Difference Image Analysis in LSST stack

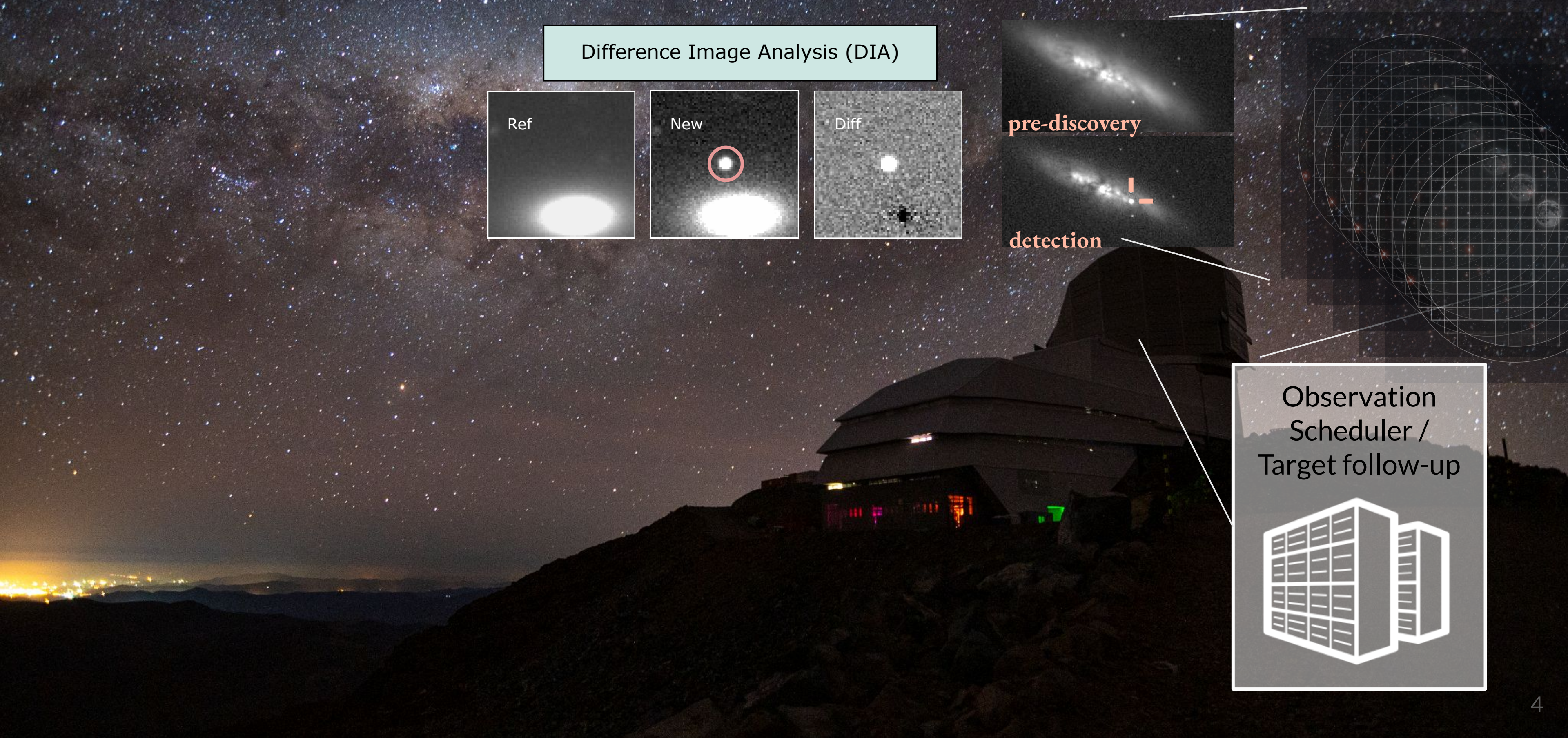
Alert Production testing & perspectives

Bruno S., Ian S. & AP Team

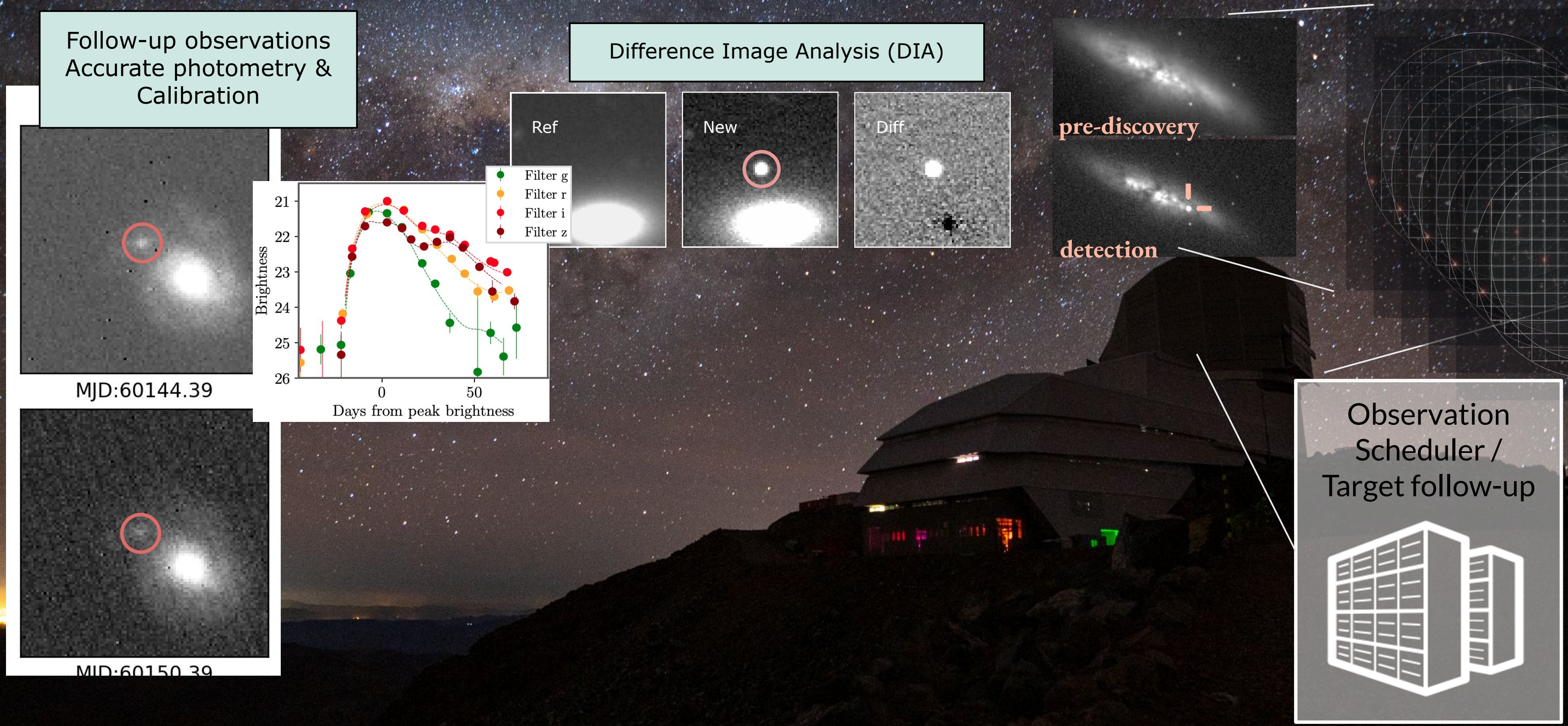
Time domain Cosmology: from pixel to Constraints

Time domain Cosmology: from pixel to Constraints

Time domain Cosmology: from pixel to Constraints

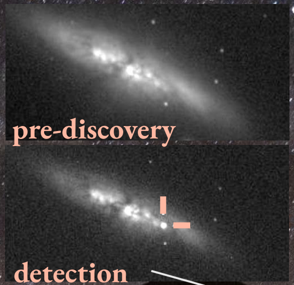


Time domain Cosmology: from pixel to Constraints



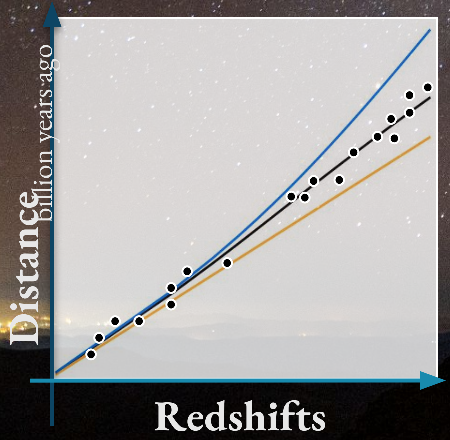
Time domain Cosmology: from pixel to Constraints

Difference Image Analysis (DIA)



