Automated discovery of structural differences in dynamical systems

Sedighe Raeisi *Osnabrück University* `sedighe.raeisi@uos.de`
Pascal Nieters *Osnabrück University* `pascal.nieters@uos.de`
Sebastian Musslick *Osnabrück University, Brown University*
`sebastian.musslick@uos.de`

Abstract

A major challenge in the natural and engineering sciences is uncovering dynamical systems that explain observed natural phenomena. Data-driven equation discovery has emerged as a powerful approach to automate this process. Existing equation discovery methods, however, typically pool data across instances of systems they seek to explain, ignoring structural differences or reducing them to parameter variations within the same equation. Yet, different instances of the same physical or biological system may be governed by different equations. The challenge is to uncover structural components that are universal across systems while distinguishing them from individual-specific variations. To capture structural differences across systems, we introduce a Bayesian hierarchical approach to equation discovery. We demonstrate across case studies in physics, ecology, and neuroscience that this method recovers mechanisms of the data-generating processes while capturing structural variability across systems. By explicitly modeling structural variability in equations, it establishes a foundation for data-driven automated model discovery at the population level, providing scientists with a tool to separate universal principles from individual differences.

Keywords: data-driven model discovery, symbolic regression, automated scientific discovery

1 Introduction

A central challenge in AI for science is the data-driven discovery of scientific models—for example, dynamical systems that characterize natural phenomena. Equation discovery refers to a family of methods that aim to automatically infer interpretable equations from noisy data [Langley, 1987, Džeroski et al., 2007, Musslick et al., 2025]. A major obstacle for such approaches is the variability across individual systems. Here, an “individual” denotes a specific instance of a system—such as a circuit, a cell, or an ecological community—that is governed by universal principles yet exhibits structural idiosyncrasies. For instance, while nerve cells share common biophysical mechanisms, different instances may recruit additional processes that uniquely shape their firing dynamics.

Most approaches to data-driven equation discovery impose a single model structure for the same type of system. If data is pooled across different instances of the same system (e.g., populations of the same species in different habitats)

AI in Science (AIS), 2025, Copenhagen, Denmark

AI in Science
Conference 2025
EU2025DK

existing algorithms ignore structural variability across system instances or reduce differences to parameter differences [Džeroski et al., 2007]. This risks oversimplification on the one hand and overfitting on the other. To address this gap, we introduce a hierarchical Bayesian framework for data-driven equation discovery that models variability in the structure of dynamical systems while enforcing sparsity for interpretability. Our method recovers universal mechanisms alongside individual-specific variations, providing an automated tool for disentangling universality and individuality in scientific models.

Related Work. Data-driven equation discovery (or symbolic regression) seeks to discovery equations from data—for example, rediscovering physical laws from measurements of physical quantities [Langley, 1987, Džeroski et al., 2007]. A widely used method is Sparse Identification of Nonlinear Dynamics (SINDy) [Brunton et al., 2016], which casts equation discovery as sparse regression over a library of (non-linear) candidate terms, yielding interpretable equations. To improve inference under noisy data regimes, a Bayesian extension of SINDy places sparsity-inducing priors (e.g., the Horseshoe [Carvalho et al., 2009, Piironen and Vehtari, 2017]) on regression terms, enabling uncertainty quantification [Hirsh et al., 2022]. Yet, most approaches to equation discovery—including Bayesian variants—represent each system with a single set of equations, treating variability across systems as noise. This conflates measurement noise with true structural differences across systems. A hierarchical Bayesian approach would allow candidate terms to vary across individual systems while sharing statistical strength at the population level. Such methods are common in fields like cognitive modeling [Wiecki et al., 2013] but have not been applied to equation discovery. Here, we extend Bayesian SINDy with hierarchical priors, enabling discovery of universal system components alongside individual variations.

2 Methods

SINDy recovers dynamical equations from time-series data by expressing derivatives $\dot{X}$ as sparse linear combinations of non-linear terms $\Theta(X)\Xi$ [Brunton et al., 2016]. Sparsity in Ξ identifies a small subset of candidate terms needed to fit the time series. Bayesian SINDy extends this approach to noisy data, inferring posterior distributions over coefficients and inducing sparsity via priors such as the regularized Horseshoe [Hirsh et al., 2022]. However, this approach assumes a single model structure if data is pooled from multiple instances of a system, treating system variability as parameter noise. To capture structural variability across system instances ("individuals"), we propose hierarchical Bayesian SINDy as illustrated in Figure 1. Each individual i is modeled as

$$\dot{X}_i = \Theta(X_i)\Xi_i + \eta Z_i, \tag{1}$$

where coefficients Ξ_i are drawn from population-level distributions over candidate equation terms. Ξ_{mean} follows a Horseshoe prior and Ξ_{std} follows a Half-Normal distribution. Observation noise is represented as ηZ, with η denoting the noise magnitude and Z a random matrix drawn from the noise distribution.

