Toward Energy-Resolved Neural-Network Emulators for Joint Inference of Neutron-Star Mass, Radius, and Multipolar Magnetic Fields with NICER

Yaswant Sai Ejjagiri[1,2,3], Greg Olmschenk[4,2,3], Constantinos Kalapotharakos[2], Jaime Raynaud[6], Wendy F. Wallace[7], Anu Kundu[2,5], Zorawar Wadiasingh[4,2,3], Demosthenes Kazanas[2]

[1]SURA, USA [2]NASA Goddard Space Flight Center, USA [3]CRESST II, USA [4]University of Maryland, College Park, USA [5]North-West University, South Africa [6]Max Planck Institute for Astrophysics, Germany [7]University of Hamburg, Germany

INTRODUCTION

NICER pulse-profile modeling of millisecond pulsars uses relativistic hot-spot emission models to **infer neutron star properties, like mass and radius**, from phase and energy resolved X-ray data. For **PSR J0030+0451**, Miller et al. (2019) and Riley et al. (2019) independently modeled the hot spots directly using parameters describing their shapes, locations, and temperatures. They found that the two hot spots lie in the same rotational hemisphere. This geometry led Kalapotharakos et al. (2021) to replace these directly parameterized hot spots with the footpoints of open magnetic-field lines in an offset dipole-plus-quadrupole model.

Bayesian inference requires exploring this forward model over many parameter combinations which can be computationally expensive. **Neural network (NN) emulators** address this limitation by learning the mapping **from model parameters to simulated light curves**. Olmschenk et al. (2025) found a *speed-up of more than 400x times* for their static-vacuum-field model allowing the NN to replace the physical model during MCMC exploration. Taiyebah et al. (2026) and Kundu et al. (2026) extended this strategy to related multipolar-field models, showing that NN emulators can accelerate inference for complex magnetic field geometries.

However, the **previous emulators only predict bolometric light curves** discarding the spectral information from the NICER observations.

GOAL

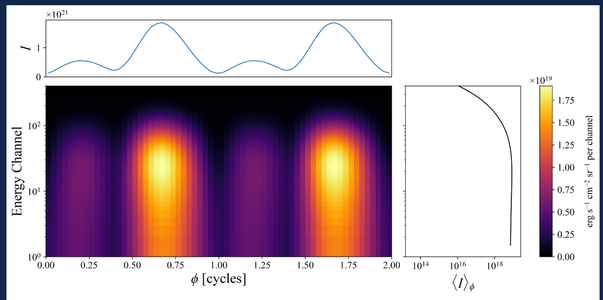
Build an **energy-resolved NN emulator** predicting the **full 2D X-ray phase-energy phaseogram** keeping NICER's spectral information while preserving the **computational speedup** of the NN-based MCMC

DATA

How is the training data simulated?

1. **Draw parameters**
   Uniform priors with physical-validity cuts
2. **Field → hot spots**
   Open-field-line footpoints of the offset dipole + quadrupole
3. **Ray trace + atmosphere**
   Light bending + Doppler & NSX hydrogen atmosphere
4. **Training sample**
   Phase–energy phaseogram (or 1D light curve), ×32 rotations

Parameters	1D	2D
Dipole + quadrupole field: offsets, directions, B_Q/B_D (11)	✓	✓
Mass M, radius R_{eq}, observer angle ζ	fixed	✓
Hot-spot temperatures log T_1, log T_2 (per polarity)	fixed	✓
Distance d, absorption N_H	—	after NN



Phaseogram and corresponding bolometric light curve (top panel) and energy spectrum (right panel) for a representative training set NS with spin frequency 205.53 Hz and an arbitrary set of θ

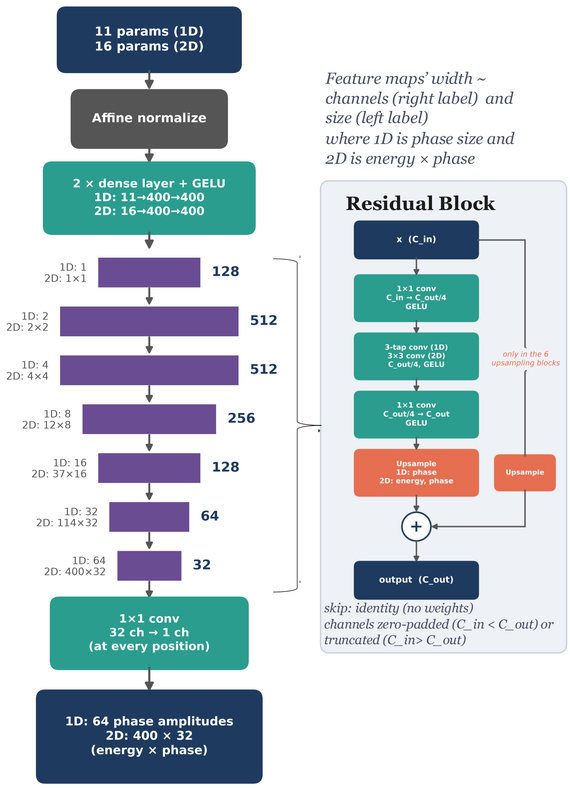
Splits	Light Curves	Phaseograms
Training	50M	50M
Validation	100K	100K
Test	100K	100K

