PLAsTiCC v2.0
Or How to Prepare Brokers & Science Collaborations for LSST

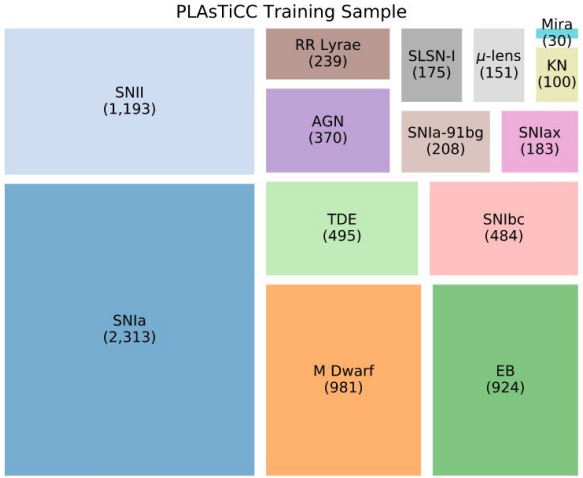
Alex Gagliano
Martine Lokken
Renée Hložek
Gautham Narayan

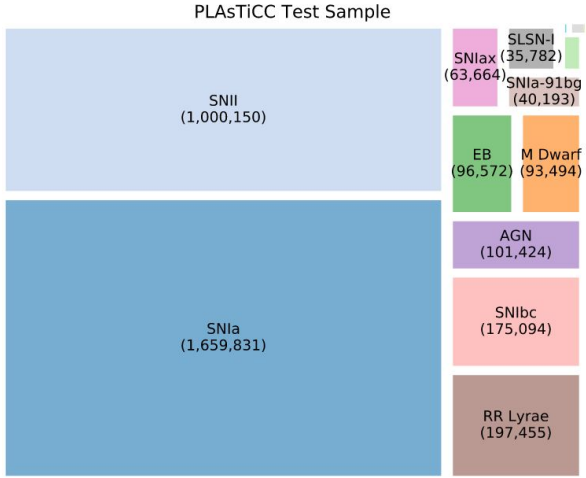
LSSTC enabling science 2021 broker II Day 3
April 15, 2021

LSST DESC
Dark Energy Science Collaboration

PLAsTiCC Version 1 (Dec. 2018 - Feb 2019)

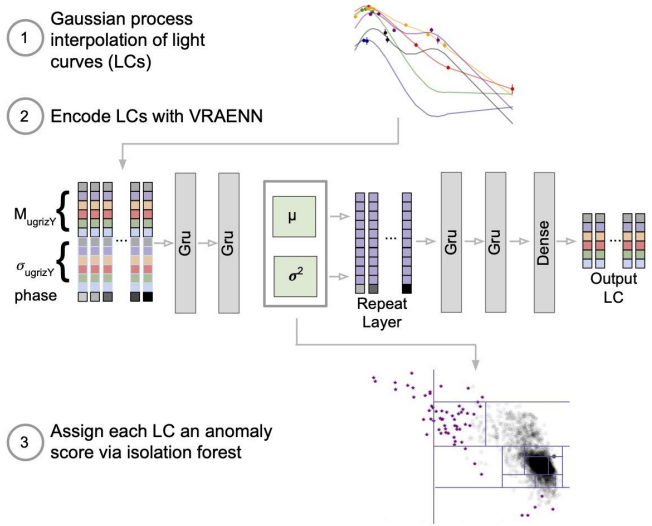
- Public Kaggle challenge for photometric classification of transients
- Data: 3M VRO-simulated *ugrizY* lightcurves
- Primary goal: setup massive time-domain simulation infrastructure, **jump start ML photometric classification efforts**





(Hložek+2020)

(Villar+2021)

