Research Square

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

Spike-Based Adaptation of EEG Foundation Models for Calibration Efficient Cross-Subject Brain-Computer Interfaces

Mohammad Alam

mohammad.z.alam@utu.fi

University of Turku

Article

Keywords: Brain–computer interface, EEG foundation model, cross subject adaptation, spiking neural network, neuromorphic computing, calibration efficient learning

Posted Date: September 14th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10757161/v1

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

Additional Declarations: No competing interests reported.

Spike-Based Adaptation of EEG Foundation Models for Calibration Efficient Cross-Subject Brain-Computer Interfaces

Mohammad Alam[0009-0009-0130-339X]

Department of Computing, ML and Software Engineering, University of Turku, Turku, Finland

mohammad.z.alam@utu.fi

Abstract. Non-invasive electroencephalography (EEG)-based brain–computer interfaces (BCIs) hold potential for assistive communication, neurorehabilitation, and human–machine interaction. However, cross-subject variability hinders deployment because most systems require subject-specific calibration or costly retraining for unseen individuals. EEG foundation models improve transferability, but downstream adaptation often remains computationally intensive for wearable BCIs. We propose a spike-based adaptation framework integrating a pretrained EEG foundation model with a lightweight neuromorphic module for cross-subject personalisation. The foundation model extracts transferable latent representations, while the event-driven spiking module captures subject-specific temporal variations through sparse computation. The framework uses limited calibration trials without modifying the pretrained backbone. It is evaluated on four EEG datasets comprising 175 participants across motor imagery, P300/ERP, SSVEP, and multi-session visual paradigms under leave-one-subject-out, cross-session, and few-shot protocols. Performance is compared with full fine-tuning, linear probing, transfer-learning approaches, lightweight adapters, and neuromorphic models. It achieved 86.7% average cross-subject accuracy and macro-F1 of 0.861, while retaining 84.9% accuracy with 10% calibration data. Compared with full fine-tuning, it reduced adaptation time by 87.6% and trainable parameters by 98.9%, required 0.42 mJ per trial for the adaptation pathway, and limited cross-session degradation to 4.3 percentage points. Subject-level improvements remained significant after Holm–Bonferroni correction ($p<0.001$). These findings support calibration-efficient, resource-aware cross-subject BCI systems by effectively bridging EEG foundation models with sparse neuromorphic processing.

Keywords: Brain–computer interface, EEG foundation model, cross subject adaptation, spiking neural network, neuromorphic computing, calibration efficient learning.

1 Introduction

Brain-computer interfaces (BCIs) based on electroencephalography (EEG) have emerged as a promising non-invasive pathway for translating neural activity into commands for communication, rehabilitation, neuroprosthetic control, and cognitive assistance [1]. Unlike invasive interfaces, EEG-based systems avoid surgical implantation

and support portable or wearable deployment in clinical and home environments. However, their wider adoption remains constrained by low signal to noise ratios, subject specific neural patterns, electrode and channel variability, session drift, and sensitivity to fatigue, attention, and recording conditions [1], [2]. Consequently, a model that performs well for one participant may degrade substantially when transferred to an unseen user. Most EEG-based BCI systems therefore require subject specific calibration or retraining before deployment, increasing setup time, computational cost, and user burden [2], [3]. This dependence is particularly problematic across motor-imagery, P300/ERP, SSVEP, and multi session visual paradigms, where calibration burdens and nonstationary neural responses directly affect usability and reliability.

Cross-subject EEG decoding has consequently become a central challenge in non-invasive BCI research. Existing approaches employ transfer learning, covariance alignment, adversarial domain adaptation, contrastive representation learning, and multi-source distribution matching to reduce divergence between source and target subjects [4], [5]. Although these methods have improved subject-independent decoding, they commonly remain tied to a particular dataset, paradigm, preprocessing pipeline, or channel montage. Many also require access to target domain samples during adaptation and repeatedly update dense task specific networks [4], [5]. Such strategies may be computationally expensive, prone to overfitting when only a few target trials are available, and difficult to compare fairly because reported improvements are often obtained under heterogeneous validation protocols. Random trial-level or segment-level splitting may further inflate performance by allowing subject-specific information to appear in both training and test partitions [6]. Genuine progress therefore requires strict leave-one-subject-out evaluation, explicit few-shot calibration settings, subject-level statistical analysis, and deployment-oriented measurements beyond classification accuracy alone — requirements that motivate the evaluation protocol adopted later in this work.

The recent emergence of EEG foundation models offers a powerful alternative to training task-specific decoders from limited datasets. LaBraM, EEGPT, and CBraMod use large-scale self-supervised pretraining to learn transferable representations across subjects, tasks, datasets, and recording conditions [7]– [9]. LaBraM learns generic EEG representations through masked neural-token prediction, EEGPT combines masked self-supervision with spatiotemporal representation learning, and CBraMod models heterogeneous spatial and temporal dependencies through a criss-cross transformer architecture. These developments reframe cross-subject decoding as an adaptation problem rather than complete task-specific relearning. Nevertheless, EEG foundation models do not automatically remove the personalization bottleneck. Their downstream use still commonly depends on full fine-tuning, linear probing, or dense task-specific heads, which may remain inefficient under wearable, real-time, and low-power constraints [10], [11]. Large pretrained backbones can also introduce substantial memory, latency, and energy demands, while aggressive fine-tuning may overwrite transferable representations when target subject data are scarce.

Neuromorphic computing provides a complementary route to efficient adaptation. Spiking neural networks process information through discrete events and stateful membrane dynamics, offering sparse computation, temporal memory, and compatibility with low-power neuromorphic hardware [12]. These properties are particularly relevant

to EEG, where discriminative information is distributed across transient oscillations, event-related potentials, and timing-dependent patterns. Recent studies have demonstrated the feasibility of SNNs for EEG based decoding, including adaptive graph spiking architectures for BCI classification and convolutional spiking models for intent detection [13], [14]. However, existing SNN-based BCI systems are typically trained end to end on relatively small, task-specific datasets, which limits their ability to generalize across subjects, sessions, and paradigms [13], [14].

Other SNN pipelines rely on converting pretrained artificial neural networks into spiking equivalents. Although conversion can reduce inference latency and spike activity, it may introduce approximation error and remains sensitive to quantization, threshold selection, and temporal coding design [15]. Directly trained deep SNNs also remain difficult to optimize because surrogate-gradient learning must handle discontinuous spike events and long-range temporal credit assignment [12], [16]. More importantly, current SNN-based EEG systems rarely exploit transferable representations already encoded by modern EEG foundation models. Conversely, existing EEG foundation-model pipelines remain predominantly dense and provide no dedicated event-driven mechanism for rapidly correcting subject-specific amplitude shifts, latency variations, noise characteristics, or session-dependent drift.

This paper addresses this still underexplored intersection through a spike-based adaptation framework for cross-subject non-invasive BCI. The central idea is to retain the general representation capability of a pretrained EEG foundation model while attaching a lightweight spiking adaptation module that learns only the subject-specific residual dynamics required for personalization. Instead of updating the full backbone, latent temporal variations are converted into sparse spike events and processed through stateful spiking units that adapt to individual differences in response timing, amplitude, and noise structure. This design separates universal representation learning from subject-specific correction, thereby reducing the number of trainable parameters and limiting catastrophic overwriting of pretrained knowledge. It also supports event-driven operation, in which computation is activated selectively by informative latent changes rather than performed uniformly over every dense activation.

The proposed framework is organized around *four* technical objectives. ***First,*** it improves cross-subject decoding under strict leave-one-subject-out evaluation across multiple EEG paradigms. ***Second,*** it reduces calibration burden by adapting from a limited number of trials from each unseen user. ***Third***, it lowers adaptation latency and computational cost through sparse spike-based updates while preserving the pretrained backbone. ***Fourth,*** it supports deployment-oriented evaluation by jointly measuring accuracy, macro-F1, calibration trials, trainable parameters, adaptation time, operation count, cross-session degradation, and estimated energy consumption. This broader evaluation is necessary because a marginal improvement in classification accuracy may not justify extensive calibration, dense retraining, or high inference cost in practical BCI deployment. The main contributions of this study are as follows:

- A neuromorphic adapter is introduced to perform event-driven subject-specific correction on top of pretrained EEG representations.
- A spike-encoding and adaptation mechanism is formulated to capture temporal deviations without retraining the entire foundation model. In addition, a cross-

subject evaluation protocol is developed across four large EEG datasets comprising 175 participants and covering motor imagery, P300/ERP, SSVEP, and multi-session visual paradigms, with few-shot calibration levels, repeated runs, confidence intervals, and subject-wise significance testing.
- The proposed method is compared against full fine-tuning, linear probing, transfer learning, dense lightweight adapters, and representative neuromorphic baselines under a unified preprocessing and validation pipeline.
- Evaluation extends beyond predictive performance to include adaptation efficiency, parameter economy, cross-session robustness, and estimated or measured energy cost, providing a more deployment-relevant basis for comparison.

The remainder of this paper is organised as follows. *Section 2* reviews related work. *Section 3* formulates the cross-subject adaptation problem. *Section 4* presents the methods and proposed spike-based adaptation framework. *Section 5* describes the data and experimental evaluation protocol. *Section 6* reports and analyses the results. *Section 7* discusses the implications, limitations, and deployment considerations. Finally, *Section 8* concludes the paper and outlines directions for future work.

2 Related Work

2.1 Cross-Subject EEG Decoding and Domain Adaptation

Non-invasive EEG-based brain–computer interfaces remain constrained by pronounced inter-subject and inter-session variability arising from anatomical differences, electrode placement, impedance, cognitive strategy, fatigue, and nonstationary neural responses [1]. Conventional subject-dependent decoders can achieve strong within-user performance, but extensive calibration limits their scalability in assistive, rehabilitative, and wearable settings [2], [3]. Earlier BCI pipelines employed filter bank common spatial patterns for motor imagery decoding [17], followed by feature normalisation, covariance alignment, Euclidean space alignment [18], transfer learning [19], and domain adaptation [4], [5]. These methods reduce source and target distribution mismatch, yet many remain task-specific, computationally dense, and dependent on target-user samples. Their performance is also sensitive to preprocessing, channel configuration, calibration size, trial duration, and evaluation design [19], [20]. Strict leave-one-subject-out evaluation is therefore essential because random trial-level splitting can leak participant-specific information into both training and testing partitions [6]. The unresolved challenge is not merely higher decoding accuracy, but rapid adaptation to unseen users using few calibration trials while maintaining cross-session stability and practical computational cost.

2.2 Transformer-Based EEG Representation Learning

Deep EEG decoding initially progressed through convolutional architectures that learned temporal and spatial features directly from signals. EEGNet demonstrated that temporal filtering, depthwise spatial convolution, and separable feature extraction

could provide compact cross-paradigm decoding [21]. Deeper convolutional models subsequently improved feature learning and visualisation of discriminative neural activity but remained largely task and dataset specific [22]. Transformers extended this progression by modelling long-range temporal dependencies and interactions across electrodes. Contemporary architectures combine temporal convolutions, patch embeddings, channel attention, multiscale windows, and hierarchical attention blocks [2], [23]. EEGEncoder, for example, integrates transformer and temporal convolutional components to capture local and global motor imagery features [23]. Nevertheless, such models are commonly trained end to end for individual datasets or paradigms, making them sensitive to changes in montage, sampling rate, and recording environment. Dense attention operations and backbone fine-tuning also increase memory and computational demand. Transformer-based decoding therefore improves representational capacity but does not independently resolve calibration-efficient personalisation.

