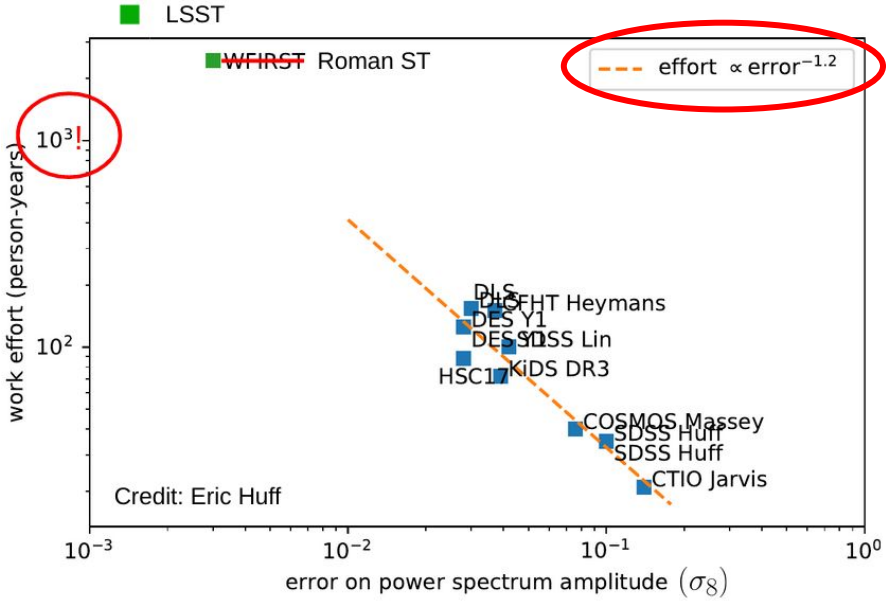
Balrog in DES Y3

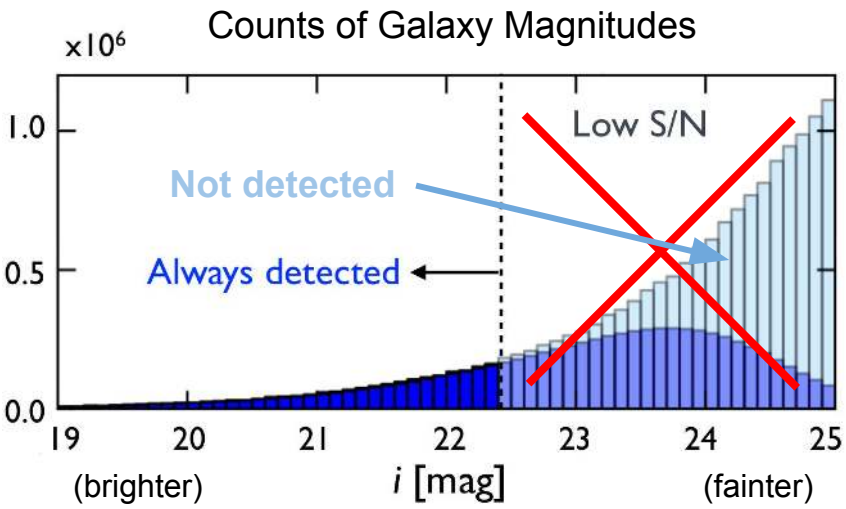
June 3, 2021

Spencer Everett presenting for lots of DES people, including:
Brian Yanny, Nikolay Kuropatkin, Eric Huff, Yuanyuan Zhang, Alex Alarcon, Sahar Allam, Alex Amon, Gary Bernstein, Katie Eckert, Jack Elvin-Poole, Daniel Gruen, Justin Myles, Judit Prat, Carles Sanchez, Erin Sheldon, and many others!

Taking a step back: ***Why*** do we need injection pipelines?



Many more ways to create a small systematic than a big one - ***power law growth in work effort!***

