Exact Cascade Uncertainty Under Pairwise Compression

A sharp law for two-seed bootstrap closure in 3-uniform hypergraphs

Aditya • NSRI Summer Research Hackathon 2026 • Solo project

Abstract (210 words)

Higher-order data are often compressed into pairwise co-occurrence counts, but this can hide how pairs are assembled into triples. We ask for the largest possible difference in deterministic cascade size between simple 3-uniform hypergraphs with the same labeled exact pair-codegree matrix. Starting from two active vertices, a triple activates its third vertex when its other two are active. We define $M_2(n)$ as the maximum closure-size gap inside any projection fiber on n vertices. A direct proof gives $M_2(n)=0$ for $2 \le n \le 5$ and $M_2(n)=n-3$ for $n \ge 6$. The upper bound follows because equal codegrees force both cascades either to stop at two or reach at least three vertices. A five-edge Pasch-trade-plus-context construction attains the bound and extends to every larger n. Fresh exhaustive enumeration of all 2^{20} simple 3-uniform hypergraphs on six labeled vertices independently checks the first ambiguous order, a maximum gap of three, and edge thresholds: four edges for different final sets, five for different sizes, and seven when non-isomorphism is additionally required on six vertices. Two separately implemented censuses agree on 69 shared summary fields. The result is a precise warning about information lost under pairwise compression for one abstract deterministic closure rule; it is not an empirical claim about real epidemics or social networks.

Research question Among labeled simple 3-uniform hypergraphs with the same exact pair-codegree matrix, how different can the final two-seed bootstrap closure sizes be?

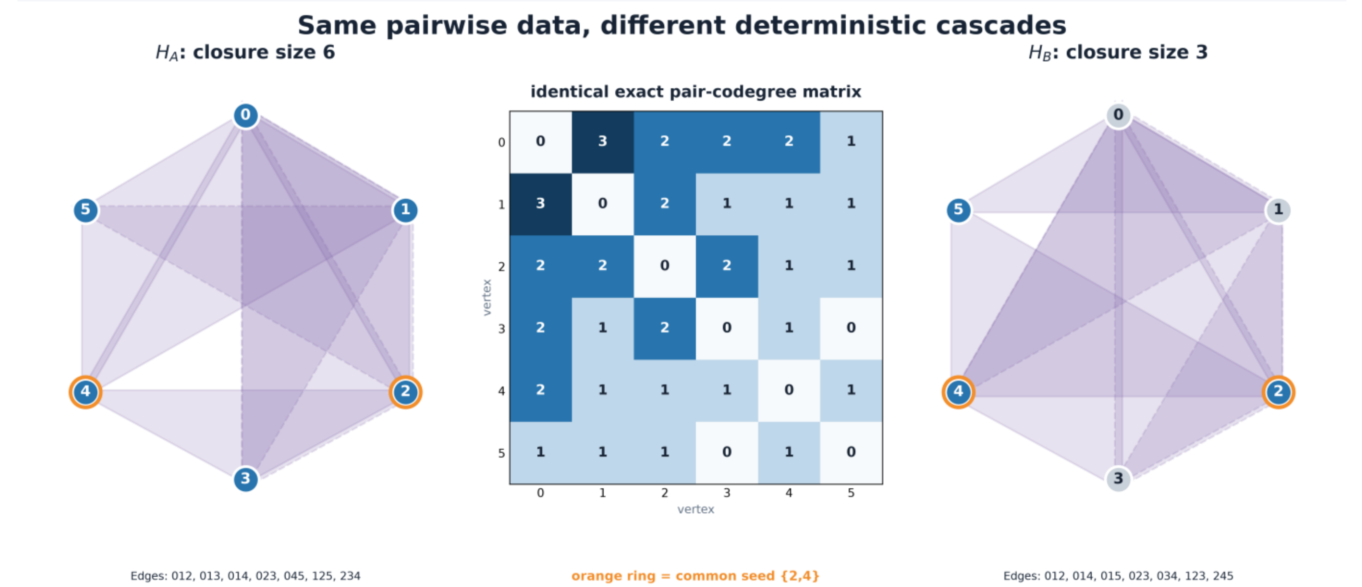

Figure 1. The supplied non-isomorphic witness. Translucent triangles are three-way hyperedges, not ordinary graph edges.

