Microlensing Data Challenge

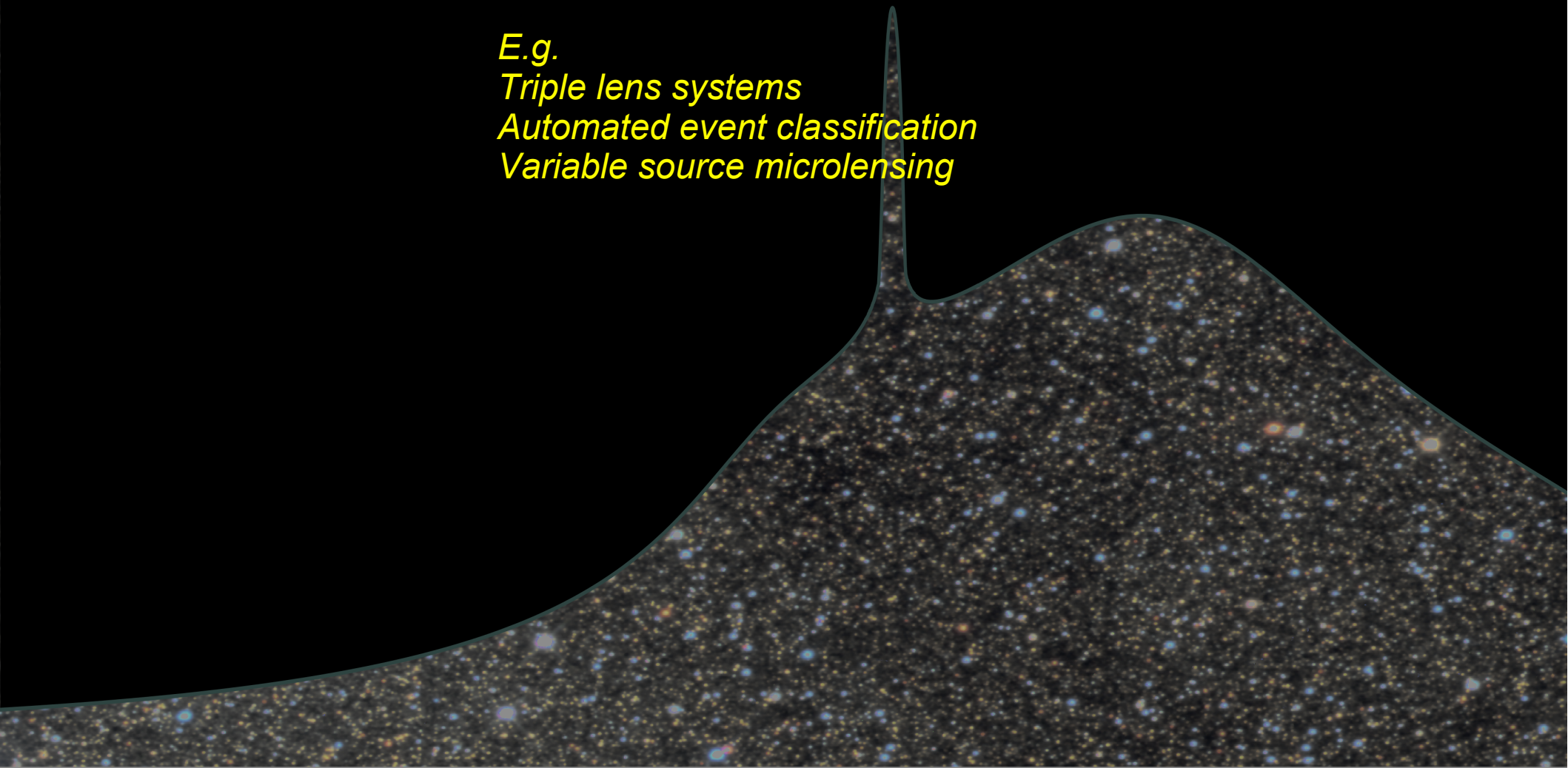
R. Street,
Las Cumbres Observatory

M. Penny, J. Yee, R. Poleski
WFIRST MicroSIT Group 2

Motivation

Though microlensing has been observed for a while, there are still a number of unsolved challenges with microlensing analysis

E.g.
Triple lens systems
Automated event classification
Variable source microlensing



Motivation

Though microlensing has been observed for a while, there are still a number of unsolved challenges with microlensing analysis

Analysis of multi-lens microlensing events has been historically time consuming, though major progress has been made in recent years

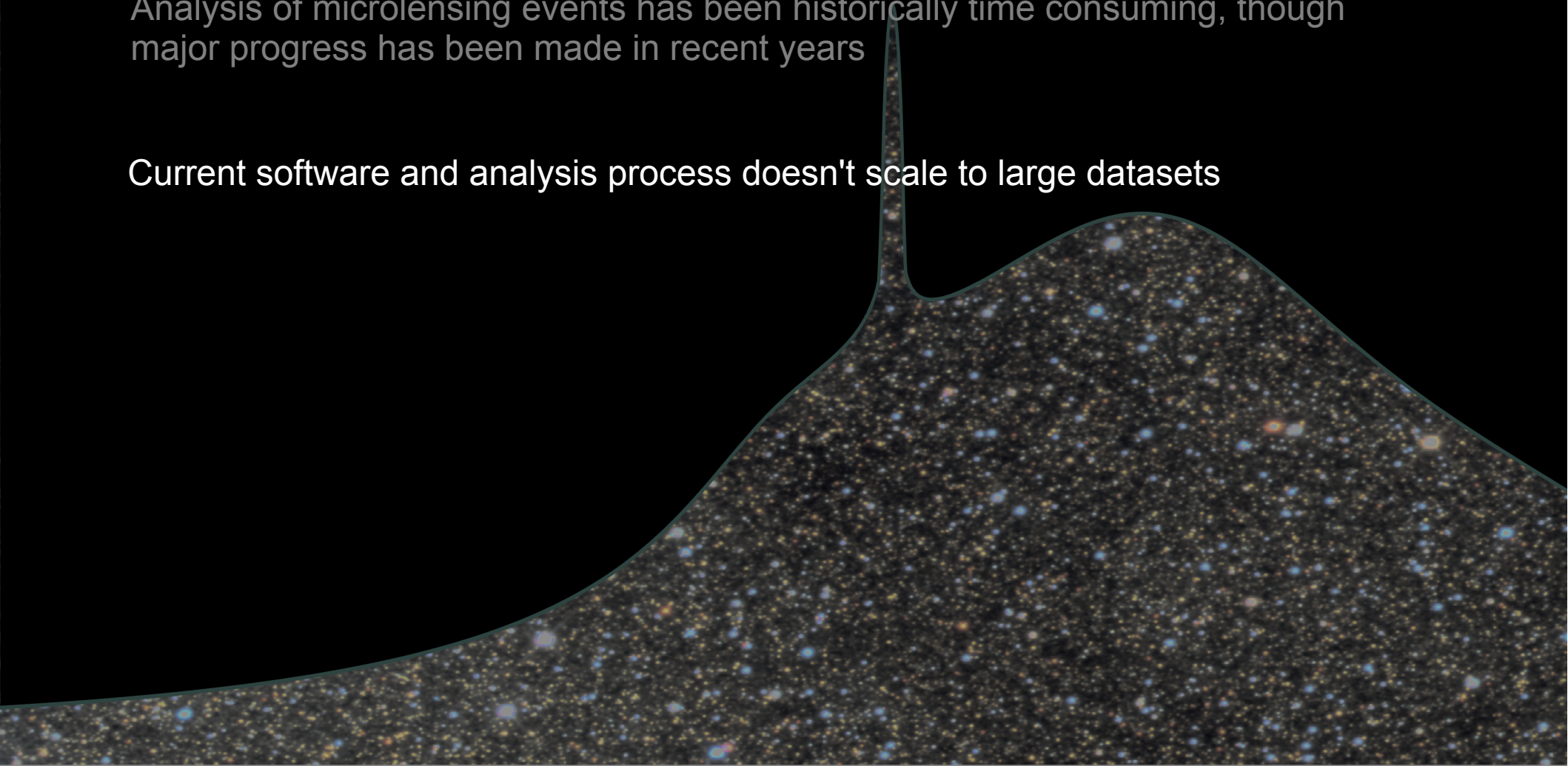
E.g.
Thorough but efficient searching of parameter space
Distinguishing binary and triple lenses
Disambiguating source variability from microlensing
...

