The Data-Driven Approach to Modeling Backgrounds in Optical Transient Searches

Seiji TOSHIKAGE in collaboration with
Masaomi Tanaka, **Shigeo S. Kimura** (Tohoku U.),
Nobuhiro Shimizu, Wataru B. Iwakiri, Shigeru Yoshida (Chiba U.),
Tomoki Morokuma (Chiba Tech), and IceCube follow-up team

Optical counterpart search

Background characterization is essential before arrival-direction analysis

1. Astrophysical transient

Explosive transient (SNe, TDEs)

→ **Simulation-driven estimate**

2. Non-transient objects & artifacts

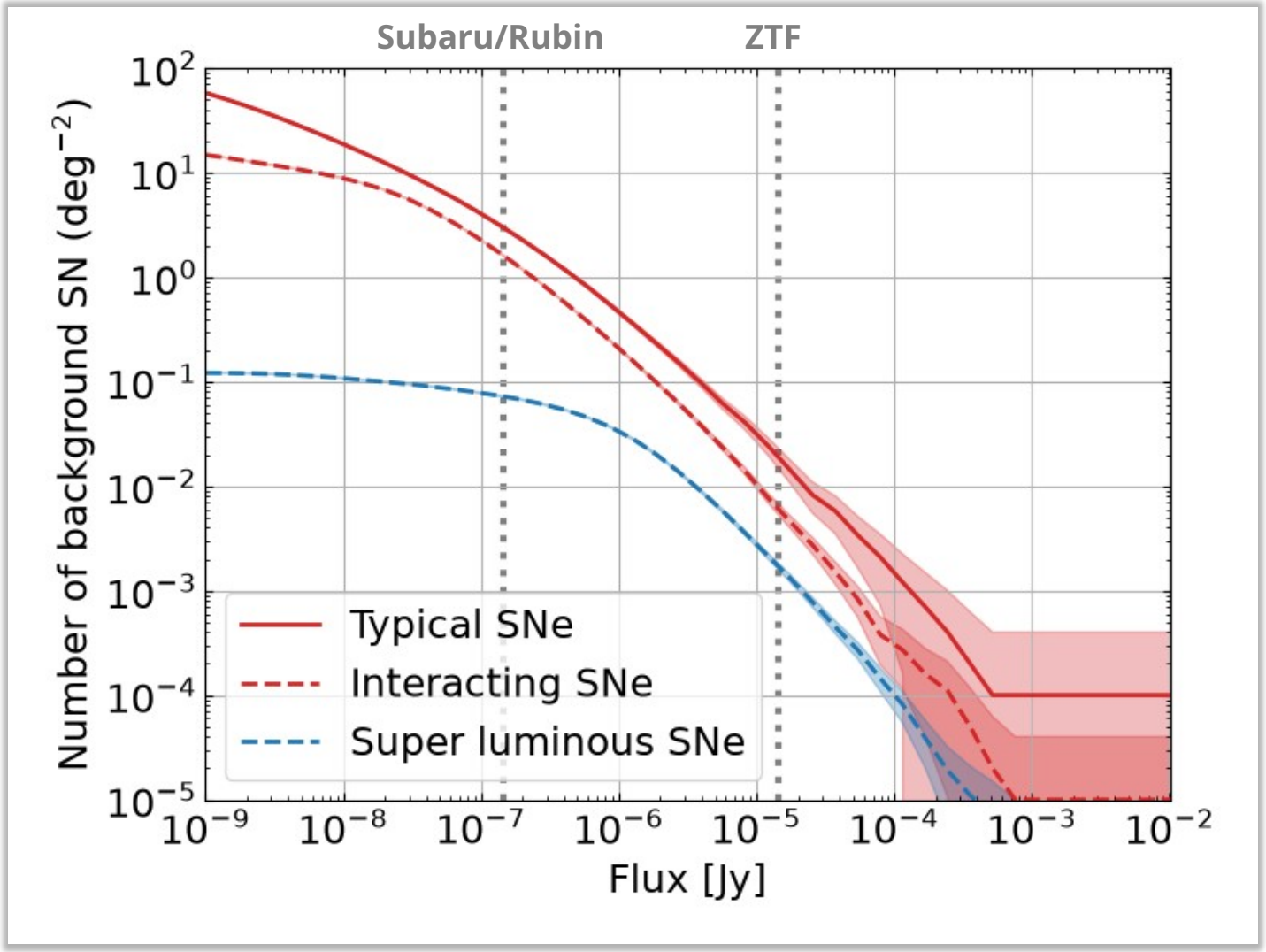
Variables 

Moving objects 

Artifacts 

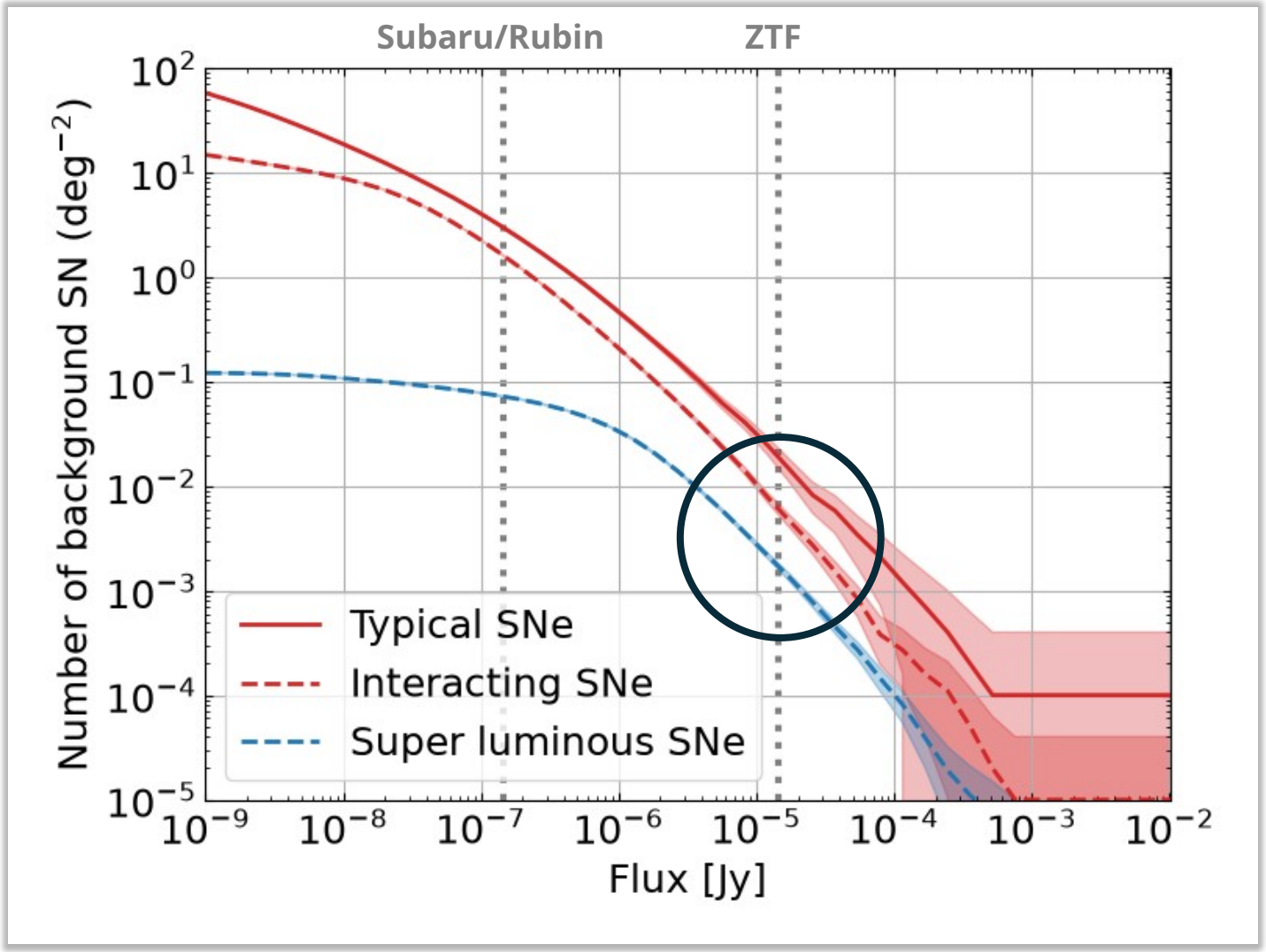
→ **Data-driven measurement**

Astrophysical transient background



Low astrophysical transient background for multiplet

Astrophysical transient background



Low astrophysical transient background for multiplet

In real data ...

Background in optical transient counterpart search

Astrophysical transient

SNe, TDEs

Difficult to simulate (poorly modeled/depend on data quality, instrument,...)
→ Data driven approach !

In real data ...

Background in optical transient counterpart search

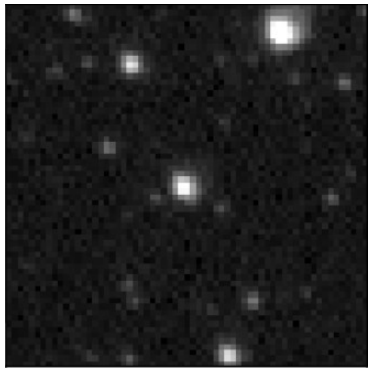
Astrophysical transient

SNe, TDEs

Astrophysical poorly-modeled background

moving objects

variable stars



Difficult to simulate (poorly modeled/depend on data quality, instrument,...)
→ Data driven approach !

In real data ...

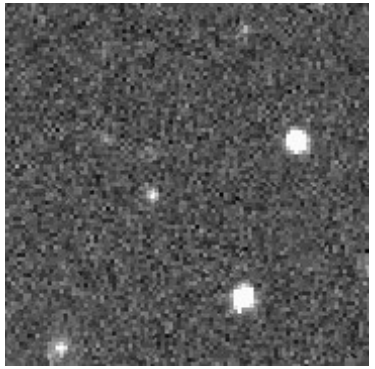
Background in optical transient counterpart search

Astrophysical transient

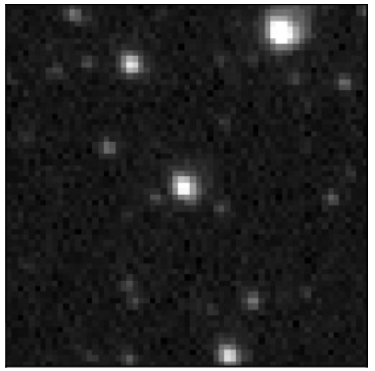
SNe, TDEs

Astrophysical poorly-modeled background

moving objects

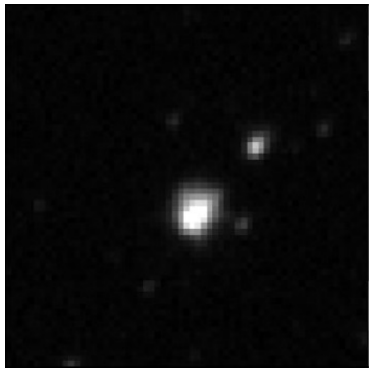


variable stars



Non-astrophysical background

artifacts

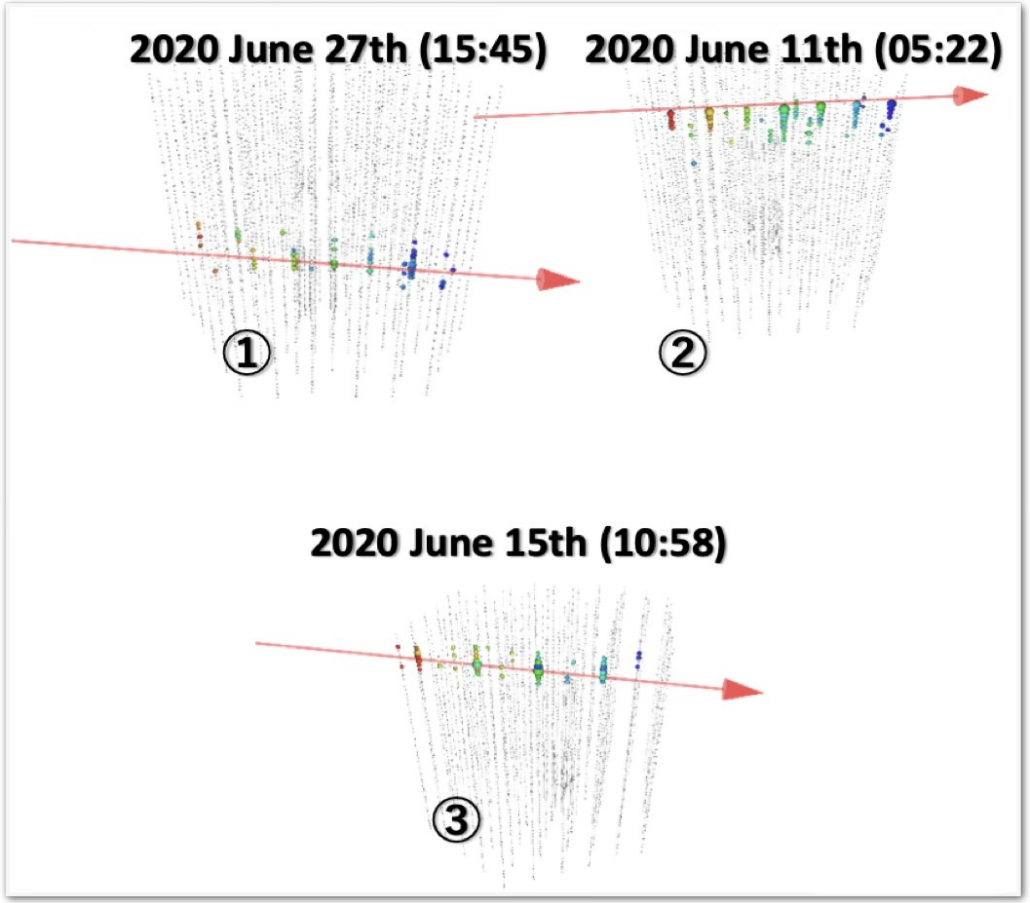


Difficult to simulate (poorly modeled/depend on data quality, instrument,...)
→ Data driven approach !

