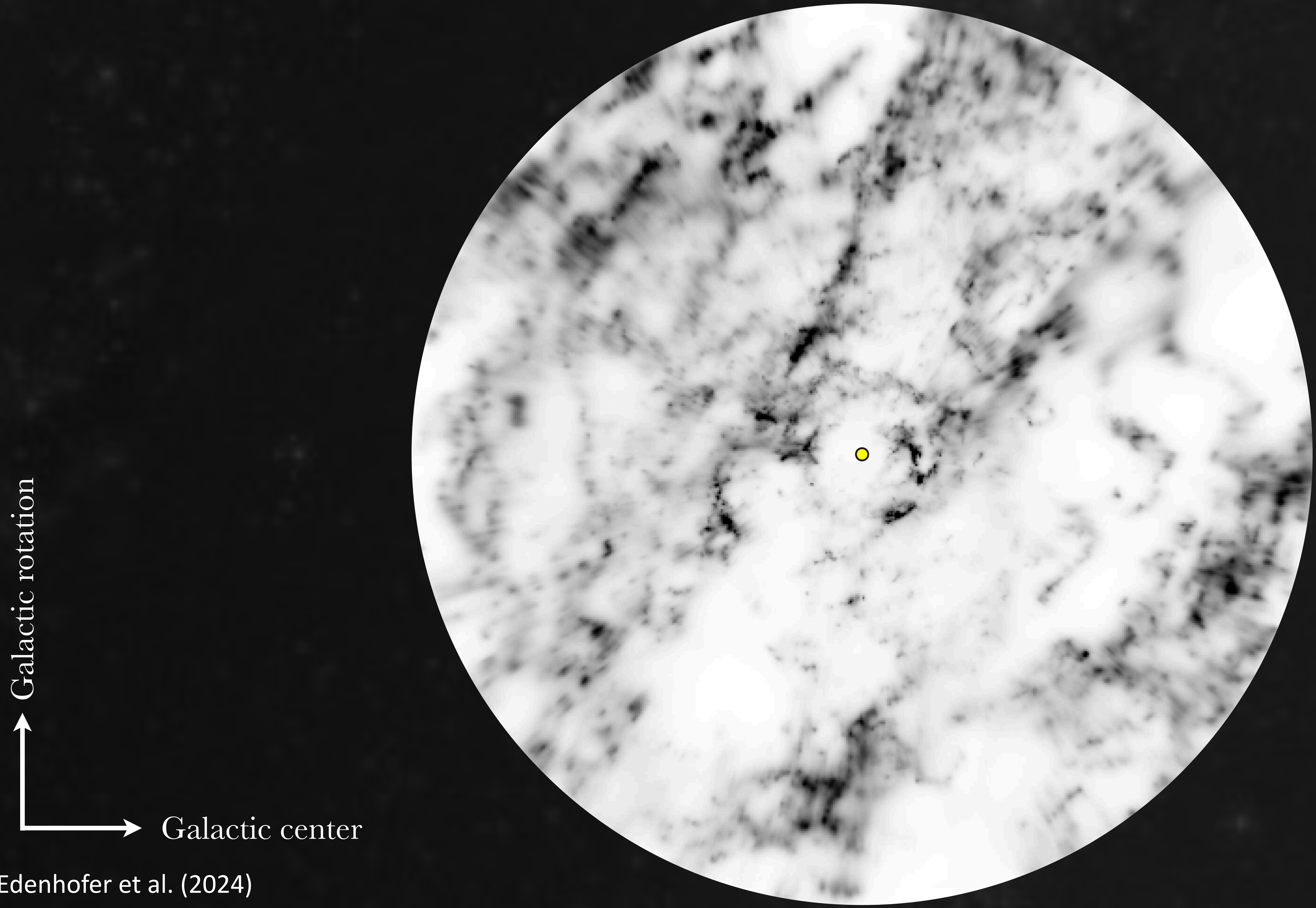
Learning with Gaps

A Domain-Adaptive SBI Framework for Mapping Young Stars from Incomplete, Multi-Survey Data

Sebastian Ratzenböck @CfA

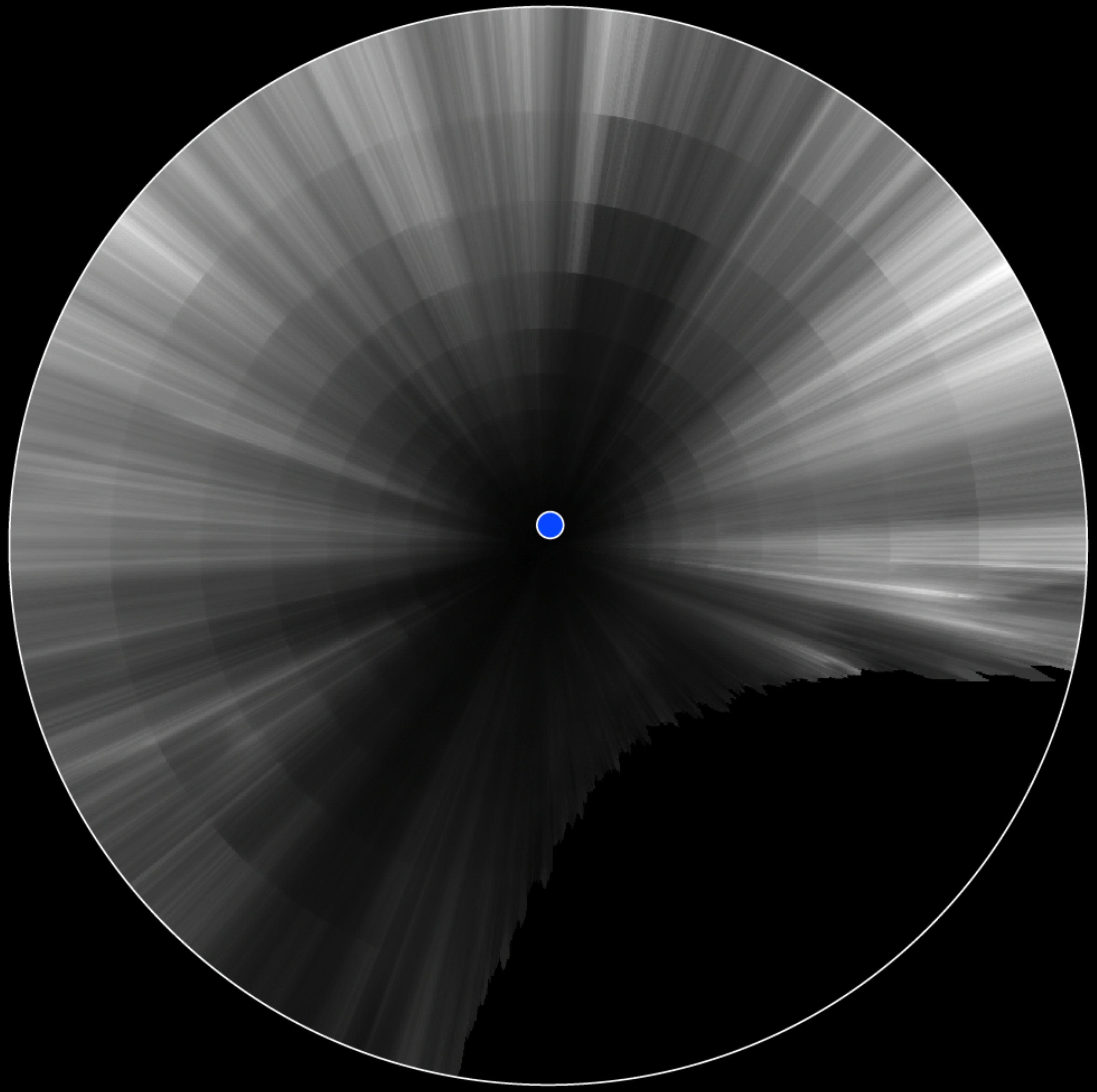
In collaboration with
Catherine Zucker, Joshua Speagle (UToronto), Phillip Cargile, Philipp Frank (Stanford), Andrew Saydjari (Princeton)

Edenhofer et al. (2024)



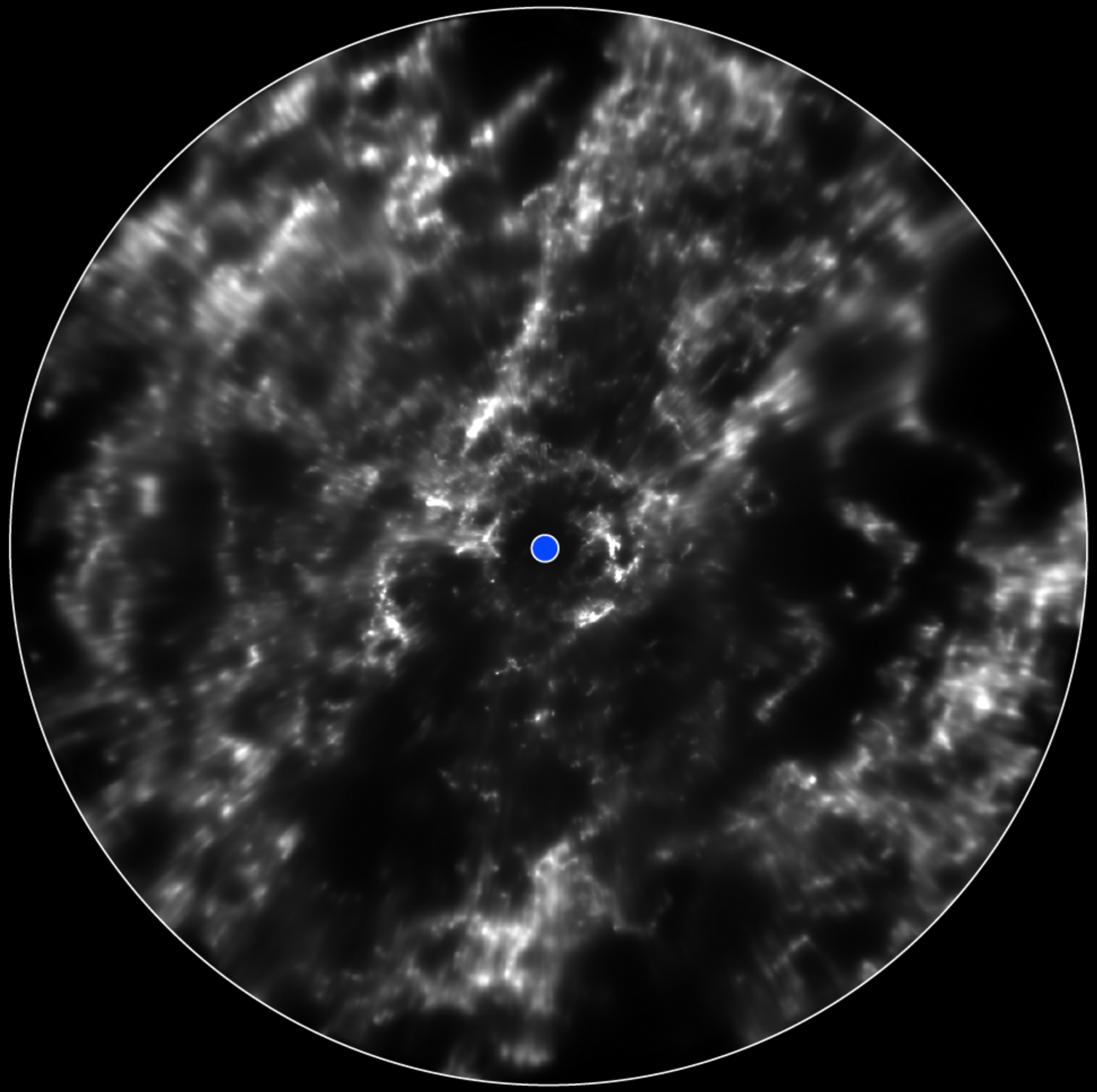
Edenhofer et al. (2024)

Solar Neighborhood **BEFORE**



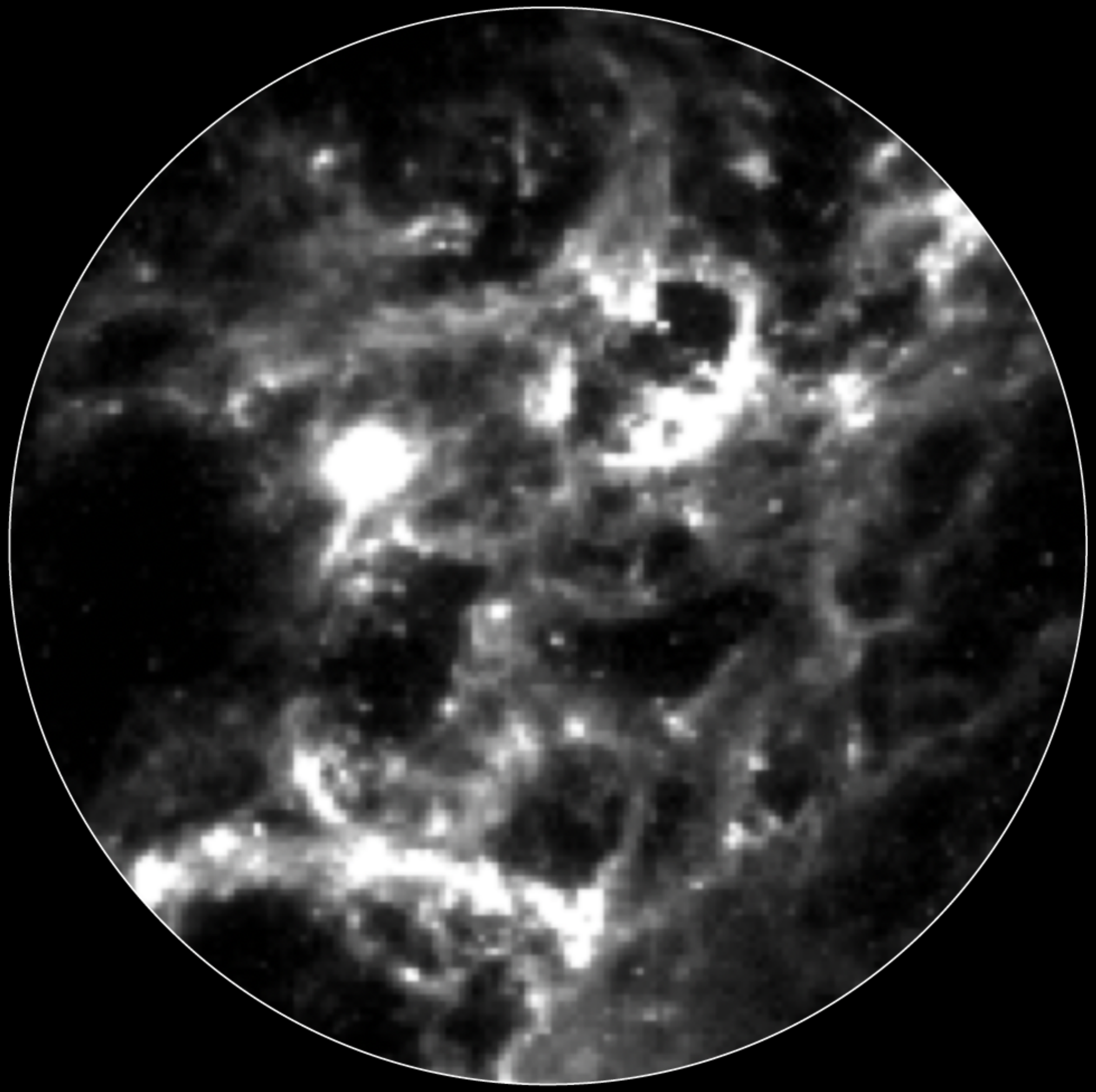
Green et al. 2018

Solar Neighborhood **NOW**



Edenhofer et al. (2024)

External Galaxy Zoom-in **NOW**



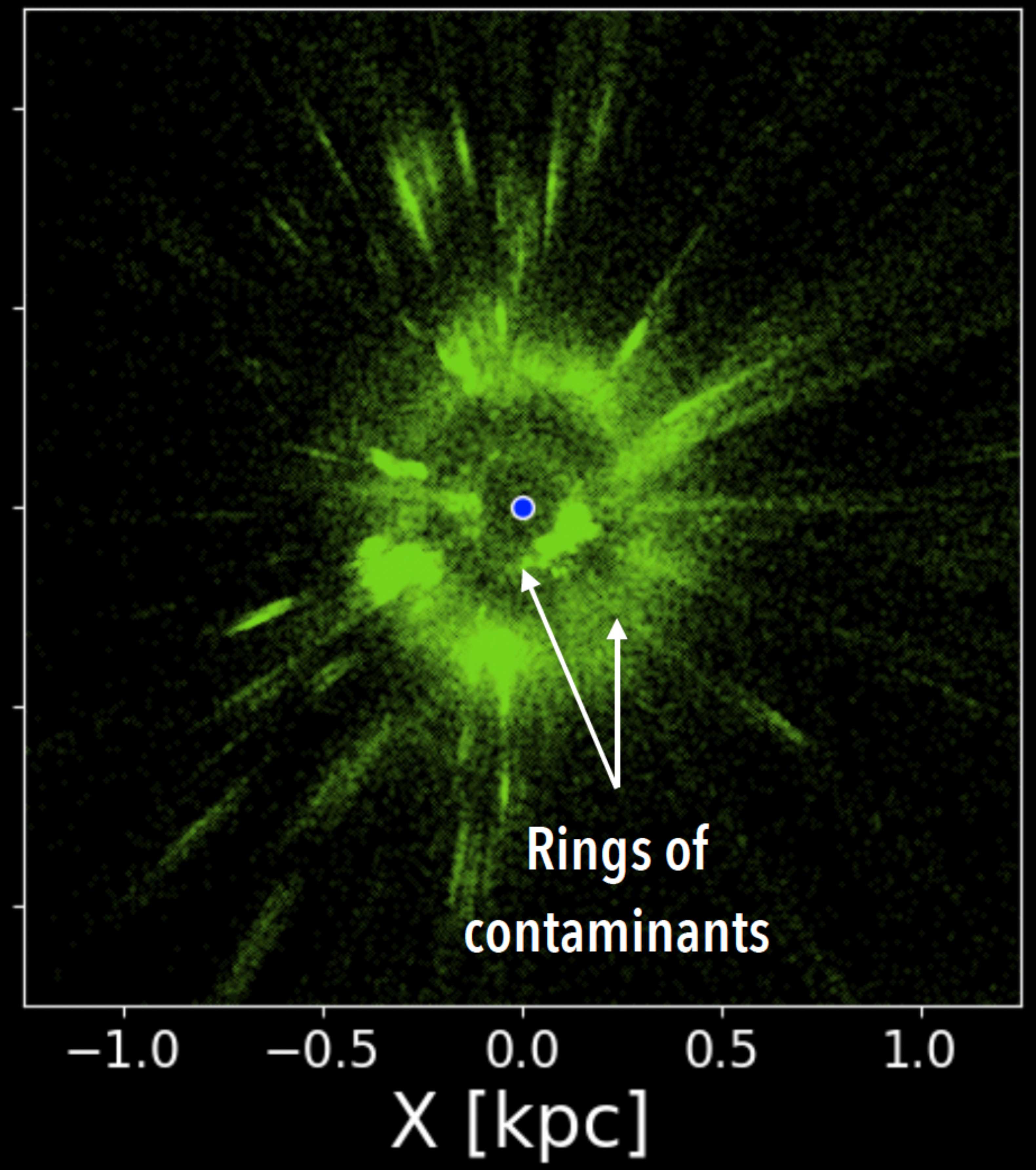
JWST 7μm zoom in

3D Dust
Young Star (SOTA Machine Learning)
Young Star (Literature Compilation)
Kounkel et al. (2023)
Rings of contaminants
Sharp boundary
Disk of contaminants
Y [kpc]
X [kpc]
X [kpc]
X [kpc]
1.0
0.5
0.0
−0.5
−1.0

Understanding Galactic baryon cycle

- YSOs connect *cloud* ↔ *stars* ↔ *feedback*

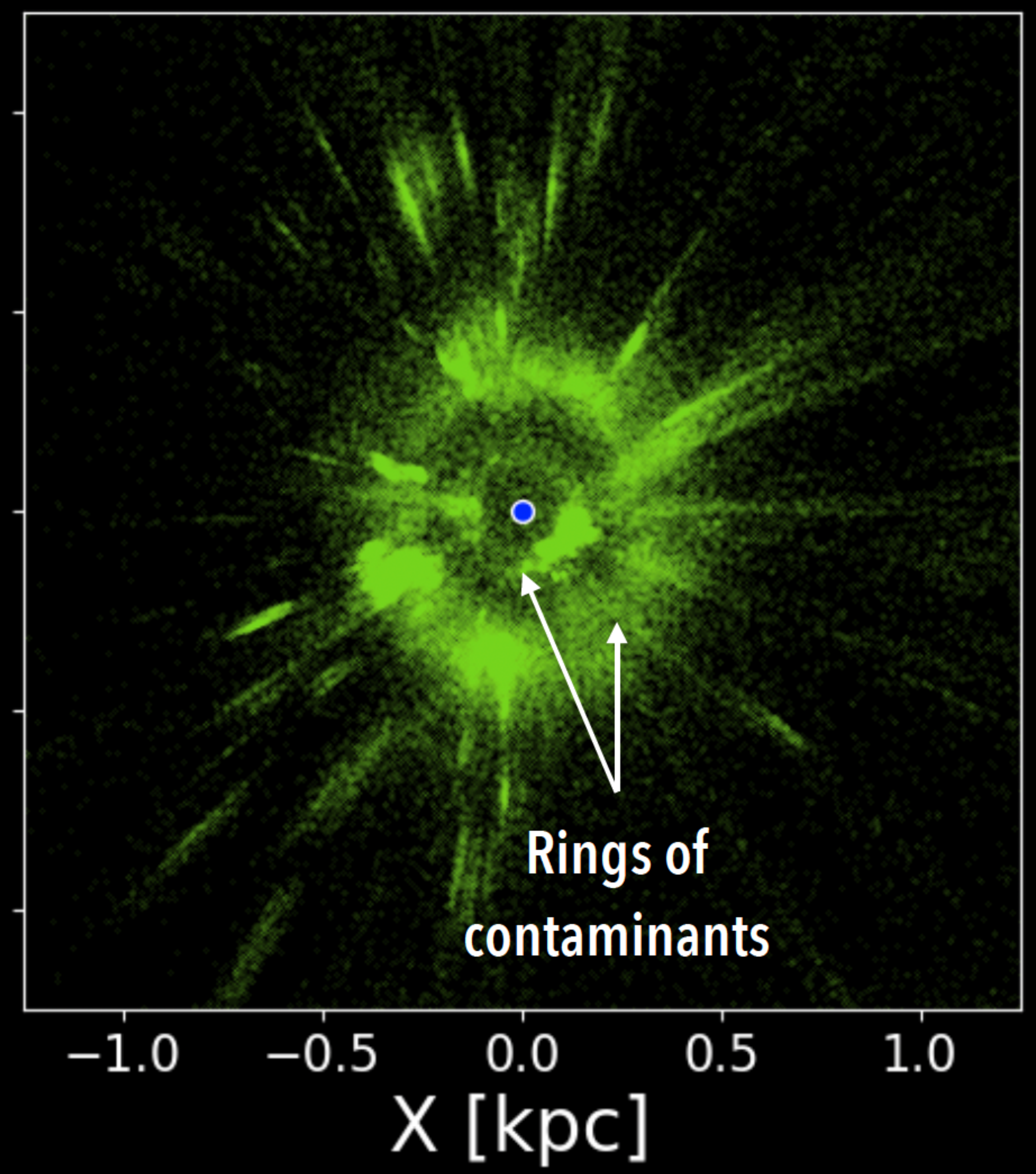
Young Star (SOTA Machine Learning)