2.3 EEG Foundation Models and General-Purpose Pretraining

EEG foundation models shift the field from task-specific supervised learning towards large-scale self-supervised pretraining on heterogeneous neural recordings. Unlike conventional CNN and transformer decoders [21]–[23], they aim to learn reusable representations before downstream adaptation. LaBraM introduced channel patch tokenisation and masked neural code prediction using approximately 2,500 hours of EEG from multiple datasets [7]. EEGPT combined masked reconstruction with spatiotemporal representation alignment to address noise, channel mismatch, and inter-subject variability [8]. CBraMod modelled spatial and temporal dependencies through criss-cross attention across diverse EEG tasks [9]. However, foundation-level performance does not guarantee calibration-free BCI operation [10]. Downstream use still relies primarily on linear probing, full fine-tuning, or dense task heads, while larger backbones impose memory, latency, and energy costs [10], [11]. More importantly, existing foundation models provide limited mechanisms for rapidly correcting subject-specific latency shifts, amplitude variations, noise, and session drift. Full fine-tuning can model these deviations but updates a large parameter space and may overfit scarce calibration data. Linear probing preserves the backbone but may lack sufficient personalisation capacity. A clear gap therefore remains between transferable population-level representation learning and efficient adaptation to individual users.

2.4 Spiking Neural Networks and Neuromorphic BCI

Spiking neural networks represent information through discrete events and stateful neuronal dynamics. Temporal, rate-based, or change-based encoding can convert EEG signals or latent features into spike trains, enabling membrane states to capture temporal accumulation and delayed responses. Deep SNNs have demonstrated strong performance with sparse neuronal activity [12], while EEG-based SNNs have incorporated adaptive graph convolution, recurrent temporal modelling, and convolutional spike processing [13], [14]. Fast-SNN improves ANN-to-SNN conversion efficiency [15], and temporally heterogeneous dendritic dynamics support multi-timescale learning [16]. Most existing SNN-based BCI systems, however, are trained on relatively small task-

specific datasets, employ fixed spike encoders, or replace an entire dense architecture with a spiking counterpart [13]–[16]. They rarely exploit representations learned by EEG foundation models and often provide limited cross-subject or cross-session validation. A more suitable integration is therefore to preserve the pretrained backbone and use a lightweight spiking adapter only for subject-specific residual correction. Its value must nevertheless be established against linear probing, dense adapters, low-rank adaptation, conventional SNNs, and full fine-tuning.

2.5 Research Gap and Positioning of the Proposed Work

Existing research addresses cross-subject transfer, EEG foundation modelling, and neuromorphic decoding largely as separate problems. Classical spatial filters support efficient handcrafted decoding [17]; compact CNNs provide task-specific feature learning [21], [22]; alignment and transfer methods reduce source and target mismatch [4], [5], [18]– [20]; foundation models provide reusable EEG representations [7]– [11]; and SNNs enable sparse temporal computation [12]– [16]. However, no established method jointly integrates a frozen EEG foundation backbone, subject-adaptive latent-to-spike encoding, neuromorphic residual correction, few-trial calibration, cross-session robustness, and deployment-oriented efficiency. The proposed framework connects these directions through a lightweight spike-based residual adapter attached to a frozen EEG foundation model. Subject-conditioned thresholds convert informative latent variations into sparse signed events, recurrent spiking dynamics estimate subject-specific corrections, and an adaptive gate controls their fusion with the transferable representation. The novelty lies in using neuromorphic dynamics for compact personalisation rather than replacing or retraining the foundation backbone. Table 1 summarises this positioning.

Table 1 Comparative Positioning of Recent Cross-subject, Foundation model, and Neuromorphic EEG Approaches

Study	Principal Method	Pretraining or Adaptation Strategy	Cross-Subject Relevance	Neuromorphic Adaptation	Deployment Consideration	Principal Gap
She *et al.* [4]	Wasserstein-distance domain adaptation network	Source–target distribution alignment	Explicit cross-subject motor imagery adaptation	No	Dense target adaptation	No foundation model representation or event-driven personalisation
Zhong *et al.* [5]	Transfer learning with target calibration	Cross-subject feature transfer and calibration	Directly addresses unseen-user adaptation	No	Requires target calibration	Does not address spike sparsity or neuromorphic efficiency
EEGEncoder [23]	Transformer–TCN motor imagery decoder	Dual-stream temporal and spatial representation learning	Supports subject-independent motor imagery decoding	No	Dense end-to-end inference	Task-specific training and no foundation model personalisation
LaBraM [7]	Large EEG foundation model	Masked neural code prediction with channel patch tokenisation	Learns transferable representations from heterogeneous EEG	No	Large pretrained backbone	Conventional downstream adaptation remains necessary
EEGPT [8]	Pretrained EEG transformer	Masked reconstruction and spatiotemporal alignment	Targets channel mismatch and inter-subject variability	No	Dense transformer processing	Few-trial and event-driven adaptation remain unresolved
CBraMod [9]	Criss-cross EEG foundation model	Separate spatial and temporal masked modelling	Evaluated across diverse EEG tasks	No	Compact relative to larger foundation models but still dense	No explicit calibration-efficient neuromorphic personalisation

Study	Principal Method	Pretraining or Adaptation Strategy	Cross-Subject Relevance	Neuromorphic Adaptation	Deployment Consideration	Principal Gap
Comparative EEG foundation model evaluation [10]	Multi-model foundation model assessment	Linear probing and downstream adaptation across BCI tasks	Examines transfer across tasks	No	Shows task-dependent value of foundation models	Standardised lightweight personalisation remains unresolved
Tóth *et al.* [11]	Fine-tuning and quantisation of EEG foundation models	Reduced-precision deployment adaptation	Supports resource-aware foundation model use	No	Targets low-power execution	Quantisation does not provide subject-adaptive temporal correction
Gong *et al.* [13]	Adaptive graph-convolutional SNN	Task-specific spike-based EEG decoding	Limited cross-subject evaluation	Yes	Sparse temporal processing	No foundation model backbone or residual personalisation
Lutes *et al.* [14]	Convolutional SNN for EEG intent detection	Spike-based task-specific classification	Application-specific EEG decoding	Yes	Event-driven inference	Limited transfer and calibration analysis
Proposed Framework	**EEG foundation model with subject-adaptive spiking residual adapter**	**Frozen backbone with signed latent event encoding, LIF residual correction, and adaptive fusion**	**Strict LOSO, few-shot, and cross-session evaluation**	**Yes**	**Targets reduced calibration, parameters, latency, operations, and energy**	**Jointly integrates transferable representation learning and neuromorphic personalisation**

3 Cross-Subject Adaptation Problem Formulation

Motivated by the limitations identified in Section 2, we formulate cross-subject EEG decoding as a calibration-constrained adaptation problem and establish the mathematical foundation of the proposed spike-based framework. The formulation progressively defines the source-target learning setting, latent representation decomposition, subject-adaptive spike generation, neuromorphic residual correction, and the optimisation objective governing efficient subject personalisation.

3.1 Cross-Subject EEG Learning Setting

Let the complete source-domain dataset be represented as

$$\mathcal{D}_s = \left\{\mathcal{D}_s^{(i)}\right\}_{i=1}^{N_s}, \qquad (1)$$

where $N_s$denotes the number of source subjects and $\mathcal{D}_s^{(i)}$is the EEG dataset associated with the i-th subject. Each source-subject dataset contains multiple labelled EEG trials and is defined as

$$\mathcal{D}_s^{(i)} = \left\{\left(X_j^{(i)}, y_j^{(i)}\right)\right\}_{j=1}^{M_i}, \qquad (2)$$

where $X_j^{(i)} \in \mathbb{R}^{C\times T}$denotes the j-th EEG trial of subject i, $C$is the number of EEG channels, $T$is the number of temporal samples, $y_j^{(i)} \in \{1, \dots, K\}$is the corresponding class label, $K$is the number of BCI classes, and $M_i$is the number of available trials. For an unseen target subject u, only a small labelled calibration set is assumed to be available:

$$\mathcal{D}_u^{\text{cal}} = \left\{\left(X_q^u, y_q^u\right)\right\}_{q=1}^{M_u^{\text{cal}}}, \qquad (3)$$

where M_u^{cal}denotes the number of calibration trials. The calibration-efficient setting requires

$$M_u^{\text{cal}} \ll M_u, \qquad (4)$$

where $M_u$is the total number of trials available for the target subject. The remaining target trials form the independent test set and are not used during adaptation. The principal challenge arises because the joint distributions of the source and target domains are not identical [4, 5,18]:

$$P_s(X, y) \neq P_u(X, y). \qquad (5)$$

This distributional mismatch reflects [19-20] inter-subject differences in cortical activation, response latency, signal amplitude, electrode impedance, cognitive strategy, and session-specific recording noise. The learning objective must therefore minimize target-subject prediction error without requiring full retraining of the pretrained model.

3.2 Foundation Model Representation Decomposition

Let F_θ denote a pretrained EEG foundation model [7, 11] parameterized by θ. Given an EEG trial X, the input is first processed by a preprocessing operator $\mathcal{P}(\cdot)$, which may include filtering, artefact attenuation, normalization, resampling, and temporal segmentation. The resulting latent representation is expressed as

$$H = F_\theta\big(\mathcal{P}(X)\big),\ H \in \mathbb{R}^{L\times d}, \qquad (6)$$

where $L$is the number of temporal latent tokens and $d$is the embedding dimension. During subject-specific calibration [10-11], the foundation-model parameters θremain frozen to preserve the transferable knowledge learned during large-scale pretraining. For the unseen subject u, the extracted latent representation is conceptually decomposed as

$$H^u = H^{\text{inv}} + \Delta^u + \varepsilon^u, \qquad (7)$$

where H^{inv}contains subject-invariant task information, Δ^urepresents systematic subject-specific deviation, and ε^udenotes transient noise and unmodelled variation. The proposed framework does not relearn H^{inv}; instead, it estimates a compact correction corresponding to Δ^u.

