Revolutionizing the inference of mass, radius, and magnetic field structure of neutron stars using NICER and Fermi-LAT data

Constantinos Kalapotharakos

NASA Goddard Space Flight Center

Collaborators

Zorawar Wadiasingh (UMD, CRESST II, NASA/GSFC)
Greg Olmschenk (UMD, CRESST II, NASA/GSFC)
Thibault Lechien (SURA, CRESST II, NASA/GSFC)
Emily Broadbent (SURA, CRESST II, NASA/GSFC)
Anu Kundu (UMBC, CRESST II, NASA/GSFC)
Dimitrios Skiathas (SURA, CRESST II, NASA/GSFC, Univ. of Patras)
Demos Kazanas (NASA/GSFC)
George Younes (NASA/GSFC)

Alice Harding (LANL)
Soumi De (LANL)
Diane Oyen (LANL)
Matthew Baring (Rice University)
Hoa Dinh (Rice University)
Christo Venter (Northwest University, SA

September 2024

Outline

- Problem overview - Motivation

- Modeling the thermal NICER X-ray light curves

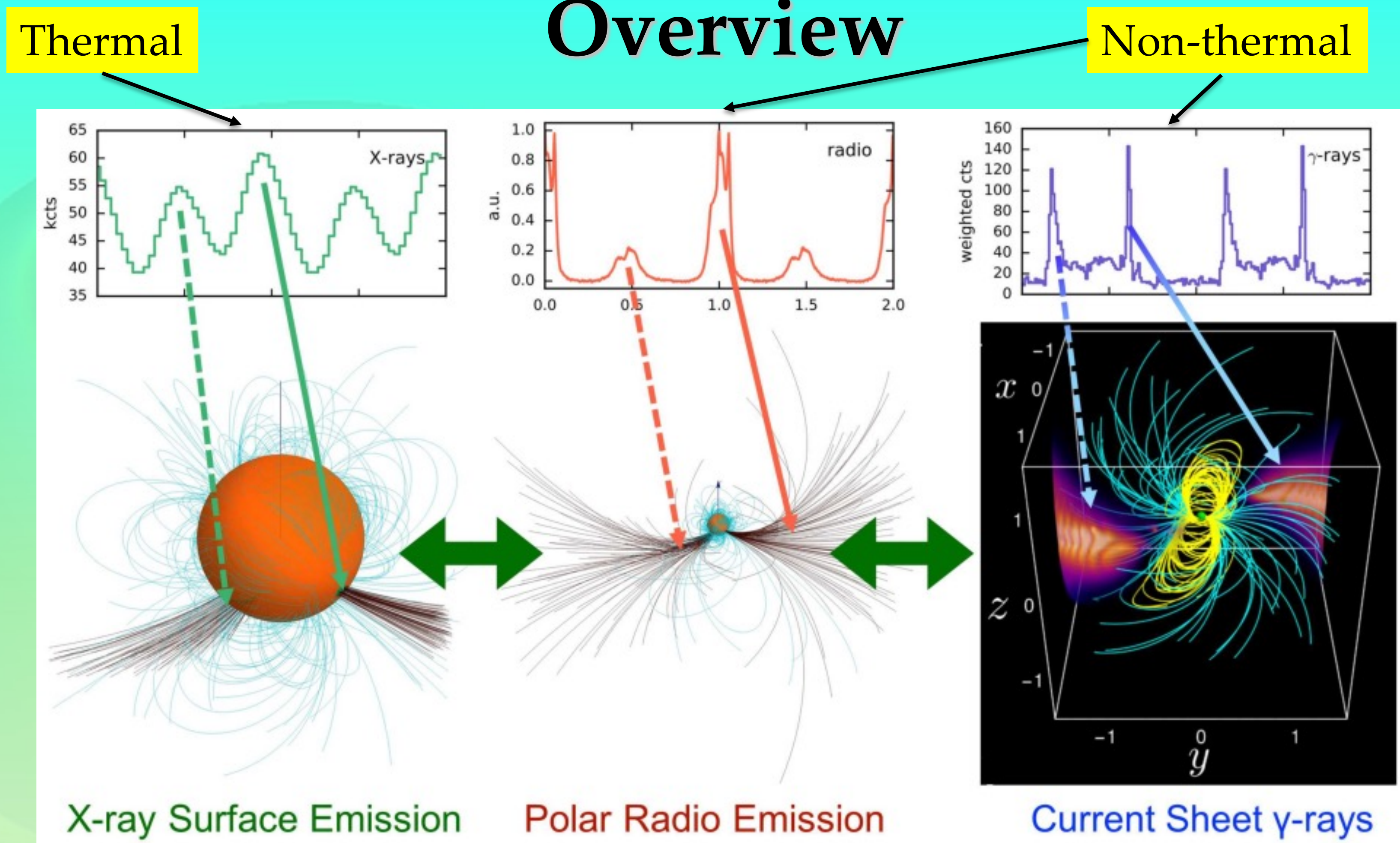
  Constraints: Multipolar magnetic fields, Mass, Radius (Equation of State)

- Revolutionizing the parameter inference by ML techniques

  From unfeasible to feasible

- Summary, future, importance

Overview

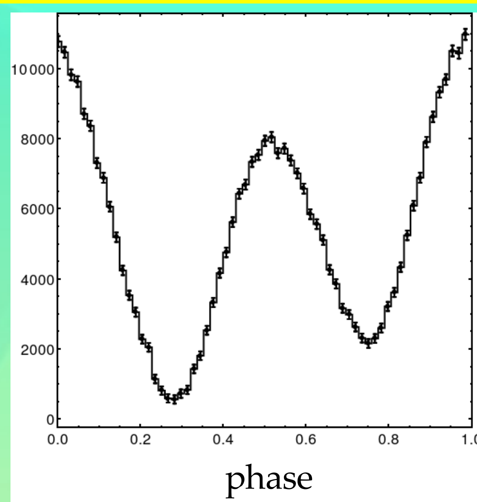

Thermal
Non-thermal
X-rays
kcts
radio
a.u.
γ-rays
weighted cts
x
z
y
X-ray Surface Emission
Polar Radio Emission
Current Sheet γ-rays

NICER results

J0030+0451

~20 parameters

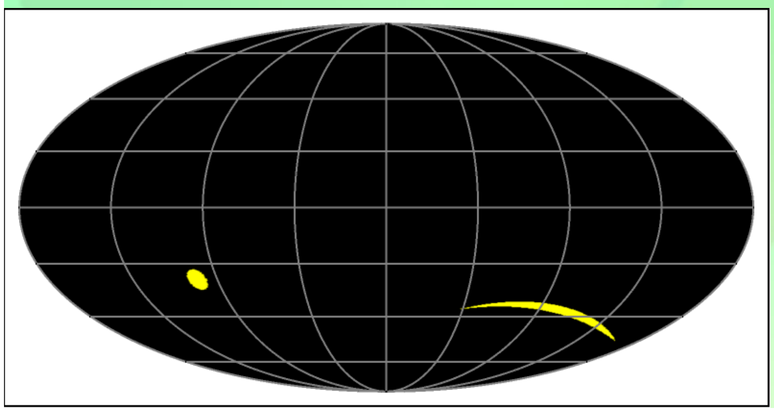
Thermal X-ray LC **Bogdanov et al. (2019)**



Riley et al. (2019)

$M = 1.34 M_{\odot}$
$R = 12.71 km$
$\zeta = 53.85°$

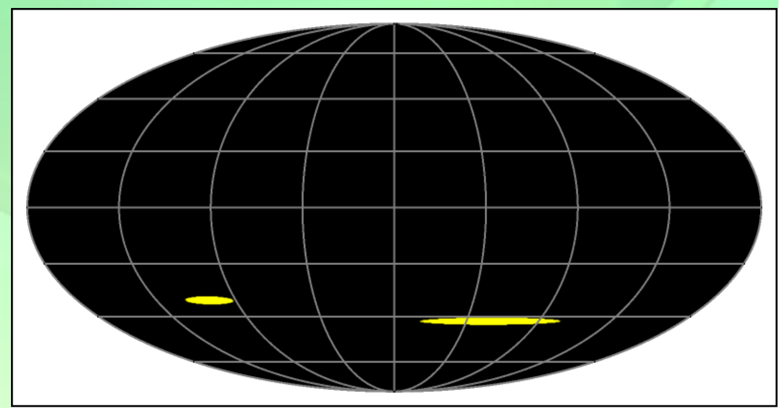
2 hotspots
same temperatures



Miller et al. (2019)

$M = 1.49 M_{\odot}$
$R = 13.64 km$
$\zeta = 47.38°$

2 hotspots
same temperatures
+
3 hotspots

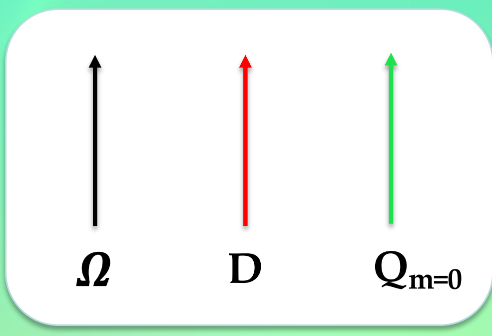


Strong evidence for multipolar magnetic field

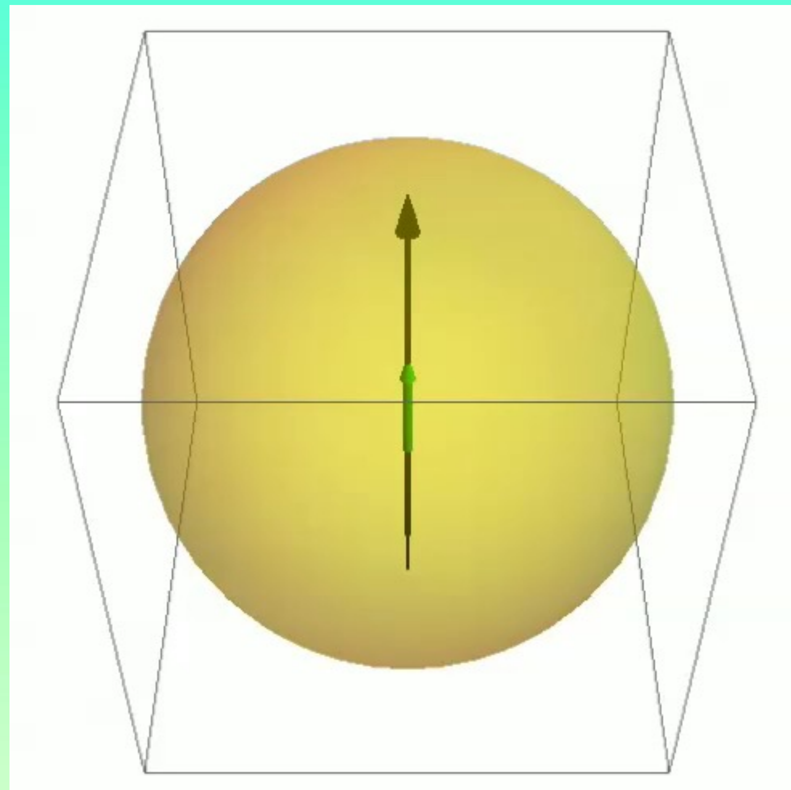
Hot spots are based on pre-selected shape patterns unrelated to magnetic field structures

Non-Dipolar Fields
Offset Dipole + Offset Quadrupole

Static Vacuum Field (SVF)



