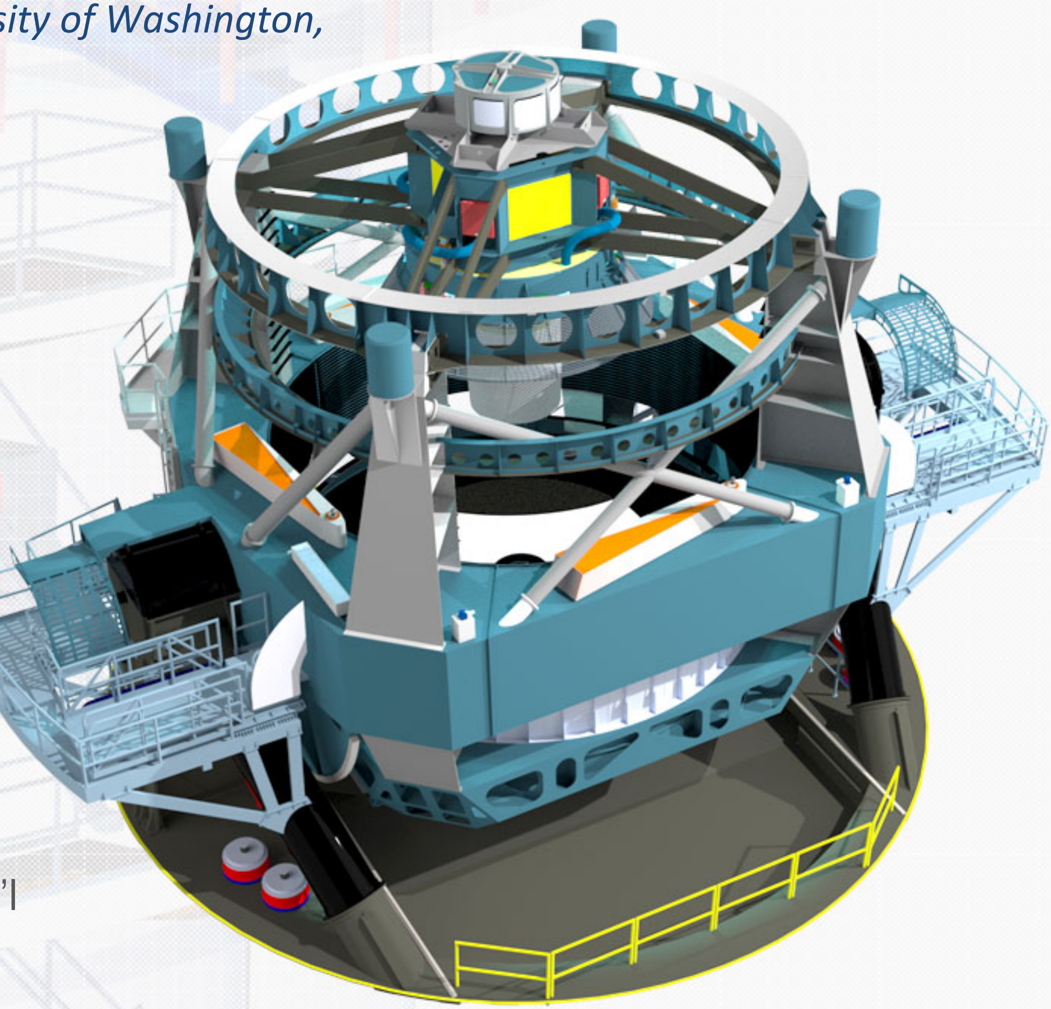
The Large Synoptic Survey Telescope: Data Challenges

Mario Juric
WRF Data Science Chair in Astronomy, University of Washington,
LSST Data Management Project Scientist

Lynne Jones
University of Washington,
LSST Performance Scientist
LSST Solar System Science Collaboration Lead



IAU 2015, HONOLULU, HAWAI'I
August 7th, 2015

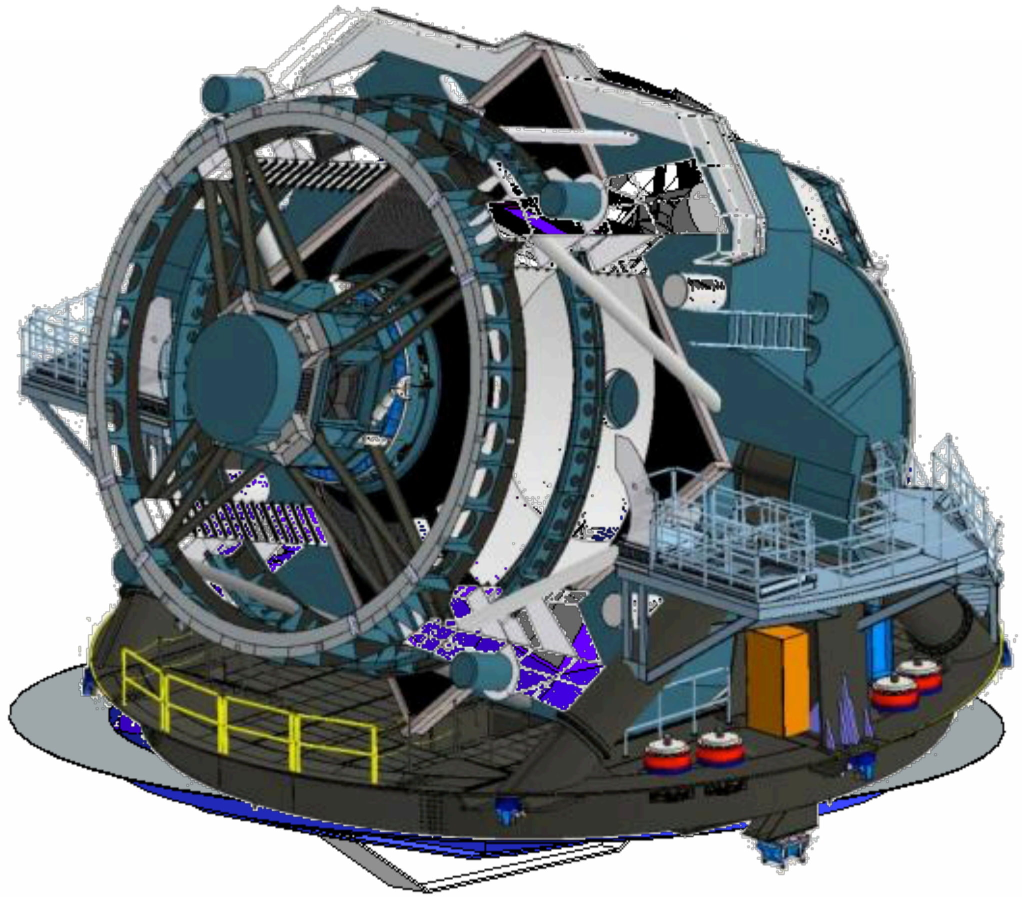
LSST: Evolution of Design and Purpose

History

1996-2000 *"Dark Matter Telescope"*
This project began as a quest to understand cosmology and the Solar System.

