ADARELATE: LEARNING TO ACQUIRE MULTISCALE EVIDENCE FOR EEG EMOTION RECOGNITION

Anonymous authors
Paper under double-blind review

ABSTRACT

Emotion-related EEG structure spans individual channels, geometry-defined regions, and relations between regions, yet the useful scale and target change with each recording. Existing static and sample-adaptive graphs generally instantiate a broad relational representation in one pass, before the value of its constituent computations is known. AdaRelate instead formulates recognition as reliability-anchored sequential hierarchical evidence acquisition. A training-fitted low-order anchor supplies initial class logits. A graph-constrained actor–critic then selects a REGION, CHANNEL, RELATION, or STOP action; only the selected shared expert is evaluated, and its evidence conditions the next decision. A strictly forward accumulator incorporates preceding windows within each trial. Across four datasets, the complete system exceeds the strongest reproduced mean for Accuracy, balanced accuracy, and macro-F1. A matched three-seed evaluation yields 66.81% Accuracy, against 65.87% for one-shot top-k and 65.64% for no-RL routing. Removing the anchor reduces Accuracy by 7.42 points, while forward history contributes 6.06 points. Evidence-nulling diagnostics show that acquired evidence changes subsequent actions and stopping decisions, and four-dataset maps reveal distinct channel, region, and relation allocations. The results establish a hierarchical EEG interface in which a reliable prediction base and sequential reinforcement learning jointly support sample-conditioned emotion recognition.

1 INTRODUCTION

Emotion recognition from EEG is intrinsically hierarchical. An electrode records local activity, groups of electrodes describe broader spatial patterns, and interactions between groups provide relational evidence that cannot be reduced to either endpoint alone. The informative level is not fixed: discriminative channels and spatial organisation vary between individuals and across sessions, trials, and local signal conditions (Zheng & Lu, 2015; Song et al., 2019; 2020a). A useful model must therefore represent channel, region, and inter-region evidence while adapting the computation path to the current sample.

Graph neural networks offer a natural representation, ranging from shared adjacency learning to instance-adaptive, temporal, frequency-aware, and multiscale graphs (Song et al., 2020b;a; Li et al., 2024c; 2025e; Zhou et al., 2025; Liu et al., 2025). Their principal limitation is operational rather than representational. A broad graph is usually constructed or transformed in one pass, so node-, region-, and relation-level computations are entangled before their sample-specific utility is known. Previously acquired evidence cannot determine which expensive computation should follow, and the model does not learn when sufficient evidence has already been obtained. Figure 1 contrasts this one-shot execution pattern with the proposed sequential alternative.

AdaRelate addresses two connected problems. On the EEG side, it makes the spatial hierarchy explicit: REGION actions open coarse geometry groups, CHANNEL actions acquire electrode-specific temporal–spectral evidence within an opened region, and RELATION actions compare two opened regions. The resulting path links local signals, regional summaries, and inter-region interactions to the evolving emotion logits. On the methodological side, graph computation becomes a constrained sequential decision process. A low-order reliability anchor supplies a stable starting prediction; a single actor–critic selects the next legal target; only the chosen shared expert runs; and STOP terminates acquisition when further evidence is not selected.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

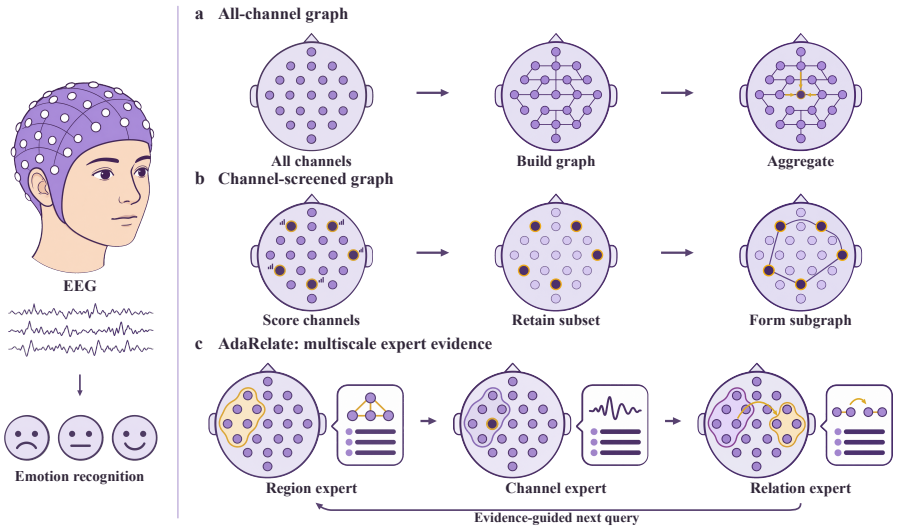

Figure 1: Three EEG graph-inference strategies. (a) All channels are retained for graph construction and message aggregation. (b) Channel screening precedes construction of a selected-channel subgraph. (c) AdaRelate queries REGION, CHANNEL, and inter-region RELATION experts, with acquired evidence conditioning the next query. The shared EEG-to-emotion task is shown on the left. Sensor positions and groups are schematic.

The acquired item updates both the recurrent state and a bounded class-logit residual, allowing later choices to respond to earlier evidence. Trial chronology provides a complementary source: the output accumulator resets at each trial and combines current and preceding windows only. This strictly forward rule exploits affective continuity while preserving the direction of information flow.

The contributions follow these two axes. First, a typed REGION–CHANNEL–RELATION grammar represents EEG evidence at three coupled spatial scales and yields an inspectable path to the emotion prediction. Second, reliability-anchored actor–critic routing resolves which legal computation to execute next and when to stop, rather than fixing the graph before evidence acquisition. Third, four-dataset recognition results are accompanied by matched routing controls, feedback-sensitive RL diagnostics, learned-stopping measurements, and channel–region–relation visualisations. The resulting evidence links recognition gains to the behaviour of the proposed decision process.

2 METHOD

2.1 SEQUENTIAL HIERARCHICAL ACQUISITION

Let $\mathbf{X}_n \in \mathbb{R}^{C\times T\times B}$ denote an EEG sample with $C = 62$ channels, $T = 10$ temporal positions, and $B = 5$ frequency bands; y_n is its emotion label. Deterministic sensor geometry partitions the channels into eight regions and induces 28 unordered region pairs. At acquisition step j, let $\mathcal{O}_j$, $\mathcal{Q}_j$, and $\mathcal{P}_j$ contain the opened regions, queried channels, and queried region pairs. The legal set is

$$\begin{aligned}\mathcal{L}_j = \{R_r : r \notin \mathcal{O}_j\} \cup \{C_c : \rho(c) \in \mathcal{O}_j,\ c \notin \mathcal{Q}_j\} \\ \cup \{E_{rs} : r, s \in \mathcal{O}_j,\ (r, s) \notin \mathcal{P}_j\} \cup \{S : j \geq K_{\min}\},\end{aligned} \tag{1}$$

where R_r, C_c, E_{rs}, and S denote REGION, CHANNEL, RELATION, and STOP actions; $\rho(c)$ returns the region of channel c, and $K_{\min}$ is the first stopping step. A channel is therefore queried only after its region is available, and a relation requires both endpoint regions. The hierarchy is an executable dependency structure rather than a post-hoc grouping.

The actor–critic scores only $\mathcal{L}_j$. A selected action invokes one shared expert and updates the acquisition state and class logits:

$$a_j \sim P_\theta(\cdot \mid \mathbf{h}_{j-1}, \mathcal{L}_j), \qquad \mathbf{e}_j = E_{a_j}(\mathbf{X}_n), \qquad (\mathbf{h}_j, \mathbf{z}_n^{(j)}) = F(\mathbf{h}_{j-1}, \mathbf{z}_n^{(j-1)}, \mathbf{e}_j). \tag{2}$$

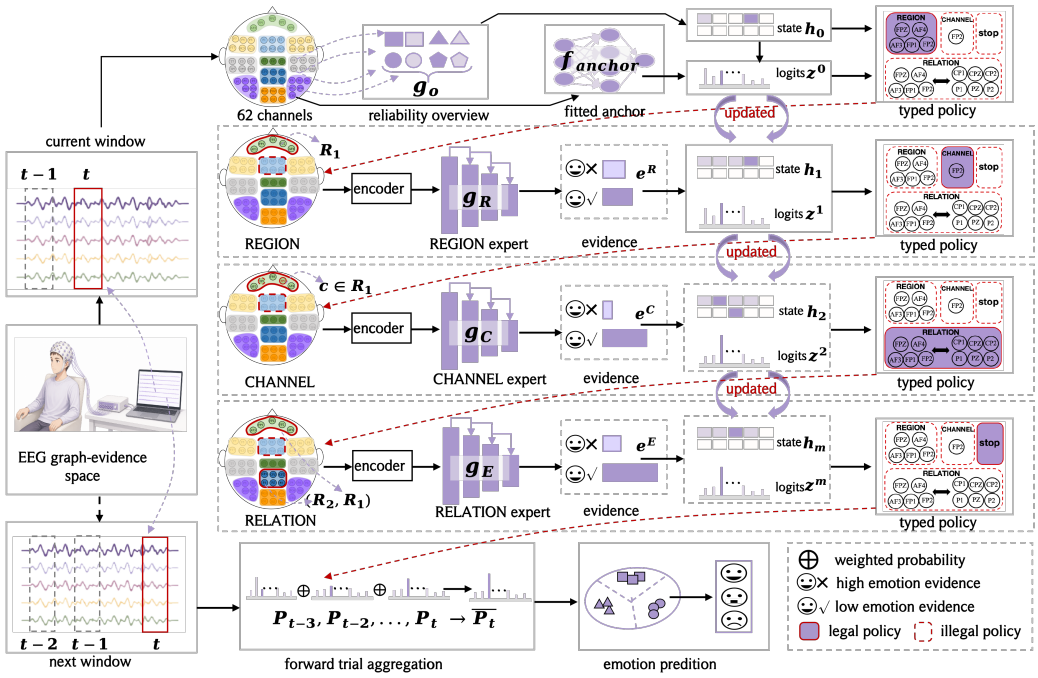

Figure 2: Detailed AdaRelate framework. A reliability overview and fitted anchor initialise the state and zero-action logits. At each step, the actor–critic selects one legal REGION, CHANNEL, RELATION, or STOP target. The selected expert alone returns evidence for the bounded state and logit update; after stopping, the output integrates current and preceding windows within the trial.

Here P_θ is the policy, E_{a_j} is the expert chosen by a_j, and F is the recurrent state–logit transition. Training samples from the policy; evaluation selects the highest-probability legal action. Acquisition ends at STOP or after six evidence actions. No deep expert is executed for an unselected candidate.

2.2 Reliability-Anchored Prediction Base

A low-cost overview describes every channel by six temporal statistics in each band, a robust and learned quality score, and its three-dimensional sensor coordinate. These descriptors initialise $\mathbf{h}_0$. A separate linear anchor is fitted to training-standardised low-order channel or region descriptors and frozen during evaluation. Its zero-action prediction is

$$\mathbf{z}_n^{(0)} = f_{\text{anchor}}(\mathbf{s}(\mathbf{X}_n)) + \alpha_0 f_{\text{overview}}(\mathbf{h}_0), \tag{3}$$

where $\mathbf{s}(\mathbf{X}_n)$ contains low-order descriptors and α_0 is learned. The anchor gives every sample a usable prediction before deep acquisition; subsequent evidence forms a bounded residual rather than replacing this reference. Appendix B specifies normalisation, reliability weighting, anchor fitting, and dataset-frozen views.

2.3 Selected-Only Evidence and Reinforcement Learning

REGION experts encode regional moments and geometry, CHANNEL experts encode the selected electrode's temporal–spectral sequence, and RELATION experts encode endpoint sequences, products, differences, geometry, and lagged correlation signatures. The experts share parameters within type. For selected evidence $\mathbf{e}_j$, a sigmoid vector gate $\mathbf{g}_j$ controls the state update, while a scalar gate η_j and positive type gain control the logit residual:

$$\begin{cases} \mathbf{h}_j = \mathbf{h}_{j-1} + \mathbf{g}_j \odot \tanh(\mathrm{GRU}(\mathbf{e}_j, \mathbf{h}_{j-1}) - \mathbf{h}_{j-1}), \\ \Delta\mathbf{r}_j = \mathrm{softplus}(\gamma_{\tau_j})\eta_j W_e \mathbf{e}_j, \qquad \mathbf{r}_j = \mathbf{r}_{j-1} + \Delta\mathbf{r}_j, \\ \mathbf{z}_n^{(j)} = \mathbf{z}_n^A + \eta_z \tanh(\mathbf{r}_j). \end{cases} \tag{4}$$

Here $\tau_j \in \{R, C, E\}$ is the selected evidence type, W_e maps evidence to class space, $\mathbf{z}_n^A$ denotes anchor logits, and $\eta_z > 0$ bounds the acquired residual. The state makes the next action evidence-conditioned; the bounded output preserves the anchor as the reference prediction.

Every prefix receives class-balanced supervision. The actor is additionally rewarded by the detached reduction in prefix cross-entropy, less the measured cost of the selected action:

$$r_j = \ell(\mathbf{z}_n^{(j-1)}, y_n) - \ell(\mathbf{z}_n^{(j)}, y_n) - \lambda_c \kappa(a_j), \tag{5}$$

where ℓ is classification loss, κ is action cost, and λ_c is its frozen weight. The critic estimates discounted return and the actor uses the resulting advantage (Mnih et al., 2016). Full prefix, distillation, actor, critic, and entropy terms appear in Appendix D.

2.4 Strictly forward trial-history output

Window probabilities are ordered by session, trial, and original index, and the accumulator resets at every trial boundary. For window i, the generic exponential form is

$$\bar{\mathbf{p}}_i = \frac{\sum_{u=1}^{i} \rho_D^{i-u} \mathbf{p}_u}{\sum_{u=1}^{i} \rho_D^{i-u}}, \tag{6}$$

where $\mathbf{p}_u$ is the final-prefix probability at preceding or current window u, and ρ_D is the dataset-frozen decay. The cumulative-mean configuration (MPED) adds an uncertainty-scaled frozen prior. Both forms exclude later windows by construction; exhaustive suffix perturbations verify the implementation. Full definitions appear in Appendix D.

3 Experiments

3.1 Evaluation design

Evaluation uses SEED, MPED, SEED-IV, and SEED-V (Zheng & Lu, 2015; Song et al., 2019; Zheng et al., 2019; Liu et al., 2022), containing 15, 23, 15, and 16 subjects. Each personalised model receives a $62 \times 10 \times 5$ differential-entropy tensor; configuration selection, fresh refitting, trial partitions, frame durations, frequency bands, and dataset-specific output rules are specified in Appendix B. Accuracy, balanced accuracy, and macro-F1 are calculated per subject. MPED folds are concatenated before metric calculation, and repeated seeds are averaged within subject before paired analysis.

Confidence intervals use subject-level bootstrap resampling (Efron & Tibshirani, 1993); paired permutation tests use 100,000 draws with Holm correction within each prespecified Accuracy family (Holm, 1979). The one-shot and no-RL controls preserve the encoder, legal vocabulary, and evidence budget while changing the routing rule. Adaptive inference permits at most six evidence actions, enables STOP after the second action, disables the training-only teacher, and records every expert call. Final numerical configurations and tests are supplied by the anonymised implementation.

3.2 Overall recognition under dataset-frozen configurations

Table 1 places the complete system beside a compact cross-section of reproduced competitors. The subset spans shared-adjacency graph learning (DGCNN), recent graph models (PGCN and HDGNet), and temporal recognition architectures (MASA-TCN and EmT) (Song et al., 2020b; Jin et al., 2024; Li & Kim, 2025; Ding et al., 2024; 2025). The complete 11-baseline, three-metric comparison is reported without filtering in Appendix F. AdaRelate reaches 72.37%, 66.85%, 66.88%, and 80.01% Accuracy on SEED, MPED, SEED-IV, and SEED-V. Relative to the strongest reproduced mean in the complete comparison, the corresponding margins are 5.32, 1.10, 2.28, and 3.65 percentage points. Balanced accuracy and macro-F1 follow the same mean ordering.

The common mean ordering across Accuracy, balanced accuracy, and macro-F1 shows that the complete-system advantage persists under overall, class-balanced, and class-wise evaluation. Because every reproduced method follows the common tensor, partition, and subject-level metric contract, the four datasets provide a consistent benchmark view. A stricter subject-paired analysis then isolates router identity.