3.3 Subject-Adaptive Latent-to-Spike Encoding

The dense latent tokens are first projected into a lower-dimensional adaptation space:

$$Z_t = H_t W_p + b_p,\ Z_t \in \mathbb{R}^{d_s}, \qquad (8)$$

where $H_t$is the t-th latent token, $W_p \in \mathbb{R}^{d\times d_s}$is a trainable projection matrix, $b_p$is the projection bias, and $d_s \ll d$is the dimension of the spiking adaptation space. This projection restricts the number of trainable parameters and isolates the personalization pathway from the foundation-model backbone. A subject-context vector is then estimated from the available calibration trials:

$$c^u = \frac{1}{M_u^{\text{cal}} L} \sum_{q=1}^{M_u^{\text{cal}}} \sum_{t=1}^{L} Z_{q,t}^u, \qquad (9)$$

where $Z_{q,t}^u$denotes the projected latent token at time step $t$for calibration trial q. The vector c^usummarizes the latent statistical profile of the unseen subject. Using this context, the spike-generation threshold is made subject dependent:

$$\tau^u = \text{softplus}(\tau_0 + W_\tau c^u), \qquad (10)$$

where $\tau_0$is a learnable base threshold, $W_\tau$maps the subject context into the spike space, and the softplus function ensures positive threshold values. Unlike fixed-threshold

encoding, this formulation adapts event sensitivity to each user. Signed spike events are generated from temporal changes in the projected latent sequence:

$$S_t^{\text{in}} = \Theta(|Z_t - Z_{t-1}| - \tau^u) \odot \text{sgn}(Z_t - Z_{t-1}), \qquad (11)$$

where $\Theta(\cdot)$is the element-wise Heaviside function, $\text{sgn}(\cdot)$preserves the direction of the latent variation, and $\odot$ denotes element-wise multiplication. Consequently, a spike is produced only when a latent change exceeds the subject-specific threshold.
The resulting activity level is measured through the spike-density ratio

$$\rho^u = \frac{\sum_{t=1}^{L} \|S_t^{\text{in}}\|_0}{Ld_s}, \qquad (12)$$

where $\|\cdot\|_0$counts the number of non-zero spike events. A lower value of ρ^uindicates sparser event generation and reduced neuromorphic computation.

3.4 Neuromorphic Residual Adaptation

The encoded spike sequence is processed by stateful leaky integration and fire neurons [12, 13, 16]. The membrane dynamics are defined as

$$U_t = \lambda U_{t-1} + W_s S_t^{\text{in}} + W_r S_{t-1}^{\text{out}} - V_{\text{th}} S_{t-1}^{\text{out}}, \qquad (13)$$

where $U_t$is the membrane potential, $\lambda \in [0, 1]$is the membrane-decay factor, $W_s$and $W_r$are feedforward and recurrent synaptic matrices, and $V_{\text{th}}$is the neuronal firing threshold. The final term resets the membrane state following an output spike. The output spike vector is obtained as

$$S_t^{\text{out}} = \Theta(U_t - V_{\text{th}}). \qquad (14)$$

Because $\Theta(\cdot)$is non-differentiable, surrogate-gradient learning is used [12, 15] during backpropagation, while discrete spike generation is retained during the forward pass. The temporal spike outputs are aggregated into a subject-specific residual representation:

$$R^u = \frac{1}{L}\sum_{t=1}^{L} S_t^{\text{out}} W_o, \qquad (15)$$

where $W_o$maps the spiking representation back to the foundation-model embedding space. The strength of the residual correction is controlled by a subject-conditioned gate:

$$\alpha^u = \sigma(w_\alpha^\top c^u + b_\alpha), \qquad (16)$$

where $\sigma(\cdot)$ is the sigmoid function, $w_\alpha$is a trainable gating vector, and $b_\alpha$is the corresponding bias. This gate enables stronger adaptation for subjects with greater latent deviation and weaker adaptation when the pretrained representation is already sufficiently aligned. The final adapted representation is therefore defined as

$$\tilde{h}^u = \text{Pool}(H^u) + \alpha^u R^u, \qquad (17)$$

where $\text{Pool}(\cdot)$aggregates the foundation-model tokens through mean pooling, attention pooling, or a designated classification token. The residual formulation preserves subject-invariant information while injecting only the required neuromorphic correction.

3.5 Calibration Constrained Learning Objective

The adapted representation $\tilde{h}^u$is passed to a lightweight classification head. During calibration, only the projection layer, adaptive threshold generator, spiking adapter,

residual gate, and classifier are updated, while the foundation-model backbone remains fixed. The complete training objective combines predictive accuracy, source–target alignment, spike sparsity, and parameter regularisation:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{cls}} + \beta_1 \mathcal{L}_{\text{align}} + \beta_2 \mathcal{L}_{\text{spike}} + \beta_3 \mathcal{L}_{\text{reg}}. \quad (18)$$

Here, $\mathcal{L}_{\text{cls}}$denotes the classification loss, $\mathcal{L}_{\text{align}}$reduces the class-conditional discrepancy between source and target representations, $\mathcal{L}_{\text{spike}}$penalises excessive neuronal firing, and $\mathcal{L}_{\text{reg}}$constrains the magnitude of the trainable adapter parameters. The coefficients β_1, β_2, and $\beta_3$control the relative contributions of the auxiliary objectives. Accordingly, the proposed formulation transforms cross-subject EEG decoding from full-model retraining into a constrained residual personalisation problem. It integrates subject-conditioned latent event generation, signed temporal spike encoding, stateful neuromorphic correction, adaptive residual gating, and sparsity-aware optimisation within a frozen EEG foundation-model pipeline**.** These mathematical components provide the basis for the architecture, training strategy, and inference workflow presented in Section 4.

4 Methods

Based on the mathematical framework established in Section 3, we present the complete spike-based adaptation framework illustrated in Fig. 1. The framework combines a frozen EEG foundation model with a lightweight neuromorphic residual pathway that learns subject-specific corrections from limited calibration data. Sparse signed spikes are generated from latent variations, processed through recurrent spiking dynamics, and fused with the transferable foundation representation before classification. By separating general representation learning from subject-specific adaptation, the framework preserves pretrained knowledge while reducing calibration and adaptation costs.

4.1 Architecture and Processing Workflow of the Proposed Framework

As shown in Fig. 1, the proposed framework comprises *four* connected stages: EEG preprocessing, foundation representation extraction, spike-based residual adaptation, and final classification. Each EEG trial is first standardized through filtering, normalization, and segmentation before being passed to the pretrained EEG foundation encoder [7-11]. The encoder remains frozen during subject adaptation and produces transferable latent representations that preserve the general knowledge acquired from large-scale EEG data [10-11]. The resulting latent sequence is then directed to the proposed adaptation pathway, whose internal operation is detailed in Fig. 1. A compact projection layer first constructs the subject-specific adaptation space, while the available calibration trials are used to estimate subject context. This context controls the adaptive firing threshold and regulates the conversion of meaningful latent variations into sparse signed spike events. The neuromorphic residual module processes these events through recurrent spiking dynamics and generates a compact subject-specific correction. An adaptive gate then controls the strength of this residual before it is fused with the original foundation representation for classification. By separating transferable

representation learning from subject-specific personalization, the framework avoids full-backbone fine-tuning and confines adaptation to a small set of trainable parameters.

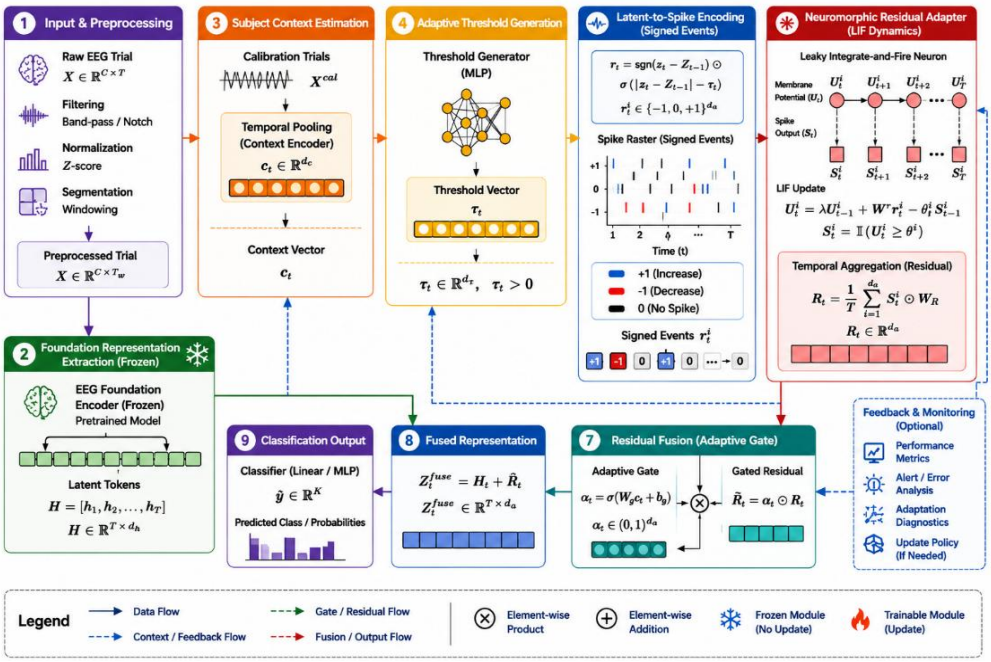


Fig. 1. Overview of the proposed framework. Pre-processed EEG trials are encoded by a frozen foundation model into latent tokens, while subject calibration trials generate a context vector that drives an adaptive threshold generator.

4.2 Subject-Context Estimation and Adaptive Spike Encoding

Following foundation representation extraction, the latent tokens are projected into the lower-dimensional adaptation space defined in Eq. (8). This projection retains task-relevant temporal information while reducing the dimensionality and number of parameters involved in subject-specific adaptation. Unlike approaches that encode raw EEG amplitudes directly into spikes [13–16], the proposed encoder operates on foundation-model representations that already capture transferable spatial, temporal, and spectral characteristics. As illustrated in Fig. 1, the available calibration trials are aggregated to estimate the subject-context vector defined in Eq. (9). This context characterises the latent response profile of the unseen target subject and conditions the adaptive firing threshold in Eq. (10). Spike sensitivity is therefore adjusted according to the target subject rather than being governed by a fixed global threshold. Meaningful temporal changes are subsequently converted into signed spike events according to Eq. (11), with positive and negative spikes preserving the direction of latent variation. Subthreshold fluctuations are suppressed, thereby limiting noise-driven activity and producing the sparse event representation quantified by Eq. (12). The resulting spike sequence provides a compact, subject-sensitive input to the neuromorphic residual module.

4.3 Neuromorphic Residual Personalization

The signed spike sequence is processed by the stateful leaky integrate and fire network defined in (13) and (14) [12,13, 16]. Membrane potentials accumulate temporally distributed evidence, while leakage and recurrent connections regulate the persistence of earlier events. This mechanism is well suited to EEG signals, in which discriminative neural responses often emerge gradually across an observation window rather than at a single instant. The output spikes are temporally aggregated through (15) to form a residual representation that captures the subject-specific deviation from the transferable foundation representation. Importantly, the neuromorphic module does not replace the pretrained feature space or relearn the complete classification function. Instead, it estimates only the correction required for the current subject. The adaptive gate in (16) further controls the contribution of this correction. Subjects already well represented by the backbone receive a limited residual adjustment, whereas larger corrections can be introduced for subjects exhibiting stronger latent mismatch. The gated residual is finally added to the pooled foundation representation according to (17), producing the personalized representation used by the classifier. This residual design constitutes the central distinction of the proposed model. Rather than combining a foundation encoder with an independent SNN classifier, the framework uses spike dynamics specifically to model subject-dependent latent deviations while retaining the general representation learned during pretraining.

