Evidence Lower Bound

for model selection in high-dimensional Bayesian Inference

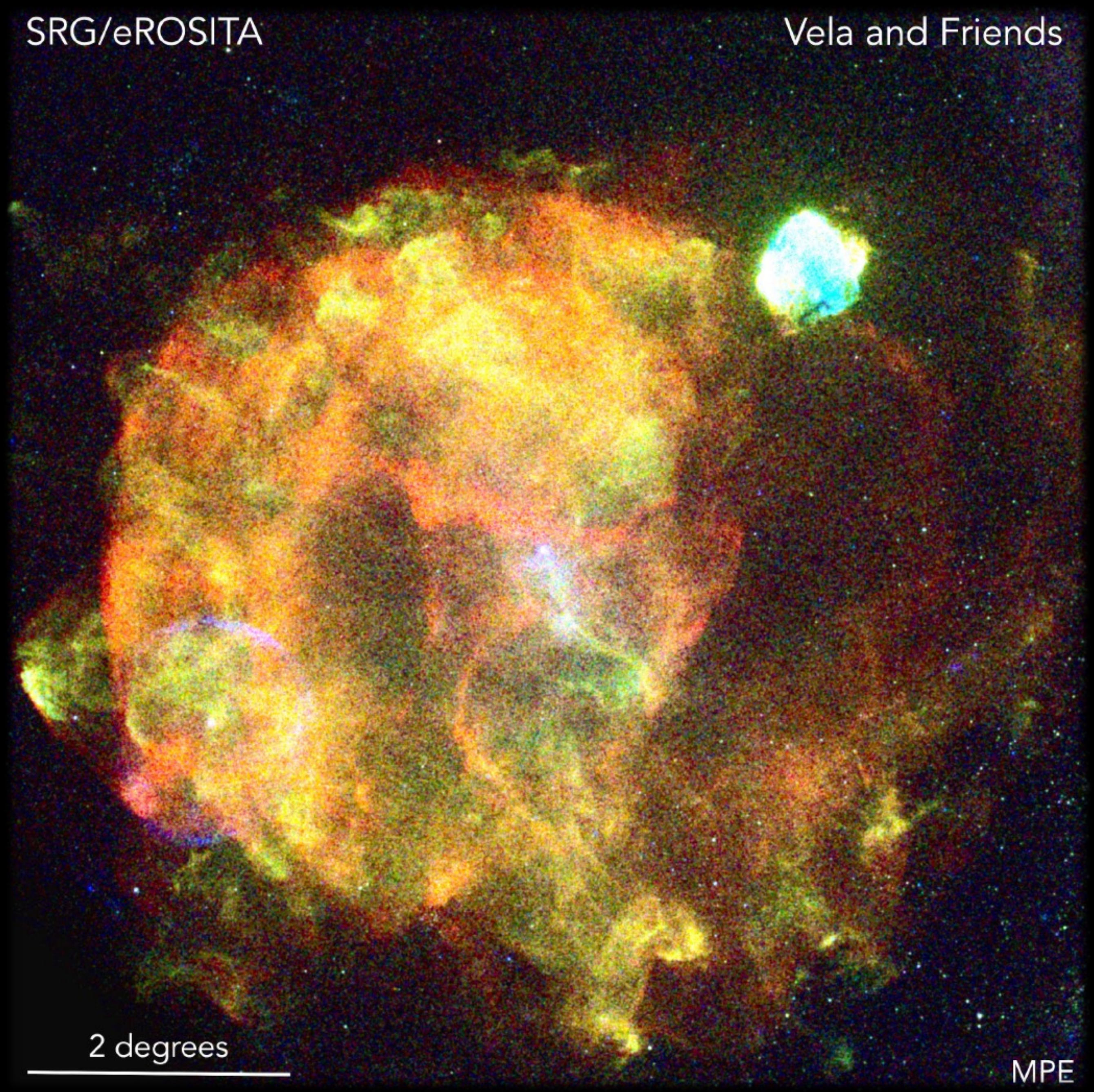
How does data challenge our models?

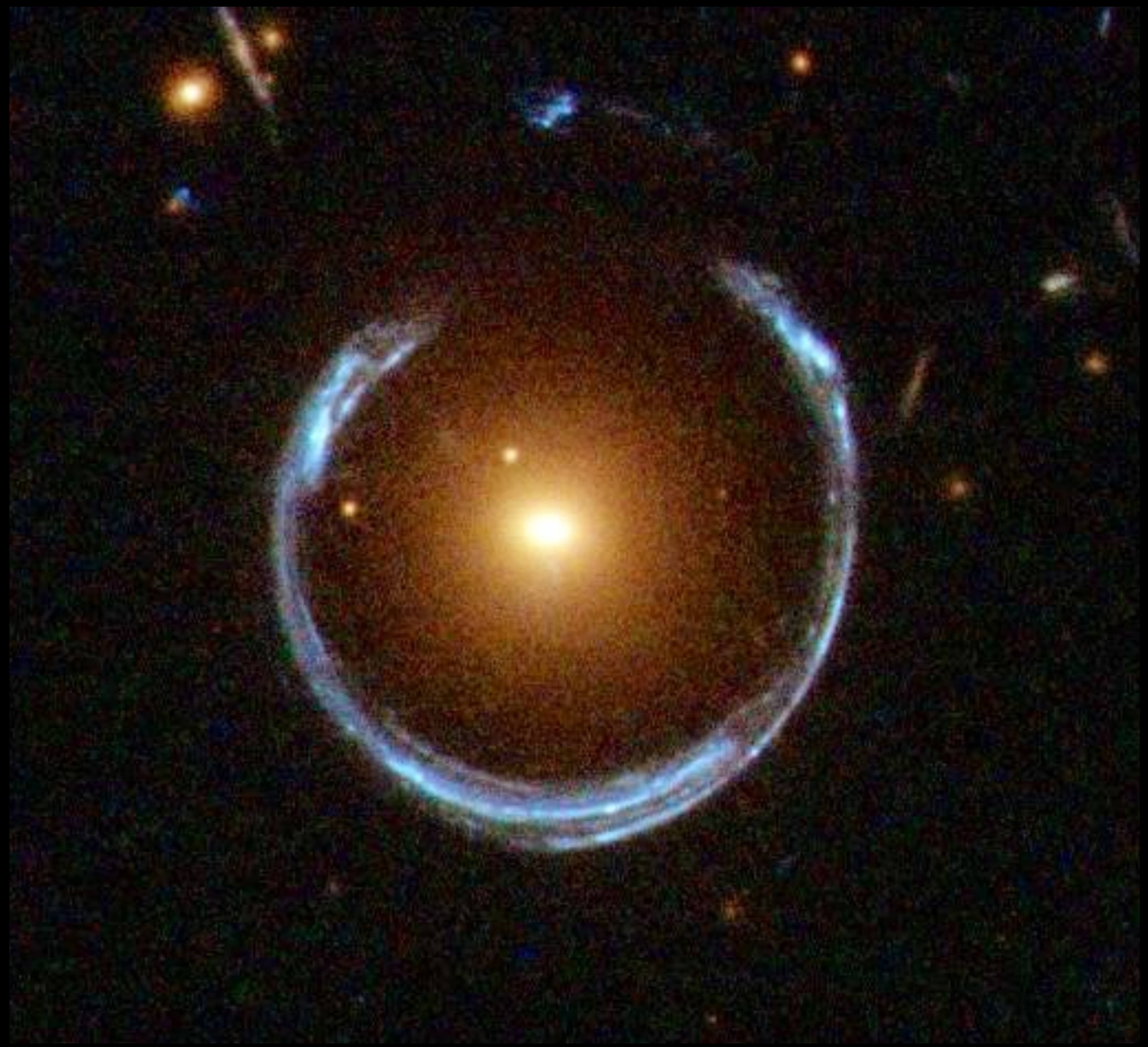
Matteo Guardiani
PhD candidate
Information field theory group
Supervisor: Torsten Enßlin

March 24th 2026
Workshop on “Model Choice in Bayesian Inference”
Center for Advanced Studies
LMU Munich

DLR
atlas
MAX PLANCK INSTITUTE FOR ASTROPHYSICS

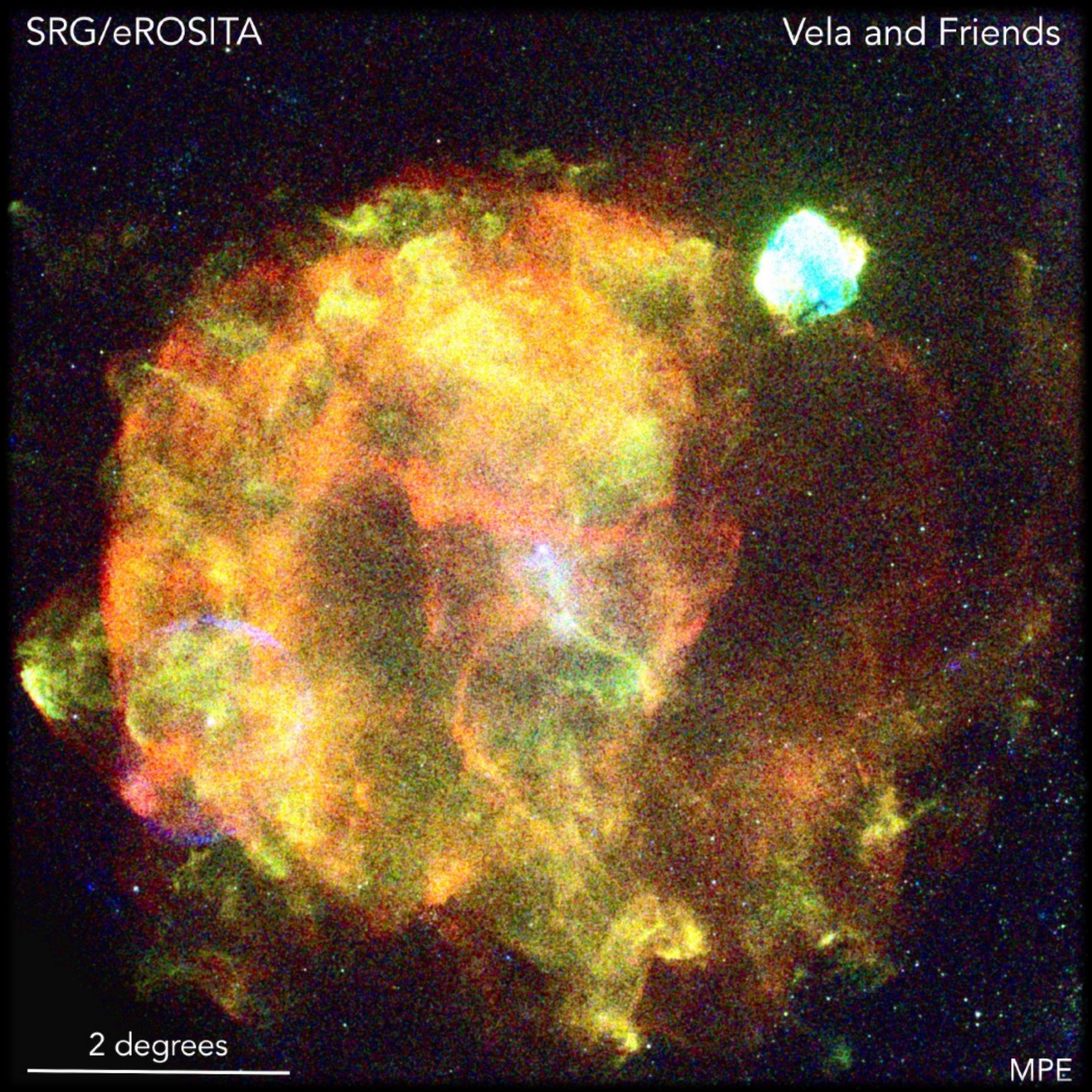
The problem

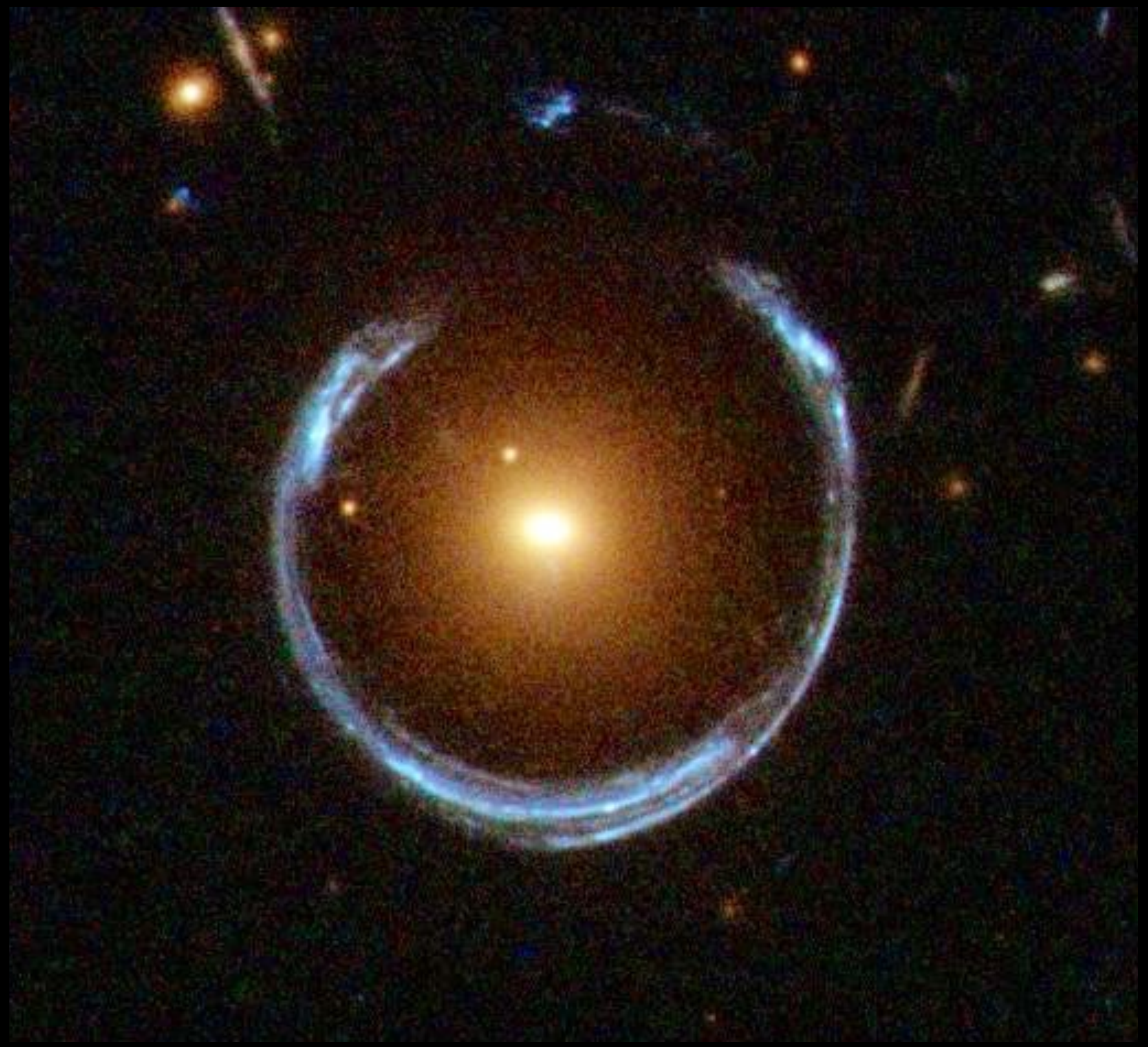




The problem

From great data come great responsibility



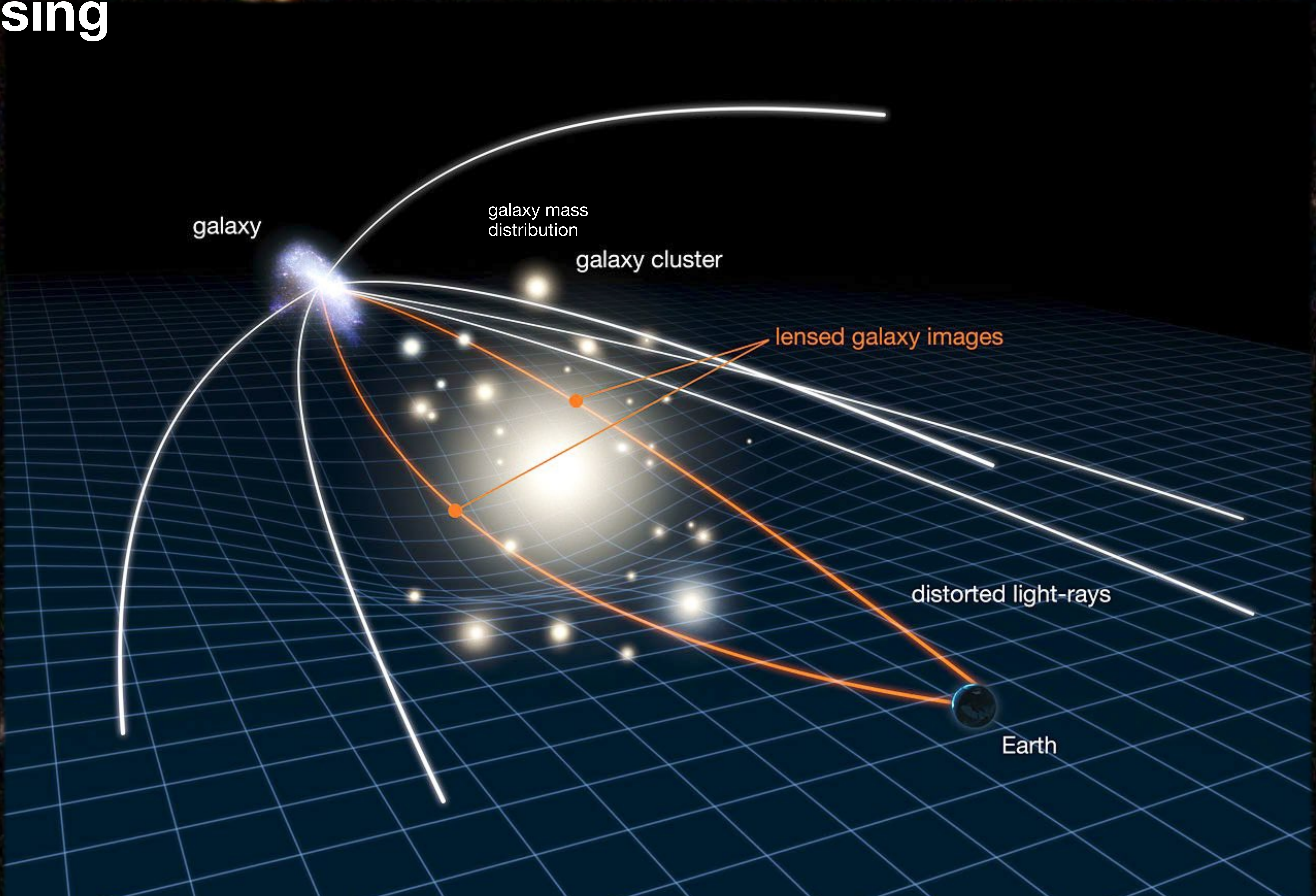


The problem

Strong lensing

The problem

Strong lensing



Modeling data with IFT

Information field theory

Bayes' Theorem

$$P(s \,|\, d) =$$

Information field theory

Bayes' Theorem

$$P(s\,|\,d) = \frac{P(s)}{}$$

Information field theory

Bayes' Theorem

$$P(s\,|\,d) = \frac{P(d\,|\,s)\,P(s)}{}$$

Information field theory

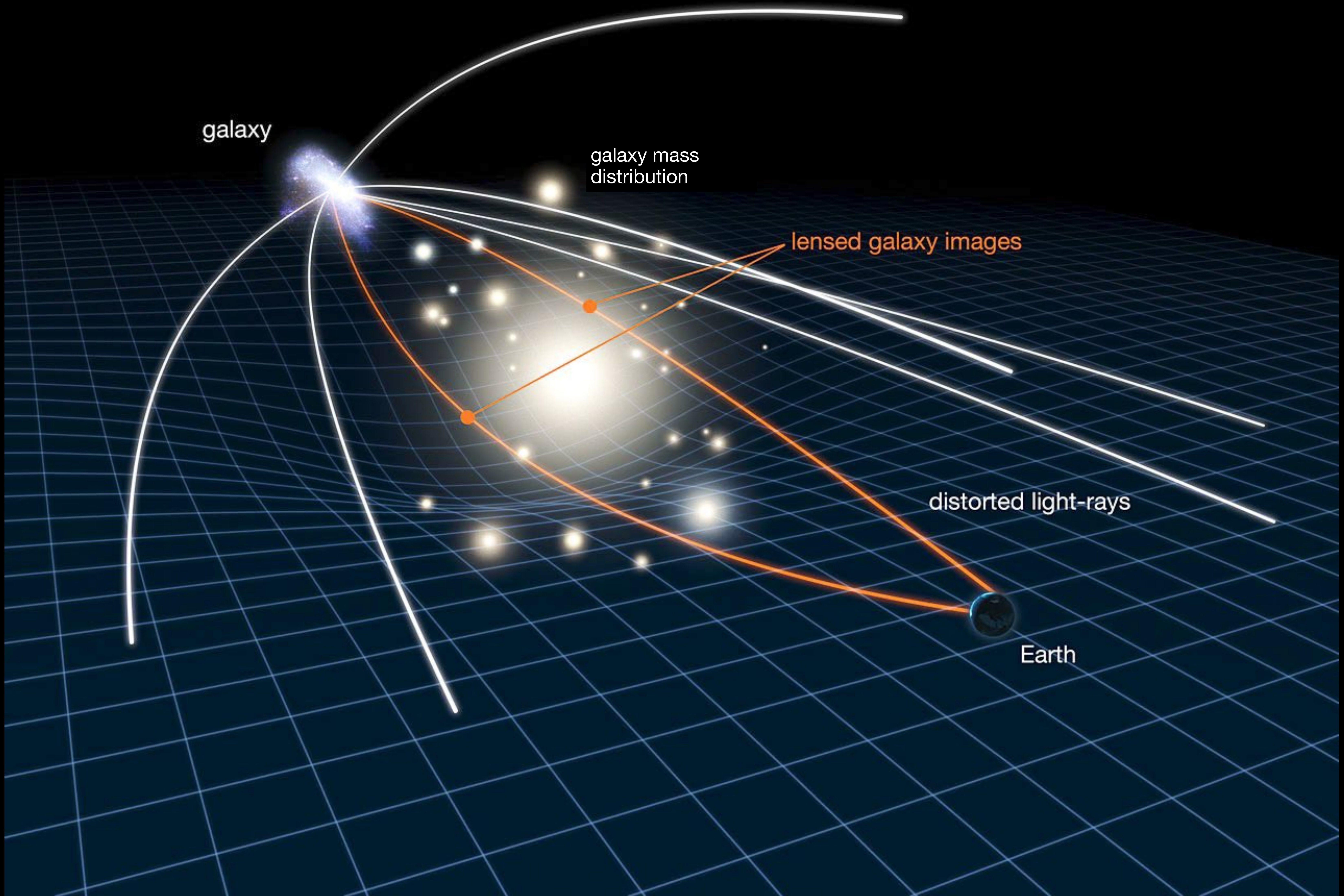
Bayes' Theorem

$$P(s\,|\,d) = \frac{P(d\,|\,s)\,P(s)}{P(d)}$$

The prior

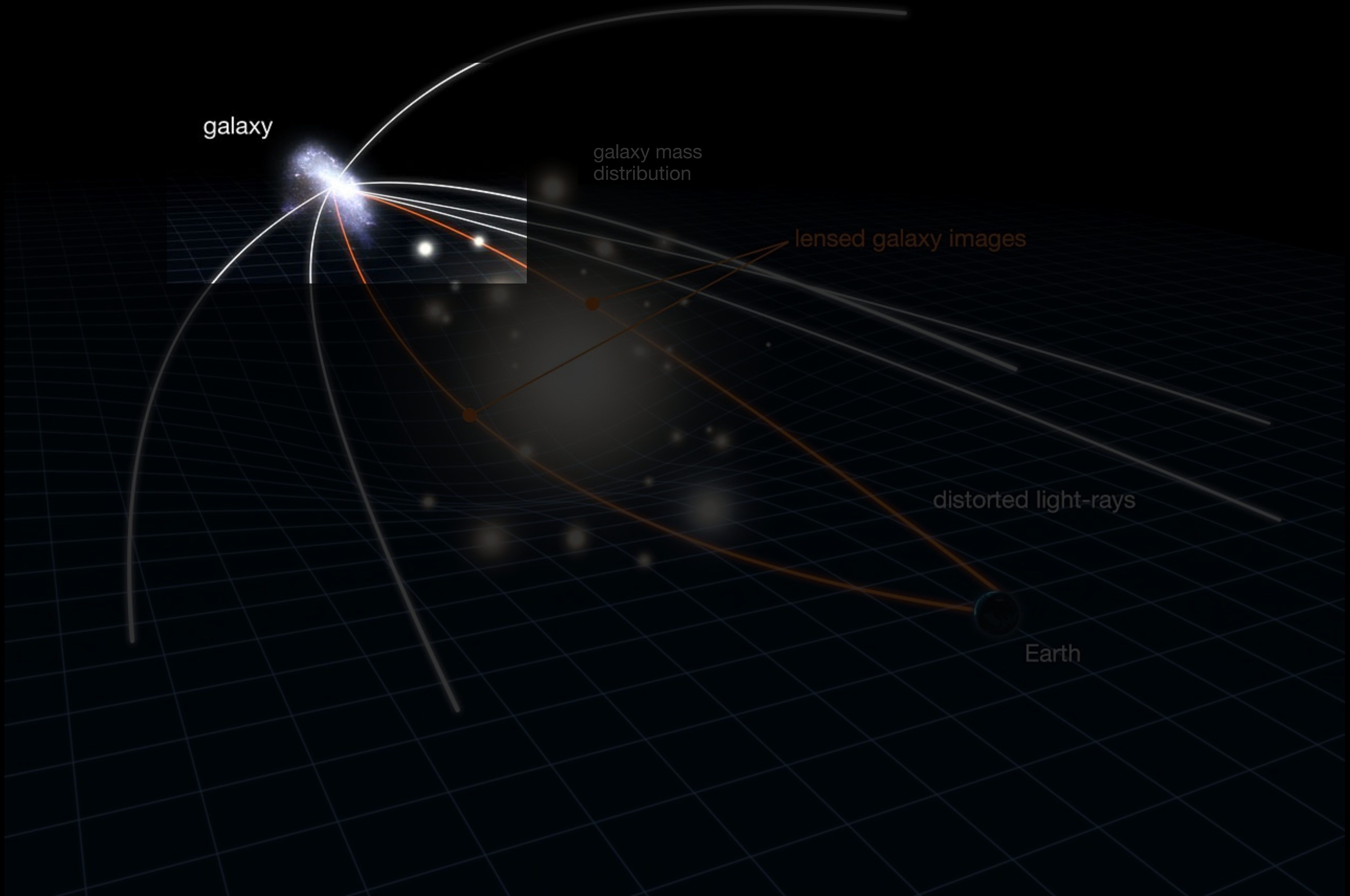
$$P(s \mid d) = \frac{P(d \mid s)\, P(s)}{P(d)}$$

