Sparse Identification of Turbulence State Dynamics for Fusion Plasmas

Lucas Perrier

Supervisors: Mathieu Seurin   Alessandro Biancalani

Introduction & Motivation

Turbulence strongly impacts confinement and transport in tokamak plasmas. Reduced dynamical models derived from gyrokinetic theory capture key mechanisms such as the interaction between drift waves and zonal flows. Chen, Lin, and White proposed a low-dimensional nonlinear model exhibiting fixed points, limit cycles, and chaos [1].

Here, we apply Sparse Identification of Nonlinear Dynamical Systems (SINDy) [2] to the Chen–Lin–White (CLW) model, first in an idealized oracle-derivative setting, and then directly from noisy state measurements, to assess robustness of equation discovery.

Dynamical System Equations

The Chen–Lin–White reduced model governs the nonlinear interaction between a coherent drift wave, its sidebands, and a zonal flow. After normalization, the system is given by [1]:

$$\dot{P} = P - 2ZS\cos C, \qquad \dot{S} = -\Gamma_d S + ZP\cos C,$$
$$\dot{Z} = -\gamma_z Z + 2PS\cos C, \qquad \dot{C} = \delta - \frac{PZ}{S}\sin C$$

Here, P is the pump drift-wave amplitude, S the sideband amplitude, Z the zonal flow amplitude, and C the relative phase. The parameters Γ_d, γ_z, and δ represent normalized damping and frequency mismatch terms.

Data and Methods

Data Generation. Training data are generated from the Chen–Lin–White (CLW) reduced turbulence model using 250 short trajectories. Each trajectory is simulated over a time horizon $T = 5.0$ with timestep $\Delta t = 0.01$ (501 samples). Initial conditions (P_0, S_0, Z_0, C_0) are drawn from $\mathcal{U}([0.5, 2.0]^4)$ and parameters are fixed to $(\Gamma_d, \gamma_z, \delta) = (2.0, 0.8, 2.0)$. In the idealized benchmark, time derivatives are obtained directly from the governing equations.

SINDy Identification. SINDy identifies governing equations by solving

$$\dot{\mathbf{x}} \approx \Theta(\mathbf{x})\Xi,$$

where $\Theta(\mathbf{x})$ contains candidate nonlinear functions. We use a physics-informed library including (P, S, Z), $\sin C$, $\cos C$, and known CLW interaction terms. Sparse coefficients are obtained via Sequentially Thresholded Least Squares (STLSQ).