Follow-up observations
Accurate photometry &
Calibration

Light-curve
standardization
Cosmology data vector

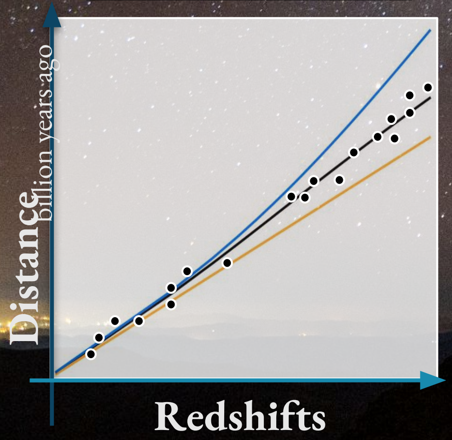


Observation
Scheduler /
Target follow-up

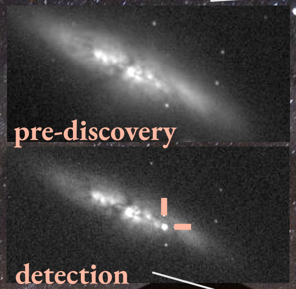
Time domain Cosmology: from pixel to Constraints

Follow-up observations
Accurate photometry &
Calibration

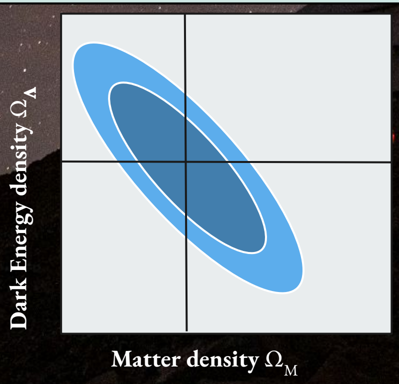
Light-curve
standardization
Cosmology data vector



Difference Image Analysis (DIA)



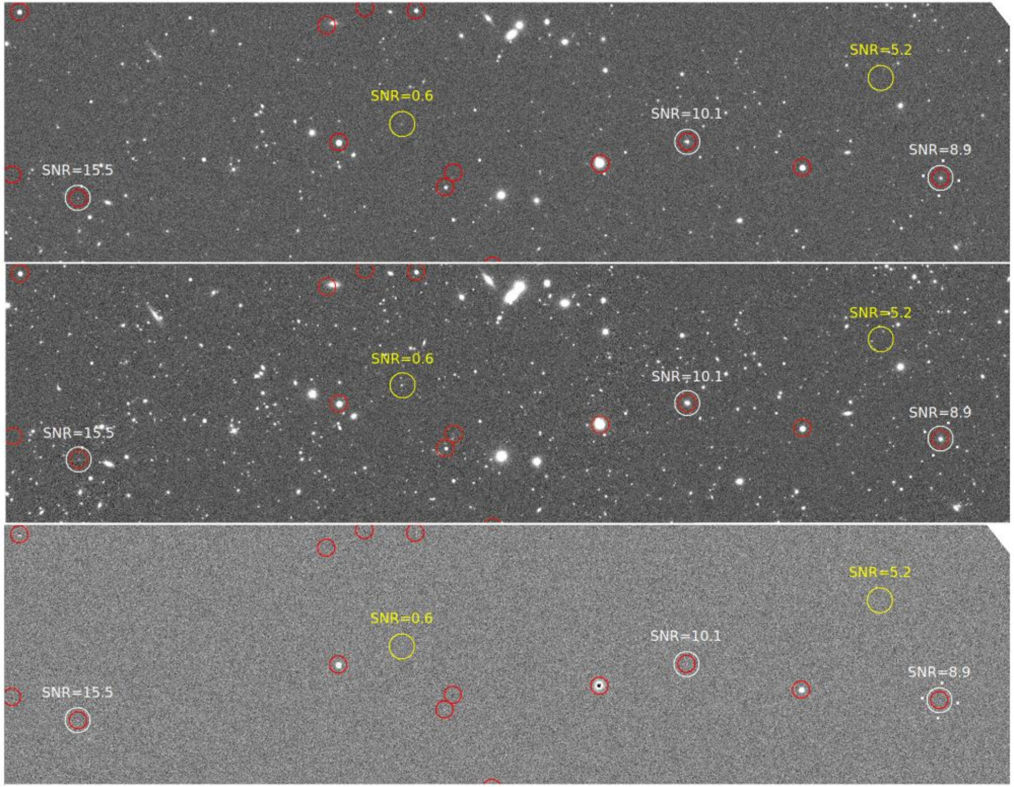
Cosmological Parameter
inference



Observation
Scheduler /
Target follow-up

Difference Image Analysis

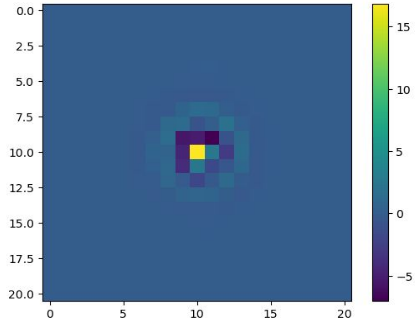
- DIA is the workhorse behind a large portion of Vera Rubin LSST Time Domain science
- DIA is the cornerstone methodology for many science cases
- LSST Data volume will make tiny or rare effects (corner cases) a massive daily problem to solve
- There is new and exciting science where the limit cases take place
  - Very low SNR transients
  - Very noisy environments
  - …

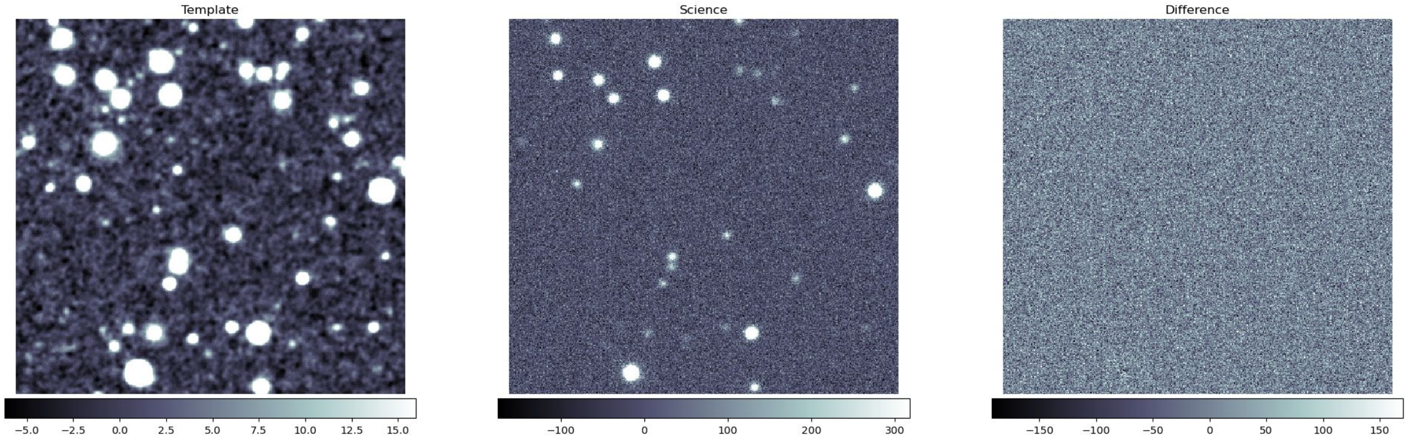


lp_diffim A&L Subtraction tasks

Difference Image Analysis consists on pixel-to-pixel comparisons.

- Nominal Alard & Lupton 98/2000 DIA algorithm
  $D = N - k * R$





lp_diffim A&L Subtraction tasks

Difference Image Analysis consists on pixel-to-pixel comparisons.

- Nominal Alard & Lupton DIA algorithm
  $D = N - k * R$

- Auto-convolution mode: perform a re-assignment of the image to transform
  $D = k * N - R$ `if` $(\mathrm{PSF}_N < \mathrm{PSF}_R)$ `else` $D = N - k * R$

lp_diffim A&L Subtraction tasks

Difference Image Analysis consists on pixel-to-pixel comparisons.

- Nominal Alard & Lupton DIA algorithm
  $D = N - k * R$

- Auto-convolution mode: perform a re-assignment of the image to transform
  $D = k * N - R$ `if` $(\mathrm{PSF}_N < \mathrm{PSF}_R)$ `else` $D = N - k * R$

- Pre-Convolution mode: convolution of N with a given known kernel v:
  $\mathrm{PSF}_{N'} > \mathrm{PSF}_R$; as result k is no longer a de-convolution kernel

