RECURRENT NEURAL NETWORK BASED LINEAR EMBEDDINGS FOR TIME EVOLUTION OF NON-LINEAR DYNAMICS

東京大学 THE UNIVERSITY OF TOKYO

Shlok Mohta | Department of Electrical Engineering and Information Systems, The University of Tokyo | Email: shlokmohta@cspp.cc.u-tokyo.ac.jp

Introduction

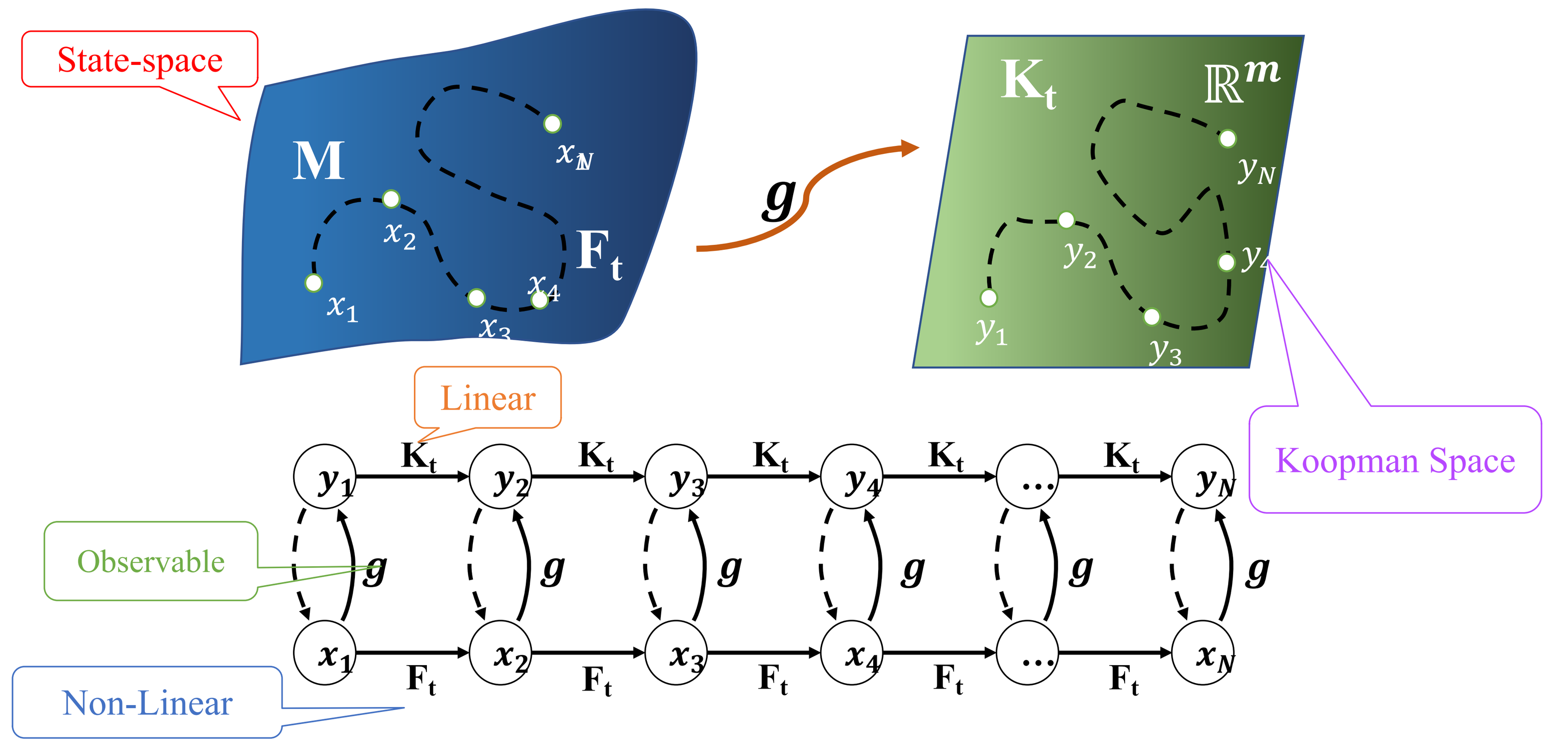
In modern dynamical system modelling, finding coordinate transformation for representing highly non-linear dynamics in terms of approximate linear dynamics has been of crucial importance for enabling non-linear control, estimation, and prediction. Recently developed interest in Koopman operator theory has shown that its eigenfunctions can provide such coordinates that intrinsically linearize the global dynamics But finding and representation of such eigenfunctions have been challenging. The present work leverages deep learning methods, specifically Recurrent Neural Networks (RNNs) for discovering the Koopman eigenfunction representations and exploit RNNs ability to model temporal dependencies, to allow multi-step evolution of the dynamics, as long forecasting for such systems still remains a major challenge. Current work is an incremental work on the network architecture, which is interpretable in terms of Koopman theory and parsimonious, allowing augmentation to the lacking interpretability to deep learning architectures, while capturing the fewest meaningful eigenfunctions. Some other challenges related to modelling such architectures are discussed in future work.

