


A Six Birds’ Eye View of Dark Energy: Closure, Route Mismatch, and Audits for Apparent Acceleration

Ioannis Tsiokos 

5 February 2026

Version 2 (revised 10 February 2026) · Zenodo v1: 10.5281/zenodo.18494480

Abstract

Empirical cosmology infers late-time acceleration by fitting homogeneous FLRW models to data products that are, by construction, coarse-grained summaries of an inhomogeneous universe. This procedure makes it difficult to separate a fundamental component from a closure correction required by the coarse-graining itself. We apply Six Birds Theory (SBT) to treat cosmological modeling as a closure package consisting of a lens (specifying what is retained), a completion (specifying how discarded structure is filled in), and audits (testing whether the packaged macro description is dynamically coherent under evolution). Operationally, we implement a minimal computational API for lenses, completions, packaging operators, idempotence defects, and route mismatch, together with provenance-tracked run bundles and PPD-style cross-probe audits. In calibrated toy systems, route mismatch is (numerically) zero in a linear control regime and becomes strictly positive when nonlinearity is turned on; in a patch-expansion proxy, a homogeneous-control case yields $Q(t) \approx 0$, whereas heterogeneous initial conditions produce a positive acceleration proxy ($Q_{\text{peak}} \approx 1.74$) and $\ddot{a}_{\mathcal{D}} > 0$ for $\sim 99\%$ of late-time samples. Using synthetic distance–redshift data from a null- Λ heterogeneous generator, homogeneous Λ CDM fitting recovers $\Omega_{\Lambda} \simeq 0.60$, and a heterogeneity-proxy rewrite term matches Λ CDM fit quality and improves held-out prediction of a heterogeneity proxy by $\sim 100\times$ in RMSE. On public DES SN5YR and DES Y6 BAO releases we demonstrate the same provenance and audit pipeline. These results motivate framing “dark energy” as an audit-driven rewrite hypothesis with staging and probe-split tests rather than a default new fluid. We are explicit that a fundamental cosmological constant is not ruled out here; current background probes prefer an approximately constant correction over the sampled redshift range and our LSS demonstrations are protocol-level under a surrogate backend pending physical $3\times 2\text{pt}$ likelihoods.

keywords: dark energy; coarse-graining; cosmological backreaction; closure; posterior predictive checks; data compression; six birds; cosmology

1 Introduction

Cosmological “dark energy” is usually introduced as an additional component in the stress–energy budget of an otherwise homogeneous Friedmann–Lemaître–Robertson–Walker (FLRW) universe. Empirically, the case for late-time acceleration is established by comparing model-predicted distance measures and clustering statistics to survey data products, and the resulting minimal phenomenological description is Λ CDM [1, 2, 3]. The conceptual step that often remains implicit is that the fitted homogeneous description is treated as a closed macro-dynamical system: the model

is evolved forward in time at the macro level, and its predictions are compared directly to observations. This paper asks what is required for that “macro closure” step to be justified, and what it means when it fails.

Six Birds Theory (SBT) provides a vocabulary designed for precisely this situation: modeling a system that is observed only through a coarse interface. In SBT, a theory is not just a dynamical law; it is a closure package consisting of (i) a lens $f : X \rightarrow Y$ specifying what is retained, (ii) a completion $U : Y \rightarrow X$ specifying how discarded information is filled in, and (iii) audits that test whether the induced macro description is coherent under evolution [4, 5]. The associated packaging operator $E = U \circ f$ is intended to be idempotent and to define the “objects” at the macro level. A central SBT diagnostic is route mismatch: if T denotes time evolution on X , then in general

$$E \circ T \neq T \circ E, \tag{1}$$

so that “evolve then package” differs from “package then evolve.” When this mismatch is non-negligible, a closed homogeneous macro law cannot be expected to emerge from micro-evolution without modification. SBT classifies such modification as an *operator rewrite*: the effective macro evolution must be changed (the “P1” primitive) to restore coherence under the chosen interface.

From this viewpoint, dark energy can be reinterpreted as a rewrite term required by closure. In the standard cosmology pipeline, the lens is determined by the data product (e.g., distance–redshift summaries, BAO dilation measures, or two-point functions), and the completion is the modeling assumption that a homogeneous FLRW spacetime—augmented by a small number of parameters—provides an adequate macro completion. If route mismatch is present, the completion must either be refined (track more macro variables), the lens must be changed (retain more information), or the macro operator must be rewritten. Λ CDM can then be viewed as one widely successful rewrite family that repairs closure under common lenses.

Scope boundary. This paper does not claim to refute a fundamental cosmological constant, nor does it propose a specific microphysical alternative. Our claim is methodological and testable: *given a particular lens and completion, a nonzero effective acceleration term can arise as the closure correction required by route mismatch*, and this correction should leave audit signatures that differ from those of a truly fundamental component. We therefore focus on (i) making the closure assumptions explicit, (ii) providing reproducible mismatch- and audit-based diagnostics, and (iii) demonstrating the protocol on public data products that do not require proprietary likelihood code. In particular, although we build a switch-ready data layer for 3×2 pt large-scale-structure (LSS) vectors, we do not present a full physical 3×2 pt likelihood reproduction in this manuscript; instead, we use probe-splitting and staging audits on the public vectors as a closure-consistency demonstration. The public DES Y6 3×2 pt cosmology likelihoods were not available at the time of this writing, and our results do not depend on them.

Related work. The idea that cosmological backreaction from inhomogeneities could mimic or contribute to apparent acceleration has a substantial literature. Buchert’s averaging formalism for dust and perfect-fluid cosmologies [6, 7] provides the foundational kinematical framework; Clarkson et al. [8] offer a comprehensive review of the averaging, backreaction, and fitting problems. Räsänen [9] and Kolb et al. [10] explored scenarios in which backreaction could be quantitatively significant, while Green & Wald [11] argued on rigorous perturbative grounds that FLRW is a good approximation under standard conditions, placing important constraints on the viability of large backreaction effects. Wiltshire [12] proposed a “timescape” model treating observers in voids and walls differently, and Buchert et al. [13] subsequently compiled an extensive status report on the

backreaction debate. Our work does not resolve this debate but contributes a distinct methodological layer: rather than computing backreaction magnitudes within a specific GR averaging scheme, we provide a model-comparison and audit framework that can be applied to any proposed macro description and tested against public data products.

Contributions. We make the following concrete, falsifiable contributions:

1. **Closure-first formulation of dark energy.** We restate cosmological inference as a packaging problem involving a lens, a completion, and an induced packaging operator, and we identify route mismatch as the mechanism that forces operator rewrite terms at the macro level [4, 5].
2. **A reproducible computational and audit toolkit.** We implement a minimal API for lenses, completions, packaging, route mismatch, and idempotence defects, together with run manifests, provenance, and PPD-style cross-probe audits that standardize evidence reporting across experiments.
3. **Mechanism and inference demonstrations.** We provide toy nonlinear and patch-cosmology proxies in which coarse-graining generates mismatch and an effective acceleration proxy; we also show, with synthetic distance–redshift data, that a null- Λ heterogeneous generator can be mis-fit by a homogeneous Λ CDM model, yielding $\Omega_\Lambda > 0$ as a closure-induced inference artifact.
4. **Rewrite without fundamental Λ .** We introduce a mismatch-derived macro rewrite term that can match Λ CDM fit quality in the synthetic setting with comparable parameter count, and whose amplitude tracks heterogeneity in the toy generator.
5. **Public-data instantiations and audits.** We demonstrate the data/provenance pipeline and cross-probe audits on public DES SN5YR distances and DES Y6 BAO α -likelihood releases [14, 15, 16], and we apply the same audit protocol to public LSS vectors from DES Y3 and KiDS-450 [17, 18, 19].

Roadmap. Section 2 summarizes the SBT primitives and their cosmology mapping. Section 3 defines the toy generators, rewrite families, and audit protocol. Section 4 reports mechanism, synthetic-inference, and public-data results, emphasizing probe splits and staging dependence as closure fingerprints. We conclude in Section 5 with a set of audit-driven predictions and a checklist for future high-fidelity likelihood releases.

2 Six Birds framework for cosmology

This section summarizes the minimal Six Birds Theory (SBT) primitives we use throughout the paper and fixes notation. We follow the closure-first formulation in [4, 5]: a macro description is not merely a reduced state; it is the output of an explicit *packaging* operator induced by an observational interface (lens) and a modeling assumption (completion). Route mismatch (non-commutation) is the key diagnostic that forces operator rewrites when a closed homogeneous macro theory does not descend.

Instantiation note (cosmology). In full cosmology, a natural choice of microstate space X would include the spacetime geometry and matter fields (or, operationally, the evolving density and velocity fields on a hypersurface), while Y corresponds to survey data products derived from that state. This manuscript does not attempt to operate directly on that full microstate space; instead, we (i) treat X and Y abstractly in the SBT definitions, and (ii) instantiate them concretely in controlled toy systems (Sections 3.1–3.2) where X is a patch ensemble or a micro vector and Y is a coarse summary (mean or volume-average). The purpose is to isolate the closure logic (packaging, mismatch, rewrite, and audits) in settings where it can be computed end-to-end.

2.1 Lens, completion, and packaging

Let X denote a microstate space (“world states”) and let Y denote an observation space (“data products”). A *lens* retains only the coarse information available to an observer, whereas a *completion* specifies how the discarded degrees of freedom are filled in by a model.

Definition 2.1 (Lens and completion). A *lens* is a map $f : X \rightarrow Y$. A *completion* is a map $U : Y \rightarrow X$.

Definition 2.2 (Packaging operator). The induced *packaging operator* is the endomap

$$E = U \circ f : X \rightarrow X. \tag{2}$$

Intuitively, $E(x)$ is the “macro-consistent representative” obtained by observing x through f and filling in missing structure via U .

In many settings (including cosmology), f is determined by the data reduction pipeline (e.g., compressed distance summaries, dilation parameters, or two-point vectors), whereas U encodes the modeling family used to interpret those summaries (e.g., an FLRW background with a small parameter set). The packaging operator formalizes the act of “fitting a homogeneous model” as a map on microstates.

2.2 Coherence: idempotence and route mismatch

SBT distinguishes two minimal coherence requirements: (i) packaging should be stable when applied repeatedly, and (ii) packaging should be compatible with time evolution. To express approximate coherence, we allow pseudometrics d_X on X and d_Y on Y (in practice implemented as simple norms or residual metrics on vectors).

Definition 2.3 (Idempotence and idempotence defect). Packaging is *idempotent* if $E(E(x)) = E(x)$ for all $x \in X$. When idempotence is only approximate, we define the *idempotence defect*

$$\delta(x) := d_X(E(E(x)), E(x)). \tag{3}$$

Let $T : X \rightarrow X$ denote one-step evolution on microstates (or T_τ evolution over a time increment τ). The central obstruction to “macro closure” is that packaging need not commute with evolution:

$$E \circ T \neq T \circ E, \tag{4}$$

as already highlighted in Eq. (1). This non-commutation is what SBT terms *route mismatch* (or holonomy).

Table 1: Six Birds primitives (P1–P6) mapped to cosmology and to the concrete artifacts used in this manuscript.

Primitive	SBT role	Cosmology interpretation / operationalization
P1	Operator rewrite	Effective correction required for macro closure (e.g., ΔH term; Λ as a rewrite family).
P2	Gating / constraints	Survey masks, selection functions, scale cuts, and admissible macro families defining feasible completions.
P3	Route mismatch	Backreaction-style non-commutation $E \circ T \neq T \circ E$; measured via $\text{RM}(x)$.
P4	Staging	Dependence on lens scale / smoothing / scale cuts; “staging fingerprints” from sweeps over cut scales.
P5	Packaging	A macro object is a fixed point of E ; “fit an FLRW model” as $E = U \circ f$.
P6	Accounting / audit	PPD-style checks: fit on one probe block, predict another; cross-probe bias and residual structure.

Definition 2.4 (Route mismatch). Given evolution T , define the (state-space) *route mismatch diagnostic*

$$\text{RM}(x) := d_X(E(T(x)), T(E(x))). \quad (5)$$

Optionally, a lens-space diagnostic $d_Y(f(E(T(x))), f(T(E(x))))$ may be used when only observables can be compared.

Remark 2.5 (Operator rewrite viewpoint). When $\text{RM}(x)$ is generically non-negligible under a chosen lens and completion, a closed macro evolution on packaged states does not descend. SBT treats the necessary repair as an *operator rewrite* (P1): introduce an effective macro operator \tilde{T} such that

$$E \circ T \approx \tilde{T} \circ E \quad (6)$$

in the regime of interest. In cosmology, ΛCDM can be interpreted as one successful rewrite family that makes a homogeneous packaged description appear dynamically coherent under common observational interfaces.

2.3 Six Birds (P1–P6) and their cosmology roles

SBT organizes closure phenomena into six recurring “birds” (P1–P6). Table 1 provides the working mapping used in this paper.