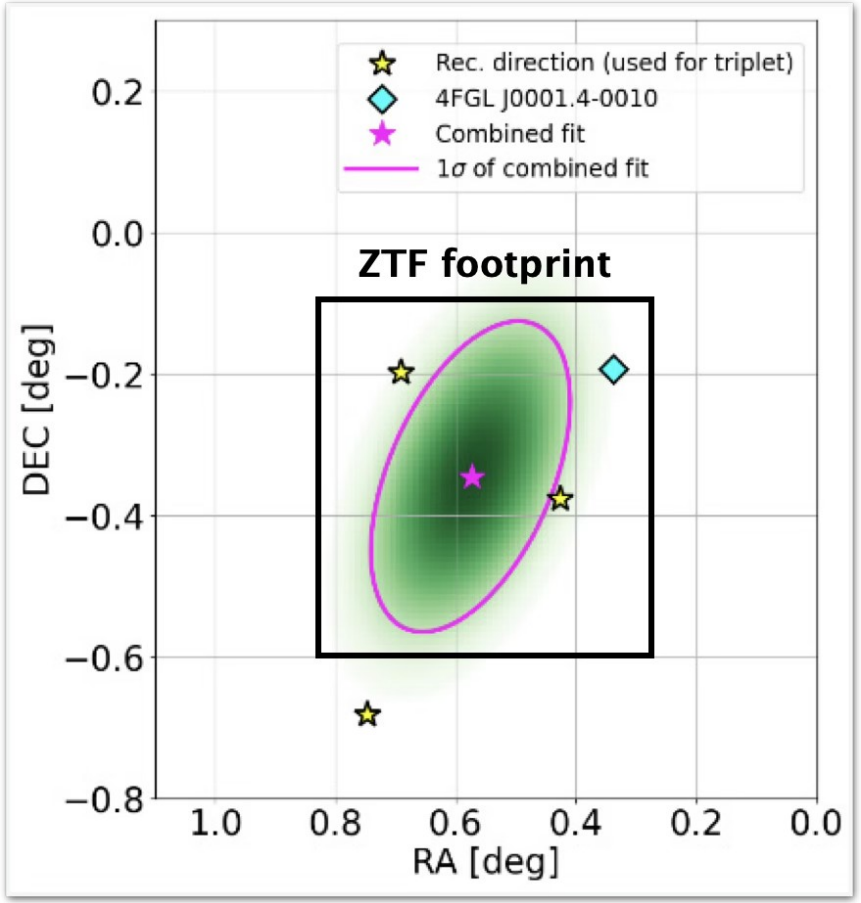
A neutrino triplet event in June 2020

IceCube collaboration 2025

Month-timescale neutrino triplet!



Serendipitous ZTF observations!

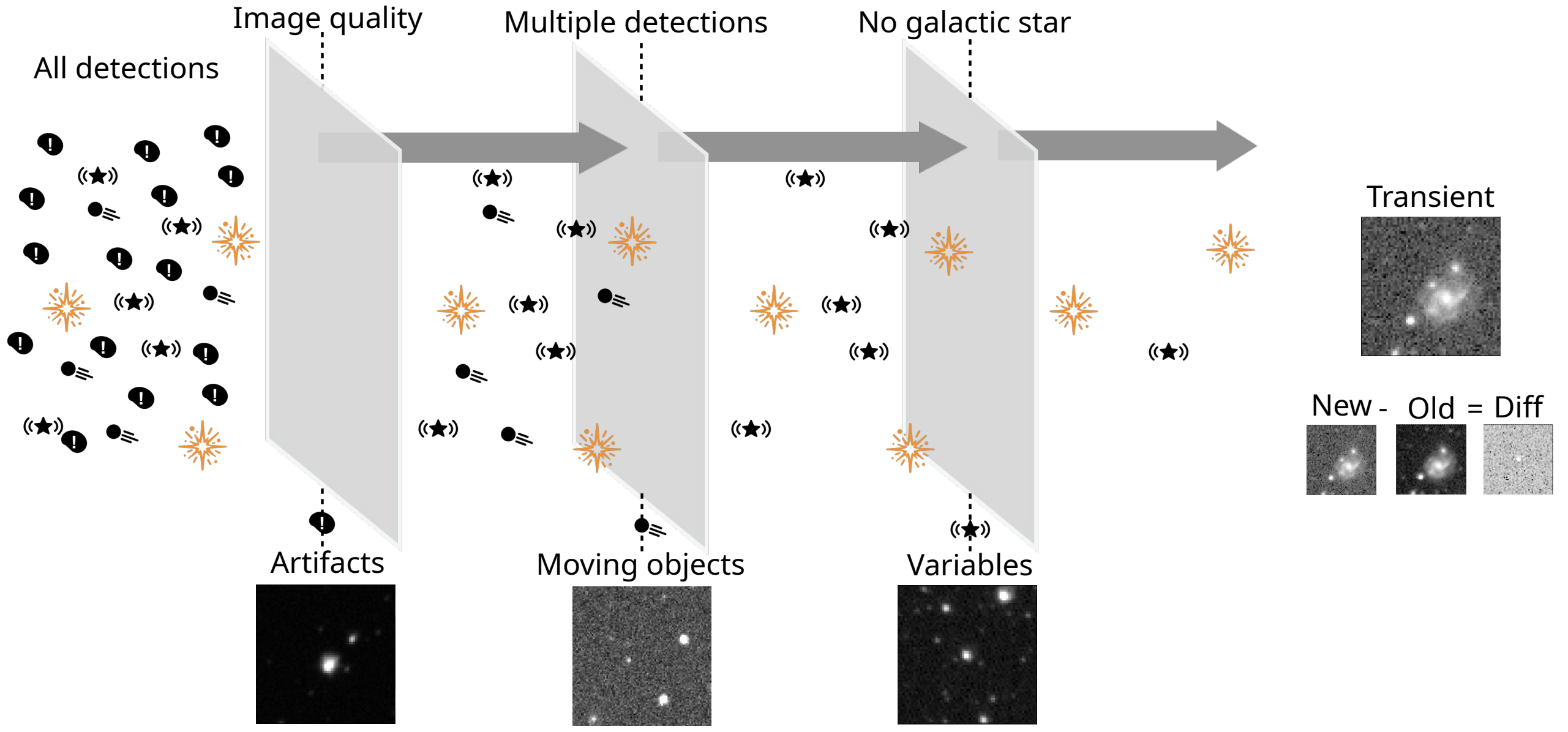


From Nobu Shimizu's slide

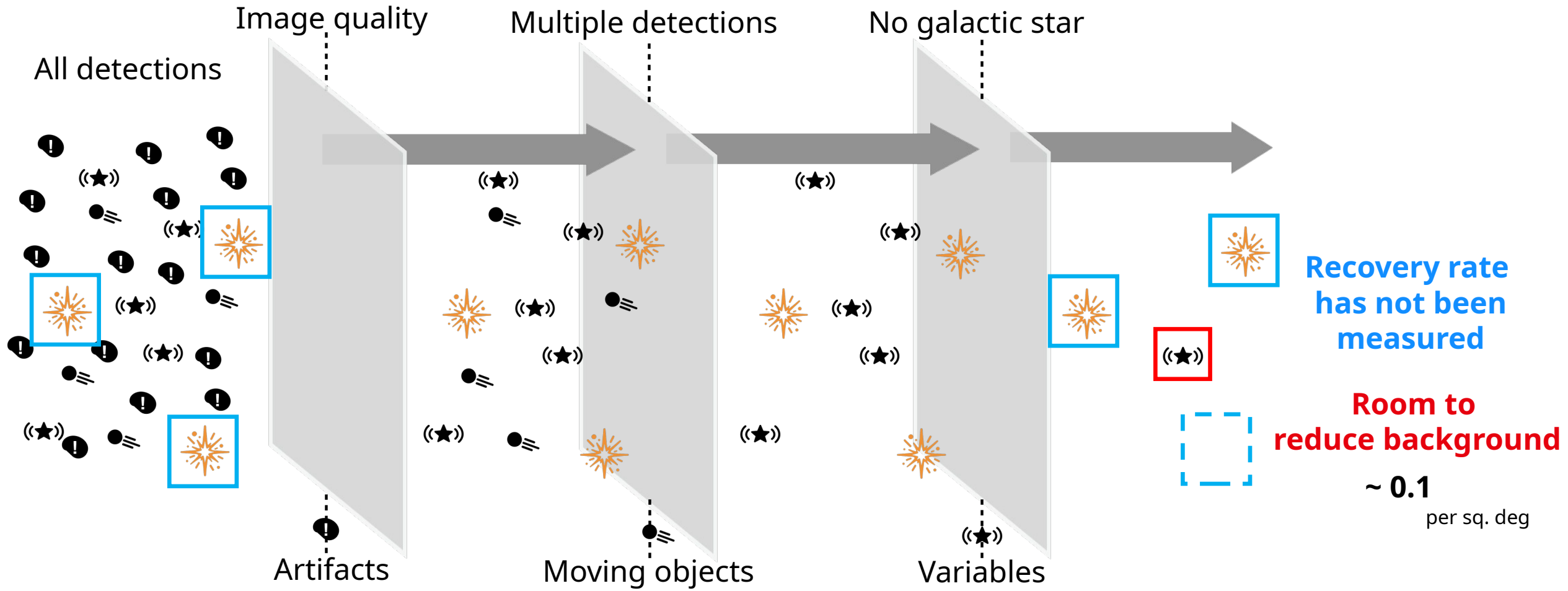
The first opportunity to search for an optical counterpart to a neutrino multiplet