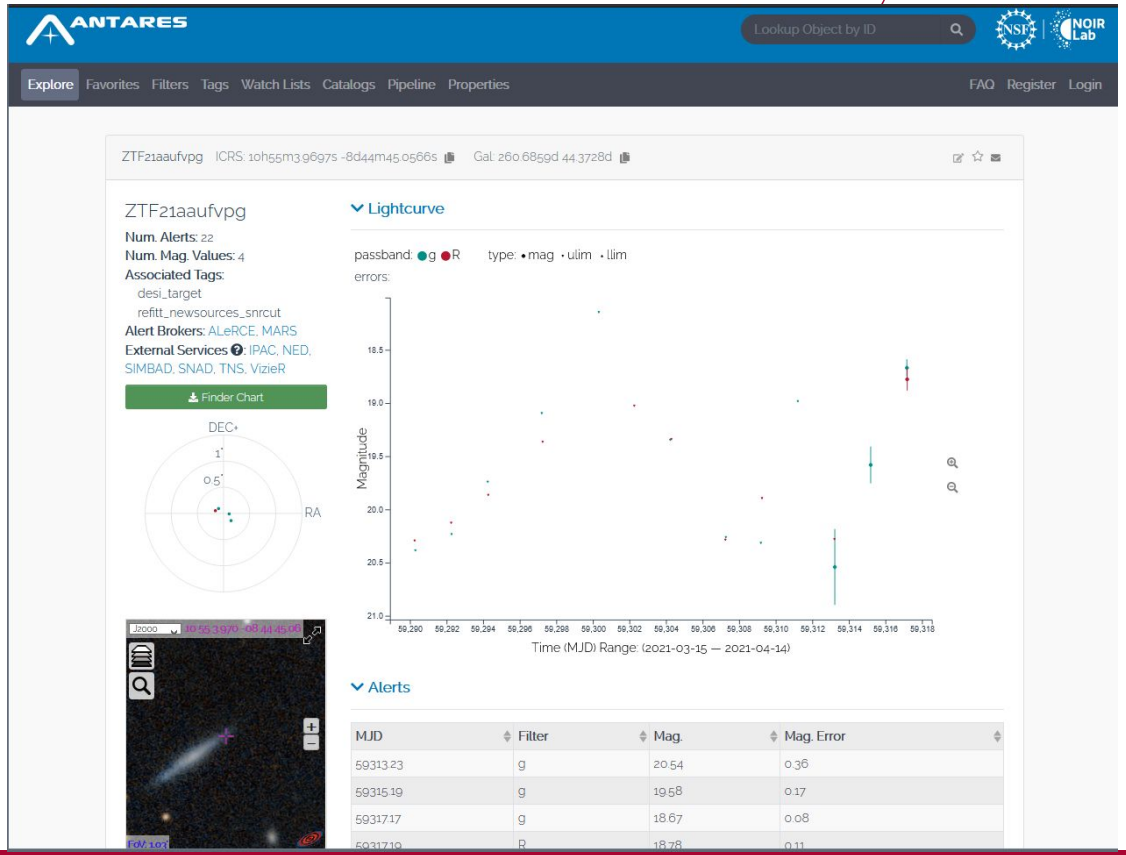


(Some known) shortcomings of v1

What tells you that 21jap at right is real?

No host galaxy information, postage stamps, or alerts

Simply, flat light curves with full phase coverage aren't a realistic model for how the science collaborations will interact with LSST data



Why broker teams should care about PLAsTiCC v2

LSST DESC
Dark Energy Science Collaboration

Goal for PLAsTiCC v2 (~Summer 2021)

To evaluate real-time Broker performance on a realistic LSST alert stream.*

Broker Roles

- Storing, processing, classifying alerts, informing follow-up
- *Potential additional roles:*
  - Collecting active source features
  - Maintaining source databases

VERIFICATION

Alert Stream

- Set of LSST-like alert packets
- must preserve environmental correlations
- Should contain a representative sample of expected events

VALIDATION

Iterative process to ensure compatibility between brokers and alert stream.

**metrics in progress!*

v2: Simulating Transients & Hosts

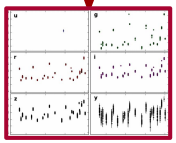
Pre-processing *Simulation* *Validation*

GHOST
16.5k SNe, host galaxies
arXiv:2008.09630
Gagliano+2021

cosmoDC2
DESC Synthetic Sky Catalog
arXiv:1907.06530
Korytov+2019

SNANA
Simulated transient photometry (with host galaxy properties)
arXiv:0908.4280
Kessler+2009

EmpiriciSN
Simulated transient parameters
arXiv:1611.00363
Holoien+2016



(Oluseyi+2012)

(Abolfathi+2021)

LIGO
Localized Event Skymaps
D. Chatterjee

Brokers

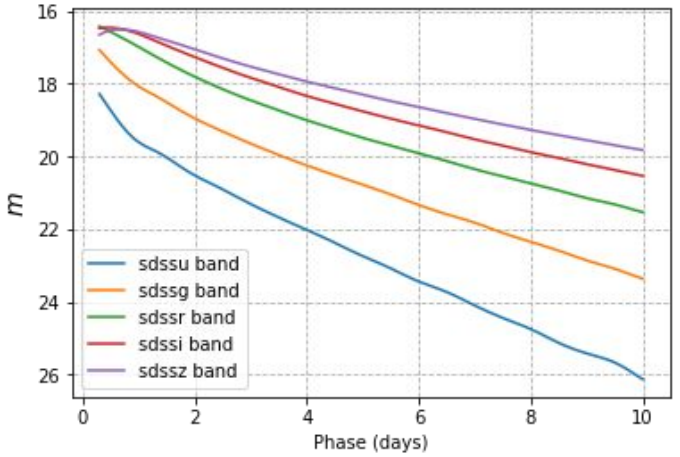
- ANTARES
- Pitt - Google
- ALeRCE
- LASAIR
- Fink
- AMPEL
- MARS
- INAF broker
- Fritz
- Babamul
- South African broker team
- NYU Anomalies
- SNAPS
- UW Genesis

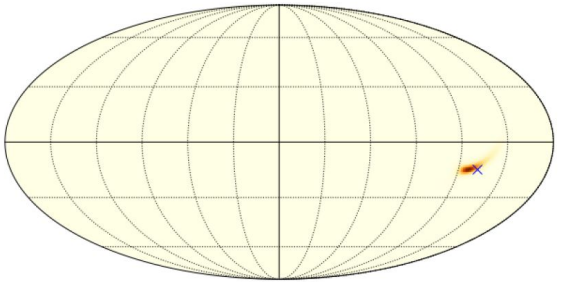
v2 Alert Stream

Gagliano & Lokken

New in v2: Updated Transient Models

- *New:* delta Scuti, Cepheids (K. Malanchev), dwarf novae (Q. Cheng)
- *Updated:*
  - **SNe Ib & c** (Vincenzi models)
  - **M-dwarfs** (V. Shah)
  - **KNe** (D. Chatterjee, Bulla models; data on github)
    - SED models added to SNANA
    - Parameterized by - ejecta mass, lanthanide fraction, observing angle