Understanding Galactic baryon cycle

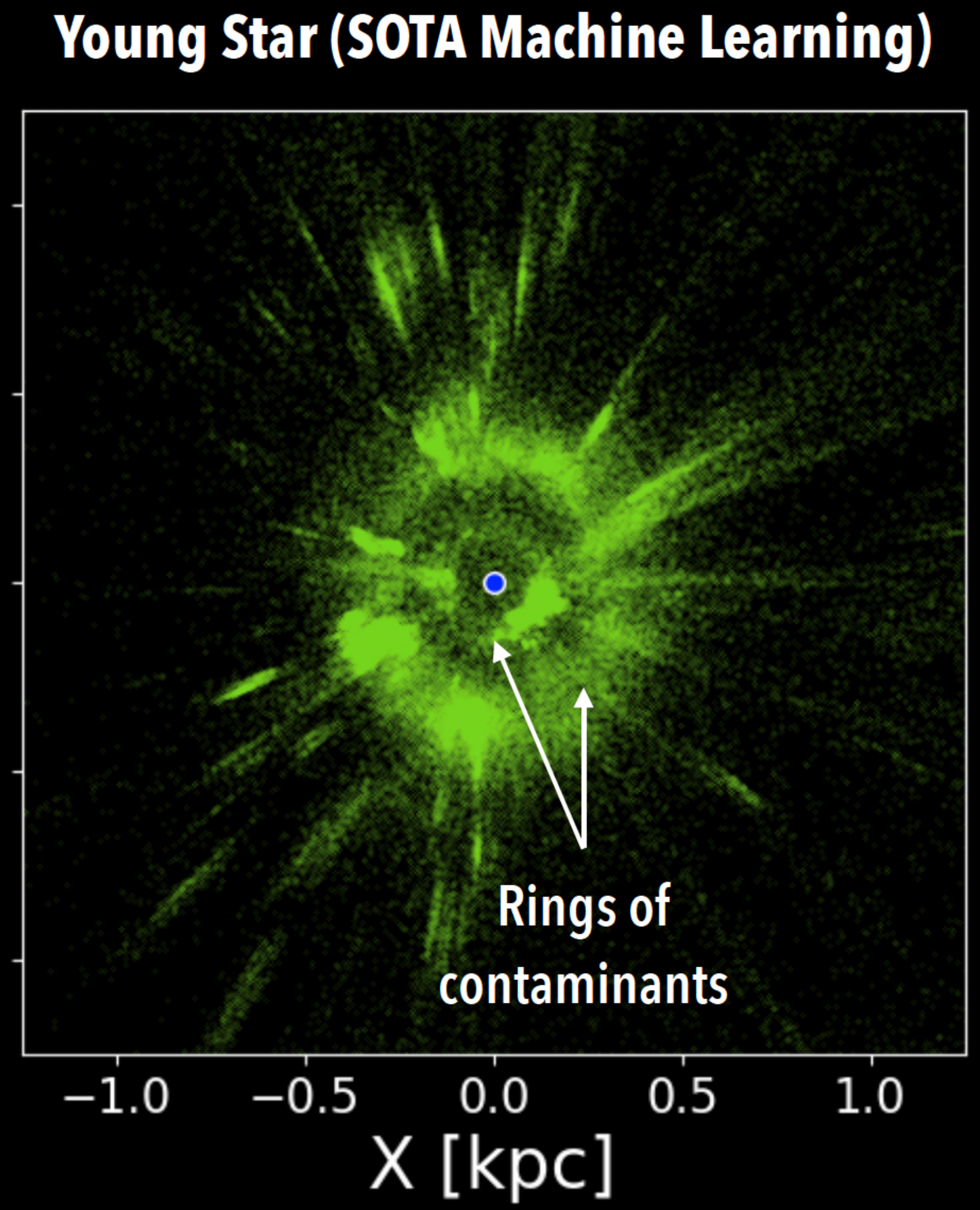
- YSOs connect *cloud* ↔ *stars* ↔ *feedback*
- How does the Milky Way convert gas into stars?

Young Star (SOTA Machine Learning)



Understanding Galactic baryon cycle

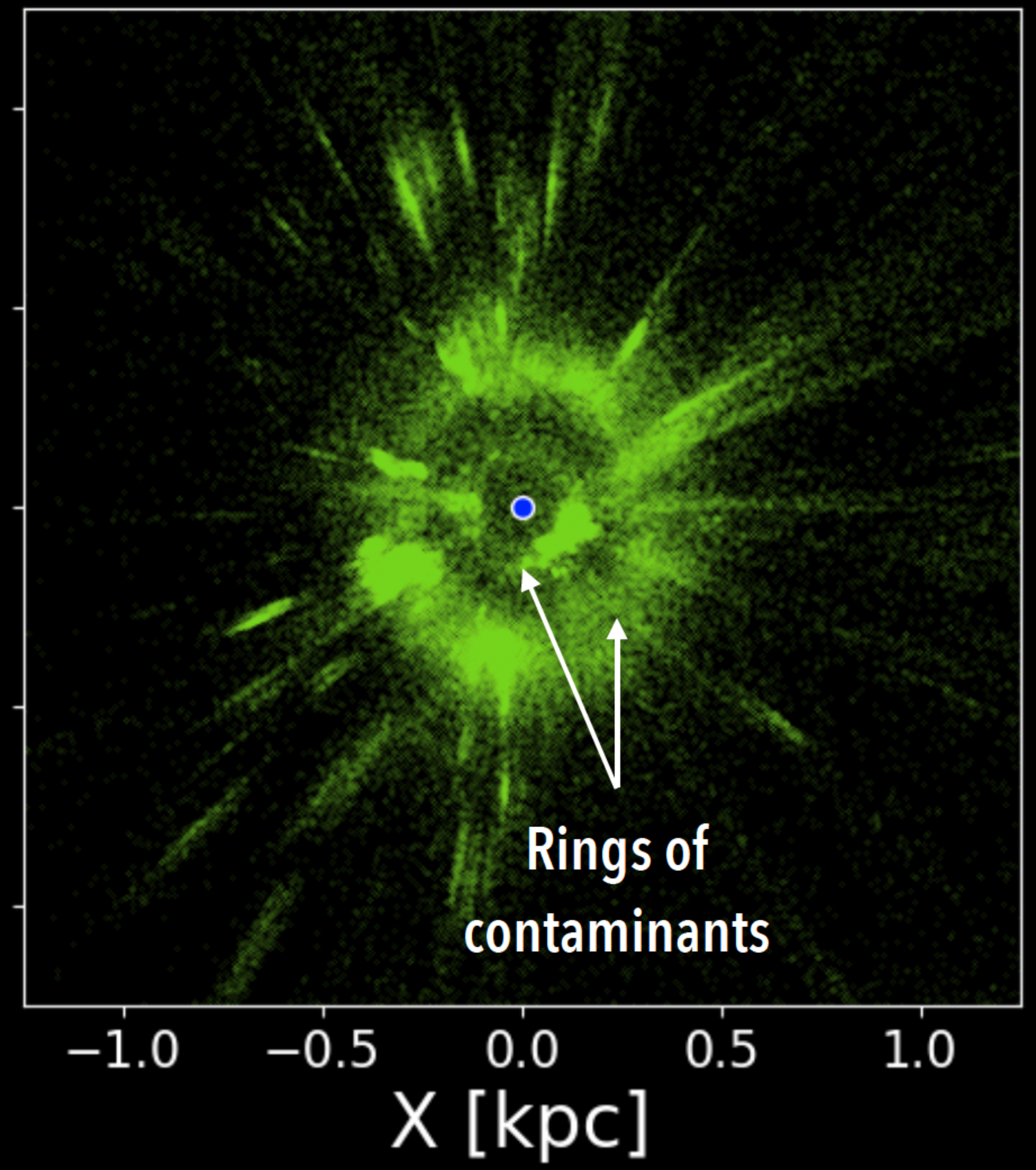
- YSOs connect *cloud* ↔ *stars* ↔ *feedback*
- How does the Milky Way convert gas into stars?
- How do stars leave their birth clouds and shape the ISM?



Understanding Galactic baryon cycle

- YSOs connect *cloud* ↔ *stars* ↔ *feedback*
- How does the Milky Way convert gas into stars?
- How do stars leave their birth clouds and shape the ISM?
- How do supernovae regulate, trigger, or suppress new generations of stars?

Young Star (SOTA Machine Learning)



Untapped potential of current methods

- Dedicated YSO models not employed

Untapped potential of current methods

- Dedicated YSO models not employed
- SED coverage with Gaia+2MASS+WISE photometry only

Untapped potential of current methods

- Dedicated YSO models not employed
- SED coverage with Gaia+2MASS+WISE photometry only
- Simple binary NN classifier and age regressor trained on mock data — no probabilistic approach

→ **apply SBI**

Challenges with “1 model does it all” approach

- Fusing surveys is hard due to different
  - resolutions & depths
  - coverage
  - instrument response
  - noise model

Challenges with “1 model does it all” approach

- Fusing surveys is hard due to different
  - resolutions & depths
  - coverage
  - instrument response
  - noise model
- Domain shift between simulation and real data

Challenges with “1 model does it all” approach

- Fusing surveys is hard due to different
  - resolutions & depths
  - coverage
  - instrument response
  - noise model
- Domain shift between simulation and real data

→ ***Domain-Adaptive*** *SBI w/* ***incomplete, multi-survey*** *data*

Model implementation
I. SBI model

Simformer: learning with incomplete data

θ

$$(X, \theta) \rightarrow \hat{X}$$

X

Shape $\quad \hat{X} \in (N, d_x + D_\theta = D_{\hat{X}})$

Inputs: θ, X

Gloeckler et al. (2024)

Simformer: learning with incomplete data

$\hat{X} \longrightarrow$ Tokenizer: $\hat{T}$

Value emb

ID emb

Conditional state emb

Shape $\hat{X} \in (N, D_{\hat{X}}) \rightarrow \hat{T} \in (N, D_{\hat{X}}, E_V + E_{\mathrm{id}} + E_C)$

Inputs: $\theta, X, M_C \in \{0,1\}^{D_X}$

Gloeckler et al. (2024)

Simformer: learning with incomplete data

$\hat{X} \longrightarrow$ Tokenizer: $\hat{T} \longrightarrow$ Transformer

$Q = \hat{T}W_Q \in (N, D_{\hat{X}}, d_v)$

Shape $\hat{T} \in (N, D_{\hat{X}}, E) \rightarrow S \propto QK^T \in (N, D_{\hat{X}}, D_{\hat{X}})$

Inputs: $\theta, X, M_C \in \{0,1\}^{D_X}$

ϕ : $W_Q, W_K, W_V \in (E, d_v)$

Gloeckler et al. (2024)

Simformer: learning with incomplete data

Attention masking: M_E

$\downarrow$

$\hat{X} \longrightarrow$ Tokenizer: $\hat{T} \longrightarrow$ Transformer

Shape $\quad S \to S + M_E$

Inputs: $\theta, X, M_C, M_E \in (N, D_{\hat{X}}, D_{\hat{X}})$

$\phi: \quad W_Q, W_K, W_V$

Gloeckler et al. (2024)

Simformer: learning with incomplete data

Attention masking: M_E

$\hat{X} \longrightarrow$ Tokenizer: $\hat{T} \longrightarrow$ Transformer $\longrightarrow$ Score (Diffusion) $\in (N, D_{\hat{X}})$

$V = \hat{T}W_v \in (D_{\hat{X}}, d_v)$

Shape $\quad s_\phi^{M_E}(\hat{x}^{M_C}) \mapsto \mathrm{softmax}(S + M_E)\, V\, P$

Inputs: θ, X, M_C, M_E

ϕ : $\quad W_Q, W_K, W_V, P \in (d_v, 1)$

Gloeckler et al. (2024)

Model implementation

II. Learning to condition and marginalize

Training

Given Score model: $s_\phi^{M_E}(\hat{x}^{M_C})$

Gloeckler et al. (2024)

Training

Given Score model: $s_{\phi}^{M_E}(\hat{x}^{M_C})$

1. Add noise acc. to noise schedule $\hat{x}_t^{M_C} = (1 - M_C)\,\hat{x}_t + M_C\,\hat{x}_0$

$$\hat{x}_t \sim p_t(\hat{x}_t \,|\, \hat{x}_0) = \mathcal{N}(\mu_t(\hat{x}_0), \sigma_t(\hat{x}_0))$$

Gloeckler et al. (2024)

Training

Given Score model: $s_\phi^{M_E}(\hat{x}^{M_C})$

1. Add noise acc. to noise schedule $\hat{x}_t^{M_C} = (1 - M_C)\,\hat{x}_t + M_C\,\hat{x}_0$

2. Compute denoising loss

$$\ell(\phi, M_C, t, \hat{\mathbf{x}}_0, \hat{\mathbf{x}}_t) = (1 - M_C) \cdot \left(s_\phi^{M_E}(\hat{\mathbf{x}}_t^{M_C}, t) - \nabla_{\hat{\mathbf{x}}_t} \log p_t(\hat{\mathbf{x}}_t | \hat{\mathbf{x}}_0) \right)$$

Gloeckler et al. (2024)

Training

Given Score model: $s_\phi^{M_E}(\hat{x}^{M_C})$

1. Add noise acc. to noise schedule $\hat{x}_t^{M_C} = (1 - M_C)\,\hat{x}_t + M_C\,\hat{x}_0$

2. Compute denoising loss

$$\ell(\phi, M_C, t, \hat{\mathbf{x}}_0, \hat{\mathbf{x}}_t) = (1 - M_C) \cdot \left(s_\phi^{M_E}(\hat{\mathbf{x}}_t^{M_C}, t) - \nabla_{\hat{\mathbf{x}}_t} \log p_t(\hat{\mathbf{x}}_t | \hat{\mathbf{x}}_0)\right)$$

$\longrightarrow$ With M_C score approximates any conditional

Gloeckler et al. (2024)

Training

Given Score model: $s_\phi^{M_E}(\hat{x}^{M_C})$

1. Add noise acc. to noise schedule $\hat{x}_t^{M_C} = (1 - M_C)\,\hat{x}_t + M_C\,\hat{x}_0$

2. Compute denoising loss

$$\ell(\phi, M_C, t, \hat{\mathbf{x}}_0, \hat{\mathbf{x}}_t) = (1 - M_C) \cdot \Big(s_\phi^{M_E}(\hat{\mathbf{x}}_t^{M_C}, t) - \nabla_{\hat{\mathbf{x}}_t} \log p_t(\hat{\mathbf{x}}_t | \hat{\mathbf{x}}_0) \Big)$$

$\longrightarrow$ With M_C score approximates any conditional

$\longrightarrow$ With M_E score approximates any marginal

Gloeckler et al. (2024)

Model implementation
III. Domain adaption

Architecture

$$\begin{tikzcd} \theta \arrow[dr] & & & \\ & (X, \theta) \arrow[r] & \text{Transformer} \arrow[r] & \text{Flow matching} \\ X \arrow[ur] & & & \end{tikzcd}$$

Input split into *simulated*, *real* & *paired* data

$$\theta$$

$$X \longrightarrow \begin{matrix} X_{\mathrm{sim}} \\ X_{\mathrm{real}} \\ X_{\mathrm{sim-real-pairs}} \end{matrix}$$

$\longrightarrow$ Transformer $\longrightarrow$ Flow matching

Modality encoders

$$X \longrightarrow \begin{matrix} X_{\mathrm{s}} \\ X_{\mathrm{r}} \\ X_{\mathrm{s-r}} \end{matrix}$$

Modality encoders: split into indiv. spectra

X_s

Apogee

Boss

Lamost

XP

Photometry

Gaia, 2Mass, …

Modality encoders: encode

Encoder

$X_s \rightarrow$

Apogee $\longrightarrow Z_s^{A}$

Boss $\longrightarrow Z_s^{B}$

Lamost $\longrightarrow Z_s^{L}$

XP $\longrightarrow Z_s^{xp}$

Photometry

Gaia, 2Mass, … $\longrightarrow Z_s^{G}$

$\vdots$

Modality encoders: contrastive loss

Encoder

X_s

Apogee $\longrightarrow Z_s^{A}$

Boss $\longrightarrow Z_s^{B}$

Lamost $\longrightarrow Z_s^{L}$

XP $\longrightarrow Z_s^{xp}$

$\left.\begin{matrix} Z_s^{B} \\ Z_s^{L} \\ Z_s^{xp} \end{matrix}\right\} Z^{pool}$

Pool embedding (mean)

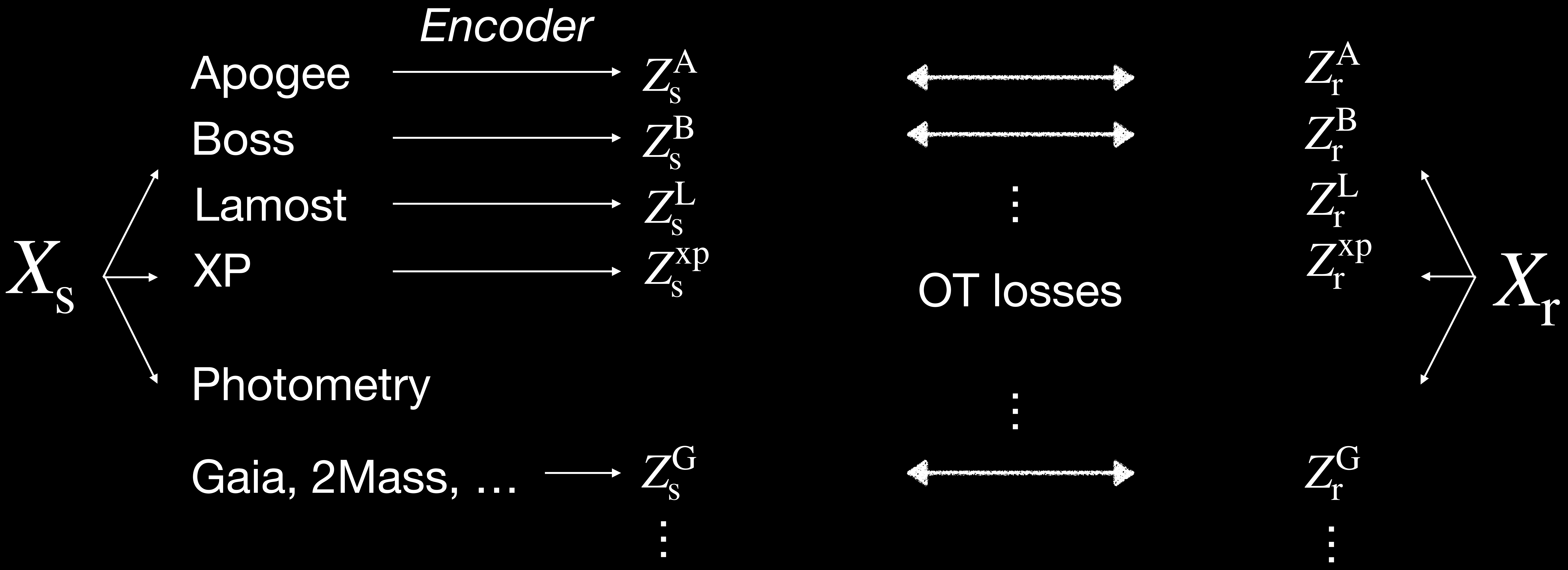
Contrastive loss: $Z^{\{xp,L,B\}}$ & Z^{pool}

