Neural ODE Processes
NDPs

Alexander Norcliffe*, Cristian Bodnar*, Ben Day*, Jacob Moss*, Pietro Liò
alex.norcliffe98@gmail.com,
[cb2015, bjd39, jm2311, pl219]@cam.ac.uk

UNIVERSITY OF CAMBRIDGE

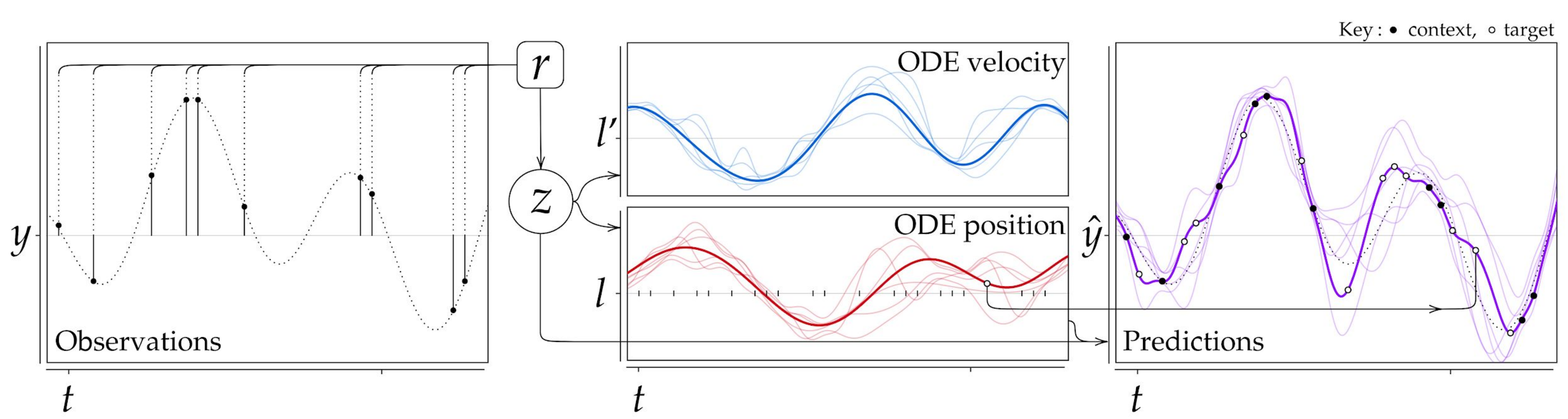
Machine Learning for the Physical Sciences 2020
NeurIPS workshop

Motivation and Contributions

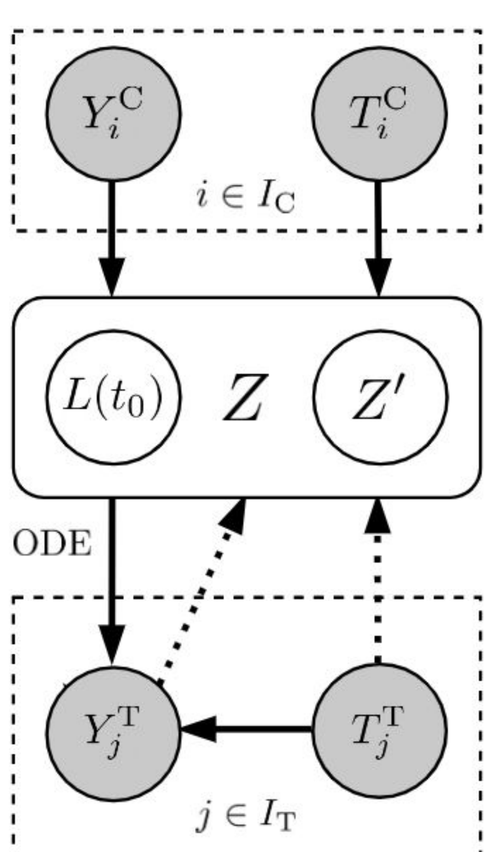
Neural ODEs (NODEs) [1] use a neural network to model the instantaneous rate of change of the state of a system. Despite their suitability for modelling dynamics governed time-series they have some drawbacks:

- They are unable to adapt at test time to incoming data points
- They do not capture uncertainty over the dynamics that model a system

To overcome these issues we introduce Neural ODE Processes (NDPs). NDPs are a new class of stochastic processes determined over a distribution of ODEs, which extends Neural Processes (NPs) [2] to use Neural ODEs.



