# Learning Interaction Strengths in Network Dynamics from a Single Snapshot


Moritz Lampert[1,2], Dominic Grün[1,2], Ingo Scholtes[1,2]

[1]Julius-Maximilians-University Würzburg, [2]Center for AI and Data Science (CAIDAS)


## Motivation

Dynamical systems are ubiquitous in many areas such as the social sciences or biology. Recent advances in machine learning enable us to learn these dynamics from data to predict how the system evolves based on previous states of the system. However, e.g. in biology, only a single snapshot of a dynamical system can be observed. This lack of temporal information limits the models' learning abilities but by incorporating prior knowledge into our model, we can still learn details about the system.

## Contributions

We empirically investigate the ability of machine learning models to learn a dynamical system based on a single snapshot of the system using inductive biases. Using a synthetic model, we show that we can learn the interaction strengths from a single snapshot and the initial state of the system. We evaluate both linear regression and message-passing-based models to learn linear and non-linear dynamics.

## Linear Dynamical System

We consider a dynamical system with interactions between nodes $v \in V$ where the node class determines the strength of the interaction. Given a graph $G = (V, E)$ without self-loops and node classes $z_v \in \{1, \dots, c\}$, the state of a node $x_v^{(t)}$ at discrete time $t$ depends on its initial state $x_v^{(0)}$ and its neighbours $N(v)$. The interaction strengths $W_{z_v,z_w}$ between node pairs $(v, w)$ specify how much the previous state $x_w^{(t-1)}$ of each neighbour $w \in N(v)$ influences $x_v^{(t)}$. The update function of this class of dynamical systems is given as follows:

$$x_v^{(t+1)} = \alpha \cdot \sum_{w \in N(v)} W_{z_v,z_w} \cdot x_w^{(t)} + (1-\alpha) \cdot x_v^{(0)},$$

with a parameter $\alpha \in [0, 1]$, where $1 - \alpha$ can be seen as a restoring force that pulls the node back to its initial state $x_v^{(0)}$.