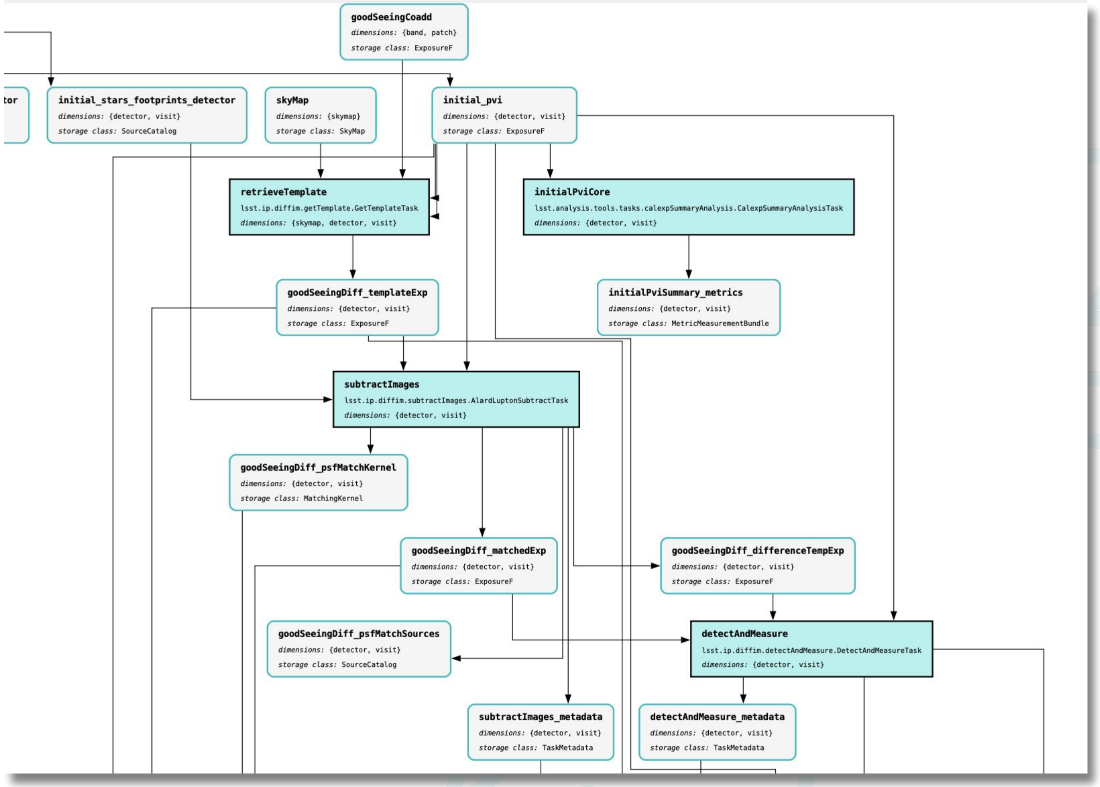
  $v * D = v * N - k * R$

Difference Image Analysis

In project we work with **pipelines**

This is best represented as a **Direct Acyclic Graph (DAG)** that represents the *flow of data through inter-dependent tasks with individual responsibilities*.

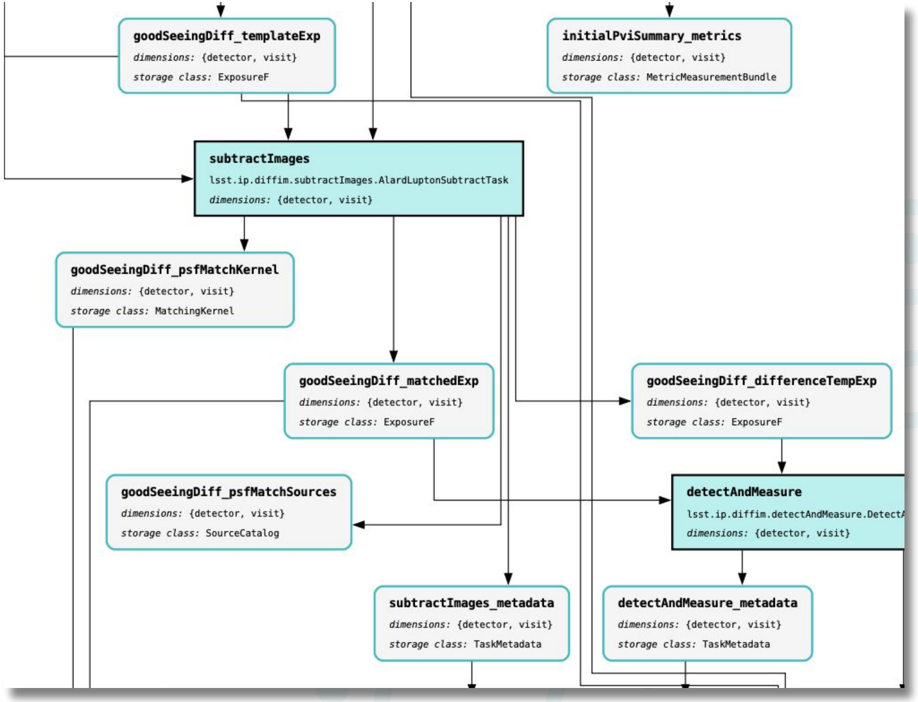
In the case of Alert Production this is the portion that includes the key ***subtractImages task***



The SubtractImages task

The outcome of this task is:

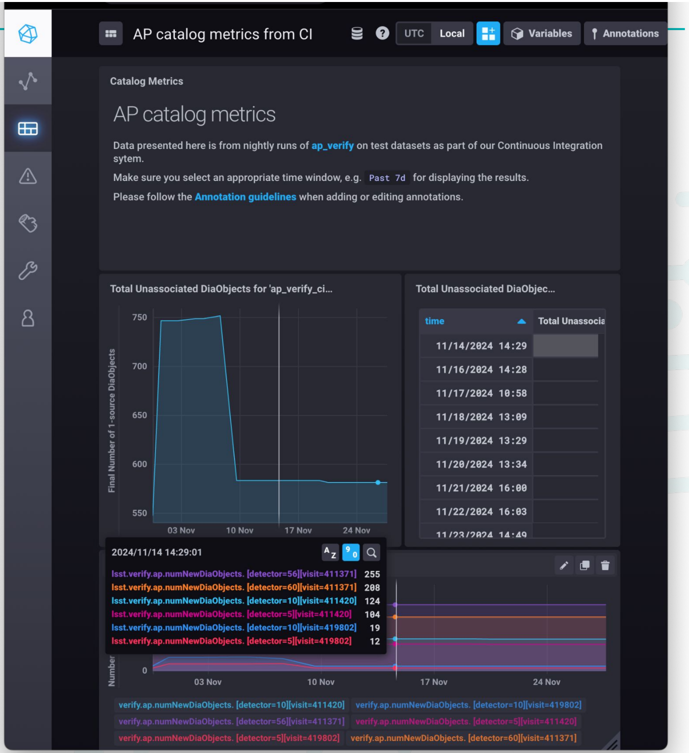
- The image difference: `differenceExp`
- The matching kernel: `psfMatchKernel`
- The template transformed to match the science exposure: `matchedExp`
- Relevant (and valuable) information: `metadata`
- The list of used sources: `psfMatchSources`



Testing DIA in the stack

For testing DIA we use several levels and scales:

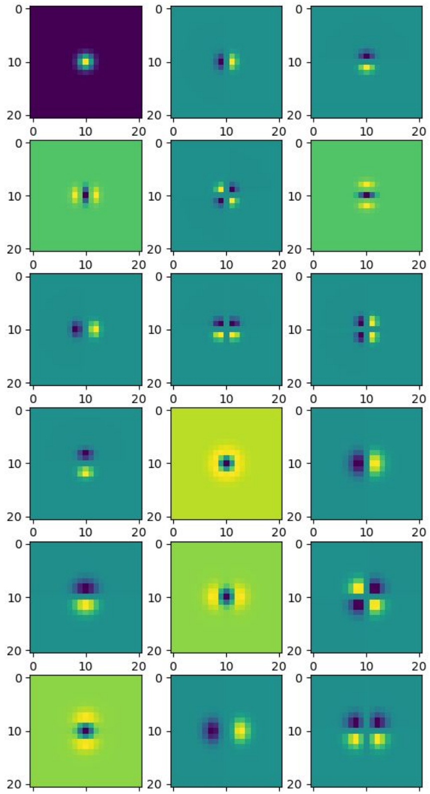
- Single unit tests that verify code integrity
- Continuous Integration tests (CI) that verify pipeline interlocking blocks
- CI with datasets (we call them “verify”) that report metrics (Chronograph)
    - HSC images
    - DECam images
    - DC2 simulation images
- There are also, development testing that we run.
- Latest test benchmarks come from Operations Rehearsal 4
    - Set of simulated images in several fields that were used for real data transfer and operations tests



Inspection of the `psfMatchKernel`

The kernel is constructed using Gaussian basis functions, modulated by low-order polynomials.
Gauss-Hermite basis functions.

