DESCRIBE-TO-ANCHOR: PROFILE-GUIDED RELATIONAL ANCHORING FOR SEMANTICALLY GROUNDED MULTIVARIATE TIME SERIES CLUSTERING

Anonymous authors
Paper under double-blind review

ABSTRACT

Deep clustering of unlabeled multivariate time series (MTS) increasingly relies on self-supervised representations derived from reconstruction, augmented views, or evolving pseudo-labels. Without an independent semantic reference, these objectives may capture incidental regularities, drift with the model, or reinforce degenerate geometry. Deterministic temporal descriptors provide stable and interpretable knowledge, but using them as inputs or representations creates a mismatch with fine-grained sequence dynamics. This paper, therefore, proposes PRISM (**P**rofile-guided **R**epresentation learning with **I**ntrinsic **S**tructure **M**atching), a Describe-to-Anchor framework that converts deterministic profiles into fixed relational guidance. Because profile prototypes capture only coarse organization, PRISM complements the prototype anchor with a neighborhood anchor that preserves local distinctions. Since descriptors omit fine-grained dynamics, a temporal patch encoder learns from raw MTS while matching both anchors. This division grounds the learned geometry in explicit temporal characteristics without limiting its representational capacity. Experiments on 30 UEA benchmark datasets demonstrate strong performance against recent baselines, including the best macro-average score and average rank under all four metrics. Statistical tests, anchor-target diagnostics, and semantic traceability analyses on two real-world datasets further validate fixed profile relations.

1 INTRODUCTION

Multivariate time series (MTS) clustering seeks latent process structure in unlabeled sensor streams (Liu et al., 2026), industrial traces (Wen et al., 2022), clinical recordings (Gupta et al., 2020), and system monitoring data (Su et al., 2019). Complex temporal dynamics and cross-variable interactions make this structure difficult to specify directly (Paparrizos & Gravano, 2015). Deep representation learning extracts fine-grained patterns from raw sequences (Cheng et al., 2026), but typically relies on reconstruction, augmented views, or evolving pseudo-labels. These self-contained objectives lack an independent semantic reference and may capture incidental regularities or reinforce unstable group relations. This limitation matters in clustering, where meaningful partitions should expose the temporal properties underlying process similarity (Huang et al., 2025; Boniol et al., 2025). Predefined distances (Palpanas, 2015), temporal profiles (Wang et al., 2011), and expert-designed indicators (Zhang & Sun, 2023) provide inspectable semantics, yet impose fixed inductive biases. Their compressed descriptions may also omit local events and cross-variable interactions (Guerrini et al., 2025), particularly in long, high-dimensional sequences (Wang et al., 2026c). Therefore, the central challenge is to introduce stable external temporal knowledge without constraining representations learned from raw MTS.

Existing deep objectives lack a fixed external reference for representation learning. Reconstruction-based methods (Zhang et al., 2024) recover raw sequences but may entangle shared waveforms with sparse group-relevant signals (Wang et al., 2026c). Contrastive methods (Meng et al., 2023) enforce view consistency but assume that transformations such as cropping and scaling preserve temporal semantics (Wang et al., 2026b).

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.

Prototype and pseudo-label objectives approach clustering directly, but their evolving predictions may reinforce early errors (Eldele et al., 2024). Despite their differences, these paradigms derive supervision internally, providing neither independent temporal semantics nor fixed external relations. Deterministic temporal profiles provide such a reference through auditable descriptors (Lubba et al., 2019) with explicit temporal properties (Bonifati et al., 2022). Nevertheless, using profiles as encoder inputs forces compressed summaries and fine-grained dynamics into a shared representation. As illustrated in Figure 1(a), fusion may bury profile semantics or suppress informative local patterns. Its sensitivity to α in Figure 1(b) further shows that this balance is dataset-dependent. The key problem is therefore not whether profiles contain useful knowledge, but how they should guide learning: fixed profile relations should shape geometry without becoming the representation itself.

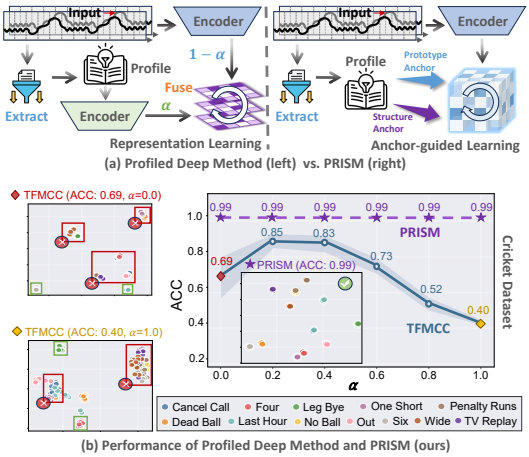


Figure 1: Motivation for Describe-to-Anchor. (a) Direct fusion mixes profiles with raw-sequence representations, whereas PRISM uses profile-derived relational anchors. (b) Fusion is sensitive to α, while PRISM matches fixed profile relations.

In light of this, we propose **PRISM** (**P**rofile-guided **R**epresentation learning with **I**ntrinsic **S**tructure **M**atching), which realizes the *Describe-to-Anchor* principle by converting deterministic temporal profiles into fixed relational targets rather than encoder inputs. The prototype anchor provides stable coarse organization, but its global abstraction may blur distinctions among nearby samples; PRISM therefore introduces a neighborhood anchor to preserve local profile relations. Since profiles cannot retain all fine-grained dynamics, a temporal patch encoder learns directly from raw MTS while using both anchors only as structural guidance. This division avoids heterogeneous feature fusion and keeps the guidance independent of encoder optimization, grounding the embedding geometry in explicit temporal semantics while preserving local patterns and cross-channel context. The contributions are summarized as follows:

- Deterministic temporal profiles are repurposed as fixed relational anchors through the *Describe-to-Anchor* principle rather than used as encoder inputs or final representations. Stable and interpretable external temporal semantics are thereby introduced without constraining raw-sequence learning.
- Complementary prototype and neighborhood anchors are combined with a temporal patch encoder to provide coarse and local guidance while retaining fine-grained dynamics. Theoretical analysis characterizes target-feedback removal, non-collapse conditions, and stability under profile perturbations.
- Target independence is preserved by fixed profile relations, which also link learned groups to explicit temporal attributes. Across 30 UEA benchmark datasets, PRISM achieves the best average scores and ranks on all four metrics; target diagnostics, real-world traceability, robustness, and efficiency studies further support its design.

2 Related Work

Fixed Temporal Priors. Predefined MTS clustering methods impose grouping structure through fixed input-space priors. Elastic distances align shifted or stretched observations across variables (Paparrizos et al., 2023), while shape-based measures compare normalized waveforms (Paparrizos & Gravano, 2015); their effectiveness depends on dataset characteristics and the assumed similarity (Paparrizos & Gravano, 2017; d'Hondt et al., 2025). Motif and salient-subsequence methods instead focus on discriminative local regions (Guerrini et al., 2025; Zhang et al., 2019), whereas feature-based methods represent sequences using compact temporal descriptors (Tiano et al., 2021), whose cross-dataset behavior has also been studied (Su et al., 2025). These priors provide auditable semantics but remain bounded by their construction: global measures may overlook short events,

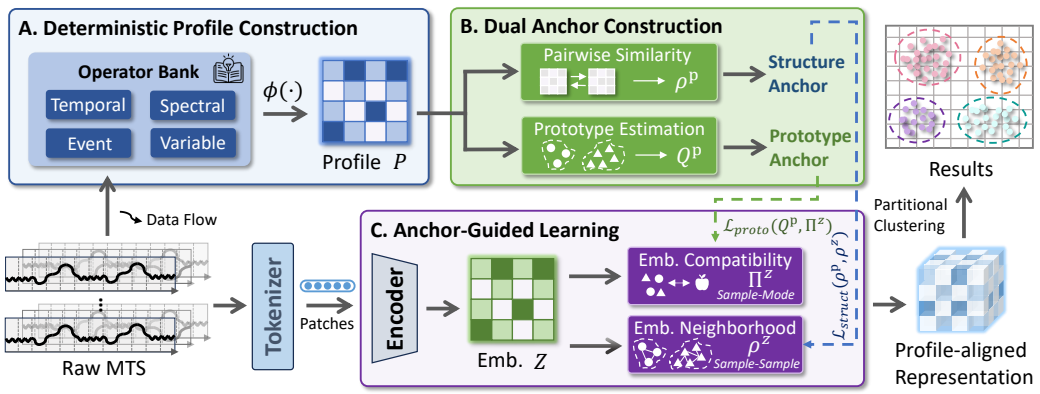


Figure 2: **Overview of PRISM.** (A) Deterministic operators transform raw MTS into the fixed profile matrix $\boldsymbol{P}$. (B) Pairwise similarity and prototype estimation construct fixed neighborhood and prototype anchors encoding local and coarse temporal relations. (C) A patch encoder matches the embedding-side relations, transferring profile semantics without feature fusion or target updates.

and compact descriptors cannot preserve unencoded local patterns or cross-variable interactions. Existing approaches also use such knowledge primarily as the similarity or representation itself. PRISM instead converts incomplete yet stable temporal profiles into fixed relational guidance for representations learned from raw MTS.

Self-Supervised Deep Clustering. Deep methods learn adaptive grouping geometry from the data or evolving model. Reconstruction and masked-modeling methods recover observed or removed values Chowdhury et al. (2022), often with richer temporal encoders or clustering losses (Senane et al., 2024; Dong et al., 2023), but may retain background behavior and noise (Cheng et al., 2026). Contrastive methods learn consistency across views (Wu et al., 2024; Liu & Chen, 2024) generated by parameterized augmentations or time-frequency transformations (Zheng et al., 2024; Zhang et al., 2022), with soft relations relaxing rigid pairwise assumptions (Lee et al., 2024). Their semantics remain determined by the selected views. Prototype-based methods encourage cluster organization (Sun et al., 2024), while uncertainty-aware variants reduce unreliable assignments (Gong et al., 2025), but evolving targets may propagate early errors. These paradigms differ in objective but provide no fixed external semantic reference independent of optimization. PRISM addresses this gap by matching embeddings to deterministic profile relations while keeping profiles outside the encoder.

3 Proposed Method

PRISM learns MTS representations from fixed, profile-derived relations rather than profile inputs. As shown in Figure 2, a prototype anchor provides coarse sample-to-mode organization, while a neighborhood anchor preserves local relations. A temporal patch encoder matches both anchors while retaining fine-grained dynamics and cross-variable context omitted by the profiles. This separation keeps guidance independent of encoder optimization and transfers explicit temporal semantics without feature fusion. Table 5 summarizes the notation.

3.1 Problem Formulation

Let $\mathcal{X} = \{\boldsymbol{X}_i\}_{i=1}^{n}$ be an unlabeled dataset of n MTS samples, where $\boldsymbol{X}_i \in \mathbb{R}^{T\times C}$ contains T time steps and C variables. An encoder $f_{\boldsymbol{\theta}}(\cdot)$ maps each sample to a d_{z}-dimensional embedding:

$$\boldsymbol{z}_i = f_{\boldsymbol{\theta}}(\boldsymbol{X}_i) \in \mathbb{R}^{d_{\mathrm{z}}}, \qquad \boldsymbol{Z} = [\boldsymbol{z}_i^{\top}]_{i=1}^{n} \in \mathbb{R}^{n\times d_{\mathrm{z}}}. \tag{1}$$

Given a target number k of clusters, let $\boldsymbol{y} = [y_1, \ldots, y_n]^{\top}$ denote the cluster assignments and $\mathcal{M} = \{\boldsymbol{\mu}_c\}_{c=1}^{k}$ the embedding-space centroids. The final partition is obtained by

$$(\hat{\boldsymbol{y}}, \hat{\mathcal{M}}) = \arg\min_{\boldsymbol{y},\, \mathcal{M}} \sum_{i=1}^{n} \|\boldsymbol{z}_i - \boldsymbol{\mu}_{y_i}\|_2^2, \qquad \text{s.t.} \quad y_i \in \{1, \ldots, k\}. \tag{2}$$

Eq. (2) partitions a given embedding space but does not determine its geometry. PRISM therefore learns $\boldsymbol{\theta}$ by matching fixed, semantically explicit profile relations at coarse and local scales while retaining fine-grained raw-MTS dynamics.

3.2 Deterministic Profile Construction

Fixed relational guidance requires an interpretable reference space independent of encoder updates. PRISM builds it from deterministic temporal descriptors (Lubba et al., 2019), which provide external knowledge rather than learned features. After standardizing each variable of $\boldsymbol{X}_i \in \mathbb{R}^{T\times C}$ over time, descriptors are computed at three scales.

(1) **Quality and global scale.** Quality descriptors record valid-value, invalid-value, and constant-variable ratios. From the mean standardized sequence, PRISM extracts trend slope, volatility statistics including mean, standard deviation, range, interquartile range, and coefficient of variation, periodicity statistics including dominant-power ratio, maximum autocorrelation, and its lag, and event statistics including event density, large changes, turning points, and spikes.

(2) **Patch scale.** Overlapping windows capture local behavior through patch count, mean and maximum energy, mean volatility, mean temporal variation, and the maximum-energy patch location. These descriptors differ from the raw patch tokens processed by the encoder in Sec. 3.4.

(3) **Variable and dependency scale.** Variables are ranked by an activity score combining variance, temporal variation, trend, and periodicity, with at most C_{a} retained. Each activity-ranked slot records its activity ratio, trend, volatility range and standard deviation, event and spike densities, and periodicity strength. Cross-variable structure is summarized by mean absolute zero-lag correlation and the number of valid pairs.

The resulting blocks form the fixed profile matrix:

$$\boldsymbol{p}_i = \mathrm{Norm}_{\mathcal{X}}\left(\mathrm{concat}\left[\boldsymbol{p}_i^{\mathrm{qua}}, \boldsymbol{p}_i^{\mathrm{glo}}, \boldsymbol{p}_i^{\mathrm{pat}}, \boldsymbol{p}_i^{\mathrm{var}}, \boldsymbol{p}_i^{\mathrm{dep}}\right]\right) \in \mathbb{R}^{d_{\mathrm{p}}}, \qquad \boldsymbol{P} = [\boldsymbol{p}_i^{\top}]_{i=1}^{n} \in \mathbb{R}^{n\times d_{\mathrm{p}}}. \quad (3)$$

Here, $\mathrm{Norm}_{\mathcal{X}}(\cdot)$ applies coordinate-wise dataset standardization followed by sample-wise ℓ_2 normalization. The fixed matrix $\boldsymbol{P}$ supplies explicit but incomplete relations, while final representations are learned from raw MTS. Appendix B specifies the operators, thresholds, and dimensions, while Appendix I.3 separates profile and representation-learning effects.

3.3 Dual Anchor Construction

A single relational scale cannot capture both coarse organization and local variation. PRISM therefore converts the fixed profile matrix $\boldsymbol{P}$ into complementary prototype and neighborhood anchors.

Prototype anchor. To encode coarse profile semantics, PRISM quantizes the normalized profiles into m unit-norm prototypes $\mathcal{A} = \{\boldsymbol{a}_r\}_{r=1}^{m}$ using spherical k-means:

$$\mathcal{A} = \underset{\{\boldsymbol{a}_r:\|\boldsymbol{a}_r\|_2=1\}_{r=1}^{m}}{\arg\min} \sum_{i=1}^{n} \min_{1\le r\le m} \|\boldsymbol{p}_i - \boldsymbol{a}_r\|_2^2 . \quad (4)$$

Because hard assignments discard graded compatibility, PRISM retains each sample's soft affinity to all prototypes:

$$q_{i,r}^{\mathrm{p}} = \frac{\exp(\mathrm{sim}(\boldsymbol{p}_i, \boldsymbol{a}_r)/\tau_{\mathrm{a}})}{\sum_{\ell=1}^{m} \exp(\mathrm{sim}(\boldsymbol{p}_i, \boldsymbol{a}_\ell)/\tau_{\mathrm{a}})}. \quad (5)$$

Here, $\mathrm{sim}(\cdot,\cdot)$ denotes cosine similarity and τ_{a} controls sharpness. The matrix $\boldsymbol{Q}^{\mathrm{p}} = [q_{i,r}^{\mathrm{p}}] \in \mathbb{R}^{n\times m}$ encodes coarse sample-to-prototype relations. The number m sets the anchor resolution and need not equal the target cluster number k.

Neighborhood anchor. Coarse prototype affinities may obscure local distinctions. For sample i, let $\mathcal{N}_i^{\mathrm{p}}$ contain the indices of its κ most similar profiles, where $s_{i,j}^{\mathrm{p}} = \mathrm{sim}(\boldsymbol{p}_i, \boldsymbol{p}_j)$ for $j \neq i$. PRISM preserves their relative strengths through

$$\rho_{i,j}^{\mathrm{p}} = \frac{\exp(s_{i,j}^{\mathrm{p}}/\tau_{\mathrm{s}})}{\sum_{\ell\in\mathcal{N}_i^{\mathrm{p}}} \exp(s_{i,\ell}^{\mathrm{p}}/\tau_{\mathrm{s}})}, \qquad j \in \mathcal{N}_i^{\mathrm{p}}, \quad (6)$$

with $\rho^{\mathrm{p}}_{i,j} = 0$ otherwise. The sparse matrix $\boldsymbol{\rho}^{\mathrm{p}} = [\rho^{\mathrm{p}}_{i,j}] \in \mathbb{R}^{n\times n}$ encodes local profile relations. Together, $\boldsymbol{Q}^{\mathrm{p}}$ and $\boldsymbol{\rho}^{\mathrm{p}}$ provide semantically explicit coarse and local targets that remain fixed during training. Appendix C provides implementation details.

3.4 Anchor-Guided Learning

Profile and embedding coordinates are not directly comparable. PRISM therefore aligns their relations while keeping the fixed profile anchors outside the encoder and independent of its optimization.

Temporal patch encoder. Because profiles omit fine-grained dynamics, PRISM learns representations directly from raw MTS. Each sample is divided into synchronized patches $\boldsymbol{R}_i = \psi(\boldsymbol{X}_i) \in \mathbb{R}^{L\times w\times C}$, where each patch retains all C variables. A shared encoder captures local dynamics, cross-variable interactions, and cross-patch context:

$$\boldsymbol{z}_i = f_{\boldsymbol{\theta}}(\boldsymbol{X}_i) = \frac{g_{\boldsymbol{\theta}}(\boldsymbol{R}_i)}{\|g_{\boldsymbol{\theta}}(\boldsymbol{R}_i)\|_2} \in \mathbb{R}^{d_{\mathrm{z}}}. \tag{7}$$

Appendix D specifies the patch construction and encoder architecture.

Embedding-side prototype compatibility. To reproduce coarse relations across different spaces, PRISM maintains a trainable embedding prototype bank $\mathcal{B}^{\mathrm{z}} = \{\boldsymbol{b}^{\mathrm{z}}_r\}_{r=1}^m$ that shares indices but not coordinates with $\mathcal{A}$. For each sample,

$$\pi^{\mathrm{z}}_{i,r} = \frac{\exp(\mathrm{sim}(\boldsymbol{z}_i, \boldsymbol{b}^{\mathrm{z}}_r)/\tau_{\mathrm{z}})}{\sum_{\ell=1}^m \exp(\mathrm{sim}(\boldsymbol{z}_i, \boldsymbol{b}^{\mathrm{z}}_\ell)/\tau_{\mathrm{z}})}. \tag{8}$$

The matrix $\boldsymbol{\Pi}^{\mathrm{z}} = [\pi^{\mathrm{z}}_{i,r}] \in \mathbb{R}^{n\times m}$ defines embedding-side affinities over the prototype indices shared with $\boldsymbol{Q}^{\mathrm{p}}$. Aligning these distributions transfers coarse profile semantics without enforcing coordinate correspondence.