This hierarchical prior allows terms to be present in some individuals but absent in others, while Horseshoe priors enforce sparsity at both individual and population levels. The resulting posterior distributions capture universal (shared) structure across systems, individual-specific variations, and even multimodal structures (e.g., subpopulations governed by distinct dynamics). We approximate these posteriors using MCMC inference. Hierarchical Bayesian SINDy thus disentangles noise-driven uncertainty from genuine structural differences, providing a framework for equation discovery in populations of systems.

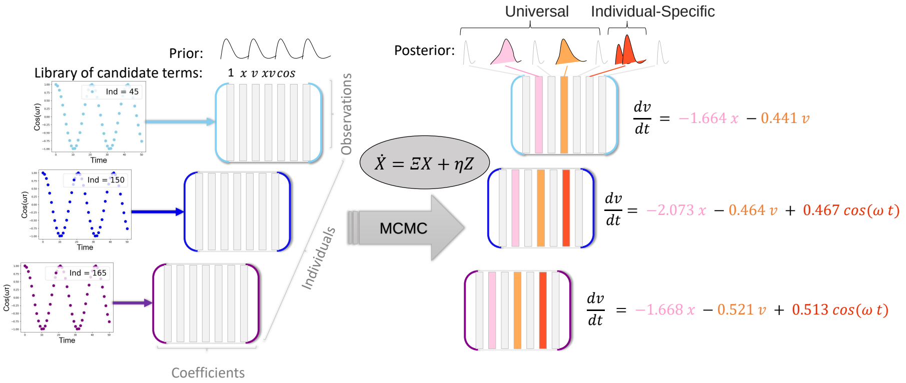


Figure 1: The schematic process of Bayesian hierarchical SINDy. The posterior distribution of coefficients of the library of candidate terms, Ξ is computed using MCMC as a sampling method, and the sparsity is leveraged by using the Horseshoe distribution as the prior of the mean of each ξ_k.

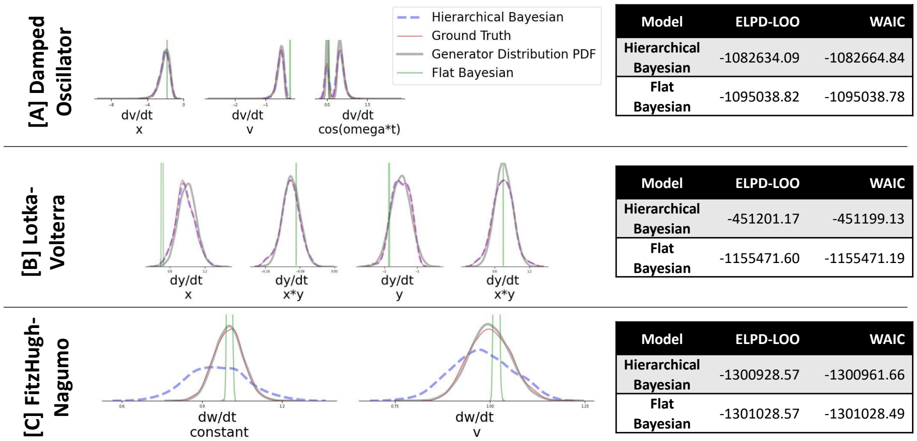
3 Experiments

To evaluate our approach, we generated synthetic datasets from three established dynamical systems in physics, neuroscience, and ecology: a damped forced harmonic oscillator [Brunton et al., 2016], the FitzHugh–Nagumo model [Prokop and Gelens, 2024], and the Lotka–Volterra system [Fasel et al., 2022] (Table 1). For each system, multiple instances were instantiated by sampling equation term coefficients from normal distributions, which in turn produced ground-truth distributions over candidate terms. This allowed us to specify which equation terms were universal (present across all instances) and which were individual-specific.

We simulated time series data from these systems and applied hierarchical Bayesian SINDy to recover governing equations. For comparison, we also fit the flat Bayesian SINDy model proposed by [Hirsh et al., 2022], which assumes a single structure across instances. Model comparison based on standard metrics (ELPD-LOO and WAIC; [Vehtari et al., 2017, 2024]) consistently favored the hierarchical approach across all systems. Moreover, posterior distributions over candidate terms closely matched the ground-truth distributions, demonstrating that hierarchical modeling recovers both universal and individual-specific components, whereas flat models conflate structural variability with noise.

Table 1: Experimental systems used to evaluate hierarchical Bayesian SINDy: [A] damped oscillator, [B] Lotka-Volterra, and [C] FitzHugh-Nagumo.