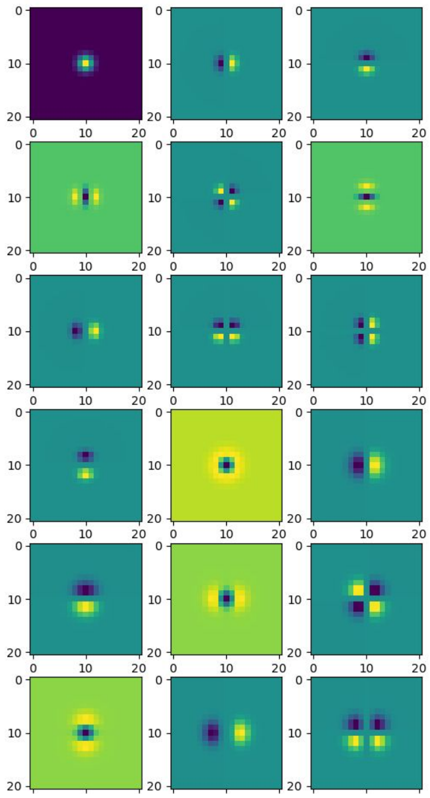
Some of the fitted components look like this:

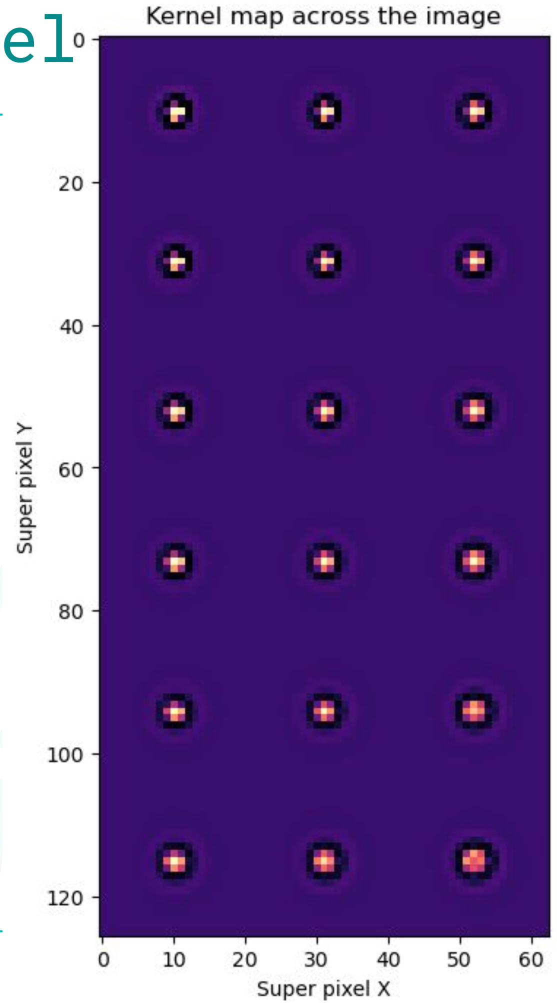


Inspection of the `psfMatchKernel`

The kernel is constructed using Gaussian basis functions, modulated by low-order polynomials.
Gauss-Hermite basis functions.

Some of the fitted components look like this:

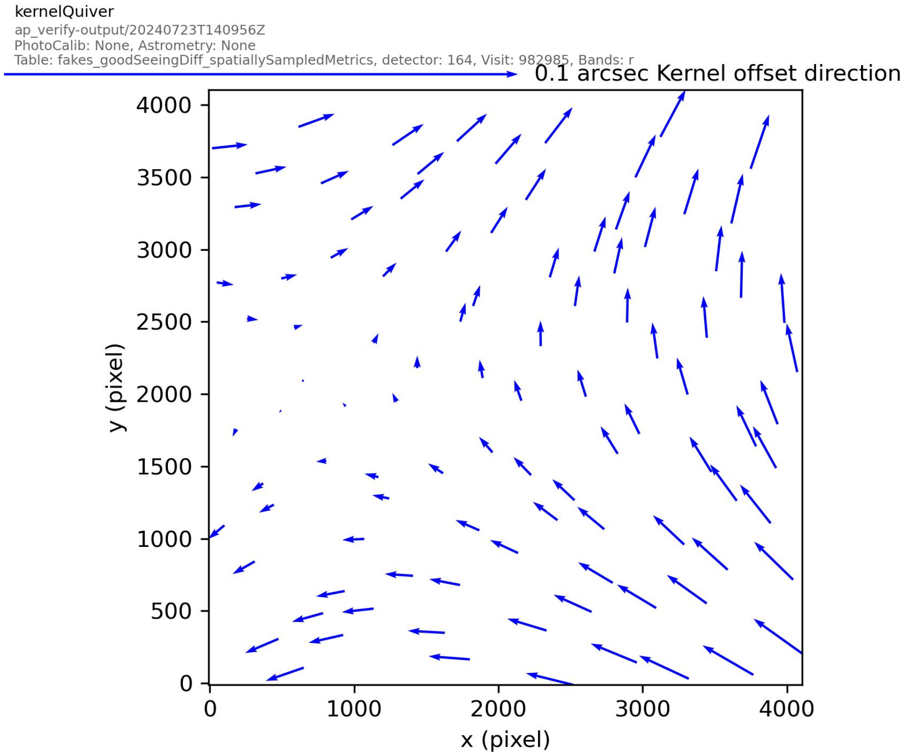




Inspection of the `psfMatchKernel`

The kernel contains information about the transformation needed to match both images.

For instance looking at its centroid we can generate maps of the smallest astrometric shift

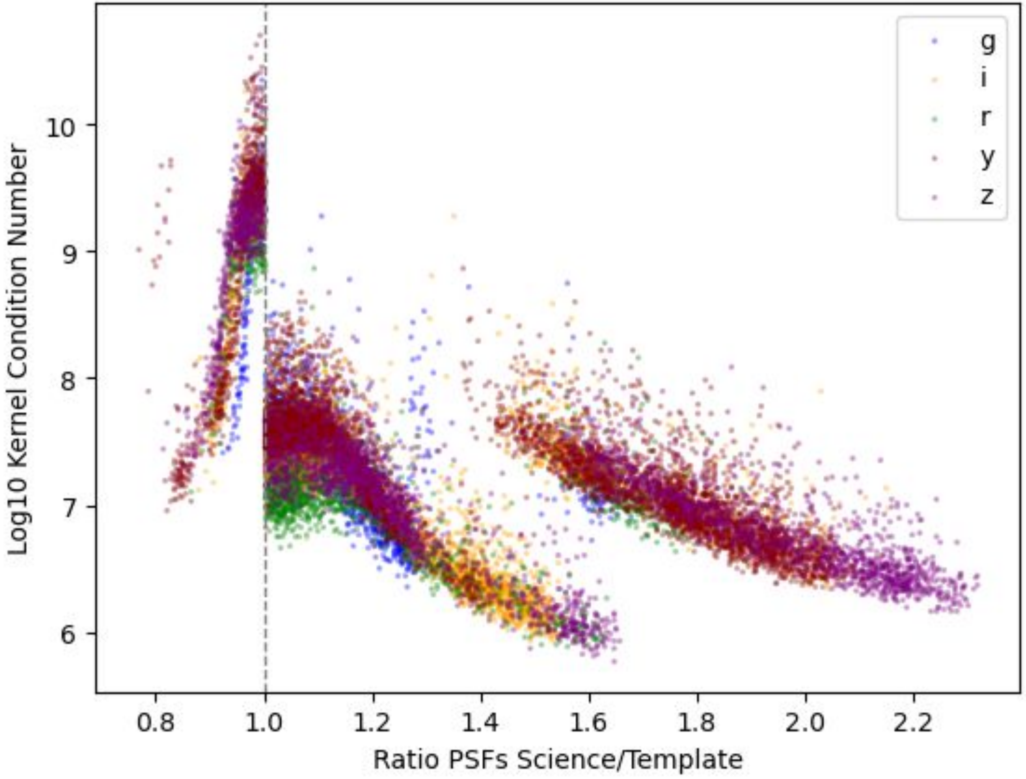


Inspection of the `metadata`

Metrics or metadata on the kernel is crucial for understanding our performance!

We inspect the dependency of two key parameters with the ratio of PSF sizes:

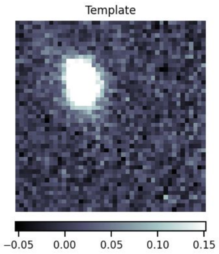
- ratioPsfs = $PSF_{Science} / PSF_{Template}$
- Condition number: the power ratio of kernel top coefficients
- SigGauss: the smallest size of the kernel basis components

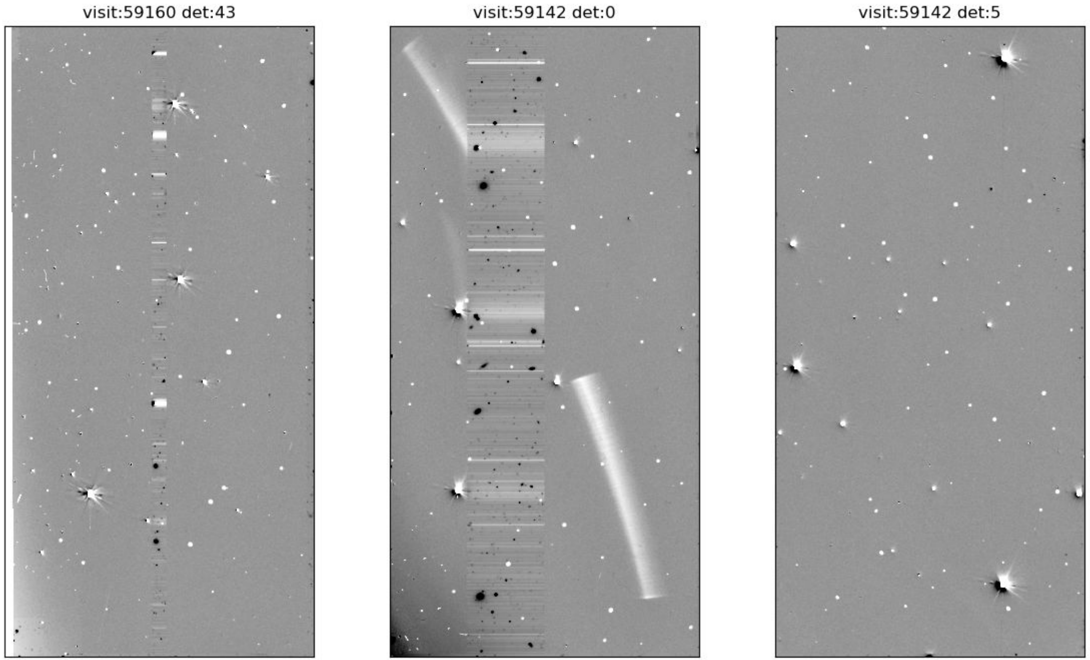


Inspecting the `differenceExp`

We check for visual artifacts, and also **inject synthetic sources**, to test recovery metrics for example

