Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Higher-Order Hodge Spectral Signatures of Local Nucleation in Dynamical Network Biomarkers

Satoshi Kanno

satoshi.kanno06@g.softbank.co.jp

SoftBank Corp. https://orcid.org/0000-0002-0885-1884

Yoshi-aki Shimada

SoftBank Corp. https://orcid.org/0000-0002-4483-1901

Kazuyuki Aihara

International Research Center for Neurointelligence, The University of Tokyo Institutes for Advanced Study, The University of Tokyo https://orcid.org/0000-0002-4602-9816

Article

Keywords:

Posted Date: October 15th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11300474/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Higher-Order Hodge Spectral Signatures of Local Nucleation in Dynamical Network Biomarkers

Satoshi Kanno,[1,*] Yoshi-aki Shimada,[1] and Kazuyuki Aihara[2]

[1] *Quantum Information Technology Department, Quantum Technology Division, Product Research and Development Division, SoftBank Corp., 1-7-1 Kaigan, Minato-ku, Tokyo 105-7529, Japan*

[2] *International Research Center for Neurointelligence, The University of Tokyo Institutes for Advanced Study, The University of Tokyo, Tokyo, Japan*

(Dated: October 8, 2026)

Complex systems can shift abruptly when they approach tipping points, but the earliest warning signs may emerge only in small parts of the system. Here we show theoretically that Dynamical Network Biomarkers (DNBs) can form through local nucleation: when the critical mode is heterogeneous, a subset of variables becomes strongly correlated first and then expands into the full DNB region. To detect this hidden precursor without searching over all possible subsets, we map correlation networks to clique complexes and analyze higher-order Hodge-Laplacian spectra. In addition, gene-expression data from TSOD and TSNO mice, such a fingerprint appears in TSOD mice before the conventional DNB signal peaks. These findings reveal a local stage of DNB formation that may provide earlier warning of critical transitions.

I. INTRODUCTION

Complex dynamical systems can undergo abrupt state transitions as they approach tipping points, and early detection of such transitions is important in a wide range of phenomena, including disease onset [1–9]. Dynamical Network Biomarker (DNB) theory[1–9], or Dynamical Network Marker (DNM) theory[5, 8, 10–14] provides a framework for de-

* satoshi.kanno06@g.softbank.co.jp

tecting pre-critical states without reconstructing detailed dynamical models by exploiting characteristic changes that arise near a critical transition: increased fluctuations and internal correlations within a specific group of variables, together with a relative decrease in correlations with variables outside the group [1–9].

However, more localized structural changes may arise before the conventional DNB signal becomes pronounced. When a single critical mode is dominant and its components are heterogeneous, variables with large contributions to the critical mode can become strongly correlated first, forming a small local structure that subsequently grows into a broader DNB region. We refer to such an early strongly correlated structure as a *local nucleus*, and view DNB formation as a process that develops from local nucleation toward a global structure.

Because a local nucleus may remain embedded in a large background network, detecting it using conventional graph statistics or by searching over unknown subsets is difficult. To address this problem, we lift the correlation network to a clique complex and use the spectra of higher-order Hodge Laplacians to detect localized higher-order structures [15–20]. In a clique complex, higher-dimensional simplices are formed only when multiple edges coexist coherently, allowing internally coherent structures to be emphasized while suppressing accidental background connections. Moreover, coherent local structures generate characteristic high-multiplicity features in the Hodge spectrum, providing a potential means of detecting a local nucleus without specifying its location in advance.

In this study, we first show theoretically that heterogeneity of the critical mode can lead to the earlier formation of a local nucleus, and then develop a framework for detecting such structures through higher-order Hodge spectra. We further demonstrate the local-to-global formation process using an analytically tractable stochastic dynamical model and apply the framework to longitudinal gene-expression data from TSOD and TSNO mice [21–23]. In TSOD mice, higher-order spectral structures consistent with a local nucleus were observed before the time point at which the conventional DNB indicator became most pronounced. These results suggest that higher-order spectral analysis can complement conventional DNB analysis by capturing an earlier local stage of DNB formation.

II. RESULTS

A. Dynamical Network Biomarkers and the Critical Mode

Consider a nonlinear dynamical system described by

$$\frac{dx}{dt} = f(x, \mu), \qquad x(t) = (x_1(t), \ldots, x_n(t))^\top,$$

where μ is a bifurcation parameter. Let $x^*(\mu)$ be a stable equilibrium and define the fluctuation around it as $\eta(t) = x(t) - x^*(\mu)$. Linearization around the equilibrium gives

$$\frac{d\eta}{dt} = J(\mu)\eta, \qquad J(\mu) = \left.\frac{\partial f}{\partial x}\right|_{x=x^*(\mu)}.$$

At a stable equilibrium, all eigenvalues of $J(\mu)$ have negative real parts. As μ changes, a critical transition can occur when the real part of a dominant eigenvalue, denoted by $\lambda_1(\mu)$, approaches zero from a negative value. In this regime, the restoring force along the corresponding eigendirection becomes weak, and fluctuations along this direction are strongly amplified. We refer to this dominant eigendirection as the *critical mode*. Detecting the approach to such a loss of stability is important because it provides a means of anticipating a large change in the state of the system through bifucation. In many applications, however, the detailed form of f, the bifurcation parameter μ, and the Jacobian matrix $J(\mu)$ are not directly available. Dynamical Network Biomarker (DNB) theory[1–9] is designed to detect the approach to the critical transition without requiring explicit knowledge of these quantities.

To describe fluctuations near the stable equilibrium, we introduce the stochastic linearized system

$$d\eta_t = J(\mu)\eta_t\, dt + \Sigma\, dW_t,$$

where W_t is a Wiener process. Suppose that only one eigenvalue $\lambda_1(\mu)$ of $J(\mu)$ becomes soft, with $\lambda_1(\mu) \to 0^-$, while all other modes remain uniformly stable. Then the covariance matrix can be written near the tipping point as

$$C(\mu) = \mathbb{E}[\eta_t \eta_t^\top] = \alpha(\mu) v v^\top + R(\mu),$$

where v is the critical mode, $R(\mu)$ is a bounded contribution from noncritical modes, and $\alpha(\mu) \to \infty$ as $\lambda_1(\mu) \to 0^-$, so that the contribution associated with the critical eigenvalue

becomes dominant as $\lambda_1 \to 0^-$. The general single-critical-mode decomposition and its derivation are given in Supplementary S1.

Let

$$S = \{i : v_i \neq 0\}, \qquad |S| = m,$$

be the support of the critical mode. For $i, j \in S$, the Pearson correlation coefficient is

$$\rho_{ij}(\alpha) = \frac{\alpha v_i v_j + R_{ij}}{\sqrt{(\alpha v_i^2 + R_{ii})(\alpha v_j^2 + R_{jj})}}.$$

Since the rank-one critical contribution dominates as $\alpha \to \infty$, one obtains $|\rho_{ij}| \to 1$ for $i, j \in S$. In contrast, for $i \in S$ and $j \notin S$, the critical contribution is absent from the j component, and under the same boundedness assumptions one obtains $|\rho_{ij}| \to 0$. These limits provide the dynamical basis for the conventional DNB conditions,

$$SD_{\text{in}} \uparrow, \qquad PCC_{\text{in}} \uparrow, \qquad PCC_{\text{out}} \downarrow .$$

A conventional DNB index may therefore be defined, for example, as

$$I_{\text{DNB}}(S) = SD_{\text{in}}(S) \frac{PCC_{\text{in}}(S)}{PCC_{\text{out}}(S) + \varepsilon},$$

and an increase in this quantity is interpreted as a pre-critical signal. A simpler index $I_{\text{DNB}}(S) = SD_{\text{in}}(S) PCC_{\text{in}}(S)$ is also widely used [1]

To obtain a graphical interpretation, we construct an unweighted Pearson correlation graph

$$G_\tau = (V, E_\tau), \qquad (i, j) \in E_\tau \quad \Longleftrightarrow \quad |\rho_{ij}| \geq \tau,$$

for a fixed threshold value $0 < \tau < 1$. Since $|\rho_{ij}| \to 1$ for all $i, j \in S$, sufficiently close to the tipping point the induced subgraph on the DNB region satisfies

$$G_\tau[S] = K_m.$$

In the ideal limit, correlations between S and the remaining variables also become sufficiently weak that the DNB region can appear as an isolated complete component. In realistic observations, however, finite sampling, measurement noise, and a finite distance from the tipping point generally prevent such perfect separation. It is therefore more natural to regard an emerging DNB region as a strongly coherent local structure that may remain embedded in the background network. We refer to such an internally coherent but not necessarily

graph-isolated structure as a *local nucleus*. This viewpoint motivates a description of DNB formation as a process in which a coherent local structure develops within the background network and subsequently grows into the full DNB region.