2000 - … *"LSST"*
Emphasizes a broad range of science from the same multi-wavelength survey data, including unique time domain exploration

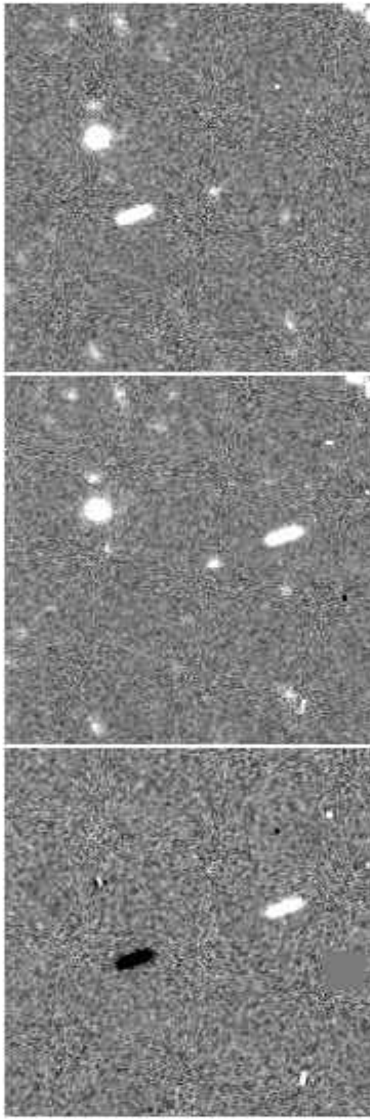
A single telescope, *a single data set*, can serve to answer a wide swath of science questions



The evolution of LSST design

Census of the Solar System is a Key Science Area

- Time domain science
  - Nova, supernova, GRBs
  - Source characterization
  - Instantaneous discovery
- **Census of the Solar System**
  - **MBAs, NEOs, Comets**
  - **KBOs, Oort Cloud**
- Mapping the Milky Way
  - Tidal streams
  - Galactic structure
- Dark energy and dark matter
  - Strong Lensing
  - Weak Lensing
  - Constraining the nature of dark energy



Exposure 1

Exposure 2

Exposure 1
-
Exposure 2

LSST From the User's Perspective: A Data Stream, a Database, and a (small) Cloud

Level 1

- A stream of ~10 million time-domain events per night, detected and transmitted to event distribution networks within 60 seconds of observation.
- A catalog of orbits for ~6 million bodies in the Solar System.

Level 2

- A catalog of ~37 billion objects (20B galaxies, 17B stars), ~7 trillion single-epoch detections ("sources"), and ~30 trillion forced sources, produced annually, accessible through online databases.
- Deep co-added images.

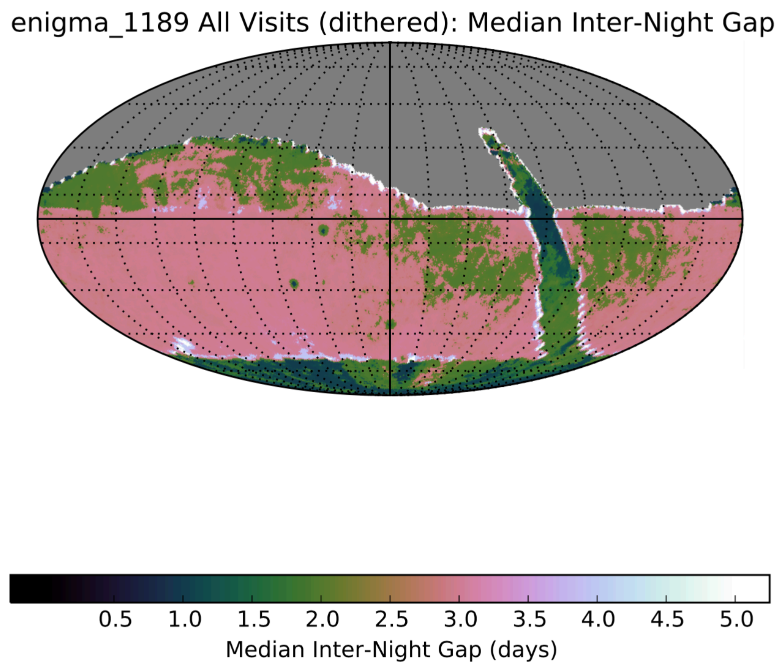
Level 3

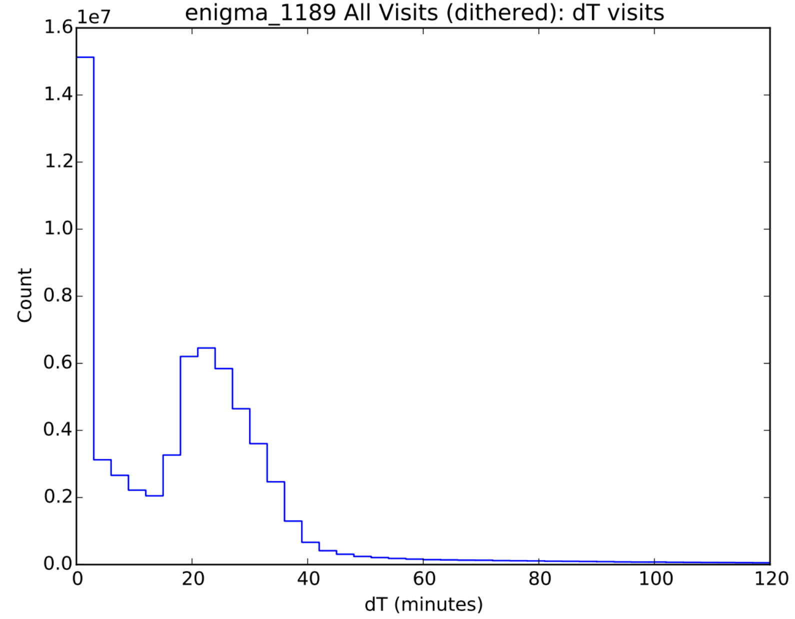
- Services and computing resources at the Data Access Centers to enable user-specified custom processing and analysis.
- Software and APIs enabling development of analysis codes.

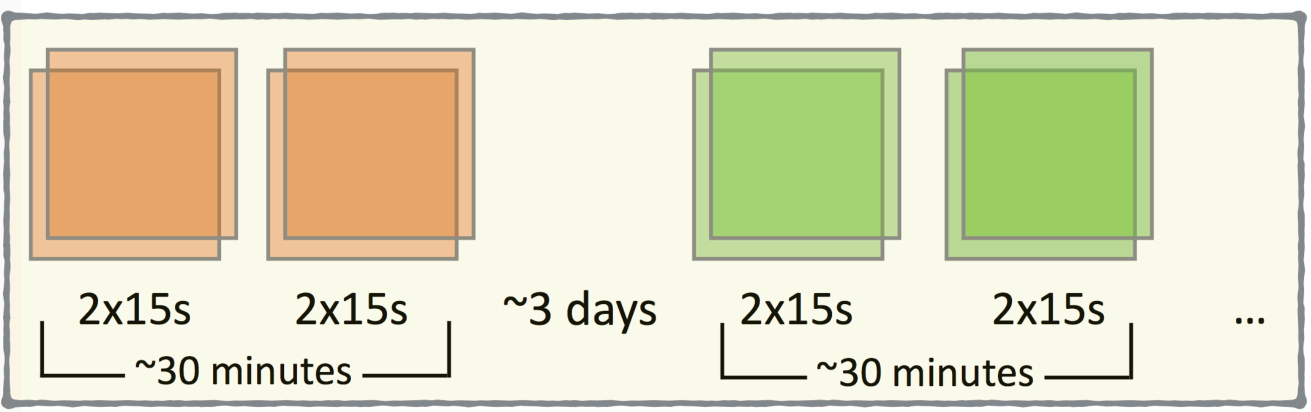
Level 1 Catalogs: Detections

- For each detection, we plan to measure:
  - Position
  - Shape (adaptive Gaussian moments; Bernstein & Jarvis 2002)
  - Model fits:
    - Point source model
      - Measure of flux and position assuming the object is a stationary point source
    - Trailed source model
      - Measure of flux, position, and direction of motion, assuming object moves sufficiently fast to trail in the image. Designed for Solar System objects.
    - Dipole model fit
      - Fit the source with a “dipole” model, a positive next to a negative point source
- Some detections will be false positives. We will flag those using a real-bogus type algorithm.