4.4 Calibration-Constrained Learning Strategy

The framework was trained in two connected stages. During source-domain training, the projection layer, threshold generator, neuromorphic adapter, residual gate, and classification head were optimised across multiple source subjects. Subject-wise episodic training was performed by treating one source subject as a temporary target, while the remaining subjects supplied transferable training information [4], [5], [19]. This procedure exposed the adaptation pathway to heterogeneous subject distributions before deployment. For an unseen target subject, the foundation encoder remained frozen, and only the lightweight adaptation parameters were updated using the limited calibration set defined in (3). This strategy limited overfitting in the large backbone and reduced the number of gradient updates required for personalisation. The total objective in (18) jointly balanced classification accuracy, source–target alignment, spike sparsity, and parameter regularisation. The classification term preserved task discrimination, the alignment term reduced representation mismatch, and the sparsity term penalised excessive neuronal firing. Regularisation further constrained the magnitude of the learned correction and discouraged unnecessary modification of the foundation representation. No target-test trial was used during calibration, threshold estimation, model selection, or early stopping. This separation ensured that the reported cross-subject performance reflected genuine adaptation to unseen users rather than target-data leakage [6].

4.5 Inference Workflow and Computational Characteristics

For each new subject, a small calibration subset was first passed through the frozen foundation encoder to estimate the subject context and initialise the adaptive threshold and residual gate. The compact adaptation module was then optimised for a limited

number of epochs. The resulting calibration and inference workflow is summarised in Fig. 2**.** Once calibration was completed, each incoming EEG trial followed the same preprocessing and foundation encoding pipeline before entering the subject-specific spike pathway. During inference, only latent changes exceeding the calibrated threshold activated the neuromorphic adapter. Sparse event generation therefore reduced the number of synaptic operations within the personalisation pathway, while the frozen backbone eliminated repeated large-scale parameter updates. The proposed design consequently improved parameter efficiency and calibration speed without discarding the representational capacity of the EEG foundation model. Because the backbone remained dense, computational and energy gains were reported separately for the adapter and the complete system to avoid overstating end-to-end neuromorphic efficiency. The framework was evaluated in Section 5 through cross-subject and cross-session experiments, calibration-efficiency analysis, parameter and computational comparisons, statistical testing, and component-wise ablation studies.

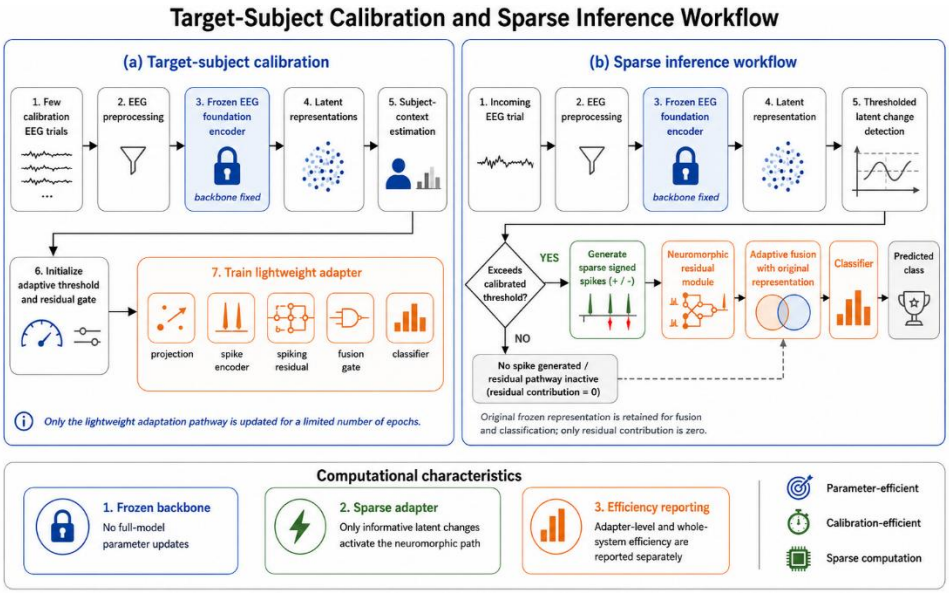


Fig. 2. Target-subject calibration and sparse inference workflow. The frozen EEG foundation encoder is retained during calibration and inference, while subject-specific adaptation is confined to the lightweight spike-based pathway. Sparse signed spikes activate the neuromorphic residual module only when calibrated latent changes exceed the adaptive threshold

5 Data and Experimental Evaluation

The experiments were conducted across multiple publicly available EEG benchmarks representing different participants, sessions, acquisition configurations, and BCI paradigms. A strict subject-independent protocol was adopted so that no EEG trial from the target participant was used during source-model training. Limited labelled trials

from the target participant were subsequently introduced only through the explicitly defined calibration procedure. All competing methods were evaluated using identical preprocessing, source–target partitions, calibration budgets, and test trials. This design ensured that the observed performance differences were attributable to the proposed subject-adaptive spike-based residual mechanism rather than to differences in data preparation or evaluation conditions.

5.1 Experimental Design and Objective Validation Strategy

The experiments were structured around the *four* research objectives established in Section 1. Cross-subject decoding was evaluated using leave-one-subject-out (LOSO) validation, in which each participant served once as an unseen target subject. Calibration efficiency was examined under progressively reduced target-subject data budgets, while the EEG foundation model backbone remained frozen. Computational efficiency was evaluated using trainable parameters, adaptation time, memory usage, operation count, spike density, inference latency, and energy consumption. Deployment robustness was assessed through cross-session evaluation and performance retention under limited calibration. Comparative, ablation, and subject-level statistical analyses were conducted using identical partitions and calibration samples across all methods.

5.2 EEG Data Sources and Cohort Characteristics

Four publicly available EEG datasets published between 2022 and 2025 were selected to evaluate the proposed framework under cross-subject and cross-session adaptation settings. Collectively, the datasets comprised 175 dataset-specific participants and covered motor imagery (MI), P300/ERP, SSVEP, motor execution, RSVP, and stimulus-evoked visual EEG paradigms. Their variation in participant cohorts, recording sessions, channel configurations, trial volumes, and acquisition conditions supported the assessment of subject variability, calibration efficiency, temporal generalisation, and robustness to recording artefacts. The dataset characteristics, experimental roles, and selection rationale are summarised in Table 2. For each dataset, an EEG trial was represented as

$$X_i^u \in \mathbb{R}^{C \times T}, \qquad (19)$$

where $u$denoted the participant, $i$denoted the trial, Crepresented the number of EEG channels, and Trepresented the number of temporal samples. The corresponding class label was denoted by y_i^u. Participant identities, session boundaries, class labels, event markers, and paradigm information were preserved throughout preprocessing and evaluation. To ensure fair cross-subject evaluation, channel configurations and sampling frequencies were harmonised only after subject-wise partitioning, thereby preventing information leakage among the training, adaptation, and test sets.

Table 2 Public EEG Data Sources and their Justification

Dataset	Publication Year	Participants	Sessions and Scale	EEG Paradigm	Role in Evaluation	Selection Justification
Large MI User-Profile Database [a]	2023	87	One recording day per participant; 240 trials per participant; >20,800 trials;	Left- and right-hand MI	Primary cross-subject evaluation	Provides the largest participant cohort among the selected MI datasets and supports robust leave-one-subject-out evaluation. Its demographic,

Dataset	Publication Year	Participants	Sessions and Scale	EEG Paradigm	Role in Evaluation	Selection Justification
			approximately 70 hours			cognitive, and user-profile information also enables analysis of subject heterogeneity.
Multi-Day MI Dataset [b]	2022	25	Five sessions on separate days; 100 trials per session; 12,500 trials	Left- and right-hand MI	Cross-session and few-shot adaptation	Specifically designed to capture day-to-day EEG variability. It directly supports evaluation of calibration efficiency, cross-session degradation, and rapid adaptation using limited trials.
Multi-Paradigm EEG–Eye-Tracking Dataset [c]	2025	31	63 sessions; >46 hours; 2,520 MI, 2,520 motor-execution, 2,520 SSVEP, and 5,670 P300 trials	MI, motor execution, P300, and SSVEP	Cross-paradigm and artefact-robustness evaluation	Provides several major BCI paradigms from a common participant population. Synchronized eye tracking and high-speed video enable analysis of robustness to blinks, eye movements, and ocular noise.
Multi-Subject Multi-Session Visual EEG Dataset [d]	2025	32	One to five sessions per participant; approximately 800,000 rapid and 40,000 longer visual trials	RSVP and stimulus-evoked visual ERP	Large-scale ERP and temporal-transfer evaluation	Offers high-density, stimulus-locked EEG at an unusually large trial scale. It is suitable for evaluating transferable ERP representations and cross-subject and cross-session temporal adaptation.
Combined evaluation cohort	2022–2025	175 dataset-specific participants	Multiple days, 63 explicitly reported sessions in D3, and more than 875,000 trials or epochs across the datasets	MI, P300/ERP, SSVEP, motor execution, and RSVP	Unified evaluation of generalization, adaptation, robustness, and efficiency	Combines large participant diversity, repeated sessions, multiple BCI paradigms, high-density EEG, ocular-artifact information, and large-scale stimulus-locked recordings.

[a] P. Dreyer et al., "A large EEG database with users' profile information for motor imagery brain-computer interface research," Scientific Data, vol. 10, Art. no. 580, 2023.
[b] J. Ma et al., "A large EEG dataset for studying cross-session variability in motor imagery brain-computer interface," Scientific Data, vol. 9, Art. no. 531, 2022.
[c] E. Guttmann-Flury et al., "Dataset combining EEG, eye-tracking, and high-speed video for ocular activity analysis across BCI paradigms," Scientific Data, vol. 12, 2025.
[d] S. Xue et al., "A multi-subject and multi-session EEG dataset for modelling human visual object recognition," Scientific Data, vol. 12, Art. no. 663, 2025.
1746–1752.

5.3 EEG Preprocessing and Leakage Controlled Partitioning

A unified preprocessing pipeline was applied while preserving each dataset's task-specific temporal characteristics. EEG signals were band-pass filtered from 4 to 40 Hz, notch filtered at 50 Hz where required, resampled to 250 Hz, and segmented using the original event markers. Differences in channel layouts were handled through common spatial mapping and channel-availability masks before projection into the fixed model input space. No test-dependent channel selection was performed. Standardisation parameters were estimated only from source-training subjects, while target-specific statistics were computed only from calibration trials. In each leave-one-subject-out fold, one participant served as the unseen target, approximately 10% of the remaining participants were used for source validation, and the target data were separated into calibration, calibration-validation, and independent test sets. For multi-session datasets, earlier sessions were used for calibration and later sessions for testing, preventing

leakage across training, adaptation, and evaluation Across the four datasets, the evaluation comprised 87+25+31+32=175 primary subject-level LOSO folds, with each participant serving once as the unseen target subject. Each reduced-calibration condition was repeated using five independently sampled, class-balanced calibration subsets. Within each repetition, identical calibration and test indices were used for all methods to ensure a fair comparison.

