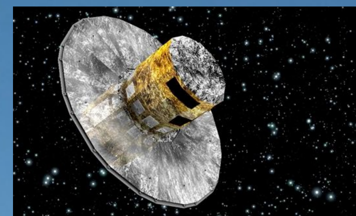
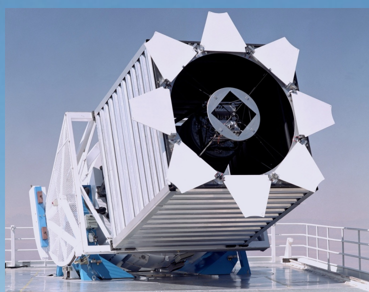
From SDSS to Gaia and LSST

Željko Ivezić
University of Washington





The Milky Way Unravelled by Gaia
Barcelona, Dec 4, 2014

What is a "Stellar Population"?

… a collection of stars with common spatial, kinematic, chemical, and/or age distributions

E.g., 8D space of

- 3 spatial coord.
- 3 velocity comp.
- [Fe/H], [α/Fe]

and then slice & dice!

Galactic Stellar Populations in the Era of the Sloan Digital Sky Survey and Other Large Surveys

Željko Ivezić,[1] Timothy C. Beers,[2] and Mario Jurić[3]

[1]Department of Astronomy, University of Washington, Seattle, Washington 98195; email: ivezic@astro.washington.edu

[2]National Optical Astronomy Observatory, Tucson, Arizona 85719, and Department of Physics & Astronomy and the Joint Institute for Nuclear Astrophysics (JINA), Michigan State University, East Lansing, Michigan 48824; email: beers@noao.edu

[3]Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138; email: mjuric@cfa.harvard.edu

Annu. Rev. Astron. Astrophys. 2012. 50:251–304

The *Annual Review of Astronomy and Astrophysics* is online at astro.annualreviews.org

This article's doi: 10.1146/annurev-astro-081811-125504

0066-4146/12/0922-0251$20.00

Keywords

methods: data analysis; stars: statistics; Galaxy: disk, halo, stellar content, structure, interstellar medium

Abstract

Studies of stellar populations, understood to mean collections of stars with common spatial, kinematic, chemical, and/or age distributions, have been reinvigorated during the past decade by the advent of large-area sky surveys such as the Sloan Digital Sky Survey, the Two-Micron All Sky Survey, the Radial Velocity Experiment, and others. We review recent analyses of these data that, together with theoretical and modeling advances, are revolutionizing our understanding of the nature of the Milky Way and galaxy formation and evolution in general. The formation of galaxies like the Milky

What is a "Stellar Population"?

Galactic Stellar Populations in the Era of the Sloan Digital Sky Survey and Other Large Surveys

Željko Ivezić,[1] Timothy C. Beers,[2] and Mario Jurić[3]

[1]Department of Astronomy, University of Washington, Seattle, Washington 98195; email: ivezic@astro.washington.edu

[2]National Optical Astronomy Observatory, Tucson, Arizona 85719, and Department of Physics & Astronomy and the Joint Institute for Nuclear Astrophysics (JINA), Michigan State University, East Lansing, Michigan 48824; email: beers@noao.edu

[3]Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138;

From Wikipedia: In statistics, a **population** is a complete set of items that share at least one property in common that is the subject of a statistical analysis.

Ann

The
onli

Thi
10.1

0066-4146/12/0922-0251$20.00

Abstract

Studies of stellar populations, understood to mean collections of stars with common spatial, kinematic, chemical, and/or age distributions, have been reinvigorated during the past decade by the advent of large-area sky surveys such as the Sloan Digital Sky Survey, the Two-Micron All Sky Survey, the Radial Velocity Experiment, and others. We review recent analyses of these data that, together with theoretical and modeling advances, are revolutionizing our understanding of the nature of the Milky Way and galaxy formation and evolution in general. The formation of galaxies like the Milky

Outline

- Brief introduction to LSST
  – science drivers
  – system design
  – opportunities for international participation

- A comparison of Gaia and LSST
  - photometric depth and precision
  - trigonometric parallax accuracy
  - proper motion accuracy

- Example: dark matter studies with stars
  a.k.a. “near-field cosmology”
  SDSS: Milky Way’s dark matter halo to ~15 kpc
  Gaia: to 15 kpc like SDSS, but much more precise and robust
  LSST: Milky Way’s dark matter halo to ~100 kpc

- A brief note about student training

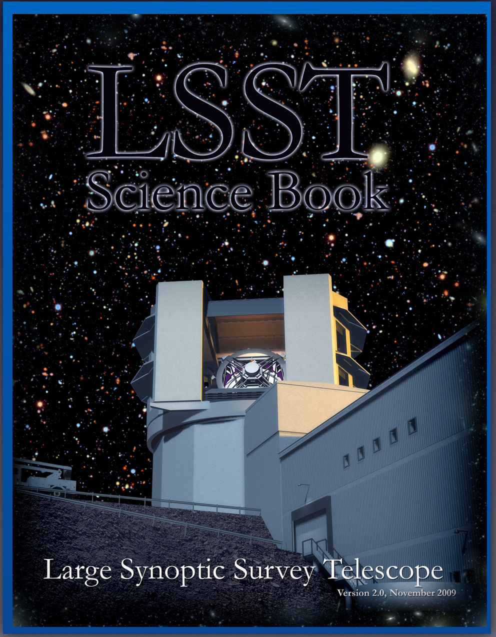
LSST Science Themes

- 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



LSST in one sentence:

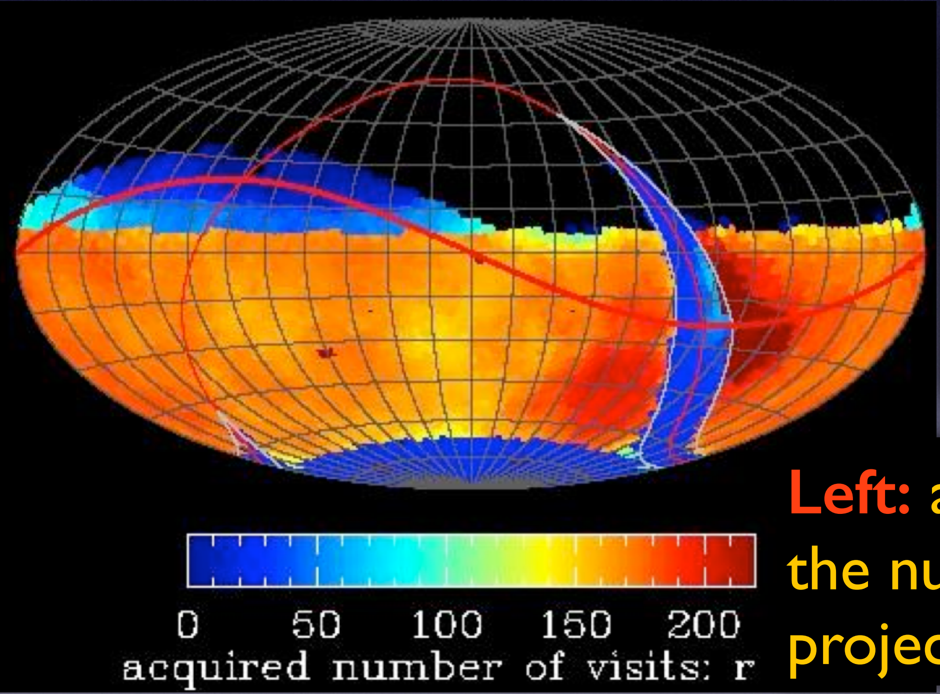
An optical/near-IR survey of half the sky in ugrizy bands to r~27.5 based on ~1000 visits over a 10-year period:

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 vs. LSST comparison: LSST=d(SDSS)/dt, LSST=SuperSDSS

