Accelerating PBH Phenomenology via Neural Operators

Tae-Geun Kim (金泰瑾)

collaborated with H. Jung[1], M. G. Park[1], J. Park[1], S. C. Park[1] and Y. Park[2]

1. *Yonsei University*
2. *Institute of Basic Science*

The 2nd AI+HEP in East Asia Workshop

2026.01.22

Primordial Black Hole

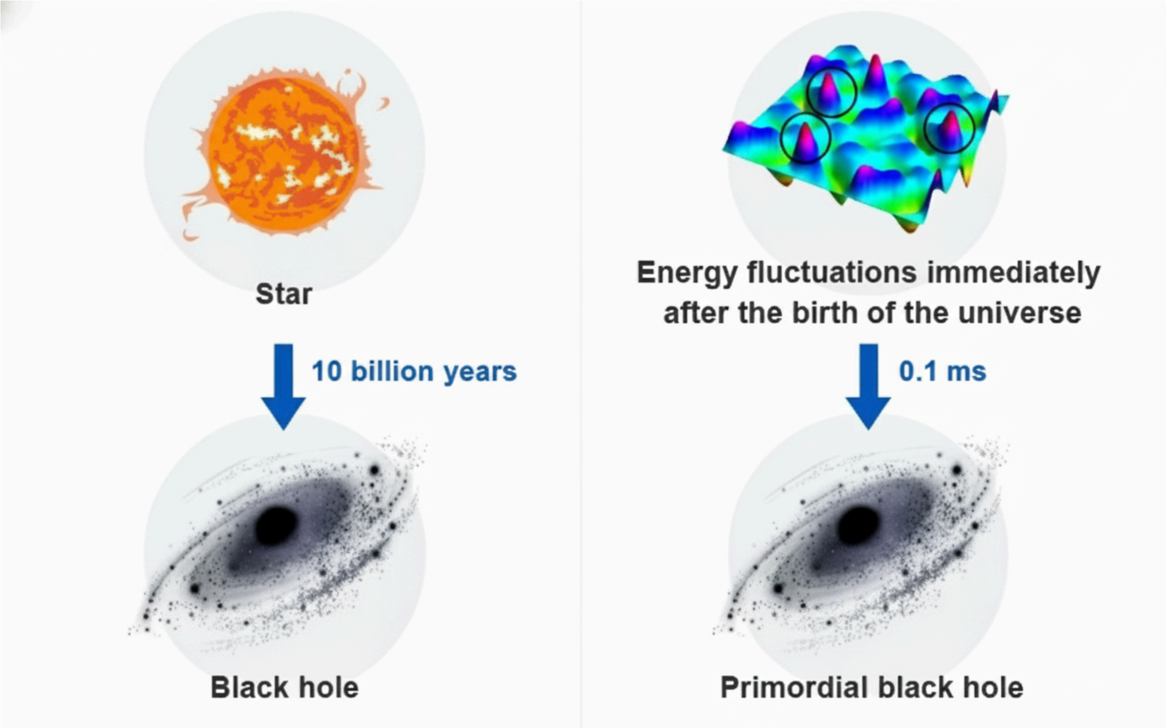


Figure 1: Differences in the formulation of black holes and primordial black holes [Science Tokyo]

- **Origin**: Formed from gravitational collapse of large primordial density fluctuations
- **Dark Matter Candidate**: A promising contender for explaining dark matter

Hawking Radiation

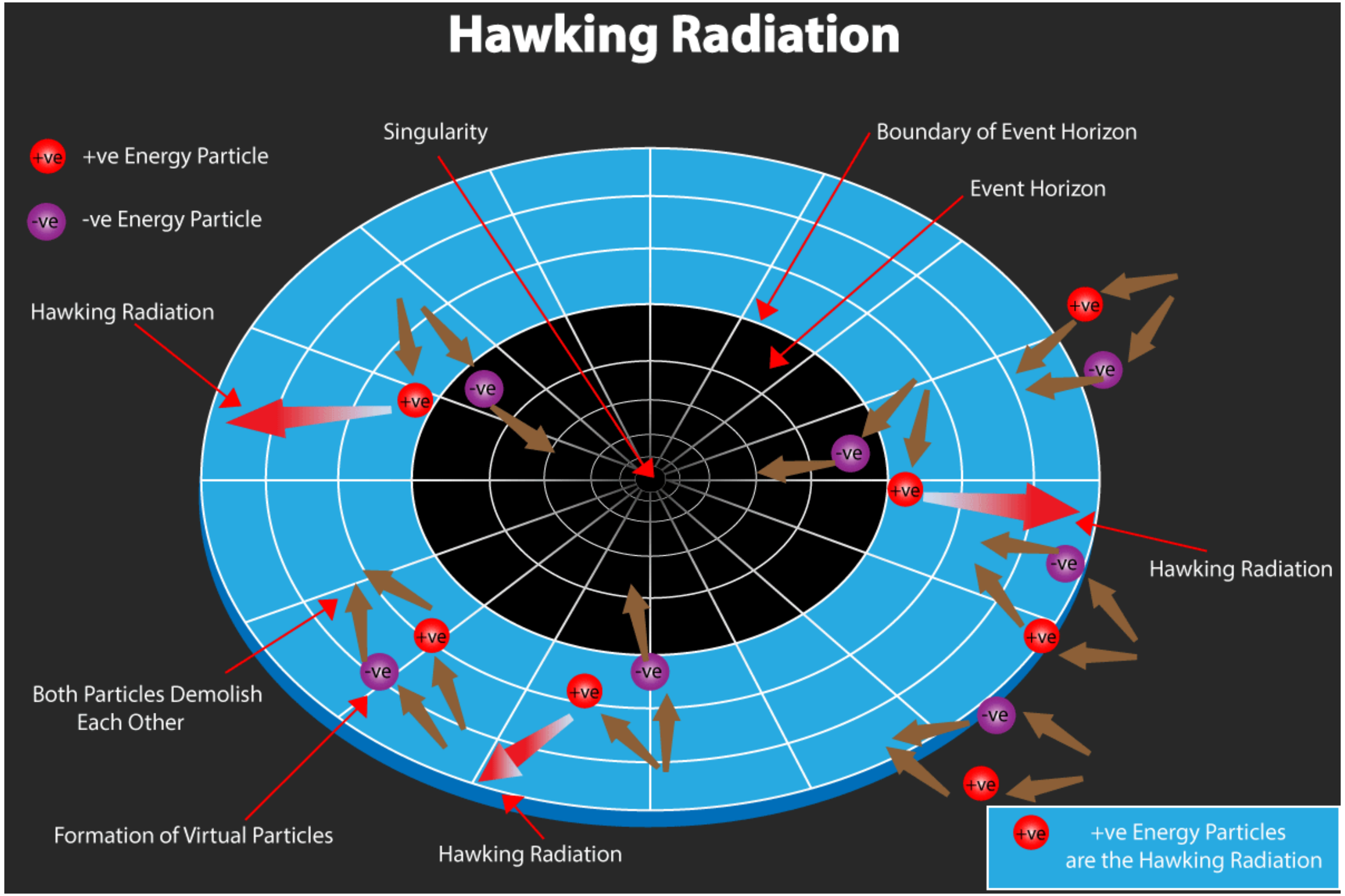


Figure 2: Illustration of Hawking radiation [Physics feed]

PBHs as Particle Factories

- **Hawking temperature of PBH**

[J. D. Beckenstein, PRD (1973), S. Hawking, Comm. Math. Phys. (1975) & PRD (1976)]

$$k_{\mathrm{B}} T_{\mathrm{PBH}} = \frac{\hbar c^3}{8\pi G M_{\mathrm{PBH}}} \sim 1.06 \left(\frac{10^{16}\mathrm{g}}{M_{\mathrm{PBH}}} \right) \mathrm{MeV} \sim 10^{10} \left(\frac{10^{16}\mathrm{g}}{M_{\mathrm{PBH}}} \right) \mathrm{K}$$

- **Emission rates of particle χ**

[A. A. Starobinsky, Sov.Phys.JETP (1973), S. A. Teukolsky (1974), D. N. Page, ApJ (1976)]
[D. Ida, K. Oda and S. C. Park, PRD (2003, 2004)]

$$\frac{\mathrm{d}^2 N_\chi}{\mathrm{d}E \mathrm{d}t} = \frac{g_\chi}{2\pi} \frac{\Gamma(E, M_{\mathrm{PBH}})}{e^{E/k_{\mathrm{B}} T_{\mathrm{PBH}}} - (-1)^{2 s_\chi}}$$

  - g_χ: Degree of freedom of χ
  - s_χ: Spin of χ
  - $\Gamma(E, M_{\mathrm{PBH}})$: Greybody factor

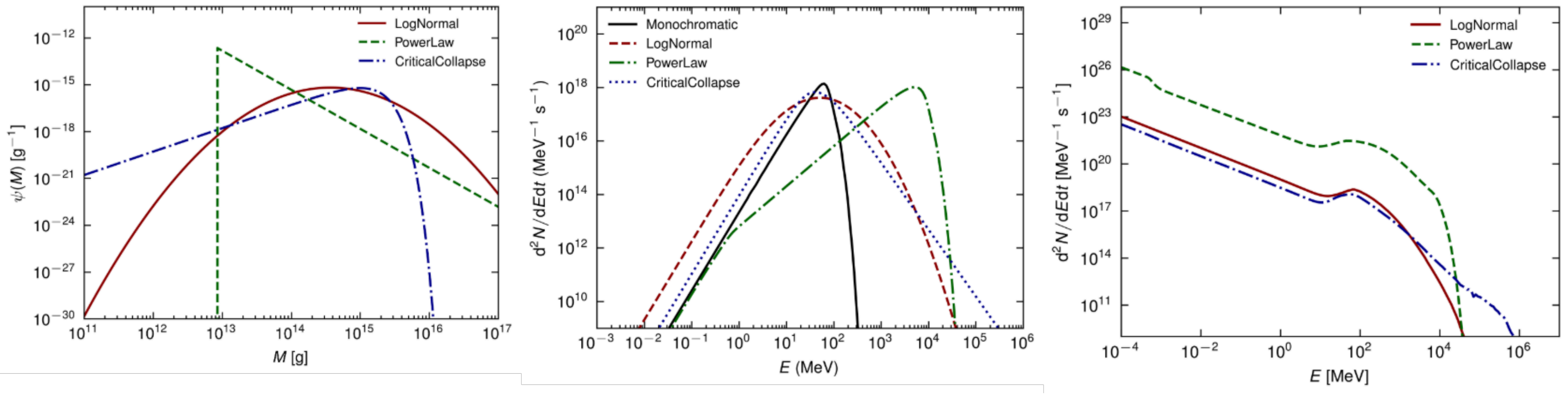
- **Lifetime of PBH**

[D. N. Page, PRD (1976)]

$$\tau_{\mathrm{PBH}} \sim 13.8 \times 10^9 \ \text{years} \left(\frac{M_{\mathrm{PBH}}}{5.1 \times 10^{14}\mathrm{g}} \right)^3$$

- **Particle Factory**: Can be a source of various particle emissions through Hawking radiation

Computation of Hawking Radiation



Secondary particles produced by the decay and hadronization of primary particles

Secondary Spectrum & BlackHawk

- **Emission rate of secondary photons**

[A. Arbey and J. Auffinger, EPJC (2019), A. Coogan et al., PRL (2021)]

$$\frac{\mathrm{d}^2 N_j^{\mathrm{sec}}}{\mathrm{d}E\mathrm{d}t} = \int_0^{+\infty} \sum_i \frac{\mathrm{d}^2 N_i^{\mathrm{pri}}}{\mathrm{d}E'\mathrm{d}t} \frac{\mathrm{d}N_j^i}{\mathrm{d}E} \mathrm{d}E'$$

$$\frac{\mathrm{d}^2 N_\gamma^{\mathrm{sec}}}{\mathrm{d}E_\gamma\mathrm{d}t} = \sum_{i=e^\pm,\mu^\pm,\pi^\pm} \int \mathrm{d}E_i \left(\frac{\mathrm{d}^2 N_i^{\mathrm{pri}}}{\mathrm{d}E_i\mathrm{d}t}\right) \frac{\mathrm{d}N_i^{\mathrm{FSR}}}{\mathrm{d}E_\gamma} + \sum_{i=\mu^\pm,\pi^0,\pi^\pm} \int \mathrm{d}E_i \left(\frac{\mathrm{d}^2 N_i^{\mathrm{pri}}}{\mathrm{d}E_i\mathrm{d}t}\right) \frac{\mathrm{d}N_i^{\mathrm{decay}}}{\mathrm{d}E_\gamma}$$

- **How to compute the secondary spectrum**
  - For low energy ($\lesssim 5\ \mathrm{GeV}$) : use `HAZMA` [A. Coogan et al., JCAP (2020)]
  - For mid energy ($5\ \mathrm{GeV} \le E \lesssim 10\mathrm{TeV}$) : use `PYTHIA` [T. Sjöstrand et al., Comput. Phys. Commun. (2008)]
  - For high energy ($\gtrsim 10\mathrm{TeV}$) : use `HDMSpectra` [C. W. Bauer et al., JHEP (2021)]

$\Rightarrow$ **BlackHawk**

[A. Arbey & J. Auffinger, EPJC (2021)]

Challenges in Computing Secondary Spectra

[A. Coogan et al., PRL (2021)]

