

Stable State Coordinates for Compact Learned Dynamics: Aircraft Vibration and Nano-Drone Identification

Daniel Schmitter

Abstract

A compact approximation of a neural predictor need not preserve its autonomous dynamics. We study an alternative in which stable input-driven state coordinates are supplied before fitting a nonlinear readout. On an F-16 ground-vibration record, replacing predicted-output history with a stable modal carrier yields mean channel RMSE 0.49899, compared with 0.48940 for a locally reproduced neural teacher, while reducing counted deployment arrays by a factor of 6.96. On a separate nano-drone personalization comparison, a direct exponential-memory program uses 1,020 times fewer retained program bytes and 65.9 times less measured compilation time than an adapted neural ensemble, but loses all four accuracy groups. The studies therefore establish conditional accuracy-resource tradeoffs, not neural-model dominance or robotic control. We derive the bounded-input stability and quadratic update identities behind the implementations, identify their supplied information, and retain negative controls showing that static approximation quality and additional spline capacity do not guarantee better dynamics.

1 Introduction

The state representation determines what a compact physical predictor can express. A memoryless nonlinear map cannot distinguish two histories that share its current input but imply different futures. Conversely, compressing an autoregressive network may change feedback stability even when its one-step approximation error is small. We investigate input-driven stable state as a boundary between these problems.

The contribution is a measured compiler design: retain causal linear memory, learn small cardinal response maps or direct finite-horizon heads, and use quadratic statistics for fitting and update admission. The two datasets support distinct conclusions. Aircraft vibration tests near-teacher free-run prediction; nano-drone data reveal a substantial resource reduction with a remaining accuracy cost.

2 Related work

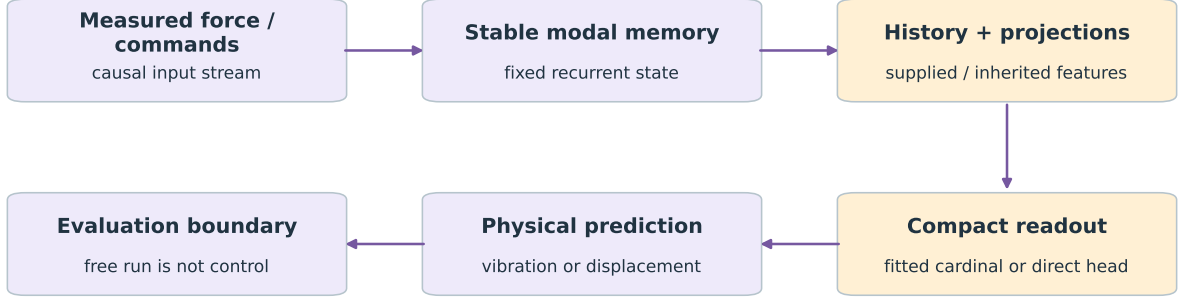
Nonlinear system identification already combines linear dynamics with nonlinear maps. The F-16 benchmark [1] and convolutional identification study [2] provide the measured system and neural comparison. The nano-drone benchmark [3] supplies multi-horizon physical errors; the released ASIA models [4] provide the later strong adapted comparator. These are not benchmarks of a robot controlled by our compiled model. Cardinal representation and conditional least squares are established tools; the tested contribution is their state-aware integration and its measured limitations.

3 Architecture and information flow

The compact systems preserve a causal state before reducing the output model. In the aircraft path, observed force drives stable modes, finite histories feed inherited projection directions, and fitted cardinal edges correct a modal carrier. Predicted output is not returned to the state recurrence. In the drone path, exact held-input exponential traces and kinematic features feed direct horizon-specific predictions. These are two distinct realizations of the same allocation principle, not one shared trained network.

Distillation can select useful directions without making its teacher part of the deployed arithmetic, but the teacher and source fitting remain part of acquisition cost. Conversely, a readout that still invokes a large encoder is not an independent compressed model. The aircraft and drone accounting must therefore distinguish inherited training, retained executable arrays, state initialization, and the trajectory task. The large resource ratio in the drone comparison is meaningful only alongside its consistently worse physical errors.

Compress the state, then fit the response



Read left to right above, then right to left below. Arrows show data or computation flow.