LSST Surveying Strategy



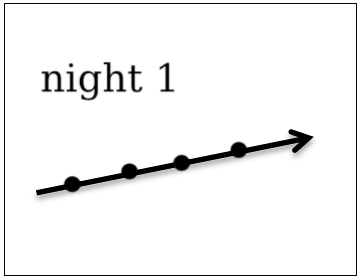




1 visit = 2x15s back-to-back exposures
2 visits per night (30-60 minutes)
Revisit ~3-4 days

Asteroid Discovery Strategies

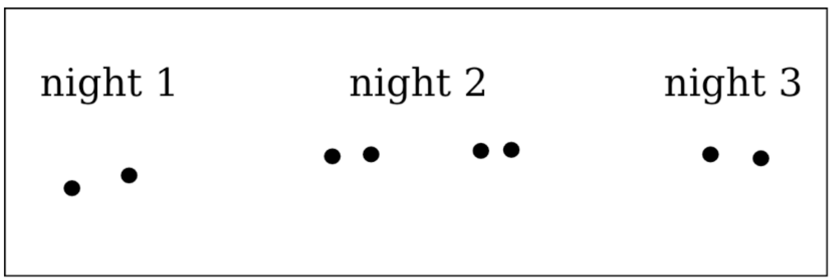
Rapid-response optimized

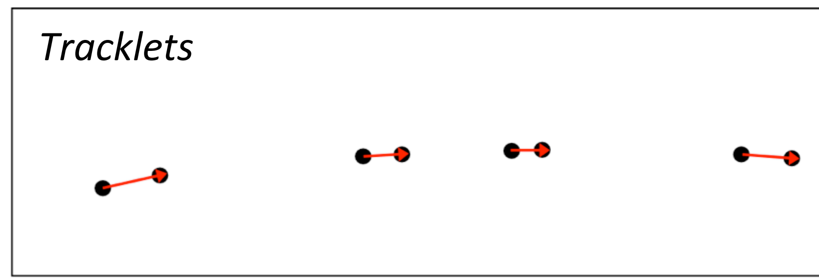


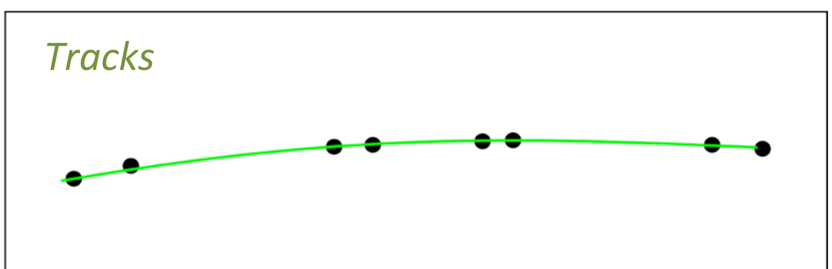
Requirement for discovery: four closely spaced detections in a single night.

Generally used by surveys today. Simple to implement and execute, mitigates quality issues with the telescope or camera, at the expense of reduced nightly survey area and orbit accuracy at discovery. Typically requires follow-up by other telescopes.

Long-term yield optimized







Requirement for discovery: three pairs taken over three nights in a short (e.g., ~14 day) period.

Planned for PanSTARRS 4 and LSST. Maximizes discovery capability, orbit accuracy, and nightly survey area, at the expense of increased software complexity and more stringent instrument design. Requires no third-party follow-up.

Why 2+2+2? Optimizes efficiency …

- More sky can be covered each night

- Minimize the need for rapid follow-up of particularly interesting objects (e.g., NEOs). Infeasible at LSST depths (r~24.5, single visit).

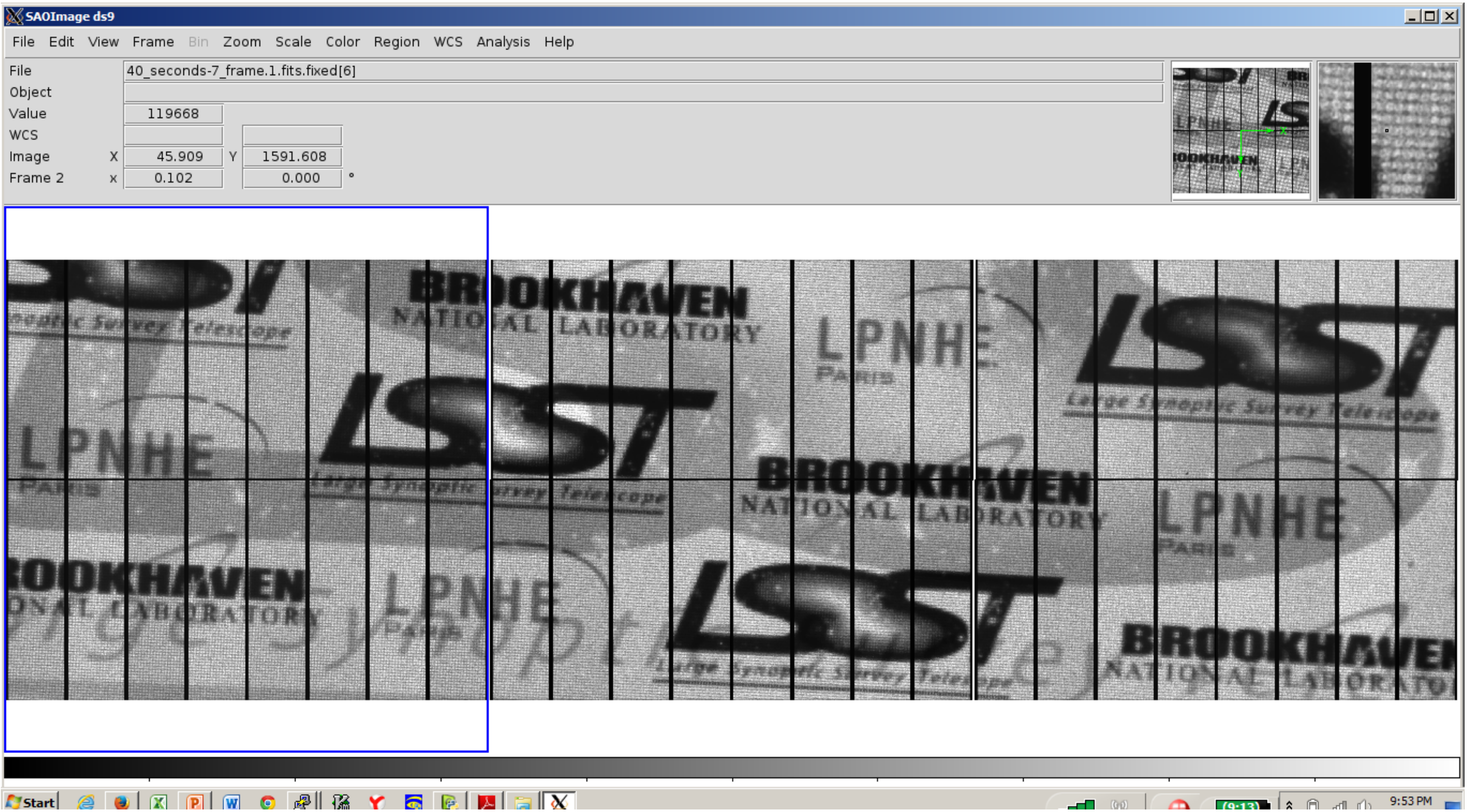
- Better estimate of orbit at discovery (maximizing the information gain from each individual measurement)

