Cosmology: from fundamental questions to computing challenges

Réza Ansari

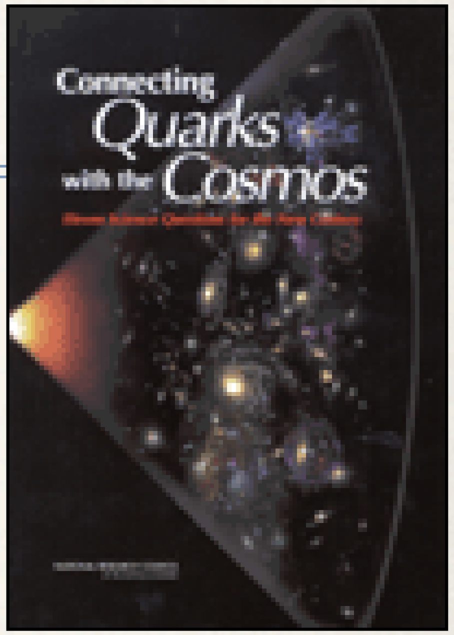
DASPAC Seminar

APC - Paris *April 2014*

- Cosmology : fundamental questions and puzzles
- Future projects: SKA & LSST
- Computing and data management challenges

Extraits de *Connecting Quarks with the Cosmos: Eleven Science Questions for the New Century (2003)*

http://www.nap.edu/books/0309074061/html/



1- What Is Dark Matter?

2- What Is the Nature of Dark Energy?

3- How Did the Universe Begin?

4- Did Einstein Have the Last Word on Gravity?

5- What Are the Masses of the Neutrinos, and How Have They Shaped the Evolution of the Universe?

6- How Do Cosmic Accelerators Work and What Are They Accelerating?

7- Are Protons Unstable?

8- What Are the New States of Matter at Exceedingly High Density and Temperature?

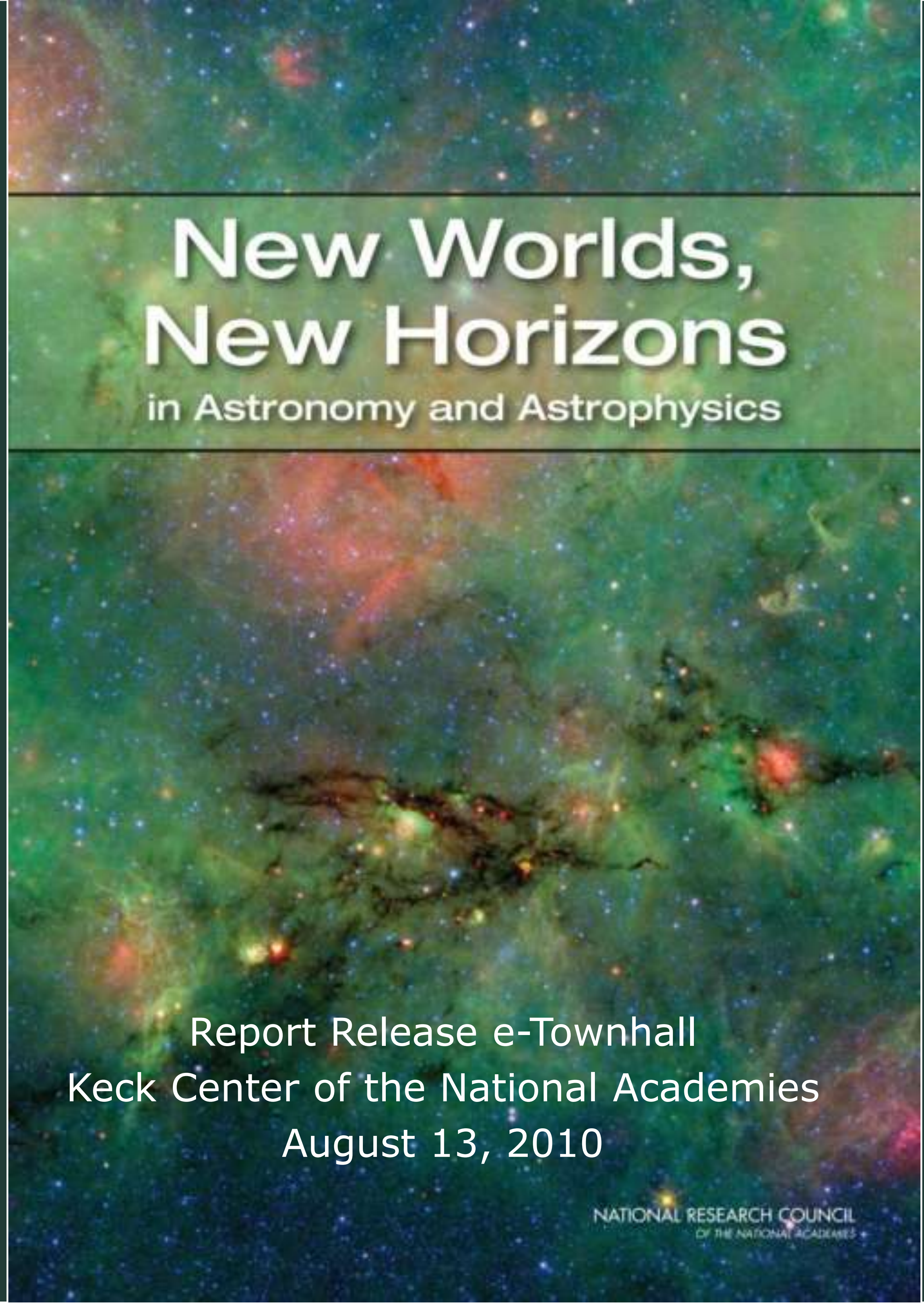
9- Are There Additional Space-Time Dimensions?

