Meta-learning using privileged information for dynamics

Ben Day, Alexander Norcliffe, Jacob Moss, Pietro Liò

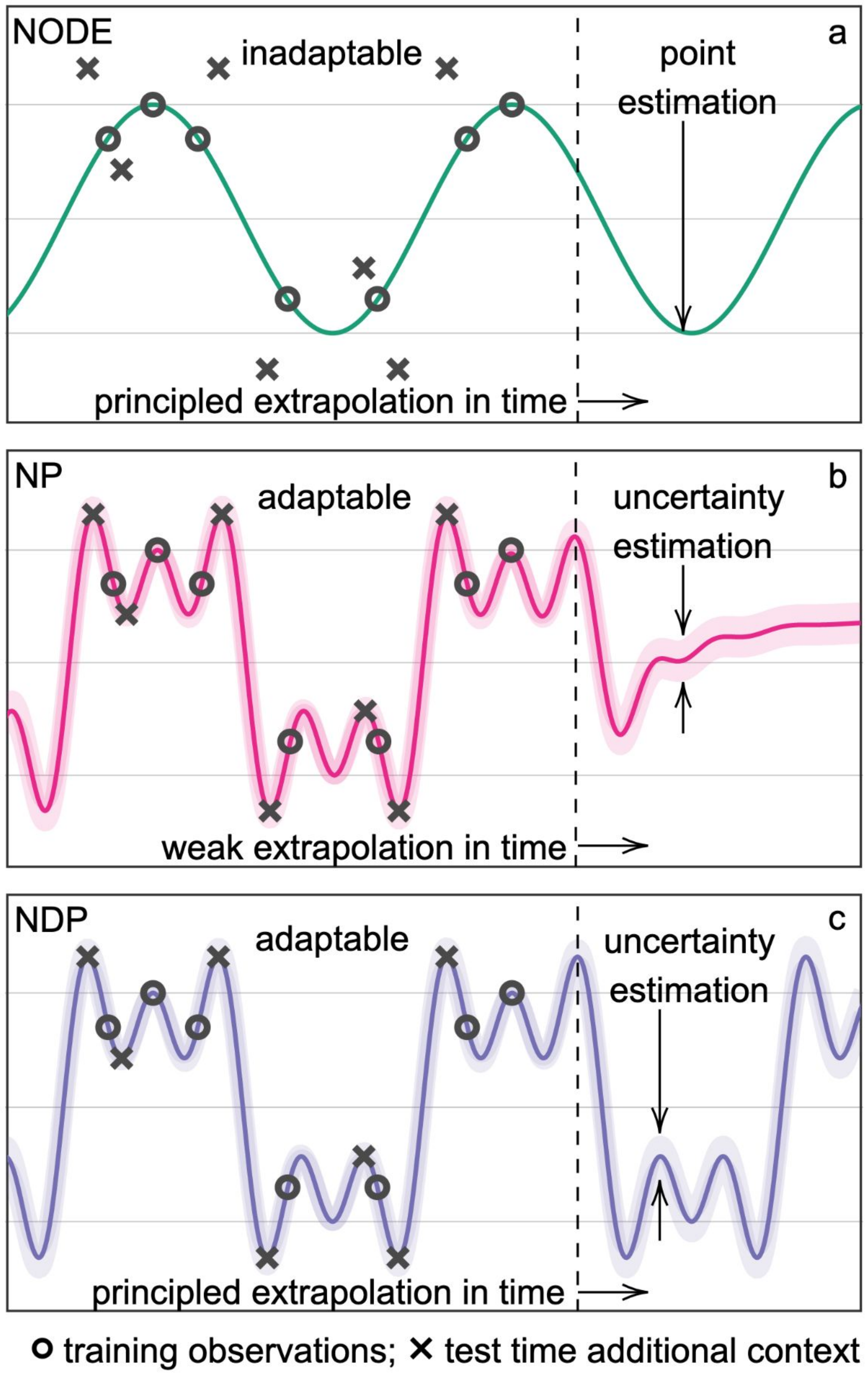
Department of Computer Science and Technology, University of Cambridge, Cambridge, UK.