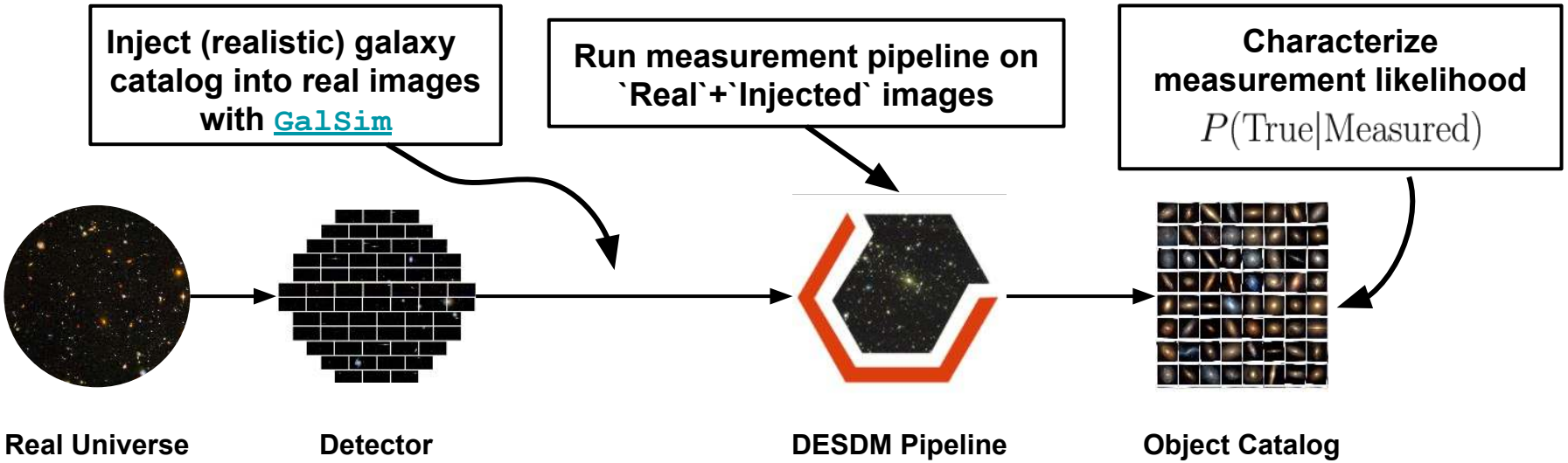


Traditionally we limit cosmology measurements to bright, highly complete subsamples - ***throwing out most of the data!***

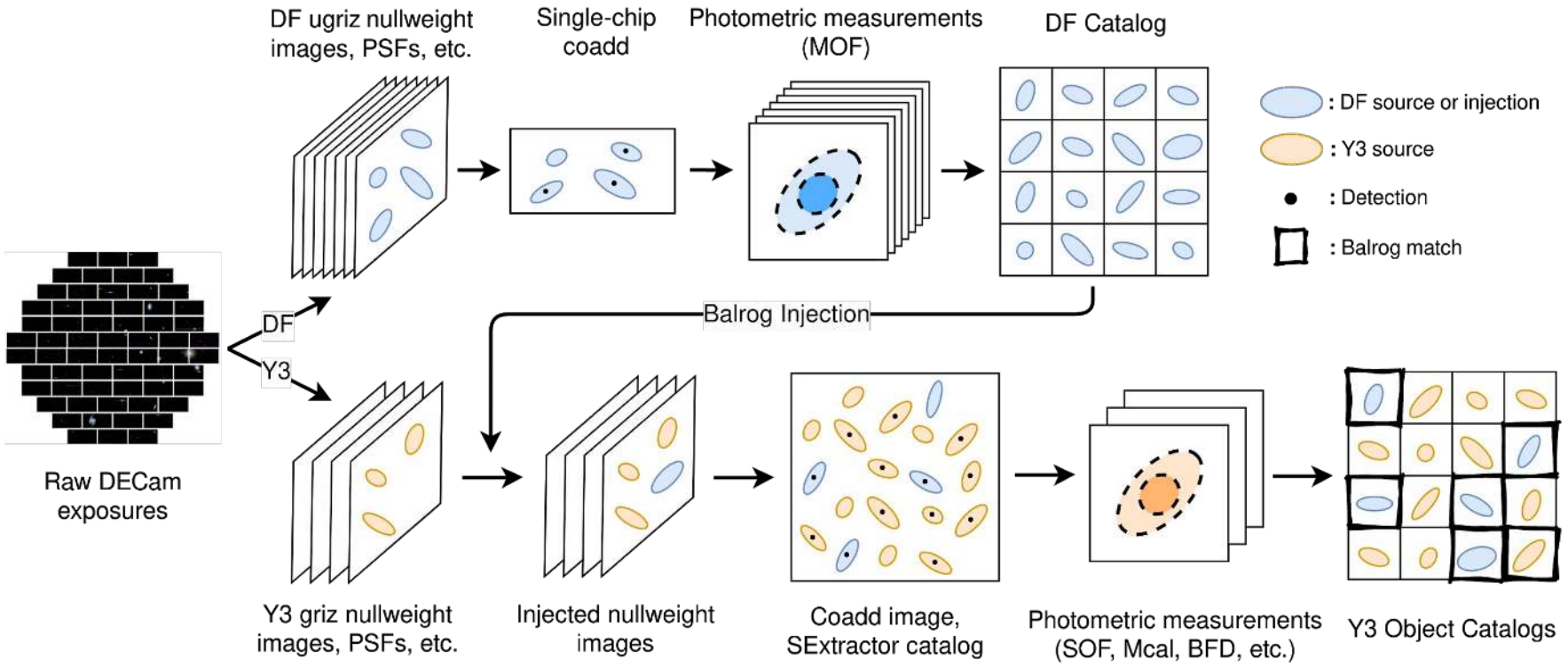
Balrog was first introduced in Suchyta et al. 2016

Characterizes the ***selection effects*** and ***measurement biases*** of the DES pipeline by injecting a realistic ensemble of fake star and galaxies into the real survey images

