BCQM VII Lab Note
Repeatability check: two consecutive N=32 cloth runs (v0.1)

Peter M. Ferguson
Independent Researcher

4 February 2026

Purpose

Assess whether stochasticity in the thread dynamics produces materially different *mesoscopic* Stage–2 outcomes when the *same* configuration is executed twice with different RNG seed ranges. Concretely, we test repeatability of: (i) core/halo split (Φ), (ii) clock quality (Q_{clock}), (iii) cloth health and size summaries, and (iv) Gate–4 localisation (hop-distance distribution on the community cloth).

Configuration (fixed across both runs)

Headline case (Stage–2 stress-test regime):

- Ensemble size: $N = 32$ (only), glue regime: $n = 0.8$ (high coherence).
- Coherence horizon: $W_{\text{coh}} = 100$.
- Cloth extraction: hits1 (`min_bin_hits`=1), bins=20, x10 epoch (burn-in 5000 epochs; measurement 15000 epochs; total steps 20000).
- Logging: `trace_threads=true` (stride 1), `include_ledger=true` to support Gate–4 hop localisation.

Two independent runs were performed:

- **repA**: `gateA3_N32_hits1_x10_bins20_n0p8_repA.yml`, seeds.start=56791, seeds.count=5, output dir `outputs_cloth/gateA3_N32_hits1_x10_bins20_n0p8_repA`.
- **repB**: `gateA3_N32_hits1_x10_bins20_n0p8_repB.yml`, seeds.start=66791, seeds.count=5, output dir `outputs_cloth/gateA3_N32_hits1_x10_bins20_n0p8_repB`.

Procedure

From the repository root:

1. Execute two scans (repA and repB) with identical parameters except for the seed range and output directory.
2. Summarise run-level metrics to CSV: `bcqm_vii_cloth/analysis/summarise_runs.py` → `csv/repeatability/repA_run_summary.csv` and `csv/repeatability/repB_run_summary.csv`.
3. Run Gate–4 hop localisation on each run: `bcqm_vii_cloth/analysis/gate4_thread_localisation.py` (all/all partition and super-graph; resolution 1.0) → `gate4_repA_hopdist_seedwise.csv` and `gate4_repB_hopdist_seedwise.csv`.
4. Compare repA and repB using `repeatability_compare.py` → `repeatability_compare_table.csv` and `fig_repeatability_repA_vs_repB.pdf`.

Outputs

- `repeatability_compare_table.csv` (summary table of mean ± sd across the 5 seeds in each run).
- `fig_repeatability_repA_vs_repB.pdf` (visual comparison of headline metrics).

Results

Run-level stability (cloth and clock)

Metric	repA (mean ± sd)	repB (mean ± sd)	$\|\Delta\|$
Core fraction	0.1820 ± 0.0125	0.1845 ± 0.0090	0.002
Clock quality Q_clock	6.200 ± 0.144	6.229 ± 0.134	0.029
Core events	16,783 ± 1,169	17,035 ± 838	252
Halo events	75,408 ± 1,083	75,303 ± 791	105
Core edges	18,750 ± 1,500	19,049 ± 1,137	299
Halo edges	436,371 ± 1,496	436,067 ± 1,134	303
Ball component size	16,783 ± 1,168	17,034 ± 838	252

Table 1: Repeatability summary (run-level metrics) for two independent seed ranges at N=32, n=0.8 (hits1).

Interpretation:

- The key Stage–2 scalars are stable across the two independent runs: Φ differs by ≈ 0.0024, and Q_{clock} differs by ≈ 0.029, both far below their respective seed-to-seed scatter within each run.
- Core/halo event and edge counts shift slightly between repA and repB but remain within the run-to-run variability; no qualitative change in the core+halo regime is observed.

Gate–4 locality (hop distances on the community cloth)

Metric	repA (mean ± sd)	repB (mean ± sd)	$\|\Delta\|$
Hop fraction d=0	0.406 ± 0.189	0.521 ± 0.192	0.115
Hop fraction d=1	0.535 ± 0.163	0.440 ± 0.148	0.094
Hop fraction d=2	0.059 ± 0.039	0.038 ± 0.056	0.020
Hop fraction d3	0.000 ± 0.000	0.000 ± 0.000	0.000
Mean hop distance	0.653 ± 0.219	0.517 ± 0.241	0.135
Mean hop distance | change	1.090 ± 0.053	1.058 ± 0.084	0.033

Table 2: Repeatability summary (Gate-4 hop/localisation metrics) on the undirected community cloth (all/all partition and super-graph).