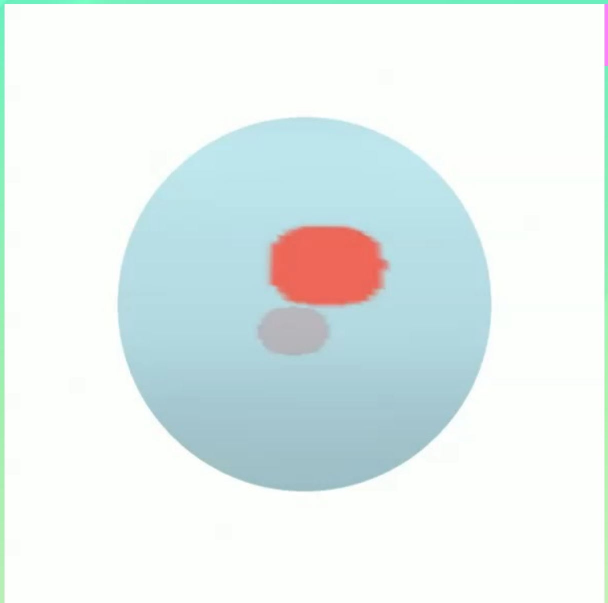
11 parameters



Non-Dipolar Fields

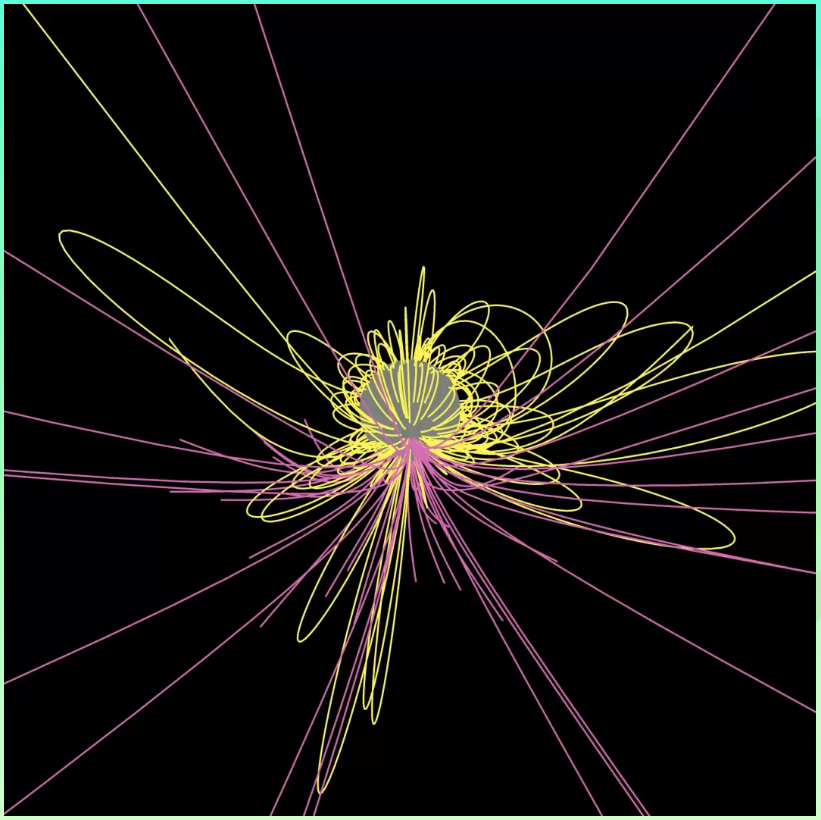
Offset Dipole + Offset Quadrupole

SVF

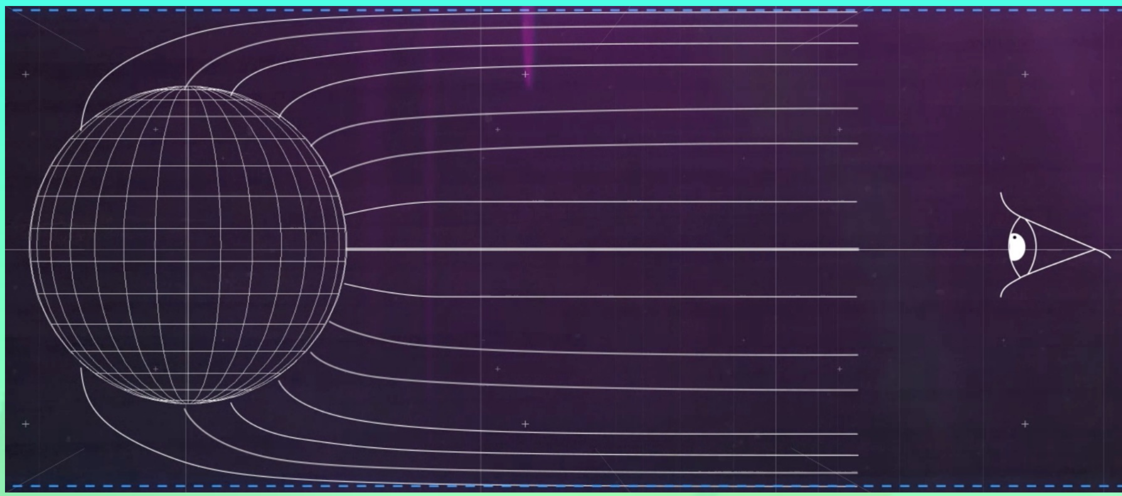


11 parameters

Central dipole ⟶ offset dipole + offset quadrupole



GR Ray Tracing



NASA video

We use Kerr metric,
but Schwarzschild metric would be adequate
for PSR J0030 spin rate

$N = 2.5\times10^6$ GIKS CODE

Mathematica, C++, Parallel, Pleiades NASA supercluster

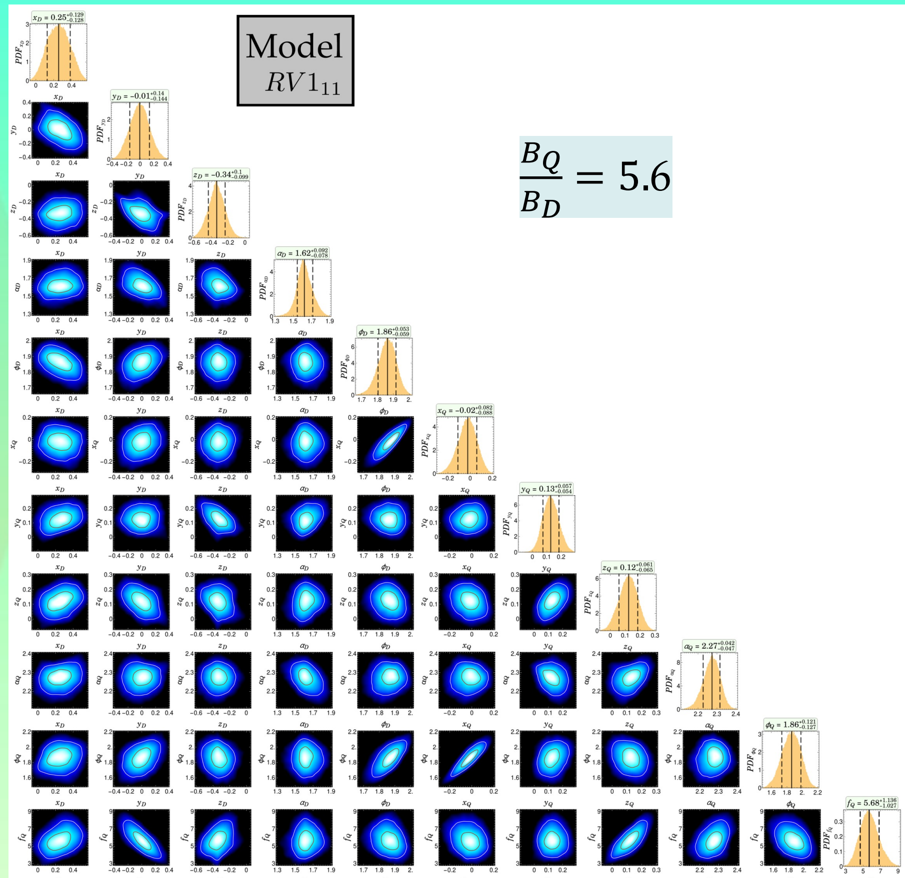
Assumptions

$M_\star, r_\star, \zeta$ from Miller et al. 2019 and Riley et al. 2019

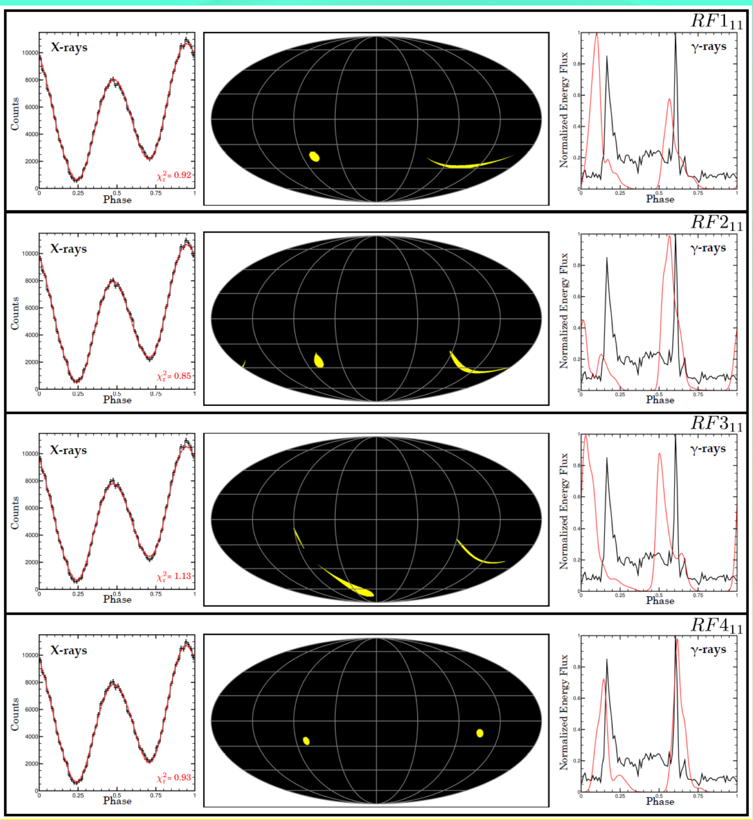
Kalapotharakos et al. 2021, see Psaltis & Johannsen 2012, Lechien et al. 2024 (in prep.)

MCMC exploration

We developed a parallel **MCMC code** implementing the *stretch move* (Goodman & Weare 2010, Foreman-Mackey et al. 2013) ensemble method.

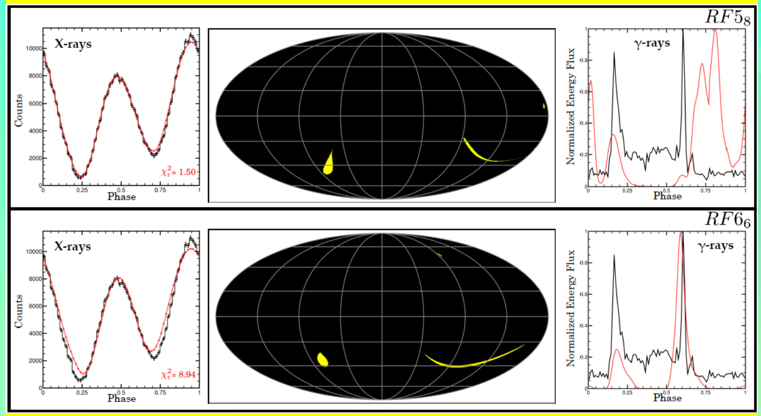


MCMC
Realistic FF/Dissipative Simulations



11 parameters

8 parameters



6 parameters

Fermi data

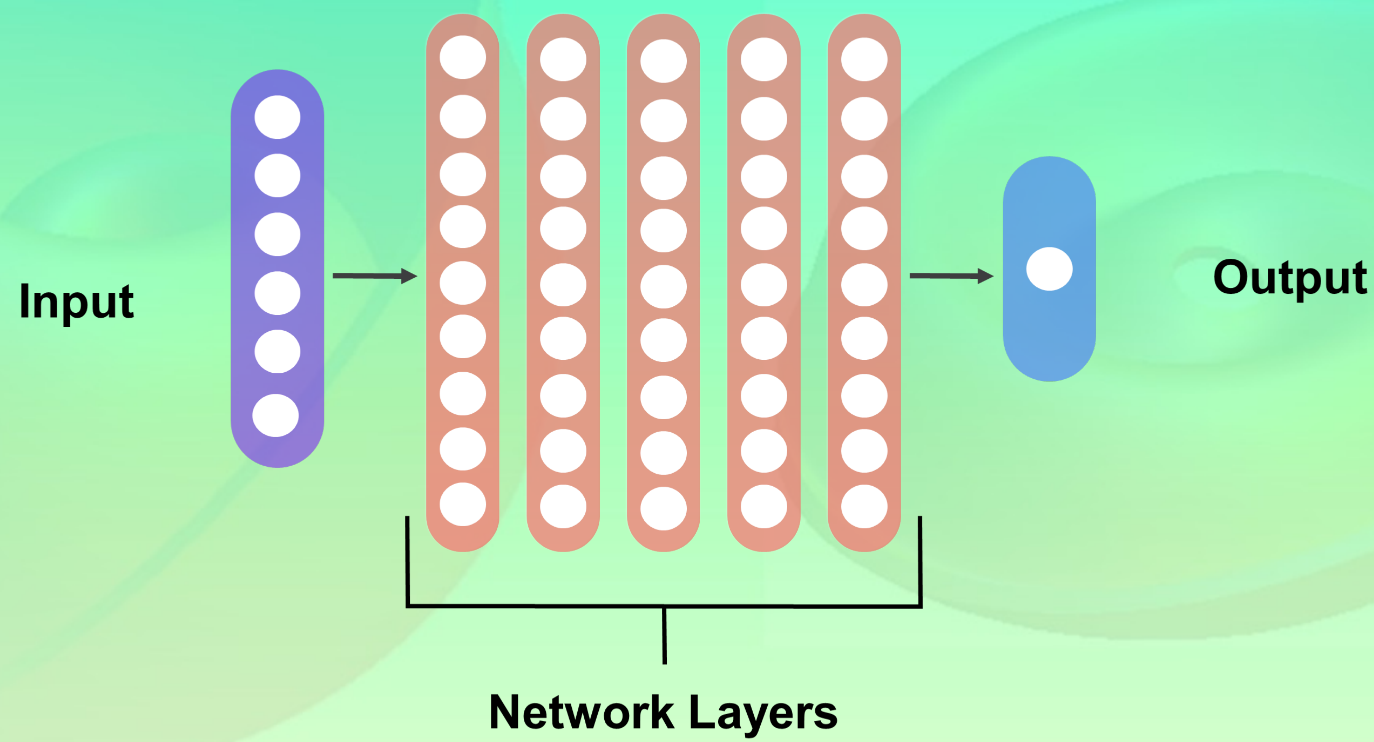
Model

FF, Riley et al., 2019 fixed parameters

Kalapotharakos et al. (2021)