Motivation

Though microlensing has been observed for a while, there are still a number of unsolved challenges with microlensing analysis

Analysis of microlensing events has been historically time consuming, though major progress has been made in recent years

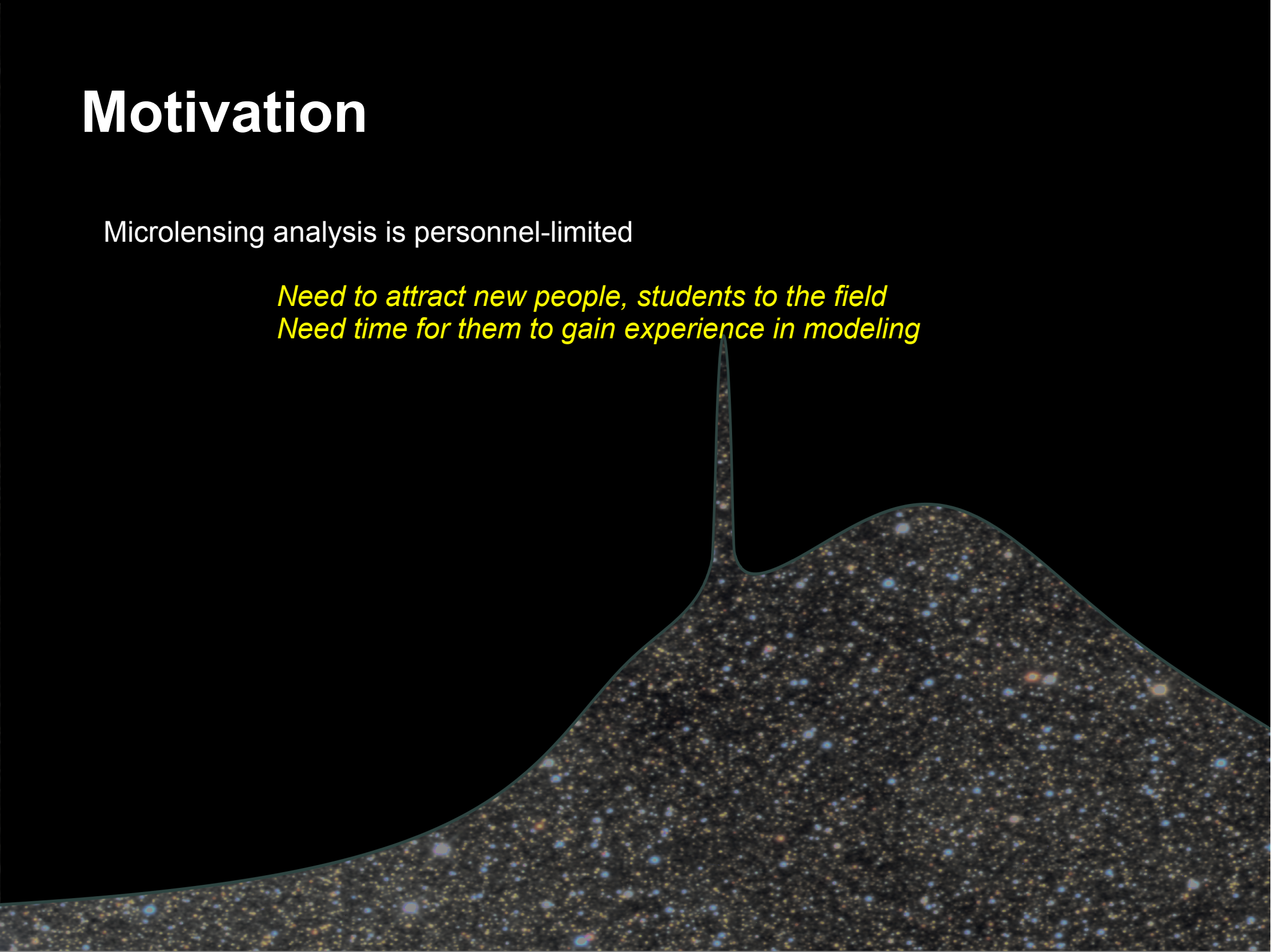
Current software and analysis process doesn't scale to large datasets



Motivation

Microlensing analysis is personnel-limited

Need to attract new people, students to the field
Need time for them to gain experience in modeling

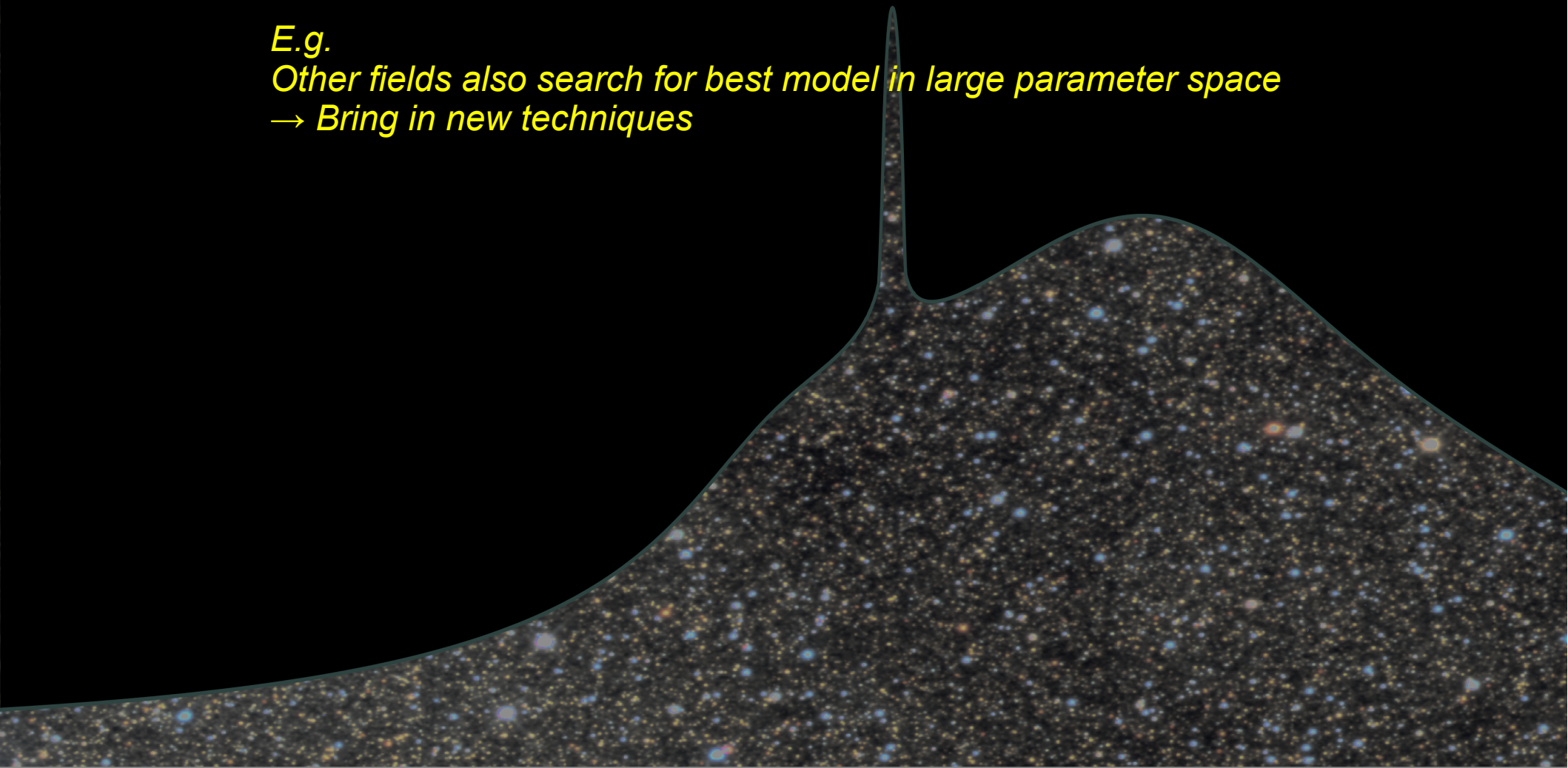


Motivation

Microlensing is personnel-limited

Would benefit from involvement of experts in mathematics, statistics, informatics

E.g.
Other fields also search for best model in large parameter space
→ Bring in new techniques



Motivation

Microlensing is personnel-limited

Would benefit from involvement of experts in mathematics, statistics, informatics

Data Challenges have been successful in stimulating engagement and innovation in other fields including exoplanets

E.g.
Radial velocity fitting challenge
https://github.com/EPRV3EvidenceChallenge/Inputs
Dumusque, X. et al. (2016), A&A, 593, 5
Dumusque, X. et al. (2017), A&A, 598, 133

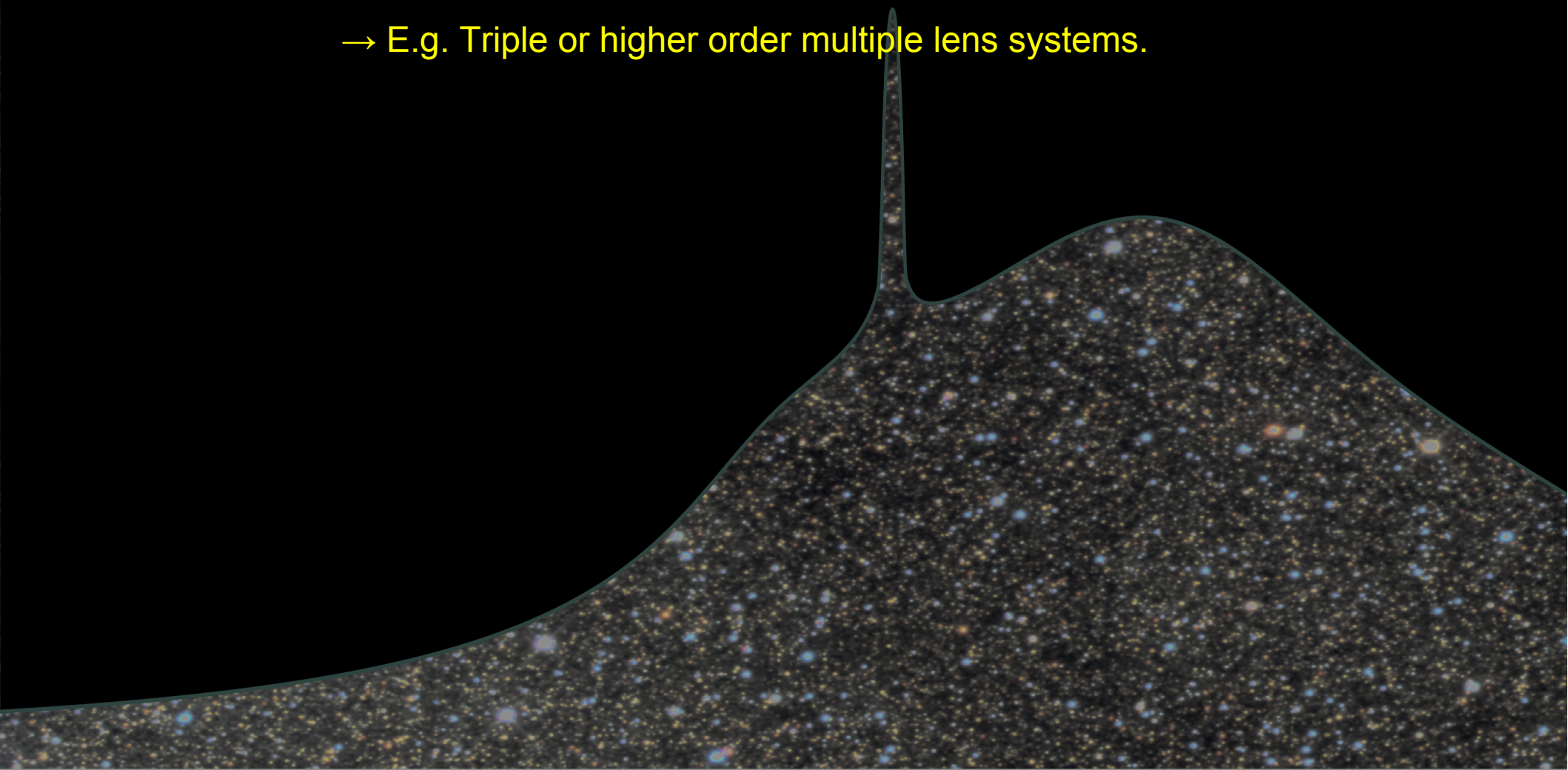
Transit detection
CoRoT analyses challenge

