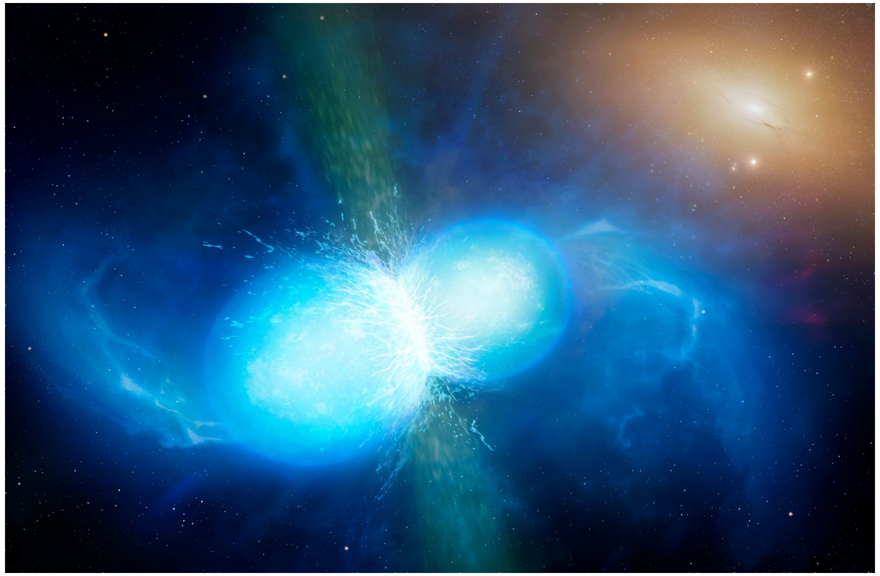
Rapid prediction of kilonova light curves from gravitational wave signals for observed binary neutron star coalescences

Xiaofei Dong[1], Ik Siong Heng[1], Gavin P. Lamb[2], Chris Messenger[1]

[1]SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, UK
[2]Astrophysics Research Institute, Liverpool John Moores University, IC2 Liverpool Science Park, 146 Brownlow Hill, Liverpool L3 5RF, UK



(Credits: University of Warwick/Mark Garlick, wiki)

