Training-run guide

exp022 · 11 August 2026 · Draft

Contents

Abstract

Exp022 defines the collection's shared training runs and checkpoint bank. It specifies the motivation, parameterization, output-layer shape, and downstream consumers for six training-run types. Five run types comprise 87 trained cells spanning reference models and controlled sweeps over spike budget, inhibitory timescale, integration timestep, and recurrent initialization. TR-06 adds three PING cells trained across variable input rates with ten spiking output LIF neurons; each class logit is the corresponding neuron's output firing rate. Together, these runs provide the checkpoints used by the collection's training-dependent experiments.

Summary

All runs map Poisson-encoded pixels through 1,024 excitatory neurons to ten spiking output LIF neurons. COBA and PING use the same input-weight mean and readout initialization

scale. COBA disables recurrent E/I coupling; PING adds a $1024 \to 256 \to 1024$ E/I feedback loop and uses stronger backward-pass gradient damping to stabilize training through that recurrent path.

Unless a run-specific table says otherwise, every cell uses the following contract.

Parameter	**Default**	**Meaning**
Dataset	MNIST	784 normalized pixels encoded as independent Poisson channels
Presentation duration	200 ms	One static digit per training presentation
Integration timestep	0.1 ms	2,000 recurrent updates per presentation
Epochs	50	Training horizon for every production cell
Seeds	42, 43, 44	Three independently initialized cells per configuration
Minibatch	256	Presentations per optimization step
Optimizer learning rate	4×10^{-4}	Shared learning rate
Input population	784	One channel per image pixel
Excitatory population	1,024	Learned stimulus representation
Inhibitory population	256	PING feedback population; silent when E/I coupling is disabled
τ_{AMPA}	2 ms	Fixed excitatory synaptic decay
τ_{GABA}	6 ms	Default inhibitory decay at the gamma operating point
Input sparsity	0.95	Sparsity of the input projection
Input-weight mean	0.9	Shared by COBA and PING so input initialization is not an architecture-specific factor
Readout initialization scale	225	Shared by COBA and PING so the initial classifier scale is matched
E/I loop strength	COBA: 0; PING: 1	The forward architectural difference under comparison
Gradient damping	COBA: 1; PING: 1,000	Backward-pass stabilization for the recurrent PING loop; it does not change forward dynamics
Surrogate slope	1	Spike-gradient surrogate parameter
Default readout	`mem-mean`	A **spiking** $1024 \to 10$ output-LIF layer. Its neurons emit spikes and reset, but classification logits are their membrane voltages averaged over time—not their spike counts

Parameter	Default	Meaning
Stored projection shapes	784×1024; 1024×256; 256×1024; 1024×10	Input→E, E→I, I→E, and E→class, in source-to-destination orientation

---

Training-run guide

TR-01 — Canonical full-data reference

The full-data COBA and PING cells are used for headline accuracy. They train on all pooled MNIST with no spike-budget penalty, so the comparison is not affected by the smaller dataset or regularization used in the sweeps. This experiment uses these cells directly.

Key parameter	Value	Why it differs
Architectures	COBA and PING	Provides a feedforward control and recurrent E/I model
Training pool	70,000 samples	Uses all pooled MNIST rather than the sweep default of 7,000
Input rate	25 Hz maximum-pixel rate	Fixed-rate collection baseline
Spike budget	Off	Measures unconstrained capacity
Cells	2 architectures $\times$ 3 seeds $= 6$	Across-seed comparison for both models
Readout shape	$1024 \rightarrow 10$ spiking LIF outputs	`mem-mean`: mean membrane voltage supplies the logits

---

TR-02 — Spike-budget sweep

This run measures the trade-off between accuracy and firing rate as the spike budget is tightened. It uses a smaller training pool to keep the multi-seed sweep manageable; only the spike cap and its penalty change across conditions. Its absolute rates should therefore not be compared directly with the full-data cells. The resulting checkpoints are used by exp024, exp025, exp037, exp038.

Key parameter	Value	Why it differs
Architectures	COBA and PING	Direct architecture comparison
Training pool	7,000 samples	Keeps the 36-cell sweep tractable
Spike budget θ_u	off, 5, 2, 1, 0.5, 0.2 spikes/ neuron/trial	Spans unconstrained through severe sparsity
Penalty strength	10^{-3} when enabled	Applies the upper-rate regularizer

Key parameter	Value	Why it differs
Cells	2 × 6 settings × 3 seeds = 36	Error bars at every frontier point
Readout shape	1024 → 10 spiking LIF outputs	`mem-mean`: mean membrane voltage supplies the logits

---

TR-03 — Inhibitory-timescale sweep

This run changes τ_{GABA} to test how the inhibitory timescale affects gamma frequency, firing rate, and accuracy. All other scientific settings are held fixed, with 6 ms as the standard condition. The resulting checkpoints are used by exp041, exp042, exp046.

