Persistent Homology–Driven Trading Strategies Under Financial Market Anomalies

Hyunjune Kim
hyunjunekim@postech.ac.kr

December 11, 2025

Abstract

This study investigates how persistent homology can be used to detect structural changes in financial time series and build trading strategies. By applying sliding-window embeddings and extracting topological features, we identify anomaly periods that match important market movements. Using these signals, we design several simple PH-based trading strategies and test them on major indices and individual stocks. The results show that the PH-driven strategies often outperform the buy-and-hold baseline, especially when the window size and other hyperparameters are properly chosen. Overall, the study suggests that persistent homology can be a useful tool for understanding market dynamics and improving quantitative trading performance.

1 Introduction

Financial markets are inherently difficult to predict. Their behavior changes through regime shifts and unexpected shocks, leading to irregular and sometimes extreme price movements. Traditional statistical models often struggle to capture these nonlinear transitions, especially during anomaly periods when market structure temporarily deviates from typical patterns. As a result, both quantitative models and human decision-making become less reliable.

To address this challenge, Topological Data Analysis (TDA), particularly persistent homology (PH), offers a way to quantify structural properties of time series beyond raw price information. Rather than treating financial data as simple numerical sequences, persistent homology extracts features such as connectivity, loops, and topological persistence, which may capture hidden structural changes in market dynamics.

In this project, financial time series are transformed into point clouds using a sliding-window method. Persistence diagrams are computed from Vietoris–Rips filtrations, and L_2 distances between diagrams are used to measure structural variation over time. A one-class Support Vector Machine (SVM) is then applied to detect anomaly periods based on these topological signals.

Using the detected anomalies data, several heuristic rule-based trading strategies are developed and tested on ten years of U.S. market data, including major indices (S&P 500, NASDAQ Composite) and selected individual stocks. The goal is to evaluate whether PH–driven anomaly information can outperform a passive buy-and-hold benchmark and offer a more robust decision framework under abnormal market conditions.

2 Related Work

Persistent Homology has its theoretical foundation in algebraic topology, and its modern computational form was established through the work [1], which formalized persistence and later introduced efficient algorithms for computing persistence diagrams. Since then, PH has been widely used in TDA to capture structural patterns in nonlinear data, offering stable representations of high-dimensional systems.

In finance, prior work applied persistent homology to market indices and showed that persistence landscapes and diagram distances increase sharply before major crashes such as the 2000 dot-com bubble and the 2008 subprime mortgage financial crisis [2]. Follow-up research [3] extended this analysis to sector ETFs, confirming that topological signals can reflect global structural changes in financial systems.

Additionally, PH has been explored for unsupervised anomaly detection in time series. Recent research [4] demonstrated that PH-based features can be used to detect abnormal temporal behavior comparable to machine learning baselines. Additional work proposed modifying persistence computation with feature weighting to better adapt PH to real-world time series [5].

However, most existing work focuses on detecting anomalies rather than turning those signals into solid financial actions. That is, persistent homology is rarely used directly for trading decision-making. Instead of stopping at anomaly detection, we integrate PH-based signals into actual heuristic and intuitive rule-based trading strategies and evaluate their performance on real market data.

3 Theoretical Background

3.1 Topological Data Analysis

TDA provides a framework for understanding the shape of data rather than focusing only on numerical values. TDA extracts global structural patterns such as connectivity and loops. This perspective is especially useful when analyzing noisy, nonlinear, or high-dimensional time series, where traditional methods may fail to capture structural transitions. TDA enables us to quantify shape evolution and structural similarity across time windows.

3.2 Persistent Homology

Homology is a central concept in algebraic topology that characterizes a space by quantifying its holes. Homology provides a quantitative description of the structural properties of a space by counting features such as connected components, loops, and voids. For a topological space X, the k-th homology $H_k(X)$ represents k-dimensional holes. Its rank is defined as the Betti number:

$$\beta_k = \mathrm{rank}(H_k(X)),$$

where β_0 counts connected components, β_1 captures loop structures, and higher orders represent high-dimensional voids. Homology serves as the mathematical foundation for persistent homology by providing a way to quantify and compare structural patterns within evolving data.

Persistent homology formalizes how topological features evolve across different geometric scales. Let (X, d) be a metric space and ϵ a filtration parameter. A Vietoris–Rips complex $VR_\epsilon(X)$ is constructed by connecting points whose pairwise distances satisfy following below:

$$d(x_i, x_j) \leq \epsilon.$$

As ϵ increases, topological features such as connected components and loops appear and eventually disappear. Their lifetimes are summarized in a persistence diagram:

$$D = \{(b_i, d_i) \mid i = 1, \dots, k\},$$

where b_i and d_i represent the birth and death ϵ of each feature. Longer persistence generally indicates meaningful structural behavior rather than noise. In this project, H_0 is mainly interpreted as clustering stability, while H_1 reflects loop-like irregularities that may correspond to structural regime shifts in financial markets.

3.3 Support Vector Machine

SVM is a supervised learning algorithm, and in anomaly detection settings, a one-class SVM is used. The model learns a boundary that encloses the majority of "normal" data points while marking distant observations as anomalies. Given training samples $\{x_1, \dots, x_n\}$, the decision function is written as:

$$f(x) = \mathrm{sign}(w \cdot \phi(x) - \rho),$$

where $\phi(x)$ is a kernel mapping and ρ is the learned margin threshold. In this project, the input to the SVM is not raw price data but the logarithmic scale of PH–based distance signal, allowing anomaly detection to depend on structural volatility.

3.4 Financial Anomaly

In financial markets, anomalies refer to short-term behavioral patterns or structural deviations that fall outside expected statistical distributions. These may occur due to macroeconomic events, geopolitical shocks, or sudden regime shifts in market sentiment. One recent example is the global tariff war, where abrupt policy changes led to elevated volatility and regime shifts. Detecting such anomalies early can provide useful signals for adjusting risk exposure.

