

SINDy-AD: Discovering governing equations from unstructured data using automatic differentiation and sparse regression

Bernat Font*, Mars Gao†, Nathan Kutz†

Discovering nonlinear dynamical systems from raw data allows modelling and predicting physical phenomena with limited a-priori knowledge of the system. Enabled by recent advances in scientific machine learning, these models provide a descriptive picture of hidden physics which might be intractable otherwise. In this context, the sparse identification of nonlinear dynamics (SINDy)¹ method has shown great success in a wide range of applications, and it has been extended to improve noise robustness² and to provide uncertainty quantification³. However, its applicability has remained limited to space-time structured data because of the traditional numerical schemes, such as finite differences, used to compute the candidate *library* terms. In this work, we present a SINDy modelling approach which leverages neural networks and automatic differentiation (AD) to address the unstructured (mesh-free) data problem, namely SINDy-AD. The framework uses a neural network as a surrogate model of a given unstructured dataset. Once the neural network is trained, the library of candidate derivative terms is generated via AD. SINDy then applies sparse regression on the library matrix to find the dominant terms of the system, as depicted in Figure 1. Tests are conducted on a set of canonical 1D systems including viscous Burgers equation, heat equation, and Korteweg–De Vries equation, and a 2-component reaction-diffusion equation. Using this approach, we find that SINDy-AD is able to discover nonlinear dynamical systems from scarce mesh-free data with up to 50% background noise, improving on previous related works^{4,5}. In this method, the neural network acts as a smooth continuous representation of the mesh-free data, and, together with an Ensemble-SINDy regression⁶, noise and data scarcity are successfully addressed. On the other hand, systems containing very high-order derivatives such as the Kuramoto–Sivashinsky equation have challenged the proposed approach. This finding demonstrates the limitations of AD in recursively computing high-order derivatives, and points towards future exploration of using AD based on polynomial expansion.

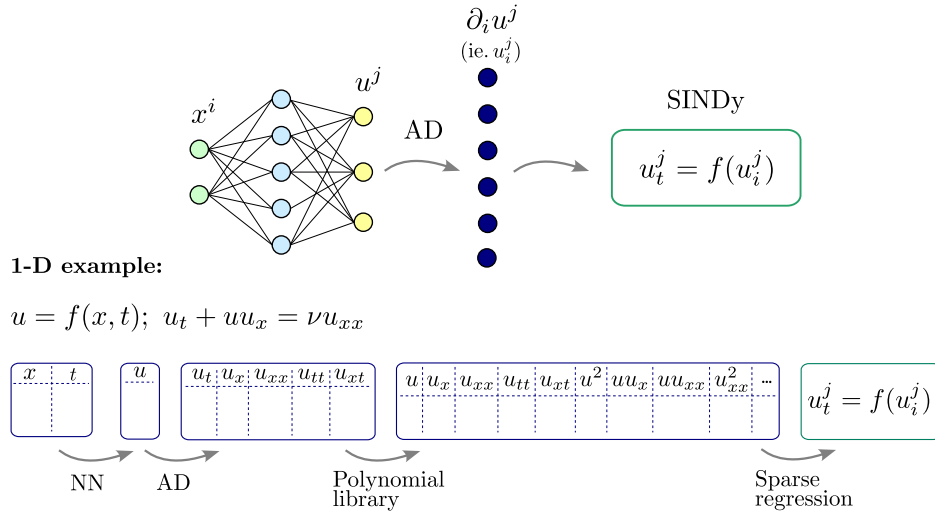


Figure 1: SINDy-AD workflow.

*Faculty of Mechanical Engineering, Delft University of Technology, Mekelweg 2, 2628 CD Delft, Netherlands

†Mechanical Engineering and Applied Mathematics, University of Washington, Seattle, WA 98195, USA

¹Brunton et al., *Proceedings of the National Academy of Sciences* **478**, 3932-3937 (2016)

²Fasel et al., *Proceedings of the Royal Society A* **478**, 20210904 (2022)

³Guo and Kutz, *Proceedings of the Royal Society A* **480**, 20230506 (2024)

⁴Both et al., *Journal of Computational Physics* **428**, 109985 (2021)

⁵Chen et al., *Nature Communications* **12**, 6136 (2021)

⁶Guo and Kutz, *Proceedings of the Royal Society A* **480**, 20230506 (2024)