Key parameter	Value	Why it differs
Architecture	PING	The manipulated quantity belongs to the recurrent inhibitory loop
Training pool	7,000 samples	Sweep-scale default
τ_{GABA}	4.5, 6, 9, 12, 18, 27 ms	Moves the inhibitory rhythm across a broad timescale range
Cells	6 settings × 3 seeds = 18	Across-seed estimate at each decay
Readout shape	1024 → 10 spiking LIF outputs	`mem-mean`: mean membrane voltage supplies the logits

---

TR-04 — Integration-timestep sweep

This run changes the integration timestep while keeping each presentation at 200 ms. It tests whether the observed dynamics depend on numerical resolution and measures the extra compute required by finer timesteps. The 0.05 ms cells need the large-memory Cambridge request. The resulting checkpoints are used by exp044.

Key parameter	Value	Why it differs
Architecture	PING	Tests the recurrent reference model
Training pool	7,000 samples	Sweep-scale default
Δt	0.05, 0.1, 0.25, 0.5, 1 ms	Changes numerical resolution while holding 200 ms physical time fixed
Steps/presentation	4,000; 2,000; 800; 400; 200	Compute and activation memory scale inversely with Δt
Cells	5 settings × 3 seeds = 15	Across-seed stability check
Readout shape	1024 → 10 spiking LIF outputs	`mem-mean`: mean membrane voltage supplies the logits

---

TR-05 — Recurrent-initialization sweep

This run tests whether training preserves the PING loop or learns a useful loop from weaker initial conditions. Recurrent initialization and trainability change together, while the feedforward network and classifier remain fixed to the PING recipe. The resulting checkpoints are used by exp049.

Key parameter	Value	Why it differs
Architecture	PING	Manipulates the recurrent loop directly
Training pool	7,000 samples	Sweep-scale default
Loop conditions	frozen PING; trainable PING init; trainable zero init; trainable 0.1 init	Separates built-in dynamics from recurrence learned during task training
Trainable projections	W_{EI} and W_{IE} only in trainable conditions	The frozen condition is the mechanistic control
Cells	4 conditions $\times$ 3 seeds = 12	Across-seed comparison
Readout shape	$1024 \rightarrow 10$ spiking LIF outputs	`mem-mean`: mean membrane voltage supplies the logits

---

TR-06 — Variable-rate streaming bank

This run trains PING across the input rates used by exp082. One rate is sampled uniformly for each presentation, and the ten output LIF neurons produce class logits from their firing rates rather than their mean membrane voltages. Expressing the logits in hertz makes them comparable across presentation durations. During streaming inference, the output neurons reset at digit boundaries while the hidden PING state continues. The resulting checkpoints are used by exp082.

Key parameter	Value	Why it differs
Architecture	PING	Target model for streaming inference
Training pool	7,000 samples	Uses the shared sweep-scale training set
Input-rate set	0.5, 0.75, 1, 1.5, 2, 3, 5, 7.5, 10, 15, 25 Hz	Denser sampling within the interval selected by exp080
Sampling rule	Uniform categorical, independently per presentation	Makes rate variation part of the training distribution
Readout	`spike-rate`	Hidden E spikes drive ten spiking LIF class neurons; each logit is that class neuron's spike count divided by presentation duration in seconds

Key parameter	Value	Why it differs
Readout shape	1024 → 10 spiking LIF outputs	Ten class neurons emit and reset throughout the presentation
Cells	1 recipe × 3 seeds = 3	Checkpoint bank expected by exp082

Results by training run

Each completed run reports one training-curve figure and one representative seed-42 raster. The figures summarize all configurations and seeds; the raster is a diagnostic example, not an across-seed statistic.