Exoplanet atmosphere spectral analysis
Hildebrant, S et al.
http://adsabs.harvard.edu/abs/2018AAS...23115803H

Data Challenge Goals

- To stimulate research effort into outstanding modeling issues

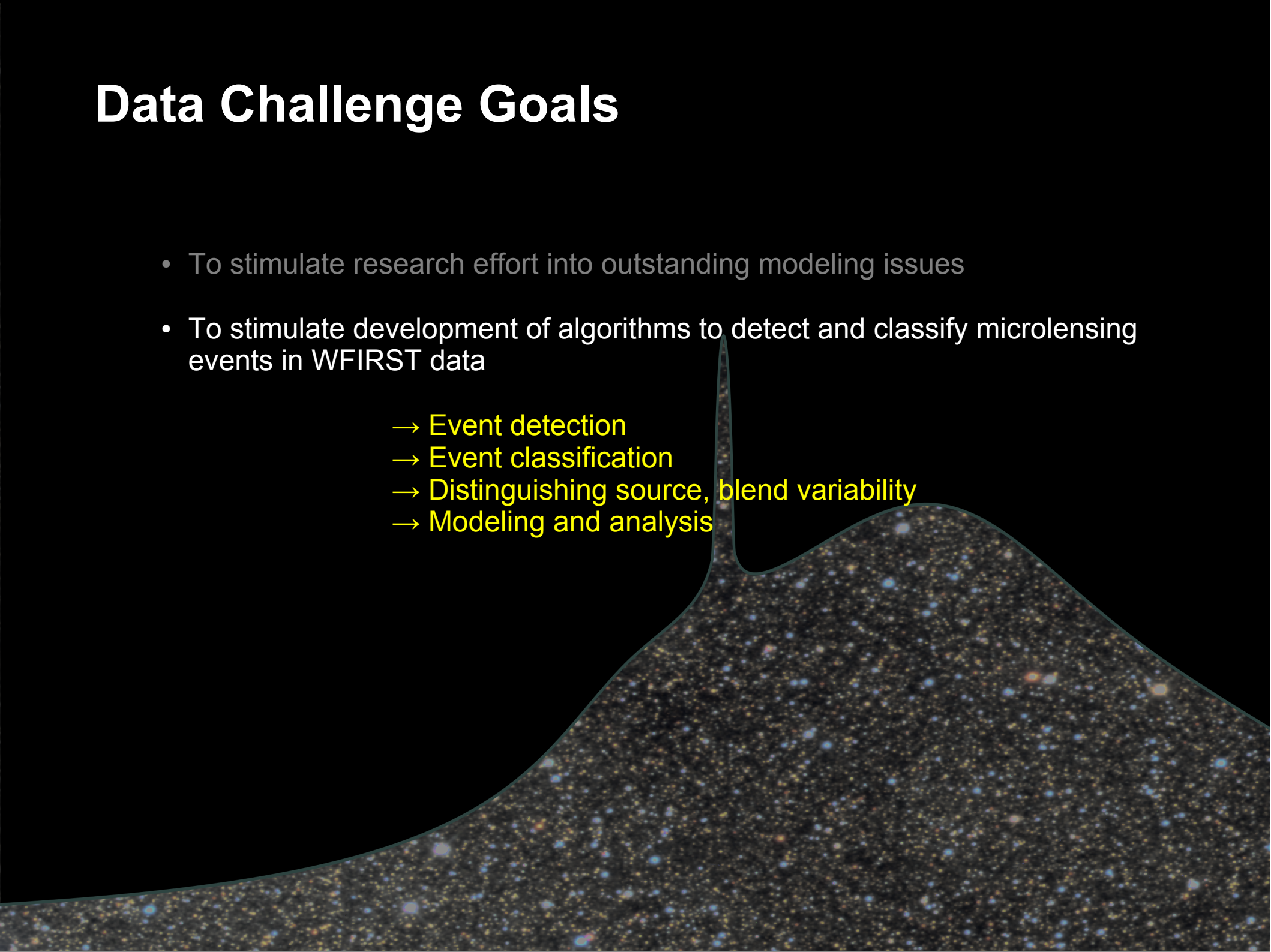
→ E.g. Triple or higher order multiple lens systems.



Data Challenge Goals

- To stimulate research effort into outstanding modeling issues
- To stimulate development of algorithms to detect and classify microlensing events in WFIRST data

→ Event detection
→ Event classification
→ Distinguishing source, blend variability
→ Modeling and analysis



Data Challenge Goals

- To stimulate research effort into outstanding modeling issues
- To stimulate development of algorithms to detect and classify microlensing events in WFIRST data
- To stimulate development of software for modeling microlensing events, capable of conducting analyses of WFIRST-scale datasets

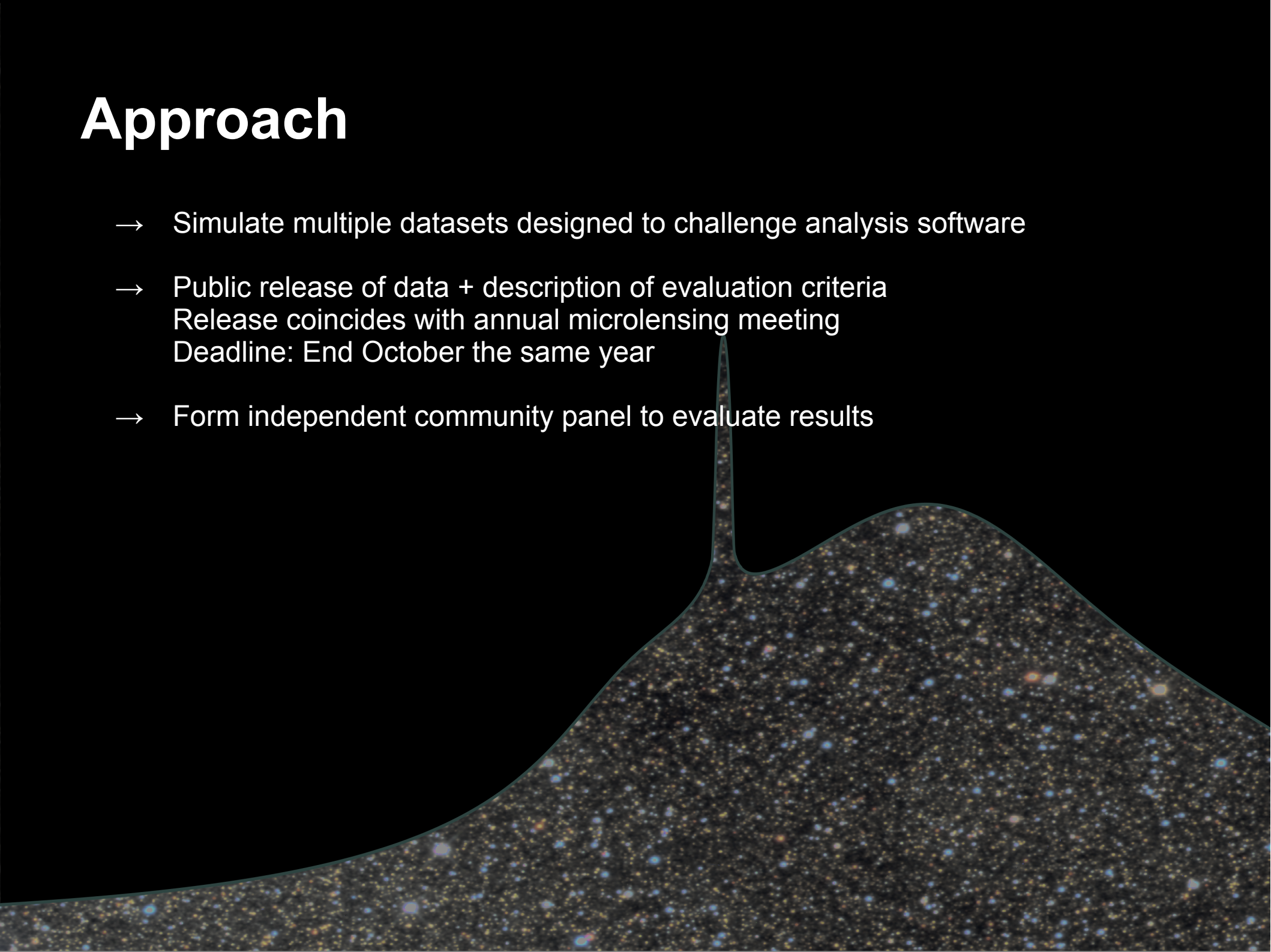
  → Hundreds of thousands of lightcurves
  → Thousands of events within a reasonable timeframe
  → Understand issues presented by WFIRST sampling, data gaps etc