Data driven approach through blind analysis

Data analysis: ZTF general filtering procedure



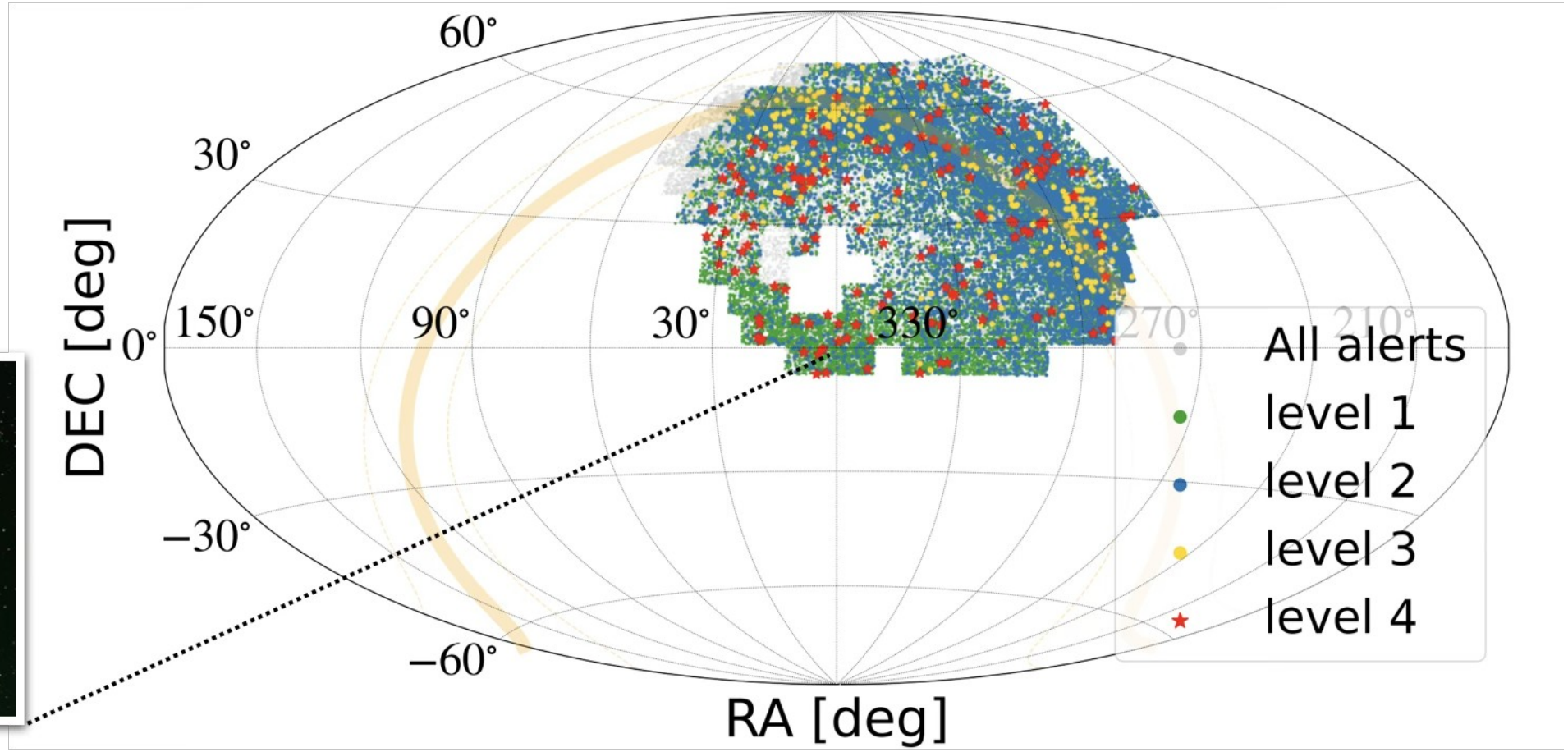
Data analysis: ZTF general filtering procedure



Optimize for month-timescale optical counterpart search

Background estimation

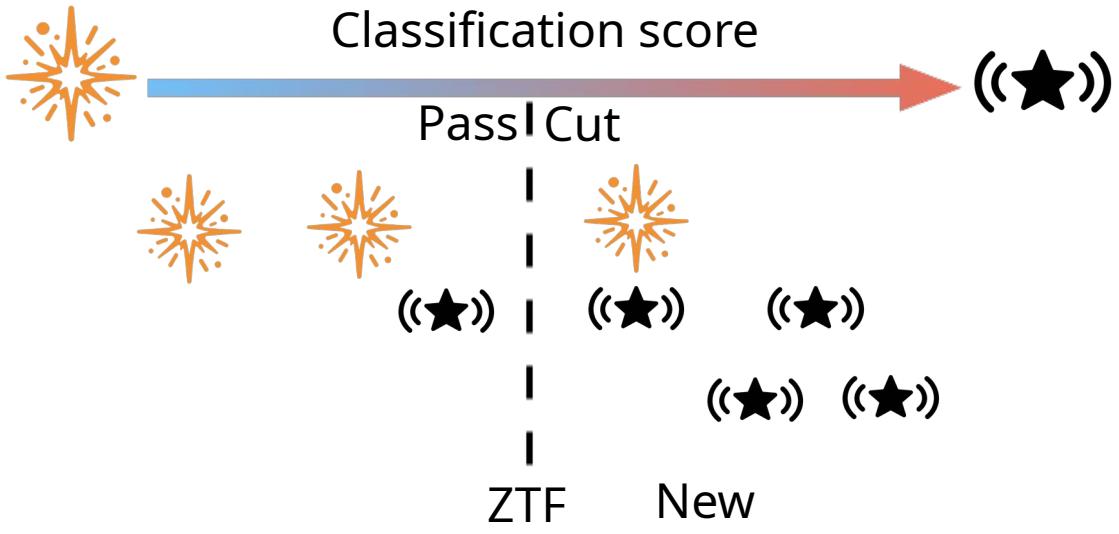
0.4 sq. deg



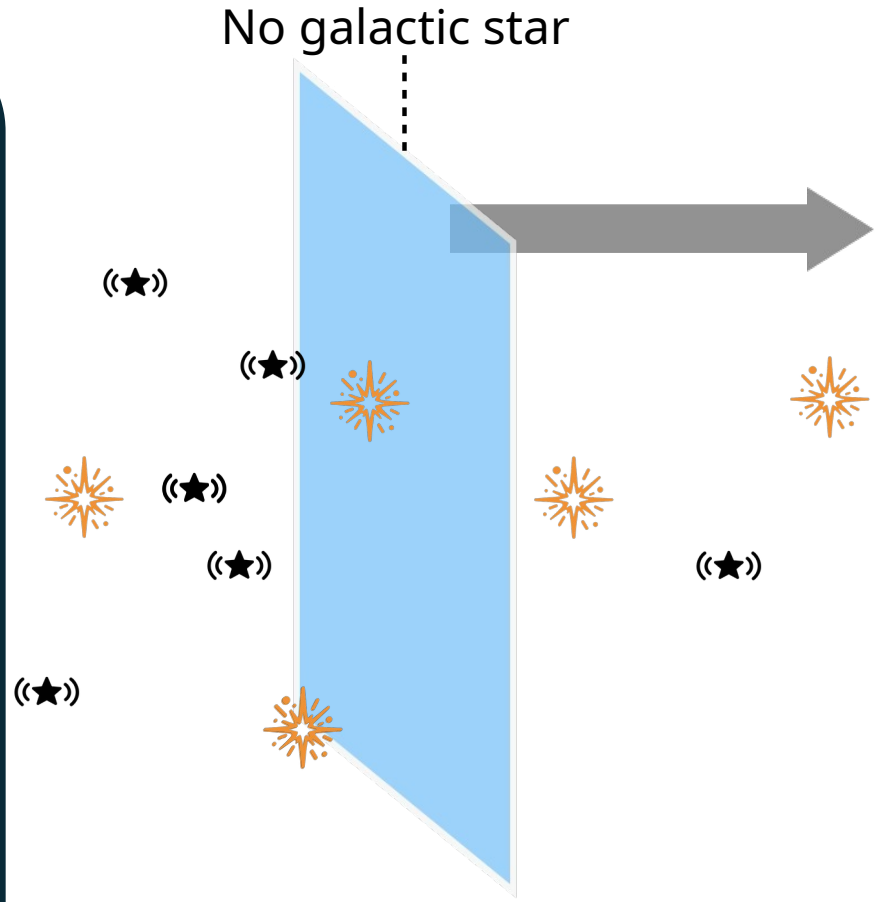
Using background area ~ 16,000 times larger than target area

Data analysis: optimized for counterpart

No galactic star criterion

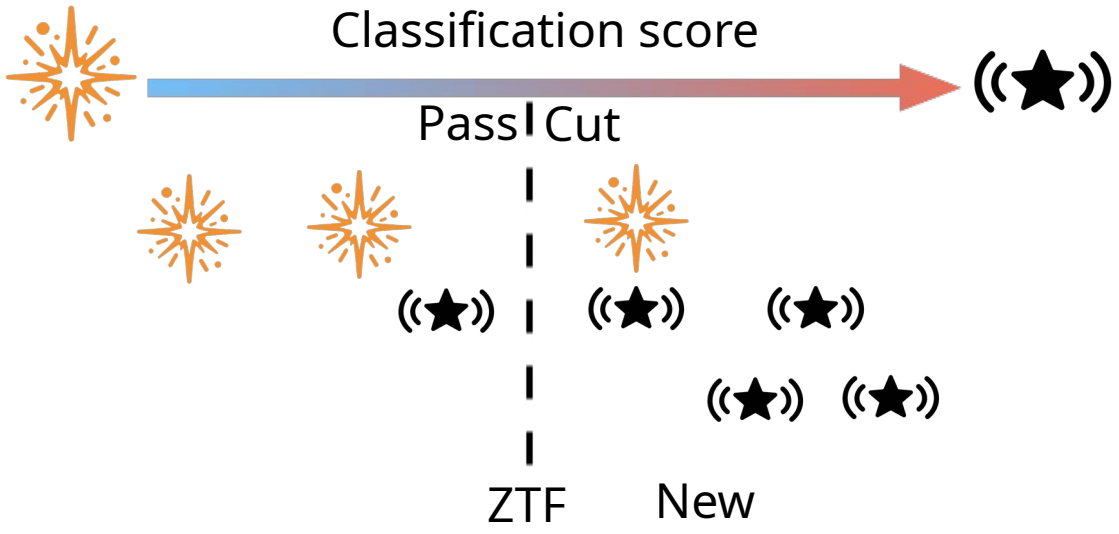


Low background rate!
BUT, low recovery rate ...

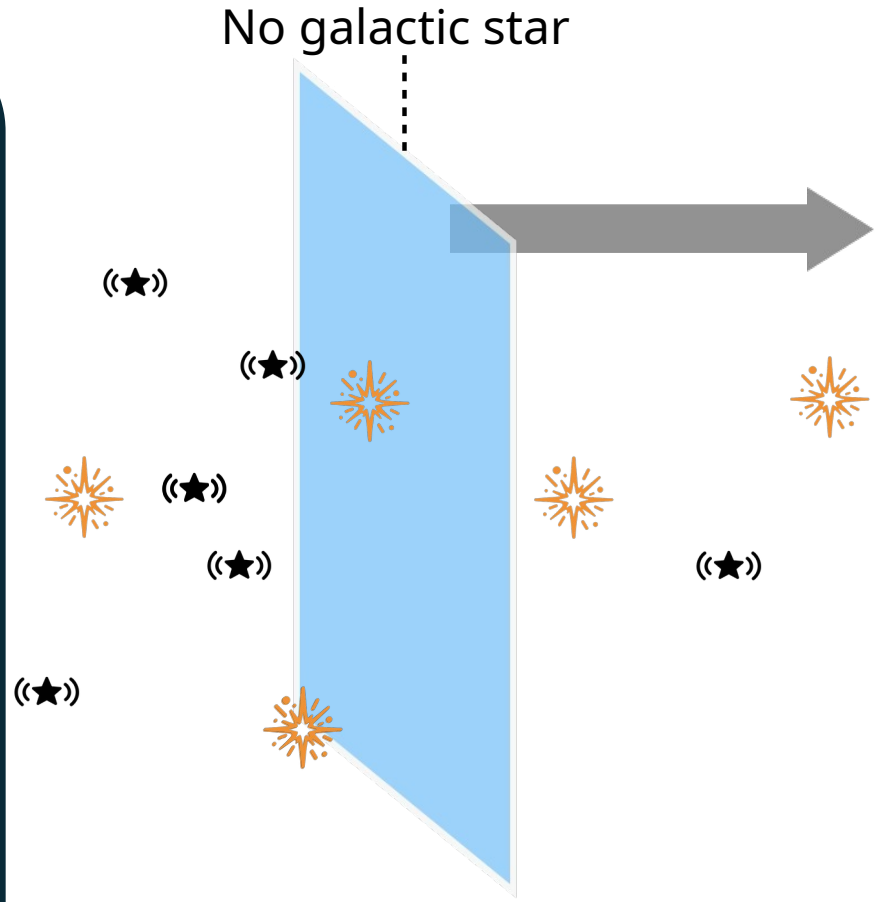


Data analysis: optimized for counterpart

No galactic star criterion

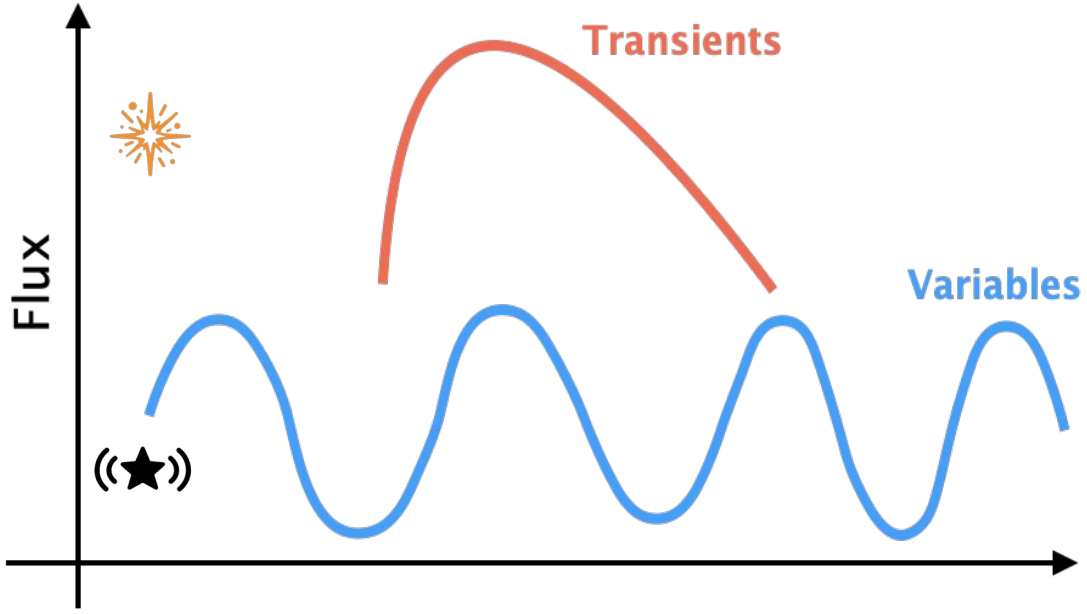


High recovery rate!
BUT, high background rate ...