TR-01 results — Canonical full-data reference

Run status: complete · 6/6 cells represented

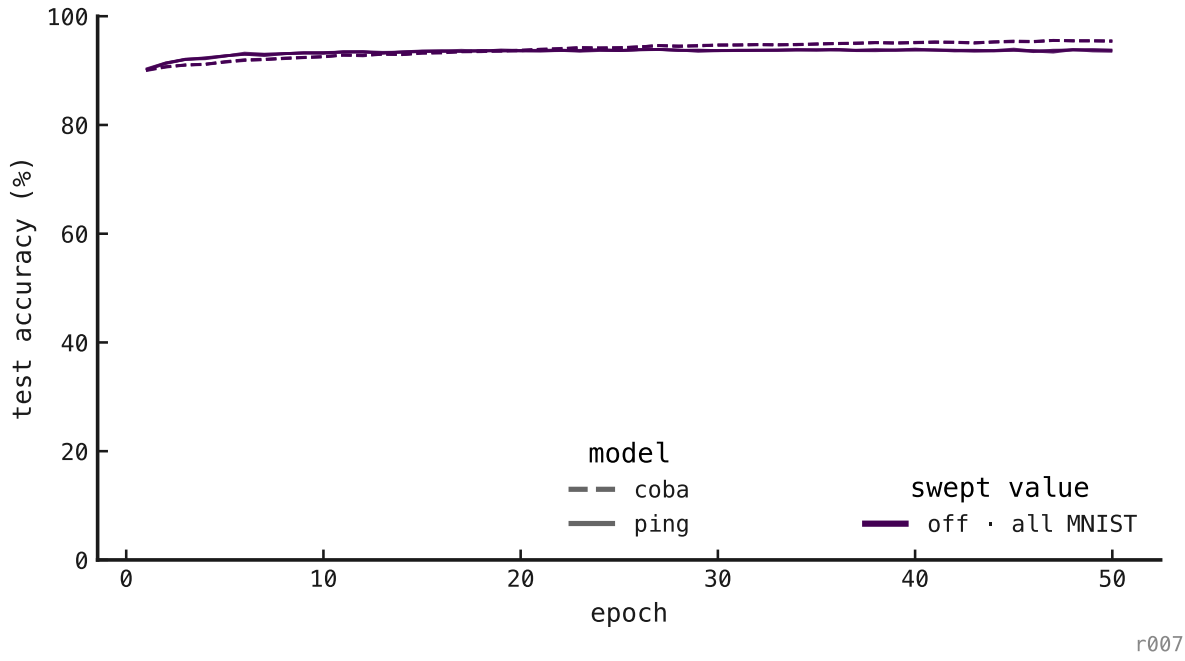

Figure 1: Training curves for the six full-data reference cells.

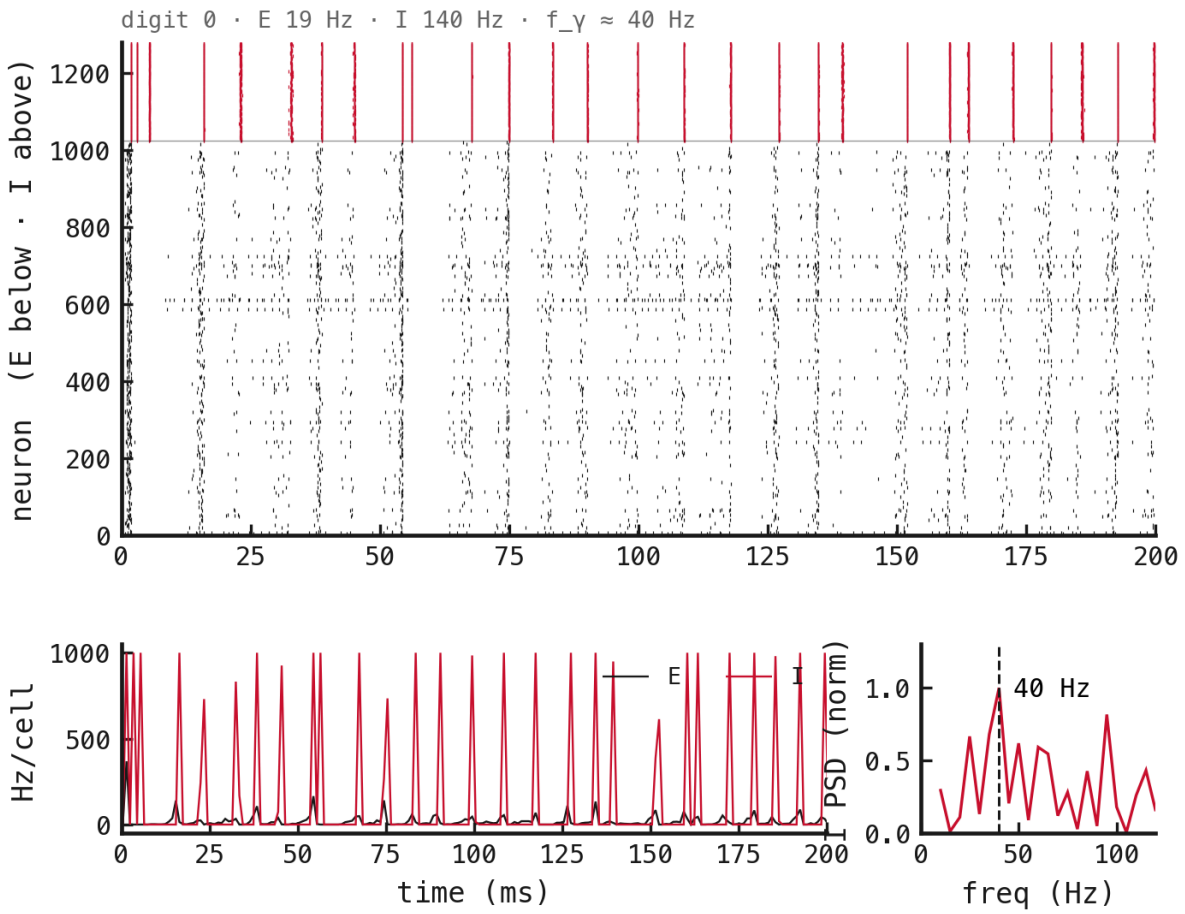

Figure 2: Sample raster: canonical PING, seed 42, held-out digit-0 probe.