Xiaofei Dong
x.dong.2@research.gla.ac.uk

Institution for Gravitational Research
University of Glasgow

GEMMA 2
Rome, Italy
18 Sep. 2024

CONTENTS

BACKGROUND

- What is kilonova
- Science with kilonova
- Motivation and goal

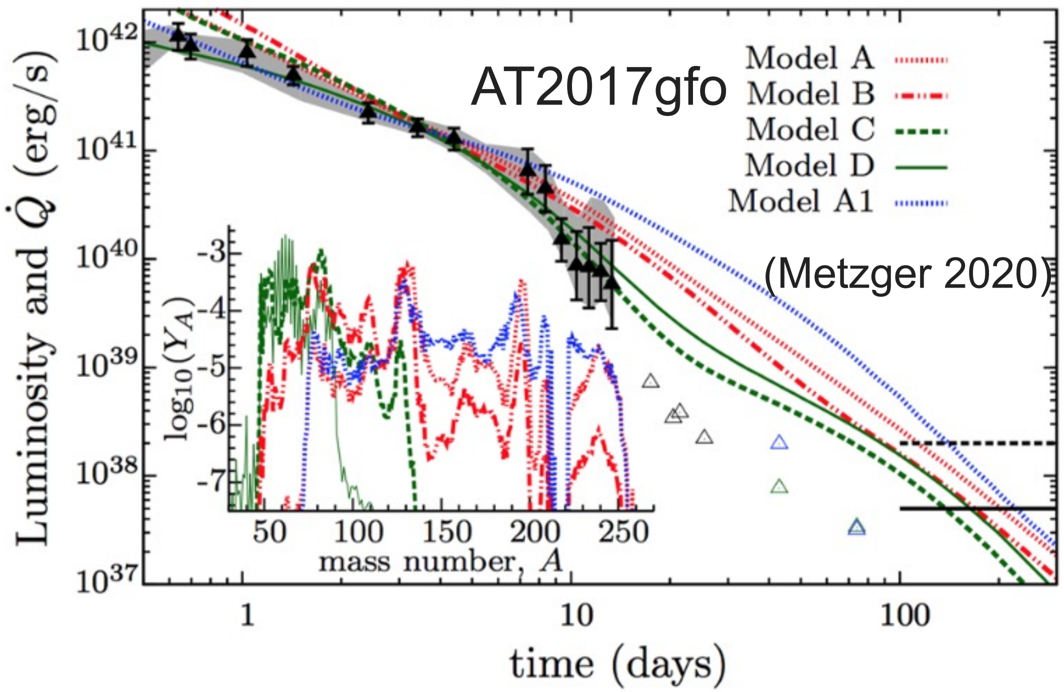
METHOD

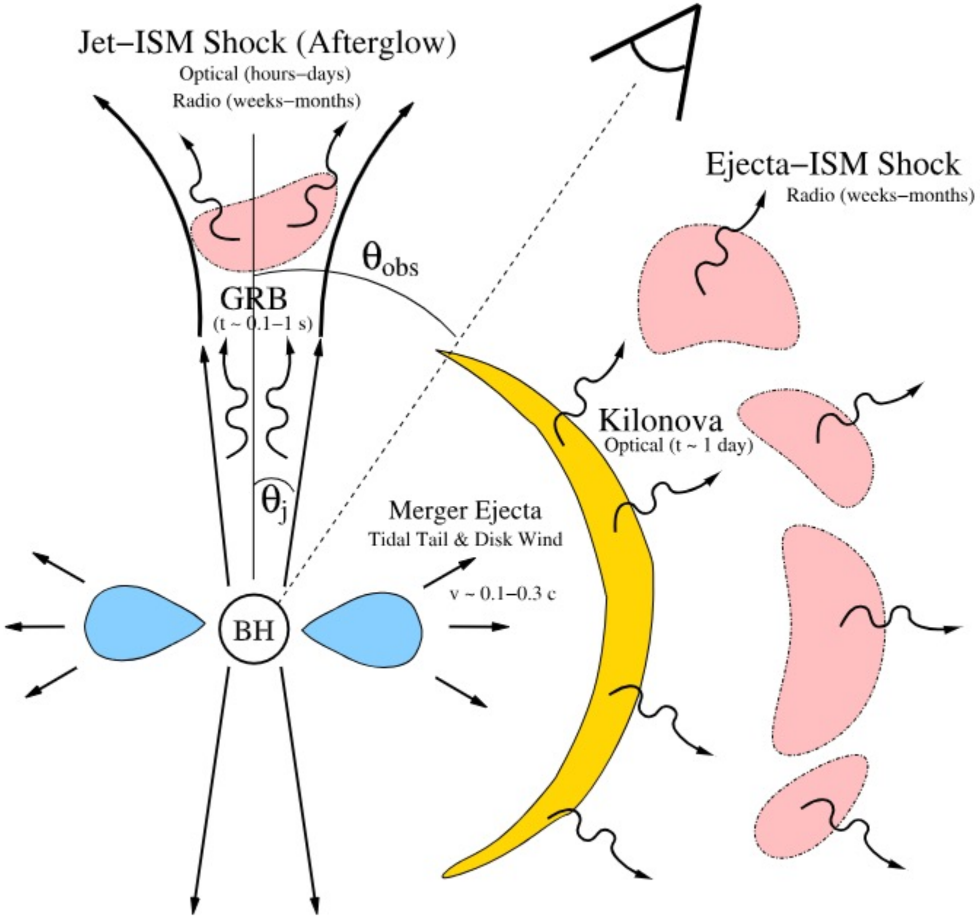
- Kilonova model
- Normalising flow

RESULTS

OUTLOOK

Background: Kilonovae



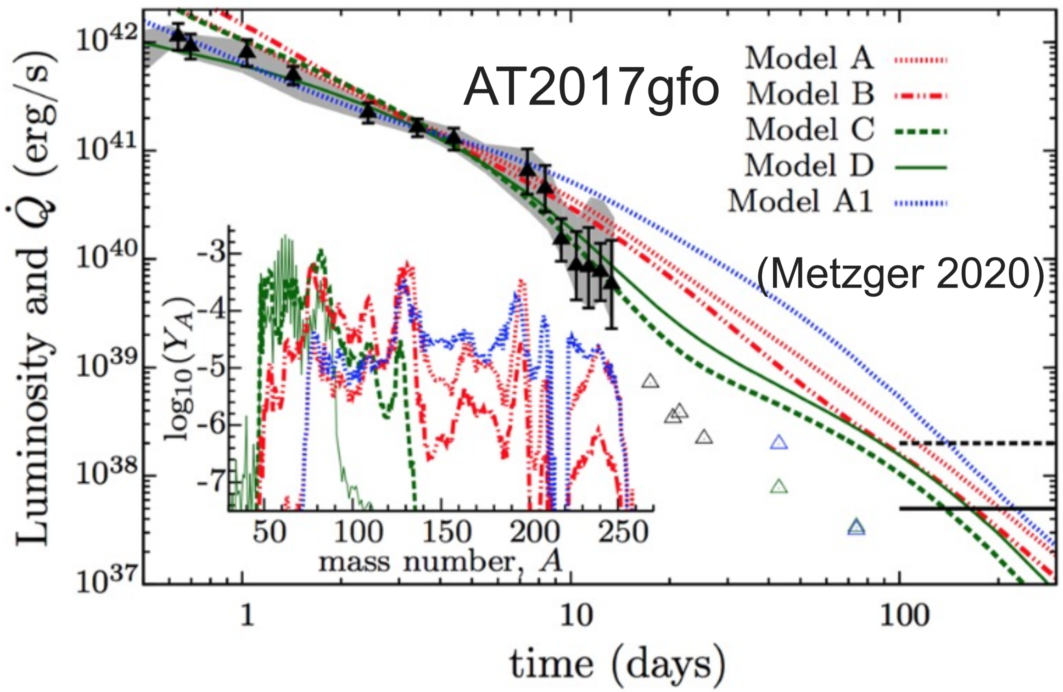


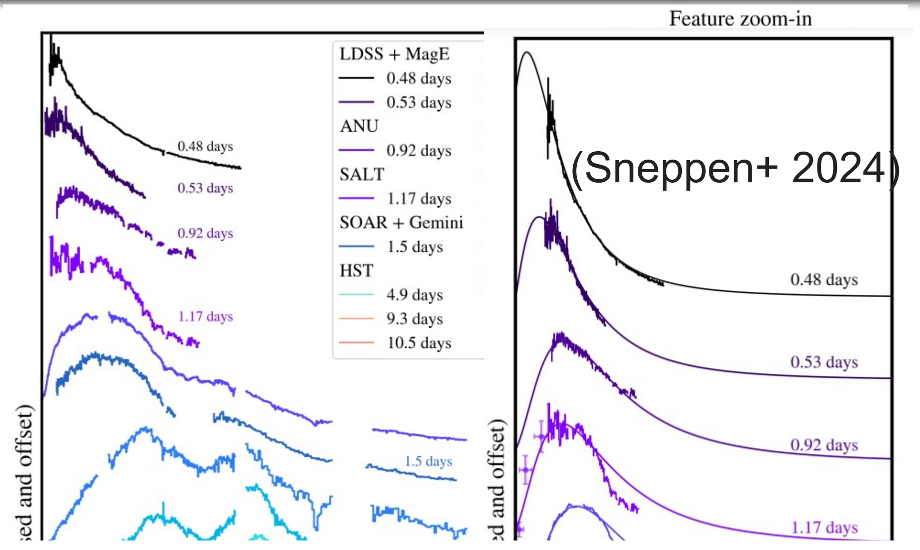
(Metzger 2020)

- $L_{\rm peak} \sim 10^{40-42}$ erg s^{-1}; optical and near-infrared
- Fast evolving: ~day to ~weeks
- Radioactive decay of the *r*-process elements
- From the merger of compact objects, involving at least one neutron star
- Associated with gravitational waves (GWs) and Gamma-ray bursts (GRBs)

Xiaofei Dong
x.dong.2@research.gla.ac.uk
University of Glasgow

Background: Kilonovae



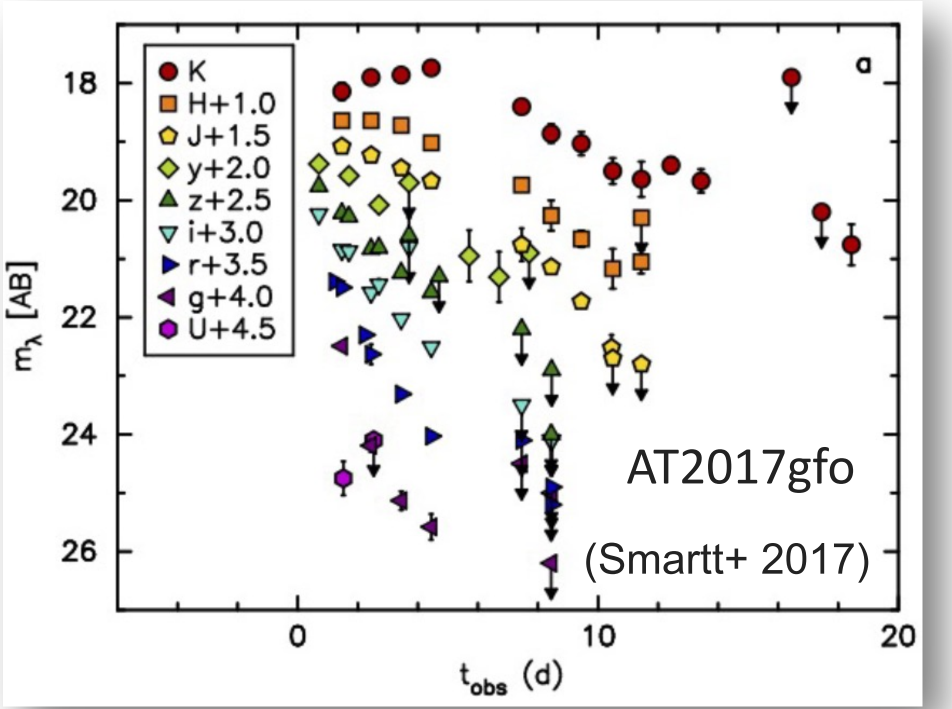


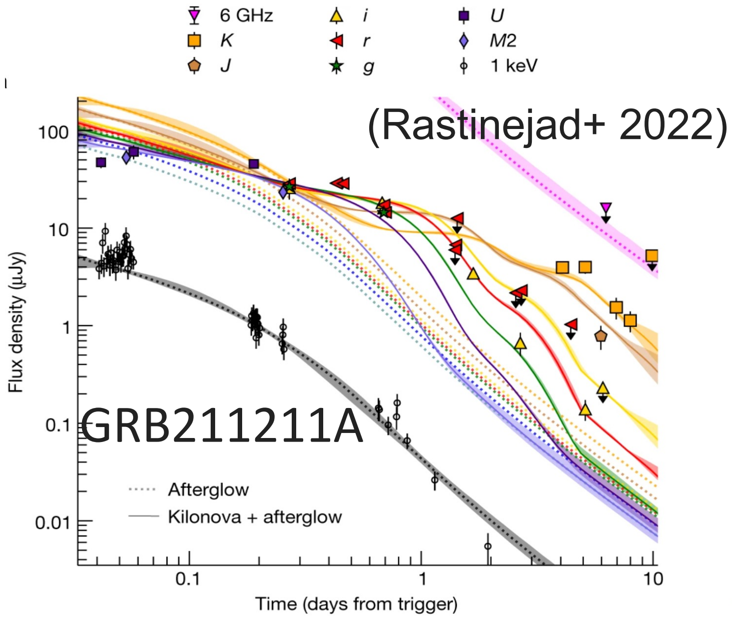
Time (days)	Telescope/Instrument	T_{obs} (K)	Publication
0.48	Magellan/LDSS	$11\,000^{+3400}_{-900}$ [1]	Shappee et al. (2017)
0.53	Magellan/MagE	9300 ± 300 [1]	Shappee et al. (2017)
0.92	ANU/WiFeS	6800 ± 200 [2]	Andreoni et al. (2017)
1.17	SALT/RSS	6400 ± 110	Buckley et al. (2018)
1.43	VLT/X-shooter	5440 ± 60	Pian et al. (2017)
1.45	VLT/X-shooter	5380 ± 60	Sneppen et al. (2023a)
1.47	SOAR/GHTS, Gemini/FLAMINGOS-2	5330 ± 60	Nicholl et al. (2017), Chornock et al. (2017)
2.42	VLT/X-shooter	3940 ± 50	Smartt et al. (2017)
3.41	VLT/X-shooter	3420 ± 40	Pian et al. (2017)
4.40	VLT/X-shooter	3330 ± 40	Smartt et al. (2017)
5.40	VLT/X-shooter	3070 ± 40	Pian et al. (2017)

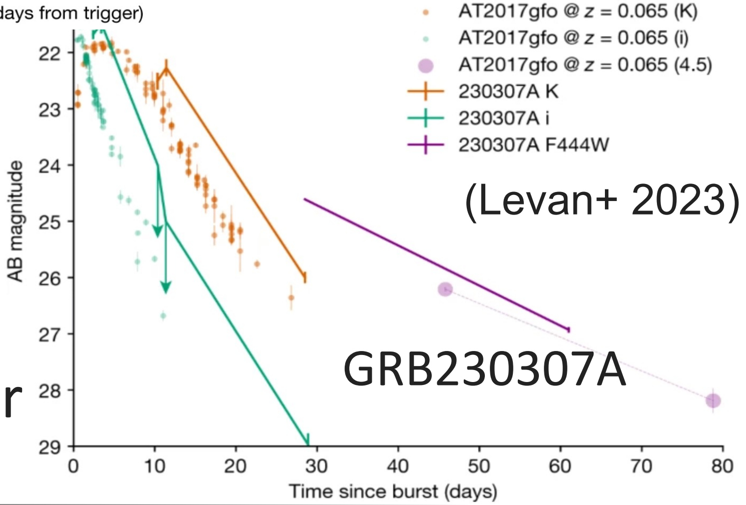
- $L_{peak} \sim 10^{40-42}$ erg s^{-1}; optical and near-infrared
- **Fast evolving**: ~day to ~weeks
- Radioactive decay of the *r*-process elements
- From the merger of compact objects, involving at least o
- Associated with gravitational waves (GWs) and Gamma-ra

