An hybrid physics-ML framework to model exoplanetary light curves

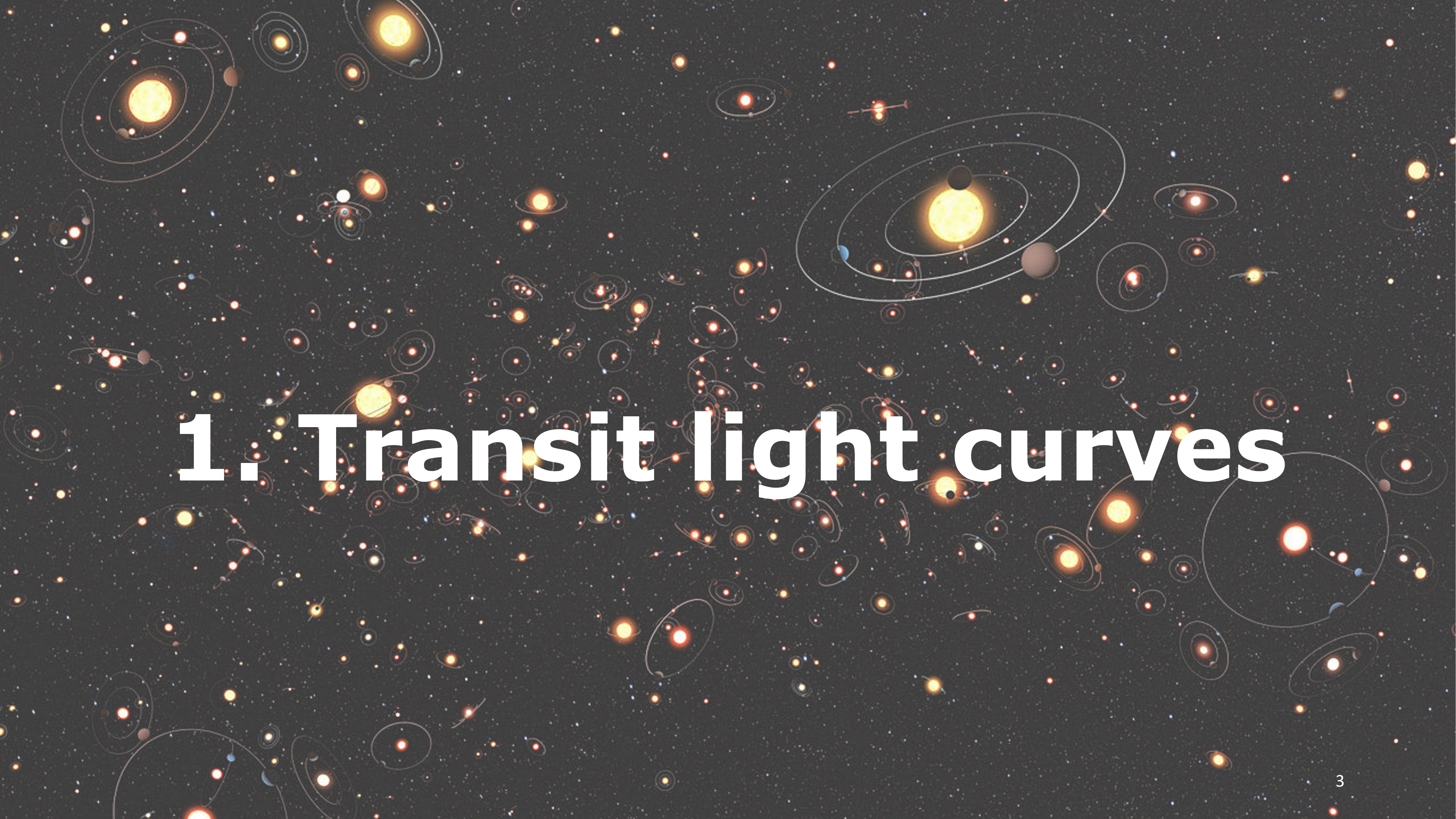
ML-IAP conference

22nd October 2022

Mario Morvan, PhD Student
University College London – exoplanets group
Centre for Doctoral Training in Data Intensive Science
Collaborators: Ingo Waldmann, Nikolaos Nikolaou, Angelos Tsiaras, Gordon Yip

