Efficient PINN Models of Cosmic Rays

A lightweight surrogate model for the γ-ray background

Eric Putney[1], Yujin Park[1], David Shih[1], Matthew R. Buckley[1], Tracy Slatyer[2], Christopher Chen[2]

[1]Rutgers NHETC, [2]MIT CTP

Pheno 2026

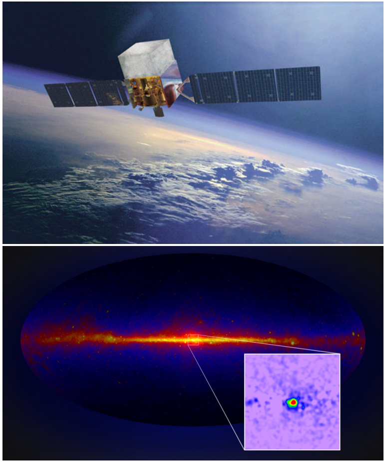
May 11

Section 1

The Galactic Center GeV Excess

The GeV Excess

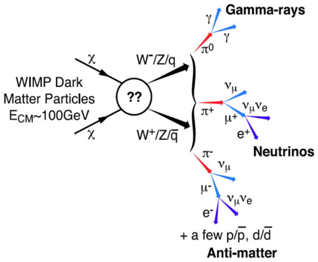
- Diffuse glow of GeV γ-rays from the inner Milky Way (MW). [Goodenough & Hooper, arXiv:0910.2998]

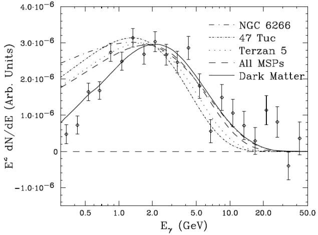


Top: NASA/Aurore Simonnet; Bottom: Goodenough, Hooper

The GeV Excess

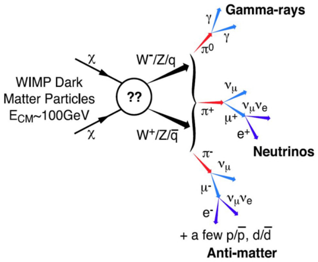
- Diffuse glow of GeV γ-rays from the inner Milky Way (MW). [Goodenough & Hooper, arXiv:0910.2998]
- Spectrum + morphology match thermal-relic WIMP annihilation ($m_\chi \sim 30$ GeV). [Daylan et al., arXiv:1402.6703]

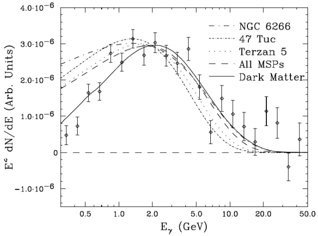




The GeV Excess

- Diffuse glow of GeV γ-rays from the inner Milky Way (MW). [Goodenough & Hooper, arXiv:0910.2998]
- Spectrum + morphology match thermal-relic WIMP annihilation ($m_\chi \sim 30$ GeV). [Daylan et al., arXiv:1402.6703]
- Holds up to decade+ of analysis. Improved diffuse models, point-source catalogs.





What else could it be?

- Primary astrophysical alternative: unresolved millisecond pulsars (MSPs).[a]

ESA & F. Ferraro (Bologna Astron. Obs.)

[a] Or PBHs...

What else could it be?

- Primary astrophysical alternative: unresolved millisecond pulsars (MSPs).[a]
- Early fits needed $\mathcal{O}(10^2\text{–}10^3)$ MSPs; now requires $\mathcal{O}(>10^4)$ in the GC[b]. [Hooper & Mohlabeng, arXiv:1506.05124]

ESA & F. Ferraro (Bologna Astron. Obs.)

[a] Or PBHs...
[b] *Four* MSPs have now been found in the GC!

What else could it be?

- Primary astrophysical alternative: unresolved millisecond pulsars (MSPs).[a]
- Early fits needed $\mathcal{O}(10^2\text{–}10^3)$ MSPs; now requires $\mathcal{O}(>10^4)$ in the GC[b]. [Hooper & Mohlabeng, arXiv:1506.05124]
- Photon-statistics tests (NP template fits, wavelets) conflicting: sensitive to the background model [Mishra-Sharma & Cranmer, arXiv:2110.06931]

ESA & F. Ferraro (Bologna Astron. Obs.)

---

[a] Or PBHs...
[b] *Four* MSPs have now been found in the GC!

Section 2

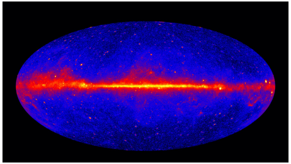
Diffuse Gamma-ray Background

Where the background comes from

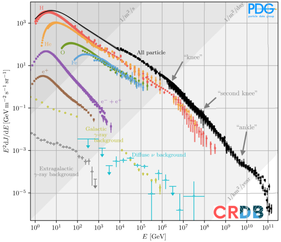
Galactic diffuse gamma rays trace the cosmic-ray (CR) population convolved with the ISM:

- π^0 decay ($p \times$ gas)
- Inverse Compton ($e^- \times$ ISRF)
- Bremsstrahlung ($e^- \times$ gas)

Modeling these requires the full 3D distribution of CRs across the Galaxy – not just the locally measured spectra.