3.5 Trading Strategy

A trading strategy is defined as a rule-based decision system for entering, exiting, or adjusting market positions. In this project, anomaly signals generated by persistent homology and one-class SVM are translated into multiple versions of intuitive trading rules. These strategies range from simple anomaly avoidance to more adaptive frameworks that adjust re-entry timing or trade execution ratio.

4 Methodology

This section summarizes the overall workflow of the project. We extract topological features from financial time series using persistent homology, convert them into landscape-based anomaly signals, and evaluate how these signals perform through simple rule-based trading strategies. The goal is to see whether PH-driven signals can provide meaningful support for decision-making in financial markets.

4.1 Pipeline

The proposed methodology converts raw financial time series into PH–driven financial anomalies, which are then used as trading signals. The overall pipeline is summarized as follows:

- **Financial Time Series Input**
  Daily closing prices are collected from major U.S. indices and selected individual stocks. Data is transformed into log-returns to ensure scale invariance and reduce trend bias.

- **Sliding-Window Embedding**
  A sliding window W is applied, and each segment is embedded using a delay-coordinate approach, forming a point cloud representation that captures local temporal dynamics.

- **Persistent Landscape Extraction**
  For each point cloud, a Vietoris–Rips filtration is constructed and persistence diagrams are computed. The diagrams are then converted into persistence landscapes to obtain a stable and comparable topological representation.

- **Landscape Distance Signal**
  Structural evolution is quantified using the L^2 distance between consecutive landscapes. This generates a new time-series signal that reflects topological regime shifts rather than price movement.

- **Anomaly Detection Layer**
  The distance signal is log-scaled and standardized before being evaluated through a one-class SVM. A Local Outlier Factor (LOF) model is used as a fallback when the SVM becomes overly sensitive or unstable. This procedure outputs binary anomaly labels aligned with the timeline.

- **Decision and Execution Layer**
  Anomaly labels are translated into rule-based trading signals. Instead of stopping at anomaly detection, the output directly adjusts market exposure to evaluate practical decision-making performance.

The goal of this pipeline is to determine whether persistent homology provides useful anomaly signals that can support more robust trading decisions, rather than relying solely on traditional financial indicators.

4.2 Strategy Design

To evaluate the usability of PH–derived anomaly signals, five increasingly adaptive trading strategies are defined. All strategies share the assumption that detected anomalies correspond to abnormal market states that require risk adjustment.

- **Strategy 0 — Passive Baseline (BH)**
  A benchmark buy-and-hold strategy with full exposure for the entire period. No anomaly information is used and this serves only as a reference baseline.
- **Strategy 1 — Full Exit on Anomaly**
  Upon detecting any anomaly, the portfolio fully moves to cash and returns to full exposure once the anomaly disappears. This represents the strictest interpretation of anomaly-driven risk control.
- **Strategy 2 — Exit on Positive Anomaly**
  Positive anomalies (topologically steep-increase regimes) trigger full exit. Negative anomalies are interpreted as potential local minima, so the strategy initiates re-entry when a negative anomaly occurs. This separates anomalies into "exit signals" and "re-entry signals."
- **Strategy 3 — Delayed Re-Entry**
  Same logic as Strategy 2, but re-entry requires a cool-down period of length T during which no anomaly occurs. This helps avoid whipsaws in highly unstable post-anomaly environments.
- **Strategy 4 — Adaptive Re-Entry**
  Instead of a fixed cool-down T, this strategy assigns an adaptive threshold based on recent volatility. Let σ_{60} be the 60-day rolling volatility and σ_{252} the long-term annual volatility. A volatility ratio determines the required heuristic anomaly-free window length:

  $$T_{\text{adaptive}} = \text{clip}\left(\left\lceil 5 \cdot \frac{\sigma_{60}}{\sigma_{252}} \right\rceil, 3,\ 15\right).$$

  Higher short-term volatility leads to longer required stabilization before re-entry. This implements a dynamic, risk-aware cool-down mechanism.
- **Strategy 5 — Partial Adjustments**
  This strategy performs partial scaling instead of full enter/exit decisions. Buy and sell fractions are selected from a grid search over the range 0.1 to 0.9, and the best-performing pair is chosen. The goal is to identify an optimal partial rule rather than relying on all-in/all-out transitions used in earlier strategies.

4.3 Evaluation Framework

Each strategy is back-tested under identical data conditions, and performance is evaluated primarily through cumulative return. The BH baseline strategy is used as the baseline to determine whether PH–driven signals provide any improvement in return generation. The focus is on whether anomaly-based exposure adjustments yield higher overall performance compared to the passive benchmark.

5 Data

This project uses daily closing price data collected from Investing.com. The full dataset covers roughly 10 years of market activity (2016.12.08–2025.12.07). Only closing prices are used, and all downstream computations (returns, embeddings, landscapes) are derived from these series.

5.1 Market Index

Two major U.S. market indices are selected to represent broad market behavior. Both datasets shares the same 10 years period mentioned above.

- **S&P 500** — A large-cap U.S. equity index capturing overall market performance.
- **Nasdaq Composite** — A tech-focused index with higher exposure to growth and innovation sectors.

These indices provide a stable baseline for understanding how persistent homology responds to macro-level structural changes.

5.2 Individual Stock

To check whether the method generalizes outside broad indices, four individual stocks are included. These also follow the same period but IonQ and Palantir has shorter public company history so that their data is for about 5 years. Additionally, they are chosen based on market relevance and partly on the author's actual portfolio.

- **AAPL (Apple Inc.)** — A global consumer electronics and software company.
- **MSFT (Microsoft Corp.)** — A leading enterprise software and cloud services provider.
- **IONQ (IonQ Inc.)** — A quantum computing company developing ion-trap computing method.
- **PLTR (Palantir Technologies)** — A data analytics and AI-oriented software firm.

Together, these assets create a small but diverse set, allowing us to evaluate whether PH–based signals behave consistently across indices, big-tech stocks, relatively small but innovating tech stocks.

