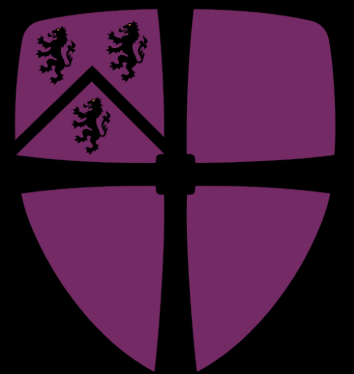


Probing dark matter substructure using gravitational microlensing

Djuna Lize Croon (IPPP Durham)

BNL remote seminar, June 2025

djuna.l.croon@durham.ac.uk | djunacroon.com



A key question for particle physics & cosmology

What on Earth is dark matter?

- Plenty of evidence!
- Likely not explained by modifications of gravity



<https://youtu.be/IEKcPfJHfNU> (see also djunacroon.com)

A key question for particle physics & cosmology

What on Earth is dark matter?

- Plenty of evidence!
- Likely not explained by modifications of gravity



<https://youtu.be/IEKcPfJHfNU> (see also djunacroon.com)

No idea, or too many ideas, about its nature:



Dark matter substructure

Two things we may agree upon...

- All of our evidence for dark matter is gravitational
- Many dark matter models feature substructure

PBHs

Boson stars

Subhalos

Miniclusters

Mirror stars



I call objects like these EDOs
(extended dark objects)

Dark matter substructure

Two things we may agree upon...

- All of our evidence for dark matter is gravitational
- Many dark matter models feature substructure

PBHs

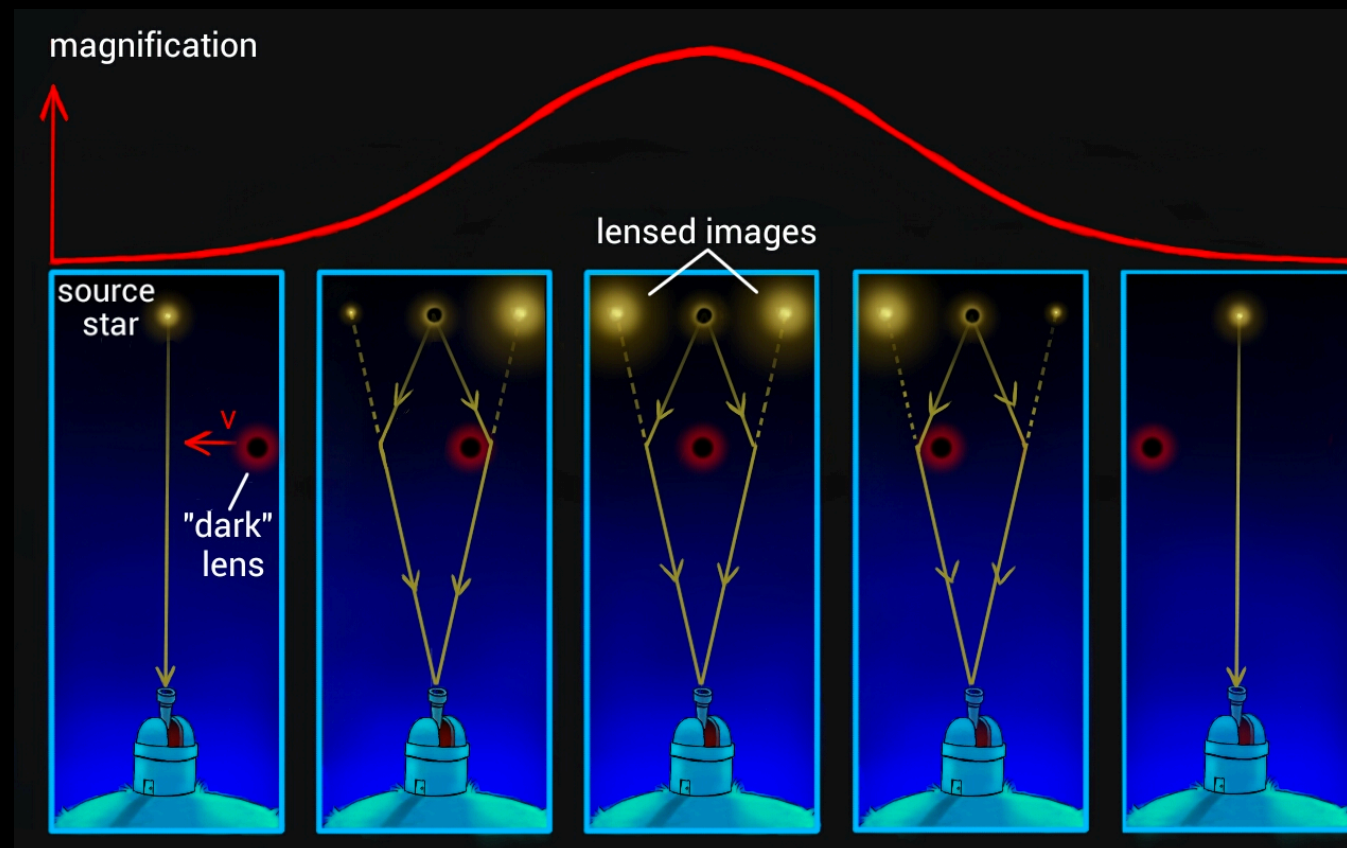
Boson stars

Subhalos

Miniclusters

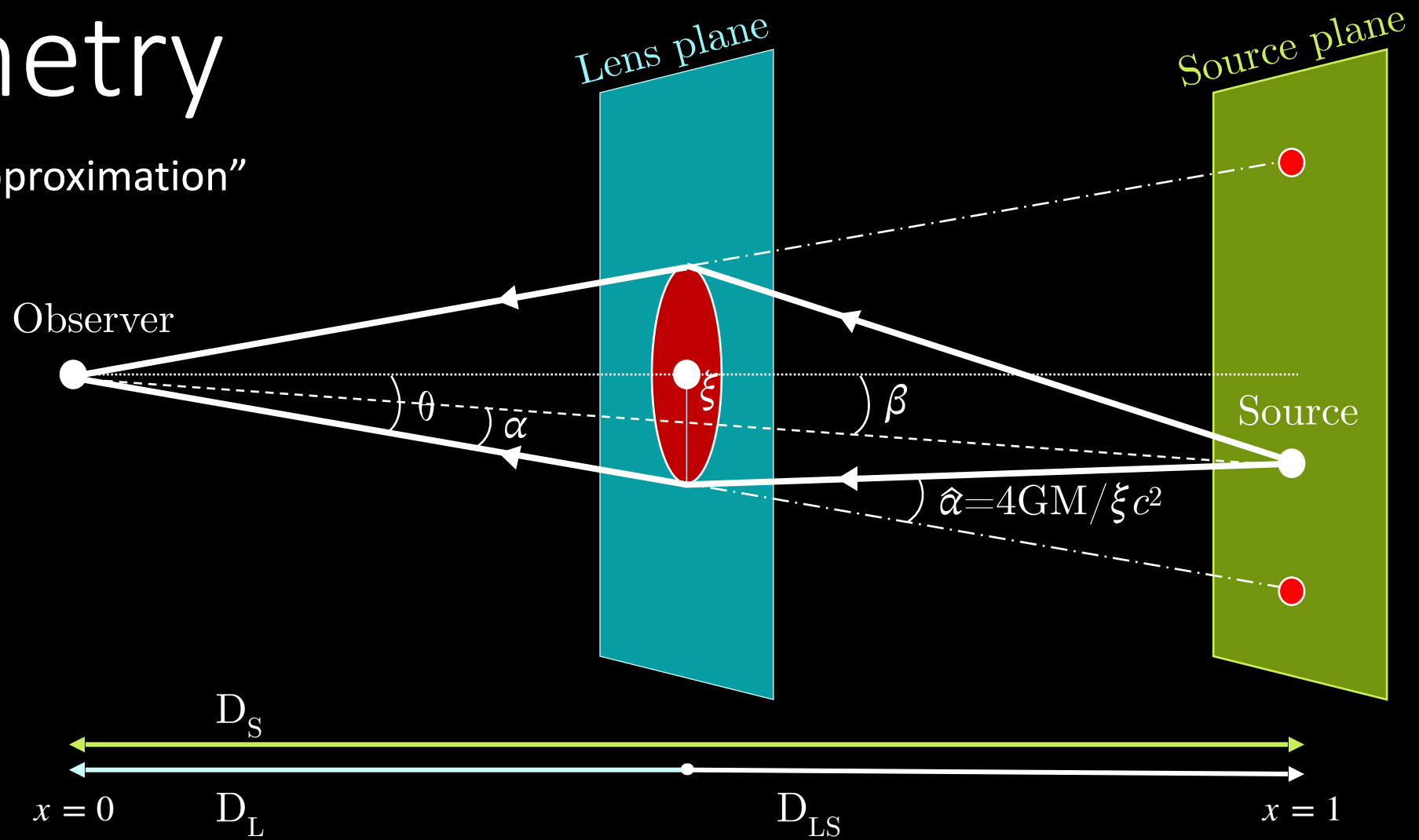
Mirror stars

- Microlensing can be used to probe such models



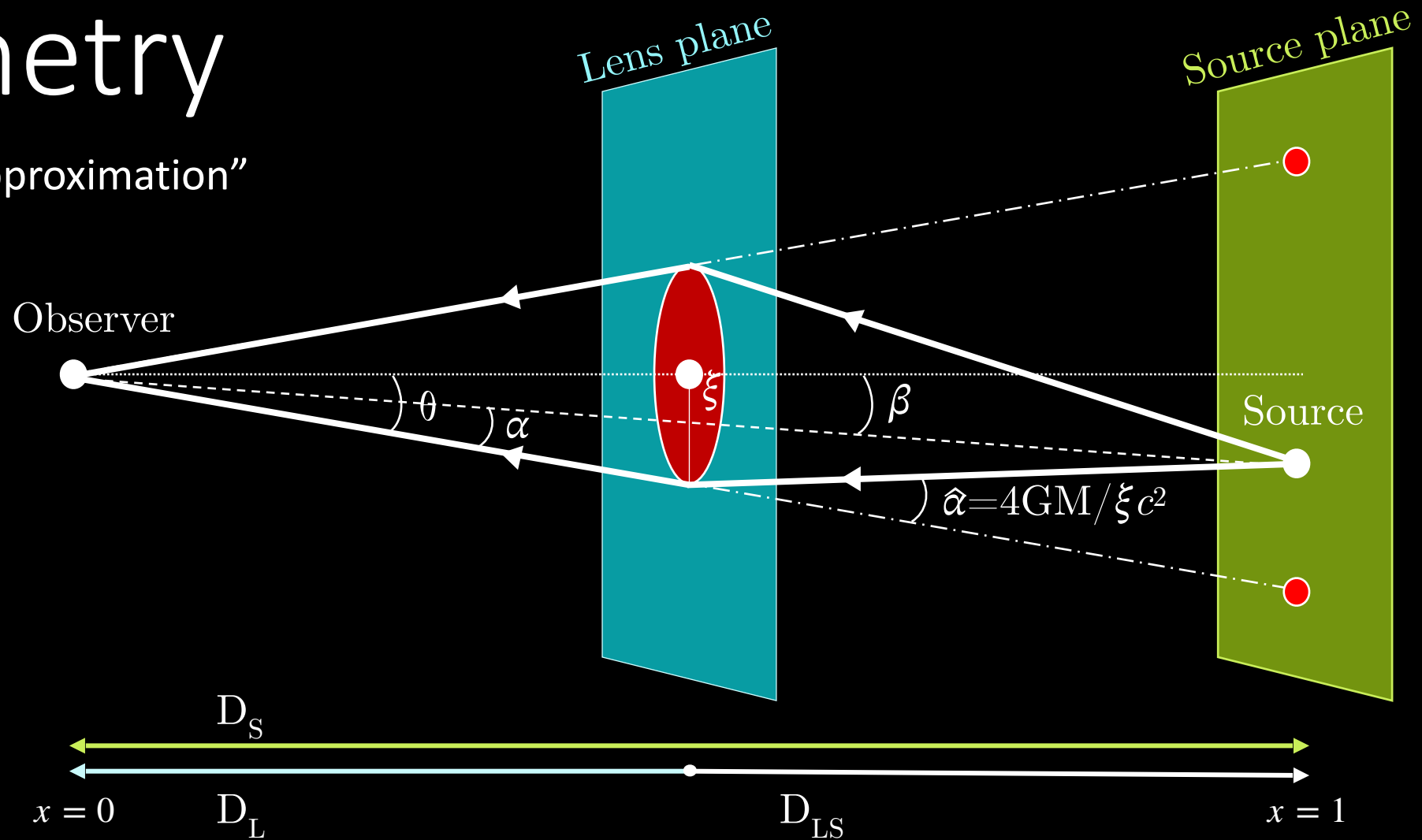
Geometry

“Thin screen approximation”



Geometry

"Thin screen approximation"

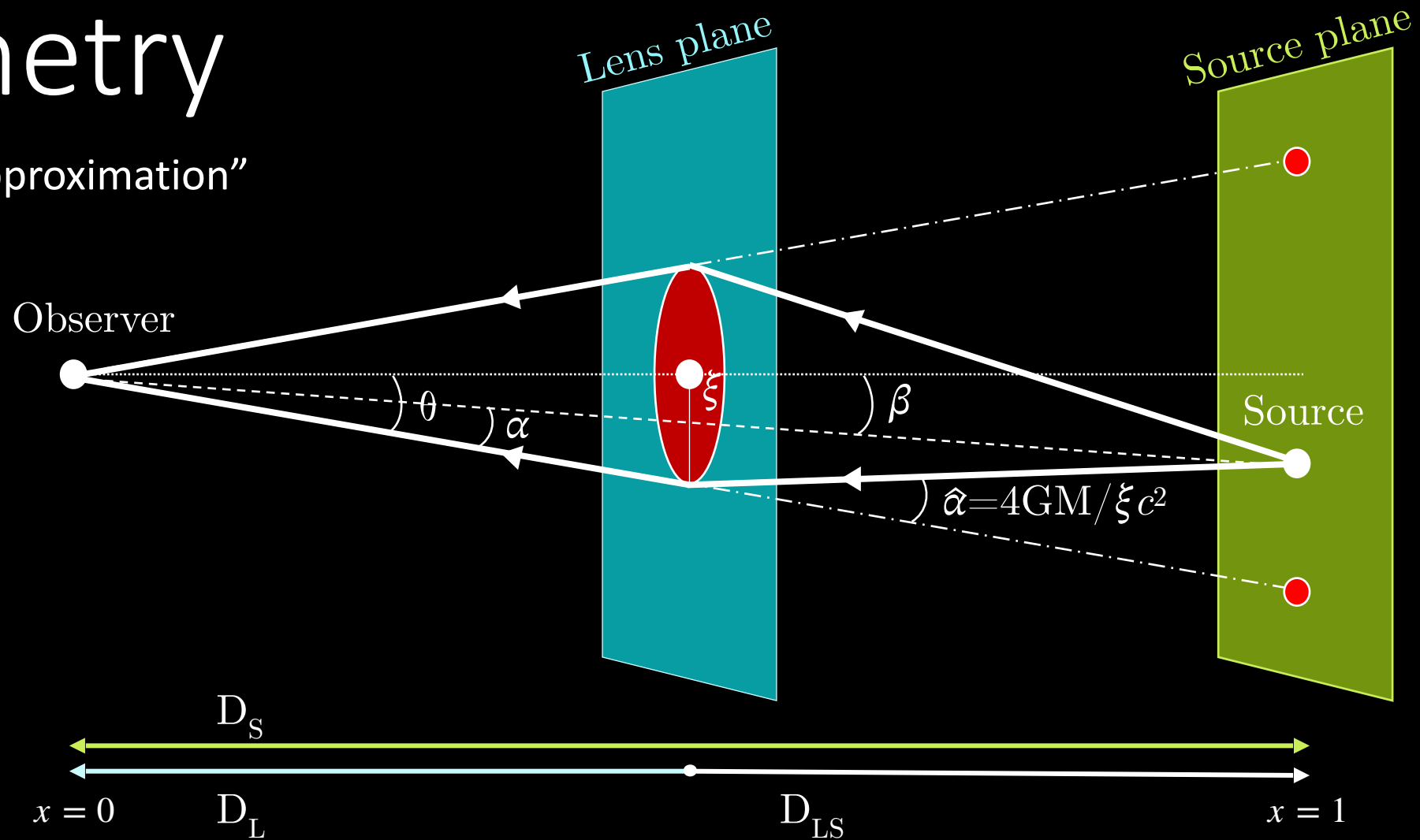


Lensing equation:
$$\beta = \theta - \frac{4GM(\theta)}{\theta D_L c^2} \frac{D_{LS}}{D_S}$$

Magnification:
$$\mu = \frac{\theta}{\beta} \frac{d\theta}{d\beta} = \sum_i \mu_i$$

Geometry

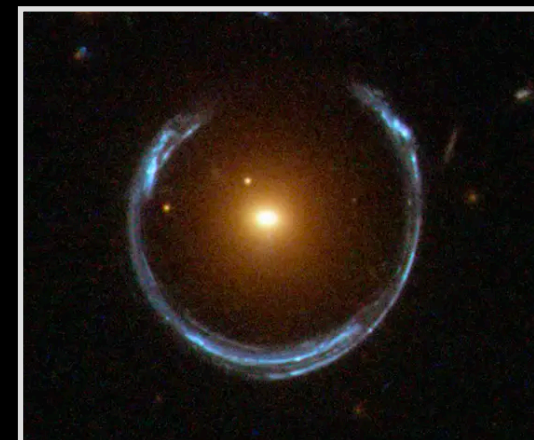
"Thin screen approximation"



Lensing equation:
$$\beta = \theta - \frac{4GM(\theta)}{\theta D_L c^2} \frac{D_{LS}}{D_S}$$

$$\beta = 0 \rightarrow \theta \equiv \theta_E = \sqrt{\frac{4GM}{c^2} \frac{D_{LS}}{D_L D_S}}$$

Magnification:
$$\mu = \frac{\theta}{\beta} \frac{d\theta}{d\beta} = \sum_i \mu_i$$

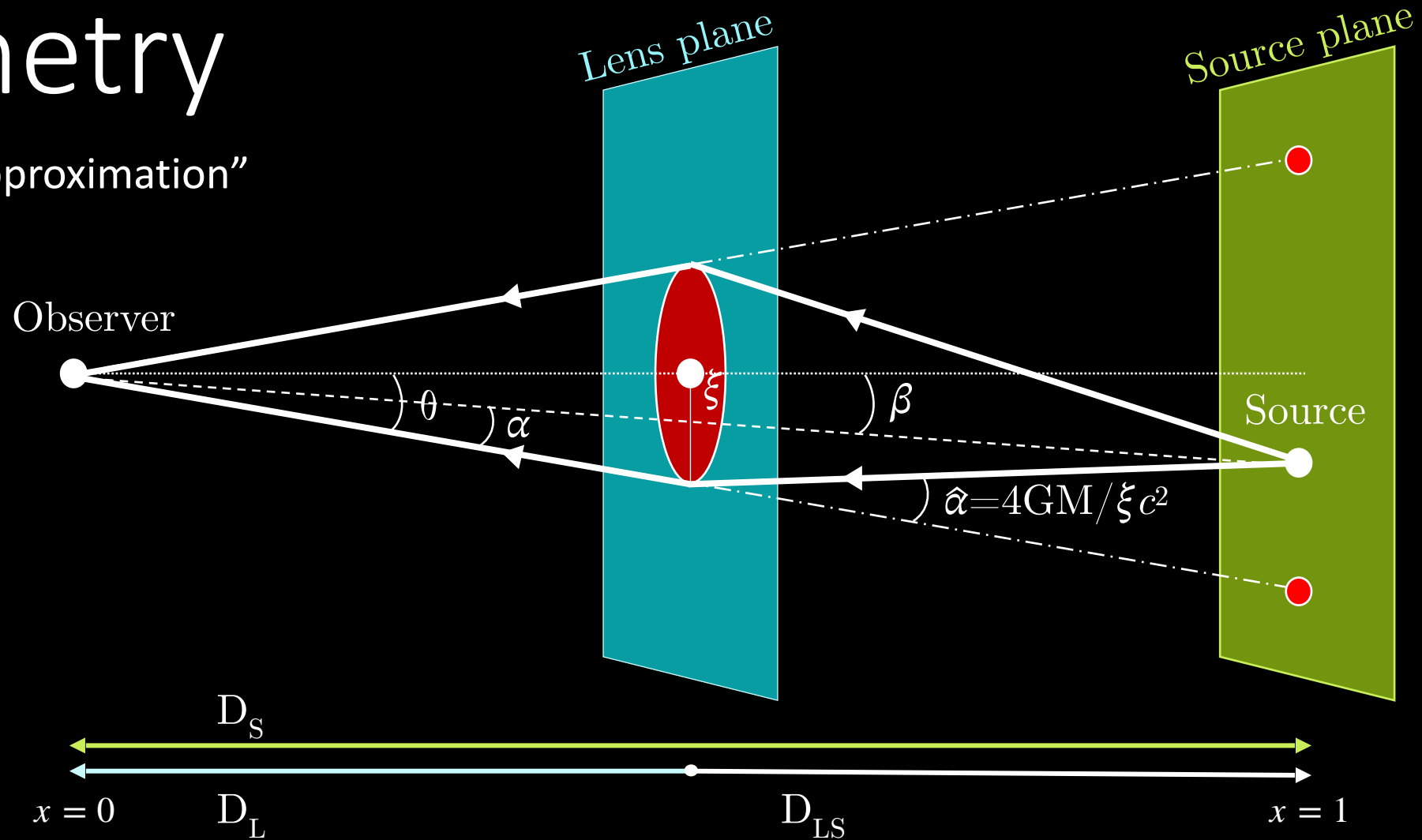


Einstein radius
 $r_E = \theta_E D_L$

Near perfect Einstein Ring with the HST

Geometry

"Thin screen approximation"



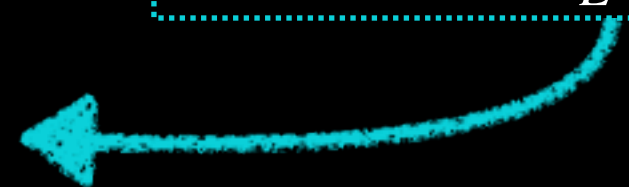
Lensing equation: $u = \tau - \frac{m(\tau)}{\tau}$



Normalise everything to θ_E

- $u \equiv \beta/\theta_E$
- $\tau \equiv \theta/\theta_E$
- $m(\tau) \equiv M(\theta_E \tau)/M$

Magnification: $\mu = \left| 1 - \frac{m(\tau)}{\tau^2} \right|^{-1} \left| 1 + \frac{m(\tau)}{\tau^2} - \frac{1}{\tau} \frac{dm(\tau)}{d\tau} \right|^{-1}$



Geometry

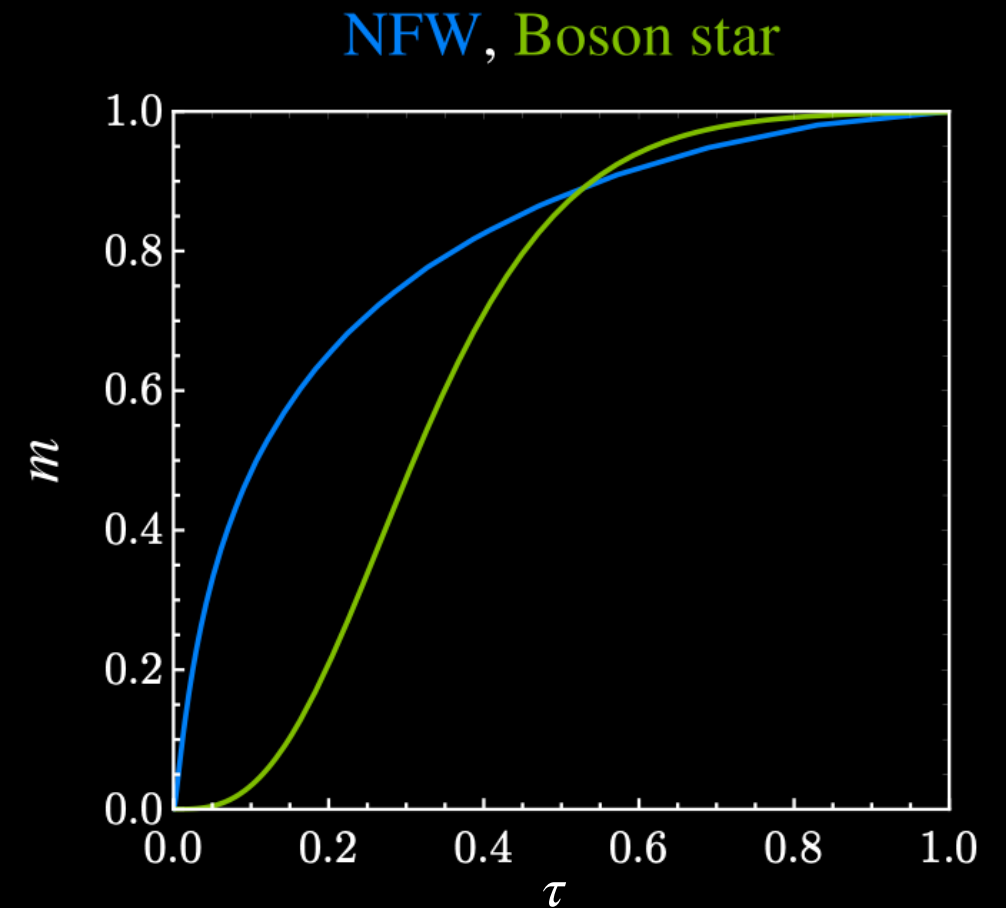
“Thin screen approximation”

$$m(\tau) = \frac{\int_0^\tau d\sigma \sigma \int_0^\infty d\lambda \rho(r_E \sqrt{\sigma^2 + \lambda^2})}{\int_0^\infty d\gamma \gamma^2 \rho(r_E \gamma)}$$