NASA / DOE / Fermi-LAT Collaboration



PDG, Review of Particle Physics 2025

The cosmic-ray transport equation

The CR flux distribution $\psi_i(\vec{r}, E)$ for each species obeys

$$\frac{\partial \psi}{\partial t} = q(\vec{r}, E) + \vec{\nabla} \cdot \left(D_{xx} \vec{\nabla} \psi - \vec{V} \psi \right) + \frac{\partial}{\partial E} E^2 D_{EE} \frac{\partial}{\partial E} \frac{1}{E^2} \psi - \frac{\partial}{\partial E} \left[\dot{E}\, \psi - \frac{E}{3} (\vec{\nabla} \cdot \vec{V}) \psi \right] - \frac{\psi}{\tau_f} - \frac{\psi}{\tau_r}$$

- Source, diffusion, convection, reacceleration, energy losses, fragmentation, decay.
- Coupled across species (primary, secondary, tertiary $e^{\pm}$, p, nuclei ...).
- Large energy range, physical scales and processes.

Gold standard numerical code is GALPROP [Strong & Moskalenko, ApJ 509, 212 (1998)].

GALPROP

- Public, finite-difference transport solver [Strong & Moskalenko, 1998].

$$\vec{\alpha}_{\mathrm{GP}}, \vec{r}, E \longrightarrow \boxed{\text{GALPROP}} \longrightarrow \psi_i$$

GALPROP

- Public, finite-difference transport solver [Strong & Moskalenko, 1998].
- User picks $\vec{\alpha}_{\mathrm{GP}}$ (diffusion, source, gas, ISRF, . . .); GALPROP evolves to steady state over $\vec{r}, E$ grid.

$$\vec{\alpha}_{\mathrm{GP}}, \vec{r}, E \longrightarrow \boxed{\text{GALPROP}} \longrightarrow \psi_i$$

GALPROP

- Public, finite-difference transport solver [Strong & Moskalenko, 1998].
- User picks $\vec{\alpha}_{\mathrm{GP}}$ (diffusion, source, gas, ISRF, ...); GALPROP evolves to steady state over $\vec{r}, E$ grid.
- However – many params, hours of CPU, $\mathcal{O}(10^2)$ GB RAM per run.

$$\vec{\alpha}_{\mathrm{GP}}, \vec{r}, E \longrightarrow \boxed{\text{GALPROP}} \longrightarrow \psi_i$$

Modeling the diffuse background (is hard)

- Many poorly-constrained CR/GALPROP parameters: diffusion, convection, source distribution, gas maps, ISRF, . . .

Modeling the diffuse background (is hard)

- Many poorly-constrained CR/GALPROP parameters: diffusion, convection, source distribution, gas maps, ISRF, . . .
- Standard practice: pick a handful of fiducial parameter sets, fit the GCE over the bracket. Template fitting has been shown to be biased! [Leane & Slatyer, arXiv:1904.08430, 2002.12371; Buschmann et al., arXiv:2002.12373]

Modeling the diffuse background (is hard)

- Many poorly-constrained CR/GALPROP parameters: diffusion, convection, source distribution, gas maps, ISRF, . . .
- Standard practice: pick a handful of fiducial parameter sets, fit the GCE over the bracket. Template fitting has been shown to be biased! [Leane & Slatyer, arXiv:1904.08430, 2002.12371; Buschmann et al., arXiv:2002.12373]
- Marginalizing over the plausible propagation parameter space takes thousands of GALPROP runs – expensive/intractable.

Modeling the diffuse background (is hard)

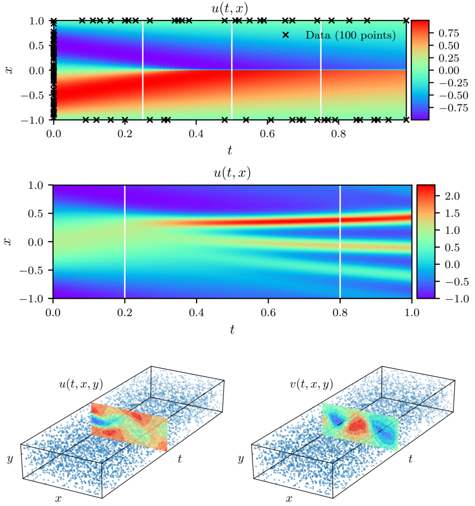
- Many poorly-constrained CR/GALPROP parameters: diffusion, convection, source distribution, gas maps, ISRF, . . .
- Standard practice: pick a handful of fiducial parameter sets, fit the GCE over the bracket. Template fitting has been shown to be biased! [Leane & Slatyer, arXiv:1904.08430, 2002.12371; Buschmann et al., arXiv:2002.12373]
- Marginalizing over the plausible propagation parameter space takes thousands of GALPROP runs – expensive/intractable.
- We need a new approach that solves the transport equation efficiently and flexibly.

