BCQM VII Path A Validation: Test Plan Beyond Stability

Peter M. Ferguson
Independent Researcher

5 February 2026 (Progress Update v0.2)

Executive Summary

Gate 1 (stability) has been cleared: community partitions show strong agreement (NMI ∼0.83–0.84, ARI moderate-to-strong), super-graph edge sets are moderately stable (Jaccard ∼0.37–0.60) with high weight correlations at $n = 0.8$ (∼0.89–0.91), and super-graph ball-growth diagnostics are stable across seeds. Gate 3 tightening via a Louvain resolution sweep confirms robustness for $\gamma \geq 1.0$ (with $\gamma = 0.5$ an outlier). Gate 4 scale-up runs at $N = 16$ and $N = 32$ (hits1, $n = 0.8$) remain connected and diagnostic-stable; at $N = 32$ a substantial halo emerges while cloth- and super-graph-level diagnostics remain stable and the clock signal remains robust (core fraction Φ drops, while Q_{clock} stays high). Gate 4 localisation remains strongly local on the community cloth when partitions/super-graphs are built on all-used edges/flows: hops are confined to $d \in \{0, 1, 2\}$ with $d \geq 3$ absent. Test 2.3 (curvature proxy) and Test 2.1 (spectral dimension) have now been executed as non-blocking diagnostics: curvature summaries are stable across seeds, while spectral-dimension curves show no robust plateau at current community sizes (finite-size / fast-mixing dominated), consistent with prior cautions.

As of this update, the BCQM VII manuscript draft is in a near-final state (compile-clean), with figures generated (Fig. 1–9) and remaining work primarily limited to figure cosmetics (legend placement and removal of embedded "Fig. x" text) and final literature-context citations.

This document outlines the remaining validation tests for Gates 2–4 to establish that the super-graph is a genuine geometric object with physical significance, and it records progress and methodological choices (notably A2: core/core vs all/all) where they affect interpretation and coverage.

Overview: Validation Gates

- **Gate 1 (Stability):** PASSED — community structure, super-graph edges, and super-graph ball-growth are stable across seeds (with robustness to Louvain resolution for $\gamma \geq 1.0$).
- **Gate 2 (Geometry):** IN PROGRESS — Test 2.4 (d_{eff} from ball growth) completed; Tests 2.1 (spectral dimension) and 2.3 (curvature) remain.
- **Gate 3 (Physics):** IN PROGRESS — thread locality on the super-graph (Gate-4 localisation proxy) implemented and tested; remaining Gate 3 tests relate to interaction/defect interpretations.
- **Gate 4 (Scaling):** IN PROGRESS — scale-up runs completed at $N = 16$ and $N = 32$ for $n = 0.8$; further N-scans remain.

1 Gate 2: The Super-Graph Has Geometric Structure

1.1 Test 2.1: Spectral Dimension

What

Calculate the spectral dimension d_s of the super-graph from the return probability of a random walk.

Method

1. Run random walks on super-graph nodes
2. Measure return probability $P(t)$ at time t
3. Plot $\log P(t)$ vs $\log t$
4. Spectral dimension: $d_s = -2 \times \text{slope}$

Why It Matters

Real emergent geometries have finite d_s even when the graph is topologically complex. CDT and other quantum gravity models predict $d_s \approx 2$ at short scales, $d_s \approx 4$ at long scales. If the super-graph shows $d_s \approx$ 2–4, that is a strong hint of emergent spacetime structure.

Pass Criterion

d_s is finite, stable across seeds (standard deviation $< 10\%$ of mean), and in physically plausible range (1.5–4.0).

Implementation Notes

- Use 1000+ random walk samples per super-graph
- Sample return probability at $t \in \{1, 2, 4, 8, 16, 32, \ldots\}$ up to $t_{\max} \sim K/2$
- Average d_s across all seeds at each (N, n) configuration

> **Finite-Size Warning:** Spectral dimension extraction suffers severe finite-size effects for $K < 50$ nodes. Random walks hit graph boundaries before establishing a clean power-law regime. At $N = 8$ ($K \approx$ 22–24), d_s estimates will be noisy and may not be physically meaningful. Do not treat poor d_s results at small N as evidence of geometry failure. Rely on Test 2.4 (ball-growth dimension) for Phase 1 validation. Consider d_s reliable only at $N \geq 32$ ($K \geq 50$).

