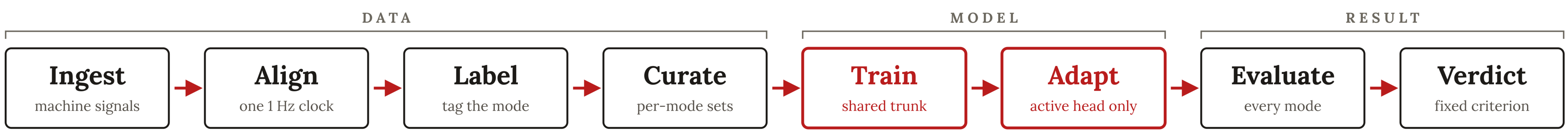
Adapting Without Forgetting
A Mode-Aware Surrogate for a Drifting FEL

Nicholas Mamais – mamais@slac.stanford.edu, E. Williams, S. Miskovich, M. Gibbs, A. Reed, D. Cesar, G. White, A. Edelen

SLAC NATIONAL ACCELERATOR LABORATORY | U.S. DEPARTMENT OF ENERGY | Stanford University

FROM MACHINE SIGNALS TO A VERDICT



THE PROBLEM: DRIFT AND FORGETTING

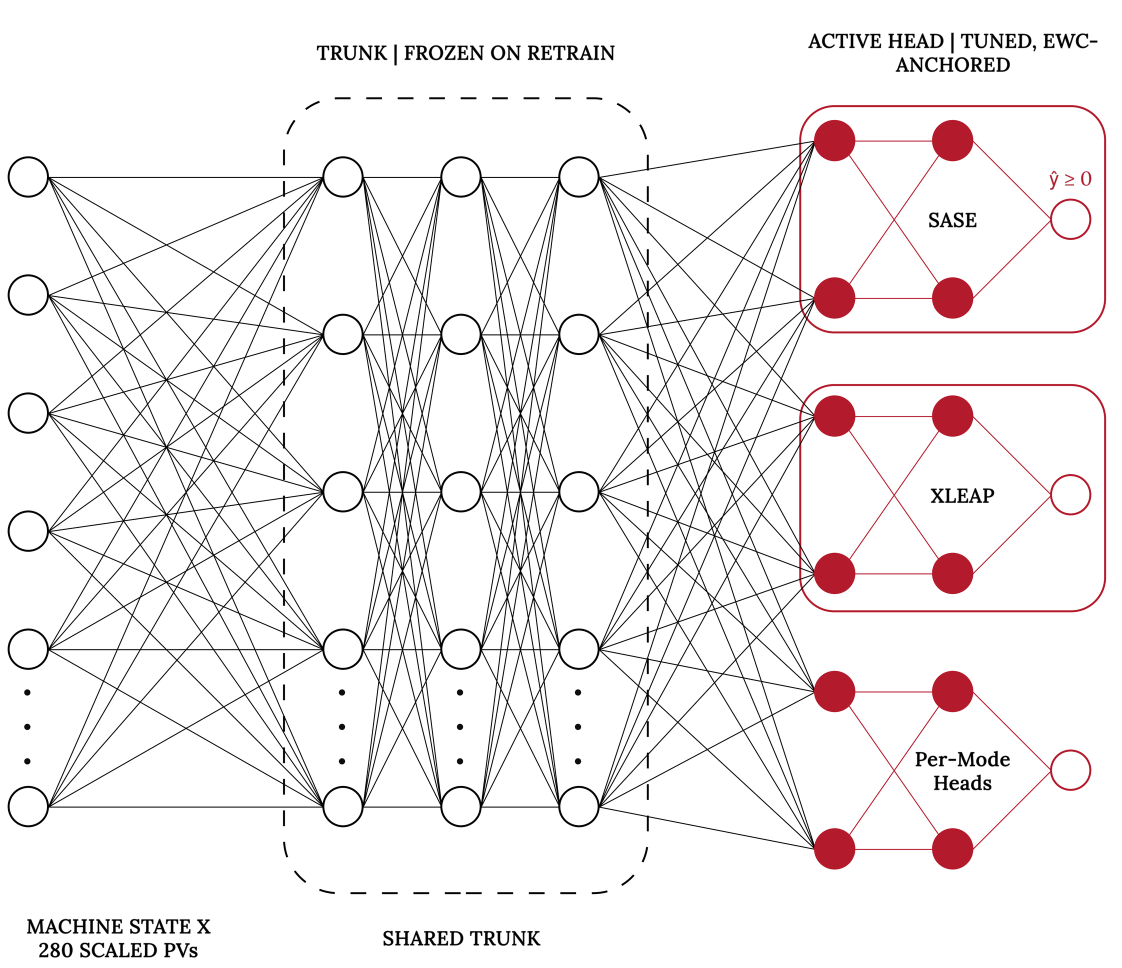
LCLS, SLAC's X-ray free-electron laser, changes operating conditions frequently and drifts continuously, so a machine-learning surrogate trained once goes stale over time. Retraining online restores model accuracy for the mode being delivered now, but degrades accuracy for the modes it no longer sees[2,3]. A useful surrogate must track the machine and retain every mode it has learned.