B. Heterogeneous Critical Modes and Early Local Nucleation

The limiting result $|\rho_{ij}| \to 1$ for $i, j \in S$ does not imply that all variables correlate at the same rate. Let the critical mode be nonuniform and order its components as

$$|v_1| \geq |v_2| \geq \cdots \geq |v_m| > 0.$$

Such nonuniformity can arise naturally in structurally heterogeneous systems, for example when state variables differ in their local dynamics, coupling strengths, or network positions[8, 24–27]. In networked dynamical systems, dominant modes can exhibit strongly nonuniform participation or become localized on particular nodes, mesoscale motifs, or network cores.

The correlation formula

$$\rho_{ij}(\alpha) = \frac{\alpha v_i v_j + R_{ij}}{\sqrt{(\alpha v_i^2 + R_{ii})(\alpha v_j^2 + R_{jj})}}$$

shows that pairs with larger $|v_i|$ and $|v_j|$ can become dominated by the critical rank-one contribution at smaller values of α. Consider a subset $Q \subsetneq S$, with $|Q| = q$, consisting of variables with relatively large critical-mode loadings. If the correlations within Q approach unity faster than the correlations between Q and $S \setminus Q$, then there exists a sufficiently high threshold value $\tau_H < 1$ and a finite range of the critical amplification parameter for which

$$|\rho_{ij}| \geq \tau_H \quad (i, j \in Q),$$

while

$$|\rho_{ij}| < \tau_H \quad (i \in Q,\ j \in S \setminus Q).$$

Consequently,

$$G_{\tau_H}[Q] = K_q, \qquad G_{\tau_H}[S] \neq K_m.$$

Thus, a small complete or nearly complete nucleus can emerge before the entire DNB region becomes globally correlated (fig 1). A more explicit sufficient condition based on the asymptotic convergence rates of the correlations is derived in Supplementary S2.

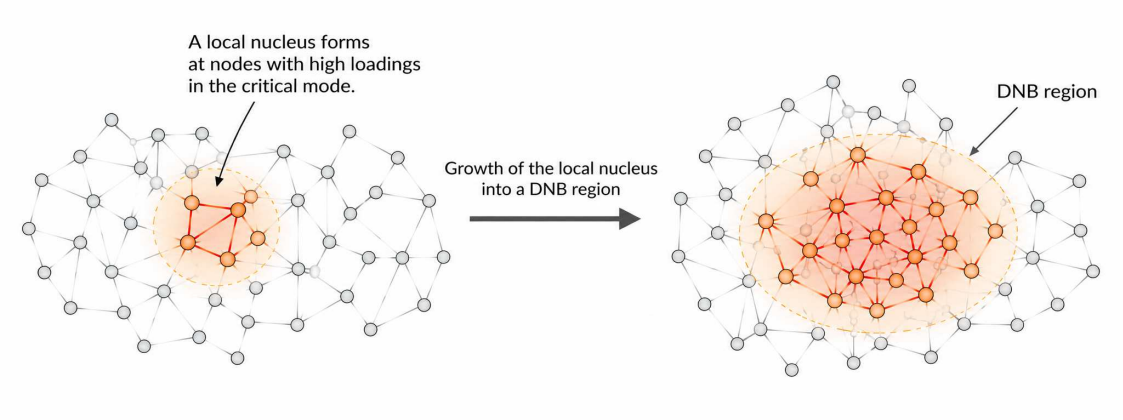


FIG. 1. Conceptual illustration of local-to-global DNB formation. A local nucleus first emerges among nodes with high critical-mode loadings and subsequently grows into the DNB region.

This observation suggests a local-to-global growth picture. Define nested subsets $Q_q = \{1, \ldots, q\}$. As the critical amplification increases, a small high-loading subset can first become coherent, after which variables with smaller critical-mode loadings are progressively incorporated. In an ideal thresholded graph, this process is represented by an increasing sequence of complete subgraphs,

$$K_{q_1} \subset K_{q_2} \subset \cdots \subset K_m.$$

In finite and noisy observations, the more appropriate interpretation is the progressive growth of an almost-complete coherent nucleus into a larger coherent region and eventually into the global DNB structure. The central point is that, under sufficient critical-mode heterogeneity, the formation time T_{nuc} of a local nucleus can precede the time T_{DNB} at which the global DNB structure becomes sufficiently developed to be detected by conventional DNB observables,

$$T_{\mathrm{nuc}} < T_{\mathrm{DNB}}.$$

Since conventional DNB quantities average fluctuations and correlations over a larger region, the contribution of a small correlated subset can be diluted when $q \ll m$. Detecting the local nucleus itself could therefore provide access to an earlier stage of critical-mode amplification.

This local-nucleation stage is not necessarily universal. If the critical mode is approximately homogeneous, many variables correlate at comparable rates and the distinction between local and global correlation becomes weak or disappears. The claim made here is therefore conditional: sufficiently heterogeneous critical-mode participation can create a

finite parameter regime in which a small coherent nucleus is already present while the full DNB region has not yet formed. The next subsection provides an exactly solvable example in which this separation can be calculated explicitly.

C. An Exactly Solvable Model of Local-to-Global DNB Formation

To provide an explicit nonlinear dynamical origin for the linear model considered below, we first introduce a bistable system undergoing a saddle-node bifurcation. Let $u \in \mathbb{R}^m$ satisfy $\|u\| = 1$, and define

$$P = uu^\top, \qquad z = u^\top x.$$

The vector u represents the critical direction. We consider the nonlinear deterministic system

$$\dot{x} = -(I - P)x + \left(\frac{1}{2}z - z^3 - h_\beta\right) u,$$

where

$$h_\beta = \frac{\beta}{3}\sqrt{\frac{3 - 2\beta}{6}}, \qquad 0 \le \beta \le 1.$$

This system can be written as a gradient flow,

$$\dot{x} = -\nabla V_\beta(x),$$

with potential

$$V_\beta(x) = \frac{1}{2}\|(I - P)x\|^2 + \Phi_\beta(z),$$

where

$$\Phi_\beta(z) = \frac{1}{4}z^4 - \frac{1}{4}z^2 + h_\beta z.$$

Thus, all directions orthogonal to u remain uniformly stable, whereas the dynamics along u is governed by a tilted double-well potential.

At $\beta = 0$, one has $h_\beta = 0$, and Φ_β is a symmetric double-well potential with two stable equilibria separated by one unstable equilibrium. As β increases, the potential becomes progressively tilted, and the stable equilibrium on the positive-u side approaches the neighboring unstable equilibrium.

