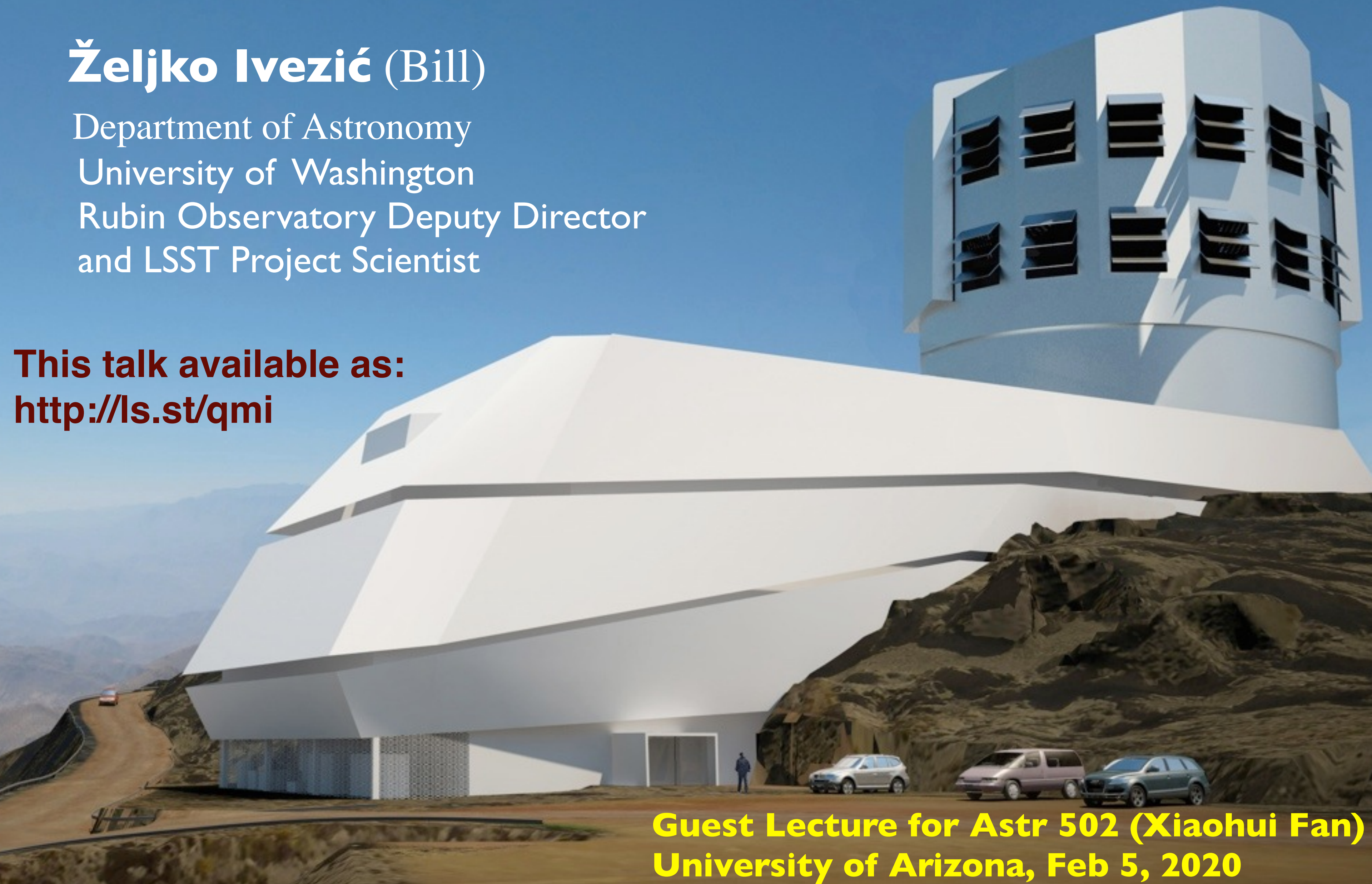
LSST overview:
the greatest movie of the Universe

Željko Ivezić (Bill)
Department of Astronomy
University of Washington
Rubin Observatory Deputy Director
and LSST Project Scientist



This talk available as:
http://ls.st/qmi

Guest Lecture for Astr 502 (Xiaohui Fan)
University of Arizona, Feb 5, 2020

Outline

- LSST overview and construction status report
  - multi-color time-resolved faint sky map
  - 20 billion stars and 20 billion galaxies
- A tour of anticipated LSST science programs
  - time domain
  - the Solar System structure
  - cosmology (dark matter and dark energy)
  - the Milky Way structure
- Data analysis challenges ahead of us
  - large data sets
  - complex analysis
  - aiming for small systematics

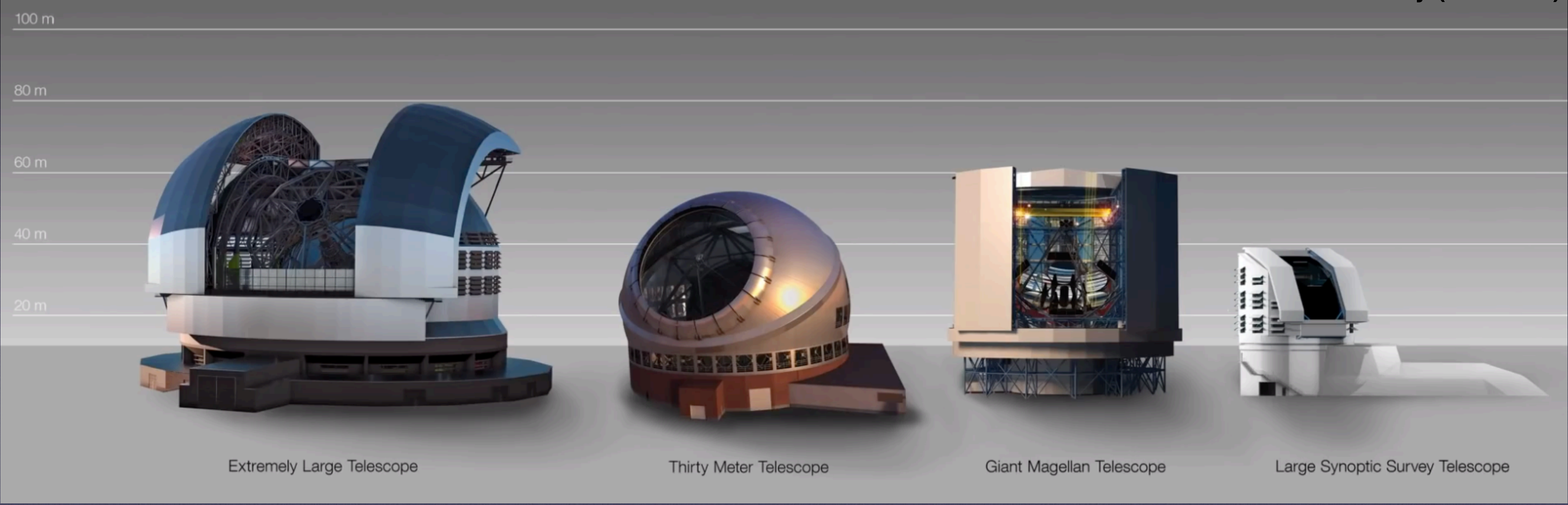
"Ask Not What Data You Need To Do Your Science, Ask What Science You Can Do With Your Data."

The era of surveys...

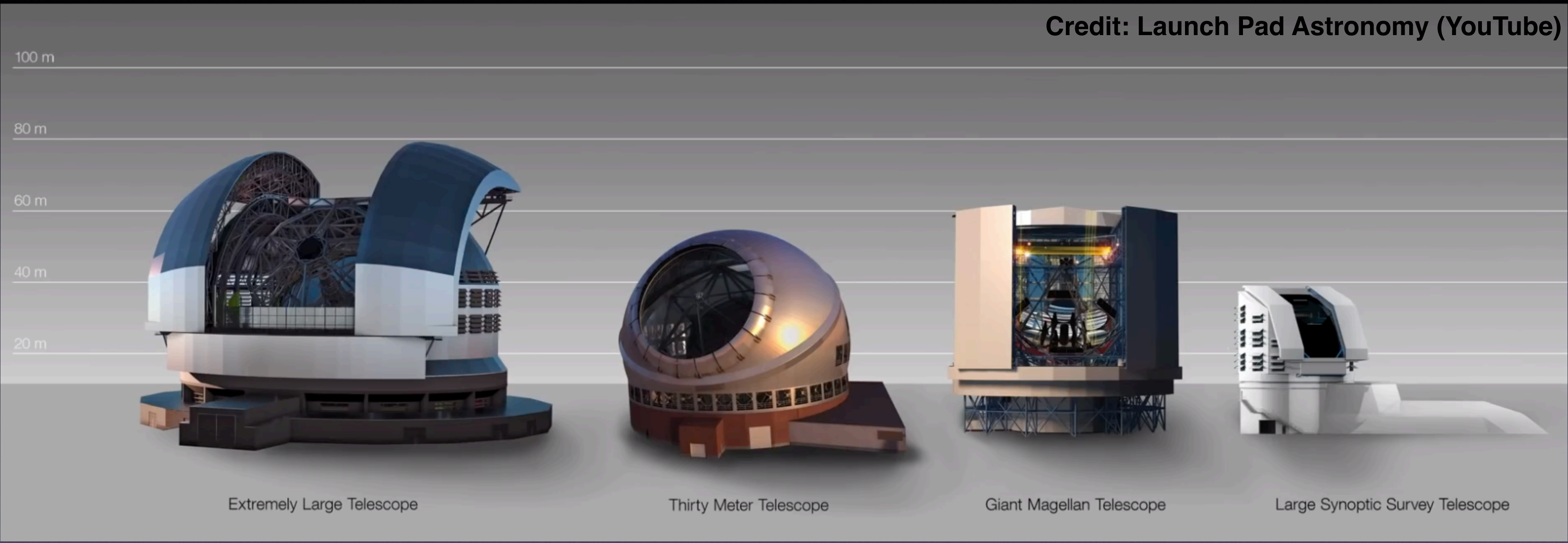
- Standard: "What data do I have to collect to (dis)prove a hypothesis"?
- Data-driven: "What theories can I test given the data I already have?"

Why do we need another sky survey, such as LSST? Is it worth the investment of 1,400,000,000 USD?

Credit: Launch Pad Astronomy (YouTube)



Why do we need another sky survey, such as LSST? Is it worth the investment of 1,400,000,000 USD?



LSST will not have the largest mirror but will have by far the largest product of the mirror area and the field-of-view size (etendue or throughput)

Why do we need another sky survey, such as LSST? Is it worth the investment of 1,400,000,000 USD?

- A large sky survey, enabled by the progress in telescope, sensor and computing technologies, can address a broad array of science questions, including pressing cosmological problems. Funding from the National Science Foundation, Department of Energy, and private gifts (Simonyi, Gates, Schmidt,…)
- Congressional mandate to NASA to catalog dangerous asteroids that might collide with the Earth.

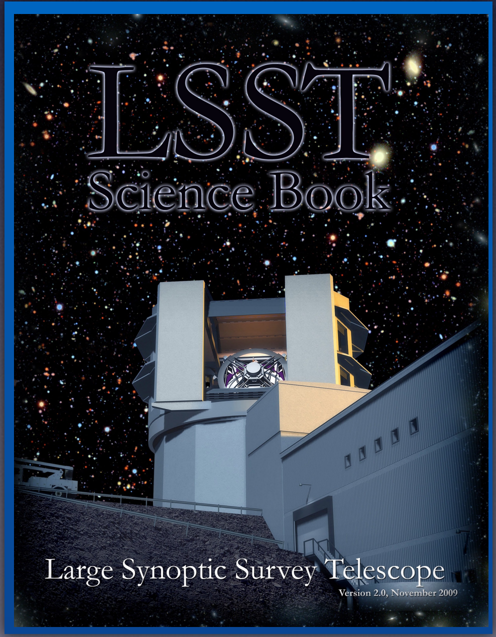
LSST Science Themes

- Dark matter, dark energy, cosmology (spatial distribution of galaxies, gravitational lensing, supernovae, quasars)
- Time domain (cosmic explosions, variable stars)
- The Solar System structure (asteroids)
- The Milky Way structure (stars)

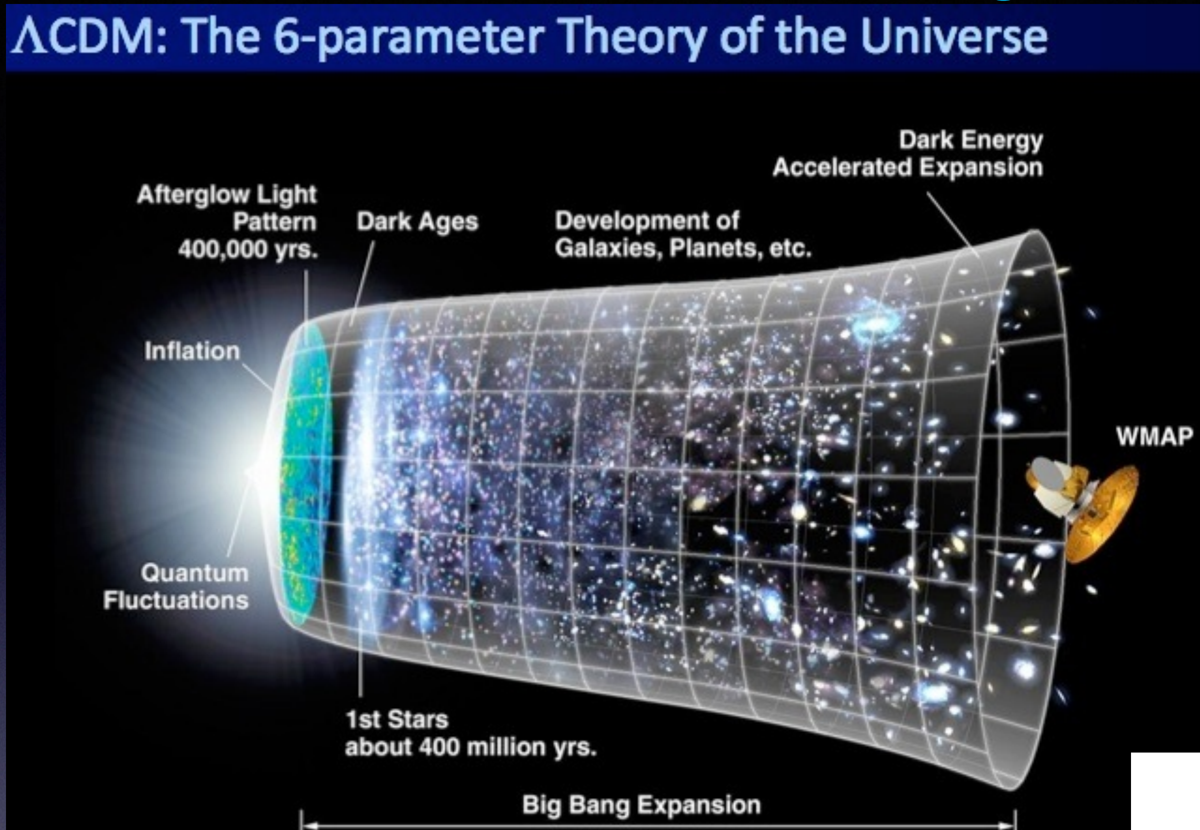
LSST Science Book: arXiv:0912.0201

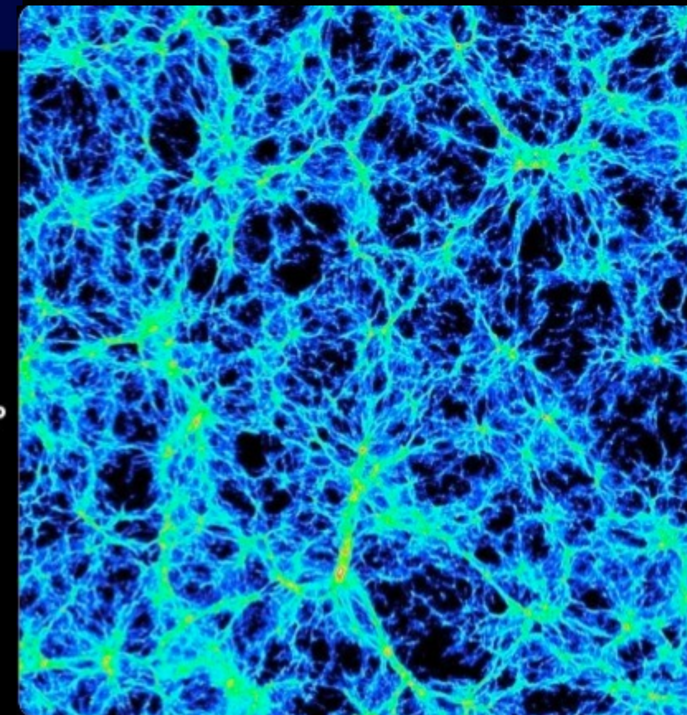
Summarizes LSST hardware, software, and observing plans, science enabled by LSST, and educational and outreach opportunities

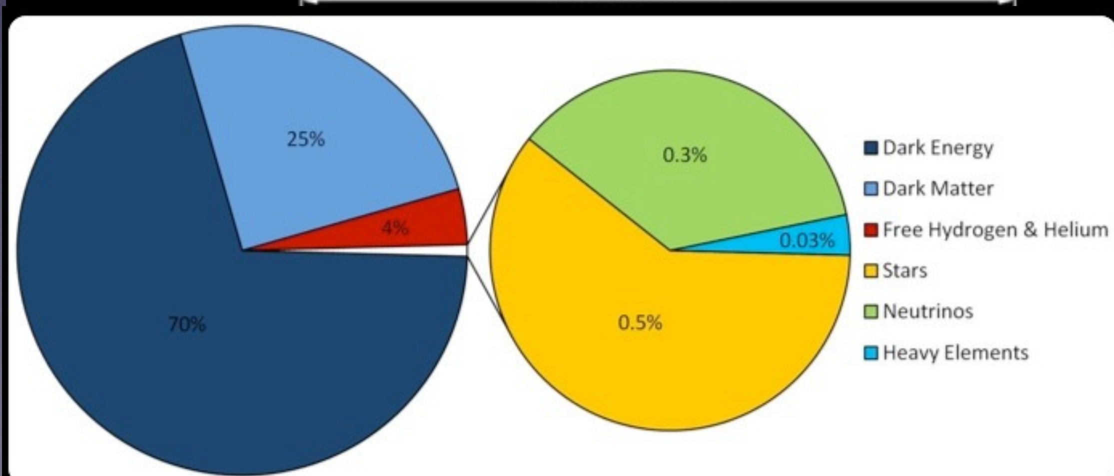
245 authors, 15 chapters, 600 pages



New Cosmological Puzzles







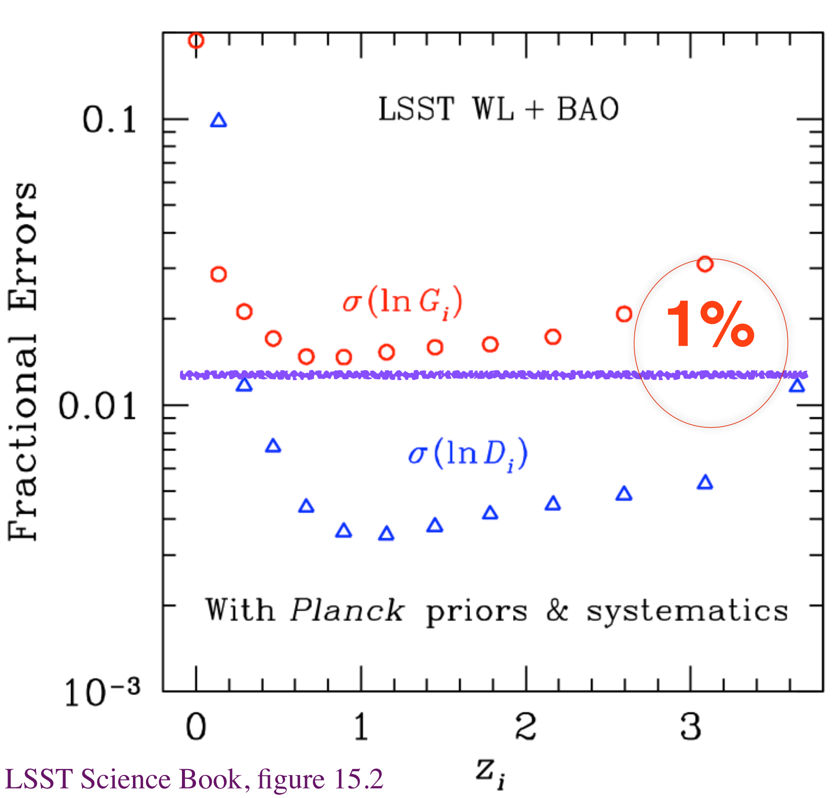
The modern cosmological models can explain all observations, but need to **postulate** dark matter and dark energy (though gravity model could be wrong, too)

