Geometric Deep Learning on the Memecoin Graph

Jovonni L. Pharr
Georgia Cyber Warfare Range / BlackMind
info@gacwr.org

November 2026

Abstract

A memecoin is usually modeled as a price series: a token is a row, its features are scalars, and a classifier predicts whether it will "moon." We argue the data is *relational, not i.i.d.*—creators launch many tokens, wallets trade many tokens, and the coordination that actually moves these markets (bot fleets, shared funding sources, self-pump rings) lives in the *graph*, not in any per-token row. This paper is a geometric-deep-learning study of that claim, built from two code bases (`GACWR/fomo`, `GACWR/meme-research`) at three nested scales that we keep rigorously apart: an outcome-independent Solana census of 5,610,591 bonding curves, 1,823,902 creators and a 1.72% graduation base rate; tabular graduation models trained on *hundreds of thousands* of tokens that predict well out of time (714,747 early-window tokens, launch-time gradient boosting ROC 0.95 with 23× top-1% lift; 1,024,000 recent tokens, logistic 0.95→ GBM 0.98→ deep-MLP 0.99, 11× → 41× lift); and the graph benchmark on which the specifically *geometric* claim is tested—10,344 mint-disjoint snapshots over 5,172 mints (test fold 2,204). We formalize a heterogeneous, temporal wallet–token–creator graph; build an operator zoo (combinatorial and symmetric-normalized Laplacians, a GCN renormalization, a signature eigenvalue-diagonal operator, and typed buy/sell channels); and train a 1-hop creator prior, spectral-moment logistic models, and a typed two-channel message-passing GNN, each under strict mint-disjoint, outcome-time-purged, token-clustered validation. Both model families fit the target well; the audit is deliberately strict about *what that signal is attributable to*, and it narrows the specifically geometric claim. **(i)** Graph structure helps only after the raw flow columns are removed—"graph-only" spectral operators reach test ROC-AUC 0.86–0.92 on the mid-cap heads while the flow baseline sits *below chance* (0.33–0.48)—yet **(ii)** against a matched DeepSets "wallet-set" control that is handed the same wallets but *no adjacency*, the paired topology advantage is within noise ($\Delta\mathrm{AP} \approx -0.01$ at the \$1M head): a graph that does not beat its own bag of wallets. **(iii)** The typed GNN *classifies* well (ROC-AUC 0.79–0.89 across all fourteen heads; test $\log_{10}$-peak regression $R^2 = 0.61$–0.73) but does not *improve* on a logistic regression over the same wallet features—on single-token stars there is no extra structure for message passing to exploit. **(iv)** The pooled 0.97 is partly a source/time confound—within a single source AUCs run 0.76–0.99—and the empty high-cap cells on one 476-token trajectory fold are a *fold-size* limit, not a dataset limit (the benchmark holds 734/574/162 positives at \$1M/\$10M/\$100M). **(v)** On the client's own market-cap tiers a deployer-graph head beats the *free* market-cap-sort rule only at \$100k/\$250k, and the deployed model (AUC 0.759) is statistically indistinguishable from a zero-parameter Brownian-volatility null (0.757). We also prove why the naive *single-token* operators degenerate: every ego graph is a *star*, whose weighted adjacency has only the two nonzero eigenvalues $\pm\sqrt{\sum_i w_i^2}$ and whose normalized-Laplacian spectrum is $\{0, 2, 1^{(n-2)}\}$, so spectral moments mostly re-encode node count; and the eigenvalue-diagonal operator is measurably *not* permutation-equivariant. The one durable, genuinely non-price signal is *cross-token wallet convergence*: survivors and collapses separate with large Cliff's δ on shared-wallet and context fractions, and coordinated buyer clusters are $\sim$237× denser than chance—but only inside a single source and only as a forward-looking snapshot quantity, never as a backtest feature. The deliverables are the validated predictive models and this surviving signal, together with a first-class catalogue of the leakage traps, nulls and degeneracies that bound them.

1 Introduction

The dominant framing of "can you predict a memecoin" is tabular: each token is an independent row $x_i \in \mathbb{R}^d$ of hand-crafted scalars (early volume, holder counts, buy/sell ratios), and one fits a logistic regression, a gradient-boosted tree, or an MLP to a binary label such as `hit_1M`. This framing contains two illusions.

The sampling illusion. A market-cap-sorted census of "tokens that reached a pool" implies a graduation rate near 52%; a cohort selected on early features plus a realized outcome—never *on* the outcome—puts the true rate at 4.2%, falling to 0.95% across a 1.25M-token, 13-month population as pump.fun saturated. Enumerating the chain directly settles it: across 5,610,591 pump.fun bonding curves, exactly 96,613 graduate, a base rate of **1.72**% measured *before any outcome exists*. Every model

trained in the predecessor program targeted a pre-filtered ∼22% event—an order of magnitude apart from the population it must generalize to (Fig. 1).

The i.i.d. illusion. Those same models were fitted on 261 tokens. The census holds 21,497× as many curves and 511× as many positives, and it is not a bag of rows: it is a *deployer graph* over 1,823,902 creators and 6,823,760 accounts. Aggregation to scalars is a lossy projection of that graph. The established finding that *concentration beats breadth*—a tightly held token outruns a broadly held one (univariate AUC 0.92, inverted)—is itself a shadow of a graph property (cohort cohesion) measured through a scalar proxy.

This paper takes the relational object seriously and asks, with the lab's usual discipline, how much the geometry actually buys. Our contributions are:

1. A formalization of the memecoin market as a heterogeneous, temporal graph—creators, tokens, wallets; `deployed`, `trades`, `funded`, `co_traded` edges—together with the as-of discipline that keeps a present-day follow snapshot from leaking into the past (§2).
2. An operator zoo and three model families (1-hop creator prior, spectral-moment logistic, typed two-channel GNN with a DeepSets *wallet-set* ablation), with exact update equations (§5).
3. Two propositions explaining why the naive geometric operators are doomed on single-token ego graphs: a *star-graph spectral degeneracy* and the *non-equivariance* of the eigenvalue-diagonal operator (§6).
4. An adversarial empirical audit (§7): where structure helps, the matched wallet-set null, the failed paired tests, the source confound, the deployer head versus the free rule, and the Brownian-volatility null.
5. The one signal that survives—cross-token wallet convergence—stated with its effect sizes and its strict scope limits (§8).

Throughout, validated predictions, a surviving non-price signal, *and* disclosed negative results are all treated as deliverables, in the tradition of this lab's learnability studies.

2 The memecoin as a graph

Figure 2 draws the trade graph as one object: the full 6,828-token, 56,711-wallet graph, with the layout resolving the heaviest cross-token hubs so the structure is legible. It is a useful first look because the thing that will matter—a small set of wallets recurring across a great many otherwise-independent tokens—is visible before any model, as the bright hubs that tie the constellation together. Everything downstream is an attempt to say precisely how much of the outcome that recurrence explains.

What the raw data already dictates. Four pictures of the committed data (Fig. 3) fix the constraints every later model must respect: price is nearly a deterministic function of reserves (A); almost all tokens have already collapsed off their peak (B); the surviving peaks are heavy-tailed and regime-shifting (C); and the labels are an as-of lattice whose geometry is exactly what makes a naive join leak the future (D).

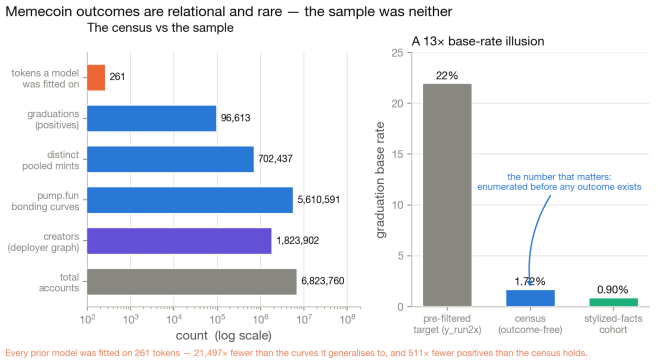


Figure 1: The data is relational and rare, and the sample was neither. Left: the outcome-independent census against the 261-token sample every prior model was fit on (log scale). Right: the graduation base rate corrected from a pre-filtered 22% to the enumerated 1.72%.