2.4 Lean-backed sanity lemmas

To avoid ambiguity about what “mismatch vanishes” means, we record the exact-equality version as a sanity check and note its formalization. Our Lean4 project (`lean/SixBirds/Defs.lean` and `lean/SixBirds/Lemmas.lean`) defines abstract types X, Y , lenses, completions, packaging, and the exact predicates corresponding to Eqs. (2)–(5). In the experiments we use pseudometric-valued diagnostics (Eqs. (3)–(5)); the Lean lemmas record the exact-commutation ideal limit that those diagnostics approximate.

Lemma 2.6 (Commutation implies zero route mismatch (exact form)). *If T and E commute as functions on X (i.e., $E \circ T = T \circ E$), then the route-mismatch predicate holds for all x (equivalently, $\text{RM}(x) = 0$ under any pseudometric compatible with equality).*

Remark 2.7 (Lean formalization references). The previous lemma is formalized as `SixBirds.Lemmas.routeMismatch_of_commute`. Similarly, exact idempotence implies zero idempotence defect, formalized as `SixBirds.Lemmas.idempotenceDefectAt_of_idempotent`. A sufficient-condition lemma for packaging commutation under homomorphism/section assumptions is formalized as `SixBirds.Lemmas.routeMismatch_packaging_of_homomorphism_section`. Proofs are omitted here; the Lean sources compile in this repository.

3 Methods

Our methodology proceeds in three layers: (i) mechanism demonstrations in controlled toy systems, (ii) synthetic inference in which a heterogeneous (null- Λ) generator is fit by homogeneous macro models, and (iii) audit protocols (PPD-style checks) that quantify cross-probe predictive coherence. This section defines the toy generators, the synthetic distance proxy construction, the rewrite families, and the audit metrics used throughout.

3.1 Toy model 1: nonlinear microdynamics under a mean lens

Toy 1 is designed to isolate the origin of route mismatch in the simplest setting: nonlinear microevolution combined with an information-losing lens.

Microstate and evolution. Let the microstate be a vector $x_t \in \mathbb{R}^n$ at discrete time t . We evolve each component with a logistic-like quadratic drift:

$$x_{t+1} = x_t + \Delta t(\alpha x_t - \beta x_t^{\odot 2}), \tag{7}$$

where $x_t^{\odot 2}$ denotes elementwise squaring, and $(\alpha, \beta, \Delta t)$ are scalar parameters. The linear case corresponds to $\beta = 0$.

Lens and completion. We take the lens to be the mean,

$$f(x) = \bar{x} := \frac{1}{n} \sum_{i=1}^n x_i, \tag{8}$$

and the completion to be the constant-vector embedding,

$$U(\bar{x}) = (\bar{x}, \dots, \bar{x}) \in \mathbb{R}^n. \tag{9}$$

The induced packaging operator $E = U \circ f$ maps any microstate to the constant vector with the same mean.

Diagnostics. Given the one-step evolution map $T(x) = x + \Delta t(\alpha x - \beta x^{\odot 2})$, we compute the route mismatch diagnostic

$$\text{RM}(x) = d_X(E(T(x)), T(E(x))), \tag{10}$$

and the idempotence defect $\delta(x)$ as in Eq. (3). In this construction E is exactly idempotent (up to floating point), so δ serves primarily as a numerical sanity check, while RM captures the failure of coarse-grained closure driven by nonlinearity.

3.2 Toy model 2: patch-expansion proxy and an acceleration diagnostic

Toy 2 is a cosmology-motivated proxy designed to capture the key structural feature relevant to dark energy inference: heterogeneous expansion histories that become “homogenized” by a volume-average lens.

Patch ensemble. We consider N patches indexed by $i = 1, \dots, N$, each with a local scale factor $a_i(t)$ and matter density $\rho_i(t)$. We impose the dust scaling

$$\rho_i(t) = \rho_{i0} a_i(t)^{-3}, \quad (11)$$

with heterogeneous initial densities ρ_{i0} .

Local expansion law. Each patch evolves via

$$\dot{a}_i(t) = a_i(t) H_i(t), \quad (12)$$

where we define a Friedmann-like closure for H_i :

$$H_i(t)^2 = \frac{g_{4\pi}}{3} \rho_i(t) + \frac{\kappa_i}{a_i(t)^2}. \quad (13)$$

Here $g_{4\pi}$ plays the role of $4\pi G$ (up to normalization), and κ_i is a patch-dependent “curvature-like” parameter used only as a proxy for how under/over-density can induce systematically different expansion rates. In the numerical experiments, ρ_{i0} and κ_i are drawn from simple heterogeneous prescriptions controlled by an amplitude parameter; the precise sampling procedure is recorded in the run configuration bundles.

Domain-average lens. Define the volume-average (domain) scale factor by

$$a_{\mathcal{D}}(t) := \left\langle a_i(t)^3 \right\rangle^{1/3}, \quad (14)$$

where $\langle \cdot \rangle$ denotes the arithmetic mean over patches. We also define the domain Hubble proxy $H_{\mathcal{D}} := \dot{a}_{\mathcal{D}}/a_{\mathcal{D}}$ and, when needed, a heterogeneity proxy such as $\text{Var}(H_i)$ across the ensemble.

Acceleration proxy. For a homogeneous dust FLRW model with the same normalization, one expects $3\ddot{a}/a + g_{4\pi}\rho \approx 0$ [6, 7, 8]. Motivated by this, we define the domain “closure term” proxy

$$Q(t) := 3 \frac{\ddot{a}_{\mathcal{D}}(t)}{a_{\mathcal{D}}(t)} + g_{4\pi} \langle \rho_i(t) \rangle. \quad (15)$$

Positive Q indicates that the domain-average evolution departs from a homogeneous dust closure in the direction of an “effective acceleration” term.

3.3 Synthetic distance–redshift mock and macro-model fits

To demonstrate how closure mismatch can produce an apparent Λ -like signal, we treat the Toy 2 domain-average history as a source of synthetic “observations” and fit homogeneous macro models to them.

Redshift mapping. We normalize the domain scale factor so that $a_{\mathcal{D}}(t_0) = 1$ at the final simulation time t_0 and define

$$1 + z(t) := a_{\mathcal{D}}(t)^{-1}. \quad (16)$$

This yields a monotone mapping between simulation time and redshift over the modeled interval.

Distance proxy. Given a macro expansion history $H(z)$, we define the comoving distance proxy (flat case) as

$$\chi(z) = c \int_0^z \frac{dz'}{H(z')}, \quad (17)$$

and luminosity distance $D_L(z) = (1+z)\chi(z)$ [20]. Synthetic measurements are generated at a grid $\{z_k\}$ with additive noise,

$$D_k^{\text{obs}} = D_L(z_k) + \epsilon_k, \quad \epsilon_k \sim \mathcal{N}(0, \sigma_k^2). \quad (18)$$

Macro models and objective. We fit two standard homogeneous background families: (i) a matter-only model (with $\Omega_\Lambda = 0$), and (ii) a Λ CDM model (with Ω_Λ free, optionally with flatness imposed). Given a model prediction $D_k(\theta)$, we minimize the Gaussian chi-square

$$\chi^2(\theta) = \sum_k \frac{(D_k^{\text{obs}} - D_k(\theta))^2}{\sigma_k^2}, \quad (19)$$

and report information criteria such as $\text{AIC} = \chi_{\text{min}}^2 + 2k$, where k is the number of fitted parameters [21, 22].

3.4 Rewrite model families

SBT interprets dark energy phenomenology as an operator rewrite required by closure under a chosen lens. In background form, this motivates a family of “rewrite” expansions in which an additive correction supplements matter-only evolution.

One- and two-parameter rewrite. Let $H_m(z)$ denote the matter-only expansion law (with any curvature assumptions specified by the macro family). We define a rewrite model by

$$H(z)^2 = H_m(z)^2 + \Delta_H(z), \quad (20)$$

with a low-dimensional parameterization of Δ_H . In the synthetic experiments we use simple families such as

$$\Delta_H(z) = A g(z), \quad g(z) = (1+z)^{-m}, \quad (21)$$

where $A \geq 0$ is an amplitude and $m \geq 0$ controls redshift dependence. In addition, when a heterogeneity proxy is available (e.g., $\text{Var}(H_i)$ in Toy 2), we use $g(z)$ as a fixed shape proportional to that proxy, leaving only an amplitude parameter to be fit. Positivity is enforced by restricting parameters so that $H(z)^2 > 0$ over the fit range.

Proxy-derived vs. phenomenological rewrites. In the synthetic patch setting (Toy 2), we can define $g(z)$ using a heterogeneity proxy measured from the generator (e.g., variance of H_i across patches mapped to redshift), so that $\Delta_H(z)$ is tied to a mismatch proxy and only an amplitude is fit. In the public-data background tests, no such generator-level heterogeneity proxy is available; there we use a simple phenomenological shape family $g(z) = (1+z)^{-m}$ to test whether a low-parameter rewrite can match Λ CDM on geometry-only probes, without claiming a unique microphysical origin.

3.5 Posterior-predictive (PPD-style) audit protocol

We follow the posterior-predictive assessment framework of [23]. SBT emphasizes audits (P6) that test whether a closure package is predictive across observational routes, not merely in-sample. We implement lightweight probe-splitting audits in which a model is fit on one “probe” (or subset of a data vector) and evaluated on another.

Train–test split. Let \mathcal{D}_A and \mathcal{D}_B denote two disjoint probe blocks (e.g., SN vs. BAO, or cosmic shear vs. galaxy clustering + galaxy–galaxy lensing). Let θ denote model parameters. We fit on \mathcal{D}_A to obtain

$$\hat{\theta}_A := \arg \min_{\theta} \chi_A^2(\theta), \quad (22)$$

and then compute predictive discrepancies on \mathcal{D}_B using $\hat{\theta}_A$.

Held-out metrics. Given residuals $r_B = y_B - y_B(\hat{\theta}_A)$ and covariance C_B for the held-out block, we report

$$\chi_{B|A}^2 = r_B^T C_B^{-1} r_B, \quad (23)$$

together with unweighted and covariance-weighted root-mean-square error (RMSE),

$$\text{RMSE}_{B|A} = \sqrt{\frac{1}{n_B} \sum_{j=1}^{n_B} r_{B,j}^2}, \quad \text{RMSE}_{B|A}^{(w)} = \sqrt{\frac{\chi_{B|A}^2}{n_B}}, \quad (24)$$

and the mean residual (bias) $\text{bias}_{B|A} = \frac{1}{n_B} \sum_j r_{B,j}$. We repeat the audit in the reverse direction ($A \leftrightarrow B$). In the context of closure, systematic predictive bias on held-out probes is interpreted as evidence that the chosen completion (macro family) is not coherent under the observational interface, motivating either staging changes (P4) or operator rewrite (P1). We refer to these as “PPD-style” audits because they mirror the posterior predictive structure (fit on one route, test on another), but in this manuscript we typically use point estimates (MAP/least-squares) and treat the resulting discrepancies as comparative diagnostics rather than calibrated posterior predictive p -values.

3.6 Public background-probe likelihoods (geometry anchors)

This manuscript uses public background probes as low-overhead geometry anchors and as substrates for cross-probe audits. We emphasize that these are *compressed* likelihood objects (a distance-modulus vector for SN; a one-dimensional $\Delta\chi^2(\alpha)$ curve for BAO), used here to demonstrate closure and audit mechanics rather than to reproduce the full survey likelihood machinery.

3.6.1 DES SN5YR distance-modulus vector likelihood

The SN observable is a distance-modulus vector μ^{obs} with total covariance

$$C_{\text{SN}} = C_{\text{sys}} + \text{diag}(\sigma_{\text{stat}}^2), \quad (25)$$

where C_{sys} is the released systematics covariance and σ_{stat} are the per-supernova statistical uncertainties provided in the release metadata. Given a background model θ and luminosity distance $D_L(z; \theta)$ (in Mpc), we define the predicted distance modulus

$$\mu^{\text{th}}(z; \theta, \mathcal{M}) = 5 \log_{10} \left(\frac{D_L(z; \theta)}{\text{Mpc}} \right) + 25 + \mathcal{M}, \quad (26)$$

where \mathcal{M} is a single intercept parameter that absorbs the absolute magnitude and any overall distance normalization [20]. Because SN distances do not determine H_0 without an absolute calibration, we fix H_0 to a reference value in the background distance calculation and treat \mathcal{M} as the nuisance intercept (shape-only constraints on θ).

The Gaussian SN objective is

$$\chi_{\text{SN}}^2(\theta, \mathcal{M}) = (\mu^{\text{obs}} - \mu^{\text{th}}(\theta, \mathcal{M}))^\top C_{\text{SN}}^{-1} (\mu^{\text{obs}} - \mu^{\text{th}}(\theta, \mathcal{M})). \quad (27)$$

We profile over \mathcal{M} analytically (equivalently: maximize the likelihood in \mathcal{M} for each θ). Writing $\Delta(\theta) = \mu^{\text{obs}} - \mu^{\text{th}}(\theta, \mathcal{M}=0)$ and letting $\mathbf{1}$ be the all-ones vector, the profiled intercept is

$$\widehat{\mathcal{M}}(\theta) = \frac{\mathbf{1}^\top C_{\text{SN}}^{-1} \Delta(\theta)}{\mathbf{1}^\top C_{\text{SN}}^{-1} \mathbf{1}}. \quad (28)$$

All reported SN-only fits in this paper use the profiled objective $\chi_{\text{SN}}^2(\theta, \widehat{\mathcal{M}}(\theta))$. We report best-fit points (MAP under bounds) rather than full posterior constraints, since the paper’s emphasis is closure/audit behavior rather than precision parameter inference.