Revolutionizing the parameter inference through Machine Learning Approaches

Neural Networks

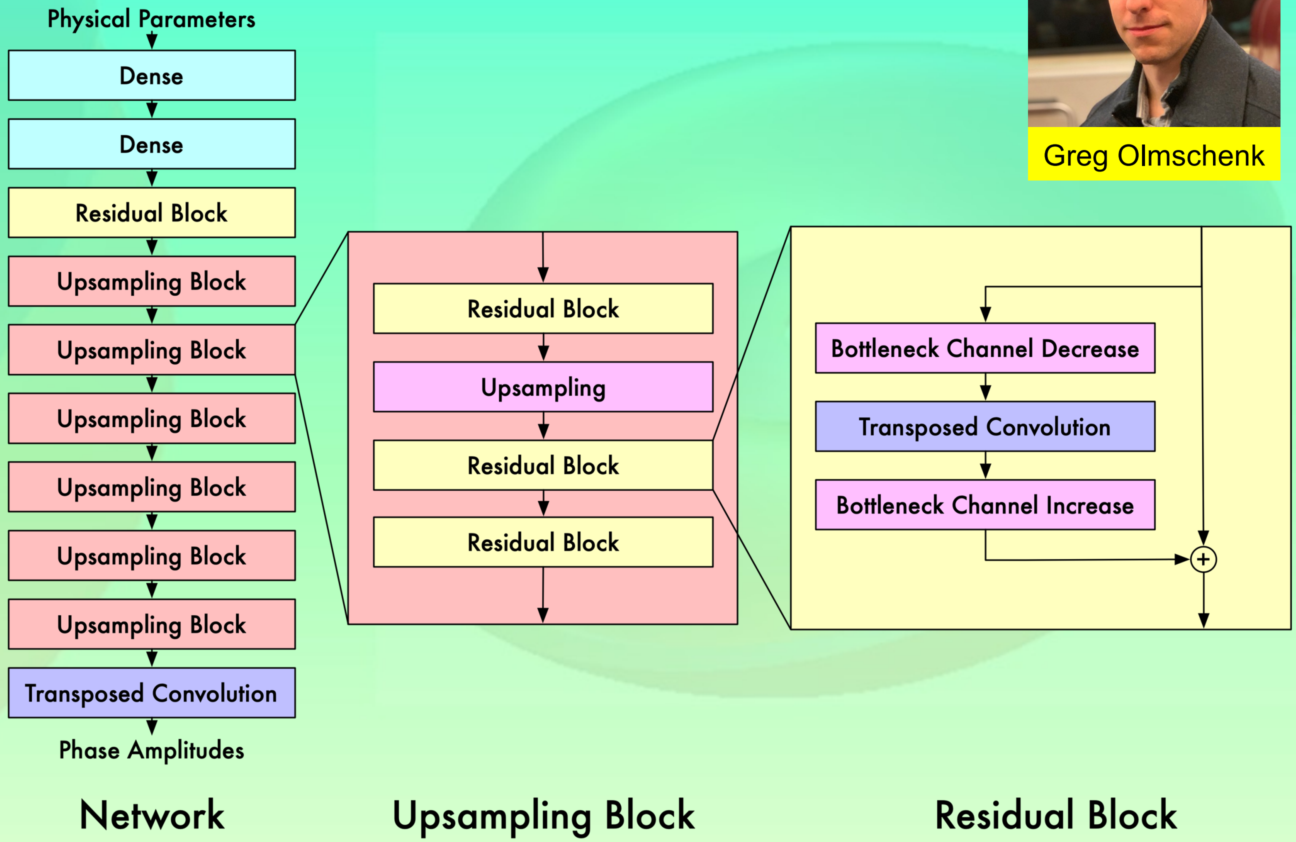


Revolutionizing the parameter inference through Machine Learning Approaches

Neural Networks

Greg Olmschenk

Our network is a ResNet-like structure but transposed and 1D. The main trainable layers are transposed convolutions which connect adjacent features in the phase dimension. Residual skip connections help enable training with a deep number of layers.

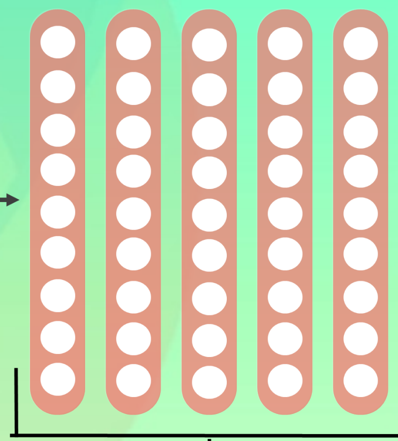


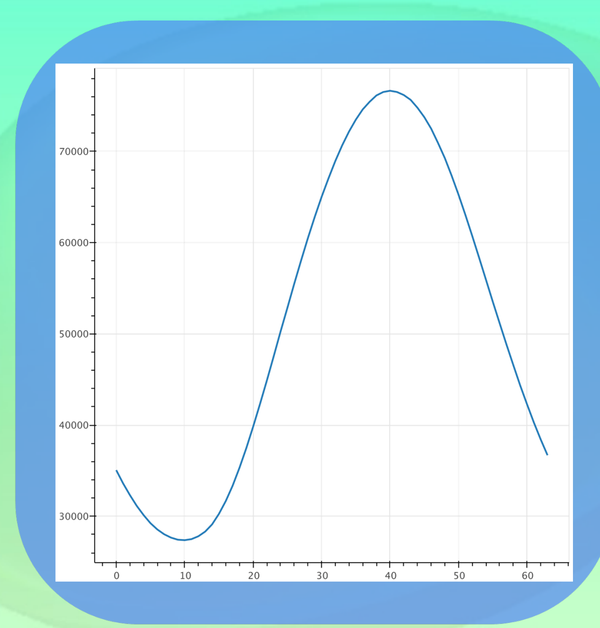
Revolutionizing the parameter inference through Machine Learning Approaches

Neural Networks

Emily Broadbent

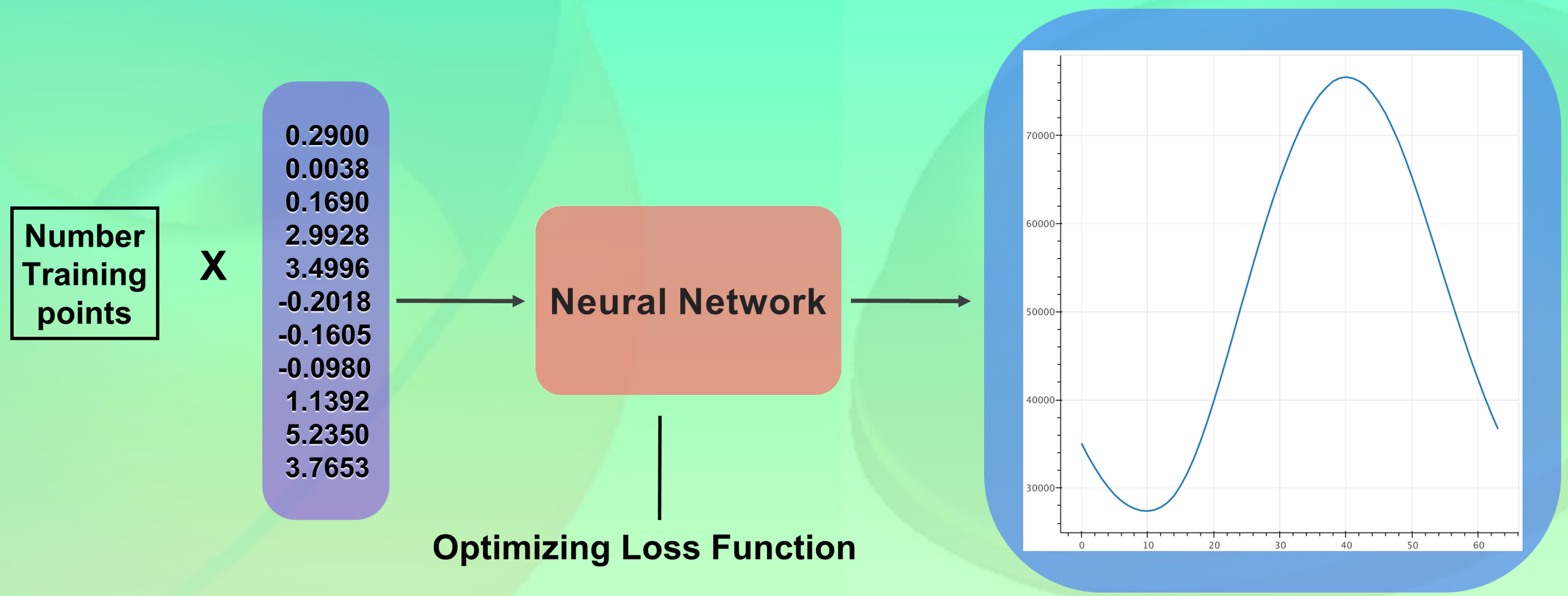
x_D	0.2900
y_D	0.0038
z_D	0.1690
a_D	2.9928
φ_D	3.4996
x_Q	-0.2018
y_Q	-0.1605
z_Q	-0.0980
a_Q	1.1392
φ_Q	5.2350
Q/D	3.7653



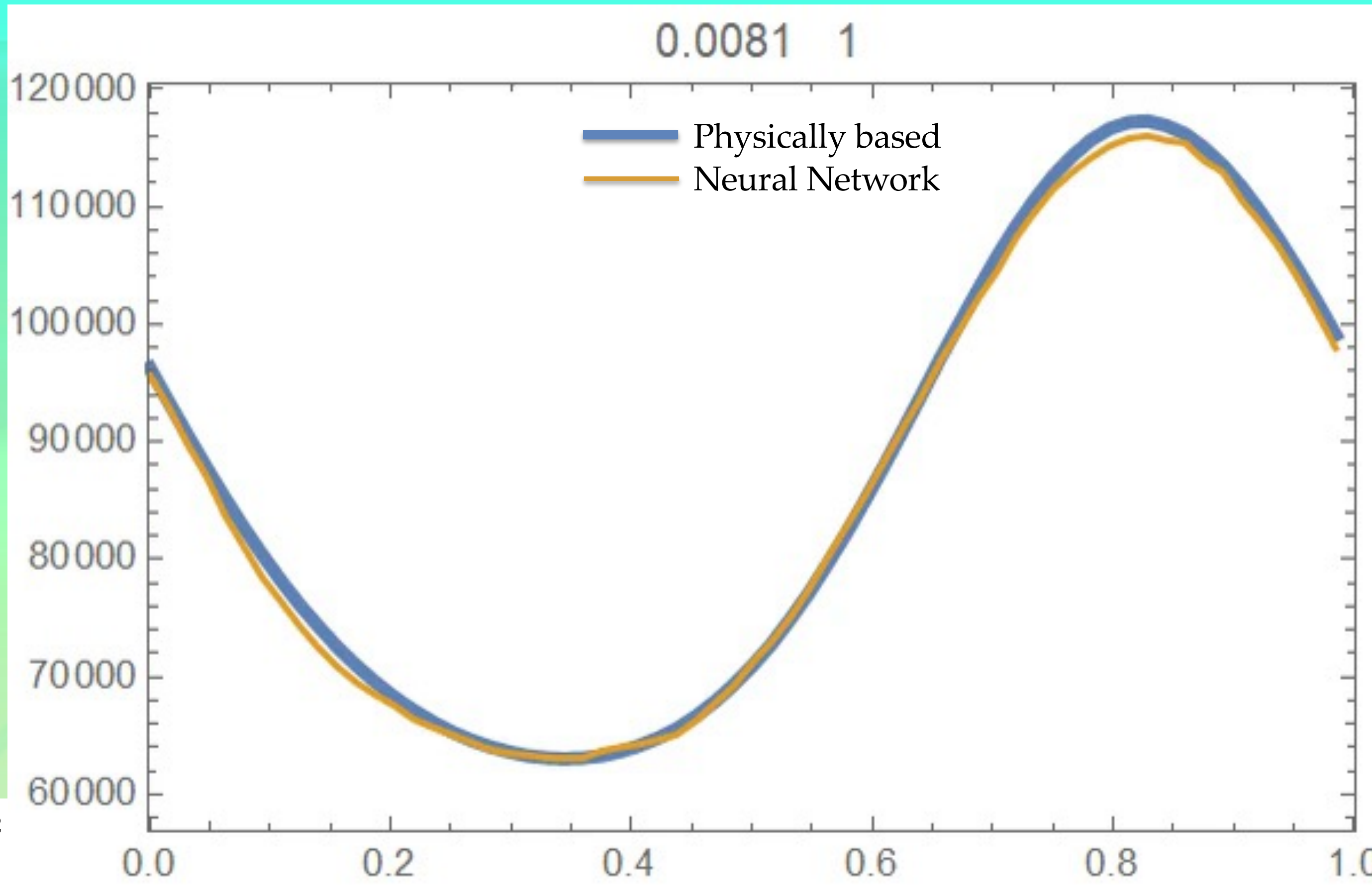


Revolutionizing the parameter inference through Machine Learning Approaches

Neural Networks



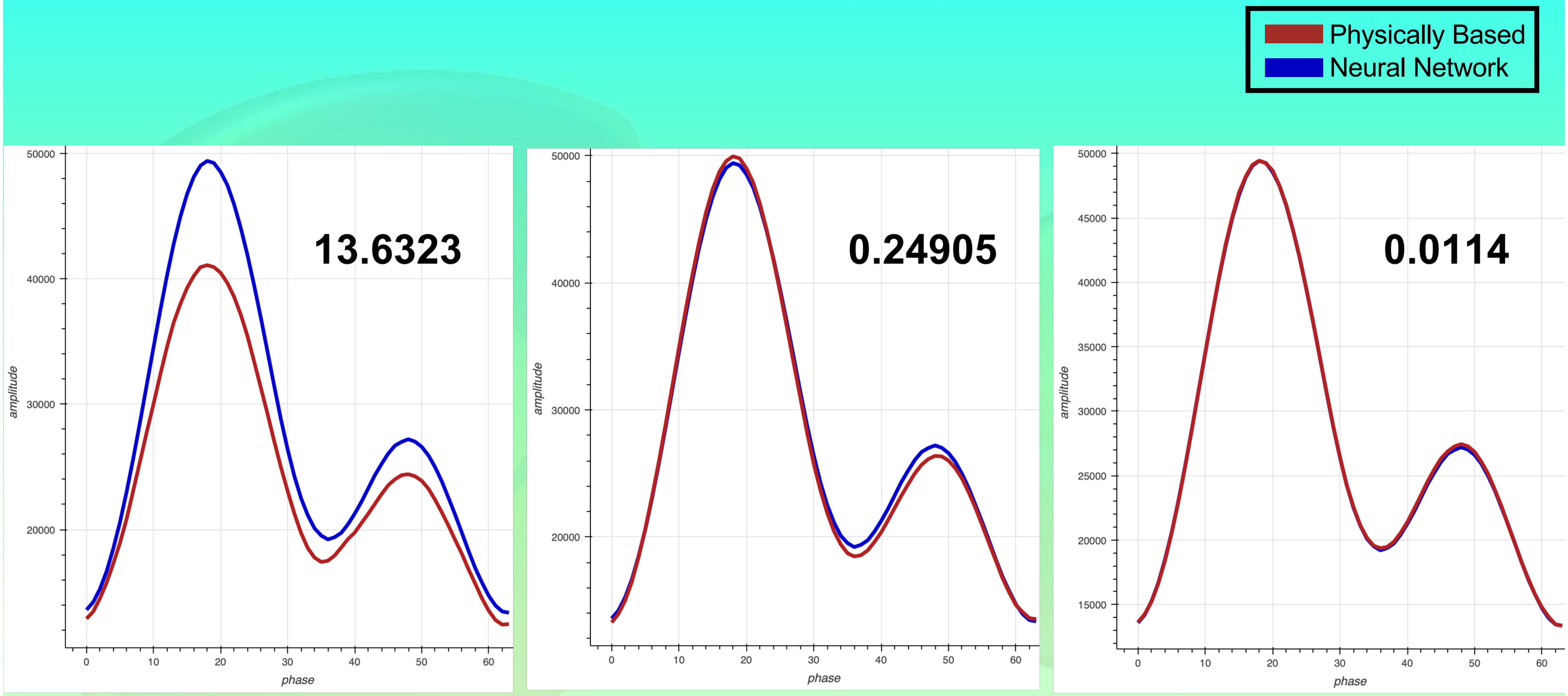
Neural Networks



$$QI = \frac{1}{Median_{PB}^2} \sum_{bins} \left(Y_{PB_i} - Y_{ML_i}\right)^2$$

Olmschenk et al. 2024 (in prep)