Xiaofei Dong
x.dong.2@research.gla.ac.uk

Background: Kilonovae





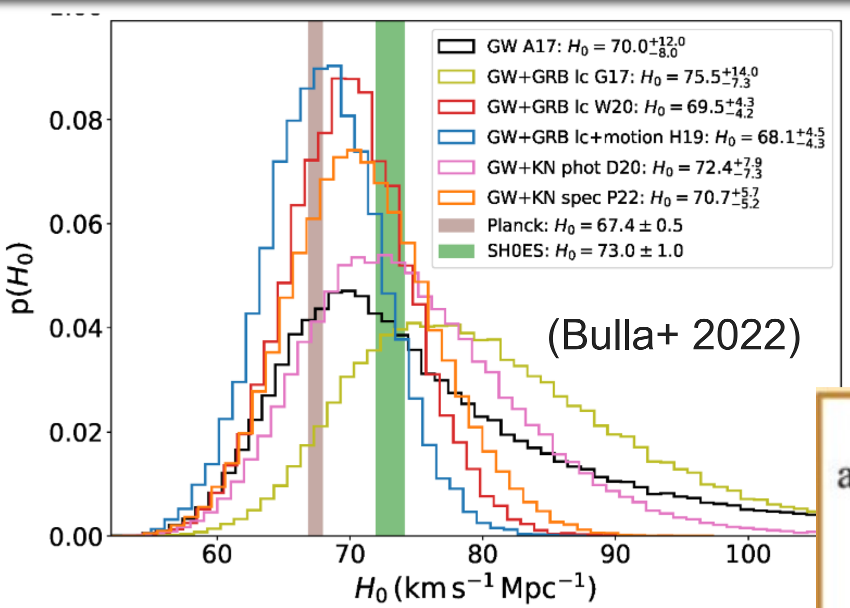


- AT2017gfo
- A few kilonova candidates associated with GRBs
- More evidence that they originate from the same BNS merger

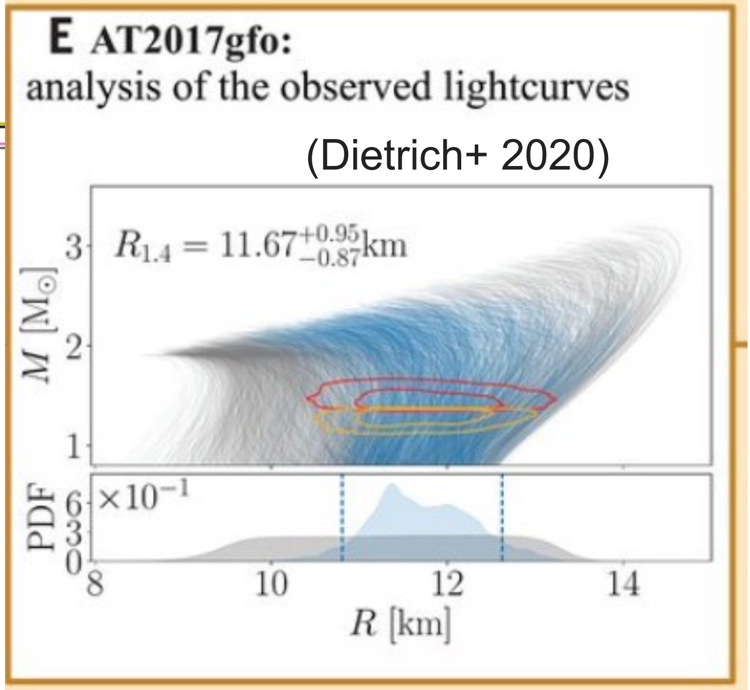
Xiaofei Dong
x.dong.2@research.gla.ac.uk

Background: Science with kilonovae

- Cosmology
  - Bright siren, H_0
- Nuclear matter physics
  - Equation of State
- Nucleosynthesis in the Universe
  - Enrichment of heavy elements from *r*-process
- Compact objects evolution
  - Supermassive/Hypermassive neutron stars
  - Black holes
- …



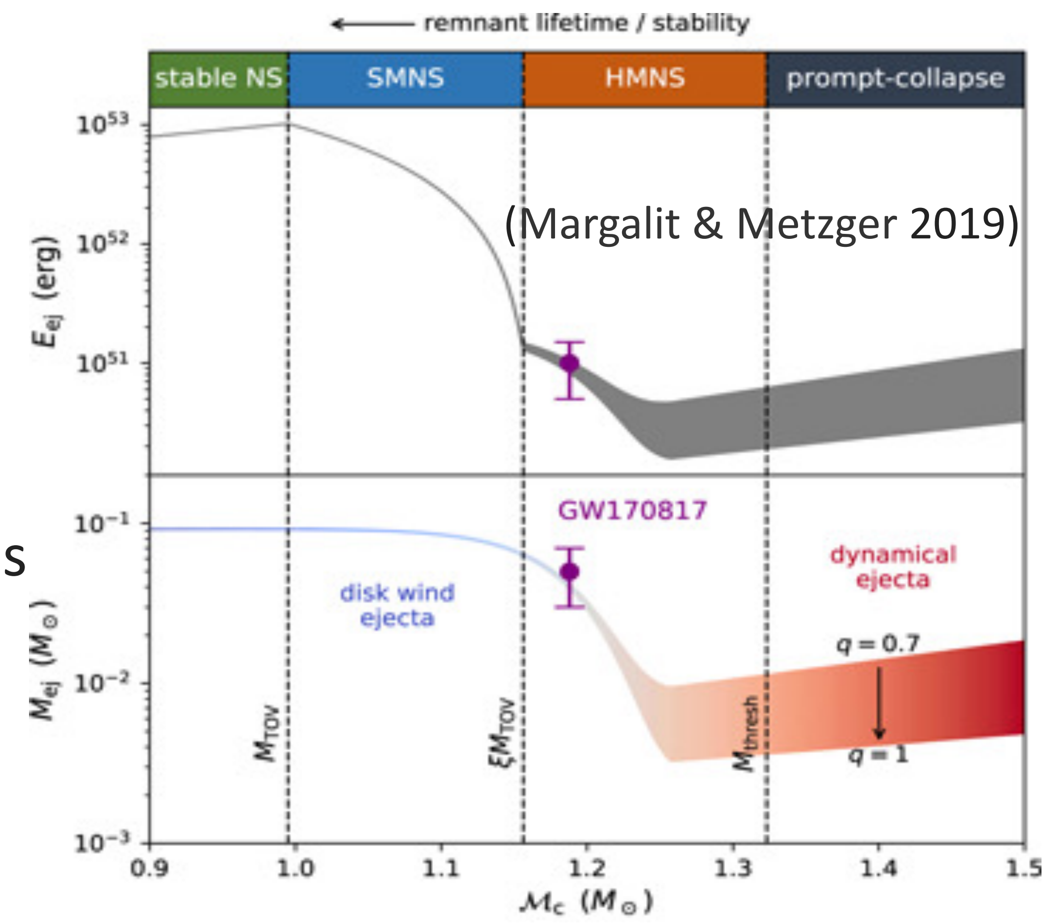
(Bulla+ 2022)



(Dietrich+ 2020)

Background: Science with kilonovae

- Cosmology
  - Bright siren, H_0
- Nuclear matter physics
  - Equation of State
- Nucleosynthesis in the Universe
  - Enrichment of heavy elements from *r*-process
- Compact objects evolution
  - Supermassive/Hypermassive neutron stars
  - Black holes
- …



Goal

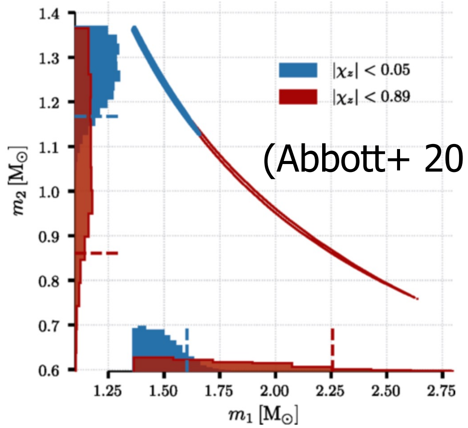
- Only one confirmed kilonova from the binary neutron star (BNS) merger
- Numerical simulations are usually time-consuming
- **Our goal**: **Rapidly predict** kilonova light curves from gravitational wave signals, while **considering model uncertainties**
- Inform EM follow-up strategy

Method

- A **machine learning (ML) tool** to rapidly produce kilonova light curves based on **GW posteriors**: mass (m) and tidal deformability (Λ)
- Factoring in **uncertainties of theoretical kilonova models.**
- Instruct EM follow-up strategy

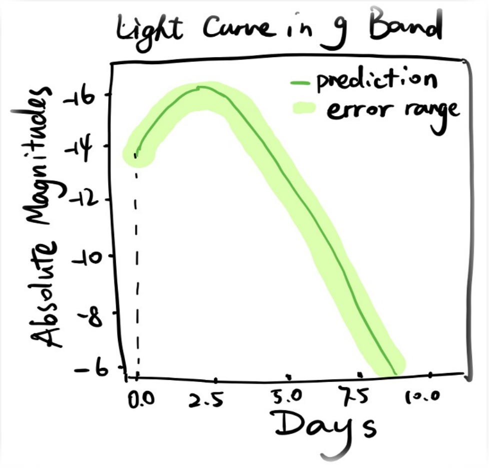
GW Observations

$[m_1, m_2, \Lambda_1, \Lambda_2]$

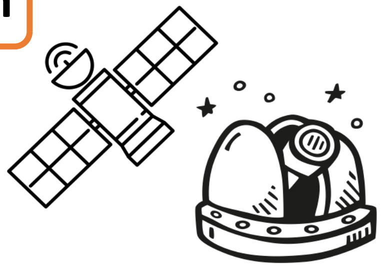


(Abbott+ 2017)