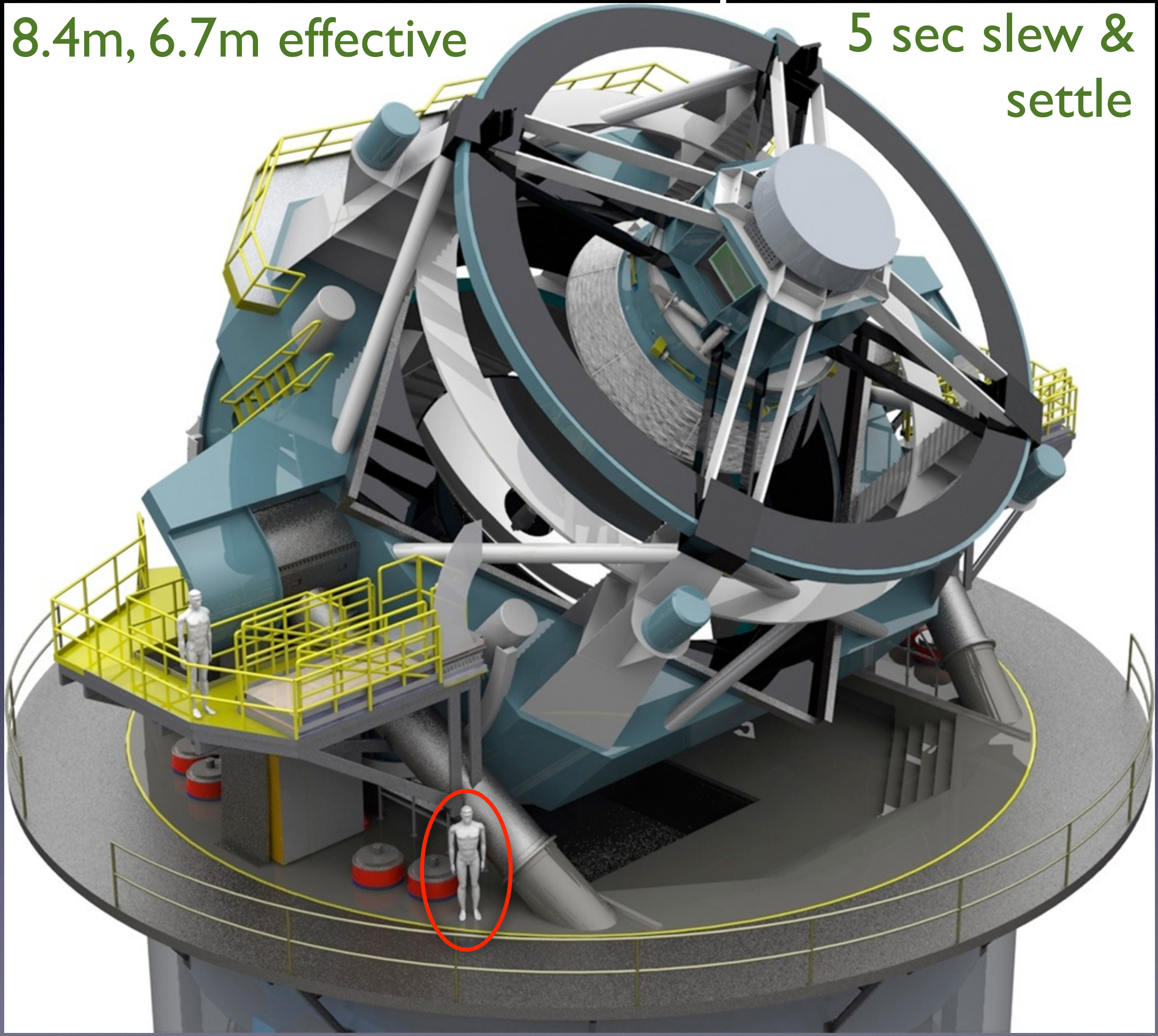
3x3 arcmin, gri

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

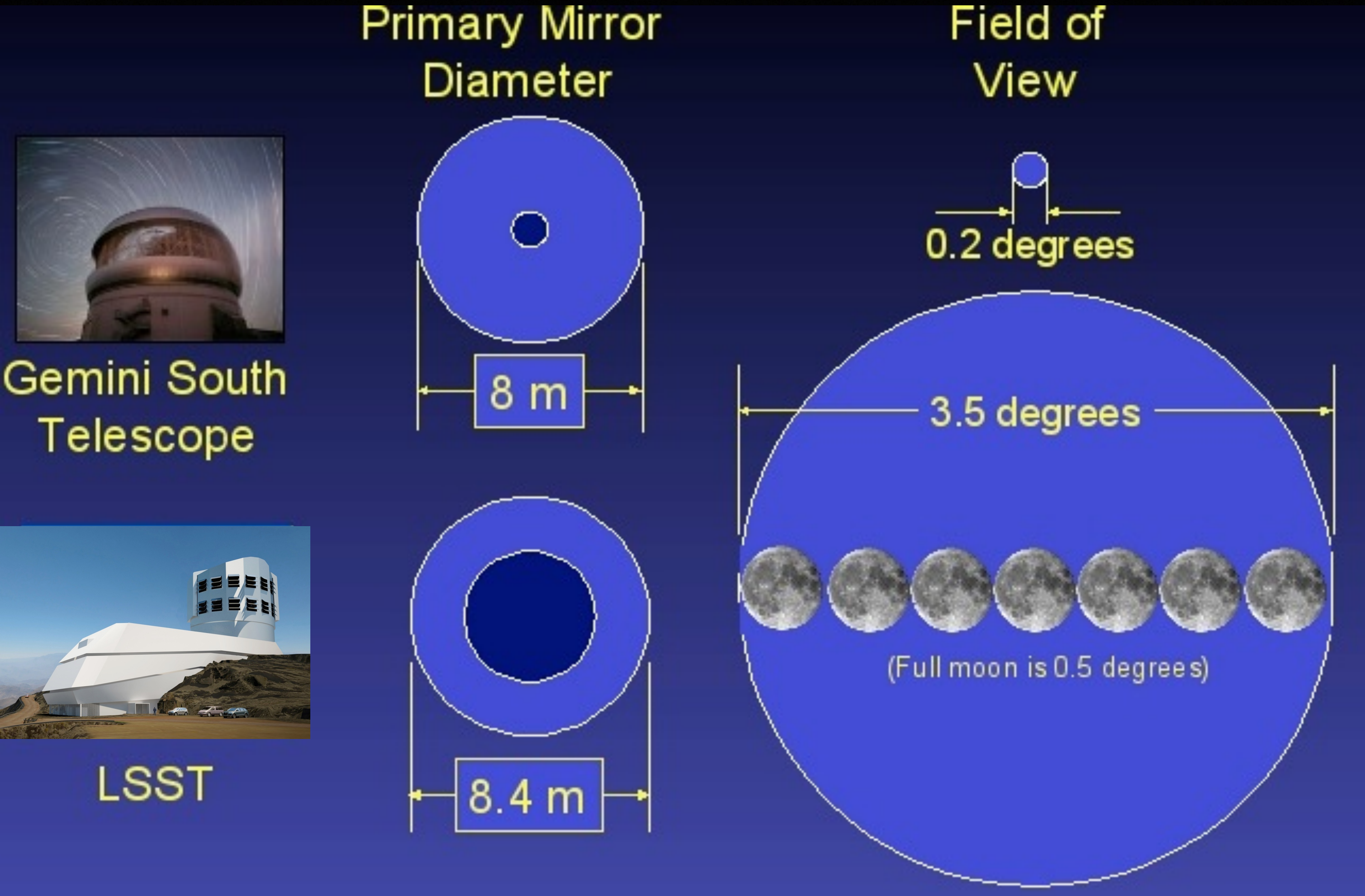
(almost) like LSST depth (but tiny area)

20x20 arcsec; lensed SDSS quasar (SDSS J1332+0347, Morokuma et al. 2007)