NEURAL NETWORK

ResNet-style CNN (named *AntidotePrototype8* from `haplo`) trained for **24-96 hrs** on 64 NVIDIA A100 GPUs

Network is made up of dense and convolutional layers (green) and residual blocks (purple). Model starts with 1 residual block and then has 6 stages where each stage has 3 blocks:

- One upsampling and then 2 same size
  - For 1D, upsample ×2 along phase
  - For 2D, upsample ×2, ×2 in stages 1-2, ×3.1, ×2 in stages 3-5, ×3.51, ×1 in stage 6



Feature maps' width ~ channels (right label) and size (left label) where 1D is phase size and 2D is energy × phase

skip: identity (no weights) channels zero-padded (C_in < C_out) or truncated (C_in> C_out)

LOSS FUNCTIONS

i: batch/sample index B: batch size, F_i: Simulated phaseogram
ϕ: phase bin index E: Energy bin index $\hat{F}_i$: Predicted phaseogram
$\langle\cdot\rangle$: Mean over phaseogram cells

1D: Median normalized squared error (MdNSE)

$$\frac{1}{B}\sum_{i=1}^{B}\log_{10}\left(\frac{\sum_\phi(\hat{y}_{i,\phi}-y_{i,\phi})^2}{[\mathrm{median}_\phi(y_{i,\phi})]^2}\right)$$

2D: Same Form For Phaseograms (MrNSE)

$$\frac{1}{B}\sum_{i=1}^{B}\log_{10}\left(\frac{\left\langle(\hat{F}_i-F_i)^2\right\rangle_{E,\phi}}{[(\max_{E,\phi}F_i+\min_{E,\phi}F_i)/2]^2}\right)$$

Alternative 2D Losses

① L_{map}: RMS-Normalized Phaseogram

$$\frac{1}{B}\sum_{i=1}^{B}\left\langle\left(\frac{\hat{F}_{i,E,\phi}-F_{i,E,\phi}}{S_2(E)}\right)^2\right\rangle_{E,\phi}$$

$S_2(E)$: RMS scale for energy bin E Set once based on the entire training set, putting all energies on one scale

② L_{bol}: Mean Bolometric

$$\frac{1}{B}\sum_{i=1}^{B}\left\langle\left(\frac{\hat{B}_{i,\phi}-B_{i,\phi}}{D_i}\right)^2\right\rangle_{\phi}$$

$B_{i,\phi}=\sum_E F_{i,E,\phi}$

$D_i=([\max_\phi B_{i,\phi}+\min_\phi B_{i,\phi}]/2)$

③ L_{spec}: Spectral Shape

$$\frac{1}{B}\sum_{i=1}^{B}\left\langle\left(\hat{C}_{i,E,\phi}-C_{i,E,\phi}\right)^2\right\rangle_{E,\phi}$$

$\hat{C}$: normalized cumulative spectrum of $\hat{F}_i$
$C_{i,E,\phi}$: normalized cumulative spectrum of F_i

④ L_{bol} + L_{spec} Combined

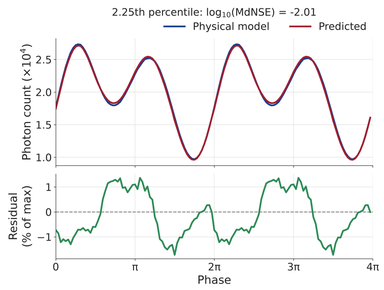
$$\frac{1}{B}\sum_{i=1}^{B}\log_{10}\left(\frac{\mathcal{L}_{\mathrm{bol},i}/R_{\mathrm{bol}}+\mathcal{L}_{\mathrm{spec},i}/R_{\mathrm{spec}}}{2}\right)$$

$R_{\mathrm{bol}}=5.70\times10^{-4}$, $R_{\mathrm{spec}}=6.45\times10^{-4}$

Reference values are set by the error of Miller et al. (2019)'s best fit to NICER J0030 data