CI dataset from HSC ap_verify_cosmos_pdr2

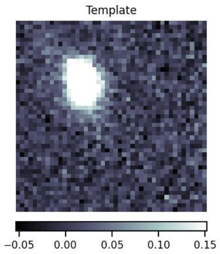


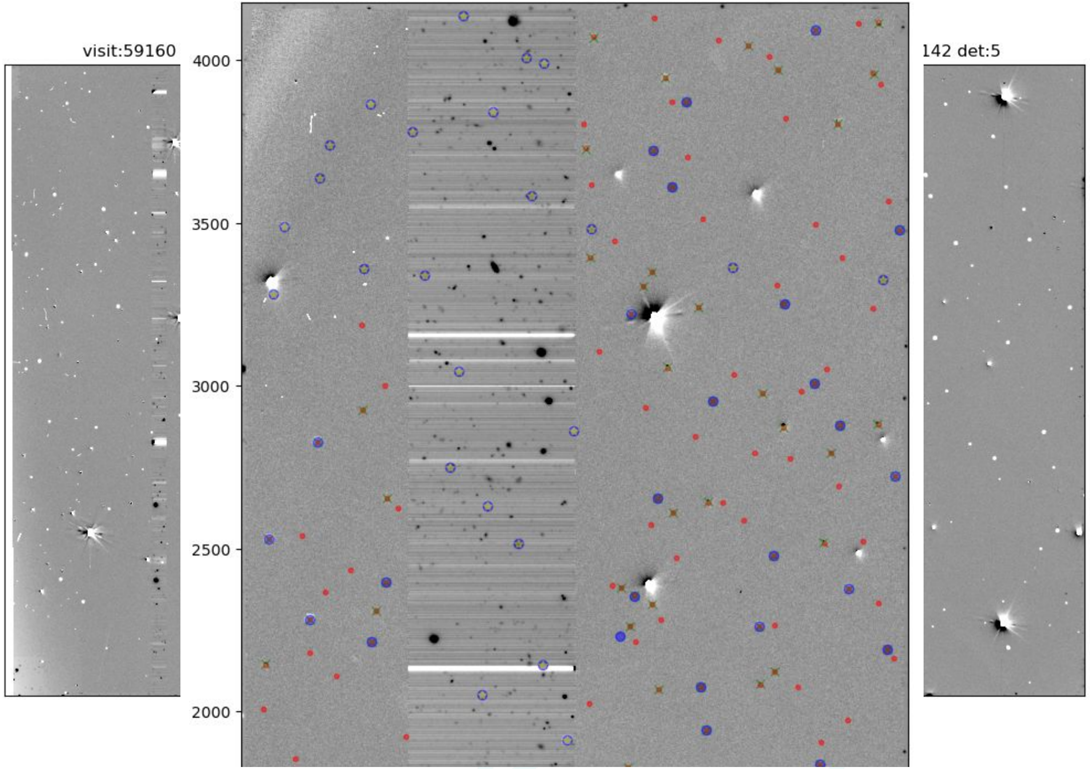


Inspecting the `differenceExp`

We check for visual artifacts, and also **inject synthetic sources**, to test recovery metrics for example

CI dataset from HSC
ap_verify_cosmos_pdr2

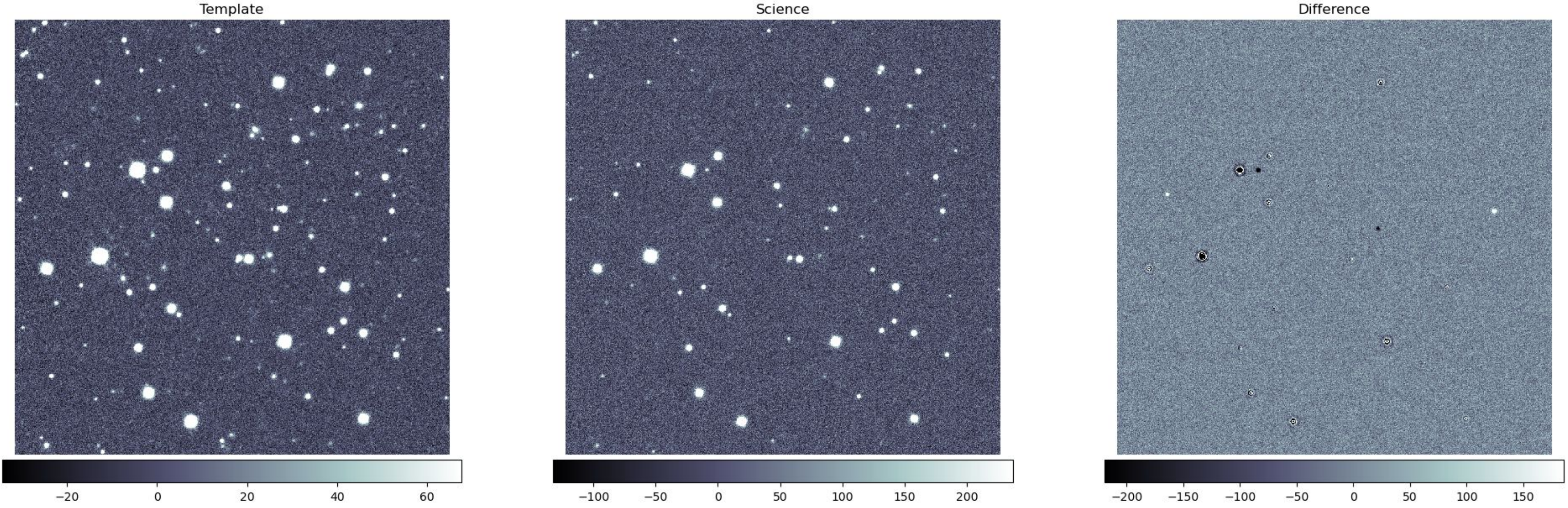




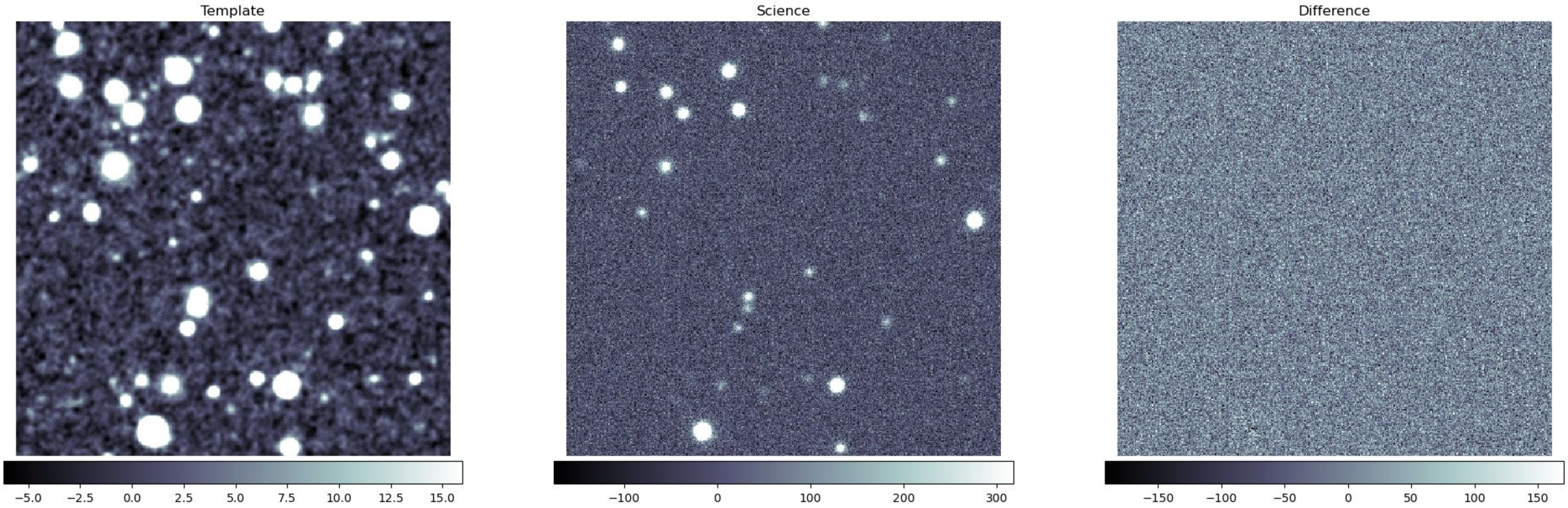
Inspecting the `differenceExp`

- Artifact zoo!
  - Deconvolution (Template PSF is worse than Science PSF)
  - Dipoles (Astrometric registering solution is imperfect, Differential Chromatic Refraction effects)
  - Saturation effects
  - Spikes! LSST will have diffraction spikes
  - Background subtraction effects
  - Correlated noise (wave ripples and granularity effects)
  - Trailed sources (not a DIA problem per se!)
  - …