Correspondence to Ben: bjd39@cam.ac.uk

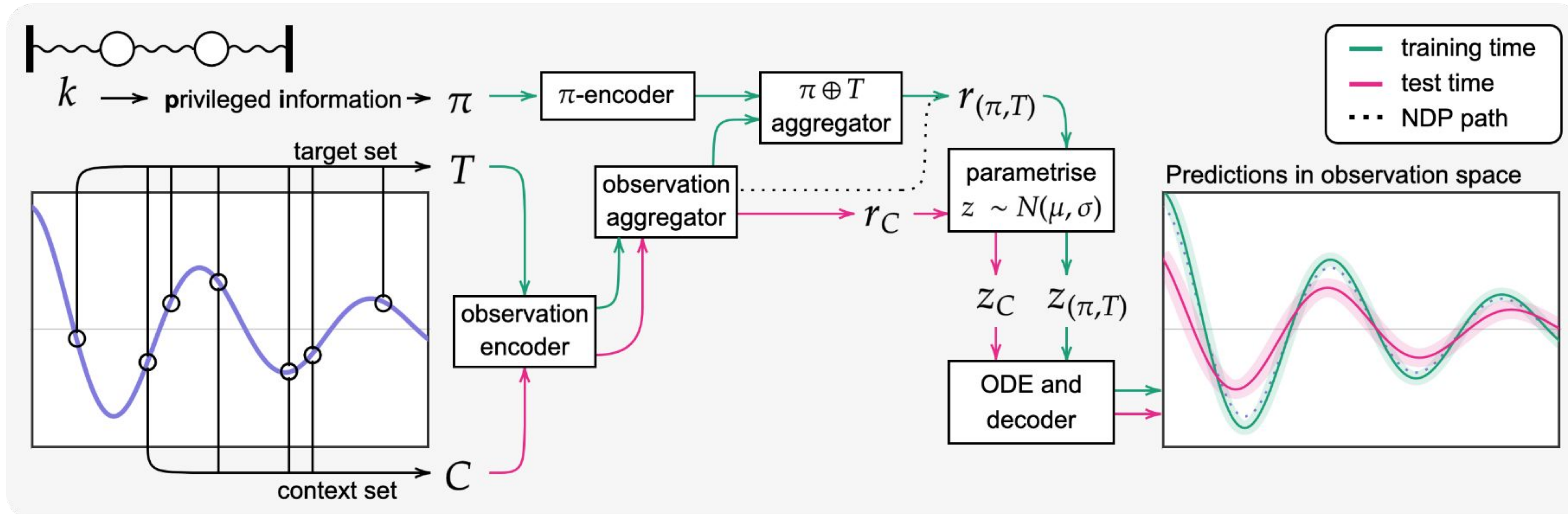
Code: github.com/bjd39/lupi-ndp

twitter: @itsmebenday, @alexnorcliffe98, @MossJacob, @pl219_Cambridge

Background: Neural ODEs (NODE), NPs, and NDPs

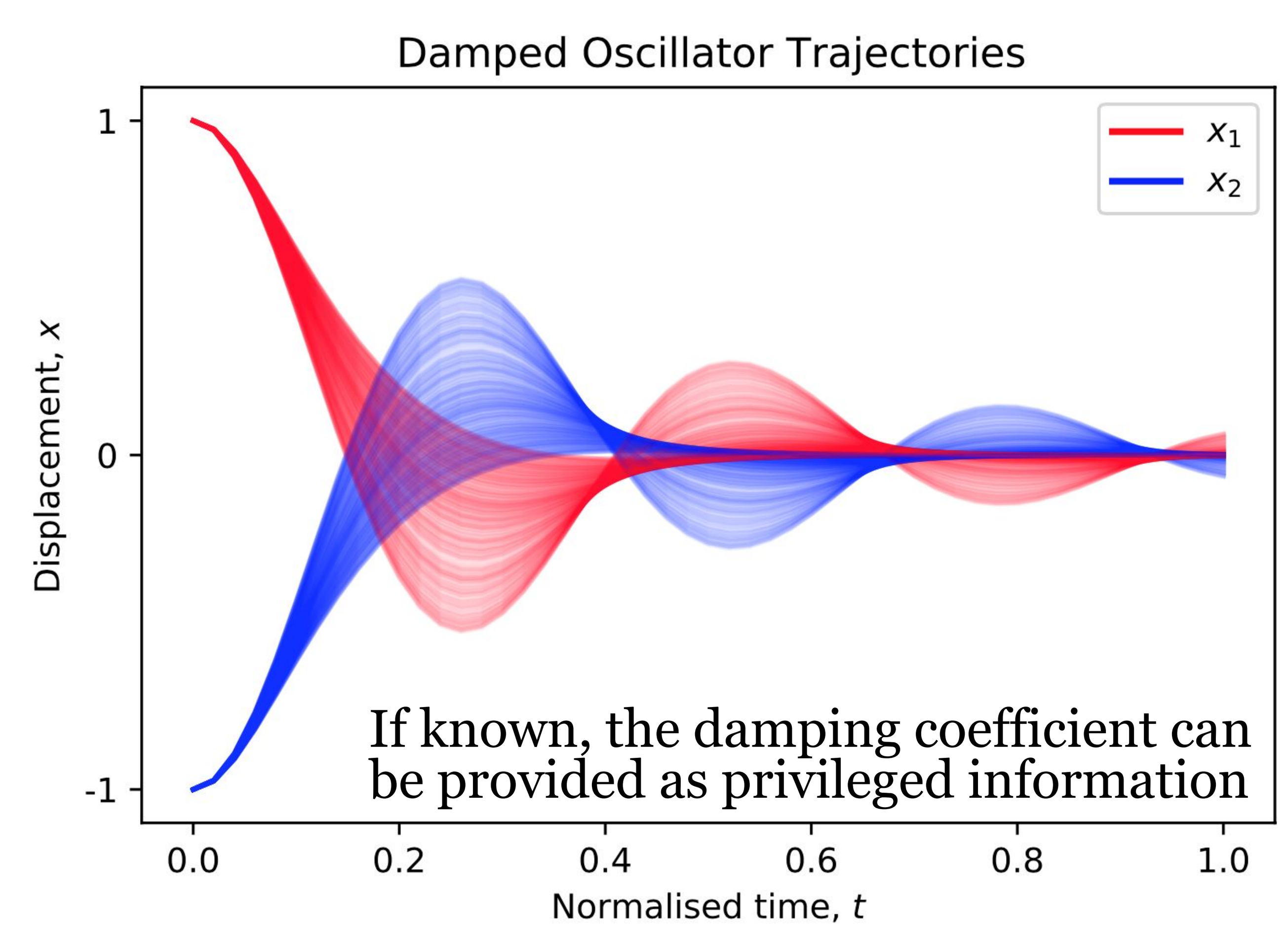


tl;dr: Can we better meta-learn dynamics if we have access to high-level descriptions at training time? (Yes.)



1. Neural ODE Processes (NDP) approach meta-learning dynamics with a latent variable model, and inherit a flexible context aggregation mechanism from the Neural Process (NP): arbitrarily sized sets of observations are aggregated into a fixed length representation. **2.** In the sciences, we often have access to high-level information in addition to observations (e.g. known conserved quantities) this is *privileged information*.