The stable equilibrium that disappears at the transition is

$$x^*_\beta = a_\beta u, \qquad a_\beta = \sqrt{\frac{3 - 2\beta}{6}}.$$

Indeed,

$$\frac{1}{2}a_\beta - a_\beta^3 - h_\beta = 0.$$

As $\beta \to 1^-$,

$$a_\beta \longrightarrow \frac{1}{\sqrt{6}},$$

and the stable equilibrium coalesces with the adjacent unstable equilibrium. At $\beta = 1$, the two equilibria therefore undergo a saddle-node bifurcation, and one of the two stable states disappears.

This saddle-node scenario provides a canonical nonlinear realization of the single-critical-mode setting underlying DNB theory. In particular, the eigenvalue associated with the u direction at x_β^* is

$$\lambda_c = \frac{1}{2} - 3a_\beta^2 = -(1-\beta),$$

and hence

$$\lambda_c \to 0^- \qquad \text{as} \qquad \beta \to 1^-.$$

In contrast, all eigenvalues in directions orthogonal to u remain equal to -1. Thus, only a single mode becomes progressively softer as the critical transition is approached, while all noncritical modes remain uniformly stable.

We now linearize the nonlinear dynamics around the disappearing stable equilibrium. Writing

$$x = x_\beta^* + \eta,$$

we obtain, to first order,

$$\dot{\eta} = J(\beta)\eta,$$

where

$$J(\beta) = -(I - P) + \left(\frac{1}{2} - 3a_\beta^2\right) P = -I + \beta u u^\top.$$

Therefore, the linearized dynamics has the critical eigenvalue

$$\lambda_c = -(1-\beta)$$

along u, whereas every direction orthogonal to u has eigenvalue -1.

To describe fluctuations around this stable equilibrium, we introduce isotropic stochastic forcing and consider the m-dimensional Ornstein–Uhlenbeck process

$$dX_t = J(\beta)X_t\,dt + \sigma\,dW_t, \qquad J(\beta) = -I + \beta u u^\top,$$

where $0 < \beta < 1$ and $\|u\| = 1$. Here, X_t denotes the fluctuation around the stable equilibrium. The vector u represents the critical mode, while β controls the proximity to the tipping point. Along u, the eigenvalue is

$$\lambda_c = -(1-\beta),$$

whereas every direction orthogonal to u has eigenvalue -1. Hence,

$$\lambda_c \to 0^- \qquad \text{as} \qquad \beta \to 1^-,$$

while all noncritical modes remain stable.

The stationary covariance is [28]

$$C = \frac{\sigma^2}{2}\left(I + \frac{\beta}{1-\beta} uu^\top\right).$$

Introducing

$$g = \frac{\beta}{1-\beta}$$

gives

$$C = \frac{\sigma^2}{2}\left(I + guu^\top\right),$$

so that g directly measures the amplification of the critical mode.

To introduce a controlled heterogeneity, let $S = \{1, \ldots, m\}$ and choose a high-loading core $Q \subset S$ with $|Q| = q < m$. We define

$$u_i = \begin{cases} \dfrac{r}{\sqrt{qr^2 + m - q}}, & i \in Q, \\ \dfrac{1}{\sqrt{qr^2 + m - q}}, & i \in S \setminus Q, \end{cases} \qquad r \geq 1,$$

where r is the heterogeneity parameter. The homogeneous case corresponds to $r = 1$, while $r > 1$ gives larger critical-mode loadings to the variables in Q. It is convenient to define the rescaled amplification parameter

$$t = g\frac{r^2}{qr^2 + m - q}.$$

The three possible pairwise correlations are then

$$\begin{cases} \rho_{HH}(t) = \frac{t}{1+t}, \\ \rho_{HL}(t) = \frac{t}{\sqrt{(1+t)(r^2+t)}}, \\ \rho_{LL}(t) = \frac{t}{r^2+t} \end{cases}$$

where HH, HL, and LL denote high–high, high–low, and low–low pairs, respectively. For every fixed $r > 1$ and $t > 0$,

$$\rho_{HH}(t) > \rho_{HL}(t) > \rho_{LL}(t).$$

Thus, the high-loading core correlate first, followed by the correlations between the core and the remaining variables, while the low-loading variables correlate last.

For a fixed threshold value $0 < \tau < 1$, the high–high pairs cross the threshold value at

$$t_H(\tau) = \frac{\tau}{1-\tau},$$

whereas the low–low pairs cross at

$$t_L(\tau, r) = r^2 \frac{\tau}{1-\tau}.$$

The high–low threshold-crossing scale t_X is defined by $\rho_{HL}(t_X) = \tau$. For $r > 1$, the three scales satisfy

$$t_H < t_X < t_L.$$

The exact expression for t_X and the derivation of this ordering are given in Supplementary S3.

The thresholded correlation graph therefore passes through distinct structural regimes. For $t < t_H$, no pair exceeds the threshold value. For $t_H \leq t < t_X$, only high–high pairs exceed the threshold value and

$$G_\tau[S] = K_q \sqcup (m-q)K_1.$$

A small local nucleus K_q is therefore present while the remaining variables are still outside the coherent structure. For $t_X \leq t < t_L$, high–low correlations also exceed the threshold value, and the correlated region expands beyond the original core. Finally, for $t \geq t_L$, all pairwise correlations exceed the threshold value and

$$G_\tau[S] = K_m.$$

The model therefore exhibits an explicit local-to-global formation process in which a small coherent nucleus appears before the complete DNB region.

The separation between local and global formation is controlled directly by the heterogeneity parameter. Since

$$\frac{t_L}{t_H} = r^2,$$

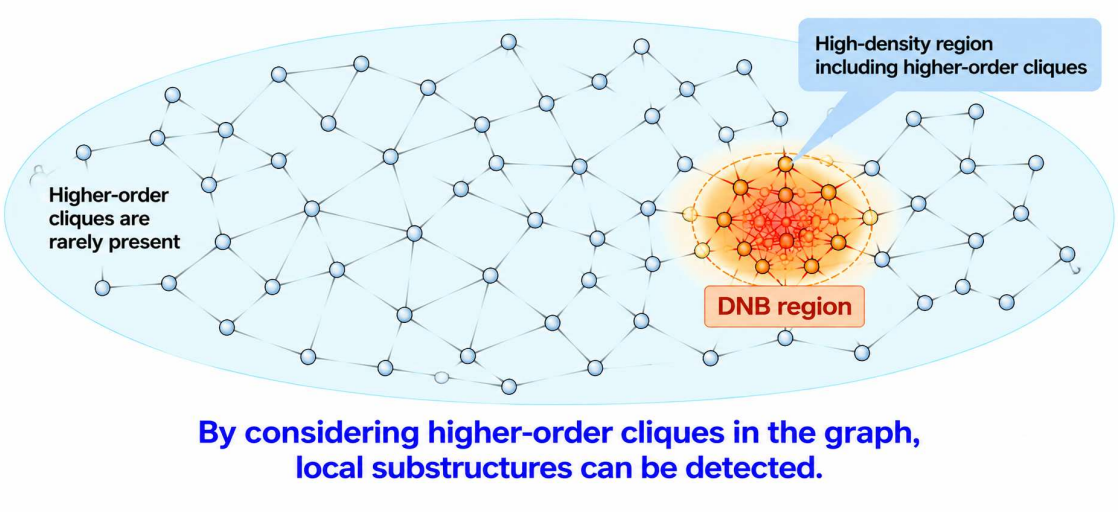


FIG. 2. Conceptual illustration of higher-order detection of a local DNB nucleus.

a larger r produces a wider interval between the formation of the small nucleus and the completion of the global DNB structure. In contrast, for $r = 1$ one has $\rho_{HH} = \rho_{HL} = \rho_{LL}$ and hence $t_H = t_X = t_L$; namely, the intermediate local-nucleation regime disappears. This exactly solvable model therefore confirms that local nucleation is not an automatic consequence of DNB formation itself, but can arise from heterogeneity of the critical mode. It also shows why detecting the small nucleus before the system reaches t_L could provide information about the approaching critical transition at an earlier stage. The next section addresses the remaining problem: how to detect such a small nucleus when its location and size are not known in advance and when it remains embedded in a much larger background network.

