PING locks E rate ≈10× below COBA

exp025 · 30 May 2026

The trained networks this entry uses are produced once in the shared training hub, exp022 (Training), and reused here rather than retrained.

Abstract

Head-to-head comparison of COBA (recurrent inhibitory loop disabled) against PING (loop active) on MNIST under matched architecture and training recipe. PING locks the hidden E rate to ≈ 12 Hz while COBA runs at ≈ 181 Hz; accuracy is within a few points across the two, so the loop buys better than an order-of-magnitude per-spike economy. The rate-vs-accuracy frontier traced by sweeping the θ_u rate regulariser shows the floor is structural, not a trade-off the optimiser can navigate away.

Method

Training recipe (canonical / medium tier):

Parameter	Value
Integration timestep Δt	0.1 ms
Trial duration T	200 ms
MNIST samples (80/20 stratified split of 2000)	1600 train / 400 test (≈ 2.9% of the 70k-sample MNIST corpus)
Epochs	100

Two configurations of the same COBANet architecture, differing only in whether the E→I→E inhibitory loop is active. **COBA** (*–ei-strength 0*) disables the loop; excitatory cells drive each other but receive no structured inhibition. **PING** (*–ei-strength 1*) enables the loop, producing pyramidal-interneuron gamma (PING) oscillations at a cadence set by τ_{AMPA} and τ_{GABA}.

Architecture. $N_E = 1024$ excitatory cells, $N_I = 256$ inhibitory cells, single hidden layer. Input: 784 channels (MNIST pixels), Poisson-encoded at 25 Hz peak rate. Readout: mem-mean (time-averaged E spike vector projected through a trained linear layer W_{out}; see ar006 for the full readout specification). Dale's law enforced.

What is trainable. Only the input weights W_{in} (784 × 1024, 95% sparse) and the readout W_{out} (1024 × 10). The recurrent weights W_{ee} (fixed at zero, since Börgers-style PING needs no E→E coupling), W_{ei}, and W_{ie} are initialised once and held *requires_grad = False*. The synaptic time constants τ_{AMPA}, τ_{GABA} are module-level constants. 813k trainable parameters out of 2.4M total.

Training. Adam, lr $= 4 \times 10^{-4}$, batch size 256, gradient norm clipped to 1.0. Cross-entropy loss on 10-class MNIST (definition in ar006). Gradient stabiliser: *–v-grad-dampen 1000* (uniform scaling of per-step voltage gradients). Three seeds (42, 43, 44) for baselines, one seed (42) for sweep cells.

Recipe difference. The only parameter that differs between COBA and PING besides *–ei-strength* is the W_{in} initialisation: COBA uses mean 0.3 (std 0.03), PING uses mean 1.2 (std 0.12). PING needs stronger input drive to reliably recruit the I-loop at init.

Spike-budget regulariser. To probe the rate axis, the training loss adds a soft upper bound on per-trial spike count. For each E cell n with mean per-trial spike count $\bar{z}_n$:

$$L_{\text{rate}} = \lambda \sum_n \text{ReLU}(\bar{z}_n - \theta_u)^2,$$

with θ_u the per-cell budget (spikes/trial) and $\lambda = 10^{-3}$; only cells over budget contribute, cells under it are free, and the total loss is cross-entropy plus L_{rate}. We sweep six budgets (off, 5, 2, 1, 0.5, 0.2 spikes/trial; at $T = 200$ ms this is no penalty, 25, 10, 5, 2.5, 1 Hz), giving twelve (model, θ_u) cells.

Rate-floor decomposition. The affine law $r_E = p \cdot f_\gamma$ (exp041, exp046) factors the E rate into per-cycle participation p and gamma frequency f_γ. Both are measured at every cell: f_γ from the Welch PSD peak of the E-population trace, p via I-burst peak detection and per-(cell, cycle) spike counting (style of exp046).

Basin and landscape probes. To test basin attractivity, four PING networks are trained with W_{in} initialised at 0.05, 0.1, 0.3, and 1.2 ($\theta_u = 0.2$ from epoch 0; Figure 3). To map the loss landscape around the operating point, each network trained at the heaviest penalty ($\theta_u = 0.2$) has its W_{in} scaled by a common scalar $s \in [0.05, 3]$ at inference with all other weights frozen, metrics averaged over the test set at 24 values of s (Figures 4–5).