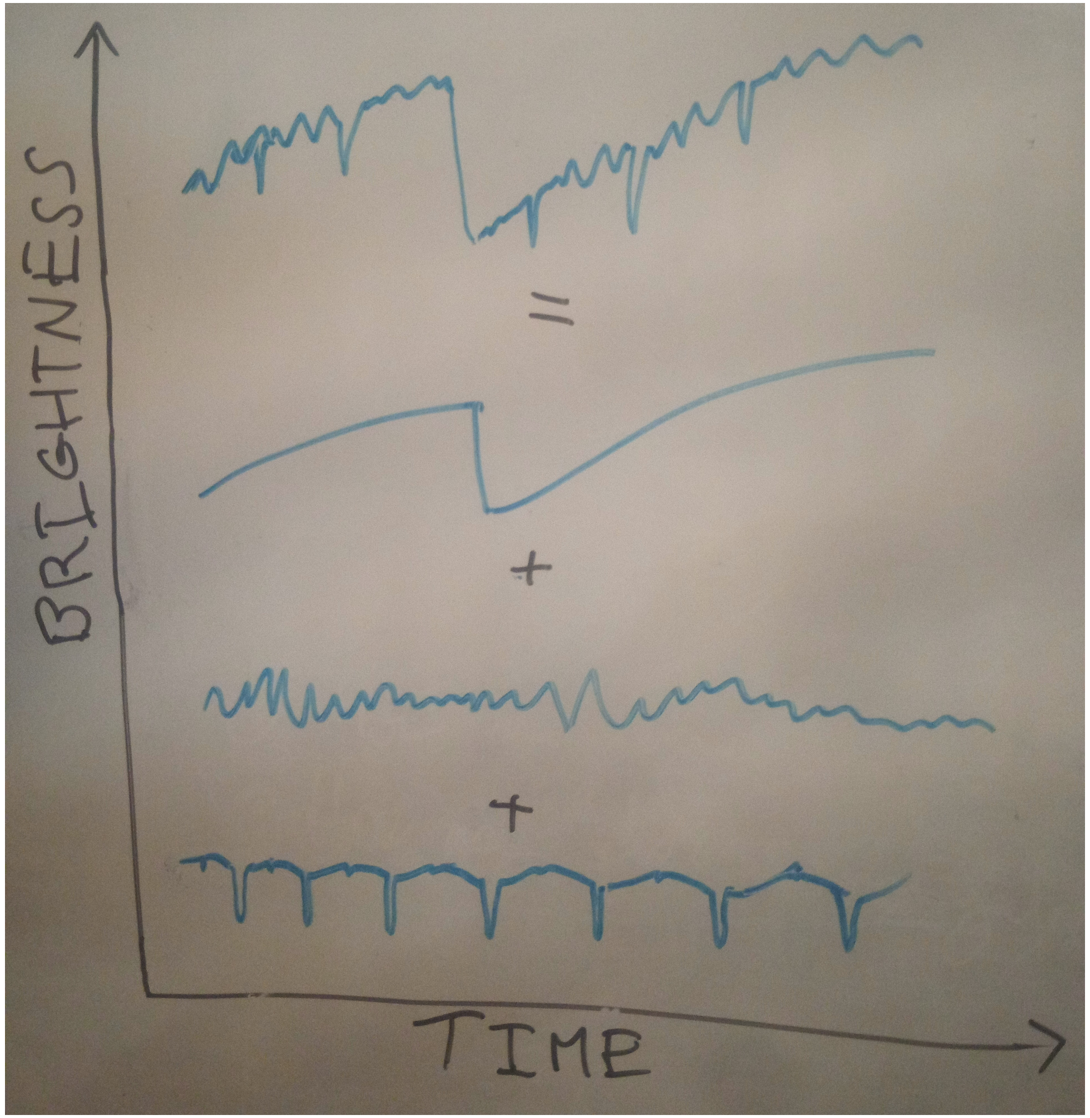
Outline

1. **Transit light curves: Data and Challenges**
2. **Can neural nets denoise light curves?**
3. **Towards hybrid DL-physics models**