The prior



galaxy
galaxy mass distribution
lensed galaxy images
distorted light-rays
Earth

The source

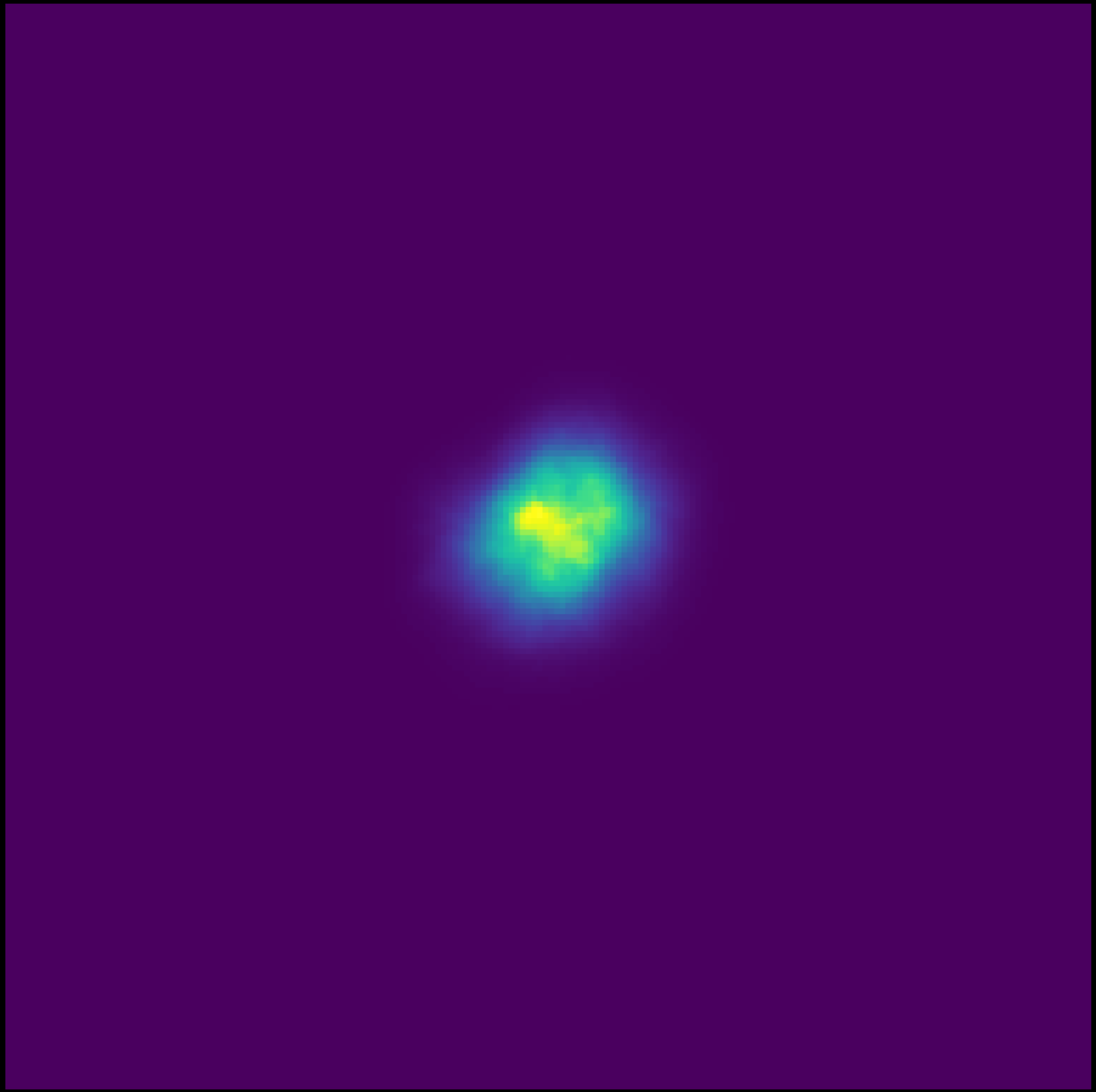


galaxy
galaxy mass
distribution
lensed galaxy images
distorted light-rays
Earth

The prior

galaxy

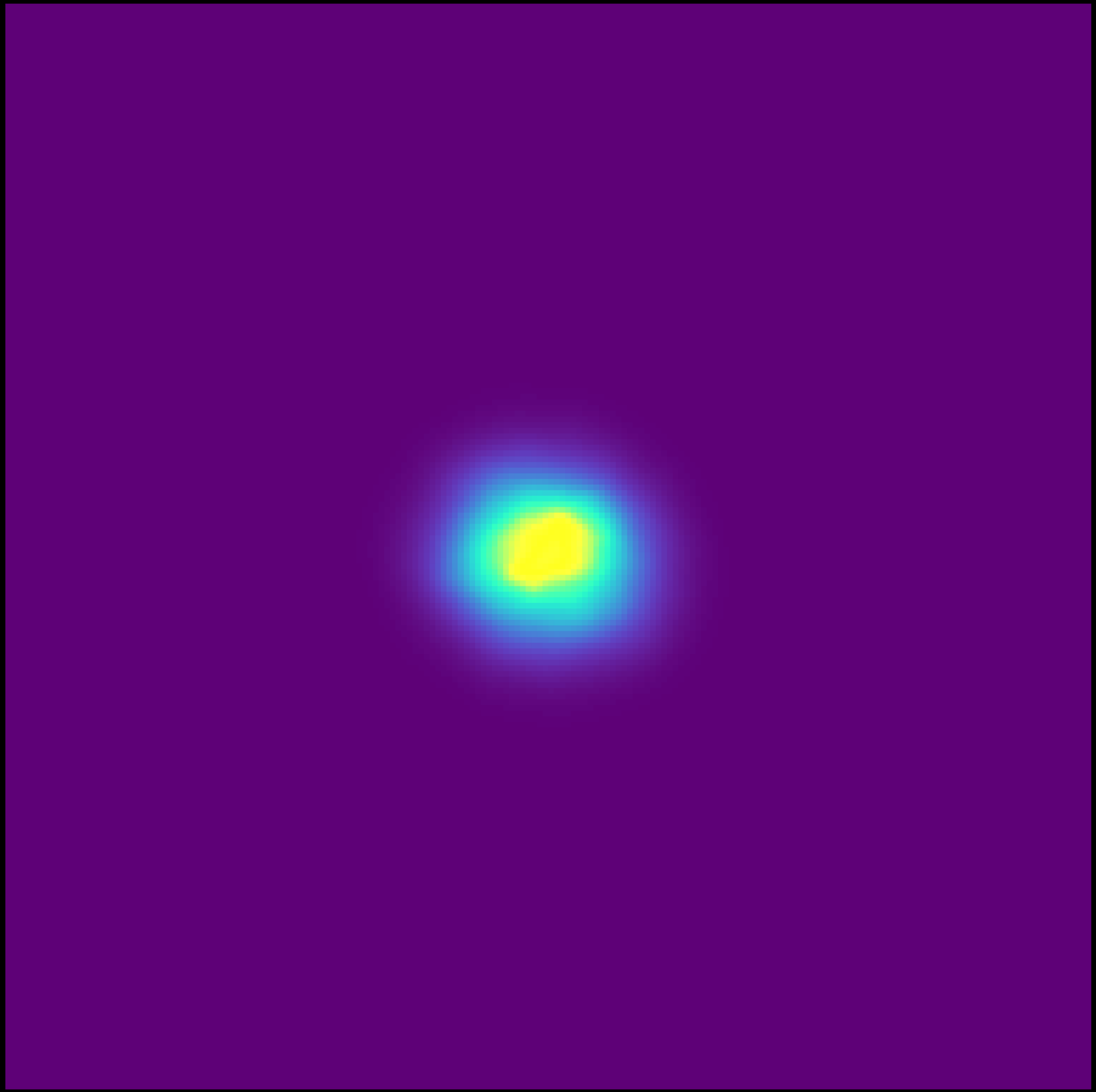
$$P(s)$$



The prior

galaxy

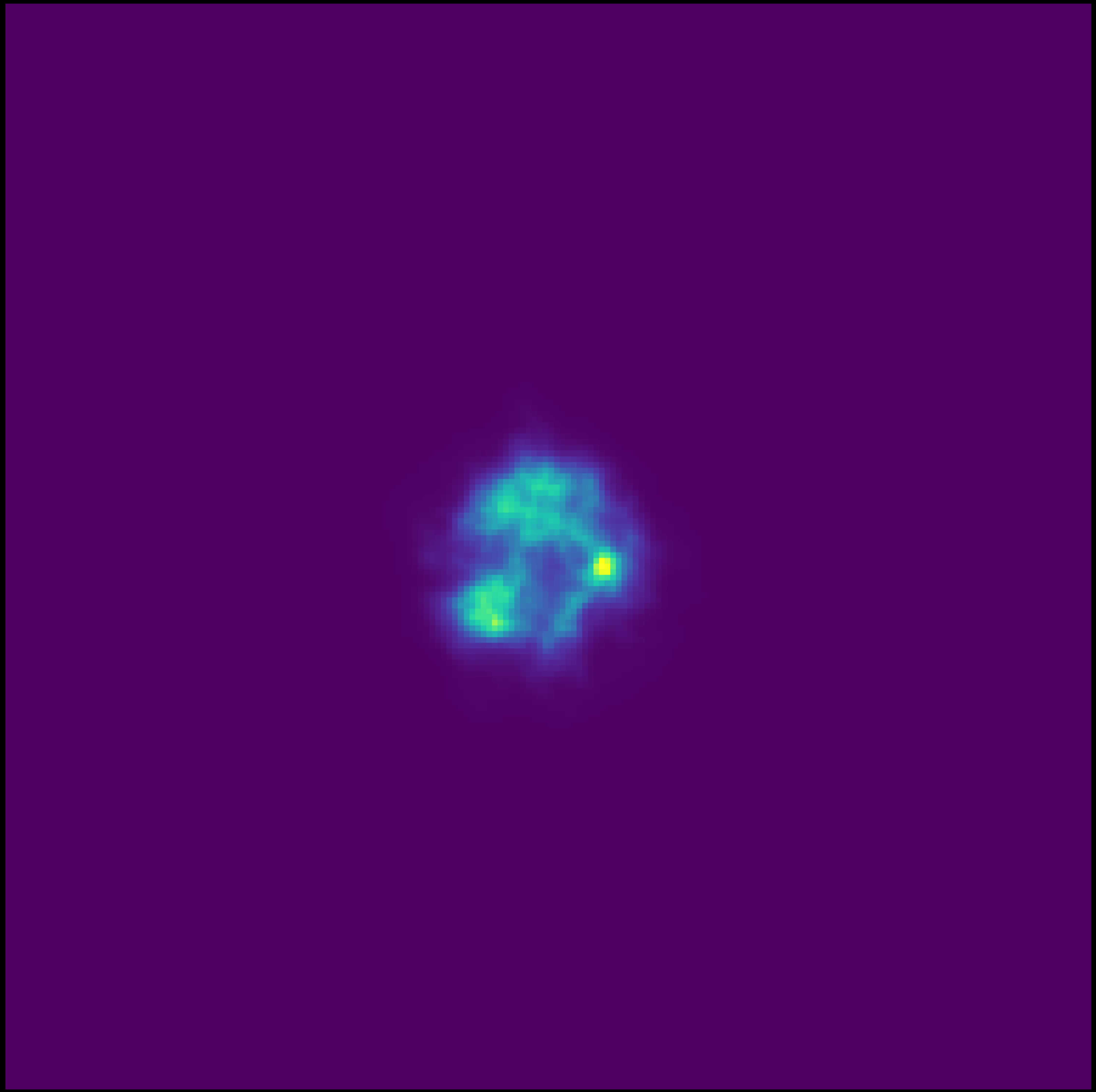
$$P(s)$$



The prior

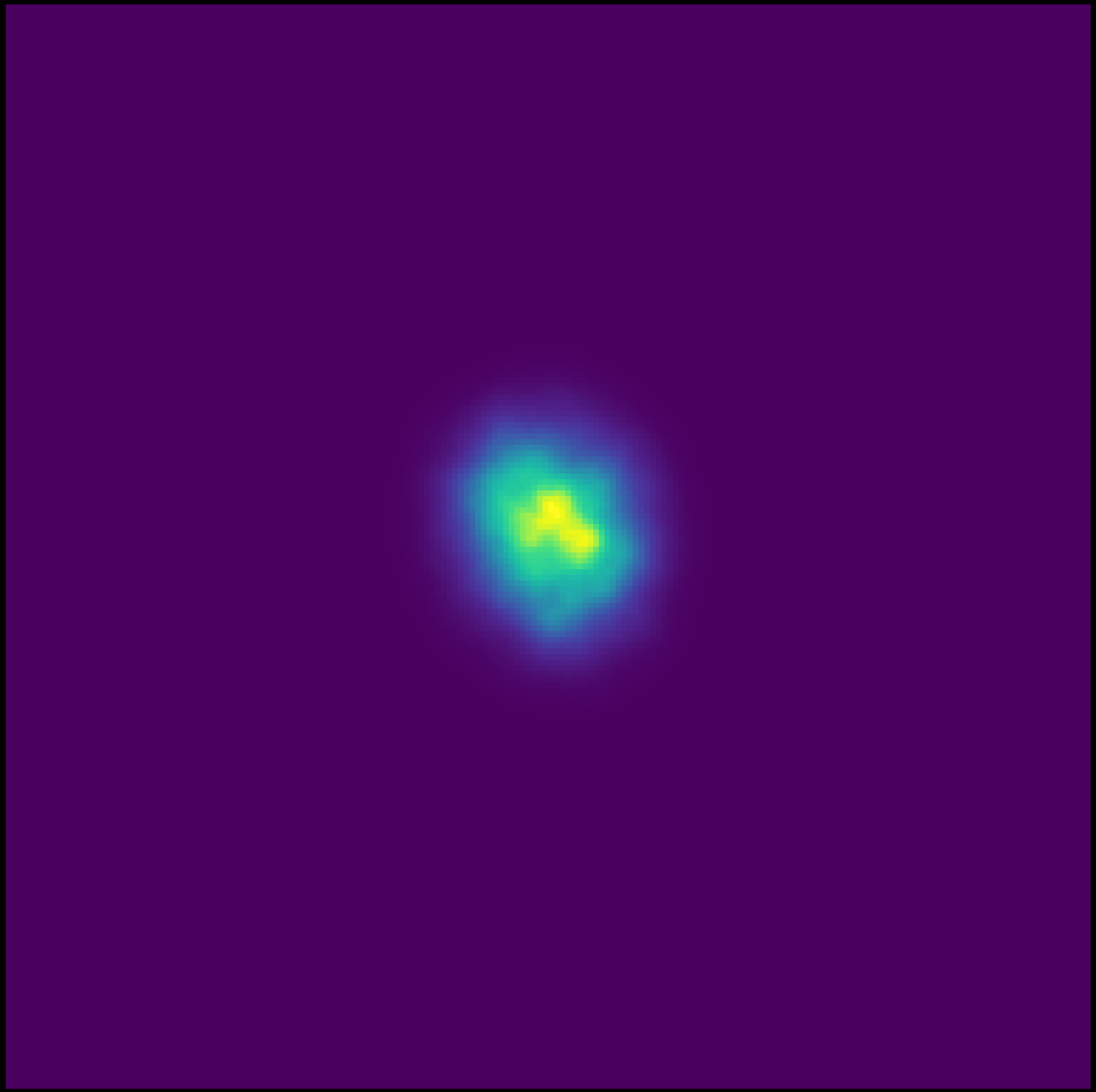
galaxy

$$P(s)$$



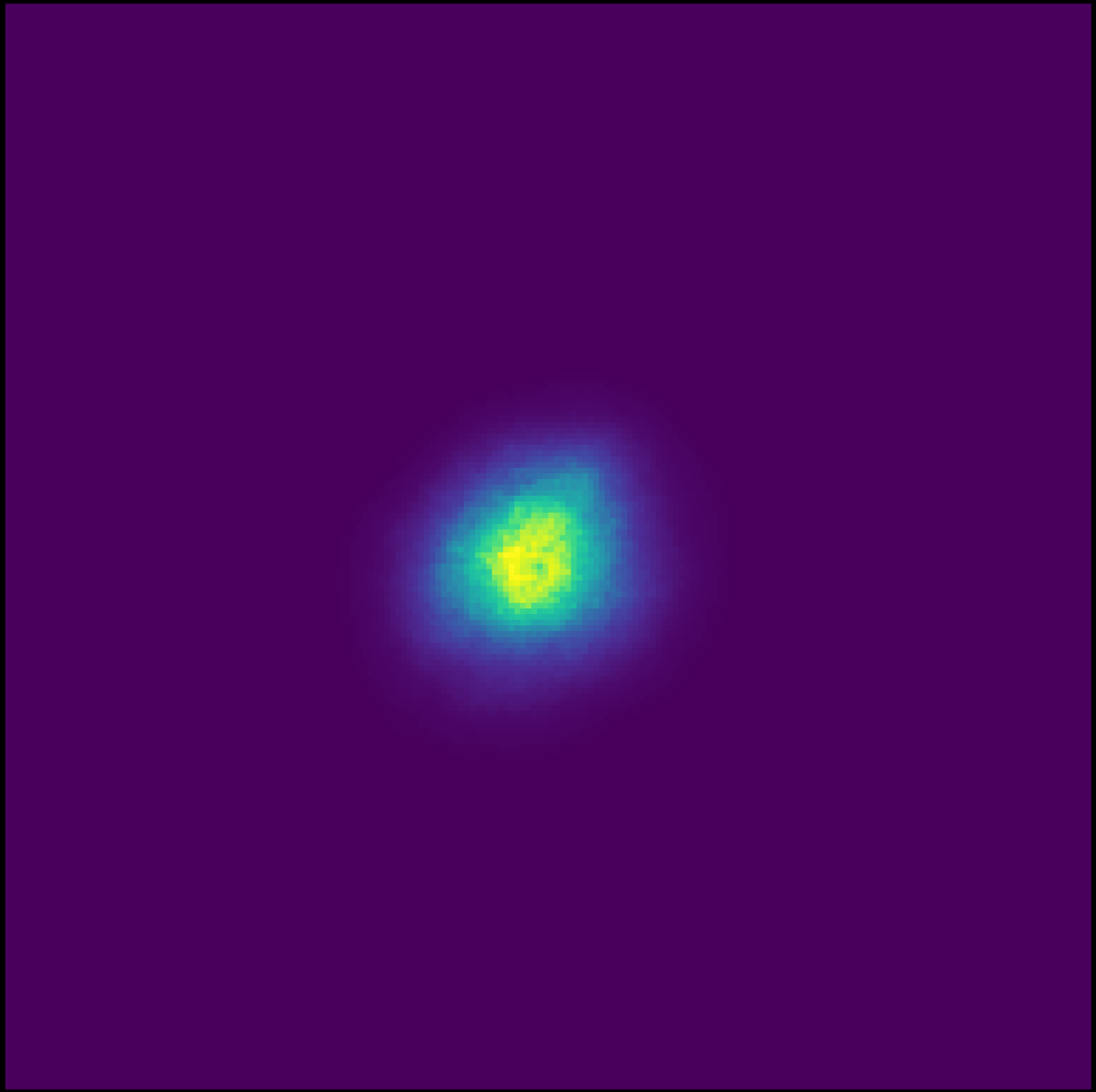
The prior
galaxy

$$P(s)$$

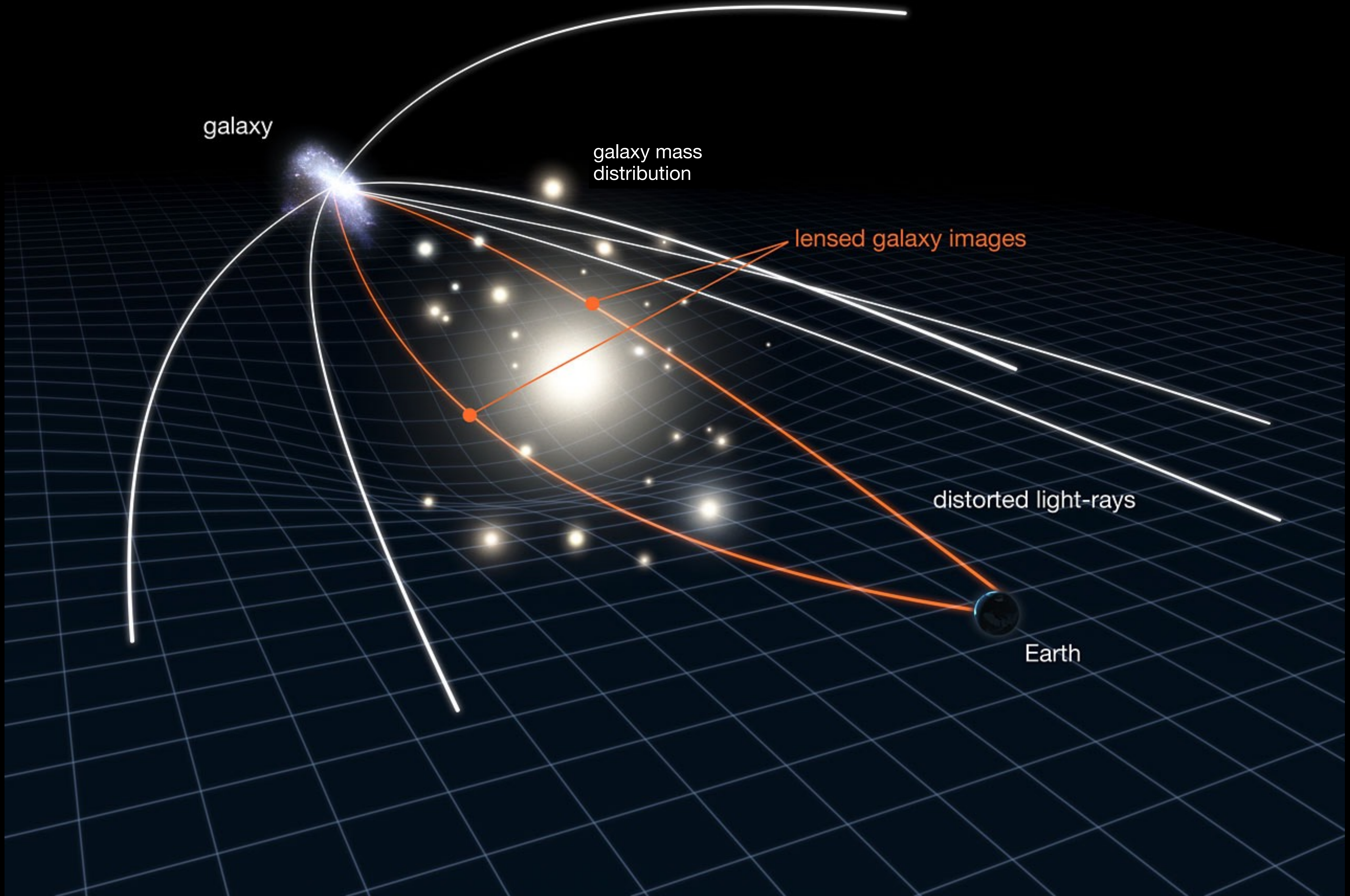


The prior
galaxy

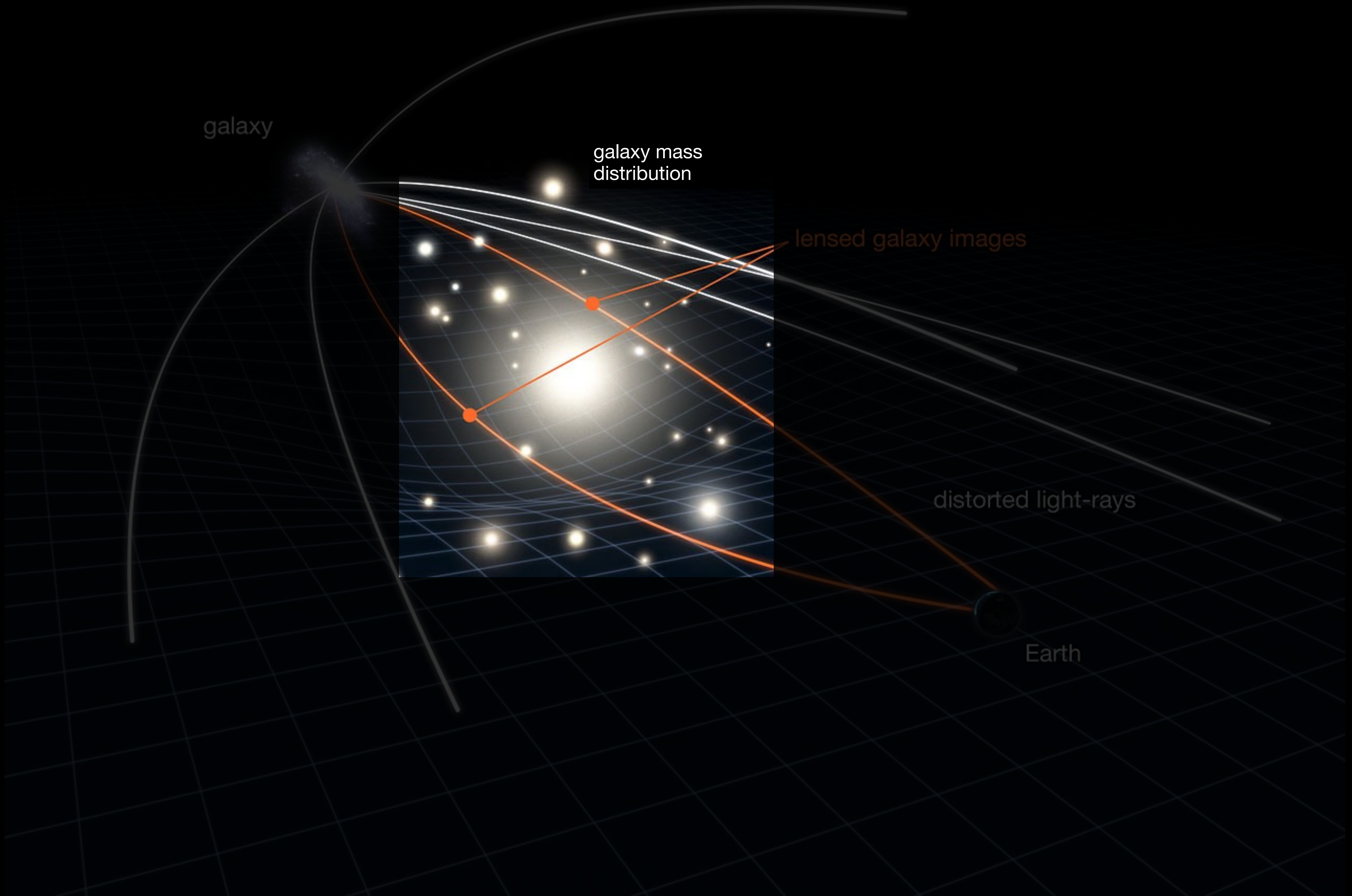
$$P(s)$$



The convergence



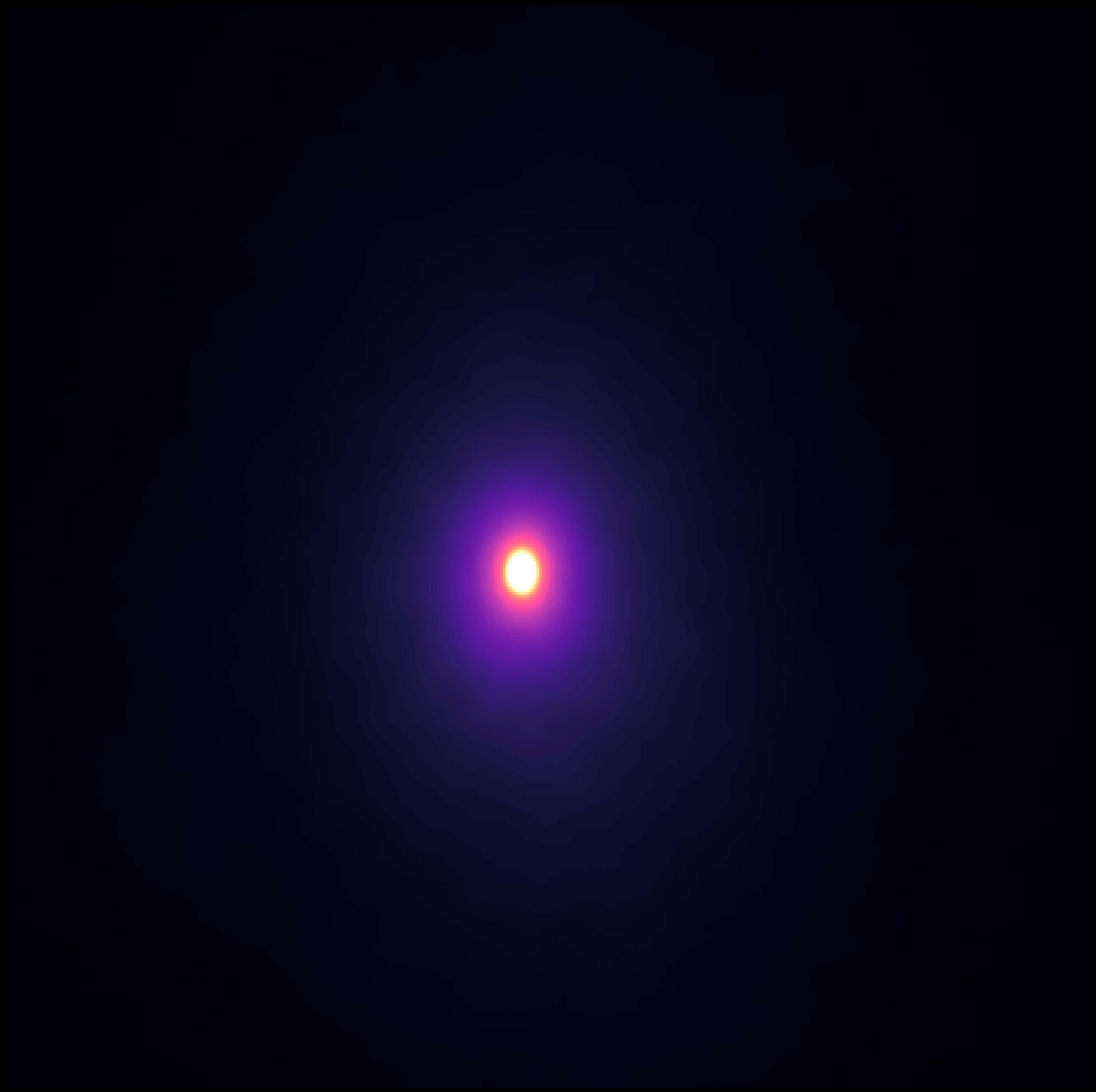
The convergence



The prior

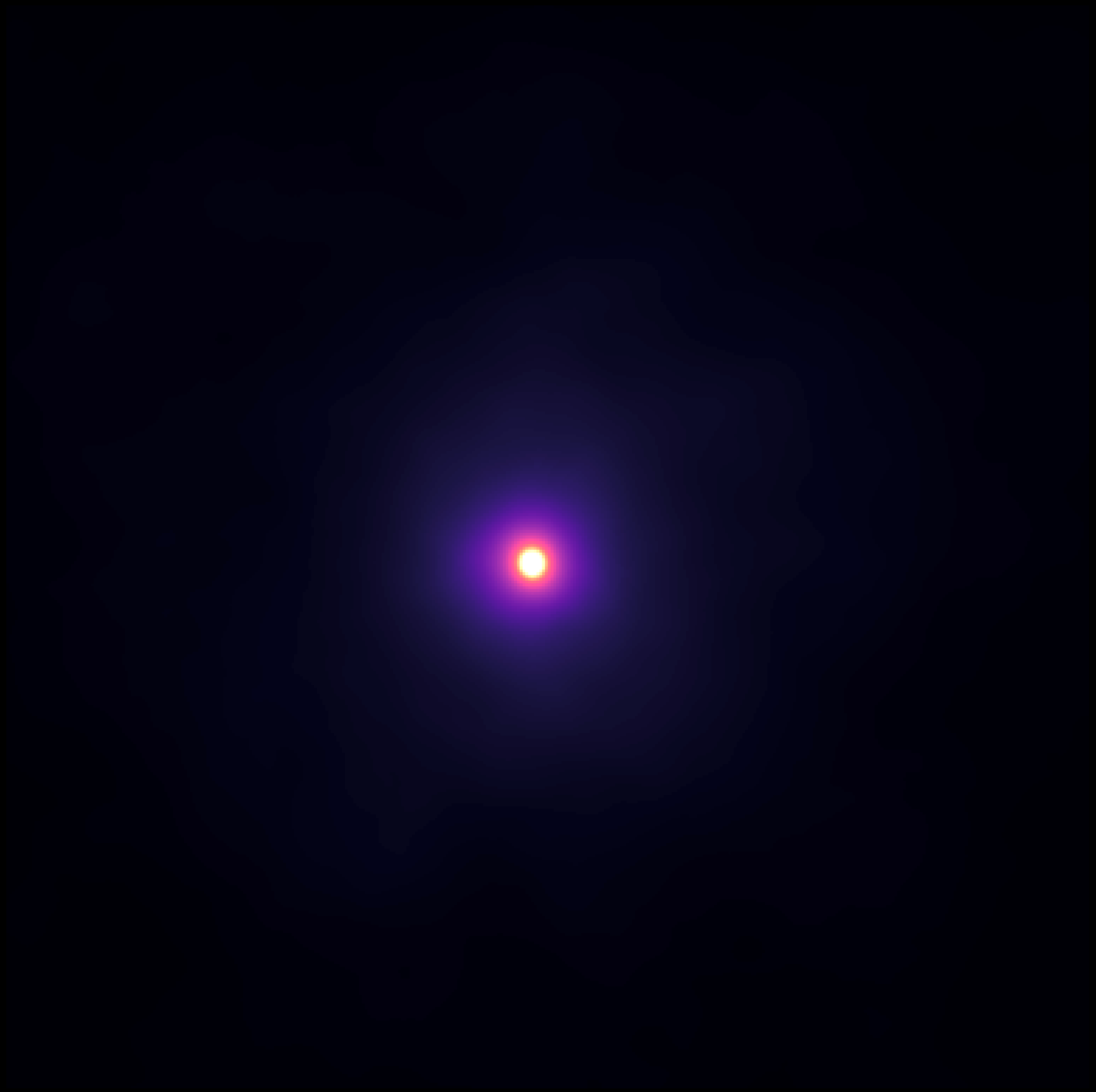
Convergence

$$P(s)$$



The prior
Convergence

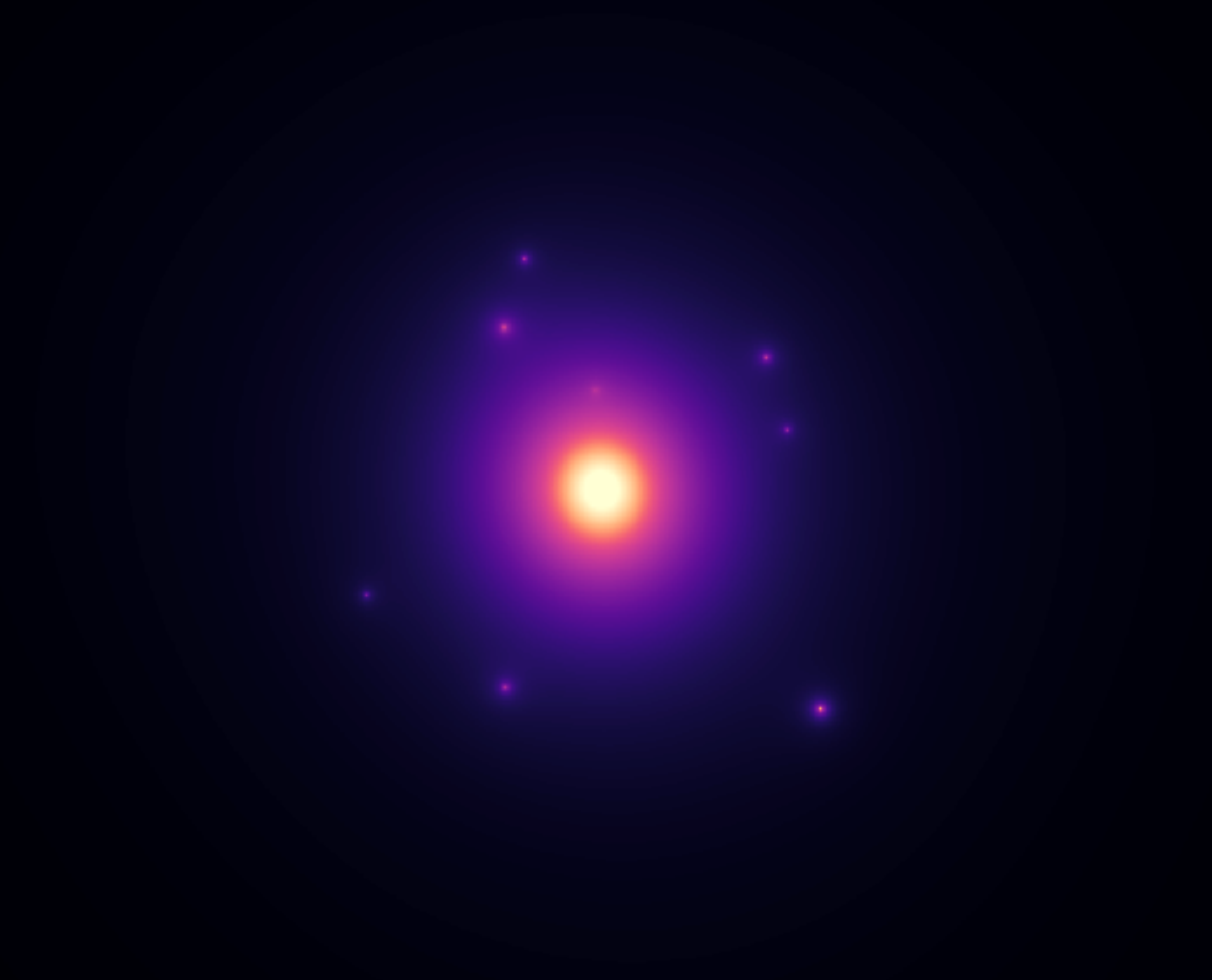
$$P(s)$$



The prior

Convergence

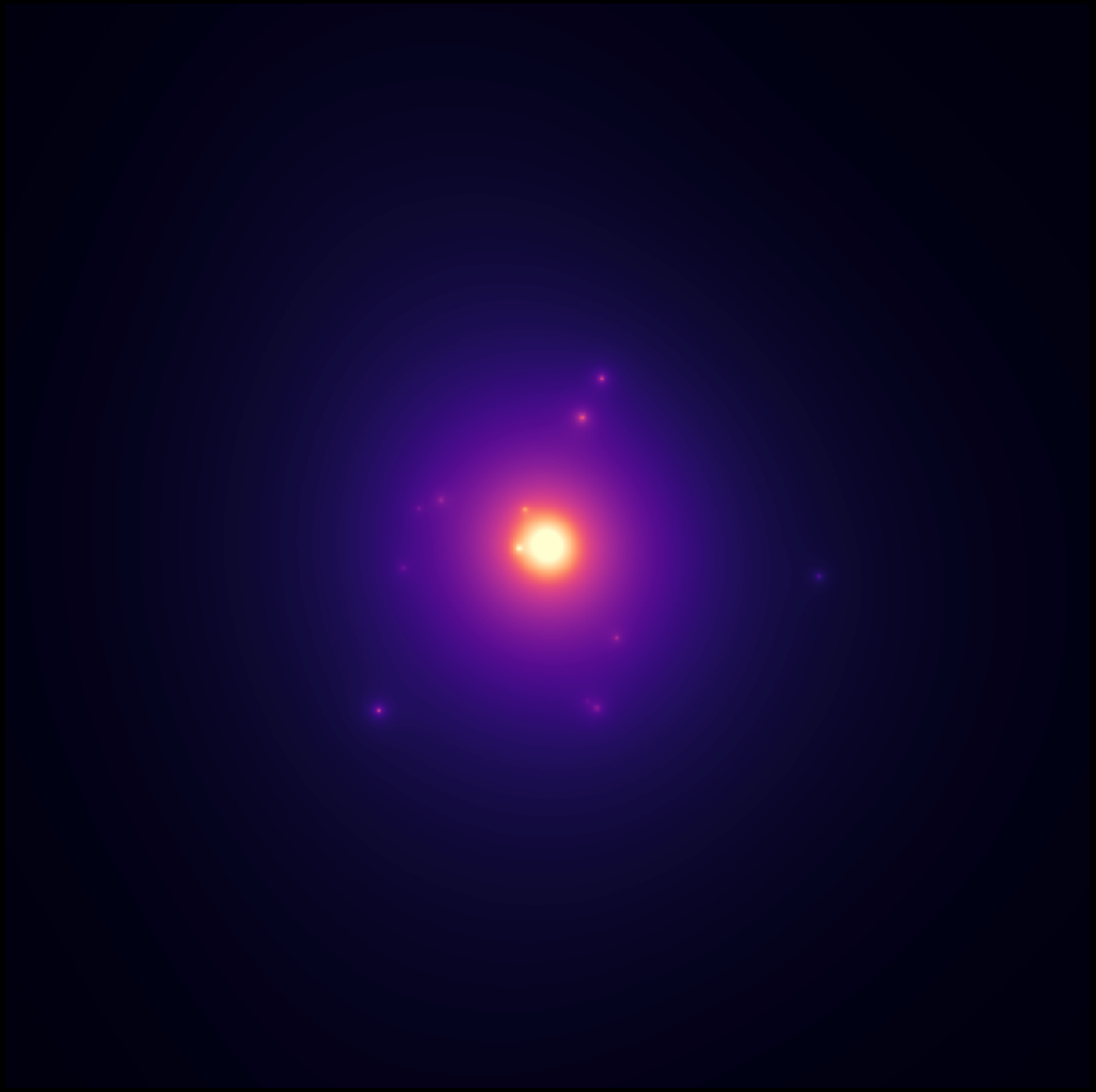
$$P(s)$$



The prior

Convergence

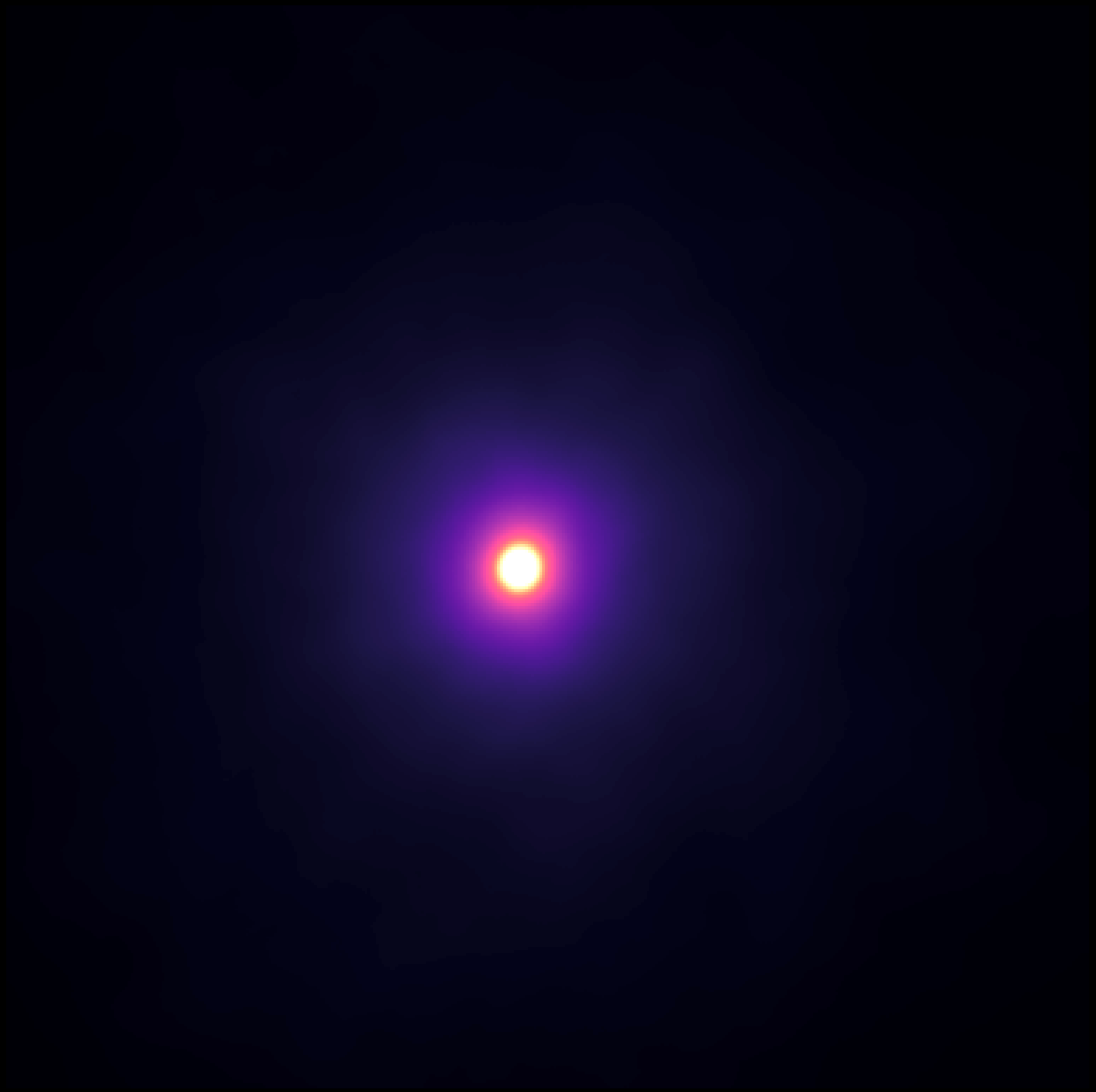
$$P(s)$$



The prior

Convergence

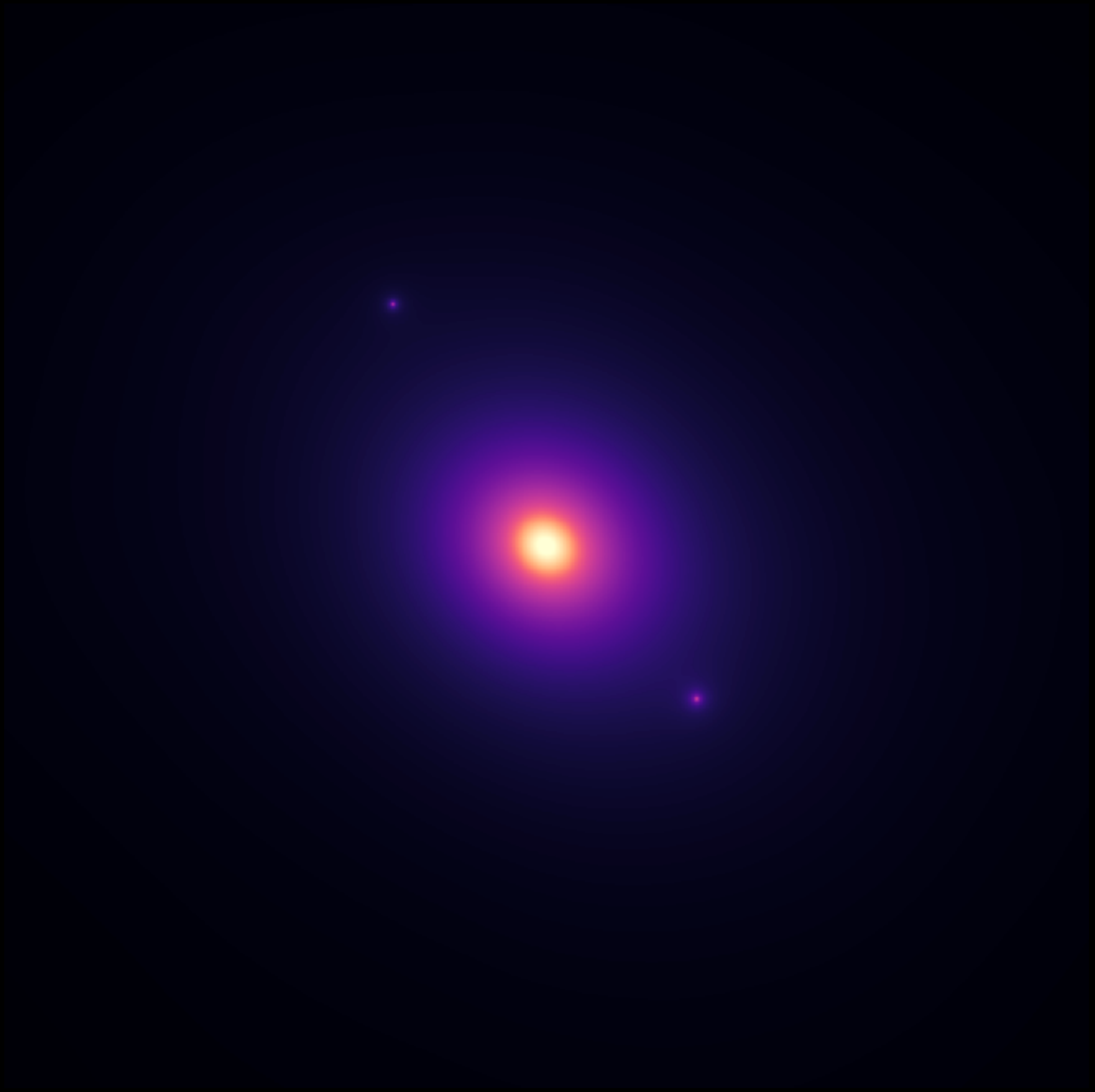
$$P(s)$$



The prior

Convergence

$$P(s)$$



The likelihood

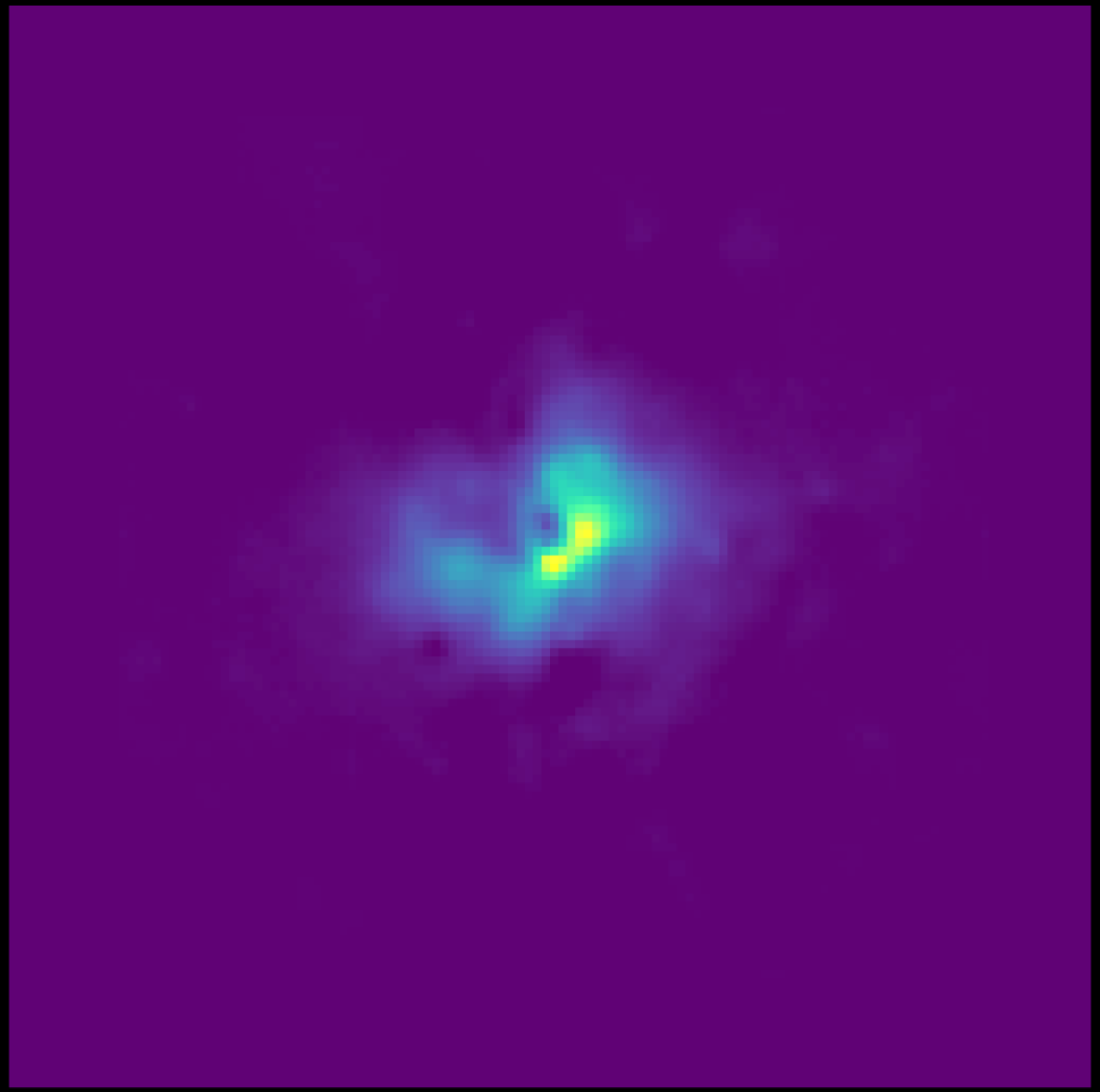
$$P(s \mid d) = \frac{P(d \mid s)\, P(s)}{P(d)}$$