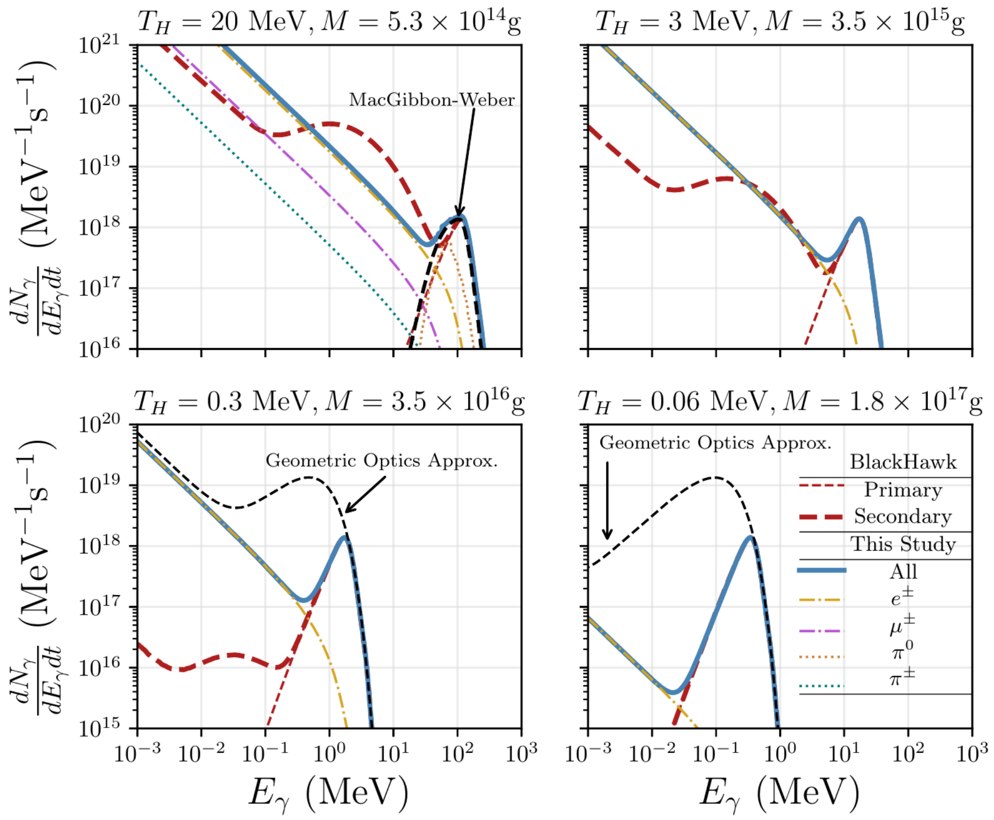


Figure 3: Comparison of the emission rates computed by HAZMA and PYTHIA

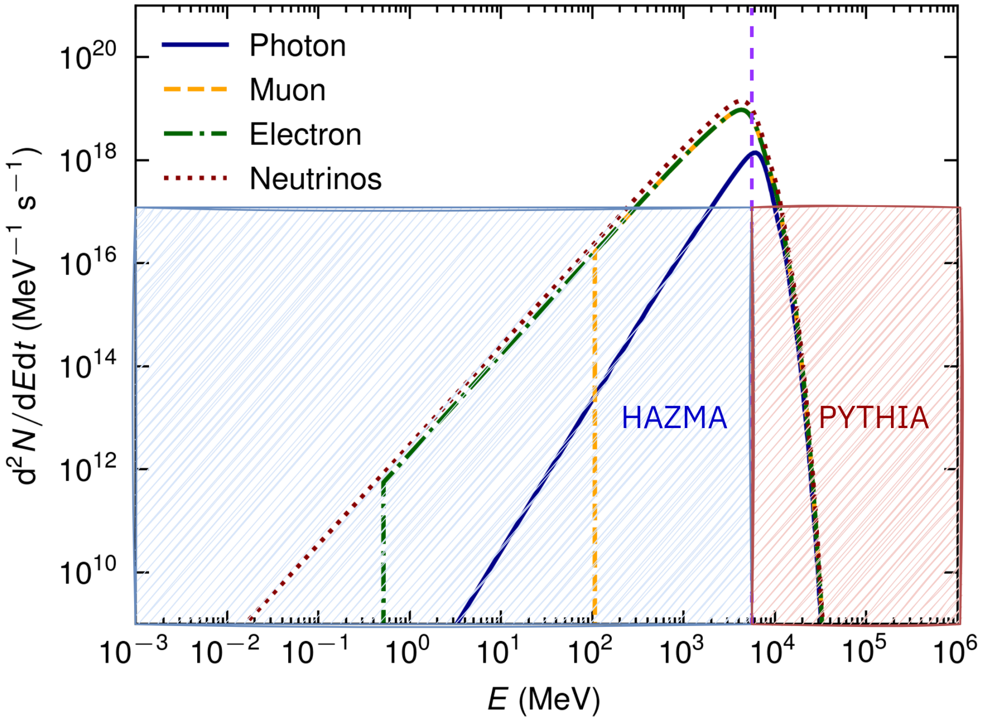
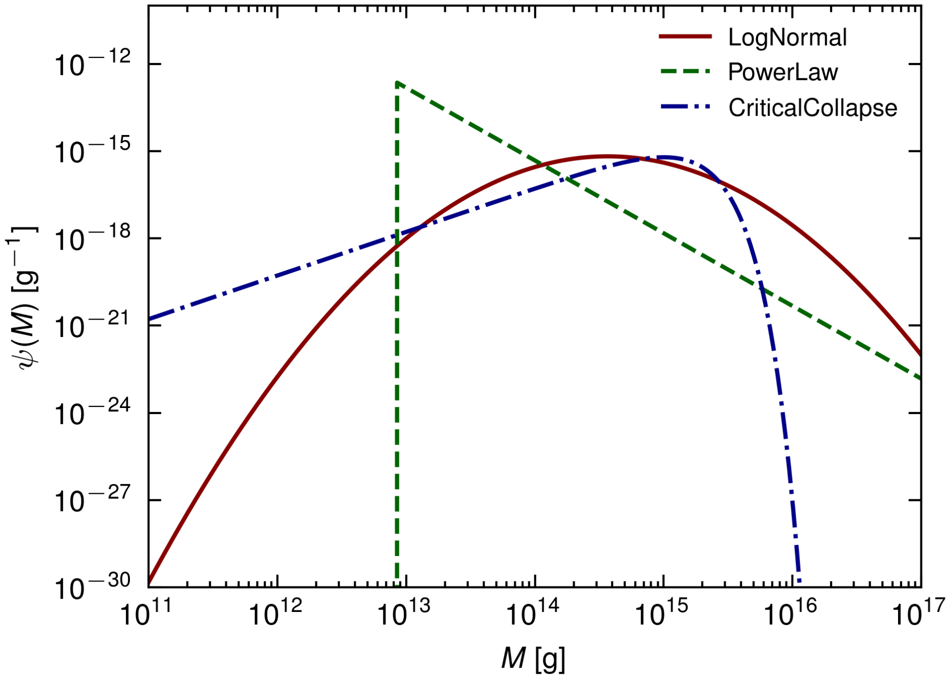


Figure 4: The primary spectrum for $M_{\rm PBH} = 10^{13}$g intersects the ROI from Hazma and Pythia.

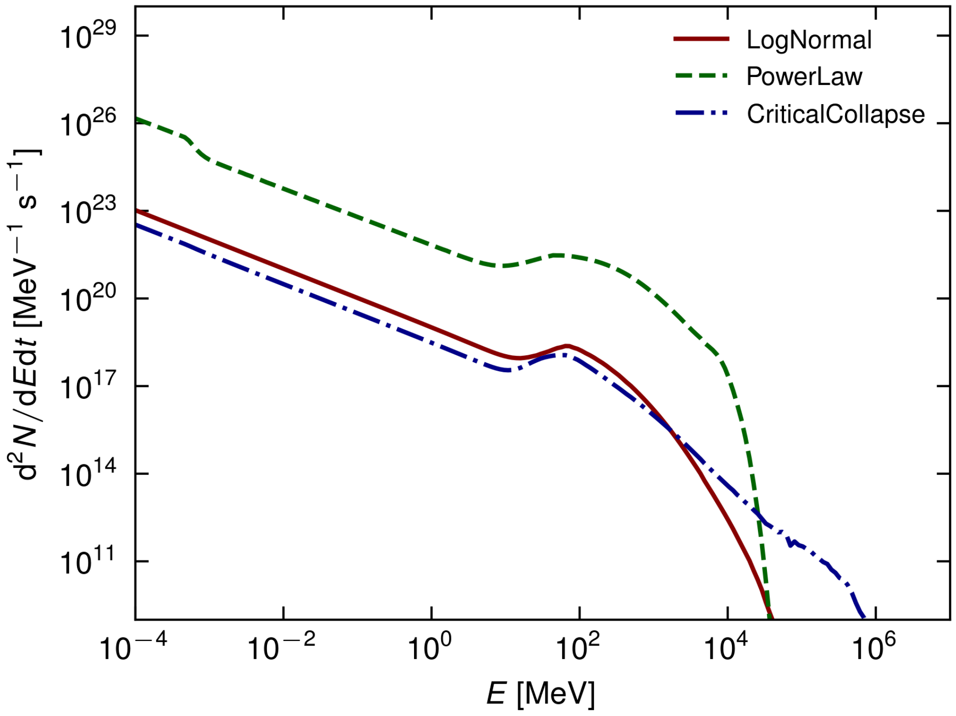
Q1: Can't we directly obtain secondary spectrum?

$$\psi(M)$$



?

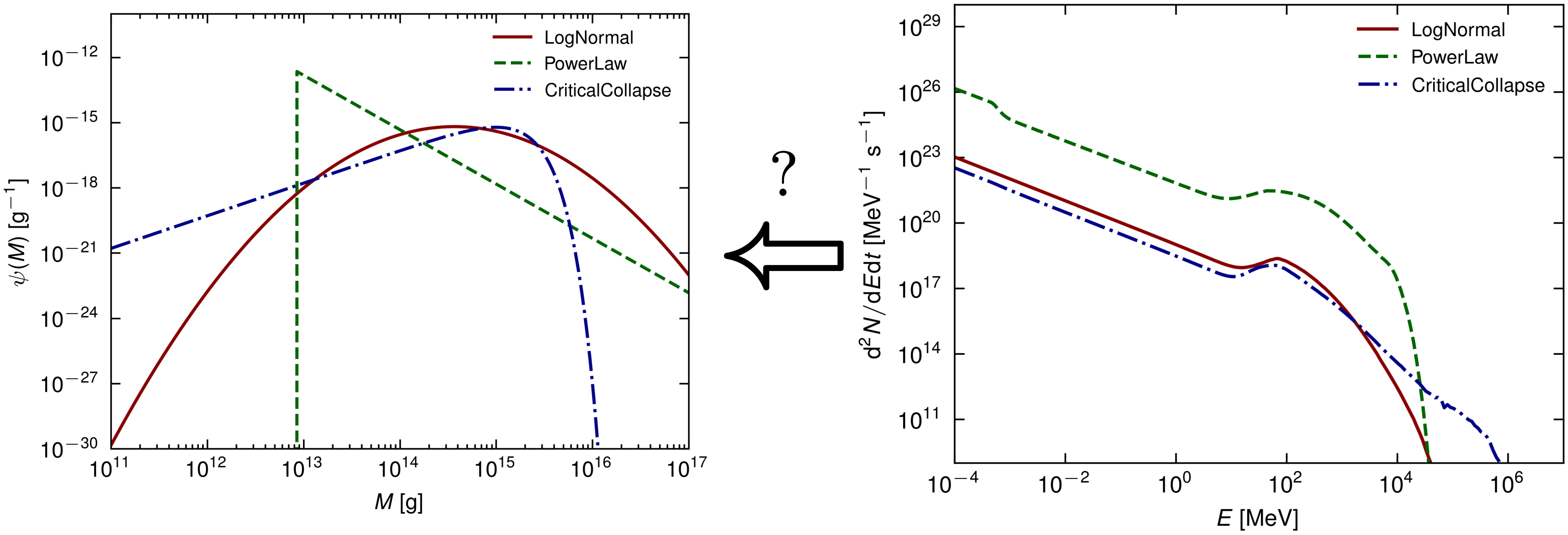
$$\frac{\mathrm{d}^2 N_\gamma^{\mathrm{total}}}{\mathrm{d}E_\gamma \mathrm{d}t}$$



Q2: How about an inverse direction?

$$\psi(M)$$

$$\frac{\mathrm{d}^2 N_\gamma^{\mathrm{total}}}{\mathrm{d}E_\gamma \mathrm{d}t}$$



Operator Formulation of Hawking Radiation

- Total photon flux is defined by convolution of the single secondary photon flux and the mass function:

$$\left(\frac{\mathrm{d}^2 N_\gamma^{\mathrm{tot}}}{\mathrm{d}E\mathrm{d}t}\right)_\psi = \int_{M_{\mathrm{min}}}^{M_{\mathrm{max}}} \frac{\mathrm{d}^2 N_\gamma^{\mathrm{sec}}}{\mathrm{d}E\mathrm{d}t}\psi(M)\mathrm{d}M$$

$$\int_{M_{\mathrm{min}}}^{M_{\mathrm{max}}} \psi(M)\mathrm{d}M = 1$$

- If we fix M_{min} and M_{max}, then this can be expresses as the linear operator:

$$\mathfrak{H} : \psi(M) \xrightarrow{\mathfrak{H}} \left(\frac{\mathrm{d}^2 N_\gamma^{\mathrm{tot}}}{\mathrm{d}E\mathrm{d}t}\right)_\psi$$