Koopman Operator Theory

Koopman operator theory is a mathematical framework for evolving a system on the *observable* functions of an infinite dimensional Hilbert Space. Koopman operator is a linear operator which acts on these system observables in the same way a flow-map of the system does, evolving the observable of the system. The following figure gives a more intuitive feeling of Koopman theory.

The main idea is to able to represent the observables of the systems in terms of the eigenvectors of the Koopman Operator. But the find the eigenvectors is a challenging task.

To better understand, think of fluid flow as an example with pressure or vorticity being the system observable (function of state-space).



Approach

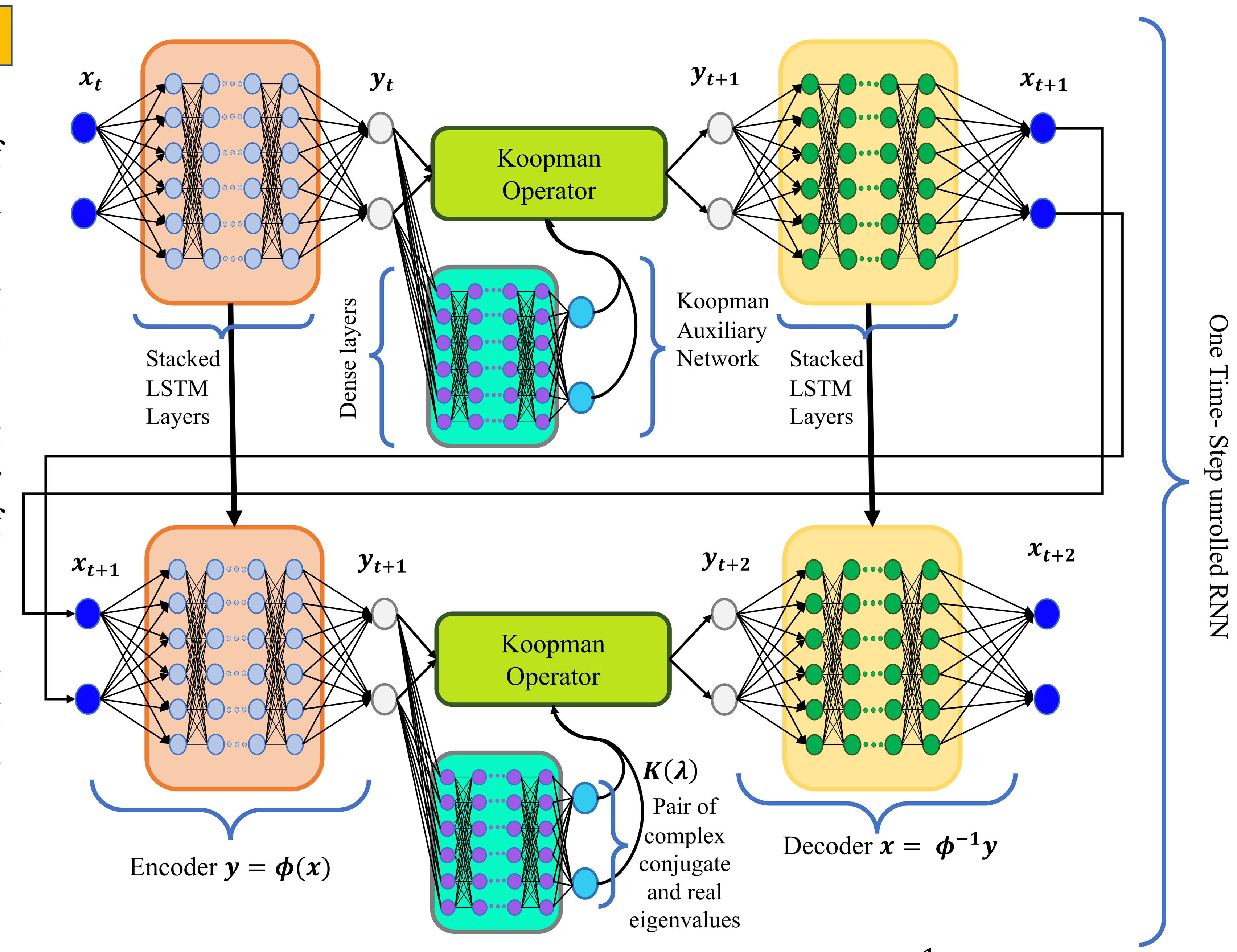
The primary goal of this work is to enable longer-time forecasting of complex dynamical systems while maintaining interpretability and the parsimony of the network. The present work builds upon the network architecture proposed by Lusch et.al.