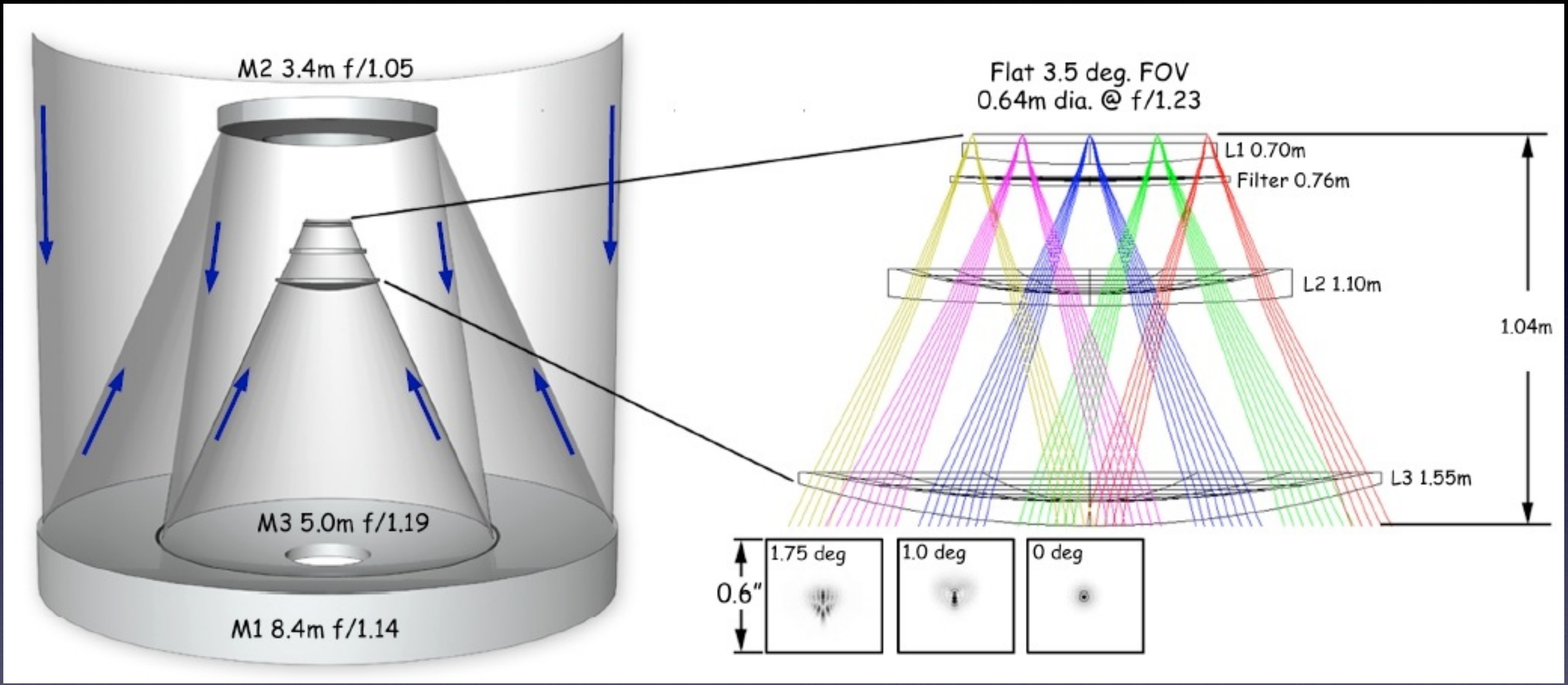
LSST Telescope



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

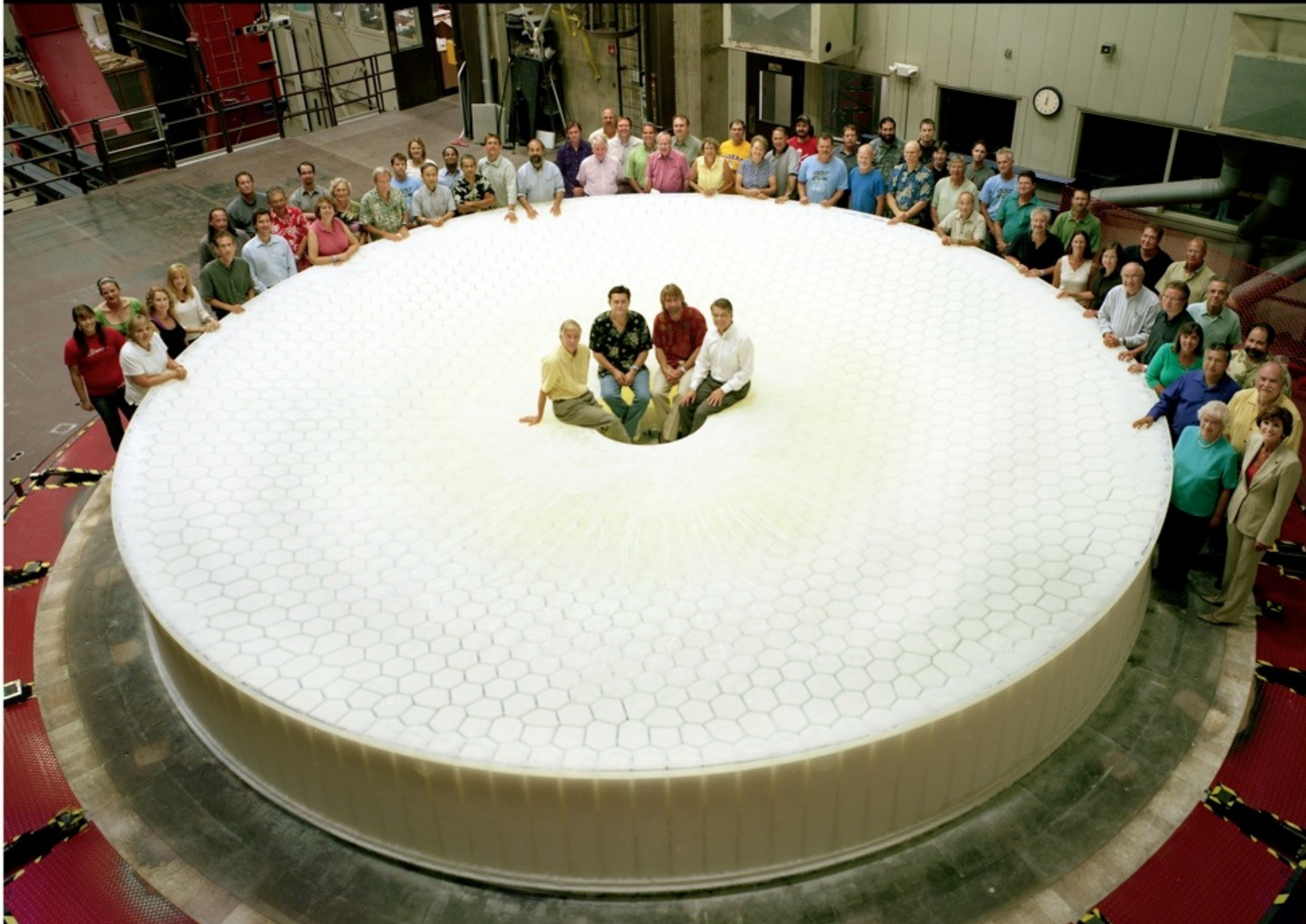


Optical Design for LSST



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

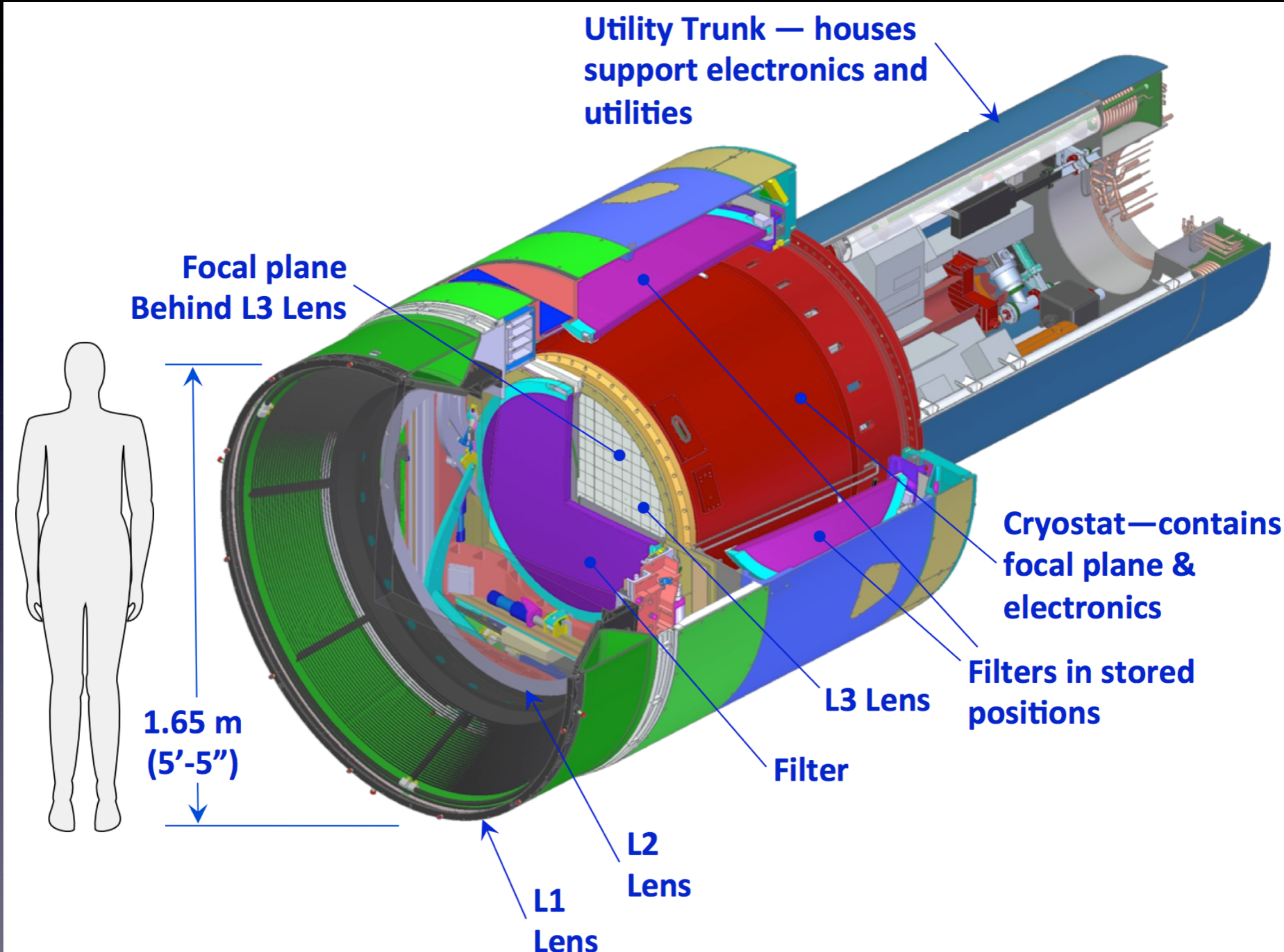


LSST Primary/Tertiary Mirror Blank
August 11, 2008, Steward Observatory Mirror Lab, Tucson, Arizona

The University of Arizona
Tucson Arizona

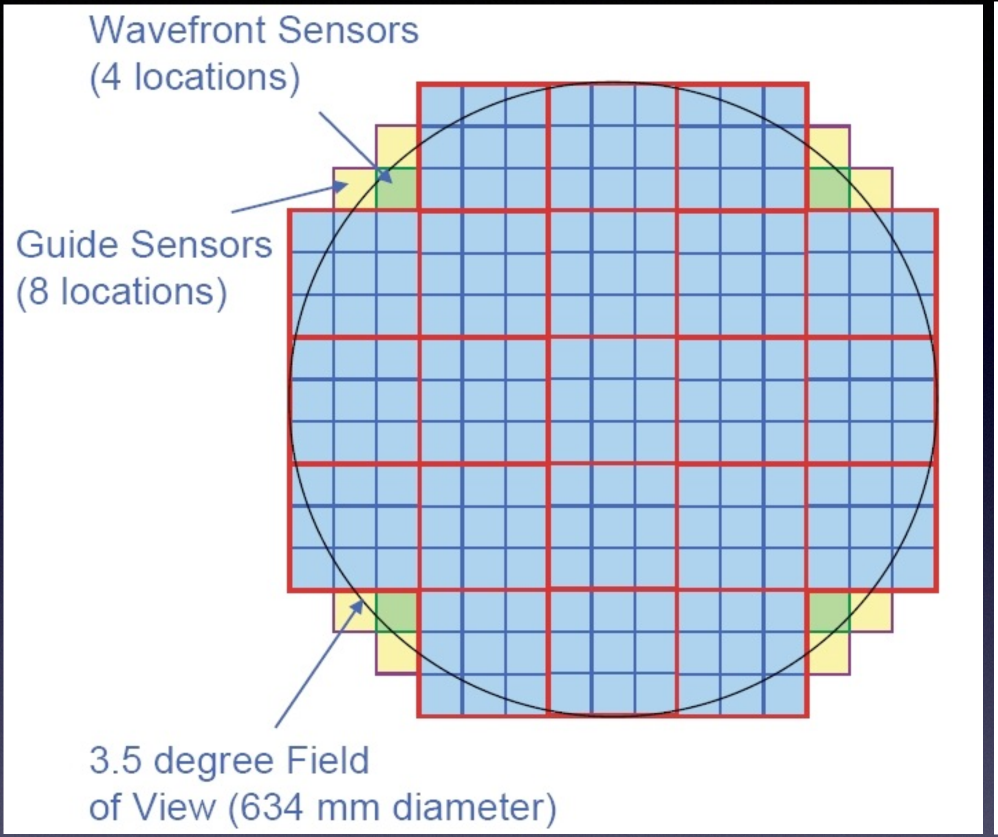
Large Synoptic Survey Telescope

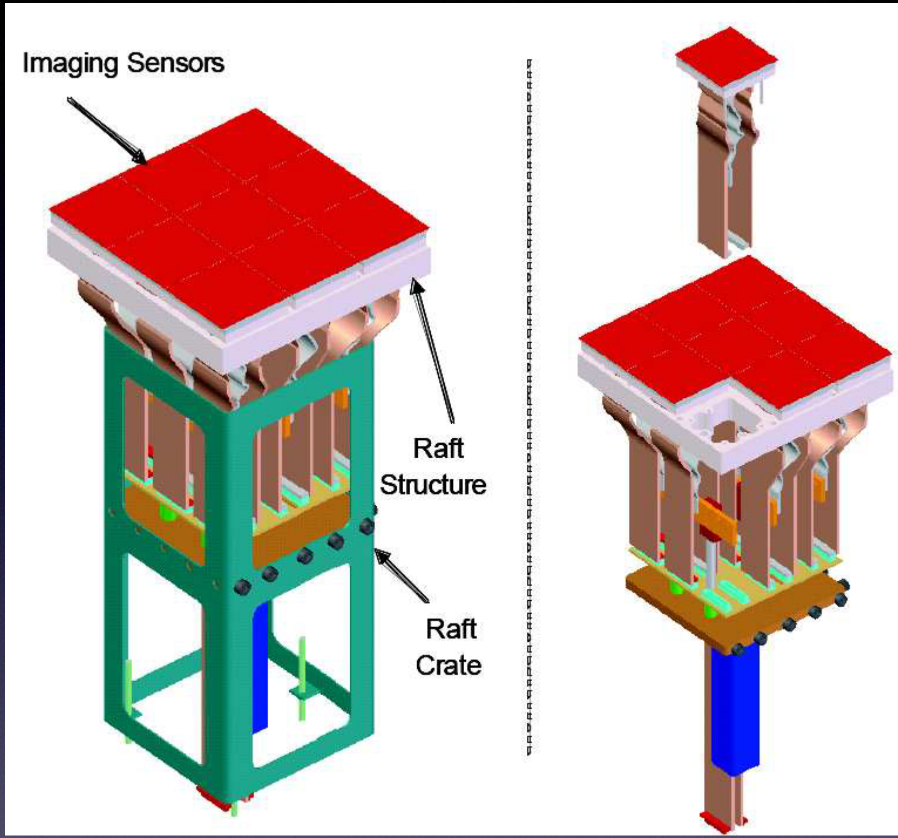
LSST camera



The largest astronomical camera: 2800 kg, 3.2 Gpix

LSST camera





Modular design: 3200 Megapix = 189 x16 Megapix CCD

9 CCDs share electronics: raft (=camera)

Problematic rafts can be replaced relatively easily

Galaxies:

- **Photometric redshifts:** 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)

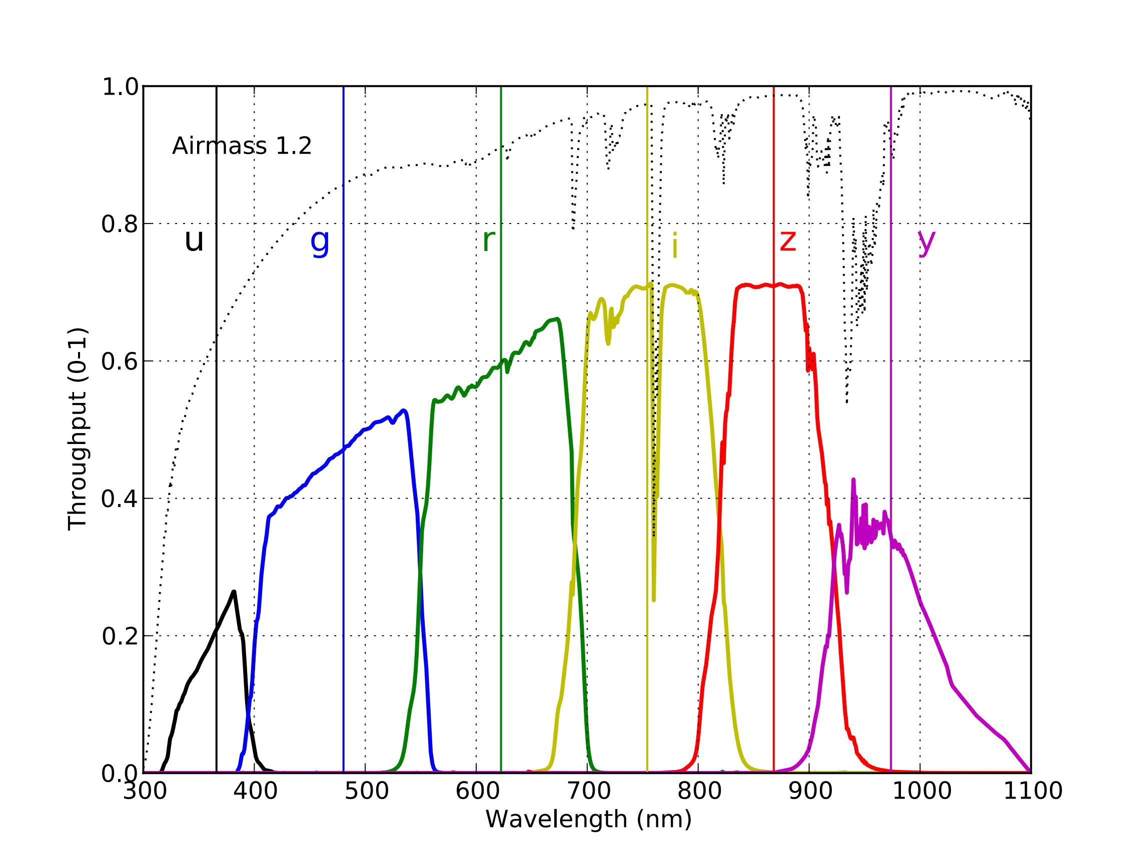


Photo-z requirements correspond to r~27.5

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)

LSST software

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 (~40 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.

