# ABSTRACT

## AUTOMATIC DEVELOPMENT AND ADAPTATION OF CONCISE NONLINEAR MODELS FOR SYSTEM IDENTIFICATION

SEPTEMBER 2016



WILLIAM G. LA CAVA

B.S., CORNELL UNIVERSITY

M.Eng., CORNELL UNIVERSITY

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Kourosh Danai



Mathematical descriptions of natural and man-made processes are the bedrock of science, and are used by humans to understand, estimate, predict and control the natural and built world around them. The goal of the field of system identification is to automate the inference of the true underlying dynamics of processes from measured observations. The crux of system identification is the challenge of identifying the dynamic model form (topology) in addition to its parameters. This dissertation proposes three methods designed to improve the ability of system identification to identify succinct nonlinear model structures.

The first is a model structure adaptation method (MSAM) that allows for the tuning of first principles models to increase their predictive ability while maintaining intelligibility. Model structure identification is achieved in the presence of parametric error through a novel means of estimating the gradient of model structure perturbations. I demonstrate MSAM's ability to identify underlying nonlinear dynamic processes from linear starting models in the presence of parametric uncertainty.

The second method, known as epigenetic linear genetic programming (ELGP), conducts symbolic regression without *a priori* knowledge of the form of the model or its parameters. ELGP proposes to incorporate a layer of genetic regulation into GP and to adapt it as a form of local search that tunes the resultant model structures to be accurate and concise. This method is tested on hundreds of dynamics problems that demonstrate the ability of epigenetic local search to improve GP.

The third method is a multidimensional GP approach (M4GP) for solving multiclass classification problems. The proposed method uses GP to conduct nonlinear feature transformations that are optimized to cluster data samples according to their classes. This method is compared to several out-of-the-box classifiers and is able to generalize to test data better than many other tested methods.

A large portion of this dissertation is devoted to the application of these methods to real world dynamic modeling problems. MSAM is applied to the restructuring of controllers to improve the closed-loop system response of nonlinear plants. ELGP is used to identify the closed-loop dynamics of an industrial scale wind turbine, and is also used to identify reduced-order models of fluid-structure interaction. Lastly, M4GP is used to identify a dynamic behavioral model of bald eagles from GPS data. The methods are analyzed alongside many other state-of-the-art system identification techniques in the context of model accuracy and conciseness.