Research Statement

Masih Haseli

Rigorous Real-Time Learning, Prediction, and Control of Complex Systems

Most dynamical systems in the natural and engineered world are inherently nonlinear, and for decades, researchers have pursued the goal of translating nonlinear system theory into practical methodologies for analysis, control, and design. Although the theoretical foundations of nonlinear system theory are deep and well established, realizing their full potential in applications remains an ongoing challenge. This difficulty stems from the fundamental distinction between linear and nonlinear frameworks: linear system theory is built upon regular algebraic structures that align naturally with computational algorithms, whereas nonlinear system theory is expressed in geometric terms that resist direct numerical implementation. Recent advances in data-driven modeling, operator-theoretic approaches, and machine learning offer a new perspective on this classical divide. By integrating rigorous system-theoretic principles with modern data-driven techniques, we can begin to develop computational frameworks that both retain theoretical fidelity and enable scalable analysis and control of complex nonlinear systems.

My research broadly aims to address the problem above. My goal is to achieve **real-time data-driven identification, prediction, and control of broad classes of nonlinear systems with theoretical guarantees**. Towards this goal, I have developed scalable methods for learning and understanding complex nonlinear behavior that draw from control theory, operator theory, and network science and all are backed by rigorous theory. My main directions towards this goal are **(*i*) developing operator-theoretic frameworks for nonlinear systems, (*ii*) building fast algorithms with rigorous guarantees, and (*iii*) applying these algorithms in real-time applications.** Figure 1 demonstrates the details of my research.

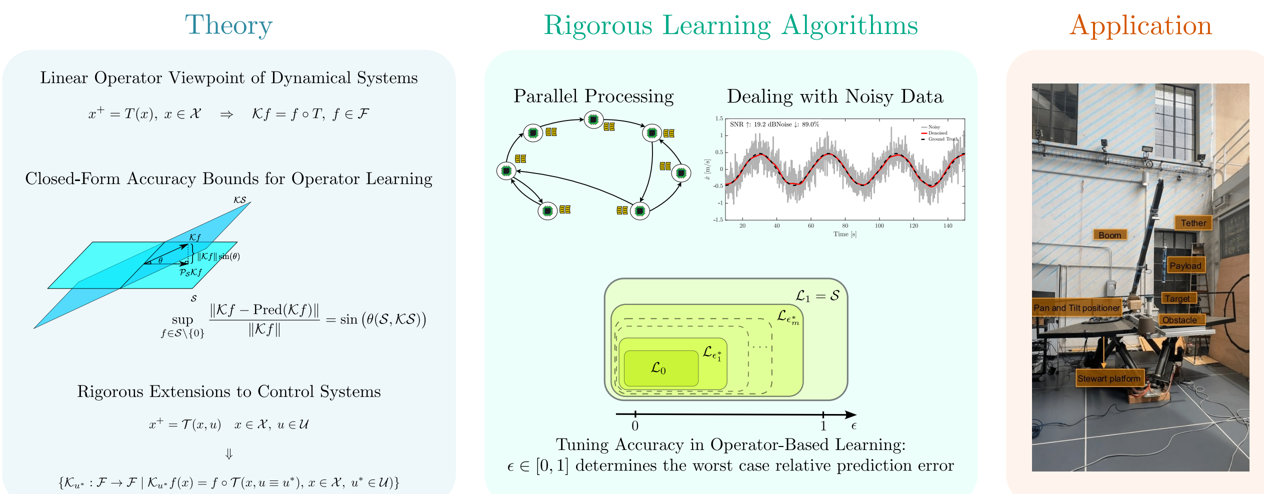

Figure 1: The interplay between operator viewpoint of dynamical systems, fast and rigorous learning algorithms, and applications.

1. Theory with Practical Foresight: Linear Operator Viewpoint

Ever since the groundbreaking work of Isaac Newton, theoretical structures have played an indispensable role in both science and engineering. Yet, the structure of a theory often determines the extent to which it can be realized in practice. Nonlinear system theory, despite its conceptual depth and elegance, poses substantial challenges for computational implementation due to its geometric nature.

A promising strategy to overcome this challenge is to describe nonlinear dynamics through linear operators acting on vector spaces of functions, rather than directly through nonlinear differential or difference equations. This operator-theoretic perspective offers a highly regular algebraic structure that aligns naturally with computational methods, providing a bridge between nonlinear dynamics and algorithmic implementation.