The likelihood

$$P(s \mid d) = \frac{P(d \mid s)\, P(s)}{P(d)}$$

The likelihood

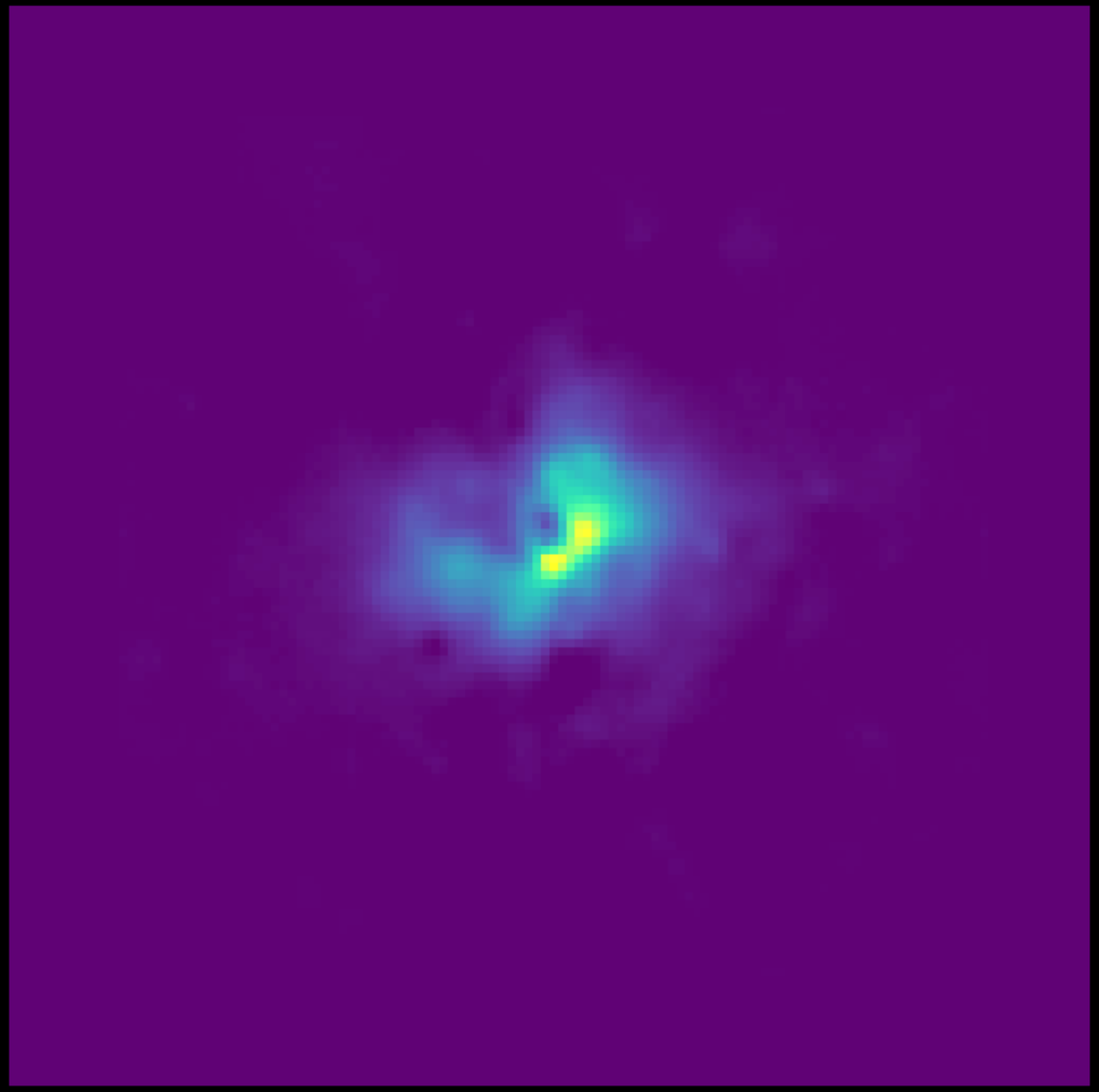
Strong lensing response



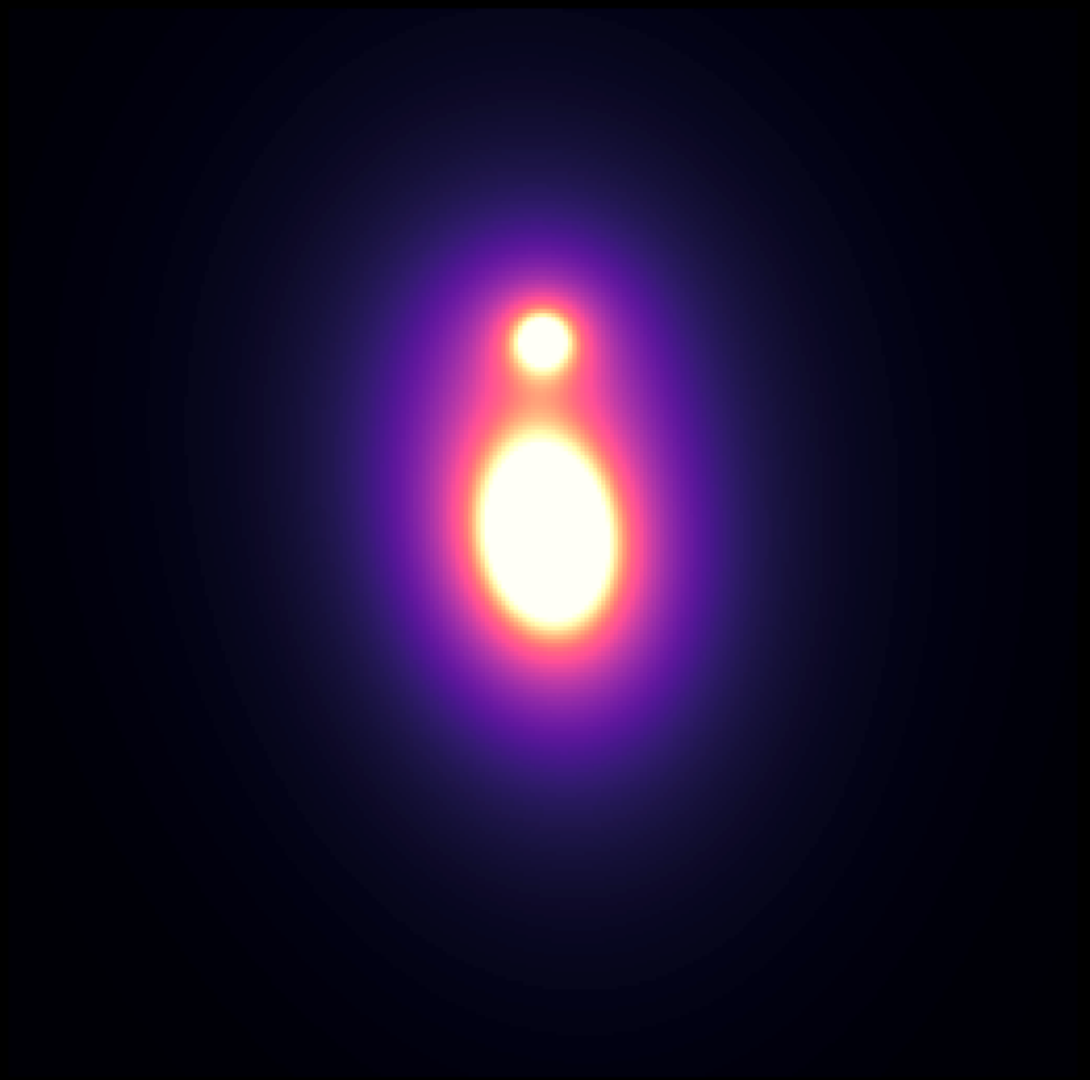
Source galaxy

The likelihood

Strong lensing response



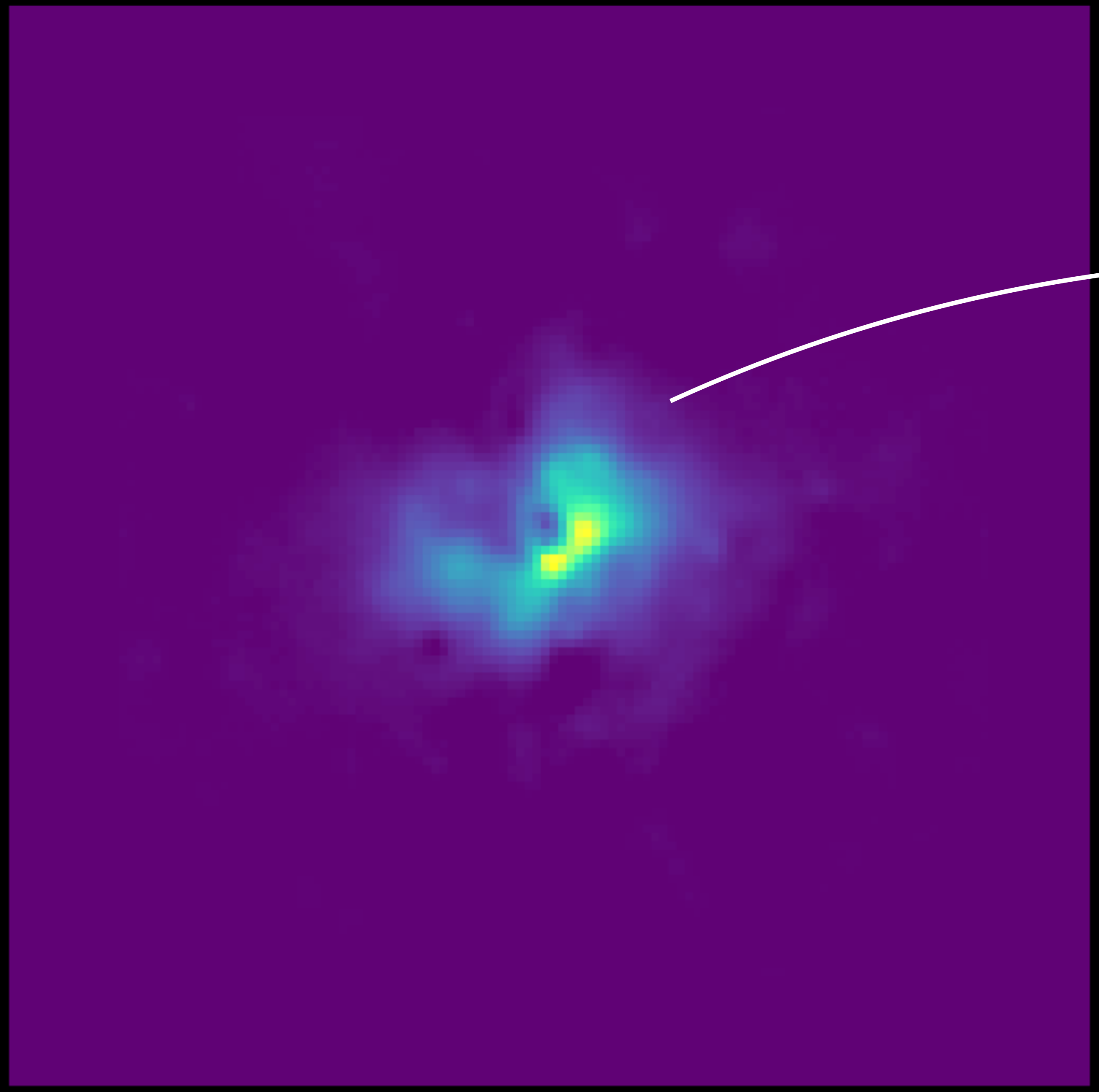
Source galaxy



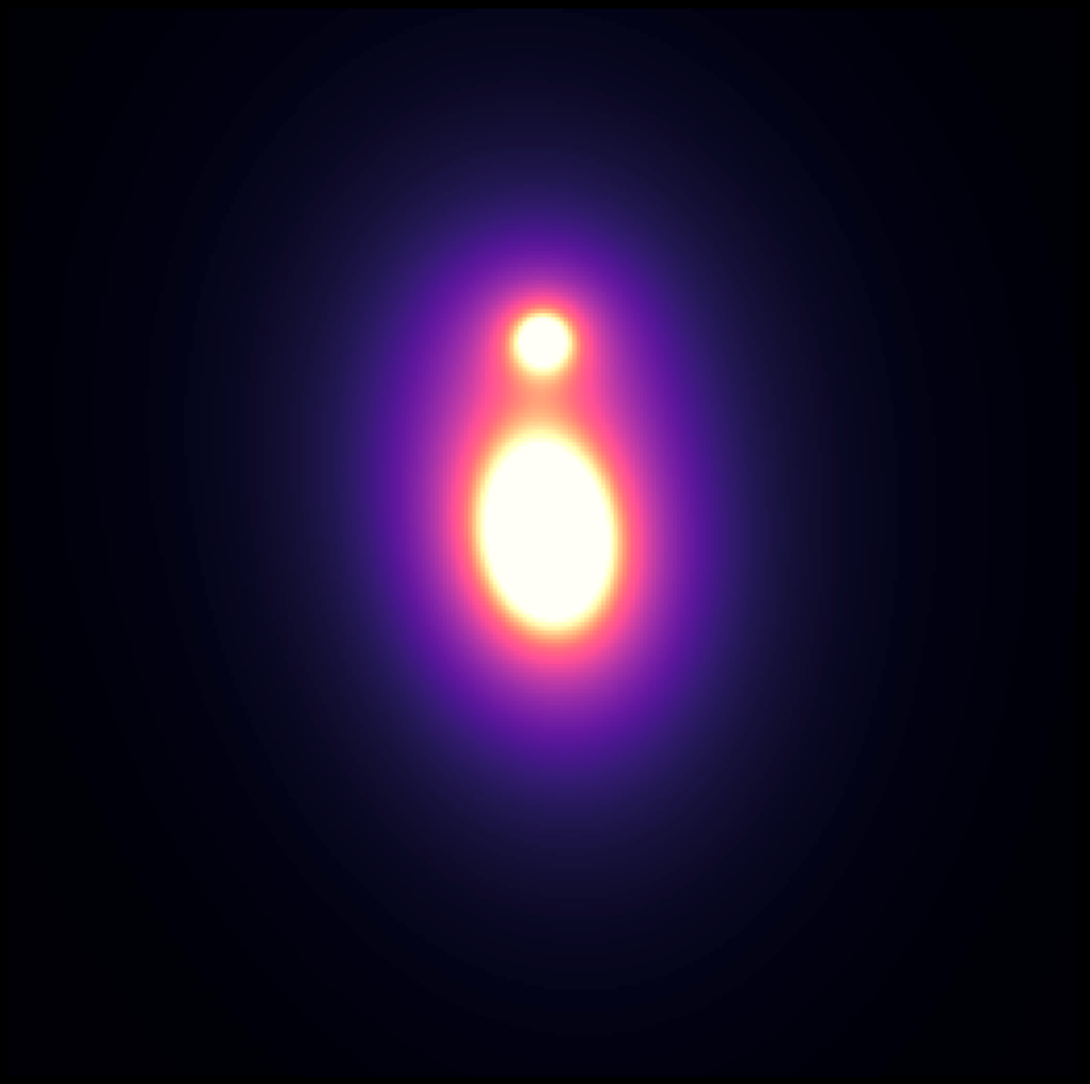
DM halo mass profile

The likelihood

Strong lensing response



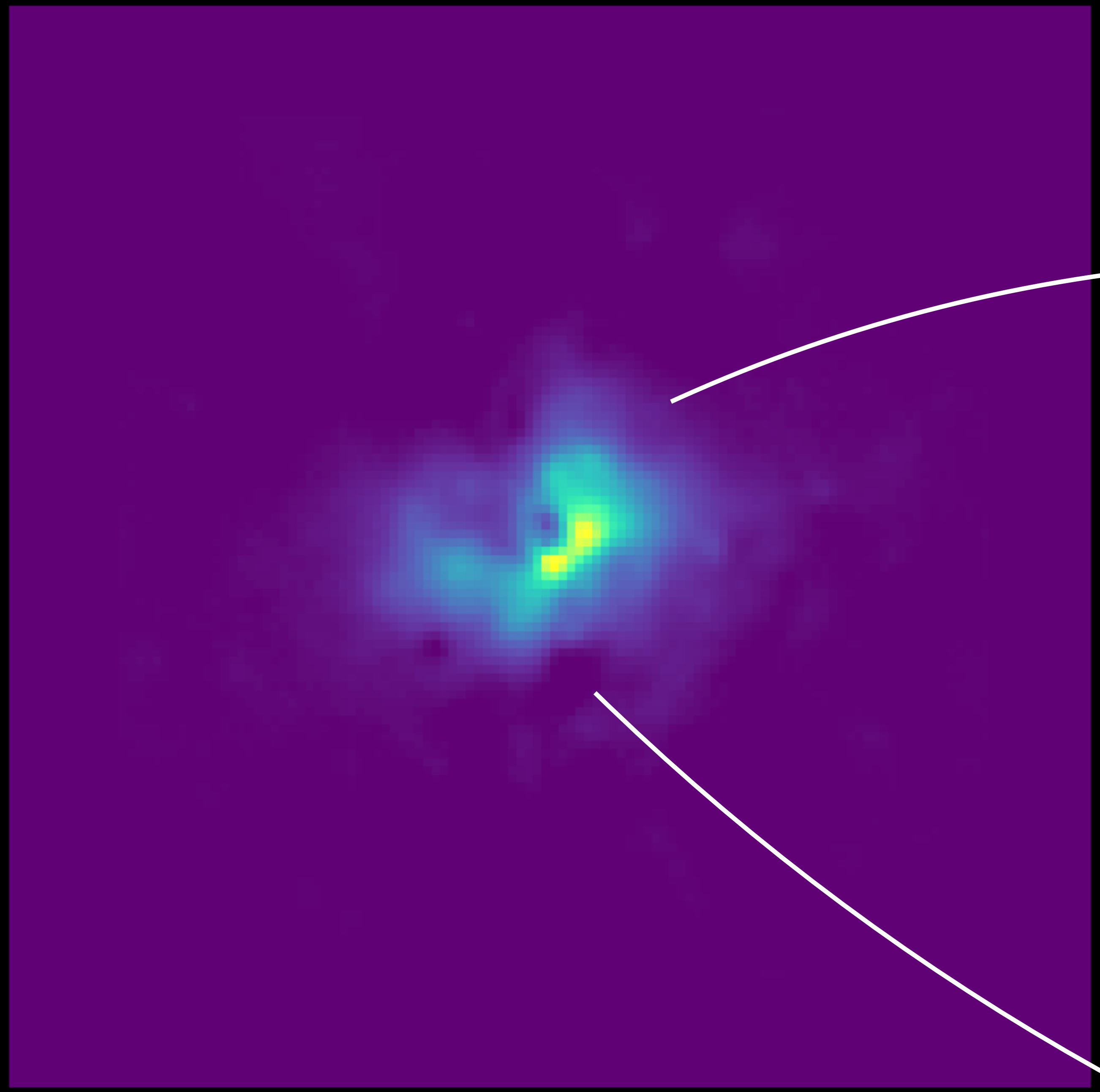
Source galaxy



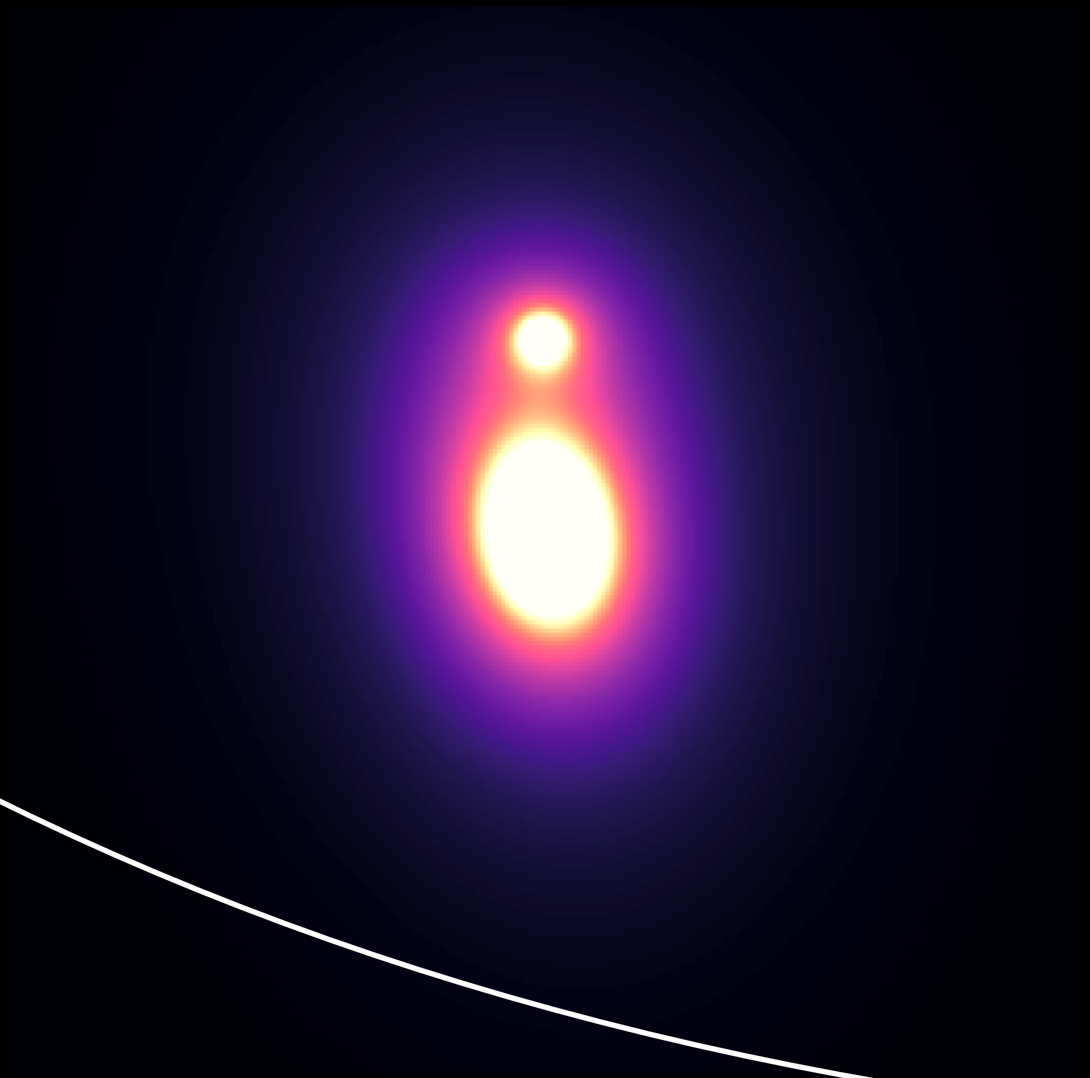
DM halo mass profile

The likelihood

Strong lensing response



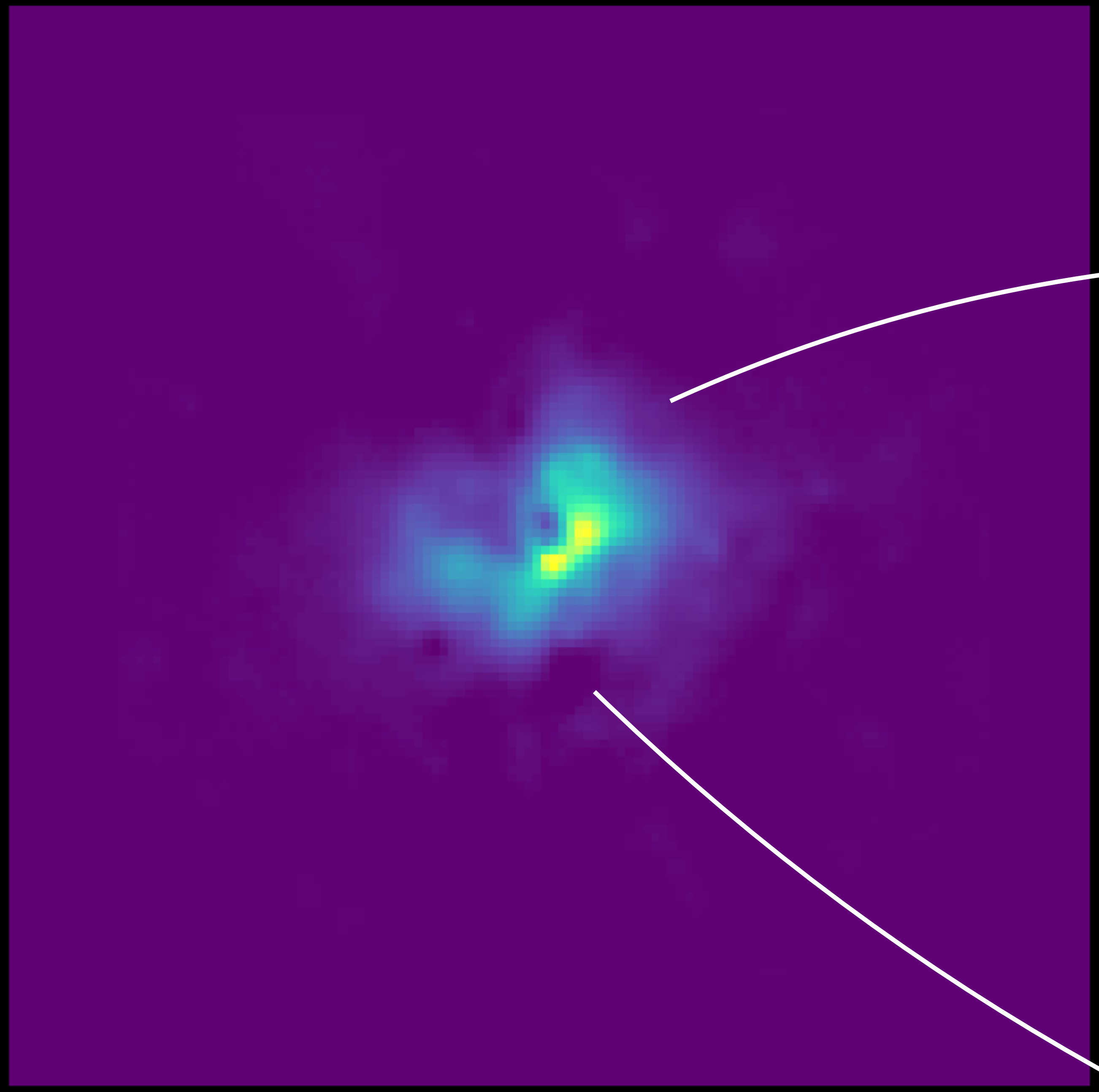
Source galaxy



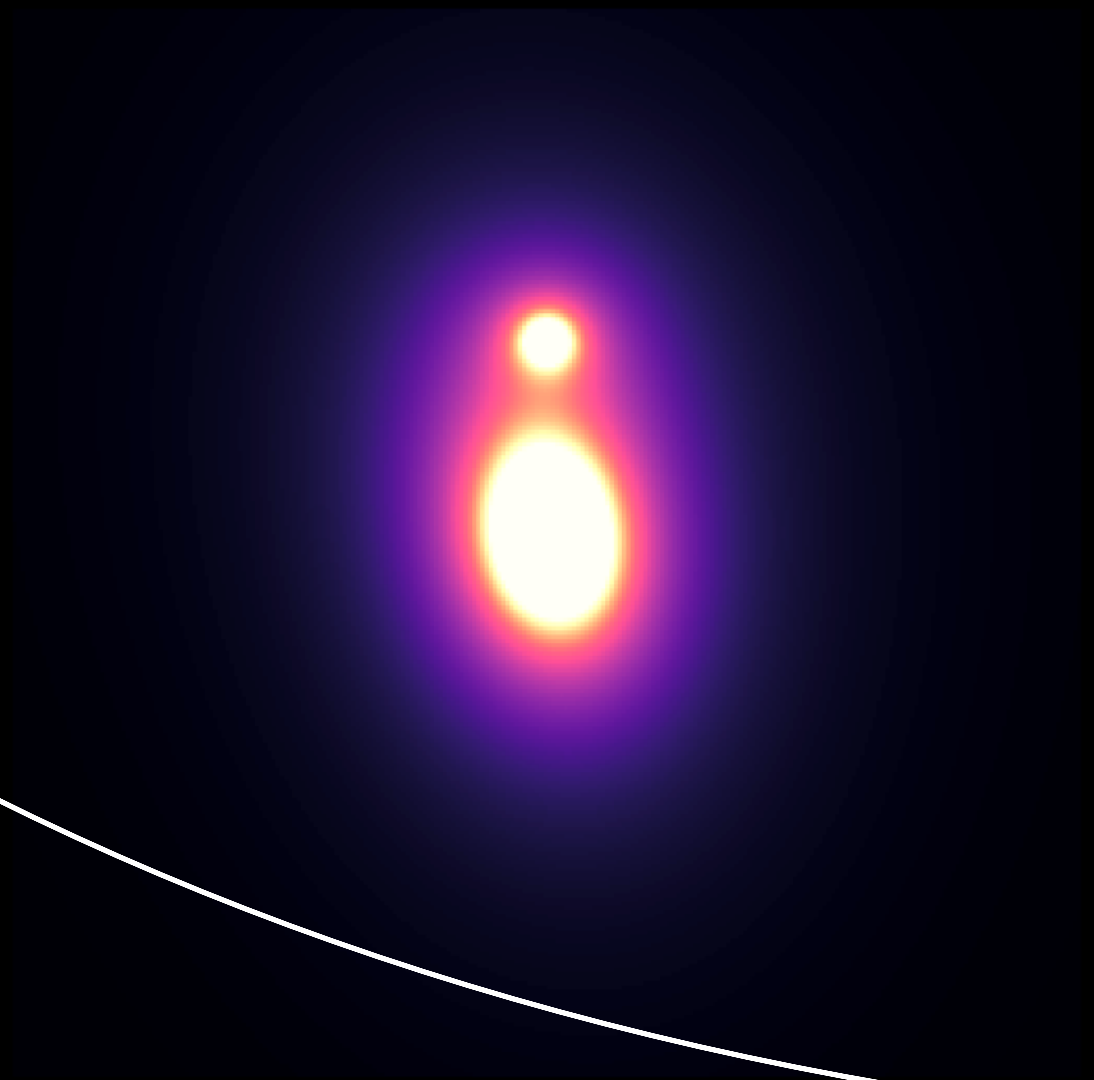
DM halo mass profile

The likelihood

Strong lensing response



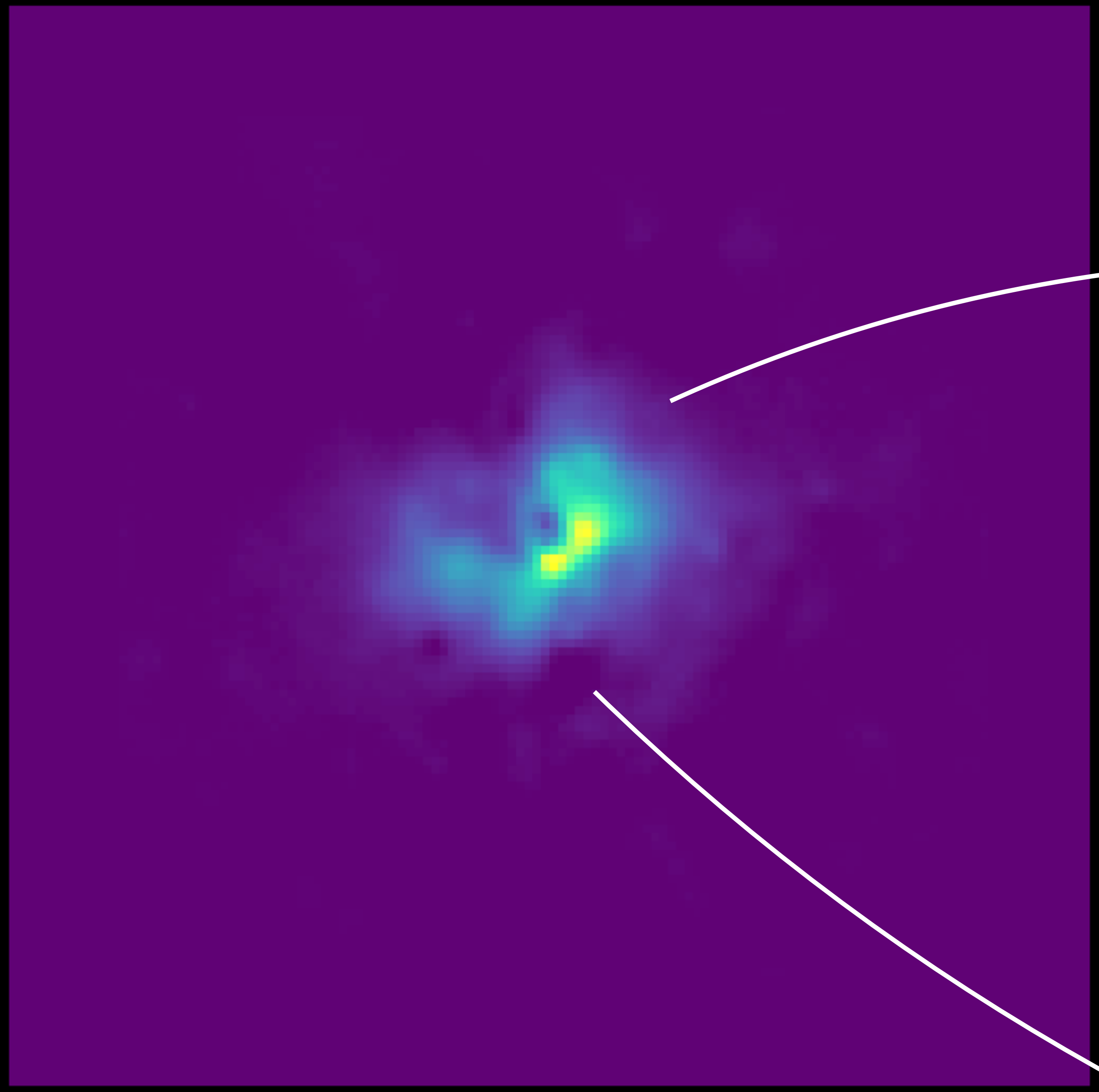
Source galaxy



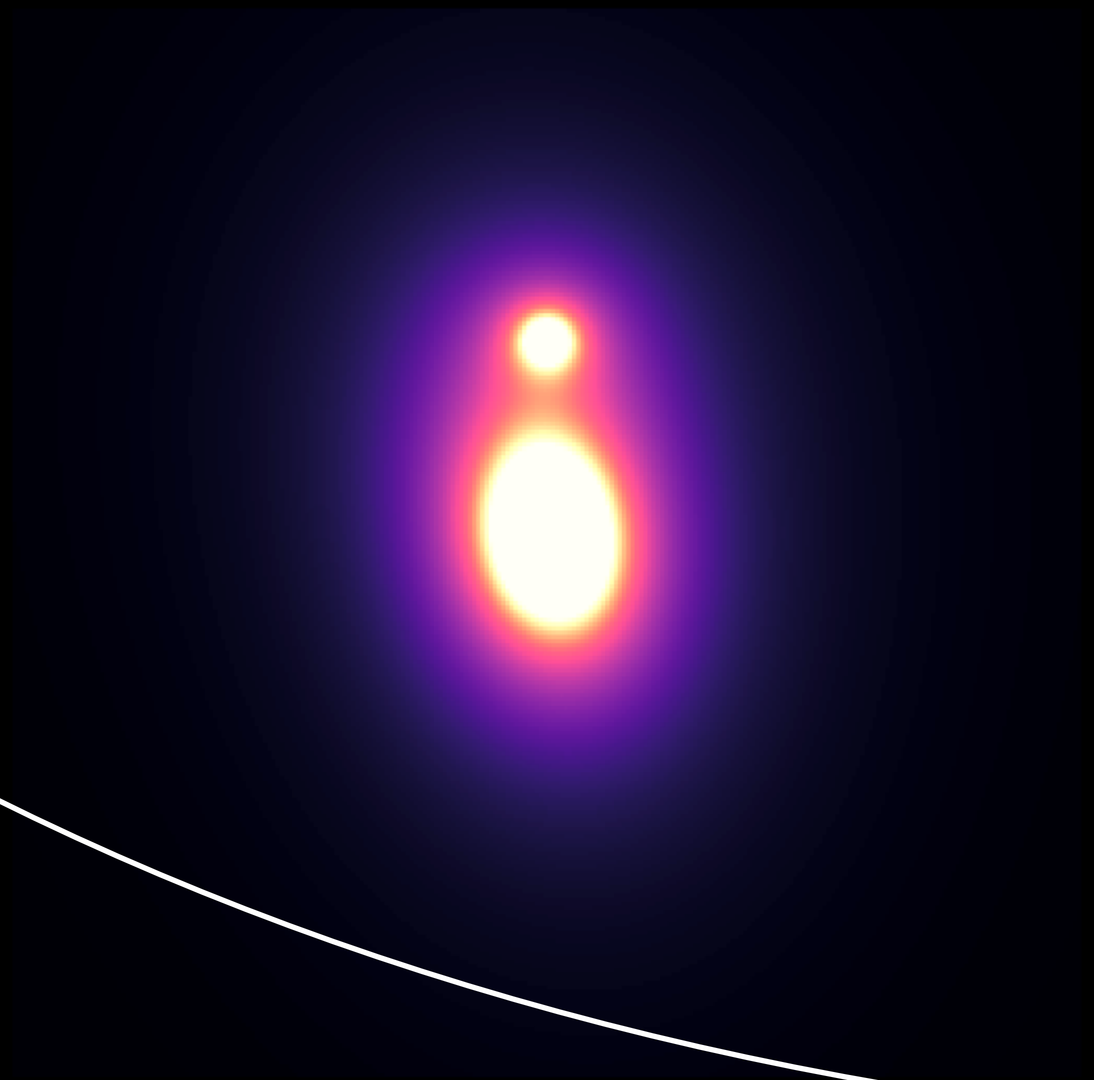
DM halo mass profile

The likelihood

Strong lensing response



Source galaxy

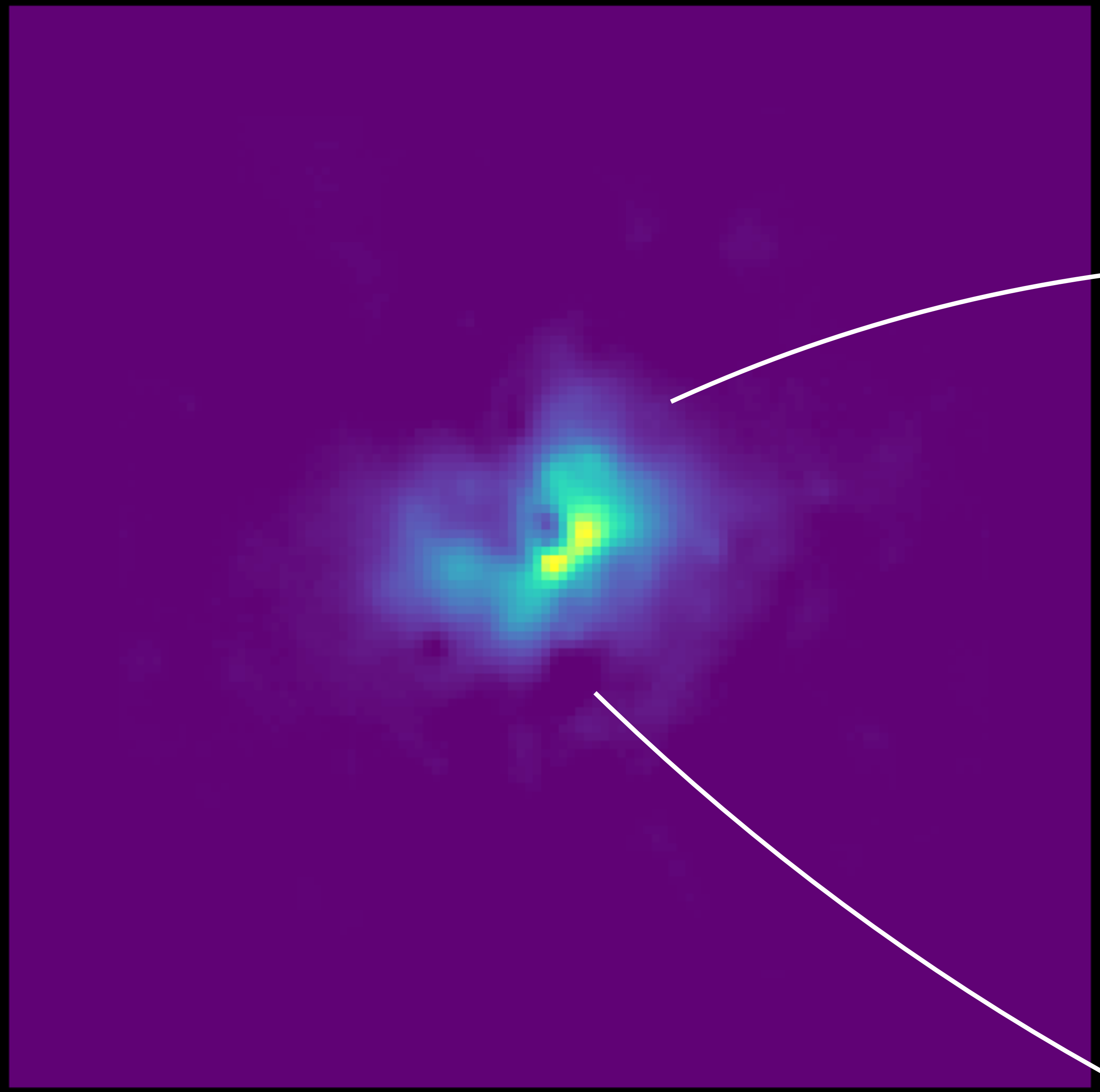


DM halo mass profile

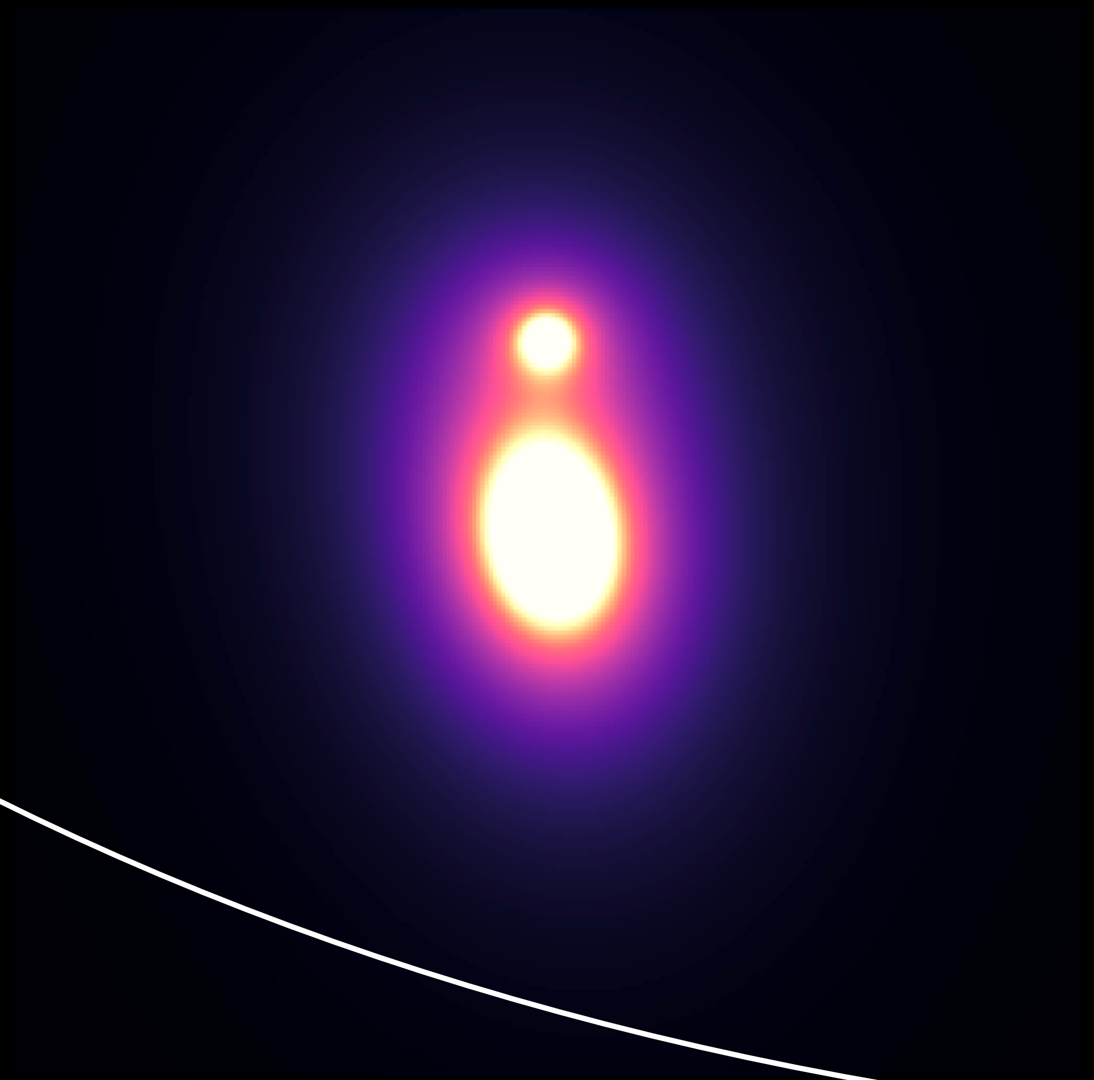
$$\nabla \cdot \alpha = 2\ \kappa$$

The likelihood

Strong lensing response

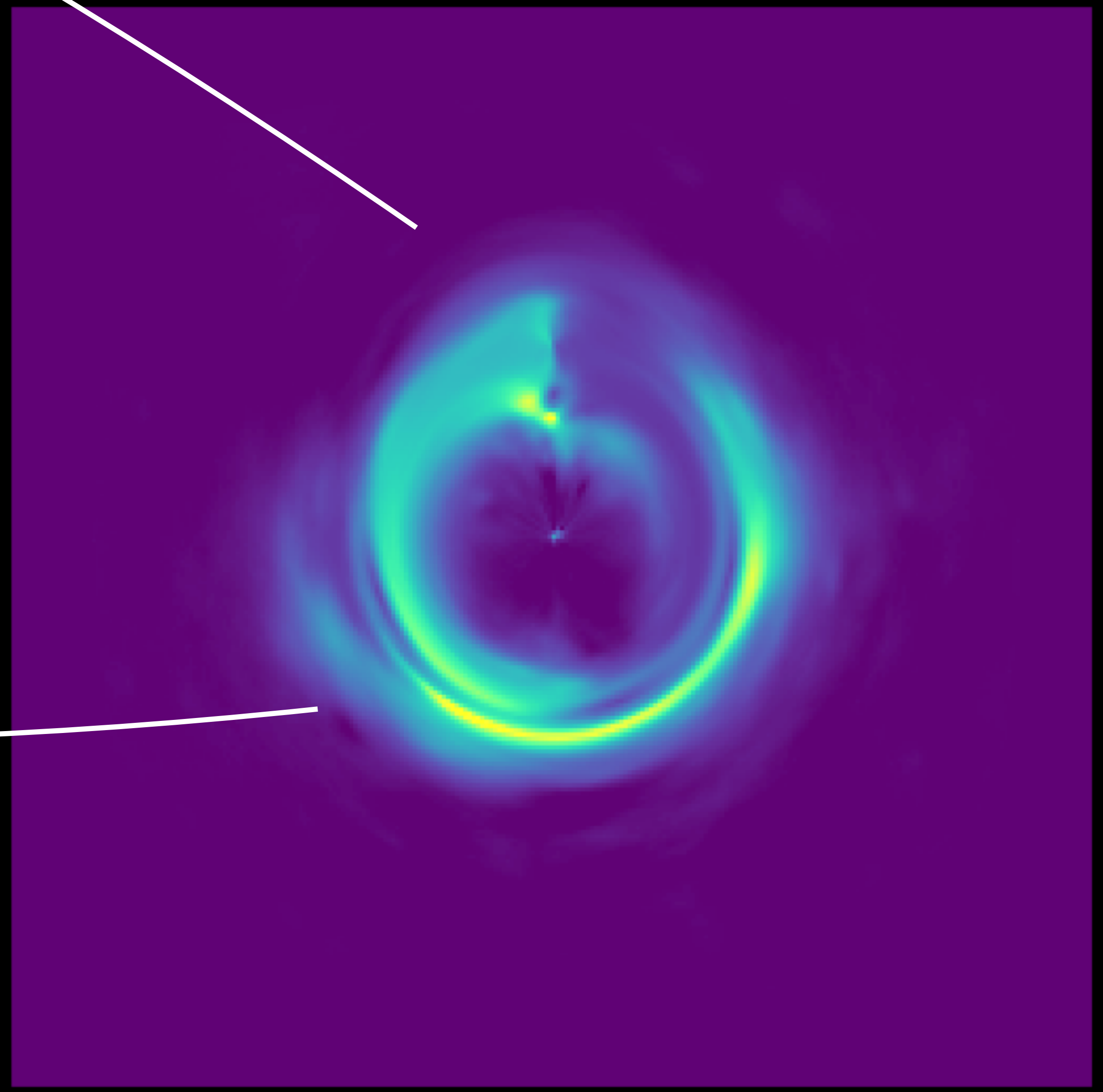


Source galaxy



DM halo mass profile

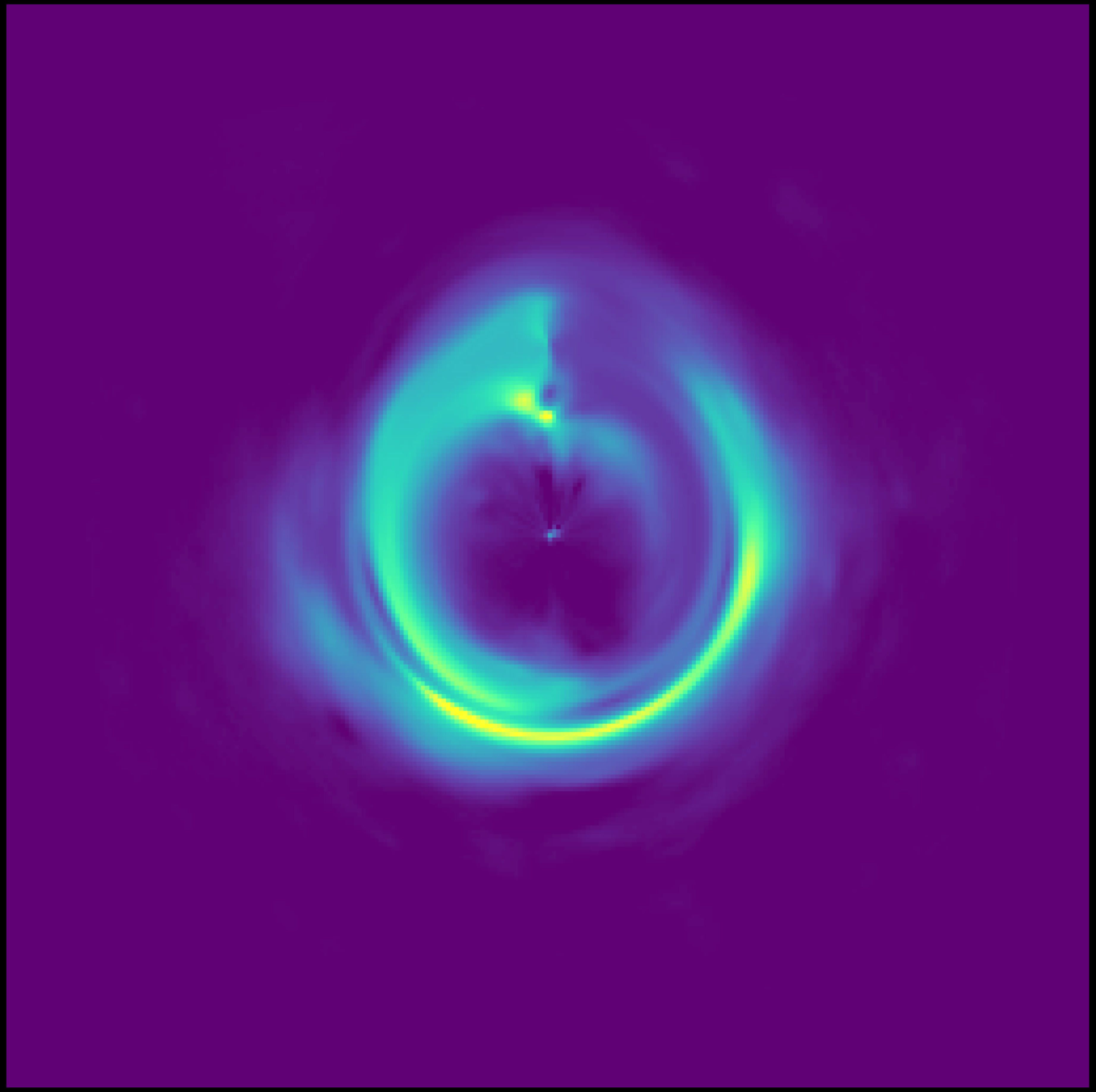
$$\nabla \cdot \alpha = 2\,\kappa$$



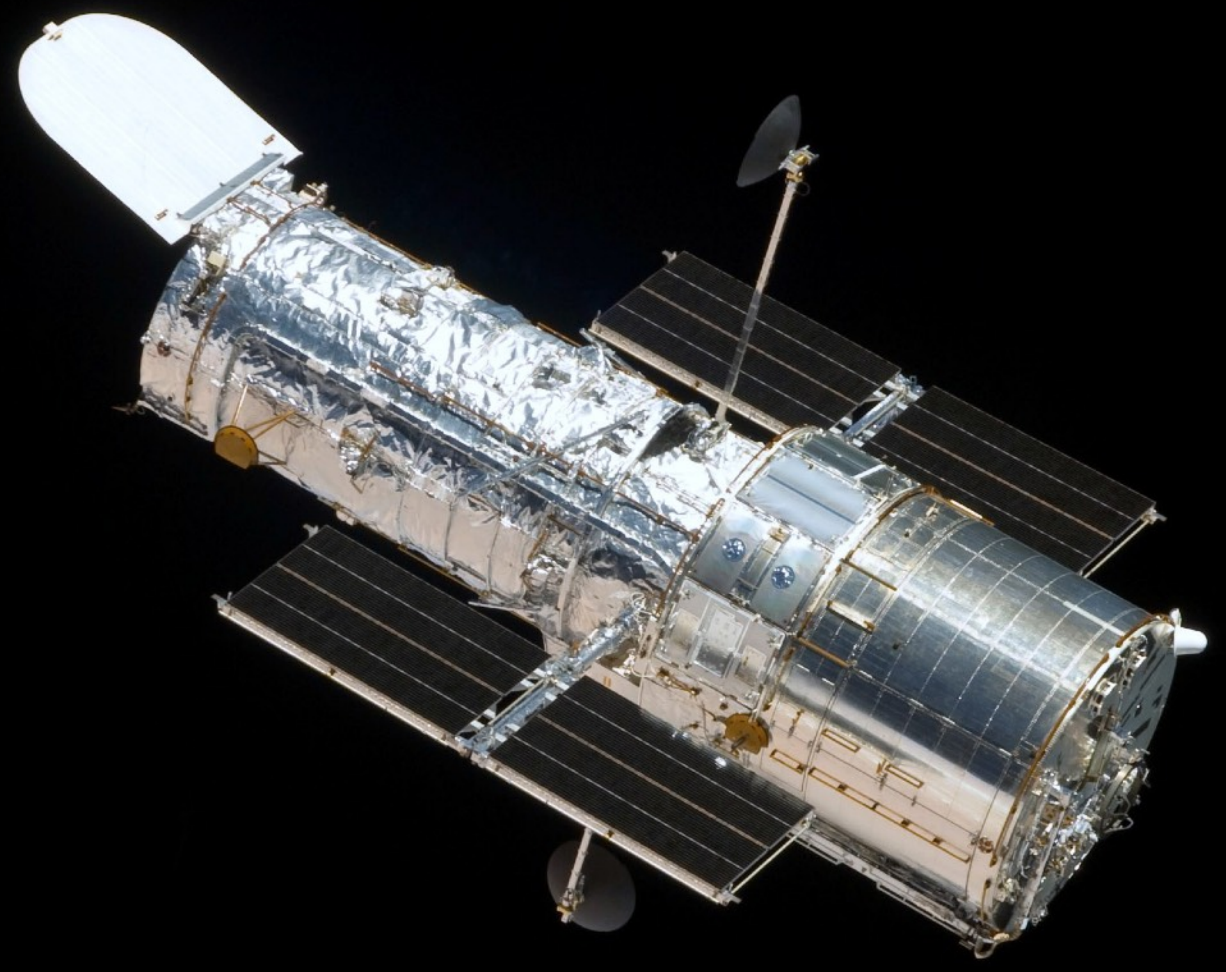
$$y = x - \alpha(\mathbf{x})$$

The likelihood

Instrument response



Lensed galaxy signal

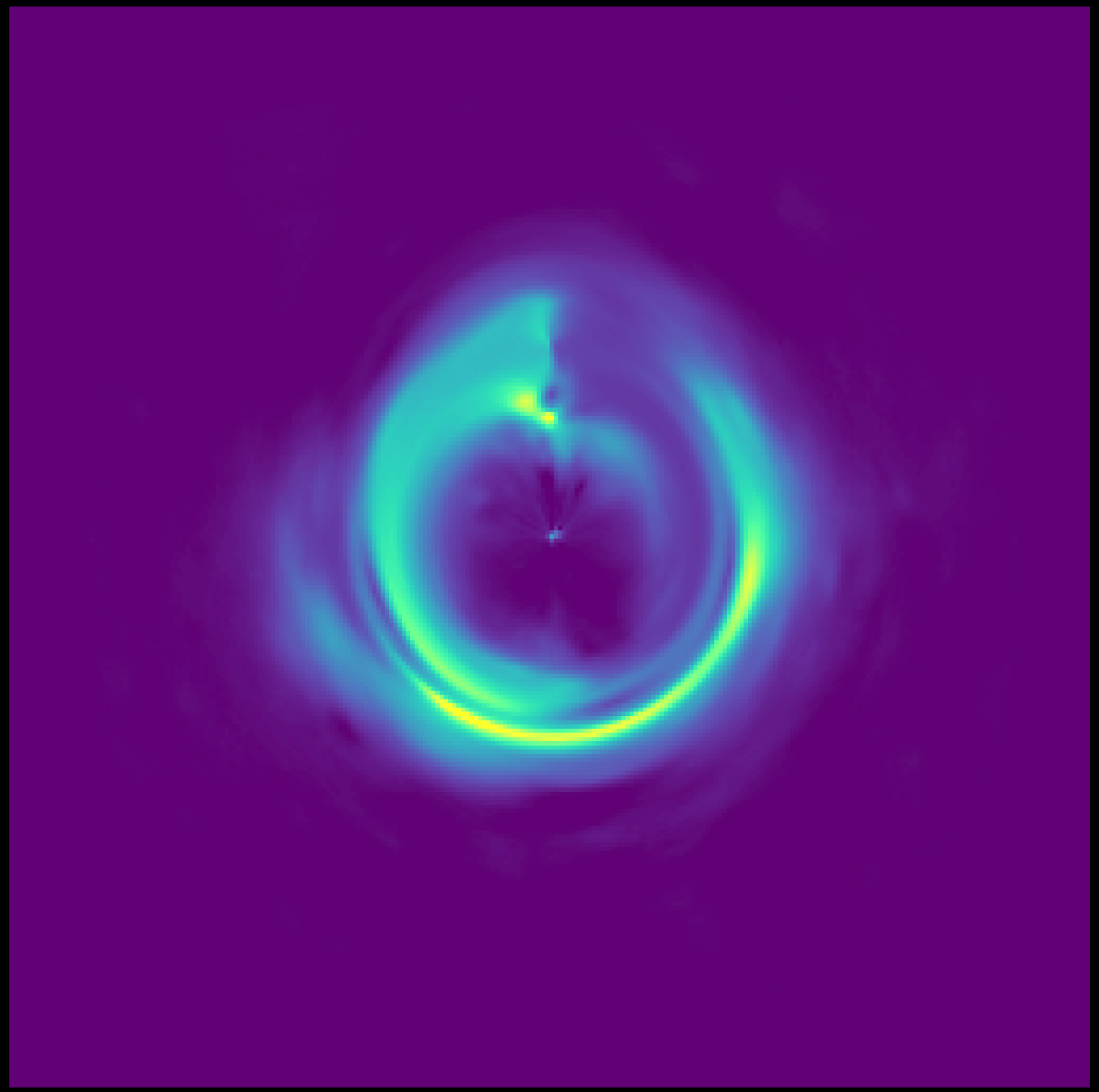


Hubble Space Telescope

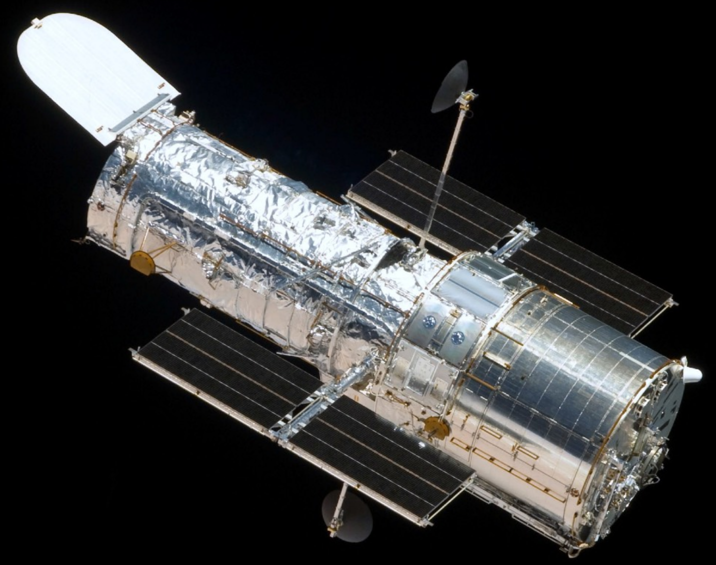
Credits @ Ruffnax (Crew of STS-125)

The likelihood

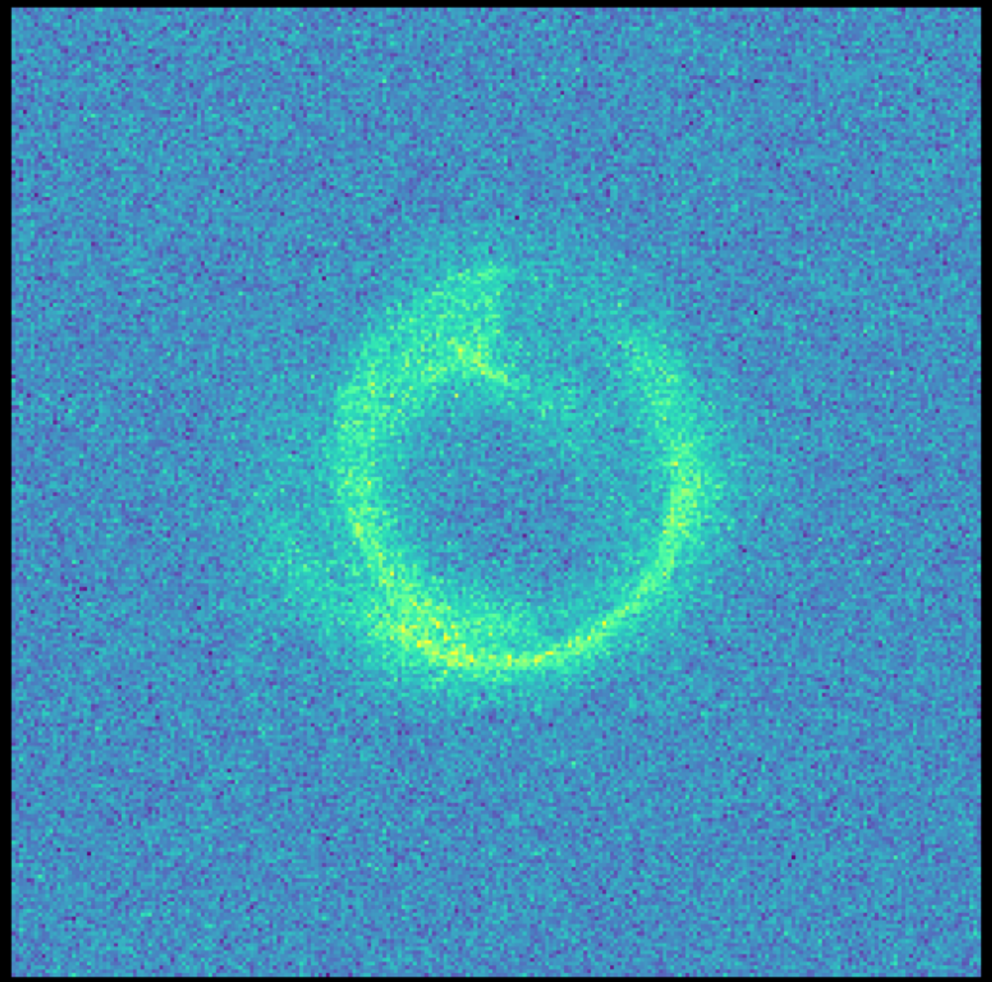
Instrument response



Lensed galaxy signal



Hubble Space Telescope

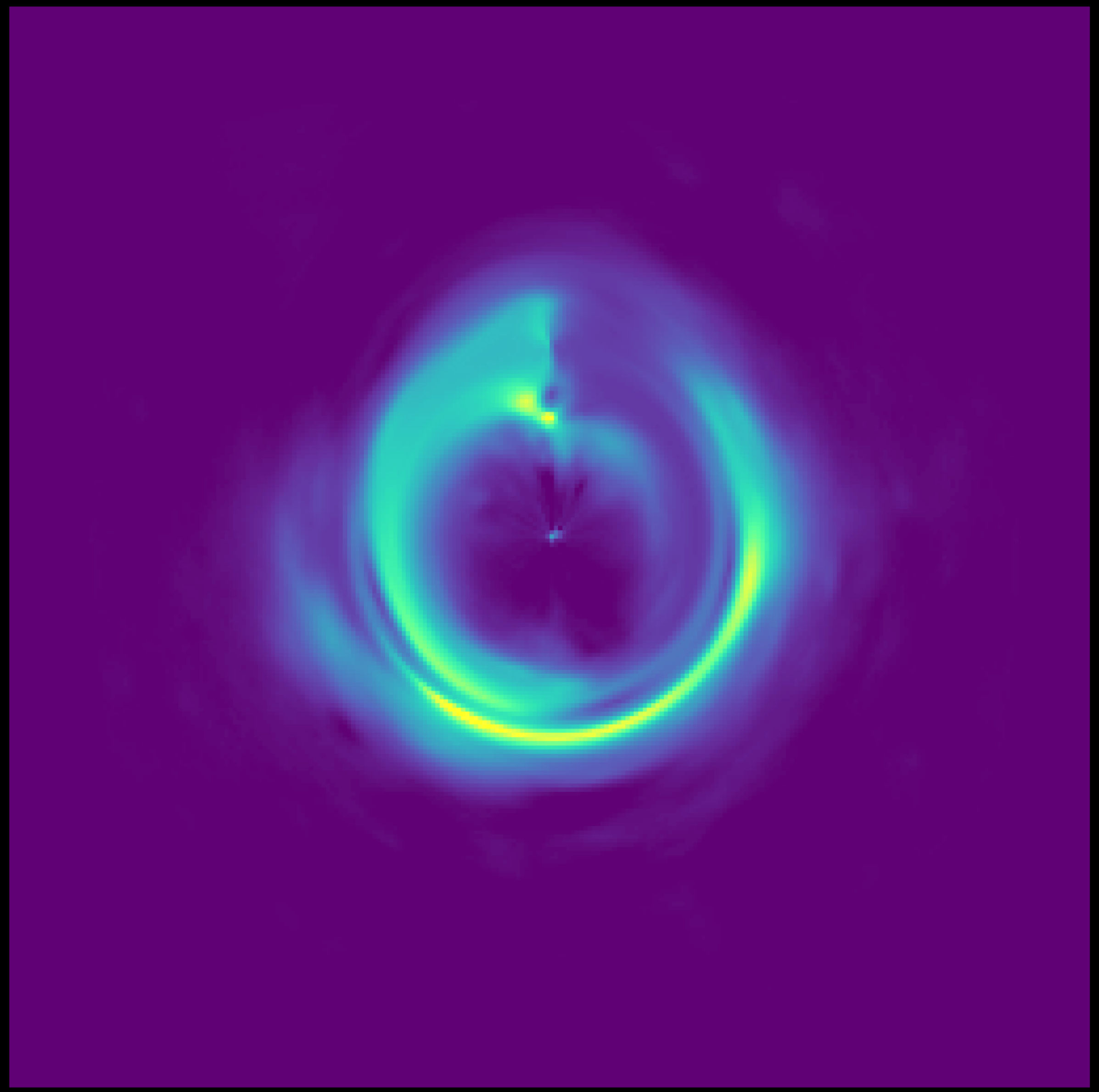


Data on Earth

Credits @ Ruffnax (Crew of STS-125)

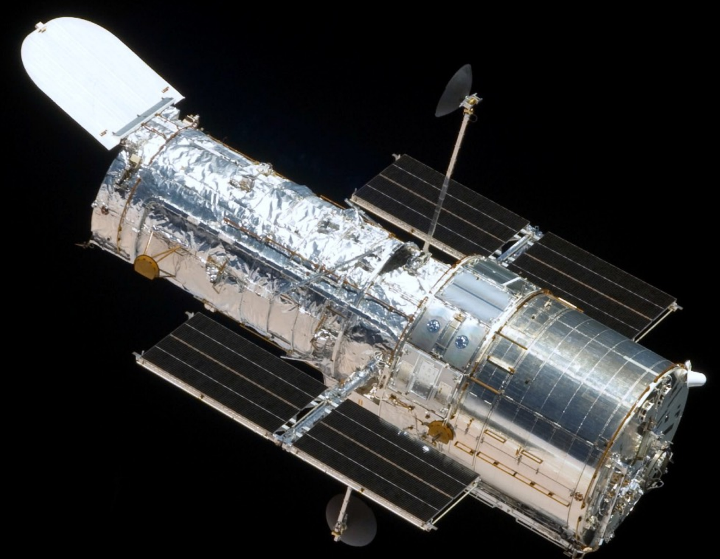
The likelihood

Instrument response



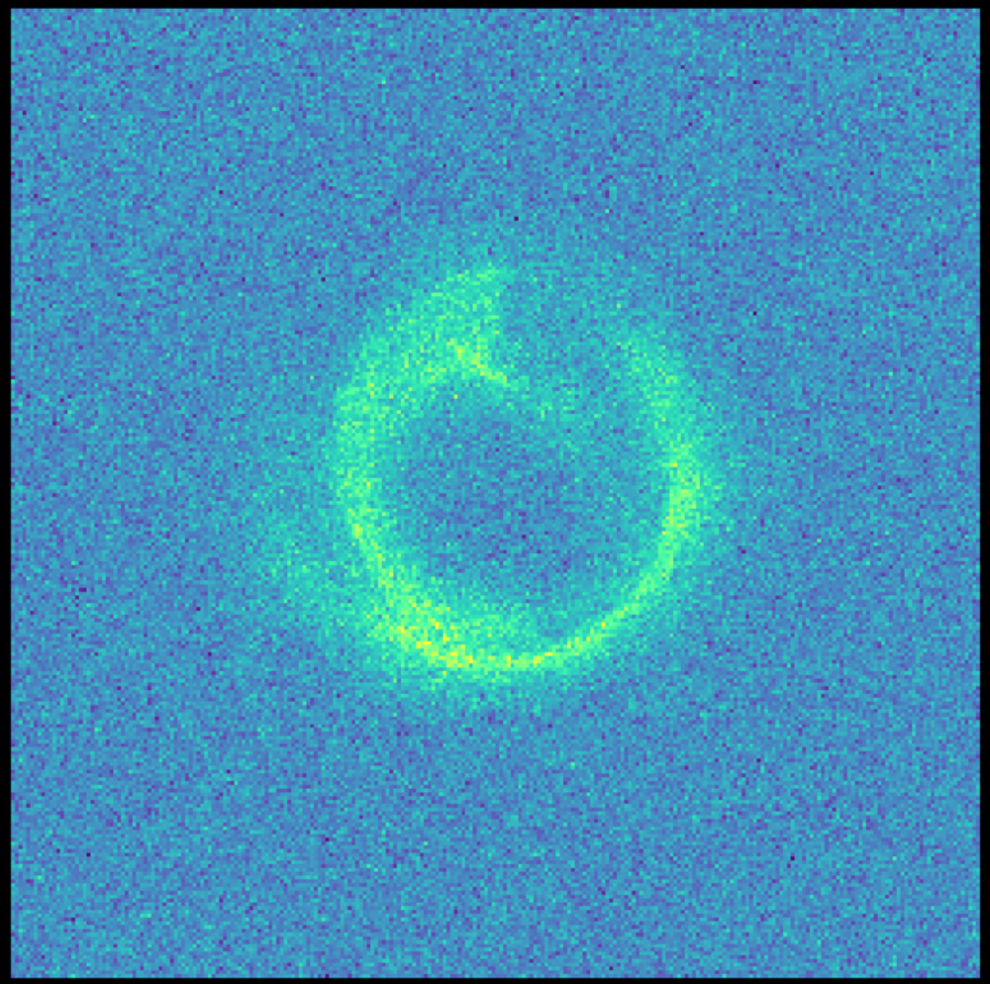
Lensed galaxy signal

$$L(s)$$

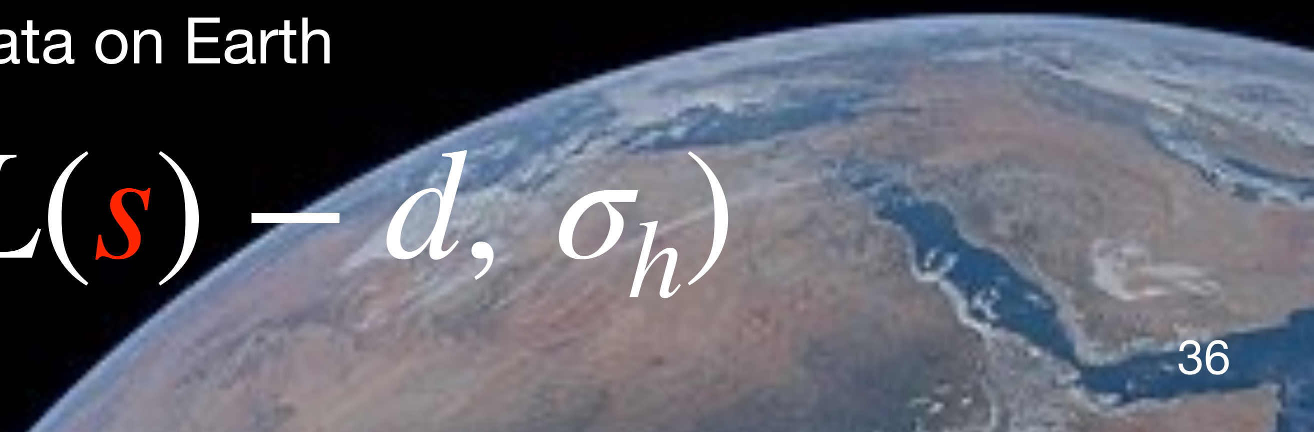


Hubble Space Telescope

$$RL(s)$$



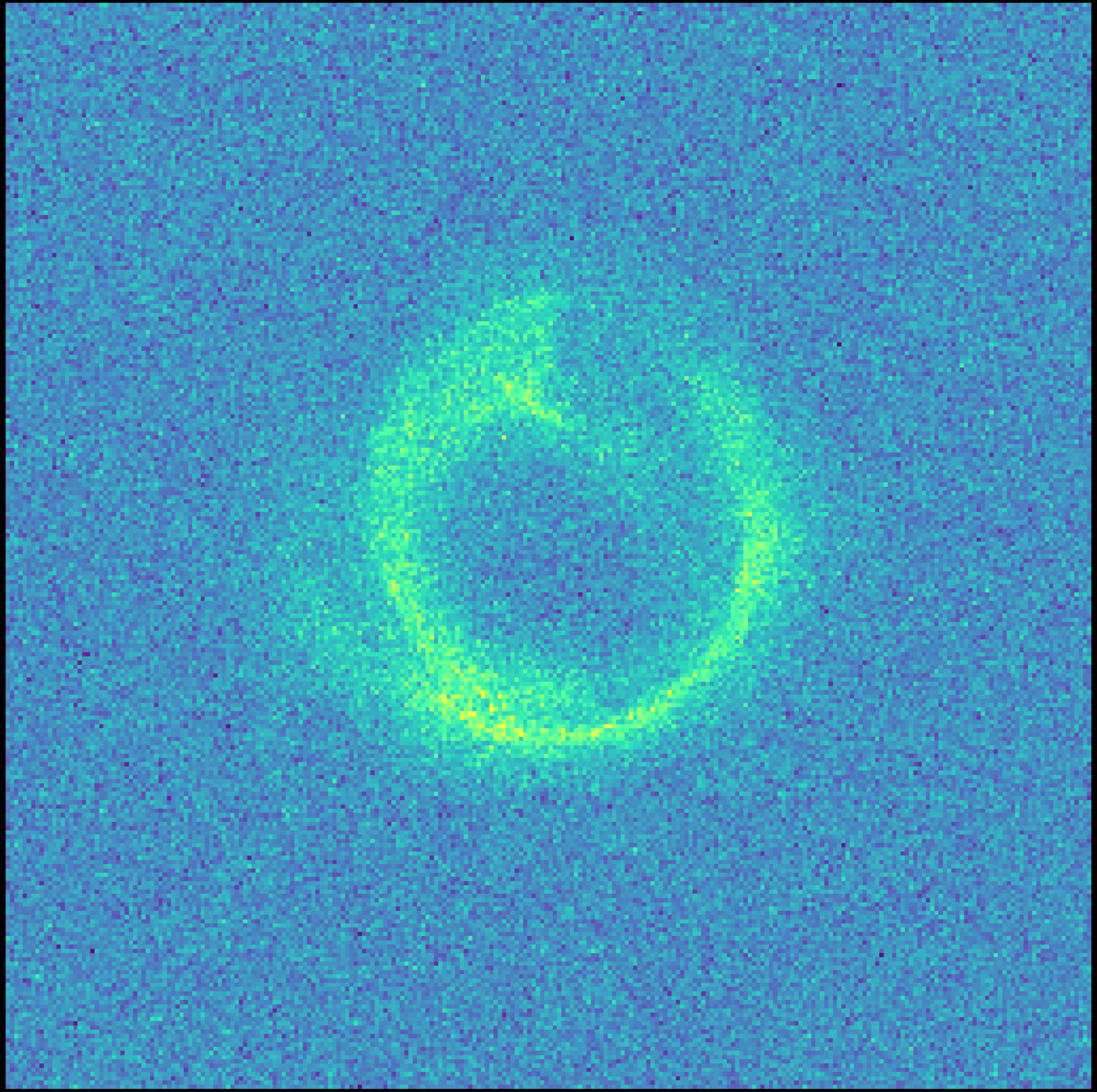
Data on Earth



$$P(d \mid s) = \mathcal{G}(RL(s) - d, \sigma_h)$$

Credits @ Ruffnax (Crew of STS-125)

