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.

# A Tensor Network Framework for Interpretable Graph Analysis of Brain Networks

**Domenico Pomarico**
Università degli Studi di Bari Aldo Moro
**Giuseppe Magnifico**
giuseppe.magnifico@uniba.it
Università degli Studi di Bari Aldo Moro
**Alessandro Grecucci**
Università degli Studi di Bari Aldo Moro
**Loredana Bellantuono**
Istituto Nazionale di Fisica Nucleare (INFN)
**Jesus M. Cortes**
Biocruces-Bizkaia Health Research Institute
**Marianna La Rocca**
Università degli Studi di Bari Aldo Moro
**Alfonso Monaco**
Università degli Studi di Bari Aldo Moro
**Marlis Ontivero-Ortega**
Università degli Studi di Bari Aldo Moro
**Alessandro Scarano**
University of Trento
**Massimo Stella**
University of Trento
**Roberto Bellotti**
Università degli Studi di Bari Aldo Moro
**Sebastiano Stramaglia**
Università degli Studi di Bari Aldo Moro
**Nicola Amoroso**
Istituto Nazionale di Fisica Nucleare (INFN)

---

**Article**

**Keywords:** Tensor Network Machine Learning, Brain Disorder Classification, Feature Interaction Networks

**Posted Date:** September 9th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-10821560/v1

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

**Additional Declarations:** No competing interests reported.

# A Tensor Network Framework for Interpretable Graph Analysis of Brain Networks

Domenico Pomarico[1,2], Giuseppe Magnifico[1,2], Alessandro Grecucci[3], Loredana Bellantuono[2,4], Jesus M. Cortes[5,6,7,8], Marianna La Rocca[1,2], Alfonso Monaco[1,2], Marlis Ontivero-Ortega[1,2], Alessandro Scarano[9], Massimo Stella[9], Roberto Bellotti[1,2], Sebastiano Stramaglia[1,2], Nicola Amoroso[2,10]

[1]Dipartimento di Fisica, Università degli Studi di Bari Aldo Moro, Bari, 70126, Italy.
[2]Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Bari, Bari, 70125, Italy.
[3]Dipartimento di Scienze della Formazione, Psicologia, Comunicazione (For.Psi.Com.), Università degli Studi di Bari Aldo Moro, Bari, 70122, Italy.
[4]Dipartimento di Biomedicina Traslazionale e Neuroscienze (DiBraiN), Università degli Studi di Bari Aldo Moro, Bari, 70124, Italy.
[5]Biocruces-Bizkaia Health Research Institute, Barakaldo, 48903, Spain.
[6]Biomedical Research Doctorate Program, University of the Basque Country, Leioa, 48940, Spain.
[7]Department of Cell Biology and Histology, University of the Basque Country, Leioa, 48940, Spain.
[8]IKERBASQUE Basque Foundation for Science, Bilbao, 48009, Spain.
[9]Department of Psychology and Cognitive Science, University of Trento, Trento, Italy.
[10]Dipartimento di Farmacia - Scienze del Farmaco, Università degli Studi di Bari Aldo Moro, Bari, 70125, Italy.

Contributing authors: domenico.pomarico@ba.infn.it; giuseppe.magnifico@uniba.it; alessandro.grecucci@uniba.it; loredana.bellantuono@uniba.it; jesus.m.cortes@gmail.com; marianna.larocca@uniba.it; alfonso.monaco@ba.infn.it; marlis.ontiveroortega@uniba.it; scaranoalex@gmail.com;

massimo.stella-1@unitn.it; roberto.bellotti@uniba.it;
sebastiano.stramaglia@ba.infn.it; nicola.amoroso@uniba.it;

**Abstract**

Identifying robust neurobiological signatures of brain disorders requires machine learning approaches that combine predictive performance with interpretable representations of feature interactions. Here we introduce a quantum-inspired framework based on tensor network machine learning that learns distributed representations of gray-matter features encoded in a Matrix Product State representation, a variational ansatz originally developed for quantum many-body systems. The trained model is then used not only as a classifier but to extract quantum connected correlations between features, which encode higher-order feature interactions and which we map onto a weighted graph. This construction allows us to track, within a single representation, both: (i) the global spectral properties of the network (capturing collective learning dynamics), and (ii) node-level centrality measures (providing interpretable signatures of individual brain regions). Using repeated train-test sampling schemes, we analyze two classification tasks on structural MRI data as examples of complex brain disorders: healthy controls versus schizophrenia and versus bipolar disorder. Node-level analysis identifies a stable set of gray-matter features, most prominently Heschl gyrus, insular cortex, and frontal regions, that act as hubs across multiple centrality measures and across resamplings. These centralities display lower variability across resamplings than Shapley values, supporting the interpretive value of the network representation. The bipolar feature set emerges as a subset of the schizophrenia one, consistent with the hierarchically organized neuroanatomical alterations reported in neuroimaging studies and offering a network-based characterization of this hierarchy.

**Keywords:** Tensor Network Machine Learning, Brain Disorder Classification, Feature Interaction Networks

# 1 Introduction

Recent advances in quantum computing have opened a stimulating frontier at the intersection of quantum technologies and machine learning [1–12]. Although current quantum processors remain limited by noise, decoherence, and restricted circuit depth, the conceptual framework of quantum information continues to inspire powerful computational models that can be implemented efficiently on classical hardware [13–21]. Among these, tensor networks, originally devised to study the properties of quantum many-body systems, such as quantum correlations and entanglement, have emerged as a versatile class of quantum-inspired machine learning tools. Their ability to encode structured, higher-order correlations within a single global model, combined with their interpretable parameterizations, has positioned them as a complementary approach to conventional deep neural networks [22–36]. Quantum-inspired approaches based on tensor networks have shown encouraging results in complex tasks such as cancer

classification, biomarker discovery, and integrative analysis of gene-expression networks [37–39]. These properties make tensor network models particularly attractive for neuroscience, where data are high-dimensional, noisy, and characterized by complex dependencies not easily captured by conventional statistical models [2, 3, 40–44].

A central pillar of contemporary computational neuroscience is the modeling of brain organization through correlation-based measures, which quantify how activity or morphology co-varies across regions [45–48]. These approaches stem from the principle that the brain operates as a distributed system in which functional and structural components express coordinated fluctuations rather than acting in isolation [49]. In functional Magnetic Resonance Imaging (fMRI), correlations between blood-oxygenation-level-dependent (BOLD) time series form the basis of functional connectivity estimates, allowing the identification of coherent networks that support perception, cognition, and affective regulation [50]. Similarly, in structural neuroimaging, covariance patterns, derived from correlations in gray- or white-matter morphology across individuals, are interpreted as proxies of shared developmental, genetic, or experience-dependent factors shaping brain architecture [51, 52]. This correlation-based perspective has proven particularly influential because it reveals network-level dependencies that cannot be captured by local voxelwise comparisons alone, providing a more holistic depiction of interregional coordination underlying both healthy and pathological brain conditions [53, 54]. However, most covariance-based approaches remain fundamentally pairwise, potentially overlooking higher-order dependencies distributed across multiple interacting regions.

Beyond mapping intrinsic connectivity, correlation based models have become widely used tools for characterizing disease-related alterations in network architecture. Changes in covariance patterns may reflect deviations in synchronized development, neurodegenerative processes, or dysregulation of large-scale circuits implicated in psychiatric and neurological conditions [52, 55]. These network-level anomalies may be more sensitive and interpretable than absolute volumetric changes, as they capture the reorganization of relationships among regions rather than their isolated deterioration [56]. As such, correlation-driven modeling has become a key conceptual bridge between traditional neuroanatomy and computational network science, informing hypotheses about disconnectivity, compensatory pathways, and circuit-level vulnerability. When combined with machine-learning approaches, and especially with models capable of representing long-range dependencies, such as tensor networks, correlation-based measures can be used as structured inputs through which learning algorithms can infer latent organization principles, identify disease-specific signatures, and generate hypotheses about large-scale organizational principles underlying brain dysfunction.

In this work, we combine these two perspectives. We train a tensor network, specifically a Matrix Product State (MPS) classifier, on structural MRI data, using it as both a predictive model and a probe of feature-level correlation structure. More in detail, we extract the quantum connected correlations of the trained MPS to define a weighted graph whose global spectral properties and node-level centralities provide, respectively, a dynamical and an interpretable view of the learning process. We apply this framework to two binary classification tasks: healthy controls versus schizophrenia, and healthy controls versus bipolar disorder. We find that a stable set

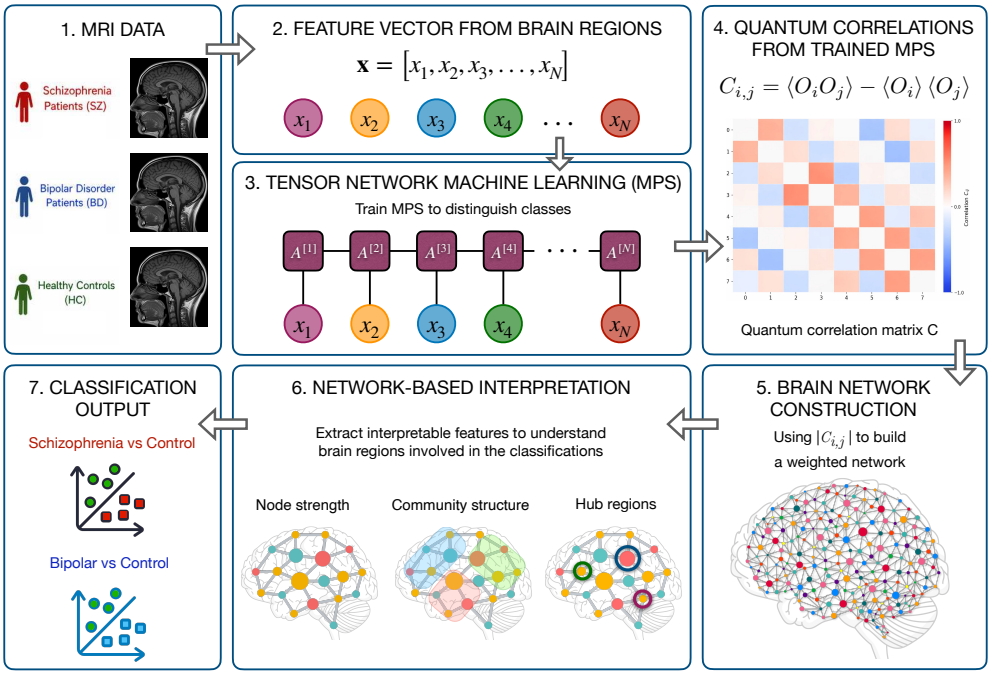



**Fig. 1** Schematic overview of the quantum-inspired tensor network framework for brain disorder classifications. (1) MRI data are considered for three groups: schizophrenia patients, bipolar disorder patients, and healthy controls. (2) Gray-matter concentration values extracted according to the AAL Atlas define a feature vector $\mathbf{x}$ for each subject. (3) Features are encoded into an MPS classifier, which is trained to distinguish diagnostic classes. (4) Quantum connected correlations are extracted from the trained MPS, yielding a pairwise correlation matrix. (5) The absolute values of the correlations are used as edge weights to construct a weighted network whose nodes are the gray-matter regions. (6) Graph-theoretic quantities are computed from the network to provide interpretable, region-level signatures of the learned representations. (7) The framework produces binary classification outputs for two diagnostic tasks: schizophrenia versus healthy controls and bipolar disorder versus healthy controls.

of gray-matter features, most prominently Heschl gyrus, insular cortex, and frontal regions, emerges consistently across resamplings. Node-level centralities show substantially lower variability across resamplings than Shapley values, indicating that the network representation captures structural features of the data more robustly than model-based explanations. The set of discriminative features associated with bipolar disorder is found to be a subset of those features identified in schizophrenia, consistent with the hierarchically organized neuroanatomical alterations reported in neuroimaging studies, and in particular with previous structural covariance analyses on the same dataset showing broader network alterations in schizophrenia than in bipolar disorder [51]. The overall pipeline of the work is summarized in Fig. 1.

The article is organized as follows. In Sec. 2, we introduce the structural MRI dataset and the tensor network learning framework used to encode gray-matter features as quantum-inspired representations, including our classification performance

compared with a classical benchmark. In Sec. 3, we present the construction of the quantum pairwise-correlation graphs that form the basis for our graph-theoretic analysis. In particular, we detail how global spectral metrics and node-level centrality measures extract interpretable neural signatures from the correlation structures, revealing stable hubs and task-dependent reorganization patterns associated with the classification of schizophrenia and bipolar disorder relative to controls. In Sec. 4, we discuss the neuroscientific implications of these findings, emphasizing how quantum-inspired machine learning combined with correlation-based graph modeling provides a principled, interpretable route for uncovering robust structural markers of brain disorders. Finally, in Sec. 5, we summarize our findings.

# 2 Materials and Methods

## 2.1 Structural MRI dataset

The dataset comprised 383 right-handed adults (ages $18-65$), split into three demographically matched groups: schizophrenia ($n = 128$), bipolar I disorder ($n = 128$), and healthy controls ($n = 127$). Patients were recruited at Spanish clinical centers (Benito Menni CASM; Mare de Déu de la Mercè; Hospital Clínic de Barcelona). Controls were drawn from hospital non-medical staff and the community and had no history of psychiatric illness or psychotropic treatment. Matching across groups was based on age, sex, and premorbid IQ (estimated via the Word Accentuation Test). Clinical characterization included PANSS for schizophrenia, and YMRS and HDRS for bipolar disorder; at scan time bipolar disorder participants included euthymic ($n = 77$), manic ($n = 28$), and depressed ($n = 23$) individuals. Common exclusion criteria were neurological disease, prior brain trauma, and alcohol/substance abuse within 12 months of imaging. All participants provided written informed consent under approval of the Comité de Ética de Investigación Clínica de las Hermanas Hospitalarias [51].