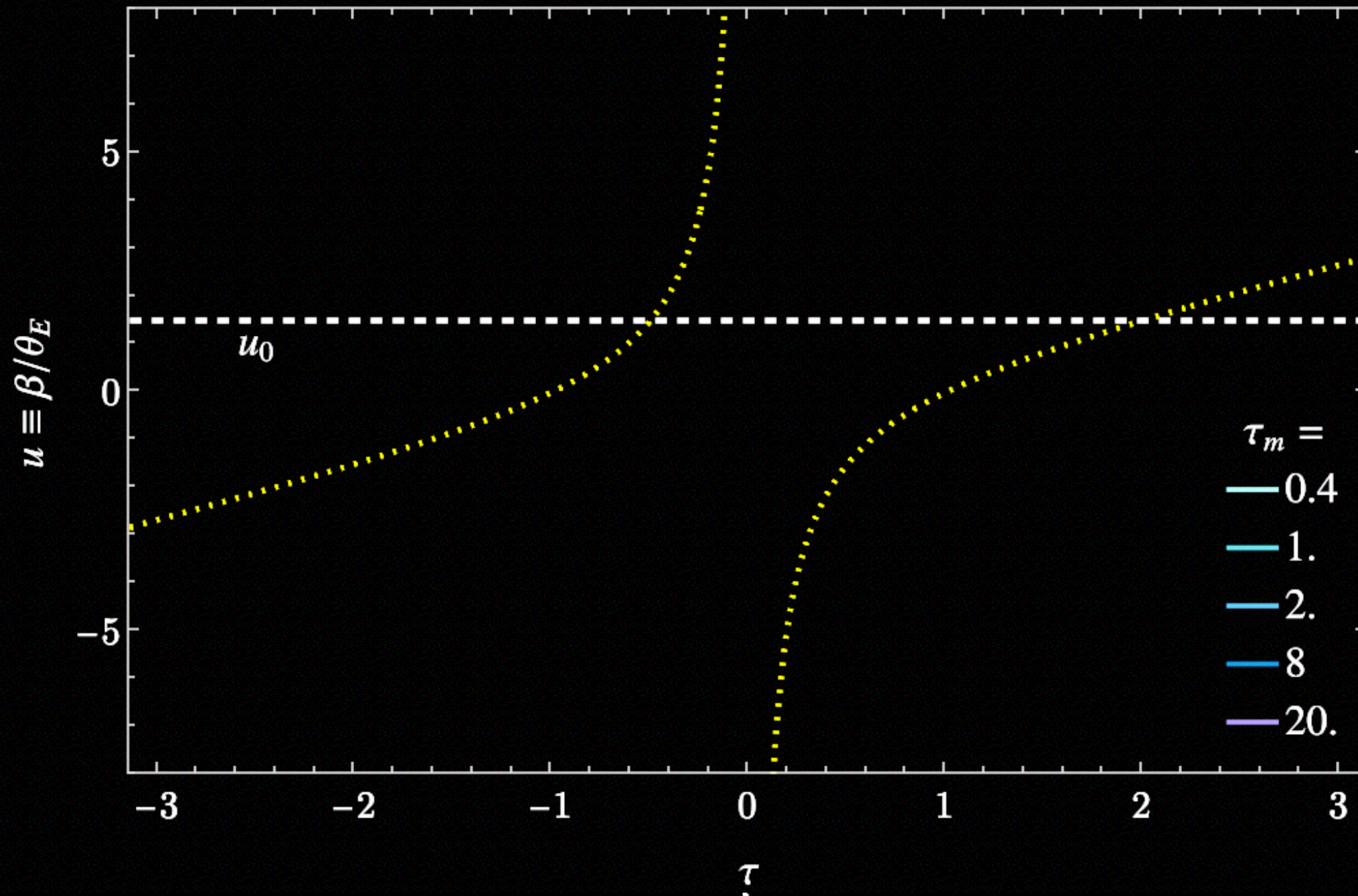
Projected lens mass distribution
Point-like lenses: $m(\tau) = 1$

Lensing equation: $u = \tau - \frac{m(\tau)}{\tau}$

$$\text{Magnification: } \mu = \left| 1 - \frac{m(\tau)}{\tau^2} \right|^{-1} \left| 1 + \frac{m(\tau)}{\tau^2} - \frac{1}{\tau} \frac{dm(\tau)}{d\tau} \right|^{-1}$$

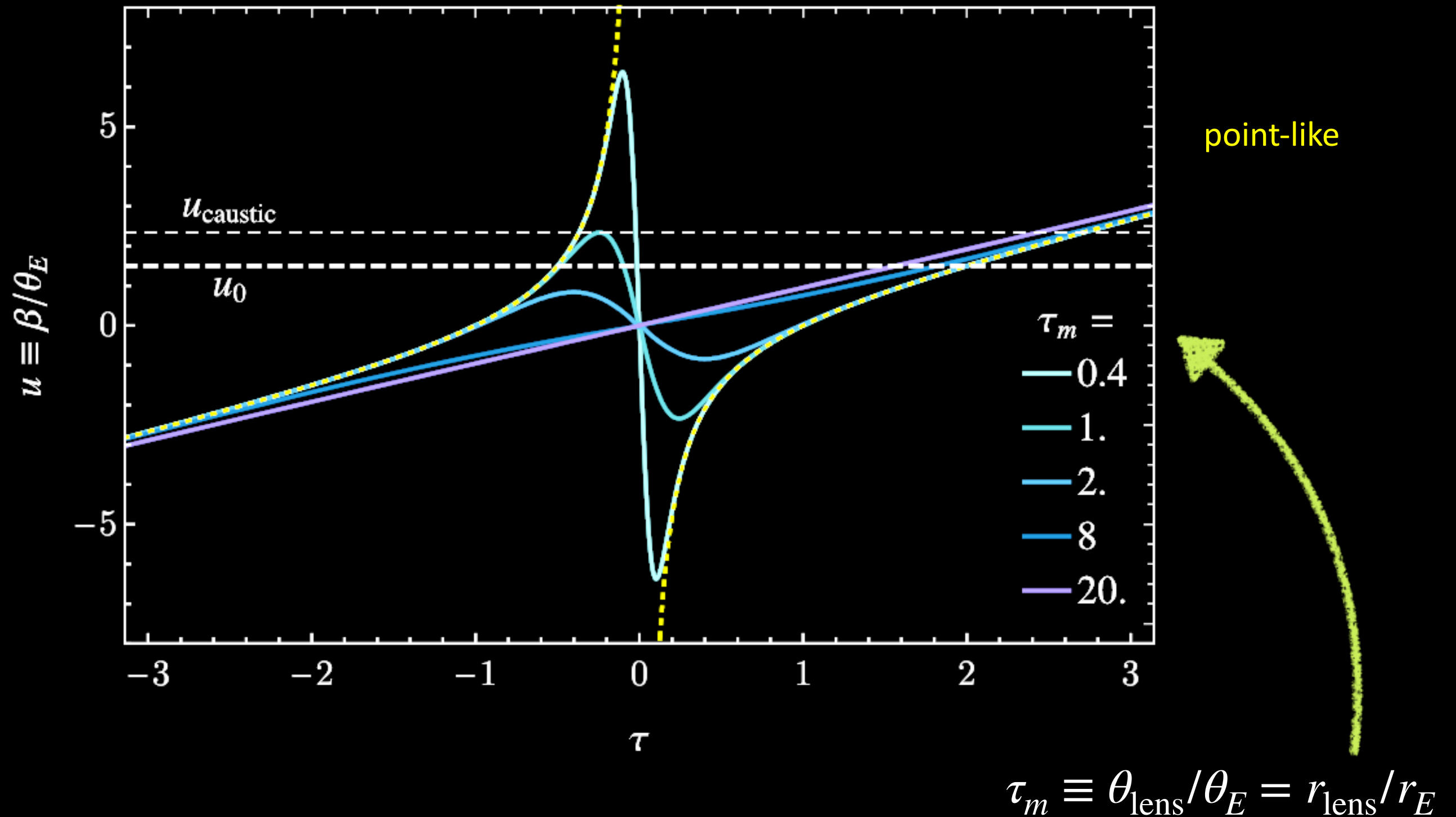


Solving the lens equation $u = \tau - \frac{m(\tau)}{\tau}$

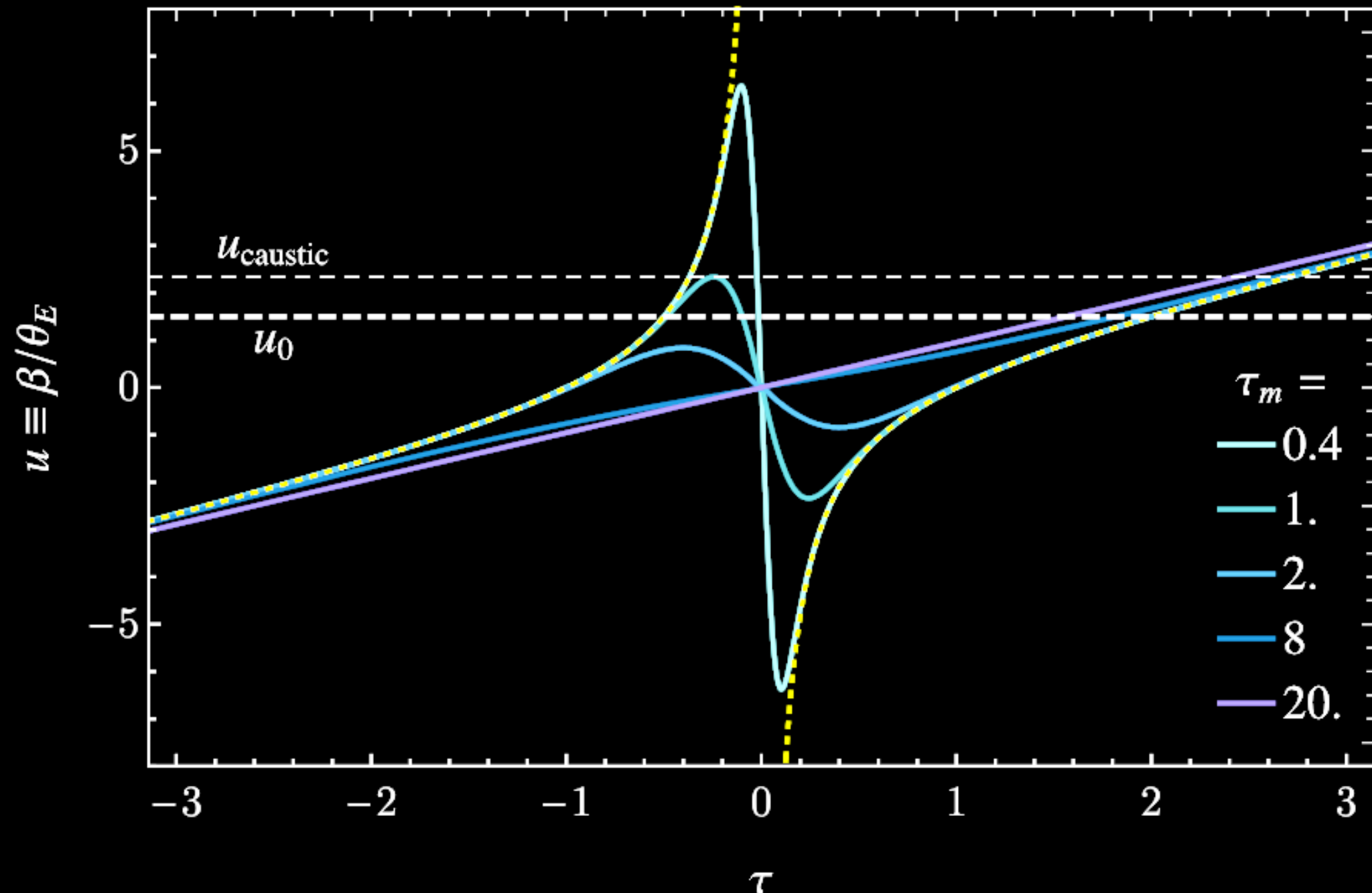


point-like

Solving the lens equation $u = \tau - \frac{m(\tau)}{\tau}$



Solving the lens equation $u = \tau - \frac{m(\tau)}{\tau}$



point-like

3 solutions, one with negligible μ_i

3 solutions, all contributing μ_i

1 solution of the point-like case

1 solution, larger μ_i than point-like

1 solution, $\mu_i \rightarrow 0$

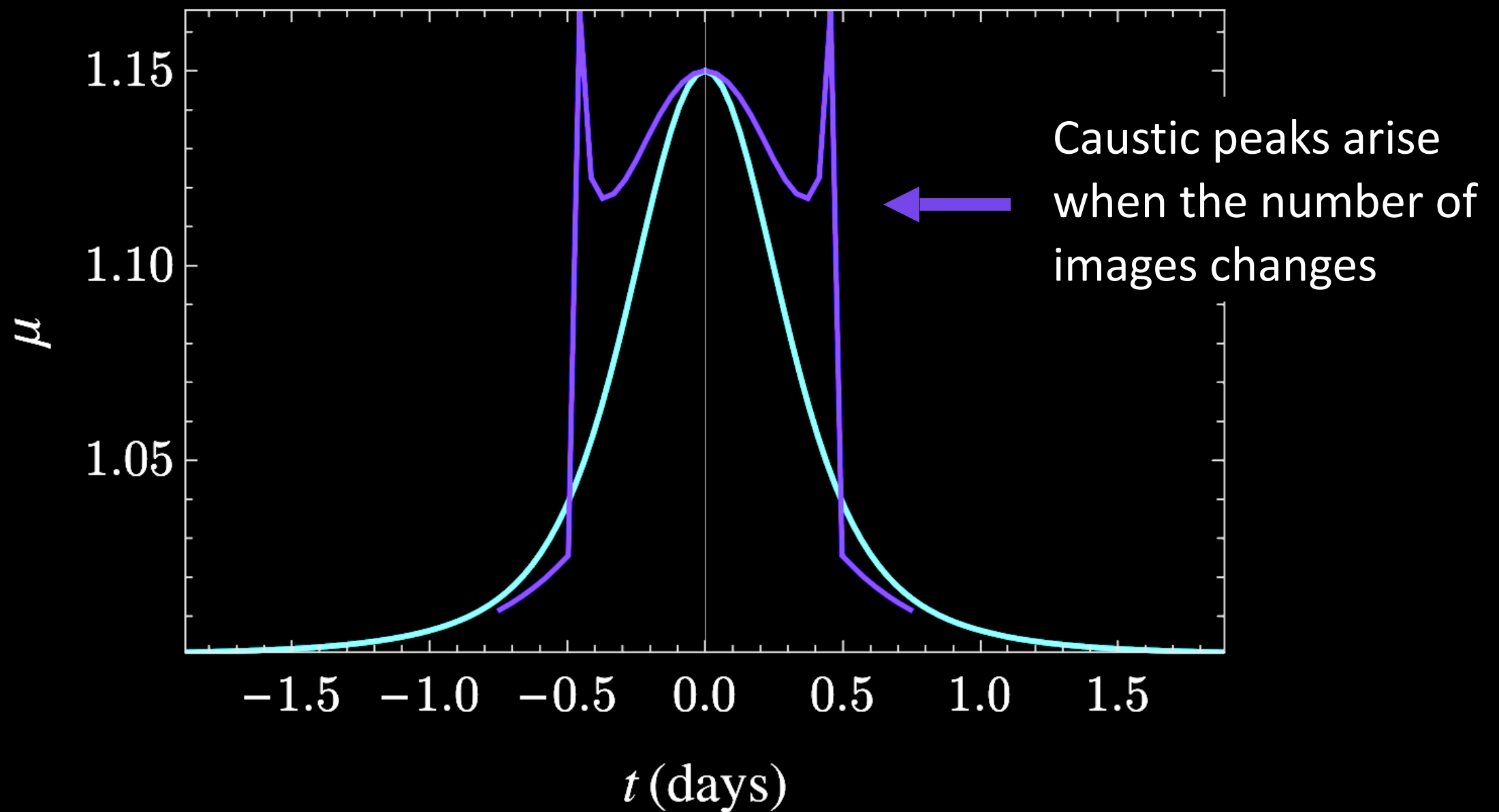
→ at u_{caustic} , number of solutions jumps from 1 to 3

Caustics

Example light curve

Boson star with $\tau_m = 1$

PBH (or $\tau_m = 0$)



Point-like lens: $m(\tau) = 1 \rightarrow \mu = \frac{u^2 + 2}{u\sqrt{u^2 + 4}}$

“Efficiency” of extended lenses

Define $u_{1.34}$ by $\mu_{\text{tot}}(u \leq u_{1.34}) > 1.34$

The threshold impact parameter

All smaller impact parameters produce
a magnification above $\mu > 1.34$

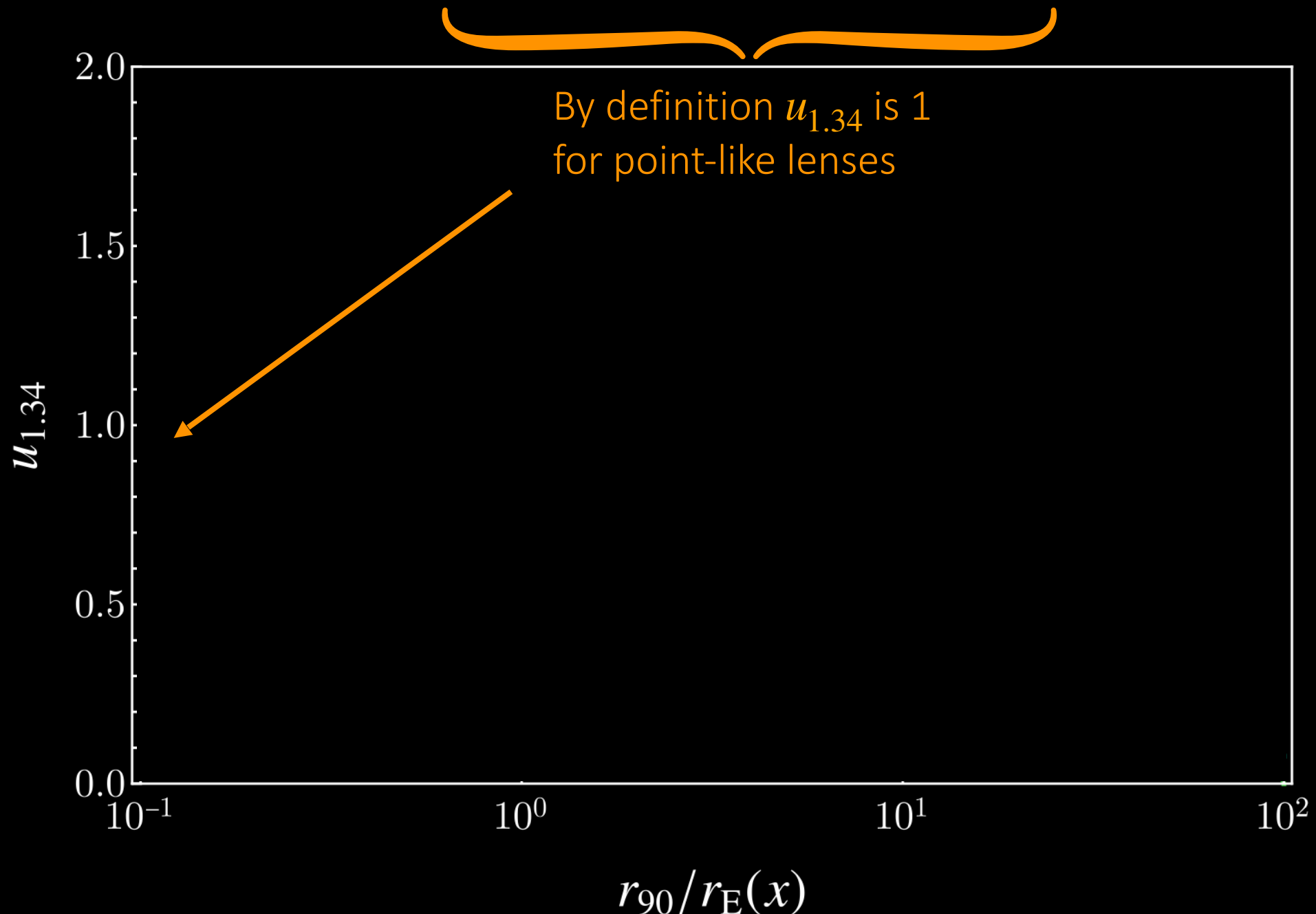
By definition $u_{1.34}$ is 1
for point-like lenses