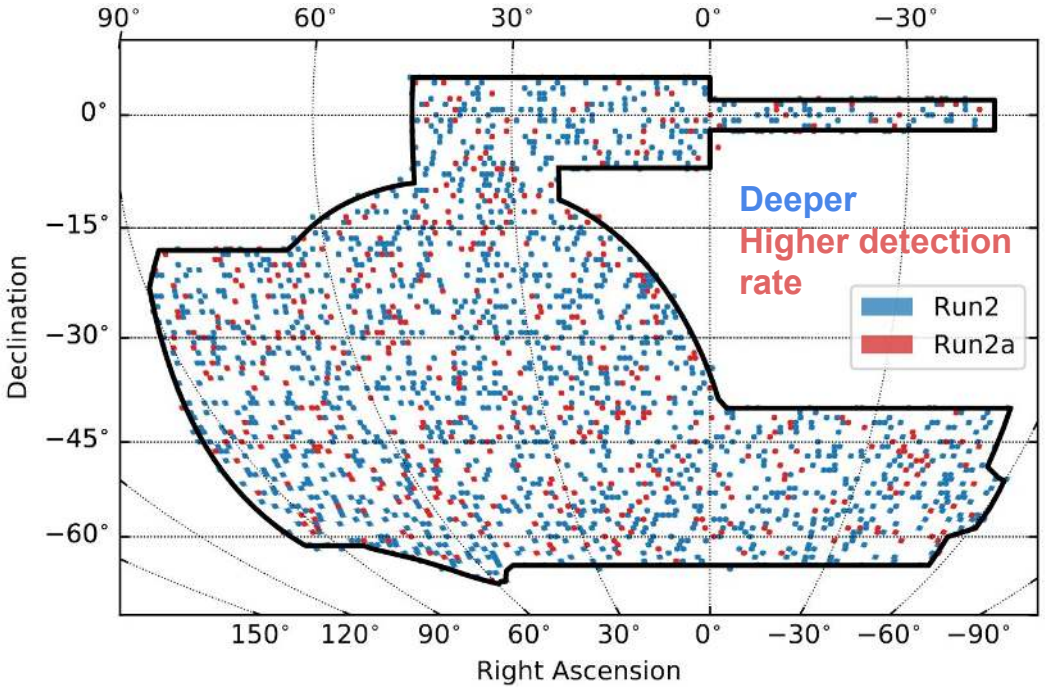
Balrog objects ***inherit difficult to model systematic effects*** that vary across the footprint



Injection Catalog Creation: Use measurements from the DES Deep Fields (DF) (Hartley & Choi et al.)



Injection: New Balrog framework that handles single-epoch processing for DES data products (Everett et al.)



~11.3 million detected from **26.4 million total injections** from fits to sources in the Deep Fields

2,000 randomly selected tiles (**~20% of full footprint**)

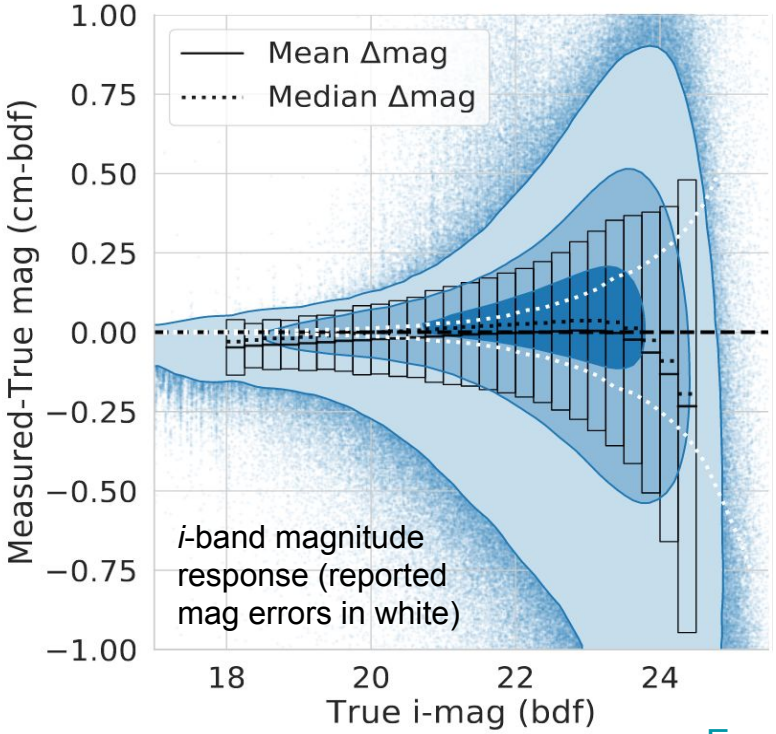
Injections on a ***hexagonal lattice*** of DF objects (80%) and simulated stars (20%)

