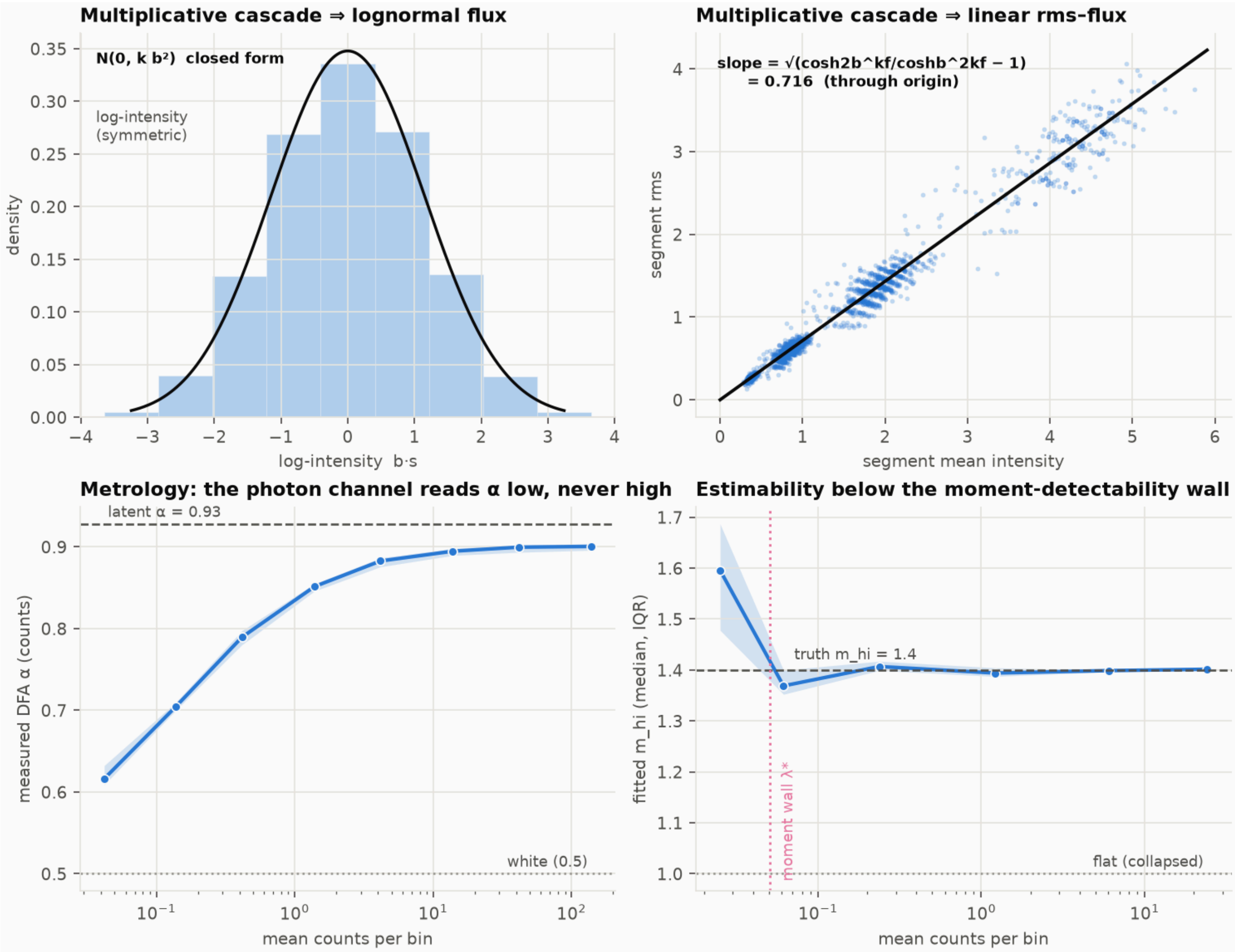
The Markov-switching multifractal as an estimable propagating-fluctuations model

A correspondence and a metrology for photon-limited variability

Evan Tabak Atlas (ORCID 0009-0007-7374-2338)

Draft — methods paper. Target venue: MNRAS or ApJ (methods); arXiv first (astro-ph.HE / astro-ph.IM). Every quantitative claim traces to a named, immutable results record under `out/` *; the closed forms are asserted against simulation in* `tests/test_msm_theory.py` *(module* `fenocosm/msm_theory.py`*). Figure:* `papers/figures/fig_correspondence.png`



papers/figures/fig_correspondence.png

, regenerable by `python -m fenocosm.paper_figures`*.*

Abstract

The broadband aperiodic variability of accreting compact objects carries three robust, model-independent fingerprints: a lognormal flux distribution, a linear relation between the local root-mean-square amplitude and the local mean flux (the “rms–flux relation”), and 1/f-family temporal persistence. The standard generative reading of these fingerprints is the *propagating-fluctuations* picture — multiplicative modulations injected across a range of radii combine to produce a lognormal, rms–flux-linked, red-noise light curve. That picture is usually realised as an unobservable latent field and summarised by its power spectrum; it is rarely fit as a *state-space model with an exact likelihood*.