“Efficiency” of extended lenses

Define $u_{1.34}$ by $\mu_{\text{tot}}(u \leq u_{1.34}) > 1.34$

The threshold impact parameter

All smaller impact parameters produce
a magnification above $\mu > 1.34$

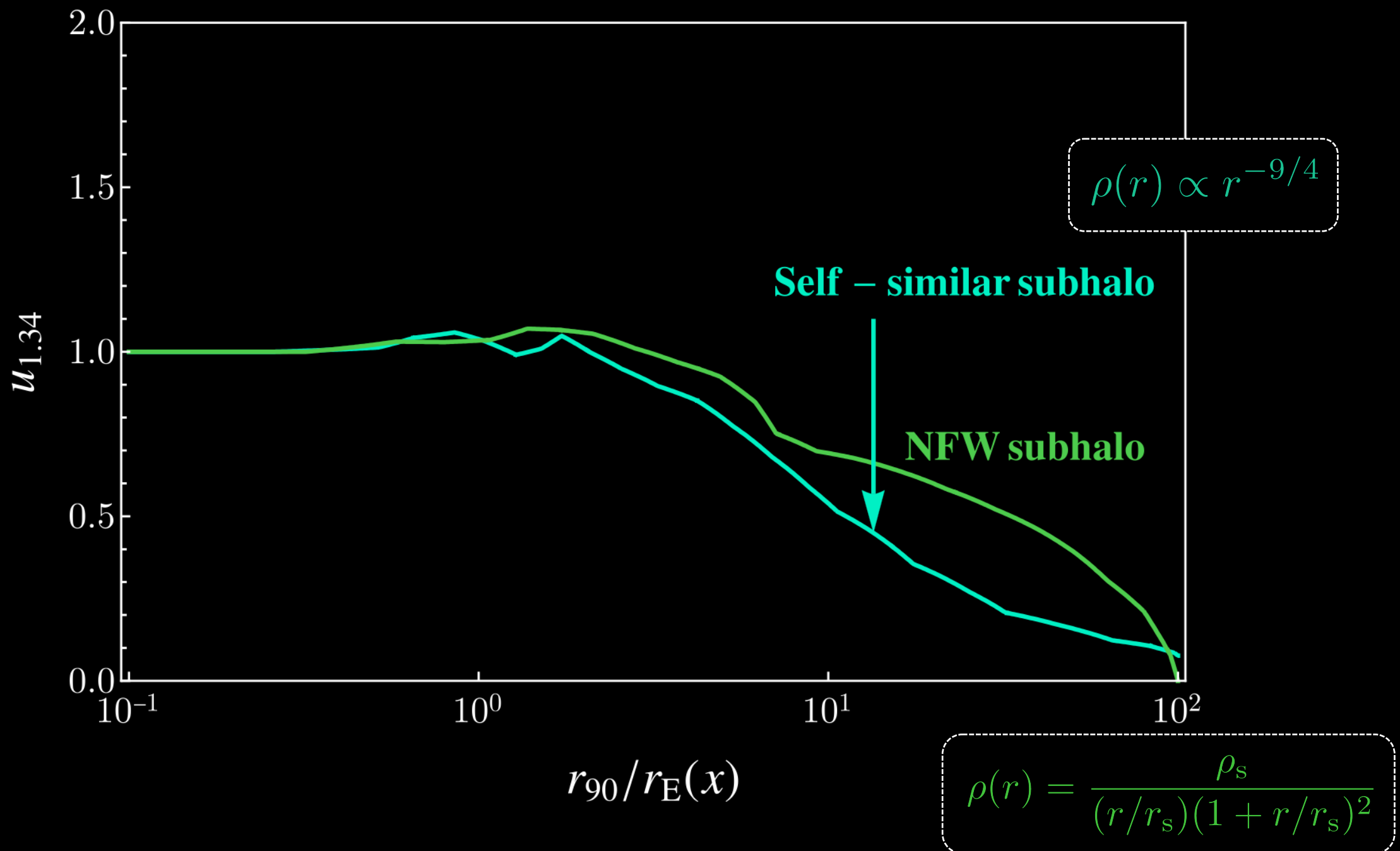


“Efficiency” of extended lenses

Define $u_{1.34}$ by $\mu_{\text{tot}}(u \leq u_{1.34}) > 1.34$

The threshold impact parameter

All smaller impact parameters produce
a magnification above $\mu > 1.34$



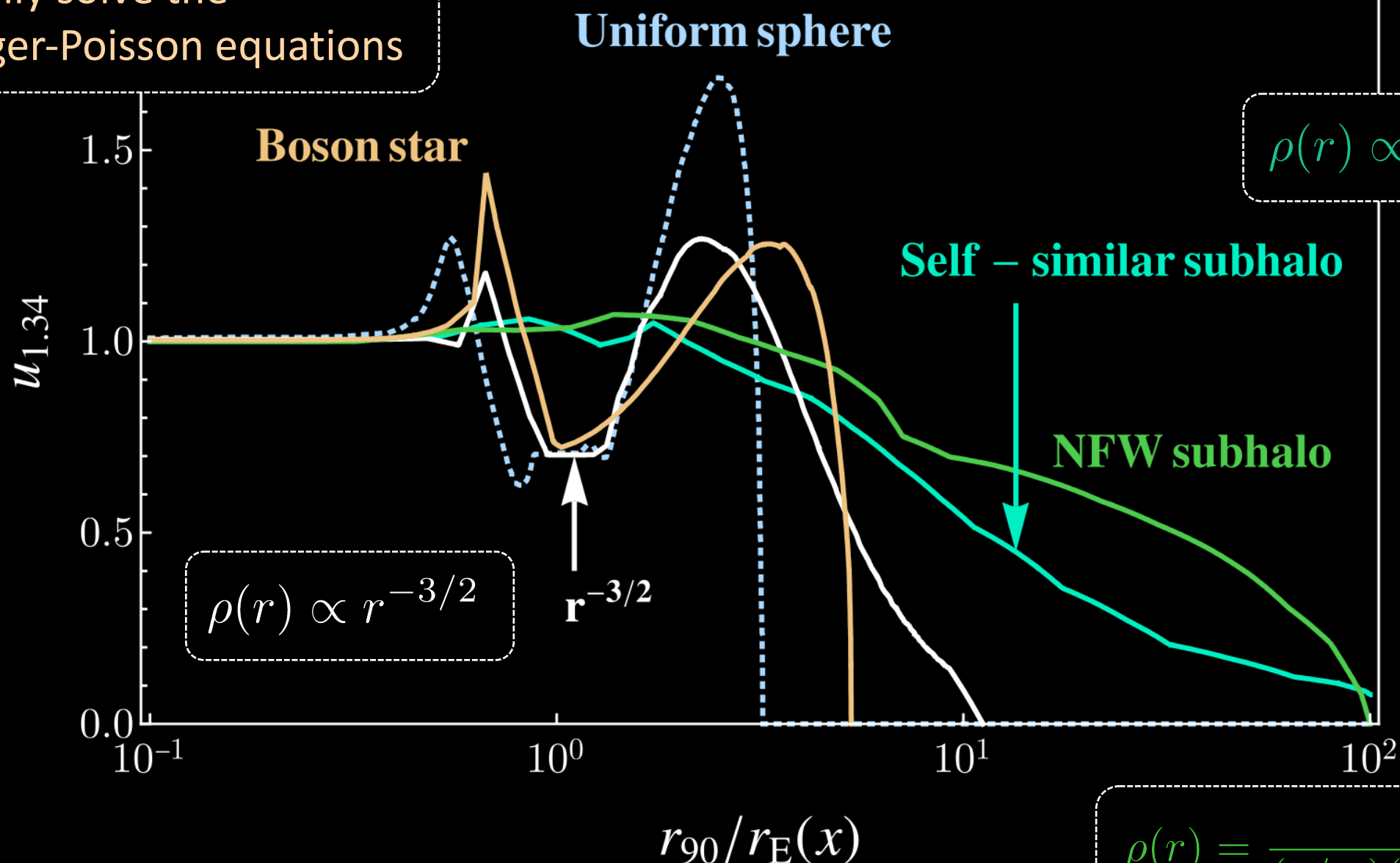
“Efficiency” of extended lenses

Define $u_{1.34}$ by $\mu_{\text{tot}}(u \leq u_{1.34}) > 1.34$

The threshold impact parameter

All smaller impact parameters produce a magnification above $\mu > 1.34$

Numerically solve the Schrodinger-Poisson equations



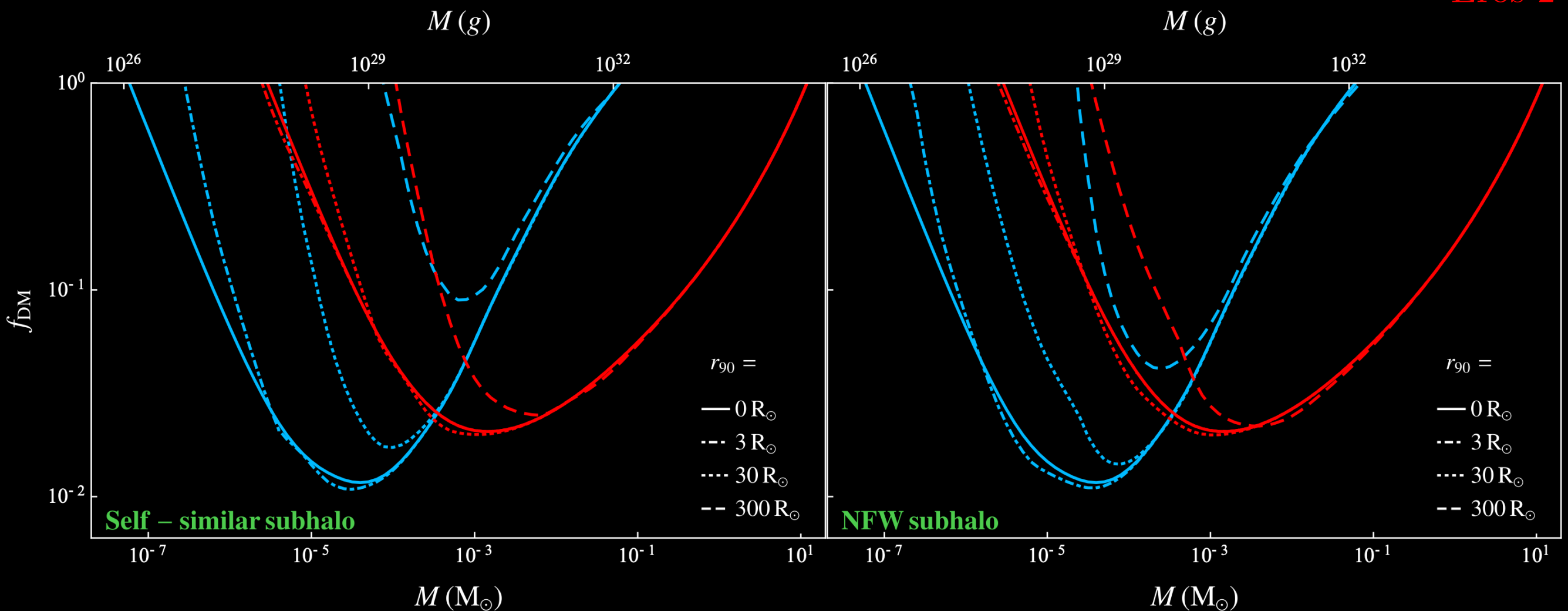
Constraints on DM fraction

Using the differential event rate, find constraints given expected number of (non-observed) events. Use derived efficiency for extended lenses.

As expected, constraints on extended objects are weaker...

Ogle-IV

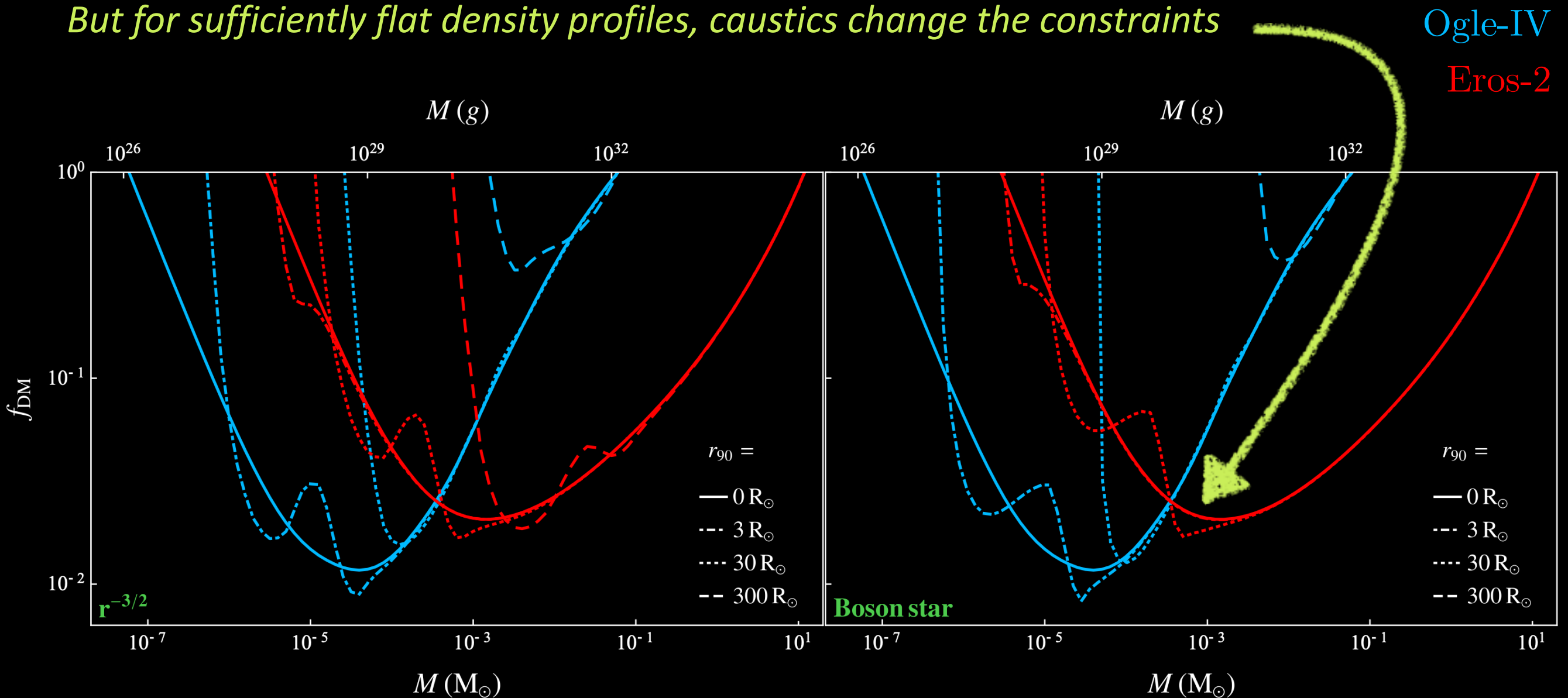
Eros-2

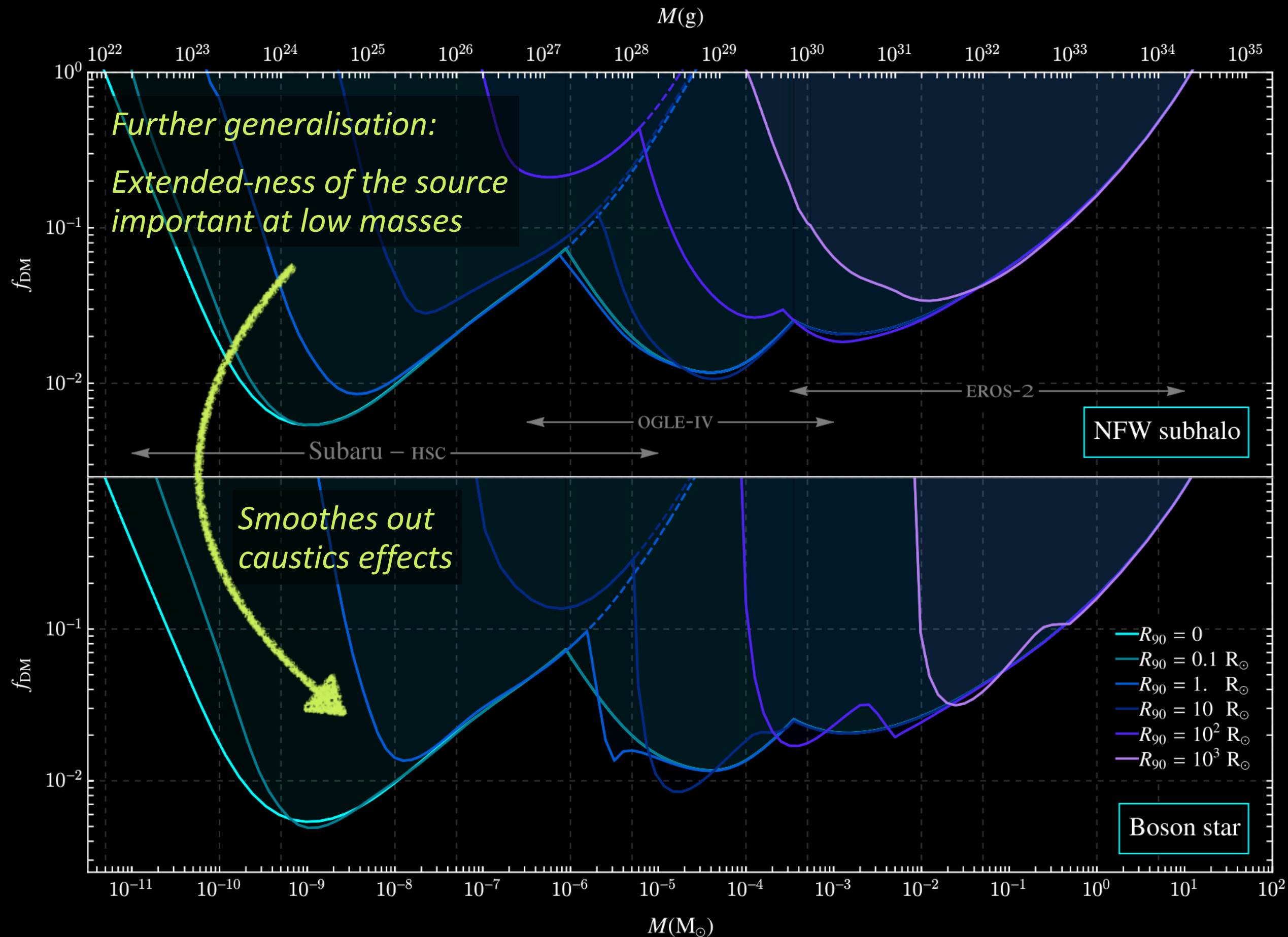


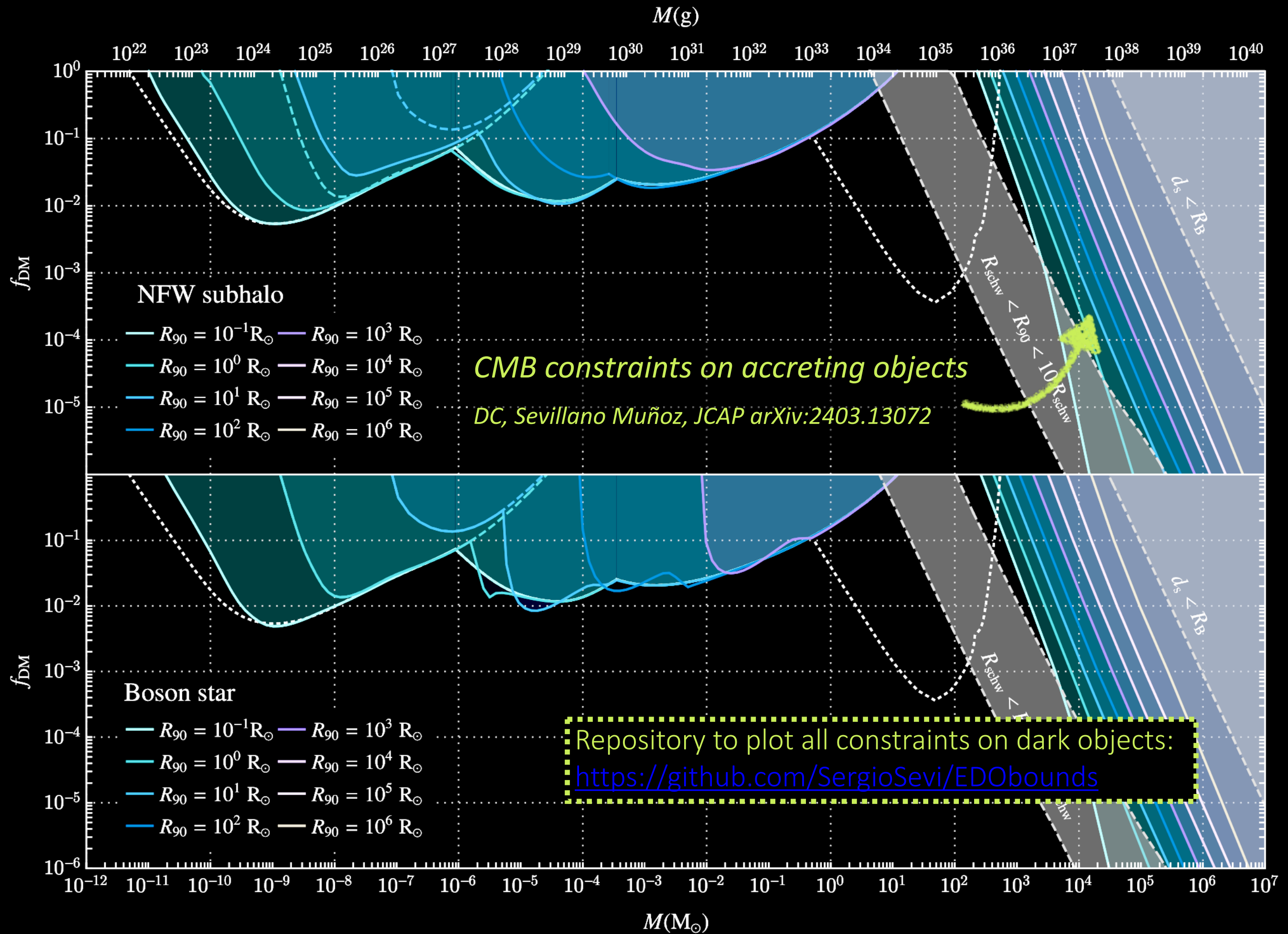
Constraints on DM fraction

Using the differential event rate, find constraints given expected number of (non-observed) events. Use derived efficiency for extended lenses.

But for sufficiently flat density profiles, caustics change the constraints



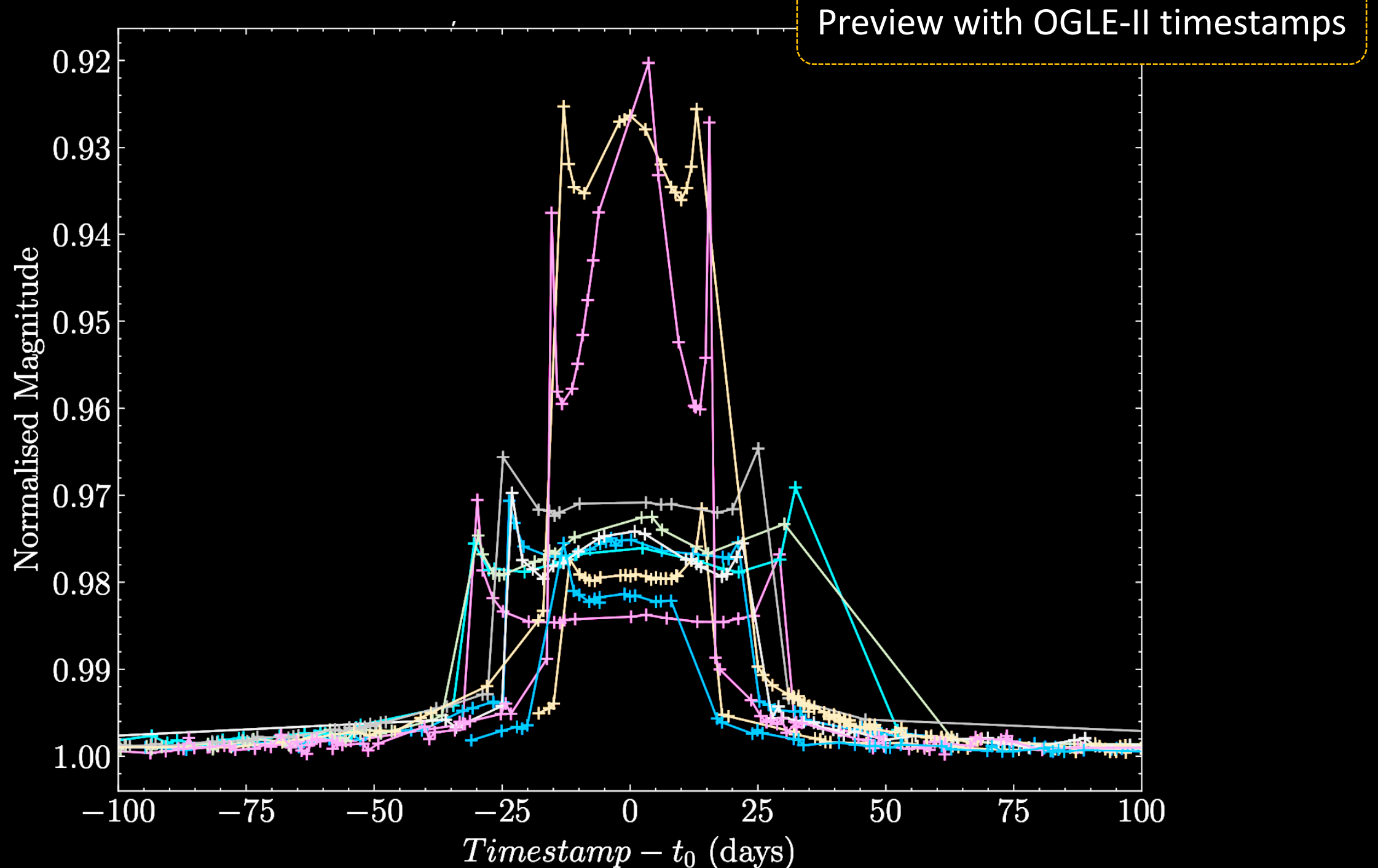




Light curves with caustics

Can we look for these explicitly?

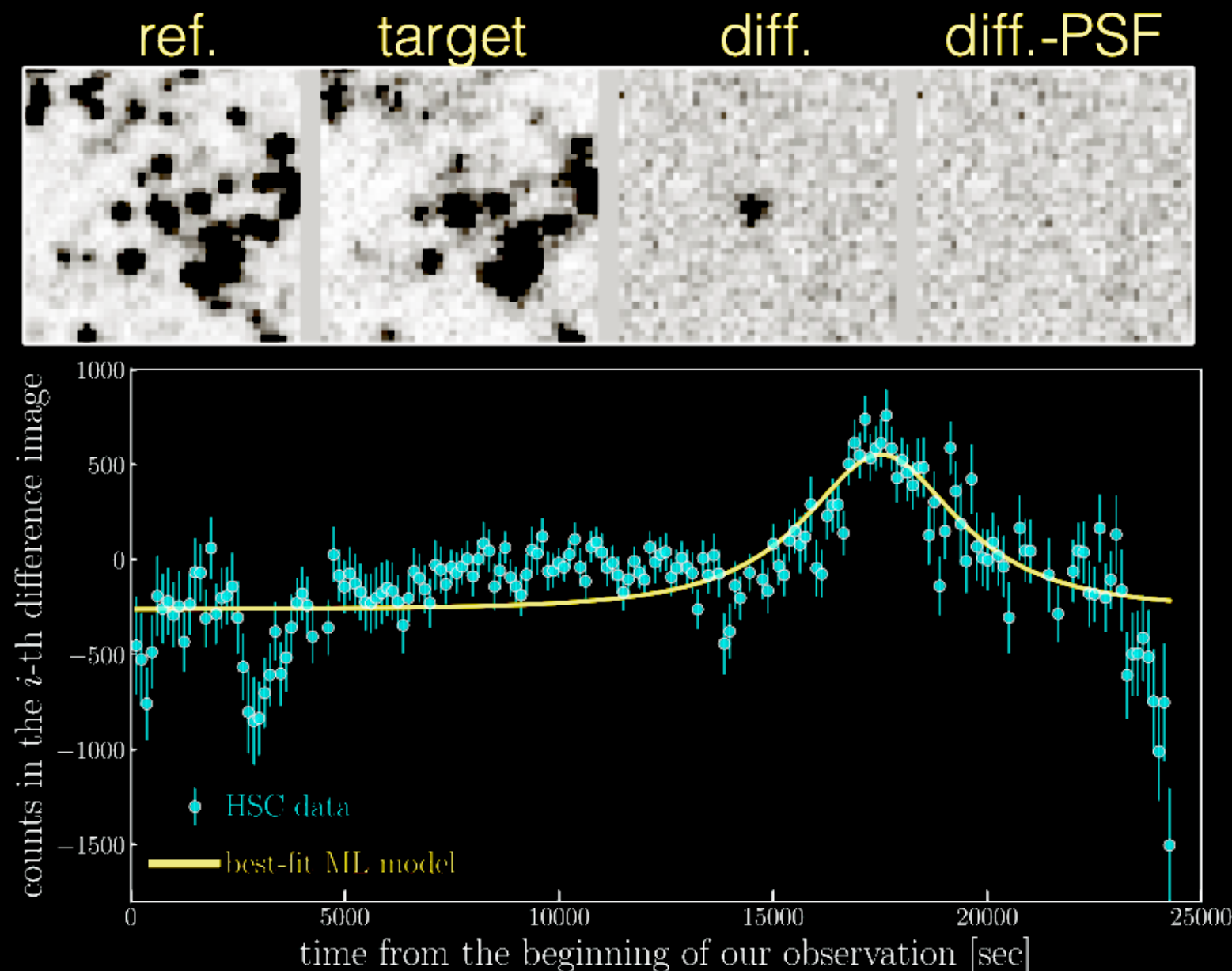
M. Crispim-Romao, DC, PRD, arXiv:2402.00107



ML + ML

Microlensing + Machine Learning

- Microlensing data is time series data
- Challenge: low-cadence data, lower signal-to-noise ratios