1. Why pairwise compression can be dynamically insufficient

A 3-uniform hypergraph records interactions among triples. Its weighted pairwise projection keeps, for every labeled pair {i,j}, the exact number c(i,j) of triples containing that pair. This is a common form of higher-order compression: it preserves all pair multiplicities but forgets which three pairs belong to the same triple [1–3]. The question is not whether some information is lost—that is established—but how much one precise deterministic outcome can vary while every recorded pair count remains fixed.

Definitions frozen before the final run

- Object: a labeled simple 3-uniform hypergraph H=([n],E), with no loops or repeated triples.
- Projection: Φ(H)=(c(i,j)) for i<j; a fiber F(P) contains all H with exactly the same labeled integer matrix P.
- Dynamics: from seed S, repeatedly activate the third vertex of any triple whose other two vertices are active. The least fixed point is clH(S).
- Metric: $\Delta_2(P)$ is the largest absolute difference in |clH(S)| over two members of F(P) and the same named two-vertex seed S. Then $M_2(n)$ is the maximum of $\Delta_2(P)$ over all projections P.

The process is equivalent to forward chaining on the three Horn implications ab→c, ac→b, bc→a supplied by every triple abc. Monotonicity on a finite vertex set makes the final closure independent of synchronous or asynchronous update order.

Main theorem

PROVED BY HUMAN-READABLE ARGUMENT $M_2(n)=0$ for 2≤n≤5, and $M_2(n)=n-3$ for every n≥6.

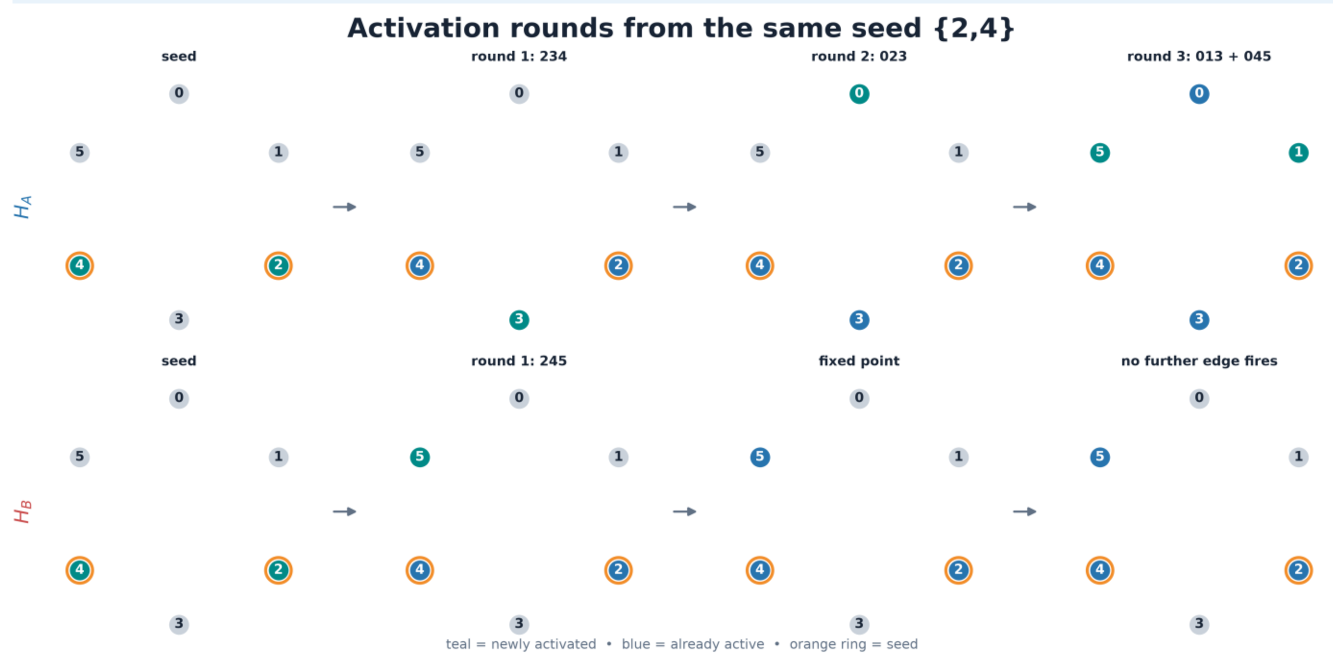

Figure 2. Explicit rounds verify the six-versus-three closure gap. The final fixed point is update-order independent.