The first figure with 'Prediction loss' below represents the primary network architecture. Given to the Koopman interpretability of the network, various 'interpretable' losses are computed during the training, described in by the figures of 'Autoencoder' and 'Linearization' losses. x_t and y_t represent the dynamics in the non-linear subspace and Koopman subspace respectively, where y_t is encoded from x_t, evolved using the Koopman operator to y_{t+1} and then decoded to x_{t+1}.

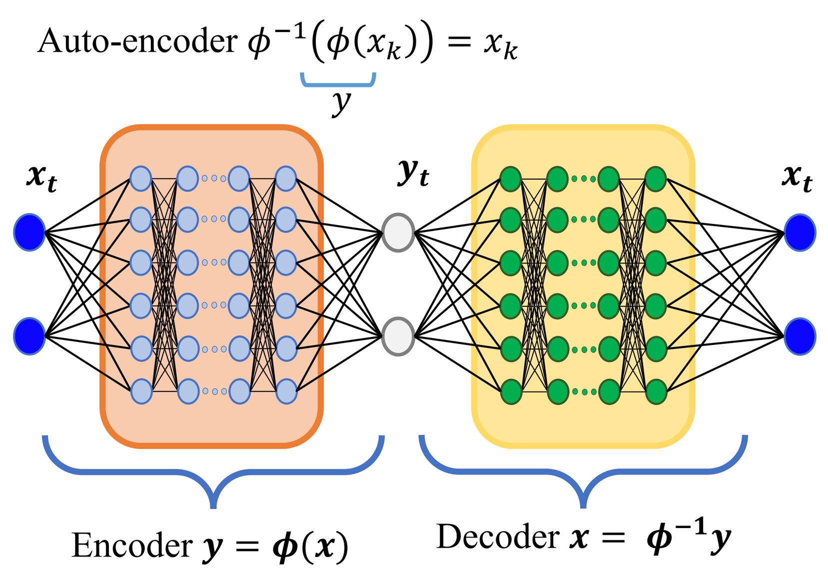
The enabling thought for the current architecture is to implement 'Long-Short Term Memory' stacked layers for the encoder-decoder layers, tackling the problem of high data-requirement and longer range forecasting of the system. RNN networks have shows great promise for tasks pertaining to time-series forecasting.