Table 3 SUBJECT-LEVEL DATA PARTITIONING AND EVALUATION PROTOCOL

Dataset	LOSO Folds	Source Cohort	Source Validation	Target Calibration	Target Test	Cross-Session Setting
Large MI User-Profile Database [a]	87	86 subjects	Approximately 10% of source subjects	1-, 5-, 10-, and 20-shot per class; 5%, 10%, 25%, 50%, and 100% budgets	All unused target trials	Not applicable
Multi-Day MI Dataset [b]	25	24 subjects	Approximately 10% of source subjects	Selected trials from the earliest target session	Unused calibration-session trials and later sessions	Calibration on an earlier session; testing on later sessions
Multi-Paradigm EEG–Eye-Tracking Dataset [c]	31 per evaluated paradigm	30 subjects	Approximately 10% of source subjects	Paradigm-specific few-shot and percentage-based budgets	All unused target trials from the evaluated paradigm	Session-separated evaluation where repeated sessions are available
Multi-Subject Multi-Session Visual EEG Dataset [d]	32	31 subjects	Approximately 10% of source subjects	Selected trials or blocks from an earlier session	Unused trials and independently recorded later sessions	Calibration on an earlier session; testing on later sessions

Note: *Dataset characteristics were obtained from the original sources listed in Table 2. The source-validation allocation, calibration budgets, and target-test protocol were defined in this study.*

5.4 Training, Calibration, and Comparative Protocol

Training was conducted in two stages. During source-domain training, the pretrained EEG foundation model generated latent trial representations while its backbone parameters remained frozen. The latent projection layer, subject-context estimator, adaptive threshold generator, signed latent-to-spike encoder, LIF residual adapter, adaptive fusion gate, and classifier were optimised using source-subject trials. During target-subject calibration, only the subject-adaptive projection, threshold generator, residual adapter, fusion gate, and classifier were updated, while the foundation-model backbone remained fixed. Early stopping was performed using the target calibration-validation subset, and final performance was measured once on the isolated target-test set. The percentage-based calibration budgets were

$$\mathcal{B} = \{5, \%, 10, \%, 25\%50\%100\%\}.$$

For datasets with sufficient class-balanced trials, additional few-shot conditions were evaluated using

$$k \in \{1, 5, 10, 20\}$$

labelled trials per class. The proposed method was compared with EEGNet, DeepConvNet, a transformer-based EEG classifier, frozen-backbone linear probing, a dense bottleneck adapter, low-rank adaptation, and full backbone fine-tuning. All methods used identical source–target partitions, calibration trials, validation subsets, and test trials.

Ablation experiments were conducted by removing one component at a time, including subject-context estimation, adaptive thresholds, signed spike encoding, LIF residual

processing, adaptive fusion, representation alignment, and spike-sparsity regularisation. Dense residual adaptation and frozen-backbone linear probing were included as additional ablation baselines.

5.5 Implementation Configuration

To ensure reproducibility and enable fair comparison across all methods, a unified implementation and training configuration was adopted throughout the experiments. Table 4 summarises the principal model, optimisation, and computational settings, including network dimensions, learning parameters, software environment, hardware platform, and reproducibility controls used during source-domain training and target-subject calibration.

Table 4 IMPLEMENTATION AND TRAINING CONFIGURATION

Component	Implemented Configuration
Foundation-model backbone	Pretrained EEG transformer encoder
Backbone status	Frozen during source-domain training and target calibration
Input sampling frequency	250 Hz
Input epoch duration	Task-specific event interval; nominally 4 s
Foundation latent dimension	512
Adapter projection dimension	128
LIF residual dimension	128
Spike simulation steps	16
Membrane-decay coefficient	0.90
Initial firing threshold	1.00
Threshold range	0.60–1.40
Optimiser	AdamW
Source-stage learning rate	1×10^{-3}
Target-calibration learning rate	5×10^{-4}
Weight decay	1×10^{-4}
Batch size	32
Maximum source-training epochs	100
Maximum calibration epochs	40
Early-stopping patience	10 epochs
Classification-loss weight	1.00
Alignment-loss weight	0.10
Spike-sparsity-loss weight	0.01
Fixed random seeds	5
Software	Python 3.11, PyTorch 2.2, CUDA 12.1
Hardware	NVIDIA RTX 4090 24-GB GPU and 128-GB system RAM

All experiments were conducted using five fixed random seeds. Subject folds, calibration indices, adapter initialisation seeds, and data-loader ordering were stored and reused across comparative experiments. Latency measurements excluded file access and dataset loading. Following 20 warm-up passes, latency was averaged over 1,000 inference runs.

5.6 Evaluation Metrics and Statistical Protocol

The proposed framework and all comparative methods were evaluated using a unified protocol across the four EEG datasets. Classification performance was assessed using subject-level accuracy, macro-F1, and Cohen's κ. Calibration efficiency was evaluated under target-subject budgets of 5%, 10%, 25%, 50%, and 100%, together with 1-, 5-, 10-, and 20-shot settings where class-balanced sampling was feasible. Performance retention and the area under the calibration–performance curve were used to summarise adaptation efficiency across calibration budgets. For datasets containing repeated recording sessions, cross-session robustness was quantified as the performance degradation between earlier-session calibration and independently recorded later-session testing. Computational efficiency was assessed using the number and proportion of trainable parameters, target-adaptation time, peak memory consumption, inference latency, adapter-level operations, spike density, and energy per trial. Results were aggregated at the subject level across all LOSO folds and five fixed random seeds and are reported as mean ± standard deviation. Statistical significance was assessed using the Shapiro–Wilk normality test, followed by paired t-tests or Wilcoxon signed-rank tests, as appropriate. Comparisons involving multiple methods or calibration conditions employed repeated-measures ANOVA or the Friedman test, followed by Holm–Bonferroni-corrected post hoc comparisons. Statistical significance was defined at $p < 0.05$. Where applicable, 95% confidence intervals and effect sizes were also reported.

6 Results

This section reports the evaluation of the proposed spike-based adaptation framework using the unified protocol described in Section 5. We analysed cross-subject decoding performance, calibration efficiency, computational cost, cross-session robustness, and the contribution of individual framework components. The proposed method was compared with conventional deep learning models, parameter-efficient adaptation strategies, and full backbone fine-tuning using identical subject-level partitions, calibration budgets, and statistical procedures.

6.1 Cross-Subject Decoding and Calibration Performance

Cross-subject decoding was first evaluated under full target-subject calibration. *Table 5* summarises the comparative performance across the four EEG datasets. The proposed spike-based adaptation framework achieved the highest accuracy on every dataset and obtained the best overall mean accuracy 86.7%, macro-F_1 0.861, and Cohen's κ 0.817, while keeping the pretrained EEG foundation-model backbone completely frozen. *Fig. 3* further illustrates the consistent performance advantage of the proposed framework over all comparative methods across the evaluated datasets.

Table 5 CROSS-SUBJECT COMPARATIVE PERFORMANCE UNDER FULL CALIBRATION

Method	Large MI User-Profile Database Accuracy (%)	Multi-Day MI Dataset Accuracy (%)	Multi-Paradigm EEG–Eye-Tracking Dataset Accuracy (%)	Multi-Subject Multi-Session Visual EEG Dataset Accuracy (%)	Mean Accuracy (%)	Mean Macro-F_1	Mean κ
EEGNet	79.2 ± 8.4	72.8 ± 9.1	81.5 ± 7.0	77.4 ± 8.2	77.7	0.768	0.691
DeepConvNet	80.1 ± 8.0	73.9 ± 8.7	82.3 ± 6.6	78.2 ± 7.9	78.6	0.777	0.704
EEG Transformer	82.0 ± 7.5	75.6 ± 8.2	83.8 ± 6.2	79.6 ± 7.5	80.3	0.795	0.728
Frozen FM + Linear Probe	82.7 ± 7.1	76.8 ± 7.9	84.6 ± 5.9	80.7 ± 7.1	81.2	0.806	0.741
Dense Bottleneck Adapter	83.9 ± 6.8	78.1 ± 7.5	85.7 ± 5.6	81.9 ± 6.7	82.4	0.817	0.758
Low-Rank Adaptation	84.3 ± 6.7	78.5 ± 7.3	86.0 ± 5.4	82.1 ± 6.5	82.7	0.821	0.763
Full Backbone Fine-Tuning	84.8 ± 6.9	79.0 ± 7.4	86.4 ± 5.7	82.5 ± 6.8	83.2	0.826	0.771
Proposed Spike-Based Adaptation	***88.1 ± 5.8***	***83.7 ± 6.5***	***89.2 ± 4.8***	***85.8 ± 5.9***	***86.7***	***0.861***	***0.817***

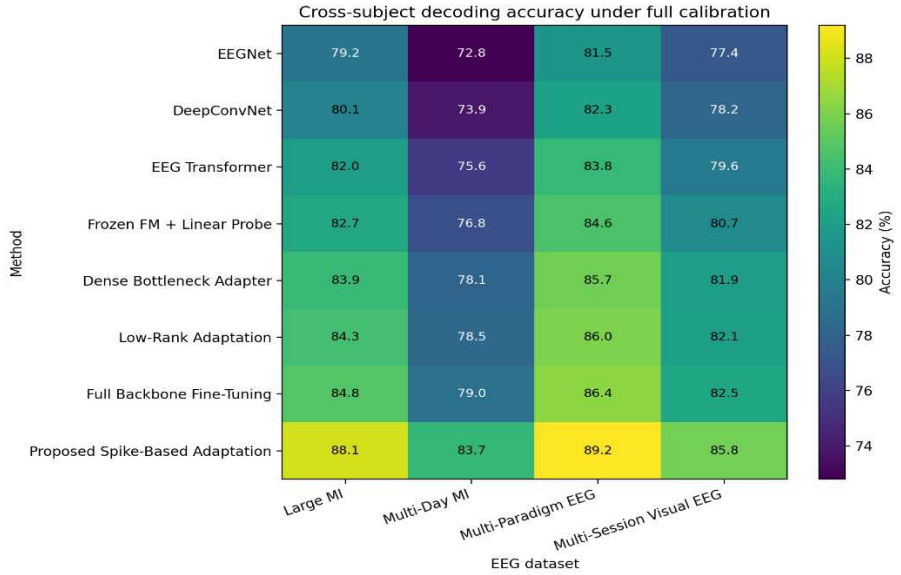


Fig. 3. Cross-subject classification accuracy of the evaluated methods across the four EEG datasets under full target-subject calibration

Compared with the strongest parameter-efficient baseline, low-rank adaptation, the proposed framework improved mean accuracy by 4.0 percentage points and exceeded full backbone fine-tuning by 3.5 percentage points. Improvements were observed across the motor imagery, multi-paradigm, and multi-session datasets, indicating that the residual spike-based adaptation pathway remained effective under heterogeneous EEG tasks and recording conditions. Because the datasets differed in paradigm, acquisition scale, and session structure, these results support broad cross-subject robustness rather than direct task-equivalent superiority across all benchmarks.