Table 1: Representative subject-level Accuracy (mean±sample SD, %). Configurations are frozen per dataset within the common evaluation framework; the appendix reports the complete 11-baseline, three-metric comparison. The AdaRelate row gives the frozen reference run, while Section 3.3 gives the matched three-seed routing study.

Method	SEED	MPED	SEED-IV	SEED-V
DGCNN	64.27±14.68	65.75±8.00	62.14±15.36	61.90±18.85
EmT	59.32±14.94	63.78±7.53	64.12±13.12	70.44±22.85
HDGNet	59.84±15.11	61.53±6.94	59.84±15.40	63.93±21.16
MASA-TCN	67.05±15.33	63.33±7.44	64.60±13.06	76.36±23.70
PGCN	62.36±17.28	63.23±7.42	61.38±15.94	64.87±24.76
AdaRelate	**72.37±15.99**	**66.85±7.95**	**66.88±10.84**	**80.01±19.89**

Figure 3 makes the EEG execution object explicit before the aggregate comparisons: a frozen validation path moves from the reliability-fitted brain state through REGION and CHANNEL evidence to adaptive termination. Section 3.3 analyses the corresponding recognition, feedback, and stopping measurements.

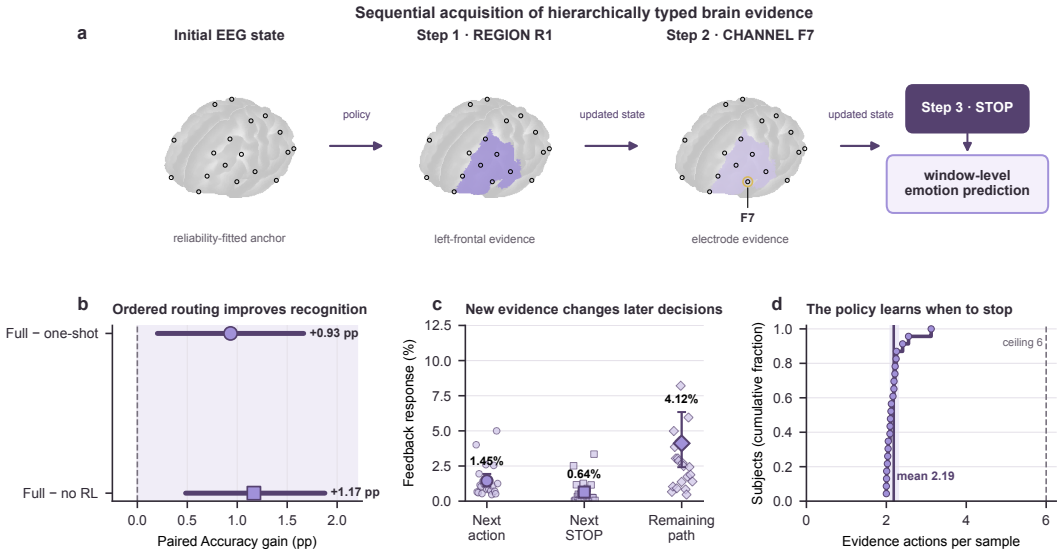

Figure 3: Brain-grounded sequential routing and matched controls. (a) A frozen anchor → REGION R1 → CHANNEL F7 → STOP path. (b) Subject-paired three-seed Accuracy gains over matched controls. (c) Later-decision changes after acquired evidence is zeroed. (d) Subject-level action-count distribution; markers and intervals report means and bootstrap 95% CIs.

Figure 4 expands the representative Accuracy table to the full three-metric comparison against the strongest reproduced mean for each dataset and metric. Every one of the 12 margins is positive. The gains range from $1.10\,\mathrm{pp}$ to $6.64\,\mathrm{pp}$ and persist for BAcc and macro-F1. The result therefore supports a recognition advantage that is not confined to the majority-weighted Accuracy score.

The frozen reference and matched routing analyses use two prespecified summaries. Table 1 reports the 66.85% seed-8635 reference-run estimate, whereas the routing study averages three seeds within subject and gives 66.81% for Full AdaRelate.

3.3 Prediction Base and Sequential-Routing Diagnostics

This section uses the subject-paired MPED analysis for component and routing diagnostics. Removing the fitted anchor reduces Accuracy from 66.85% to 59.43%, a Full-minus-ablation contribution of $7.42\,\mathrm{pp}$ (95% CI $[4.45, 10.08]$; Holm-adjusted $p = 0.0003$). Forward history contributes $6.06\,\mathrm{pp}$

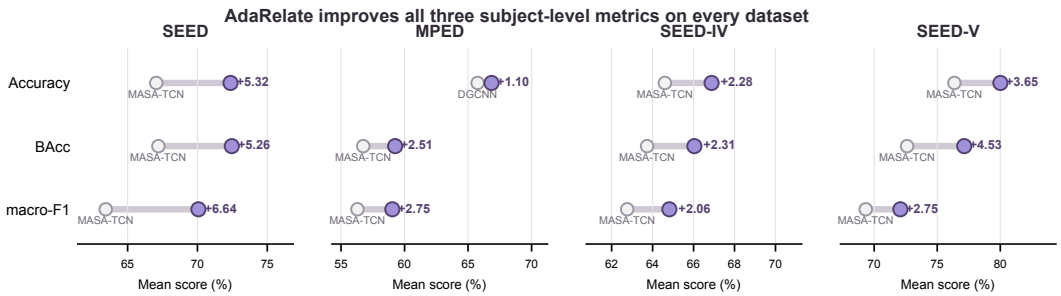

Figure 4: Three-metric recognition advantage on four datasets. Each grey marker is the strongest reproduced mean for that dataset and metric; the purple marker is AdaRelate, and the label gives the absolute margin in percentage points. Full method-by-method means and subject-level standard deviations appear in Appendix F.

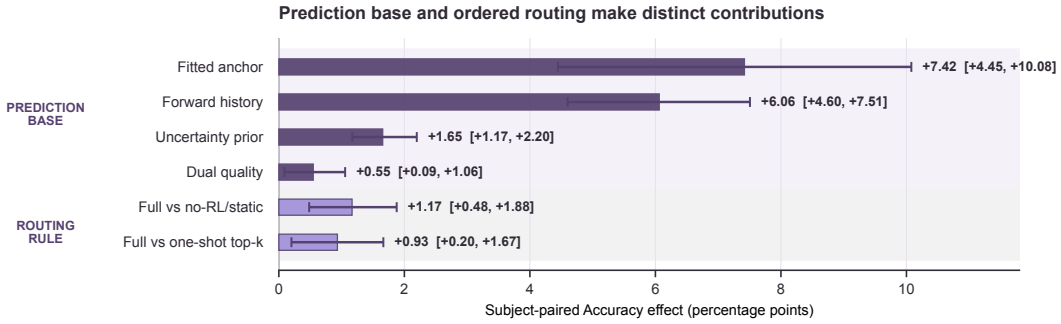

Figure 5: Prediction-base and matched routing effects. Bars give subject-paired Accuracy effects with bootstrap 95% CIs. Component effects use the component-positive orientation; routing effects compare Full AdaRelate with the labelled controls. Contrasts have distinct references and are not summed.

over window-only output without the uncertainty prior (95% CI $[4.60, 7.51]$; adjusted $p < 0.0001$). The uncertainty prior contributes 1.65 pp, while learned-and-robust quality adds 0.55 pp over uniform quality. All 857,808 suffix perturbations leave earlier outputs unchanged, confirming the implemented forward rule.

Figure 5 places these component effects beside the matched routing contrasts on a common percentage-point axis. Each contrast has its own labelled reference condition and is interpreted separately. The ordering reveals a deliberate division of labour: the anchor establishes a reliable low-order decision surface, forward history supplies the principal contextual refinement, and reliability and prior terms make smaller adjustments to the starting state and output. The sequential policy then determines how typed deep evidence is assembled around that prediction base.

This separation is central to the method. The actor–critic receives a stable reference and optimises the ordered acquisition of residual evidence. The anchor defines the starting decision surface, while the policy remains free to change evidence type, target identity, path length, and stopping time as the recurrent state evolves. The component hierarchy therefore connects the two technical axes of the paper: reliability stabilises prediction, while reinforcement learning makes hierarchical graph computation conditional on the evidence already acquired.

Figure 3 grounds the sequential policy in its EEG execution path and then measures its task contribution. Panel a follows one frozen trajectory across the complete cortical surface: the reliability-fitted anchor initialises the state, the policy opens REGION R1, acquires CHANNEL F7, and selects STOP before producing the window-level emotion prediction. Panel b defines task contribution as the subject-paired Accuracy difference from a routing-matched control. Averaging seeds 8635–8637 within subject, Full AdaRelate reaches 66.81% Accuracy, compared with 65.87% for one-shot top-k and 65.64% for no-RL/static routing. The gains are 0.93 pp (95% CI $[0.20, 1.67]$; adjusted

$p = 0.0251$) and 1.17 pp (95% CI $[0.48, 1.88]$; adjusted $p = 0.0086$). Because encoder, vocabulary, and evidence budget remain matched, the comparison isolates the benefit of ordered actor–critic routing.

The remaining panels explain the routing behaviour. When newly acquired evidence is zeroed, the next action changes in 1.45% of eligible Full trajectories, the next STOP decision changes in 0.64%, and the remaining path has a 4.12% normalised edit distance. These rates directly quantify feedback use. The policy also uses 2.19 evidence actions per sample on average against the common ceiling of six. The linked brain path and three diagnostics therefore show that the policy improves emotion recognition, responds to acquired EEG evidence, and learns when to terminate acquisition. Type-resolved target transitions and exact-mask comparisons appear in Appendix J.

3.4 Sensor-grounded typed evidence

Using MPED as the representative case, Figure 6 places the three acquired evidence types in a common EEG coordinate system. CHANNEL utility is displayed at the observed electrodes, REGION utility is assigned to the eight sensor-geometry groups on the same 62-channel scalp montage, and RELATION utility is attached to routed region pairs. The largest channel utilities occur at P4, F8, and PO3. The regional branch assigns positive mean utility across the eight groups, led by the left-anterior, right-anterior, and right-lateral partitions. The relation branch is used more selectively and identifies a small set of beneficial inter-region computations.

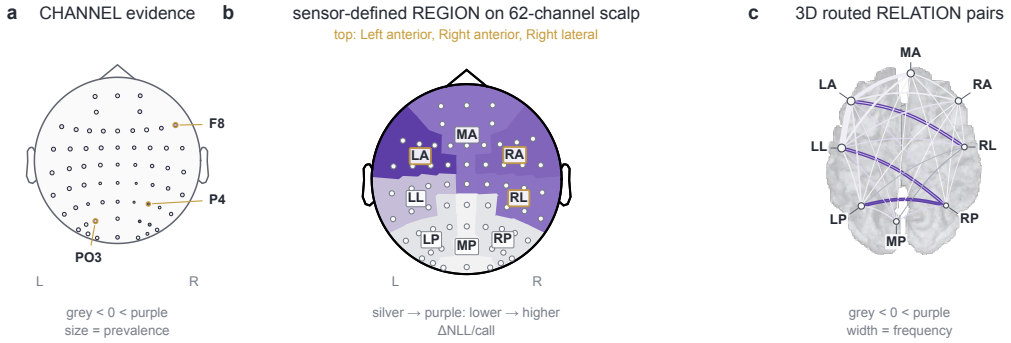

Figure 6: Sensor-grounded typed EEG evidence. (a) CHANNEL colour denotes mean per-call ΔNLL and marker size denotes subject prevalence; gold rings label the three largest utilities. (b) Eight sensor-defined REGION groups are rendered on the measured 62-channel MNE scalp montage; colour denotes mean per-call ΔNLL. (c) Routed RELATION pairs use colour for per-call ΔNLL and width for occurrence frequency. The REGION map is a sensor-space model diagnostic, not a source-localised cortical estimate. Positive ΔNLL denotes an improvement in true-class probability.

The visualisation makes the EEG hierarchy operational rather than nominal. Electrodes, sensor-defined regions, and region pairs correspond to distinct expert calls, yet remain comparable through the same realised-utility measure. This organisation exposes how a decision path moves from coarse spatial evidence to a selected sensor and, where useful, to a relation. Complete four-dataset allocation maps and class-stratified channel views appear in Appendix H.

3.5 Typed contribution and state transitions

On the same representative case, Figure 7 measures the contribution of each evidence type before resolving the ordered transitions. CHANNEL calls provide the largest mean per-call utility ($+0.0066$ ΔNLL), followed by REGION ($+0.0038$) and RELATION ($+0.0009$). REGION→CHANNEL is both frequent (31.8%) and strongly positive ($+0.0066$), while START→REGION contributes $+0.0036$ and accounts for another 31.8% of transitions. CHANNEL→STOP forms 30.1% of transitions. The dominant route therefore follows the intended hierarchy: initialise a regional context, acquire selected electrode evidence, and terminate once the state is sufficient.

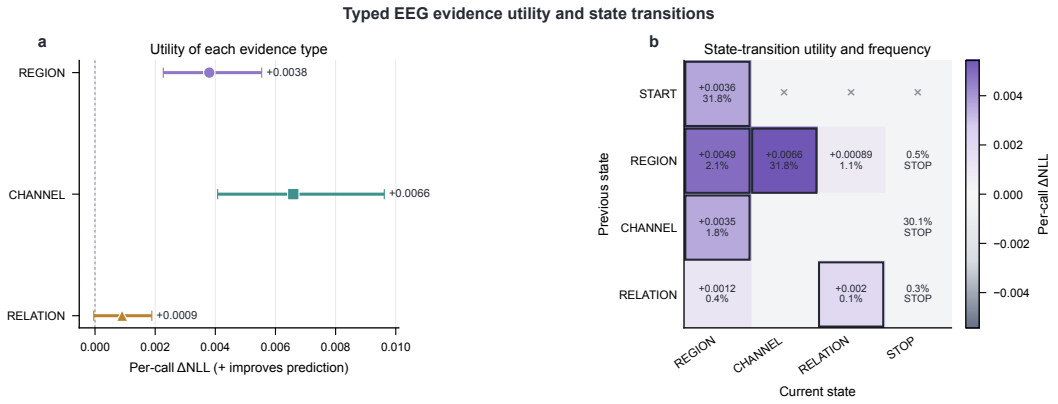

Figure 7: Typed evidence utility and transition structure. (a) Mean per-call ΔNLL for REGION, CHANNEL, and RELATION evidence; intervals are subject-bootstrap 95% CIs. (b) Cell colour gives per-call ΔNLL and text gives transition frequency. STOP cells report frequency only; outlined utility cells have 95% CIs excluding zero.

Together with the matched Accuracy gains in Figure 3b, these measurements connect the method's two central claims. The actor–critic improves the final recognition result, while the utility-resolved path shows which hierarchical state changes carry predictive evidence. Complete four-dataset utility, ablation, and transition analyses are reported in Appendix J.

3.6 Evidence synthesis

Across the four datasets, all 12 margins over the strongest reproduced means are positive and range from $1.10\,\mathrm{pp}$ to $6.64\,\mathrm{pp}$. Accuracy does not carry this ordering alone: balanced accuracy and macro-F1 retain it under class-balanced and class-wise evaluation. The comparison also covers three-, four-, and five-class vocabularies drawn from distinct recording populations, showing that the typed interface adapts to several affective label spaces without imposing one sensor graph on every recording.

Matched controls separate the responsibilities of the prediction base and the router. The $7.42\,\mathrm{pp}$ anchor effect establishes a reliable initial decision surface; the $6.06\,\mathrm{pp}$ forward-history effect shows how much chronologically available context contributes to that base. With encoder, vocabulary, and evidence budget held constant, ordered actor–critic routing then adds $0.93\,\mathrm{pp}$ over one-shot top-k acquisition and $1.17\,\mathrm{pp}$ over no-RL/static routing. The architecture assigns a different role to each component: the anchor stabilises the initial estimate, the forward accumulator retains preceding evidence, and the policy selects the next deep computation. Their measured effects identify a coordinated prediction-and-control design rather than a single monolithic graph encoder.