Figure 1: Compress the state, then fit the response. Boxes distinguish supplied information, fitted components, and the quantity evaluated. Arrows show computation or data dependence, not a newly trained deep network.

4 Stable input-driven construction

Let a finite collection of modal states follow stable linear recurrences. At each time, concatenate a finite history of the input and modal outputs, project it along fixed directions, and evaluate compact cardinal response functions. The output is the modal carrier plus the summed corrections:

$$z_{k+1} = Az_k + Bu_k, \quad \rho(A) < 1, \quad (1)$$

$$q_e[k] = w_e^T \ell_k + b_e, \quad \hat{y}_k = Cz_k + \sum_e f_e(q_e[k]). \quad (2)$$

Proposition 1. For bounded inputs, finite initial state, fixed finite projections, positive finite normalization scales, and bounded cardinal readouts, this predictor has bounded output. Indeed, a finite-dimensional stable matrix admits a summable impulse response. Its convolution with bounded input and its decaying initial response are bounded. Finite histories and fixed projections preserve boundedness, as does the final readout. This is a bounded-input statement, not an incremental-gain, passivity, closed-loop safety, or approximation-error certificate.

Crucially, predicted physical outputs are not fed back into the aircraft compiler’s modal carrier. A nonlinear autoregressive surrogate instead needs a condition on the complete feedback map. Small teacher-forced error does not supply such a condition. Likewise, physical position without velocity is generally not a sufficient state: at one position, opposite velocities give opposite immediate motion.

5 Fitting and quadratic admission

For fixed features Φ and output targets Y , the squared-error fit retains a Gram G and right-hand side B . Continuous mass or derivative Grams add a declared functional regularizer. Only that regularizer inherits appropriate basis symmetry; the empirical Gram is not assumed circulant.

$$(G + \lambda Q)W = B, \quad G = \Phi^T \Phi, \quad B = \Phi^T Y, \quad (3)$$

$$\Delta L(\eta) = 2\eta \langle \Phi W - Y, \Phi \Delta W \rangle + \eta^2 \|\Phi \Delta W\|^2. \quad (4)$$

The second identity evaluates an interpolation between a retained core and a candidate correction on fixed admission observations. Its coefficients can be computed from the same quadratic statistics. It proves the value of the declared empirical loss along that line; it does not imply no harm on unobserved trajectories or permit changing features without additional historical cross-statistics.

6 Experimental methods

The F-16 compiler uses SpecialOddMSine Level-2 measurements. Seven realizations build the readout-selection statistics and normalization; zero-based realization seven selects among four fixed ridge values. The final readout fits the first eight realizations; realization eight is internal evaluation, followed by the separately packaged official test. The modal carrier and inherited teacher directions use their recorded source fitting procedures. Thus the inner ridge-selection realization is not an untouched test of the entire feature-learning pipeline.

The carrier uses 159 stable complex modes. Sixteen response-selected directions inherited from a reproduced 256-unit sigmoid network project 64-tap force and carrier histories. Each vector-valued cubic cardinal edge has seventeen coefficients per output. The teacher has 66,563 binary32 parameters and initializes autoregressive evaluation with 64 measured outputs. The compiled predictor uses no measured-output initialization. Both inherited training and the different initialization rules are part of the comparison.

The nano-drone study uses fifteen pinned 100-Hz flights. Nine Square/Random/Chirp flights fit the source and three disjoint repeats select ridge. Three Melon flights are used in order for adaptation, admission, and confirmation. Six exact held-input actuator recurrences use time constants 0.02, 0.05, 0.10, 0.25, 0.50 and 1.00 seconds. Physical rotor mixtures, kinematic carriers, and horizon coordinates form a 126-feature direct displacement program for horizons one through fifty. Direct predictions do not recursively feed predicted states into this head.

$$z_j[k+1] = e^{-\Delta t/\tau_j} z_j[k] + (1 - e^{-\Delta t/\tau_j}) u[k]. \quad (5)$$

The later neural comparison reproduces the released three-fold ASIA ensemble and adapts it on the same Melon run-one sparse endpoint task, with run-two admission. Run-three evaluation uses its 130 nonoverlapping starts. The recorded comparison is post-exposure: earlier operator results had already opened this target. Operator and neural label counts are 3,250 and 3,380, respectively, not exactly equal. Timings and program bytes are distinct from host RSS, teacher-training cost, and physical energy.