Injected objects are reddened to account for local galactic extinction

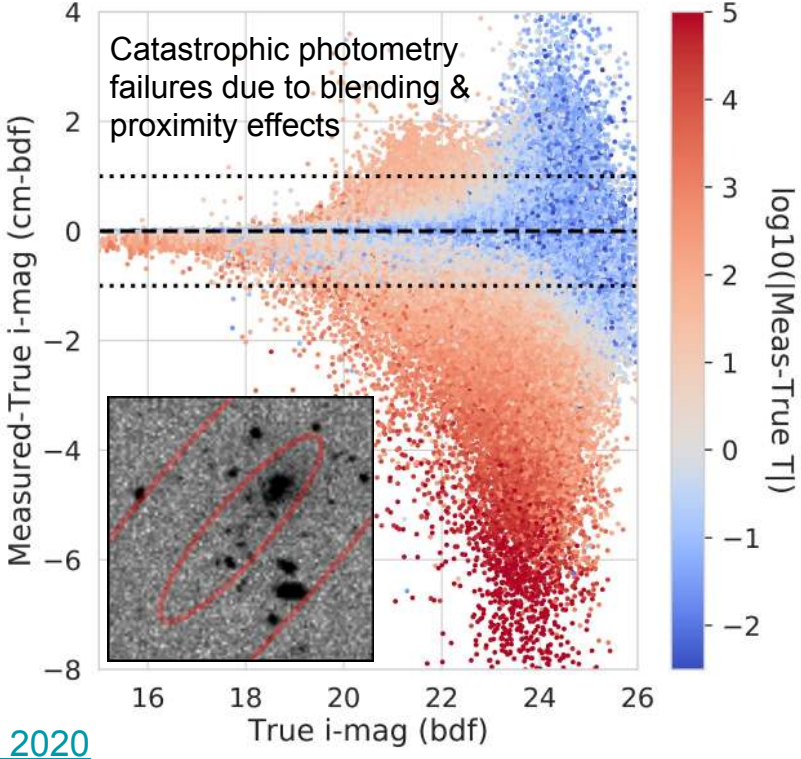
~80s CPU time per recovered Balrog injection; ***12 times more expensive than Y1!***

What can we use injection pipelines for?

Photometric Calibration & Characterization



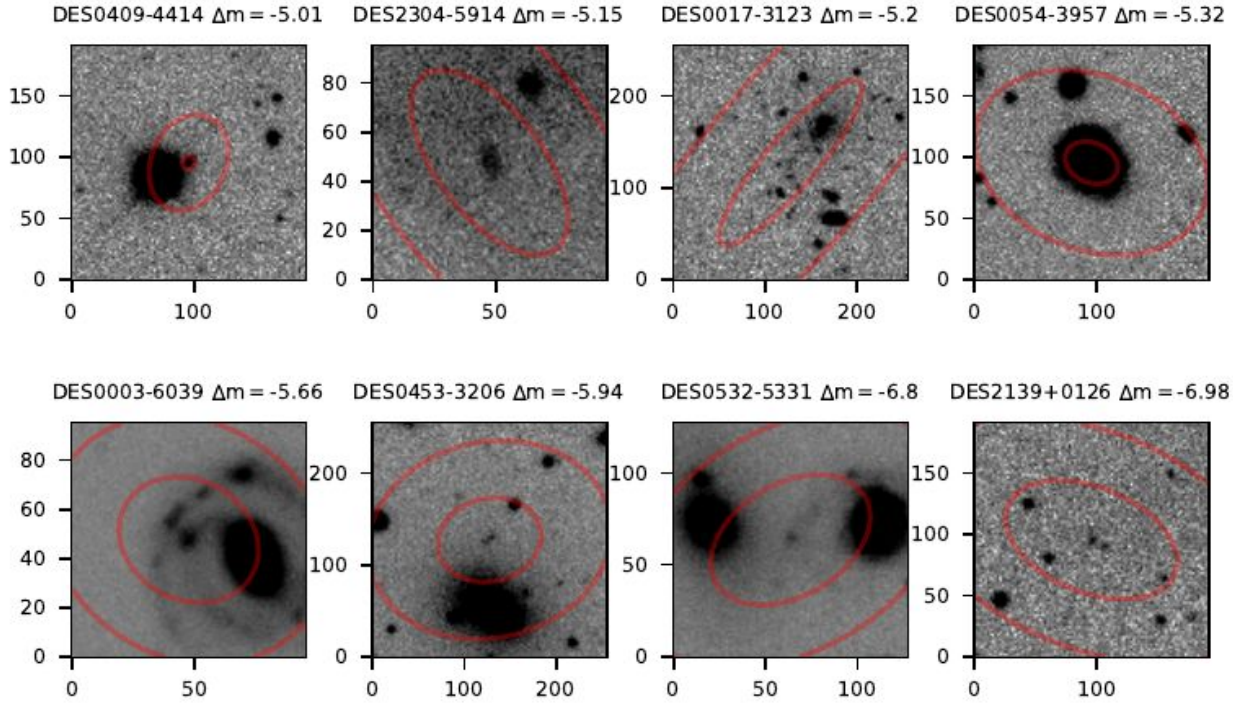
Diagnostics, Outlier Discovery, & Pipeline Feedback



Everett et al. 2020

Catastrophic photometry failures proportionally affect bright objects more, but are far more extreme for fainter (and thus usually smaller) galaxies; it is ***heavily dependent on the local conditions***