Neural Networks

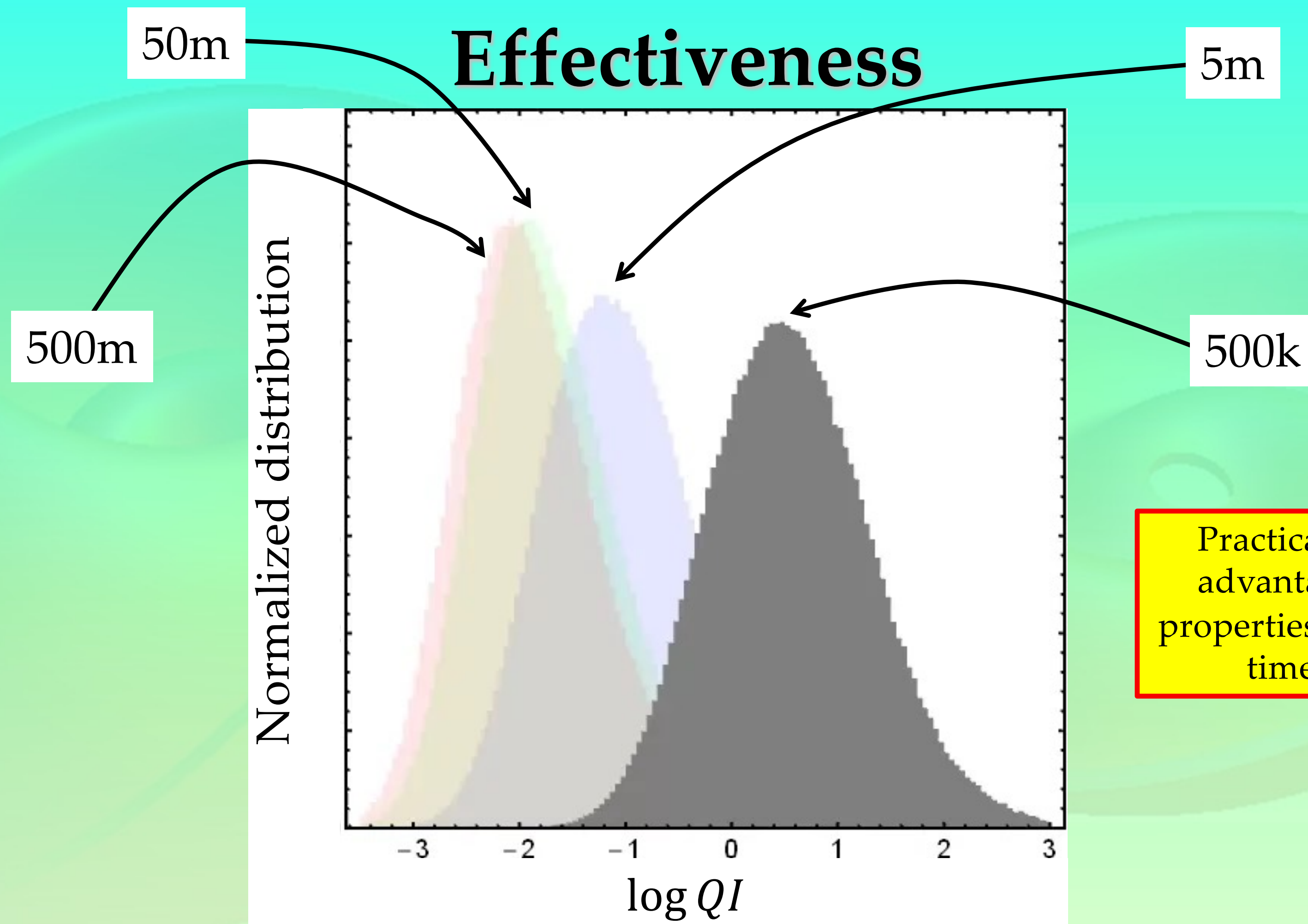


500k Training Data

5M Training Data

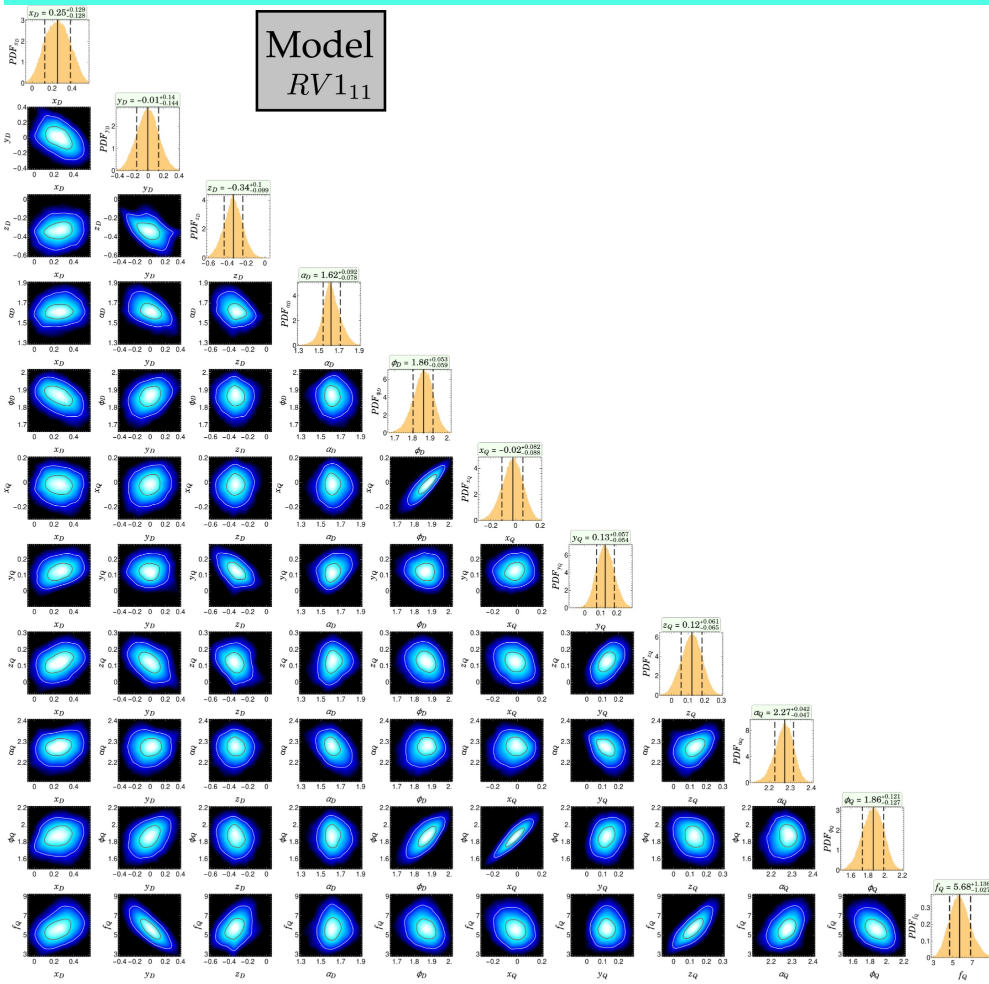
50M Training Data

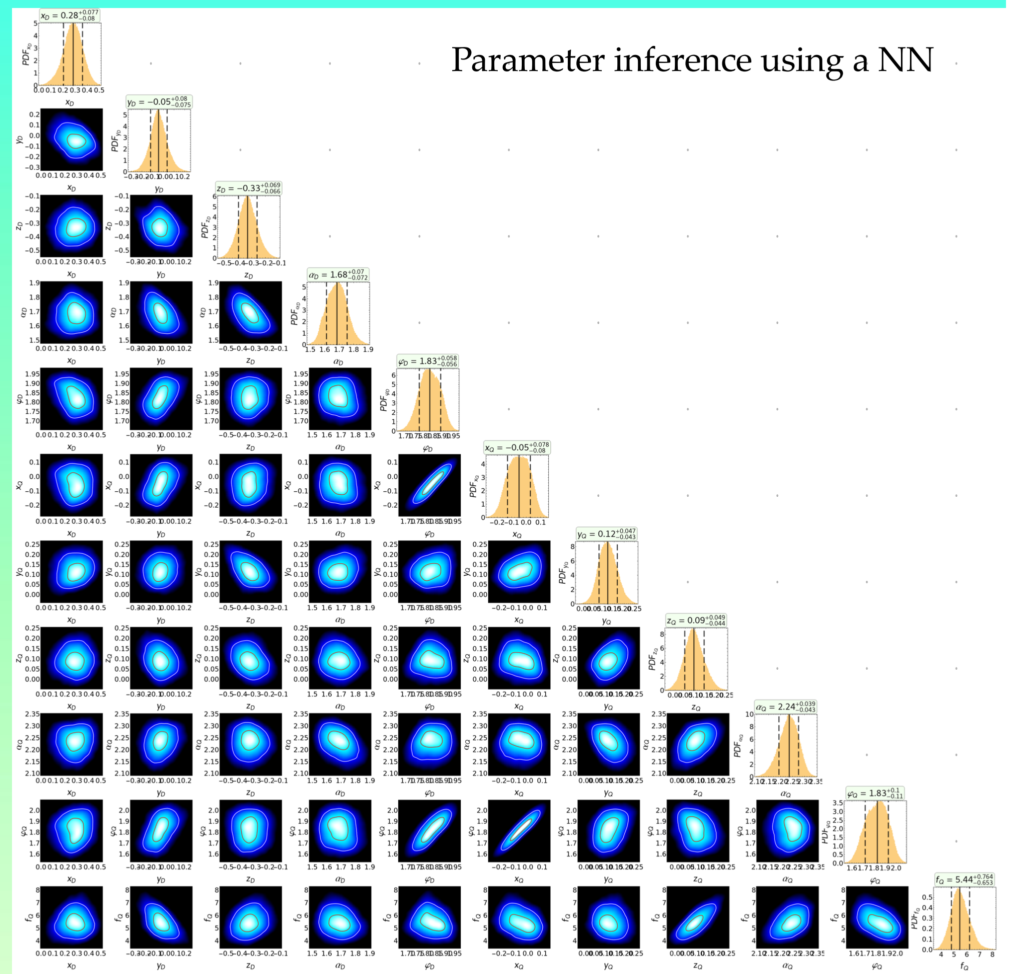
Neural Networks Effectiveness



Practically, taking the advantage of rotation properties, we calculate 64 times less data.