1. Transit light curves

Transit light curves



Raw Transit Light Curve

=

Instrumental noise:

- Read-out
- Electron trapping (Spitzer, JWST…)
- Dark current
- Intra/inter-pixels variations

+

Astrophysical noise

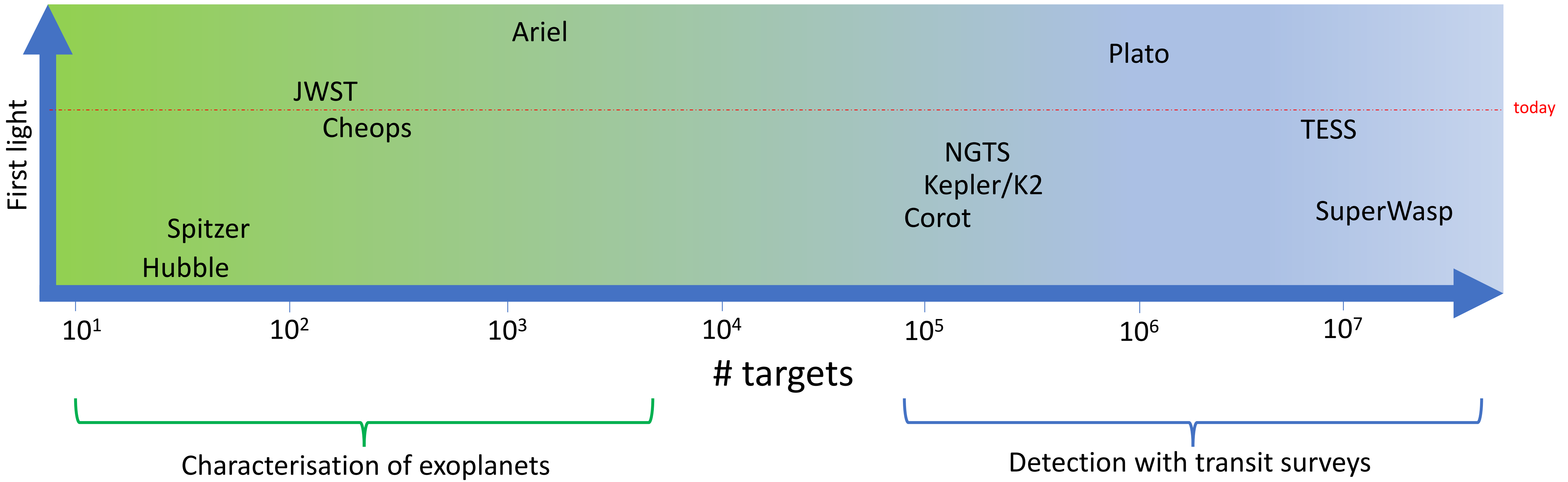
- Background and foreground light
- Stellar variability
- Solar system reflexions
- Unresolved binaries

+

Planetary signals

- Primary transit
- Secondary transit
- Phase curves

Detection and characterisation: Towards exoplanet surveys



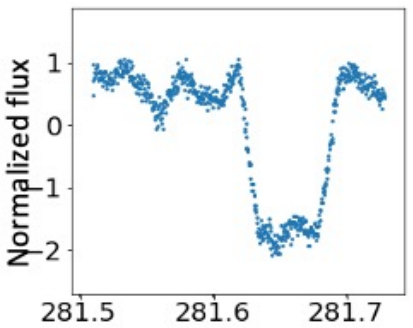
Detrending Light curves

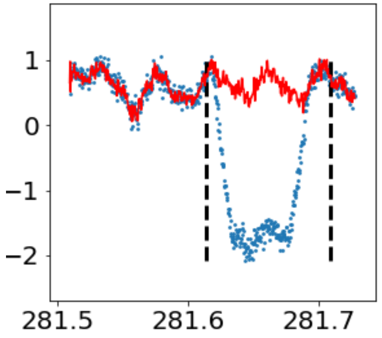
- Detrending ~ Temporal Denoising
- Processing milestone for detection and characterisation
- Previous techniques either:
  - Either local, linear or non-scalable models
  - Fail to provide a general consensus on how to best detrend LCs