We distinguish *three* graphs and keep them rigorously apart, because conflating them is the single most common way to manufacture a leak.

Entities. A *creator* deploys tokens; a *token* is a mint on a chain; a *wallet* trades tokens; an *event* $(\text{wallet}, \text{token}, t)$ is a timestamped trade, the unit of prediction. Edges come in four relation types:

$$\underbrace{\text{creator} \to \text{token}}_{\texttt{deployed}}, \quad \underbrace{\text{wallet} \to \text{token}}_{\texttt{trades}\ (\pm)},$$
$$\underbrace{\text{wallet} \to \text{wallet}}_{\texttt{funded}}, \quad \underbrace{\text{wallet} \leftrightarrow \text{wallet}}_{\texttt{co_traded}}.$$

`trades` carries a sign (buy / sell); `funded` is a native-token transfer *before* the recipient's first buy, the seeding signature that distinguishes an organic launch from a dusting or a self-transfer; `co_traded` is the shared-token overlap. Figure 4 draws the object.

The central graph question. When several leaders buy the same token, are they independent actors (organic discovery) or a connected cluster (coordination)? A token bought by eight wallets that all follow or fund each other is a different object from one bought by eight strangers, and no market-cap feature can tell them apart. This is the hypothesis the whole program exists to test.

The as-of discipline. The follow graph is a *single present-day snapshot*; it contains edges created *after* historical trades, so joining it onto past events leaks the future. Every such column is prefixed `gsnap_` to make the

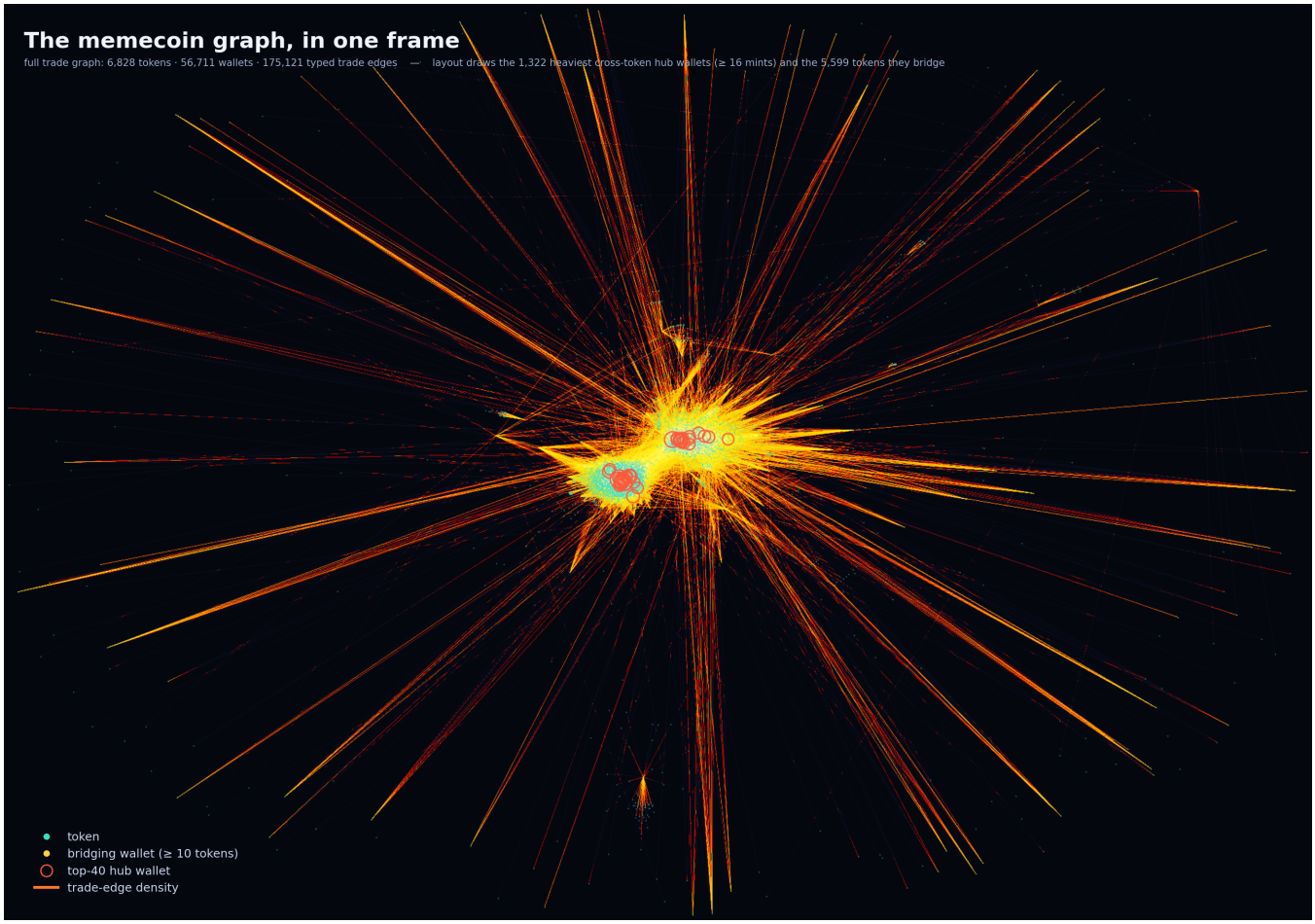


Figure 2: The whole object at once. One Datashader rendering of the bipartite wallet–token trade graph assembled from every committed panel snapshot—the full graph holds 6,828 tokens, 56,711 wallets and 175,121 typed trade edges (roughly 1.9M SOL of flow)—with edge density shaded by a histogram-equalised count. For a legible force layout we draw the heaviest *cross-token* hub wallets (those touching ≥ 16 mints, about 1,300 of them) and the 5,599 tokens they bridge; the single-token leaf halo is omitted. The structure is immediate: dense hubs are wallets that recur across hundreds or thousands of launches (ringed, the top-40 by cross-token degree; one touches 3,026 mints), stitching otherwise unrelated tokens into shared components. This recurrence—not any per-token wiring—is the signal that survives every control in ğ8.

hazard impossible to forget: these quantities are valid for describing the present network and for forward-looking scoring, **never for a backtest**. Likewise we train and evaluate at the *token* level, not the row level: the intra-token correlation is ICC $= 0.638$ with a design effect of 46.7, so all intervals are token-clustered bootstraps and all splits are mint-disjoint and out-of-time.

Three scales, kept apart. The program operates at three nested scales, and conflating them is the second-most-common way to manufacture a result. The *census* enumerates 5,610,591 bonding curves to fix the 1.72% base rate; no model is trained on it. The *tabular graduation models*—creator history, socials, launch timing, early-window microstructure—are trained on *hundreds of thousands* of tokens (714,747 in the deployed early-window predictor, 1,024,000 in the recent store, $\sim$ 1.3M across 13 monthly cohorts over 613k creators) and predict graduation *well* out of time: the early-window launch-time model reaches gradient-boosting ROC 0.95 (23× top-1% lift), and on the recent store logistic 0.95 rises through GBM 0.98 to a deep MLP at 0.99 (11× → 41× lift)—far above the creator-prior-only baseline of ≈ 0.66. The *graph benchmark*, by contrast, is the smallest scale: 5,172 mints carry the dual-snapshot trade prefix and peak labels a geometric model needs, giving 10,344 mint-disjoint snapshots (train 5,998 / val 2,142 / test 2,204). Every GNN result in this paper lives at that third scale; the "hundreds of thousands" belong to the tabular family and are never silently upgraded into evidence for the graph.