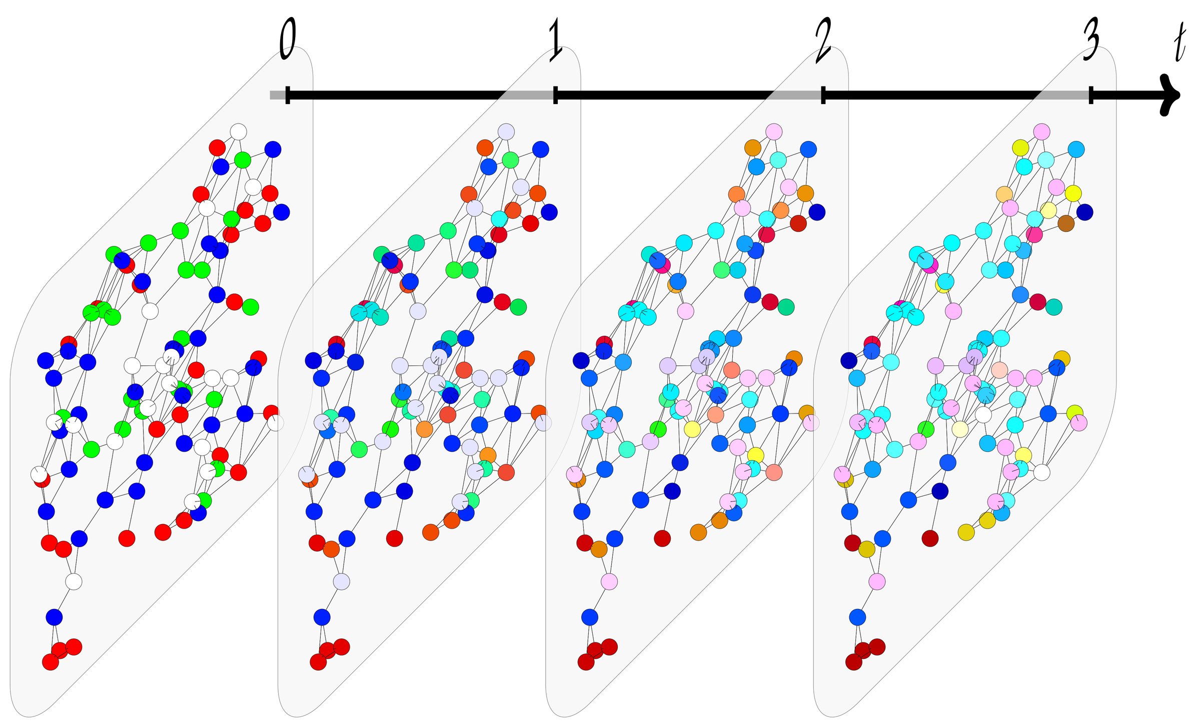



Figure 1: Linear dynamical system with $x_v^{(t)} \in \mathbb{R}^3$ visualised as RGB-channels and $c = 4$ classes for four time steps. We use $\alpha = 0.1$, resulting in a small influence of the neighbours on the node state.

## Learning Linear Interaction Strengths

Given the state $x_v^{(t)}$ for each node $v \in V$ at time step $t$, we aim to learn the interaction strengths $W_{z_v,z_w}$ between node types $z_v$ and $z_w$. For each node $v$, we treat its current state $x_v^{(t)}$ as the target variable. As input features, we use the initial state $x_v^{(0)}$ of node $v$, and the set of current states $\mathcal{X}_{N(v)}^{(t)} = \{x_w^{(t)} \mid w \in N(v)\}$ of all neighbouring nodes $w \in N(v)$. Assuming a linear model of the form:

$$x_v^{(t)} = \alpha \cdot \sum_{x_w^{(t)} \in \mathcal{X}_{N(v)}^{(t)}} W_{z_v,z_w} \cdot x_w^{(t)} + (1-\alpha) \cdot x_v^{(0)},$$

we can estimate $W_{z_v,z_w}$ by solving a linear regression problem for each node $v$.
Alternatively, we can model the same dynamics using message passing, where the prediction for node $v$ at time $t$ is given by:

$$\mathrm{MP}_\alpha\left(x_v^{(0)}, \mathcal{X}_{N(v)}^{(t)}\right) = \alpha \cdot \sum_{x_w^{(t)} \in \mathcal{X}_{N(v)}^{(t)}} W_{z_v,z_w} \cdot x_w^{(t)} + (1-\alpha) \cdot x_v^{(0)}.$$

The interaction weights $W_{z_v,z_w}$ are treated as trainable parameters, which are optimized using gradient descent to minimise the mean squared error (MSE).