Trajectory measurements connect the routing gain to its execution. The policy invokes 2.19 evidence actions per sample on average, and evidence nulling changes the next action, the next STOP decision, or the remaining path. Most transitions follow START→REGION, REGION→CHANNEL, and CHANNEL→STOP. Per-call utility is largest for CHANNEL evidence, followed by REGION and the more selectively invoked RELATION computations; the spatial views place these quantities at electrodes, sensor-defined scalp regions, and routed region pairs. Node, region, and relation thus denote separately executed operations, each with an observable contribution and an explicit position in the decision sequence.

These findings support both axes of the paper. For EEG emotion recognition, AdaRelate represents channels, regions, and inter-region relations as separately executed evidence operations. For adaptive computation, it learns the order and stopping point of those operations from previously acquired evidence. Unlike one-shot graph inference, which fixes computation before its sample-specific value is known, the proposed predictor retains a stable reference while adapting evidence depth, target, and type. Recognition margins, matched routing controls, transition-resolved utility, and sensor-space

maps jointly show that entangled multiscale evidence and fixed computation are replaced by an observable, evidence-conditioned sequence.

4 Discussion

Variation across people, sessions, and local signal conditions makes EEG emotion recognition a computation-allocation problem as well as a representation problem. AdaRelate avoids assuming that one densely instantiated graph is equally informative for every recording. Its reliability-fitted anchor absorbs predictable low-order structure, while the policy spends the remaining evidence budget on sample-specific region, channel, or relation computation.

The formulation differs from both one-shot adaptive graphs and generic active observation. Its observations are typed EEG computations rather than interchangeable crops or tokens; CHANNEL and RELATION become legal only after their regional prerequisites are opened; and each selected expert changes the state used to choose the next target and STOP. Table 2 collects the evidence for this prediction-and-control account.

Table 2: Claim–evidence ledger for the subject-paired matched analysis ($n = 23$, four folds merged). Accuracy contrasts use subject means after three-seed averaging where applicable.

Design role	Frozen observation	Supported conclusion
Fitted anchor	Full minus w/o anchor: $+7.42$ pp, CI $[4.45, 10.08]$, $p = 0.0003$	Principal measured source of the stable prediction base
Forward history	History without prior minus window only: $+6.06$ pp, CI $[4.60, 7.51]$, $p < 0.0001$	Past trial context supplies the main output refinement
Sequential order	Full minus one-shot: $+0.93$ pp, CI $[0.20, 1.67]$, $p = 0.0251$	Ordered acquisition improves matched fixed ranking
RL selection	Full minus no-RL: $+1.17$ pp, CI $[0.48, 1.88]$, $p = 0.0086$	Actor–critic improves matched static routing
Feedback use	Next action 1.45%; next STOP 0.64%; remaining-path edit 4.12%	Acquired evidence changes later decisions
Adaptive stopping	2.19 evidence actions per sample; ceiling 6	The learned policy usually terminates before the maximum horizon
Future exclusion	857,808 suffix checks; maximum error 0	Executed output does not read later windows

The learned graph is therefore an execution-level evidence graph: its nodes, typed edges, and order record which classifier computations were invoked and how they changed the prediction. Brainspace maps expose this allocation rather than estimating causal neural connectivity. Under this interpretation, one REGION–CHANNEL–RELATION–STOP interface links recognition, graph structure, reinforcement learning, and auditability without assigning physiological meaning beyond measured discriminative evidence.

5 Conclusion

AdaRelate formulates hierarchical EEG graph computation as reliability-anchored REGION, CHANNEL, RELATION, and STOP decisions. Across four datasets, it exceeds the strongest reproduced mean for Accuracy, balanced accuracy, and macro-F1; matched three-seed comparisons favour ordered routing over one-shot and supervised-static alternatives. Feedback-sensitive trajectories and multiscale maps link recognition to an inspectable, sample-conditioned control process. The result is traceable graph inference whose targets and depth adapt to the current recording.

AI USE STATEMENT

The human-portrait components in Figures 1 and 2 are AI-generated. Their scientific content, annotations, and interpretation are specified and verified by the authors.

ETHICS STATEMENT

No new human-participant data are collected. The study uses released EEG datasets under their stated access conditions. Ethical approval, consent, and data governance for the original recordings remain those reported by the dataset providers (Zheng & Lu, 2015; Song et al., 2019; Zheng et al., 2019; Liu et al., 2022).

REPRODUCIBILITY STATEMENT

The anonymised implementation is available at `https://anonymous.4open.science/r/adarelate-5CE4/`. It contains the source code, concise setup instructions, final configurations, and formal tests, but excludes datasets, result artefacts, and manuscript files. Dataset licensing requires users to obtain the EEG data from their providers. The appendix specifies the processing, split, statistical, and interpretation contracts used in the reported analyses.

REFERENCES

Bianna Chen, C. L. Philip Chen, and Tong Zhang. GDDN: Graph domain disentanglement network for generalizable EEG emotion recognition. *IEEE Transactions on Affective Computing*, 15(3): 1739–1753, 2024. doi: 10.1109/TAFFC.2024.3371540.

Bianna Chen, C. L. Philip Chen, and Tong Zhang. UGAN: Uncertainty-guided graph augmentation network for EEG emotion recognition. *IEEE Transactions on Computational Social Systems*, 12 (2):695–707, 2025a. doi: 10.1109/TCSS.2024.3488201.

C. L. Philip Chen, Bianna Chen, and Tong Zhang. AdamGraph: Adaptive attention-modulated graph network for EEG emotion recognition. *IEEE Transactions on Cybernetics*, 55(5):2038–2051, 2025b. doi: 10.1109/TCYB.2025.3550191.

Victor-Alexandru Darvariu, Stephen Hailes, and Mirco Musolesi. Goal-directed graph construction using reinforcement learning. *Proceedings of the Royal Society A*, 477(2254):20210168, 2021. doi: 10.1098/rspa.2021.0168.

Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. Convolutional neural networks on graphs with fast localized spectral filtering. In *Advances in Neural Information Processing Systems*, volume 29, 2016.

Yi Ding, Su Zhang, Chuangao Tang, and Cuntai Guan. MASA-TCN: Multi-anchor space-aware temporal convolutional neural networks for continuous and discrete EEG emotion recognition. *IEEE Journal of Biomedical and Health Informatics*, 28(7):3953–3964, 2024. doi: 10.1109/JBHI.2024.3392564.

Yi Ding, Chengxuan Tong, Shuailei Zhang, Muyun Jiang, Yong Li, Kevin JunLiang Lim, and Cuntai Guan. EmT: A novel transformer for generalized cross-subject EEG emotion recognition. *IEEE Transactions on Neural Networks and Learning Systems*, 36(6):10381–10393, 2025. doi: 10.1109/TNNLS.2025.3552603.

Bradley Efron and Robert J. Tibshirani. *An Introduction to the Bootstrap*. Chapman and Hall/CRC, 1993.

Hongxiang Gao, Xingyao Wang, Zhenghua Chen, Min Wu, Zhipeng Cai, Lulu Zhao, Jianqing Li, and Chengyu Liu. Graph convolutional network with connectivity uncertainty for EEG-based emotion recognition. *IEEE Journal of Biomedical and Health Informatics*, 28(10):5917–5928, 2024. doi: 10.1109/JBHI.2024.3416944.

Wei Gu, Wenxin Chen, Bao-Liang Lu, and Wei-Long Zheng. BrainTune: An artifact-aware meta-contrastive framework for fast and lasting personalization of EEG emotion models. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 48:1189–1196, 2026. URL `https://escholarship.org/uc/item/2mt8635v`.

Yun Gu, Xinyue Zhong, Cheng Qu, Chuanjun Liu, and Bin Chen. A domain generative graph network for EEG-based emotion recognition. *IEEE Journal of Biomedical and Health Informatics*, 27(5): 2377–2386, 2023. doi: 10.1109/JBHI.2023.3242090.

Sture Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979.

Ming Jin, Changde Du, Huiguang He, Ting Cai, and Jinpeng Li. PGCN: Pyramidal graph convolutional network for EEG emotion recognition. *IEEE Transactions on Multimedia*, 26:9070–9082, 2024. doi: 10.1109/TMM.2024.3385676.

Dominik Klepl, Min Wu, and Fei He. Graph neural network-based EEG classification: A survey. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 32:493–503, 2024. doi: 10.1109/TNSRE.2024.3355750.

Chao Li, Ning Bian, Ziping Zhao, Haishuai Wang, and Björn W. Schuller. Multi-view domain-adaptive representation learning for EEG-based emotion recognition. *Information Fusion*, 104: 102156, 2024a. doi: 10.1016/j.inffus.2023.102156.

Chao Li, Ning Bian, Ziping Zhao, Haishuai Wang, and Björn W. Schuller. Multi-view domain-adaptive representation learning for EEG-based emotion recognition. *Information Fusion*, 104: 102156, 2024b. doi: 10.1016/j.inffus.2023.102156.

Chao Li, Feng Wang, Ziping Zhao, Haishuai Wang, and Björn W. Schuller. Attention-based temporal graph representation learning for EEG-based emotion recognition. *IEEE Journal of Biomedical and Health Informatics*, 2024c. doi: 10.1109/JBHI.2024.3395622.

Cunbo Li, Tian Tang, Yue Pan, Lei Yang, Shuhan Zhang, Zhaojin Chen, Peiyang Li, Dongrui Gao, Huafu Chen, Fali Li, Dezhong Yao, Zehong Cao, and Peng Xu. An efficient graph learning system for emotion recognition inspired by the cognitive prior graph of EEG brain network. *IEEE Transactions on Neural Networks and Learning Systems*, 36(4):7130–7144, 2025a. doi: 10.1109/TNNLS.2024.3405663.

Cunbo Li, Shuhan Zhang, Yufeng Mu, Lei Yang, Yueheng Peng, Fali Li, Yangsong Zhang, Zhen Liang, Zehong Cao, Feng Wan, Dezhong Yao, Peiyang Li, and Peng Xu. Emotion recognition by learning the manifold of fused multiscale information of EEG signals. *IEEE Transactions on Affective Computing*, 16(3):2172–2188, 2025b. doi: 10.1109/TAFFC.2025.3555226.

Dongdong Li, Zhishuo Jin, Yujun Shen, Zhe Wang, and Suo Jiang. EEG emotion recognition based on an implicit emotion regulatory mechanism. *IEEE Transactions on Artificial Intelligence*, 6 (11):3005–3017, 2025c. doi: 10.1109/TAI.2025.3560593.

Hanyu Li and Byung Hyung Kim. Hierarchical dynamic local–global-graph representation learning for EEG emotion recognition. *IEEE Transactions on Instrumentation and Measurement*, 74:1–14, 2025. doi: 10.1109/TIM.2025.3635321.

Qilin Li, Tong Zhang, C. L. Philip Chen, Ke Yi, and Long Chen. Residual GCB-Net: Residual graph convolutional broad network on emotion recognition. *IEEE Transactions on Cognitive and Developmental Systems*, 15(4):1673–1685, 2023a. doi: 10.1109/TCDS.2022.3147839.

Qilin Li, Tong Zhang, C. L. Philip Chen, Xiaowei Zhang, and Bin Hu. DGC-Link: Dual-gate chebyshev linkage network on EEG emotion recognition. *IEEE Transactions on Affective Computing*, 16(4):3499–3511, 2025d. doi: 10.1109/TAFFC.2025.3602823.

Wei Li, Mingming Wang, Junyi Zhu, and Aiguo Song. EEG-based emotion recognition using trainable adjacency relation driven graph convolutional network. *IEEE Transactions on Cognitive and Developmental Systems*, 15(4):1656–1672, 2023b. doi: 10.1109/TCDS.2023.3270170.

Yueyang Li, Shengyu Gong, Weiming Zeng, Nizhuan Wang, and Wai Ting Siok. FreqDGT: Frequency-adaptive dynamic graph networks with transformer for cross-subject EEG emotion recognition, 2025e.

Hanwen Liu, Yifeng Gong, Zuwei Yan, Zeheng Zhuang, and Jiaxuan Lu. MSGM: A multi-scale spatiotemporal graph mamba for EEG emotion recognition, 2025.

Wei Liu, Jie-Lin Qiu, Wei-Long Zheng, and Bao-Liang Lu. Comparing recognition performance and robustness of multimodal deep learning models for multimodal emotion recognition. *IEEE Transactions on Cognitive and Developmental Systems*, 14(2):715–729, 2022. doi: 10.1109/TCDS.2021.3071170.

Volodymyr Mnih, Adrià Puigdomènech Badia, Mehdi Mirza, Alex Graves, Timothy Lillicrap, Tim Harley, David Silver, and Koray Kavukcuoglu. Asynchronous methods for deep reinforcement learning. In *Proceedings of the 33rd International Conference on Machine Learning*, pp. 1928–1937, 2016.

Tongjie Pan, Yalan Ye, Yangwuyong Zhang, Kunshu Xiao, and Hecheng Cai. Online multi-hypergraph fusion learning for cross-subject emotion recognition. *Information Fusion*, 108:102338, 2024. doi: 10.1016/j.inffus.2024.102338.

Shadi Sartipi, Mastaneh Torkamani-Azar, and Mujdat Cetin. A hybrid end-to-end spatiotemporal attention neural network with graph-smooth signals for EEG emotion recognition. *IEEE Transactions on Cognitive and Developmental Systems*, 16(2):732–743, 2024. doi: 10.1109/TCDS.2023.3293321.

Tengfei Song, Wenming Zheng, Cheng Lu, Yuan Zong, Xilei Zhang, and Zhen Cui. MPED: A multi-modal physiological emotion database for discrete emotion recognition. *IEEE Access*, 7: 12177–12191, 2019. doi: 10.1109/ACCESS.2019.2891579.

Tengfei Song, Suyuan Liu, Wenming Zheng, Yuan Zong, and Zhen Cui. Instance-adaptive graph for EEG emotion recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pp. 2701–2708, 2020a. doi: 10.1609/aaai.v34i03.5656.

Tengfei Song, Wenming Zheng, Peng Song, and Zhen Cui. EEG emotion recognition using dynamical graph convolutional neural networks. *IEEE Transactions on Affective Computing*, 11(3):532–541, 2020b. doi: 10.1109/TAFFC.2018.2817622.

Tengfei Song, Suyuan Liu, Wenming Zheng, Yuan Zong, Zhen Cui, Yang Li, and Xiaoyan Zhou. Variational instance-adaptive graph for EEG emotion recognition. *IEEE Transactions on Affective Computing*, 14(1):343–356, 2023. doi: 10.1109/TAFFC.2021.3064940.

Yilin Wang, Li Zhang, and Yan Zhang. Band-level adaptive fusion network for cross-subject EEG emotion recognition. *IEEE Transactions on Instrumentation and Measurement*, 74:1–12, 2025a. doi: 10.1109/TIM.2025.3544334.

Yulin Wang, Zhaoxi Chen, Haojun Jiang, Shiji Song, Yizeng Han, and Gao Huang. Adaptive focus for efficient video recognition. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 16249–16258, 2021. doi: 10.1109/ICCV48922.2021.01594.

Yulin Wang, Yang Yue, Yuanze Lin, Haojun Jiang, Zihang Lai, Victor Kulikov, Nikita Orlov, Humphrey Shi, and Gao Huang. AdaFocus V2: End-to-end training of spatial dynamic networks for video recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 20062–20072, 2022. doi: 10.1109/CVPR52688.2022.01943.

Yulin Wang, Yang Yue, Huanqian Wang, Haojun Jiang, Yizeng Han, Zanlin Ni, Yifan Pu, Minglei Shi, Rui Lu, Qisen Yang, Andrew Zhao, Zhuofan Xia, Shiji Song, and Gao Huang. Emulating human-like adaptive vision for efficient and flexible machine visual perception. *Nature Machine Intelligence*, 2025b. doi: 10.1038/s42256-025-01130-7.

Mengqi Wu, C. L. Philip Chen, Bianna Chen, and Tong Zhang. Grop: Graph orthogonal purification network for EEG emotion recognition. *IEEE Transactions on Affective Computing*, 16(1): 319–332, 2025. doi: 10.1109/TAFFC.2024.3433613.

Yushun Xiao, Wenming Zheng, and Guoying Zhao. Dynamical causal graph neural network for EEG emotion recognition. *IEEE Transactions on Affective Computing*, 16(4):2803–2815, 2025. doi: 10.1109/TAFFC.2025.3582740.