	System Equations	**Coefficient Distributions**	**Candidate Terms**	**Samples**
[A]	$\dot{x} = v$ $\dot{v} = -\frac{k}{m}x - \frac{c}{m}v + \frac{F}{m}\cos(\omega t)$	$k, c, F \sim$ hierarchical priors	$\{Const., x, v, \cos(\omega t), x^2, v^2, xv\}$	500 individuals, 1001 steps each
[B]	$\dot{X} = \alpha X - \beta XY$ $\dot{Y} = \delta XY - \gamma Y$	$\alpha, \beta, \gamma, \delta \sim$ hierarchical priors	$\{Const. X, Y, XY$ $X^2, Y^2\}$	100 individuals, 2002 steps each
[C]	$\dot{v} = v - \frac{v^3}{3} - w - I$ $\dot{w} = ab_0 + ab_1 v - aw$	$a, b_0, b_1, I \sim$ hierarchical priors	$\{Const., v, v^3, w, v^2, w^2, vw\}$	600 individuals, 1001 steps each



[A] Damped Oscillator

Model	ELPD-LOO	WAIC
Hierarchical Bayesian	-1082634.09	-1082664.84
Flat Bayesian	-1095038.82	-1095038.78

[B] Lotka-Volterra

Model	ELPD-LOO	WAIC
Hierarchical Bayesian	-451201.17	-451199.13
Flat Bayesian	-1155471.60	-1155471.19

[C] FitzHugh-Nagumo

Model	ELPD-LOO	WAIC
Hierarchical Bayesian	-1300928.57	-1300961.66
Flat Bayesian	-1301028.57	-1301028.49

Figure 2: Model fit and recovery of equation term distributions across use cases, for both flat and hierarchical Bayesian SINDy.

4 Conclusion

We introduced hierarchical Bayesian SINDy for automated discovery of dynamical systems that accounts for structural variability across individuals. Across three domains—physics, ecology, and neuroscience—our method recovered the correct governing terms while capturing full distributions of coefficients. Unlike prior approaches, it separates universal mechanisms from individual-specific variations rather than conflating them with noise. This enables population-level representations of dynamics and provides an approach for uncovering structural variability in equations describing complex systems.

5 Acknowledgements

S.R. received support from the Alexander von Humboldt Foundation. The calculations were performed on DFG-funded cluster, 456666331. P.N received support from Volkswagen Foundation.

References

S. L. Brunton, J. L. Proctor, and J. N. 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.

C. M. Carvalho, N. G. Polson, and J. G. Scott. Handling sparsity via the horseshoe. In *Artificial intelligence and statistics*, pages 73–80. PMLR, 2009.

S. Džeroski, P. Langley, and L. Todorovski. Computational discovery of scientific knowledge. In *Computational discovery of scientific knowledge: introduction, techniques, and applications in environmental and life sciences*, pages 1–14. Springer, 2007.

U. Fasel, J. N. Kutz, B. W. Brunton, and S. L. Brunton. Ensemble-sindy: Robust sparse model discovery in the low-data, high-noise limit, with active learning and control. *Proceedings of the Royal Society A*, 478(2260):20210904, 2022.

S. M. Hirsh, D. A. Barajas-Solano, and J. N. Kutz. Sparsifying priors for bayesian uncertainty quantification in model discovery. *Royal Society open science*, 9(2):211823, 2022.

P. Langley. *Scientific discovery: Computational explorations of the creative processes*. MIT press, 1987. ISBN 0-262-62052-9.

S. Musslick, L. K. Bartlett, S. H. Chandramouli, M. Dubova, F. Gobet, T. L. Griffiths, J. Hullman, R. D. King, J. N. Kutz, C. G. Lucas, S. Mahesh, F. Pestill, S. J. Sloman, and W. R. Holmes. Automating the practice of science: Opportunities, challenges, and implications. *Proceedings of the National Academy of Sciences*, 122(0):e2401238121, 2025.

J. Piironen and A. Vehtari. Sparsity information and regularization in the horseshoe and other shrinkage priors. *Electronic Journal of Statistics*, 11(2): 5018, 2017.

B. Prokop and L. Gelens. From biological data to oscillator models using sindy. *iScience*, 27(4):109316, 2024.

A. Vehtari, A. Gelman, and J. Gabry. Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and computing*, 27 (5):1413–1432, 2017.

A. Vehtari, D. Simpson, A. Gelman, Y. Yao, and J. Gabry. Pareto smoothed importance sampling. *Journal of Machine Learning Research*, 25(72):1–58, 2024.

T. V. Wiecki, I. Sofer, and M. J. Frank. Hddm: Hierarchical bayesian estimation of the drift-diffusion model in python. *Frontiers in neuroinformatics*, 7:14, 2013.