However, this connection is far from straightforward. The operator representations are typically infinite dimensional, and finite-dimensional approximations must be constructed with care to retain essential dynamical properties and theoretical guarantees. In particular, understanding and quantifying **accuracy bounds** for such approximations is critical to ensure that computational models faithfully represent the underlying nonlinear phenomena. The central challenge, and the focus of my research, is to develop **rigorous and computationally tractable frameworks** that achieve provable accuracy while enabling scalable analysis and control.

Studying Complex Systems via Koopman Operator. Studying complex systems through geometric means becomes immediately difficult in high-dimensional state spaces. This issue can be addressed by representing the nonlinear dynamics with a linear operator (known as the *Koopman operator*) acting on a vector space of functions. This approach yields a regular algebraic structure that allows one to bypass the geometric complexities of the state space. The spectral properties of the Koopman operator enable a systematic study of the system's behavior. Specifically, the eigenfunctions of the Koopman operator have the following property: their values evolve **linearly** in time along the system trajectories (cf. Figure 2). **I have established theoretical guarantees, in the form of necessary and sufficient conditions, that enable the accurate identification of Koopman eigenfunctions and their associated eigenvalues [1, 2]. These guarantees ensure a precise decomposition and prediction of the system's behavior. Verifying them is straightforward, requiring only a comparison between two differently projected models obtained through small regression problems.**

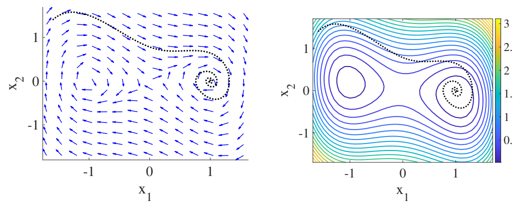

Figure 2: A planar system's vector field and a trajectory (left) and the level sets of an approximated Koopman eigenfunction (right). The eigenfunction's value **evolves linearly in time** on the trajectories allowing for **linear computations for nonlinear prediction**.

Providing Tight and Objective Accuracy Bounds on Linear Koopman-based Models. In data-driven modeling, providing accuracy bounds is essential for ensuring both performance and safety. Traditional accuracy measures for Koopman-based models predominantly depend on the quality of a basis for the corresponding finite-dimensional space. An example of this is the residual error in the widely popular Extended Dynamic Mode Decomposition (EDMD) method[1]. However, such accuracy measures only evaluate a finite number of functions, ignoring the uncountably many others present in the space. This oversight can lead to skewed conclusions based on coordinate selection rather than genuine model accuracy. To address this crucial issue, I have developed the **"temporal consistency index"**, an objective accuracy measure that provides a **tight upper bound on the worst-case prediction error over the entire vector space** [3]. The temporal consistency index **can be calculated in closed form** and has significant implications for **performance guarantees, and safety evaluation. Moreover, the consistency index can be used as a loss function in optimization-based learning. Interestingly, this is equivalent to a robust minimax optimization where we have a closed-form for the max part.** In addition, I have developed a generalization of temporal consistency index termed **"invariance proximity"** suited for **general inner-product spaces** [4, 5]. This measure has also been used in efficient algebraic algorithms for system identification [6].

Rigorous Extension of the Koopman Theory to Control: Koopman Control Family and Universal Forms. Control applications are prevalent across a wide range of engineering disciplines. Consequently, any theoretical framework we construct must adequately address control systems. However, extending the Koopman operator viewpoint to control systems with rigorous theory has proven more challenging than expected, primarily due to the inherent difference between the concepts of 'input' and 'state'. While the state emerges from the system's dynamics and is influenced by the history of the system's behavior, the input in an open-loop system is arbitrary and does not conform to a dynamic rule. This distinction has led to numerous adaptations of the Koopman operator theory for control, many of which depend on the unique structure of the specific system. This approach often results in a partial loss of structure and necessitates case-by-case studies that are challenging to implement. To address this issue, I have developed a comprehensive mathematical framework termed **"Koopman Control Family"** which fully encapsulates the behavior of general (**not necessarily control affine**) nonlinear control systems [4]. **Interestingly, this framework leads to a universal finite-dimensional form termed the *"input-state separable"* model, which takes the widely used Koopman-based linear, bilinear, and switched linear models as special cases.** Moreover, we recently showed that Koopman Control Family is equivalent to another general extension of Koopman theory to control systems [13, 14]; thus providing a unified theoretical framework for Koopman-based control.