6 Analysis

6.1 Hyperparameter Selection

Choosing appropriate hyperparameters is important because the PH-based signal can change a lot depending on how the embedding and detection stages are configured.

Sliding-Window Embedding: The main hyperparameter is the window size W, which determines how much historical structure is captured at once. Since too-small windows underestimate trend structure and too-large windows smooth out meaningful local variations, we test $W \in \{15, 30, 45, 60, 75, 90\}$ in 15-day increments. Other embedding settings were fixed following standard practices: dimension $\mathrm{dim} = 3$, delay $\tau = 2$, and step size $= 1$.

Anomaly Detection (One-Class SVM): For the PH-derived distance series, we use a One-Class SVM with $\nu = 0.05$ to control the anomaly proportion and $\gamma = \texttt{auto}$ by the sklearn SVM package, which adapts the kernel width to the data scale.

Strategy Hyperparameters: Strategy 5 includes partial scaling of position sizes. To choose reasonable buy/sell fractions, we perform a grid search in 0.1 increments for both parameters. When both fractions equal 1.0, the strategy becomes equivalent to Strategy 4.

6.2 Main Analysis: S&P 500

In this section, we analyze how the topological features extracted from the S&P 500 time series behave under different market conditions. The goal is mainly to check (1) whether the persistent homology structures actually reflect meaningful regime shifts, (2) whether the anomaly detector reacts in a stable and interpretable way, and (3) how much the PH-driven strategies overperforms compared with the baseline.

6.2.1 Persistent Diagram and Barcode

To understand the overall shape of the signal, Figure 1 shows an example persistence diagram and barcode computed from the S&P 500 time-delay embedding with window size $W = 45$. Barcode and diagram shows that H_0 intervals disappear in various time, forming many different sizes of the bars. This is common in noisy high-dimensional embeddings, where points connect to each other and become a single component. In contrast, only a few H_1 intervals appear, and they dissapear quickly but some of them last noticeably longer than the rest.

Overall, the results suggest that the S&P 500 embedding has relatively low topological complexity, with only a few H_1 features. These persistent loops may represent mild oscillatory or recurrent patterns rather than simple upward or downward movements, which is consistent with market conditions during periods of changing volatility.

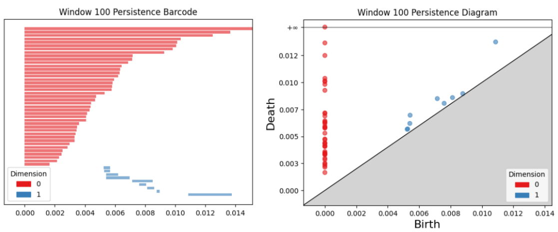

Figure 1: Example Persistent Diagram and Barcode for S&P 500 ($W = 45$)

6.2.2 Full Landscape Overview

Before moving to anomaly detection, we also check how the full set of persistent landscapes evolves over time. Figure 2 shows the landscape values stacked across all windows (again using $W = 45$). The smooth segments generally correspond to stable trend periods, while the rough and rapidly changing areas tend to appear right before volatility expansion.

The landscape plot basically helps us confirm that the distance signal is not random noise. The structure changes gradually most of the time, and only shows sharp shape transitions during market stress. This behavior is important because the SVM should ideally react only to these non-smooth transitions.

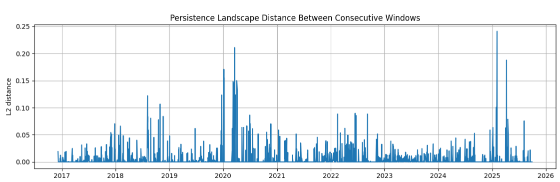

Figure 2: Persistent landscape evolution over time for S&P 500

6.2.3 SVM Score and Anomaly Detection

Figure 3 shows the one-class SVM decision score applied to the landscape distance series. Lower decision scores correspond to higher anomaly likelihood. That is, larger absolute value of decision scores represent more huge financial anomalies. The highlighted red segments on the Figure 3 indicate timestamps where the model flags an anomaly.

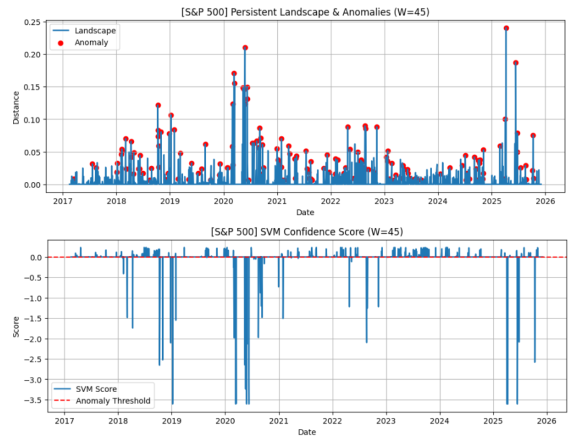

Figure 3: Landscape distance, SVM decision score, and detected anomalies for S&P 500

Figure 4 shows the S&P 500 price series with detected anomalies marked in red. Most anomalies appear near local reversals or during volatility build-up periods, confirming that the detector is reacting to non-stable market segments rather than random noise.

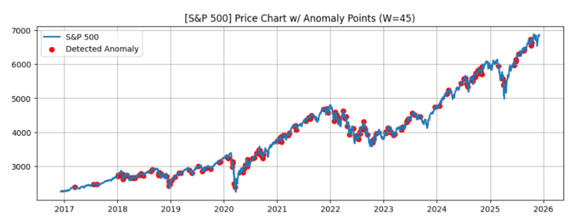

Figure 4: S&P 500 price with anomaly points

6.2.4 Event Mapping for Top Anomalies

To understand whether the detected anomalies correspond to meaningful market events, Table 1 lists the top five anomalous windows based on the SVM decision score (using $W = 45$ for the S&P 500). We then briefly map each window to major market news or structural shifts that happened around that time.