Results

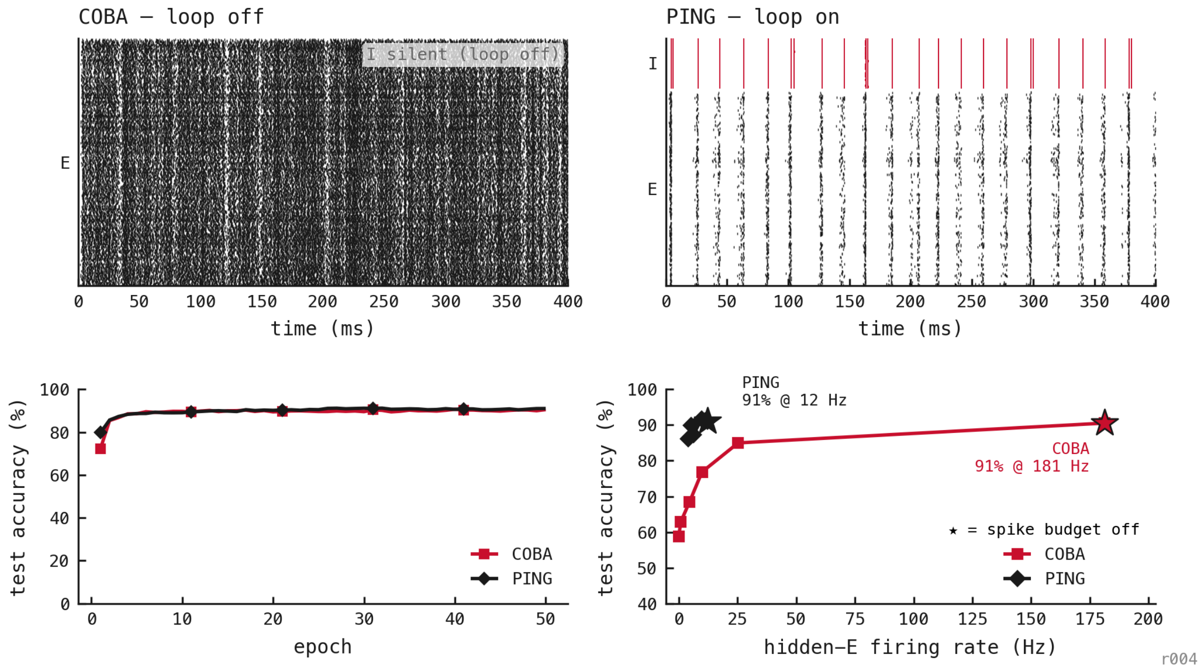

Figure 1: The headline comparison in one 2×2 frame (rasters, learning, and the accuracy–rate frontier together). **Top:** trained-baseline single-trial rasters on the same MNIST digit 0 input. COBA (*–ei-strength 0*) fires densely and asynchronously at ≈ 181 Hz (I silent, loop off); PING (*–ei-strength 1*) fires in gamma bands at ≈ 22 ms cadence (≈ 12 Hz per E cell) with synchronous I bursts (red) above E (black), on the same architecture, parameter count, and recipe, with only PING's recurrent E↔I matrices non-zero. **Bottom left:** test accuracy per epoch (mean over three seeds); both reach ≈ 91%, so both learn the task. **Bottom right:** the accuracy–rate frontier across the spike-budget penalty θ_u. PING sits up-and-left of COBA, the same accuracy at a fraction of the hidden-E rate, with the θ_u-off operating points starred and labelled (PING 91% at ≈ 12 Hz, COBA 91% at ≈ 181 Hz). COBA red, PING black; E black and I red in the rasters. *The headline of this entry: gamma gating buys an order-of-magnitude per-spike economy at matched accuracy.*

