

A hypergraph neural network-based framework for identification of Parkinson's disease subtypes

Meili Lu(卢梅丽) and Hongbao Lu(卢宏保)

Citation: Chin. Phys. B, 2026, 35 (**7**): 078703. DOI: 10.1088/1674-1056/ae2677
http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Evolutionary hypergraph dismantling via deep reinforcement learning

Junjie Qian(钱俊杰), Wenlan Wang(王文蓝), Hanyun Wang(王涵韵), Qiqi Wang(王萁淇), Yao Zhang(张瑶), and Huijia Li(李慧嘉)
Chin. Phys. B, 2026, 35 (**6**): 068901. DOI: 10.1088/1674-1056/ae5a12

A node importance prediction algorithm based on graph attention and contrastive learning

Jun Ai(艾均), Yuming Zhang(张玉明), Zhan Su(苏湛), Chenye Guo(郭晨晔), and Mingsong Li(李铭松)
Chin. Phys. B, 2026, 35 (**5**): 058901. DOI: 10.1088/1674-1056/ae0680

Six-degree gravity centrality for detecting influential nodes in networks

Jianbo Wang(王建波), Bohang Lin(林渤杭), Zhanwei Du(杜占玮), Ping Li(李平), and Xiao-Ke Xu(许小可)
Chin. Phys. B, 2025, 34 (**8**): 088902. DOI: 10.1088/1674-1056/adec62

Epidemic threshold influenced by non-pharmaceutical interventions in residential university environments

Zechao Lu(卢泽超), Shengmei Zhao(赵生妹), Huazhong Shu(束华中), and Long-Yan Gong(巩龙延)
Chin. Phys. B, 2024, 33 (**2**): 028707. DOI: 10.1088/1674-1056/ace2b0

Detecting overlapping communities in networks via dominant label propagation

Sun He-Li (孙鹤立), Huang Jian-Bin (黄健斌), Tian Yong-Qiang (田勇强), Song Qin-Bao (宋擒豹), Liu Huai-Liang (刘怀亮)
Chin. Phys. B, 2015, 24 (**1**): 018703. DOI: 10.1088/1674-1056/24/1/018703

A hypergraph neural network-based framework for identification of Parkinson's disease subtypes

Meili Lu(卢梅丽)† and Hongbao Lu(卢宏保)

School of Information Technology Engineering, Tianjin University of Technology and Education, Tianjin 300000, *China*

(Received 7 November 2025; revised manuscript received 20 November 2025; accepted manuscript online 2 December 2025)

Parkinson's disease (PD) exhibits significant phenotypic heterogeneity, which complicates clinical management and underscores the need for precise subtyping. Existing subtyping approaches often rely on a single modality, such as clinical assessments, failing to capture the complex, multi-faceted nature of the disease. This paper proposes a novel computational framework that integrates multi-modal data, specifically preprocessed functional MRI, DNA methylation, and clinical behavioral assessments, for PD subtyping. The methodology involves constructing individual hypergraphs for each modality using K-nearest neighbors (KNN), followed by the integration of these hypergraphs into a unified, multi-modal hypergraph using similarity network fusion (SNF). This consolidated hypergraph is then processed via a hypergraph neural network (HGNN) utilizing hyperedge convolution to cluster patients into distinct subtypes. Our experimental results demonstrate that this approach effectively identifies PD subtypes with significant clinical and biological relevance. We provide a comprehensive analysis of the model's performance and further validate the reliability of the identified subtypes through post-hoc statistical tests. This study highlights the potential of graph-based machine learning in disentangling disease heterogeneity, paving the way for personalized therapeutic strategies and improved patient outcomes.

Keywords: hypergraph neural network, similarity network fusion, Parkinson's disease subtypes, multi-modal data

PACS: 89.75.Hc, 87.19.lx, 89.70.–a

DOI: 10.1088/1674-1056/ae2677 **CSTR:** 32038.14.CPB.ae2677

1. Introduction

Parkinson's disease (PD) is a chronic neurological disorder that primarily affects middle-aged and elderly populations. The primary pathological alteration in PD involves the degeneration of the nigrostriatal dopaminergic pathway, leading to reduced levels of striatal dopamine neurotransmitters. This, in turn, gives rise to a series of motor and non-motor symptoms.[1] These symptoms contribute to a decline in patients' quality of life and impaired cognitive performance, not only affecting their ability to perform daily living activities but also potentially exacerbating the deterioration of motor function and increasing the morbidity and mortality rates among the PD population.[2] With the advent of the aging era, the substantial PD population has become a prominent public health issue. However, the phenotypic heterogeneity observed within the PD population poses significant challenges for clinical treatment and prognosis.[3] For instance, the motor symptoms of PD exhibit considerable diversity, encompassing not only classic manifestations such as tremor, rigidity, and bradykinesia, but also abnormalities in posture and gait. Beyond motor symptoms, PD also includes non-motor symptoms — such as cognitive impairment, depression, and sleep disorders — that significantly impact patients' quality of life.[4] The broad spectrum of motor and non-motor symptoms suggests the existence of distinct disease subtypes within PD, which may respond differently to treatments. Therefore, accurate identification of PD subtypes could enable precise clinical management and personalized interventions, thereby alleviating the social and economic burdens associated with aging.

From a computational perspective, patient subtyping is essentially a clustering problem. Currently, most subtyping schemes for PD are primarily based on single variables. For example, based on clinical presentations, PD can be classified into tremor-dominant, rigidity-dominant, and mixed subtypes.[5] However, in addition to clinical behavioral assessments, other types of data also play a non-negligible role in subtyping Parkinson's patients.[6] For instance, functional magnetic resonance imaging (fMRI) is a non-invasive brain imaging technique used to measure and map brain activity.[7] Markello *et al.* found that MRI data help capture patient subgroups with significant differences in PD pathophysiology.[8] DNA methylation, an important epigenetic modification, primarily involves the addition of methyl ($–CH_3$) groups to certain bases of the DNA molecule.[9] At the molecular level, DNA methylation — a vital epigenetic process involving the attachment of methyl ($–CH_3$) groups to specific DNA bases — offers another dimension. Research employing methylation-specific polymerase chain reaction (MSP) by Lin