50th and 95th percentile ***measured*** flux contours



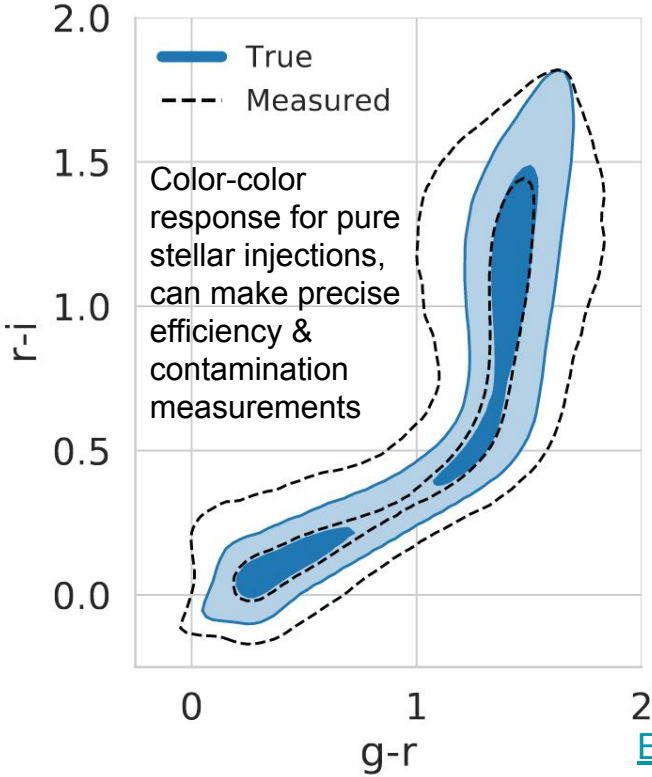
"Super-Spreaders"

Extreme biases (***up to factors of ~2,000 in flux!***) can occur for:

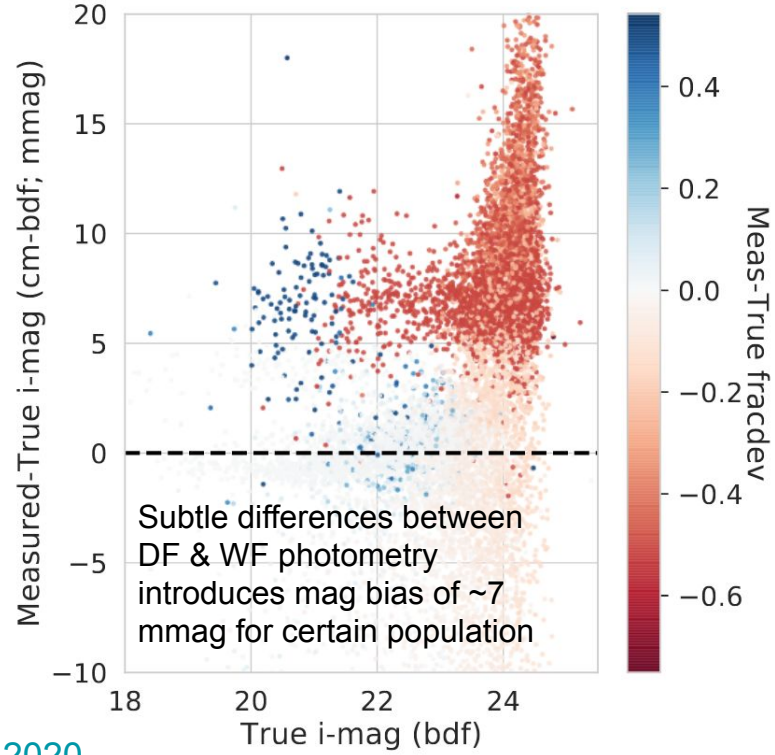
- ***Dense fields*** like galaxy clusters
- Regions with ***image artifacts*** such as scattered light
- Injections ***near very large or bright sources*** such as saturated stars

What can we use injection pipelines for?

Photometric Calibration & Characterization



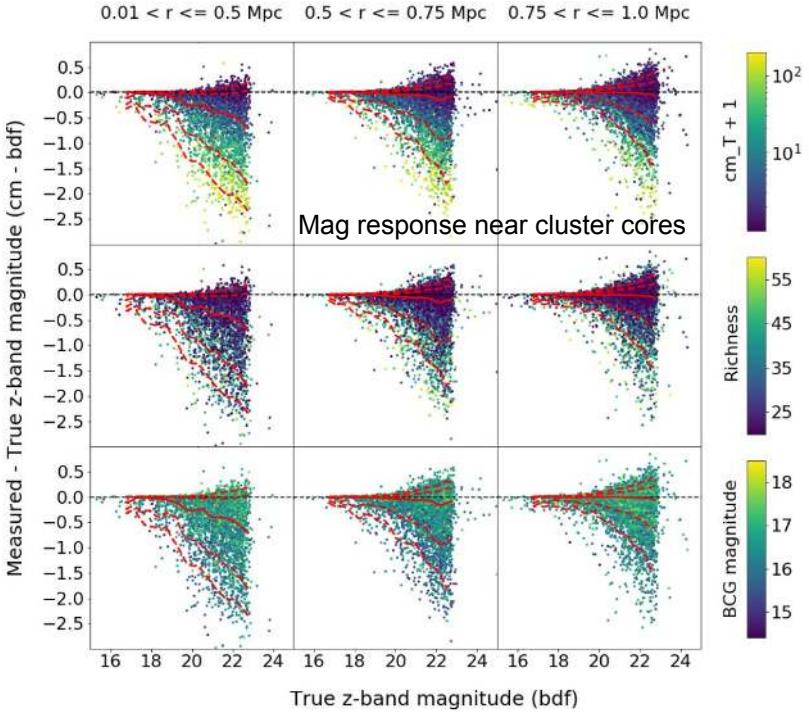
Diagnostics, Outlier Discovery, & Pipeline Feedback



Everett et al. 2020

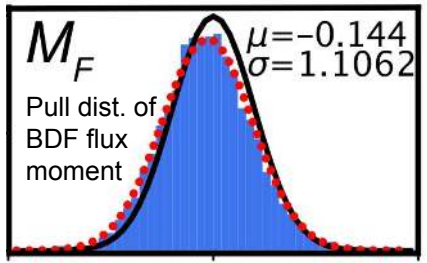
What can we use injection pipelines for?