From the Subaru collaboration, arXiv:1701.02151

ML + ML

Microlensing + Machine Learning

- Microlensing data is time series data
- Challenge: low-cadence data, lower signal-to-noise ratios
- MicroLIA: use a Random Forest (RF) algorithm to find microlensing event (and distinguish from other events)

Godines et al, arXiv:2004.14347

ML + ML

Microlensing + Machine Learning

- Microlensing data is time series data
- Challenge: low-cadence data, lower signal-to-noise ratios
- MicroLIA: use a Random Forest (RF) algorithm to find microlensing event (and distinguish from other events)

Our adaptations:

- Implement **boson star** and **NFW** light curves with $0.5 < \tau_m < 5$
 - Instead of an RF, we use a histogram-based gradient boosted classifier (HBGC) to improve speed
 - Add criterium $\mu \geq 1.34$
- (... and a few fixes)

Complete datasets not available

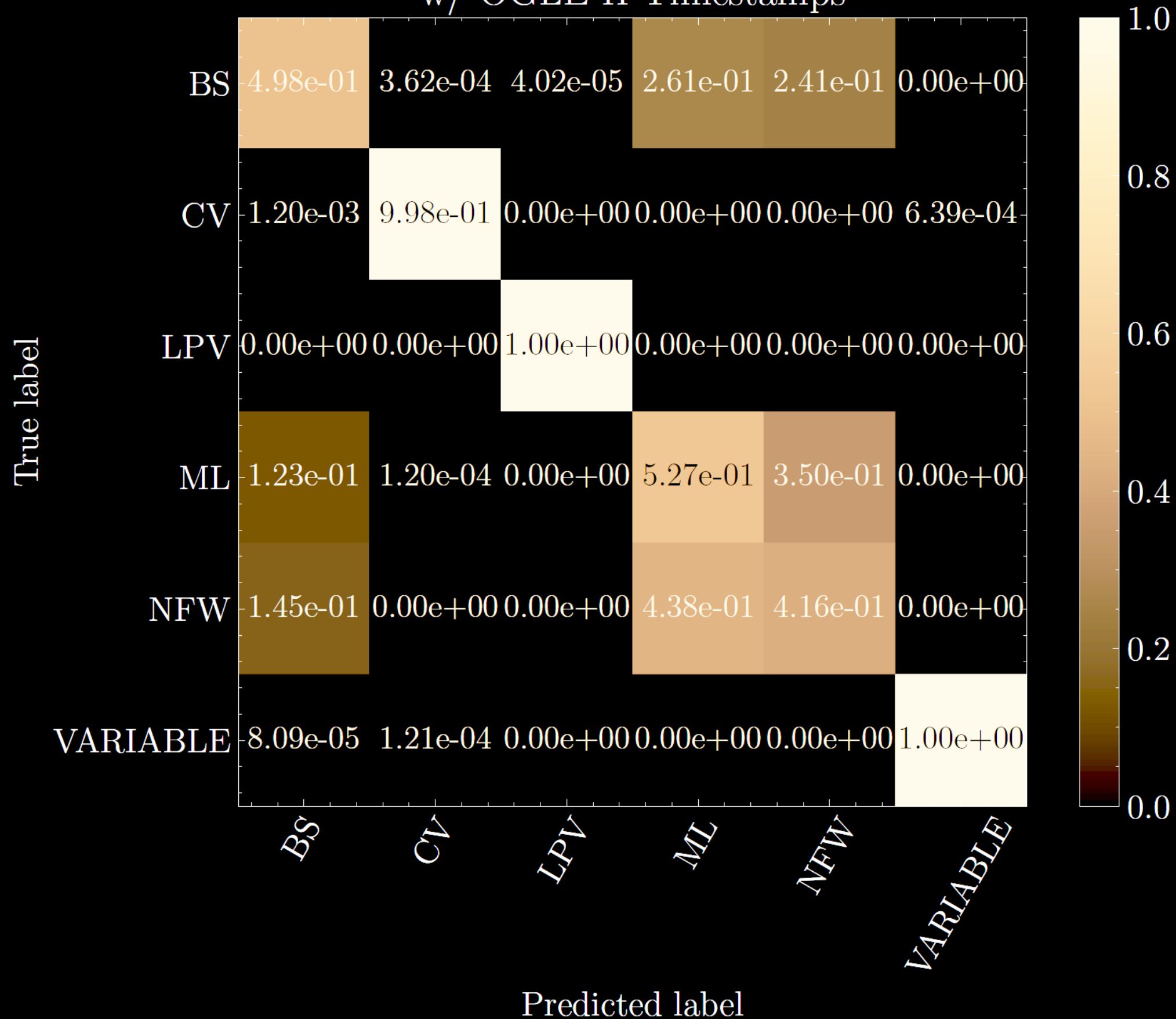
Table 1
Selection Criteria for High-quality Microlensing Events in OGLE GVS Fields

Criteria	Remarks	Number
All stars in databases		1,856,529,265
$\chi^2_{\text{out}}/\text{dof} \leq 2.0$ $n_{\text{DIA}} \geq 3$ $\chi_{3+} = \sum_i (F_i - F_{\text{base}})/\sigma_i \geq 32$	No variability outside a window centered on the event (duration of the window depends on the field) Centroid of the additional flux coincides with the source star centroid Significance of the bump	23,618
$A \geq 0.1$ mag $n_{\text{bump}} = 1$	Rejecting low-amplitude variables Rejecting objects with multiple bumps	18,397
$\chi^2_{\text{fit}}/\text{dof} \leq 2.0$ $\chi^2_{\text{fit},t_E}/\text{dof} \leq 2.0$ $\sigma(t_E)/t_E < 0.5$ $t_{\text{min}} \leq t_0 \leq t_{\text{max}}$ $u_0 \leq 1$ $t_E \leq 500$ d $A \geq 0.4$ mag if $t_E \geq 100$ days $I_s \leq 21.0$ $F_b > -F_{\text{min}}$	Fit quality: χ^2 for all data χ^2 for $ t - t_0 < t_E$ Einstein timescale is well measured Event peaked between t_{min} and t_{max} , which are moments of the first and last observation of a given field Maximum impact parameter Maximum timescale Long-timescale events should have high amplitudes Maximum I -band source magnitude Maximum negative blend flux, corresponding to $I = 20.5$ mag star	460

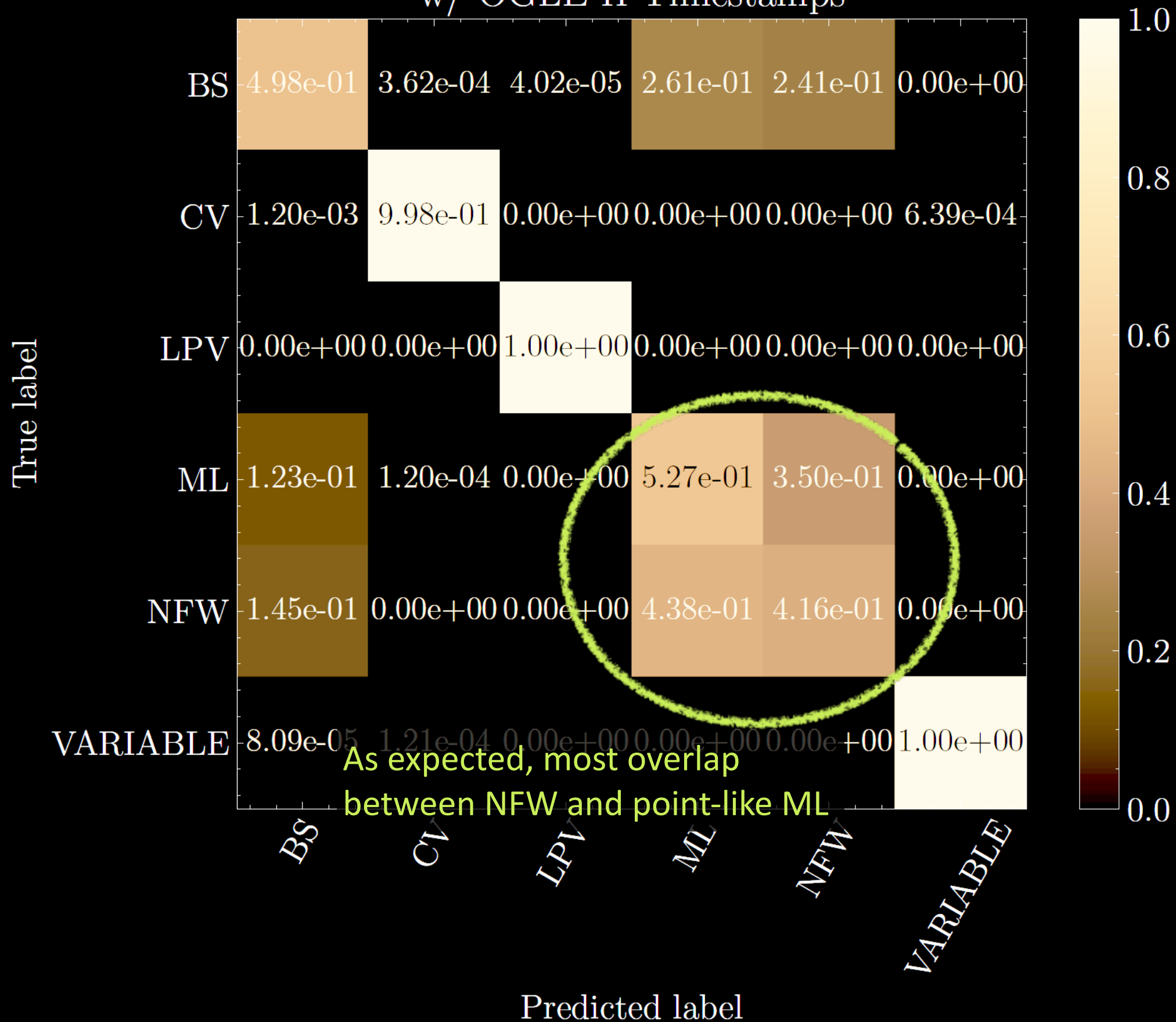
 **Reject events with multiple bumps**

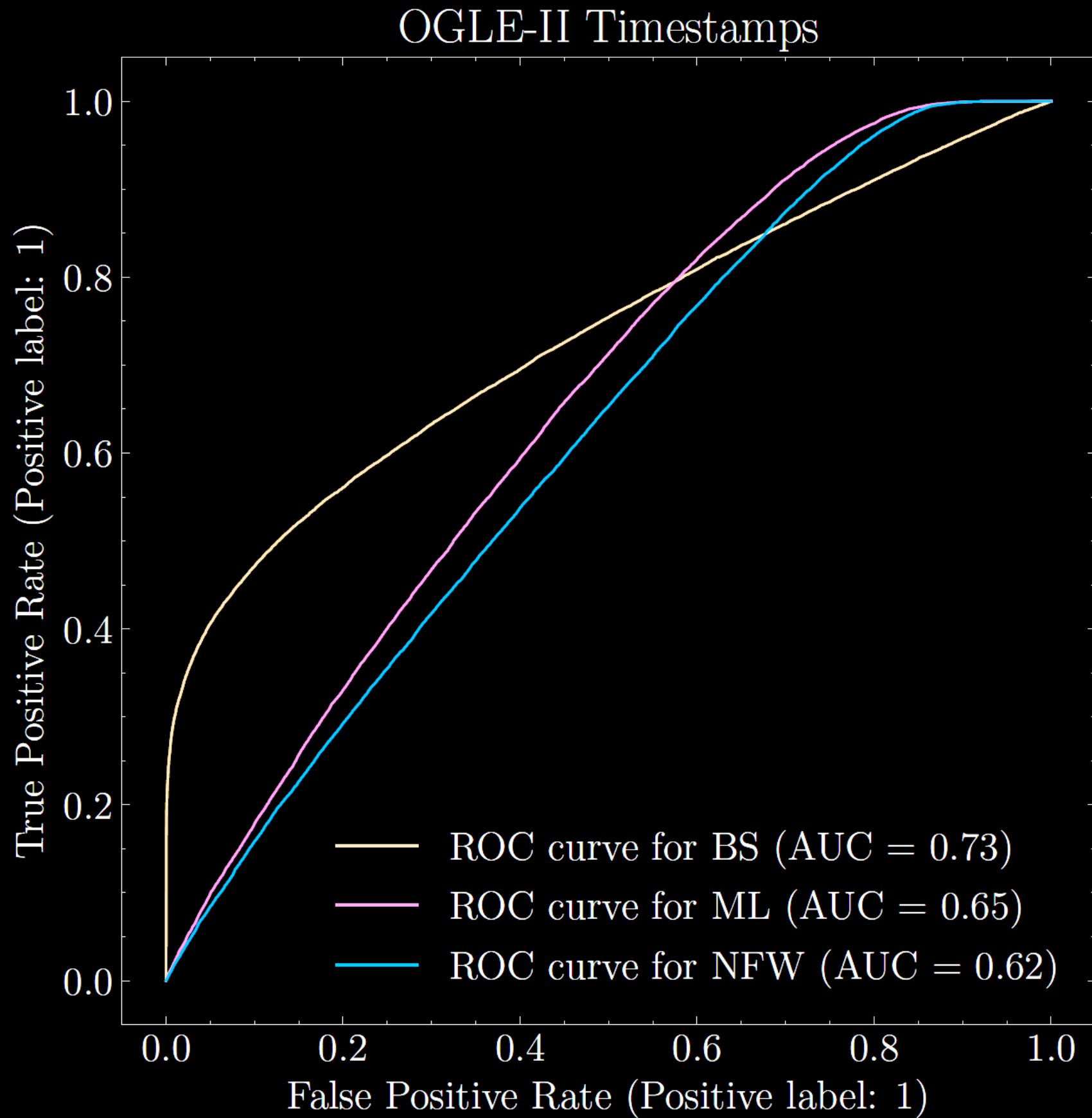
So for now... generating and injecting events

Confusion Matrix All vs All
w/ OGLE-II Timestamps

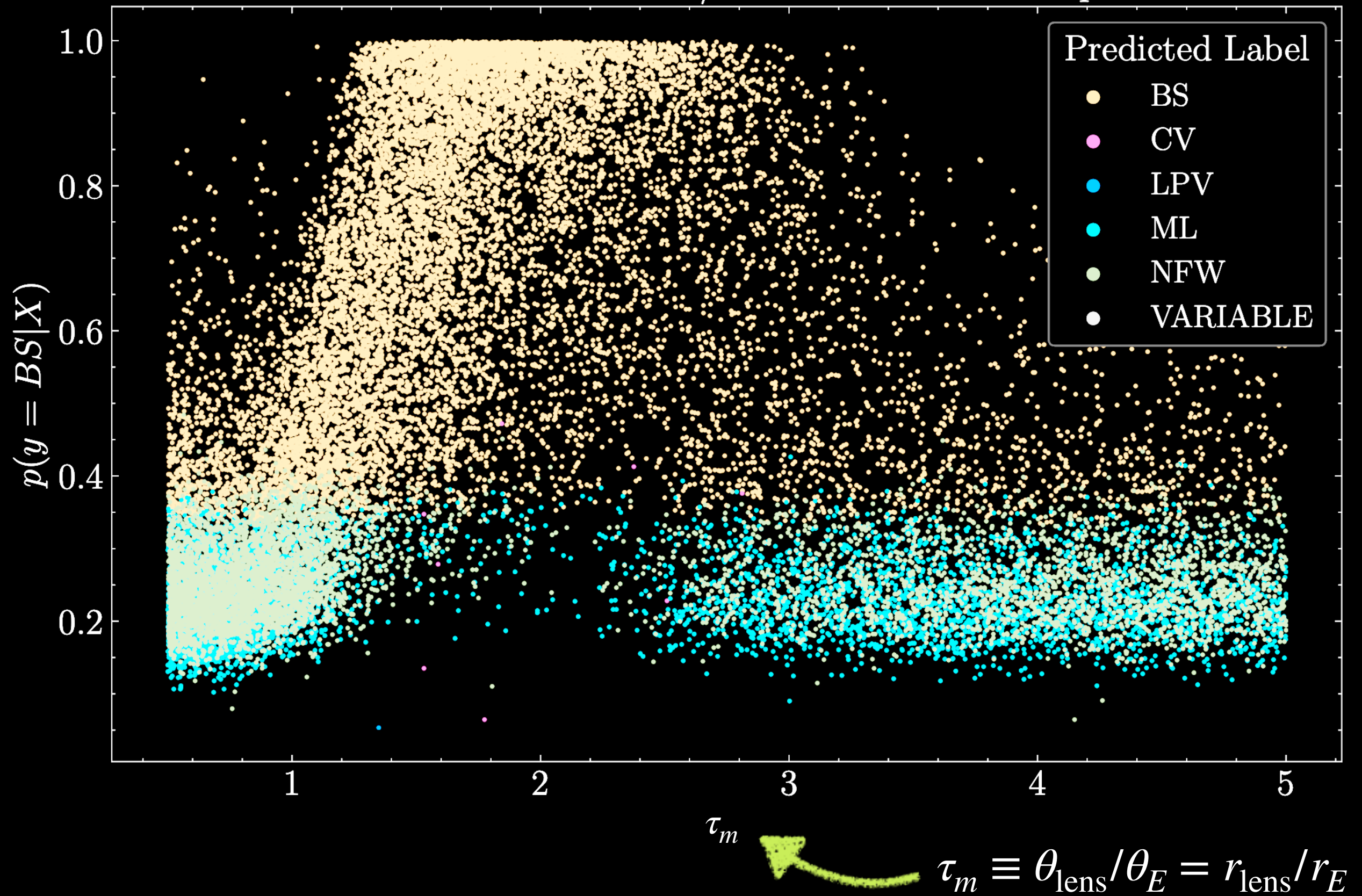


Confusion Matrix All vs All w/ OGLE-II Timestamps

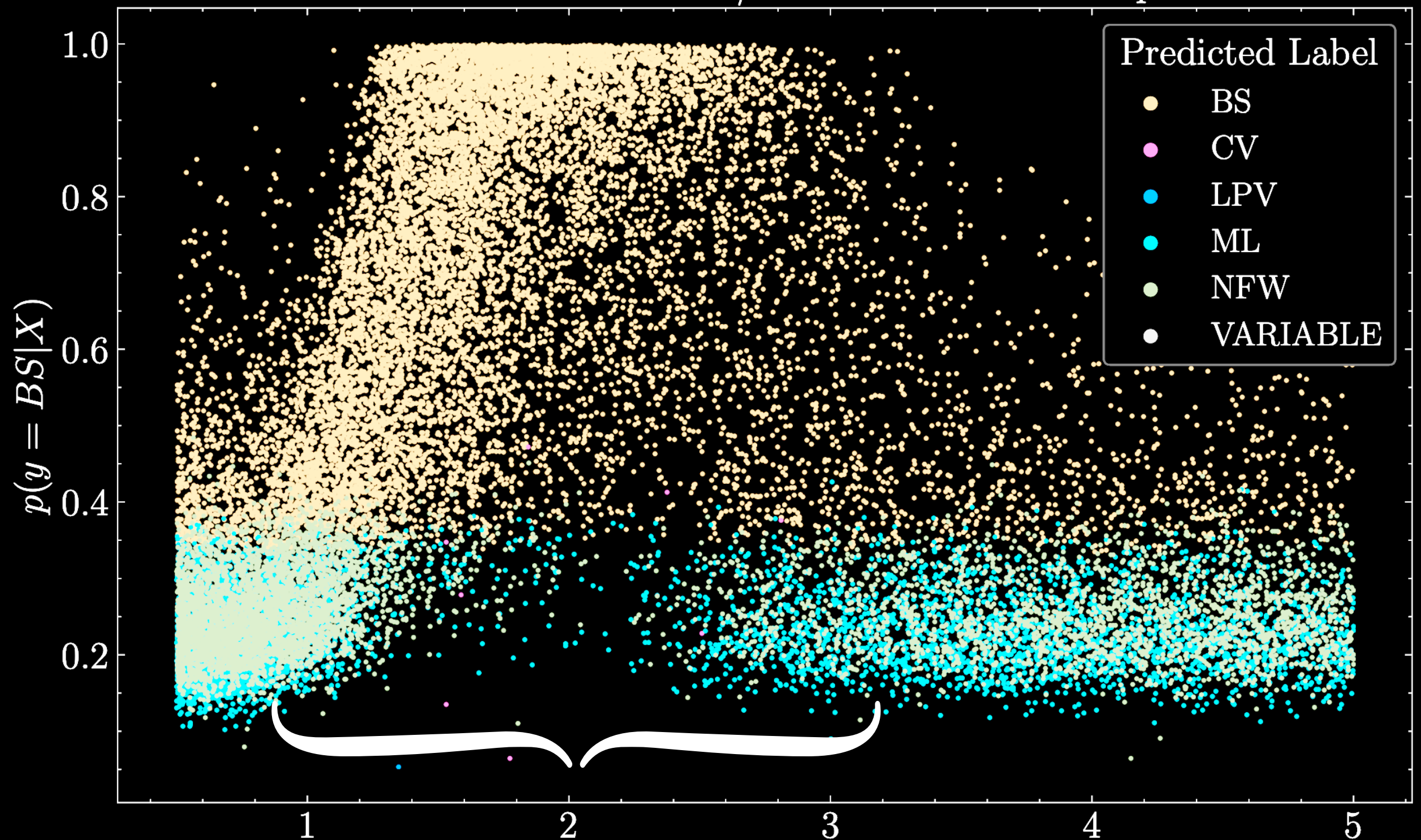




Boson Star Events w/ OGLE-II Timestamps



Boson Star Events w/ OGLE-II Timestamps



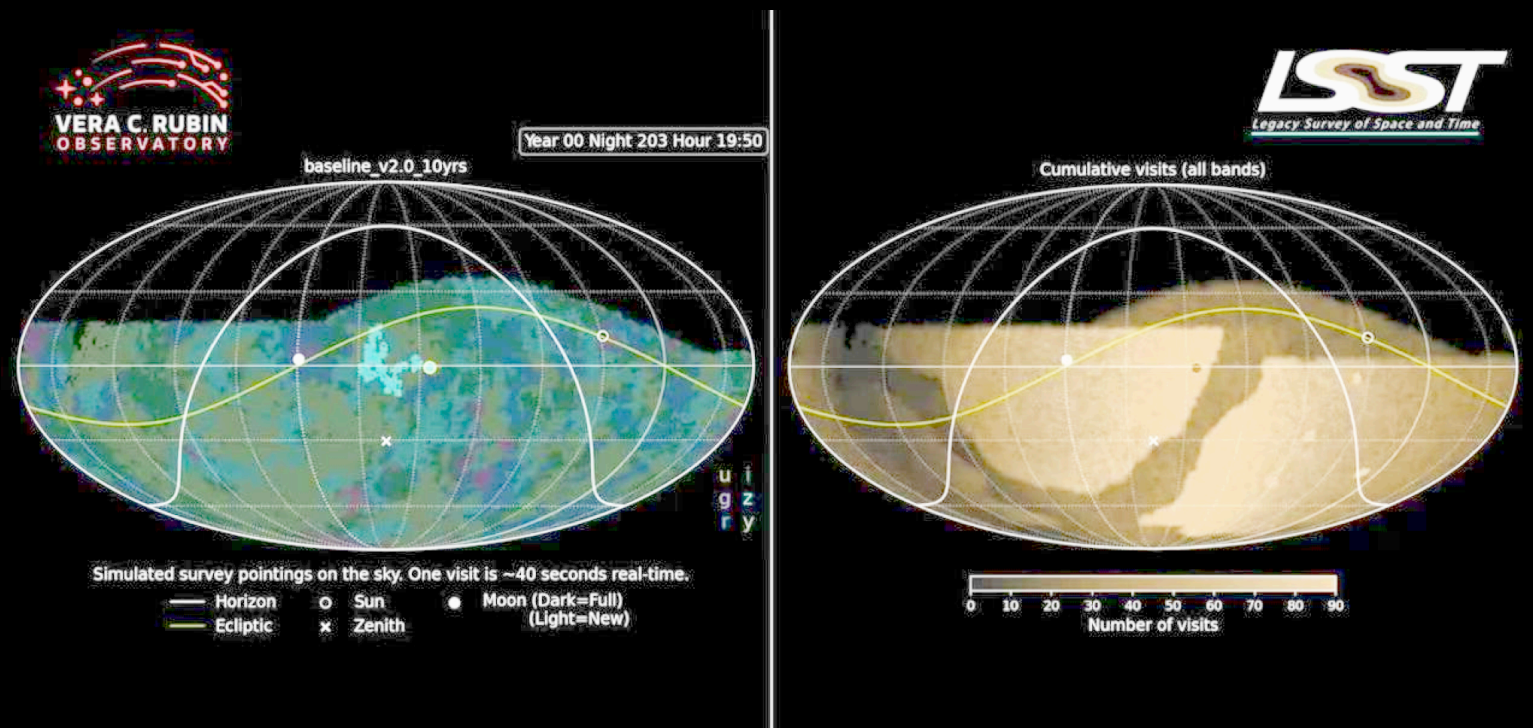
Indeed, the most probable
detections are for $0.8 < \tau_m < 3$

$$\tau_m \equiv \theta_{\text{lens}}/\theta_E = r_{\text{lens}}/r_E$$

LSST by Rubin

The Legacy Survey of Space and Time

- 10-year survey by the Vera C. Rubin Observatory
- Large field of view and rapid survey speed
- Relatively high cadence observations, allowing frequent monitoring of $\sim 10^9$ stars
- However, sparse — *develop follow up strategies*