Research Question

Can a surrogate adapt to drifting the active mode without degrading the modes it isn't currently seeing?

THE APPROACH & MODEL

Hundreds of EPICS control signals, aligned to one clock and tagged with the mode actually delivered. A shared trunk feeds one head per mode; while streaming, only the active head updates, and only when its error spikes.



One shared trunk over 280 scaled machine signals, one output head per operating mode. Only the active head is tuned.

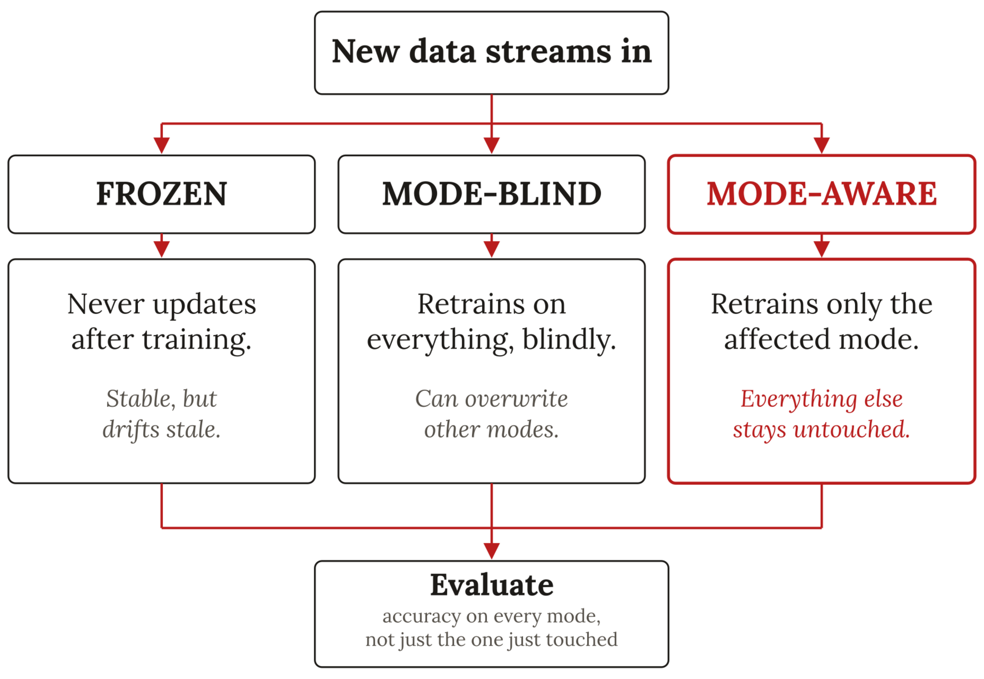
HOW WE KNOW IF IT WORKS

Hypothesis: adapting to the mode being delivered improves prediction accuracy over a non-adaptive model.