TR-02 results — Spike-budget sweep

Run status: complete · 36/36 cells represented

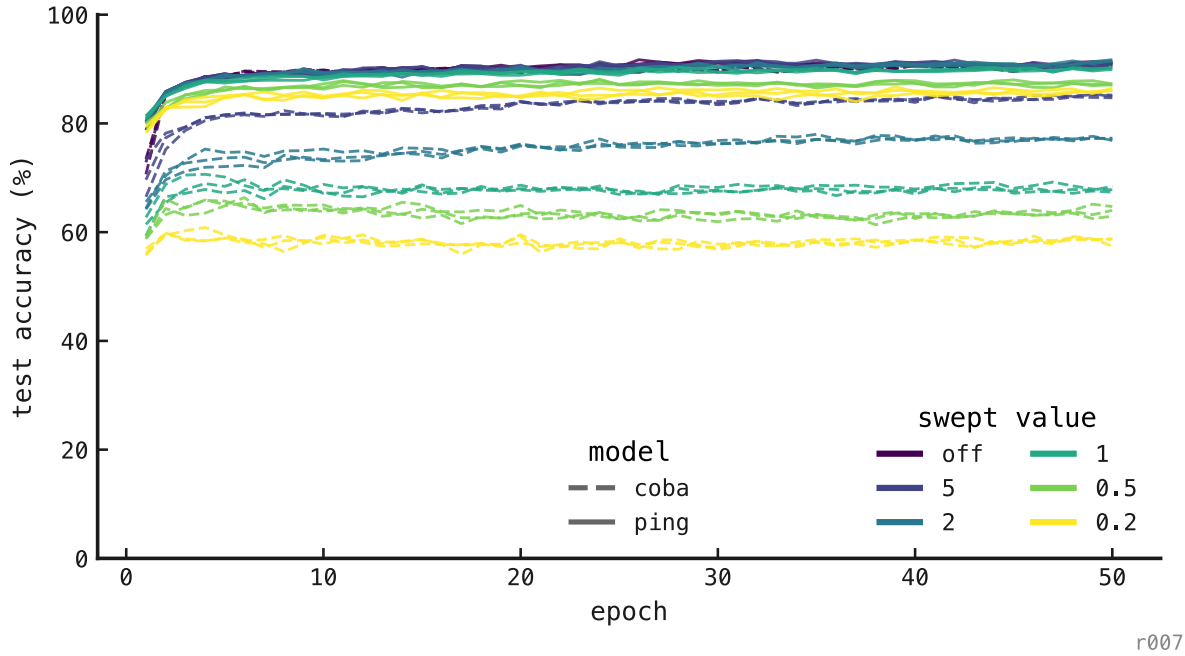

Figure 3: Training curves across both architectures, six spike budgets, and three seeds.