Rapid prediction using a **machine learning model**



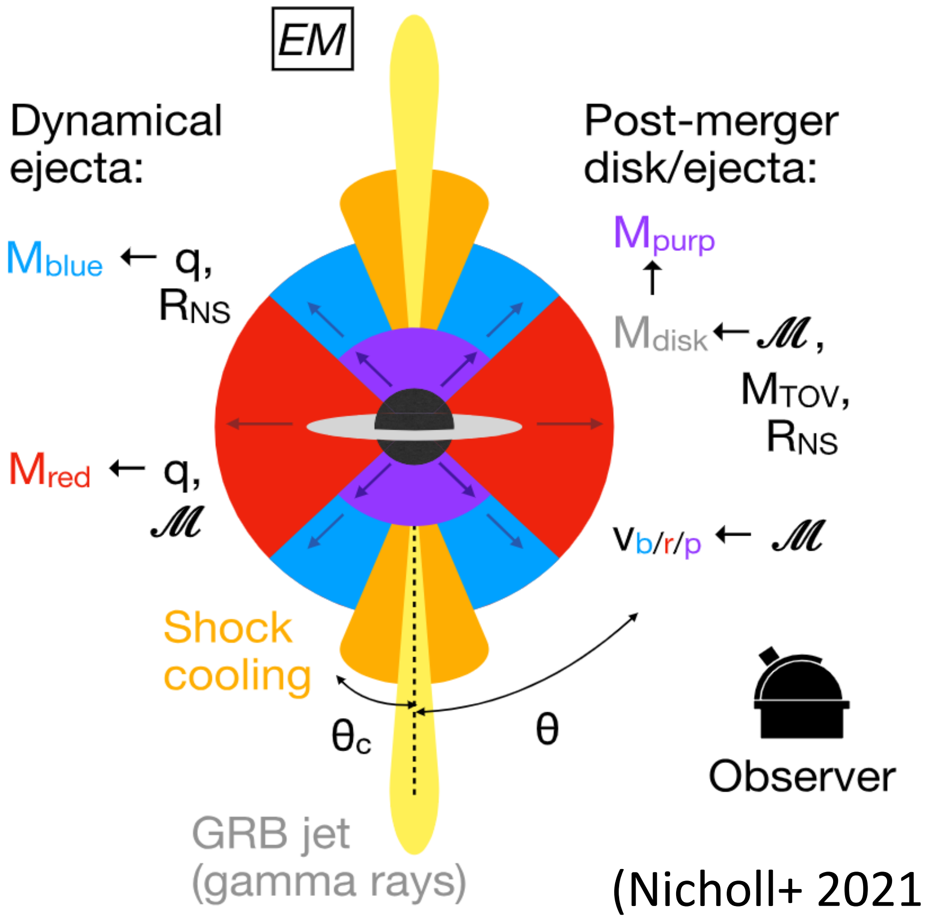
Instruction



EM follow-up

Kilonova model

- BNS model (Nicholl+ 2021)



(Nicholl+ 2021)

Intrinsic

Mass and velocity of the dynamical and disk ejecta

$[m_{\mathrm{ej,red}}, v_{\mathrm{ej,red}}]$

$[m_{\mathrm{ej,blue}}, v_{\mathrm{ej,blue}}]$

Extrinsic

Opacity κ (electron fraction Y_{e})

Viewing angle θ

Shocked cocoon fraction α

Kilonova light curves

Generating kilonova light curves from the theoretical KN model

- BNS model (Nicholl+ 2021)

Intrinsic parameters

$[m_1, m_2, \Lambda_1, \Lambda_2]$

⇩

$[m_{\mathrm{ej}}, v_{\mathrm{ej}}]$

Extrinsic parameters (model-dependent)

Opacity κ (electron fraction Y_{e})
Viewing angle θ
Shocked cocoon fraction α
…

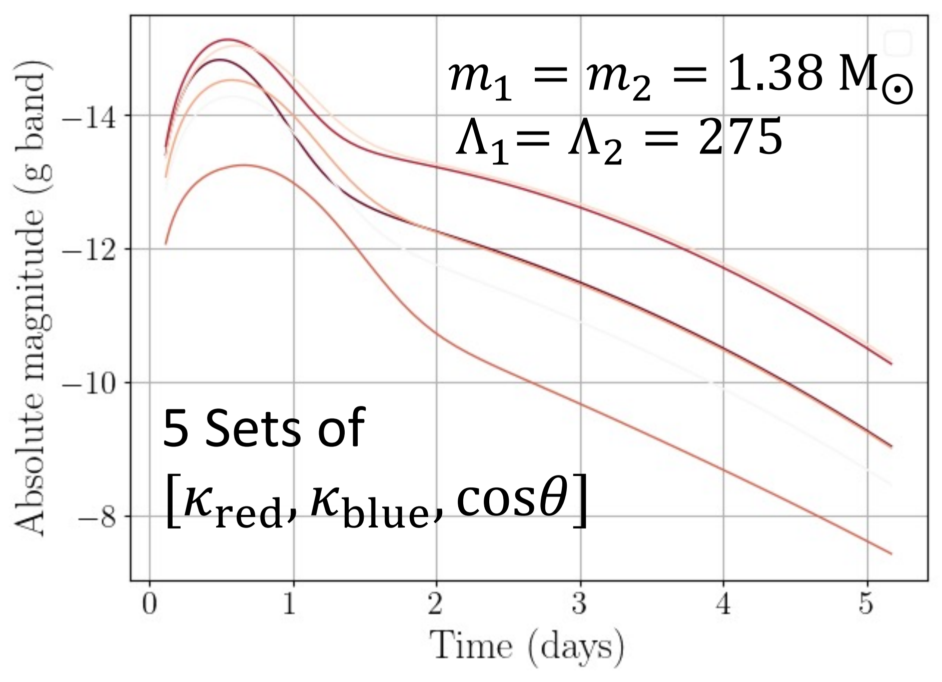
⇩

- $\kappa_{\mathrm{red}} \sim U(1.0, 20)$
- $\kappa_{\mathrm{blue}} \sim U(0.01, 1.0)$
- $\cos\theta \sim U(0.0, 1.0)$

Uncertainties!

⇨ Embedded into light curves

Kilonova light curves generated using REDBACK (Sarin+ 2024)

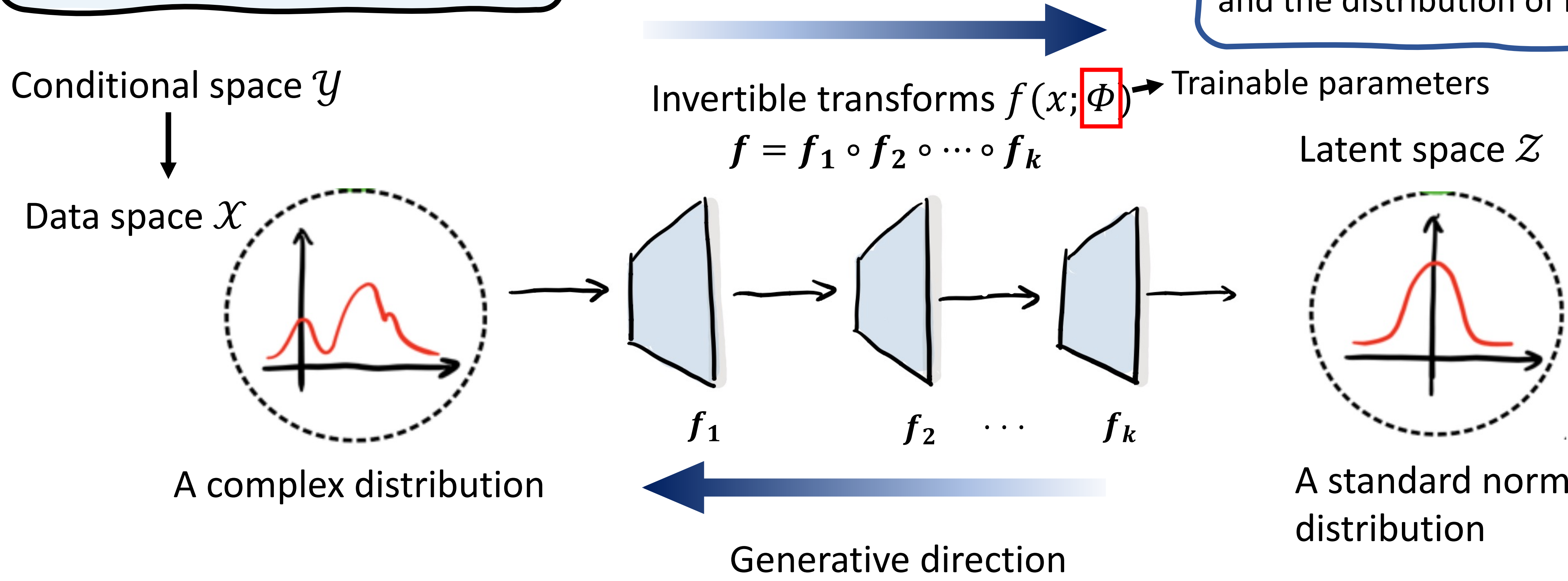


Xiaofei Dong
x.dong.2@research.gla.ac.uk

Method: What is a normalising flow?