Photometric Calibration & Characterization

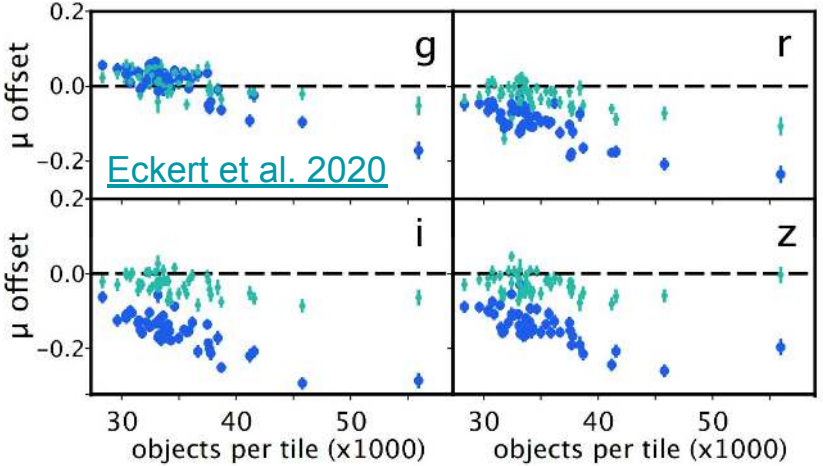


Everett et al. 2020

Diagnostics, Outlier Discovery, & Pipeline Feedback



Discovered noise from undetected sources in DES calibrated images; up to 30% underestimate of pixel variance. Also found evidence of sky over-subtraction in *riz*



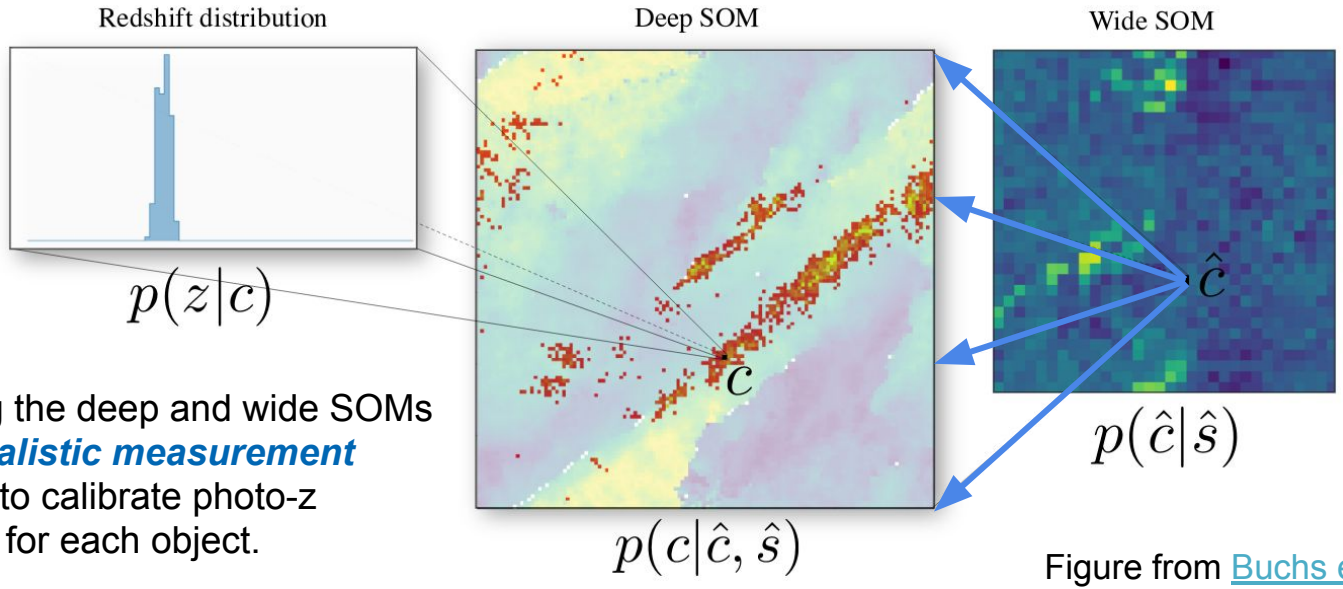
Eckert et al. 2020

Main science driver for Y3 Balrog: **Photometric Redshifts**

New approach uses self-organizing maps (SOM) to link high-dimensional COSMOS Deep Field colors into galaxy *phenotypes* mapped to an equivalent *riz* Wide Field SOM by Balrog:

Myles & Alarcon et al. 2020

$$p(z|\hat{c},\hat{s}) \approx \sum_c \underbrace{p(z|c)}_{\text{Redshift}} \underbrace{p(c)}_{\text{Deep}} \underbrace{\frac{p(c,\hat{c})}{p(c)p(\hat{c})}}_{\texttt{Balrog}} \underbrace{p(\hat{c})}_{\text{Wide}},$$



Connecting the deep and wide SOMs requires ***realistic measurement likelihood*** to calibrate photo-z distribution for each object.