Above: Schematic diagram of Neural ODE Processes.
Left: Observations from a time series, the context set, are encoded and aggregated to form a representation of the time-series. This parametrises the global latent variable ***z***.
Middle: Samples of ***z*** are drawn, which initialise and condition a latent ODE, with each sample producing a coherent trajectory.
Right: Predictions at target times are made by decoding the latent ODE.



Left: Graphical model of NDPs. Dark nodes represent random variables, light nodes represent hidden variables.

To predict a trajectory from a context set:

1. Encode each context point
2. Aggregate the encoded representations
3. Parametrise the distribution of z
4. Integrate the latent ODE to get $l(t)$ at all times
5. Decode the latent state to get the prediction

Our main contributions are:

- Introduce NDPs, a stochastic process for dynamical systems.
- Show NDPs can adapt to new data points and capture uncertainty.
- Demonstrate its applicability to low-dimensional dynamical systems.
- Show they scale up to high-dimensional time-series with latent dynamics.

Predator Prey Dynamics

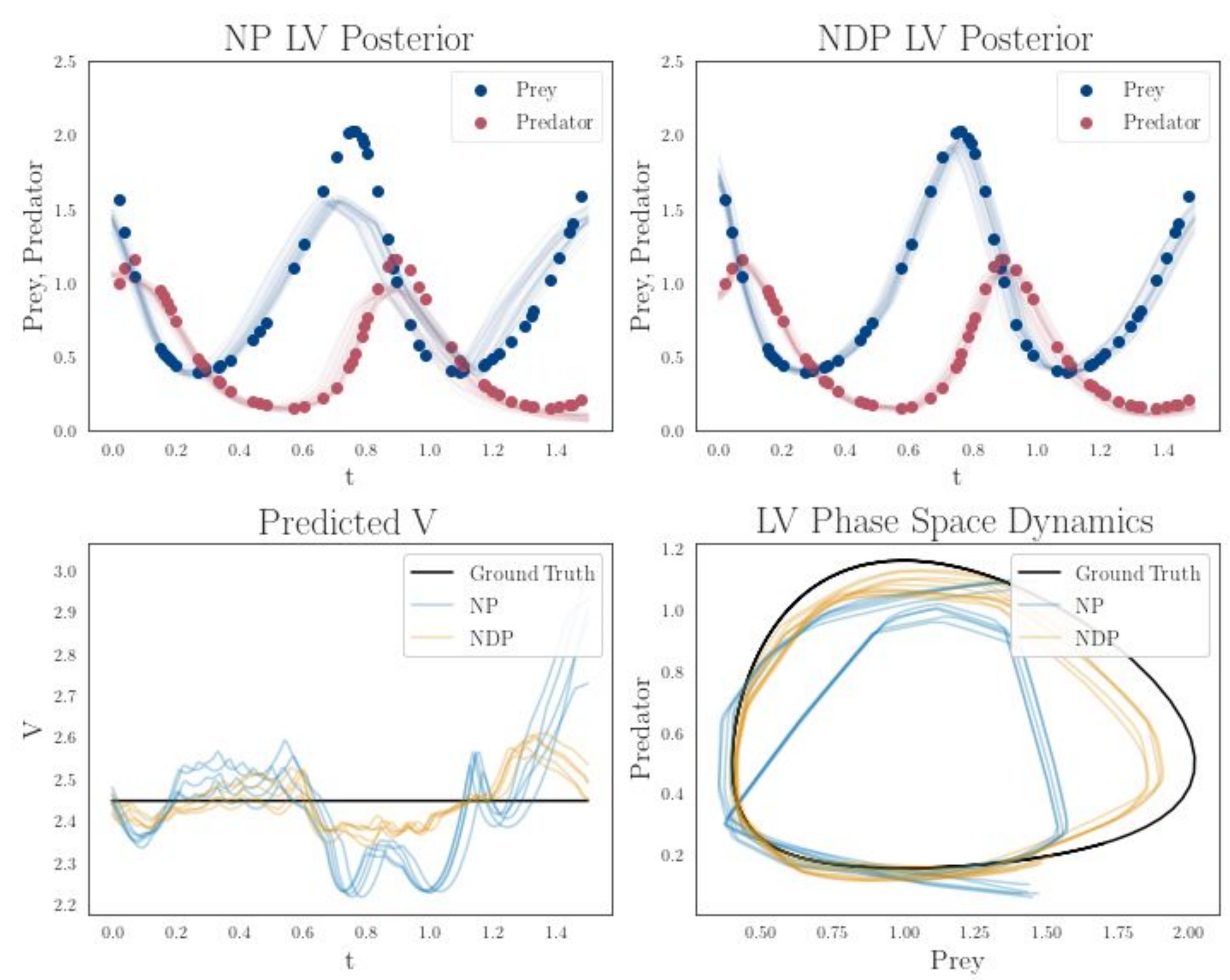
We considered modelling the Lotka-Volterra equations, for a predator-prey system. The population of the prey *u*, and predator *v*, change according to the ODEs:

$$\frac{du}{dt} = \alpha u - \beta uv \qquad \frac{dv}{dt} = \delta uv - \gamma v$$

Intuitively the prey population goes up, unless the predator population is high and the predator population will go down unless the prey population is high. We use the values (α, β, γ, δ) = (2/3, 4/3, 1, 1).

We consider the conserved quantity *V* to see how well the models learn the LV solution:

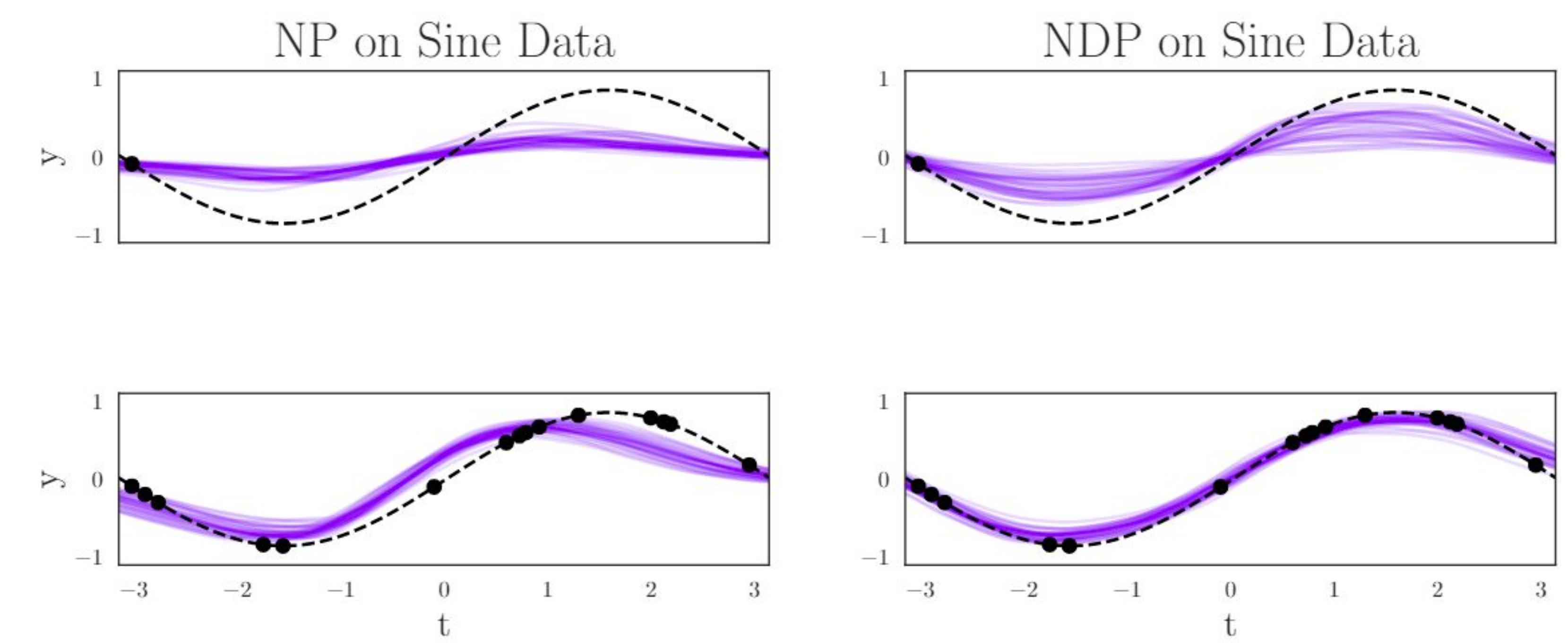
$$V = \delta u - \gamma \ln(u) + \beta v - \alpha \ln(v)$$