Operator Formulation of Hawking Radiation

- Total photon flux is defined by convolution of the single secondary photon flux and the mass function:

$$\left(\frac{\mathrm{d}^2 N_\gamma^{\mathrm{tot}}}{\mathrm{d}E\mathrm{d}t}\right)_\psi = \int_{M_{\min}}^{M_{\max}} \frac{\mathrm{d}^2 N_\gamma^{\mathrm{sec}}}{\mathrm{d}E\mathrm{d}t}\psi(M)\mathrm{d}M$$

$$\int_{M_{\min}}^{M_{\max}} \psi(M)\mathrm{d}M = 1$$

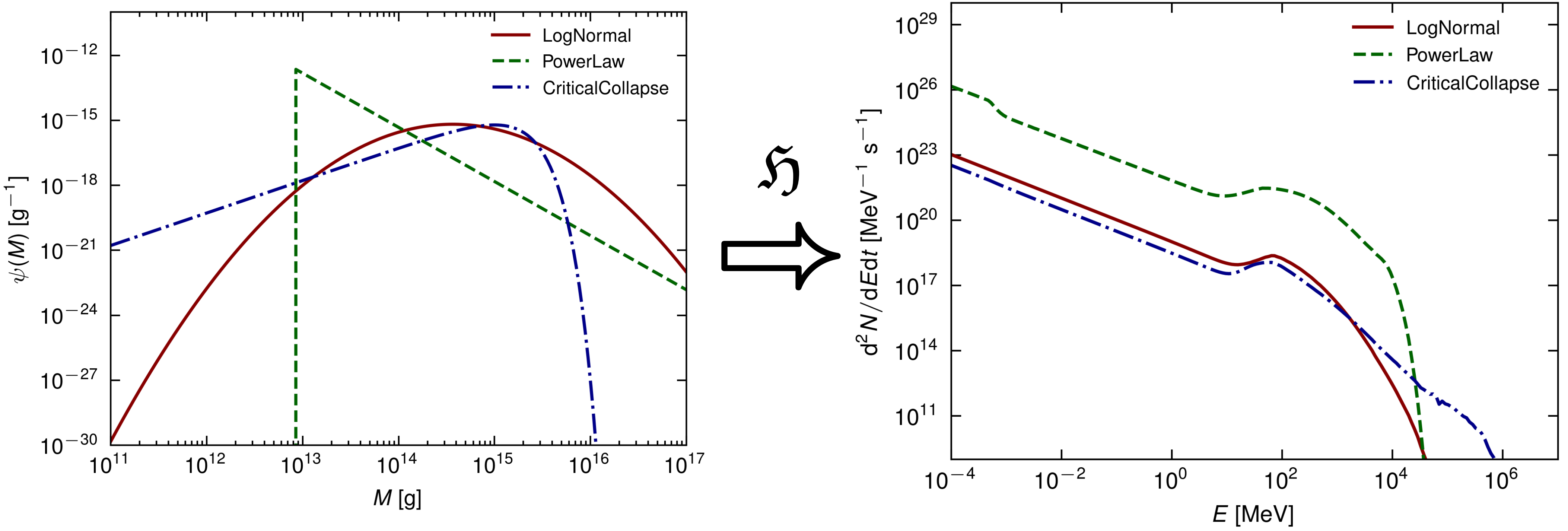
- If we fix $M_{\min}$ and $M_{\max}$, then this can be expresses as the linear operator:

$$\mathfrak{H} : \psi(M) \xrightarrow{\mathfrak{H}} \left(\frac{\mathrm{d}^2 N_\gamma^{\mathrm{tot}}}{\mathrm{d}E\mathrm{d}t}\right)_\psi$$

We call this operator the **Hawking Operator**.

Hawking Operator

$$\psi(M) \qquad \xRightarrow{\mathfrak{H}} \qquad \frac{\mathrm{d}^2 N_\gamma^{\mathrm{total}}}{\mathrm{d}E_\gamma \mathrm{d}t}$$

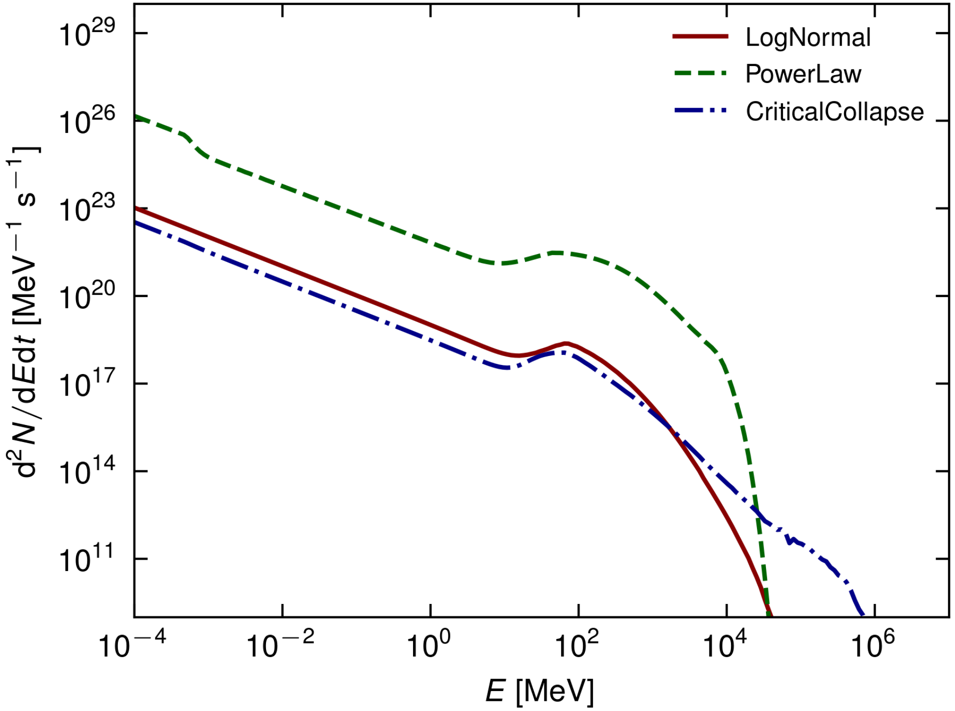


Inverse Hawking Operator

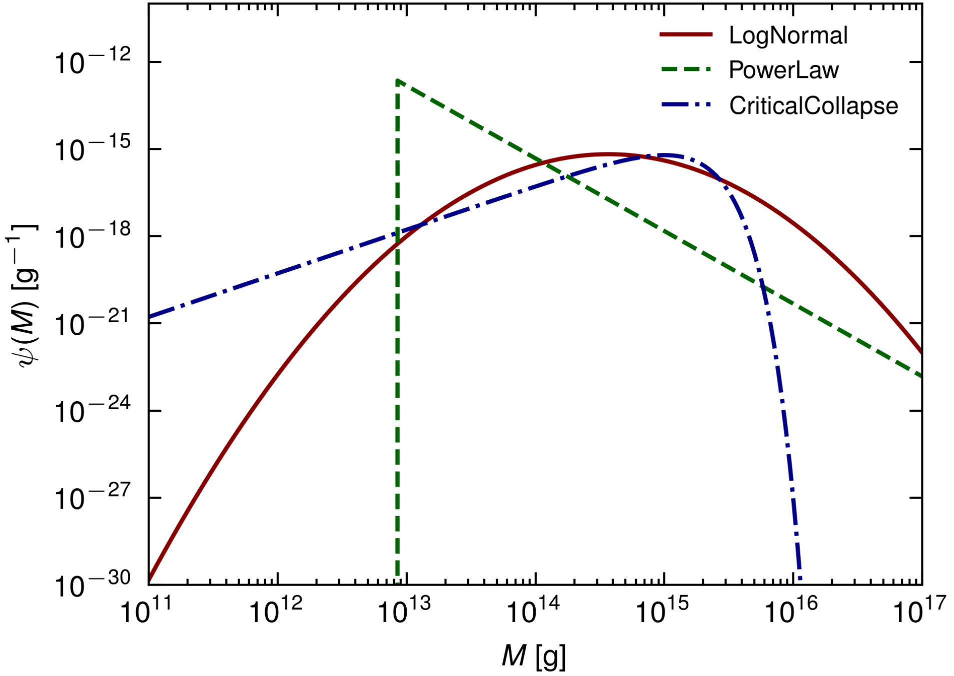
Is it well-defined?

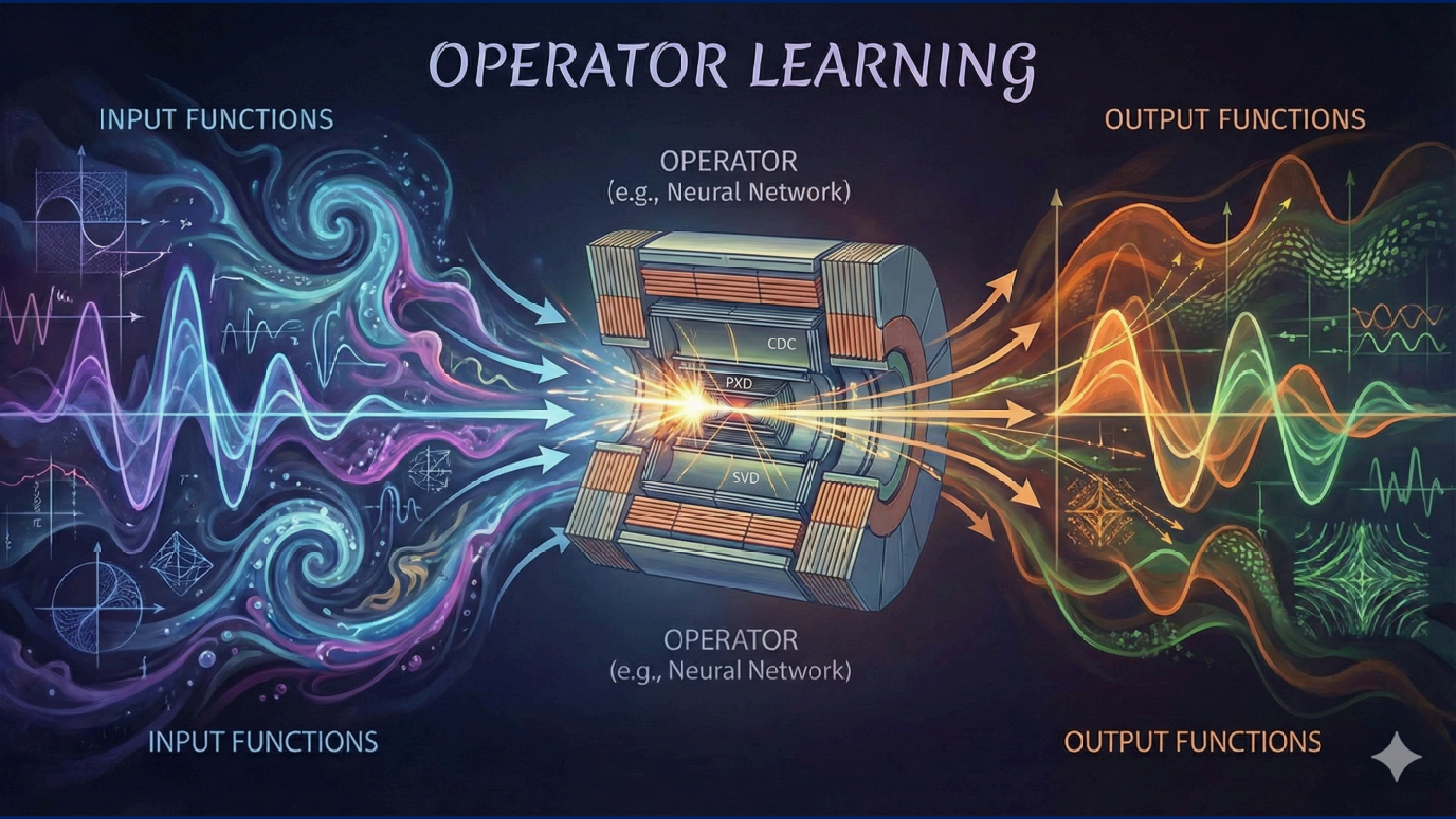
$$\frac{\mathrm{d}^2 N_\gamma^{\mathrm{total}}}{\mathrm{d}E_\gamma \mathrm{d}t}$$

$$\psi(M)$$



$$\mathfrak{H}^{-1} \Longrightarrow$$




OPERATOR LEARNING
INPUT FUNCTIONS
OUTPUT FUNCTIONS
OPERATOR
(e.g., Neural Network)
CDC
PXD
SVD
OPERATOR
(e.g., Neural Network)
INPUT FUNCTIONS
OUTPUT FUNCTIONS

Two Pillars of Operator Learning

Article | Published: 18 March 2021

Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators

Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang & George Em Karniadakis

Nature Machine Intelligence **3**, 218–229 (2021) | Cite this article

59k Accesses | **2235** Citations | **193** Altmetric | Metrics

Deep Operator Network (DeepONet)

[L. Lu et al., Nat. Mach. Intell. (2021)]

Fourier Neural Operator (FNO)

[Z. Li et al., ICLR (2021)]