We show that the **Markov-switching multifractal (MSM)**, a multiplicative regime-cascade introduced in econometrics by Calvet & Fisher, is an estimable member of the propagating-fluctuations class. Coupled to a Poisson counting channel (the **MSM-Cox** process) it (i) reproduces all three accretion fingerprints against matched surrogate controls; (ii) admits *closed forms* for the lognormal marginal and the linear rms–flux slope, each a pure function of the cascade contrast, which we derive and verify by simulation; and (iii) is recoverable by an exact 2^k-state forward-filter maximum-likelihood estimator from as few as ~500 photons — below the count rate at which the same signal is detectable in the count moments.

Our substantive methodological contribution is a **metrology**: an *attenuation atlas* that quantifies, as a calibration curve over the observable count rate, how the Poisson channel biases the measured persistence exponent downward. The bias is strictly negative, monotone, and saturating — persistence “reads low, never high” — reaching ≈0.3 in the exponent at photon-starved rates. Any cross-source comparison of persistence at different count rates is confounded without this correction. We close with the estimability result and its misspecification reading: an apparent likelihood collapse on real streams is a symptom of a wrong *emission* model, not of photon starvation, and can be diagnosed by validating the emission channel rather than abandoning the fit.

All results are reproducible from a fresh clone behind a deterministic apparatus gate.

---

1. Introduction

1.1 Three fingerprints and one generative picture

Three properties recur across the aperiodic X-ray variability of X-ray binaries (XRBs) and the optical variability of active galactic nuclei (AGN):

1. **Lognormality.** The flux distribution is skewed and approximately lognormal; the logarithm of the flux is close to Gaussian (Uttley, McHardy & Vaughan 2005).
2. **The rms–flux relation.** Measured on short segments, the variability amplitude (rms) scales *linearly* with the mean flux of the segment, with a near-zero intercept (Uttley & McHardy 2001; Uttley, McHardy & Vaughan 2005). This is a signature of *multiplicative*, not additive, variability: additive noise has an rms independent of the mean.
3. **1/f-family persistence.** The power spectrum is red, often approximately $1/f^{\alpha}$ over a broad band, with detrended-fluctuation-analysis (DFA) exponents $\alpha > 0.5$.

The dominant generative reading is the **propagating-fluctuations** model of Lyubarskii (1997): independent multiplicative fluctuations, stirred in at a range of disc radii on their local viscous timescales, multiply together as matter drifts inward. A product of many independent modulations is lognormal by the multiplicative central limit theorem; because the modulations multiply, the amplitude scales with the level, giving the rms–flux relation; and a superposition of fluctuations over a wide range of timescales gives broadband red noise. The picture is quantitatively successful, but it is normally expressed as a latent field summarised by its power spectral density (PSD) and compared to data at the level of the PSD or a small set of variability statistics.

1.2 The gap: an estimable member of the class

What the field's default *fitted* models capture is narrower. The damped random walk (DRW; a continuous-time AR(1)/Ornstein–Uhlenbeck process) introduced to quasar light curves by Kelly, Bechtold & Siemiginowska (2009) is a *linear-Gaussian* state-space model with an exact Kalman likelihood. It captures red noise with a single break timescale, but it is additive: it produces neither a lognormal marginal nor an rms–flux relation. Surrogate-based methods that *do* reproduce the lognormal marginal and the rms–flux relation — the amplitude- adjusted Fourier transform (Timmer & König 1995) and its non-linear extension (Emmanoulopoulos, McHardy & Papadakis 2013) — are *generators* for hypothesis testing, not *estimable* forward models with a per-datum likelihood.

There is thus a gap between (a) the propagating-fluctuations *picture*, which explains all three fingerprints multiplicatively, and (b) the *estimable* models routinely fit to data, which are linear-Gaussian and reproduce only the red noise. This paper fills the gap with a specific, tractable multiplicative model.