Section 3

PINNs

Physics-Informed Neural Networks

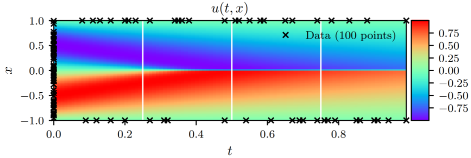
- Represent solution to PDE $\mathcal{D}[\psi_\theta] = 0$ as a neural network. [Raissi et al., arXiv:1711.10561 & 1711.10566]
  - Derivatives of ψ by *autodiff*.

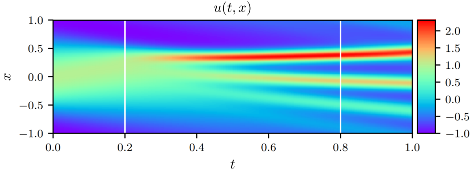


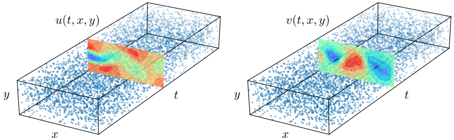
Physics-Informed Neural Networks

- Represent solution to PDE $\mathcal{D}[\psi_\theta] = 0$ as a neural network. [Raissi et al., arXiv:1711.10561 & 1711.10566]
  - Derivatives of ψ by *autodiff*.
- Loss = MSE of PDE residual on random collocation points:

$$\mathcal{L} = \frac{1}{N} \sum_{i \in V} \left| \mathcal{D}[\psi_\theta]_i \right|^2 + \lambda_{\partial V} \, \mathcal{L}_{\partial V} \, .$$







PINNs as a faster “GALPROP”

1. Cheap, arbitrary inference after training. Learned flux is a continuous function of space, energy. Evaluating ψ anywhere is a quick forward pass.

PINNs as a faster "GALPROP"

1. Cheap, arbitrary inference after training. Learned flux is a continuous function of space, energy. Evaluating ψ anywhere is a quick forward pass.
2. Set of solutions over α_{GP}. Pass CR parameters as inputs, $\psi_\theta(\vec{x};\, \vec{\alpha}_{\mathrm{GP}})$ – *one trained model could replace a full set of GALPROP runs.*

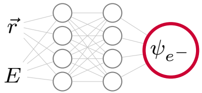
Section 4

Results

Primary electrons: diffusion + source

$$0 = q(\vec{r}, E) + D_{xx}\nabla^2\psi$$

- Single species (e^-), steady state.

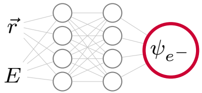


Diffusion	**GALPROP**	**PINN**
	Solve: ~1 hr	Train: 20 mins
Inference	~41 ms/10k pts	2.1 ms/10k pts
Memory	20 GB RAM	1 GB VRAM

Primary electrons: diffusion + source

$$0 = q(\vec{r}, E) + D_{xx}\nabla^2\psi$$

- Single species (e^-), steady state.
- Broken power-law source injection over $10^2 - 10^9$ MeV.

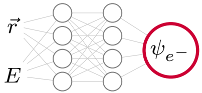


Diffusion	**GALPROP**	**PINN**
	Solve: $\sim$1 hr	Train: 20 mins
Inference	$\sim$41 ms/10k pts	2.1 ms/10k pts
Memory	20 GB RAM	1 GB VRAM

Primary electrons: diffusion + source

$$0 = q(\vec{r}, E) + D_{xx}\nabla^2\psi$$

- Single species (e^-), steady state.
- Broken power-law source injection over $10^2 - 10^9$ MeV.
- PINN outputs $\log_{10}\left[\psi(\vec{r}, E)\, E^2\right]$.



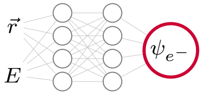
Diffusion	**GALPROP**	**PINN**
	Solve: $\sim$1 hr	Train: 20 mins
Inference	$\sim$41 ms/10k pts	2.1 ms/10k pts
Memory	20 GB RAM	1 GB VRAM