7 Results

Table 1: F-16 official-record mean channel RMSE; local teacher evaluation is retrospective.

Predictor	Mean RMSE	Counted deployment bytes
Modal/cardinal compiler	0.49899	38,248
Local neural teacher	0.48940	266,252
Published MLP / TCN / LSTM	0.48 / 0.63 / 0.74	Not remeasured

The compiler is 1.96% worse in mean RMSE than the local teacher and 6.96 times smaller in the stated array accounting. Sparse/dense basis, accumulated/concatenated statistics, and independent solve discrepancies are below 10^{-10} . Whole/chunked prediction and repeat prediction are bitwise identical in the recorded check. Direct teacher-forced pruning instead gives free-run RMSE 1.73593, demonstrating that a good static fit is insufficient.

Table 2: Matched nano-drone cumulative errors across fifty horizons, in the benchmark’s four physical groups.

Model	Position	Velocity	Orientation	Angular velocity
ASIA sparse admitted	1.6051	4.8836	2.1640	13.8756
Exact-memory direct program	1.9305	7.1858	2.7783	17.0201

The neural ensemble wins every accuracy group. The operator retains 24,288 program bytes versus approximately 24.79 MB for source plus adapted neural programs, and compiles in 1.538 seconds versus 101.33 seconds of target optimization. Those ratios do not establish equal capability. A preceding ablation worsens the operator score when exponential-memory columns are removed, but improves it when broad cardinal horizon/state/input columns are removed. Memory is load-bearing in that panel; generic basis capacity is not.

A later teacher-consolidation experiment combines recurrent translation with direct angular outputs and multiple teacher execution Grams. It reduces the mean physical error ratio to the adapted ensemble to 1.1218 while remaining about 314 times smaller. This is additional post-exposure mechanism evidence, not independent confirmation or a replacement for the failed matched accuracy comparison.

8 Discussion

These experiments support compiling the dynamical state rather than merely pruning a nonlinear graph. They also show why a resource advantage must be presented as a frontier rather than a universal win. Stable input-driven coordinates limit error amplification, but may omit genuinely necessary output feedback or hidden state. A direct finite-horizon predictor must be revalidated if used recursively or embedded in a controller.

The source artifacts contain exact split definitions, full error curves, normalization, and resource accounting. The publication supplement preserves their hashes and result files without redistributing raw benchmark data. No new aircraft fit, drone training, or physical experiment was performed for this paper reconstruction. Subsequent work would need independent systems and matched full deployment costs before making a broad edge-intelligence claim.

9 Application boundary and research implication

The practical target is a low-cost model component for monitoring, simulation, or adaptation. Closing a feedback loop changes the input distribution and adds the controller to the stability question. Neither bounded-input state nor successful recorded-flight prediction proves safe flight. The failed static-pruning and no-memory controls make the state representation a substantive part of the contribution.

10 Conclusion

Stable physical memory and coefficient-domain fitting can yield compact executable predictors with useful bounded tradeoffs. The aircraft compiler approaches a reproduced teacher without output feedback; the nano-drone comparison remains less accurate than its adapted neural control. Neither result demonstrates autonomous flight or a universal replacement for neural dynamics.

References

- [1] M. Schoukens and J.-P. Noël. F-16 Aircraft Benchmark Based on Ground Vibration Test Data. Nonlinear System Identification Benchmarks Workshop, 2017. Dataset version 1. <https://doi.org/10.4121/12954911.v1>
- [2] C. Andersson, A. H. Ribeiro, K. Tiels, N. Wahlström, and T. B. Schön. Deep Convolutional Networks in System Identification. 2019. <https://arxiv.org/abs/1909.01730>
- [3] R. Busetto, E. Cereda, M. Forgione, G. Maroni, D. Piga, and D. Palossi. Nonlinear System Identification Nano-drone Benchmark. 2025 preprint; Control Engineering Practice 172, 106871, 2026. <https://arxiv.org/abs/2512.14450>
- [4] D. Piga and M. Forgione. ASIA: an Autonomous System Identification Agent. 2026. <https://arxiv.org/abs/2605.10480>