D. Higher-Order Spectral Detection of a Local Nucleus

The local nucleus described in the previous section need not be completely separated from the background network, making it difficult to detect using ordinary graph structure alone. In contrast, many mutually consistent edges are concentrated within the local nucleus, leading to the formation of higher-order cliques, whereas such higher-order structures are less likely to arise in sparse background regions (fig 2). We therefore lift the correlation graph to a clique complex and use the spectra of higher-order Hodge Laplacians to detect the local nucleus. A quantitative analysis showing that clique-complex lifting suppresses

incoherent background connections is provided in Supplementary S4.

Ideally, if a local nucleus consisting of q vertices forms a full simplex Δ^{q-1} and is separated from the background at the higher-order level, the pth Hodge Laplacian satisfies

$$L_p(\Delta^{q-1}) = qI_{\binom{q}{p+1}}.$$

Its spectrum therefore contains the eigenvalue

$$\lambda = q$$

with multiplicity

$$\binom{q}{p+1}.$$

Thus, a coherent local structure appears in the spectrum through its characteristic scale q together with a high eigenvalue multiplicity.

In practice, a local nucleus need not form a complete simplex. We therefore consider a local structure consisting of a coherent q-vertex simplex core Δ^{q-1} together with an incomplete surrounding structure $\mathcal{L}$. Let

$$a_j(\eta) = \#\left\{\mu \in \operatorname{Spec} \widetilde{L}_j(\mathcal{L}) : 0 \le \mu \le \eta\right\}$$

denote the number of low-lying eigenvalues of the reduced jth Hodge Laplacian of $\mathcal{L}$. The theoretical multiplicity of eigenvalues generated near $\lambda \simeq q$ in the pth Hodge spectrum is then

$$M_{p,q}(\eta) = \sum_{j=-1}^{p} a_j(\eta)\binom{q}{p-j}.$$

We refer to this cross-order binomial multiplicity pattern as a *binomial spectral fingerprint*.

The full simplex is recovered as a special case, for which

$$M_{p,q} = \binom{q}{p+1}.$$

If, for example, $a_{-1} = 0, a_0 = 1, a_1 = 0, a_2 = 0, \cdots$, then

$$M_{p,q} = \binom{q}{p}.$$

These values provide the theoretical spectral fingerprint used in the empirical analysis below.

When coupling to the background or perturbations of the internal structure are present, the exact eigenvalue degeneracy is generally lifted. However, if these perturbations are

sufficiently small, the corresponding eigenvalues remain concentrated as a high-multiplicity cluster near $\lambda \simeq q$. Thus, the cross-order multiplicity pattern in the higher-order Hodge spectra can be used to detect not only an ideal complete local nucleus but also an imperfect one. Detailed derivations and perturbation analyses are provided in Supplementary S4.

E. Empirical Detection of Local Nucleation in TSOD Mice

We next applied the higher-order spectral framework introduced in the preceding section to longitudinal gene-expression data from TSOD mice and examined whether a smaller local higher-order structure could be detected before the conventional DNB signal became pronounced. We used the previously identified 147-gene DNB region [21] as a fixed analysis domain. The purpose of the present analysis was not to identify a new DNB gene set, but to determine whether a local coherent structure had already formed within a region known to subsequently exhibit a pronounced DNB signal.

We first examined the temporal behavior of the conventional DNB indicator in this 147-gene region. The DNB score in the TSOD group was 0.337 at week 3 and 0.439 at week 4, but increased sharply to 2.218 at week 5 and subsequently decreased again at weeks 6 and 7. At week 5, the within-region fluctuations and correlations also increased, consistent with the conventional DNB conditions (fig 3). We therefore regarded week 5 as the time point at which the conventional DNB signal of the 147-gene region became most pronounced in this dataset.

We then examined whether a smaller coherent structure was already present at week 4, one week before the DNB peak, using the higher-order Hodge spectra. Both the second- and third-order Hodge Laplacians exhibited a pronounced concentration of eigenvalues around the common characteristic scale

$$\lambda \simeq 14.$$

In the second-order Hodge spectrum, the near-14 cluster contained

$$N_{2,14}(0.5) = 91$$

eigenvalues, whereas the corresponding cluster in the third-order Hodge spectrum contained

$$N_{3,14}(0.5) = 364$$

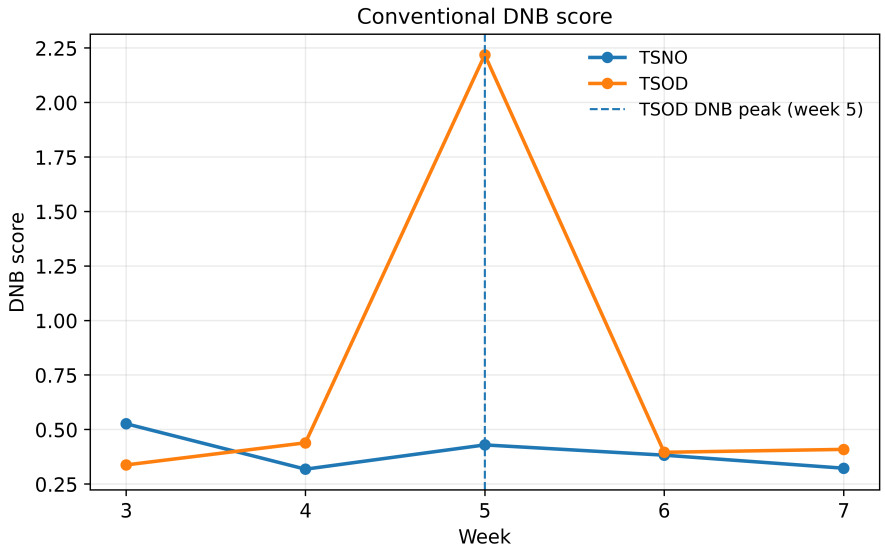


FIG. 3. Conventional DNB score from week 3 to week 7 in TSOD and TSNO mice. The DNB score was evaluated using the fixed 147-gene DNB region previously identified by Koizumi et al. [21]. A pronounced peak is observed at week 5 in the TSOD group, whereas the score remains substantially lower at week 4.

eigenvalues. In both cases, the cluster was followed by a clear spectral separation from the subsequent eigenvalues (fig 4).

These multiplicities coincide with the binomial spectral fingerprint derived in the preceding section for a local structure consisting of a coherent core with $q = 14$ and $a_{-1} = 0, a_0 = 1, a_1 = 0, a_2 = 0, \cdots,$

$$M_{2,14} = \binom{14}{2} = 91, \qquad M_{3,14} = \binom{14}{3} = 364.$$

Thus, at two independent Hodge orders, not only did the spectral locations indicate the same characteristic scale $q = 14$, but the corresponding multiplicities simultaneously agreed with the theoretically predicted cross-order binomial pattern.

These results indicate that a higher-order coherent structure substantially smaller than the full 147-gene DNB region was already present at week 4, before the conventional DNB signal reached its maximum at week 5. However, the spectral information alone does not uniquely identify a specific set of 14 genes, nor does it imply that the local structure is an exact 14-vertex clique. We therefore interpret the observed structure as a *spectral local-nucleus candidate* with characteristic scale $q = 14$.