LSST by Rubin: simulations

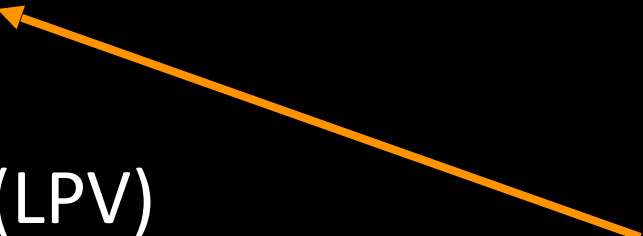
<https://zenodo.org/records/15005108>

Miguel Crispim-Romao, DC, Daniel Godines, arXiv:2503.09699

Simulated, using rubinsim, 7 classes of observations:

- *Constant*
- Mira long-period variables (LPV)
- RR Lyrae and Cepheid Variables (RRLyrae)
- point-like microlensing (ML)
- binary microlensing
- microlensing by NFW-subhalos
- microlensing by boson stars (BS)

Cadence:
(baseline_v2.0_10yrs)

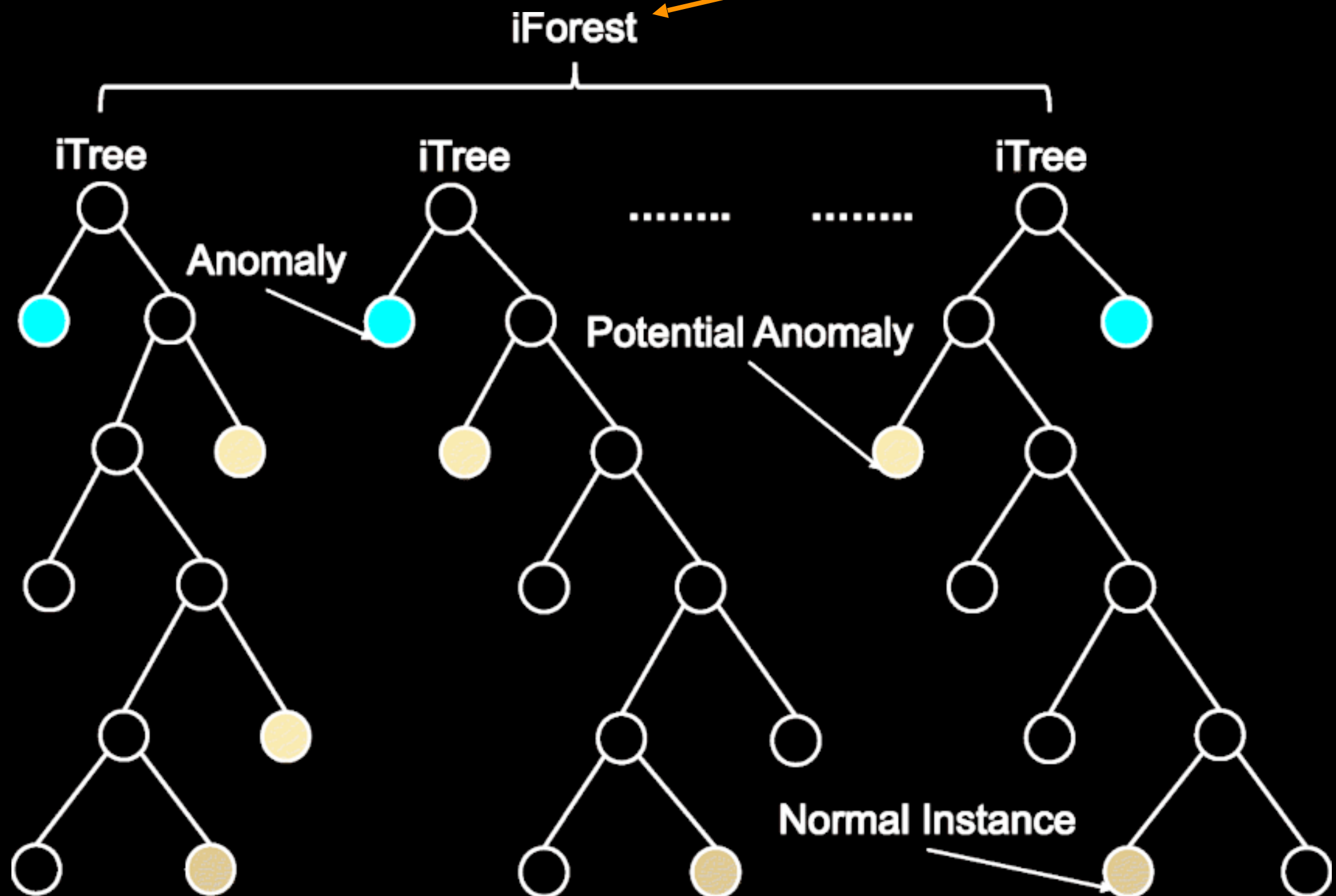


parameter	min	max	spacing
t_E (days)	0	100	linear
u_0	0	3	linear
τ_m	0.05	5	logarithmic

LSST by Rubin: anomaly detection

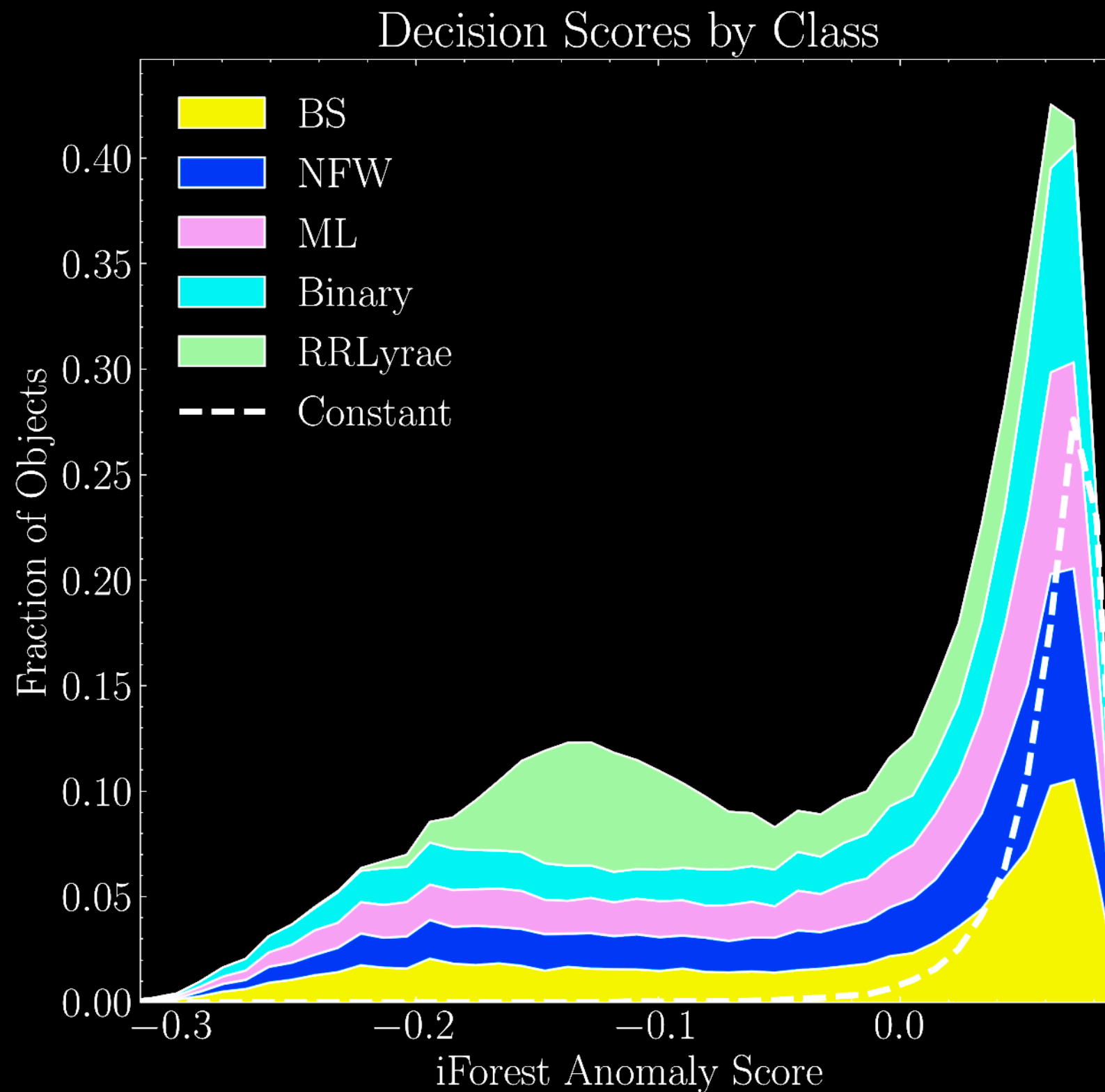
Miguel Crispim-Romao, DC, Daniel Godines, arXiv:2503.09699

Train on signal-less light curves



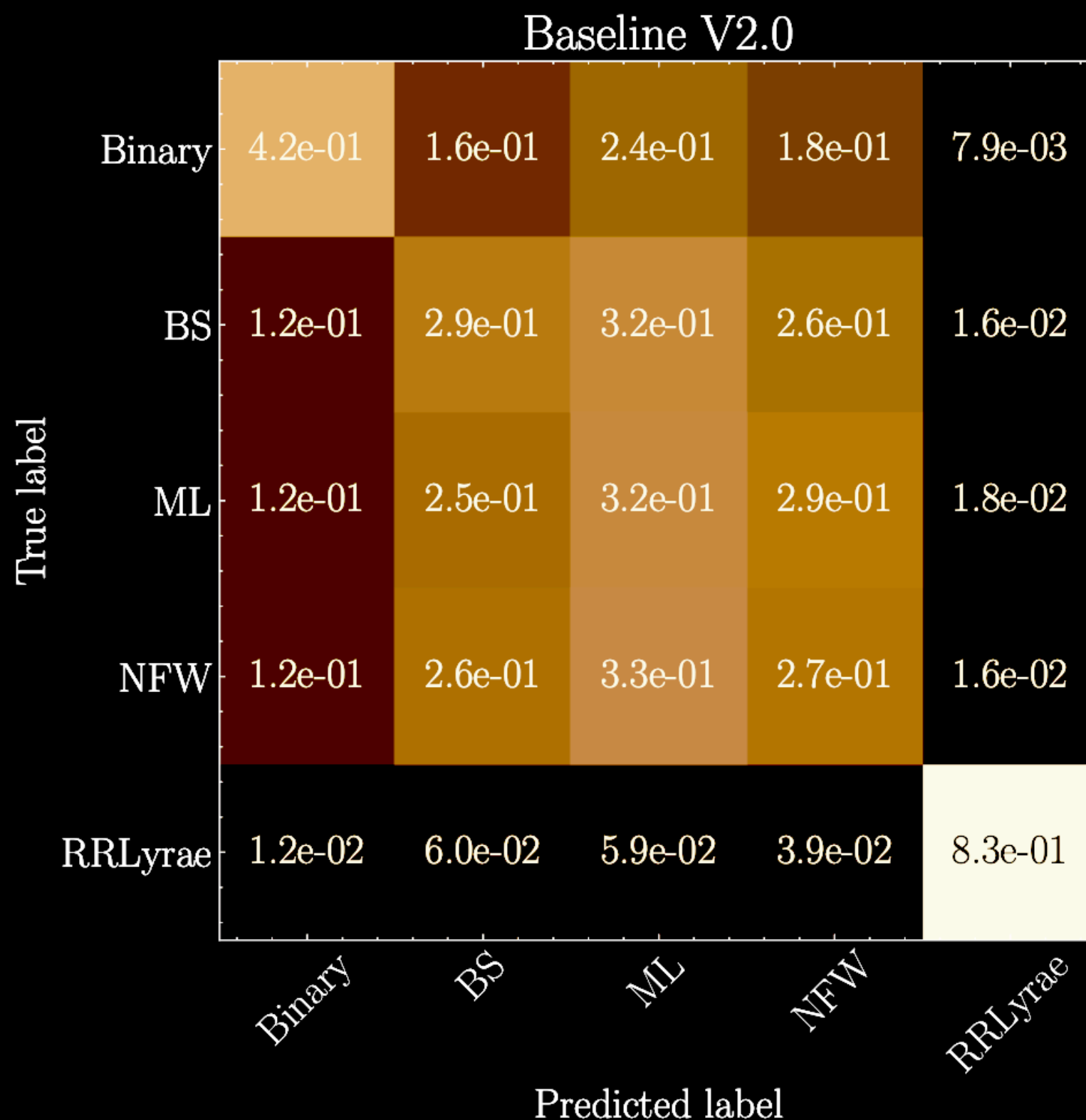
LSST by Rubin: offline analysis

Miguel Crispim-Romao, DC, Daniel Godines, arXiv:2503.09699



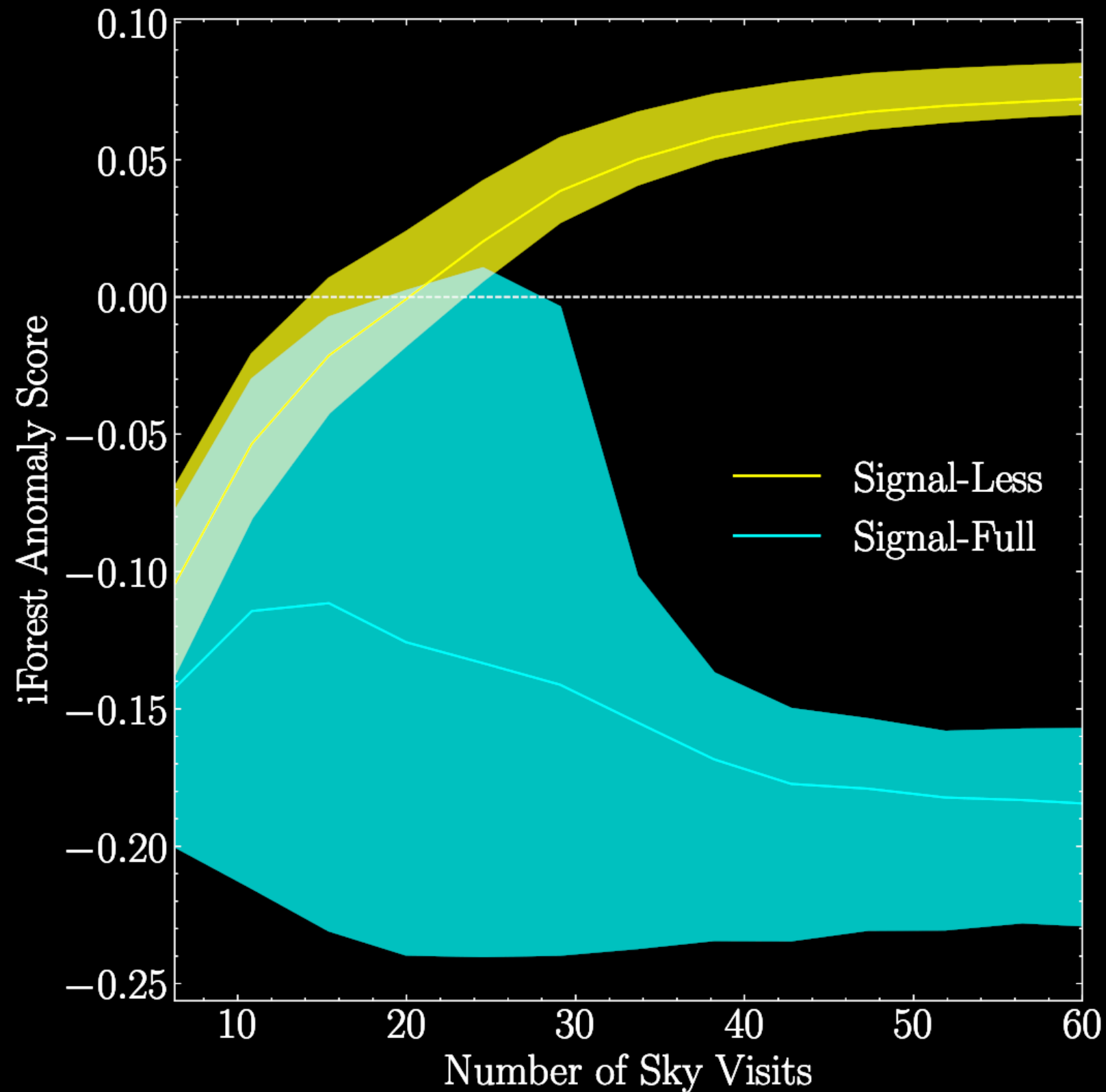
LSST by Rubin: offline analysis

Miguel Crispim-Romao, DC, Daniel Godines, arXiv:2503.09699



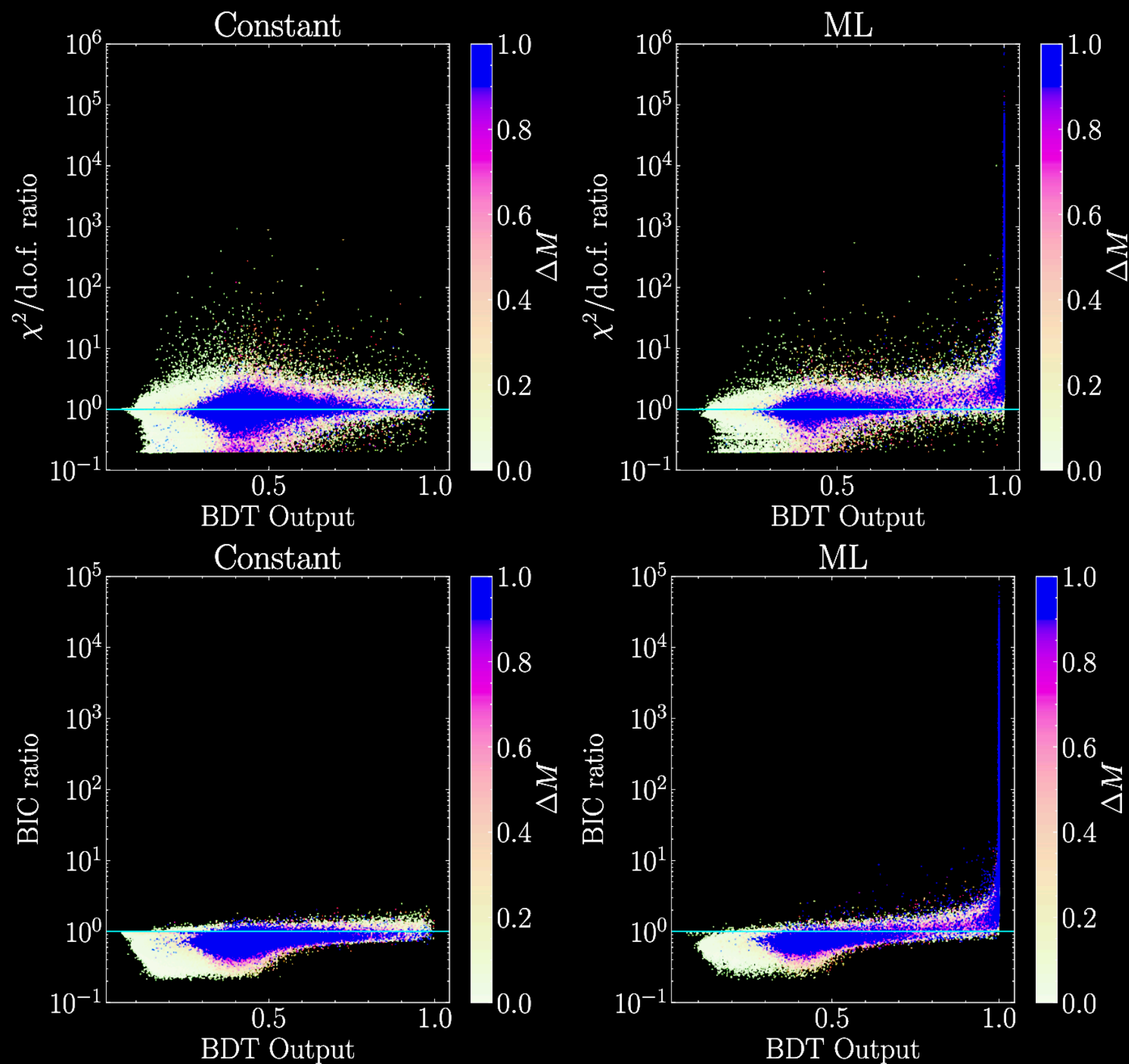
LSST by Rubin: online analysis

Miguel Crispim-Romao, DC, Daniel Godines, arXiv:2503.09699



LSST by Rubin: projections

Miguel Crispim-Romao, DC, Benedict Crossey, Daniel Godines, arXiv:2506.XXXX

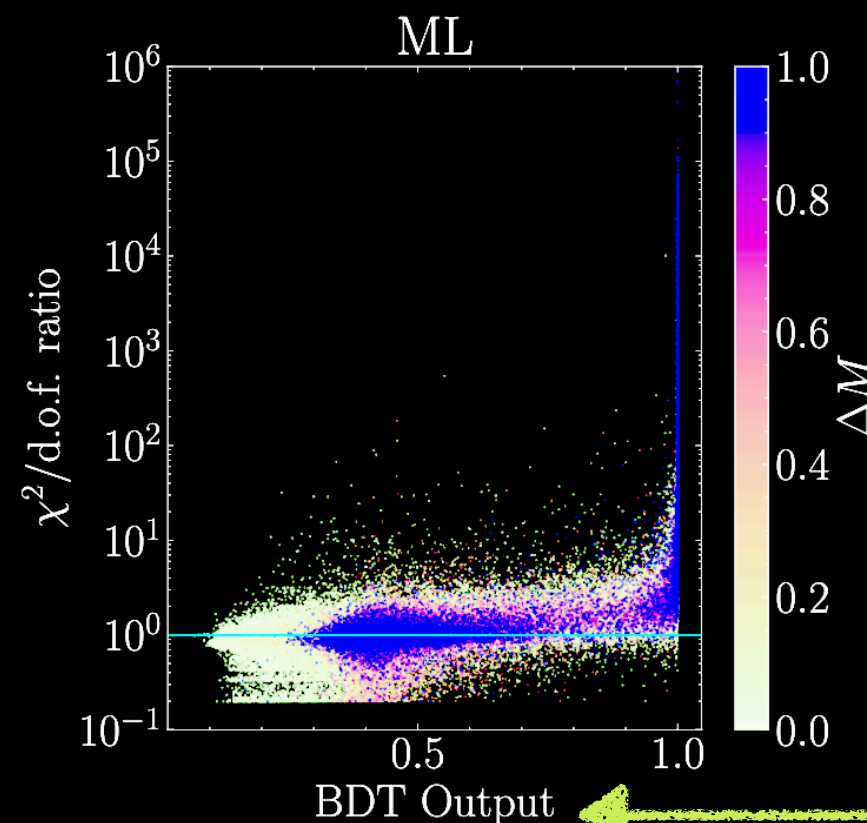
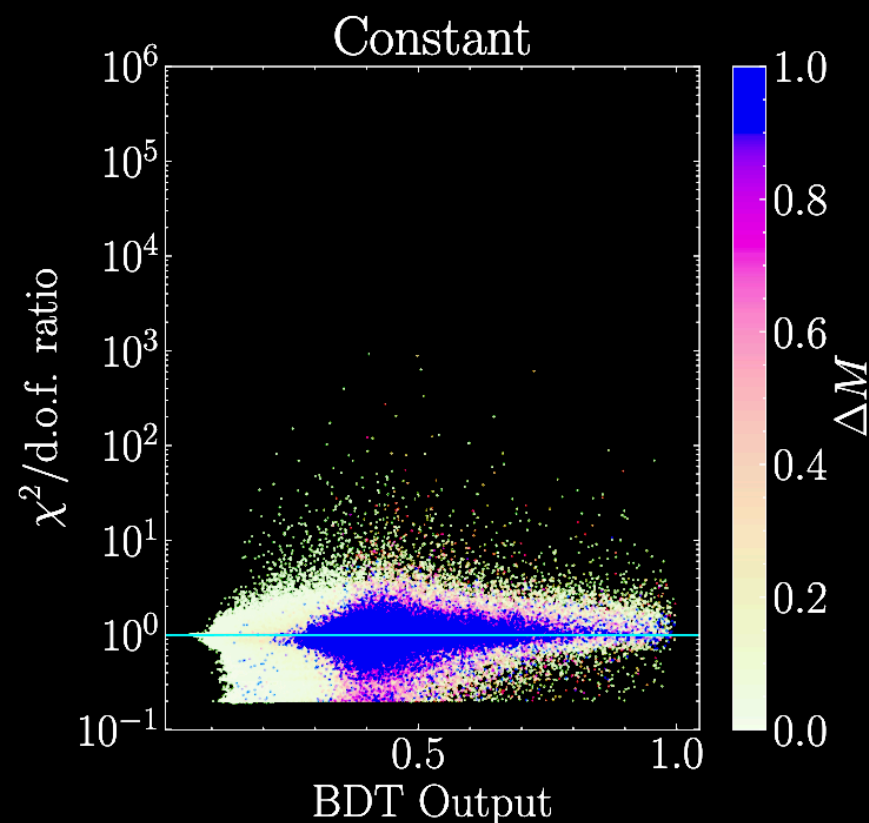