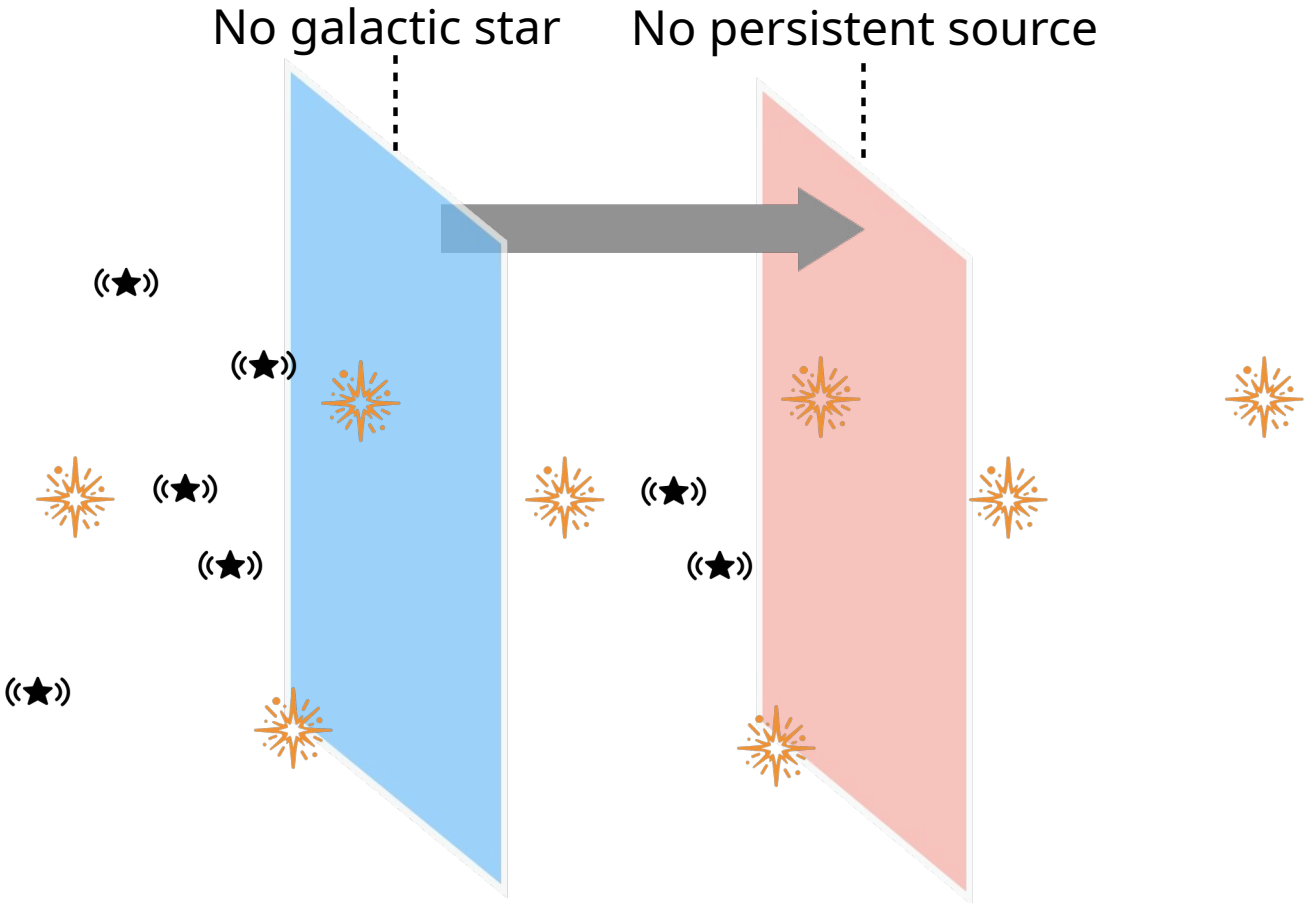


Data analysis: optimized for counterpart

No persistent source



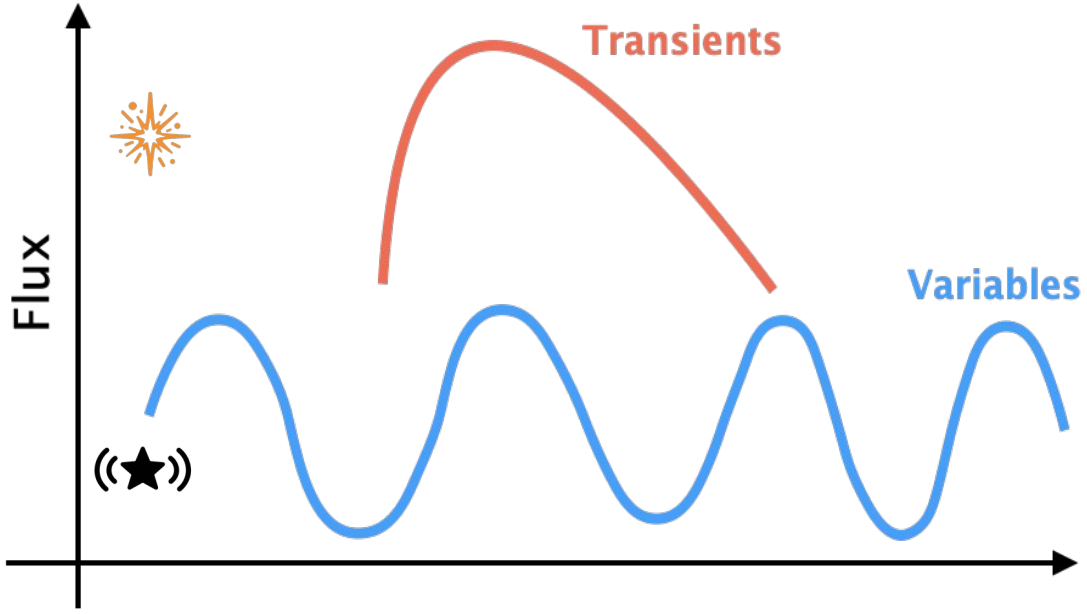
High recovery rate!
AND, low background rate !



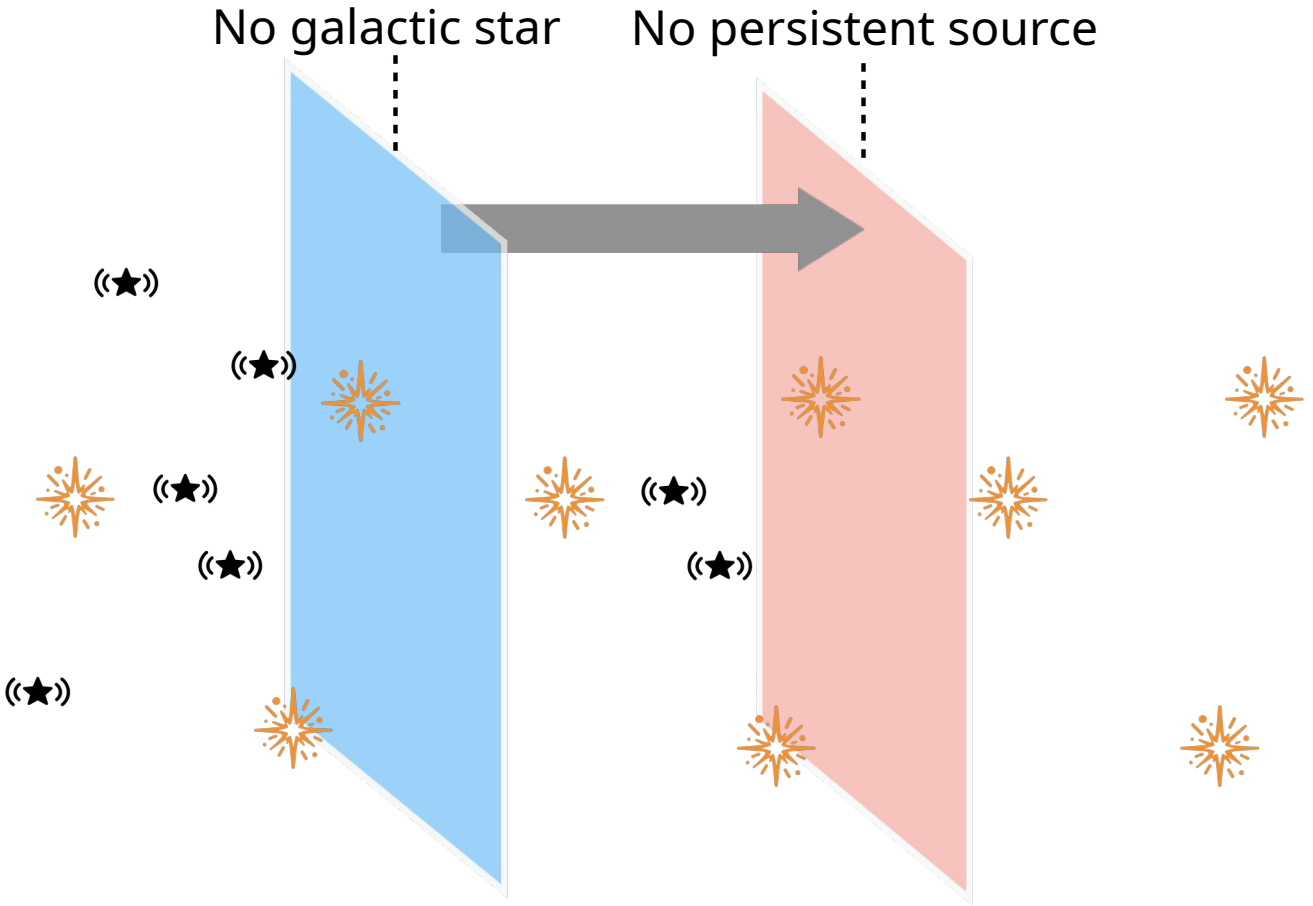
High recovery rate!
BUT, high background rate ...