Primary electrons: diffusion + source

$$0 = q(\vec{r}, E) + D_{xx}\nabla^2\psi$$

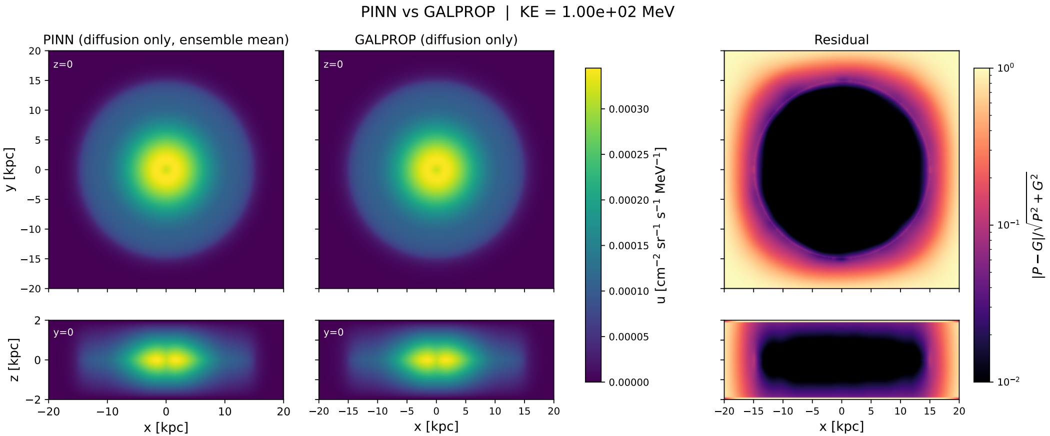
- Single species (e^-), steady state.
- Broken power-law source injection over $10^2 - 10^9$ MeV.
- PINN outputs $\log_{10}\left[\psi(\vec{r}, E)\, E^2\right]$.
- Benchmark against GALPROP[a]

[a] The PINN *never* sees GALPROP during training.



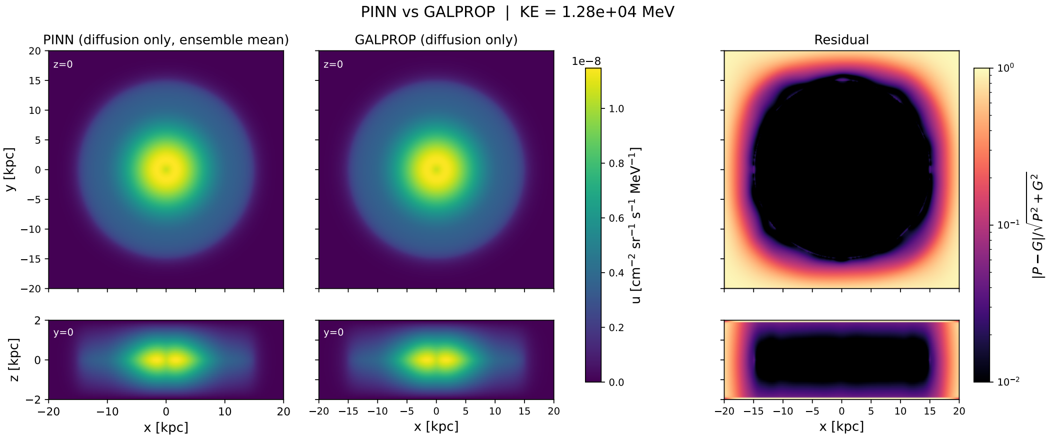
Diffusion	**GALPROP**	**PINN**
	Solve: $\sim$1 hr	Train: 20 mins
Inference	$\sim$41 ms/10k pts	2.1 ms/10k pts
Memory	20 GB RAM	1 GB VRAM

Diffusion-only: morphology



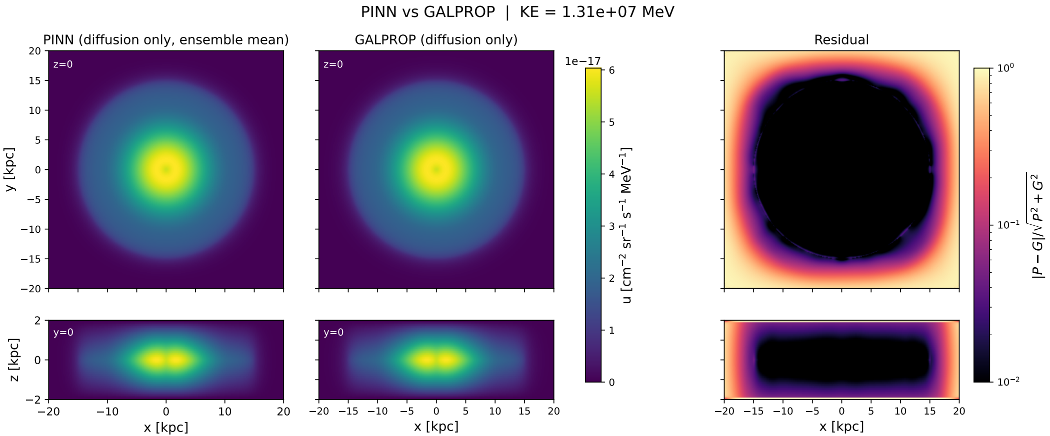
Sub-percent agreement with GALPROP across the *entire* energy range.

