

# *Skeptical review: Dynamic Weighted Peptide Network Analysis for Characterizing and Predicting Aggregate Stability*

---

## Summary

The manuscript introduces a dynamic peptide-level interaction network to characterize aggregation in an MD simulation of 30 KYFIL pentapeptides (Sec. 2.1–2.4). Peptides are nodes; inter-peptide contacts (hydrophobic, aromatic, hydrogen-bond-like) define edges whose weights are built from contact counts combined with fixed coefficients (Sec. 2.2–2.3). Over 100–1435 ns, the authors compute time-resolved global and LCC-focused graph descriptors (e.g., density, connected components, LCC size, Laplacian/Fiedler value) for both weighted and binarized graphs and correlate these with LCC radius of gyration and a proposed packing score (Sec. 2.5, Sec. 3.1–3.4). A composite order parameter  $OP_{LCC}$  is proposed as a stability/transition descriptor (Sec. 2.6, Sec. 3.5–3.6). The overall workflow is well organized and potentially useful, but several key definitions (weights, “weighted density”, packing score,  $OP_{LCC}$ ), reproducibility details (MD protocol and PBC handling), and statistical treatments (time-series autocorrelation; confounding by LCC size) currently limit interpretability and support for the stronger “stability/prediction” claims.

## Strengths

- Clear motivation for using graph-theoretic descriptions of aggregation dynamics and a coherent narrative centered on the LCC (Introduction, Sec. 3.1–3.6).
- Well-structured pipeline from trajectory frames to weighted/binary adjacency matrices and Laplacian spectral analysis (Sec. 2.2–2.4, Sec. 2.8).
- Long time-resolved analysis window (100–1435 ns) with dense sampling that enables detailed time-series and distributional characterizations (Sec. 3.1–3.3).
- Systematic side-by-side comparison of weighted vs. binary metrics, which is helpful for assessing what extra information weights may provide (Sec. 3.3–3.4).
- The LCC-focused analysis is a sensible choice for aggregation problems and makes the metrics more physically interpretable than whole-system spectral metrics in a frequently disconnected system (Sec. 3.1–3.2).
- Figures generally provide a comprehensive view of dynamics across multiple metrics and are organized to support cross-comparisons (Sec. 3).

## Major issues

1. **MD simulation protocol and trajectory preprocessing are insufficiently specified for reproducibility and for assessing physical realism (Sec. 2.1).** Key metadata are missing (force field, water model, thermostat/barostat and targets, timestep/constraints, nonbonded cutoffs and PME settings, box size/PBC, ionic strength/countersions, preparation/equilibration). In addition, the analysis depends critically on how periodic boundary conditions are handled for distance-based contacts, connected components, and LCC radius of gyration, but PBC imaging/unwrapping/centering is not described (Sec. 2.1, Sec. 2.5.1).

*Recommendation:* Expand Sec. 2.1 (and Sec. 2.5.1 where relevant) to include a complete MD methods block: force field and water model; thermostat/barostat types and parameters; timestep and bond constraints; electrostatics method and cutoffs; box size and PBC; ions/ionic strength; how peptides were initially placed; equilibration steps and durations. Explicitly state the analyzed time range, the frame-saving interval/stride, and the number of frames used (the implied  $\sim 20$  ps/frame from 66,771 frames over 1335 ns should be stated). Add a clear description of PBC handling in analysis (e.g., minimum-image

distances, cluster-based unwrapping, recentering on the LCC) and confirm that LCC identification and  $R_g$  are robust to PBC artifacts. If relying on prior work for the trajectory, cite it but still summarize the essential parameters here.

- 2. Edge-weight construction (contact definitions + fixed coefficients) is central to the paper but is currently heuristic and not validated; key conclusions about the added value of weighted metrics depend on these choices (Sec. 2.2–2.3; Sec. 3.3–3.5).** The chosen coefficients ( $w_H = 1.0$ ,  $w_A = 1.5$ ,  $w_{HB} = 2.0$ ) and distance cutoffs are not calibrated, and weights based on raw contact counts can conflate interaction multiplicity with atom-list size and cutoff artifacts. Additionally, it is unclear whether contact types are mutually exclusive: aromatic contacts may be counted in addition to hydrophobic contacts for the same peptide pair, inflating weights for aromatic-rich interactions (Sec. 2.2.1–2.2.3).

*Recommendation:* In Sec. 2.2–2.3, (i) explicitly state whether a given atom pair/peptide pair can contribute to multiple contact types simultaneously (hydrophobic + aromatic + H-bond) and justify this choice; and (ii) provide a stronger rationale for cutoffs and coefficients with biomolecular/MD references where possible. Add a robustness/sensitivity analysis (Appendix or Sec. 3.3–3.5): test at least 2–3 alternative weighting schemes (e.g., equal weights; per-contact-type  $z$ -scoring; normalization by the number of eligible atom pairs; or a bounded transform such as  $w/(w + c)$ ) and modest cutoff variations (e.g., 4/5/6 Å; H-bond 3.2–3.8 Å). Recompute a minimal set of headline outputs (LCC Fiedler and “density” distributions; correlations with  $R_g$ /packing;  $OP_{LCC}$  time series) to identify which conclusions are robust vs. weight-definition-dependent. If computational cost is limiting, do this on a representative subset of frames and state the limitation.

