Pulsar population synthesis with multi-modal machine learning

Dr. Vanessa Graber (graber@ice.csic.es)

in collaboration with Michele Ronchi, Celsa Pardo Araujo, Nanda Rea & José Pons
We have detected ~3,000 pulsars, with ~2,000 being isolated. They have diverse characteristics, visible in different parts of electromagnetic spectrum → neutron star zoo.

Galactic core-collapse supernova rate is insufficient for independent formation of these classes (Keane & Kramer 2008) → evolutionary links (e.g., Viganò et al. 2013)?

<table>
<thead>
<tr>
<th>⍺_{PSR}, n_e</th>
<th>PSRs</th>
<th>RRATs</th>
<th>XDINSs</th>
<th>Magnetars</th>
<th>Total</th>
<th>CCSN rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>FK06, NE2001</td>
<td>2.8 ± 0.5</td>
<td>5.6^{+4.3}_{-3.3}</td>
<td>2.1 ± 1.0</td>
<td>0.3^{+1.2}_{-0.2}</td>
<td>10.8^{+7.0}_{-5.0}</td>
<td>1.9 ± 1.1</td>
</tr>
<tr>
<td>L+06, NE2001</td>
<td>1.4 ± 0.2</td>
<td>2.8^{+1.6}_{-1.6}</td>
<td>2.1 ± 1.0</td>
<td>0.3^{+1.2}_{-0.2}</td>
<td>6.6^{+4.0}_{-3.0}</td>
<td>1.9 ± 1.1</td>
</tr>
<tr>
<td>L+06, TC93</td>
<td>1.1 ± 0.2</td>
<td>2.2^{+1.7}_{-1.7}</td>
<td>2.1 ± 1.0</td>
<td>0.3^{+1.2}_{-0.2}</td>
<td>5.7^{+4.1}_{-2.7}</td>
<td>1.9 ± 1.1</td>
</tr>
<tr>
<td>V+04, NE2001</td>
<td>1.6 ± 0.3</td>
<td>3.2^{+2.5}_{-1.9}</td>
<td>2.1 ± 1.0</td>
<td>0.3^{+1.2}_{-0.2}</td>
<td>7.2^{+5.0}_{-3.4}</td>
<td>1.9 ± 1.1</td>
</tr>
<tr>
<td>V+04, TC93</td>
<td>1.1 ± 0.2</td>
<td>2.2^{+1.7}_{-1.3}</td>
<td>2.1 ± 1.0</td>
<td>0.3^{+1.2}_{-0.2}</td>
<td>5.7^{+4.1}_{-2.7}</td>
<td>1.9 ± 1.1</td>
</tr>
</tbody>
</table>

NS birth vs. Galactic CCSN rates (Keane & Manchester 2008).
Background

- We have detected ~3,000 pulsars, with ~2,000 being isolated. They have diverse characteristics, visible in different parts of electromagnetic spectrum → neutron star zoo.

- Galactic core-collapse supernova rate is insufficient for independent formation of these classes (Keane & Kramer 2008) → evolutionary links (e.g., Viganò et al. 2013)?

NS birth vs. Galactic CCSN rates (Keane & Manchester 2008).

grabr@ice.csic.es
Pulsar Population Synthesis

- We only observe a very small fraction of the $\sim10^9$ neutron stars expected in the Milky Way. Instead of looking at individual pulsars, population synthesis focuses on the entire population (e.g. Faucher-Giguère & Kaspi 2006, Lorimer et al. 2006, Gullón et al. 2014, Cieślar et al. 2020):
  - Model birth properties with a Monte-Carlo approach.
  - Evolve parameters forward in time.
  - Apply filters to mimic observational biases/limits.
  - Compare mock samples to observations to constrain input.

- Method provides e.g. birth rates and initial $v/P/B$ distributions.

  focus on Deep Learning (DL) for the comparison

graber@ice.csic.es
Focus on **positional & velocity evolution** and follow earlier works using an updated spiral-arm (Yao et al. 2017) and Galactic model (Marchetti et al. 2019) plus rigid Galaxy rotation.

- Evolve $10^5$ stars for (up to) $10^7$ yr varying exponential **scale height** $h_c$ for birth distance from plane **dispersion** $\sigma_k$ of Maxwell distribution for kick velocities to obtain 16,384 mock populations. Extract images (labelled by $h_c, \sigma_k$) as **training data** for a convolutional neural network (CNN).

Galactic evolution tracks for $h_c = 0.18$ kpc, $\sigma_k = 265$ km/s.

graber@ice.csic.es
Focus on **positional & velocity evolution** and follow earlier works using an updated spiral-arm (Yao et al. 2017) and Galactic model (Marchetti et al. 2019) plus rigid Galaxy rotation.

Evolve $10^5$ stars for (up to) $10^7$ yr varying exponential scale height $h_c$ for birth distance from plane and dispersion $\sigma_k$ of Maxwell distribution for kick velocities to obtain 16,384 mock populations. Extract images (labelled by $h_c, \sigma_k$) as **training data** for a convolutional neural network (CNN).

