Astronomical surveys and LSST

Astr 323, Lecture 9
Spring 2014, University of Washington
• Astronomical surveys
• LSST system summary
  o Science Themes
  o System Characteristics
• LSST science examples
  o Extragalactic astronomy and cosmology
  o The Milky Way and the Local Group
  o Time Domain
• The road ahead
Context: modern observational methods in astronomy and astrophysics:

- **Large telescopes (~10m):** faint objects, especially spectroscopy

The Keck telescopes on Mauna Kea (Hawaii)
Context: modern observational methods in astronomy and astrophysics:

- **Telescopes above the atmosphere**: high angular resolution (e.g., the Hubble Space Telescope) and other wavelength regions (X-ray, radio, infrared)

The HST in orbit and an example of a galaxy image
Context: modern observational methods in astronomy and astrophysics:

- **Large telescopes (~10m):** faint objects, especially spectroscopy
- **Telescopes above the atmosphere:** high angular resolution (e.g., the Hubble Space Telescope) and other wavelength regions (X-ray, radio, infrared)
- **Large sky surveys:** digital sensor technology, information technology, automated data processing and data distribution

Key point: modern sky surveys make all their data (images and catalogs) publicly available
Why are sky maps useful?

- **Sky map**: a list of all detected objects (stars, galaxies, ...) – measured parameters (size, color, brightness, ...)

- **The utility of sky maps**:
  - Discoveries of new objects: “Is this a new asteroid, or is it already cataloged?”
  - Object classification: “What types of galaxies exist?”
  - Statistical population studies: “Do quasars change their properties with time?”
  - Search for unusual objects: “Is this star very weird?”
  - Cosmological measurements: “How fast does the Universe expand?”

“Science-ready database”: measurements can be (simply) analyzed without the need for (complex) image processing.
The last decade: Sloan Digital Sky Survey

- Digital sky survey with a 120 Megapix CCD camera
- Precise measurements for 400,000,000 objects
- Revolution in astronomy: public databases
Astronomy “from your armchair”, from everywhere
“Navigation” around the sky...

SDSS DR7
ra: 18.87667 deg
dec: -0.86083 deg
scale: 1.5845 arcsec/pix
image zoom: 1:16

18.87667, -0.86083

Selected object:
ra: 18.87684
dec: -0.86098
type: GALAXY
u: 14.82
g: 13.74
r: 13.19
i: 12.91
z: 12.93
Additional, more detailed, information...
“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?”
SDSS: one US Library of Congress worth of data

LSST: one SDSS per night, or all the words ever printed!
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)

These drivers not only require similar hardware and software systems, but also motivate a uniform cadence: about 90% of time will be spent on a uniform survey.
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 $r \sim 27.5$ (36 nJy) based on 1000 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 vs. LSST comparison: \( \text{LSST} = \frac{d(\text{SDSS})}{dt}, \text{LSST} = \text{SuperSDSS} \)

3x3 arcmin, gri

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

Deep Lens Survey \((r \sim 26)\)  is (almost) like LSST depth (but tiny area)

\(20 \times 20\) arcsec; lensed SDSS quasar \(\text{J1332+0347}, \text{Morokuma et al. 2007}\)
Required system characteristics

- Large primary mirror (at least 6m) to go faint and to enable short exposures (30 s)
- Agile telescope (5 sec for slew and settle)
- Large field of view to enable fast surveying
- Impeccable image quality (weak lensing)
- Camera with 3200 Mpix
- Sophisticated software (20,000 GB/night, 20 billion objects, 20 trillion measurements)
LSST system
Telescope
Camera
Software
LSST Telescope

8.4m, 6.7m effective

5 sec slew & settle
The field-of-view comparison: Gemini vs. LSST

- **Primary Mirror Diameter**
  - Gemini South Telescope: 8 m
  - LSST: 8.4 m

- **Field of View**
  - Gemini South Telescope: 0.2 degrees
  - LSST: 3.5 degrees

(Full moon is 0.5 degrees)
Three-mirror design (Paul-Baker system) enables large field of view with excellent image quality: delivered image quality is dominated by atmospheric seeing
The largest astronomical camera: 2800 kg, 3.2 Gpix
LSST camera

Modular design: 3200 Megapix = 189 x 16 Megapix CCD
9 CCDs share electronics: raft (=camera)
Problematic rafts can be replaced relatively easily
LSST Software

TBD Site
Science Centers
Data Quality Analysis

Archive Site
Archive Center
Nightly Reprocessing
Data Release Production
Long-term Storage (copy 2)

Data Access Center
Data Access and User Services
120 – 330 TFLOPS
35 – 250 PB

Headquarters Site
Headquarters Facility
Observatory Management
Science Operations
Education and Public Outreach

Base Site
Base Facility
Data Access Center
Alert Production
Long-term storage (copy 1)
60 TFLOPS
35 – 250 PB

Summit Site
Summit Facility
Telescope and Camera
Software: the subsystem with the highest risk

- 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 new code (C++/python)
The Forward Process.