3.6.2 DES Y6 BAO one-dimensional α -likelihood

The DES Y6 BAO public release provides a one-dimensional likelihood curve in the dilation parameter α as a $\Delta\chi^2(\alpha)$ function for a fixed template/fiducial analysis. We treat this as a compressed geometry anchor. In standard BAO-template language, α is a dimensionless dilation parameter relating the observed BAO scale to the template prediction; a common theoretical mapping has the form

$$\alpha_{\text{th}}(\theta) = \frac{D_M(z_{\text{eff}}; \theta)/r_d(\theta)}{D_M^{\text{fid}}(z_{\text{eff}})/r_d^{\text{fid}}}, \quad (29)$$

where D_M is the transverse comoving distance and r_d is the sound horizon at drag epoch. We adopt the standard definition of r_d as the drag-epoch sound horizon and cite the canonical transfer-function treatment [24]. In this manuscript’s minimal pipeline, we *hold the sound-horizon ratio fixed* ($r_d(\theta) = r_d^{\text{fid}}$) and use

$$\alpha_{\text{th}}(\theta) = \frac{D_M(z_{\text{eff}}; \theta)}{D_M^{\text{fid}}(z_{\text{eff}})}, \quad (30)$$

so that α_{th} depends primarily on the shape parameters (e.g., Ω_m and curvature assumptions) and is insensitive to the absolute H_0 scale. The BAO contribution to the objective is then obtained by interpolation of the released curve:

$$\chi_{\text{BAO}}^2(\theta) = \Delta\chi^2(\alpha_{\text{th}}(\theta)). \quad (31)$$

This treatment is intentionally narrow: it reproduces the released one-dimensional likelihood object, which is used for audits and low-overhead geometry anchoring, but it does not replace a full BAO analysis in which $r_d(\theta)$ is modeled self-consistently and anisotropic observables are used.

3.7 Public LSS vectors and surrogate audit backend

For LSS audits, we ingest public two-point data vectors and covariances and construct a per-element block index (probe/statistic/bin/angle) used to define masks and probe splits. Let $y \in \mathbb{R}^n$ be a data vector with covariance C . Given a boolean mask $m \in \{0, 1\}^n$, define the masked residual and covariance by selecting the corresponding indices, yielding $r_m \in \mathbb{R}^{n_m}$ and $C_m \in \mathbb{R}^{n_m \times n_m}$.

Surrogate theory family. Because no physical 3×2 pt theory engine is used in this manuscript (Limitation 4.1), we define a low-rank surrogate prediction family as a linear correction applied to a fixed template basis:

$$y^{\text{th}}(\phi) = y^{\text{ref}} + T\phi, \quad (32)$$

where y^{ref} is a fixed reference vector, $T \in \mathbb{R}^{n \times p}$ is a fixed template matrix whose columns are predetermined functions of the block metadata (e.g., probe- and statistic-dependent shape vectors), and $\phi \in \mathbb{R}^p$ are fitted coefficients. In the manuscript tables we use a one-parameter baseline ($p = 1$, “ Λ CDM-like”) and a two-parameter extension ($p = 2$, “rewrite-like”).

Train on one probe, test on another. Given a training mask m_A (Probe A), we fit ϕ by minimizing the masked chi-square

$$\hat{\phi}_A := \arg \min_{\phi} (y_A - y_A^{\text{th}}(\phi))^{\text{T}} C_A^{-1} (y_A - y_A^{\text{th}}(\phi)), \quad (33)$$

and evaluate held-out discrepancy on mask m_B (Probe B) using Eq. (23). Crucially, the basis T is fixed *before* the split, and only the coefficients ϕ are fit using only training indices; this prevents leakage from the held-out block into the fitted coefficients. Absolute chi-square values in this surrogate setting depend on the (arbitrary) normalization of T ; therefore, in this manuscript we interpret held-out metrics as protocol diagnostics and compare models only within the same surrogate construction.

4 Results

We present results in three layers aligned with the closure-first narrative: (i) mechanism demonstrations (route mismatch and effective acceleration proxies), (ii) synthetic inference showing how a homogeneous model can infer $\Omega_{\Lambda} > 0$ from a null- Λ heterogeneous generator, and (iii) audit signatures (PPD-style) and staging dependence. Public-probe results are reported in Sections 4.4 and 4.5.

4.1 Mechanism: mismatch from nonlinearity and coarse-graining

4.1.1 Toy 1: route mismatch vanishes in the linear case and grows with nonlinearity

Toy 1 isolates the core mathematical mechanism: nonlinear micro-evolution combined with an information-losing lens yields non-commutation between packaging and evolution. Table 2 summarizes the maximal observed route mismatch RM_{max} for three nonlinearity strengths. In the linear case $\beta = 0$, route mismatch is numerically zero; when $\beta > 0$ it becomes strictly positive and increases with nonlinearity (e.g., $\text{RM}_{\text{max}} \approx 2.03 \times 10^{-2}$ at $\beta = 0.5$ and $\text{RM}_{\text{max}} \approx 8.10 \times 10^{-2}$ at $\beta = 2.0$; Table 2). As expected for the mean-lens packaging in this construction, the idempotence defect remains at the numerical floor ($\delta_{\text{max}} \sim 10^{-14}$), indicating that the observed mismatch is not a packaging-instability artifact.

4.1.2 Toy 2: heterogeneity induces a positive closure term proxy $Q(t)$

Toy 2 is a patch-expansion proxy in which each patch follows a local decelerating law, but the domain-average history under the volume-average lens can depart from homogeneous dust closure. Table 3 reports the diagnostic $Q(t)$ defined in Eq. (15). Under homogeneous initial conditions, Q

Table 2: Toy 1 mismatch summary.

metric	value
beta_linear_case_rm_max	0
beta_small_rm_max	0.0202599
beta_large_rm_max	0.0810395
delta_max_overall	7.02167e-15

Table 3: Toy 2 patch model summary.

metric	value
homogeneous_Q_max_abs	3.10862e-15
hetero_Q_peak	1.74203
hetero_Q_mean_late	3.25539e-05
hetero_frac_ddot_aD_pos_late	0.987552

is consistent with numerical zero (maximum absolute value at machine precision), whereas heterogeneous initial conditions yield a strictly positive peak ($Q_{\text{peak}} \approx 1.74$) and a late-time regime in which $\ddot{a}_{\mathcal{D}} > 0$ for most sampled times (late-time positive fraction ≈ 0.988 ; Table 3). This illustrates the minimal “apparent acceleration” mechanism: heterogeneity plus coarse-grained packaging can produce an effective acceleration proxy even when no fundamental Λ is present in the generator. We emphasize that $Q(t)$ here is a closure diagnostic defined by Eq. (15) for this proxy model; it is not the GR-averaging kinematical backreaction term, and we do not claim Toy 2 constitutes a relativistic backreaction calculation.

4.2 Synthetic inference: $\Omega_{\Lambda} > 0$ as a closure-induced artifact

We next treat the Toy 2 domain-average history as a generator of synthetic distance–redshift data (Section 3.3) and fit homogeneous macro models to that mock observation. This isolates the inference question: what does a homogeneous model infer when the underlying generator is heterogeneous but observed through a coarse lens?

4.2.1 Homogeneous fits infer $\Omega_{\Lambda} \simeq 0.6$ from null- Λ synthetic data

For one heterogeneous regime, a Λ CDM background fit yields $\Omega_{\Lambda} \simeq 0.60$ even though the data were generated from a null- Λ heterogeneous toy universe. Moreover, the Λ CDM fit improves the objective dramatically compared to a matter-only homogeneous model (information-criterion improvement $\Delta\text{AIC} \sim 8.2 \times 10^3$ in this mock; see Table 4 for the corresponding chi-square and AIC values). This is the closure-induced inference artifact predicted by the packaging viewpoint: when packaging does not commute with micro-evolution, a homogeneous completion introduces an effective rewrite term when forced to match lensed summaries.

4.2.2 Rewrite term matches Λ CDM fit quality and tracks heterogeneity

We next replace “fit Λ ” with a heterogeneity-proxy rewrite family (Section 3.4) in which $H^2(z) = H_{\text{m}}(z)^2 + \Delta_H(z)$ and Δ_H is parameterized by a low-dimensional proxy shape. Table 4 compares matter-only, Λ CDM, and rewrite fits in the synthetic setting. In the canonical run, the rewrite

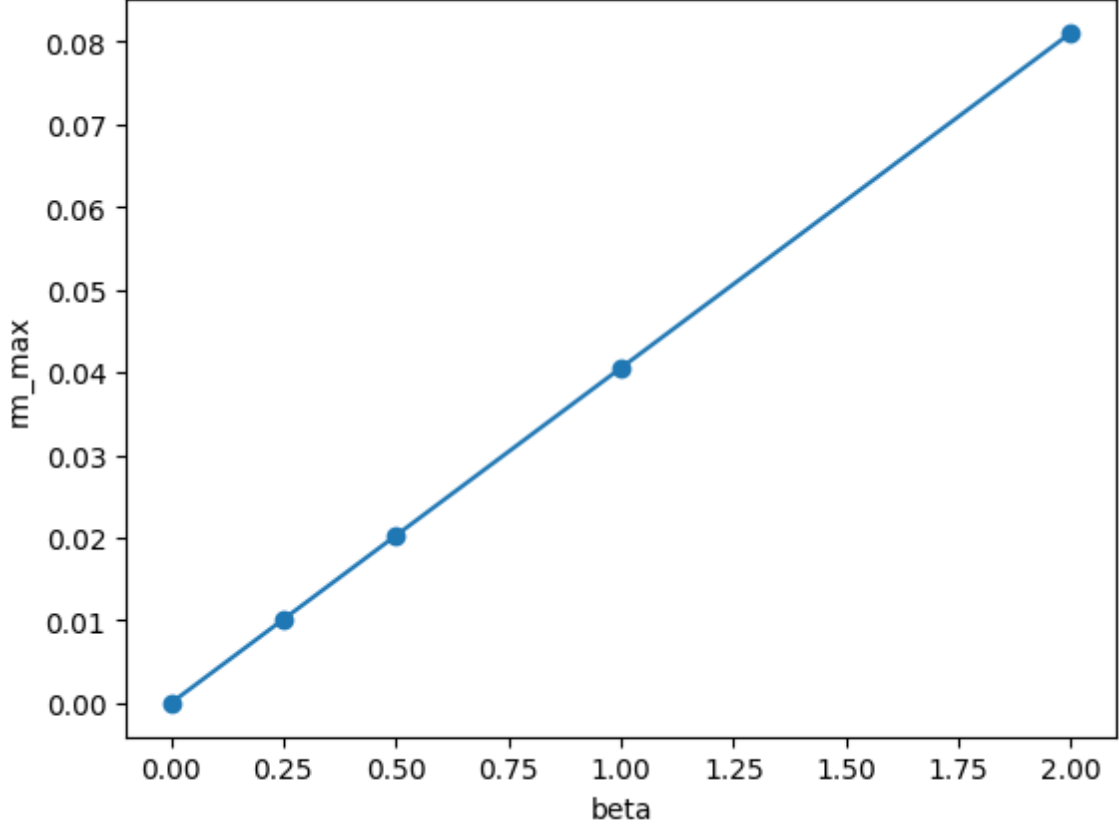


Figure 1: Toy 1: route mismatch summary as a function of nonlinearity strength β under a mean lens and constant-vector completion. The linear case exhibits $\text{RM} \approx 0$, whereas nonlinear evolution produces nonzero mismatch.

model achieves chi-square and AIC comparable to ΛCDM with the same parameter count ($k = 2$ in both cases), whereas matter-only is strongly disfavored.

In addition to fit quality, the closure-first interpretation demands a mechanistic connection between the rewrite amplitude and heterogeneity. In a sweep over heterogeneity amplitudes in the toy generator, the best-fit rewrite amplitude increases monotonically with the heterogeneity control parameter (Spearman $\rho = 1.0$ in the coarse sweep reported in the experiment logs), consistent with Δ_H functioning as a packaging-induced correction rather than as a free fluid component.