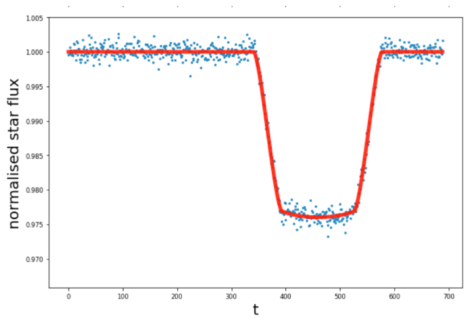
What next?

2. Can neural nets detrend light curves?

Transit Light Curve Detrending (TLCD) LSTM model in a nutshell





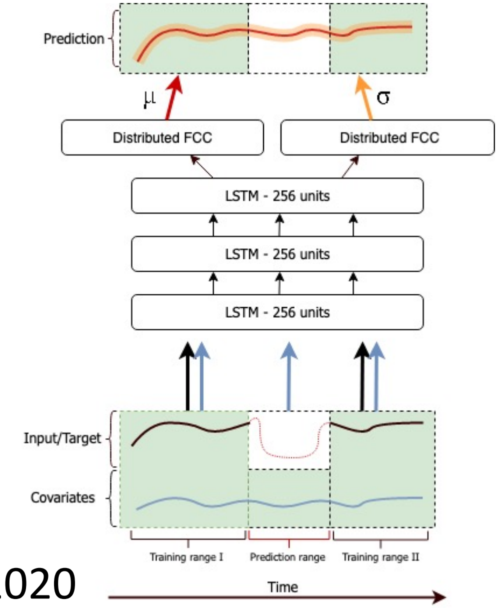


Transit parameters
$(R_P/R_S, a/R_S, P, i \ldots)$

1) Trend Prediction using TLCD-LSTM model

2) Trend removal

3) Transit-model Fitting



Autoregressive probabilistic multi-layer LSTM to learn the noise structure.
Inspired from DeepAR (Salinas et al. 2017) and BRITS (Cao et al. 2018) models:

- **gaussian likelihood loss** on next point prediction
- Input is **masked during transit**
- **reinjecting predictions** in autoregressive way predict noise during transit
- Applied to Spitzer data successfully

Morvan et al 2020

The state and future of time-series deep modelling

Towards global non-linear scalable detrending models?

- DL hegemony in image processing since 2012 (AlexNet)
- DL hegemony in NLP since 2017 (transformers)
- Has DL surpassed traditional methods for Time Series?
  - Mostly yes for forecasting/classifying many Time Series
  - Still lagging behind images and language models
  - Unsupervised and Generative TS modelling still infant

