

Supplementary Information: Topological Differences in Early Rumor Propagation

Xuanliang Li¹, Zhaoman Lin², Pengcheng Xu^{2*}

¹School of Financial Mathematics and Statistics, Guangdong University of Finance, Guangzhou, China.

^{2*}School of Data Science and Artificial Intelligence, Guangdong University of Finance, Guangzhou, China.

*Corresponding author(s). E-mail(s): topoxu@gdudf.edu.cn;

Overview

This document contains two supplementary tables that provide detailed experimental results supporting the main text. Table S1 reports the full two-sample permutation test results based on topological representation distances, and Table S2 reports the multi-core detection rates under 12 PRC parameter configurations.

S1. Full Permutation Test Results

Table S1 presents the complete set of two-sample permutation test results for all 10 “window–homology dimension” combinations under three types of topological representation distances. The cell format is “ p_{adj}/T ”, where p_{adj} is the FDR-BH corrected significance level and $T = \bar{d}_{\text{between}}/\bar{d}_{\text{within}}$ is the ratio of the mean between-group distance to the mean within-group distance. Bold values indicate significance after FDR-BH correction ($p_{\text{adj}} < 0.05$); $T > 1$ indicates that the between-group distance exceeds the within-group distance.

The max-life-difference proxy and persistence image L^1 distance serve as practical proxies for the bottleneck and Wasserstein distances, respectively. The main text summarizes the key findings (23/30 significant; all H_0 tests significant; persistence image L^1 significant in all 10 combinations); this table provides the complete numerical results for readers interested in the specific values.

Table S1 Two-sample permutation test results based on topological representation distances

Window-Dimension	Max-Life-Difference Proxy	Persistence Image L^1 Distance	Persistence Landscape L^2 Distance
50, H_0	0.0082 /1.048	0.0082 /1.037	0.0033 /1.044
50, H_1	0.0033 /1.045	0.0033 /1.131	0.8124/0.994
100, H_0	0.0033 /1.059	0.0033 /1.044	0.0033 /1.055
100, H_1	0.0033 /1.061	0.0033 /1.166	0.0898/1.013
200, H_0	0.0033 /1.057	0.0033 /1.038	0.0082 /1.047
200, H_1	0.2354/1.004	0.0033 /1.142	0.2352/1.005
500, H_0	0.0063 /1.060	0.0033 /1.065	0.0033 /1.045
500, H_1	0.0982/1.013	0.0033 /1.218	0.0391 /1.024
1000, H_0	0.0033 /1.116	0.0033 /1.097	0.0033 /1.079
1000, H_1	0.0990/1.012	0.0033 /1.225	0.1308/1.017

Note: Bold indicates significance after FDR-BH correction ($p_{\text{adj}} < 0.05$); $T > 1$ indicates that the mean between-group distance exceeds the mean within-group distance; the max-life-difference proxy and persistence image L^1 distance correspond to proxy implementations of the bottleneck and Wasserstein distances, respectively. All tests use $B = 500$ permutations.

S2. PRC Multi-Core Calibration

Table S2 reports the multi-core detection rates for the rumor and non-rumor groups under 12 PRC parameter grids. The four parameter grids with $\mathbf{n} = 4$ yield zero multi-core detection, indicating that the baseline `numpart` = 2 and coarse partition settings cannot separate locally active propagation segments. The remaining 8 configurations with $\mathbf{n} \in \{6, 8\}$ show detectable multi-core structures, and across all 8, the rumor group consistently exhibits a higher $K_{\text{core}} \geq 2$ proportion than the non-rumor group. This direction is consistent across parameter combinations, supporting the interpretation that rumor propagation is more likely to exhibit multi-core local outbreaks.

Table S2 Multi-core detection rates and rumor vs. non-rumor differences under 12 PRC parameter grids

Configuration	$K_{\text{core}} \geq 2$ (%)	Rumor – Non-rumor
n6.m3.soft	37.88	+0.0434
n8.m3.soft	24.75	+0.0900
n6.m5.soft	17.63	+0.0625
n8.m5.soft	16.50	+0.0750
n8.m3.strict	5.38	+0.0475
n8.m5.strict	3.38	+0.0425
n6.m3.strict	1.75	+0.0150
n6.m5.strict	1.50	+0.0099
n4.m3.soft	0.00	—
n4.m5.soft	0.00	—
n4.m3.strict	0.00	—
n4.m5.strict	0.00	—

Note: All 12 grids are compared on the same 800 event–window samples. The four $n=4$ configurations yield zero multi-core detection, so no rumor/non-rumor difference is defined. Multi-core detection is sensitive to partition resolution: the `numpart=2` baseline and $n=4$ settings are too coarse to separate locally active segments; at $n=6$ or $n=8$, multi-core structures become detectable, with the `soft` criterion consistently stronger than `strict`. The Spearman correlation between K_{core} and TP_{H_1} is 0.2907 across all samples and reaches 0.4648 within the rumor group at `window_50`.