Higher-Order Interactions in Financial Networks: An Ablation Study with Simplicial Complexes

João Cláudio Nunes Carvalho, Daniel Alencar Barros Tavares

https://doi.org/10.1590/SciELOPreprints.17512

Submitted on: 2026-08-18
Posted on: 2026-09-03 (version 1)
(YYYY-MM-DD)

Higher-Order Interactions in Financial Networks: An Ablation Study with Simplicial Complexes

João Cláudio Nunes Carvalho — ORCID: https://orcid.org/0000-0001-8619-0869
Daniel Alencar Barros Tavares — ORCID: https://orcid.org/0000-0003-2029-6454
Instituto Federal de Educação, Ciência e Tecnologia do Ceará, Brazil {joao.carvalho, daniel.tavares}@ifce.edu.br

---

Abstract. This paper investigates whether explicitly modeling higher-order interactions improves prediction in financial-network settings when the data-generating mechanism depends on group synergy. We formally define the prediction problem over a known simplicial complex and conduct a controlled ablation study across three topological levels: (i) a 0-simplex baseline using only node-level history, (ii) a 1-simplex graph model using self-history and neighbor aggregation, and (iii) a 2-simplex simplicial model that incorporates a triangle-face interaction term. Using synthetic data in which a target node depends on the multiplicative synergy between two other nodes, we show that the simplicial model consistently outperforms pairwise baselines. At the target node, the reduction in MSE reaches 96.1% in the main scenario and 95.7% ± 0.6% across 20 random seeds, with strong statistical significance (Wilcoxon signed-rank test, $p \approx 1.9 \times 10^{-6}$; paired t-test, $p \approx 4.4 \times 10^{-18}$). Additional analyses using MAE, R^2, correlation, sensitivity to interaction strength, noise and lag depth, a second 2-simplex, a nonlinear MLP baseline and a one-layer GCN baseline indicate that the performance gain is localized in higher-order structures and attributable to the simplicial topology rather than to generic nonlinear capacity. A set of fair pair-identification baselines — including a linear model with the trigger units as separate features, an MLP that can learn the interaction, and a model with products of all unit pairs — shows that the 2-simplex provides both the identification of the relevant interaction and a substantial parsimony advantage (7 features versus 409 for comparable error). A semi-synthetic experiment using real quarterly US macroeconomic series as triggers reproduces the same pattern (99.9% MSE reduction at the target, Wilcoxon $p \approx 1.9 \times 10^{-6}$). The full experimental pipeline, including code, data generation and figures, is publicly available for reproduction. These results provide quantitative evidence for the use of Topological Deep Learning and simplicial complexes in network problems where systemic propagation depends on group-level effects.

Keywords: ablation study, graph mining, higher-order networks, relational learning, simplicial complexes, topological machine learning

---

1. Introduction

In learning over networks, many prediction and mining tasks depend not only on node attributes or pairwise relations, but on higher-order relational patterns [17, 5]. When the behavior of a node depends on the conjunction between two specific neighbors, edge-based models tend to collapse the relevant structural information, because permutation-invariant neighborhood aggregation operators (sum, mean, maximum) preserve neither the identity of the interacting pair nor the product of their states [16, 23]. This scenario is central to graph mining, relational learning, topological machine learning, and higher-order networks [1, 3].

This paper addresses the following research question: **when the generating dynamics depend on multiplicative synergy between two units, does a model with 2-simplices improve prediction over models that use only 0- and 1-simplices?** To answer it in a controlled manner, we conduct an ablation study on synthetic data, comparing three levels of topological complexity: a baseline with individual history (0-simplex), a graph model with neighbor aggregation (1-simplex), and a simplicial model with an explicit triangle-face term (2-simplex). The experimental design isolates the role of the topological representation, separating it from the effect of the architecture and the capacity of the regressor.

The main result is that, when the target variable depends explicitly on the product of two other units, including the 2-simplex term reduces the mean squared error at the affected unit by 95.7% ± 0.6% across 20 random seeds, with strong statistical significance (paired Wilcoxon test, $p \approx 1.9 \times 10^{-6}$). The gain is localized at the higher-order structure and remains present across different metrics, seeds, noise levels, interaction strengths, and regressor specifications. Neither a nonlinear MLP nor a one-layer GCN, both without the face term, reaches comparable performance.

The contributions of this paper are five. First, we formalize the node-level prediction problem over a known simplicial complex and the information sets of each topological level (Section 3). Second, we mechanically isolate a type of dependency that graph models do not represent well under standard aggregation. Third, we propose a reproducible experimental protocol, with public code, to compare 0-, 1-, and 2-simplices in predictive tasks over networks. Fourth, we provide quantitative evidence — including fair baselines in which the trigger pair is identified and the interaction can be learned — that the benefit of the 2-simplex combines structural identification of the relevant interaction with representational parsimony, rather than mere nonlinear flexibility. Fifth, we show that the pattern reproduces in a semi-synthetic scenario in which the triggers are real macroeconomic time series.

As applied motivation, the studied phenomenon appears in closed financial networks, dependency chains, and other relational systems in which specific groups of nodes concentrate risk or influence [4, 2]. In systemic risk, for example, the joint deterioration of two institutions can amplify the effect on a third in a non-additive way, a mechanism that the financial-contagion literature describes as depending on co-occurrence rather than on isolated bilateral channels [4]. Still, the focus of this paper is less on the domain and more on the methodological value of modeling higher-order interactions explicitly, assessing when they produce measurable gains in prediction and interpretation.

The remainder of this paper is organized as follows. Section 2 reviews related work, including the state of the art in graph neural networks and simplicial neural networks. Section 3 presents formal definitions and the problem formulation. Section 4 describes the experimental methodology, the algorithms, the software packages, the hardware, and the processing times. Section 5 presents the results. Section 6 discusses interpretations and limitations. Section 7 concludes.

2. Related Work

2.1. Motifs, substructures, and higher-order organization

The first line of research motivating this study comes from graph mining and relational learning, in particular from work that incorporates motifs, triangles, and local substructures as auxiliary signals for predictive tasks on graphs. Milo et al. [17] showed that motifs — recurring subgraphs with frequency above expectation — act as building blocks of complex networks. Benson, Gleich, and Leskovec [5] extended this perspective by proposing motif-based clustering methods, demonstrating that higher-order organization reveals structure that pairwise connectivity does not capture. Battaglia et al. [1] systematized the role of relational inductive biases in deep learning over graphs. Although useful, these approaches usually preserve the central aggregation mechanism typical of GNNs and do not explicitly isolate when a group interaction is structurally necessary — precisely the gap that our ablation design seeks to fill.

2.2. Graph neural networks

The second line comprises graph neural networks (GNNs), today the dominant family in relational learning. The GCN of Kipf and Welling [16] popularized the simplified spectral convolution with normalized neighbor aggregation; GATs [22] introduced attention in the aggregation; and the message-passing framework [10] unified these variants as iterative message exchange between adjacent nodes. Xu et al. [23] characterized the expressive power of these architectures, showing that aggregation-based GNNs are bounded by the first-order Weisfeiler-Lehman test. This limitation is directly relevant to our study: permutation-invariant aggregation operators over the neighborhood (sum, mean, maximum) cannot, in a single layer, represent the product between two specific neighbors, because aggregation discards the identity of the pair. Our control experiment with a one-layer GCN (Section 5.6) empirically verifies this theoretical prediction.

2.3. Hypergraphs and networks beyond pairwise interactions

The third line uses higher-order networks and hypergraphs to model relations beyond pairs. Battiston et al. [3] comprehensively reviewed the structure and dynamics of networks with interactions beyond pairwise. Hypergraph Neural Networks [9] and HyperGCN [24] extend graph convolution to hyperedges. These models enlarge the relational space but do not necessarily offer the same geometric interpretation as a simplicial complex, in which the closure relation among vertices, edges, and faces is part of the representation itself.

2.4. Simplicial neural networks and Topological Deep Learning

Finally, work on Topological Deep Learning and simplicial neural networks shows that structures such as simplices and cell complexes can enrich representation and learning. Ebli, Defferrard, and Spreemann [8] proposed the first Simplicial Neural Networks based on the Hodge Laplacian; Schaub et al. [21] formalized random walks on simplicial complexes and the normalized Hodge Laplacian; Roddenberry, Glaze, and Segarra [19, 20] built simplicial networks with equivariance principles for trajectory prediction; Bodnar et al. [6] extended the Weisfeiler-Lehman test to simplicial message passing, proving expressivity gains over GNNs; Bodnar et al. [7] generalized the result to cell complexes; Giusti et al. [11] added simplicial attention mechanisms; and Hajij et al. [13] consolidated the research program of Topological Deep Learning, with software support in the TopoX/TopoModelX suite [14].

