A Hierarchical Bayesian Framework for Cosmology using AGN Variability

Júlia Laguna-Miralles [1,3] Vasily Belokurov [1] Miles Cranmer [1,2,3]

[1]Institute of Astronomy, University of Cambridge [2]DAMTP, University of Cambridge

[3]Kavli Institute for Cosmology, University of Cambridge

arXiv:2606.28501

From AGN Variability to Cosmology

A new route to a high-z Hubble diagram.

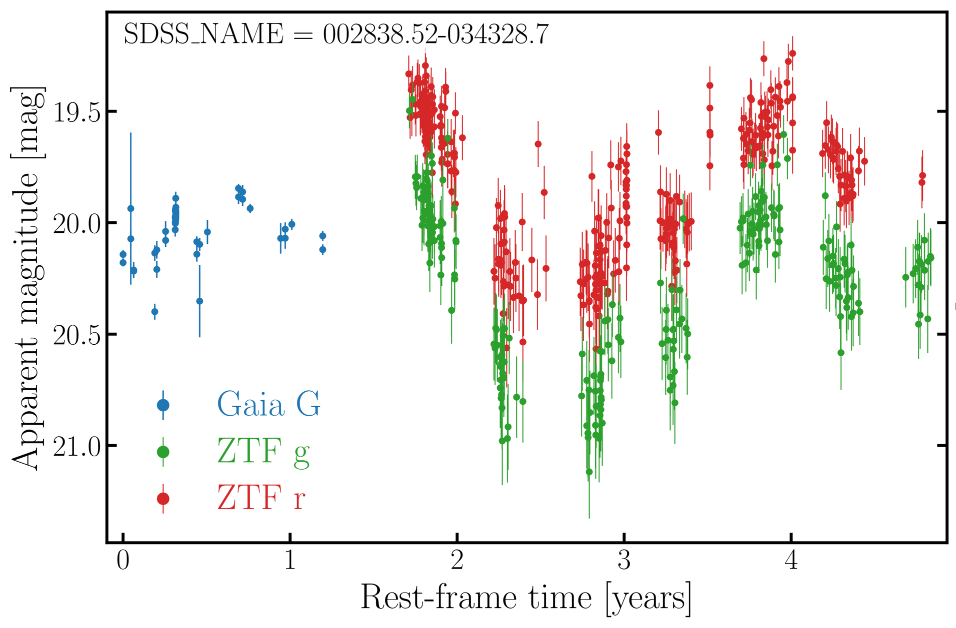
- **The $z > 2$ gap:** Type Ia SNe thin out, leaving the expansion history across the dark energy onset constrained mainly by BAO and the CMB.
- **AGN can fill it:** Abundant to $z \sim 4$, but not standard candles: a distance needs an observable tracking intrinsic luminosity.
- **UV–optical variability is that observable:** it anti-correlates with luminosity and requires only epoch photometry, scalating to entire surveys.
- **The challenge is statistical:** The relation is broad and wavelength-dependent: calibration requires joint inference of luminosity and cosmology from noisy survey light curves.

Active Galactic Nuclei (AGN)

Artist's impression of quasar J0529−4351.

Credit: ESO/M. Kornmesser

Light curves in different bands



We model average apparent magnitude m and variability.

Light curves available for $> 10^5$ AGN.

Hierarchical Bayesian Model

Multi-band light curves
Gaia G, ZTF g, ZTF r

↓

Stage 1: Light curve modelling
Gaussian Process fit, once per AGN

↓

Stage 2: Joint population inference
Variability–luminosity relation $\boldsymbol{\theta}_{\rm VL}$, intrinsic scatter $\sigma_{\rm int}$ and cosmology $\boldsymbol{\theta}_{\rm cos}$

↓

High-z Hubble diagram
Ω_m, w_0, w_a

Each light curve is fit once, then reused for every population fit — a survey-scale approach scalable to 10^5–10^6 sources.

Light Curve Modelling

We model each AGN light curve with a **Damped Random Walk** kernel [1]:

$$k(\Delta t) = \sigma^2 \exp\left(-\frac{|\Delta t|}{\tau}\right),$$

where Δt is the time lag, σ^2 is the long-timescale variability amplitude, and τ is the characteristic decorrelation timescale.