Data analysis: optimized for counterpart

No persistent source

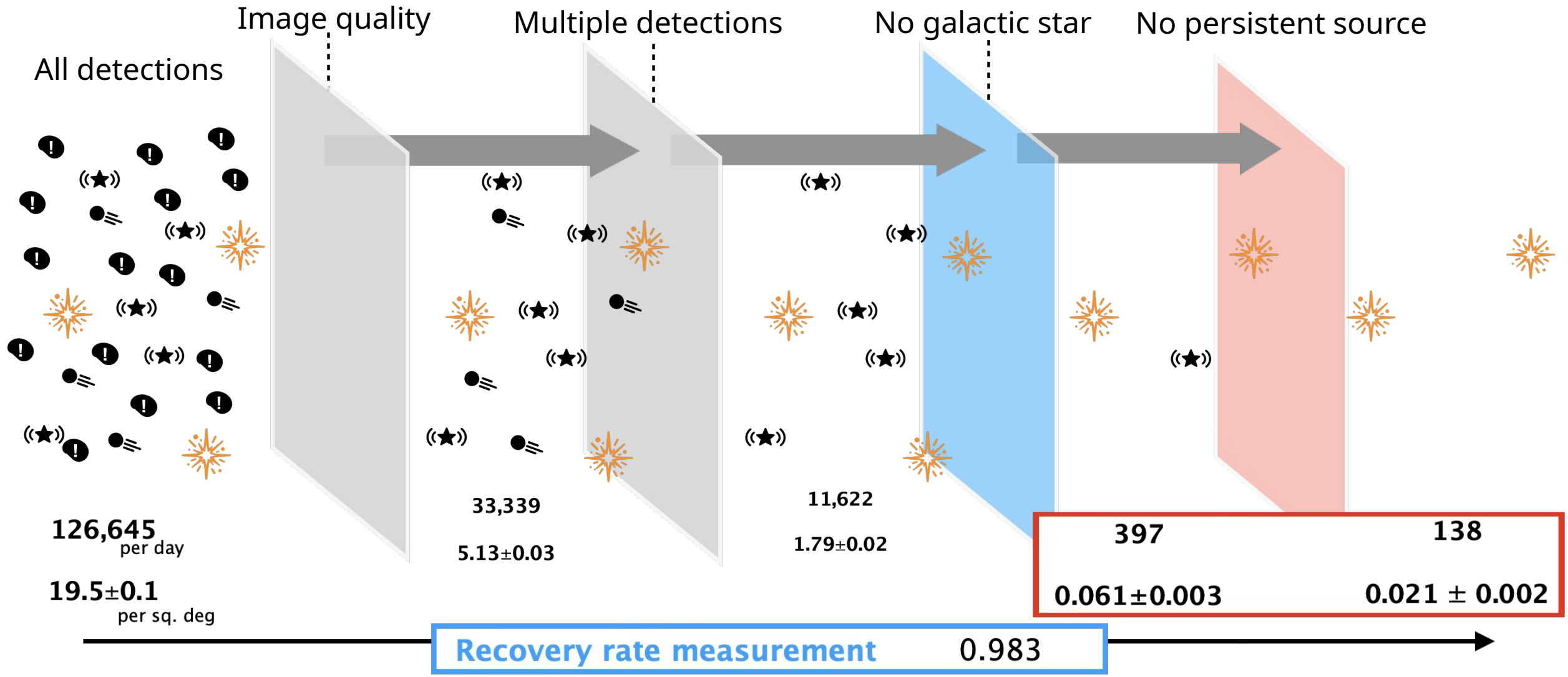


High recovery rate!
AND, low background rate !



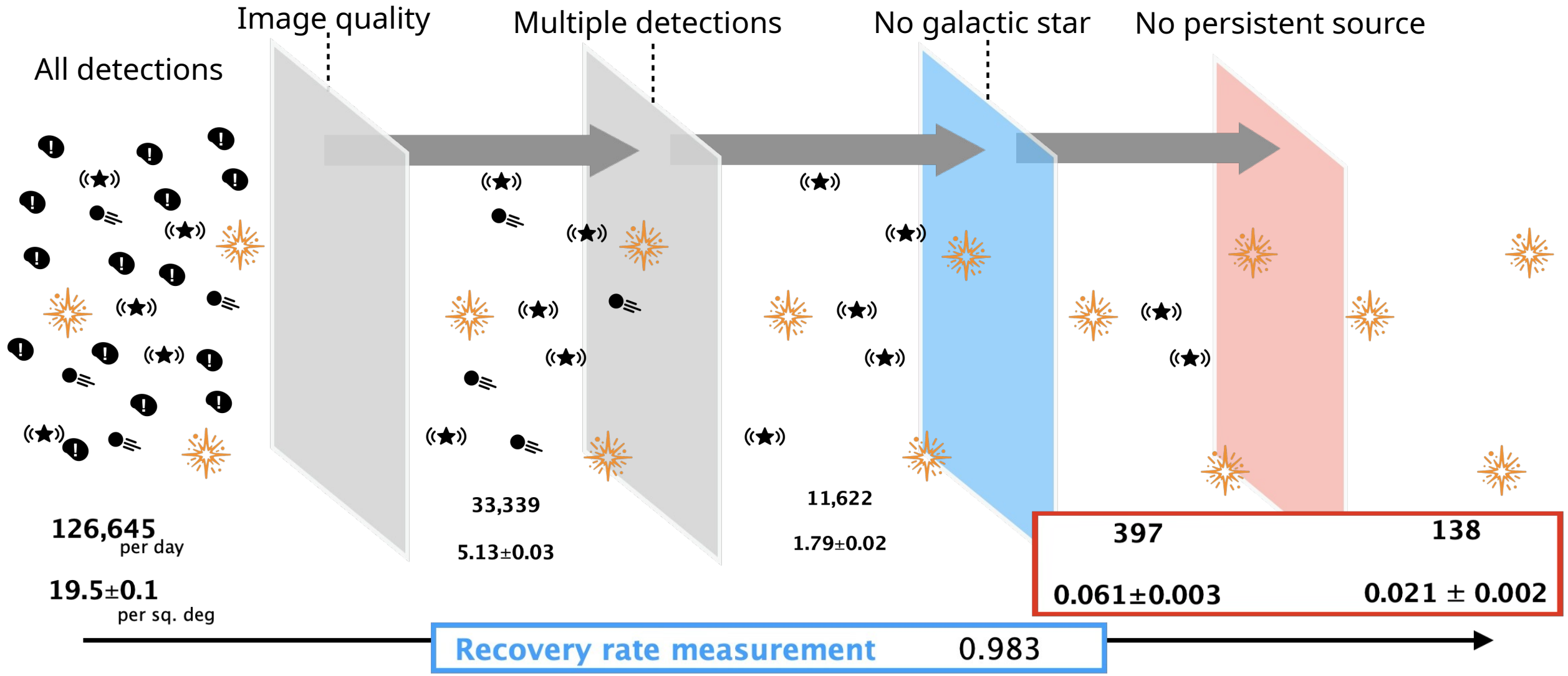
High recovery rate!
BUT, high background rate ...

Data analysis:



Predefined criteria optimized through data-driven measurements

Data analysis:



Predefined criteria optimized through data-driven measurements

Finally, unblind the data!

Applied pre-fixed criteria

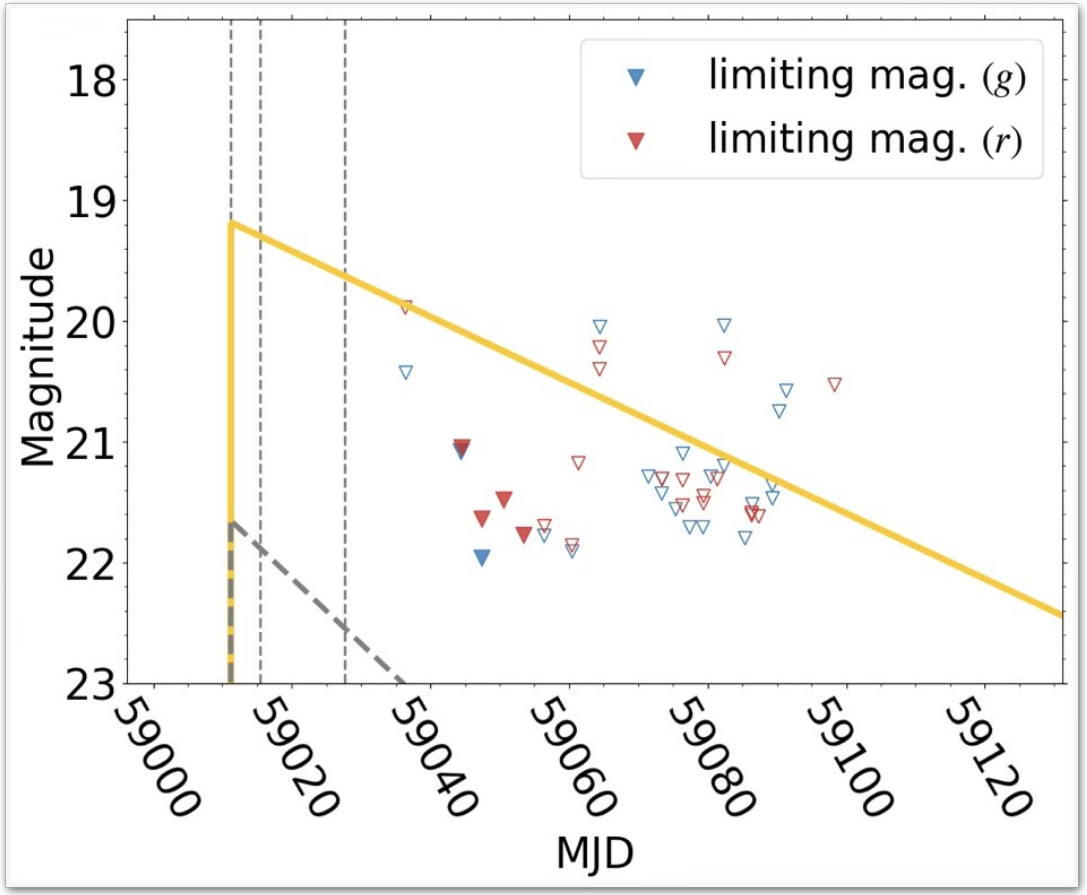
Finally, unblind the data!

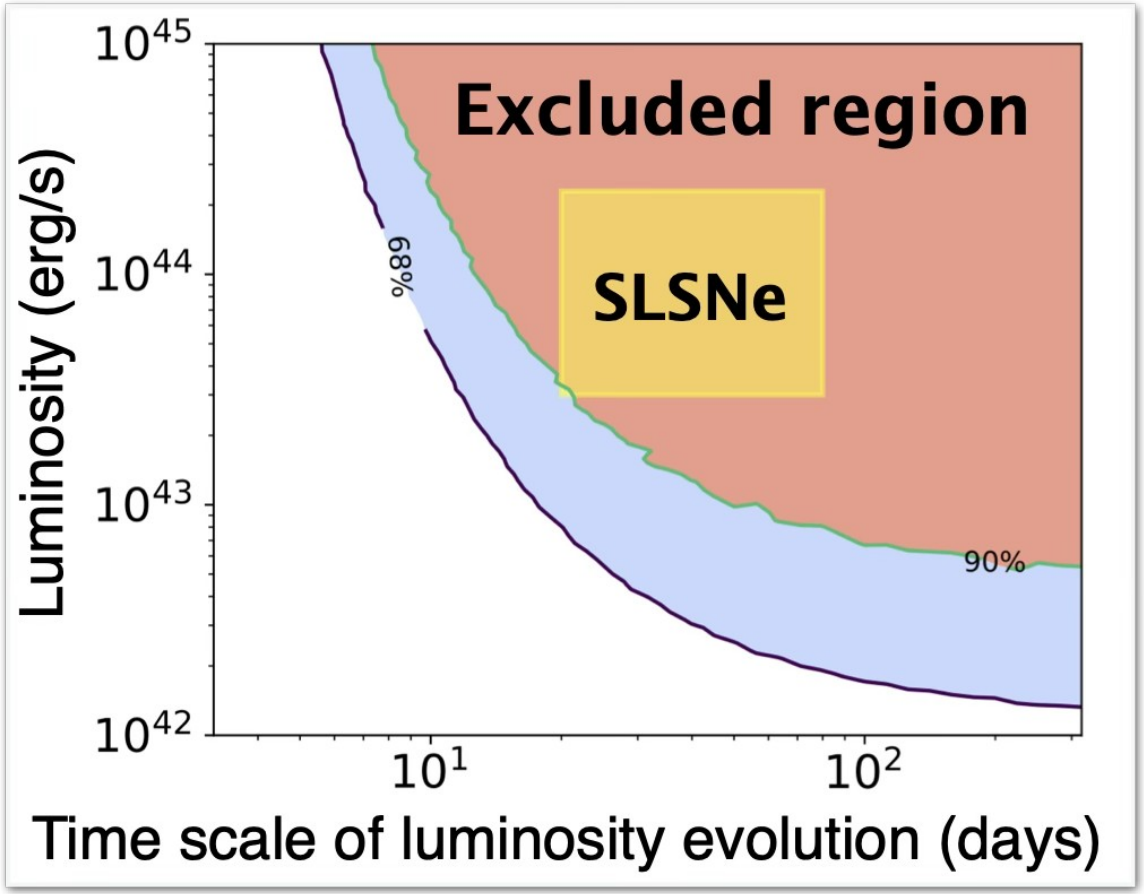
Applied pre-fixed criteria

No candidate was found in the multiplet direction

Results: Constraints from non-detection

Confirmation of low-background&high-recovery rate a priori





One multiplet search can put meaningful constraints on source properties

Summary

- Background characterization in optical counterpart search
  - Poorly modeled background (e.g., artifacts, moving objects)