Table 1: Top 5 anomaly windows based on SVM score

Rank	Date	Score	Distance
1	2025-06-13	-3.61	0.19
2	2025-04-09	-3.61	0.24
3	2020-05-26	-3.61	0.21
4	2020-03-12	-3.61	0.17
5	2020-06-11	-3.61	0.13

These top 5 windows have similar SVM score value as -3.61 but these are rounded value and they are actually slightly different. Below is a short summary of what occurred during each anomaly window:

- **Rank 1 (2025-06-13)**: U.S. PPI came in below expectations and Oracle surged on record earnings, triggering a strong tech-led rally.
- **Rank 2 (2025-04-09)**: President Trump's announcement of upcoming tariff escalations, pushing the S&P 500 to its lowest level in a year.
- **Rank 3 (2020-05-26)**: A strong rebound driven by optimism about economic reopening, vaccine progress, and continued Federal Reserve stimulus during the COVID-19 recession.
- **Rank 4 (2020-03-12)**: Panic-driven selloff that followed the WHO's pandemic declaration, leading to the historic 'Black Thursday' market collapse.
- **Rank 5 (2020-06-11)**: This anomaly coincides with the Fed reaffirming near-zero interest rates and the NASDAQ surpassing 10,000 for the first time.

6.2.5 Performances of Strategies

Among all tested window sizes from 15 to 90, $W = 45$ shows the best overall performance for the S&P 500 dataset. Thus, we present the detailed results for $W = 45$ here, while full results for all other window sizes are provided in Appendix A. Additionally, we set initial capital as 1000$ so that we can measure how much each strategies earned. Figures below summarize the performances of Strategies 1 through 5.

We begin with the outcome of Strategy 1 shown in Figure 5. As a result, the baseline BH made $2,918.92 in 10 years starting from $1,000. Strategy 1 made $2,812.70 and underperforms the BH baseline by -3.64%. This is mainly because the strategy fully liquidates its position whenever an anomaly is detected and re-enters only after the system returns to a normal state. In practice, anomaly detection often occurs after the drawdown has already started, meaning the strategy tends to exit the market at disadvantageous moments. Overall, this simple anomaly-avoidance approach behaves similarly to BH baseline but with slightly worse overall performance.

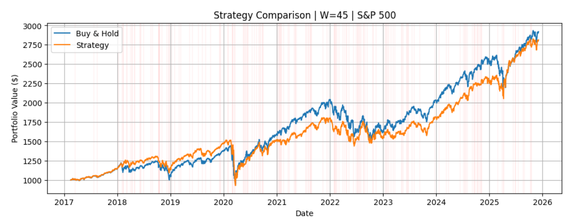

Figure 5: Performance of strategy 1 on S&P 500

Next one is the outcome of Strategy 2 shown in Figure 6. Strategy 2 made $2,859.12 so that it underperforms the baseline by -2.05%. Although this approach improves slightly compared with Strategy 1, simply selling during upward anomalies does not generate a meaningful edge. Upward anomalies often occur during strong momentum phases, so exiting positions at those points can cut profitable trends and result in limited performance gains.

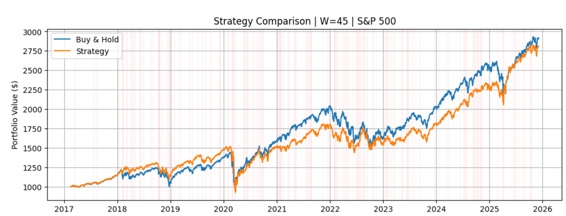

Figure 6: Performance of strategy 2 on S&P 500

Finally, we have our first outperforming strategy. As shown in Figure 7, Strategy 3 made $3,324.54, outperforming the baseline by 13.90%. This strategy incorporates a re-entry delay, allowing the model to wait until the market stabilizes before taking new positions. By avoiding premature re-entry during volatile periods, it captures more sustained trends. This leads to a clear improvement over the BH baseline.

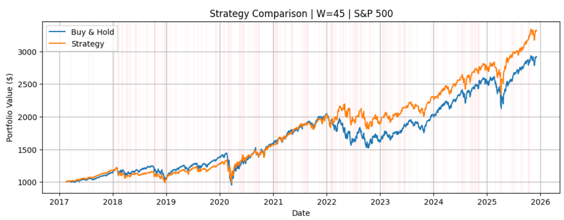

Figure 7: Performance of strategy 3 on S&P 500

Next strategy is the best for S&P 500 when $W = 45$. As shown in Figure 8, Strategy 4 made $3,536.35, outperforming the baseline by 21.15%. Because its re-entry timing is adaptively determined using volatility, the strategy more effectively aligns with actual market conditions. This adaptive mechanism allows it to re-enter during calmer periods and avoid unstable phases so that Strategy 4 can achieve a stronger overall performance.

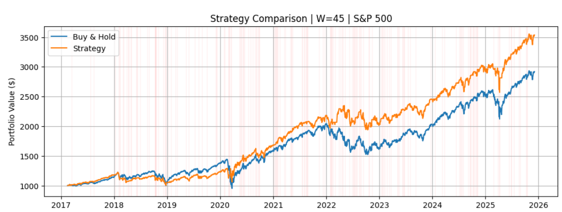

Figure 8: Performance of strategy 4 on S&P 500

As shown in Figure 9, Strategy 5 made $3,481.73, outperforming the baseline by 19.28%. The optimal buy and sell ratio are 0.9 each (it is from grid search from 0.1 to 0.9 and 1.0 is not included because it is totally same with Strategy 4). Although this strategy employs partial buying and selling, such incremental scaling is less advantageous for a large, steadily upward-trending index like the S&P 500. In this setting, the full buy–and–sell approach of Strategy 4 captures upside movements more decisively, resulting in comparatively better performance.

