ПРУЖИНСКАЯ М.В.[1], КОРНИЛОВ М.В.[1,2], МАЛАНЧЕВ К.Л.[1,2], ИШИДА Э.[3], МОНДОН Ф.[3], ВОЛЬНОВА А.А.[4], КОРОЛЕВ В.С.[5]

[1] ГАИШ, МГУ имени М.В. Ломоносова, Москва, Россия
[2] Национальный исследовательский университет «Высшая школа экономики», Москва, Россия
[3] Университет Клермон-Овернь, Клермон-Ферран, Франция
[4] Институт космических исследований РАН, Москва, Россия
[5] Центральный аэрогидродинамический институт имени профессора Н. Е. Жуковского, Жуковский, Россия

ПОИСК АНОМАЛИЙ В ОТКРЫТОМ КАТАЛОГЕ СВЕРХНОВЫХ ЗВЕЗД МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ

SNAD — SUPERNOVA ANOMALY DETECTION







https://pruzhinskaya.com/snad/

SNAD — SUPERNOVA ANOMALY DETECTION



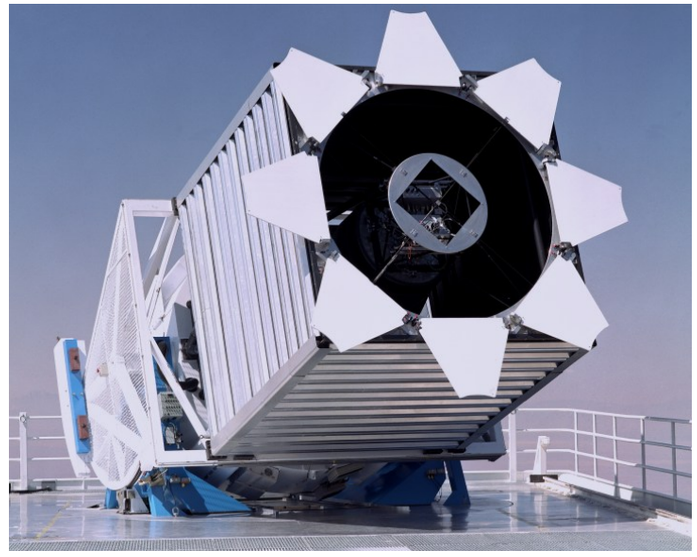


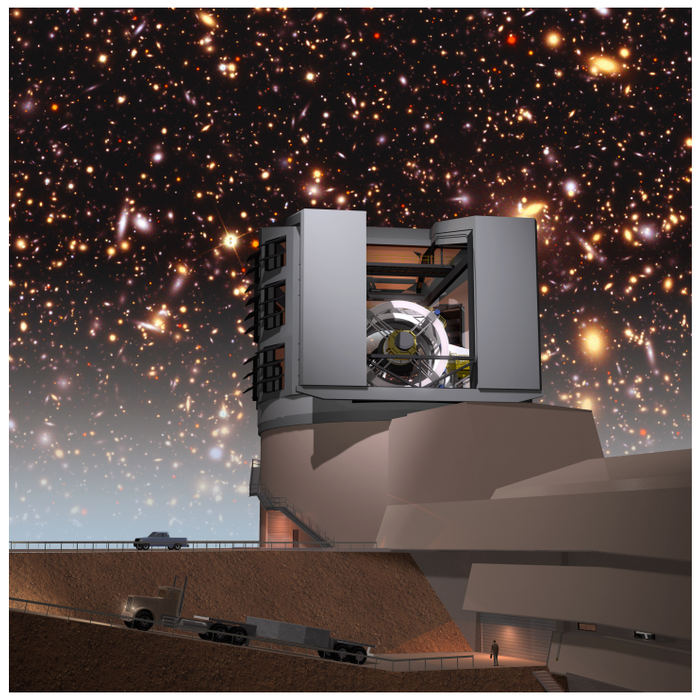


https://pruzhinskaya.com/snad/

БОЛЬШИЕ ДАННЫЕ В АСТРОНОМИИ

- SDSS
  - ~500 миллионов объектов
  - ~3 миллионов со спектрами
- LSST
  - 10 PB в ночь
  - 10 лет работы – база данных Google
  - 10^6 сверхновых в год