Approach

→ Simulate multiple datasets designed to challenge analysis software

→ Public release of data + description of evaluation criteria
Release coincides with annual microlensing meeting
Deadline: End October the same year

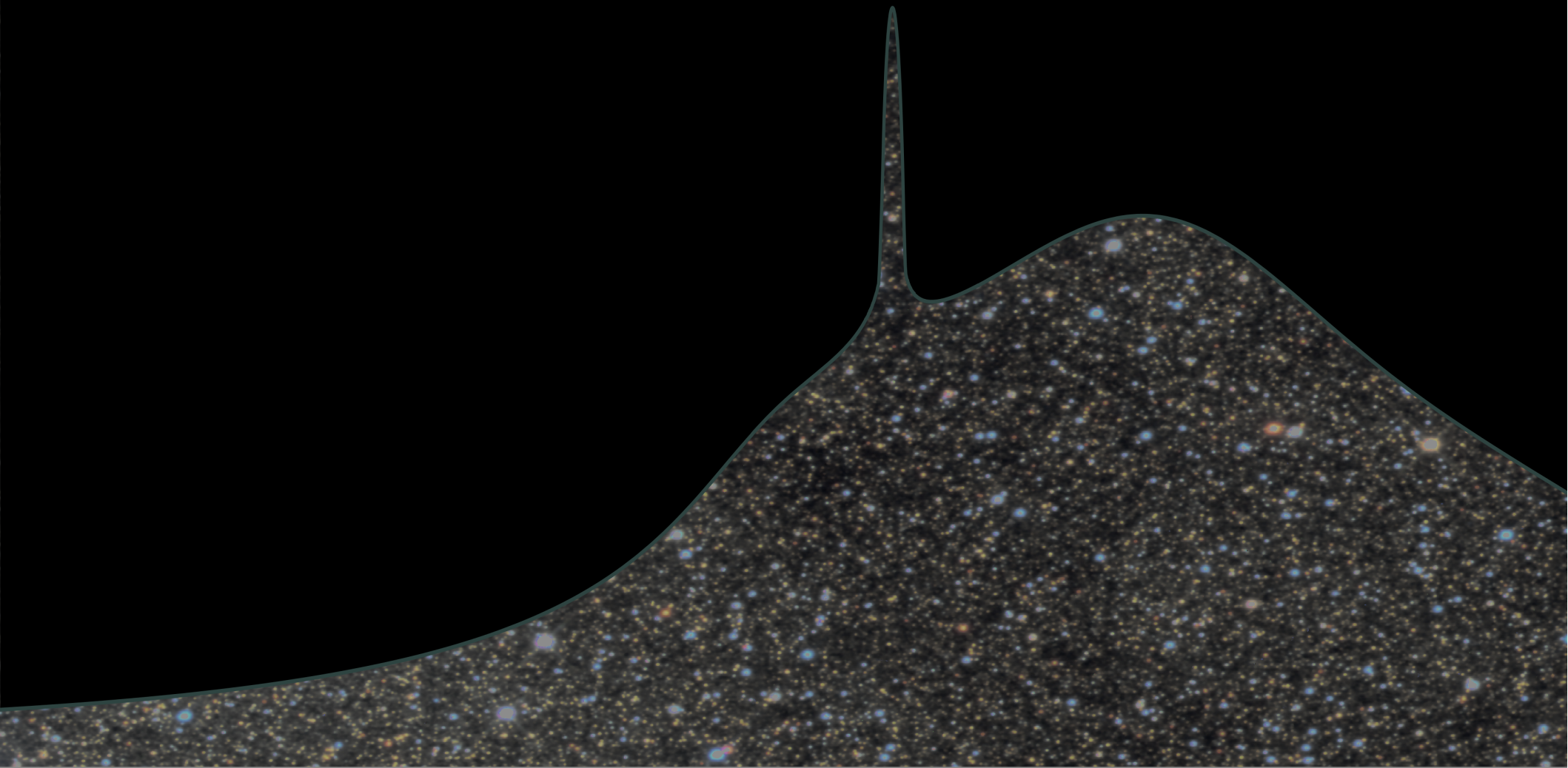
→ Form independent community panel to evaluate results



Simulated Data

- Matthew Penny has simulated the first set of lightcurves

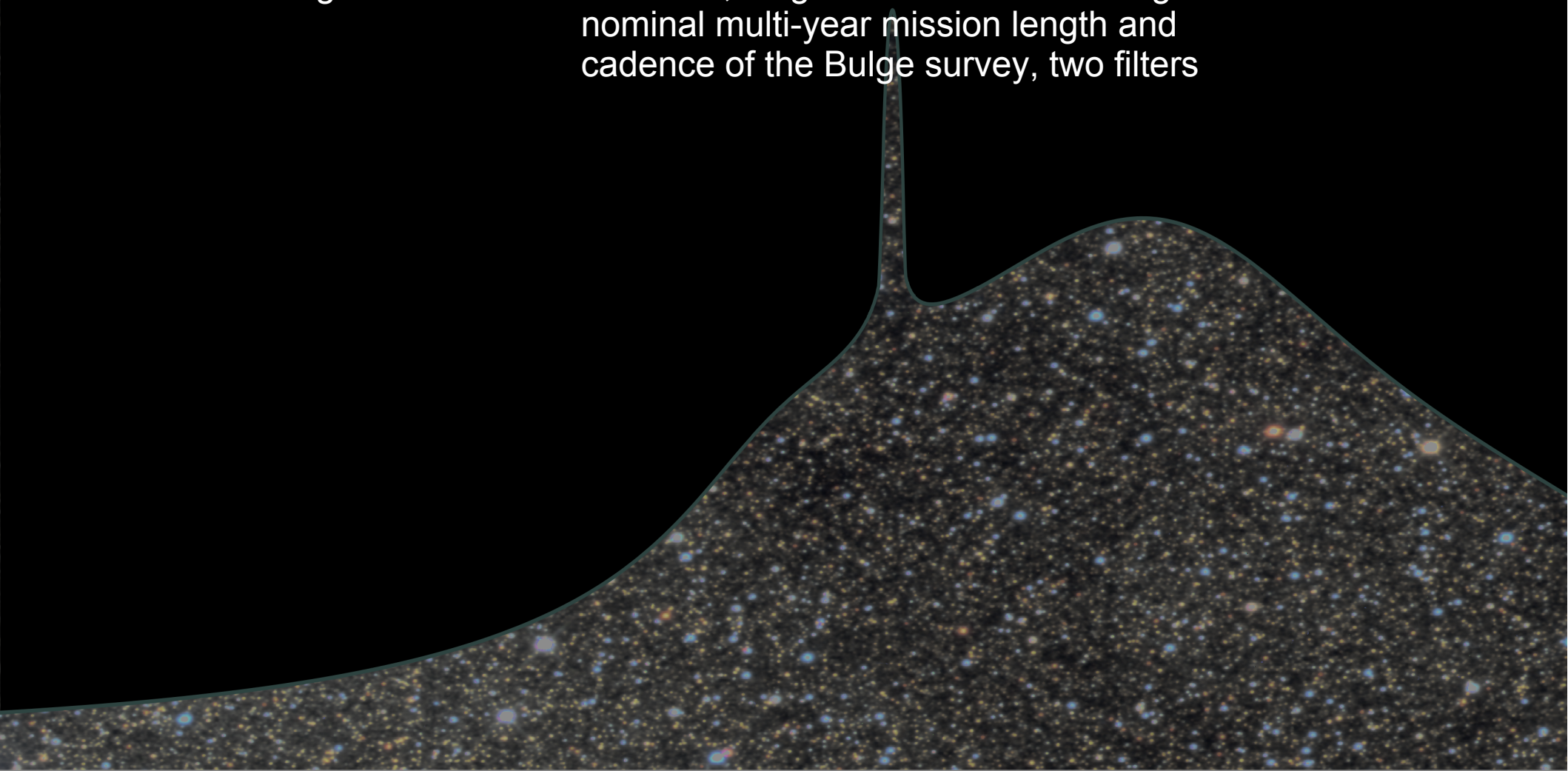
  He will keep the table of input parameters private until the submission deadline