- 3. “Weighted density” is not consistently defined and is not a bounded density, yet is compared conceptually to binary density and used multiplicatively in  $OP_{LCC}$  (Sec. 2.4.2; Sec. 3.1–3.5).** The manuscript mixes (a) language about normalization to a maximum possible sum and (b) formulas dividing  $\sum A_w$  by  $N_p(N_p - 1)$  despite also describing the number of undirected pairs as  $N_p(N_p - 1)/2$ . With weights  $> 1$  (from contact counts and coefficients), the quantity can exceed 1 and its magnitude becomes scale-dependent, complicating interpretation and comparisons.

*Recommendation:* Provide one precise definition in Sec. 2.4.2: clarify whether edges are undirected and counted once (sum over  $i < j$ ) or twice (sum over  $i \neq j$ ), and specify the exact denominator accordingly. Then either (i) rename the metric to something scale-appropriate (e.g., “mean edge weight per possible pair” or “average weighted adjacency”) and interpret it as such in Sec. 3; or (ii) define a properly normalized weighted density bounded in  $[0, 1]$  by specifying a maximum-per-pair weight (or using a bounded transform). Ensure the same definition is used for system and LCC versions and update  $OP_{LCC}$  accordingly. Include an edge-weight distribution (or typical per-pair weight ranges) in Sec. 3.1–3.2 or Supplementary material to make the magnitude interpretable.

- 4. Laplacian/Fiedler value reporting and interpretation need correction and normalization (Sec. 2.4.1; Sec. 3.1; Sec. 3.4).** (i) The manuscript states Laplacian eigenvalues are non-negative but reports a negative mean Fiedler value ( $\lambda_1 < 0$ ), indicating either numerical handling issues or a mismatch between the matrix used and the stated definition. (ii) Magnitude differences between weighted and binary Fiedler values are interpreted as “greater sensitivity,” but Laplacian eigenvalues scale linearly with uniform rescaling of weights; therefore larger values alone are not evidence of greater structural sensitivity. (iii) Whole-system Fiedler is largely uninformative if the graph is disconnected most of the time.

*Recommendation:* First, verify the Laplacian construction (symmetry, nonnegative weights, zero diagonal) and eigen-solver conventions; clip small negative eigenvalues to 0 within a stated tolerance and correct any text that implies true negativity (Sec. 2.4.1, Sec. 3.1). Second, to compare weighted vs. binary meaningfully, report scale-invariant quantities: e.g., coefficient of variation of  $\lambda_1$ , normalized fluctuations, or

use a normalized Laplacian (and clearly define it). Third, emphasize LCC-based spectral analysis (or per-component summaries) over system-wide Fiedler when disconnection is common, and explain why the chosen spectral quantity is physically informative for aggregation stability/fragmentation.

5. **The LCC packing score is inconsistent between Methods and Results and is not adequately justified/validated as a packing/compactness proxy (Sec. 2.5.2 vs Sec. 3.2.1 and captions).** Sec. 2.5.2 defines  $S_{\text{LCC}}/R_g^4$ , while Results and multiple captions use  $S_{\text{LCC}}/R_g^3$  with stated units of peptides/ $\text{\AA}^3$ . This affects reported values and correlations throughout Sec. 3.2–3.3 and any downstream interpretation of “packing.”

*Recommendation:* Resolve the exponent inconsistency and make the definition uniform everywhere (Sec. 2.5.2; Sec. 3.2–3.3; figure captions such as Figs. 7–10, 17, 19–21). Confirm which formula was actually used in computation and recompute/update correlations if needed. Add a brief physical rationale for the chosen exponent (likely  $R_g^3$  as a volume-like proxy). To strengthen validation, compare against at least one more conventional compactness/packing descriptor (e.g., contacts per peptide, SASA per peptide, convex-hull/alpha-shape volume density, coordination number) and discuss limitations of using a peptide-count-based density surrogate.

6. **Correlation analysis over MD time series is statistically overstated and likely confounded (Sec. 2.5; Sec. 3.3).** Pearson  $r$  and naive  $p$ -values computed over  $\sim 66\text{k}$  frames ignore strong temporal autocorrelation, yielding artificially tiny  $p$ -values. In addition, several relationships are likely driven by LCC size  $S_{\text{LCC}}$  (a confounder), because both graph metrics and geometric measures ( $R_g$ , packing surrogate) scale with size.

*Recommendation:* Update Sec. 2.5 and Sec. 3.3 to treat these as time-series: estimate autocorrelation times (or use block averaging) and compute an effective sample size; report 95% confidence intervals for  $r$  using block bootstrap (or equivalent) and adjust significance claims accordingly. Add at least one confounding control: partial correlations controlling for  $S_{\text{LCC}}$ , or size-stratified analyses (compute correlations within fixed  $S_{\text{LCC}}$  bins) to distinguish “compactness at fixed size” from “size-driven” correlations. Where possible, supplement Pearson with rank-based correlations (Spearman) to check robustness to nonlinearity/outliers.

7.  **$OP_{\text{LCC}}$  is not fully defined in a reproducible way and is not validated as a stability/transition or predictive indicator (Sec. 2.6; Sec. 3.5–3.6).** The text states components may be “normalized if necessary” but does not specify the implemented normalization. Multiplying three correlated, scale-dependent quantities (especially if using the current weighted “density”) risks domination by one factor and makes interpretation difficult. The manuscript also hints at prediction/stability, but evidence is primarily contemporaneous correlation and descriptive tracking.