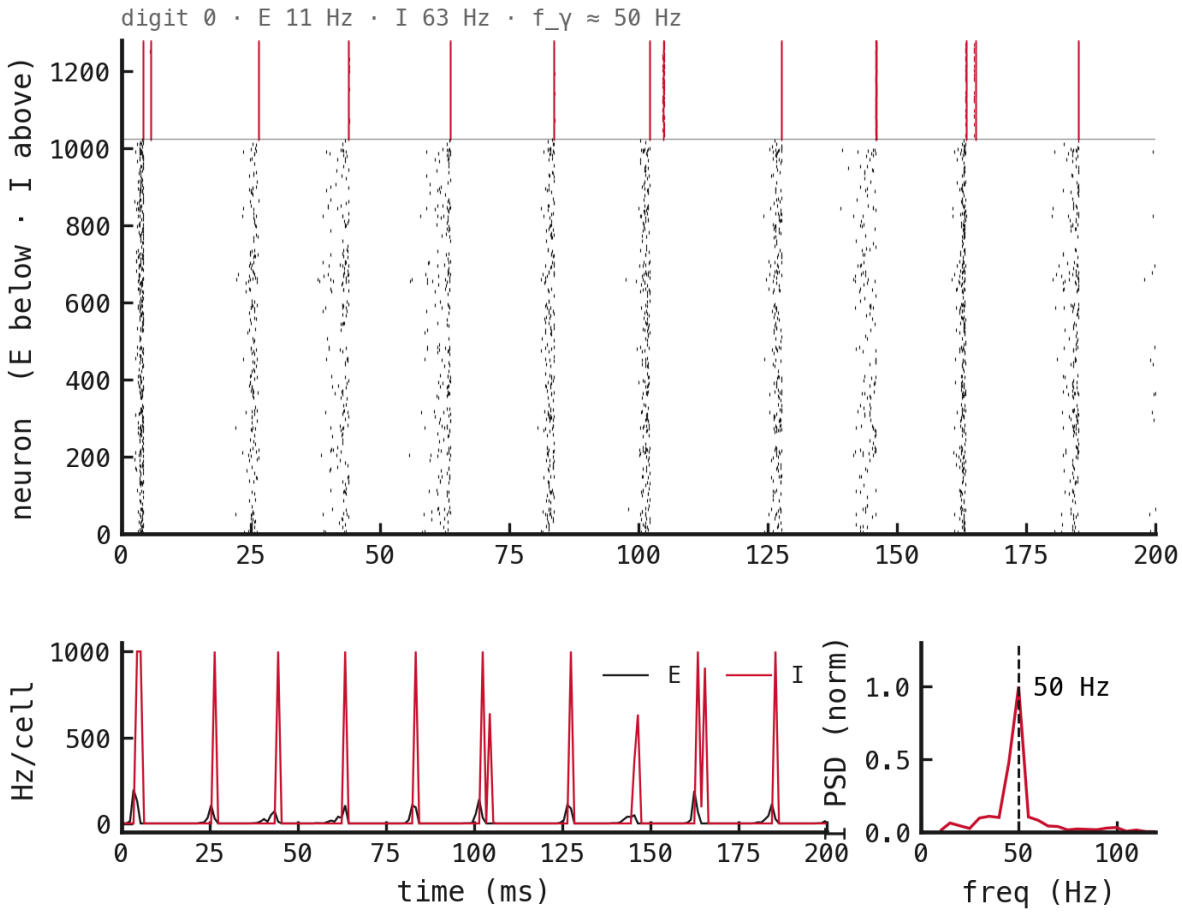

Figure 4: Sample raster: PING with spike budget off, seed 42.

TR-03 results — Inhibitory-timescale sweep

Run status: complete · 18/18 cells represented

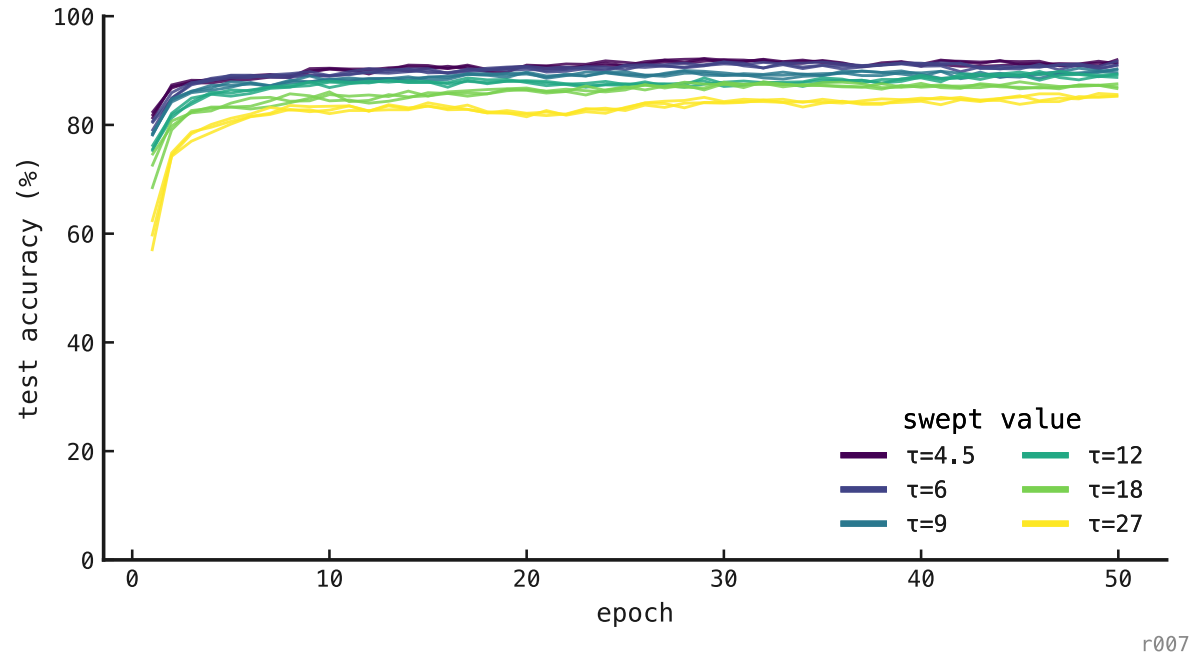

Figure 5: Training curves across six τ_{GABA} values and three seeds.