For the baselines of current surveys, σ^2 and τ are strongly degenerate. We show that the short-timescale variability measure can be robustly recovered instead (Figure 3):

$$D \equiv \frac{\sigma^2}{\tau}\ [\mathrm{mag}^2\,\mathrm{day}^{-1}]$$

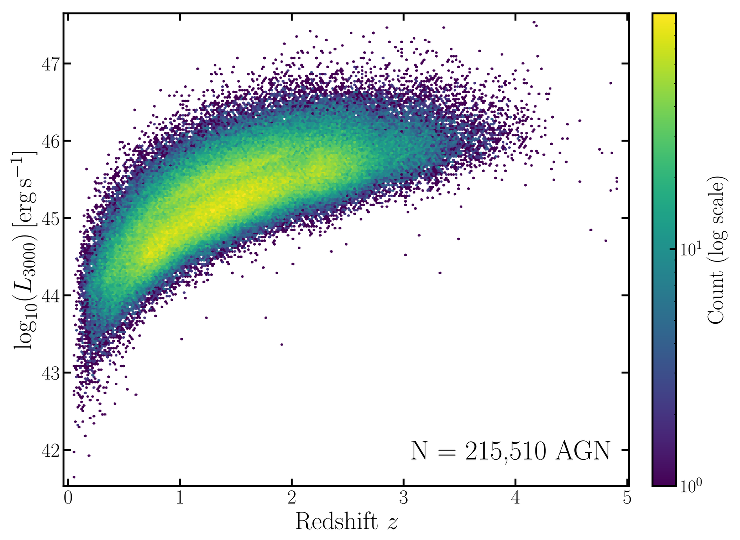


Figure 1. **Luminosity–redshift distribution:** SDSS AGN with Gaia G and ZTF g, r light curves.

Framework: Variability as a Cosmological Probe

Why variability tracks luminosity

Driver	Trend	Physical origin
Time lag $\Delta t_{\rm rf}$	longer → stronger	Variability power accumulates with timescale, saturating at τ.
Wavelength $\lambda_{\rm rf}$	bluer → stronger	Short λ probes the hot inner disc: smaller radii, faster and larger fluctuations.
Luminosity L	brighter → weaker	The emitting region grows with L, lengthening τ, so less power accumulates at fixed $\Delta t_{\rm rf}$.

The variability–luminosity relation is a size–luminosity relation observed in the time domain [2], tied to the accretion-disc scaling that underpins reverberation mapping.

We build an empirically based model

Hyperparameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_{\rm cos},\ \boldsymbol{\theta}_{\rm VL},\ \sigma_{\rm int})$ $\quad \boldsymbol{\theta}_{\rm VL} = (\eta_D,\ d_D,\ \alpha_D)$

Stage 1 — once per AGN: light curve $y_i \longrightarrow$ posterior draws $(m_i,\ D_i)$

Stage 2 — for each sampled cosmology

$$m_i,\ z_i,\ \boldsymbol{\theta}_{\rm cos} \implies \ell_i \equiv \log_{10} L_{3000,i}$$

$$\log_{10} D_i \mid m_i, z_i, \boldsymbol{\theta} \sim \mathcal{N}\left[\mu_{D,i}(\boldsymbol{\theta}),\ \sigma^2_{\rm int}\right]$$

$$\mu_{D,i}(\boldsymbol{\theta}) = \eta_D + d_D\ \ell_i + \alpha_D\ \log_{10}\lambda_{{\rm rf},i} \qquad \text{(simplified)}$$

Estimating ℓ_i requires a distance, so $\boldsymbol{\theta}_{\rm cos}$ calibrates the very relation used to constrain it → we infer them jointly.