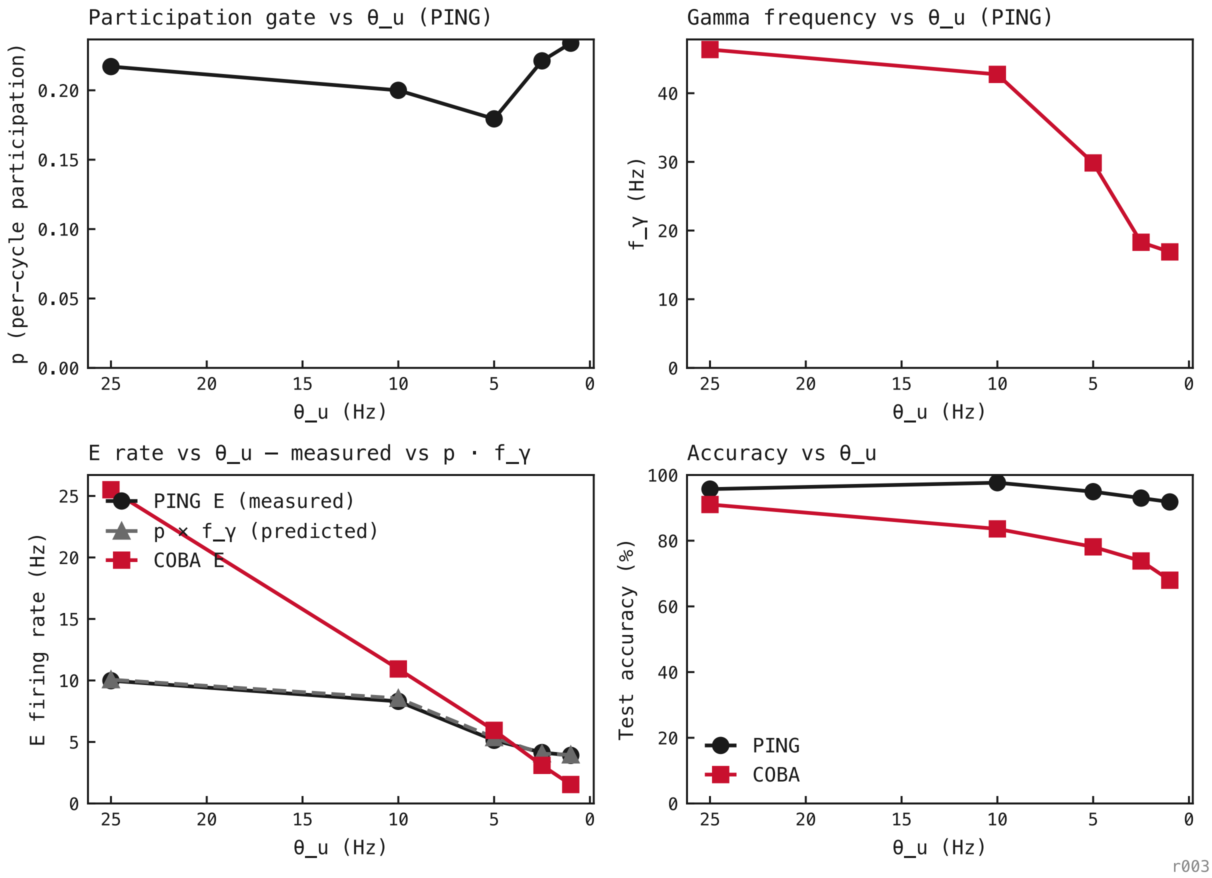

Figure 2: Five (PING, θ_u) sweep cells, 256 test trials each. $\boldsymbol{p}$ **stays in 0.18–0.23** across the entire sweep (the architecture protects the participation gate), while $\boldsymbol{f_\gamma}$ **slides from ≈ 46 Hz to ≈ 17 Hz** as the penalty tightens. The grey dashed $p \cdot f_\gamma$ curve overlays the measured E rate within 4%; the rate change is entirely in f_γ. PING accuracy holds at 92–98% the whole way; COBA's collapses from 91% to 68%.