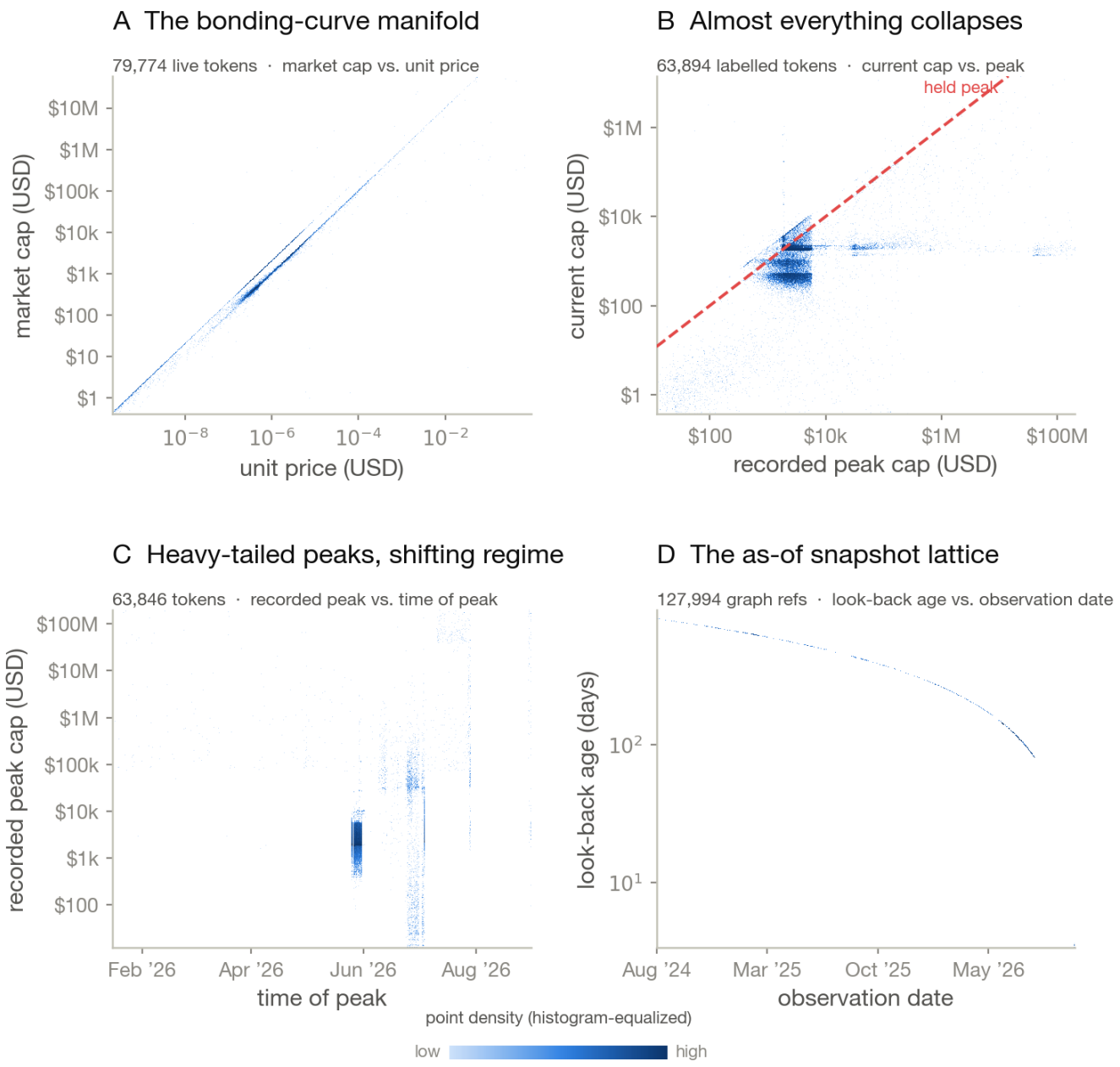


Figure 3: What the raw data dictates before any model, rendered with Datashader directly from the committed recovery tables (no sampling). **A:** every live token sits on a near-deterministic market-cap–price manifold (79,774 tokens), so a "price" feature mostly restates market cap. **B:** against their own recorded peak, almost all tokens have already collapsed—the mass sits far below the held-peak diagonal, pinned to a sub-$10k floor (63,894 labelled tokens). **C:** the peaks are heavy-tailed and regime-shifting, spanning five orders of magnitude in discrete collection batches. **D:** labels form an *as-of* lattice—each token observed at a fixed calendar date looking back a fixed horizon (127,994 graph references)—the structure a careless snapshot join turns into a future leak.

3 The ego graph as the unit of computation

The heterogeneous graph of §2 is never materialized whole. Every model instead reads a *bounded ego graph* centered on one token, because scoring is *inductive and per-token*—a fresh mint must be scored from its own neighborhood with no retraining—and the serve-time budget is fixed. We use two ego constructions, for two different jobs.

(i) The training ego graph (CAP = 24). The benchmark input is a star on 24 nodes: the token as node 0 and its top 23 wallets by known-prefix SOL volume. Each observed trade contributes a weight $\log(1 + \text{SOL}) + 0.1$ into one of two *typed, symmetrically normalized* adjacen-

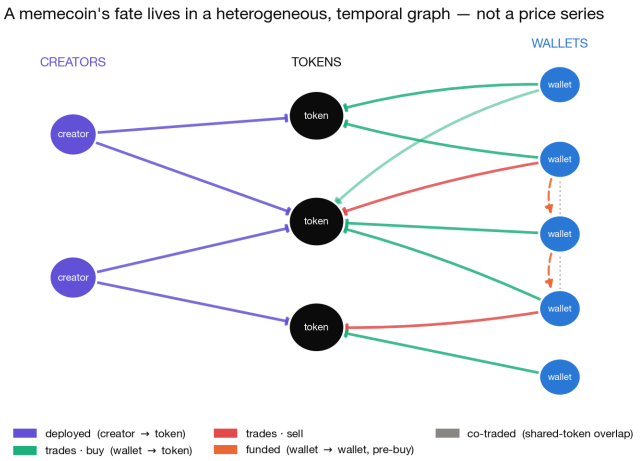


Figure 4: The heterogeneous, temporal memecoin graph. A token inherits signal from its creator's prior launches and from the track records of the wallets that trade it; coordination lives in the wallet–wallet `funded` and `co_traded` edges, which price features cannot see.

cies (BUY, SELL); no wallet–wallet edge is ever fabricated. Each node carries 14 features: four flow measures (buy count, sell count, buy SOL, sell SOL), each kept as a level, a trailing-60 s rate and a difference-of-rates acceleration ($4\times3=12$), plus two node-type indicators. At readout we concatenate a panel of global graph summaries— buyers, sellers, buy-wallet HHI and top-1 share, shared-wallet and neighbor-token counts, bipartite four-cycles, wallet-group sizes and signed SOL/USD flow—so set-level structure the star cannot carry on its edges still reaches the head. This is the object fed to the typed GNN (M2) and its wallet-set ablation (M2′).

(ii) The deployed ego view (radius 50). At serve time the product exposes a second, larger ego network, `_ego(mint, radius_buyers=50)`: the token's top 50 wallets by SOL size, each annotated with its *token degree* (how many of our tracked tokens the wallet also trades) and a pairwise *co-trade* overlap, together with token-level counts of multi-token wallets. This is deliberately a descriptive *view*, not a classifier: a high token degree or co-trade overlap *flags* shared activity for an analyst—the code says in as many words that it is not proof of a common controller. It is how the coordination hypothesis is triaged live, and it is where the surviving convergence signal of §8 is read off.