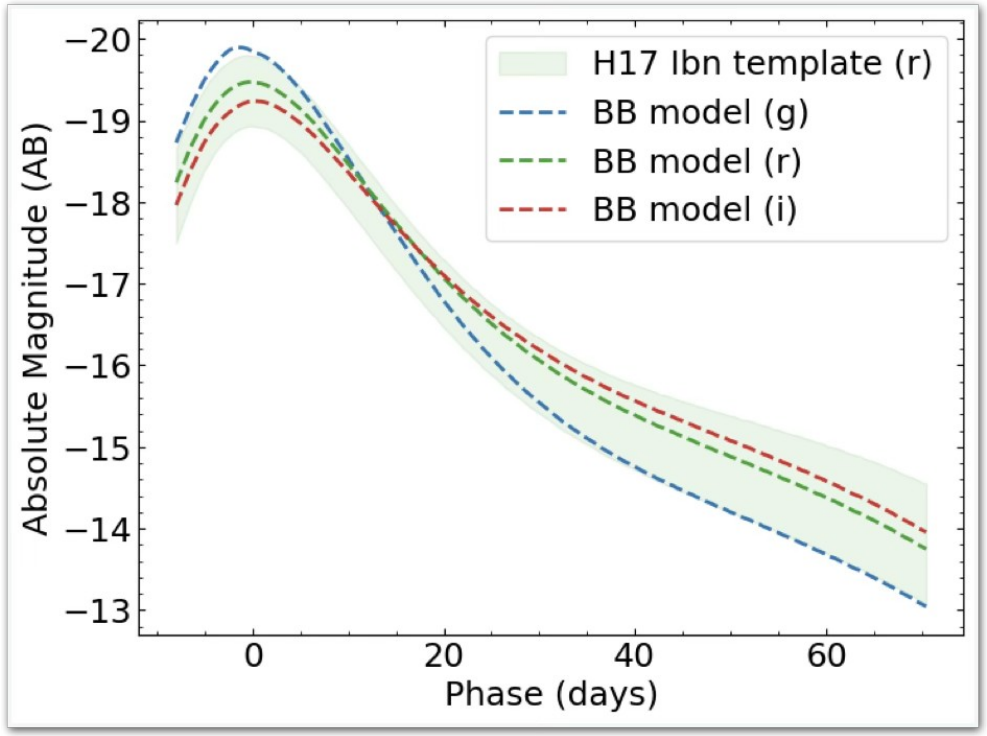
    → data-driven approach is needed
- We did a pilot work for IceCube triplet event in 2020
  - Prefixed criteria through data-driven approach
    - High recovery rate ~ 0.983 for typical source
    - Low background rate ~ 0.02 per sq. deg per epoch
  - Demonstrated that non-detection gives stringent constraints on source properties

Appendix

Simulation-driven estimate 1

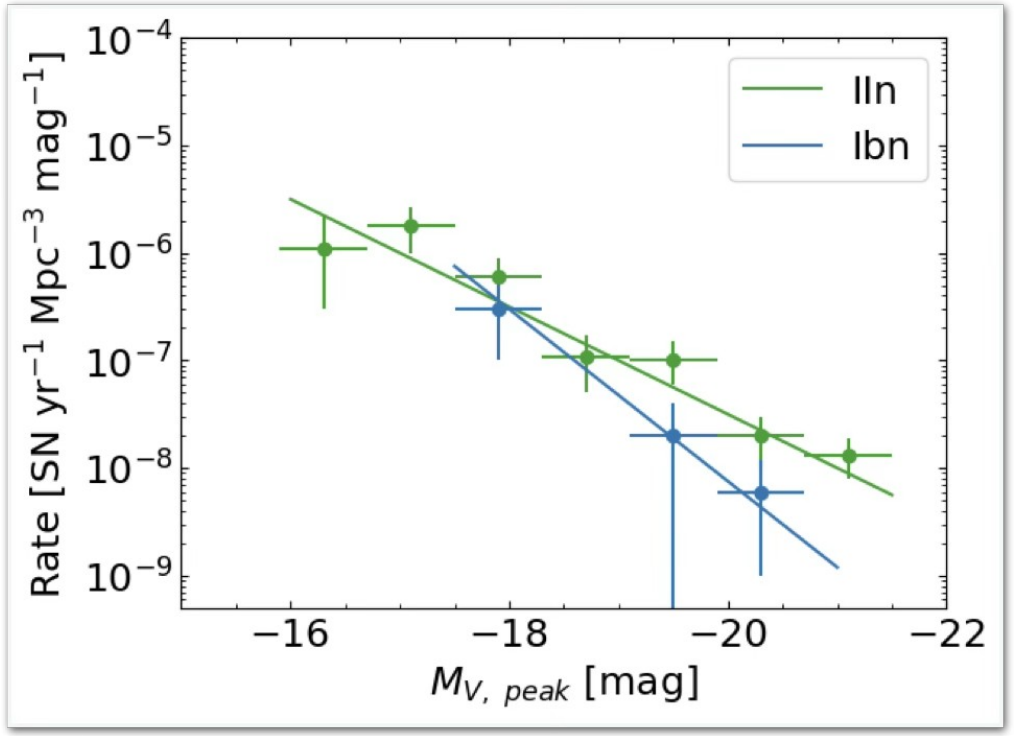
Mock observations using sncosmo A Python Library for Supernova Cosmology based on