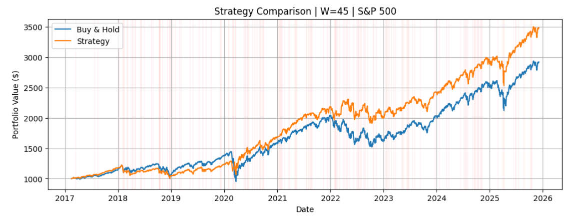

Figure 9: Performance of strategy 5 on S&P 500

Overall result is summarized in Table 2. The results show that simple anomaly-avoidance strategies (S_1, S_2) provide limited benefit and can even underperform due to premature exits. In contrast, strategies that incorporate delayed or volatility-adaptive re-entry (S_3, S_4) demonstrate clear improvements, with Strategy 4 achieving the best performance. Partial scaling (S_5) also works well but remains slightly less effective than full re-entry in a consistently upward-trending index like the S&P 500.

Table 2: Performance overview (S&P 500, $W = 45$)

Strategy	Outcome	Difference	Ratio
S_1	$2,812.70	$-106.22	-3.64%
S_2	$2,859.12	$-59.8	-2.05%
S_3	$3,324.54	$405.62	13.90%
$\boldsymbol{S_4}$	**$3,536.35**	**$617.43**	**21.15%**
S_5	$3,481.73	$562.81	19.28%

6.2.6 Window Size Interpretations

Across all window sizes, $W = 45$ emerges as the most effective, suggesting that this scale best captures the characteristic dynamics of the S&P 500. As shown in the summary Table 3, Strategies 3 and 4 consistently perform well for window sizes up to around $W \approx 60$, whereas beyond $W \approx 75$ the simpler Strategy 2 unexpectedly becomes more competitive. This pattern indicates that medium-range windows paired with adaptive re-entry logic (particularly Strategy 4) align most closely with the market's structural behavior, making the combination of

$W = 45$ and Strategy 4 the most representative for the S&P 500.

Table 3: Window sizes comparison (S&P 500)

W	15	30	45	60	75	90
Best	S_3	S_2	$\boldsymbol{S_4}$	S_4	S_2	S_2
Outcome	11.42%	4.3%	**21.15%**	18.72%	0.6%	6.4%

6.3 NASDAQ

Same pipeline as S&P 500 applied to NASDAQ Composite. Figure 10 shows the best case for NASDAQ Composite data, which is $W = 90$ and Strategy 4. Starting with $1,000 of initial capital, the BH baseline made $3,976.83 for 10 years. Strategy 4 made $5,239.43, outperforming the baseline by 31.75%, which is quite much strong.

Figure 10: Performance of strategy 4 on NASDAQ

As you can see on Table 4, Strategy 4 consistently performed well across most window sizes for NASDAQ, and the best overall result appeared at $W = 90$. This suggests that capturing longer-term structural changes is more effective for the tech-heavy NASDAQ Composite. In other words, larger windows seem to better reflect the broader momentum cycles characteristic of technology-driven markets. In contrast, not too bad result appeared at $W = 30$ and $W = 45$. This is because NASDAQ is also influenced by short-term technological issues. That is, NASDAQ is driven by not only long-term tech tendencies but also short-term issues like tech companies earnings.

Table 4: Window sizes comparison (NASDAQ)

W	15	30	45	60	75	90
Best	S_1	S_4	S_4	S_2	S_5	$\boldsymbol{S_4}$
Outcome	6.03%	23.96%	21.16%	8.98%	12.11%	**31.75%**

6.4 Individual Stock

Based on author's portfolio, Apple, Microsoft, IonQ, Palantir are selected for a experiment. Those stocks can divided into two groups, 'The Big Tech' and 'The Rising Star'.

6.4.1 Apple (AAPL)

In case of Apple, $W = 45$ and ironically Strategy 1 is the best case. Figure 11 shows the result. Starting with $1,000 of initial capital, the BH baseline made $8,285.12 for 10 years. Strategy 1 made $11,345.27, outperforming the baseline by 36.94%.

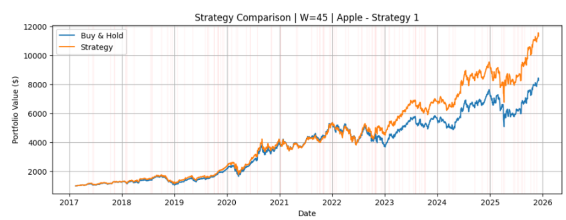

Figure 11: Performance of strategy 1 on AAPL

As summarized in Table 5, AAPL does not exhibit a consistent pattern across window sizes, showing high variability in which strategy performs best. Interestingly, the simplest approach—Strategy 1—produced the strongest result. This may reflect Apple's tendency toward sharp jumps and drops around product releases or macro-level tech news (i.e. trade negotiations with China), where quickly exiting during detected anomalies provides meaningful protection and leads to superior long-term performance.

Table 5: Window sizes comparison (AAPL)

W	15	30	45	60	75	90
Best	S_1	S_5	$\boldsymbol{S_1}$	S_3	S_5	S_1
Outcome	10.15%	-8.5%	**36.94%**	4.96%	-3.24%	19.70%

6.4.2 Microsoft (MSFT)

In case of Microsoft, $W = 90$ and Strategy 3 is the best case. Figure 12 shows the result. Starting with \$1,000 of initial capital, the BH baseline made \$7,241.57 for 10 years. Strategy 3 made \$15,978.12, outperforming the baseline by 113.35%, which is extremely powerful.

Figure 12: Performance of strategy 3 on MSFT

As shown in Table 6, Microsoft consistently favors Strategy 3, and its performance improves as the window size increases. This indicates that Microsoft benefits more from larger windows that capture broader market structure rather than short-term fluctuations. It is because Microsoft is a highly stable big-tech stock with strong long-term trends such as AI trend these days. The combination of long-horizon embedding and delayed re-entry appears to align well with Microsoft's gradual, momentum-driven growth pattern.

Table 6: Window sizes comparison (MSFT)

W	15	30	45	60	75	90
Best	S_2	S_3	S_2	S_3	S_3	$\boldsymbol{S_3}$
Outcome	9.28%	63.11%	31.70%	31.01%	106.68%	**113.35%**

6.4.3 IonQ (IONQ)

In case of IonQ, $W = 30$ and Strategy 5 is the best case. Figure 13 shows the result. Starting with \$1,000 of initial capital, the BH baseline made \$4,308.42 for about 5 years (IonQ became public about 5 years ago). Strategy 5 made \$19,943.35, outperforming the baseline by 362.89%, which is extraordinary powerful result.