- Better characterization of the survey selection function

… at the expense of increased complexity and quality requirements.

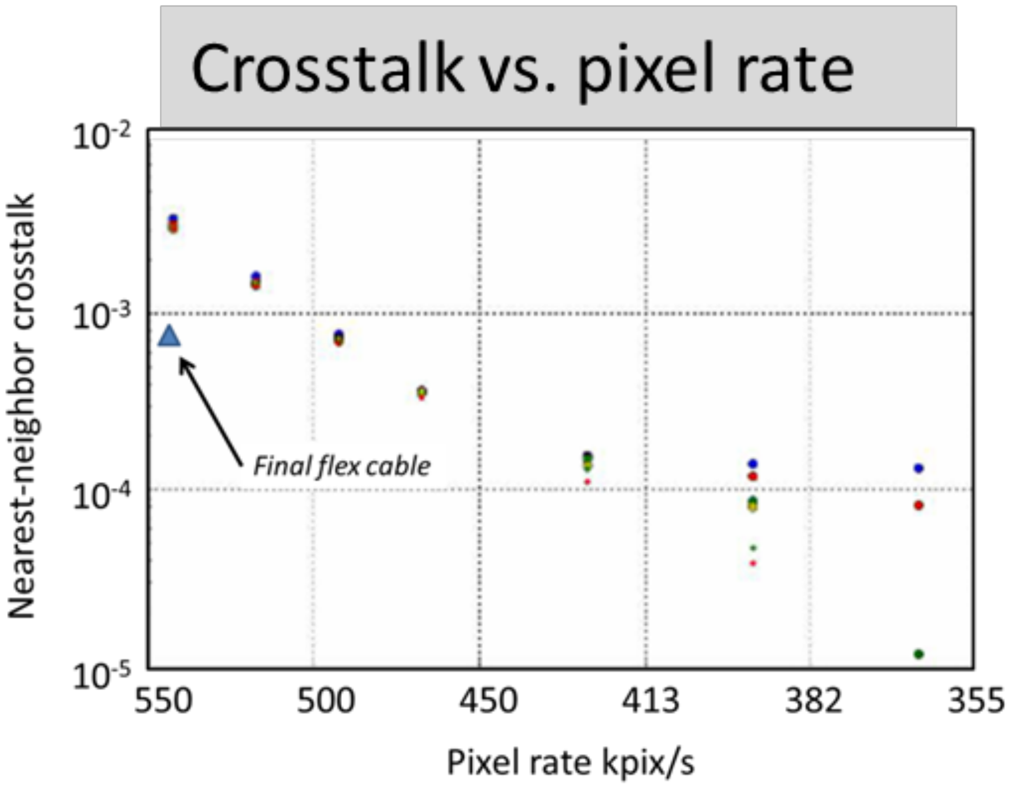
- Increased complexity:
    - Software
    - Cadence
- Increased computational demands
- Probabilistic thinking
- High quality demands to avoid/mitigate false positives
    - Camera
    - Optical system
    - Image differencing software
    - *Note: LSST is driven in these areas even harder by other science cases (e.g., transient and dark energy science)*

LSST CCDs: Within spec, clean



Credit: LSST Camera Team

Small, stable, cross-talk, no charge persistence



Amp-amp cross-talk well within spec (2e-3) and stable.

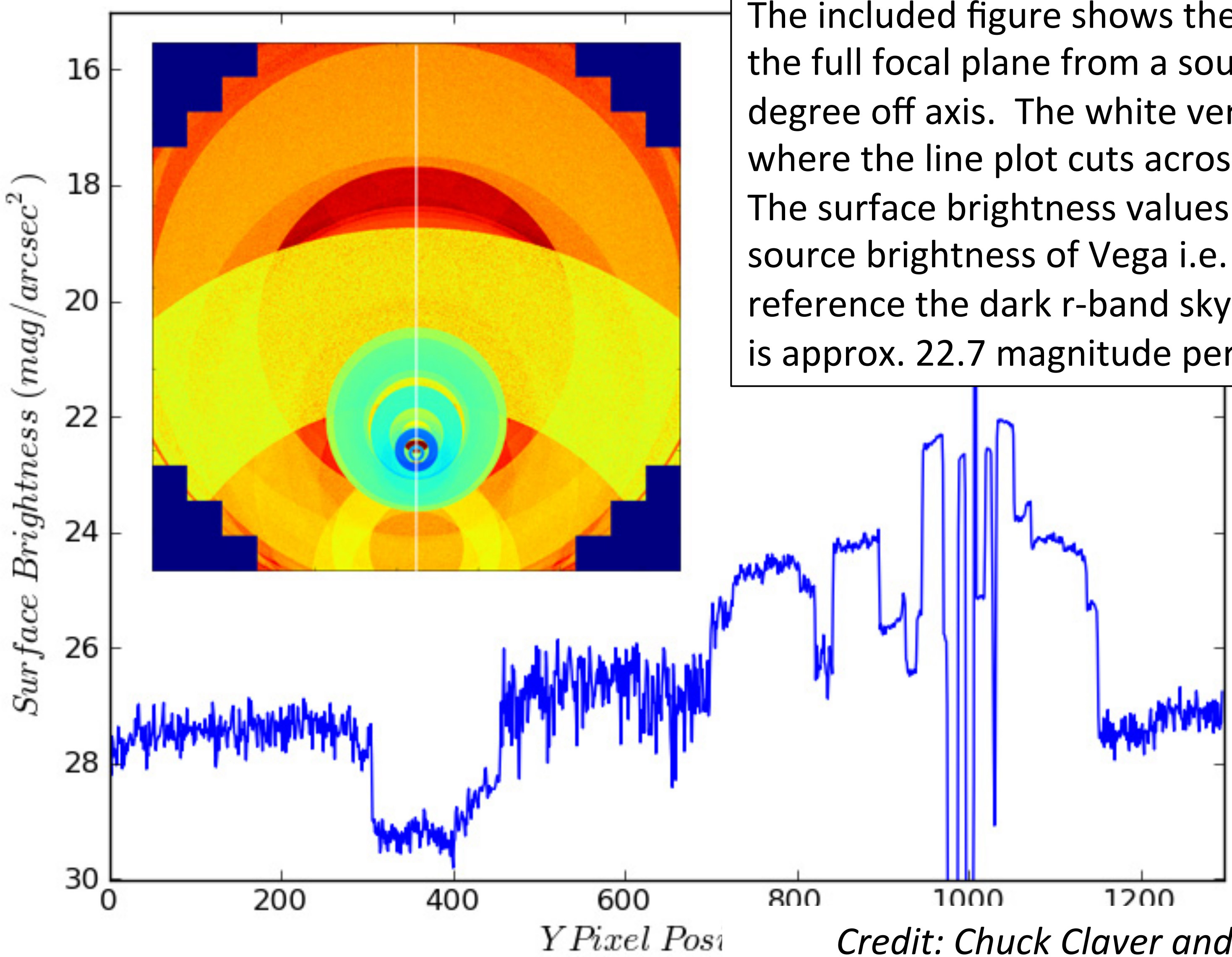
No measureable CCD-CCD crosstalk down to limit of few 10^{-5}.