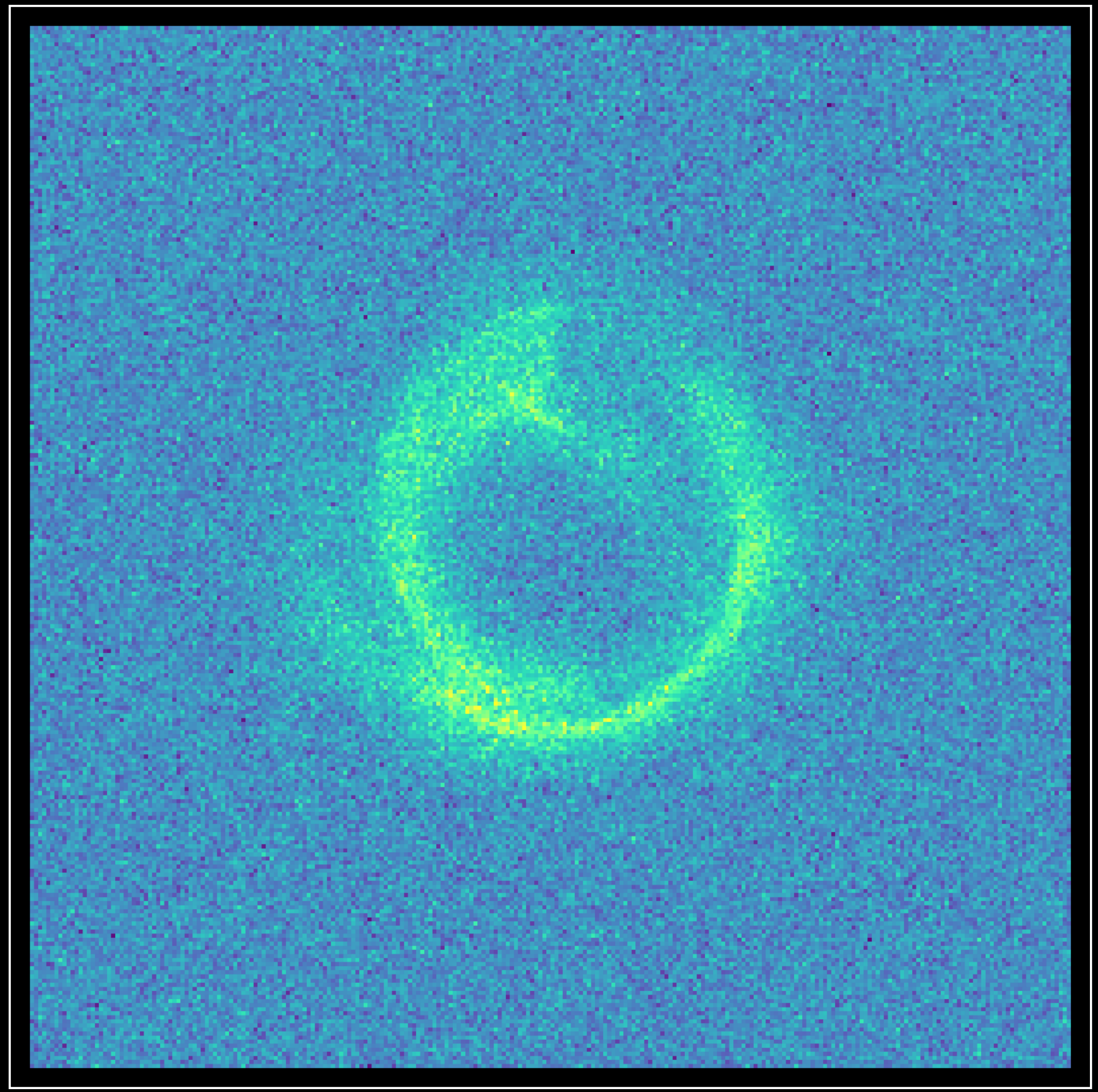
The data



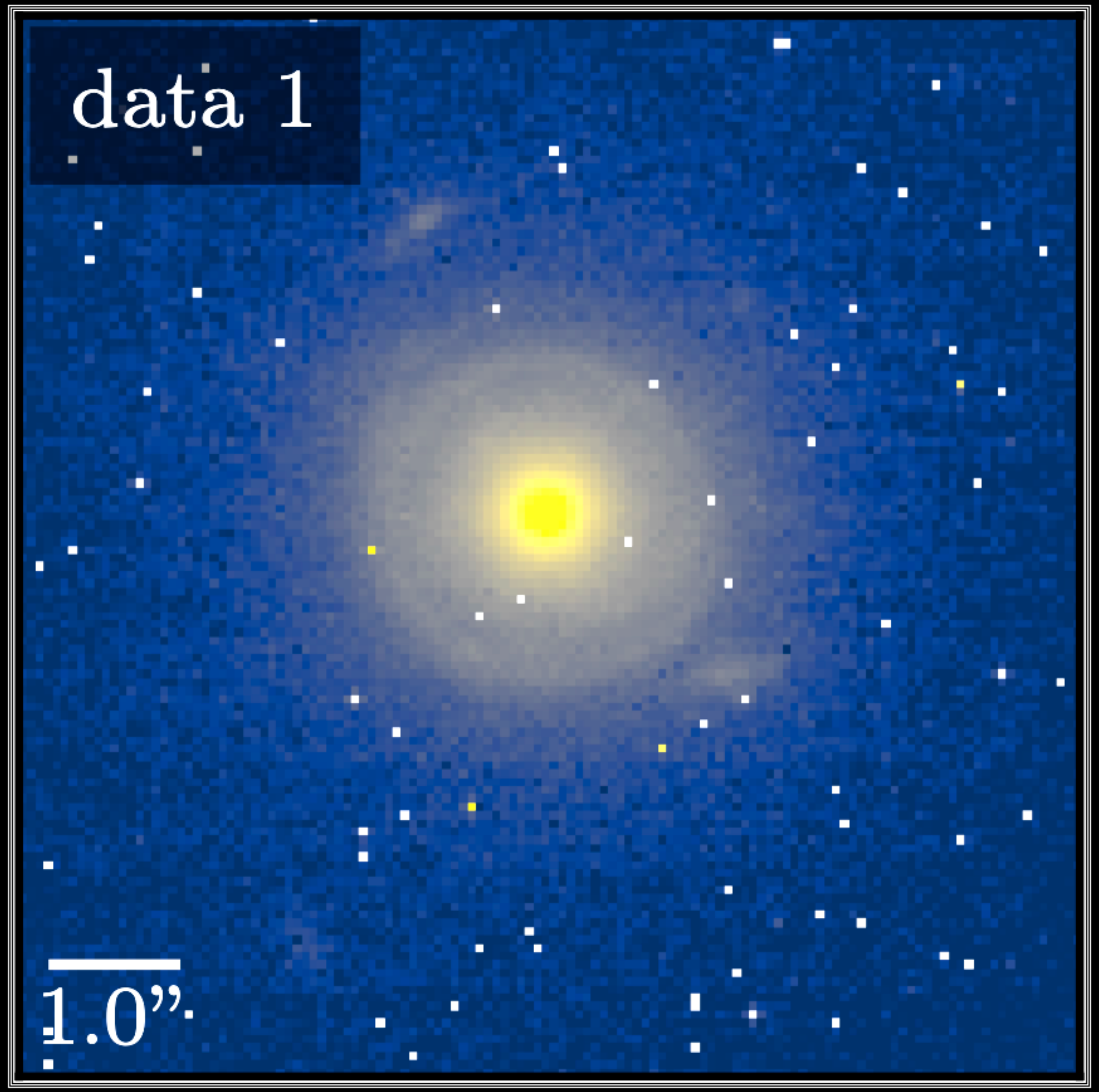
Simulated HST data

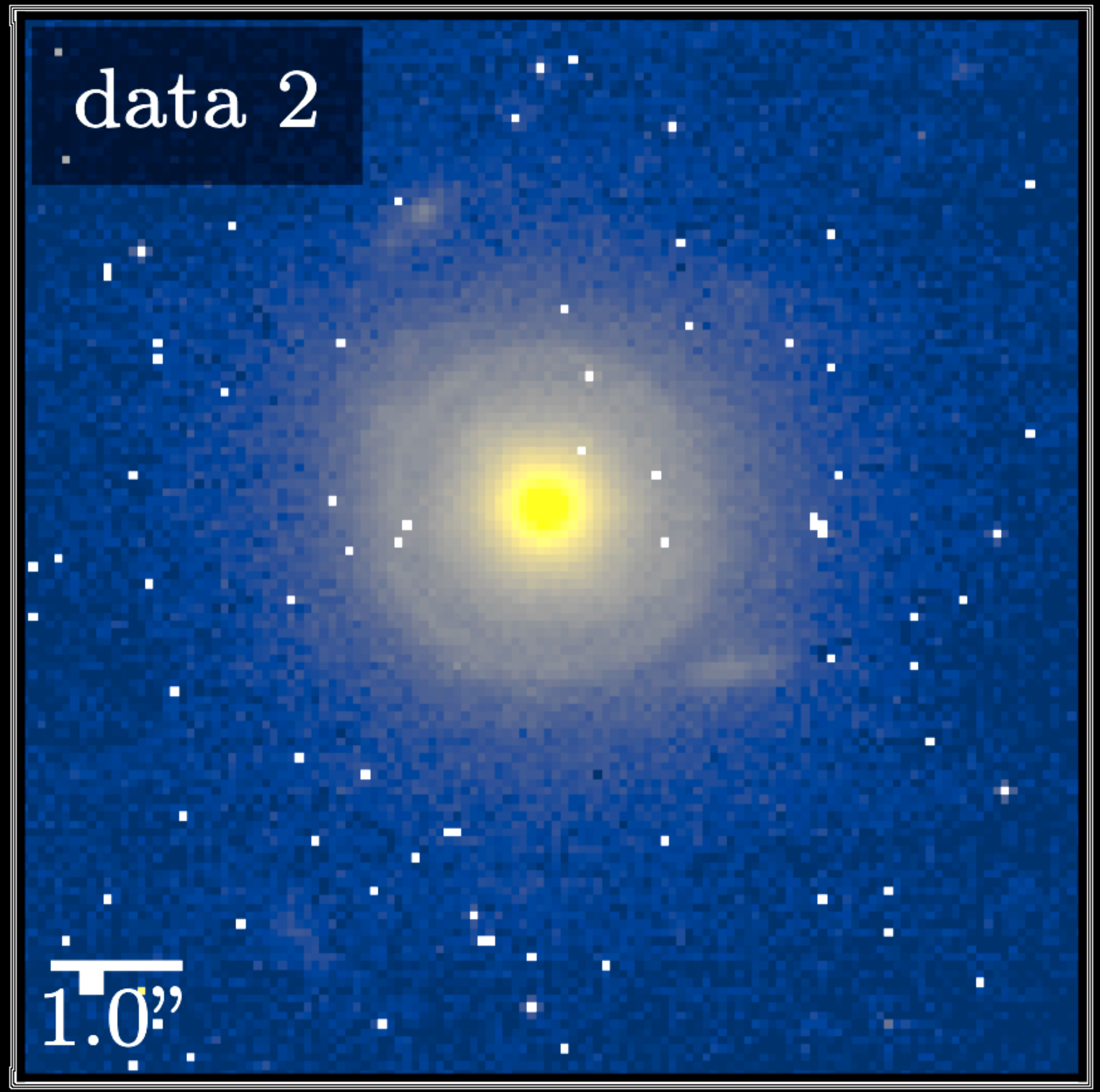
The likelihood

Instrument response



HST synthetic data



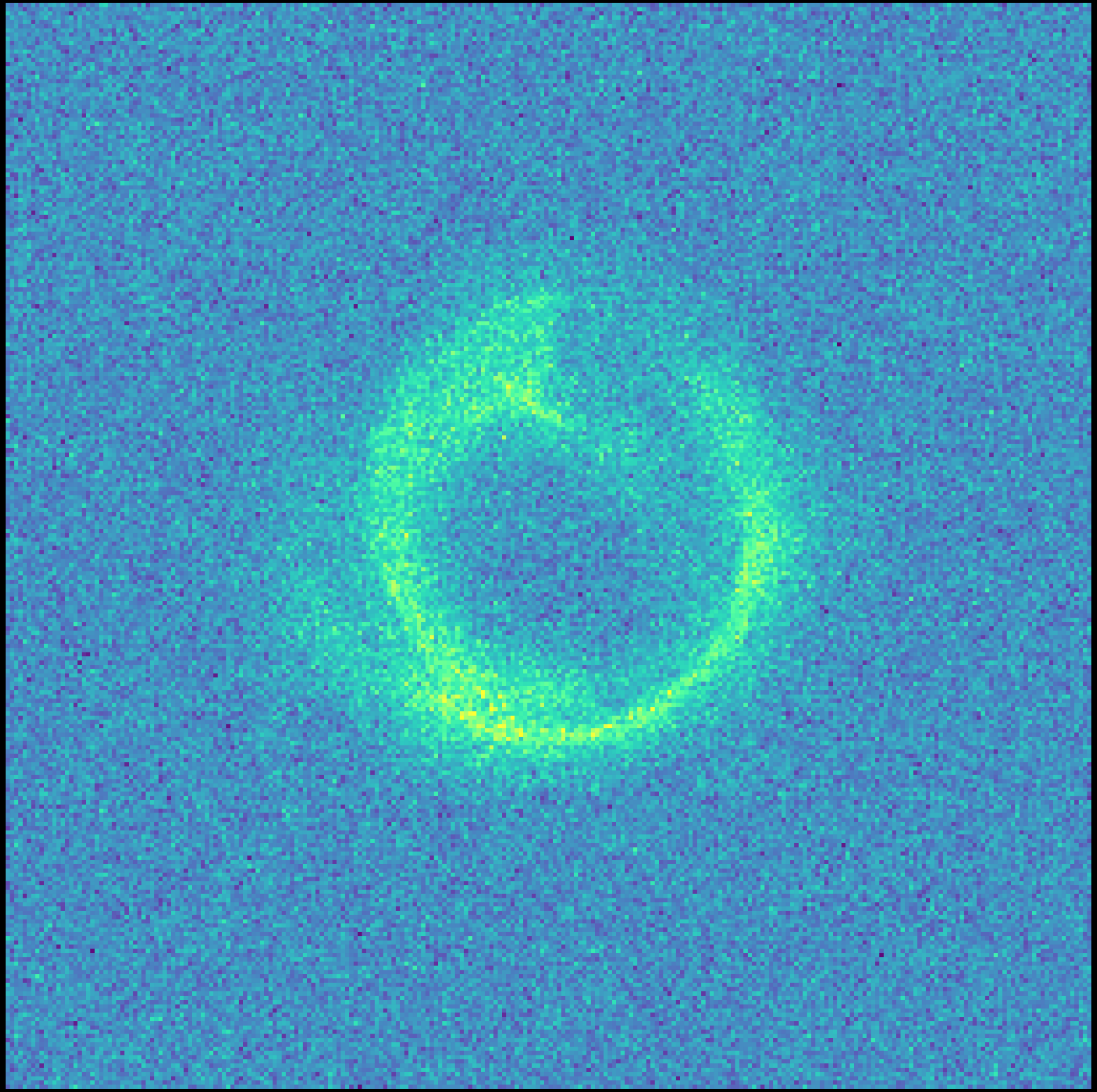


Real JWST lensing data

The Posterior

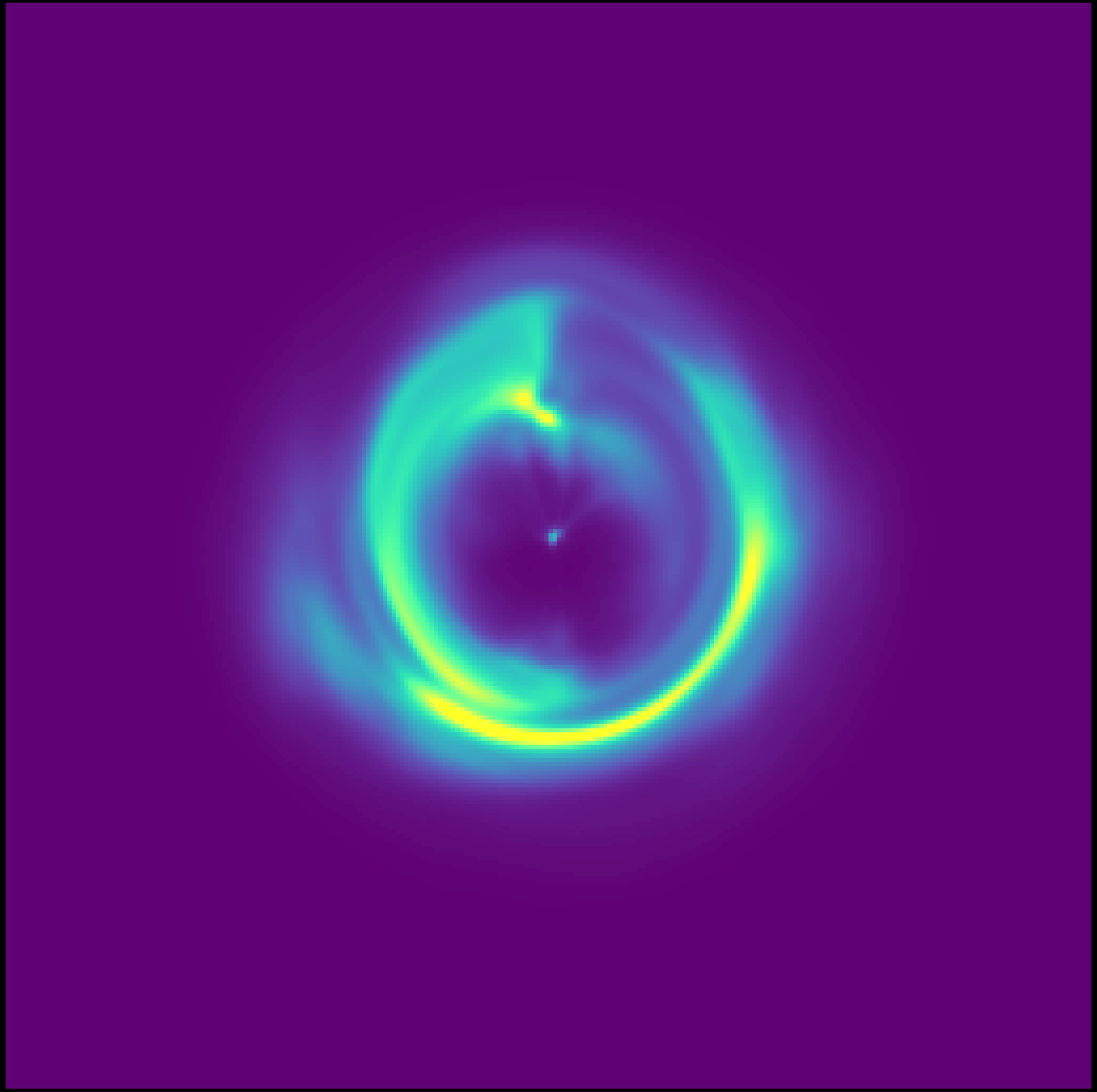
$$P(s \mid d) = \frac{P(d \mid s)\, P(s)}{P(d)}$$

The data



Simulated HST data

The reconstruction

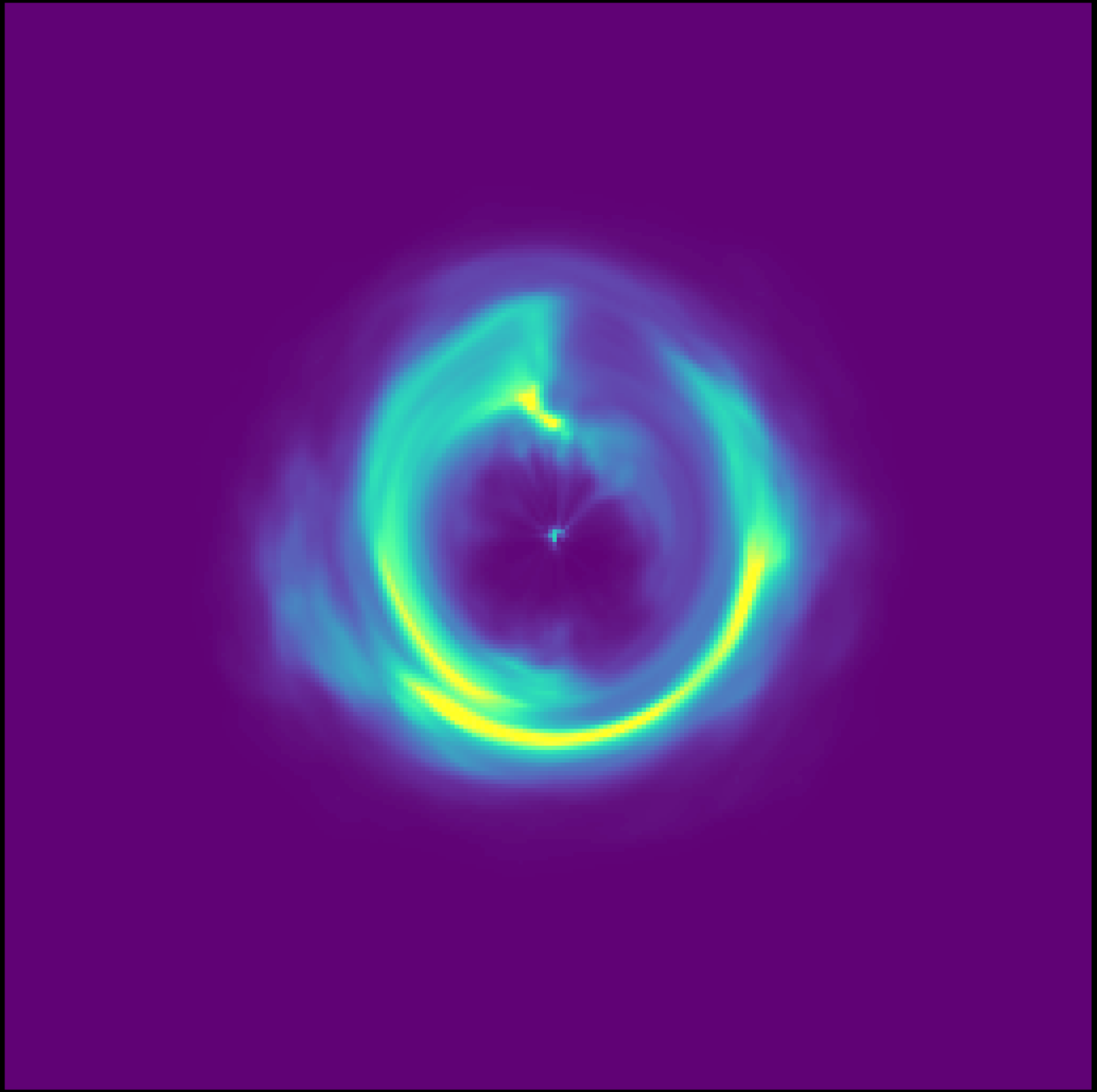


Lensed source light

Inference

Posterior (Ls)

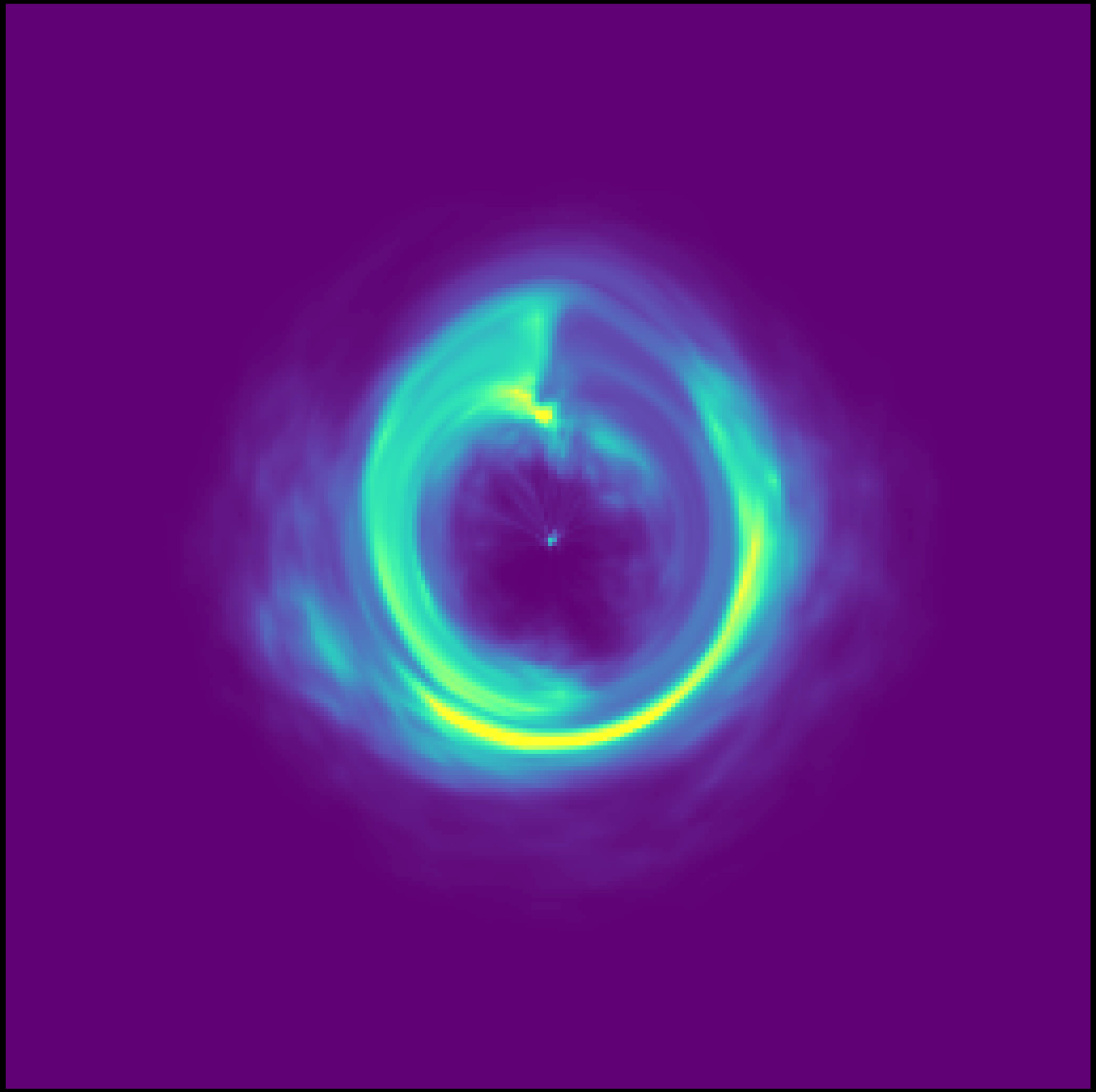
$$P(s \mid d)$$



Inference

Posterior (Ls)

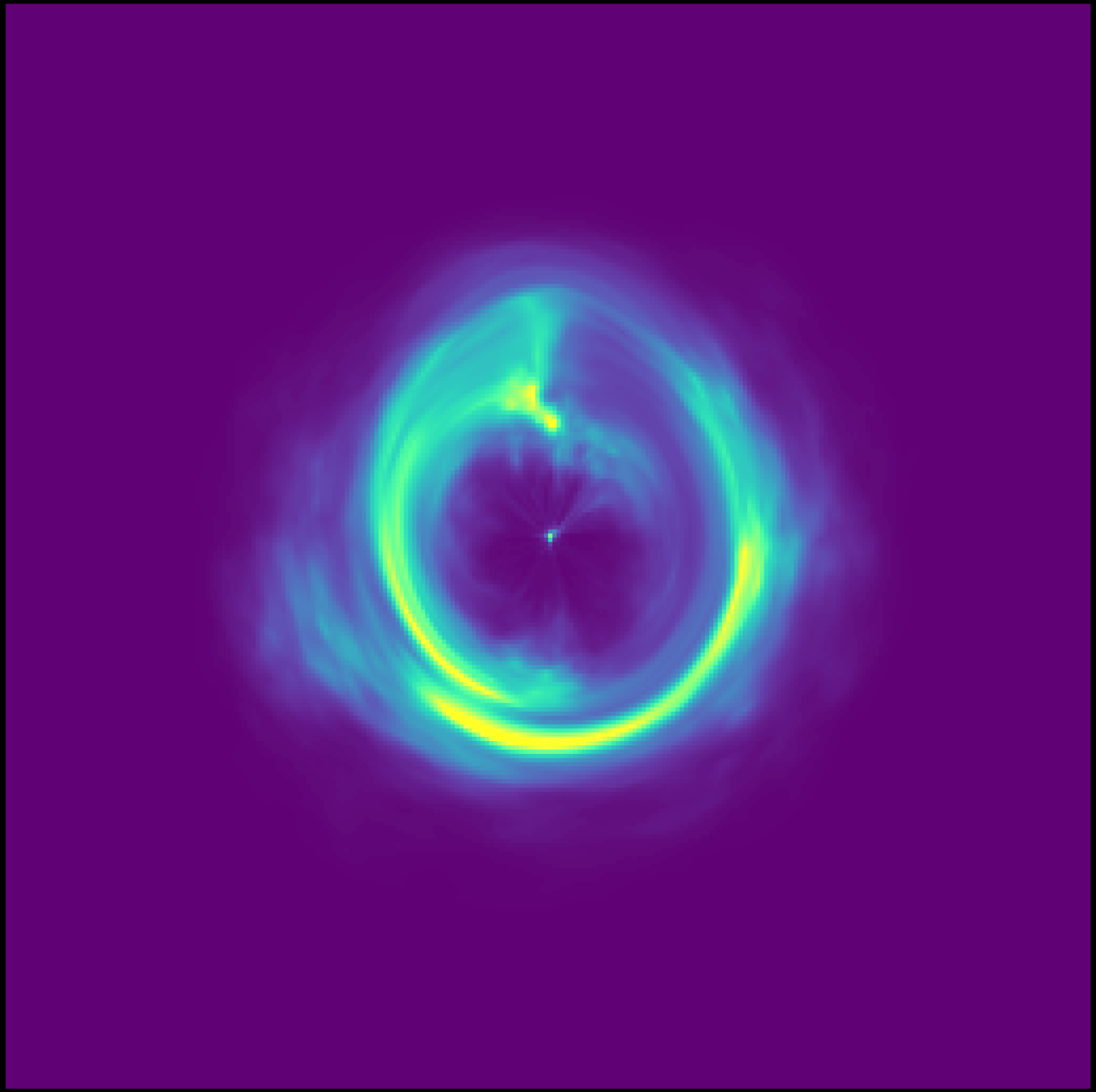
$$P(s \mid d)$$



Inference

Posterior (Ls)

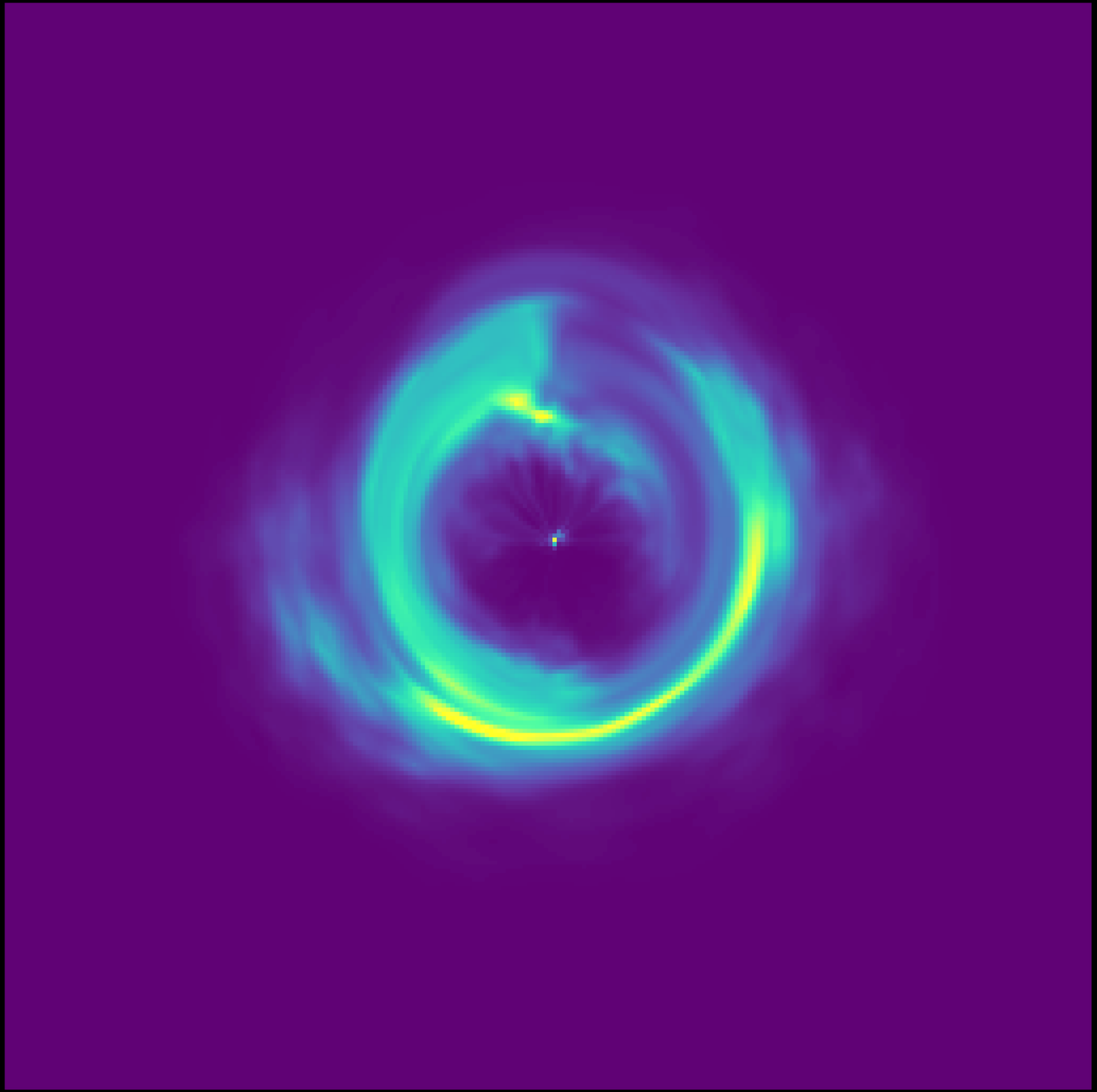
$$P(s \mid d)$$



Inference

Posterior (Ls)

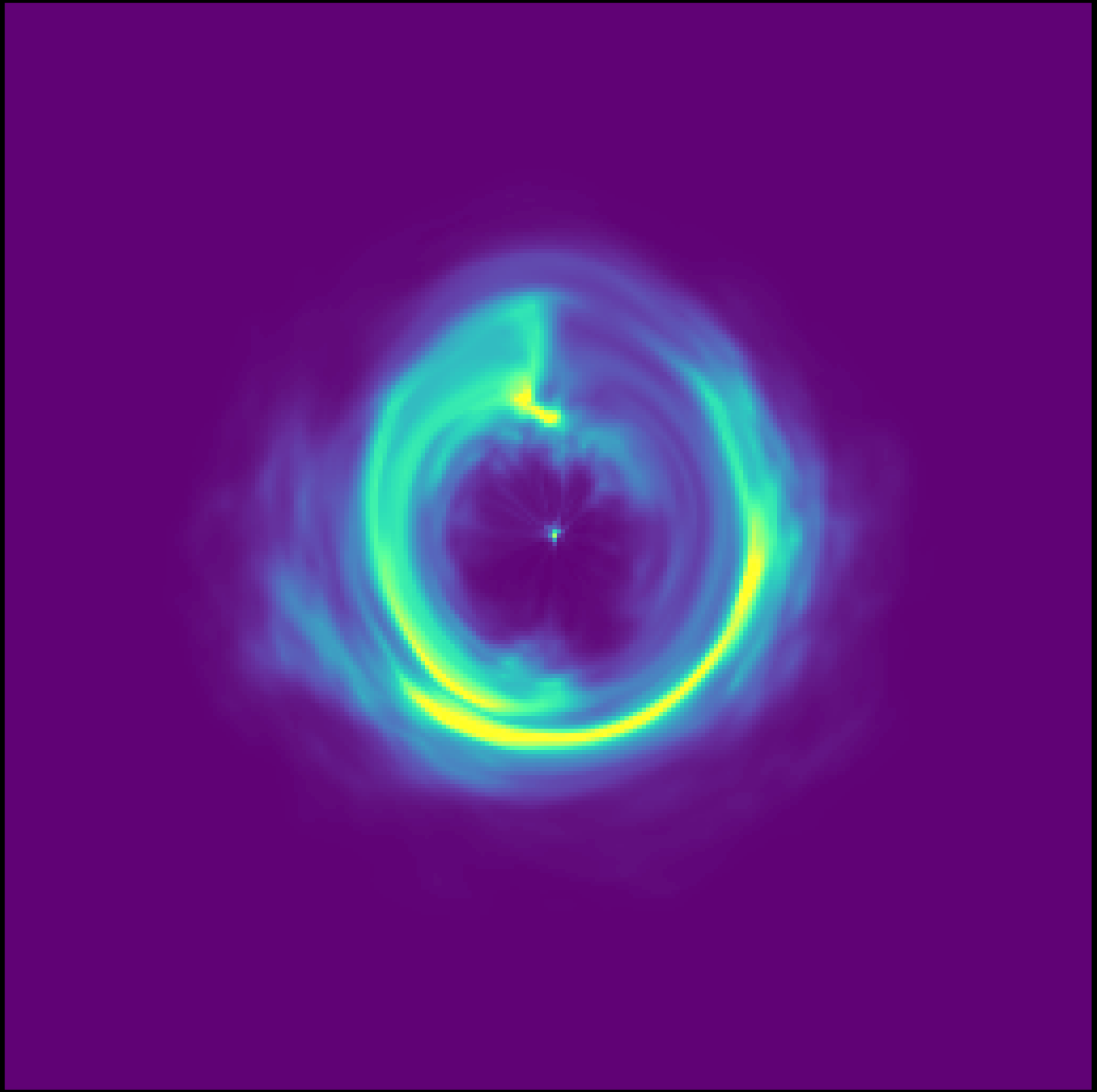
$$P(s \,|\, d)$$



Inference

Posterior (Ls)

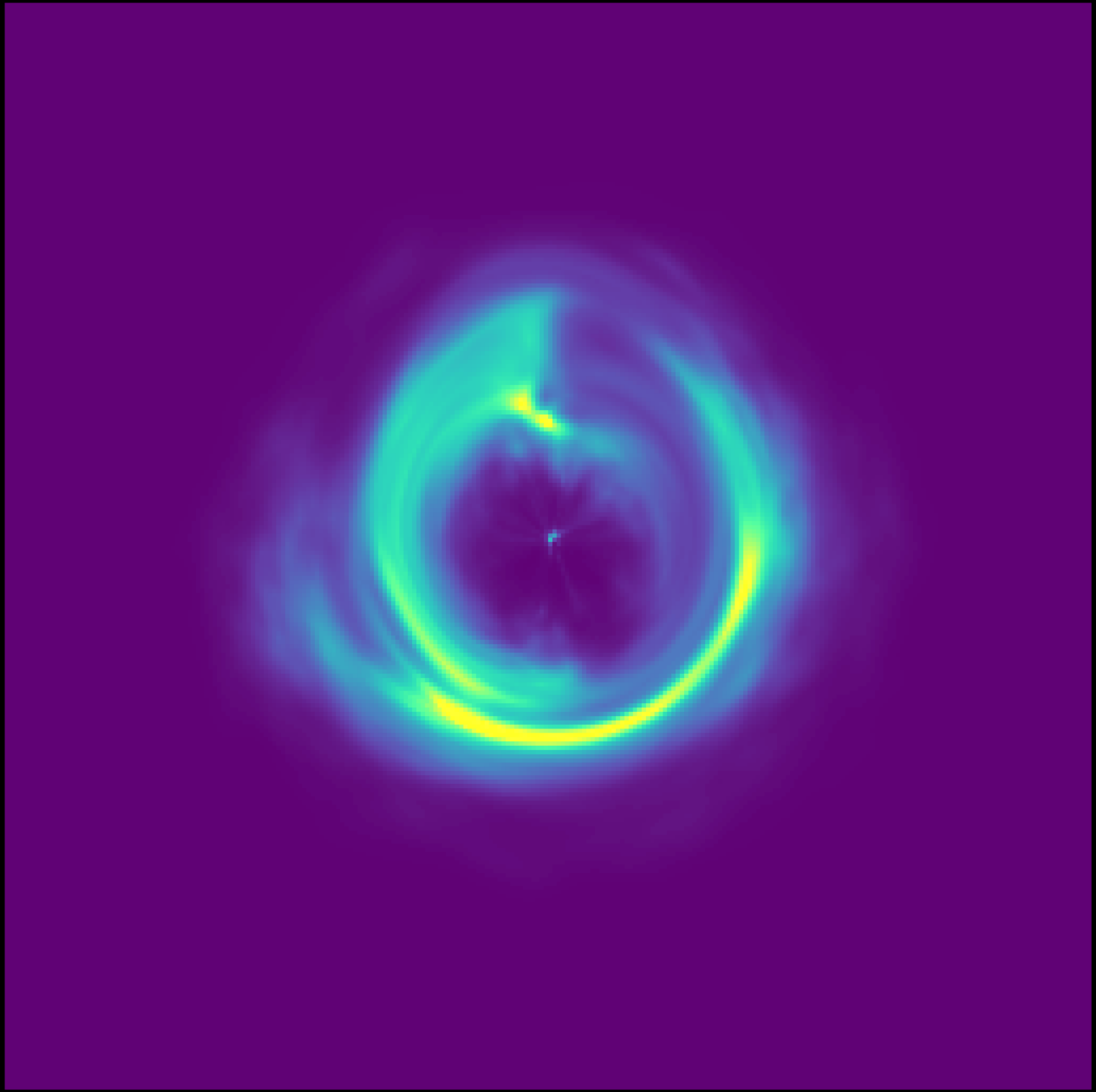
$$P(s \mid d)$$



Inference

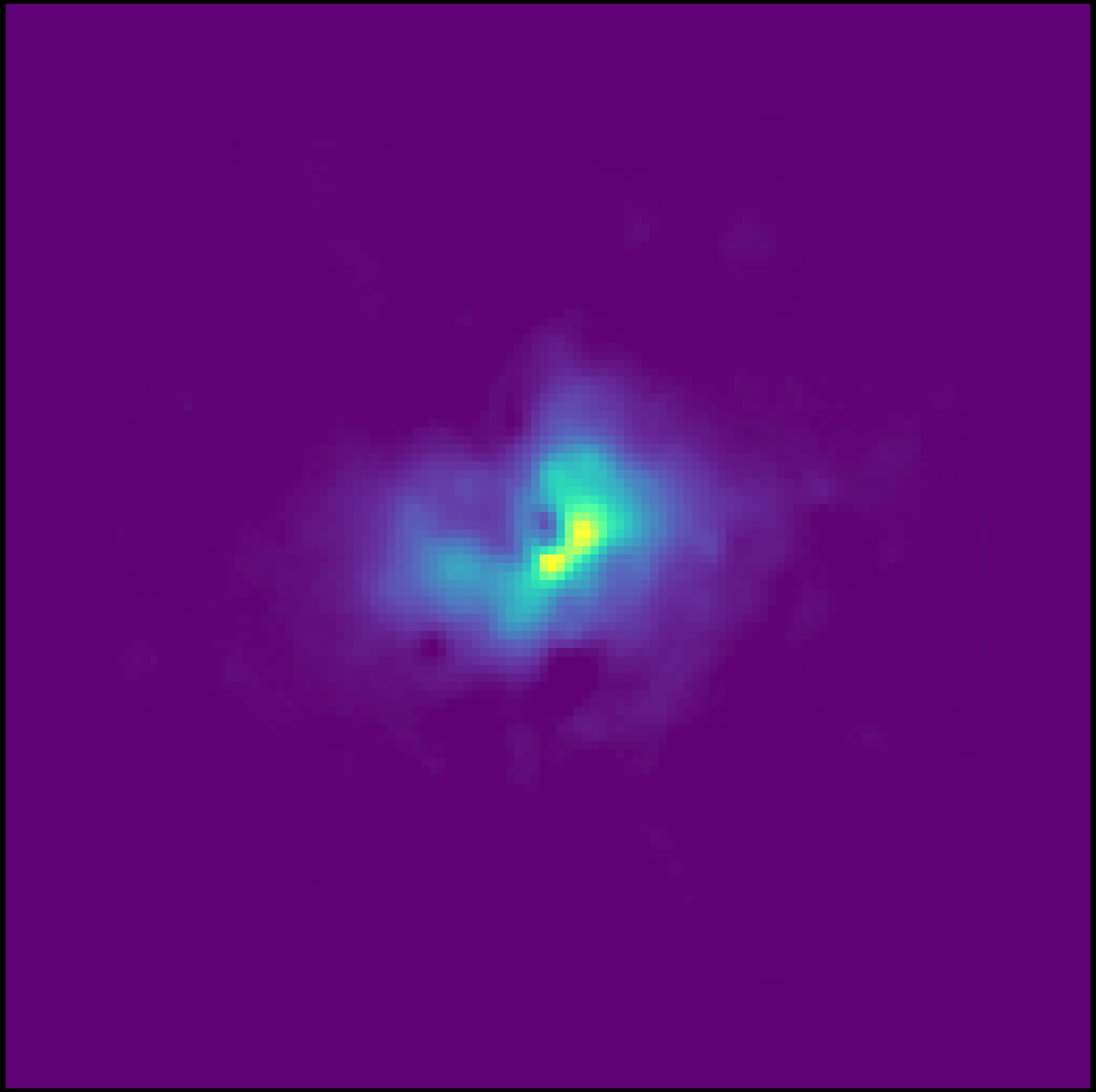
Posterior (Ls)

$$P(s\,|\,d)$$



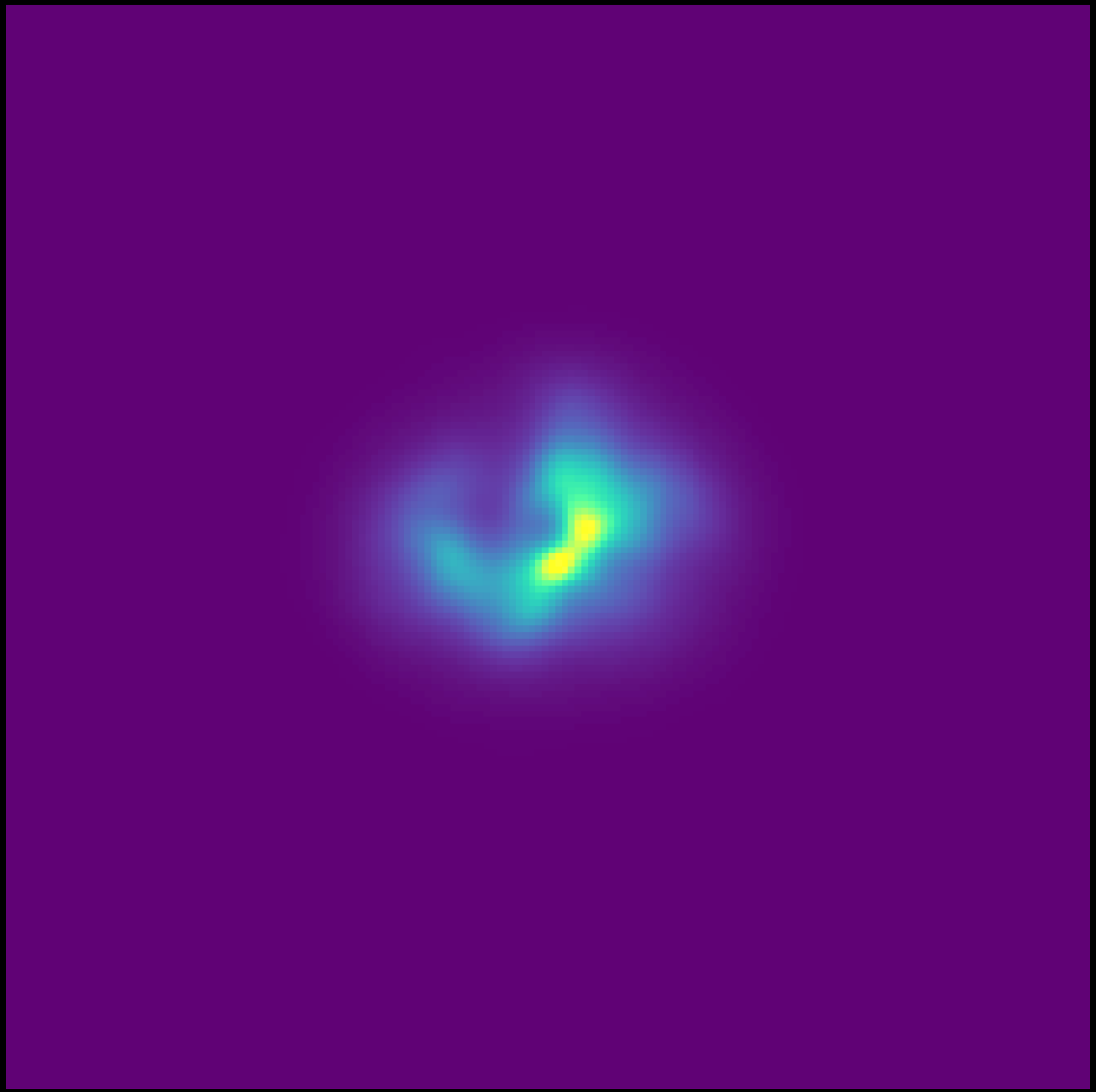
The source

Ground truth



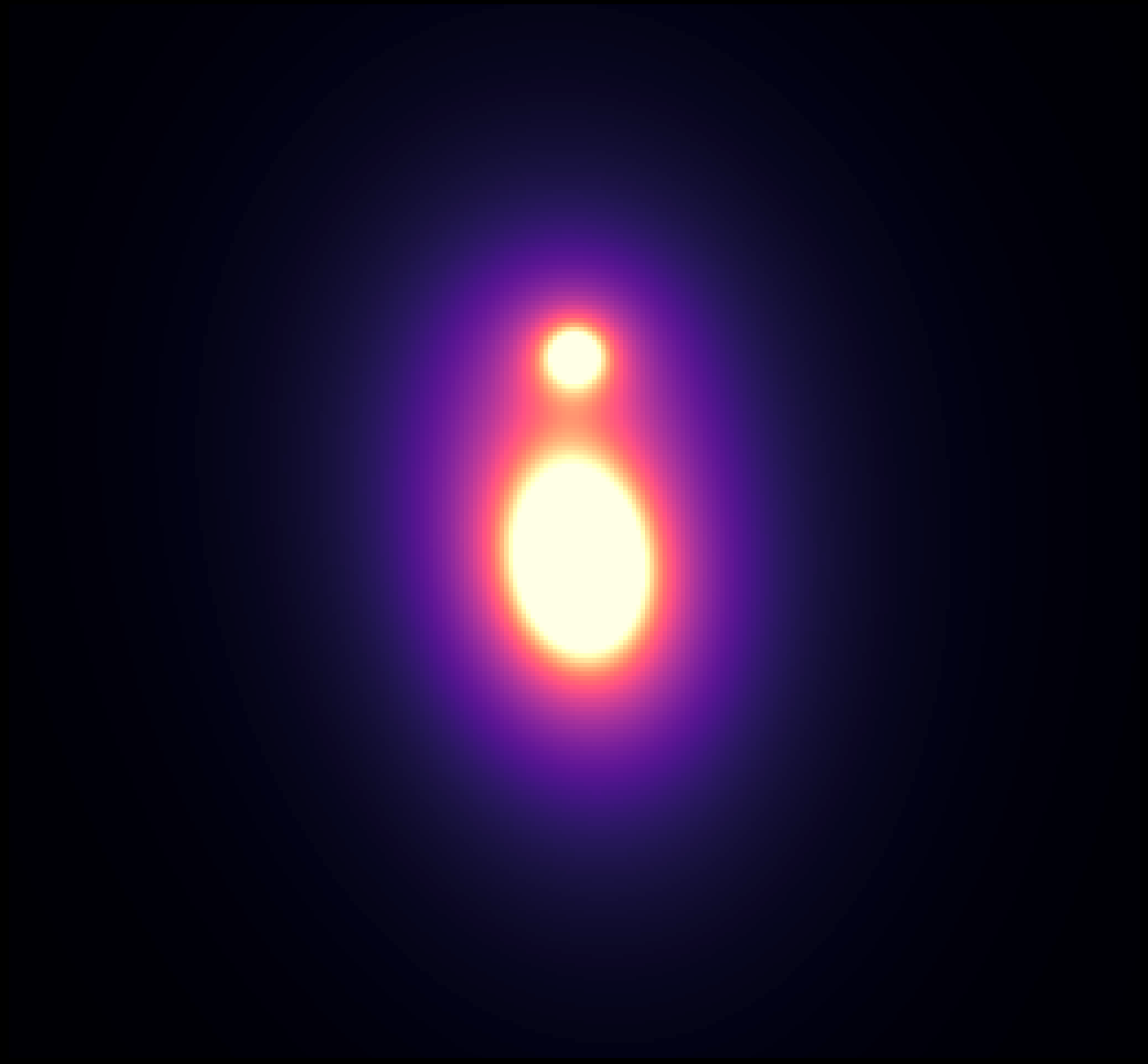
The source

Reconstruction



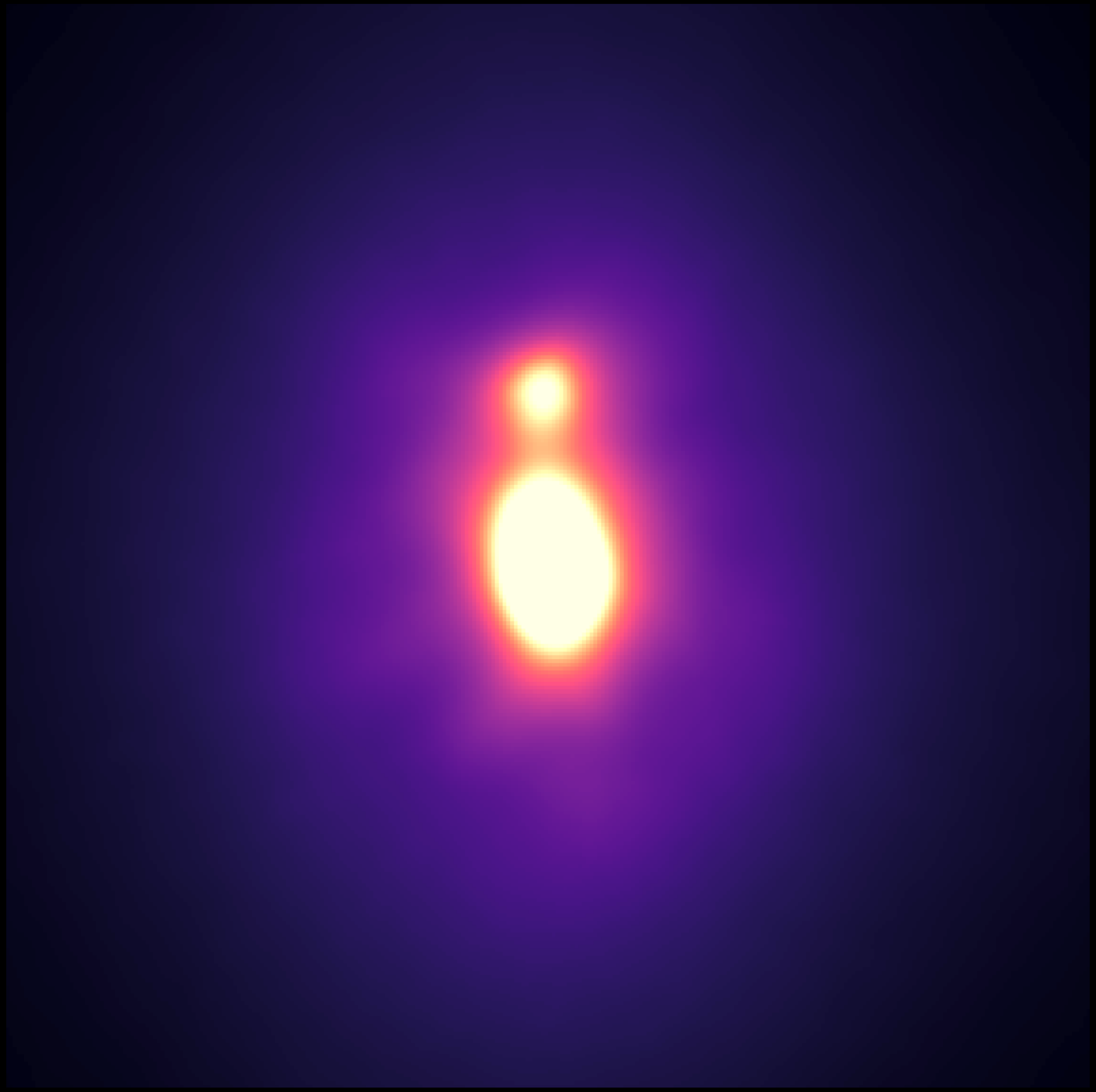
The convergence

Ground truth



The convergence

Reconstruction



So what’s the problem?

The problem

The problem

Fields!

The problem

Fields!

-> allow for **non-trivial degeneracies!**

The problem

Fields!