*Recommendation:* In Sec. 2.6, state the exact implemented formula, including any normalization (e.g., min–max over the analyzed window,  $z$ -scores, or scaling to theoretical bounds) and justify the multiplicative form versus alternatives (e.g., weighted sum or log-sum). In Sec. 3.5–3.6, quantify redundancy (pairwise correlations among OP components) and show whether  $OP_{\text{LCC}}$  adds information beyond  $S_{\text{LCC}}$  alone (or beyond the best single metric) using, e.g., variance explained, mutual information, or simple regression comparisons. If retaining predictive language, operationally define “events” (e.g., fragmentation when  $S_{\text{LCC}}$  drops by  $\geq k$  within  $\Delta t$  or when  $N_{\text{c}}$  increases) and test lead–lag/early-warning behavior (time-lag correlations, ROC/AUC for event classification). Otherwise, reframe  $OP_{\text{LCC}}$  as a descriptive composite indicator and temper claims accordingly.

8. **The manuscript’s broader positioning (prediction/general stability claims and generality across systems) is stronger than what is demonstrated (Abstract; Introduction; Sec. 3.6; Sec. 4).** Results are shown for a single peptide sequence, one set of conditions, and apparently one trajectory; this limits generality, and without an event-based test, “prediction” is not established.

*Recommendation:* Align claims with evidence across Abstract/Introduction/Sec. 3.6/Sec. 4: emphasize characterization and proof-of-concept unless predictive tests are added. Explicitly acknowledge single-system/single-trajectory limitations. If feasible, add a minimal external check: a second trajectory (different initial configuration) or a nearby condition (concentration/temperature) and show that qualitative conclusions (weighted vs binary behavior, correlation signs,  $OP_{LCC}$  behavior) persist. If additional simulations are not feasible, present a concrete future-work plan identifying what must be tested to claim generality (other sequences, force fields, concentrations). Consider adjusting the title if it currently implies prediction.

9. **Related work and citations are not well aligned with biomolecular network analysis; the framing is skewed toward astronomy/cosmology graph applications, weakening novelty assessment and scholarly positioning (Sec. 1; Sec. 2.4; References).**

*Recommendation:* Revise Sec. 1 (and optionally add a short Related Work paragraph near Sec. 2.4) to cite and discuss relevant biomolecular literature: residue/contact networks in MD, hydrogen-bond networks, protein/peptide aggregation network analyses, and prior uses of spectral graph measures/community structure in biophysics. Then clearly state what is novel here (peptide-as-node granularity, multi-contact-type weighting, long time-resolved LCC spectral tracking). Update the References accordingly and ensure citations are domain-appropriate.

10. **Figure interpretability and cross-figure consistency issues reduce actionability: inconsistent definitions/notation across captions and text (e.g., packing score exponent; density definition;  $\lambda_1$  notation), unclear units/time labels in some places, missing plot metadata (bin widths, sample sizes), and overplotting/redundancy that obscures trends (Sec. 3; multiple figure captions).**

*Recommendation:* Standardize notation and units across all figures and text ( $R_g$  vs  $R_{g,LCC}$ ;  $\lambda_{1,LCC}$ ;  $S_{LCC}$ ; packing score definition; weighted vs binary density naming). Ensure each caption is self-contained: report analysis window, number of frames, bin widths/normalization for histograms, and any smoothing/subsampling. Reduce overplotting via transparency, density plots, or summary overlays (mean/median with IQR). Merge or move redundant plots to Supplementary material and add one panel/table summarizing key quantitative comparisons (means/SD/CV) between weighted and binary metrics (Sec. 3.4).

## Minor issues

1. Hydrogen-bond and aromatic contact definitions are purely distance-based and may overcount non-specific close approaches, especially in dense aggregates; aromatic interactions lack orientation criteria and H-bonds lack angular criteria (Sec. 2.2.2–2.2.3).

*Recommendation:* Add a short limitations paragraph in Sec. 2.2.2–2.2.3 stating what is (and is not) captured by these simplified definitions. If feasible, include a small robustness check with a standard H-bond angle criterion and/or a stricter distance cutoff, and report whether aggregate-level conclusions (LCC metrics/correlations) change qualitatively.

2. The binary graph definition depends on the union of all contact types via  $A_b(t) = 1$  if  $A_w(t) > 0$ , but this dependence is easy to miss and complicates interpretation of “binary contacts” (Sec. 2.3).

*Recommendation:* Make explicit in Sec. 2.3 that the binary adjacency represents the existence of any of the defined interaction types (hydrophobic/aromatic/H-bond-like) rather than a single peptide–peptide distance criterion. Consider reporting per-type binary graphs in Supplementary material to help interpret which interaction class drives topology changes.

3. The rationale for discarding early frames ( $< 100$  ns) is only briefly stated and not demonstrated (Sec. 2.1; Sec. 3.1).

*Recommendation:* In Sec. 2.1 (or Supplementary material), show a short plot of  $S_{LCC}$  and/or  $N_{cc}$  over the full trajectory including 0–100 ns and briefly justify why 100 ns is an appropriate cutoff (equilibration vs early assembly transient).

4. Several additional metrics appear in figures/text (average degree, clustering coefficient, betweenness) without clear definitions or a clear role in the narrative (e.g., Fig. 1 bottom; references to Fig. 18 in Sec. 3.1/Sec. 3.4).

*Recommendation:* Either define these metrics concisely in Sec. 2.4 and add brief interpretation in Sec. 3, or move them to Supplementary material and clearly state they are exploratory/secondary.