Light curve template



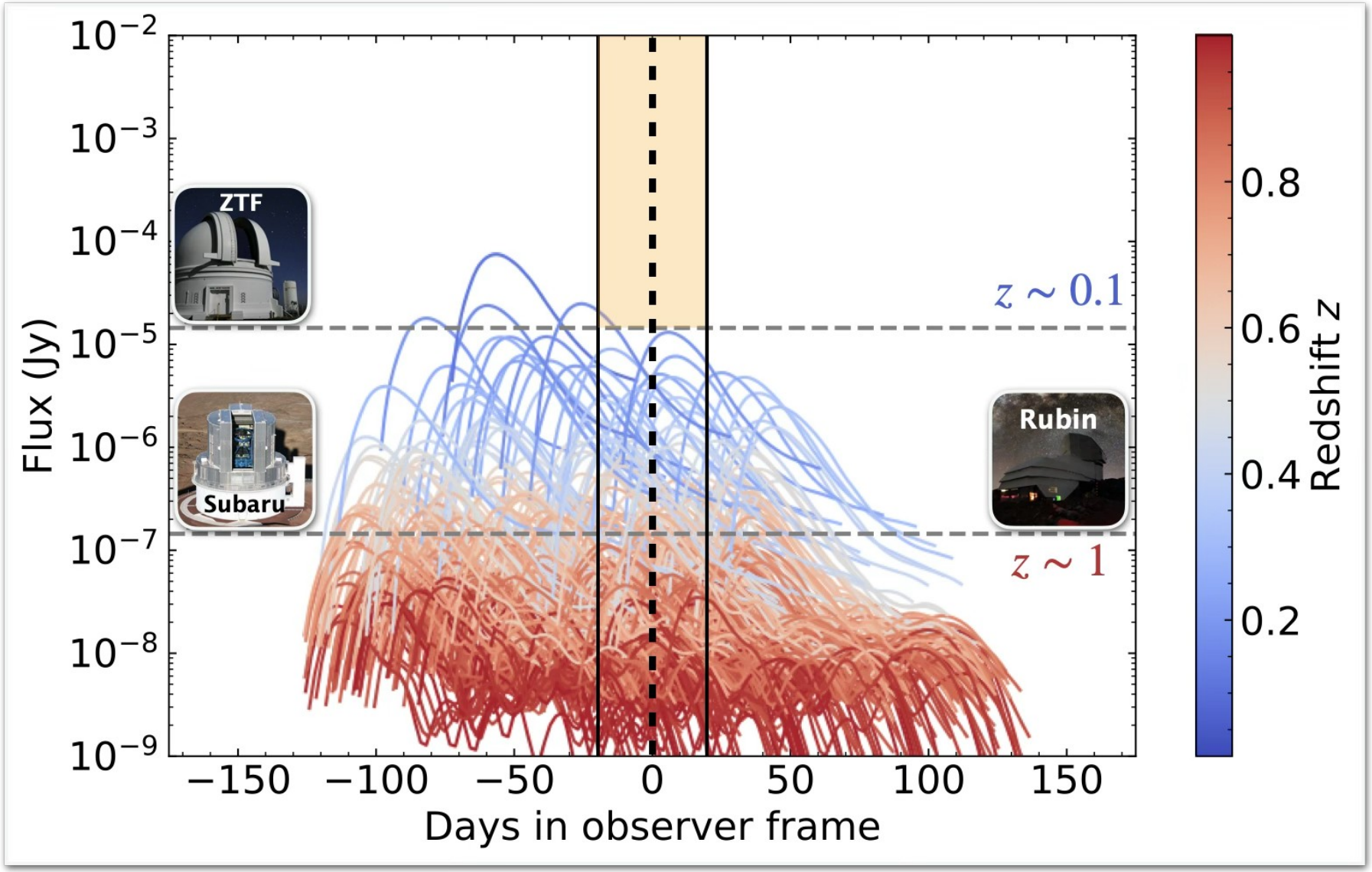
Template example for interacting SNe

Event rate&Redshift evolution

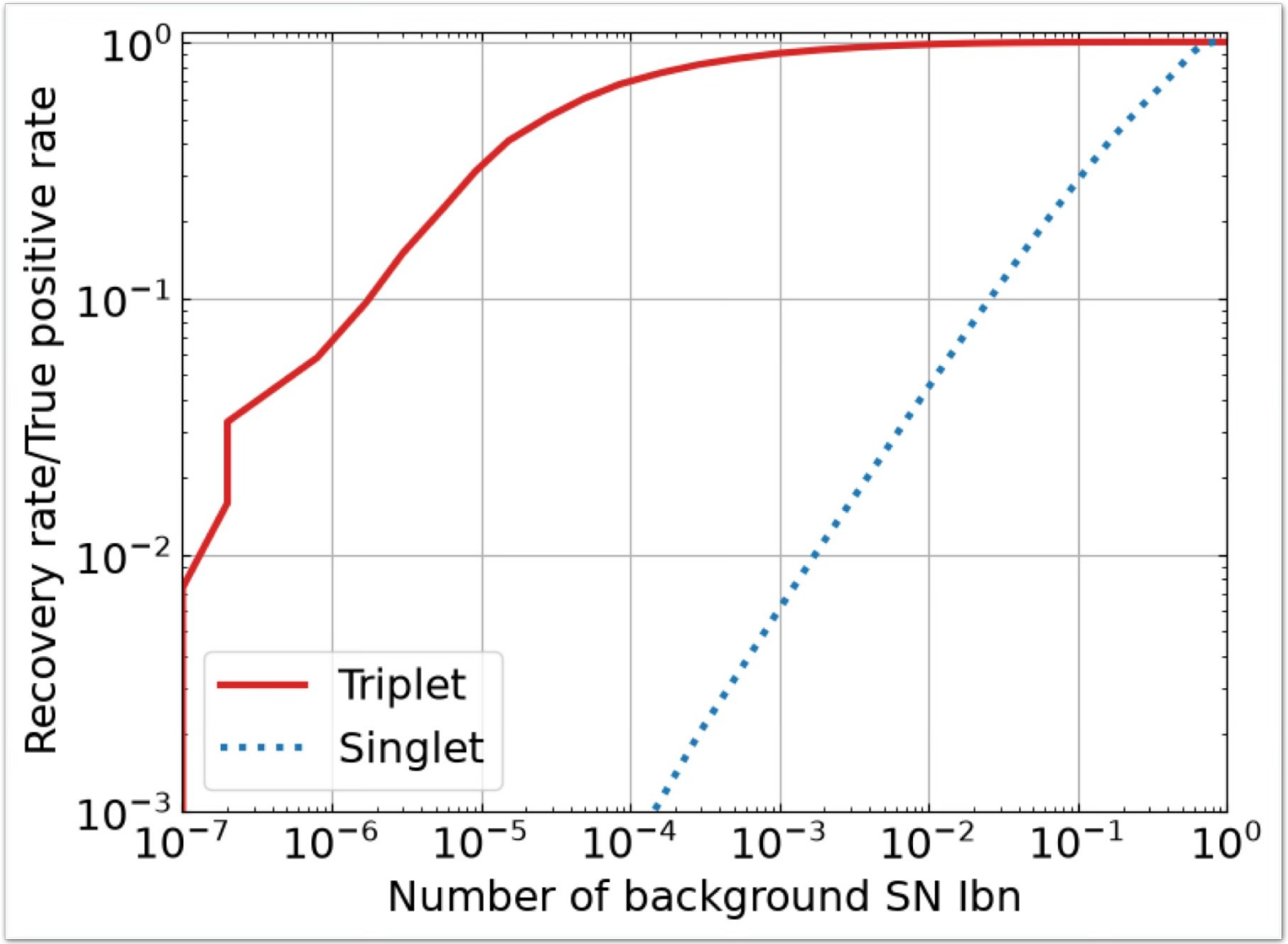


Pessi et al. 2025 + Madau&Fragos 2017

Simulation-driven estimate 2



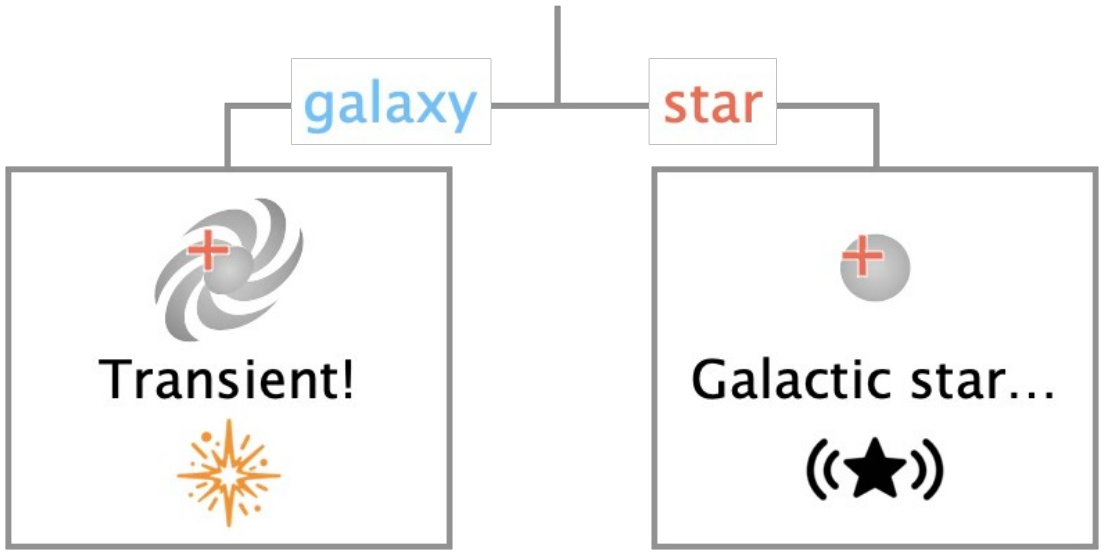
Recovery rate vs Background rate

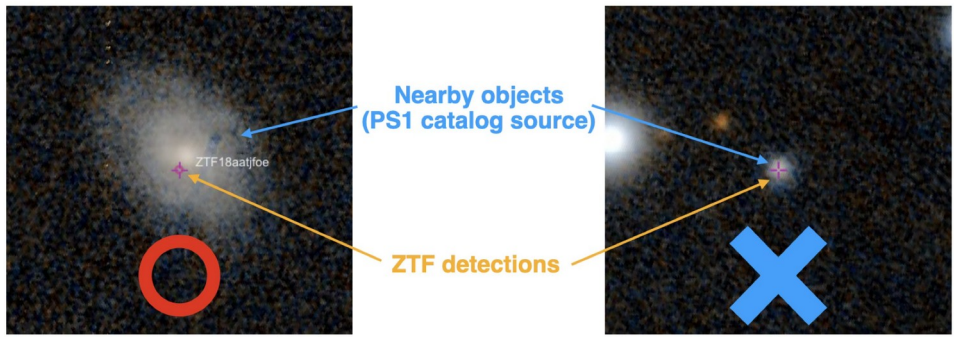


No galactic star criterion

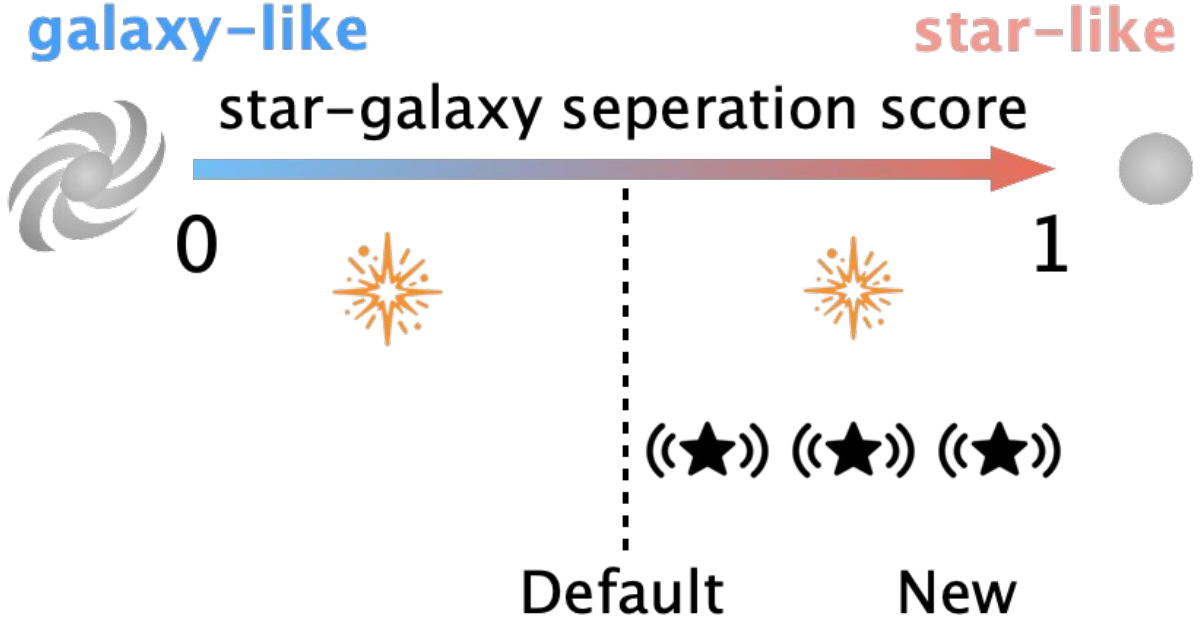
No galactic star criterion

"What is the nearest object ?"

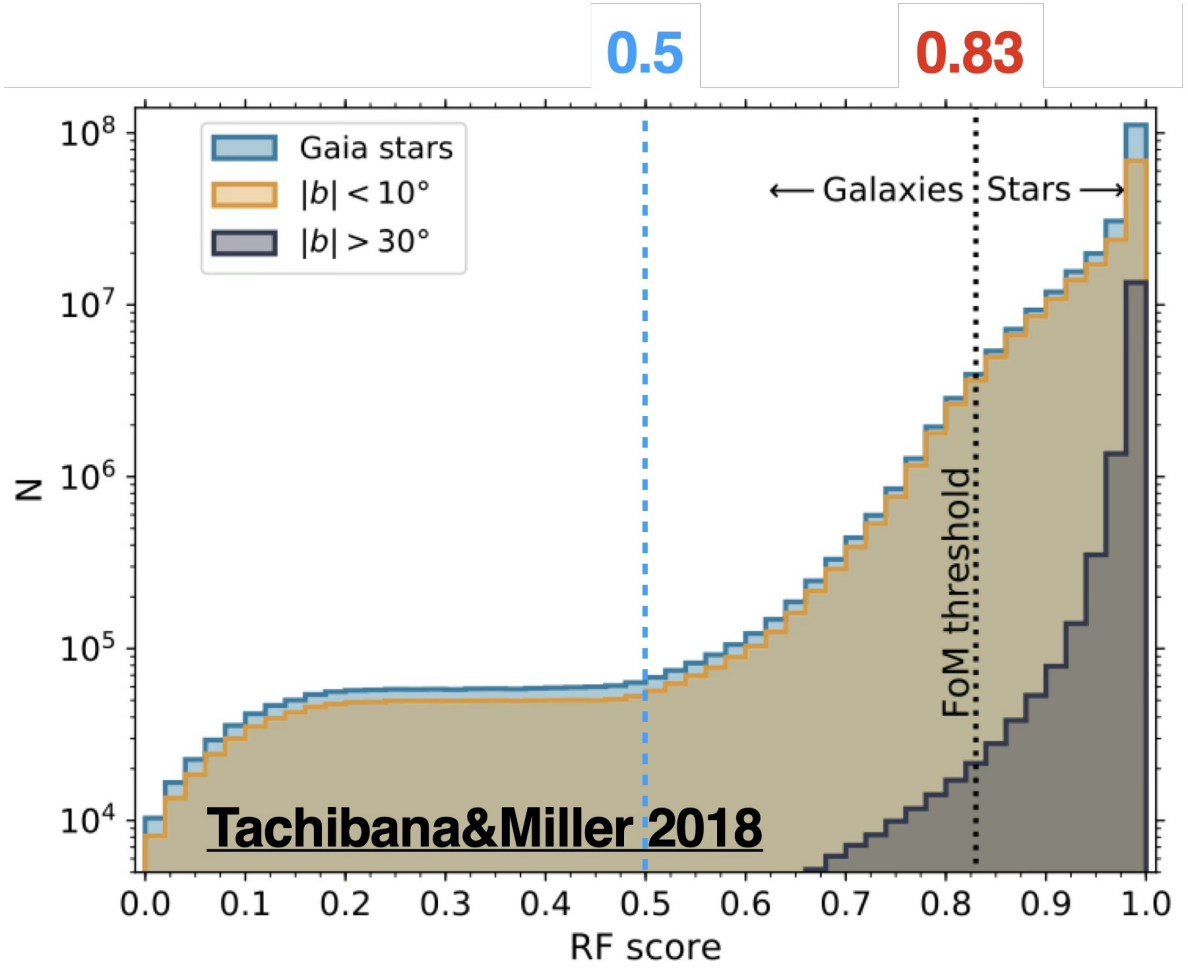




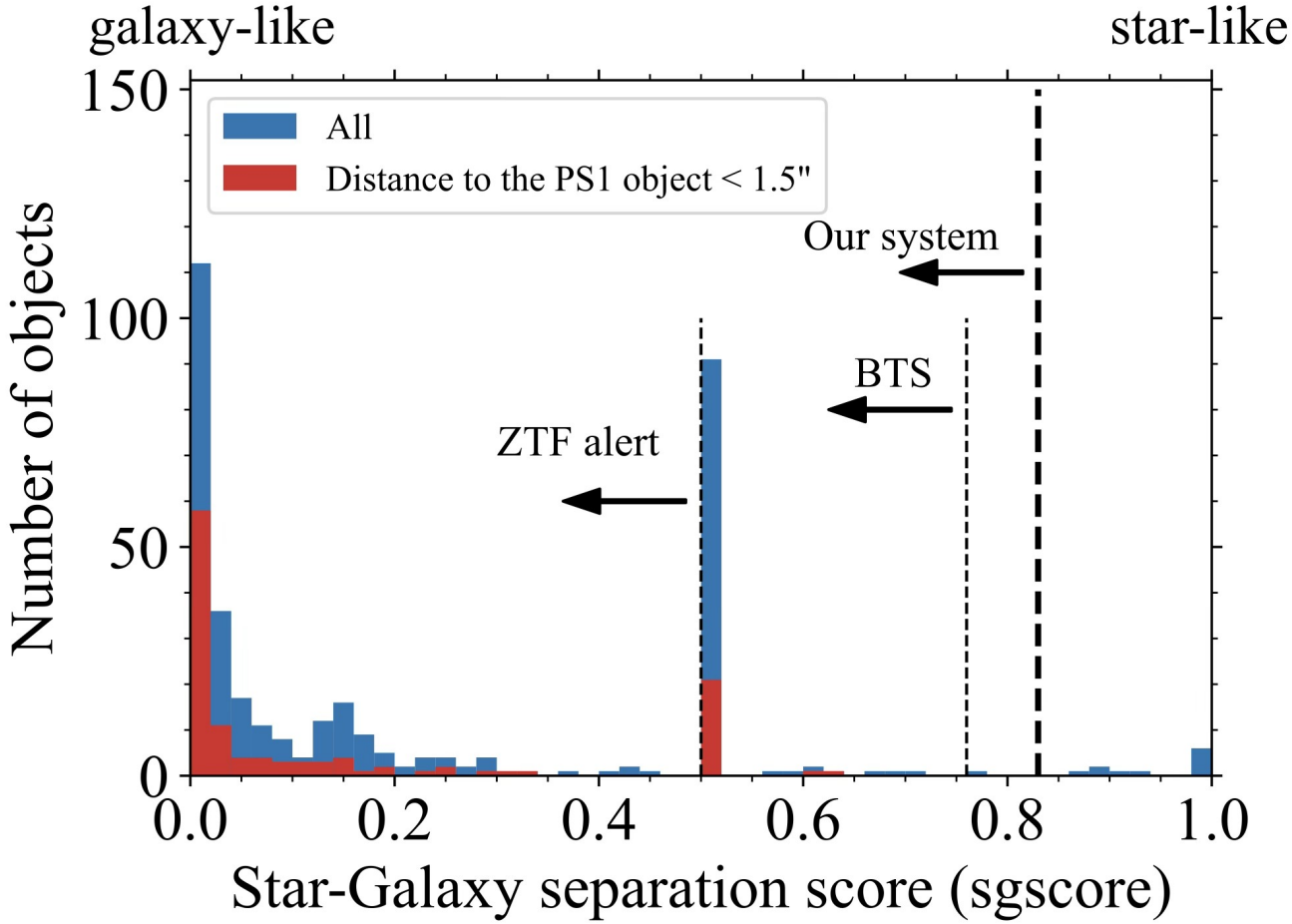
How to distinguish galaxy/star ?