Useful, and self-limiting. The ego graph is the right unit—local, inductive, cheap—and it is *also* the origin of the degeneracy proved in §6. A single-token ego graph with no fabricated edges is a star; and ranking wallets by SOL and keeping the top 23 discards the long tail, so a token with 25 buyers and one with 2,500 can present the *same* selected neighborhood. The construction that makes scoring tractable is the same one that caps what a purely spectral reading can see—which is exactly why the frontier model (§5) replaces the single-token star with a bounded 2-hop, multi-token neighborhood.

4 Graph construction and operators

Edge weights. Each observation contributes $w_e = \log(1+\text{amount}_{\text{SOL}})+0.1$, kept in separate BUY and SELL channels and separated by quote unit so SOL and USD are never summed. Writing $\mathbf{A}_r$ for the resulting per-relation adjacency and $\mathbf{D}_r$ its degree diagonal, we symmetrically normalize $\widehat{\mathbf{A}}_r = \mathbf{D}_r^{-1/2}\mathbf{A}_r\mathbf{D}_r^{-1/2}$. *Weighting artifact: $\sum\log(1+v)$ is neither total nor log-total SOL—one 10-SOL trade scores ≈ 2.5 while ten 1-SOL observations score ≈ 7.9, so capture fragmentation changes the weights.*

Operator zoo. On the (symmetrized) adjacency $\mathbf{A} = \text{sym}(\mathbf{A}_{\text{buy}} + \mathbf{A}_{\text{sell}})$ with degree $\mathbf{D}$ we form

$$\mathbf{L} = \mathbf{D} - \mathbf{A}, \qquad \mathbf{L}_{\text{sym}} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}, \tag{1}$$

$$\widehat{\mathbf{A}}_{\text{GCN}} = \widetilde{\mathbf{D}}^{-1/2}(\mathbf{A}+\mathbf{I})\widetilde{\mathbf{D}}^{-1/2}, \quad \mathbf{A}_{\text{dir}} = \mathbf{A}_{\text{buy}} + \mathbf{A}_{\text{sell}}^{\top}. \tag{2}$$

The lab's signature construction is the *eigenvalue-diagonal* operator

$$\mathbf{B} = \mathbf{A} + \text{diag}\big(\lambda_{\pi(1)}, \ldots, \lambda_{\pi(n)}\big), \qquad \lambda = \text{eig}(\mathbf{A}), \tag{3}$$

i.e. the ordinary adjacency with the graph's own eigenvalues written onto its diagonal under an assignment π. (3) is to be sharply distinguished from the spectral decomposition $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\top}$; putting eigenvalues on the *diagonal* is not a change of basis, and §6 shows the cost.

Spectral-moment feature map. Where we want a fixed-length summary of an operator $\mathbf{M}$ without training a GNN, we compress it to eight moments

$$\phi(\mathbf{M}) = \big[\,\overline{\text{Re}\,\lambda},\ \text{std}\,\text{Re}\,\lambda,\ \max|\lambda|,\ \overline{|\lambda|},\ \max|\text{Im}\,\lambda|,\ \|\mathbf{M}\|_F,\ \text{Re}\,\text{tr}\,\mathbf{M},\ \text{Re}\,\text{tr}\,\mathbf{M}^2\,\big]. \tag{4}$$

Temporal derivatives. For each graph summary $f(t)$ we keep the level and two *backward* differences (no future leak), a rate and an acceleration,

$$v(t) = \frac{f(t) - f(t^-)}{t - t^-}, \qquad a(t) = \frac{v(t) - v(t^-)}{\bar{t} - \bar{t}^-}, \tag{5}$$

at step-1 and step-5 trailing scales, so a 40-summary panel occupies 200 numeric and 200 mask positions.

5 Models

(M0) The 1-hop creator prior. The degenerate geometric baseline is a single message pass on the `deployed`

graph: a token's score is its creator's leakage-safe running graduation rate over *prior* siblings,

$$s(\text{token}) = \mathbb{E}\big[\,\text{grad} \mid \text{creator},\ t' < t\,\big], \tag{6}$$

implemented as the expanding mean of the creator's *shifted* prior-sibling outcomes, which already reaches graduation AUC ≈ 0.62.

(M1) Spectral-moment logistic / ridge. Stack $\phi(\mathbf{M})$ of (4) over the operator zoo and fit a logistic (classification) or ridge (log-peak regression). No message passing; the geometry enters only through the spectrum.

(M2) Typed two-channel GNN. With node states $\mathbf{h}^{(0)} = \mathbf{W}_{\text{proj}} x$, two message-passing layers update

$$\mathbf{h}^{(k+1)} = \text{ReLU}\Big(\mathbf{W}^{(k)}_{\text{own}}\mathbf{h}^{(k)} + \sum_{r\in\{\text{buy,sell}\}} \mathbf{W}^{(k)}_{r}\,\widehat{\mathbf{A}}_r\,\mathbf{h}^{(k)}\Big)\odot\mathbf{m}, \tag{7}$$

with a boolean real-node mask $\mathbf{m}$, and read out $\hat{g} = \mathbf{W}_{\text{head}}\big[\,x_{\text{glob}} \,\|\, \mathbf{h}^{(2)}_0 \,\|\, \overline{\mathbf{h}^{(2)}}\,\big]$. The production model has $12\to16$ projection, 16-wide messages, a 24-unit head, and $\sim$2.5k parameters; heads are Platt-calibrated on validation only.

(M2′) The wallet-set (DeepSets) ablation. The decisive control replaces the adjacency in (7) with the identity, $\widehat{\mathbf{A}}_r\mathbf{h}^{(k)} \mapsto \mathbf{h}^{(k)}$, so the model sees the *same multiset of wallets* with the *same features* but *no topology*. The question "does the graph add anything beyond the bag of wallets?" is exactly M2 vs. M2′.

(M3) The frontier (specified, not yet built). The principled generalization is a relation-typed, temporal GNN (R-GCN / heterogeneous GraphSAGE, 2–3 hops) over {wallet, token, creator} with all edges masked to $t \le t_0 + T$, so a sybil fleet funded from a common source collapses to a shared embedding and a creator's history propagates as literal multi-hop message passing. It is unbuilt because the data is not yet there (§9).

6 Why the naive single-token operators degenerate

A single-token ego graph with no invented wallet–wallet edges is a *star*: one hub (the token) joined to $n-1$ leaves (wallets), nothing else. Stars are spectrally almost empty.

Proposition 1 (Star adjacency degeneracy)**.** *Let* $\mathbf{A}$ *be the weighted adjacency of a star on* n *nodes with hub–leaf weights* $w_1, \dots, w_{n-1}$ *and zero diagonal. Then* $\mathbf{A}$ *has eigenvalues* $\pm\sqrt{\sum_i w_i^2}$ *and* 0 *with multiplicity* $n-2$.

Proof. Only the hub row and column are nonzero, so $\operatorname{rank}\mathbf{A} = 2$ and 0 is an eigenvalue of multiplicity $n-2$. Since $\operatorname{tr}\mathbf{A} = 0$ the two remaining eigenvalues are $\pm\mu$, and $\operatorname{tr}\mathbf{A}^2 = 2\sum_i w_i^2 = 2\mu^2$ gives $\mu = \sqrt{\sum_i w_i^2}$. □

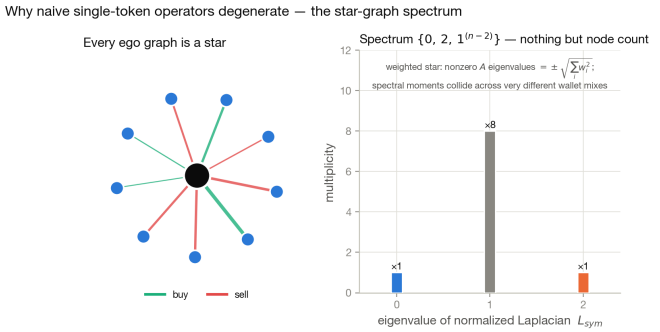


Figure 5: The star-graph degeneracy. Every single-token ego graph is a star (left); its normalized-Laplacian spectrum is $\{0, 2, 1^{(n-2)}\}$ (right), so spectral summaries re-encode node count and little else.