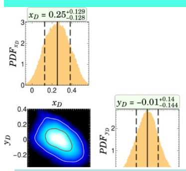
Neural Networks



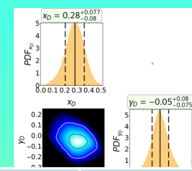


Parameter inference using a NN

Neural Networks

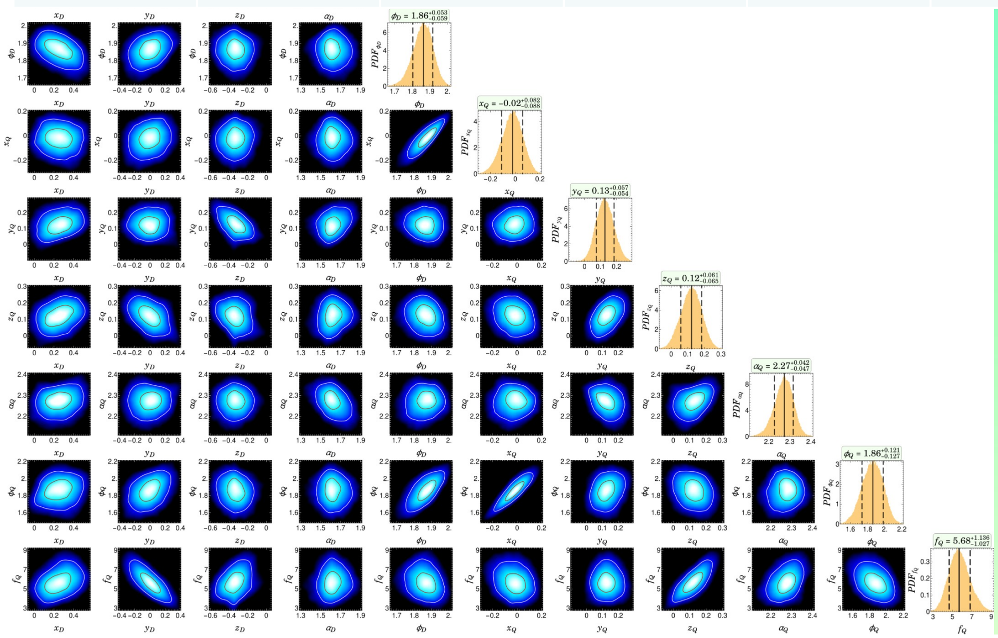


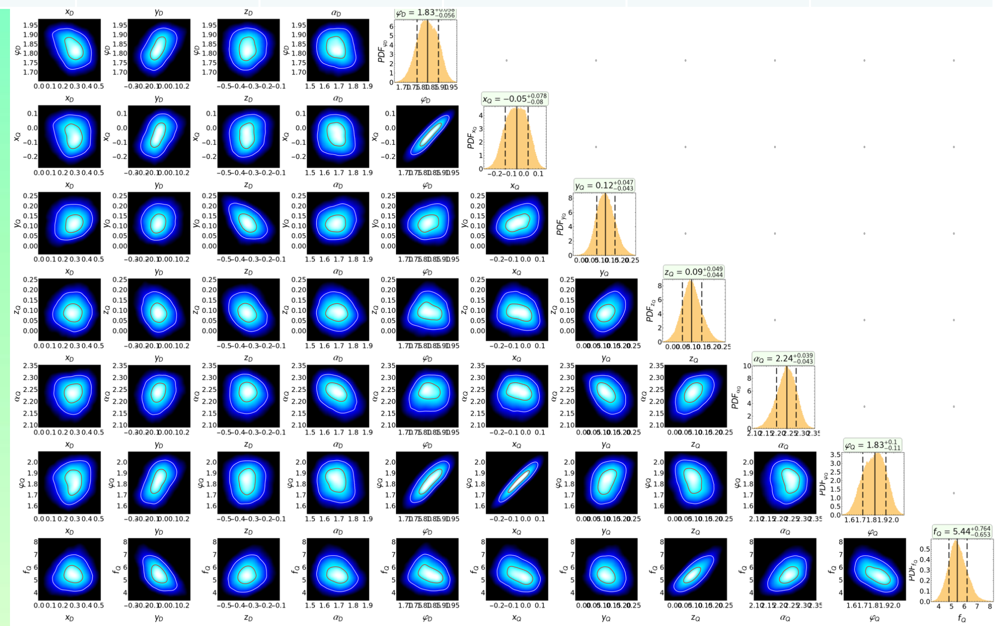
Model
$RV1_{11}$



Parameter inference using a NN

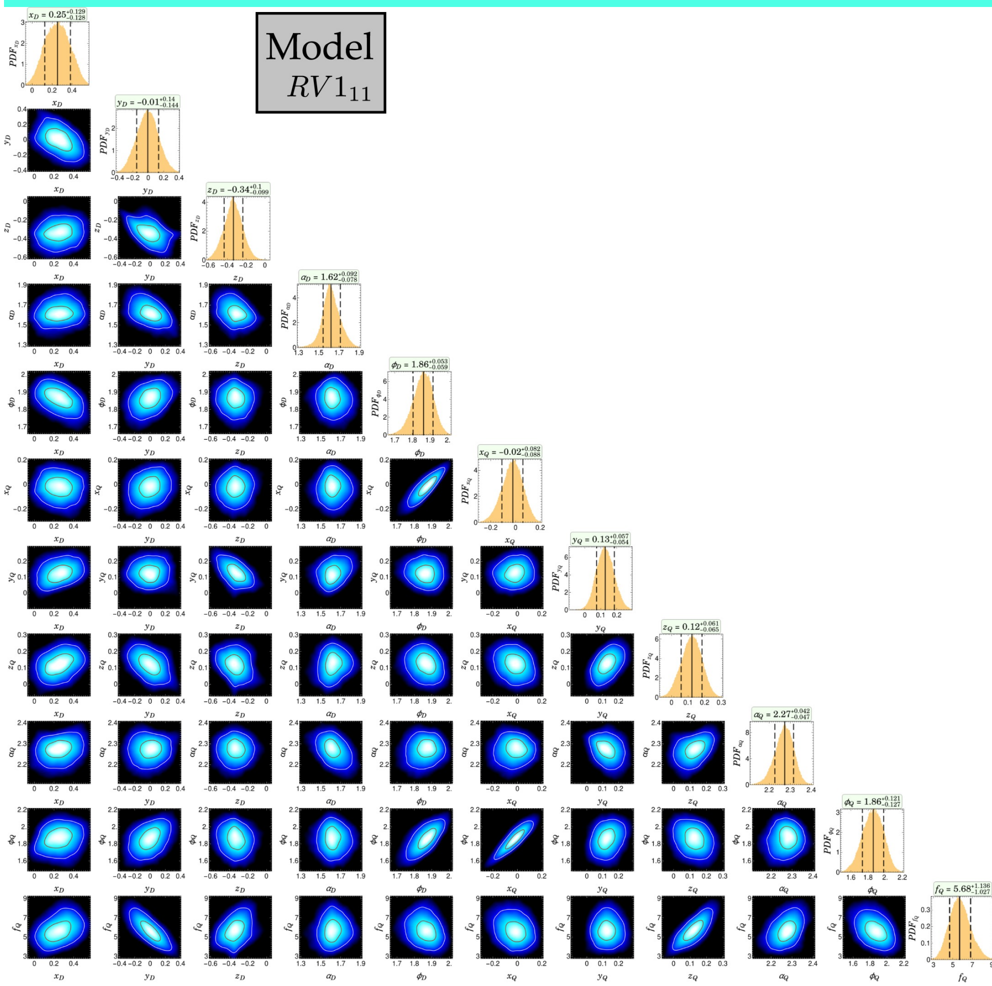
x_D	y_D	z_D	a_D	φ_D	x_Q	y_Q	z_Q	a_Q	φ_Q	Q/D
$0.25^{+0.129}_{-0.128}$	$-0.01^{+0.140}_{-0.144}$	$-0.34^{+0.1}_{-0.099}$	$1.62^{+0.092}_{-0.078}$	$1.86^{+0.053}_{-0.059}$	$-0.02^{+0.082}_{-0.088}$	$0.13^{+0.057}_{-0.054}$	$0.12^{+0.061}_{-0.065}$	$2.27^{+0.042}_{-0.047}$	$1.86^{+0.121}_{-0.127}$	$5.68^{+1.36}_{-1.027}$
$0.28^{+0.077}_{-0.080}$	$-0.05^{+0.080}_{-0.075}$	$-0.33^{+0.069}_{-0.066}$	$1.68^{+0.070}_{-0.072}$	$1.83^{+0.058}_{-0.056}$	$-0.05^{+0.078}_{-0.080}$	$0.12^{+0.047}_{-0.043}$	$0.09^{+0.049}_{-0.044}$	$2.24^{+0.039}_{-0.043}$	$1.83^{+0.100}_{-0.110}$	$5.44^{+0.764}_{-0.653}$





Neural Networks

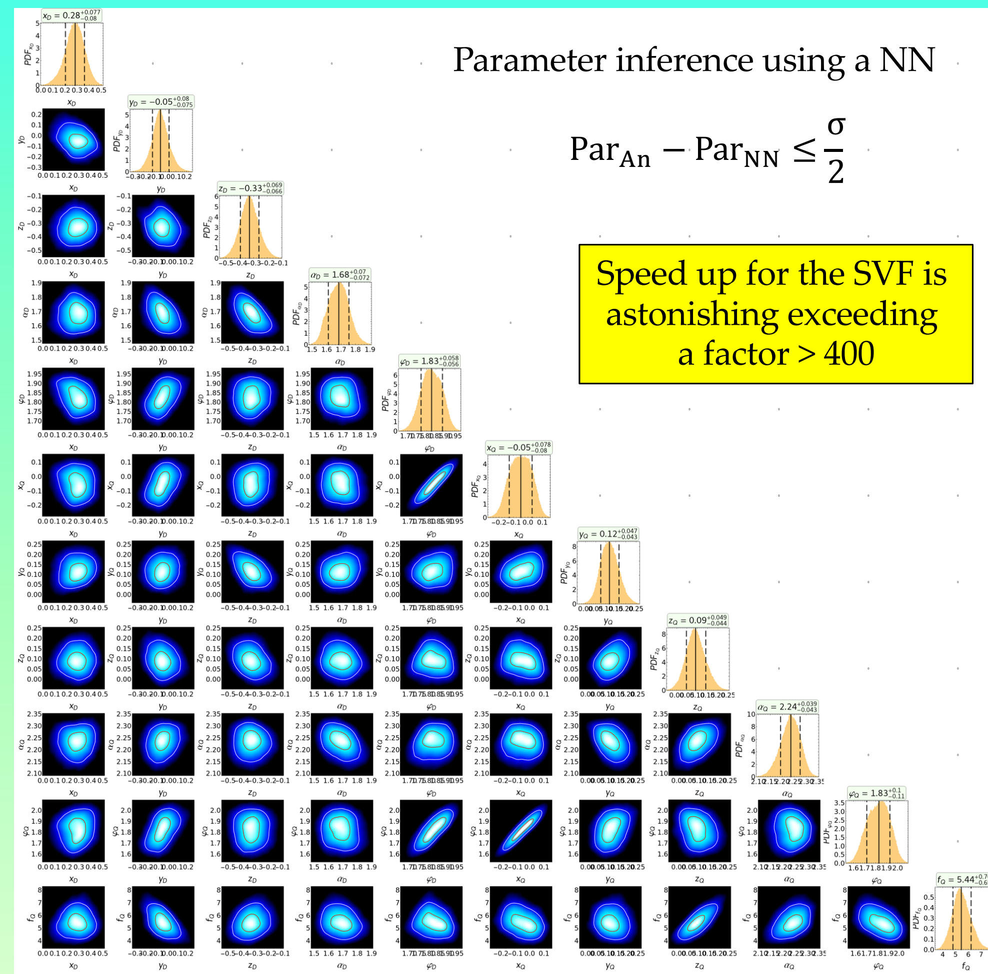
Model
$RV1_{11}$



Parameter inference using a NN

$$\mathrm{Par}_{\mathrm{An}} - \mathrm{Par}_{\mathrm{NN}} \leq \frac{\sigma}{2}$$

Speed up for the SVF is astonishing exceeding a factor > 400



Statistical Tests

F_N = NN distribution
F_P = Physical distribution

Kolmogorov Smirnov	$KS = \sup_{V} \lvert F_N - F_P \rvert$
Wasserstein Difference	$W = \int_V \lvert F_N - F_P \rvert \, dx$
Jensen-Shannon Divergence	$JSD = \frac{1}{2} KLD(F_N \lvert\lvert F_P) + \frac{1}{2} KLD(F_P \lvert\lvert F_N)$

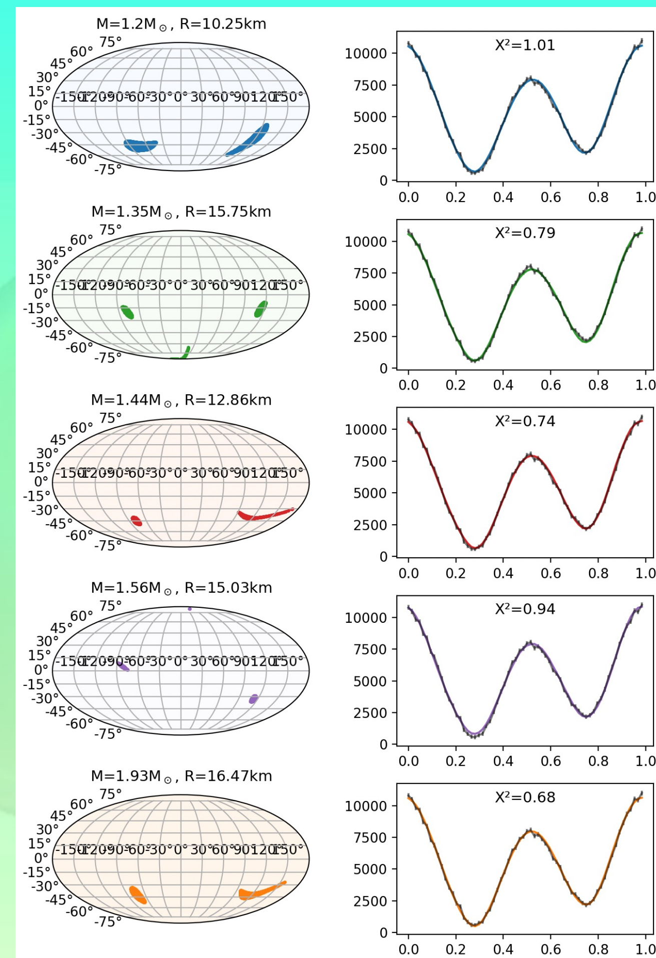
Kullback-Leibler Divergence $KLD = \sum_V F_N \log \frac{F_N}{F_P}$