Diffusion-only: morphology



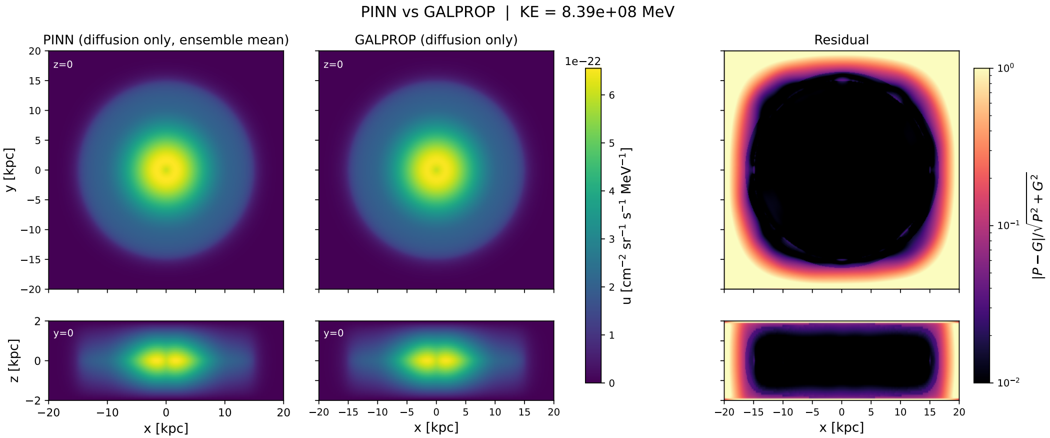
Sub-percent agreement with GALPROP across the *entire* energy range.

Diffusion-only: morphology



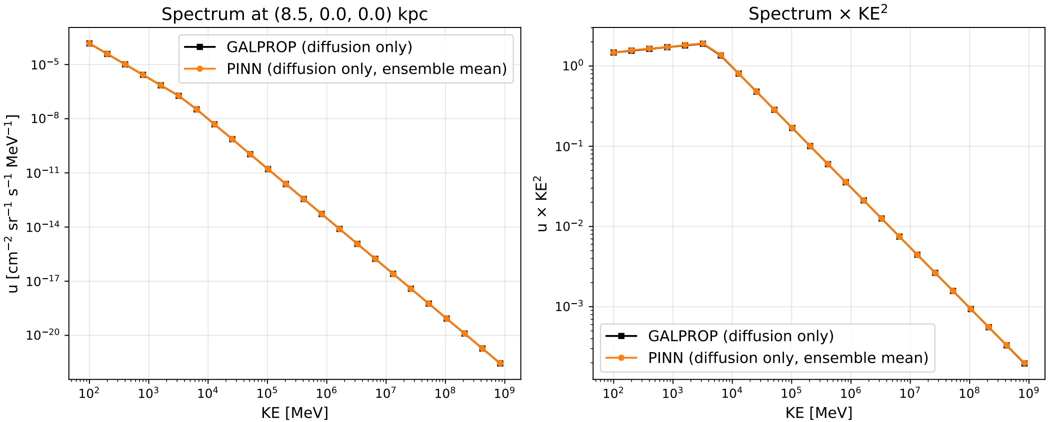
Sub-percent agreement with GALPROP across the *entire* energy range.

Diffusion-only: morphology



Sub-percent agreement with GALPROP across the *entire* energy range.

Diffusion-only: spectrum



Sub-percent agreement with GALPROP across the *entire* energy range.

Adding energy losses (ICS off the CMB)

$$0 = q + D_{xx} \nabla^2 \psi - \frac{\partial}{\partial E}(\dot{E}\,\psi)$$

- Loss term dominates at high E – fine cancellations with diffusion.

ICS losses	**GALPROP**	**PINN**
	Solve: 1 hr	Train: 1 hr
Inference	41 ms/10k pts	2.1 ms/10k pts
Memory	183 GB RAM	3 GB VRAM

Adding energy losses (ICS off the CMB)

$$0 = q + D_{xx}\nabla^2\psi - \frac{\partial}{\partial E}(\dot{E}\,\psi)$$

- Loss term dominates at high E – fine cancellations with diffusion.
- Warm start from diff-only using curriculum learning.

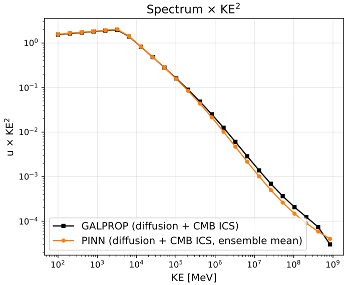
ICS losses	**GALPROP**	**PINN**
	Solve: 1 hr	Train: 1 hr
Inference	41 ms/10k pts	2.1 ms/10k pts
Memory	183 GB RAM	3 GB VRAM

Adding energy losses (ICS off the CMB)

$$0 = q + D_{xx}\nabla^2\psi - \frac{\partial}{\partial E}(\dot{E}\,\psi)$$

- Loss term dominates at high E – fine cancellations with diffusion.
- Warm start from diff-only using curriculum learning.
- Chased a high-E mismatch for months...