Results

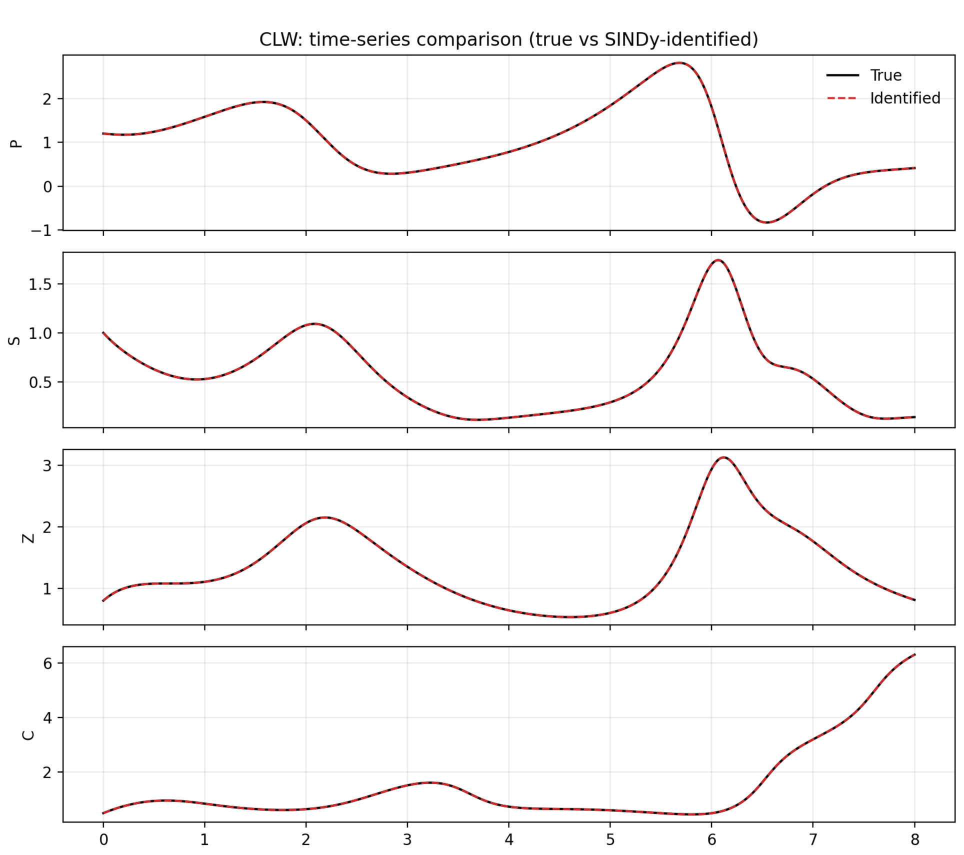

Figure 1: Short-horizon time-series comparison between the true CLW system (black) and the SINDy-identified model (red dashed) using oracle derivatives. Exact overlap demonstrates algebraic recovery of the governing equations.

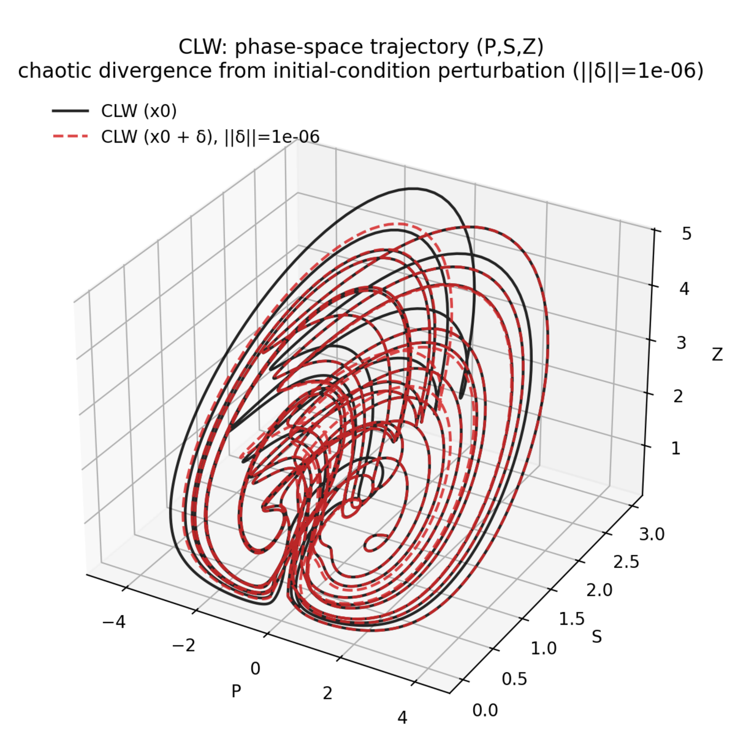

Figure 2: Phase-space trajectories in (P, S, Z) space. The SINDy model is initialized with a small phase perturbation, illustrating chaotic sensitivity rather than model error.

Justification of short-horizon evaluation. The phase-space divergence demonstrates that the CLW system exhibits sensitive dependence on initial conditions. Even with an exact governing model, long-horizon trajectory prediction is therefore fundamentally unreliable. For this reason, we focus on short-horizon trajectory overlays as the appropriate validation metric for equation discovery, where agreement reflects correct recovery of the underlying dynamics rather than incidental long-term forecasting accuracy.

Robustness to Measurement Noise

We next identify the CLW system end-to-end from noisy state measurements. Time derivatives are estimated internally using smoothed finite differences.