Simulated Data

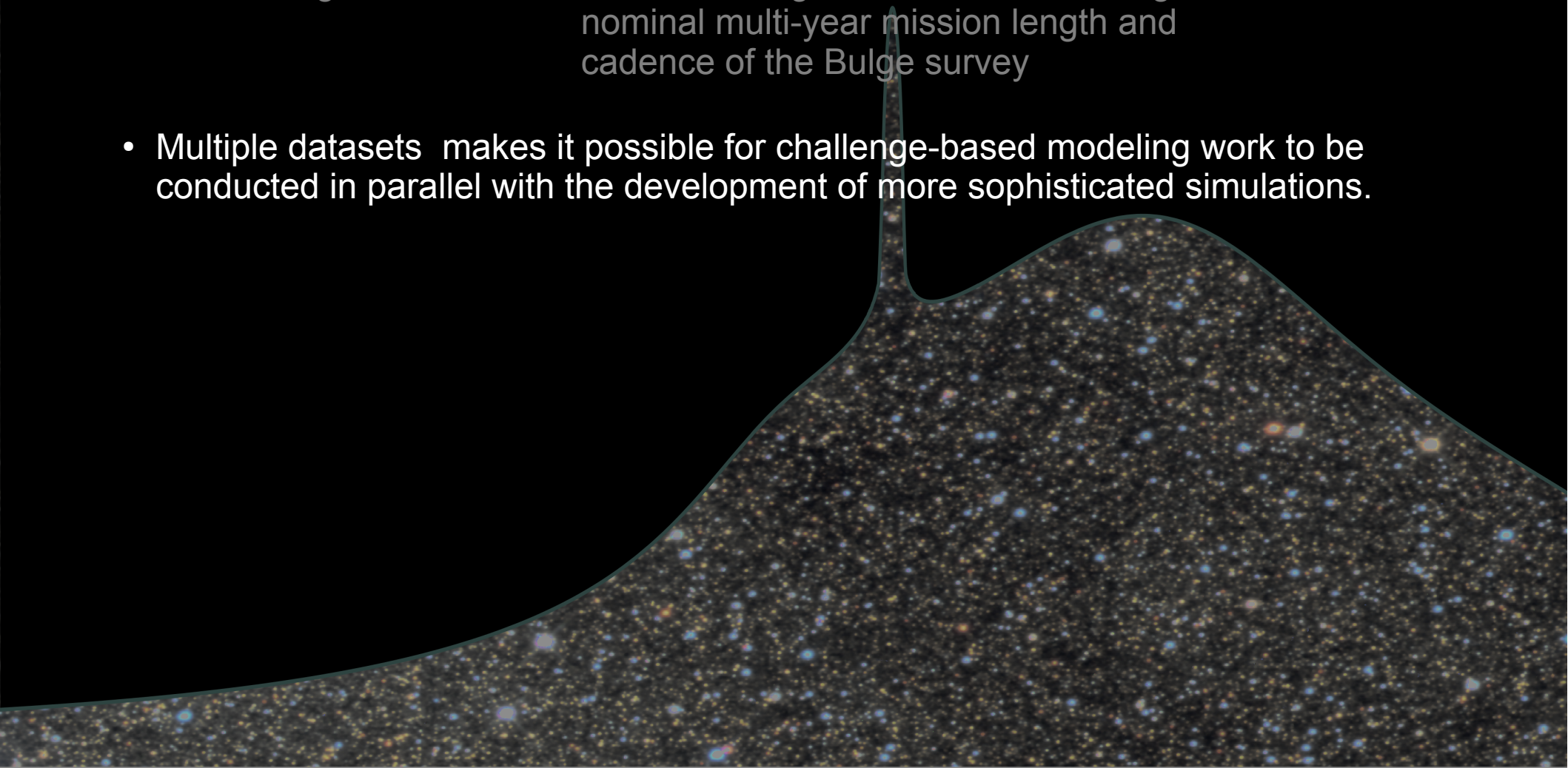
- Matthew Penny has simulated the first set of lightcurves

- WFIRST lightcurves = Cadence, length and noise mimicking the nominal multi-year mission length and cadence of the Bulge survey, two filters



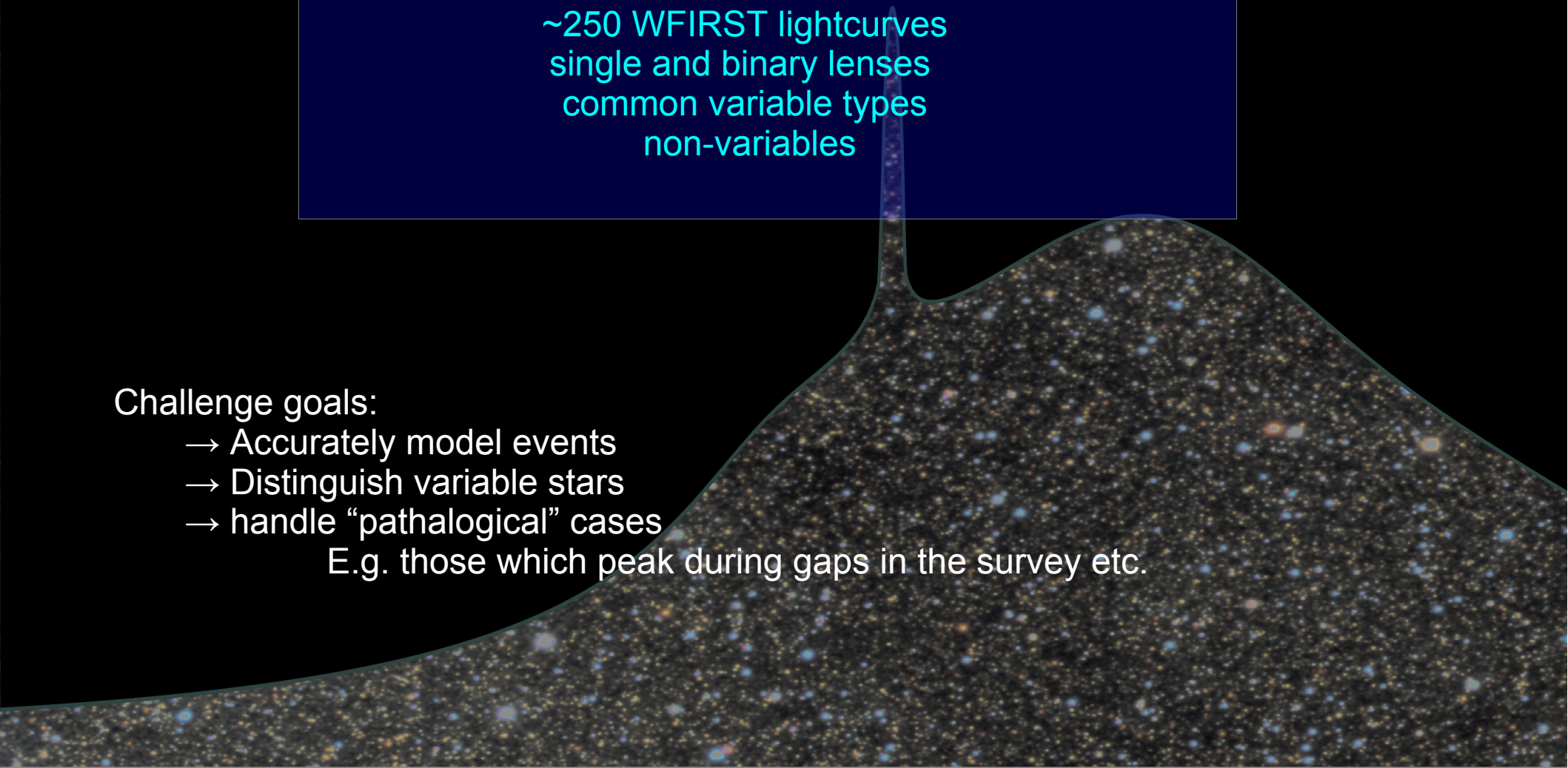
Simulated Data

- Matthew Penny has simulated the first set of lightcurves
- WFIRST lightcurves = Cadence, length and noise mimicking the nominal multi-year mission length and cadence of the Bulge survey
- Multiple datasets makes it possible for challenge-based modeling work to be conducted in parallel with the development of more sophisticated simulations.



Dataset 1

~250 WFIRST lightcurves
single and binary lenses
common variable types
non-variables



Challenge goals:

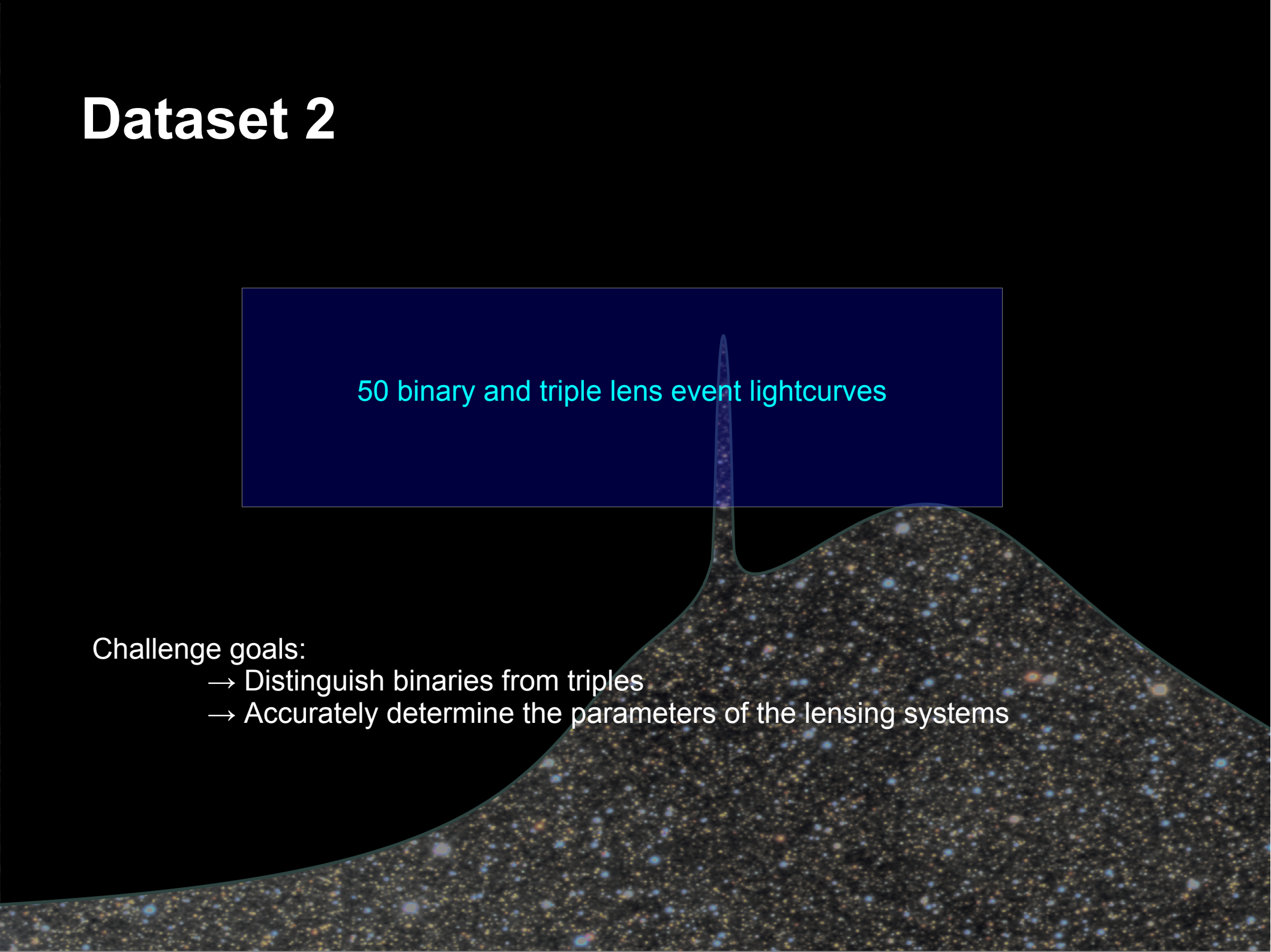
→ Accurately model events

→ Distinguish variable stars

→ handle “pathalogical” cases

E.g. those which peak during gaps in the survey etc.