Modern Cosmological Probes

- Cosmic Microwave Background (CMB) (the state of the Universe at the recombination epoch, at redshift ~1000)
- Weak Lensing: growth of structure
- Galaxy Clustering: growth of structure
- Baryon Acoustic Oscillations: standard ruler
- Supernovae: standard candle

Except for CMB, for precise cosmological measurements need to detect and precisely measure properties of billions of galaxies and millions of supernovae

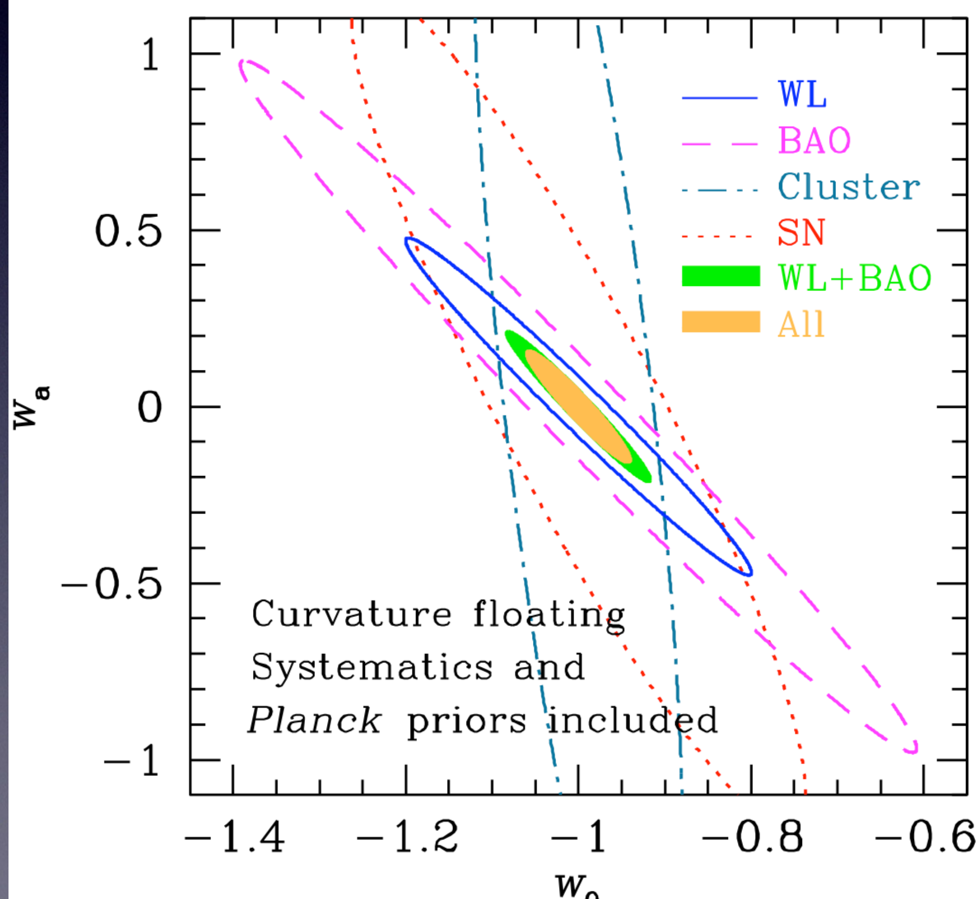
Cosmology with LSST: high precision measurements



LSST Science Book, figure 15.2

"Was Einstein right?"

- Multiple probes is the key!



LSST Science Book, figure 15.3

By simultaneously measuring growth of structure and curvature of the Universe, LSST data will tell us whether the recent acceleration is due to **dark energy or modified gravity.**

National Academies of Sciences, Engineering, and Medicine 2019. *Finding Hazardous Asteroids Using Infrared and Visible Wavelength Telescopes*. Washington, DC: The National Academies Press.

https://doi.org/10.17226/25476.

Recommendation: Objects smaller than 140 meters in diameter can pose a local damage threat. When they are detected, their orbits and physical properties should be determined, and the objects should be monitored insofar as possible.

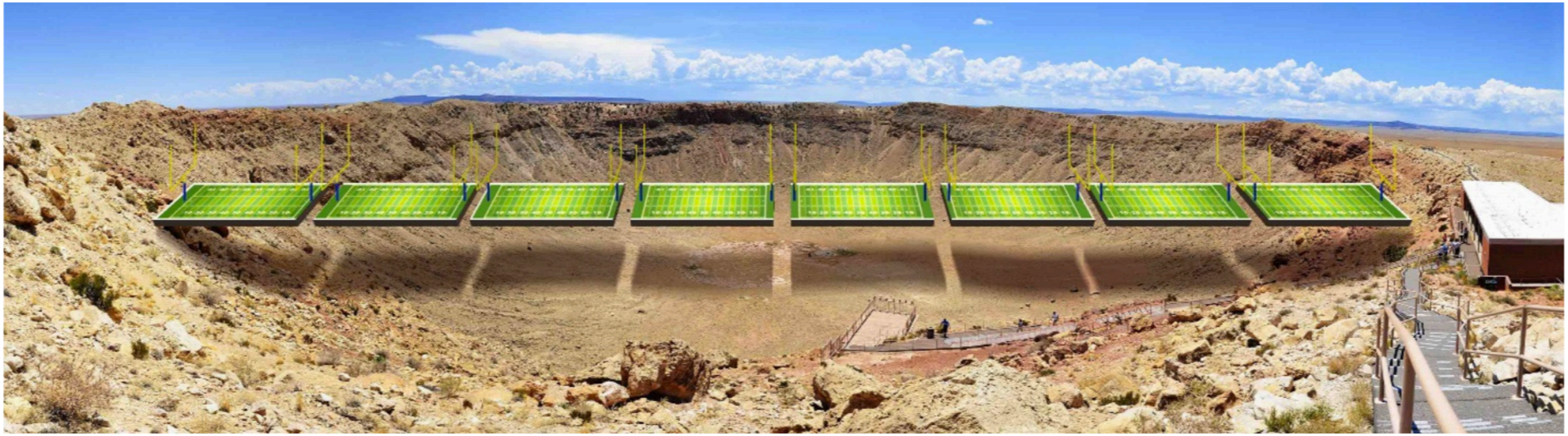


FIGURE S.1 An illustration showing Arizona's Meteor Crater with football fields superimposed to provide a sense of scale. This crater was created approximately 50,000 years ago by a nickel-iron asteroid estimated to have been 50 meters in diameter.

estimated to have been [illegible] meters in diameter.

TABLE 1.1 The Likely Consequences of an Asteroid Impact as a Function of Asteroid Size

Characteristic Diameter of Impacting Object	Approximate Average Impact Interval (years)	Estimated Number of Objects	Energy Released (megatons TNT)	Estimated Damage or Comparable Event
25-30 m	80-180	2.6-5.5 million	2	Fireball, airburst, shockwave, minor damage
50 m	1,500	>~310,000	10	Local damage comparable to that of largest existing thermonuclear weapon
140 m	20,000	~24,000	~500	Destruction on regional/ national scale
300-500 m	≥64,000-130,000	3,500-7200	≤10,000	Destruction on continental scale
1 km	520,000	~900	80,000	Global effects, many millions dead
10 km	120 million	4	80 million	Complete extinction of the human species

LSST will be sited in Central Chile - Cerro Pachon

LSST: a digital color movie of the Universe...

3.6×10^{-31} erg/s/cm^2/Hz
36 nJy

LSST in one sentence:

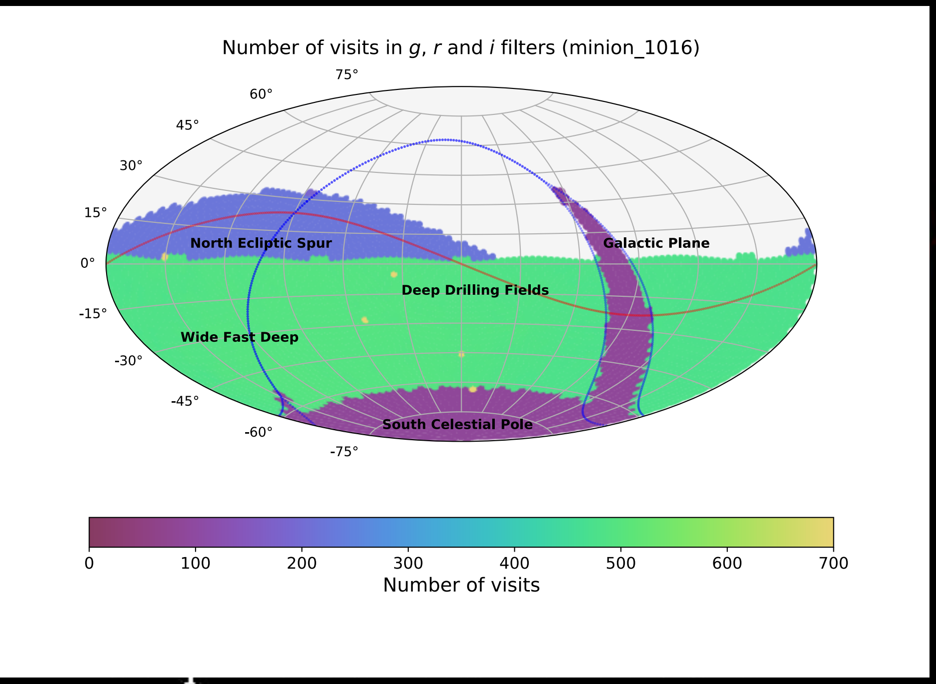
An optical/near-IR survey of half the sky in ugrizy bands to r~27.5 based on ~1000 visits over a 10-year period:

More information at www.lsst.org and arXiv:0805.2366

A catalog of 20 billion stars and 20 billion galaxies with exquisite photometry, astrometry and image quality!

Basic idea behind LSST: **a uniform sky survey**

- 90% of time will be spent on a uniform survey: every 3-4 nights, the whole observable sky will be scanned twice per night
- after 10 years, half of the sky will be imaged about 1000 times (in 6 bandpasses, ugrizy): a digital color movie of the sky
- ~100 PB of data: about a billion 16 Mpix images, enabling **measurements for 40 billion objects**



LSST in one sentence:
An optical/near-IR survey of half the sky in ugrizy bands to r~27.5 (36 nJy) based on 825 visits over a 10-year period: deep wide fast.

Left: a 10-year simulation of LSST survey: the number of visits in the r band (Aitoff projection of eq. coordinates)

SDSS
gri
3.5'x3.5'
r~22.5

3 arcmin is 1/10 of the full Moon's diameter

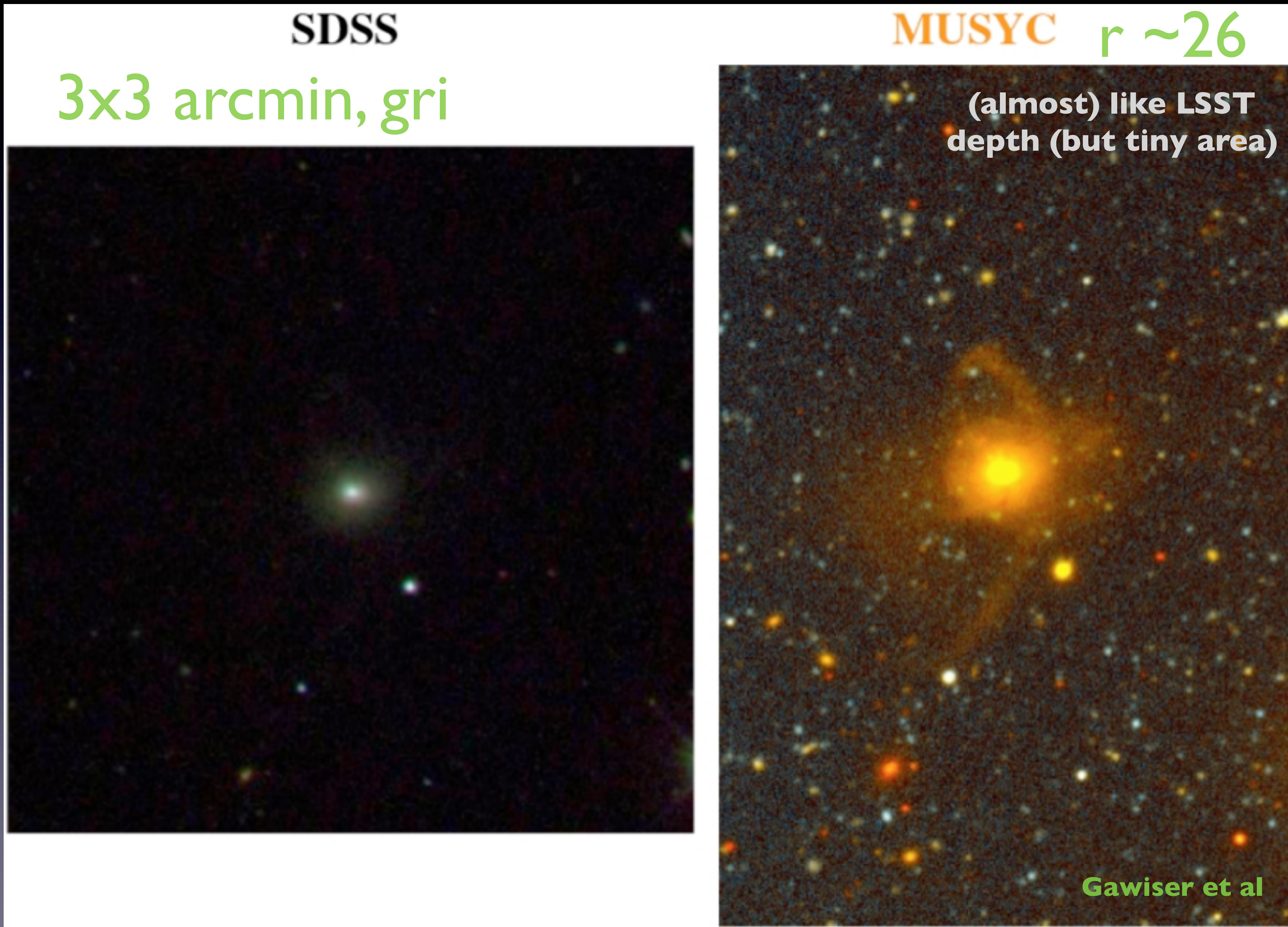
HSC
gri
3.5’x3.5’
r~27

3 arcmin is 1/10 of the full Moon’s diameter

like LSST depth (but tiny area)