Star-galaxy separation score



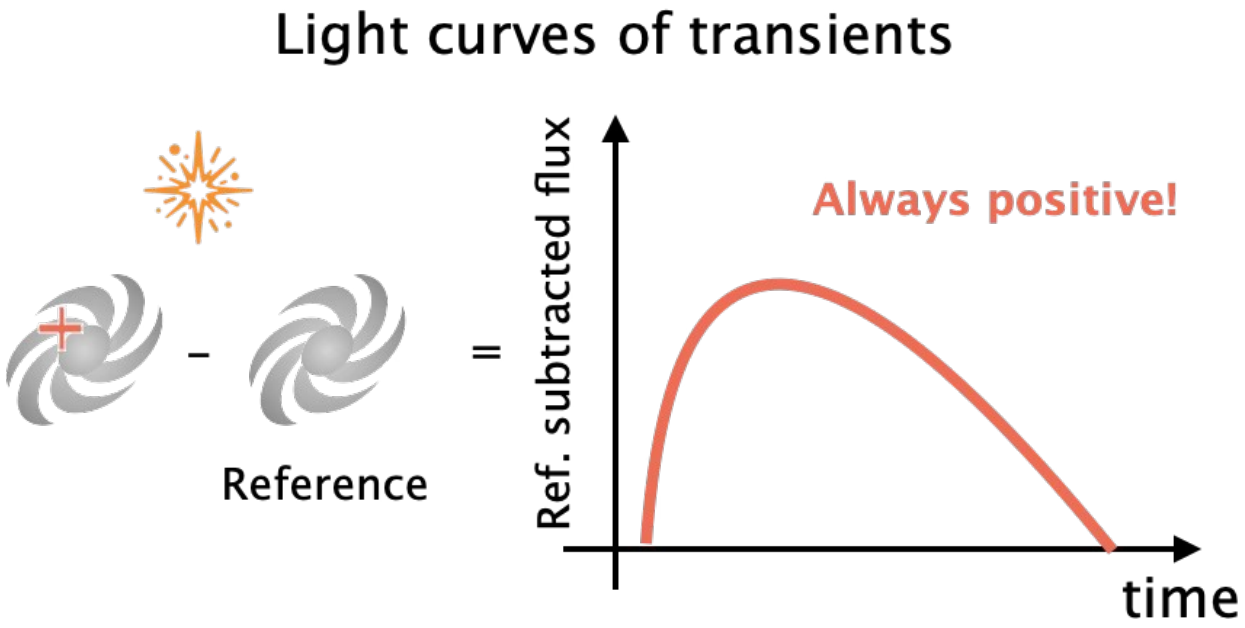
Distribution of sgscore for stars

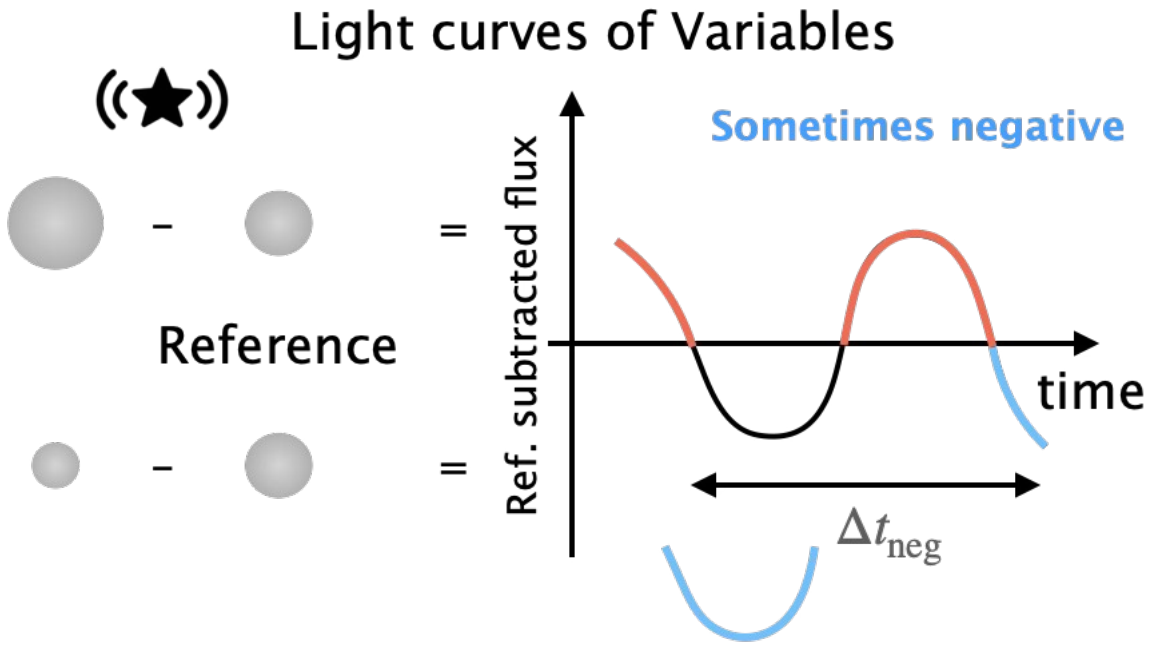


Distribution of sgscore for SNe

Light curve cut

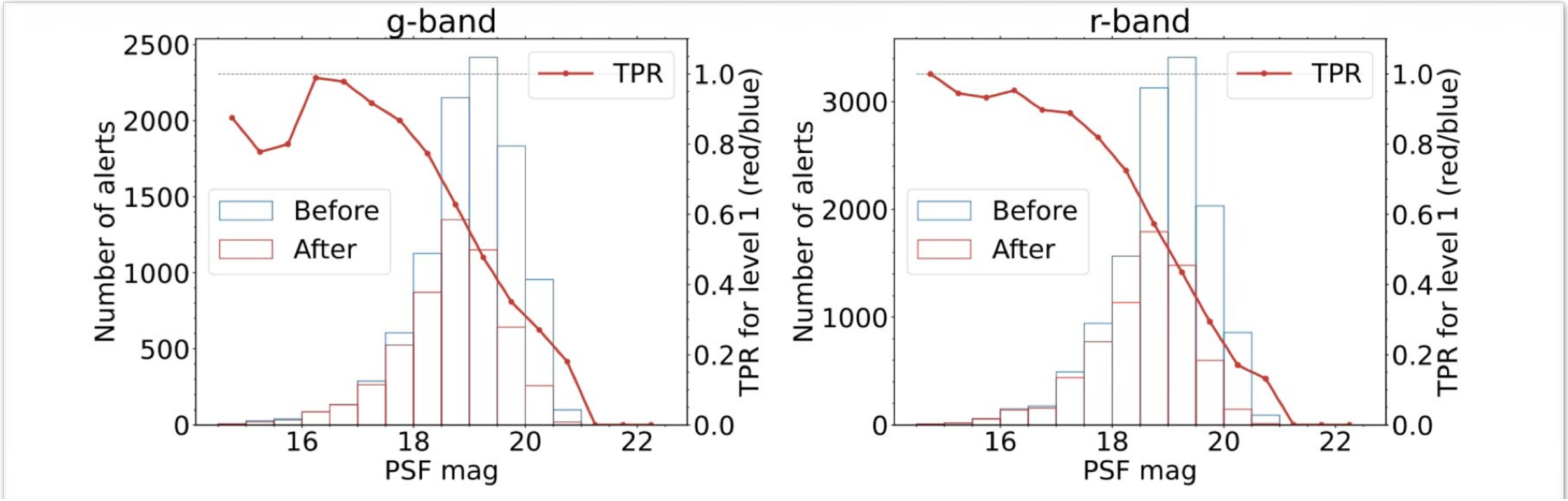
No persistent source criterion



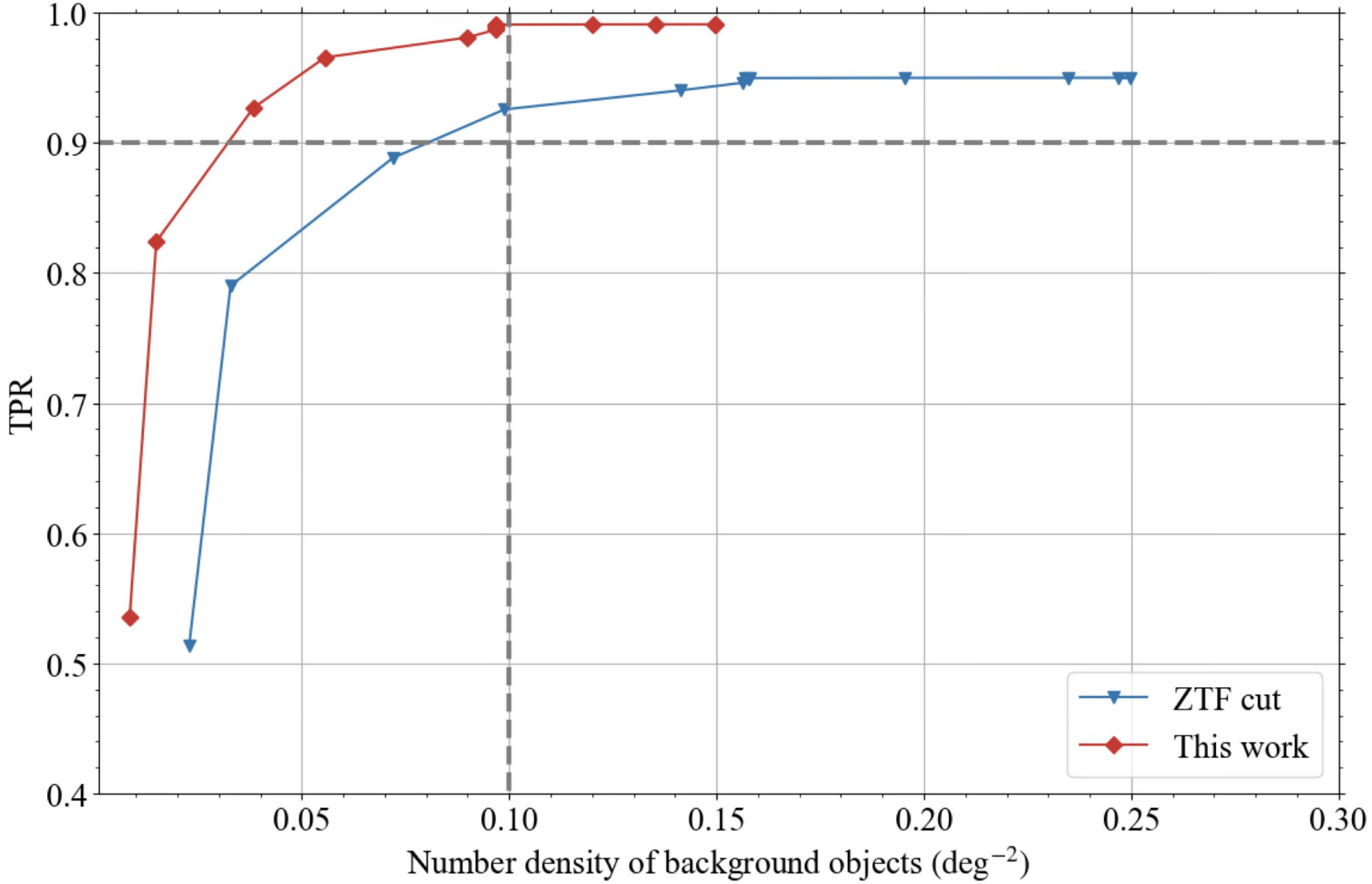


cut objects with negative durations $\Delta t_{neg} > 2$ yr

Flux dependence of recovery rate



Recovery rate vs Background rate



Recovery rate vs background rate