Neural Networks

Thibault Lechien

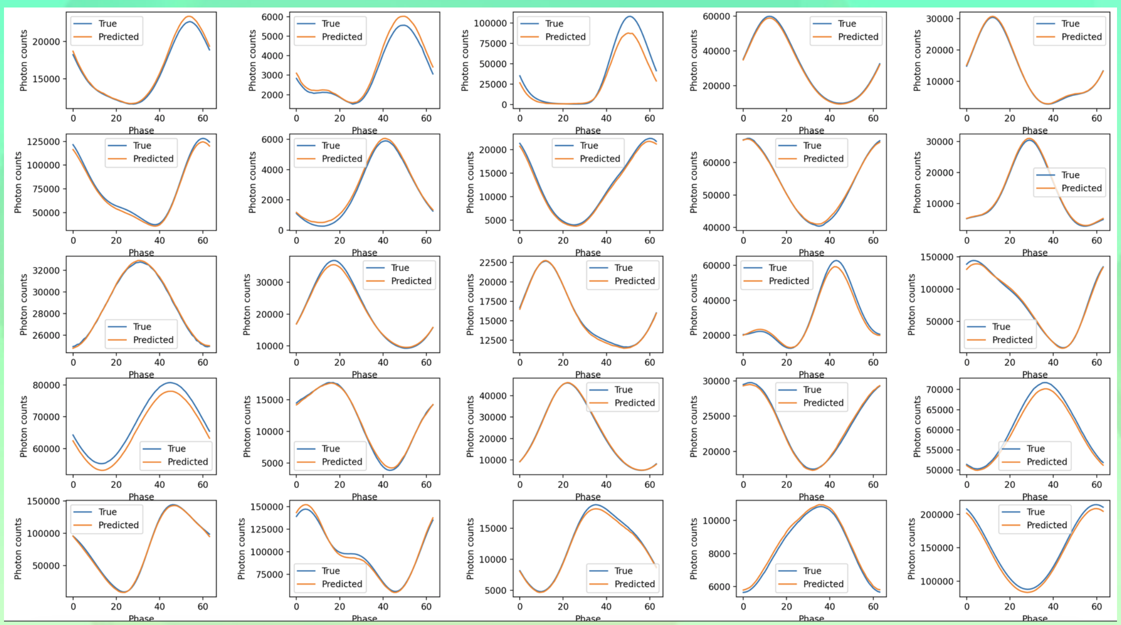
We have already expanded the parameter space considering free stellar mass and radii using again NNs

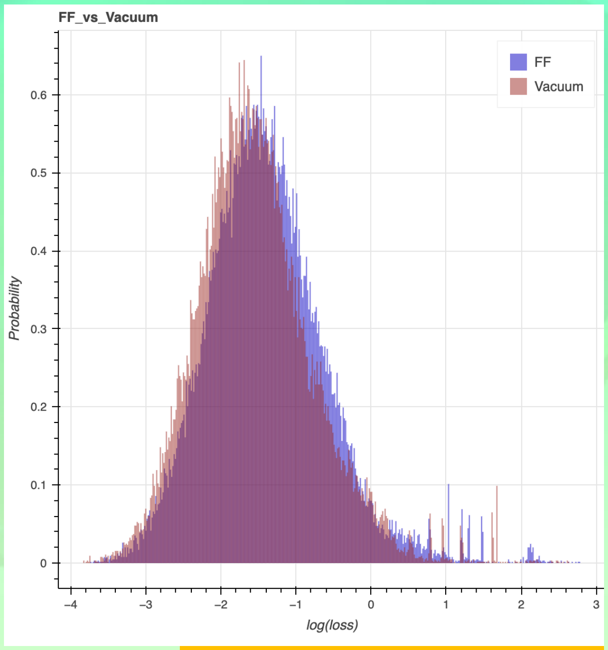
- Degenerate solutions that fit the NICER X-ray light curve
- Need a way to further constrain the parameters



Neural Networks
Vacuum to Force Free

Training NNs from scratch using FF light curves is too expensive.
However, using the NN trained on a large dataset of vacuum model data and fine-tuning it with a smaller dataset of force-free model data shows promising results.



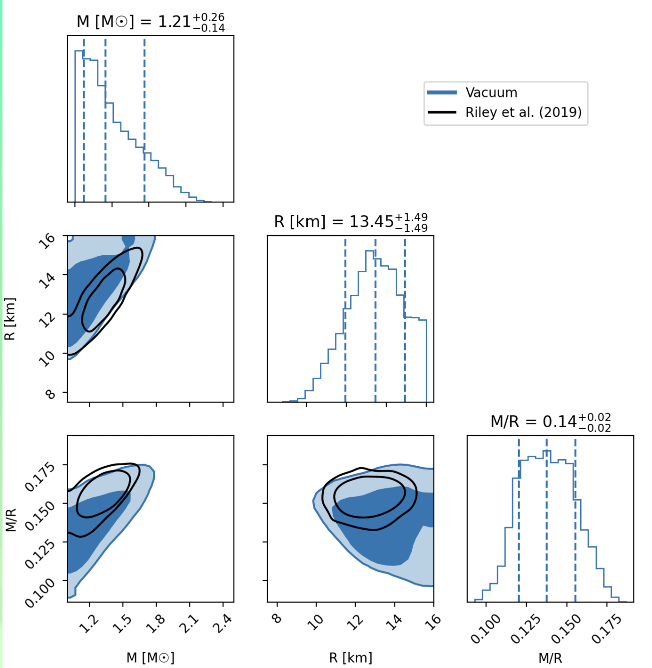


Lechien et al. (in prep.)

Neural Networks
Vacuum to Force Free

Training NNs from scratch using FF light curves is too expensive.
However, using the NN trained on a large dataset of vacuum model data and fine-tuning it with a smaller dataset of force-free model data shows promising results.

Vacuum NN

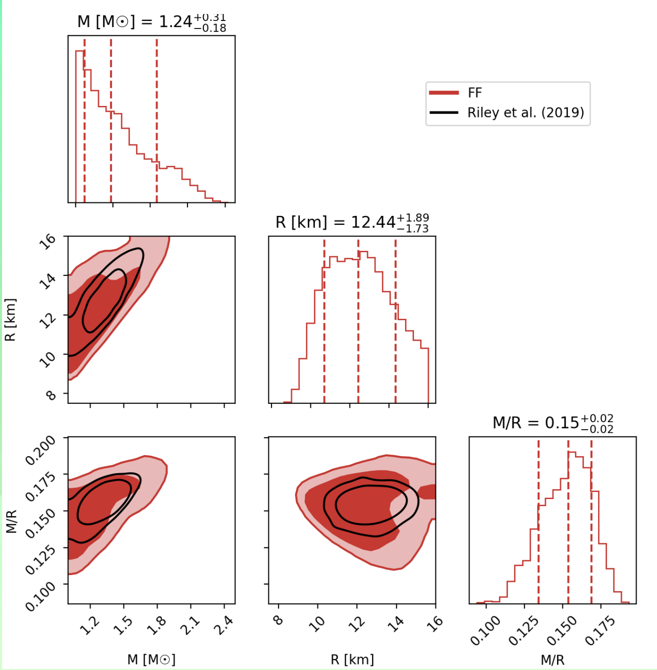


Preliminary results

Ultranest runs

Lechien et al. (in prep.)

FF NN



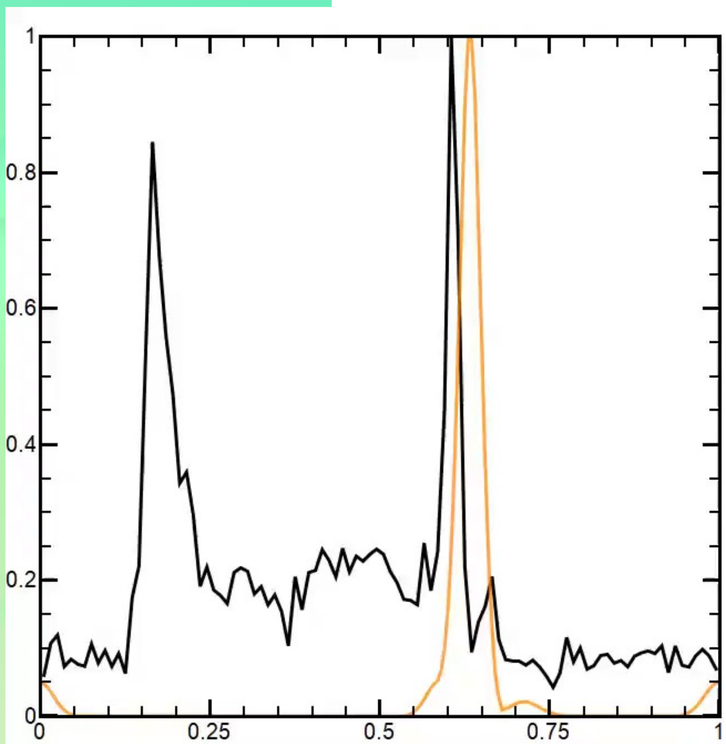
Neural Networks

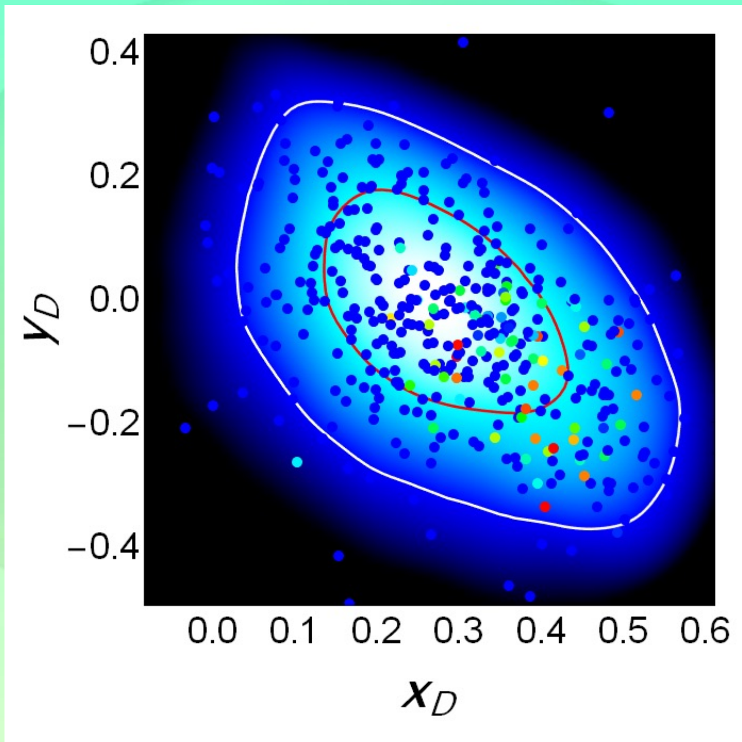
Requesting the field structure to fit the Fermi γ-ray light curve eliminates considerably the number of viable solutions

A sample of 420 cases from the posterior distribution

A metric that considers the phase difference of the peaks and the peak amplitudes

We have already started producing training datasets of model gamma-ray light curves to train NNs that emulate the model gamma-ray light curves.





Preliminary results

Neural Networks
Vacuum Retarded Multipolar Components

So far, vacuum means static vacuum with offsets.
We have started implementing vacuum retarded multipolar components, which are also analytic.

$$\boldsymbol{B}(r,\theta,\phi,t) = \sum_{l=1}^{\infty}\sum_{m=-l}^{l}\left[\boldsymbol{B_r}(r,\theta,\phi,t) + \boldsymbol{B_\theta}(r,\theta,\phi,t) + \boldsymbol{B_\phi}(r,\theta,\phi,t)\right]$$

where,

Petri 2012, 2015

$$\boldsymbol{B_r} = -\frac{\sqrt{l(l+1)}}{r} f^B_{lm} Y_{lm}(\theta) e^{im\phi} e^{-im\Omega t}$$

$$\boldsymbol{B_\theta} = -\frac{\partial_r(r f^B_{lm})}{r\sqrt{l(l+1)}}\partial_\theta(Y_{lm}(\theta))e^{im\phi}e^{-im\Omega t} + \frac{i\mu_0 m\Omega f^D_{lm}}{\sin\theta\sqrt{l(l+1)}} im Y_{lm}(\theta)e^{im\phi}e^{-im\Omega t}$$