Photometry

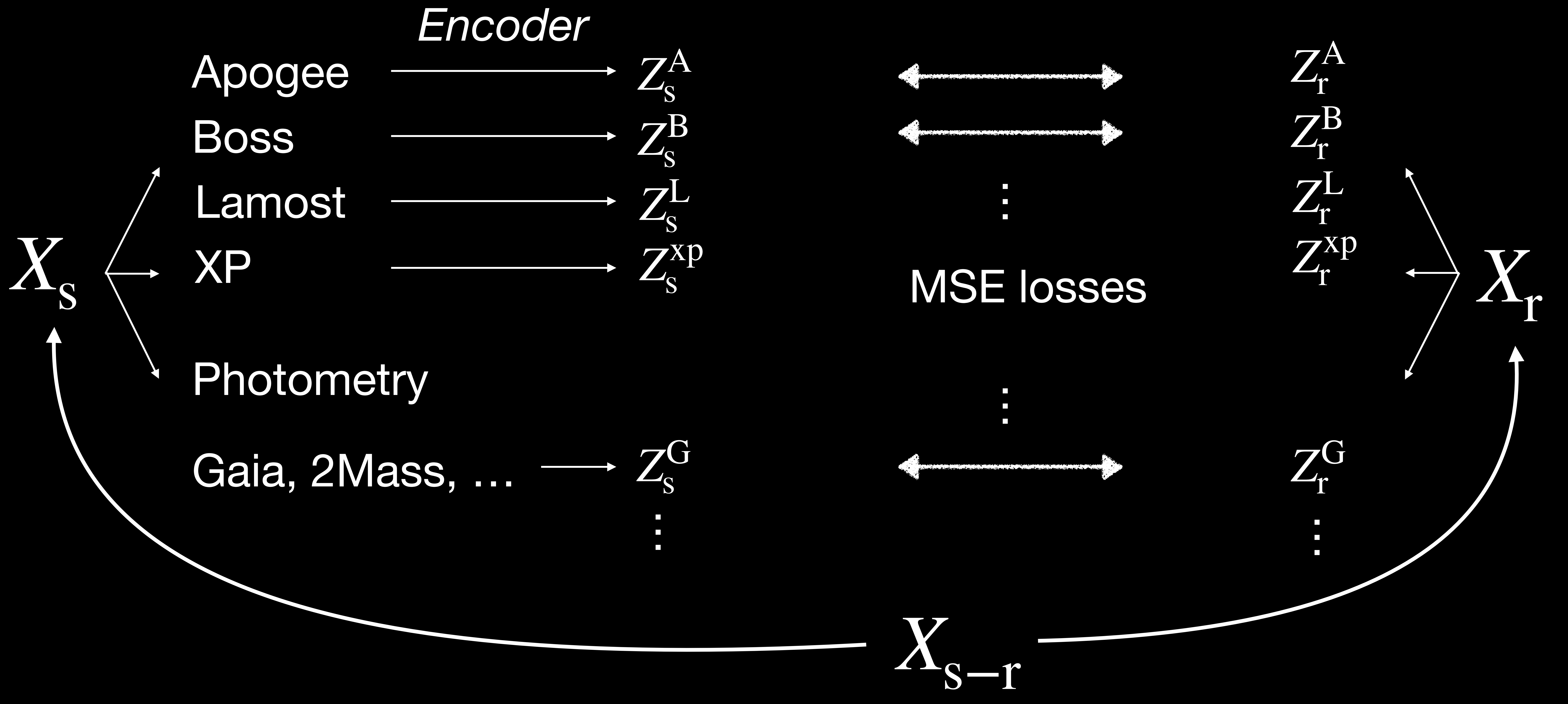
Gaia, 2Mass, … $\longrightarrow Z_s^{G}$

$\vdots$

Modality encoders: alignment loss



Modality encoders: mse loss on pairs



Final model

θ

$(\theta, Z) \longrightarrow$ Transformer $\longrightarrow$ Flow matching

Z^{A}

$X \longrightarrow Z^{\mathrm{pool}}$

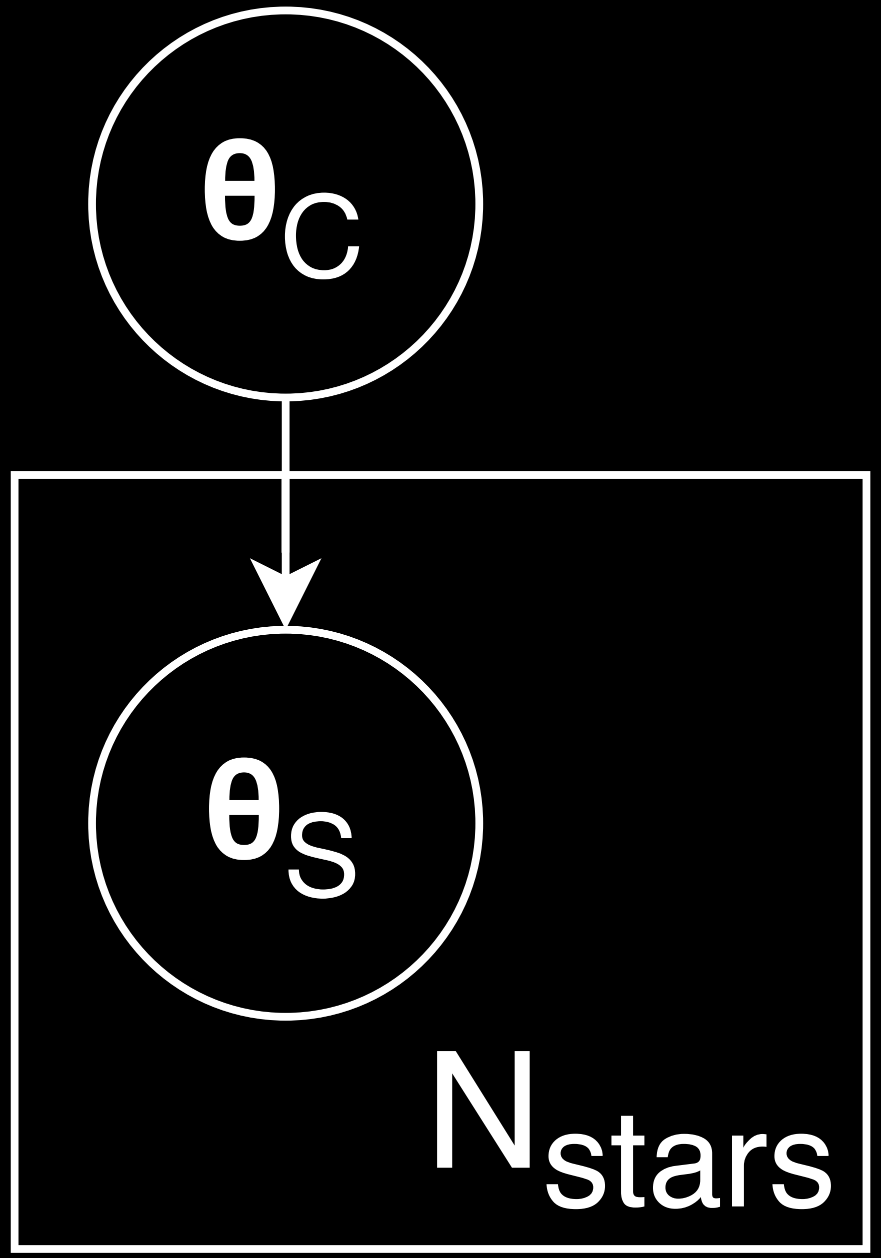
Z^{phot}

Loss

$$\ell(\phi, M_C, t, \hat{\mathbf{x}}_0, \hat{\mathbf{x}}_t) + \mathscr{L}_C + \mathscr{L}_{OT} + \mathscr{L}_{R-S}$$

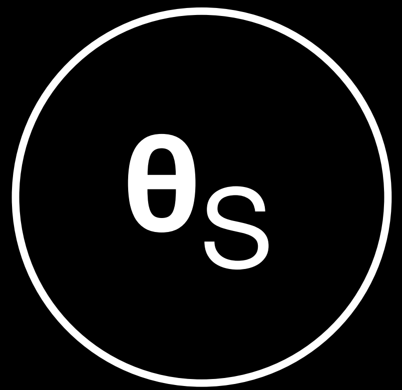
Forward model

Clusters/young stars

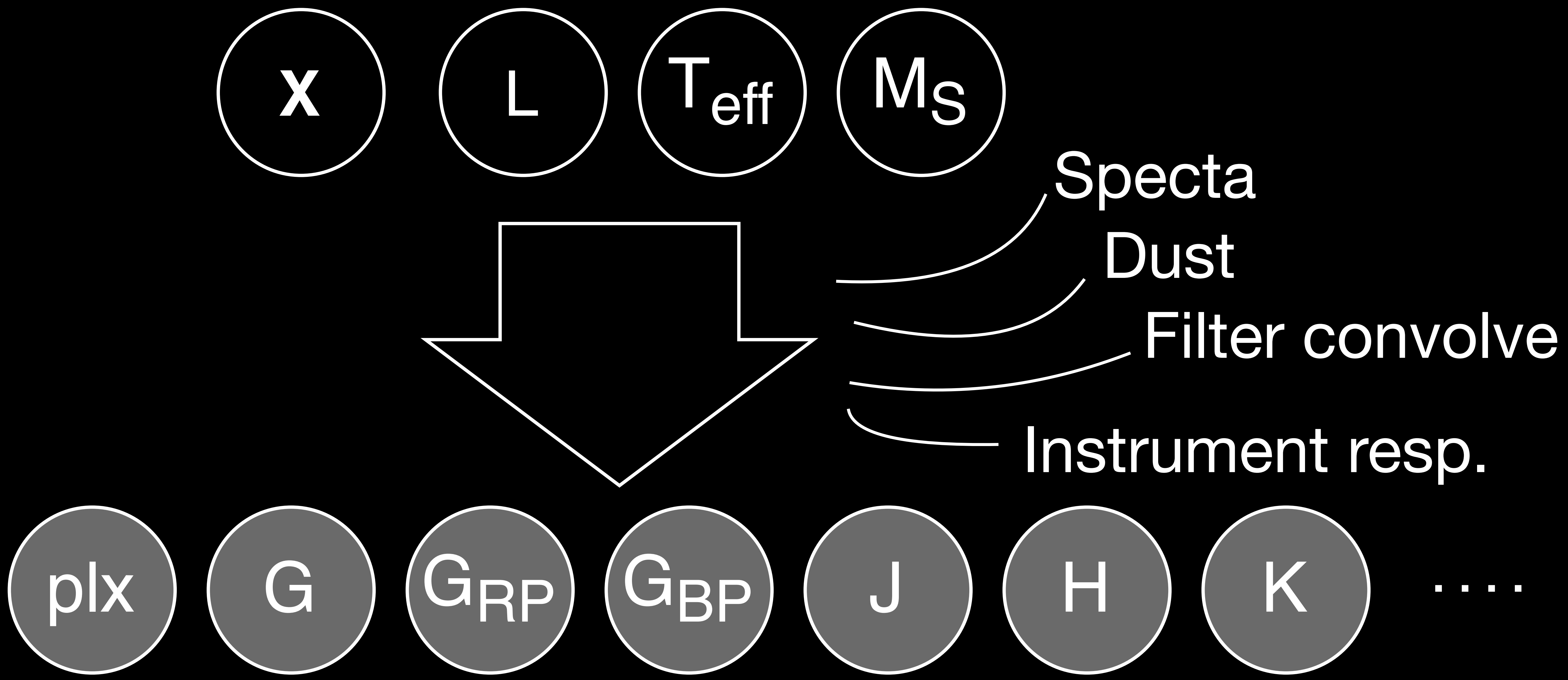


Milky Way model

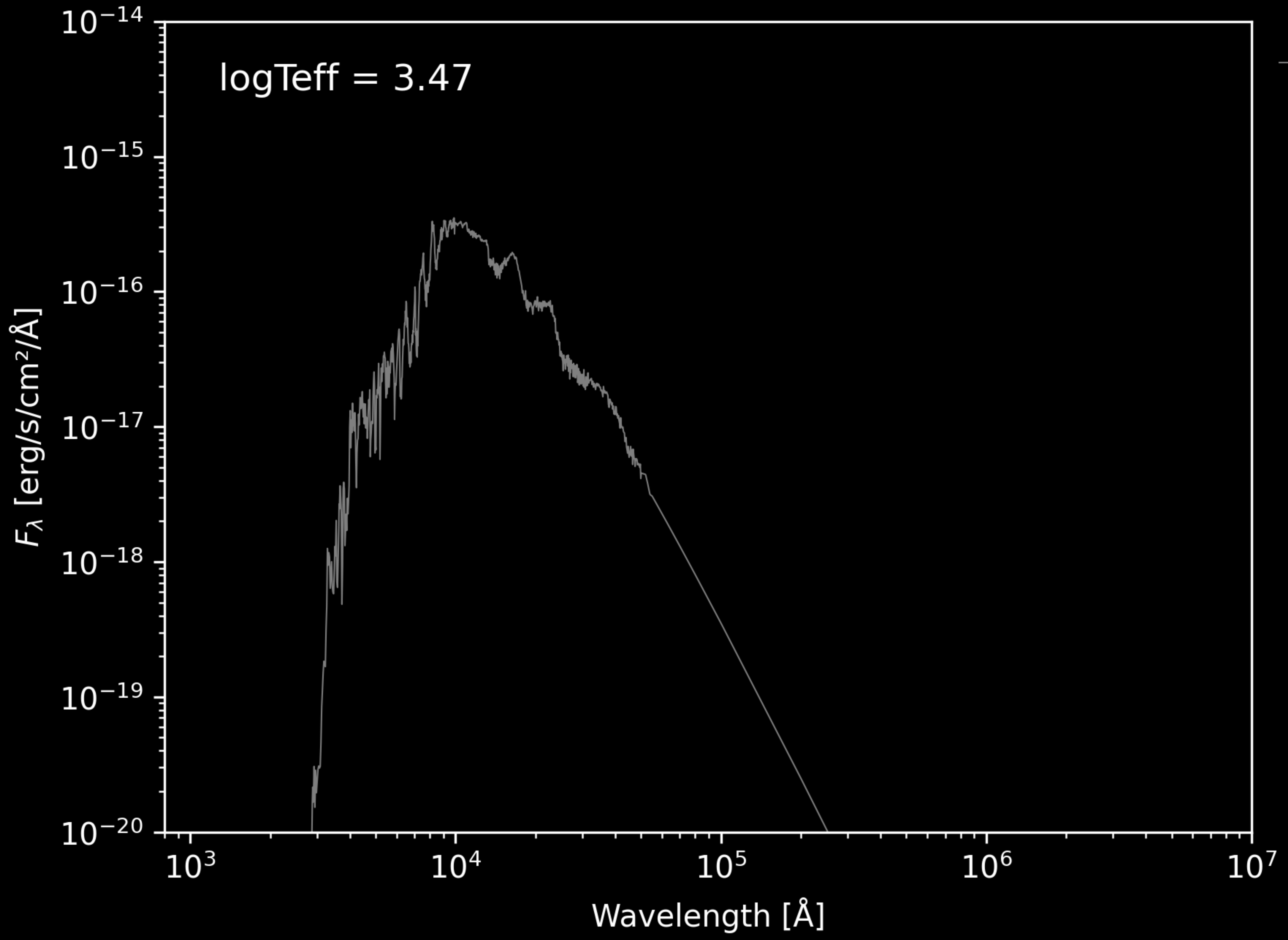
- *galaxia* code
- thin+thick+halo+bulge



Fwd model

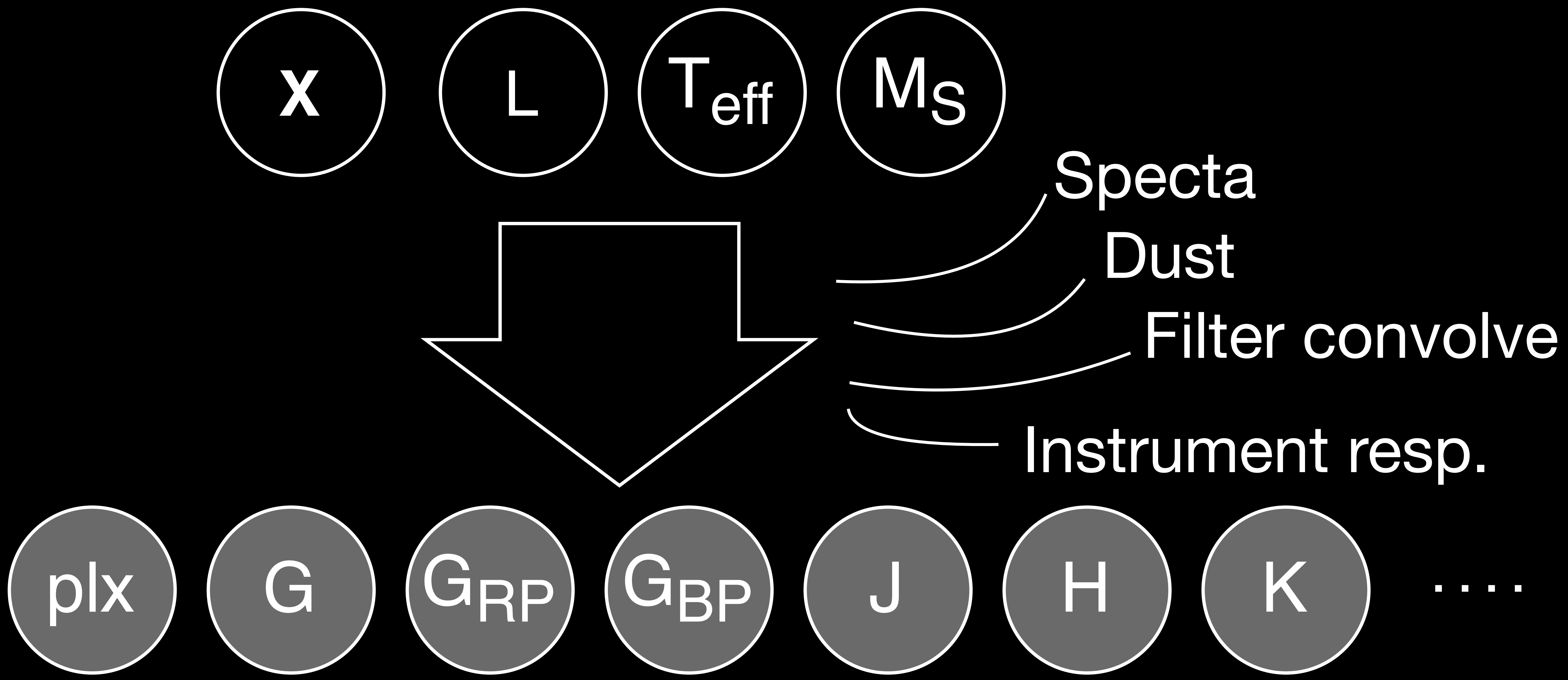


x
L
T_{eff}
M_S
Specta
Dust
Filter convolve
Instrument resp.
plx
G
G_{RP}
G_{BP}
J
H
K
....

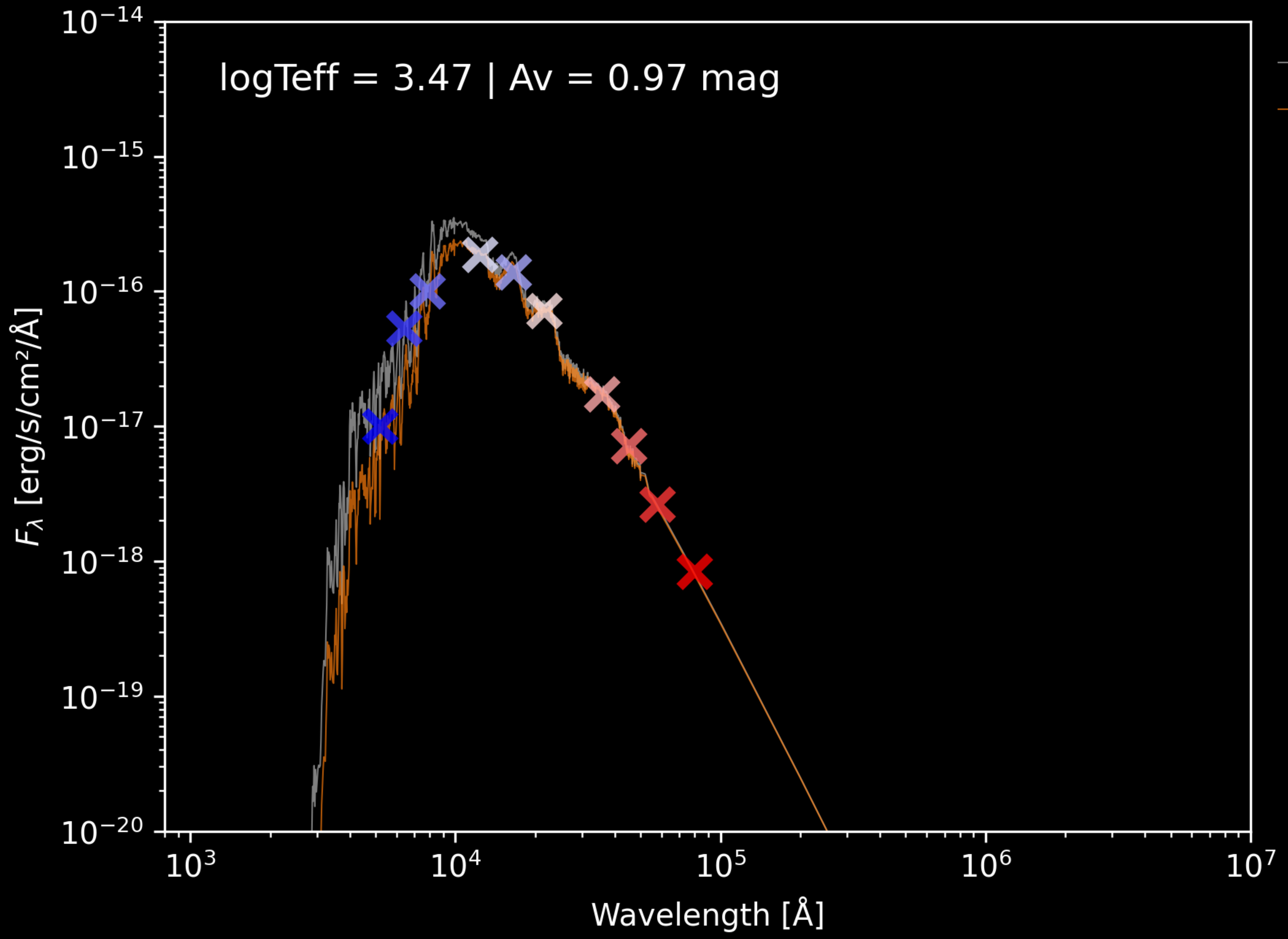


logTeff = 3.47
F_λ [erg/s/cm²/Å]
Wavelength [Å]
10^{-14}
10^{-15}
10^{-16}
10^{-17}
10^{-18}
10^{-19}
10^{-20}
10^3
10^4
10^5
10^6
10^7