†Corresponding author. E-mail: meililu@tute.edu.cn

© 2026 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

et al. indicated that abnormal expression of clock genes in PD may be linked to methylation changes in their promoter regions.[10] This suggests that mtDNA copy number could potentially serve as a biomarker for predicting PD onset and disease progression. These findings indicate that multi-modal data provide insights into PD heterogeneity from different perspectives.[11] Thus, to obtain a more comprehensive and accurate representation of PD subtypes, it is necessary to integrate multiple data sources beyond clinical assessments to derive more biologically meaningful subtypes.[12]

With the rapid development of machine learning in the medical field, disease prediction based on machine learning and multi-modal data has become a current research focus. However, existing studies still have certain limitations, such as underutilization of correlations among multi-modal data and insufficient consideration of inter-patient relationships. Graph models offer a general framework for integrating multi-modal information and discovering relationships among patients. Graph-based methods, particularly graph convolutional networks (GCNs), have been widely applied in various biomedical and disease prediction domains.[14] Recently, to effectively handle multi-modal data and capture high-order data associations, a hypergraph neural network approach has been proposed. Hypergraph neural networks can model high-order data correlations and are easily integrated with multi-modal data, offering greater flexibility in multi-modal data processing. For example, when dealing with multi-modal data, hypergraphs can utilize the data to generate different types of hyperedges, which can then be directly concatenated into a unified hypergraph for downstream tasks. Hypergraph modeling can be divided into two steps: hyperedge group construction and fusion. Current fusion strategies mainly include equivalent fusion and adaptive fusion. Equivalent fusion directly concatenates the association matrices from different hyperedge groups, while adaptive fusion associates each hyperedge group with a learnable parameter to adaptively identify the importance of different hyperedge groups to the overall hypergraph.[14] However, inter-sample connections across different modalities may introduce noise and adversely affect hypergraph performance. Similarity network fusion (SNF) is a method that can integratively process multiple data types, is robust to noise, and is suitable for scenarios with small sample sizes.[15] It has been widely used in cancer subtyping based on multi-omics data. In summary, this paper proposes an improved multi-modal data processing framework based on hypergraph neural networks for PD subtyping.[16] Specifically, we first construct hyperedge groups to capture the high-order associations within each modality. Then, we adopt an SNF-based fusion strategy to effectively integrate the high-order correlations from multiple modalities into a unified hypergraph. Finally, hypergraph convolution methods are employed for learning and classification.

The structure of this paper is as follows: Section 2 introduces the improved hypergraph construction and training methodology. Section 3 presents the analysis of model parameters and performance. Section 4 provides a post-hoc statistical analysis of the identified subtypes. Finally, the discussion and relevant conclusions of this study are provided.

2. Materials and methodology

2.1. Data acquisition and cleansing

The complete dataset was procured through the Parkinson's progression markers initiative (PPMI), a multi-institutional consortium aggregating participant-derived inputs from nine transatlantic research centers, with full open-access availability. Data used in the preparation of this article were obtained on [2025-4-28] from the Parkinson's Progression Markers Initiative (PPMI) database (www.ppmi-info.org/access-data-specimens/download-data), RRID:SCR 006431. For up-to-date information on the study, visit www.ppmi-info.org.

PPMI, a public-private partnership, is funded by the Michael J. Fox Foundation for Parkinson's Research, and funding partners, including a consortium of industry players, non-profit organizations, and private individuals.

Collected between 2012 and 2019, the dataset comprises resting-state fMRI scans, DNA methylation cell composition, and clinical behavioral assessments from 76 participants, totaling 384 scanning sessions. During the MRI data procurement phase, the acquired data were retrieved in ZIP file format, subsequently uncompressed and transformed into the NIfTI format, adhering to the Brain Imaging Data Structure (BIDS) standardization criteria.[17] We employed the fMRIPrep pipeline (version 23.2.1) to subject the blood-oxygen-level-dependent (BOLD) cerebral blood flow data to a series of preprocessing steps.[18] In particular, the preprocessing sequence involved the estimation of head motion, slice timing correction, organization of the postprocessed BOLD data within the original space, alignment of EPI images with T1W images, resampling of BOLD data into the standardized space, estimation of confounding variables, and aggregation of T2*-weighted echo sequences. To delve deeper into the analysis, we leveraged the Harvard–Oxford brain atlas to delineate 48 regions of interest (ROIs) spanning the entire cerebrum[19] and subsequently utilized NiLearn software to derive BOLD time series data from these designated ROIs.[20] When analyzing the DNA methylation cell composition data and patient motion assessment data, we excluded a number of patients ($n = 8$) with incomplete or aberrant data entries. For a minor cohort of

patients ($n = 1$) missing clinical behavior assessment scores, we conducted mean estimations. The resulting dataset incorporated three-modal information from a total of 67 patients. Figure 1 depicts the sex distribution among patients stratified by age group.

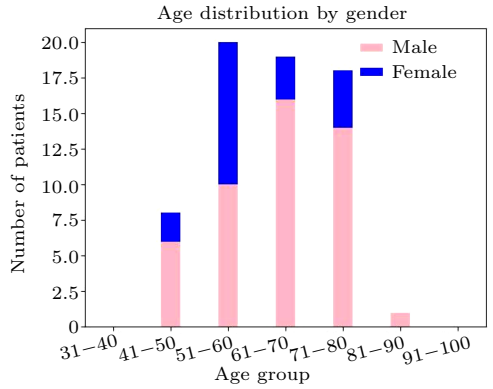


Fig. 1. Age and sex distributions of the patients.

2.2. A PD subtype identification framework based on improved hypergraph neural network

Hypergraph learning, which connects any number of nodes via hyperedges, is inherently capable of modeling complex high-order relationships in multi-modal data. It can uniformly represent heterogeneous information from different modalities (such as visual, textual, and statistical features) as high-order structures, offering significant advantages in cross-modal feature fusion, dynamic structure optimization, computational efficiency, and scalability. Building upon the hypergraph neural network (HGNN),[13] this paper introduces improvements to the hypergraph fusion strategy. The overall framework of the model, as shown in Fig. 2, primarily consists of hypergraph modeling and hypergraph learning modules. Among these, hypergraph modeling comprises hyperedge group generation and hypergraph fusion.