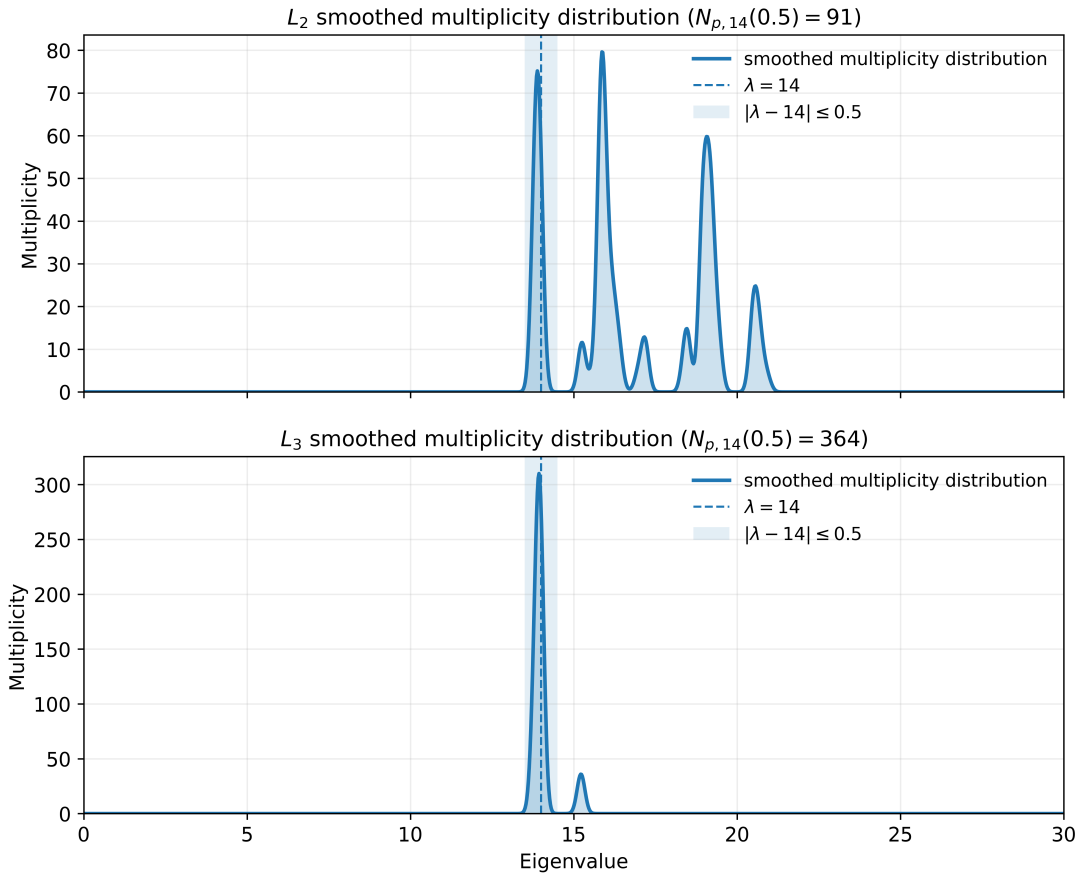


FIG. 4. Smoothed spectral multiplicity distributions of the second- and third-order Hodge Laplacians for TSOD mice at week 4. The dashed vertical line indicates the characteristic spectral scale $\lambda = 14$, and the shaded region denotes the interval $|\lambda - 14| \leq 0.5$ used to evaluate the near-14 multiplicity. For L_2, the interval contains $N_{2,14}(0.5) = 91$ eigenvalues, matching the theoretical binomial prediction $M_{2,14} = \binom{14}{2} = 91$. For L_3, the corresponding multiplicity is $N_{3,14}(0.5) = 364$, matching $M_{3,14} = \binom{14}{3} = 364$. Thus, both Hodge orders exhibit a spectral concentration around the same characteristic scale $q = 14$ with multiplicities consistent with the predicted cross-order binomial spectral fingerprint.

Taken together, these results show that, within the previously established DNB region, a more localized higher-order structure can be detected as a binomial fingerprint in the Hodge spectrum before the conventional DNB signal becomes pronounced. This observation is consistent with the local-nucleation scenario proposed in this study, in which heterogeneity of the critical mode leads to the prior formation of a local nucleus that subsequently develops into a larger DNB structure. Temporal comparisons across all time points in the TSOD and

TSNO groups are provided in the next subsection.

It should be noted that the present analysis is a proof-of-principle study in which the previously identified 147-gene DNB region was used as a fixed analysis domain to examine whether an unknown local nucleus existed within it, rather than an attempt to identify the DNB region itself at an earlier stage from the whole transcriptome. In addition, the characteristic scale $q = 14$ obtained from the spectrum does not uniquely identify a specific set of 14 genes. The present result should therefore be interpreted as showing that, within a previously established DNB region, a smaller higher-order coherent structure can form before the conventional DNB signal becomes most pronounced. In this sense, the binomial spectral fingerprint observed at week 4 provides empirical support for the local-nucleation scenario proposed in this study.

F. Temporal Specificity of the $q \simeq 14$ Local-Nucleus Spectral Fingerprint

To examine whether the binomial spectral fingerprint observed at TSOD week 4 represents a temporally distinctive structure, we fixed the correlation threshold at $\tau = 0.995$, the characteristic scale at $q = 14$, and the spectral window at $\delta = 0.5$, and applied exactly the same analysis to all time points from week 3 to week 7 in both TSOD and TSNO mice (fig 5). For each condition, we calculated the near-14 multiplicities of the second- and third-order Hodge Laplacians, $N_{2,14}$ and $N_{3,14}$, and compared them with the theoretical values $M_{2,14} = 91$ and $M_{3,14} = 364$. We further defined

$$R_2 = \frac{N_{2,14}}{91}, \qquad R_3 = \frac{N_{3,14}}{364},$$

to evaluate the consistency with the theoretical fingerprint across the two Hodge orders (fig 6).

At TSOD week 3, $(R_2, R_3) \simeq (0.68, 0.24)$, whereas at week 4,

$$N_{2,14} = 91, \qquad N_{3,14} = 364,$$

giving $(R_2, R_3) = (1, 1)$ and therefore exact agreement with the theoretical binomial fingerprint at both Hodge orders. At week 5, $N_{2,14} = 90$ and $N_{3,14} = 355$, corresponding to $(R_2, R_3) \simeq (0.989, 0.975)$, indicating that the same cross-order structure was largely preserved. In contrast, the corresponding near-14 multiplicities were no longer observed at weeks 6 and 7.