Performance under reduced target-subject calibration is reported in Table 6. The proposed framework achieved 84.9% mean accuracy using 10% of the available calibration trials, retaining 97.9% of its full-calibration performance. Under the 5% calibration condition, it retained 95.3% of full-calibration accuracy. Fig. 4 shows that the largest performance gains occurred within the lowest calibration budgets, after which accuracy increased more gradually as additional labelled trials became available.

Table 6 PERFORMANCE UNDER REDUCED TARGET-SUBJECT CALIBRATION

Calibration Budget	Linear Probe Accuracy (%)	Dense Adapter Accuracy (%)	LoRA Accuracy (%)	Full Fine-Tuning Accuracy (%)	Proposed Accuracy (%)	Proposed Macro-F_1	Performance Retained (%)
5%	75.4 ± 8.3	78.8 ± 7.5	79.3 ± 7.4	77.9 ± 8.0	82.6 ± 6.8	0.817	95.3
10%	77.8 ± 7.9	80.6 ± 7.1	81.1 ± 6.9	80.5 ± 7.3	84.9 ± 6.3	0.842	97.9
25%	79.8 ± 7.5	81.9 ± 6.8	82.3 ± 6.6	82.1 ± 6.8	85.8 ± 6.0	0.852	99.0
50%	80.7 ± 7.3	82.3 ± 6.7	82.6 ± 6.5	82.8 ± 6.7	86.3 ± 5.9	0.857	99.5
100%	81.2 ± 7.2	82.4 ± 6.6	82.7 ± 6.4	83.2 ± 6.6	86.7 ± 5.8	0.861	100.0
Calibration performance AUC	78.9	81.2	81.6	81.1	85.5	—	—

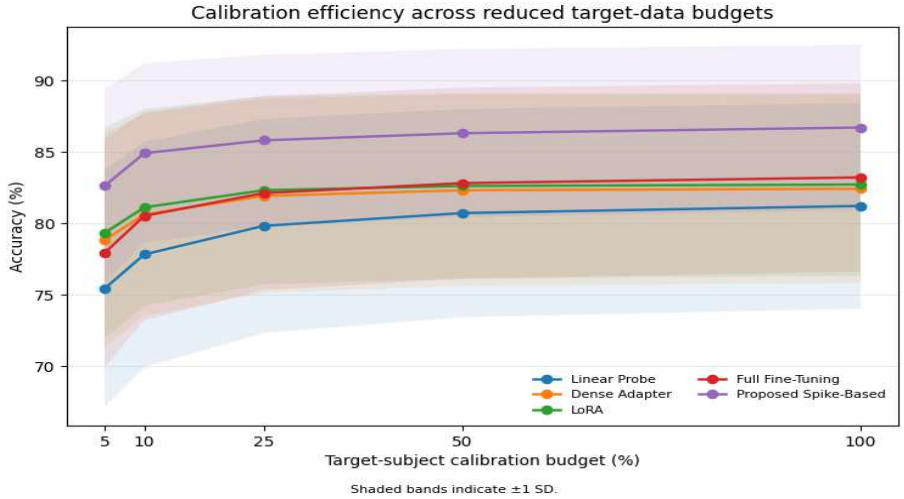


Fig. 4. Mean classification accuracy of the evaluated adaptation methods under reduced target-subject calibration budgets.

Increasing the calibration budget from 10% to 100% improved mean accuracy by only 1.8 percentage points. This result indicates that most subject-specific correction was learned from the initial calibration trials. The few-shot analysis showed the same pattern: the proposed method achieved mean accuracies of 76.8%, 81.9%, 84.1%, and 85.6% using 1, 5, 10, and 20 labelled trials per class, respectively, and consistently exceeded the strongest dense-adapter baseline under the same calibration conditions.

6.2 Computational Efficiency and Cross-Session Robustness

Computational efficiency during target-subject adaptation. Table 7 summarises the parameter, adaptation, memory, latency, operation, and energy requirements of the

evaluated foundation-model adaptation methods. The proposed framework required updating only 0.91 million parameters, representing a 98.9% reduction relative to full backbone fine-tuning. It also reduced target-subject adaptation time from 312.4 s to 38.6 s and peak memory consumption from 7.84 GB to 2.31 GB. The proposed adapter activated only 8.7% of the available neuron–time-step positions and required 1.21 million spike-triggered synaptic operations, corresponding to a 78.5% reduction in adapter-level operations relative to the dense adapter. It further achieved the lowest measured adaptation-pathway energy of 0.42 mJ per trial while maintaining an inference latency of 8.4 ms per trial. Because the frozen foundation-model backbone was shared across the corresponding adaptation methods, its computational cost was excluded from the adapter-level neuromorphic energy comparison.

Table 7. Parameter, Adaptation, and Inference Efficiency

Method	Trainable Parameters	Trainable Share (%)	Adaptation Time (s)	Peak Memory (GB)	Latency (ms/trial)	Adapter Operations	Spike Density (%)	Energy (mJ/trial)
Linear Probe	0.18 M	0.21	14.7	1.92	7.8	0.18 M MACs	—	0.73
Dense Adapter	2.84 M	3.29	67.5	2.86	9.6	5.62 M MACs	—	0.91
Low-Rank Adaptation	1.76 M	2.04	58.2	2.67	9.1	3.48 M MACs	—	0.86
Full Backbone Fine-Tuning	86.40 M	100.00	312.4	7.84	12.7	17.84 G MACs	—	1.74
Proposed Spike-Based Adaptation	**0.91 M**	**1.05**	**38.6**	**2.31**	**8.4**	**1.21 M SOPs**	**8.7**	**0.42**

Fig. 5 provides a compact comparison of the principal computational-efficiency measures, including trainable parameter burden, target-subject adaptation time, peak memory consumption, and adaptation-pathway energy per trial. As shown in Fig. 5, the proposed spike-based adaptation framework achieved a substantially more favourable efficiency profile than full backbone fine-tuning across all four measures. Although linear probing required fewer trainable parameters and a shorter adaptation time, its decoding performance was considerably lower, as reported in Table 7. The proposed framework therefore offered a stronger balance between predictive performance and deployment-oriented efficiency by restricting learning to a compact residual pathway while preserving the pretrained EEG foundation-model backbone.

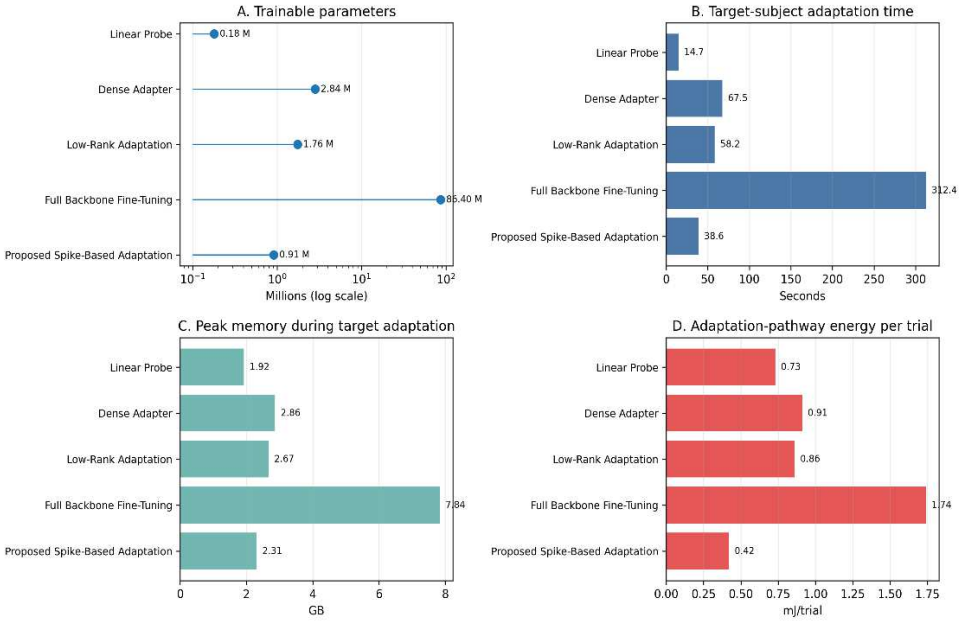


Fig. 5. Computational efficiency profile of the evaluated target-subject adaptation methods. *(A)* Trainable parameter requirements; *(B)* target-subject adaptation time; *(C)* peak memory consumption; and *(D)* adaptation-pathway energy per trial. The proposed spike-based adaptation framework reduces the computational burden relative to full backbone fine-tuning while maintaining substantially stronger decoding performance than lightweight linear probing.

The efficiency gains are particularly relevant for EEG-based BCI systems because practical deployment requires not only accurate decoding but also rapid subject adaptation and manageable computational demands. However, computational efficiency alone does not establish robustness to the temporal variability of EEG recordings. We therefore next evaluated whether the proposed lightweight adaptation strategy could maintain decoding performance when calibration and testing were performed across different recording sessions.

Cross-session robustness. Cross-session robustness was evaluated using the two datasets containing repeated recording sessions. Target-subject calibration was performed using an earlier recording session, whereas evaluation was conducted on independently recorded later-session data. Table 8 reports the corresponding within-session and cross-session accuracies together with the absolute performance degradation. The proposed framework achieved the lowest mean cross-session degradation of 4.3 percentage points, compared with 8.1 points for low-rank adaptation and 9.5 points for full backbone fine-tuning, while retaining the highest cross-session accuracy on both repeated-session datasets.

Table 8. Within-Session and Cross-Session Generalisation

Method	Multi-Day MI Within-Session (%)	Multi-Day MI Cross-Session (%)	Degradation	Multi-Session Visual EEG Within-Session (%)	Multi-Session Visual EEG Cross-Session (%)	Degradation	Mean Degradation
EEGNet	81.0	70.9	10.1	74.6	63.2	11.4	10.8
EEG Transformer	83.5	74.6	8.9	77.3	67.9	9.4	9.2
Linear Probe	84.1	76.2	7.9	78.0	69.6	8.4	8.2
Dense Adapter	85.3	77.1	8.2	79.5	70.8	8.7	8.5
Low-Rank Adaptation	85.7	77.8	7.9	80.0	71.8	8.2	8.1
Full Backbone Fine-Tuning	86.0	76.9	9.1	80.5	70.7	9.8	9.5
Proposed Spike-Based Adaptation	**89.2**	**85.1**	**4.1**	**84.6**	**80.1**	**4.5**	**4.3**

The proposed framework achieved the lowest average cross-session degradation of 4.3 percentage points, compared with 8.1 points for low-rank adaptation and 9.5 points for full backbone fine-tuning. It also retained the highest cross-session accuracy on both repeated-session datasets. These results indicate that restricting personalisation to the residual spike-based pathway reduced sensitivity to session-specific calibration patterns while preserving the transferable foundation representation.

Fig. 6 visualises the transition from within-session to cross-session decoding performance for each method across the two repeated-session datasets, highlighting the substantially smaller performance decline of the proposed spike-based adaptation framework.