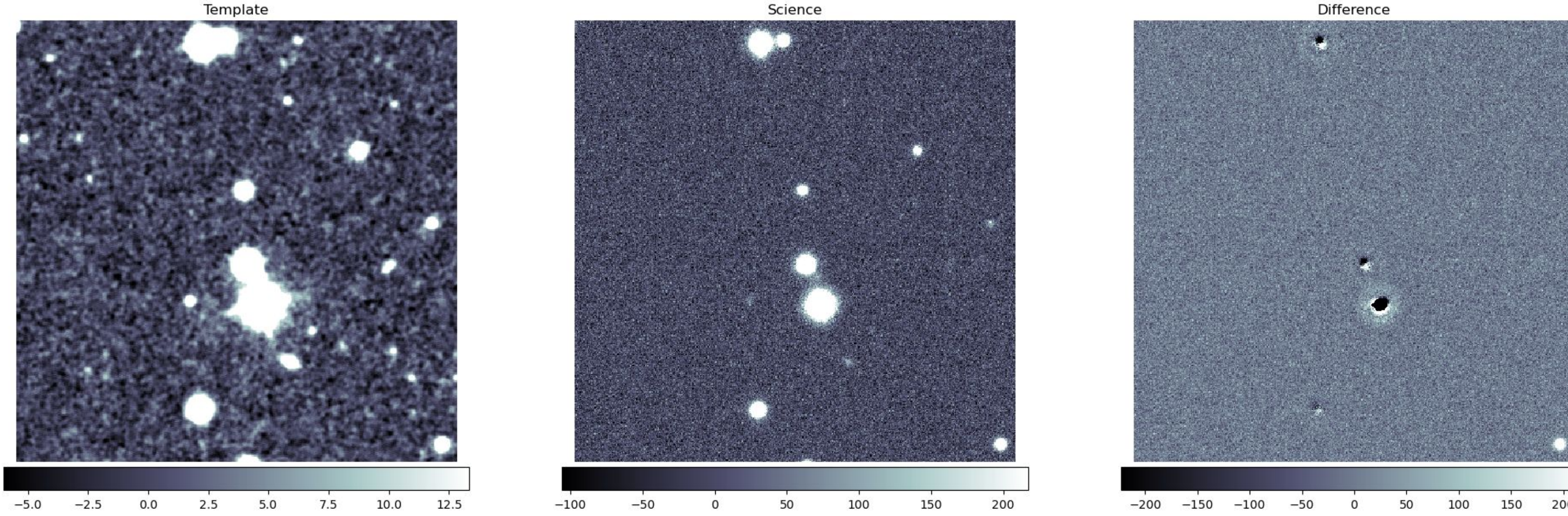
OR4 Dataset inspection



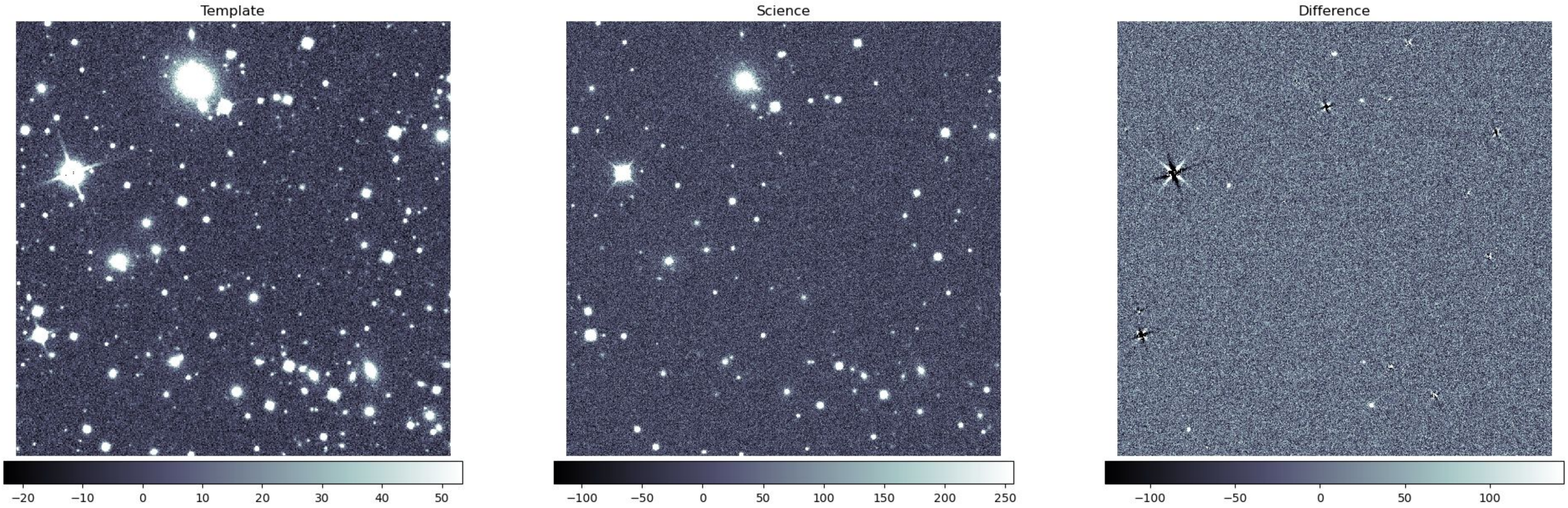
OR4 Dataset inspection



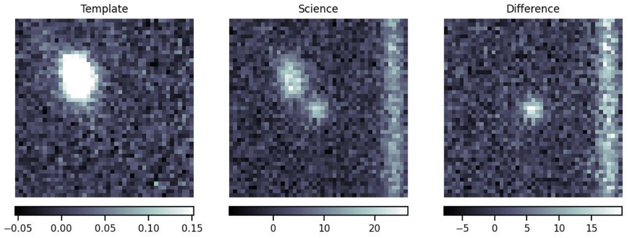
OR4 Dataset inspection

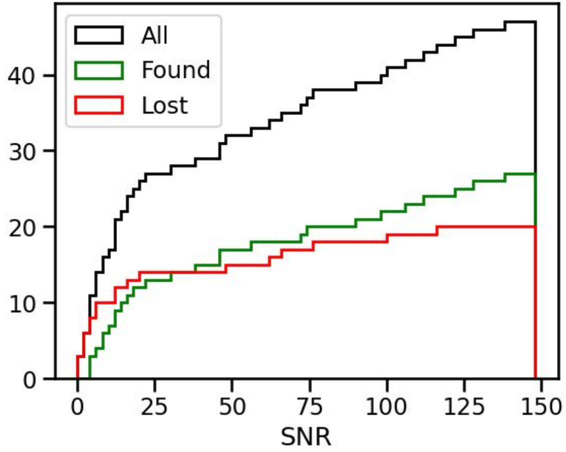


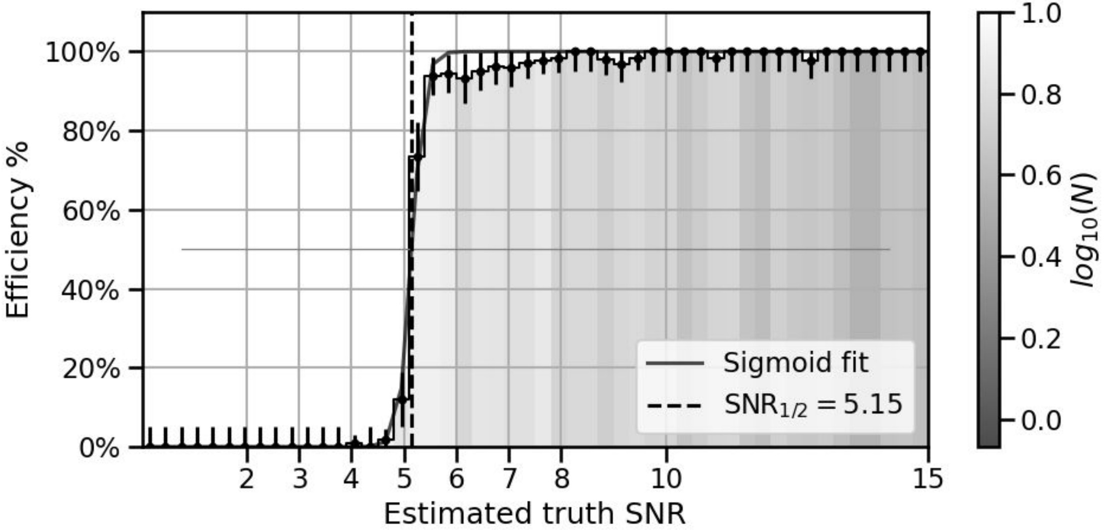
OR4 Dataset inspection



Synthetic sources: HSC example



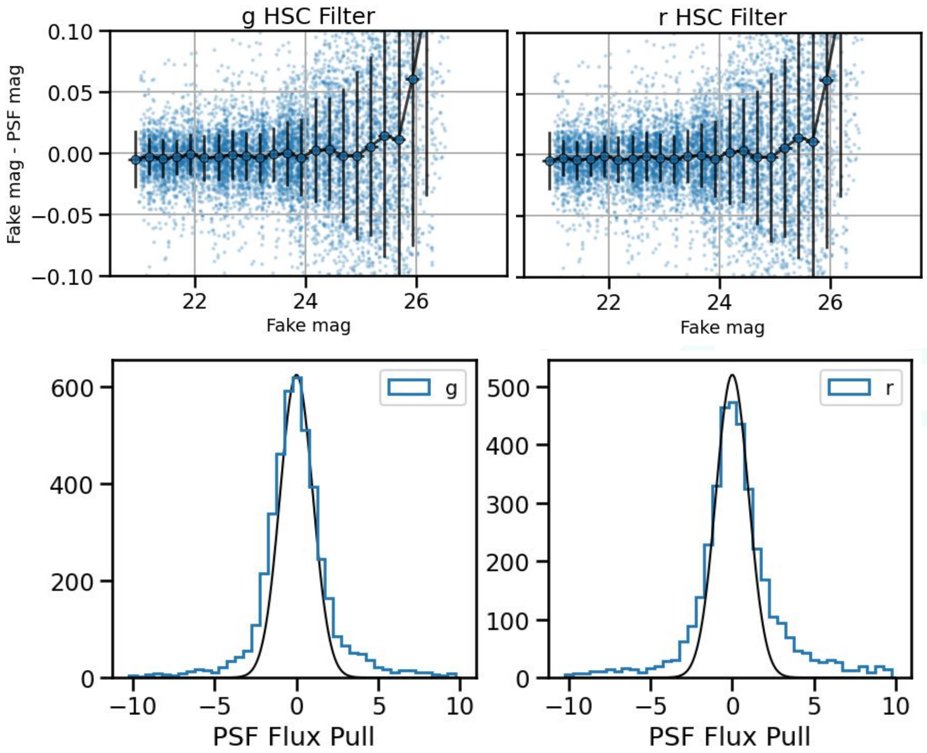




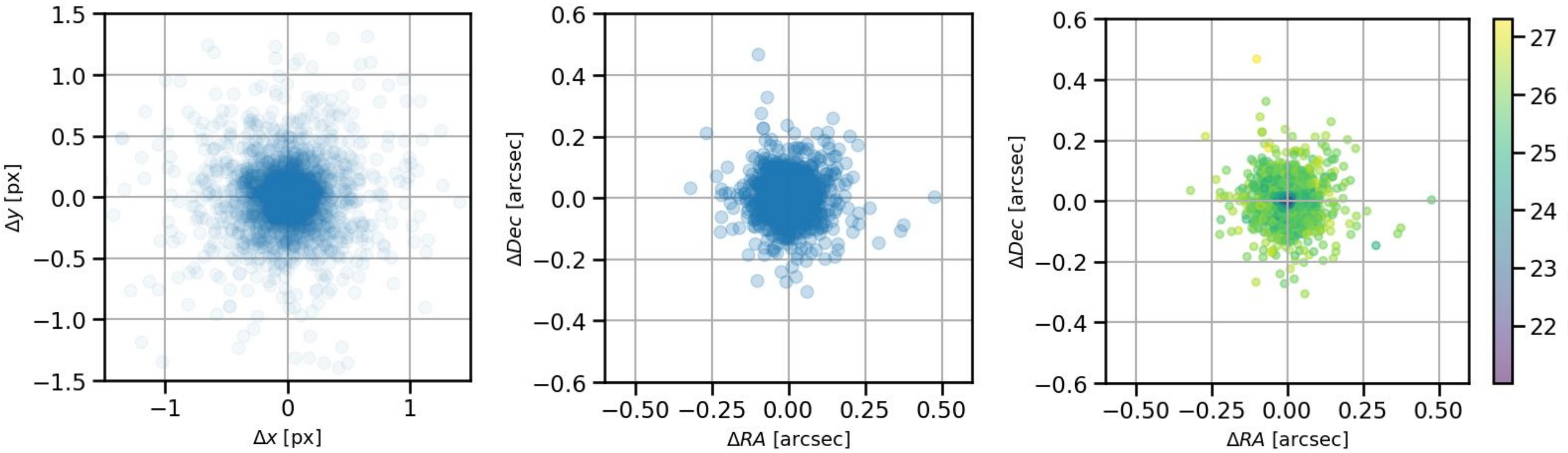
We define the value of $S/N_{1/2}$:= SNR at which Recall = 50%.
(Lower $SNR_{/2}$ is more sensitive)

Synthetic sources: HSC example

With fakes we can also test measurement results, like for example photometric measurements.
PSF photometry as well as Aperture
As function of fake magnitudes or even astrometric centroid distance to true centroid.



Synthetic sources: HSC example