Chatterjee

New in v2: Realistic Host Galaxies

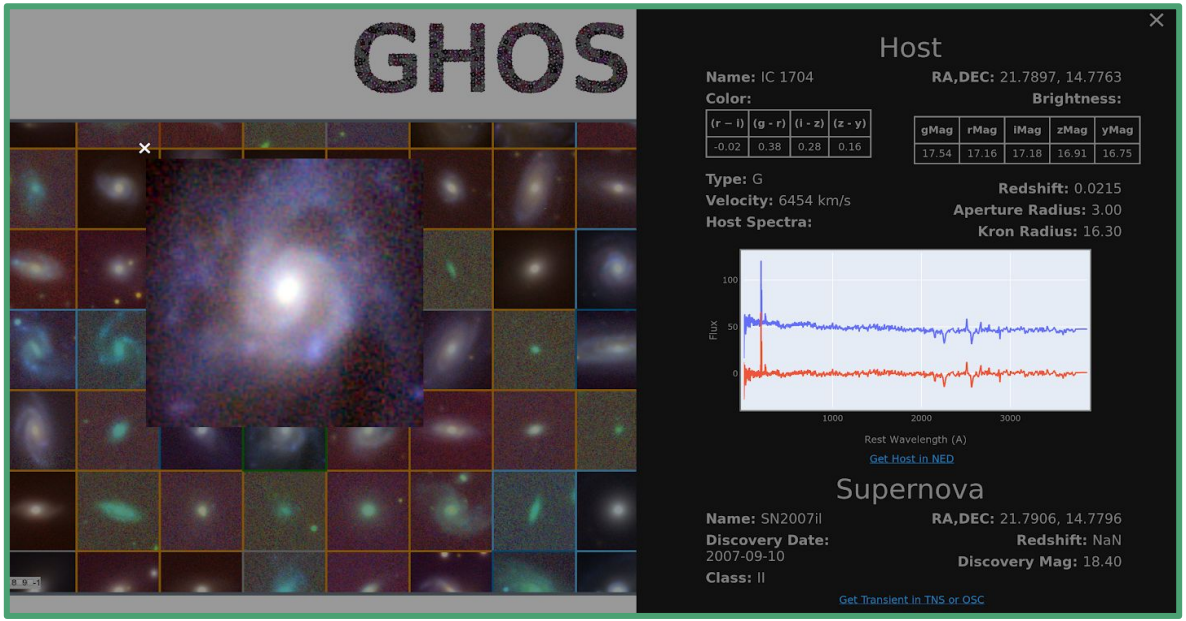
1. *Transient abundances for galaxies of different properties*
2. *Correlations of transient properties with galaxy properties.*
3. Hostless transients
   - ***Observational:*** High-*z* SNe at the magnitude limit
   - ***Intrinsic:*** SNe without host galaxies
   - ***Algorithmic:*** Limitations in host associations
4. Transient position with respect to galaxy
5. Posterior PDFs for photo-zs (To replace point estimates)

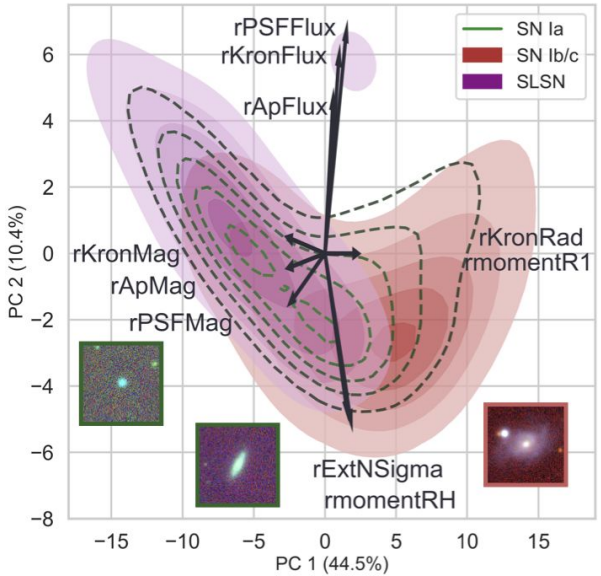
GHOST

Galaxies **HO**sting **S**upernovae and other **T**ransients

LSST DESC Dark Energy Science Collaboration

http://ghost.rubin.science/





(Gagliano+2021)