LSST by Rubin: projections

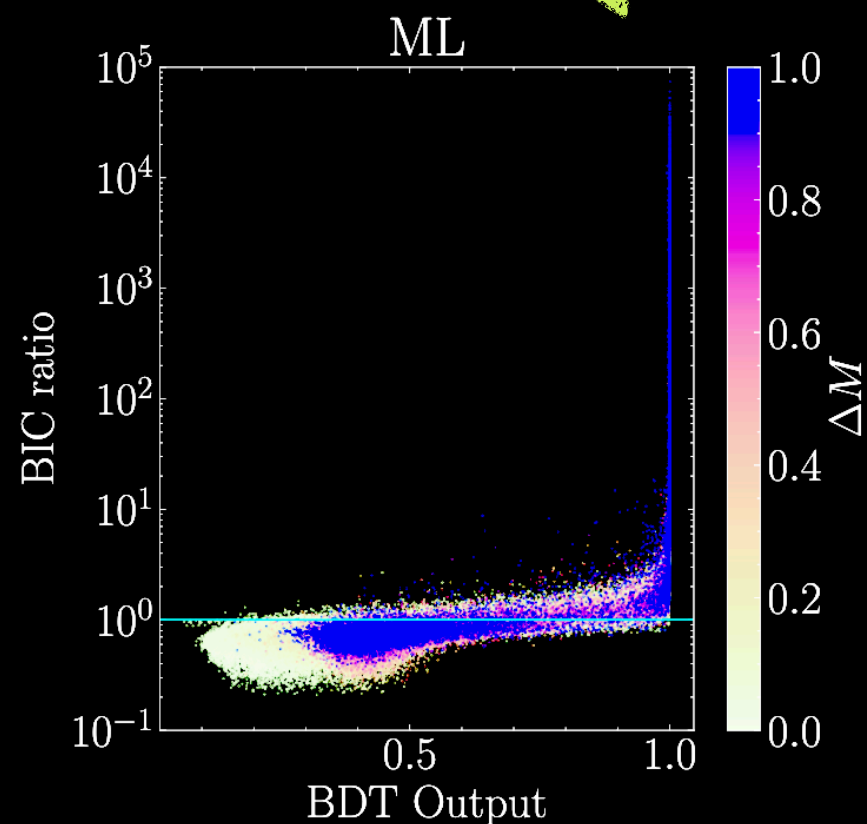
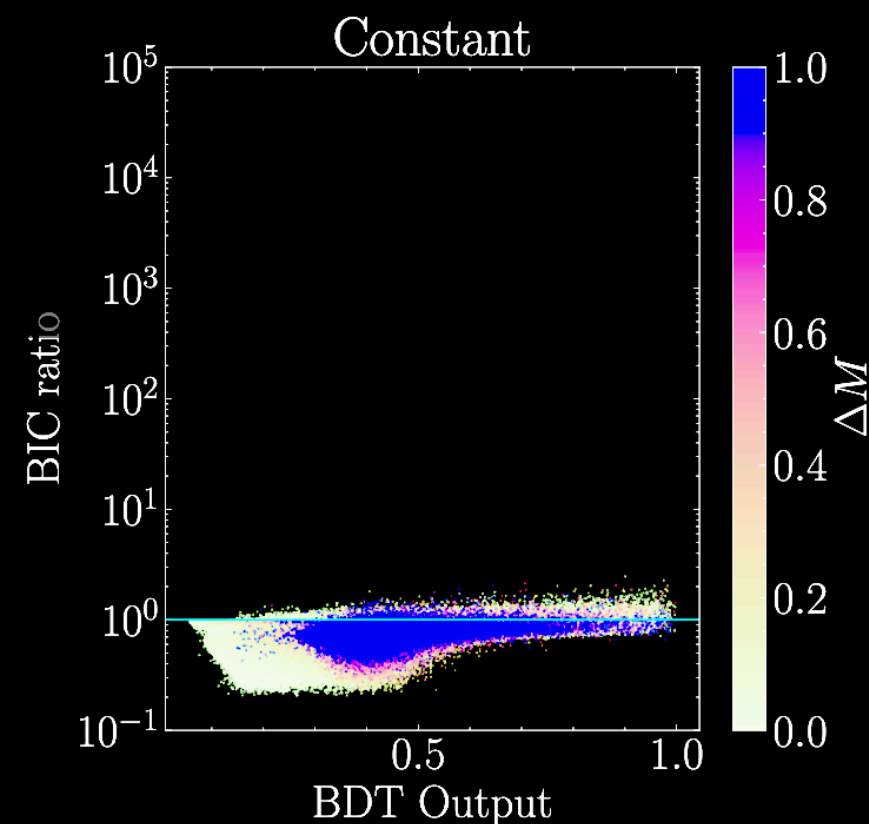
Miguel Crispim-Romao, DC, Benedict Crossey, Daniel Godines, arXiv:2506.XXXX

Goal: design cuts which optimise efficiency while minimising false positives

Typically used



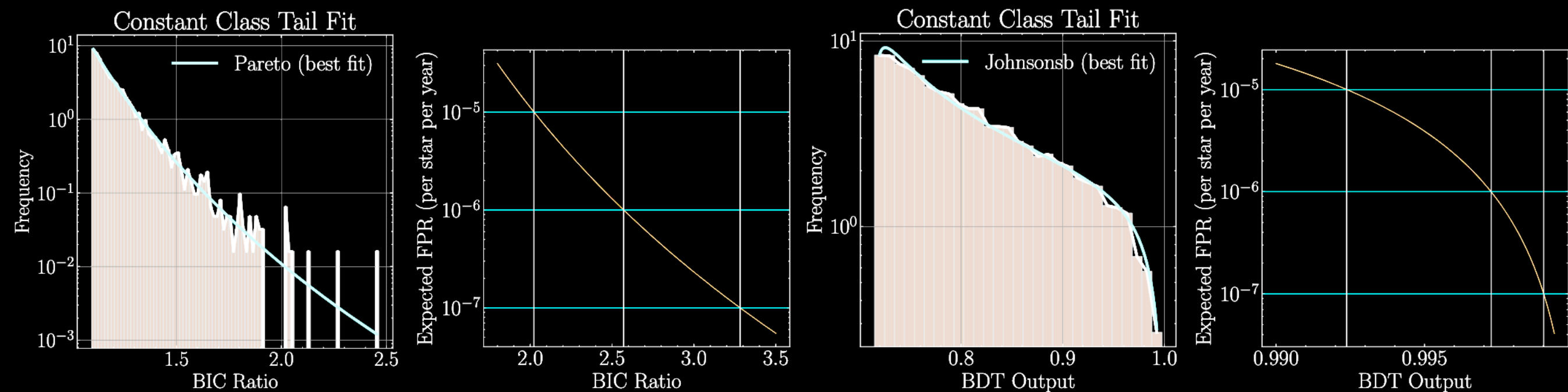
BIC



BDT

LSST by Rubin: projections

Miguel Crispim-Romao, DC, Benedict Crossey, Daniel Godines, arXiv:2506.XXXX



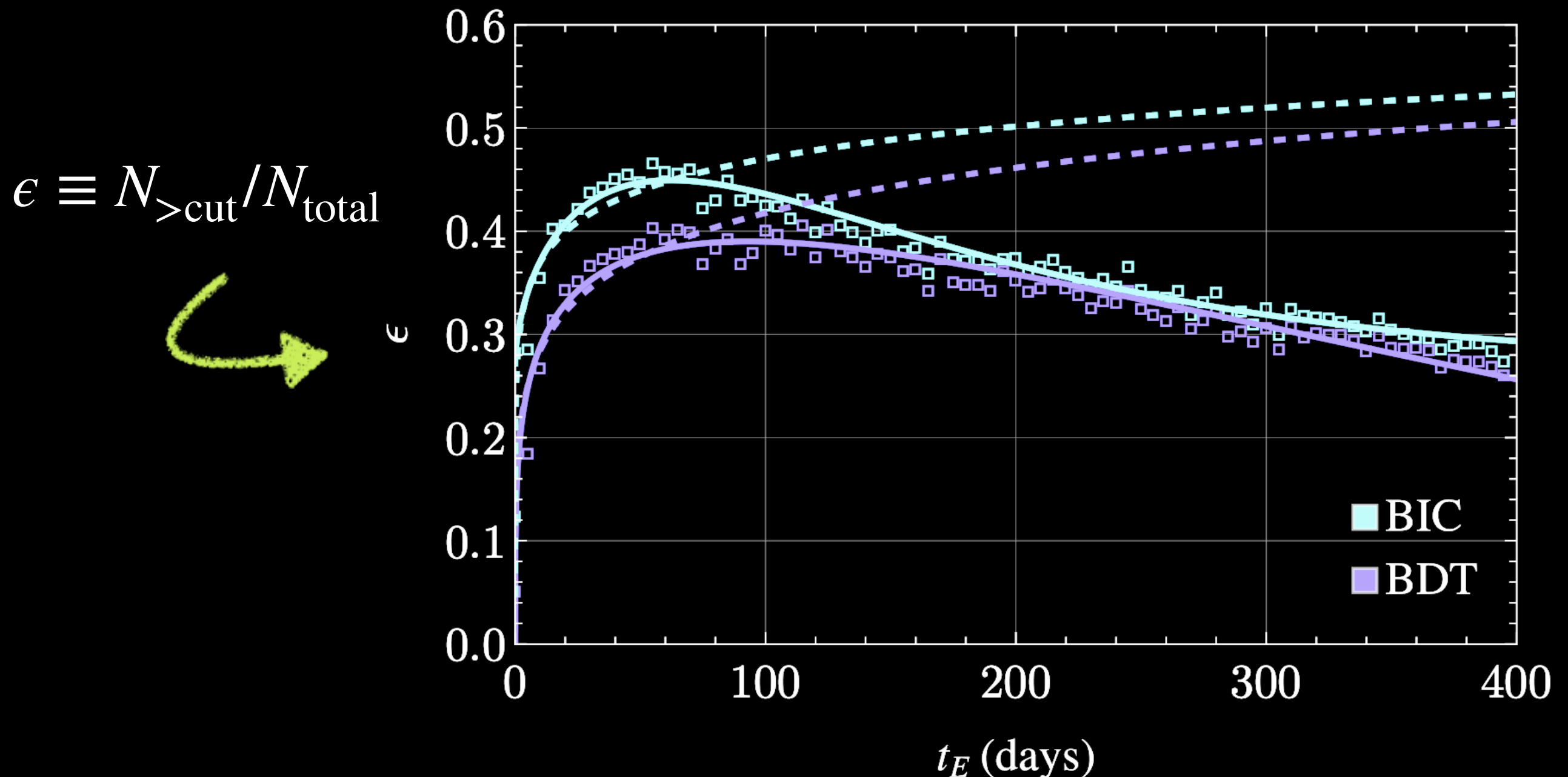
For competitive constraints,
need $FPR < 10^{-7}$

cut	FPR	$\bar{\epsilon}$
BIC ratio > 3.28	10^{-7}	0.38
BDT > 0.999	10^{-7}	0.34
$\chi^2/\text{d.o.f. ratio} > 10$	3.5×10^{-4}	0.30
$\chi^2/\text{d.o.f. ratio} > 10, \tilde{u}_0 < 1$	1.1×10^{-4}	0.20

LSST by Rubin: projections

Miguel Crispim-Romao, DC, Benedict Crossey, Daniel Godines, arXiv:2506.XXXX

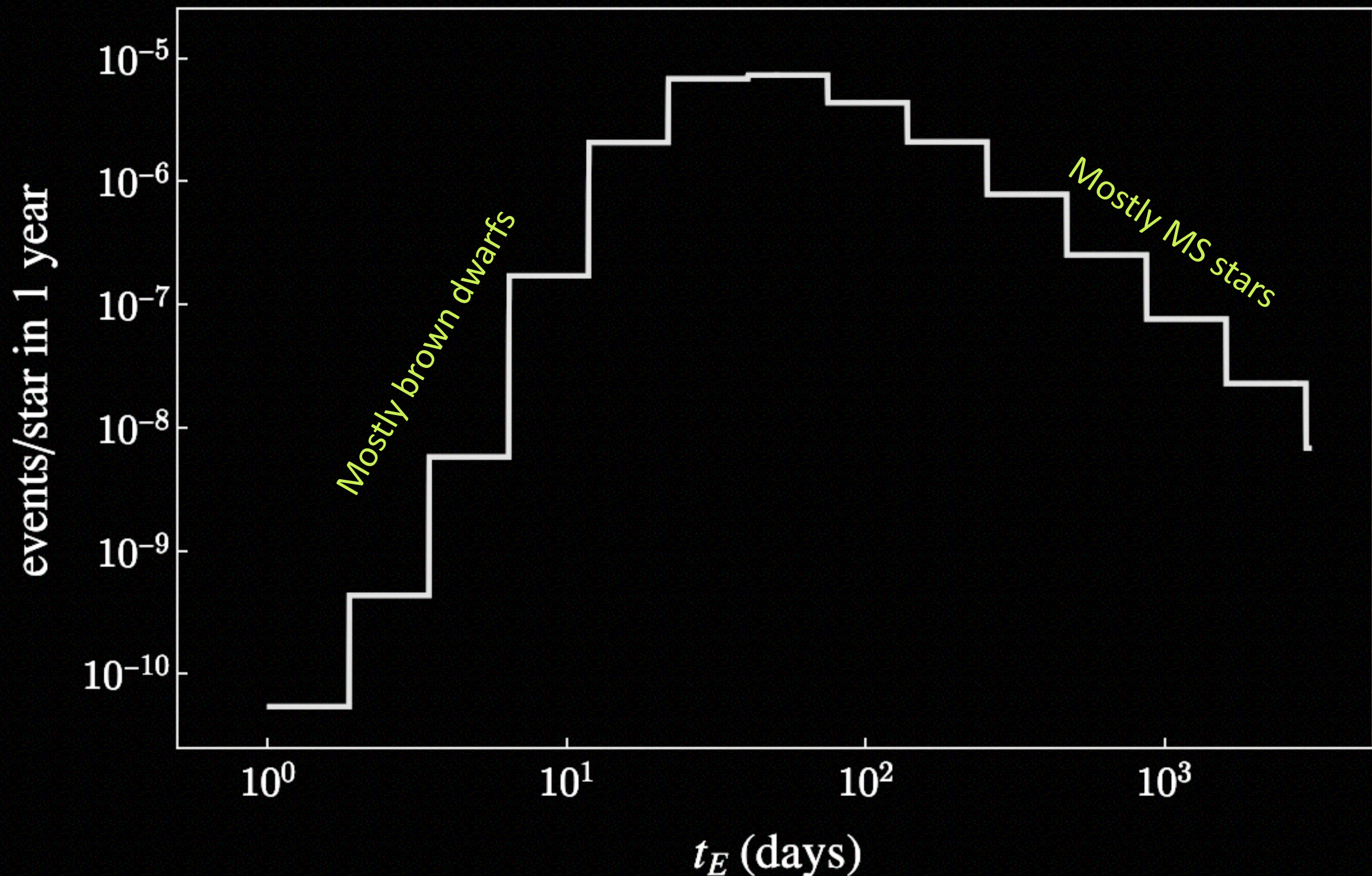
Define a cut in false positives — what fraction of events is identified?



LSST by Rubin: projections

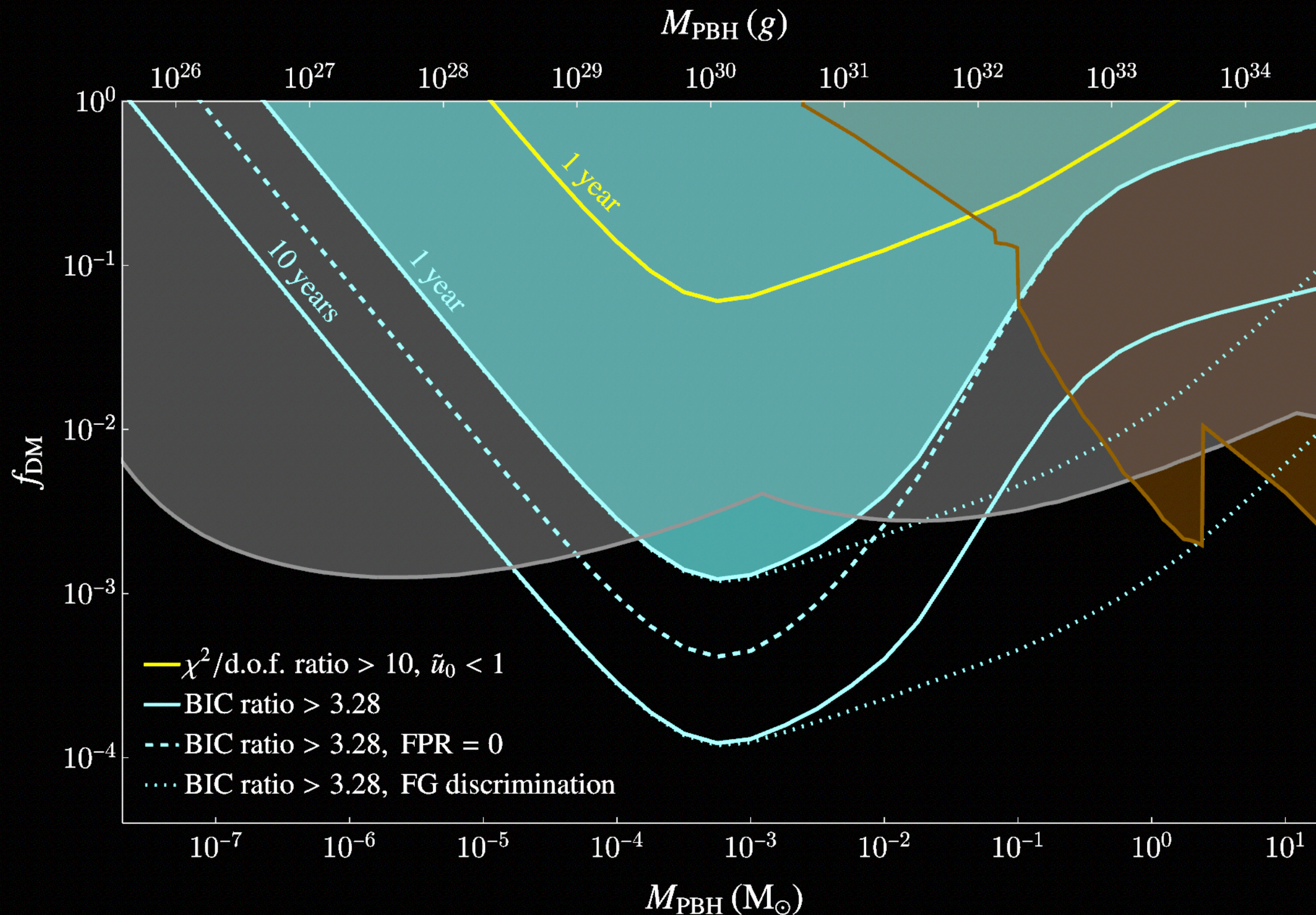
Miguel Crispim-Romao, DC, Benedict Crossey, Daniel Godines, arXiv:2506.XXXX

Foreground



LSST by Rubin: projections

Miguel Crispim-Romao, DC, Benedict Crossey, Daniel Godines, arXiv:2506.XXXX



EDOs and the early Universe

- Ultracompact mini haloes (UCMH, $\rho \sim r^{-3/2}$) are formed from the collapse of **primordial overdensities**
- The non-observation of UCMH can therefore be used to draw conclusions about the **primordial power spectrum**
- This has been done for e.g. PTAs and WIMPs (= model-dependent)

Clark, Lewis, Scott, MNRAS, arXiv:1509.02938
Bringmann, Scott, Akrami, PRD, arXiv:1110.2484
- But we now have far more gravitational probes...

The primordial power spectrum

- Primordial curvature perturbations with amplitude $\mathcal{P}_{\mathcal{R}}(k)$ determine the variance $\sigma_{\chi,H}(R)$ of CDM density fluctuations at horizon entry
 - If $\delta_{\chi}(R)$ exceeds a threshold $\delta_{\chi}^{\min}(R) \sim 10^{-3}$, the region collapses into an UCMH (much smaller than for PBHs)
 - If $\sigma_{\chi,H}(R)$ is too large many regions will exceed $\delta_{\chi}^{\min}(R)$

The primordial power spectrum

- Primordial curvature perturbations with amplitude $\mathcal{P}_{\mathcal{R}}(k)$ determine the variance $\sigma_{\chi,H}(R)$ of CDM density fluctuations at horizon entry
 - If $\delta_{\chi}(R)$ exceeds a threshold $\delta_{\chi}^{\min}(R) \sim 10^{-3}$, the region collapses into an UCMH
 - If $\sigma_{\chi,H}(R)$ is too large many regions will exceed $\delta_{\chi}^{\min}(R)$
- From f_{DM} we **work backward**:
 - f_{DM} sets a max collapse probability $\beta_{\max}(R)$ at redshift z_c
 - In Gaussian theory, $\beta(R) \sim \exp[-\delta_{\min}^2/(2\sigma^2(R))]$. Thus β_{\max} fixes the largest allowable $\sigma(R)$
 - Since $\sigma^2(R)$ is essentially an integral over $\mathcal{P}_{\mathcal{R}}(k)$ around $k \sim 1/R$, limiting $\sigma(R) \Rightarrow$ upper limit on $\mathcal{P}_{\mathcal{R}}(k)$

The primordial power spectrum

- Assume a generalised power spectrum

- EDOs formed with

$$\frac{R(z_c)}{\text{pc}} = 0.019 \left(\frac{1000}{z_c + 1} \right) \left(\frac{M(z_c)}{M_\odot} \right)^{1/3} \quad \text{with } M(z_c) = \frac{z_{\text{eq}} + 1}{z_c + 1} M_i$$

The primordial power spectrum

- Assume a generalised power spectrum

- EDOs formed with

$$\frac{R(z_c)}{\text{pc}} = 0.019 \left(\frac{1000}{z_c + 1} \right) \left(\frac{M(z_c)}{M_\odot} \right)^{1/3} \quad \text{with } M(z_c) = \frac{z_{\text{eq}} + 1}{z_c + 1} M_i$$



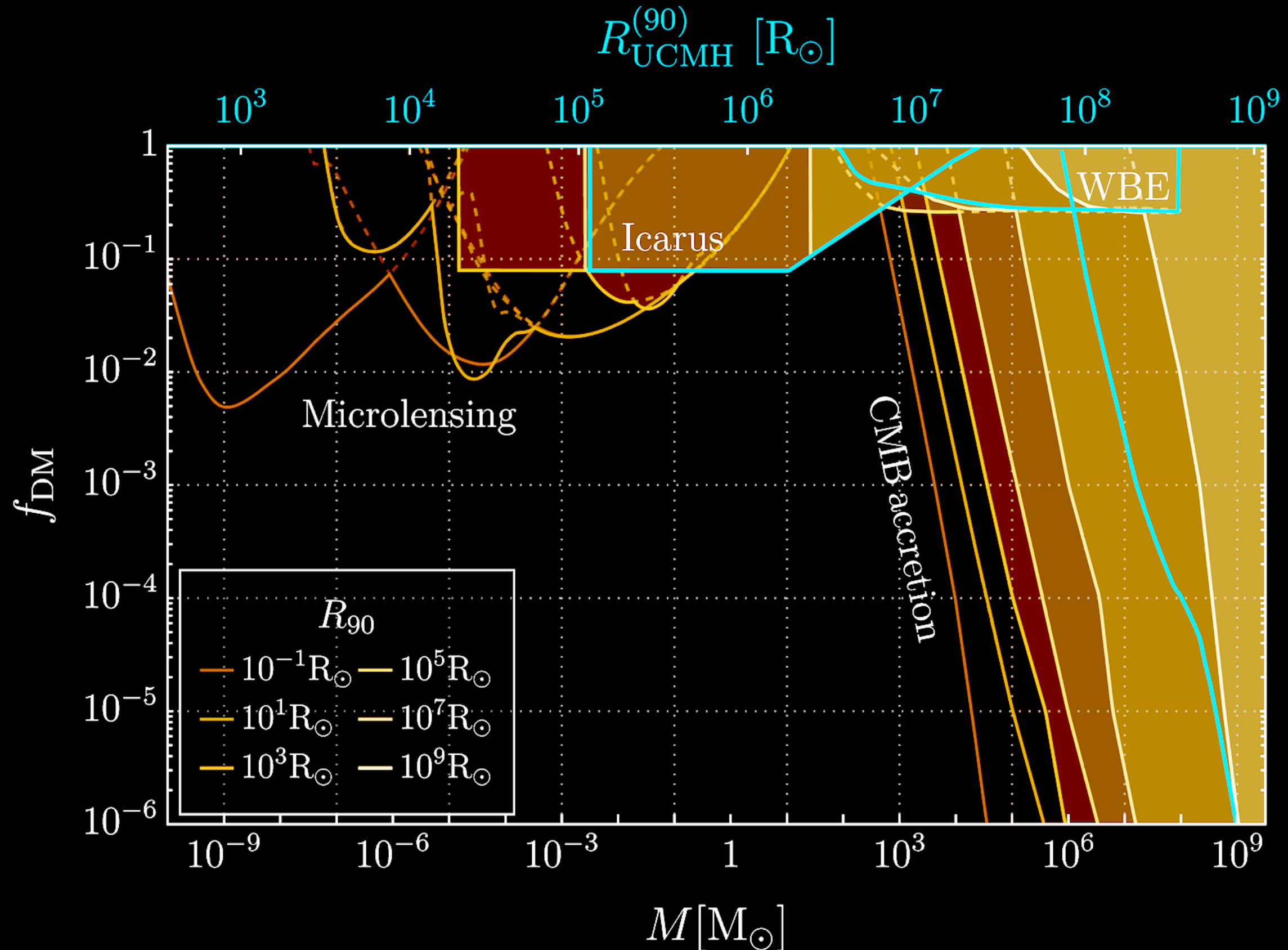
This means not all EDO constraints map to primordial power spectrum constraints