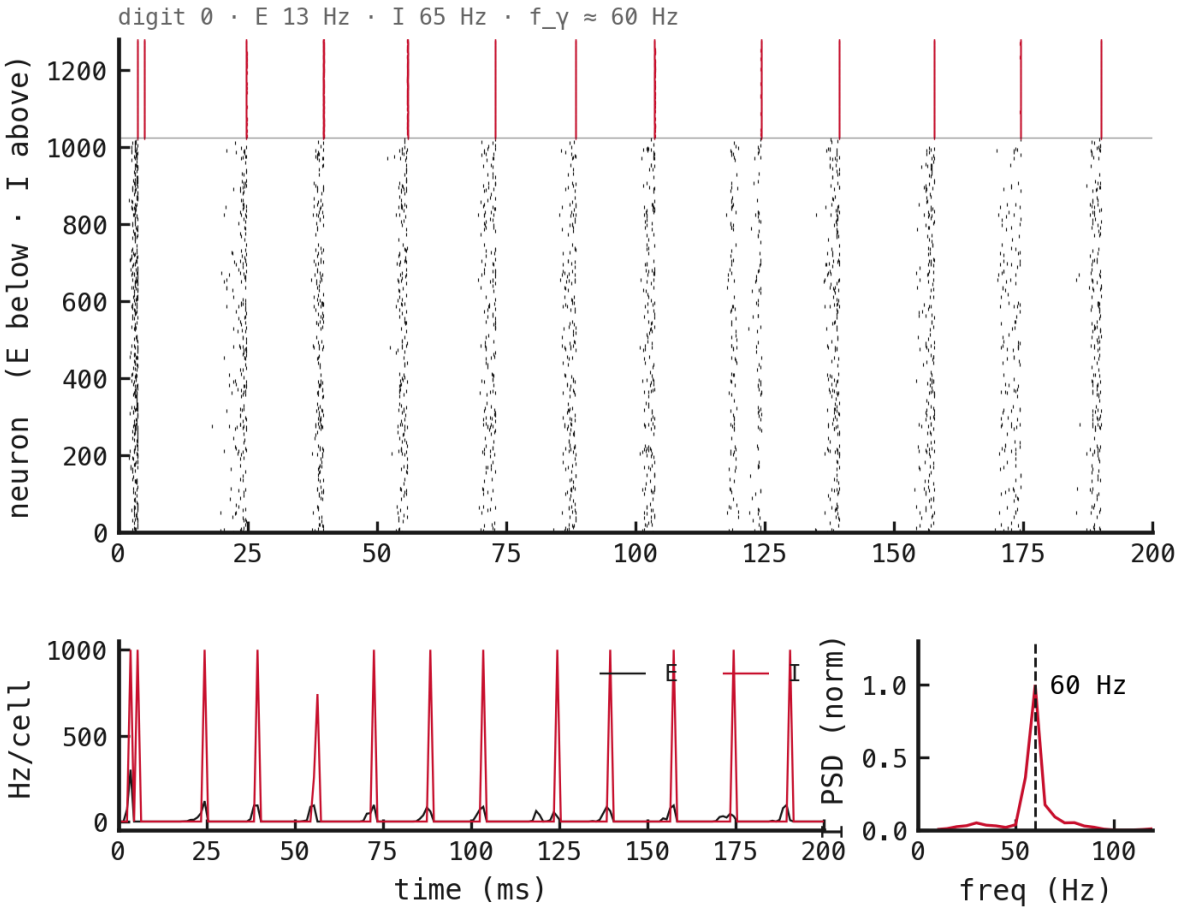

Figure 6: Sample raster: PING at $\tau_{\text{GABA}} = 6$ ms, seed 42.

TR-04 results — Integration-timestep sweep
Run status: complete · 15/15 cells represented

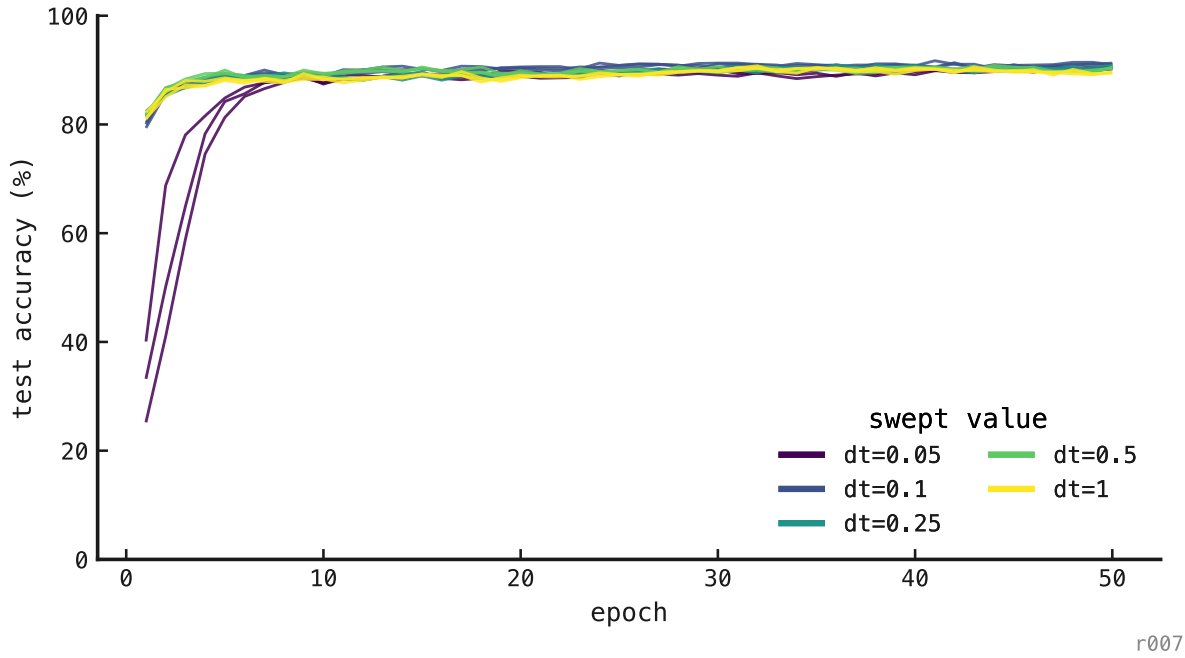

Figure 7: Training curves across five integration timesteps and three seeds.