4.2.3 Synthetic PPD audit: rewrite improves prediction of a held-out heterogeneity proxy

Fit quality on a single probe is not a closure audit. We therefore implement a PPD-style test (Section 3.5) in which the model is fit to the distance–redshift probe and evaluated on a distinct held-out observable: a heterogeneity proxy derived from the patch ensemble (e.g., variance of H_i). Across multiple random seeds in a heterogeneous regime, the rewrite model reduces the held-out predictive RMSE by roughly two orders of magnitude relative to a homogeneous ΛCDM completion (median RMSE ratio $\text{RMSE}_{\text{rewrite}}/\text{RMSE}_{\Lambda\text{CDM}} \approx 8.8 \times 10^{-3}$ in the reported sweep). This behavior is consistent with a closure correction tied to heterogeneity: the rewrite term improves cross-route predictivity, not only in-sample distances.

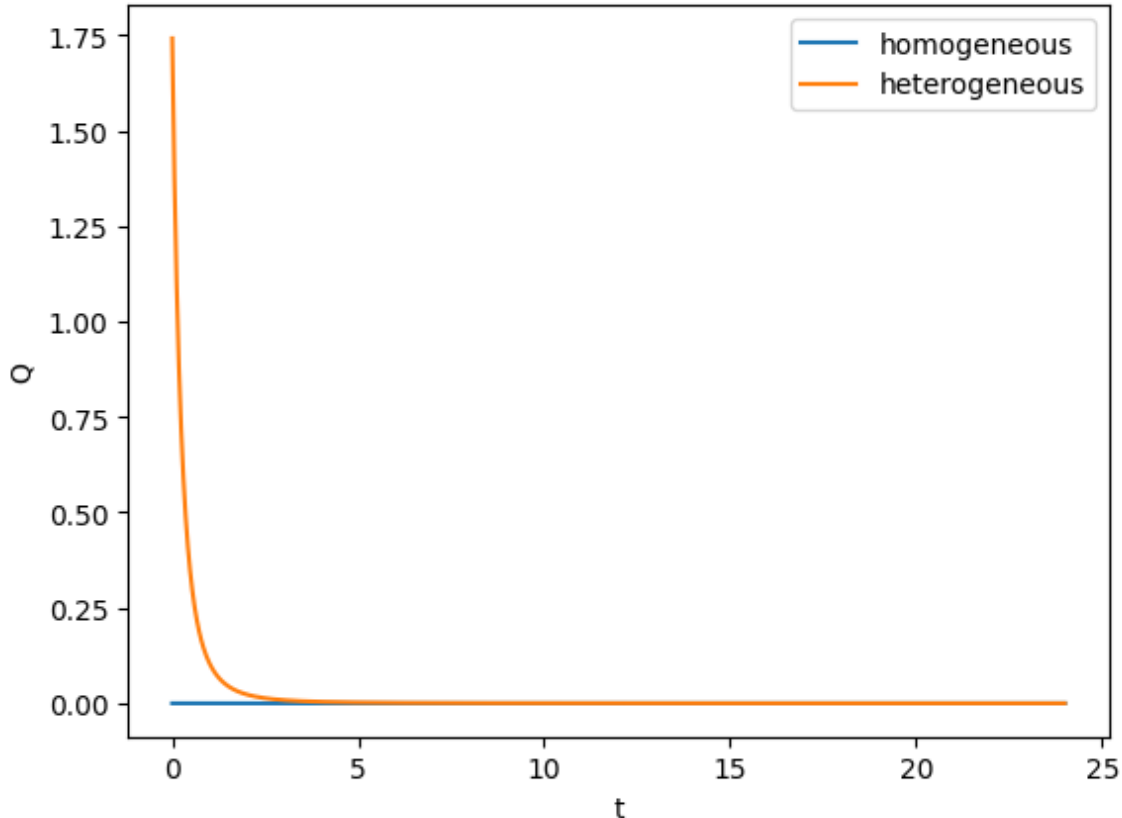


Figure 2: Toy 2: closure term proxy $Q(t)$ (Eq. (15)) becomes positive in heterogeneous regimes, indicating departure of the domain-average evolution from homogeneous dust closure.

4.3 Staging dependence as a closure fingerprint

SBT predicts that closure diagnostics can depend on the staging scale (P4): changing the lens scale or the resolution at which packaging is enforced can change mismatch and the required rewrite. We demonstrate this effect in a toy “scale sweep” in which the lens aggregates patches into domains of varying size L before packaging. In the recorded sweep, the late-time mean mismatch varies with L (e.g., $\text{RM}(L=1) \approx 0$ whereas $\text{RM}(L_{\max}) \approx 1.3 \times 10^{-4}$), and the smallest idempotence defect occurs at a specific staging choice (L minimizing δ in that sweep). These staging signatures motivate treating scale-cut and smoothing dependence in real analyses as potential closure fingerprints rather than as mere nuisances: if a correction term is packaging-induced, it should exhibit systematic staging dependence beyond what is expected from known measurement systematics.

4.4 Public background probes: DES SN5YR and DES Y6 BAO

We now demonstrate the closure-first pipeline and audit protocol on public, lightweight background probes: the DES SN5YR distance-modulus vector and covariance release, and the DES Y6 BAO α -likelihood release. These probes anchor geometry with minimal modeling overhead and do not require full large-scale-structure ($3 \times 2\text{pt}$) likelihood machinery.

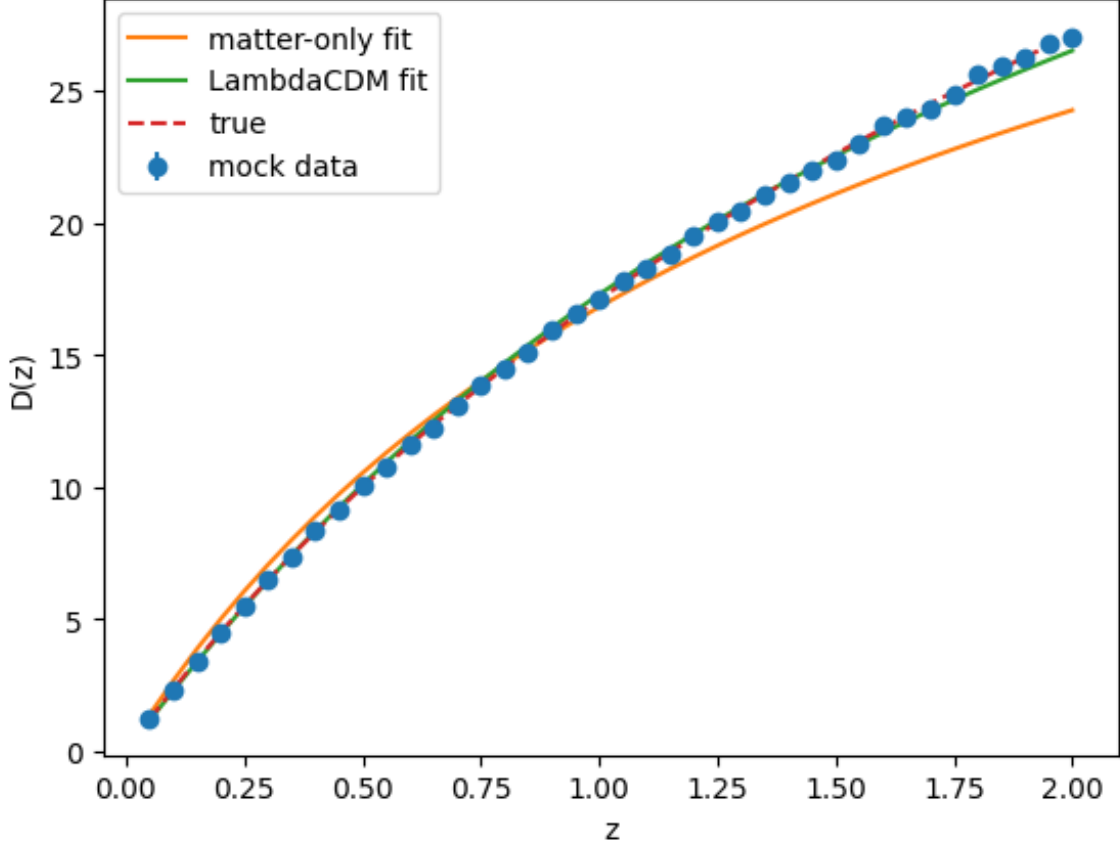


Figure 3: Synthetic distance–redshift mock from a heterogeneous (null- Λ) generator fit by homogeneous macro models. The best-fit homogeneous Λ CDM model infers $\Omega_\Lambda > 0$ and matches the mock more closely than a matter-only homogeneous model.

4.4.1 DES SN5YR likelihood sanity and SN-only fits

The DES SN5YR public release provides a Hubble-diagram distance vector and covariance components. A crucial sanity step is constructing the *total* covariance correctly: the “STATONLY” covariance component is zero by design in this release, and statistical uncertainties are provided separately as per-object distance-modulus errors. The total covariance is therefore the sum of the released systematics covariance and the diagonal statistical term. With this construction, the covariance is symmetric and positive-definite, and SN-only fits yield a reasonable goodness-of-fit (chi-square per datum $\simeq 0.90$ in our baseline configuration). Because SN distances do not fix H_0 without an absolute calibration, we fix H_0 in the distance calculation and absorb the normalization into the profiled intercept \mathcal{M} , so SN constrains shape parameters rather than absolute scale.

Fitting a flat Λ CDM background to SN-only yields a best-fit matter density $\Omega_m \simeq 0.35$. Allowing a constant equation-of-state parameter (flat w CDM) yields a best-fit $w \simeq -0.82$ with a small improvement in chi-square, consistent with the well-known SN-only degeneracy between Ω_m and w in the absence of external anchors.

Table 4: Synthetic inference comparison.

model	chi2	aic	param
matter_only	8588.54	8590.54	
lcdm	379.609	383.609	Omega_L=0.602429
rewrite	378.593	382.593	A=1509.91

4.4.2 DES Y6 BAO α -likelihood fit

For DES Y6 BAO, the public release provides a one-dimensional likelihood curve in the dilation parameter α (reported as a $\Delta\chi^2(\alpha)$ curve in the combined likelihood). In our implementation, we fit a homogeneous background by mapping (H_0, Ω_m) to a predicted α at the effective redshift and evaluating $\Delta\chi^2$ by interpolation. The combined curve peaks at $\alpha_{\text{hat}} \simeq 0.9568$ with an approximate width $\sigma_\alpha \simeq 0.0126$ in our Gaussian approximation to the local curvature at the peak. In our BAO-only fit, the ΔAIC between flat ΛCDM and an $\Omega_\Lambda=0$ homogeneous fit is $\approx 5.0 \times 10^{-2}$ (negligible). Because this is a single compressed observable, we treat BAO-only ΔAIC differences of $\ll 1$ as negligible and do not interpret them as evidence for or against Λ ; the curve serves as a geometry anchor for audits.

4.4.3 Background PPD audit: fit SN, predict BAO (and reverse)

Fit quality on a single probe is not, by itself, a closure audit. We therefore perform a probe-splitting posterior-predictive check (Section 3.5) between SN and BAO. Training on SN and predicting BAO yields a mean bias in α of $\text{bias}_\alpha \simeq 2.58 \times 10^{-2}$, corresponding to a z -score of $\simeq 2.04$ relative to the inferred σ_α . Training on BAO and predicting SN yields a mean distance-modulus bias of $\text{bias}_\mu \simeq 4.76 \times 10^{-2}$ and a covariance-weighted held-out RMSE of order unity (in units of $\sqrt{\chi^2/n}$ for the SN vector). These numbers quantify the expected limitation of fitting one compressed geometry probe and extrapolating to another in the absence of a shared completion that is coherent across both.

4.4.4 Rewrite-background vs. ΛCDM on SN and SN+BAO

We finally test a staged, phenomenological rewrite family directly on public background probes, without committing to a detailed inhomogeneous micro-model. Specifically, we fit a background rewrite model of the form $H^2(z) = H_m^2(z) + A g(z)$ with $g(z) = (1+z)^{-m}$ (Section 3.4). Table 5 summarizes fits for SN-only, BAO-only, and the combined SN+BAO objective. For combined fits we use

$$\chi_{\text{tot}}^2(\theta) = \chi_{\text{SN}}^2(\theta, \widehat{\mathcal{M}}(\theta)) + \chi_{\text{BAO}}^2(\theta). \quad (34)$$

In both SN-only and SN+BAO fits, the rewrite model matches the best-fit chi-square of ΛCDM but pays the expected information-criterion penalty for the extra parameter (typically $\Delta\text{AIC} \approx 2$). Moreover, the best-fit shape parameter saturates at $m \simeq 0$ in these background-only fits, so the rewrite term collapses to a constant contribution over the probed redshift range. This outcome is notable: with current background probes alone, the data prefer a nearly constant closure correction, consistent with why ΛCDM is phenomenologically successful for geometry. In the closure-first view, time variation (or additional state variables) would be demanded only if audits across multiple probes and staging choices were to require it.