Astrometric measurement of our Fakes position. In pixels, sky coordinates, and in combination with the brightness of the source

Open-Conclusions

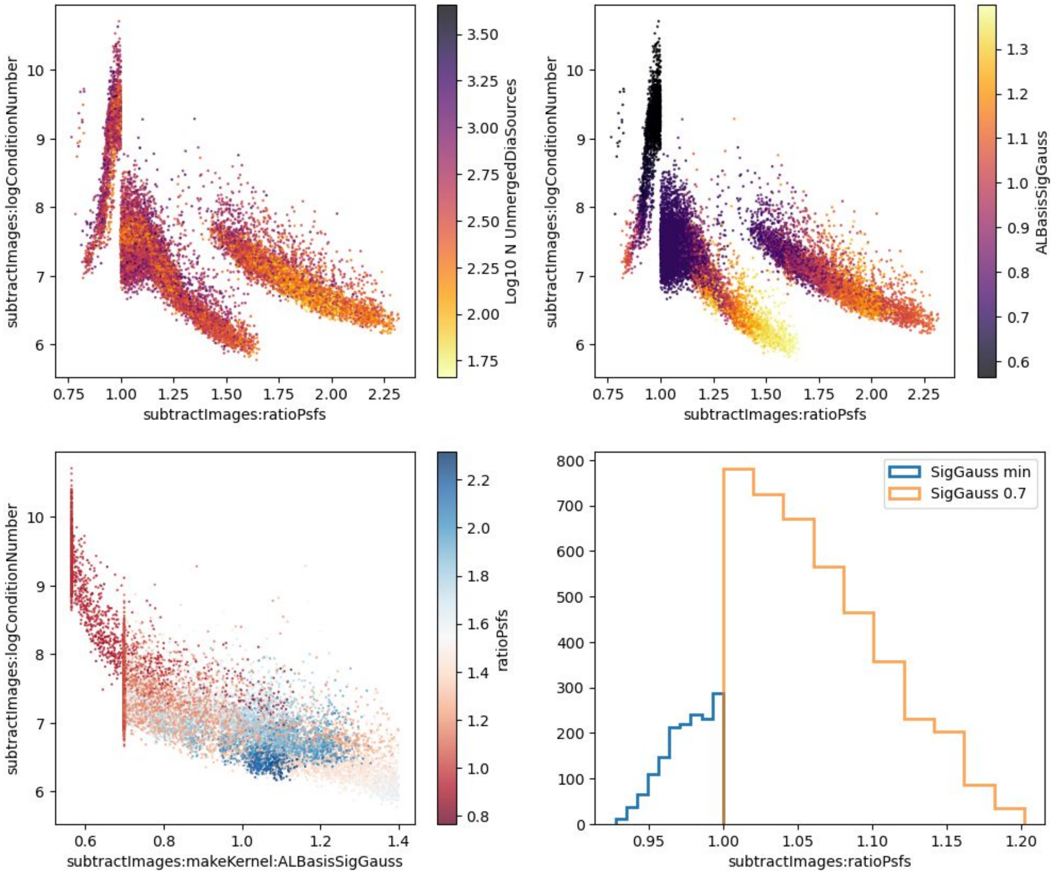
- We are conducting **regular testing and development of DIA algorithms** in the LSST Stack, we use diverse testing methodologies, and diverse datasets
- Valuable information both from visual inspection, low-level and high-level metrics
- The **synthetic source injection verification is a fundamental tool** to uncover problems and have a controlled scenario to assess improvements
- *We are currently applying all of these into LSST ComCam data and preliminary results are already internally in discussion.*
- We are moving forward with specific improvements for the DIA “Flavors”, more specifically auto-convolution
- We make use of the OR4 and future OR5 datasets as our benchmark for adopting improvements.