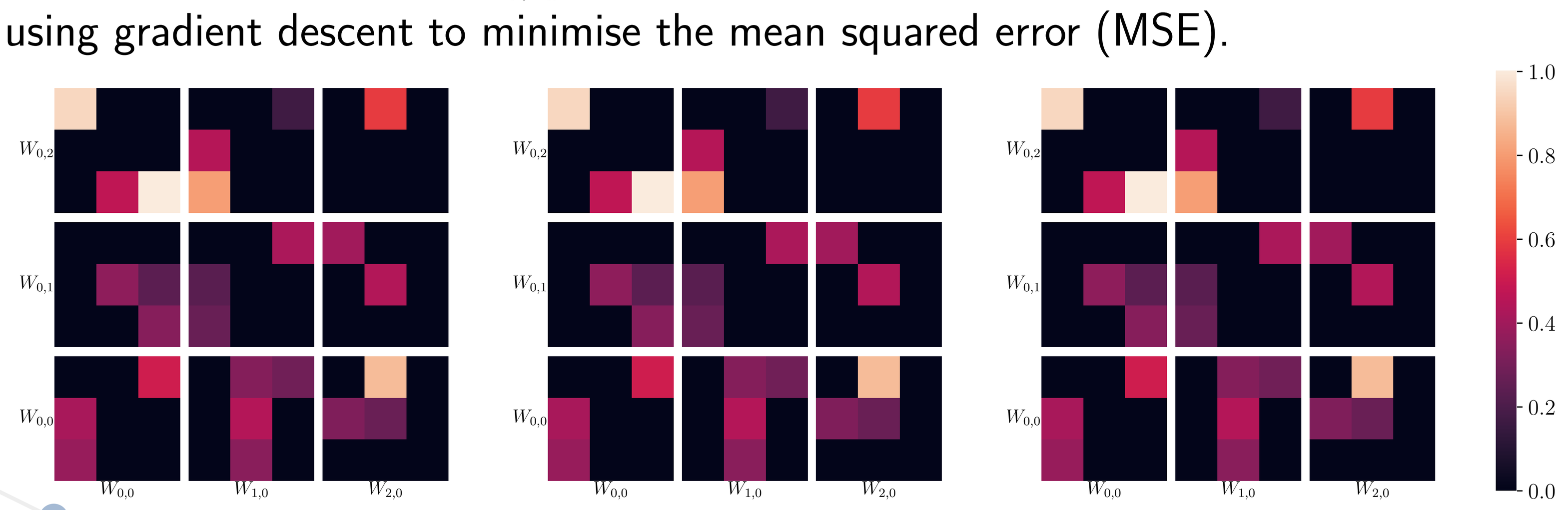


Figure 2: True interaction strengths $W_{z_v,z_w}$ (left) and reconstructed weights using linear regression (middle) and an MP-based model (right), learned on the above linear dynamical system at $t = 15$ based on a network similar to Figure 1 with $|V| = 2500$. The values are normalised to $[0, 1]$.

## Learning Non-Linear Interaction Strengths

We extend our linear to a non-linear dynamical system by adding a non-linearity $\sigma$:

$$x_v^{(t+1)} = \alpha \cdot \sum_{w \in N(v)} \sigma\left(W_{z_v,z_w} \cdot x_w^{(t)}\right) + (1-\alpha) \cdot x_v^{(0)} \text{ with } \sigma(x) = \frac{1}{1+e^{-x}}$$

We can add the non-linear function to the MP-based model:

$$\mathrm{MP}_\alpha\left(x_v^{(0)}, \mathcal{X}_{N(v)}^{(t)}\right) = \alpha \cdot \sum_{x_w^{(t)} \in \mathcal{X}_{N(v)}^{(t)}} \sigma\left(W_{z_v,z_w} \cdot x_w^{(t)}\right) + (1-\alpha) \cdot x_v^{(0)}$$

