}$ and the same set of scale points (or windows) the induced predictive distributions/states agree (or are equivalent under an explicitly specified invertible reparameterization), then we write $\mathsf{V} \sim \mathsf{V}'$. Coordinate changes are allowed within an equivalence class, but the transformation rule and its effect on $\Delta_{\mathrm{I}}, \Delta_{\mathrm{II}}, \Gamma$ must be given explicitly.

- **N6 (Cross-viewpoint comparison):** if $\mathsf{V} \not\sim \mathsf{V}'$, then $\Delta_{\mathrm{I}}^{(\mathsf{V})}, \Delta_{\mathrm{II}}^{(\mathsf{V})}, \Gamma^{(\mathsf{V})}$ should not be directly compared as the same scalar. Cross-viewpoint comparison must first specify a mapping from V to V' (e.g. interface projection, regulator transformation, or reconstruction-rule transformation), and make explicit the transformation law of the objects under the mapping.

After this section, the endpoint divergence, the path divergence/length, and the gap term appearing in later sections will all be treated as structural objects relative to the fixed viewpoint V; their numerical values and negligibility (e.g. whether $\Gamma^{(\mathsf{V})}$ is negligible under a given resolution) constitute testable statements about whether "intermediate-scale structure can be compressed into endpoints" under that viewpoint.

Optional: spiral coordinate diagram (visualization only) All definitions and inequalities in what follows are carried out under the viewpoint V locked in Sec. 2; as an *optional* geometric representation, the reader may embed several real scalar axes into the plane via a logarithmic spiral coordinate diagram to aid intuition, but this diagram is not a component of V and does not enter any derivation (see Appendix F for the definition).

3 Route I: Definition and Basic Properties of the Endpoint Divergence

Under the viewpoint locked in Sec. 2,

$$\mathsf{V} = (\mathcal{O}, \text{protocol}, R, S, \text{threshold rule}, \mathcal{M}, \pi, D, \text{Agg}),$$

this section gives the definition of the Route I (endpoint-type) divergence quantity $\Delta_{\mathrm{I}}^{(\mathsf{V})}$ and summarizes its basic structural properties. Unless otherwise stated, we omit the superscript (V) in this section.

3.1 Endpoint objects: observable distributions/states at scale μ

Within the viewpoint V, the full effective description at scale μ is denoted by T_μ. Since we allow a truncated model class $\mathcal{M}$ and fix a reconstruction map π, we define the "comparable object" at scale μ as the truncated description

$$\widehat{T}_\mu := \pi(T_\mu) \in \mathcal{M}. \tag{3.1}$$

The meaning of this convention is that we compare scale objects expressed under the same interface $\mathcal{O}$ and within the same model class $\mathcal{M}$, thereby avoiding mixing incomparable operator-tower details with comparable coordinate differences into a single scalar quantification (normative clause N4).

For any $O \in \mathcal{O}$, "truncated description + fixed observation protocol" induces a predictive object:

- If the interface corresponds to a classical observational distribution, we write

$$P_\mu^{(O)} \equiv P_{\widehat{T}_\mu}^{(O)}. \tag{3.2}$$

- If the interface corresponds to a comparison of quantum states (or reduced density matrices), we write

$$\rho_{\mu}^{(O)} \equiv \rho_{\widehat{T}_\mu}^{(O)}. \tag{3.3}$$

In what follows we uniformly denote by $X_\mu^{(O)}$ the "predictive object at scale μ under interface O" (a distribution or a state), and by $D(\cdot\|\cdot)$ the relative-entropy-type divergence fixed in Sec. 2 (e.g. classical KL divergence or quantum relative entropy).

3.2 Definition of Route I: endpoint divergence Δ_{I}

Given two scales μ_a, μ_b (typically $\mu_a = \mu_0, \mu_b = \mu_1$), the Route I endpoint divergence is defined as the "worst-case" distinguishable information difference over the interface set $\mathcal{O}$:

$$\boxed{\Delta_{\mathrm{I}}(\mu_a \to \mu_b) := \sup_{O\in\mathcal{O}} D\Big(X_{\mu_a}^{(O)} \parallel X_{\mu_b}^{(O)}\Big)} \tag{3.4}$$

[4].

If the reader prefers "multi-interface aggregation" rather than sup, one may replace sup by a fixed aggregation operator Agg, for example

$$\Delta_{\mathrm{I}}^{\mathrm{Agg}}(\mu_a \to \mu_b) := \underset{O\in\mathcal{O}}{\mathrm{Agg}}\ D\Big(X_{\mu_a}^{(O)} \| X_{\mu_b}^{(O)}\Big)\,. \tag{3.5}$$

However, the aggregation rule must be treated as part of the viewpoint V and kept unchanged throughout (normative clauses N1 and N4). To avoid introducing additional degrees of freedom in later derivations, the main text defaults to the sup form in equation (3.4).

Interpretation Δ_{I} compares predictive objects only at the endpoint scales and does not include historical information at intermediate scales, hence it is an endpoint-type divergence. In Sec. 5, Δ_{I} will be related structurally to path-type objects $\Delta_{\mathrm{path}}, \Gamma_{\mathrm{path}}$ and their computable proxies $\Delta_{\mathrm{II}}, \Gamma_{\mathrm{geo}}$.

In Route I we have written $\Delta_{\mathrm{I}}(\mu_a \to \mu_b)$ as a divergence between endpoint objects along a single scale-indexed family. In the strict bridge of Sec. 5 the diagnostic task is stated for a comparison between two scale-chain distributions $\mathbb{P}_Z$ and $\mathbb{Q}_Z$; the corresponding endpoint quantity is the endpoint-pair marginal divergence $\Delta_{\mathrm{I}}^{(\mathrm{pair})}(0 \to N)$ defined in (5.6). When $\mathbb{Q}_Z$ is instantiated as an endpoint-only (or endpoint-matched) surrogate for $\mathbb{P}_Z$ under the same locked viewpoint, the Route I endpoint divergence can be read as the specialization of that endpoint-pair divergence to the chosen construction.

3.3 Basic properties: non-negativity, the zero point, and viewpoint dependence

Proposition 3.1 (Non-negativity)**.** *For any pair of scales (μ_a, μ_b), one has*

$$\Delta_{\mathrm{I}}(\mu_a \to \mu_b) \geq 0. \tag{3.6}$$

[4]By default we aggregate differences across interfaces using sup. This symbol may be replaced by Agg in different parts, but the choice must be treated as a fixed component of the viewpoint V and must not be switched across sections (see Sec. 2, clauses N1 and N4).

Proof. By the non-negativity of the relative-entropy-type divergence $D(\cdot\|\cdot)$, for each $O \in \mathcal{O}$ we have $D\left(X_{\mu_a}^{(O)} \| X_{\mu_b}^{(O)}\right) \geq 0$; taking the supremum preserves non-negativity. $\square$

Proposition 3.2 (Viewpoint meaning of the zero point)**.** *If $\Delta_{\mathrm{I}}(\mu_a \to \mu_b) = 0$, then under the viewpoint V, for all $O \in \mathcal{O}$ one has*

$$D\left(X_{\mu_a}^{(O)} \| X_{\mu_b}^{(O)}\right) = 0, \tag{3.7}$$

i.e. the endpoints are indistinguishable in the sense of the interface $\mathcal{O}$.

Remark 3.1. This conclusion is viewpoint-relative: it only states indistinguishability under the fixed interface $\mathcal{O}$, fixed model class $\mathcal{M}$, and fixed divergence D, and does not imply indistinguishability under a finer interface or a richer model class. If D is taken to be the standard KL divergence or quantum relative entropy[6, 18–20], the zero point typically corresponds to (almost-everywhere) identical distributions or the same quantum state; however, the subsequent arguments in this paper only rely on the structural property of "uniqueness of the zero point."

Proposition 3.3 (Viewpoint dependence)**.** *Δ_{I} depends on all components of the viewpoint V. In particular:*

- *Changing the interface set $\mathcal{O}$ or the observation protocol changes the definition of $X_{\mu}^{(O)}$, and hence changes Δ_{I};*
- *Changing the regulator R, scheme S, or threshold rule changes the parameterization and predictive objects, and may change the densitization/differencing treatment;*
- *Changing the model class $\mathcal{M}$ or the reconstruction map π changes the endpoint objects $\widehat{T}_{\mu}$, and hence changes the "comparable part" of the divergence;*
- *Changing the divergence functional D changes the numerical value of the divergence and may change the range of applicability of structural properties such as monotonicity.*

Therefore, numerical comparisons of Δ_{I} are meaningful only within the same viewpoint V; cross-viewpoint comparisons must first specify a viewpoint mapping and transformation law (see Sec. 2, normative clause N6).

3.4 Endpoint divergence and coarse-graining monotonicity (preparation for Sec. 5)

We record here a structural fact that will be used in Sec. 5. Within an admissible viewpoint, if scale advancement can be represented at the interface level as a class of coarse-graining maps (channels) Φ_k acting on predictive objects, i.e. for each $O \in \mathcal{O}$ there exists a map such that

$$X_{\mu_{k+1}}^{(O)} = \Phi_k\left(X_{\mu_k}^{(O)}\right), \tag{3.8}$$

and the divergence functional D satisfies monotonicity under coarse-graining (data processing inequality, DPI)[7, 21, 22],

$$D(X\|Y) \ \geq \ D(\Phi(X)\|\Phi(Y)), \tag{3.9}$$

then under the constraint of the "same scale-advancement mechanism" (Sec. 2, normative clause N3), the endpoint divergence has a compression property consistent with the scale direction.

We do not expand the concrete form here; Sec. 5 will use equation (3.9) to establish the upper-bound relation "the endpoint divergence does not exceed the path divergence Δ_{path},"[5] and will introduce the gap term to quantify the historical distinguishable information lost under endpoint compression.

Summary Under a fixed viewpoint V, Route I rewrites theoretical differences between two endpoint scales as the worst-case interface-readable information difference Δ_{I}. The next section will induce an information metric via a local expansion of Δ_{I}, thereby obtaining Route II's information speed Ψ and path length Δ_{II}, and laying the foundation for the endpoint–path structural relations in Sec. 5.

4 Route I $\to$ Route II: Locally Induced Metric, Information Speed, and Path Length

Under the viewpoint V fixed in Sec. 2, this section localizes the endpoint divergence (Route I) from Sec. 3 in the "small-step / neighboring-scale" limit, thereby inducing a local quadratic form on parameter space (a local version of an information metric, or more generally a local structure). On this basis we define the Route II information speed Ψ and the path length Δ_{II}. Unless otherwise stated, we continue to omit the superscript (V).

4.1 Parameterization and the local comparability window (Condition G_{bridge})

Within the model class $\mathcal{M}$, we parameterize the truncated description $\widehat{T}_\mu = M(g(\mu))$ by a finite-dimensional parameter vector $g = (g^1, \dots, g^d) \in \mathbb{R}^d$. Accordingly, evolution along the scale axis γ can be written as a parameter trajectory

$$g(\gamma): \ [\gamma_0, \gamma_1] \to \mathbb{R}^d. \tag{4.1}$$

To ensure that the relative-entropy-type divergence in Route I admits a second-order expansion and induces a local quadratic form, we work in a *local comparability window*. For later reference, we label this window condition as "Condition G_{bridge}":

Condition G_{bridge} (local second-order window) Within a sufficiently small scale step $\gamma \to \gamma + \mathrm{d}\gamma$, the following hold:

1. Viewpoint elements are locked: the interface and protocol $\mathcal{O}$ are fixed, and the regulator R and scheme S are fixed (normative clauses N1–N2);
2. The threshold rule is not triggered within the step (or a fixed segmentation rule has been adopted so that the "active degrees of freedom" remain consistent within the step);
3. Local regularity: for each $O \in \mathcal{O}$, the predictive object $X^{(O)}(g)$ varies sufficiently smoothly with g within the window, so that the relative-entropy-type divergence admits a second-order approximation under parameter perturbations;

[5] The path divergence Δ_{path} mentioned here is cited in advance as terminology; its strict definition will be given in Sec. 5.

4. Aggregation regularity (if Agg aggregates over multiple interfaces): the second-order remainder in the local expansion is controlled uniformly over the interfaces being aggregated (in particular, when $\mathcal{O}$ is finite or reduced to a finite representative set), so that aggregation preserves the $o(\|\mathrm{d}g\|^2)$ order;

5. Truncation consistency: the model class $\mathcal{M}$ and reconstruction map π remain unchanged within the window (normative clause N4).

Condition G_{bridge} is the technical premise under which the Route I $\to$ Route II bridge holds; its failure typically corresponds to threshold crossing, strongly nonlinear segments, or situations where the model-class truncation is insufficient to stably capture interface-sensitive directions.

4.2 Local expansion of Route I and the induced quadratic form (local form of an information metric)

For any fixed interface $O \in \mathcal{O}$, consider the predictive object $X^{(O)}(g)$ and a parameter perturbation $g \mapsto g + \mathrm{d}g$. Within the window of Condition G_{bridge}, the relative-entropy-type divergence admits the second-order expansion

$$D\Big(X^{(O)}(g) \parallel X^{(O)}(g+\mathrm{d}g)\Big) = \frac{1}{2}\, G^{(O)}_{ij}(g)\, \mathrm{d}g^i\, \mathrm{d}g^j \; + \; o\big(\|\mathrm{d}g\|^2\big)\,, \tag{4.2}$$

where the coefficient tensor $G^{(O)}_{ij}(g)$ is the local quadratic form induced by the second-order term of the relative-entropy expansion under that interface (the classical case corresponds to the Fisher information matrix; the quantum case corresponds to a local information quadratic form induced by relative entropy) [7–10].

Worst-case induced local quadratic form via interface aggregation Since Δ_{I} in Sec. 3 aggregates via $\sup_{O\in\mathcal{O}}$, we correspondingly define, under the viewpoint, the local quadratic form

$$Q_g(\mathrm{d}g) := \sup_{O\in\mathcal{O}} \; \mathrm{d}g^{\top} G^{(O)}(g)\, \mathrm{d}g \; = \; \sup_{O\in\mathcal{O}} \; G^{(O)}_{ij}(g)\, \mathrm{d}g^i\, \mathrm{d}g^j. \tag{4.3}$$

It characterizes the "worst-case local sensitivity" over the admissible interface set $\mathcal{O}$. If the main text adopts an aggregation operator Agg instead of sup, then the sup in equation (4.3) should be consistently replaced by the same aggregation rule; this choice is part of the viewpoint V (normative clauses N1 and N4).

Remark 4.1 (On sup and smoothness). Taking sup over an interface family may cause $Q_g(\cdot)$ to exhibit a piecewise structure in $\mathrm{d}g$, and hence it need not be a smooth quadratic form in general. At the v0 level we treat Q_g as a local structure induced by interface aggregation: in concrete applications one may (i) choose finitely many representative interfaces and take the maximum, or (ii) first fix a set of sufficient-statistic interfaces so that Q_g becomes piecewise smooth. These implementations do not change the definition chain of Route II; their differences belong to the viewpoint-implementation level and should be reported for reproducibility (normative clauses N1 and N4).

From equations (4.2) and (4.3), within the window of Condition G_{bridge}, the small-step

Route I divergence satisfies

$$\Delta_{\mathrm{I}}(\gamma \to \gamma + \mathrm{d}\gamma) = \sup_{O\in\mathcal{O}} D\Big(X^{(O)}(\gamma) \| X^{(O)}(\gamma+\mathrm{d}\gamma)\Big) = \frac{1}{2}\, Q_{g(\gamma)}(\mathrm{d}g) + \ o\big(\|\mathrm{d}g\|^2\big)\,, \tag{4.4}$$

where $\mathrm{d}g = g(\gamma + \mathrm{d}\gamma) - g(\gamma)$.

4.3 RG flow and the information speed Ψ

Under a fixed scheme and threshold rule, the parameter trajectory satisfies an RG-type evolution equation

$$\frac{\mathrm{d}g^i}{\mathrm{d}\gamma} = \beta^i(g), \tag{4.5}$$

where $\beta(g)$ is the beta function (or more generally the scale-flow vector field) in the coordinates of the model class. Thus within a small step one has $\mathrm{d}g = \beta(g(\gamma))\,\mathrm{d}\gamma$.

Information line element and speed Using equation (4.3), define the local line element induced by Q_g (more precisely: the local "line element squared")

$$\mathrm{d}s^2 := Q_{g(\gamma)}(\mathrm{d}g), \qquad \mathrm{d}s := \sqrt{Q_{g(\gamma)}(\mathrm{d}g)}. \tag{4.6}$$

Substituting $\mathrm{d}g = \beta\,\mathrm{d}\gamma$ into equation (4.6) yields the information speed

$$\boxed{\Psi(\gamma) := \frac{\mathrm{d}s}{\mathrm{d}\gamma} = \sqrt{Q_{g(\gamma)}\big(\beta(g(\gamma))\big)} = \sqrt{\sup_{O\in\mathcal{O}}\ \beta^i(g(\gamma))\, G^{(O)}_{ij}(g(\gamma))\, \beta^j(g(\gamma))}.} \tag{4.7}$$

Remark 4.2 (Small-step relation to Δ_{I}). From equations (4.4) and (4.6), within the window of Condition G_{bridge},

$$\Delta_{\mathrm{I}}(\gamma \to \gamma + \mathrm{d}\gamma) = \tfrac{1}{2}\,\mathrm{d}s^2 + o(\mathrm{d}\gamma^2),$$

i.e. the endpoint divergence provides the second-order coefficient of the local line element squared in the small-step limit; Route II then defines the path length by a first-order accumulation of $\mathrm{d}s$.

4.4 Route II path length Δ_{II}, distance dist_G, and calculus relations

In the continuous-scale limit, define the Route II path length (full-trajectory divergence accumulation) along the RG trajectory by

$$\boxed{\Delta_{\mathrm{II}}(\gamma_0 \to \gamma_1) := \int_{\gamma_0}^{\gamma_1} \Psi(\gamma)\,\mathrm{d}\gamma = \int_{\gamma_0}^{\gamma_1} \sqrt{Q_{g(\gamma)}\big(\beta(g(\gamma))\big)}\ \mathrm{d}\gamma.} \tag{4.8}$$

Equation (4.8) immediately yields the "calculus relation" of Route II:

$$\frac{\mathrm{d}}{\mathrm{d}\gamma}\Delta_{\mathrm{II}}(\gamma_0 \to \gamma) = \Psi(\gamma), \tag{4.9}$$

where $\Delta_{\mathrm{II}}(\gamma_0 \to \gamma)$ is viewed as a cumulative length function with upper limit γ.

Endpoint distance induced by Q_g (for later use of d_I) To compare endpoints and full trajectories under the same induced structure, we also introduce the endpoint distance dist_G induced by Q_g: for any piecewise C^1 curve $c : [0, 1] \to \mathbb{R}^d$, define its length functional

$$L_G(c) := \int_0^1 \sqrt{Q_{c(t)}\big(\dot{c}(t)\big)}\,\mathrm{d}t. \tag{4.10}$$

For two endpoints g_a, g_b, define the distance

$$\boxed{\mathrm{dist}_G(g_a, g_b) := \inf_{c(0)=g_a,\, c(1)=g_b} L_G(c).} \tag{4.11}$$

When $Q_g(\mathrm{d}g)$ is given in $\mathrm{d}g$ by a smooth positive-definite bilinear form, dist_G reduces to the geodesic distance under the corresponding Riemannian metric; in general dist_G is defined as the endpoint distance induced by Q_g [10, 11].