Embedding-side neighborhood structure. Selecting neighbors from evolving embeddings would make the local support model-dependent. PRISM instead evaluates affinities on the fixed profile neighborhoods:

$$s^{\mathrm{z}}_{i,j} = \mathrm{sim}(\boldsymbol{z}_i, \boldsymbol{z}_j), \qquad \rho^{\mathrm{z}}_{i,j} = \frac{\exp(s^{\mathrm{z}}_{i,j}/\tau_{\mathrm{s}})}{\sum_{\ell\in\mathcal{N}^{\mathrm{p}}_i} \exp(s^{\mathrm{z}}_{i,\ell}/\tau_{\mathrm{s}})}, \quad j \in \mathcal{N}^{\mathrm{p}}_i. \tag{9}$$

Setting $\rho^{\mathrm{z}}_{i,j} = 0$ outside $\mathcal{N}^{\mathrm{p}}_i$ gives the embedding-side neighborhood matrix $\boldsymbol{\rho}^{\mathrm{z}}$. The fixed anchors $(\boldsymbol{Q}^{\mathrm{p}}, \boldsymbol{\rho}^{\mathrm{p}})$ supervise their learned counterparts $(\boldsymbol{\Pi}^{\mathrm{z}}, \boldsymbol{\rho}^{\mathrm{z}})$, defining coarse and local alignment without target updates.

3.5 Optimization and Complexity

PRISM aligns the learned prototype and neighborhood relations with their fixed profile targets through

$$\mathcal{L} = \mathcal{L}_{\mathrm{proto}} + \lambda_{\mathrm{struct}}\mathcal{L}_{\mathrm{struct}}, \qquad \mathcal{L}_{\mathrm{proto}} = -\frac{1}{n}\sum_{i=1}^n\sum_{r=1}^m q^{\mathrm{p}}_{i,r}\log\pi^{\mathrm{z}}_{i,r}, \qquad \mathcal{L}_{\mathrm{struct}} = -\frac{1}{n}\sum_{i=1}^n\sum_{j\in\mathcal{N}^{\mathrm{p}}_i}\rho^{\mathrm{p}}_{i,j}\log\rho^{\mathrm{z}}_{i,j}. \tag{10}$$

The prototype term transfers coarse sample-to-prototype organization, while the structure term preserves relative affinities over the fixed profile neighborhoods. The coefficient $\lambda_{\mathrm{struct}} \geq 0$ balances the two scales. Restricting structure matching to κ neighbors avoids dense pairwise optimization.

For a seed batch $\mathcal{I}_B$, PRISM encodes the augmented set $\mathcal{V}_B = \mathcal{I}_B \cup \bigcup_{i\in\mathcal{I}_B}\mathcal{N}^{\mathrm{p}}_i$ and retains gradients for all included embeddings. The losses update the encoder $f_{\boldsymbol{\theta}}$ and embedding prototype bank $\mathcal{B}^{\mathrm{z}}$, whereas the profile anchors remain fixed and receive no gradients. After training, $\mathcal{B}^{\mathrm{z}}$ is discarded and the embeddings $\boldsymbol{Z}$ are partitioned using Eq. (2). The profile prototypes $\mathcal{A}$ construct the fixed targets, the embedding prototypes $\mathcal{B}^{\mathrm{z}}$ mediate alignment, and the final centroids $\mathcal{M}$ are estimated only from $\boldsymbol{Z}$. Consequently, the anchor resolution m need not equal the target cluster number k. Algorithm 1 in Appendix A summarizes the procedure.

Complexity. Profiling costs $O(nCT\log T + nTC_{\mathrm{a}}^2)$, prototype estimation costs $O(I_{\mathrm{p}}nmd_{\mathrm{p}})$, and exact neighborhood construction costs $O(n^2 d_{\mathrm{p}})$ time and $O(n^2)$ temporary memory when the similarity matrix is materialized. The persistent profiles and anchors require $O(nd_{\mathrm{p}} + nm + n\kappa)$

space. Beyond encoding $\mathcal{V}_B$, each update adds $O(B(m+\kappa)d_{\mathrm{z}})$ for relational alignment, giving $O(E\lceil n/B\rceil[\mathcal{C}_{\mathrm{enc}}(|\mathcal{V}_B|, L, w, C) + B(m+\kappa)d_{\mathrm{z}}])$ over E epochs. Final k-means costs $O(I_{\mathrm{k}}nkd_{\mathrm{z}})$. For fixed m and κ, alignment scales linearly with n, while approximate neighbor search can replace quadratic preprocessing at larger scales. Appendix F provides the complete analysis.

4 EXPERIMENTS

Our evaluation addresses five research questions: **RQ1:** Does PRISM improve MTS clustering compared with recent representation-learning and clustering methods? **RQ2:** Should temporal profiles serve as relational anchors rather than representations or encoder inputs, and are both relational scales necessary? **RQ3:** Do the gains require both the semantic content and optimization independence of fixed profile anchors? **RQ4:** How does fixed profile guidance shape assignment dynamics and prototype utilization throughout training? **RQ5:** Can learned groups be traced to explicit, interpretable temporal properties?

4.1 EXPERIMENTAL SETTINGS

Datasets. This paper evaluates PRISM on 30 datasets from the UEA MTS archive (Bagnall et al., 2018), spanning audio, ECG, EEG, gesture, human activity recognition, magnetoencephalography, motion, sensor, simulation, and traffic domains. MHEALTH and UCI-HAR are additionally used to examine the semantic traceability of learned groups in real-world human-activity data. Dataset statistics and preprocessing details are provided in Appendix G.

Baselines and metrics. We compare PRISM with nine recent methods, including general-purpose representation models TimesURL (Liu & Chen, 2024), UNITS (Gao et al., 2024), FEI (Fu & Hu, 2025), and GTM (He et al., 2026); deep clustering methods MVCIMTS (Li et al., 2025), TFMCC (Wang et al., 2026a), FCACC (Wang et al., 2026b), and MUFASA (Li & Paparrizos, 2026); and the interpretable structural method k-Graph (Boniol et al., 2025). Clustering quality is evaluated using accuracy (ACC), macro F1, normalized mutual information (NMI), and adjusted Rand index (ARI), with higher values indicating better performance.

Protocol and implementation. Official implementations and recommended configurations are used whenever available. Methods that output embeddings are followed by k-means with the dataset-specified cluster number. Ground-truth labels are used only for evaluation. We report the mean and standard deviation over five random seeds, together with unweighted average scores and ranks across the 30 UEA datasets. PRISM configurations, hardware, and reproducibility details are reported in Appendices D and G.

4.2 OVERALL CLUSTERING PERFORMANCE

Table 1 shows that PRISM achieves the highest macro-average score and lowest average rank under all four metrics. Its overall rank is 3.379, compared with 4.471 for FCACC, the strongest baseline by overall rank. Relative to the best competing average for each metric, PRISM improves ACC, F1, NMI, and ARI by 0.063, 0.065, 0.056, and 0.081, respectively. These gains span both label-matching metrics (ACC and F1) and permutation-invariant partition metrics (NMI and ARI), while the corresponding rank improvements indicate that the aggregate advantage is not driven solely by a few datasets. Using datasets as independent units, Friedman tests reject equal performance among the compared methods for every metric ($p < 0.00001$). Holm-corrected two-sided Wilcoxon signed-rank tests further show that PRISM significantly outperforms $7/9$, $9/9$, $6/9$, and $6/9$ baselines under ACC, F1, NMI, and ARI, respectively. The consistent improvements across label-matching and permutation-invariant metrics indicate that fixed profile relations improve clustering geometry while preserving fine-grained information learned from raw sequences. Complete dataset-level results and adjusted p-values are reported in Appendix I.1.

4.3 PROFILES AS RELATIONAL ANCHORS

Table 2 evaluates whether temporal profiles should serve as representations, encoder inputs, or fixed relational anchors. Profile-only preserves explicit temporal semantics but remains constrained by compressed descriptors, whereas lower Profile-input performance suggests direct fusion entangles

Table 1: Aggregate clustering results over all 30 datasets. Avg. denotes the unweighted macro average across datasets, and Rank denotes the average rank among all compared methods. Lower ranks are better. The best and second-best results are highlighted in dark and light blue, respectively.

Methods	ACC		F1		NMI		ARI		Overall
	Avg.	Rank	Avg.	Rank	Avg.	Rank	Avg.	Rank	Rank
TimesURL (AAAI'24)	0.409	5.833	0.412	5.283	0.219	5.533	0.143	5.883	5.633
UNITS (NeurIPS'24)	0.398	5.750	0.382	5.617	0.160	6.600	0.084	6.650	6.154
FEI (AAAI'25)	0.451	4.633	0.445	4.750	0.257	5.050	0.163	4.933	4.842
MVCIMTS (INFFUS'25)	0.297	7.683	0.201	9.117	0.070	7.833	0.032	7.783	8.104
k-Graph (TKDE'25)	0.409	5.850	0.392	5.933	0.210	6.000	0.124	6.050	5.958
TFMCC (AAAI'26)	0.464	5.217	0.449	4.433	0.309	4.067	0.201	4.217	4.483
FCACC (PR'26)	0.481	4.650	0.474	4.700	0.301	4.267	0.213	4.267	4.471
GTM (ICLR'26)	0.356	6.917	0.332	7.233	0.137	7.350	0.073	7.017	7.129
MUFASA (SIGMOD'26)	0.458	5.033	0.430	4.900	0.276	4.733	0.208	4.717	4.846
PRISM (Ours)	**0.544**	**3.433**	**0.539**	**3.033**	**0.365**	**3.567**	**0.294**	**3.483**	**3.379**

Table 2: Core ablations averaged over 30 UEA datasets. "Proto." and "Neigh." denote the prototype and neighborhood anchors, respectively. Higher values are better.

Variant	Profile role	Proto.	Neigh.	ACC	F1	NMI	ARI
Profile-only	Representation	–	–	0.472	0.465	0.287	0.200
Profile-input	Encoder input	–	–	0.411	0.397	0.219	0.130
w/o neighborhood	Anchor	✓	–	0.519	0.512	0.345	0.261
w/o prototype	Anchor	–	✓	0.526	0.520	0.349	0.271
Full PRISM	**Anchor**	✓	✓	**0.544**	**0.539**	**0.365**	**0.294**

coarse summaries with fine-grained sequence dynamics. Using profiles as fixed relational anchors instead yields the strongest aggregate performance. Removing either the prototype or neighborhood anchor degrades all four metrics, confirming that coarse sample-to-mode organization and local sample relations provide complementary guidance. Full PRISM is best or tied best in 80 of 120 dataset-metric cases for profile-role comparisons and 65 of 120 cases for anchor-component comparisons. Complete results are provided in Appendices I.3 and I.4.

4.4 Semantic Content and Target Independence

Table 3 isolates two essential properties of relational guidance: semantic content and optimization independence. Random and shuffled controls remain fixed but replace or permute profile relations, whereas learnable and EMA variants preserve profile initialization while allowing targets to follow the evolving model. Fixed profile anchors achieve the highest macro average and lowest average rank under all four metrics. Together, these comparisons show that fixedness without profile semantics is insufficient and that semantic initialization becomes less effective when target relations evolve with representation optimization. This evidence directly supports PRISM's use of profiles as an independently constructed semantic reference. Variant definitions and epoch-wise training dynamics are provided in Appendix I.5.

4.5 Optimization Dynamics under Fixed Guidance

Table 4 compares assignment evolution and prototype utilization under different targets. Pseudo-labeling has the lowest drift, but lower effective utilization and ACC suggest premature locking rather than better convergence. Fixed profile guidance achieves the highest active- and effective-prototype ratios and ACC with second-best midpoint-to-final consistency. Relative to learnable anchors, it reduces drift from 0.196 to 0.189, increases effective-prototype utilization from 0.656

Table 3: Anchor-target analysis on 10 diagnostic UEA datasets. Results are macro averages over datasets and five seeds, whereas ranks are averaged across datasets. Lower ranks are better. Best and second-best results are bold and underlined, respectively.

Anchor target	Results				Avg. Rank			
	ACC	F1	NMI	ARI	ACC	F1	NMI	ARI
Fixed controls without intact profile semantics								
Random fixed	0.629	0.627	0.481	0.409	4.150	3.900	4.400	4.500
Shuffled fixed	0.605	0.594	0.466	0.385	5.400	5.700	4.700	4.700
Profile-initialized adaptive targets								
Learnable-P	<u>0.670</u>	<u>0.664</u>	<u>0.539</u>	<u>0.473</u>	3.950	4.100	<u>3.400</u>	3.800
Learnable-N	0.660	0.655	0.517	0.452	<u>3.300</u>	<u>3.500</u>	4.000	<u>3.600</u>
Learnable-Both	0.644	0.632	0.509	0.440	3.800	3.900	3.700	3.700
EMA-updated	0.633	0.630	0.486	0.411	4.500	4.300	4.700	4.700
Fixed profile (PRISM)	**0.684**	**0.681**	**0.540**	**0.482**	**2.900**	**2.600**	**3.100**	**3.000**

Table 4: Training dynamics on 10 diagnostic UEA datasets. Values are macro averages over datasets and five seeds. $\bar{D}_{\text{assign}}$ measures adjacent-epoch assignment drift, C_{mid} midpoint-to-final NMI, and Act./Eff. normalized prototype utilization. Best and second-best results are bold and underlined, respectively.

Target	$\bar{D}_{\text{assign}} \downarrow$	$C_{\text{mid}} \uparrow$	Act.↑	Eff.↑	ACC↑
Learnable anchor	0.196	0.750	0.864	0.656	0.634
EMA anchor	<u>0.178</u>	**0.773**	0.897	<u>0.832</u>	<u>0.640</u>
Pseudo-label	**0.116**	0.762	<u>0.929</u>	0.737	0.630
Fixed profile (PRISM)	0.189	<u>0.769</u>	**0.957**	**0.850**	**0.678**

to 0.850, and improves ACC from 0.634 to 0.678. Although EMA targets yield lower drift and higher final consistency, they use fewer prototypes and produce lower clustering quality. Fixed semantic guidance therefore maintains assignment plasticity while promoting broad prototype use. Appendix I.5 reports the corresponding epoch-wise trajectories.

4.6 Semantic Traceability of Learned Groups

Figure 3 traces the learned organization through raw-sequence and descriptor views of explicit temporal properties. For each embedding-side prototype, the left panels summarize its assigned samples across channels: colored curves show group mean, shaded regions show within-group variation, and gray curves show representative sequences. Titles report the aligned prototype index, post-hoc majority activity, and purity. Right panels decompose the profile prototypes into explicit temporal-property compositions. Repetitive activity groups emphasize periodicity and volatility, whereas localized or irregular movements rely more strongly on event density, patch variation, or cross-channel coupling. Labels are used only after training. Together, these views connect learned groups to observable patterns and profile attributes, with formal diagnostics in Appendix H.

4.7 Robustness and Efficiency

Figure 4 summarizes robustness to input noise, sensitivity to the embedding dimension, and end-to-end scalability. PRISM maintains consistently high macro F1 as the noise level increases and remains stable across embedding dimensions. Moreover, its runtime grows smoothly with the sample size and remains substantially lower than those of the compared deep clustering methods. These results indicate that profile-guided relational alignment is robust to input perturbations and representation capacity without introducing substantial computational overhead. Dataset-level results and additional scalability analyses are provided in Appendix I.6.

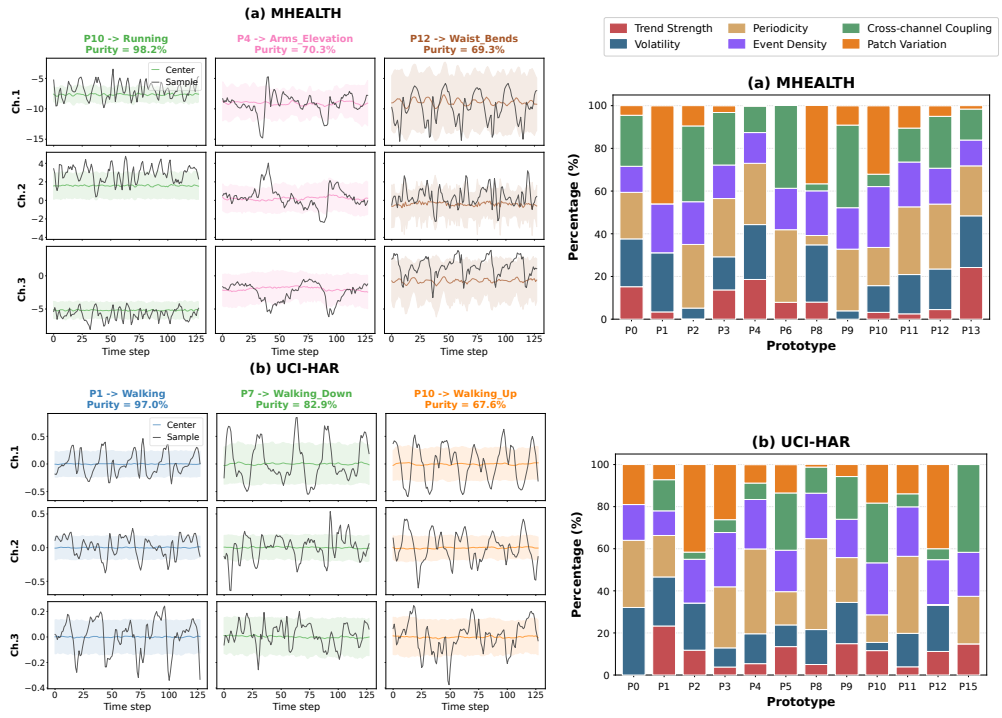


Figure 3: **Semantic traceability on MHEALTH and UCI-HAR.** Left panels compare learned prototypes with representative three-channel sequences. Right panels show their temporal-property compositions. Activity names and purity are computed post hoc without training labels.

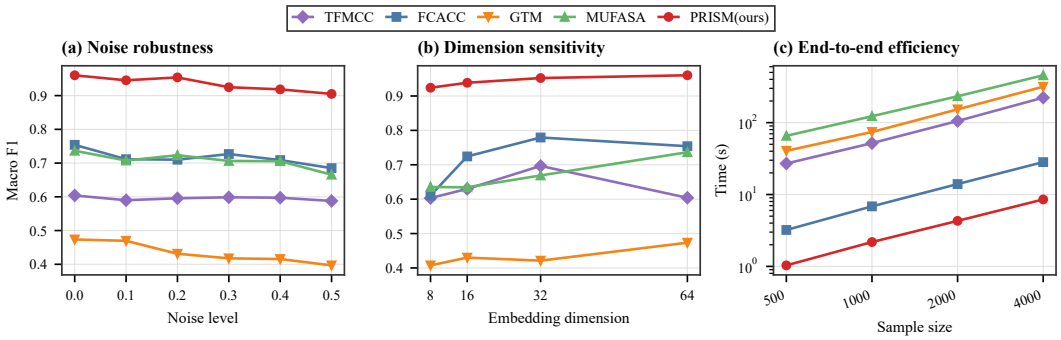


Figure 4: **Robustness, dimension sensitivity, and efficiency.** Panels (a) and (b) report macro F1 averaged over four datasets, while panel (c) reports runtime as sample size increases.

5 Conclusion

MTS clustering requires temporal guidance that is semantically explicit and independent of representation optimization. Therefore, PRISM converts deterministic temporal profiles into fixed prototype and neighborhood anchors, while a patch encoder retains omitted fine-grained dynamics. Theoretical analysis characterizes target-feedback removal, conditions excluding global collapse, and stability under profile perturbations. Across 30 UEA datasets, PRISM achieves the best macro-average score and average rank under ACC, F1, NMI, and ARI. Further analyses validate dual-scale guidance, anchor semantics, target independence, optimization dynamics, and semantic traceability of learned groups. Its effectiveness nevertheless depends on whether the selected descriptors capture group-relevant temporal properties, and exact neighborhood construction remains quadratic. Future work will explore adaptive descriptor selection and approximate anchor construction while retaining an independently inspectable semantic reference.

AI USE STATEMENT

OpenAI's ChatGPT was used for language editing, manuscript organization, concision, LaTeX formatting, and feedback on the presentation of technical arguments, experiments, and results. It was not used to generate experimental data or cluster labels, execute experiments, or fabricate or selectively report results. All AI-assisted suggestions concerning mathematical statements, experimental design, and interpretation were independently evaluated against the authors' derivations, implementations, and empirical evidence, and were revised or rejected where necessary. All references, equations, code, and reported results were verified by the authors. The authors take full responsibility for the final content of this paper.