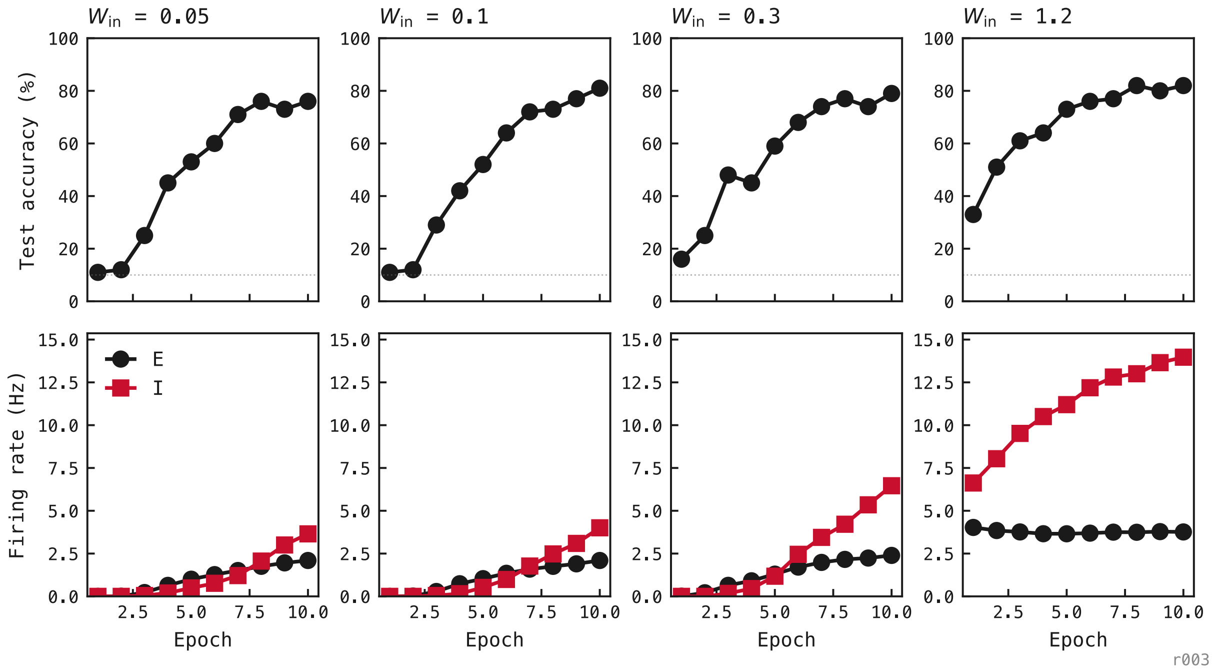

Figure 3: Per-epoch training traces from four PING networks (seed 42, $\theta_u = 0.2$ from epoch 0), one per column. Top: test accuracy. Bottom: test-set E (black) and I (red) firing rates. Recruitment is W_{in}-ordered: at $W_{\text{in}} = 0.05$ and 0.1 the I population stays silent for the first ≈ 8 epochs and engages only weakly in the final one or two; at $W_{\text{in}} = 0.3$ the loop recruits by epoch 7; at $W_{\text{in}} = 1.2$ it is active from epoch 1. The lower the input drive, the later the loop crosses the recruitment threshold. Final accuracies: 76% / 81% / 79% / 82%; final I rates: 6.9 / 2.9 / 10 / 20 Hz.