The normalising flow machine learning method

Goal: minimise the '**distance**' between the learned distribution and the distribution of input data.

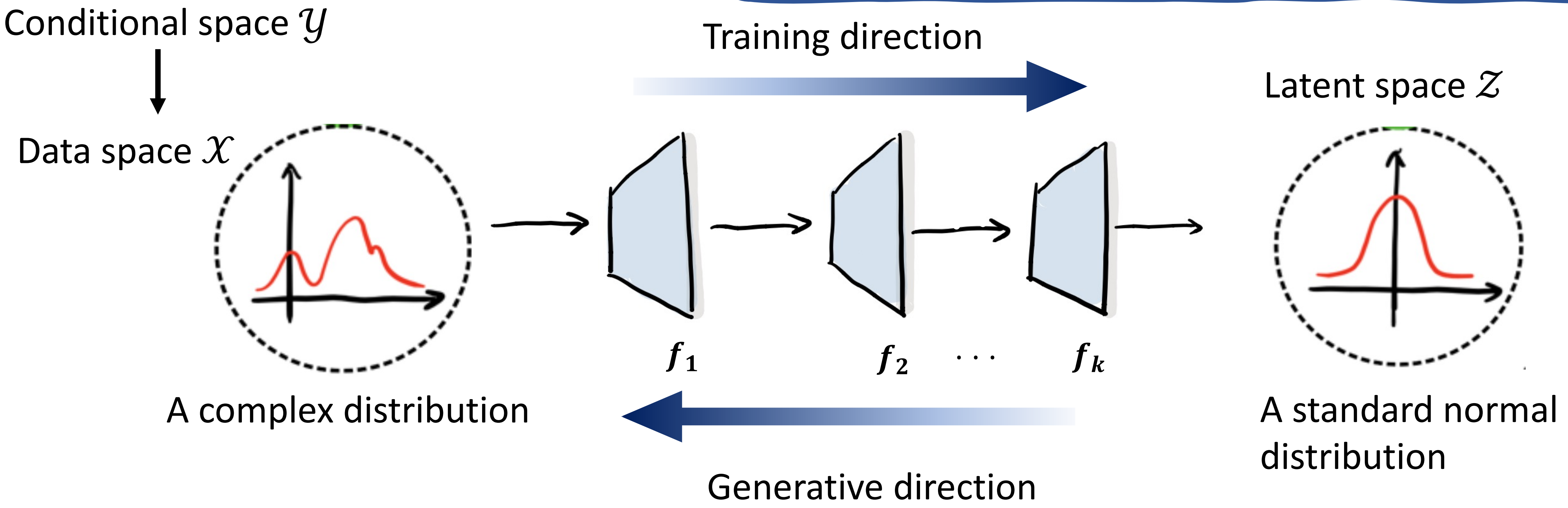


Method: What is a normalising flow?

The normalising flow machine learning method

'Distance' : KL divergence

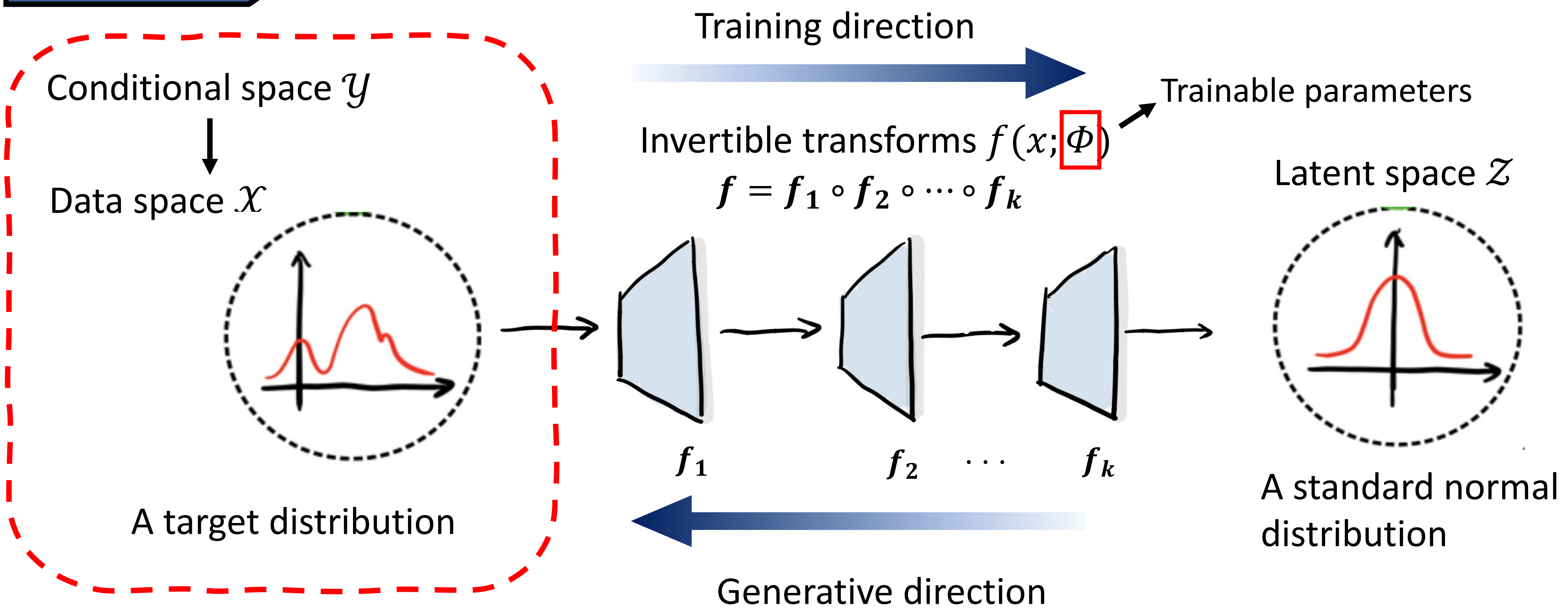
$$\text{Loss} \approx -\frac{1}{K}\sum_{k=1}^{K} \ln p_z\big(f(x|y;\phi)\big) + \ln\left|\det\frac{\partial f(x|y;\phi)}{\partial x}\right|$$



Xiaofei Dong
x.dong.2@research.gla.ac.uk

Method: What is a normalising flow?

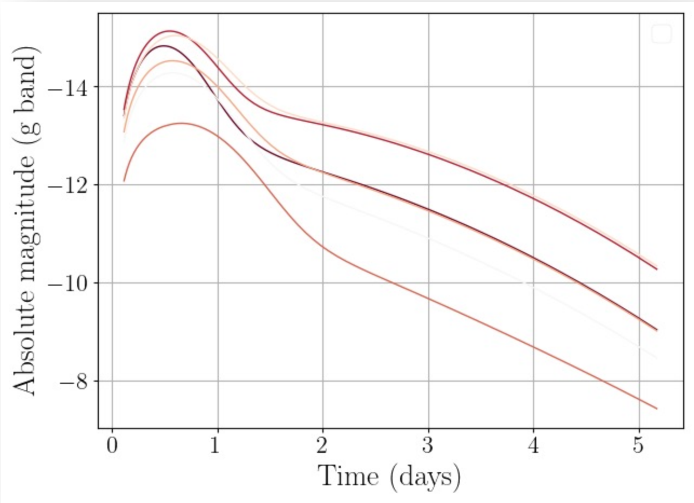
Training



Method: Training process

Training

Training data space $\mathcal{X}$:
~10^6 light curves, each contains 30 data points



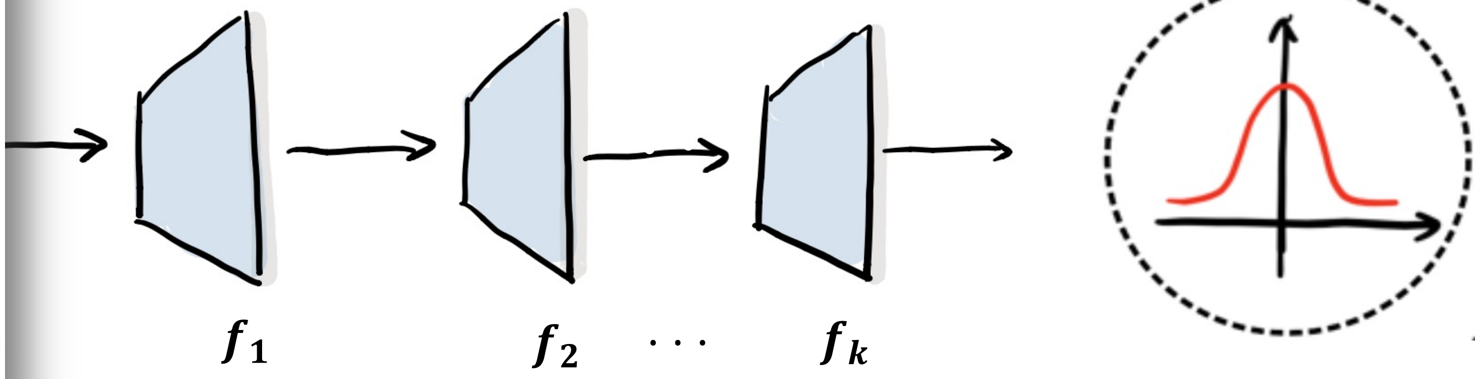
Conditional space $\mathcal{Y}$:
$[m_1, m_2, \Lambda_1, \Lambda_2]$