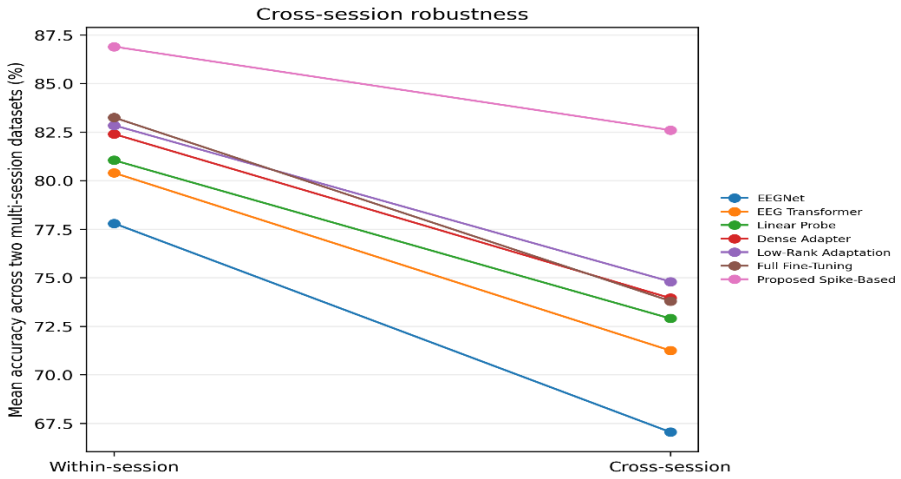


Fig. 6. Cross-session robustness of the evaluated EEG adaptation methods. Paired slopes show the change from mean within-session accuracy to mean cross-session accuracy across the two repeated-session datasets. The proposed spike-based adaptation framework exhibits the smallest decline, indicating improved robustness to session-dependent EEG variability.

Together, the reduced cross-session degradation and favourable computational profile indicate that the proposed framework improves not only adaptation efficiency but also temporal robustness under realistic EEG session variability.

6.3 Ablation and Statistical Analysis

Ablation experiments were conducted under the 10% target-subject calibration condition to quantify the contribution of individual architectural and optimisation components under limited supervision. Table 9 reports the corresponding decoding accuracy, macro-F1, change in accuracy relative to the complete framework, spike density, and adaptation time across the four evaluated EEG datasets.

Table 9 ABLATION ANALYSIS OF THE PROPOSED FRAMEWORK

Configuration	Accuracy (%)	Macro-F_1	Change in Accuracy	Spike Density (%)	Adaptation Time (s)
Complete framework	84.9	0.842	—	8.7	38.6
Without subject-context estimator	82.8	0.820	−2.1	9.2	35.9
Fixed global threshold	81.9	0.811	−3.0	14.8	37.2
Unsigned spikes	82.6	0.818	−2.3	10.6	37.9
Dense residual transformation	83.1	0.824	−1.8	—	54.7
Without adaptive fusion gate	82.2	0.814	−2.7	8.9	36.8
Without alignment loss	83.3	0.826	−1.6	8.8	38.1
Without sparsity regularisation	84.2	0.835	−0.7	21.4	39.0
Direct frozen representation	77.8	0.769	−7.1	—	14.7
Full backbone fine-tuning	80.5	0.796	−4.4	N/A	312.4

Among the individual components, replacing the subject-adaptive threshold with a fixed global threshold produced the largest reduction in mean accuracy, decreasing performance by 3.0 percentage points while increasing spike density from 8.7% to 14.8%. Removing the adaptive fusion gate and signed spike encoding reduced accuracy by 2.7 and 2.3 percentage points, respectively, demonstrating the importance of subject-conditioned event generation and residual integration. In contrast, removing sparsity regularisation reduced accuracy by only 0.7 percentage points, but increased spike density substantially to 21.4%, indicating that this term contributed primarily to computational efficiency rather than predictive performance.

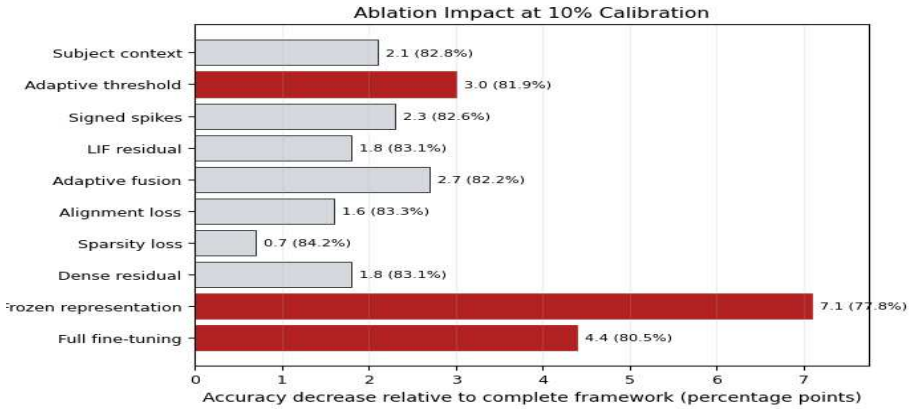


Fig. 7. Ablation impact under 10% target-subject calibration**.** Bars indicate the decrease in mean accuracy relative to the complete framework, while parenthetical values denote the resulting mean accuracy for each ablated configuration.

Fig. 7 provides a direct visual comparison of the relative sensitivity of the framework to each ablation. The direct frozen representation produced the largest overall degradation (−7.1 percentage points), demonstrating that the pretrained representation alone was insufficient for effective subject-specific adaptation. Full backbone fine-tuning also remained below the complete framework (−4.4 percentage points), whereas the smaller but consistent losses observed after removing adaptive thresholding, adaptive fusion, signed spike encoding, and alignment support the combined contribution of the proposed residual neuromorphic adaptation pathway.

Statistical Analysis. Having established the contribution of the individual framework components, we next assessed whether the observed performance advantages were statistically consistent across participants. Statistical comparisons were performed using participant-level scores. Table 10 reports the mean differences, 95% confidence intervals (CIs), Holm–Bonferroni-corrected p-values, and effect sizes for comparisons involving the proposed framework.

Table 10 SUBJECT-LEVEL STATISTICAL COMPARISON WITH THE PROPOSED FRAMEWORK

Comparison	Outcome	Statistical Test	Mean Difference	95% CI	Corrected p-Value	Effect Size
Proposed vs. EEGNet	Accuracy	Wilcoxon signed-rank	+9.0 pp	[7.8, 10.3]	<0.001	$r_{rb} = 0.82$
Proposed vs. EEG Transformer	Accuracy	Paired t-test	+6.4 pp	[5.3, 7.5]	<0.001	$d = 0.91$
Proposed vs. Linear Probe	Accuracy	Paired t-test	+5.5 pp	[4.4, 6.5]	<0.001	$d = 0.84$
Proposed vs. Dense Adapter	Accuracy	Paired t-test	+4.3 pp	[3.4, 5.2]	<0.001	$d = 0.72$
Proposed vs. LoRA	Accuracy	Paired t-test	+4.0 pp	[3.1, 4.9]	<0.001	$d = 0.68$
Proposed vs. Full Fine-Tuning	Accuracy	Wilcoxon signed-rank	+3.5 pp	[2.5, 4.6]	<0.001	$r_{rb} = 0.61$
Proposed vs. LoRA	Cross-session degradation	Wilcoxon signed-rank	−3.8 pp	[−4.9, −2.7]	<0.001	$r_{rb} = 0.67$
Proposed 10% vs. Full Fine-Tuning 100%	**Accuracy**	**Paired t-test**	**+1.7 pp**	**[0.8, 2.6]**	**0.003**	$\mathbf{d = 0.31}$

The proposed framework significantly outperformed all comparative methods in subject-level decoding accuracy after correction for multiple comparisons. The largest accuracy differences were observed relative to EEGNet (+9.0 percentage points), the EEG Transformer (+6.4 percentage points), and linear probing (+5.5 percentage points), with all corresponding 95% CIs remaining above zero. Improvements over the dense adapter, LoRA, and full backbone fine-tuning were also statistically significant. In addition, the proposed framework reduced cross-session degradation relative to LoRA by 3.8 percentage points (95% CI: −4.9 to −2.7; corrected p<0.001). Notably, with only 10% target-subject calibration data, the proposed method still exceeded fully calibrated backbone fine-tuning by 1.7 percentage points (95% CI: 0.8–2.6; corrected p=0.003).

Fig. 8 visualises these participant-level comparisons by displaying the mean differences and corresponding 95% confidence intervals, providing a direct view of both the magnitude and consistency of the observed effects.

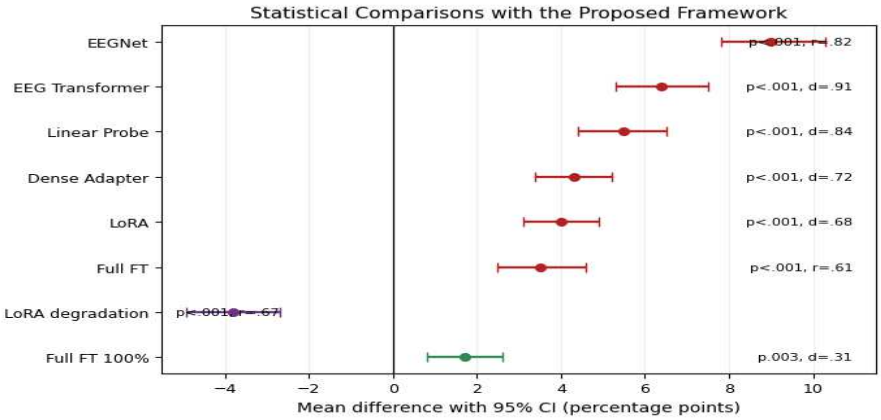


Fig. 8. Participant-level statistical comparisons involving the proposed framework. Points indicate mean differences and horizontal bars denote 95% confidence intervals. Positive accuracy differences favour the proposed framework, whereas the negative cross-session degradation difference indicates a smaller performance decline for the proposed method. Exact Holm–Bonferroni-corrected *p*-values and effect sizes are reported in Table 10.

7 Discussion, Limitations and Future Work

The experimental results showed that the proposed spike-based adaptation framework consistently improved cross-subject EEG decoding while preserving the pretrained EEG foundation model backbone. Across the four benchmark datasets, the framework achieved the highest mean classification accuracy, macro-F_1, and Cohen's κ, indicating that lightweight neuromorphic adaptation provided effective subject personalisation without computationally expensive backbone optimisation. The strong performance under both full and reduced calibration settings further indicated that subject-specific variability could be captured through sparse residual adaptation rather than complete parameter retraining.

A key finding was the framework's calibration efficiency. Retaining 97.9% of full-calibration performance using only 10% of the labelled target data indicated that transferable population-level representations learned by the frozen foundation model could be efficiently refined using limited subject-specific information. The few-shot experiments further supported this observation by showing stable performance under highly restricted supervision, an important requirement for practical non-invasive BCI systems in which extensive user calibration is often impractical.