5. Comparison claims such as “weighted metrics are more sensitive” are mostly qualitative (Sec. 2.8; Sec. 3.4) and can be confounded by scaling differences.

*Recommendation:* In Sec. 3.4, add scale-aware quantitative comparisons: e.g., coefficient of variation, normalized RMS fluctuations, or within-size-bin fluctuations for weighted vs binary metrics (especially for  $\lambda_1$  and density/mean-edge-weight).

6. Matrix conventions are not fully explicit (e.g., whether  $A_{w,ii} = 0$  and whether degree sums exclude the diagonal), which matters for strict reproducibility (Sec. 2.3; Sec. 2.4.1).

*Recommendation:* State explicitly that  $A_{w,ii}(t) = 0$  (and  $A_{b,ii}(t) = 0$ ) and write degree definitions with  $j \neq i$  (or state the diagonal is zero so including it is harmless).

## Very minor issues

1. Minor typographical and formatting inconsistencies (e.g., broken word “frame\nwork” in the Introduction; stray Markdown-like “#” in headings; inconsistent math spacing and variable formatting) (Sec. 1; Sec. 2.5.2; Sec. 3).

*Recommendation:* Proofread and standardize headings and math typography throughout; fix line-break artifacts and remove stray formatting characters.

2. Figure referencing/caption micro-inconsistencies (panel naming, occasional out-of-order references, slight notation drift) (Sec. 3.2–3.4).

*Recommendation:* Ensure figures are cited in numerical order, unify “top/bottom” (or “left/right”) panel references, and harmonize variable names/units across captions.

3. References formatting is inconsistent and the bibliography contains stylistic artifacts (e.g., stray leading dashes; mixed DOI/arXiv presentation) (References).

*Recommendation:* Clean the References to a single journal style and ensure added biomolecular-network citations follow the same format.

4. Accessibility/print-readiness: some plots appear dense with small fonts and potentially non-colorblind-safe palettes (Sec. 3 figures).

*Recommendation:* Increase font sizes/line weights, consider vector exports, and use colorblind-safe palettes; reduce visual clutter with transparency or summary overlays.

## Key statements and references

- **Aromatic interactions indicative of  $\pi$ – $\pi$  stacking between Phenylalanine (PHE) and Tyrosine (TYR) side chains were operationally defined using a 5.0 Å heavy-atom distance cutoff between aromatic-ring atoms, following prior work on  $\pi$  and hydrophobic interactions in biomolecular phase behavior and large molecular assemblies, rather than more detailed centroid/normal-vector criteria adopted elsewhere for  $\pi$ – $\pi$  stacking analysis.**

- *Reference(s)*: Das, S., Lin, Y.-H., Vernon, R. M., Forman-Kay, J. D., & Chan, H. S. 2020, Bensberg, M., Eckhoff, M., Husstein, R. T., et al. 2025
- • The dynamic weighted adjacency matrix  $A_w(t)$  of size  $N_p \times N_p$  with  $N_p = 30$  peptides was constructed for each simulation frame by defining edge weights as a linear combination of contact counts of different types,  $A_w(t)_{ij} = w_H C_H(t) + w_A C_A(t)_{ij} + w$ , using a graph-based interaction-weighting strategy previously applied in network analyses of correlated systems and epidemics on networks.  $C_{HB}(t)_{ij}$
- *Reference(s)*: Silva, D. H., Ferreira, S. C., Cota, W., Pastor-Satorras, R., & Castellano, C. 2019, Jagvaral, Y., Lanasse, F., Singh, S., et al. 2022
- • Binary adjacency matrices  $A_b(t)$  were derived from the weighted matrices by setting  $A_b(t)_{ij} = 1$  if  $A_w(t)_{ij} > 0$  and 0 otherwise, adopting a binary-contact graph construction analogous to prior recurrence- and graph-based classification approaches used for complex dynamical systems and close binary stars.
- *Reference(s)*: Shrivastava, A., & Li, P. 2014, George, S. V., Misra, R., & Ambika, G. 2019
- • For each frame, the weighted graph Laplacian  $L_w(t) = D_w(t) - A_w(t)$  was computed, where  $D_w(t)$  is the diagonal weighted degree matrix with entries  $D_w(t)_{ii} = \sum_j A_w(t)_{ij}$ , and the spectrum of  $L_w(t)$  was analyzed to obtain the Fiedler value  $\lambda_1$  as a measure of algebraic connectivity and fragmentation propensity, following established spectral graph-theoretic methodologies for large graphs and weighted Laplacian problems.  $A_w(t)_{ij}$
- *Reference(s)*: Granzio, D., Ru, R., Zohren, S., et al. 2019, Druskin, V., Mamonov, A. V., & Zaslavsky, M. 2021, Drábek, P., Ho, K., & Sarkar, A. 2018
- • Time series of weighted and binary adjacency matrices, graph-theoretical metrics (including Fiedler values, densities, and connected components), and associated physical properties were stored in HDF5 format and analyzed using a pipeline inspired by prior work that applied minimum-spanning-tree and recurrence-based graph methods to astrophysical time series and filamentary structures, leveraging these graph constructs as unsupervised descriptors of complex systems.
- *Reference(s)*: García, C. R., Torres, D. F., Zhu-Ge, J.-M., & Zhang, B. 2024, Gilpin, W. 2024, Strey, S.-G., Castronovo, A., & Elumalai, K. 2024

## Mathematical consistency audit

This section audits **symbolic/analytic** mathematical consistency (algebra, derivations, dimensional/unit checks, definition consistency).