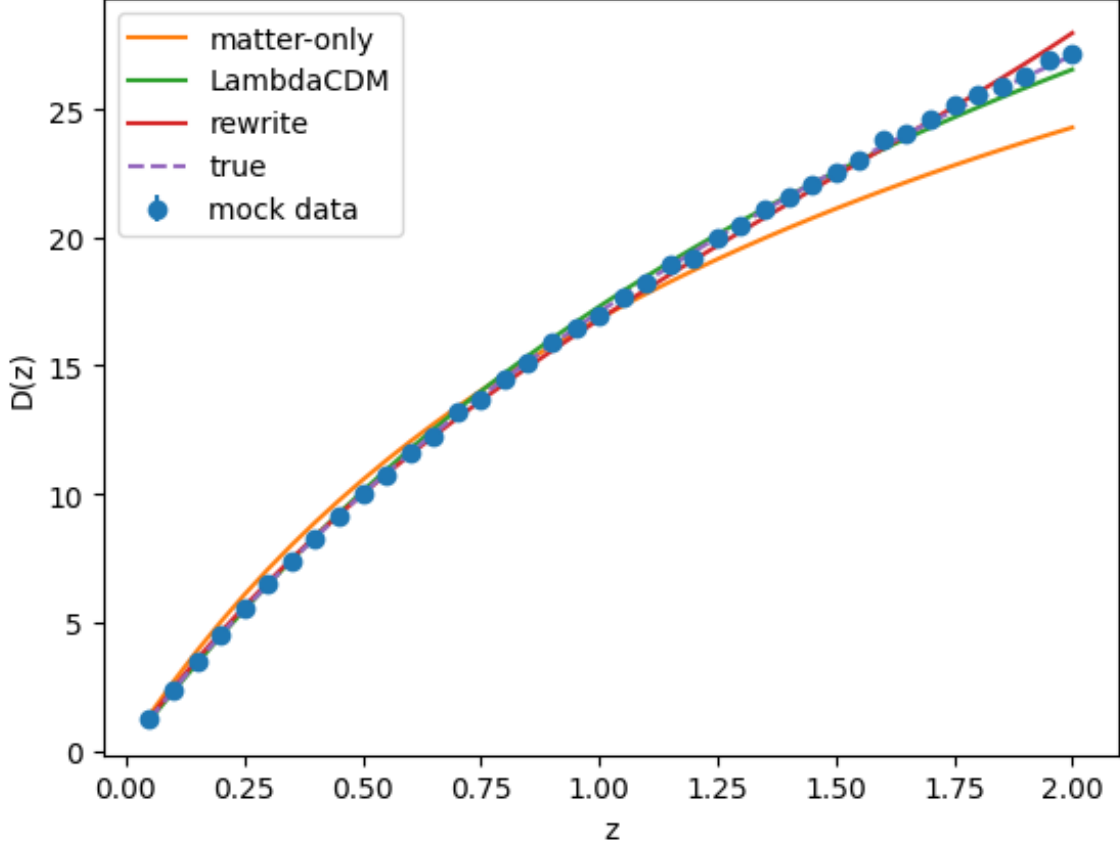


Figure 4: Synthetic distance–redshift mock: best-fit rewrite model compared to best-fit Λ CDM. In the synthetic setting, a heterogeneity-proxy rewrite term can reproduce the same distance fit quality as Λ CDM without positing a fundamental cosmological constant in the generator.

4.5 LSS data-layer and audit protocol (surrogate theory)

We next demonstrate the Six Birds audit protocol (P6) and data-layer staging machinery on public large-scale-structure (LSS) two-point vectors. This section is intentionally framed as an *audit protocol demonstration*: we ingest public data vectors, construct a block map for probe splitting, and run cross-probe predictivity checks. A full physical 3×2 pt theory backend (e.g., CCL/Firecrown/pyccl) is not required to demonstrate the audit mechanics, but it *is* required to interpret these LSS results as cosmological parameter constraints. We therefore separate the protocol and I/O layer (which is definitive here) from the physics backend (explicitly limited here).

Limitation 4.1 (Surrogate LSS backend).

The DES Y3 and KiDS-450 LSS results in this manuscript use a *surrogate* theory backend rather than a physical 3×2 pt forward model. Concretely, we define deterministic probe-dependent “template residuals,” fit low-dimensional correction coefficients, and then evaluate held-out discrepancies (chi-square/RMSE) under the public covariance. This approach establishes the audit protocol, probe splitting, masking, and reproducible bookkeeping under realistic vector/covariance dimensions, but it does *not* constitute a physical reproduction of the DES Y3 3×2 pt likelihood [17].

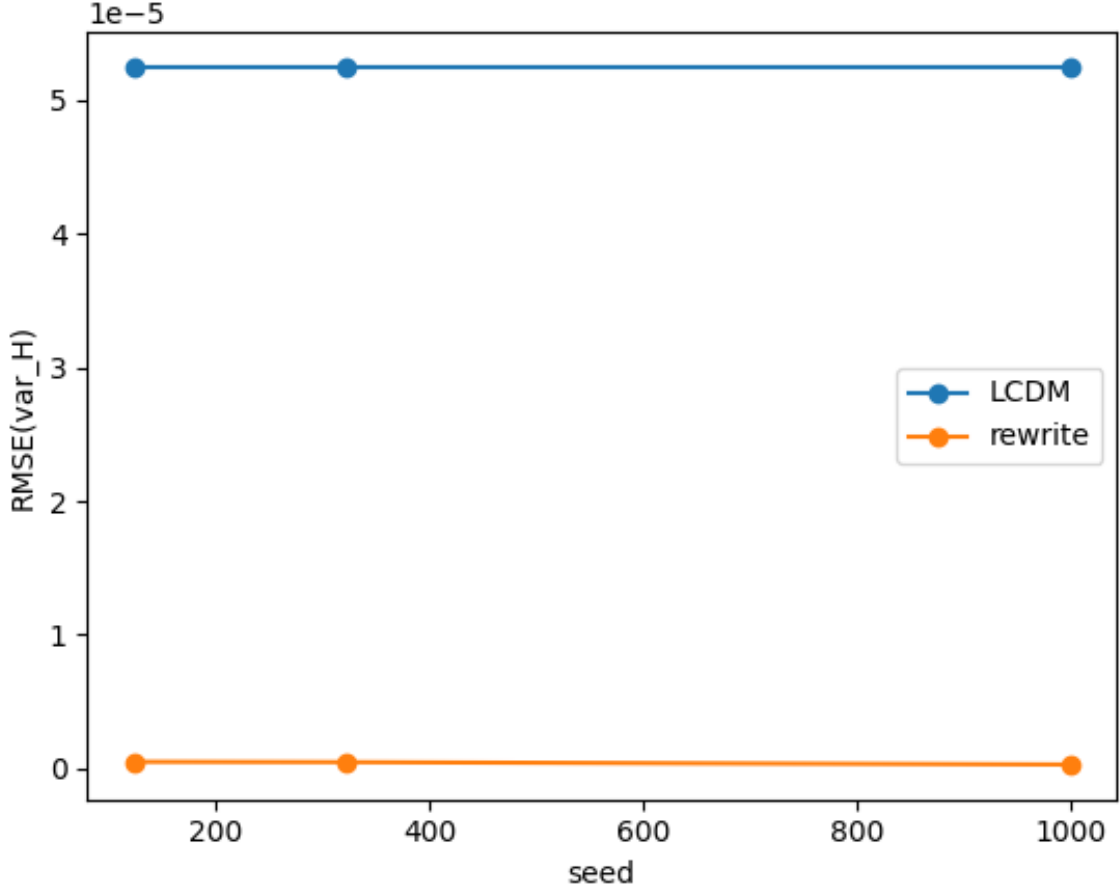


Figure 5: Synthetic PPD-style audit: predictive RMSE on a held-out heterogeneity proxy across multiple seeds in a heterogeneous regime. The rewrite model substantially reduces predictive error relative to a homogeneous Λ CDM fit on the distance probe.

The repository includes a plug-in backend interface and an environment probe; in the baseline environment, the physical backend import fails (e.g., `pycc1` unavailable), motivating this staged presentation.

4.5.1 DES Y3 2pt block map enables probe splits

We ingest the public DES Y3 2pt (maglim) data vector and covariance and construct a per-point block index assigning each element to a probe/statistic/bin combination. For the maglim vector used here ($n = 1000$), the resulting decomposition contains $n_{\text{shear}} = 400$ cosmic-shear points, $n_{\text{clust}} = 120$ galaxy-clustering points, and $n_{\text{ggl}} = 480$ galaxy-galaxy-lensing points (with no unclassified remainder). This block map is the prerequisite for meaningful probe splits and for staging operations such as scale cuts and block bootstraps. Numerically, we apply probe masks by selecting sub-vectors and sub-covariances and compute χ^2 via Cholesky-based linear solves (without explicit matrix inversion). In the baseline Y3 maglim run, the covariance is symmetric positive-definite with `chol_success = true` and `chol_jitter_used = 0.0`, while the estimated minimum eigenvalue and condition number are $\lambda_{\text{min}} \approx 4.5 \times 10^{-15}$ and $\kappa \approx 8.7 \times 10^9$, indicating ill-conditioning but stable solves under masked Cholesky.

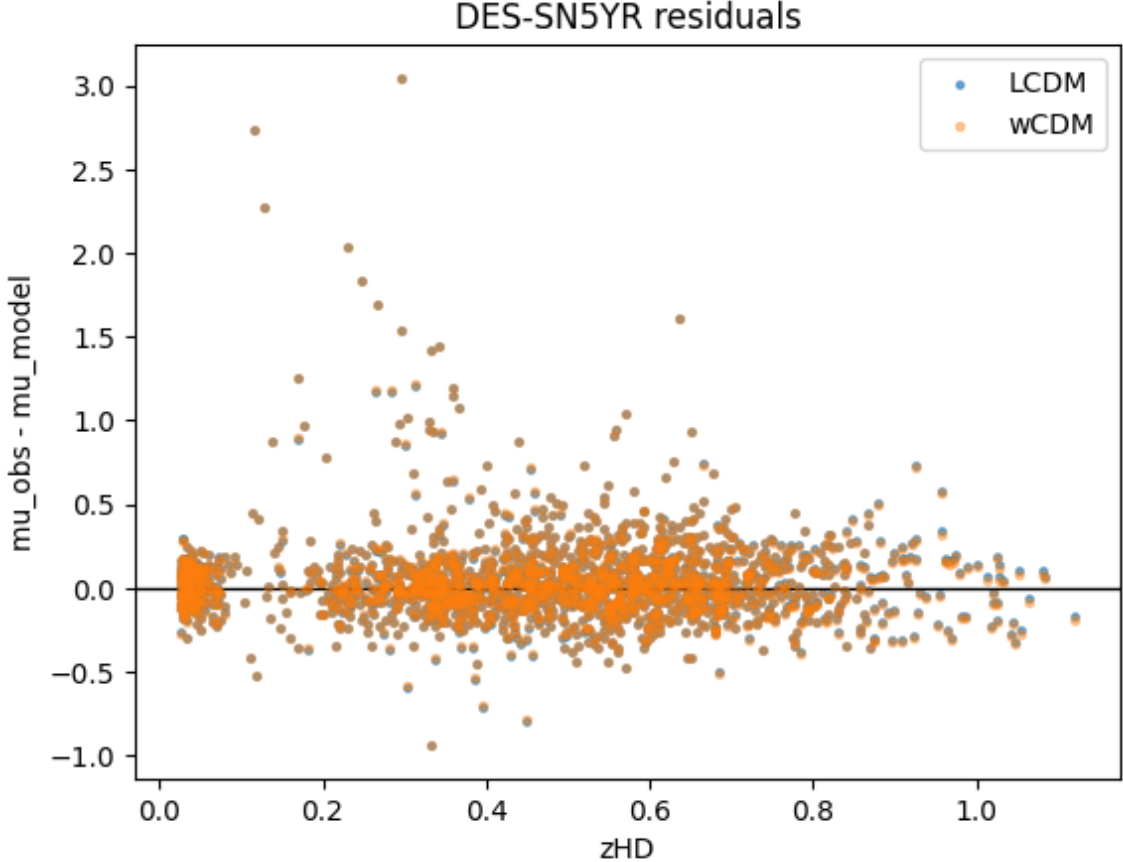


Figure 6: DES SN5YR: Hubble-diagram residuals for a best-fit homogeneous background model under the public distance-vector likelihood. Residual structure is shown relative to the fitted macro completion.

4.5.2 Probe-split PPD on DES Y3 vectors (surrogate backend)

Using the block map, we define Probe A as cosmic shear and Probe B as clustering+galaxy-galaxy lensing. We then run the PPD-style audit (Section 3.5): fit a low-dimensional macro correction on one probe block and evaluate predictive discrepancy on the held-out block. Table 6 reports held-out chi-square values for two macro families in the surrogate setting: a baseline “ Λ CDM-like” one-parameter correction and a “rewrite-like” two-parameter correction. Because the backend is surrogate (Limitation 4.1), the absolute scale of chi-square is not interpreted as a cosmological goodness-of-fit; instead we treat held-out discrepancy as an audit diagnostic that can be transported unchanged to a physical backend when available. The absolute χ^2 values are small because the surrogate template normalization is arbitrary; we therefore interpret only relative held-out comparisons and the reproducible train \rightarrow test protocol artifact.