Visualising NPs and NDPs learning the Lotka Volterra equations. We see that NDPs can learn trajectories that are closer to the ground truth, both in phase space and in conserving the quantity *V*.

Adapting to New Data Points

We trained the models on various 1D systems: sine curves, exponentials, straight lines, harmonic oscillators. Below we show how they perform on sine curves as they are given more data points.

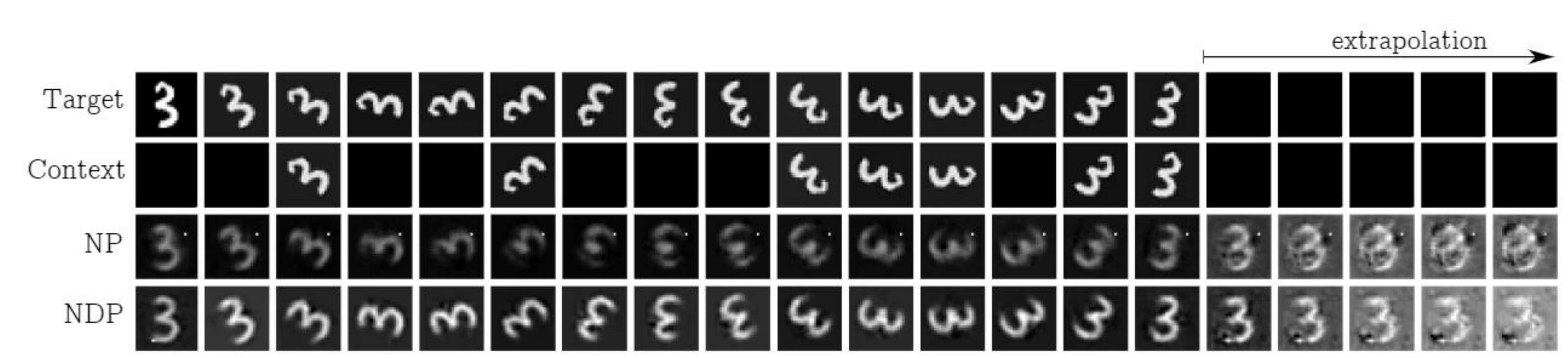


The models trained on sine curves adapting to new data. We see that NDPs are able to adapt to the new data better than NPs.

Rotating MNIST Digits

To test our model on high dimensional data, we look at the Rotating MNIST digits dataset [3], which consists of MNIST 3s being rotated over 16 frames.

We give NDPs a small number of whole frames as the context, and then let it predict the intermediate frames and extrapolate.



We compare the performance of NDPs and NPs on the rotating MNIST task. We see that NDPs are able to interpolate well with only a small number of context frames. Both model struggle to extrapolate.

References

[1] Chen, R.T., Rubanova, Y., Bettencourt, J. and Duvenaud, D.K., 2018. Neural ordinary differential equations. In Advances in neural information processing systems (pp. 6571-6583).
[2] Garnelo, M., Schwarz, J., Rosenbaum, D., Viola, F., Rezende, D.J., Eslami, S.M. and Teh, Y.W., 2018. Neural processes. arXiv preprint arXiv:1807.01622.
[3] Casale, F.P., Dalca, A., Saglietti, L., Listgarten, J. and Fusi, N., 2018. Gaussian process prior variational autoencoders. In Advances in Neural Information Processing Systems (pp. 10369-10380).