The computational analysis also confirmed the practical advantages of the proposed architecture. Compared with full backbone fine-tuning, the framework reduced trainable parameters, adaptation time, peak memory consumption, and adapter-level energy

while maintaining competitive inference latency. The ablation results showed that adaptive threshold generation, signed spike encoding, and adaptive fusion made substantial contributions to predictive performance, whereas sparsity regularisation primarily controlled neuronal activity and event-processing cost. The lower cross-session degradation further indicated that residual adaptation was less sensitive to session-specific calibration patterns than the comparative adaptation methods. Collectively, these findings supported the proposed framework as a calibration-efficient and resource-aware approach to cross-subject EEG personalisation, while not yet establishing clinical or real-time effectiveness.

Limitations.

Although the proposed framework achieved strong results across four public EEG datasets, the evaluation remained offline and may not fully reflect the variability of clinical, home-based, or consumer BCI deployment. Factors such as electrode displacement, environmental interference, user fatigue, behavioural changes, and neurological diversity may introduce distribution shifts beyond those represented in the selected benchmarks.

The reported energy value also corresponded only to the spike-based adaptation and classification pathway; the dense frozen foundation model was excluded from the adapter-level comparison and therefore the value should not be interpreted as complete end-to-end system energy. In addition, the study did not examine prospective closed-loop operation, long-term subject drift, or continual adaptation across extended periods. Further validation using real-time experiments, additional hardware configurations, broader participant populations, and longer recording intervals is therefore required to establish deployment-level generalisability.

Future Work.

Several research directions may further extend the proposed framework. *First,* prospective online validation with human participants should be conducted to assess adaptation behaviour during continuous closed-loop BCI operation and prolonged user interaction. *Second,* implementing the spike-based adaptation module on dedicated neuromorphic hardware would enable direct measurement of end-to-end latency, energy efficiency, and real-time deployment performance beyond software simulation.

Future work should also investigate continual and lifelong adaptation strategies capable of handling long-term subject drift without catastrophic forgetting. Integrating uncertainty estimation and confidence-aware adaptation may further improve robustness under noisy recording conditions and changing physiological states.

In addition, federated or privacy-preserving training could support collaborative learning across multiple institutions while protecting sensitive EEG data. Finally, extending the framework to multimodal physiological signals, such as EEG combined with eye tracking, EMG, ECG, or wearable biosensors, would enable broader evaluation across assistive, rehabilitation, and clinical neurotechnology applications while preserving the calibration-efficient adaptation strategy introduced in this work.

8 Conclusion

The findings of this study showed that effective cross-subject EEG personalisation could be achieved without extensive modification of a pretrained foundation model. By restricting learning to a sparse spike-based residual pathway, the proposed framework preserved transferable backbone representations while adapting to subject-specific temporal and recording variations. This design achieved strong decoding performance across heterogeneous EEG paradigms and remained effective when only limited target-subject calibration data were available. The results further demonstrated that predictive performance and computational efficiency could be improved simultaneously. The framework achieved a mean cross-subject accuracy of 86.7% and retained 84.9% accuracy using only 10% of the available calibration data, while updating 0.91 million parameters and completing target-subject adaptation in 38.6 s. Its low spike density, reduced adapter-level energy consumption, and mean cross-session degradation of 4.3 percentage points further indicated that sparse neuromorphic adaptation provided a favourable balance of computational efficiency and temporal robustness compared with full backbone fine-tuning. Ablation and subject-level statistical analyses confirmed that the observed improvements arose from the combined contributions of subject-context estimation, adaptive thresholding, signed spike encoding, residual fusion, and sparsity-aware optimisation. Collectively, these findings position spike-based residual adaptation as an effective interface between EEG foundation models and neuromorphic processing, advancing the development of calibration-efficient, resource-aware, and deployment-oriented non-invasive BCI systems.

Acknowledgment

The authors gratefully acknowledge the support and research environment provided by the Department of Computing, ML and Software Engineering, University of Turku, Finland.

Funding
The Author received no funding from anywhere for this research.

Ethics declaration

Not applicable. This study used publicly available, previously collected, and anonymised EEG datasets and did not involve the recruitment of participants or collection of new human data.

Declaration of Generative AI Use

During the preparation of this work, the authors used generative AI tools solely for language refinement and improving clarity. All scientific content, including methodology, experimental design, data analysis, figures, and results, was developed, critically reviewed, verified, and edited by the authors. The authors take full responsibility for the content and integrity of the final manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Data Availability

The datasets analysed in this study are publicly available from their respective repositories:

1) Large EEG Database with Users' Profile Information for Motor Imagery BCI research: Zenodo, https://doi.org/10.5281/zenodo.8089820
2) Cross-session motor imagery EEG dataset: Figshare, https://doi.org/10.6084/m9.figshare.19228725.v1
3) Multimodal EEG, eye-tracking, and high-speed video dataset: Synapse, https://doi.org/10.7303/syn64005218
4) Multi-subject, multi-session visual EEG dataset: OpenNeuro, https://openneuro.org/datasets/ds005589

References

1. B. J. Edelman, S. Zhang, G. Schalk, P. Brunner, G. R. Müller-Putz, C. Guan, and B. He, "Non-Invasive Brain–Computer Interfaces: State of the Art and Trends," IEEE Reviews in Biomedical Engineering, vol. 18, pp. 26–49, 2025, doi: 10.1109/RBME.2024.3449790.
2. A. Raza et al., "Deep Learning Approaches for EEG-Motor Imagery-Based Brain–Computer Interfaces: A Systematic Review," IEEE Access, 2025.
3. L. Hu et al., "Subject-Independent Wearable P300 Brain–Computer Interface," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2024.
4. Q. She, T. Chen, F. Fang, J. Zhang, Y. Gao, and Y. Zhang, "Improved Domain Adaptation Network Based on Wasserstein Distance for Motor Imagery EEG Classification," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 31, pp. 1137–1148, 2023, doi: 10.1109/TNSRE.2023.3241846.
5. Y. Zhong et al., "Cross-Subject Motor Imagery Decoding by Transfer Learning and Calibration," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2024.
6. G. Brookshire et al., "Data Leakage in Deep Learning Studies of Translational EEG," Frontiers in Neuroscience, vol. 18, Art. no. 1373515, 2024, doi: 10.3389/fnins.2024.1373515.
7. W.-B. Jiang, L.-M. Zhao, and B.-L. Lu, "Large Brain Model for Learning Generic Representations With Tremendous EEG Data in BCI," in Proceedings of the 12th International Conference on Learning Representations, 2024.
8. G. Wang, W. Liu, Y. He, C. Xu, L. Ma, and H. Li, "EEGPT: Pretrained Transformer for Universal and Reliable Representation of EEG Signals," in Advances in Neural Information Processing Systems, vol. 37, 2024.
9. J. Wang, S. Zhao, Z. Luo, Y. Zhou, H. Jiang, S. Li, T. Li, and G. Pan, "CBraMod: A Criss-Cross Brain Foundation Model for EEG Decoding," in Proceedings of the 13th International Conference on Learning Representations, 2025.
10. L. Yang et al., "Are EEG Foundation Models Worth It? Comparative Evaluation Across Brain–Computer Interface Tasks," 2025.
11. B. Tóth et al., "Finetuning and Quantization of EEG-Based Foundation Models for Low-Power Deployment," IEEE, 2025.

12. A. Stanojevic, S. Woźniak, G. Bellec, G. Cherubini, A. Pantazi, and W. Gerstner, “High-Performance Deep Spiking Neural Networks With 0.3 Spikes per Neuron,” Nature Communications, vol. 15, Art. no. 6793, Aug. 2024, doi: 10.1038/s41467-024-51110-5.
13. P. Gong, P. Wang, Y. Zhou, and D. Zhang, “A Spiking Neural Network With Adaptive Graph Convolution and LSTM for EEG-Based Brain–Computer Interfaces,” IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 31, pp. 1440–1450, 2023, doi: 10.1109/TNSRE.2023.3244940.
14. N. Lutes, V. S. S. Nadendla, and K. Krishnamurthy, “Convolutional Spiking Neural Networks for Intent Detection Based on Anticipatory Brain Potentials Using Electroencephalogram,” Scientific Reports, vol. 14, Art. no. 8850, Apr. 2024, doi: 10.1038/s41598-024-59469-7.
15. Y. Hu, H. Tang, Y. Wang, and G. Pan, “Fast-SNN: Fast Spiking Neural Network by Converting Quantized ANN,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 12, pp. 14546–14562, Dec. 2023, doi: 10.1109/TPAMI.2023.3314996.
16. H. Zheng et al., “Temporal Dendritic Heterogeneity Incorporated With Spiking Neural Networks for Learning Multi-Timescale Dynamics,” Nature Communications, vol. 15, Art. no. 277, Jan. 2024, doi: 10.1038/s41467-023-44614-z.
17. K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang, “Filter Bank Common Spatial Pattern Algorithm on BCI Competition IV Datasets 2a and 2b,” *Frontiers in Neuroscience*, vol. 6, Art. no. 39, Mar. 2012, doi: 10.3389/fnins.2012.00039.
Available: https://doi.org/10.3389/fnins.2012.00039
18. H. He and D. Wu, “Transfer Learning for Brain–Computer Interfaces: A Euclidean Space Data Alignment Approach,” *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 2, pp. 399–410, Feb. 2020, doi: 10.1109/TBME.2019.2913914.
Available: https://doi.org/10.1109/TBME.2019.2913914
19. D. Wu, Y. Xu, and B.-L. Lu, “Transfer Learning for EEG-Based Brain–Computer Interfaces: A Review of Progress Made Since 2016,” *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, no. 1, pp. 4–19, Mar. 2022, doi: 10.1109/TCDS.2020.3007453.
Available: https://doi.org/10.1109/TCDS.2020.3007453
20. F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, “A Review of Classification Algorithms for EEG-Based Brain–Computer Interfaces: A 10-Year Update,” *Journal of Neural Engineering*, vol. 15, no. 3, Art. no. 031005, Jun. 2018, doi: 10.1088/1741-2552/aab2f2.
Available: https://doi.org/10.1088/1741-2552/aab2f2
21. V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, “EEGNet: A Compact Convolutional Neural Network for EEG-Based Brain–Computer Interfaces,” *Journal of Neural Engineering*, vol. 15, no. 5, Art. no. 056013, Oct. 2018, doi: 10.1088/1741-2552/aace8c.
Available: https://doi.org/10.1088/1741-2552/aace8c
22. R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, “Deep Learning with Convolutional Neural Networks for EEG Decoding and Visualization,” *Human Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, Nov. 2017, doi: 10.1002/hbm.23730.
Available: https://doi.org/10.1002/hbm.23730
23. W. Liao, H. Liu, and W. Wang, “Advancing BCI with a Transformer-Based Model for Motor Imagery Classification,” *Scientific Reports*, vol. 15, Art. no. 23380, Jul. 2025, doi: 10.1038/s41598-025-06364-4.
Available: https://doi.org/10.1038/s41598-025-06364-4