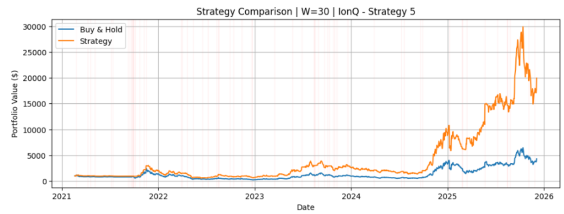

Figure 13: Performance of strategy 5 on IONQ

As shown in Table 7, IonQ consistently favors Strategy 5, with especially strong performance at smaller window sizes. This behavior reflects IonQ's extremely high volatility. That is, short windows around $W = 30$ capture rapid trend shifts more effectively than longer-term structures. As a result, the partial buy–sell mechanism of Strategy 5 aligns well with IonQ's abrupt swings, leading to the exceptional outperformance observed.

Table 7: Window sizes comparison (IONQ)

W	15	30	45	60	75	90
Best	S_3	$\boldsymbol{S_5}$	S_5	S_5	S_5	S_5
Outcome	64.24%	**362.89%**	89.74%	14.23%	40.44%	9.80%

6.4.4 Palantir (PLTR)

In case of Palantir, $W = 45$ and Strategy 4 is the best case. Figure 14 shows the result. Starting with \$1,000 of initial capital, the BH baseline made \$6,147.89 for about 5 years (Palantir became public about 5 years ago). Strategy 4 made \$18,107.58, outperforming the baseline by 194.53%, which is extremely powerful.

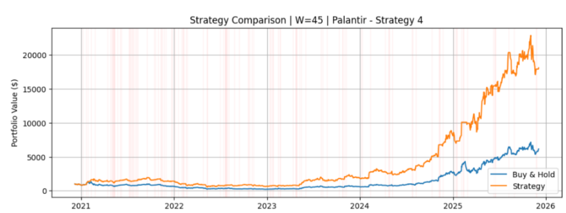

Figure 14: Performance of strategy 4 on PLTR

For Palantir, Strategy 5 generally showed strong and stable performance, with Strategy 4 also performing well in specific cases. Compared to IonQ, Palantir achieved better results with larger window sizes, likely because its volatility is noticeably lower. This makes mid-term structural patterns more important than short-term fluctuations. In particular, at $W = 45$, Strategy 4 delivered the highest return, achieving a 194.53% outperformance over the baseline across the 5-year period. In contrast, very small windows failed to capture meaningful temporal structure and therefore produced weaker results.

Table 8: Window sizes comparison (PLTR)

W	15	30	45	60	75	90
Best	S_5	S_5	S_4	S_5	S_4	S_5
Outcome	-0.46%	-2.00%	194.53%	-6.00%	32.99%	56.03%

7 Discussion

7.1 PH-driven Approach

Across all datasets and window sizes, the PH–driven strategies demonstrate that topological signals can provide meaningful trading advantages when properly configured. Although the simplest approaches (Strategies 1 and 2) often underperformed or behaved similarly to the BH baseline, the more advanced strategies consistently showed stronger results. These strategies attempt delayed re-entry, volatility-aware timing, and partial scaling so that we can avoid unstable conditions. As a result, they frequently outperformed for both broad indices (S&P 500, NASDAQ) and individual stocks.

A key finding is that strategy performance is highly sensitive to the choice of window size used in the sliding-window embedding. Each asset exhibits its own characteristic time scale: medium-sized windows (around $W = 45$) worked best for the S&P 500, very large windows ($W = 90$) were optimal for NASDAQ and MSFT, while smaller window ($W = 30$) aligned better with highly volatile stocks such as IonQ. This suggests that PH–based signals do not offer a universal configuration, but when window size is adapted to each asset's structural dynamics, PH-driven strategies achieve strong and sometimes dramatic levels of outperformance. Selecting asset-specific embedding horizons appears to be a future work for building robust PH-based trading systems.

7.2 Market Index

For the major market indices, the PH-based strategies showed clear and consistent patterns in the S&P 500 and NASDAQ. Medium or large window sizes captured the main market cycles well, which helped Strategies 3–5 avoid large drawdowns and re-enter during more stable phases. The S&P 500 produced stable improvements, while the NASDAQ required larger windows ($W = 90$) because of its stronger momentum. Overall, the results suggest that persistent homology features can reflect meaningful structural changes at the index level.

7.3 Individual Stock

For individual stocks, the PH-based strategies reacted strongly to each stock's volatility and trend behavior. Stocks like IonQ and Palantir, which have clear short term steep movements, worked better with smaller windows that capture shorter structural changes. More stable stocks such as Apple or Microsoft showed clearer PH signals when the window size matched their longer term tendency. Across all names, Strategies 4 and 5 were the most effective and often outperformed the BH baseline. This shows that persistent homology signals can adapt well not only to market indices but also to a wide range of individual equities.

8 Future Work

Future work should explore how can we optimize hyperparameter choices such as window size and the ν value in the One-Class SVM. Our experiments showed that these hyperparameters have a large influence on performance, so selecting them in a more systematic way is important. It would also be useful to test other machine-learning classifiers instead of SVM to see if they provide more stable or more flexible anomaly detection for persistent homology features.

In addition, the trading strategies used in this study are still simple and mainly focus on basic rules. Designing more detailed strategies that use the PH-based signals could lead to better performance. For example, combining multiple persistent homology features and even traditional trading subindexes, or applying more adaptive strategies may help build more robust and practical PH-driven trading systems.

Furthermore, because the window sizes are different and we only use the closing price, the BH baseline sometimes varies across windows. This issue is especially noticeable for IonQ and Palantir, since both companies became publicly traded only about five years ago. As a result, the early high-volatility period is included in the first window for some window sizes but not for others. Future work should address this problem so that the baseline remains consistent across all window sizes, particularly when the initial highly volatile period is involved.