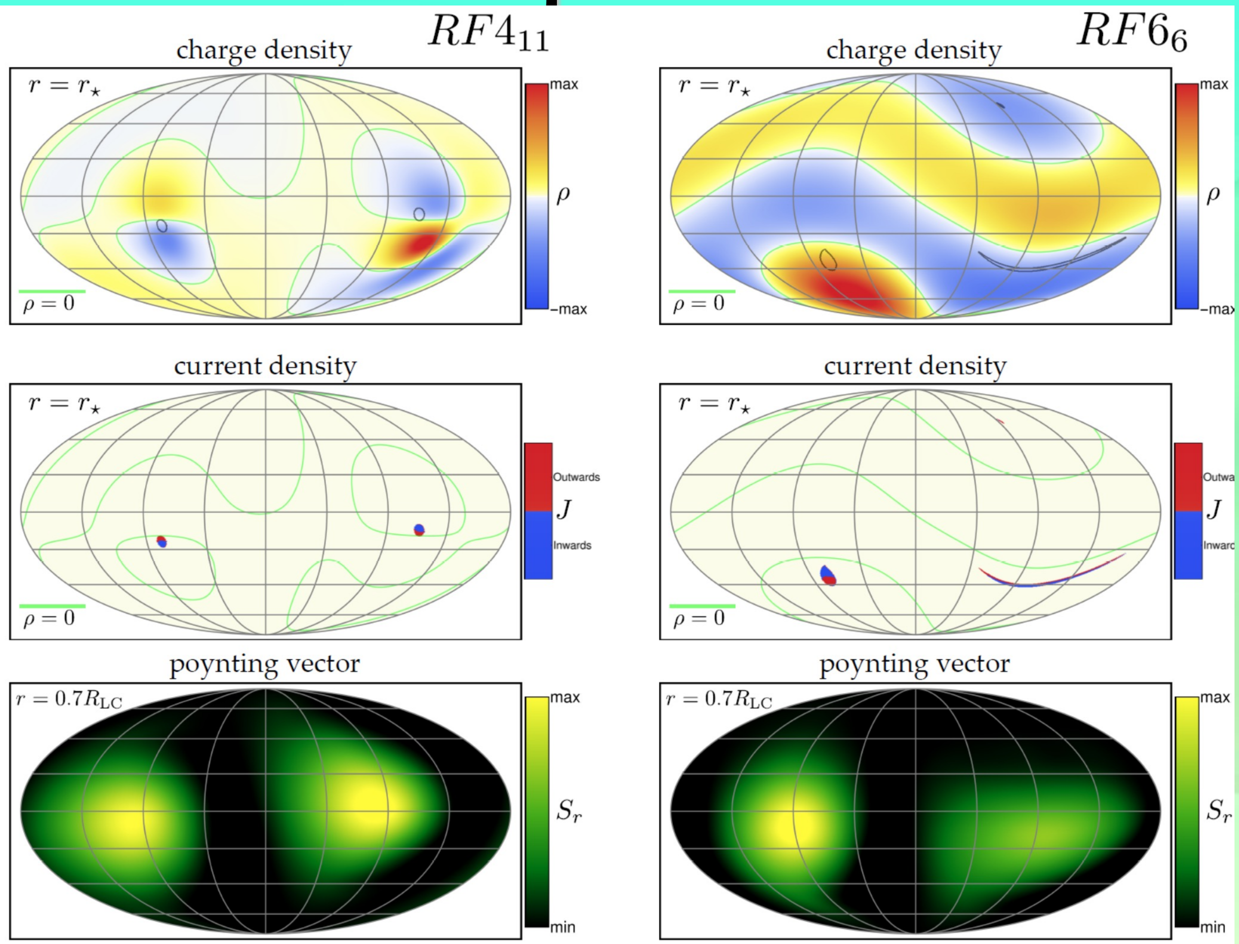
$$\boldsymbol{B_\phi} = -\frac{\partial_r(r f^B_{lm})}{r\sin\theta\sqrt{l(l+1)}} im Y_{lm}(\theta)e^{im\phi}e^{-im\Omega t} - \frac{i\mu_0 m\Omega f^D_{lm}}{\sqrt{l(l+1)}}\partial_\theta(Y_{lm}(\theta))e^{im\phi}e^{-im\Omega t}$$

Kundu et al. (in prep.)

The advantage with the vacuum retarded fields is that, besides their analytic nature, they are closer to the FF field structures.

So, the vacuum retarded fields will provide advanced NN initial conditions for the fine-tuning training of the FF NNs

Hot Spots
Polar Caps vs. Current Structures



- Temperature distribution
- Pair production

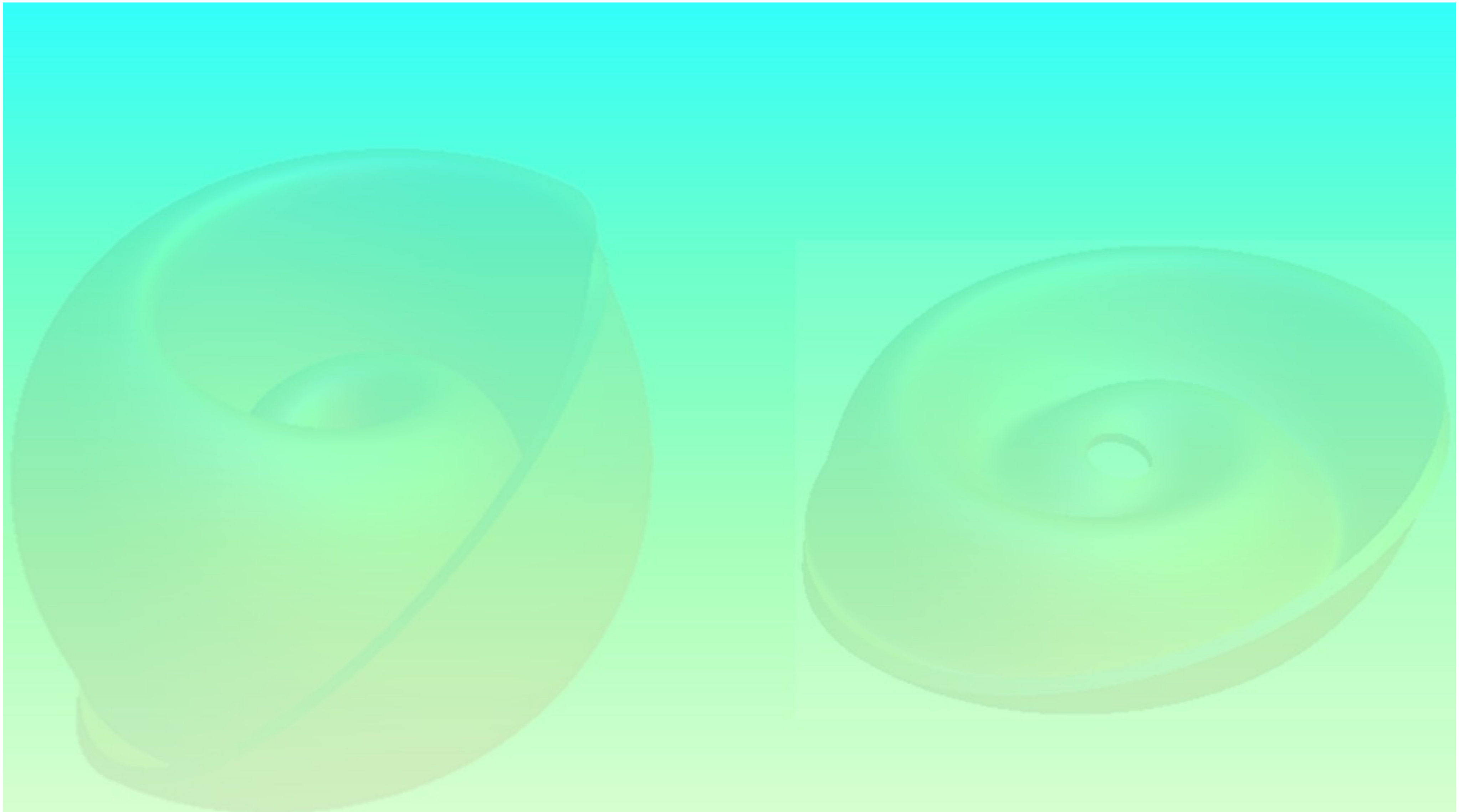
Rocket effect

Summary, Future, Importance

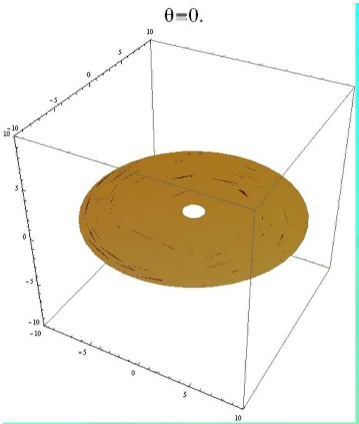
- **NICER thermal X-ray light curves of MSPs can provide meaningful constraints not only on masses & radii, i.e., EoS, but also field structures.**
- **A proper and self-consistent treatment requires the consideration of hot spots in agreement with realistic field structures.**
- **Even though the existing data incorporate the necessary information, the actual physical modeling is computationally demanding making the study unfeasible.**
- **Machine learning techniques revolutionize the parameter inference making the global study feasible.**
- **Incorporating the Fermi-LAT data and gamma-ray models into this study (doable with ML) will provide more robust and stricter constraints.**

Next steps: Incorporate energy dependent light curves, deviations from sphericity, explore more efficient NN structures & loss functions, explore normalizing flows & generative adversarial networks (GANs), hot-spots not only on polar caps but also on current structures, magnetars

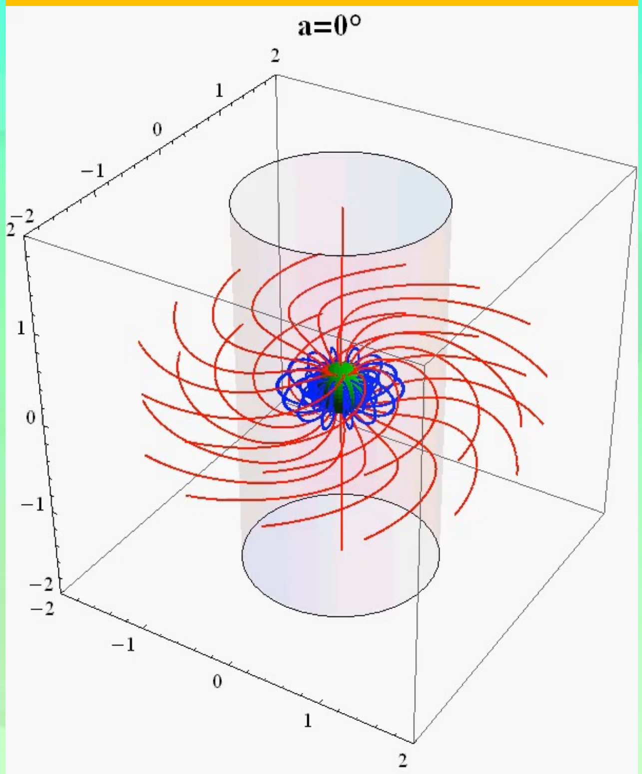
Broader scientific impact: Radio emission region, pair production efficiency in MSPs, generation and evolution of magnetic fields



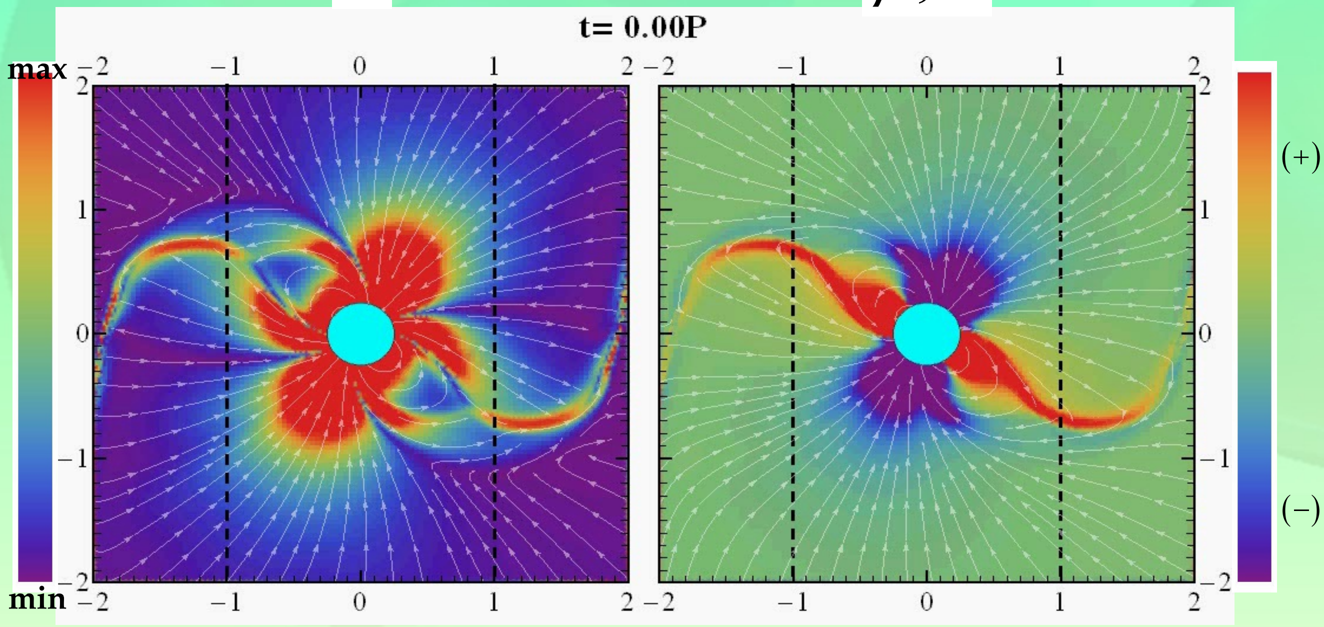
FFE Solutions



Kalapotharakos & Contopoulos 2009; Kalapotharakos et al. 2012



J $\alpha = 45°$ ρ, B



GR Ray Tracing

We developed a **GR-code** (GIKS) that follows the photon trajectories in the full **Kerr-metric**

Kalapotharakos et al. 2021, see Psaltis & Johannsen 2012

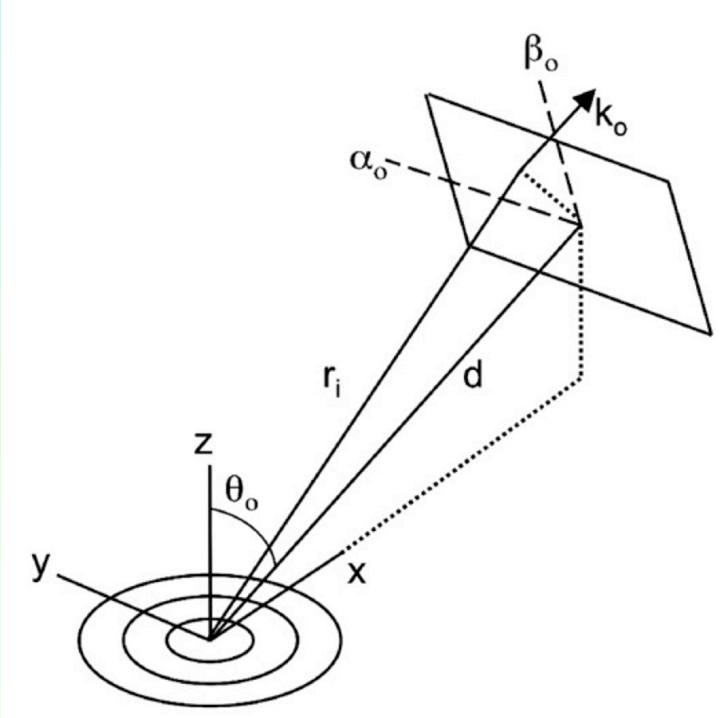
We use Kerr metric,
but Schwarzschild metric would be adequate
for PSR J0030 spin rate