2.5. Positioning of this work

Our contribution is empirical and methodological: instead of proposing yet another architecture or merely applying a more expressive model, we build a scenario in which the generating mechanism depends explicitly on a multiplicative interaction between two units, allowing the predictive gain to be attributed to the 2-simplex in a controlled, comparable, and statistically testable way. The study therefore serves as a methodological assessment of when higher-order representations are necessary in relational learning problems — a type of evidence that complements theoretical expressivity results [6, 23] with direct measurement of the gain in prediction.

3. Theoretical Background and Problem Formulation

3.1. Definitions

Definition 1 (Abstract simplicial complex). Let V be a finite set of vertices. An abstract simplicial complex K is a family of non-empty finite subsets of V, called simplices, such that: (i) for every $v \in V$, $\{v\} \in K$; and (ii) if $\sigma \in K$ and $\tau \subseteq \sigma$ with $\tau \neq \emptyset$, then $\tau \in K$ (closure under inclusion). A simplex σ with $|\sigma| = k + 1$ vertices is a k-simplex; every proper non-empty subset of σ is a face of σ.

Definition 2 (Skeleton). The k-skeleton of K, denoted $K^{(k)}$, is the subcomplex formed by all simplices of K with dimension at most k. In particular, $K^{(1)}$ is a graph $G = (V, E)$, with E the set of 1-simplices.

In this paper, 0-simplices are the units (nodes), 1-simplices are the edges, and 2-simplices are the triangle faces. We denote by $\Delta(u) = \{\sigma \in K : \dim(\sigma) = 2, u \in \sigma\}$ the set of 2-simplices that contain unit u, and by N(u) the neighborhood of u in $K^{(1)}$.

Definition 3 (Node-level prediction over a known complex). Let X ∈ ℝ^{T×N} be a matrix of time series of N units observed over T time steps, with x_u(t) the value of unit u at time t, and let K be a known simplicial complex over the units. Given a one-step horizon and a lag depth L, the problem consists of learning, for each unit u, a function f_u such that

x̂_u(t) = f_u(I_u(t − 1)),

where $I\_u(t-1)$ is the information set available up to $t-1$, determined by the topological level of the model. Quality is measured by the mean squared error (MSE) on a temporal test partition.

3.2. Information sets per topological level

The ablation study compares three nested information sets, $I^0 \subset I^1 \subset I^2$, one per topological level:

- **0-simplex (baseline):** $I^0\_u(t-1) = \{ x\_u(t-\ell) : \ell = 1, \ldots, L \}$ — only the unit's own history;
- **1-simplex (graph):** $I^1\_u(t-1) = I^0\_u(t-1) \cup \{ m\_u(t-\ell) : \ell = 1, \ldots, L \}$, with $m\_u(t) = (1/|N(u)|)\ \Sigma\_\{v \in N(u)\}\ x\_v(t)$ — adds the neighbor mean, the typical GNN aggregation operator;
- **2-simplex (simplicial):** $I^2\_u(t-1) = I^1\_u(t-1) \cup \{ x\_a(t-1) \cdot x\_b(t-1) : \{u, a, b\} \in \Delta(u) \}$ — adds, for each triangle containing u, the product of the other two vertices at lag 1, i.e., the explicit face term.

Level	Structure	Information used
0-simplex	node	individual history
1-simplex	edges	history + neighbor aggregation
2-simplex	triangle face	history + neighbors + group interaction

Table 1. Topological levels compared in the ablation study.

Since the three models use the same regressor (Ridge) and differ only in the information set, any systematic performance difference is attributable to the topological representation, not to the capacity of the estimator.

3.3. Hypotheses

The experimental design tests three hypotheses. **H1 (localized gain):** when the generating mechanism of a unit contains multiplicative synergy between two neighbors, the 2-simplex model reduces the MSE at that unit relative to levels 0 and 1. **H2 (structural specificity):** the gain concentrates at the units belonging to 2-simplices active in the generating mechanism and does not spread to the others. **H3 (topological attribution):** nonlinear regressors without the face term (MLP, GCN) do not recover the gain, indicating that the cause is the topology of the representation rather than nonlinear capacity.

4. Methodology

4.1. Synthetic data generation

We generate time series for $N = 30$ units over $T = 300$ time steps. Units 5 and 6 receive sinusoidal periodic signals with distinct frequencies and random phases, serving as triggers of the higher-order interaction:

$x_g(t) = \sin(\omega_g t + \varphi_g) + \eta_g(t),\ g \in \{5, 6\},$

with $\eta_g(t) \sim N(0, \sigma^2)$. Unit 7 is constructed to depend explicitly on the product of the two units at the previous lag:

$$\mathbf{x_7(t) = \varepsilon_t + \alpha \cdot x_5(t-1) \cdot x_6(t-1),} \quad (1)$$

where $\alpha = 5$ in the main scenario and $\varepsilon_t \sim N(0, \sigma^2)$ is low-variance noise ($\sigma = 0.1$). This mechanism simulates a synergy effect: the risk at unit 7 increases when units 5 and 6 deteriorate jointly. The remaining 27 units follow AR(1) processes with coefficient 0.5, weak coupling (0.1) to the chain neighbor, and the same noise level.

In the main scenario, the complex contains the triangle {5, 6, 7} as the only 2-simplex, and the 1-skeleton links the remaining nodes in a chain, connected to the triangle to keep the graph connected (Figure 3). In the generalization experiment, we introduce a second 2-simplex {10, 11, 12} with an analogous mechanism at unit 12. We assume the simplicial complex to be known (oracle) to isolate the effect of the representation; the lifting problem — inferring the higher-order structure from observational data — is left as future work and discussed in Section 6.3. This oracle assumption is methodologically important: it avoids conflating two distinct problems, namely **discovering** the higher-order structure and **using** that structure to predict. The study thus cleanly measures the marginal value of the 2-simplex when the relevant topology is known.

4.2. Compared models

We compare three main supervised models and two controls.

Baseline, Graph, and Simplicial. The three topological levels use Ridge regression (ℓ^2 regularization, $\alpha_{ridge} = 1.0$) over the information sets of Section 3.2, with $L = 3$ lags in the main scenario. The choice of a single linear regressor makes the comparison transparent: the models differ only in their input features.

MLP control. A multilayer perceptron (two hidden layers of 64 units, ReLU activation, Adam optimizer, up to 3000 epochs) trained on the graph-level features. The goal is to verify whether a nonlinear regressor, without the adequate topology, would suffice to recover the synergy.

GCN control. A one-layer graph convolutional network (16 hidden units, tanh activation, symmetrically normalized adjacency with self-loops, Adam optimizer with rate 10^{-2}, 800 epochs, own NumPy implementation), trained end-to-end for node-level regression on the lags of all nodes. The goal is to verify whether the learned aggregation of a standard GNN recovers the group interaction.

Fair pair-identification baselines. A legitimate criticism of the design so far is that the simplicial model receives exactly the product $x_5 \cdot x_6$, the same term used to generate x_7, so the gain could merely reflect the injection of the correct feature. To separate the two effects — identifying the relevant interaction and representing it — we add three baselines. **Pair-Ridge** receives the unit's own history and the separate lags of x_5 and x_6 (identified pair, without the product): if the limitation were merely one of information access, this model should suffice. **Pair-MLP** applies the MLP described above to these same features: with the pair identified, a nonlinear approximator can, in principle, learn the interaction. **AllPairs-Ridge** receives the unit's own history and the lag-1 products of all pairs of units in the network (409 features for N = 30): it represents the scenario in which the form of the interaction is known, but the topology — which pair matters — is not.