1.3 The Markov-switching multifractal

The **Markov-switching multifractal** (MSM) of Calvet & Fisher (2001, 2004, 2008) is a multiplicative regime cascade. The (log-)intensity is built from *k* independent two-state components ("layers") whose switching rates lie on a geometric ladder spanning fast to slow. Each layer contributes a multiplicative factor; the product is the instantaneous intensity. The MSM is exactly the discrete, estimable realisation of "many multiplicative modulations across a range of timescales": each layer is one radius/timescale in the propagating- fluctuations picture, and the geometric ladder is the range of viscous frequencies.

Crucially, because the latent state lives on a finite 2^k-state space, the MSM admits an **exact forward-algorithm likelihood** — the same object a hidden Markov model provides — and hence maximum-likelihood estimation, BIC model selection, and out-of-sample scoring. Coupling the cascade to a Poisson counting channel,

$$ N_t \sim \mathrm{Poisson}\!\big(\lambda_0\, G_t^{\kappa}\, \mathrm{d}t\big), \qquad G_t = m_{\rm hi}^{\,s_t}, \qquad s_t = \sum_{i=1}^{k}\sigma_{i,t}, \qquad \sigma_{i,t}\in{-1,+1}, $$

gives the **MSM-Cox** process: a doubly-stochastic (Cox) point process whose latent intensity is a multiplicative cascade. Here `m_hi` is the per-layer multiplier, `κ` the intensity exponent (fixed by an identifiability convention), and λ_0 the baseline rate. The log-intensity contrast per unit net state is `b = κ ln m_hi`, so `G_t^κ = exp(b s_t)`. This is the model we study.

1.4 Contributions

1. **A correspondence (§4).** We derive, as *closed forms*, the two multiplicative fingerprints of the MSM-Cox that no linear-Gaussian model produces — the lognormal marginal and the *linear rms–flux slope* — each a pure function of the cascade contrast `b` and layer count. We verify them against simulation in the test suite, and place them beside the vendored second-order theory (mean/Fano; the sum-over-scales autocovariance that gives the 1/f band). We confirm all three fingerprints on MSM-Cox light curves against matched *phase-randomized-linear* and *shuffled* surrogates.

2. **A metrology (§5).** We build the **photon-channel attenuation atlas**: a calibration curve giving the downward bias of the measured DFA exponent as a function of the observable count rate. This is the paper's principal deliverable for practitioners — a correction to apply before comparing persistence exponents across sources of different brightness.

3. **An estimability result and a misspecification reading (§6).** The exact filter recovers the cascade contrast from ~500 photons — down to the moment- detectability wall, where the count moments are already blind — and we show that an apparent estimator "collapse" reflects emission-model *misspecification*, not photon starvation, with a concrete diagnostic.

The MSM-Cox generator, the DFA/MFDFA estimator mirror, and the exact forward algorithm are vendored, with provenance headers, from an existing CC BY 4.0 analysis package (see PROVENANCE.md); the closed forms of §4, the attenuation atlas of §5, the estimability battery of §6, and all figures are new work first written for this repository.

---

2. The MSM-Cox model

We use the notation of the vendored theory module (`vendor/msm_cox_hawkes/theory.py`).

- **Layers.** *k* independent components. Layer *i* carries a signed toggle `σ_{i,t} ∈ {−1, +1}` that flips with per-tick probability `p_i`; the stationary distribution is uniform (`P(σ=+1)=P(σ=−1)=1/2`), and the per-lag autocorrelation is `ρ_i^L = (1 − 2p_i)^L`.
- **Rate ladder.** The `p_i` lie on a geometric ladder, `f_i = f_slow · span^{(i−1)/(k−1)}`, `p_i = 1 − e^{−f_i dt}`, from a slow layer (long dwell) to a fast layer (rapid flicker).
- **Multiplier and intensity.** `G_t = m_hi^{s_t}` with net state `s_t = Σ_i σ_{i,t}`; the Cox intensity is `λ_t = λ₀ G_t^κ = λ₀ exp(b s_t)` with contrast `b = κ ln m_hi`. Per-layer factor `Y_{i,t} = exp(b σ_{i,t})`, so `E[Y_i^r] = cosh(rb)`.
- **Counting channel.** `N_t ~ Poisson(λ_t dt)`.

The canonical synthetic operating point ("astronomy-like") used throughout the signature and atlas experiments is `k=6, m_hi=1.4, κ=1, f_slow=2×10⁻³/bin, span=300, λ₀=5/bin, dt=1`; every pinned constant lives in `fenocosm/protocol.py` and is recorded with the record it produces.