Stellar density and velocities in ICRS coordinates.

graber@ice.csic.es
Focus on **positional & velocity evolution** and follow earlier works using an updated spiral-arm (Yao et al. 2017) and Galactic model (Marchetti et al. 2019) plus rigid Galaxy rotation.

- Evolve $10^5$ stars for (up to) $10^7$ yr varying exponential scale height $h_c$ for birth distance from plane and dispersion $\sigma_k$ of Maxwell distribution for kick velocities to obtain 16,384 mock populations. Extract images (labelled by $h_c, \sigma_k$) as training data for a convolutional neural network (CNN).

Stellar density and velocities in ICRS coordinates. Examine the labels $(h_c, \sigma_k)$ from the 3 images = regression problem.
DL Proof-of-concept Study: Results

CNN architecture and results from Ronchi et al. (2021).

- We use RMSE as loss function & validation metric, Kaiming initialisation, Adam for gradient-descent optimisation, and 80 / 20% for the training / validation process.

graber@ice.csic.es
We use RMSE as loss function & validation metric, Kaiming initialisation, Adam for gradient-descent optimisation, and 80/20% for the training/validation process.

MRE = 0.061 gives ME = 0.01 kpc at $h_c = 0.18$ kpc
MRE = 0.039 gives ME = 10 km/s at $\sigma_k = 265$ km/s

graber@ice.csic.es
DL Proof-of-concept Study: Selection Biases

- Analyse the **CNN’s predictive power** as a function of available data points (i.e. NSs) by resampling our fiducial simulation to **incorporate selection biases** from NSs with **proper-motion measurements**.

![Graph 1](image1.png)
![Graph 2](image2.png)

graber@ice.csic.es
DL Proof-of-concept Study: Selection Biases

- Analyse the **CNN’s predictive power** as a function of available data points (i.e. NSs) by resampling our fiducial simulation to **incorporate selection biases** from NSs with proper-motion measurements.

Increasing the number of detected neutron stars and accurately **classifying them with SKA** will allow us to better constrain the pulsar population.
Magneto-rotational Evolution & Emission

- Model **additional information** to constrain population properties further:
  - predict $P$ and $\dot{P}$ by modelling **magneto-rotational evolution**

- Combine with an **emission model** and **survey detectability limits** to determine observability. This produces mock **$PP$ populations**, which we compare to the observed pulsar sample with CNNs.

Evolution tracks in $P\dot{P}$-plane. Objects are born on top left.

graber@ice.csic.es
Magneto-rotational Evolution & Emission

- Model **additional information** to constrain population properties further:
  - predict $P$ and $\dot{P}$ by modelling **magneto-rotational evolution**
- Combine with an **emission model** and **survey detectability limits** to determine observability. This produces mock $PP$ populations, which we compare to the observed pulsar sample with CNNs.

**inference** information about the **initial $P$ & $B$ distributions** and magnetic field evolution
Multi-modal DL

● Different parts of the pulsar population are observed in different wavebands: observations provide complementary information about the same underlying neutron star population.

● In computer science, this concept is known as multi-modal learning.

“Batman raises the stakes in his war on crime. With the help of Lt. Jim Gordon and District Attorney Harvey Dent, Batman sets out to dismantle the remaining criminal organizations that plague the streets. The partnership proves to be effective, but they soon find themselves prey to a reign of chaos unleashed by a rising criminal mastermind known to the terrified citizens of Gotham as the Joker.” - TMDB
Different parts of the pulsar population are observed in different wavebands: observations provide complementary information about the same underlying neutron star population.

In computer science, this concept is known as multi-modal learning.

Multi-modal learning architecture combining information from movie poster plus description (credit: TMDB) to predict the genre of a movie.
Conclusions

Deep Learning with CNNs is a promising tool to infer birth properties from the current population.

Pulsar population synthesis constrains birth rates of different neutron star classes and birth properties of the entire population.

Complementary information from different wavelengths can be combined in a multi-modal network.

Stay tuned for updates and/or come and talk to us!!

grabber@ice.csic.es
Back-up Slides
CNN validation results for $h_c$:
RMSE = 0.038 kpc & MRE = 0.061.

CNN validation results for $\sigma_k$:
RMSE = 8.8 km/s & MRE = 0.039.
Ronchi et al. (2021) - Degeneracy

- Histogramming **distances from the Galactic plane** for current mock pulsar populations, we note a **degeneracy between $h_c$ and $\sigma_k$**: large scale heights combined with small velocity dispersions lead to same outcomes as small scale heights with large velocity dispersions.

- **CNN recovers the degeneracy!**

We find an anticorrelation between the residuals in $h_c$ and $\sigma_k$. 

graberg@ice.csic.es
To incorporate selection effects & observational biases, we use a **phenomenological approach** to reanalyse the CNN performance.

- Use proper motion and distance estimates for 216 isolated pulsars to deduce **empirical selection function** $f(d_\odot)$ and **resample population** with $h_c = 0.18$ kpc and $\sigma_k = 265$ km/s.

216 isolated NSs with proper motions.

---

Ronchi et al. (2021) - Selection Function

$f(d_\odot) = d_\odot^{-1} \exp(-d_\odot/2)$