ETHICS STATEMENT

This study uses only publicly available benchmark datasets and does not involve new data collection, human-subject recruitment, or access to personally identifiable information. We follow the original dataset licenses and use labels only for evaluation. Although PRISM is intended for exploratory analysis, clustering outcomes may reflect biases in the data or predefined temporal descriptors. Deployments in clinical, industrial, or other high-stakes settings should therefore include domain-specific validation and human oversight. We are unaware of any conflicts of interest or other significant ethical concerns.

REPRODUCIBILITY STATEMENT

Reproducibility is supported through complete definitions, implementation details, and released code. Algorithm 1 summarizes the training procedure, while Appendices B–D specify profile construction, anchor estimation, encoder architecture, and optimization. Appendix E provides the assumptions and proofs, and Appendix G documents datasets, evaluation protocols, random seeds, and hardware. Appendix H defines the diagnostic and interpretability measures, while Appendix I reports complete results, statistical tests, ablations, anchor analyses, training dynamics, robustness, sensitivity, and efficiency. The anonymized implementation and configurations are available at `https://anonymous.4open.science/r/PRISM-demo`.

REFERENCES

Anthony Bagnall, Hoang Anh Dau, Jason Lines, Michael Flynn, James Large, Aaron Bostrom, Paul Southam, and Eamonn Keogh. The uea multivariate time series classification archive, 2018. URL `https://arxiv.org/abs/1811.00075`.

Angela Bonifati, Francesco Del Buono, Francesco Guerra, and Donato Tiano. Time2feat: learning interpretable representations for multivariate time series clustering. *Proceedings of the VLDB Endowment*, 16(2):193–201, 2022.

Paul Boniol, Donato Tiano, Angela Bonifati, and Themis Palpanas. k-graph: A graph embedding for interpretable time series clustering. *IEEE Transactions on Knowledge and Data Engineering*, 37(5):2680–2694, 2025.

Mingyue Cheng, Xiaoyu Tao, Zhiding Liu, Qi Liu, Hao Zhang, Rujiao Zhang, and Enhong Chen. Timemae: Self-supervised representations of time series with decoupled masked autoencoders. In *Proceedings of the ACM International Conference on Web Search and Data Mining (WSDM)*, pp. 498–508, New York, NY, USA, 2026. Association for Computing Machinery.

Ranak Roy Chowdhury, Xiyuan Zhang, Jingbo Shang, Rajesh K. Gupta, and Dezhi Hong. Tarnet: Task-aware reconstruction for time-series transformer. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD)*, pp. 212–220, New York, NY, USA, 2022. Association for Computing Machinery.

Jens E. d'Hondt, Haojun Li, Fan Yang, Odysseas Papapetrou, and John Paparrizos. A structured study of multivariate time-series distance measures. *Proceedings of the ACM on Management of Data*, 3(3):1–29, 2025.

Jiaxiang Dong, Haixu Wu, Haoran Zhang, Li Zhang, Jianmin Wang, and Mingsheng Long. Simmtm: A simple pre-training framework for masked time-series modeling. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, pp. 29996–30025, Red Hook, NY, USA, 2023. Curran Associates, Inc.

Emadeldeen Eldele, Mohamed Ragab, Zhenghua Chen, Min Wu, Chee-Keong Kwoh, and Xiaoli Li. Label-efficient time series representation learning: A review. *IEEE Transactions on Artificial Intelligence*, 5(12):6027–6042, 2024.

En Fu and Yanyan Hu. Frequency-masked embedding inference: A non-contrastive approach for time series representation learning. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, pp. 16639–16647, Washington, DC, USA, 2025. AAAI Press.

Shanghua Gao, Teddy Koker, Owen Queen, Thomas Hartvigsen, Theodoros Tsiligkaridis, and Marinka Zitnik. Units: A unified multi-task time series model. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, volume 37, pp. 140589–140631, Red Hook, NY, USA, 2024. Curran Associates, Inc.

Peiliang Gong, Mohamed Ragab, Min Wu, Zhenghua Chen, Yongyi Su, Xiaoli Li, and Daoqiang Zhang. Augmented contrastive clustering with uncertainty-aware prototyping for time series test time adaptation. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD)*, pp. 390–401, New York, NY, USA, 2025. Association for Computing Machinery.

Valerio Guerrini, Thibaut Germain, Charles Truong, Laurent Oudre, and Paul Boniol. Time series motif discovery: A comprehensive evaluation. *Proceedings of the VLDB Endowment*, 18(7): 2226–2239, 2025.

Priyanka Gupta, Pankaj Malhotra, Jyoti Narwariya, Lovekesh Vig, and Gautam Shroff. Transfer learning for clinical time series analysis using deep neural networks. *Journal of Healthcare Informatics Research*, 4(2):112–137, 2020.

Cheng He, Xu Huang, Gangwei Jiang, Zhaoyi Li, Defu Lian, Hong Xie, Enhong Chen, Xijie Liang, Zengrong Zheng, and Patrick P. C. Lee. GTM: A general time-series model for enhanced representation learning of time-series data. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2026. URL `https://openreview.net/forum?id=PWM6FERWz9`.

Zheng Huang, Hao Hao, and Lun Du. Exploring the explainability of time series clustering: A review of methods and practices. In *Proceedings of the ACM International Conference on Web Search and Data Mining (WSDM)*, pp. 1005–1007, New York, NY, USA, 2025. Association for Computing Machinery.

Seunghan Lee, Taeyoung Park, and Kibok Lee. Soft contrastive learning for time series. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2024. URL `https://openreview.net/forum?id=pAsQSWlDUf`.

Haojun Li and John Paparrizos. Mufasa: Fast and accurate multivariate time-series clustering. *Proceedings of the ACM on Management of Data*, 4(3):1–29, 2026.

Yurui Li, Mingjing Du, Xiang Jiang, and Nan Zhang. Contrastive learning-based multi-view clustering for incomplete multivariate time series. *Information Fusion*, 117:102812, 2025.

Jiexi Liu and Songcan Chen. Timesurl: Self-supervised contrastive learning for universal time series representation learning. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, pp. 13918–13926, Washington, DC, USA, 2024. AAAI Press.

Xin Liu, Dongzhou Cheng, Zenan Fu, Lei Zhang, Hao Wu, and Aiguo Song. Sensor-prompt tuning: Aligning time series foundational models with motion sensors for few-shot activity recognition. *IEEE Transactions on Mobile Computing*, 25(7):11227–11243, 2026.

Carl H. Lubba, Sarab S. Sethi, Philip Knaute, Simon R. Schultz, Ben D. Fulcher, and Nick S. Jones. catch22: Canonical time-series characteristics. *Data Mining and Knowledge Discovery*, 33:1821–1852, 2019.

Qianwen Meng, Hangwei Qian, Yong Liu, Lizhen Cui, Yonghui Xu, and Zhiqi Shen. Mhccl: Masked hierarchical cluster-wise contrastive learning for multivariate time series. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, pp. 9153–9161, Washington, DC, USA, 2023. AAAI Press.

Themis Palpanas. Data series management: The road to big sequence analytics. *ACM SIGMOD Record*, 44(2):47–52, 2015.

John Paparrizos and Luis Gravano. k-shape: Efficient and accurate clustering of time series. In *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD)*, pp. 1855–1870, New York, NY, USA, 2015. Association for Computing Machinery.

John Paparrizos and Luis Gravano. Fast and accurate time-series clustering. *ACM Transactions on Database Systems*, 42(2):1–49, 2017.

John Paparrizos, Kaize Wu, Aaron Elmore, Christos Faloutsos, and Michael J. Franklin. Accelerating similarity search for elastic measures: A study and new generalization of lower bounding distances. *Proceedings of the VLDB Endowment*, 16(8):2019–2032, 2023.

Zineb Senane, Lele Cao, Valentin Leonhard Buchner, Yusuke Tashiro, Lei You, Pawel Andrzej Herman, Mats Nordahl, Ruibo Tu, and Vilhelm von Ehrenheim. Self-supervised learning of time series representation via diffusion process and imputation-interpolation-forecasting mask. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD)*, pp. 2560–2571, New York, NY, USA, 2024. Association for Computing Machinery.

Ya Su, Youjian Zhao, Chenhao Niu, Rong Liu, Wei Sun, and Dan Pei. Robust anomaly detection for multivariate time series through stochastic recurrent neural network. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD)*, pp. 2828–2837, New York, NY, USA, 2019. Association for Computing Machinery.

Yunxiang Su, Kenny Ye Liang, and Shaoxu Song. In-database time series clustering. *Proceedings of the ACM on Management of Data*, 3(1):1–26, 2025.

Chenxi Sun, Hongyan Li, Yaliang Li, and Shenda Hong. Test: Text prototype aligned embedding to activate llm's ability for time series. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2024. URL https://openreview.net/forum?id=Tuh4nZVb0g.

Donato Tiano, Angela Bonifati, and Raymond Ng. Featts: Feature-based time series clustering. In *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD)*, pp. 2784–2788, New York, NY, USA, 2021. Association for Computing Machinery.

Congyu Wang, Mingjing Du, and Xiang Jiang. Time-frequency augmented multi-level contrastive clustering for time series. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, pp. 26142–26150, Washington, DC, USA, 2026a. AAAI Press.

Congyu Wang, Mingjing Du, Xiang Jiang, and Yongquan Dong. Fuzzy cluster-aware contrastive clustering for time series. *Pattern Recognition*, 173:112899, 2026b.

Peng Wang, Haixun Wang, and Wei Wang. Finding semantics in time series. In *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD)*, pp. 385–396, New York, NY, USA, 2011. Association for Computing Machinery.

Yuxuan Wang, Haixu Wu, Jiaxiang Dong, Yong Liu, Chen Wang, Mingsheng Long, and Jianmin Wang. Deep time series models: A comprehensive survey and benchmark. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–20, 2026c.

Qingsong Wen, Linxiao Yang, Tian Zhou, and Liang Sun. Robust time series analysis and applications: An industrial perspective. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD)*, pp. 4836–4837, New York, NY, USA, 2022. Association for Computing Machinery.

Yu Wu, Ting Dang, Dimitris Spathis, Hong Jia, and Cecilia Mascolo. Statiocl: Contrastive learning for time series via non-stationary and temporal contrast. In *Proceedings of the ACM International Conference on Information and Knowledge Management (CIKM)*, pp. 2575–2584, New York, NY, USA, 2024. Association for Computing Machinery.

Nan Zhang and Shiliang Sun. Multiview unsupervised shapelet learning for multivariate time series clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4):4981–4996, 2023.

Qin Zhang, Jia Wu, Peng Zhang, Guodong Long, and Chengqi Zhang. Salient subsequence learning for time series clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41 (9):2193–2207, 2019.

Wenrui Zhang, Ling Yang, Shijia Geng, and Shenda Hong. Self-supervised time series representation learning via cross reconstruction transformer. *IEEE Transactions on Neural Networks and Learning Systems*, 35(11):16129–16138, 2024.

Xiang Zhang, Ziyuan Zhao, Theodoros Tsiligkaridis, and Marinka Zitnik. Self-supervised contrastive pre-training for time series via time-frequency consistency. In *Proceedings of the Annual Conference on Neural Information Processing Systems (NeurIPS)*, pp. 3988–4003, Red Hook, NY, USA, 2022. Curran Associates, Inc.

Xu Zheng, Tianchun Wang, Wei Cheng, Aitian Ma, Haifeng Chen, Mo Sha, and Dongsheng Luo. Parametric augmentation for time series contrastive learning. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2024. URL `https://openreview.net/forum?id=EIPLdFy3vp`.

A Notation and Complete Procedure

Table 5 summarizes the notation, and Algorithm 1 presents the complete procedure. Profiles and their relational anchors are constructed once before representation learning and remain fixed throughout training. Only the temporal encoder and embedding prototype bank are jointly optimized, preserving target independence from encoder updates.

B Deterministic Profile Details

This section specifies the deterministic operators used in Sec. 3.2. All descriptors are computed without labels or trainable parameters and provide the fixed semantic reference from which PRISM constructs its relational anchors.

B.1 Preprocessing and Global Sequences

For sample $\boldsymbol{X}_i$, the quality statistics are computed before preprocessing. Invalid entries are then replaced by their variable-wise temporal means. Variable c is standardized over time as

$$\widetilde{X}_{i,t,c} = \frac{X_{i,t,c} - \mu_{i,c}}{\sigma_{i,c} + \epsilon}, \tag{11}$$

where $\mu_{i,c}$ and $\sigma_{i,c}$ are computed over valid timestamps. Constant variables are mapped to zero after standardization. PRISM forms a signed global sequence and a nonnegative activity sequence:

$$g_{i,t} = \frac{1}{C}\sum_{c=1}^{C} \widetilde{X}_{i,t,c}, \qquad u_{i,t} = \frac{1}{C}\sum_{c=1}^{C} \left|\widetilde{X}_{i,t,c}\right|. \tag{12}$$

The signed sequence $\boldsymbol{g}_i$ describes collective temporal variation, while $\boldsymbol{u}_i$ measures scale-independent activity magnitude.

Table 5: Main notation used in PRISM.

Symbol	Description
$\mathcal{X}, n$	Unlabeled MTS dataset and number of samples.
$\boldsymbol{X}_i$	The i-th MTS sample in $\mathbb{R}^{T\times C}$.
T, C, C_{a}, k	Numbers of timestamps, variables, retained active variables, and target clusters.
$\boldsymbol{p}_i, \boldsymbol{P}, d_{\mathrm{p}}$	Temporal profile, profile matrix, and profile dimension.
$\mathcal{A}, \boldsymbol{a}_r, m$	Profile prototypes, the r-th prototype, and prototype number.
$\boldsymbol{Q}^{\mathrm{p}}$	Fixed profile-side prototype anchor.
$\mathcal{N}_i^{\mathrm{p}}, \kappa$	Profile neighborhood of sample i and its size.
$\boldsymbol{\rho}^{\mathrm{p}}$	Fixed profile-side neighborhood anchor.
$f_{\boldsymbol{\theta}}$	Temporal patch encoder with parameters $\boldsymbol{\theta}$.
$\boldsymbol{z}_i, \boldsymbol{Z}, d_{\mathrm{z}}$	Sample embedding, embedding matrix, and embedding dimension.
$\mathcal{B}^{\mathrm{z}}, \boldsymbol{b}_r^{\mathrm{z}}$	Trainable embedding prototype bank and its r-th prototype.
$\boldsymbol{\Pi}^{\mathrm{z}}$	Embedding-side prototype compatibility matrix.
$\boldsymbol{\rho}^{\mathrm{z}}$	Embedding-side neighborhood distribution.
$\mathcal{I}_B, \mathcal{V}_B$	Seed indices and neighbor-augmented indices of a mini-batch.

Algorithm 1 PRISM: Describe-to-Anchor Representation Learning

Input: Dataset $\mathcal{X}$, cluster number k, prototype number m, neighborhood size κ, epochs E, and batch size B

Output: Assignments $\hat{\boldsymbol{y}}$ and embeddings $\boldsymbol{Z}$

Construct the profile matrix $\boldsymbol{P}$ using Eq. (3);
Estimate $\mathcal{A}$ and construct $\boldsymbol{Q}^{\mathrm{p}}$ using Eqs. (4) and (5);
Construct $\{\mathcal{N}_i^{\mathrm{p}}\}_{i=1}^n$ and $\boldsymbol{\rho}^{\mathrm{p}}$ using Eq. (6);
Initialize the encoder $f_{\boldsymbol{\theta}}$ and embedding prototype bank $\mathcal{B}^{\mathrm{z}}$;
for $e = 1$ to E **do**
  for each seed batch $\mathcal{I}_B$ **do**
    Form $\mathcal{V}_B = \mathcal{I}_B \cup \bigcup_{i\in\mathcal{I}_B} \mathcal{N}_i^{\mathrm{p}}$;
    Encode $\{\boldsymbol{X}_j : j \in \mathcal{V}_B\}$ with gradients retained for all embeddings;
    Compute $\boldsymbol{\Pi}^{\mathrm{z}}$ and $\boldsymbol{\rho}^{\mathrm{z}}$ on the seed samples and their fixed profile neighbors;
    Update $\boldsymbol{\theta}$ and $\mathcal{B}^{\mathrm{z}}$ using Eq. (10);
  end for
end for
Encode all samples to obtain $\boldsymbol{Z} = [\boldsymbol{z}_i^\top]_{i=1}^n$;
Apply Eq. (2) to $\boldsymbol{Z}$ and obtain $\hat{\boldsymbol{y}}$;
return $\hat{\boldsymbol{y}}$ and $\boldsymbol{Z}$.

B.2 Quality and Global Descriptors

Quality. The quality block records the proportions of valid values, invalid values, and variables that remain constant over time.

Trend. A least-squares line is fitted to $\boldsymbol{g}_i$ using timestamps normalized to $[0, 1]$. Its slope forms the numeric trend descriptor. For qualitative inspection, $\boldsymbol{g}_i$ is also divided into at most five equal-length segments whose successive mean changes are described as increasing, decreasing, or stable. This symbolic summary is not included in the profile vector.

Volatility. To avoid the zero-mean degeneracy induced by temporal standardization, the volatility block is computed from the activity sequence $\boldsymbol{u}_i$. It contains the mean, standard deviation, range, interquartile range, and coefficient of variation

$$\mathrm{CV}(\boldsymbol{u}_i) = \frac{\mathrm{std}(\boldsymbol{u}_i)}{\mathrm{mean}(\boldsymbol{u}_i) + \epsilon}. \tag{13}$$

Periodicity. After removing its temporal mean, let $P_{i,h}$ denote the Fourier power of $\boldsymbol{g}_i$ at frequency index h. The dominant-power ratio is

$$r_i^{\mathrm{fft}} = \frac{\max_{h\geq 1} P_{i,h}}{\sum_{h\geq 1} P_{i,h} + \epsilon}. \tag{14}$$

PRISM additionally computes the absolute autocorrelation over $1 \le \ell \le \lfloor T/2 \rfloor$. Its maximum value and corresponding lag form the remaining periodicity descriptors.

Events. For temporal differences $\Delta g_{i,t} = g_{i,t+1} - g_{i,t}$, define the robust change score

$$r^{\Delta}_{i,t} = \frac{0.6745\,(\Delta g_{i,t} - \mathrm{median}(\Delta \boldsymbol{g}_i))}{\mathrm{MAD}(\Delta \boldsymbol{g}_i) + \epsilon}. \tag{15}$$

Event density is the proportion of timestamps satisfying $|r^{\Delta}_{i,t}| > 3$. A large change is separately detected when

$$|\Delta g_{i,t}| \ge Q_{0.95}\left(\{|\Delta g_{i,s}|\}_{s=1}^{T-1}\right), \tag{16}$$

where $Q_{0.95}$ denotes the 95th percentile. A turning point occurs when two consecutive temporal differences have opposite signs. Spikes are detected using

$$z^{\mathrm{rob}}_{i,t} = \frac{0.6745\,(g_{i,t} - \mathrm{median}(\boldsymbol{g}_i))}{\mathrm{MAD}(\boldsymbol{g}_i) + \epsilon}, \tag{17}$$

with $|z^{\mathrm{rob}}_{i,t}| > 3$. The event block contains event density, large-change count, turning-point count, and spike density.

B.3 Patch Descriptors

The profile divides $\widetilde{\boldsymbol{X}}_i$ into overlapping windows of length $w_{\mathrm{p}} = 16$ and stride $s_{\mathrm{p}} = 8$. The number of candidate windows is

$$K^{\mathrm{all}}_i = 1 + \left\lfloor \frac{T - \min(T, w_{\mathrm{p}})}{s_{\mathrm{p}}} \right\rfloor, \qquad K_i = \min\{K^{\mathrm{all}}_i, K_{\max}\}, \quad K_{\max} = 128. \tag{18}$$

When $K^{\mathrm{all}}_i > K_{\max}$, windows are selected at uniformly spaced indices to preserve full-sequence coverage. For patch $\boldsymbol{B}_{i,\ell} \in \mathbb{R}^{w_i \times C}$, where $w_i = \min(T, w_{\mathrm{p}})$, its energy, volatility, and temporal variation are

$$\begin{aligned} E_{i,\ell} &= \frac{\|\boldsymbol{B}_{i,\ell}\|^2_{\mathrm{F}}}{w_i C}, \qquad V_{i,\ell} = \frac{1}{C}\sum_{c=1}^{C} \mathrm{std}(\boldsymbol{B}_{i,\ell,:,c}), \\ D_{i,\ell} &= \frac{\sum_{t=1}^{w_i-1}\sum_{c=1}^{C} |B_{i,\ell,t+1,c} - B_{i,\ell,t,c}|}{(w_i - 1)C}. \end{aligned} \tag{19}$$

When $w_i = 1$, $D_{i,\ell}$ is set to zero. The patch block contains the patch count, mean and maximum energy, mean volatility, mean temporal variation, and the index, start, and end of the maximum-energy patch. These deterministic summaries are distinct from the raw patch tokens processed by the temporal encoder.