Interpretation:

- Locality is robust: the hop distribution is strictly confined to $d \in \{0, 1, 2\}$ in both runs (no $d \geq 3$ events).
- The mixture between $d = 0$ and $d = 1$ varies between the two runs at the present sample size, while maintaining strong locality. This suggests that Stage–2 wording should prefer "dominated by $d \in \{0, 1\}$" rather than "dominated by $d = 1$" unless further averaging (more seeds or longer logging) is performed.

Figure

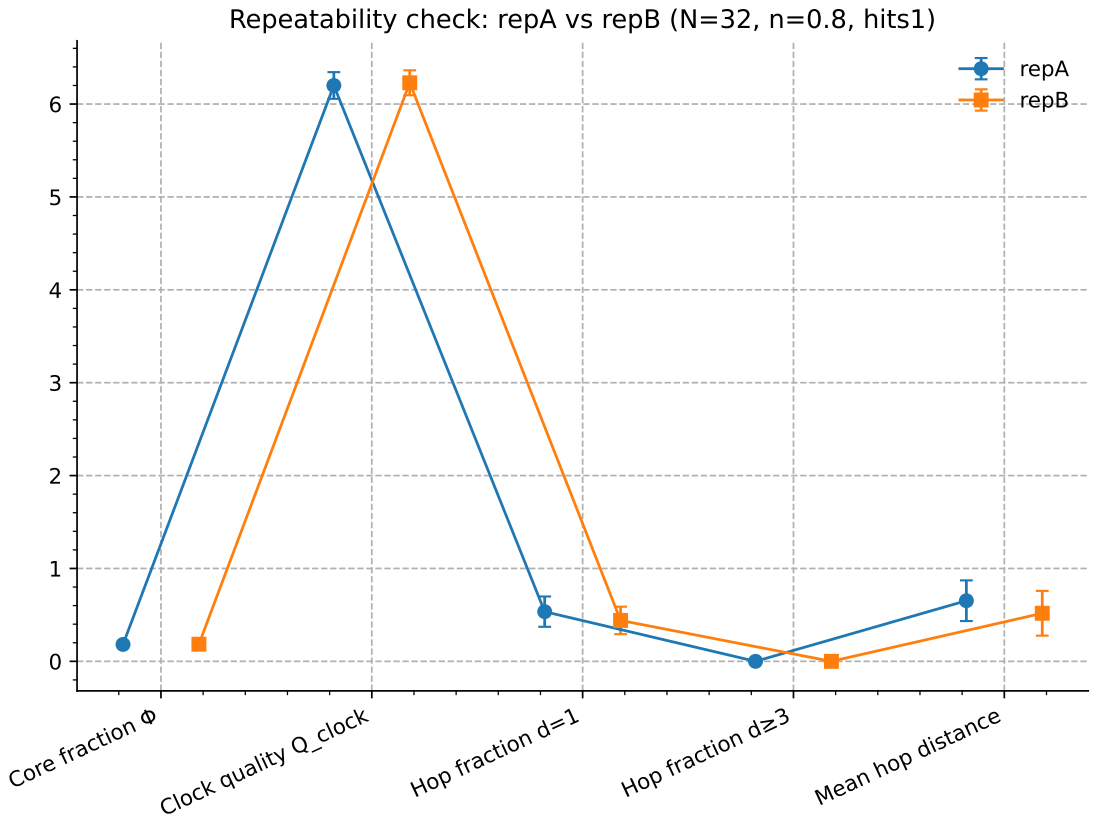

Figure 1: Repeatability check: repA versus repB for headline metrics at N=32, n=0.8 (hits1).

Conclusion

Two independent executions of the headline Stage–2 configuration (N=32, n=0.8, hits1) yield materially consistent mesoscopic outputs. The core/halo fraction Φ and clock quality Q_{clock} are repeatable across seed ranges, supporting the Stage–2 claim that the coarse cloth object and clock signal are robust to microscopic stochasticity. Gate–4 results confirm strict locality on the community cloth ($d \geq 3$ absent), with some variability in the d=0 versus d=1 mixture at the present sample size.