-> allow for **non-trivial degeneracies!**

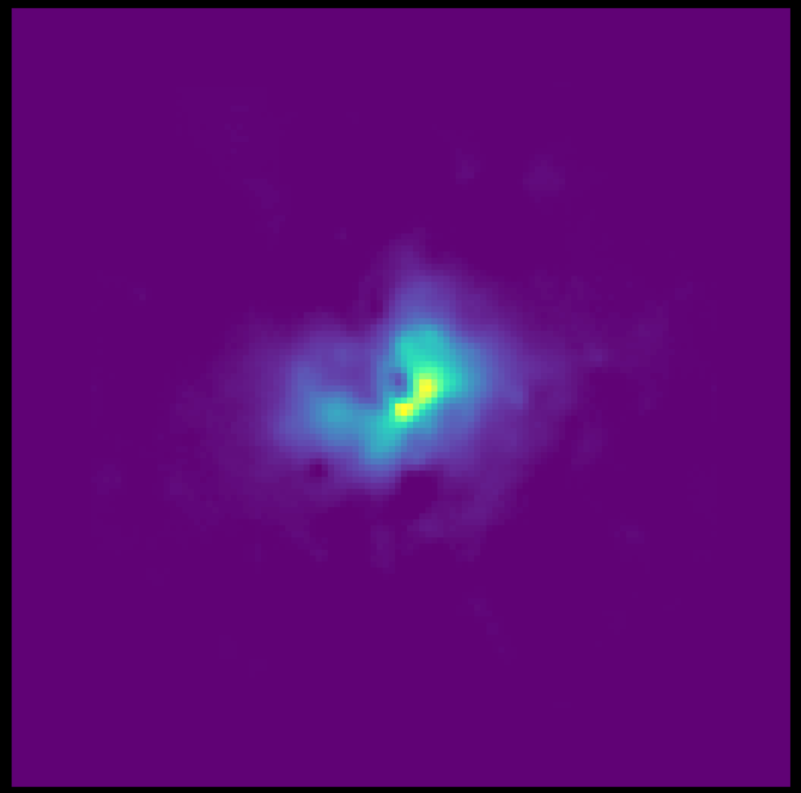
$$M_1(\theta_1)$$

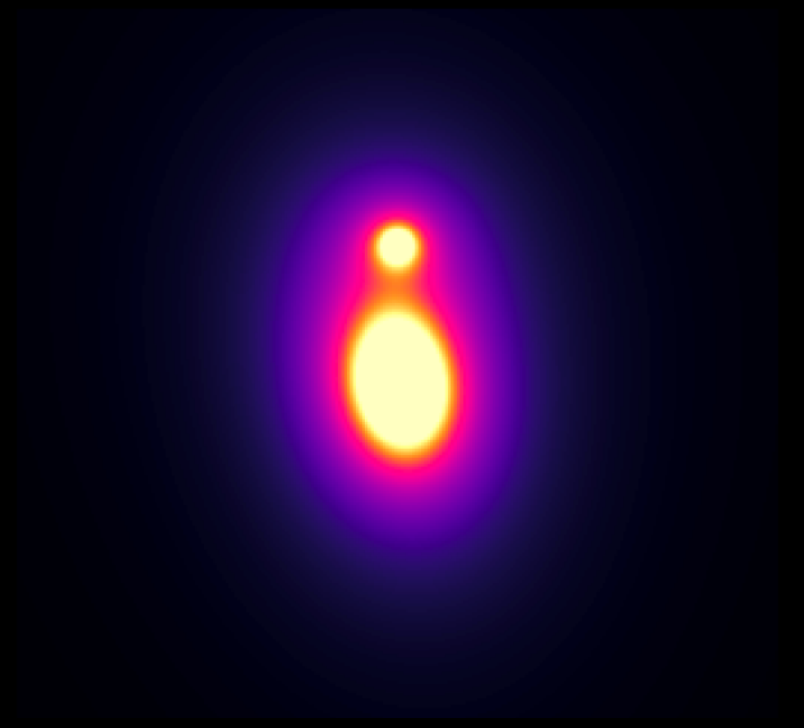
The problem

Fields!

-> allow for **non-trivial degeneracies!**

$M_1(\theta_1)$





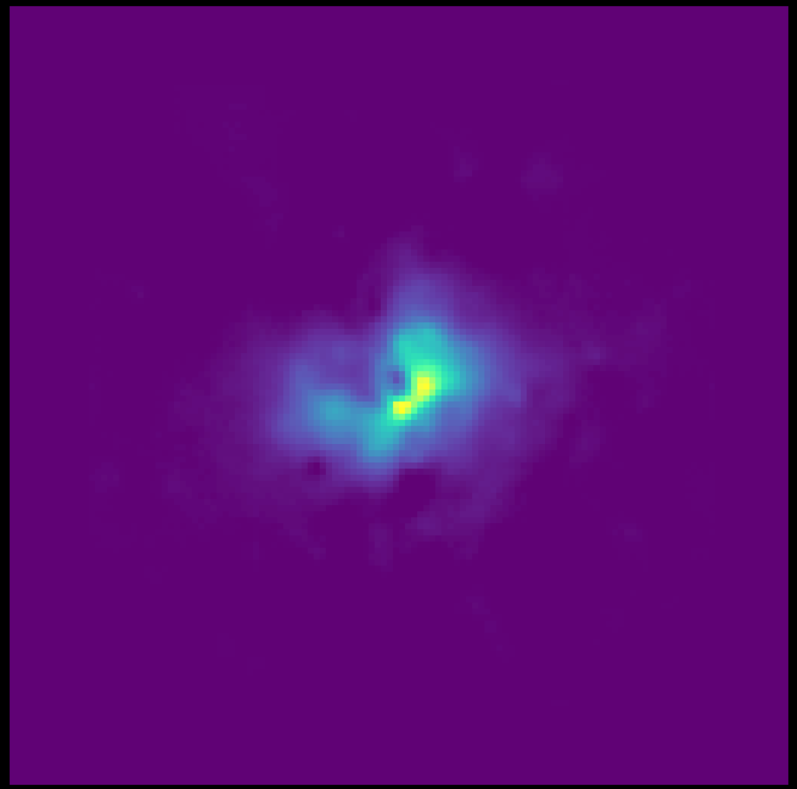
The problem

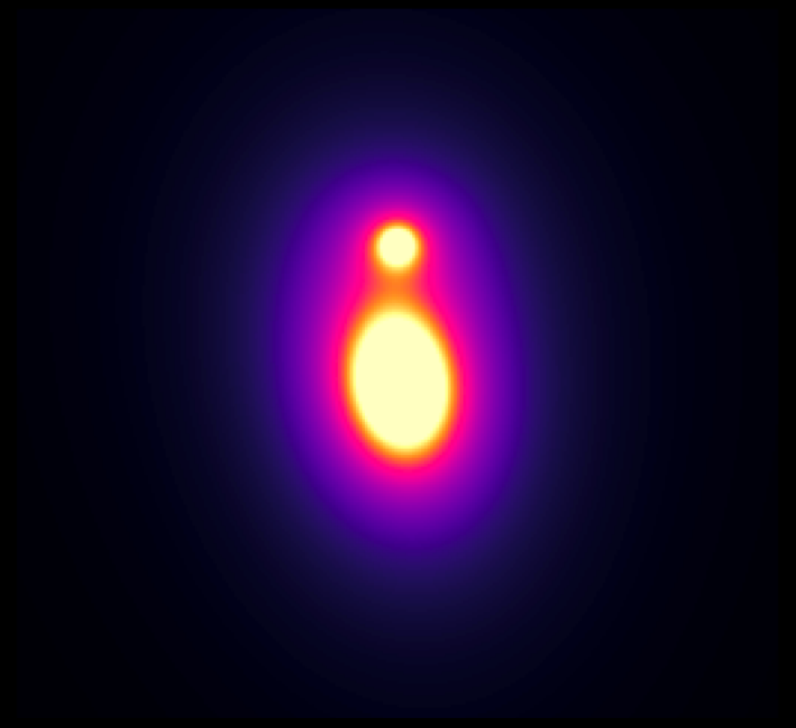
Fields!

-> allow for **non-trivial degeneracies!**

$M_1(\theta_1)$

$M_2(\theta_2)$



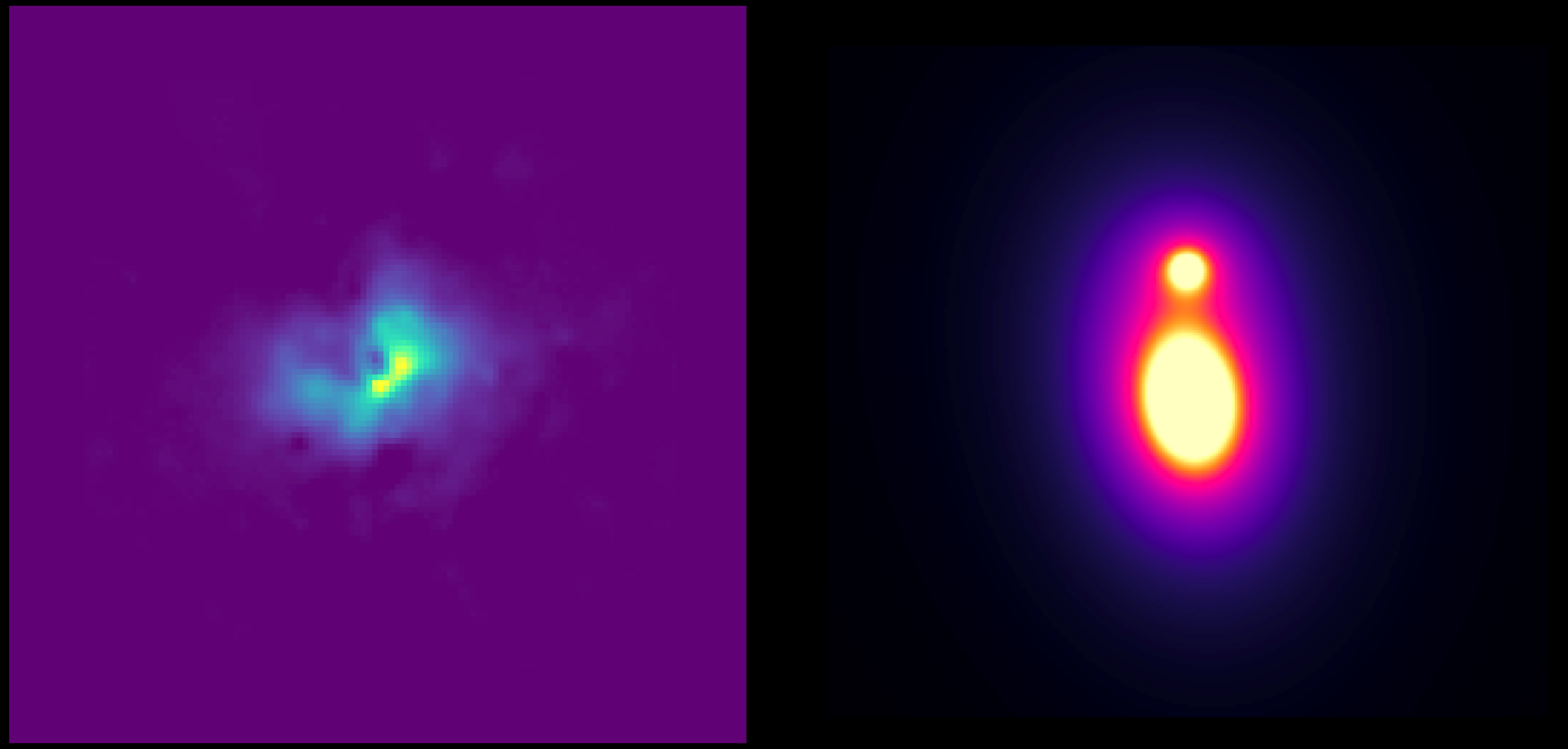


The problem

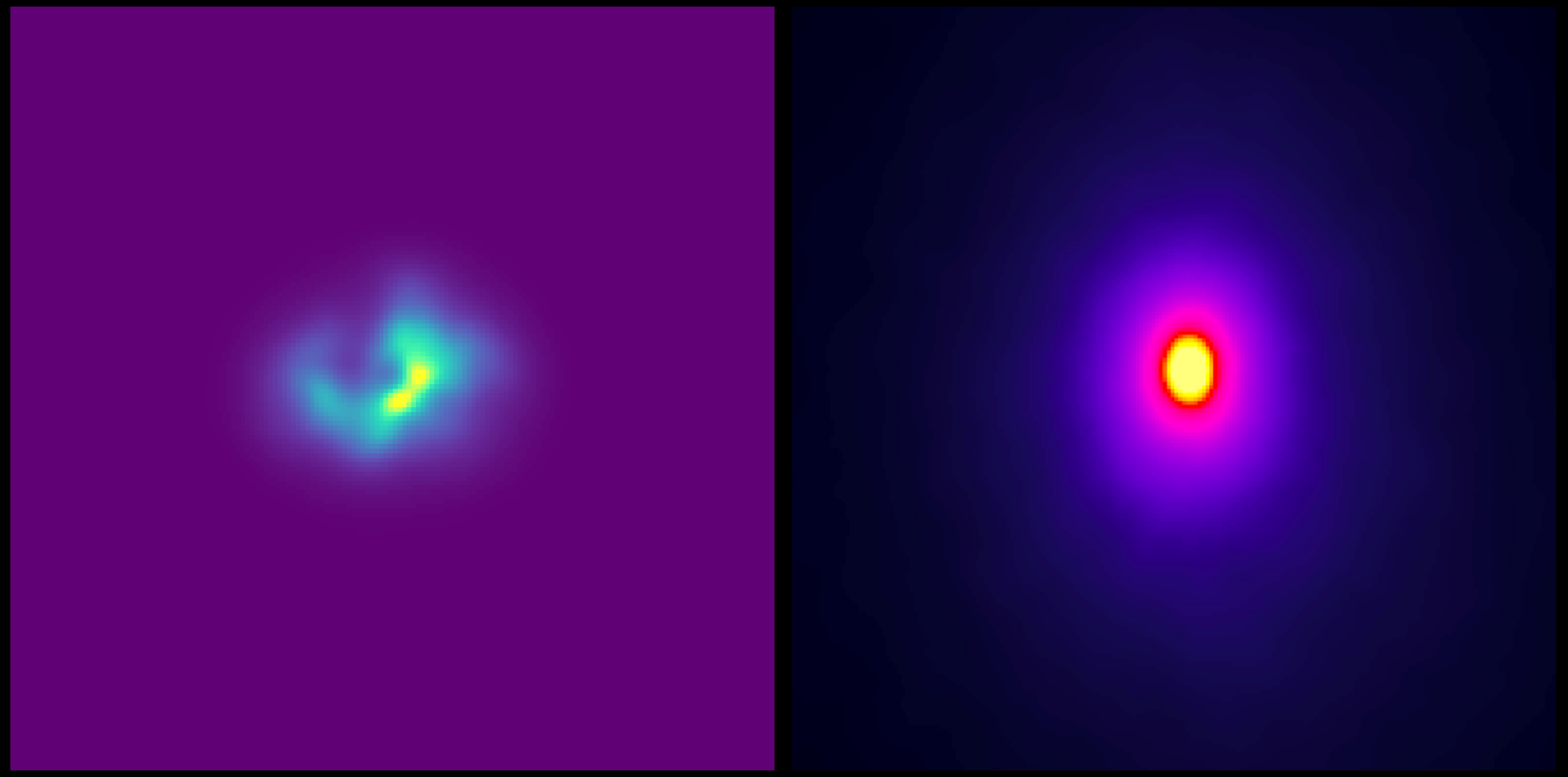
Fields!

-> allow for **non-trivial degeneracies!**

$M_1(\theta_1)$



$M_2(\theta_2)$

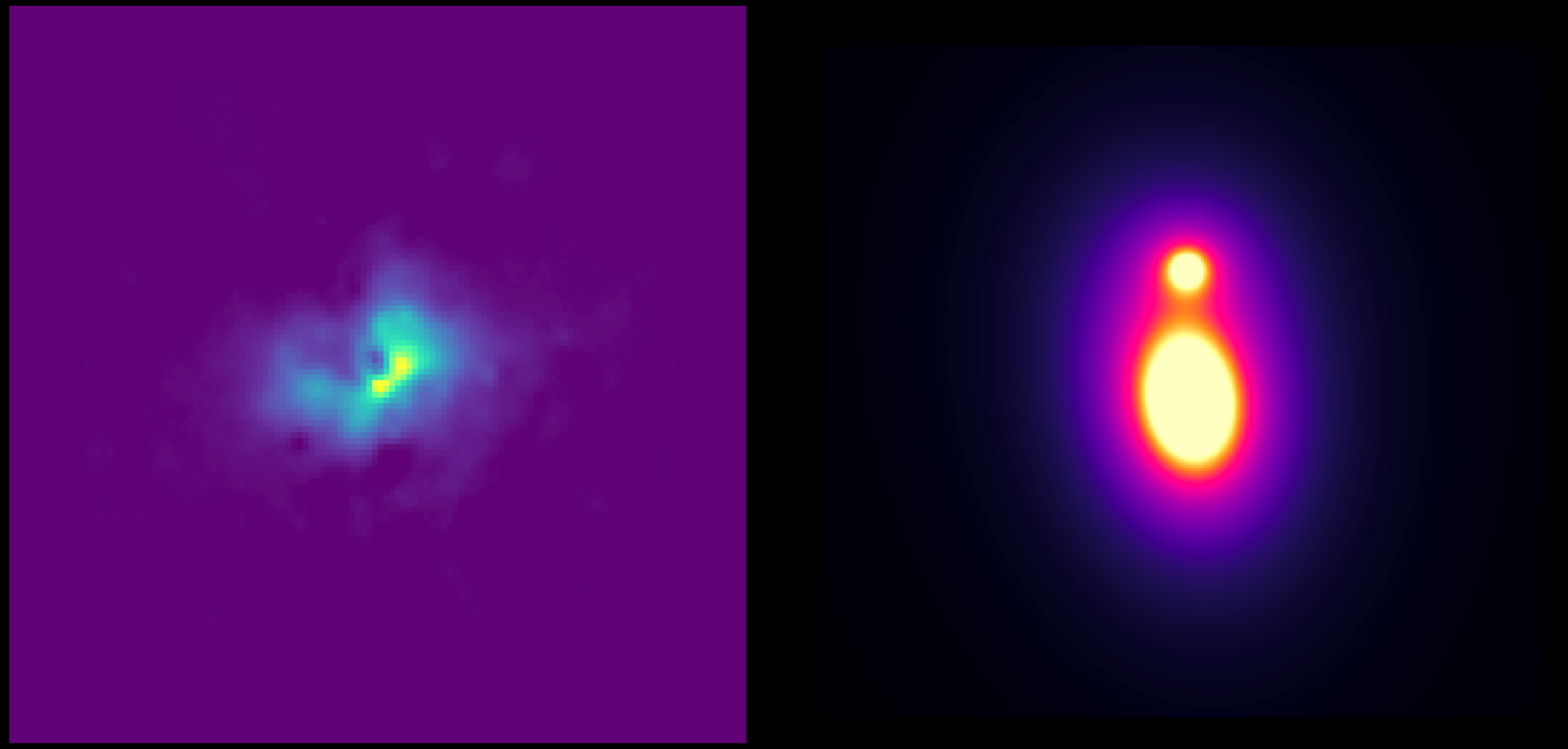


The problem

Fields!

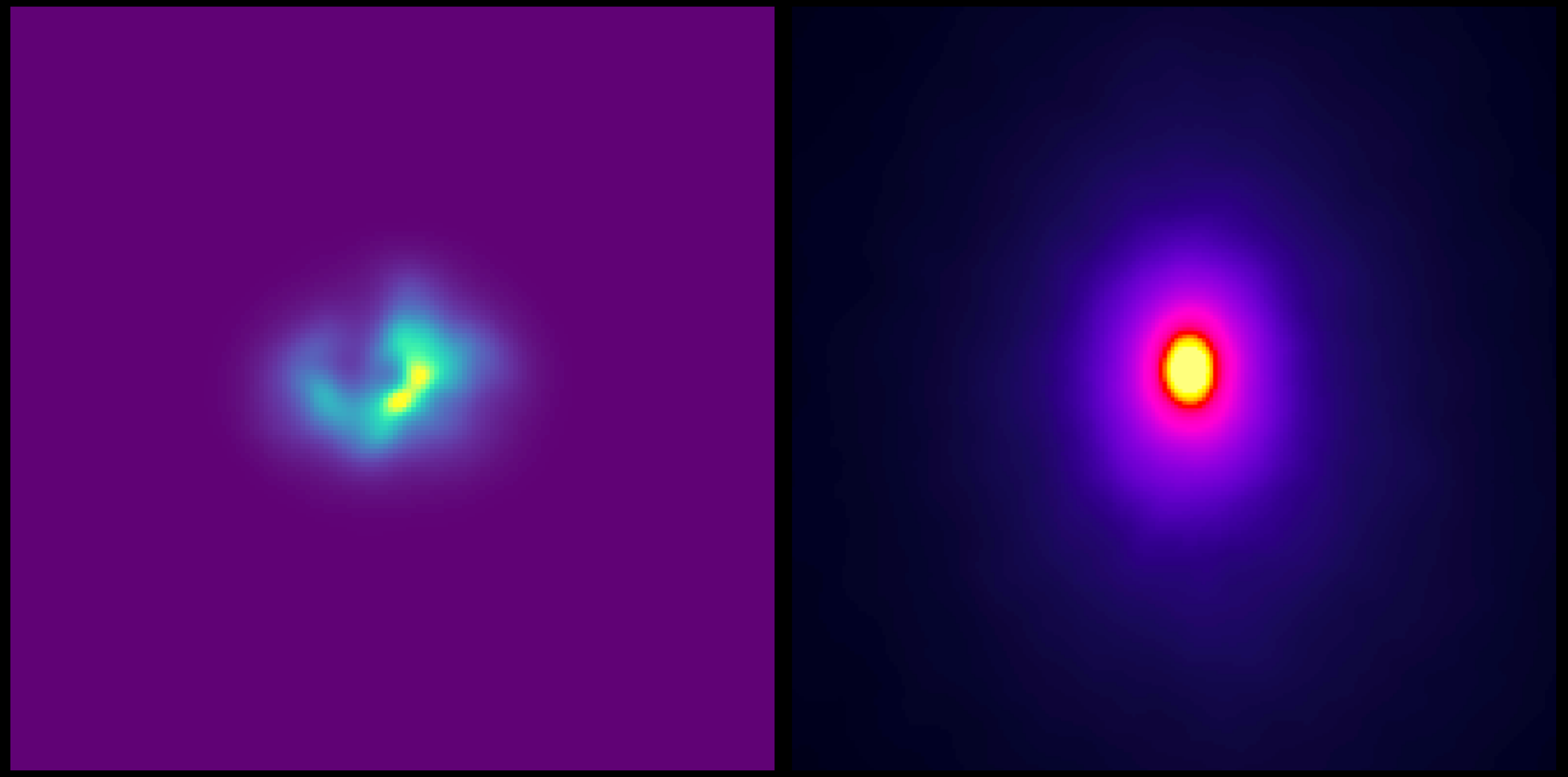
-> allow for **non-trivial degeneracies!**

$M_1(\theta_1)$



$\chi^2_{\mathrm{red}} \simeq 1.01$

$M_2(\theta_2)$

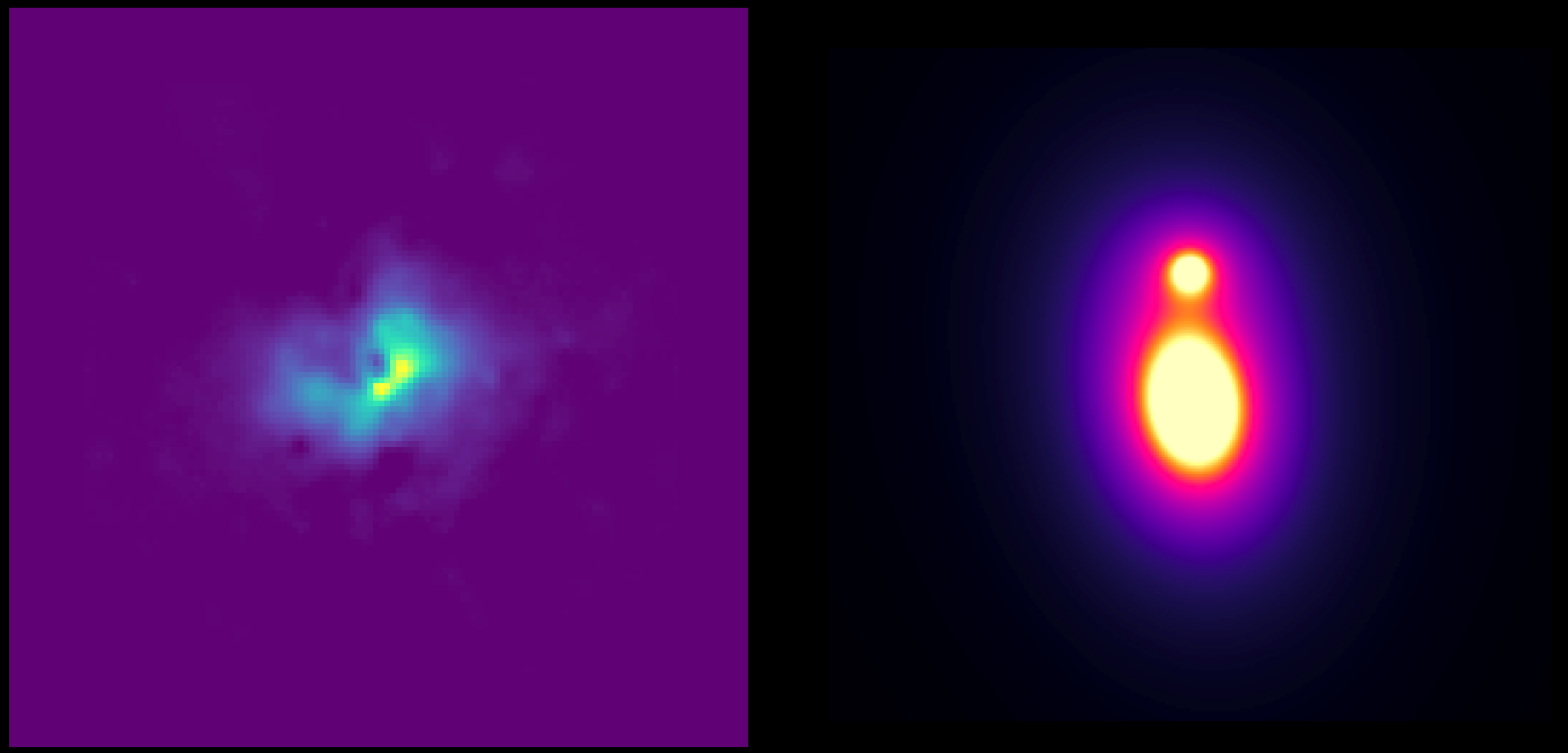


The problem

Fields!

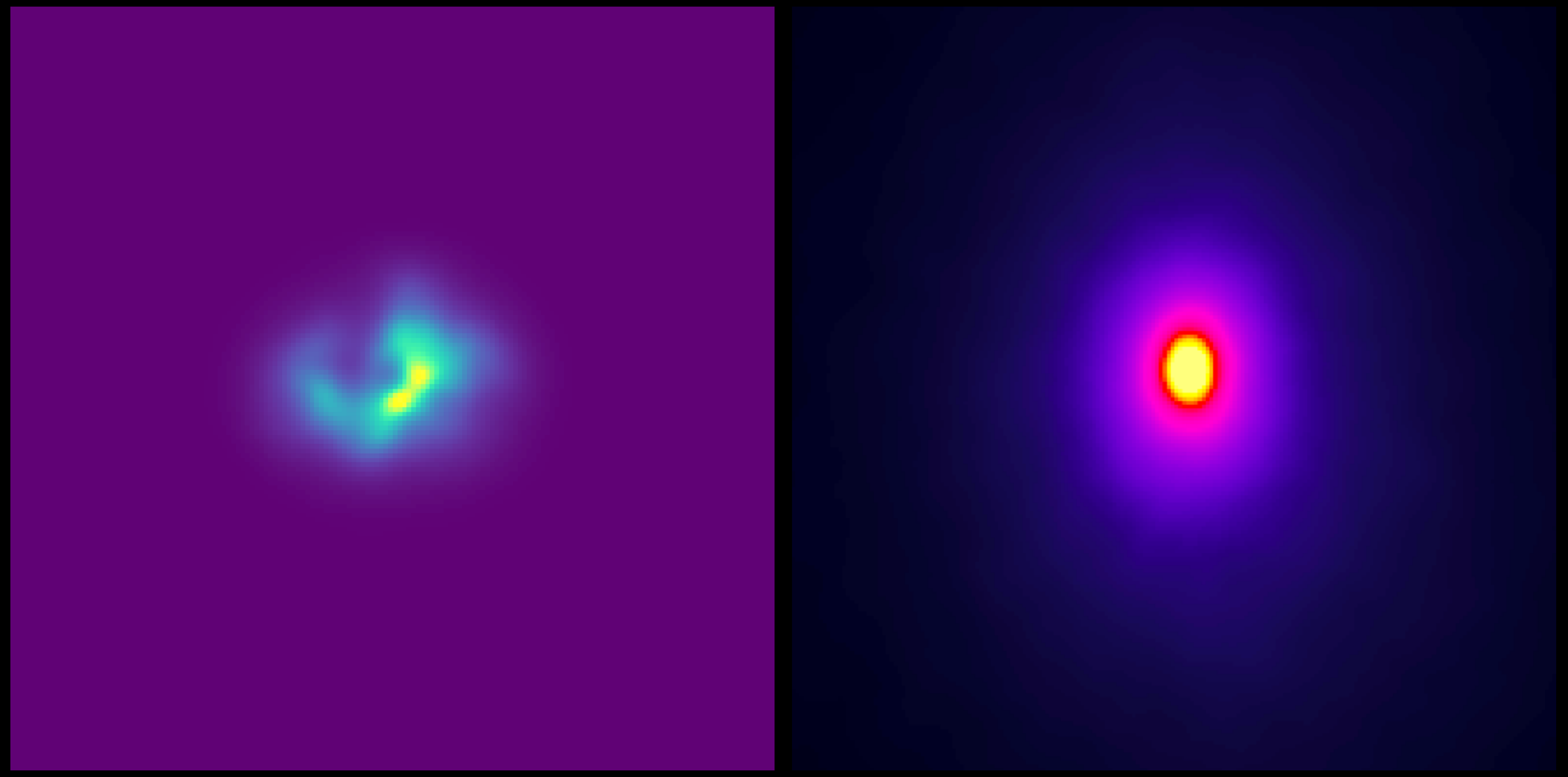
-> allow for **non-trivial degeneracies!**

$M_1(\theta_1)$



$\chi^2_{\text{red}} \simeq 1.01$

$M_2(\theta_2)$



$\chi^2_{\text{red}} \simeq 1.02$

Model choice

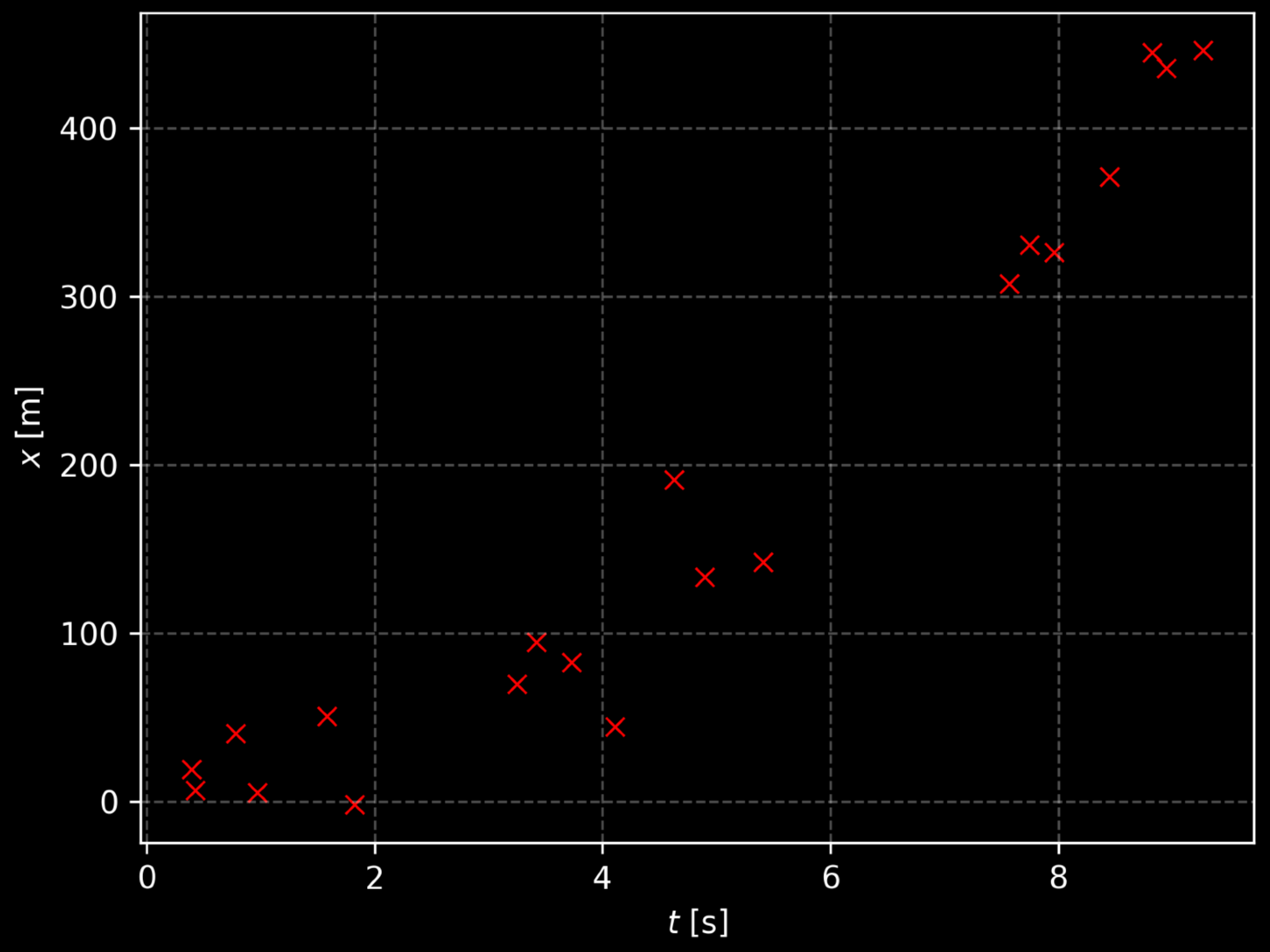
Model choice

The solution

Can we let the data decide?

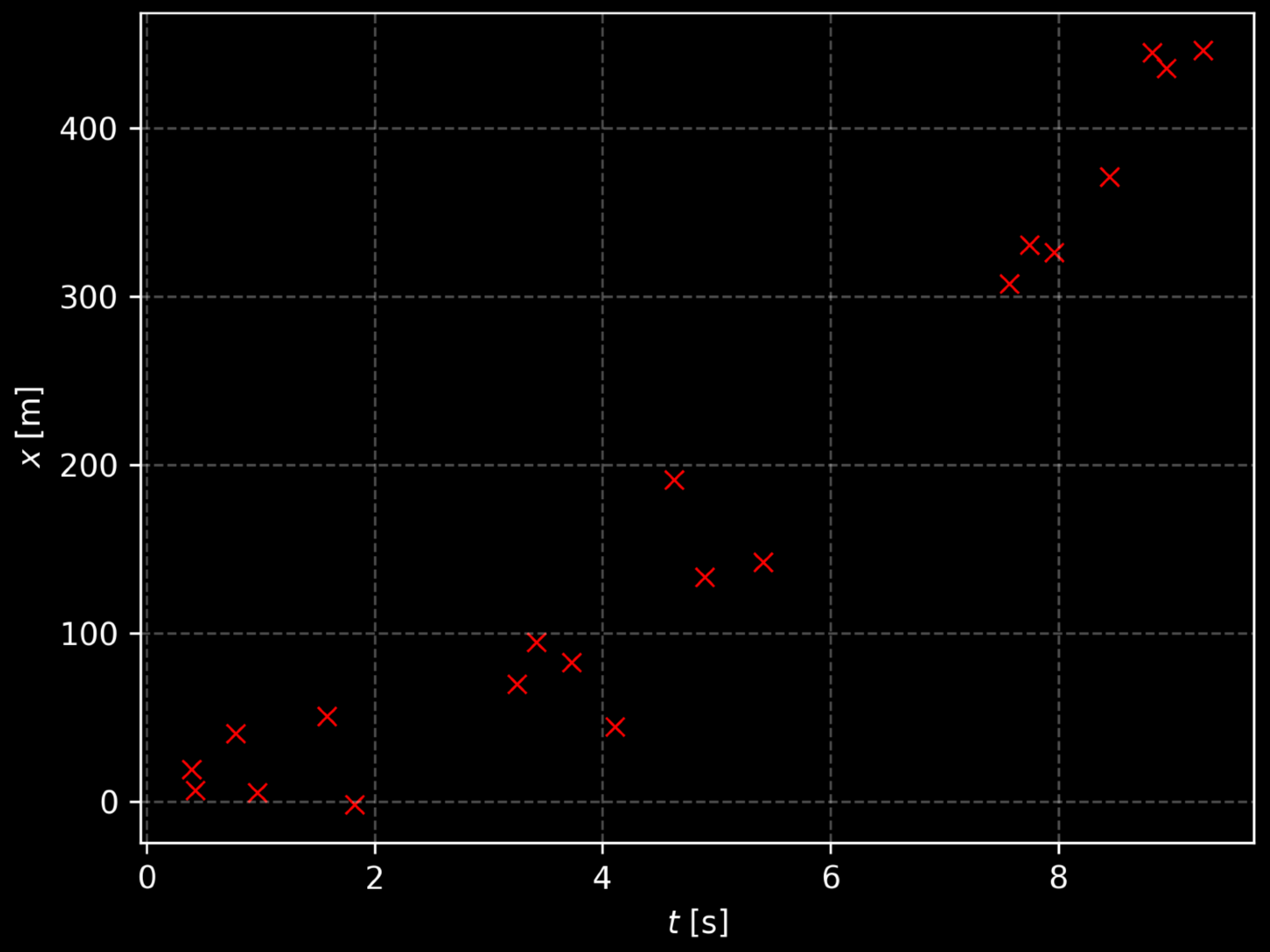
The solution

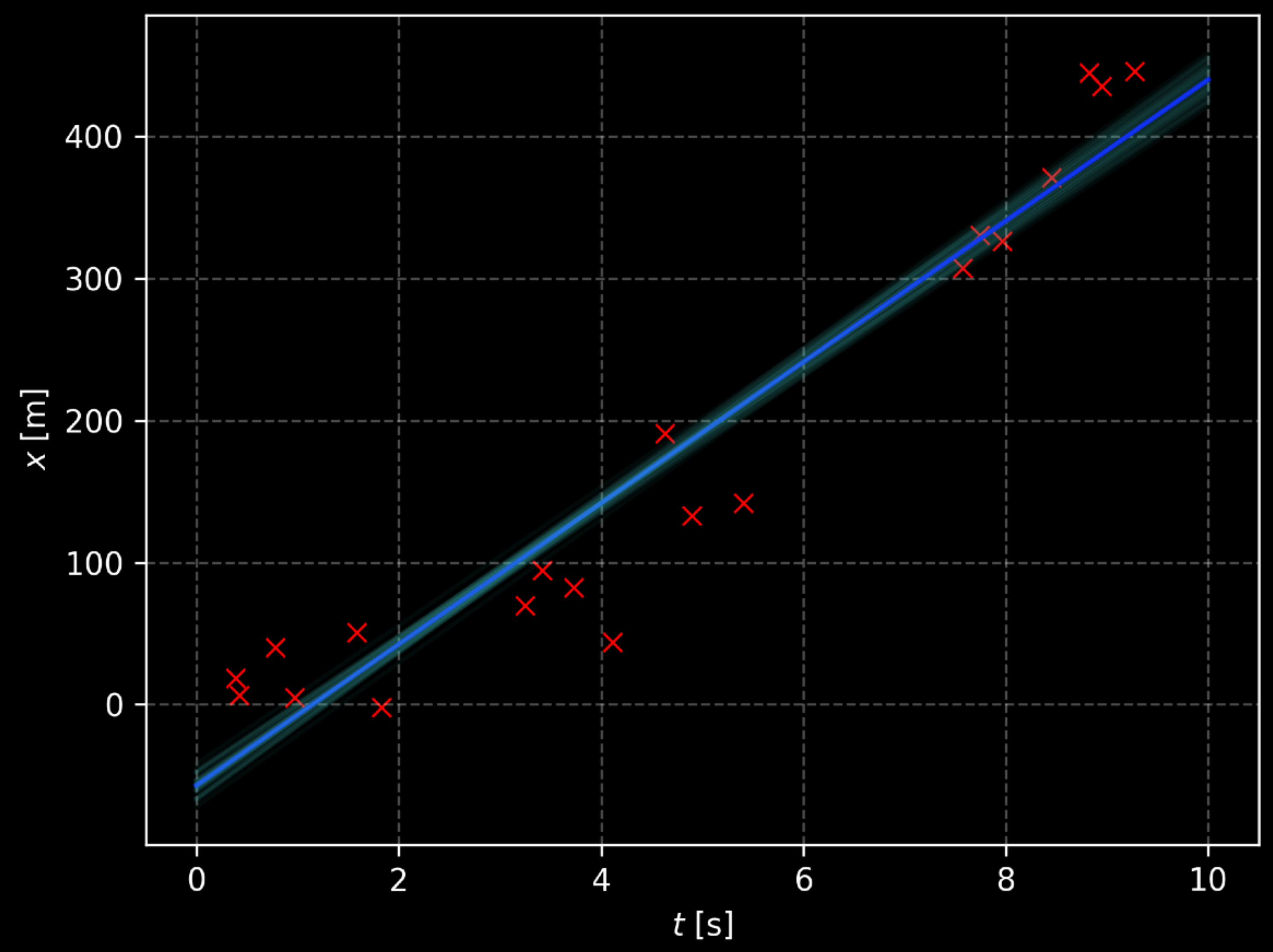
Can we let the data decide?



The solution

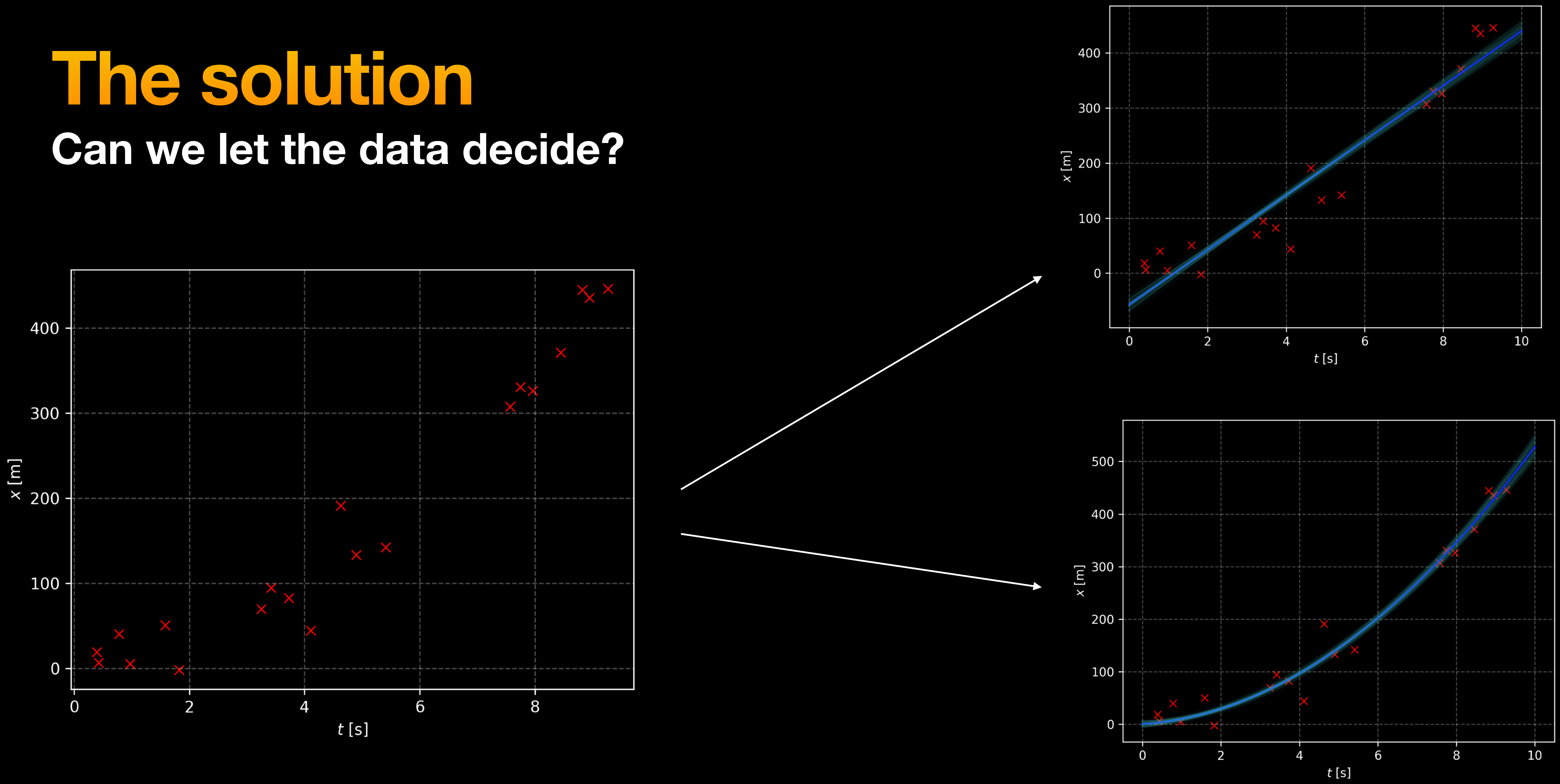
Can we let the data decide?





The solution

Can we let the data decide?

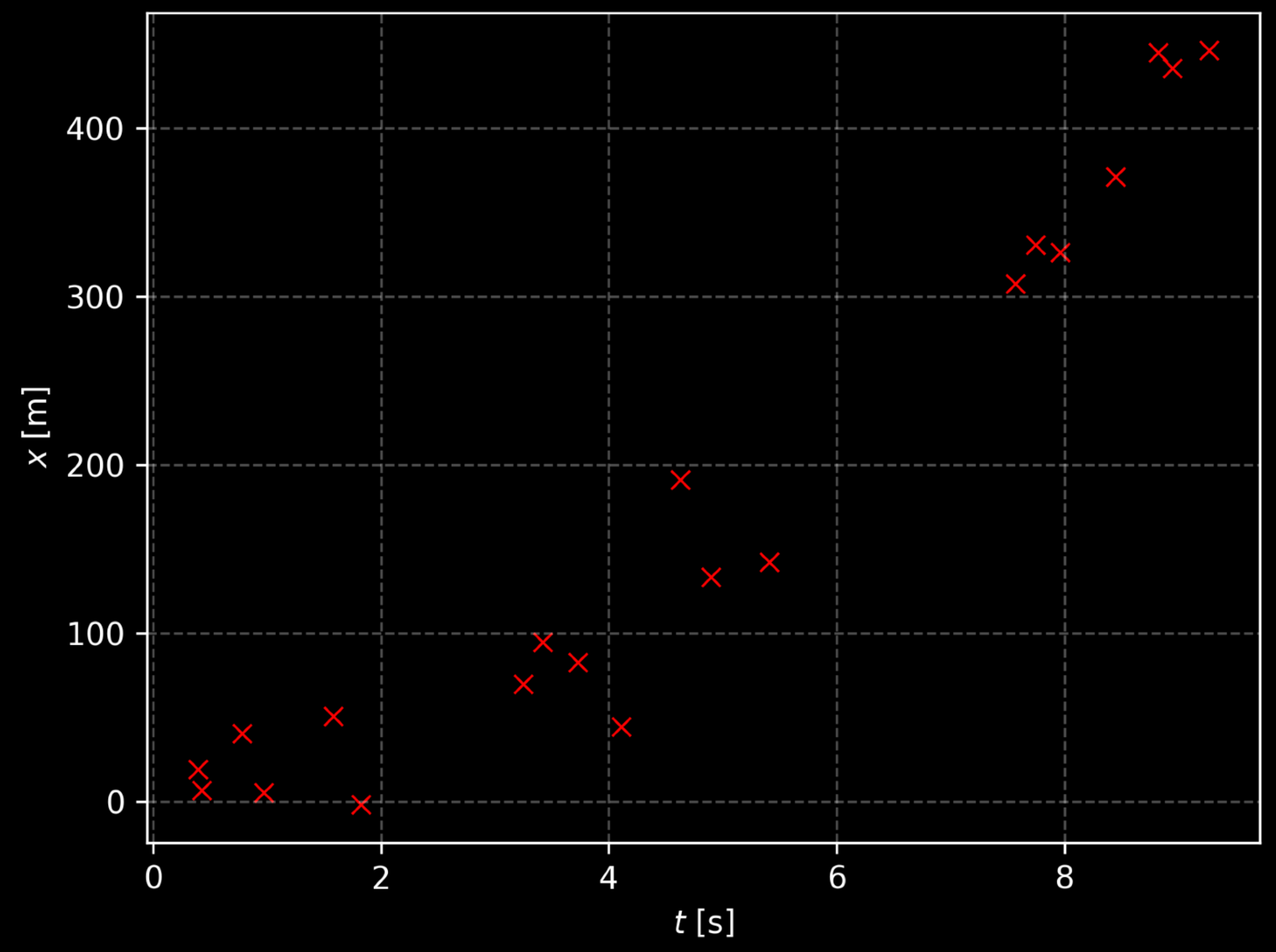


The solution

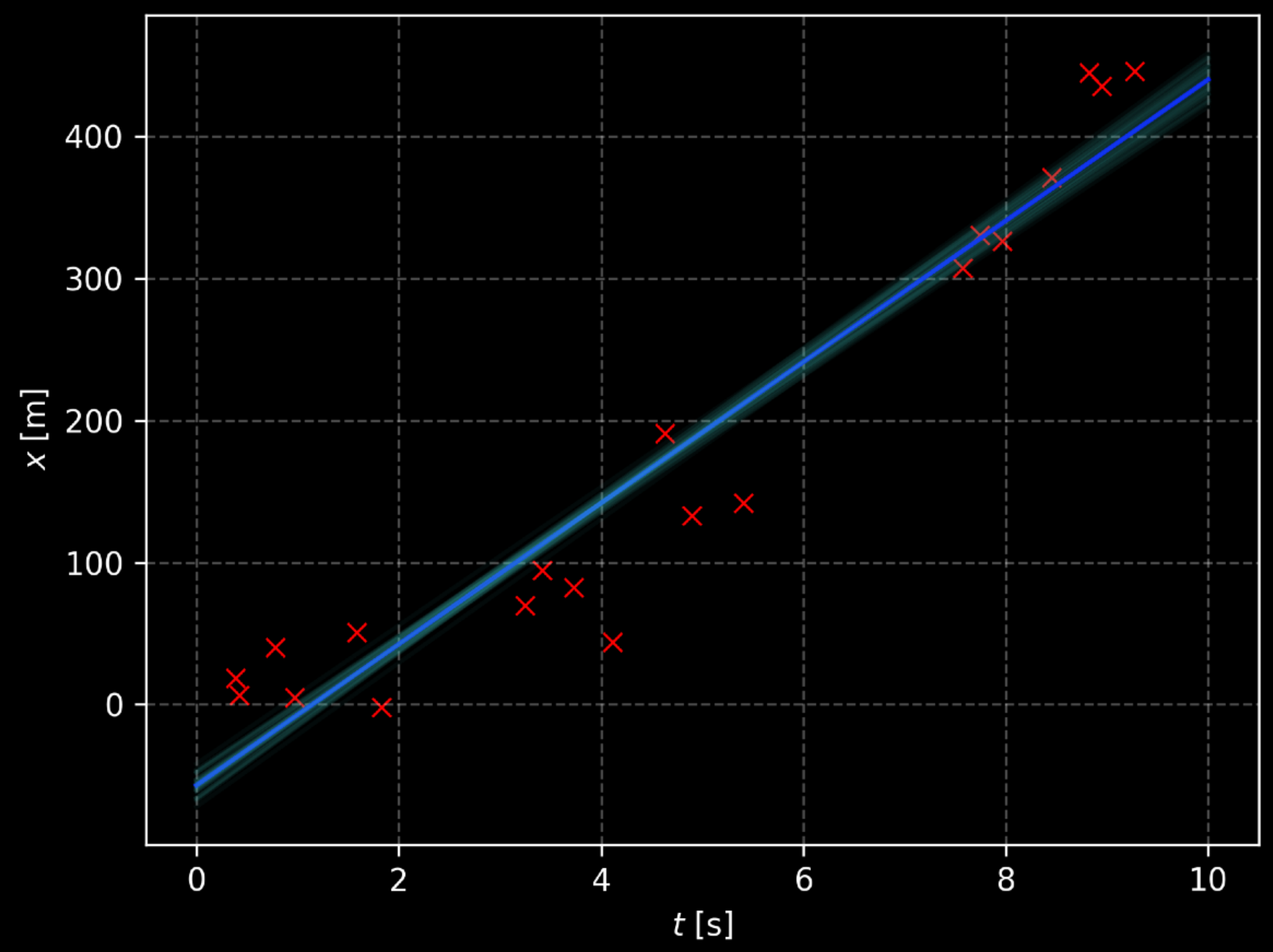
Can we let the data decide?

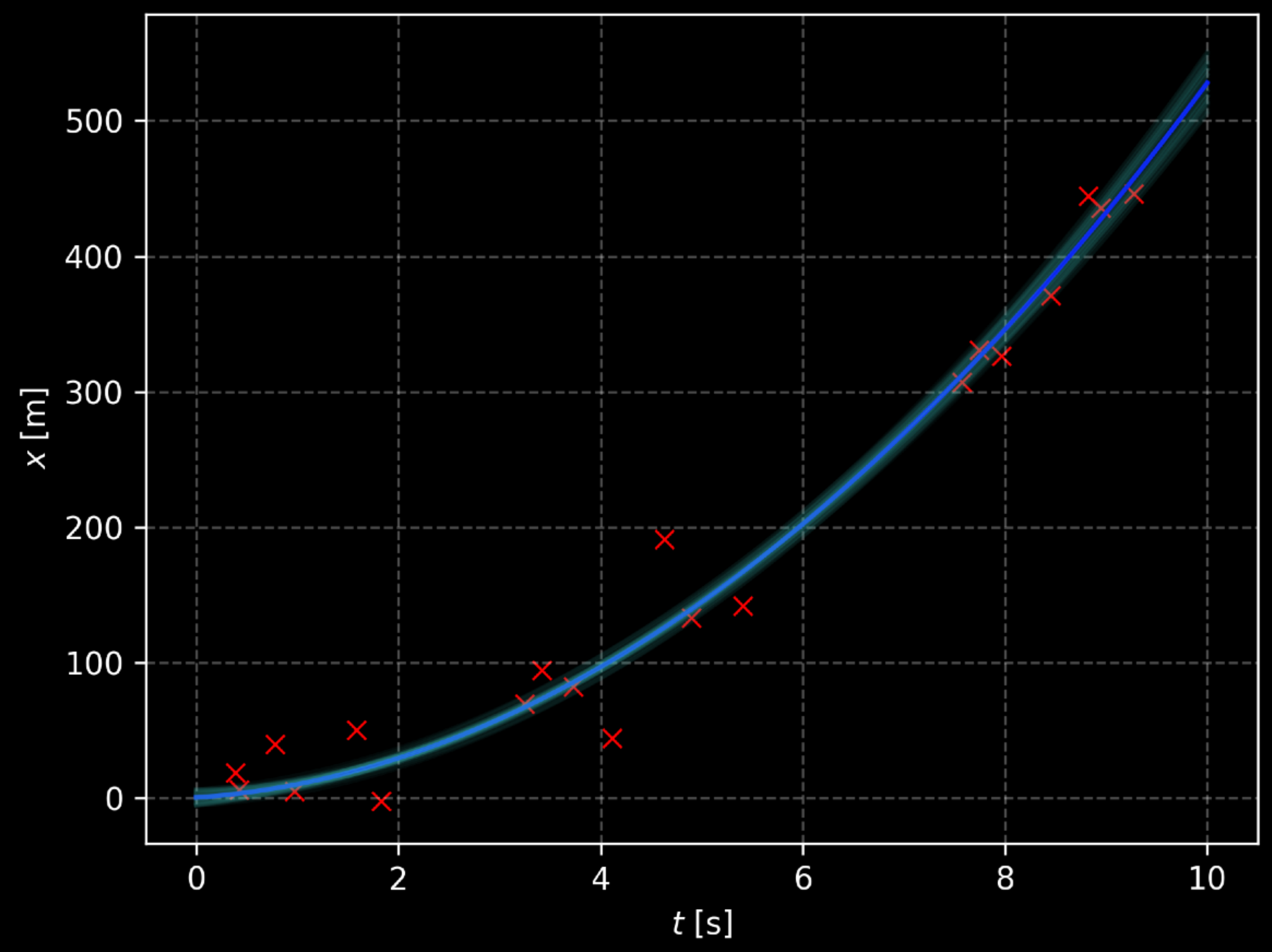
The solution

Can we let the data decide?