Huachao Yan, Kailing Guo, Xiaofen Xing, and Xiangmin Xu. Bridge graph attention based graph convolution network with multi-scale transformer for EEG emotion recognition. *IEEE Transactions on Affective Computing*, 15(4):2042–2054, 2024. doi: 10.1109/TAFFC.2024.3394873.

Hua Yang, C. L. Philip Chen, Bianna Chen, and Tong Zhang. Improving the interpretability through maximizing mutual information for EEG emotion recognition. *IEEE Transactions on Affective Computing*, 16(2):744–757, 2025. doi: 10.1109/TAFFC.2024.3463469.

Mengqing Ye, C. L. Philip Chen, Wenming Zheng, and Tong Zhang. Adaptive dual-space network with multigraph fusion for EEG-based emotion recognition. *IEEE Transactions on Computational Social Systems*, 11(5):5763–5774, 2024. doi: 10.1109/TCSS.2024.3386621.

Yongtao Zhang, Yue Pan, Yulin Zhang, Linling Li, Li Zhang, Gan Huang, Zhen Liang, and Zhiguo Zhang. Unsupervised time-aware sampling network with deep reinforcement learning for EEG-based emotion recognition, 2022.

Wei-Long Zheng and Bao-Liang Lu. Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Transactions on Autonomous Mental Development*, 7(3):162–175, 2015. doi: 10.1109/TAMD.2015.2431497.

Wei-Long Zheng, Wei Liu, Yifei Lu, Bao-Liang Lu, and Andrzej Cichocki. EmotionMeter: A multimodal framework for recognizing human emotions. *IEEE Transactions on Cybernetics*, 49 (3):1110–1122, 2019. doi: 10.1109/TCYB.2018.2797176.

Rushuang Zhou, Zhiguo Zhang, Hong Fu, Li Zhang, Linling Li, Gan Huang, Fali Li, Xin Yang, Yining Dong, Yuan-Ting Zhang, and Zhen Liang. PR-PL: A novel prototypical representation based pairwise learning framework for emotion recognition using EEG signals. *IEEE Transactions on Affective Computing*, 15(2):657–670, 2024. doi: 10.1109/TAFFC.2023.3288118.

Yaru Zhou, Xueying Zhang, Ying Sun, Guijun Chen, Lixia Huang, and Haifeng Li. EEG emotion recognition based on dynamic temporal causal graph convolutional network. *Knowledge-Based Systems*, 323:113752, 2025. doi: 10.1016/j.knosys.2025.113752.

A Related work and positioning

A.1 Static and sample-adaptive EEG graphs

Early EEG graph models treat channels as nodes and learn a common adjacency or spectral filter, while a recent review traces the subsequent move towards dynamic, heterogeneous, and multi-view formulations (Defferrard et al., 2016; Song et al., 2020b; Klepl et al., 2024). Instance-adaptive graph learning relaxes population-level sharing by predicting sample-specific connections (Song et al., 2020a; 2023; Li et al., 2023b). Temporal and multiscale models then combine spatial relations with ordered segments, adaptive neighbourhoods, causal structure, or frequency-specific representations (Li et al., 2024c; Zhou et al., 2025; Li et al., 2025e; Liu et al., 2025; Sartipi et al., 2024; Yan et al., 2024; Jin et al., 2024; Xiao et al., 2025). These studies establish the value of sample-adaptive relational representations. AdaRelate advances the execution model by determining which typed computation is instantiated next before evaluating the remaining high-capacity candidates.

Cross-session variation motivates domain-generalised graphs, dual-space representations, connectivity uncertainty, uncertainty-aware adjacency, and multi-view alignment (Gu et al., 2023; Chen et al., 2024; Gao et al., 2024; Ye et al., 2024; Chen et al., 2025a; Li et al., 2024b; Pan et al., 2024). Topology, bands, manifolds, and class-relevant structure are also regularised explicitly (Wu et al., 2025; Chen et al., 2025b; Wang et al., 2025a; Yang et al., 2025; Li et al., 2025b;c). These methods improve the representation from which a classifier reasons. The proposed action path addresses a complementary conditional-computation problem by separating evidence already encoded at high capacity from candidates whose experts remain uncalled.

A.2 Active observation and stopping

In active perception, computation depends on the current observation. AdaFocus selects informative image regions with a recurrent policy, and its differentiable successor retains conditional observation while simplifying training (Wang et al., 2021; 2022). AdaptiveNN extends the formulation to sequential fixation, accumulated state, and learned stopping (Wang et al., 2025b). For EEG, TAS-Net applies deep reinforcement learning to temporal-fragment selection (Zhang et al., 2022). AdaRelate changes the decision object from where to look to which prerequisite-constrained EEG computation to instantiate next: a region, a channel within an opened region, a relation between opened regions, or STOP.

A.3 Sequential graph construction

Graph construction can itself be formulated as a decision process. Goal-directed graph construction uses reinforcement learning to add edges in pursuit of a global structural objective (Darvariu et al., 2021). In AdaRelate, the graph is instead a transient record of classifier computations, and each action invokes an evidence expert tied to the current EEG sample. Classification improvement and action cost therefore train a policy over a typed, reliability-anchored EEG hierarchy. This formulation combines the representational motivation of adaptive EEG graphs with the evidence-conditioned ordering and stopping mechanisms of active observation.

B Notation, data contract, and estimands

B.1 Notation and statistical unit

The appendix specifies the implementation and evaluation contract. A sample index is denoted by n, channel by c, temporal position by t, frequency band by b, geometry region by r or s, class by k, and acquisition step by j. The fixed dimensions are $C = 62$, $T = 10$, $B = 5$, $R = 8$, hidden width $d = 48$, and maximum horizon $K_{\max} = 6$. The action types are REGION (R), CHANNEL (C),

RELATION (E), and STOP (S). A dynamic evidence graph is the typed action path executed for a sample, providing a sparse and inspectable record of classifier computation.

The subject is the independent unit for every inferential result. SEED, SEED-IV, and SEED-V contribute one outer value per subject. MPED contributes four folds per subject, but their frozen predictions are concatenated before Accuracy, balanced accuracy, and macro-F1 are calculated. Fold-wise values are never entered as independent observations. When three seeds are available, each subject's metric is first averaged over seeds; the paired contrast is then formed across the 23 subject means. This order prevents either folds or repeated random initialisations from inflating the effective sample size.

The input contract uses released or officially derived differential-entropy features at their dataset-native temporal resolution. SEED and MPED use non-overlapping one-second frames, whereas SEED-IV and SEED-V use non-overlapping four-second Hanning windows. The five bands are delta (1–4 Hz), theta (4–8 Hz), alpha (8–14 Hz), beta (14–31 Hz), and gamma (31–50 Hz). A length-ten sliding view with unit stride therefore spans 10 seconds for SEED and MPED and 40 seconds for SEED-IV and SEED-V. Each view retains ten consecutive temporal positions and yields a $62 \times 10 \times 5$ sample. Sliding views never cross a trial boundary. Channel–band standardisation is fitted on the selection-training partition and is refitted on training plus development data only after all choices are frozen.

SEED uses all trials from session 1, trials 1–12 from session 2, and trials 1–3 from session 3 for selection training. Session-2 trials 13–15 form development data, and session-3 trials 4–15 form the outer partition. For SEED-IV and SEED-V, selection training contains all of session 1, every session-2 trial except the final occurrence of each class, and the first session-3 trial of each class. The held-out final session-2 occurrence of each class forms development data; all remaining session-3 trials form the outer partition. The rule is defined by class occurrence rather than a dataset-specific list of trial numbers because the class orders differ.

MPED contains four trials for each of seven discrete emotions. Following its three-class protocol, joy and funny are mapped to Positive; anger, sadness, fear, and disgust are mapped to Negative; and neutral remains Neutral (Song et al., 2019). In outer fold $f \in \{1, \ldots, 4\}$, the fth trial of every emotion is assigned to the outer partition, the next trial in cyclic order is assigned to development, and the remaining two trials are assigned to selection training. Each fold therefore contains 14 training, seven development, and seven outer trials. Fresh refitting combines the training and development partitions. The four outer prediction sets are subsequently concatenated within subject before the three reported metrics are calculated.

All reported baselines, ablations, paired controls, replay checks, and repeated-seed analyses follow the same fixed evaluation protocol. The anonymous repository contains the corresponding implementation and configurations.

B.2 Input standardisation and low-cost channel sketch

For training-partition mean μ_{cb} and standard deviation σ_{cb}, channel–band normalisation is

$$\widetilde{X}_{nctb} = \frac{X_{nctb} - \mu_{cb}}{\max(\sigma_{cb}, \epsilon)}. \tag{7}$$

Here $\epsilon > 0$ prevents division by zero. The overview then describes each channel and band by six temporal statistics. With τ_t evenly spaced on $[-1, 1]$, the sketch is

$$\begin{cases} \mathcal{S}(\mathbf{x}) = \left[\mathrm{mean}_t\, x_t, \mathrm{sd}_t\, x_t, \max\limits_t x_t, \min\limits_t x_t, T^{-1}\sum_t \tau_t x_t, \mathrm{rms}_{t\geq 2}(x_t - x_{t-1})\right], \\ \mathbf{m}_{nc} = \mathrm{concat}_{b=1}^{B}\, \mathcal{S}(\widetilde{\mathbf{x}}_{ncb}), \qquad \widetilde{\mathbf{x}}_{ncb} = (\widetilde{X}_{nctb})_{t=1}^{T}. \end{cases} \tag{8}$$

The first two components of $\mathcal{S}(\widetilde{\mathbf{x}}_{ncb})$ are denoted by $\bar{X}_{ncb}$ and s_{ncb}. The statistics retain level, dispersion, extrema, slow change, and local variation at linear channel complexity. They provide low-cost candidate descriptors derived from the observed signal. The shared channel token combines the sketch with the standardised three-dimensional sensor coordinate $\widetilde{\mathbf{p}}_c$:

$$\mathbf{h}_{nc}^{\mathrm{ch}} = \mathrm{LN}(g_m(\mathbf{m}_{nc}) + g_p(\widetilde{\mathbf{p}}_c)) \in \mathbb{R}^d. \tag{9}$$

The maps g_m and g_p are learned projections and LN denotes layer normalisation. Target identity enters through the current signal and continuous geometry, with projection and expert parameters shared across channels. The candidate count is consequently distinct from the number of expert networks.

B.3 Reliability weighting and the initial state

The learned quality branch is complemented by a robust within-sample dispersion score. Let $v_{nc} = B^{-1}\sum_b s_{ncb}$ and let δ_{nc} be its absolute deviation from the channel median. The implemented reliability is

$$\begin{cases} q^L_{nc} = \sigma(f_q(\mathbf{h}^{\mathrm{ch}}_{nc})), \qquad \tau_n = \max\{\mathrm{med}_i\, \delta_{ni}, \tau_{\min}\}, \\ q^R_{nc} = \mathrm{clip}_{[q_{\min},1]} \exp\!\left(-\dfrac{\delta_{nc}}{\beta_q \tau_n}\right), \qquad r_{nc} = \tfrac{1}{2}(q^L_{nc} + q^R_{nc}), \\ \alpha_{nc} = r_{nc} \big/ \sum_i r_{ni}. \end{cases} \tag{10}$$

Here q^L_{nc} and q^R_{nc} are the learned and robust quality scores, r_{nc} is their mean, and α_{nc} is the corresponding channel-normalised reliability weight. The positive constants $\tau_{\min}$, $q_{\min}$, and β_q set the dispersion floor, reliability floor, and robust-decay scale; all are frozen on development data. The lower clip retains a non-zero candidate weight before routing. Learned quality adapts to task-relevant signal patterns, whereas robust quality responds to unusually flat or variable channels without label input. Their complementary roles yield the measured MPED refinement of the starting state.

The fitted linear anchor acts on a training-standardised low-order descriptor $\widetilde{\boldsymbol{\phi}}_n$. With selection-training set $\mathcal{D}_{\mathrm{tr}}$, sample weight ν_n, class weight ω_{y_n}, and regularisation coefficient λ_A, its training-boundary fit is

$$(\widehat{W}, \widehat{\mathbf{b}}) = \arg\min_{W,\mathbf{b}} \sum_{n\in\mathcal{D}_{\mathrm{tr}}} \nu_n \omega_{y_n}\, \mathrm{CE}(W\widetilde{\boldsymbol{\phi}}_n + \mathbf{b}, y_n) + \lambda_A \|W\|_F^2. \tag{11}$$

The fitted coefficients and descriptor statistics are frozen for evaluation. The initial neural state combines an unweighted channel mean, a reliability-weighted mean, a channel-wise maximum, and a low-order stable projection:

$$\begin{cases} \mathbf{e}^{\mathrm{stb}}_n = W_s\, \mathrm{LN}\big(\mathrm{vec}_{c,b}\{(\delta_0 + r_{nc})[\bar{X}_{ncb}, s_{ncb}]\}\big)\,, \\ \mathbf{h}_0 = g_o\!\left([\mathrm{mean}_c\, \mathbf{h}^{\mathrm{ch}}_{nc}; \sum_c \alpha_{nc}\mathbf{h}^{\mathrm{ch}}_{nc}; \max_c \mathbf{h}^{\mathrm{ch}}_{nc}; \mathbf{e}^{\mathrm{stb}}_n]\right). \end{cases} \tag{12}$$

Here $\bar{X}_{ncb}$ and s_{ncb} denote the mean and standard deviation from $\mathcal{S}$, $\mathrm{vec}_{c,b}$ vectorises channel–band entries, and $\delta_0 > 0$ is a development-frozen weight floor. The maps W_s and g_o are learned. The four summaries are complementary but remain low-cost. The fitted anchor in Equation 11 is estimated separately and frozen. Consequently, the initial neural state and the zero-action linear prediction can be audited independently even though both use low-order evidence.

C Candidate hierarchy and selected-only experts

C.1 Region and relation candidate tokens

Let $M \in \{0,1\}^{R\times C}$ denote deterministic geometry-region membership and $\bar{M}_{rc} = M_{rc}/\sum_i M_{ri}$. A cheap region token is the corresponding channel-token mean,

$$\mathbf{h}^{\mathrm{reg}}_{nr} = \sum_{c=1}^{C} \bar{M}_{rc}\mathbf{h}^{\mathrm{ch}}_{nc}, \qquad r = 1, \ldots, R. \tag{13}$$

Only membership is fixed; the token varies with the sample. For an unordered region pair $r < s$, the candidate relation token is

$$\mathbf{h}^{\mathrm{rel}}_{n,rs} = g_{rel}([\mathbf{h}^{\mathrm{reg}}_{nr}; \mathbf{h}^{\mathrm{reg}}_{ns}; \mathbf{h}^{\mathrm{reg}}_{nr} \odot \mathbf{h}^{\mathrm{reg}}_{ns}; |\mathbf{h}^{\mathrm{reg}}_{nr} - \mathbf{h}^{\mathrm{reg}}_{ns}|])\,. \tag{14}$$

The shared map g_{rel} projects pairwise token features for policy scoring. Deep relation evidence is computed after selection, allowing all 28 candidates to be valued through cheap tokens before one relation expert is invoked. This separation between candidate valuation and acquired evidence defines the conditional-computation contract.

The four datasets use the same action grammar with development-frozen stable views adapted to their recording distributions. MPED uses an 80-dimensional geometry-region anchor. SEED and SEED-V retain channel statistics and coarse regional correlation summaries; SEED-IV uses a channel-statistic-dominant anchor. These fixed instances combine a shared routing interface with dataset-appropriate prediction bases.

C.2 Action vocabulary, legality, and policy

The complete candidate vocabulary contains eight regions, 62 channels, 28 unordered region pairs, and STOP:

$$\mathcal{A} = \{a_r^R\}_{r=1}^{8} \cup \{a_c^C\}_{c=1}^{62} \cup \{a_{rs}^E\}_{r<s} \cup \{a^S\}, \qquad |\mathcal{A}| = 99. \tag{15}$$

Let $o_r^{(j)}$, $v_c^{(j)}$, and $w_{rs}^{(j)}$ indicate whether region r, channel c, and pair (r, s) have already been acquired before step j. The binary legality indicators are

$$\begin{cases} m_r^{R,(j)} = 1 - o_r^{(j)}, & m_c^{C,(j)} = o_{\rho(c)}^{(j)}(1 - v_c^{(j)}), \\ m_{rs}^{E,(j)} = o_r^{(j)} o_s^{(j)}(1 - w_{rs}^{(j)}), & m^{S,(j)} = \mathbb{1}[j \geq K_{\min}]. \end{cases} \tag{16}$$

The map $\rho(c)$ returns the geometry region containing channel c, and $K_{\min}$ is the first permitted stopping step. The mask prevents duplicate targets, permits a channel only after its region is open, and permits a relation only after both endpoints are open.