3. Composition of transit physics with DL model

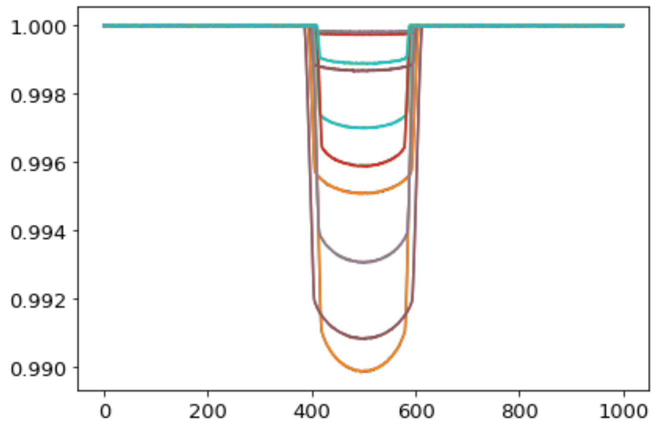
A differentiable transit model

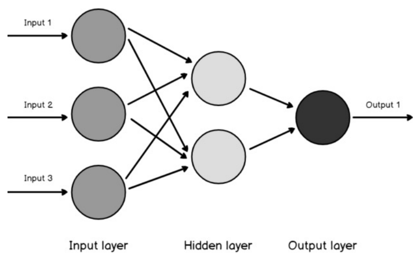
- Objective: embed a transit model in DL framework
- Requirements:
  - Differentiable → joint optim
  - Vectorisable and GPU-compatible → efficiency
  - Deep-Learning framework → convenience

→ PyLightcurve-torch

Morvan et al 2021

Experiment: inverse transit problem





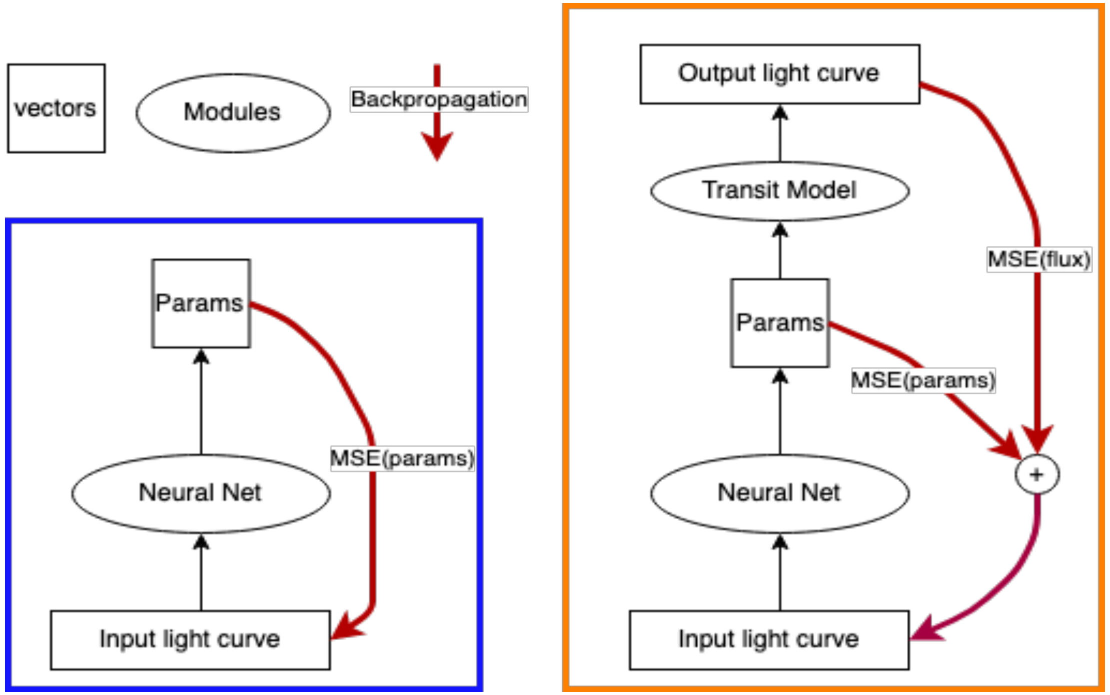
Transit. Parameters

- ***Data***: 2000 synthetic transit light curves with random transit parameters and Gaussian noise
- ***Metric***: Mean Squared Error on transit parameter **Rp/Rs**
- ***Architecture*:** 1D CNN