??



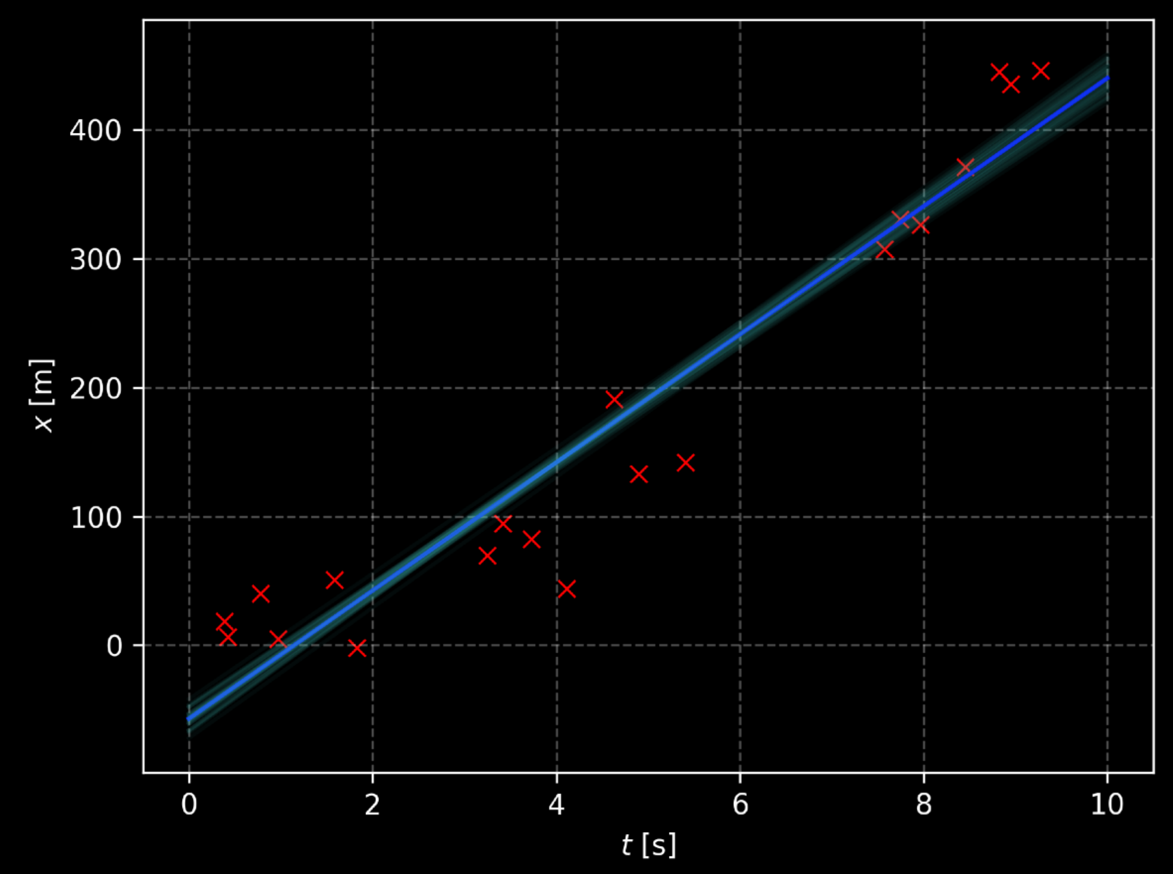


...

Comparing models

Comparing models

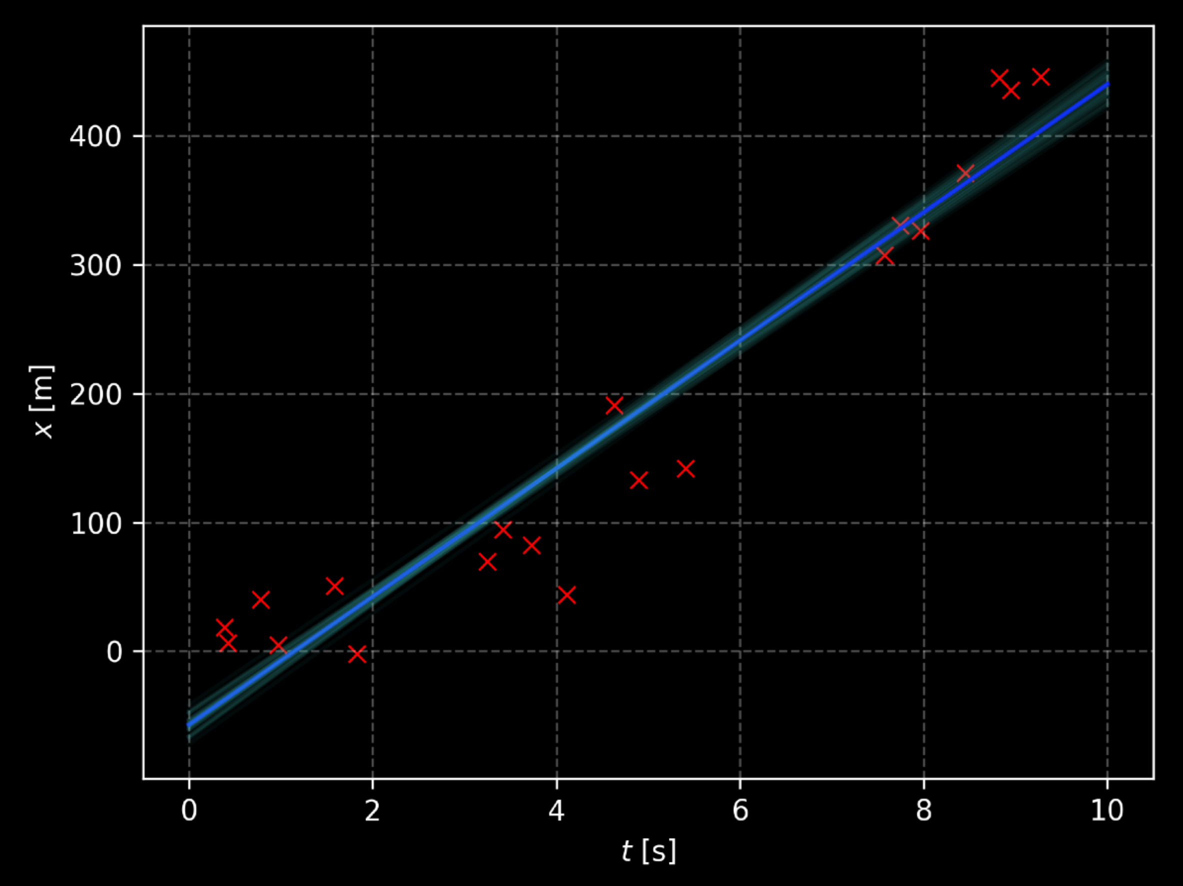
$M_1(\theta_1)$



Comparing models

$M_1(\theta_1)$

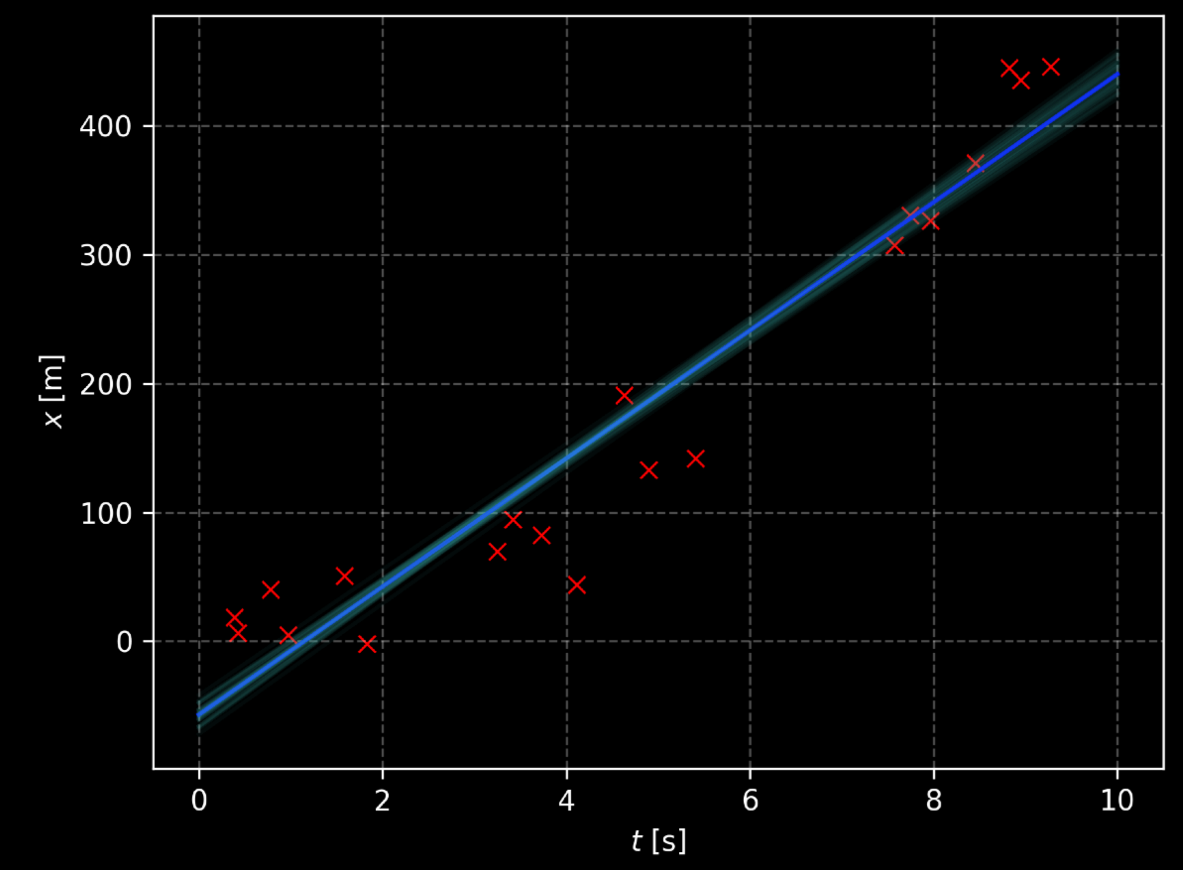
$$p(M_1 \,|\, d) = \frac{p(d \,|\, M_1)\, p(M_1)}{p(d)}$$



Comparing models

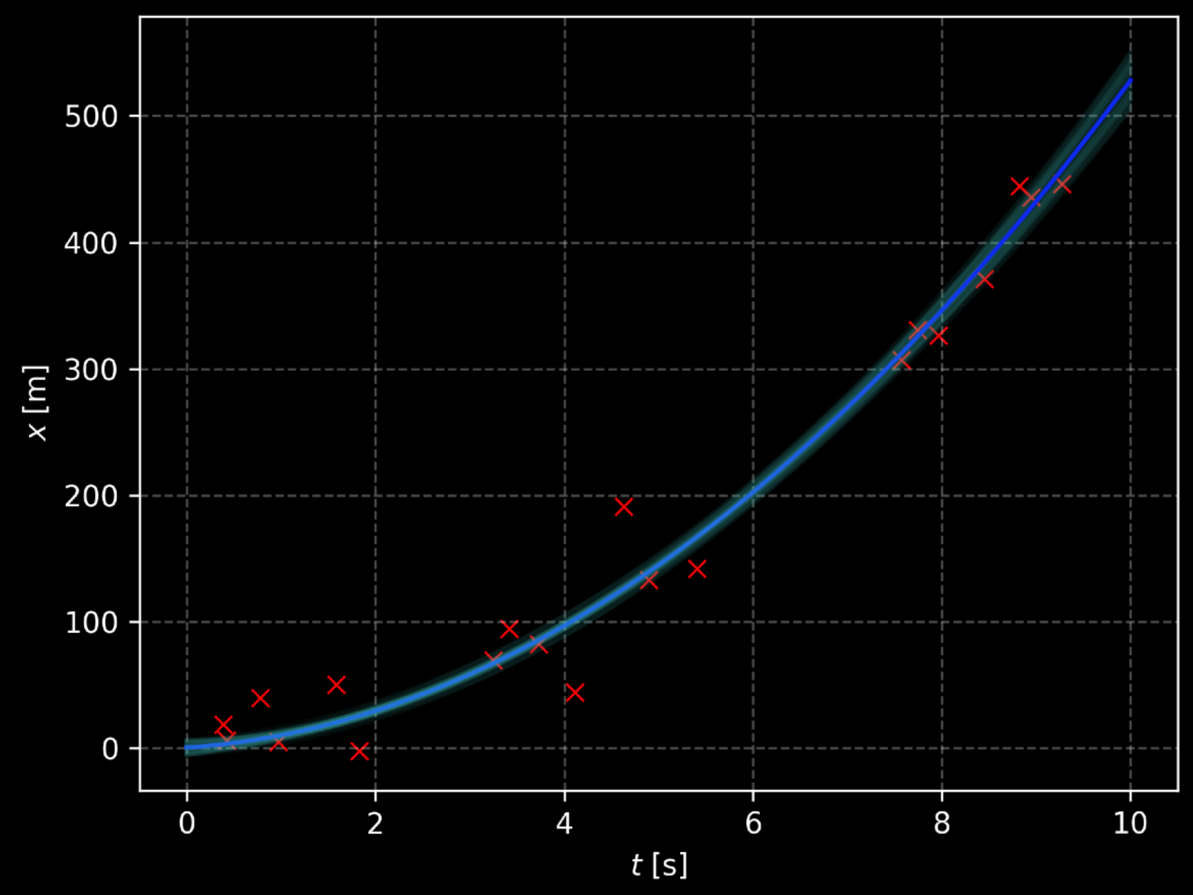
$M_1(\theta_1)$

$$p(M_1 \mid d) = \frac{p(d \mid M_1)\, p(M_1)}{p(d)}$$



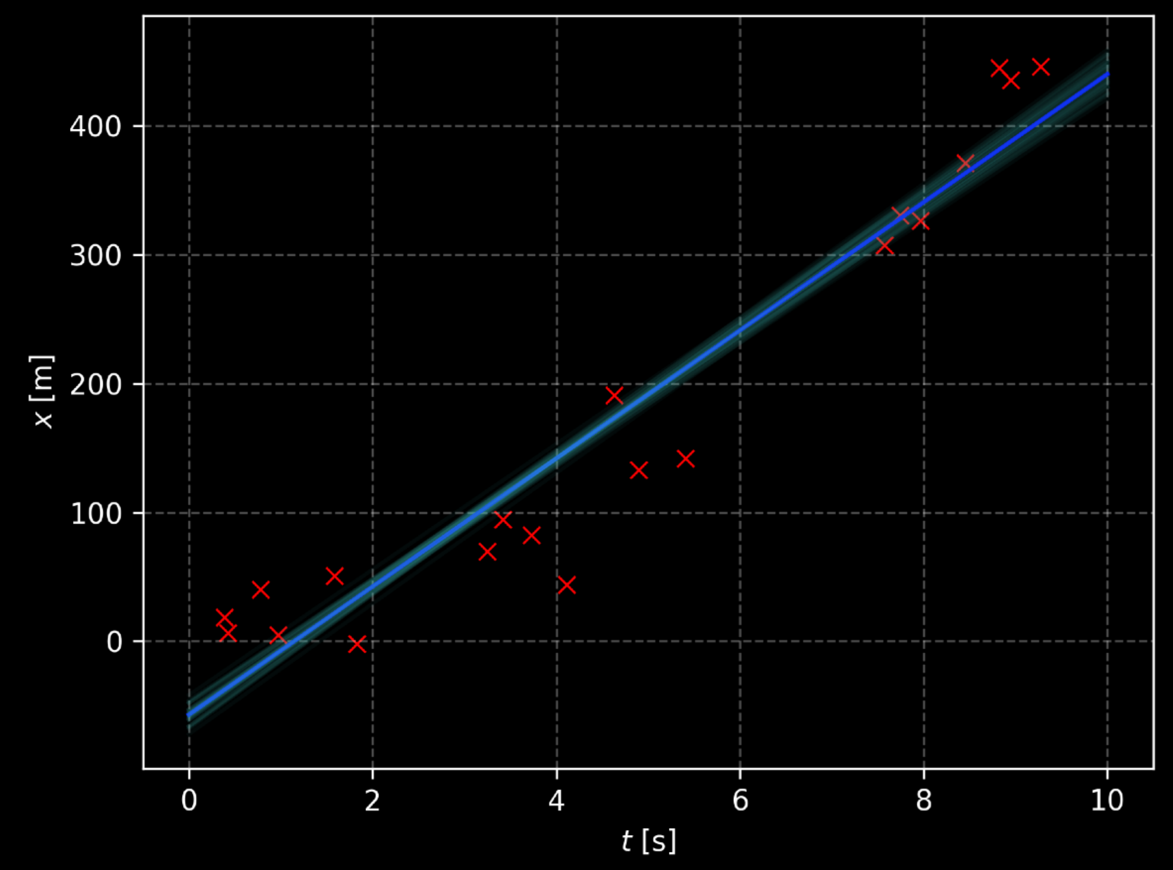
$M_2(\theta_2)$

$$p(M_2 \mid d) = \frac{p(d \mid M_2)\, p(M_2)}{p(d)}$$

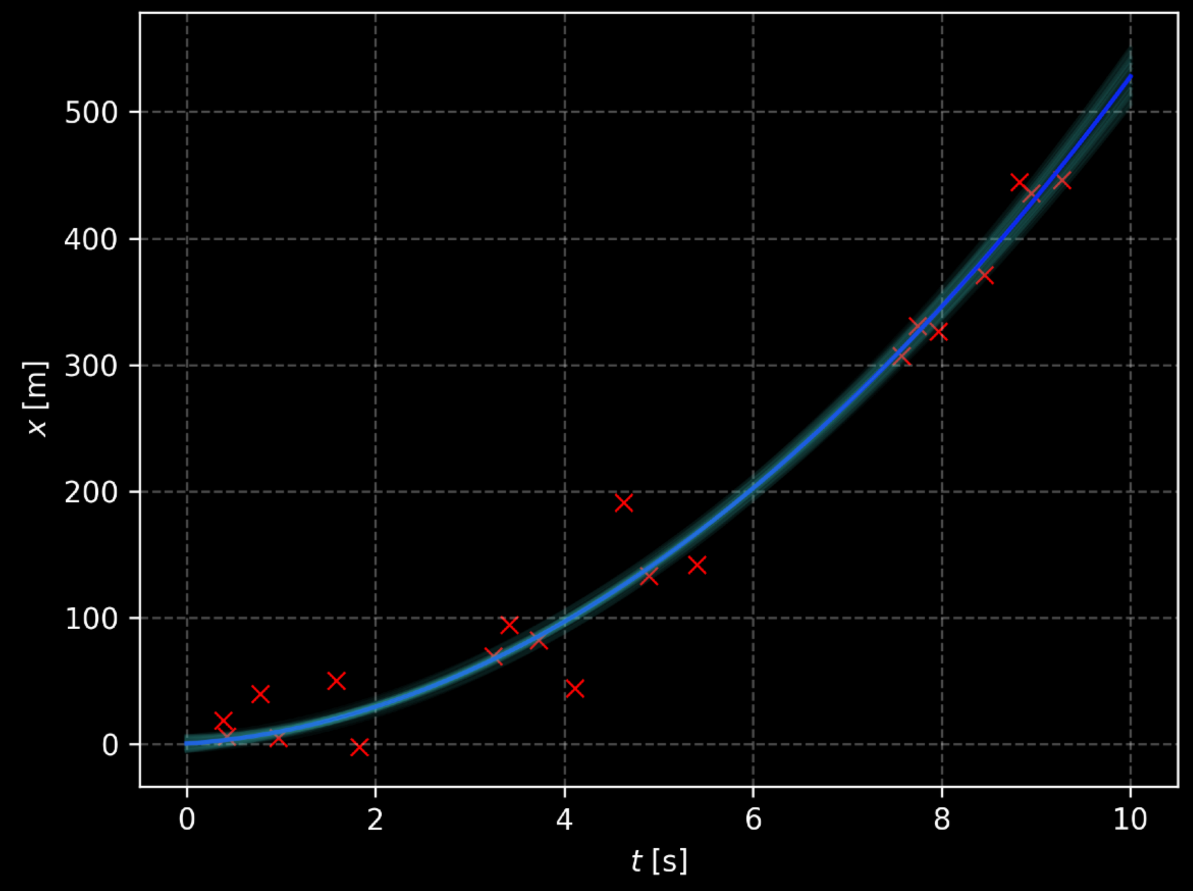


Comparing models

$M_1(\theta_1)$ $$p(M_1 | d) = \frac{p(d | M_1)\, p(M_1)}{p(d)}$$



$M_2(\theta_2)$ $$p(M_2 | d) = \frac{p(d | M_2)\, p(M_2)}{p(d)}$$

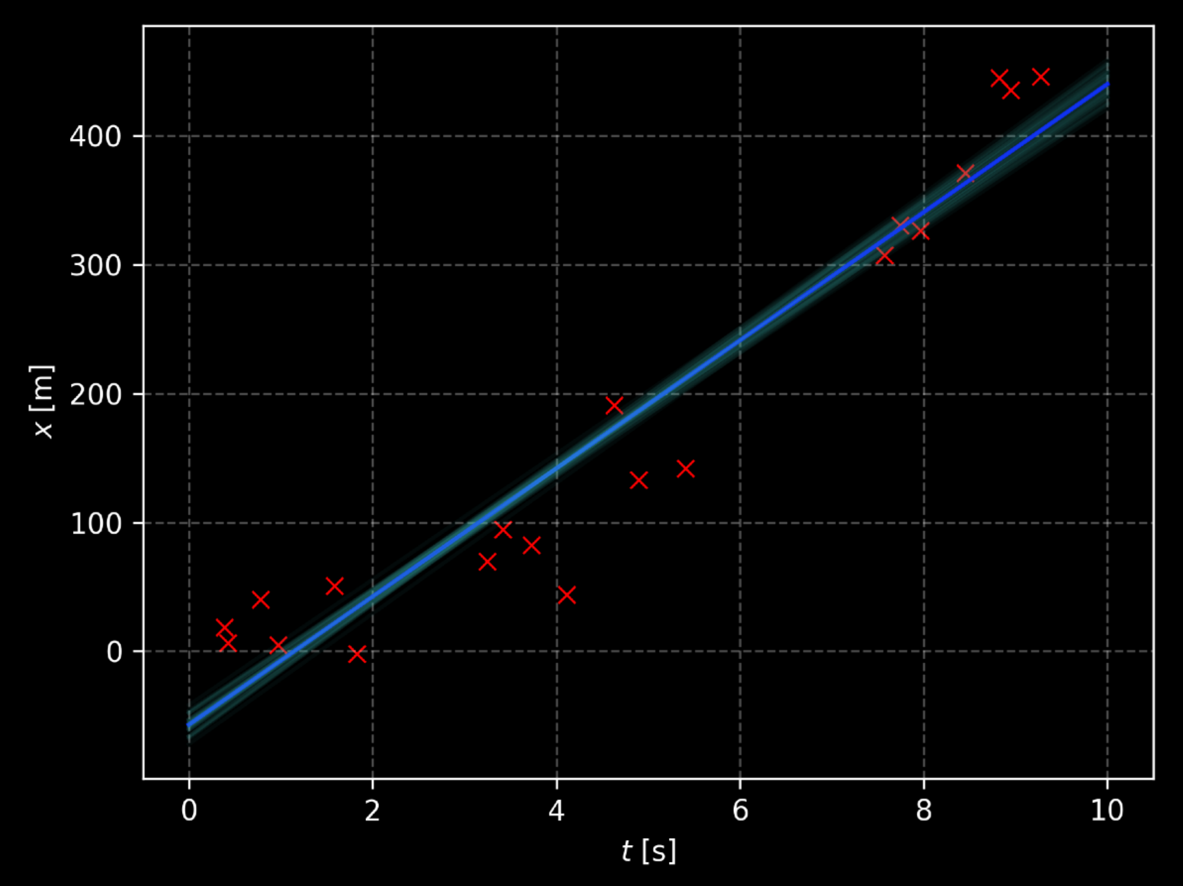


$$\mathcal{O}(M_1, M_2) := \frac{p(d | M_1)\, \cancel{p(M_1)}}{p(d | M_2)\, \cancel{p(M_2)}} = \frac{p(d | M_1)}{p(d | M_2)}$$

Comparing models

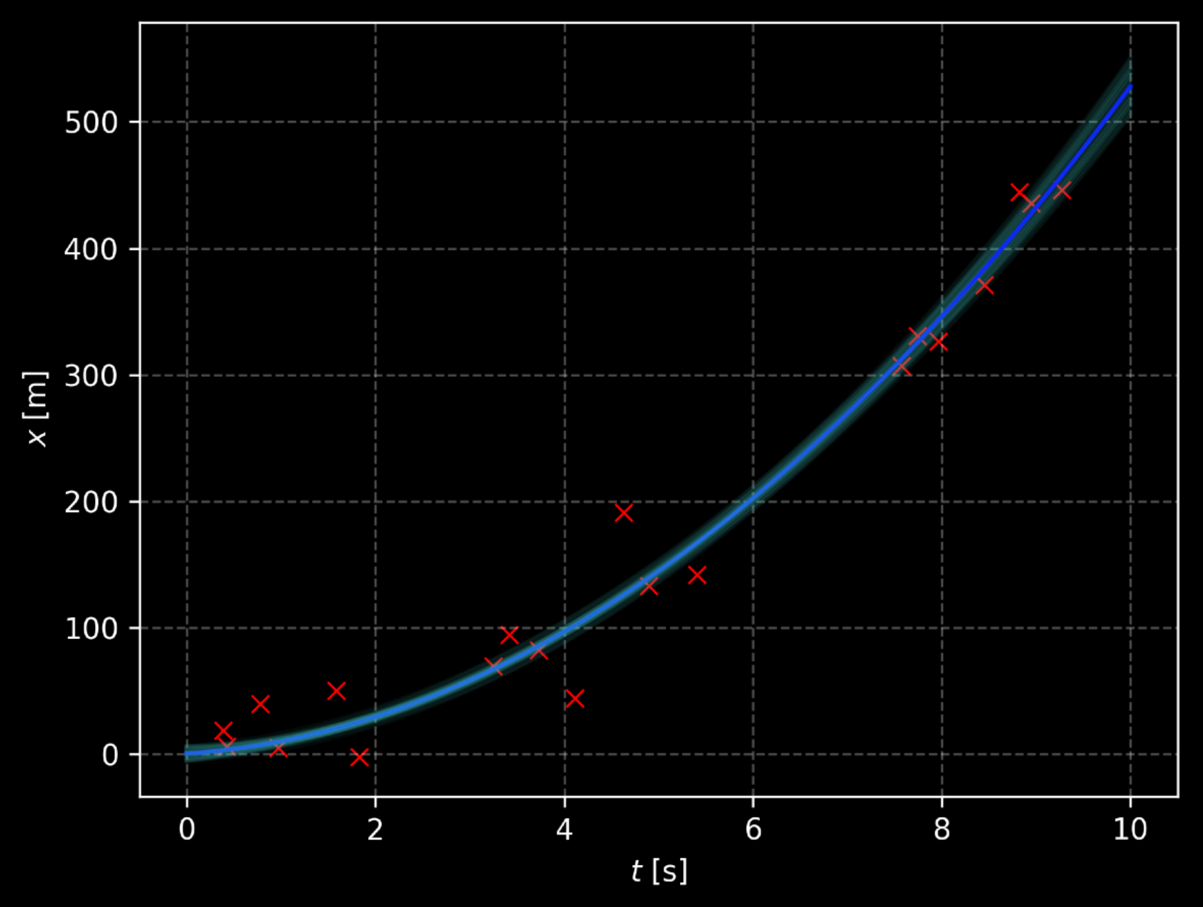
$M_1(\theta_1)$

$$p(M_1 | d) = \frac{p(d | M_1)\, p(M_1)}{p(d)}$$



$M_2(\theta_2)$

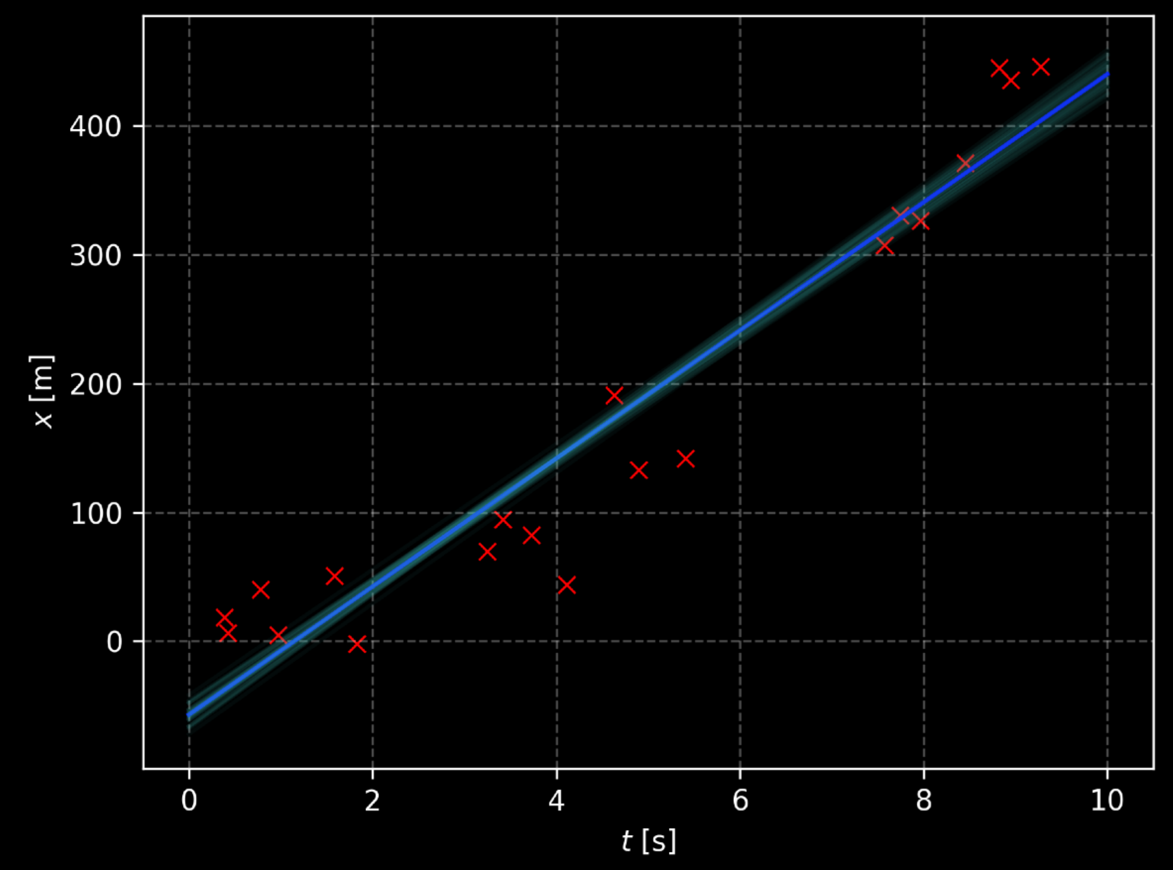
$$p(M_2 | d) = \frac{p(d | M_2)\, p(M_2)}{p(d)}$$



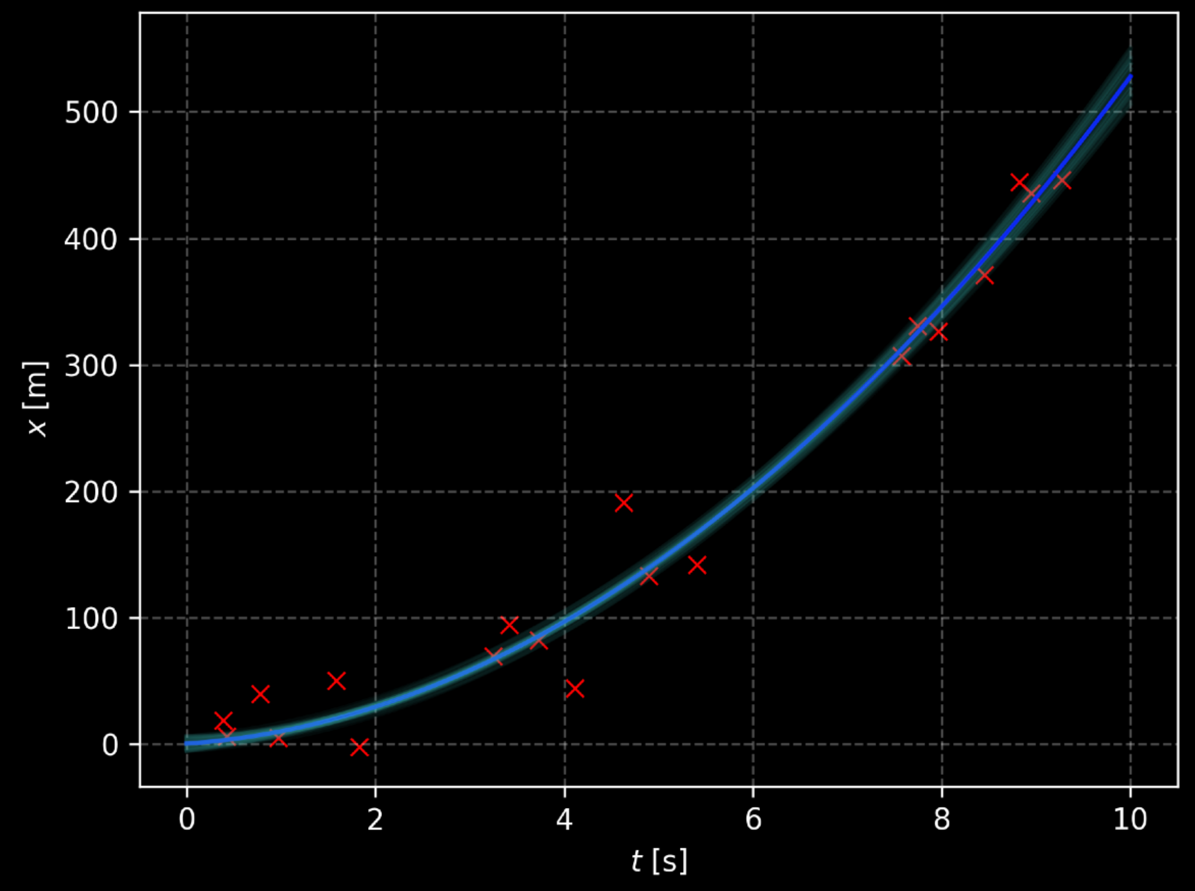
$$\mathcal{O}(M_1, M_2) := \frac{p(d | M_1)\, \cancel{p(M_1)}}{p(d | M_2)\, \cancel{p(M_2)}} = \frac{p(d | M_1)}{p(d | M_2)} \simeq e^{-24}$$

Comparing models

$$M_1(\theta_1) \qquad p(M_1\,|\,d) = \frac{p(d\,|\,M_1)\,p(M_1)}{p(d)}$$



$$M_2(\theta_2) \qquad p(M_2\,|\,d) = \frac{p(d\,|\,M_2)\,p(M_2)}{p(d)}$$



$$\mathcal{O}(M_1, M_2) := \frac{p(d\,|\,M_1)\,\cancel{p(M_1)}}{p(d\,|\,M_2)\,\cancel{p(M_2)}} = \frac{p(d\,|\,M_1)}{p(d\,|\,M_2)} \simeq e^{-24}$$

So what’s the problem?

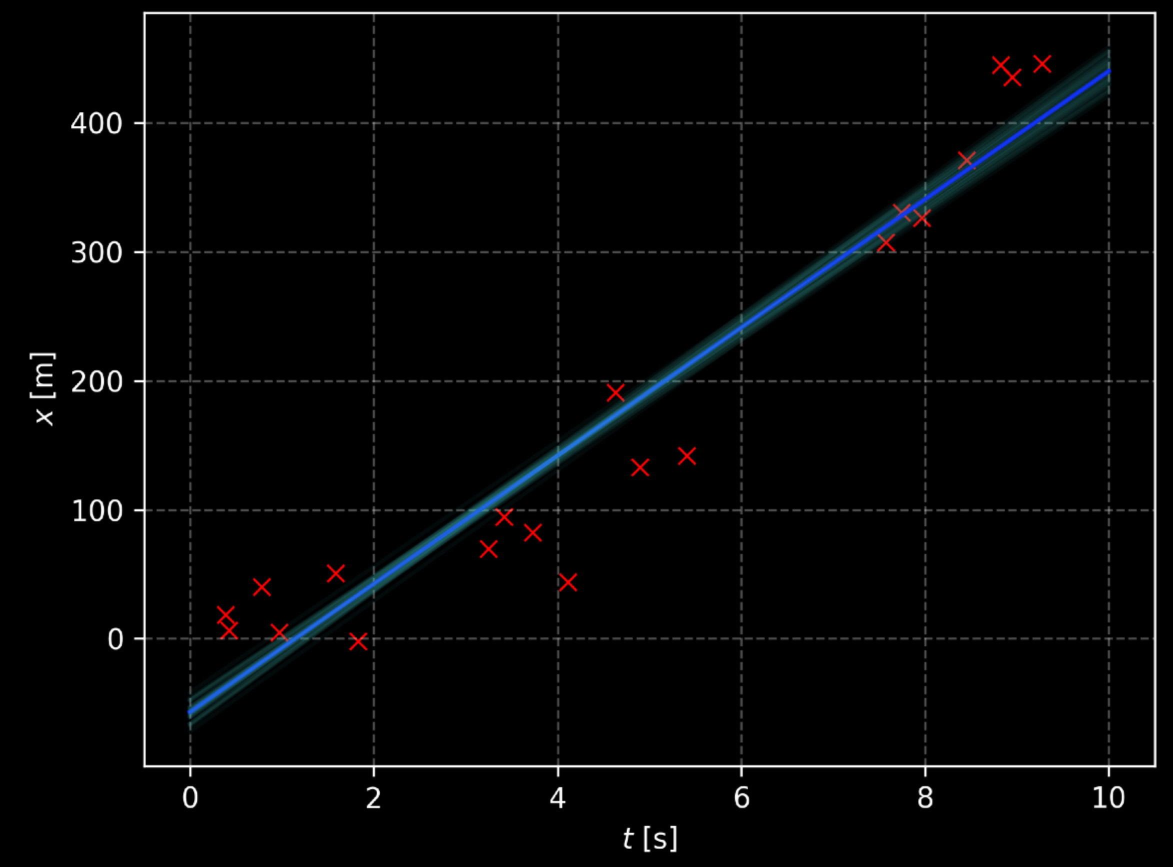
Evidence in HD

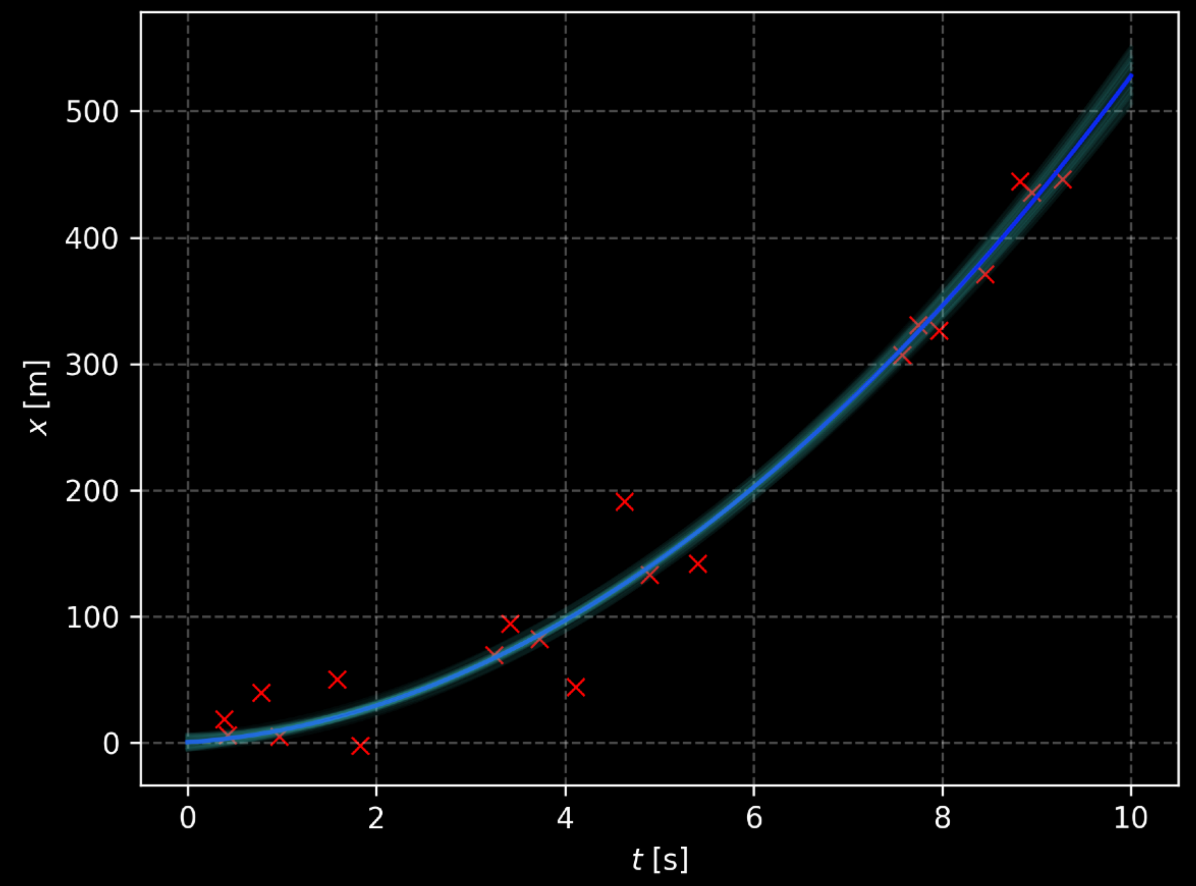
Evidence in HD

$$p(d \,|\, M) = \int p(d \,|\, \theta, M)\, p(\theta \,|\, M)\, \mathrm{d}\theta$$

Evidence in HD

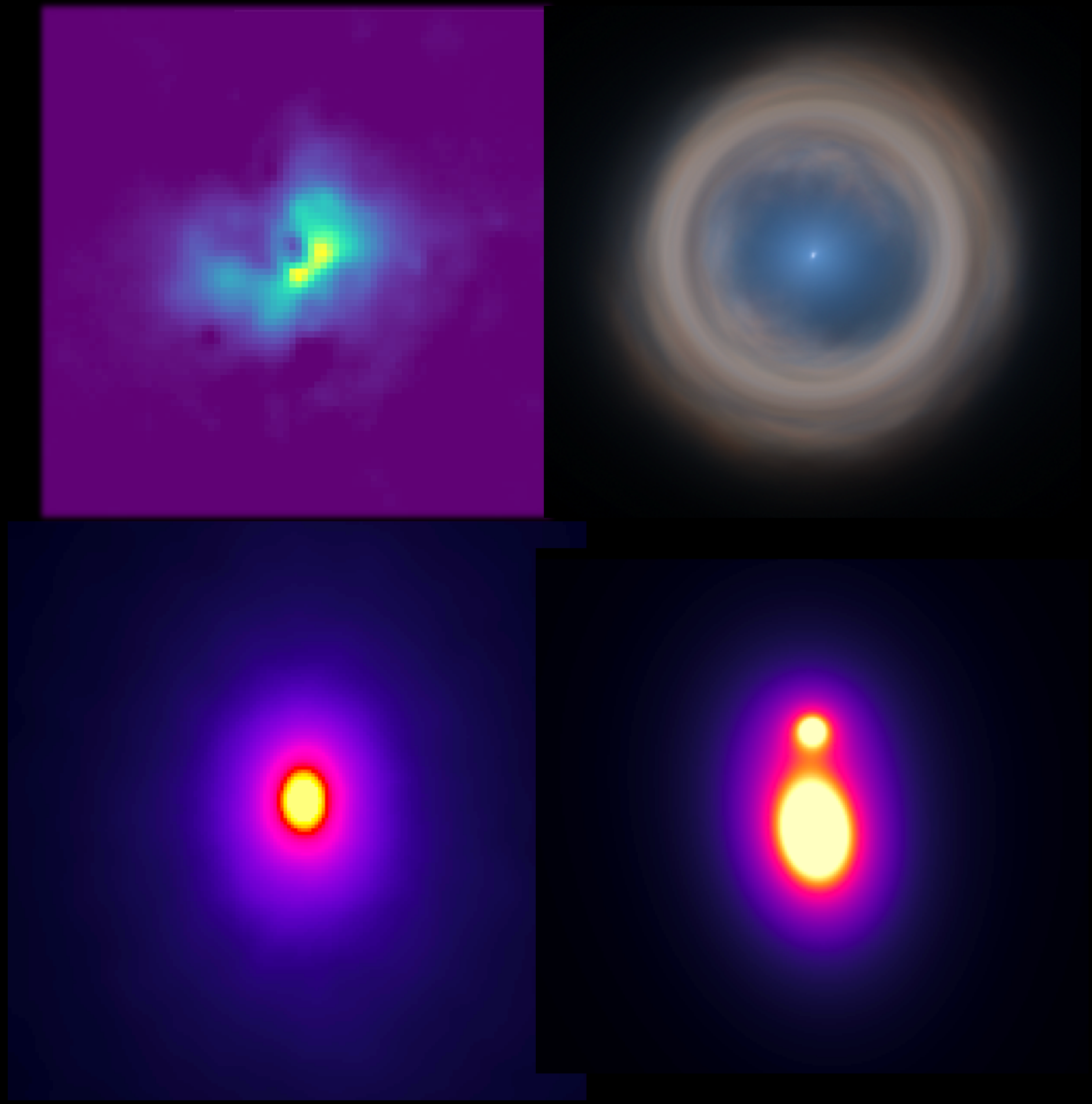
$$p(d\,|\,M) = \int p(d\,|\,\theta, M)\, p(\theta\,|\,M)\, \mathrm{d}\theta$$





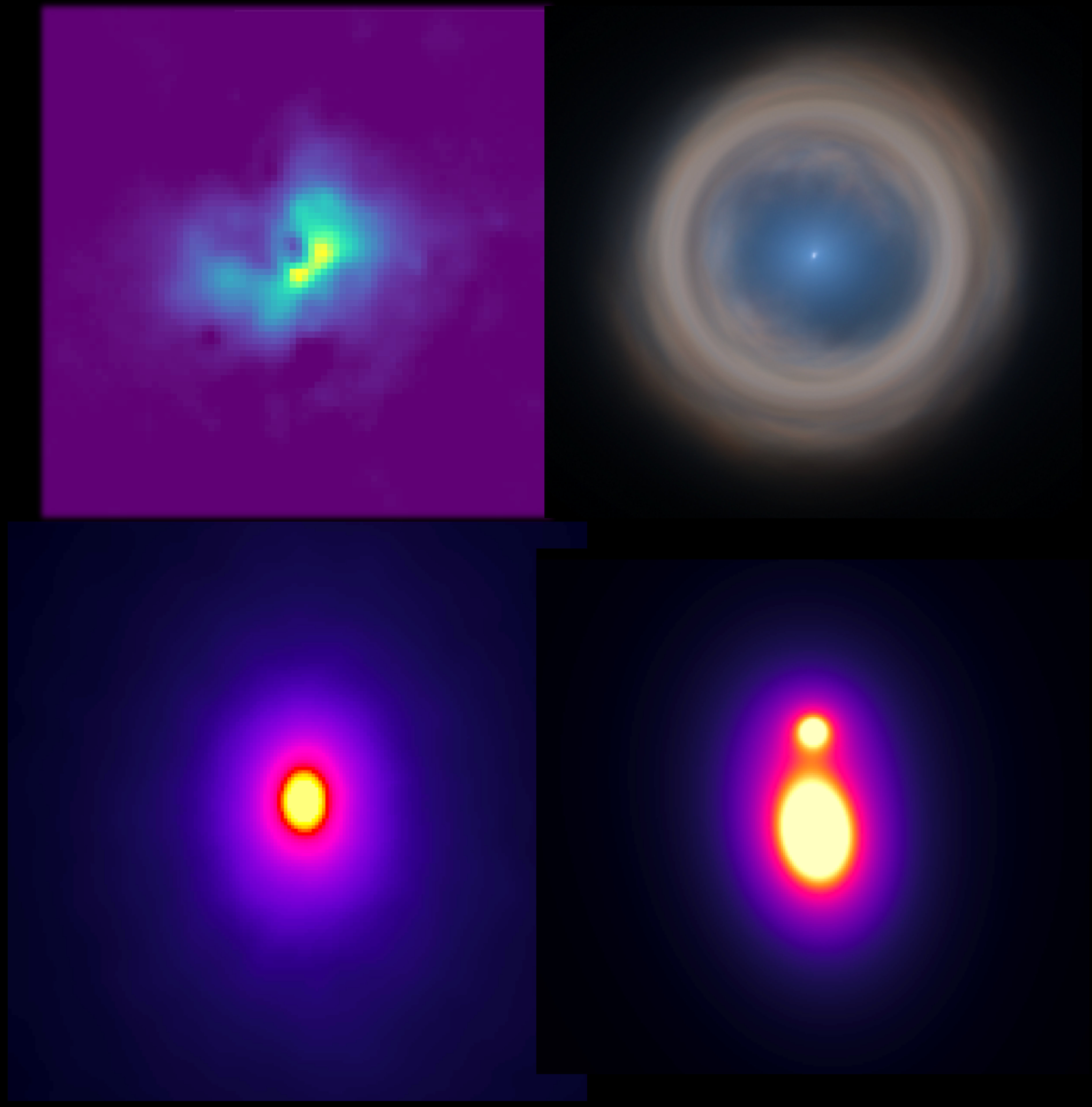
Evidence in HD

$$p(d \mid M) = \int p(d \mid \theta, M)\, p(\theta \mid M)\, \mathrm{d}\theta$$



Evidence in HD

$$p(d\,|\,M) = \int p(d\,|\,\theta, M)\, p(\theta\,|\,M)\, \mathrm{d}\theta$$



$\text{dim}(\theta) \sim 10^6 - 10^9$

Model choice

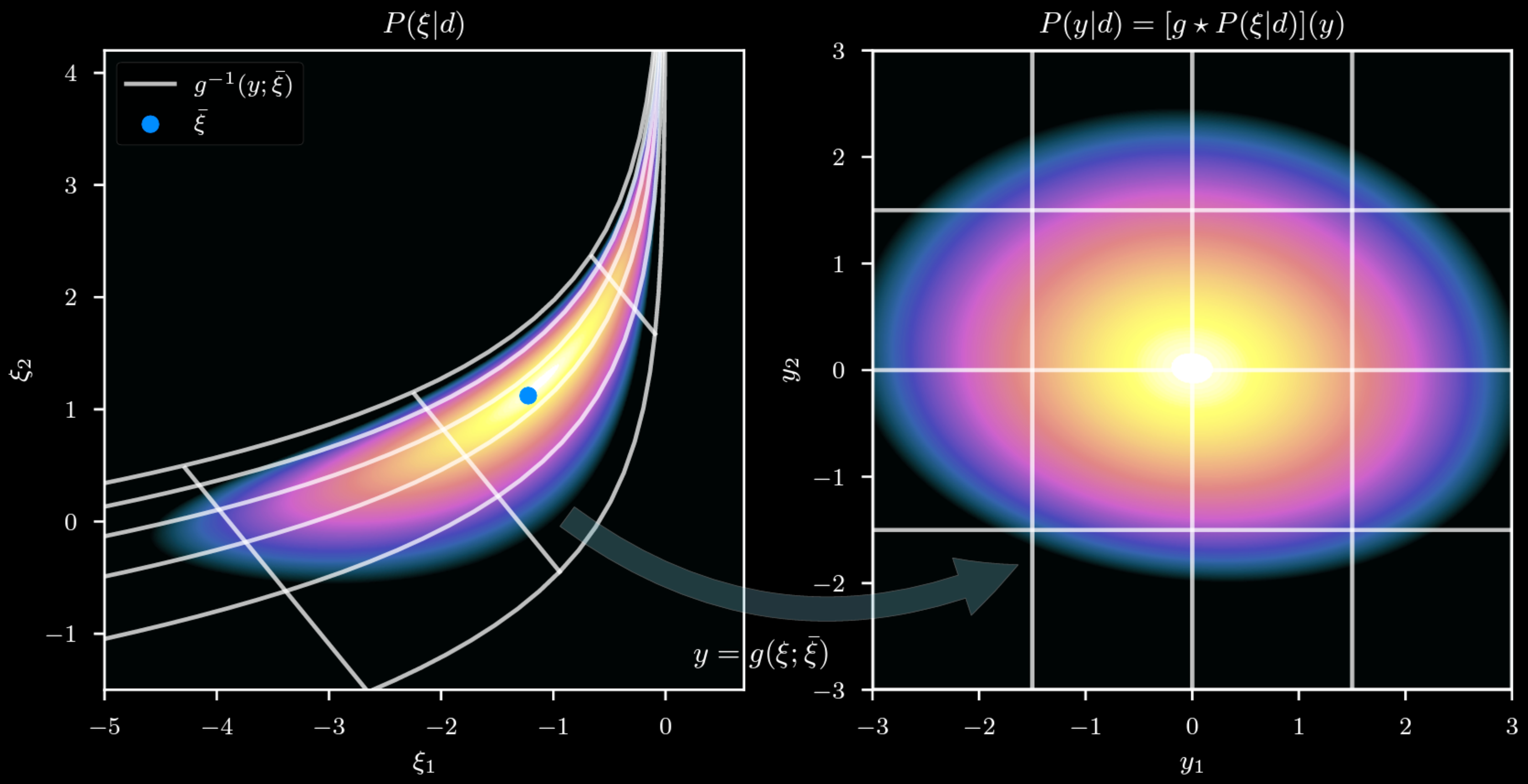
For fields :)

Model choice

For fields :)

Inference

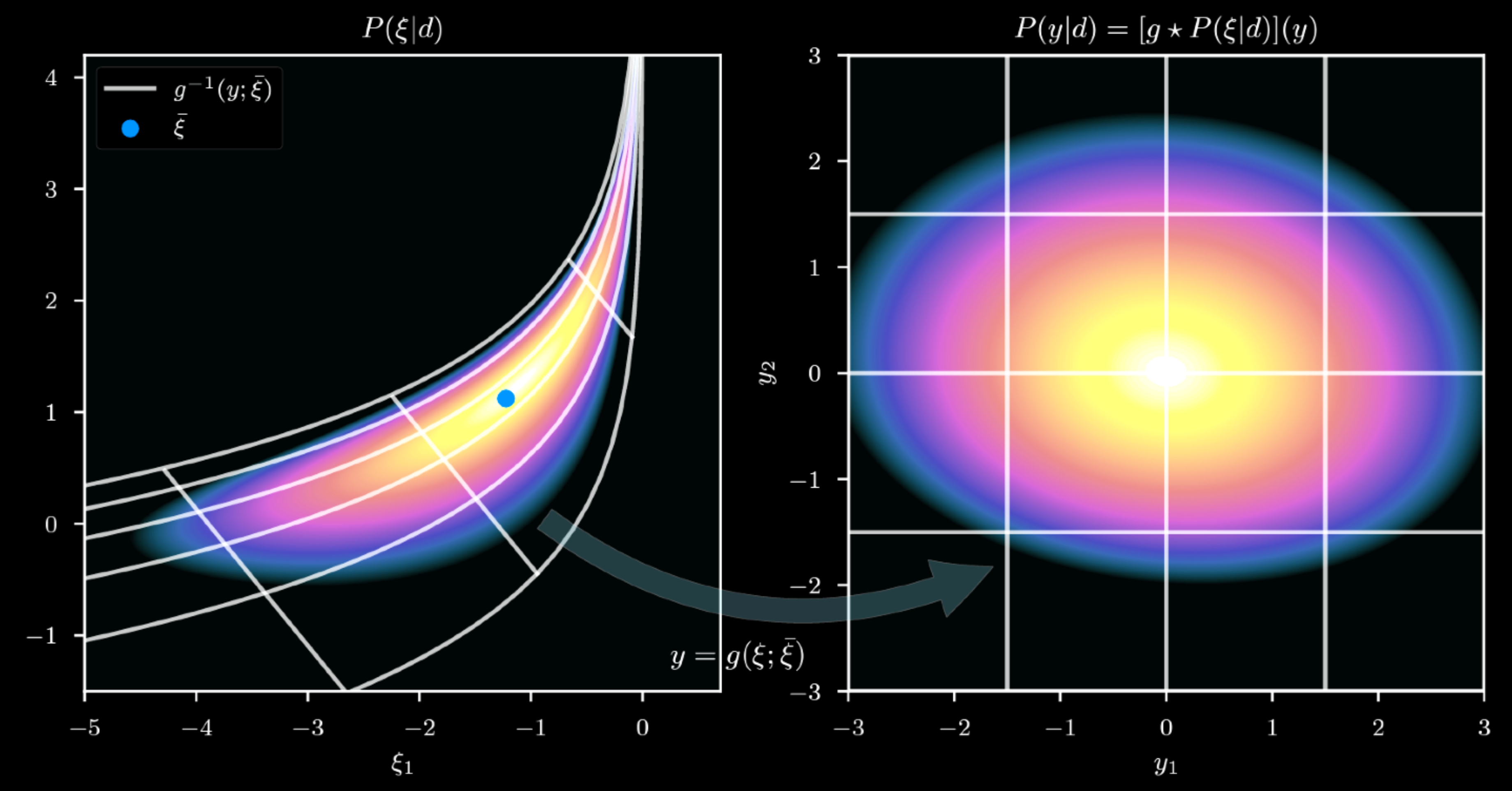
geometric Variational Inference



Credits @ Frank, P.; Leike, R.; Enßlin, T.A. Geometric Variational Inference. *Entropy* **2021**, *23*, 853.