Fourier Neural Operator For Parametric PDEs

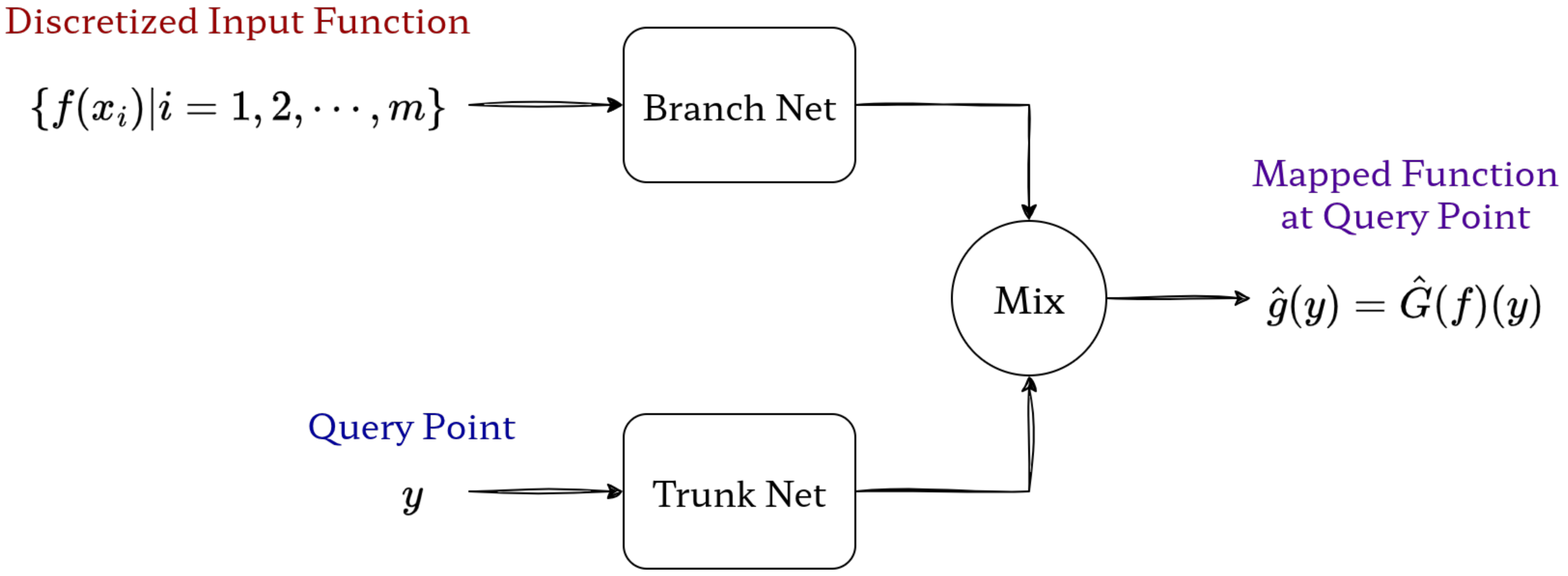
ICLR 2021

Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, Anima Anandkumar

Caltech

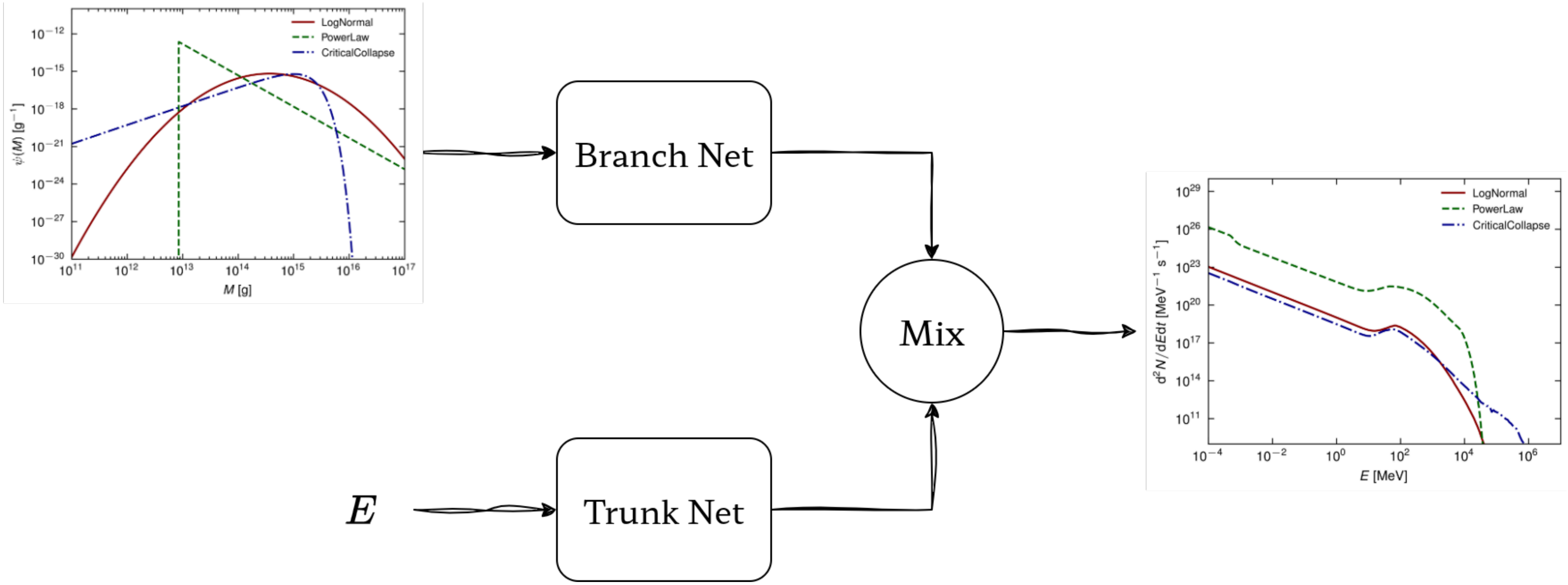
DeepONet Architecture

- Consider an operator $G : \mathcal{F} \to \mathcal{G}$, where $f(x) \in \mathcal{F}$ and $g(y) \in \mathcal{G}$ are functions.



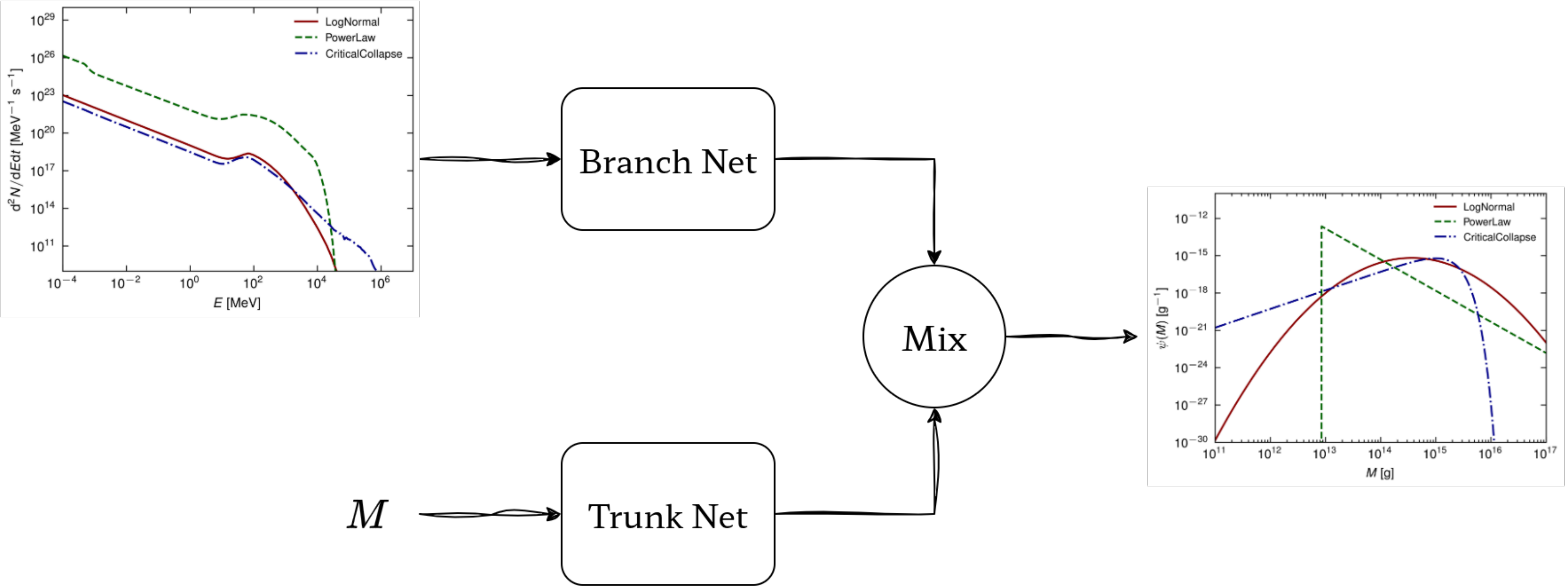
Neural Hawking Operator

- The Hawking operator is defined as $\mathfrak{H} : \psi(M) \mapsto \mathrm{d}^2N/\mathrm{d}E\mathrm{d}t$



Neural Inverse Hawking Operator

- The Inverse Hawking operator is defined as $\mathfrak{H}^{-1} : \mathrm{d}^2N/\mathrm{d}E\mathrm{d}t \mapsto \psi(M)$

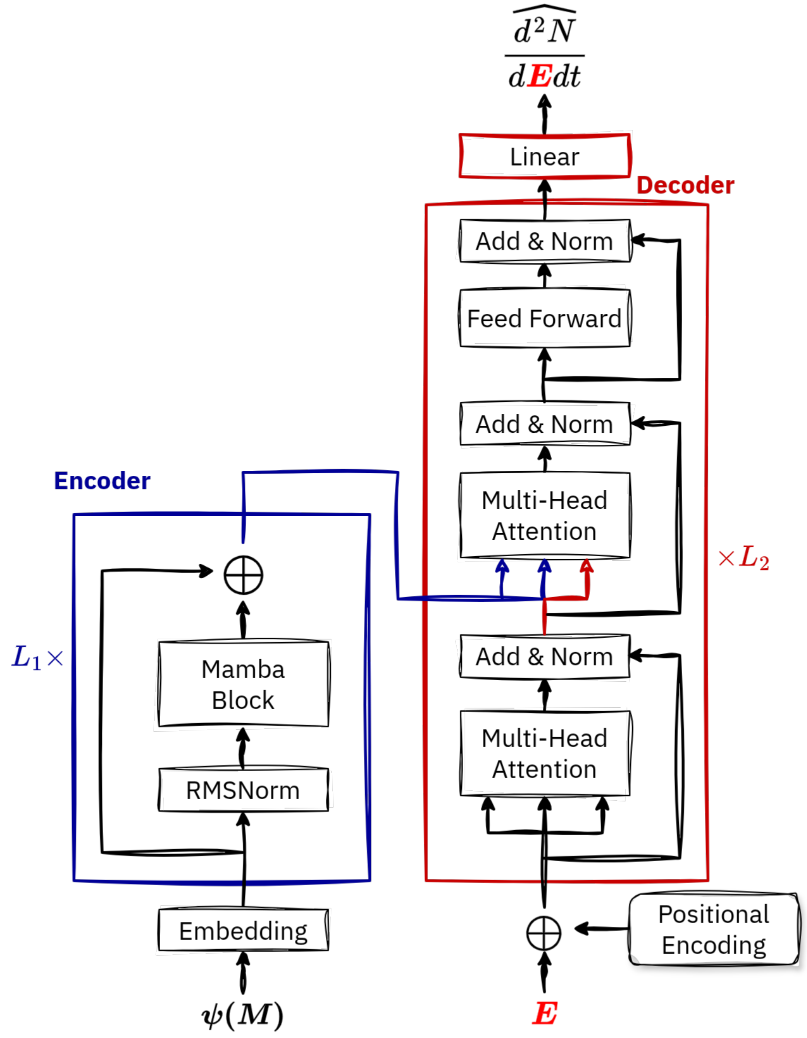


More Precise: *MambONet*

MambONet (Mamba Operator Network)

: Mamba + Multi-Head Attention + Transformer

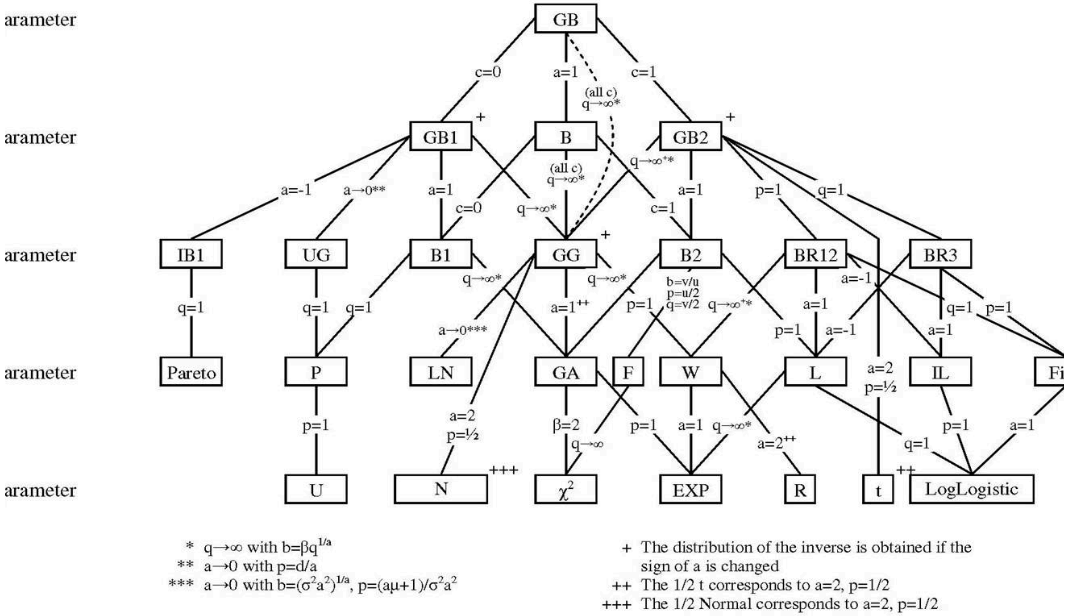
[**T.-G. Kim** & S. C. Park, arXiv:2410.20951]