4.3. Experimental protocol

Algorithm 1 summarizes the pipeline.

```
Algorithm 1 — Ablation study pipeline
Input: configuration (N, T, α, σ, triangles, seed), lags L, test fraction 25%
1: X ← generate_scenario(configuration)          // Eq. (1) and Section 4.1
2: for each level ∈ {baseline, graph, simplicial}:
3:    for each unit u ∈ {0, …, N−1}:
4:       (Φ, y) ← build_features(X, K, u, level, L)
5:       (Φ_tr, Φ_te, y_tr, y_te) ← temporal_split(Φ, y)   // no shuffling
6:       f ← Ridge(α_ridge = 1).fit(Φ_tr, y_tr)
7:       metrics[level][u] ← {MSE, MAE, R², correlation}(f(Φ_te), y_te)
8: repeat steps 1–7 for 20 seeds; apply paired Wilcoxon and paired t-test
9: repeat over grids of α, σ, L and for the two-2-simplex scenario
Output: per-unit metrics, statistical tests, sensitivity curves
```

Semi-synthetic scenario with real series. To address the limitation of purely synthetic data, we add an experiment in which the network units are ten real quarterly US macroeconomic series (1959–2009, 202 observations after transformation), obtained from the macrodata dataset of the statsmodels package: real GDP, real consumption, real investment, government spending, disposable income, CPI, M1, Treasury bill rate, unemployment, and inflation. Level series receive a log difference and rates receive a first difference; all are standardized. The consumption and investment series — whose co-movement is economically documented — act as triggers, and a synthetic target unit is added with the same mechanism of Eq. (1). Only the target and its noise are synthetic; all the dynamics of the triggers and of the remaining units come from observed data. The 20 repetitions vary the seed of the synthetic noise, keeping the real series fixed.

The temporal partition uses the first 75% of the time steps for training and the final 25% for testing (223 and 74 samples, respectively, with L = 3, in the synthetic scenario; 149 and 50 in the semi-synthetic one), without shuffling, respecting temporal order. The statistical significance of the gain at the target unit is assessed with the paired Wilcoxon signed-rank test (non-parametric) and the paired t-test over the MSEs of 20 random seeds (seeds 0 to 19). The reported metrics are MSE, MAE, R^2, and Pearson correlation.

4.4. Reproducibility: software, hardware, and processing times

All code, generated data, and figure scripts are publicly available at: **https://github.com/joaoclau ablation** (repository with installation instructions and single-command execution, `python run_all.py`). All random seeds are fixed; the numbers in this paper correspond exactly to the script's output.

The implementation uses Python 3.11.15 with the packages NumPy 2.4.4 [15], SciPy

1.17.1, scikit-learn 1.8.0 [18], NetworkX 3.6.1 [12], statsmodels 0.14, and Matplotlib 3.10.9. The experiments were executed on a Linux virtual machine with 2 vCPUs (Intel Xeon at 2.80 GHz) and 8 GB of RAM, without GPU. Table 2 presents the wall-clock times of each experiment; the full run takes about 60 seconds, which makes the protocol accessible for replication and extension.

Experiment	Time (s)
Main scenario (seed 0, 30 units × 3 levels)	0.27
Twenty seeds with paired tests	5.24
Sensitivity to interaction strength α (4 values × 5 seeds)	5.47
Sensitivity to noise σ (4 values × 5 seeds)	5.28
Sensitivity to lag depth L (4 values × 5 seeds)	5.21
Two-2-simplex scenario	0.25
MLP control	0.93
GCN control (800 epochs)	3.99
Fair baselines (20 seeds)	17.97
Semi-synthetic (20 noise seeds)	12.13
Total	**59.3**

Table 2. Measured processing times (2 vCPUs Intel Xeon 2.80 GHz, 8 GB RAM).

5. Results

5.1. Main result

Table 3 summarizes the central finding in the main scenario (seed 0). The strongest gain occurs at unit 7, precisely where the higher-order interaction was inserted in the generating mechanism. At the target unit, the simplicial model reduces the MSE from 0.533 (graph) to 0.021, a 96.1% reduction. The mean global MSE over the 30 units drops from 0.0292 to 0.0121 (a 58.6% reduction), a drop mostly explained by unit 7 itself, which already anticipates the localized nature of the gain.

Model	Global MSE	MSE unit 7	MAE unit 7	R^2 unit 7	Corr. unit 7
Baseline (0-simplex)	0.0293	0.536	0.588	0.860	0.929
Graph (1-simplex)	0.0292	0.533	0.587	0.861	0.929
Simplicial (2-simplex)	**0.0121**	**0.021**	**0.116**	**0.995**	**0.997**

Table 3. Main result (seed 0, L = 3, test with 74 samples). The 2-simplex gain concentrates at the unit affected by the group interaction.

Two aspects deserve emphasis. First, the graph model is practically indistinguishable from the baseline at unit 7 (0.533 versus 0.536): the neighbor mean adds no useful information about the product $x_5 \cdot x_6$, confirming that aggregating neighbors by mean is not equivalent to representing the conjunction between two specific neighbors. Second, the jump in R^2 from 0.861 to 0.995 and in correlation from 0.929 to 0.997 shows that the improvement is not an artifact of a single metric. The critical information lies in the **conjunction** between 5 and 6, and this conjunction is lost when the model sees only an aggregated neighborhood signal.

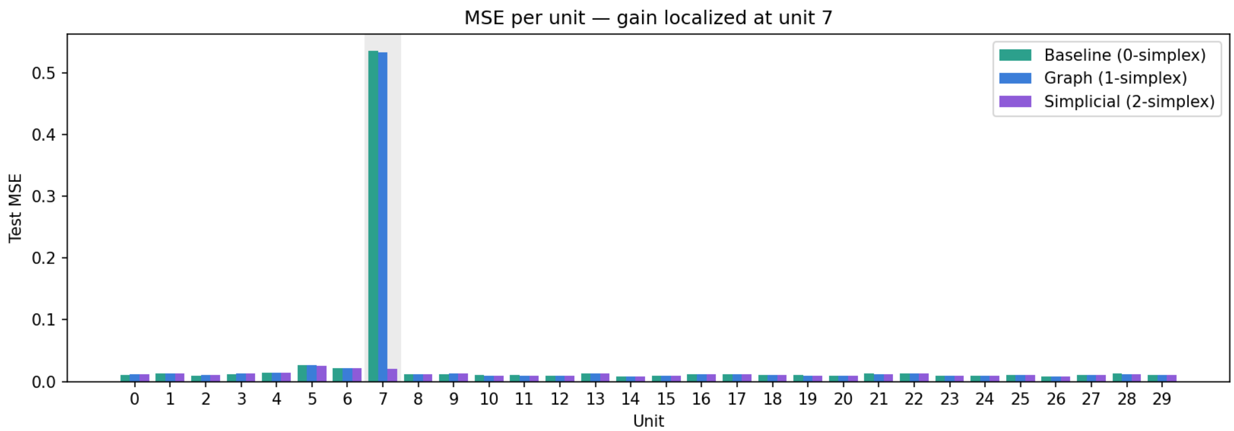

Figure 1: Figure 1

Figure 1. Test MSE per unit. The simplicial model's gain concentrates at unit 7, where the 2-simplex is present; at the remaining units the three levels are equivalent.

5.2. Robustness across multiple seeds

To verify that the result does not depend on a particular noise realization, we repeated the main scenario over 20 random seeds. The graph model's MSE at unit 7 was 0.592 ± 0.081 (mean ± standard deviation), against 0.025 ± 0.004 for the simplicial model. The relative improvement was **95.7% ± 0.6%**, ranging from 94.5% to 96.6% — the simplicial model won in all 20 seeds. The paired Wilcoxon test rejects the null hypothesis of equal distributions ($W = 0$, $p \approx 1.9 \times 10^{-6}$, the smallest possible p-value for $n = 20$), and the paired t-test confirms it ($p \approx 4.4 \times 10^{-18}$). At the network level, the reduction of the mean global MSE was 61.4% (range across seeds: 55.4% to 66.5%), again dominated by the target unit.

Figure 2. Distribution of the MSE at unit 7 over 20 seeds. The simplicial model maintains a consistent, statistically significant advantage.

Figure 3. Synthetic network highlighting the main 2-simplex {5, 6, 7}.