B.4 Variable and Dependency Descriptors

For variable c, PRISM computes variance $v_{i,c}$, mean absolute temporal variation $e_{i,c}$, absolute trend $t_{i,c}$, and periodicity strength $f_{i,c}$. Each quantity is normalized across variables:

$$\widehat{u}_{i,c} = \frac{u_{i,c}}{\sum_{c'=1}^{C} u_{i,c'} + \epsilon}, \qquad u \in \{v, e, t, f\}. \tag{20}$$

The activity score is

$$a_{i,c} = 0.50\widehat{v}_{i,c} + 0.20\widehat{e}_{i,c} + 0.15\widehat{t}_{i,c} + 0.15\widehat{f}_{i,c}. \tag{21}$$

Variables are ranked by $a_{i,c}$. PRISM retains at most $C_{\mathrm{a}} = 5$ variables, stopping earlier when their cumulative activity reaches 0.85. Descriptor slots are aligned by activity rank rather than channel identity. Each slot records the activity ratio, trend slope, volatility standard deviation and range, event and spike densities, and periodicity strength. Unused slots are zero-padded.

Let $\mathcal{A}^{\mathrm{v}}_i$ denote the retained variables and $\mathcal{V}^{\mathrm{pair}}_i$ the retained variable pairs whose correlations are finite and well-defined. Cross-variable dependency is summarized by

$$d^{\mathrm{dep}}_i = \frac{1}{|\mathcal{V}^{\mathrm{pair}}_i|} \sum_{(c,c') \in \mathcal{V}^{\mathrm{pair}}_i} \left|\mathrm{corr}(\boldsymbol{X}_{i,:,c}, \boldsymbol{X}_{i,:,c'})\right|, \tag{22}$$

Table 6: Composition of the deterministic temporal profile. Dimensions follow from the listed descriptors and are not tunable hyperparameters.

Block	Dimension	Descriptors
Quality	3	Valid, invalid, and constant ratios
Global trend	1	Linear slope
Global volatility	5	Mean, std., range, IQR, and CV
Global periodicity	3	Power ratio, autocorrelation, and lag
Global events	4	Density, changes, turns, and spikes
Patch	8	Count, energy, volatility, variation, and location
Dependency	2	Correlation and valid-pair count
Variable	$7C_{\mathrm{a}}$	Seven descriptors per activity-ranked slot
Total	$26 + 7C_{\mathrm{a}}$	61 when $C_{\mathrm{a}} = 5$

with $d_i^{\mathrm{dep}} = 0$ when $|\mathcal{V}_i^{\mathrm{pair}}| = 0$. The dependency block contains d_i^{dep} and the valid-pair count $|\mathcal{V}_i^{\mathrm{pair}}|$.

Table 6 summarizes the resulting profile. After concatenation, each coordinate is standardized across $\mathcal{X}$ and each profile is ℓ_2-normalized according to Eq. (3). This normalization prevents descriptor scale from dominating the cosine relations used to construct the fixed anchors.

C Anchor Construction Details

Spherical k-means is applied to the unit-normalized profiles using k-means++ initialization and multiple restarts. Empty prototypes are reinitialized with the profile having the largest current quantization error, and the solution with the lowest objective is retained. The resulting centers are normalized to define $\mathcal{A} = \{\boldsymbol{a}_r\}_{r=1}^m$, after which Eq. (5) constructs the fixed prototype anchor $\boldsymbol{Q}^{\mathrm{p}}$. Its auxiliary hard assignment is

$$u_i = \arg\max_{1 \le r \le m} q_{i,r}^{\mathrm{p}}. \tag{23}$$

To preserve the separation between profile and embedding coordinates, profile centers are not copied into the embedding space. Instead, the encoder is first applied to obtain initial embeddings $\boldsymbol{z}_i^{(0)}$, and each embedding prototype is initialized from the normalized mean of embeddings sharing the same auxiliary index:

$$\boldsymbol{b}_r^{\mathrm{z},(0)} = \frac{\sum_{i:u_i=r} \boldsymbol{z}_i^{(0)}}{\left\| \sum_{i:u_i=r} \boldsymbol{z}_i^{(0)} \right\|_2}, \qquad r = 1, \ldots, m. \tag{24}$$

Empty embedding groups are assigned random unit-norm vectors. The auxiliary assignments are used only for initialization and never as final cluster labels.

For the neighborhood anchor, self-relations are excluded before the κ largest profile similarities are selected. Ties at the κ-th position are resolved by sample index, making the support deterministic. Equation (6) normalizes the retained similarities, and only their indices and weights are stored. The persistent neighborhood anchor therefore requires $O(n\kappa)$ rather than $O(n^2)$ space. Both $\boldsymbol{Q}^{\mathrm{p}}$ and $\boldsymbol{\rho}^{\mathrm{p}}$ are constructed without labels or encoder outputs and remain unchanged throughout representation learning.

D Encoder and Training Details

D.1 Temporal Patch Encoder

Each raw MTS is divided into synchronized patches using length $w = 16$, stride $s = 8$, and a maximum of $L_{\max} = 128$ patches. Complete windows are extracted when $T \ge w$. If their number exceeds $L_{\max}$, uniformly spaced windows are retained to preserve full-sequence coverage. Sequences shorter than w are zero-padded to form one patch, and padded values are excluded by the corresponding mask.

Each patch is represented through three complementary views comprising raw values, first-order temporal differences, and Fourier magnitudes. Separate linear projections map the views into a shared hidden space, where learned softmax weights combine them. A channel-relation token derived from cross-variable correlations supplies channel context, while positional embeddings preserve patch order.

The resulting token sequence is processed by a depthwise local mixer followed by four pre-normalized Transformer layers. Each layer uses four attention heads, hidden dimension 128, feed-forward dimension 512, GELU activation, and dropout 0.1. Masked mean pooling and attentive pooling summarize the patch tokens. Their outputs are combined and passed through a two-layer projection head to produce a $d_{\mathrm{z}} = 128$ representation. Sample-wise ℓ_2 normalization then gives the embedding in Eq. (7).

D.2 Anchor-Guided Optimization

The encoder and embedding prototype bank are optimized jointly with AdamW using learning rate 10^{-3} and weight decay 10^{-4}. Unless otherwise stated, the temperatures and structural weight are

$$\tau_{\mathrm{a}} = \tau_{\mathrm{z}} = 0.2, \qquad \tau_{\mathrm{s}} = 0.1, \qquad \lambda_{\mathrm{struct}} = 0.2. \tag{25}$$

The default prototype number is determined from the dataset size:

$$m = \min\left\{128,\ \max\left(2, \mathrm{round}(\sqrt{n})\right)\right\}. \tag{26}$$

The default seed-batch size is $B = 64$. The neighborhood size κ, epoch count, and any memory-driven adjustments to patch length or batch size are specified in the released dataset configurations.

The embedding prototype bank is initialized using Eq. (24). For each seed batch $\mathcal{I}_B$, PRISM constructs the neighbor-augmented set $\mathcal{V}_B$ and retains gradients for both seed and neighbor embeddings. Prototype compatibility is evaluated for the seed samples, while neighborhood compatibility is evaluated between each seed and its fixed profile neighbors. The losses in Eq. (10) update only the encoder parameters $\boldsymbol{\theta}$ and embedding prototypes $\mathcal{B}^{\mathrm{z}}$. The profile anchors $\boldsymbol{Q}^{\mathrm{p}}$ and $\boldsymbol{\rho}^{\mathrm{p}}$ are detached from optimization and remain unchanged throughout training.

After training, $\mathcal{B}^{\mathrm{z}}$ is discarded and the encoder produces the final embedding matrix $\boldsymbol{Z}$. Final k-means uses the dataset-specified cluster number k, ten initializations, and at most 300 iterations. All reported results are averaged over seeds $\{2022, 2023, 2024, 2025, 2026\}$. Dataset-specific settings are provided in the released configuration files.

E Theoretical Analysis

This section formalizes three properties of fixed relational guidance. The results characterize target independence, non-collapse conditions, and anchor sensitivity rather than recovery of unknown ground-truth clusters.

E.1 Target Independence

Let $\boldsymbol{A}$ denote a relational target and consider an objective $\ell(\boldsymbol{\theta}, \boldsymbol{A})$. When a differentiable target is generated by the current model, $\boldsymbol{A} = \boldsymbol{A}(\boldsymbol{\theta})$, its total derivative contains both the direct optimization term and a target-feedback term:

$$\frac{\mathrm{d}\ell}{\mathrm{d}\boldsymbol{\theta}} = \frac{\partial \ell}{\partial \boldsymbol{\theta}} + \frac{\partial \ell}{\partial \boldsymbol{A}} \frac{\partial \boldsymbol{A}}{\partial \boldsymbol{\theta}}. \tag{27}$$

Stop-gradient implementations remove the second term within one update, but their targets may still change between iterations because they are recomputed from evolving model predictions.

Proposition 1. For PRISM, let

$$\boldsymbol{A} = (\boldsymbol{Q}^{\mathrm{p}}, \boldsymbol{\rho}^{\mathrm{p}}). \tag{28}$$

Then $\partial \boldsymbol{A} / \partial \boldsymbol{\theta} = 0$ and $\boldsymbol{A}^{(e)} = \boldsymbol{A}$ for every training epoch e.

Proof. Both anchors are deterministic functions of the fixed profile matrix $\boldsymbol{P}$ and are constructed before encoder optimization. Neither updates to $\boldsymbol{\theta}$ nor updates to the embedding prototype bank $\mathcal{B}^{\mathrm{z}}$

alter their values. Hence, $\boldsymbol{A}$ has no gradient path through the encoder and remains unchanged across epochs. □

Proposition 1 distinguishes target independence from assignment stability. The profile-side supervision has zero target drift, while the embedding-side relations and cluster assignments may continue to evolve as the encoder improves.

E.2 Conditions Excluding Global Collapse

Because the profile-side distributions are fixed, the two cross-entropy terms can be written as

$$\begin{aligned}\mathcal{L}_{\text{proto}} &= \frac{1}{n}\sum_{i=1}^{n} \mathrm{KL}\left(\boldsymbol{q}^{\mathrm{p}}_{i,:} \| \boldsymbol{\pi}^{\mathrm{z}}_{i,:}\right) + \text{const.},\\ \mathcal{L}_{\text{struct}} &= \frac{1}{n}\sum_{i=1}^{n} \mathrm{KL}\left(\boldsymbol{\rho}^{\mathrm{p}}_{i,:} \| \boldsymbol{\rho}^{\mathrm{z}}_{i,:}\right) + \text{const.},\end{aligned} \tag{29}$$

where the constant terms are the entropies of the fixed targets and do not depend on the learned parameters.

Proposition 2. Assume that the encoder and embedding prototype bank can jointly realize the fixed anchor distributions. A globally collapsed representation $\boldsymbol{z}_1 = \cdots = \boldsymbol{z}_n$ is not a global minimizer of Eq. (10) if either

1. two samples have distinct prototype targets, or
2. $\lambda_{\text{struct}} > 0$ and at least one neighborhood target is nonuniform.

Proof. Under global collapse, all samples have identical similarities to the embedding prototype bank and therefore share the same prototype distribution. This distribution cannot match two distinct rows of $\boldsymbol{Q}^{\mathrm{p}}$. Moreover, all pairwise embedding similarities are equal, so $\boldsymbol{\rho}^{\mathrm{z}}_{i,:}$ is uniform over $\mathcal{N}^{\mathrm{p}}_i$. It cannot match a nonuniform row of $\boldsymbol{\rho}^{\mathrm{p}}$. Under either condition, at least one positively weighted KL divergence in Eq. (29) is strictly positive, whereas joint realizability permits all such divergences to reach zero. □

The prototype anchor excludes collapse when coarse relations differ across samples. The neighborhood anchor provides an additional constraint whenever coarse prototype compatibility alone cannot distinguish locally related samples.

E.3 Stability under Profile Perturbations

Let $\mathrm{softmax}_\tau(\boldsymbol{v})$ denote softmax with temperature τ. Its Jacobian gives

$$\left\|\mathrm{softmax}_\tau(\boldsymbol{v}) - \mathrm{softmax}_\tau(\boldsymbol{v}')\right\|_1 \le \frac{2}{\tau}\left\|\boldsymbol{v} - \boldsymbol{v}'\right\|_\infty. \tag{30}$$

Proposition 3. Suppose the original and perturbed profiles are unit-normalized and satisfy

$$\left\|\boldsymbol{p}'_i - \boldsymbol{p}_i\right\|_2 \le \varepsilon. \tag{31}$$

After aligning prototype indices, suppose the corresponding unit-norm prototypes satisfy

$$\left\|\boldsymbol{a}'_r - \boldsymbol{a}_r\right\|_2 \le \varepsilon_{\mathrm{a}}. \tag{32}$$

If the prototype correspondence and neighborhood supports remain unchanged, then

$$\left\|\boldsymbol{q}^{\mathrm{p}'}_{i,:} - \boldsymbol{q}^{\mathrm{p}}_{i,:}\right\|_1 \le \frac{2(\varepsilon+\varepsilon_{\mathrm{a}})}{\tau_{\mathrm{a}}}, \qquad \left\|\boldsymbol{\rho}^{\mathrm{p}'}_{i,:} - \boldsymbol{\rho}^{\mathrm{p}}_{i,:}\right\|_1 \le \frac{4\varepsilon}{\tau_{\mathrm{s}}}. \tag{33}$$

Proof. For unit-norm vectors,

$$\left|\mathrm{sim}(\boldsymbol{p}'_i, \boldsymbol{a}'_r) - \mathrm{sim}(\boldsymbol{p}_i, \boldsymbol{a}_r)\right| \le \varepsilon + \varepsilon_{\mathrm{a}}. \tag{34}$$

Similarly, perturbing both samples changes their cosine similarity by at most 2ε. Applying Eq. (30) to the prototype and neighborhood logits yields the two bounds in Eq. (33). □

Proposition 3 bounds changes in anchor weights under small profile perturbations. Its stable-support condition is essential because perturbations that change prototype correspondence or nearest-neighbor membership can produce discrete structural changes. The result is an anchor-sensitivity guarantee and does not assume recovery of the ground-truth partition.

F Complexity Details

Profile preprocessing. Computing the temporal and spectral descriptors requires $O(nCT \log T)$ time. Dependency estimation among at most C_{a} retained variables adds $O(nTC_{\mathrm{a}}^2)$, giving a total profiling cost of

$$O\big(nCT \log T + nTC_{\mathrm{a}}^2\big)\,. \tag{35}$$

The resulting profile matrix occupies $O(nd_{\mathrm{p}})$ space.

Anchor construction. Spherical prototype estimation over I_{p} iterations costs $O(I_{\mathrm{p}}nmd_{\mathrm{p}})$ time. Computing the prototype anchor $\boldsymbol{Q}^{\mathrm{p}}$ requires an additional $O(nmd_{\mathrm{p}})$ time and $O(nm)$ storage. Exact neighborhood search requires $O(n^2 d_{\mathrm{p}})$ time. Naively materializing the complete similarity matrix uses $O(n^2)$ temporary memory, while blockwise exact search reduces this requirement without changing the quadratic time complexity. Only the selected neighbor indices and weights are retained, so the persistent neighborhood anchor occupies $O(n\kappa)$ space.

Anchor-guided training. For a seed batch of size B, let $\mathcal{C}_{\mathrm{enc}}(|\mathcal{V}_B|, L, w, C)$ denote the cost of encoding its neighbor-augmented sample set. Computing embedding-side prototype compatibility costs $O(Bmd_{\mathrm{z}})$, while evaluating affinities over the fixed neighborhood support costs $O(B\kappa d_{\mathrm{z}})$. Training for E epochs therefore requires

$$O\Big(E\left\lceil \frac{n}{B} \right\rceil \left[\mathcal{C}_{\mathrm{enc}}(|\mathcal{V}_B|, L, w, C) + B(m+\kappa)d_{\mathrm{z}}\right]\Big)\,. \tag{36}$$

No dense sample-to-sample matrix is constructed during optimization because neighborhood matching is restricted to the fixed sparse support.

Storage and final clustering. Excluding the raw data and encoder parameters, the persistent profiles, anchors, and embedding prototype bank require

$$O(nd_{\mathrm{p}} + nm + n\kappa + md_{\mathrm{z}}) \tag{37}$$

space. Storing all final embeddings adds $O(nd_{\mathrm{z}})$. Final k-means over I_{k} iterations costs $O(I_{\mathrm{k}}nkd_{\mathrm{z}})$ time. For fixed m and κ, the relational alignment cost grows linearly with n. Exact neighborhood search is the only quadratic stage and can be replaced by approximate nearest-neighbor search for larger datasets.

G Experimental Protocol

G.1 Datasets and Evaluation

The evaluation uses 30 datasets from the UEA MTS archive and two real-world human-activity datasets, MHEALTH and UCI-HAR. The UEA collection spans audio, electrocardiography, electroencephalography, gesture, human activity recognition, magnetoencephalography, motion, sensor, simulation, and traffic domains. Following the transductive clustering setting, the official training and test partitions of each UEA dataset are combined into one unlabeled repository. MHEALTH and UCI-HAR are used separately to evaluate the semantic traceability of learned groups.

Table 7 summarizes the dataset statistics, where n, T, C, and k^* denote the numbers of samples, timestamps, variables, and ground-truth clusters. UEA rows report averages within each domain, while the two real-world datasets are listed individually. Ground-truth labels are excluded from profiling, anchor construction, representation learning, and clustering. They are accessed only for post-hoc evaluation.

Clustering quality is evaluated using clustering accuracy (ACC), macro F1, normalized mutual information (NMI), and adjusted Rand index (ARI). ACC uses optimal one-to-one label matching, and macro F1 is computed under the resulting correspondence. NMI and ARI evaluate partition agreement independently of label identities. Results are reported as the mean and standard deviation over seeds $\{2022, 2023, 2024, 2025, 2026\}$. Per-dataset ranks are computed among all methods using average ranks for ties and are then macro-averaged without weighting datasets by sample size.

G.2 Baselines and Fair Comparison

The nine baselines cover three complementary methodological families.

Table 7: Dataset statistics. UEA rows report group averages, MHEALTH and UCI-HAR are listed individually.

Type	Group/Dataset (#)	n	T	C	k^*
UEA benchmark groups	AUDIO (5)	3323.00	188.20	288.40	6.60
	HAR (4)	147.50	956.75	11.50	5.75
	ECG (1)	30.00	640.00	2.00	3.00
	MEG (2)	4940.50	231.00	77.00	3.00
	EEG (5)	402.00	1060.80	21.60	2.40
	MOTION (5)	3136.80	3694.00	4.60	17.20
	GESTURE (4)	524.00	119.00	8.25	4.00
	SENSOR (2)	25262.00	914.50	101.50	7.00
	SIMUL (1)	4925.00	36.00	36.00	14.00
	TRAFFIC (1)	440.00	144.00	963.00	7.00
Real	MHEALTH (1)	1800.00	128.00	21.00	12.00
	UCI-HAR (1)	10299.00	128.00	9.00	6.00

General-purpose representation learning. TimesURL (Liu & Chen, 2024), UNITS (Gao et al., 2024), FEI (Fu & Hu, 2025), and GTM (He et al., 2026) learn transferable time-series representations without being designed exclusively for clustering. They test whether general self-supervised objectives alone produce suitable grouping geometry.

Deep MTS clustering. MVCIMTS (Li et al., 2025), TFMCC (Wang et al., 2026a), FCACC (Wang et al., 2026b), and MUFASA (Li & Paparrizos, 2026) incorporate clustering-oriented objectives, multiview consistency, frequency information, or adaptive grouping mechanisms. They provide the closest comparison with PRISM's clustering-specific representation learning.

Interpretable structural clustering. k-Graph (Boniol et al., 2025) constructs explicit graph-based time-series relations and therefore serves as a data-centric, structurally interpretable baseline.