For candidate token $\mathbf{k}_v$ and state $\mathbf{h}_{j-1}$, the policy score and legal-set probability are

$$\begin{cases} q_{jv} = \langle W_q \mathbf{h}_{j-1}, W_k \mathbf{k}_v \rangle + b_{\tau(v)} + b_v^{\mathrm{prior}}, \\ P_\theta(a_j = v \mid \mathbf{h}_{j-1}, \mathcal{L}_j) = \dfrac{\exp(q_{jv})}{\sum_{u \in \mathcal{L}_j} \exp(q_{ju})}, \qquad v \in \mathcal{L}_j. \end{cases} \tag{17}$$

Here W_q and W_k are learned projections, $\tau(v)$ returns the action type, and $b_{\tau(v)}$ and b_v^{prior} are learned type and candidate biases. The evaluation rule selects the largest legal probability.

C.3 Deep expert definitions

A CHANNEL action encodes only the selected electrode. Five bands are first projected at each temporal position, followed by a depthwise temporal convolution, pointwise mixing, and attention pooling:

$$\begin{cases} \mathbf{u}_{nct} = \mathrm{GELU}(W_b\, \mathrm{LN}(\widetilde{\mathbf{X}}_{nct:})), \\ \mathbf{v}_{nct} = \mathbf{u}_{nct} + \mathrm{PWConv}(\mathrm{GELU}(\mathrm{DWConv}_3(\mathbf{u}_{nc1:T})))_t, \\ \beta_{nct} = \mathrm{softmax}_t(\mathbf{w}_a^\top \mathbf{v}_{nct}), \qquad \mathbf{e}_{nc}^C = g_C\left(\sum_t \beta_{nct} \mathbf{v}_{nct} + g_p(\widetilde{\mathbf{p}}_c)\right). \end{cases} \tag{18}$$

Here W_b projects the five-band input, DWConv_3 and PWConv are depthwise temporal and pointwise convolutions, $\mathbf{w}_a$ produces attention weights β_{nct}, and g_C is the shared channel-evidence readout. One channel action integrates all five bands, so the channel-level topography records the complete spectral evidence used by each selected electrode.

For a selected region r with channel set $\mathcal{C}_r$, let $\boldsymbol{\mu}_{nr}$, $\boldsymbol{\sigma}_{nr}$, and $\mathbf{u}_{nr}^{\max}$ denote the element-wise channel mean, standard deviation, and maximum of $\widetilde{\mathbf{X}}$, and let $\bar{\mathbf{p}}_r$ be the mean sensor coordinate. The REGION expert is

$$\mathbf{e}_{nr}^R = g_R([\mathrm{vec}\, \boldsymbol{\mu}_{nr}; \mathrm{vec}\, \boldsymbol{\sigma}_{nr}; \mathrm{vec}\, \mathbf{u}_{nr}^{\max}; \bar{\mathbf{p}}_r]). \tag{19}$$

The expert preserves regional consistency and peak responses through one shared readout across scalp sectors.

The RELATION expert becomes available only after both endpoint regions are open. It combines the two regional sequences, their product and absolute difference, endpoint geometry, and synchronous plus one-step correlation summaries:

$$\begin{cases} \boldsymbol{\kappa}_{n,rs} = [\mathrm{corr}(\boldsymbol{\mu}_{nr}, \boldsymbol{\mu}_{ns}); \mathrm{corr}(\boldsymbol{\mu}_{nr,1:T-1}, \boldsymbol{\mu}_{ns,2:T}); \mathrm{corr}(\boldsymbol{\mu}_{ns,1:T-1}, \boldsymbol{\mu}_{nr,2:T})], \\ \mathbf{e}^{E}_{n,rs} = g_E([\mathrm{vec}\,\boldsymbol{\mu}_{nr}; \mathrm{vec}\,\boldsymbol{\mu}_{ns}; \mathrm{vec}(\boldsymbol{\mu}_{nr} \odot \boldsymbol{\mu}_{ns}); \mathrm{vec}\,|\boldsymbol{\mu}_{nr} - \boldsymbol{\mu}_{ns}|; \bar{\mathbf{p}}_r; \bar{\mathbf{p}}_s; \boldsymbol{\kappa}_{n,rs}]). \end{cases} \tag{20}$$

The operator corr is evaluated bandwise and its outputs form the signature $\boldsymbol{\kappa}_{n,rs}$; g_E is the shared relation-evidence readout. The lagged terms encode temporal ordering within the feature window, enriching the discriminative relation signature. Call counters establish a one-to-one correspondence between selected targets and deep-expert evaluations. Shared expert parameters make the number of networks independent of the 99 candidate identities.

C.4 Prefix composition

The selected evidence receives a type embedding. A scalar gate and a positive type-specific gain produce its class-space increment:

$$\eta_j = \sigma(f_g([\mathbf{h}_{j-1}; \mathbf{e}_j])), \qquad \Delta\mathbf{z}_j = \mathrm{softplus}(\gamma_{\tau_j})\eta_j W_e \mathbf{e}_j. \tag{21}$$

Here f_g is the scalar gate, $\tau_j \in \{R, C, E\}$ is the selected type, γ_{τ_j} is its learned gain, and W_e projects evidence to class space. A vector gate controls the recurrent state transition:

$$\begin{cases} \widetilde{\mathbf{h}}_j = \mathrm{GRU}(\mathbf{e}_j, \mathbf{h}_{j-1}), \\ \mathbf{g}_j = \sigma(W_g[\mathbf{h}_{j-1}; \mathbf{e}_j]), \\ \mathbf{h}_j = \mathbf{h}_{j-1} + \mathbf{g}_j \odot \tanh(\widetilde{\mathbf{h}}_j - \mathbf{h}_{j-1}). \end{cases} \tag{22}$$

In the bounded-residual configuration, the overview contribution and acquired increments form the prefix residual:

$$\begin{cases} \mathbf{r}^{(0)}_n = \mathrm{softplus}(\gamma_o) W_o \mathbf{h}_0, \\ \mathbf{r}^{(j)}_n = \mathbf{r}^{(0)}_n + \displaystyle\sum_{i=1}^{j} \Delta\mathbf{z}^{(i)}_n, \\ \mathbf{z}^{(j)}_n = \mathbf{z}^{A}_n + \eta_z \tanh(\mathbf{r}^{(j)}_n). \end{cases} \tag{23}$$

Here γ_o is a learned overview gain, W_o maps the overview state to logits, $\mathbf{z}^A_n$ denotes anchor logits, and $\eta_z > 0$ bounds the residual magnitude. Every prefix $j = 0, \ldots, K_{\max}$ therefore provides a prediction. The anchor remains an explicit reference and evidence supplies a bounded residual. The same classifier consequently supports exact-budget and STOP-based inference across budgets.

D Training objective and deployment closure

D.1 Training-only teacher and prefix supervision

The training-only teacher shares the channel and region experts. Separate attention pools their dense outputs, and a learned projection adds the resulting teacher residual to the frozen anchor logits. It broadens training coverage without maintaining independent expert parameters or enumerating deep relation experts, then remains disabled for development and outer inference.

For a mini-batch of N samples, all prefixes receive class-weighted, label-smoothed cross-entropy, denoted by $L^{CE}_{n,j}$. The safety term penalises a prefix that exceeds the zero-action loss:

$$\begin{cases} \mathcal{L}_{pre} = N^{-1} \displaystyle\sum_{n=1}^{N} \sum_{j=0}^{K_{\max}} w_j L^{CE}_{n,j}, \qquad \sum_j w_j = 1, \\ \mathcal{L}_{safe} = \dfrac{1}{N K_{\max}} \displaystyle\sum_{n,j \geq 1} \max(0, L^{CE}_{n,j} - L^{CE}_{n,0}). \end{cases} \tag{24}$$

The development-frozen weights $w_j \geq 0$ are non-decreasing with the prefix length, reflecting the availability of more evidence while retaining a usable low-budget predictor. Denoting detached teacher logits by $\mathbf{z}_n^T$, the distillation term is

$$\mathcal{L}_{KD} = \frac{T_d^2}{N(K_{\max}+1)} \sum_{n=1}^{N} \sum_{j=0}^{K_{\max}} \mathrm{KL}\Big[\mathrm{softmax}(\mathbf{z}_n^T/T_d) \parallel \mathrm{softmax}(\mathbf{z}_n^{(j)}/T_d)\Big], \tag{25}$$

where $T_d > 0$ is the distillation temperature and the teacher distribution is detached. Prefix supervision, safety, and distillation form the optimisation package used in every complete-system evaluation.

D.2 Actor–critic terms and checkpoint protection

The immediate reward appears in Equation 5. Backward recursion gives the discounted return and state-value advantage:

$$\begin{cases} G_n^{(j)} = r_n^{(j)} + \gamma G_n^{(j+1)}, & 0 \leq \gamma < 1, \\ A_n^{(j)} = G_n^{(j)} - V_\psi(\mathbf{h}_n^{(j-1)}). \end{cases} \tag{26}$$

Here $G_n^{(j)}$ and $A_n^{(j)}$ are the return and advantage, V_ψ is the critic, and γ is the development-frozen discount factor. The recursion terminates with $G_n^{(J_n+1)} = 0$, where J_n is the final active step for sample n. Only active sample–step pairs Ω enter the policy losses:

$$\begin{cases} \mathcal{L}_{actor} = -|\Omega|^{-1} \displaystyle\sum_{(n,j)\in\Omega} \log P_\theta(a_n^{(j)} \mid \mathbf{h}_n^{(j-1)}, \mathcal{L}_n^{(j)})\, \mathrm{stopgrad}(A_n^{(j)}), \\ \mathcal{L}_{critic} = |\Omega|^{-1} \displaystyle\sum_{(n,j)\in\Omega} (V_\psi(\mathbf{h}_n^{(j-1)}) - \mathrm{stopgrad}(G_n^{(j)}))^2, \\ \mathcal{H} = |\Omega|^{-1} \displaystyle\sum_{(n,j)\in\Omega} H[P_\theta(\cdot \mid \mathbf{h}_n^{(j-1)}, \mathcal{L}_n^{(j)})]. \end{cases} \tag{27}$$

The operator $\mathrm{stopgrad}$ blocks gradient propagation through its argument, and $H[\cdot]$ denotes categorical entropy. The implemented loss separates predictive supervision from reinforcement learning:

$$\mathcal{L} = \mathcal{L}_{\mathrm{sup}} + \mathcal{L}_{\mathrm{RL}}. \tag{28}$$

The two groups are

$$\begin{cases} \mathcal{L}_{\mathrm{sup}} = \mathcal{L}_{pre} + \lambda_{safe}\mathcal{L}_{safe} + \lambda_T \mathcal{L}_T + \lambda_{KD}\mathcal{L}_{KD}, \\ \mathcal{L}_{\mathrm{RL}} = \lambda_{\mathrm{act}}\mathcal{L}_{actor} + \lambda_{\mathrm{val}}\mathcal{L}_{critic} - \lambda_H \mathcal{H}. \end{cases} \tag{29}$$

Here $\mathcal{L}_T$ denotes the teacher's class-weighted classification loss. Every λ is a non-negative development-frozen weight, with its numerical value recorded in the executable configurations. A development checkpoint is chosen by strictly forward exact-$K = 6$ balanced accuracy; the training guard restores the last accepted state when this criterion deteriorates beyond its frozen tolerance. Fresh refitting starts after the development configuration is frozen. The anonymised repository records all numerical weights, optimiser settings, dataset-specific anchor views, forward accumulators, and prior parameters.

D.3 Strictly forward output

Let q index a trial, and let $J_{qt} \leq K_{\max}$ be the final valid acquisition prefix for its window at position t. The corresponding probability is $\mathbf{p}_{qt} = \mathrm{softmax}(\mathbf{z}_{qt}^{(J_{qt})})$. SEED, SEED-IV, and SEED-V use a dataset-specific frozen decay ρ_D; MPED uses the cumulative mean:

$$\begin{cases} \mathbf{s}_{qt} = \mathbf{p}_{qt}, & t = 1, \\ \mathbf{s}_{qt} = \rho_D \mathbf{s}_{q,t-1} + (1-\rho_D)\mathbf{p}_{qt}, & t > 1, \\ \bar{\mathbf{p}}_{qt} = t^{-1} \displaystyle\sum_{i=1}^{t} \mathbf{p}_{qi}. \end{cases} \tag{30}$$

Here $\mathbf{s}_{qt}$ is the exponential accumulator and $\bar{\mathbf{p}}_{qt}$ is the cumulative mean. Both forms reset at a new $(\text{session}, \text{trial})$ key. MPED applies an uncertainty-conditioned, development-frozen logit prior after the cumulative mean:

$$\begin{cases} u_{qt} = [1 - (\bar{p}_{qt,(1)} - \bar{p}_{qt,(2)})]^{\xi}, \\ \widetilde{\mathbf{p}}_{qt} = \mathrm{softmax}(\log(\bar{\mathbf{p}}_{qt} + \epsilon_p) + u_{qt}\mathbf{b}_D), \qquad \widehat{y}_{qt} = \arg\max_k \widetilde{p}_{qt,k}. \end{cases} \tag{31}$$

Here $\bar{p}_{qt,(1)}$ and $\bar{p}_{qt,(2)}$ are the largest two cumulative probabilities, $\xi > 0$ is the frozen uncertainty exponent, $\epsilon_p > 0$ stabilises the logarithm, and $\mathbf{b}_D$ is the development-fitted class-logit prior. The prior depends exclusively on current predictive uncertainty and the development-fitted vector $\mathbf{b}_D$. When the cumulative margin is large, u_{qt} decreases and the current evidence dominates. The history-only and prior-only controls show that trial history is the principal measured source, while the prior supplies a smaller correction.

Exhaustive invariance tests perturb every suffix after an earlier prediction and recompute all output conditions. Across 857,808 checks, the maximum absolute change in any earlier probability or score is zero, directly verifying the strictly forward execution contract.

E Metrics and inferential procedure

E.1 Subject-level metrics

Accuracy, balanced accuracy (BAcc), and macro-F1 are calculated for each subject using their standard definitions. Tables report the cross-subject mean and sample standard deviation. BAcc assigns equal weight to class recalls, whereas macro-F1 averages class-wise F1 scores. Agreement across their reproduced mean rankings confirms the ordering under overall, class-balanced, and class-wise summaries.

For a paired comparison, M_s^A and M_s^B denote the metric for subject s under conditions A and B. Let $d_s = M_s^A - M_s^B$ after all fold merging and, where applicable, seed averaging, and let S be the number of paired subjects. The estimand and standardised paired effect are

$$\begin{cases} \widehat{\Delta} = S^{-1} \sum_{s=1}^{S} d_s, \\ d_z = \widehat{\Delta} \Big/ \sqrt{(S-1)^{-1} \sum_s (d_s - \widehat{\Delta})^2}. \end{cases} \tag{32}$$

Subject bootstrap confidence intervals resample the S paired subjects with replacement. Paired permutation tests randomly flip the sign of each d_s and compare the absolute permuted mean with $|\widehat{\Delta}|$. The reported analyses use 100,000 permutations. Holm correction is applied within each prespecified Accuracy family, with secondary summaries reported separately (Holm, 1979; Efron & Tibshirani, 1993).

E.2 Evidence hierarchy

The evidence hierarchy combines four-dataset recognition, subject-paired MPED controls, and direct mechanism audits. Reproduced means evaluate the complete system within a common framework with dataset-frozen configurations. The matched controls compare routers with a common encoder, vocabulary, and evidence budget, providing evidence for ordered reinforcement-learned routing. Mechanism audits verify execution properties through future invariance, hard-call sparsity, legal-mask matching, and path disagreement. Together, these levels connect recognition performance, component attribution, and observable model behaviour.

