Doctoral Program in Electrical, Electronics and Communications Engineering
(XXXVIII cycle)

Physics-informed identification of nonlinear dynamical systems

A unified framework combining domain knowledge and data-driven sparse modeling

Cesare Donati

Politecnico di Torino
2026

Abstract

This thesis investigates the problem of physics-informed identification of nonlinear dynamical systems, proposing a unified framework that integrates partial physical knowledge with data-driven models to achieve interpretable, reliable, and scalable identification. The research focuses on the explicit combination of known physical equations with a corrective black-box component that accounts for unmodeled dynamics. In this context, the proposed methodology builds upon the notion of off-white models. These correspond to grey-box representations where part of the model structure is derived from first principles and physics, while unknown parameters must be identified from data.

The first contribution is the formulation of a *multi-step identification framework* in which the cumulative prediction error over extended horizons is minimized. Unlike classical single-step identification approaches, which often yield models that perform well only for short-term predictions, the proposed multi-step formulation provides consistency and reliability over longer horizons, a critical aspect in control and forecasting applications.

A second contribution is the introduction of a *sparse black-box augmentation* strategy that complements the known physical dynamics. The black-box component, initially expressed as a sparse combination of basis functions, captures only the

missing dynamics required to model discrepancies. This formulation enables the joint estimation of physical parameters and residual dynamics, preventing the bias typically introduced by alternative identification procedures. Theoretical results are established, providing explicit bounds on the parametric estimation error and conditions for maximum sparsity recovery, thus ensuring both interpretability and accuracy.

The framework is then extended to address the challenge of *non-uniform observations*, a frequent feature of real-world data due to missing samples, multiple experimental runs, or temporally aggregated measurements. New formulations and bounds are derived to quantify the effect of these irregularities on parameter estimation, demonstrating the robustness of the method across heterogeneous data collection settings.

To overcome the limitations associated with the manual selection of basis functions, a *kernel-based extension* is then introduced. By embedding the model residuals in a reproducing kernel Hilbert space, the framework enables a nonparametric, data-adaptive correction of the nominal physical model. This extension preserves the interpretability of the physical parameters while removing the dependence on predefined function dictionaries. The kernel-based formulation is further developed in the state-space setting, integrating kernel approximations with state reconstruction techniques such as unscented Kalman filtering and smoothing, resulting in enhanced predictive and simulation accuracy.

The proposed methods are empirically validated through a range of representative benchmarks and real-world-inspired case studies, including spacecraft inertia identification, cascade tank systems, continuous stirred-tank reactors, and ecological population dynamics, demonstrating robustness, interpretability, and improved long-horizon predictive performance under realistic data conditions.

Altogether, this thesis contributes a general and theoretically grounded approach to physics-informed identification, spanning sparse and kernel-based regularization, multi-step optimization, and non-uniform data handling. The proposed framework bridges the gap between traditional system identification and modern machine learning paradigms, establishing new foundations for a structured and interpretable modeling of nonlinear dynamical systems.