LSST software

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)

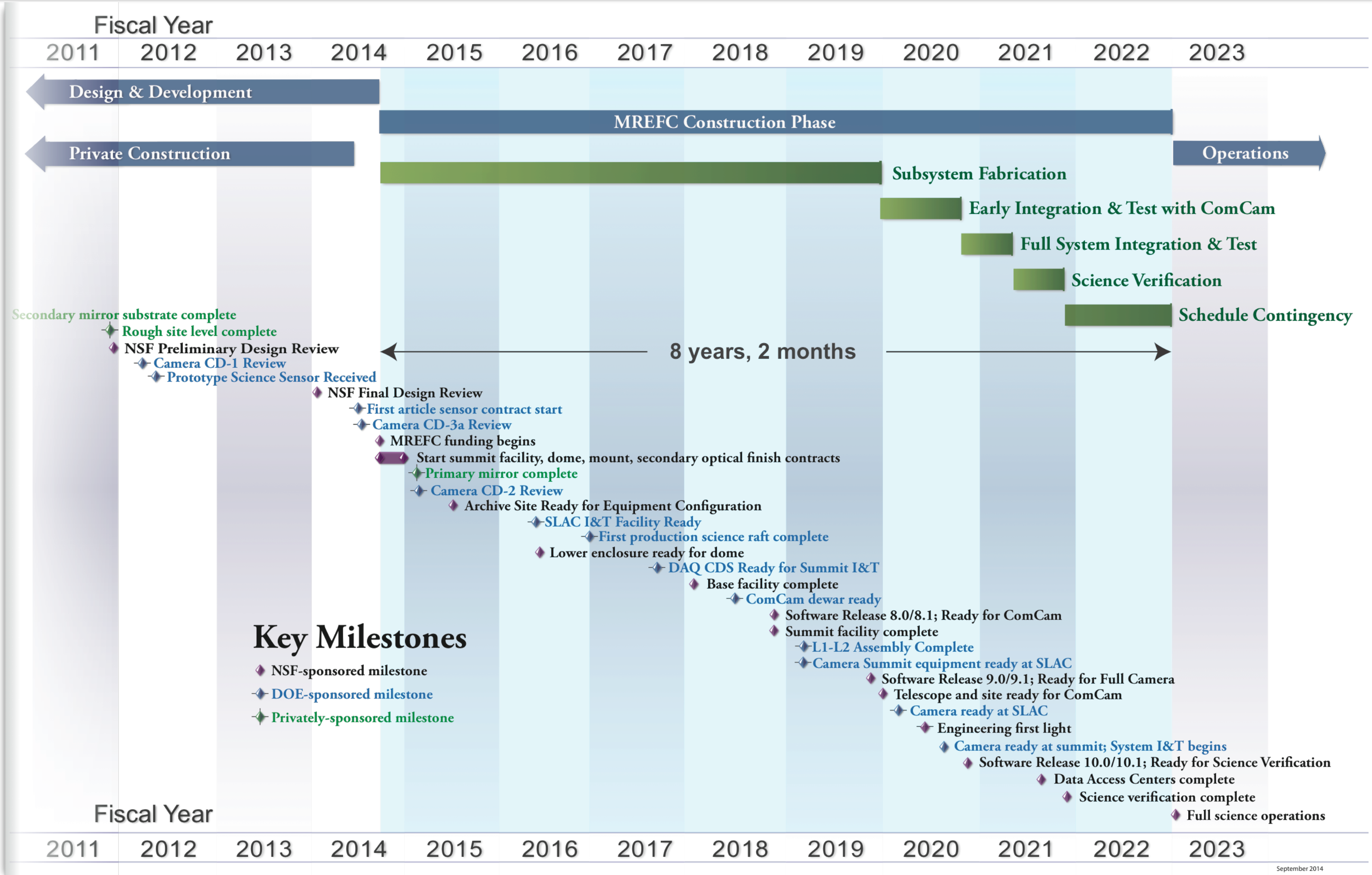
3) Produce catalogs ("model parameters") of detected objects (~40 billion)

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.

Seeking a postdoc? LSST is hiring this year!

Integrated Project Schedule



First data: ~2019 with the commissioning camera (one raft = nine 4kx4k CCDs)

Science Verification with the full camera: ~2021

First public Data Release: ~2022

International Participation in LSST

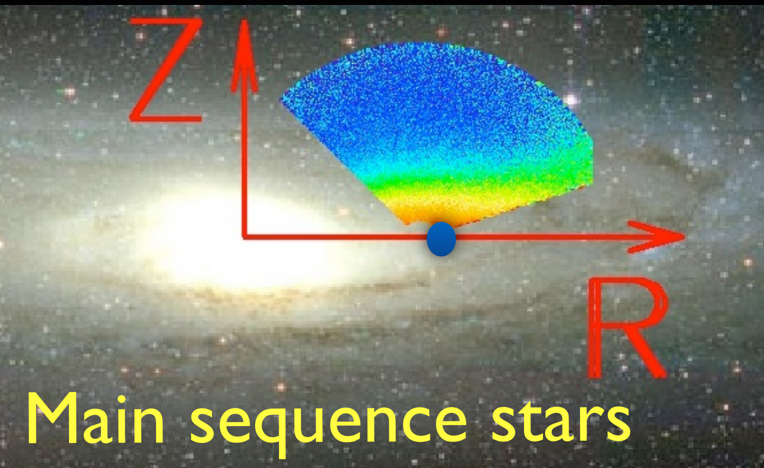
- LSST construction funding is primarily coming from US federal agencies: NSF and DOE (and ~10% from private gifts)
- Unlimited access to data will be granted to US and Chilean scientists, and international partners who signed MOUs with LSST (for others: 2 year delay, except for transients)
- LSST is seeking international partners to support a fraction of operations cost (M$10 per year); the cost per PI (a senior researcher plus two postdocs and graduate students) is very small: $20,000 per year, starting in 2019 (plus DM cost)
- LSST will be a *world unique* scientific facility; the scientific exploitation of LSST data would certainly benefit from greater international participation. International collaborators bring:
  - New creative ideas for innovative investigations with LSST
  - Access to corollary facilities that can enhance the science of LSST
  - Key skills to collaborations that are preparing for some of the more challenging LSST analyses

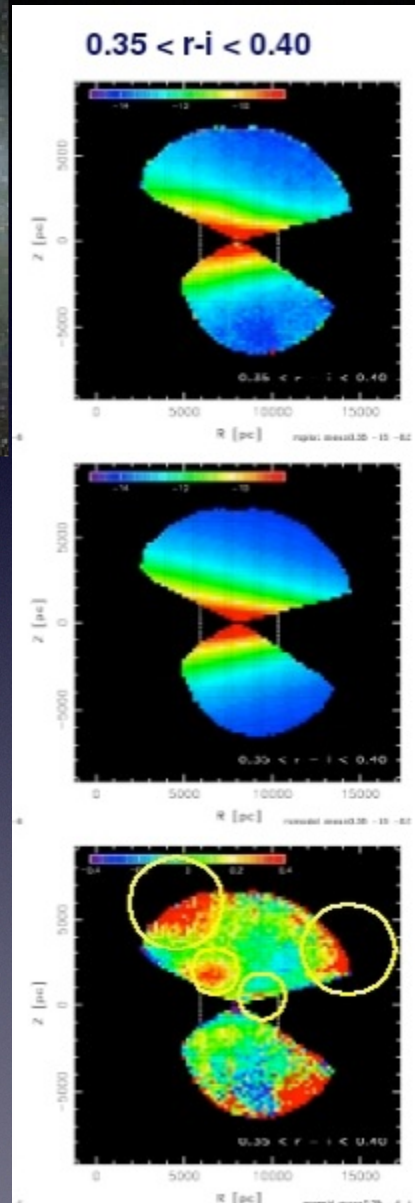
International Participation in LSST

- LSST construction funding is primarily coming from US federal agencies: NSF and DOE (and ~10% from private gifts)
- Unlimited access to data will be granted to US and Chilean scientists, and international partners who signed MOUs with LSST (for others: 2 year delay, except for transients)
- LSST is seeking international partners to support a fraction of operations cost (M$10 per year); the cost per PI (a senior researcher plus two postdocs and graduate students) is very small: $20,000 per year, starting in 2019 (plus DM cost)
- LSST will be a *world unique* scientific facility; the scientific exploitation of LSST data would certainly benefit from greater international participation. International collaborators bring:
  - New creative ideas for innovative investigations with LSST
  - Access to corollary facilities that can enhance the science of LSST
  - Key skills to collaborations that are preparing for some of the more challenging LSST analyses

Individuals and institutions with Gaia experience will have a major advantage in using LSST data!

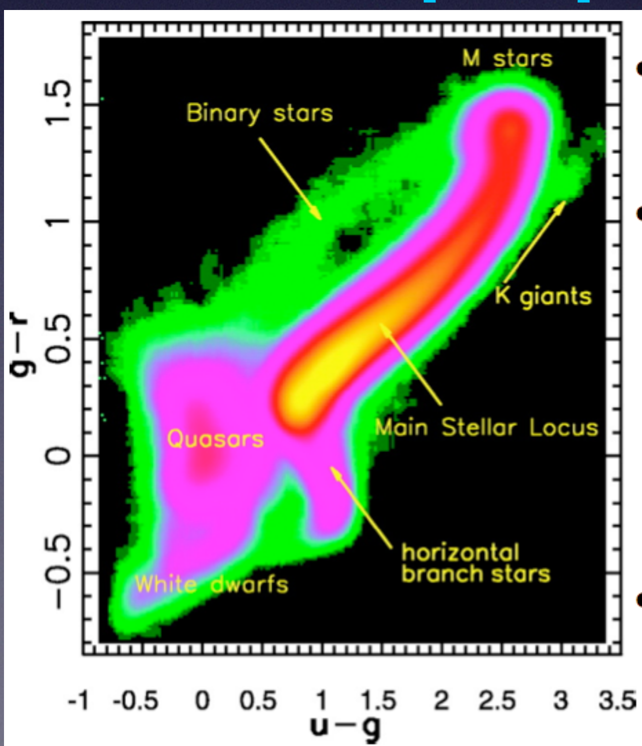
The Milky Way structure: 20 billion stars, time domain massive statistical studies!