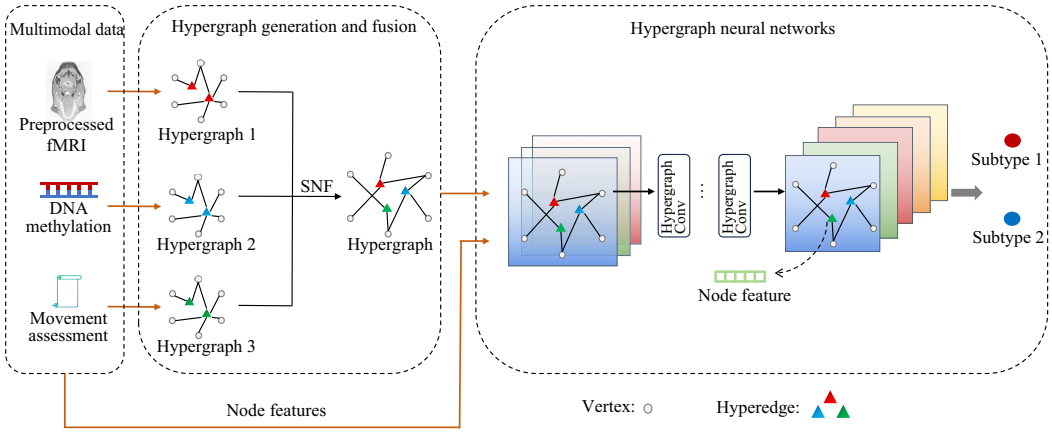


Fig. 2. The overall workflow of this study is structured as follows. First, individual hypergraphs are constructed from three preprocessed data modalities, functional MRI, DNA methylation, and clinical behavioral assessments, using K-nearest neighbors (KNN). These modality-specific hypergraphs are then integrated into a unified, multi-modal hypergraph via SNF. Finally, hyperedge convolution is performed using an HGNN to cluster patients into two distinct Parkinson's disease subtypes. A comprehensive analysis is subsequently conducted to elucidate the differences between these two subtypes.

2.2.1. Construction of hypergraphs

In the hypergraph construction method, the generation of hyperedges in this study is primarily based on similarity relationships between data samples. For each vertex in the dataset, its neighboring relations are determined by calculating the Euclidean distance between its feature vector and those of other vertices. Subsequently, a hyperedge is constructed for each vertex, comprising the vertex itself and its K-nearest neighbors in the feature space. Through this construction approach, a hypergraph structure containing N hyperedges is ultimately formed, where each hyperedge connects $K + 1$ vertices, thereby effectively capturing high-order local similarities within the data. The hypergraph construction process is illustrated in Fig. 3.

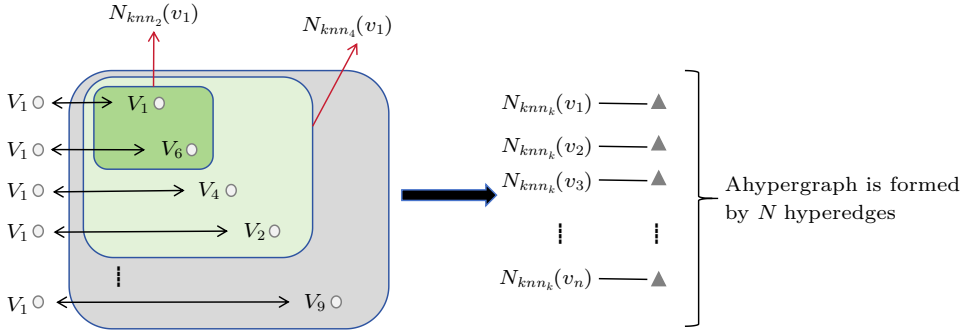


Fig. 3. The construction process of hypergraph, N_{knn} denotes the hyperedge of a node.

2.2.2. Hypergraphs fusion based on SNF

Similarity network fusion (SNF) is a sophisticated approach in which separate similarity networks are established for each exploitable data type within a sample set, followed by seamless integration into a singular network that encapsulates the entirety of the data.[13] For example, in constructing an encompassing perspective of a particular group of Parkinson's disease patients, the SNF method calculates and amalgamates patient similarity networks derived from each distinct data type, thereby harnessing the complementary aspects of the various datasets. This integrated approach substantially outperforms the analysis of single datasets. We utilized SNF to consolidate neuroimaging data, DNA methylation cell composition data, and clinical behavior assessment data from Parkinson's disease patient datasets.

Exemplifying the patient network, we formalize the inter-patient similarity structure as a graph $G(V,E)$. Within this network representation, vertex set V comprises the clinical cohort, whereas edge set E encapsulates dyadic similarity measures between subjects. The edge weight distribution is encoded in an n-dimensional similarity matrix W. Transformation from the distance space to the similarity space employs a scaled exponential kernel function:

$$W(i,j) = \exp\left[-\frac{p^2(x_i,x_j)}{\mu\varepsilon_{i,j}}\right]. \tag{1}$$

Here, the Euclidean distance between patients x_i and x_j is represented by the function $p(x_i,x_j)$. The scaling parameter μ serves as an adjustable hyperparameter, whereas $\varepsilon_{i,j}$ mitigates scaling variance through the following calculation

$$\varepsilon_{i,j} = \frac{\bar{p}(x_i,N_i)+\bar{p}(x_j,N_j)+\bar{p}(x_i,y_j)}{3}, \tag{2}$$

where $\bar{p}(x_i,N_i)$ represents the distance between x_i and its neighborhood points.

For multimodal data integration, the primary procedure entails normalizing all the constituent similarity networks. Standard normalization methods, however, are vulnerable to numerical instability due to diagonal self-similarity entries. The enhanced normalization scheme operates as follows:

$$p(i,j) = \begin{cases} \frac{\mathrm{W}(i,j)}{2\sum_{k\neq i}W(i,k)}, & j\neq i, \\ 1/2, & j=i. \end{cases} \tag{3}$$

We propose that proximity-driven similarity metrics are more reliable than their distance-based counterparts. By implementing graph diffusion processes, similarity measures propagate to nonadjacent nodes, thereby establishing an affinity matrix manifesting enhanced sparsity characteristics:

$$S(i,j) = \begin{cases} \dfrac{W(i,j)}{\sum_{k\in N_i}W(i,k)}, & j\in N_i, \\ 0, & \text{otherwise}. \end{cases} \tag{4}$$

Consequently, matrix P comprehensively encodes global affinity relationships across all patient pairs, whereas matrix S specifically captures localized neighborhood similarities restricted to K-nearest neighbors.