HPC Aspect

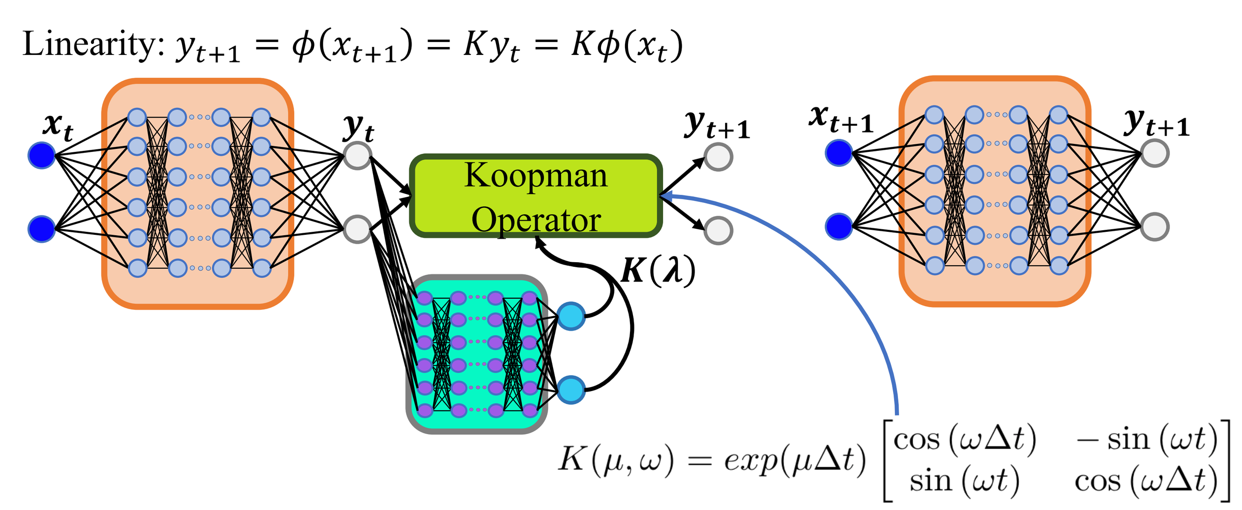
Given the large hyperparameter space of the network, faster and parallelized training is crucial for obtaining the best fit. Though the current architecture is written using TensorFlow API, for efficient utilization of supercomputer resources, packages such as 'Horovod' and 'mpi4py' are being explored for distributed and simultaneous training.



Prediction Loss: $||x_{t+1} - \phi^{-1}K\phi(x)||$



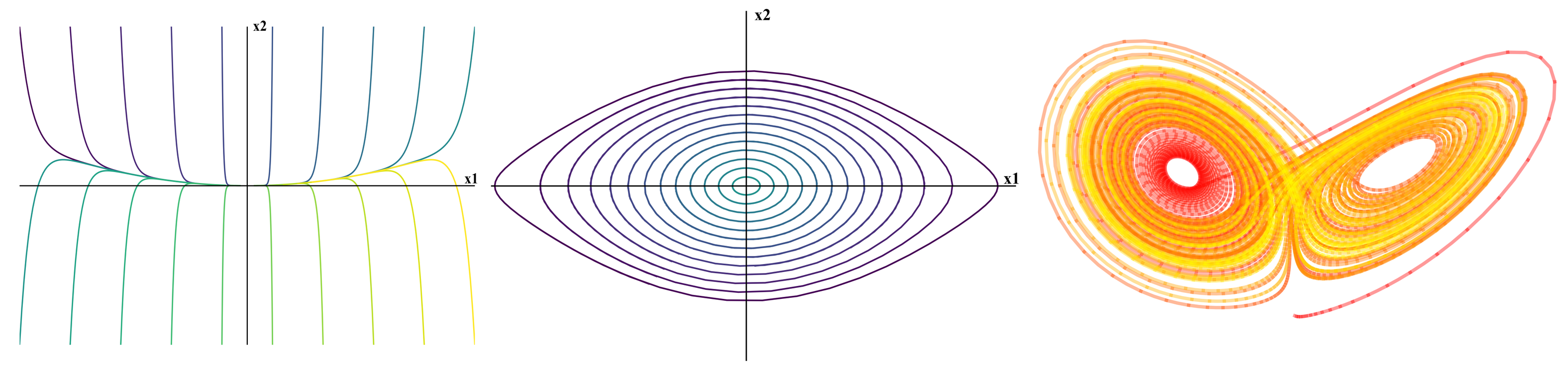
Autoencoder Loss: $||x - \phi^{-1}\phi(x)||$



