Observability-Limited Parameter Identification in Physics-Informed Neural Networks: Stability Analysis for Quadrotor System Identification

Sreejita Chatterjee[1]

***Abstract*—Physics-Informed Neural Networks (PINNs) enable simultaneous dynamics learning and parameter identification by embedding governing equations into neural network training. We present a systematic analysis of PINN-based system identification for 6-DOF quadrotor dynamics, addressing two challenges: autoregressive prediction stability and parameter observability limits. We show that architectural modifications improving single-step prediction can destabilize multi-step rollouts by 100–1,000,000×, with direct implications for model predictive control. We identify modular architecture decoupling and Fourier feature extrapolation as primary failure mechanisms. For parameter identification, we characterize observability using Fisher Information: mass and motor coefficients achieve 0% error due to strong translational dynamics gradients, while inertia parameters saturate at 5% error due to weak cross-coupling at small angles ($\pm 20°$). A curriculum-based training methodology achieves 51× stability improvement. Experiments with aggressive maneuvers (± 45–$60°$) paradoxically degraded identification due to simulator-model mismatch, demonstrating that excitation must match model fidelity. These results establish practical bounds for PINN-based system identification in control applications.**

I. INTRODUCTION

System identification is fundamental to model-based control of dynamical systems. Physics-Informed Neural Networks (PINNs) offer a learning-based approach that embeds governing equations—such as Newton-Euler dynamics—directly into neural network training [1]. This enables simultaneous dynamics learning and parameter identification with physical consistency guarantees.

For control applications, however, two critical challenges arise. First, model predictive control (MPC) requires stable *multi-step* predictions, yet PINNs are typically evaluated on *single-step* accuracy. We demonstrate that these metrics can contradict: architectures improving single-step accuracy by 2–10× may destabilize 100-step rollouts by 10^2–$10^6\times$. Second, parameter identification accuracy is fundamentally limited by *observability*—the information content in measured trajectories about unknown parameters.

The contributions of this paper are:

1) Stability analysis of autoregressive PINNs, identifying failure mechanisms in modular and Fourier architectures (Sec. III).
2) Fisher Information-based observability analysis explaining differential identification accuracy across parameters (Sec. IV).
3) A curriculum-based training methodology achieving 51× stability improvement with accurate parameter identification (Sec. V).
4) Characterization of model mismatch effects on identification from aggressive maneuvers (Sec. VI).

II. PROBLEM FORMULATION

A. Quadrotor Dynamics

Consider a 6-DOF quadrotor with state $\mathbf{x} \in \mathbb{R}^{12}$:

$$\mathbf{x} = [x, y, z, \phi, \theta, \psi, p, q, r, v_x, v_y, v_z]^T \tag{1}$$

and control input $\mathbf{u} = [T, \tau_x, \tau_y, \tau_z]^T$. The dynamics follow Newton-Euler equations with unknown parameters $\boldsymbol{\theta} = [m, J_{xx}, J_{yy}, J_{zz}, k_t, k_q]^T$:

Rotational dynamics:

$$\dot{p} = \frac{(J_{yy} - J_{zz})qr}{J_{xx}} + \frac{\tau_x}{J_{xx}} \tag{2}$$

Translational dynamics:

$$\dot{v}_z = -\frac{T\cos\theta\cos\phi}{m} + g - c_d v_z |v_z| \tag{3}$$

B. PINN Formulation

The PINN predicts next state: $\hat{\mathbf{x}}_{t+1} = g_\phi(\mathbf{x}_t, \mathbf{u}_t)$ with learnable parameters $\hat{\boldsymbol{\theta}}$. Training minimizes:

$$\mathcal{L} = \mathcal{L}_{\text{data}} + \lambda_p \mathcal{L}_{\text{physics}} \tag{4}$$

where $\mathcal{L}_{\text{physics}}$ enforces (2)–(3).

C. Autoregressive Rollout

For control, predictions recursively feed as inputs:

$$\hat{\mathbf{x}}_{t+k} = g_\phi^{(k)}(\mathbf{x}_t, \mathbf{u}_{t:t+k-1}) \tag{5}$$

Stability requires bounded error growth over K steps.

III. STABILITY ANALYSIS

A. Experimental Setup

We compare four PINN architectures:

- **Baseline**: Monolithic 5-layer MLP
- **Modular**: Separate translation/rotation subnetworks
- **Fourier**: Periodic encoding of angular states
- **Proposed**: Curriculum-trained monolithic

All share identical physics constraints; only architecture differs.

[1]S. Chatterjee is with [Department], [University], [Address]. email@institution.edu

TABLE I
SINGLE-STEP VS. 100-STEP PERFORMANCE

Model	1-Step MAE z (m)	ϕ (rad)	100-Step MAE z (m)	ϕ (rad)
Baseline	0.087	0.0008	1.49	0.018
Modular	0.041	0.0005	30.0	0.24
Fourier	**0.009**	**0.0001**	5.2M	8,596
Proposed	0.026	0.0002	**0.029**	**0.001**

B. Main Result

Table I reveals inverse correlation between single-step accuracy and autoregressive stability.

C. Failure Mode I: Gradient Decoupling

The modular architecture separates:

- Translational module: predicts z, v_z
- Rotational module: predicts $\phi, \theta, \psi, p, q, r$

This breaks physical coupling in (3). During autoregressive rollout, errors in ϕ, θ (rotation module) cause thrust projection errors in $\ddot{z}$ (translation module), but gradients do not flow between modules to enable coordinated correction.