F Complete reproduced baseline comparisons

The main table retains representative graph and temporal methods. Tables 3–6 reorganise the complete comparison by dataset, with Accuracy, balanced accuracy (BAcc), and macro-F1 reported in adjacent

columns at body font size. This arrangement supports within-dataset reading without requiring readers to move between metric-specific tables. Method labels follow their published names. DGCNN, MASA-TCN, EmT, HDGNet, and PGCN denote their published architectures (Song et al., 2020b; Ding et al., 2024; 2025; Li & Kim, 2025; Jin et al., 2024).

The remaining labels are also public methods. BF-GCN-DE denotes the BF-GCN graph-learning system evaluated on the common differential-entropy input; BrainTune denotes the artefact-aware meta-contrastive framework; CADD-DCCNN is the cross-attention dilated causal convolutional network with a domain discriminator (Li et al., 2025a; Gu et al., 2026; Li et al., 2024a). DGC-Link retains the published dual-gate Chebyshev linkage architecture, PR-PL is prototypical-representation pairwise learning, and RGCB denotes Residual GCB-Net (Li et al., 2025d; Zhou et al., 2024; Li et al., 2023a). Author implementations and reported optimisation defaults are retained where available; interface adapters align input layout and classifier dimension with the common data and split contract. The resulting 12-row comparison enables complete three-metric evaluation within each dataset.

Table 3: Complete SEED subject-level comparison (mean±sample SD, %).

Method	Accuracy	BAcc	macro-F1
AdaRelate	**72.37±15.99**	**72.46±16.01**	**70.07±18.04**
BF-GCN-DE	55.95±13.84	56.10±13.48	50.96±16.49
BrainTune	55.62±14.44	55.32±14.53	51.94±16.13
CADD-DCCNN	36.86±13.54	36.73±13.57	27.75±16.13
DGC-Link	54.12±17.53	54.04±17.65	49.00±20.35
DGCNN	64.27±14.68	64.31±14.61	60.97±16.62
EmT	59.32±14.94	59.45±14.81	55.09±17.56
HDGNet	59.84±15.11	60.01±15.11	56.58±17.85
MASA-TCN	67.05±15.33	67.20±15.27	63.43±17.80
PGCN	62.36±17.28	62.42±17.10	58.55±20.07
PR-PL	52.95±16.82	52.79±17.08	45.95±21.82
RGCB	57.37±12.47	57.55±12.53	52.29±15.67

On SEED, the complete system leads the strongest reproduced mean by $5.32\,\mathrm{pp}$ in Accuracy, $5.26\,\mathrm{pp}$ in BAcc, and $6.64\,\mathrm{pp}$ in macro-F1. The aligned margins show that the advantage persists across overall, class-balanced, and class-wise evaluation.

Table 4: Complete MPED subject-level comparison (mean±sample SD, %). The four folds are merged within subject before metric computation. The AdaRelate row is the frozen seed-8635 reference estimate and is distinct from the three-seed matched estimate used to test routing identity.

Method	Accuracy	BAcc	macro-F1
AdaRelate	**66.85±7.95**	**59.26±10.86**	**59.04±11.15**
BF-GCN-DE	61.42±6.80	51.54±8.92	51.56±9.60
BrainTune	60.31±5.64	42.60±9.60	39.86±12.55
CADD-DCCNN	55.92±2.50	37.08±4.09	33.01±6.57
DGC-Link	56.81±6.51	39.65±4.50	37.12±6.52
DGCNN	65.75±8.00	55.24±11.22	55.25±12.05
EmT	63.78±7.53	56.44±10.32	55.87±10.29
HDGNet	61.53±6.94	52.12±8.31	51.88±8.77
MASA-TCN	63.33±7.44	56.75±10.43	56.29±10.48
PGCN	63.23±7.42	53.96±9.65	54.16±10.09
PR-PL	62.62±7.29	55.97±10.57	55.27±10.35
RGCB	65.00±6.85	56.37±10.57	56.11±10.83

MPED yields gains of $1.10\,\mathrm{pp}$ in Accuracy, $2.51\,\mathrm{pp}$ in BAcc, and $2.75\,\mathrm{pp}$ in macro-F1. AdaRelate's 66.85% entry is the seed-8635 frozen reference estimate; the matched-router analysis uses the three-

seed subject mean reported in the main text. The aligned benchmark and paired-control results provide complementary system- and routing-level evidence.

Table 5: Complete SEED-IV subject-level comparison (mean±sample SD, %).

Method	Accuracy	BAcc	macro-F1
AdaRelate	**66.88±10.84**	**66.04±11.47**	**64.82±12.79**
BF-GCN-DE	50.68±11.96	49.50±13.07	46.91±15.73
BrainTune	41.93±16.69	40.39±17.94	31.33±22.78
CADD-DCCNN	34.27±7.52	33.81±9.06	23.48±11.01
DGC-Link	34.16±8.78	32.11±9.36	22.27±14.14
DGCNN	62.14±15.36	60.75±16.24	57.67±19.61
EmT	64.12±13.12	62.70±13.78	61.08±14.83
HDGNet	59.84±15.40	59.31±16.16	57.15±18.10
MASA-TCN	64.60±13.06	63.73±13.76	62.76±13.60
PGCN	61.38±15.94	60.56±16.83	58.50±18.32
PR-PL	58.64±15.94	57.71±16.21	54.42±20.24
RGCB	61.31±12.27	60.15±12.71	58.23±13.35

On SEED-IV, the corresponding margins are $2.28\,\mathrm{pp}$, $2.31\,\mathrm{pp}$, and $2.06\,\mathrm{pp}$. Their similar scale supports metric-consistent mean behaviour, while the reported dispersions retain the full subject-level distribution.

Table 6: Complete SEED-V subject-level comparison (mean±sample SD, %).

Method	Accuracy	BAcc	macro-F1
AdaRelate	**80.01±19.89**	**77.14±20.51**	**72.08±24.10**
BF-GCN-DE	56.51±22.94	58.72±22.86	49.60±25.12
BrainTune	34.71±22.85	36.56±18.62	24.50±22.24
CADD-DCCNN	25.87±20.58	27.54±16.67	16.83±17.08
DGC-Link	36.63±19.12	37.77±16.92	28.04±19.49
DGCNN	61.90±18.85	60.73±17.70	51.93±19.97
EmT	70.44±22.85	67.23±23.16	61.50±25.81
HDGNet	63.93±21.16	63.60±21.13	54.53±23.76
MASA-TCN	76.36±23.70	72.61±24.18	69.33±25.46
PGCN	64.87±24.76	65.13±24.47	59.03±26.85
PR-PL	63.78±19.31	63.95±19.95	55.09±21.98
RGCB	63.14±16.60	62.24±15.51	54.77±17.23

SEED-V gives margins of $3.65\,\mathrm{pp}$ in Accuracy, $4.53\,\mathrm{pp}$ in BAcc, and $2.75\,\mathrm{pp}$ in macro-F1. Across the four dataset-specific tables, all three metrics retain the same mean ordering. This consistency supports the complete-system contribution, while the paired router controls provide the focused attribution to sequential reinforcement learning.

G Budget, structural, and intervention controls

G.1 Anytime budgets and adaptive stopping

Tables 7 and 8 report the full frozen Accuracy sweep. Exact-K replays truncate the stored path after a fixed number of acquired actions, while adaptive inference follows the STOP rule.

Table 7: Early exact-budget and adaptive Accuracy means (%). Adaptive is listed here for comparison; its mean action count appears in Table 8.

Dataset	$K0$	$K1$	$K2$	$K3$	Adaptive
SEED	72.37	72.37	72.37	72.37	72.37
MPED	66.22	66.49	66.76	66.79	66.77
SEED-IV	67.11	67.04	67.02	66.97	66.95
SEED-V	80.01	80.01	80.01	80.01	80.01

Table 8: Later exact-budget Accuracy means, $K6 - K0$ contrast, and adaptive action count.

Dataset	$K4$	$K5$	$K6$	$K6 - K0$ (pp)	Adaptive actions
SEED	72.37	72.37	72.37	0.00	5.19
MPED	66.77	66.80	66.85	+0.63	2.19
SEED-IV	66.95	66.93	66.88	-0.23	5.33
SEED-V	80.01	80.01	80.01	0.00	5.99

The sweep reveals two operating regimes. The fitted anchor already determines the final class on SEED and SEED-V and remains dominant on SEED-IV, while MPED gains $0.63\,\mathrm{pp}$ from six acquisitions, with a 95% CI of $[-0.15, 1.58]$ and Holm-adjusted $p = 0.1763$. The matched one-shot and no-RL controls then show that organising the same structured budget through sequential reinforcement learning improves Accuracy.

Adaptive stopping averages 2.19 actions on MPED and remains close to six on SEED-V. STOP therefore provides an anytime interface that adjusts path length to the dataset while retaining a common maximum computation horizon.

G.2 Action-space and structural controls

The relation controls retrain the relevant variant under the same outer protocol. Removing RELATION masks this action type throughout training, refitting, and evaluation. Dense relations evaluate all 28 relation experts. Table 9 reports variant-minus-Full Accuracy differences at seed 8635.

Table 9: Relation-action controls. Values are Accuracy differences in percentage points with subject-bootstrap 95% CIs.

Dataset	No relation − Full	Dense relations − Full
SEED	0.00 $[0.00, 0.00]$	0.00 $[0.00, 0.00]$
MPED	-0.27 $[-1.05, 0.42]$	+0.24 $[-0.46, 0.91]$
SEED-IV	-0.02 $[-0.37, 0.30]$	-0.02 $[-0.16, 0.14]$
SEED-V	0.00 $[0.00, 0.00]$	0.00 $[0.00, 0.00]$

Relation controls locate the principal predictive contribution in the coupled REGION–CHANNEL–RELATION–STOP system. Relation targets supply endpoint-aware options within the typed grammar and leave auditable calls in every executed path.

The three-seed MPED study also evaluates graph-consistent legality. Full reaches 66.81% Accuracy and the unmasked control reaches 66.74%, giving $+0.07\,\mathrm{pp}$ in the Full-minus-control orientation (95% CI $[-0.20, 0.34]$; $p = 0.6165$). Separately, 33.50% of unconstrained targets violate the Full grammar and shift the action-type composition. The mask therefore supplies a structural interface that guarantees valid, auditable dependencies while preserving the predictive profile.

G.3 Evidence interventions

Inference-time interventions distinguish the selected path from the values carried by its acquired evidence. Table 10 uses the MPED exact-$K6$ checkpoint. "Evidence null" retains the path interface but removes the acquired residual. The same-type replacements substitute a randomly chosen or previously unselected target of the selected action type.

Table 10: MPED inference interventions. Accuracy is a subject mean; Δ is intervention-minus-original. Holm correction covers the three prespecified interventions.

Condition	Accuracy (%)	Δ (pp)	95% CI (pp)	Holm p
Original	66.85	0.00	—	—
Evidence null	66.22	-0.63	$[-1.59, 0.14]$	0.3790
Random same type	66.82	-0.03	$[-0.06, 0.00]$	0.3790
Unselected same type	66.82	-0.03	$[-0.08, 0.01]$	0.3790

Evidence nulling increases cross-entropy by 0.0185 (Holm-adjusted $p < 0.0001$), whereas same-type substitutions produce much smaller changes. Acquired evidence therefore refines probability quality even when the stable anchor preserves the final class label. The result directly supports the stability–adaptation account: the anchor fixes many decisions, while the residual improves the predictive distribution near the decision boundary.

G.4 Conditional-computation efficiency

Table 11 averages four profiles per mode: the first frozen subject/fold case from each of SEED, MPED, SEED-IV, and SEED-V. Each profile starts from the same device-resident $62 \times 10 \times 5$ tensor and times the complete model forward, including the common overview, anchor, and class-output path. Host-to-device transfer, feature construction, and the separately timed strictly forward integration are excluded. Dense relations retains the common prediction path but replaces sequential policy execution with one vectorised evaluation of all 28 candidate relations. Operator counts cover profiler-supported multiply–accumulate equivalents. The GPU is an NVIDIA A100 80GB PCIe with CUDA 12.4 and PyTorch 2.5.1. Batch-one latency uses 50 warm-up and 200 synchronised iterations.

Table 11: Mean efficiency over one frozen subject/fold case from each dataset. GPU timings exclude data transfer and the separately profiled forward trial accumulator. MAC equivalents include only profiler-supported operations.

Mode	Actions	MAC eq. (M)	GPU B1 (ms)	B64 samples/s
Overview ($K0$)	0.00	2.64	18.20	3296
Adaptive	4.39	2.70	23.17	2480
Exact $K6$	6.00	2.75	25.41	2438
One-shot top-k	6.00	1.56	23.05	2564
Dense relations	28.00	1.51	4.65	13,389

Adaptive inference reduces expert calls from 6.00 to 4.39 relative to exact $K6$ and reduces measured batch-one latency from 25.41 to 23.17 ms. The dense-relation profile further shows that vectorisation strongly affects hardware throughput. Together, the profiles distinguish AdaRelate's verified conditional execution and call reduction from implementation-specific scheduling effects, providing a clear basis for subsequent kernel optimisation.

H Class-stratified adaptive channel-query analysis

The main paper concentrates on realised utility in a representative sensor-space view. Figure 8 supplies the complementary four-dataset execution-frequency view, resolving how often the frozen policy

invokes CHANNEL, REGION, and RELATION targets. Figure 9 then compares emotion-conditioned channel profiles. These behavioural summaries precede the complete class-wise topographies and typed-evidence maps below.

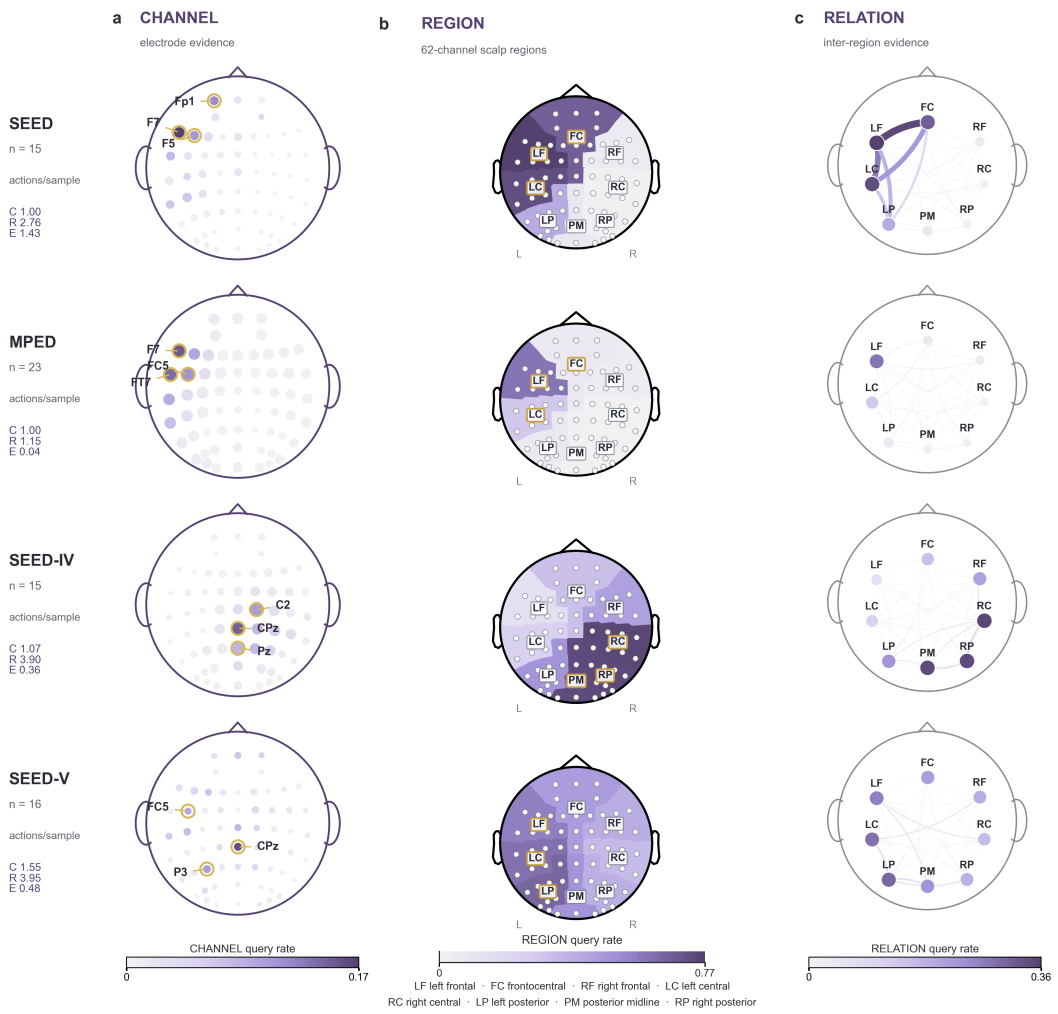

Figure 8: Four-dataset hierarchical evidence allocation. Each row contains a CHANNEL scalp map, an MNE scalp map of the eight sensor-defined REGION groups on the measured 62-channel montage, and a RELATION network. Colour gives target query rate, and row labels report mean CHANNEL/REGION/RELATION actions per sample. Channel maps display all 62 electrodes without interpolation; gold rings and adjacent labels identify the three largest mean rates. REGION shading is a sensor-space model allocation, not a source-localised cortical estimate.