Training direction

Trainable parameters

Invertible transforms $f(x; \Phi)$

$\boldsymbol{f} = \boldsymbol{f_1} \circ \boldsymbol{f_2} \circ \cdots \circ \boldsymbol{f_k}$



Generative direction

Latent space $\mathcal{Z}$

A standard normal distribution

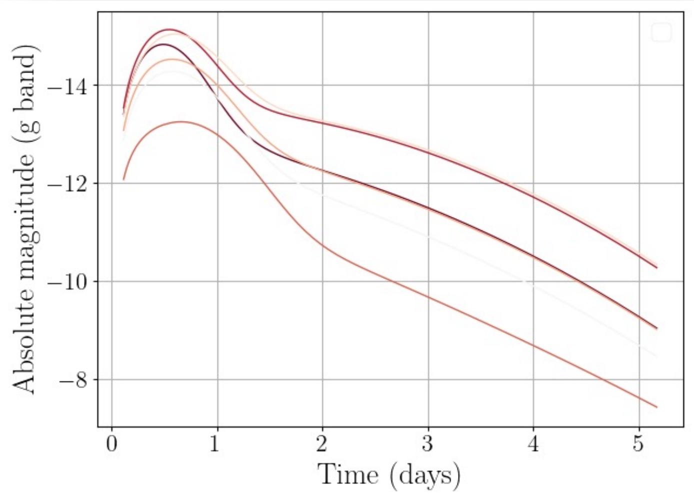
Method: Training process

Training

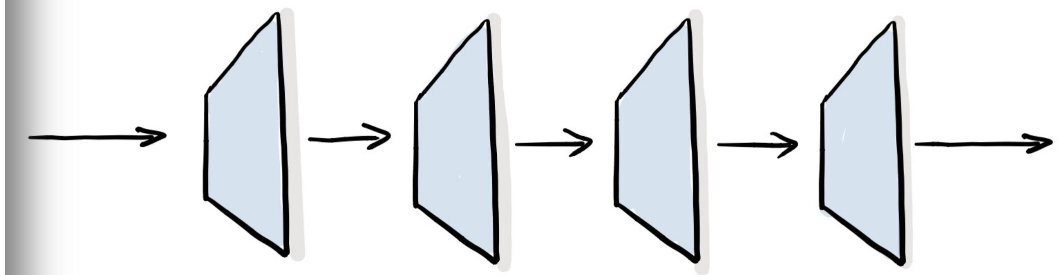
~4 hours training on a GPU

We use GLASFLOW, a normalising flow collection (Williams+, 2024). https://github.com/uofgravity/glasflow

Training data space $\mathcal{X}$:
~10^6 light curves, each contains 30 data points

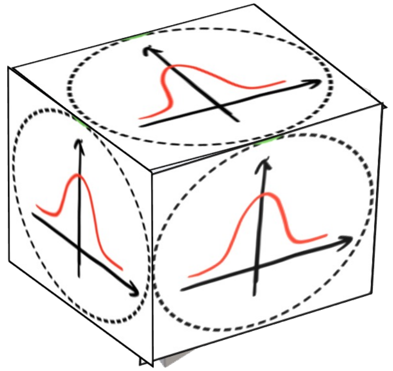


Forward mapping through $f: p(x|y) \rightarrow p(z|y)$



4 transforms: $f = f_1 \circ f_2 \circ f_3 \circ f_4$

Latent space $\mathcal{Z}$



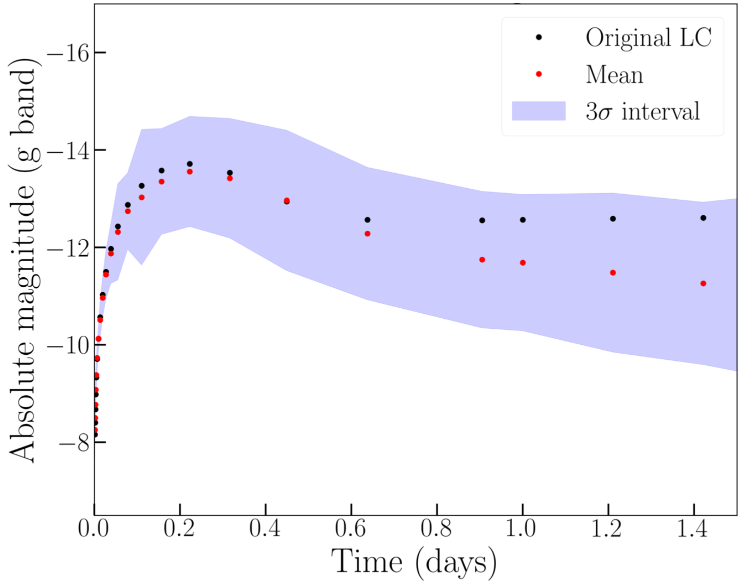
Multi-dimensional standard normal distribution

Conditional space $\mathcal{Y}$:
$[m_1, m_2, \Lambda_1, \Lambda_2]$

Method: Predicting process

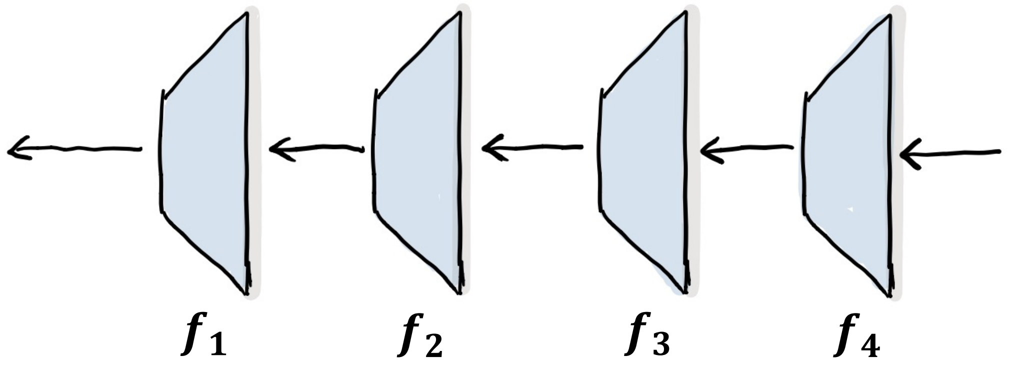
Predicting

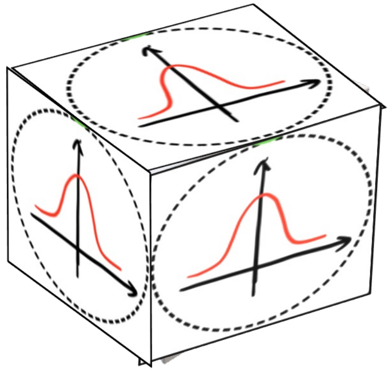
Data space $\mathcal{X}$:
KN light curves
with uncertainties



~ seconds

Apply inverse mapping through f^{-1}: $p(z|y) \longrightarrow p(x|y)$





Latent space $\mathcal{Z}$

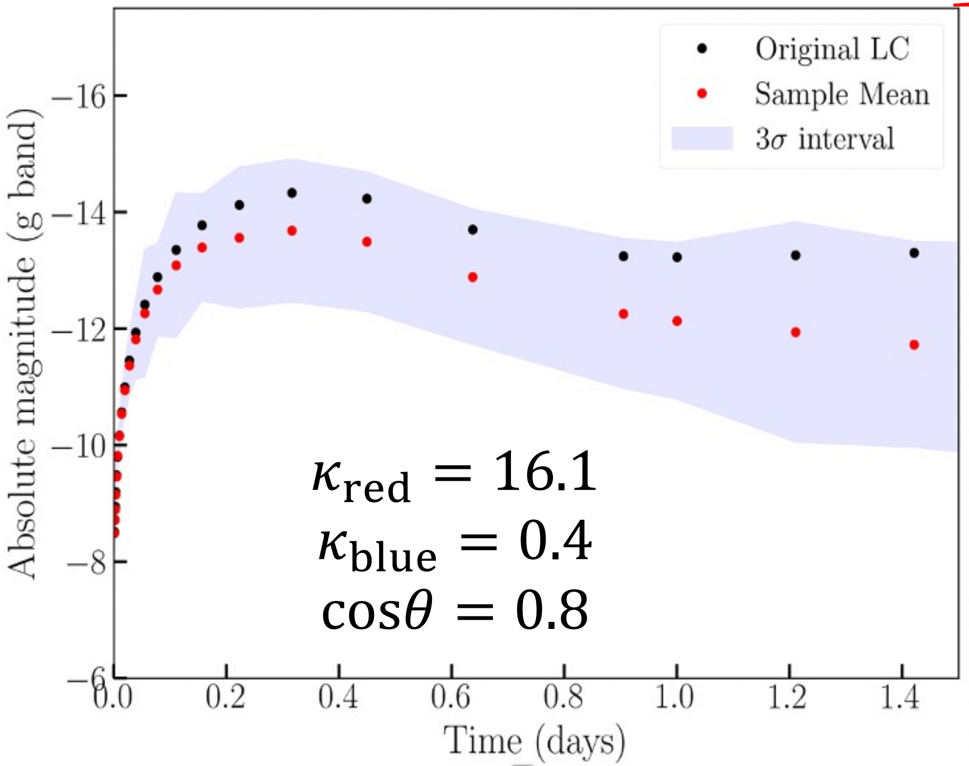
4 transforms: $\boldsymbol{f^{-1} = f_4^{-1} \circ f_3^{-1} \circ f_2^{-1} \circ f_1^{-1}}$