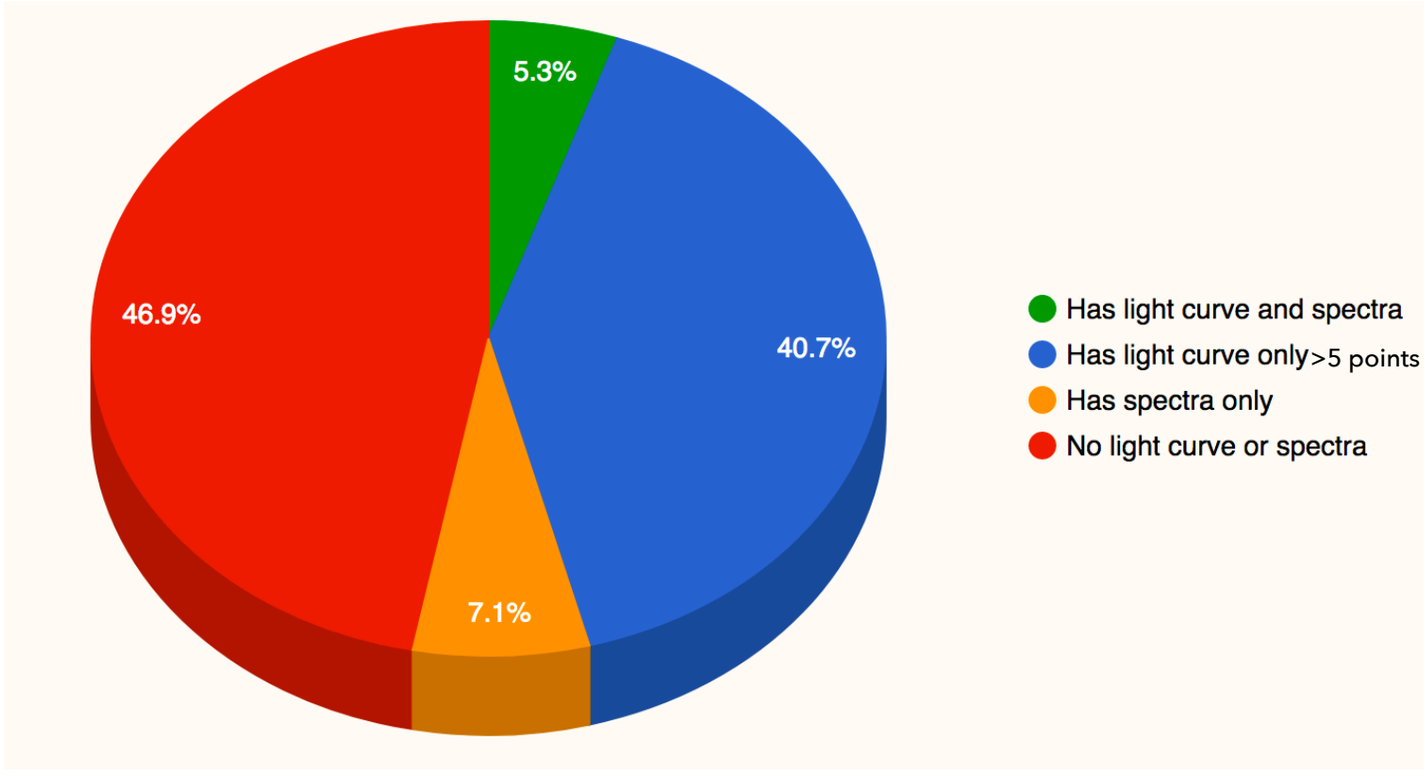


МАШИННОЕ ОБУЧЕНИЕ

- «с учителем» и «без учителя»
- активно используется в медицине, генетике, экологии
- редко в астрономии:
  - классификация (Kessler et al. 2010; Ishida & de Souza 2013; Lochner et al. 2016)
  - регрессия (Hildebrandt et al. 2010; Cavuoti et al. 2015)
- поиск аномалий особенно важен для широкопольных транзиентных обзоров