2. Proof strategy and edge-minimum hierarchy

Upper bound: why the smaller cascade cannot stay at two

Fix a seed {u,v}. A first activation exists exactly when some triple contains that exact pair, which is equivalent to $c(u,v)>0$. Two fiber members have the same $c(u,v)$. Therefore either both closures remain size two, giving gap zero, or both have size at least three. In the second case both sizes lie between 3 and n, so their difference is at most $n-3$. This proves $M_2(n)\leq n-3$.

Lower bound: a five-edge core that extends

Let A_5={013,045,125,234,023} and B_5={015,034,123,245,023}. The four differing triples form a pair-balanced Pasch trade, so $\Phi(A_5)=\Phi(B_5)$. From seed {2,4}, A_5 activates 3, then 0, then 1 and 5; B_5 activates only 5. The closures have sizes six and three. For every new vertex x, add the same triple 01x to both sides. The first branch already activates 0 and 1 and therefore every x; the second never activates either. The gap becomes $n-3$.

Why nothing happens before six vertices

On five vertices, let m be one third of the sum of all pair codegrees, and let d(i) be one half of the codegree sum around vertex i. If T is the three-vertex set left after removing a and b, then $1[T\in E]=m-d(a)-d(b)+c(a,b)$: the right side counts edges avoiding a and b, and T is the only possible such triple. Hence the projection reconstructs every edge through $n=5$.

The 4/5/7 hierarchy

Canceling common edges from a projection collision leaves a pair-balanced trade. Every nonzero simple 2-(v,3) trade has at least four blocks per side; at volume four it is a Pasch trade up to relabeling [5]. A pure Pasch trade can change which third vertex activates, but both closures have the same size. One common context edge is therefore necessary and sufficient for a size difference: four edges first allow different final sets, and five first allow different sizes. The unqualified claim “seven edges are minimal” is false. Seven is the independently verified threshold only after additionally requiring non-isomorphic members on six vertices.

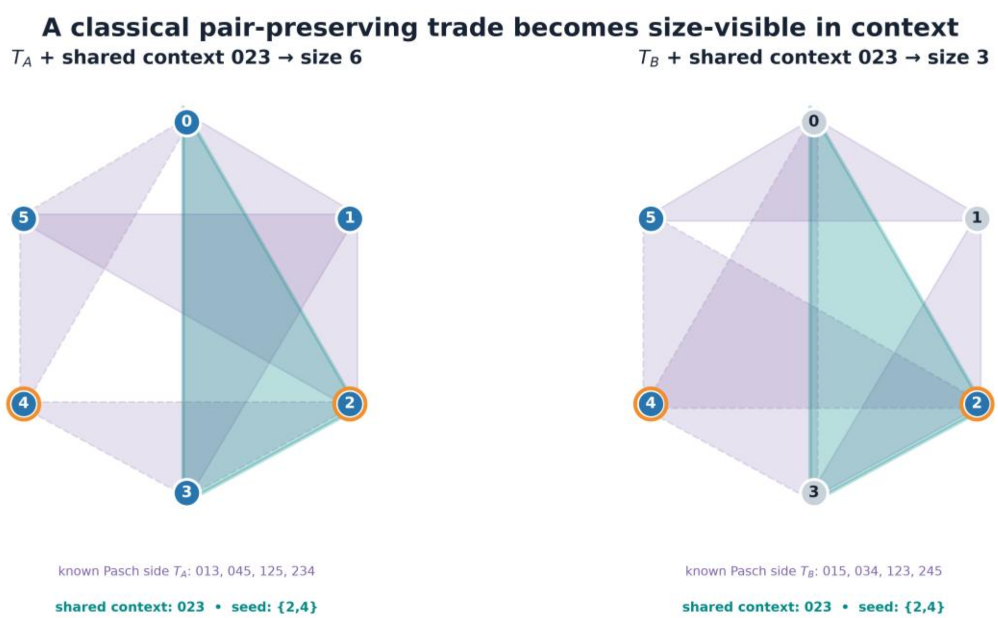

Figure 3. The Pasch trade is classical; the theorem analyzes its exact closure effect in context.

3. Fresh computation as a finite proof and regression check

The all-n theorem is deductive. Computation serves a different purpose: it completely audits the finite n=6 search space, verifies qualified minimum claims, and provides independently checkable certificates. There are C(6,3)=20 possible triples and therefore 2^{20}=1,048,576 labeled simple hypergraphs.