Corollary 1 (Spectral moments see only node count)**.** *For an unweighted star, the normalized-Laplacian spectrum is exactly* $\{0,\ 2,\ 1^{(n-2)}\}$ *and the normalized-adjacency spectrum is* $\{-1,\ +1,\ 0^{(n-2)}\}$. *Hence the moment map* (4) *of* $\mathbf{L}_{\text{sym}}$ *is a function of* n *alone, and of* $\mathbf{A}$ *a function of* $(n, \sum_i w_i^2)$ *alone: two ego graphs with the same node count and weight energy but entirely different wallet behaviour collide before the learner sees them.*

This is not a worst case; it is the *typical* case, because every ego graph in the corpus is a star (Fig. 5). It predicts exactly the collapse we observe for `normalized_L` in §7.2.

Proposition 2 (Non-equivariance of the eigenvalue-diagonal operator)**.** *The map* $\mathbf{A} \mapsto \mathbf{B} = \mathbf{A} + \operatorname{diag}(\lambda_\pi)$ *of* (3) *is permutation-equivariant iff the assignment* π *commutes with relabeling, i.e.* $\pi(P\mathbf{A}P^\top) = P\,\pi(\mathbf{A})$ *for every permutation* P. *The implemented index-order assignment does not, so there exists* P *with* $\mathbf{B}(P\mathbf{A}P^\top) \neq P\,\mathbf{B}(\mathbf{A})\,P^\top$.

Measured on held-out ego graphs, the operator-equivariance error of (3) is many orders of magnitude above the $\sim 10^{-14}$ residual of $\mathbf{L}$ and $\mathbf{L}_{\text{sym}}$—the eigenvalue-diagonal assignment is, as proved, genuinely not permutation-equivariant. A non-equivariant operator makes the model's output depend on wallet *ordering*, an artifact with no market meaning; we therefore refuse to rank it as a winner wherever it appears to lead.

7 Experiments and results

Setup. The graph benchmark is 10,344 mint-disjoint snapshots over 5,172 mints (train 5,998 / val 2,142 / test 2,204), split chronological 60/20/20, with an *outcome-time-purged* variant that removes snapshots whose label horizon overlaps the test window. The prediction target is a nested family of binary market-cap "peak" heads—did the future observed peak reach $\ge$ {\$50k, …, \$100M}—plus a continuous $\log_{10}$-peak regression. Heads are nested, not multiclass: a token above \$1M also satisfies \$50k. We report ROC-AUC, PR-AUC/average precision, lift, Brier and calibration—never accuracy.

Table 1: The large tabular graduation stores and their out-of-time performance—the "hundreds of thousands" the program is actually built on. All are token-disjoint, train-earlier/test-later. (The census of 5,610,591 curves and the 1,398,948-token, 13-cohort population fix base rates; no model is trained on them.)

token store	n tokens	base	best OOT ROC (lift@1%)
early-window (launch-time)	714,747	2.96%	0.95 GBM (23×)
recent store	1,024,000	1.96%	0.99 deep-MLP (40×)

7.1 The outcome is learnable: performance across scales

Before the controls, the plain fact: the target is *learnable*, and the models fit it well—at *every* scale, including the largest. The tabular graduation family, trained on hundreds of thousands of tokens, predicts graduation strongly out of time (Table 1): the deployed early-window model reaches gradient-boosting ROC 0.95 with 23× top-1% lift on 714,747 tokens, and on the 1,024,000-token recent store a logistic 0.95 rises through GBM 0.98 to a deep MLP at 0.99. On the graph side, Table 2 collects the within-fold numbers: the typed two-channel GNN reaches test ROC-AUC 0.79–0.89 across all fourteen nested peak heads on the mint-disjoint, three-seed split, and the Pons peak-tier panel (n = 24,232 test) reaches 0.94–0.99 on observed-peak and current-cap tiers. That these targets are learnable is *not* the claim this paper contests. The controls that follow (ğ7.2–ğ8) ask the harder questions—what the signal is *attributable* to (the wallet set, not its wiring), and whether the specifically geometric gains *survive* source and time shift—and those are where the geometric story narrows. Read these tables as the ceiling the rest of the paper interrogates.

7.2 Where structure helps, and where it is harmful

Figure 6 is the learnability map across the operator zoo and seven thresholds. Two facts stand out. First, the raw-flow baseline is *at or below chance* (0.33–0.48) on this mint-disjoint split: the six flow columns that a tabular model would lean on do not survive source/era shift. Second, *zeroing* those columns ("graph-only") sharply *improves* the spectral operators— `standard_L` climbs to 0.86–0.92 on the mid-caps—while `normalized_L` collapses to 0.17–0.24, exactly the star-graph degeneracy of Corollary 1 made visible. Structure helps, but only once the actively harmful flow features are removed, and only for operators the degeneracy spares.

7.3 The topology control

A graph model that beats a flow baseline has proven only that *wallets* carry signal, not that their *wiring* does. The correct control is M2′: the same wallets, no adjacency. On a matched $1M comparison (same split, seed and 8-node budget), topology and set are a dead heat (Fig. 7): the paired difference is $\Delta\text{AP} = -0.011$ $[-0.028, +0.001]$. We therefore state plainly that *this pilot does not establish a topology advantage*; the apparent graph-versus-flow gap conflates input set, model family and node budget.

7.4 The learned GNN does not improve on a simple baseline

The typed two-channel GNN (7) *classifies well*—it carries the same 0.79–0.89 test AUCs of Table 2. The question here is narrower: does its message-passing *add* anything over a plain logistic regression on the same wallet features? Validation-selected per head and compared by a token-paired bootstrap, its ΔROC-AUC against the logistic is either indistinguishable from zero or mildly *negative* at every threshold (Fig. 8): $[-0.105, -0.022]$ at $250k, $[-0.060, -0.010]$ at $1M, $[-0.054, -0.003]$ at $10M. On the continuous target the GNN is, by contrast, a solid regressor—test $\log_{10}$-peak $R^2 = 0.61$–0.73 on the benchmark (it collapses only on a tiny 10-token end-to-end prefix, where no method fits). Node-permutation outputs match to $\sim 10^{-8}$, confirming the network is equivariant: it is faithful, and it simply does not *need* the adjacency that a bag-of-wallets model lacks.

7.5 The source confound

The headline that makes memecoin graph-learning look solved—pooled AUC $\sim$ 0.97—is a source/time artifact. Reported source-stratified, the same models give resident-$10M 0.987 but resident-$500k–$1M only 0.76–0.77 (Fig. 9, left). On *one* thin single-source trajectory fold (476 tokens) the average precision barely clears prevalence ($50k flow 0.406 vs. GNN 0.437; $250k 0.082 vs. 0.098), and the $1M/$10M/$100M heads have no positives *in that fold*—a fold-size limit, not a dataset limit: the benchmark itself holds 734/574/162 positives at those tiers (Fig. 9, right). Part of the pooled headline lives in the mixing of a legacy ATH-labelled source with a resident-trajectory source, not purely in the graph.

7.6 The deployer head and the Brownian null

Folding a nine-feature deployer block (prior-launch counts, best/median multiples, fraction $\geq 2\times$ /5×) into nine market-cap tier heads, over 1.33M rows and 65,138 tokens, we score "hits per 15 tokens inspected" against the *free* market-cap-sort rule. The learned head wins only at the bottom: $+0.14$ $[+0.11, +0.18]$ at $100k and $+0.05$ $[+0.02, +0.08]$ at $250k, but -0.04 at $1M, -0.11 at $5M and -0.09 at $10M (Fig. 10). Above $250k, sorting by one column you already have beats the model. *A commit message claimed the deployer block "im-*

Table 2: Predictive performance across model families and outcome scales—the within-fold learning numbers. *Left:* the typed two-channel GNN on the graph benchmark (three-seed test, mint-disjoint, n=2,204; ± is seed SD). *Right:* the tabular peak-tier models on the Pons panel (gradient-boosted, n=24,232 test), for both observed-peak and current-cap tiers. AUC=ROC-AUC, AP/PR=average precision, base=positive rate. Both families fit the target well; whether that signal is topological, and whether it survives source/time shift, are the questions of ğ7.2–ğ8, not of this table.