Second-order structure (vendored theory, for context). The count mean and Fano factor are closed forms,

$$ E[N] = \lambda_0\,\cosh(b)^k, \qquad \mathrm{Fano}[N] = 1 + \lambda_0!\left(\frac{\cosh(2b)^k}{\cosh(b)^k} - \cosh(b)^k\right), $$

and, because $\ln G^\kappa = b \sum_i \sigma_i$ is a sum of independent layers, the log-intensity autocovariance is an **exact sum over scales**, $\mathrm{Cov}[\ln G^\kappa_t, \ln G^\kappa_{t+L}] = b^2 \sum_i (1 - 2p_i)^L$. A geometric ladder of rates superposes exponentials with geometrically spaced decay constants, which is ≈1/f (power law) over the band between the fast and slow cutoffs — the standard mechanism for broadband red noise. These are checked `statement → closed form → simulation` in the vendored module and are the basis of the apparatus gate's closed-form mean/Fano test. The **new** closed forms of §4 concern the *marginal* and a *cross-moment*, which the second-order theory does not touch.

---

3. Method: signatures, controls, estimators

All exponents are measured on a single estimator mirror (DFA-1 with dyadic scale ladders; MFDFA width) so that "the same ruler" is applied to model, control, and data. The three signatures are operationalised as:

- **S1 — lognormality.** On window-averaged flux, compare a Gaussian fit to the flux against a Gaussian fit to the log-flux; report the skewness of each and the per-point log-likelihood advantage of the lognormal model, with Kolmogorov–Smirnov (KS) distances.
- **S2 — the rms–flux relation.** Segment the series; per segment compute the mean flux and the *excess* (measurement-noise-corrected) rms; bin segments by mean flux and fit $\mathrm{rms} = a \cdot \mathrm{mean} + c$. For raw counts the correction subtracts the Poisson variance (= the segment mean); for rate data it subtracts the mean squared measurement error. Report the slope a, the binned correlation, and the segment-level rank correlation.
- **S3 — persistence.** DFA-1 exponent α of the raw count series.

Matched controls. For each seed the *same* latent intensity path is processed two ways before Poisson observation: a **phase-randomized linear equivalent** (amplitude spectrum preserved, phases randomized — the Timmer & König linear surrogate; same PSD, Gaussianized marginal, multiplicative phase structure destroyed) and a **shuffle** (marginal preserved, temporal order destroyed). A signature that survives in the MSM-Cox but not in these controls is attributable to the *multiplicative cascade structure*, not to the spectrum or the marginal alone.

---

4. The correspondence: closed forms for the two multiplicative fingerprints

(Record: `out/v0/`, F1. Closed forms and their simulation checks: `fenocosm/msm_theory.py`, `tests/test_msm_theory.py`.)

4.1 Lognormal marginal (S1)

The log-intensity is a rescaled sum of *k* independent signed toggles: $\ln \lambda_t = \ln \lambda_0 + b\, s_t$, with $s_t = \sum_{i=1}^{k} \sigma_{i,t}$ and each $\sigma_{i,t} \in \{-1,+1\}$ stationary-uniform. Hence $s_t = 2B - k$ with $B \sim \mathrm{Binomial}(k, 1/2)$, and the central moments of the log-intensity are exact:

$$ \mathrm{Var}[\ln\lambda] = k\,b^2, \qquad \mathrm{Skew}[\ln\lambda] = 0 \ \text{(exactly, all odd central moments vanish)}, \qquad \mathrm{ExKurt}[\ln\lambda] = -\tfrac{2}{k}. $$

So the log-intensity is **exactly log-symmetric at every k**, with a variance set by the cascade contrast, and by the central limit theorem it tends to a Gaussian as *k* grows (the discreteness, measured by the `−2/k` excess kurtosis, is the only departure). The intensity itself is therefore *asymptotically lognormal* — right-skewed on the level, symmetric in the log. The level moments are also closed, `E[(G^κ)^r] = cosh(rb)^k`, giving a strictly positive level skewness that tends to the lognormal value `(e^v+2)√(e^v−1)`, `v = k b^2`.

Simulation check (`check_lognormal`, seed 101, `k=8, m_hi=1.5, T=3×10^5`): log-intensity variance closed-form 1.3152 vs simulated 1.3234; skewness 0 vs −0.003; excess kurtosis −0.2500 vs −0.2496; level skewness closed-form 4.137 vs simulated 4.145. Asserted in `test_lognormal_marginal`.