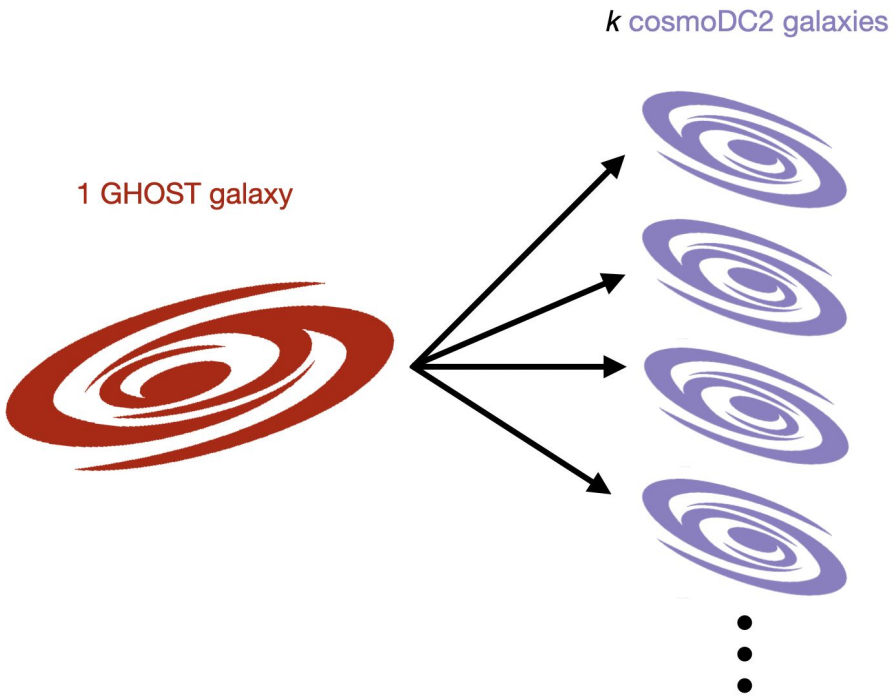
- 16,228 SNe-host galaxy pairs: 78% of unique events reported on TNS/OSC.
- PS1,NED photometric & derived properties (color, redshift, radial moments)

Matching GHOST to CosmoDC2

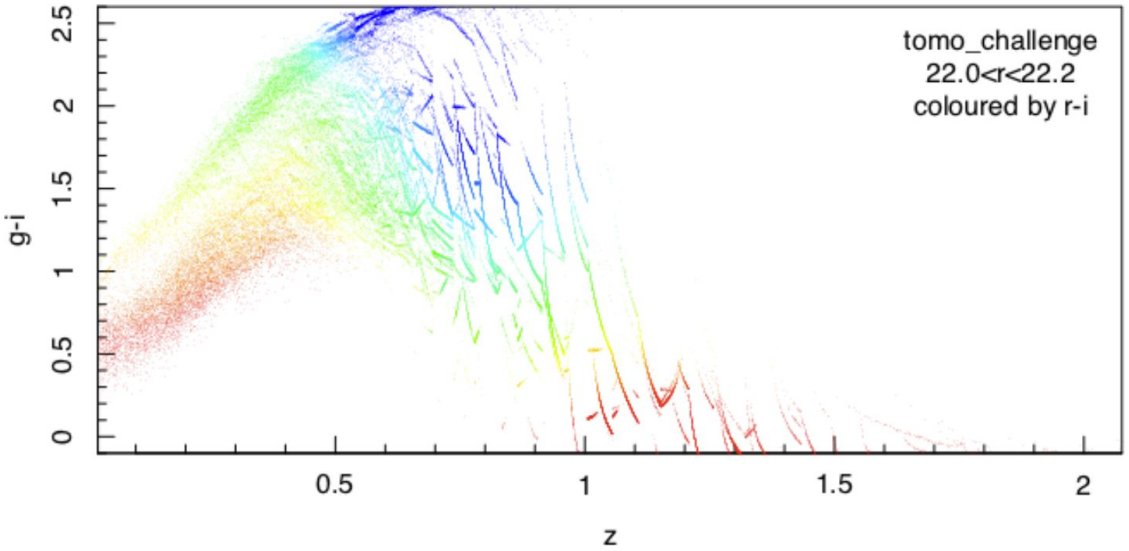
1. ***Select redshift-independent properties:***
   - Rest frame absolute magnitudes (R, I)
   - Rest frame colors (g-r, i-z)
   - Galaxy ellipticity

2. ***Normalize all properties:***
   - remove mean and scale to unit variance
   - Down-weight redshift

3. **Find *k* nearest neighbors:**
   - scikit-learn Nearest Neighbors
   - 6D parameter space
   - *k*=1000



Discretized DC2 Properties



Problems:

- Discreteness in DC2 colors (from underlying SED model) leads to unrealistic distribution of host galaxies.

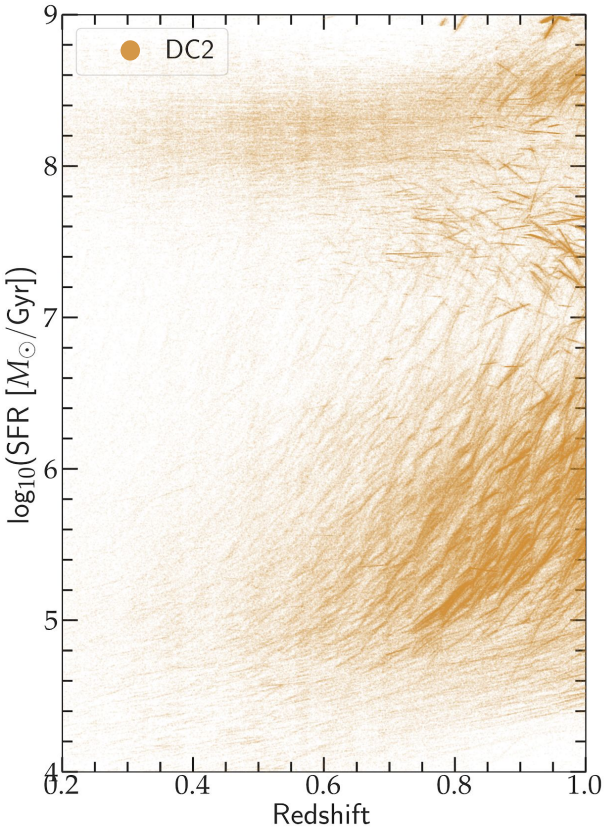
- Oversampled galaxies (e.g. `imblearn`) have no postage stamps