5.3. Temporal analysis and visual interpretation

Beyond the per-metric summaries, two visualizations help interpret the models' behavior at the target unit. In the predicted-versus-actual plot (Figure 4), the simplicial model concentrates closer to the 45° line, indicating better local calibration. In the time series (Figure 5), it tracks with higher fidelity the peaks and valleys induced by

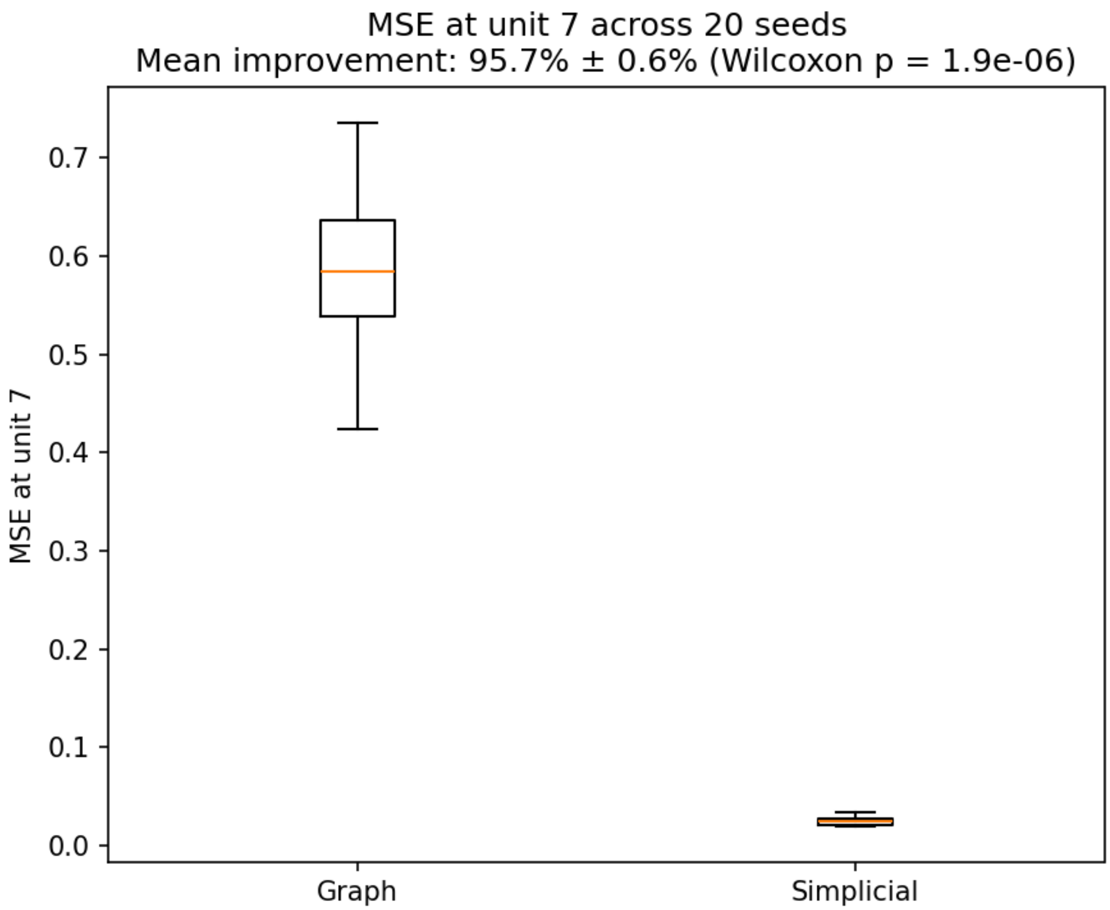

Figure 2: Figure 2

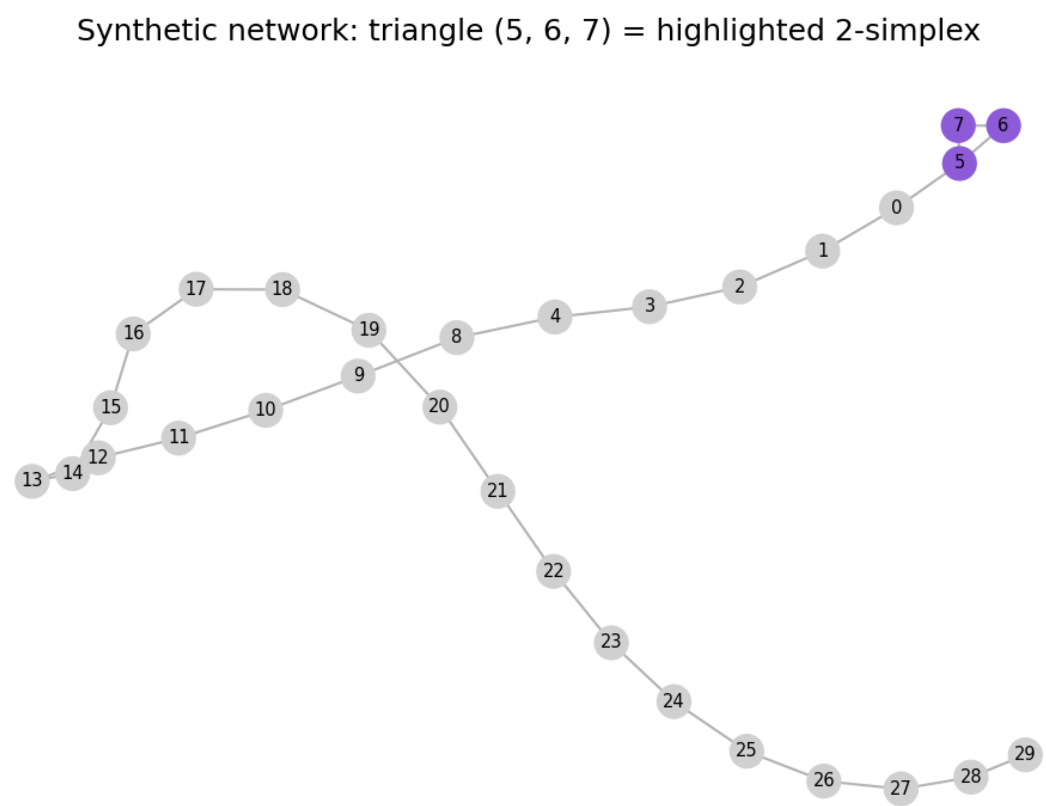

Figure 3: Figure 3

the synergy between units 5 and 6, while models based only on pairwise relations tend to smooth or shift those events.

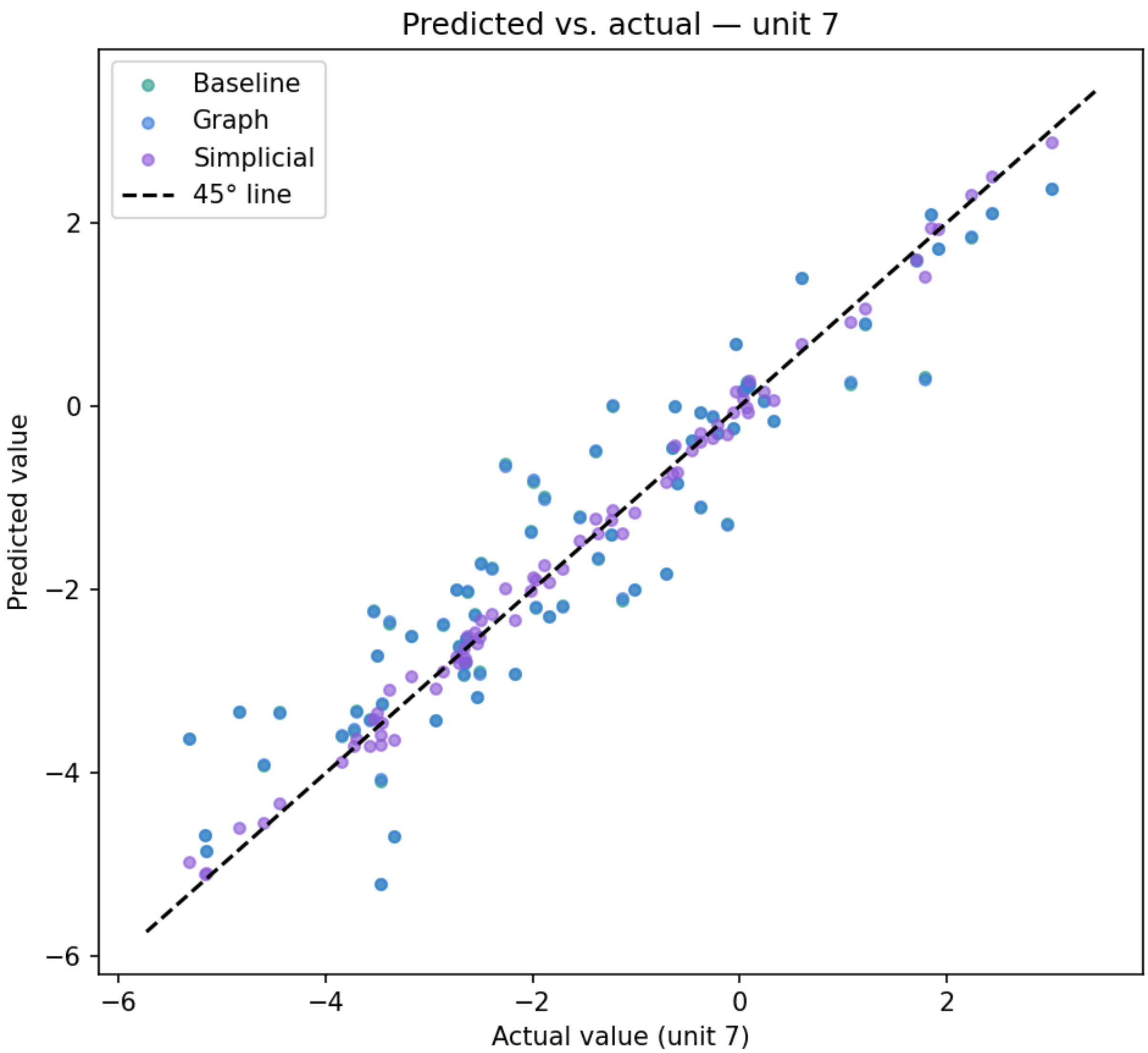

Figure 4: Figure 4

Figure 4. Predicted versus actual at unit 7. The simplicial model stays closer to the ideal line.