10- How Were the Elements from Iron to Uranium Made?

11- Is a New Theory of Matter and Light Needed at the Highest Energies?

2003 →

Quoting from 2010 decadal survey

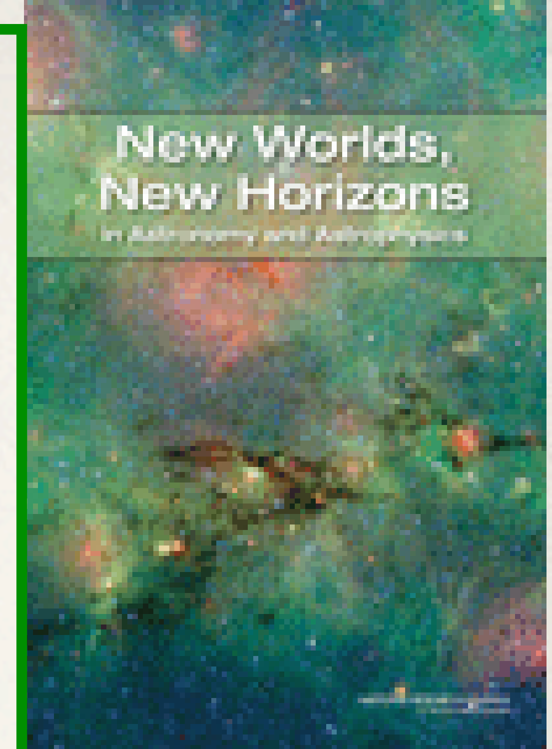


DISCOVERY

New technologies, observing strategies, theories, and computations open vistas on the universe and provide opportunities for transformational comprehension, i.e. discovery.

Science frontier discovery areas are:

- ***Identification and characterization of nearby habitable exoplanets***
- ***Gravitational wave astronomy***
- ***Time-domain astronomy***
- ***Astrometry***
- ***The epoch of reionization***



New Worls, New Horizons
Astro 2010 decadal survey

ORIGINS

Study of the origin and evolution of astronomical objects including planets, stars, galaxies, and the universe itself can elucidate our origins.

Science frontier questions in this category are:

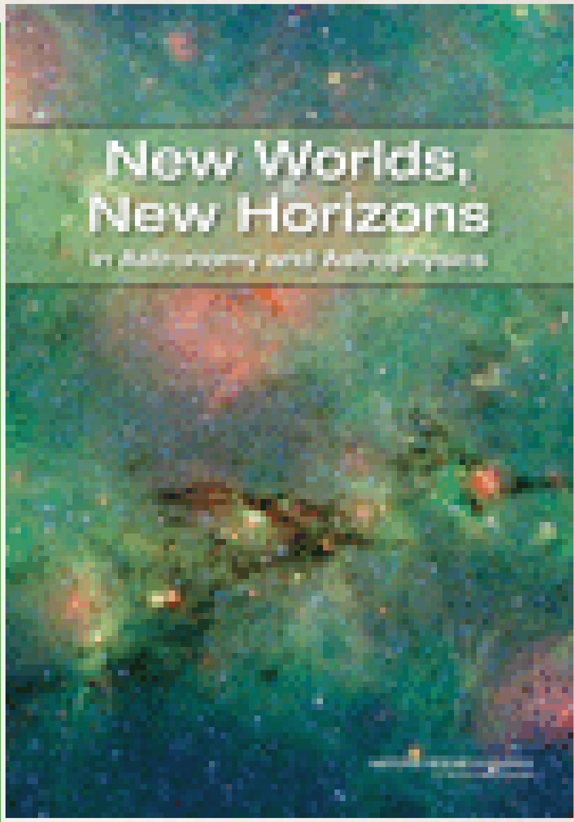
- ***How did the universe begin?***
- ***What were the first objects to light up the universe and when did they do it?***
- ***How do cosmic structures form and evolve?***
- ***What are the connections between dark and luminous matter?***
- ***What is the fossil record of galaxy assembly and evolution from the first stars to the present?***
- ***How do stars and black holes form?***
- ***How do circumstellar disks evolve and form planetary systems?***

UNDERSTANDING THE COSMIC ORDER

When known physical laws interact, often in complex ways, outcomes of great astrophysical interest and impact result and their study improves our understanding of the cosmic order.

Science frontier questions in this category are:

- ***How do baryons cycle in and out of galaxies and what do they do while they are there?***
- ***What are the flows of matter and energy in the circumgalactic medium?***
- ***What controls the mass-energy-chemical cycles within galaxies?***
- ***How do black holes work and influence their surroundings?***
- ***How do rotation and magnetic fields affect stars?***
- ***How do massive stars end their lives?***
- ***What are the progenitors of Type Ia supernovas and how do they explode?***
- ***How diverse are planetary systems and can we identify the telltale signs of life on an exoplanet?***



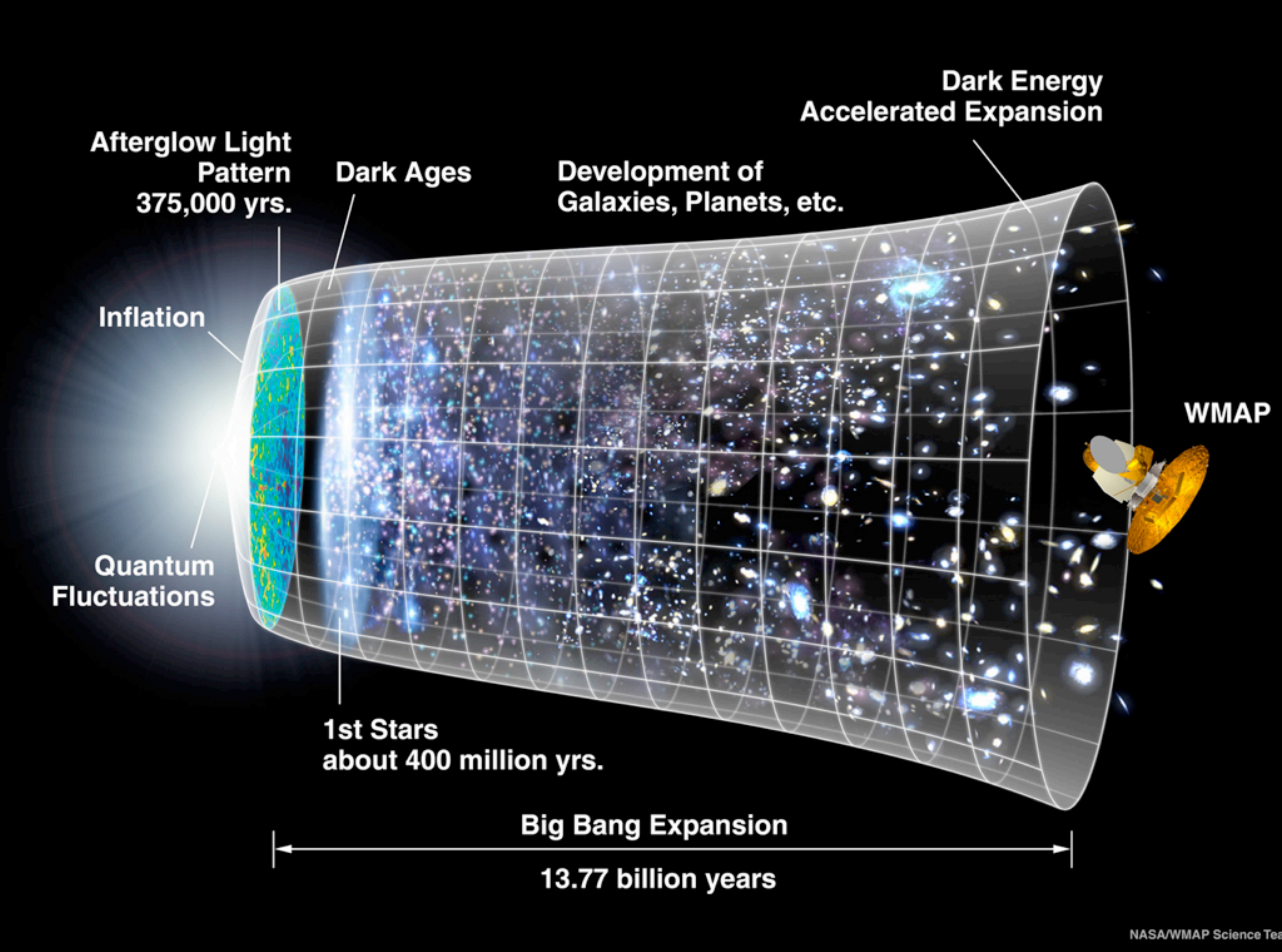
New Worls, New Horizons
Astro 2010 decadal survey