Progress (as of 3 February 2026)

Spectral-dimension estimates were attempted on the core/core community super-graphs using Monte-Carlo return probabilities and log-derivative evaluation of $d_s(t)$ over $t \in [10, 80]$. While the mean effective values are in-family and consistent at $n = 0.8$, the within-run variability remains large (no clear plateau), and the low-n cases are strongly finite-size/noise dominated. We therefore treat Test 2.1 as executed but not yet a robust "dimension" diagnostic at the present super-graph sizes. A planned refinement is to replace Monte-Carlo return estimates with an exact transition-matrix evaluation of $P_0(t)$ (noise-free), and then re-assess whether any plateau emerges.

Progress (as of 3 February 2026)

Test 2.1 was executed on the community super-graph using an *exact* return probability computation $P_0(t) = \frac{1}{K}\mathrm{tr}(P^t) = \frac{1}{K}\sum_i \lambda_i^t$ (noise-free), followed by the standard log-derivative estimate $d_s(t) = -2\, d\log P_0(t)/d\log t$. At the present super-graph sizes ($K \sim$ 20–80), $d_s(t)$ does not exhibit a robust plateau over a broad diffusion-time range: in the high-coherence regime ($n = 0.8$, $K \approx 22$) the walk mixes rapidly and $d_s(t)$ drifts toward 0, while at lower coherence ($n = 0.4$) a short early-time transient is visible but decays with t and remains finite-size dominated. Accordingly, Test 2.1 is recorded as executed but not yet a stable "dimension" diagnostic at Stage–2; ball-growth and curvature proxies remain the primary geometry diagnostics.

1.2 Test 2.3: Curvature Proxy on the Community Super-Graph (Forman–Ricci)

What

Compute curvature-like diagnostics on the community super-graph built from Path A runs. We use graph-native proxies: (i) unweighted Forman edge curvature $F(e) = 4-\deg(u)-\deg(v)$ and (ii) augmented Forman curvature $F_{\mathrm{aug}}(e) = F(e)+3T(e)$, where $T(e)$ is the number of triangles containing edge e. We also report transitivity and mean clustering coefficient as triangle-closure sanity checks.

Progress (as of 3 February 2026)

This test has been executed on the pivot baseline ensemble (hits1, x10 epoch, bins=20; $N \in \{4, 8\}$, $n \in \{0.4, 0.8\}$) using `bcqm_vii_cloth/analysis/curvature_supergraph_forman.py`. Results show a clear regime dependence and an important "graph-choice" sensitivity:

- At $n = 0.8$, the community super-graph is sparse and triangle-free (transitivity = 0, clustering = 0) for both $N = 4$ and $N = 8$; core/core and all/all are identical and $F_{\mathrm{aug}} = F$.
- At $n = 0.4$, core/core retains nonzero triangle closure (e.g. for $N = 8$: transitivity ≈ 0.088, clustering ≈ 0.055), and F_{aug} differs substantially from F, whereas all/all collapses the community graph ($K \approx 24$) and eliminates triangle closure (transitivity = 0, clustering = 0).

Interpretation: halo edges act as shortcut closures at low n, collapsing the mesoscopic super-graph and suppressing loop/triangle structure. Therefore curvature proxies should be evaluated on core/core super-graphs by default (Gate 2), while all/all remains appropriate for localisation coverage (Gate 4) with coverage explicitly reported.

Artefacts

CSV outputs:

- `csv/curvature/curvature_pivot_core_supergraph_curvature_summary.csv` and `..._runs.csv`
- `csv/curvature/curvature_pivot_all_supergraph_curvature_summary.csv` and `..._runs.csv`

1.3 Test 2.4: Effective Dimensionality from Ball Growth

What

Extract an "effective" scaling exponent from ball-growth curves by fitting $\log|B(r)|$ versus $\log r$ over an intermediate window:

$$|B(r)| \sim r^{d_{\text{eff}}}.$$

Method

1. Use existing ball-growth data: cloth-level ball growth from the `RUN_METRICS` key `cloth/ball_growth`; optionally, super-graph ball growth computed on the community super-graph.
2. Choose a contiguous scaling window that avoids very small r and avoids saturation ($|B(r)| \approx |C|$).
3. Fit $\log|B(r)|$ versus $\log r$ in that window; slope gives d_{eff}; report R^2.