Subsequent similarity network integration proceeds iteratively until convergence. The core iterative procedure of SNF propagates updated similarity matrices across modalities through cross-network diffusion:

$$P_{t+1}^{(1)} = S^{(1)} \times P_t^{(2)} \times \left(S^{(1)}\right)^{\mathrm{T}}, \tag{5}$$

$$P_{t+1}^{(2)} = S^{(2)} \times P_t^{(1)} \times \left(S^{(2)}\right)^{\mathrm{T}}. \tag{6}$$

Among these, the state matrix of the first data type after t iterations is given by $P_{t+1}^{(1)}$, and the similarity matrix corresponding to the second data type is denoted by $P_{t+1}^{(2)}$. Each state matrix update iteration initiates bidirectional cross-diffusion processes. Upon completing t iterations, the fused state matrix is derived via

$$P^{(c)} = \frac{P_t^{(1)}+P_t^{(2)}}{2}. \tag{7}$$

When $m>2$, the following equation is obtained from Eqs. (5) and (6)

$$P^{(v)} = S^{(v)} \times \left(\frac{\sum_{k\neq v}P^{(k)}}{m-1}\right) \times \left(S^{(v)}\right)^{\mathrm{T}}, \quad v=1,2,3,\ldots,m. \tag{8}$$

Every iterative sequence requires renormalizing the updated matrix via the previously stated equation. Algorithm termination triggers when either matrix convergence is achieved or the iteration count attains a preset threshold, typically set at 20 by empirical convention.

2.2.3. Hypergraph learning module

This paper adopts the hypergraph neural network (HGNN) framework for data representation learning. The framework enables modeling of high-order relationships between data through hypergraph structures, making it particularly suitable for multimodal and complex-structured data. The core of the hypergraph neural network architecture is the spectral convolution operation on hypergraphs, which achieves feature propagation and aggregation through the hypergraph Laplacian operator. The hypergraph Laplacian operator, Δ, is defined as:

$$\Delta = I - D_v^{-1/2}HWD_e^{-1}H^{\mathrm{T}}D_v^{-1/2}, \tag{9}$$

where I is the identity matrix, Δ is a positive semidefinite matrix, which can be decomposed to derive the spectral convolution. In the spectral domain, the convolution operation between the signal $x\in R^n$ and the filter g is approximated by Chebyshev polynomials

$$g*x \approx \sum_{k=0}^{K}\theta_k T_k(\tilde{\Delta})x, \tag{10}$$

where T_k is the k-th Chebyshev polynomial, $\tilde{\Delta} = \frac{2}{\lambda_{max}}\Delta - I$ is the scaled Laplace operator, and $\lambda_{\max} \approx 2$ for simplification of calculation, take $K = 1$ for the first-order approximation:

$$g * x \approx \theta_0 x - \theta_1 D_v^{-1/2} H W D_e^{-1} H^{\mathrm{T}} D_v^{-1/2} x. \tag{11}$$

In order to prevent the problem of overfitting, a single parameter θ is set, which is defined as

$$\begin{cases} \theta_1 = \frac{1}{2}\theta, \\ \theta_0 = \frac{1}{2}\theta D_v^{-1/2} H D_e^{-1} H^{\mathrm{T}} D_v^{-1/2}. \end{cases} \tag{12}$$

The convolution operation is further simplified as

$$g * x \approx \theta D_v^{-1/2} H (W + 1) D_e^{-1} H^{\mathrm{T}} D_v^{-1/2} x. \tag{13}$$

In practical implementation, the hyper-edge weight matrix WW is initialized as a unit matrix, indicating that all hyper-edge weights are initially equal. For hypergraphs with feature matrices $X \in R^{n*c}$, the single-layer hyper-edge convolution operation in HGNN is defined as

$$X^{(l+1)} = \sigma(D_v^{-1/2} H W D_e^{-1} H^{\mathrm{T}} D_v^{-1/2} X^{(l)} \Theta^{(l)}. \tag{14}$$

Among them, $X^{(l)}$ is the input feature of the l-th layer, $X^{(0)} = X$; $\Theta^{(l)} \in R^{c_{\mathrm{in}} * c_{\mathrm{out}}}$ as a learnable filter parameter; σ is a nonlinear activation function. This formula realizes the "node-edge-node" feature transformation mechanism, which can effectively capture the higher-order correlations in hypergraphs.

Model parameters: The HGNN model comprises the following key parameters: (i) Hyper-edge weight matrix W: It can be a fixed unit matrix or treated as a learnable parameter. (ii) Convolution layer parameters $\Theta^{(l)}$: They represent kernel weights per layer, with dimensions determined by input and output feature dimensions. (iii) Hypergraph structure parameters: association matrix H and degree matrix D_v, D_e, which are generated by the data preprocessing stage.

3. Model parameters and performance analysis

3.1. Selection of parameters and number of clusters

In the hypergraph construction process, the parameter K determines the number of nearest neighbors for each vertex (*i.e.*, each sample) to form a hyperedge, thereby governing the vertex count per hyperedge. A smaller K value yields more localized connections, resulting in a sparser graph structure, whereas a larger K value generates denser connections, leading to a more interconnected graph. The selection of K significantly impacts model performance: an excessively small K may lead to insufficient information capture and ineffective modeling of global relationships, while an overly large K may introduce substantial noise that adversely affects clustering structure and effectiveness.

To determine the optimal value, we compared results using the Calinski–Harabasz (CH) index and Davies–Bouldin (DB) index.[21] The CH index calculates the ratio between inter-cluster separation and intra-cluster compactness, where higher values indicate better clustering performance. Conversely, the DB index measures the worst-case ratio of intra-cluster to inter-cluster distance, with lower values denoting superior clustering quality. The impact of different K values on clustering evaluation is illustrated in Fig. 4, where subplot B presents the CH and DB index results, demonstrating that the optimal K value is 4.

Subsequently, we investigated the effect of cluster number selection on clustering performance. Subplot (c) in Fig. 3 displays the clustering evaluation results for cluster numbers ranging from 2 to 5. Based on the combined results of the CH and DB indices, the optimal number of clusters was determined to be 2.