SEED follows the most relation-intensive paths, whereas MPED concentrates its shorter trajectory on REGION and CHANNEL decisions. SEED-IV and SEED-V allocate more actions to coarse regions before querying selected sensors. The common grammar therefore supports dataset-conditioned evidence depth without changing the legal action interface.

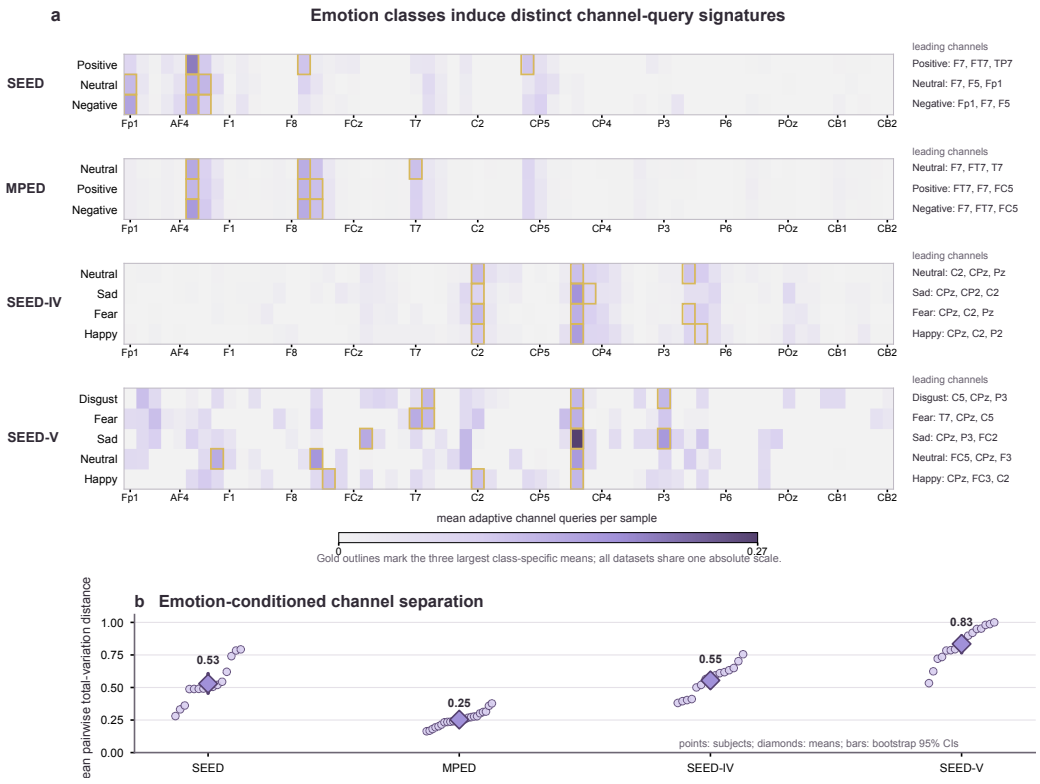

Figure 9: Emotion-conditioned channel-query signatures. (a) Subject-averaged true-class query rates on a common absolute scale; gold cells identify the three largest row means. (b) Subject-level mean pairwise total-variation distance between normalised class profiles. Diamonds and bars report means and bootstrap 95% CIs.

Mean pairwise total-variation distances are 0.53, 0.25, 0.55, and 0.83 for SEED, MPED, SEED-IV, and SEED-V. The router consequently retains reusable channel cores while increasing spatial differentiation when the emotion vocabulary and learned execution paths require it. The following full topographies preserve all 62 channels, use one absolute colour scale, and encode subject prevalence by marker area.

H.1 SEED

SEED retains a left frontal or frontotemporal concentration in all three classes, with class-dependent leading channels. The ringed sites are F7, FT7, and TP7 for Positive trials and Fp1, F7, and F5 for Neutral and Negative trials. The accompanying subject points show that top-ten concentration varies materially between participants even when the class mean is high.

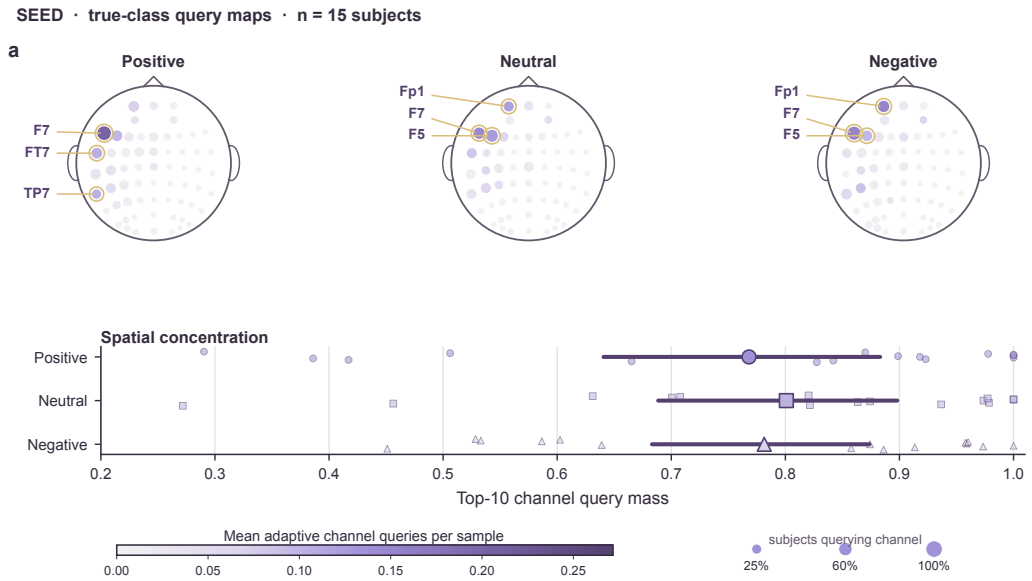

Figure 10: SEED true-class adaptive channel-query maps and subject-level spatial concentration. Colour gives mean channel queries per sample; marker area gives the fraction of subjects querying the channel. Gold rings and adjacent names identify the three largest class-specific means. Points in the lower panel are subjects; large markers and intervals are the subject mean and bootstrap 95% CI of top-ten query mass.

The ringed set changes between Positive and the other two classes, and the full 62-channel intensity field also changes. The trained classifier therefore uses distinct channel-query patterns across the true-class strata, demonstrating class-conditioned allocation within a shared policy.

H.2 THREE-CLASS DATASET

MPED shows a consistent left frontotemporal region with class-dependent ordering. For Neutral, the ringed channels are F7, FT7, and T7; for Positive and Negative, they are F7, FC5, and FT7. The maps are derived after the four folds are merged within each subject, preserving $n = 23$ as the independent unit. Larger markers indicate subject prevalence rather than additional window-level observations.

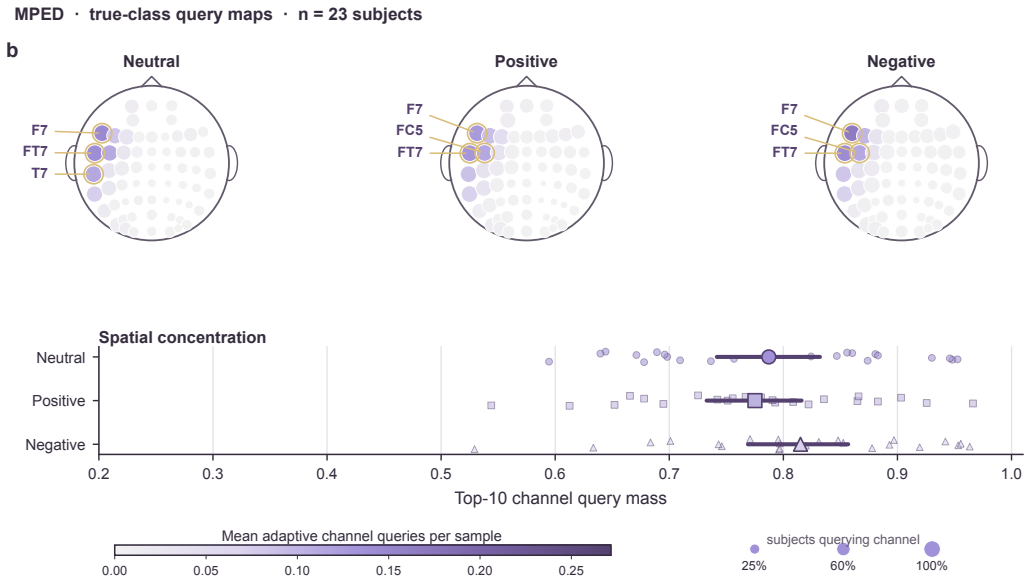

Figure 11: MPED true-class adaptive channel-query maps. The four outer folds are merged within participant before subject-level spatial summaries are formed. Channel names are placed beside every gold-ringed electrode with short leaders.

The top-ten mass distributions are broad and include participants with weaker concentration than the class mean. This heterogeneity complements the 56.51% exact-mask target-disagreement rate and distinguishes participant-specific paths from the class-average map. The map reports where the frozen router spends channel calls, whereas the matched control establishes whether sequential reinforcement learning improves Accuracy.

H.3 SEED-IV

SEED-IV moves the largest rates towards central and centroparietal sites. C2 and CPz appear in all four ringed sets; the third channel is Pz for Neutral and Fear, CP2 for Sad, and P2 for Happy. Shared central sites therefore coexist with class-specific ordering, revealing a stable computational core with class-conditioned allocation.

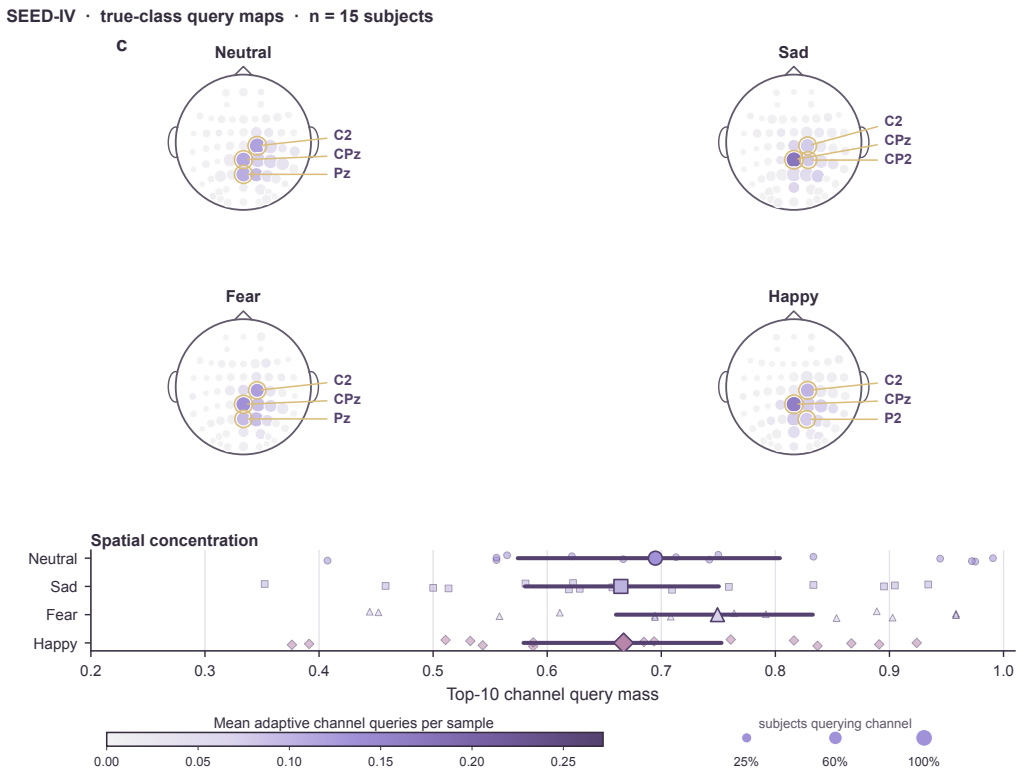

Figure 12: SEED-IV true-class adaptive channel-query maps and subject-level top-ten query mass. The four class maps occupy two rows at full single-column width. Direct channel labels replace a detached legend so the ringed sites can be read locally.

The subject distributions show overlapping concentration across the four classes. Together with the class-dependent ringed channels, this pattern shows that routing concentration and encoded evidence play complementary roles: the policy can reuse central candidates while changing their ordering and evidence content.

H.4 SEED-V

SEED-V produces the most diverse ringed sets. Disgust emphasises C5, P3, and CPz; Fear emphasises C5, T7, and CPz; Sad highlights FC2, P3, and CPz; Neutral highlights F3, FC5, and CPz; Happy highlights FC3, C2, and CPz. CPz is common, but its surrounding channels vary across emotion categories.

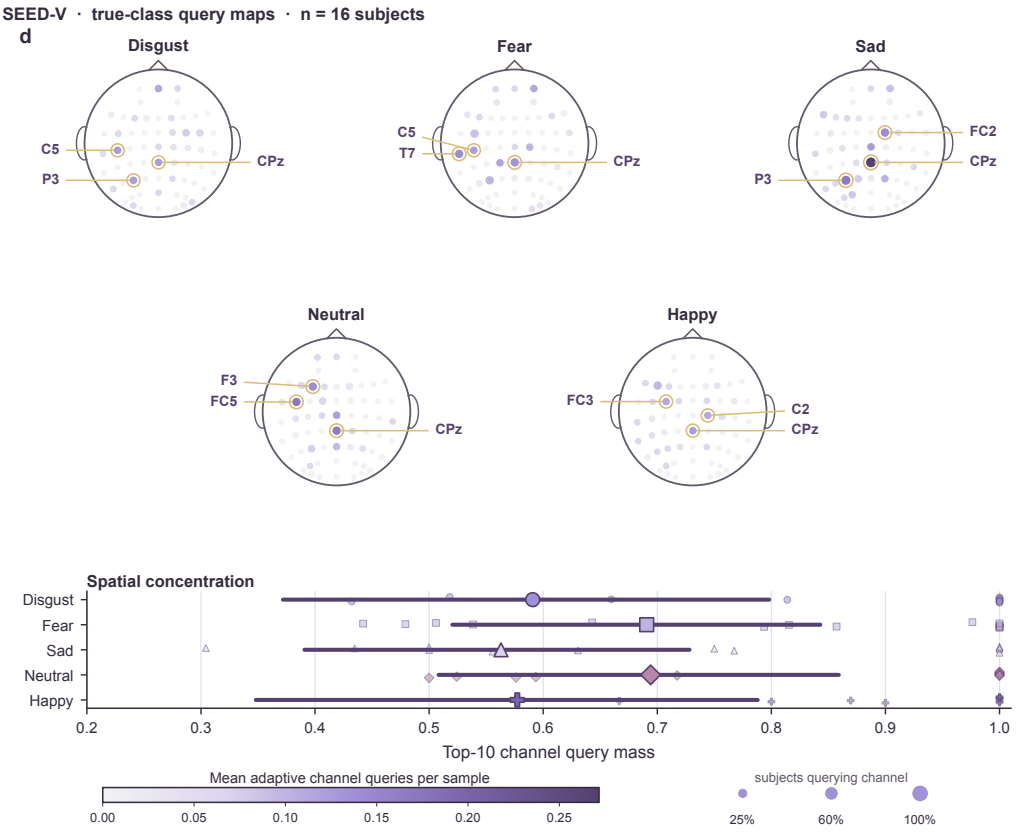

Figure 13: SEED-V true-class adaptive channel-query maps and subject-level spatial concentration. The five maps use two rows, followed by a full-width distribution panel. Colour and size retain the same absolute encodings as Figures 10–12.