$N = 2.5\times10^6$

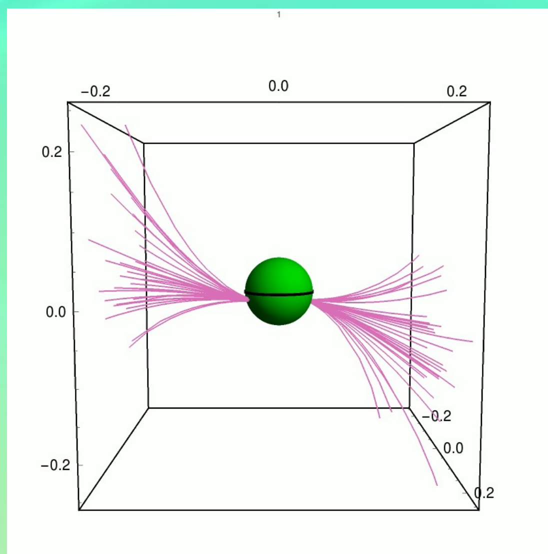
Mathematica, C++, Parallel, Pleiades NASA supercluster

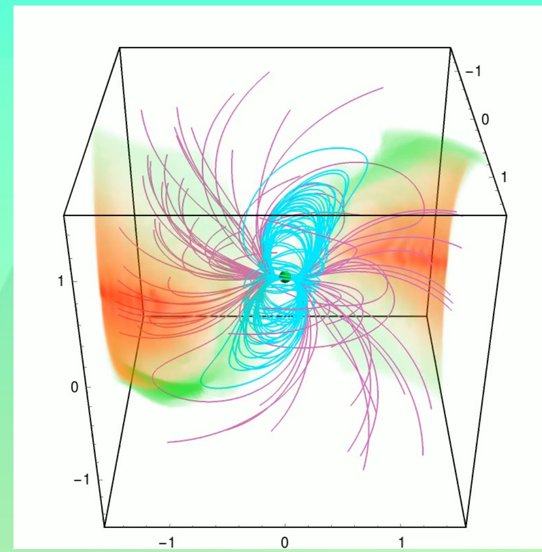
Assumptions

$M_\star, r_\star, \zeta$ from Miller et al. 2019 and Riley et al. 2019



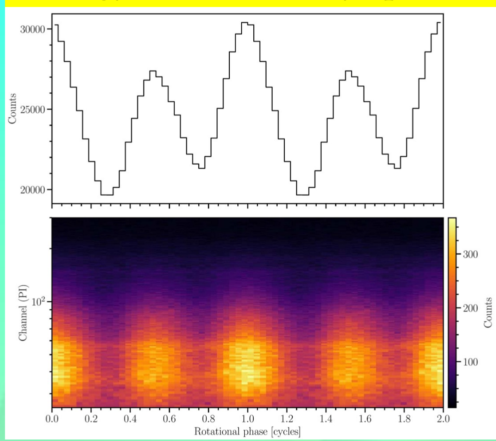
Lechien et al. 2024 (in prep.)


−0.2
0.0
0.2
0.2
0.0
−0.2
−0.2
0.0
0.2

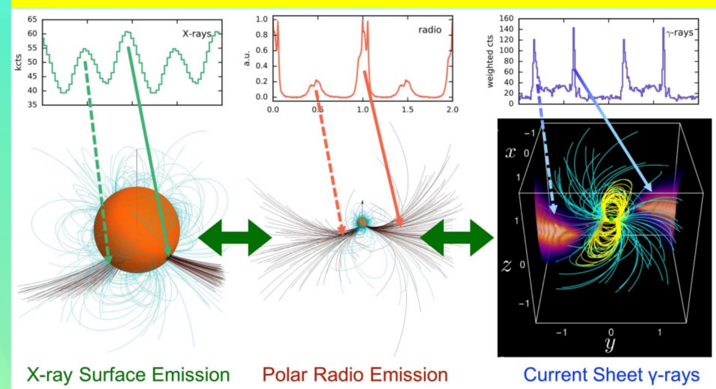

−1
0
1
1
0
−1
−1
0
1

Future

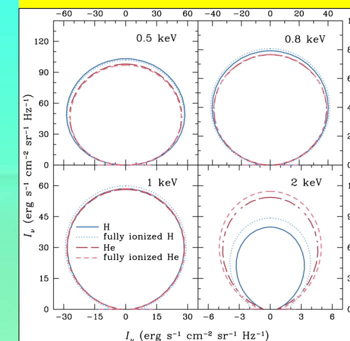
Energy resolved X-ray spectra



X-ray, γ-ray, & radio data
Advanced vacuum (**Petri**) & FF fields

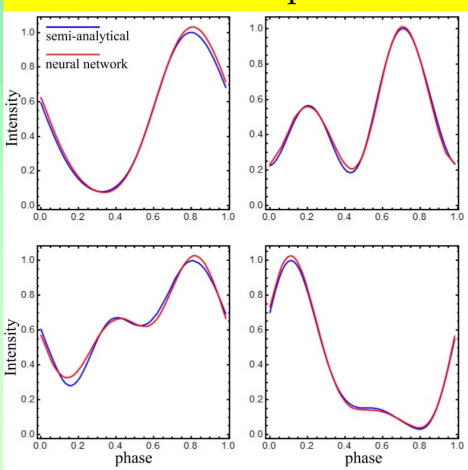


Advanced realistic atmosphere models



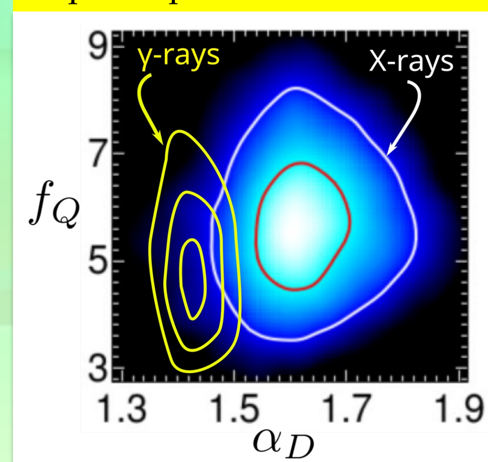
Magnetars

ML techniques

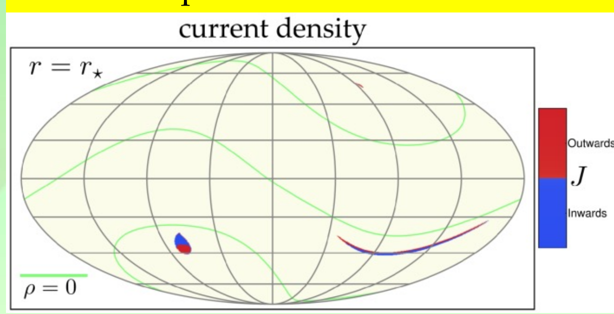


Self-consistent M_*, r_*, and B-field determination will shift the median values and reduce uncertainties

Superior parameter inferences



Temperature distributions



Atmosphere model

The reconstruction of X-ray LCs (i.e., the intensity at each phase) requires the incorporation of the **Doppler boosting** and an **atmosphere model.**

Miller et al. 2019 and Riley et al. 2019 used the same atmosphere model (i.e., pure H^+) even though they used slightly different energy channels.

Atmosphere Model
$\mathbf{I} = \mathcal{F}(\vartheta_z, E)$
$\mathbf{I} \propto \cos^n(\vartheta_z)$

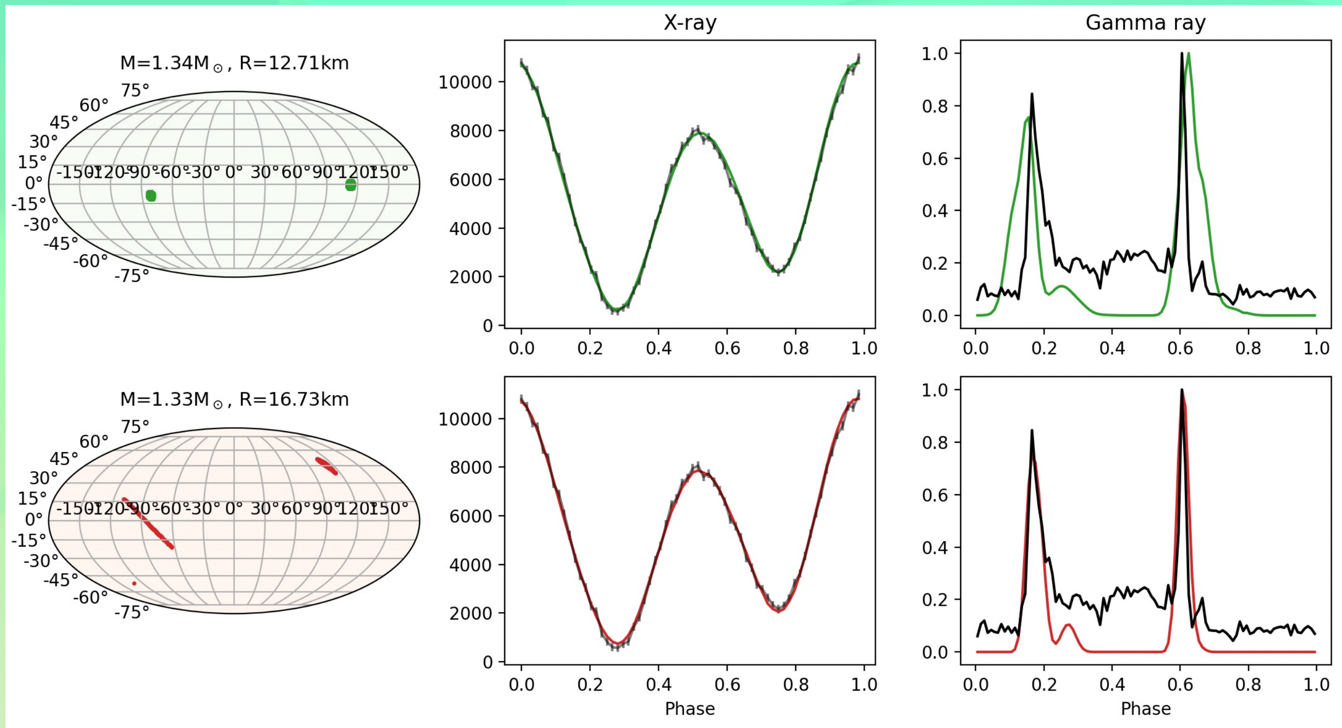
Assuming the Riley et al. HSs, we were able to reproduce the X-ray LC for $n \approx 1$

Assuming the Miller et al. HSs, we were able to reproduce the X-ray LC for $n \approx 0.65$

Neural Networks

Requesting the field structure to fit the Fermi γ-ray light curve eliminates thousands of solutions down to ≈ 2

We have already started producing training datasets of model gamma-ray light curves to train NNs that emulate the model gamma-ray light curves.



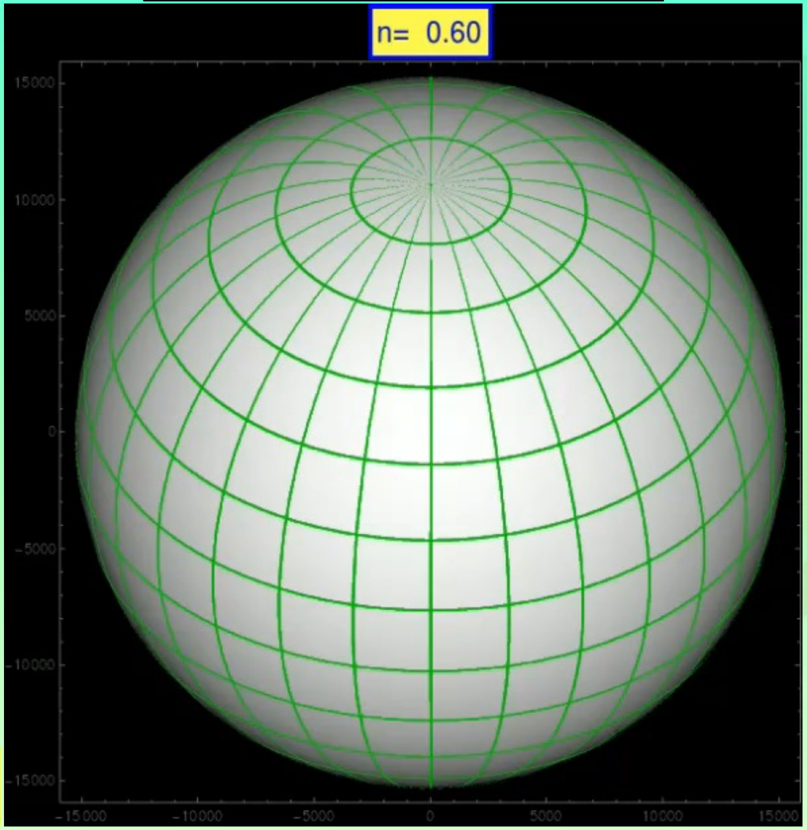
Atmosphere model

Atmosphere Model
$\mathbf{I} = \mathcal{F}(\vartheta_z, E)$
$\mathbf{I} \propto \cos^n(\vartheta_z)$

For the Riley et al. HSs, we were able to reproduce the X-ray LC for $n \approx 1$

For the Miller et al. HSs, we were able to reproduce the X-ray LC for $n \approx 0.65$

Projection on the observer plane



$\zeta = 53.85°$