Inference

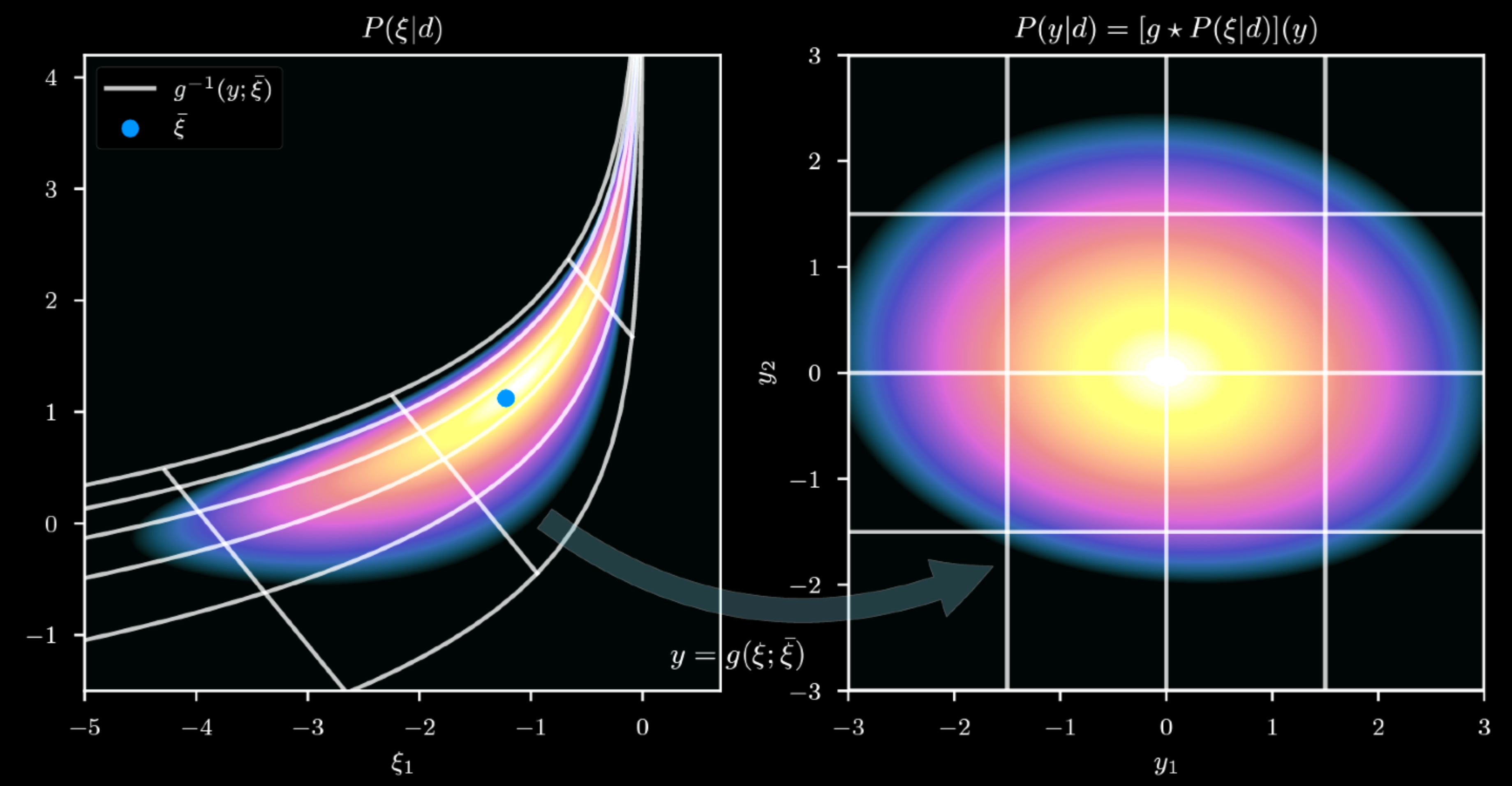
geometric Variational Inference



Inference

geometric Variational Inference

$$D_{\mathrm{KL}}(q_\phi \,\|\, p) = \int_\Theta q_\phi(\theta)\,\log \frac{q_\phi(\theta)}{p(\theta \mid d, M)}\,\mathrm{d}\theta$$

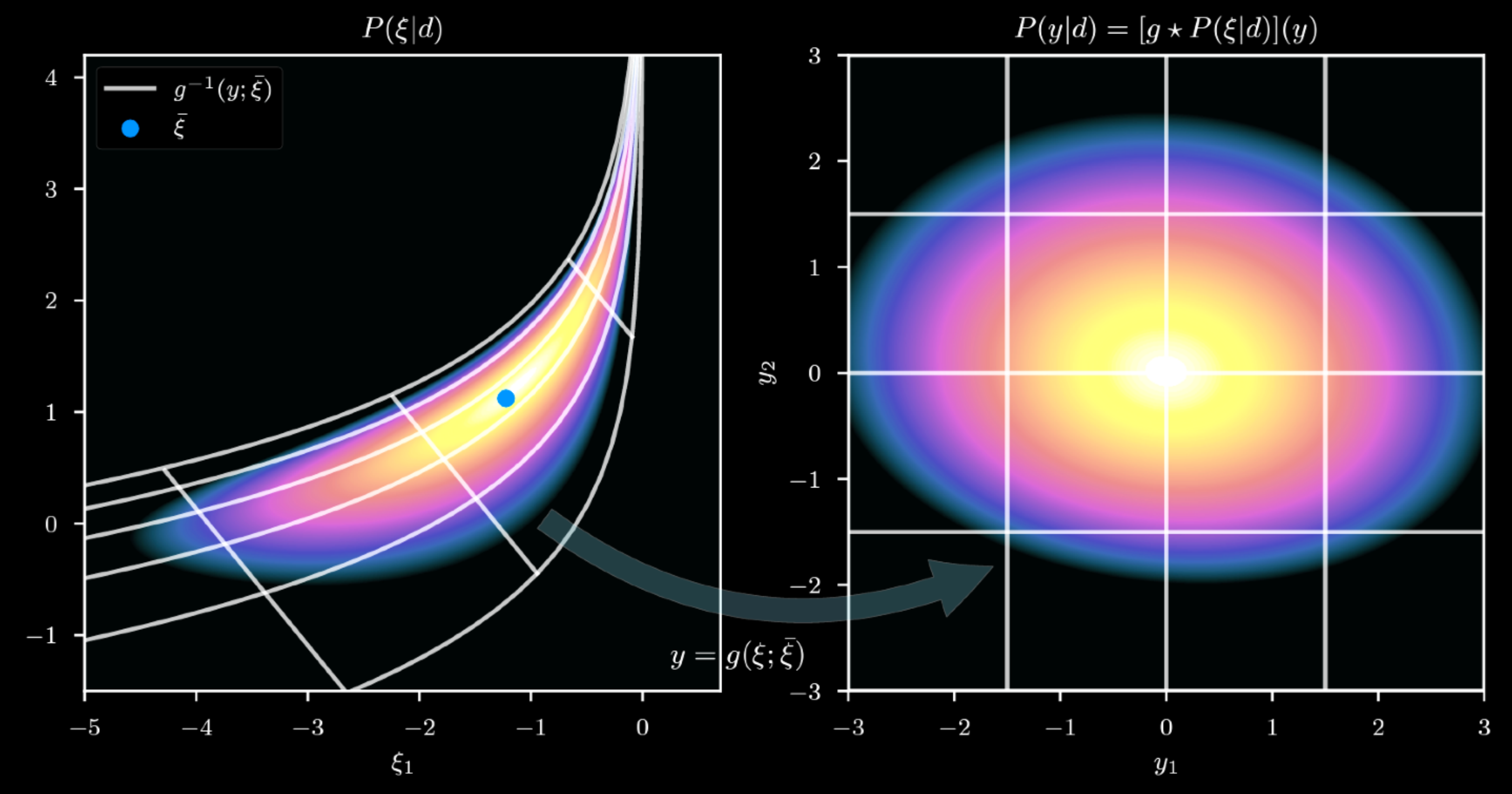


Inference

geometric Variational Inference

$$D_{\mathrm{KL}}(q_\phi \,\|\, p) = \int_\Theta q_\phi(\theta)\,\log\frac{q_\phi(\theta)}{p(\theta \mid d, M)}\,\mathrm{d}\theta$$

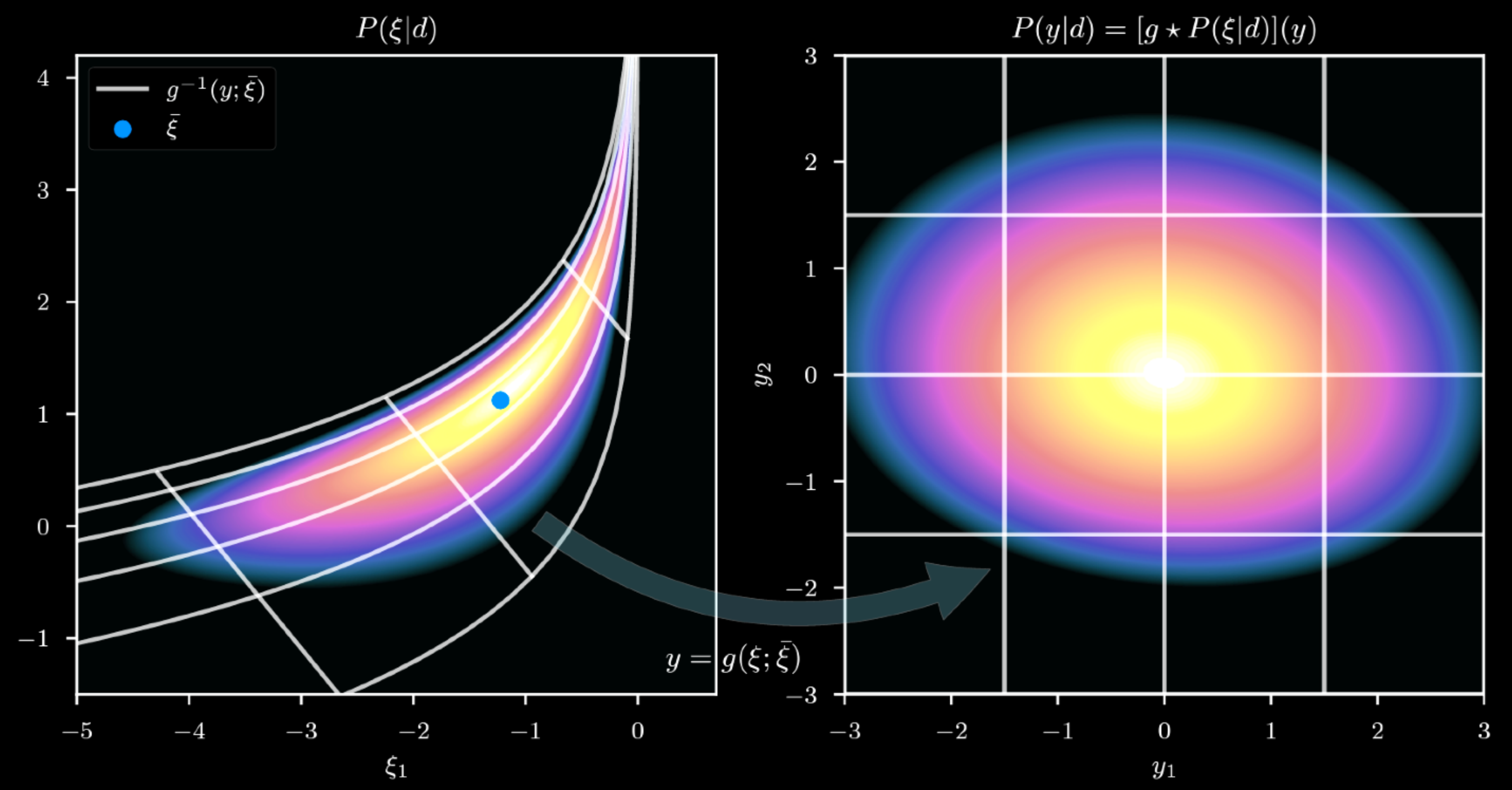
$$= \mathbb{E}_{q_\phi}\left[\log\frac{q_\phi(\theta)}{p(\theta \mid d, M)}\right]$$



Inference

geometric Variational Inference

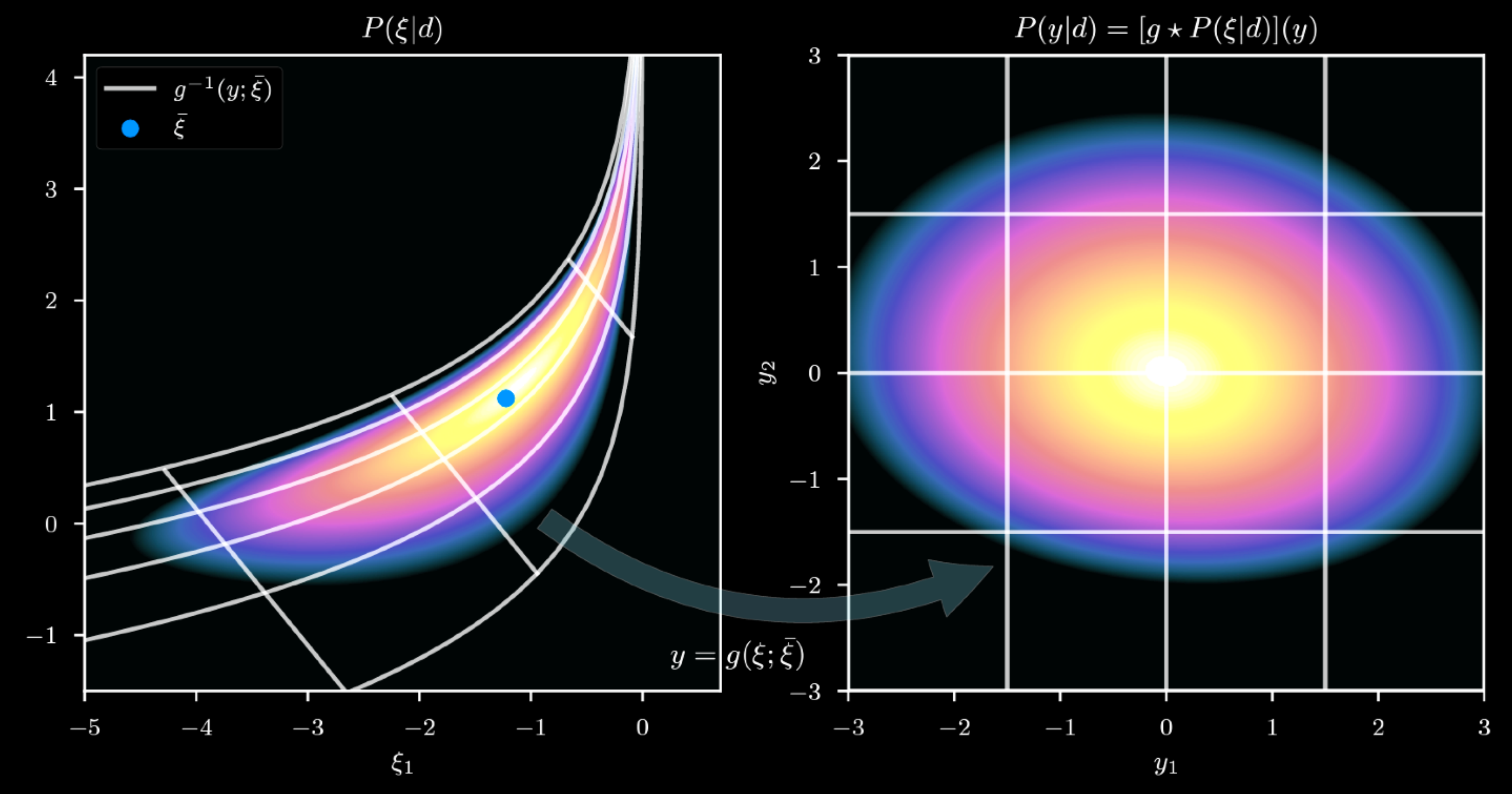
$$
\begin{aligned}
D_{\mathrm{KL}}(q_\phi \,\|\, p) &= \int_\Theta q_\phi(\theta)\,\log \frac{q_\phi(\theta)}{p(\theta \mid d, M)}\,\mathrm{d}\theta \\
&= \mathbb{E}_{q_\phi}\left[\log \frac{q_\phi(\theta)}{p(\theta \mid d, M)}\right] \\
&= \mathbb{E}_{q_\phi}\left[\log q_\phi(\theta)\right] - \mathbb{E}_{q_\phi}\left[\log p(\theta \mid d, M)\right]
\end{aligned}
$$



Inference

geometric Variational Inference

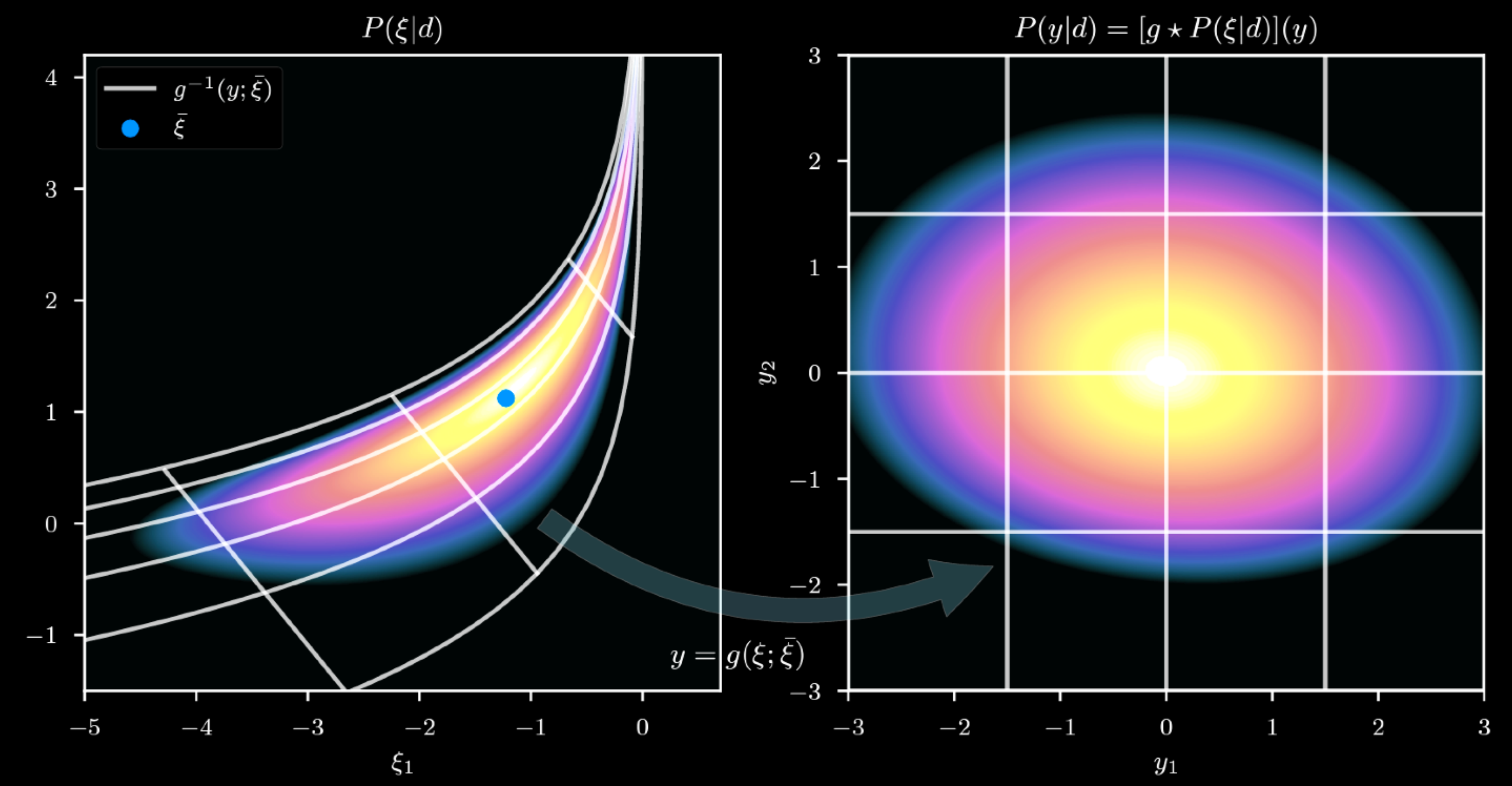
$$
\begin{aligned}
D_{\mathrm{KL}}(q_\phi \,\|\, p) &= \int_\Theta q_\phi(\theta)\, \log \frac{q_\phi(\theta)}{p(\theta \mid d, M)}\, \mathrm{d}\theta \\
&= \mathbb{E}_{q_\phi}\left[\log \frac{q_\phi(\theta)}{p(\theta \mid d, M)}\right] \\
&= \mathbb{E}_{q_\phi}\left[\log q_\phi(\theta)\right] - \mathbb{E}_{q_\phi}\left[\log p(\theta \mid d, M)\right] \\
&= \mathbb{E}_{q_\phi}\left[\log q_\phi(\theta)\right] - \mathbb{E}_{q_\phi}\left[\log \frac{p(\theta, d \mid M)}{p(d \mid M)}\right]
\end{aligned}
$$



Inference

geometric Variational Inference

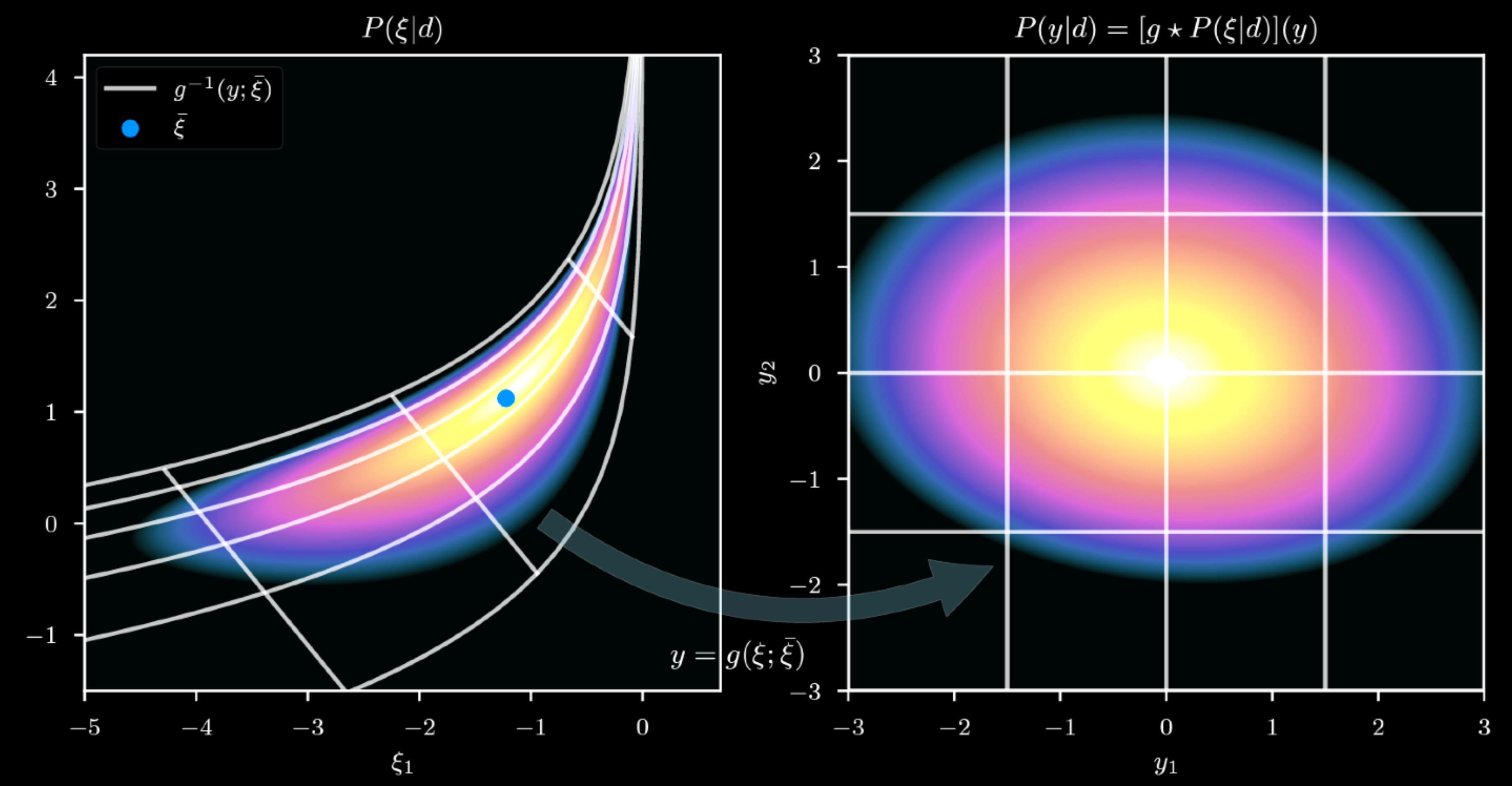
$$
\begin{aligned}
D_{\mathrm{KL}}(q_\phi \,\|\, p) &= \int_\Theta q_\phi(\theta)\, \log \frac{q_\phi(\theta)}{p(\theta \mid d, M)}\, \mathrm{d}\theta \\
&= \mathbb{E}_{q_\phi}\left[\log \frac{q_\phi(\theta)}{p(\theta \mid d, M)}\right] \\
&= \mathbb{E}_{q_\phi}\left[\log q_\phi(\theta)\right] - \mathbb{E}_{q_\phi}\left[\log p(\theta \mid d, M)\right] \\
&= \mathbb{E}_{q_\phi}\left[\log q_\phi(\theta)\right] - \mathbb{E}_{q_\phi}\left[\log \frac{p(\theta, d \mid M)}{p(d \mid M)}\right] \\
&= \mathbb{E}_{q_\phi}\left[\log q_\phi(\theta)\right] - \mathbb{E}_{q_\phi}\left[\log p(\theta, d \mid M)\right] + \log p(d \mid M)
\end{aligned}
$$



Inference

geometric Variational Inference

$$
\begin{aligned}
D_{\mathrm{KL}}(q_\phi \,\|\, p) &= \int_\Theta q_\phi(\theta)\, \log \frac{q_\phi(\theta)}{p(\theta \mid d, M)}\, \mathrm{d}\theta \\
&= \mathbb{E}_{q_\phi}\left[\log \frac{q_\phi(\theta)}{p(\theta \mid d, M)}\right] \\
&= \mathbb{E}_{q_\phi}\left[\log q_\phi(\theta)\right] - \mathbb{E}_{q_\phi}\left[\log p(\theta \mid d, M)\right] \\
&= \mathbb{E}_{q_\phi}\left[\log q_\phi(\theta)\right] - \mathbb{E}_{q_\phi}\left[\log \frac{p(\theta, d \mid M)}{p(d \mid M)}\right] \\
&= \mathbb{E}_{q_\phi}\left[\log q_\phi(\theta)\right] - \mathbb{E}_{q_\phi}\left[\log p(\theta, d \mid M)\right] + \log p(d \mid M) \\
&= -\,\mathbb{E}_{q_\phi}\left[\mathscr{H}_q(\theta)\right] + \mathbb{E}_{q_\phi}\left[\mathscr{H}_p(\theta, d \mid M)\right] + \log p(d \mid M)\,.
\end{aligned}
$$



Inference

geometric Variational Inference

Inference

geometric Variational Inference

$$D_{\mathrm{KL}}(q_\phi \,\|\, p) = -\,\mathbb{E}_{q_\phi}\left[\mathscr{H}_q(\theta)\right] + \mathbb{E}_{q_\phi}\left[\mathscr{H}_p(\theta, d \mid M)\right] + \log p(d \mid M)$$

Inference

geometric Variational Inference

$$D_{\mathrm{KL}}(q_\phi \,\|\, p) = \underbrace{-\,\mathbb{E}_{q_\phi}\left[\mathscr{H}_q(\theta)\right] + \mathbb{E}_{q_\phi}\left[\mathscr{H}_p(\theta, d \mid M)\right]}_{-\mathrm{ELBO}} + \log p(d \mid M)$$

Inference

geometric Variational Inference

$$D_{\mathrm{KL}}(q_\phi \,\|\, p) = \underbrace{-\,\mathbb{E}_{q_\phi}\Big[\mathscr{H}_q(\theta)\Big] + \mathbb{E}_{q_\phi}\Big[\mathscr{H}_p(\theta, d \mid M)\Big]}_{-\mathrm{ELBO}} + \log p(d \mid M)$$

$$\log p(d \mid M) = D_{\mathrm{KL}}(q_\phi \,\|\, p) + \mathrm{ELBO}$$

Inference

geometric Variational Inference

$$D_{\mathrm{KL}}(q_\phi \,\|\, p) = \underbrace{-\,\mathbb{E}_{q_\phi}\Big[\mathscr{H}_q(\theta)\Big] + \mathbb{E}_{q_\phi}\Big[\mathscr{H}_p(\theta, d \mid M)\Big]}_{-\mathrm{ELBO}} + \log p(d \mid M)$$

$$\log p(d \mid M) \leq \mathrm{ELBO}$$

Comparing models

Fields

Comparing models

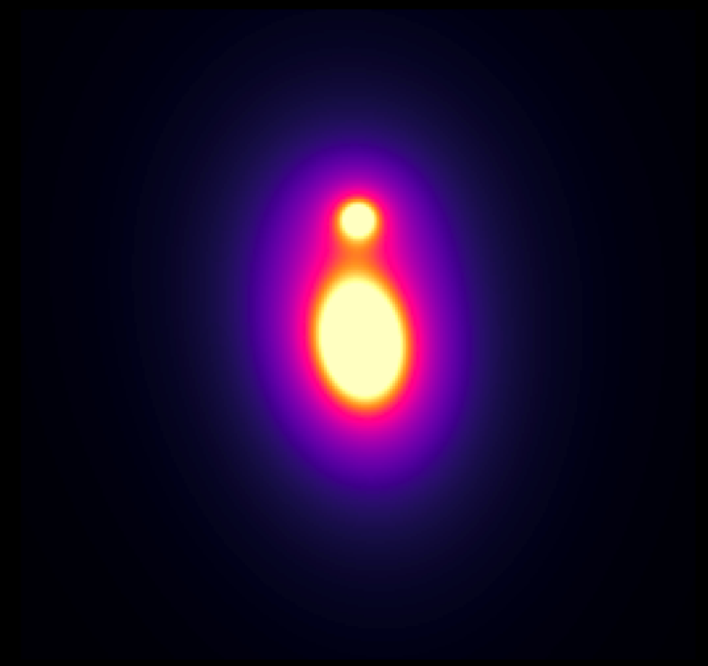
Fields

$$\log \mathcal{O}(M_1, M_2 \,|\, d) = \left[\mathrm{ELBO}_1 - \mathrm{ELBO}_2\right] + \left[\mathcal{D}_{\mathrm{KL},1} - \mathcal{D}_{\mathrm{KL},2}\right] + \log \frac{p(M_1)}{p(M_2)}$$

Comparing models

Fields

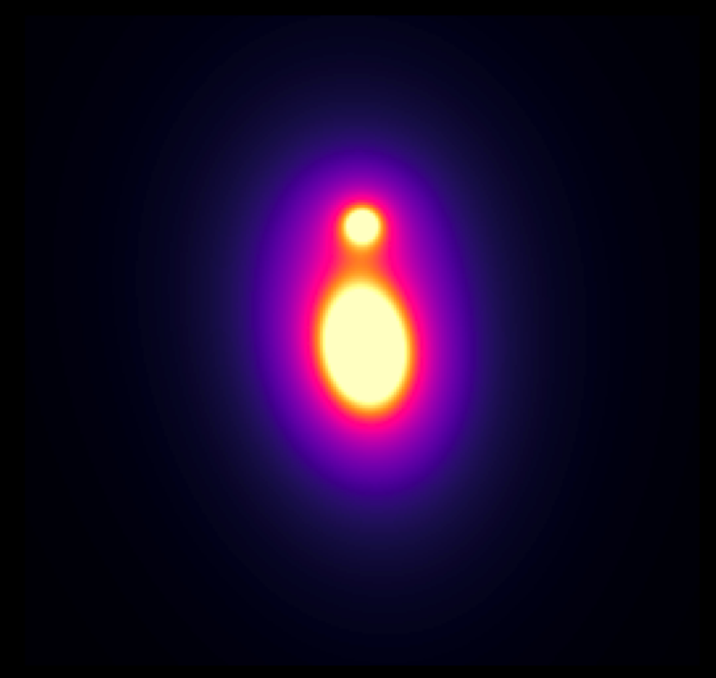
$$\log \mathcal{O}(M_1, M_2 \,|\, d) = \left[\mathrm{ELBO}_1 - \mathrm{ELBO}_2\right] + \left[\mathscr{D}_{\mathrm{KL},1} - \mathscr{D}_{\mathrm{KL},2}\right] + \log \frac{p(M_1)}{p(M_2)}$$

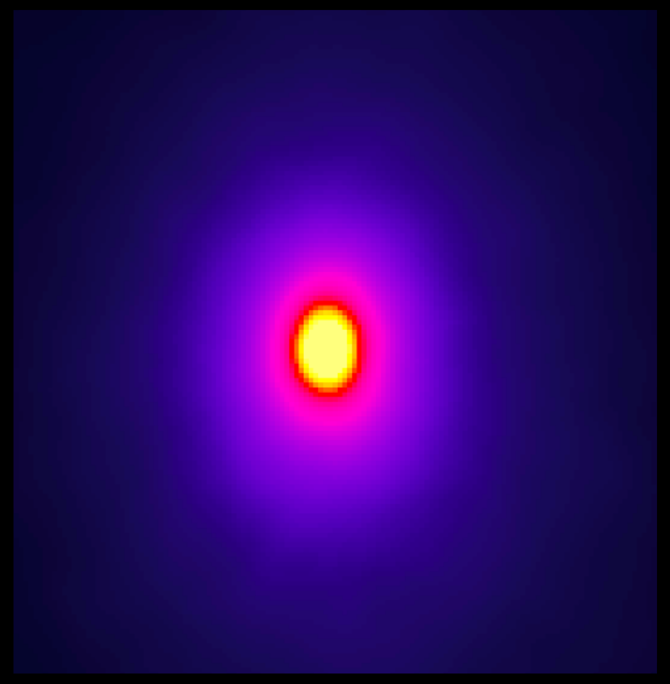


Comparing models

Fields

$$\log \mathcal{O}(M_1, M_2 \,|\, d) = \left[\mathrm{ELBO}_1 - \mathrm{ELBO}_2\right] + \left[\mathscr{D}_{\mathrm{KL},1} - \mathscr{D}_{\mathrm{KL},2}\right] + \log \frac{p(M_1)}{p(M_2)}$$

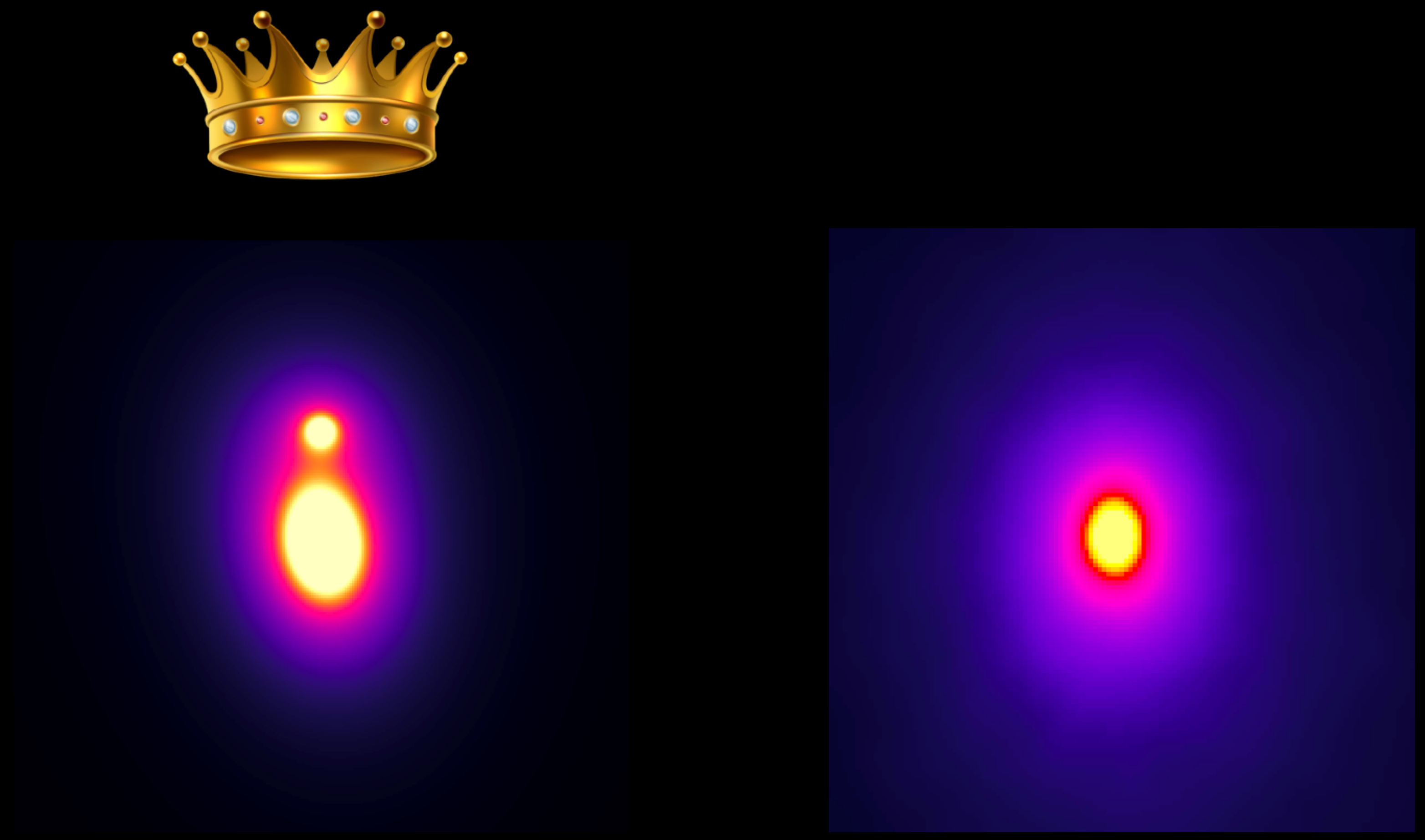




Comparing models

Fields

$$\log \mathcal{O}(M_1, M_2 \,|\, d) = \left[\mathrm{ELBO}_1 - \mathrm{ELBO}_2\right] + \left[\mathcal{D}_{\mathrm{KL},1} - \mathcal{D}_{\mathrm{KL},2}\right] + \log \frac{p(M_1)}{p(M_2)}$$

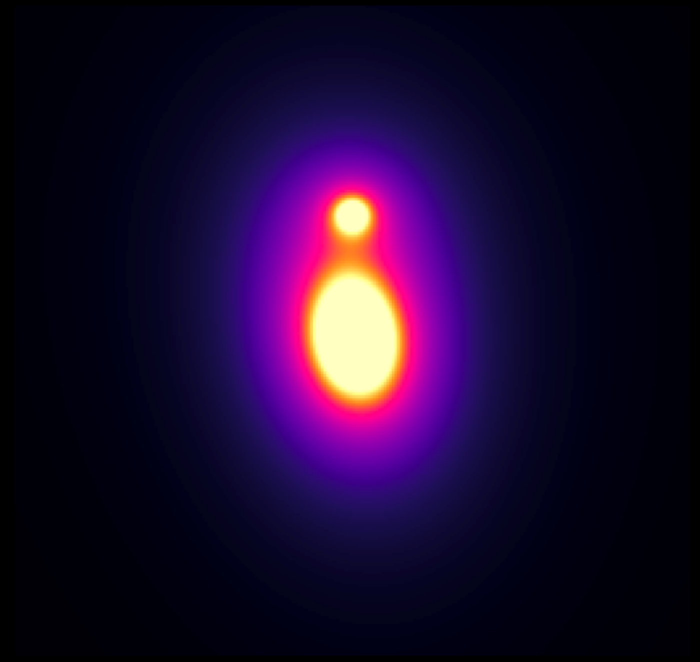


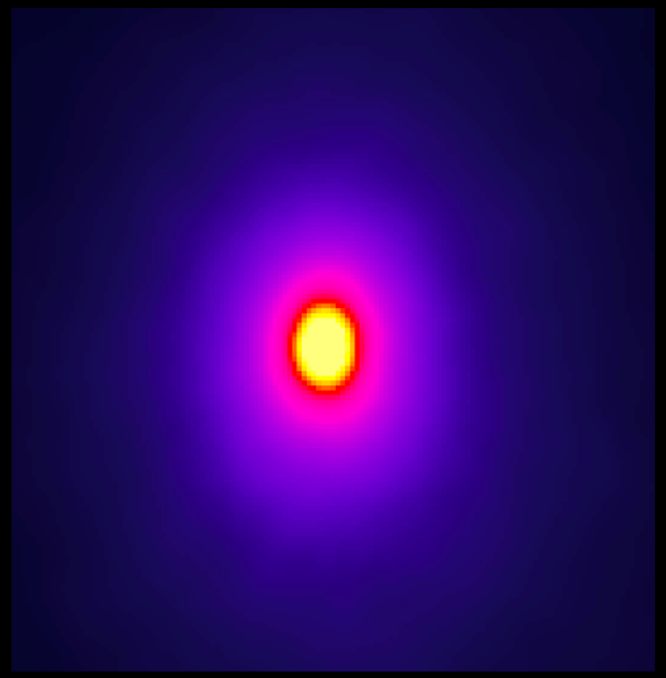
$$\log \mathcal{O}(M_1, M_2 \,|\, d) \simeq 28$$

Comparing models

Fields

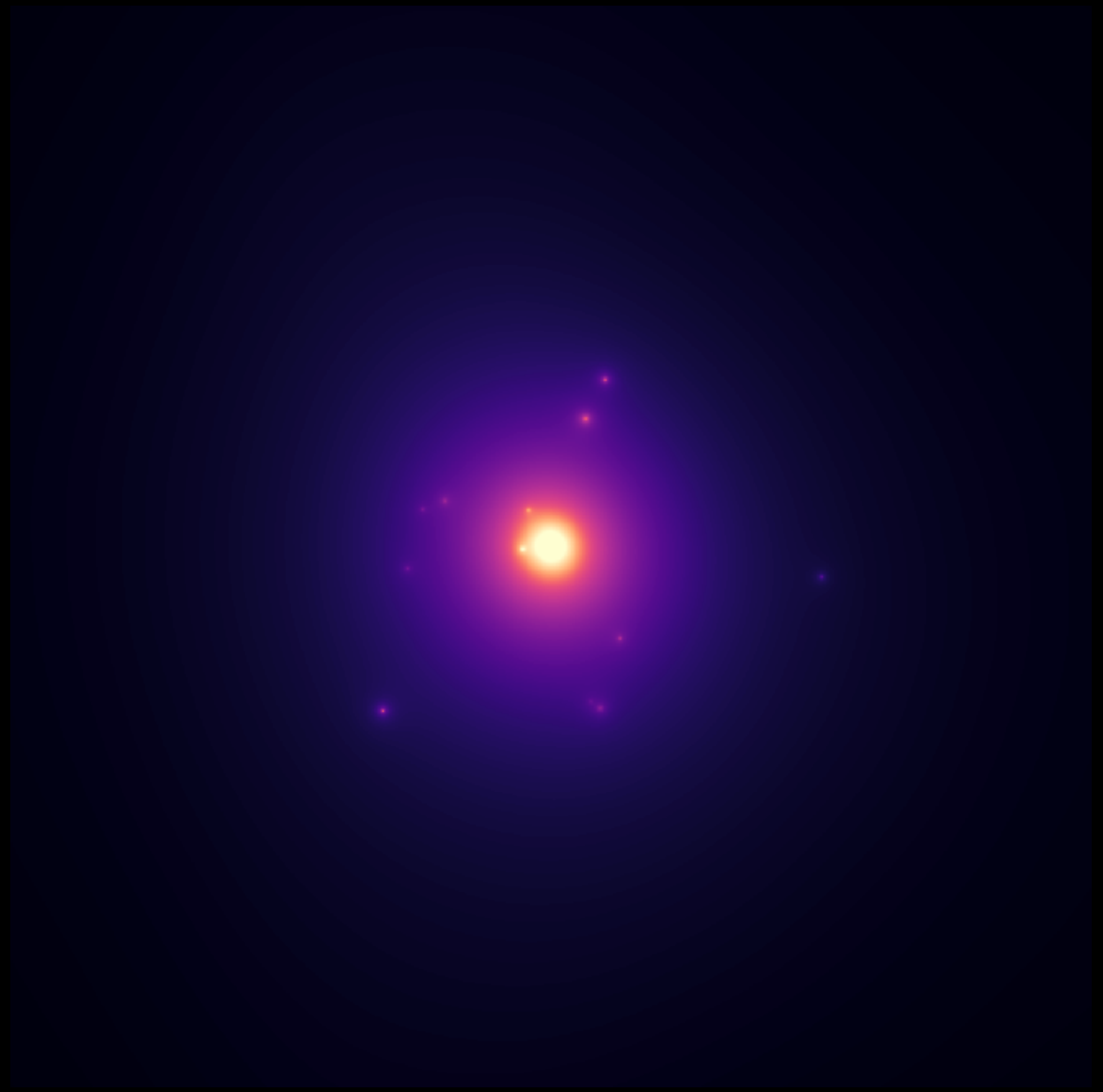
$$\log \mathcal{O}(M_1, M_2 \,|\, d) = \left[\text{ELBO}_1 - \text{ELBO}_2\right] + \left[\mathcal{D}_{\text{KL},1} - \mathcal{D}_{\text{KL},2}\right] + \log \frac{p(M_1)}{p(M_2)}$$