**Galaxies:** Intrinsic galaxy shapes to measured image:

- Intrinsic galaxy (shape unknown)
- Gravitational lensing causes a shear ($g$)
- Atmosphere and telescope cause a convolution
- Detectors measure a pixelated image
- Image also contains noise

**Stars:** Point sources to star images:

- Intrinsic star (point source)
- Atmosphere and telescope cause a convolution
- Detectors measure a pixelated image
- Image also contains noise
LSST imaging processing: an example

- A raw data frame. The difference in bias levels from the two amplifiers is visible.
- Bias-corrected frame with saturated pixels, bad columns, and cosmic rays masked in green.
- Frame corrected for saturated pixels, bad columns, and cosmic rays.
- Bright object detections marked in blue.
- Faint object detections marked in red.
- Measured objects, masked and enclosed in boxes. Small empty boxes are objects detected only in some other band.
- Measured objects in the data frame.
- Reconstructed image using postage stamps of individual objects and sky background from binned image.
LSST image processing pipelines

Data Products

Eng/Fac Data Archive
Calibration Data Products
Image Archive
Source Catalog
Object Catalog
Orbit Catalog
Alert Archive

Most users will use these
A comparison of LSST data processing with other software projects:

- **Complexities we have to deal with in DM**
  - Very high data volumes (transfer, ingest, and especially query)
  - Advances in scale of algorithms for photometry, astrometry, PSF estimation, moving object detection, shape measurement of faint galaxies
  - Provenance recording and reprocessing
  - Evolution of algorithms and technology

- **Complexities we DON’T have to deal with in DM**
  - Tens of thousands of simultaneous users (e.g. online stores)
  - Fusion of remote sensing data from many sources (e.g. earthquake prediction systems)
  - Millisecond or faster time constraints (e.g. flight control systems)
  - Very deeply nested multi-level transactions (e.g. banking OLTP systems)
  - Severe operating environment-driven hardware limitations (e.g. space-borne instruments)
  - Processing that is highly coupled across entire data set with large amount of inter-process communication (e.g. geophysics 3D Kirchhoff migration)
Statistical analysis of a massive LSST dataset

• A large (100 PB) database and sophisticated analysis tools: for each of 20 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 20 billion points
• Characterization of known objects
• Classification of new populations
• Discoveries of unusual objects

Clustering, classification, outliers
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 (the state of the Universe at the recombination epoch, at redshift $\sim 1000$)
- Weak Lensing: growth of structure
- Galaxy Clustering: growth of structure
- Baryon Acoustic Oscillations: standard ruler
- Supernovae: standard candle

 Except for CMB, measuring $H(z)$ and growth of structure $G(z)$

$$H(z) \sim \frac{d[\ln(a)]}{dt}, \quad G(z) = a^{-1} \delta \rho_m / \rho_m, \quad \text{with} \quad a(z) = (1+z)^{-1}$$

LSST is designed to reach Stage IV level from DETF report which kinda means “It will be awesome and better than anything today!”
Cosmology with LSST SNe: is the cosmic acceleration the same in all directions?

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


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

- Photometric redshifts: random errors smaller than 0.02, bias below 0.003, fewer than 10% $>3\sigma$ 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 $r \sim 27.5$

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)
Cosmology with LSST

- Derived from 4 billion galaxies (i<25.3, SNR>20) with accurate photo-z and shape measurements
- Measuring distances and growth of structure with a percent accuracy for 0.5 < z < 3
- SNe will provide a high angular resolution probe of homogeneity and curvature

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: high precision measurements**

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

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

- $E_G$ combines 3 measures of large-scale structure: galaxy-galaxy lensing ($\phi + \psi$), galaxy clustering ($\phi$) 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!)

Reyes et al. (2010, Nature 464, 256)
Extragalactic astronomy: galaxies

- About 10 billion galaxies, with 4 billion in a “gold” sample defined by $i<25.3$
- The “gold” sample extends to redshifts of $>2.5$: evolution

SDSS: snapshot at $z\sim0$

LSST: a galaxy evolution movie to $z\sim2.5$
Extragalactic astronomy: galaxies

SDSS

MUSYC

Gawiser et al
Extragalactic astronomy: quasars

- About 10 million quasars will be discovered using variability, colors, and the lack of proper motions. Really?? SDSS: yes!
- The sample will include $M_i = -23$ objects even at redshifts beyond 3.
- Quasar variability studies will be based on millions of light curves with 1000 observations over 10 yrs.

Top: absolute magnitude vs. redshift diagram for quasars.