ОТКРЫТЫЙ КАТАЛОГ СВЕРХНОВЫХ ЗВЁЗД

- https://sne.space (Guillochon et al. 2017)
- ~55 000 сверхновых и кандидатов в сверхновые
- ~600 000 фотометрических измерений
- ~20 000 спектров

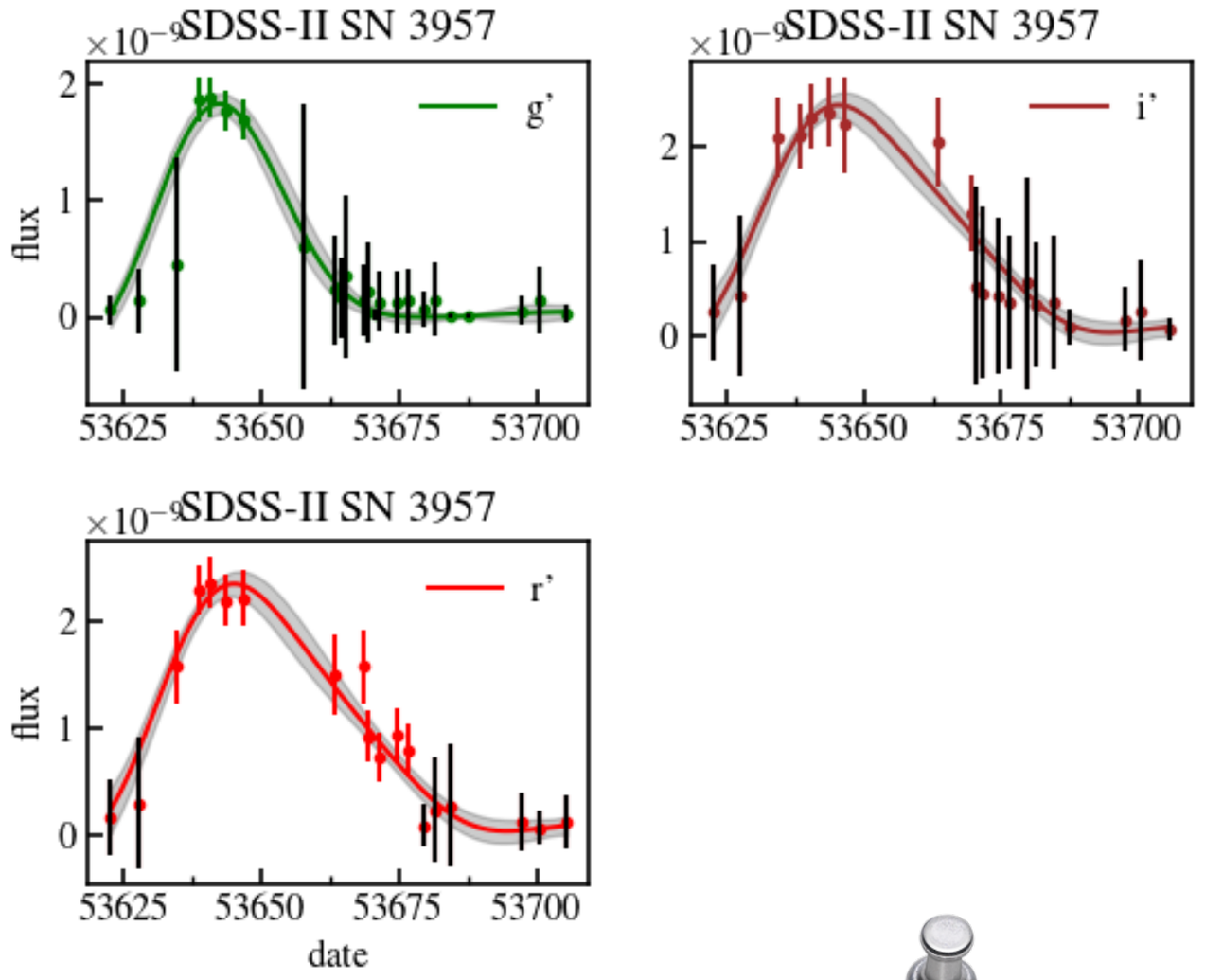


ЧТО ТАКОЕ АНОМАЛИЯ?

- артефакты в данных:
  - ошибки измерений
  - ошибки при переносе данных из источника
- ошибки классификации
- редкие классы объектов (сверхъяркие СН, СН-ГВ, килоновые)
- новые объекты – новая физика

РАБОТА С ДАННЫМИ

- Проблемы:
  - разные фильтры