High-resolution structural MRI was acquired on a $1.5T$ GE Signa system (General Electric, Milwaukee, WI). The T1-weighted protocol used 180 axial slices, 1 $mm$ slice thickness (no gap), $512\times512$ matrix, and $0.5\times0.5\times1$ $mm^3$ voxel size, with $TE = 4\,ms$, $TR = 2000\,ms$, and flip angle $15^\circ$.

Structural images underwent a standard pipeline: unified tissue segmentation in SPM12 to derive gray-matter partial-volume maps; brain extraction per BET; and spatial normalization to the MNI152 (2 $mm$) template using FSL registration tools. Deformation fields from normalization were then applied to generate normalized gray-matter images, which were down-sampled to $4 \times 4 \times 4$ $mm^3$ to reduce computational burden while preserving the large-scale spatial covariance structure relevant for network-level analyses. Grey matter concentration for each region was extracted according to the AAL Atlas. This regional representation was used to characterize distributed structural variation across subjects rather than localized voxel-wise abnormalities.

## 2.2 Tensor network machine learning

We employ Matrix Product States (MPS), a tensor network ansatz that provides a compact and controllable parameterization of high-dimensional model tensors while preserving sufficient expressive power for supervised classification tasks. An MPS represents a tensor with many indices as a contracted sequence of three-index tensors arranged along a one-dimensional chain. The bond dimension $\chi$ of the MPS controls the maximum allowed quantum correlations and entanglement across the chain and therefore the representational capacity [57–60].

Given an input sample with $N$ features, we embed the data into a product state of $N$ qubits arranged on a one-dimensional lattice. Each qubit encodes a single feature through a fixed local embedding $|\phi(x^{(j)})\rangle = \big(\cos(x^{(j)}),\ \sin(x^{(j)})\big)^{\intercal} \in \mathbb{R}^2, j = 1, \ldots, N$, which is the standard sinusoidal feature map of MPS-based classification problems [22]. For a binary classification task with class index $\ell \in \{0, 1\}$, the MPS defines the predictor as the contraction of the weight tensor $W^{\ell}$ with the product feature state

$$f_W^{\ell}(x) = \sum_{s_1,\ldots,s_N} W^{\ell}_{s_1,\ldots,s_N}\ |\phi(x^{(1)})_{s_1}\rangle \otimes \cdots \otimes |\phi(x^{(N)})_{s_N}\rangle\,, \tag{1}$$

and is trained by minimizing the mean-squared-error cost

$$\mathcal{C}(W) = \frac{1}{2N_T} \sum_{\omega=1}^{N_T} \sum_{\ell} \big(f_W^{\ell}(x_\omega) - y_\omega^{\ell}\big)^2\,, \tag{2}$$

where $N_T$ is the number of training samples and $y_\omega^{\ell}$ denotes the one-hot encoded label. The predicted class corresponds to the index $\ell$ maximizing the magnitude of $f_W^{\ell}(x)$ [22, 35].

Training proceeds via gradient descent using a two-site update scheme: at each step, two adjacent MPS tensors are jointly optimized and then split via a truncated singular-value decomposition that enforces the bond-dimension constraint and retains the most informative components of the model [22–36, 61]. The two-site optimization moves back and forth along the one-dimensional lattice; a full pass that returns the optimization window to its initial position defines a *sweep*, conceptually analogous to a training epoch. The resulting training dynamics can exhibit a dynamical entanglement transition that shares qualitative features with the grokking phenomenon recently reported in tensor network classifiers [62, 63], and that we analyze in Appendix A.

To track the learning dynamics, we compute several quantum observables on the MPS: the reduced density matrix in label space $\varrho^{(\ell)}$, the local magnetizations $\langle \sigma_Z^{k,i} \rangle$ for feature $i$ and label subspace $k \in \{0, 1\}$, and the pairwise correlation functions $\langle \sigma_Z^{k,i} \sigma_Z^{k',j} \rangle$. From these, we define the connected correlation measure

$$C_{i,j}^{(k,k')} = \frac{\langle \sigma_Z^{k,i} \sigma_Z^{k',j} \rangle - \langle \sigma_Z^{k,i} \rangle \langle \sigma_Z^{k',j} \rangle}{\sqrt{\varrho_{k,k}^{(\ell)} \varrho_{k',k'}^{(\ell)}}}, \qquad k, k' \in \{0, 1\}, \tag{3}$$

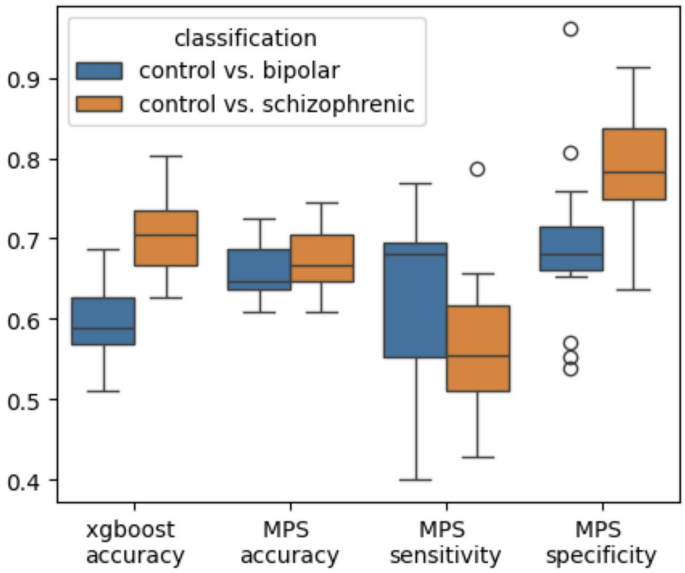



**Fig. 2** Boxplot comparison of classifier performance across the two diagnostic tasks. The distribution of performance metrics are obtained for an XGBoost classifier and the Matrix Product State (MPS) machine-learning model when distinguishing control vs. bipolar subjects (blue) and control vs. schizophrenic subjects (orange). Shown are boxplots for XGBoost accuracy, MPS accuracy, MPS sensitivity, and MPS specificity.

which leads to the identification of feature pairs that contribute most prominently to the classification process.

Finally, we characterize the entanglement structure along the lattice by evaluating the class-resolved von Neumann entropies across the generic bond $(i, i+1)$, that is the quantity $S^{\ell}(i) = -\mathrm{Tr}(\varrho^{(\ell)}_{1:i} \ln \varrho^{(\ell)}_{1:i})$, where $\varrho^{(\ell)}_{1:i}$ is the reduced density matrix of the MPS partition comprising sites $1, 2, ..., i$ [64, 65]. Monitoring these quantities across training reveals the transition from the initial volume-law entanglement of the random MPS towards a sub-volume scaling regime, accompanied by spectral reorganization in the reduced density matrices that reflects emerging feature selectivity [62, 63]. We define the transition sweep as the sweep at which these internal observables, i.e. magnetization patterns, reduced density matrix coherence, entanglement entropy, undergo simultaneous sharp reorganization.

## 2.3 Classification performance

We compared the MPS classifier with a tree-ensemble benchmark (XGBoost), chosen as a solid, hyperparameter-robust baseline, on two binary learning tasks: controls vs. bipolar disorder and controls vs. schizophrenia. Both tasks were derived from $N = 30$ gray-matter features selected via Boruta ranking [66], and performance was assessed over 15 random train/test resamples (test proportion 0.20). Fig. 2 reports the resulting distributions of accuracy for both methods, together with sensitivity and specificity for MPS, evaluated at the best performing sweep during training.

For XGBoost, the control vs. schizophrenia task yielded clearly higher accuracy than control vs. bipolar, with a median around 0.71 (interquartile range IQR $\approx [0.67, 0.74]$) versus 0.59 (IQR $\approx [0.57, 0.63]$). MPS exhibited more balanced performance across tasks: for bipolar, the median accuracy was 0.65 (IQR $\approx [0.64, 0.69]$), exceeding XGBoost by roughly six percentage points; for schizophrenia, MPS reached 0.67 (IQR $\approx [0.65, 0.71]$), slightly below XGBoost. The bipolar classification task exhibits a statistically significant difference in accuracy distributions ($p \approx 0.001 < 0.01$), whereas the schizophrenia classification task does not show statistically significant discrimination ($p \approx 0.1 > 0.01$). Two patterns emerge: (i) the control vs. schizophrenia separation is intrinsically easier than control vs. bipolar for the given features, as both classifiers reach higher accuracy on the former; and (ii) MPS captures structure that benefits the bipolar task in particular, where it improves on the tree-ensemble baseline.

For MPS, sensitivity (recall of the patient class) was higher in the bipolar task (median 0.68, IQR $\approx [0.55, 0.70]$) than in the schizophrenia task (median 0.56, IQR $\approx [0.51, 0.62]$, with occasional higher outliers). Specificity displayed the opposite asymmetry, favoring the schizophrenia task (median 0.78, IQR $\approx [0.75, 0.84]$, with high outliers around 0.90) over the bipolar one (median 0.68, IQR $\approx [0.66, 0.71]$, with a few low outliers near 0.55).

This pattern suggests that MPS tends to reduce false positives on the schizophrenia task (strong specificity) at the expense of higher false negatives (lower sensitivity), whereas performance on the bipolar task is more balanced between sensitivity and specificity. Considering the interquartile ranges, accuracy distributions for both models are comparatively tight, indicating stable central performance under data splits. Sensitivity displays greater dispersion, particularly for MPS, implying that fold composition materially affects recall of the patient class. Overall, MPS achieves accuracy that improves over the XGBoost baseline on the harder bipolar task and matches it on the schizophrenia task.

A central motivation for using quantum-inspired tensor network architectures is their capability to yield interpretable representations of feature interactions. As discussed in Appendix B, the single-site entanglement entropy displays a broadly uniform pattern across features in both tasks, with consistently large values and limited variability across runs. Its uniformity indicates that the stable features are all significantly integrated into the collective quantum correlations of the classification structure, providing a quantum-inspired validation of the feature selection: the stable selected regions are not only statistically discriminative but also exhibit a high level of quantum correlations in the global representations. This stability supports the use of quantum-inspired observables as reliable measures of feature integration in the learned representation, and motivates the network-based analysis presented in the following sections, where we move from single-site quantum properties to pairwise quantum correlations between features, constructing a weighted brain network whose topology encodes the learned classification structure.

# 3 Quantum pairwise correlation graph

The connected correlation measure defined in Eq. (3) assigns, to each pair of features $(i, j)$ and each pair of label indices $(k, k')$, a real-valued quantity $C_{i,j}^{(k,k')}$ computed from the trained MPS. We interpret the absolute value of these quantum correlations $\left|C_{i,j}^{(k,k')}\right|$ as edge weights of a weighted graph whose nodes are the gray-matter features; these weights represent learned interaction strengths within the classifier rather than direct anatomical or functional connectivity. (Fig. 1, steps 4–5). This defines three graphs of interest: two diagonal ones, $\left|C^{(0,0)}\right|$ and $\left|C^{(1,1)}\right|$, capturing intra-class correlations for the control and pathological subspaces respectively, which we refer to as the control mask and the pathology-associated mask; and the off-diagonal $\left|C^{(0,1)}\right|$, capturing cross-class correlations, which we refer to as the confounding mask. The three graphs evolve during training and encode complementary aspects of the learned representation.

In the following sections, we analyze the evolution of the three graphs during training through global spectral observables, which capture collective reorganization of the correlation structure, and through node-level centrality measures at the transition sweep, which identify the regions that act as hubs of the learned representation.

## 3.1 Global spectral observables

Once we have established that the MPS classifier captures discriminative structure in both tasks, we can examine how this structure emerges during training in the quantum pairwise correlation network. We monitored a set of global graph-theoretic and spectral observables derived from the correlation matrices $C^{(0,0)}$, $C^{(1,1)}$, $C^{(0,1)}$. Further details on the definition of the global graph-theoretic and spectral observables are provided in Appendix C.

The temporal evolution of these observables across training sweeps reveals clear signatures of the dynamical transition, an abrupt reorganization of the internal correlation structure, coinciding with a specific transition sweep for each task, where the model starts to generalize properly. This transition is clearly visible in Fig. 3, which shows the four metrics we analyze: the largest eigenvalue of the adjacency matrix, the spectral entropy, the algebraic connectivity of the Laplacian, and the Laplacian spectral ratio as a function of the sweep number. Further details on this dynamics are discussed in Appendix A.

Together, these quantities probe complementary aspects of the network topology, including global coherence, spectral complexity, connectivity robustness, and synchronization capacity. We refer to global spectral quantities as topological invariants of the weighted graph in the sense that they are invariant under node relabeling and depend only on the underlying connectivity structure.