On real MSM-Cox light curves (record v0, F1, 8 seeds × 2^{17} ticks, medians): the flux is right-skewed while the log-flux is nearly symmetric — skew(flux) **+1.55 → skew(log-flux) −0.10** — and the lognormal model beats the normal by **+0.237** in per-point log-likelihood, with the KS distance to a Gaussian falling from 0.121 (flux) to **0.017** (log-flux). The phase-randomized linear equivalent, by construction, does *not* become lognormal (its per-point lognormal advantage is **−0.122**, i.e. the normal model is preferred, and its residual flux skew is only +0.49). Lognormality is a property of the multiplicative structure, not of the shared spectrum.

4.2 The linear rms–flux relation (S2)

The rms–flux relation is the multiplicative fingerprint *par excellence*, and the MSM-Cox yields its slope in closed form. Partition the *k* layers, relative to a measurement segment, into a **fast** set (the `k_fast` layers that decorrelate *within* the segment) and a **slow** set (frozen across it). Over a segment in which the slow layers are frozen at net state `s_slow`, the intensity factorises:

$$ \lambda_t = \underbrace{\lambda_0\,e^{b\,s_{\rm slow}}}_{A \ (\text{local level, fixed})} \cdot\ e^{b\, s_{\rm fast},t} $$

Because the fast layers are independent with `E[e^{bσ}]=cosh(b)` and `E[e^{2bσ}]=cosh(2b)`, the within-segment mean and standard deviation are

$$ E[\lambda \mid s_{\rm slow}] = A\,\cosh(b)^{k_{\rm fast}}, \qquad \mathrm{SD}[\lambda \mid s_{\rm slow}] = A\,\sqrt{\cosh(2b)^{k_{\rm fast}} - \cosh(b)^{2k_{\rm fast}}}. $$

Their ratio is **independent of the local level `A`**:

$$ \boxed{\ \text{rms–flux slope} \;=\; \sqrt{\dfrac{\cosh(2b)^{k_{\rm fast}}}{\cosh(b)^{2k_{\rm fast}}} - 1}\ } $$

Thus the local rms is *linear in the local mean, through the origin*, with a slope that is a pure function of the contrast `b` and the number of fast layers. An additive (linear-Gaussian) process has the opposite property — its local rms is independent of the local mean, i.e. slope 0 — which is exactly why the phase-randomized control fails this test.

Simulation check (`check_rmsflux_linear`, seed 202; three fast layers at `p=0.5` over three frozen slow layers, `m_hi=1.5`): closed-form slope 0.716 vs fitted 0.710, intercept +0.008, correlation 0.987. Observed through the Poisson channel with the excess-rms estimator (`check_rmsflux_poisson`): recovered slope 0.710, binned correlation 1.000. Asserted in `test_rmsflux_linear_latent` and `test_rmsflux_slope_poisson_estimator`.

On real MSM-Cox light curves (record v0, F1, medians): the rms–flux slope is **0.674** with binned correlation **0.998**. The phase-randomized linear equivalent gives a much smaller slope (**0.177**, rank-r 0.62); the shuffle's apparent slope (1.135) comes entirely from its heavy-tailed marginal and is disqualified by its α = 0.5 (see §4.3). Only the multiplicative cascade carries the rms–flux relation with a steep slope *and* the correct persistence.

4.3 1/f persistence (S3) and the joint verdict

The DFA-1 exponent of the MSM-Cox counts is **α = 0.859** (median), matched by the phase-randomized linear equivalent (**0.855**, which shares the spectrum by construction) and destroyed by the shuffle (**0.500**, white). The persistence is real red noise, not an artifact of the marginal.

Taken together (Table 1), only the multiplicative cascade carries **all three** fingerprints jointly: the linear equivalent keeps the persistence (α 0.855) but loses the lognormality (Δℓℓ −0.122) and most of the rms–flux slope (0.177); the shuffle keeps a heavy marginal but is white and has no genuine rms–flux relation. This is the Uttley–McHardy–Vaughan multiplicativity argument, reproduced on a generator with an exact likelihood (8/8 seeds).