  - неравномерное покрытие по времени
  - пределы
  - отсутствие ошибок
  - …



- gri, g′r′i′, BRI ⟶ gri фильтры
- минимум 3 измерения в каждом фильтре (bin = 3 дня)
- Фит Гауссовыми процессами (Kornilov et al. 2019, in prep.) , https://gp-multistate-kernel.readthedocs.io

1999 объектов

ПОИСК АНОМАЛИЙ

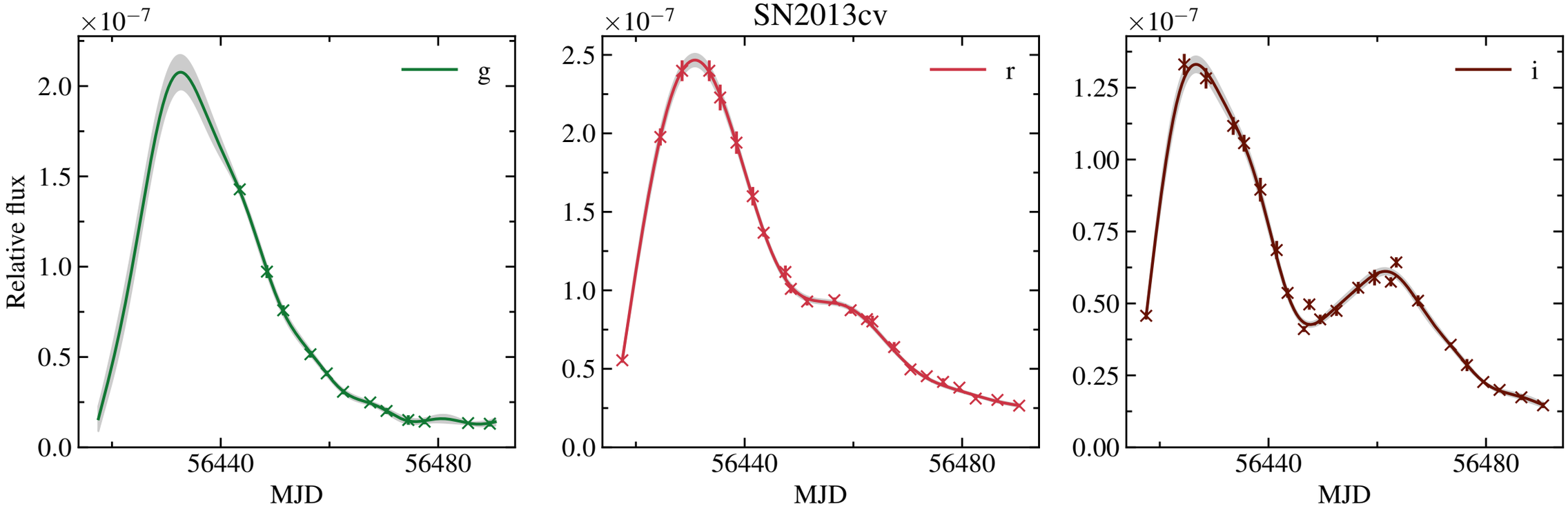
- Итого 373 признака (features):
  - 363 фотометрических точки (121 день, 3 фильтра)
  - 10 параметров фита ГП
- Снижение размерности
- Алгоритм «Изолирующий лес» («Isolation forest»)