The five-class view makes the portability claim more precise. The same router interface accommodates a different emotion vocabulary and shifts its channel-use pattern, while the fitted prediction base remains dominant for this dataset. The maps demonstrate non-collapsed execution; the SEED-V Accuracy margin remains a property of the complete system.

I Path and Mechanism Audit

On MPED, Figure 14 combines an aggregate region-pair network, a deterministic action path, and the adaptive-length distribution. The network aggregates region and relation query rates across 23 subjects. Anonymous nodes organise the typed computation graph, and edge thickness records the frequency with which the trained policy invokes each relation expert.

The centre panel shows the sample path selected by the smallest original index under the frozen rule: REGION R1, CHANNEL F7, then STOP. This deterministic example makes the action grammar and hard-call sequence concrete, while the final panel shows how subject-level adaptive path length departs from exact $K = 6$.

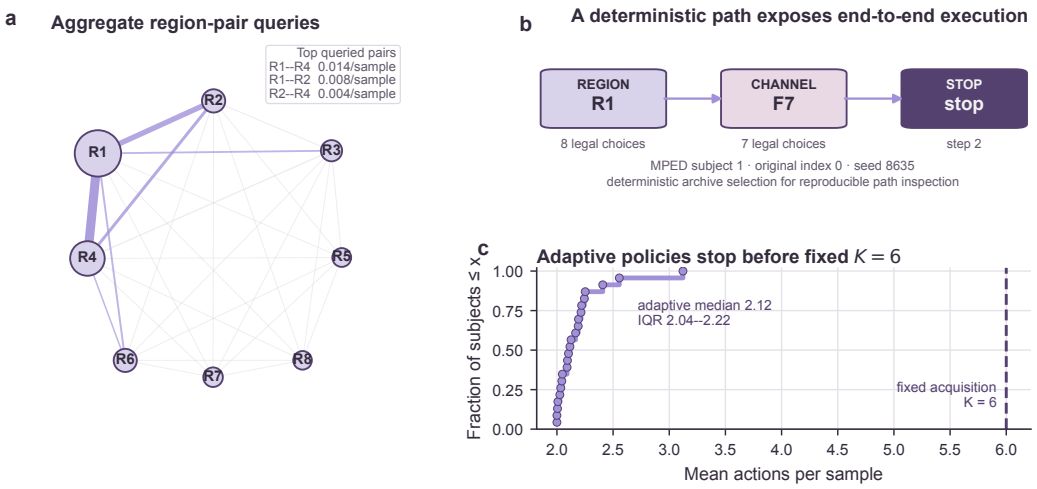

Figure 14: Mechanism audit. (a) Aggregate region-pair query network. (b) Deterministically selected sample path, REGION R1 $\rightarrow$ CHANNEL F7 $\rightarrow$ STOP. (c) Empirical cumulative distribution of subject-level adaptive actions per sample with exact $K = 6$ as reference.

Table 12: Full action composition in the mechanism audit. Calls are averaged over 23 subjects after fold merging; STOP terminates a path and is excluded from evidence-call counts.

Quantity	Mean per sample	Operational role
REGION calls	1.15	opens geometry-region evidence
CHANNEL calls	1.00	invokes one selected channel expert
RELATION calls	0.04	invokes one legal region-pair expert
Total evidence actions	2.19	adaptive path length before termination

Table 12 shows that REGION and CHANNEL calls dominate, with few RELATION calls. Zeroing newly acquired evidence changes the next Full action in 1.45% of eligible trajectories, compared with 0.00% for one-shot top-k, and thus records feedback-sensitive replanning. Sample-conditioning extends across the path: even with an identical exact legal mask, 56.51% of matched pairs choose a different target. The policy therefore adapts target identity to the sample state through both initial conditioning and acquired evidence.

J STATE-TRANSITION CONTRIBUTION AND ROUTING GEOMETRY

Using MPED as the matched case, the three evidence states contribute at complementary levels. Figure 15a–b holds candidate availability fixed and measures whether target identity changes between samples sharing the same subject, fold, seed, step, action type, and exact legal-mask signature. Full AdaRelate yields target-disagreement rates of 64.9%, 51.5%, and 14.9% for CHANNEL, REGION, and RELATION, with corresponding normalised target entropies of 0.66, 0.48, and 0.22. The result locates the strongest sample-conditioned variation in node and region selection while retaining non-zero relational variation.

Panel c measures how often each state is entered. REGION accounts for 53%, 52%, 73%, and 66% of acquired evidence actions on SEED, MPED, SEED-IV, and SEED-V; the corresponding CHANNEL shares are 19%, 46%, 20%, and 26%. RELATION use is largest on SEED and more selective on the other datasets. These shifts quantify how one action grammar changes its computational scale with the recording distribution. Panel d supplies the task criterion: Full AdaRelate improves Accuracy by $0.93\,\mathrm{pp}$ over one-shot top-k and by $1.17\,\mathrm{pp}$ over no-RL/static routing. Target diversity

and recognition contribution are therefore reported separately: panels a–c resolve what changes along the path, while panel d identifies the routing rule that performs better.

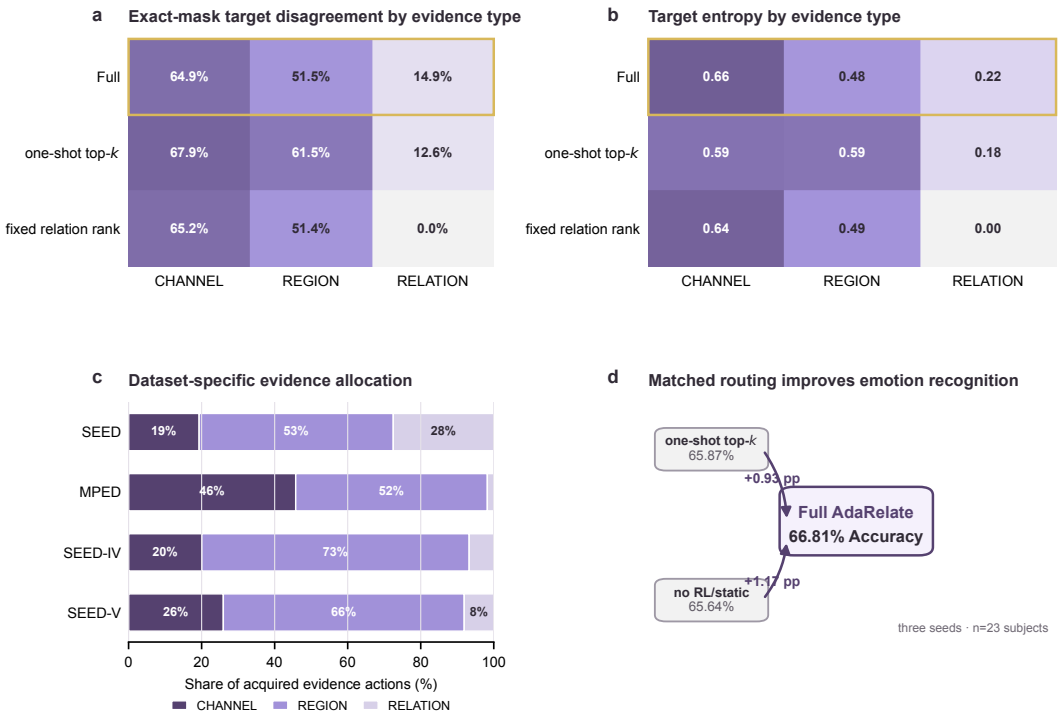

Figure 15: Typed state transitions and routing contribution. (a) Pairwise target disagreement under an exact legal-mask signature. (b) Normalised target entropy within CHANNEL, REGION, and RELATION decisions. Gold outlines identify Full AdaRelate. (c) Action-type shares on all four datasets. (d) Three-seed Accuracy and subject-paired Full-minus-control differences for the matched case.

Figure 16 separates local evidence utility, total realised utility, and prediction-transition rates for every evidence type and dataset. The matched case exhibits the clearest positive REGION and CHANNEL effects. The remaining datasets place useful context around that result: the identical interface can preserve the anchor decision while adjusting confidence by different amounts and in different directions.

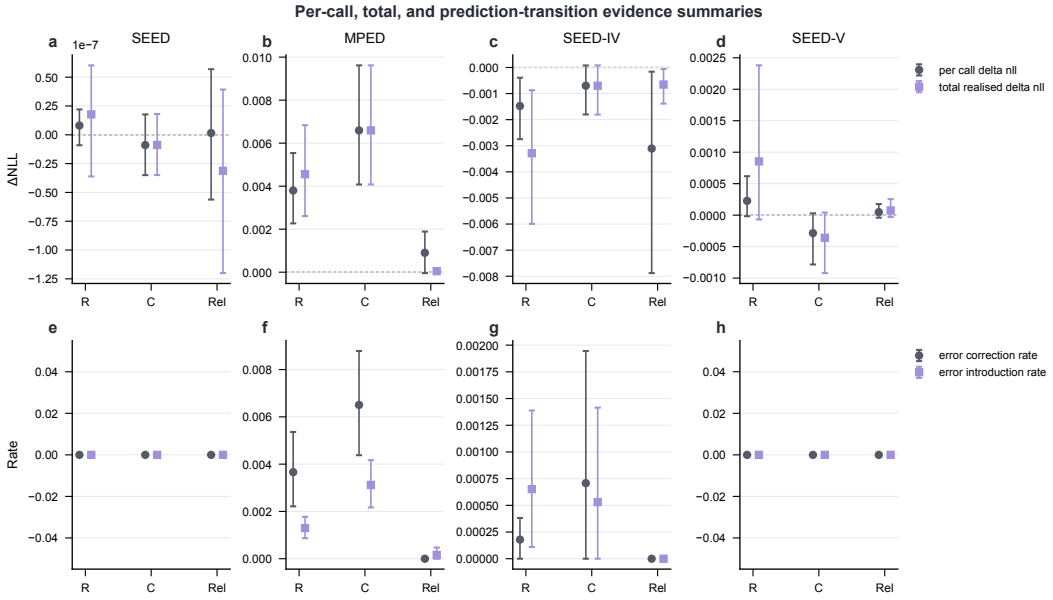

Figure 16: Complete evidence-utility and prediction-transition summaries. Panels a–d report per-call and total realised ΔNLL for REGION (R), CHANNEL (C), and RELATION (Rel). Panels e–h report error-correction and error-introduction rates. Points are subject means and intervals are subject-bootstrap 95% CIs.

The frozen-route replay in Figure 17 holds every action identity, order, legality mask, and STOP decision fixed, then zeros one evidence type. This intervention tests the evidence carried by an executed path without changing the route itself. Probability-level effects most clearly resolve the matched-case REGION and CHANNEL contributions, including confidence refinements that preserve the final class label.

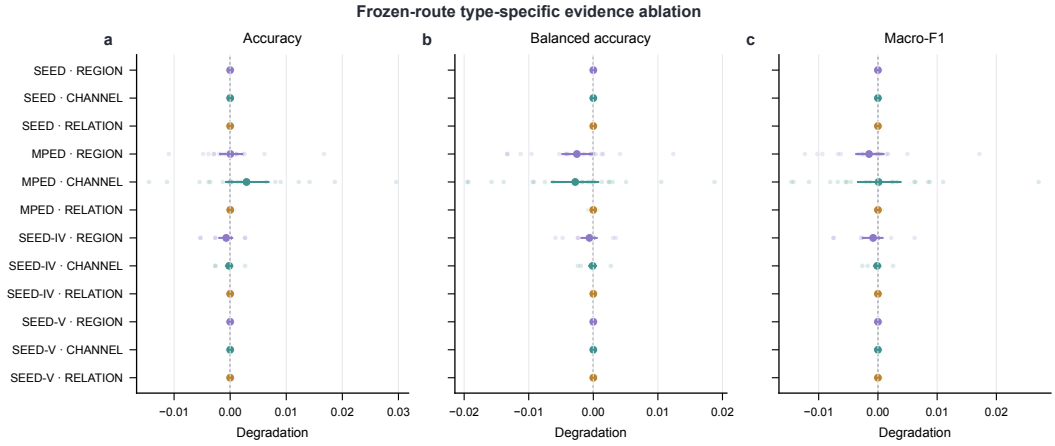

Figure 17: Frozen-route type-specific evidence ablation on all four datasets. Small points are subjects; large markers and intervals report the subject mean and bootstrap 95% CI for Accuracy, balanced accuracy, and macro-F1 degradation after the selected evidence type is zeroed. Positive degradation denotes a reduction in performance.

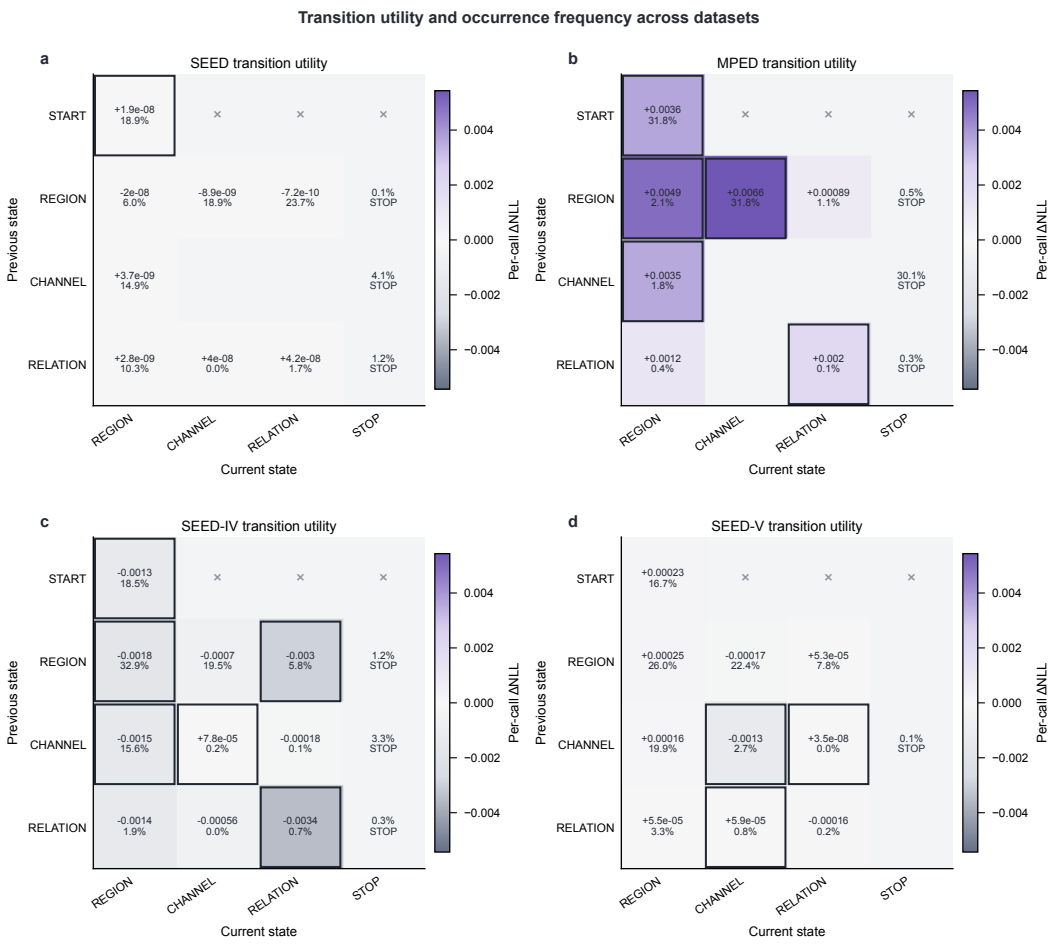

Figure 18: State-transition utility and occurrence frequency on all four datasets. Cell colour gives mean per-call ΔNLL and text gives transition frequency. STOP cells report frequency only; $\times$ denotes a transition excluded by the typed action grammar. Outlined utility cells have 95% CIs excluding zero.

The transition matrices show how one legal grammar produces distinct execution regimes. The matched case concentrates on START$\rightarrow$REGION$\rightarrow$CHANNEL$\rightarrow$STOP, whereas the other datasets distribute more calls among repeated REGION states and selective RELATION transitions. The full view therefore links dataset-conditioned path geometry to the compact utility analysis in the main paper.