Fwd model

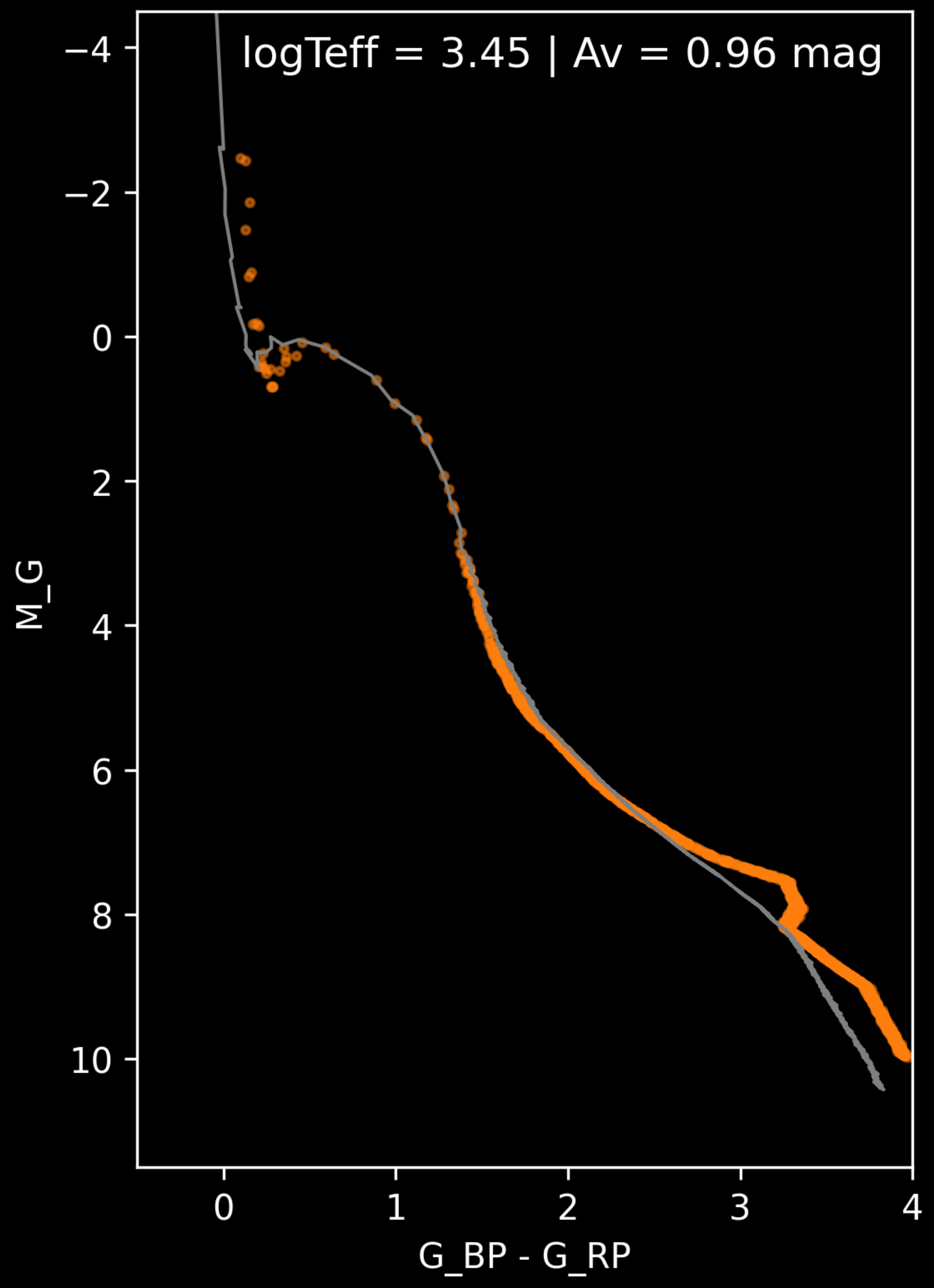


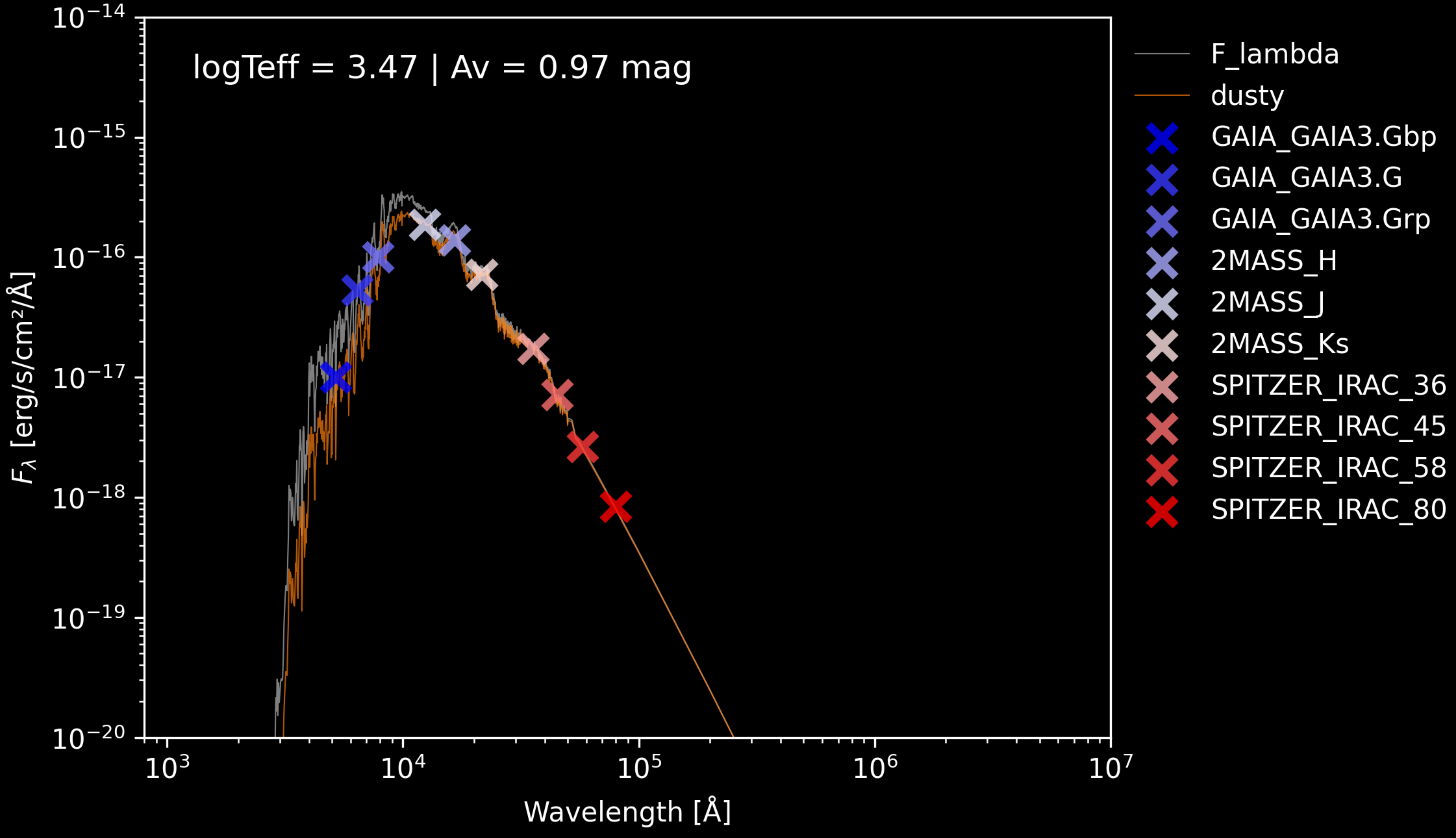
x
L
T_{eff}
M_S
Specta
Dust
Filter convolve
Instrument resp.
plx
G
G_{RP}
G_{BP}
J
H
K
....



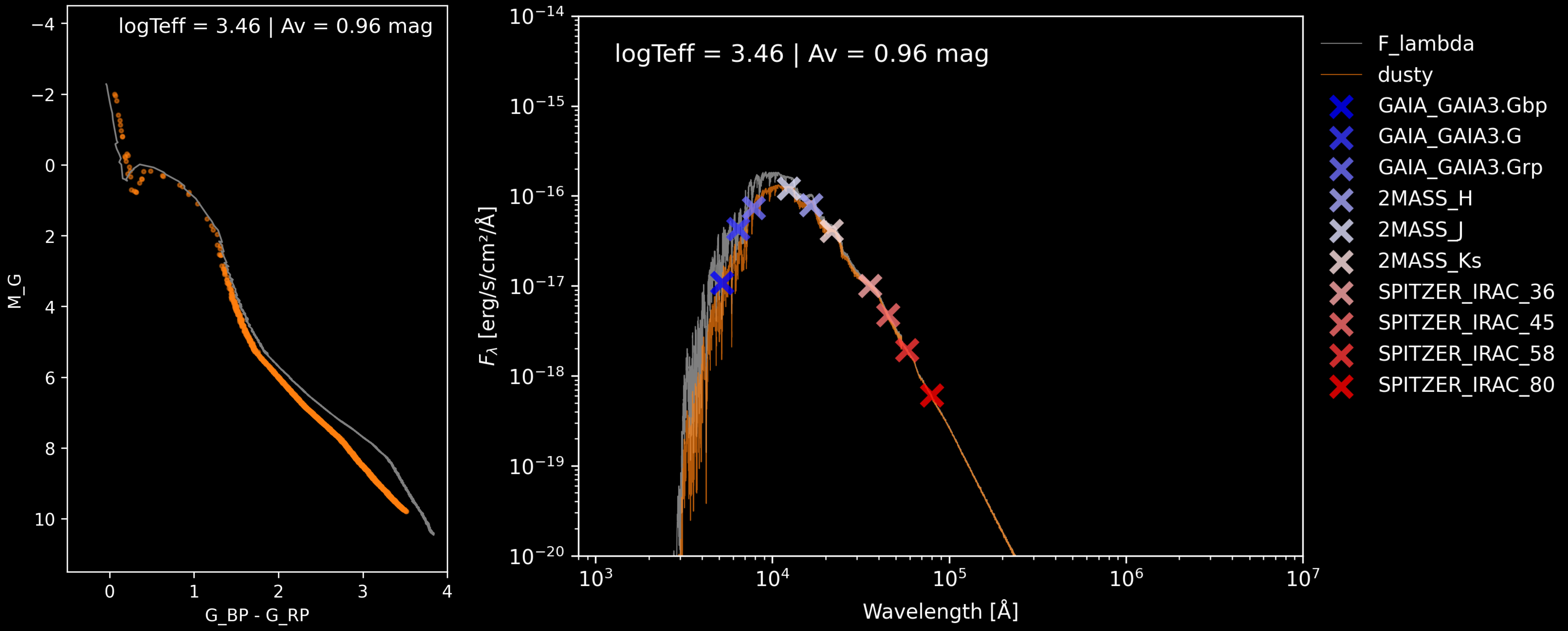
logTeff = 3.47 | Av = 0.97 mag
F_λ [erg/s/cm²/Å]
Wavelength [Å]
10^{-14}
10^{-15}
10^{-16}
10^{-17}
10^{-18}
10^{-19}
10^{-20}
10^3
10^4
10^5
10^6
10^7

Model differences (BaSeL)

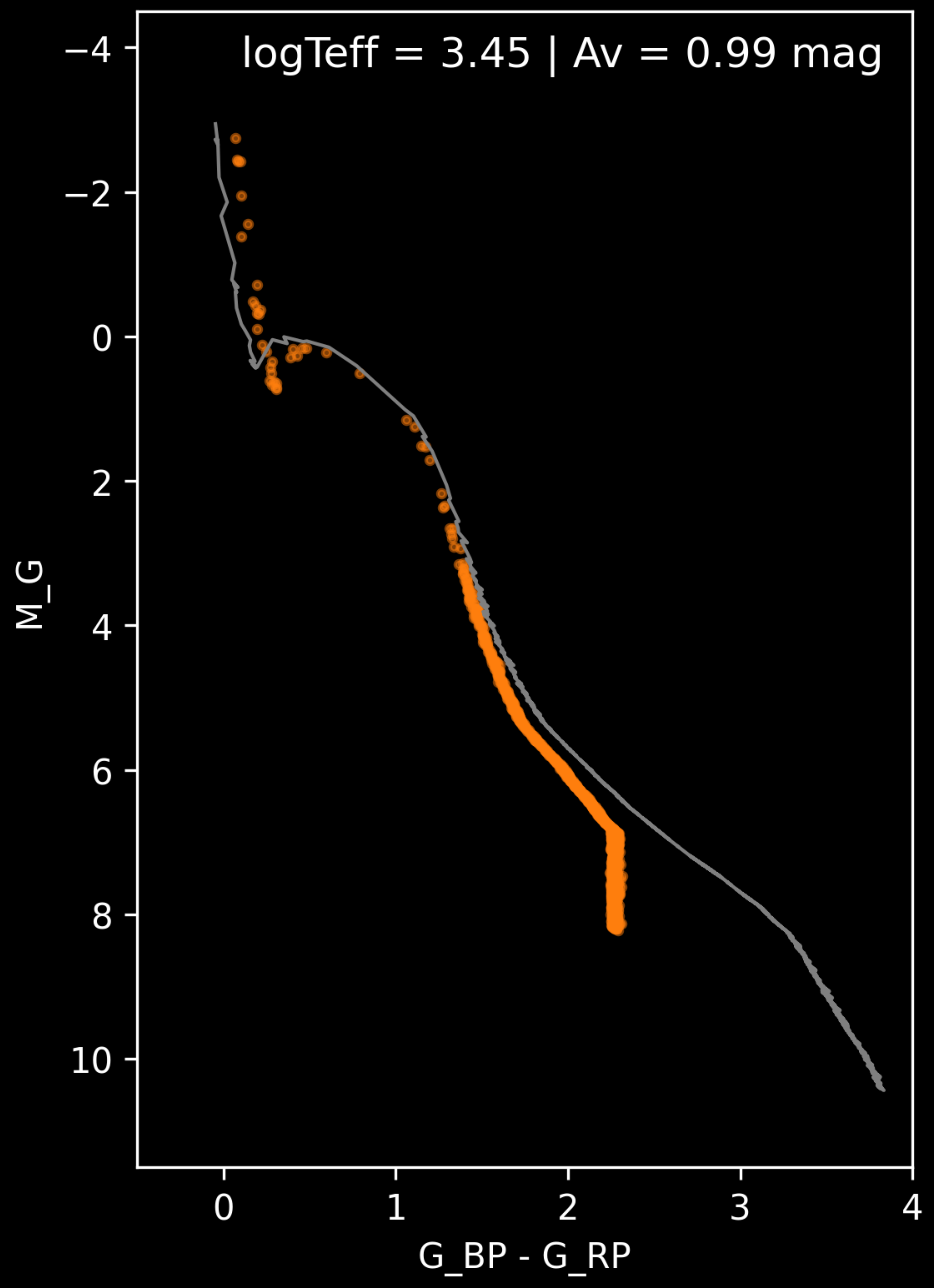


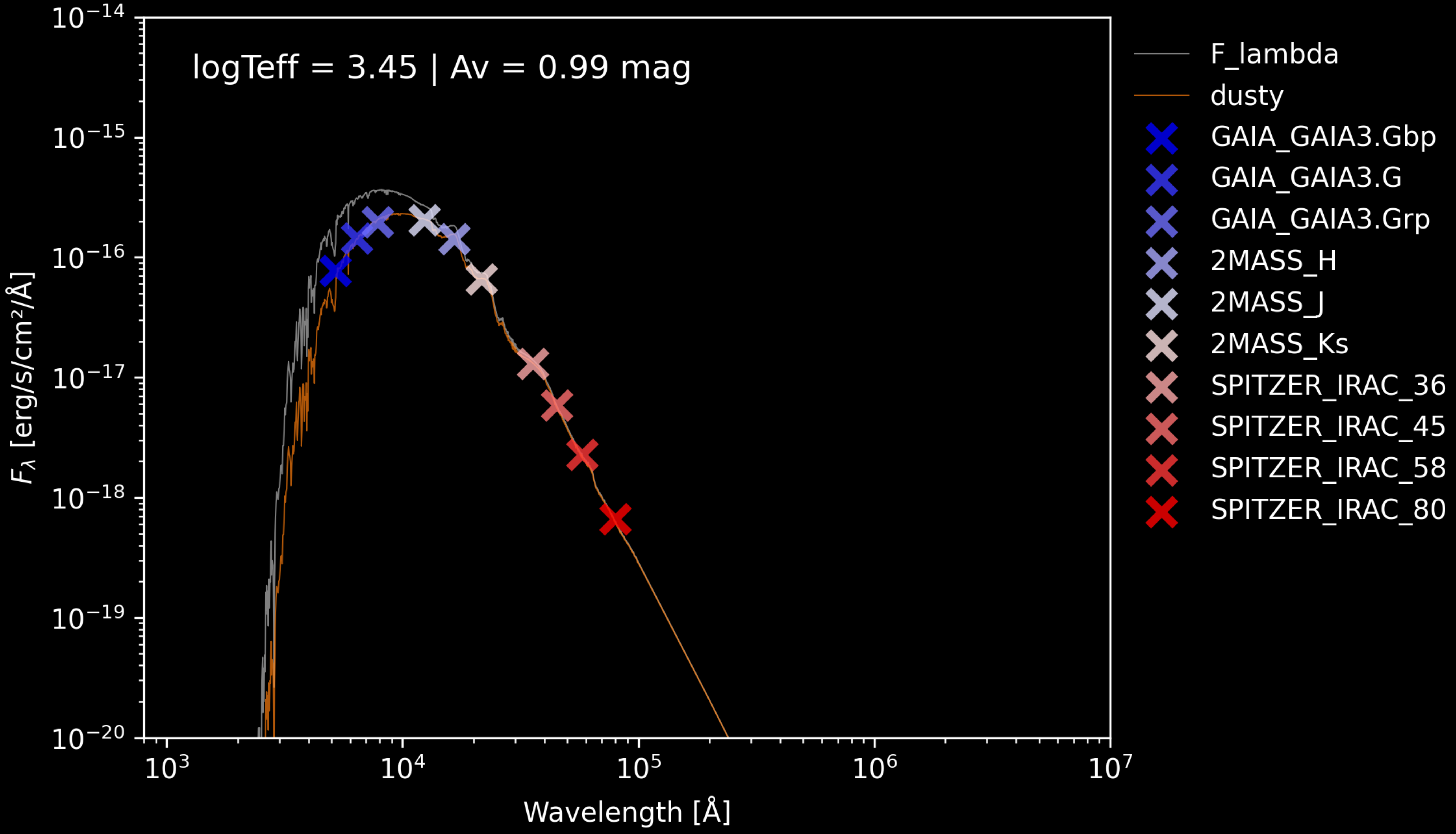


Model differences (BTSettl)



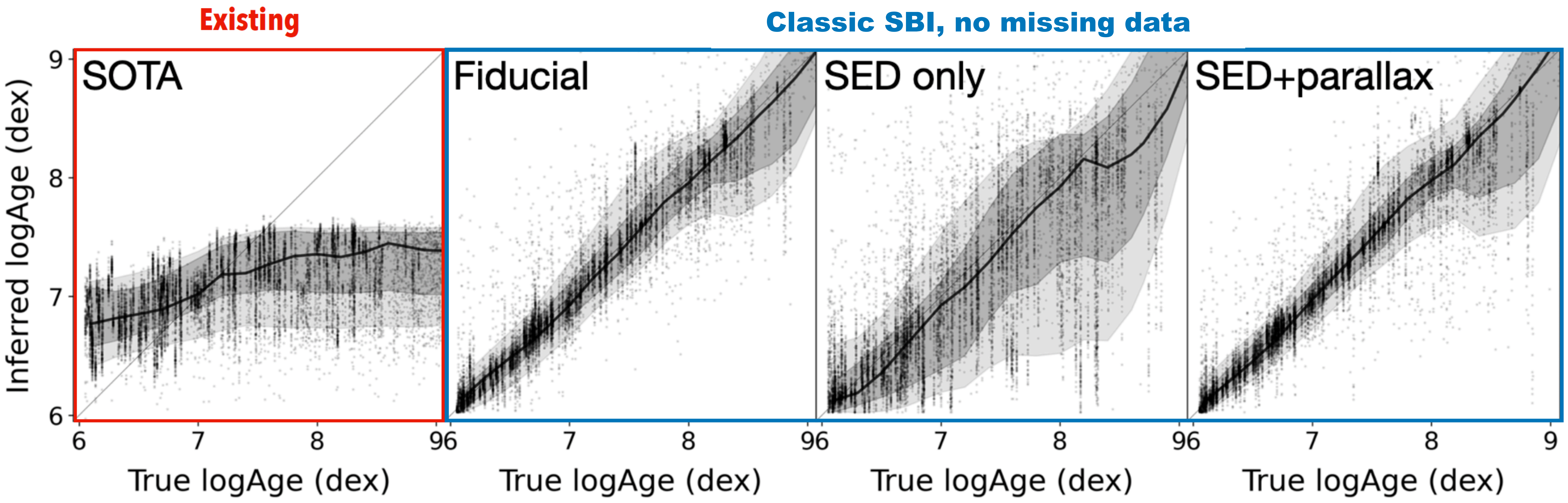
Model differences (Kurucz)