Figure 5. Actual and predicted time series at unit 7 (test samples).

5.4. Sensitivity to interaction strength, noise, and lags

Three sensitivity analyses quantify the robustness of the gain, each with 5 seeds per grid point (Figures 6 to 8).

Interaction strength α. The improvement at unit 7 grows with α: 91.1% ± 0.9% for

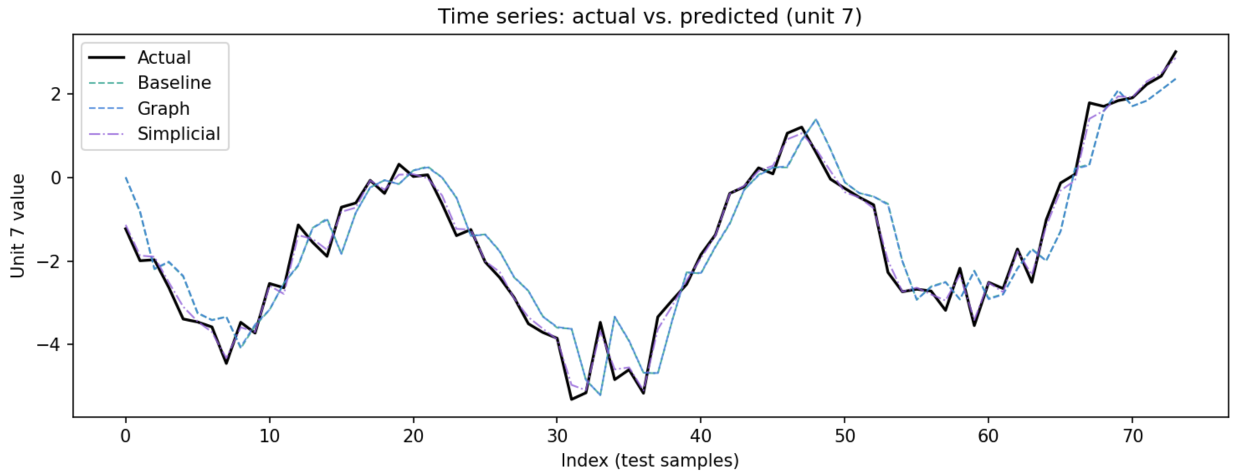

Figure 5: Figure 5

α = 2.5; 95.7% ± 0.5% for α = 5; 96.7% ± 0.5% for α = 7.5; and 97.0% ± 0.4% for α = 10. The pattern is consistent with the central hypothesis: the benefit of the 2-simplex widens as the group interaction comes to dominate data generation.

Noise level σ. The advantage remains high across the whole grid: 86.1% ± 2.2% for σ = 0.05; 95.7% ± 0.5% for σ = 0.1; 97.1% ± 0.4% for σ = 0.2; and 97.0% ± 0.5% for σ = 0.4. At no noise level does the face term stop being identifiable; the relative gain is smaller in the very-low-noise regime, in which pairwise models already exploit the smoothness of the series, and stabilizes in the noisier regimes, in which the multiplicative term dominates the target's variance.

Lag depth L. The gain is stable with respect to history depth: 96.3% ± 0.4% (L = 1), 96.1% ± 0.4% (L = 2), 95.7% ± 0.5% (L = 3), and 94.4% ± 0.6% (L = 5). Increasing individual or neighborhood history does not substitute for the face term: the 2-simplex gain holds because it acts on the **type** of dependency represented, not on the **amount** of history given to the predictor.

Figure 6. Sensitivity to the strength of the multiplicative interaction (mean ± sd over 5 seeds).

Figure 7. Sensitivity to the generator's noise level.

Figure 8. Sensitivity to the predictor's lag depth.

5.5. Generalization: two 2-simplices

Introducing a second 2-simplex {10, 11, 12}, with an analogous generating mechanism at unit 12, reproduces the same pattern. At unit 7, the MSE drops from 0.542 (graph) to 0.026 (a 95.2% reduction); at unit 12, from 0.595 to 0.027 (a 95.4% reduction). The gain reappears exactly at the units whose generating mechanism includes a group interaction, and only at those, corroborating hypothesis H2 of structural specificity.