No measurable charge persistence in any conditions (vendor A) and with optimized clocking (vendor B)

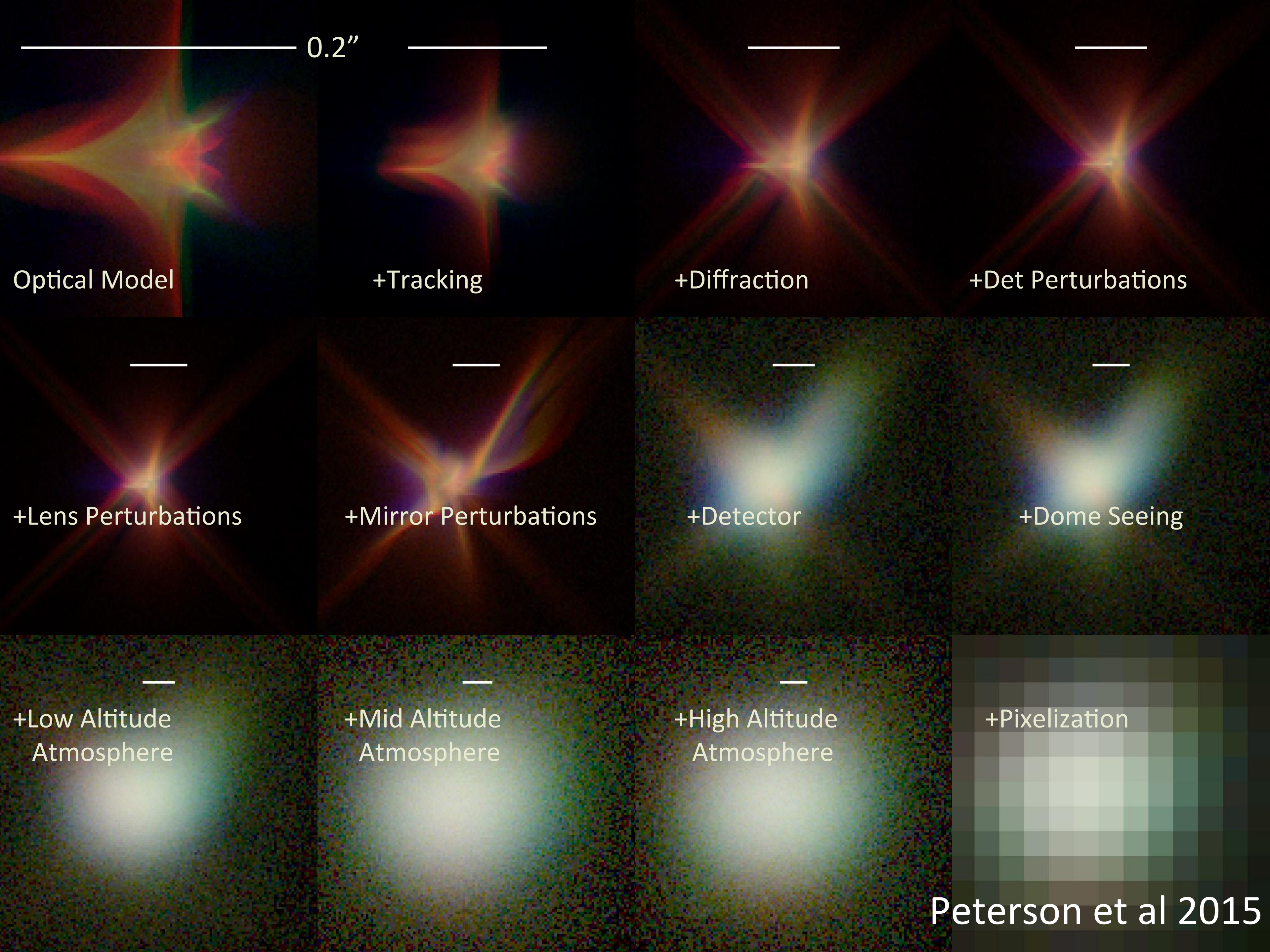
Credit: LSST Camera Team

No optical ghosts, well characterized optical system

The included figure shows the ghost pattern over the full focal plane from a source located 1 degree off axis. The white vertical line indicates where the line plot cuts across the field of view. The surface brightness values are scalled for a source brightness of Vega i.e. M_r=0. For reference the dark r-band sky surface brightness is approx. 22.7 magnitude per sq. arcsec.



Credit: Chuck Claver and the LSST T&S Team


0.2"
Optical Model
+Tracking
+Diffraction
+Det Perturbations
+Lens Perturbations
+Mirror Perturbations
+Detector
+Dome Seeing
+Low Altitude Atmosphere
+Mid Altitude Atmosphere
+High Altitude Atmosphere
+Pixelization
Peterson et al 2015