In Sec. 5 we will use

$$d_I := \mathrm{dist}_G\big(g(\gamma_0), g(\gamma_1)\big)$$

as an endpoint-equivalent length, and within the local window of Condition G_{bridge} provide an operational approximation relating it to $\sqrt{2\Delta_{\mathrm{I}}}$.

Discrete-scale approximation If a discrete scale grid $\{\gamma_k\}$ is used, one may use the Riemann-sum approximation

$$\Delta_{\mathrm{II}}(\gamma_0 \to \gamma_N) \approx \sum_{k=0}^{N-1} \Psi(\gamma_k)\,\Delta\gamma_k, \qquad \Psi(\gamma_k) = \sqrt{Q_{g(\gamma_k)}\big(\beta(g(\gamma_k))\big)}. \tag{4.12}$$

4.5 Scope and failure modes (preparation for interpreting the gap term Γ)

The bridge in this section relies on Condition G_{bridge} (local second-order window) and viewpoint locking (Sec. 2, clauses N1–N4), and hence is mainly applicable to:

- threshold-free windows, or windows where threshold handling has been fixed and segmentation is explicit;
- truncations whose model class is sufficient to stably capture the main sensitivity directions of the interface $\mathcal{O}$;
- local variations that do not invalidate the second-order approximation.

In contrast, when the following situations occur, they should be treated as outside the window of Condition G_{bridge} or reorganized according to segmentation rules:

1. the model class is too narrow, so that directions generated at intermediate scales are strongly projected out, making the local structure unstable within the window;
2. the effective sensitivity directions of the interface family $\mathcal{O}$ switch across different scale segments, making Q_g induced by sup strongly piecewise;
3. threshold crossing changes the content of active degrees of freedom, so that β and the local quadratic form must be defined piecewise under the prescribed threshold rule.

These situations will be incorporated in Sec. 5 into the structural meaning of the gap term (Γ_{path} and its computable proxy Γ_{geo}): the gap term quantifies the difference between "endpoint compression" and "full-trajectory accumulation," i.e. the distinguishable information in intermediate-scale history discarded by an endpoint description.

Summary This section induces a local quadratic form Q_g from a local relative-entropy expansion of Route I, then defines the information speed Ψ via the RG vector field β, and obtains the Route II path length Δ_{II} by integrating along the scale axis. We also introduce the endpoint distance dist_G induced by Q_g, providing a unified "length scale" for later endpoint–path matching and for defining the geometric gap Γ_{geo}. The next section will build on this to give the endpoint–path structural inequality and gap decomposition, and interpret the testable meaning of the gap term.

5 Consistency Bridge: Endpoint–Path Inequalities and the Gap Term Γ

Under a fixed viewpoint V, this section places the endpoint divergence Δ_{I} from Sec. 3 and the path length Δ_{II} from Sec. 4 within a single structural framework, derives controlled endpoint–path inequalities (rather than presupposing equality), and introduces gap terms (collectively denoted by Γ) to quantify the cross-scale historical differences compressed away by an "endpoint-only" description relative to a "full scale-evolution" description. Unless otherwise stated, we continue to omit the superscript (V).

5.1 Path objects: scale chains and path distributions (or process states)

To capture the "full trajectory" rather than only endpoints, we adopt a discrete scale grid

$$\mu_0 < \mu_1 < \cdots < \mu_N,$$

and view scale evolution as a path process. Under a fixed viewpoint V, for each scale μ_k and for the comparable object induced by a chosen interface implementation (including observation protocol, regulator/scheme, threshold rule, model class, and reconstruction), Sec. 3 yields a predictive object at that scale point (a distribution in the classical case, a state in the quantum case). For the strict path construction, let X_k denote the corresponding protocol readout random variable (or its sufficient statistic) at scale μ_k, whose law is induced by that predictive object.

To place the endpoint divergence and the full-trajectory divergence into the same relative-entropy framework, we introduce the path variable

$$Z := (X_0, X_1, \ldots, X_N). \tag{5.1}$$

Note: this section gives the formal definition of the path divergence Δ_{path} anticipated at the end of Sec. 3, and establishes its structural correspondence with the endpoint quantity Δ_{I}.

For two scale flows being compared (e.g. two sets of initial conditions, two RG trajectories, or two reconstruction sequences), under the same scale-advancement mechanism (Sec. 2, N3) and the same interface implementation they induce two path-process objects

$$\mathbb{P}_Z, \qquad \mathbb{Q}_Z. \tag{5.2}$$

In the diagnostic use case emphasized in this paper, $\mathbb{P}_Z$ may represent the scale-chain induced by a less-truncated (or higher-dimensional) description, while $\mathbb{Q}_Z$ represents an endpoint-matched but more aggressively coarse-grained/truncated surrogate under the same viewpoint. We do not require writing all details of $\mathbb{P}_Z, \mathbb{Q}_Z$ explicitly; we only require that, under the fixed viewpoint, the joint object over the scale chain can be treated as a joint distribution (classical case) or a process state (quantum case) for which relative entropy is defined. For multiple interfaces, there are two equivalent implementations: either take X_k to be the joint object of the interface-family readouts (convenient for a strict "mother object"), or construct a separate Z for each representative interface and report the interface choice for reproducibility (see N1, N4).

5.2 The strict mother object for Route II: path divergence Δ_{path}

On the above path objects, define the strict version of the "full-trajectory divergence" by

$$\boxed{\Delta_{\text{path}}(0 \to N) := D(\mathbb{P}_Z \,\|\, \mathbb{Q}_Z)\,.} \tag{5.3}$$

It measures the distinguishable information difference, under the viewpoint V, between two scale flows viewed as "full process objects." This quantity is the strict mother object for the endpoint–path structural relations; its subsequent decomposition follows directly from the chain rule for relative entropy [6, 7].

Remark 5.1 (Strict object and computable proxy). The Δ_{II} in Sec. 4 is an information-geometric length obtained by integrating along the scale axis. Within Condition G_{bridge} (local second-order window), it can be used as a computable proxy of Δ_{path} insofar as the scale-chain process admits a stepwise localization compatible with the second-order expansion (so that local KL elements can be accumulated into an effective length). In general, this paper does *not* claim an identity $\Delta_{\text{II}} = \Delta_{\text{path}}$; the strict semantics remain anchored by the chain-rule object Γ_{path} in equation (5.10), while Δ_{II} and Γ_{geo} serve as operational proxies for reproducible diagnosis. This section provides both the strict endpoint–path inequality (based on Δ_{path}) and the operational matching with Δ_{II} (for practical calibration).

5.3 Endpoint divergence as a coarse-graining of path divergence: an endpoint upper bound (DPI)

Let the endpoint map (coarse-graining) Π retain only the two ends of the path variable:

$$\Pi(Z) = (X_0, X_N). \tag{5.4}$$

Then the endpoint marginals satisfy

$$\Pi(\mathbb{P}_Z) = \mathbb{P}_{X_0,X_N}, \qquad \Pi(\mathbb{Q}_Z) = \mathbb{Q}_{X_0,X_N}.$$

If the divergence D satisfies monotonicity under coarse-graining (data processing inequality, DPI), then

$$D(\mathbb{P}_Z \| \mathbb{Q}_Z) \;\geq\; D(\mathbb{P}_{X_0,X_N} \| \mathbb{Q}_{X_0,X_N}). \tag{5.5}$$

Alignment with the Route I endpoint divergence (implementation level) Under a fixed interface implementation/aggregation rule, the right-hand side can be identified as the endpoint-pair marginal divergence induced by the strict path object. To avoid ambiguity with the Route I endpoint divergence Δ_{I} of Sec. 3, define

$$\boxed{\Delta_{\mathrm{I}}^{(\mathrm{pair})}(0 \to N) \;:=\; D(\mathbb{P}_{X_0,X_N} \| \mathbb{Q}_{X_0,X_N}).} \tag{5.6}$$

In the single-interface instantiation (or when the viewpoint takes the endpoint object to be the endpoint pair), $\Delta_{\mathrm{I}}^{(\mathrm{pair})}$ coincides with the Route I endpoint divergence Δ_{I}. If Route I aggregates over multiple interfaces (e.g. via a sup over $\mathcal{O}$), then Δ_{I} is obtained by applying the same aggregation rule to the corresponding endpoint-pair divergences; in that case, equation (5.5) applies to each interface projection separately, yielding the corresponding endpoint upper bounds. This belongs to the viewpoint-implementation level and should be reported for reproducibility (see N1, N4).

Therefore, from equations (5.5) and (5.6) we obtain the endpoint–path upper-bound relation:

$$\boxed{\Delta_{\mathrm{I}}^{(\mathrm{pair})}(0 \to N) \;\leq\; \Delta_{\mathrm{path}}(0 \to N).} \tag{5.7}$$

Equation (5.7) expresses the basic endpoint–path inequality: retaining only endpoints does not increase distinguishable information difference; the endpoint-pair marginal divergence is at most the full-trajectory divergence [7, 21, 22].

5.4 Gap decomposition: $\Delta_{\mathrm{path}} = \Delta_{\mathrm{I}}^{(\mathrm{pair})} + \Gamma_{\mathrm{path}}$

Decompose the path variable into endpoints

$$E := (X_0, X_N)$$

and history

$$H := (X_1, \ldots, X_{N-1}).$$

The chain-rule decomposition of relative entropy gives

$$D(\mathbb{P}_{E,H} \| \mathbb{Q}_{E,H}) = D(\mathbb{P}_E \| \mathbb{Q}_E) + \mathbb{E}\big[\, D(\mathbb{P}_{H|E} \| \mathbb{Q}_{H|E}) \,\big]_{\mathbb{P}_E}. \tag{5.8}$$

The endpoint term is exactly the endpoint-pair marginal divergence in equation (5.6),

$$\Delta_{\mathrm{I}}^{(\mathrm{pair})}(0 \to N) := D(\mathbb{P}_E \| \mathbb{Q}_E).$$

Accordingly, define the path gap term

$$\boxed{\Gamma_{\mathrm{path}}(0 \to N) := \mathbb{E}\big[\, D(\mathbb{P}_{H|E} \| \mathbb{Q}_{H|E}) \,\big]_{\mathbb{P}_E} \;\geq\; 0,} \tag{5.9}$$

and obtain the exact identity

$$\boxed{\Delta_{\text{path}}(0 \to N) = \Delta_{\text{I}}^{(\text{pair})}(0 \to N) + \Gamma_{\text{path}}(0 \to N).} \tag{5.10}$$

Equation (5.10) gives the structural core of the endpoint–path relation: the difference between endpoint and path divergences is not an "error," but a strictly nonnegative structural term Γ_{path} with a clear conditional meaning; it measures the "historical distinguishable information" compressed away by the coarse-graining that retains only the endpoints $E = (X_0, X_N)$.

Remark 5.2 (Notation (the Ψ–Δ alignment)). We use Γ as a collective name for "endpoint–trajectory difference terms," so that the main endpoint/path notations $\Delta_{\text{I}}, \Delta_{\text{II}}$ remain clear. At the strict level the endpoint-pair marginal divergence $\Delta_{\text{I}}^{(\text{pair})}$ appears together with the chain-rule residual Γ_{path} (see equation (5.10)); at the computable level we obtain the geometric proxy $\Gamma_{\text{geo}} = \Delta_{\text{II}} - d_I$. In the author's Ψ–Δ semantics, these difference terms are not additional independent primitives, but can be regarded as components (decomposition residuals) of the same underlying gap Δ under this decomposition. Thus Γ is a notational convention introduced to highlight the chain-rule decomposition, nonnegativity, and equality condition; its value and its negligibility under a given resolution both depend on the fixed-viewpoint clauses of V.

Equality condition (structural statement) From equation (5.9), $\Gamma_{\text{path}} = 0$ holds if and only if $D(\mathbb{P}_{H|E} \| \mathbb{Q}_{H|E}) = 0$ holds $\mathbb{P}_E$-almost surely, i.e. the two path processes have indistinguishable conditional history distributions given the endpoints $E = (X_0, X_N)$. Equivalently, under the viewpoint V, the endpoints E are sufficient statistics for distinguishing the two scale flows: the history provides no additional distinguishing information [6, 22].

5.5 Matching with Route II: the geometric gap Γ_{geo}

The Γ_{path} defined above is a strict object, but in generic field-theoretic problems it is often difficult to compute directly. Sec. 4 provides a computable proxy for Route II: the information speed $\Psi(\gamma)$ and the path length Δ_{II}.

To place the endpoint divergence Δ_{I} and the path length Δ_{II} on the same "length scale," we adopt the endpoint information distance induced by Route II (Sec. 4, equation (4.11)) to define an endpoint-equivalent length:

$$\boxed{d_I(\mu_0 \to \mu_1) := \operatorname{dist}_G\big(g(\gamma_0),\, g(\gamma_1)\big).} \tag{5.11}$$

Here dist_G is the endpoint distance induced by Q_g; when Q_g is given by a smooth positive-definite bilinear form, it reduces to the corresponding geodesic distance.

Local equivalence (relation to $\sqrt{2\Delta_{\text{I}}}$) Within the local second-order window of Condition G_{bridge}, if the divergence satisfies the local relation

$$D \approx \frac{1}{2}\,\mathrm{d}s^2,$$

then for sufficiently close endpoints one may use the operational approximation

$$d_I(\mu_0 \to \mu_1) \;\approx\; \sqrt{2\,\Delta_{\text{I}}(\mu_0 \to \mu_1)}. \tag{5.12}$$

When an explicit computable expression is needed, we use equation (5.12) as an approximation within the local window; in general we use dist_G in equation (5.11) to avoid misreading the local equivalence as a global identity.

Accordingly, define the geometric gap

$$\boxed{\Gamma_{\mathrm{geo}}(\mu_0 \to \mu_1) := \Delta_{\mathrm{II}}(\mu_0 \to \mu_1) - d_I(\mu_0 \to \mu_1).} \tag{5.13}$$

Interpretation Δ_{II} is the accumulated length along the scale axis (full trajectory), while d_I is the endpoint distance under the same induced structure (endpoints); thus Γ_{geo} measures the "extra length of the trajectory relative to an endpoint-equivalent connection," and is used to diagnose whether intermediate-scale structure is compressible into endpoints at a given resolution.

Remark 5.3 (On nonnegativity and proxy character). If within the window of Condition G_{bridge}, Δ_{II} is defined as the curve length induced by Q_g, and d_I is taken as the corresponding endpoint distance dist_G, then by "curve length is no smaller than endpoint distance" one has

$$\Gamma_{\mathrm{geo}} \geq 0.$$

Outside the window of Condition G_{bridge}, or when using the local approximation equation (5.12), Γ_{geo} is treated in this paper as a computable diagnostic proxy: its vanishing typically corresponds to a one-dimensional sufficient-statistic case or a geodesic case, while a significant deviation from zero indicates that endpoint compression discards non-negligible intermediate-scale structure. The strict structure remains anchored by the chain-rule object Γ_{path} in equation (5.10).

Theorem 5.1 (Reparameterization invariance of Γ_{geo})**.** *Fix a viewpoint* $\vee$ *and assume Condition* G_{geo} *holds (Sec. 8.2), namely: Route II yields a* C^1 *trajectory* $\gamma \mapsto g(\gamma)$ *on a parameter manifold equipped with a single positive-definite metric field* $G(g)$*, so that* Δ_{II} *is the* G*-arc length and* $d_I = \mathrm{dist}_G$ *is the induced endpoint distance. Then the geometric proxy gap* $\Gamma_{\mathrm{geo}} := \Delta_{\mathrm{II}} - d_I$ *is invariant under:*

1. *any smooth change of coupling coordinates* $\tilde{g} = \phi(g)$ *with the covariant metric transformation* $\tilde{G}(\tilde{g}) = \left(J^{-1}\right)^{\top} G(g)\left(J^{-1}\right)$*, where* $J = \partial\tilde{g}/\partial g$*;*
2. *any monotone reparameterization of the scale variable* $\tilde{\gamma} = \tilde{\gamma}(\gamma)$*.*

Proof. Both Δ_{II} (arc length) and $d_I = \mathrm{dist}_G$ (the induced geodesic distance) are invariants of the underlying Riemannian structure $(\mathcal{M}, G)$, independent of coordinate charts and curve parameterizations. Therefore their difference Γ_{geo} is invariant as well. □

Theorem 5.2 (Zero-gap characterization)**.** *Under the hypotheses of Theorem 5.1, one has* $\Gamma_{\mathrm{geo}} \geq 0$*. Moreover,* $\Gamma_{\mathrm{geo}} = 0$ *holds if and only if the trajectory segment* $g([\gamma_0, \gamma_1])$ *realizes the endpoint distance* d_I*, i.e. it is a minimizing geodesic between the endpoints (up to monotone reparameterization).*

Proof. For any rectifiable curve, its length is at least the induced distance between its endpoints; this gives $\Gamma_{\mathrm{geo}} \geq 0$. Equality holds exactly when the curve is distance-realizing, i.e. when it is a minimizing geodesic segment (up to reparameterization). □

5.6 Testable statements: criteria for negligible/non-negligible Γ

We summarize the meaning of the gap term as testable statements (all stated under a fixed viewpoint V and a given resolution):

- If under the given resolution and model-class truncation the gap term can be treated as negligible (e.g. $\Gamma_{\text{path}} \approx 0$ or its proxy $\Gamma_{\text{geo}} \approx 0$), then the endpoint divergence is sufficient to represent the full-trajectory divergence: intermediate-scale history provides no additional distinguishable information, and the system is endpoint-compressible under the viewpoint.

- If the gap term is significantly greater than zero (e.g. $\Gamma_{\text{path}} > 0$ or $\Gamma_{\text{geo}} > 0$ with a non-negligible deviation), then the endpoint divergence systematically underestimates the full-trajectory divergence: intermediate scales contribute non-negligible structural content. Common sources include (i) insufficient model-class truncation (directions are projected out), (ii) multi-direction interface sensitivity that cannot be stably carried by the given parameter dimension, or (iii) threshold crossing / piecewise changes of active degrees of freedom that modify the process structure in segments.

In Sec. 6 (one-loop, threshold-free QED window), we will exhibit a baseline case with $\Gamma_{\text{geo}} \approx 0$; in Sec. 7 (EFT-extended two-dimensional truncation), we will present a nontrivial example where $\Gamma_{\text{geo}} > 0$ typically appears after an irrelevant direction enters, thereby advancing the endpoint–path gap from a concept to a reproducible structural output.

Summary This section provides two layers of endpoint–path relations: at the strict level, $\Delta_{\text{I}}^{(\text{pair})} \leq \Delta_{\text{path}}$ (DPI), and $\Delta_{\text{path}} = \Delta_{\text{I}}^{(\text{pair})} + \Gamma_{\text{path}}$ (chain-rule decomposition, $\Gamma_{\text{path}} \geq 0$); at the computable level, when the local bridge under Condition G_{bridge} is valid, the endpoint quantity can be converted into an equivalent length d_I, and the geometric gap $\Gamma_{\text{geo}} = \Delta_{\text{II}} - d_I$ is defined to diagnose whether "intermediate-scale structure is compressible into endpoints." These structural relations serve as the common theoretical basis for the subsequent one-loop QED calibration and the EFT examples.