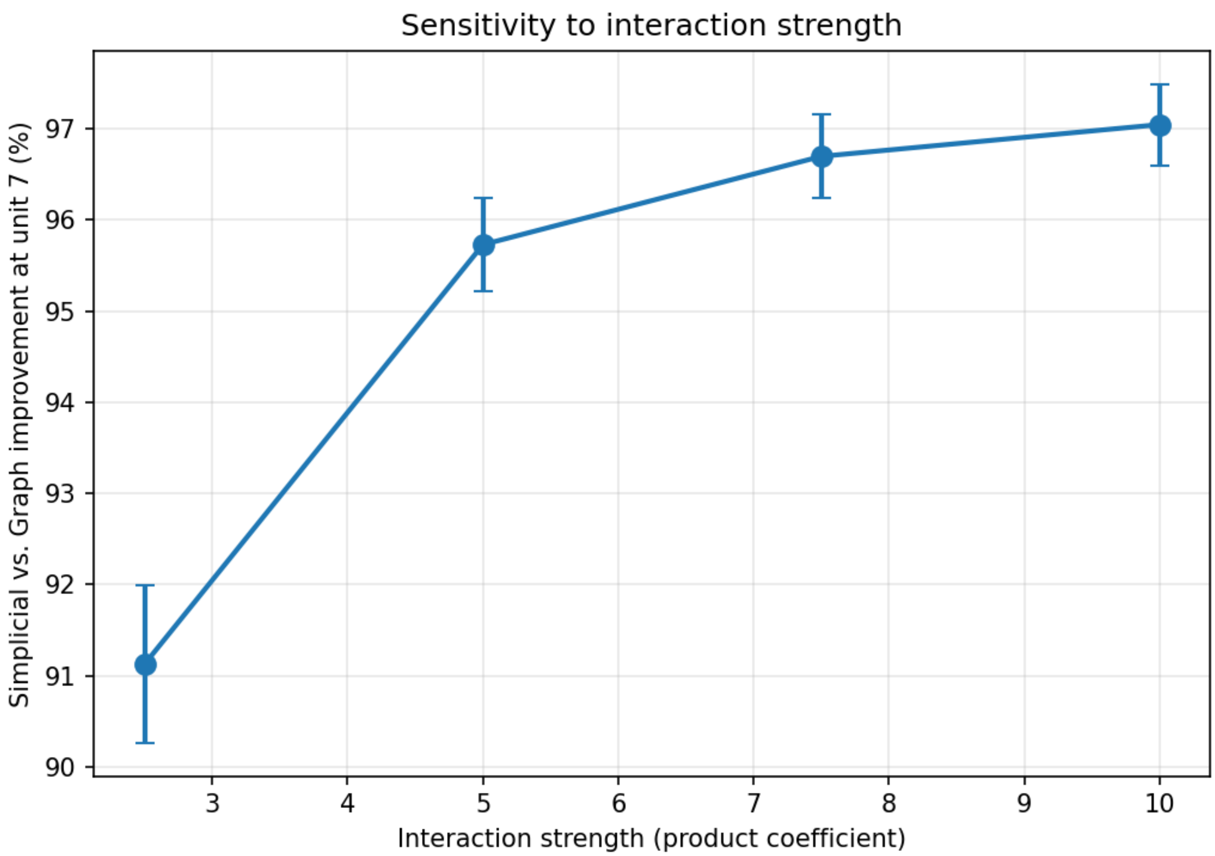

Figure 6: Figure 6

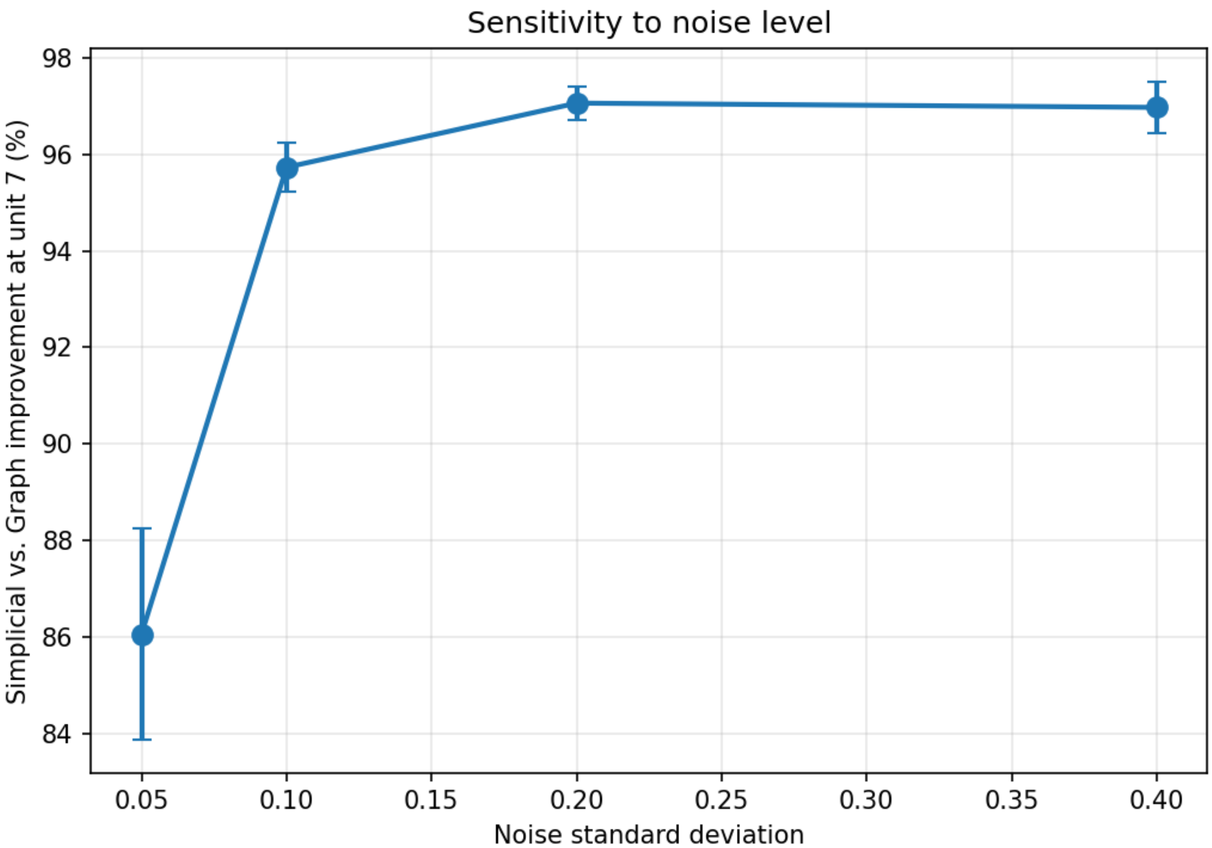

Figure 7: Figure 7

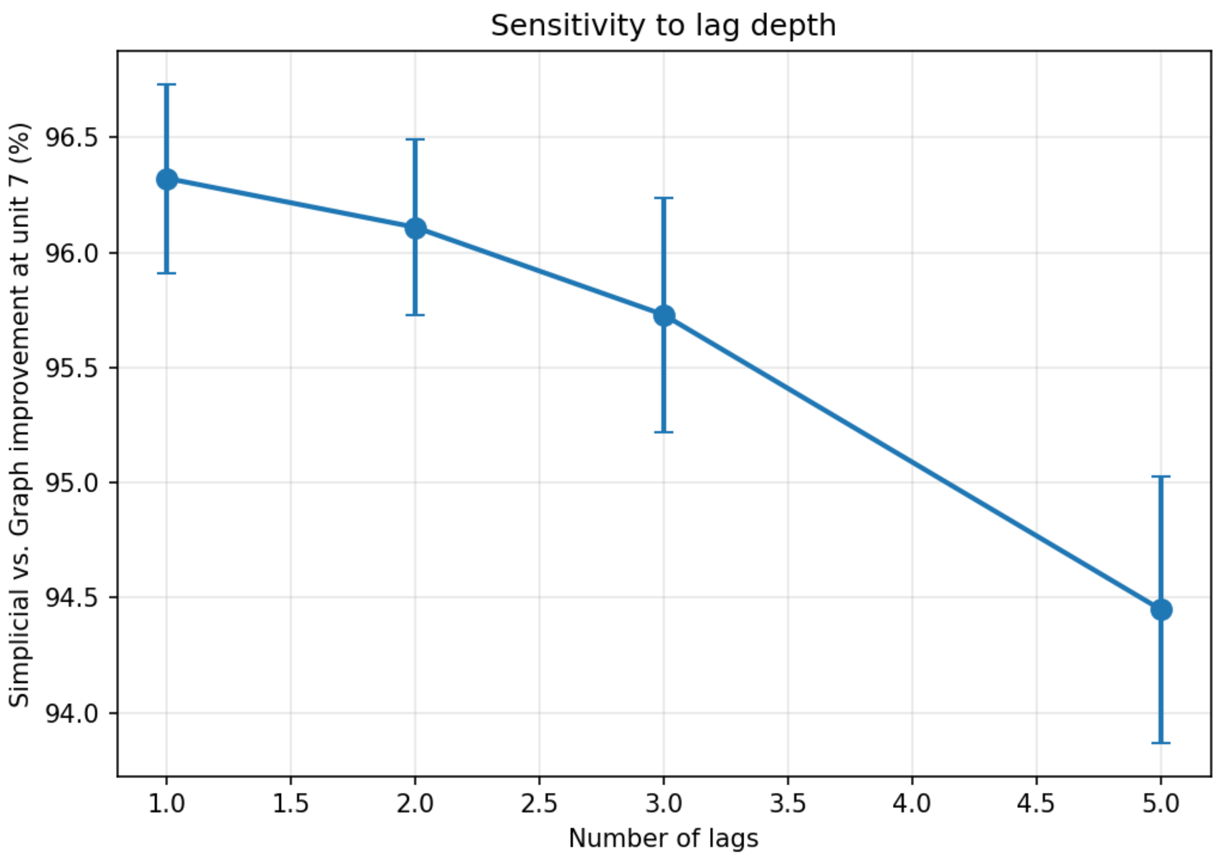

Figure 8: Figure 8

5.6. Controls: nonlinear capacity without the correct topology

MLP. The MLP trained on graph features reaches an MSE of 0.470 at unit 7 (R^2 = 0.877). It outperforms the graph Ridge (0.533), as expected from a nonlinear approximator with access to the triggers' lags via the neighborhood mean, but remains an order of magnitude behind the simplicial model (0.021). Adding generic nonlinearity to the regressor is therefore not enough; without the correct topological structure, the model remains at a substantial disadvantage.

One-layer GCN. The end-to-end trained GCN obtains an MSE of 2.874 at unit 7 — worse, in fact, than the individual baseline. The result is consistent with the theoretical expressivity analysis [23] and with the graph-regression analysis: without an explicit second-order operator between specific neighbors, an aggregation-based GNN with parameters shared across nodes tends to collapse neighborhood information and does not reproduce the multiplicative synergy. We note that this comparison uses a minimal architecture (1 layer, 16 hidden units); deeper GNNs or attention-based ones could reduce the gap, but the contrast with the simplicial model — which solves the problem with a single additional linear term — illustrates the cost of representing group interactions with pairwise operators.

Together, the controls support hypothesis H3: the observed gain is **attributable to the topology** of the representation, not to the capacity of the estimator.

5.7. Fair baselines: pair identification versus feature injection

A legitimate objection to the design so far is that the simplicial model receives exactly the product $x_5 \cdot x_6$ present in the generating mechanism. The pair-identification baselines separate the two components of the gain. Table 4 presents the results over 20 seeds.