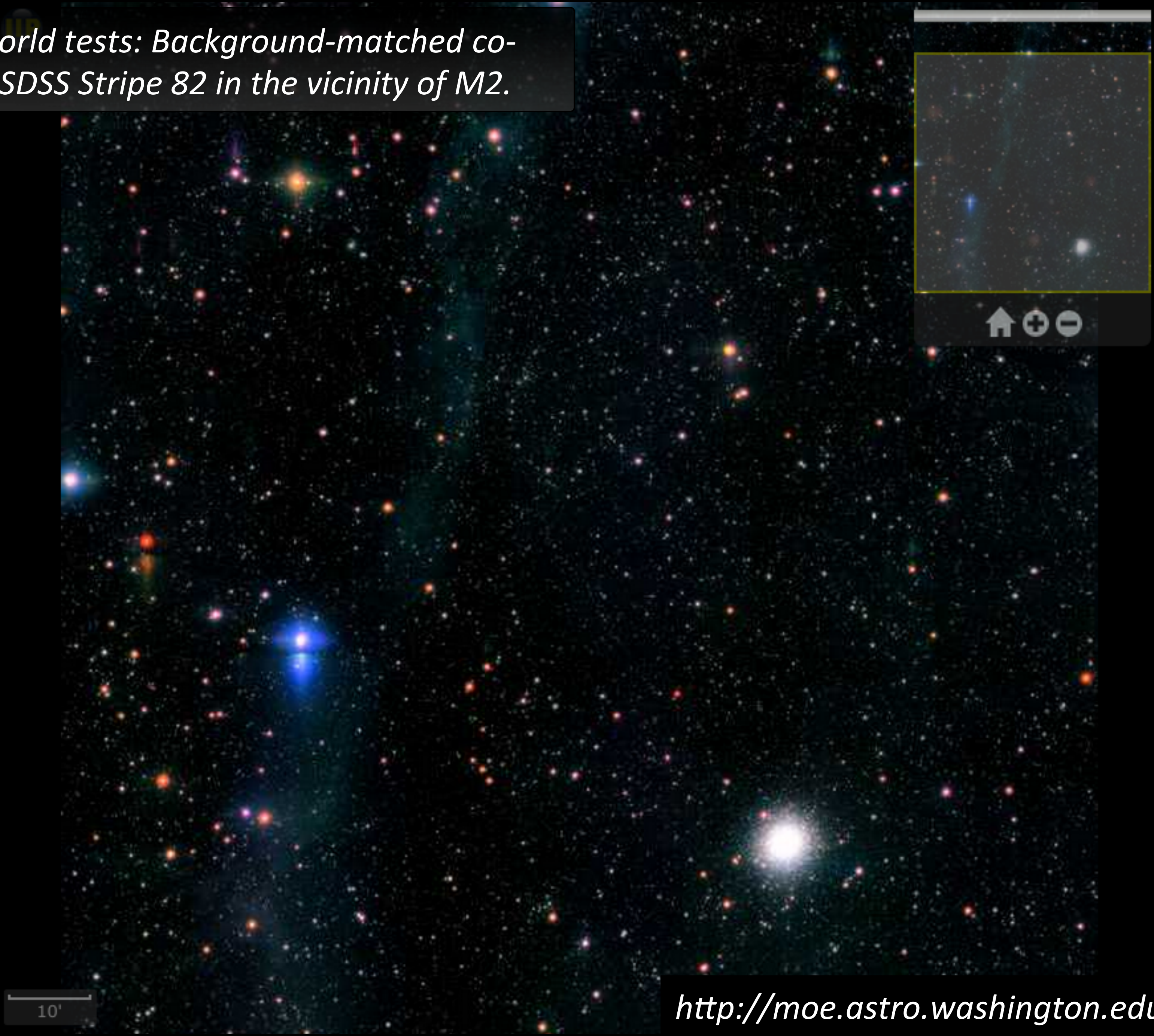

−4°54′48.0″
55′00.0″
12.0″
24.0″
36.0″
Dec (J2000)
49.00ˢ
48.00ˢ
47.00ˢ
0ʰ05ᵐ46.00ˢ
RA (J2000)

A Significant Software Effort

- **8-year prototyping effort**
  - 8 software releases (Data Challenges)
  - Status: A rapidly maturing state-of-the art astronomical data reduction system
    - ~SDSS level quality of reductions
    - ***Most recently tested by building co-adds using SDSS Stripe 82 data***
    - ***Used in production by the Hyper Suprime-Cam Survey on Subaru***

- **Prototyped Features:**
  - Instrumental signature removal
  - Single-frame processing
  - Point source photometry
  - Extended source photometry (model fitting)
  - Deblender
  - Co-addition of images
  - **Image differencing (early prototype!)**
  - **MOPS (early prototype!)**
  - Object characterization on multi-epoch data (StackFit/MultiFit)
  - …

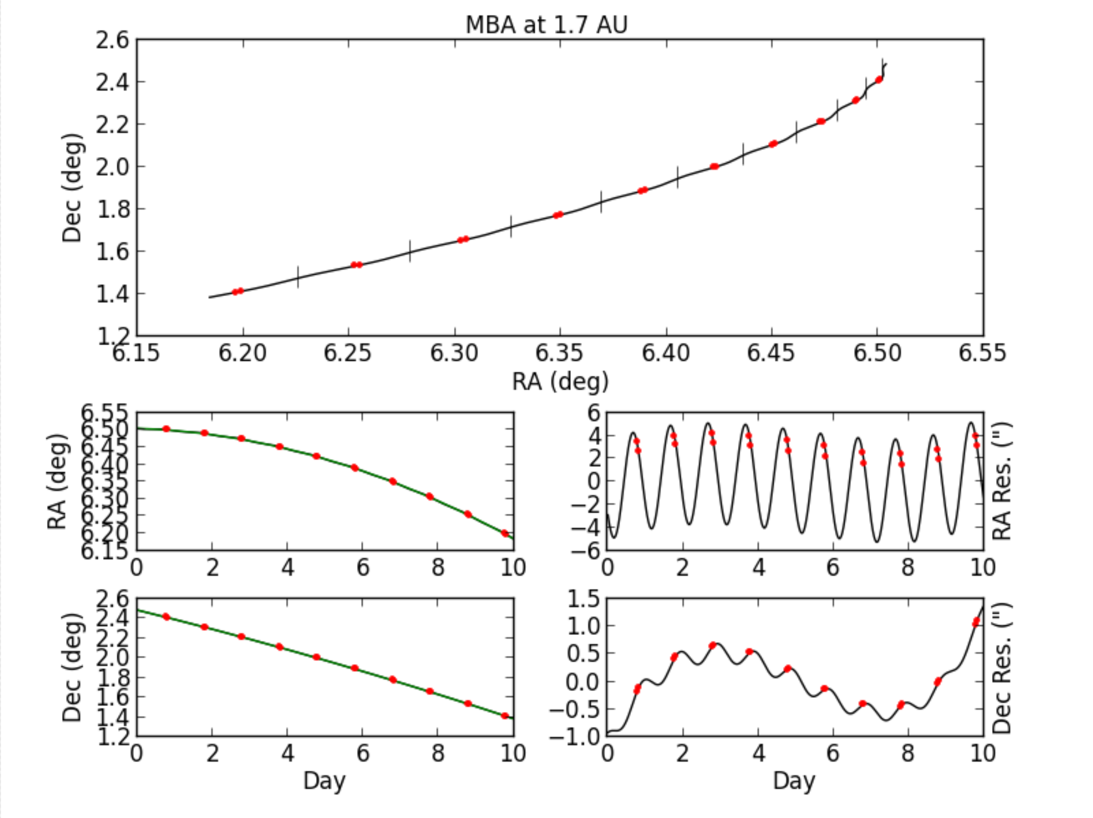
Real-world tests: Background-matched co-add of SDSS Stripe 82 in the vicinity of M2.



http://moe.astro.washington.edu/sdss/

The work on MOPS has started:
Early MOPS Improvements – Topocentric correction

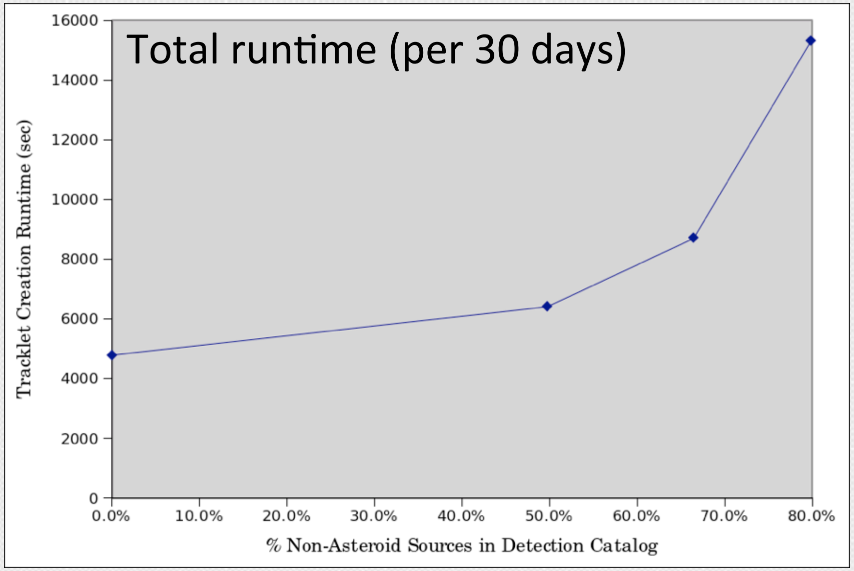
- Many tracks will be random alignments, because of (necessarily) generous matching criteria
- Full orbital solutions are costly: try to cut down the number of candidates by computing the topocentric correction, and reject tracks that don't meet empirically determined thresholds
- This reduces number of false tracks at a given completeness for true tracks dramatically - ~50x