→ Taking advantage of the aggregation flexibility, we extend NDPs to use additional information within the Learning Using Privileged Information (LUPI) setting (training only) and find general improvements.



	Varying stiffness, $k \sim \mathrm{U}(0.2, 1)$			Varying damping, $c \sim \mathrm{U}(0.5, 2)$		
Model	MSE ↓	Calib. error ↓	Sharp. ↓	MSE ↓	Calib. error ↓	Sharp. ↓
NoPI	1.05 ± 0.05	0.51 ± 0.02	6.88	2.82 ± 0.29	0.84 ± 0.04	2.15
LUPI	**0.93 ± 0.04**	**0.47 ± 0.02**	6.57	**2.39 ± 0.09**	**0.37 ± 0.02**	4.71
NoPI*	0.16 ± 0.02	2.69 ± 0.02	1.00	0.56 ± 0.02	1.56 ± 0.03	0.93
LUPI*	**0.06 ± 0.01**	**0.91 ± 0.02**	1.10	**0.25 ± 0.01**	**0.73 ± 0.03**	1.18

	L-V, $u_0 \sim \mathrm{U}(0.2, 1), v_0 \sim \mathrm{U}(0.1, 0.5)$		
Model	MSE ↓	Calib. error ↓	Sharp. ↓
NoPI	6.44 ± 0.44	2.19 ± 0.05	2.23
LUPI	**1.82 ± 0.13**	**0.90 ± 0.04**	3.44
NoPI*	5.24 ± 0.30	2.89 ± 0.04	1.37
LUPI*	**0.73 ± 0.02**	**1.23 ± 0.04**	1.48

* indicates 'training mode' i.e. additional observations and privileged information provided