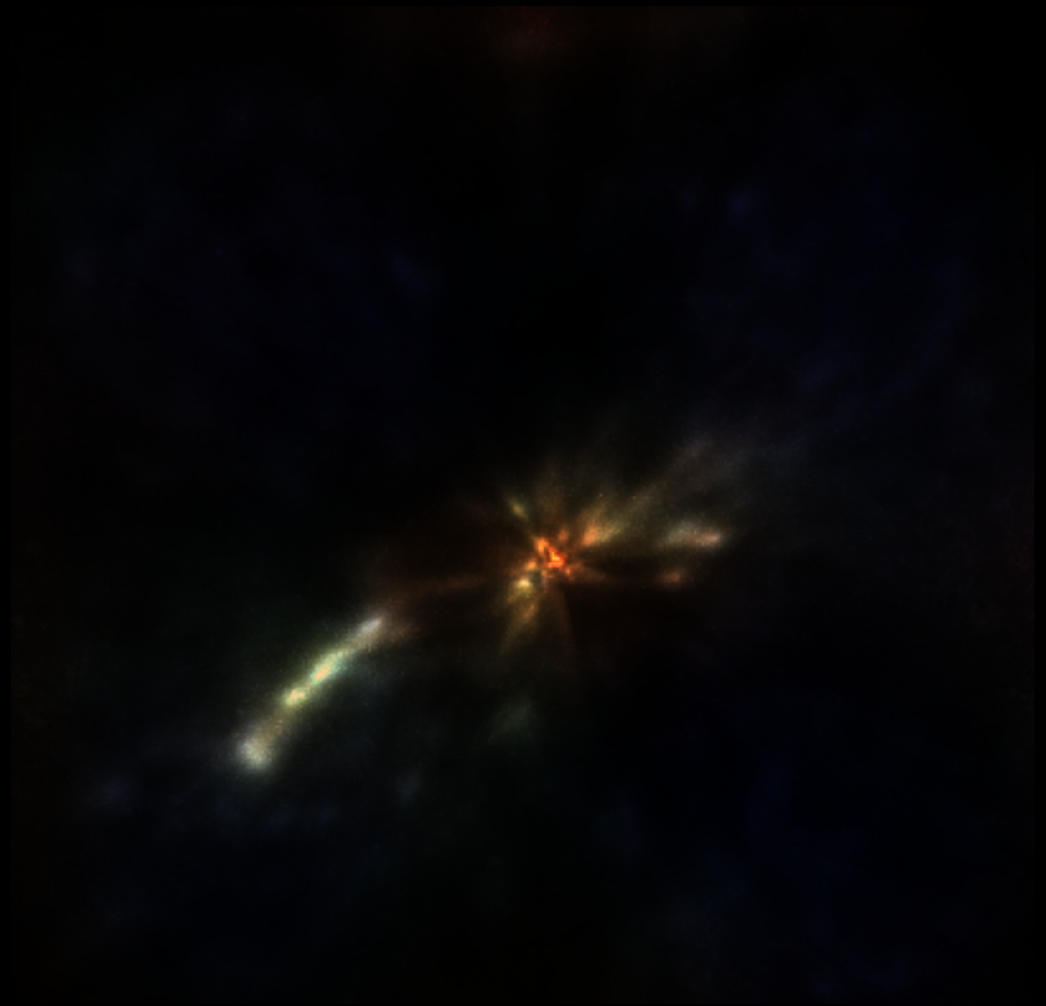


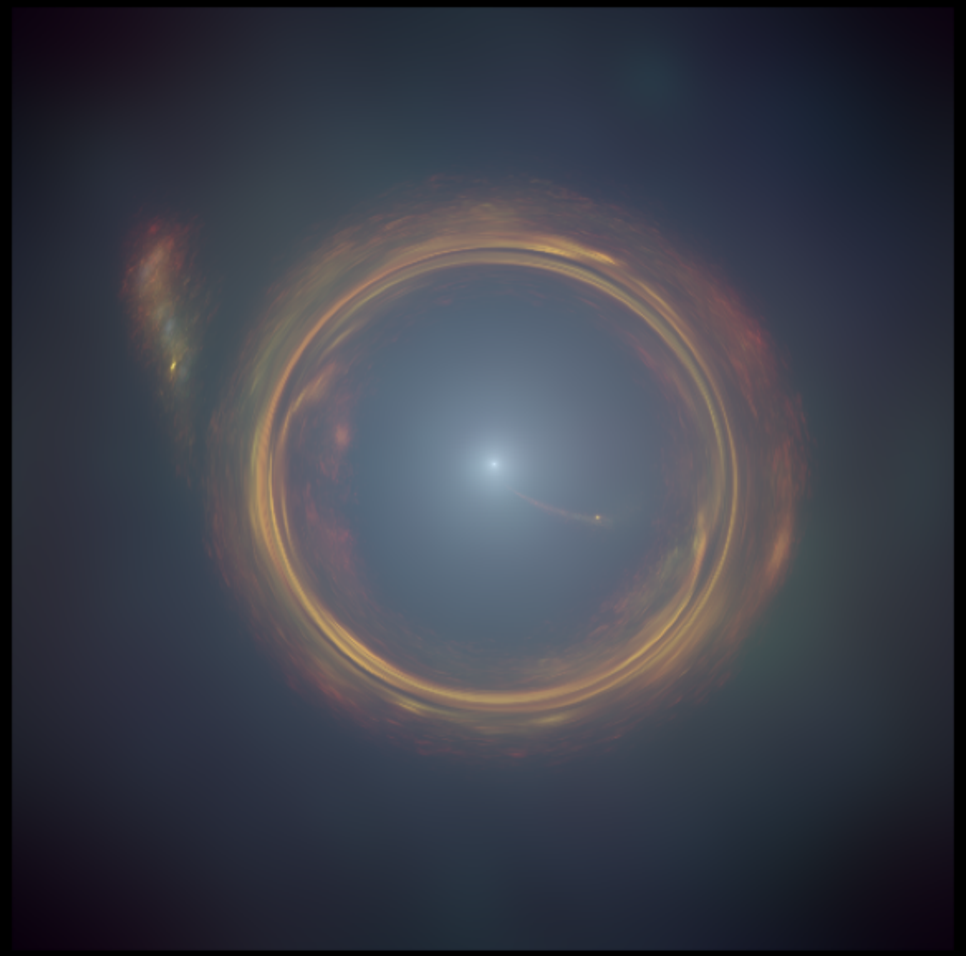


$$\log \mathcal{O}(M_1, M_2 \,|\, d) \simeq 28$$

Take home

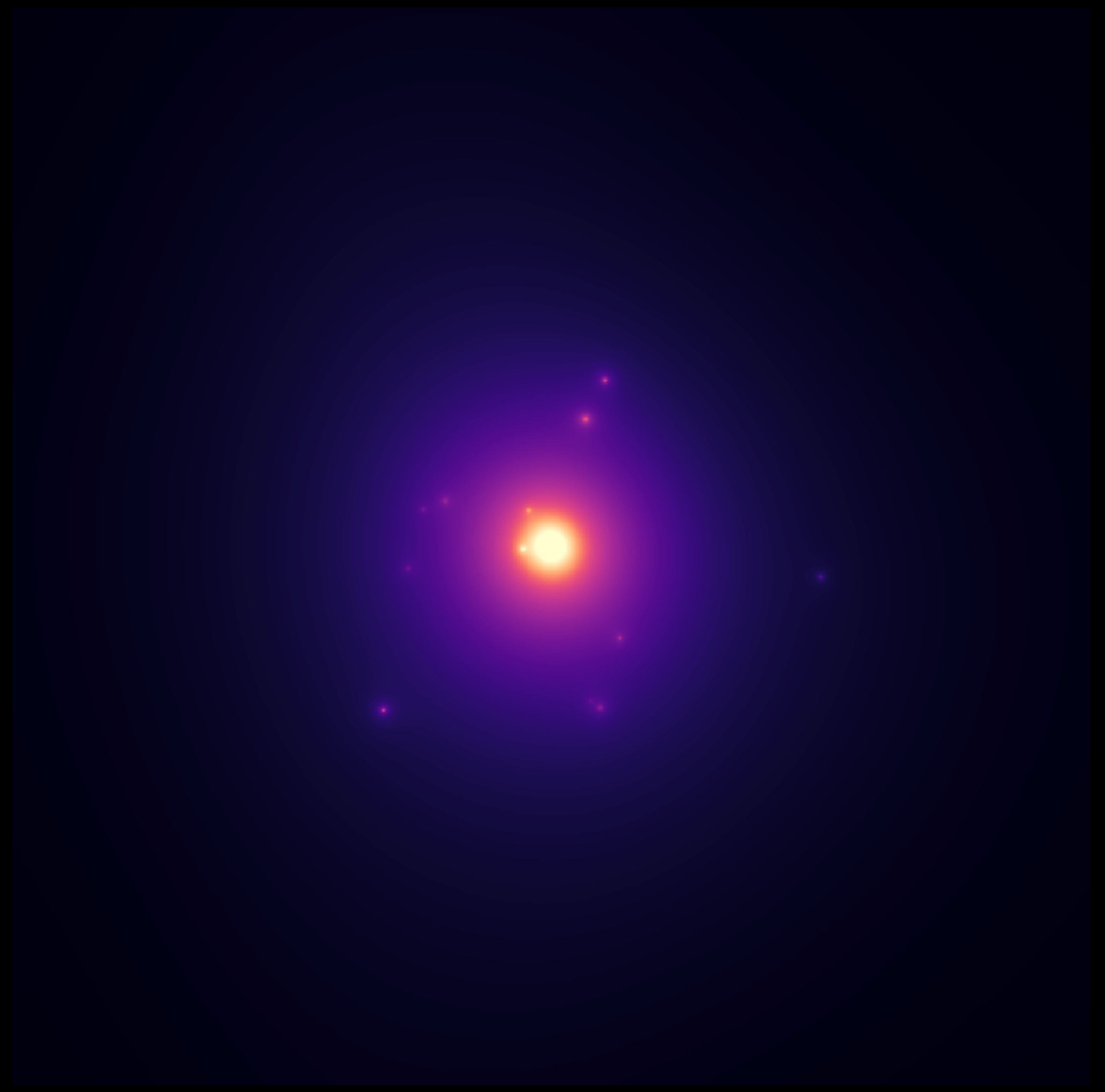


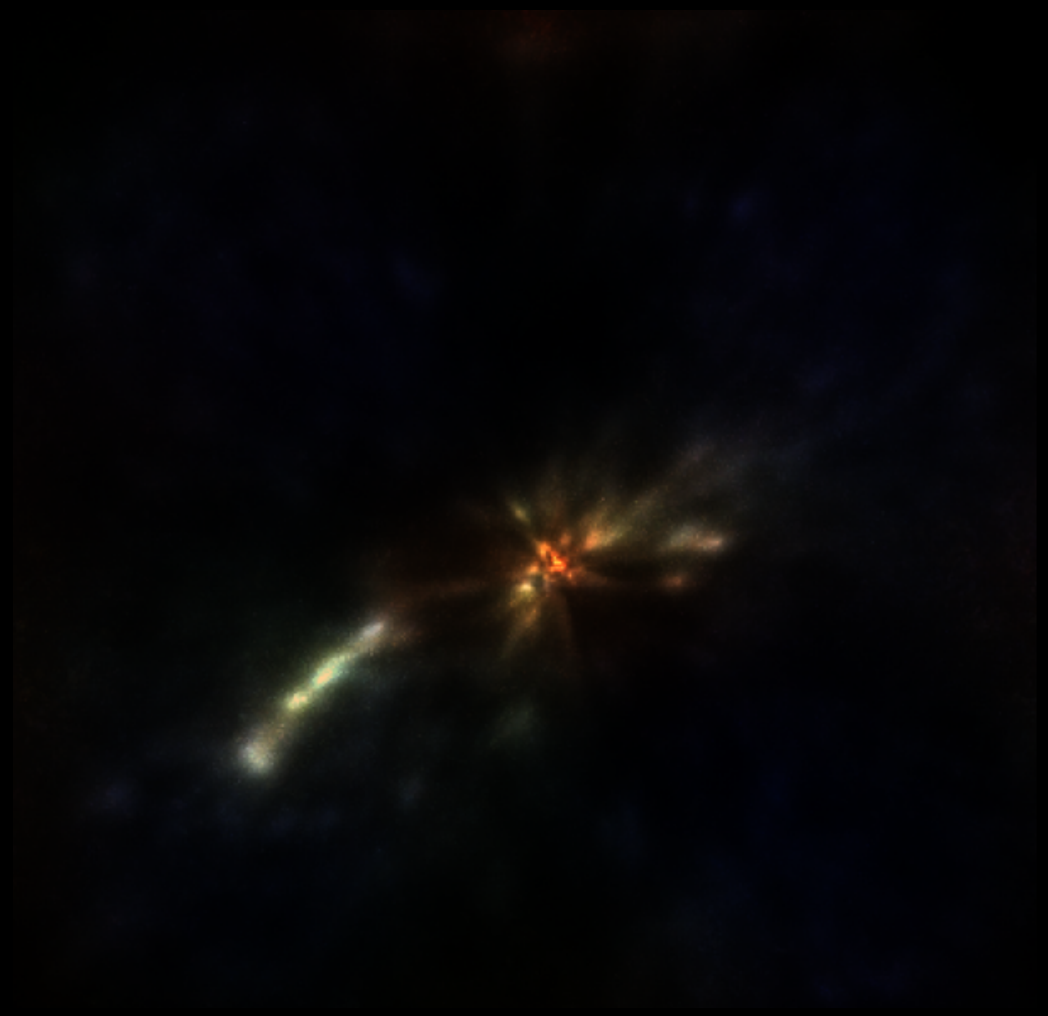


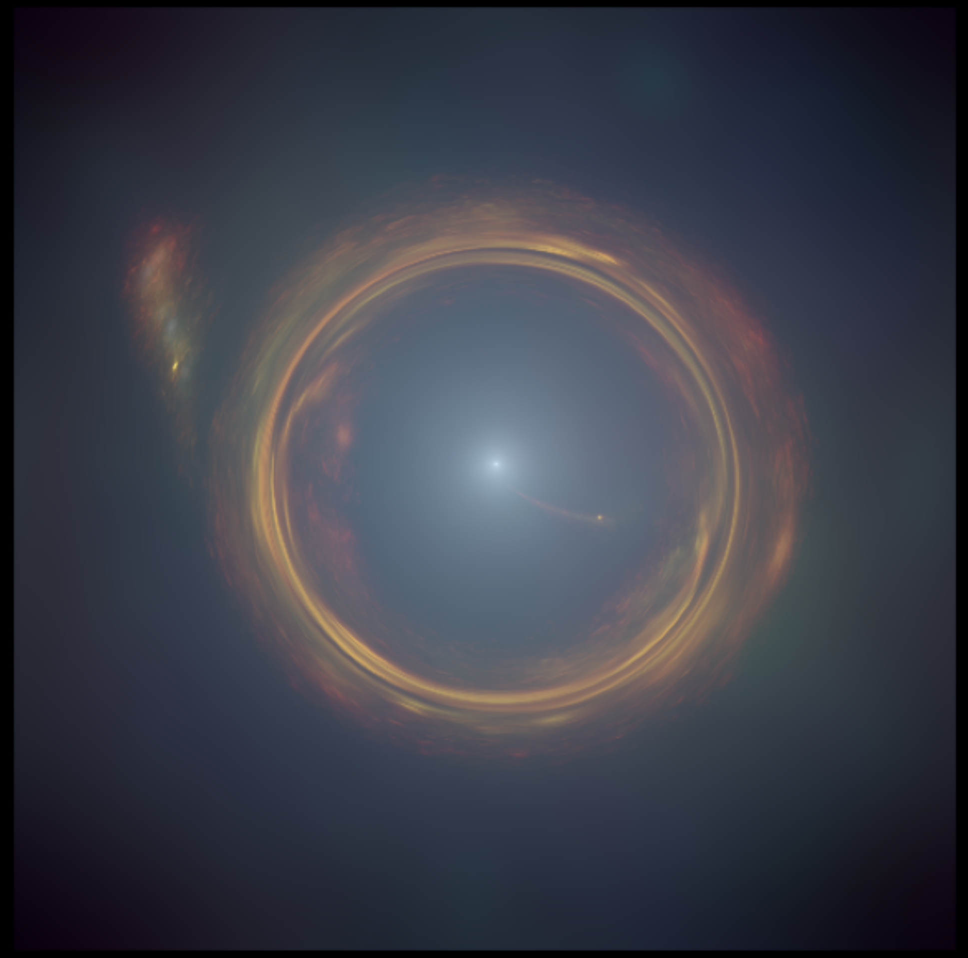


Take home

Model choice in high dimension is possible!



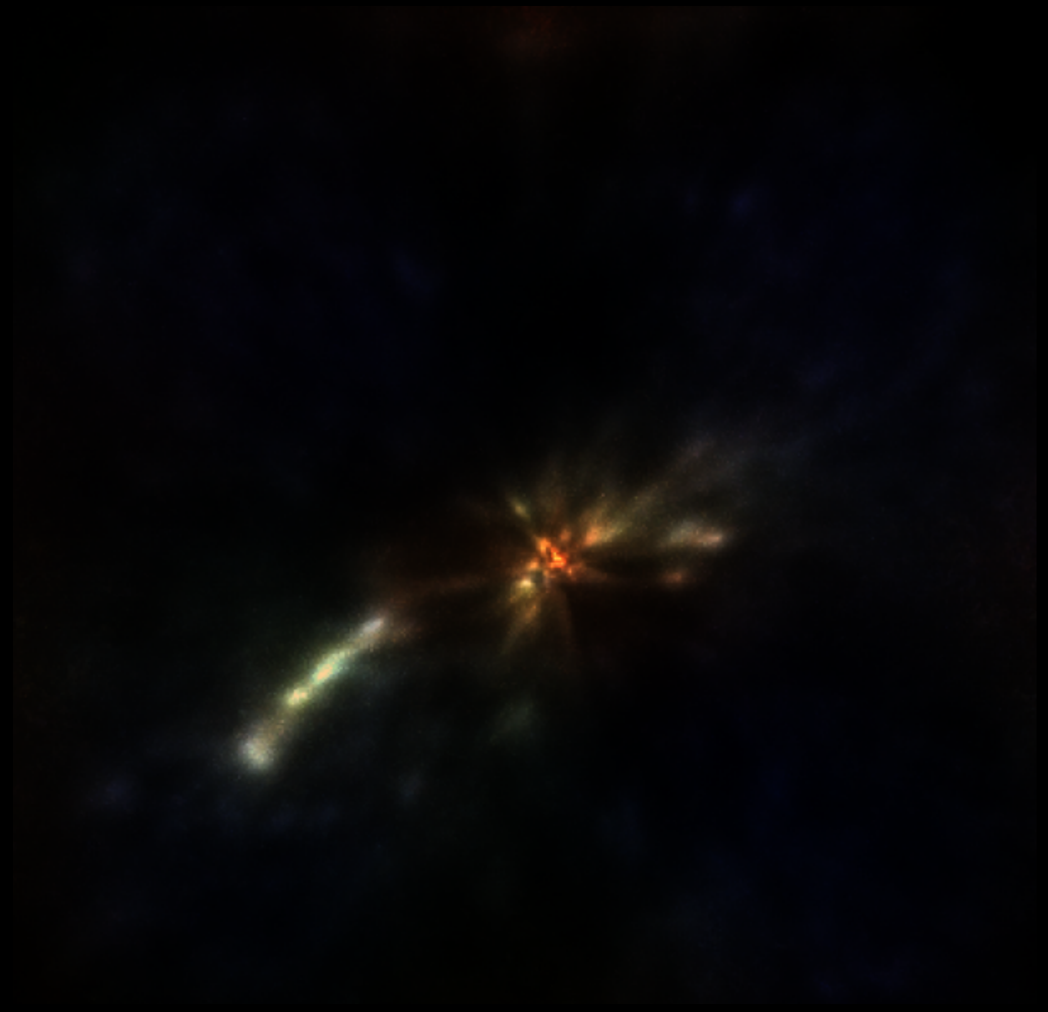


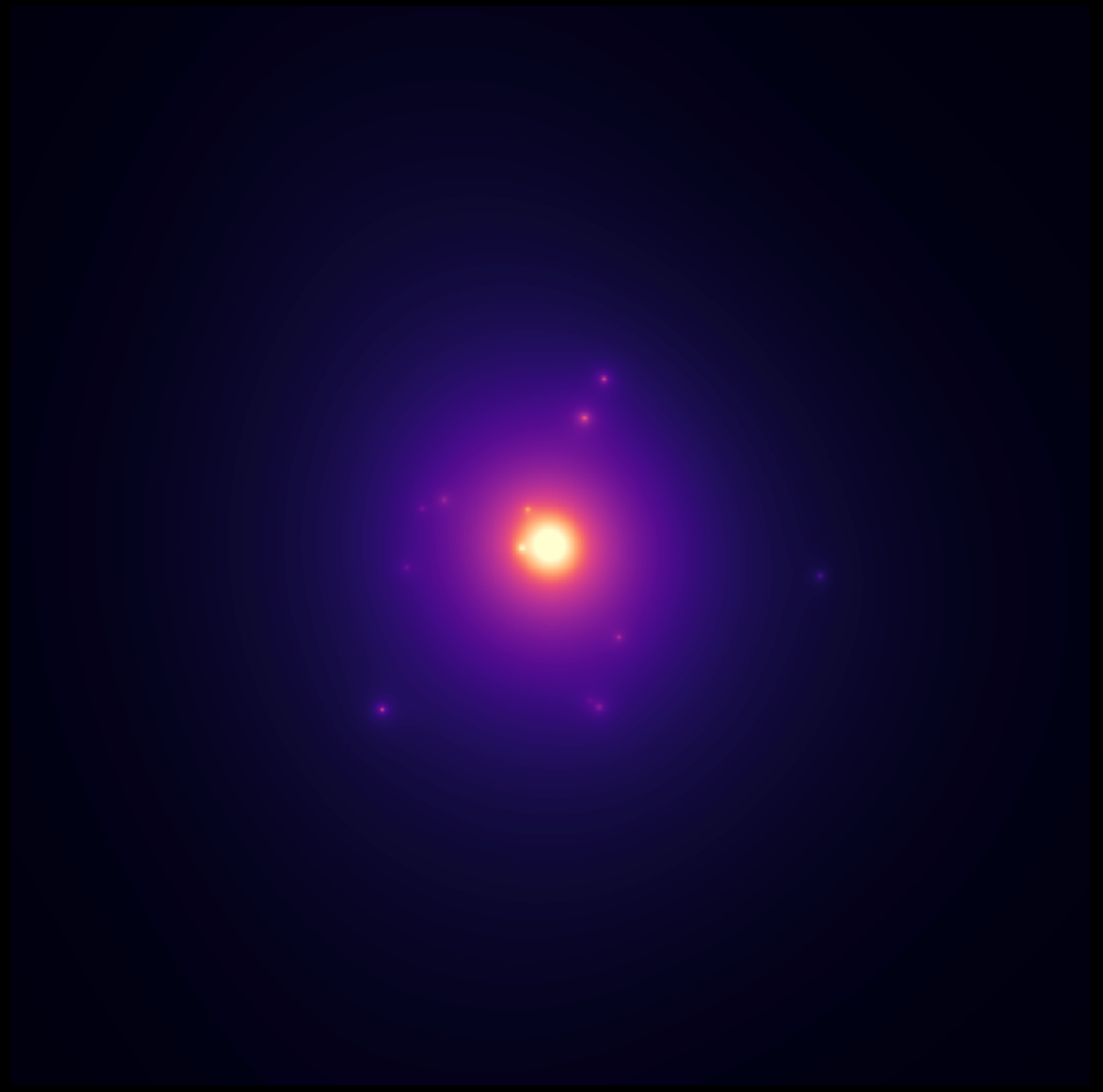


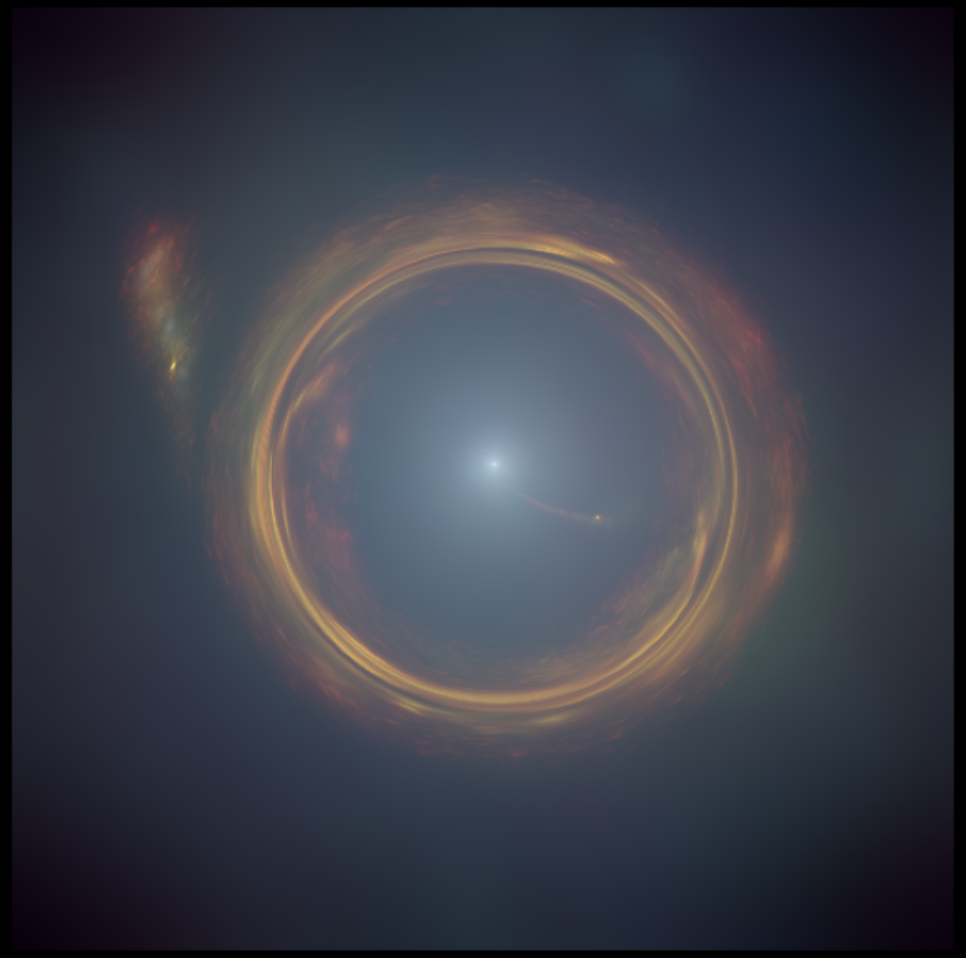
Take home

Model choice in high dimension is possible!

Applications:





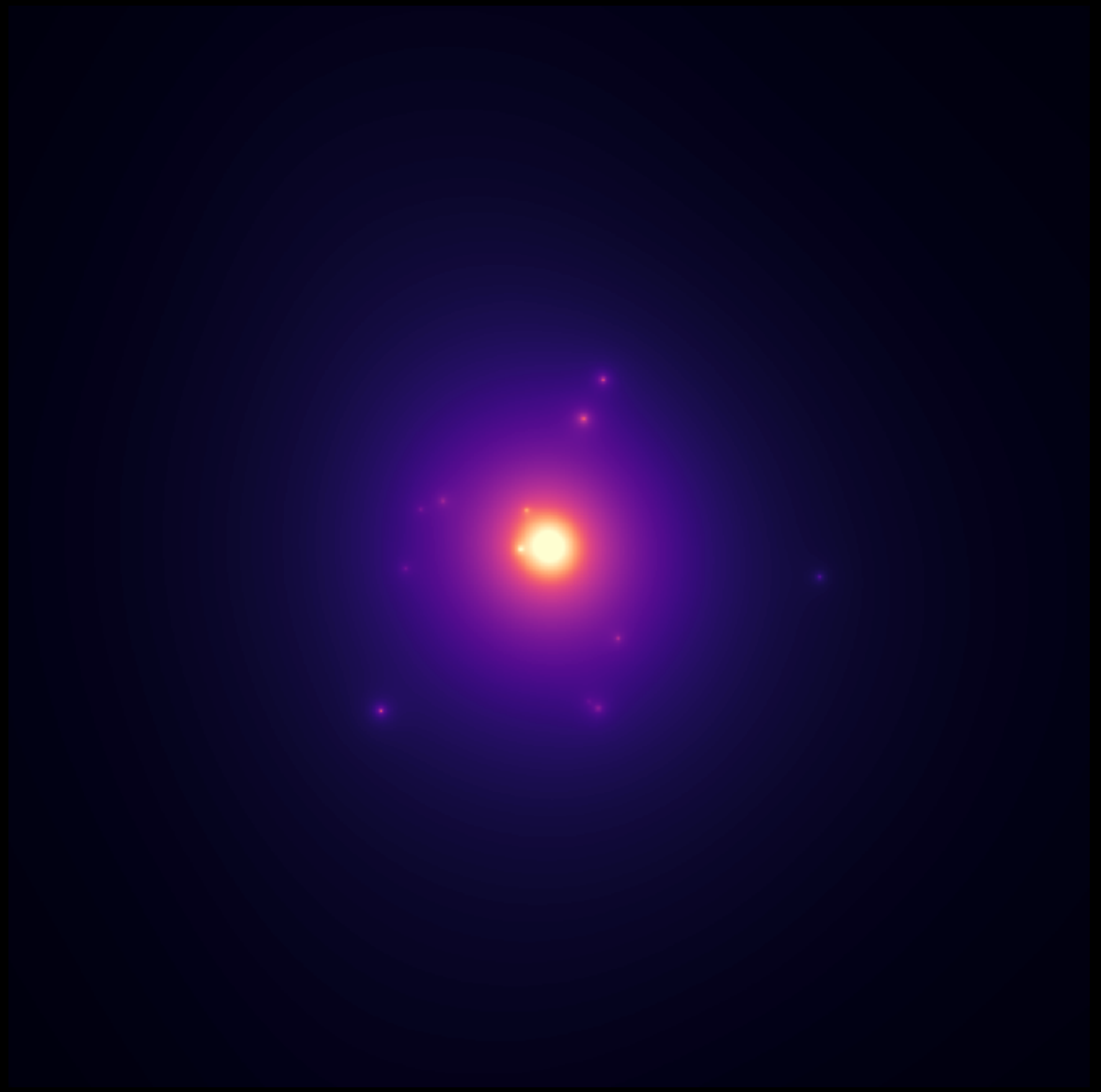


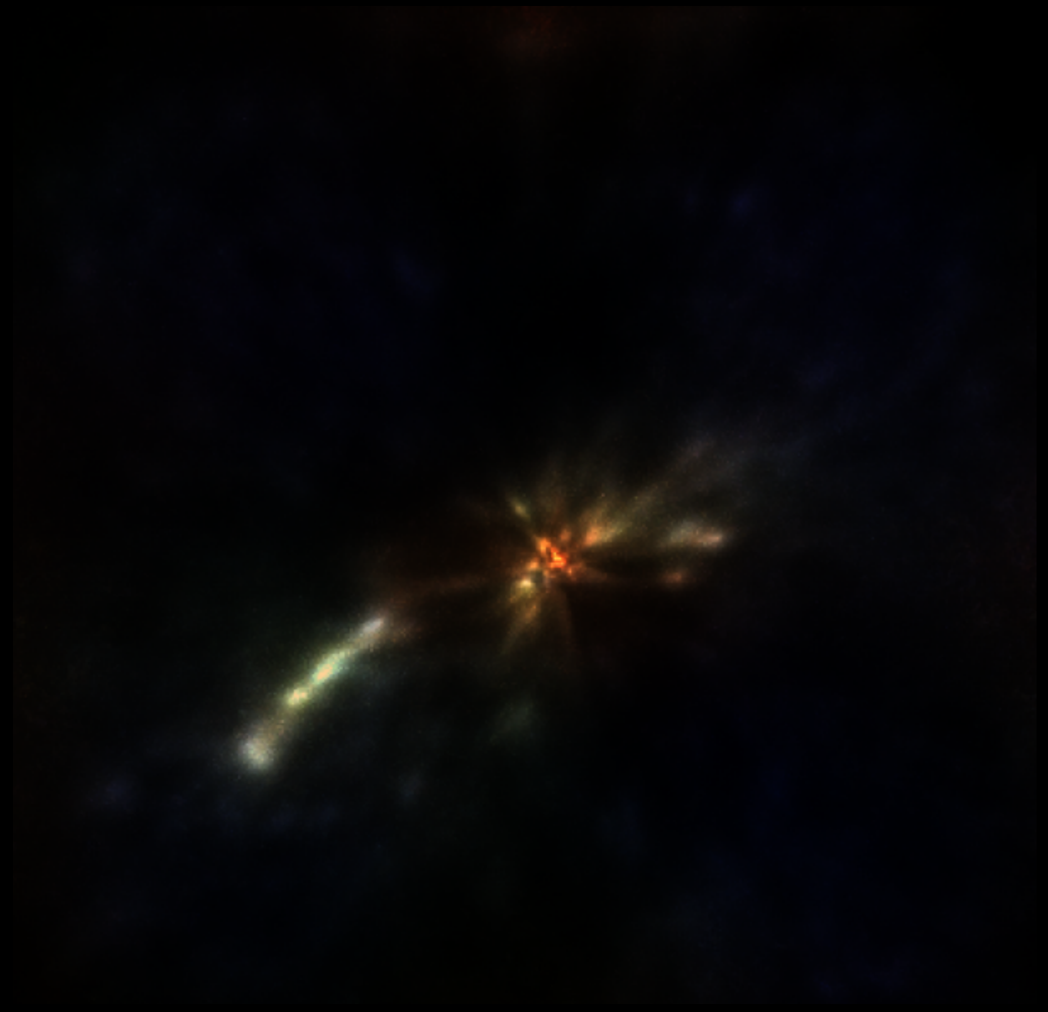
Take home

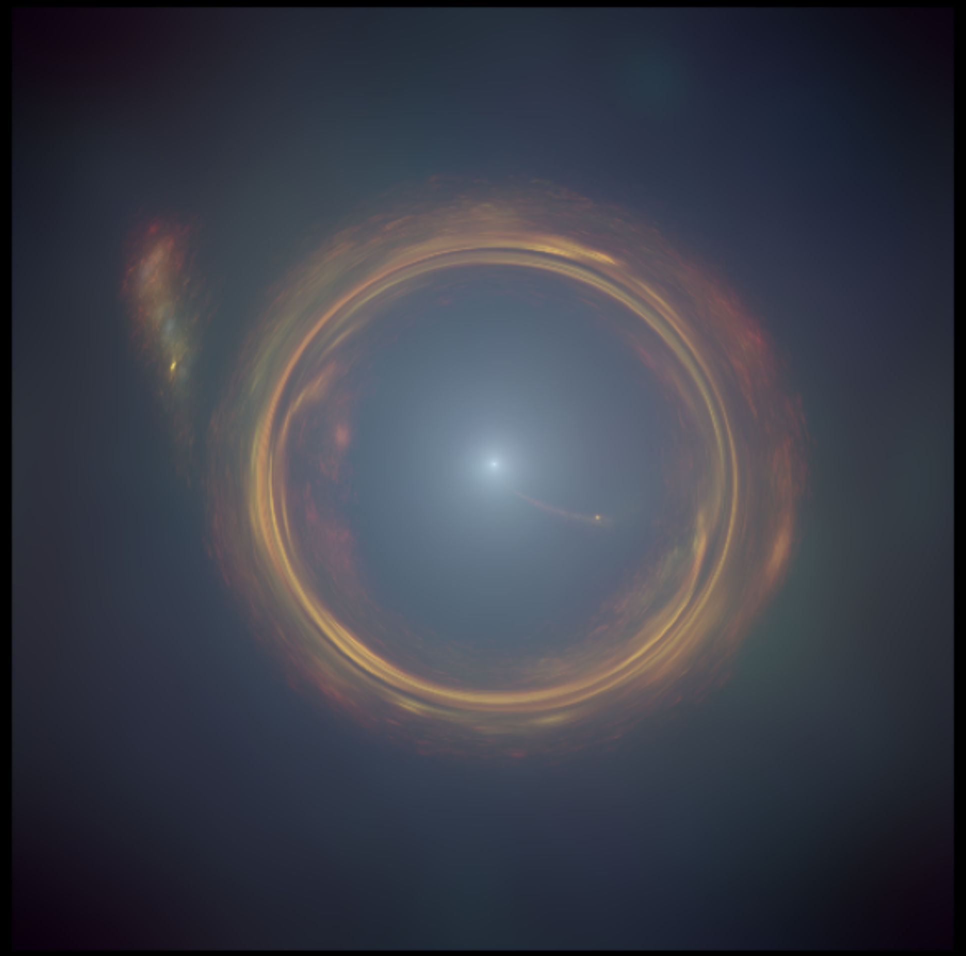
Model choice in high dimension is possible!

Applications:

- Astrophysics





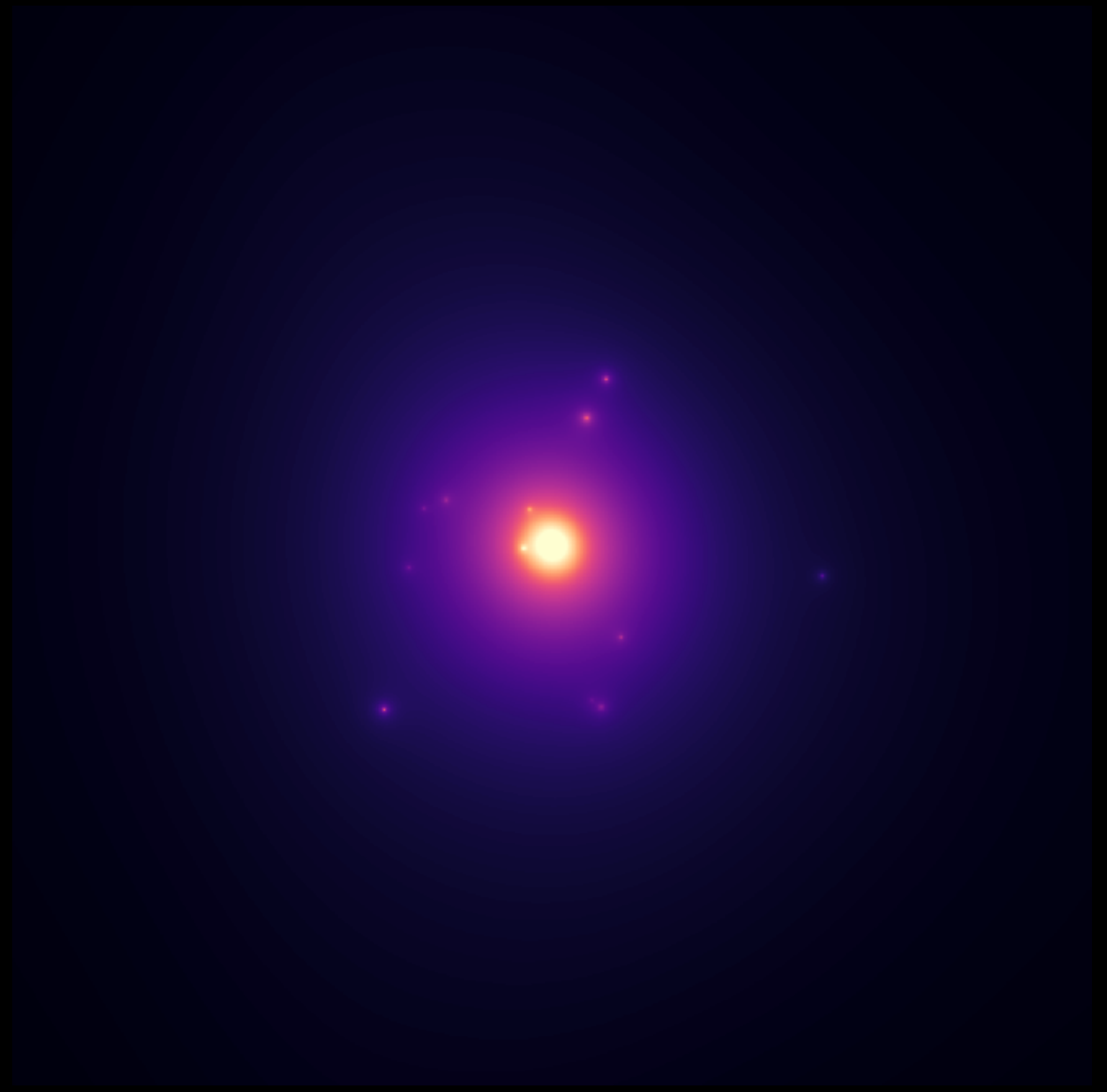


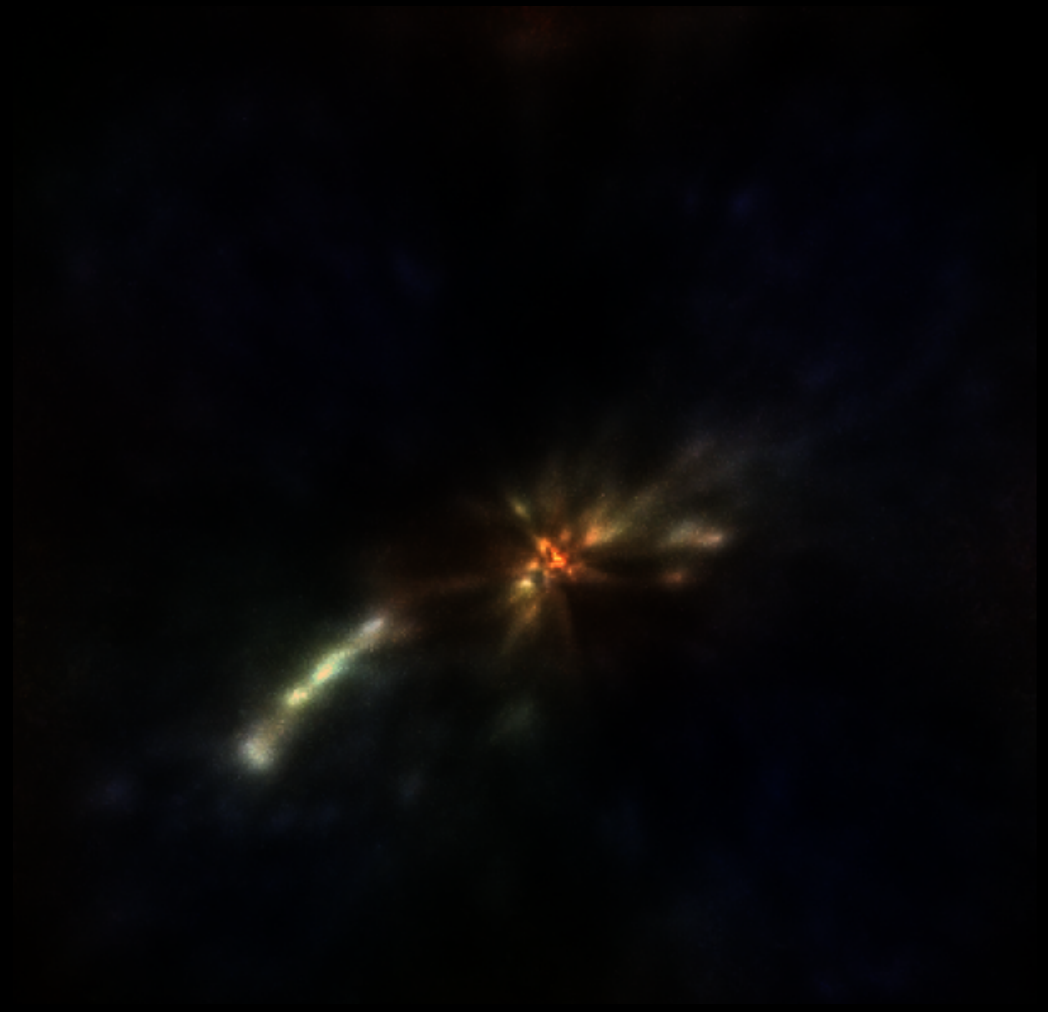
Take home

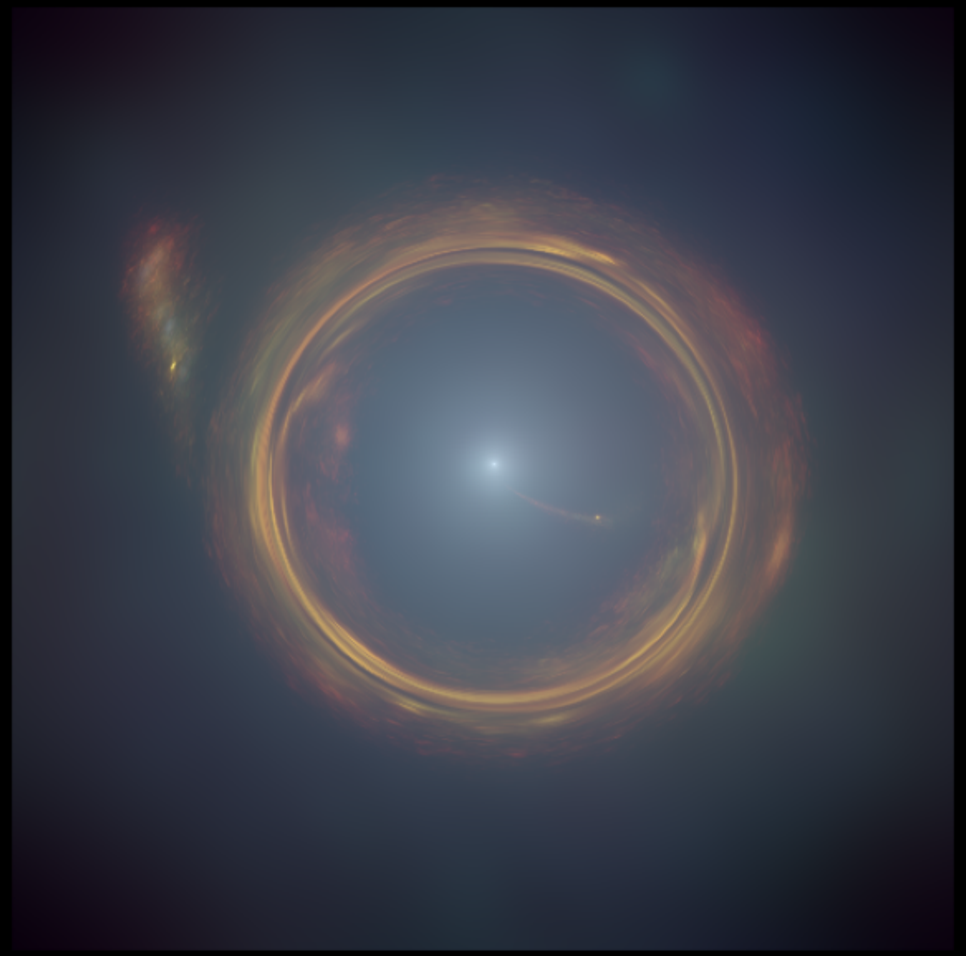
Model choice in high dimension is possible!

Applications:

- Astrophysics
- Causal inference





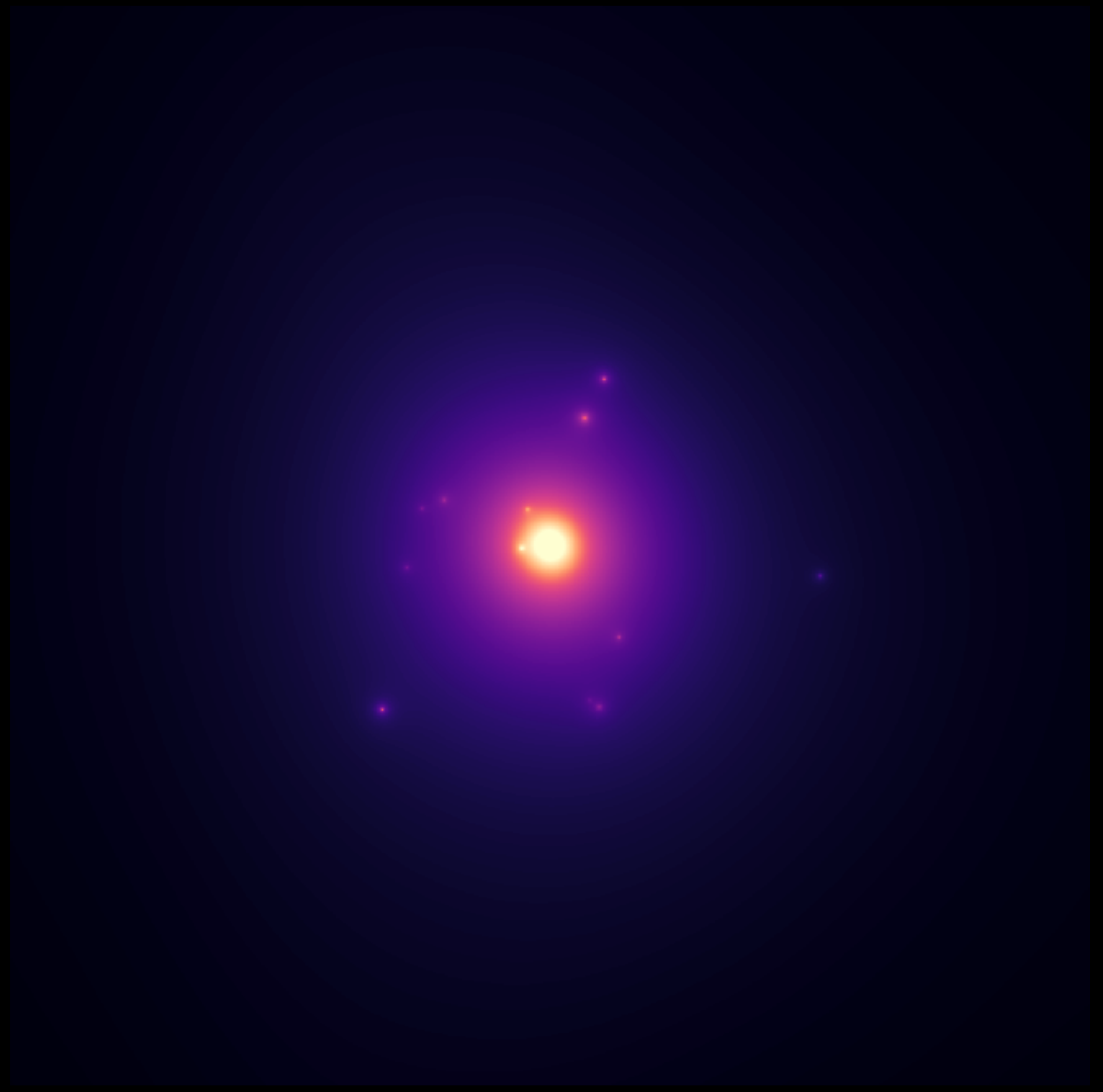


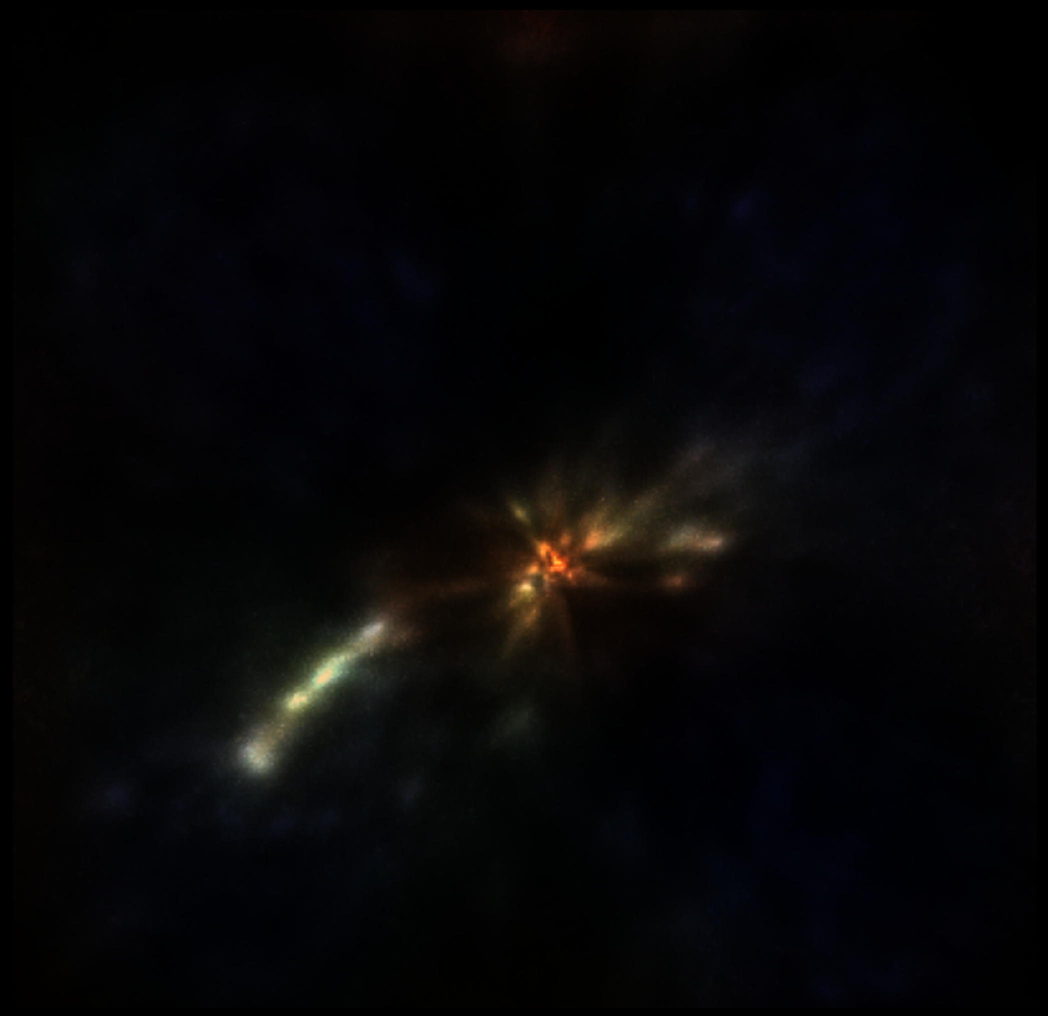
Take home

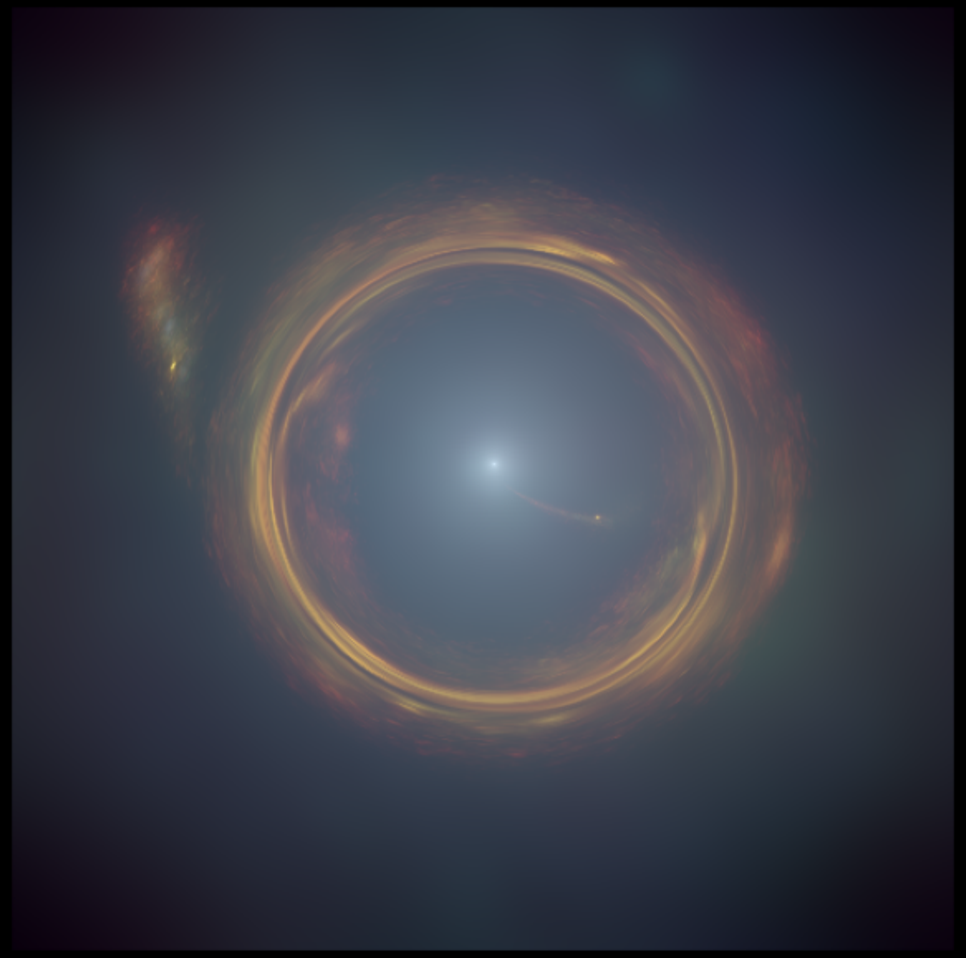
Model choice in high dimension is possible!

Applications:

- Astrophysics
- Causal inference
- …







Some advertisement…

- **NIFTy (Numerical Information Field Theory)**
  High-dimensional inference & Gaussian processes + ELBO

Some advertisement…

- **NIFTy (Numerical Information Field Theory)**
  High-dimensional inference & Gaussian processes + ELBO
- **J-UBIK (Universal Bayesian Imaging Kit)**
  Sky models & astronomical instrument responses

Some advertisement…

- **NIFTy (Numerical Information Field Theory)**
  High-dimensional inference & Gaussian processes + ELBO
- **J-UBIK (Universal Bayesian Imaging Kit)**
  Sky models & astronomical instrument responses
- **LensCharm**
  Charming strong-lensing signal reconstructions

Some advertisement…

- **NIFTy (Numerical Information Field Theory)**
  High-dimensional inference & Gaussian processes + ELBO
- **J-UBIK (Universal Bayesian Imaging Kit)**
  Sky models & astronomical instrument responses
- **LensCharm**
  Charming strong-lensing signal reconstructions

https://github.com/NIFTy-PPL/NIFTy
https://github.com/NIFTy-PPL/J-UBIK

Some advertisement…

- **NIFTy (Numerical Information Field Theory)**
  High-dimensional inference & Gaussian processes + ELBO
- **J-UBIK (Universal Bayesian Imaging Kit)**
  Sky models & astronomical instrument responses
- **LensCharm**
  Charming strong-lensing signal reconstructions

https://github.com/NIFTy-PPL/NIFTy
https://github.com/NIFTy-PPL/J-UBIK

GitLab https://gitlab.mpcdf.mpg.de/ift/lenscharm


Thank you!
matteani@mpa-garching.mpg.de
MAX PLANCK INSTITUTE
FOR ASTROPHYSICS
DLR
atlas