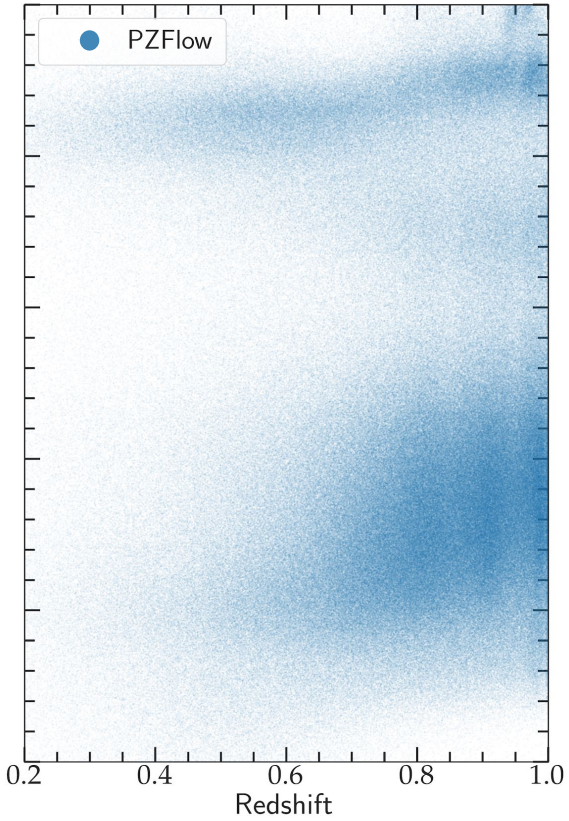
Predicting Properties using PZFlow

Solution: Oversample using normalizing flows (`PZflow`)

Properties conditioned on DC2 color, magnitude, size, ellipticity

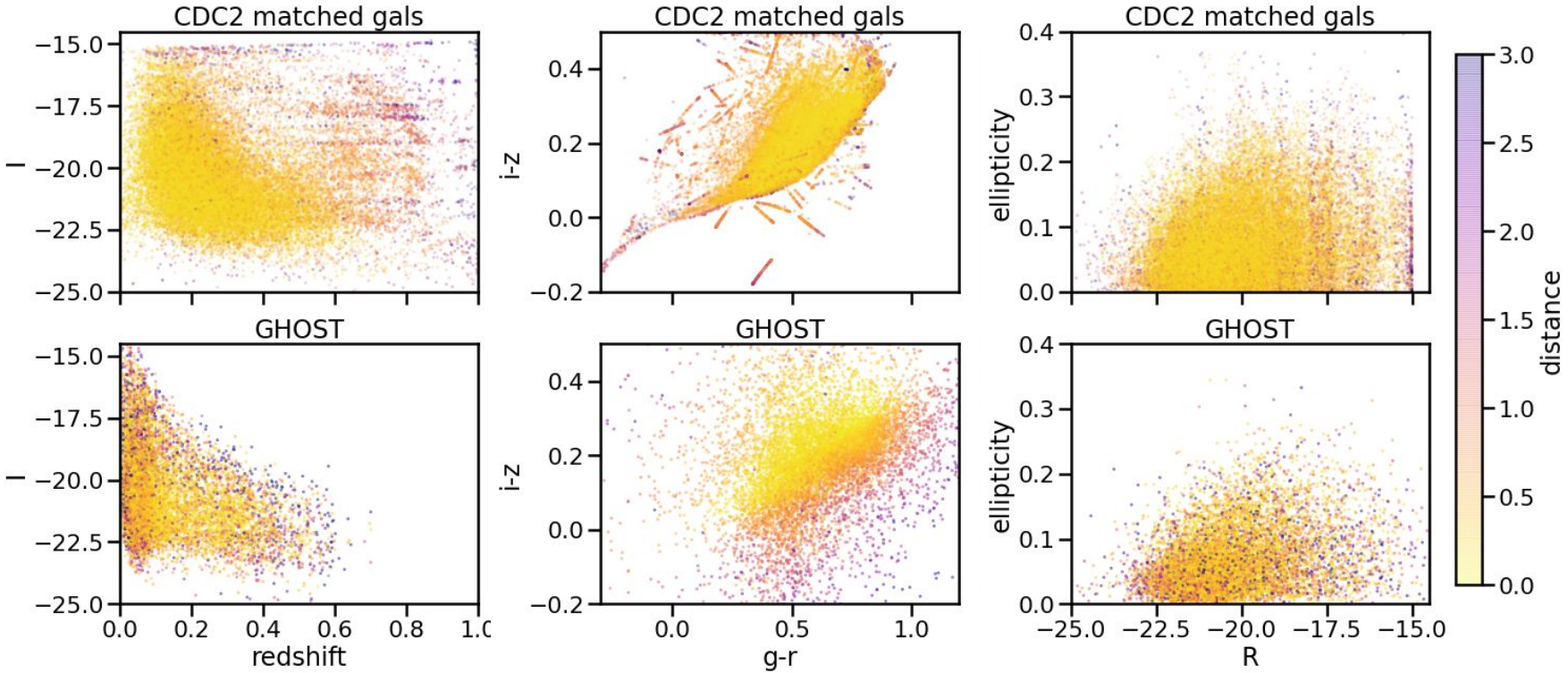
Distributions smoother than DC2 -- no need to re-match new galaxies