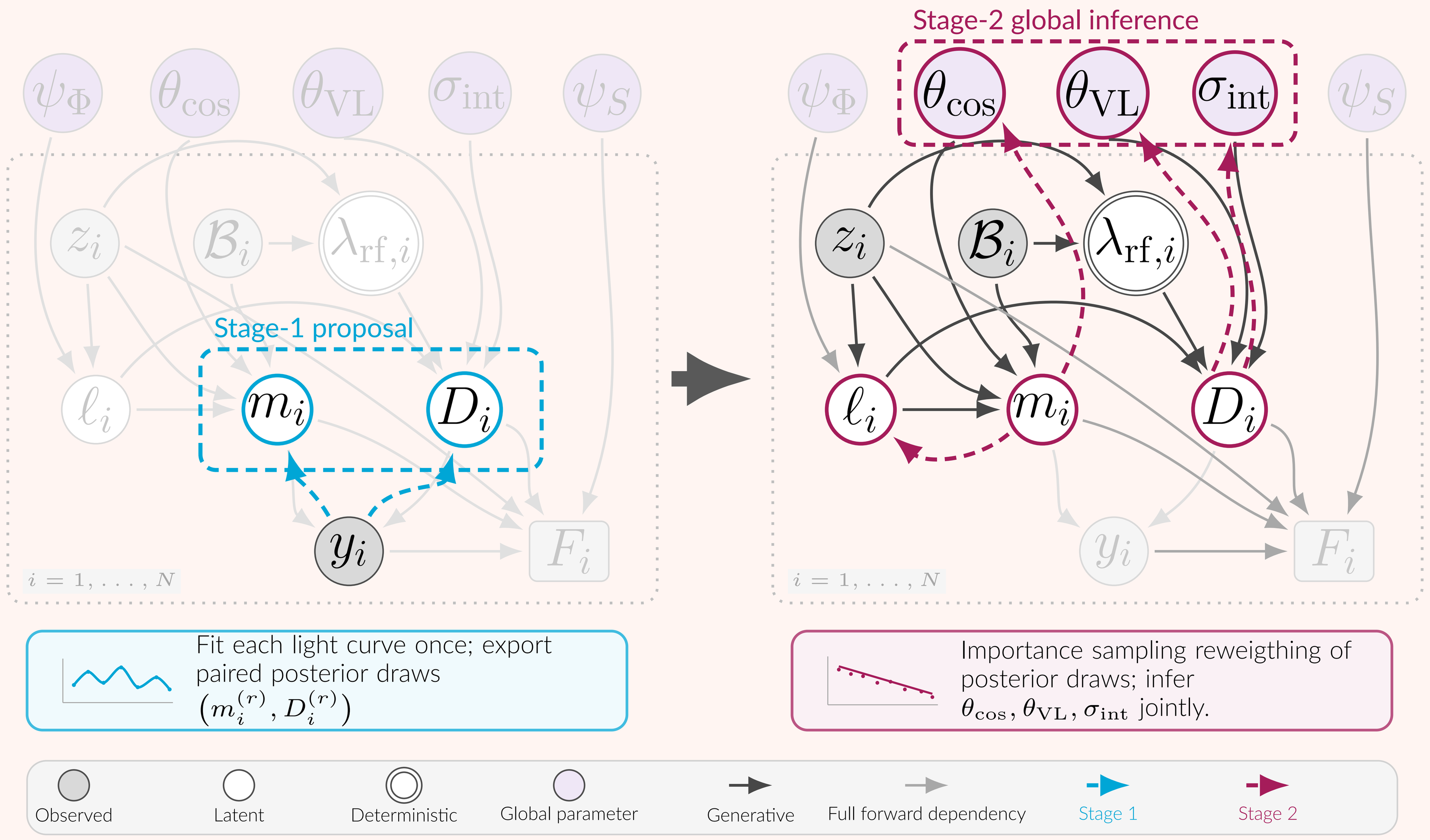


Figure 2. Probabilistic graphical model for AGN variability in grey, with the two-stage inference overlaid in colour: per-object light-curve fits (left), then importance-reweighted population inference (right).