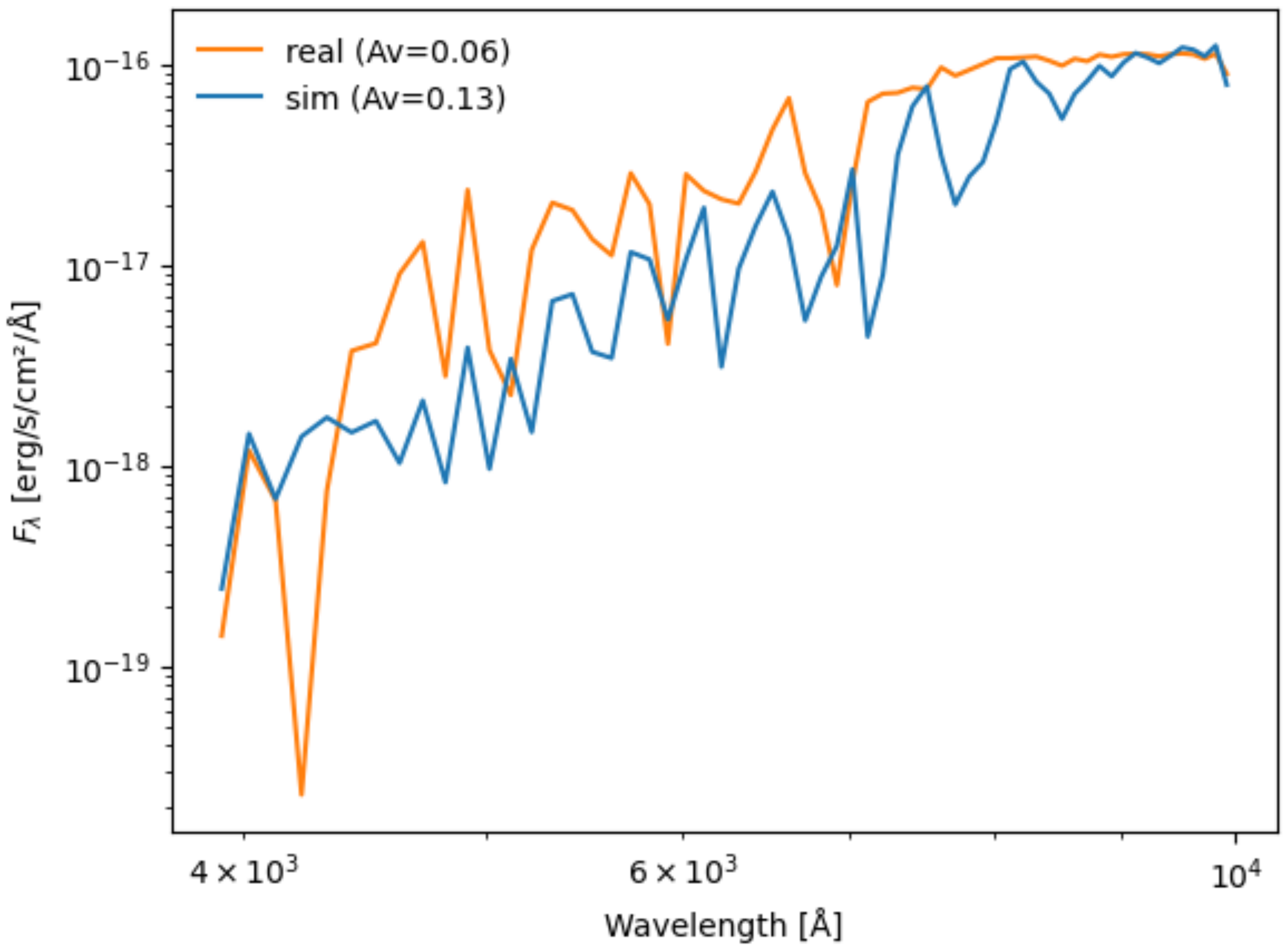


Results

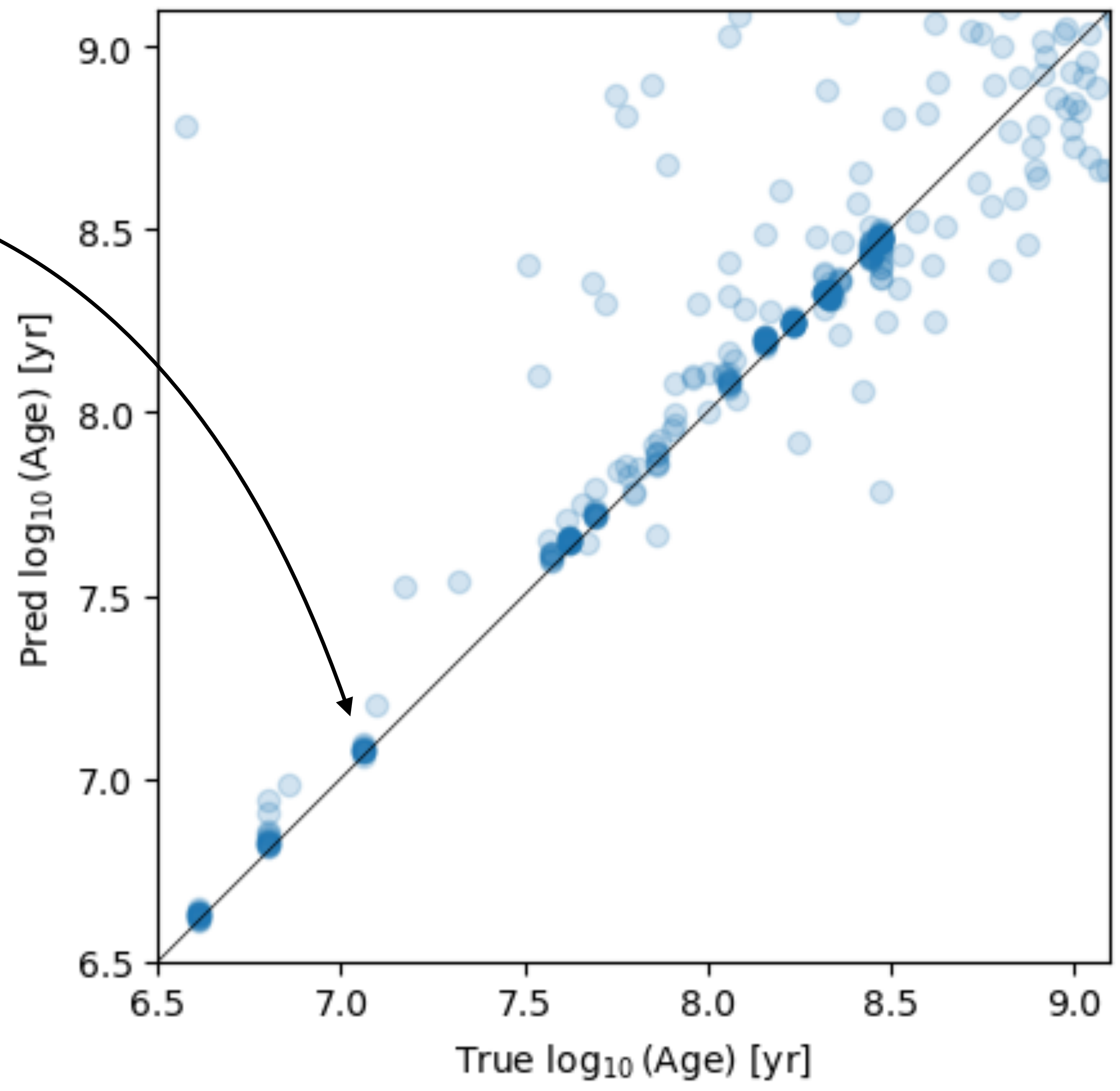
Pilot study: NPE



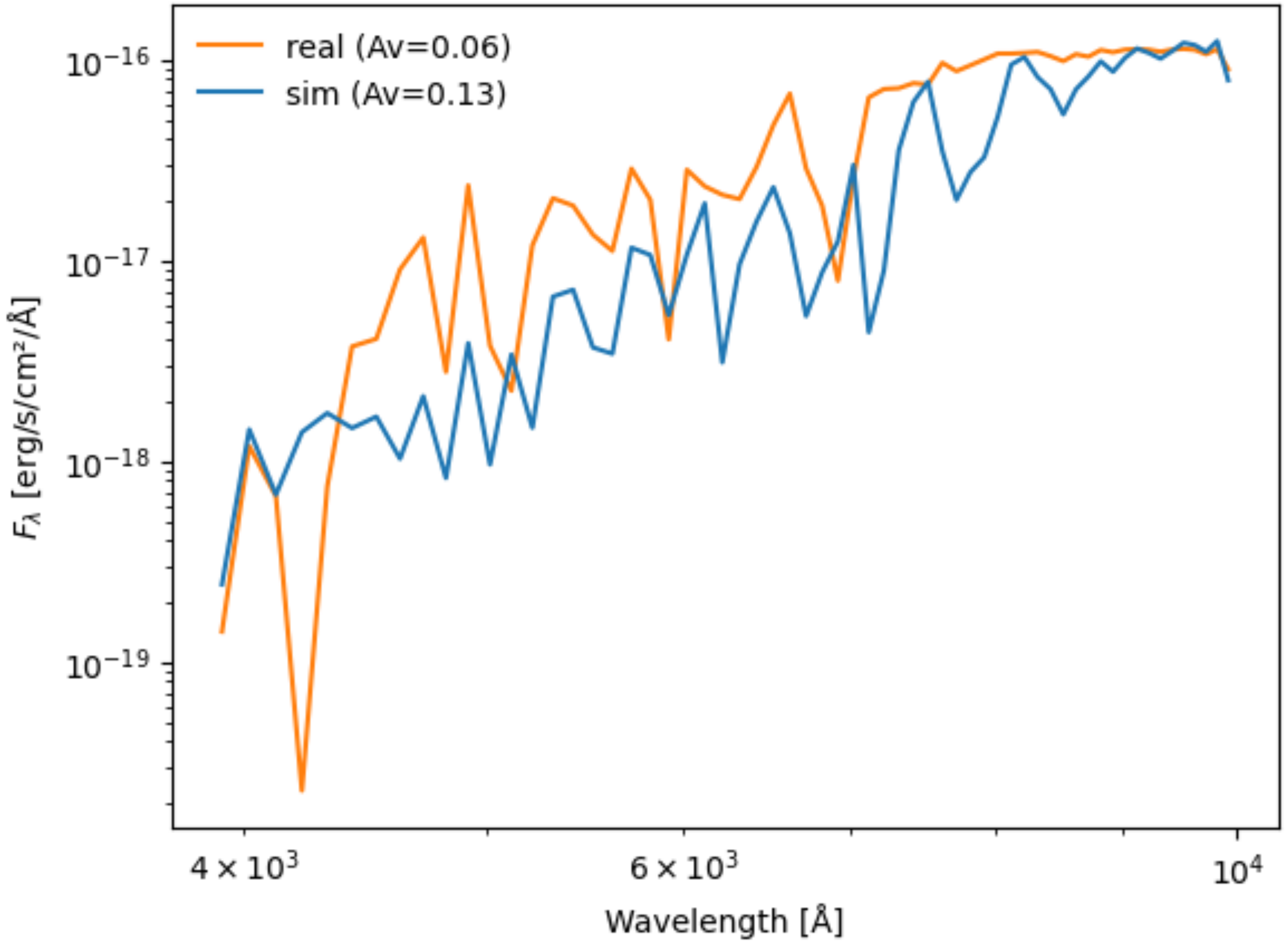
Updated pipeline + 50% missing + XP spectra



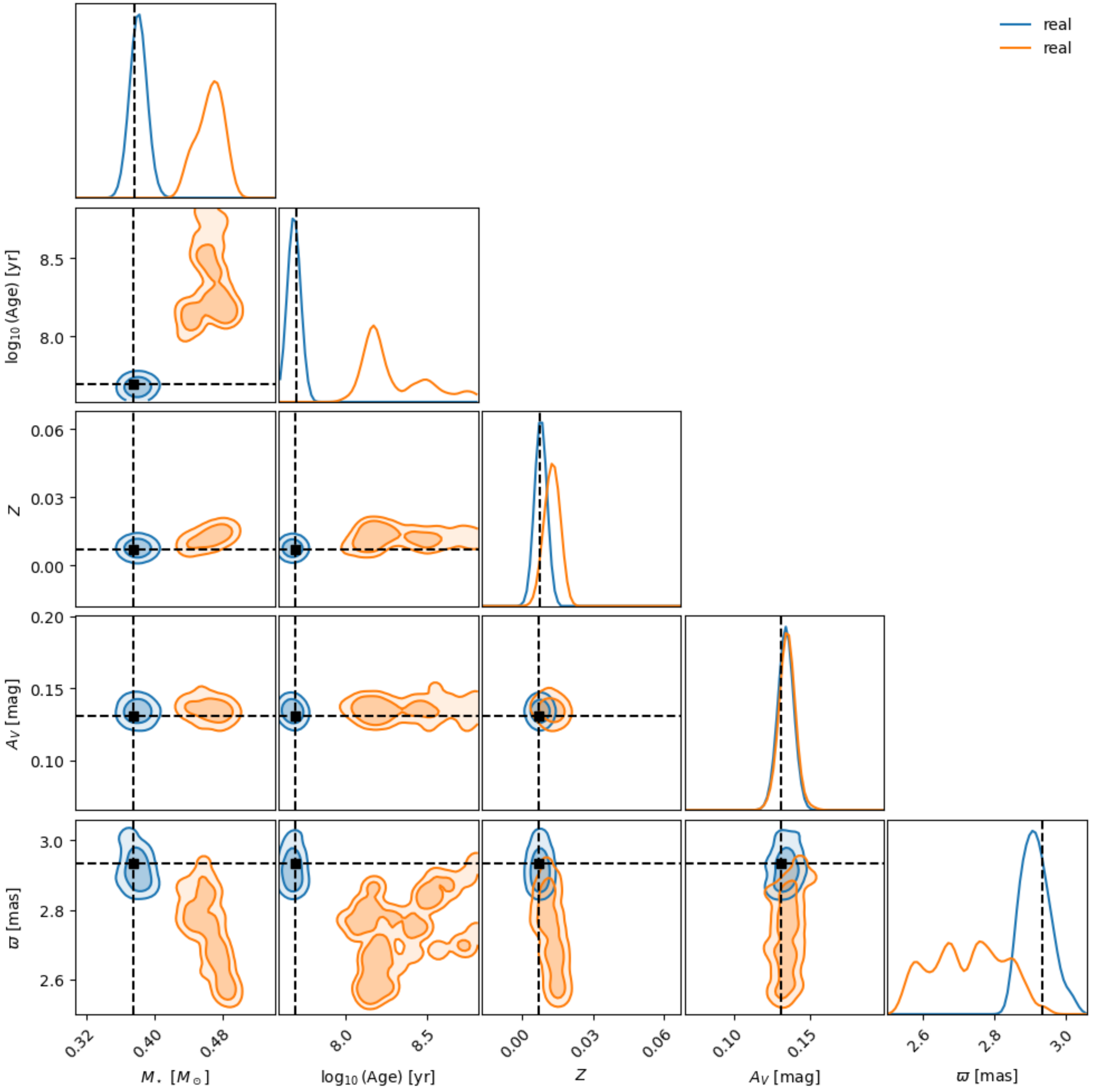
Posterior mean **sim**



Without DA

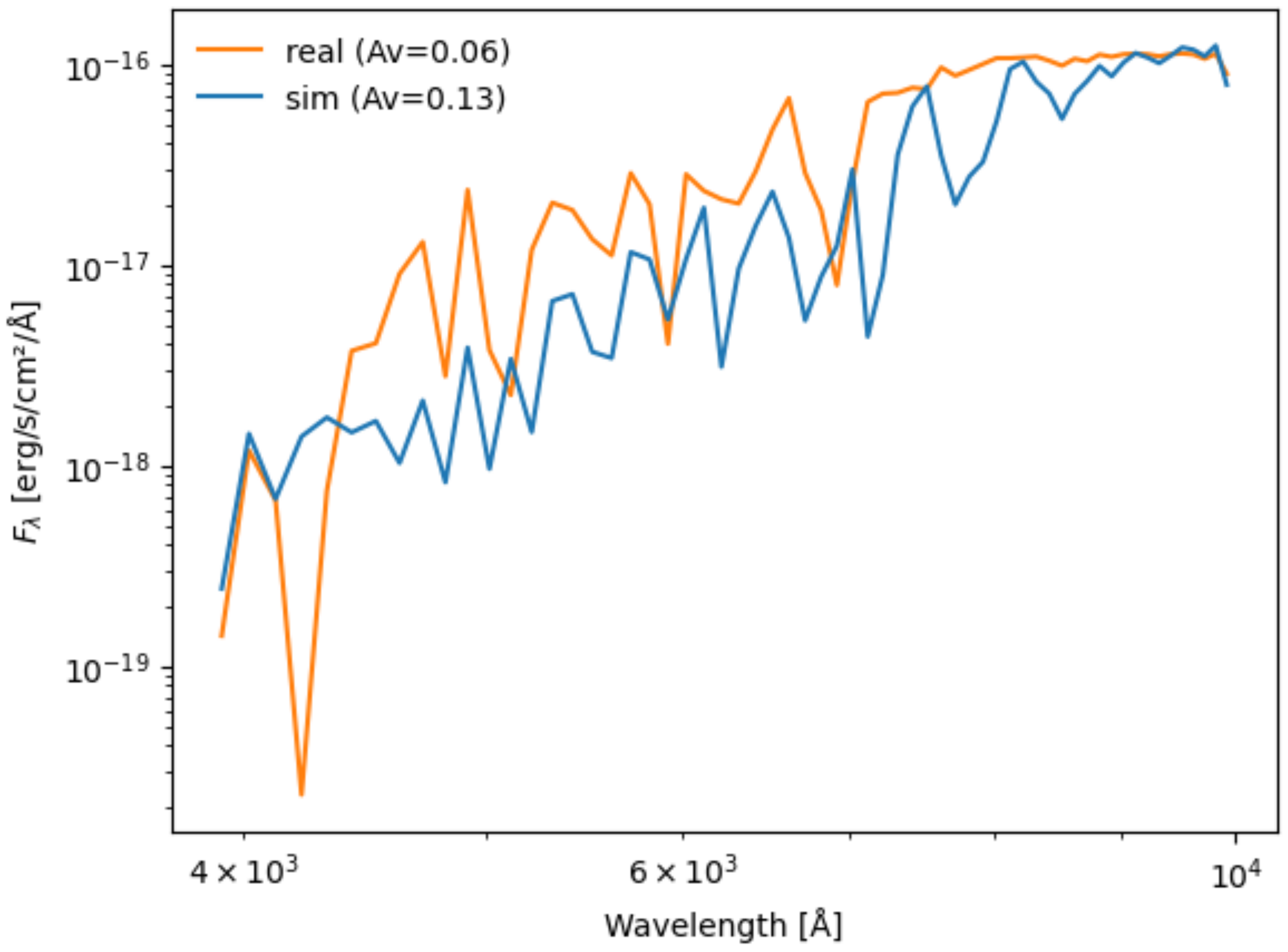


real (Av=0.06)
sim (Av=0.13)
F_λ [erg/s/cm²/Å]
10^{-16}
10^{-17}
10^{-18}
10^{-19}
4×10^3
6×10^3
10^4
Wavelength [Å]

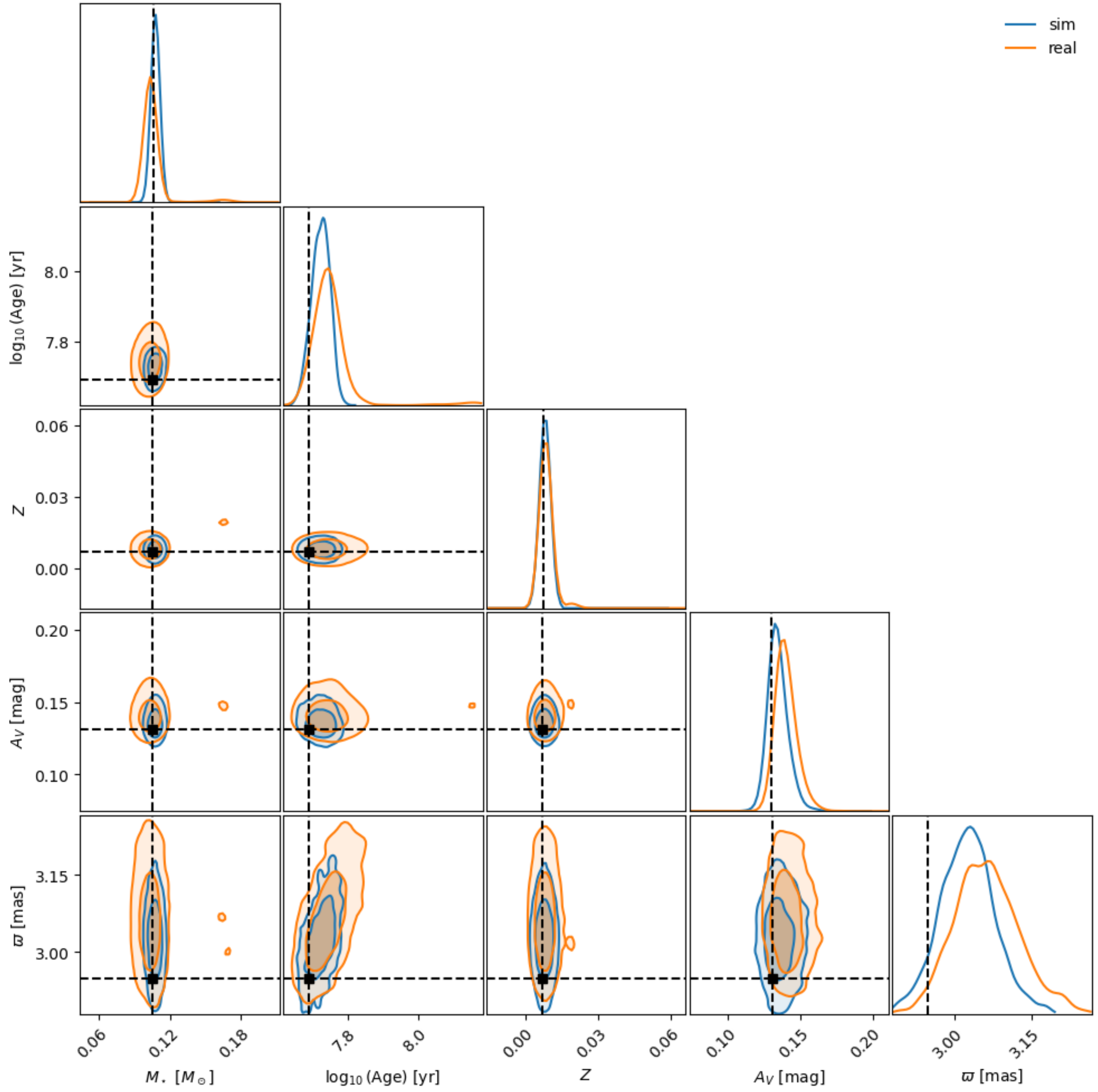


real
real
$\log_{10}$(Age) [yr]
Z
A_V [mag]
ϖ [mas]
$M_\star$ [$M_\odot$]
$\log_{10}$(Age) [yr]
Z
A_V [mag]
ϖ [mas]