Future Work. My theoretical contributions mentioned above open several interesting avenues of research in various domains. **I will create an interdisciplinary effort to incorporate expertise in control theory, operator theory, reinforcement learning, and formal methods to provide scalable algorithms with theoretical performance and safety guarantees for real-time deployment.**

Complex System Modeling and Analysis: I strive to develop a unified theoretical framework, suitable for computer implementation, that merges insights from differential geometry, graph theory, and model reduction. This approach aims to decompose complex systems into components with similar algebraic properties.

[1] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, "A Data-driven Approximation of the Koopman Operator: Extending Dynamic Mode Decomposition," *Journal of Nonlinear Science*, vol. 25, no. 6, pp. 1307–1346, 2015.

For instance, one could decompose complex graphs into sub-graphs with similar spatial or temporal features or can express a nonlinear vector field as a sum of distinct vector fields corresponding to specific behaviors (e.g. based on invariant set decomposition or Koopman eigenfunctions). My ultimate goal is to merge these insights with formal methods and operator theoretic frameworks to create algorithms for system analysis with rigorous accuracy and performance guarantees. Moreover, I aim to apply this framework to applications in *cyber-physical systems, biological systems, and traffic networks.*

Safety Guarantees for Autonomous Systems: Safety guarantees for autonomous systems (e.g. self-driving cars, and autonomous robots) are of utmost importance. Current safety certificates, often based on control barrier functions or reachability concepts, typically focus on nonlinear control affine systems. However, challenges arise since: (i) many systems are not control affine, (ii) system dynamics might be unknown, (iii) deriving control barrier functions is not straightforward, and (iv) finding reachable sets is computationally intensive. My objective is to integrate existing safety and reachability approaches with algebraic and operator-theoretic concepts. This integration aims to address strong nonlinearities in the model and develop fast computational methods for real-time processing while handling model and data uncertainties.

Robust Learning, Interpretable Models, and Neural-Network Pruning: While neural networks play a pivotal role in autonomous systems, significant challenges persist within dynamical systems and controls. My goal is to provide mathematically rigorous structures for neural-network based methodologies, leading to the *natural decomposition of complex phenomena into smaller, understandable subsystems.* I also aim to hone in on robust learning strategies that can navigate uncertainty and ensure the accuracy of trained networks. I also seek to explore the domain of neural network pruning. Leveraging my previous research on complex system modeling, I plan to offer accurate yet considerably more compact models for trained neural networks. This approach is especially beneficial for large language models whose inference is energy consuming.

2. Fast and Reliable Data-Driven Algorithms with Guarantees

Given the rapid development of autonomous systems, the importance of scalable algorithms that come with performance guarantees cannot be overemphasized. Building on my prior theoretical contributions, I have developed **efficient data-driven algorithms designed to discover features of underlying dynamics, all while providing rigorous theoretical guarantees.** As I will explain next, these methods are not only computationally and memory efficient but are also well-suited to parallel processing hardware such as GPUs.

Dealing with Measurement Noise in Mode Decompositions. Real-world data is generally noisy and it is crucial for our algorithms to take this into account. The commonly used EDMD method for Koopman-based identification presumes exact data and relies on data preprocessing for proper results. Moreover, EDMD operates in a lifted space created by passing the data through nonlinear functions, resulting in noise distortion. This significantly complicates the problem since the noise is *not identically distributed*. In my work [7], I address this problem by taking the noise distortion into account and create an optimization-based algorithm that recovers the correct modes from noisy data and is accompanied by theoretical guarantees. In addition, in [15], we provided a different algorithm for real-time learning and prediction of robotic systems with theoretical guarantees and hardware validation.

Using Algebraic Structure to Speed Up Data-driven Algorithms. Given their linear temporal evolution, the Koopman eigenfunctions are pivotal in numerous engineering applications, including *stability analysis of attractors, identification of invariant sets, as well as the analysis of complex systems such as fluid flows, biological systems, and traffic networks.* **The first step to enable such important applications is to correctly identify the Koopman eigenfunctions.** Building on my earlier theoretical investigations into the algebraic structures associated with the Koopman operator, I have developed the algebraic Symmetric Subspace Decomposition (SSD) algorithm that efficiently searches through a vector space of functions, a.k.a. the search space, to find the largest subspace invariant under the Koopman operator (on such subspaces linear Koopman-based models are exact). SSD adopts an algebraic technique based on successive orthogonal decompositions of the search space, establishing an efficient method (with **linear complexity in the size of the data set**) to uncover a maximal linear model for the system's dynamics that accurately explains the system components. **SSD converges in finite time and enjoys rigorous theoretical guarantees regarding the correct extraction of the system's information**.

