Does Market Topology Add Information Beyond Correlation?

1. Question and Mechanism

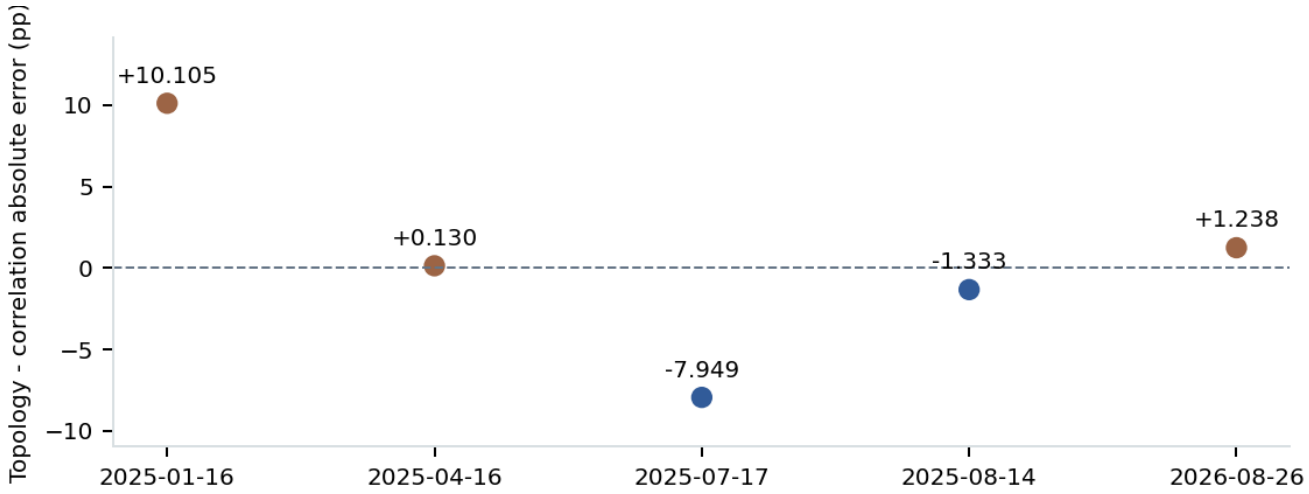
Persistent homology compresses asset relationships into the birth and death of components and loops. Losing asset identity may help compare similar shapes across baskets, while discarding which specific assets are linked. Labelled correlation retrieval is the benchmark for testing whether this compression improves later-return comparisons.

2. Fixed Experiment

Data and Source	Marketstack; AAPL, AMZN, META, MSFT, NVDA, QQQ, SPY, TSLA
Dates and Adjustment	2023-10-13 to 2026-10-12; provider-adjusted OHLCV; 653 aligned observations
Protocol	30 prices; next 20 observations; evaluation stride 20
Past-Only Rule	References and their outcomes must finish before the query begins. At most six separated references, at least three; stability retention threshold 0.5.

Ninety shared dates were excluded for incomplete adjusted records. Horizons count retained observations and have irregular calendar spacing. This data-quality limitation prevents a clean forecasting-advantage claim.

3. Primary Result: the Same Paired Dates



Each point is one paired evaluation date. Negative favors topology. No smoothing.

Across five paired dates, topology MAE was 15.673 pp versus 15.234 pp for correlation: delta +0.438 pp. This small sample does not establish an advantage.

Copyright 2026 Elaine (Yiyao) Liu / 刘一尧 | Personal research project / 个人研究项目

4. Coverage, Ablation and H_1

Method	Scored / Attempted	Coverage	Four-Way MAE (pp)
$H_0 + H_1$	7/23	30.4%	15.673
H_0 only	7/23	30.4%	15.673
H_1 only	22/23	95.7%	15.015
Labelled correlation	11/23	47.8%	15.234

The four-method intersection has 5 date(s). All topology-scored dates: 7/23 (30.4%), MAE 13.998 pp. Topology-correlation paired dates: 5/23 (21.7%). Errors from these two sample sets should be read separately.

H_1 appeared in either channel in 125/624 valid trailing 30-price windows (20.0%): linear 57, rank 88. Full and H_0-only predictions were identical on all 7 shared scored dates, with paired delta MAE 0.000 pp. No predictive increment from H_1 was observed under this protocol.

5. What the Evidence Supports

This run demonstrates auditable representation, benchmark comparison and abstention. Five pairs are below the 20-origin threshold for uncertainty intervals. Frequent abstention may select special market states. Today’s basket, calendar gaps, provider revisions and settings chosen after inspecting data limit generalization. Independent holdout data and transaction costs remain untested.

6. How to Challenge the Hypothesis

First obtain a complete trading-calendar feed with consistent adjustments. Predeclare the universe, time interval and gates. On untouched holdout data, compare correlation, H_0 and full topology, reporting coverage, paired errors and costs. Accept a stronger correlation baseline or no H_1 increment if the evidence supports it.

7. Reproduce and Inspect

Project contribution: research workflow, trading-calendar and horizon audits, benchmark and ablation comparisons, independent numerical checks, and reproducible result review. Export, import and replay on the website to compare data, parameters, references, predictions and errors. Tolerance: absolute 1e-10, relative 1e-9.

The separate fixed-trading-horizon version rejects missing exact 20-session endpoints: 2/23 shared scores (8.7%), insufficient evidence. The original retained-observation archive is unchanged. Do not rank errors across these different date sets. Independent checks passed for 63 diagrams and 84 distance cases.

Dataset SHA-256: 7263ff324a2c7c38f0c34a87c2cbb7bfed906fd49fe2b8eeb76778685423b47b
Case: META 2026-08-31 / 2026-10-12
homology.secondorder.tools/method.html

Mathematical sources: Edelsbrunner & Harer (2010); Zomorodian & Carlsson (2005); Cohen-Steiner, Edelsbrunner & Harer (2007). The method page contains nine full references.

Copyright 2026 Elaine (Yiyao) Liu / 刘一尧 | Personal research project / 个人研究项目