Graph benchmark — typed GNN (3-seed test)

peak ≥	base	ROC-AUC	AP
$50k	0.607	0.829 ± .038	0.803 ± .044
$100k	0.525	0.857 ± .054	0.783 ± .075
$250k	0.439	0.885 ± .045	0.767 ± .093
$500k	0.399	0.880 ± .049	0.729 ± .088
$750k	0.348	0.862 ± .040	0.660 ± .070
$1M	0.333	0.864 ± .038	0.652 ± .070
$2.5M	0.308	0.873 ± .035	0.657 ± .049
$5M	0.282	0.871 ± .021	0.616 ± .022
$7.5M	0.267	0.872 ± .018	0.594 ± .027
$10M	0.260	0.869 ± .020	0.576 ± .029
$25M	0.248	0.864 ± .010	0.543 ± .023
$50M	0.166	0.824 ± .019	0.355 ± .025
$75M	0.095	0.792 ± .010	0.206 ± .005
$100M	0.073	0.790 ± .014	0.163 ± .004

Tabular peak-tier — Pons panel (n=24,232)

tier ≥	base	ROC-AUC	PR-AUC
observed-peak tier			
$10k	0.1284	0.973	0.908
$25k	0.0342	0.955	0.689
$50k	0.0092	0.961	0.579
$100k	0.0040	0.956	0.303
$250k	0.0018	0.960	0.179
$500k	0.0012	0.968	0.110
$1M	0.0004	0.993	0.595
current-cap tier			
$10k	0.0095	0.940	0.209
$25k	0.0037	0.942	0.119
$50k	0.0017	0.939	0.099
$100k	0.0010	0.981	0.072
$250k	0.0006	0.981	0.092

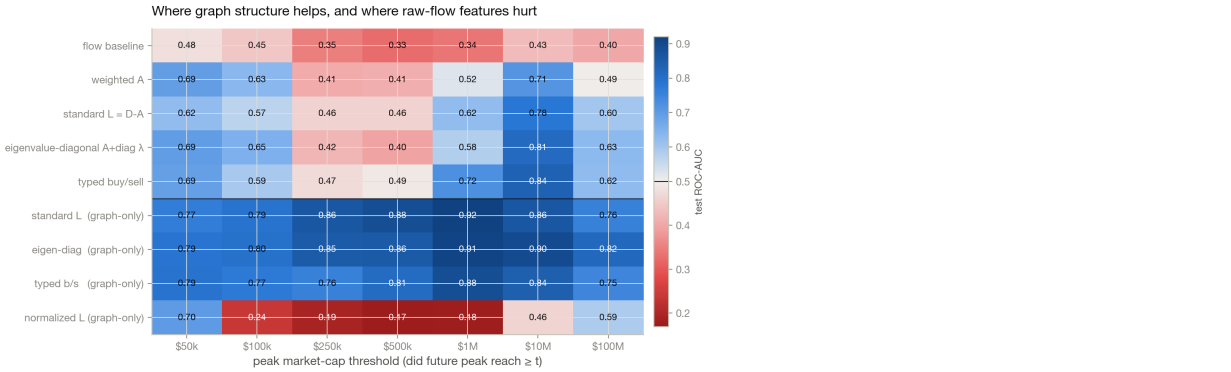


Figure 6: Test ROC-AUC across the operator zoo (rows) and seven peak thresholds (columns), mint-disjoint and outcome-time-purged. Blue beats chance (0.50); red is below it. Removing the six raw-flow columns ("graph-only," below the rule) lifts the spectral operators; `normalized_L` collapses on star graphs.

proves $500k/$1M/$5M and hurts none"; the committed numbers contradict it. We report the numbers.

Worse, the deployed production scorer (AUC 0.759, top-5% lift 2.53×) is statistically indistinguishable from a *zero-parameter* Brownian-volatility null (0.757): much of the apparent signal is just "volatile things move more" (Fig. 11). Reframing the target from market-cap level to size-of-*move* lifts an EVM-first pool to AUC 0.827 and lift 4.18×, but on live tokens that reframe *loses* to the deployed model and clears the Brownian null in 0 of 9 learnability cells. A 103-column single-feature sweep beats the null 0 times; a 336-cell cost-aware backtest has 0 cells with P&L confidence above zero; moonshots specifically ($\geq 31.6\times$, $\geq 100\times$) score AUC 0.444 and 0.407—below a coin flip. The leakage-free ceiling is ≈ 0.76 AUC after a lifetime-count leak is removed.

7.7 The out-of-time learnability map

For completeness, Table 3 gives the out-of-time numbers where a genuine (if thin) edge does survive: the mid-cap graduation-magnitude heads on real backfill. They are good classifiers and poor rankers of magnitude, consistent with an information ceiling in which most of the log-peak variance is simply unlearnable.

8 The one signal that survives

Strip away everything that fails a control, and one structural phenomenon remains, and it is exactly the coordination the program was built to find. Among tokens past $100k on the resident source, survivors (retained $\geq 50\%$ of peak) and collapses (rugged to $< 10\%$) separate with large Cliff's δ on precisely the cross-token quan-

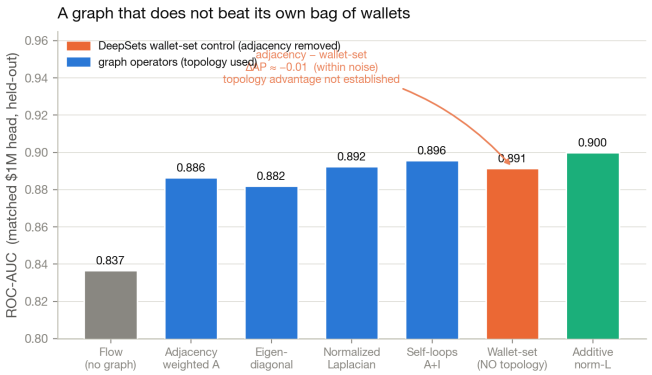


Figure 7: Matched $1M head. The DeepSets wallet-set control (orange), handed the same wallets with the adjacency removed, matches the graph operators; the paired adjacency−set advantage is -0.011 $[-0.028, +0.001]$.

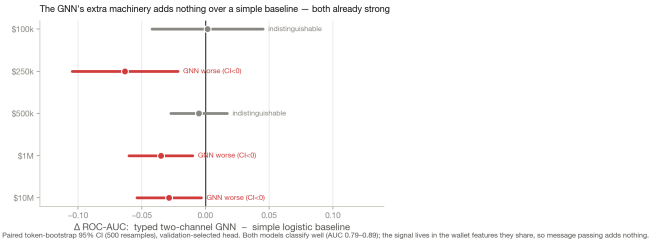


Figure 8: Paired token-bootstrap 95% CI of ΔROC-AUC (typed GNN minus simple logistic). Right of zero would favour the GNN; no interval does.