Official implementations and recommended configurations are used whenever available. All methods receive the same preprocessed unlabeled repositories and dataset-specified cluster number k. Methods that output representations but not assignments are followed by the same k-means protocol used for PRISM. Native clustering outputs are retained for clustering-specific methods. No labels are used during training or representation construction.

G.3 Computing Environment

PRISM is implemented in Python 3.12 with PyTorch 2.8.0 and CUDA 12.8 on Ubuntu 22.04. Experiments are conducted on one NVIDIA RTX 5090 GPU with 32 GB memory and an Intel Xeon Platinum 8470Q processor using 25 virtual CPU cores. Runtime comparisons use the same hardware environment and timing scope for all methods.

H Additional Diagnostics

H.1 Target Drift

Target drift measures whether the relational supervision itself changes during optimization. For a row-normalized target $\boldsymbol{A}^{(e)}$ at epoch e, we define

$$D_{\text{target}}^{(e)} = \frac{1}{n}\sum_{i=1}^{n} \text{JS}\left(\boldsymbol{A}_{i,:}^{(e-1)}, \boldsymbol{A}_{i,:}^{(e)}\right), \tag{38}$$

where JS denotes Jensen–Shannon divergence. Prototype and neighborhood targets are evaluated separately on their corresponding indexed supports. If an adaptive method changes a neighborhood support, the two distributions are embedded into their union support and assigned zero weight outside their original supports.

For PRISM, $\boldsymbol{A} \in \{\boldsymbol{Q}^{\mathrm{p}}, \boldsymbol{\rho}^{\mathrm{p}}\}$ is constructed once before representation learning. Consequently,

$$D_{\mathrm{target}}^{(e)} = 0, \qquad e = 1, \ldots, E. \tag{39}$$

This property concerns the supervision rather than the learned assignments. Embedding-side relations and cluster assignments may evolve throughout training, but the semantic reference against which they are optimized remains unchanged. Assignment dynamics are analyzed separately in Table 4 and Figure 5.

H.2 Semantic Alignment

We quantify whether the learned geometry preserves the temporal relations encoded by the profiles. Let $\mathcal{S}$ denote a fixed set of sample pairs. Semantic alignment is measured by the Spearman rank correlation between profile and embedding similarities:

$$R_{\mathrm{sem}} = \mathrm{Spearman}\left(\{s_{i,j}^{\mathrm{p}}\}_{(i,j)\in\mathcal{S}}, \{s_{i,j}^{\mathrm{z}}\}_{(i,j)\in\mathcal{S}}\right). \tag{40}$$

Higher R_{sem} indicates stronger preservation of explicit temporal relations without requiring correspondence between profile and embedding coordinates. All pairs are used when feasible. For larger datasets, $\mathcal{S}$ is a uniformly sampled pair set fixed across methods and random seeds.

H.3 Descriptor-Block Attribution

Because profiles are ℓ_2-normalized, their cosine similarity equals their coordinate-wise inner product. The signed contribution of descriptor block b to the profile relation between samples i and j is therefore

$$\Gamma_{i,j}^{(b)} = \sum_{h\in\mathcal{D}_b} p_{i,h} p_{j,h}, \qquad s_{i,j}^{\mathrm{p}} = \sum_{b} \Gamma_{i,j}^{(b)}, \tag{41}$$

where $\mathcal{D}_b$ contains the normalized profile coordinates assigned to block b.

For a learned group $\mathcal{G}_c = \{i : \hat{y}_i = c\}$, its nonnegative block composition is obtained by normalizing the absolute within-group contributions:

$$\omega_c^{(b)} = \frac{\displaystyle\sum_{\substack{i<j \\ i,j\in\mathcal{G}_c}} \left|\Gamma_{i,j}^{(b)}\right|}{\displaystyle\sum_{b'} \sum_{\substack{i<j \\ i,j\in\mathcal{G}_c}} \left|\Gamma_{i,j}^{(b')}\right| + \epsilon}. \tag{42}$$

Thus, $\sum_b \omega_c^{(b)} = 1$, and $\omega_c^{(b)}$ measures the relative strength of temporal property b within the learned group. This decomposition produces the temporal-property compositions in Figure 3. It uses only profiles and learned assignments. Ground-truth activity labels are introduced afterward solely to name the displayed groups and compute purity.

I Additional Experimental Results

I.1 Complete Clustering Results

Tables 8 and 9 report the mean and standard deviation over five seeds on all 30 UEA datasets. Among the 120 dataset-metric comparisons, PRISM achieves the highest mean in 39 cases and the second-highest in another 23. It ranks first under all four metrics on DuckDuckGeese, JapaneseVowels, PhonemeSpectra, Epilepsy, BasicMotions, Cricket, ArticularyWordRecognition, and EigenWorms. These datasets span audio, EEG, human activity, motion, and sensor domains, showing that the aggregate advantage is not driven by a single domain or a few isolated datasets.

PRISM is less competitive on AtrialFibrillation, FingerMovements, MotorImagery, and HandMovementDirection. These cases primarily involve subtle group differences that are not fully expressed by the selected trend, volatility, periodicity, event, patch, or dependency descriptors. They therefore expose a boundary of fixed profile guidance: its benefit depends on whether the deterministic profile captures group-relevant temporal variation. Statistical comparisons across all datasets are reported in Appendix I.2.

Table 8: Top-performing clustering results on all thirty datasets evaluated by ACC (↑), F1 (↑), NMI (↑), and ARI (↑). Best and Second Best results are highlighted by cell background.

Group	Dataset	Metric	TimesURL (AAAI'24)	UNITS (NeurIPS'24)	FEI (AAAI'25)	MVCIMTS (INFFUS'25)	k-Graph (TKDE'25)	TFMCC (AAAI'26)	FCACC (PR'26)	GTM (ICLR'26)	MUFASA (SIGMOD'26)	PRISM (ours)
AUDIO	DuckDuckGeese	ACC	0.322 ± 0.015	0.296 ± 0.014	0.322 ± 0.012	0.262 ± 0.052	0.328 ± 0.016	0.472 ± 0.017	0.307 ± 0.013	0.362 ± 0.032	0.462 ± 0.095	0.514 ± 0.039
		F1	0.264 ± 0.018	0.287 ± 0.007	0.301 ± 0.035	0.165 ± 0.058	0.302 ± 0.011	0.339 ± 0.019	0.261 ± 0.020	0.355 ± 0.034	0.449 ± 0.092	0.506 ± 0.040
		NMI	0.173 ± 0.035	0.068 ± 0.018	0.103 ± 0.009	0.034 ± 0.030	0.096 ± 0.010	0.279 ± 0.027	0.080 ± 0.018	0.154 ± 0.048	0.274 ± 0.115	0.308 ± 0.058
		ARI	0.037 ± 0.011	0.006 ± 0.011	0.037 ± 0.009	0.007 ± 0.017	0.022 ± 0.009	0.178 ± 0.025	0.023 ± 0.011	0.066 ± 0.037	0.190 ± 0.107	0.202 ± 0.046
	Heartbeat	ACC	0.715 ± 0.011	0.624 ± 0.024	0.709 ± 0.000	0.676 ± 0.050	0.730 ± 0.003	0.708 ± 0.048	0.584 ± 0.008	0.534 ± 0.029	0.569 ± 0.042	0.619 ± 0.012
		F1	0.792 ± 0.076	0.535 ± 0.019	0.830 ± 0.000	0.459 ± 0.048	0.546 ± 0.007	0.693 ± 0.067	0.569 ± 0.005	0.497 ± 0.019	0.512 ± 0.034	0.597 ± 0.009
		NMI	0.030 ± 0.021	0.007 ± 0.008	0.020 ± 0.000	0.002 ± 0.002	0.034 ± 0.004	0.046 ± 0.007	0.039 ± 0.002	0.001 ± 0.002	0.006 ± 0.007	0.053 ± 0.005
		ARI	0.022 ± 0.071	0.024 ± 0.011	−0.014 ± 0.000	0.002 ± 0.010	0.087 ± 0.007	0.097 ± 0.031	0.026 ± 0.006	0.003 ± 0.006	0.008 ± 0.014	0.054 ± 0.011
	JapaneseVowels	ACC	0.184 ± 0.000	0.262 ± 0.027	0.530 ± 0.018	0.184 ± 0.000	0.337 ± 0.020	0.477 ± 0.064	0.629 ± 0.053	0.184 ± 0.000	0.186 ± 0.000	0.786 ± 0.056
		F1	0.311 ± 0.000	0.276 ± 0.028	0.546 ± 0.026	0.035 ± 0.000	0.307 ± 0.024	0.427 ± 0.035	0.662 ± 0.030	0.035 ± 0.000	0.038 ± 0.000	0.786 ± 0.059
		NMI	0.000 ± 0.000	0.164 ± 0.021	0.490 ± 0.014	0.000 ± 0.000	0.282 ± 0.011	0.525 ± 0.040	0.595 ± 0.013	0.000 ± 0.000	0.003 ± 0.000	0.753 ± 0.037
		ARI	0.000 ± 0.000	0.065 ± 0.017	0.326 ± 0.028	0.000 ± 0.000	0.151 ± 0.021	0.317 ± 0.047	0.477 ± 0.039	0.000 ± 0.000	0.000 ± 0.000	0.664 ± 0.055
	PhonemeSpectra	ACC	0.087 ± 0.002	0.081 ± 0.007	0.128 ± 0.002	0.045 ± 0.006	0.092 ± 0.003	0.112 ± 0.004	0.115 ± 0.000	0.063 ± 0.001	0.100 ± 0.004	0.146 ± 0.006
		F1	0.085 ± 0.002	0.079 ± 0.008	0.124 ± 0.004	0.010 ± 0.003	0.091 ± 0.003	0.105 ± 0.007	0.116 ± 0.000	0.060 ± 0.001	0.096 ± 0.004	0.145 ± 0.006
		NMI	0.117 ± 0.007	0.093 ± 0.019	0.208 ± 0.004	0.037 ± 0.014	0.109 ± 0.004	0.204 ± 0.003	0.179 ± 0.002	0.045 ± 0.002	0.135 ± 0.005	0.237 ± 0.004
		ARI	0.019 ± 0.002	0.014 ± 0.004	0.045 ± 0.001	0.005 ± 0.003	0.018 ± 0.002	0.042 ± 0.005	0.034 ± 0.001	0.004 ± 0.000	0.026 ± 0.001	0.053 ± 0.002
	SpokenArabic Digits	ACC	0.100 ± 0.000	0.281 ± 0.031	0.461 ± 0.070	0.189 ± 0.045	0.369 ± 0.005	0.421 ± 0.031	0.490 ± 0.026	0.100 ± 0.000	0.100 ± 0.000	0.494 ± 0.071
		F1	0.182 ± 0.000	0.282 ± 0.029	0.453 ± 0.070	0.120 ± 0.052	0.371 ± 0.008	0.420 ± 0.030	0.494 ± 0.029	0.018 ± 0.000	0.018 ± 0.000	0.492 ± 0.072
		NMI	0.000 ± 0.000	0.181 ± 0.032	0.381 ± 0.057	0.103 ± 0.053	0.269 ± 0.014	0.471 ± 0.052	0.474 ± 0.013	0.000 ± 0.000	0.000 ± 0.000	0.437 ± 0.056
		ARI	0.000 ± 0.000	0.095 ± 0.023	0.265 ± 0.061	0.053 ± 0.028	0.172 ± 0.007	0.293 ± 0.051	0.337 ± 0.019	0.000 ± 0.000	0.000 ± 0.000	0.311 ± 0.072
ECG	Atrial Fibrillation	ACC	0.460 ± 0.049	0.450 ± 0.017	0.387 ± 0.016	0.409 ± 0.044	0.440 ± 0.039	0.433 ± 0.056	0.433 ± 0.000	0.467 ± 0.042	0.453 ± 0.065	0.420 ± 0.016
		F1	0.417 ± 0.028	0.434 ± 0.009	0.354 ± 0.019	0.349 ± 0.056	0.429 ± 0.041	0.404 ± 0.019	0.430 ± 0.003	0.459 ± 0.039	0.450 ± 0.064	0.417 ± 0.017
		NMI	0.086 ± 0.034	0.064 ± 0.007	0.035 ± 0.024	0.144 ± 0.082	0.114 ± 0.075	0.054 ± 0.035	0.101 ± 0.004	0.073 ± 0.028	0.058 ± 0.039	0.028 ± 0.008
		ARI	0.014 ± 0.034	−0.013 ± 0.001	−0.036 ± 0.009	0.106 ± 0.060	0.004 ± 0.046	0.006 ± 0.028	−0.006 ± 0.001	0.005 ± 0.026	−0.004 ± 0.047	−0.042 ± 0.010
EEG	Epilepsy	ACC	0.516 ± 0.019	0.498 ± 0.065	0.636 ± 0.081	0.295 ± 0.032	0.617 ± 0.032	0.495 ± 0.048	0.428 ± 0.030	0.387 ± 0.040	0.436 ± 0.034	0.921 ± 0.037
		F1	0.498 ± 0.008	0.509 ± 0.093	0.620 ± 0.082	0.188 ± 0.024	0.620 ± 0.026	0.415 ± 0.055	0.429 ± 0.030	0.371 ± 0.043	0.431 ± 0.036	0.919 ± 0.039
		NMI	0.336 ± 0.080	0.313 ± 0.085	0.408 ± 0.077	0.021 ± 0.034	0.401 ± 0.042	0.303 ± 0.079	0.140 ± 0.035	0.102 ± 0.035	0.183 ± 0.049	0.822 ± 0.069
		ARI	0.229 ± 0.054	0.232 ± 0.073	0.383 ± 0.080	0.008 ± 0.018	0.328 ± 0.053	0.199 ± 0.064	0.096 ± 0.021	0.081 ± 0.043	0.132 ± 0.032	0.815 ± 0.076
	Finger Movements	ACC	0.519 ± 0.008	0.513 ± 0.005	0.516 ± 0.013	0.526 ± 0.004	0.526 ± 0.003	0.507 ± 0.005	0.525 ± 0.005	0.518 ± 0.009	0.510 ± 0.000	0.522 ± 0.005
		F1	0.516 ± 0.006	0.512 ± 0.005	0.505 ± 0.018	0.523 ± 0.002	0.500 ± 0.002	0.600 ± 0.060	0.495 ± 0.011	0.498 ± 0.021	0.464 ± 0.002	0.519 ± 0.006
		NMI	0.001 ± 0.001	0.001 ± 0.000	0.001 ± 0.002	0.002 ± 0.001	0.003 ± 0.001	0.002 ± 0.004	0.003 ± 0.001	0.001 ± 0.001	0.001 ± 0.000	0.002 ± 0.001
		ARI	−0.001 ± 0.001	−0.002 ± 0.000	−0.001 ± 0.002	0.000 ± 0.001	0.001 ± 0.001	−0.001 ± 0.001	0.001 ± 0.001	0.000 ± 0.001	−0.001 ± 0.000	0.000 ± 0.001
	MotorImagery	ACC	0.513 ± 0.005	0.533 ± 0.017	0.506 ± 0.004	0.511 ± 0.006	0.516 ± 0.011	0.505 ± 0.007	0.502 ± 0.001	0.536 ± 0.015	0.529 ± 0.000	0.520 ± 0.003
		F1	0.440 ± 0.034	0.518 ± 0.027	0.379 ± 0.021	0.457 ± 0.042	0.509 ± 0.011	0.659 ± 0.071	0.351 ± 0.002	0.508 ± 0.032	0.517 ± 0.000	0.519 ± 0.003
		NMI	0.001 ± 0.000	0.004 ± 0.004	0.001 ± 0.001	0.002 ± 0.002	0.001 ± 0.001	0.002 ± 0.004	0.000 ± 0.000	0.005 ± 0.003	0.003 ± 0.000	0.001 ± 0.000
		ARI	−0.001 ± 0.000	0.003 ± 0.006	0.000 ± 0.000	−0.001 ± 0.001	−0.001 ± 0.001	0.000 ± 0.000	0.000 ± 0.000	0.004 ± 0.004	0.001 ± 0.000	−0.001 ± 0.000
	SelfRegulation SCP1	ACC	0.734 ± 0.006	0.581 ± 0.040	0.610 ± 0.047	0.554 ± 0.034	0.559 ± 0.009	0.600 ± 0.087	0.536 ± 0.003	0.545 ± 0.031	0.617 ± 0.022	0.711 ± 0.027
		F1	0.718 ± 0.008	0.568 ± 0.037	0.608 ± 0.049	0.472 ± 0.096	0.558 ± 0.009	0.537 ± 0.132	0.508 ± 0.004	0.543 ± 0.032	0.598 ± 0.012	0.710 ± 0.028
		NMI	0.249 ± 0.004	0.033 ± 0.030	0.042 ± 0.029	0.026 ± 0.014	0.010 ± 0.003	0.072 ± 0.070	0.006 ± 0.001	0.009 ± 0.009	0.064 ± 0.042	0.138 ± 0.037
		ARI	0.218 ± 0.011	0.031 ± 0.030	0.055 ± 0.039	0.015 ± 0.012	0.012 ± 0.004	0.070 ± 0.104	0.004 ± 0.001	0.010 ± 0.013	0.055 ± 0.019	0.180 ± 0.049
	SelfRegulation SCP2	ACC	0.543 ± 0.005	0.581 ± 0.040	0.552 ± 0.010	0.535 ± 0.021	0.505 ± 0.007	0.503 ± 0.000	0.516 ± 0.005	0.526 ± 0.016	0.511 ± 0.012	0.521 ± 0.005
		F1	0.511 ± 0.011	0.568 ± 0.037	0.545 ± 0.008	0.530 ± 0.024	0.504 ± 0.007	0.707 ± 0.000	0.491 ± 0.006	0.524 ± 0.017	0.475 ± 0.038	0.520 ± 0.006
		NMI	0.008 ± 0.002	0.033 ± 0.030	0.009 ± 0.003	0.005 ± 0.005	0.000 ± 0.000	0.000 ± 0.000	0.001 ± 0.001	0.003 ± 0.003	0.003 ± 0.004	0.001 ± 0.001
		ARI	0.006 ± 0.002	0.031 ± 0.030	0.009 ± 0.004	0.004 ± 0.006	−0.002 ± 0.001	0.000 ± 0.000	−0.001 ± 0.001	0.001 ± 0.004	−0.001 ± 0.002	−0.001 ± 0.001
GESTURE	ERing	ACC	0.743 ± 0.116	0.409 ± 0.064	0.712 ± 0.092	0.367 ± 0.156	0.538 ± 0.042	0.861 ± 0.022	0.633 ± 0.036	0.418 ± 0.059	0.751 ± 0.025	0.848 ± 0.039
		F1	0.786 ± 0.084	0.416 ± 0.070	0.751 ± 0.081	0.268 ± 0.165	0.547 ± 0.040	0.756 ± 0.033	0.631 ± 0.039	0.412 ± 0.059	0.719 ± 0.013	0.848 ± 0.040
		NMI	0.741 ± 0.031	0.253 ± 0.080	0.656 ± 0.077	0.305 ± 0.211	0.384 ± 0.014	0.782 ± 0.032	0.566 ± 0.010	0.236 ± 0.048	0.751 ± 0.042	0.740 ± 0.046
		ARI	0.647 ± 0.076	0.158 ± 0.065	0.567 ± 0.096	0.190 ± 0.160	0.282 ± 0.032	0.705 ± 0.042	0.423 ± 0.020	0.160 ± 0.047	0.661 ± 0.035	0.691 ± 0.062
	Libras	ACC	0.270 ± 0.015	0.284 ± 0.012	0.432 ± 0.026	0.146 ± 0.042	0.384 ± 0.021	0.354 ± 0.010	0.382 ± 0.005	0.389 ± 0.032	0.244 ± 0.048	0.316 ± 0.033
		F1	0.277 ± 0.017	0.285 ± 0.016	0.444 ± 0.030	0.063 ± 0.030	0.370 ± 0.024	0.263 ± 0.023	0.393 ± 0.019	0.377 ± 0.037	0.159 ± 0.038	0.315 ± 0.033
		NMI	0.310 ± 0.018	0.356 ± 0.015	0.506 ± 0.019	0.131 ± 0.078	0.513 ± 0.006	0.470 ± 0.025	0.531 ± 0.009	0.470 ± 0.033	0.307 ± 0.073	0.387 ± 0.027
		ARI	0.090 ± 0.017	0.125 ± 0.015	0.242 ± 0.014	0.037 ± 0.025	0.253 ± 0.013	0.202 ± 0.023	0.248 ± 0.008	0.214 ± 0.028	0.110 ± 0.056	0.138 ± 0.022
	NATOPS	ACC	0.352 ± 0.081	0.297 ± 0.025	0.435 ± 0.027	0.327 ± 0.092	0.382 ± 0.019	0.567 ± 0.086	0.654 ± 0.098	0.486 ± 0.051	0.671 ± 0.045	0.657 ± 0.039
		F1	0.213 ± 0.069	0.295 ± 0.025	0.437 ± 0.060	0.205 ± 0.087	0.387 ± 0.017	0.517 ± 0.082	0.653 ± 0.097	0.464 ± 0.054	0.668 ± 0.056	0.661 ± 0.042
		NMI	0.411 ± 0.161	0.110 ± 0.032	0.390 ± 0.046	0.288 ± 0.167	0.190 ± 0.021	0.557 ± 0.072	0.645 ± 0.081	0.435 ± 0.055	0.663 ± 0.044	0.629 ± 0.033
		ARI	0.255 ± 0.126	0.055 ± 0.022	0.256 ± 0.046	0.160 ± 0.119	0.115 ± 0.021	0.418 ± 0.098	0.509 ± 0.116	0.306 ± 0.059	0.541 ± 0.045	0.489 ± 0.046
	UWaveGesture Library	ACC	0.456 ± 0.042	0.521 ± 0.015	0.366 ± 0.033	0.125 ± 0.000	0.330 ± 0.022	0.461 ± 0.099	0.545 ± 0.036	0.307 ± 0.029	0.846 ± 0.070	0.554 ± 0.028
		F1	0.454 ± 0.049	0.513 ± 0.017	0.363 ± 0.025	0.028 ± 0.000	0.332 ± 0.021	0.347 ± 0.063	0.537 ± 0.037	0.295 ± 0.028	0.841 ± 0.075	0.547 ± 0.024
		NMI	0.346 ± 0.026	0.002 ± 0.003	0.239 ± 0.011	0.000 ± 0.000	0.166 ± 0.028	0.389 ± 0.086	0.463 ± 0.041	0.176 ± 0.018	0.785 ± 0.042	0.499 ± 0.026
		ARI	0.227 ± 0.034	0.000 ± 0.003	0.140 ± 0.019	0.000 ± 0.000	0.096 ± 0.025	0.169 ± 0.117	0.359 ± 0.047	0.090 ± 0.015	0.725 ± 0.076	0.369 ± 0.028