~100 аномалий

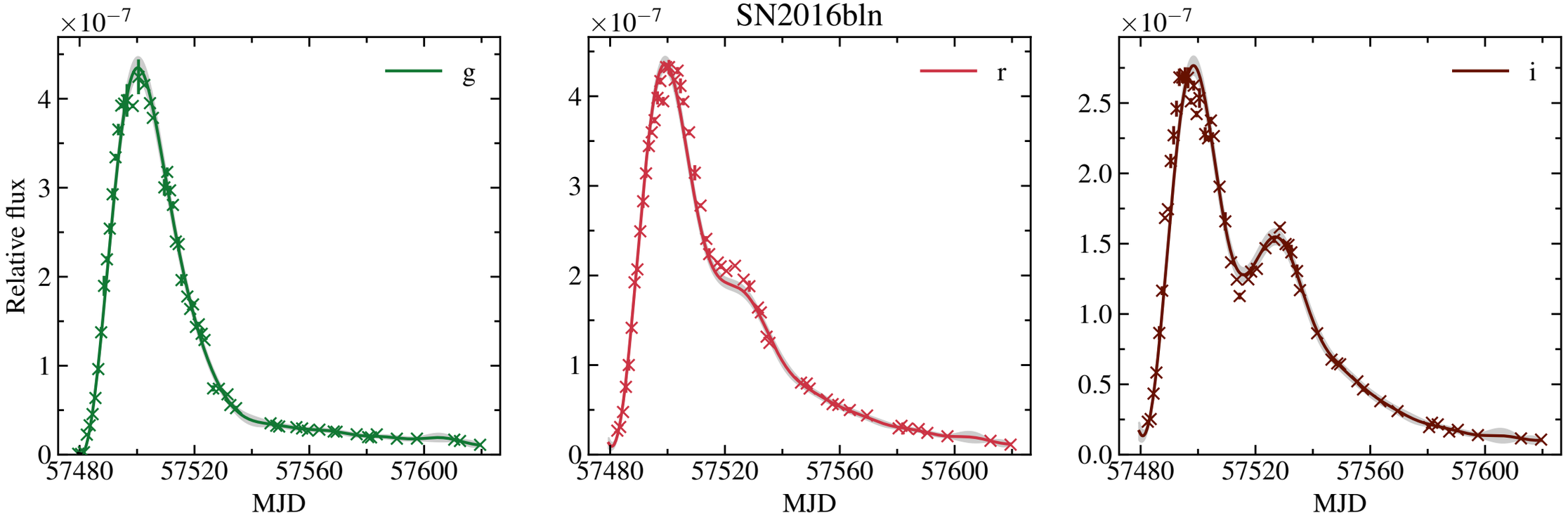
СВЕРХНОВЫЕ ТИПА 91-T

- SN 2013cv (Cao et al. 2016)