Table 3: Out-of-time learnability (real backfill, n = 7,030, token-disjoint, train-earlier/test-later). HistGBM unless noted; base rates are cohort rates.

head	base	ROC	PR	lift@1%
`hit_250k`	0.056	0.906	0.476	13.5×
`hit_500k`	0.039	0.928†	–	–
`hit_1m`	0.023	0.912	0.359	23.9×
log-peak (reg.)	best Spearman 0.381, R^2 0.480 (MLP)			
survival ≥ 1h	100% censored—hazard undefined			

†deep MLP. EVM retrain (n = 341): ROC 0.886/0.860/0.829; the breadth thesis does *not* transfer to absolute tiers (−traders/trade AUC 0.39/0.37/0.34).

tities: $\delta = +0.70$ on *context fraction*, +0.65 on *shared-wallet fraction*, +0.37 on sell fraction and wallet-event HHI (Fig. 12). Define, for a token with buyer set $\mathcal{B}$ present in the follow graph,

$$I(\text{token}) = \frac{\#\{\text{connected components of } \mathcal{B}\}}{|\mathcal{B}|} \in \left[\tfrac{1}{n}, 1\right], \tag{8}$$

$I = 1$ meaning every buyer is an island (organic discovery) and $I = \frac{1}{n}$ meaning they collapse to a single cluster (coordination). At the population level, 65% of tokens have an all-islands buyer set, while the 35% carrying any follow edge have a buyer subgraph $\sim 237\times$ denser than the ambient follow graph—a self-pump signature that is *invisible to every price feature*. Order flow confirms the mechanism: buys are self-exciting (a Hawkes branching ratio $\eta \to 1$ marks a token going critical), and concentration beats breadth because a high top-buyer share raises the effective coordination of the Keynesian beauty contest and the persistence of informed order flow.

Two scope limits bound the claim. The separation is *source-conditional*: on the legacy source the same δ is $\approx +0.01$ (no separation). And the 237× and the convergence features are `gsnap_` snapshot quantities—valid for forward-looking scoring, never as backtest inputs. The signal is real, it is non-price, and it is narrow.

9 Verdict and limitations

Both model families predict the target well, and a genuine non-price coordination signal survives every control. The tabular graduation family reaches out-of-time ROC up to 0.95–0.99 on hundreds of thousands of tokens; the graph benchmark's typed GNN classifies at 0.79–0.89 across fourteen heads. What the audit shows is *not* that prediction fails but that the specifically *geometric* gains are narrow. Under matched controls: topology does not beat a bag of wallets; the GNN does not improve on a logistic over the same inputs; part of the pooled 0.97 is a source/time confound; the deployed scorer sits near a volatility null; and the signature eigenvalue-diagonal operator is not equivariant. The durable positives are concrete: strong, deployable graduation prediction at scale, a locatable edge at $100k–$250k, a genuinely non-price coordination signal with a large within-source effect, and a proof of why the naive single-token operators degenerate.

Three limitations are structural, not incidental. **The graph is not yet built.** The `funded` graph has 0 collected rows; the follow graph was silently discarded before being rewired; only 2 of 104 buy events carry a full 1-hour forward window. The frontier heterogeneous temporal GNN (§5) is specified but unbuildable until the firehose backfill lands. **The ego graph is a star.** Until multi-token, multi-hop neighborhoods replace single-token stars, the spectral degeneracy of §6 caps what any spectral method can see. **The economics are tail-driven.** A cost-aware backtest entering the top decile at $T = 5$ minutes returns +11.1% per trade at 3% cost but wins only 28–38% of the time with a median trade of −2% to −7%: modest, tail-driven, and execution-sensitive. This is not a money printer, and a prior shipped $1M head was found degenerate (every real coin scored 0) against fixture-fabricated base rates—do not deploy capital on it.

10 Related work

Our framing is the geometric-deep-learning program of unifying architectures by their symmetries [1], instan-

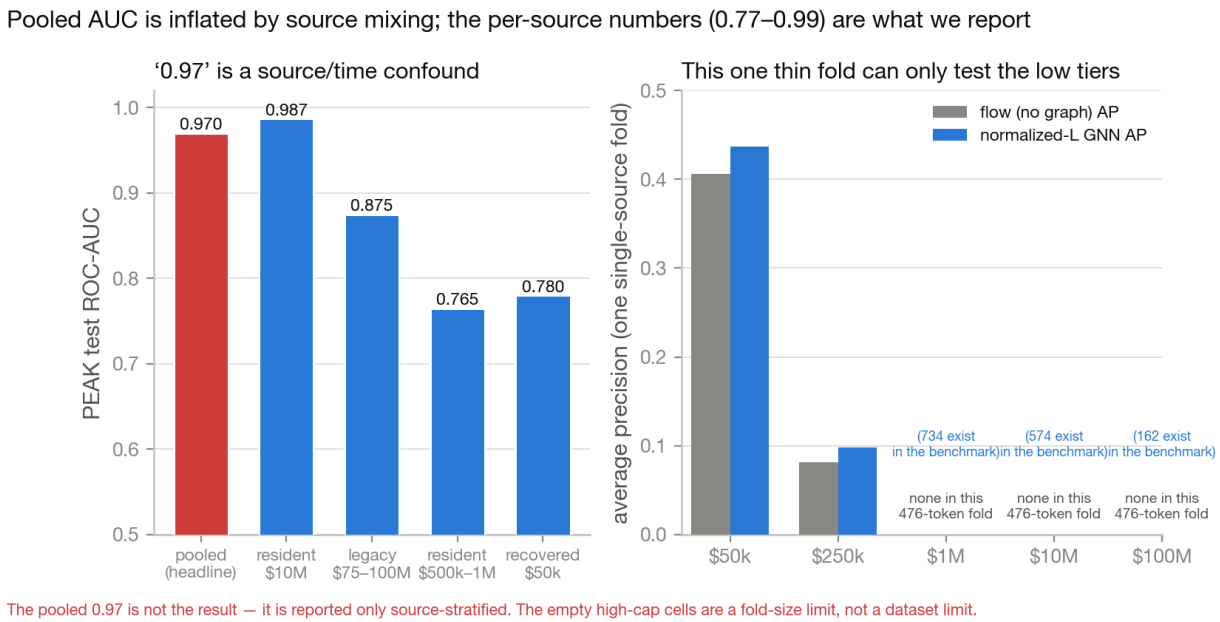


Figure 9: Left: the pooled 0.97 is a confound; source-stratified AUCs are far lower and are the only numbers we report. Right: on one source the edge nearly vanishes and the high-cap heads become undefined for lack of positives.

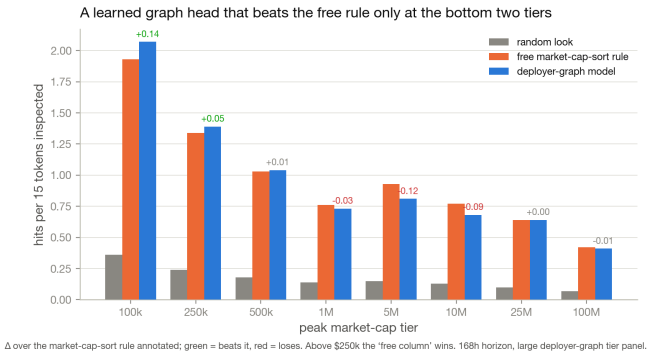


Figure 10: A deployer-graph tier head versus the free market-cap-sort rule and a random look. It beats the free rule only at $100k and $250k (Δ annotated) and loses above.

tiated with graph convolution [2], inductive neighborhood sampling [3], attention [4], and relation-typed message passing [5]; the wallet-set control is a Deep Sets permutation-invariant encoder [6], and the expressivity ceiling of 1-2-hop message passing is the Weisfeiler–Leman bound of GIN [7]. The temporal masking follows temporal graph networks [8]. The spectral arguments are standard spectral graph theory [9]. The market mechanisms are Kyle's informed order flow [10] and Keynes's beauty contest [11]; self-exciting trade arrivals are Hawkes processes [12]; effect sizes use Cliff's δ [13]; and the heavy-tail caution is the Clauset–Shalizi–Newman discipline on power-law claims [14].