Figure from Buchs et al. 2019

Secondary science driver for Y3 Balrog: **Lens Magnification**

Magnification is correlated large-scale structure and should be taken into account in the modeling of the two-point correlation functions

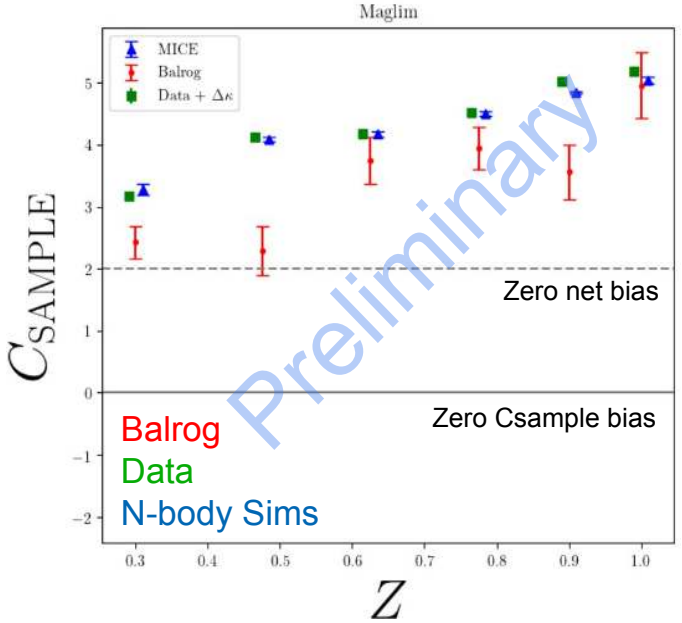
$$\delta_g^{\rm obs} = \delta_g^{\rm int} + \delta_g^{\rm mag} = \delta^{\rm int} + [C_{\rm area} + C_{\rm sample}] \cdot \delta\kappa$$

There are competing effects; a ***geometric suppression*** factor ($C_{\rm area} = -2$) and an ***boost in detection of faint sources*** which increases the local number density:

$$C_{\rm sample}\delta\kappa \approx \frac{n_{\rm int}(F, \kappa = \delta\kappa)}{n_{\rm int}(F, \kappa = 0)}$$

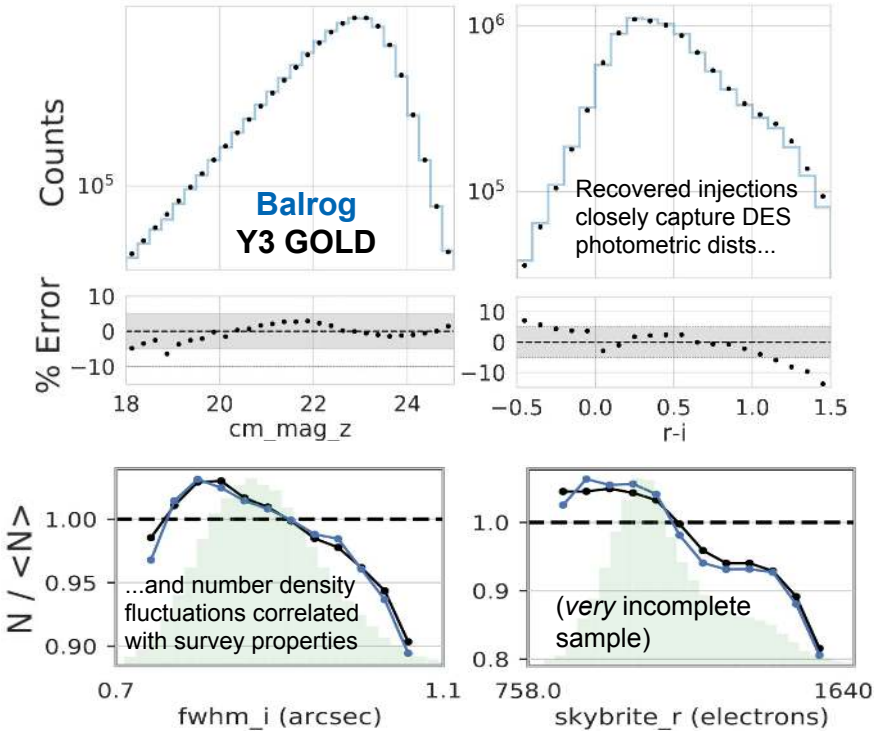
Balrog can measure the change in density of lens (or other!) galaxies as a function of observing systematics + magnification with slightly modified (& repeated) runs