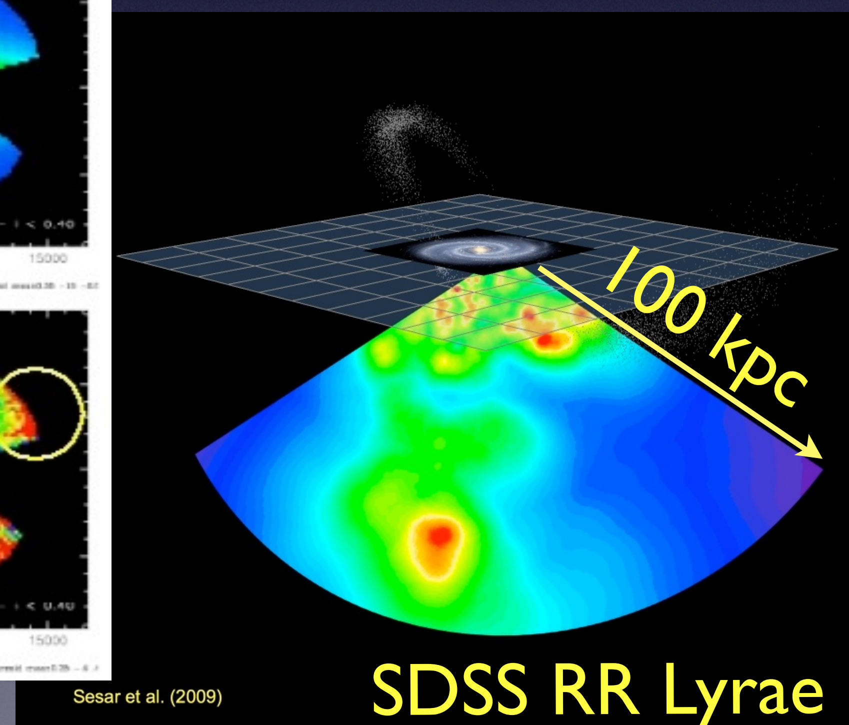


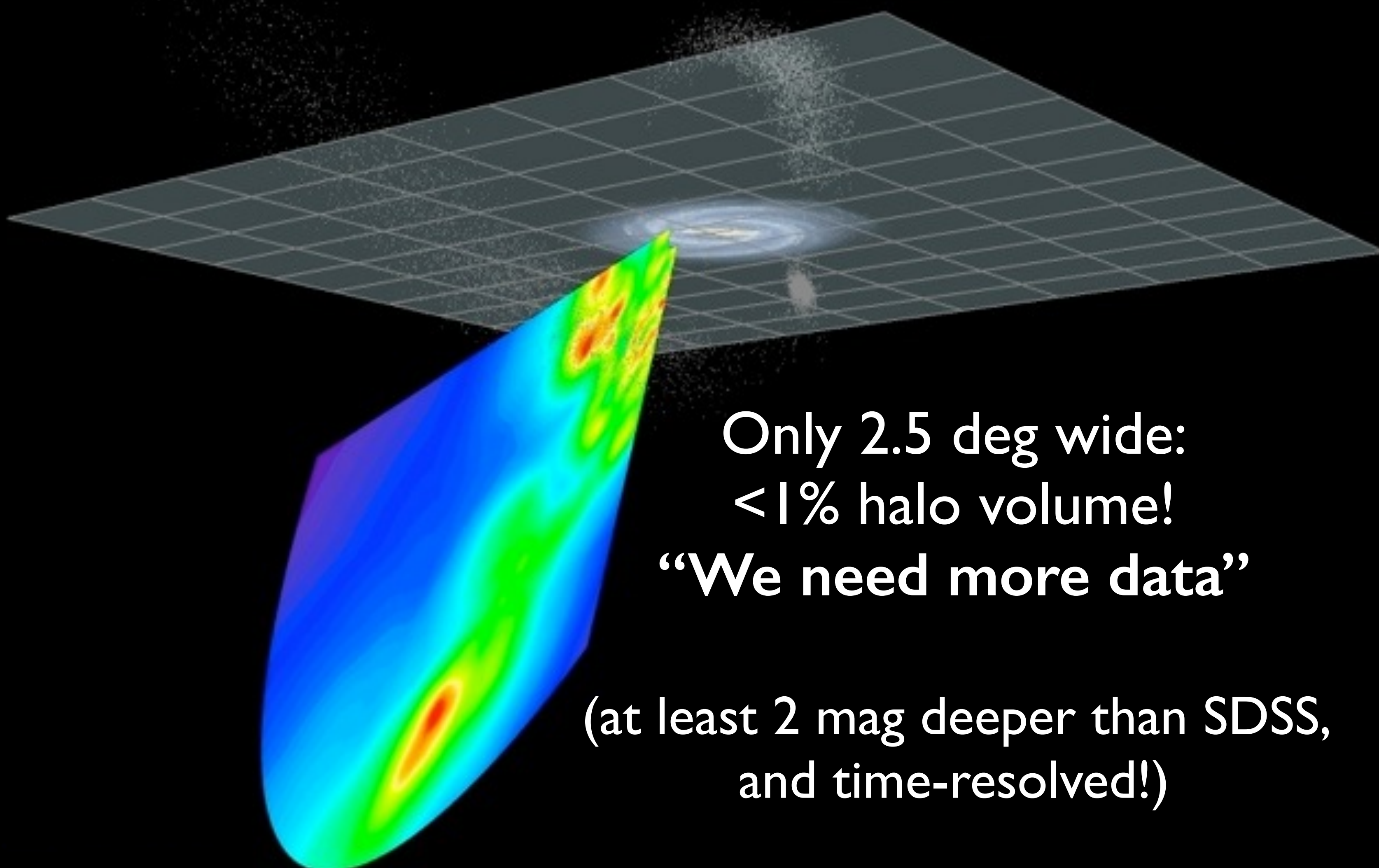


Compared to SDSS: LSST can "see" about 40 times more stars, 10 times further away and over twice as large sky area

Distance and [Fe/H]:





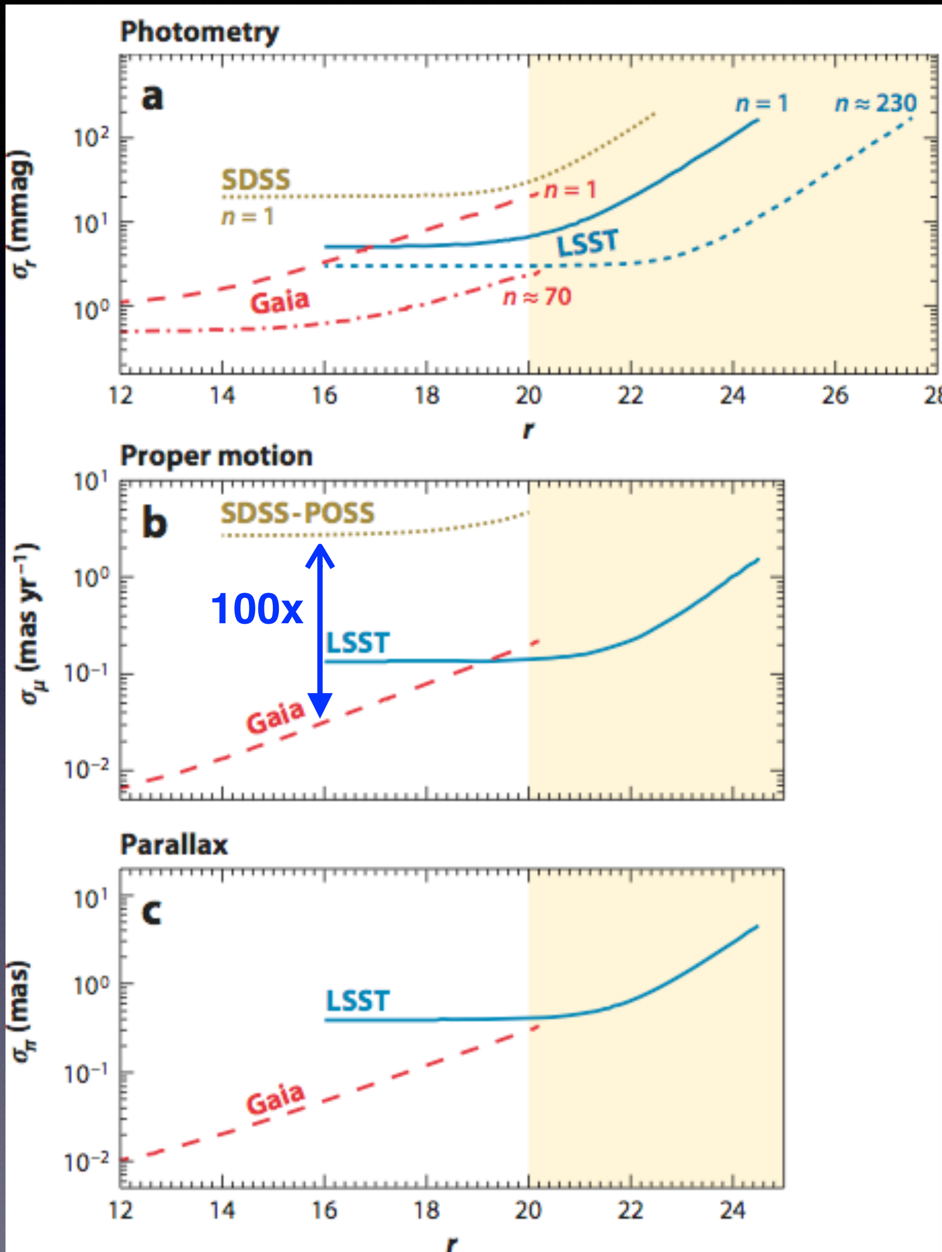


Only 2.5 deg wide:
<1% halo volume!
“We need more data”

(at least 2 mag deeper than SDSS,
and time-resolved!)

Sesar et al. (2010)

Gaia vs. LSST comparison



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

In the context of Gaia, the LSST can be thought of as its deep complement.

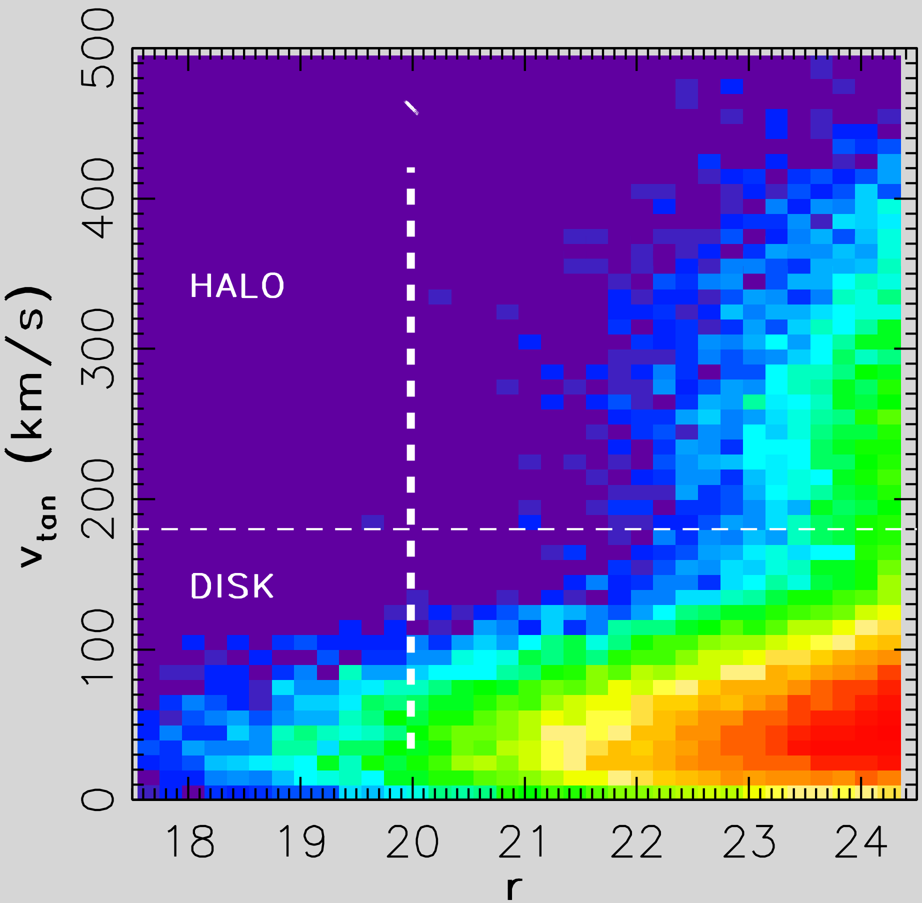
- **Gaia:** excellent astrometry (and photometry), but only to r < 20
- **LSST:** photometry to r < 27.5 and time resolved measurements to r < 24.5
- Complementarity of the two surveys: photometric, proper motion and trigonometric parallax errors are similar around r=20 Calibration of LSST!