- $f_{\text{DM}} < 1$ for



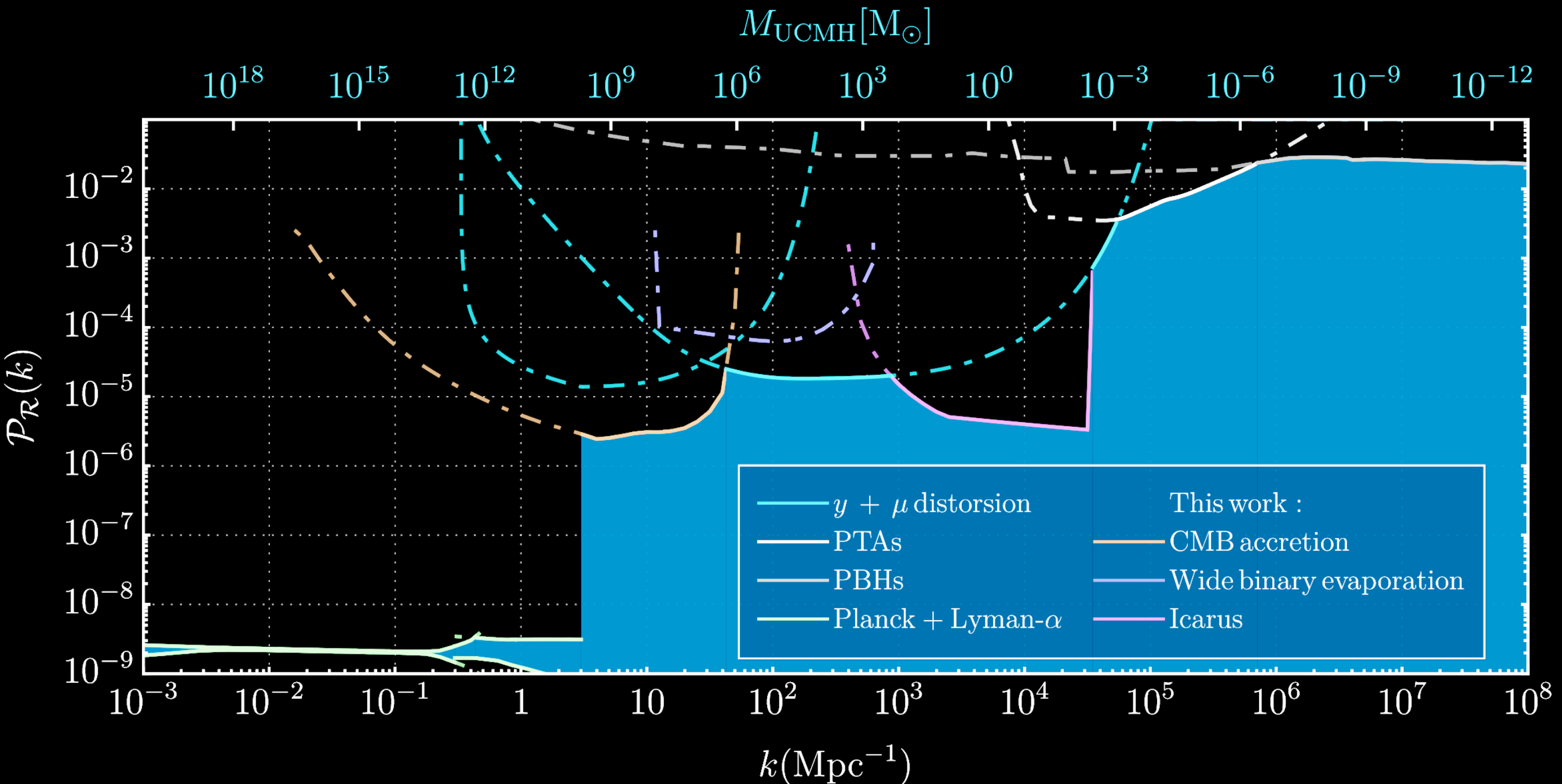
- CMB accretion (generalised to larger EDOs)
- Wide binary evaporation
- “ICARUS” microlensing (generalised from PBHs to EDOs)

The primordial power spectrum



The primordial power spectrum

$$\text{for UCMH collapsing at } z_c, \quad \frac{R(z_c)}{\text{pc}} = 0.019 \left(\frac{1000}{z_c + 1} \right) \left(\frac{M(z_c)}{M_\odot} \right)^{1/3} \quad \text{with } M(z_c) = \frac{z_{\text{eq}} + 1}{z_c + 1} M_i$$



To conclude,

- All of our current evidence for Dark Matter is gravitational; many dark matter models feature substructure
- Microlensing provides a way to look for dark matter substructure of a large range of sizes and masses
 - Extended objects may give **unique microlensing signatures**
 - Non-observation can be used to derive **constraints**

To conclude,

- All of our current evidence for Dark Matter is gravitational; many dark matter models feature substructure
- Microlensing provides a way to look for dark matter substructure of a large range of sizes and masses
 - Extended objects may give **unique microlensing signatures**
 - Non-observation can be used to derive **constraints**
- Microlensing signatures of extended objects can be distinguished using **machine learning**
 - Anomaly detection **finds events early** in wide-field surveys

To conclude,

- All of our current evidence for Dark Matter is gravitational; many dark matter models feature substructure
- Microlensing provides a way to look for dark matter substructure of a large range of sizes and masses
 - Extended objects may give **unique microlensing signatures**
 - Non-observation can be used to derive **constraints**
- Microlensing signatures of extended objects can be distinguished using **machine learning**
 - Anomaly detection **finds events early** in wide-field surveys
- Gravitational constraints on EDOs lead to constraints on the primordial power spectrum

Many opportunities for future work!

Thank you!

...ask me anything you like!

djuna.l.croon@durham.ac.uk | djunacroon.com

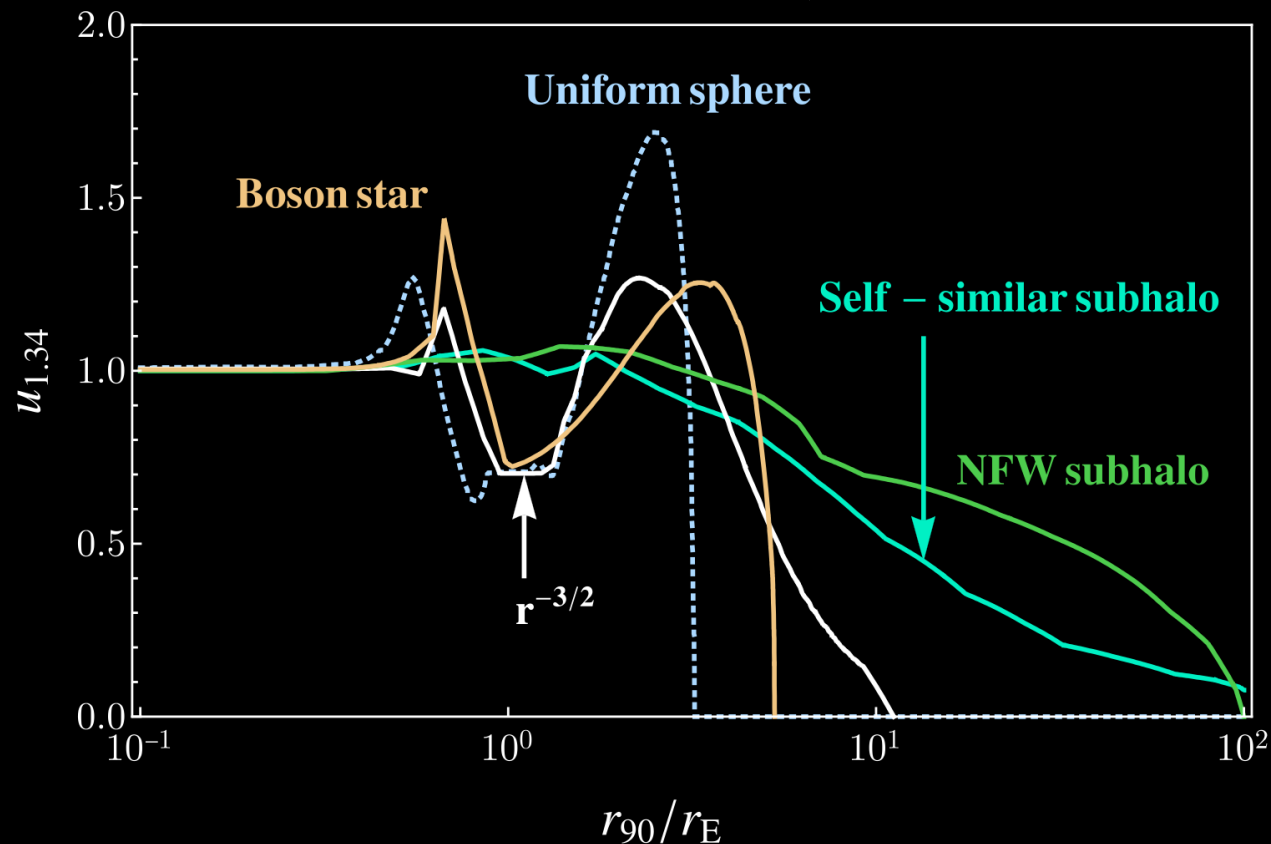
Back up slides

Extended sources: $r_E = \theta_E D_L \sim r_S$

Same procedure as before, but now $u_{1.34}$ is a function of both r_{90} and r_S

DC, D. McKeen, N. Raj, Z. Wang, PRD, arXiv:2007.12697 [astro-ph.CO]

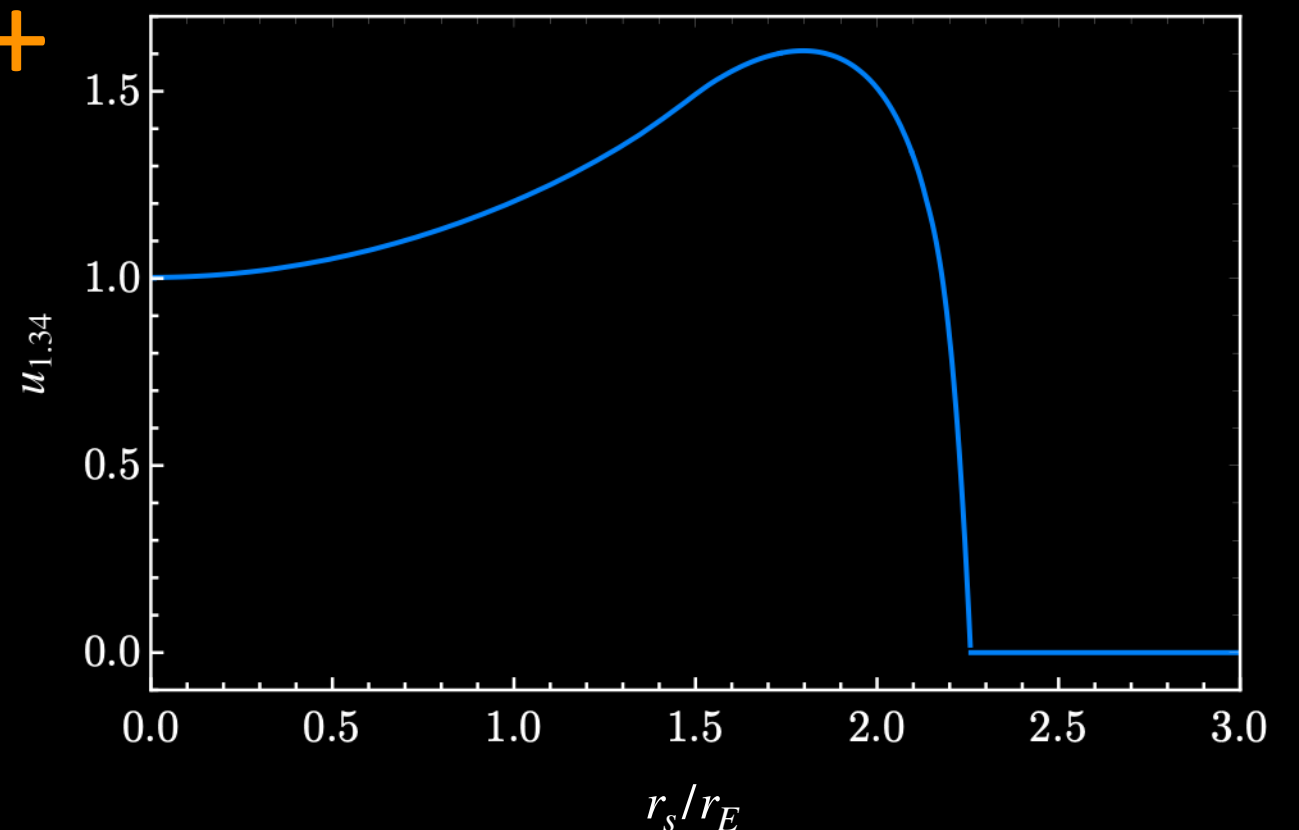
From before: Pointsource, extended lens



For point-like lenses, see for example,
 Witt and Mao, *Astrophys. J* (1994);
 Montero-Camacho, Fang, Vasquez, Silva, Hirata,
 [JCAP, arXiv:1906.05950];
 Smyth, Profumo, English, Jeltema, McKinnon,
 Guhathakurta [PRD, arXiv:1910.01285];

+

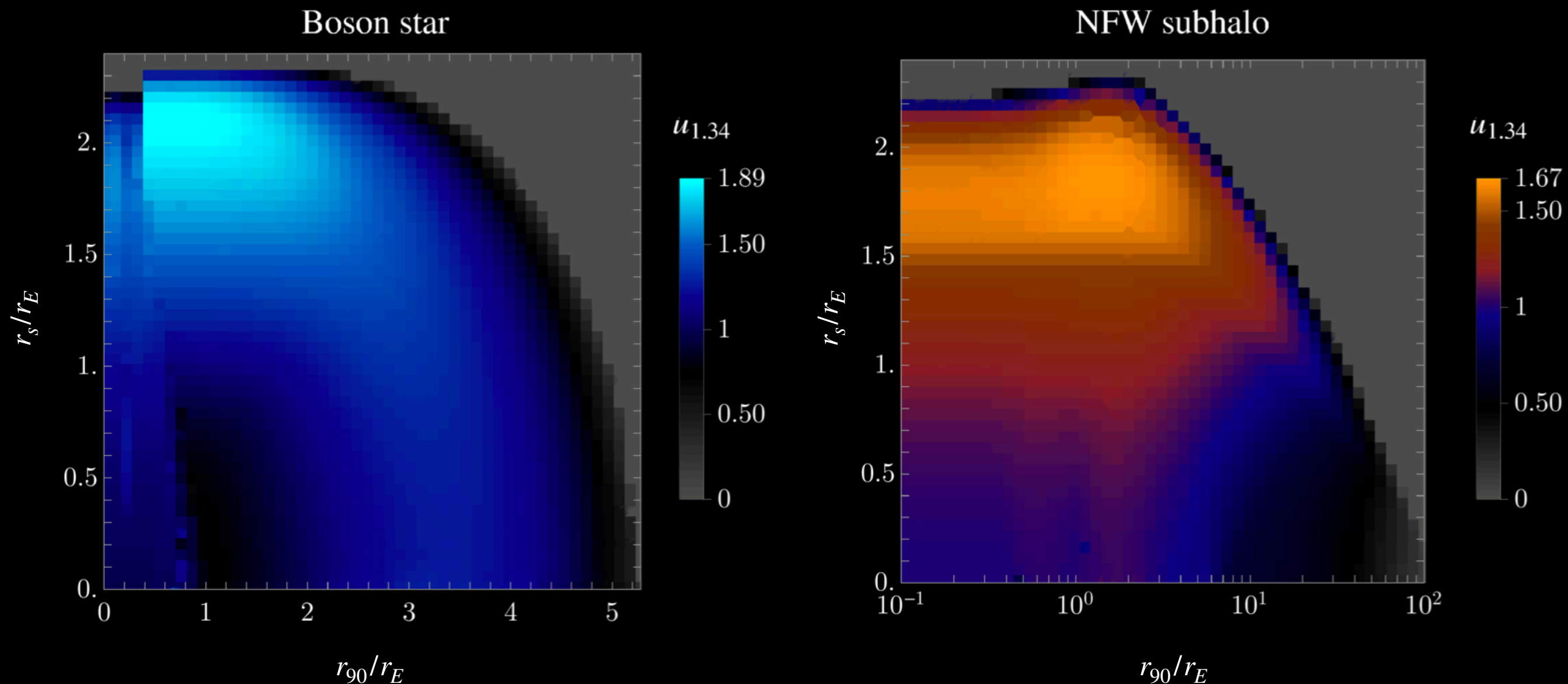
Point-like lens, extended source



Extended sources: $r_E = \theta_E D_L \sim r_S$

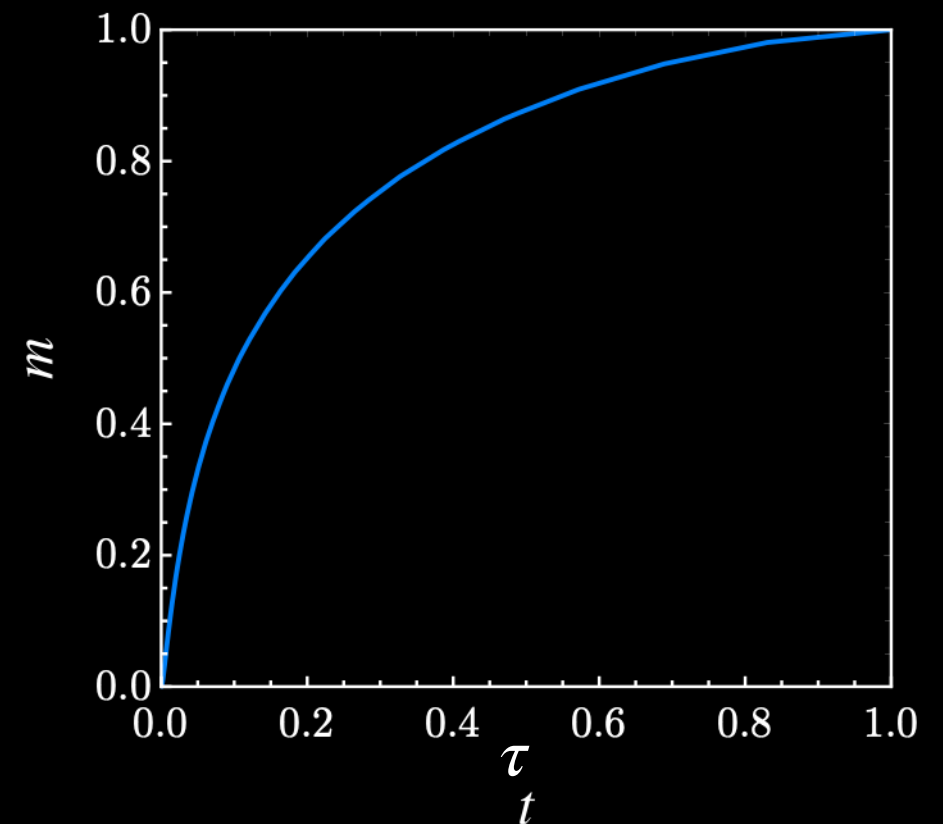
Same procedure as before, but now $u_{1.34}$ is a function of both r_{90} and r_S

DC, D. McKeen, N. Raj, Z. Wang, PRD, arXiv:2007.12697 [astro-ph.CO]



Case study 1: NFW-halo mass profile

- Well-known halo profile: $\rho(r) = \frac{\rho_s}{(r/r_s)(1 + r/r_s)^2}$
- As the mass inclosed formally diverges, we cut it off at $R_{\text{cut}} = 100R_{\text{sc}}$
- Enclosed mass $\propto \log(\kappa + 1) - (\kappa/(\kappa + 1))$ where $\kappa = R_{\text{cut}}/R_{\text{sc}}$
- Computing $m(\tau)$ is then a trivial exercise:



Case study 2: Boson star mass profile

- The Schrodinger-Poisson equation,

$$\mu\Psi = -\frac{1}{2m_\phi}\left(\Psi'' + \frac{2}{r}\Psi'\right) + m_\phi\Phi\Psi$$

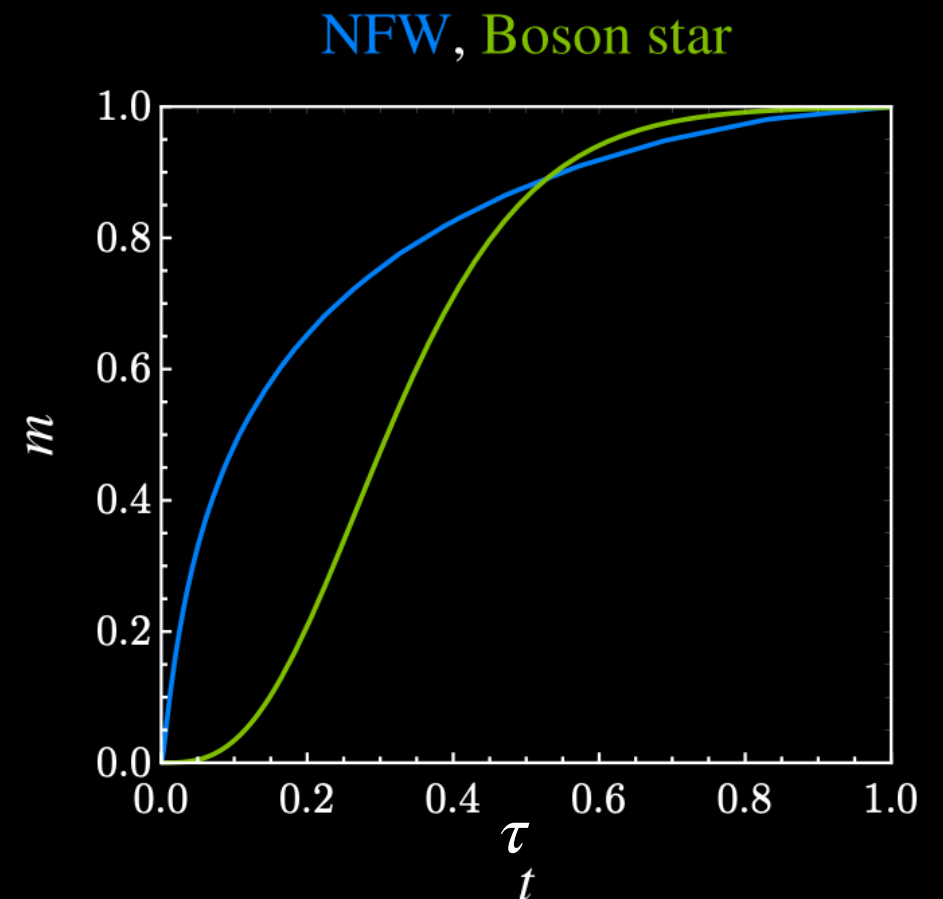
Describes the radial distribution

describes a *spherically symmetric ground state of a free scalar field in the non-relativistic limit*

- The mass enclosed is given by

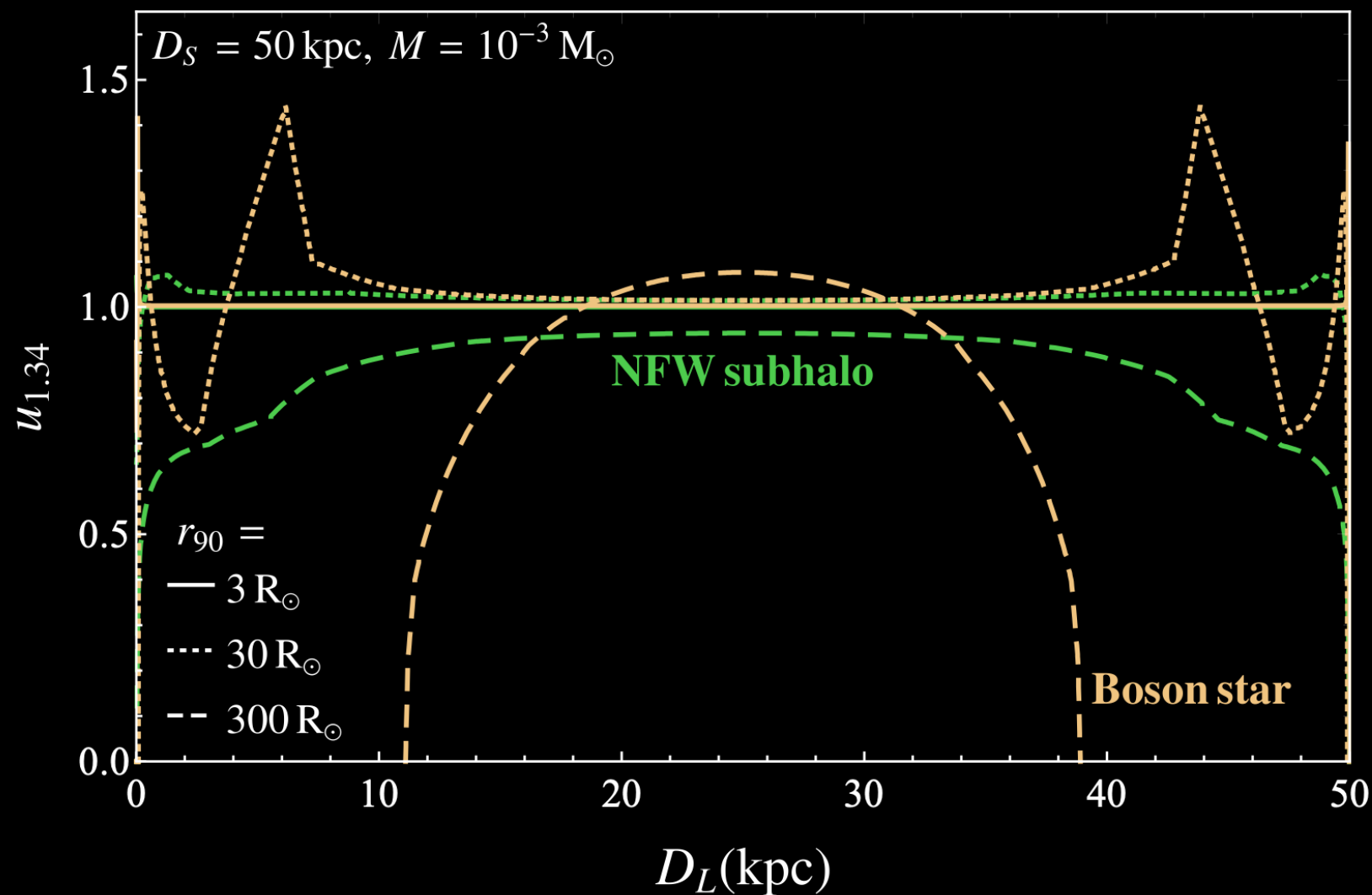
$$M_{\text{BS}}(r) = \frac{1}{m_\phi G} \int_0^{m_\phi r} dy y^2 \Psi^2(y)$$

from which $m(\tau)$ may be computed



Caustics

Consequence: the Einstein tube is not a tube; not ellipsoidal



→ Depending on the source, experiments may be more or less sensitive to extended objects compared to point sources in different locations