Dataset 2



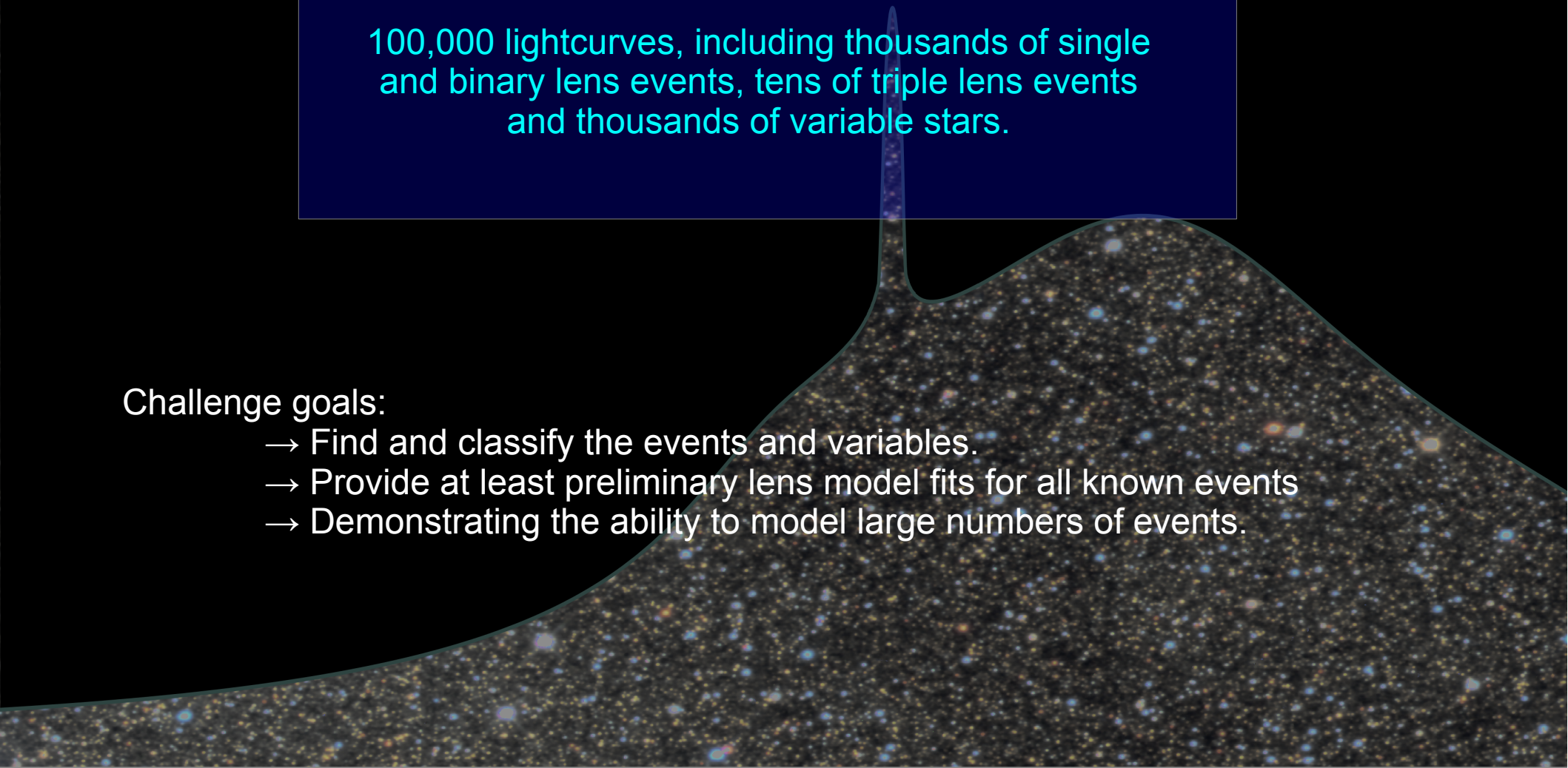
Challenge goals:

→ Distinguish binaries from triples

→ Accurately determine the parameters of the lensing systems

Dataset 3

100,000 lightcurves, including thousands of single and binary lens events, tens of triple lens events and thousands of variable stars.



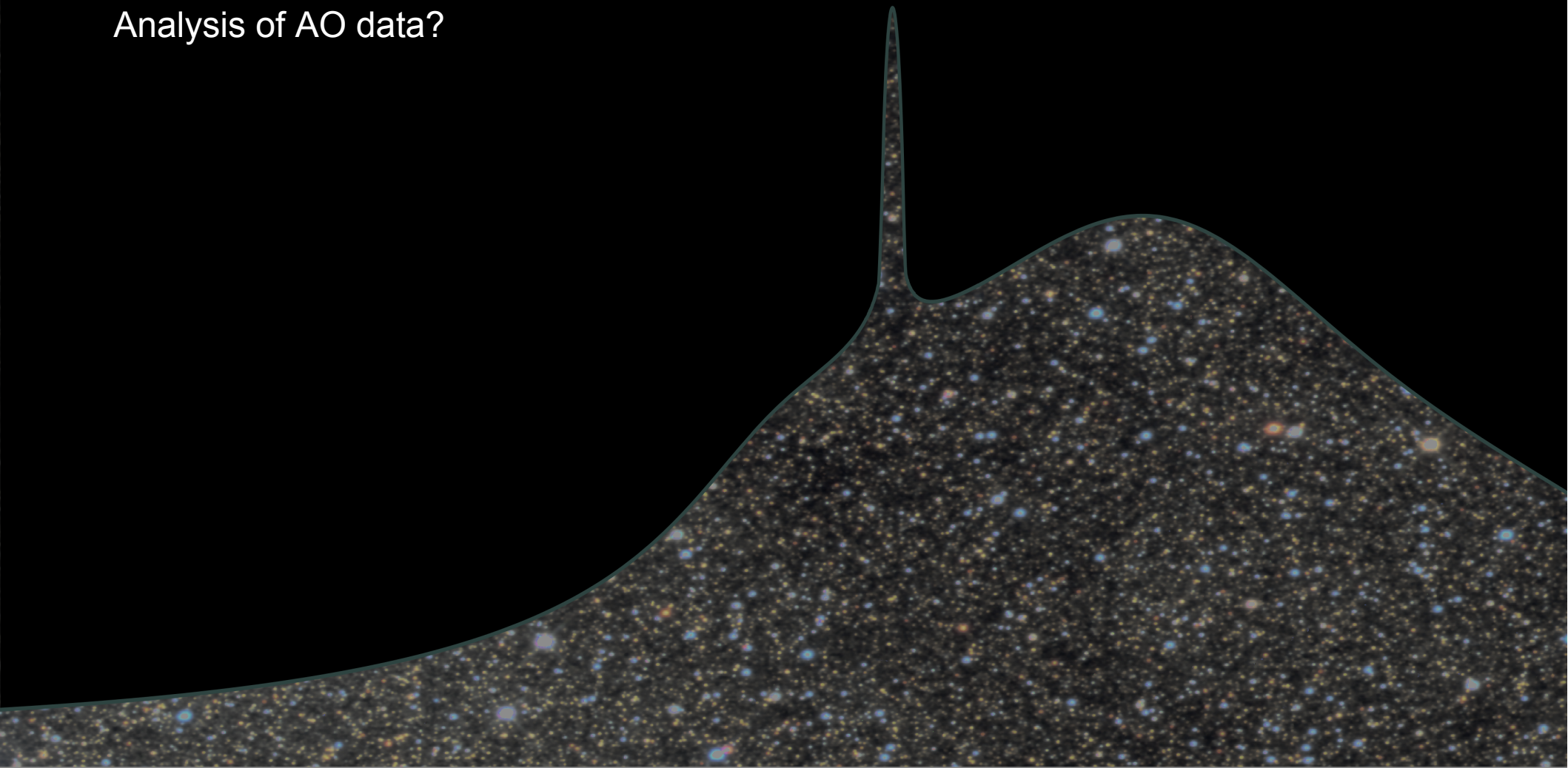
Challenge goals:

→ Find and classify the events and variables.
→ Provide at least preliminary lens model fits for all known events
→ Demonstrating the ability to model large numbers of events.

Possible Future Challenges?

Photometry in heavily blended fields?

Analysis of AO data?