4.5.3 Geometry anchors (SN+BAO) shift held-out discrepancies

To test sensitivity to external geometry information, we repeat the DES Y3 probe-split audit while adding public background-probe anchors (DES SN5YR distances and DES Y6 BAO α -likelihood) to the training objective. Table 7 compares held-out chi-square values with and without these anchors.

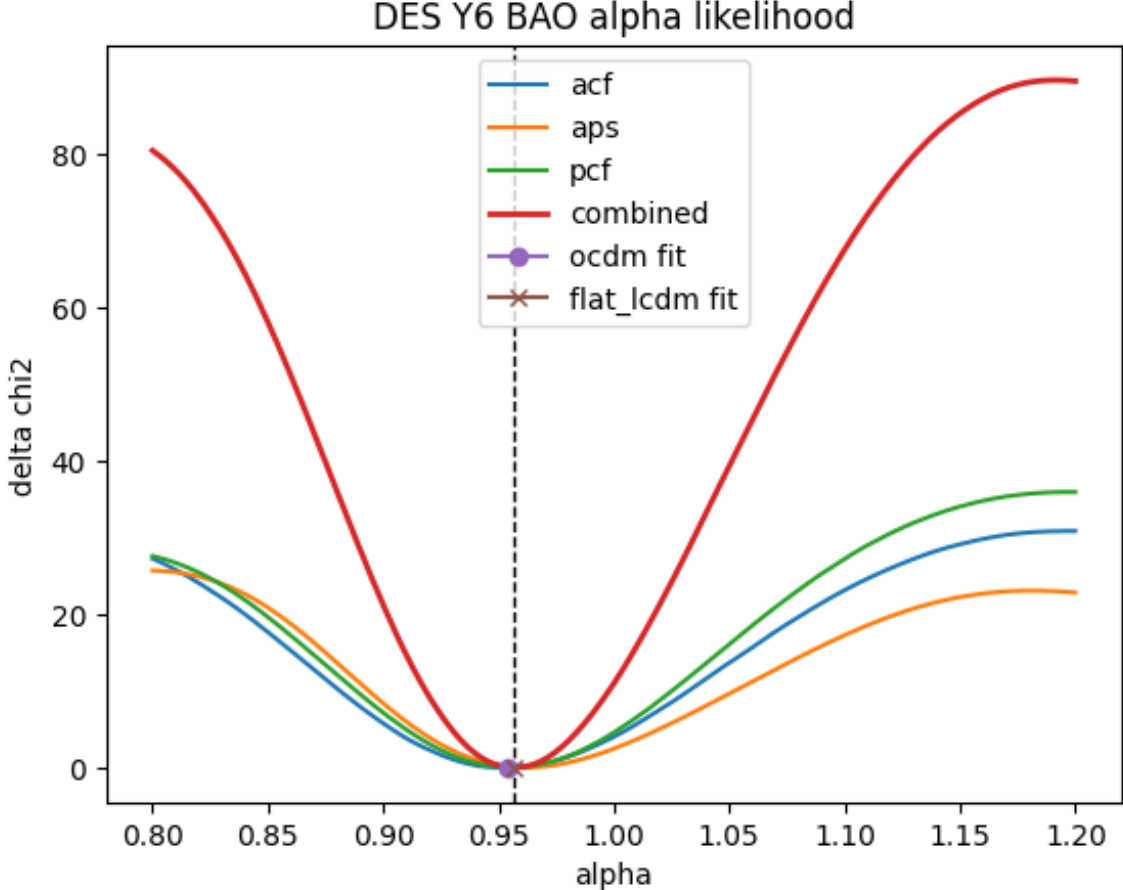


Figure 7: DES Y6 BAO: public one-dimensional $\Delta\chi^2(\alpha)$ curve (combined). In our pipeline, background parameters are mapped to a predicted α at the effective redshift and evaluated by interpolation.

In this staged setting, the anchors increase the held-out chi-square for both macro families and reduce the directional differences. This behavior is expected: adding shared geometry constraints partially fixes the degrees of freedom that the surrogate correction would otherwise have used to absorb cross-probe discrepancies. In these anchored runs we use

$$\chi_{\text{train}}^2 = \chi_{\text{LSS,train}}^2 + \chi_{\text{SN}}^2 + \chi_{\text{BAO}}^2,$$

and assume independence (no modeled cross-covariances) between the LSS and background-probe likelihoods.

4.5.4 KiDS-450 replication (bandpower vectors)

To demonstrate portability beyond DES, we replicate the same probe-split audit protocol on a public KiDS-450 3×2 pt bandpower data product (cosmic shear bandpowers, galaxy–galaxy lensing bandpowers, and clustering bandpowers) [18, 19]. The ingested vector has $n = 100$ with an even split between shear (Probe A; $n_A = 50$) and ggl+clustering (Probe B; $n_B = 50$). Table 8 reports held-out chi-square values under the same surrogate-backend macro families. Although the absolute

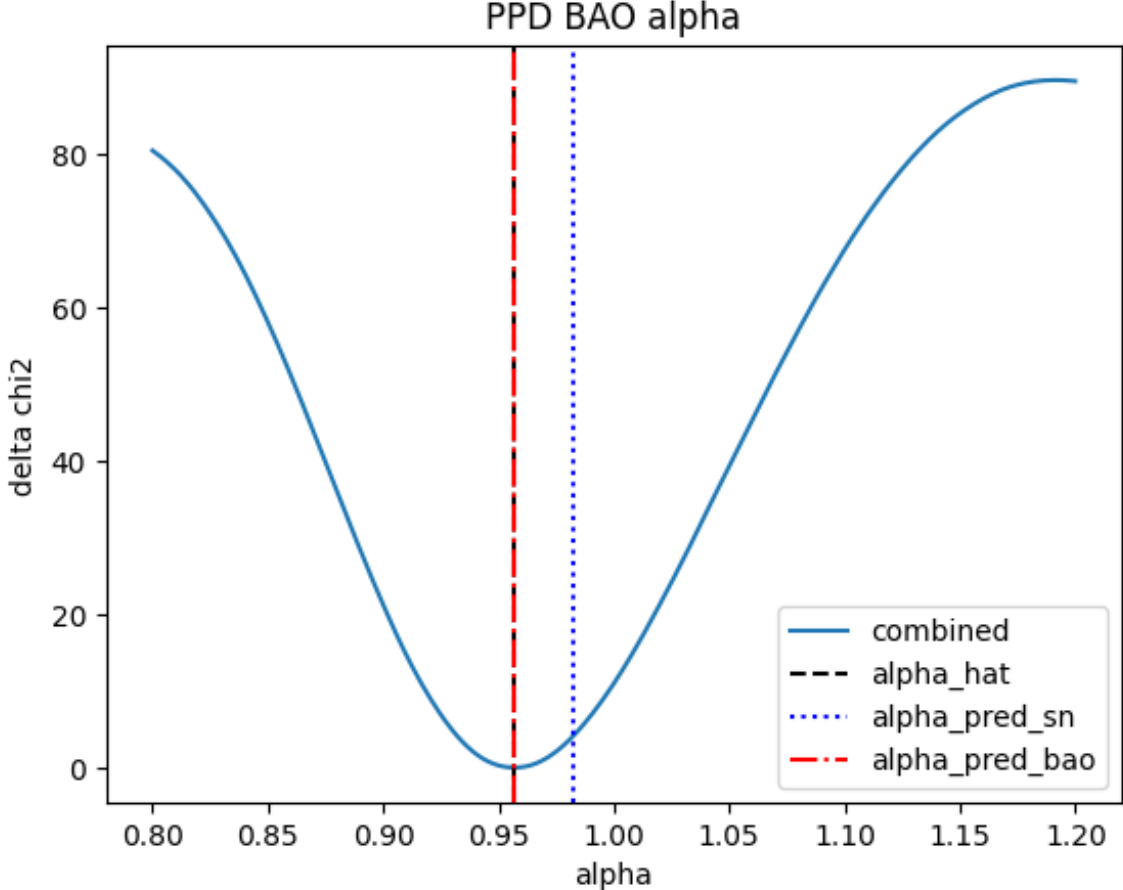


Figure 8: Background PPD audit example: BAO α prediction evaluated under a model trained on SN distances. The discrepancy is summarized by bias, RMSE, and z -score metrics.

values remain backend-dependent, the protocol yields a consistent summary artifact (train–test directions, held-out discrepancy, and reproducible provenance) that can be compared across surveys.

4.5.5 Robustness suite: multiple splits, block bootstrap, covariance jitter

Finally, we stress-test the audit protocol on DES Y3 by varying the split definition (shear versus rest, even/odd tomographic-bin proxy, low/high tomographic-bin proxy), subsampling probe-block keys (bootstrap-like block subsampling without replacement), and perturbing the covariance with small diagonal jitter. Across 30 requested scenarios, 26 completed successfully with 4 skipped due to insufficient points after subsampling (0 failed). For the canonical shear \rightarrow (clust+ggl) direction, the distribution of held-out improvement $\Delta\chi^2 = \chi_{\Lambda\text{CDM-like}}^2 - \chi_{\text{rewrite-like}}^2$ has median $\simeq 5.7 \times 10^{-3}$ and interquartile range $\simeq 1.94 \times 10^{-2}$ in this surrogate setting. The small magnitude reflects the surrogate scale of residuals; the stability of the sign and distribution across splits and jitter levels demonstrates that the evidence artifacts (tables, CSVs, provenance bundles) are robust to modest protocol perturbations, as required for P6-style accounting.

Staging note. We also performed a scale-cut (staging) sweep on the DES Y3 vectors to demonstrate the mechanics of P4 staging under the block map. Under the surrogate backend, the staging

Table 5: Background fits summary.

dataset	model	chi2	aic
SN	lcdm	1639.53	1645.53
SN	rewrite	1639.53	1647.53
SN+BAO	lcdm	1642.63	1648.63
SN+BAO	rewrite	1642.63	1650.63
BAO	lcdm	6.23778e-09	2

Table 6: DES Y3 (maglim) probe-split PPD audit under a surrogate LSS backend. Probe A is shear ($n_A = 400$) and Probe B is clustering+ggl ($n_B = 600$). Reported values are held-out χ^2 (Eq. (23)).

Direction	Model	Held-out block	χ^2_{test}
A→B	Λ CDM-like	B (clust+ggl)	0.1165
A→B	rewrite-like	B (clust+ggl)	0.0885
B→A	Λ CDM-like	A (shear)	0.0674
B→A	rewrite-like	A (shear)	0.0305

sweep is best treated as a protocol demonstration rather than a physics result; we therefore summarize it in Appendix A.

5 Discussion

The standard interpretation of Λ CDM treats dark energy as an additional component of the cosmic energy budget. The Six Birds Theory (SBT) interpretation offered here is distinct in character: it treats “dark energy” as the name we give to an effective correction required to obtain a closed macro description under a chosen observational interface. This section clarifies what that means, what it predicts, and how it can be tested.

5.1 Dark energy as a rewrite term induced by mismatch

In the closure-first formulation (Section 2), the macro model is not identified with a micro law; it is identified with a *packaged* description $E = U \circ f$ together with an effective macro evolution \tilde{T} satisfying Eq. (6). When route mismatch is negligible, a homogeneous macro model can descend without further correction. When route mismatch is non-negligible (Eq. (5)), an operator rewrite is forced if one insists on a closed macro description under the given lens.

The toy results make this logic explicit: nonlinear micro-evolution under coarse-graining produces nonzero mismatch (Toy 1), and a heterogeneous patch ensemble can yield an effective acceleration proxy in the domain-average history (Toy 2). The synthetic inference results then show the observational consequence: when a null- Λ heterogeneous generator is fit with a homogeneous completion, a Λ -like term is inferred (Section 4.2.1). In this sense, Λ CDM can be viewed as a highly successful rewrite family for common cosmological lenses, rather than necessarily as direct evidence for a fundamental fluid. This mismatch-based framing should not be conflated with GR backreaction claims: our Toy 2 proxy is a closure diagnostic, not a relativistic averaging calculation.

Table 7: DES Y3 probe-split PPD audit under surrogate LSS backend, comparing held-out chi-square without geometry anchors vs. with SN+BAO anchors added to the training objective.

Direction	Model	χ_{test}^2 (no anchors)	χ_{test}^2 (SN+BAO anchors)
A→B	Λ CDM-like	0.6750	1.0798
A→B	rewrite-like	0.9967	1.0820
B→A	Λ CDM-like	0.7186	1.2329
B→A	rewrite-like	0.4985	1.2458

Table 8: KiDS-450 bandpower probe-split PPD audit under a surrogate LSS backend (Probe A: shear; Probe B: ggl+clustering).