Obtaining constraints


To obtain limits, we have to account for the observed events

- EROS-2: 3.9 events at 90% CL
- OGLE-IV: $\mathcal{O}(1000)$ astrophysical events, Poissonian 90% CL:
 $\kappa = 4.61$

Bin events in t_E

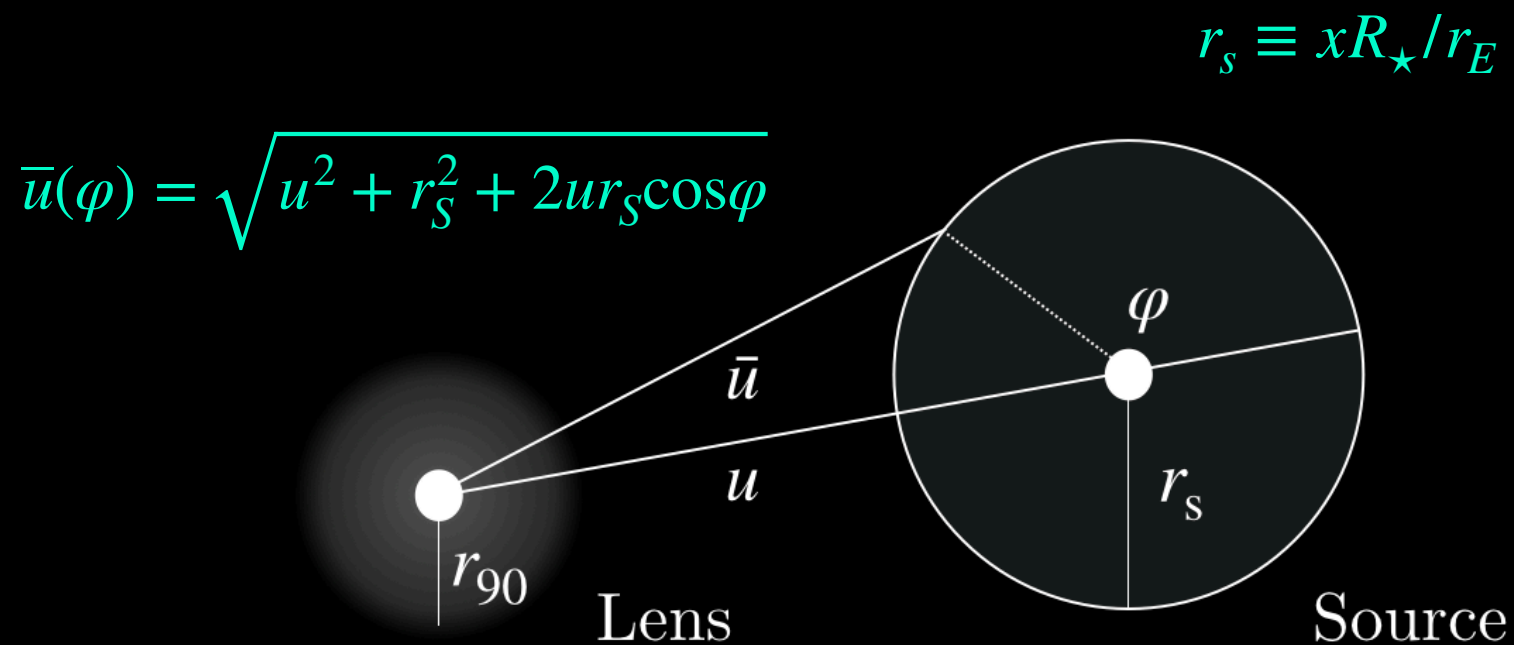
$$\kappa = 2 \sum_{i=1}^{N_{\text{bins}}} \left[N_i^{\text{FG}} - N_i^{\text{SIG}} + N_i^{\text{SIG}} \ln \frac{N_i^{\text{SIG}}}{N_i^{\text{FG}}} \right]$$

$N_i^{\text{SIG}} \equiv N_i^{\text{FG}} + N_i^{\text{DM}}$



Lensing geometry

- Up to this point, we have assumed that the sources are point-like light sources (a good approximation for EROS/OGLE)
- This approximation breaks down when $r_E = \theta_E D_L \sim r_S$
- Geometry in the lens plane:



Lensing equation:

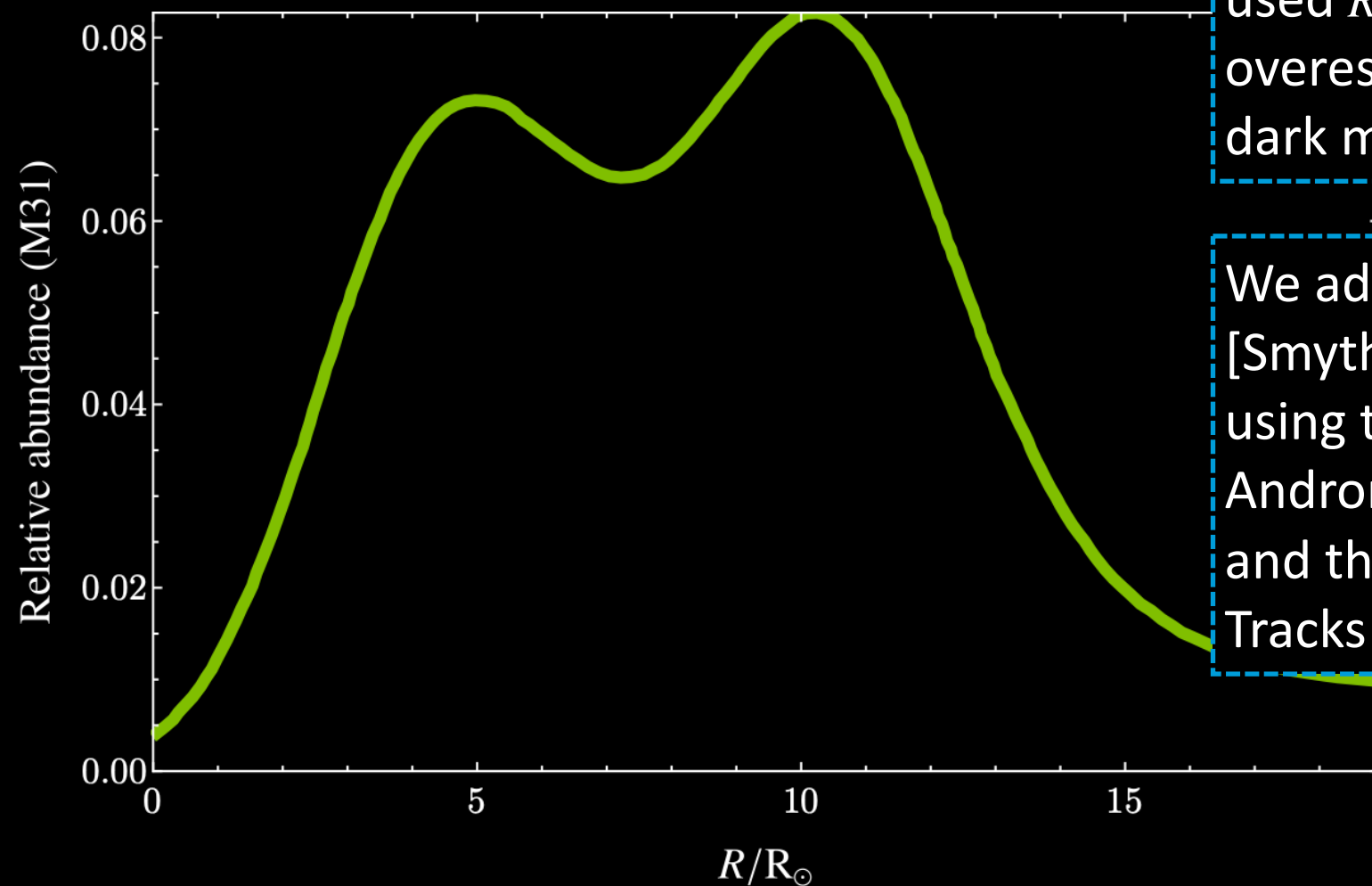
$$\bar{u}(\varphi) = \tau(\varphi) - \frac{m(\tau(\varphi))}{\tau(\varphi)}$$

Image
parity

Image
position

$$\mu_i = \eta \frac{1}{\pi r_s^2} \int_0^{2\pi} d\varphi \frac{1}{2} \tau_i^2(\varphi)$$

Star sizes in M31



Initially, the Subaru-HSC collaboration used $R = R_{\odot}$ for all stars, but this overestimates the constraints on the dark matter fraction

We adopt the distribution derived in [Smyth et al., PRD, arXiv:1910.01285] using the Panchromatic Hubble Andromeda Treasury star catalogue and the MESA Isochrones and Stellar Tracks stellar evolution package

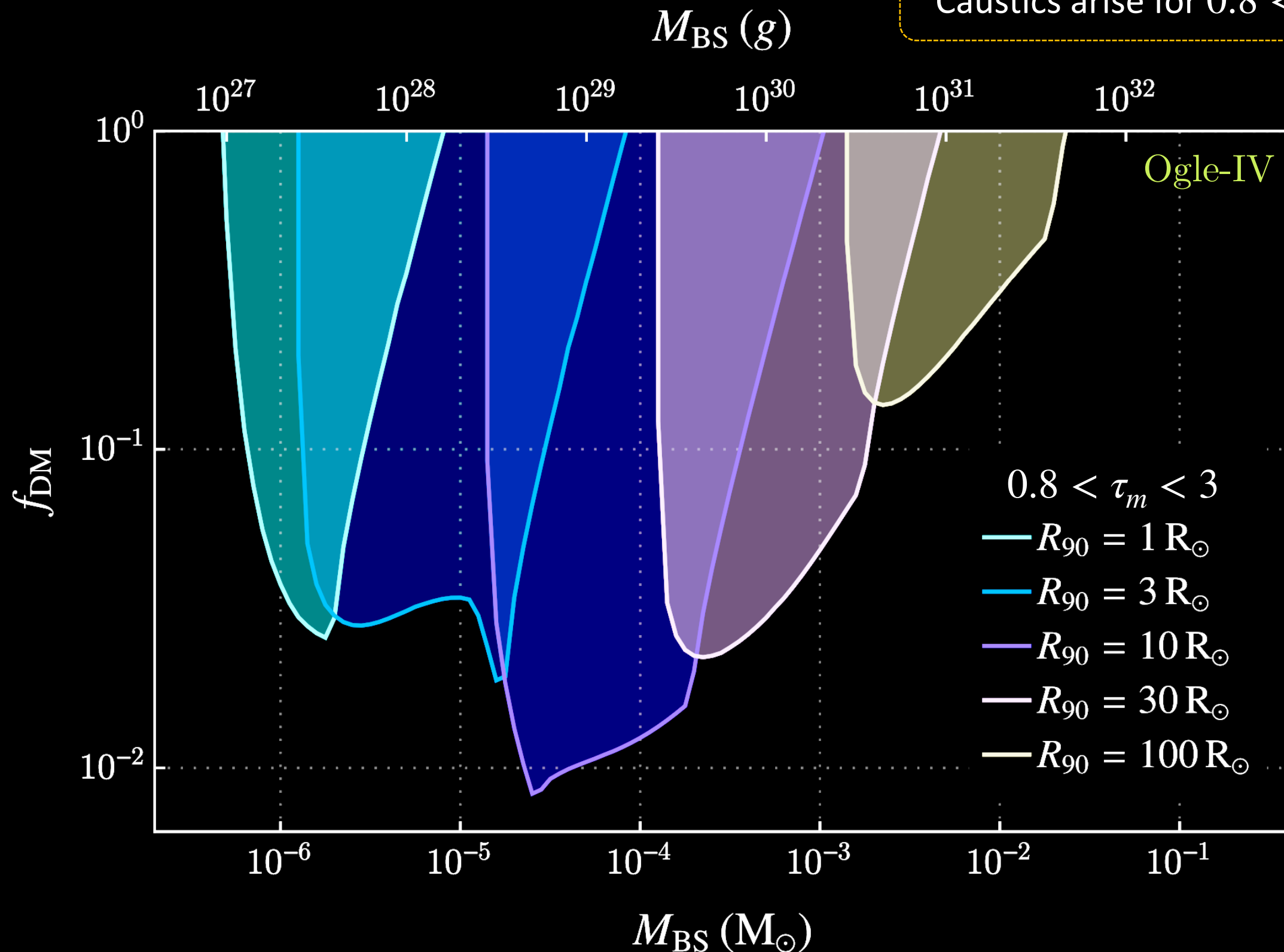


$$N_{\text{events}} = N_{\star} T_{\text{obs}} \int dt_{\text{E}} \int dR_{\star} \int_0^1 dx \frac{d^2\Gamma}{dx dt_{\text{E}}} \frac{dn}{dR_{\star}}$$

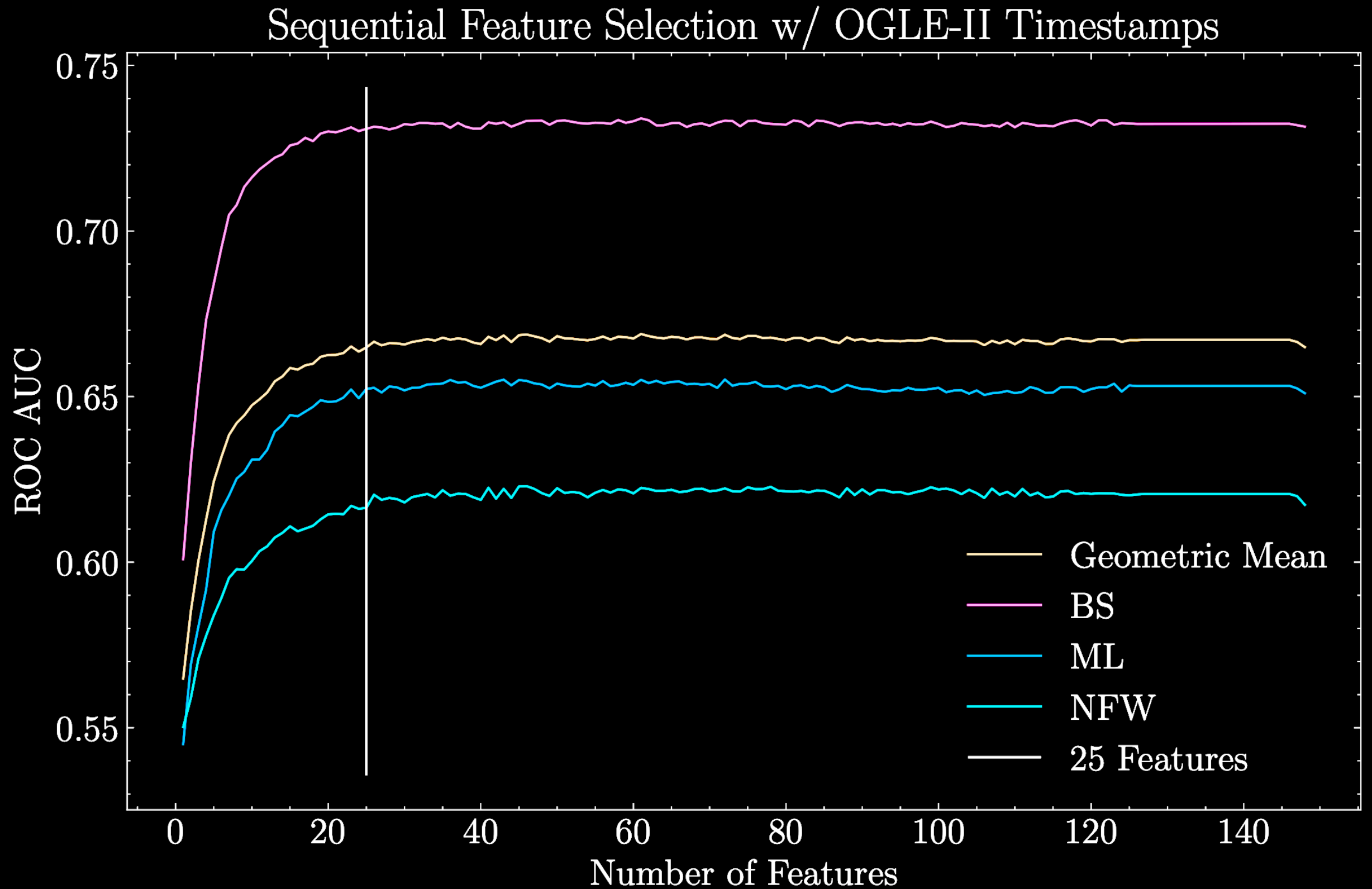
Opportunities for positive detection

M. Crispim-Romao, DC, PRD, arXiv:2402.00107

Caustics arise for $0.8 < \tau_m < 3$



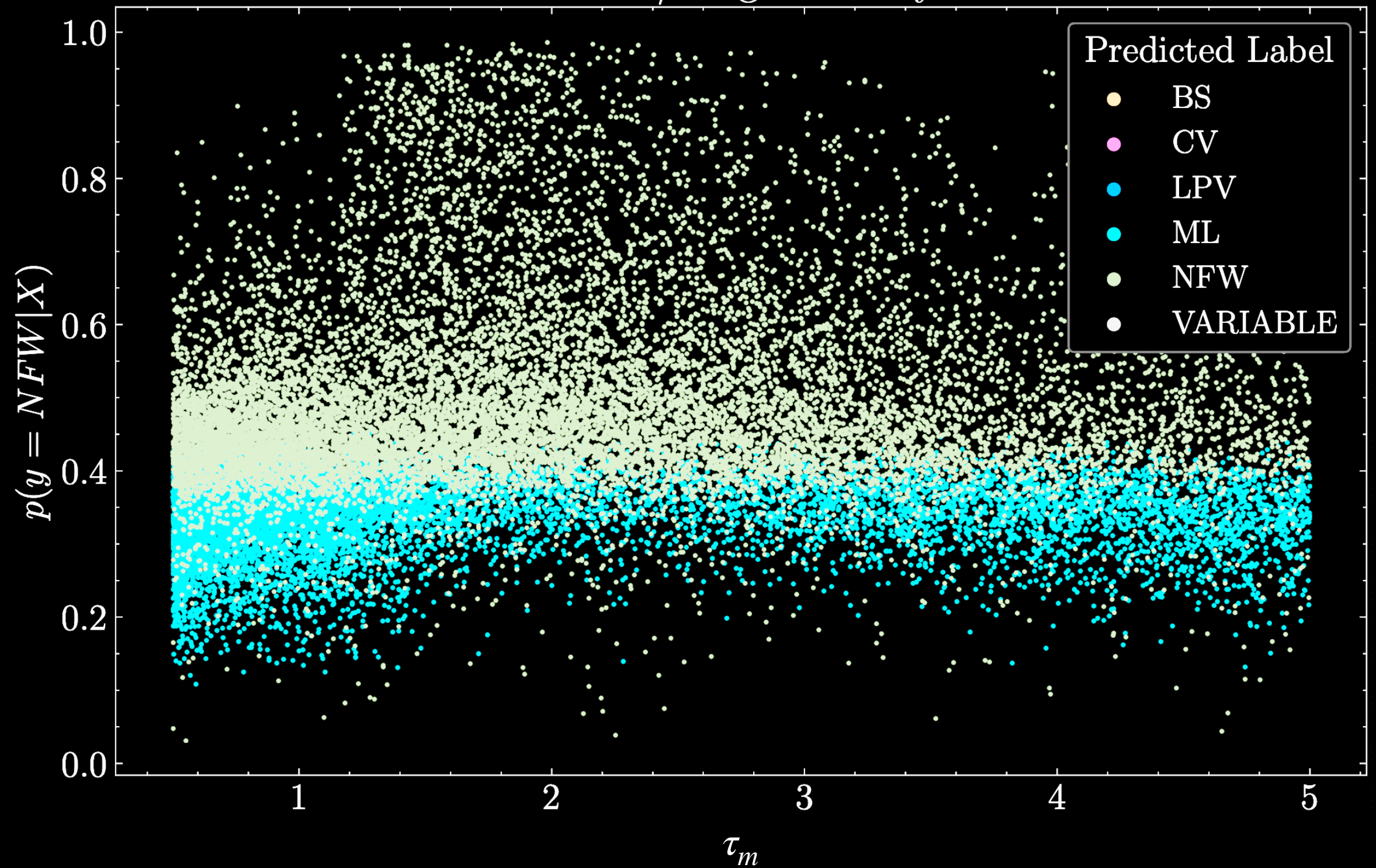
Feature importance

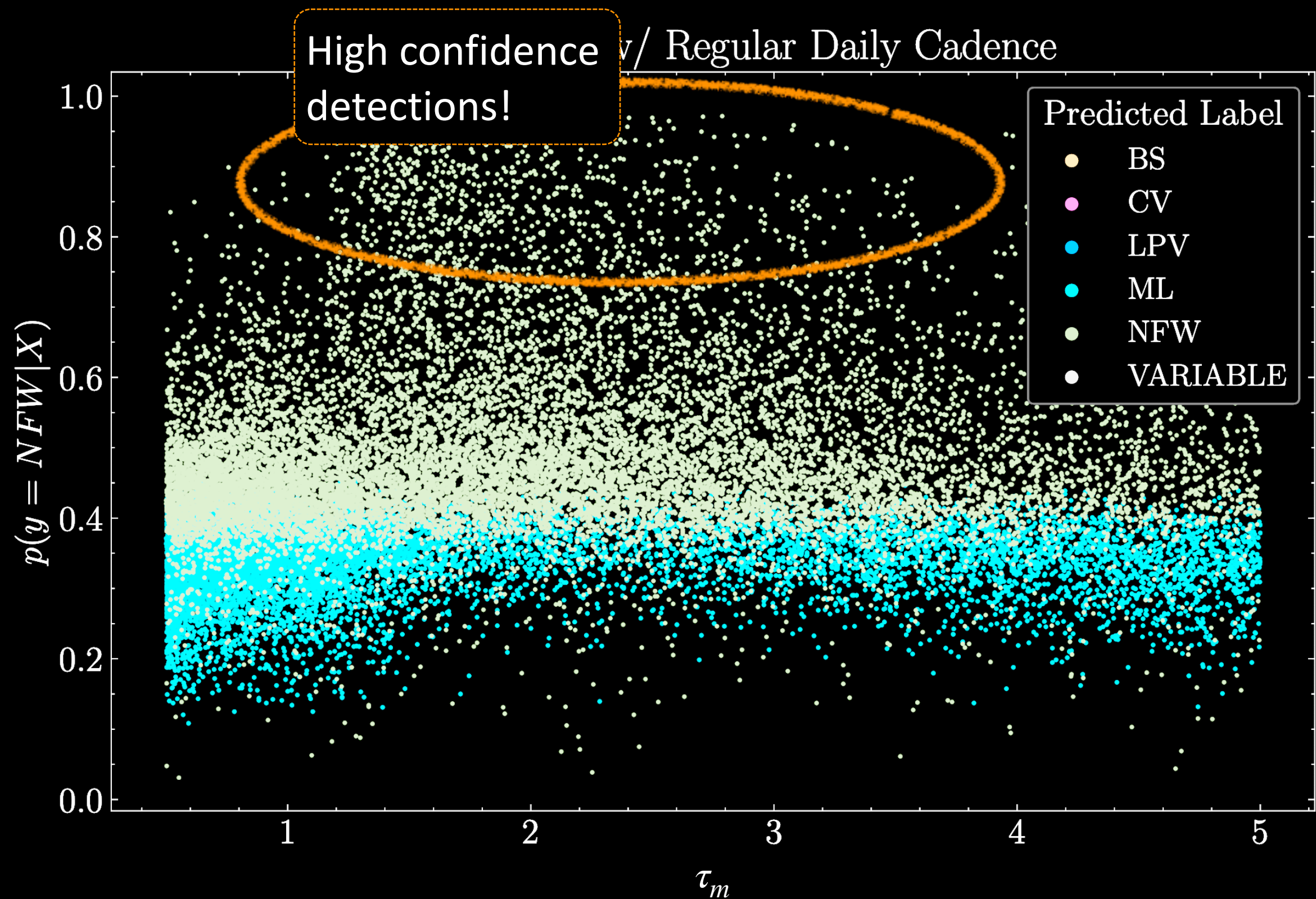


Let's dream...

- The OGLE time steps are quite irregular
- Many different factors play a role...
 - Observational Constraints (weather, moon phase, ...)
 - Resource Allocation
 - Target Prioritization
 - Technical Maintenance and Downtime
- But it is interesting what the effect of cadence (ir)regularity is on the observational prospects
- So, let us imagine for a moment that we could achieve perfect daily cadence

NFW Events w/ Regular Daily Cadence





... only observed if regular cadence is achieved