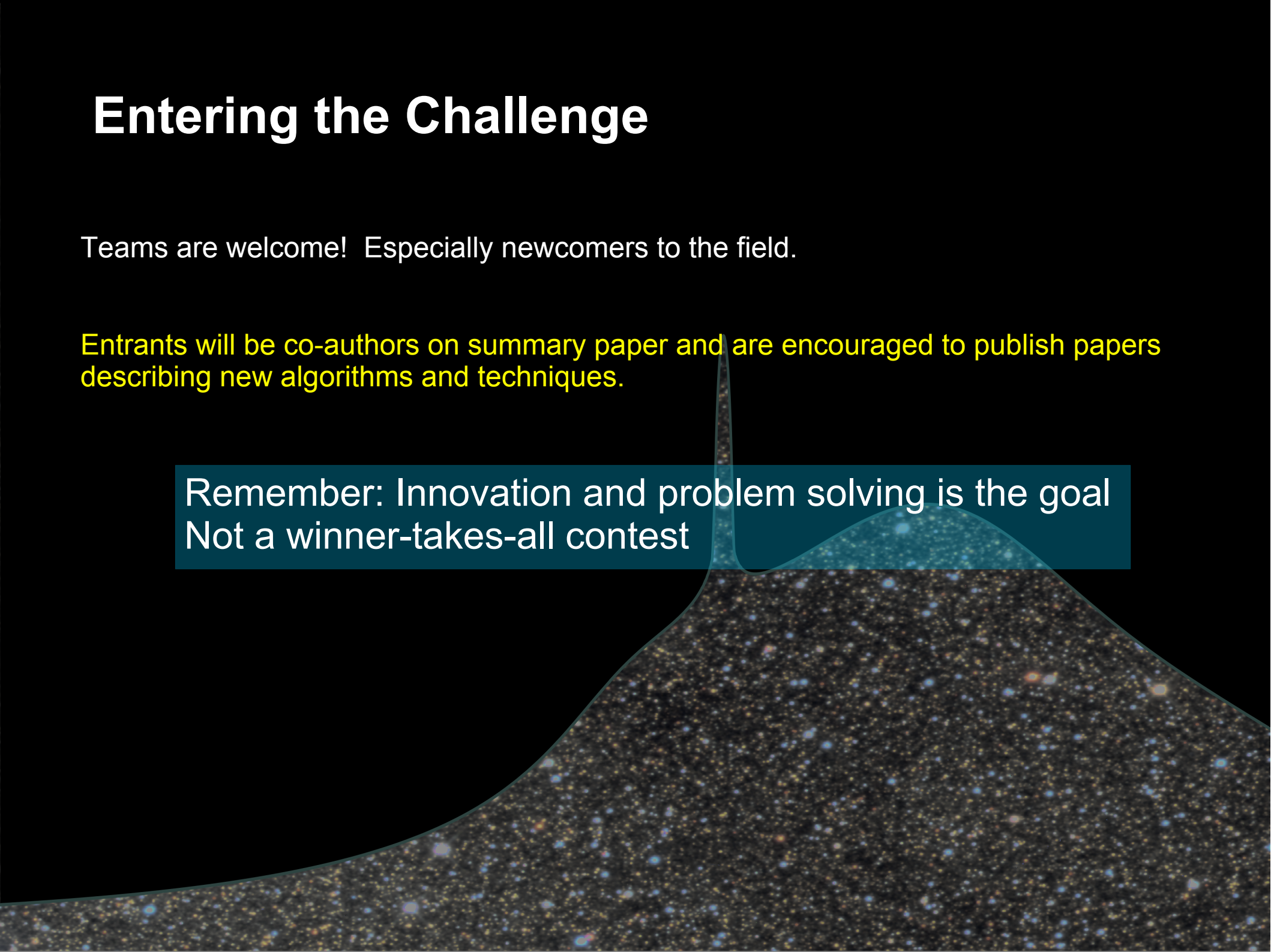


Entering the Challenge

Teams are welcome! Especially newcomers to the field.

Entrants will be co-authors on summary paper and are encouraged to publish papers describing new algorithms and techniques.

Remember: Innovation and problem solving is the goal
Not a winner-takes-all contest

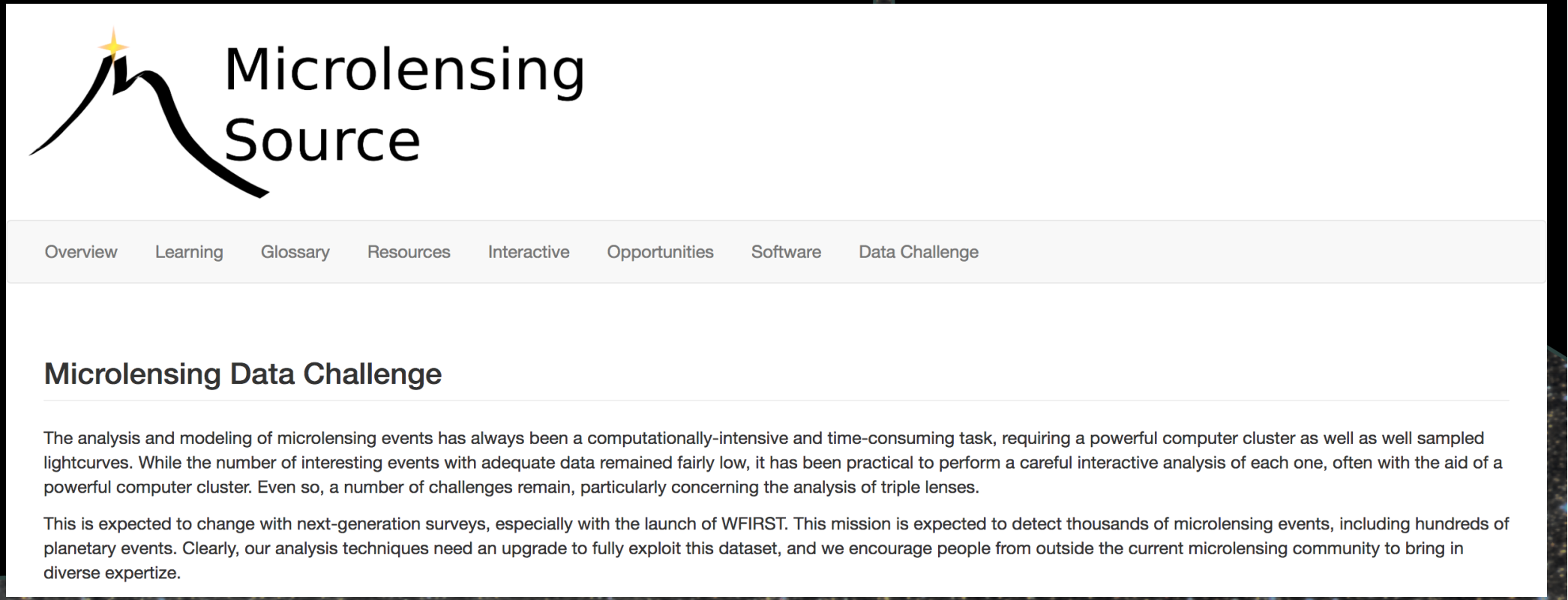


Entering the Challenge

http://microlensing-source.org/data-challenge

Mailing list: microlensing-data-challenge@lco.global

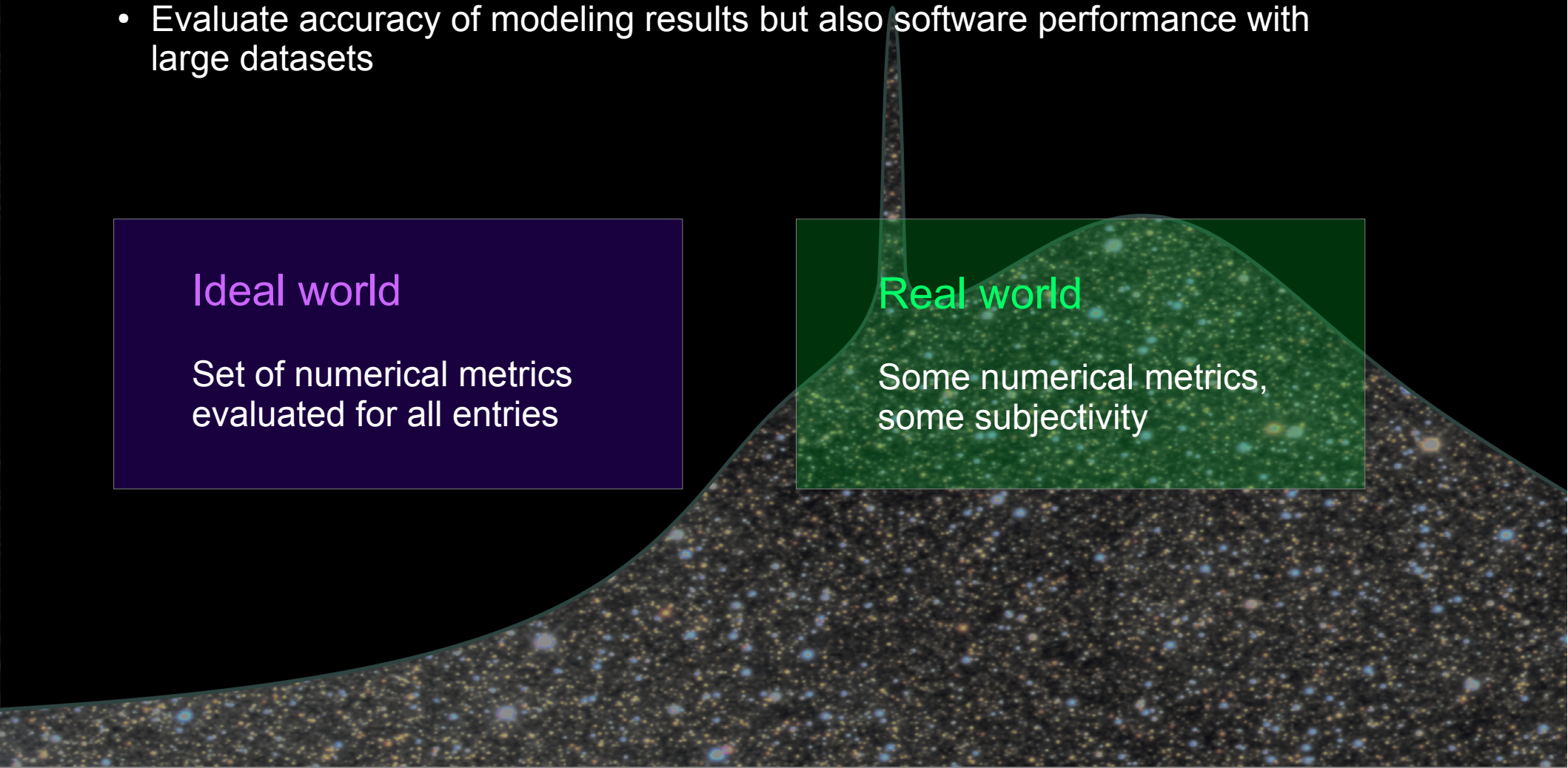
Github organization: https://github.com/microlensing-data-challenge