LSST will deliver 5 million such images

Extragalactic astronomy: faint surface brightness limit



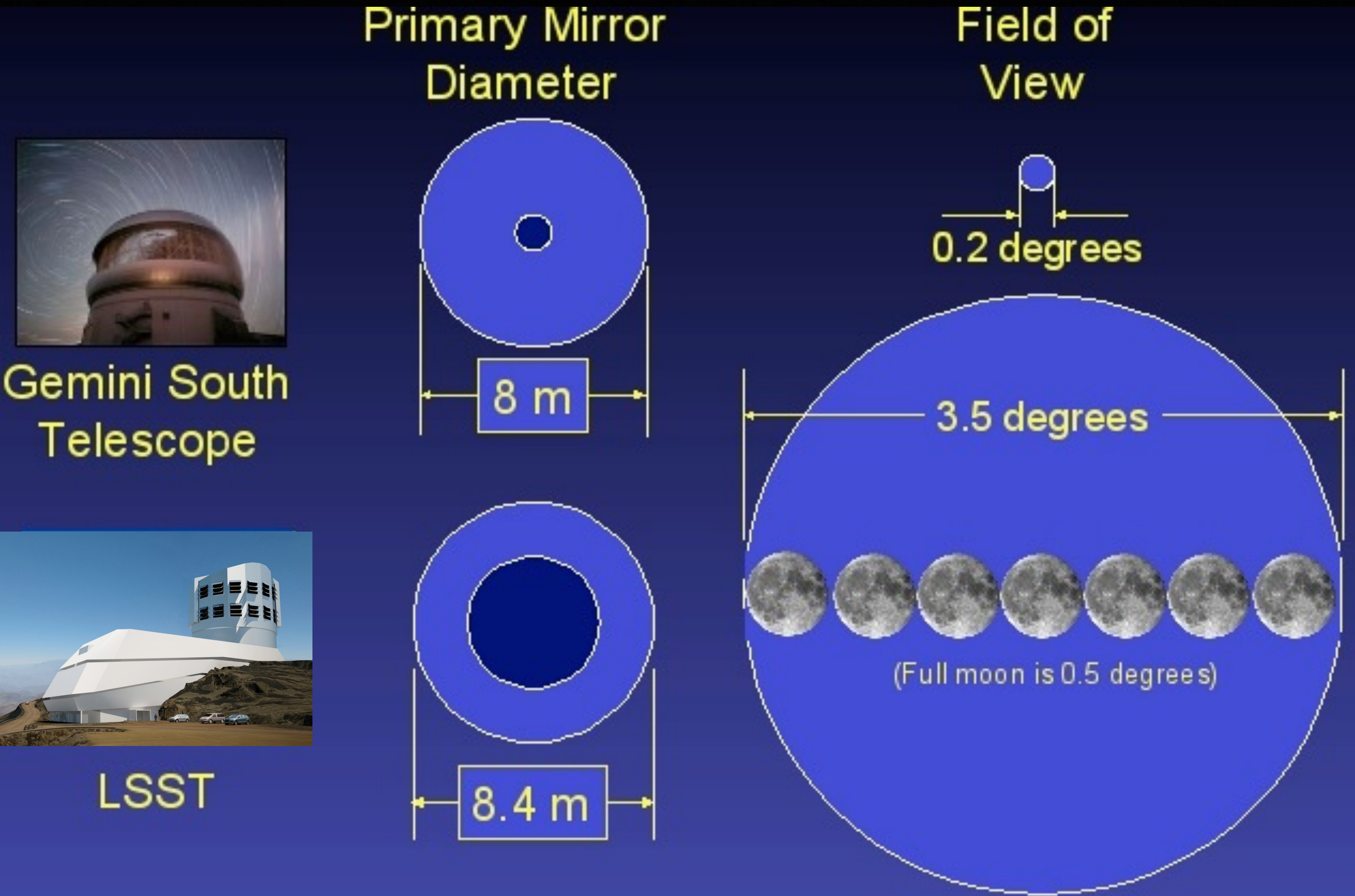
First light: 2021
Operations: 2022

8.4m, 6.7m effective

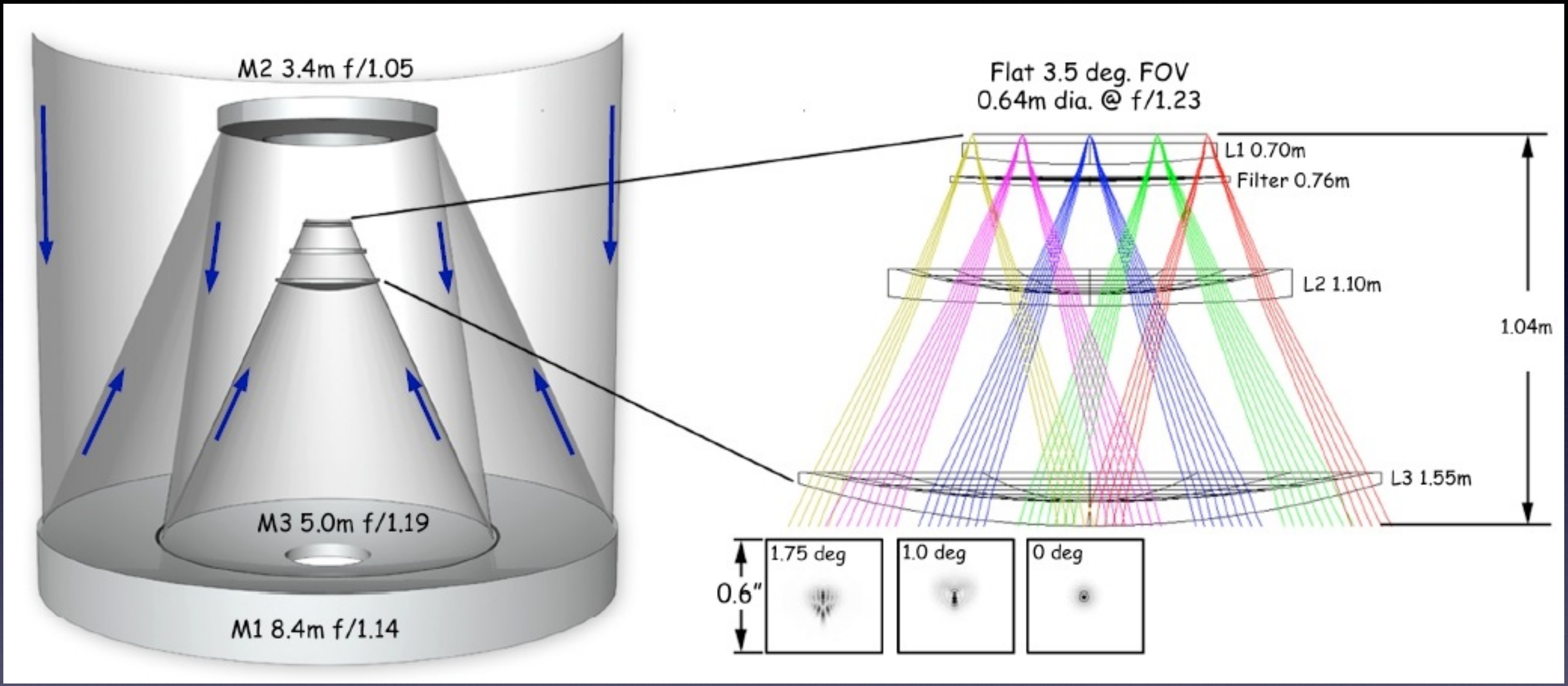
5 sec slew & settle

Telescope Mount Assembly before going from Spain to Chile

The field-of-view comparison: Gemini vs. LSST

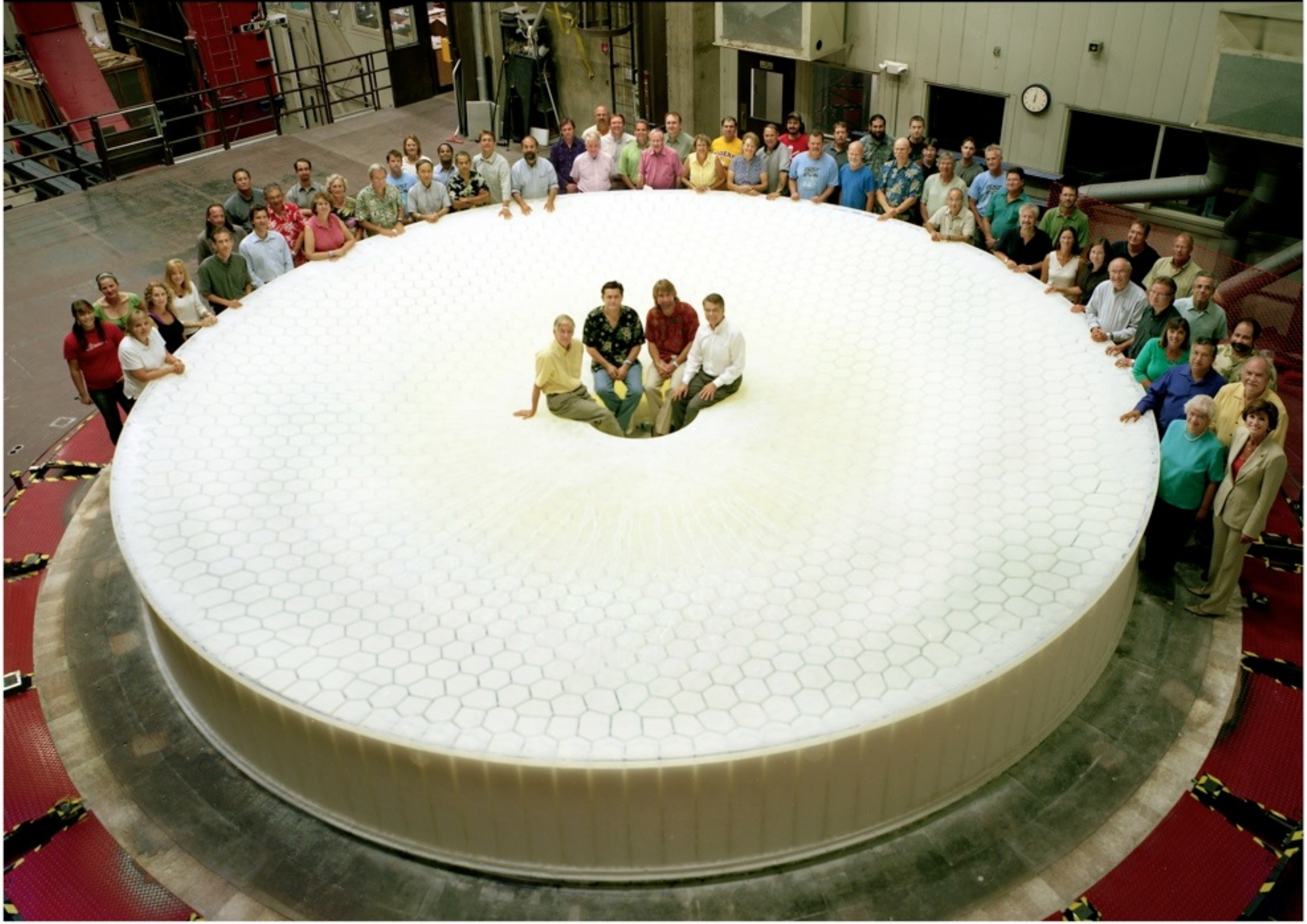


Optical Design for LSST



Three-mirror design (Paul-Baker system)
enables large field of view with excellent image quality:
delivered image quality is dominated by atmospheric seeing

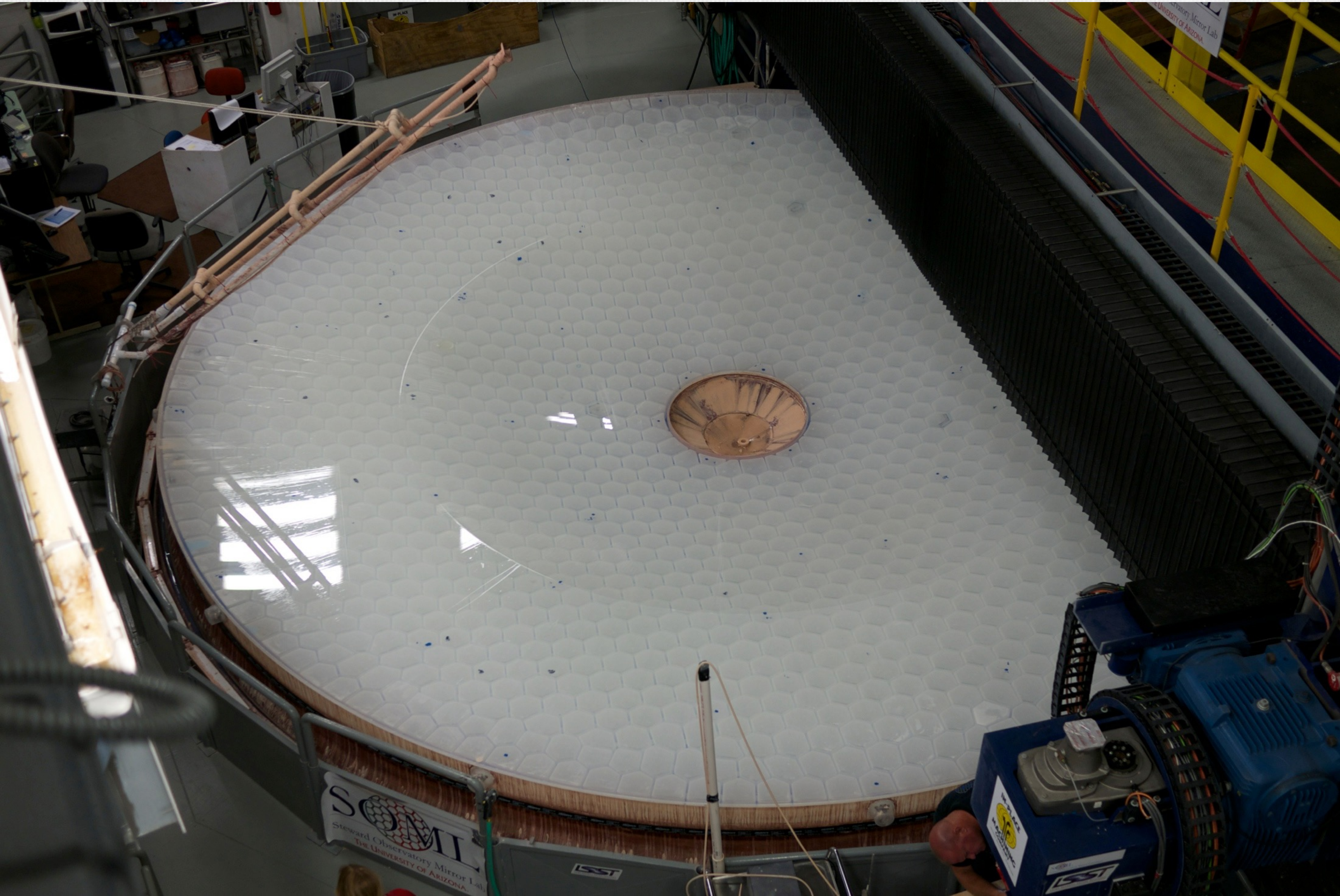
Large Synoptic Survey Telescope



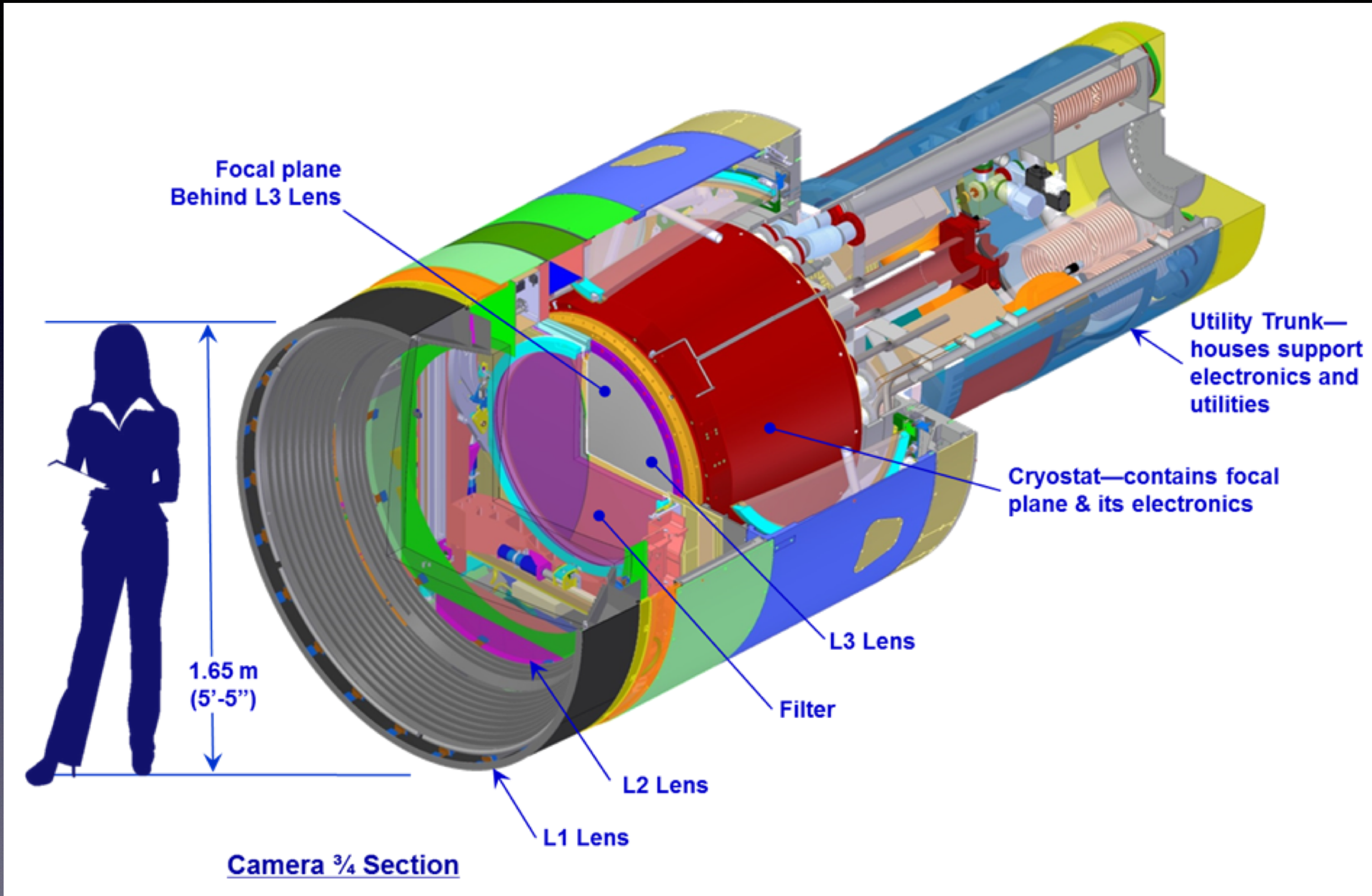
LSST Primary/Tertiary Mirror Blank
August 11, 2008, Steward Observatory Mirror Lab, Tucson, Arizona

The University of Arizona. Tucson Arizona SOML

Large Synoptic Survey Telescope

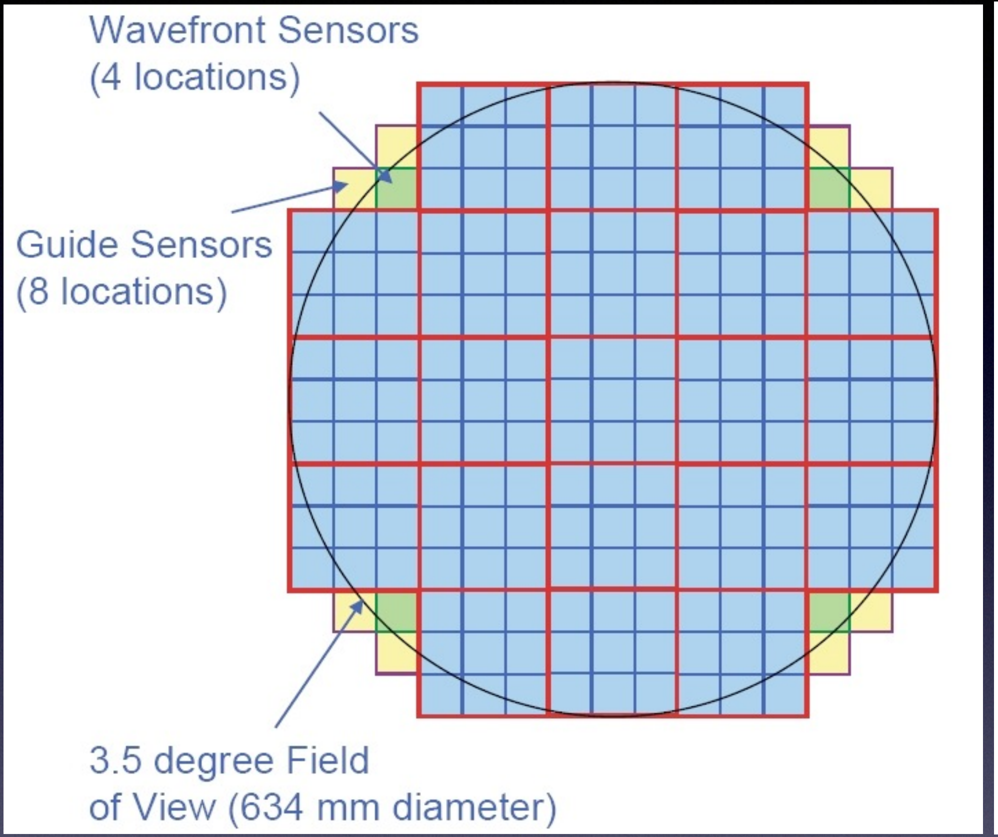

SOML
Steward Observatory Mirror Lab
THE UNIVERSITY OF ARIZONA

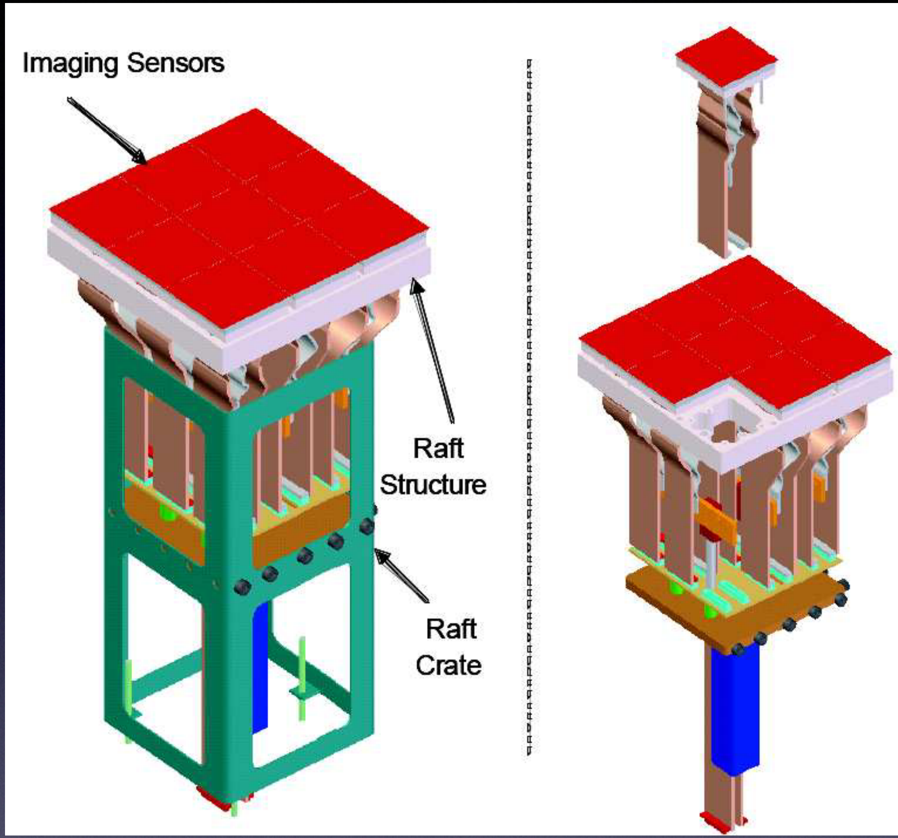
LSST camera



The largest astronomical camera: 2800 kg, 3200 Megapix

LSST camera





Modular design: 3200 Megapix = 189 x16 Megapix CCD
9 CCDs share electronics: raft (=camera)
Problematic rafts can be replaced relatively easily

LSST Science Sensor procurement (~200 CCDs) is complete!