Test: across 20 independent trials (5 random seeds x 4 stream start points), the adaptive surrogate had to beat a non-adaptive baseline on two metrics: prediction error (RMSE) and distributional skill. We required it to pass in at least 15 0f the 20.

THE EXPERIMENT

Three regimes share the same trunk and the same warm start, differing only in head structure and update policy.

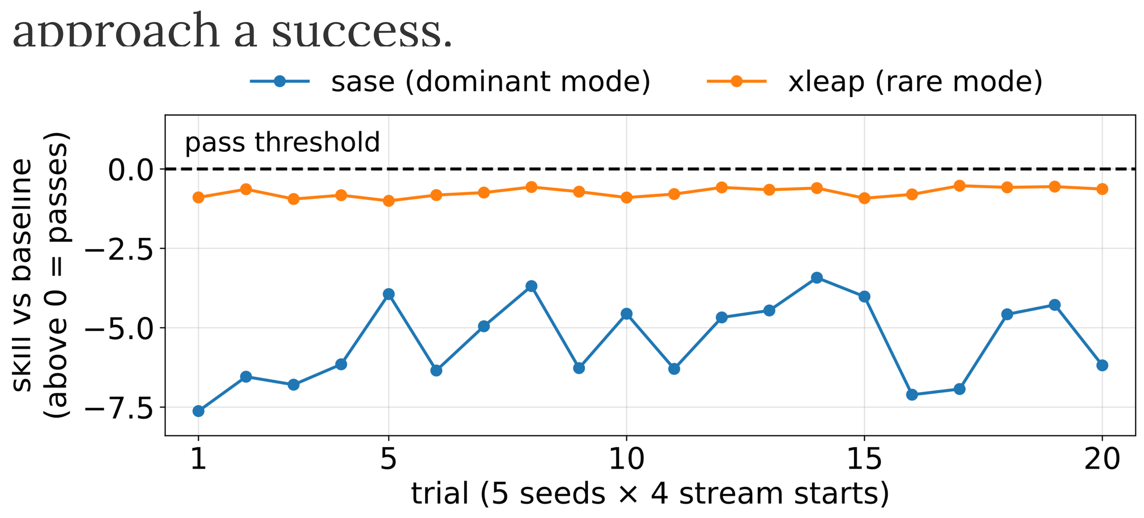


Only the update rule differs. Mode-aware confines each retrain to the head for the mode that drifted.