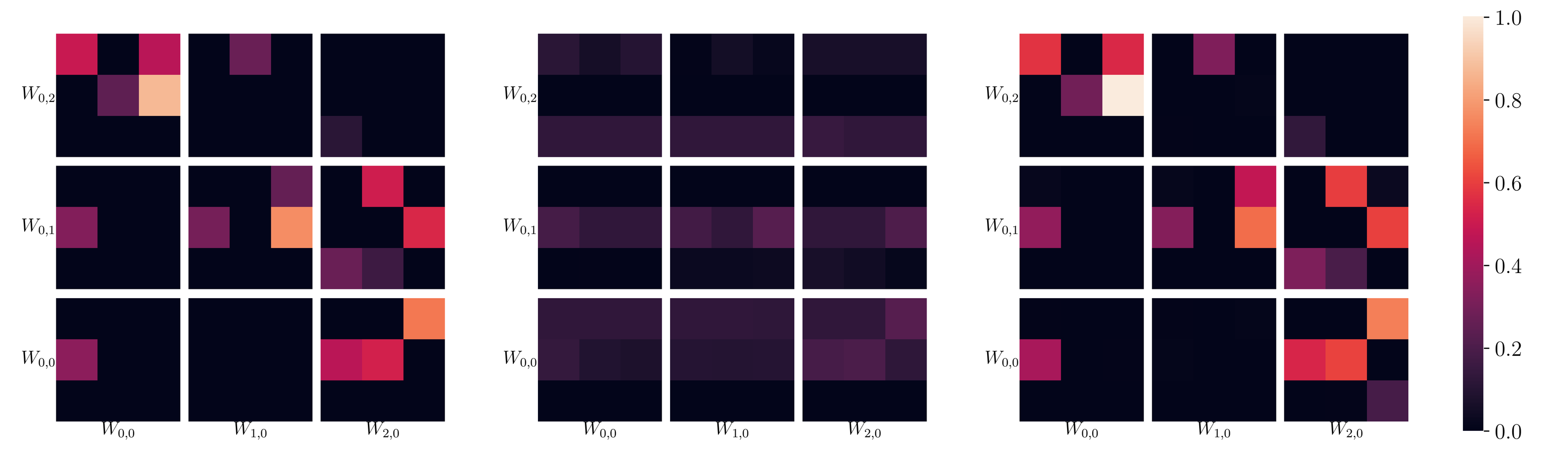
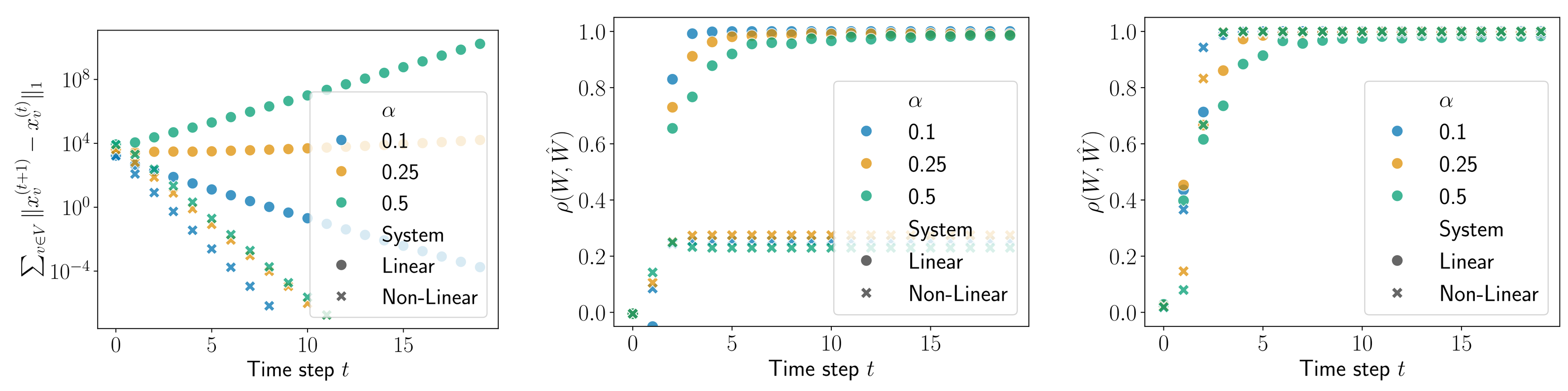


Figure 3: As in Figure 2 but for a non-linear dynamical system with true interaction strengths $W_{z_v,z_w}$ (left) and reconstructed weights using linear regression (middle) and an MP-based model (right).

## Stationarity

We show empirically that we are able to learn the interaction strengths of a dynamical system with knowledge about $x_v^{(t)}$ only at a single timestamp $t$. Intuition suggests that we can replace $x_v^{(t+1)}$ with $x_v^{(t)}$ if $|x_v^{(t+1)} - x_v^{(t)}|$ is small, i.e. the system is stationary. However, we show in Figure 4 that the models are able to learn the interaction strengths even if the system is not stationary.