**Maths relevance:** substantial

The paper’s analytic content centers on constructing time-dependent weighted and binary graphs from contact counts, defining standard Laplacian/spectral quantities (Fiedler value), global connectivity metrics (connected components, LCC size), a weighted-density-like metric, and composite/physical derived quantities (packing score, combined order parameter). Most definitions are standard, but there are internal inconsistencies in the definition/normalization of weighted density and an explicit contradiction between the claimed non-negativity of Laplacian eigenvalues and the reported negative Fiedler values.

**Checked items**

1. ✓ **Weighted adjacency from contact counts** (Sec. 2.3, p.3:  $A_w(t)_{ij} = w_H C_H(t) + w_A C_A(t)_{ij} + w$ )  $C_{HB}(t)_{ij}$

- **Claim:** Defines the weighted interaction strength between peptides  $i$  and  $j$  as a linear combination of contact-type counts with fixed coefficients.
  - **Checks:** algebra, symbol/definition consistency
  - **Verdict:** PASS; confidence: high; impact: moderate
  - **Assumptions/inputs:** Contact counts  $C_H, C_A, C_{HB}$  are nonnegative integers., Weights  $w_H, w_A, w_{HB}$  are fixed scalars., Contacts are symmetric between  $i$  and  $j$ .
  - **Notes:** The linear combination is well-formed and produces nonnegative weights if all  $w$ 's are positive, consistent with later use of  $A_w(t)_{ij} > 0$  as a connection criterion.
2. ✓ **Symmetry of weighted adjacency** (Sec. 2.3, p.3: 'The matrix  $A_w(t)$  is symmetric ( $A_w(t)_{ij} = A_w(t)_{ji}$ )')
- **Claim:** Weighted adjacency is symmetric because interactions are reciprocal.
  - **Checks:** definition consistency
  - **Verdict:** PASS; confidence: medium; impact: minor
  - **Assumptions/inputs:** Contact counting procedure is symmetric under  $i \leftrightarrow j$  (same distance criterion both directions).
  - **Notes:** True if the implemented counting is symmetric; the paper asserts reciprocity but does not show the counting algorithm. Analytically plausible given distance-based contacts.
3. ✓ **Binary adjacency induced by weighted adjacency** (Sec. 2.3, p.3: ' $A_b(t)_{ij} = \text{if } A_w(t)_{ij} > 0 \text{ otherwise}$ ')  $> 0$ , and  $A_b(t)_{ij}$
- **Claim:** Defines a binary graph by thresholding the weighted adjacency at zero.
  - **Checks:** definition consistency, logical implication
  - **Verdict:** PASS; confidence: high; impact: moderate
  - **Assumptions/inputs:**  $A_w(t)_{ij}$  is real-valued and comparable to 0.
  - **Notes:** Consistent with  $A_w$  being nonnegative from contact counts;  $A_b$  captures existence of any (weighted) contact.
4. ✓ **Weighted Laplacian construction** (Sec. 2.4.1, p.3-4: ' $L_w(t) = D_w(t) - A_w(t)$ ')
- **Claim:** Defines the weighted graph Laplacian as degree matrix minus adjacency matrix.
  - **Checks:** algebra, symbol/definition consistency
  - **Verdict:** PASS; confidence: high; impact: critical
  - **Assumptions/inputs:**  $D_w(t)$  is diagonal of weighted degrees.,  $A_w(t)$  is symmetric.
  - **Notes:** Standard definition; later spectral claims rely on this.
5. △ **Weighted degree definition (diagonal of  $D_w$ )** (Sec. 2.4.1, p.4: ' $D_w(t)_{ii} = \sum_j A_w(t)_{ij}$ ')  $A_w(t)_{ij}$
- **Claim:** Defines the weighted degree of node  $i$  as the sum of incident edge weights.
  - **Checks:** definition consistency, notation consistency
  - **Verdict:** UNCERTAIN; confidence: medium; impact: minor
  - **Assumptions/inputs:**  $A_w(t)_{ii} = 0$  (typical; not stated).
  - **Notes:** The formula includes  $j = i$ . This is fine only if  $A_w(t)_{ii}$  is guaranteed to be 0; the paper does not explicitly define the diagonal of  $A_w(t)$ . If  $A_w(t)_{ii} \neq 0$ , this would introduce self-loops and alter  $L_w$  and  $\lambda_1$  interpretation.
6. ✘ **Eigenvalue ordering and non-negativity claim** (Sec. 2.4.1, p.4: 'The eigenvalues are non-negative and can be ordered as  $0 = \lambda_0 \leq \lambda_1 \leq \dots$ ')
- **Claim:** States Laplacian eigenvalues are nonnegative and ordered with  $\lambda_0 = 0$ .
  - **Checks:** internal logical consistency, cross-check with later reported values