Direction	Model	Held-out block	χ_{test}^2
A→B	Λ CDM-like	B (ggl+clust)	0.0560
A→B	rewrite-like	B (ggl+clust)	0.0495
B→A	Λ CDM-like	A (shear)	0.0270
B→A	rewrite-like	A (shear)	0.0205

Non-uniqueness and why audits matter. One may object that one can always introduce an arbitrary function $\Delta_H(z)$ to fit geometry data. We agree: rewrites are generally non-unique. The novelty in the SBT framing is not the existence of a flexible correction, but the discipline imposed by packaging and audits: a rewrite hypothesis must (i) be tied to a lens/completion choice, (ii) improve cross-route predictivity under probe splits (P6), and (iii) exhibit staging dependence consistent with a packaging-induced correction rather than with a fundamental component (P4).

5.2 Why SBT is stricter than “alternative interpretation”

SBT is not merely interpretive; it is prescriptive about what a valid macro claim requires. Three features are central.

Stability (P5). Packaging should be coherent: repeated application should not change the packaged object (idempotence), and the idempotence defect (Eq. (3)) provides a basic sanity check ensuring that numerical results are not driven by unstable projection.

Novelty vs. refit (P1/P3). A good macro model is not only one that fits a chosen probe; it must also clarify whether its success arises from true descent (small mismatch) or from rewrite (large mismatch but repaired by a correction). The synthetic experiments illustrate this distinction: matter-only closure fails badly on the mock, whereas Λ CDM and the rewrite family both repair closure on the distance probe but differ in how they perform on held-out observables.

Directionality and accounting (P6). SBT emphasizes audits: fit one route, predict another. A model that “explains away” a discrepancy by adding parameters but fails systematically in posterior-predictive checks is not a coherent closure package. This is why we treat probe-splitting and cross-probe bias/RMSE as primary evidence artifacts, not optional diagnostics.

5.3 Staging predictions: scale dependence and probe splits (P4)

If a correction term is packaging-induced, it is expected to depend on staging choices: what the lens retains, how the completion is parameterized, and which scales are included in the data vector. Conversely, a truly fundamental cosmological constant should not exhibit structured dependence on reasonable staging changes, beyond known observational systematics.

In practice this motivates two families of tests: (i) *scale-cut or smoothing sweeps*, in which one varies the effective lens scale and re-runs inference/audits, and (ii) *probe splits*, in which one fits on a subset of observables and predicts another. We demonstrated the mechanics of these tests in both synthetic settings and in public-vector LSS audits, with an explicit limitation that the current LSS results use a surrogate backend (Limitation 4.1) and should not be read as physical parameter constraints. The staging sweep on DES Y3 vectors is therefore recorded as supplementary protocol evidence (Appendix A), while physical staging interpretation is deferred to a backend-enabled analysis. Operationally, a staging fingerprint would be a structured, reproducible dependence of the inferred correction (or of cross-probe predictive bias) on lens-scale choices (e.g., scale cuts, smoothing windows, summary-statistic definitions) that (a) persists across reasonable reanalyses, (b) changes directionally with probe splits, and (c) correlates with independent heterogeneity proxies when available. Absent these controls, staging dependence can be an observational systematic rather than a closure signature; our emphasis is therefore on probe-split audits and robustness suites rather than on any single scale-sweep curve.

5.4 How public-data audits can falsify (or support) rewrite families

A closure-first interpretation becomes scientifically useful only if it generates falsifiable consequences. The most direct falsification route is cross-probe predictivity: a rewrite family that repairs closure on one probe should not introduce systematic predictive bias on an independent probe once both are constrained consistently.

Our public-probe results illustrate the current status: background-only fits prefer a nearly constant rewrite correction (the power-law shape parameter saturates at $m \simeq 0$), which is consistent with the limited lever arm of SN+BAO geometry over the sampled redshift range. In this regime, a fundamental Λ and a constant closure correction are observationally difficult to distinguish on geometry alone. The discriminating power must therefore come from (i) additional probes sensitive to growth and inhomogeneity, and (ii) audit signatures under probe splits and staging changes. We also acknowledge the active debate on averaging and backreaction in relativistic cosmology and treat our closure-first claims as audit-driven rather than definitive statements about GR dynamics [8, 11].

5.5 Predictions and pre-registered checklists for forthcoming likelihood releases

The repository underlying this manuscript is designed to make these tests “switch-ready” as new public likelihoods and data products appear. In particular, once the DES Y6 3×2pt likelihood products are public, the following checklist can be executed with minimal code changes (dataset configuration swap plus backend selection), producing a reproducible manifest bundle for each run:

- **(P6) Probe-split PPD on physical theory vectors:** fit shear-only and predict clustering + ggl; then reverse. Compare held-out χ^2 and bias patterns between Λ CDM/ w CDM and a rewrite-bridge family.
- **(P4) Staging sweeps:** repeat the above while varying standard scale cuts (and, where

applicable, alternative summary statistics) and test for structured dependence of the inferred rewrite amplitude/shape.

- **(P1) Rewrite parsimony:** compare information criteria (AIC/BIC) for Λ CDM vs. low-parameter rewrite families; require that any added flexibility is justified by predictive improvement, not just in-sample fit.
- **(P3) Mismatch proxies:** when heterogeneity proxies are available (e.g., variance-like summary diagnostics, lensing-systematics splits), test whether inferred rewrite amplitudes correlate with them as predicted by closure-induced corrections.
- **(P2) Gating sensitivity:** repeat audits under alternative masks/selection/gating choices (tomographic splits, redshift cuts) to separate true staging fingerprints from selection-driven artifacts.
- **(Reproducibility) Evidence bundles:** for each run, record the configuration, metrics, plots, and provenance, and publish the figure/table provenance mapping used in the manuscript.

A key point is that these predictions do not require committing to a single microphysical alternative to Λ ; they require treating closure as a hypothesis with observable audit consequences.

5.6 Scope of claims

We emphasize that the synthetic demonstrations in this paper show a *mechanism*: that coarse-graining and route mismatch can generate an effective acceleration term under packaging. They do not, by themselves, prove that the real universe is null- Λ . The main contribution is the audit protocol—a set of falsifiable, probe-split, and staging-dependent tests that can distinguish a packaging-induced correction from a fundamental component as higher-fidelity data products become available. These tests are designed to be applied regardless of the outcome: if Λ is fundamental, audits should show no structured staging dependence or cross-probe bias beyond known systematics; if Λ is (partly) packaging-induced, the audit signatures described here should be detectable.

5.7 Limitations

We close with explicit limitations, which also identify the highest-value next steps.

- **Background-probe degeneracy:** with current SN+BAO geometry alone, the best-fit rewrite family collapses to a constant-like correction ($m \simeq 0$), making it difficult to distinguish a fundamental Λ from a constant closure correction on the basis of geometry alone.
- **Physical LSS backend unavailable in baseline environment:** the DES Y3 and KiDS LSS results here demonstrate audit protocol mechanics on public vectors but do not constitute a physical 3×2 pt likelihood reproduction (Limitation 4.1).
- **DES Y6 3×2 pt public artifacts not yet available:** the pipeline is switch-ready, but this manuscript does not report Y6 3×2 pt constraints because the relevant likelihood/data products were not public at the time of writing.
- **Rewrite family is phenomenological:** the rewrite term used here is intentionally low-dimensional, designed to test closure hypotheses; mapping $\Delta_H(z)$ to specific inhomogeneity

models is a separate step that should be constrained by audits across probes and staging. Additionally, the BAO α -likelihood is used in its released compressed form with a fixed sound-horizon ratio (full $r_d(\theta)$ modeling is deferred).

6 Conclusion

This paper has presented a Six Birds Theory (SBT) framing of dark energy as a closure/rewrite term induced by coarse-graining and route mismatch between packaging and evolution. We defined the minimal primitives (lens, completion, packaging, mismatch diagnostics), demonstrated the mechanism in toy systems, showed that homogeneous inference can yield $\Omega_\Lambda > 0$ from null- Λ heterogeneous synthetic data, and introduced a heterogeneity-proxy rewrite family that matches Λ CDM fit quality in that synthetic setting while improving cross-probe posterior-predictive performance.

On public data products, we implemented reproducible pipelines and audits for DES SN5YR and DES Y6 BAO background probes and established an LSS audit protocol on public vectors (DES Y3 and KiDS-450) with explicit staging and robustness machinery. The central message is not that Λ is false, but that *closure is a testable hypothesis*: if a correction term is packaging-induced, it should leave staging and cross-probe audit signatures that can be checked systematically as higher-fidelity likelihood releases become available.

Acknowledgements

This work used publicly released data products from the Dark Energy Survey (DES) and KiDS collaborations. All remaining errors are our own.

Declarations

Competing interests. The author declares no competing interests.

Funding. No external funding was received for this research.

Ethics approval. Not applicable; no human or animal participants or personal data were involved.

Data availability. All observational data used in this study are publicly available. DES SN5YR distance-modulus data and covariance: Zenodo 10.5281/zenodo.12720778 [15]. DES Y6 BAO α -likelihood: published with [16]. DES Y3 3×2 pt data vectors: published with [17]. KiDS-450 bandpower products: <http://kids.strw.leidenuniv.nl/> [18, 19]. Synthetic data and the seeds used to generate them are included in the code repository (see below).

Code availability. The full analysis code, experiment runners, and manuscript build system are available at <https://github.com/ioannist/six-birds-de>. A permanent archived snapshot is deposited on Zenodo: 10.5281/zenodo.18595154.

Author contributions. I. Tsiokos: Conceptualization, Methodology, Software, Investigation, Data curation, Formal analysis, Visualization, Writing (all stages).

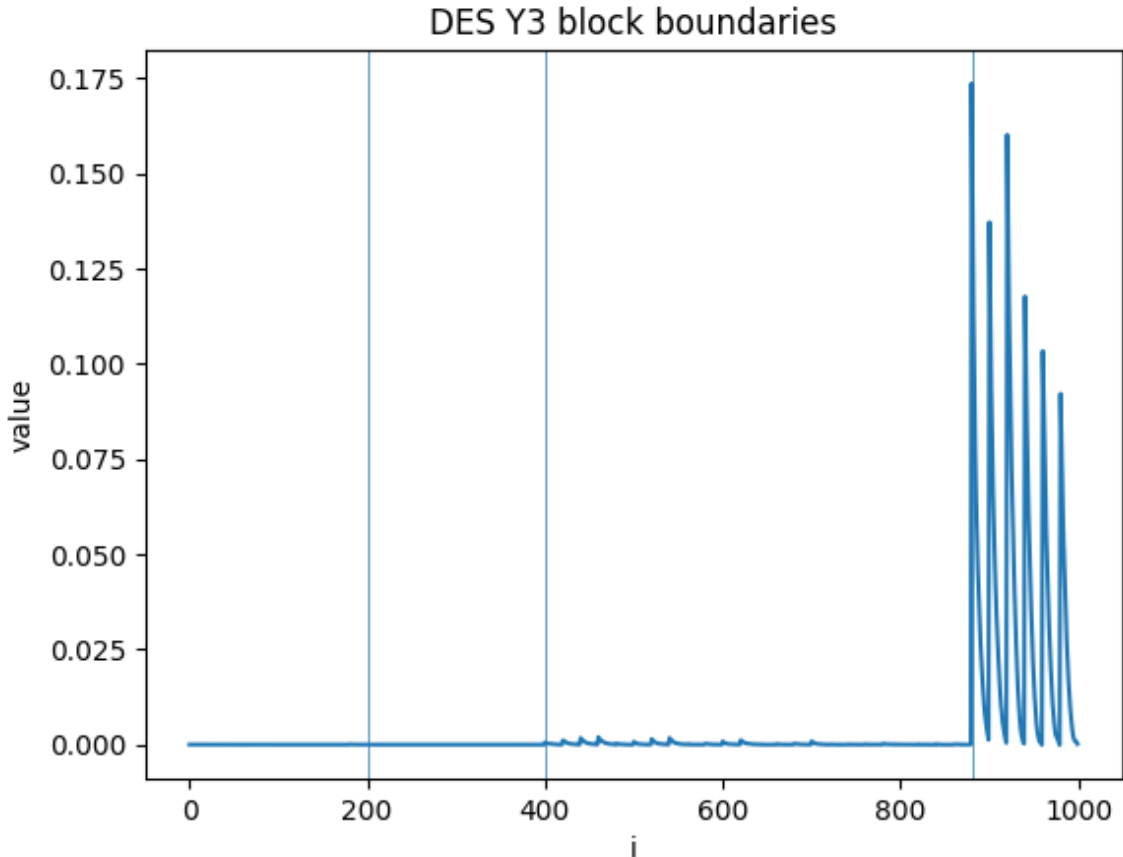


Figure 9: DES Y3 2pt vector diagnostic: block boundary visualization used to validate the per-element block index (probe/statistic segments) that defines our probe splits and masks.

Use of AI assistants. Portions of the codebase were written and edited with the assistance of ChatGPT (OpenAI) and Claude Code (Anthropic). All AI-generated code was reviewed, tested, and validated by the author. The scientific content, analysis design, and interpretation are solely the author’s.