Model	Features	MSE unit 7 (mean ± sd)
Graph (Ridge)	6	0.592 ± 0.081
Identified pair (Ridge)	9	0.603 ± 0.080
Identified pair (MLP)	9	0.198 ± 0.062
All pairs (Ridge)	409	0.037 ± 0.007
Simplicial (Ridge)	7	**0.025 ± 0.004**

Table 4. Fair baselines at unit 7 (20 seeds). The 2-simplex combines interaction identification with representational parsimony.

Three conclusions emerge. First, identifying the pair is not enough for a linear model: Pair-Ridge, even receiving the separate lags of x_5 and x_6, is statistically indistinguishable from the graph model (0.603 versus 0.592), because no linear combination of individual lags represents a product. Second, the interaction can be partially learned: Pair-MLP reduces the MSE to 0.198, confirming that a nonlinear approximator with the pair identified recovers part of the synergy — but it remains about eight times behind the simplicial model, with larger variance across seeds. Third, knowing the form of the interaction without knowing the topology has a measurable cost: AllPairs-Ridge reaches 0.037, close to the simplicial model, but needs 409 features against 7 — a 58-fold parsimony factor — and still shows a 45% higher error, reflecting the dilution of regularization over hundreds of irrelevant products. In summary, the value of the 2-simplex is not merely providing the correct feature: it is providing the **structural information of which interaction matters**, allowing a parsimonious, interpretable linear model to achieve the best performance.

Figure 9: Figure 9

Figure 9. Fair baselines at unit 7 (log scale, mean ± sd over 20 seeds, number of features annotated).

5.8. Semi-synthetic experiment with real series

Table 5 presents the results of the semi-synthetic scenario, in which the triggers are the real US consumption and investment series and only the target unit is synthetic. The pattern of the main study reproduces in full: the simplicial model reduces the MSE at the target unit from 16.00 (graph) to 0.011, a gain of 99.93% ± 0.02% (paired Wilcoxon over 20 noise seeds, $p \approx 1.9 \times 10^{-6}$). The simplicial model's error approaches the theoretical floor defined by the noise variance ($\sigma^2 = 0.01$), indicating nearly complete recovery of the mechanism. The fair baselines preserve the same ordering as in the synthetic scenario — Pair-Ridge does not capture the product (19.81), Pair-MLP captures only part of it (9.29 ± 2.99, with high variance), and AllPairs comes close (0.058), but with an error about five times larger than the simplicial model's. The absolute MSE is larger than in the synthetic scenario because the product of standardized real series has heavy tails, making the target more volatile; the relevant comparison is the relative one, among models within the same scenario.

Model	MSE at target unit (mean ± sd)
Baseline (0-simplex)	20.36 ± 0.15
Graph (1-simplex)	16.00 ± 0.12
Identified pair (Ridge)	19.81 ± 0.14
Identified pair (MLP)	9.29 ± 2.99
All pairs (Ridge)	0.058 ± 0.026
Simplicial (2-simplex)	**0.0111 ± 0.0025**

Table 5. Semi-synthetic scenario: real triggers (consumption × investment, 1959–2009), synthetic target, 20 noise seeds.

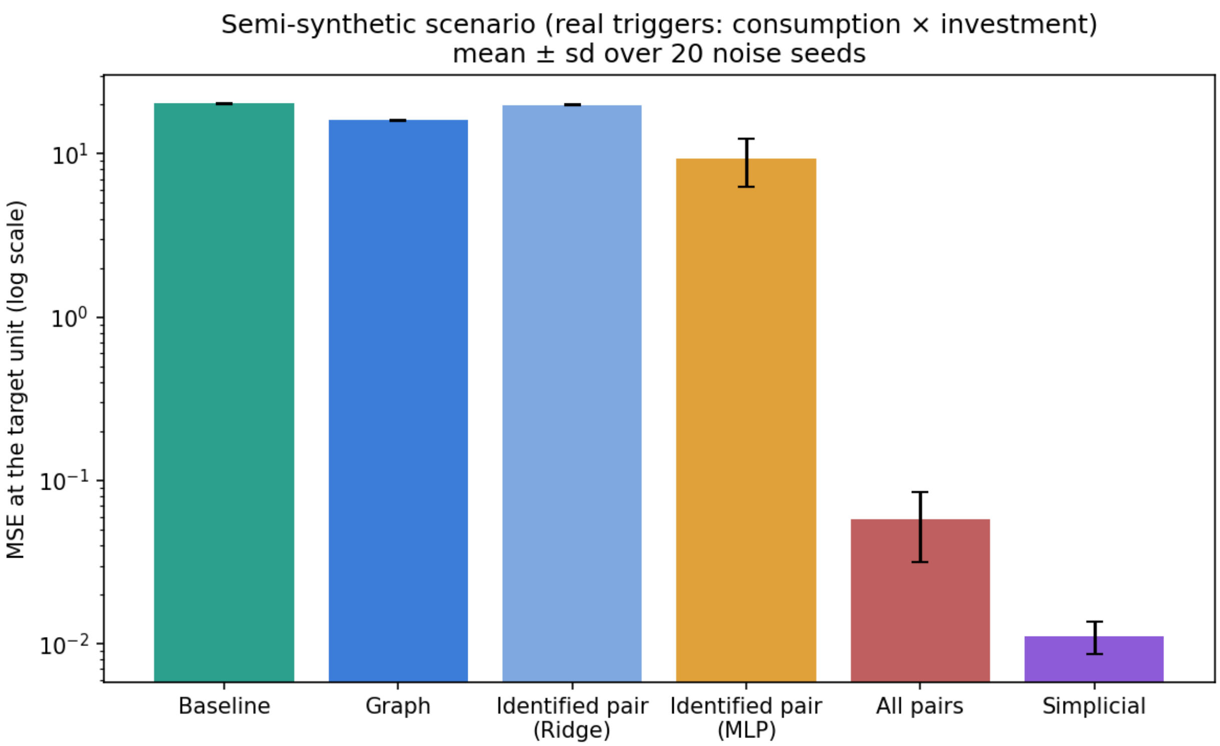

Figure 10: Figure 10

Figure 10. MSE at the target unit of the semi-synthetic scenario (log scale).

We record the nature of the design precisely: the 20 repetitions vary only the target's synthetic noise, since the real series are unique; the experiment demonstrates that the result is not an artifact of the regularity of the synthetic sinusoidal signals, but it does not constitute full validation on entirely observed data, in which the synergy itself would have to be real and unknown.

6. Discussion

6.1. Interpretation of the results

The results support three interpretations. First, **graph models fail when the dependency is higher-order**: neighbor mean or sum recovers neither the identity of the interacting pair nor its product — at unit 7, the graph level was statistically indistinguishable from the individual baseline. Second, the gain is **localized**: since most of the network contains no 2-simplices active in the generating mechanism, the global improvement (58.6% in the main seed; 61.4% on average) is almost entirely explained by the target units, and the remaining units are unaffected. Third, the gain is **attributable to the topology**: neither the MLP nor the GCN, both without the explicit face term, reaches the same performance. The fair baselines refine this reading: the benefit of the 2-simplex decomposes into structural identification (knowing which pair interacts, something AllPairs must compensate for with 409 features) and functional form (the explicit product, which Pair-MLP must learn imperfectly from separate lags). The 2-simplex delivers both components simultaneously, with the lowest error and the most parsimonious, interpretable representation among all evaluated models.

From the perspective of learning over networks, this suggests that certain modeling errors arise when higher-order dependencies are forced to fit into pairwise aggregation operators. Explicitly representing critical triangles can improve not only prediction, but also structural interpretability and the assessment of which relational patterns actually carry predictive signal.

6.2. Limitations

The conclusions must be read within the limits of the experimental design. First, the main study uses **synthetic data** with a known generating mechanism; the magnitude of the reported gains (around 95% at the target unit) reflects a scenario in which the multiplicative synergy dominates the target's variance. The semi-synthetic experiment mitigates this limitation by using real triggers, but the target remains synthetic: on entirely observed data, in which the synergy is real, unknown, and possibly unstable over time, the improvement ceiling will be lower. Second, we assume the **simplicial complex to be known** (oracle); in real settings, the higher-order structure must be inferred, and inference errors may consume part of the gain — as suggested by the performance difference between the simplicial model and AllPairs, which quantifies the cost of not knowing the topology. Third, the generator uses **multiplicative synergy at lag 1**; other functional forms of group interaction (joint thresholds, minima, heterogeneous delays) deserve their own study. Fourth, the nonlinear controls use minimal architectures, sufficient for the attribution argument but not representative of the state of the art in deep GNNs. Even so, precisely because it controls these factors, the experiment offers a useful evaluation methodology for future comparisons among GNNs, SNNs, and other relational learning models.