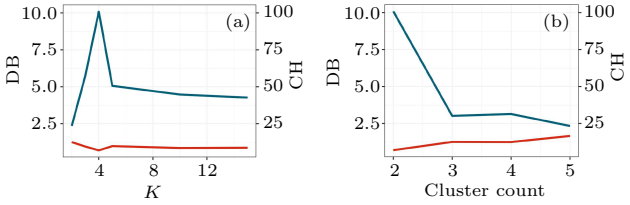


Fig. 4. Clustering performance index for different values of k and number of clusters. Subpanel (a) shows the Kalinsky–Habbash index and Davis–Balding index obtained with different parameter K, and subpanel (b) shows the Kalinsky–Habbash index and Davis–Balding index obtained with different number of clusters. The blue line represents the value of the Kalinsky–Habbash index, and the goal is to find the maximum value of the Kalinsky–Habbash index; The red line represents the Davis–Balding exponent, and the goal is to find the minimum value of the Kalinsky–Habbash exponent.

3.2. Ablation experiment

We individually constructed hypergraphs for each of the three modalities and performed clustering using HGNN, obtaining the corresponding CH index and DB index for each modality. These results were compared with those obtained from the fused multi-modal approach.

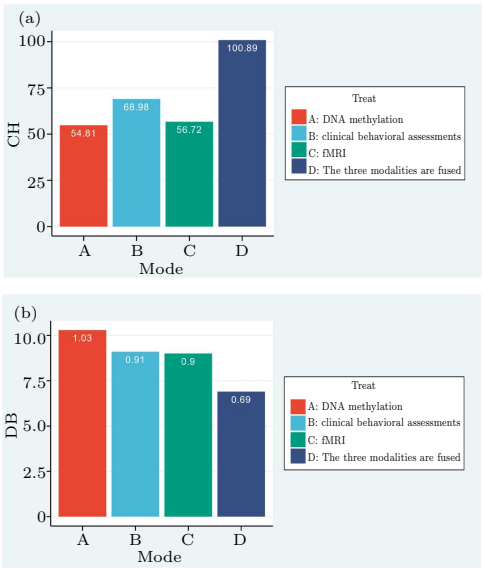


Fig. 5. The clustering effect of fMRI, DNA methylation, and clinical behavior assessment respectively and the clustering effect of the three modes and the fusion. Subfigure (a) represents the value of Kalinsky–Habbash index, and the purpose is to find the maximum value of Kalinsky–Habbash index. Subplot (b) represents the Davis–Balding exponent value, and the goal is to find the minimum value of the Kalinsky–Habbash exponent.

As shown in Fig. 5, the CH and DB indices demonstrate that the performance achieved through multi-modal fusion is consistently superior to that of any single modality.

3.3. Comparison of model performance

For multimodal data fusion, we employed both feature concatenation and SNF methods. For patient clustering, we evaluated four approaches: spectral clustering, graph convolutional network (GCN), hypergraph neural network (HGNN), and general hypergraph neural network (HGNN+). It should be noted that spectral clustering and GCN utilize conventional graphs as input, whereas HGNN and HGNN+ operate on hypergraph structures.

We have selected this series of multimodal data fusion and patient clustering methods with the aim of rigorously exploring and validating the most effective technical pathway for uncovering potential patient subtype structures through a systematic, baseline-to-advanced comparative framework.

At the data fusion level, we employed two parallel strategies: feature concatenation serves as a simple, intuitive, and strong baseline method to evaluate the effectiveness of direct fusion of raw features, whereas similarity network fusion (SNF), by operating at the sample relationship level, more robustly handles the heterogeneity of multimodal data and enhances complementarity, potentially yielding more biologically interpretable results. Comparing these two constructs a complete validation chain from "raw feature fusion" to "sample relationship fusion," aimed at demonstrating the superiority and necessity of the latter.

At the clustering level, we designed a progressive evaluation pipeline: We first employed spectral clustering, a fundamental yet powerful method for handling complex graph structures, as a direct and classical application to the fused network generated by SNF. We then introduced graph convolutional networks (GCNs) for deep representation learning; their advantage lies in simultaneously integrating node features and graph structure, learning deep embeddings that account for patient "social relationships" through neighbor aggregation, with the expectation of achieving clustering results superior to traditional methods. Going a step further, we adopted HGNNs to move beyond the pairwise relationship modeling of GCNs and capture higher-order associations defined by multiple features collectively, which better aligns with the nature of complex co-occurrence relationships in medicine. Finally, we used its enhanced version, HGNN+, to explore the performance ceiling of current hypergraph learning techniques, leveraging its greater flexibility and representational power to unearth deeper insights.

This progression from traditional spectral clustering to deep learning-based GCNs, and further to higher-order modeling with HGNN/HGNN+, systematically evaluates the efficacy of relationship modeling methods at varying complexities. Together, they form a comprehensive research approach that moves from simple to complex, shallow to deep, and pairwise to higher-order relationships, ensuring we identify the optimal technical combination for this specific patient subtyping problem.

Through comparative analysis of CH index and DB index across different methodological combinations, we determined that the integration of SNF with HGNN yielded optimal performance, as demonstrated in Table 1. The abbreviations in the chart are: feature stitching (FC), graph convolutional networks (GCN), spectral clustering (SC).

Table 1. Clustering performance indicators under different models, the best results are shown in bold and the second best results are underlined.

Model	DB	CH
snf+HGNN	<u>0.69</u>	**100.89**
FC+HGNN	1.02	39.05
snf+HGNN+	1.41	27.70
FC+HGNN+	1.26	34.29
snf+GCN	**0.16**	34.13
FC+GCN	1.06	<u>42.50</u>
snf+SC	0.92	2.18

4. Post Hoc statistical analysis of subtypes

4.1. Statistical method

SPSS26.0 software was used for data sorting and statistical analysis. Measurement data were expressed as mean ± standard deviation ($x \pm s$), Shapiro–Wilk (SW) test, and Levene test were used to consider the normality and homogeneity of variance of the data, and the normal distribution and homogeneity of variance were tested by ANOVA. Mann–Whitney U test was used for measurement data that did not conform to normal distribution. A p-value less than 0.05 indicates a statistically significant difference.

4.2. PD subtypes and biological differences

By adjusting the parameter k and the number of clusters, we obtained two different clusters: subtype 1 (48 cases, 71.6%) and subtype 2 (19 cases, 28.4%).