A Supplementary: DES Y3 scale-cut staging sweep under surrogate backend

This appendix records the DES Y3 scale-cut sweep performed as a staging (P4) demonstration on the public 2pt vectors. We progressively remove small-angle (or high- ℓ) points via a minimum-scale cut and re-run the probe-split PPD evaluation for each cut. In the reported sweep, the optimal cut value coincides for both macro families and both directions ($\operatorname{argmin} x_{\min} \approx 25.41$ in the run configuration), and the resulting curves should be interpreted cautiously because the surrogate backend can induce degeneracies not present in a physical theory engine. In particular, identical optima across models are plausibly an artifact of the surrogate degeneracy and should not be interpreted as a physical staging signature. We therefore treat this sweep as confirming that the *staging machinery is in place* (block map \rightarrow scale-cut mask \rightarrow refit \rightarrow held-out metrics), while deferring physical interpretation of staging dependence to a future backend-enabled analysis.

B Reproducibility

This appendix describes how to reproduce the evidence artifacts (plots, tables, and metrics) used in the manuscript. The repository is designed so that (i) external datasets are fetched through a registry with provenance and caching, (ii) each experiment produces a self-describing run bundle (configuration, metrics, provenance, plots), and (iii) the manuscript uses only *vendored* figures and tables that are tracked in the repository. The repository is available at <https://github.com/ioannist/six-birds-de> [25].

B.1 Paper build

From the repository root, the manuscript PDF is built by:

```
make paper
```

The build writes `docs/paper/build/sixbirds_dark_energy.pdf`.

B.2 External datasets: registry, caching, and provenance

External datasets are managed through a registry and fetch utility to prevent silent drift and to preserve provenance. Dataset fetches are described by a YAML registry (see `data/registry.yaml`) and executed via:

```
python scripts/fetch_data.py --dataset <dataset_key>
```

Each fetched dataset is stored under `data/raw/<dataset_key>/` and accompanied by a provenance record (URL, expected hash if provided, observed hash, byte count, timestamp, platform, and Python version). Re-fetching a dataset results in a cache hit when hashes match.

B.3 Experiment run bundles (manifest system)

Every experiment script writes a run bundle of the form `results/<exp_name>/<timestamp>_<gitsha>/` containing a resolved `config.yaml`, quantitative `metrics.json`, `provenance.json`, and at least one plot or table. Each run bundle also includes a short `RUN_LOG.md` summary used as lab notes (see `docs/experiments/`).

B.4 From run bundles to paper artifacts (vendoring)

The manuscript includes figures and quantitative tables only from tracked “vendored” artifacts: `docs/paper/figures/` and `docs/paper/tables/`. Vendoring is performed via:

```
python scripts/vendor_figures.py
python scripts/vendor_tables.py
```

Each vendoring step writes a provenance YAML file recording, for each included artifact, the originating run folder, source file, output path, hash, and timestamp: `docs/paper/figures/provenance.yaml` and `docs/paper/tables/provenance.yaml`. This ensures that the manuscript can be audited against the exact experimental outputs used at the time of writing.

B.5 One-command evidence suites and metrics aggregation

For convenience, the repository provides umbrella targets that regenerate the public-data evidence suites (assuming required datasets have been fetched):

```
make exp-public-evidence-background
make exp-public-evidence-lss
```

A repository-wide aggregation of experiment metrics can be generated by:

```
make collect-metrics
```

which writes `results/_aggregate/metrics_table.csv` as a flattened table over all discovered `metrics.json` files.

B.6 How to reproduce all figures

Starting from a clean clone:

```
git clone https://github.com/ioannist/six-birds-de.git
cd six-birds-de
pip install -e ".[dev]"
python scripts/fetch_data.py --all           # download public datasets
make exp-all                               # run all experiments
python scripts/vendor_figures.py           # copy figures to paper dir
python scripts/vendor_tables.py           # copy tables to paper dir
make paper                                 # build PDF
```

Environment: Python \geq 3.10; tested on Ubuntu 22.04 with NumPy 1.26, SciPy 1.12, Matplotlib 3.8, Astropy 6.0. Random seeds are set at the top of each experiment runner (see `run_*.py` files). Runtime: the full experiment suite completes in under 30 minutes on a 4-core laptop; no GPU is required.

B.7 Evidence mapping

Table 9 maps each plot and quantitative table in the manuscript to its generating script/command, dataset key(s), and the run bundle used, together with the tracked vendored artifact included in the paper build.

Table 9: Evidence map for plots and quantitative tables used in the manuscript. “Dataset key(s)” refer to entries in the repository data registry (synthetic experiments use N/A). Run folders are manifest bundles under `results/`. Vended artifacts are the tracked files included by the paper build.

Artifact	Script / command	Dataset key(s)	Run folder and vended file
Fig. 1	<code>experiments/run_toy1.py</code>	N/A (synthetic)	<code>results/toy1/20260202T175625Z_nogit</code> <code>docs/paper/figures/fig_toy1_rm_summary_vs_beta.png</code>
Table 2	<code>scripts/vendor_tables.py</code>	N/A (synthetic)	<code>results/toy1/20260202T175625Z_nogit</code> <code>docs/paper/tables/tab_toy1_summary.tex</code>
Fig. 2	<code>experiments/run_toy2_patch.py</code>	N/A (synthetic)	<code>results/toy2_patch/20260202T182630Z_nogit</code> <code>docs/paper/figures/fig_toy2_Q_vs_time.png</code>
Table 3	<code>scripts/vendor_tables.py</code>	N/A (synthetic)	<code>results/toy2_patch/20260202T182630Z_nogit</code> <code>docs/paper/tables/tab_toy2_summary.tex</code>
Fig. 3	<code>experiments/run_infer_distance.py</code>	N/A (synthetic)	<code>results/infer_distance/20260202T190430Z_nogit</code> <code>docs/paper/figures/fig_infer_distance_fit.png</code>
Fig. 4	<code>experiments/run_rewrite_vs_lambda.py</code>	N/A (synthetic)	<code>results/rewrite_vs_lambda/20260202T194715Z_nogit</code> <code>docs/paper/figures/fig_rewrite_vs_lambda_distance_fit.png</code>
Table 4	<code>scripts/vendor_tables.py</code>	N/A (synthetic)	<code>results/rewrite_vs_lambda/20260202T194715Z_nogit</code> <code>docs/paper/tables/tab_synth_inference_compare.tex</code>
Fig. 5	<code>experiments/run_ppd_synthetic.py</code>	N/A (synthetic)	<code>results/ppd_synthetic/20260202T201136Z_nogit</code> <code>docs/paper/figures/fig_ppd_synth_rmse_across_seeds.png</code>
Fig. 6	<code>experiments/run_fit_des_sn5yr.py</code>	<code>des_sn5yr_distances</code>	<code>results/fit_des_sn5yr/20260203T092216Z_nogit</code> <code>docs/paper/figures/fig_des_sn5yr_mu_residuals.png</code>
Fig. 7	<code>experiments/run_fit_des_y6_bao.py</code>	<code>des_y6_bao_release</code>	<code>results/fit_des_y6_bao/20260203T075814Z_nogit</code> <code>docs/paper/figures/fig_des_y6_bao_alpha_likelihood.png</code>
Fig. 8	<code>experiments/run_ppd_background.py</code>	<code>des_sn5yr_distances,</code> <code>des_y6_bao_release</code>	<code>results/ppd_background/20260203T094137Z_nogit</code> <code>docs/paper/figures/fig_ppd_background_bao_alpha.png</code>
Table 5	<code>scripts/vendor_tables.py</code>	<code>des_sn5yr_distances,</code> <code>des_y6_bao_release</code>	<code>results/rewrite_background_vs_lcdm/20260203T101804Z_nogit;</code> <code>results/fit_des_y6_bao/20260203T075814Z_nogit</code> <code>docs/paper/tables/tab_background_fits.tex</code>
Table 6	<code>experiments/run_y3_ppd_probe_split.py</code>	<code>des_y3a2_datavectors_2pt</code>	<code>results/y3_ppd_probe_split/20260204T070918Z_nogit</code>
Fig. 9	<code>experiments/run_des_y3_blockmap.py</code>	<code>des_y3a2_datavectors_2pt</code>	<code>results/des_y3_blockmap/20260204T045551Z_nogit</code> <code>docs/paper/figures/fig_y3_block_boundaries.png</code>
Table 7	<code>experiments/run_y3_ppd_with_geometry.py</code>	<code>des_y3a2_datavectors_2pt,</code> <code>des_sn5yr_distances,</code> <code>des_y6_bao_release</code>	<code>results/y3_ppd_with_geometry/20260204T080238Z_nogit</code>
Table 8	<code>experiments/run_kids450_ppd_probe_split.py</code>	<code>kids450_vuitert18_3x2pt_bandpowers</code>	<code>results/kids450_ppd_probe_split/20260204T083352Z_nogit</code>

References

- [1] Adam G. Riess, Alexei V. Filippenko, Peter Challis, et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116(3):1009–1038, 1998.
- [2] S. Perlmutter, G. Aldering, G. Goldhaber, et al. Measurements of omega and lambda from 42 high-redshift supernovae. *The Astrophysical Journal*, 517(2):565–586, 1999.
- [3] David H. Weinberg, Michael J. Mortonson, Daniel J. Eisenstein, Christopher Hirata, Adam G. Riess, and Eduardo Rozo. Observational probes of cosmic acceleration. *Physics Reports*, 530(2):87–255, 2013.
- [4] Ioannis Tsiokos. Six birds: Foundations of emergence calculus, 2026. arXiv:2602.00134.
- [5] Ioannis Tsiokos. To wake a stone with six birds: A life is a theory, 2026.
- [6] Thomas Buchert. On average properties of inhomogeneous fluids in general relativity: Dust cosmologies. *General Relativity and Gravitation*, 32:105–125, 2000.
- [7] Thomas Buchert. On average properties of inhomogeneous fluids in general relativity: Perfect fluid cosmologies. *General Relativity and Gravitation*, 33:1381–1405, 2001.
- [8] Chris Clarkson, George F. R. Ellis, Julien Larena, and Obinna Umeh. Does the growth of structure affect our dynamical models of the universe? the averaging, backreaction and fitting problems in cosmology. *Reports on Progress in Physics*, 74(11):112901, 2011.
- [9] Syksy Räsänen. Backreaction in the Lemaitre–Tolman–Bondi model. *Journal of Cosmology and Astroparticle Physics*, 2004(11):010, 2004.
- [10] Edward W. Kolb, Sabino Matarrese, and Antonio Riotto. On cosmic acceleration without dark energy. *New Journal of Physics*, 8:322, 2006.
- [11] Stephen R. Green and Robert M. Wald. How well is our universe described by an FLRW model? *Classical and Quantum Gravity*, 31(23):234003, 2014.
- [12] David L. Wiltshire. Cosmic clocks, cosmic variance and cosmic averages. *New Journal of Physics*, 9:377, 2007.
- [13] Thomas Buchert, Mauro Carfora, George F. R. Ellis, Edward W. Kolb, Malcolm A. H. MacCallum, Jan J. Ostrowski, Syksy Räsänen, Boudewijn F. Roukema, Lars Andersson, Alan A. Coley, and David L. Wiltshire. Is there proof that backreaction of inhomogeneities is irrelevant in cosmology? *Classical and Quantum Gravity*, 32(21):215021, 2015.
- [14] DES Collaboration. The dark energy survey: Cosmology results with ~1500 new high-redshift type ia supernovae using the full 5-year dataset. *The Astrophysical Journal Letters*, 973(1):L14, 2024.
- [15] The DES SN Working Group. Des supernova 5yr data release, July 2024. Version v1.
- [16] DES Collaboration. Dark energy survey: A 2.1% measurement of the angular baryonic acoustic oscillation scale at redshift $z_{\text{eff}} = 0.85$ from the final dataset. *Physical Review D*, 110:063515, 2024.

- [17] DES Collaboration. Dark energy survey year 3 results: Cosmological constraints from galaxy clustering and weak lensing. *Physical Review D*, 105:023520, 2022.
- [18] Edo van Uitert, Benjamin Joachimi, Shahab Joudaki, Alexandra Amon, Catherine Heymans, Fabian Köhlinger, Marika Asgari, Chris Blake, Ami Choi, Thomas Erben, et al. Kids+gama: Cosmology constraints from a joint analysis of cosmic shear, galaxy–galaxy lensing and angular clustering. *Monthly Notices of the Royal Astronomical Society*, 476(4):4662–4689, 2018.
- [19] Kids survey data products, 2026. Accessed 2026-02-04.
- [20] David W. Hogg. Distance measures in cosmology. *arXiv e-prints*, 1999.
- [21] Hirotugu Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.
- [22] Gideon Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464, 1978.
- [23] Andrew Gelman, Xiao-Li Meng, and Hal Stern. Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*, 6:733–807, 1996.
- [24] Daniel J. Eisenstein and Wayne Hu. Baryonic features in the matter transfer function. *The Astrophysical Journal*, 496:605–614, 1998.
- [25] Ioannis Tsiokos. Six birds: Dark energy repository, 2026. Source code and reproducibility bundle. Zenodo archive: <https://doi.org/10.5281/zenodo.18595154>.