With DA



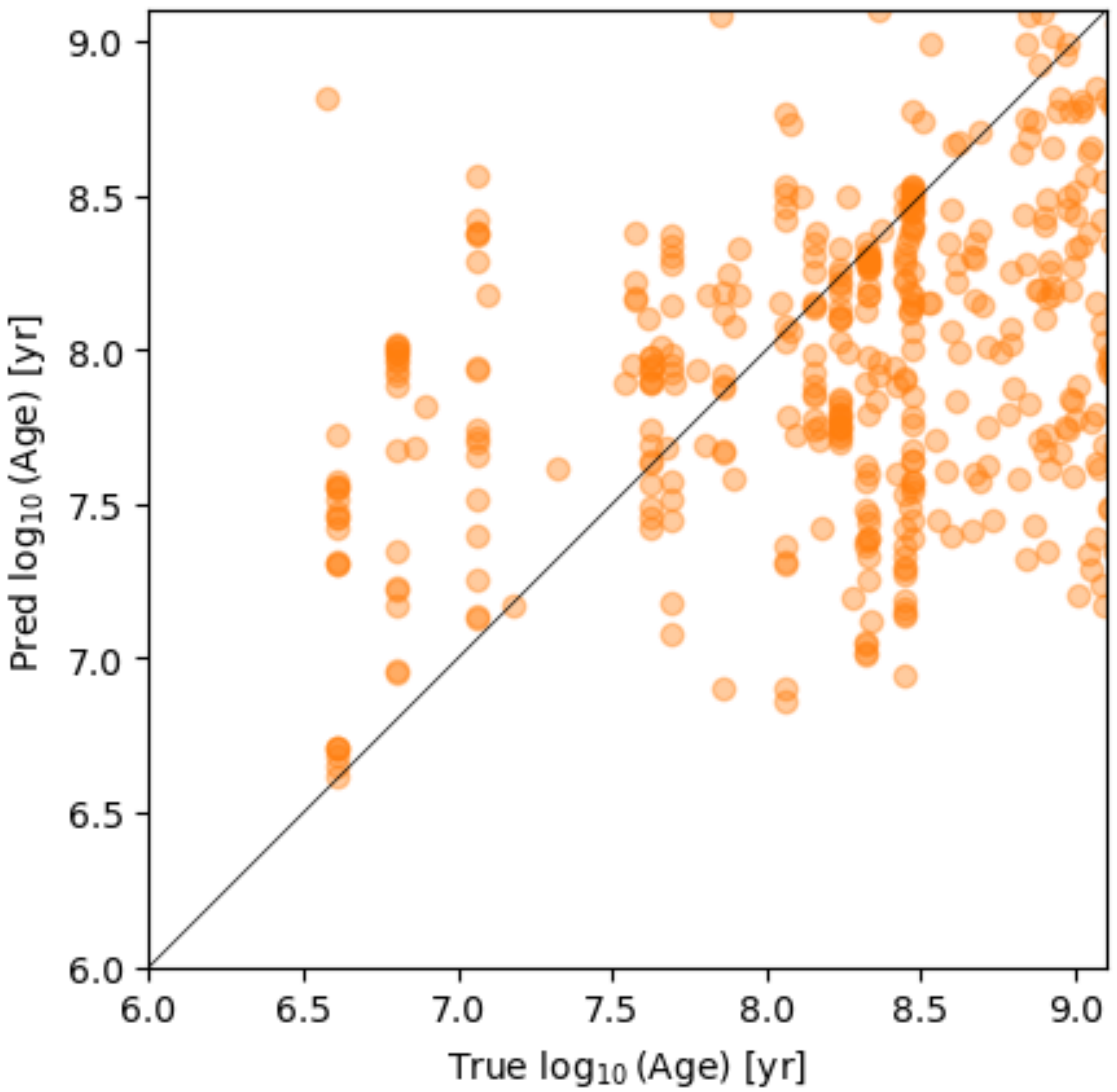
real (Av=0.06)
sim (Av=0.13)
F_λ [erg/s/cm²/Å]
Wavelength [Å]



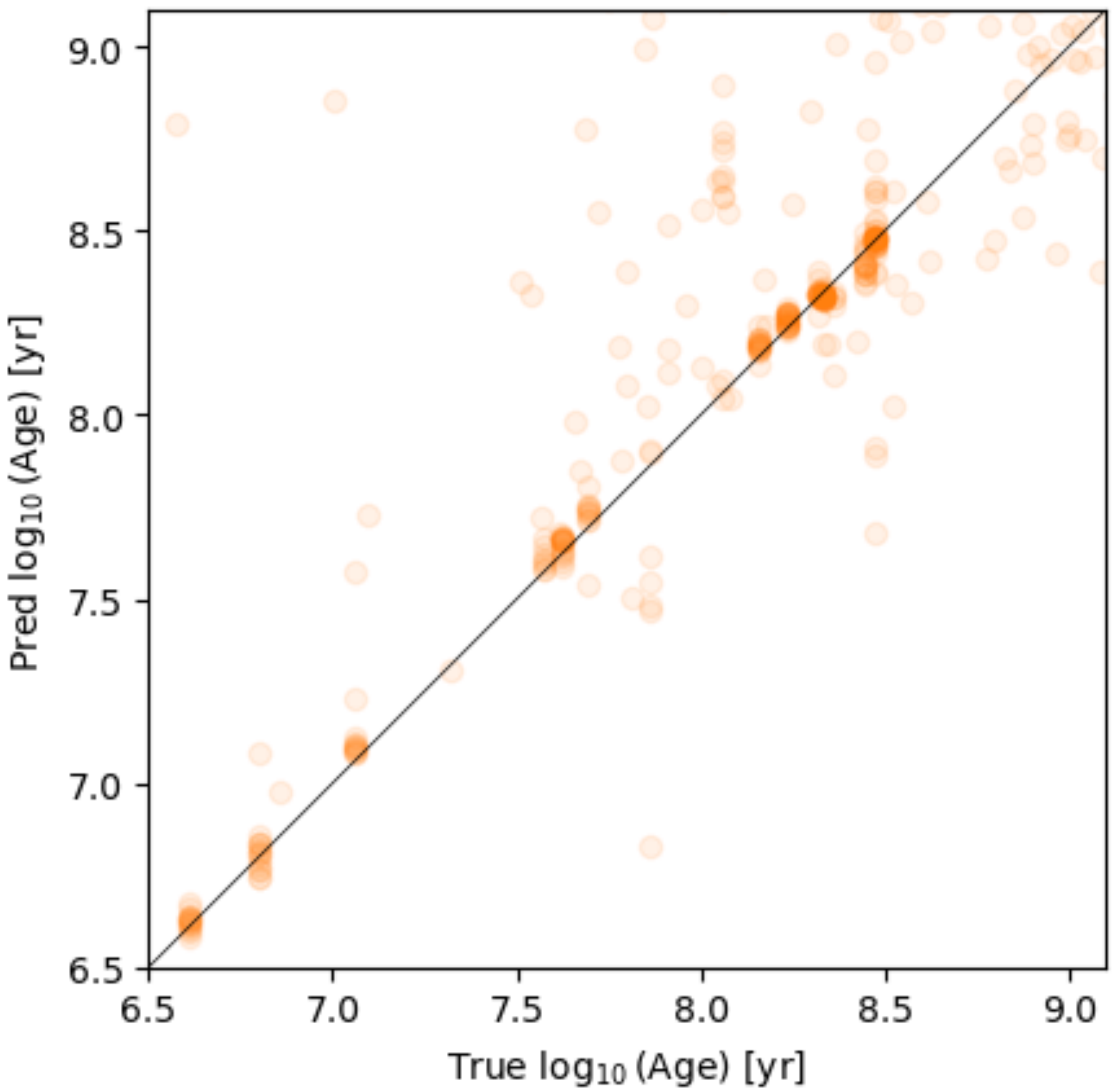
sim
real
$\log_{10}$(Age) [yr]
Z
A_V [mag]
ϖ [mas]
M_* [$M_\odot$]

Predictions

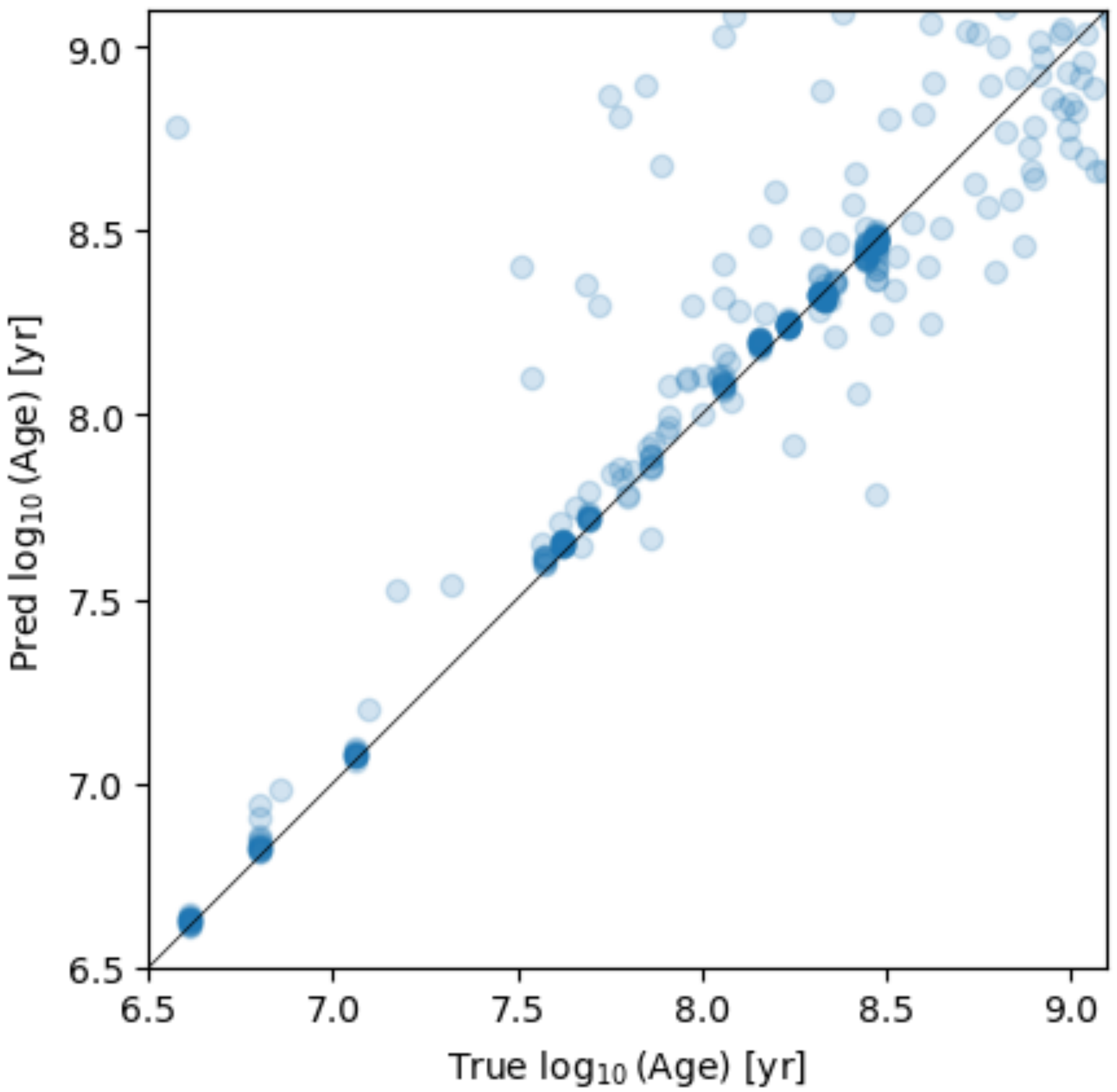
Posterior mean **real**



Posterior mean **real**



Posterior mean **sim**



no domain adaption

with domain adaption

Summary

- Combine **flow matching** + **transformer model** to learn arbitrary **conditionals** and **marginals**
- Add OT + pair + contrastive loss to close domain gap
- Obtain promising results on simulations

Backup

Model implementation

II. Learning to condition and marginalize

Conditional properties: 3D example

$$\ell(\phi, M_C, t, \hat{\mathbf{x}}_0, \hat{\mathbf{x}}_t) = (1 - M_C) \cdot \left(s_\phi^{M_E}(\hat{\mathbf{x}}_t^{M_C}, t) - \nabla_{\hat{\mathbf{x}}_t} \log p_t(\hat{\mathbf{x}}_t | \hat{\mathbf{x}}_0) \right)$$

$$\log p(x, y, z) = \log p(x, y \mid z) + \log p(z)$$

Conditional properties: 3D example

$$\ell(\phi, M_C, t, \hat{\mathbf{x}}_0, \hat{\mathbf{x}}_t) = (1 - M_C) \cdot \left(s_\phi^{M_E}(\hat{\mathbf{x}}_t^{M_C}, t) - \nabla_{\hat{\mathbf{x}}_t} \log p_t(\hat{\mathbf{x}}_t | \hat{\mathbf{x}}_0) \right)$$

$$\log p(x, y, z) = \log p(x, y \mid z) + \log p(z)$$

$$\nabla \log p(x, y, z) = \nabla \log p(x, y \mid z) + \nabla \log p(z)$$

Conditional properties: 3D example

$$\ell(\phi, M_C, t, \hat{\mathbf{x}}_0, \hat{\mathbf{x}}_t) = (1 - M_C) \cdot \left(s_\phi^{M_E}(\hat{\mathbf{x}}_t^{M_C}, t) - \nabla_{\hat{\mathbf{x}}_t} \log p_t(\hat{\mathbf{x}}_t | \hat{\mathbf{x}}_0) \right)$$

$$\log p(x, y, z) = \log p(x, y \mid z) + \log p(z)$$

$$\nabla \log p(x, y, z) = \nabla \log p(x, y \mid z) + \nabla \log p(z)$$

$$\nabla_{x,y} \log p(x, y \mid z) + \nabla_{x,y} \log p(z) = \nabla_{x,y} \log p(x, y \mid z)$$

Conditional properties: 3D example

$$\ell(\phi, M_C, t, \hat{\mathbf{x}}_0, \hat{\mathbf{x}}_t) = (1 - M_C) \cdot \left(s_\phi^{M_E}(\hat{\mathbf{x}}_t^{M_C}, t) - \nabla_{\hat{\mathbf{x}}_t} \log p_t(\hat{\mathbf{x}}_t | \hat{\mathbf{x}}_0) \right)$$

$$\log p(x, y, z) = \log p(x, y \mid z) + \log p(z)$$

$$\nabla \log p(x, y, z) = \nabla \log p(x, y \mid z) + \nabla \log p(z)$$

$$\nabla_{x,y} \log p(x, y \mid z) + \nabla_{x,y} \log p(z) = \nabla_{x,y} \log p(x, y \mid z)$$

$\nabla_z \log p(x, y \mid z) + \nabla_z \log p(z) \leftarrow$ Is set to 0 due to $(1 - M_C)$ in ℓ

Marginalization properties

Attention masking: M_E

$\hat{X} \longrightarrow$ Tokenizer: $\hat{T} \longrightarrow$ Transformer

Marginalization properties

Attention masking: M_E

$$\hat{X} \longrightarrow \text{Tokenizer: } \hat{T} \longrightarrow \text{Transformer}$$

$$\hat{T} \in (N, D_{\hat{X}}, E) \rightarrow S \propto QK^T \in (N, D_{\hat{X}}, D_{\hat{X}})$$

$$\text{softmax}(S + M_E)\, V P \in (N, D_{\hat{X}})$$

$$M_E = \begin{pmatrix} 0 & 0 & -\infty \\ 0 & 0 & -\infty \\ -\infty & -\infty & 0 \end{pmatrix} \longrightarrow s_\phi^{M_E}(\hat{x}^{M_C}) = \begin{pmatrix} f(x,y) \\ f(x,y) \\ f(z) \end{pmatrix}$$

Marginalization properties

Attention masking: M_E

$$\downarrow$$

$$\hat{X} \longrightarrow \text{Tokenizer: } \hat{T} \longrightarrow \text{Transformer}$$

$$\hat{T} \in (N, D_{\hat{X}}, E) \rightarrow S \propto QK^T \in (N, D_{\hat{X}}, D_{\hat{X}})$$

$$\mathrm{softmax}(S + M_E)\, V P \in (N, D_{\hat{X}})$$

$$M_E = \begin{pmatrix} 0 & 0 & -\infty \\ 0 & 0 & -\infty \\ -\infty & -\infty & 0 \end{pmatrix} \longrightarrow s_\phi^{M_E}(\hat{x}^{M_C}) = \begin{pmatrix} f(x,y) \\ f(x,y) \\ f(z) \end{pmatrix}$$

$$\ell(\phi, M_C, t, \hat{\mathbf{x}}_0, \hat{\mathbf{x}}_t) = (1 - M_C) \cdot \left(s_\phi^{M_E}(\hat{\mathbf{x}}_t^{M_C}, t) - \nabla_{\hat{\mathbf{x}}_t} \log p_t(\hat{\mathbf{x}}_t | \hat{\mathbf{x}}_0) \right) \propto \begin{pmatrix} f(x,y) - \Delta x \\ f(x,y) - \Delta y \\ f(z) - \Delta z \end{pmatrix}$$

$$\propto \hat{x}_0 - \hat{x}_t$$

Marginalization properties

$$\ell(\phi, M_C, t, \hat{\mathbf{x}}_0, \hat{\mathbf{x}}_t) = (1 - M_C) \cdot \left(s_\phi^{M_E}(\hat{\mathbf{x}}_t^{M_C}, t) - \nabla_{\hat{\mathbf{x}}_t} \log p_t(\hat{\mathbf{x}}_t | \hat{\mathbf{x}}_0) \right) \propto \begin{pmatrix} f(x, y) - \Delta x \\ f(x, y) - \Delta y \\ f(z) - \Delta z \end{pmatrix}$$

$$\propto \hat{x}_0 - \hat{x}_t$$

$$\mathscr{L} \propto (f(x, y) - \Delta x)^2 + (f(x, y) - \Delta y)^2 + (f(z) - \Delta z)^2$$

Marginalization properties

$$\ell(\phi, M_C, t, \hat{\mathbf{x}}_0, \hat{\mathbf{x}}_t) = (1 - M_C) \cdot \left(s_\phi^{M_E}(\hat{\mathbf{x}}_t^{M_C}, t) - \nabla_{\hat{\mathbf{x}}_t} \log p_t(\hat{\mathbf{x}}_t | \hat{\mathbf{x}}_0) \right) \propto \begin{pmatrix} f(x, y) - \Delta x \\ f(x, y) - \Delta y \\ f(z) - \Delta z \end{pmatrix}$$

$$\propto \hat{x}_0 - \hat{x}_t$$

$$\mathscr{L} \propto (f(x, y) - \Delta x)^2 + (f(x, y) - \Delta y)^2 + (f(z) - \Delta z)^2 =$$

$$= \nabla \log p(x, y) + \nabla \log p(z) = \nabla \log p(x, y) p(z)$$

YSO models

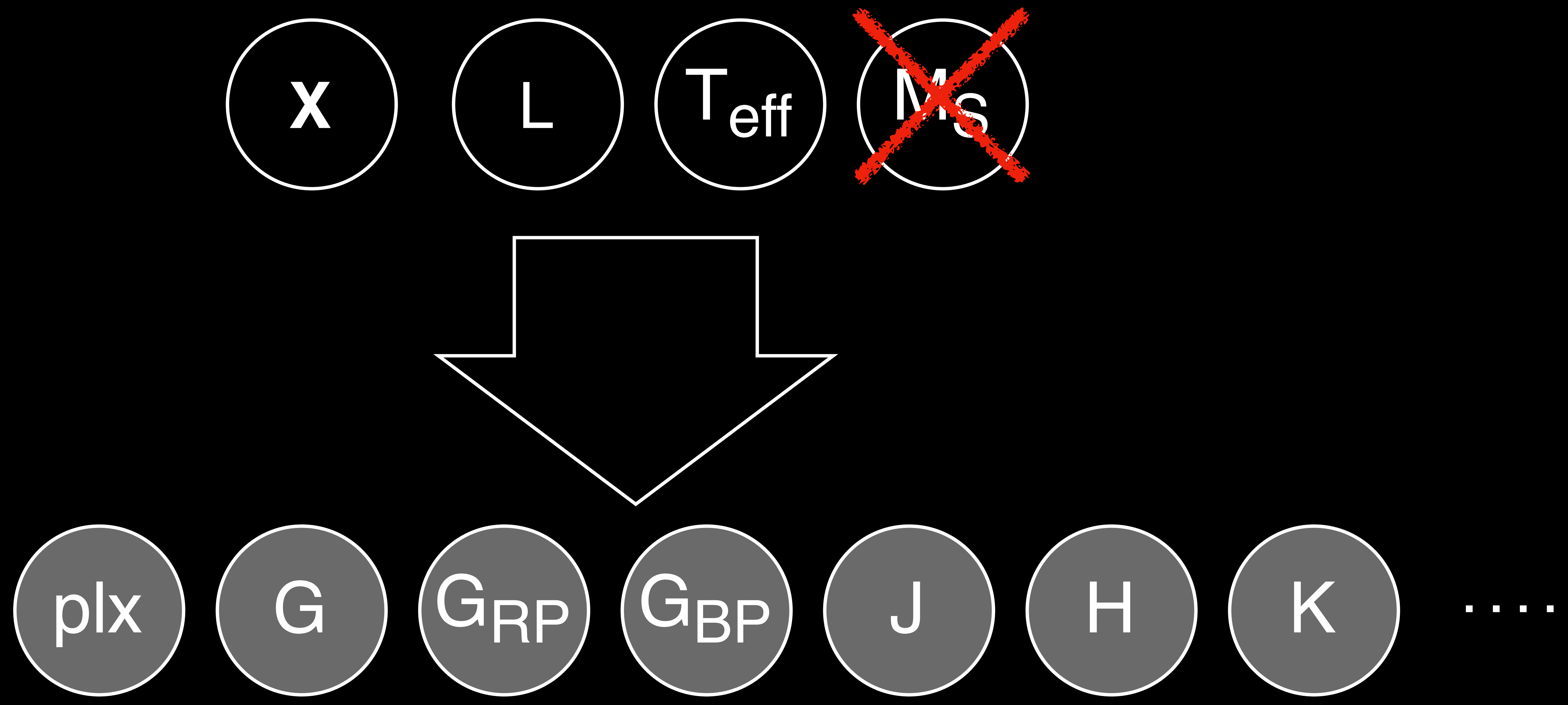
Fwd model: YSO (Robitaille17+Richardson+24)