Morvan et al 2021

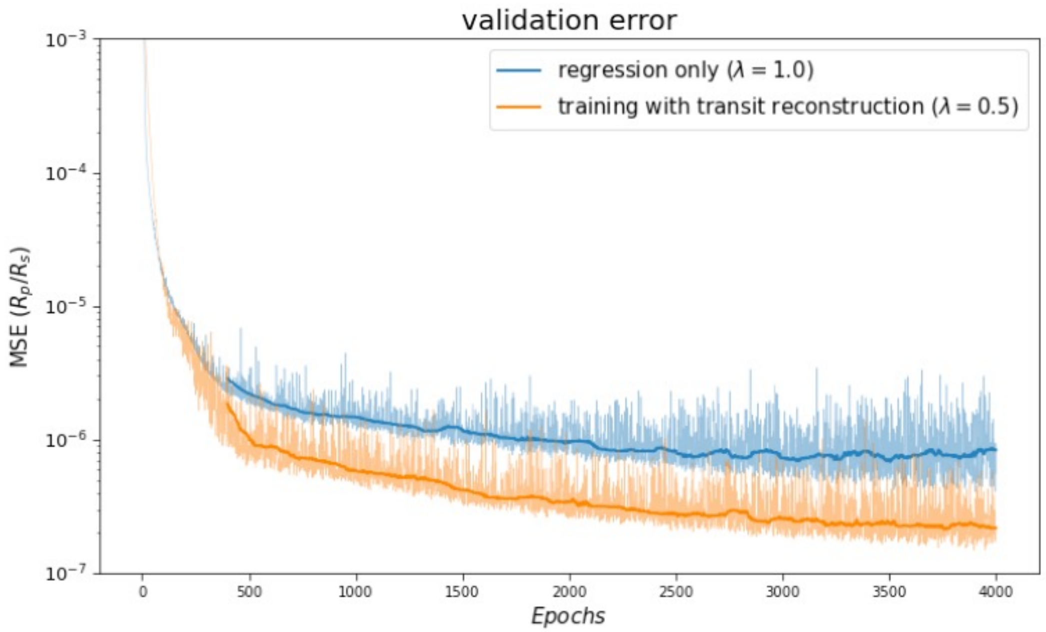
Experiment: inverse transit problem

Loss: **Regression** loss for Model 1 and **Regression + Reconstruction** for Model 2



$\mathcal{L}_{regression}$

$\mathcal{L}_2 = \lambda \mathcal{L}_{regression} + (1 - \lambda)\mathcal{L}_{reconstruction}$