Linearization loss: $||\phi(x_{t+1}) - K\phi(x)||$

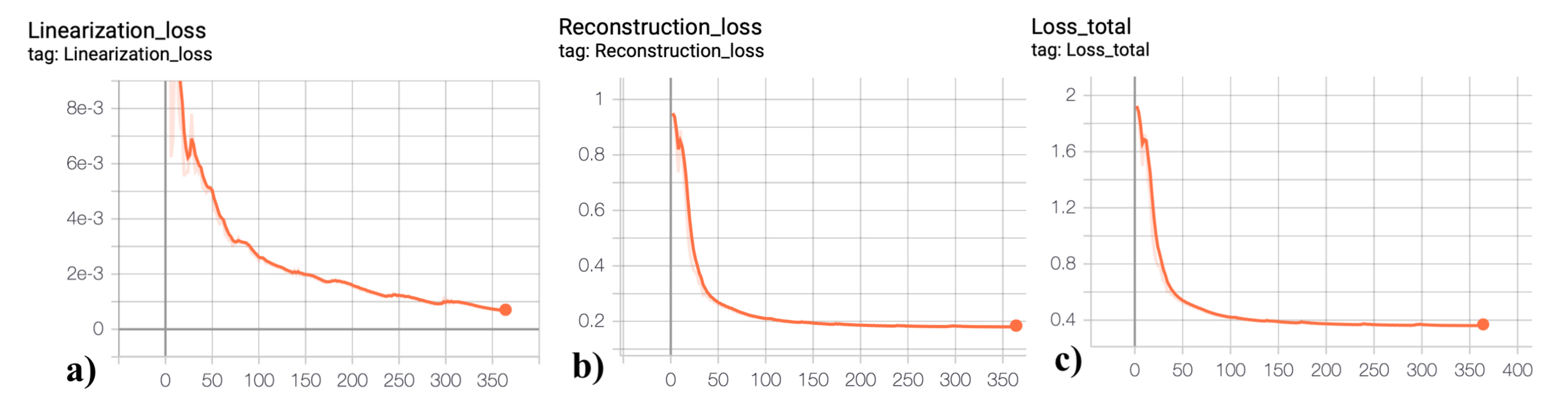
Dataset

The dynamical systems considered here consists of cases with discrete and continuous eigenvalues for the Koopman eigenvectors , presented by dynamics in the first and second figure (non-linear dynamics) from the right. To evaluate the long range prediction capability of the network, the classic Lorenz system has been considered.



Preliminary Results

As this work is still in its early phase, the following figure shows various loss as described in the 'approach'. As it can be seen here, the network is still an underfit, even though the architecture was trained 75% of data size used by previous research for the non-linear pendulum case, showing promise for less data-intensive architecture. This was trained on a P100 GPU on a test node of Reedbush supercoumpter at University of Tokyo. Work is in progress to parallelly run multiple instances for hyperparameter searching.



Future Work

In regards with the current plan:

- Comparison with current data-driven approaches for dynamics evolution.
- Initial State Initializer for RNN layers

Further advancements:

- Bifurcation parameter estimation
- Automatic detection of required eigenfunction
- Systems with higher complexity e.g. Turbulent flows

References

- Bethany Lusch, J Nathan Kutz, and Steven L Brunton. Deep learning for universal linear embeddings of nonlinear dynamics.
- Igor Mezic. 2005. Spectral properties of dynamical systems, model reduction and decompositions.
- Hassan Arbabi. Introduction to Koopman operator theory of dynamical systems.
- N. Mohajerin. Multi-Step Prediction of Dynamic Systems with RNNs