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).

Dwarfs in LSST

White dwarfs: LF is age probe

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



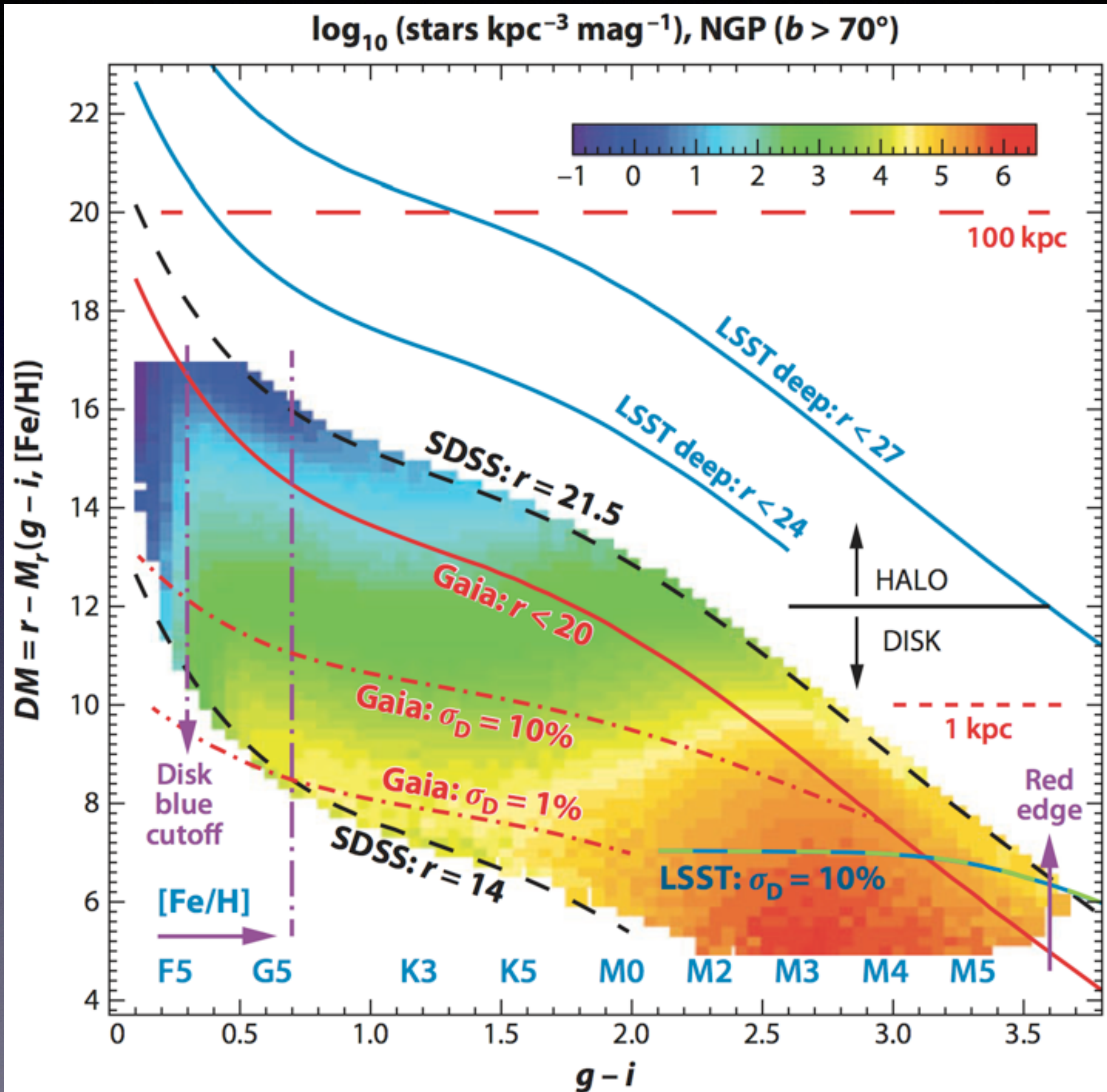
L / T dwarfs: L dwarfs are dime a dozen: 200,000 in LSST with proper motion and trigonometric parallax measurements

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])

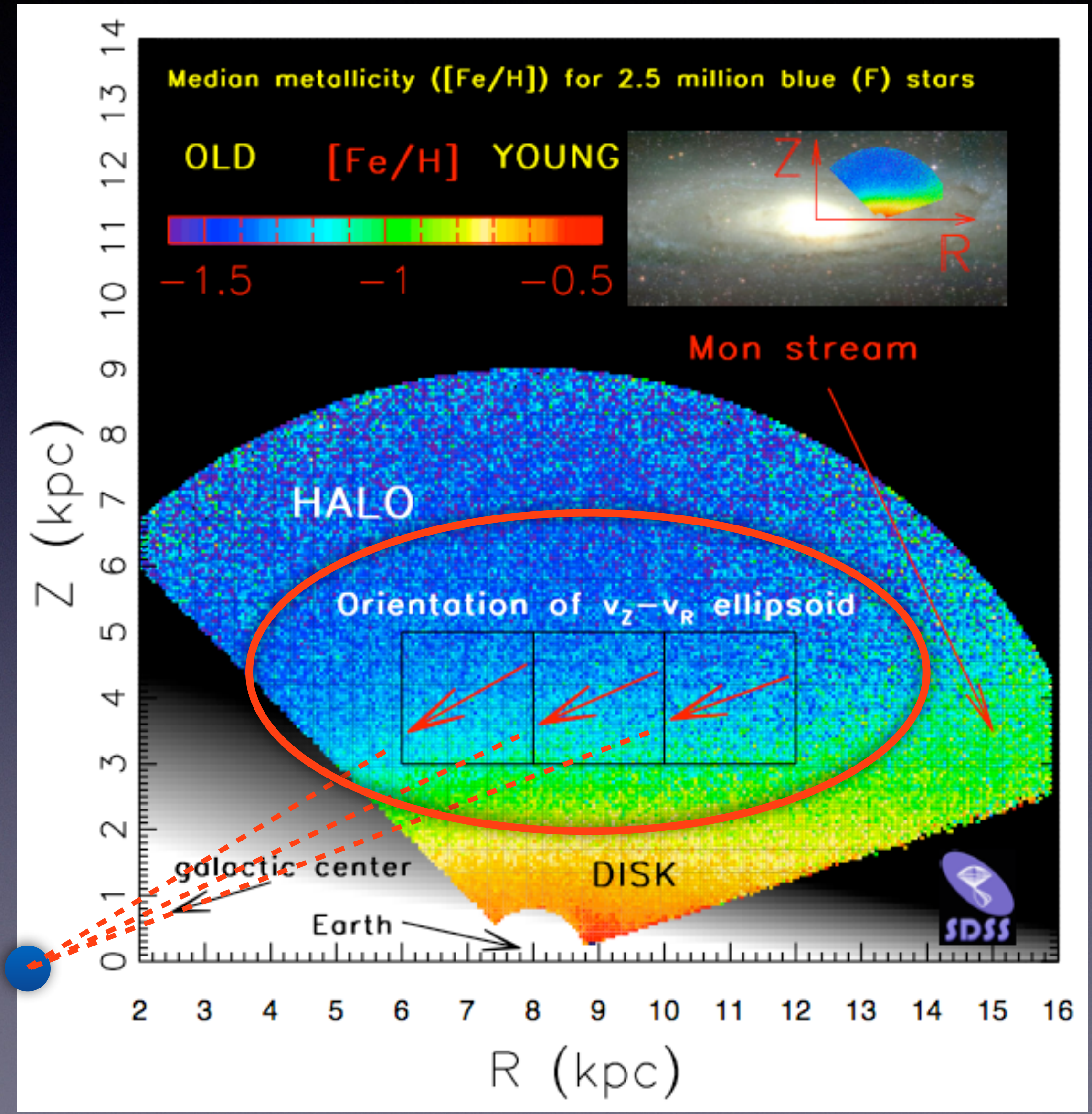
Comparison of SDSS, Gaia and LSST for main sequence stars:



Main points:

1) even Gaia will benefit from D via photom. parallax
2) LSST will have trig. parallax for very red faint stars
3) LSST will reach much further

Velocity distribution for (nearby) halo stars



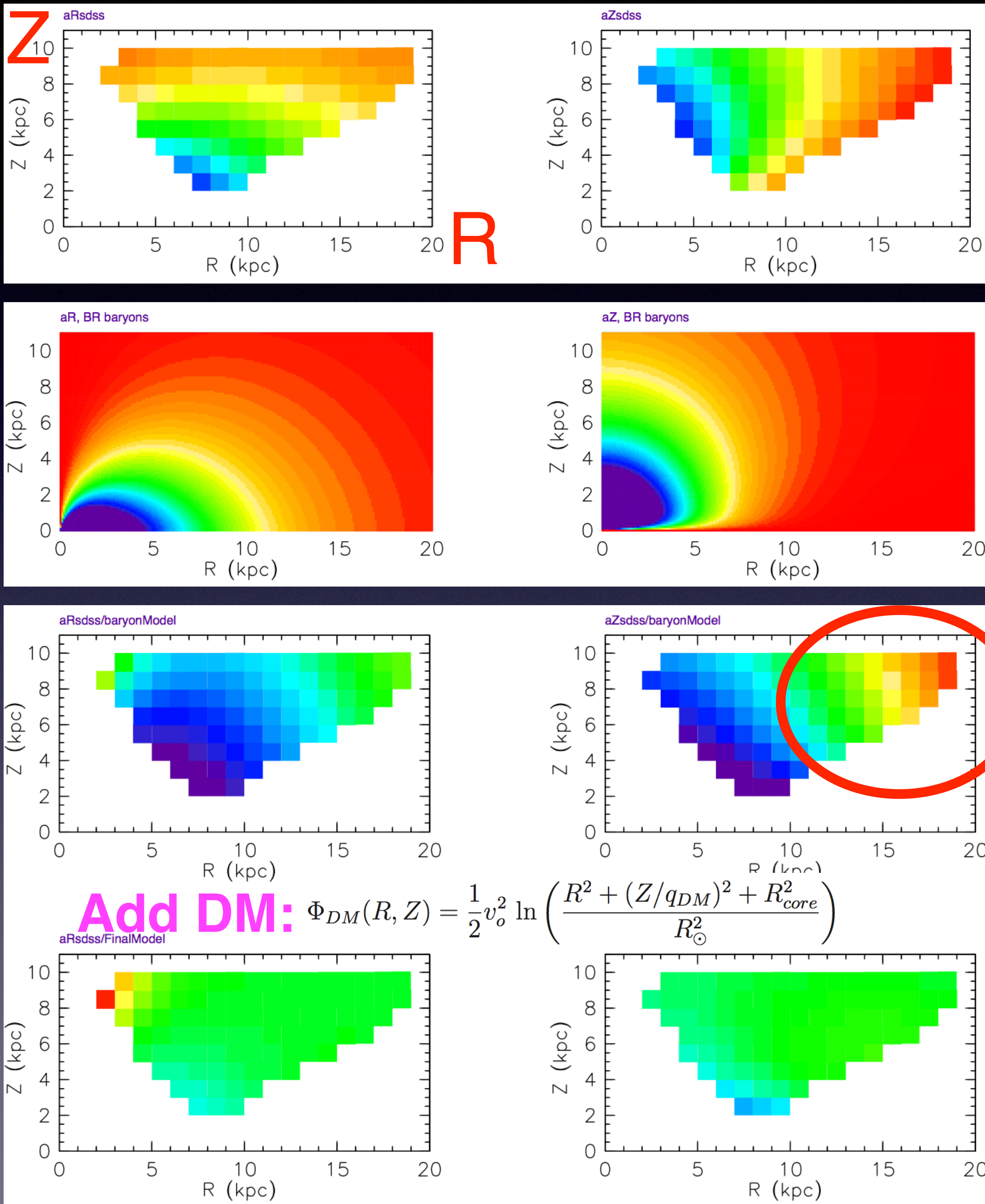
Kinematics of halo stars based on SDSS-POSS proper motions:

velocity ellipsoid is nearly invariant in spherical coordinate system

Bond et al. (2010, ApJ, 716, 1)

A side note: we should redo this map with Gaia!

aR aZ



Add DM: $\Phi_{DM}(R,Z) = \frac{1}{2} v_o^2 \ln \left(\frac{R^2 + (Z/q_{DM})^2 + R_{core}^2}{R_\odot^2} \right)$

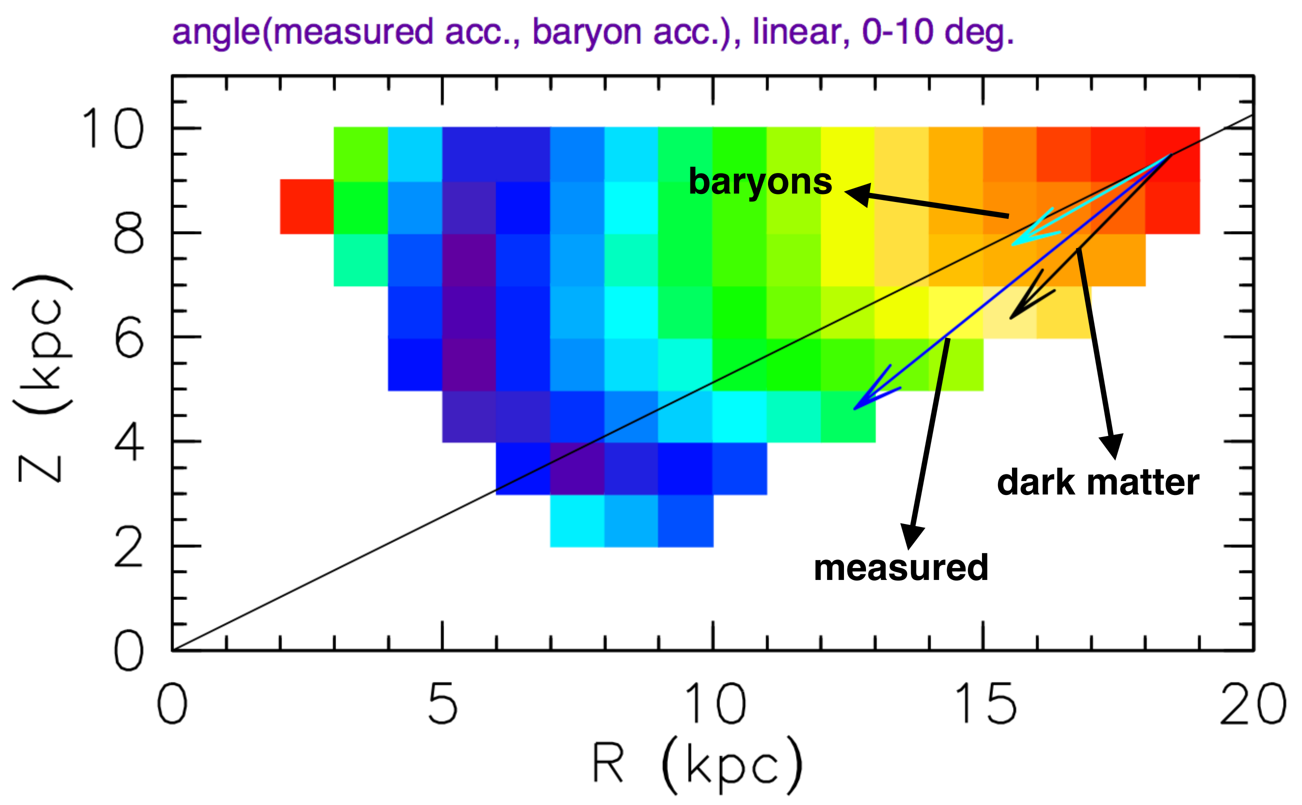
SDSS, halo, total
(Loebman et al. 2012)

Baryons (SDSS, disk)
(Bovy & Rix, 2013)

Up to 3 times stronger acc.!

SDSS measured over baryon model

DM halo is oblate!
qPot=0.7±0.1
qRho=0.4±0.1
(Loebman et al. 2014)

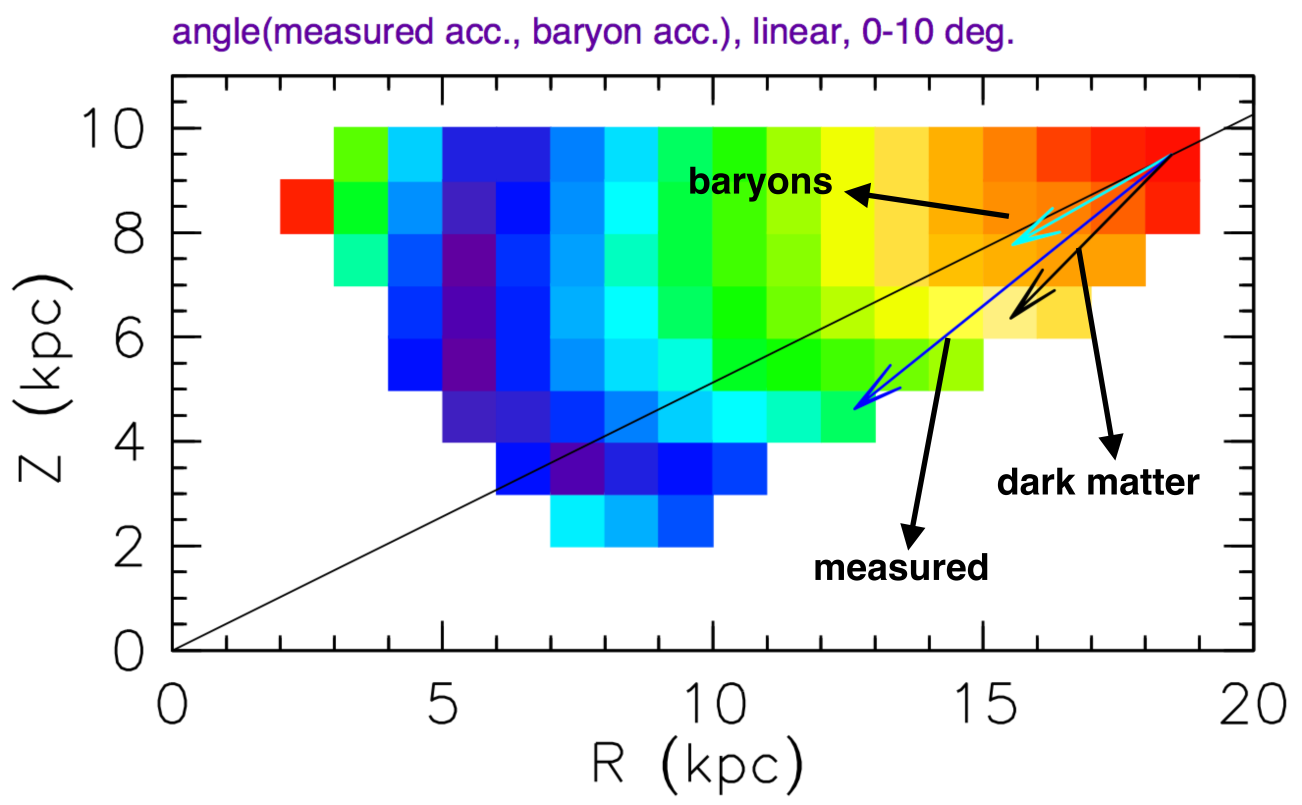


Acceler. due to baryons and measured acc. don't point in the same direction:

1) DM halo can't be spherical

2) MOND does not work

Strong constraints because of 2D acceleration measurements!



Acceler. due to baryons and measured acc. don't point in the same direction:

1) DM halo can't be spherical

2) MOND does not work

Strong constraints because of 2D acceleration measurements!

This is an example of a science program where progress of survey astronomy from SDSS to Gaia and to LSST is clearly evident.

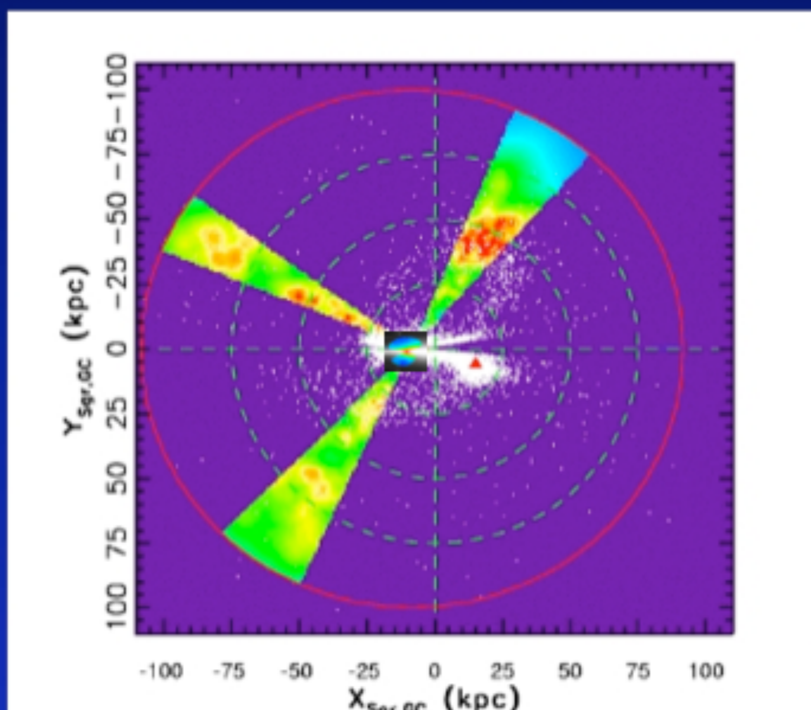
Individuals and institutions with Gaia experience will have a major advantage in using LSST data!