The largest eigenvalue of the adjacency matrix provides a measure of the global strength and coherence of correlations within the network [67]. As shown in panels (a-c), this quantity exhibits a marked increase during training sweeps for both classification tasks, signaling the progressive emergence of a dominant correlation mode within the learned representation. Importantly, growth occurs earlier for the schizophrenia task than for the bipolar task. This suggests that both control and

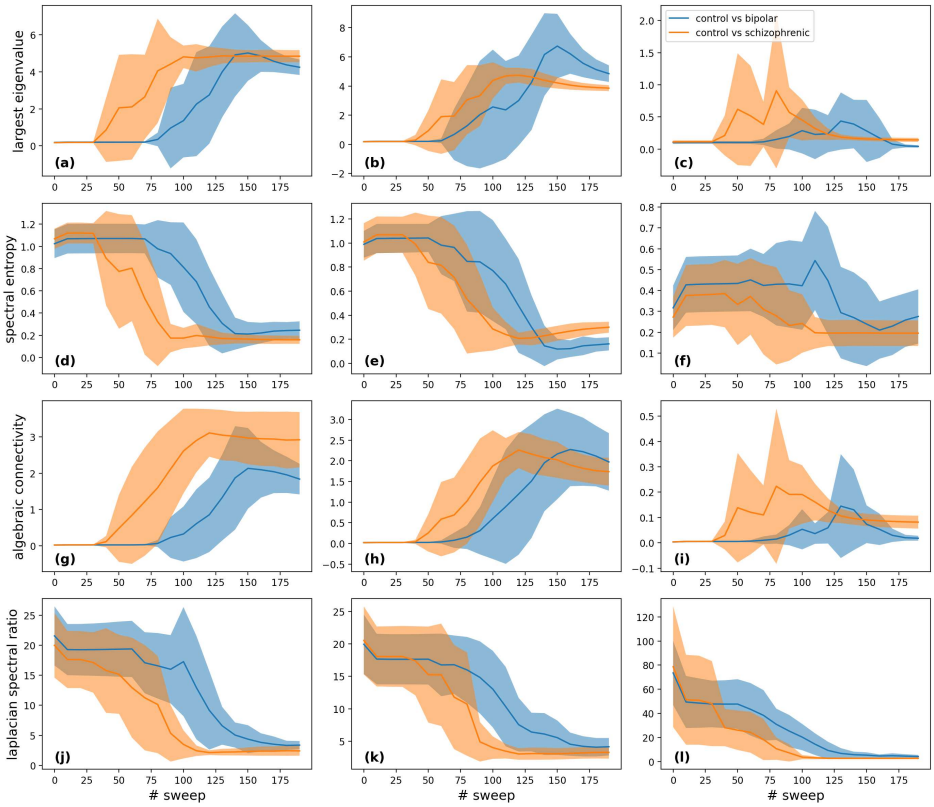



**Fig. 3** Evolution of graph-theoretic and spectral metrics computed from the pairwise quantum correlations matrices $\left|C_{i,j}^{(k,k')}\right|$ during training. Each row reports a different network metric: largest eigenvalue (a–c), spectral entropy (d–f), algebraic connectivity (g–i), and Laplacian spectral ratio (j–l), tracked over training sweeps. Columns correspond to (left) the control mask $\left|C^{(0,0)}\right|$, (center) the pathological mask $\left|C^{(1,1)}\right|$, and (right) confounding mask $\left|C^{(0,1)}\right|$. For each panel, curves show the mean metric value across sweeps, and shaded regions indicate the standard deviation across the train-test resampling runs. The two binary classification tasks are shown in blue (control vs. bipolar) and orange (control vs. schizophrenia).

pathological masks associated with schizophrenia classification develop structured correlations earlier during training. In contrast, bipolar classification requires a longer training period before the network reorganizes, producing a delayed increase in the leading eigenvalue. After reaching a peak, the eigenvalue tends to stabilize, indicating that the system transitions from a phase dominated by growing correlations to a more stable configuration in which correlations are redistributed across multiple modes. Notably, the confounding mask case in panel (c) exhibits a distinct non-monotonic evolution of the largest eigenvalue, characterized by a transient increase followed by a decline, suggesting the formation of short-lived collective correlations that do not stabilize into a persistent global structure.

The spectral entropy of the adjacency spectrum quantifies the diversity of modes that contribute to the network structure [68], with lower values indicating dominance

by a few modes. Panels (d–f) show a pronounced drop during training for both diagonal masks, a spectral condensation in which the correlation structure collapses onto a small number of dominant modes. As with the largest eigenvalue, the drop occurs earlier for the schizophrenia task, providing consistent evidence that the transition is delayed in the bipolar task. The confounding mask in panel (f) instead maintains relatively comparable and broadly distributed entropy values throughout the training, indicating that no comparable spectral condensation takes place.

The algebraic connectivity measures the global connectedness and robustness of the network [69, 70]. Larger values indicate stronger integration and greater resistance to partitioning into disconnected communities. Panels (g–i) reveal a rapid increase in algebraic connectivity during training for both diagonal masks. This suggests that the network transitions from a relatively fragmented correlation structure to a highly connected configuration. The timing of this increase again differs between the two classification problems: the rise occurs earlier for the schizophrenia task and later for the bipolar one, a temporal shift that provides a topological signature of the delayed organization. In contrast, the confounding masks display a non-monotonic evolution of the algebraic connectivity, with a transient increase followed by a reduction, indicating the temporary formation of integrated correlation pathways that subsequently weaken rather than stabilizing.

The Laplacian spectral ratio characterizes the spectral spread of the Laplacian and the dynamical stability of the network [71–73], with small ratios indicating improved synchronization capacity and structural regularity. Panels (j–l) show a dramatic reduction of this ratio as training progresses, indicating that the network becomes more homogeneous and dynamically stable. The transition again occurs earlier in the schizophrenia classification than in the bipolar one. The convergence of the ratio toward small values suggests that both tasks reach a regime of strong global coordination, but the onset of this regime differs across training sweeps. It is possible that this difference reflects the greater extent of cortical alterations in schizophrenia compared to bipolar disorder. In particular, Grecucci et al. [51], using the same dataset, showed that schizophrenia was characterized by multiple significantly altered structural networks relative to controls, whereas bipolar disorder exhibited a much more limited pattern of network differences. This asymmetry in network-level disruption may contribute to the earlier emergence of structured correlations observed in the schizophrenia classification task, as observed at the level of effective tensor network representations.

These spectral observables provide a consistent and coordinated description of the training dynamics. Importantly, these dynamics should be interpreted as changes in the internal organization of the learned representation rather than as temporal changes in biological brain networks. Initially, the quantum correlation networks display relatively weak structure imposed by random initial conditions, characterized by moderate eigenvalues, high spectral entropy, low algebraic connectivity, and large Laplacian spectral ratios. As training progresses, the networks undergo a sharp topological reorganization marked by the dominant eigenmode growing, the spectral entropy collapsing, the algebraic connectivity rising, and the Laplacian spectrum compressing.

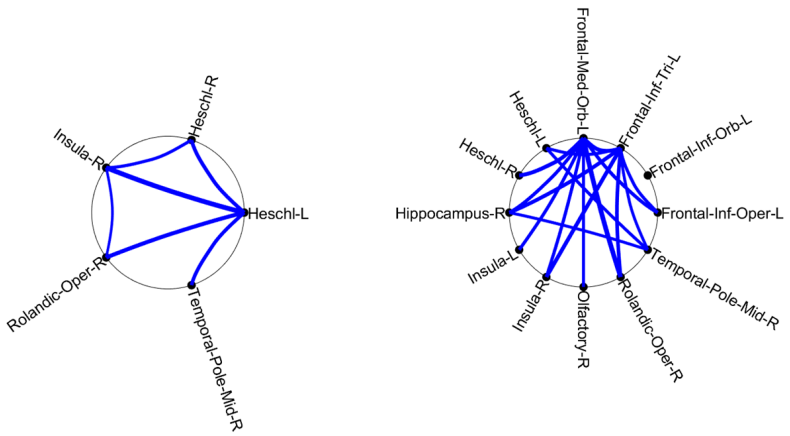



**Fig. 4** Circular plot of pairwise regions with significant differences in correlation values between controls and patients groups: bipolar disorder (left), and schizophrenia (right).

This coordinated set of changes identifies the transition in which the internal representations become globally structured and coherent. Crucially, the timing of this transition differs between classification tasks: the pathological masks associated with schizophrenia reach the organized regime earlier, while the bipolar case exhibits a systematic delay in spectral reorganization. The confounding cross cases display weaker and less structured transitions, indicating that the spectral signatures observed in the other panels are specifically associated with meaningful pathological discrimination.

By capturing large-scale reorganization of the learned correlation structure, these observables show how the MPS classifier, through its quantum correlations, progressively assembles the gray-matter features into a coherent network of brain regions whose topology encodes the discriminative signatures of the classification task.

## 3.2 Node-level interpretable quantities

Once the global topological reorganization of the correlation network obtained from quantum observables has been characterized, we identified the individual brain regions that act as hubs of this reorganization, focusing on the same stable feature sets considered in Appendix B: 12 regions for schizophrenia and 5 for bipolar, i.e. the features consistently selected by Boruta across all 15 resampling runs. Importantly, the five features selected in the bipolar classification are a subset of the twelve features identified in the schizophrenia task, indicating that the latter classification involves a broader but overlapping network of discriminative regions, in line with previous findings [51].

Figure 4 shows the pairwise regions that exhibit statistically significant differences in correlation values between the control and patient groups after Bonferroni correction for multiple comparisons. Before the two-sample t-tests, the correlation networks from all training resampling were transformed into Fisher's z-scores. Notably, all significant connections reflect increased correlation between regions in both patient groups, i.e. bipolar disorder and schizophrenia, compared to controls (see Fig. D5 in Appendix D). These increases should be interpreted within the learned interaction topology of

the classifier rather than as direct evidence of increased biological connectivity. This pattern suggests that the quantum-inspired algorithm focuses on pathological feature correlations. The left Heschl gyrus for the bipolar disorder task, and the frontal medial orbital and the frontal inferior triangular gyrus for the schizophrenia task, are the regions with the highest number of significant connections indicating that these regions emerge as highly central nodes within the learned interaction structure associated with classification.

Now we investigate the node-level importance, by focusing on graph-theoretic centralities which provide a topological characterization of how individual brain regions contribute to the classification. Fig. 5 reports five node-level centrality measures of the correlation networks derived from trained quantum states, i.e. strength, eigenvector centrality, PageRank, current-flow betweenness, and closeness, defined in Appendix C, computed over the 15 resampling runs.

Panels (a–e) correspond to the classification controls vs. schizophrenic patients and focus on the 12 features shared across all runs. Panels (f–j) correspond to controls vs. bipolar patients and focus on the 5 features consistently shared across the same number of samplings. For each feature, we compare the stability of its graph-theoretic centralities across runs with that of the Shapley attribution scores [74, 75] derived from the XGBoost tree-ensemble baseline introduced in Sec. 2.3. The Shapley values, shown in Fig. B4(b,d) in Appendix B, provide a complementary, classically computed view of feature importance (independent of the quantum correlations of the MPS).

Strength (panels a and f) quantifies the total connectivity of a node within the correlation network, i.e. how strongly a feature co-varies with the others. In the schizophrenia task, the pathological mask exhibits systematically larger strengths than the control mask for several brain regions, most prominently the Heschl gyrus, the frontal inferior opercular and triangular regions, and the insula; conversely, the hippocampus emerges as a relatively more stable integrative node in the control mask. In the bipolar task, the five shared features display the same pattern: elevated strength for the pathological mask relative to the control. The confounding mask shows markedly reduced strength across all regions, indicating weak and inconsistent cross-class correlations. The narrow error bars of the pathological masks across the 15 resamplings confirm the stability of these observations, which are consistent with previous findings on the same dataset [51] reporting more extensive structural covariance alterations in schizophrenia than in bipolar disorder.

Panels (b-d) and (g-i) report eigenvector centrality, PageRank, and current-flow betweenness. Despite their different formal definitions, capturing, respectively, embedding in densely connected neighborhoods, diffusion accessibility, and mediation of correlation flow, the three measures produce consistent rankings across nodes and masks. In the schizophrenia task, the pathological mask assigns the highest centrality to the frontal inferior triangular gyrus, the temporal pole, the Heschl gyrus, and the Rolandic operculum; the control mask additionally emphasizes the hippocampus. In the bipolar task, the shared features retain moderate-to-high centrality values, with an overall organization less polarized than in the schizophrenia case. The confounding mask displays lower and less structured centrality patterns throughout. The agreement among three measures with distinct analytical origins, together with the narrow

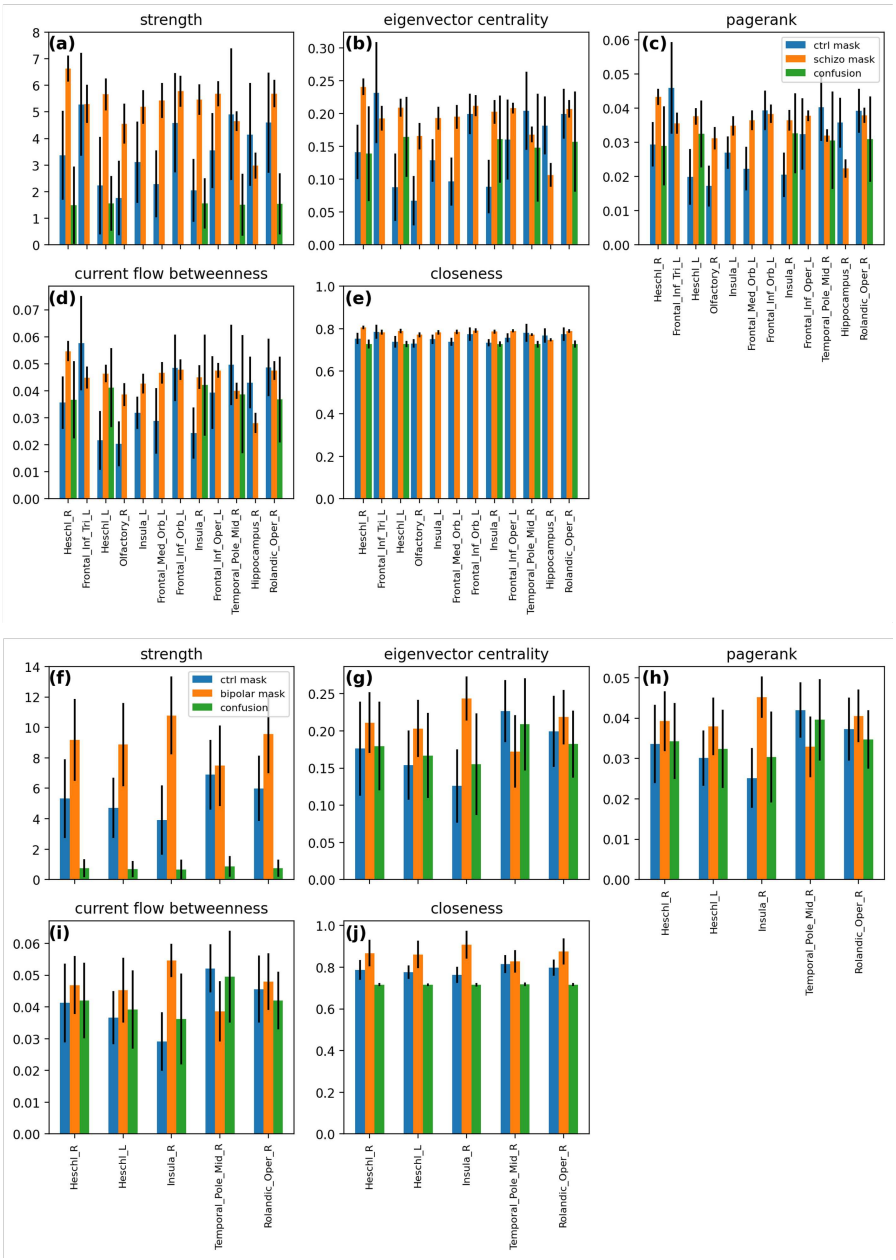



**Fig. 5** Comparison of node-level graph metrics derived from pairwise correlation matrices at the transition sweep for the control vs. schizophrenia classification in (a-e) and control vs. bipolar in (f-j). Panels report node-level strength, eigenvector centrality, PageRank, current-flow betweenness, and closeness computed from the control, pathological, and confounding pairwise correlation matrices (blue, orange, and green bars, respectively). The analysis focuses exclusively on the brain regions that were consistently selected across all 15 resampling runs following Boruta feature-ranking, ensuring that only the most stable and high-importance regions are retained.

error bars across the 15 resamplings, indicates a stable correlation backbone where the identified regions jointly act as hubs, diffusion centers, and bridges of correlation flow within the network.