Progress (as of 3 February 2026)

- **Cloth (hits1, $n = 0.8$, A3 scale-ups):** for $N = 16$ and $N = 32$, no admissible intermediate scaling window is found; d_{eff} is therefore *undefined* (reported as `NaN`) for all seeds. This remains true under loosened window parameters. Interpretation: the cloth core saturates at very small radius (often $r \leq 2$), consistent with a small-world/shortcut-rich cloth in this high-coherence regime.
- **Super-graph (pivot baseline):** a short but consistent scaling window exists (typically $r \approx 3$–6), yielding stable $d_{\text{eff}} \approx 0.78$–$0.79$ with excellent fit quality ($R^2 \approx 0.99999$) across quadrants.

Why It Matters

This test clarifies where "dimension-like" scaling is meaningful. In the high-n cloth regime, rapid saturation implies no usable power-law window. At the community super-graph level, the sparse coarse graph admits an intermediate window, making d_{eff} a meaningful diagnostic at that scale. This supports the Stage–2 pivot: geometry-like scaling behaviour can emerge more cleanly on coarse objects than on molecular edge sets.

Revised pass criterion

- For super-graphs: d_{eff} exists with high R^2 and is stable across seeds (low standard deviation relative to mean).
- For cloth cores in high-connectivity regimes: "no window found" is treated as a diagnostic signature of rapid saturation (small-world cloth), not a failure.

Implementation Notes

- The analysis is performed by `bcqm_vii_cloth/analysis/d_eff_ball_growth.py`.
- Outputs are recorded in the local `csv/` folder (see Evidence Manifest).

2 Gate 3: Super-Graph Connects to Physical Observables

2.1 Test 3.1: Thread Bundles as "Particles" on the Cloth

What

Threads (sequences of events from BCQM IV/V) should map to trajectories on the super-graph that are local in the emergent geometry.

Method

1. Take a thread (sequence of events)
2. Map each event to its community $\rightarrow$ sequence of super-graph nodes
3. Check if thread-induced paths on super-graph are "local" (consecutive super-graph nodes are adjacent or close)
4. Compute "localization index":

$$\text{LOC} = \frac{\text{\# thread steps with super-graph hop distance} \leq 1}{\text{total \# thread steps}}$$

5. Compare to random baseline: shuffle community assignments and recompute LOC

Why It Matters

This is the most important test in the validation plan. If threads are physical (particle worldlines), they should be local on the emergent geometry. Random threads would hop randomly across the super-graph. Physical threads should trace out geodesic-like paths on the cloth. **This test directly establishes whether particles "live on" the emergent spacetime**—the core compatibility between quantum mechanics (threads) and general relativity (super-graph geometry).

Pass Criterion

Threads show strong localization: LOC > 0.70, and significantly above random baseline ($p < 0.001$ in permutation test).

Implementation Notes

- Identify threads using BCQM V thread-formation algorithm
- For each thread, record sequence of communities visited
- Compute super-graph hop distance between consecutive community pairs
- Generate 1000 random permutations of community labels to establish baseline

2.2 Test 3.2: Cross-Link Density as "Interaction Strength"

What

Super-graph edge weights (number of Path A cross-links between communities) should correlate with physical interaction between threads in those communities.

Method

1. Identify "interacting" threads: threads that share events or are causally linked
2. For each thread pair, record their home communities (C_i, C_j)
3. Check if these communities are connected in super-graph, and record edge weight w_{ij}
4. Compute correlation (Spearman ρ) between super-graph edge weight and thread interaction frequency

Why It Matters

If the super-graph encodes spatial proximity, then threads in nearby communities (high w_{ij}) should interact more frequently. This tests whether cloth-level adjacency $\rightarrow$ micro-level interaction.

Pass Criterion

Significant positive correlation: Spearman $\rho > 0.5$, $p < 0.01$.

Implementation Notes

- Define "interaction": threads that share ≥ 1 event, or have causal links (event from thread A in causal past of event from thread B)
- Aggregate interaction counts across all thread pairs in same community pair
- Report scatterplot: $x = w_{ij}$, $y =$ interaction count

2.3 Test 3.3: Metric Reconstruction from Cross-Link Data

What

Use super-graph edge weights to define a proto-metric $g_{\mu\nu}$ on the cloth via metric embedding.

Method

1. Convert edge weight w_{ij} to distance: $d_{ij} = 1/w_{ij}$ (or $d_{ij} = -\log w_{ij}$)
2. Use metric embedding (MDS, Isomap, or force-directed layout) to assign coordinates $\{x_i\}$ to communities
3. Check embedding quality: stress (residual error), embedding dimension d_{embed}
4. Test if resulting metric has sensible properties: triangle inequality, smoothness

Why It Matters

This is the "metric extraction" step needed for BCQM VII/VIII. If Path A cross-links encode spatial proximity, this should produce a consistent metric. If it fails, Path A might not be encoding geometry correctly.