6 One-Loop QED Matching Test: Baseline Behavior of Γ in a Threshold-Free Window

This section selects a threshold-free window (the active set of degrees of freedom is unchanged; regulator/scheme are fixed), and in the one-loop QED approximation instantiates the objects from Secs. 3–5 into reproducible calculations: the endpoint divergence Δ_{I}, the information speed Ψ, the path length Δ_{II}, and the computable proxy of the gap term Γ_{geo} (with the relation to the strict gap Γ_{path} stated when needed). The goal is to provide a baseline case: under a one-dimensional truncation and a stable interface, the gap can be treated as negligible at the given resolution, so that endpoints and the full trajectory are consistent within this window.

6.1 Physical window and truncation: single-coupling, threshold-free QED effective description

Take the scale coordinate

$$\gamma = \ln(\mu/\mu_\star), \tag{6.0}$$

where $\mu_\star$ is the reference scale for this section (it may be chosen equal to the reference scale in Sec. 2; endpoint differences do not depend on this choice).

Choose an energy window

$$m_{\text{low}} \ll \mu \ll m_{\text{high}}, \tag{6.1a}$$

so that within this window the set of charged species counted as active remains unchanged (no new mass threshold is crossed), and the RG equation does not switch piecewise on this interval. This is equivalent to fixing the "threshold/active set" as a clause of the viewpoint V: this section only considers a single-segment window under that clause.

Within this window, adopt the minimal truncated model class

$$\mathcal{M}_1 = \{\text{keep only one effective parameter } \alpha(\mu) \text{ or an equivalent coordinate } \theta(\mu)\}, \qquad \theta := \alpha^{-1}. \tag{6.1b}$$

The purpose of introducing θ is to write the one-loop QED running as a linear flow (see the next subsection), thereby constraining the relation between "path geometry" and "endpoint difference" to the minimal number of degrees of freedom.

6.2 One-loop QED running: $\beta(\alpha)$ and the endpoint linear law

The one-loop QED β-function can be written as (with $\alpha = e^2/(4\pi)$) [12–15]

$$\boxed{\frac{\mathrm{d}\alpha}{\mathrm{d}\ln\mu} = \beta(\alpha) = b\,\alpha^2, \qquad b = \frac{2}{3\pi}\sum_{f\in\text{active}} n_f\, Q_f^2 \;> 0,} \tag{6.2}$$

where the sum runs over the charged Dirac fermions treated as active in this window; Q_f is the charge number, and n_f is a degeneracy factor (e.g. a color factor, if applicable, absorbed into n_f). This section uses only two structural facts: b is constant within a threshold-free window, and $\beta(\alpha) > 0$.

Changing coordinates to $\theta = \alpha^{-1}$, we have

$$\frac{\mathrm{d}\theta}{\mathrm{d}\ln\mu} = \frac{\mathrm{d}(\alpha^{-1})}{\mathrm{d}\ln\mu} = -b. \tag{6.3}$$

Hence in a threshold-free window we obtain the endpoint linear law

$$\boxed{\theta(\mu_1) - \theta(\mu_0) = -b\ln\frac{\mu_1}{\mu_0}.} \tag{6.4}$$

This shows that in the θ coordinate, the RG flow is a constant-velocity straight line. (The formal Landau pole occurs at scales far above the window considered here and does not enter the present threshold-free one-loop test.)

6.3 A minimal implementation of the viewpoint V: single interface, statistical object, and induced structure

To instantiate the objects in Secs. 3–5 into closed, computable form, this section fixes a "minimal interface implementation" by taking the interface set to be a single interface,

$$\mathcal{O} = \{O\}, \tag{6.5a}$$

so as to avoid the non-smooth technical burden from multi-interface aggregation (e.g. $\sup_{O\in\mathcal{O}}$).

Let the interface output an estimator Y for θ. At a fixed resolution, we model its conditional distribution by a local Gaussian model

$$P_\theta(y) = \mathcal{N}(\theta, \sigma^2). \tag{6.5b}$$

Here σ makes the "given protocol and resolution" explicit within the viewpoint V: σ is fixed by the protocol (statistics, sample size, systematic error budget, etc.), and represents the "given resolution" in this section.

The endpoint KL divergence for (6.5b) is[7]

$$D(P_{\theta_0} \| P_{\theta_1}) = \frac{(\theta_0 - \theta_1)^2}{2\sigma^2}. \tag{6.6}$$

Its local second-order term yields the Fisher information (local quadratic form/metric) [8, 9]

$$G(\theta) = \frac{1}{\sigma^2} \qquad \text{(constant)}. \tag{6.7}$$

Therefore this section satisfies Condition G_{bridge} (local second-order window) from Sec. 4, and in this model it may be treated as holding everywhere within the window.

Remark 6.1. If one instead uses a concrete physical process as the interface (e.g. a scattering cross section or angular distribution at fixed momentum transfer), then σ and G are determined by the shape of the cross section and statistical errors, and may vary with scale. This section uses (6.5b) as a normalized baseline interface; its role is to realize the meaning of the proxy gap Γ_{geo} in a reproducible closed calculation without introducing additional degrees of freedom.

6.4 Endpoint–path consistency: explicit calculations of $\Delta_{\mathrm{I}}, \Psi, \Delta_{\mathrm{II}}, d_I, \Gamma_{\mathrm{geo}}$

(i) Route I: endpoint divergence Δ_{I} From the definition in Sec. 3 (no sup is needed in the single-interface case) and (6.6), we obtain

$$\boxed{\Delta_{\mathrm{I}}(\mu_0 \to \mu_1) = \frac{(\theta_0 - \theta_1)^2}{2\sigma^2}.} \tag{6.8}$$

(ii) Route II: information speed Ψ and path length Δ_{II} By the definition chain in Sec. 4, in the one-dimensional, constant-quadratic-form case,

$$\Psi(\gamma) = \sqrt{G(\theta)} \left| \frac{\mathrm{d}\theta}{\mathrm{d}\gamma} \right| = \frac{1}{\sigma} |-b| = \frac{b}{\sigma} \qquad \text{(constant)}. \tag{6.9}$$

Hence

$$\boxed{\Delta_{\mathrm{II}}(\mu_0 \to \mu_1) = \int_{\gamma_0}^{\gamma_1} \Psi(\gamma)\, \mathrm{d}\gamma = \frac{b}{\sigma}\, |\gamma_1 - \gamma_0| = \frac{|\theta_0 - \theta_1|}{\sigma}.} \tag{6.10}$$

The last equality uses the endpoint linear law (6.4), and notes that $|\gamma_1 - \gamma_0| = \left|\ln \frac{\mu_1}{\mu_0}\right|$.

(iii) Endpoint-equivalent length d_I and the baseline gap Γ_{geo} By the strict definition in Sec. 5, the endpoint-equivalent length is taken as the endpoint distance induced by the Route II structure,

$$d_I := \mathrm{dist}_G\big(\theta(\mu_0), \theta(\mu_1)\big).$$

In the present one-dimensional constant-metric model, dist_G reduces to a straight-line length, so

$$\boxed{d_I(\mu_0 \to \mu_1) = \frac{|\theta_0 - \theta_1|}{\sigma}.} \tag{6.11}$$

Moreover, in this model the local equivalence relation not only holds but becomes an exact identity:

$$d_I = \sqrt{2\Delta_{\mathrm{I}}}.$$

Therefore, by the definition in Sec. 5, the baseline value of the geometric gap is

$$\boxed{\Gamma_{\mathrm{geo}} = \Delta_{\mathrm{II}} - d_I = \frac{|\theta_0 - \theta_1|}{\sigma} - \frac{|\theta_0 - \theta_1|}{\sigma} = 0.} \tag{6.12}$$

Thus, under the baseline conditions "threshold-free window + one-dimensional truncation $\mathcal{M}_1$ + stable interface resolution,"

$$\boxed{\Gamma_{\mathrm{geo}}(\mu_0 \to \mu_1) = 0 \quad \text{(treatable as endpoint–path consistency at this resolution).}} \tag{6.13}$$

This conclusion corresponds to a clear structural condition: when the viewpoint compresses the system to a single effective coordinate and the induced structure does not switch piecewise within the window, the RG trajectory coincides with the endpoint geodesic connection, and hence "endpoint-only" and "full scale process" are consistent within the window.

6.5 When a nonzero Γ appears: trigger conditions from baseline to nontrivial cases

The reason that $\Gamma_{\mathrm{geo}} = 0$ here is a geometric degeneration from "one-dimensional flow + constant induced structure." To avoid notational confusion, in this section Γ by default refers to the computable proxy Γ_{geo}; the strict gap Γ_{path} is defined in Sec. 5. To obtain an observable $\Gamma > 0$ in the same physical context, it is necessary to break at least one of the following conditions:

1. **Threshold crossing:** the active-species set changes, making (6.2)–(6.4) piecewise; endpoint compression discards segment information, providing a nonzero source for the strict gap Γ_{path}. If the interface/induced structure also switches piecewise at thresholds (or threshold-segment labels are explicitly included in the comparable coordinate system), then Γ_{geo} will also respond nontrivially to the segmented structure.

2. **Model-class extension:** extend $\mathcal{M}_1$ to $\mathcal{M}_2$ (e.g. by adding the first EFT irrelevant-operator coefficient u), so that the flow enters a two- (or higher-) dimensional parameter space; in the generic case the RG trajectory does not coincide with the endpoint geodesic, producing a systematic difference between Δ_{II} (full-trajectory accumulation) and $d_I :=$ dist_G (endpoint-equivalent length), hence a natural source of $\Gamma_{\text{geo}} > 0$.

3. **Switching of interface-sensitive directions:** even if the model class is unchanged, if the observation protocol makes the induced structure strongly scale-dependent within the window (e.g. the dominant statistical error source changes with μ, so that σ or the equivalent G exhibits strong scale dependence), then the difference between the "full-trajectory accumulation" Δ_{II} and the "endpoint length" d_I can also become nonzero.

The next section advances to a nontrivial case with minimal additional cost: still within the QED/EFT context, we take $\mathcal{M}_1 \to \mathcal{M}_2$ (introducing one irrelevant direction or matching parameter), and provide an externally reproducible example with $\Gamma_{\text{geo}} > 0$, thereby advancing "gap diagnosis" from a baseline test to a structural output.

7 EFT Extension Example: A Nonzero Gap Γ Under a Two-Dimensional Truncation

Under the same threshold-free window setup as in Sec. 6, this section extends the truncated model class from one dimension $\mathcal{M}_1 = \{\alpha\}$ to two dimensions $\mathcal{M}_2 = \{\alpha, u\}$, where u denotes a dimensionless coefficient of an EFT irrelevant direction. The purpose of this extension is not to introduce new physical assumptions, but to explicitly include, within the comparable coordinate system, an additional structural direction that may be generated/carried by intermediate scales under Wilsonian scale advancement [1, 2, 16]. This yields a reproducible structural conclusion: at a fixed viewpoint resolution, one generally finds

$$\Gamma_{\text{geo}} > 0, \tag{7.0}$$

i.e. endpoint compression is insufficient to represent full-trajectory accumulation.

To remain consistent with the scale convention in Sec. 2, we continue to use $\gamma = \ln(\mu/\mu_0)$ as the scale coordinate.

To more clearly represent the coarse-graining direction from UV to IR in this section, we introduce $\ell := -\gamma$ as a new scale variable. This variable is numerically equivalent to γ but with the opposite direction, i.e. ℓ increases monotonically as μ decreases. All path-related expressions in this section will be written with ℓ as the parameter; the definition chain is completely consistent with the earlier γ-formulation.

At the same time, to make the parameter along the coarse-graining direction (decreasing resolution) monotonically increasing, we introduce

$$\ell := -\gamma = \ln(\mu_0/\mu), \qquad \ell \in [0, T],\ T > 0, \tag{7.0a}$$

to describe the coarse-graining step from UV toward IR. The following is only a change of variables and does not alter the earlier definition chain.

7.1 Choosing a physical irrelevant direction u

In the EFT description of QED (within a fixed window), one may introduce a higher-dimensional operator direction of dimension $d > 4$ as a representative irrelevant direction [3, 4]. To give a minimal and general example, this section does not bind to a unique operator form, and only requires:

1. This direction is irrelevant in the Wilsonian sense, with canonical scaling exponent

$$p := d - 4 > 0, \tag{7.1}$$

   and it is suppressed in the IR under coarse-graining (increasing ℓ);

2. Under the chosen observation interface (including protocol and resolution), it has observable sensitivity to some statistical direction (otherwise it would be projected to an indistinguishable direction at the interface level, and the gap could be treated as negligible under that viewpoint).

Let the Wilson coefficient along this direction be $C_d(\mu)$, and define the dimensionless coordinate

$$u(\mu) := \mu^p \, C_d(\mu), \qquad p > 0. \tag{7.2}$$

In the minimal computable example of this section, we only use the structural facts $p > 0$ and the monotone decay behavior of u, and do not rely on finer operator details.

7.2 Two-dimensional model class and scale flow (minimal RG approximation)

Take the two-dimensional truncated model class

$$\mathcal{M}_2 = \{\theta(\mu), u(\mu)\}, \qquad \theta := \alpha^{-1}. \tag{7.3}$$

In a threshold-free window, one-loop QED gives (consistent with Sec. 6)

$$\frac{\mathrm{d}\theta}{\mathrm{d}\ell} = b, \qquad b > 0 \text{ is a constant within the window.} \tag{7.4}$$

(With $\ell = -\gamma$, this constant is $b = b$ in the notation of Sec. 6.)For the irrelevant direction u, under the minimal approximation "canonical scaling dominates and higher-order mixing terms with α are neglected," take

$$\frac{\mathrm{d}u}{\mathrm{d}\ell} = -p\,u, \qquad p > 0, \tag{7.5}$$

so that

$$\theta(\ell) = \theta_0 + b\ell, \qquad u(\ell) = u_0\, e^{-p\ell}. \tag{7.6}$$

The geometric meaning of (7.6) is that, in the (θ, u) plane, the scale trajectory is generally a curved curve (unless $u_0 = 0$ or in degenerate cases).

Remark 7.1. If a more specific EFT is needed, one may add mixing terms such as $O(\alpha\, u)$ in (7.5). The mechanism for "$\Gamma_{\mathrm{geo}} > 0$" in this section does not depend on these higher-order details; the decisive factors are $\dim(\mathcal{M}) \geq 2$ and that the trajectory is generically not collinear/affine with respect to the endpoint geodesic connection.

7.3 A two-dimensional interface implementation of the viewpoint V: making both θ and u readable

To make the gap a reproducible quantity, this section fixes a two-dimensional interface implementation (or a single interface containing a two-dimensional sufficient statistic). Abstractly: at each scale point, the observation protocol outputs two types of statistical readouts (Y_θ, Y_u), which are sensitive to θ and u, respectively, and at a given resolution can be modeled by local Gaussian models:

$$Y_\theta \mid \theta \sim \mathcal{N}(\theta, \sigma_\theta^2), \qquad Y_u \mid u \sim \mathcal{N}(u, \sigma_u^2), \tag{7.7}$$

where $(\sigma_\theta, \sigma_u)$ are fixed by the protocol (sample size, systematic errors, fitting strategy, etc.) and form part of the viewpoint V. To minimize degrees of freedom, we take the two readouts to be approximately independent under this interface implementation; if correlations exist, the induced structure will acquire off-diagonal terms, but the gap mechanism itself is unchanged.

The local quadratic form induced by (7.7) is taken in this section to be a constant diagonal form

$$\mathrm{d}s^2 = \frac{\mathrm{d}\theta^2}{\sigma_\theta^2} + \frac{\mathrm{d}u^2}{\sigma_u^2}, \qquad G = \mathrm{diag}\big(\sigma_\theta^{-2}, \sigma_u^{-2}\big)\,. \tag{7.8}$$

7.4 Explicit calculations: Δ_{I}, d_I, Δ_{II}, and Γ_{geo}

Let the coarse-graining window length be $T := \ell_1 - \ell_0 > 0$, and take $\ell_0 = 0,\ \ell_1 = T$. From (7.6) the endpoint differences are

$$\Delta\theta = \theta(T) - \theta(0) = bT, \qquad \Delta u = u(T) - u(0) = u_0\left(e^{-pT} - 1\right). \tag{7.9}$$

(i) Route I: endpoint divergence Δ_{I} For Gaussian families with the same covariance, the closed form of KL divergence gives the endpoint divergence

$$\boxed{\Delta_{\mathrm{I}} = \frac{1}{2}\left(\frac{(\Delta\theta)^2}{\sigma_\theta^2} + \frac{(\Delta u)^2}{\sigma_u^2}\right).} \tag{7.10}$$

(ii) Endpoint-equivalent length $d_I := \mathrm{dist}_G$ By the definition in Sec. 5, the endpoint-equivalent length is the endpoint distance induced by the present structure, $d_I := \mathrm{dist}_G\big((\theta_0, u_0), (\theta_T, u_T)\big)$. Under the constant diagonal metric (7.8), dist_G reduces to a scaled Euclidean distance, hence

$$\boxed{d_I = \sqrt{\frac{(bT)^2}{\sigma_\theta^2} + \frac{u_0^2(1 - e^{-pT})^2}{\sigma_u^2}}\,.} \tag{7.11}$$

Note that in this "constant metric + equal-covariance Gaussian family" model, indeed $d_I = \sqrt{2\Delta_{\mathrm{I}}}$; this is an equivalence relation in a baseline implementation, not an identity in general settings.

(iii) Route II: information speed and path length Δ_{II} From (7.4), (7.5), and (7.8), the information speed is

$$\Psi(\ell) = \sqrt{\frac{1}{\sigma_\theta^2}\left(\frac{\mathrm{d}\theta}{\mathrm{d}\ell}\right)^2 + \frac{1}{\sigma_u^2}\left(\frac{\mathrm{d}u}{\mathrm{d}\ell}\right)^2} = \sqrt{\frac{b^2}{\sigma_\theta^2} + \frac{p^2u_0^2e^{-2p\ell}}{\sigma_u^2}}. \tag{7.12}$$

The path length is

$$\boxed{\Delta_{\rm II} = \int_0^T \Psi(\ell)\,\mathrm{d}\ell = \int_0^T \sqrt{\frac{b^2}{\sigma_\theta^2} + \frac{p^2 u_0^2 e^{-2p\ell}}{\sigma_u^2}}\,\mathrm{d}\ell.} \tag{7.13}$$

This integral can be computed numerically in a stable way; if a closed form is needed, it can be written as a combination of logarithms and radicals via an elementary substitution (recommended to place in an appendix; not expanded in the main text).