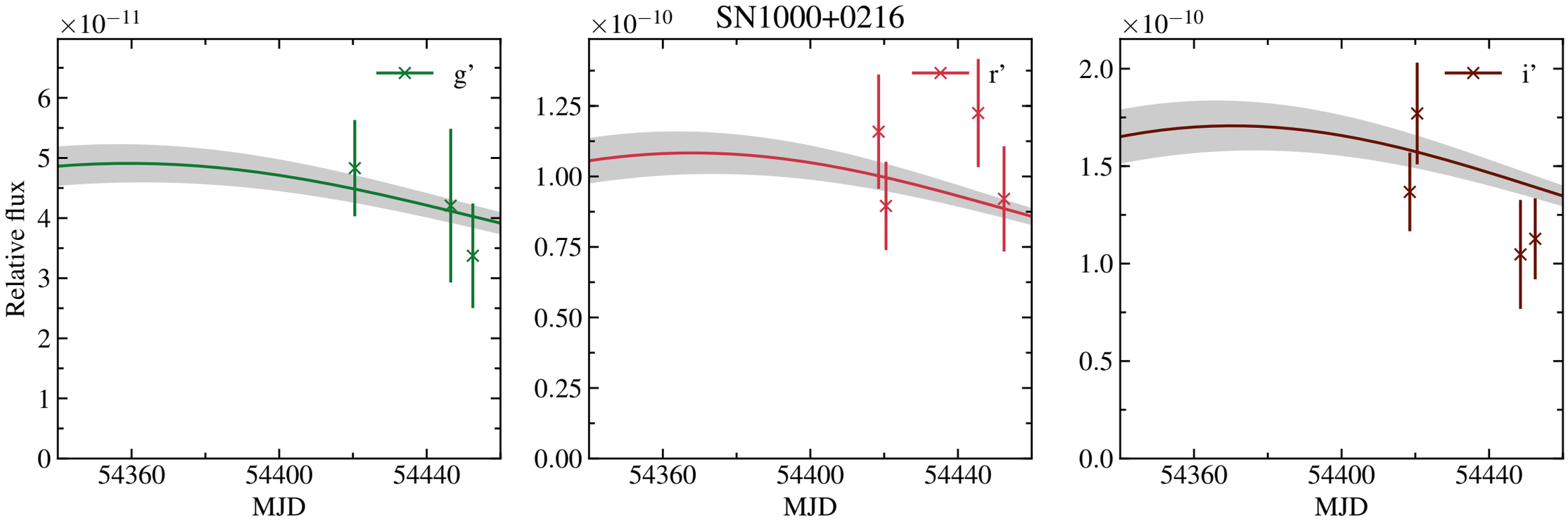
СВЕРХНОВЫЕ ТИПА 91-T

- SN 2016bln (Cenko et al. 2016, ATel 8909)