Census output	Fresh value	Evidence status
Labeled hypergraphs	1,048,576	complete enumeration
Distinct projections	995,085	two implementations agree
Non-singleton fibers	44,521	two implementations agree
Fibers with size disagreement	2,750	two implementations agree
Maximum two-seed size gap	3	two implementations agree
Shared cross-check fields	69; zero mismatches	isolated rerun

Method. The primary implementation enumerates hypergraphs as 20-bit masks, hashes 15-entry projection tuples, and computes all two-seed closures only inside collision fibers. A separately written verifier uses a different direct representation and independently enumerates the same full space. A Z3 model separately checks edge-cap satisfiability, and five machine-readable certificates are recomputed rather than trusted. The final isolated environment executed both test suites, both censuses, the solver checks, the 69-field comparison, and certificate verification in 74.51 seconds.

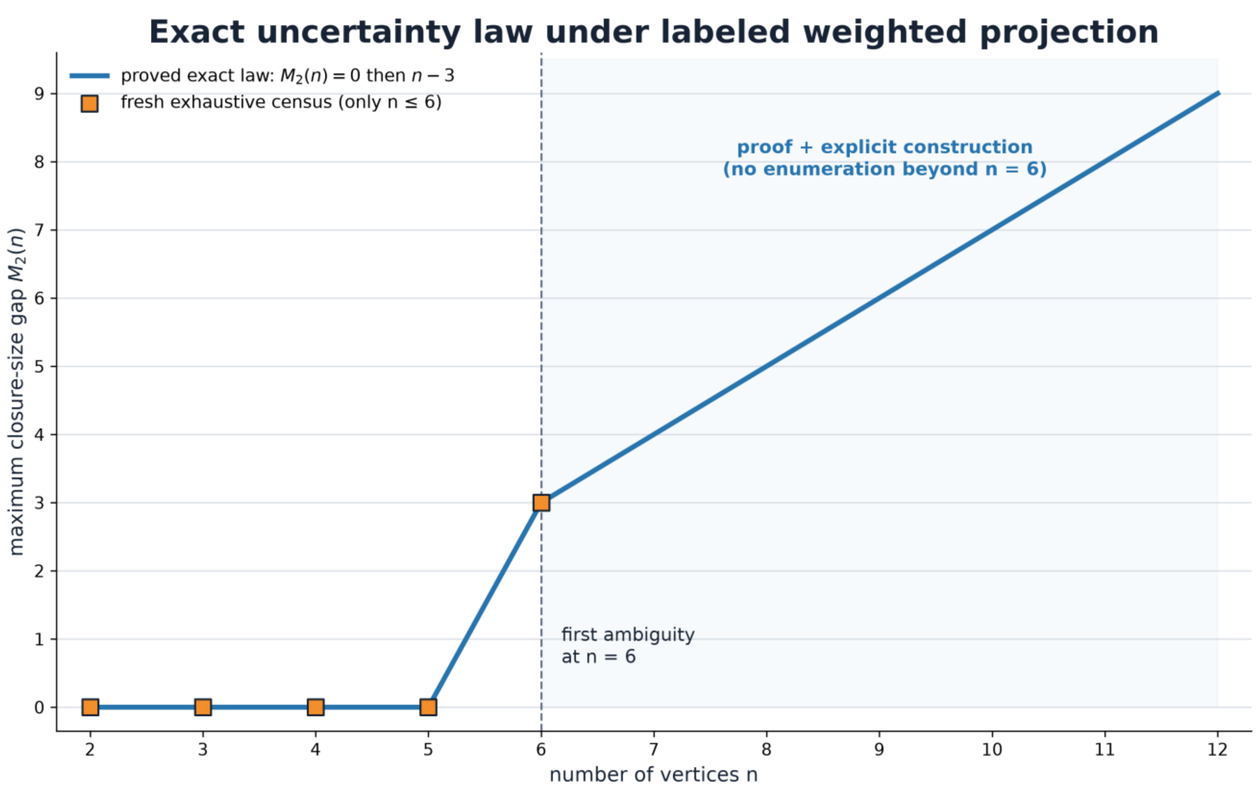

Figure 4. Exhaustive markers stop at n=6. Values for n>6 come only from the proved upper bound and explicit family.