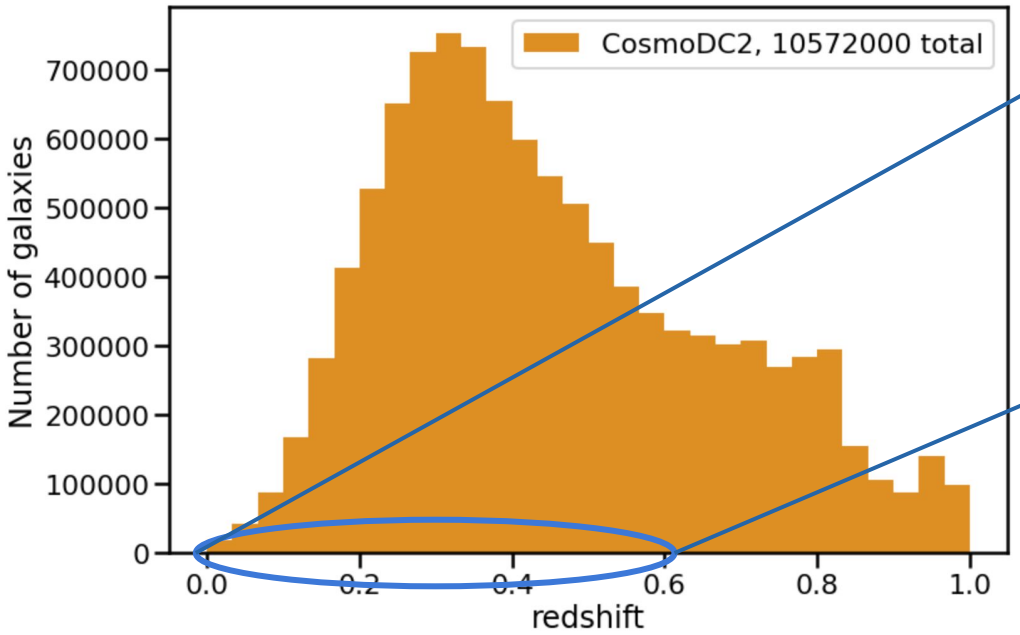


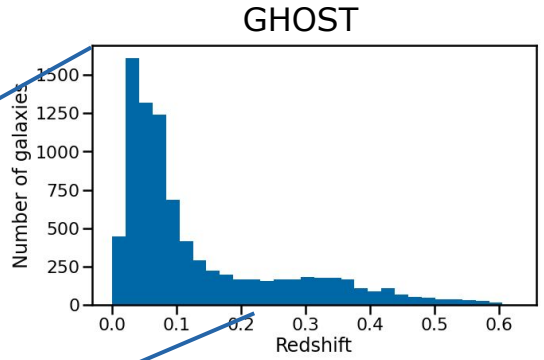
Validating Simulated SNe + Hosts

1. *Observed GHOST vs cosmoDC2 host galaxy properties*



Matching GHOST to CosmoDC2





Scaling up by 1000x

Host properties available in V2

LSST DESC
Dark Energy Science Collaboration

Host Galaxy Properties in Final Catalog

Definite:

- LSST grizY magnitudes, colors
- Photo-zs
- Apparent morphology, e.g. 1st and 2nd order moments, half-light radius
- Offset of transient from host nucleus

Possible:

- Directional light radius
- Local host-galaxy surface brightness (at the location of the transient)
- Derived morphology, e.g. Sersic index
- Postage stamps!

Unlikely (but we have them):

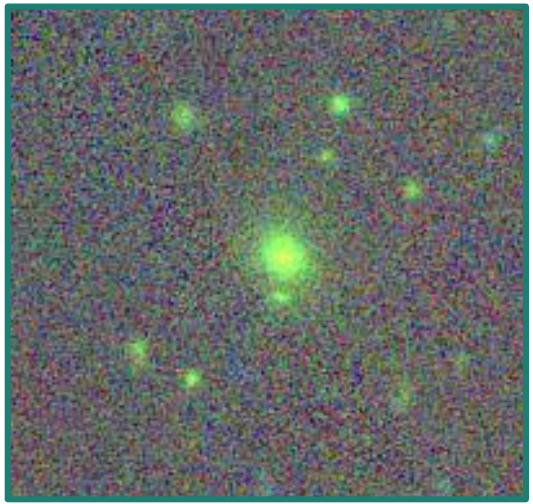
- Derived properties e.g. Star Formation Rate, Stellar Mass, etc.