- **Verdict:** FAIL; confidence: high; impact: critical
  - **Assumptions/inputs:**  $L_w(t)$  is a (standard) Laplacian built from a symmetric nonnegative adjacency.
  - **Notes:** This statement conflicts with Results Sec. 3.1 (p.6) reporting a negative mean weighted system Fiedler value (mean  $-5.205 \times 10^{-6}$ ). Under the stated construction,  $\lambda_1$  should be  $\geq 0$ ; negative reporting requires clarification (numerical tolerance vs different matrix).
7. ✓ **Fiedler value interpretation for connected/disconnected graphs** (Sec. 2.4.1, p.4: 'For a connected graph,  $\lambda_1 > 0$  ... For a disconnected graph,  $\lambda_1 = 0$ ')
- **Claim:** Uses  $\lambda_1$  as algebraic connectivity with standard connectedness implications.
  - **Checks:** logical implication
  - **Verdict:** PASS; confidence: high; impact: moderate
  - **Assumptions/inputs:** Standard Laplacian definition is used.
  - **Notes:** These implications match the stated Laplacian framework.
8. ✗ **Weighted density definition (formula vs conceptual normalization)** (Sec. 2.4.2, p.4: bullet 'Weighted Graph Density')
- **Claim:** Defines weighted density as (i) sum of edge weights divided by the maximum possible sum if every pair had weight 1.0, and (ii) 'more simply' sum of off-diagonal entries of  $A_w$  divided by  $N_p(N_p - 1)$ .
  - **Checks:** algebra, normalization/constraints, definition consistency
  - **Verdict:** FAIL; confidence: high; impact: critical
  - **Assumptions/inputs:**  $A_w$  is symmetric; diagonal excluded.
  - **Notes:** The two descriptions are not consistent when weights can exceed 1 due to multiple contacts ( $A_w$  is a weighted sum of counts). If the maximum edge weight is 1.0, the density should be  $\leq 1$ ; however later reported weighted densities exceed 1 (system  $\approx 1.234$ ; LCC  $\approx 8.039$ ). Also, for undirected graphs the phrase 'maximum possible sum' depends on whether edges are counted once or twice; the stated denominator  $N_p(N_p - 1)$  corresponds to counting ordered pairs, while many later phrases reference unordered pairs.
9. △ **Binary density interpretation** (Sec. 3.2.2, p.7: 'binary density ... ratio of existing edges to possible edges')
- **Claim:** Binary density is a standard bounded density in  $[0, 1]$ .
  - **Checks:** definition consistency, normalization/constraints
  - **Verdict:** UNCERTAIN; confidence: medium; impact: moderate
  - **Assumptions/inputs:** Undirected simple graph on  $S_{\text{icc}}$  nodes., No self-loops.
  - **Notes:** Earlier (Sec. 2.4.2) the 'more simply' formula divides by  $N_p(N_p - 1)$ , which would equal  $2E/(n(n - 1))$  for undirected graphs (still bounded by 1). But later text alternates between possible pairs  $n(n - 1)/2$  and possible ordered pairs  $n(n - 1)$ , leaving an unresolved factor-of-2 ambiguity in what was actually computed.
10. ✓ **Connected components equivalence for weighted vs binary graphs** (Sec. 3.1, p.5: 'connectivity definition used for component analysis is binary (an edge weight  $> 0$  constitutes a connection), thus these metrics are identical')
- **Claim:**  $N_{\text{cc}}$  and  $S_{\text{icc}}$  are identical for weighted and binary representations.
  - **Checks:** logical implication, definition consistency
  - **Verdict:** PASS; confidence: high; impact: minor
  - **Assumptions/inputs:** Connected components in the weighted graph are computed using threshold  $A_w > 0$ , not by weighted notions of connectivity.

- **Notes:** Given the stated threshold rule,  $N_{cc}$  and  $S_{lcc}$  depend only on whether  $A_w > 0$ , matching  $A_b$ .
11. ✓ **Packing score definition and units** (Sec. 2.5.2, p.4 and repeated in Results p.6-8: 'Packing Score =  $S_{lcc}/R_g^3$ ')
  - **Claim:** Defines a packing score as peptide count divided by cubic radius of gyration.
  - **Checks:** dimensional/units consistency, notation consistency
  - **Verdict:** PASS; confidence: high; impact: minor
  - **Assumptions/inputs:**  $S_{lcc}$  is a dimensionless integer count.,  $R_g$  has units of length (Å).
  - **Notes:** Units follow directly: peptides/Å<sup>3</sup>. No algebraic issues.
12. △ **Combined order parameter definition completeness** (Sec. 2.6, p.4: ' $OP_{LCC}(t) = S_{lcc}(t) \times \lambda_{1,LCC}(t) \times \text{Density}_{LCC}(t)$ ')
  - **Claim:** Proposes an order parameter as a product of LCC size, Fiedler value, and density, with optional normalization.
  - **Checks:** algebra, definition completeness
  - **Verdict:** UNCERTAIN; confidence: medium; impact: moderate
  - **Assumptions/inputs:** All three factors are computed on the same LCC subgraph at time  $t$ .
  - **Notes:** The product is algebraically fine, but 'Components were normalized if necessary' is not specified. Without an explicit normalization, the proposed  $OP$  is not uniquely defined/reproducible from the text alone, and inherits the density-definition ambiguity.
13. ✗ **System pair-count statement vs density formula** (Sec. 3.1, p.6: 'system-wide densities... across all possible peptide pairs  $(N_p(N_p - 1)/2)$ ' contrasted with Sec. 2.4.2 density denominator  $N_p(N_p - 1)$ )
  - **Claim:** Describes normalization in terms of possible unordered pairs  $n(n - 1)/2$ .
  - **Checks:** symbol/definition consistency
  - **Verdict:** FAIL; confidence: high; impact: moderate
  - **Assumptions/inputs:** Undirected interactions between distinct peptides.
  - **Notes:** There is an internal factor-of-2 inconsistency between the described count of possible pairs  $n(n - 1)/2$  and the earlier stated normalization by  $n(n - 1)$ . One of these descriptions (or the implemented calculation) must be corrected to avoid mis-scaled density values.