(iv) Geometric gap: $\Gamma_{\rm geo} \geq 0$, and generically strictly positive By the definition in Sec. 5,

$$\boxed{\Gamma_{\rm geo} := \Delta_{\rm II} - d_I.} \tag{7.14}$$

Under the constant-metric setting (7.8), rescale coordinates as

$$x(\ell) := \theta(\ell)/\sigma_\theta, \qquad y(\ell) := u(\ell)/\sigma_u,$$

then $\Delta_{\rm II}$ is the Euclidean arc length of the curve $(x(\ell), y(\ell))$, while d_I is the chord length between endpoints under the same Euclidean metric. Therefore one must have

$$\boxed{\Gamma_{\rm geo} \geq 0.} \tag{7.15}$$

Moreover, the necessary and sufficient condition for $\Gamma_{\rm geo} = 0$ is that the image of the curve lies on the endpoint line segment and does not turn back (the tangent direction does not reverse). In this section $x(\ell) = x_0 + (b/\sigma_\theta)\ell$ is monotonically increasing, so there is no turning back; thus $\Gamma_{\rm geo} = 0$ is equivalent to the curve being a straight segment in the (x, y) plane, i.e.

$$y(\ell) \text{ as a function of } x(\ell) \text{ is affine (a straight line).} \tag{7.16}$$

For the exponentially decaying trajectory in (7.6), $u(\ell) = u_0 e^{-p\ell}$, this condition fails unless $u_0 = 0$ (degenerating back to one dimension) or $p = 0$ (no longer an irrelevant direction). Hence we obtain the minimal nontriviality statement of this section:

$$\boxed{u_0 \neq 0,\ p > 0,\ T > 0 \implies \Gamma_{\rm geo} > 0.} \tag{7.17}$$

Theorem 7.1 (Strict positivity of $\Gamma_{\rm geo}$ in the minimal two-dimensional truncation)**.** *In the two-dimensional implementation of Sec. 7 with constant diagonal metric* (7.8) *and minimal flow* (7.6) *(equivalently* (7.4)*–*(7.5)*), the geometric proxy gap satisfies, for any window length $T > 0$,*

$$u_0 \neq 0,\ p > 0 \implies \Gamma_{\rm geo}(T) > 0.$$

Moreover, for $T > 0$, equality holds only in the degenerate cases $u_0 = 0$ (one-dimensional reduction) or $p = 0$ (no longer an irrelevant direction), see Appendix B, (B.19)*.*

This shows that, within the same physical system and the same threshold-free window, merely including one EFT irrelevant direction in the model class and making it readable at the interface produces a structural difference between endpoint compression and full-trajectory accumulation.

7.5 Small-quantity approximation: when Γ_{geo} is negligible

To make "negligible / non-negligible" an operational criterion, introduce the slope ratio

$$s(\ell) := \frac{\dot{y}(\ell)}{\dot{x}(\ell)} = \frac{(\dot{u}/\sigma_u)}{\left(\dot{\theta}/\sigma_\theta\right)} = -\frac{p\,\sigma_\theta}{b\,\sigma_u}\,u_0\,e^{-p\ell}, \qquad (\dot{\ } := \mathrm{d}/\mathrm{d}\ell). \tag{7.18}$$

When $|s(\ell)| \ll 1$ (i.e. the irrelevant direction is uniformly weak at the given resolution), one may perform a second-order expansion of the arc length and obtain the approximate structure

$$\Gamma_{\text{geo}} \approx \frac{b}{2\sigma_\theta}\left[\int_0^T s(\ell)^2\,\mathrm{d}\ell - \frac{1}{T}\left(\int_0^T s(\ell)\,\mathrm{d}\ell\right)^2\right] = \frac{bT}{2\sigma_\theta}\,\mathrm{Var}_{[0,T]}[s(\ell)] \ \geq 0. \tag{7.19}$$

Proposition 7.1 (Quadratic regime: variance control and small-T behavior)**.** *Assume $|s(\ell)| \ll 1$ on $[0,T]$ so that the second-order expansion leading to (7.19) is valid. Then the leading contribution to Γ_{geo} is controlled by the scale variance of the slope ratio $s(\ell)$, namely (7.19). For the exponential trajectory (7.18), writing $s(\ell) = -K\,u_0\,e^{-p\ell}$ with $K := p\sigma_\theta/(b\sigma_u)$, one obtains*

$$\Gamma_{\text{geo}} \ \approx\ \frac{bT}{2\sigma_\theta}\,K^2 u_0^2\left[\frac{1-e^{-2pT}}{2pT} - \left(\frac{1-e^{-pT}}{pT}\right)^2\right] \ \geq 0. \tag{7.20}$$

In particular, as $T \to 0$,

$$\Gamma_{\text{geo}} \ = \ \frac{p^4\,\sigma_\theta}{24\,b\,\sigma_u^2}\,u_0^2\,T^3 \ + \ O(T^4) \qquad (\textit{within the quadratic regime}). \tag{7.21}$$

Equation (7.19) provides a direct criterion: under this viewpoint, the leading contribution to Γ_{geo} is controlled by the scale fluctuation (variance) of $s(\ell)$. If $s(\ell)$ is approximately constant (or approximately 0), then Γ_{geo} can be treated as negligible; if $s(\ell)$ varies significantly with ℓ (e.g. exponential decay), then Γ_{geo} should not be compressed to 0 by an endpoint description.

7.6 Reproducible procedure (baseline–nontrivial contrast with Sec. 6)

Note: this section uses a single joint interface that reads out the statistics (Y_θ, Y_u) for (θ, u), and therefore does not use the $\sup_{O\in\mathcal{O}}$ aggregation rule defined in Sec. 3. The error structure of this joint interface is represented by a two-dimensional covariance matrix (diagonal approximation in this section); the aggregation rule is locked as part of the viewpoint V. For reproducibility, one should fully report the interface choice, resolution parameters, and the structure of the error model.

Under a fixed viewpoint V, the minimal reproducible procedure corresponding to this example is:

1. Fix the window: choose a threshold-free window and fix the scheme and the regulator-handling convention;
2. Fix the model class: take $\mathcal{M}_2 = \{\theta, u\}$, and lock the definition $u = \mu^p C_d(\mu)$;
3. Fix the interface: choose statistical directions that can read out both θ and u, and specify the resolution $(\sigma_\theta, \sigma_u)$;

4. Fit the scale flow: within the window, estimate the linear drift of $\theta(\ell)$ and the decay of $u(\ell)$ (or its more general form);
5. Compute: compute Δ_{I}, d_I, Δ_{II}, Γ_{geo} by (7.10)–(7.14) and report error propagation;
6. Criterion: test whether Γ_{geo} is negligible at the given resolution; if it is not negligible, then "endpoint $\approx$ trajectory" does not hold under this viewpoint, and intermediate-scale structure must enter the model class via explicit directions (e.g. u).

Summary Section 6 provides a one-dimensional baseline case with $\Gamma_{\mathrm{geo}} = 0$, corresponding to endpoint-compressibility; this section provides a minimal physical example with $\Gamma_{\mathrm{geo}} > 0$ in the same window via the two-dimensional truncation $\mathcal{M}_2 = \{\alpha, u\}$, corresponding to intermediate-scale structure not being fully representable by endpoints. This contrast yields a structural statement: the nonzero gap term is not rhetorical, but a testable output determined jointly by the viewpoint resolution, the model-class dimension, and the geometry of the scale flow.

8 Discussion: Measurement Semantics and Constraint Norms for Γ

This section does only three things: (i) state the strict semantics of Γ; (ii) provide a reproducible estimation and decision protocol; and (iii) write "$\Gamma \approx 0$ / $\Gamma > 0$" as falsifiable statements. Throughout, we continue to work under the same fixed viewpoint from Sec. 2,

$$\mathsf{V} = (\mathcal{O}, \text{protocol}, R, S, \text{threshold rule}, \mathcal{M}, \pi, D, \mathrm{Agg}) \tag{8.1}$$

and changing any component of V changes the problem itself.

8.1 Strict objects: endpoint divergence, path divergence, and gap decomposition

Endpoint divergence (Route I) Given endpoint scales (μ_0, μ_1), define the endpoint divergence by

$$\Delta_{\mathrm{I}}(\mu_0 \to \mu_1) := \operatorname*{Agg}_{O \in \mathcal{O}} D\Big(X^{(O)}_{\mu_0} \| X^{(O)}_{\mu_1}\Big), \tag{8.2}$$

where Agg is the interface aggregation rule locked in the viewpoint V (identity in the single-interface case; in the multi-interface case one may take sup or a fixed weighting, but it must be fixed in V and kept consistent throughout).

Path divergence and gap (strict version) Take a discrete scale grid $\mu_0 < \mu_1 < \cdots < \mu_N$, and let the path variable be

$$Z := (X_0, X_1, \ldots, X_N), \tag{8.3}$$

where X_k denotes the corresponding protocol readout random variable (or its sufficient statistic) at scale μ_k, whose law is induced by the predictive object fixed by V. For two scale flows being compared, under the same scale-advancement mechanism (Sec. 2, N3) and the same viewpoint V, they induce two path processes

$$\mathbb{P}_Z, \qquad \mathbb{Q}_Z. \tag{8.4}$$

Define the strict path divergence by

$$\Delta_{\text{path}}(0 \to N) := D(\mathbb{P}_Z \| \mathbb{Q}_Z)\,. \tag{8.5}$$

Chain-rule decomposition and the strict gap Decompose the path variable into endpoints $E := (X_0, X_N)$ and history $H := (X_1, \dots, X_{N-1})$. The chain rule for relative entropy gives the identity

$$\Delta_{\text{path}}(0 \to N) = D(\mathbb{P}_E \| \mathbb{Q}_E) + \mathbb{E}_{\mathbb{P}_E}\Big[D(\mathbb{P}_{H|E} \| \mathbb{Q}_{H|E})\Big]. \tag{8.6}$$

Accordingly define the strict path gap term

$$\Gamma_{\text{path}}(0 \to N) := \mathbb{E}_{\mathbb{P}_E}\Big[D(\mathbb{P}_{H|E} \| \mathbb{Q}_{H|E})\Big] \;\geq 0, \tag{8.7}$$

and rewrite (8.6) as

$$\boxed{\Delta_{\text{path}}(0 \to N) = \Delta_{\mathrm{I}}^{(\text{pair})}(0 \to N) + \Gamma_{\text{path}}(0 \to N), \qquad \Delta_{\mathrm{I}}^{(\text{pair})}(0 \to N) := D(\mathbb{P}_E \| \mathbb{Q}_E).} \tag{8.8}$$

Endpoint upper bound (DPI) By the data processing inequality (endpoint projection is a coarse-graining), one must have

$$\Delta_{\mathrm{I}}^{(\text{pair})}(0 \to N) \leq \Delta_{\text{path}}(0 \to N). \tag{8.9}$$

Equality condition (unambiguous version)

$$\Gamma_{\text{path}}(0 \to N) = 0 \quad \Longleftrightarrow \quad D(\mathbb{P}_{H|E} \| \mathbb{Q}_{H|E}) = 0 \;\; \mathbb{P}_E\text{-a.s.}. \tag{8.10}$$

That is: under the viewpoint V, the endpoints E are sufficient statistics for distinguishing the two path processes; the history H provides no additional distinguishable information.

Notation convention (alignment of gap terminology) We use Γ as a collective name for the "difference terms lost by endpoint compression." At the strict level this difference is quantified by Γ_{path} (see (8.7)–(8.8)); at the computable level we use the geometric proxy Γ_{geo} (see the next subsection). Importantly, Γ is not an additional independent quantity: it is a decomposition residual of the same underlying gap Δ associated with an endpoint/trajectory split. Consequently, both the numerical value of Γ and the decision "negligible / non-negligible" are meaningful only under a fixed viewpoint V together with the chosen resolution clause.

8.2 Computable proxy: domain and decision semantics of Γ_{geo} (locked)

To avoid "uncertain sign properties of the proxy quantity," we use the geometric gap Γ_{geo} only when all of the following conditions hold.

Condition G_{geo} (domain of validity for the geometric proxy) There exist parameter coordinates $g(\gamma) \in \mathbb{R}^d$ and a (piecewise) C^1 positive-definite metric (or, more generally, a positive-definite quadratic form) $G(g)$, such that

$$\mathrm{d}s^2 = \mathrm{d}g^\top G(g)\,\mathrm{d}g, \qquad \Delta_{\mathrm{II}}(\gamma_0 \to \gamma_1) := \int_{\gamma_0}^{\gamma_1} \sqrt{\dot{g}^\top G(g)\,\dot{g}}\,d\gamma \tag{8.11}$$

If G is induced via an aggregated local quadratic form (e.g. by $\mathrm{Agg} = \sup_{O\in\mathcal{O}}$ over interfaces), the aggregation rule must select a single positive-definite matrix field $G(g)$ on the window (at least piecewise C^1), so that dist_G and geodesicity are well-defined; otherwise the proxy should be interpreted in the more general length-space/Finsler sense, which we do not pursue here. Δ_{II} is defined as the arc length under this structure, and the endpoint-equivalent length d_I is defined as the endpoint distance under the same structure,

$$d_I(\gamma_0 \to \gamma_1) := \mathrm{dist}_G\big(g(\gamma_0), g(\gamma_1)\big). \tag{8.12}$$

Here dist_G denotes the *global* geodesic distance induced by G (an infimum over all connecting curves), and is not defined as the length of (nor restricted to) the observed discrete trajectory segment $\{g(\gamma_k)\}$ on the window.

Under Condition G_{geo}, define

$$\boxed{\Gamma_{\mathrm{geo}} := \Delta_{\mathrm{II}} - d_I.} \tag{8.13}$$

Equivalently, Γ_{geo} is the *geodesic excess* of the trajectory segment under G: the arc length Δ_{II} minus the endpoint geodesic distance d_I.

Then the sign and equality condition of Γ_{geo} have strict semantics:

$$\boxed{\begin{aligned} &\Gamma_{\mathrm{geo}} \geq 0, \\ \Gamma_{\mathrm{geo}} = 0 \iff\ & g(\gamma) \text{ is a geodesic (under } G\text{) connecting the two endpoints} \\ & \text{(affine under arc-length parametrization).} \end{aligned}} \tag{8.14}$$

The QED/EFT implementations in this paper satisfy Condition G_{geo} Sections 6–7 use equal-covariance Gaussian interfaces (or an equivalent Fisher implementation), so that G is a constant diagonal metric and the endpoint distance is a scaled Euclidean distance. In this implementation, $\sqrt{2\Delta_{\mathrm{I}}}$ coincides exactly with dist_G, hence the relation

$$d_I = \sqrt{2\Delta_{\mathrm{I}}}$$

appearing in Secs. 6–7 is an identity rather than an approximation.

Realization clause (theoretical objects vs. computed estimates). Equations (8.11)–(8.14) define the *theoretical* objects Δ_{II}, $d_I = \mathrm{dist}_G$, and $\Gamma_{\mathrm{geo}} := \Delta_{II} - d_I$, where dist_G is the global geodesic distance induced by G:

$$\mathrm{dist}_G(u, v) := \inf_{x(0)=u,\, x(1)=v} \int_0^1 \sqrt{x'(s)^\top G(x(s)) x'(s)}\,ds.$$

In computations, one reports $\widehat{\Delta}_{II}$, $\widehat{d}_I$, and $\widehat{\Gamma_{\mathrm{geo}}} := \widehat{\Delta}_{II} - \widehat{d}_I$ as *estimates* of the corresponding theoretical quantities under a locked viewpoint V. The numerical procedures used to produce these estimates are part of V (hence must be reported), but they do not redefine the objects.

Details of a reproducible instantiation are given in Appendix D.

8.3 Estimation and Decision Protocol (turning "≈ 0" into a decidable statement)

This subsection provides a *reproducible* protocol that is *comparable across works*. All decisions are made under the same viewpoint $\mathsf{V} = (\mathcal{O}, \text{protocol}, R, S, \text{threshold rule}, \mathcal{M}, \pi, D, \text{Agg})$; changing any component changes the problem itself. When the geometric proxy is used, the reported quantities are *estimates* of the corresponding theoretical objects defined in (8.11)–(8.14); the numerical procedures that produce these estimates belong to V and must be reported, but they do not redefine the objects (see Appendix D for a reproducible instantiation template).

Inputs (must be reported; otherwise cross-work comparison as the same scalar is not allowed)

1. **Window and coordinates:** the window $[\mu_0, \mu_1]$ and the choice of scale coordinate ($\gamma = \ln(\mu/\mu_0)$ or $\ell = -\gamma$). If a scan over window length T is performed (e.g. $\ell \in [0, T]$ or $\gamma_1 - \gamma_0 = T$), one must report the rule that generates the window family $\{W(T)\}$ (e.g. fixing μ_0 and varying $\mu_1(T)$), and the discrete grid $\{T_j\}_{j=1}^m$ (or the equivalent $\{\mu_{1,j}\}$).

2. **Complete list of viewpoint clauses:** $\mathcal{O}$ and protocol, R, S, the threshold rule, the model class $\mathcal{M}$, reconstruction π, the divergence/distance function D, and the interface aggregation rule Agg (including an explicit statement of how regulator dependence is handled: density-normalization / differencing / explicit retention, and the corresponding implementation convention).

3. **If using the geometric proxy:** one must (i) declare that Condition G_{geo} holds (see Sec. 8 above, (8.11)–(8.14)), and (ii) report the viewpoint-locked procedures used to produce the estimates $\widehat{\Delta_{\text{II}}}$ and $\widehat{d_I}$ (hence $\widehat{\Gamma_{\text{geo}}}$). These procedures are part of V: changing them changes the scalar being reported, even if the symbolic formula $\Gamma_{\text{geo}} = \Delta_{\text{II}} - d_I$ is unchanged. (For avoidance of doubt: $\widehat{d_I}$ is intended as an approximation to the geodesic distance dist_G defined by the induced metric G, rather than a path-length restricted to the observed discrete trajectory.)

4. **Uncertainty model (locked):** one must provide a fixed method for generating uncertainty that propagates to $\widehat{\Gamma_{\text{geo}}}$ (choose one and lock it):

   (a) analytic error propagation (must report the covariance matrix / linearization rule); or

   (b) bootstrap / resampling (must report the resampling unit, the number of resamples B, and the interval construction rule) [23].

5. **Significance level (locked):** lock a global one-sided significance level $\alpha \in (0, 1)$. If a window scan (multiple testing) is performed, one must also lock a multiple-comparison correction rule; this paper uses Bonferroni by default[24, 25]:

$$\alpha_j := \alpha/m, \qquad j = 1, \dots, m. \tag{8.15}$$

(If other corrections such as Holm are used, they must also be locked in the inputs and kept consistent throughout.)

Outputs (must be reported) For each window (or each T_j), one must report

$$\widehat{\Delta_{\mathrm{I}}}, \qquad \widehat{\Delta_{\mathrm{II}}}, \qquad \hat{d}_I, \qquad \widehat{\Gamma_{\mathrm{geo}}}, \qquad \mathrm{CI}_L\big(\widehat{\Gamma_{\mathrm{geo}}}; 1-\alpha\big), \tag{8.16}$$

where $\mathrm{CI}_L(\cdot; 1-\alpha)$ denotes a *one-sided lower confidence bound*. If a window scan is performed, replace it by $\mathrm{CI}_L(\cdot; 1-\alpha_j)$ (see (8.15)).

Decision rule (null-hypothesis test, locked) Define the null and alternative:

$$H_0: \ \Gamma_{\mathrm{geo}} = 0, \qquad H_1: \ \Gamma_{\mathrm{geo}} > 0. \tag{8.17}$$

One-sided decision rule:

$$\text{if } \mathrm{CI}_L\big(\widehat{\Gamma_{\mathrm{geo}}}; 1-\alpha\big) > 0, \text{ reject } H_0; \qquad \text{otherwise do not reject } H_0. \tag{8.18}$$

If a window scan is performed, replace α by α_j.