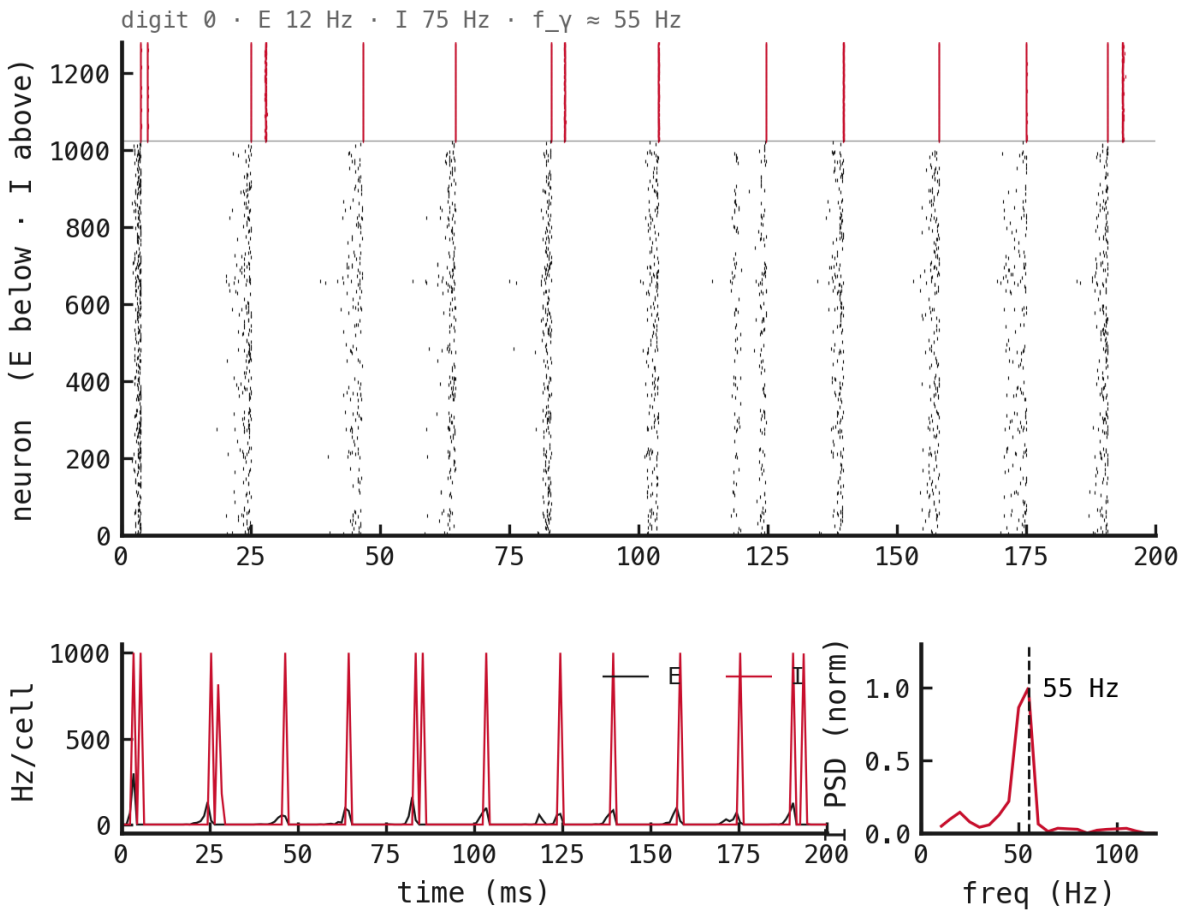

Figure 8: Sample raster: PING at $\Delta t = 0.1$ ms, seed 42.

TR-05 results — Recurrent-initialization sweep
Run status: complete · 12/12 cells represented