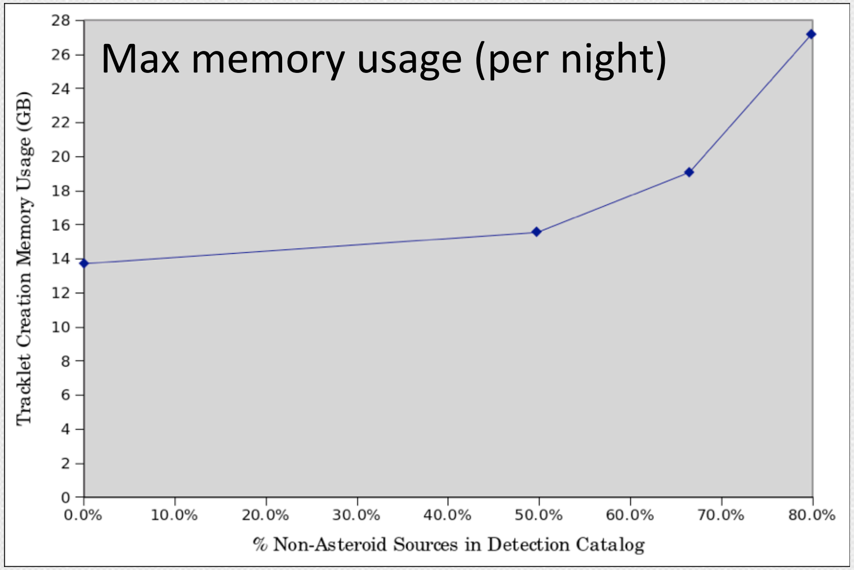


Early MOPS Prototype Performance

- Test while increasing amounts of false positives, ranging from 0 to 80% false positives
- Time and memory requirements scale highly non-linearly
  - Between O(n^2) and O(n^3) scaling
- Positives:
  - Did not lose significant numbers of objects as noise increased
  - Even at high noise, this is tractable for LSST

Credit: Lynne Jones and the LSST MOPS team





LSST and Planetary Defense

- **As deployed and funded (NSF+DOE), LSST is primarily a science-driven mission**
  - Existing cadence is optimized to maximize the overall science returns (incl. Solar System science), rather than NEO/PHA discovery completeness
- **As deployed and funded, the survey is not optimized for rapid discovery and follow-up of all types of moving objects.**
  - Note: LSST will enable rapid identification and follow-up of ***trailed*** objects (within 60 seconds of discovery)
- **If deployed with a planetary-defense optimized cadence, the NEO yields could be improved**
  - Early simulations indicate 90% is achievable (***more work is needed!***)
  - Other science goals would be affected (including Solar System science)
- **The LSST and the IR space-base missions would result in highly complementary science returns**
  - Colors, albedos, etc… A boon for Solar System science!

Lots of work ahead of us!

- LSST development is on track for first light in ~2019 and start of operations in ~2022

- Lots of hard work devoted to hardware improvements is beginning to pay off!

- Still, many challenges remain, especially in the software area!

- 7 years of further software development ahead of us
  - Already 8 years of software R&D, prototype pipelines in place

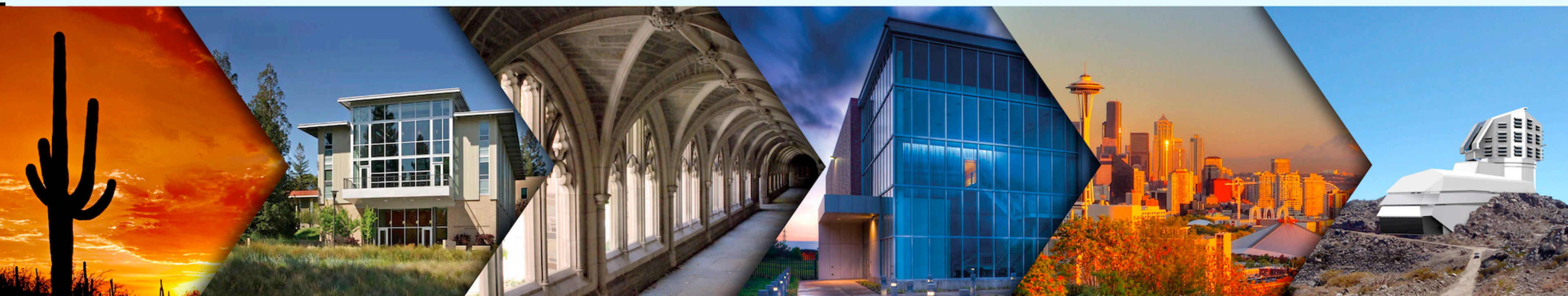
- Planing to run LSST moving object pipelines on large scale simulations and real data even before LSST commissioning
  - E.g., DECam, PS1 (?)

LSST IS HIRING

Large Synoptic Survey Telescope

LSST.ORG

WE'RE SEEKING TOP TALENT TO WORK IN A TEAM ENVIRONMENT THAT INSPIRES EXCELLENCE.



JOIN US IN:

LSST HEADQUARTERS
TUCSON, AZ

SLAC/STANFORD
MENLO PARK, CA

PRINCETON UNIVERSITY
PRINCETON, NJ

NCSA / UIUC
URBANA-CHAMPAIGN, IL

UNIVERSITY OF WASHINGTON
SEATTLE, WA

LSST OBSERVATORY SITE
CERRO PACHÓN, CHILE

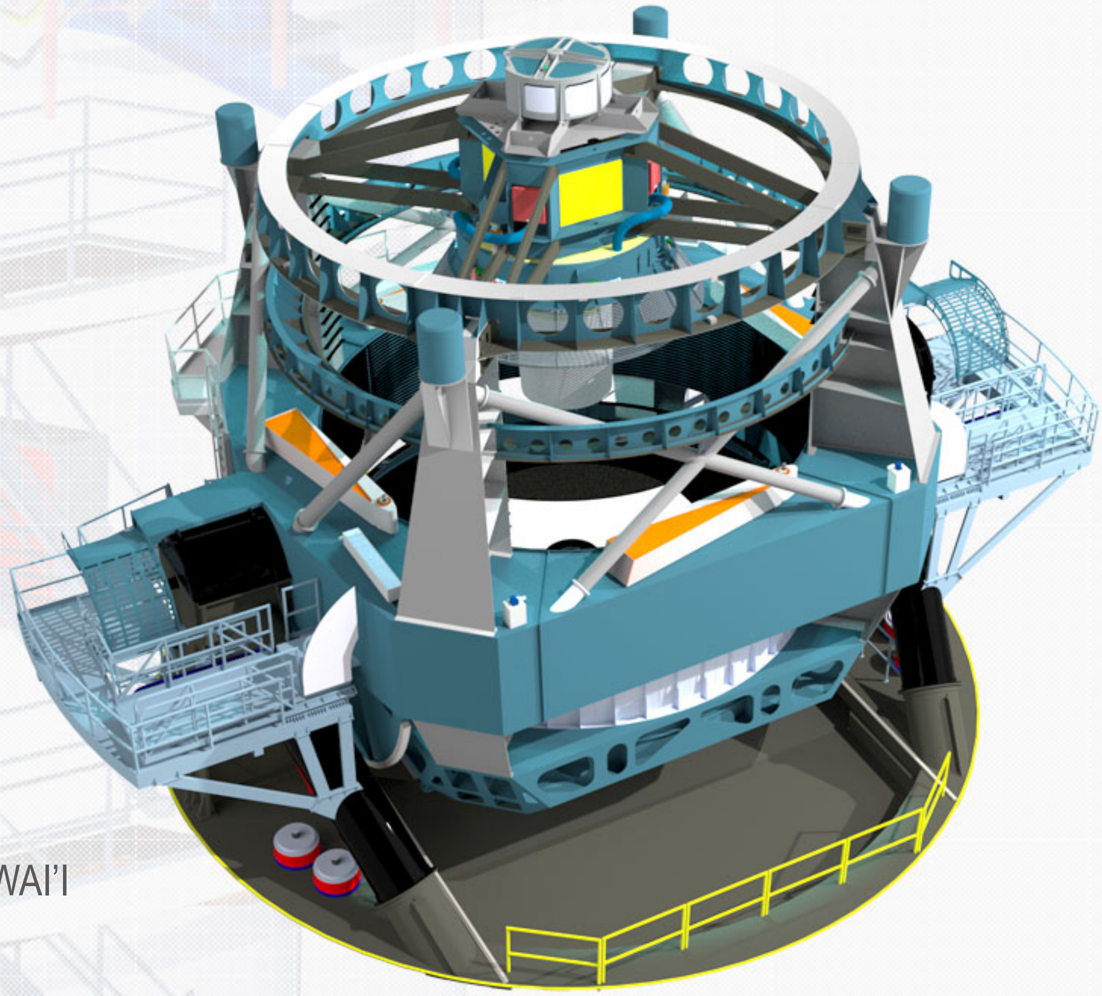
LSST.ORG/HIRING

NSF

U.S. DEPARTMENT OF ENERGY | Office of Science

CHARLES AND LISA SIMONYI FUND
FOR ARTS AND SCIENCES

Thank You!



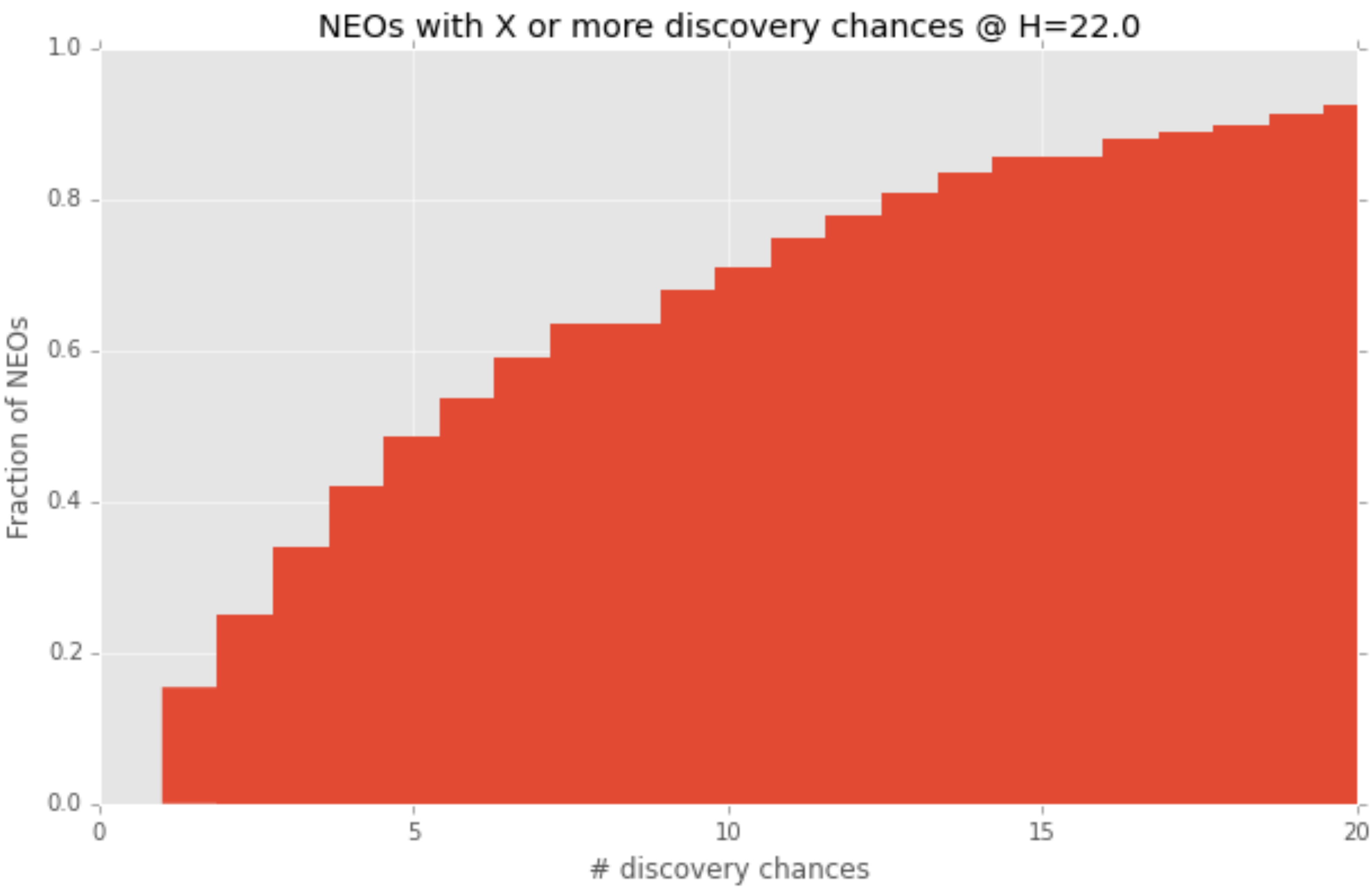
IAU 2015, HONOLULU, HAWAI'I

August 4th, 2015

Worst-case, fallback scenarios

- **CAVEAT: VERY EARLY, UNOPTIMIZED, RESULTS**
  - **(literally 16 hrs old!)**
- Baseline 2+2+2:
  - 73% completeness to H=22
- 2+2+2+2:
  - See next slide
- 3+3+3:
  - 56% completeness to H=22
- 4+4+4:
  - 43% completeness to H=22
- Again, these are likely lower limits, as the cadence optimizer doesn't take into account where the asteroids are on the sky.

Multiple discovery opportunities for large fraction of NEOs!



Very few are discoverable only once