Break location (first-detection window) Under a fixed viewpoint V, fixed resolution clauses (including the uncertainty model / interval construction), and a locked significance level (including the multiple-testing correction), parameterize the window length by $T > 0$ and take a discrete grid $0 < T_1 < \cdots < T_m$. For each T_j, compute $\widehat{\Gamma_{\mathrm{geo}}}(T_j)$ and $\mathrm{CI}_L\big(\widehat{\Gamma_{\mathrm{geo}}}(T_j); 1-\alpha_j\big)$.

Define "a gap is detected at T_j" by

$$\mathrm{Det}(T_j): \quad \mathrm{CI}_L\big(\widehat{\Gamma_{\mathrm{geo}}}(T_j); 1-\alpha_j\big) > 0.$$

Then the discrete implementation of the first-detection scale is defined by

$$\widehat{T}_\delta \ := \ \min\{T_j : \ \mathrm{Det}(T_j)\}, \tag{8.19}$$

and one reports a resolution-limited bracketing interval using adjacent grid points (e.g. $[T_{j-1}, T_j]$). If bootstrap is used to obtain an empirical distribution of $\widehat{T}_\delta$, one may report $[T_\delta^-, T_\delta^+]$ via quantiles, but this rule must be locked in the inputs.

Endpoint–path formulation of the structural consistency axiom: an intrinsic residual across viewpoints

This subsection formulates "whether endpoint compression can be removed by choosing a viewpoint" as a reproducible structural statement. Recall: under a fixed viewpoint V, the strict gap $\Gamma_{\mathrm{path}}^{(\mathsf{V})} \geq 0$ is given by the chain-rule decomposition of relative entropy, and $\Gamma_{\mathrm{path}}^{(\mathsf{V})} = 0$ if and only if the endpoint variable $E = (X_0, X_N)$ is a sufficient statistic for the path difference under that viewpoint (Appendix A).

Admissible viewpoint family $\mathfrak{V}$ (admissible viewpoints) Let $\mathfrak{V}$ denote a set of *admissible* viewpoints. Here "admissible" only means:

1. it satisfies the normative clauses in Sec. 2 (interface/protocol, regulator and scheme, threshold rule, model class and reconstruction, divergence function and aggregation operator, etc. are all reportable and locked within each viewpoint);

2. it has the *readability* stipulated by the research goal (it can read out the claimed parameter directions / statistical directions);

3. it excludes degenerate resolutions: it does not allow an artificially zero gap by making the interface arbitrarily coarse (e.g. one may require a uniform lower bound $G(\gamma) \succeq G_{\min} \succ 0$, or an equivalent resolution lower-bound clause).

Changing to a viewpoint $\mathsf{V} \notin \mathfrak{V}$ is treated as changing the problem itself, not as a different writing of the same object.

Definition 8.1 (Intrinsic residual across viewpoints)**.** Under a fixed scale window $[\mu_0, \mu_1]$ and a fixed admissible viewpoint family $\mathfrak{V}$, define the cross-viewpoint infimum of the strict gap by

$$\Gamma_*(\mu_0 \to \mu_1) \;:=\; \inf_{\mathsf{V}\in\mathfrak{V}} \Gamma^{(\mathsf{V})}_{\text{path}}(\mu_0 \to \mu_1) \;\geq 0. \tag{8.20}$$

When the main text uses the geometric proxy gap $\Gamma^{(\mathsf{V})}_{\text{geo}}$ for computable diagnostics, one may correspondingly define

$$\Gamma^{\text{geo}}_*(\mu_0 \to \mu_1) \;:=\; \inf_{\mathsf{V}\in\mathfrak{V}} \Gamma^{(\mathsf{V})}_{\text{geo}}(\mu_0 \to \mu_1) \;\geq 0, \tag{8.21}$$

and emphasize that Γ^{geo}_* serves only as a computable proxy for Γ_*; its interpretation should be anchored to the chain-rule decomposition for the strict object Γ_{path}.

Proposition 8.1 (Cross-viewpoint lower-bound criterion (endpoint–path version))**.** *Under a fixed scale window and a fixed admissible viewpoint family $\mathfrak{V}$:*

1. *If there exists a viewpoint $\mathsf{V} \in \mathfrak{V}$ such that $\Gamma^{(\mathsf{V})}_{\text{path}}(\mu_0 \to \mu_1) = 0$, then $\Gamma_*(\mu_0 \to \mu_1) = 0$.*

2. *If $\Gamma_*(\mu_0 \to \mu_1) > 0$, then for every $\mathsf{V} \in \mathfrak{V}$ one has $\Gamma^{(\mathsf{V})}_{\text{path}}(\mu_0 \to \mu_1) \geq \Gamma_*(\mu_0 \to \mu_1) > 0$; in this case, endpoint compression cannot eliminate the full-trajectory historical information under any admissible viewpoint.*

3. *If $\Gamma_*(\mu_0 \to \mu_1) = 0$ but attainability of the infimum is not claimed, then its strict meaning is: for every $\varepsilon > 0$ there exists $\mathsf{V}_\varepsilon \in \mathfrak{V}$ such that $\Gamma^{(\mathsf{V}_\varepsilon)}_{\text{path}}(\mu_0 \to \mu_1) \leq \varepsilon$; that is, the gap can be made* arbitrarily small *within the viewpoint family, but there need not exist a single viewpoint for which it is strictly zero.*

Remark (comparison with Secs. 6–7) The baseline case $\Gamma_{\text{geo}} = 0$ in Sec. 6 corresponds to an admissible viewpoint implementation under which, at the given resolution and model class, the system effectively degenerates to a one-dimensional readable coordinate (geodesic degeneration), hence endpoints and the full trajectory agree. In Sec. 7, after including an EFT irrelevant direction in the model class and making it readable at the interface, one typically obtains $\Gamma_{\text{geo}} > 0$, meaning endpoint compression is no longer sufficient. This distinction emphasizes that once the viewpoint clauses are locked, the decision "negligible / non-negligible" for Γ becomes a testable output.

8.4 Falsifiable statements

This subsection lists several statements. Each statement is phrased so that, once objects are locked, it can be decided as true/false; unless otherwise stated, viewpoint superscripts are omitted below.

T0 (cross-viewpoint lower bound: falsifiability of intrinsic residual) Lock an allowed viewpoint family $\mathfrak{V}_{\mathrm{adm}} \subseteq \mathfrak{V}$, and impose *hard constraints* to exclude degenerate resolutions (e.g. require $G(\gamma) \succeq G_{\min} \succ 0$). Define the (proxy-level) minimal gap

$$\underline{\Gamma}_{\mathrm{geo}}([\mu_0, \mu_1]) := \inf_{\mathsf{V} \in \mathfrak{V}_{\mathrm{adm}}} \Gamma^{(\mathsf{V})}_{\mathrm{geo}}([\mu_0, \mu_1]). \tag{8.22}$$

If $\underline{\Gamma}_{\mathrm{geo}}([\mu_0, \mu_1]) > 0$, then the window is said to have an *intrinsic residual lower bound* under $\mathfrak{V}_{\mathrm{adm}}$ and the resolution lower-bound constraint; i.e. for every allowed viewpoint $\mathsf{V} \in \mathfrak{V}_{\mathrm{adm}}$, one has $\Gamma^{(\mathsf{V})}_{\mathrm{geo}} \geq \underline{\Gamma}_{\mathrm{geo}} > 0$. Conversely, if there exists an allowed viewpoint such that $\Gamma^{(\mathsf{V})}_{\mathrm{geo}} = 0$, then necessarily $\underline{\Gamma}_{\mathrm{geo}} = 0$.

T1 (sufficient-statistic window: geometric criterion) Under a locked viewpoint V and assuming Condition G_{geo} holds,

$$\begin{aligned} \Gamma_{\mathrm{geo}} = 0 \iff g(\gamma) &\text{ is a geodesic under the induced metric } (G), \\ &\text{(equivalently: arc length = endpoint geodesic distance).} \end{aligned} \tag{8.23}$$

Therefore, testing T1 is exactly the one-sided test $H_0 : \Gamma_{\mathrm{geo}} = 0$ described in Subsec. 8.3.

T2 (two-dimensional EFT irrelevant direction: necessary and sufficient condition for $\Gamma_{\mathrm{geo}} > 0$ and closed-form output) In a threshold-free window, Sec. 7 takes the two-dimensional truncation $\mathcal{M}_2 = \{\theta, u\}$ and induces a constant diagonal metric under an equal-covariance Gaussian interface,

$$\mathrm{d}s^2 = \frac{\mathrm{d}\theta^2}{\sigma_\theta^2} + \frac{\mathrm{d}u^2}{\sigma_u^2}. \tag{8.24}$$

Let $\ell \in [0, T]$ be the UV$\to$IR coarse-graining parameter, and take the scale flow

$$\frac{\mathrm{d}\theta}{\mathrm{d}\ell} = b, \quad b > 0; \qquad \frac{\mathrm{d}u}{\mathrm{d}\ell} = -p\,u, \quad p > 0; \qquad \theta(\ell) = \theta_0 + b\ell, \quad u(\ell) = u_0 e^{-p\ell}. \tag{8.25}$$

Then under this viewpoint implementation,

$$\Gamma_{\mathrm{geo}}(T) = 0 \iff (u_0 = 0) \text{ or } (p = 0); \qquad u_0 \neq 0,\ p > 0,\ T > 0 \implies \Gamma_{\mathrm{geo}}(T) > 0. \tag{8.26}$$

For a closed-form, reproducible output, define

$$a := \frac{b}{\sigma_\theta} > 0, \qquad c := \frac{p\,u_0}{\sigma_u}.$$

Then

$$\begin{aligned}\Delta_{\mathrm{II}}(T) &= \int_0^T \sqrt{a^2 + c^2 e^{-2p\ell}}\ \mathrm{d}\ell \\ &= aT + \frac{1}{p}\left[\sqrt{a^2+c^2} - \sqrt{a^2 + c^2 e^{-2pT}} + a\ln\frac{a + \sqrt{a^2 + c^2 e^{-2pT}}}{a + \sqrt{a^2+c^2}}\right],\end{aligned} \tag{8.27}$$

and the endpoint geodesic distance is

$$d_I(T) = \sqrt{(aT)^2 + \left(\frac{u_0}{\sigma_u}(1 - e^{-pT})\right)^2}, \tag{8.28}$$

hence

$$\Gamma_{\mathrm{geo}}(T) = \Delta_{\mathrm{II}}(T) - d_I(T) \tag{8.29}$$

is a closed-form, reproducible object under this implementation.

Moreover, the two limits of $\Gamma_{\mathrm{geo}}(T)$ can be given directly:

$$\Gamma_{\mathrm{geo}}(T) = \frac{a^2}{24} \cdot \frac{\left(\frac{u_0}{\sigma_u}\right)^2 p^4}{\left(a^2 + \left(\frac{u_0}{\sigma_u}\right)^2 p^2\right)^{3/2}} T^3 + O(T^4), \qquad (T \downarrow 0), \tag{8.30}$$

$$\lim_{T\to\infty} \Gamma_{\mathrm{geo}}(T) = \frac{1}{p}\left[\sqrt{a^2 + \left(\frac{u_0}{\sigma_u}\right)^2 p^2} - a + a\ln\frac{2a}{a + \sqrt{a^2 + \left(\frac{u_0}{\sigma_u}\right)^2 p^2}}\right] > 0 \quad (u_0 \neq 0, p > 0). \tag{8.31}$$

T3 (threshold crossing stitching: if used, it must be tested by this clause) If the window crosses a threshold and the threshold rule is locked, one must also report the gaps on the three windows:

$$\Gamma_{\mathrm{geo}}([\mu_0, \mu_*])\ , \quad \Gamma_{\mathrm{geo}}([\mu_*, \mu_1])\ , \quad \Gamma_{\mathrm{geo}}([\mu_0, \mu_1]), \tag{8.32}$$

where μ_* is the threshold point (or stitching point). The test requirement of T3 is: all three must be computed under the same viewpoint V, and the threshold matching conditions must be reported; if any item is missing, one must not claim a "Γ behavior on threshold-crossing windows."

8.5 Checklist (minimum clauses for reproducibility and cross-work comparability)

Any work applying the divergence and gap quantities of this paper must at least report:

1. the window $[\mu_0, \mu_1]$ and coordinates (γ or ℓ), and the window-scan grid (if used);
2. the concrete implementation of $\mathcal{O}$ and the protocol (statistics, energy/momentum points, normalization, sampling);
3. the regulator R and scheme S, and the choice among density-normalization / differencing / explicit retention of regulator dependence;

4. the threshold rule and matching conditions (if thresholds are involved);
5. the model class $\mathcal{M}$ and reconstruction π (including the uniqueness criterion);
6. the divergence/distance function D and the interface aggregation rule Agg;
7. if using Γ_{geo}: one must provide G, $g(\gamma)$, $d_I = \text{dist}_G$, and declare Condition G_{geo} holds;
8. the uncertainty model and interval construction (including the locked α and multiple-testing correction rule, if a scan is used).

If any item is missing, the result must be classified as a cross-viewpoint output and must not be used for direct comparison as the same scalar with other works.

9 Conclusion

This paper proposes a *viewpoint-relative* cross-scale gap framework. Under fixed observational interfaces, regulator choices, threshold rules, and a model class, it rewrites "structural differences along scale evolution" into distinguishable information differences, and establishes a reproducible structural relationship between the endpoint-type (Route I) and the full-trajectory-type (Route II) descriptions. The basic convention is: once the viewpoint

$$\mathsf{V} = (\mathcal{O}, \text{protocol}, R, S, \text{threshold rule}, \mathcal{M}, \pi, D, \text{Agg})$$

is locked, all subsequent gap quantities, length quantities, and inequalities are defined under this viewpoint; changing any component is treated as changing the problem itself, rather than as a different expression of the same object.

Under a fixed viewpoint V, this paper defines the endpoint gap Δ_{I} (Route I). Meanwhile, when the local second-order bridging is valid and Condition G_{geo} holds (or an equivalent single-interface/Fisher implementation), an information speed Ψ is generated from the induced quadratic form/metric and the RG flow field, and the computable path length (Route II) is defined by

$$\Delta_{\text{II}} = \int \Psi \, d\gamma, \qquad \Psi = \sqrt{\dot{g}^{\top} G(g)\, \dot{g}} \quad \text{(under Condition } G_{\text{bridge}}\text{)}.$$

At the strict level, this paper defines the full-trajectory gap via a path process object,

$$\Delta_{\text{path}} := D(\mathbb{P}_Z \| \mathbb{Q}_Z),$$

and uses the monotonicity of relative-entropy-type distances under coarse-graining maps and the chain-rule decomposition to obtain the endpoint–path relationship: the endpoint-pair marginal divergence does not exceed the full-trajectory gap, and there exists a nonnegative gap term such that

$$\Delta_{\text{path}} = \Delta_{\text{I}}^{(\text{pair})} + \Gamma_{\text{path}}, \qquad \Gamma_{\text{path}} \geq 0.$$

At the computable level, on the domain where Condition G_{geo} applies, this paper introduces a geometric proxy

$$\Gamma_{\text{geo}} = \Delta_{\text{II}} - d_I, \qquad d_I = \text{dist}_G\big(g(\gamma_0), g(\gamma_1)\big),$$

and in the implementations of Secs. 6–7 takes $d_I = \sqrt{2\Delta_{\mathrm{I}}}$ (an identity rather than an approximation), thereby elevating "whether endpoint compression suffices to represent full-trajectory accumulation" to an estimable and falsifiable structural statement. It should be emphasized that $\Gamma \approx 0$ only means *under the given viewpoint and resolution the endpoint description is sufficiently informative (the gap is compressible)*, and does not correspond to the disappearance of an underlying gap.

On the physical alignment side, this paper provides two complementary examples in a threshold-free window in one-loop QED, thereby completing a precise calibration of the framework and a minimal nontrivial test. First, under a one-dimensional truncation and a stable interface resolution, it obtains the baseline behavior $\Gamma_{\mathrm{geo}} = 0$, demonstrating consistency between endpoints and the full trajectory at that viewpoint resolution. Second, within the same window, after extending to a two-dimensional truncation that includes an EFT irrelevant direction, one generally obtains a nontrivial gap $\Gamma_{\mathrm{geo}} > 0$, demonstrating the mechanism by which, when intermediate-scale structure is readable under the given interface, endpoint compression systematically underestimates full-trajectory accumulation. Accordingly, the contribution is to provide an operational, viewpoint-locked gap diagnostic together with (i) an invariant meaning and a sharp zero-gap criterion for the geometric proxy (Theorems 5.1–5.2); and (ii) closed-form and asymptotic calibrations in standard QED/EFT windows (Secs. 6–7 and Appendix B).

Future work can proceed along three directions. First, replace the minimal Gaussian implementation with more specific physical interfaces (e.g. real scattering distributions, lattice observables, or constructible process states) to obtain direct estimates or bounds for the strict gap Γ_{path}. Second, incorporate threshold-crossing stitching and multi-coupling flows into a unified workflow, so that the gap term becomes a diagnostic tool for "piecewise-structure incompressibility," and promote the threshold rule from a technical detail to a reportable and testable structural input. Third, under higher-order corrections or richer EFT bases, systematically study scaling laws of Γ as functions of window length and resolution, thereby advancing the gap diagnostic into a comparable ruler and providing an operational boundary characterization for the "validity domain of endpoint descriptions" in cross-scale modeling.

Appendix A: Relative-Entropy Chain Rule, DPI, $\Gamma_{\mathrm{path}} \geq 0$, and the Equality Condition

This appendix states two structural facts used in the main text: (i) the data processing inequality (DPI), and (ii) the chain-rule decomposition (chain rule) for relative-entropy-type divergences (KL-type). Together they imply the conclusions used in Sec. 5 and Sec. 8: the strict gap term is nonnegative ($\Gamma_{\mathrm{path}} \geq 0$) and the corresponding equality condition. All statements are made under the fixed viewpoint locked in Sec. 2,

$$\mathsf{V} = (\mathcal{O}, \text{protocol}, R, S, \text{threshold rule}, \mathcal{M}, \pi, D, \mathrm{Agg}),$$

and viewpoint superscripts are omitted for brevity.

A.1 Data processing inequality (DPI)

Let X, Y be two predictive objects on the same space (in the classical case, probability distributions; in the quantum case, density operators/states), and let Φ be a coarse-graining map (in the classical case, a Markov kernel/stochastic map; in the quantum case, a CPTP map). For a relative-entropy-type divergence $D(\cdot\|\cdot)$ (classical KL or quantum relative entropy), one has

$$D(X\|Y) \;\geq\; D\big(\Phi(X)\|\Phi(Y)\big). \tag{A.1}$$

Equation (A.1) is DPI: coarse-graining does not increase distinguishable information difference.