Table 1 — Accretion signatures on MSM-Cox vs matched controls (record v0, F1; 8-seed medians).

statistic (median)	MSM-Cox	linear equiv.	shuffled
DFA α (counts)	**0.859**	0.855	0.500
rms–flux slope (binned r)	**0.674** (0.998)	0.177	1.135*
flux skew → log-flux skew	+1.55 → **−0.10**	+0.49 → —	—
per-point Δℓℓ lognormal − normal	**+0.237**	−0.122	—
KS: normal → lognormal	0.121 → **0.017**	—	—

* the shuffle's apparent slope is a marginal artifact, disqualified by α = 0.5.

The upper-left panels of `figures/fig_correspondence.png` show the two closed forms of §4.1–4.2 against simulation.

---

5. The metrology: the photon-channel attenuation atlas

(Record: `out/v0/` *, F2,* `x2_channel.json` *.)*

This is the paper's principal methodological deliverable. Persistence exponents are routinely compared across sources and epochs, but the Poisson counting channel — sparse photons on a bright-to-faint gradient — biases the measured exponent. We quantify that bias as a **calibration curve over the observable count rate**, holding the latent process fixed.

Construction. For each of 8 seeds we simulate one latent cascade path (`k=6, m_hi=1.4` , latent log-intensity DFA α = **0.928**) and observe it through the Poisson channel at rates spanning 3.5 decades, from 0.04 to 139 counts/bin. We measure the DFA-1 exponent of each count series.

Table 2 — The attenuation atlas (record v0, F2; 8-seed medians).

counts/bin	0.04	0.14	0.42	1.4	4.2	14	42	139
measured α	0.616	0.704	0.789	0.851	0.883	0.895	0.900	0.900
(latent α = 0.928)								

The law. The bias is **strictly downward, monotone, and saturating**: the measured exponent approaches the latent value from below as the rate rises, settling ≈0.03 below latent above ~10 counts/bin, and falling to ~0.62 (a ≈0.3 deficit) at photon-starved rates. Persistence **"reads low, never high"**: white Poisson noise dilutes the red latent signal, flattening the fluctuation function toward α = 0.5, and it can never *inflate* the exponent. The lower-left panel of `figures/fig_correspondence.png` is this curve.

Why it matters. A persistence exponent measured on a photon-starved series can be biased low by up to ~0.3. Any cross-source comparison of α at *different* count rates is confounded unless the rate-dependent bias is corrected — the atlas is the correction. Concretely: two sources with identical latent persistence but count rates of 0.1 and 10 counts/bin would be measured at α ≈ 0.70 and 0.90, a spurious 0.2 difference that is entirely instrumental.

Gaps. Realistic monitoring has duty-cycle gaps. Under an orbit-like mask (62.5% duty) at 3 counts/bin, the two common quick treatments both bias α mildly *downward* relative to the ungapped truth (α = 0.883): segment concatenation by **−0.014** (to 0.868) and linear interpolation by **−0.009** (to 0.874). Gaps at this duty are a second-order effect next to the count-rate bias, but they add in the same (downward) direction.

This atlas is why every persistence exponent reported by this program travels with its count rate and the "read-low" caveat (a standing rule of the repository).

---

6. Estimability and the misspecification reading

(Record: `out/v0/` *, F3,* `x3_estimability.json` *.)*

The natural worry about fitting a rich latent model to sparse photons is that the likelihood goes flat — that below some rate the data cannot distinguish a cascade from white noise. We test this directly. We simulate true MSM-Cox data (`k=4, m_hi=1.4, κ=1, T=8192`) at falling count rates and recover the contrast by the exact-filter maximum-likelihood estimator (κ fixed at 1 for identifiability; 6 seeds per rate). "Collapse" would show as the fitted `m̂_hi` falling to 1 (a flat cascade).

Table 3 — ML recovery vs photon rate (record v0, F3; 6-seed medians).

counts/bin	0.025	0.061	0.24	1.2	6.0	24
Fano	1.01	1.02	1.13	1.58	3.96	12.9
m̂_hi (median)	1.596	**1.369**	1.407	1.394	1.399	1.401
fraction collapsed (< 1.05)	0/6	0/6	0/6	0/6	0/6	0/6

No collapse. Recovery of the true contrast (`m_hi = 1.4`) holds down to **0.06 counts/bin (~500 photons total)** — right down to the count rate at which the signal has already vanished from the *moments*. The closed-form overdispersion available for detection is `Fano − 1 = λ₀(cosh(2b)^k/cosh(b)^k − cosh(b)^k)` ; setting it equal to the sampling noise `2√(2/T)` predicts a moment-detectability wall at **λ* ≈ 0.051 counts/bin** (marked in the lower-right panel of `figures/fig_correspondence.png`). At the tested 0.06 counts/bin — essentially at that wall — the measured Fano excess is only ~0.02, already below the sampling floor and so undetectable in the moments; yet the *timing* of the sparse events still constrains the regime persistence, and the exact filter reads out `m̂_hi = 1.37`. The estimator sees structure the moments cannot.