4. What is new, what is known, and what the result means

Adversarial novelty finding

The broad story is prior art. Weighted 2-sections and reconstruction are established [1,6,7]; exact node co-occurrence fibers ("twin hypergraphs") are studied by LaRock and Lambiotte [2]; Pasch trades are classical [5]; and Salova and D'Souza already exhibit non-isomorphic six-node, seven-edge hypergraphs with the same weighted projection but different synchronization dynamics [3]. Exact source recovery showed that their construction uses the same labeled Pasch trade as our supplied witness, with a different context. The witness mechanism must not be marketed as new.

CONJECTURAL NOVELTY Our bounded primary-source search through July 13, 2026 found no paper defining or determining the worst-case two-seed closure-size gap inside a labeled exact-codegree fiber. Non-discovery is not proof of priority.

Limitations

- One abstract deterministic unique-healthy-vertex H-bootstrap rule is analyzed; other contagion, timing, stochastic, or threshold rules may behave differently.
- The primary invariant uses labeled exact integer codegrees and the same named seed. Binary projections, unlabeled matrices, multisets, and other seed sizes define different problems.
- Only $n \leq 6$ was enumerated. Larger-n values are theorem-derived, not computational observations.
- No empirical network or real epidemic is modeled. The result is a representational warning, not a claim about real-world effect size.
- The six-vertex non-isomorphic threshold is a computational proof under the stated exhaustive-search and software assumptions; detailed independent-only orbit totals are not used centrally.

Conclusion and impact

Exact pairwise co-occurrence counts can determine every triple through five vertices yet lose $n-3$ vertices of cascade-size information from six onward. The result converts a qualitative warning—"projection can matter"—into a sharp, reproducible extremal law for a precisely defined closure process. It also separates three notions reviewers often conflate: collision, final-set disagreement, and final-size disagreement. The practical lesson is narrow but rigorous: when a process genuinely depends on triples, an exact pairwise projection need not be a sufficient statistic for its outcome.

Evidence status

The theorem and unrestricted four-/five-edge minima are PROVED BY HUMAN-READABLE ARGUMENT. The six-vertex census, supplied witness, certificates, and qualified seven-edge threshold are EXECUTED AND VERIFIED. Priority of the exact invariant remains CONJECTURAL.

References

[1] R. Janczewski, P. Obszarski, and K. Turowski, "Weighted 2-sections and hypergraph reconstruction," Theoretical Computer Science 915 (2022), 11–25. https://doi.org/10.1016/j.tcs.2022.02.016

[2] T. LaRock and R. Lambiotte, "Exploring the Non-uniqueness of Node Co-occurrence Matrices of Hypergraphs," arXiv:2506.01479v2 (2025). https://arxiv.org/abs/2506.01479

[3] A. Salova and R. M. D'Souza, "Analyzing states beyond full synchronization on hypergraphs requires methods beyond projected networks," arXiv:2107.13712 (2021). https://arxiv.org/abs/2107.13712

[4] N. Morrison and J. A. Noel, "A sharp threshold for bootstrap percolation in a random hypergraph," Electronic Journal of Probability 26, Article 97 (2021), 1–85. https://doi.org/10.1214/21-EJP650

[5] Z. Eslami and B. Tayfeh-Rezaie, "On 2-(v,3) trades of minimum volume," Australasian Journal of Combinatorics 19 (1999), 239–251. https://ajc.maths.uq.edu.au/pdf/19/ocr-ajc-v19-p239.pdf

[6] G. Bresler, C. Guo, Y. Polyanskiy, and A. Yao, "Partial and Exact Recovery of a Random Hypergraph from its Graph Projection," Proceedings of COLT 2025, PMLR 291, 534–593. https://proceedings.mlr.press/v291/bresler25a.html

[7] J.-G. Young, G. Petri, and T. P. Peixoto, "Hypergraph reconstruction from network data," Communications Physics 4, 135 (2021). https://doi.org/10.1038/s42005-021-00637-w

[8] M. Lucas et al., "Reducibility of higher-order networks from dynamics," Nature Communications 17, 1551 (2026). https://doi.org/10.1038/s41467-025-68273-4

[9] M. Xie et al., "Transformability of dynamics on higher-order networks," Communications Physics 9, 149 (2026). https://doi.org/10.1038/s42005-026-02555-1