Filter complement

- **Photometric redshifts for galaxies:** random errors smaller than 0.02, bias below 0.003, fewer than 10% >3σ outliers
- These photo-z requirements are one of the primary drivers for the photometric depth and accuracy of the main LSST survey (and the definition of filter complement)

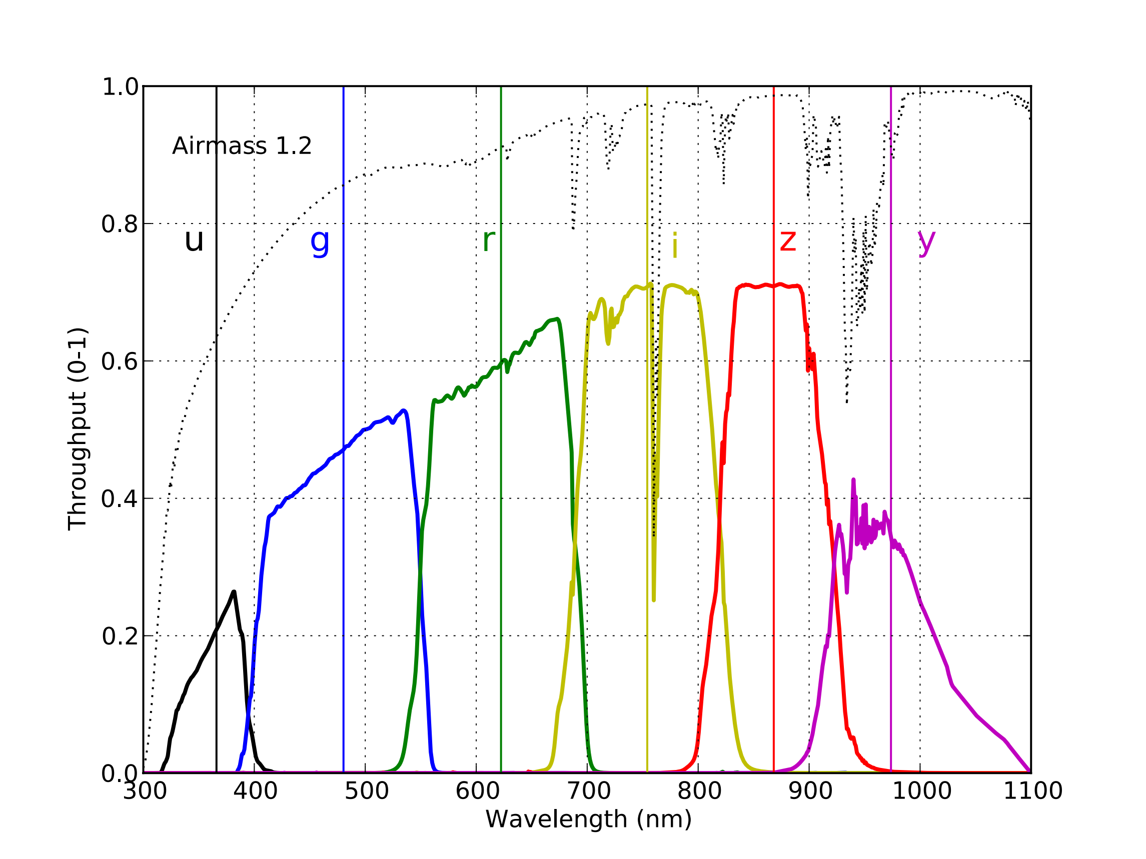


Photo-z requirements correspond to r~27.5 with the following per band time allocations:

u: 8%; g: 10%

r: 22%; i: 22%

z: 19%; y: 19%

Consistent with other science themes (stars)

Large Synoptic Survey Telescope
LSST Operations:
Sites & Data Flows
French Site
Satellite Processing Center
Data Release Production
Long-term Storage (copy 3)
Archive Site
Archive Center
Alert Production
Data Release Production
Calibration Products Production
EPO Infrastructure
Long-term Storage (copy 2)
Data Access Center
Data Access and User Services
HQ Site
Science Operations
Observatory Management
Education & Public Outreach
Base Site
Base Center
Long-term storage (copy 1)
Data Access Center
Data Access & User Services
Summit Site
Telescope & Camera
Data Acquisition
Crosstalk Correction
National Center for
Supercomputing Applications
United States
Mexico
Gulf of Mexico
Cuba
Dominican Republic
Puerto Rico
Guatemala
Honduras
Nicaragua
Costa Rica
Panama
Venezuela
Guyana
Suriname
French Guiana
Colombia
Ecuador
Brazil
Peru
Bolivia
Paraguay
Chile
La Serena
Cerro Pachón
Uruguay
Argentina
Google

At the highest level, LSST objectives are:

1) Obtain about 5.5 million images, with 189 CCDs (4k x 4k) in the focal plane; this is about **a billion 16 Megapixel images of the sky**

2) Calibrate these images (and provide other metadata)

3) Produce catalogs ("model parameters") of detected objects (37 billion)

4) **Serve** images, catalogs and all other metadata, that is, **LSST data products to LSST users**

The ultimate deliverable of LSST is not just the telescope, nor the camera, but **the fully reduced science-ready data** as well. **Software!**

LSST Data Management System ("software")

- 20 TB of data to process every day (~one SDSS/day)
- 1000 measurements for 40 billion objects during 10 years
- Existing tools and methods (e.g. SDSS) do not scale up to LSST data volume and rate (100 PB!)

- About 5-10 million lines of code (C++/python)

LSST data products are organized into three main categories:

Prompt Data Products

Real Time Difference Image Analysis (DIA)

- A stream of ~10 million time-domain events per night (Alerts), transmitted to event distribution networks within 60s of camera readout.
- Images, Object and Source catalogs derived from DIA, and an orbit catalog for ~6 million Solar System bodies within 24h.
- Enables discovery and rapid follow-up of time domain events

Data Release Data Products

Reduced single-epoch & deep co-added images, catalogs, reprocessed DIA products

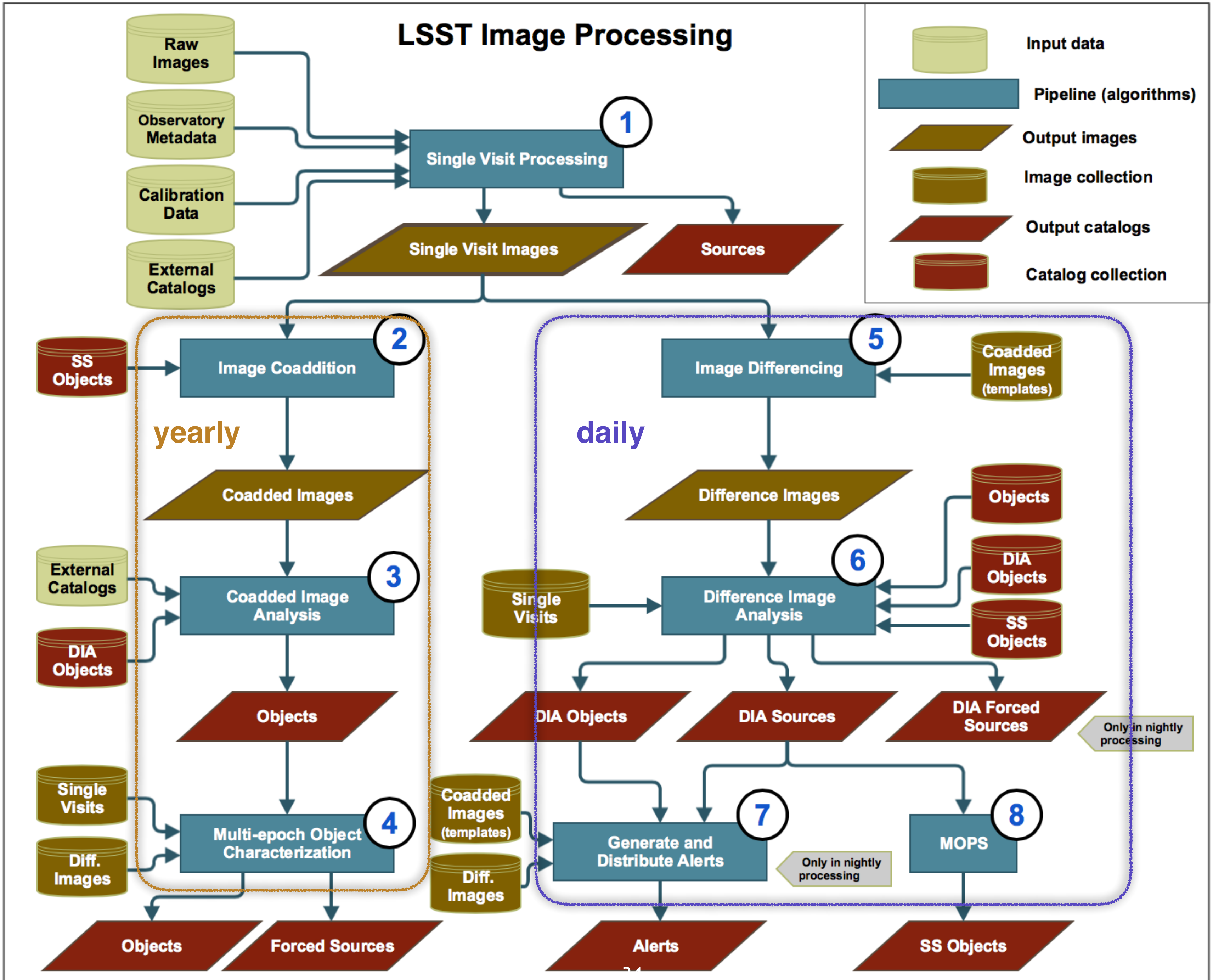
- Catalogs of ~37 billion objects (20 billion galaxies, 17 billion stars), ~7 trillion sources and ~30 trillion forced source measurements.
- 11 Data Releases, produced ~annually over 10 years of operation
- Accessible via the LSST Science Platform & LSST Data Access Centers.

User Generated Data Products

User-produced derived, added-value data products

- Deep KBO/NEO, variable star classifications, shear maps, etc …
- Enabled by services & computing resources at the LSST DACs and via the LSST Science Platform (LSP).
- 10% of LSST computing resources will be allocated for User Generated data product storage & processing.

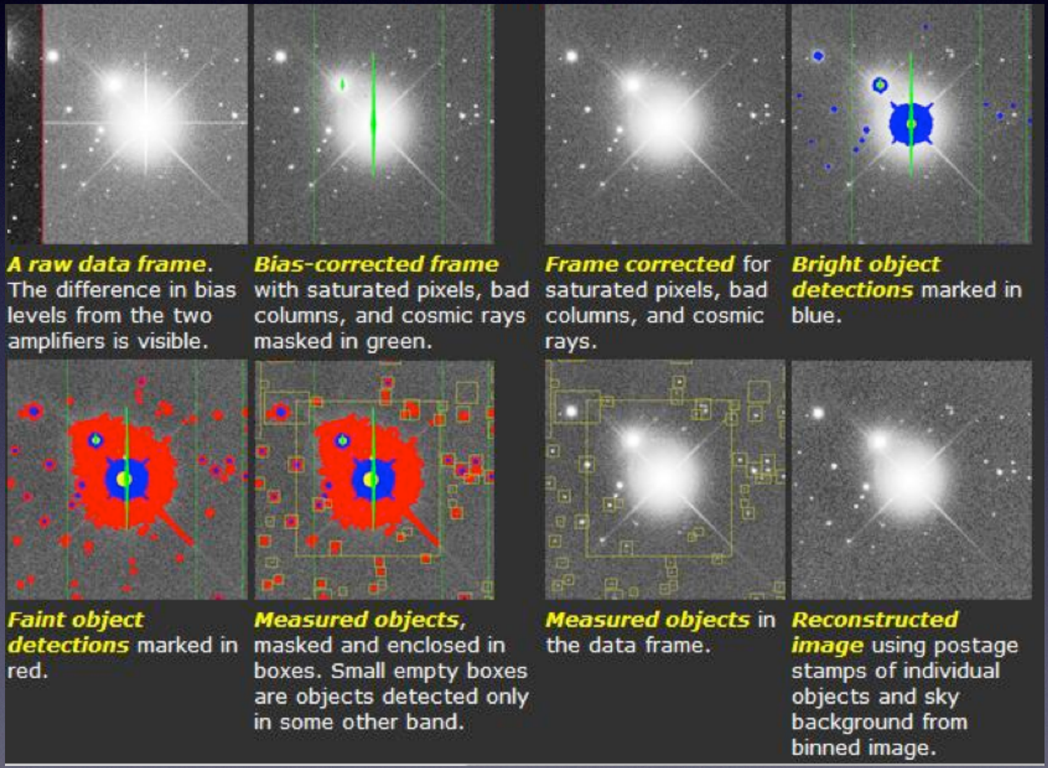
LSST Data Products: see http://ls.st/dpdd



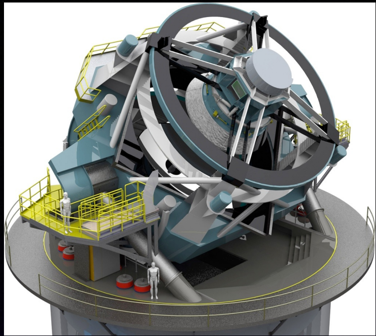
LSST Image Processing
Input data
Pipeline (algorithms)
Output images
Image collection
Output catalogs
Catalog collection
Raw Images
Observatory Metadata
Calibration Data
External Catalogs
1
Single Visit Processing
Single Visit Images
Sources
yearly
2
SS Objects
Image Coaddition
Coadded Images
3
External Catalogs
DIA Objects
Coadded Image Analysis
Objects
4
Single Visits
Diff. Images
Multi-epoch Object Characterization
Objects
Forced Sources
daily
5
Image Differencing
Coadded Images (templates)
Difference Images
6
Objects
DIA Objects
SS Objects
Single Visits
Difference Image Analysis
DIA Objects
DIA Sources
DIA Forced Sources
Only in nightly processing
Coadded Images (templates)
Diff. Images
7
Generate and Distribute Alerts
Only in nightly processing
8
MOPS
Alerts
SS Objects

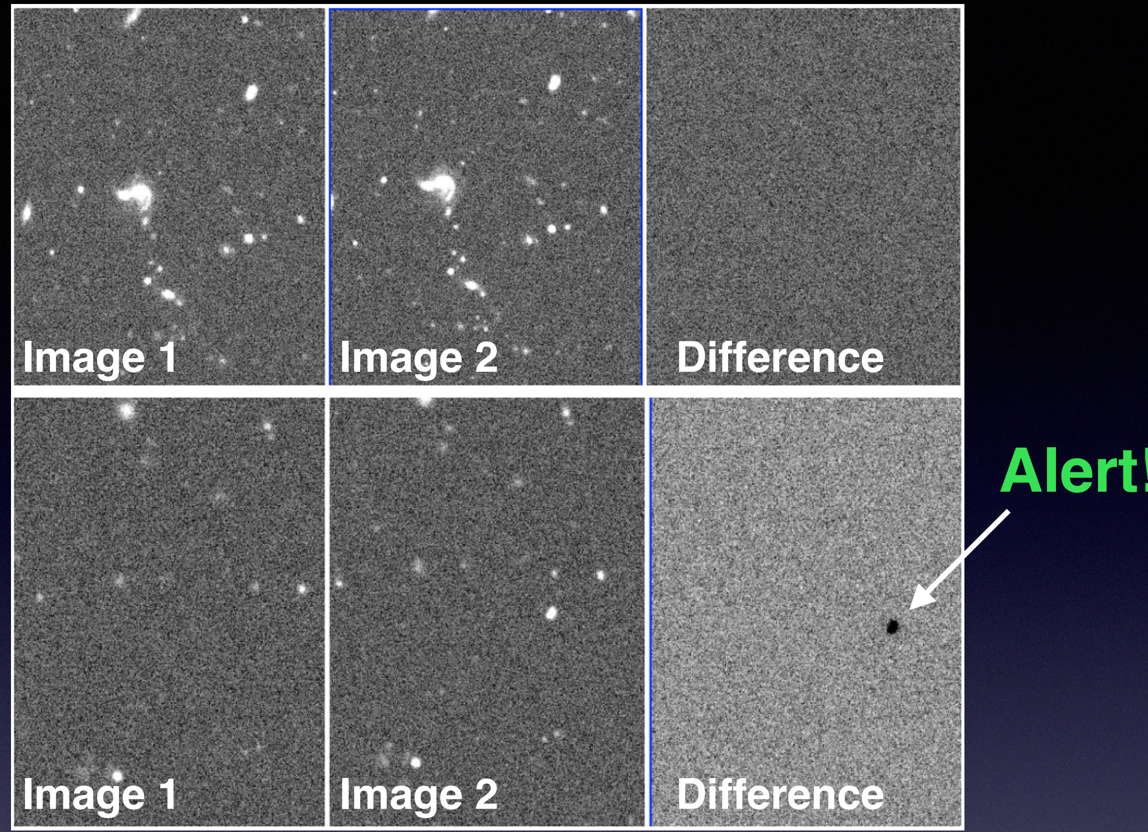
- Astronomical catalogs:
  - a list of all detected objects (stars, galaxies, ...)
  - measured parameters (size, color, brightness,...)

Basic steps in astronomical image processing (example: Sloan Digital Sky Survey):



All these (complicated) steps are already done: "science-ready database"





Additional "followup" data obtained to:

- confirmation and classification
- provide better temporal resolution
- use different filters/wavelengths
- obtain spectra (distance!)
- other measurements (e.g. polarimetry)

~10 billion alerts

Alerts can trigger "Followup" observations:

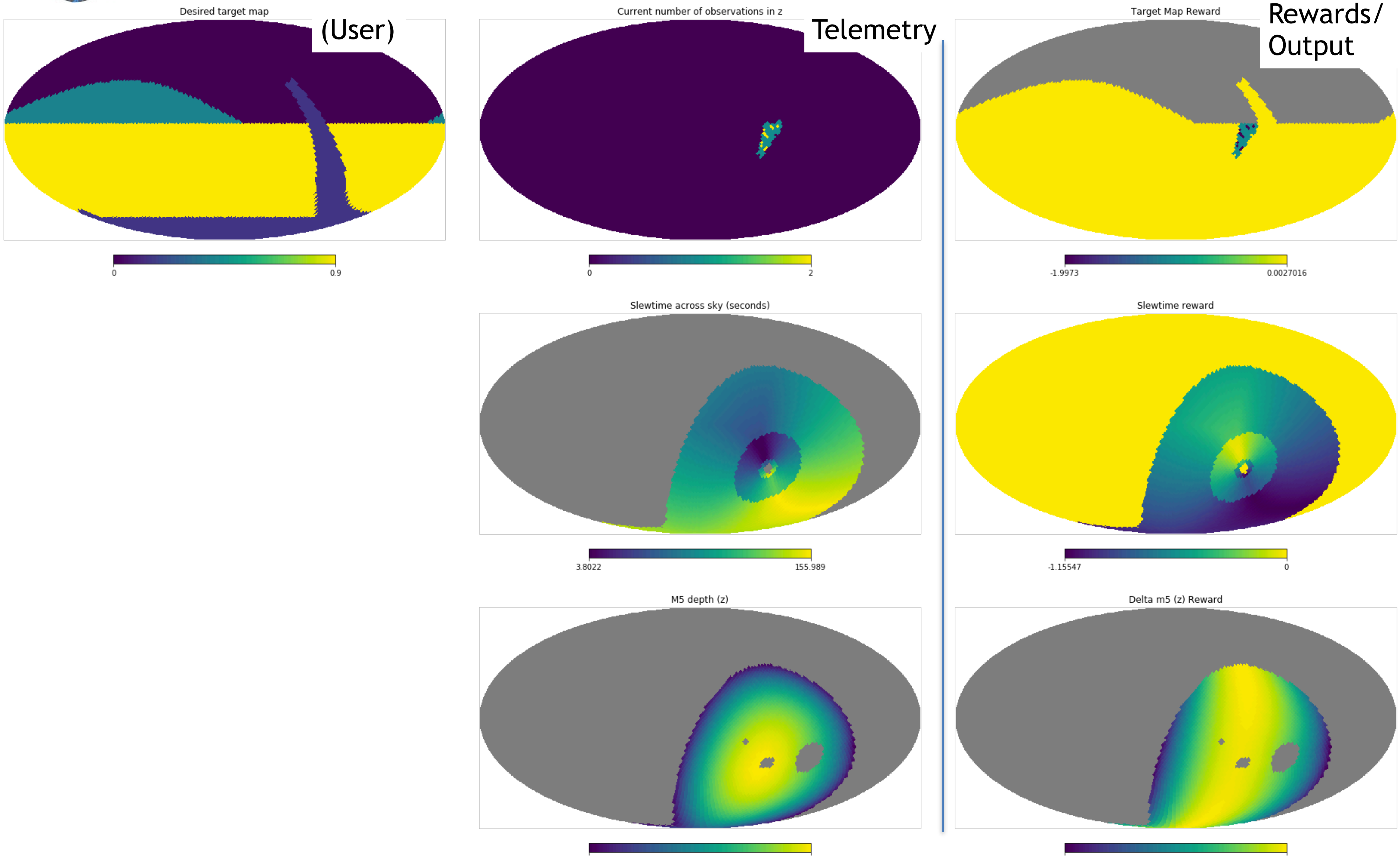
Automated scheduling of LSST observations

- LSST will have to make about 2.5 million decisions about where to point the telescope (2 sky coordinates), what filter to use and how long to keep the shutter open (exposure time); even with a fixed exposure time, each time there are about 100,000 options.