Closeness quantifies how efficiently information can spread from a node to the rest of the graph via shortest paths [76], as expressed by panels (e) and (j). The closeness values show limited variation between masks and across runs. This behavior indicates that, although the relative strengths of correlations change, the global geometric accessibility of these regions remains relatively stable within the correlation network. The highest values tend to occur for Heschl gyrus, frontal inferior triangular gyrus and rolandic operculum. This indicates that these regions occupy central geometric positions in the correlation network, and this property remains relatively stable regardless of the mask.

The network-based centrality measures display substantially lower variability across the 15 resampling runs than the Shapley values reported in Appendix B (Fig. B4), confirming that the topology emerging from the quantum correlations of the MPS provides a more robust characterization of the underlying relational structure among features.

# 4 Discussion

The node-level analysis of the quantum pairwise correlation networks offers a direct interpretable link between the topological organization of gray matter features and the neurobiological substrates underlying schizophrenia and bipolar disorder. Importantly, these networks should be interpreted as model-derived interaction structures reflecting distributed classification-relevant organization rather than direct anatomical connectivity. By focusing on features that remain stable across 15 independent training–test samplings, the analysis isolates regions whose importance does not depend on a specific realization of the dataset but instead reflects persistent structural relationships in the gray matter correlation network within the learned representation, which is consistent with the broader perspective that structure-function interplay is fundamental to understanding disease mechanisms [77].

This pattern aligns with previous findings on the same dataset, where these regions participated in distributed structural covariance networks, with a broader involvement in schizophrenia and a more circumscribed engagement in bipolar disorder [51]. Importantly, these regions should be interpreted as components of distributed networks rather than isolated loci, suggesting that the identified features reflect stable hubs within a shared structural backbone within the learned interaction topology.

A key observation is the nested structure of the discriminative feature sets. The schizophrenia classification task identifies a broader set of twelve stable regions, whereas the bipolar classification isolates a reduced subset of five regions, all of which are contained within the schizophrenia feature set. This hierarchical organization is consistent with previous findings on the same dataset [51], where schizophrenia was associated with a more extensive pattern of structural covariance alterations across multiple networks, while bipolar disorder exhibited a more circumscribed configuration

embedded within a partially overlapping set of regions. These results are also supported by large meta-analyses that reported that both disorders involve gray matter reductions in frontal and insular regions [78, 79], but schizophrenia typically displays more extensive cortical, sensorimotor and cognitive alterations compared with bipolar disorder [80, 81], pointing to a broader principle of heterogeneity in brain organization that extends to other neurodevelopmental disorders [82].

Several of the most stable nodes identified in our analysis correspond to cortical regions repeatedly implicated in the pathophysiology of psychotic disorders:

(i) The presence of Heschl gyrus among the most stable nodes is consistent with extensive evidence linking anomalies of the primary auditory cortex to schizophrenia. Structural MRI studies show significant reductions in gray matter volume in Heschl gyrus in patients with schizophrenia compared with both healthy controls and bipolar disorder patients [83, 84]. These alterations are believed to contribute to the abnormal auditory processing and hallucination-related symptoms that characterize the disorder [85]. The elevated network centrality of this region within the learned interaction network is consistent with the prominent involvement of auditory cortical systems reported in schizophrenia literature.

(ii) Another recurrent region in the shared feature sets is the insula, which plays a key role in interoception, emotional regulation, and integration of sensory information. Neuroimaging meta-analyses have consistently reported decreased gray matter volume and altered connectivity in the insular cortex in both schizophrenia and bipolar disorder [78, 86]. The insula is a major node of the salience network, which coordinates switching between cognitive control networks and the default mode network. Importantly, the centrality of the insula within the identified networks supports its role as a transdiagnostic hub, mediating the integration between internally and externally oriented processes, and potentially contributing to both psychotic and affective dysregulation [87].

(iii) Several frontal areas appearing among the highlighted features, including inferior frontal and orbitofrontal regions, are associated with executive control, emotional regulation, and social cognition [88]. Structural and functional abnormalities in frontal cortex are among the most consistent findings in both schizophrenia and bipolar disorder, although again the alterations tend to be more pronounced and widespread in schizophrenia [78, 79]. The differential engagement of frontal regions is consistent with a hierarchical organization of network alterations, in which schizophrenia is characterized by a more extensive disruption of frontally mediated control systems, whereas bipolar disorder involves a more selective modulation within the same circuitry, as previously observed in structural covariance analyses on the same dataset [51]. Taken together, these findings suggest that the identified regions should not be interpreted as isolated loci, but rather as key nodes within distributed networks, whose coordinated disruption defines the structural organization of psychotic and affective disorders.

The node-level centrality measures provide complementary perspectives on how these brain regions participate in the gray matter correlation network:

(i) Strength and eigenvector centrality highlight nodes embedded within dense correlation structures. The consistently elevated values for several regions in the schizophrenia task suggest the presence of strongly coupled correlation hubs, indicating that structural changes in these regions are tightly linked with alterations across multiple cortical areas. In the bipolar task, the same regions remain important, showing comparable centrality values. This is consistent with neuroimaging literature indicating that bipolar disorder shares core structural alterations with schizophrenia but with less extensive network disruption [78]. This interpretation is further supported by previous analyses on the same dataset [51], where schizophrenia was associated with multiple altered structural covariance networks, whereas bipolar disorder showed a more circumscribed pattern within overlapping regions.

(ii) PageRank centrality reflects how correlations propagate through the network via indirect pathways. The relatively small variability across runs suggests that these regions maintain stable diffusion accessibility, implying that they form part of a persistent communication backbone within the correlation network.

(iii) Current-flow betweenness identifies nodes that mediate the flow of correlations between distant regions. The presence of elevated values for certain temporal and frontal regions indicates that they act as topological bridges within the learned interaction graph between otherwise weakly connected cortical modules, a property that may reflect abnormal integration between sensory and cognitive processing circuits in psychosis [80, 81].

(iv) Closeness centrality measures the global accessibility of nodes. The relatively small differences observed across tasks indicate that although local connectivity patterns change, the global geometric position of these nodes within the network remains stable.

A consistent pattern emerges across both quantum-inspired measures analyzed in this work: the single-site entanglement entropy (Appendix B) and the network centrality measures (Sec. 3.2) both display substantially higher stability across resampling runs than Shapley values (Appendix B). While the stability of the entanglement entropy reflects the structural properties of the global MPS representation, the stability of network centralities has a distinct origin: centrality measures depend on the second-order statistical structure of the data (pairwise correlations), which is relatively stable under repeated sampling, whereas Shapley values depend on the specific trained model and contextual learned feature interactions. As a result, centrality measures capture structural properties of the feature network, while Shapley values capture model-specific attribution patterns that may fluctuate depending on training conditions. The fact that the same brain regions emerge as stable network hubs despite fluctuations in Shapley scores reinforces the interpretation that these areas represent robust neuroanatomical correlates of the classification tasks.

The node-level analysis supports a model in which the bipolar classification corresponds to a reduced core subnetwork embedded within a broader schizophrenia-related structural network. This hierarchical relationship aligns with neuroimaging studies indicating that schizophrenia typically involves more widespread cortical alterations,

whereas bipolar disorder tends to affect a more restricted set of limbic and frontal regions [89]. This broader hierarchical perspective also resonates with morphology-based approaches revealing how brain network organization is structured in autism [90].

Within this framework, the five shared regions identified in the bipolar classification may represent a core structural module common to both disorders, potentially associated with cognitive deficits and sensory integration. The additional regions identified in the schizophrenia task may instead reflect more extensive disruptions of sensory, cognitive, and associative networks, consistent with the broader symptomatology of schizophrenia.

Insular and fronto-temporal regions define a shared structural backbone across both disorders [91], but with different weights: while the insula reflects a relatively stronger involvement in affective processing typical of bipolar disorder, fronto-temporal and auditory regions exhibit a more pronounced and widespread centrality in schizophrenia, consistent with its broader cortical dysconnectivity profile [92]. These characterizations partially support the asymmetry observed in Fig. 2 about sensitivity and specificity.

The results suggest that graph-theoretic analysis of the learned gray matter correlations networks, obtained from quantum observables, provides a robust framework for identifying stable neurobiological signatures of brain disorders. By focusing on network topology rather than individual feature contributions, the approach reveals hierarchical relationships between disorders and highlights regions that emerge as structural hubs within the pathological network. In this sense, the integration of graph-based and model-based interpretability may represent a promising direction for developing more robust and interpretable neuroimaging markers.

# 5 Conclusion

We introduced a tensor network machine-learning framework that maps gray-matter features into quantum-inspired correlation structures, enabling their interpretation as weighted graphs. Crucially, the quantum correlations encoded in this framework should not be interpreted as standard measures of statistical covariance. While conventional structural covariance networks rely on pairwise correlations that quantify isolated co-fluctuations, the MPS representation learns dependencies jointly within a single structured model. In this setting, the relationship between any two regions depends on the full configuration of features, rather than being estimated independently. As a result, although the computed quantum observables take the form of two-point correlations, they are derived from a global model and therefore implicitly capture context-dependent, higher-order interactions. This provides a more expressive and integrated description of feature interactions reflected in the network topology.

Within this unified representation, learning is not only assessed in terms of predictive performance, but also characterized through the emergence of structured correlations, providing direct access to the relational organization of brain features.

By combining MPS with graph-theoretic analysis, we showed that the learning dynamics are reflected in a sharp topological reorganization of the correlation network. This transition is characterized by global spectral observables, which reveal the formation of coherent collective modes, and by node-level centrality measures, which identify a stable set of brain regions acting as hubs of the learned representation.

By focusing on network topology rather than isolated feature contributions, we highlighted the emergence of a hierarchical organization of discriminative features, in which the regions relevant for bipolar disorder form a subset of those identified for schizophrenia. This nested structure provides a network-based characterization of the relationship between the two disorders and is consistent with the broader and more distributed structural alterations observed in schizophrenia [51].

Importantly, observables encoding quantum correlations, both the entanglement entropy and the graph-based centrality measures, exhibit higher stability across train-test resampling than model-based attribution measures, underscoring the value of quantum-inspired representations for identifying robust neurobiological markers.

Overall, our results demonstrate that tensor-network models, when coupled with interpretable graph-theoretic analysis, offer a principled framework for uncovering hierarchical and disorder-specific brain signatures, providing a bridge between quantum-inspired machine learning and neuroscience. More broadly, these findings support the view that brain disorders may be more effectively characterized through distributed organizational patterns emerging within learned representations than through isolated regional alterations alone.

**Acknowledgements.** G.M. acknowledges support from INFN through the project "QUANTUM".

## Funding Declarations

Authors were supported by the Italian funding within the "Budget MUR - Dipartimenti di Eccellenza 2023 - 2027" (Law 232, 11 December 2016) - Quantum Sensing and Modelling for One-Health (QuaSiModO), CUP:H97G23000100001.

## Data Availability

The data analysed in the present study are publicly and freely available through the Zenodo repository at https://doi.org/10.5281/zenodo.460878. A detailed description of the dataset, including participant characteristics, MRI acquisition parameters, and data features, is provided in [93].

## Additional information

The authors declare no competing interests. The code is available upon reasonable request.

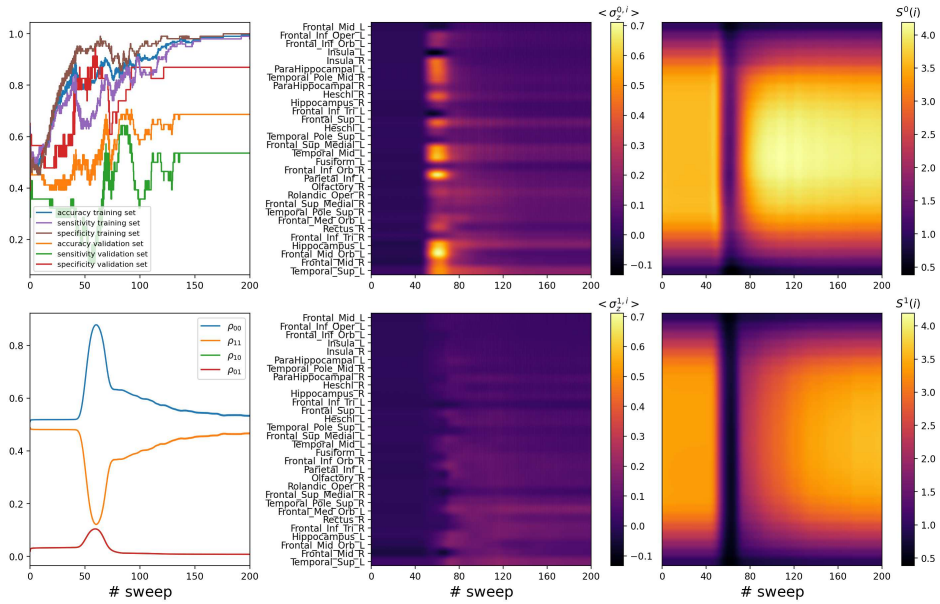



**Fig. A1** Left panels: evaluation metrics for training and independent sets (top) and reduced density matrix in the label subspace (bottom). Central panels: magnetization patterns extracted by the MPS masks for the control class (top) and schizophrenic patients (bottom). Right panels: entanglement entropy profiles $S^{(0)}(i)$ (controls, top) and $S^{(1)}(i)$ (patients, bottom) along the one-dimensional lattice.