I.2 Statistical Significance

Statistical tests treat datasets as paired, equally weighted evaluation units. For each metric, a Friedman omnibus test compares all ten methods under the null hypothesis that their performance ranks are equal across datasets. The null hypothesis is rejected for ACC, F1, NMI, and ARI, with p-values of 2.02×10^{-6}, 5.25×10^{-14}, 5.35×10^{-10}, and 1.79×10^{-9}, respectively.

Following each omnibus test, PRISM is compared with every baseline using a two-sided Wilcoxon signed-rank test over the 30 paired dataset results. Holm correction is applied separately within each metric across the nine comparisons. As reported in Table 10, PRISM significantly outperforms $7/9$, $9/9$, $6/9$, and $6/9$ baselines under ACC, F1, NMI, and ARI, respectively, at the adjusted significance level $\alpha = 0.05$.

I.3 Profile Utilization Ablation

Tables 11 and 12 compare three roles for the same temporal profiles. Profile-only applies k-means directly to $\boldsymbol{P}$, isolating the clustering capacity of the descriptors. Profile-input fuses profile vec-

Table 9: Top-performing clustering results on all thirty UEA datasets (continued).

Group	Dataset	Metric	TimesURL (AAAI'24)	UNITS (NeurIPS'24)	FEI (AAAI'25)	MVCIMTS (INFFUS'25)	k-Graph (TKDE'25)	TFMCC (AAAI'26)	FCACC (PR'26)	GTM (ICLR'26)	MUFASA (SIGMOD'26)	PRISM (ours)
HAR	BasicMotions	ACC	0.770 ± 0.158	0.544 ± 0.081	0.677 ± 0.044	0.297 ± 0.083	0.887 ± 0.132	0.985 ± 0.012	0.667 ± 0.010	0.730 ± 0.109	0.870 ± 0.095	1.000 ± 0.000
		F1	0.863 ± 0.084	0.571 ± 0.041	0.735 ± 0.072	0.152 ± 0.079	0.879 ± 0.150	0.970 ± 0.024	0.823 ± 0.010	0.731 ± 0.106	0.869 ± 0.096	1.000 ± 0.000
		NMI	0.849 ± 0.063	0.390 ± 0.111	0.679 ± 0.032	0.101 ± 0.177	0.863 ± 0.058	0.963 ± 0.025	0.805 ± 0.002	0.557 ± 0.163	0.798 ± 0.108	1.000 ± 0.000
		ARI	0.735 ± 0.145	0.299 ± 0.128	0.532 ± 0.035	0.069 ± 0.138	0.817 ± 0.122	0.961 ± 0.031	0.629 ± 0.004	0.453 ± 0.190	0.751 ± 0.146	1.000 ± 0.000
	Cricket	ACC	0.672 ± 0.081	0.492 ± 0.099	0.574 ± 0.083	0.117 ± 0.035	0.459 ± 0.022	0.692 ± 0.066	0.941 ± 0.029	0.453 ± 0.026	0.880 ± 0.096	0.987 ± 0.003
		F1	0.756 ± 0.056	0.541 ± 0.099	0.603 ± 0.059	0.042 ± 0.025	0.455 ± 0.022	0.632 ± 0.070	0.960 ± 0.002	0.439 ± 0.038	0.864 ± 0.112	0.987 ± 0.003
		NMI	0.796 ± 0.039	0.595 ± 0.084	0.679 ± 0.058	0.082 ± 0.096	0.475 ± 0.023	0.786 ± 0.046	0.951 ± 0.002	0.519 ± 0.010	0.940 ± 0.028	0.980 ± 0.004
		ARI	0.603 ± 0.078	0.336 ± 0.107	0.447 ± 0.087	0.017 ± 0.030	0.253 ± 0.025	0.584 ± 0.083	0.905 ± 0.016	0.258 ± 0.019	0.850 ± 0.103	0.980 ± 0.004
	RacketSports	ACC	0.504 ± 0.048	0.374 ± 0.037	0.426 ± 0.025	0.313 ± 0.030	0.355 ± 0.032	0.621 ± 0.040	0.417 ± 0.006	0.392 ± 0.015	0.715 ± 0.030	0.541 ± 0.101
		F1	0.527 ± 0.035	0.367 ± 0.038	0.407 ± 0.033	0.210 ± 0.051	0.353 ± 0.034	0.533 ± 0.033	0.414 ± 0.006	0.390 ± 0.013	0.714 ± 0.031	0.542 ± 0.105
		NMI	0.328 ± 0.042	0.082 ± 0.039	0.125 ± 0.029	0.069 ± 0.037	0.060 ± 0.035	0.469 ± 0.034	0.140 ± 0.004	0.102 ± 0.039	0.608 ± 0.042	0.355 ± 0.115
		ARI	0.239 ± 0.044	0.049 ± 0.028	0.092 ± 0.020	0.013 ± 0.019	0.040 ± 0.029	0.372 ± 0.044	0.088 ± 0.004	0.059 ± 0.017	0.530 ± 0.046	0.283 ± 0.123
	StandWalkJump	ACC	0.393 ± 0.018	0.541 ± 0.055	0.526 ± 0.072	0.333 ± 0.000	0.437 ± 0.036	0.444 ± 0.041	0.474 ± 0.015	0.459 ± 0.050	0.541 ± 0.056	0.430 ± 0.030
		F1	0.480 ± 0.024	0.520 ± 0.059	0.521 ± 0.077	0.167 ± 0.000	0.434 ± 0.041	0.418 ± 0.063	0.466 ± 0.019	0.455 ± 0.053	0.526 ± 0.057	0.417 ± 0.034
		NMI	0.072 ± 0.012	0.215 ± 0.147	0.122 ± 0.060	0.000 ± 0.000	0.087 ± 0.049	0.114 ± 0.042	0.100 ± 0.024	0.111 ± 0.083	0.156 ± 0.095	0.028 ± 0.008
		ARI	−0.015 ± 0.015	0.103 ± 0.126	0.037 ± 0.058	0.000 ± 0.000	−0.015 ± 0.031	0.014 ± 0.039	0.034 ± 0.024	0.018 ± 0.062	0.062 ± 0.067	0.001 ± 0.027
MEG	FaceDetection	ACC	0.500 ± 0.000	0.505 ± 0.004	0.508 ± 0.003	0.507 ± 0.004	0.504 ± 0.002	0.505 ± 0.003	0.513 ± 0.003	0.505 ± 0.003	0.503 ± 0.002	0.502 ± 0.003
		F1	0.334 ± 0.000	0.503 ± 0.006	0.505 ± 0.004	0.502 ± 0.007	0.503 ± 0.002	0.611 ± 0.059	0.512 ± 0.003	0.503 ± 0.004	0.503 ± 0.002	0.501 ± 0.004
		NMI	0.003 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.001 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
		ARI	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.001 ± 0.000	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
	HandMovement Direction	ACC	0.312 ± 0.010	0.285 ± 0.008	0.302 ± 0.009	0.302 ± 0.012	0.303 ± 0.010	0.325 ± 0.016	0.283 ± 0.006	0.296 ± 0.009	0.324 ± 0.014	0.294 ± 0.004
		F1	0.309 ± 0.010	0.280 ± 0.010	0.287 ± 0.011	0.267 ± 0.021	0.300 ± 0.013	0.333 ± 0.019	0.211 ± 0.040	0.291 ± 0.006	0.319 ± 0.013	0.291 ± 0.003
		NMI	0.021 ± 0.004	0.007 ± 0.002	0.018 ± 0.005	0.021 ± 0.010	0.014 ± 0.005	0.029 ± 0.011	0.017 ± 0.010	0.016 ± 0.006	0.029 ± 0.008	0.013 ± 0.003
		ARI	0.005 ± 0.004	−0.007 ± 0.001	0.000 ± 0.004	0.001 ± 0.003	0.000 ± 0.003	0.011 ± 0.010	−0.005 ± 0.004	0.002 ± 0.006	0.014 ± 0.008	−0.001 ± 0.002
MOTION	ArticularyWord Recognition	ACC	0.693 ± 0.052	0.295 ± 0.020	0.471 ± 0.073	0.050 ± 0.012	0.284 ± 0.016	0.503 ± 0.025	0.733 ± 0.004	0.355 ± 0.038	0.808 ± 0.044	0.939 ± 0.010
		F1	0.743 ± 0.041	0.311 ± 0.024	0.497 ± 0.082	0.011 ± 0.010	0.277 ± 0.017	0.399 ± 0.024	0.804 ± 0.005	0.353 ± 0.035	0.782 ± 0.045	0.935 ± 0.011
		NMI	0.822 ± 0.015	0.428 ± 0.020	0.619 ± 0.074	0.031 ± 0.043	0.395 ± 0.010	0.701 ± 0.016	0.841 ± 0.006	0.471 ± 0.031	0.920 ± 0.019	0.959 ± 0.005
		ARI	0.621 ± 0.034	0.140 ± 0.019	0.335 ± 0.084	0.002 ± 0.003	0.122 ± 0.009	0.447 ± 0.021	0.664 ± 0.004	0.183 ± 0.030	0.785 ± 0.041	0.907 ± 0.010
	Character Trajectories	ACC	0.065 ± 0.000	0.453 ± 0.027	0.656 ± 0.076	0.087 ± 0.032	0.450 ± 0.017	0.368 ± 0.045	0.798 ± 0.001	0.065 ± 0.000	0.065 ± 0.000	0.800 ± 0.021
		F1	0.122 ± 0.000	0.170 ± 0.037	0.668 ± 0.081	0.020 ± 0.017	0.435 ± 0.015	0.341 ± 0.028	0.858 ± 0.001	0.006 ± 0.000	0.007 ± 0.000	0.792 ± 0.019
		NMI	0.000 ± 0.000	0.178 ± 0.069	0.736 ± 0.048	0.068 ± 0.111	0.535 ± 0.012	0.576 ± 0.042	0.873 ± 0.001	0.000 ± 0.000	0.001 ± 0.000	0.835 ± 0.015
		ARI	0.000 ± 0.000	0.053 ± 0.028	0.569 ± 0.088	0.018 ± 0.035	0.321 ± 0.019	0.191 ± 0.047	0.770 ± 0.001	0.000 ± 0.000	0.000 ± 0.000	0.720 ± 0.030
	EigenWorms	ACC	0.366 ± 0.051	0.453 ± 0.037	0.418 ± 0.036	0.432 ± 0.002	0.314 ± 0.024	0.380 ± 0.021	0.378 ± 0.028	0.361 ± 0.018	0.283 ± 0.023	0.545 ± 0.033
		F1	0.303 ± 0.081	0.441 ± 0.060	0.390 ± 0.060	0.154 ± 0.017	0.292 ± 0.018	0.360 ± 0.024	0.353 ± 0.023	0.281 ± 0.011	0.265 ± 0.028	0.519 ± 0.035
		NMI	0.070 ± 0.057	0.287 ± 0.092	0.183 ± 0.015	0.030 ± 0.009	0.085 ± 0.018	0.173 ± 0.024	0.164 ± 0.028	0.065 ± 0.005	0.051 ± 0.017	0.350 ± 0.017
		ARI	0.051 ± 0.070	0.169 ± 0.049	0.145 ± 0.023	0.010 ± 0.005	0.040 ± 0.016	0.071 ± 0.027	0.097 ± 0.024	0.044 ± 0.012	0.006 ± 0.011	0.270 ± 0.043
	Handwriting	ACC	0.149 ± 0.007	0.294 ± 0.007	0.178 ± 0.010	0.088 ± 0.016	0.171 ± 0.010	0.109 ± 0.021	0.243 ± 0.004	0.126 ± 0.003	0.292 ± 0.003	0.207 ± 0.006
		F1	0.148 ± 0.008	0.288 ± 0.007	0.174 ± 0.007	0.035 ± 0.013	0.168 ± 0.010	0.121 ± 0.017	0.242 ± 0.008	0.126 ± 0.002	0.293 ± 0.007	0.205 ± 0.006
		NMI	0.201 ± 0.004	0.010 ± 0.005	0.250 ± 0.024	0.064 ± 0.027	0.226 ± 0.010	0.123 ± 0.051	0.372 ± 0.006	0.152 ± 0.005	0.461 ± 0.009	0.302 ± 0.009
		ARI	0.028 ± 0.001	−0.005 ± 0.003	0.049 ± 0.011	0.008 ± 0.005	0.043 ± 0.004	0.019 ± 0.009	0.111 ± 0.004	0.013 ± 0.002	0.162 ± 0.003	0.077 ± 0.006
	PenDigits	ACC	0.365 ± 0.015	0.469 ± 0.052	0.348 ± 0.035	0.267 ± 0.051	0.461 ± 0.017	0.655 ± 0.016	0.425 ± 0.041	0.413 ± 0.067	0.552 ± 0.049	0.523 ± 0.015
		F1	0.375 ± 0.015	0.454 ± 0.051	0.355 ± 0.032	0.171 ± 0.069	0.461 ± 0.018	0.569 ± 0.017	0.412 ± 0.030	0.397 ± 0.075	0.497 ± 0.040	0.523 ± 0.015
		NMI	0.261 ± 0.017	0.435 ± 0.063	0.306 ± 0.017	0.208 ± 0.056	0.442 ± 0.006	0.658 ± 0.009	0.323 ± 0.020	0.341 ± 0.062	0.588 ± 0.030	0.511 ± 0.018
		ARI	0.153 ± 0.011	0.282 ± 0.066	0.150 ± 0.012	0.103 ± 0.043	0.284 ± 0.009	0.520 ± 0.019	0.222 ± 0.029	0.215 ± 0.056	0.419 ± 0.051	0.378 ± 0.037
SENSOR	Ethanol Concentration	ACC	0.296 ± 0.004	0.269 ± 0.003	0.290 ± 0.002	0.296 ± 0.017	0.292 ± 0.010	0.263 ± 0.010	0.299 ± 0.003	0.279 ± 0.005	0.279 ± 0.006	0.316 ± 0.005
		F1	0.275 ± 0.007	0.209 ± 0.006	0.274 ± 0.009	0.244 ± 0.031	0.288 ± 0.011	0.468 ± 0.019	0.276 ± 0.003	0.253 ± 0.009	0.258 ± 0.005	0.316 ± 0.006
		NMI	0.020 ± 0.003	0.003 ± 0.002	0.012 ± 0.002	0.018 ± 0.011	0.007 ± 0.003	0.012 ± 0.010	0.015 ± 0.001	0.005 ± 0.001	0.005 ± 0.001	0.047 ± 0.003
		ARI	0.004 ± 0.001	−0.001 ± 0.001	0.003 ± 0.001	0.007 ± 0.007	0.001 ± 0.002	0.000 ± 0.001	0.001 ± 0.000	−0.001 ± 0.001	−0.001 ± 0.001	0.030 ± 0.002
	InsectWingbeat	ACC	0.100 ± 0.000	0.132 ± 0.002	0.154 ± 0.004	0.100 ± 0.000	0.100 ± 0.021	0.110 ± 0.002	0.165 ± 0.002	0.102 ± 0.003	0.101 ± 0.006	0.134 ± 0.002
		F1	0.182 ± 0.000	0.126 ± 0.011	0.112 ± 0.008	0.018 ± 0.000	0.014 ± 0.020	0.096 ± 0.007	0.162 ± 0.002	0.026 ± 0.008	0.117 ± 0.003	0.121 ± 0.004
		NMI	0.000 ± 0.000	0.015 ± 0.002	0.033 ± 0.002	0.000 ± 0.000	0.012 ± 0.006	0.188 ± 0.006	0.043 ± 0.001	0.001 ± 0.001	0.031 ± 0.006	0.015 ± 0.002
		ARI	0.000 ± 0.000	0.007 ± 0.001	0.013 ± 0.002	0.000 ± 0.000	0.004 ± 0.002	0.038 ± 0.005	0.022 ± 0.001	0.000 ± 0.000	0.016 ± 0.003	0.006 ± 0.001
SIMUL	LSST	ACC	0.329 ± 0.003	0.180 ± 0.010	0.333 ± 0.004	0.323 ± 0.009	0.158 ± 0.011	0.249 ± 0.017	0.343 ± 0.002	0.122 ± 0.004	0.307 ± 0.015	0.253 ± 0.005
		F1	0.239 ± 0.016	0.138 ± 0.008	0.169 ± 0.014	0.052 ± 0.022	0.115 ± 0.005	0.211 ± 0.015	0.192 ± 0.013	0.098 ± 0.007	0.226 ± 0.016	0.220 ± 0.010
		NMI	0.140 ± 0.004	0.054 ± 0.013	0.134 ± 0.004	0.024 ± 0.030	0.076 ± 0.006	0.207 ± 0.020	0.119 ± 0.002	0.031 ± 0.006	0.269 ± 0.014	0.261 ± 0.011
		ARI	0.025 ± 0.006	0.021 ± 0.007	0.086 ± 0.011	0.007 ± 0.009	0.022 ± 0.004	0.063 ± 0.031	0.045 ± 0.001	0.010 ± 0.002	0.142 ± 0.018	0.099 ± 0.006
TRAFFIC	PEMS-SF	ACC	0.238 ± 0.010	0.441 ± 0.047	0.358 ± 0.037	0.251 ± 0.090	0.427 ± 0.017	0.227 ± 0.018	0.486 ± 0.056	0.197 ± 0.003	0.243 ± 0.017	0.308 ± 0.013
		F1	0.237 ± 0.015	0.468 ± 0.059	0.378 ± 0.033	0.118 ± 0.073	0.424 ± 0.021	0.226 ± 0.020	0.511 ± 0.045	0.194 ± 0.004	0.236 ± 0.015	0.302 ± 0.016
		NMI	0.174 ± 0.023	0.417 ± 0.044	0.334 ± 0.028	0.272 ± 0.223	0.459 ± 0.020	0.127 ± 0.017	0.451 ± 0.042	0.031 ± 0.005	0.186 ± 0.080	0.273 ± 0.023
		ARI	0.065 ± 0.008	0.247 ± 0.043	0.167 ± 0.031	0.133 ± 0.111	0.256 ± 0.012	0.037 ± 0.016	0.291 ± 0.048	0.002 ± 0.002	0.069 ± 0.044	0.149 ± 0.027

Table 10: Holm-adjusted p-values from two-sided Wilcoxon signed-rank tests comparing PRISM with each baseline across 30 UEA datasets. Bold values indicate statistical significance at $p < 0.05$.