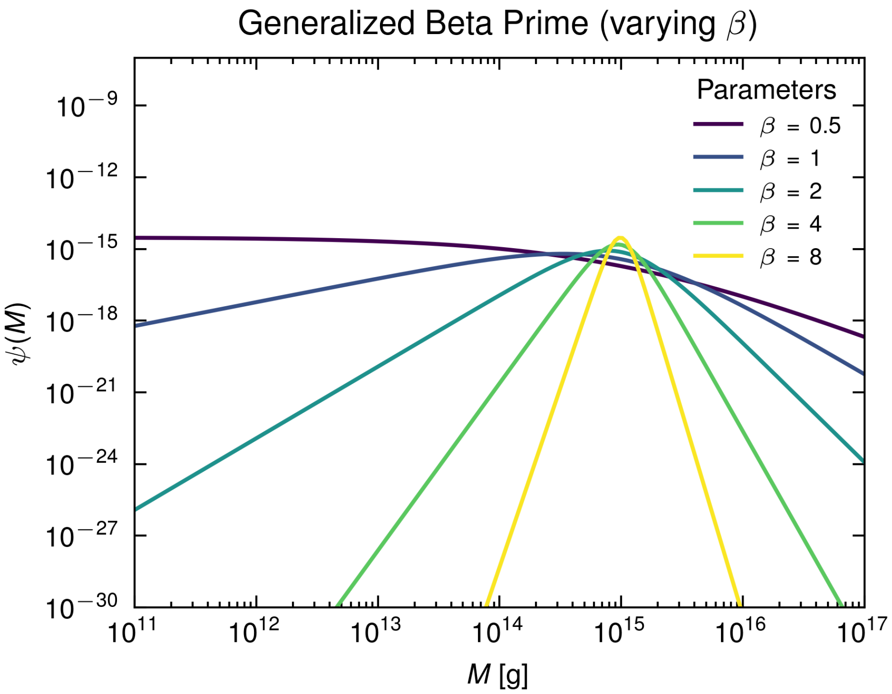


Prepare Data - *Generalized Beta Prime Distribution*

- **Generalized Beta Prime Distribution** for generating various PBH mass functions: [McDonald et al. (1995)]

$$\psi(M|M_s, \alpha, \beta, \gamma) = \frac{\beta}{M_s B(\alpha, \gamma)} \left(\frac{M}{M_s}\right)^{\alpha\beta - 1} \left(1 + \left(\frac{M}{M_s}\right)^{\beta}\right)^{-(\alpha+\gamma)}$$

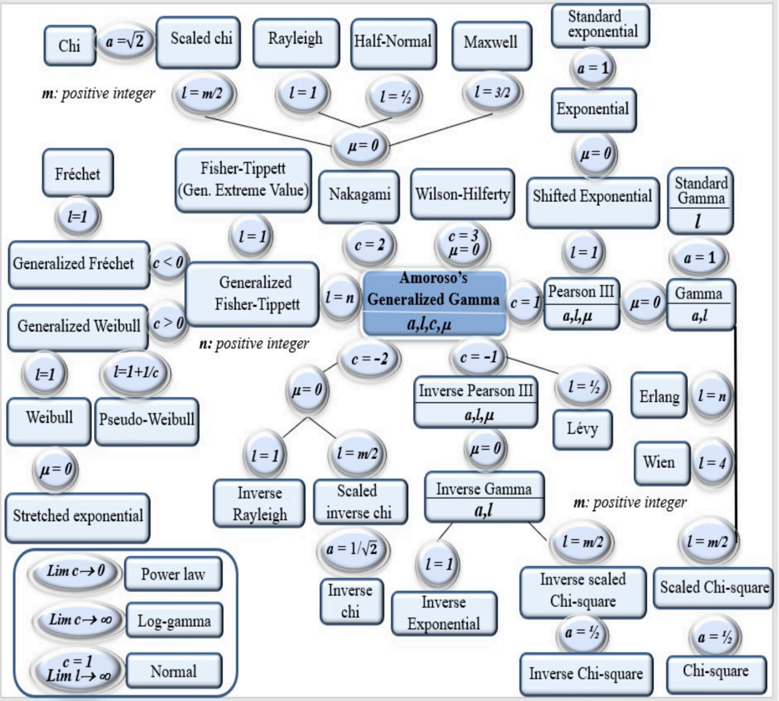


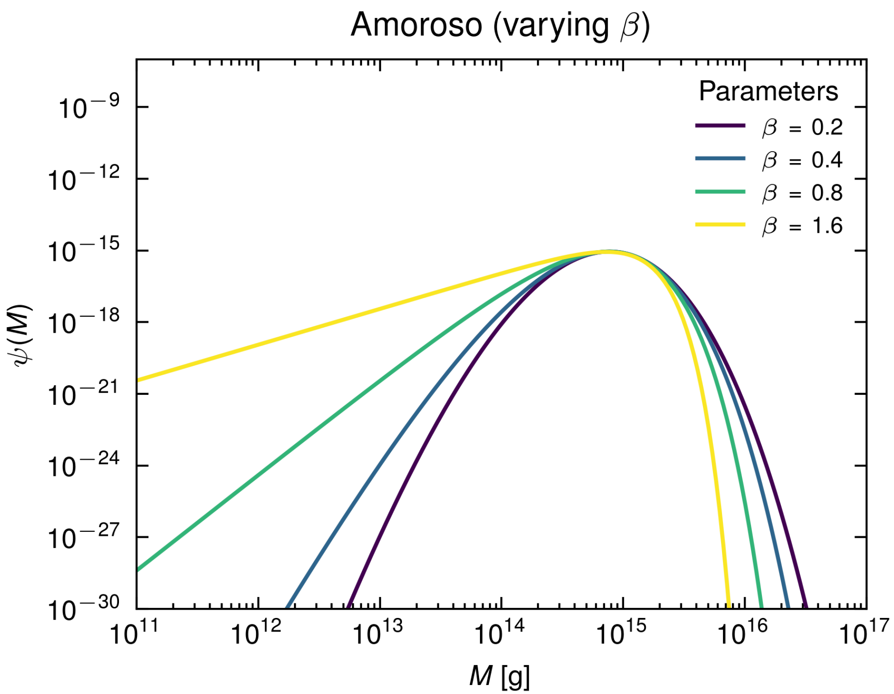


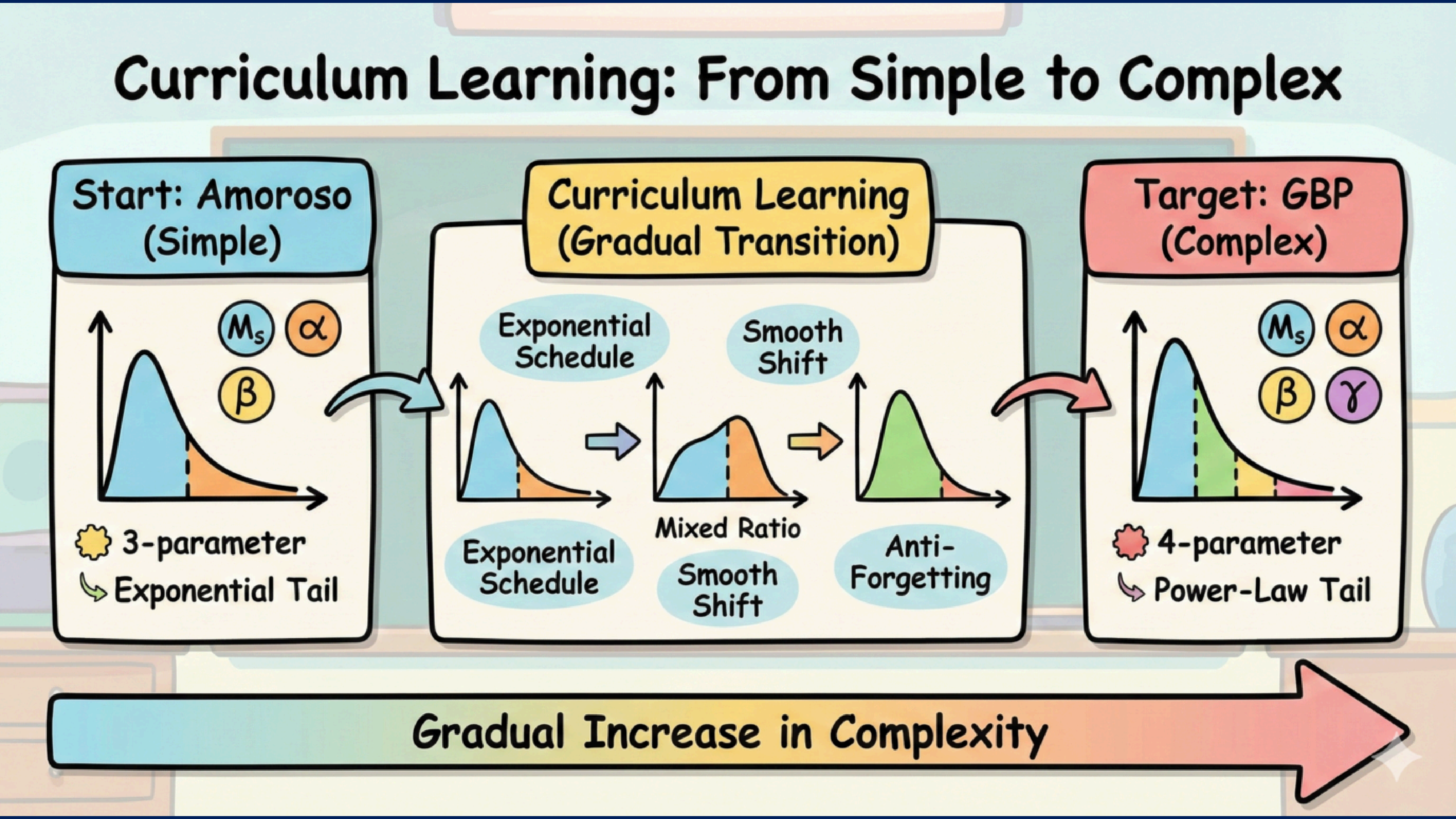
Prepare Data - *Amoroso Distribution*

- **Amoroso distribution** for generating various PBH mass functions: [G. E. Crooks (2010); C. Combes et al., MCS (2022)]

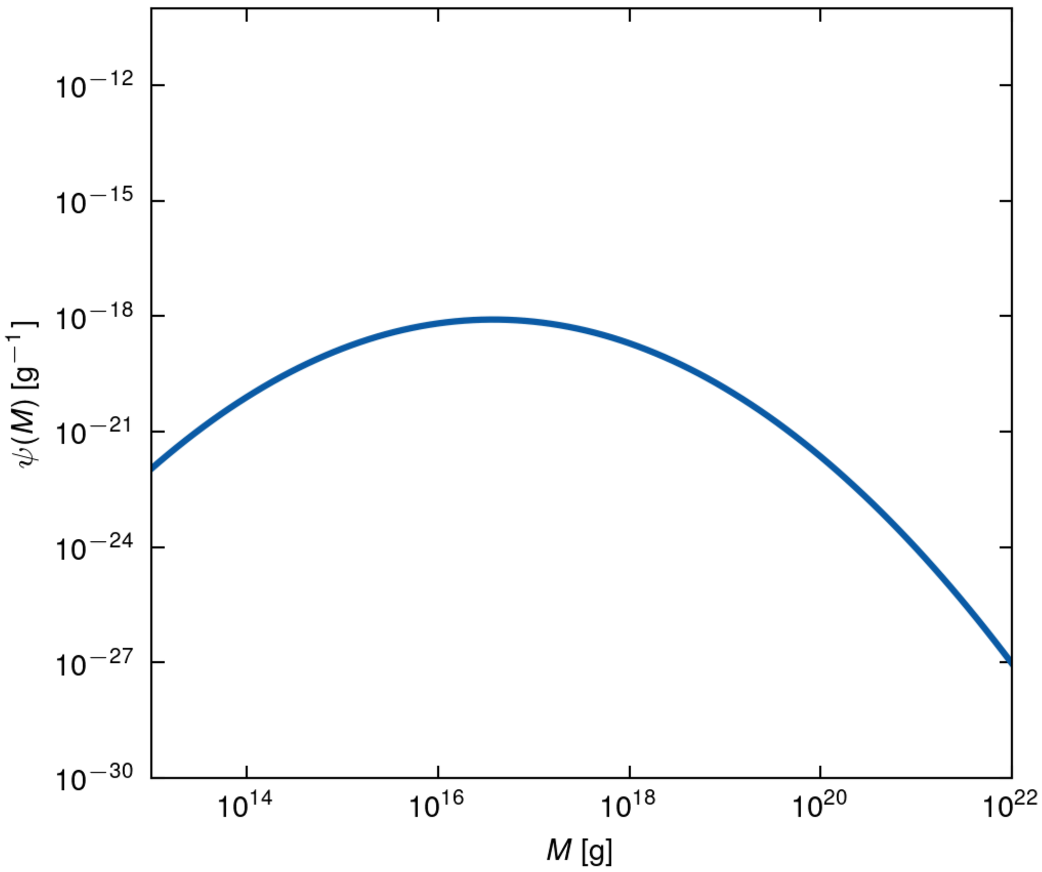
$$\psi(M|M_s,\alpha,\beta) = \frac{1}{\Gamma(\alpha)}\frac{|\beta|}{M_s}\left(\frac{M}{M_s}\right)^{\alpha\beta-1}\exp\left(-\left(\frac{M}{M_s}\right)^{\beta}\right)$$