ICS losses	**GALPROP**	**PINN**
	Solve: 1 hr	Train: 1 hr
Inference	41 ms/10k pts	2.1 ms/10k pts
Memory	183 GB RAM	3 GB VRAM

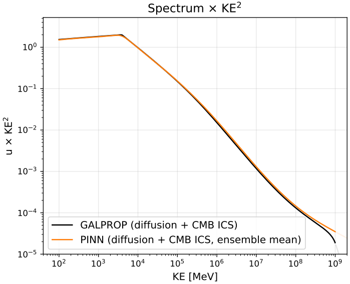


Adding energy losses (ICS off the CMB)

$$0 = q + D_{xx}\nabla^2\psi - \frac{\partial}{\partial E}(\dot{E}\,\psi)$$

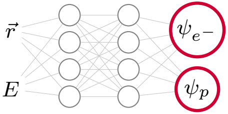
- Loss term dominates at high E – fine cancellations with diffusion.
- Warm start from diff-only using curriculum learning.
- Chased a high-E mismatch for months...
- ...GALPROP's energy grid was under-resolved. The PINN was right!

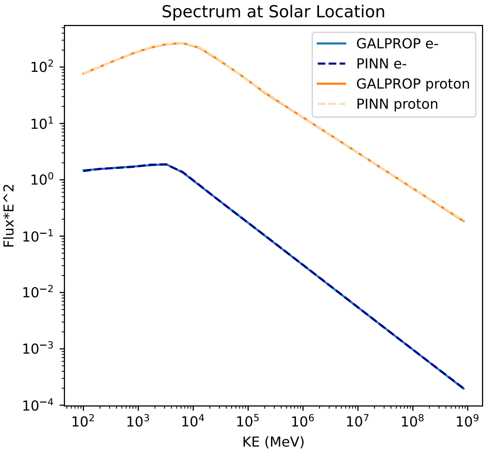
ICS losses	**GALPROP**	**PINN**
	1 hr	Train: 1 hr
Inference	41 ms/10k pts	2.1 ms/10k pts
Memory	183 GB RAM	3 GB VRAM



Multi-species: e^- + p

- Add protons as a *second output node* – wire in to another copy of transport eq. and solve simul. w/ e^-

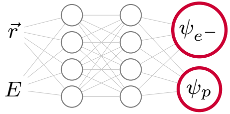


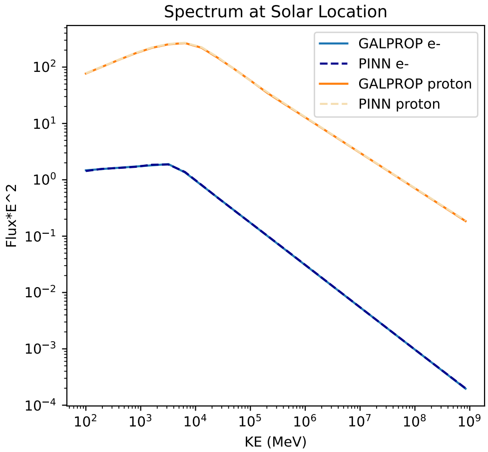


[a] IC doesn't matter much for p

Multi-species: e^- + p

- Add protons as a *second output node* – wire in to another copy of transport eq. and solve simul. w/ e^-
- Sub-percent agreement w/ GALPROP (diffusion-only).[a]





[a] IC doesn't matter much for p

Next: toward a full GALPROP surrogate

Now ✓
e^-, p
diffusion + ICS

Secondaries
$\bar{p}$, e^+, nuclei

γ-ray map
fold against
gas + ISRF

Condition on $\vec{\alpha}_{\mathrm{GP}}$
one net,
every GALPROP run

Marginalize GCE
over BG
systematics

Next: toward a full GALPROP surrogate

Now ✓
e^-, p
diffusion + ICS

Secondaries
$\bar{p}$, e^+, nuclei

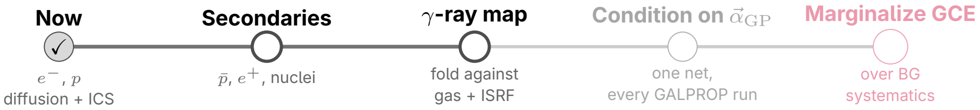
γ-ray map
fold against
gas + ISRF

Condition on $\vec{\alpha}_{\mathrm{GP}}$
one net,
every GALPROP run

Marginalize GCE
over BG
systematics

Coupled-species solver: $\bar{p}$, e^+, and nuclear cascades.

Next: toward a full GALPROP surrogate



γ-ray emissivities are a downstream integral of ψ against gas density and ISRF

Next: toward a full GALPROP surrogate

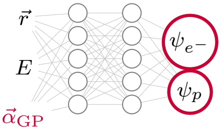
Now — e^-, p diffusion + ICS

Secondaries — $\bar{p}$, e^+, nuclei

γ-ray map — fold against gas + ISRF

Condition on $\vec{\alpha}_{\rm GP}$ — one net, every GALPROP run

Marginalize GCE — over BG systematics



Parameter-conditional surrogate: one trained network spans the entire GALPROP parameter scan. Helps us marginalize over background systematics for the GCE.