Baseline	ACC	F1	NMI	ARI
TimesURL	**0.001**	**0.003**	**0.002**	$\mathbf{5.94 \times 10^{-4}}$
UNITS	**0.009**	**0.003**	**0.001**	**0.002**
FEI	**0.015**	**0.006**	**0.003**	**0.002**
MVCIMTS	$\mathbf{5.38 \times 10^{-5}}$	$\mathbf{2.35 \times 10^{-7}}$	$\mathbf{2.27 \times 10^{-4}}$	$\mathbf{3.23 \times 10^{-4}}$
k-Graph	**0.003**	$\mathbf{6.79 \times 10^{-4}}$	**0.002**	**0.002**
TFMCC	**0.016**	**0.032**	0.172	0.105
FCACC	0.054	**0.032**	0.172	0.105
GTM	$\mathbf{4.63 \times 10^{-4}}$	$\mathbf{1.72 \times 10^{-4}}$	$\mathbf{6.35 \times 10^{-4}}$	$\mathbf{5.94 \times 10^{-4}}$
MUFASA	0.054	**0.032**	0.172	0.122

tors with raw-sequence features inside the encoder, evaluating direct feature-level integration. Full PRISM instead keeps profiles outside the encoder and converts them into fixed relational anchors. Its stronger aggregate performance supports using profiles as external semantic guidance rather than complete representations or fused inputs.

Table 11: Profile utilization ablation on all thirty datasets. Profile-only clusters profile vectors directly; Profile-input feeds them into the encoder; Full PRISM uses profiles to guide raw-MTS representation learning. Bold denotes the best variant per metric; ‡ marks a decrease relative to Full PRISM.

Group	Dataset	Variant	ACC	F1	NMI	ARI
AUDIO	DuckDuckGeese	Full PRISM	0.514 ± 0.039	0.506 ± 0.040	0.308 ± 0.058	0.202 ± 0.046
		Profile-only	0.374 ± 0.015‡	0.373 ± 0.015‡	0.142 ± 0.015‡	0.062 ± 0.010‡
		Profile-input	0.294 ± 0.014‡	0.288 ± 0.016‡	0.071 ± 0.007‡	0.006 ± 0.005‡
	Heartbeat	Full PRISM	0.619 ± 0.012	0.597 ± 0.009	0.053 ± 0.005	0.054 ± 0.011
		Profile-only	0.613 ± 0.001‡	0.588 ± 0.001‡	0.041 ± 0.000‡	0.049 ± 0.001‡
		Profile-input	0.623 ± 0.007	0.593 ± 0.005‡	0.041 ± 0.002‡	0.057 ± 0.007
	JapaneseVowels	Full PRISM	0.786 ± 0.056	0.786 ± 0.059	0.753 ± 0.037	0.664 ± 0.055
		Profile-only	0.286 ± 0.015‡	0.275 ± 0.009‡	0.180 ± 0.008‡	0.098 ± 0.013‡
		Profile-input	0.345 ± 0.013‡	0.339 ± 0.014‡	0.308 ± 0.021‡	0.163 ± 0.017‡
	PhonemeSpectra	Full PRISM	0.146 ± 0.006	0.145 ± 0.006	0.237 ± 0.004	0.053 ± 0.002
		Profile-only	0.139 ± 0.003‡	0.139 ± 0.003‡	0.209 ± 0.001‡	0.047 ± 0.001‡
		Profile-input	0.124 ± 0.003‡	0.123 ± 0.003‡	0.187 ± 0.010‡	0.038 ± 0.003‡
	SpokenArabic Digits	Full PRISM	0.494 ± 0.071	0.492 ± 0.072	0.437 ± 0.056	0.311 ± 0.072
		Profile-only	0.477 ± 0.000‡	0.472 ± 0.000‡	0.389 ± 0.001‡	0.281 ± 0.001‡
		Profile-input	0.369 ± 0.028‡	0.369 ± 0.027‡	0.287 ± 0.022‡	0.173 ± 0.023‡
ECG	Atrial Fibrillation	Full PRISM	0.420 ± 0.016	0.417 ± 0.017	0.028 ± 0.008	-0.042 ± 0.010
		Profile-only	0.413 ± 0.016‡	0.401 ± 0.014‡	0.043 ± 0.010	-0.018 ± 0.011
		Profile-input	0.447 ± 0.034	0.440 ± 0.033	0.065 ± 0.018	0.005 ± 0.024
EEG	Epilepsy	Full PRISM	0.921 ± 0.037	0.919 ± 0.039	0.822 ± 0.069	0.815 ± 0.076
		Profile-only	0.464 ± 0.004‡	0.460 ± 0.005‡	0.358 ± 0.003‡	0.235 ± 0.002‡
		Profile-input	0.564 ± 0.034‡	0.552 ± 0.034‡	0.372 ± 0.027‡	0.312 ± 0.018‡
	Finger Movements	Full PRISM	0.522 ± 0.005	0.519 ± 0.006	0.002 ± 0.001	0.000 ± 0.001
		Profile-only	0.516 ± 0.002‡	0.516 ± 0.002‡	0.001 ± 0.000‡	-0.001 ± 0.000‡
		Profile-input	0.509 ± 0.012‡	0.507 ± 0.011‡	0.001 ± 0.001‡	-0.001 ± 0.002‡
	MotorImagery	Full PRISM	0.520 ± 0.003	0.519 ± 0.003	0.001 ± 0.000	-0.001 ± 0.000
		Profile-only	0.529 ± 0.000	0.528 ± 0.000	0.003 ± 0.000	0.001 ± 0.000
		Profile-input	0.514 ± 0.006‡	0.513 ± 0.006‡	0.001 ± 0.001‡	-0.002 ± 0.001‡
	SelfRegulation SCP1	Full PRISM	0.711 ± 0.027	0.710 ± 0.028	0.138 ± 0.037	0.180 ± 0.049
		Profile-only	0.576 ± 0.000‡	0.576 ± 0.000‡	0.017 ± 0.000‡	0.021 ± 0.000‡
		Profile-input	0.556 ± 0.013‡	0.554 ± 0.014‡	0.010 ± 0.004‡	0.011 ± 0.005‡
	SelfRegulation SCP2	Full PRISM	0.521 ± 0.005	0.520 ± 0.006	0.001 ± 0.001	-0.001 ± 0.001
		Profile-only	0.503 ± 0.000‡	0.502 ± 0.000‡	0.000 ± 0.000‡	-0.003 ± 0.000‡
		Profile-input	0.520 ± 0.006‡	0.517 ± 0.007‡	0.001 ± 0.001	-0.001 ± 0.001‡
GESTURE	ERing	Full PRISM	0.848 ± 0.039	0.848 ± 0.040	0.740 ± 0.046	0.691 ± 0.062
		Profile-only	0.897 ± 0.007	0.895 ± 0.007	0.786 ± 0.009	0.774 ± 0.012
		Profile-input	0.498 ± 0.004‡	0.475 ± 0.003‡	0.408 ± 0.006‡	0.329 ± 0.005‡
	Libras	Full PRISM	0.316 ± 0.033	0.315 ± 0.033	0.387 ± 0.027	0.138 ± 0.022
		Profile-only	0.426 ± 0.017	0.415 ± 0.016	0.526 ± 0.021	0.271 ± 0.023
		Profile-input	0.263 ± 0.013‡	0.245 ± 0.012‡	0.357 ± 0.011‡	0.126 ± 0.012‡
	NATOPS	Full PRISM	0.657 ± 0.039	0.661 ± 0.042	0.629 ± 0.033	0.489 ± 0.046
		Profile-only	0.572 ± 0.038‡	0.556 ± 0.043‡	0.618 ± 0.016‡	0.477 ± 0.012‡
		Profile-input	0.351 ± 0.034‡	0.345 ± 0.035‡	0.224 ± 0.053‡	0.141 ± 0.042‡
	UWaveGesture Library	Full PRISM	0.554 ± 0.028	0.547 ± 0.024	0.499 ± 0.026	0.369 ± 0.028
		Profile-only	0.426 ± 0.031‡	0.420 ± 0.031‡	0.331 ± 0.019‡	0.209 ± 0.020‡
		Profile-input	0.421 ± 0.026‡	0.410 ± 0.028‡	0.378 ± 0.026‡	0.228 ± 0.025‡

I.4 Anchor Component Ablation

To isolate the two relational scales, each anchor is removed while the encoder, remaining objective, and training protocol are unchanged. Tables 13 and 14 report the complete results. Removing the neighborhood anchor reduces average ACC, F1, NMI, and ARI from 0.544, 0.539, 0.365, and 0.294 to 0.519, 0.512, 0.345, and 0.261, respectively. Removing the prototype anchor yields 0.526, 0.520, 0.349, and 0.271. The larger degradation without neighborhood guidance shows that preserving local profile relations is particularly important, while the consistent losses without prototype guidance confirm the complementary value of coarse sample-to-mode organization. Although either

Table 12: Profile utilization ablation on all thirty datasets (continued).

Group	Dataset	Variant	ACC	F1	NMI	ARI
HAR	BasicMotions	Full PRISM	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
		Profile-only	0.925 ± 0.000‡	0.923 ± 0.000‡	0.881 ± 0.000‡	0.826 ± 0.000‡
		Profile-input	0.847 ± 0.030‡	0.846 ± 0.030‡	0.696 ± 0.046‡	0.647 ± 0.052‡
	Cricket	Full PRISM	0.987 ± 0.003	0.987 ± 0.003	0.980 ± 0.004	0.980 ± 0.004
		Profile-only	0.776 ± 0.007‡	0.767 ± 0.009‡	0.800 ± 0.016‡	0.661 ± 0.020‡
		Profile-input	0.511 ± 0.039‡	0.492 ± 0.042‡	0.561 ± 0.030‡	0.333 ± 0.042‡
	RacketSports	Full PRISM	0.541 ± 0.101	0.542 ± 0.105	0.355 ± 0.115	0.283 ± 0.123
		Profile-only	0.426 ± 0.001‡	0.415 ± 0.002‡	0.126 ± 0.004‡	0.092 ± 0.001‡
		Profile-input	0.423 ± 0.009‡	0.402 ± 0.018‡	0.156 ± 0.012‡	0.125 ± 0.010‡
	StandWalkJump	Full PRISM	0.430 ± 0.030	0.417 ± 0.034	0.028 ± 0.008	0.001 ± 0.027
		Profile-only	0.541 ± 0.018	0.545 ± 0.020	0.194 ± 0.032	0.090 ± 0.033
		Profile-input	0.511 ± 0.049	0.489 ± 0.059	0.115 ± 0.044	0.035 ± 0.030
MEG	FaceDetection	Full PRISM	0.502 ± 0.003	0.501 ± 0.004	0.000 ± 0.000	0.000 ± 0.000
		Profile-only	0.504 ± 0.000	0.504 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
		Profile-input	0.507 ± 0.004	0.506 ± 0.005	0.000 ± 0.000	0.000 ± 0.000
	HandMovement Direction	Full PRISM	0.294 ± 0.004	0.291 ± 0.003	0.013 ± 0.003	-0.001 ± 0.002
		Profile-only	0.293 ± 0.009‡	0.291 ± 0.009	0.010 ± 0.002‡	-0.005 ± 0.002‡
		Profile-input	0.298 ± 0.021	0.294 ± 0.022	0.012 ± 0.006‡	-0.003 ± 0.005‡
MOTION	ArticularyWord Recognition	Full PRISM	0.939 ± 0.010	0.935 ± 0.011	0.959 ± 0.005	0.907 ± 0.010
		Profile-only	0.496 ± 0.020‡	0.484 ± 0.019‡	0.597 ± 0.014‡	0.331 ± 0.022‡
		Profile-input	0.243 ± 0.024‡	0.228 ± 0.023‡	0.377 ± 0.023‡	0.104 ± 0.019‡
	Character Trajectories	Full PRISM	0.800 ± 0.021	0.792 ± 0.019	0.835 ± 0.015	0.720 ± 0.030
		Profile-only	0.591 ± 0.006‡	0.572 ± 0.007‡	0.652 ± 0.007‡	0.485 ± 0.009‡
		Profile-input	0.490 ± 0.030‡	0.471 ± 0.030‡	0.528 ± 0.021‡	0.353 ± 0.025‡
	EigenWorms	Full PRISM	0.545 ± 0.033	0.519 ± 0.035	0.350 ± 0.017	0.270 ± 0.043
		Profile-only	0.602 ± 0.039	0.592 ± 0.034	0.381 ± 0.034	0.296 ± 0.055
		Profile-input	0.430 ± 0.022‡	0.377 ± 0.033‡	0.212 ± 0.015‡	0.146 ± 0.028‡
	Handwriting	Full PRISM	0.207 ± 0.006	0.205 ± 0.006	0.302 ± 0.009	0.077 ± 0.006
		Profile-only	0.165 ± 0.006‡	0.160 ± 0.007‡	0.231 ± 0.005‡	0.043 ± 0.001‡
		Profile-input	0.152 ± 0.006‡	0.147 ± 0.006‡	0.190 ± 0.007‡	0.033 ± 0.003‡
	PenDigits	Full PRISM	0.523 ± 0.015	0.523 ± 0.015	0.511 ± 0.018	0.378 ± 0.037
		Profile-only	0.512 ± 0.001‡	0.506 ± 0.000‡	0.460 ± 0.001‡	0.353 ± 0.001‡
		Profile-input	0.380 ± 0.027‡	0.384 ± 0.028‡	0.357 ± 0.026‡	0.201 ± 0.023‡
SENSOR	Ethanol Concentration	Full PRISM	0.316 ± 0.005	0.316 ± 0.006	0.047 ± 0.003	0.030 ± 0.002
		Profile-only	0.298 ± 0.000‡	0.297 ± 0.000‡	0.012 ± 0.000‡	0.005 ± 0.000‡
		Profile-input	0.283 ± 0.008‡	0.245 ± 0.008‡	0.011 ± 0.002‡	0.001 ± 0.002‡
	InsectWingbeat	Full PRISM	0.134 ± 0.002	0.121 ± 0.004	0.015 ± 0.002	0.006 ± 0.001
		Profile-only	0.163 ± 0.001	0.162 ± 0.002	0.037 ± 0.001	0.018 ± 0.001
		Profile-input	0.217 ± 0.013	0.214 ± 0.011	0.072 ± 0.012	0.029 ± 0.007
SIMUL	LSST	Full PRISM	0.253 ± 0.005	0.220 ± 0.010	0.261 ± 0.011	0.099 ± 0.006
		Profile-only	0.211 ± 0.004‡	0.192 ± 0.004‡	0.194 ± 0.001‡	0.062 ± 0.001‡
		Profile-input	0.194 ± 0.013‡	0.167 ± 0.011‡	0.159 ± 0.012‡	0.053 ± 0.007‡
TRAFFIC	PEMS-SF	Full PRISM	0.308 ± 0.013	0.302 ± 0.016	0.273 ± 0.023	0.149 ± 0.027
		Profile-only	0.431 ± 0.012	0.429 ± 0.012	0.389 ± 0.005	0.232 ± 0.008
		Profile-input	0.446 ± 0.020	0.372 ± 0.034	0.414 ± 0.061	0.263 ± 0.049

single-anchor variant can perform better on individual datasets, Full PRISM achieves the strongest aggregate results under all four metrics, supporting dual-scale relational alignment.

I.5 Anchor Semantics and Stability

We examine whether PRISM benefits from the semantic content and optimization independence of its anchors. The analysis covers AtrialFibrillation, BasicMotions, Cricket, DuckDuckGeese, ERing, Epilepsy, FaceDetection, JapaneseVowels, PEMS-SF, and UWaveGestureLibrary. Random and Shuffled replace the profile relations with fixed controls, whereas Learnable-P, Learnable-N, and Learnable-Both allow the prototype, neighborhood, or both targets to evolve during optimization. EMA updates the targets using an exponential moving average of model predictions.

Table 13: Component ablation study of PRISM on all thirty multivariate time-series datasets used in the main comparison table. Results are reported as mean ± standard deviation over five runs. Bold values indicate the best result based on unrounded values for each dataset and metric; ‡ marks a decrease relative to Full PRISM.

Group	Dataset	Variant	ACC	F1	NMI	ARI
AUDIO	DuckDuckGeese	Full PRISM	0.514 ± 0.039	0.506 ± 0.040	0.308 ± 0.058	0.202 ± 0.046
		w/o neighborhood	0.490 ± 0.026‡	0.484 ± 0.027‡	0.297 ± 0.030‡	0.190 ± 0.030‡
		w/o prototype	0.495 ± 0.035‡	0.488 ± 0.038‡	0.323 ± 0.034	0.206 ± 0.029
	Heartbeat	Full PRISM	0.619 ± 0.012	0.597 ± 0.009	0.053 ± 0.005	0.054 ± 0.011
		w/o neighborhood	0.555 ± 0.035‡	0.527 ± 0.067‡	0.046 ± 0.018‡	0.004 ± 0.024‡
		w/o prototype	0.622 ± 0.029	0.600 ± 0.020	0.057 ± 0.005	0.058 ± 0.028
	JapaneseVowels	Full PRISM	0.786 ± 0.056	0.786 ± 0.059	0.753 ± 0.037	0.664 ± 0.055
		w/o neighborhood	0.708 ± 0.029‡	0.701 ± 0.030‡	0.649 ± 0.021‡	0.546 ± 0.014‡
		w/o prototype	0.721 ± 0.042‡	0.719 ± 0.047‡	0.677 ± 0.038‡	0.572 ± 0.055‡
	PhonemeSpectra	Full PRISM	0.146 ± 0.006	0.145 ± 0.006	0.237 ± 0.004	0.053 ± 0.002
		w/o neighborhood	0.148 ± 0.005	0.147 ± 0.006	0.237 ± 0.006	0.053 ± 0.002‡
		w/o prototype	0.129 ± 0.003‡	0.128 ± 0.002‡	0.204 ± 0.003‡	0.042 ± 0.001‡
	SpokenArabic Digits	Full PRISM	0.494 ± 0.071	0.492 ± 0.072	0.437 ± 0.056	0.311 ± 0.072
		w/o neighborhood	0.403 ± 0.050‡	0.402 ± 0.050‡	0.354 ± 0.041‡	0.218 ± 0.041‡
		w/o prototype	0.489 ± 0.045‡	0.488 ± 0.045‡	0.458 ± 0.044	0.309 ± 0.048‡
ECG	Atrial Fibrillation	Full PRISM	0.420 ± 0.016	0.417 ± 0.017	0.028 ± 0.008	−0.042 ± 0.010
		w/o neighborhood	0.400 ± 0.030‡	0.397 ± 0.030‡	0.022 ± 0.013‡	−0.046 ± 0.014‡
		w/o prototype	0.385 ± 0.015‡	0.382 ± 0.016‡	0.022 ± 0.015‡	−0.051 ± 0.013‡
EEG	Epilepsy	Full PRISM	0.921 ± 0.037	0.919 ± 0.039	0.822 ± 0.069	0.815 ± 0.076
		w/o neighborhood	0.899 ± 0.031‡	0.895 ± 0.034‡	0.790 ± 0.055‡	0.771 ± 0.061‡
		w/o prototype	0.785 ± 0.088‡	0.778 ± 0.087‡	0.639 ± 0.080‡	0.600 ± 0.107‡
	Finger Movements	Full PRISM	0.522 ± 0.005	0.519 ± 0.006	0.002 ± 0.001	0.000 ± 0.001
		w/o neighborhood	0.516 ± 0.005‡	0.513 ± 0.005‡	0.001 ± 0.001‡	−0.001 ± 0.001‡
		w/o prototype	0.511 ± 0.009‡	0.510 ± 0.009‡	0.001 ± 0.001‡	−0.002 ± 0.001‡
	MotorImagery	Full PRISM	0.520 ± 0.003	0.519 ± 0.003	0.001 ± 0.000	−0.001 ± 0.000
		w/o neighborhood	0.530 ± 0.014	0.529 ± 0.013	0.003 ± 0.003	0.002 ± 0.004
		w/o prototype	0.521 ± 0.011	0.520 ± 0.011	0.002 ± 0.001	0.000 ± 0.002
	SelfRegulation SCP1	Full PRISM	0.711 ± 0.027	0.710 ± 0.028	0.138 ± 0.037	0.180 ± 0.049
		w/o neighborhood	0.680 ± 0.035‡	0.680 ± 0.034‡	0.101 ± 0.042‡	0.133 ± 0.053‡
		w/o prototype	0.706 ± 0.049‡	0.704 ± 0.050‡	0.140 ± 0.062	0.179 ± 0.083‡
	SelfRegulation SCP2	Full PRISM	0.521 ± 0.005	0.520 ± 0.006	0.001 ± 0.001	−0.001 ± 0.001
		w/o neighborhood	0.531 ± 0.005	0.530 ± 0.006	0.003 ± 0.001	0.001 ± 0.001
		w/o prototype	0.539 ± 0.011	0.538 ± 0.011	0.005 ± 0.002	0.004 ± 0.003
GESTURE	ERing	Full PRISM	0.848 ± 0.039	0.848 ± 0.040	0.740 ± 0.046	0.691 ± 0.062
		w/o neighborhood	0.758 ± 0.036‡	0.756 ± 0.036‡	0.649 ± 0.018‡	0.565 ± 0.032‡
		w/o prototype	0.810 ± 0.100‡	0.807 ± 0.102‡	0.742 ± 0.115	0.671 ± 0.154‡
	Libras	Full PRISM	0.316 ± 0.033	0.315 ± 0.033	0.387 ± 0.027	0.138 ± 0.022
		w/o neighborhood	0.267 ± 0.027‡	0.260 ± 0.027‡	0.310 ± 0.035‡	0.094 ± 0.021‡
		w/o prototype	0.303 ± 0.023‡	0.300 ± 0.021‡	0.369 ± 0.026‡	0.125 ± 0.021‡
	NATOPS	Full PRISM	0.657 ± 0.039	0.661 ± 0.042	0.629 ± 0.033	0.489 ± 0.046
		w/o neighborhood	0.663 ± 0.013	0.663 ± 0.015	0.651 ± 0.030	0.517 ± 0.025
		w/o prototype	0.654 ± 0.051‡	0.657 ± 0.053‡	0.608 ± 0.050‡	0.483 ± 0.060‡
	UWaveGesture Library	Full PRISM	0.554 ± 0.028	0.547 ± 0.024	0.499 ± 0.026	0.369 ± 0.028
		w/o neighborhood	0.493 ± 0.027‡	0.485 ± 0.025‡	0.421 ± 0.022‡	0.294 ± 0.025‡
		w/o prototype	0.633 ± 0.126	0.626 ± 0.128	0.577 ± 0.108	0.463 ± 0.148

