PSLIF: A Primary-Supplementary LIF Neuron for Spiking Neural Networks

Supplementary Material

A. Distinction between PSLIF and Multi-Compartment Neuron Models

To clearly illustrate the differences, we provide a concise comparison between PSLIF and representative multi-compartment neuron models.

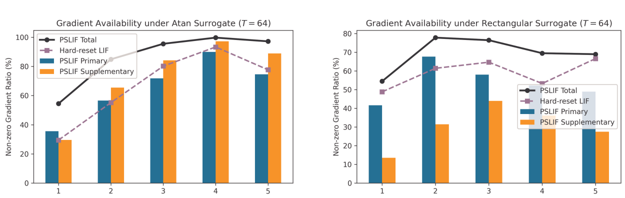
Table S1. Comparison between PSLIF and representative multi-compartment models.

Model	Method	LIF Float Ops	Fire function
LIF	✗	✓	1
TC-LIF	coupling	✗	1
LM-H	hierarchical	✗	1
CLIF	complementary	✗	1
PSLIF	interaction	✓	2

- PSLIF is, to our knowledge, the first neuron model to explicitly construct a **serial composition** of Heaviside spiking functions, using sequential firing to decouple surplus accumulation and enable an auxiliary gradient pathway—an approach not adopted in prior CLIF, TC-LIF or LM-H models.
- PSLIF achieves surplus reuse **without** introducing additional learnable parameters per neuron, preserving parameter efficiency.
- PSLIF is **topologically simpler** (only one additional state variable and no extra synaptic connections), yet still captures a form of "carry-over" dynamics that mitigates information loss during reset.

B. Empirical Evidence of Improved Gradient Propagation in PSLIF

To empirically validate the theoretical analysis of gradient propagation, we conduct ablation experiments on gradient availability using a VGG5 network with $T = 64$. We report the layer-wise ratio of non-zero gradients ($|\partial\mathcal{L}/\partial u| > 10^{-6}$ under the ATan surrogate).



(a) Atan surrogate

(b) Rect surrogate

Figure S1. Layer-wise gradient availability ($T = 64$). Bars denote primary and supplementary gradient paths in PSLIF, while lines compare the total PSLIF gradient with hard-reset LIF.

Direct temporal gradient profiling in long-horizon SNNs is technically challenging; thus, we use layer-wise non-zero gradient ratios as a practical proxy. Our theory guarantees non-zero temporal recursion paths rather than universally non-zero local gradients, and local sparsity remains expected. By preventing zero gradients along the temporal dimension from collapsing the entire gradient chain, PSLIF increases layer-wise gradient availability in practice, consistent with our theoretical analysis. Across different surrogate gradients, PSLIF consistently outperforms hard-reset LIF, indicating a structural rather than surrogate-specific improvement.

C. Supplementary Experiments on Fairness and Efficiency

To comprehensively evaluate the efficiency of PSLIF, we provide additional experimental results under different timestep settings. All latency, energy, and training cost comparisons in the main paper are conducted under the same $T = 4$ setting, which is in fact **unfavorable** to PSLIF. Despite this constraint, PSLIF still achieves lower energy consumption than CLIF.

Table S2. Comparison with multi-compartment methods ($T = 2$ for PSLIF and $T = 4$ for others)

Tasks	PSLIF(2x2)	TC-LIF	STC-LIF	LM-H
CIFAR100 (%)	79.80	77.09	78.60	79.41
Parameters (M)	12.54	12.54	12.54	12.54
Time (s/epoch)	115	180	180	165

Table S3. The SOP energy consumption

Neuron	time step	ACs	MACs	SOP Cost
ReLU	1	0	549.18M	2526.71μJ
LIF	4	62.46M	2.22M	66.46μJ
CLIF	4	57.47M	6.68M	82.48μJ
PSLIF	2x2	54.22M	2.22M	**59.05**μJ

To evaluate our PSLIF method from multiple perspectives, we supplemented the experimental data with $T = 2$ under the same network construction conditions. PSLIF still achieves 79.8% accuracy on CIFAR-100, while the training time and energy consumption are approximately halved compared with the baselines. These results demonstrate the efficiency advantage of PSLIF.

Furthermore, since PSLIF does not introduce any non-LIF modules, it can be directly implemented by reusing standard LIF hardware primitives, avoiding explicit temporal unfolding.