# Author contributions statement

Conceptualization: DP, GM, AG; Software: DP; Data Curation: DP, AG; Methodology: DP; Formal analysis and investigation: DP, GM, MO; Visualization: DP, GM, MO; Funding acquisition: RB, SS, AM, NA; Resources: DP, AM, NA; Supervision: DP, GM, AG; Writing - original draft preparation: DP, GM, AG; Writing - review and editing: DP, GM, AG, LB, JC, MLR, AM, MO, AS, MS, RB, SS, NA.

# Appendix A Training dynamics of the MPS classifier

In this appendix, we characterize the learning dynamics observed in a representative run of each of the two binary classification tasks involving clinical data: *schizophrenic vs. control subjects* and *bipolar vs. control subjects*. For each task, we track the evolution of the evaluation metrics, the reduced density matrix in label space, the magnetization patterns extracted by the two trained MPS weight tensors with bond dimension $\chi = 400$, and the entanglement entropy profiles along the one-dimensional lattice.

## A.1 Schizophrenic patients vs. controls

In the schizophrenia–control task, shown in Fig. A1, the left column displays the evolution of accuracy, sensitivity, and specificity for both training and independent

test sets. The evaluation metrics show a delayed generalization, stabilizing at high values only after a transient regime of fluctuations. Around a relatively early sweep number, which we refer to as the transition sweep, the sensitivity and specificity of the independent test set start to improve and stabilize. This stabilization marks the onset of effective generalization.

This delayed generalization is accompanied by a sharp transition in the internal observables of the trained state, with a phenomenology qualitatively similar to the grokking transition reported in tensor-network classifiers [62]. In particular, the reduced density matrix in the label subspace, shown below the metrics, exhibits a sharp reorganization: its off-diagonal terms decay abruptly around the same sweep, indicating the emergence of two dynamically decoupled masks, and signaling a clear reorganization of the classification subspace.

The central column illustrates the local magnetization patterns for the two trained quantum states: the upper panel corresponds to control subjects, the lower to schizophrenic patients. Features relevant for each class become sharply activated near the transition sweep, forming stable and class-specific magnetization profiles.

The right column presents the bipartite entanglement entropy $S^{(\ell)}(i)$ across the one-dimensional lattice for each class. Starting from a volume-law profile associated with a random initialization, the entropy rapidly collapses into a sub-volume scaling after the transition sweep. This transition is consistent with the reduced number of effective degrees of freedom required once the model has extracted the minimal feature set necessary for generalization.

The schizophrenia-control classification undergoes this transition at a comparatively *small* number of sweeps, indicating that the model can quickly isolate a concise set of discriminative features.

## A.2 Bipolar patients vs. controls

The bipolar-control task, shown in Fig. A2, exhibits the same qualitative phenomenology but with a markedly different timescale. In the left column, the accuracy and sensitivity curves for the independent set reveal a delayed generalization phase: the sudden improvement occurs only after a considerably larger number of sweeps compared to the schizophrenia-control case, indicating that the model must explore a more complex landscape before isolating class-specific features.

The magnetization patterns shown in the central panels corroborate this observation. In contrast to the schizophrenia task, where a small number of regions dominate early, the bipolar-control masks develop discernible patterns only after extended training. This suggests that the discriminative structure of the data is either more diffuse or encoded in more subtle correlations that require more optimization steps to uncover.

The entanglement entropy profiles in the right column again begin in a volume-law regime. However, the collapse to a sub-volume profile occurs later, precisely aligned with the delayed transition point. The fact that the volume-to-sub-volume transition is postponed reflects the need for the MPS to retain higher entanglement for a longer period while navigating a larger or more entangled feature manifold before locking onto the relevant discriminatory subspace.

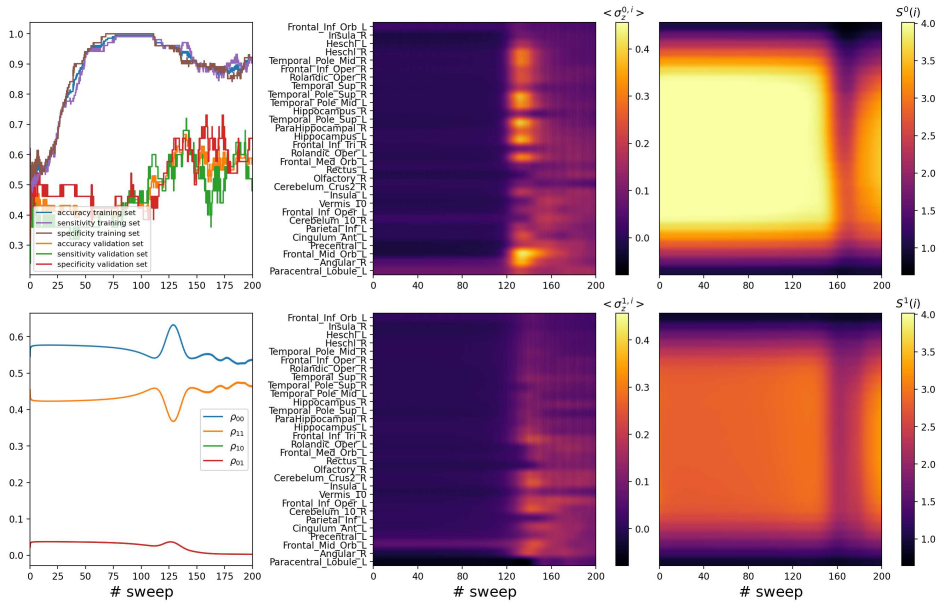



**Fig. A2** Left panels: evaluation metrics for training and independent sets (top) and reduced density matrix in label space (bottom). Central panels: magnetization patterns for the control class (top) and bipolar patients (bottom). Right panels: entanglement entropy profiles $S^{(0)}(i)$ (controls, top) and $S^{(1)}(i)$ (patients, bottom).

## A.3 Comparative interpretation

The two sets of panels illustrate that although both tasks display the characteristic signatures of delayed generalization, i.e. decay of label space coherence, emergence of stable magnetization, and an entanglement entropy transition, the *timescales* of these transitions differ markedly.

The schizophrenia vs. control task reaches high generalization after relatively few sweeps, indicating that its feature structure is easier to isolate within the MPS architecture. In contrast, the bipolar–control task requires significantly more sweeps, implying that its discriminatory features are harder to extract, more entangled, or less spatially localized along the one-dimensional lattice.

To assess the stability of these observations against dataset variability, we performed 15 independent training–test resampling runs for each diagnostic task. The resulting histograms of transition sweeps, shown in Fig. A3, provide a quantitative measure of the variability of the transition times for both the control and pathology quantum states.

For the schizophrenia–control comparison, the distribution of transition sweeps (orange bars) is consistently shifted toward earlier values. In the control histogram, the schizophrenia transitions cluster sharply in the range $\sim$ 60-110 sweeps, with a narrow spread and no heavy tail. The analogous histogram for the state encoding the pathology class exhibits the same early-transition behaviour, indicating that the

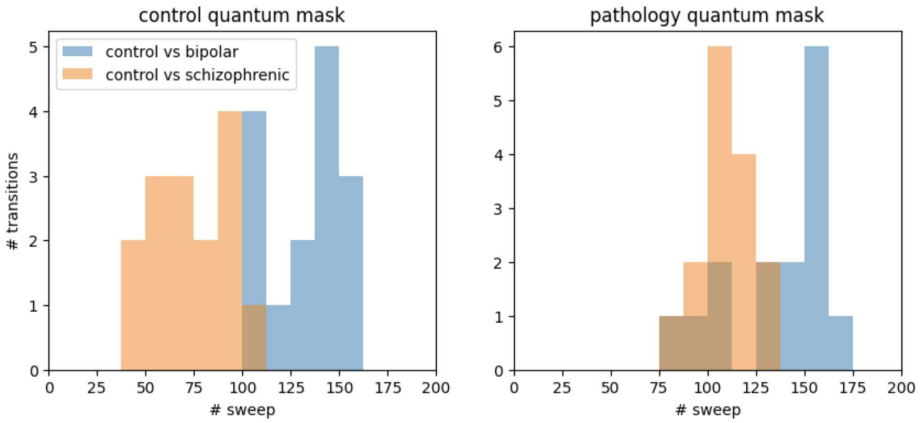



**Fig. A3** Distribution of transition sweeps collected over 15 independent training–test resampling runs for the two diagnostic tasks. Left: transition–sweep histograms for the control quantum state in the bipolar–control (blue) and schizophrenia–control (orange) tasks. Right: corresponding histograms for the pathology quantum states. The schizophrenia–control task consistently exhibits earlier and more tightly clustered transitions, while the bipolar–control task shows later and more variable transition times, reflecting the differing complexity of the underlying feature structures.

onset of stable magnetization and entanglement reorganization is reproducible across resampling and not driven by a specific partition of the data.

In contrast, the bipolar–control task (blue bars) exhibits systematically later transitions and a markedly broader spread. Both the control and pathology states show transitions concentrated primarily between $\sim$ 120-170 sweeps, with higher variance and a noticeable right-skew. This broader distribution signals that the transition dynamics are more sensitive to resampling, consistent with the interpretation that bipolar features are intrinsically harder to disentangle and require more extensive training for the masks to isolate class-specific signatures.

The histogram analysis reinforces the comparative picture: schizophrenia–control classifications consistently undergo transitions early and with low variance across resampling, whereas bipolar–control classifications transition substantially later, with higher trial-to-trial variability. This behaviour mirrors the differences already evident at the level of magnetization patterns, reduced density matrix decoherence, and entanglement-entropy transitions, confirming that the observed timescale separation is a robust phenomenon rather than a single-run artefact.

# Appendix B Quantum-inspired interpretability

In order to extract feature information from the trained quantum-inspired model, we consider the single-site entanglement entropy [26, 94]. For each site $i$ of the MPS, we compute the reduced density matrix $\varrho_i$, obtained by tracing out all other sites from the trained MPS and summing over the class-label index so that $\varrho_i$ encodes information from both classes simultaneously. The single-site entanglement entropy is then defined as $S_i = -\mathrm{Tr}\,(\varrho_i \ln \varrho_i)$, which quantifies the degree of quantum correlation between the feature $i$ and the other features within the MPS representation. Since each feature

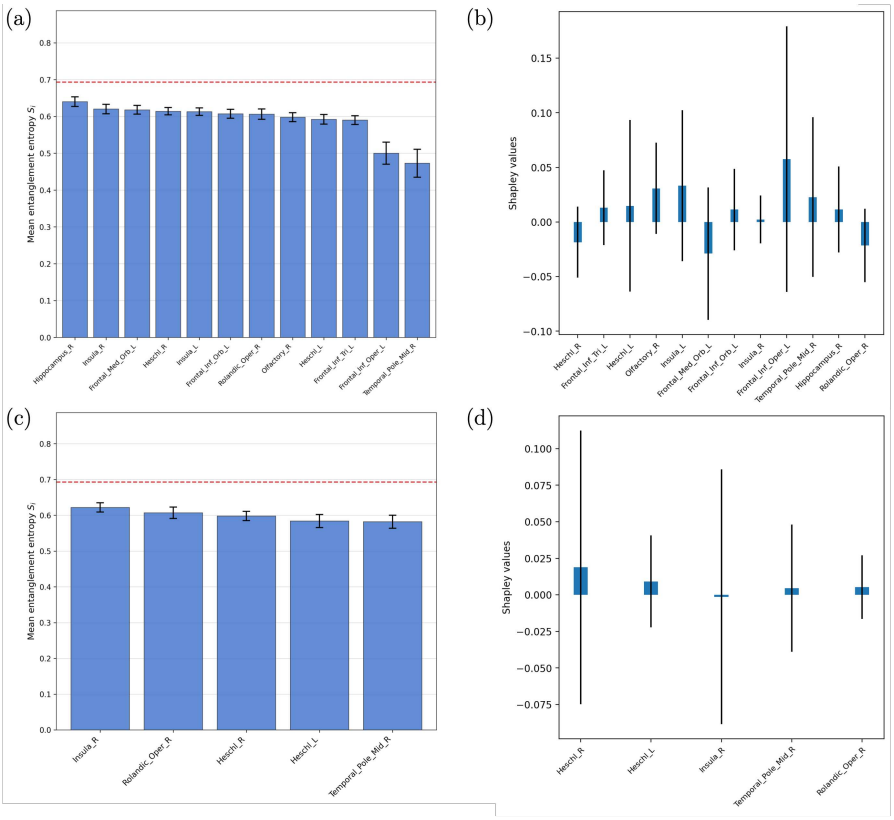



**Fig. B4** Quantum entropy and standard interpretability measures for the control vs. schizophrenia classification (panels a–b) and control vs. bipolar classification (panels c–d). Panels (a) and (c) show the mean single-site entanglement entropy $S_i$ computed from the trained MPS, averaged over the 15 resampling runs, for the features consistently selected by Boruta across all runs. The red dashed line indicates the maximum achievable value of the single-site entanglement entropy. Panels (b) and (d) show the Shapley values assigned to the same subset of regions, reflecting each region's contribution to the XGBoost classifier predictive decision.

is encoded as a qubit, $S_i$ is bounded above by $\ln 2 \approx 0.693$. Features with higher entanglement entropy are more deeply integrated into the collective quantum correlations that define the learned representation. Importantly, the entropy values should be interpreted as properties of the learned representation rather than direct neurobiological measures. By construction, this measure reflects representational integration rather than marginal contribution to the classification output, which is the focus of attribution methods such as Shapley values [74, 75]. We compute the single-site entanglement entropy from the trained MPS and the Shapley values from the XGBoost baseline, obtaining both quantities for the same set of features and treating them as complementary probes of the learned structures.