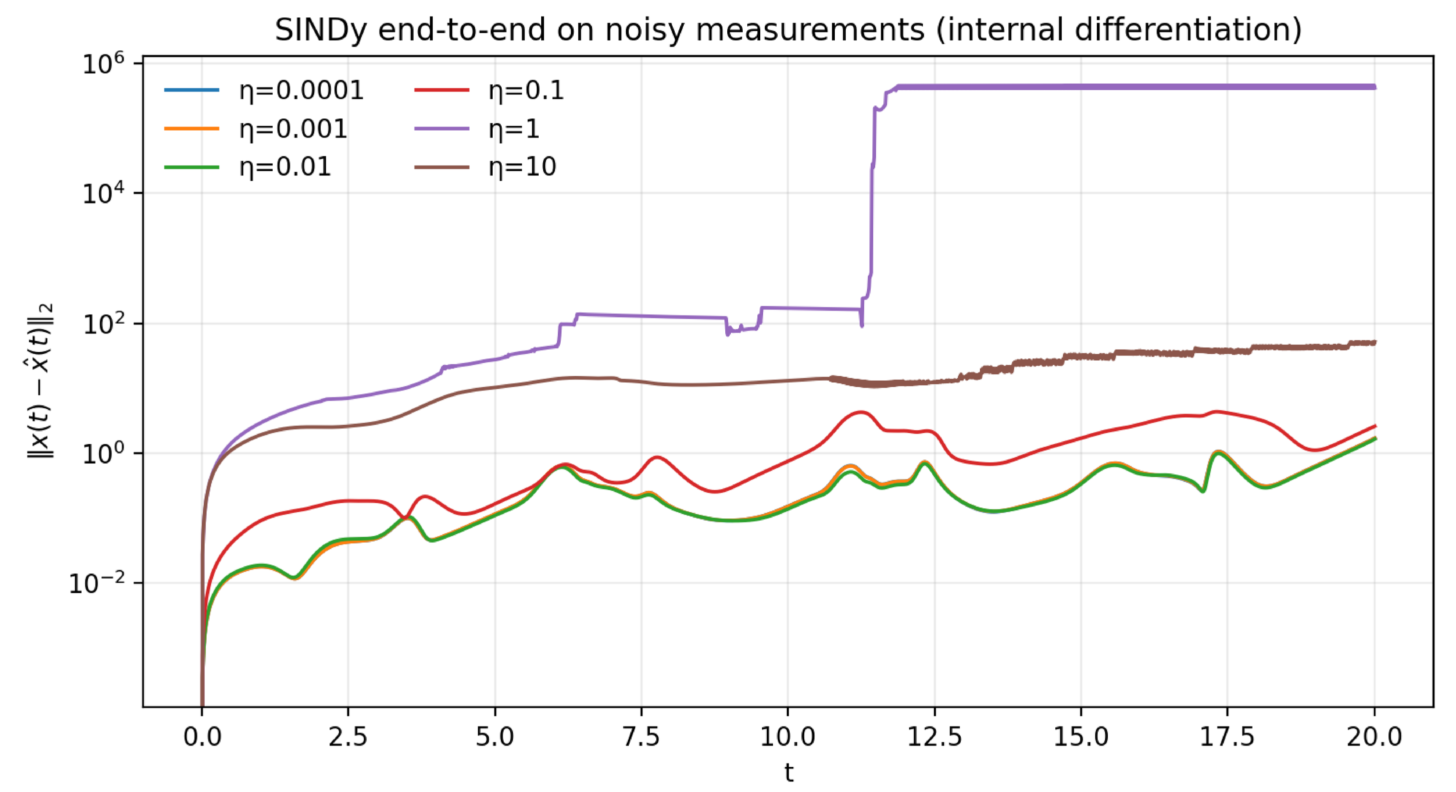

Figure 3: Trajectory error $\|\mathbf{x}(t) - \hat{\mathbf{x}}(t)\|_2$ versus time for increasing noise levels η. Error growth increases smoothly with noise.