11 Conclusion

Treating a memecoin as a node in a graph rather than a row in a table is the right idea, and the models built on it predict well: graduation is forecast out of time at ROC 0.95–0.99 on hundreds of thousands of tokens, and a real, price-invisible coordination signal survives every control. What this study pins down is *where the specifically geometric edge lives and where it does not*: the representation (graph vs. set) matters less than its proponents hope, the surviving phenomenon (coordination) is real but narrow, and the naive single-token operators degenerate for reasons we can prove. The deliverable is the validated predictive family, the surviving non-price signal, and a reproducible map of the nulls that bound them—together with a clear statement of the multi-hop, multi-token structure that would let geometric deep learning on these markets clear its own nulls.

References

[1] M. M. Bronstein, J. Bruna, T. Cohen, P. Veličković. *Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges.* arXiv:2104.13478, 2021.

[2] T. N. Kipf, M. Welling. *Semi-Supervised Classification with Graph Convolutional Networks.* ICLR, 2017.

[3] W. L. Hamilton, R. Ying, J. Leskovec. *Inductive Representation Learning on Large Graphs* (GraphSAGE). NeurIPS, 2017.

[4] P. Veličković et al. *Graph Attention Networks.* ICLR, 2018.

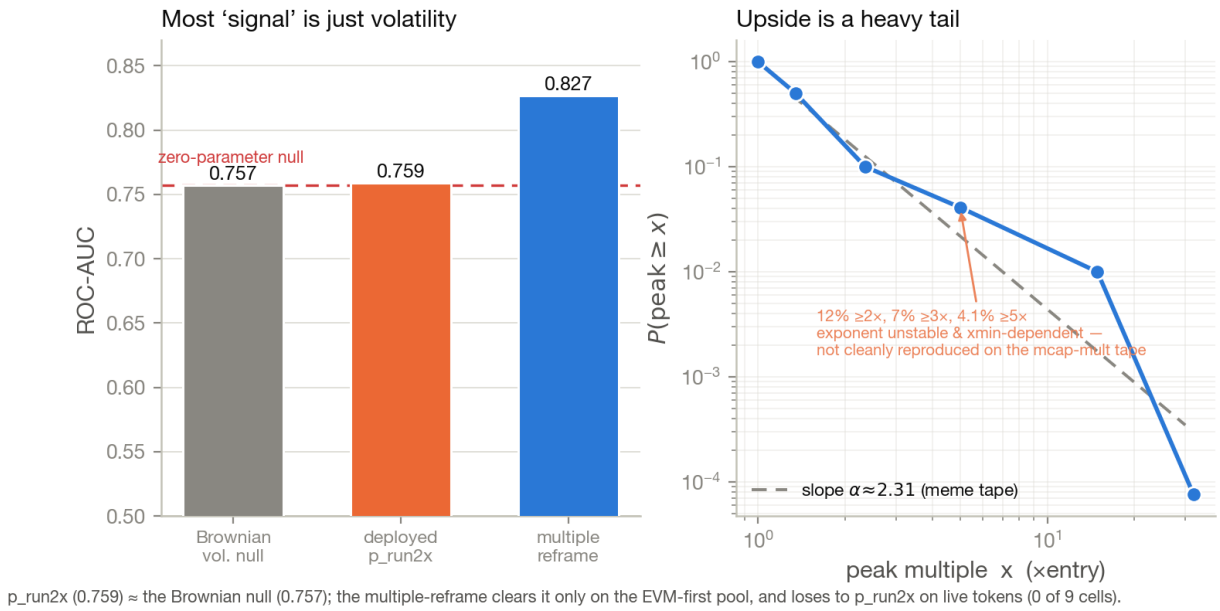


Figure 11: Two nulls. Left: the deployed model (0.759) sits on the zero-parameter Brownian-volatility null (0.757); the size-of-move reframe clears it only on the EVM-first pool. Right: upside is a heavy tail whose exponent is unstable and $x_{\min}$-dependent—the clean $\alpha \approx 2.31$ is not reproduced on the market-cap-multiple tape, and the far tail cannot be called.

[5] M. Schlichtkrull et al. *Modeling Relational Data with Graph Convolutional Networks* (R-GCN). ESWC, 2018.

[6] M. Zaheer et al. *Deep Sets.* NeurIPS, 2017.

[7] K. Xu, W. Hu, J. Leskovec, S. Jegelka. *How Powerful are Graph Neural Networks?* (GIN). ICLR, 2019.

[8] E. Rossi et al. *Temporal Graph Networks for Deep Learning on Dynamic Graphs.* arXiv:2006.10637, 2020.

[9] F. R. K. Chung. *Spectral Graph Theory.* AMS, 1997.

[10] A. S. Kyle. *Continuous Auctions and Insider Trading.* Econometrica 53(6), 1985.

[11] J. M. Keynes. *The General Theory of Employment, Interest and Money*, Ch. 12. 1936.

[12] A. G. Hawkes. *Spectra of Some Self-Exciting and Mutually Exciting Point Processes.* Biometrika 58(1), 1971.

[13] N. Cliff. *Dominance Statistics: Ordinal Analyses to Answer Ordinal Questions.* Psychological Bulletin 114(3), 1993.

[14] A. Clauset, C. R. Shalizi, M. E. J. Newman. *Power-Law Distributions in Empirical Data.* SIAM Review 51(4), 2009.

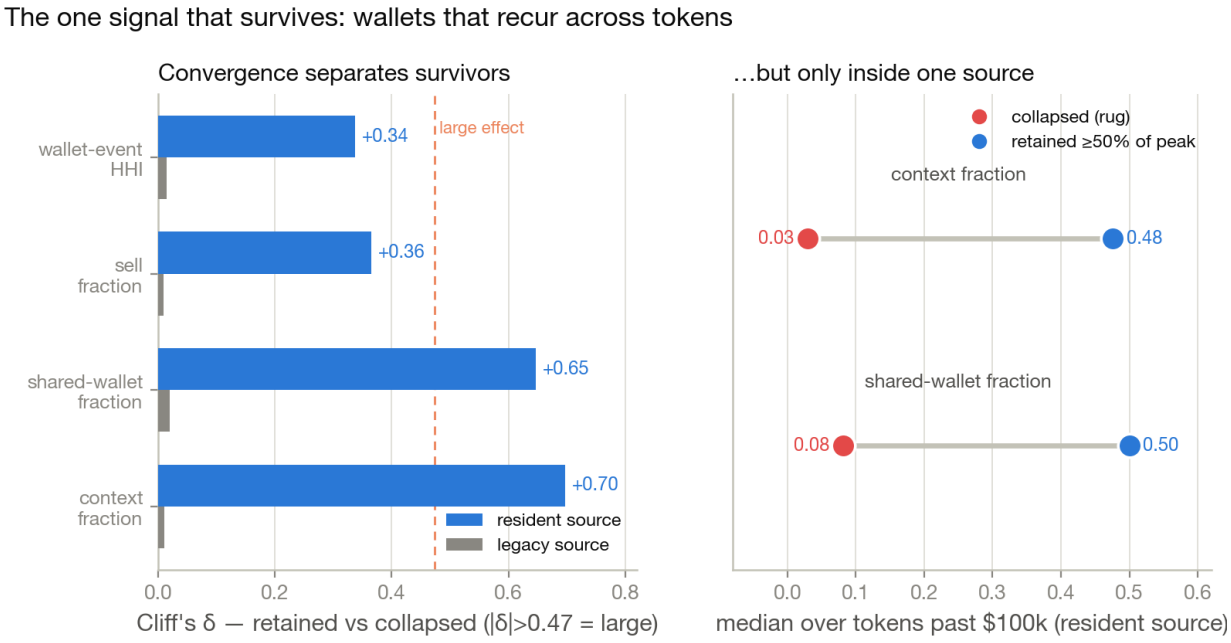


Figure 12: Cross-token wallet convergence. Left: survivors and collapses separate with large Cliff's δ on shared-wallet and context fractions on the resident source, and not at all on the legacy source. Right: median shared-wallet and context fractions for retained vs. collapsed tokens.