Endpoint–path upper bound (source of $\Delta_{\mathrm{I}}^{(\mathrm{pair})} \leq \Delta_{\mathrm{path}}$ in the main text) Let the joint path variable be (with X_k denoting the fixed-viewpoint protocol readout variable, or its sufficient statistic, at scale μ_k)

$$Z := (X_0, X_1, \ldots, X_N),$$

and fix the endpoint projection (coarse-graining) map Π that retains only the two endpoints:

$$\Pi(Z) := (X_0, X_N) =: E.$$

Applying (A.1) to two path-process objects $\mathbb{P}_Z, \mathbb{Q}_Z$ with $\Phi = \Pi$ yields

$$D(\mathbb{P}_Z\|\mathbb{Q}_Z) \;\geq\; D(\mathbb{P}_E\|\mathbb{Q}_E). \tag{A.2}$$

In the notation of the main text, the left-hand side is the strict path divergence $\Delta_{\mathrm{path}}(0 \to N)$, and the right-hand side is the endpoint-pair marginal divergence $\Delta_{\mathrm{I}}^{(\mathrm{pair})}(0 \to N) := D(\mathbb{P}_E\|\mathbb{Q}_E)$, hence the endpoint–path bound $\Delta_{\mathrm{I}}^{(\mathrm{pair})}(0 \to N) \leq \Delta_{\mathrm{path}}(0 \to N)$.

A.2 Relative-entropy chain rule (classical form)

Decompose the same joint variable Z into endpoints and history:

$$E := (X_0, X_N), \qquad H := (X_1, \ldots, X_{N-1}), \qquad Z \equiv (E, H).$$

In the classical case, if the relative entropy is finite and conditional distributions are well-defined (e.g. $\mathbb{P}_{E,H} \ll \mathbb{Q}_{E,H}$), then the KL relative entropy satisfies the chain-rule decomposition

$$D(\mathbb{P}_{E,H}\|\mathbb{Q}_{E,H}) = D(\mathbb{P}_E\|\mathbb{Q}_E) + \mathbb{E}_{\mathbb{P}_E}\Big[D(\mathbb{P}_{H|E}\|\mathbb{Q}_{H|E})\Big]. \tag{A.3}$$

Accordingly, define the strict gap term by

$$\Gamma_{\mathrm{path}}(0 \to N) := \mathbb{E}_{\mathbb{P}_E}\Big[D(\mathbb{P}_{H|E}\|\mathbb{Q}_{H|E})\Big] \;\geq\; 0, \tag{A.4}$$

and obtain the exact identity

$$\Delta_{\mathrm{path}}(0 \to N) = \Delta_{\mathrm{I}}^{(\mathrm{pair})}(0 \to N) + \Gamma_{\mathrm{path}}(0 \to N), \tag{A.5}$$

where

$$\Delta_{\text{path}}(0 \to N) := D(\mathbb{P}_{E,H} \| \mathbb{Q}_{E,H}) = D(\mathbb{P}_Z \| \mathbb{Q}_Z), \qquad \Delta_{\mathrm{I}}^{(\text{pair})}(0 \to N) := D(\mathbb{P}_E \| \mathbb{Q}_E).$$

The nonnegativity in (A.4) follows from the nonnegativity of relative entropy and the fact that expectation preserves nonnegativity.

How the quantum case is used (structural note; no additional technical burden in the main text) The main text allows quantum predictive objects. In that case, the "interface + protocol" in the viewpoint V specifies a fixed readable statistical object (e.g. a measurement-outcome distribution or its sufficient statistics), so that at the interface level one can reduce to classical distributions and apply the same structural chain (A.1)–(A.5). The main text uses Γ_{path} only through this structural meaning (whether endpoint compression discards distinguishable history information) and does not require additional technical details of the quantum chain rule.

A.3 Equality condition: $\Gamma_{\text{path}} = 0$ and endpoint sufficiency

From (A.4), one immediately has

$$\Gamma_{\text{path}}(0 \to N) = 0 \quad \Longleftrightarrow \quad D(\mathbb{P}_{H|E} \| \mathbb{Q}_{H|E}) = 0 \ \text{ holds in the } \mathbb{P}_E\text{-almost-sure sense.} \tag{A.6}$$

In the classical case, by the uniqueness of the zero point of $D(\cdot\|\cdot)$, (A.6) is equivalent to

$$\mathbb{P}_{H|E} = \mathbb{Q}_{H|E} \quad (\mathbb{P}_E\text{-a.s.}). \tag{A.7}$$

This means that, conditional on the endpoints $E = (X_0, X_N)$, the two path processes are indistinguishable on the history component H. Therefore, under the fixed viewpoint V, the endpoint variable E is sufficient for distinguishing the two scale flows: the history H provides no additional distinguishable information.

Equality condition for DPI (recoverability statement; not required) For DPI, equality typically corresponds to the endpoint coarse-graining being a (recoverable) sufficient compression on the relevant object family: there exists a recovery map $\mathcal{R}$ such that $(\mathcal{R} \circ \Pi)$ recovers (or approximately recovers) the joint object on the relevant family, hence DPI saturates. The structural judgement in the main text only needs the semantics of "endpoint sufficiency / recoverability" and does not rely on further technical characterization of such recovery maps.

Appendix B: Closed Forms in 1D/2D for the Gaussian Interface / Fisher Metric (Computational Details for Sec. 6 and Sec. 7)

Under the fixed viewpoint V locked in the main text, this appendix records closed-form or directly computable expressions for the Gaussian interface (equivalently, a Fisher realization) used in Sec. 6 and Sec. 7: the endpoint divergence Δ_{I}, the induced metric G, the information speed Ψ, the path length Δ_{II}, and the geometric gap Γ_{geo}. Viewpoint superscripts are omitted throughout.

B.1 One-dimensional interface: $Y \mid \theta \sim \mathcal{N}(\theta, \sigma^2)$

Assume the observation protocol outputs a scalar statistic Y at a fixed resolution, modeled under parameter θ by a Gaussian family $P_\theta = \mathcal{N}(\theta, \sigma^2)$, where $\sigma > 0$ is the resolution parameter locked by the protocol (and hence part of the viewpoint terms).

Endpoint KL (Route I) For two endpoints θ_0, θ_1,

$$D(P_{\theta_0} \| P_{\theta_1}) = \frac{(\theta_0 - \theta_1)^2}{2\sigma^2}. \tag{B.1}$$

Hence, the endpoint divergence (single-interface case) is

$$\Delta_{\mathrm{I}}(\theta_0 \to \theta_1) = \frac{(\theta_0 - \theta_1)^2}{2\sigma^2}. \tag{B.2}$$

Induced metric (Fisher) and line element From the second-order expansion of the equal-covariance Gaussian family (equivalently, Fisher information), the induced metric is constant:

$$G(\theta) = \frac{1}{\sigma^2}, \qquad \mathrm{d}s^2 = \frac{\mathrm{d}\theta^2}{\sigma^2}. \tag{B.3}$$

Information speed and path length (Route II) Let the scale flow be parametrized by a scale parameter t (in the main text, one may take $t = \gamma$ or $t = \ell$). Then

$$\Psi(t) = \frac{\mathrm{d}s}{\mathrm{d}t} = \frac{1}{\sigma}\left|\frac{\mathrm{d}\theta}{\mathrm{d}t}\right|, \qquad \Delta_{\mathrm{II}}(t_0 \to t_1) = \int_{t_0}^{t_1} \Psi(t)\,\mathrm{d}t = \frac{1}{\sigma}\int_{t_0}^{t_1}\left|\frac{\mathrm{d}\theta}{\mathrm{d}t}\right|\mathrm{d}t. \tag{B.4}$$

Endpoint distance and geometric gap The endpoint-equivalent distance (as defined in the main text) is

$$d_I(\theta_0 \to \theta_1) = \sqrt{2\Delta_{\mathrm{I}}} = \frac{|\theta_1 - \theta_0|}{\sigma}. \tag{B.5}$$

Therefore,

$$\Gamma_{\mathrm{geo}} = \Delta_{\mathrm{II}} - d_I = \frac{1}{\sigma}\left(\int_{t_0}^{t_1}\left|\frac{\mathrm{d}\theta}{\mathrm{d}t}\right|\mathrm{d}t - |\theta_1 - \theta_0|\right) \ \geq \ 0, \tag{B.6}$$

with equality if and only if $\theta(t)$ does not "turn back" on the window (i.e., no sign reversal of $\mathrm{d}\theta/\mathrm{d}t$, equivalently monotonicity on the window). In Sec. 6, one-loop QED on a threshold-free window gives a linear monotone drift of θ, hence $\Gamma_{\mathrm{geo}} = 0$ in that baseline case.

B.2 Two-dimensional interface: independent Gaussians $(Y_\theta, Y_u) \mid (\theta, u)$ (the minimal implementation locked in the main text)

Assume the observation protocol outputs a two-dimensional statistic (Y_θ, Y_u) satisfying

$$Y_\theta \mid \theta \sim \mathcal{N}(\theta, \sigma_\theta^2), \qquad Y_u \mid u \sim \mathcal{N}(u, \sigma_u^2), \qquad \text{and independent given } (\theta, u), \tag{B.7}$$

where $\sigma_\theta, \sigma_u > 0$ are resolution parameters locked by the protocol (and hence part of the viewpoint terms).

Endpoint KL and endpoint distance (Route I) For endpoint differences $\Delta\theta = \theta_1 - \theta_0$ and $\Delta u = u_1 - u_0$, the KL divergence for equal-covariance Gaussians gives

$$\Delta_{\mathrm{I}} = \frac{1}{2}\left(\frac{(\Delta\theta)^2}{\sigma_\theta^2} + \frac{(\Delta u)^2}{\sigma_u^2}\right), \qquad d_I = \sqrt{2\Delta_{\mathrm{I}}} = \sqrt{\frac{(\Delta\theta)^2}{\sigma_\theta^2} + \frac{(\Delta u)^2}{\sigma_u^2}}. \tag{B.8}$$

Induced metric and line element The Fisher information yields a constant diagonal metric

$$G = \mathrm{diag}\left(\sigma_\theta^{-2}, \sigma_u^{-2}\right), \qquad \mathrm{d}s^2 = \frac{\mathrm{d}\theta^2}{\sigma_\theta^2} + \frac{\mathrm{d}u^2}{\sigma_u^2}. \tag{B.9}$$

Information speed and path length (Route II) Parametrize the scale flow by the coarse-graining parameter ℓ (dots denote $\mathrm{d}/\mathrm{d}\ell$). Then

$$\Psi(\ell) = \sqrt{\frac{\dot\theta(\ell)^2}{\sigma_\theta^2} + \frac{\dot u(\ell)^2}{\sigma_u^2}}, \qquad \Delta_{\mathrm{II}}(0 \to T) = \int_0^T \Psi(\ell)\,\mathrm{d}\ell. \tag{B.10}$$

B.3 Minimal 2D flow in Sec. 7: closed forms for Δ_{II} and Γ_{geo}

For the minimal scale flow used in Sec. 7,

$$\dot\theta = b, \qquad \dot u = -p\,u, \qquad u(\ell) = u_0 e^{-p\ell}, \qquad b > 0,\ p > 0,\ \ell \in [0, T], \tag{B.11}$$

equation (B.10) gives the information speed

$$\Psi(\ell) = \sqrt{\frac{b^2}{\sigma_\theta^2} + \frac{p^2 u_0^2 e^{-2p\ell}}{\sigma_u^2}}. \tag{B.12}$$

Hence the path length is

$$\Delta_{\mathrm{II}}(0 \to T) = \int_0^T \sqrt{\frac{b^2}{\sigma_\theta^2} + \frac{p^2 u_0^2 e^{-2p\ell}}{\sigma_u^2}}\,\mathrm{d}\ell. \tag{B.13}$$

Closed form (identity) Define

$$a := \frac{b}{\sigma_\theta} > 0, \qquad c := \frac{p\,u_0}{\sigma_u} \in \mathbb{R}, \tag{B.14}$$

then (B.13) integrates to

$$\Delta_{\mathrm{II}}(0 \to T) = aT + \frac{1}{p}\left[\sqrt{a^2 + c^2} - \sqrt{a^2 + c^2 e^{-2pT}} + a \ln \frac{a + \sqrt{a^2 + c^2 e^{-2pT}}}{a + \sqrt{a^2 + c^2}}\right]. \tag{B.15}$$

This expression degenerates continuously to $\Delta_{\mathrm{II}} = aT$ at $c = 0$ (consistent with the 1D baseline), with no need for case splitting.

Endpoint distance and geometric gap (closed form) From (B.11), the endpoint differences are

$$\Delta\theta = bT, \qquad \Delta u = u_0(e^{-pT} - 1).$$

Substituting into (B.8) gives the endpoint distance

$$d_I(T) = \sqrt{\left(\frac{bT}{\sigma_\theta}\right)^2 + \left(\frac{u_0}{\sigma_u}(1 - e^{-pT})\right)^2}. \tag{B.16}$$

Therefore,

$$\Gamma_{\mathrm{geo}}(T) = \Delta_{\mathrm{II}}(0 \to T) - d_I(T), \tag{B.17}$$

where $\Delta_{\mathrm{II}}(0 \to T)$ is given by (B.15) and $d_I(T)$ is given by (B.16).

Nonnegativity and equality condition (locked) Under the constant metric $G = \mathrm{diag}(\sigma_\theta^{-2}, \sigma_u^{-2})$, Δ_{II} is the arc length of the trajectory and d_I is the endpoint geodesic distance. Hence for any $T \geq 0$,

$$\Gamma_{\mathrm{geo}}(T) \geq 0. \tag{B.18}$$

Moreover, equality holds if and only if the trajectory is a straight line without turn-back in the scaled coordinates

$$x(\ell) := \theta(\ell)/\sigma_\theta, \qquad y(\ell) := u(\ell)/\sigma_u.$$

For the exponentially decaying trajectory in (B.11), for $T > 0$ this straight-line condition holds only in the degenerate cases:

$$\Gamma_{\mathrm{geo}}(T) = 0 \iff (u_0 = 0) \text{ or } (p = 0). \tag{B.19}$$

Hence under the nontrivial setting $u_0 \neq 0,\ p > 0$ in the main text, one has $\Gamma_{\mathrm{geo}}(T) > 0$ for any $T > 0$.

Appendix C: Choice of EFT Direction and the Dimensionless Convention ($p = d - 4$ and the Conditions for Neglecting Mixing Terms)

This appendix explains the origin of the dimensionless normalization and the scaling exponent p for the irrelevant direction u used in Sec. 7, and clarifies the scope and testable conditions under which one may neglect mixing terms with the principal direction (e.g. α or the equivalent coordinate $\theta = \alpha^{-1}$). All statements are made under the fixed viewpoint V (including the scale-coordinate convention) locked in the main text; viewpoint superscripts are omitted for brevity.

C.1 Dimensionless normalization and the canonical scaling exponent

Consider a local operator $\mathcal{O}_d$ in a four-dimensional EFT, with engineering dimension $d > 4$. Suppose the effective action in some basis contains the term

$$S_{\mathrm{EFT}} \supset \int d^4x \, C_d(\mu) \, \mathcal{O}_d(x).$$

Since $[d^4x] = -4$ and $[\mathcal{O}_d] = d$, the coefficient must satisfy

$$[C_d] = 4 - d$$

so that the action is dimensionless. Hence define the dimensionless Wilson coordinate

$$u(\mu) := \mu^{\,d-4} C_d(\mu), \qquad p := d - 4 > 0. \tag{C.1}$$

In Sec. 7, the coarse-graining parameter is

$$\ell := \ln(\mu_0/\mu), \qquad \ell \uparrow \iff \mu \downarrow,$$

so that increasing ℓ corresponds to UV→IR coarse-graining. In the minimal approximation that retains only canonical scaling and neglects anomalous dimensions and basis mixing, the irrelevant direction obeys

$$\frac{du}{d\ell} = -p\,u, \qquad \Rightarrow \qquad u(\ell) = u_0\, e^{-p\ell}, \tag{C.2}$$

which encodes the Wilsonian structure that irrelevant directions are suppressed in the IR when $p > 0$.

C.2 A more general form with mixing and a testable condition for "neglecting mixing terms"

More generally, in a fixed operator basis an irrelevant direction typically receives anomalous-dimension corrections and mixes with other directions. In the semantics of a truncation of dimension ≥ 2 (principal direction $\theta = \alpha^{-1}$ together with an EFT family $\{u, u_j\}$), the scale flow of u can be written as

$$\frac{du}{d\ell} = -\big(p + \gamma_u(\alpha)\big)\, u + r(\alpha, \theta) + \sum_{j \neq u} m_{uj}(\alpha)\, u_j + \cdots, \tag{C.3}$$

where $\gamma_u(\alpha)$ denotes the anomalous-dimension correction, $m_{uj}(\alpha)$ denotes linear mixing with other EFT directions u_j, and $r(\alpha, \theta)$ denotes a possible inhomogeneous source term (e.g. induced by matching/redefinitions). Here u_j is of the same type as u: dimensionless Wilson coordinates. The main text does not expand their physical meaning, and uses them only as placeholders for "additional structure directions".

The meaning of using the minimal flow (C.2) in Sec. 7 is that, on the chosen window and at the chosen accuracy, all terms in (C.3) other than $-p\,u$ contribute to $du/d\ell$ at a controllably negligible level. To state this as a reproducible condition, collect the neglected part in (C.3) into a residual term

$$\mathcal{R}_u(\ell) := -\gamma_u(\alpha(\ell))\, u(\ell) + r(\alpha(\ell), \theta(\ell)) + \sum_{j \neq u} m_{uj}(\alpha(\ell))\, u_j(\ell) + \cdots.$$

A sufficient, "hard" criterion that does not degenerate as $u(\ell) \to 0$ is: on the window $\ell \in [0, T]$

there exists a small parameter $\varepsilon \ll 1$ such that

$$\sup_{\ell\in[0,T]} \frac{|\mathcal{R}_u(\ell)|}{p\,|u(\ell)| + \sigma_u} \;\le\; \varepsilon. \tag{C.4}$$

Here $\sigma_u > 0$ is the interface-resolution parameter in Sec. 7 (see (7.7)); it provides a normalization floor in the indistinguishable regime $|u(\ell)| \lesssim \sigma_u$, avoiding denominator degeneracy that would occur with a purely relative-error criterion. This term is fully fixed by the "protocol/resolution" clauses of the viewpoint V, and does not introduce new physical degrees of freedom.

It should be emphasized that even if (C.4) fails, the structural conclusion in Sec. 7 that "a generic 2D trajectory is non-geodesic $\Rightarrow \Gamma_{\rm geo} > 0$" typically still holds. What changes is the specific functional form and numerical magnitude of Ψ and $\Delta_{\rm II}$, not the mechanism that a non-affine / non-straight scale process in dimension ≥ 2 triggers a gap. In that case, one should treat the actual flow in (C.3) as part of the viewpoint clauses and recompute $\Psi, \Delta_{\rm II}, \Gamma_{\rm geo}$ accordingly.

C.3 Physical options: Euler–Heisenberg and four-fermion contact terms (illustrative)

To make the irrelevant direction u in Sec. 7 more physically explicit, we list two common EFT directions as optional illustrations. They are not required for the derivations; their role is to instantiate the "additional structure direction at intermediate scales" as standard EFT coefficients, so that the 2D trajectory and the mechanism for $\Gamma_{\rm geo} > 0$ do not rely on an abstract parameter.