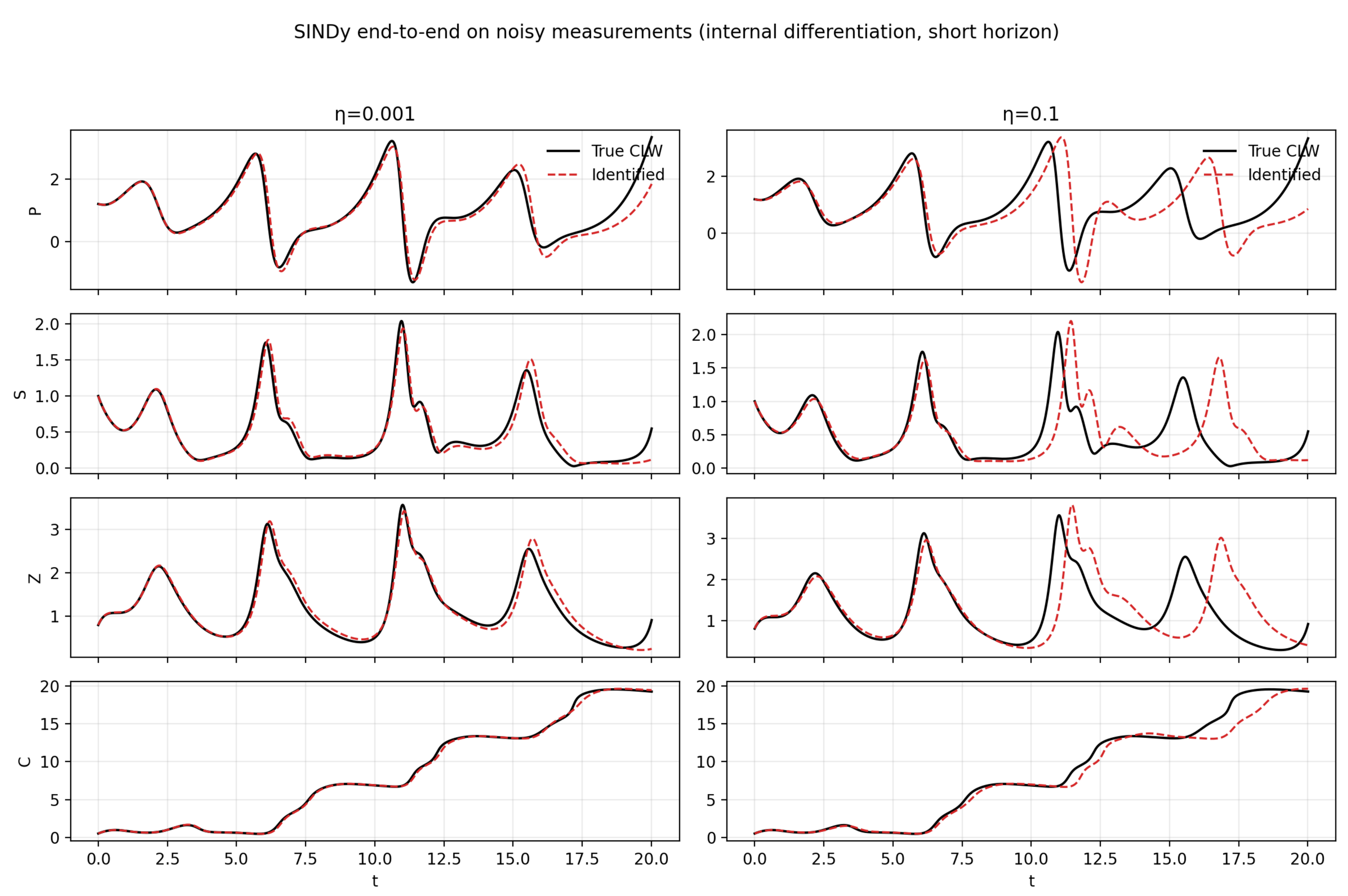

Figure 4: Short-horizon trajectories identified from noisy measurements. Left: $\eta = 10^{-3}$. Right: $\eta = 10^{-1}$. Dominant dynamics are preserved at low noise.

Assumptions and Limitations

- **Derivatives:** Oracle derivatives are used in the benchmark; noisy experiments rely on internal numerical differentiation.
- **Library:** Candidate functions are chosen to include the true CLW interaction terms.
- **Validation:** Long-horizon behavior is assessed qualitatively due to chaos.

Conclusions

SINDy exactly recovers the CLW governing equations in an idealized oracle-derivative setting. When applied end-to-end to noisy measurements, identification accuracy degrades smoothly with increasing noise while preserving short-time dynamics at low noise levels. These results highlight both the potential and limitations of data-driven equation discovery in chaotic plasma systems.

References

[1] Liu Chen, Zhihong Lin, and Roscoe White. Excitation of zonal flow by drift waves in toroidal plasmas. *Physics of Plasmas*, 7(8):3129–3132, 2000.

[2] Steven L. Brunton, Joshua L. Proctor, and J. Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15):3932–3937, 2016.