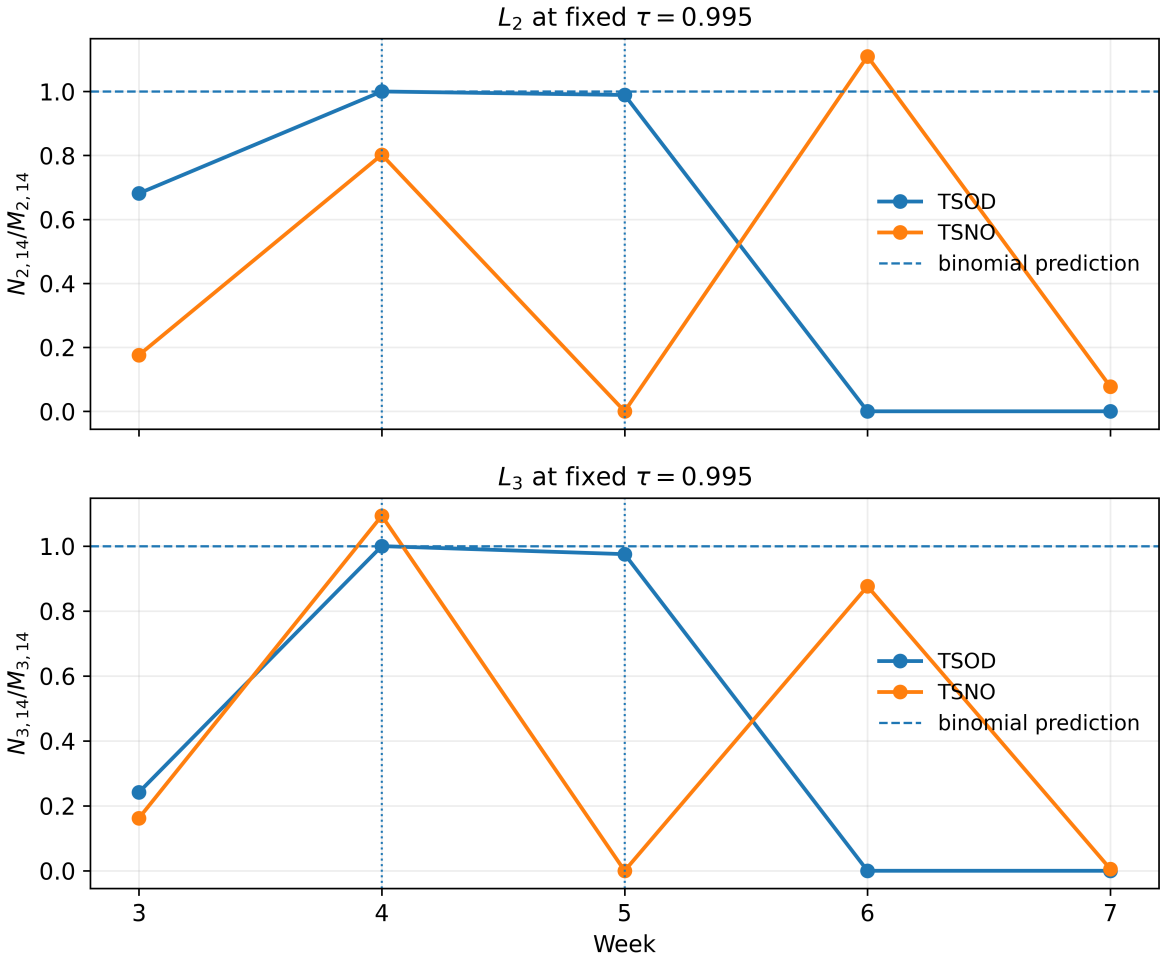


FIG. 5. Temporal comparison of the fixed-q spectral fingerprint in TSOD and TSNO mice from week 3 to week 7. The upper and lower panels show the normalized near-14 multiplicities $N_{2,14}/M_{2,14}$ and $N_{3,14}/M_{3,14}$, respectively, at the fixed correlation threshold $\tau = 0.995$. The dashed horizontal line indicates the theoretical binomial prediction. TSOD week 4 exactly matches the theoretical multiplicity at both Hodge orders, whereas TSOD week 5 remains close to the prediction.

In the TSNO group, no time point from week 3 to week 7 simultaneously satisfied $(R_2, R_3) = (1, 1)$ at both Hodge orders. Thus, under the fixed analysis conditions, TSOD week 4 was the only condition showing exact agreement with the theoretical 91/364 binomial fingerprint at both Hodge orders, while the same structure remained close to the theoretical prediction at week 5 during the DNB peak. This temporal ordering is consistent with

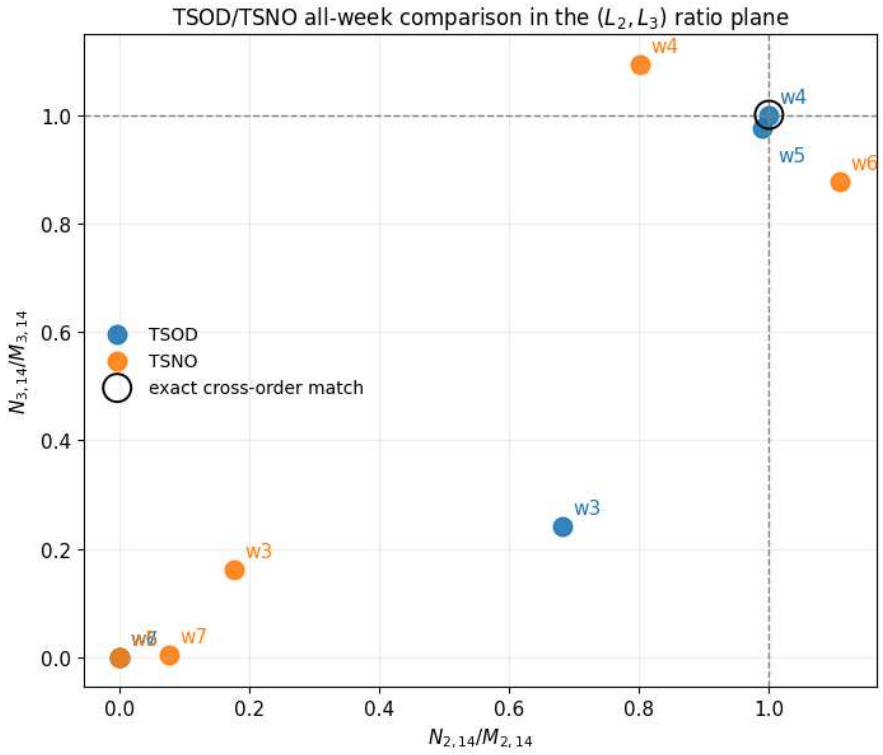


FIG. 6. Cross-order consistency of the $q = 14$ binomial spectral fingerprint across TSOD and TSNO mice. Each point represents one strain–week condition in the plane defined by $R_2 = N_{2,14}/M_{2,14}$ and $R_3 = N_{3,14}/M_{3,14}$. The point $(1, 1)$ represents simultaneous agreement with the theoretical multiplicities at both Hodge orders. TSOD week 4 lies exactly at $(1, 1)$, whereas TSOD week 5 remains close to this point. TSNO conditions deviate from the theoretical point in at least one Hodge order.

the local-to-global scenario in which a local nucleus forms before the larger DNB structure becomes most pronounced.

III. SUMMARY AND DISCUSSION

In this study, we reformulated the formation of Dynamical Network Biomarkers (DNBs) not only in terms of conventional statistical indicators but also as the emergence of local geometric structures in correlation networks, and proposed a framework for detecting this process using clique complexes and higher-order Hodge Laplacians. In particular, we focused on the possibility that, before the entire DNB region becomes strongly correlated, a subset

of variables first forms a highly coherent local nucleus and subsequently grows into a larger DNB structure.

Our theoretical analysis showed that, in a regime dominated by a single critical mode, heterogeneity in the critical-mode loadings can produce differences in the rates at which correlations increase, allowing variables with large loadings to form a local nucleus before the entire DNB region becomes coherent. This local-to-global formation process was further demonstrated explicitly using an analytically tractable model. Because a local nucleus may remain embedded in the background network, it can be difficult to detect from ordinary graph structure alone, whereas clique-complex lifting emphasizes mutually consistent higher-order structures. An ideal local simplex of q vertices produces a high-multiplicity Hodge eigenvalue at $\lambda = q$, and even in the presence of background coupling or internal imperfections, this signature can persist as a near-q spectral cluster or a cross-order binomial spectral fingerprint.

In the gene-expression data from TSOD mice, the conventional DNB signal of the previously identified 147-gene region became pronounced at week 5, whereas a higher-order spectral structure with characteristic scale $q = 14$ was already observed at week 4. The multiplicities observed in the second- and third-order Hodge spectra were consistent with the theoretically predicted binomial spectral fingerprint, supporting the possibility that a smaller local coherent structure forms before the conventional DNB peak. Higher-order Hodge analysis should therefore be viewed not as a replacement for conventional DNB analysis, but as an additional observable for capturing an earlier local stage of the DNB formation process.

Future work should test the generality of local nucleation and higher-order spectral signatures across different diseases, biological species, and datasets involving critical transitions outside biological systems. In addition, because the number of higher-order simplices in a clique complex grows combinatorially, application to larger datasets will require computational improvements such as sparse matrix representations, local simplex enumeration, and iterative eigensolvers. Spectral analysis of large-scale Hodge Laplacians may also become a potential application of quantum approaches [29–38] to linear algebra and spectral estimation, although substantial challenges remain, including efficient encoding, state preparation, and extraction of the required spectral information.