4.2.1. Clinical features

We employed statistical tests to determine the differences between the two identified subtypes. To date, subtyping of PD has largely been based on the observation of clinical features. Relying primarily on a priori assumptions such as age and tremor, and employing a dichotomous classification based on single characteristics, the disease has been categorized into subtypes such as early-onset *versus* late-onset, and tremor-dominant *versus* postural instability/gait difficulty-dominant. The Unified Parkinson's Disease Rating Scale (UPDRS), internationally recognized as the standard assessment tool, com-

prises four distinct parts: MDS–UPDRS part I evaluates non-motor experiences of daily living, MDS–UPDRS part II assesses motor experiences of daily living, MDS–UPDRS part III examines motor function, and MDS–UPDRS part IV evaluates motor complications. Each item is scored on a five-point scale (0–4), where higher scores indicate more severe PD symptoms.[22] Detailed results are presented in Table 2.

No significant differences were observed between the two subtypes in terms of mean age or gender distribution. Regarding the total MDS–UPDRS part I scores, no statistically significant difference was found between the subtypes (1.17 ± 1.31 *versus* 1.11 ± 1.33, $p = 0.9$). However, significant differences emerged in MDS–UPDRS part II total scores (5.69 ± 4.14 *versus* 2.32 ± 3.06, $p < 0.001$), with subtype 1 demonstrating higher scores in 6 out of 13 items, including speech, dressing, hygiene, tremor impact, rising from chairs, and walking. For MDS-UPDRS part III, substantial differences were observed in both total scores (19.65 ± 9.90 *versus* 8.42 ± 8.84, $p < 0.001$) and most individual items. Specifically, subtype 1 showed significantly higher scores than subtype 2 in items related to posture, gait, balance, and tremor, including rising from chairs, gait, freezing of gait, postural stability, posture, postural and kinetic tremors in both upper limbs, rest tremor amplitudes in all extremities, and jaw tremor. No significant difference was detected in MDS–UPDRS part IV total scores between the subtypes (1.38 ± 2.84 *versus* 2.42 ± 3.73, $p = 0.108$). Collectively, these UPDRS assessments indicate that subtype 1 exhibits more severe PD symptoms compared to the milder manifestations characteristic of subtype 2.

Table 2. Demographic differences between subtypes 1 and 2, as well as variations in the unified Parkinson's disease rating scale (UPDRS).

Subtype	Subtype 1 ($n = 48$)	Subtype 2 ($n = 19$)	*P*
Gender			
Male	34 (70.83%)	14 (73.68%)	
Female	14 (29.17%)	5 (29.32%)	
Average Age	60.08 ± 10.98	61.28 ± 10.23	0.9
Non-motor symptoms			
UPDRSI NP1RTOT	1.17 ± 1.31	1.11 ± 1.33	0.368
Motor symptoms			
UPDRSII NP2SPCH	0.46 ± 0.58	0.16 ± 0.50	**0.023**
UPDRSII NP2DRES	0.54 ± 0.68	0.21 ± 0.54	**0.041**
UPDRSII NP2HYGN	0.23 ± 0.42	0.00 ± 0.00	**0.023**
UPDRSII NP2TRMR	0.44 ± 0.54	0.11 ± 0.32	**0.001**
UPDRSII NP2RISE	0.40 ± 0.54	0.05 ± 0.23	**0.015**
UPDRSII NP2WALK	0.06 ± 0.24	0.05 ± 0.23	**0.009**
UPDRII NP2PTOT	5.69 ± 4.14	2.32 ± 3.06	$<$ **0.001**
UPDRSIII NP3RISNG	0.13 ± 0.33	0.00 ± 0.00	0.109
UPDRSIII NP3GAIT	0.42 ± 0.54	0.05 ± 0.23	**0.006**
UPDRSIII NP3FRZGT	0.02 ± 0.14	0.00 ± 0.00	0.529
UPDRSIII NP3PSTBL	0.04 ± 0.20	0.00 ± 0.00	0.37
UPDRSIII NP3POSTR	0.44 ± 0.54	0.11 ± 0.32	**0.015**
UPDRSIII NP3PTRMR	0.27 ± 0.57	0.16 ± 0.37	0.503
UPDRSIII NP3PTRML	0.38 ± 0.61	0.26 ± 0.56	0.427
UPDRSIII NP3KTRMR	0.27 ± 0.45	0.16 ± 0.37	0.332
UPDRSIII NP3KTRML	0.40 ± 0.64	0.16 ± 0.37	0.169
UPDRSIII NP3RTARU	0.52 ± 0.85	0.37 ± 0.83	0.42
UPDRSIII NP3RTALU	0.65 ± 0.96	0.37 ± 0.96	0.205
UPDRSIII NP3RTARL	0.15 ± 0.46	0.11 ± 0.32	0.969
UPDRSIII NP3RTALL	0.13 ± 0.39	0.11 ± 0.46	0.546
UPDRSIII NP3RTALJ	0.10 ± 0.37	0.05 ± 0.23	0.658
UPDRSIII NP3RTCON	1.38 ± 1.30	0.58 ± 1.02	**0.019**
UPDRSIII NP3TOT	19.65 ± 9.90	8.42 ± 8.84	$<$ **0.001**
UPDRSIV NP4TOT	1.38 ± 2.84	2.42 ± 3.73	0.108

4.2.2. Neuroimaging feature

For the fMRI modality, we computed and compared graph metrics of functional connectivity matrices for subtype 1, subtype 2, and healthy controls (HC), as summarized in Table 3. The analysis revealed that subtype 1 demonstrated lower values in clustering coefficient, *k*-core, assortativity, global efficiency, and local efficiency compared to both subtype 2 and HC. In contrast, subtype 1 exhibited higher modularity than the other two groups. The substantially higher *k*-core value observed in subtype 2 indicates generally stronger functional matrix integrity compared to subtype 1. This observation is visually corroborated in Fig. 6, where subtype 1 dis-

plays weaker functional connectivity strength relative to subtype 2.

Table 3. Graphic indices of fMRI for subtype 1, subtype 2, and healthy controls.