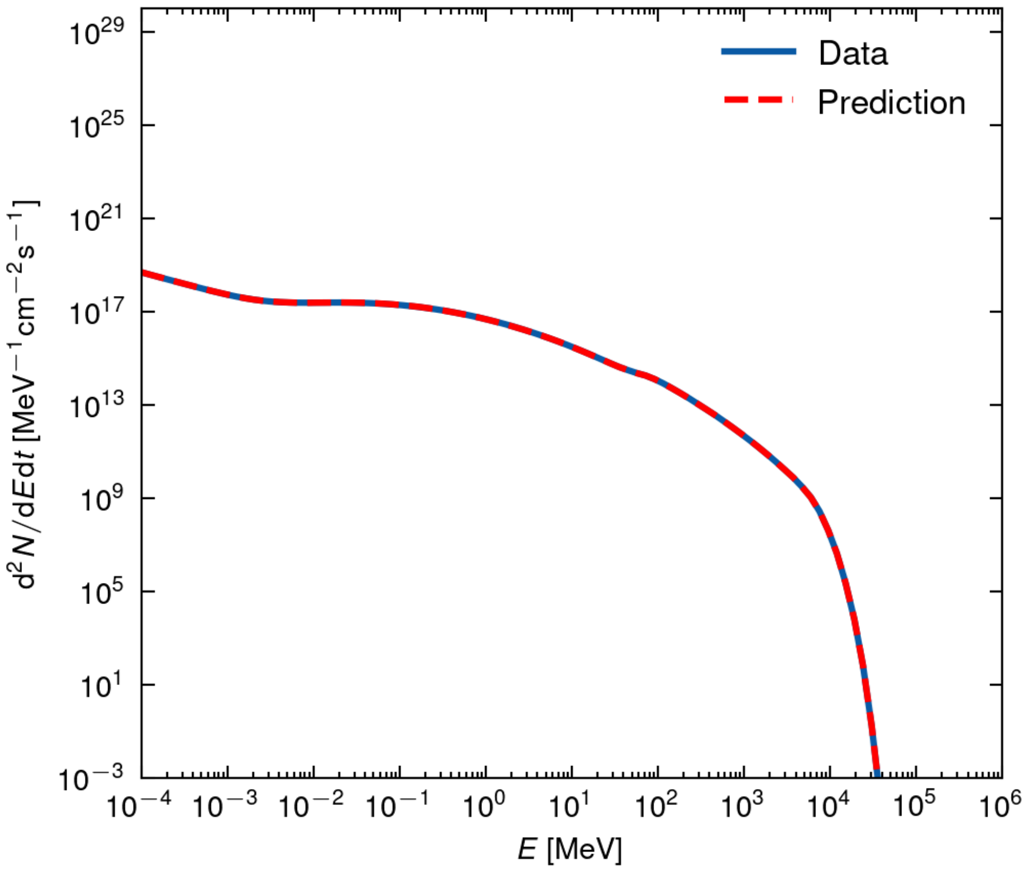




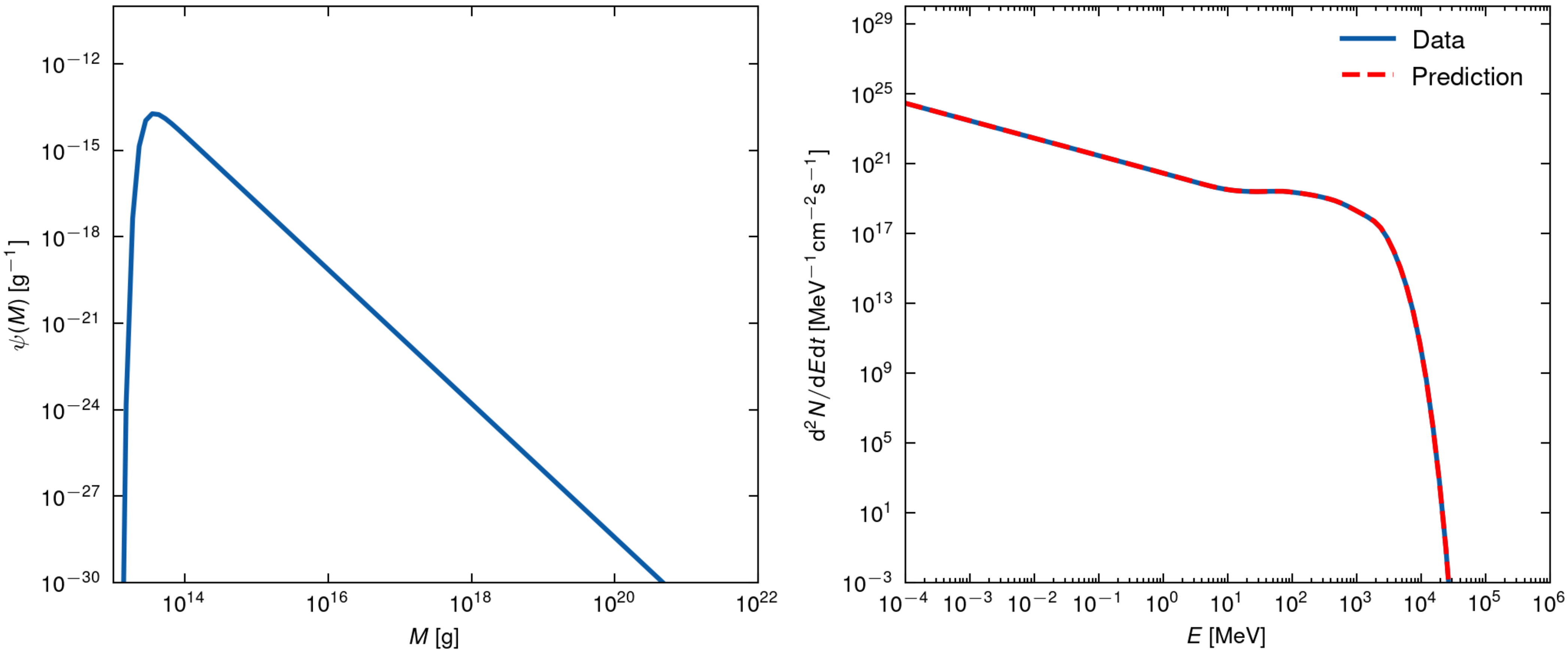

Curriculum Learning: From Simple to Complex
Start: Amoroso (Simple)
M_s
α
β
3-parameter
Exponential Tail
Curriculum Learning (Gradual Transition)
Exponential Schedule
Smooth Shift
Mixed Ratio
Exponential Schedule
Smooth Shift
Anti-Forgetting
Target: GBP (Complex)
M_s
α
β
γ
4-parameter
Power-Law Tail
Gradual Increase in Complexity