In Fig. B4, we report the mean single-site entanglement entropy $S_i$ (panels a and c), and the Shapley values (panel b and d) for the features consistently selected by

Boruta across all 15 resampling runs, i.e. 12 stable regions for the schizophrenia task and 5 for the bipolar task. Both quantities are averaged over the 15 runs.

A key observation emerges from the comparison with the Shapley values, which show greater variability between runs for both classification tasks. This indicates that the single-site entanglement entropy is more stable across resampling runs than the model-based attribution scores. While Shapley values depend on the specific trained model and learned decision boundary, the entanglement entropy remains robust under data splits.

# Appendix C Graph theory global and node-level quantities

We introduce in this appendix the weighted graph quantities employed in the analysis of the structural organization underlying the quantum masks $W^\ell$ adopted for the binary classification tasks.

Let $G = (V, E)$ be a weighted undirected graph with $|V| = N$ nodes and weighted adjacency matrix $A = (a_{i,j}) \in \mathbb{R}^{N\times N}$, where $a_{i,j} = |C_{i,j}| \geq 0$ encodes the strength of the interaction between nodes $i$ and $j$, where $C_{i,j}$ is introduced in Eq. (3).

## C.1 Global quantities

### C.1.1 Largest eigenvalue of the weighted adjacency matrix

The spectral radius of $A$,

$$\lambda_N(A) = \max\{\,\mathrm{spec}(A)\,\}, \tag{C1}$$

provides a global descriptor of the graph's connectivity strength. For non-negative symmetric matrices, $\lambda_N(A)$ coincides with the largest real eigenvalue and is associated with a strictly non-negative eigenvector by the Perron-Frobenius theorem. In weighted graphs this quantity captures the presence of strongly connected hubs or dense subgraphs and is often interpreted as a proxy for the network's overall interaction intensity [67].

### C.1.2 Spectral entropy of the adjacency matrix

To quantify the heterogeneity of the adjacency spectrum, we define the normalized squared eigenvalues

$$p_i = \frac{\lambda_i^2}{\sum_{j=1}^N \lambda_j^2}, \qquad i = 1, \ldots, N, \tag{C2}$$

where $\{\lambda_i\}$ are the eigenvalues of $A$. The spectral entropy is then given by [68]

$$H_{\mathrm{spec}}(A) = -\sum_{i=1}^{N} p_i \log p_i. \tag{C3}$$

A low spectral entropy indicates that a small number of eigenmodes dominate the connectivity structure (e.g., the presence of few influential hubs), whereas higher entropy reflects a more uniform distribution of topological scales.

### C.1.3 Algebraic connectivity

Let $D = \mathrm{diag}(d_1, \ldots, d_N)$ denote the degree matrix with $d_i = \sum_j a_{ij}$, and define the weighted Laplacian

$$L = D - A. \tag{C4}$$

Since $L$ is symmetric and positive semidefinite, its eigenvalues satisfy

$$0 = \mu_1 \leq \mu_2 \leq \cdots \leq \mu_N. \tag{C5}$$

The second smallest eigenvalue, $\mu_2$, is called the *algebraic connectivity* (or Fiedler value) [69, 70]:

$$\alpha(G) = \mu_2. \tag{C6}$$

This quantity measures the robustness of the network's global connectivity: larger values indicate that the graph is more difficult to partition and possesses more redundant or distributed pathways across its nodes.

### C.1.4 Laplacian spectral ratio

To capture both the connectivity strength and the spectral spread of the Laplacian, we define the Laplacian ratio as [71–73]

$$R_L(G) = \frac{\mu_N}{\mu_2}, \tag{C7}$$

where $\mu_N$ is the largest eigenvalue of $L$. A high Laplacian ratio reflects a network with tightly clustered or highly heterogeneous connectivity patterns, while smaller values indicate a more homogeneous connectivity profile. This ratio is particularly informative when comparing networks that share similar densities but differ in structural organization, as it incorporates both global cohesion ($\mu_2$) and maximal variation across nodes ($\mu_N$).

## C.2 Node-level quantities

### C.2.1 Node strength

For a weighted, undirected network with adjacency matrix $A = (a_{ij})$, the *strength* of node $i$ is the sum of incident edge weights [95]

$$s_i \;=\; \sum_{j=1}^{N} a_{i,j}. \tag{C8}$$

The vector $\boldsymbol{s} = (s_1, \ldots, s_N)^\intercal$ equals $A\mathbf{1}$, where $\mathbf{1}$ is the all-ones vector. Node strength generalizes degree to weighted graphs and quantifies the total interaction intensity of a node in terms of its connectivity.

### C.2.2 Eigenvector centrality

Eigenvector centrality assigns high score to nodes that are connected to other highly central nodes, according to each component value $x_i$, $i = 1, \dots, N$. It is defined as the (entrywise nonnegative) eigenvector associated with the largest eigenvalue of $A$ [76]

$$A\,\boldsymbol{x} \;=\; \lambda_N(A)\,\boldsymbol{x}, \qquad x_i \geq 0, \quad \|\boldsymbol{x}\|_2 = 1. \tag{C9}$$

In weighted graphs, edges with larger $a_{ij}$ contribute more strongly to the mutual reinforcement among neighboring centralities.

### C.2.3 PageRank

Let $P$ be a row-stochastic transition matrix derived from $A$. A possible choice for undirected weighted graphs is $P_{i,j} = a_{i,j}/s_i$, which performs a random walk proportional to edge weights. We can introduce a centrality satisfying a steady state condition $\boldsymbol{\pi}^\intercal = \boldsymbol{\pi}^\intercal P$, yielding the normalized strength $\pi_i = s_i / \sum_k s_k$.

The PageRank vector [96] is deduced by introducing a uniform fallback because dense weighted graphs are often nearly fully connected and easily present dominant eigenmodes. A damping factor $\alpha \in (0, 1)$ weights the inclusion of a personalization vector $\boldsymbol{v}$ (assumed uniform in our analysis, $\boldsymbol{v} = \frac{1}{N}\mathbf{1}$), that biases the random surfer to restart at specific, preferred nodes, thus customizing importance scores, essential for personalized recommendations, user-specific search and community detection. The steady state satisfies

$$\boldsymbol{\pi}^\intercal = \alpha \boldsymbol{\pi}^\intercal P + (1-\alpha)\boldsymbol{v}^\intercal, \tag{C10}$$

where larger $\alpha$ emphasizes network structure, while smaller $\alpha$ increases the influence of $\boldsymbol{v}$. In our implementation the damping factor $\alpha = 0.85$, ruling the transition matrix $\widetilde{P} = \alpha P + \dfrac{1-\alpha}{N}\mathbf{1}\mathbf{1}^\intercal$.

### C.2.4 Current-flow betweenness

Current-flow betweenness [97, 98] measures how often a node carries flow when one unit of current is injected at a source node $s$ and extracted at a target one $t$, assuming each edge $(i, j)$ has conductance $a_{i,j}$ (thus resistance $r_{i,j} = 1/a_{i,j}$ for $a_{i,j} > 0$). Let $L$ be the weighted Laplacian introduced in Eq. (C4) and $L^+$ its Moore-Penrose pseudoinverse.

Ohm's law on edge $(i, j)$ reads $I^{s,t}_{i,j} = a_{i,j}(v^{s,t}_i - v^{s,t}_j)$, where $v^{s,t}_i$ is the voltage at node $i$ and $I^{s,t}_{i,j}$ the current from $i$ to $j$. Kirchhoff's current law at each node $i$ yields the external current injection $\sum_j I^{s,t}_{i,j} = b^{s,t}_i$, and the application of Ohm's law results in

$$v^{s,t}_i \sum_j a_{i,j} - \sum_j a_{i,j} v^{s,t}_j = \sum_j L_{i,j} v^{s,t}_j = b^{s,t}_i. \tag{C11}$$

For a given pair $(s, t)$ with $b^{s,t}_i = e^s_i - e^t_i$, the node potential vector is $v^{s,t}_i = L^+(e^s_i - e^t_i)$, where $e_i$ is the $i$-th canonical basis vector. The net current through node $i$ induced by the unit $s \to t$ injection can be expressed in terms of potential differences

across incident edges. The current-flow betweenness of node $i$ is then the average of its absolute through-current over all ordered pairs:

$$\mathrm{B}(i) \;=\; \frac{1}{N(N-1)} \sum_{\substack{s,t\in V\\ s\neq t}} \frac{1}{2}\sum_{j} a_{ij} \left| L^{+}(e_i^s - e_i^t) - L^{+}(e_j^s - e_j^t) \right|, \tag{C12}$$

where the inner sum aggregates the magnitude of current traversing node $i$ via edges $(i, j)$. In practice, stochastic sampling over $(s, t)$ pairs can be used to estimate B efficiently in large networks.

### C.2.5 Closeness

We define a geometric distance between nodes from the edge weights as

$$d_{i,j} \;=\; \sqrt{2(1 - a_{i,j})}\,, \qquad d_{i,i} = 0, \tag{C13}$$

so that larger weights imply shorter distances. This Mantegna distance formula comes from embedding standardized variables on a unit hypersphere. The weighted closeness centrality of node $i$ is defined as [99, 100]

$$c_i \;=\; \frac{N-1}{\sum_{j\neq i} d_{i,j}}. \tag{C14}$$

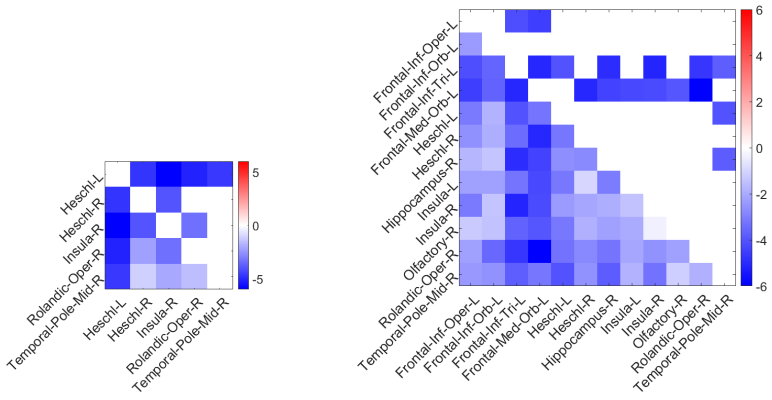
where for pairs without a direct edge, we set $d_{i,j}$ to the length of the *shortest path* under Eq. (C13), i.e. the minimal path sum of edge-wise distances along any $i \to j$ route in the graph (finite if $j$ is reachable from $i$, and $+\infty$ otherwise).

# Appendix D Supplementary figures

# References


[1] Arrazola, J.M., Bergholm, V., Brádler, K., Bromley, T.R., Collins, M.J., Dhand, I., Fumagalli, A., Gerrits, T., Goussev, A., Helt, L.G., Hundal, J., Isacsson, T., Israel, R.B., Izaac, J., Jahangiri, S., Janik, R., Killoran, N., Kumar, S.P., Lavoie, J., Lita, A.E., Mahler, D.H., Menotti, M., Morrison, B., Nam, S.W., Neuhaus, L., Qi, H.Y., Quesada, N., Repingon, A., Sabapathy, K.K., Schuld, M., Su, D., Swinarton, J., Száva, A., Tan, K., Tan, P., Vaidya, V.D., Vernon, Z., Zabaneh, Z., Zhang, Y.: Quantum circuits with many photons on a programmable nanophotonic chip. Nature **591**(7848), 54–60 (2021) https://doi.org/10.1038/s41586-021-03202-1

[2] Banchi, L., Fingerhuth, M., Babej, T., Ing, C., Arrazola, J.M.: Molecular docking with gaussian boson sampling. Science




(a) (b)

**Fig. D5** T-values for group differences in correlation networks: controls versus bipolar disorder (a) and schizophrenia (b). The upper triangle of each matrix indicates the pairwise regions with statistically significant differences.