In summary, this study provides a framework in which DNB formation is viewed not only as the emergence of a completed global structure but also as a geometric process involving an

earlier stage of local nucleation. The theoretical analysis and TSOD results suggest that a small coherent structure can form before the conventional DNB signal and that information about this structure can be reflected in higher-order Hodge spectra. Further validation using independent datasets and the development of computational methods for large-scale higher-order Laplacians may allow this framework to evolve into a more general approach for detecting earlier structural changes associated with critical transitions.

IV. METHODS

A. Dataset and Analysis Domain

We used the longitudinal gene-expression dataset GSE112653 reported by Koizumi et al. The analysis included TSOD and TSNO mice from week 3 to week 7. In the TSOD group, the numbers of samples were five at weeks 3, 4, 5, and 7, and four at week 6. In the TSNO group, five samples were available at each time point from week 3 to week 7.

The purpose of the present study was not to identify a new DNB gene set from the whole transcriptome, but to examine whether a smaller local coherent structure forms within a previously identified DNB region before the conventional DNB signal becomes pronounced. We therefore used the previously identified 147-gene DNB region as a fixed analysis domain. The 147 genes were not re-selected separately for individual strains or time points and were used consistently throughout all analyses.

For the higher-order spectral analysis, expression matrices organized by strain and time point were used as input. No additional normalization or transformation was applied to the input expression matrices in the present analysis.

B. Conventional DNB Analysis

The conventional DNB signal was evaluated for the fixed 147-gene region from week 3 to week 7. We used SD_{in} to represent fluctuations within the DNB region, PCC_{in} to represent the absolute Pearson correlation within the region, and PCC_{out} to represent the absolute Pearson correlation between the DNB region and variables outside the region. The DNB score was evaluated as

$$I_{\mathrm{DNB}} = SD_{\mathrm{in}} \frac{PCC_{\mathrm{in}}}{PCC_{\mathrm{out}} + \varepsilon},$$

where ε is a small positive constant introduced to avoid division by zero. The same 147-gene set was used at all time points, and the time point at which the DNB score was largest in the TSOD group was used as the time at which the conventional DNB signal was most pronounced for comparison with the higher-order spectral analysis.

C. Construction of Correlation Networks and Clique Complexes

For each strain and time point, Pearson correlation coefficients were calculated between the 147 genes. Let $|\rho_{ij}|$ denote the absolute Pearson correlation coefficient between genes i and j. An undirected edge was placed between two genes when

$$|\rho_{ij}| \geq \tau.$$

The main analysis used the correlation threshold

$$\tau = 0.995.$$

Self-loops were excluded. When a Pearson correlation coefficient was numerically undefined, its value was treated as zero.

A clique complex was constructed from the resulting binary correlation graph. A set of $p+1$ mutually connected vertices was treated as a p-simplex. To construct the second- and third-order Hodge Laplacians, triangles, tetrahedra, and five-vertex cliques were enumerated.

An orientation was assigned to each simplex by fixing the vertex ordering, and the boundary matrices

$$B_2, \qquad B_3, \qquad B_4$$

were constructed. The second- and third-order Hodge Laplacians were defined as

$$L_2 = B_2^\top B_2 + B_3 B_3^\top,$$

and

$$L_3 = B_3^\top B_3 + B_4 B_4^\top,$$

respectively. The consistency of the implementation was verified by confirming

$$B_2 B_3 = 0, \qquad B_3 B_4 = 0.$$

The boundary matrices and Hodge Laplacians were stored as sparse matrices.

D. Computation of Hodge Spectra

In the main analysis, the spectra of the second- and third-order Hodge Laplacians were analyzed for the TSOD week-4 correlation network. To compare the observed spectra with the theoretical binomial spectral fingerprint, eigenvalues near the characteristic scale

$$q = 14$$

were evaluated.

For large Hodge Laplacians, up to 400 eigenvalues closest to $\lambda = 14$ were computed using a sparse eigensolver with the shift-invert method. The number 400 was chosen to exceed the largest theoretical multiplicity considered in the present analysis,

$$M_{3,14} = 364,$$

so that the theoretical eigenvalue cluster would not be truncated. For smaller matrices, all eigenvalues were computed directly.

To avoid numerical instability caused by choosing a shift exactly equal to an eigenvalue, the shift-invert calculation primarily used

$$\sigma = 14 + 10^{-6}.$$

The convergence tolerance was set to 10^{-10}, with a maximum of 100000 iterations. If numerical factorization failed at the primary shift, a deterministic fallback shift

$$\sigma = 14 + 10^{-4}$$

was used.

E. Near-q Multiplicity and the Binomial Spectral Fingerprint

Let $\lambda_j^{(p)}$ denote the eigenvalues of the pth Hodge Laplacian. We defined the number of eigenvalues near the characteristic scale q as

$$N_{p,q}(\delta) = \#\left\{ j : |\lambda_j^{(p)} - q| \le \delta \right\}.$$

In the main analysis, we used

$$q = 14, \qquad \delta = 0.5,$$

and evaluated

$$N_{2,14}(0.5), \qquad N_{3,14}(0.5).$$

Thus, for each Hodge order, we counted the eigenvalues satisfying

$$13.5 \leq \lambda \leq 14.5.$$

For the degree-zero attachment profile derived in the preceding theoretical section,

$$a_{-1} = 0, \qquad a_0 = 1, \qquad a_j = 0 \quad (j \neq 0),$$

the theoretical binomial spectral fingerprint is

$$M_{p,q} = \binom{q}{p}.$$

Accordingly, for $q = 14$,

$$M_{2,14} = \binom{14}{2} = 91, \qquad M_{3,14} = \binom{14}{3} = 364.$$

The observed values of $N_{2,14}(0.5)$ and $N_{3,14}(0.5)$ were compared with these theoretical multiplicities to assess whether the same characteristic scale and the corresponding binomial multiplicities appeared simultaneously across different Hodge orders.

F. Numerical Implementation

All analyses were performed in Python. NumPy and pandas were used for numerical arrays and data processing, SciPy was used for sparse-matrix construction and eigenvalue computation, and Matplotlib was used for visualization. Eigenvalues of large Hodge Laplacians were computed using the ARPACK-based `scipy.sparse.linalg.eigsh` routine with the shift-invert method.

DATA AVAILABILITY

The gene-expression data analysed in this study are publicly available from the NCBI Gene Expression Omnibus under accession number GSE112653.

CODE AVAILABILITY

The custom code used for the numerical and spectral analyses in this study is available from the corresponding author upon reasonable request.

[1] L. Chen, R. Liu, Z.-P. Liu, M. Li, and K. Aihara, Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers, Scientific reports **2**, 342 (2012).

[2] R. Liu, K. Aihara, and L. Chen, Dynamical network biomarkers for identifying critical transitions and their driving networks of biologic processes, Quantitative Biology **1**, 105 (2013).

[3] R. Liu, X. Wang, K. Aihara, and L. Chen, Early diagnosis of complex diseases by molecular biomarkers, network biomarkers, and dynamical network biomarkers, Medicinal research reviews **34**, 455 (2014).

[4] R. Liu, X. Yu, X. Liu, D. Xu, K. Aihara, and L. Chen, Identifying critical transitions of complex diseases based on a single sample, Bioinformatics **30**, 1579 (2014).

[5] R. Liu, P. Chen, K. Aihara, and L. Chen, Identifying early-warning signals of critical transitions with strong noise by dynamical network markers, Scientific reports **5**, 17501 (2015).