Pass Criterion

- Metric embedding produces stable coordinates: stress < 0.15
- Embedding dimension $d_{\text{embed}} \leq 4$
- Coordinates stable across seeds (Procrustes alignment shows > 0.8 similarity)
- Reconstructed metric approximately satisfies triangle inequality ($\eta > 0.90$)

Implementation Notes

- Use `scikit-learn.manifold.MDS` or similar
- Try embedding dimensions $d \in \{2, 3, 4\}$, report which gives lowest stress
- For stability check: embed super-graphs from different seeds, align via Procrustes, compute coordinate RMSD
- **Note:** This test subsumes the non-trivial version of Test 2.2 (weighted triangle inequality)

3 Gate 4: Scaling Behavior with N

3.1 Test 4.1: Community Count Scaling

What

Determine how K (community count) scales with N at fixed n.

Method

1. Plot K vs N for $N \in \{4, 8, 16, 24, 32, 48, 64, 96, 128\}$ at fixed $n \in \{0.6, 0.8\}$
2. Fit power law: $K = A \cdot N^\alpha$
3. Extract scaling exponent α and goodness-of-fit R^2

Why It Matters

If the super-graph is a "finite-size cloth", expect $K \sim N^\alpha$ with $\alpha < 1$ (sublinear). If every event is its own community, $K \sim N$ (bad—no coarse-graining). If communities merge well, expect $\alpha \approx 0.5$–0.7 (good coarse-graining).

Pass Criterion

$K \sim N^\alpha$ with $\alpha < 0.8$, stable fit ($R^2 > 0.95$) across n regimes.

Implementation Notes

- Average K across seeds at each N
- Fit on log-log plot: $\log K$ vs $\log N$
- Report separate α for each n regime

3.2 Test 4.2: Spectral Dimension Convergence

What

Measure d_s at multiple N, check for convergence as $N \to \infty$.

Method

1. Compute d_s (from Test 2.1) at each $N \in \{4, 8, 16, 32, 64, 128\}$
2. Plot $d_s(N)$ vs N
3. Fit to asymptotic form: $d_s(N) = d_s^\infty + A/N^\beta$
4. Check if d_s flattens (derivative $\to 0$) by $N = 128$

Why It Matters

True emergent geometry should have d_s that stabilizes at large N. If d_s keeps drifting, the super-graph might be a system-size artifact rather than a genuine geometric object.

Pass Criterion

$d_s(N)$ flattens: $|d_s(128) - d_s(64)| < 0.2$. Converges to asymptotic value $d_s^\infty \in [1.5, 4.5]$.

Implementation Notes

- Use same random walk protocol at all N
- Report error bars (standard deviation across seeds) at each N

3.3 Test 4.3: Super-Graph Diameter Scaling

What

Measure graph diameter $D_{\max}$ of super-graph giant component as function of N.

Method

1. For each N, compute diameter $D_{\max}$ = longest shortest path in super-graph
2. Plot $\log D_{\max}$ vs $\log N$ (or $\log K$)
3. Fit scaling law: $D_{\max} \sim N^\gamma$ or $D_{\max} \sim \log N$

Why It Matters

Expected scaling laws:

- Small-world network: $D_{\max} \sim \log K \sim \log N$
- Spatial network in d dimensions: $D_{\max} \sim N^{1/d}$

If super-graph has emergent dimensionality d, diameter should scale like $N^{1/d}$. This tests whether the cloth has spatial structure vs just being well-connected.

Pass Criterion

Clear scaling law ($R^2 > 0.90$), consistent with $d = 2$–4. Example: $D_{\max} \sim N^{1/3}$ suggests $d \approx 3$.

Implementation Notes

- Use NetworkX `diameter()` or `eccentricity()` functions
- If disconnected, compute diameter of giant component only
- Average across seeds at each N

4 Implementation: 3-Phase Plan (Revised)

4.1 Phase 1: Core Validation (Weeks 1–2)

Priority tests (revised based on methodological review):

1. **Test 2.4: Ball-growth effective dimension** (most robust for small K, data already available)
2. **Test 3.1: Thread localization** (most important conceptual validation—establishes physics connection)
3. **Test 4.1: Community count scaling** (validates coarse-graining is occurring)
4. **Test 2.1: Spectral dimension** (OPTIONAL—informative but noisy at $N = 8$)

Tests DEFERRED from Phase 1:

- triangle inequality (unweighted)—tautological for unweighted graphs; merge into Test 3.3 in Phase 3

Datasets:

- Use existing $N \in \{4, 8\}$, $n \in \{0.4, 0.8\}$, 5 seeds each
- Add $N = 16$ if time permits (from scaling benchmark)

Deliverable:

- Short report: "Super-graph is geometric and physical"
- 3 core measures: d_{eff} (geometry), LOC (physics), α (coarse-graining)
- Decision point: If all three pass, proceed to Phase 2. If Test 3.1 (thread localization) fails, major red flag—geometry may not be where physics lives.