Memory-Efficient Processing for Large and Streaming Data Sets. In many practical applications, we encounter massive data sets that require substantial memory for processing, or we deal with streaming data sets originating from real-time sampling schemes, requiring immediate processing upon receipt. To address these challenges, I have developed the Streaming SSD algorithm that starts with a small portion of the data and iteratively updates the solution as more data is received [1]. This method also handles massive data sets by processing small portions of data at a time. **This approach enables the algorithm to function within a fixed and substantially reduced memory footprint, regardless of the data set's size, making it an ideal candidate for embedded systems performing real-time computations.**

Parallel Processing. To take advantage of parallel computing hardware, I have developed an extension of the SSD algorithm to speed up the computation of Koopman-invariant subspaces and accurate lifted linear models [8, 9]. This method breaks down the subspace search problem into smaller pieces and solves them simultaneously using multiple processors communicating through a network. The processors achieve consensus on the correct solution after finite iterations. The time complexity of the Parallel SSD (P-SSD) algorithm is inversely related to the number of processors; therefore, **it gets linearly faster by increasing the number of processors (cf. Figure 3)**. In addition to enjoying all the theoretical properties of SSD, **P-SSD is compatible with a broad spectrum of communication networks (even time-varying directed networks) and is provably robust against packet drops and communication failures.**

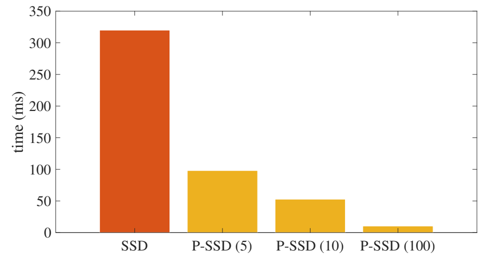

Figure 3: Time taken for the SSD [1, 2] and P-SSD [8, 9] with 5, 10, and 100 processors to accurately compute Koopman eigenfunctions and linear Koopman-based models using 5×10^5 data points taken from a polynomial system.

Algorithms with Tunable Accuracy and Expressiveness. Exact finite-dimensional Koopman invariant subspaces can encapsulate crucial information about a system, including the stability of attractors, invariant sets, and basins of attraction. However, exact finite-dimensional invariant subspaces that capture complete system information **rarely** exist. In many applications, absolute precision is not a necessity, and moderately accurate models are often sufficient. Thus, allowing minor inaccuracies in the model to capture more information or to introduce structure in the model for enhanced computational efficiency is a reasonable and widely accepted practice. To address this, I have developed the Tunable Symmetric Subspace Decomposition (T-SSD) algorithm that identifies Koopman-based models meeting a predefined level of accuracy [10, 11]. **As described in Figure 4, T-SSD grants the user full control over the model's accuracy level, with a parameter ranging from zero (representing the exact model) to one (no control over accuracy).**

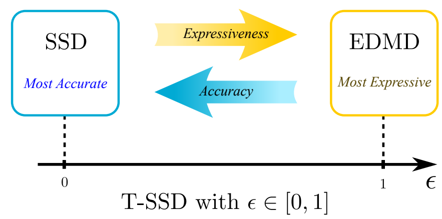

Figure 4: The parameter ϵ in T-SSD [10, 11] allows the user to balance the model's accuracy and expressiveness. T-SSD with $\epsilon = 0$ gives an exact model (equivalent to SSD [1, 2]) and with $\epsilon = 1$ gives the most expressive model (equivalent to EDMD).

Future Work. My past research paves the way for a myriad of applications, which I will outline here.

Real-Time Path Planning and Control in Unstructured Environments: My focus lies in predictive control of nonlinear systems, ensuring **real-time execution on embedded systems with theoretical performance guarantees**. I aim to build on my past research on *algebraic methods for nonlinear systems combined with geometric control and formal methods* to create software packages enabling a diverse set of applications including **robotics (particularly unstructured and soft robots), real-time path planning and control in complex and unstructured environments (e.g. self-driving cars)**.

Fast Computational Methods for Stability Analysis: A key application in control theory lies in the stability analysis of nonlinear systems and determining basins of attractions. It is well-known that Koopman eigenfunctions can be employed to derive Lyapunov functions for evaluating the stability of attractors in nonlinear systems, and this typically can be tackled in finite dimensions. However, the success of these methods largely depends on the accurate identification of appropriate eigenfunctions, a challenge I have rigorously tackled in my past research. Therefore, an efficient computational solution for the construction of Lyapunov functions is achievable. I plan to further investigate this area, aiming to introduce efficient algebraic methods to determine the stability of attractors and find their basin of attraction.