Results - *Log Normal*

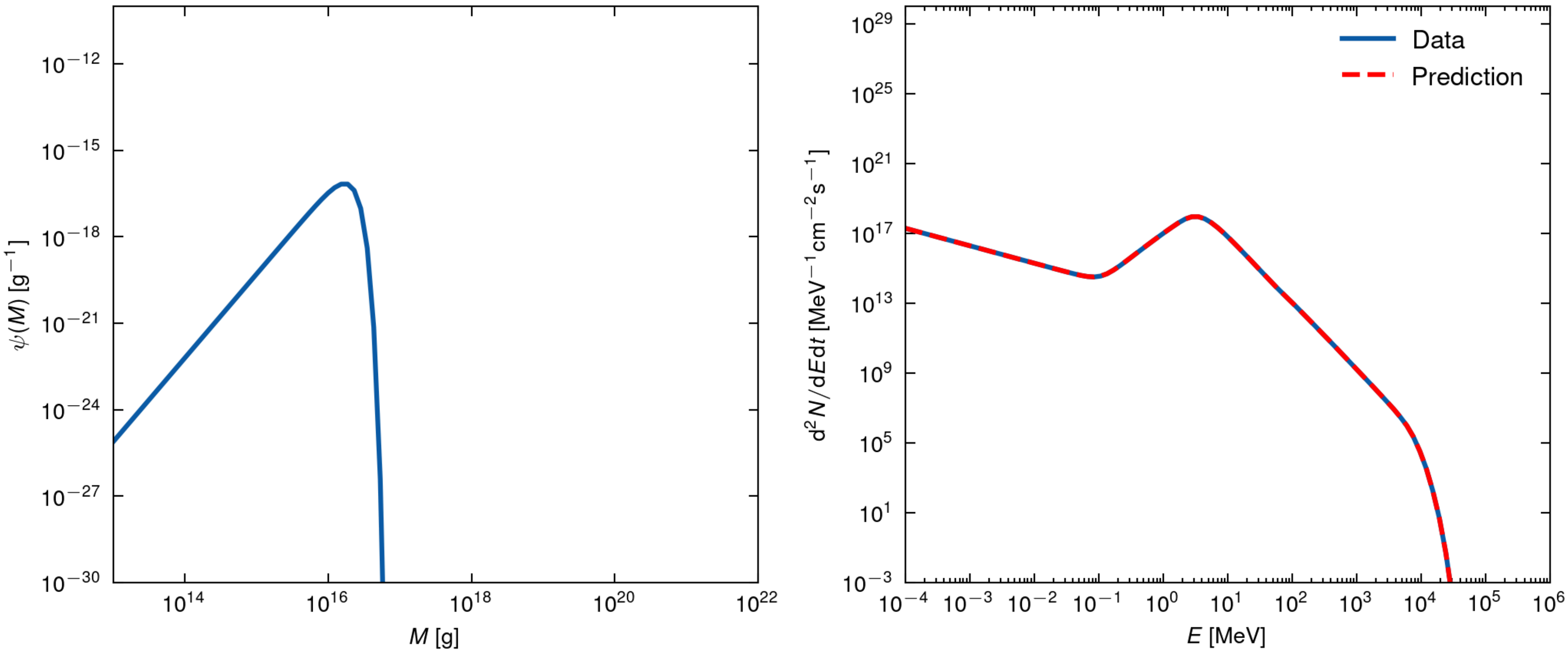




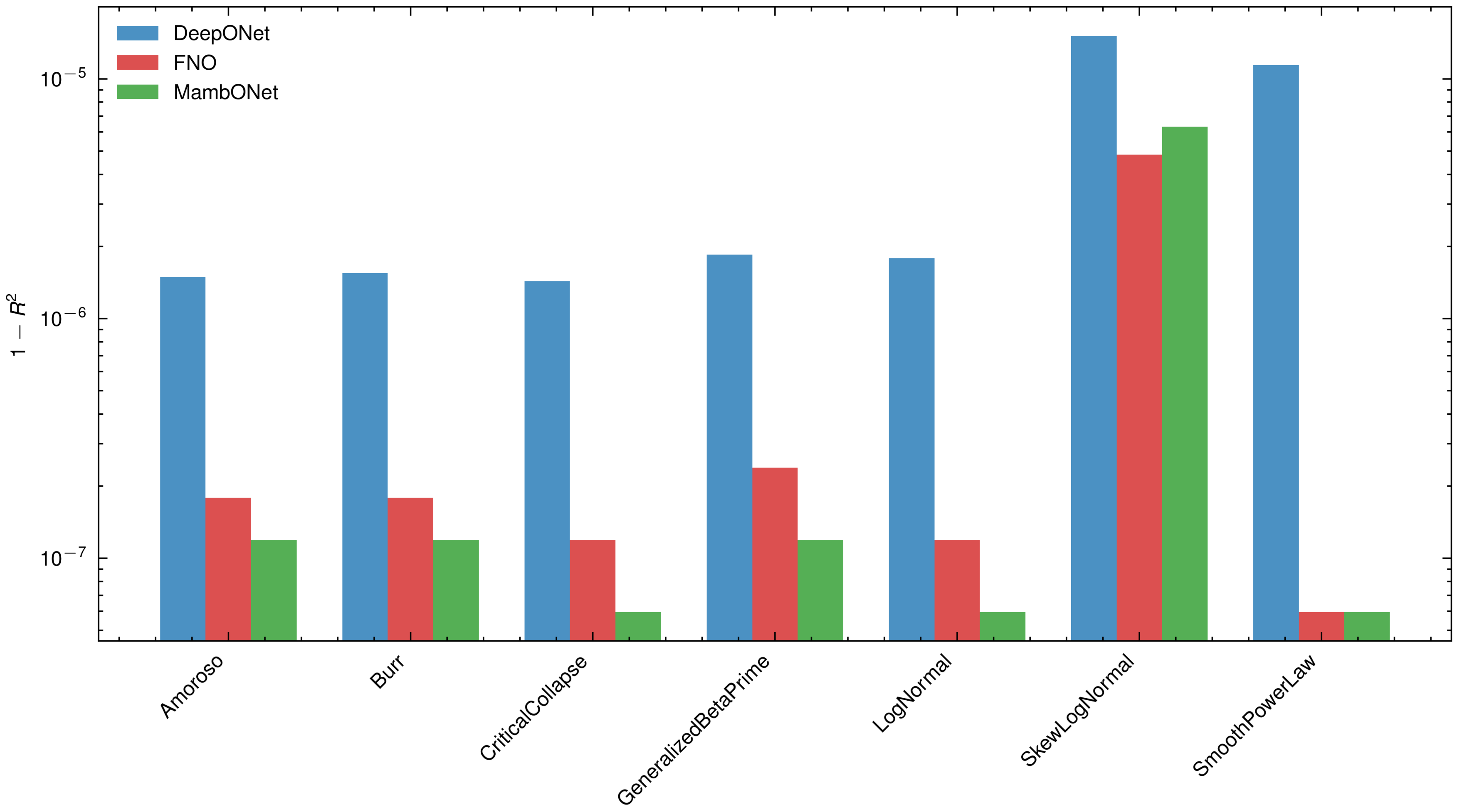
Results - *Smooth Power Law*



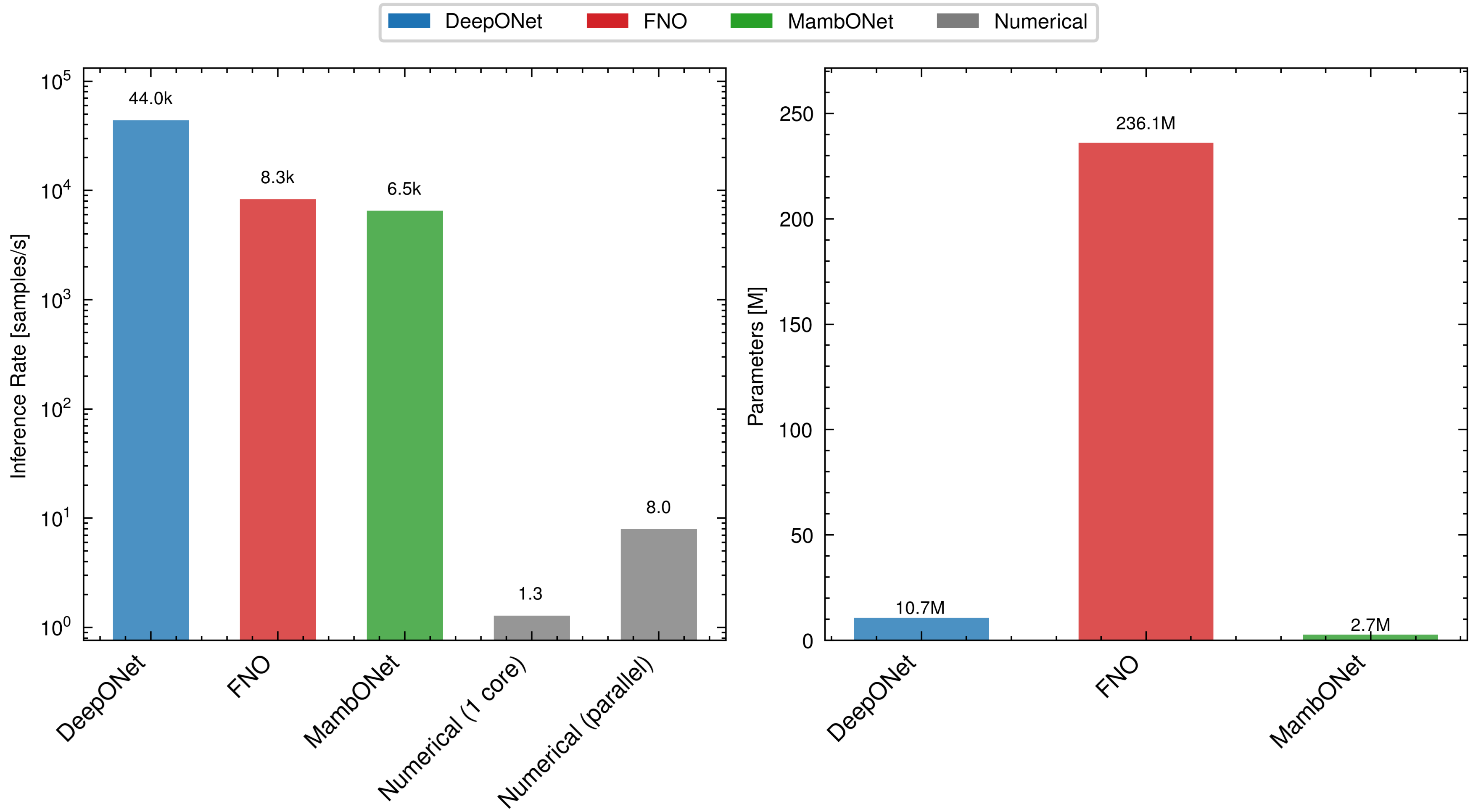
Results - *Critical Collapse*



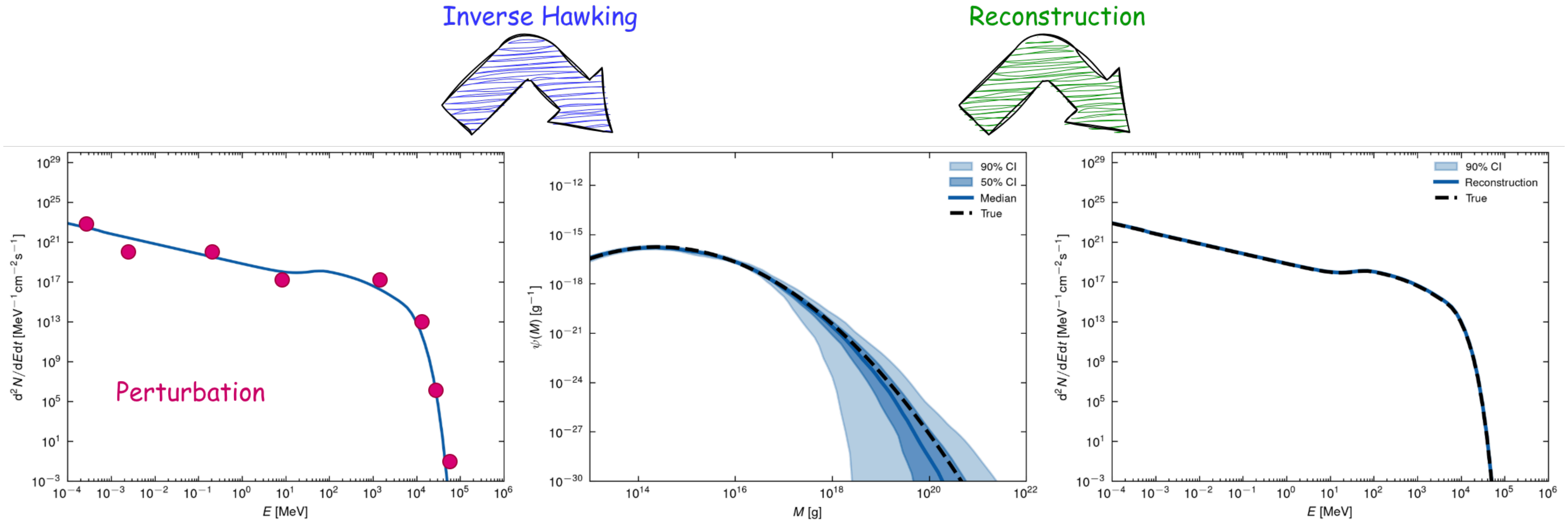
Results - *Comparison of Models*



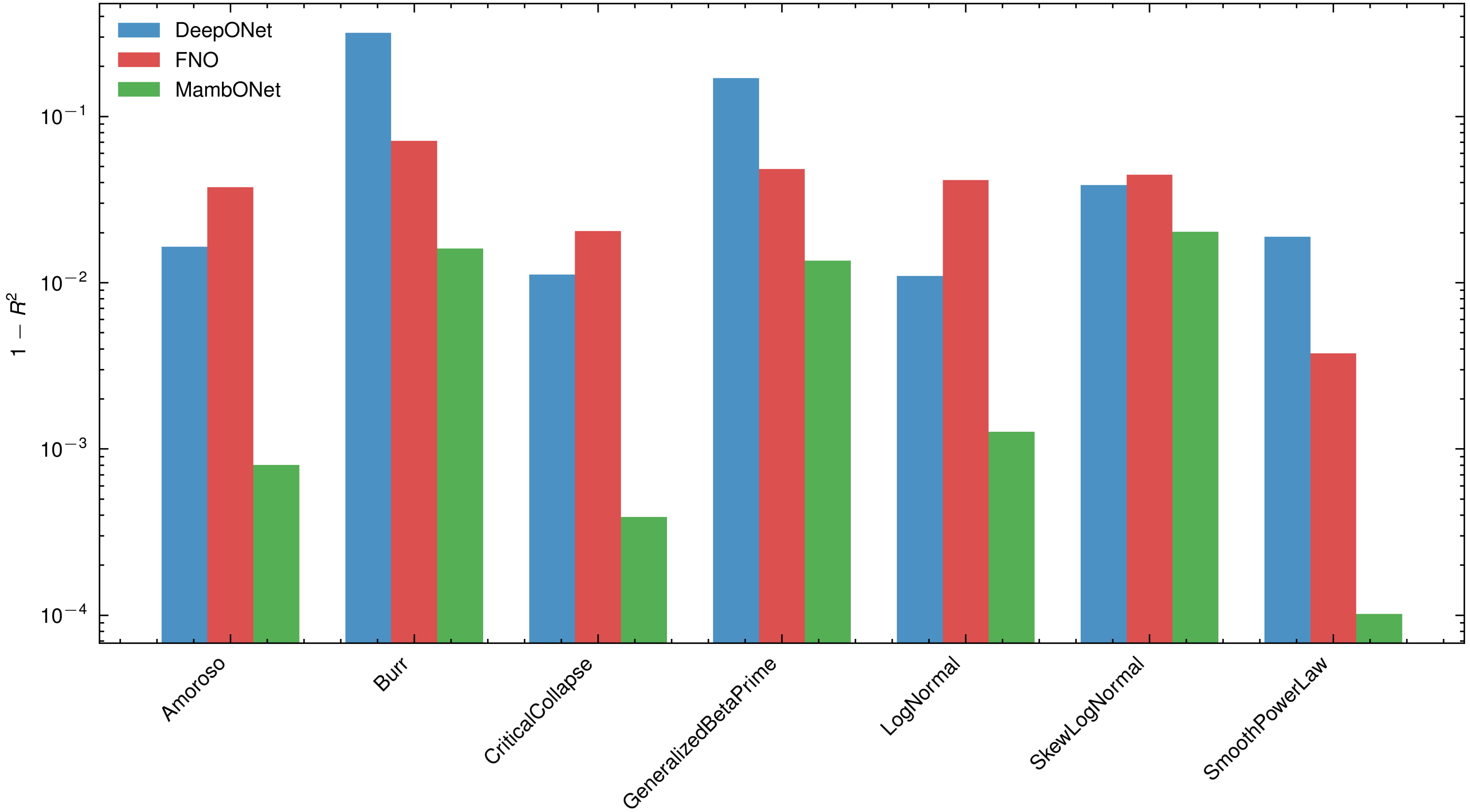
Results - *Efficiency Comparison*



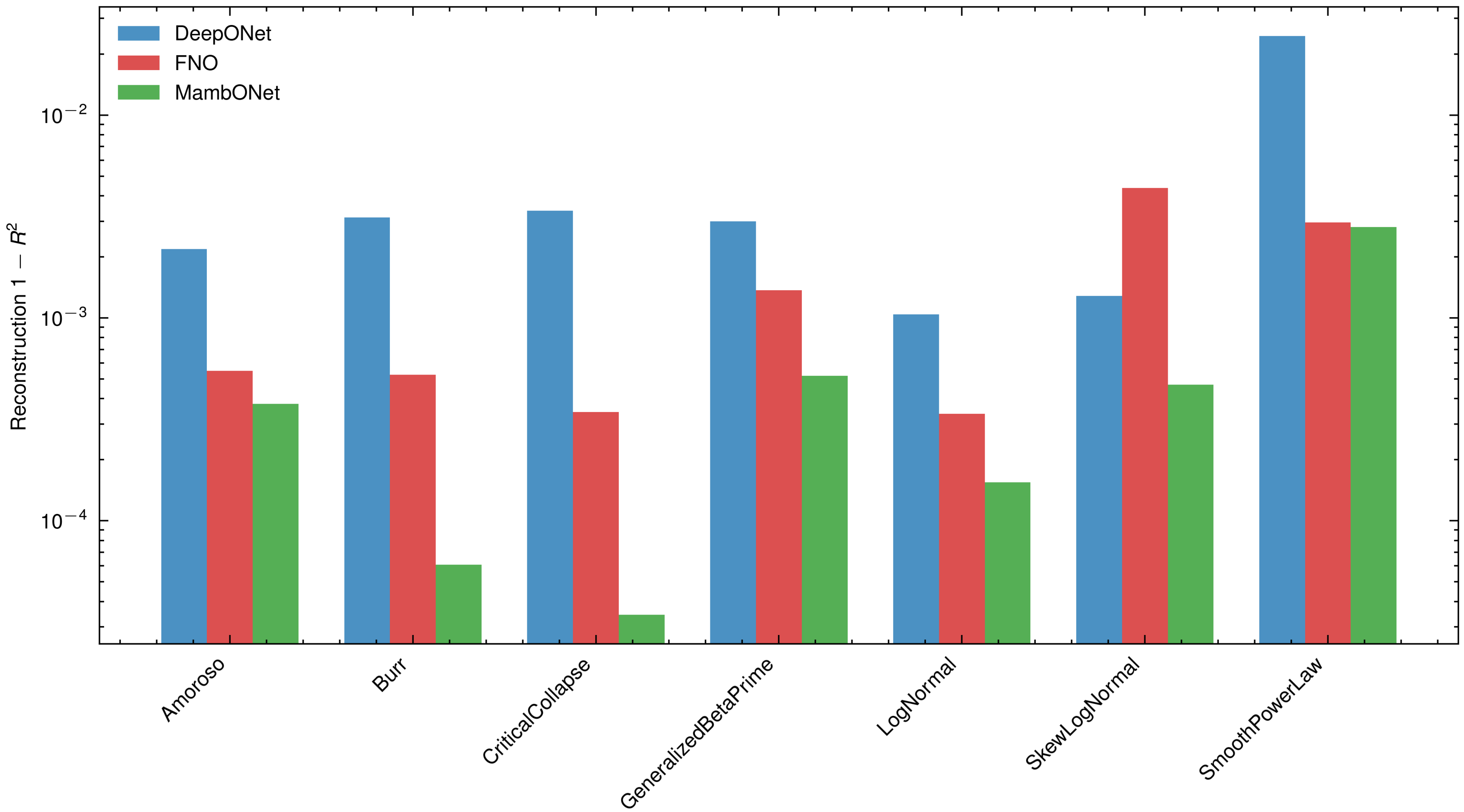
Inverse Hawking - *Uncertainty Quantification*



Inverse Hawking - *Comparison of Models*



Inverse Hawking - *Comparison of Models (Reconstruction)*



Summary

- **New Paradigm: Neural Hawking Operator**
  Established a unified framework ($\mathfrak{H}$) that replaces the discontinuous "patchwork" of traditional simulation tools with a continuous, end-to-end differentiable model.

- **Precision Meets Efficiency**
  MambONet achieves scientific-grade accuracy ($1 - R^2 < 10^{-6}$) while being significantly more lightweight (2.7M params) than SOTA baselines.

- **Accelerating Discovery**
  Delivers $\mathcal{O}(10^3)$ speedup compared to numerical methods, enabling rapid hypothesis testing

- **Toward Inverse Phenomenology**
  Successfully demonstrated the inverse mapping ($\mathfrak{H}^{-1}$), opening a direct path from gamma-ray observations to PBH mass distribution discovery.