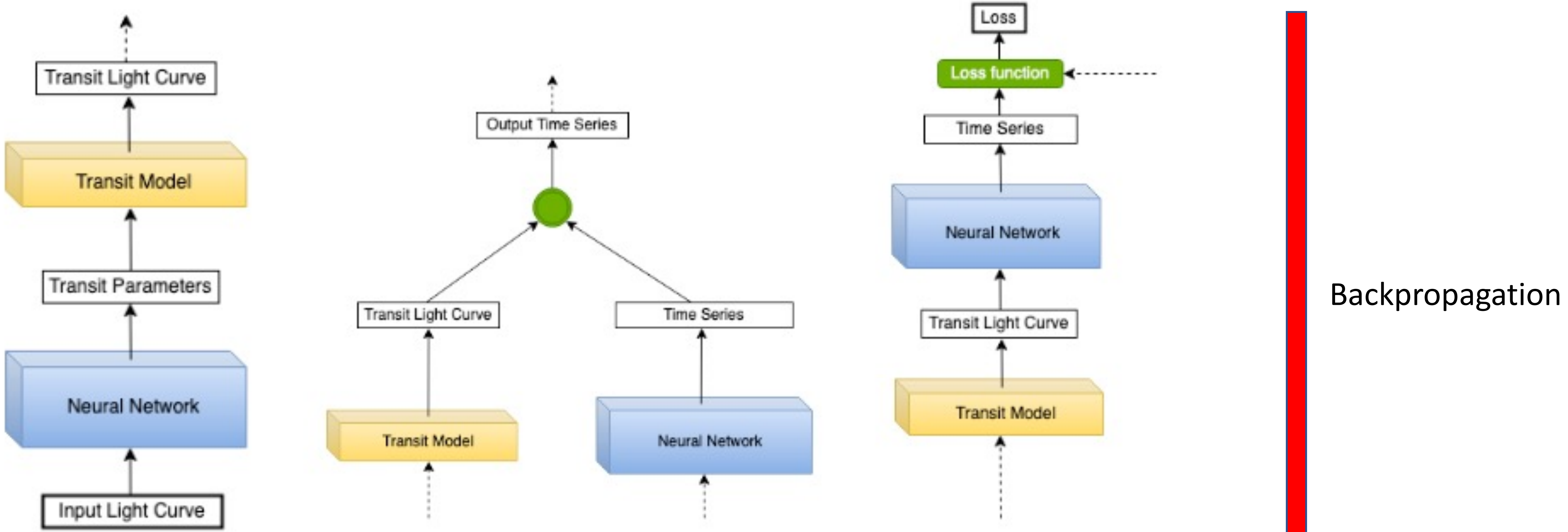


Learning curves

- Test performance improved by the use of reconstruction transit loss
- Bonus: use of physics model improve generalisability

Morvan et al 2021

Hybrid DL-transit structures



(a) Neural net. in input

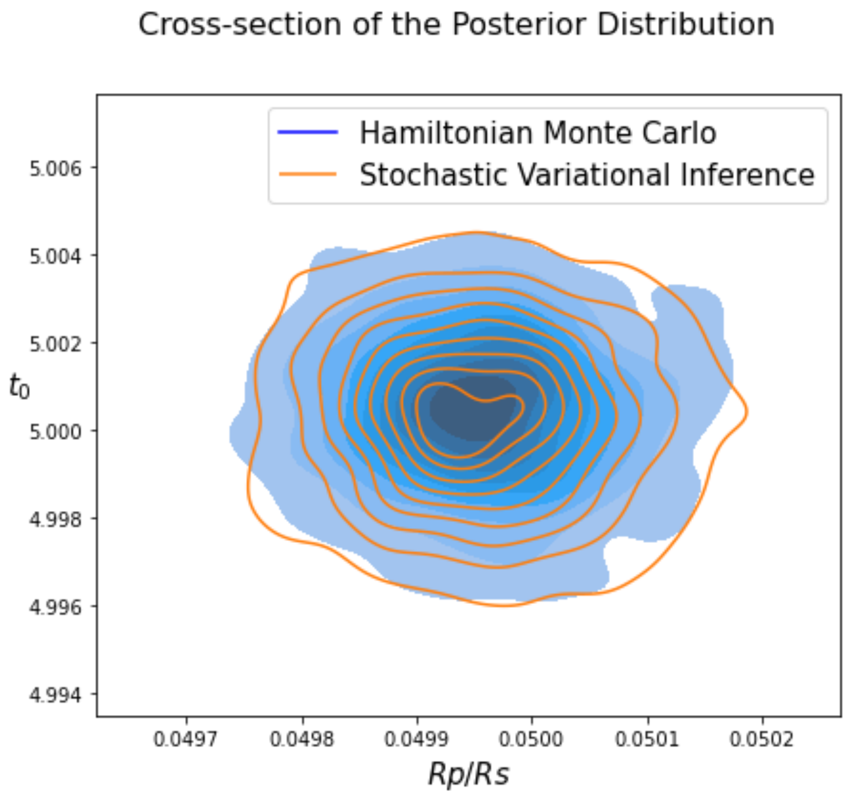
(b) Neural net. in parallel

(c) Neural net. in output

Ref: M2 2020

Morvan et al 2021

Exact VS Approximate Bayesian Inference



Posterior distributions for a transit inference problem using *PylightCurve-Torch* and *Pyro*

Pros of SVI:

- More efficient and **scalable** than MCMC
- Optimised along with Neural Nets

Caveats:

- True **posterior** not necessarily captured
- SVI can suffer from **under-dispersion**

Conclusions

- **Neural Nets flexible** to model instrumental and stellar noise
- Inference: **differentiable transit model** combined w neural nets, loosing exact Bayesian inference but gaining end-to-end
- Future work on design & test of **end-to-end hybrid** models
- …Pending progress in Generative Models of Time Series…

…To be continued…

Thank you!