Generic Algebraic Solvers for Real-Time Nonlinear Computation: Nonlinear systems often result in optimization formulations that are non-convex, challenging to solve, and lack tight theoretical guarantees. My long-term vision is to design generic algebraic software packages that can do fast computations for a variety of tasks **(e.g., real-time modeling and decomposition, stability and reachability analysis, path planning and control) and are accompanied by rigorous theoretical guarantees.**

References

[1] **M. Haseli** and J. Cortés, "Learning Koopman Eigenfunctions and Invariant Subspaces from Data: Symmetric Subspace Decomposition," *IEEE Transactions on Automatic Control*, vol. 67, no. 7, pp. 3442-3457, 2022.

[2] **M. Haseli** and J. Cortés, "Efficient Identification of Linear Evolutions in Nonlinear Vector Fields: Koopman Invariant Subspaces," in *IEEE Conference on Decision and Control*, Nice, France, 2019, pp. 1746-1751.

[3] **M. Haseli** and J. Cortés, "Temporal Forward-Backward Consistency, not Residual Error, Measures the Prediction Accuracy of Extended Dynamic Mode Decomposition," *IEEE Control Systems Letters*, vol. 7, pp. 649-654, 2023.
IEEE Control Systems Letters Outstanding Paper Award

[4] **M. Haseli** and J. Cortés, "Modeling Nonlinear Control Systems via Koopman Control Family: Universal Forms and Subspace Invariance Proximity," *Automatica*, 2025, to appear.

[5] **M. Haseli** and J. Cortés, "Efficient Computation of Invariance Proximity: Closed-Form Error Bounds for Finite-Dimensional Koopman-Based Models," *Systems and Control Letters*, 2023, submitted.

[6] **M. Haseli** and J. Cortés, "Recursive forward-backward EDMD: guaranteed algebraic search for Koopman invariant subspaces," *IEEE Access* 13 (2025), 61006-61025.

[7] **M. Haseli** and J. Cortés, "Approximating the Koopman Operator using Noisy Data: Noise-Resilient Extended Dynamic Mode Decomposition," in *American Control Conference*, Philadelphia, PA, 2019, pp. 5499-5504.

[8] **M. Haseli** and J. Cortés, "Fast Identification of Koopman-Invariant Subspaces: Parallel Symmetric Subspace Decomposition," in *American Control Conference*, Denver, CO, 2020, pp. 4545-4550.

[9] **M. Haseli** and J. Cortés, "Parallel Learning of Koopman Eigenfunctions and Invariant Subspaces for Accurate Long-Term Prediction," *IEEE Transactions on Control of Network Systems*, vol. 8, no. 4, pp. 1833-1845, 2021.

[10] **M. Haseli** and J. Cortés, "Data-Driven Approximation of Koopman-Invariant Subspaces with Tunable Accuracy," in *American Control Conference*, New Orleans, LA, 2021, pp. 470-475.
Best Student Paper Award

[11] **M. Haseli** and J. Cortés, "Generalizing Dynamic Mode Decomposition: Balancing Accuracy and Expressiveness in Koopman Approximations," *Automatica*, vol. 153, p. 111001, 2023.

[12] L. Shi, **M. Haseli**, G. Mamakoukas, D. Bruder, I. Abraham, T. D. Murphey, J. Cortés, K. Karydis, "Koopman Operators in Robot Learning," *IEEE Transactions on Robotics*, 2024, submitted.

[13] **M. Haseli**, I. Mezić, J. Cortés, "Koopman Operator Extensions for Control: Bridging Infinite Input Sequences and Operator Families," *Proceedings of the IEEE Conference on Decision and Control*, Rio de Janeiro, Brazil, 2025, to appear.

[14] **M. Haseli**, I. Mezić, J. Cortés, "Two Roads to Koopman Operator Theory for Control: Infinite Input Sequences and Operator Families," *IEEE Transactions on Automatic Control*, 2025, submitted.

[15] S. B. Kombo, **M. Haseli**, Skylar X. Wei, J. W. Burdick, "Real-Time Learning of Predictive Dynamic Obstacle Models for Robotic Motion Planning," *Proceedings of the IEEE International Conference on Robotics and Automation*, 2026, submitted.