The failure mode is not flattening. At the lowest rate (0.025 counts/bin, below λ*) the estimator does not collapse to `m̂ → 1`; it *blows up upward* (median m̂ 1.596, with a wide inter-quartile range) as the variance estimate becomes unstable. Photon starvation makes the estimate *noisy and high-biased*, never spuriously flat. This matters because a spurious *low* reading would masquerade as "no cascade"; the actual failure is loud, not silent.

The misspecification reading. This reframes a phenomenon seen when the same family is fit to real streams under a *mismatched emission model*: the likelihood appeared to collapse. The estimability result says that collapse is **not** photon starvation — on *well-specified* data the filter recovers the cascade from far fewer photons than the moments need. The collapse was emission-model *misspecification*. The operational lesson for astronomy is therefore:

1. **Fit with the exact filter**, not a moment proxy; it extracts regime persistence from sparse event timing.
2. **Validate the emission channel** as the misspecification alarm — resynthesise from the fitted model and check that the surrogate reproduces the observed marginal and rms–flux structure (a statistical Turing test). A failed resynthesis, not a flat likelihood, is the true signal that the *emission* model is wrong.

6.1 The scope of "no collapse": which arena this result is about

(Cross-reference added 2026-08-30; records `out/v32/` and `out/v35/`.)

Everything in §6 is measured in **one arena**: a *binned* Cox channel — a bright, regularly sampled count series of fixed length `T = 8192`, fitted by the exact filter with a correctly specified emission model, where the only thing varied is the photon rate. Within that arena the result stands as stated: the contrast is recovered down to ~0.06 counts/bin, and the failure mode at lower rates is upward blow-up, not flattening.

It is **not** a statement about the **event-altitude** arena — a sparse point process read through a real observation process (an exposure mask, an activity window, a detection threshold), where what varies is the *event count* `N` rather than the rate at fixed span. That arena has its own floor curve, measured separately in records `v20` (synthetic) and `v28` (real-anchored), and its own distinction between two questions this section does not separate:

- **Selection** — is regime structure decidable at all? There the floor is `N*_selection ≈ 800` events at the canonical injected amplitude, and below it a verdict is an estimability statement, not a world statement.
- **Recovery** — is the fitted ladder a measurement of the injected one? Record `v32` read that floor as lying above the whole grid and, on that reading, retroactively demoted every fitted ladder span in this program from a measurement to a model-selection quantity — a reading in visible tension with §6's "no collapse". Record **`v35`** resolves the tension in this section's favour and against `v32`'s: `v32`'s recovery half was measured under a likelihood that omitted the generator's own activity window, and refitting the same sealed seeds with that covariate restores multiplier recovery *inside* the grid. `v32`'s `N*_recovery` is withdrawn as a misspecification artifact; `N*_selection ≈ 800` stands.

The two arenas remain distinct, and the honest joint reading is the one this paper's §6 already gestures at: **a recovery failure is a statement about the likelihood you wrote down, not about the photons.** In §6 the culprit was the emission model; in `v32` it was an omitted covariate. Neither is photon starvation.

---

7. Reproducibility

Every number above is traced to an immutable, versioned results record under `out/`; the founding record is `out/v0/`, which pins the exact environment, seeds, and per-artifact command lines. The apparatus gate (`./verify.sh`) runs from a fresh clone with only `requirements.txt` installed — DFA anchors, the closed-form mean/Fano check, the provenance-firewall import check, per-experiment smokes, and a ledger citation check that enforces that every headline number in the claim surfaces traces to a record. The **closed forms of §4** are asserted against simulation, at fixed seeds so they cannot flake, in `tests/test_msm_theory.py` (run by `pytest` and in CI). The figure regenerates with `python -m fenocosm.paper_figures`.

```
python3 -m venv venv && venv/bin/pip install -r requirements-dev.txt
./verify.sh                                   # apparatus gate
venv/bin/python -m fenocosm.msm_theory     # §4 closed forms vs simulation
pytest                                        # full fast gate incl. the §4 tests
venv/bin/python -m fenocosm.paper_figures  # papers/figures/*.png
```

---

8. Discussion and limitations

The MSM-Cox is one estimable member of the propagating-fluctuations class, not *the* accretion model. Its value here is methodological: it makes the three multiplicative fingerprints *fittable* with an exact likelihood, so that they can be tested by model comparison rather than by summary statistics, and it comes with a metrology (the atlas) and a recovery guarantee (the estimability battery).