- Mass not input parameter
  - Unlink to evo. tracks

➡ logg = 4

➡ Small impact for $T_{\text{eff}} \in [3,20]$ kK

- Low resolution (R~15)



Fwd model: YSO (Robitaille17+Richardson+24)

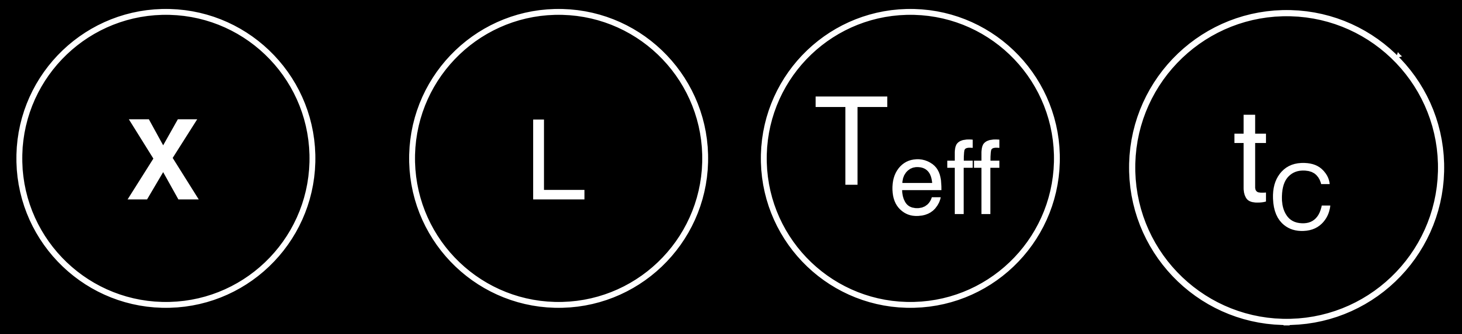
- Mass not input parameter
  - Unlink to evo. tracks
  - ➡ logg = 4
  - ➡ Small impact for $T_{\mathrm{eff}} \in [3,20]$ kK
- Low resolution (R~15)
- Introduces many additional parameters

X L T_{eff}

Parameter	Symbol	Minimum	Maximum	Sampling
Stellar radius	$R_\star$	$0.1\,R_\odot$	$100\,R_\odot$	Log
Stellar temperature	$T_\star$	2000 K	30 000 K	Log
Disk mass [dust]	M_{disk}	$10^{-8}\,M_\odot$	$0.1\,M_\odot$	Log
Disk inner radius	$R_{\mathrm{min}}^{\mathrm{disk}}$	R_{sub}	$1000\,R_{\mathrm{sub}}$	Log
Disk outer radius	$R_{\mathrm{max}}^{\mathrm{disk}}$	50 AU	5000 AU	Log
Disk flaring power	β	1	1.3	Linear
Disk surface density power	p	−2	0	Linear
Disk scaleheight	$h_{100\mathrm{AU}}$	1 AU	20 AU	Log
Envelope density [dust]	ρ_0^{env}	10^{-24} g/cm^3	10^{-16} g/cm^3	Log
Envelope density power	γ	−2	−1	Linear
Envelope centrifugal radius	R_{c}	50 AU	5000 AU	Log
Cavity density [dust]	ρ_0^{cav}	10^{-23} g/cm^3	10^{-20} g/cm^3	Log
Cavity opening angle	θ_0	0°	60°	Linear
Cavity power	c	1	2	Linear

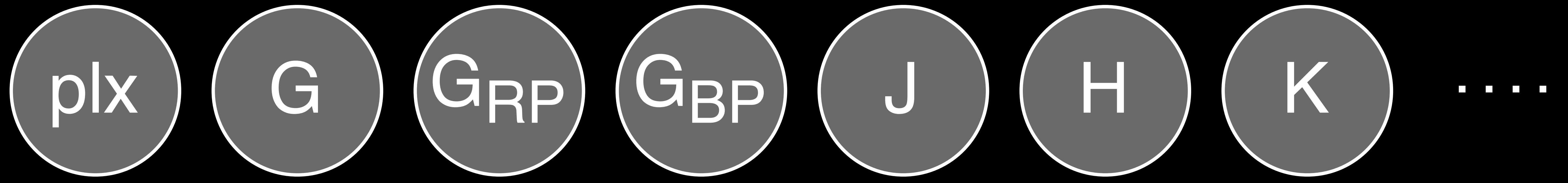
plx G G_{RP} G_{BP} J H K ....

Fwd model: YSO (Robitaille17+Richardson+24)

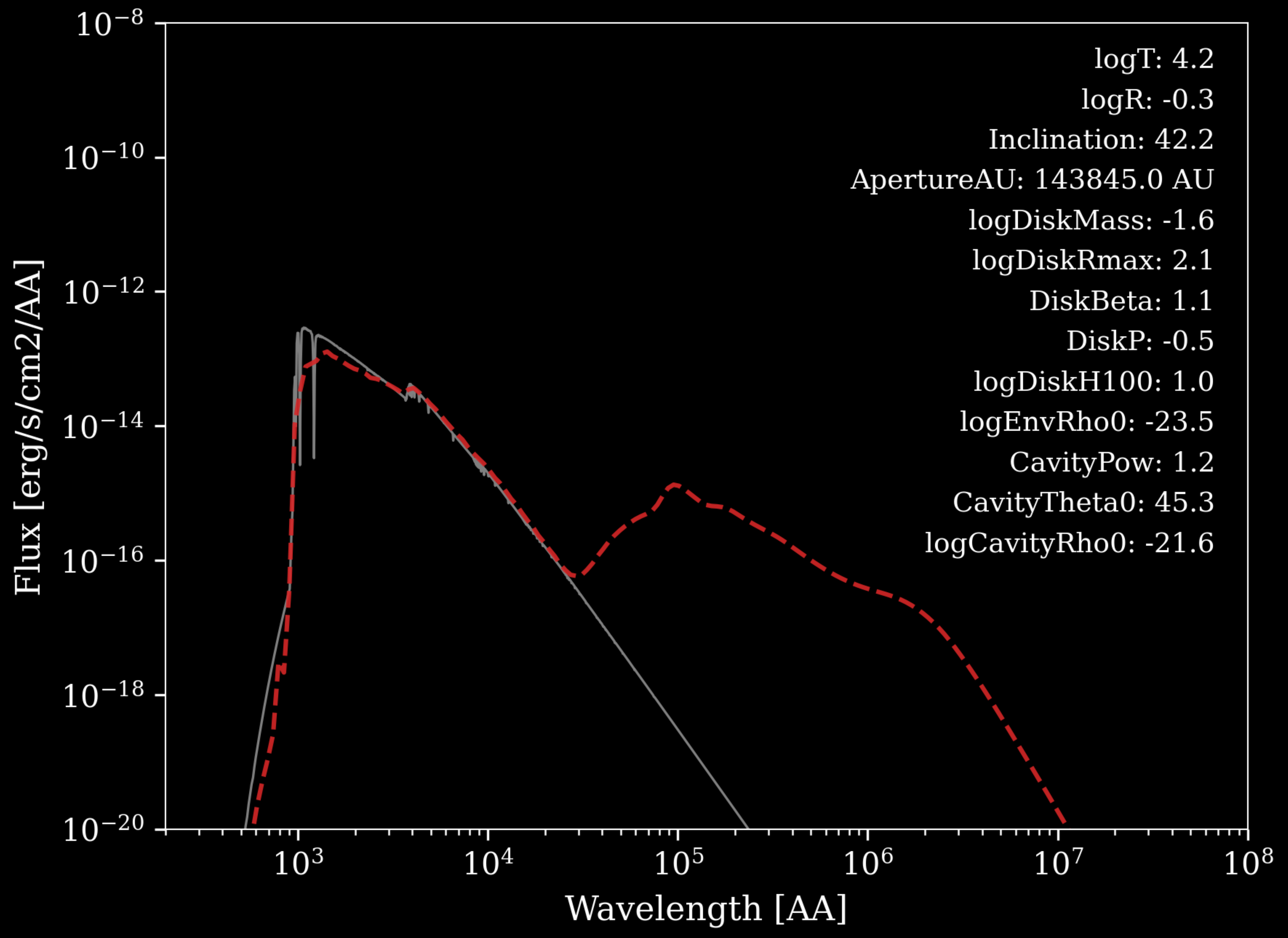


- Mass not input parameter
  - Unlink to evo. tracks
  - ➡ logg = 4
  - ➡ Small impact for $T_{\mathrm{eff}} \in [3,20]$ kK
- Low resolution (R~15)
- Introduces many additional parameters

Parameter	Symbol	Minimum	Maximum	Sampling
Stellar radius	$R_\star$	$0.1\,R_\odot$	$100\,R_\odot$	Log
Stellar temperature	$T_\star$	2000 K	30 000 K	Log
Disk mass [dust]	M_{disk}	$10^{-8}\,M_\odot$	$0.1\,M_\odot$	Log
Disk inner radius	$R_{\mathrm{min}}^{\mathrm{disk}}$	R_{sub}	$1000\,R_{\mathrm{sub}}$	Log
Disk outer radius	$R_{\mathrm{max}}^{\mathrm{disk}}$	50 AU	5000 AU	Log
Disk flaring power	β	1	1.3	Linear
Disk surface density power	p	−2	0	Linear
Disk scaleheight	$h_{100\mathrm{AU}}$	1 AU	20 AU	Log
Envelope density [dust]	ρ_0^{env}	10^{-24} g/cm^3	10^{-16} g/cm^3	Log
Envelope density power	γ	−2	−1	Linear
Envelope centrifugal radius	R_c	50 AU	5000 AU	Log
Cavity density [dust]	ρ_0^{cav}	10^{-23} g/cm^3	10^{-20} g/cm^3	Log
Cavity opening angle	θ_0	0°	60°	Linear
Cavity power	c	1	2	Linear

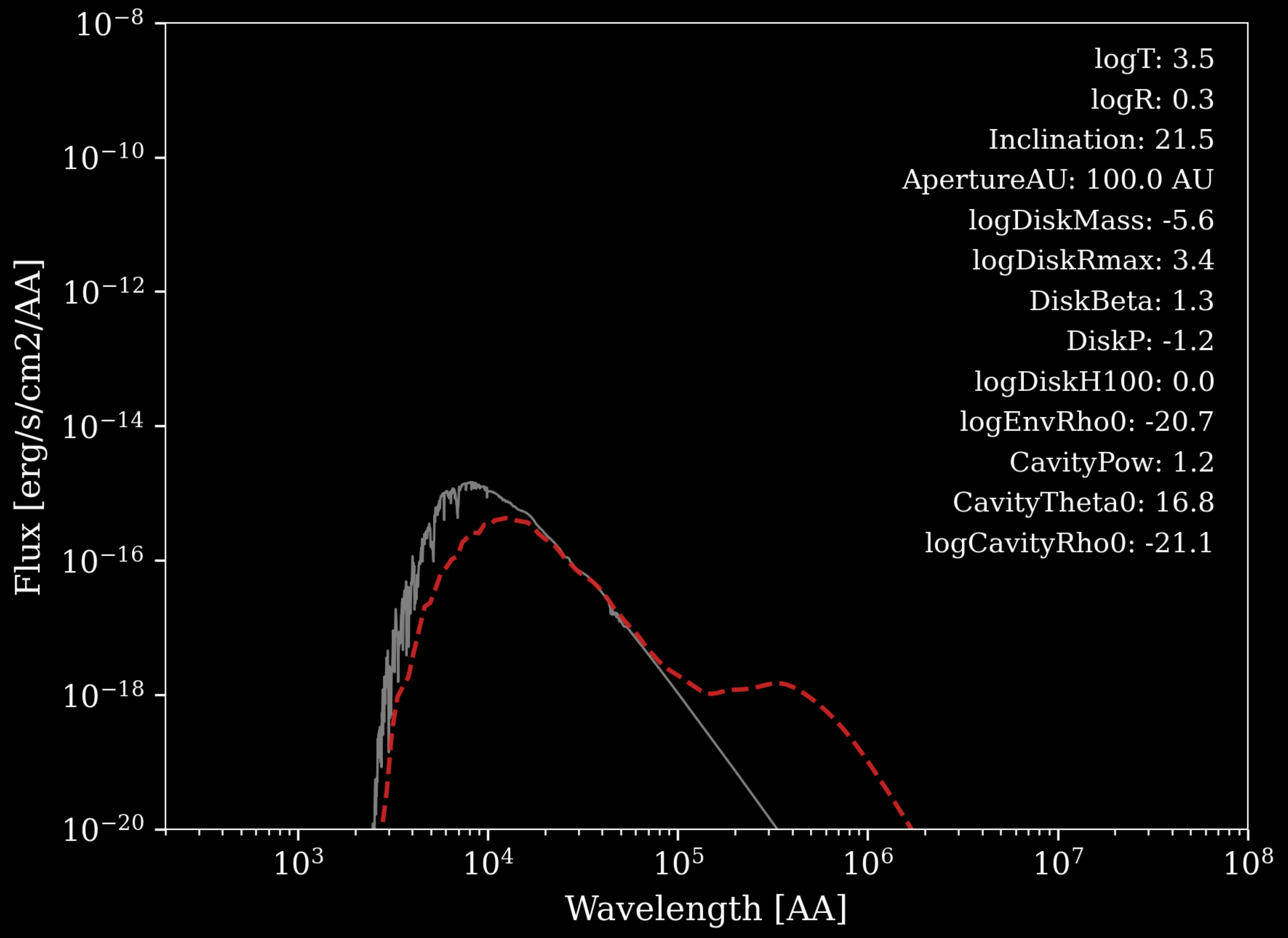


Fwd model: YSO (Robitaille17+Richardson+24)



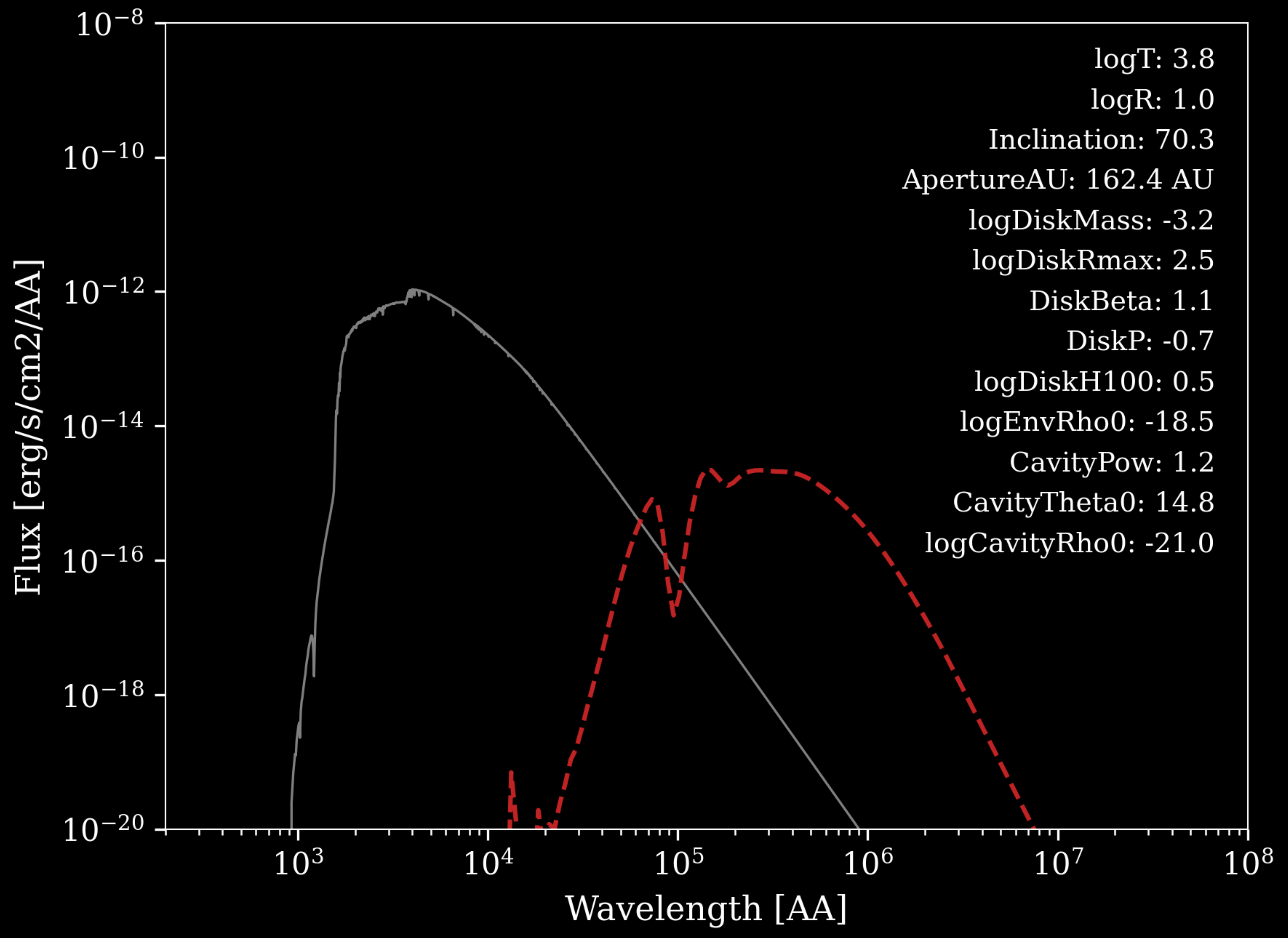
Parameter	Symbol	Minimum	Maximum	Sampling
Stellar radius	$R_\star$	$0.1\,R_\odot$	$100\,R_\odot$	Log
Stellar temperature	$T_\star$	2000 K	30 000 K	Log
Disk mass [dust]	$M_{\rm disk}$	$10^{-8}\,M_\odot$	$0.1\,M_\odot$	Log
Disk inner radius	$R_{\rm min}^{\rm disk}$	$R_{\rm sub}$	$1000\,R_{\rm sub}$	Log
Disk outer radius	$R_{\rm max}^{\rm disk}$	50 AU	5000 AU	Log
Disk flaring power	β	1	1.3	Linear
Disk surface density power	p	−2	0	Linear
Disk scaleheight	$h_{100\rm AU}$	1 AU	20 AU	Log
Envelope density [dust]	$\rho_0^{\rm env}$	10^{-24} g/cm^3	10^{-16} g/cm^3	Log
Envelope density power	γ	−2	−1	Linear
Envelope centrifugal radius	$R_{\rm c}$	50 AU	5000 AU	Log
Cavity density [dust]	$\rho_0^{\rm cav}$	10^{-23} g/cm^3	10^{-20} g/cm^3	Log
Cavity opening angle	θ_0	0°	60°	Linear
Cavity power	c	1	2	Linear

Fwd model: YSO (Robitaille17+Richardson+24)



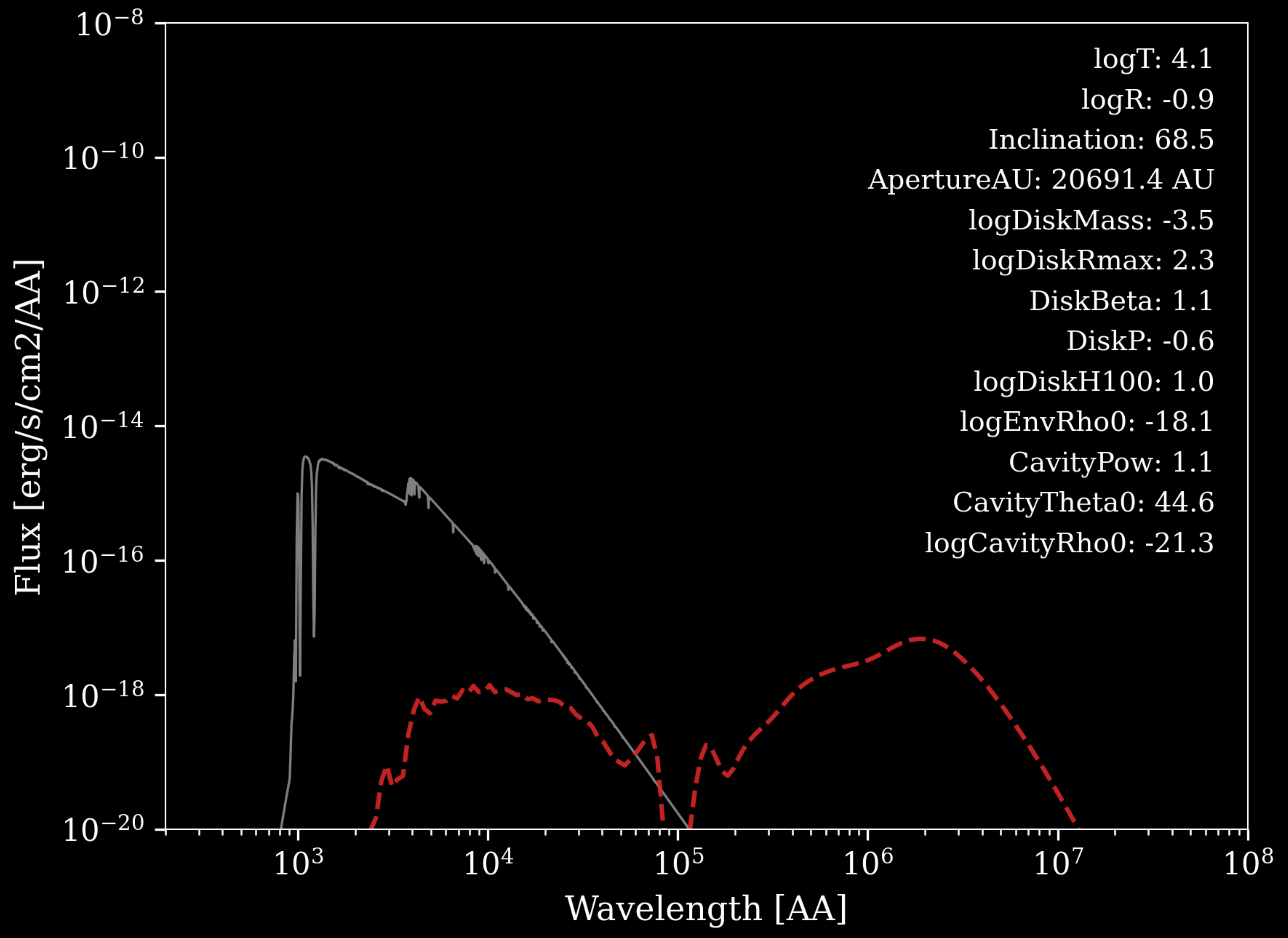
Parameter	Symbol	Minimum	Maximum	Sampling
Stellar radius	$R_\star$	$0.1\,R_\odot$	$100\,R_\odot$	Log
Stellar temperature	$T_\star$	2000 K	30 000 K	Log
Disk mass [dust]	$M_{\rm disk}$	$10^{-8}\,M_\odot$	$0.1\,M_\odot$	Log
Disk inner radius	$R_{\rm min}^{\rm disk}$	$R_{\rm sub}$	$1000\,R_{\rm sub}$	Log
Disk outer radius	$R_{\rm max}^{\rm disk}$	50 AU	5000 AU	Log
Disk flaring power	β	1	1.3	Linear
Disk surface density power	p	-2	0	Linear
Disk scaleheight	$h_{100\rm AU}$	1 AU	20 AU	Log
Envelope density [dust]	$\rho_0^{\rm env}$	10^{-24} g/cm^3	10^{-16} g/cm^3	Log
Envelope density power	γ	-2	-1	Linear
Envelope centrifugal radius	$R_{\rm c}$	50 AU	5000 AU	Log
Cavity density [dust]	$\rho_0^{\rm cav}$	10^{-23} g/cm^3	10^{-20} g/cm^3	Log
Cavity opening angle	θ_0	$0°$	$60°$	Linear
Cavity power	c	1	2	Linear