The large blue circle: the ~400 kpc limit of future LSST studies based on RR Lyrae

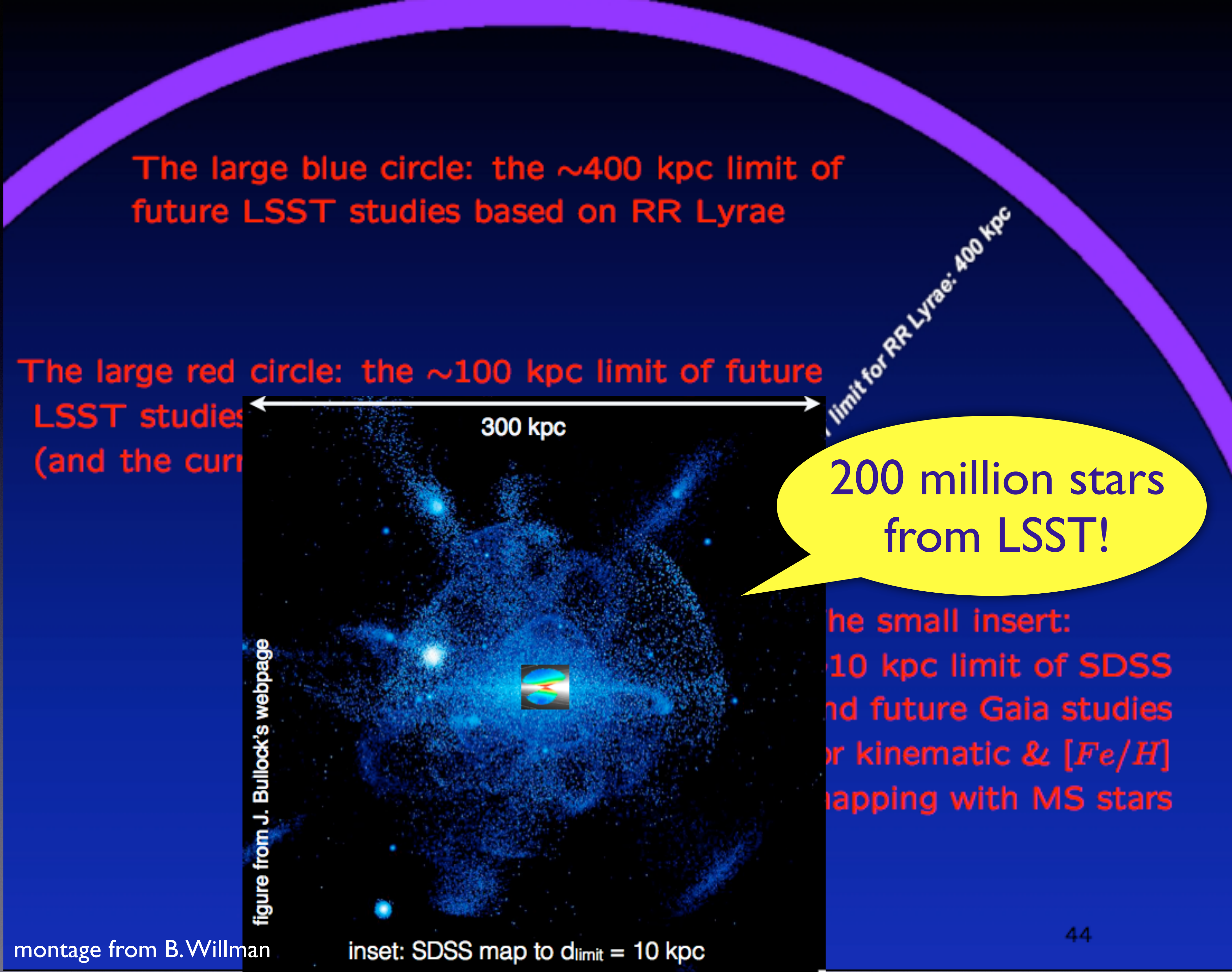
LSST limit for RR Lyrae: 400 kpc

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

6D information from LSST: 3D spatial, 2 velocities, [Fe/H]



The small insert: ~10 kpc limit of SDSS and future Gaia studies for kinematic & $[Fe/H]$ mapping with MS stars



The large blue circle: the ~400 kpc limit of
future LSST studies based on RR Lyrae
limit for RR Lyrae: 400 kpc
The large red circle: the ~100 kpc limit of future
LSST studies
(and the curr
300 kpc
200 million stars
from LSST!
he small insert:
10 kpc limit of SDSS
nd future Gaia studies
r kinematic & [Fe/H]
apping with MS stars
figure from J. Bullock's webpage
inset: SDSS map to dlimit = 10 kpc
montage from B. Willman
44

Is there a dual halo?

As always with controversial topics, we need to carefully define what we are talking about…

Existing data: “The MW stellar halo at radii <30 kpc is fairly smooth, with [Fe/H] centered on -1.5, and with no net rotation (at the measurement accuracy of 10-20 km/s).

Is there a dual halo?

As always with controversial topics, we need to carefully define what we are talking about…

Existing data: “The MW stellar halo at radii <30 kpc is fairly smooth, with [Fe/H] centered on -1.5, and with no net rotation (at the measurement accuracy of 10-20 km/s).

- **In a broad sense:** “Can we assume that the behavior seen at <30 kpc extends to the 30-100 kpc range? No, we cannot! There are enough data to directly and strongly reject this hypothesis!

- **In a narrow sense:** “Is the halo [Fe/H] distribution more metal poor and is the net rotation retrograde in the 30-100 kpc range?” Beers et al.: there is indirect (kinematic) evidence for a tentative Yes Schoenrich et al.: no, the suggested evidence is bogus.

Is there a dual halo?

As always with controversial topics, we need to carefully define what we are talking about…

Existing data: “The MW stellar halo at radii <30 kpc is fairly smooth, with [Fe/H] centered on -1.5, and with no net rotation (at the measurement accuracy of 10-20 km/s).

- **In a broad sense:** “Can we assume that the behavior seen at <30 kpc extends to the 30-100 kpc range? No, we cannot! There are enough data to directly and strongly reject this hypothesis!

- **In a narrow sense:** “Is the halo [Fe/H] distribution more metal poor and is the net rotation retrograde in the 30-100 kpc range?”
  Beers et al.: there is indirect (kinematic) evidence for a tentative Yes
  Schoenrich et al.: no, the suggested evidence is bogus.

For a direct test, we need [Fe/H] and rotational velocity measurements for many stars to beyond 30 kpc: we need proper motions and u-g colors to r < 24 or so. **LSST!**

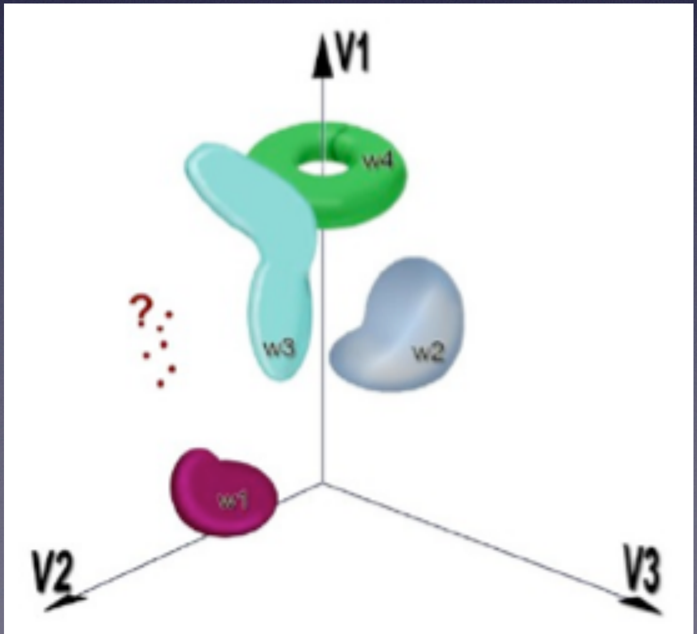
"Ask Not What Data You Need To Do Your Science, Ask What Science You Can Do With Your Data."

The era of surveys... But are our curricula reflecting it?

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

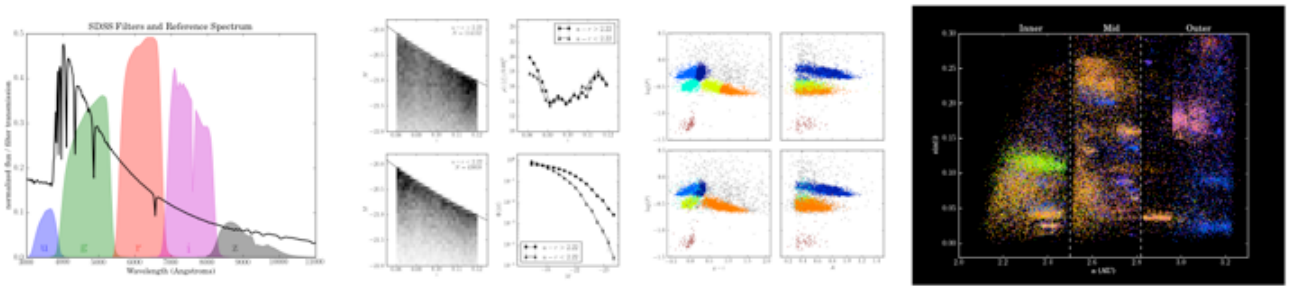
Clustering, classification, outliers

astroML

Home | User Guide | Book Figures | Examples Plots

Google™ Custom Search

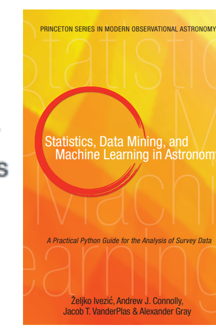
AstroML: Machine Learning and Data Mining for Astronomy



AstroML is a Python module for machine learning and data mining built on numpy, scipy, scikit-learn, and matplotlib, and distributed under the 3-clause BSD license. It contains a growing library of statistical and machine learning routines for analyzing astronomical data in python, loaders for several open astronomical datasets, and a large suite of examples of analyzing and visualizing astronomical datasets.

Downloads

- Released Versions: Python Package Index
- Bleeding-edge Source: github



The goal of astroML is to provide a community repository for fast Python implementations of common tools and routines used for statistical data analysis in astronomy and astrophysics, to provide a uniform and easy-to-use interface to freely available astronomical datasets. We hope this package will be useful to researchers and students of astronomy. The astroML project was started in 2012 to accompany the book **Statistics, Data Mining, and Machine Learning in Astronomy** by Zeljko Ivezic, Andrew Connolly, Jacob VanderPlas, and Alex Gray, to be published in late 2013. The table of contents is available here: `here(pdf)`.

User Guide

1. Introduction

- 1.1. Philosophy

News

October 2012: astroML 0.1 has been released! Get the source on Github

Our Introduction to astroML paper received the CIDU 2012 best paper award.

Links

astroML Mailing List

GitHub Issue Tracker

Videos

Scipy 2012 (15 minute talk)

Citing

If you use the software, please consider citing astroML.

Open source!
www.astroML.org

Despite the unprecedented performance of Gaia for $r < 20$, the LSST will enable major discoveries with its deep $r > 20$ sky coverage. Individuals and institutions with Gaia experience will have a major advantage in using LSST data!

"If You Liked SDSS, You will Love Gaia and LSST!"