Future work should also apply this pipeline to a wider range of assets. In this study, we focused only on the S&P 500, NASDAQ, and a small set of individual stocks such as AAPL, MSFT, IONQ, and PLTR. Testing the method on assets with different volatility patterns, sectors, and market structures would help determine how general the PH-based signals are. This broader evaluation could show whether the approach works well across many types of financial time series or if it is more effective only for certain categories of assets.

9 Conclusion

In this study, we showed that persistent homology can capture meaningful structural changes in financial time series. By applying sliding-window embeddings and analyzing the resulting topological features, we were able to detect periods that behave like anomalies or regime shifts. These changes often aligned with important market transitions, confirming that persistent homology provides useful information about the underlying dynamics.

Using these PH-based signals, we designed several simple trading strategies and tested them across major indices and individual stocks. The results showed that the PH-driven strategies often performed better than the buy-and-hold baseline, especially when the hyperparameters were chosen appropriately. This demonstrates that persistent homology features can be a practical tool for building effective trading rules and may offer a new direction for quantitative investment research.

10 References

1 Edelsbrunner et. al., "Topological Persistence and Simplification," IEEE, 2000.

2 Gidea et. al., "Topological Data Analysis of Financial Time Series: Landscapes of Crashes," arXiv preprint arXiv:1703.04385, 2017.

3 Gidea et. al., "Topological Signatures in US Indices and Sector ETFs," Journal of Mathematical Finance, 2020.

4 Bois et. al., "A Persistent Homology-Based Algorithm for Unsupervised Time-Series Anomaly Detection," TMLR, 2024.

5 Heo et. al., "Persistent Homology of Featured Time Series Data and Its Applications," AIMS Mathematics, 2024.

A Full Experiment Results

As discussed in the Analysis section, each asset showed a different optimal window size that maximized the performance of the PH-driven strategies. In this appendix, I present the full experimental results across all window sizes and all five strategies. By examining the complete tables, we can clearly observe that PH-based approaches tend to work especially well at specific window sizes that match the structural behavior of each index or stock.

Table 9: Full Experiment Results for S&P 500

S / W	15	30	45	60	75	90
$\mathbf{S_1}$	$3,282.23 8.07%	$2,924.60 -1.97%	$2,812.70 -3.64%	$2,805.13 -3.26%	$2,587.00 -10.66%	$2,867.49 -1.78%
$\mathbf{S_2}$	$3,288.07 8.26%	$3,113.15 4.35%	$2,859.12 -2.05%	$3,035.74 4.70%	$2,914.41 0.65%	$3,106.59 6.41%
$\mathbf{S_3}$	$3,383.98 11.42%	$2,392.87 -19.79%	$3,324.54 13.90%	$3,241.39 11.79%	$2,756.62 -4.80%	$2,490.24 -14.70%
$\mathbf{S_4}$	$2,879.25 -5.20%	$2,690.39 -9.82%	**$3,536.35** **21.15%**	$3,442.42 18.72%	$2,785.91 -3.79%	$2,613.01 -10.50%
$\mathbf{S_5}$	$3,093.87 1.87%	$3,019.05 1.19%	$3,481.73 19.28%	$3,397.18 17.16%	$2,877.73 -0.62%	$2,884.15 -1.21%

For the S&P 500, the full results Table 9 clearly show that PH-driven strategies perform best when using medium-sized windows, particularly around $W = 45$ to $W = 60$. As observed earlier, strategies that incorporate delayed or adaptive re-entry (Strategies 3–5) consistently outperform the baseline within this window size range. These window sizes seem to capture the market's structural patterns more effectively, allowing the topological signal to identify meaningful regime shifts.

Table 10: Full Experiment Results for NASDAQ Composite

S / W	15	30	45	60	75	90
$\mathbf{S_1}$	$4,590.62 6.03%	$3,480.20 -16.25%	$3,193.39 -20.94%	$4,084.51 1.46%	$4,007.86 0.85%	$3,939.88 -0.93%
$\mathbf{S_2}$	$4,460.16 3.02%	$4,485.62 7.94%	$3,588.93 -11.14%	$4,387.14 8.98%	$4,259.15 7.17%	$4,140.48 4.12%
$\mathbf{S_3}$	$3,971.93 -8.26%	$4,878.93 17.41%	$4,823.02 19.41%	$3,979.19 -1.15%	$3,936.70 -0.95%	$4,699.35 18.17%
$\mathbf{S_4}$	$4,096.09 -5.39%	$5,151.10 23.96%	$4,893.63 21.16%	$3,941.59 -2.09%	$4,016.07 1.05%	**$5,239.43** **31.75%**
$\mathbf{S_5}$	$4,300.30 -0.67%	$5,078.70 22.22%	$4,653.73 15.22%	$3,866.59 -3.95%	$4,455.79 12.12%	$5,103.46 28.33%

As you can see from Table 10for the NASDAQ Composite, Strategies 3–5 generally perform best, with especially strong results at both relatively small windows (around $W = 30$–45) and the largest window, $W = 90$. This pattern likely reflects the dual nature of NASDAQ as a tech-heavy index: it reacts quickly to short-term volatility driven by news and earnings cycles, yet it also follows strong long-term momentum trends over extended periods. As a result, both shorter windows that capture rapid fluctuations and longer windows that reflect broader structural movements tend to align well with NASDAQ's overall behavior.