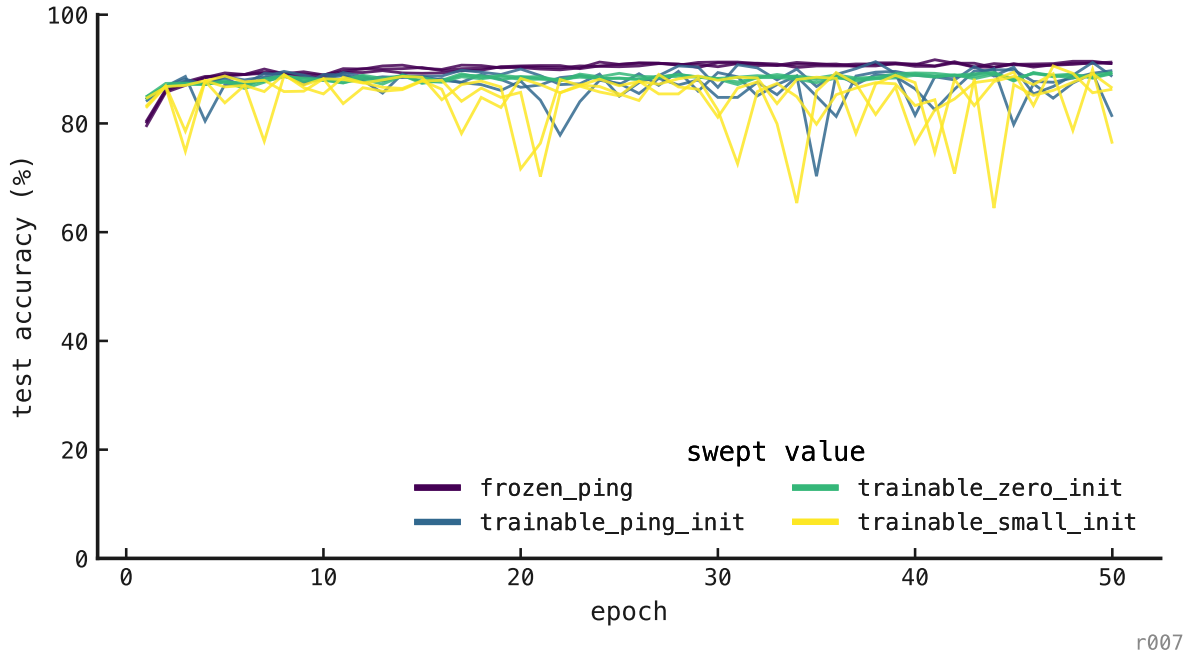

Figure 9: Training curves across four recurrent-loop conditions and three seeds.

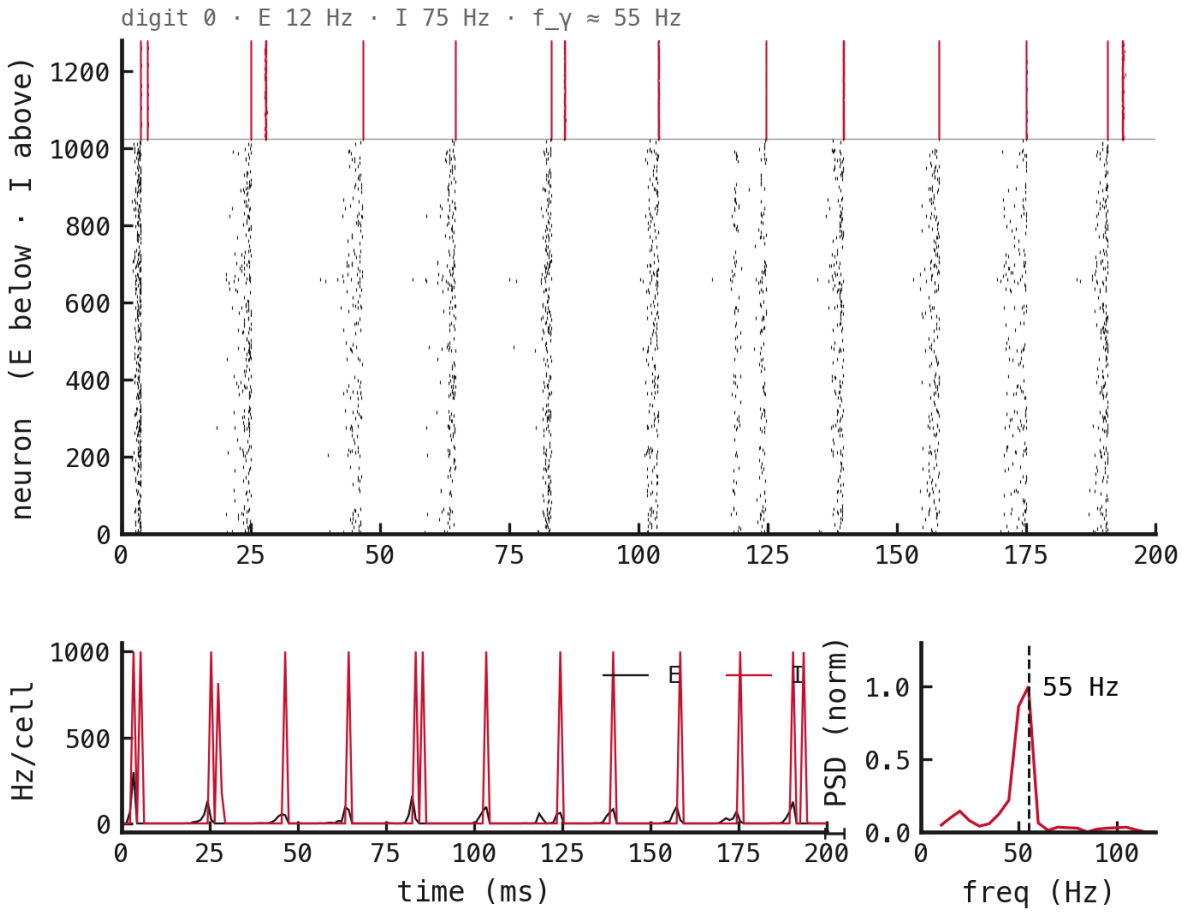

Figure 10: Sample raster: frozen PING recurrent loop, seed 42.

TR-06 results — Variable-rate streaming bank
Run status: pending · 0/3 cells trained

TODO — training curves. Add the three variable-rate learning curves after the Cambridge jobs complete.

TODO — sample raster. Add a seed-42 E/I/output raster at a declared held-out rate, plus low- and high-rate diagnostics if one raster hides a rate-dependent failure.