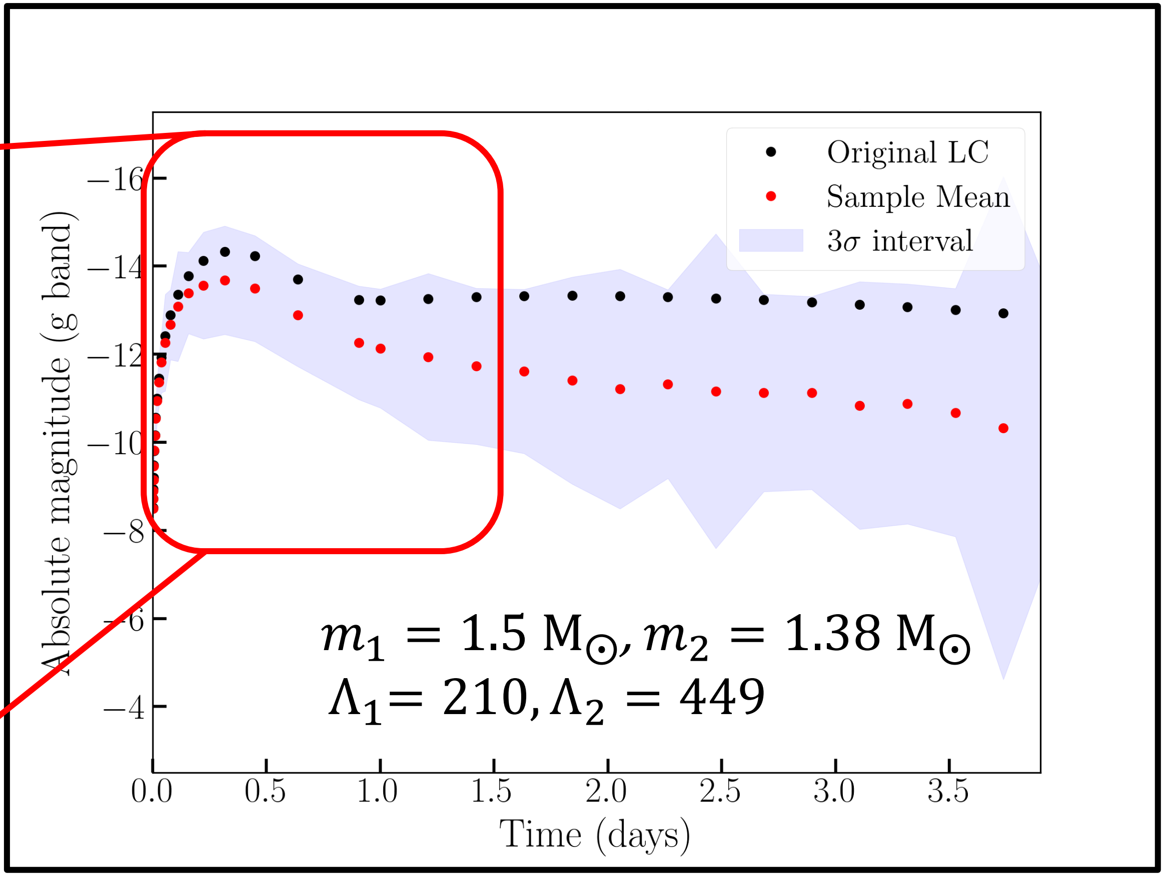
Conditional labels $\mathcal{Y}$:
$[m_1, m_2, \Lambda_1, \Lambda_2]$

Xiaofei Dong
x.dong.2@research.gla.ac.uk

Results

Training LC in g band



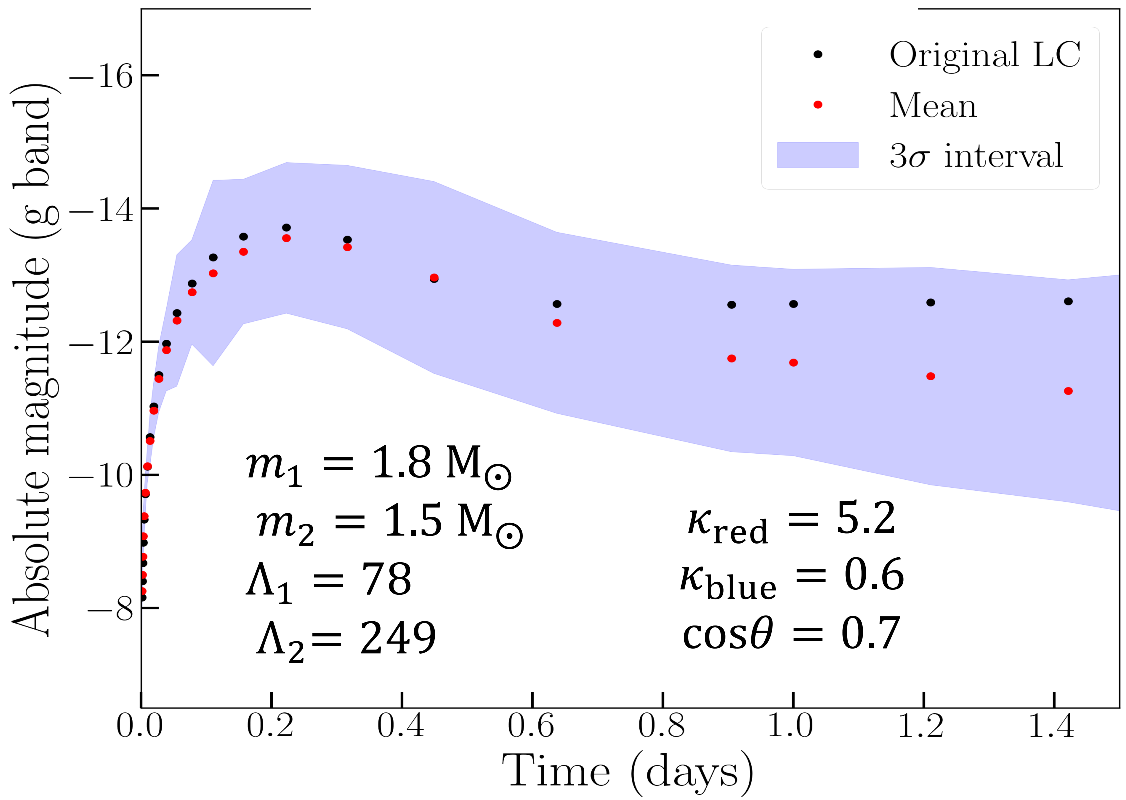


- The predicted light curve for training data falls in the error range.

The mean light curve (red dots) of 1000 predictions (the translucent purple range) in the g band.

Results

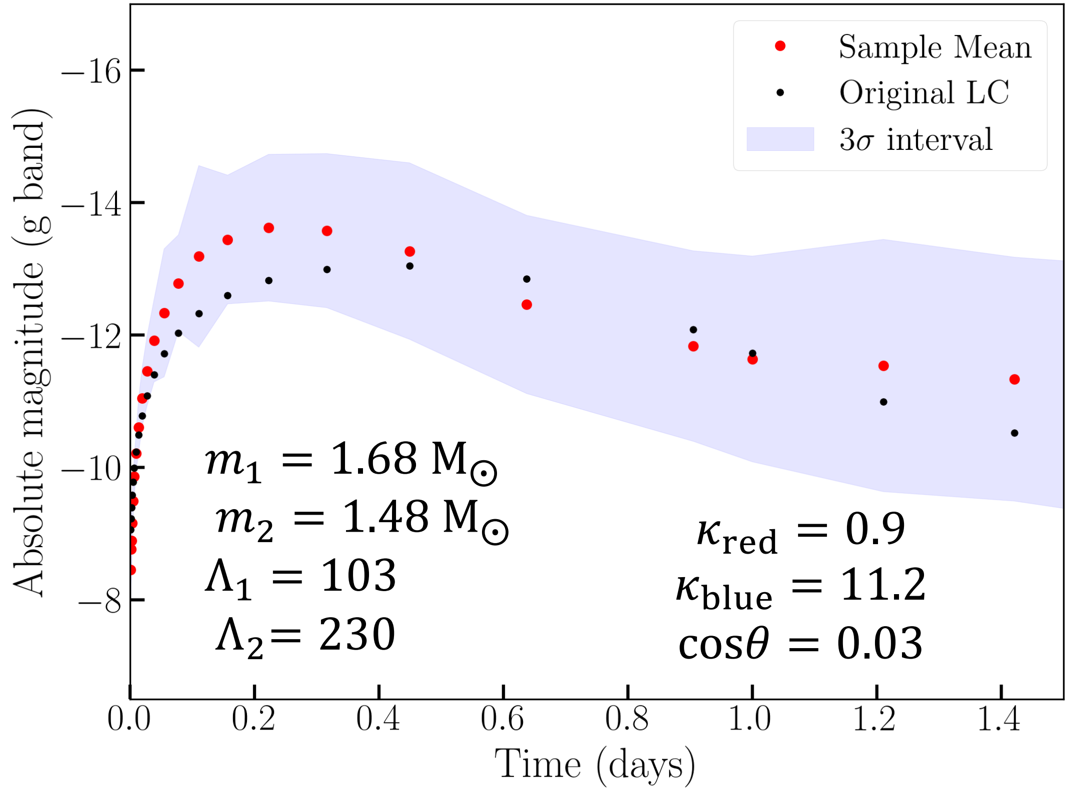
Testing curve in g band



- Predictions on a set of **unseen** intrinsic parameters $[m_1, m_2, \Lambda_1, \Lambda_2]$.
- The uncertainties from extrinsic parameters are embedded in the error range
- The predicted magnitude span is ~2 mag during the peak for the BNS model