Table 11: Full Experiment Results for AAPL

S / W	15	30	45	60	75	90
S_1	\$10,647.60 10.16%	\$7,289.90 -20.87%	**\$11,345.27** **36.94%**	\$8,050.21 -0.57%	\$6,248.17 -19.91%	\$9,445.74 19.70%
S_2	\$9,229.85 -4.51%	\$6,988.75 -24.14%	\$9,406.55 13.54%	\$7,964.11 -1.63%	\$6,250.02 -19.89%	\$8,318.29 5.41%
S_3	\$9,839.26 1.79%	\$4,937.50 -46.40%	\$7,560.02 -8.75%	\$8,498.10 4.96%	\$6,519.89 -16.43%	\$6,207.43 -21.34%
S_4	\$8,324.56 -13.88%	\$4,289.81 -53.43%	\$7,208.69 -12.99%	\$8,491.01 4.87%	\$6,520.46 -16.42%	\$6,681.19 -15.34%
S_5	\$9,566.77 -1.03%	\$8,429.54 -8.50%	\$8,323.50 0.46%	\$8,436.77 4.20%	\$7,549.00 -3.24%	\$7,761.61 -1.65%

As you can see from Table 11, for Apple, the overall pattern is less consistent compared to the major indices, and no single window size dominates across all strategies. However, one interesting exception appears in Strategy 1, where the $W = 45$ window yields the strongest performance by a large margin.

Table 12: Full Experiment Results for MSFT

S / W	15	30	45	60	75	90
S_1	\$6,871.71 -10.57%	\$5,567.64 -26.27%	\$6,310.38 -15.31%	\$6,045.46 -18.62%	\$5,905.44 -19.30%	\$7,731.43 6.76%
S_2	\$8,396.64 9.28%	\$8,640.29 14.43%	\$9,813.57 31.70%	\$8,389.85 12.94%	\$8,692.61 18.79%	\$11,619.14 60.45%
S_3	\$7,003.41 -8.85%	\$12,316.49 63.11%	\$8,633.16 15.86%	\$9,732.29 31.01%	\$15,124.27 106.68%	\$15,450.05 113.35%
S_4	\$7,089.64 -7.73%	\$11,541.08 52.84%	\$7,319.63 -1.77%	\$9,210.18 23.99%	\$14,593.51 99.43%	**\$15,978.12** **120.64%**
S_5	\$7,611.79 -0.93%	\$11,059.00 46.46%	\$7,639.26 2.52%	\$9,025.73 21.50%	\$13,186.79 80.21%	\$14,505.33 100.31%

As you can see from Table 12, all strategies except Strategy 1 show strong performance, and Strategies 4 and 5 in particular generate the highest returns across most window sizes for Microsoft. A clear trend appears where larger windows (especially $W = 75$ and $W = 90$) lead to significantly better outcomes. This makes sense given Microsoft's profile as a large, stable tech company. That is, Microsoft tends to move by long-term structural trends and persistent momentum play a more important role than short-term fluctuations.

As you can see from Table 13 almost all strategies outperform the baseline for IonQ, and Strategy 5 consistently delivers the strongest results, including the highest overall outperforming rate among all assets tested. Because IonQ is an extremely high-volatility stock with steep short-term movements, smaller window sizes such as $W = 15$ and $W = 30$ capture these rapid fluctuations especially well. This makes the PH-driven approach particularly very effective for assets with strong short-term dynamics like IonQ.

For Palantir, the results Table 14 show a pattern similar to IonQ, with Strategies 3–5 generally outperforming the baseline by large margins. However, unlike IonQ, the strongest performance emerges at a slightly

Table 13: Full Experiment Results for IONQ

S / W	15	30	45	60	75	90
$\mathbf{S_1}$	\$5,693.70 18.22%	\$8,736.96 102.79%	\$7,633.68 60.45%	\$3,502.82 -31.62%	\$6,122.76 17.85%	\$3,887.64 -29.01%
$\mathbf{S_2}$	\$4,881.04 1.35%	\$6,604.33 53.29%	\$6,266.12 31.71%	\$4,851.59 -5.29%	\$6,878.36 32.39%	\$4,763.78 -13.01%
$\mathbf{S_3}$	\$7,910.10 64.24%	\$13,949.21 223.77%	\$5,557.15 16.81%	\$4,734.18 -7.58%	\$6,635.05 27.71%	\$4,786.09 -12.60%
$\mathbf{S_4}$	\$5,621.88 16.73%	\$19,019.74 341.46%	\$5,174.94 8.77%	\$4,079.18 -20.37%	\$6,388.80 22.97%	\$4,384.71 -19.93%
$\mathbf{S_5}$	\$6,056.31 25.75%	**\$19,943.35** **362.89%**	\$9,026.96 89.74%	\$5,851.28 14.23%	\$7,296.67 40.44%	\$6,001.92 9.60%

Table 14: Full Experiment Results for PLTR

S / W	15	30	45	60	75	90
$\mathbf{S_1}$	\$9,418.55 -49.76%	\$4,994.46 -55.65%	\$4,826.58 -21.49%	\$4,937.94 68.44%	\$6,056.10 88.43%	\$5,969.46 11.09%
$\mathbf{S_2}$	\$13,550.35 -27.72%	\$6,177.32 -45.14%	\$6,689.88 8.82%	\$3,742.49 51.87%	\$6,345.79 92.66%	\$5,834.65 8.58%
$\mathbf{S_3}$	\$14,638.89 -21.92%	\$4,932.47 -56.20%	\$1,3247.00 115.47%	\$4,717.31 65.38%	\$9,019.82 131.71%	\$7,455.10 38.74%
$\mathbf{S_4}$	\$11,397.36 -39.21%	\$5,272.94 -53.17%	**\$18,107.58** **194.53%**	\$4,833.26 66.99%	\$9,107.89 132.99%	\$8,322.65 54.88%
$\mathbf{S_5}$	\$18,662.65 -0.46%	\$11,036.61 -1.99%	\$15,737.65 155.98%	\$6,782.07 94.00%	\$8,905.97 130.05%	\$8,384.59 56.03%

larger window size, especially around $W = 45$. Even in smaller window sizes, every strategies failed to capture structural patterns so that they underperformed a lot. This likely reflects the fact that Palantir, while still volatile, is a larger and more established company than IonQ. So we can say medium-range structural patterns matter more than purely short-term fluctuations for Palantir. Overall, the PH-driven strategies remain highly effective for Palantir, particularly in the mid-window range.