Section 5

Conclusions

Summary

- GCE is limited by diffuse-background systematics; GALPROP is expensive.

Thanks for listening!

Summary

- GCE is limited by diffuse-background systematics; GALPROP is expensive.
- PINNs solve the CR transport equation directly: $< 1\%$ agreement with GALPROP. Cheap, more robust convergence.

Thanks for listening!

Summary

- GCE is limited by diffuse-background systematics; GALPROP is expensive.
- PINNs solve the CR transport equation directly: $< 1\%$ agreement with GALPROP. Cheap, more robust convergence.
- Parameter-conditional PINNs replace discrete templates with one model; marginalize over CR propagation uncertainties.

Thanks for listening!

Summary

- GCE is limited by diffuse-background systematics; GALPROP is expensive.
- PINNs solve the CR transport equation directly: $< 1\%$ agreement with GALPROP. Cheap, more robust convergence.
- Parameter-conditional PINNs replace discrete templates with one model; marginalize over CR propagation uncertainties.
- Downstream: cleaner GCE spectrum/morphology and sharper DM vs. MSP tests; broader uses for $e^{\pm}$ excesses, antiprotons, secondary nuclei.

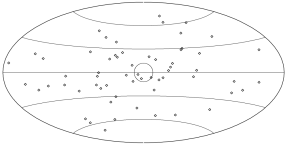
Thanks for listening!

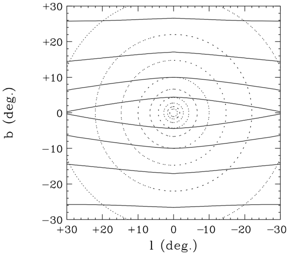
Summary

- GCE is limited by diffuse-background systematics; GALPROP is expensive.
- PINNs solve the CR transport equation directly: $< 1\%$ agreement with GALPROP. Cheap, more robust convergence.
- Parameter-conditional PINNs replace discrete templates with one model; marginalize over CR propagation uncertainties.
- Downstream: cleaner GCE spectrum/morphology and sharper DM vs. MSP tests; broader uses for $e^{\pm}$ excesses, antiprotons, secondary nuclei.
- Upshot: PINNs can accelerate or even improve special-purpose co HEP/Astronomy!

Thanks for listening!

Backup: MSP morphology



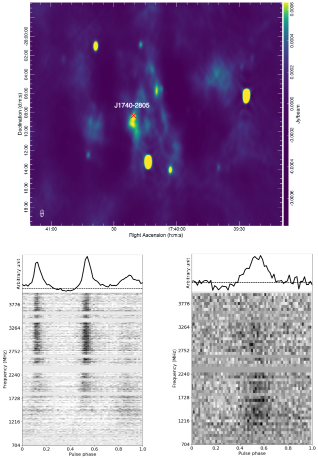


Hooper & Goodenough, arXiv:1407.5625

- A GC MSP population gives a different “boxy” morphology that traces the bulge + thick disk, rather than annihilating DM halo.

Backup: Two new bulge MSPs

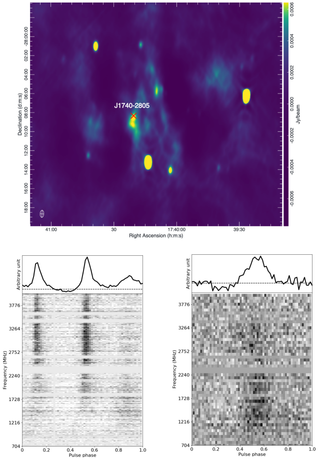
- Berteaud et al. [arXiv:2512.16699]: deep MeerKAT/GBT follow-up of X-ray candidates in the inner Galaxy.



Berteaud et al., arXiv:2512.16699

Backup: Two new bulge MSPs

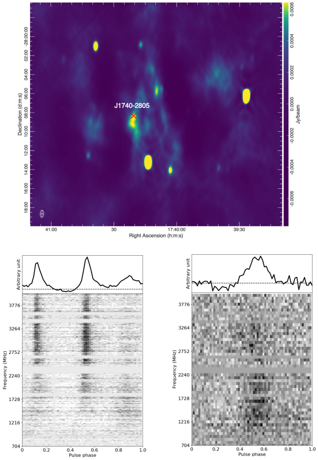
- Berteaud et al. [arXiv:2512.16699]: deep MeerKAT/GBT follow-up of X-ray candidates in the inner Galaxy.
- Two new MSPs in the bulge. 4 total now!
  - PSR J1740–2805
  - PSR J1740–28



Berteaud et al., arXiv:2512.16699

Backup: Two new bulge MSPs

- Berteaud et al. [arXiv:2512.16699]: deep MeerKAT/GBT follow-up of X-ray candidates in the inner Galaxy.
- Two new MSPs in the bulge. 4 total now!
  - PSR J1740–2805
  - PSR J1740–28
- Consistent with prediction [Hooper, Sweden 2023] for MSP detection rates in/around the bulge.