Option 1: Euler–Heisenberg-type F^4 direction ($d = 8$) In a low-energy effective description of QED, one may take a representative $d = 8$ F^4-type operator, e.g.

$$\mathcal{O}_8 \sim (F_{\mu\nu}F^{\mu\nu})^2 \quad \text{or an equivalent gauge-invariant basis.}$$

If its Wilson coefficient is denoted by $C_8(\mu)$, then the dimensionless coordinate is

$$u(\mu) = \mu^4 C_8(\mu), \qquad p = d - 4 = 4.$$

The advantage is that it connects directly to QED effective nonlinear responses (e.g. vacuum nonlinearity / light–light scattering), making it straightforward to express "interface sensitivity to u" in terms of concrete observable combinations.

Option 2: four-fermion contact term ($d = 6$) In a more general effective theory, four-fermion operators are among the most common irrelevant directions, e.g.

$$\mathcal{O}_6 \sim (\bar{\psi}\gamma_\mu\psi)(\bar{\psi}\gamma^\mu\psi) \quad \text{or a problem-matched four-fermion basis.}$$

If the coefficient is $C_6(\mu)$, then

$$u(\mu) = \mu^2 C_6(\mu), \qquad p = d - 4 = 2.$$

Even if it is not dominant in some pure-QED windows, it still provides a clear example of an "irrelevant but readable" direction, enabling a reproducible numerical contrast for the nonzero behavior of Γ.

Synchronized locking of viewpoint clauses The common feature of these options is that they instantiate the "extra structure direction" as explicit EFT coefficients. The paper allows replacing the concrete meaning of u across different physical problems; once replaced, it must be synchronized and locked in the viewpoint V, including (i) the statistics (or sufficient statistics) through which the interface reads u, and (ii) the error model/covariance (thereby locking the induced metric G). Otherwise, the magnitude of Γ becomes a cross-viewpoint output and loses reproducibility.

Appendix D: Operational Workflow Pseudocode (Input–Output Interface Style)

This appendix provides a reproducible "input–output" pseudocode framework. Its purpose is not to constrain implementation details, but to pin down a checkable workflow by explicitly listing the viewpoint elements V, the Route I/II computation chain, and the diagnostic procedure for the geometric-proxy gap Γ_{geo}. Unless stated otherwise, viewpoint superscripts are omitted.

D.1 Input specification (must be locked and reported)

(I) Scale window and discrete grid

- Window: $[\mu_0, \mu_1]$;
- Grid: $\{\mu_k\}_{k=0}^{N}$ (strictly monotone), together with a locked choice of scale coordinate.

The main text allows two equivalent coordinates; you must choose *exactly one* and keep it consistent throughout the full pipeline:

- Choose $\gamma := \ln(\mu/\mu_0)$: then

$$\gamma_k := \ln(\mu_k/\mu_0), \qquad \Delta\gamma_k := \gamma_{k+1} - \gamma_k > 0.$$

- Choose $\ell := -\gamma = \ln(\mu_0/\mu)$: then

$$\ell_k := \ln(\mu_0/\mu_k), \qquad \Delta\ell_k := \ell_{k+1} - \ell_k > 0.$$

In the pseudocode below we use a unified symbol t for the locked scale coordinate (either $t = \gamma$ or $t = \ell$), and write $t_k,\ \Delta t_k$.

(II) Viewpoint object V (full tuple, including the aggregation operator) The viewpoint must be locked and reported as the full tuple

$$\mathsf{V} = (\mathcal{O},\ \text{protocol},\ R,\ S,\ \text{threshold rule},\ \mathcal{M},\ \pi,\ D,\ \text{Agg}).$$

Here Agg is the interface-aggregation operator (e.g. sup or a fixed weighted aggregation). It must be treated as a component of V and kept fixed throughout the paper.

(III) Interface outputs / predictive objects For each scale point k and each interface $O \in \mathcal{O}$, obtain or generate the interface output / predictive object

$$X_k^{(O)} \equiv X_{\mu_k}^{(O)}.$$

The concrete form of $X_k^{(O)}$ may be a distribution, a quantum state, or a sufficient-statistics representation under the protocol; however, it must be consistent throughout the pipeline and consistent with the "protocol/regularization/scheme/model class/reconstruction" clauses in V.

(IV) Parameterization and the applicability domain for the geometric proxy (Condition $\mathrm{G}_{\mathrm{geo}}$) If you intend to compute the geometric-proxy gap Γ_{geo}, the input must explicitly lock:

- Parameter coordinates: $g(t) \in \mathbb{R}^d$ (with d the model-class dimension, which must be fixed);
- A positive-definite metric: $G(g)$ (it may be constant diagonal or g-dependent, but must be positive definite);
- The endpoint-distance computation rule: $d_I = \mathrm{dist}_G\big(g(t_0), g(t_N)\big)$ (the algorithm or closed form must be fixed).

Emphasis: *only* in the equal-covariance Gaussian / Fisher constant-metric implementation used in Secs. 6–7 may one further use the identity $d_I = \sqrt{2\Delta_{\mathrm{I}}}$; under general Condition $\mathrm{G}_{\mathrm{geo}}$ one must not treat $\sqrt{2\Delta_{\mathrm{I}}}$ as the definition of d_I.

(V) Local-metric estimator and flow/velocity estimator (must be fixed)

- Metric estimator: specify the method for estimating or defining $G(g)$ on g (e.g. Fisher information, inverse covariance, etc.), and record it as part of V;
- Flow/velocity estimator: specify how to compute $\dot{g}(t) := \frac{dg}{dt}$ (either from a theoretical β function or from finite-difference fitting), and record it in the protocol clause.

D.2 Pseudocode: Route I / Route II / Γ_{geo}

Implementation clause (locked numerical conventions for computable and comparable Γ_{geo}). Whenever $\Gamma_{\mathrm{geo}} := \Delta_{II} - d_I$ is used for diagnostics or hypothesis testing, the viewpoint specification V must be augmented by an explicit set of numerical conventions. Once fixed, these conventions are treated as part of the locked viewpoint and must be reported; otherwise the resulting scalar(s) must be classified as cross-viewpoint outputs and are not comparable as the same quantity.

1. **Divergence computability (mandatory).** The protocol/output representation must make $D(\cdot\|\cdot)$ finite on all evaluated endpoint pairs. For KL/relative-entropy-type divergences, the support convention, normalization, log base, and any ε-smoothing / truncation rule must be locked and reported.

2. **SPD enforcement for the metric (mandatory when G is estimated).** If $G(g)$ is obtained from an estimator rather than a closed form, an explicit SPD-enforcement map (e.g. symmetrization followed by eigenvalue flooring / regularization with a fixed $\varepsilon_{\mathrm{SPD}}$) must be locked in V so that all square-root expressions are well-defined.

3. **Discrete rule for Δ_{II} (mandatory in computations).** On a discrete grid, Δ_{II} must be computed by a fully specified quadrature/midpoint rule locked in V. A standard admissible form is the discrete geometric length

$$\widehat{\Delta}_{II} := \sum_{k=0}^{N-1} \sqrt{\Delta g_k^\top G_{\mathrm{mid},k}\, \Delta g_k}, \qquad \Delta g_k := g_{k+1} - g_k,$$

with a locked midpoint rule $G_{\mathrm{mid},k}$ (e.g. arithmetic midpoint).

4. **Endpoint distance d_I under the same metric (mandatory).** In general, $d_I = \mathrm{dist}_G(g(\gamma_0), g(\gamma_1))$ denotes the *global* geodesic distance induced by the same G, and it must *not* be approximated by reusing the observed discrete path $\{g_k\}$ as the minimizing path (neither directly nor via a graph restricted to $\{g_k\}$). Any numerical approximation method (discretization level, initialization/multi-start schedule, optimizer, stopping rule, tolerances, and any randomness/seed) must be locked and reported as part of V. In the special equal-covariance Gaussian/Fisher constant-metric implementation used in Secs. 6–7, dist_G reduces to a scaled Euclidean distance and $d_I = \sqrt{2\Delta_I}$ holds identically.

5. **No clipping; numerical-validity gate (mandatory).** One must not clip negative $\widehat{\Gamma}_{\mathrm{geo}}$ to zero. If $\widehat{\Gamma}_{\mathrm{geo}} < 0$ (or close to 0 within tolerance), a locked numerical-validity check must be applied (grid refinement for $\widehat{\Delta}_{II}$ and solver tightening/refinement for $\widehat{d_I}$). The refinement rule and acceptance tolerances (e.g. $\tau_\gamma, \tau_{\mathrm{refine}}$) must be locked and reported; if the check fails, the result must be reported as numerically unresolved under the locked tolerances.

6. **Reference workflow.** Appendix D provides a standardized operational pseudocode that instantiates the above clauses; any alternative implementation is admissible only if it makes the same conventions equally explicit and locks them into V.

```
# PSEUDOCODE (LOCKED-VIEWPOINT, ATTACK-RESISTANT VERSION)

Input (ALL items below are part of the locked viewpoint Vw and must be reported
verbatim):
  window [muScale_0, muScale_1]
  grid {muScale_k}_{k=0..N} (strictly monotone; N fixed)
  scale coordinate choice (record only): t = gscale or t = ell
  viewpoint Vw = (
    O_set, protocol, Reg, Sch, threshold_rule, model_class M,
    Rec, output_representation_map, support_sets {S_O},
    divergence_spec (D, eps_D, base, normalization),
    aggregator_spec (Agg, weights, order),
    metric_spec (G_rule, estimator_params, SPD_projection_rule),
    path_length_spec (midpoint_rule, quadrature_rule),
    geodesic_distance_spec (DistAlg, discretization_L, multi_start_schedule,
```

```
                             optimizer, tolerances, max_iters, seed),
    numerical_tolerances (tau_SPD, tau_dist, tau_gamma, tau_refine),
    randomness (seed; must be fixed)
  )
  interface outputs/predictions X_k^O for each k and O (in the representation fixed
  above)

Hard conventions (non-negotiable; violations invalidate T1/T2 usage):
  (C1) D must be finite on all encountered pairs (X_0^O, X_N^O). If D is KL/relative
  entropy,
       enforce eps-smoothing on the fixed support S_O:
         normalize: p <- p / sum_{i in S_O} p_i, q <- q / sum_{i in S_O} q_i
         smooth:    p~ <- (p + eps_D) / (1 + eps_D*|S_O|),  q~ <- (q + eps_D) / (1 +
         eps_D*|S_O|)
         D(p||q) := sum_{i in S_O} p~_i * log( p~_i / q~_i )   (log base fixed in
         Vw)
       If D is not KL, an equally explicit finiteness-preserving convention must be
       fixed in Vw.
  (C2) Any estimator-produced metric matrix must be symmetrized and projected to SPD
       with an explicit rule in Vw (below is the default rule).
  (C3) dist_G MUST represent the global geodesic distance under the induced metric
  G; it MUST NOT
       reuse the observed discrete path {g_hat[k]} as the minimization path (neither
       as graph edges,
       nor as a restricted candidate family). Only endpoints may be used as boundary
       conditions.
  (C4) Gamma_geo is not allowed to be "clipped" to 0. Negative values are treated as
  numerical
       error and must trigger refinement checks under a fixed criterion in Vw (see
       Step 9).

Precompute:
  Record t_k and Deltat_k according to the chosen coordinate (gscale or ell).
  NOTE: Route II is implemented in a parameterization-robust discrete geometric form
  (Step 5),
        so Deltat_k does not affect Delta_II numerically.

# Step 1: Reconstruction / parameterization on the model class (locked Rec, locked
protocol)
for k = 0..N:
    g_hat[k] = Rec( Vw, {X_k^O}_{O in O_set} )          # g_hat[k] in R^d, d fixed
    by M

# Step 2: Endpoint fissure (Route I) with finite divergence convention (locked D,
locked eps_D)
Delta_I = Agg_{O in O_set}  D( X_0^O || X_N^O )         # Agg and any weights are
fixed in Vw

# Step 3: Induced local metric along the path (Condition G_geo) with mandatory SPD
projection
```

```
# Default SPD projection rule (must be used unless Vw specifies an alternative
explicit rule):
#   A <- G_rule(g_hat[k]) (closed form or estimator output)
#   A <- 0.5*(A + A^T)
#   eigendecompose A = Q diag(lambda) Q^T
#   lambda_i <- max(lambda_i, eps_SPD), where eps_SPD is fixed in Vw (e.g., eps_SPD
= tau_SPD * trace(A)/d)
#   G_hat[k] <- Q diag(lambda) Q^T
for k = 0..N:
    A = G_rule( g_hat[k] )
    A = 0.5 * ( A + A^T )
    G_hat[k] = SPD_Project( A; eps_SPD )                # eps_SPD and projector are
    fixed in Vw

# Step 4: Optional velocity record (not used to compute Delta_II, only for
diagnostics)
for k = 0..N-1:
    gdot_hat[k] = ( g_hat[k+1] - g_hat[k] ) / Deltat_k  # recorded; rule locked in
    Vw

# Step 5: Information path length (Route II) in discrete geometric form
(parameterization-robust)
# Delta_II := sum_k sqrt( (Δg_k)^T G_mid(k) (Δg_k) )
# where Δg_k = g_hat[k+1]-g_hat[k], and G_mid is fixed (default: arithmetic
midpoint).
Delta_II = 0
for k = 0..N-1:
    G_mid  = 0.5 * ( G_hat[k] + G_hat[k+1] )            # fixed midpoint rule in Vw
    Delta_g = g_hat[k+1] - g_hat[k]
    Psi_ds  = sqrt( Delta_g^T * G_mid * Delta_g )        # always real due to SPD
    Delta_II += Psi_ds

# Step 6: Endpoint distance d_I under the SAME metric (Condition G_geo): GLOBAL
geodesic distance
# Default dist_G implementation (must be used unless Vw specifies an alternative
explicit algorithm):
#   Compute d_I = inf_{path x(s)} _0^1 sqrt( x'(s)^T G(x(s)) x'(s) ) ds, x(0)=g0,
x(1)=gN.
#   Practical solver (fixed in Vw):
#     - Discretize path with L internal knots: x_0=g0, x_L=gN, variables
x_1..x_{L-1}.
#     - Objective: J(x_1..x_{L-1}) = sum_{j=0..L-1} sqrt( (Δx_j)^T
G_mid(x_j,x_{j+1}) (Δx_j) ),
#       where Δx_j=x_{j+1}-x_j and G_mid uses the SAME midpoint rule as Step 5.
#     - Optimize J with a fixed optimizer (e.g., L-BFGS) + fixed tolerances and
max_iters.
#     - Multi-start: initialize with K fixed initializations using ONLY endpoints
(no use of {g_hat[k]}),
#       e.g. straight-line + K-1 deterministic perturbations; pick the minimum J.
#     - Output d_I = min_{starts} J*.
g0 = g_hat[0]
```

```
gN = g_hat[N]
d_I = dist_G_global( g0, gN; G_rule, SPD_Project, midpoint_rule,
                     L, K, optimizer, tau_dist, max_iters, seed )

# Step 7: Gap (geometric proxy)
Gamma_geo = Delta_II - d_I

# Step 8: Uncertainty propagation (mandatory fields if uncertainty is reported)
# If uncertainty is computed, it must include contributions from:
#   (i) Rec/protocol estimation (g_hat),
#   (ii) metric estimation + SPD projection (G_hat),
#   (iii) discretization (N for Delta_II; L/K/optimizer tolerance for d_I),
#   (iv) any smoothing eps_D for D.
# Otherwise, explicitly state "uncertainty not computed" (do not omit silently).

# Step 9: Mandatory numerical validity checks (do NOT clip Gamma_geo)
# Goal: ensure that sign claims are not artifacts of discretization/solver slack.
# Fixed acceptance criterion (locked in Vw):
#   - Refinement: increase N -> N' (grid refinement) and L -> L' (geodesic
discretization refinement),
#     tighten tau_dist, recompute Gamma_geo'.
#   - Accept if BOTH:
#       (A) |Gamma_geo' - Gamma_geo| <= tau_refine
#       (B) Gamma_geo >= -tau_gamma  AND  Gamma_geo' >= -tau_gamma
#     Otherwise, declare "numerically unresolved at locked tolerances" and report
both values.
Validate_or_Refine( Vw, Gamma_geo, Delta_II, d_I )

Output (must include the full locked viewpoint specification Vw):
  Delta_I, Delta_II, d_I, Gamma_geo,
  numerical validity report (Step 9),
  uncertainty report (if computed),
  and the complete Vw checklist (all locked clauses + seeds + tolerances).
```

Implementation notes (hard constraints that do not introduce extra degrees of freedom)

- Agg must not be switched across different windows / different experiments;
- the model-class dimension d must not vary with scale within a single result chain;
- if $\dot{g}$ is obtained by finite differences, the differencing rule (forward/central/smoothing/fitting) must be written into the protocol and fixed;
- if Γ_{geo} is used, then the implementations of $g(t)$, $G(g)$, and dist_G must be written into the viewpoint clauses (Condition G_{geo});
- if the window crosses a threshold, one must first segment the window according to the locked threshold rule, then run the above workflow within each segment, and report the stitching/matching conditions.

D.3 Output and reporting requirements (minimum reproducibility package)

(O1) Numerical outputs (required) At minimum, report

$$\Delta_{\rm I}, \qquad \Delta_{\rm II}, \qquad d_I = \mathrm{dist}_G\big(g(t_0), g(t_N)\big), \qquad \Gamma_{\rm geo} = \Delta_{\rm II} - d_I,$$

together with (when available) error bars / confidence intervals and an explicit uncertainty-propagation description (or the bootstrap scheme and the number of resamples). If the main text uses estimator notation, report

$$\widehat{\Delta_{\rm I}}, \qquad \widehat{\Delta_{\rm II}}, \qquad \widehat{d}_I, \qquad \widehat{\Gamma_{\rm geo}},$$

and their interval estimates in a manner consistent with the locked rules in Sec. 8.3.

(O1′) Fracture-location output (mandatory when the condition applies) If a window-length scan $T \mapsto W(T)$ is performed and the main-text "gap detected" decision rule is used, then one must additionally report the first-detection scale $\widehat{T_\delta}$ and its uncertainty interval (bracketing interval or quantile interval), and simultaneously report: the parameterization of T, the generation rule for the window family $\{W(T)\}$, the discrete grid $\{T_j\}$ (or equivalently $\{\mu_{1,j}\}$), the significance level and multiple-testing correction rule (if used), and the interval-construction rule.

(O2) Viewpoint elements (required) Attach a complete list of viewpoint elements (the main-text checklist), at least including: the window/grid; $\mathcal{O}$+protocol; regularization handling R; scheme S; the threshold rule; model class $\mathcal{M}$; reconstruction π; divergence D; aggregation Agg; and the error model / covariance (including resolution parameters such as σ_u) and the significance level (if hypothesis-testing decisions are made). If $\Gamma_{\rm geo}$ is used, explicitly report the computation rules for $g(t)$, $G(g)$, and $d_I = \mathrm{dist}_G$ (Condition $\mathrm{G}_{\rm geo}$).

Appendix E: Window Dependence of the QED One-Loop Coefficient b and the Piecewise Rule for "Active Species" (Normalization Note)

This appendix provides no derivations. It only gives a standardized way of writing "how to choose a window / how to splice across thresholds" for Sec. 6, so that threshold handling is promoted from an implicit assumption to an explicit clause of the viewpoint V. Unless stated otherwise, viewpoint superscripts are omitted; the scale variable is written uniformly as μ.