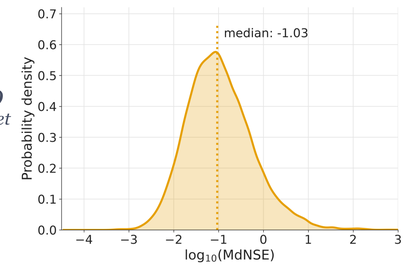
We calculate **baseline** values for each 2D loss function from the **Poisson uncertainty of NICER's J0030 counts**, $\sigma=\sqrt{N_T+N_B}$, per phase-energy cell (Kalapotharakos et al. 2021), using Miller et al. (2019)'s best fit as the reference phaseogram

RESULTS: 1D MODEL (24HRS)

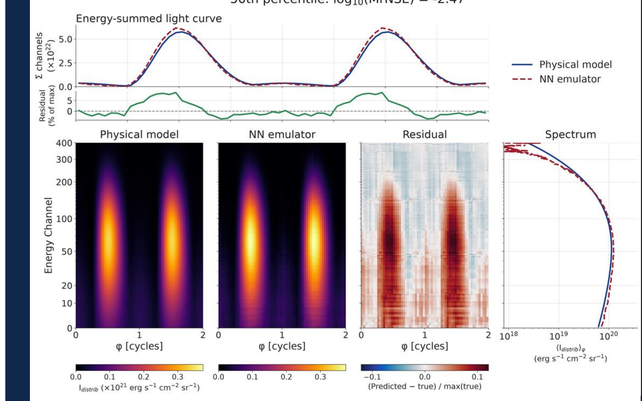


*Held-out test light curve at the 2.25th percentile of fit quality where **residuals (<2%)** shown as a % of the light curve's peak*

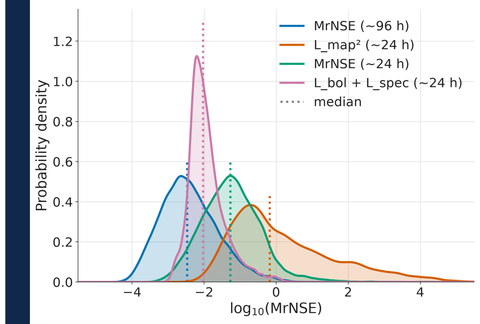
*MdNSE scores **(median = -1.03)** over the 100k test set aligning with Olmschenk et al. (2025)'s results*



RESULTS: 2D MODEL

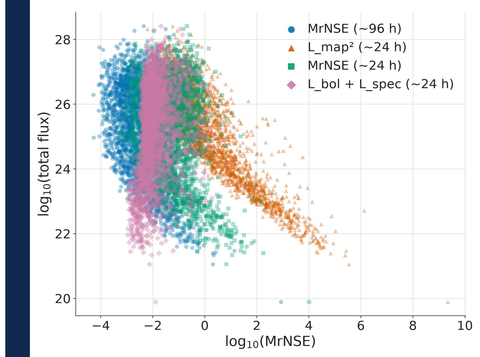


*Held-out test phaseogram at the 50th percentile of fit quality and **residuals stay within ~10%** of the peak across phase and energy*



*MrNSE distribution on held-out test samples for each 2D loss function, with training time. **Lower is better** & dotted lines are medians*

FAILURE MODES



*MrNSE vs. total flux of each held-out test sample for each 2D loss function. Both L_{map} and L_{MrNSE} get **steadily worse as maps get fainter** while $L_{bol,spec}$ has the opposite trend*

- **Faint, cool maps** struggle under L_{map} and L_{MrNSE}, where *bright cells* dominate
- **Empty phase bins** (hot spot eclipsed) and **sharp spectral drops** in data

FUTURE WORK

1) Bayesian inference: plug the 2D NN into an MCMC chain to obtain J0030 posteriors, *jointly constraining* mass, radius, viewing geometry and multipolar magnetic-field structure

2) Force-free fields: fine-tune the emulator on force-free magnetosphere simulations instead of static vacuum

3) Joint inference with γ-rays: Add phase-aligned Fermi-LAT γ-ray light curves within the same field geometry

References: *Kalapotharakos, C., et al. 2021, ApJ, 907, 63; Kundu, A., et al. 2026, arXiv:2604.19534; Miller, M. C., et al. 2019, ApJL, 887, L24; Olmschenk, G., et al. 2025, AJ, arXiv:2501.12383; Riley, T. E., et al. 2019, ApJL, 887, L21; Taiyebah, F., et al. 2026, arXiv:2606.30886.* **This material is based upon work supported by NASA under award number 80GSFC24M0006. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of NASA. Computing resources were provided by the NASA High-End Computing Program through the NASA Advanced Supercomputing Division at Ames Research Center.*