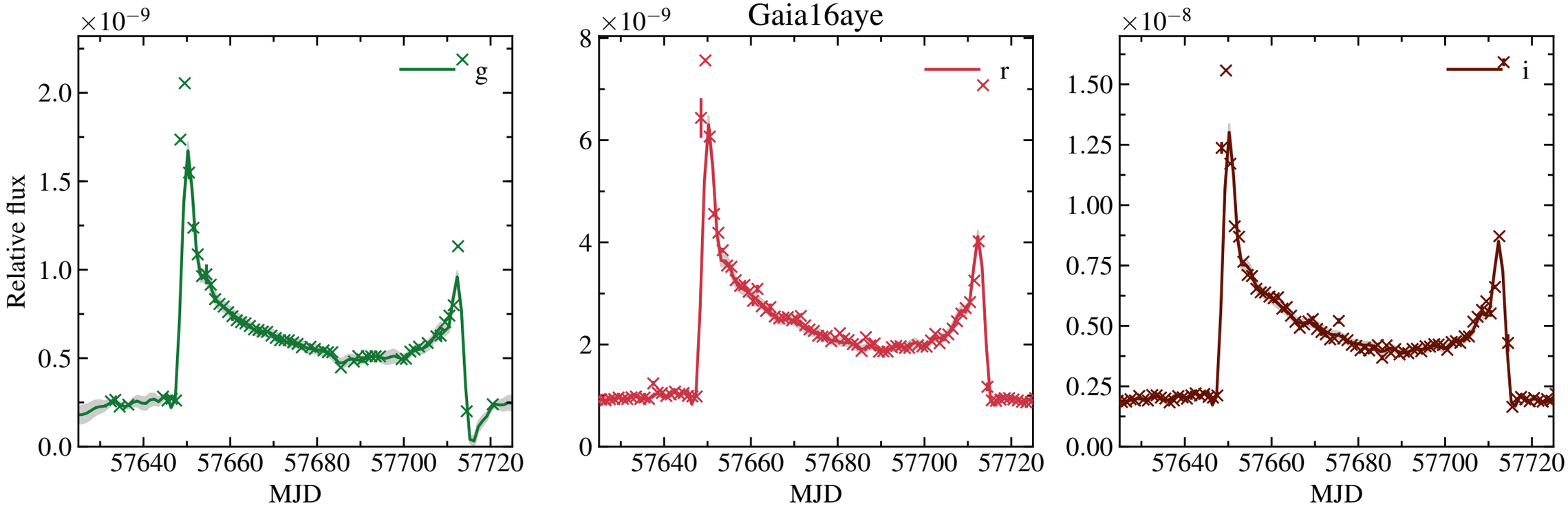
СВЕРХЪЯРКИЕ СВЕРХНОВЫЕ

- SN1000+0216 (SLSN-II, Cooke et al. 2012)



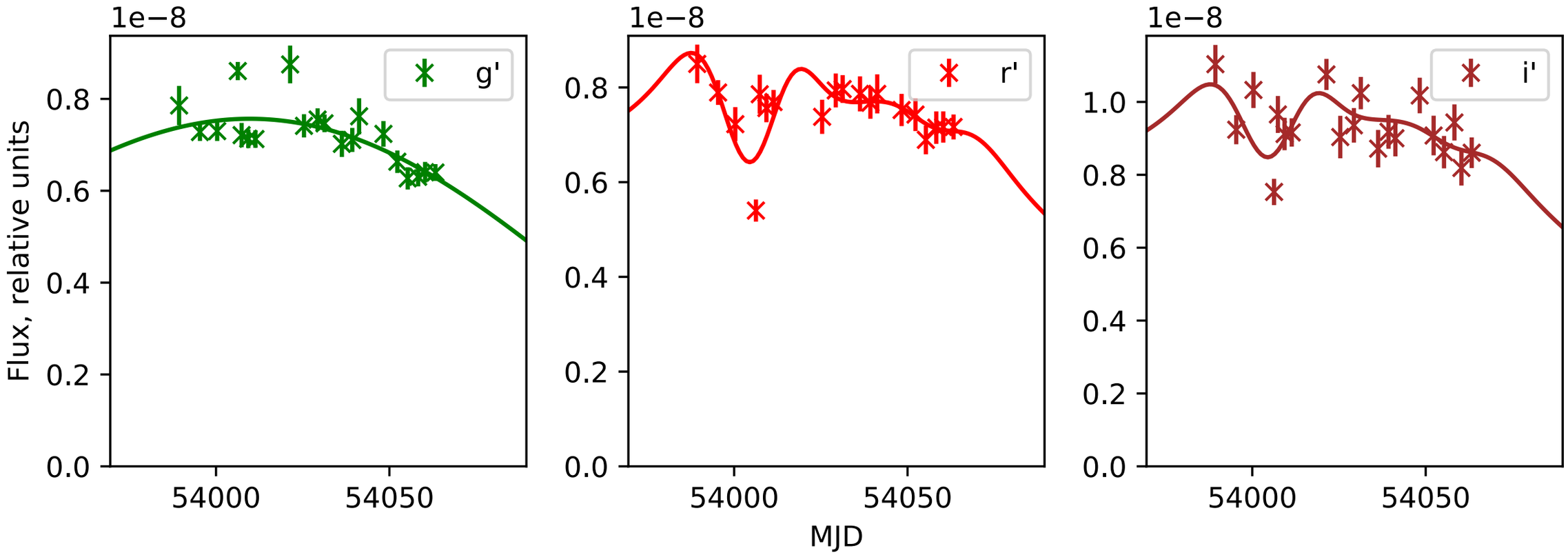
ДВОЙНОЕ МИКРОЛИНЗИРОВАНИЕ

- Gaia16aye (Wyrzykowski et al. 2016, ATel 9507)



АКТИВНЫЕ ЯДРА ГАЛАКТИК

- SN 2006kg (Sako et al., 2018)



СПАСИБО ЗА ВНИМАНИЕ!