E.1 Window-wise expression of the one-loop coefficient (piecewise by the active set)

In the QED one-loop approximation, the running of the electromagnetic coupling α can be written as

$$\frac{d\alpha}{d\ln\mu} = b(\mu)\,\alpha^2, \qquad b(\mu) = \frac{2}{3\pi} \sum_{f\in\mathrm{active}(\mu)} w_f\, Q_f^2. \tag{E.1}$$

Here active(μ) denotes the set of electrically charged Dirac fermions treated as "active" at scale μ (locked by the threshold rule in the viewpoint V); Q_f is the charge in units of e; and w_f is an optional multiplicity factor (e.g. color degeneracy). If not needed, one sets $w_f \equiv 1$ and *locks this choice* in the viewpoint clauses.

Introduce the coordinate consistent with the main text,

$$\theta(\mu) := \alpha^{-1}(\mu),$$

then (E.1) is equivalent to

$$\frac{d\theta}{d\ln\mu} = -b(\mu). \tag{E.2}$$

No-threshold window (constant within a segment) If, on the window $[\mu_0, \mu_1]$, the set active(μ) does not change with μ, then $b(\mu) \equiv b$ is a constant, and hence

$$\theta(\mu_1) - \theta(\mu_0) = -b\,\ln\frac{\mu_1}{\mu_0}. \tag{E.3}$$

Equation (E.3) is the "within-segment endpoint linear law" used in Sec. 6: in the θ coordinate, the one-loop flow is a constant-speed straight line.

Interface with the ℓ coordinate (if the main text uses ℓ) If the main text uses $\ell := \ln(\mu_0/\mu)$, then $d/d\ell = -\,d/d\ln\mu$, and (E.2) becomes

$$\frac{d\theta}{d\ell} = b(\mu(\ell)), \tag{E.4}$$

which yields the within-segment affine form $\theta(\ell) = \theta_0 + b\,\ell$ on a no-threshold segment. (The piecewise rules below are coordinate-independent; one only needs to lock the coordinate choice in V and keep it consistent throughout.)

E.2 Piecewise rule and matching clauses (threshold splicing)

When the scale window crosses mass thresholds, threshold handling must be written as an explicit clause of the viewpoint V. We adopt the following normalized writing. We do not claim any particular decoupling scheme is "optimal"; we only require that the clause be *locked* and *checkable*.

(T1) Threshold segmentation: partition the window into intervals with a fixed active set Given a collection of threshold scales $\{m_j\}_{j=1}^{J}$ (typically of the same order as particle masses), take the subset that falls in the open interval (μ_0, μ_1) and sort by scale:

$$\mu_0 < \tau_1 < \tau_2 < \cdots < \tau_{J'} < \mu_1, \qquad \{\tau_j\} = \{m_j\} \cap (\mu_0, \mu_1).$$

Define the segment endpoints by

$$\mu_{(0)} := \mu_0, \qquad \mu_{(j)} := \tau_j \ (j = 1, \ldots, J'), \qquad \mu_{(J'+1)} := \mu_1,$$

and the segments

$$I_j := [\mu_{(j-1)}, \mu_{(j)}], \qquad j = 1, \ldots, J' + 1.$$

The core requirement of the threshold rule is: on each segment I_j, the active set active(μ) is constant, so that

$$b(\mu) \equiv b_j \quad (\mu \in I_j), \qquad b_j = \frac{2}{3\pi} \sum_{f \in \mathrm{active}_j} w_f Q_f^2, \tag{E.5}$$

and each segment must report active$_j$ (or equivalently report b_j together with the locked conventions for w_f, Q_f).

(T2) Matching point and matching rule: lock the change of degrees of freedom as a clause For adjacent segments I_j and I_{j+1}, choose a matching point near their interface,

$$\mu_j^{\mathrm{match}} \sim \tau_j,$$

and impose a matching rule at that point. Both the selection rule for μ_j^{match} and the matching form must be written (and locked) in the threshold clause of V, for example:

- Continuous matching (minimal convention): $\alpha_j(\mu_j^{\mathrm{match}}) = \alpha_{j+1}(\mu_j^{\mathrm{match}})$, equivalently $\theta_j(\mu_j^{\mathrm{match}}) = \theta_{j+1}(\mu_j^{\mathrm{match}})$;
- Matching with a correction: $\alpha_{j+1}(\mu_j^{\mathrm{match}}) = \alpha_j(\mu_j^{\mathrm{match}}) + \delta\alpha_j$, where the definition, perturbative order, and negligibility conditions of $\delta\alpha_j$ must be locked; equivalently, in the θ coordinate one may write

  $$\theta_{j+1}(\mu_j^{\mathrm{match}}) = \theta_j(\mu_j^{\mathrm{match}}) + \Delta\theta_j^{\mathrm{match}},$$

  where $\Delta\theta_j^{\mathrm{match}}$ is fully determined by the chosen matching rule.

The "no-threshold window" baseline used in the main text corresponds to $J' = 0$, so there are no matching points and no matching rule.

(T3) Within-segment endpoint relations and full-window splicing On each segment $I_j = [\mu_{(j-1)}, \mu_{(j)}]$, (E.2) and (E.5) give the within-segment endpoint relation

$$\theta(\mu_{(j)}) - \theta(\mu_{(j-1)}) = -b_j \ln \frac{\mu_{(j)}}{\mu_{(j-1)}}, \qquad j = 1, \ldots, J' + 1. \tag{E.6}$$

By connecting adjacent segments using the matching clause (T2), the full-window endpoint relation can be written in the normalized form "sum of segment contributions + matching corrections":

$$\theta(\mu_1) - \theta(\mu_0) = -\sum_{j=1}^{J'+1} b_j \ln \frac{\mu_{(j)}}{\mu_{(j-1)}} + \sum_{j=1}^{J'} \Delta\theta_j^{\mathrm{match}}. \tag{E.7}$$

Under continuous matching, $\Delta\theta_j^{\mathrm{match}} = 0$, and (E.7) reduces to a pure sum over piecewise-constant coefficients b_j.

Within-segment "constant-speed straight line" vs. cross-threshold "kinks" Equation (E.6) shows that within each segment, θ is affine in $\ln \mu$ (or in ℓ). Across thresholds, changes in b_j (and possible $\Delta\theta_j^{\mathrm{match}}$) produce kinks/jumps in the trajectory. Therefore, "constant-speed straight line in θ at one loop" should be understood as *piecewise* constant-speed straight lines, rather than a single global straight line on a cross-threshold window.

E.3 Relation to the gap Γ (thresholds as viewpoint input, not an implicit assumption)

Once segmentation and matching are introduced, additional structural inputs enter at the viewpoint level: the window is no longer described by a single constant b, but is determined jointly by

$$\{I_j\}, \quad \{\mathrm{active}_j\} \text{ (equivalently } \{b_j\}), \quad \{\mu_j^{\mathrm{match}}, \text{ matching rule}\}.$$

Structural meaning (kept strictly within scope)

- At the level of strict objects: if your "history variable / process object" treats the segment labels, changes of active sets, or matching corrections (equivalently: process information from the locked threshold rule) as distinguishable information, then endpoint compression may discard such process information, providing a structural source for $\Gamma_{\mathrm{path}} > 0$.
- At the level of the geometric proxy: if the geometric coordinate is chosen as one-dimensional $g(t) = \theta(t)$ and θ is monotone on the window, then in the minimal one-dimensional constant-metric implementation one may still obtain $\Gamma_{\mathrm{geo}} = 0$ (the degenerate case "arc length = endpoint distance"). To make threshold structure explicit in Γ_{geo}, the viewpoint clauses must encode the "threshold structure" into the readable coordinates themselves (e.g. by enlarging $g(t)$ to include additional readable directions/effective degrees of freedom, or by following the main-text T3 rule to report segment-wise gaps alongside the full-window result).

Reporting requirement (mandatory for cross-threshold windows) Therefore, if the main text reports Γ on a cross-threshold window (either the strict Γ_{path} or the geometric proxy Γ_{geo}), one must simultaneously report the three clause-level inputs above:

$$\{I_j\}, \quad \{\mathrm{active}_j\} \text{ (or equivalently } \{b_j\}), \quad \{\mu_j^{\mathrm{match}}, \text{ matching rule}\}.$$

Missing any of these items makes the numerical comparison of Γ degenerate into a cross-viewpoint output and hence non-checkable (consistent with the viewpoint-locking clauses in Sec. 2).

Appendix F: Log-spiral chart

This appendix provides an *optional* geometric representation: it embeds a scalar variable on the *positive real axis* into a two-dimensional plane (or the complex plane) via a logarithmic-spiral map. Emphasis: this embedding does *not* modify the fixed viewpoint V in the main text, and it does *not* change any definition of fracture/length quantities. It is only a *visual coordinate encoding* for trajectories and windows.

F.1 One-dimensional spiral embedding: definition, invertibility, and scale–phase coupling

Domain (locked) The spiral embedding in this appendix is defined on $x \in \mathbb{R}_{>0}$. If a reader wishes to use this representation for variables that can take negative values (e.g. a general real parameter), one must first choose a *monotone* positive-valued map (e.g. $x = \exp(\xi)$, or $x = \xi + c$ with $\xi > -c$), and state that map explicitly as an additional visualization convention; the main text neither adopts nor depends on such conventions.

Spiral embedding Fix a reference scale $x_{\mathrm{ref}} > 0$ and a spiral coupling constant $\kappa \in \mathbb{R}$. Define the map

$$\mathsf{S}_\kappa : \ \mathbb{R}_{>0} \to \mathbb{C}, \qquad \mathsf{S}_\kappa(x) := x \ \exp\bigl(i\kappa \ln(x/x_{\mathrm{ref}})\bigr). \tag{F.1}$$

In polar coordinates,

$$r = x, \qquad \theta = \kappa \ln(x/x_{\mathrm{ref}}), \tag{F.2}$$

so $\mathsf{S}_\kappa(\mathbb{R}_{>0})$ is a logarithmic (equiangular) spiral. The map is one-to-one on $\mathbb{R}_{>0}$: since $|\mathsf{S}_\kappa(x)| = x$, the endpoint scale can be recovered directly. Hence the embedding introduces no information loss at the endpoint level; it only encodes "scale change" simultaneously as radial and phase variation.

Scale–phase coupling identity For any $c > 0$,

$$\mathsf{S}_\kappa(cx) = c\, e^{i\kappa \ln c}\, \mathsf{S}_\kappa(x). \tag{F.3}$$

That is, a multiplicative rescaling $x \mapsto cx$ is encoded as "radial dilation by c + rotation by angle $\kappa \ln c$".

Euclidean arc length (for image geometry only) Equip the complex plane with the standard Euclidean metric $|dz|$ (note: this is *not* the Fisher/information metric of the main text). From

$$\frac{d}{dx}\mathsf{S}_\kappa(x) = \exp\bigl(i\kappa \ln(x/x_{\mathrm{ref}})\bigr)\,(1 + i\kappa) \tag{F.4}$$

we obtain

$$\bigl|d\mathsf{S}_\kappa(x)\bigr|^2 = |1 + i\kappa|^2\, dx^2 = (1 + \kappa^2)\, dx^2, \qquad \Rightarrow \qquad \mathrm{Len}_{\mathbb{C}}\bigl([x_0, x_1]\bigr) = \sqrt{1+\kappa^2}\, |x_1 - x_0|. \tag{F.5}$$

Thus, in the sense of *Euclidean arc length in the image*, this embedding is equivalent to a constant rescaling of the original real axis; its "nontriviality" comes entirely from the scale–phase coupling in (F.3), not from introducing any new metric structure used in the main text.

F.2 Multi-dimensional (componentwise) spiralization: visualization-only encoding

If $g = (g^1, \ldots, g^d)$ has positive components ($g \in (\mathbb{R}_{>0})^d$), fix parameter tuples

$$\kappa = (\kappa_1, \ldots, \kappa_d) \in \mathbb{R}^d, \qquad x_{\mathrm{ref}} = (x_{\mathrm{ref},1}, \ldots, x_{\mathrm{ref},d}) \in (\mathbb{R}_{>0})^d, \tag{F.6}$$

and define the componentwise spiral embedding

$$\mathsf{S}_\kappa(g) := \big(\mathsf{S}_{\kappa_1}(g^1), \dots, \mathsf{S}_{\kappa_d}(g^d)\big) \in \mathbb{C}^d. \tag{F.7}$$

This embedding does not preserve linear structure in general (except when $\kappa = 0$). Therefore it should be understood purely as a *chart / visualization encoding.* To promote it to a new inner-product/metric geometry, one would have to specify a new metric and write it into the main-text viewpoint V; this paper does not do so. All quantities in the main text remain defined by the locked G, D, and the associated $\Delta_{\mathrm{I}}, \Delta_{\mathrm{II}}, \Gamma_{\mathrm{geo}}$.

F.3 Direct interface with the scale variable (optional)

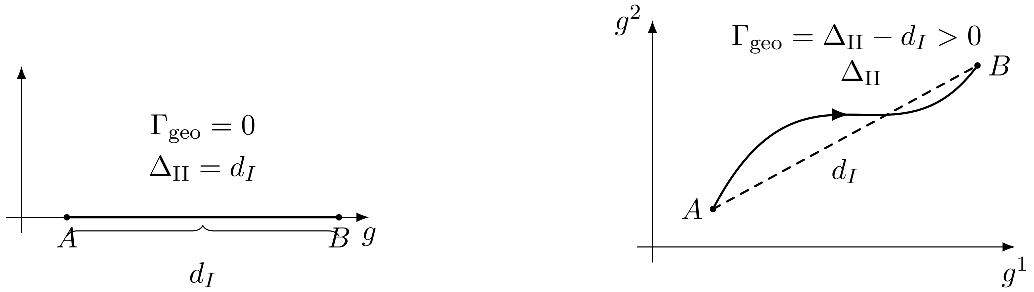
Take $x = \mu$ and $x_{\mathrm{ref}} = \mu_0$, then

$$\mathsf{S}_\kappa(\mu) = \mu\, e^{i\kappa \ln(\mu/\mu_0)}. \tag{F.8}$$

This encodes the logarithmic scale coordinate $\gamma = \ln(\mu/\mu_0)$ as a phase variable $\theta = \kappa\gamma$. The encoding is only used to draw spiral images of scale windows and scale steps; it does not change any definition or decision rule in the main text.

F.4 Image-level schematic for the endpoint–path gap (representation only; not part of V)

This subsection provides a purely schematic figure to visually correspond to Δ_{II} (path length / arc length), d_I (endpoint distance), and $\Gamma_{\mathrm{geo}} = \Delta_{\mathrm{II}} - d_I$. Emphasis: this schematic does not enter the viewpoint V and does not participate in any derivation. In the figure, "arc length / chord length" only expresses the geometric intuition "path length vs. endpoint distance"; the actual computation remains governed by the metric G and the defining formulas locked in the main text.



(a) Baseline: a straight trajectory (path length = endpoint distance).

(b) Nontrivial: a curved trajectory (path length > endpoint distance).

Figure 1: Schematic visualization of the endpoint–path geometric gap (representation only; not part of V).

References

[1] Kenneth G. Wilson. Renormalization group and strong interactions. *Physical Review D*, 3 (8):1818–1846, 1971. doi: 10.1103/PhysRevD.3.1818.

[2] Kenneth G. Wilson and John Kogut. The renormalization group and the ϵ expansion. *Physics Reports*, 12(2):75–200, 1974. doi: 10.1016/0370-1573(74)90023-4.

[3] Steven Weinberg. Phenomenological lagrangians. *Physica A*, 96(1–2):327–340, 1979. doi: 10.1016/0378-4371(79)90223-1.

[4] Howard Georgi. Effective field theory. *Annual Review of Nuclear and Particle Science*, 43: 209–252, 1993. doi: 10.1146/annurev.ns.43.120193.001233.

[5] Leo P. Kadanoff. Scaling laws for ising models near t_c. *Physics Physique Fizika*, 2(6): 263–272, 1966. doi: 10.1103/PhysicsPhysiqueFizika.2.263.

[6] Solomon Kullback and Richard A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951. doi: 10.1214/aoms/1177729694.

[7] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, 2 edition, 2006. ISBN 9780471241959.

[8] Ronald A. Fisher. Theory of statistical estimation. *Mathematical Proceedings of the Cambridge Philosophical Society*, 22(5):700–725, 1925. doi: 10.1017/S0305004100009580.

[9] C. Radhakrishna Rao. Information and the accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, 37:81–91, 1945.

[10] Shun-ichi Amari and Hiroshi Nagaoka. *Methods of Information Geometry*. American Mathematical Society and Oxford University Press, 2000. ISBN 9780821805302.

[11] N. N. Cencov. *Statistical Decision Rules and Optimal Inference*, volume 53 of *Translations of Mathematical Monographs*. American Mathematical Society, 1982. ISBN 9780821845028.

[12] Murray Gell-Mann and Francis E. Low. Quantum electrodynamics at small distances. *Physical Review*, 95(5):1300–1312, 1954. doi: 10.1103/PhysRev.95.1300.

[13] Richard P. Feynman. Space-time approach to quantum electrodynamics. *Physical Review*, 76(6):769–789, 1949. doi: 10.1103/PhysRev.76.769.

[14] Freeman J. Dyson. The radiation theories of tomonaga, schwinger, and feynman. *Physical Review*, 75(3):486–502, 1949. doi: 10.1103/PhysRev.75.486.

[15] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Addison-Wesley, 1995. ISBN 9780201503975.

[16] Joseph Polchinski. Renormalization and effective lagrangians. *Nuclear Physics B*, 231(2): 269–295, 1984. doi: 10.1016/0550-3213(84)90287-6.

[17] Thomas Appelquist and J. Carazzone. Infrared singularities and massive fields. *Physical Review D*, 11(10):2856–2861, 1975. doi: 10.1103/PhysRevD.11.2856.

[18] Hisaharu Umegaki. Conditional expectation in an operator algebra. iv. entropy and information. *Kodai Mathematical Seminar Reports*, 14(2):59–85, 1962.

[19] Huzihiro Araki. Relative entropy of states of von neumann algebras. *Publications of the Research Institute for Mathematical Sciences*, 11(3):809–833, 1976. doi: 10.2977/prims/1195191148.

[20] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 10th anniversary edition, 2010. ISBN 9781107002173.

[21] Göran Lindblad. Completely positive maps and entropy inequalities. *Communications in Mathematical Physics*, 40(2):147–151, 1975. doi: 10.1007/BF01609396.

[22] Dénes Petz. Sufficient subalgebras and the relative entropy of states of a von neumann algebra. *Communications in Mathematical Physics*, 105(1):123–131, 1986. doi: 10.1007/BF01212345.

[23] Bradley Efron and Robert J. Tibshirani. *An Introduction to the Bootstrap*. Chapman & Hall/CRC, 1994. ISBN 9780412042317.

[24] Olive Jean Dunn. Multiple comparisons among means. *Journal of the American Statistical Association*, 56(293):52–64, 1961. doi: 10.1080/01621459.1961.10482090.

[25] Sture Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979. doi: 10.2307/4615733. URL https://www.jstor.org/stable/4615733.