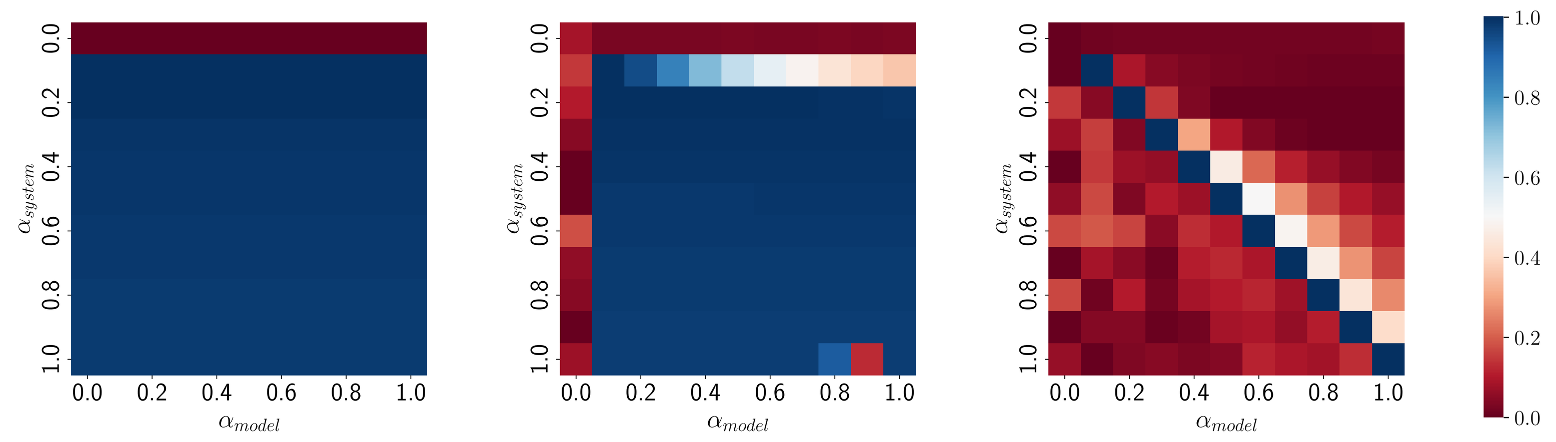



(a): Stationarity. (b): Regression models. (c): MP-based models.

Figure 4: Node state differences between $t$ and $t + 1$ (left) indicating the stationarity of the system. Pearson correlation between true and learned interaction strengths of the models (middle, right) showing that they are able to learn the interaction strengths regardless of the stationarity of the system.

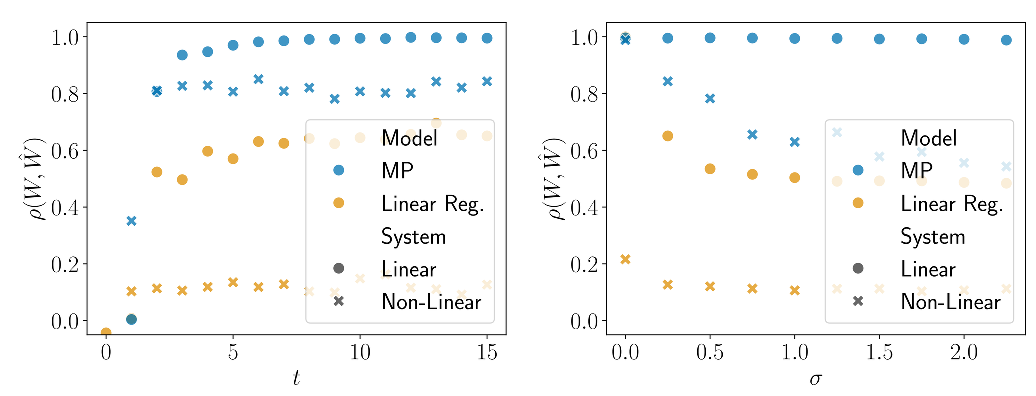
## Prior Knowledge

A big limitation of our approach is that we need to have prior knowledge about the system. In Figure 5-7, we empirically investigate the influence of partially wrong prior knowledge. In particular, we investigate deviating values of $\alpha$, noisy initial states $x_v^{(0)}$, and using a linear model to learn from a non-linear system and vice versa.



(a): Lin. system & lin. reg. (b): Lin. system & MP. (c): Non-lin. system & MP.

Figure 5: Pearson correlation with mismatched $\alpha$ values for the dynamical system and the learning model.




(a): Across $t$ ($\sigma = 0.25$). (b): Across $\sigma$ ($t = 15$).

Figure 6: Pearson correlation using $x_v^{(0)}$ with added noise sampled from a Normal distribution with $\mathcal{N}(0, \sigma)$ during learning.

| | Lin. | Non-Lin. |
|---|---|---|
| Lin. Reg. | 0.99 | 0.22 |
| MP-Lin. | 0.99 | 0.86 |
| MP-Non-Lin. | 0.02 | 0.96 |

Figure 7: Pearson correlation for dynamical system and learning model combinations: linear vs. non-linear.

## Conclusions

Our work demonstrates how we can encode inductive biases into machine learning models, enabling them to learn details of a dynamical system from a single snapshot in time series data. This approach can be useful in scenarios where we have prior knowledge about the class of the dynamical system but lack specific parameters.