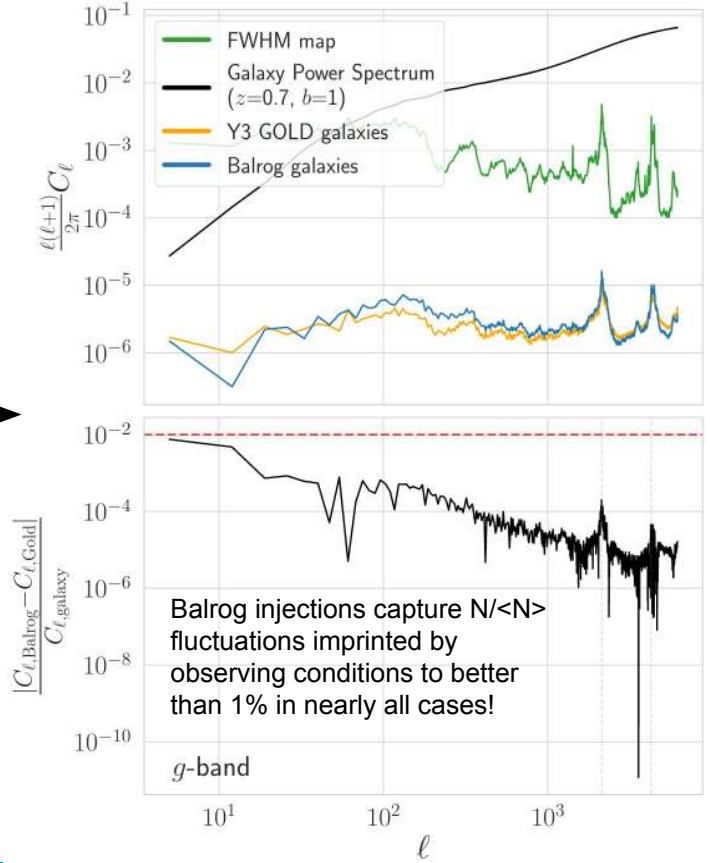
Full measurement of $C_{\rm sample}$ for both DES clustering samples upcoming in Elvin-Poole et al. (in prep)



(If time) What can we use injection pipelines for?

Systematics Tests & Sample Randoms





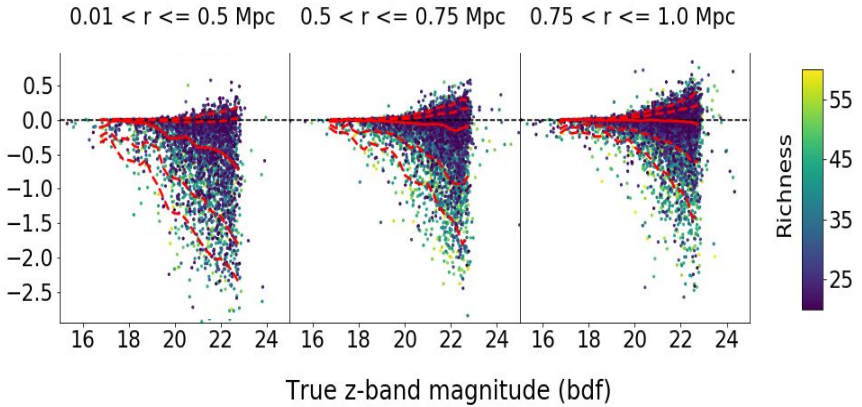
Everett et al. 2020

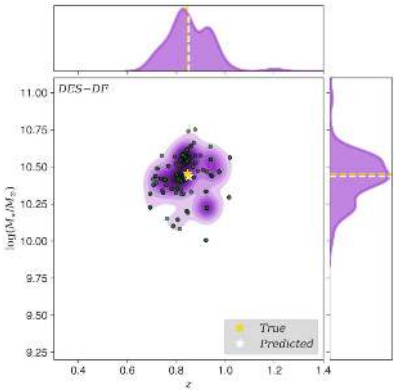
Far more diagnostic and calibration potential! Just a few additional applications that are either published or ongoing:

Investigating ***photometric response near cluster cores*** to constrain dark matter properties and improve cluster cosmology calibration:

Everett et al. (2020)

Masegian & Zhang et al. (in prep.)



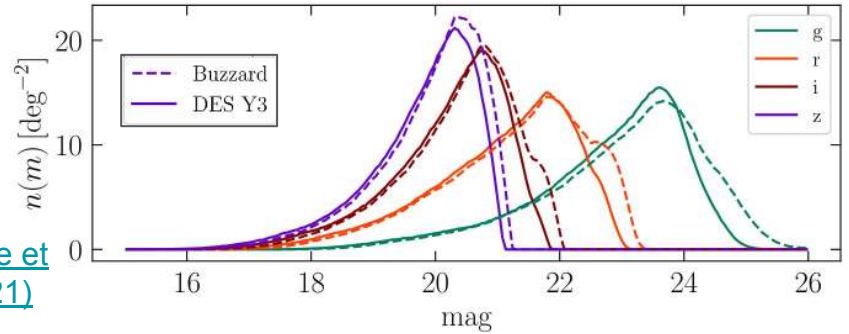


Training ML methods for accurate joint redshift-stellar mass probability distributions

Mucesh et al. (2021)

Accurate ***error estimation in end-to-end cosmology sims***

DeRose et al. (2021)



This is just the beginning of what we can do to fully incorporate object injection pipelines into our pipeline diagnostics & cosmological analyses!

The DES Y3 Balrog paper submitted to arXiv in December - check it out!

Full DES Y3 results & papers are being compiled here

Extremely computationally expensive to make large samples (rerun full measurement pipeline multiple times), so working on an ***emulation approach*** that uses the full runs as a training set. Also working to incorporate postage-stamp injections

Lots of lessons learned on how to run this kind of analysis at survey-scale. Many of these will be needed for effective object injection in LSST where the computational feasibility issues are ***much*** more difficult.

Message me anytime on slack or email (sweveret@ucsc..edu - switching to JPL email soon)