- optimal decision depends on observing conditions (sky brightness, including lunar contribution, atmospheric "seeing"), the system properties (imaging sensitivity as a function of filter and exposure time), as well as survey progress (number of images, their time sampling and achieved signal-to-noise ratio as functions of sky position and filter)

- too complicated for an astronomer to handle: instead, an algorithm will autonomously schedule observations by performing cost-benefit analysis. Astronomers will track the performance and modify this algorithm as needed (and sometimes override it for other reasons)

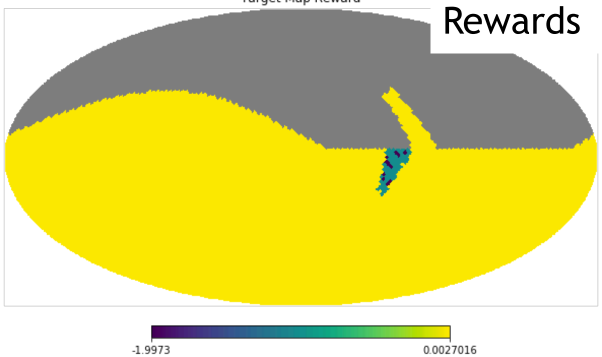
Examples of “Basis Functions”



Examples of “Basis Functions”

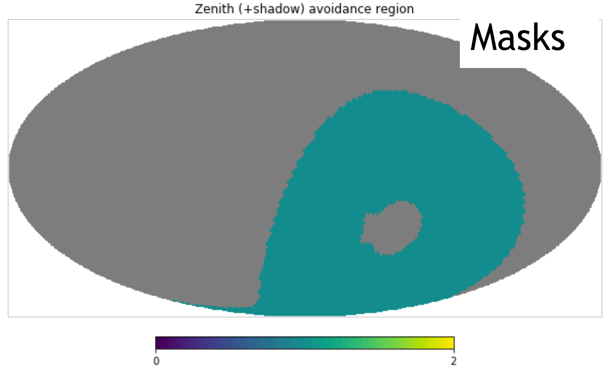
Rewards

Target Map Reward

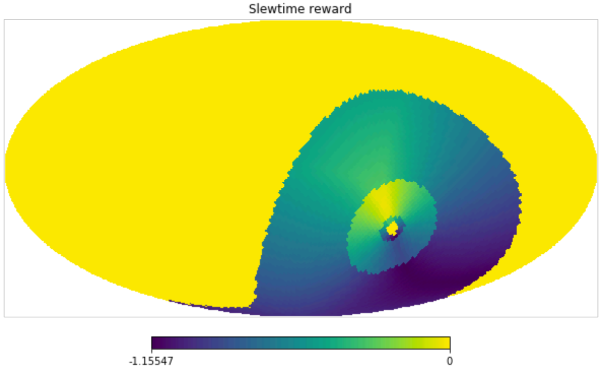


Masks

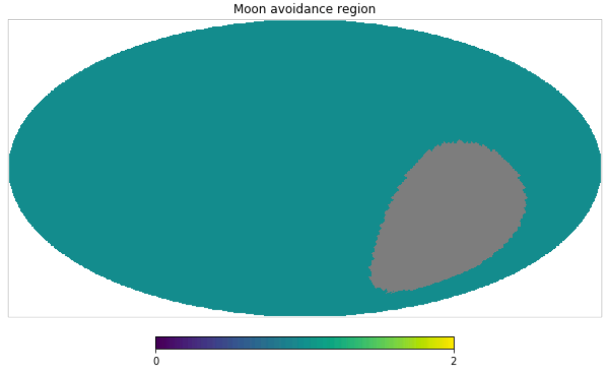
Zenith (+shadow) avoidance region



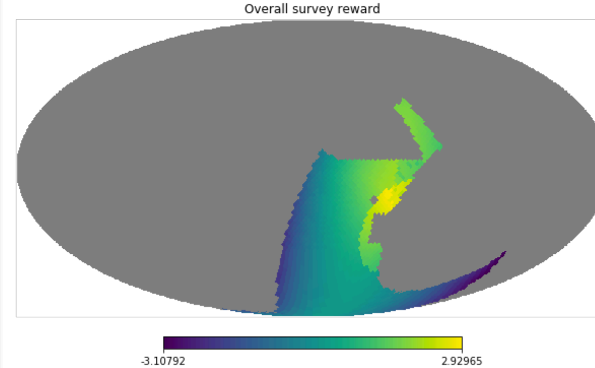
Slewtime reward



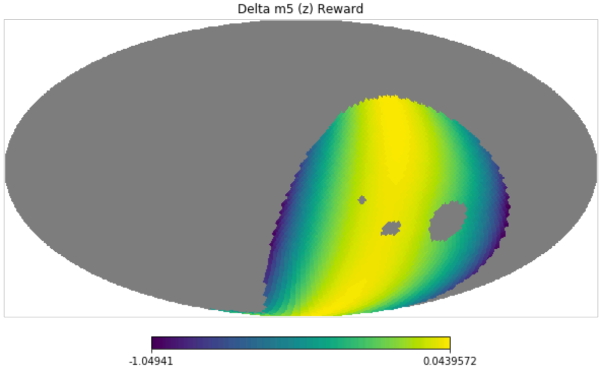
Moon avoidance region



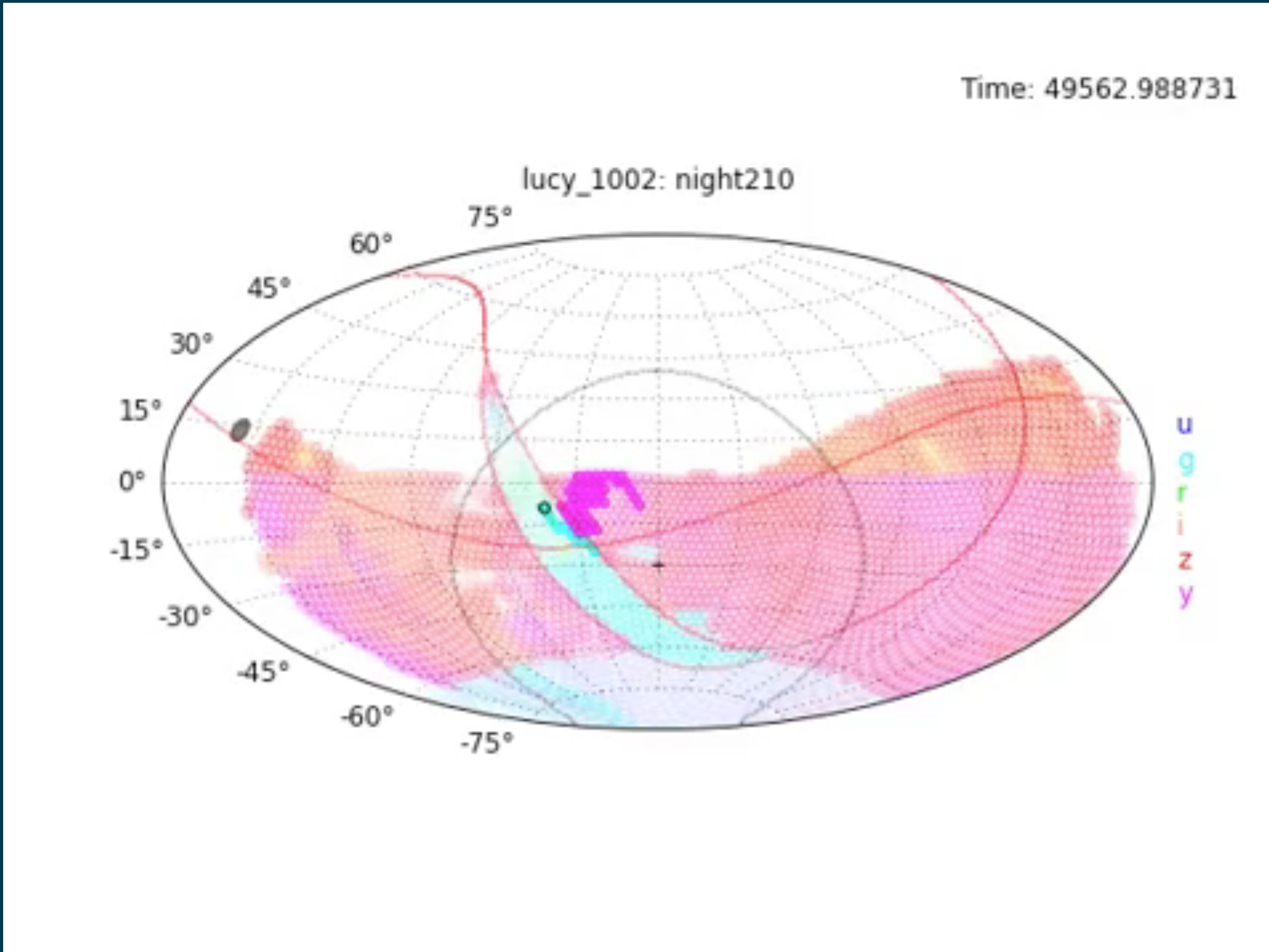
Overall survey reward



Delta m5 (z) Reward



Next pointing @ position with highest reward


Time: 49562.988731
lucy_1002: night210
75°
60°
45°
30°
15°
0°
-15°
-30°
-45°
-60°
-75°
u
g
r
i
z
y

Outline

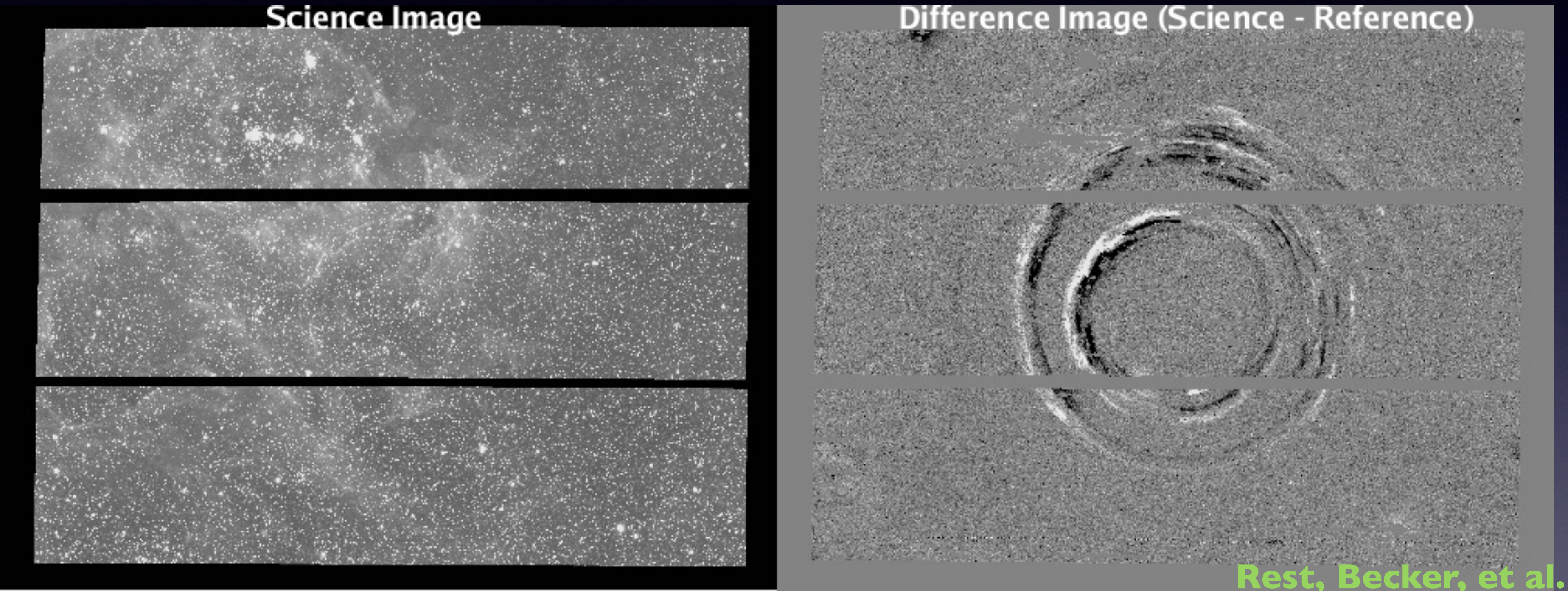
- LSST overview and construction status report
  - multi-color time-resolved faint sky map
  - 20 billion stars and 20 billion galaxies
- **A tour of anticipated LSST science programs**
  - time domain
  - the Solar System structure
  - cosmology (dark matter and dark energy)
  - the Milky Way structure
- Data analysis challenges ahead of us
  - large data sets
  - complex analysis
  - aiming for small systematics

Time Domain: objects changing in time

positions: asteroids and stellar proper motions

brightness: cosmic explosions and variable stars

Not only point sources - echo of a supernova explosion:



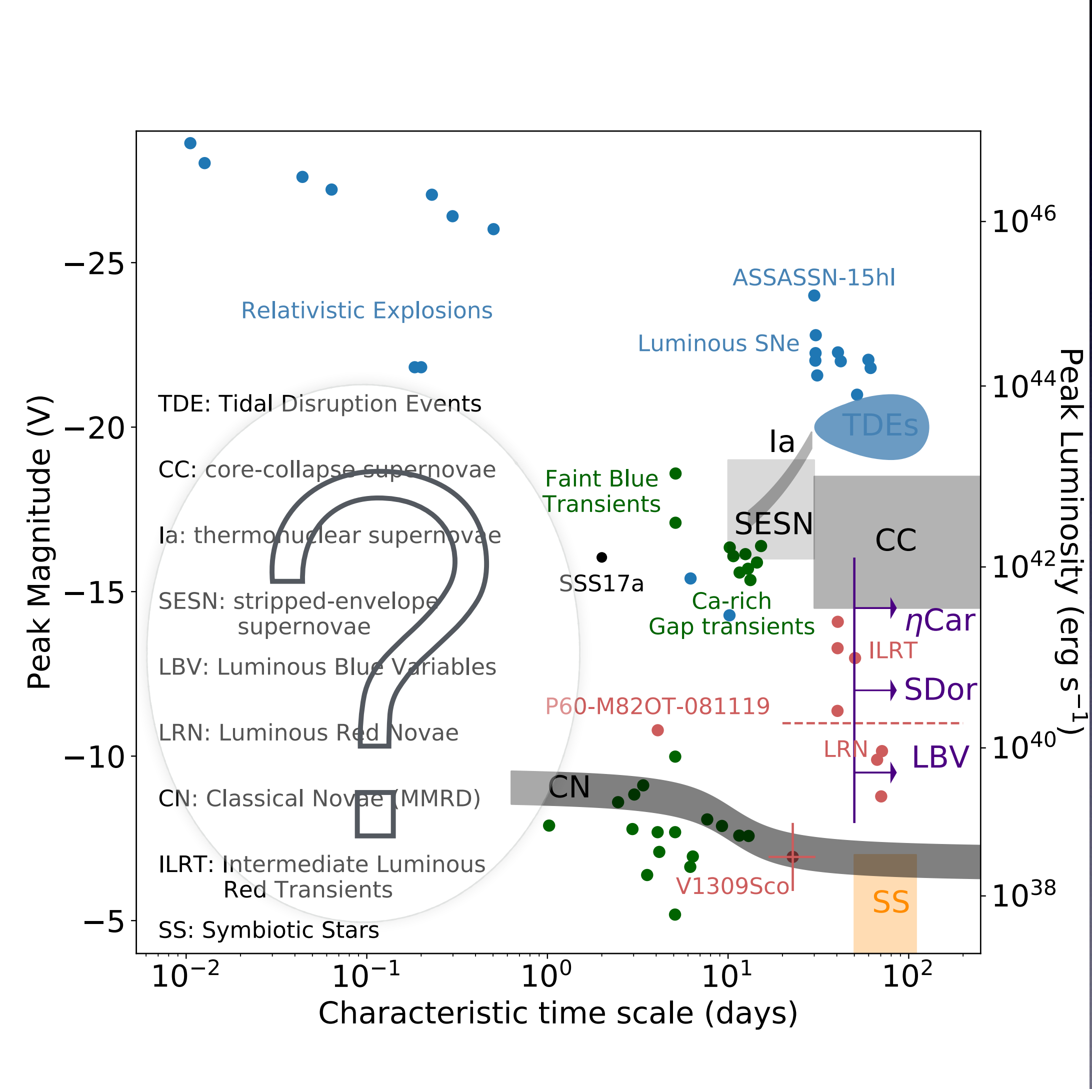
As many variable stars from LSST, as all stars from SDSS

Web stream with data for transients within 60 seconds.

Time Domain: objects changing in time

positions: asteroids and stellar proper motions

brightness: cosmic explosions and variable stars

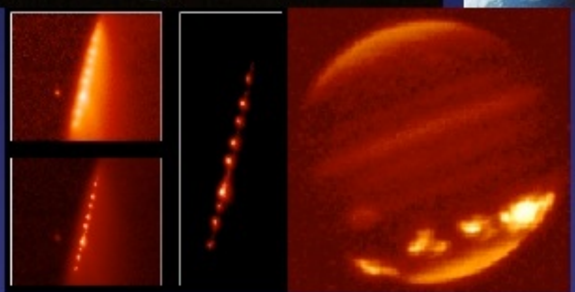


LSST will extend time-volume space a hundred times over current surveys (new classes of object?): **multi-messenger astrophysics**

known unknowns
unknown unknowns

Killer asteroids: the impact probability is not 0!

photomontage!