Results

Stage 1: light curve fits

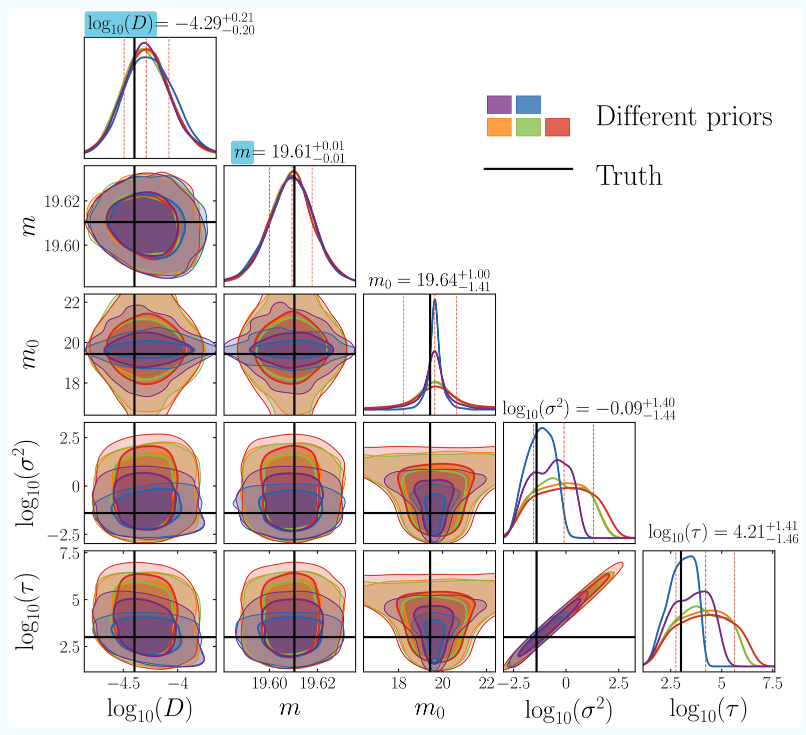


Figure 3. **What survives short baselines.** DRW posteriors for one mock Gaia G light curve under different priors. Individually, σ^2 and τ are degenerate and strongly prior-dependent; the short-term measure D is recovered robustly and almost independently of the prior. Point estimates are biased in this regime, so the *full posterior draws* — not summary statistics — are passed to Stage 2, propagating light curve uncertainty into the cosmological fit.

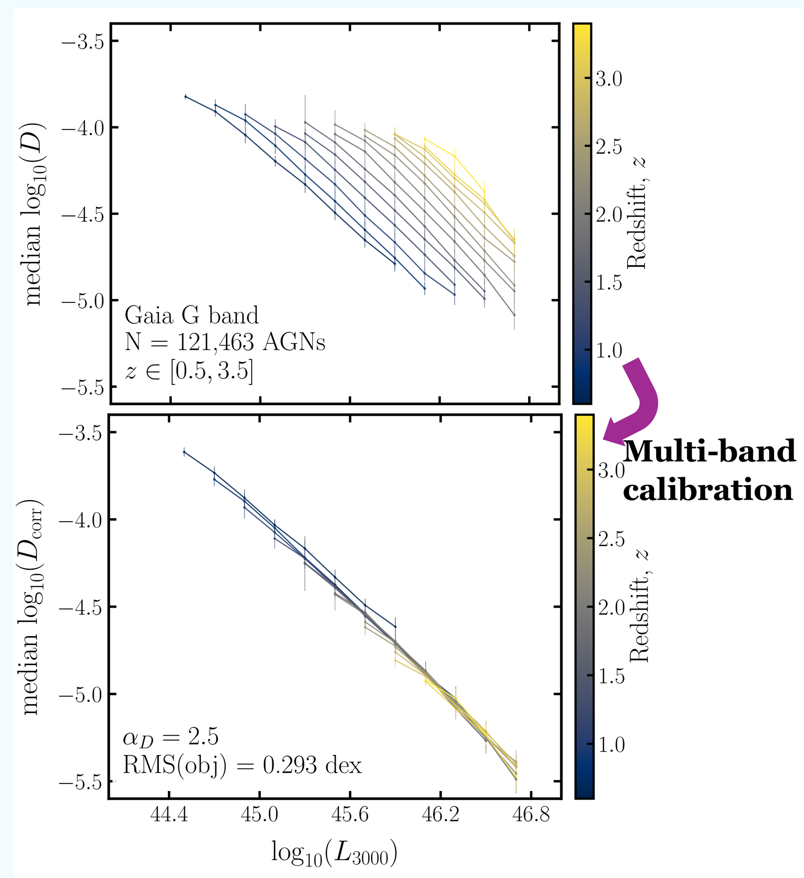
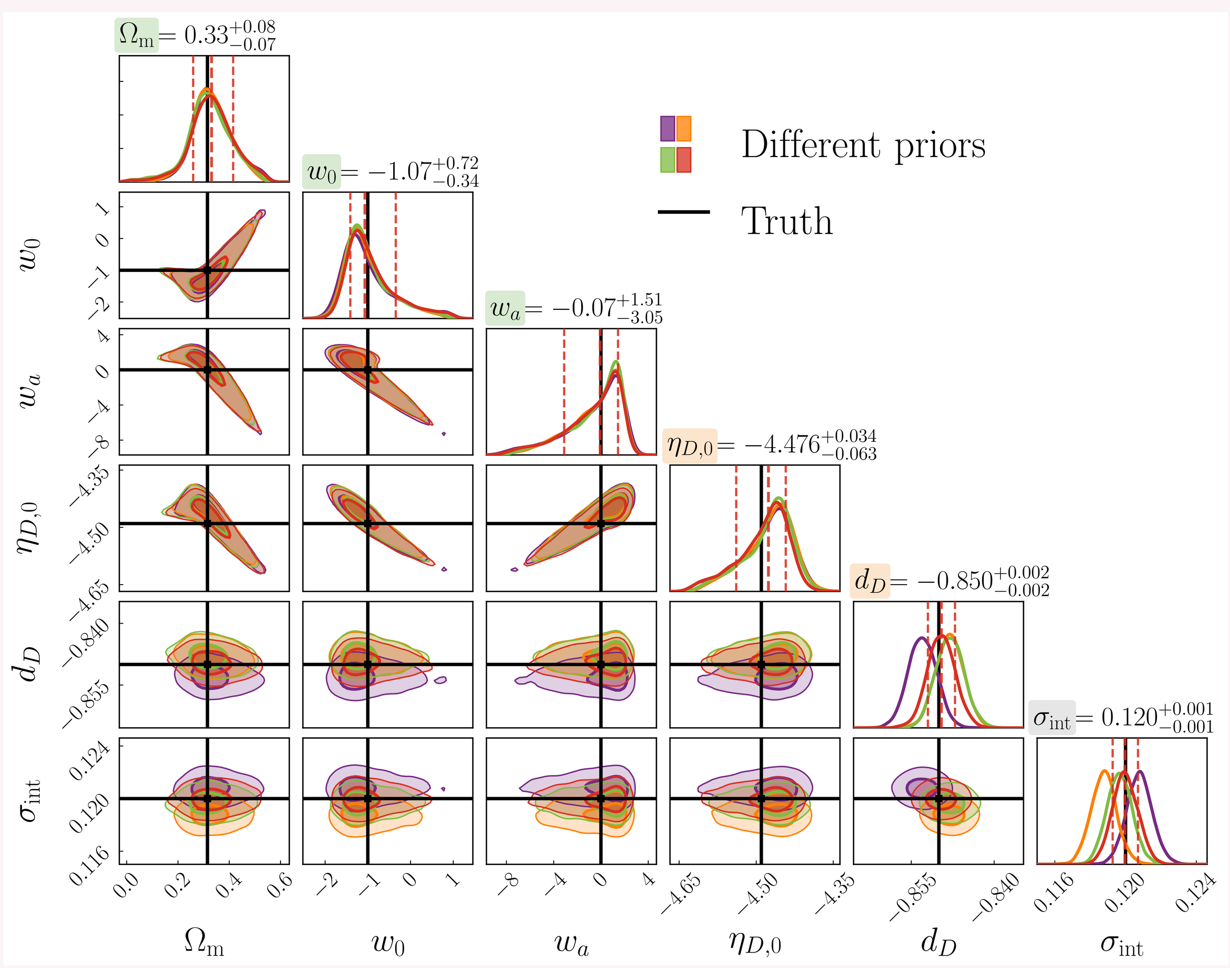