Results

Prediction in g band



- Predictions on a set of **unseen** intrinsic parameters from **a different EoS**
- Reasonable predictions for intrinsic parameter sets generated within the same prior range as the training data
- Include more EoSs…

Summary and outlook

- These initial results indicate the promising potential of normalising flow as an alternative method for **rapidly predicting** kilonova light curves
- Taking into account model **uncertainties**, more flexible
- Model-independent: can be applied to any kilonova model

Future focuses

- Increase complexities: more numerical models (e.g., Bulla+ 2023), multiple EoSs…
- Test the model on multi-messenger observations of GW170817
- Explore collaborations with other ML tools, such as DINGO (Maximilian+ 2021), for rapid mass-lambda posteriors
- Long-term goal: possible implementation in LVK public alerts

Backup 1: GW posteriors to ejecta properties

$$q = M_2/M_1$$

$$\mathcal{M} = (M_1 M_2)^{3/5}(M_1+M_2)^{-1/5}$$

$$C_{\mathrm{NS},i} = 0.36 - 0.0355 \ln(\Lambda_i) + 0.000705 \ln(\Lambda_i)^2.$$

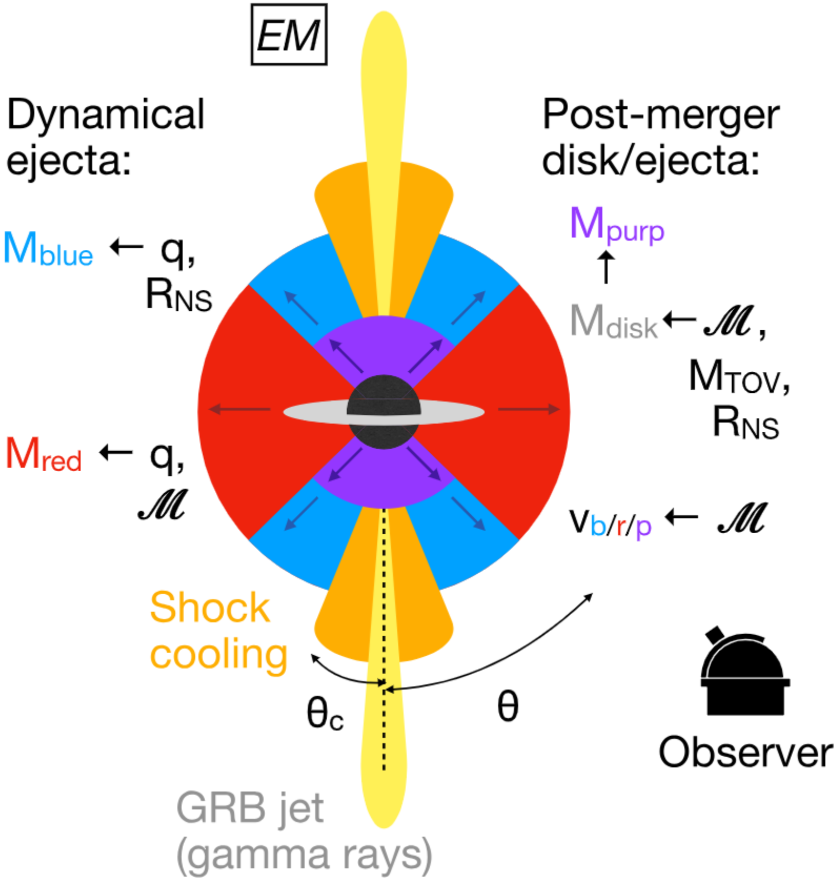
$$M_i^* = M_i + 0.08 M_i^2.$$

$$\frac{M_{\mathrm{ej}}^{\mathrm{fit}}}{10^{-3} M_\odot} = \left[a \left(\frac{M_2}{M_1}\right)^{1/3} \left(\frac{1-2C_{\mathrm{NS},1}}{C_{\mathrm{NS},1}}\right) + b \left(\frac{M_2}{M_1}\right)^n + c \left(1 - \frac{M_1}{M_1^*}\right) \right] M_1^* + (1 \leftrightarrow 2) + d$$

$$M_{\mathrm{thr}} = (2.38 - 3.606 \cdot M_{\mathrm{TOV}}/R_{\mathrm{max}}) \times M_{\mathrm{TOV}}$$

$$R_{\mathrm{rem}} = 11.2 \mathcal{M} \cdot (\tilde{\Lambda}/800)^{1/6}$$

$$\log_{10}(m_{\mathrm{disc}}[M_{\mathrm{tot}}/M_{\mathrm{thr}}]) = \max\left(-3, a\left(1 + b \tanh\left[\frac{c - M_{\mathrm{tot}}/M_{\mathrm{thr}}}{d}\right]\right)\right)$$



(Nicholl+ 2021)

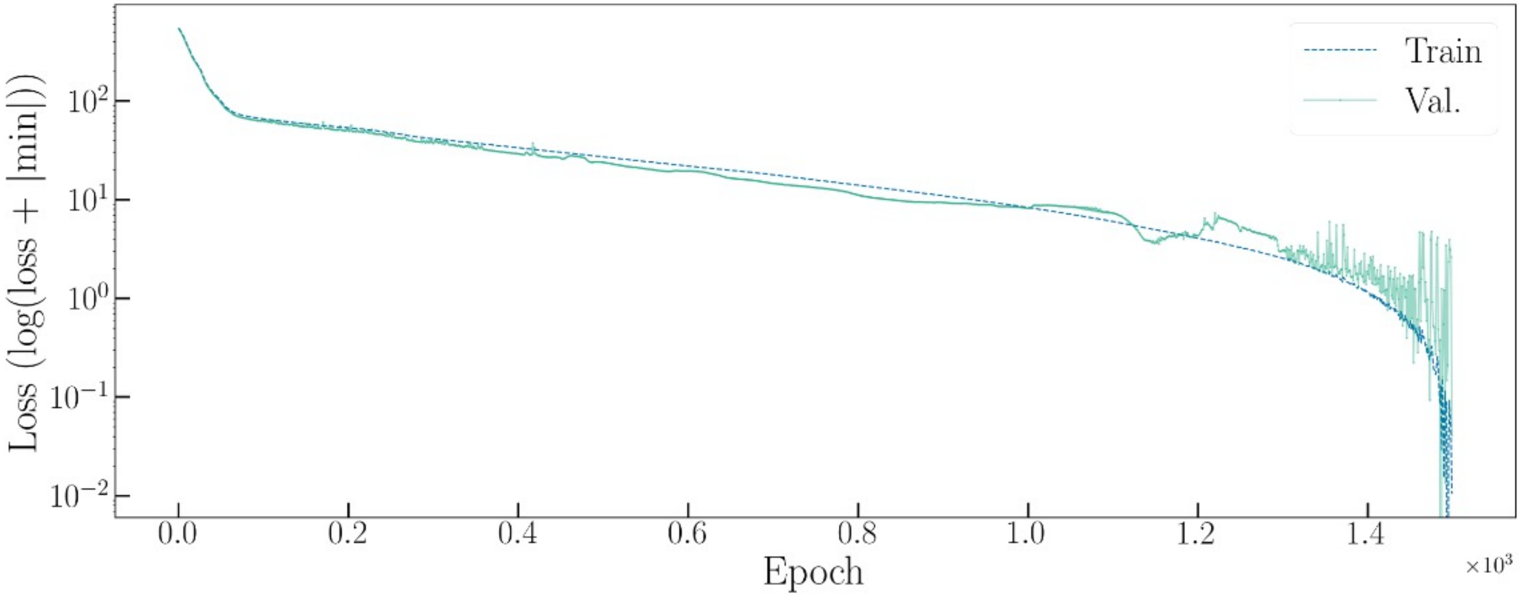
Backup 2: Maths of normalising flow

KL Divergence

- a non-symmetric distance measure between two distributions.

If conditions A in the conditional space $\mathcal{Y}$

$$p_x^*(x) \longrightarrow p_x^*(x|A)$$



$$D_{\mathrm{KL}}[p_x^*(x) \| p_x(x)] = \int_{-\infty}^{+\infty} p_x^*(x) \ln\left(\frac{p_x^*(x)}{p_x(x)}\right) dx$$

$$= -\mathbb{E}_{p_x^*(x)}\left[\ln p_z(f(x;\phi)) + \ln\left|\det \frac{\partial f(x;\phi)}{\partial x}\right|\right] + \text{const.}$$