Back-Up Slides

Inspection of the `metadata`

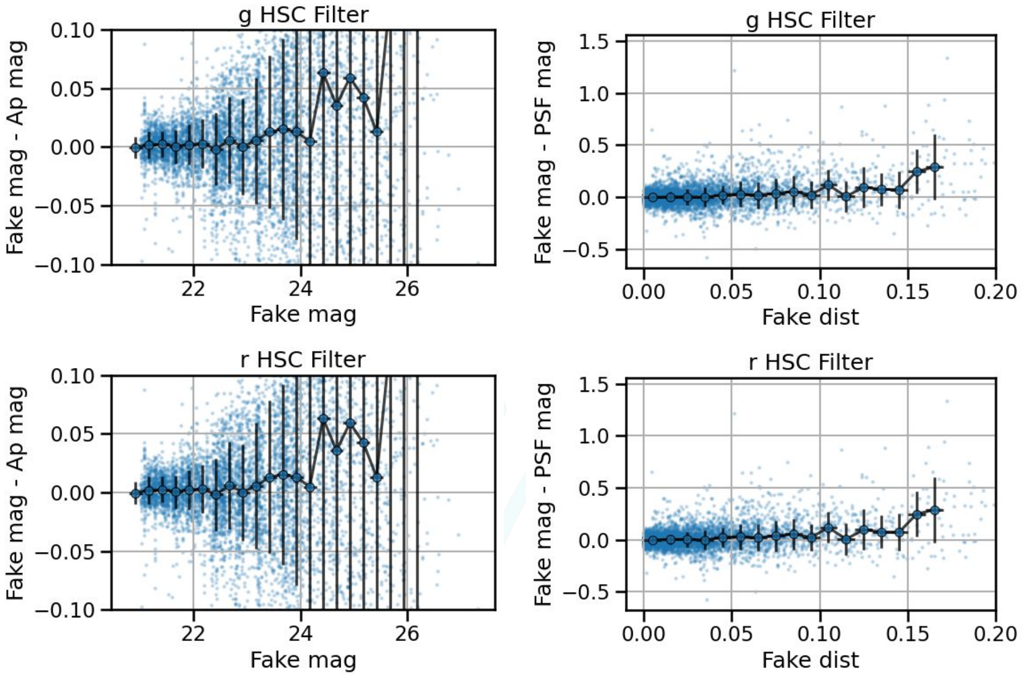
Metrics or metadata on the kernel is crucial for understanding our performance!
We inspect the dependency of two key parameters with the ratio of PSF sizes:

- ratioPsfs = $PSF_{Science} / PSF_{Template}$
- Condition number: the power ratio of kernel top coefficients
- SigGauss: the smallest size of the kernel basis components

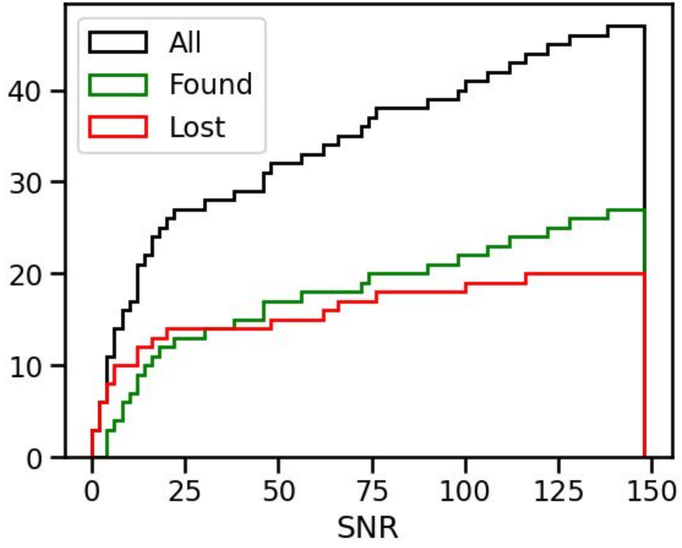


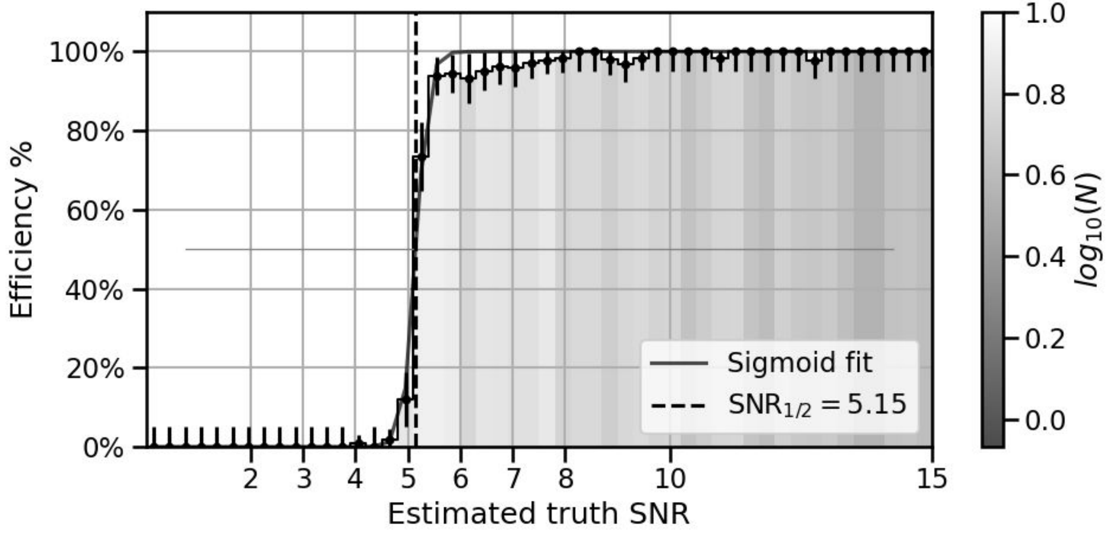
Synthetic sources: HSC example

With fakes we can also test measurement results, like for example photometric measurements.
PSF photometry as well as Aperture
As function of fake magnitudes or even astrometric centroid distance to true centroid.



HSC Processing of several images

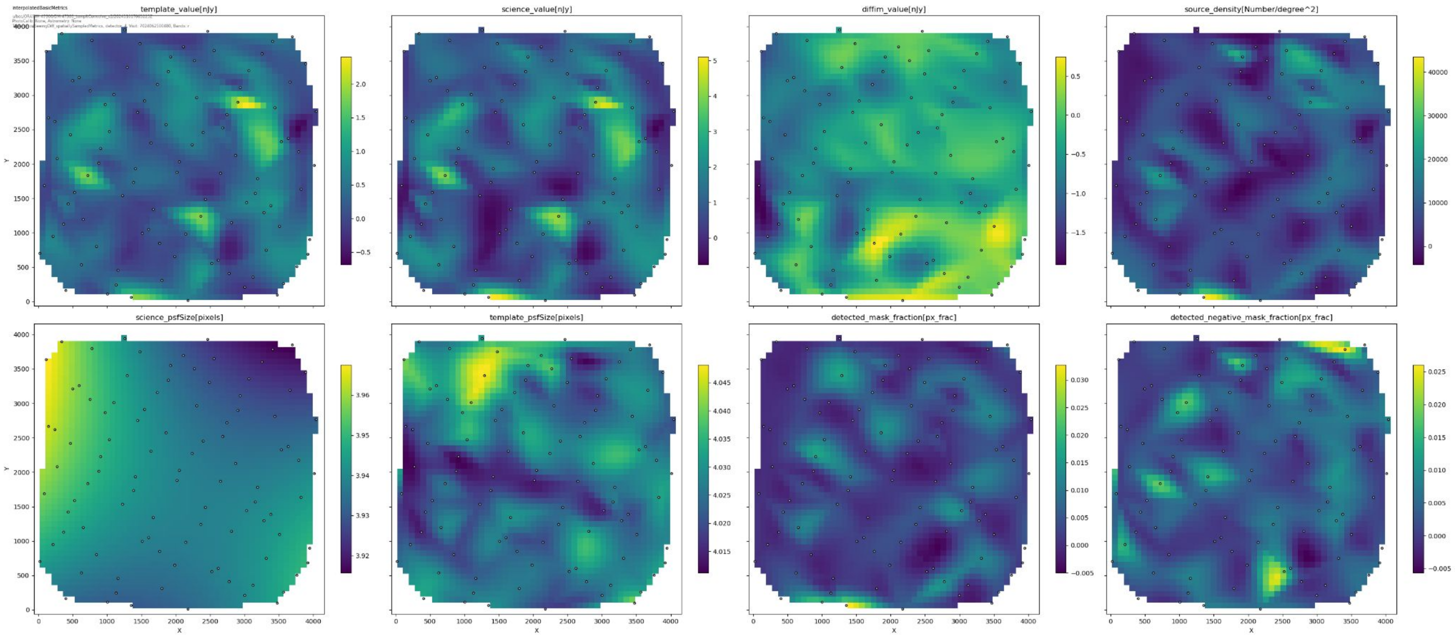




Recovery fraction per S/N bin

Testing DIA in the stack

These tests produce metrics, and analysis tools plots that are going to be used during LSST Survey in production phase.



lp_diffim Subtraction tasks

Things to keep in mind about broad perspective:

- The **difference image** subtraction task **is a link in a chain** inside a **very complex system**
-

lp_diffim Subtraction tasks

Things to keep in mind about broad perspective:

- The **difference image** subtraction task **is a link in a chain** inside a **very complex system**
- The main objectives of this system are always the detection of transients but **many more things depend on it**
  - The subtraction results yield products that are used for **measurements**
  - The **subtraction metadata** is also valuable piece of information – each implementation might handle this differently
-

lp_diffim Subtraction tasks

Things to keep in mind about broad perspective:

- The **difference image** subtraction task **is a link in a chain** inside a **very complex system**
- The main objectives of this system are always the detection of transients but **many more things depend on it**
  - The subtraction results yield products that are used for **measurements**
  - The **subtraction metadata** is also valuable piece of information – each implementation might handle this differently
- The core algorithm is to find a transformation kernel that is best suited for finding pixel flux changes, but implementation has to **deal** with many more **things such as masks** – this means **extra development** (such as Zogy case)