Table 15 separates the effects of semantic content and target stability. Random and shuffled anchors remain fixed but consistently underperform the profile-derived targets, showing that fixedness alone does not provide useful guidance. Making either or both profile-side targets adaptive also reduces aggregate performance, while EMA updating produces no consistent improvement. Fixed profile anchors obtain the highest macro average under all four metrics and the best average ranks of 2.90 and 3.10 under ACC and NMI, respectively. These results support using profiles as an independently constructed semantic reference rather than self-generated training targets.

Table 14: Component ablation study of PRISM on all thirty datasets (continued).

Group	Dataset	Variant	ACC	F1	NMI	ARI
HAR	BasicMotions	Full PRISM	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
		w/o neighborhood	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
		w/o prototype	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
	Cricket	Full PRISM	0.987 ± 0.003	0.987 ± 0.003	0.980 ± 0.004	0.980 ± 0.004
		w/o neighborhood	0.970 ± 0.013‡	0.969 ± 0.014‡	0.961 ± 0.014‡	0.937 ± 0.025‡
		w/o prototype	0.916 ± 0.052‡	0.908 ± 0.059‡	0.943 ± 0.025‡	0.874 ± 0.065‡
	RacketSports	Full PRISM	0.541 ± 0.101	0.542 ± 0.105	0.355 ± 0.115	0.283 ± 0.123
		w/o neighborhood	0.548 ± 0.098	0.551 ± 0.099	0.338 ± 0.106‡	0.278 ± 0.108‡
		w/o prototype	0.569 ± 0.037	0.564 ± 0.043	0.430 ± 0.055	0.339 ± 0.055
	StandWalkJump	Full PRISM	0.430 ± 0.030	0.417 ± 0.034	0.028 ± 0.008	0.001 ± 0.027
		w/o neighborhood	0.430 ± 0.030	0.413 ± 0.044‡	0.071 ± 0.026	-0.011 ± 0.015‡
		w/o prototype	0.415 ± 0.028‡	0.400 ± 0.032‡	0.049 ± 0.024	-0.019 ± 0.025‡
MEG	FaceDetection	Full PRISM	0.502 ± 0.003	0.501 ± 0.004	0.000 ± 0.000	0.000 ± 0.000
		w/o neighborhood	0.503 ± 0.002	0.501 ± 0.003	0.000 ± 0.000	0.000 ± 0.000
		w/o prototype	0.506 ± 0.003	0.502 ± 0.005	0.000 ± 0.000	0.000 ± 0.000
	HandMovement Direction	Full PRISM	0.294 ± 0.004	0.291 ± 0.003	0.013 ± 0.003	-0.001 ± 0.002
		w/o neighborhood	0.305 ± 0.019	0.303 ± 0.019	0.016 ± 0.009	0.002 ± 0.008
		w/o prototype	0.287 ± 0.006‡	0.282 ± 0.008‡	0.006 ± 0.002‡	-0.006 ± 0.002‡
MOTION	ArticularyWord Recognition	Full PRISM	0.939 ± 0.010	0.935 ± 0.011	0.959 ± 0.005	0.907 ± 0.010
		w/o neighborhood	0.673 ± 0.029‡	0.651 ± 0.033‡	0.762 ± 0.018‡	0.550 ± 0.030‡
		w/o prototype	0.872 ± 0.048‡	0.865 ± 0.051‡	0.914 ± 0.030‡	0.817 ± 0.061‡
	Character Trajectories	Full PRISM	0.800 ± 0.021	0.792 ± 0.019	0.835 ± 0.015	0.720 ± 0.030
		w/o neighborhood	0.776 ± 0.011‡	0.770 ± 0.009‡	0.823 ± 0.004‡	0.696 ± 0.016‡
		w/o prototype	0.699 ± 0.033‡	0.691 ± 0.033‡	0.767 ± 0.012‡	0.600 ± 0.026‡
	EigenWorms	Full PRISM	0.545 ± 0.033	0.519 ± 0.035	0.350 ± 0.017	0.270 ± 0.043
		w/o neighborhood	0.552 ± 0.031	0.527 ± 0.028	0.371 ± 0.049	0.284 ± 0.051
		w/o prototype	0.482 ± 0.021‡	0.458 ± 0.018‡	0.250 ± 0.017‡	0.173 ± 0.022‡
	Handwriting	Full PRISM	0.207 ± 0.006	0.205 ± 0.006	0.302 ± 0.009	0.077 ± 0.006
		w/o neighborhood	0.194 ± 0.009‡	0.192 ± 0.009‡	0.287 ± 0.007‡	0.067 ± 0.005‡
		w/o prototype	0.197 ± 0.018‡	0.195 ± 0.018‡	0.267 ± 0.023‡	0.059 ± 0.012‡
	PenDigits	Full PRISM	0.523 ± 0.015	0.523 ± 0.015	0.511 ± 0.018	0.378 ± 0.037
		w/o neighborhood	0.530 ± 0.053	0.528 ± 0.051	0.515 ± 0.032	0.357 ± 0.046‡
		w/o prototype	0.618 ± 0.059	0.618 ± 0.058	0.588 ± 0.045	0.451 ± 0.060
SENSOR	Ethanol Concentration	Full PRISM	0.316 ± 0.005	0.316 ± 0.006	0.047 ± 0.003	0.030 ± 0.002
		w/o neighborhood	0.304 ± 0.017‡	0.304 ± 0.019‡	0.032 ± 0.017‡	0.018 ± 0.011‡
		w/o prototype	0.311 ± 0.009‡	0.310 ± 0.010‡	0.041 ± 0.005‡	0.028 ± 0.003‡
	InsectWingbeat	Full PRISM	0.134 ± 0.002	0.121 ± 0.004	0.015 ± 0.002	0.006 ± 0.001
		w/o neighborhood	0.131 ± 0.002‡	0.118 ± 0.004‡	0.014 ± 0.002‡	0.005 ± 0.001‡
		w/o prototype	0.129 ± 0.003‡	0.116 ± 0.007‡	0.014 ± 0.002‡	0.005 ± 0.001‡
SIMUL	LSST	Full PRISM	0.253 ± 0.005	0.220 ± 0.010	0.261 ± 0.011	0.099 ± 0.006
		w/o neighborhood	0.255 ± 0.008	0.225 ± 0.009	0.258 ± 0.014‡	0.098 ± 0.006‡
		w/o prototype	0.199 ± 0.010‡	0.170 ± 0.007‡	0.143 ± 0.002‡	0.048 ± 0.003‡
TRAFFIC	PEMS-SF	Full PRISM	0.308 ± 0.013	0.302 ± 0.016	0.273 ± 0.023	0.149 ± 0.027
		w/o neighborhood	0.356 ± 0.013	0.335 ± 0.014	0.361 ± 0.014	0.224 ± 0.035
		w/o prototype	0.282 ± 0.028‡	0.282 ± 0.025‡	0.227 ± 0.051‡	0.110 ± 0.035‡

I.6 Robustness and Efficiency

Noise robustness. Figure 6 reports dataset-level macro F1 under increasing input noise. PRISM remains consistently competitive across all noise levels, indicating that fixed relational guidance is robust to local input perturbations.

Embedding-dimension sensitivity. Figure 7 evaluates different embedding dimensions. PRISM remains stable across the tested dimensions, showing that its gains do not depend on a narrowly selected representation size.

Scalability. Figure 8 reports end-to-end runtime as sample size, sequence length, and variable count increase. PRISM scales favorably along all three dimensions, while several deep clustering baselines incur substantially higher costs.

Table 15: Anchor analysis over 10 diagnostic datasets. Results are macro averages over datasets and five seeds. Higher values are better.

Anchor variant	ACC	F1	NMI	ARI
Random fixed	0.629	0.627	0.481	0.409
Shuffled fixed	0.605	0.594	0.466	0.385
Learnable-P	0.670	0.664	0.539	0.473
Learnable-N	0.660	0.655	0.517	0.452
Learnable-Both	0.644	0.632	0.509	0.440
EMA-updated	0.633	0.630	0.486	0.411
Fixed profile anchors	**0.684**	**0.681**	**0.540**	**0.482**

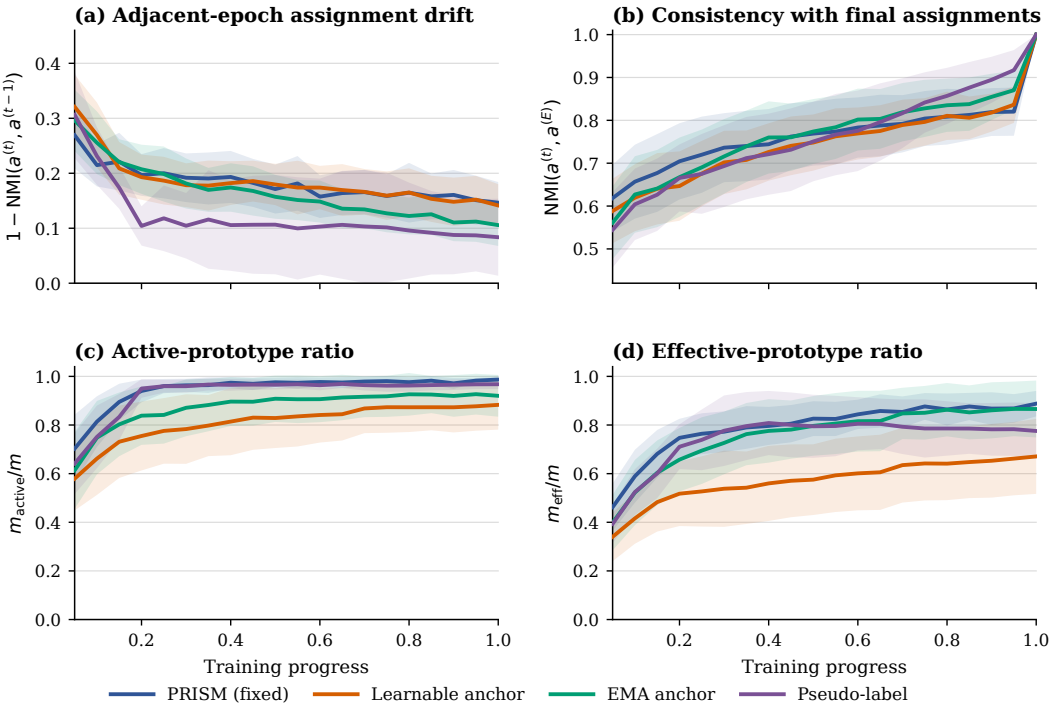


Figure 5: **Training dynamics on diagnostic datasets.** Curves report macro means over nine UEA datasets and five seeds, with shaded regions denoting 95% confidence intervals across datasets after seed averaging. Assignment drift is measured by $1 - \mathrm{NMI}(\boldsymbol{a}^{(t)}, \boldsymbol{a}^{(t-1)})$, while final consistency measures $\mathrm{NMI}(\boldsymbol{a}^{(t)}, \boldsymbol{a}^{(E)})$. Active and effective prototype counts are normalized by the prototype number m. BasicMotions is excluded from the trajectories because it contains only three training epochs and is retained in the aggregate summary table.

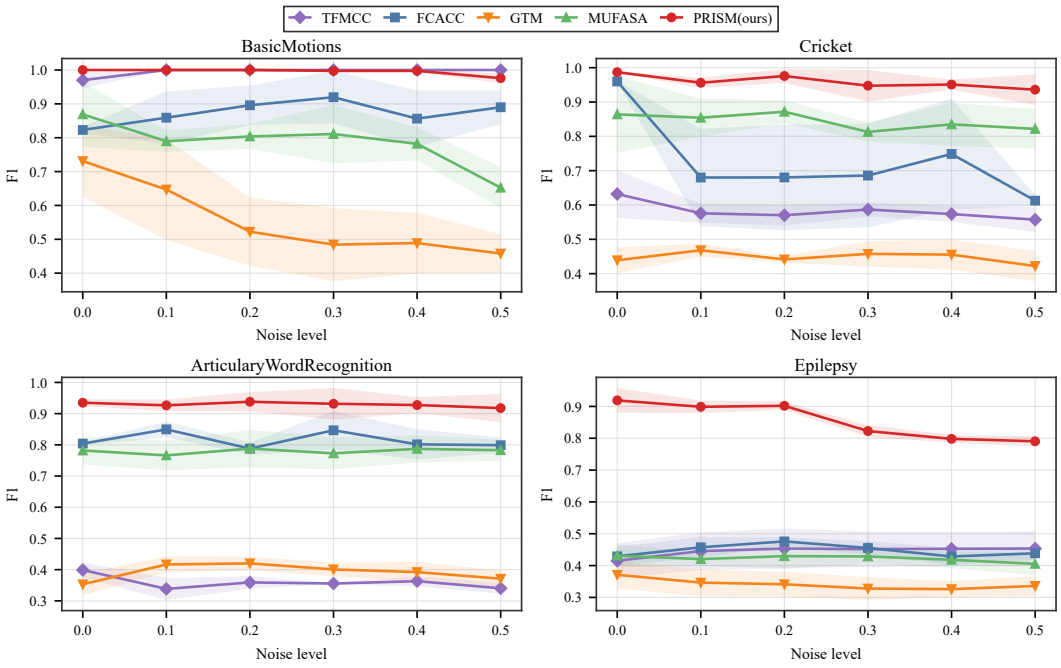


Figure 6: Macro F1 under increasing input noise. Curves and shaded regions denote the mean and standard deviation over five seeds.

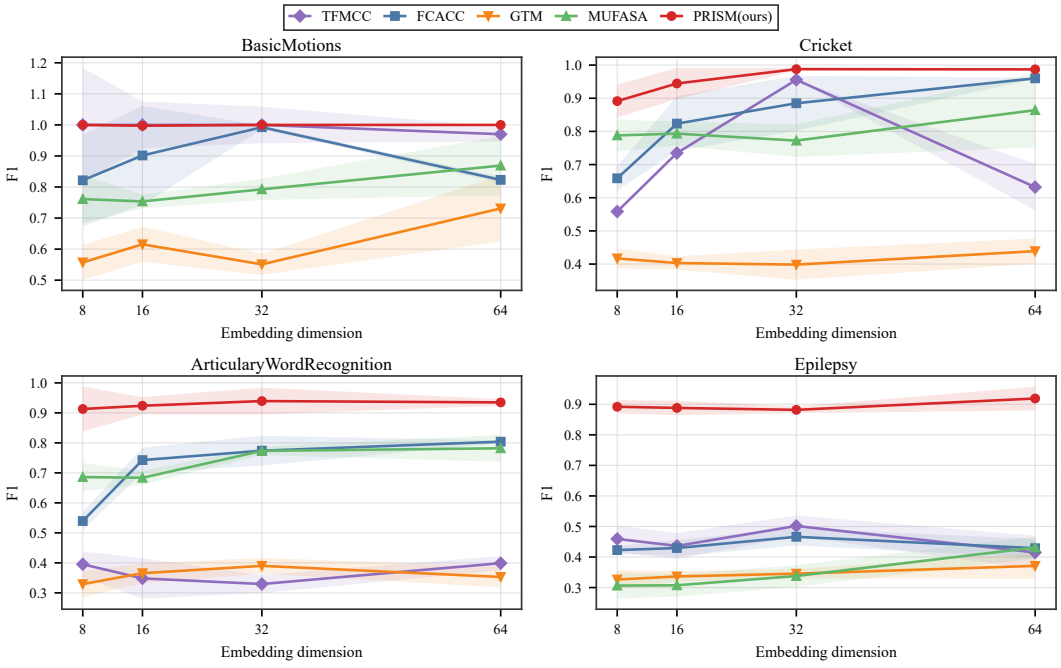


Figure 7: Macro F1 across embedding dimensions. Curves and shaded regions denote the mean and standard deviation over five seeds.

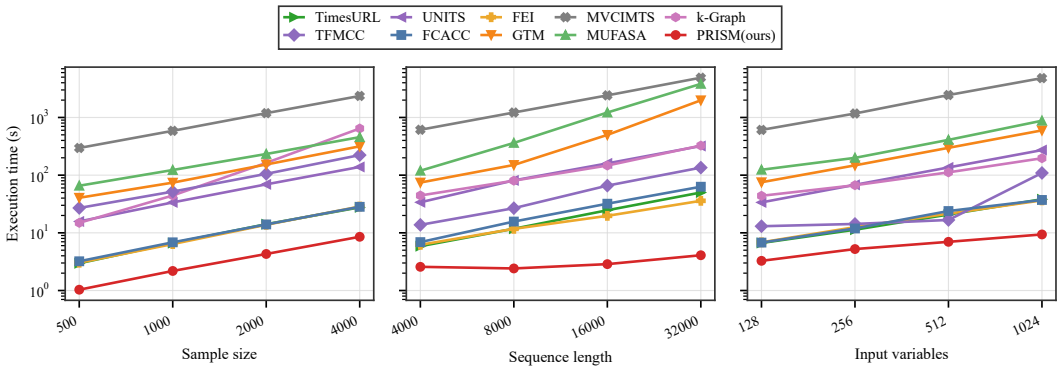


Figure 8: End-to-end runtime with increasing sample size, sequence length, and variable count. The vertical axis is logarithmic.