D. Failure Mode II: Fourier Extrapolation

Fourier encoding: $\gamma(\theta) = [\sin(\omega_k \theta), \cos(\omega_k \theta)]_{k=1}^{K}$

For high frequencies ω_K:

$$\|\gamma(\theta + \epsilon) - \gamma(\theta)\| \propto \omega_K |\epsilon| \tag{6}$$

Small state drift causes large feature-space discontinuities. During rollout, this creates exponential feedback: drift $\rightarrow$ feature jump $\rightarrow$ poor prediction $\rightarrow$ larger drift.

IV. OBSERVABILITY ANALYSIS

A. Parameter Sensitivity

From (2), the sensitivity to J_{xx}:

$$\frac{\partial \dot{p}}{\partial J_{xx}} = -\frac{\tau_x}{J_{xx}^2} + \frac{(J_{yy} - J_{zz})}{J_{xx}^2} qr \tag{7}$$

At small angles ($|\phi|, |\theta| < 20°$), angular rates $|q|, |r| < 0.5$ rad/s, making the cross-coupling term in (7) negligible: $|qr| \approx O(10^{-2})$.

For mass, from (3):

$$\frac{\partial \dot{v}_z}{\partial m} = \frac{T \cos\theta \cos\phi}{m^2} \tag{8}$$

Mass sensitivity couples directly to easily-measured vertical acceleration, providing strong gradient signal even at hover.

B. Fisher Information Analysis

The Fisher Information Matrix element for parameter θ_i:

$$\mathcal{I}_{ii} = \mathbb{E}\left[\left(\frac{\partial \log p(\mathbf{y}|\boldsymbol{\theta})}{\partial \theta_i}\right)^2\right] \tag{9}$$

The Cramér-Rao bound establishes:

$$\text{Var}(\hat{\theta}_i) \geq \frac{1}{\mathcal{I}_{ii}} \tag{10}$$

When output sensitivity $\partial \mathbf{y}/\partial J_{xx}$ is small (as at small angles), $\mathcal{I}(J_{xx})$ decreases and estimation variance increases.

TABLE II
PARAMETER IDENTIFICATION RESULTS

Parameter	True	Learned	Error
Mass m	0.068 kg	0.0680 kg	0.0%
k_t	0.0100	0.0100	0.0%
k_q	7.83e-4	7.83e-4	0.0%
J_{xx}	6.86e-5	7.21e-5	5.0%
J_{yy}	9.20e-5	9.66e-5	5.0%
J_{zz}	1.37e-4	1.43e-4	5.0%

TABLE III
ABLATION STUDY: 100-STEP POSITION MAE

Configuration	MAE (m)	Improvement
Baseline	1.49	–
+ Curriculum	0.82	45%
+ Scheduled sampling	0.45	70%
+ Dropout	0.12	92%
+ Energy conservation	**0.029**	**98%**

C. Experimental Validation

Table II shows identification results matching theoretical predictions.

Strong-observable parameters (mass, k_t, k_q) achieve 0% error. Weak-observable parameters (inertias) saturate at 5%, consistent with Fisher Information bounds.

V. PROPOSED METHODOLOGY

A. Curriculum Learning

We progressively extend training rollout horizon:

$$K(e) = \begin{cases} 5 & e < 50 \\ 10 & 50 \leq e < 100 \\ 25 & 100 \leq e < 150 \\ 50 & e \geq 150 \end{cases} \tag{11}$$

B. Scheduled Sampling

Replace ground truth with predictions with probability $p(e)$ increasing from 0 to 0.3 over training, bridging train-test distribution gap.

C. Physics-Consistent Regularization

Energy conservation:

$$\mathcal{L}_{\text{energy}} = \left(\frac{dE}{dt} - P_{\text{thrust}} - P_{\text{torque}} + P_{\text{drag}}\right)^2 \tag{12}$$

Temporal smoothness:

$$\mathcal{L}_{\text{smooth}} = \sum_i \text{ReLU}\left(\left|\frac{d\hat{x}_i}{dt}\right| - v_{\max,i}\right)^2 \tag{13}$$

D. Results

Table III shows ablation results. All components necessary; full combination achieves 51× improvement.

VI. MODEL MISMATCH ANALYSIS

A. Aggressive Trajectory Experiment

To improve inertia observability, we generated trajectories with ± 45–60° attitudes to excite cross-coupling terms in (7).

Expected: Stronger qr terms $\rightarrow$ better J_{xx} gradients $\rightarrow$ lower error.

Observed: Inertia errors *increased* from 5% to 46%.

B. Root Cause

The simulator uses linearized drag assumptions:

$$F_{\text{drag}} = c_d \mathbf{v}|\mathbf{v}| \tag{14}$$

At large angles, nonlinear aerodynamics (blade flapping, gyroscopic effects) dominate. The PINN learns "effective" parameters that fit the invalid high-angle data but degrade prediction in the valid operating envelope.

C. Implication

Increased excitation improves observability only when model fidelity matches the operational regime. For PINN-based system identification:

$$\text{Identification accuracy} \propto \min(\text{Observability}, \text{Model fidelity}) \tag{15}$$

VII. CONCLUSIONS

We presented systematic analysis of PINN-based quadrotor system identification, addressing autoregressive stability and parameter observability. Key findings:

1) Single-step accuracy does not predict autoregressive stability; modular and Fourier architectures can destabilize rollouts by 10^2–$10^6\times$.
2) Parameter identification accuracy is bounded by Fisher Information-based observability limits (5% for inertias at small angles).
3) Curriculum learning with scheduled sampling achieves $51\times$ stability improvement while maintaining accurate identification.
4) Aggressive excitation without matched model fidelity degrades rather than improves identification.

Future work includes real-world validation and MPC integration.

REFERENCES

[1] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019.