	Subtype 1	Subtype 2	Healthy controls
Clustering coefficient	0.720	0.873	0.794
k core	9.729	24.167	16.604
Degree of modulus	0.370	0.120	0.248
Network assortativity	0.187	0.259	0.193
Global efficiency	0.608	0.820	0.737
Local efficiency	0.854	0.935	0.896

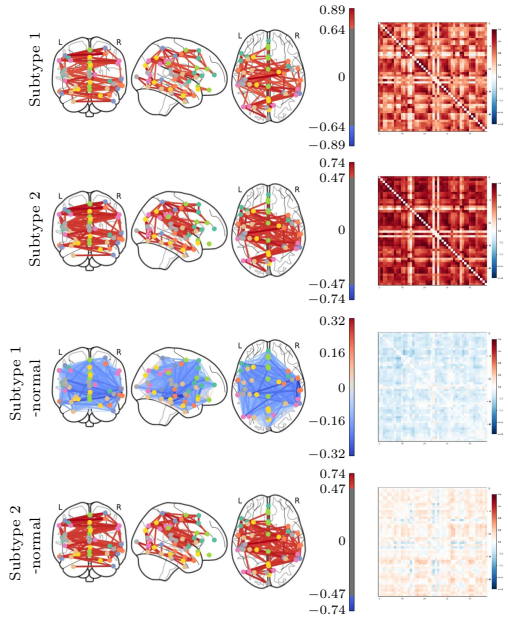


Fig. 6. The average functional connectivity matrix and brain connectivity map of subtypes 1 and 2 and the functional connectivity matrix and brain connectivity map obtained by subtracting the functional connectivity matrix of normal subjects from subtypes 1 and 2. Only the top 5% of connections with strength were plotted in the brain connectivity map. Blue edges indicate decreased connectivity, and red edges indicate increased connectivity.

Furthermore, the differential matrix obtained by subtracting the average functional connectivity matrix of HC from that of subtype 1 appears predominantly blue, suggesting widespread reduction in functional connectivity strength in subtype 1 compared to healthy individuals. Conversely, the differential matrix between subtype 2 and HC shows primarily red coloration, indicating generally enhanced functional connectivity strength in subtype 2 relative to healthy controls.

4.2.3. DNA methylation features

For the DNA methylation-based cell composition data, we employed non-parametric statistical analysis (Mann–Whitney *U* test) to examine six cell-type components for significant differences between subtype 1 and subtype 2.[23] The results are summarized in Table 4.

Table 4. Mann–Whitney *U* test results for the DNA methylation data of subtypes 1 and 2.

	Subtype 1	Subtype 2	*u*	*p*
CD8T	0.083	0.103	438	0.802
CD4T	0.105	0.133	655	**0.006**
NK	0.047	0.050	292	**0.023**
Bcell	0.049	0.064	487	0.666
Mono	0.097	0.109	584	0.075
Gran	0.637	0.562	335	0.092

The analysis revealed statistically significant differences ($p < 0.05$) in CD4 + T cells and natural killer (NK) cells between the two subtypes. Specifically, subtype 1 demonstrated substantially lower mean values in CD4 + T cells compared to subtype 2, while showing marginally reduced mean values in NK cells. Additionally, violin plots were utilized to visually represent the distribution differences in all six methylation-based cell composition metrics between the two subtypes (Fig. 7). These plots clearly illustrate distinct distribution patterns across all examined cell-type components between the identified subtypes, with particular emphasis on the differentially distributed features.

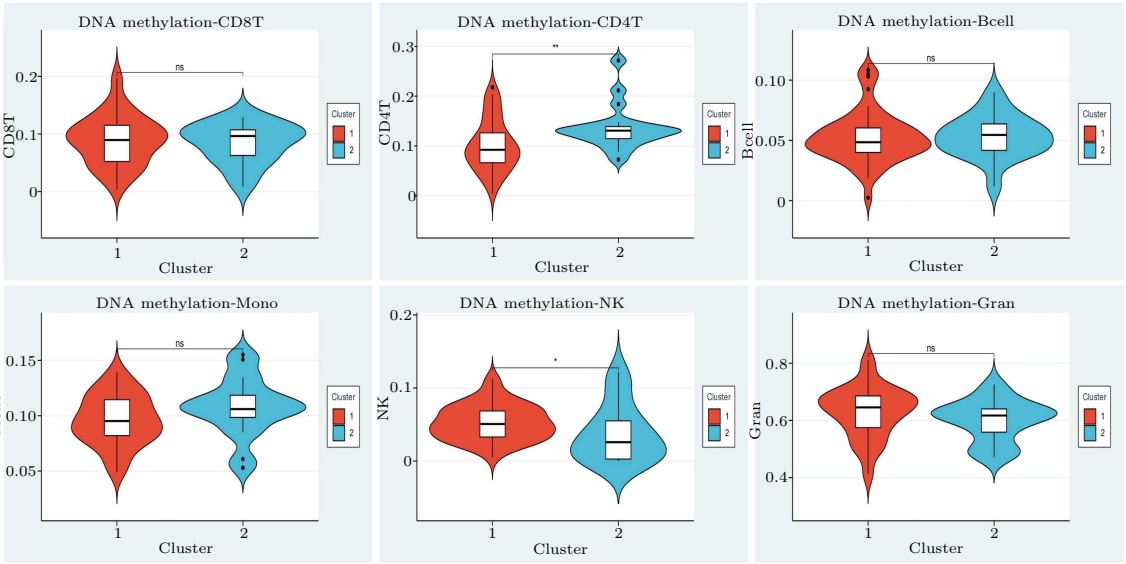


Fig. 7. Violin plots showing kernel density estimates of the distribution of DNA methylation features, the more concentrated the dataset, the fatter the plot. The black horizontal line in the box represents the median of each data item, and the upper and lower edges in the white box represent the upper and lower quartiles in the data set.

5. Discussion

In this study, we constructed hypergraphs from three modalities — MRI, DNA methylation, and clinical behavioral assessments — of Parkinson's disease (PD) patients. These modality-specific hypergraphs were integrated into a unified multi-modal hypergraph using similarity network fusion (SNF), and patient clustering was performed via a hypergraph neural network (HGNN). By evaluating the Calinski–Harabasz index and Davies–Bouldin index across different settings, we determined the optimal hyperparameters: a fusion neighborhood size K of 20, a hyperedge neighborhood size K of 4, and a cluster number of 2.

This process identified two distinct PD subtypes: Subtype 1, comprising 48 patients with severe symptoms, and Subtype 2, comprising 19 patients with mild symptoms. We further compared clustering performance between individual modalities and their fusion, demonstrating that multi-modal integration consistently outperformed any single modality. Among various models evaluated — including HGNN, HGNN+, GCN, and SNF — the combination of SNF and HGNN yielded significantly superior clustering results.