Advances **6**(23), 1950 (2020) https://doi.org/10.1126/sciadv.aax1950 https://www.science.org/doi/pdf/10.1126/sciadv.aax1950

[3] Yu, S., Zhong, Z.-P., Fang, Y., Patel, R.B., Li, Q.-P., Liu, W., Li, Z., Xu, L., Sagona-Stophel, S., Mer, E., Thomas, S.E., Meng, Y., Li, Z.-P., Yang, Y.-Z., Wang, Z.-A., Guo, N.-J., Zhang, W.-H., Tranmer, G.K., Dong, Y., Wang, Y.-T., Tang, J.-S., Li, C.-F., Walmsley, I.A., Guo, G.-C.: A universal programmable gaussian boson sampler for drug discovery. Nature Computational Science **3**(10), 839–848 (2023) https://doi.org/10.1038/s43588-023-00526-y

[4] Vakili, M.G., Gorgulla, C., Nigam, A., Bezrukov, D., Varoli, D., Aliper, A., Polykovsky, D., Das, K.M.P., Snider, J., Lyakisheva, A., Mansob, A.H., Yao, Z., Bitar, L., Radchenko, E., Ding, X., Liu, J., Meng, F., Ren, F., Cao, Y., Stagljar, I., Aspuru-Guzik, A., Zhavoronkov, A.: Quantum Computing-Enhanced Algorithm Unveils Novel Inhibitors for KRAS (2024). https://arxiv.org/abs/2402.08210

[5] Benedetti, M., Garcia-Pintos, D., Perdomo, O., Leyton-Ortega, V., Nam, Y., Perdomo-Ortiz, A.: A generative modeling approach for benchmarking and training shallow quantum circuits. npj Quantum Information **5**(1), 45 (2019) https://doi.org/10.1038/s41534-019-0157-8

[6] Hibat-Allah, M., Mauri, M., Carrasquilla, J., Perdomo-Ortiz, A.: A framework for demonstrating practical quantum advantage: comparing quantum against classical generative models. Communications Physics **7**(1), 68 (2024) https://doi.org/10.1038/s42005-024-01552-6

[7] Gili, K., Hibat-Allah, M., Mauri, M., Ballance, C., Perdomo-Ortiz, A.: Do quantum circuit born machines generalize? Quantum Science and Technology **8**(3), 035021 (2023) https://doi.org/10.1088/2058-9565/acd578

[8] Caro, M.C., Huang, H., Cerezo, M., Sharma, K., Sornborger, A., Cincio, L., Coles, P.J.: Generalization in quantum machine learning from few training data. Nat. Commun. **13**, 4919 (2022) https://doi.org/10.1038/s41467-022-32550-3

[9] Gibbs, J., Holmes, Z., Caro, M.C., Ezzell, N., Huang, H.-Y., Cincio, L., Sornborger, A.T., Coles, P.J.: Dynamical simulation via quantum machine learning with provable generalization. Phys. Rev. Res. **6**, 013241 (2024) https://doi.org/10.1103/PhysRevResearch.6.013241

[10] Peters, E., Schuld, M.: Generalization despite overfitting in quantum machine learning models. Quantum **7**, 1210 (2023) https://doi.org/10.22331/q-2023-12-20-1210

[11] Bowles, J., Wright, V.J., Farkas, M., Killoran, N., Schuld, M.: Contextuality and inductive bias in quantum machine learning (2023). https://arxiv.org/abs/2302.01365

[12] Gil-Fuster, E., Eisert, J., Bravo-Prieto, C.: Understanding quantum machine learning also requires rethinking generalization. Nat. Commun. **15**(2277) (2024) https://doi.org/10.1038/s41467-024-45882-z

[13] Pomarico, D., Cosmai, L., Facchi, P., Lupo, C., Pascazio, S., Pepe, F.V.: Dynamical Quantum Phase Transitions of the Schwinger Model: Real-Time Dynamics on IBM Quantum. Entropy **25**(4), 608 (2023) https://doi.org/10.3390/e25040608

[14] Pomarico, D., Pandey, M., Cioli, R., Dell'Anna, F., Pascazio, S., Pepe, F.V., Facchi, P., Ercolessi, E.: Quantum error mitigation in optimized circuits for particle-density correlations in real-time dynamics of the Schwinger model. Entropy **27**(4) (2025) https://doi.org/10.3390/e27040427

[15] Pomarico, D., Dell'Anna, F., Cioli, R., Pascazio, S., Pepe, F.V., Facchi, P., Ercolessi, E.: Digital rydberg simulation of dynamical quantum phase transitions in the schwinger model. Phys. Rev. Res. **8**, 013177 (2026) https://doi.org/10.1103/b233-qtvm

[16] Cincio, L., Rudinger, K., Sarovar, M., Coles, P.J.: Machine learning of noise-resilient quantum circuits. PRX Quantum **2**, 010324 (2021) https://doi.org/10.1103/PRXQuantum.2.010324

[17] Ran, S.-J.: Encoding of matrix product states into quantum circuits of one- and two-qubit gates. Phys. Rev. A **101**, 032310 (2020) https://doi.org/10.1103/PhysRevA.101.032310

[18] Rudolph, M.S., Chen, J., Miller, J., Acharya, A., Perdomo-Ortiz, A.: Decomposition of matrix product states into shallow quantum circuits. Quantum Science and Technology **9**(1), 015012 (2023) https://doi.org/10.1088/2058-9565/ad04e6

[19] Rudolph, M.S., Miller, J., Motlagh, D., Chen, J., Acharya, A., Perdomo-Ortiz, A.: Synergistic pretraining of parametrized quantum circuits via tensor networks. Nat. Commun. **14**, 8367 (2023) https://doi.org/10.1038/s41467-023-43908-6

[20] Schuhmacher, J., Ballarin, M., Baiardi, A., Magnifico, G., Tacchino, F., Montangero, S., Tavernelli, I.: Hybrid tree tensor networks for quantum simulation. PRX Quantum **6**, 010320 (2025) https://doi.org/10.1103/PRXQuantum.6.010320

[21] Khosrojerdi, M., Pereira, J.L., Cuccoli, A., Banchi, L.: Learning to classify quantum phases of matter with a few measurements. Quantum Science and Technology **10**(2), 025006 (2025) https://doi.org/10.1088/2058-9565/ada79b

[22] Stoudenmire, E.M., Schwab, D.J.: Supervised Learning with Quantum-Inspired Tensor Networks (2017). https://arxiv.org/abs/1605.05775

[23] Efthymiou, S., Hidary, J., Leichenauer, S.: TensorNetwork for Machine Learning (2019). https://arxiv.org/abs/1906.06329

[24] Huggins, W., Patil, P., Mitchell, B., Whaley, K.B., Stoudenmire, E.M.: Towards quantum machine learning with tensor networks. Quantum Science and Technology **4**(2), 024001 (2019) https://doi.org/10.1088/2058-9565/aaea94

[25] Liao, H., Convy, I., Yang, Z., Whaley, K.B.: Decohering tensor network quantum machine learning models. Quantum Machine Intelligence **5**(7) (2023) https://doi.org/10.1007/s42484-022-00095-9

[26] Felser, T., Trenti, M., Sestini, L., Gianelle, A., Zuliani, D., Lucchesi, D., Montangero, S.: Quantum-inspired machine learning on high-energy physics data. npj Quantum Inf. **7**(111) (2021) https://doi.org/10.1038/s41534-021-00443-w

[27] Dborin, J., Barratt, F., Wimalaweera, V., Wright, L., Green, A.G.: Matrix product state pre-training for quantum machine learning. Quantum Science and Technology **7**(3), 035014 (2022) https://doi.org/10.1088/2058-9565/ac7073

[28] Ballarin, M., Mangini, S., Montangero, S., Macchiavello, C., Mengoni, R.: Entanglement entropy production in Quantum Neural Networks. Quantum **7**, 1023 (2023) https://doi.org/10.22331/q-2023-05-31-1023

[29] Collura, M., Dell'Anna, L., Felser, T., Montangero, S.: On the descriptive power of neural-networks as constrained tensor networks with exponentially large bond dimension. SciPost Phys. Core **4**, 1 (2021) https://doi.org/10.21468/SciPostPhysCore.4.1.001

[30] Glasser, I., Pancotti, N., August, M., Rodriguez, I.D., Cirac, J.I.: Neural-network quantum states, string-bond states, and chiral topological states. Phys. Rev. X **8**, 011006 (2018) https://doi.org/10.1103/PhysRevX.8.011006

[31] Glasser, I., Pancotti, N., Cirac, J.I.: From probabilistic graphical models to generalized tensor networks for supervised learning. IEEE Access **8**, 68169–68182 (2020) https://doi.org/10.1109/ACCESS.2020.2986279

[32] Sun, Z.-Z., Peng, C., Liu, D., Ran, S.-J., Su, G.: Generative tensor network classification model for supervised machine learning. Phys. Rev. B **101**, 075135 (2020) https://doi.org/10.1103/PhysRevB.101.075135

[33] Gallego, A.J., Orús, R.: Language design as information renormalization. Springer Nature Computer Science **3**, 140 (2022) https://doi.org/10.1007/s42979-021-01002-y

[34] Cheng, S., Wang, L., Xiang, T., Zhang, P.: Tree tensor networks for generative modeling. Phys. Rev. B **99**, 155131 (2019) https://doi.org/10.1103/PhysRevB.99.155131

[35] Chen, H., Barthel, T.: Machine learning with tree tensor networks, CP rank constraints, and tensor dropout (2023). https://arxiv.org/abs/2305.19440

[36] Dilip, R., Liu, Y.-J., Smith, A., Pollmann, F.: Data compression for quantum machine learning. Phys. Rev. Res. **4**, 043007 (2022) https://doi.org/10.1103/PhysRevResearch.4.043007

[37] Pomarico, D., Monaco, A., Amoroso, N., Bellantuono, L., Lacalamita, A., La Rocca, M., Maggipinto, T., Pantaleo, E., Tangaro, S., Stramaglia, S., Bellotti, R.: Emerging generalization advantage of quantum-inspired machine learning in the diagnosis of hepatocellular carcinoma. Discover Applied Sciences **7**(3), 205 (2025) https://doi.org/10.1007/s42452-025-06638-6

[38] Repetto, V., Ceroni, E.G., Buonaiuto, G., D'Aurizio, R.: Quantum enhanced stratification of Breast Cancer: exploring quantum expressivity for real omics data (2024). https://arxiv.org/abs/2409.14089

[39] Pomarico, D., Fanizzi, A., Amoroso, N., Bellotti, R., Biafora, A., Bove, S., Didonna, V., La Forgia, D., Pastena, M.I., Tamborra, P., Zito, A., Lorusso, V., Massafra, R.: A Proposal of Quantum-Inspired Machine Learning for Medical Purposes: An Application Case. Mathematics **9**(410) (2021) https://doi.org/10.3390/math9040410

[40] Moutinho, J.a.P., Melo, A., Coutinho, B., Kovács, I.A., Omar, Y.: Quantum link prediction in complex networks. Phys. Rev. A **107**, 032605 (2023) https://doi.org/10.1103/PhysRevA.107.032605

[41] Magano, D., Moutinho, J., Coutinho, B.: On the quantum simulation of complex networks. SciPost Phys. Core **6**, 058 (2023) https://doi.org/10.21468/SciPostPhysCore.6.3.058

[42] Batra, K., Zorn, K.M., Foil, D.H., Minerali, E., Gawriljuk, V.O., Lane, T.R., Ekins, S.: Quantum Machine Learning Algorithms for Drug Discovery Applications. J. Chem. Inf. Model. **61**(6), 2641–2647 (2021) https://doi.org/10.1021/acs.jcim.1c00166

[43] Massafra, R., Comes, M.C., Bove, S., Didonna, V., Diotaiuti, S., Giotta, F., Latorre, A., La Forgia, D., Nardone, A., Pomarico, D., Ressa, C.M., Rizzo, A., Tamborra, P., Zito, A., Lorusso, V., Fanizzi, A.: A machine learning ensemble approach for 5- and 10-year breast cancer invasive disease event classification. PLOS ONE **17**(9), 1–17 (2022) https://doi.org/10.1371/journal.pone.0274691

[44] Fanizzi, A., Pomarico, D., Rizzo, A., Bove, S., Comes, M.C., Didonna, V., Giotta, F., La Forgia, D., Latorre, A., Pastena, M.I., Petruzzellis, N., Rinaldi, L., Tamborra, P., Zito, A., Lorusso, V., Massafra, R.: Machine learning survival models trained on clinical data to identify high risk patients with hormone responsive her2 negative breast cancer. Scientific Reports **13**(1), 8575 (2023) https://doi.org/10.1038/s41598-023-35344-9

[45] Amoroso, N., La Rocca, M., Monaco, A., Bellotti, R., Tangaro, S.: Complex networks reveal early MRI markers of parkinson's disease. Medical Image Analysis **48**, 12–24 (2018) https://doi.org/10.1016/j.media.2018.05.004

[46] Amoroso, N., La Rocca, M., Bellantuono, L., Diacono, D., Fanizzi, A., Lella, E., Lombardi, A., Maggipinto, T., Monaco, A., Tangaro, S., Bellotti, R.: Deep learning and multiplex networks for accurate modeling of brain age. Frontiers in Aging Neuroscience **11** (2019) https://doi.org/10.3389/fnagi.2019.00115

[47] La Rocca, M., Barisano, G., Garner, R., Ruf, S.F., Amoroso, N., Monti, M., Vespa, P., Bellotti, R., Erdoğmuş, D., Toga, A.W., *et al.*: Functional connectivity alterations in traumatic brain injury patients with late seizures. Neurobiology of Disease **179**, 106053 (2023)

[48] Rizzi, R., Jornkokgoud, K., Ghomroudi, P.A., Stella, M., Grecucci, A.: The dark side of the mood: structural and functional fronto-insular and cerebellar alterations classify major depression. medRxiv (2025) https://doi.org/10.1101/2025.04.09.25325506

[49] Jimenez-Marin, A., Diez, I., Erramuzpe, A., Stramaglia, S., Bonifazi, P., Cortes, J.M.: Open datasets and code for multi-scale relations on structure, function and neuro-genetics in the human brain. Scientific Data **11**(1), 256 (2024) https://doi.org/10.1038/s41597-024-03060-2

[50] Lombardi, A., Guaragnella, C., Amoroso, N., Monaco, A., Fazio, L., Taurisano,

P., Pergola, G., Blasi, G., Bertolino, A., Bellotti, R., Tangaro, S.: Modelling cognitive loads in schizophrenia by means of new functional dynamic indexes. NeuroImage **195**, 150–164 (2019) https://doi.org/10.1016/j.neuroimage.2019.03.055

[51] Grecucci, A., Scarano, A., Bruno, F., Salvato, G., Yi, X., Stella, M.: Covarying gray and white matter networks characterize schizophrenia and bipolar disorders on a continuum: A data fusion machine learning approach and a brain network analysis. Journal of Affective Disorders **389**, 119708 (2025) https://doi.org/10.1016/j.jad.2025.119708