Figure 4. **Variability-luminosity anticorrelation: why a single band is not enough.** Top: At fixed observed band, rest-frame wavelength $\lambda_{\rm rf}$ shifts blueward with redshift, imprinting redshift-dependent structure on the raw relation. Bottom: toy variability wavelength correction with α_D, calibrating its value requires multi-band light curves or multi-epoch spectra — in one band it is degenerate with cosmology.

Stage 2: population inference

Figure 5. **Full closure test.** Mock light curves are simulated from a known cosmology and passed through both stages, with selection effects included. Cosmological, variability–luminosity and intrinsic-scatter parameters are recovered jointly, with injected values shown for reference. Repeated across mock realisations and quality cuts, with different broad Stage-1 priors (as per legend) and wide top-hat Stage-2 priors.



Summary

A validated, survey-scale route to high-z distances. End-to-end closure tests jointly recover the injected cosmology, relation and scatter at Gaia-like cadence and noise.

- **A robust observable.** $D = \sigma^2/\tau$ is recoverable from current survey baselines.
- **Wavelength matters.** Chromatic variability reshapes the relation and, in a single band, mimics cosmological evolution.
- **Demonstrated at scale.** Each light curve is fit once and reused for every cosmology — validated on $\sim 10^5$ sources, scalable to 10^6.
- **Hubble diagram shape.** H_0 is exactly degenerate with the intercept $\eta_{D,0}$; an external anchor is required, from Type Ia SNe or low-redshift AGN.
- **Next.** Multi-band light curves or multi-epoch spectra are required to break the wavelength–cosmology degeneracy.

References

1. Kelly+2009(ApJ 698, pp. 895–910), 2. Caplar+2017(ApJ 834, p. 111).

JLM acknowledges support from "la Caixa" Foundation (ID 100010434; LCF/BQ/EU24/12060091).

jl2268@cam.ac.uk