4.2 Phase 2: Extended Validation (Weeks 3–4)

Priority tests:

1. Test 3.2: Cross-link density vs interaction (validates Path A mechanism)
2. Test 2.1: Spectral dimension at larger N (if deferred from Phase 1)
3. Test 4.2: Spectral dimension convergence (requires multiple N runs)

Datasets:

- Extend to $N \in \{4, 8, 16, 32\}$ for scaling tests
- Use 5–10 seeds per N at $n \in \{0.6, 0.8\}$

Deliverable:

- Report: "Path A mechanism connects cloth to micro-dynamics"
- Evidence that cross-link strength predicts thread interactions
- Decision point: If tests pass, Path A is validated. Proceed to full N-scan or Phase 3 deep validation.

4.3 Phase 3: Advanced Validation (Weeks 5–6, Optional)

Deep tests (if Phase 1–2 pass):

1. Test 2.3: Curvature distribution (claim geometric richness)
2. Test 3.3: Metric reconstruction (extract explicit $g_{\mu\nu}$, includes weighted triangle inequality check)
3. Test 4.3: Super-graph diameter scaling (spatial vs small-world)

Datasets:

- Full N-scan: $N \in \{8, 16, 24, 32, 48, 64\}$ (stop before $N = 128$ for now)
- 10–30 seeds per N depending on computational cost

Deliverable:

- Full validation report: "Path A as geometric mechanism"
- Complete evidence package for BCQM VI paper
- Readiness for $N = 128$ production runs and BCQM VII metric extraction

5 Success Criteria by Strength

5.1 Minimal Viable Result (Sufficient for BCQM VI)

- Effective dimension d_{eff} from ball growth in physical range ($d_{\text{eff}} \in [1.5, 4.5]$, std dev $< 15\%$)
- Thread localization significantly above random (LOC > 0.7, $p < 0.001$)
- Community count scaling sublinear: $K \sim N^{\alpha}$ with $\alpha < 0.8$

Interpretation: Path A creates stable coarse-grained structure with geometric properties that particles respect. Sufficient to publish as "mechanism exploration" paper.

5.2 Strong Result (Supports BCQM VII Continuation)

All minimal criteria, plus:

- Spectral dimension d_s finite and stable at $N \geq 32$ ($d_s \in [1.5, 4]$, std dev $< 10\%$)
- Cross-link density predicts thread interaction: Spearman $\rho > 0.5$
- Triangle inequality holds for weighted distances: $\eta > 0.95$

Interpretation: Super-graph is genuine emergent geometry encoding physical proximity. Strong justification for metric extraction in BCQM VII.

5.3 Publication-Grade Result

All strong criteria, plus:

- Non-trivial curvature distribution (not flat, stable across seeds)
- Successful metric embedding: stress < 0.15, $d_{\text{embed}} \leq 4$
- Clear scaling laws with N: d_s converges, diameter scales as $N^{1/d}$

Interpretation: Path A generates emergent spacetime with identifiable dimensionality, curvature, and metric structure. Publication target: PRL, PRD, or similar high-impact venue.

6 Timeline and Resource Requirements

6.1 Computational Resources

- **Phase 1–2:** Can run on existing hardware (laptop or single Mac Mini)
- **Phase 3:** May require Mac Mini cluster if extending to $N = 64$
- **Scaling benchmark:** Run before committing to Mac Mini purchase

6.2 Timeline

Phase	Duration
Phase 1 (Core validation)	2 weeks
Decision point 1	2 days
Phase 2 (Extended validation)	2 weeks
Decision point 2	2 days
Phase 3 (Advanced validation, optional)	2 weeks
Final report & BCQM VI draft outline	1 week
Total	**6–9 weeks**

6.3 Decision Points

After Phase 1:

- If core tests pass (geometry + physics + coarse-graining) → proceed to Phase 2
- If Test 3.1 (thread localization) fails → **major concern**—super-graph may be geometric artifact not connected to physics; diagnose or consider pivot
- If Test 2.4 or 4.1 fail → diagnose (implementation bug? wrong coarse-graining? Path A insufficient?)