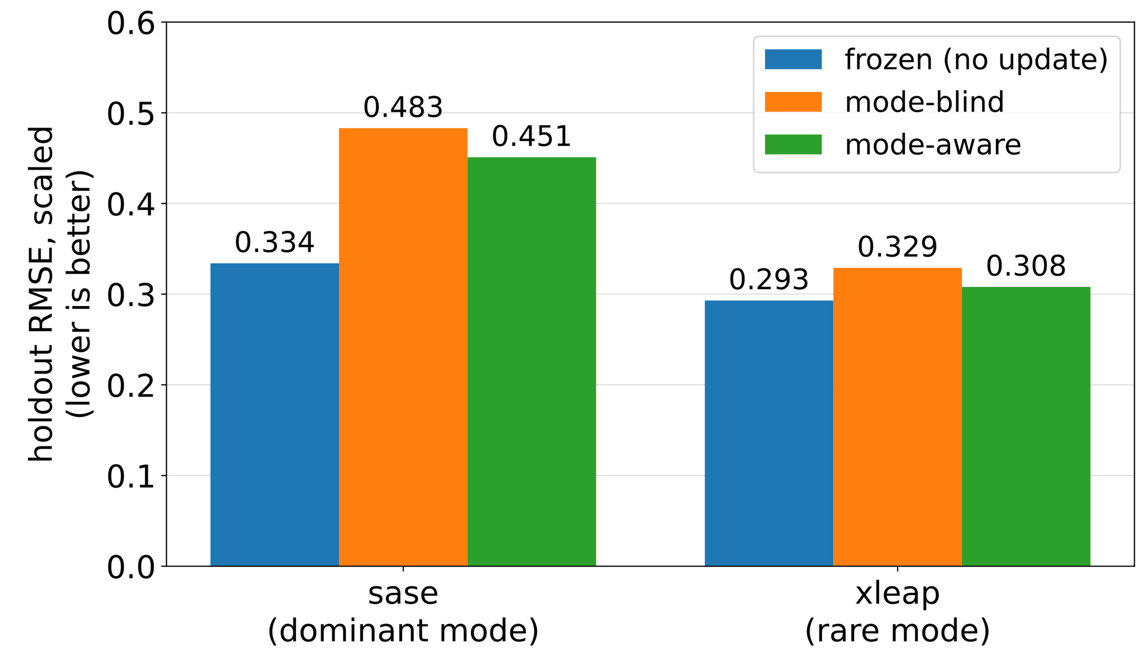
RESULTS

PRELIMINARY: RESEARCH IN PROGRESS

Judged against the criterion we fixed before running anything, this iteration of the model does not meet the standard we set for calling the approach a success.



Skill = 1 − CRPS_model / CRPS_baseline. CRPS, Continuous Ranked Probability Score, measures how well the entire predicted probability distribution matches the measured value rather than just the mean; lower is better. Every trial falls short, and none of the 20 crosses the line. The margin is far smaller on the rare mode.



Streaming error over the registered run, median of 20 trials. Both adaptive regimes score worse than frozen, on both modes.

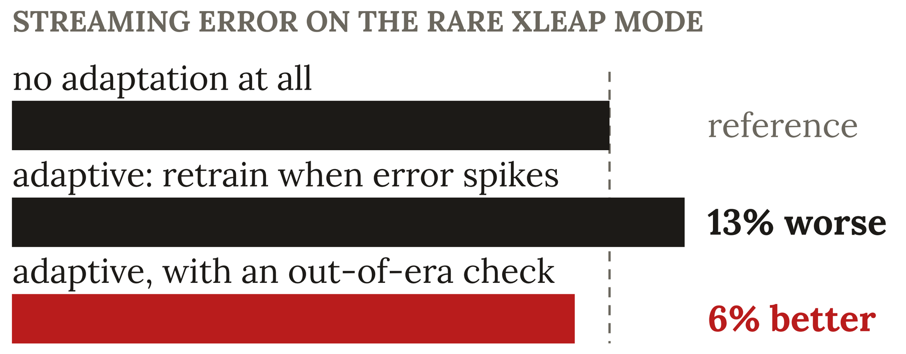
CONCLUSIONS AND NEXT STEPS

Adaptive modeling did not beat the non-adaptive baseline under the registered criterion. Judging each retrain against data from outside the current period changed that on the rare XLEAP mode, where the adaptive model went from 13% worse than never updating to 6% better. On the dominant SASE mode it improved from 45% worse to 7% worse, but never became better than not updating.

WHAT'S NEXT

Score within one continuous run, tune the retrain cadence, and grow a library of specialist models instead of overwriting one [4].

STREAMING ERROR ON THE RARE XLEAP MODE



Median over 20 trials, lower is better. Preliminary.

ACKNOWLEDGEMENTS

The author thanks TID-ID-CSE-ML (Ernest William's Group) for their support and feedback. This work was carried out through the LCLS Summer Internship Program. Work supported by the U.S. Department of Energy.

REFERENCES

1. Duris et al., Nat. Photonics 14, 30 (2020). 2. Kirkpatrick et al., PNAS 114, 3521 (2017). 3. Lopez-Paz & Ranzato, NeurIPS (2017). 4. Rajput et al., Mach. Learn.: Sci. Technol. 6, 031001 (2025).