[6] X. Liu, X. Chang, R. Liu, X. Yu, L. Chen, and K. Aihara, Quantifying critical states of complex diseases using single-sample dynamic network biomarkers, PLoS computational biology **13**, e1005633 (2017).

[7] X. Liu, X. Chang, S. Leng, H. Tang, K. Aihara, and L. Chen, Detection for disease tipping points by landscape dynamic network biomarkers, National science review **6**, 775 (2019).

[8] T. Matsumori, H. Sakai, and K. Aihara, Early-warning signals using dynamical network markers selected by covariance, Physical Review E **100**, 052303 (2019).

[9] K. Aihara, R. Liu, K. Koizumi, X. Liu, and L. Chen, Dynamical network biomarkers: Theory and applications, Gene **808**, 145997 (2022).

[10] M. A. S. Kamal, M. Oku, T. Hayakawa, J.-i. Imura, and K. Aihara, Early detection of a traffic flow breakdown in the freeway based on dynamical network markers, International Journal of Intelligent Transportation Systems Research **18**, 422 (2020).

[11] P. Chen, E. Chen, L. Chen, X. J. Zhou, and R. Liu, Detecting early-warning signals of influenza outbreak based on dynamic network marker, Journal of cellular and molecular medicine **23**, 395 (2019).

[12] K. Yang, J. Xie, R. Xie, Y. Pan, R. Liu, and P. Chen, Real-time forecast of influenza outbreak using dynamic network marker based on minimum spanning tree, BioMed Research International **2020**, 7351398 (2020).

[13] Á. Fragua, A. Jiménez-Martín, and A. Mateos, Complex network analysis techniques for the early detection of the outbreak of pandemics transmitted through air traffic, Scientific Reports **13**, 18174 (2023).

[14] M. I. Ito, H. Hasada, Y. Honma, T. Ohnishi, T. Watanabe, and K. Aihara, Identifying dynamical network markers of financial market instability, arXiv preprint arXiv:2604.21297 (2026).

[15] D. Horak and J. Jost, Spectra of combinatorial laplace operators on simplicial complexes, Advances in Mathematics **244**, 303 (2013).

[16] S. Mukherjee and J. Steenbergen, Random walks on simplicial complexes and harmonics, Random structures & algorithms **49**, 379 (2016).

[17] O. Parzanchevski and R. Rosenthal, Simplicial complexes: spectrum, homology and random walks, Random Structures & Algorithms **50**, 225 (2017).

[18] M. T. Schaub, A. R. Benson, P. Horn, G. Lippner, and A. Jadbabaie, Random walks on simplicial complexes and the normalized hodge 1-laplacian, SIAM Review **62**, 353 (2020).

[19] L.-H. Lim, Hodge laplacians on graphs, Siam Review **62**, 685 (2020).

[20] S. Barbarossa and S. Sardellitti, Topological signal processing over simplicial complexes, IEEE Transactions on Signal Processing **68**, 2992 (2020).

[21] K. Koizumi, M. Oku, S. Hayashi, A. Inujima, N. Shibahara, L. Chen, Y. Igarashi, K. Tobe, S. Saito, M. Kadowaki, *et al.*, Identifying pre-disease signals before metabolic syndrome in mice by dynamical network biomarkers, Scientific Reports **9**, 8767 (2019).

[22] I. Hirayama, Z. Yi, S. Izumi, I. Arai, W. Suzuki, Y. Nagamachi, H. Kuwano, T. Takeuchi, and T. Izumi, Genetic analysis of obese diabetes in the tsod mouse., Diabetes **48**, 1183 (1999).

[23] W. SUZUKI, S. IIZUKA, M. TABUCHI, S. FUNO, T. YANAGISAWA, M. KIMURA, T. SATO, T. ENDO, and H. KAWAMURA, A new mouse model of spontaneous diabetes-derived from ddy strain, Experimental Animals **48**, 181 (1999).

[24] H. Aufderheide, L. Rudolf, and T. Gross, Mesoscale symmetries explain dynamical equivalence of food webs, New Journal of Physics **14**, 105014 (2012).

[25] A.-L. Do, J. Höfener, and T. Gross, Engineering mesoscale structures with distinct dynamical implications, New Journal of Physics **14**, 115022 (2012).

[26] R. Pastor-Satorras and C. Castellano, Distinct types of eigenvector localization in networks, Scientific reports **6**, 18847 (2016).

[27] S. Chen, E. B. O'Dea, J. M. Drake, and B. I. Epureanu, Eigenvalues of the covariance matrix as early warning signals for critical transitions in ecological systems, Scientific reports **9**, 2572 (2019).

[28] M. Oku and K. Aihara, On the covariance matrix of the stationary distribution of a noisy dynamical system, Nonlinear Theory and Its Applications, IEICE **9**, 166 (2018).

[29] S. Lloyd, S. Garnerone, and P. Zanardi, Quantum algorithms for topological and geometric analysis of data, Nature communications **7**, 10138 (2016).

[30] C. Gyurik, C. Cade, and V. Dunjko, Towards quantum advantage via topological data analysis, Quantum **6**, 855 (2022).

[31] I. Y. Akhalwaya, A. Bhayat, A. Connolly, S. Herbert, L. Horesh, J. Sorci, and S. Ubaru, Comparing quantum and classical monte carlo algorithms for estimating betti numbers of clique complexes, arXiv preprint arXiv:2408.16934 (2024).

[32] H. Yamauchi, S. Kanno, Y. Sato, H. Tezuka, Y.-a. Shimada, E. Kaminishi, and N. Yamamoto, Quantum spectroscopy of topological dynamics via a supersymmetric hamiltonian, arXiv preprint arXiv:2511.23169 (2025).

[33] I. Akhalwaya, S. Ubaru, K. Clarkson, M. Squillante, V. Jejjala, Y.-H. He, K. Naidoo, V. Kalantzis, and L. Horesh, Topological data analysis on noisy quantum computers, in *International Conference on Learning Representations*, Vol. 2024 (2024) pp. 34945–34978.

[34] M. Crichigno and T. Kohler, Clique homology is qma 1-hard, Nature Communications **15**, 9846 (2024).

[35] C. Cade and P. M. Crichigno, Complexity of supersymmetric systems and the cohomology problem, Quantum **8**, 1325 (2024).

[36] A. Schmidhuber and S. Lloyd, Complexity-theoretic limitations on quantum algorithms for topological data analysis, PRX Quantum **4**, 040349 (2023).

[37] S. Apers, S. Gribling, S. Sen, and D. Szabó, A (simple) classical algorithm for estimating betti numbers, Quantum **7**, 1202 (2023).

[38] A. Connolly, A. Bhayat, C. Kargı, C. Krawchuk, D. M. Ramo, G. Prokopiou, H. Goto, I. Williams, I. Akhalwaya, K. Yamamoto, K. Xie, P. Kalidindi, R. Nishimura, S. Kanno, S. Maekawa, S. McAdams, S. Clark, T. M. Soini, and Y. aki Shimada, *Quantum Computing Frontiers — A Use-Case Timeline for Quantum Chemistry and Topological Data Analysis —*, White Paper (SoftBank Corp. & Quantinuum, 2026).

ACKNOWLEDGEMENTS

K.A. was supported by the Japan Science and Technology Agency (JST) Moonshot Research and Development Program (Grant No. JPMJMS2021).

AUTHOR CONTRIBUTIONS

S.K. developed the theoretical framework and performed the numerical analyses. K.A. contributed substantially to the conception of the study, development of the theoretical framework, interpretation of the results, and supervision of the overall scientific direction of the work. Y.S. contributed to the refinement of the theoretical framework, interpretation of the results, and critical review of the manuscript. All authors contributed to discussions, critically revised the manuscript, and approved the final version.

COMPETING INTERESTS

S.K. and Y.S. are employees of SoftBank Corp. K.A. declares no competing interests.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- opennaturecommcode.zip