To better characterize the two subtypes, differences across the three modalities were analyzed. In clinical behavioral assessments using the UPDRS scale, significant distinctions were observed across all four subscales, particularly in MDS–UPDRSIII, where subtype 1 scored markedly higher than subtype 2, indicating more severe motor impairment. Analysis of fMRI data revealed that subtype 1 exhibited weaker functional connectivity compared to subtype 2, as reflected in graph-theoretic measures such as clustering coefficient, k-core, modularity, assortativity, global efficiency, and local efficiency. Furthermore, when each subtype was compared to healthy controls, subtype 1 showed broadly reduced functional connectivity strength, whereas subtype 2 displayed generally elevated connectivity — a finding with potential diagnostic relevance for PD subtyping using MRI data.

In DNA methylation-based cell composition analysis, non-parametric statistical testing (Mann–Whitney U test) and violin plots revealed distributional differences between subtypes across all six cell types. Significant differences ($p < 0.05$) were detected in CD4 + T cells and NK cells. However, these differences were not pronounced enough to robustly support diagnostic discrimination between the two subtypes based solely on CD4 + T or NK cell values.

This study has several notable limitations. First, the clinical assessment relied solely on the UPDRS scale and failed to incorporate other relevant clinical information. Furthermore, the analysis was entirely data-driven and did not integrate clinical expertise, whereas combining both approaches could improve the accuracy of subtype classification. Second, and one of the primary limitations of this study, is the relatively small sample size ($n = 67$). This directly constrains the statistical power of the model, potentially failing to identify true but subtle group differences. Additionally, data-driven models with small sample sizes are more susceptible to random noise in the data, increasing the risk of overfitting and potentially rendering the identified subtype structures unstable. As a result, the current findings should be regarded as exploratory. Future work must validate this classification framework in larger, independent, prospective cohorts to assess its generalizability and clinical applicability.

6. Conclusion

In this study, we proposed a refined hypergraph neural network approach that integrates three modalities — MRI, DNA methylation, and clinical behavioral assessments — from Parkinson's disease patients. The framework constructs a unified hypergraph capturing high-order intermodal relationships and applies hypergraph learning via HGNN to cluster patients, thereby identifying two clinically distinct PD subtypes. Post-hoc statistical analyses demonstrated pronounced disparities between the two subtypes across clinical behavior, DNA methylation patterns, and fMRI-based functional connectivity features. For instance, subtype 1 exhibited more severe motor and non-motor impairments, whereas subtype 2 displayed milder symptoms. Notably, from a functional connectivity perspective, subtype 1 showed overall weaker connectivity compared to healthy controls, while subtype 2 exhibited stronger connectivity. Moreover, the significantly higher k-core values observed in subtype 2 suggest a more densely interconnected functional network architecture. These findings provide novel insights for leveraging MRI data to assist in the stratification of Parkinson's disease subtypes.

References

[1] Venkateshwarla R R 2021 *IP Indian J. Neurosci.* **7** 148
[2] Kandharkar S S, Kashid N A, Kanase S B and Mane D 2021 undefined
[3] Heimrich K G, Schönenberg A, Santos-García D, Mir P and Prell T 2023 *J. Clin. Med.* **12** 2573
[4] Deuschl G and Krack P 2020 *Lancet Neurol.* **19** 472
[5] Storey K, Foster E, Kalishman S and Doty T 2021 *Arch. Phys. Med. Rehabil.* **102** e37
[6] Pérez-Lloret S, Chevalier G, Bordet S, Barbar H, Capani F, Udovin L and Otero-Losada M 2023 undefined
[7] Wolters A F, Michielse S, Kuijf M L, Defebvre L, Lopes R, Dujardin K and Leentjens A F G 2022 *Parkinsonism Relat. Disord.* **105** 32
[8] Markello R D, Shafiei G, Tremblay C, Postuma R B, Dagher A and Mišić B 2021 *npj Parkinsons Dis.* **7** 6
[9] Yakovenko E V, Fedotova E Y and Illarioshkin S N 2020 *Ann. Clin. Exp. Neurol.* **14** 56
[10] Lin Q, Ding H, Zhang Z, Gu Z, Ma J, Chen L, Chan P and Cai Y 2012 *Neurosci. Lett.* **507** 147
[11] Shah A, Nguyen T A K, Peterman K, Khawaldeh S, Debove I, Shah S A, Torrecillos F, Tan H, Pogosyan A, Lachenmayer M L, Michelis J P, Brown P, Pollo C, Krack P, Nowacki A and Tinkhauser G 2023 *Neuromodulation* **26** 320

[12] Acosta J, Falcone G J, Rajpurkar P and Topol E J 2022 *Nat. Med.* **28** 1773
[13] Feng Y, You H, Zhang Z, Ji R and Gao Y 2019 *Proc. AAAI Conf. Artif. Intell.* **33** 3558
[14] Gao Y, Feng Y, Ji S and Li R 2022 *IEEE Trans. Pattern Anal. Mach. Intell.* **45** 3181
[15] Wang B, Mezlini A M, Demir F, Fiume M, Tu Z, Brudno M, Haibe-Kains B and Goldenberg A 2014 *Nat. Methods* **11** 333
[16] Yang K 2019 *Pure Math.* **9** 1115
[17] Vanden Bulcke C, Wynen M, Detobel J, La Rosa F, Absinta M, Dricot L, Macq B, Cuadra M B and Maggi P 2022 *Neuroimage Clin.* **36** 103252
[18] Estéban O, Markiewicz C, Blair R, Moodie C, Isik A I, Erramuzpe A, Kent J D, Goncalves M, DuPré E, Snyder M, Oya H, Ghosh S, Wright J, Durnez J, Poldrack R A and Gorgolewski K J 2018 bioRxiv
[19] Guimaraes A A R, Abdou T and Başar A 2019 undefined
[20] Estéban O, Gorgolewski K J, Blair R, Berleant S L, Moodie C and Poldrack R A 2017 *F1000Res.* **6** 1498
[21] Yue G, Deng A, Qu Y, Cui H and Wang X 2023 *Inf. Sci.* **633** 182
[22] Wang X J, Gu H G, Jia Y B, *et al.* 2023 *Chin. Phys. B* **32** 090502
[23] Ni H J, Song Z J, Qin J L, *et al.* 2025 *Chin. Phys. B* **34** 080504