Honest limitations of this methods paper:

- **κ is fixed** by an identifiability convention (`κ ln m_hi` is a ridge); the contrast `b`, not `κ` and `m_hi` separately, is what is estimated.
- The atlas and signature panels are **synthetic**, at a single canonical operating point; the point is calibration and correspondence, not a population measurement. The application to 21 years of Swift/BAT data — where a hierarchy of regime timescales is tested against the DRW and stronger linear competitors, with bootstrap significance and cadence robustness — is the companion result paper (Paper B, `papers/02-tournament-grs1915-cygx1.md`).
- The estimability wall λ* is a **moment-detectability** heuristic, not a Cramér–Rao bound; it is used only to bracket the regime where the filter beats the moments, which it does.
- DFA is one persistence estimator; the atlas is specific to DFA-1 and its scale ladder (though the qualitative "reads low, never high" law is generic to dilution by white counting noise).

---

9. References

(Verified bibliographic details and DOIs; see the citation list. arXiv: astro-ph.HE / astro-ph.IM.)

- **Calvet, L. E., & Fisher, A. J. 2001**, "Forecasting multifractal volatility," *Journal of Econometrics*, 105, 27–58. doi:10.1016/S0304-4076(01)00069-0
- **Calvet, L. E., & Fisher, A. J. 2004**, "How to Forecast Long-Run Volatility: Regime Switching and the Estimation of Multifractal Processes," *Journal of Financial Econometrics*, 2, 49–83. doi:10.1093/jjfinec/nbh003
- **Calvet, L. E., & Fisher, A. J. 2008**, *Multifractal Volatility: Theory, Forecasting, and Pricing* (Academic Press / Elsevier). ISBN 978-0-12-150013-9.
- **Emmanoulopoulos, D., McHardy, I. M., & Papadakis, I. E. 2013**, "Generating artificial light curves: revisited and updated," *MNRAS*, 433, 907–927. doi:10.1093/mnras/stt764
- **Kelly, B. C., Bechtold, J., & Siemiginowska, A. 2009**, "Are the Variations in Quasar Optical Flux Driven by Thermal Fluctuations?," *ApJ*, 698, 895–910. doi:10.1088/0004-637X/698/1/895
- **Krimm, H. A., et al. 2013**, "The Swift/BAT Hard X-ray Transient Monitor," *ApJS*, 209, 14. doi:10.1088/0067-0049/209/1/14
- **Lyubarskii, Yu. E. 1997**, "Flicker noise in accretion discs," *MNRAS*, 292, 679–685. doi:10.1093/mnras/292.3.679
- **Timmer, J., & König, M. 1995**, "On generating power law noise," *A&A*, 300, 707–710. (pre-DOI-era A&A; bibcode 1995A&A…300..707T)
- **Uttley, P., & McHardy, I. M. 2001**, "The flux-dependent amplitude of broadband noise variability in X-ray binaries and active galaxies," *MNRAS*, 323, L26–L30. doi:10.1046/j.1365-8711.2001.04496.x
- **Uttley, P., McHardy, I. M., & Vaughan, S. 2005**, "Non-linear X-ray variability in X-ray binaries and active galaxies," *MNRAS*, 359, 345–362. doi:10.1111/j.1365-2966.2005.08886.x

---

Data availability & acknowledgments

The synthetic experiments (§4–6) require no external data and reproduce from seeds. The latent-vs-observed atlas and the estimability battery are generated by `fenocosm.x2` and `fenocosm.x3`.

Deposit and DOIs. The source repository is **private and remains private**, so this paper's materials are cited by DOI rather than by repository path. The paper itself is deposited at 10.5281/zenodo.22180462, and the analysis package that regenerates every figure and number in it as a software snapshot at 10.5281/zenodo.22180503 — each *(published 2026-08-31)*; the DOIs are live and resolve. The founding record `out/v0/`, which every §4–6 number traces to, travels inside that snapshot's companion record deposits; the event-altitude counterpart of §6's estimability reading is record `v32` as corrected by record `v35` (see the §6 scope note).

 The companion result paper uses public Swift/BAT transient-monitor light curves (Krimm et al. 2013); results provided by the Swift/BAT Hard X-ray Transient Monitor team.