Possible Uses of Host Information

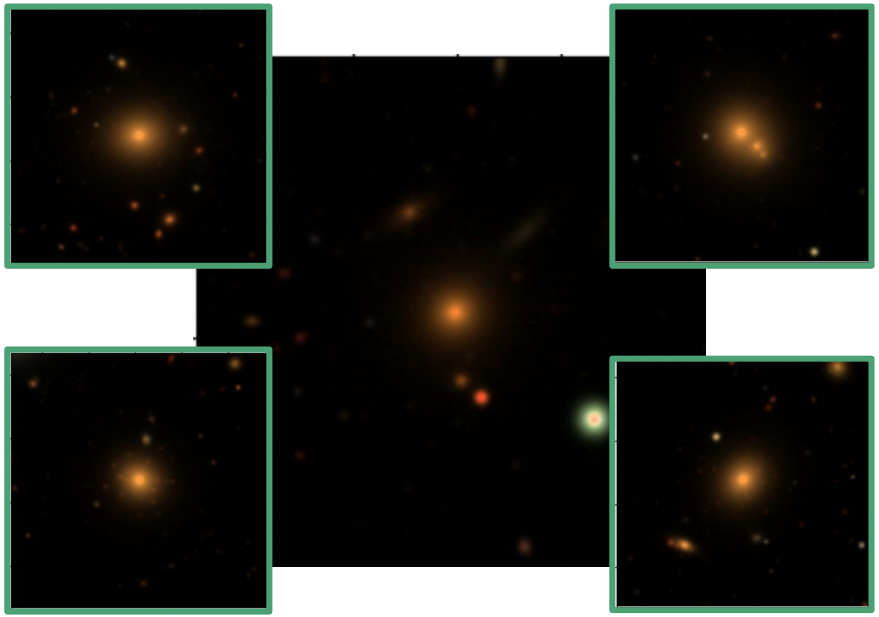
- Transient separation from host, other correlations of environment with type
- Potentially realistic template postage stamps of galaxies!
  - Real test for e.g. ALeRCE's stamp_classifier
  - You should be able to use the contextual information in v2 to improve classification accuracy over v1
- Ultimately we want to generate fully realistic DIA stamps
- Test real-bogus, subtraction pipeline, flux recovery as a function of z, host-SN contrast, magnitude
- Measure the impact of sensor anomalies, photometric calibration, PSF issues on classification with a dataset for which truth is known

DC2 Co-Add Postage stamps

- Can retrieve postage stamps from DC2 (with DC2 caveats)

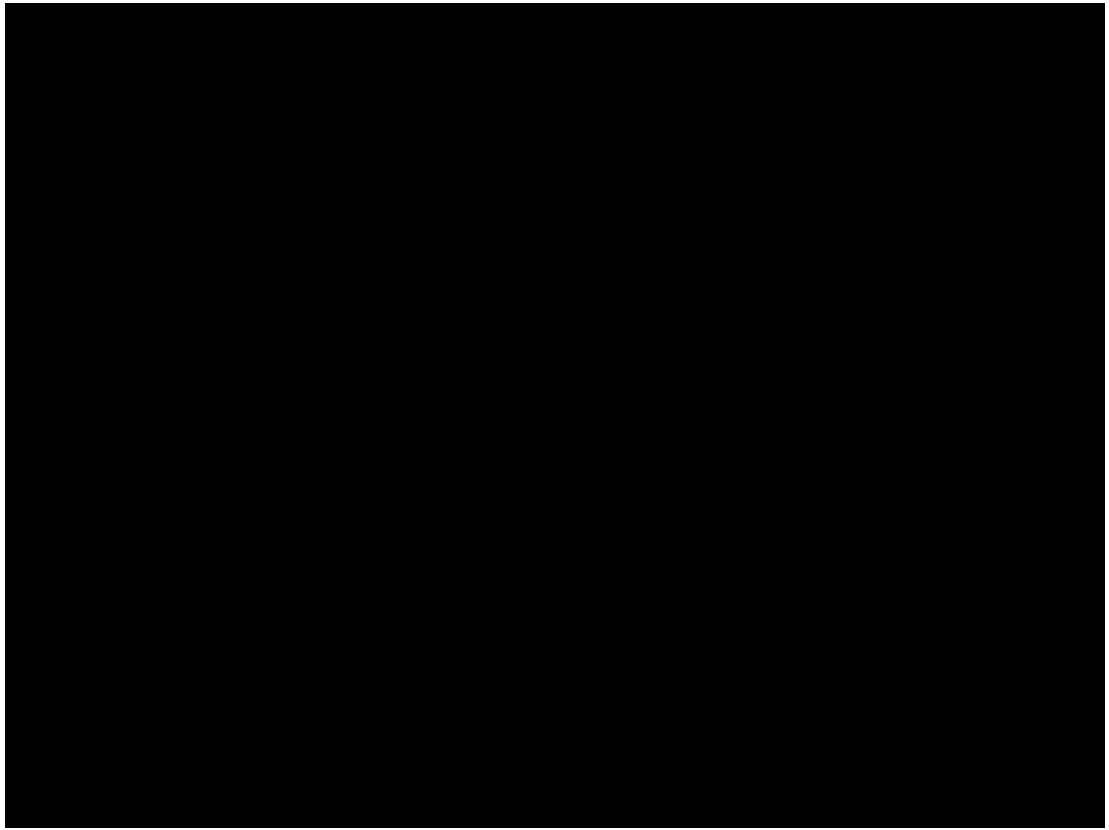


GHOST galaxy, PS1



Matched DC2 simulated galaxies

Simulating Host Postage Stamps



Based on code by Prof. Carlos Scheidegger (U. Arizona, ANTARES)

