The Compression Scaling Law (CSL) — A Plain-Language Explainer

What problem are we solving?

We want a simple way to tell whether a time series contains deeper, multi-step patterns that go beyond basic variability and simple correlations. CSL does this by comparing how well the data compresses versus a matched surrogate that keeps amplitudes and spectrum but removes higher-order structure.

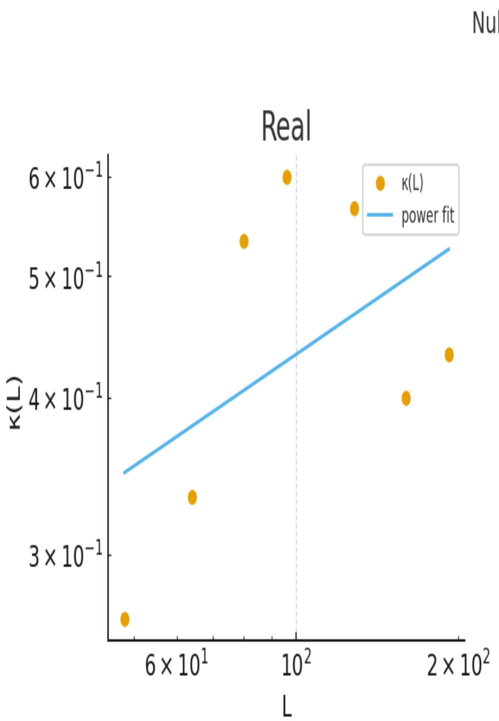
Key idea

If the extra compressibility of a signal grows with window size like a power law, the slope (encoded by α) becomes a robust index of hidden order.

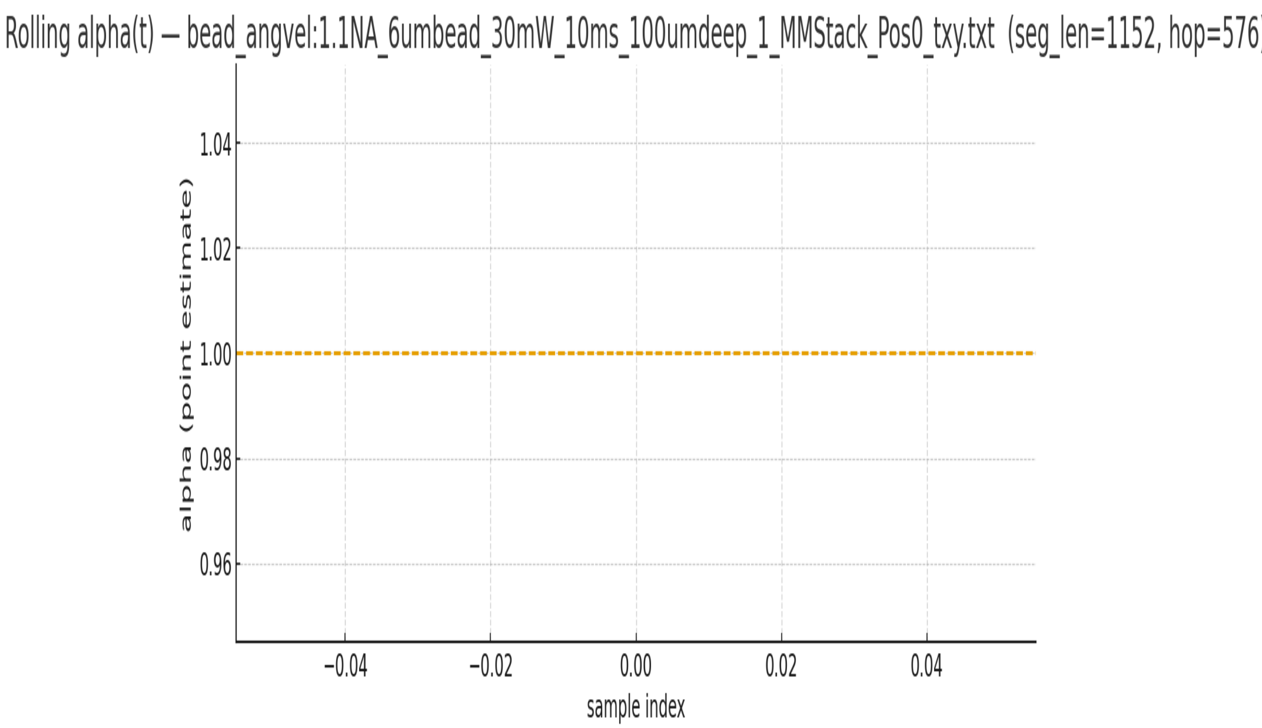
How it works

1) Window the series • 2) Quantize • 3) Compress • 4) Build surrogate • 5) Compress • 6) Contrast code lengths • 7) Fit power law to $\kappa(L)$.

Forest plot: α (95% CI) across datasets and null controls.



Null demonstrations: Real ENSO vs i.i.d. and AR(1) baselines.



Rolling $\alpha(t)$: regime changes appear as shifts below/near α=1.

Reading α

$\alpha \approx 1 \rightarrow$ near-null; $\alpha < 1 \rightarrow$ stronger, scale-reinforcing extra-order; $\alpha > 1 \rightarrow$ rare/divergent.

Why it’s robust

Surrogates remove marginals and linear correlations; slope is stable across lossless coders and quantizers.