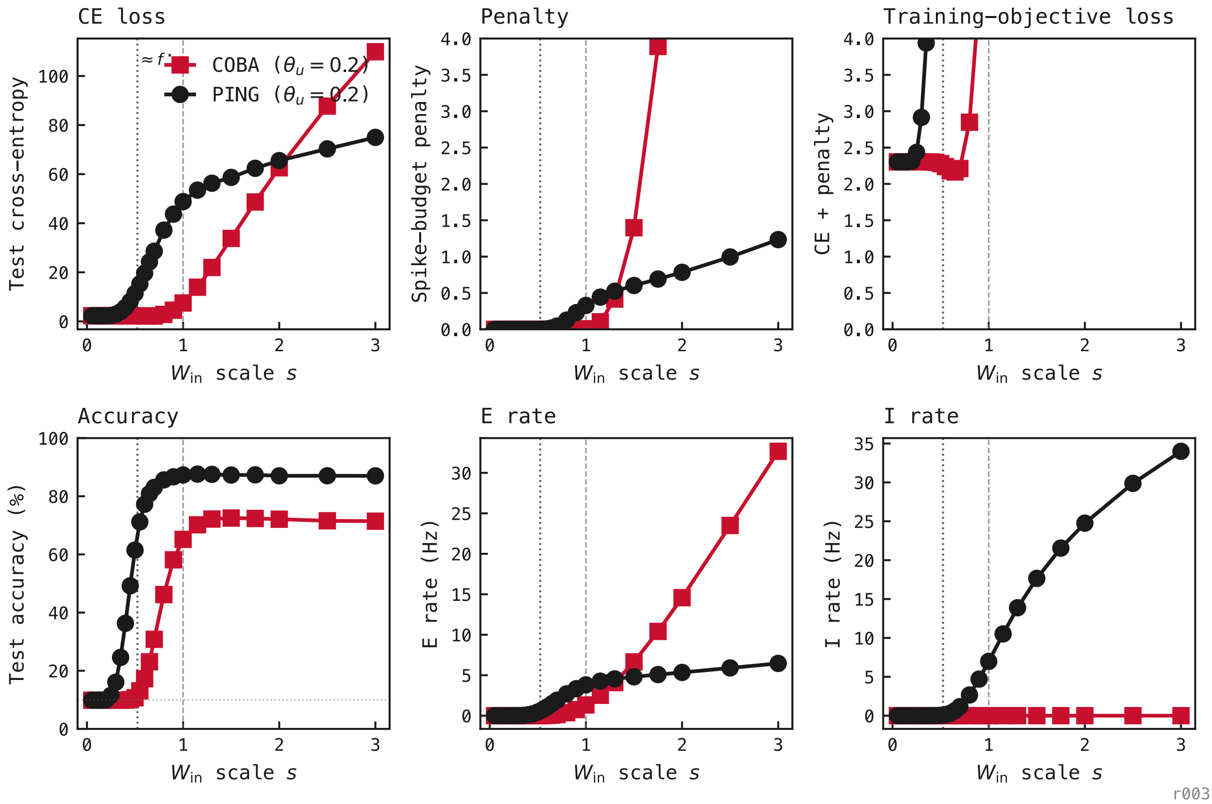

Figure 4: Inference-time W_{in} scale sweep on the two networks trained under the heaviest penalty ($\theta_u = 0.2$); every W_{in} weight multiplied by a common scalar s, all other weights frozen, 24 values of $s \in [0.05, 3]$. Top row: CE loss, spike-budget penalty L_{rate}, total objective CE + L_{rate}. Bottom row: test accuracy (chance dotted), E rate, I rate. PING black, COBA red. Vertical dashed line at $s = 1$ marks the trained operating point; dotted line marks $\approx f^*$, PING's recruitment cliff. Loss panels clipped at 4; COBA's penalty reaches ≈ 44 at $s = 3$ (rate^2 scaling).

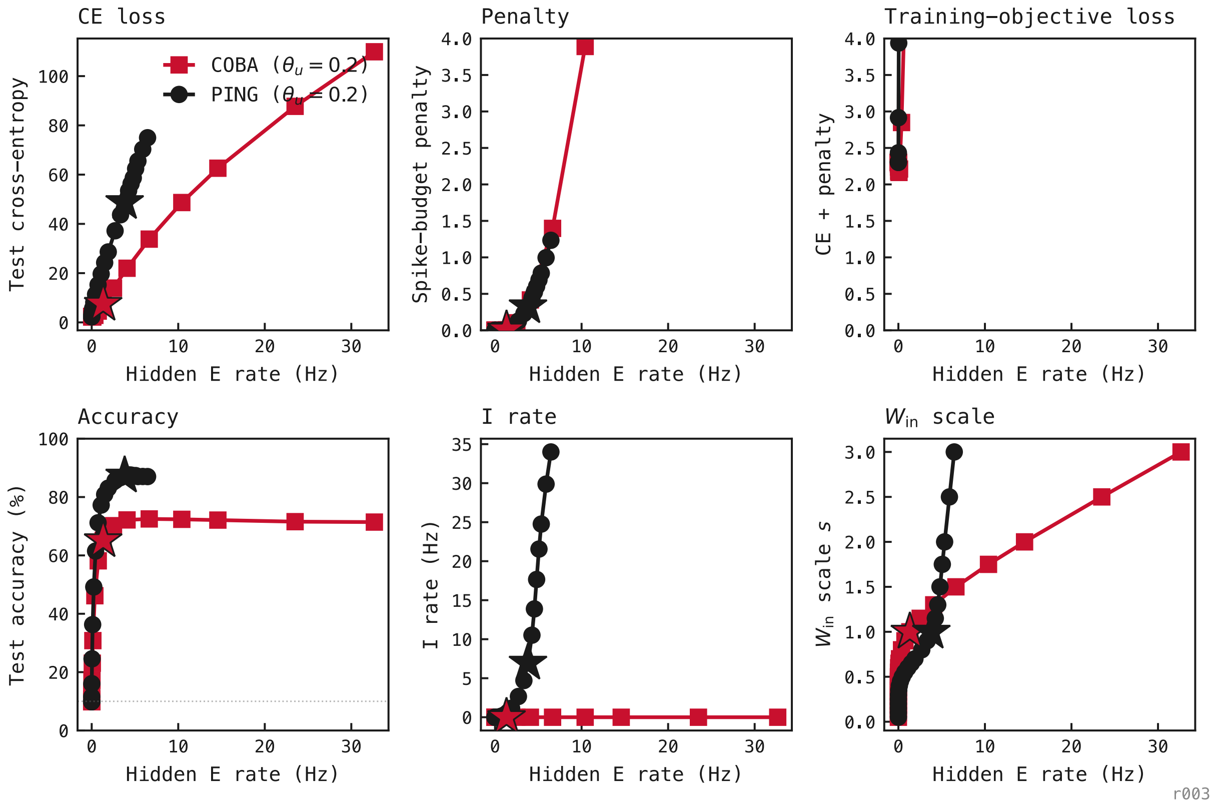

Figure 5: The same 24-point sweep as Figure 4, re-projected with hidden E rate on the x-axis. Filled stars mark each cell's trained operating point ($s = 1$): PING at $E \approx 3.8$ Hz, COBA at $E \approx 1.3$ Hz. PING's accuracy reaches its plateau ($\approx 88\%$) just past its trained operating point; COBA climbs slowly and only reaches $\approx 71\%$ even at $E \approx 33$ Hz.