[52] Yi, X., Wang, X., Ma, M., Zhang, J., Wu, F., Huang, H., Xiao, Q., Liu, X., Grecucci, A.: Disease and state dependent neural markers in adolescents with bd. understanding the neural bases of mania, depression and remission in a data fusion approach. Psychiatry Research: Neuroimaging **357**, 112113 (2026) https://doi.org/10.1016/j.pscychresns.2025.112113

[53] Amoroso, N., Quarto, S., La Rocca, M., Tangaro, S., Monaco, A., Bellotti, R.: An eXplainability Artificial Intelligence approach to brain connectivity in Alzheimer's disease. Frontiers in Aging Neuroscience **15** (2023) https://doi.org/10.3389/fnagi.2023.1238065

[54] Diez, I., García-Moreno, F., Carral-Sainz, N., Stramaglia, S., Nieto-Reyes, A., D'Amato, M., Cortes, J.M., Bonifazi, P.: Connectivity of the adult human brain with sequential neurogenesis of circuits and transcriptomics signatures. Nature Communications **17**(1), 1035 (2025) https://doi.org/10.1038/s41467-025-67785-3

[55] Jalali, A., Rizzi, R., Ahmadi Ghomroudi, P., Grecucci, A.: Integrating structural and functional brain features to classify major depressive disorder: a multi modal approach. Journal of Affective Disorders **394**, 120528 (2026) https://doi.org/10.1016/j.jad.2025.120528

[56] Yi, X., Wang, X., Zhang, J., Wu, F., Shen, L., Wang, L., Xiao, Q., Grecucci, A.: Structural alterations related to emotion dysregulation, anxiety and self-harm in adolescents borderline personality disorder. a source-based morphometry study. Journal of Affective Disorders **398**, 121035 (2026) https://doi.org/10.1016/j.jad.2025.121035

[57] White, S.R.: Density matrix formulation for quantum renormalization groups. Phys. Rev. Lett. **69**, 2863–2866 (1992) https://doi.org/10.1103/PhysRevLett.69.2863

[58] Schollwöck, U.: The density-matrix renormalization group. Rev. Mod. Phys. **77**, 259–315 (2005) https://doi.org/10.1103/RevModPhys.77.259

[59] Schollwöck, U.: The density-matrix renormalization group in the age of matrix

product states. Annals of Physics **326**(1), 96–192 (2011) https://doi.org/10.1016/j.aop.2010.09.012 . January 2011 Special Issue

[60] Cirac, J.I., Pérez-García, D., Schuch, N., Verstraete, F.: Matrix product states and projected entangled pair states: Concepts, symmetries, theorems. Rev. Mod. Phys. **93**, 045003 (2021) https://doi.org/10.1103/RevModPhys.93.045003

[61] Guo, E., Draper, D.: Neural Tangent Kernel of Matrix Product States: Convergence and Applications (2021). https://arxiv.org/abs/2111.14046

[62] Pomarico, D., Monaco, A., Magnifico, G., Lacalamita, A., Pantaleo, E., Bellantuono, L., Tangaro, S., Maggipinto, T., La Rocca, M., Picardi, E., Amoroso, N., Pesole, G., Stramaglia, S., Bellotti, R.: Grokking as an entanglement transition in tensor network machine learning (2025). https://arxiv.org/abs/2503.10483

[63] Pomarico, D., Cilli, R., Monaco, A., Bellantuono, L., La Rocca, M., Maggipinto, T., Magnifico, G., Ontivero Ortega, M., Pantaleo, E., Tangaro, S., Stramaglia, S., Bellotti, R., Amoroso, N.: Transfer entropy and O-information to detect grokking in tensor network multi-class classification problems. Technologies **13**(10) (2025) https://doi.org/10.3390/technologies13100438

[64] Amico, L., Fazio, R., Osterloh, A., Vedral, V.: Entanglement in many-body systems. Rev. Mod. Phys. **80**, 517–576 (2008) https://doi.org/10.1103/RevModPhys.80.517

[65] Eisert, J., Cramer, M., Plenio, M.B.: Colloquium: Area laws for the entanglement entropy. Rev. Mod. Phys. **82**, 277–306 (2010) https://doi.org/10.1103/RevModPhys.82.277

[66] Kursa, M.B., Rudnicki, W.R.: Feature selection with the boruta package. Journal of Statistical Software **36**(11), 1–13 (2010) https://doi.org/10.18637/jss.v036.i11

[67] Chung, F.R.K.: Spectral Graph Theory. Regional Conference Series in Mathematics, vol. 92. American Mathematical Society, ??? (1997)

[68] Santhoshkumar, C.G., Bindhu, T.K.: Entropy of the adjacency spectrum: A new graph invariant. Journal of Advances in Mathematics and Computer Science **40**(12), 156–167 (2025) https://doi.org/10.9734/jamcs/2025/v40i122081

[69] Fiedler, M.: Algebraic connectivity of graphs. Czechoslovak Mathematical Journal **23**(2), 298–305 (1973)

[70] de Abreu, N.M.M.: Old and new results on algebraic connectivity of graphs. Linear Algebra and its Applications **423**(1), 53–73 (2007) https://doi.org/10.1016/j.laa.2006.08.017 . Special Issue devoted to papers presented at the Aveiro Workshop on Graph Spectra

[71] Barahona, M., Pecora, L.M.: Synchronization in small-world systems. Phys. Rev. Lett. **89**, 054101 (2002) https://doi.org/10.1103/PhysRevLett.89.054101

[72] McGraw, P.N., Menzinger, M.: Laplacian spectra as a diagnostic tool for network structure and dynamics. Phys. Rev. E **77**, 031102 (2008) https://doi.org/10.1103/PhysRevE.77.031102

[73] You, Z., Liu, B.: On the laplacian spectral ratio of connected graphs. Applied Mathematics Letters **25**(10), 1245–1250 (2012) https://doi.org/10.1016/j.aml.2011.09.071

[74] Shapley, L.S.: Notes on the n-person game II: The value of an n-person game (1951)

[75] Lundberg, S.M., Lee, S.-I.: A unified approach to interpreting model predictions. Advances in neural information processing systems **30** (2017)

[76] Newman, M.E.J.: Mathematics of Networks, pp. 8525–8533. Palgrave Macmillan UK, London (2018). https://doi.org/10.1057/978-1-349-95189-5_2565

[77] Santucci, F., Jimenez-Marin, A., Gabrielli, A., Bonifazi, P., Ibáñez-Berganza, M., Gili, T., Cortes, J.M.: Partial correlation as a tool for mapping functional-structural correspondence in human brain connectivity. Network Neuroscience **9**(3), 1065–1086 (2025) https://doi.org/10.1162/NETN.a.22 https://direct.mit.edu/netn/article-pdf/9/3/1065/2530747/netn.a.22.pdf

[78] Qi, Z., Wang, J., Gong, J., Su, T., Fu, S., Huang, L., Wang, Y.: Common and specific patterns of functional and structural brain alterations in schizophrenia and bipolar disorder: a multimodal voxel-based meta-analysis. Journal of Psychiatry and Neuroscience **47**(1), 32–47 (2022) https://doi.org/10.1503/jpn.210111 https://doi.org/10.1503/jpn.210111

[79] Birur, B., Kraguljac, N.V., Shelton, R.C., Lahti, A.C.: Brain structure, function, and neurochemistry in schizophrenia and bipolar disorder—a systematic review of the magnetic resonance neuroimaging literature. npj Schizophrenia **3**(1) (2017) https://doi.org/10.1038/s41537-017-0013-9

[80] Suzuki, M., Nohara, S., Hagino, H., Kurokawa, K., Yotsutsuji, T., Kawasaki, Y., Takahashi, T., Matsui, M., Watanabe, N., Seto, H., Kurachi, M.: Regional changes in brain gray and white matter in patients with schizophrenia demonstrated with voxel-based analysis of mri. Schizophrenia Research **55**(1), 41–54 (2002) https://doi.org/10.1016/S0920-9964(01)00224-9

[81] Sorella, S., Lapomarda, G., Messina, I., Frederickson, J.J., Siugzdaite, R., Job, R., Grecucci, A.: Testing the expanded continuum hypothesis of schizophrenia and bipolar disorder. neural and psychological evidence for shared and distinct mechanisms. NeuroImage: Clinical **23**, 101854 (2019) https://doi.org/10.1016/

j.nicl.2019.101854

[82] Rasero, J., Jimenez-Marin, A., Diez, I., Toro, R., Hasan, M.T., Cortes, J.M.: The neurogenetics of functional connectivity alterations in autism: Insights from subtyping in 657 individuals. Biological Psychiatry **94**(10), 804–813 (2023) https://doi.org/10.1016/j.biopsych.2023.04.014

[83] Salisbury, D.F., Kuroki, N., Kasai, K., Shenton, M.E., McCarley, R.W.: Progressive and interrelated functional and structural evidence of post-onset brain reduction in schizophrenia. Archives of General Psychiatry **64**(5), 521–529 (2007) https://doi.org/10.1001/archpsyc.64.5.521

[84] Hirayasu, Y., McCarley, R.W., Salisbury, D.F., Tanaka, S., Kwon, J.S., Frumin, M., Snyderman, D., Yurgelun-Todd, D., Kikinis, R., Jolesz, F.A., Shenton, M.E.: Planum temporale and heschl gyrus volume reduction in schizophrenia: A magnetic resonance imaging study of first-episode patients. Archives of General Psychiatry **57**(7), 692–699 (2000) https://doi.org/10.1001/archpsyc.57.7.692

[85] Hugdahl, K., Løberg, E.-M., Specht, K., Steen, V.M., Wageningen, H.v., Jørgensen, H.A.: Auditory hallucinations in schizophrenia: the role of cognitive, brain structural and genetic disturbances in the left temporal lobe. Frontiers in Human Neuroscience **2** (2008) https://doi.org/10.3389/neuro.09.006.2007

[86] Wu, H., Dai, G., Aizezi, M., Tang, J., Zou, K., Wu, Y., Wu, X.: Gray matter reduction in bilateral insula mediating adverse psychiatric effects of body mass index in schizophrenia. BMC Psychiatry **22**(1) (2022) https://doi.org/10.1186/s12888-022-04285-4

[87] Moran, L.V., Tagamets, M.A., Sampath, H., O'Donnell, A., Stein, E.A., Kochunov, P., Hong, L.E.: Disruption of anterior insula modulation of large-scale brain networks in schizophrenia. Biological Psychiatry **74** (2013) https://doi.org/10.1016/j.biopsych.2013.02.029

[88] Friedman, N.P., Robbins, T.W.: The role of prefrontal cortex in cognitive control and executive function. Neuropsychopharmacology **47**(1), 72–89 (2021)

[89] Dobri, M.L., Diaz, A.P., Selvaraj, S., Quevedo, J., Walss-Bass, C., Soares, J.C., Sanches, M.: The limits between schizophrenia and bipolar disorder: What do magnetic resonance findings tell us? Behavioral Sciences **12**(3), 78 (2022) https://doi.org/10.3390/bs12030078

[90] Safari, N., Levy, I., Zhang, A., Agarwal, C., Cortes, J.M., Pelphrey, K.A., Van Horn, J.D., Rasero, J.: Brain morphology network alterations in adolescents with autism spectrum disorder: a sex-stratified study. bioRxiv (2026) https://doi.org/10.1101/2025.08.28.672884

[91] Ellison-Wright, I., Bullmore, E.: Anatomy of bipolar disorder and schizophrenia:

A meta-analysis. Schizophrenia Research **117**(1), 1–12 (2010) https://doi.org/10.1016/j.schres.2009.12.022

[92] Song, J., Han, D.H., Kim, S.M., Hong, J.S., Min, K.J., Cheong, J.H., Kim, B.N.: Differences in gray matter volume corresponding to delusion and hallucination in patients with schizophrenia compared with patients who have bipolar disorder. Neuropsychiatric Disease and Treatment **11**, 1211–1219 (2015) https://doi.org/10.2147/NDT.S80438

[93] Salvador, R., Radua, J., Canales-Rodríguez, E.J., Solanes, A., Sarró, S., Goikolea, J.M., Valiente, A., Monté, G.C., Natividad, M.d.C., Guerrero-Pedraza, A., Moro, N., Fernández-Corcuera, P., Amann, B.L., Maristany, T., Vieta, E., McKenna, P.J., Pomarol-Clotet, E.: Evaluation of machine learning algorithms and structural features for optimal mri-based diagnostic prediction in psychosis. PLOS ONE **12**(4), 1–24 (2017) https://doi.org/10.1371/journal.pone.0175683

[94] Aizpurua, B., Palmer, S., Orús, R.: Tensor networks for explainable machine learning in cybersecurity. Neurocomputing **639**, 130211 (2025) https://doi.org/10.1016/j.neucom.2025.130211

[95] Barrat, A., Barthélemy, M., Pastor-Satorras, R., Vespignani, A.: The architecture of complex weighted networks. Proceedings of the National Academy of Sciences **101**(11), 3747–3752 (2004) https://doi.org/10.1073/pnas.0400087101 https://www.pnas.org/doi/pdf/10.1073/pnas.0400087101

[96] Page, L., Brin, S., Motwani, R., Winograd, T.: The pagerank citation ranking : Bringing order to the web (1999)

[97] Brandes, U., Fleischer, D.: Centrality measures based on current flow. In: Diekert, V., Durand, B. (eds.) STACS 2005, pp. 533–544. Springer, Berlin, Heidelberg (2005)

[98] Bozzo, E., Franceschet, M.: Resistance distance, closeness, and betweenness. Social Networks **35**(3), 460–469 (2013) https://doi.org/10.1016/j.socnet.2013.05.003

[99] Bavelas, A.: Communication patterns in task-oriented groups. The Journal of the Acoustical Society of America **22**(6), 725–730 (1950) https://doi.org/10.1121/1.1906679

[100] Sabidussi, G.: The centrality index of a graph. Psychometrika **31**(4), 581–603 (1966) https://doi.org/10.1007/BF02289527