After Phase 2:

- If extended tests pass → Path A validated, proceed to full N-scan or Phase 3 deep validation
- If physics tests fail → super-graph is geometric but not physical; consider pivot to alternative cross-link mechanism

7 Next Immediate Actions (Revised)

1. **Figures (paper polish):**
   Adjust legend placement and remove embedded "Fig. x" text in the generated figure PDFs; confirm captions match plotted content.

2. **Literature context:**
   Add a short, scoped "Relation to other work" paragraph (graph/emergent-geometry models) and ensure core Stage–2 claims are appropriately cited in the main text without expanding scope.

3. **Citations and cross-references:**
   Audit figure callouts, labels, and citation placement for the Gates and Discussion (no new results; presentation-grade polish only).

4. **Optional robustness cross-check: Leiden vs Louvain:**
   Treat Leiden as an optional algorithm-independence check on the baseline cloth partition stability window; record as a non-blocking appendix note if performed.

Evidence Manifest

All run folders, CSV artefacts, and commands to regenerate them are recorded in `docs/EVIDENCE_MANIFEST.md` (or the local evidence manifest file). The current Stage–2 evidence set includes:

- Baseline and A3 scale-up run directories under `outputs_cloth/` (including $N = 16$ and $N = 32$ at $n = 0.8$, hits1).
- Consolidated analysis tables under `csv/` (partition stability, super-graph stability, ball-growth d_{eff}, curvature, spectral-dimension curves, and Gate–4 localisation outputs).
- Figure generator scripts (matplotlib) and rendered PDFs under `figures/` (Fig. 1–9), including the schematics and the CSV-driven plots.
- Repeatability check artefacts (repA/repB) demonstrating stability of Φ and Q_{clock} and strict locality ($d \geq 3$ absent), with the $d = 0$ versus $d = 1$ mix variable at current sample size.

Conclusion

This test plan (revised after methodological review) provides a systematic path to validate that the super-graph is not just stable (Gate 1), but genuinely geometric (Gate 2), physically meaningful (Gate 3), and scalable (Gate 4).

Key revisions:

- Prioritized Test 3.1 (thread localization) as the conceptual linchpin—it directly tests whether thread trajectories respect the emergent coarse geometry (super-graph locality), without assuming a manifold.
- Promoted Test 2.4 (ball-growth dimension) as the primary Phase 1 geometry diagnostic—more robust than spectral dimension at small K.
- Flagged triangle inequality (unweighted) as tautological for hop-distance graphs; deferred to Phase 3 (weighted/embedded-distance variants only).

- Added finite-size warnings for Test 2.1 (spectral dimension) at small super-graph size.

Phase 1 status (completed on current datasets). Test 2.4 and Test 3.1 have been executed on existing runs, and Test 4.1 scaling has been initiated via A3 scale-ups:

- **Test 2.4 (effective dimensionality from ball growth):** cloth-level d_{eff} is *undefined* in the high-connectivity hits1 regime (rapid saturation at $r \leq 2$; no admissible scaling window), which is treated as a diagnostic signature of a shortcut-rich cloth; super-graph d_{eff} is defined and stable with excellent fit quality in a short window.

- **Test 3.1 (thread localization):** traced thread motion is strongly local on the community super-graph (0–2 hops only) in both high- and low-coherence regimes when the partition/super-graph are built on all-used edges/flows; core-only constructions can fail coverage and are therefore not suitable for trajectory claims at larger N.

- **Test 4.1 (scaling):** A3 stress tests at $N = 16$ and $N = 32$ (hits1, $n = 0.8$) preserve connected cloth cores and stable diagnostic behaviour; at $N = 32$ a substantial halo emerges while locality on the super-graph persists.

Completion of Phase 1 establishes that Path A produces a reproducible coarse geometry object (communities and super-graph) that thread trajectories respect—sufficient to justify continuing and drafting BCQM VII. Phase 2 (publication-grade result) will extend geometry validation (e.g. curvature proxies) and formalise scaling/robustness across regimes and larger N.

Status: Phase 1 validation is completed on current datasets (Tests 2.4, 2.3, 2.1, and Gate–4 localisation; A3 scale-ups). BCQM VII drafting is substantially complete; remaining work is presentation-grade polish (figures and scoped literature context) plus an optional Leiden cross-check for algorithm-independence.