Supplements

Greybody Factor

- A radial equation for massless field around a rotating black hole: [S. A. Teukolsky, ApJ (1973)]

$$\Delta^{-s}\frac{\mathrm{d}}{\mathrm{d}r}\left(\Delta^{s+1}\frac{\mathrm{d}R}{\mathrm{d}r}\right)+\left(\frac{K^2-2is(r-M)K}{\Delta}+4is\omega r-\lambda\right)R=0$$

  - $\Delta = r^2 - 2Mr$
  - $K = r^2\omega$
  - $\lambda = l(l+1) - s(s+1)$

- Choose tortoise coordinate r_*, then it becomes Schrodinger-like equation: [S. Chandrasekhar, Proc. Roy. Soc. Lond. (1975)]

$$\Lambda^2\phi = V(r)\phi$$

  - $r_* = r + 2M\ln(r/2M - 1)$

- At the far field, the solution is:

$$\phi(r_*) \sim Ae^{i\omega r_*} + Be^{-i\omega r_*}, r_* \to +\infty$$

  - Greybody factor: $\Gamma = 1/|B|^2$

[A. Arbey & J. Auffinger, EPJC (2021)]

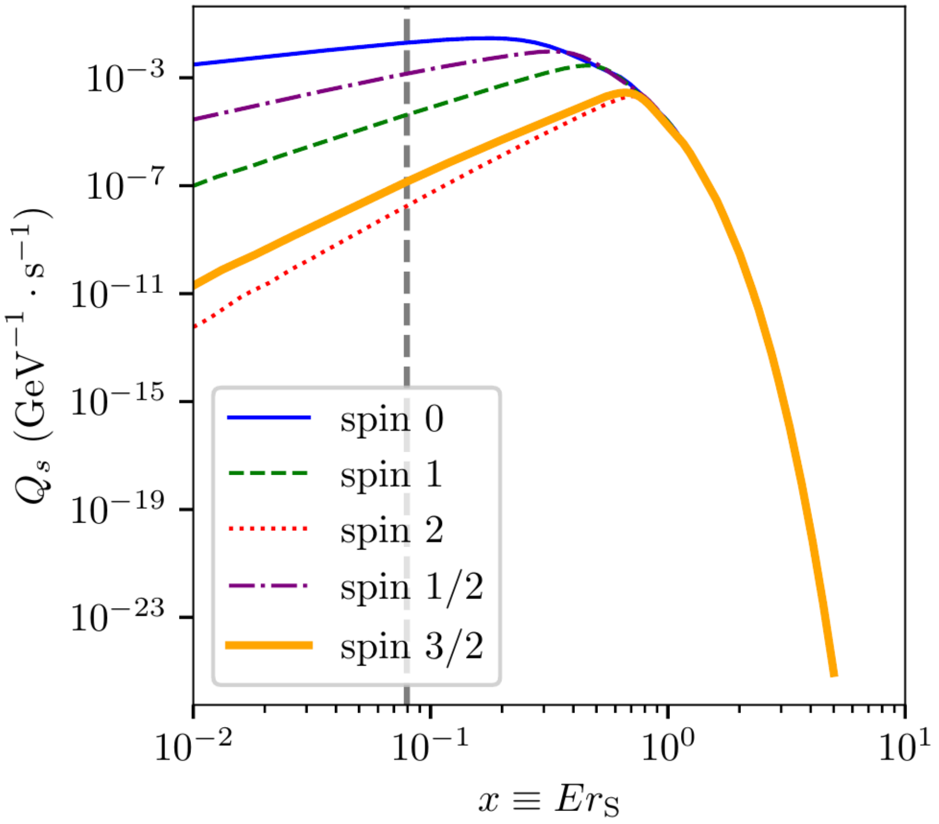


Figure 5: Hawking spectra of massless particles of various spins from a PBH

PBHs as Particle Factories

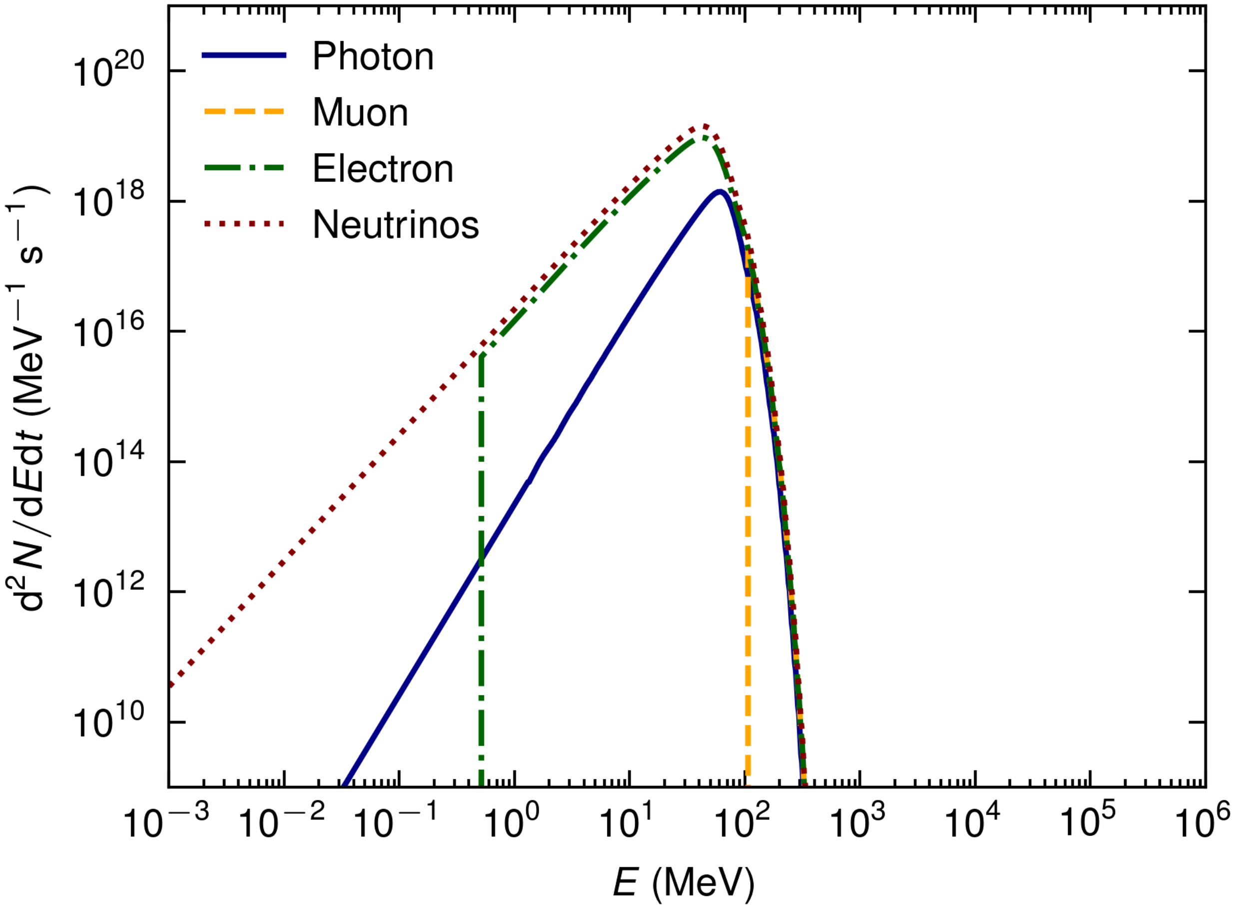


Figure 6: The emission rates of photons, muons, electrons and neutrinos from a PBH with $M_{\rm PBH} = 10^{15}$g

Input Energy Limit of PYTHIA

`parm` **Beams:eMinPert** (default = **10.**; minimum = 10.)
The lowest CM energy that collisions are allowed to have. The highest is set by the full energy at initialization, as calculated in the respective options above. If you do not want to generate events with a higher collision energy than this you can save some initialization time by raising this number accordingly, so that fewer grid energies need to be used. You should not lower the value below the default one, however, since the perturbative MPI machinery cannot be made to work at lower energies. If you do want to generate events with lower energies, it is possible in a nonperturbative framework, see next.

Figure 7: Excerpt from the PYTHIA manual indicating the lower limit of primary particle energy for accurate simulation

Is the Inverse Hawking Operator Well-Defined?

- **The Question**: Given an observed total spectrum $\Phi(E)$, can we uniquely determine the original PBH mass distribution $\psi(M)$?
- **The Analogy**: The Hawking operator $\mathfrak{H}$, mapping a mass function to a spectrum, behaves much like a **Laplace Transform**.

$$\psi(M) \xrightarrow{\text{Hawking Op. } \mathfrak{H}} \Phi(E)$$

$$f(t) \xrightarrow{\text{Laplace } \mathcal{L}} F(s)$$

- Since the Laplace transform is injective (one-to-one), we expect the Hawking operator to be as well, meaning its inverse $\mathfrak{H}^{-1}$ should be well-defined.

Sketch of Proof

- The operator is an integral transform: $\Phi(E) = \int K(E,M)\psi(M)\,\mathrm{d}M$ where the kernel $K(E,M)$ is the spectrum from a single PBH of mass M.
- Injectivity is guaranteed if the set of kernel functions $\{K(E,M)\}_E$ is **dense** (i.e., they can form a basis for any reasonable $\psi(M)$).
- At high energies ($ME \gg 1$), the kernel simplifies to:

$$K(E,M) \approx \text{poly}(M,E) \cdot e^{-8\pi ME}$$

- Two theorems confirm this:
  - Müntz-Szász: Confirms that sets of exponentials like $\{e^{-c_n M}\}$ are dense.
  - Paley-Wiener: Ensures that small perturbations don't break this property.

Conclusion: The operator is injective.

Training Details

- A dataset of **100,000 PBH instances** with varied mass functions was generated via a custom Rust code, which simulates the Hawking radiation process. We divided the dataset into training (80%), validation (15%), and test (5%) sets.
- We utilized several neural operator architectures (e.g., DeepONet, FNO, MambONet) implemented in PyTorch.
- For optimizer, we used **SPlus** optimizer [K. Frans et al., arXiv:2506.07254] and for learning rate scheduler, we used the **ExpHyperbolicLR** [**T.-G.Kim**, arXiv:2407.15200]
- Hyperparameters for each architecture were optimized using **Optuna's TPE sampler** over 100 trials (1 trial includes 10 epochs), targeting minimum validation loss.
- The best models were trained using the discovered parameters for 250 epochs with a batch size of 100.
- For both training and inference, we used single NVIDIA RTX 5090 GPU with 32GB of memory.

Inference Rate vs Batch Size

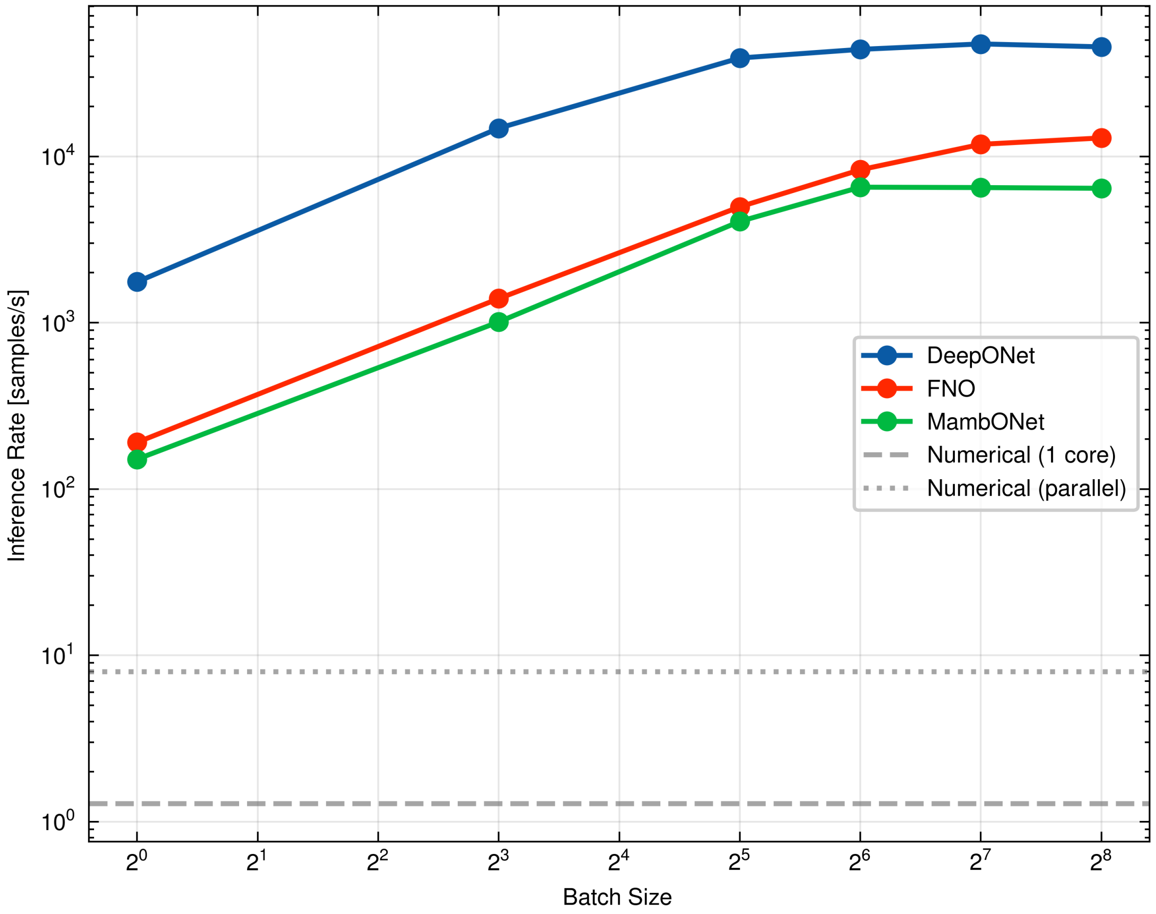


Figure 8: Inference rate (samples/second) with varying batch sizes