- Plans to rapidly generate host images from catalog-level properties

- *May* be included in v2 alert stream -- stay tuned

Current Status

LSST DESC
Dark Energy Science Collaboration

Running SNANA
for each transient class

Relative rates of transient classes represented in inputs

User inputs library of host & transient properties (class, SALT2 fits)

Generates SN light curves based on *z*-dependent rate models

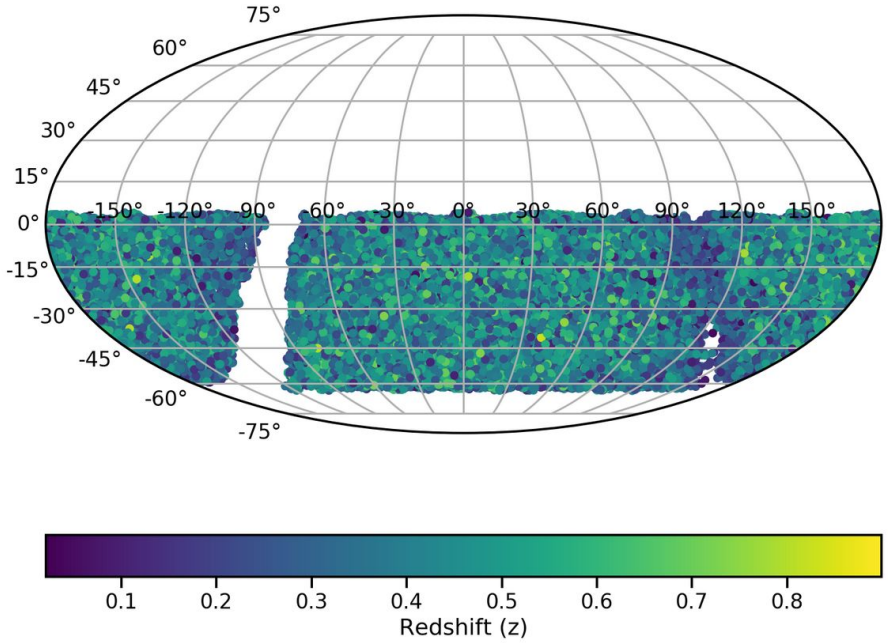
For each SN at *z*, selects host galaxy within nearby *z* window (depends on host properties)

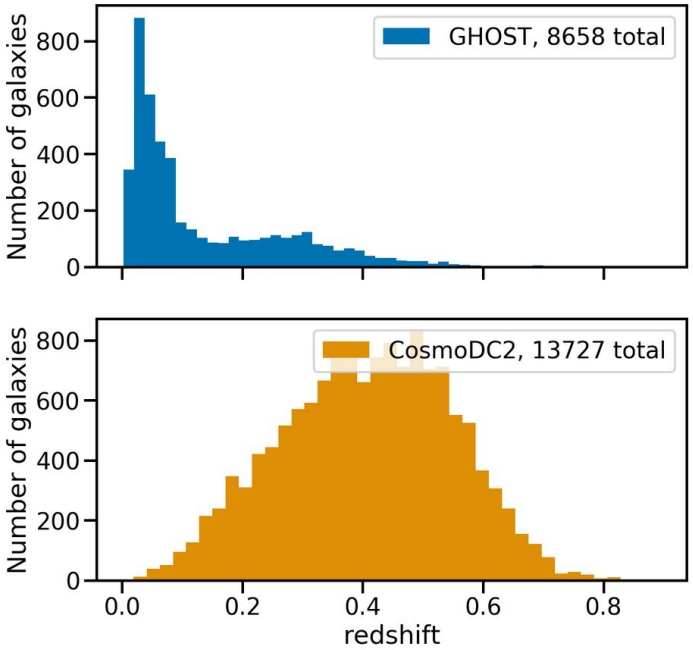
Each SN is assigned location in/near host and 'observations' are taken

SN light curves, positions, host association (and spectroscopy)

See demo from Renee/Alex this afternoon!

LSST-Simulated SNe Ia with SNANA



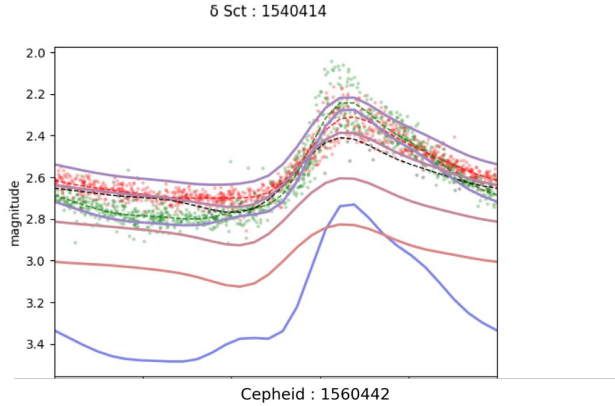


First sims run for GHOST/DC2-matched SNe Ia - more transients to come!

CAN DO THIS FOR OTHER SURVEYS!!!

What can PLAsTiCC v2 do for you

- Framework in place for generating photometry, spectroscopy for simulated transients+hosts as an alert stream **for any survey**
- Future: model host galaxy correlations for non-SN events (KNe, TDEs, etc) and generate with SNANA
- **If you are a broker team/wide-field survey and want to try to use PLAsTiCC as a training/validation set, please get in touch!**
- The super-secret goal of v2 isn't just to prepare broker teams and science collaborations for LSST alerts, but to get them to **work together**



Malanchev