6.3. Inferring higher-order structure in real settings

In real applications, the critical step prior to using the 2-simplex is discovering **which** triangles matter. Three families of methods are natural candidates. The first clusters units by similarity of dynamics — for example, partitional clustering such as k-means or hierarchical agglomerative methods over co-movement features — and tests faces within the resulting groups. The second uses direct statistical criteria, such as the significance of the product term in sparsity-regularized regressions, sweeping the triangles of the 1-skeleton. The third employs hypergraph reconstruction methods from relational data, such as the Bayesian framework of Young, Petri, and Peixoto [25]. A promising agenda combines topology inference with predictive learning: first detect candidate faces from correlation, co-movement, clustering, or domain knowledge, and then assess — with this paper's ablation protocol — whether those faces improve prediction, interpretation, and intervention policies.

7. Conclusion

This paper showed, through a controlled and fully reproducible ablation study, that incorporating 2-simplices reduces prediction error when the target variable depends on multiplicative synergy among other units. The simplicial model reduced the MSE at the target unit by 95.7% ± 0.6% over 20 seeds (paired Wilcoxon, $p \approx 1.9 \times 10^{-6}$), with a stable gain under variations of interaction strength (91% to 97%), noise level (86% to 97%), and lag depth (94% to 96%), and reproducible at a second 2-simplex inserted in another region of the network (95.4%). The controls with MLP and a one-layer GCN and the fair pair-identification baselines indicate that the benefit stems from the topological structure of the representation — identification of the relevant interaction combined with parsimony (7 features against 409 for the all-pairs model) — and not from the estimator's nonlinear capacity. The semi-synthetic scenario, with real macroeconomic series as triggers, reproduced the same pattern with a 99.9% MSE reduction at the target unit.

More broadly, the paper contributes to graph mining, relational learning, and topological machine learning by providing an experimental setting — with public code, fixed seeds, and execution times on the order of 60 seconds — in which the usefulness of higher-order representation can be tested clearly, reproducibly, and quantitatively. Instead of evaluating only an architecture, we evaluate the adequacy of the relational structure used by the model itself.

As next steps, we highlight: applying the protocol to fully observed data from financial networks and other domains with documented group effects; investigating lifting strategies to detect 2-simplices from observational data, including clustering and hypergraph reconstruction methods; extending the comparison to deep simplicial neural networks in libraries such as TopoModelX [14]; and studying other functional forms of group interaction. This bridge between controlled experiment, semi-synthetic scenario, and observational application is the natural step toward consolidating the use of simplicial complexes in learning problems over networks.

Ethical considerations

This research does not involve human subjects, animals, personal data, or sensitive content. The synthetic data are generated by the code itself with fixed seeds, and the macroeconomic series used in the semi-synthetic scenario come from a public, aggregate database (macrodata, statsmodels), with no individual-level information. No ethics committee approval is therefore required.

Acknowledgements

Not applicable.

Funding

This research received no external funding.

Authors' contributions

JCNC conceived the study, implemented the experimental pipeline, and wrote the manuscript. DABT contributed to the experimental design, the analysis of results, and the revision of the manuscript. Both authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Declaration on the Use of Artificial Intelligence

During the preparation of this work, the authors used an artificial intelligence assistant (Claude, Anthropic) to support the implementation of the experimental code, the execution and tabulation of the analyses, and the drafting and language editing of the manuscript. All research questions, methodological decisions, and interpretations were defined by the authors, who reviewed and validated the code, the results, and the entire text, and who take full responsibility for the content of this publication.

Data Availability Statement

The complete source code, data generation scripts, figure scripts, numerical results (JSON), and reproduction instructions are available at: **https://github.com/joaoclaudio82/simpl** **ablation**. The full run requires only Python 3.11+ and the packages listed in `requirements.txt`, and completes in about 60 seconds on conventional hardware without a GPU.

References

[1] Peter W. Battaglia et al. Relational inductive biases, deep learning, and graph networks. *arXiv preprint arXiv:1806.01261*, 2018.

[2] Marco Bardoscia, Paolo Barucca, Stefano Battiston, Fabio Caccioli, Giulio Cimini, Diego Garlaschelli, Fabio Saracco, Tiziano Squartini, and Guido Caldarelli. The physics of financial networks. *Nature Reviews Physics*, 3:490–507, 2021.

[3] Federico Battiston, Giulia Cencetti, Iacopo Iacopini, Vito Latora, Maxime Lucas, Alice Patania, Jean-Gabriel Young, and Giovanni Petri. Networks beyond pairwise interactions: Structure and dynamics. *Physics Reports*, 874:1–92, 2020.

[4] Stefano Battiston, Michelangelo Puliga, Rahul Kaushik, Paolo Tasca, and Guido Caldarelli. DebtRank: Too central to fail? Financial networks, the FED and systemic risk. *Scientific Reports*, 2:541, 2012.

[5] Austin R. Benson, David F. Gleich, and Jure Leskovec. Higher-order organization of complex networks. *Science*, 353(6295):163–166, 2016.

[6] Cristian Bodnar, Fabrizio Frasca, Yuguang Wang, Nina Otter, Guido F. Montúfar, Pietro Liò, and Michael Bronstein. Weisfeiler and Lehman go topological: Message passing simplicial networks. In *Proceedings of the 38th International Conference on Machine Learning (ICML)*, pages 1026–1037, 2021.

[7] Cristian Bodnar, Fabrizio Frasca, Yifan Wang, Nina Otter, Kimon Fountoulakis, Pietro Liò, and Michael Bronstein. Weisfeiler and Lehman go cellular: CW networks. In *Advances in Neural Information Processing Systems*, volume 35, pages 2625–2640, 2022.

[8] Stefania Ebli, Michaël Defferrard, and Gard Spreemann. Simplicial neural networks. In *NeurIPS Workshop on Topological Data Analysis and Beyond*, 2020.

[9] Yifan Feng, Haoxuan You, Zizhao Zhang, Rongrong Ji, and Yue Gao. Hypergraph neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 3558–3565, 2019.

[10] Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. Neural message passing for quantum chemistry. In *Proceedings of the 34th International Conference on Machine Learning (ICML)*, pages 1263–1272, 2017.

[11] Lorenzo Giusti, Claudio Battiloro, Paolo Di Lorenzo, Stefania Sardellitti, and Sergio Barbarossa. Simplicial attention neural networks. *arXiv preprint arXiv:2203.07485*, 2022.

[12] Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th Python in Science Conference (SciPy)*, pages 11–15, 2008.

[13] Mustafa Hajij, Karthikeyan Ramamurthy, Tolga Birdal, Tamal K. Dey, Sayan Mukherjee, Shreyas Samaga, Dennis R. Livesay, Meera Walters, and Paul Rosen. Topological deep learning: Going beyond graph data. *arXiv preprint arXiv:2206.00606*, 2022.

[14] Mustafa Hajij et al. TopoX: A suite of Python packages for machine learning on topological domains. *Journal of Machine Learning Research*, 25(263):1–8, 2024.

[15] Charles R. Harris et al. Array programming with NumPy. *Nature*, 585:357–362, 2020.

[16] Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations (ICLR)*, 2017.

[17] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon. Network motifs: Simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002.

[18] Fabian Pedregosa et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[19] T. Mitchell Roddenberry, Nicholas Glaze, and Santiago Segarra. Principled simplicial neural networks for trajectory prediction. In *Proceedings of the 38th International Conference on Machine Learning (ICML)*, pages 9020–9029, 2021.

[20] T. Mitchell Roddenberry, Nicholas Glaze, and Santiago Segarra. Simplicial complex neural networks. In *Topological, Algebraic and Geometric Learning Workshops*, 2021.

[21] Michael T. Schaub, Austin R. Benson, Paul Horn, Gabor Lippner, and Ali Jadbabaie. Random walks on simplicial complexes and the normalized Hodge 1-Laplacian. *SIAM Review*, 62(2):353–391, 2020.

[22] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In *International Conference on Learning Representations (ICLR)*, 2018.

[23] Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks? In *International Conference on Learning Representations (ICLR)*, 2019.

[24] Naganand Yadati, Madhav Nimishakavi, Prateek Yadav, Vikas Nitin, Anand Louis, and Partha Talukdar. HyperGCN: A new method for training graph convolutional networks on hypergraphs. In *Advances in Neural Information Processing Systems*, volume 32, pages 1511–1522, 2019.

[25] Jean-Gabriel Young, Giovanni Petri, and Tiago P. Peixoto. Hypergraph reconstruction from network data. *Communications Physics*, 4:135, 2021.