Evaluating Results

- Committee of people not entering the challenge

- Evaluate accuracy of modeling results but also software performance with large datasets



Ideal world

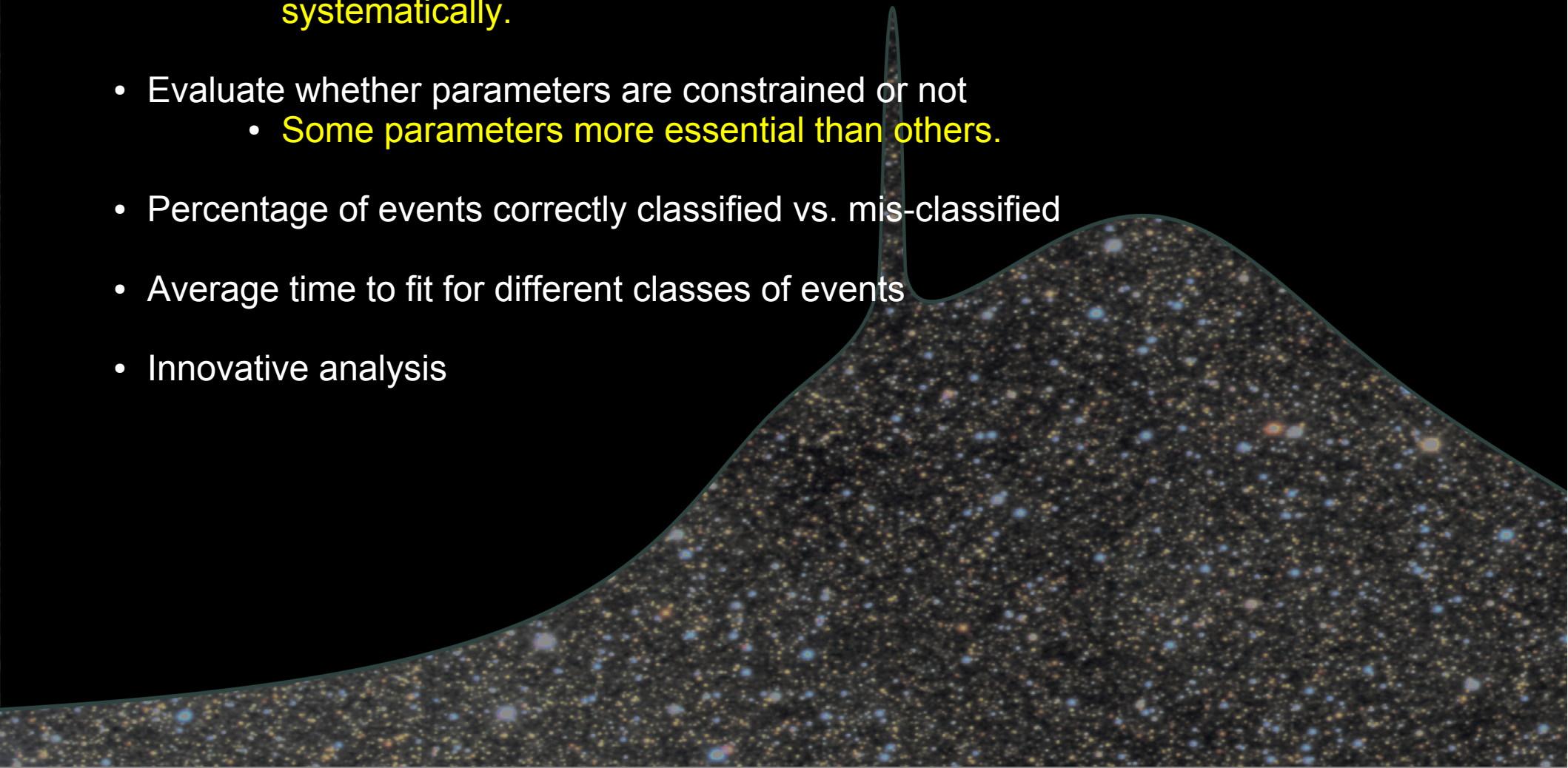
Set of numerical metrics evaluated for all entries

Real world

Some numerical metrics, some subjectivity

Evaluating Results

- Entries will be anonymized
- Compare fitted model parameters with the input parameters for the same event.
    - Identify any regions of parameter space where they diverge systematically.
- Evaluate whether parameters are constrained or not
    - Some parameters more essential than others.
- Percentage of events correctly classified vs. mis-classified
- Average time to fit for different classes of events
- Innovative analysis



Entry Requirements

- Analyze the test dataset
- Submit before the deadline
- Submitted data products must include (but not limited to):

→ Summary table of fitted event parameters with uncertainties – in format specified

→ Technical specifications of the computing resources used

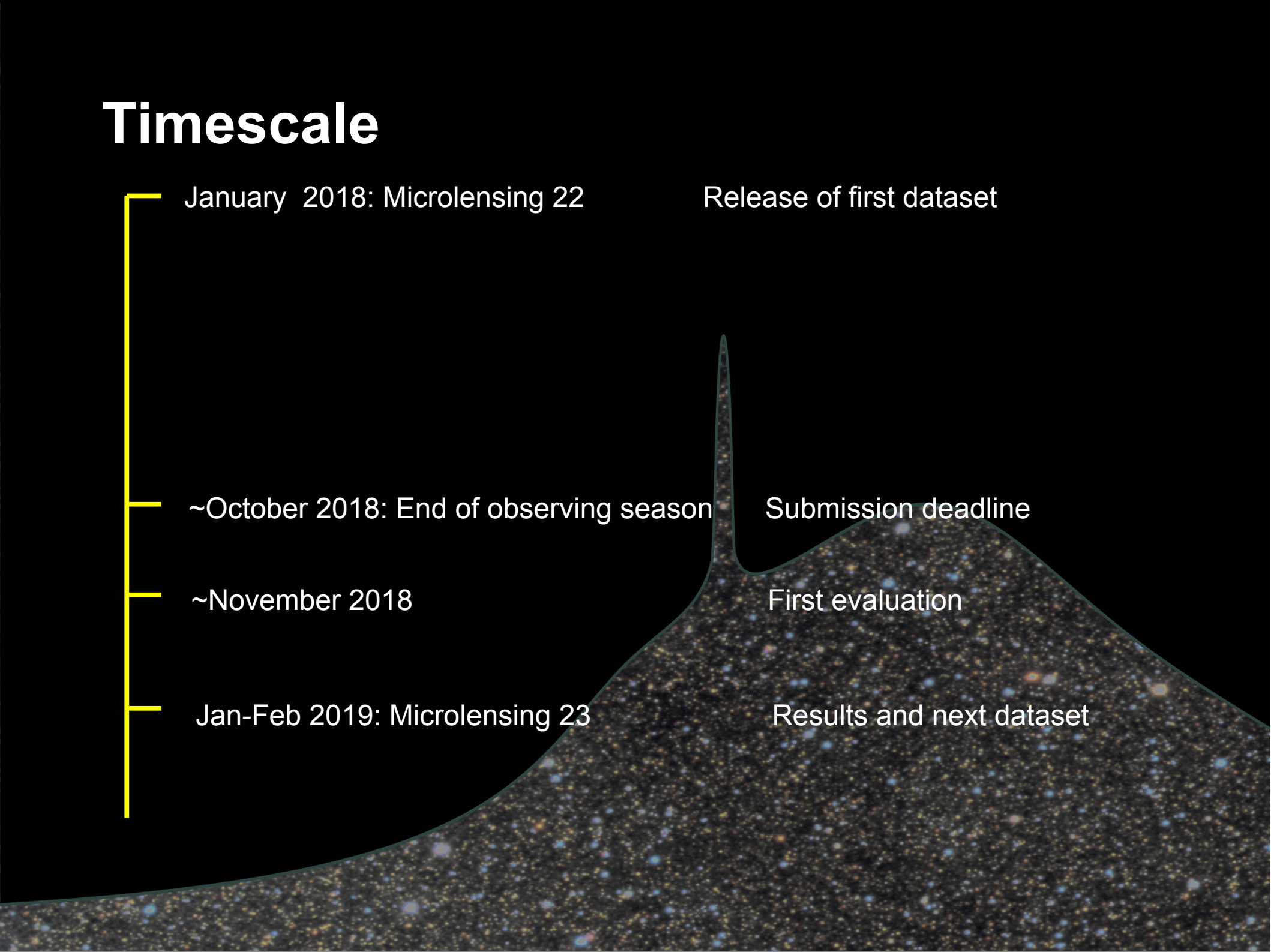
→ Description of software used including the language(s), libraries or package dependencies.

→ Time taken to model each event

→ Plots of the lightcurves with the fitted models overlaid and residuals.

→ Plots of the lens plane geometry, caustic structures and source trajectory.

Timescale


January 2018: Microlensing 22
Release of first dataset
~October 2018: End of observing season
Submission deadline
~November 2018
First evaluation
Jan-Feb 2019: Microlensing 23
Results and next dataset

Summary

- First challenge dataset is now available
  http://microlensing-source.org/data-challenge
- Deadline: Oct 31, 2018
- Invite participants, newcomers welcome
- Results will be published and presented at the next microlensing meeting