Shoemaker-Levy 9
(1994)

Tunguska
(1908)

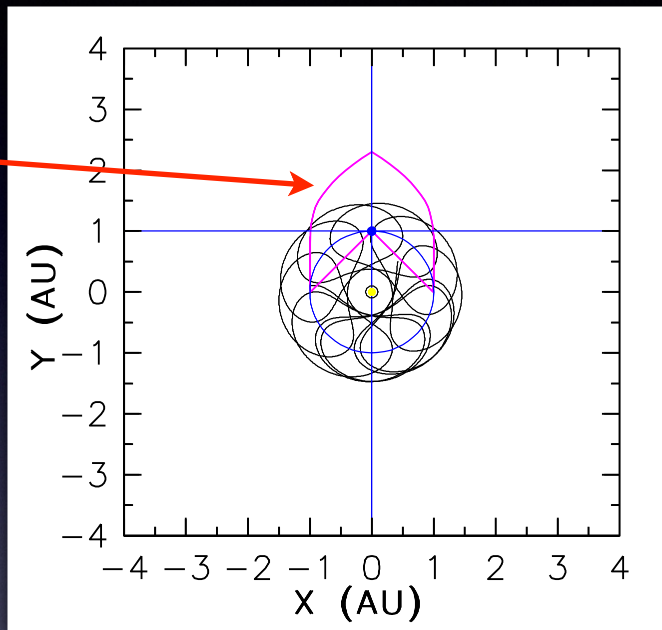
Asteroids larger than 140m collide with Earth every 20,000 years on average. Typical impact energy of such a collision is 500 Megaton TNT (50x largest bomb)

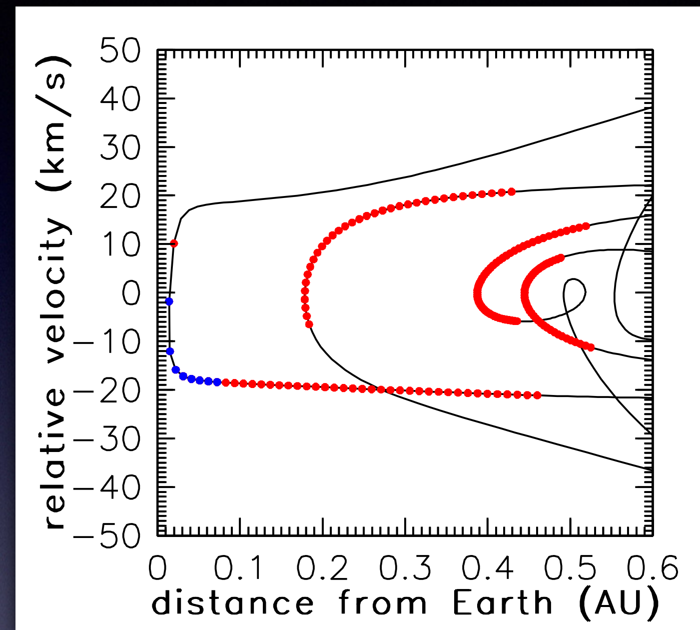
LSST is the only survey capable of delivering completeness specified in the 2005 USA Congressional NEO mandate to NASA (to find 90% NEOs larger than 140m)

The Barringer Crater, Arizona: about 40m object 50,000 yr. ago

NEO Impact Warning Time

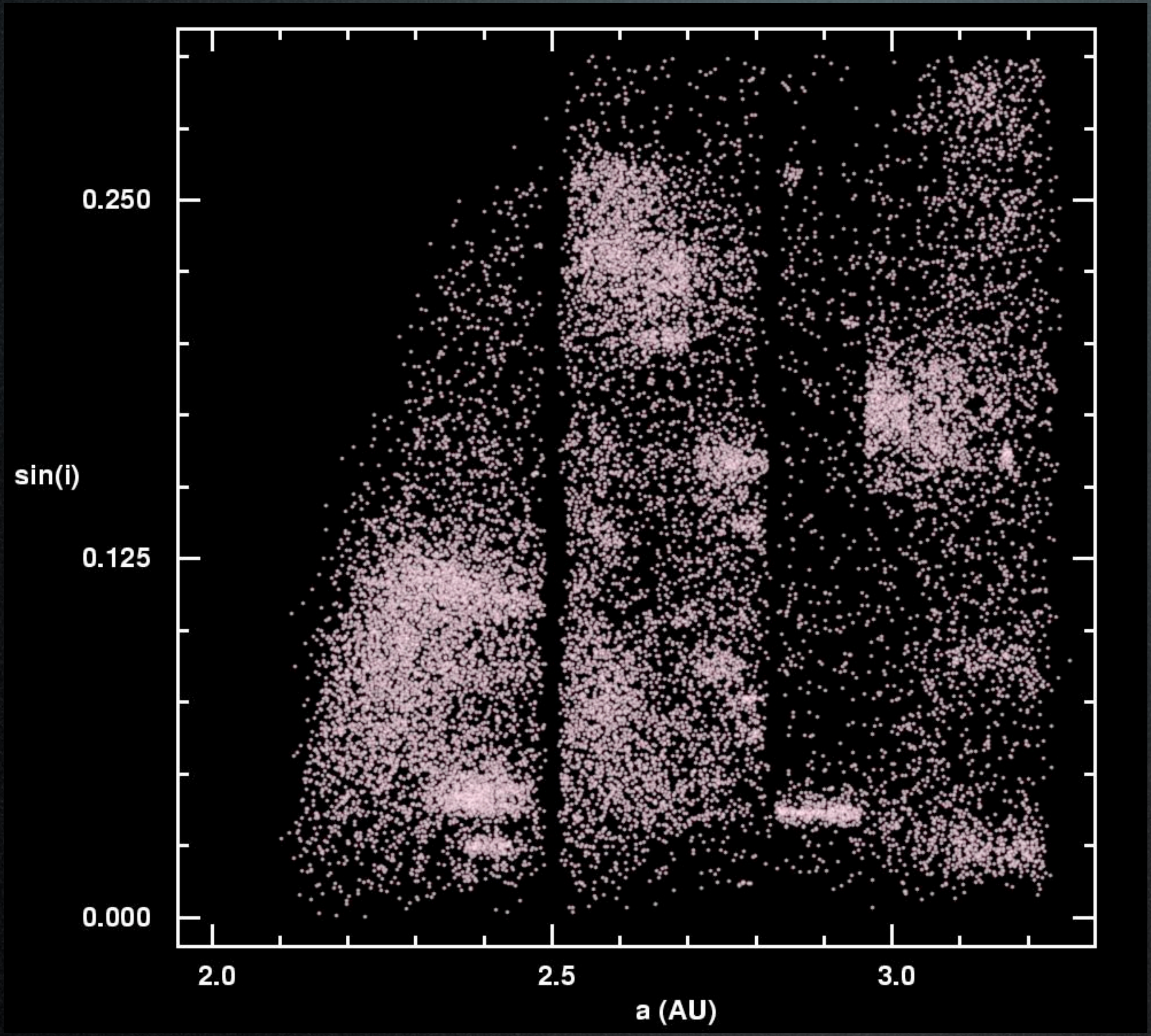
Detection volume for 140m objects





- **For 45m objects, LSST's warning time would be between 1-3 months depending on the orbit.**
- This model PHA is detected 39 days before 'impact' (vs. 5 days for systems with V<20), as well as during its 3 prior close approaches.
- Red dots indicate where the object could be detected by LSST (r=24.5) vs Blue dots, where V<20.

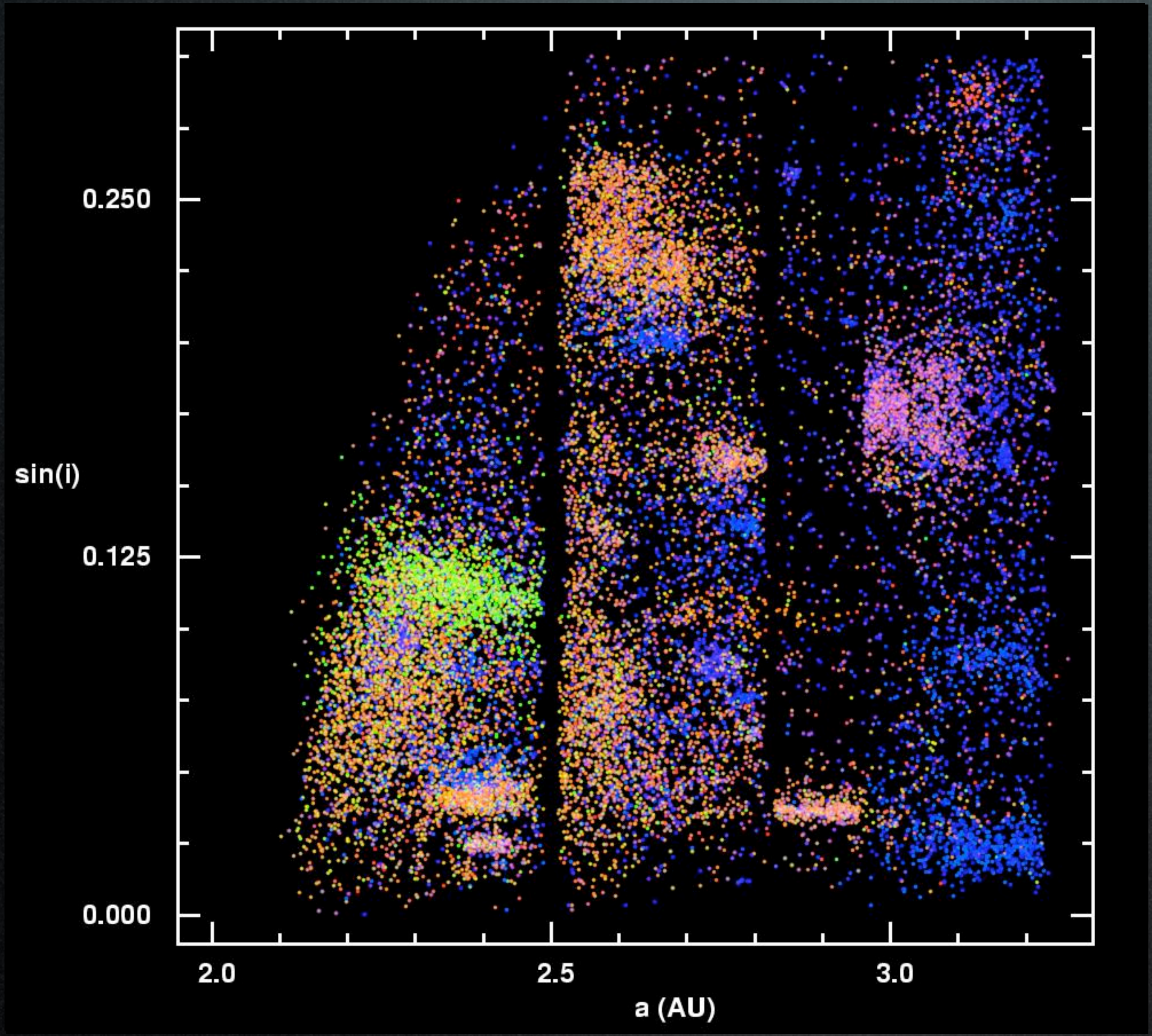
Main-belt Inventory



30,000
Asteroids with
SDSS colors and
proper
orbital elements

(Ivezic, Juric, Lupton 2002)

Main-belt Inventory



30,000
Asteroids with
SDSS colors and
proper
orbital elements

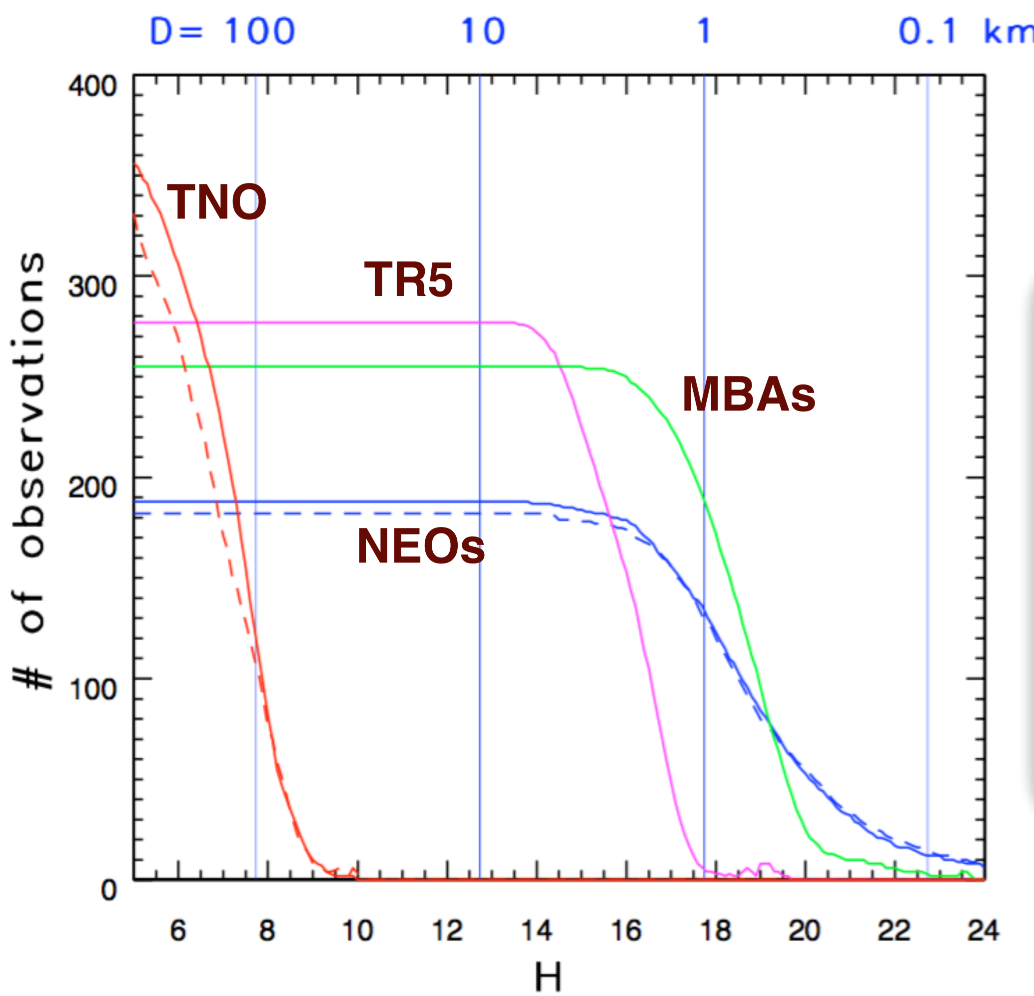
(Ivezic, Juric, Lupton 2002)

Color-coded with
SDSS colors

Colors help with the definition of asteroid families.
LSST will also provide color light curves!

From Chapter 5 in the LSST Science Book: Solar System science

- **MBAs: 5.5 million**
- **TR5: 280,000**
- **NEAs: 100,000**
- **TNOs: 40,000**

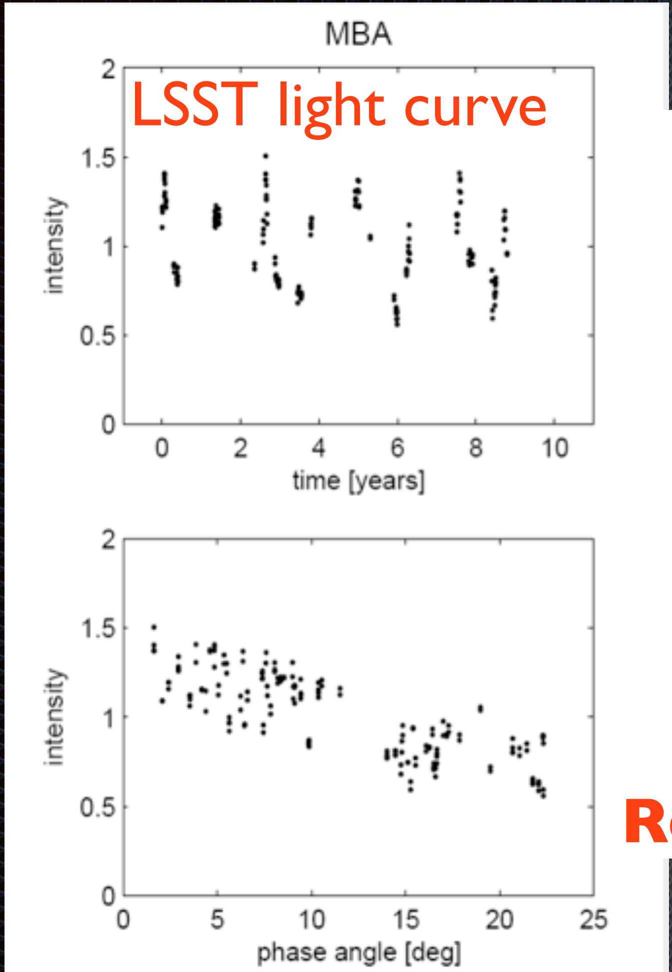


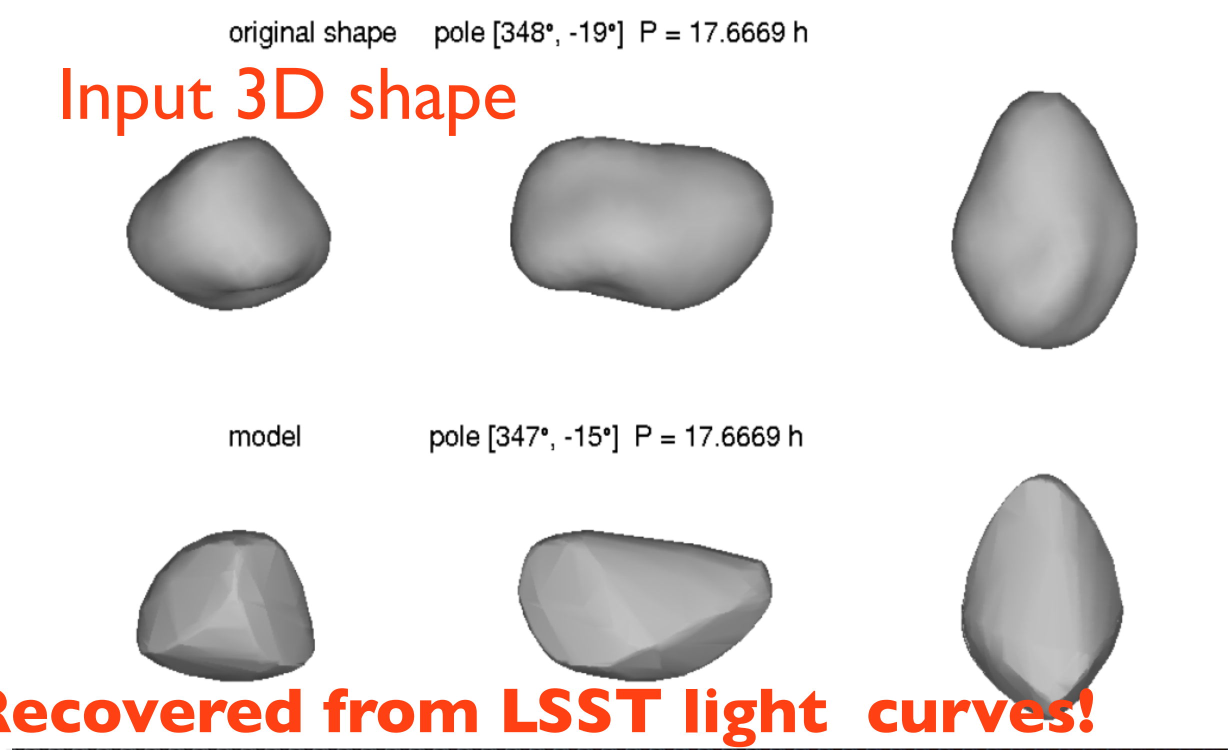
It's not just the number of newly discovered asteroids (5-6 million) but LSST observations will be unprecendented: hundreds of detections per object!

Figure 5.4: The median number of expected LSST detections of a given object as a function of H for dominant populations of Solar System bodies. Solid lines correspond to classical TNOs (red), Jovian Trojans (magenta), MBAs (green), and NEAs (blue). The red dashed line corresponds to Scattered Disk Objects, and the blue dashed line to PHAs. Nights with only one detection are not counted.

Sparse Lightcurve Inversion to determine asteroid shapes

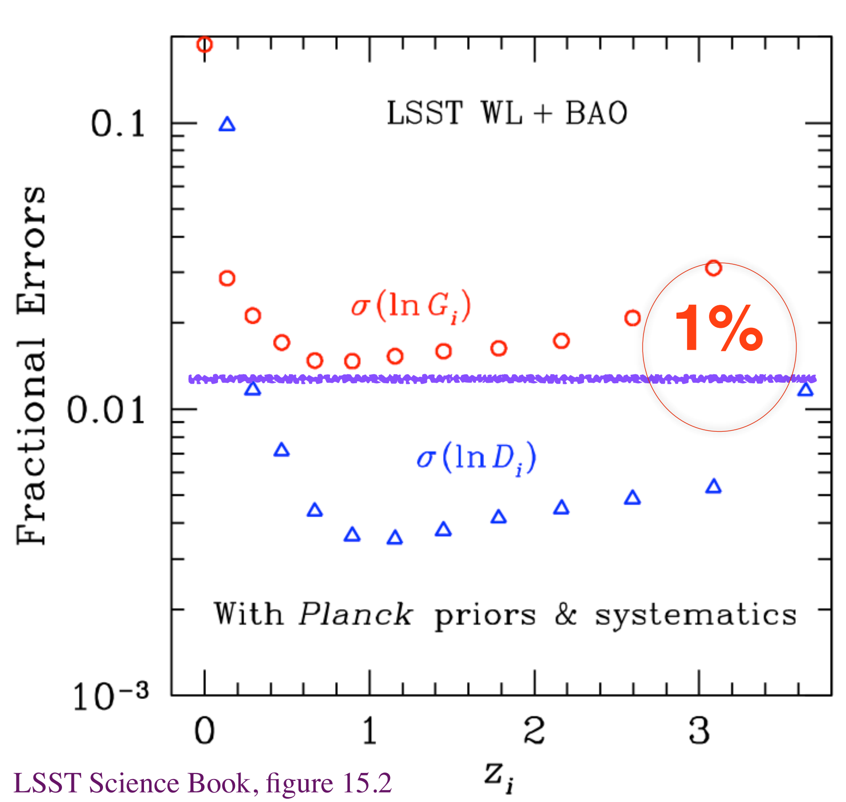
Durech et al. (2007)





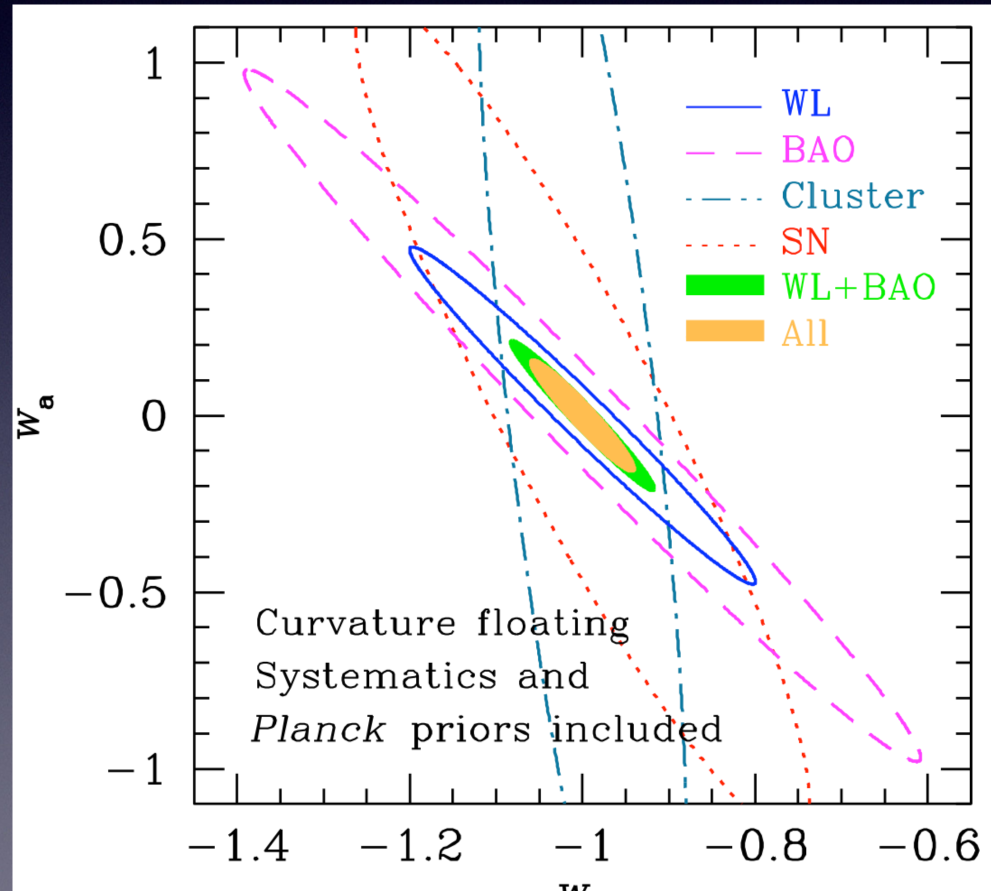
About 100 observations over >5 years are needed

Cosmology with LSST: high precision measurements



LSST Science Book, figure 15.2

- Measuring distances, H(z), and growth of structure, G(z), with a percent accuracy for 0.5 < z < 3
- Multiple probes is the key!