## Limitations

- The PDF text provides few explicit equation numbers; locations are referenced by section/page and quoted formulas instead.
- Several computations are described procedurally (e.g., contact counting) without explicit mathematical definitions sufficient to fully verify symmetry, diagonal conventions, or exact normalization choices.
- No step-by-step derivations are provided for correlations/statistical measures; this audit therefore focuses only on the analytic correctness and consistency of the definitions used.

## Numerical results audit

This section audits **numerical/empirical** consistency: reported metrics, experimental design, baseline comparisons, statistical evidence, leakage risks, and reproducibility.

All automated numeric checks passed. One theoretically motivated flag remains: the weighted system Fiedler value mean is slightly negative. Several other checks were sanity/consistency checks (rates, bounds, and scale checks) and support internal coherence of the reported summary statistics, but do not replace recomputation from raw per-frame data.

## Checked items

1. ✓ **C1\_time\_window\_duration** (Results §3 (page 5): '1335.42 ns simulation window, spanning from 100 ns ... end of the 1435.42 ns trajectory')
  - **Claim:** The analyzed window duration (1335.42 ns) should equal 1435.42 ns – 100 ns.
  - **Checks:** difference\_equals\_reported\_duration
  - **Verdict:** PASS
  - **Notes:** Computed  $1435.42 - 100.0 = 1335.42$ ; matches reported duration exactly.
2. ✓ **C2\_frames\_per\_ns** (Results §3 (page 5): '1335.42 ns ... comprising 66771 frames')
  - **Claim:** Implied sampling rate: frames per ns equals  $66771/1335.42$ .
  - **Checks:** rate\_computation
  - **Verdict:** PASS
  - **Notes:** Implied frames\_per\_ns= 50.0 and ns\_per\_frame= 0.02; no target rate was stated to compare against.
3. ✓ **C3\_binary\_density\_possible\_edges** (Methods §2.4.2 (page 4) and Results §3.1 (page 6):  $N_p = 30$  and 'binary system density averaged 0.07402')
  - **Claim:** For  $N_p = 30$ , possible undirected edges are  $N_p(N_p - 1)/2 = 435$ ; expected mean number of binary edges present  $\approx \text{density} \times 435$ .
  - **Checks:** derived\_count\_from\_density
  - **Verdict:** PASS
  - **Notes:** Possible edges computed as 435; implied mean edges= 32.1987, which is within  $[0, 435]$ .
4. ✓ **C4\_weighted\_density\_possible\_edges\_normalization** (Methods §2.4.2 (page 4): density defined as  $\text{sum}(A_w \text{ offdiag})/N_p(N_p - 1)$ ; Results §3.1 (page 6): 'weighted system density averaged 1.234')
  - **Claim:** If weighted density is defined as sum of off-diagonal weights divided by  $N_p(N_p - 1)$ , then implied mean off-diagonal sum is  $\text{density} \times N_p(N_p - 1)$ .
  - **Checks:** derived\_total\_weight\_from\_density
  - **Verdict:** PASS
  - **Notes:** Computed denom= 870; implied sum of off-diagonal weights= 1073.58; no raw target available for verification.
5. ✓ **C5\_lcc\_density\_binary\_possible\_edges** (Results §3.2.2 (page 7): 'binary LCC density averaged 0.4959')
  - **Claim:** Binary LCC density must be in  $[0, 1]$ ; 0.4959 satisfies. Also compute implied mean fraction of possible edges within LCC.
  - **Checks:** range\_check\_and\_interpretation
  - **Verdict:** PASS
  - **Notes:** Bound check passed:  $0 \leq 0.4959 \leq 1$ .
6. ✓ **C6\_lcc\_density\_weighted\_nonnegative** (Results §3.2.2 (page 7): 'weighted LCC density averaged 8.039')
  - **Claim:** Weighted density should be nonnegative;  $8.039 \geq 0$ . Also compare magnitude ratio vs binary LCC density for reported means.
  - **Checks:** nonnegativity\_and\_ratio
  - **Verdict:** PASS
  - **Notes:** Nonnegativity passed; computed ratio weighted/binary= 16.210929622907845.
7. ✓ **C7\_weighted\_system\_fiedler\_mean\_sd\_sign** (Results §3.1 (page 6): 'weighted system Fiedler value ... (mean  $-5.205 \times 10^{-6}$ , standard deviation  $2.82 \times 10^{-6}$ )')