Fwd model: YSO (Robitaille17+Richardson+24)



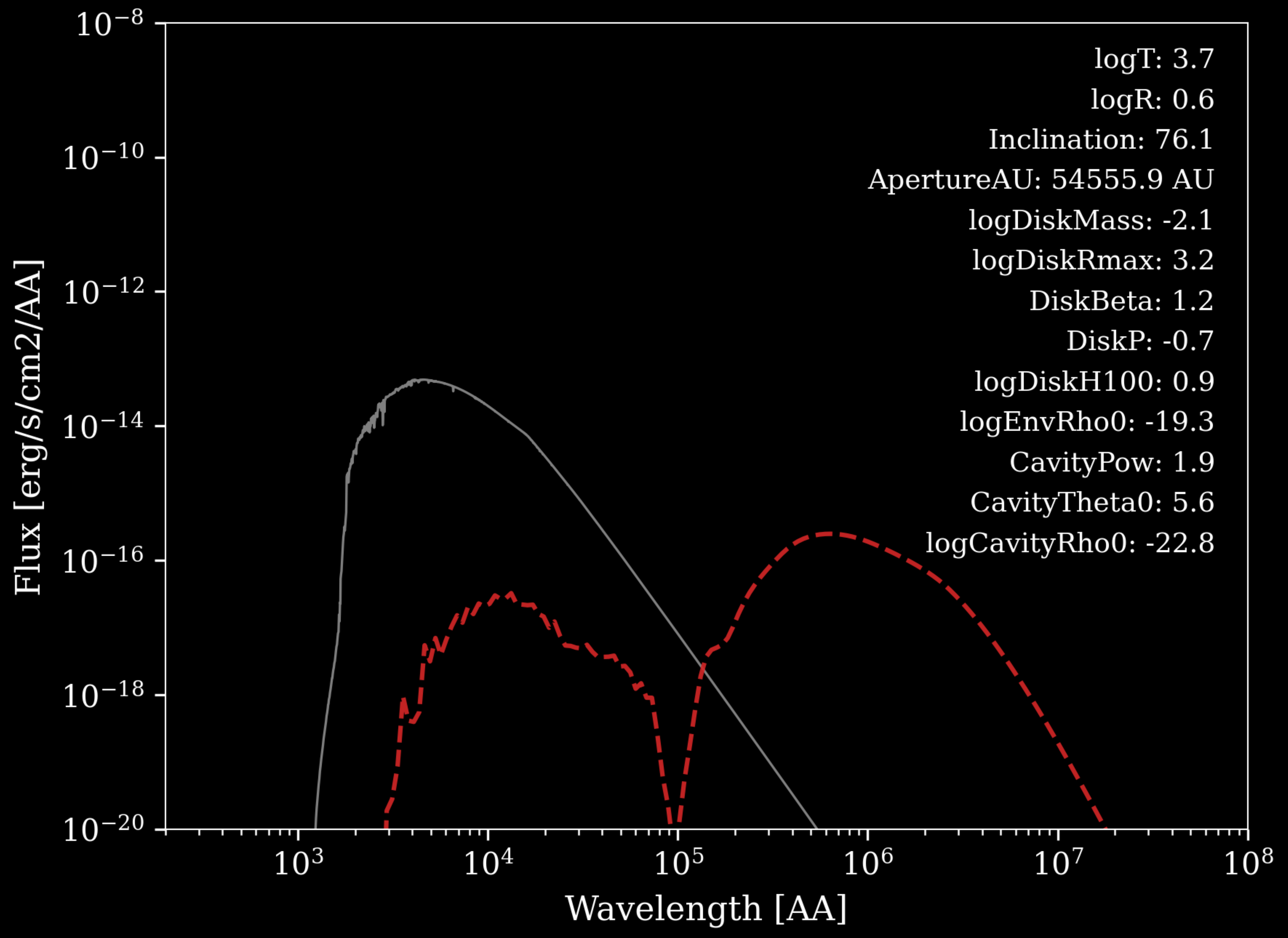
Parameter	Symbol	Minimum	Maximum	Sampling
Stellar radius	$R_\star$	$0.1\,R_\odot$	$100\,R_\odot$	Log
Stellar temperature	$T_\star$	2000 K	30 000 K	Log
Disk mass [dust]	$M_{\rm disk}$	$10^{-8}\,M_\odot$	$0.1\,M_\odot$	Log
Disk inner radius	$R^{\rm disk}_{\rm min}$	$R_{\rm sub}$	$1000\,R_{\rm sub}$	Log
Disk outer radius	$R^{\rm disk}_{\rm max}$	50 AU	5000 AU	Log
Disk flaring power	β	1	1.3	Linear
Disk surface density power	p	-2	0	Linear
Disk scaleheight	$h_{\rm 100AU}$	1 AU	20 AU	Log
Envelope density [dust]	$\rho^{\rm env}_0$	10^{-24} g/cm^3	10^{-16} g/cm^3	Log
Envelope density power	γ	-2	-1	Linear
Envelope centrifugal radius	$R_{\rm c}$	50 AU	5000 AU	Log
Cavity density [dust]	$\rho^{\rm cav}_0$	10^{-23} g/cm^3	10^{-20} g/cm^3	Log
Cavity opening angle	θ_0	$0°$	$60°$	Linear
Cavity power	c	1	2	Linear

Fwd model: YSO (Robitaille17+Richardson+24)



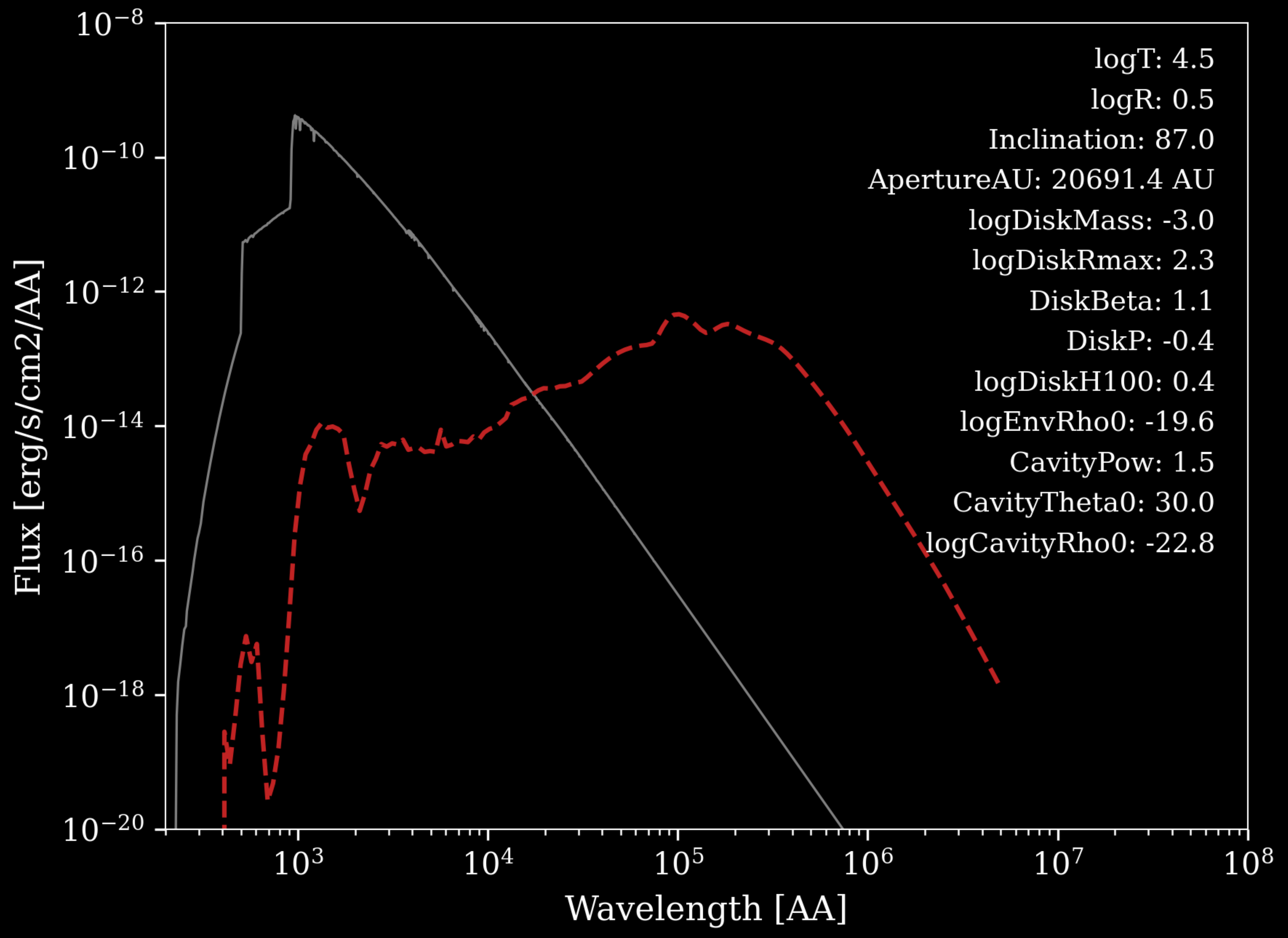
Parameter	Symbol	Minimum	Maximum	Sampling
Stellar radius	$R_\star$	$0.1\,R_\odot$	$100\,R_\odot$	Log
Stellar temperature	$T_\star$	2000 K	30 000 K	Log
Disk mass [dust]	$M_{\rm disk}$	$10^{-8}\,M_\odot$	$0.1\,M_\odot$	Log
Disk inner radius	$R^{\rm disk}_{\rm min}$	$R_{\rm sub}$	$1000\,R_{\rm sub}$	Log
Disk outer radius	$R^{\rm disk}_{\rm max}$	50 AU	5000 AU	Log
Disk flaring power	β	1	1.3	Linear
Disk surface density power	p	−2	0	Linear
Disk scaleheight	$h_{100\rm AU}$	1 AU	20 AU	Log
Envelope density [dust]	$\rho_0^{\rm env}$	10^{-24} g/cm^3	10^{-16} g/cm^3	Log
Envelope density power	γ	−2	−1	Linear
Envelope centrifugal radius	R_c	50 AU	5000 AU	Log
Cavity density [dust]	$\rho_0^{\rm cav}$	10^{-23} g/cm^3	10^{-20} g/cm^3	Log
Cavity opening angle	θ_0	0°	60°	Linear
Cavity power	c	1	2	Linear

Fwd model: YSO (Robitaille17+Richardson+24)



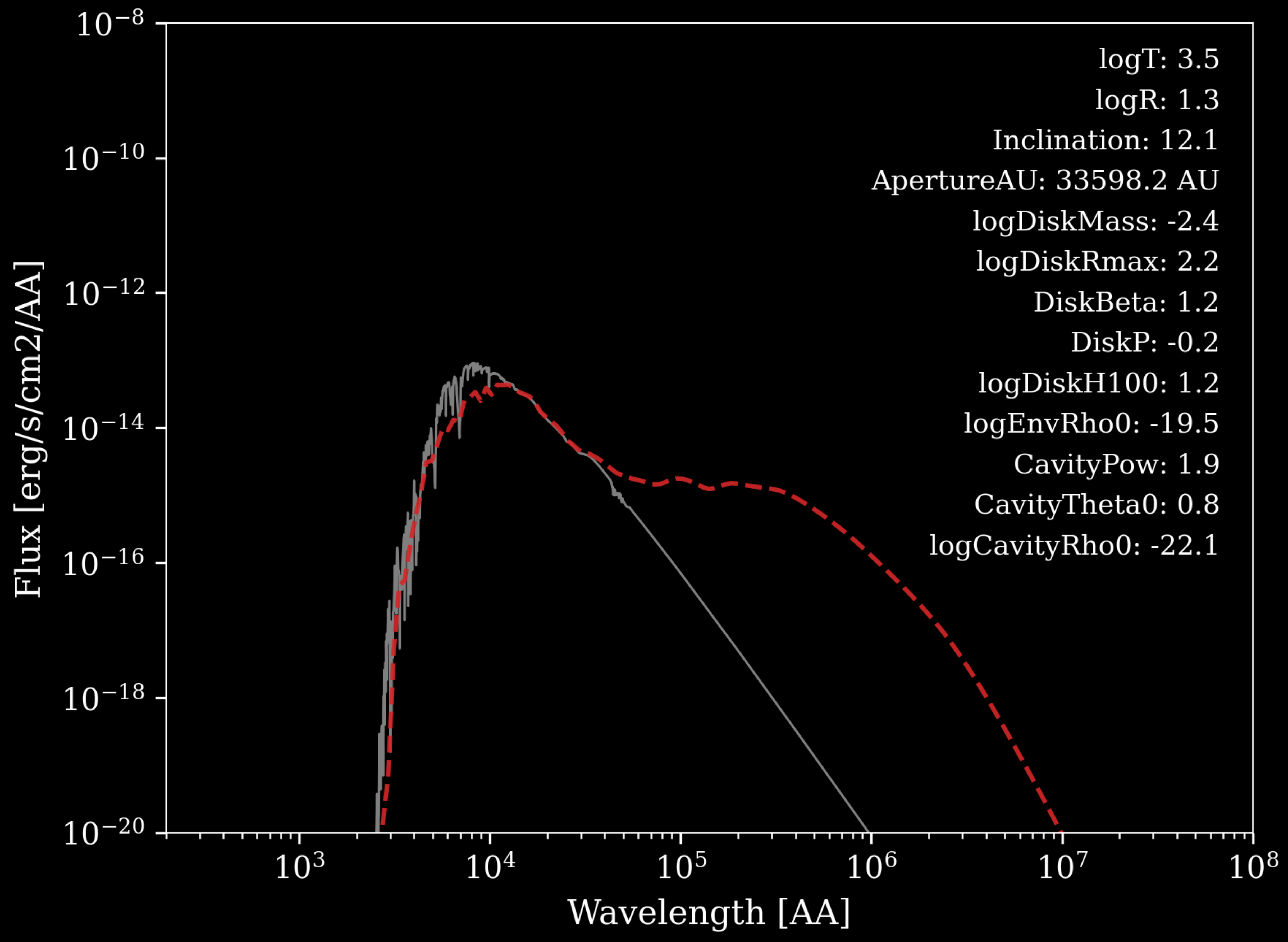
Parameter	Symbol	Minimum	Maximum	Sampling
Stellar radius	$R_\star$	$0.1\,R_\odot$	$100\,R_\odot$	Log
Stellar temperature	$T_\star$	2000 K	30 000 K	Log
Disk mass [dust]	$M_{\rm disk}$	$10^{-8}\,M_\odot$	$0.1\,M_\odot$	Log
Disk inner radius	$R^{\rm disk}_{\rm min}$	$R_{\rm sub}$	$1000\,R_{\rm sub}$	Log
Disk outer radius	$R^{\rm disk}_{\rm max}$	50 AU	5000 AU	Log
Disk flaring power	β	1	1.3	Linear
Disk surface density power	p	−2	0	Linear
Disk scaleheight	$h_{100\rm AU}$	1 AU	20 AU	Log
Envelope density [dust]	$\rho_0^{\rm env}$	10^{-24} g/cm^3	10^{-16} g/cm^3	Log
Envelope density power	γ	−2	−1	Linear
Envelope centrifugal radius	$R_{\rm c}$	50 AU	5000 AU	Log
Cavity density [dust]	$\rho_0^{\rm cav}$	10^{-23} g/cm^3	10^{-20} g/cm^3	Log
Cavity opening angle	θ_0	0°	60°	Linear
Cavity power	c	1	2	Linear

Fwd model: YSO (Robitaille17+Richardson+24)



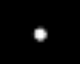
Parameter	Symbol	Minimum	Maximum	Sampling
Stellar radius	$R_\star$	$0.1\,R_\odot$	$100\,R_\odot$	Log
Stellar temperature	$T_\star$	2000 K	30 000 K	Log
Disk mass [dust]	$M_{\rm disk}$	$10^{-8}\,M_\odot$	$0.1\,M_\odot$	Log
Disk inner radius	$R^{\rm disk}_{\rm min}$	$R_{\rm sub}$	$1000\,R_{\rm sub}$	Log
Disk outer radius	$R^{\rm disk}_{\rm max}$	50 AU	5000 AU	Log
Disk flaring power	β	1	1.3	Linear
Disk surface density power	p	−2	0	Linear
Disk scaleheight	$h_{\rm 100AU}$	1 AU	20 AU	Log
Envelope density [dust]	$\rho^{\rm env}_0$	10^{-24} g/cm^3	10^{-16} g/cm^3	Log
Envelope density power	γ	−2	−1	Linear
Envelope centrifugal radius	R_c	50 AU	5000 AU	Log
Cavity density [dust]	$\rho^{\rm cav}_0$	10^{-23} g/cm^3	10^{-20} g/cm^3	Log
Cavity opening angle	θ_0	0°	60°	Linear
Cavity power	c	1	2	Linear

Fwd model: YSO (Robitaille17+Richardson+24)



Parameter	Symbol	Minimum	Maximum	Sampling
Stellar radius	$R_\star$	$0.1\,R_\odot$	$100\,R_\odot$	Log
Stellar temperature	$T_\star$	2000 K	30 000 K	Log
Disk mass [dust]	$M_{\rm disk}$	$10^{-8}\,M_\odot$	$0.1\,M_\odot$	Log
Disk inner radius	$R^{\rm disk}_{\rm min}$	$R_{\rm sub}$	$1000\,R_{\rm sub}$	Log
Disk outer radius	$R^{\rm disk}_{\rm max}$	50 AU	5000 AU	Log
Disk flaring power	β	1	1.3	Linear
Disk surface density power	p	-2	0	Linear
Disk scaleheight	$h_{100\rm AU}$	1 AU	20 AU	Log
Envelope density [dust]	$\rho_0^{\rm env}$	10^{-24} g/cm^3	10^{-16} g/cm^3	Log
Envelope density power	γ	-2	-1	Linear
Envelope centrifugal radius	$R_{\rm c}$	50 AU	5000 AU	Log
Cavity density [dust]	$\rho_0^{\rm cav}$	10^{-23} g/cm^3	10^{-20} g/cm^3	Log
Cavity opening angle	θ_0	0°	60°	Linear
Cavity power	c	1	2	Linear

YSO (Robitaille17+Richardson+24)

Model set	Icon	Star	Disk	Envelope	Cavity	Ambient	Inner radius	Variables	Models
s-s-i		yes	...	...	...	...	...	2	10 000
sp-s-i		yes	passive	...	...	...	R_{sub}	7	10 000
sp-h-i		yes	passive	...	...	...	variable	8	10 000
s-smi		yes	...	...	...	yes	R_{sub}	2	10 000
sp-smi		yes	passive	...	...	yes	R_{sub}	7	10 000
sp-hmi		yes	passive	...	...	yes	variable	8	10 000

sp-smi		yes	passive	...	...	yes	R_{sub}	7	10 000
sp-hmi		yes	passive	...	...	yes	variable	8	10 000
s-p-smi		yes	...	power-law	...	yes	R_{sub}	4	10 000
s-p-hmi		yes	...	power-law	...	yes	variable	5	10 000
s-pbsmi		yes	...	power-law	yes	yes	R_{sub}	7	10 000
s-pbhmi		yes	...	power-law	yes	yes	variable	8	10 000
s-u-smi		yes	...	Ulrich	...	yes	R_{sub}	4	10 000
s-u-hmi		yes	...	Ulrich	...	yes	variable	5	10 000
s-ubsmi		yes	...	Ulrich	yes	yes	R_{sub}	7	10 000
s-ubhmi		yes	...	Ulrich	yes	yes	variable	8	10 000

s-pbhmi		yes	...	power-law	yes	yes	variable	8	10 000
s-u-smi		yes	...	Ulrich	...	yes	R_{sub}	4	10 000
s-u-hmi		yes	...	Ulrich	...	yes	variable	5	10 000
s-ubsmi		yes	...	Ulrich	yes	yes	R_{sub}	7	10 000
s-ubhmi		yes	...	Ulrich	yes	yes	variable	8	10 000
spu-smi		yes	passive	Ulrich	...	yes	R_{sub}	8	10 000
spu-hmi		yes	passive	Ulrich	...	yes	variable	9	10 000
spubsmi		yes	passive	Ulrich	yes	yes	R_{sub}	11	40 000
spubhmi		yes	passive	Ulrich	yes	yes	variable	12	80 000

YSO (Robitaille17+Richardson+24)

sp-h-i		yes	passive	...	...	...	variable	8	10 000
s-smi		yes	...	...	...	yes	R_{sub}	2	10 000
sp-hmi		yes	passive	...	...	yes	variable	8	10 000
spubhmi		yes	passive	Ulrich	yes	yes	variable	12	80 000