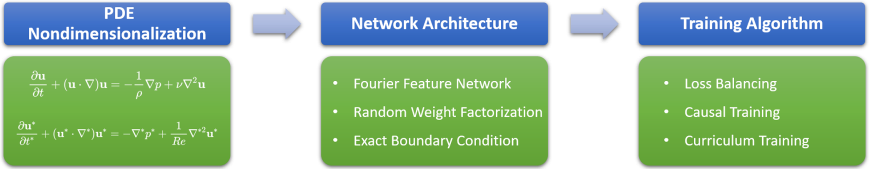
Berteaud et al., arXiv:2512.16699

Backup: Equilibrium for free

- GALPROP *time-evolves* initial conditions until $\partial_t\psi = 0$.
- Longest timescale (low-E diffusion): $\gtrsim 10^8$ yr. Shortest timescale (high-E fast processes): $\lesssim 10$ yr.
- The PINN solves the steady-state equation *directly*; no time integration needed.
- Cheaper convergence from the PINN. More robust in some scenarios.

$$\cancel{\frac{\partial\psi}{\partial t}} = q + D_{xx}\nabla^2\psi - \frac{\partial}{\partial E}(\dot{E}\,\psi)$$

Backup: Training PINNs

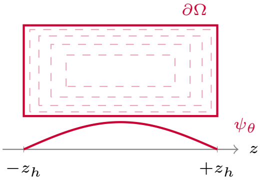


Wang, Sankaran & Perdikaris, arXiv:2308.08468

- Non-dimensionalization of fields and PDE coefficients to $\mathcal{O}(1)$.
- Fourier features; random weight factorization; modified MLP / PirateNet architectures.
- Loss balancing; curriculum learning; adaptive collocation-point sampling.

Backup: Hard-coding the boundary condition

- CR halo: $\psi = 0$ on the boundary $\partial\Omega$ (free-escape).



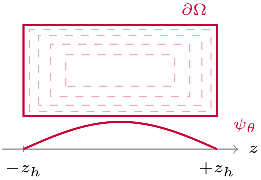
ADF level sets (dashed); ψ_θ vanishes at $\pm z_h$.

Backup: Hard-coding the boundary condition

- CR halo: $\psi = 0$ on the boundary $\partial\Omega$ (free-escape).
- Factor the network as

$$\psi_\theta(\vec{r}, E) = \phi(\vec{r})\,\mathcal{N}_\theta(\vec{r}, E),$$

ϕ an approximate distance function (ADF): smooth, positive in Ω, zero on $\partial\Omega$.



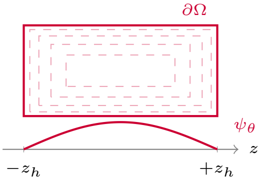
ADF level sets (dashed); ψ_θ vanishes at $\pm z_h$.

Backup: Hard-coding the boundary condition

- CR halo: $\psi = 0$ on the boundary $\partial\Omega$ (free-escape).
- Factor the network as

$$\psi_\theta(\vec{r}, E) = \phi(\vec{r})\,\mathcal{N}_\theta(\vec{r}, E),$$

  ϕ an approximate distance function (ADF): smooth, positive in Ω, zero on $\partial\Omega$.
- BC satisfied *by construction* – no boundary-penalty term to balance.



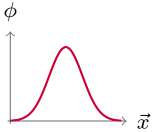
ADF level sets (dashed); ψ_θ vanishes at $\pm z_h$.

Backup: factorized PINN architecture

Spectrum spans $\sim$8 decades; morphology varies slowly. Factor the solution:

$$\log_{10}\psi(\vec{x},E) = \log_{10}\phi(\vec{x},E) + \log_{10} S(E).$$

- $S(E)$ – global spectrum, energy only.



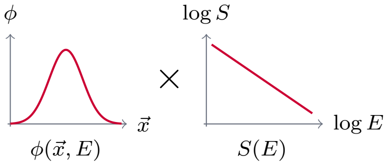
$\phi(\vec{x},E)$

Backup: factorized PINN architecture

Spectrum spans $\sim$8 decades; morphology varies slowly. Factor the solution:

$$\log_{10}\psi(\vec{x},E) = \log_{10}\phi(\vec{x},E) + \log_{10} S(E).$$

- $S(E)$ – global spectrum, energy only.
- $\phi(\vec{x},E)$ – morphology, may carry residual E dependence.



Backup: factorized PINN architecture

Spectrum spans $\sim$8 decades; morphology varies slowly. Factor the solution:

$$\log_{10}\psi(\vec{x},E) = \log_{10}\phi(\vec{x},E) + \log_{10} S(E).$$

- $S(E)$ – global spectrum, energy only.
- $\phi(\vec{x},E)$ – morphology, may carry residual E dependence.
- Two PINNs trained jointly $\Rightarrow$ sub-percent agreement across all 8 decades.