- **Claim:** For a Laplacian, eigenvalues (including Fiedler value) are non-negative; a negative mean suggests numerical sign/rounding or definition inconsistency.
  - **Checks:** theoretical\_nonnegativity\_flag
  - **Verdict:** PASS
  - **Notes:** Mean is negative but within the stated absolute tolerance; computed  $z = \text{mean}/\text{sd} = -1.8457446808510638$ .
8. ✓ **C8\_ncc\_range\_contains\_mean\_sd** (Results §3.1 (page 5): 'average  $N_{\text{cc}}$  of 15.2 (standard deviation 2.28), ranging from 8 to 24')
- **Claim:** Check whether  $\text{mean} \pm 3\text{sd}$  falls within (or near) the stated min/max range; also ensure  $\text{min} \leq \text{mean} \leq \text{max}$ .
  - **Checks:** range\_vs\_mean\_sd
  - **Verdict:** PASS
  - **Notes:** Ordering holds ( $8 \leq 15.2 \leq 24$ ). Heuristic  $3\sigma$  interval: [8.36, 22.04], within reported range.
9. ✓ **C9\_slcc\_range\_contains\_mean\_sd** (Results §3.1 (page 5): ' $S_{\text{lcc}}$  ... averaged 10.67 (standard deviation 3.076), varying between 2 and 19 ... out of total 30')
- **Claim:** Check  $\text{min} \leq \text{mean} \leq \text{max}$ ;  $\text{max} \leq \text{total peptides}$ ; and  $\text{mean} \pm 3\text{sd}$  vs stated range (heuristic).
  - **Checks:** range\_vs\_mean\_sd\_and\_total\_bound
  - **Verdict:** PASS
  - **Notes:** Core bounds pass ( $2 \leq 10.67 \leq 19 \leq 30$ ). Heuristic  $3\sigma$  interval: [1.442, 19.898], not fully within [2, 19] (heuristic only).
10. ✓ **C10\_rg\_range\_contains\_mean\_sd** (Results §3.2.1 (page 6): ' $R_g$  ... averaging 13.28 Å (standard deviation 1.539 Å), with values ranging from 7.466 Å to 21.21 Å')
- **Claim:** Check  $\text{min} \leq \text{mean} \leq \text{max}$  and  $\text{mean} \pm 3\text{sd}$  vs min/max (heuristic).
  - **Checks:** range\_vs\_mean\_sd
  - **Verdict:** PASS
  - **Notes:** Ordering holds ( $7.466 \leq 13.28 \leq 21.21$ ). Heuristic  $3\sigma$  interval: [8.663, 17.897], within reported range.
11. ✓ **C11\_packing\_score\_range\_contains\_mean\_sd** (Results §3.2.1 (page 6): 'packing score averaged 0.004531 ... (standard deviation 0.0009041) ... ranging from 0.000955 to 0.007483')
- **Claim:** Check  $\text{min} \leq \text{mean} \leq \text{max}$  and  $\text{mean} \pm 3\text{sd}$  vs min/max (heuristic).
  - **Checks:** range\_vs\_mean\_sd
  - **Verdict:** PASS
  - **Notes:** Ordering holds ( $0.000955 \leq 0.004531 \leq 0.007483$ ). Heuristic  $3\sigma$  interval: [0.0018187, 0.0072433], within reported range.
12. ✓ **C12\_order\_parameter\_weighted\_mean\_vs\_components\_product** (Methods §2.6 (page 4) defines  $OP_{\text{LCC}} = S_{\text{lcc}} \times \lambda_1 \times \text{Density}$ ; Results §3.5 (page 11): weighted OP mean 975.2; component means:  $S_{\text{lcc}}$  mean 10.67 (page 5), weighted LCC Fiedler mean 11.37 (page 7), weighted LCC density mean 8.039 (page 7))
- **Claim:** If OP is a per-frame product,  $\text{mean}(\text{OP})$  is not generally equal to product of means, but product-of-means provides a quick sanity scale check against reported mean 975.2.
  - **Checks:** order\_of\_magnitude\_sanity\_product\_of\_means
  - **Verdict:** PASS

- **Notes:** Product-of-means= 975.2745980999999; ratio to reported= 1.0000764951804757. This is a scale/sanity check (mean(product) need not equal product(means)).
13. ✓ **C13\_order\_parameter\_binary\_mean\_vs\_components\_product** (Methods §2.6 (page 4) defines  $OP_{LCC} = S_{lcc} \times \lambda_1 \times \text{Density}$ ; Results §3.5 (page 11): binary OP mean 6.619; component means:  $S_{lcc}$  mean 10.67 (page 5), binary LCC Fiedler mean 1.262 (page 7), binary LCC density mean 0.4959 (page 7))
- **Claim:** Binary OP product-of-means provides a quick scale check against reported mean 6.619.
  - **Checks:** order\_of\_magnitude\_sanity\_product\_of\_means
  - **Verdict:** PASS
  - **Notes:** Product-of-means= 6.677561286; ratio to reported= 1.0088474521831092. This is a scale/sanity check (mean(product) need not equal product(means)).
14. ✓ **C14\_correlation\_coefficients\_in\_range** (Results §3.3 (pages 8-9): multiple Pearson  $r$  values reported (e.g., -0.556, -0.594, -0.697, -0.858, 0.788, 0.631, 0.687, 0.229, 0.376, 0.144))
- **Claim:** All reported Pearson correlation coefficients must lie within  $[-1, 1]$ .
  - **Checks:** bounds\_check\_multiple\_values
  - **Verdict:** PASS
  - **Notes:** All listed  $r$  values are within  $[-1, 1]$ .

#### Limitations

- Only parsed PDF text was available; no underlying datasets (trajectory, adjacency matrices, per-frame metrics) were provided, preventing recomputation of most reported means/SDs and all correlations.
- Figures are present but numeric extraction from plots/pixels was not used per instruction; only explicit numbers in the text were considered.
- Some checks (e.g., product-of-means vs mean-of-product for order parameters) are sanity/scale checks and cannot conclusively validate reported statistics without raw data.