LSST Science Book, figure 15.3

By simultaneously measuring growth of structure and curvature, LSST data will tell us whether the recent acceleration is due to **dark energy or modified gravity.**

Cosmology with LSST: dark energy vs. modified gravity

- Even for a model with modified gravity, it is possible to assume that GR is correct and always find DE with suitable w(z) to explain data for H(z). 😱

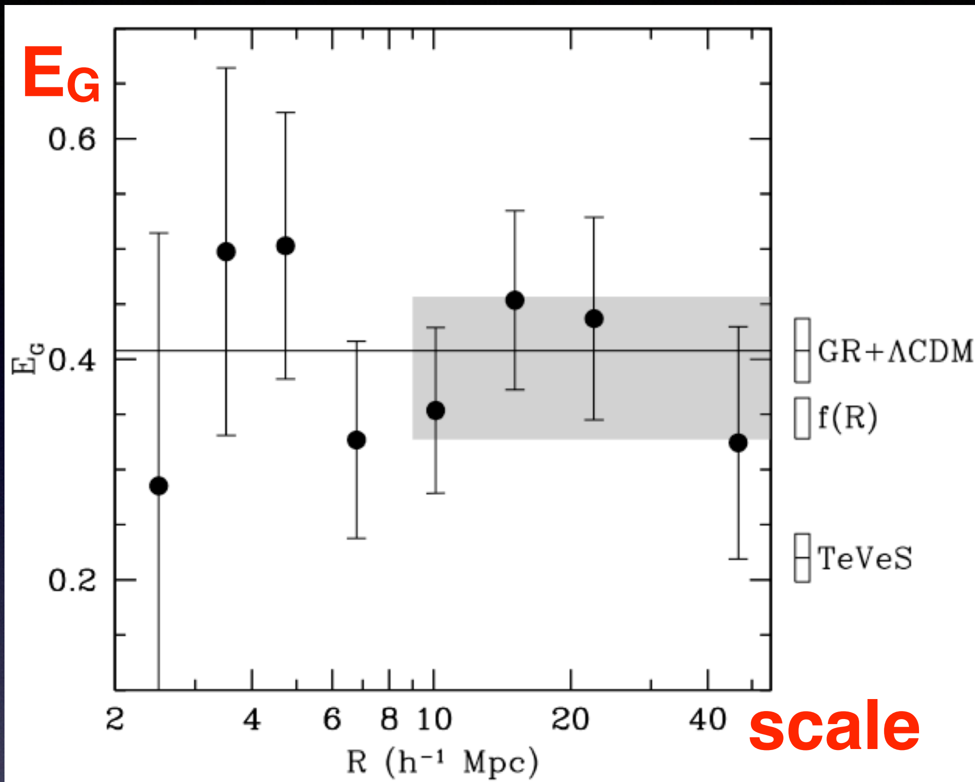
- However, the growth of structure will be different and thus when both H(z) and G(z) are measured, the degeneracy can be broken and DE vs. modified gravity models distinguished (Jain & Zhu 2008, PhysRevD 78, 063503)

$$ds^2 = -(1+2\psi)\ dt^2 + (1-2\phi)\ a^2(t)\ d\vec{x}^2$$

- φ is the curvature perturbation and ψ is the potential pertur.

- In General Relativity φ = ψ in the absence of anisotropic stresses. A metric theory of gravity relates the two potentials above to the perturbed energy-momentum tensor. 😄
  φ and ψ can be constrained with astronomical observations.

Cosmology with LSST: dark energy vs. modified gravity

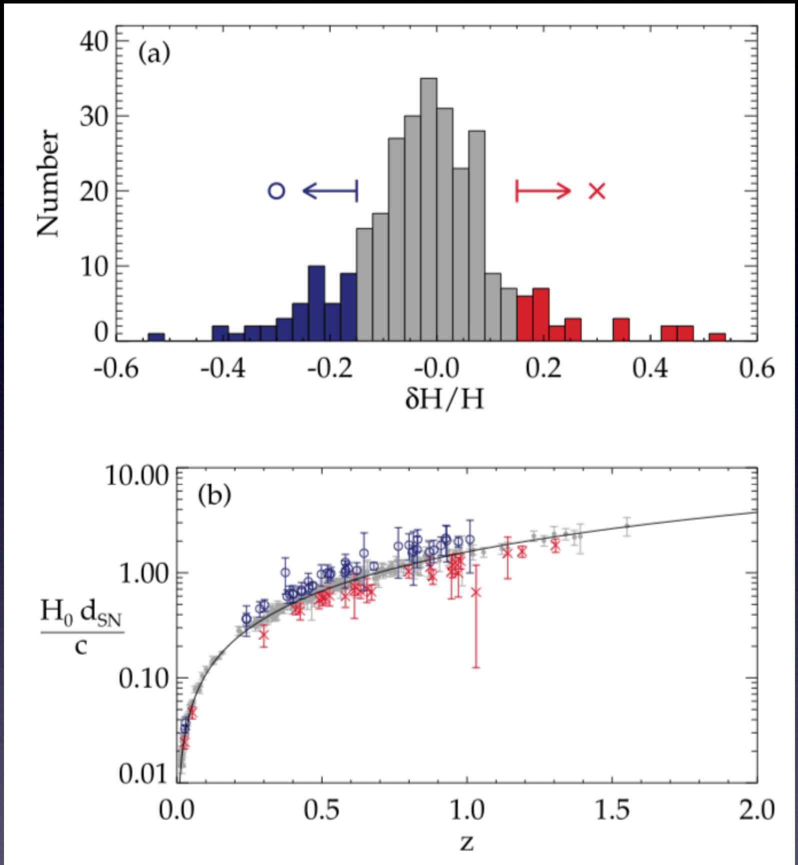


Reyes et al. (2010, Nature 464, 256)

- **E_G** combines 3 measures of large-scale structure: galaxy-galaxy lensing ($\varphi+\psi$), galaxy clustering (φ) and galaxy velocities (from galaxy redshifts; measures G(z))
- SDSS data enabled a test of GR at 15% level: it passed!
- SDSS data already excludes a model within the tensor-vector-scalar gravity theory, which modifies both Newtonian and Einstein gravity.
- Five times better precision needed to rule out f(R)

LSST will measure E_G about 10 times more precisely and will be able to rule out a large class of modified gravity theories (or GR!)

Cosmology with LSST SNe: is the cosmic acceleration the same in all directions?



Cooke & Lynden-Bell (2009, MNRAS 401, 1409)

Is there spatial structure in the SNe distance modulus residuals for the concordance model?

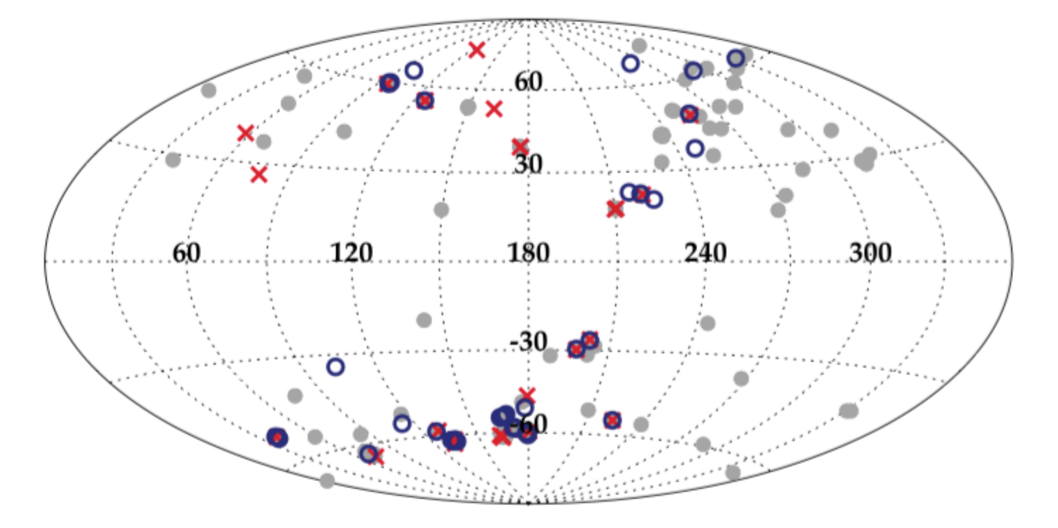


Figure 1. A projection of the spatial distribution of the Union SNe Ia sample in Galactic coordinates. Note the relative uniformity of the points, except around the Galactic plane. The symbols correspond to those in Fig. 2, and are explained in Section 3.1.

- Even a single supernova represents a cosmological measurement!
- LSST will obtain light curves for several million Type Ia supernovae!

LSST, WFIRST and Euclid are highly complementary missions.

WFIRST:

2,200 sq.deg

$m_5 \sim 27$

$r_{1/2} \sim 0.12$ arcsec

~2025-2031

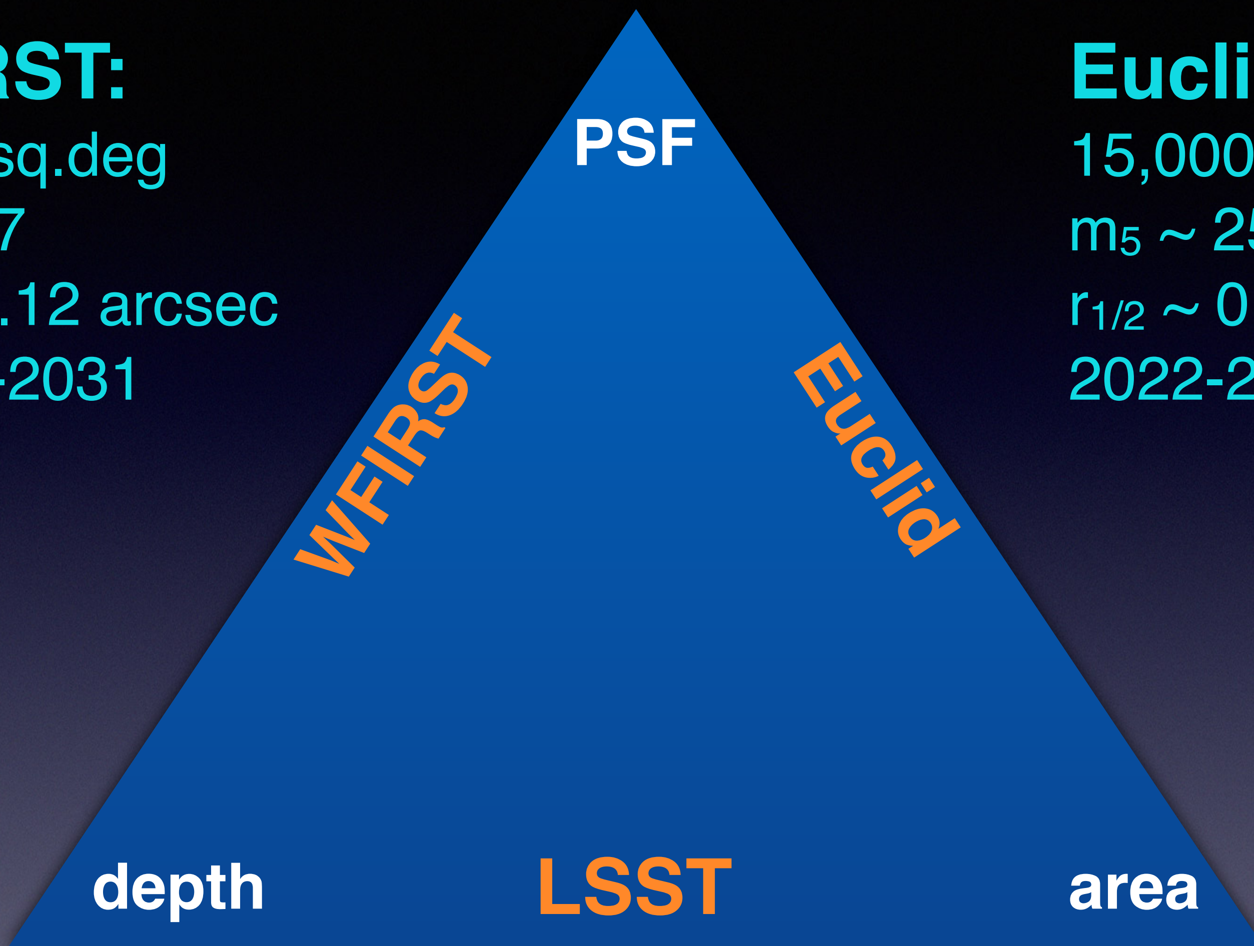
Euclid:

15,000 sq.deg

$m_5 \sim 25$

$r_{1/2} \sim 0.13$ arcsec

2022-2028



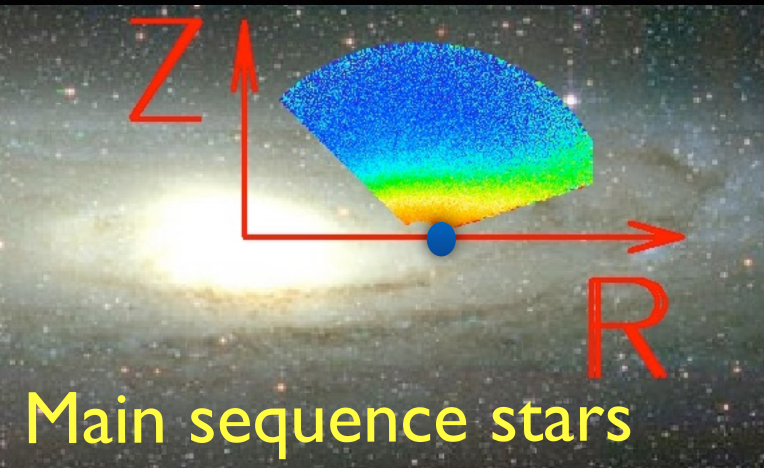
18,000 sq.deg

$m_5 \sim 27$

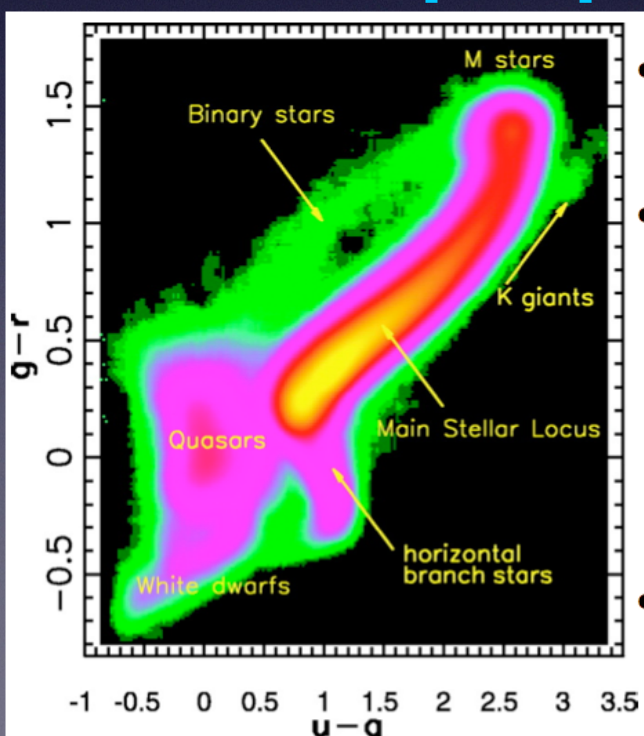
$r_{1/2} \sim 0.4$ arcsec

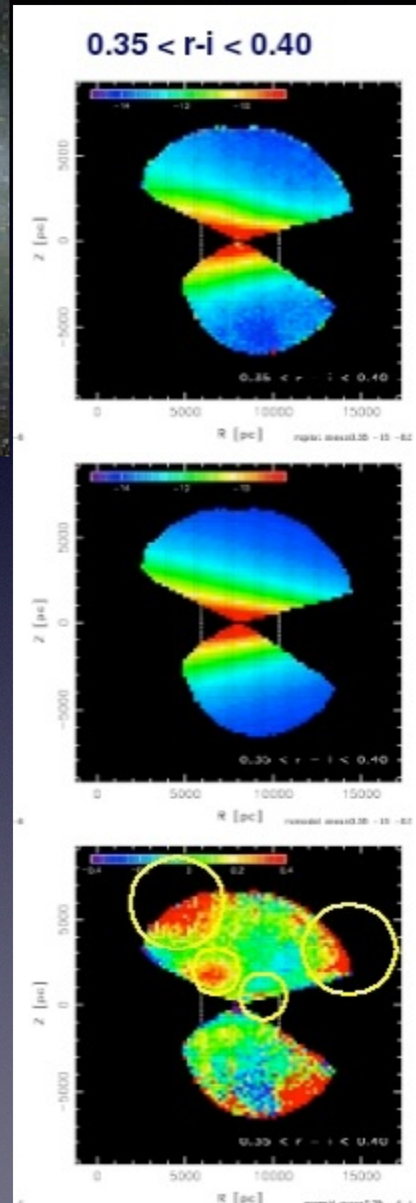
2022-2032

The Milky Way structure: 20 billion stars, time domain massive statistical studies!