Today: ~31 quasars with $6 < z < 7.5$ Reionization studies!
LSST will detect ~10,000 quasars with $6 < z < 7.5$!
The Highest Redshift Quasar at $z=7.085$ from UKIDSS

Mortlock et al. 2011

Such a quasar would be detected by LSST as a $z$-band dropout (multi-epoch data will greatly help with false positives)

LSST will discover about 1,000 quasars with $z>7$

Today: one quasar with $z>7$
The Milky Way structure: 10 billion stars, time domain

Compared to SDSS: LSST can “see” 10 times further away and over twice as large an area.

Distance and [Fe/H]:

SDSS RR Lyrae
The large blue circle: the \( \sim 400 \) kpc limit of future LSST studies based on RR Lyrae

The large red circle: the \( \sim 100 \) kpc limit of future LSST studies based on main-sequence stars (and the current limit for RR Lyrae studies)

The small insert: \( \sim 10 \) kpc limit of SDSS and future Gaia studies for kinematic \& \([Fe/H]\) mapping with MS stars
Time Domain: objects changing in time
positions: asteroids and stellar proper motions
brightness: cosmic explosions and variable stars
Time Domain: objects changing in time

positions: asteroids and stellar proper motions

brightness: cosmic explosions and variable stars

For example:

SDSS demonstrated that asteroid families have distinct colors: chemical composition

LSST will turn this diagram into a movie (millions of asteroids)
Killer asteroids: the impact probability is not 0

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

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

Becker et al.
The impact of LSST on other wavelengths, and vice versa:
1) Science Results (e.g. galaxy/AGN evolution)
2) Tools and Methods (e.g. massive databases [radio])
3) Supplemental data (coeval, identification, physical processes)
Also non-EM: e.g. Advanced LIGO
• LSST system summary
  ○ Science Themes
  ○ System Characteristics

• LSST science examples
  ○ Extragalactic astronomy and cosmology
  ○ The Milky Way and the Local Group
  ○ Time Domain

• The road ahead
Priorities:

- **Spaced-based:**
  1. *Wide-Field Infrared Survey Telescope* **WFIRST**
  2. *The Explorer Program* **“rapid response”**
  3. *Laser Interferometer Space Antenna* **LISA**
  4. *International X-ray Observatory* **IXO**

- **Ground-based:**
  1. *Large Synoptic Survey Telescope* **LSST**
  2. *Mid-scale Innovations Program* **“rapid response”**
  3. *Giant Segmented Mirror Telescope (30m)* **GSMT**
  4. *Atmospheric Čerenkov Telescope Array (γ)* **ACTA**
  5. *Cerro Chajnantor Atacama Telescope (submm)* **CCAT**
Why LSST?
The top rank accorded to LSST is a result of:

1. “its compelling science case and capacity to address so many of the science goals of this survey”, [and]
2. “its readiness for submission to the MREFC process as informed by its technical maturity, the survey’s assessment of risk, and appraised construction and operations costs.”

Also: “education and public outreach”

Bill Gates: “LSST will be the ultimate network peripheral device to the Universe”

Google Sky, World Wide Telescope, ...
Why LSST?

The top rank accorded to LSST is a result of:

1. “its compelling science case and capacity to address so many of the science goals of this survey”, [and]
2. “its readiness for submission to the MREFC process as informed by its technical maturity, the survey’s assessment of risk, and appraised construction and operations costs.”

Also: “education and public outreach”

LSST Science Book

Summarizes the basic parameters of the LSST hardware, software, and observing plans, discusses educational and outreach opportunities, and describes a broad range of science that LSST will revolutionize

245 authors, 15 chapters, 600 pages
LSST Timeline

- **Estimate:** survey operations begin in 2019 (if MREFC in FY2014)
  - Primary/Tertiary Mirror being polished, have secondary mirror blank
  - Sensor development program delivered first prototype sensors
  - Processing pipelines under construction, hand-in-hand with simulations of Operations, Images, Catalogs
- **Cost:** ~$850M in $2011
  - contributions from NSF, DOE and private gifts
How to spend a billion dollars?
Half for construction, half for 10 years of operations

Total Project Cost: 455M 2009USD

- Commissioning: $12M
- Project Management: $17M
- Education & Outreach: $9M
- Contingency: $101M
- Telescope: $151M
- Camera: $86M
- Data Management: $78M

LSST Construction Component Cost

First light: around 2019 (if federal constr. in FY2014)
El Penon: Mar 8, 2011

At 8:56:00 the first blast was detonated on the El Penon summit in preparation for the LSST...
Useful LSST Links:

Main public website:  
http://www.lsst.org/

Science Requirements Document:  
http://www.lsst.org/files/docs/SRD.pdf

Overview paper:  
http://www.lsst.org/files/docs/overviewV2.0.pdf

LSST Science Book:  
http://www.lsst.org/lsst/scibook