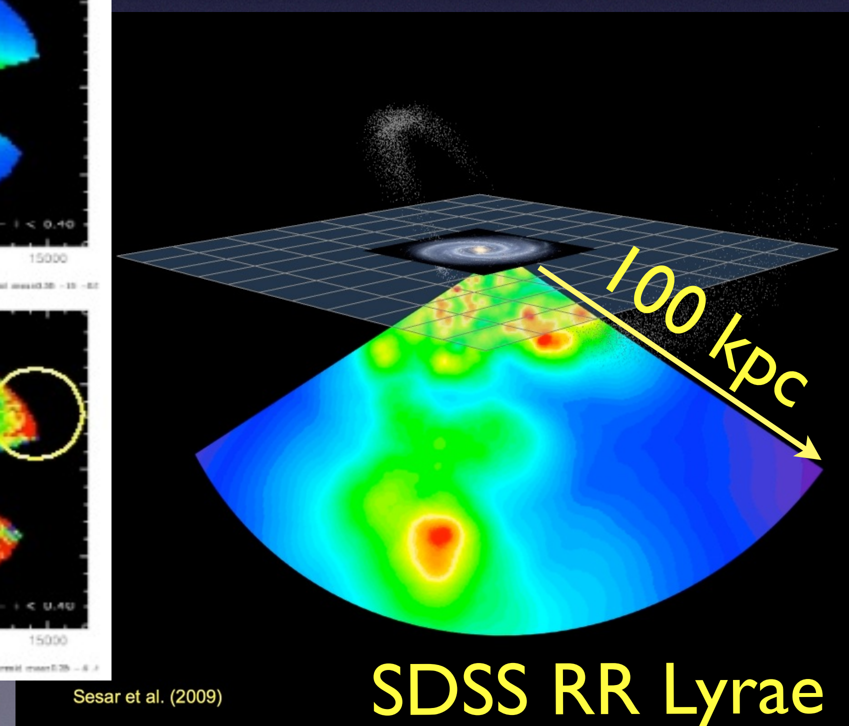


Distance and [Fe/H]:

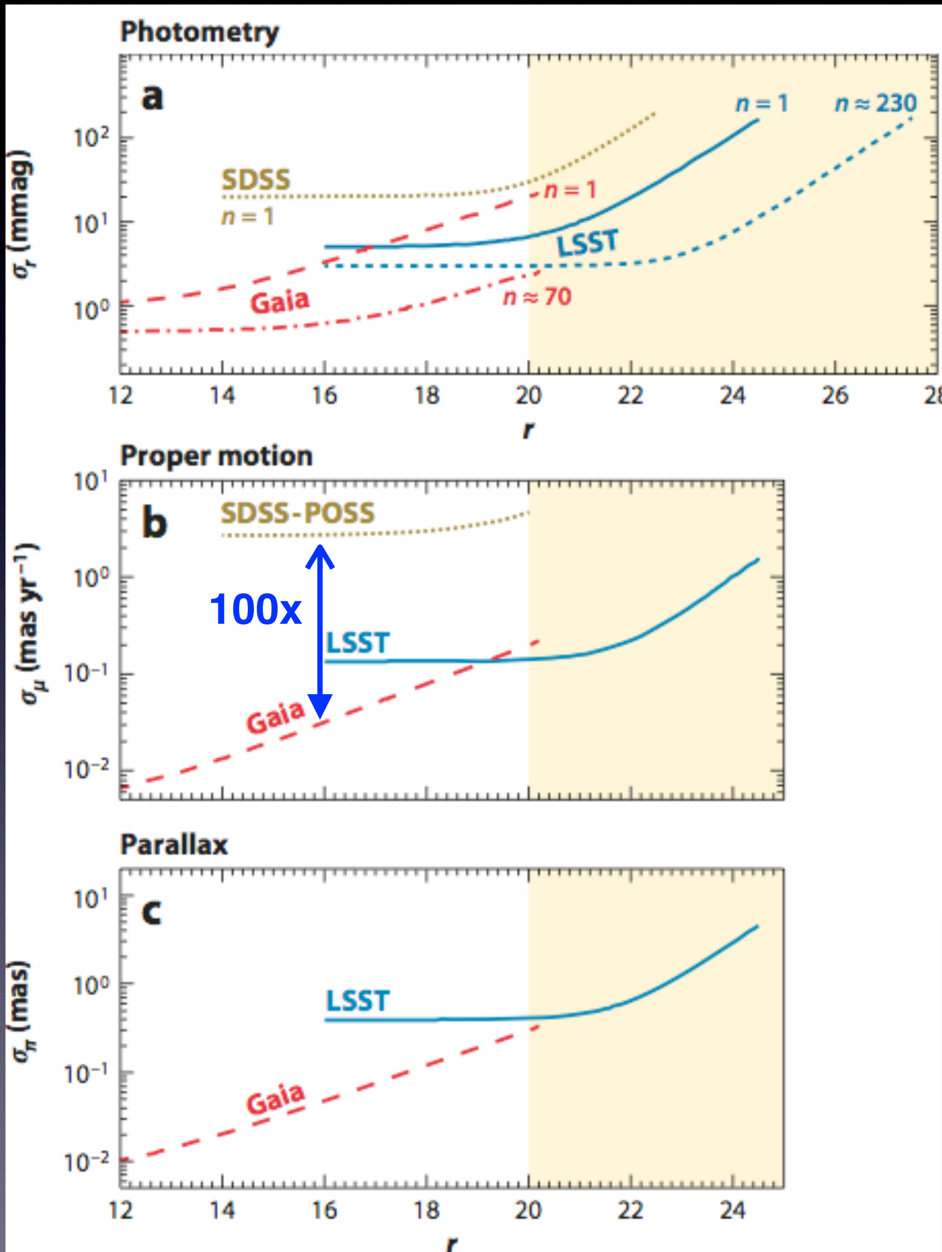




Compared to SDSS: LSST can "see" about 40 times more stars, 10 times further away and over twice as large sky area



Gaia vs. LSST comparison

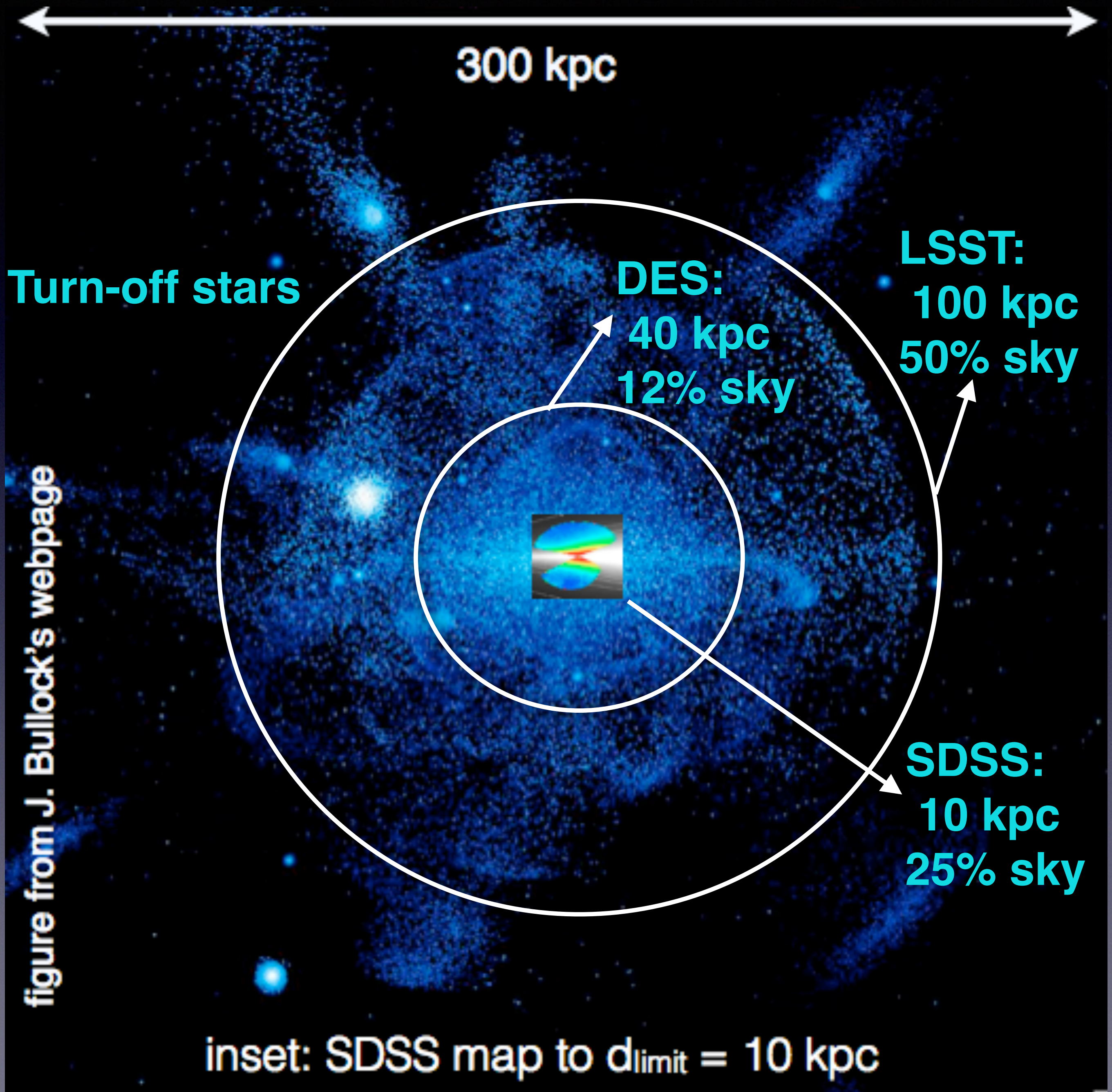


Ivezić, Beers, Jurić 2012, ARA&A, 50, 251

- **Gaia:** excellent astrometry (and photometry), but only to r < 20
- **LSST:** photometry to r < 27.5 and time resolved measurements to r < 24.5
- Complementarity of the two surveys: photometric, proper motion and trigonometric parallax errors are similar around r=20

The Milky Way disk "belongs" to Gaia, and the halo to LSST (plus very faint and/or very red sources, such as white dwarfs and LT(Y) dwarfs).

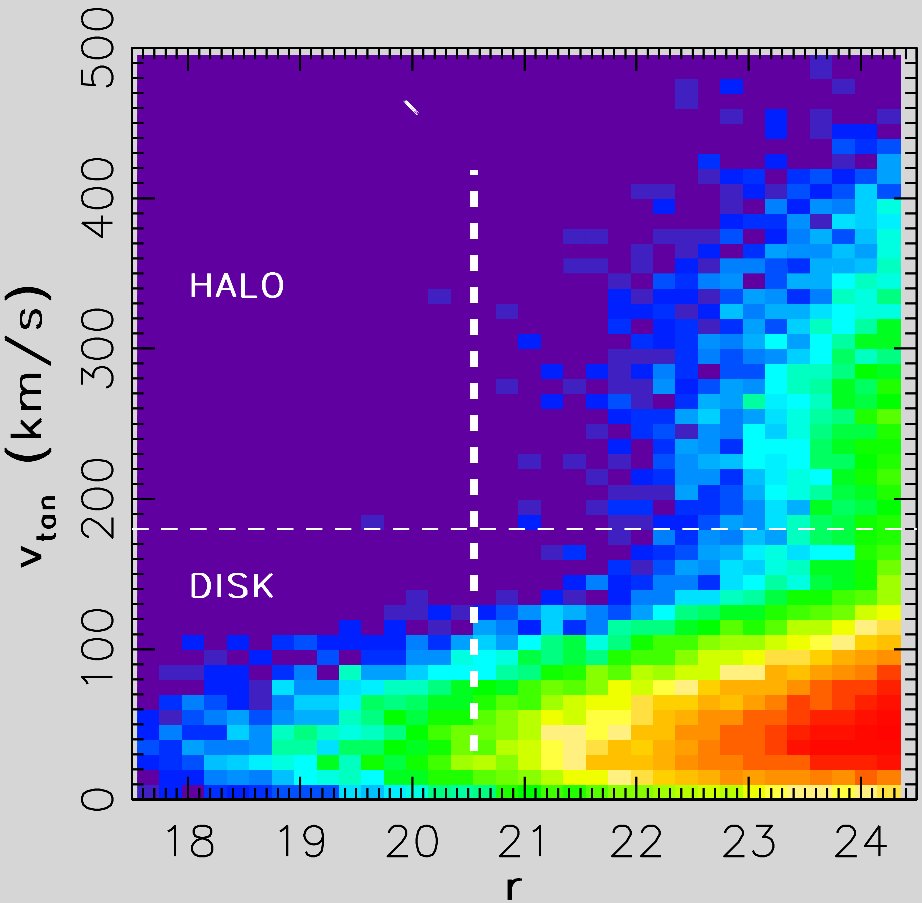
Milky Way science with coadded LSST data



Dwarfs in LSST

White dwarfs: LF is age probe

~400,000 halo white dwarfs from LSST (10 million total):



L / T dwarfs: L dwarfs are dime a dozen: 200,000 in LSST with proper motion and trigonometric parallax measurements

Simulations predict 2400 T dwarfs with >5σ proper motion and parallax measurements

Compared to UKIDSS, 5 times larger sample of T dwarfs, with parallaxes and 10-20 times more accurate proper motions

(~100 Y dwarfs [model based])

Outline

- LSST overview and construction status report
  - multi-color time-resolved faint sky map
  - 20 billion stars and 20 billion galaxies
- A tour of anticipated LSST science programs
  - time domain
  - the Solar System structure
  - cosmology (dark matter and dark energy)
  - the Milky Way structure
- **Data analysis challenges ahead of us**
  - large data sets
  - complex analysis
  - aiming for small systematics

Data analysis challenges in the era of Big Data

1) Large data volume (petabytes)

2) Large numbers of objects (billions)

3) Highly multi-dimensional spaces (thousands)

4) Unknown statistical distributions

5) Time-series data (irregular sampling)

6) Heteroscedastic errors, truncated, censored and missing data

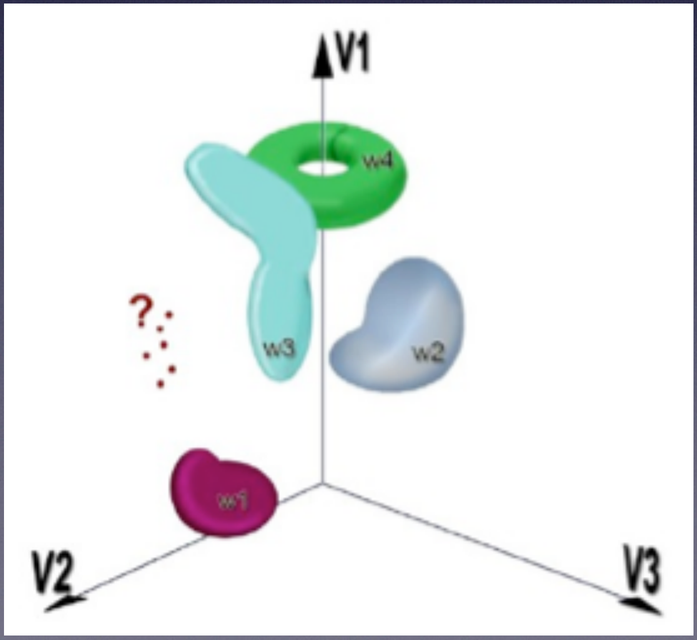
7) Unreliable quantities (e.g. unknown systematics and random errors)

The bottleneck will not be data availability but instead our ability to extract useful and reliable information from data.

Statistical analysis of a massive LSST dataset

- A large (100 PB) database and sophisticated analysis tools: for each of 40 billion objects there will be about 1000 measurements (each with a few dozen measured parameters)

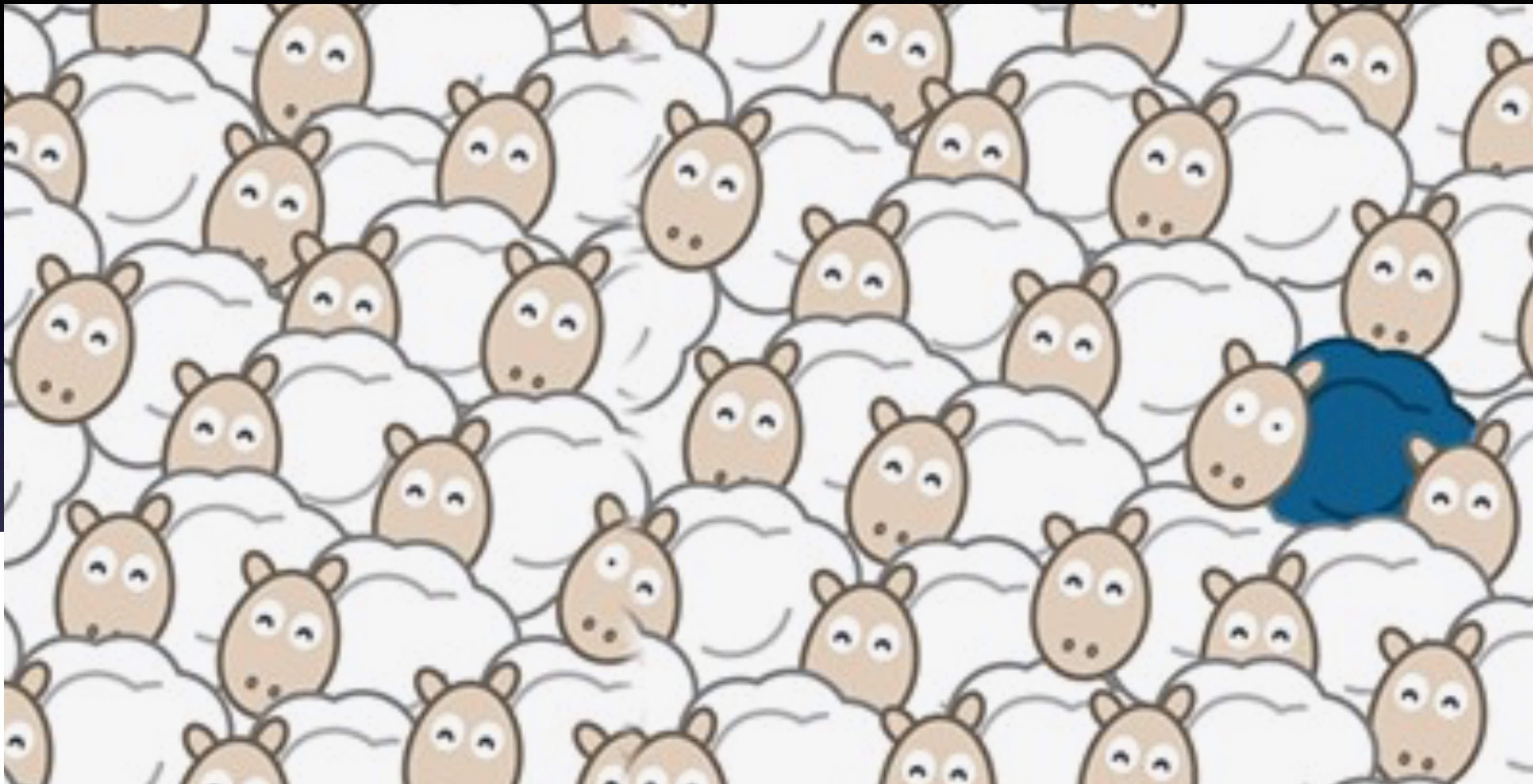
Data mining and knowledge discovery

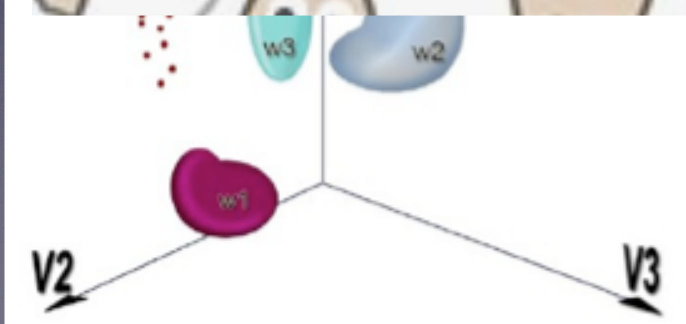


- 10,000-D space with 40 billion points
- Characterization of known objects
- Classification of new populations
- Discoveries of unusual objects

Clustering, classification, outliers

Statistical analysis of a massive LSST dataset





- Classification of new populations
- Discoveries of unusual objects

Clustering, classification, outliers

Summary

- **LSST: a 10-year survey starting in Oct 2022**
- **multi-color time-resolved faint sky map**
  - 20 billion galaxies (median redshift ~1)
  - 20 billion stars (to the edge of the Milky Way)
  - 10 billion alerts (across the Universe)
  - “millions and millions” of SNe, quasars, asteroids…
- **data analysis challenges: waiting for you!**

This talk available as: http://ls.st/qmi

More details:

LSST overview paper: arXiv:0805.2366

LSST Data Products Definition Document: ls.st/dpdd

LSST Science Requirements Document: ls.st/srd

Total Solar Eclipse over
Cerro Pachon, July 2, 2019
Photo: Kevin Reil