FRONTIERS OF KNOWLEDGE

New fundamental physics, chemistry, and biology can be revealed by astronomical measurements, experiments, or theory and hence push the frontiers of human knowledge.

Science frontier questions in this category are:

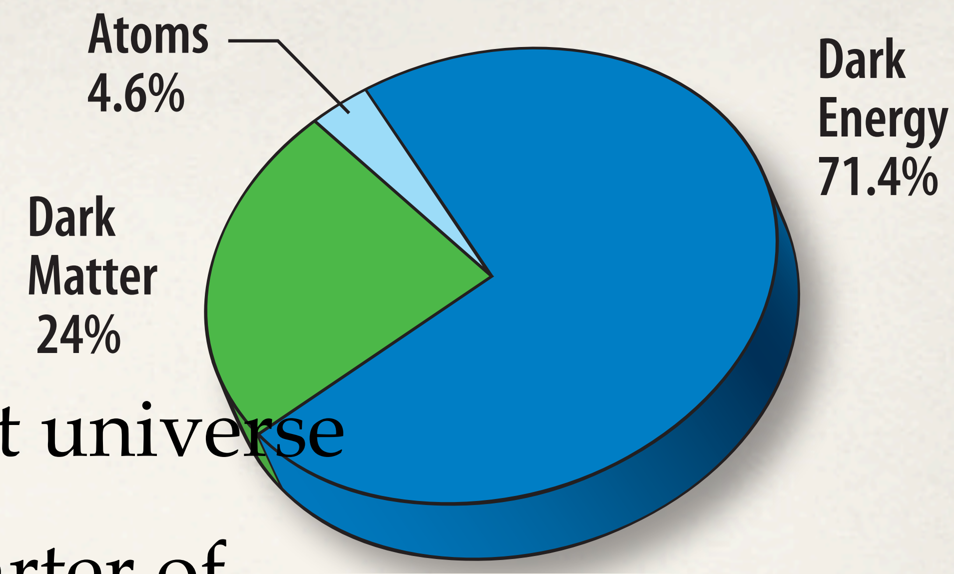
- ***Why is the universe accelerating?***
- ***What is dark matter?***
- ***What are the properties of the neutrinos?***
- ***What controls the masses, spins and radii of compact stellar remnants?***


Dark Energy
Accelerated Expansion
Afterglow Light
Pattern
375,000 yrs.
Dark Ages
Development of
Galaxies, Planets, etc.
Inflation
WMAP
Quantum
Fluctuations
1st Stars
about 400 million yrs.
Big Bang Expansion
13.77 billion years
NASA/WMAP Science Tea

Cosmology: main questions and tools

- Energy and matter content of the universe (Dark matter, dark energy)
- Structure formation and evolution
- Primordial cosmology: inflation …
- Primordial nucleosynthesis
- Formation and evolution of galaxies and stars
- Cosmic microwave background: temperature and polarisation C(l) spectrum
- Statistical properties of large scale structures
- Geometrical probes: $d_A(z)$, $d_L(z)$ … : SNIa , Clusters, BAO …

Dark Energy (or Λ)



WMAP - 9y

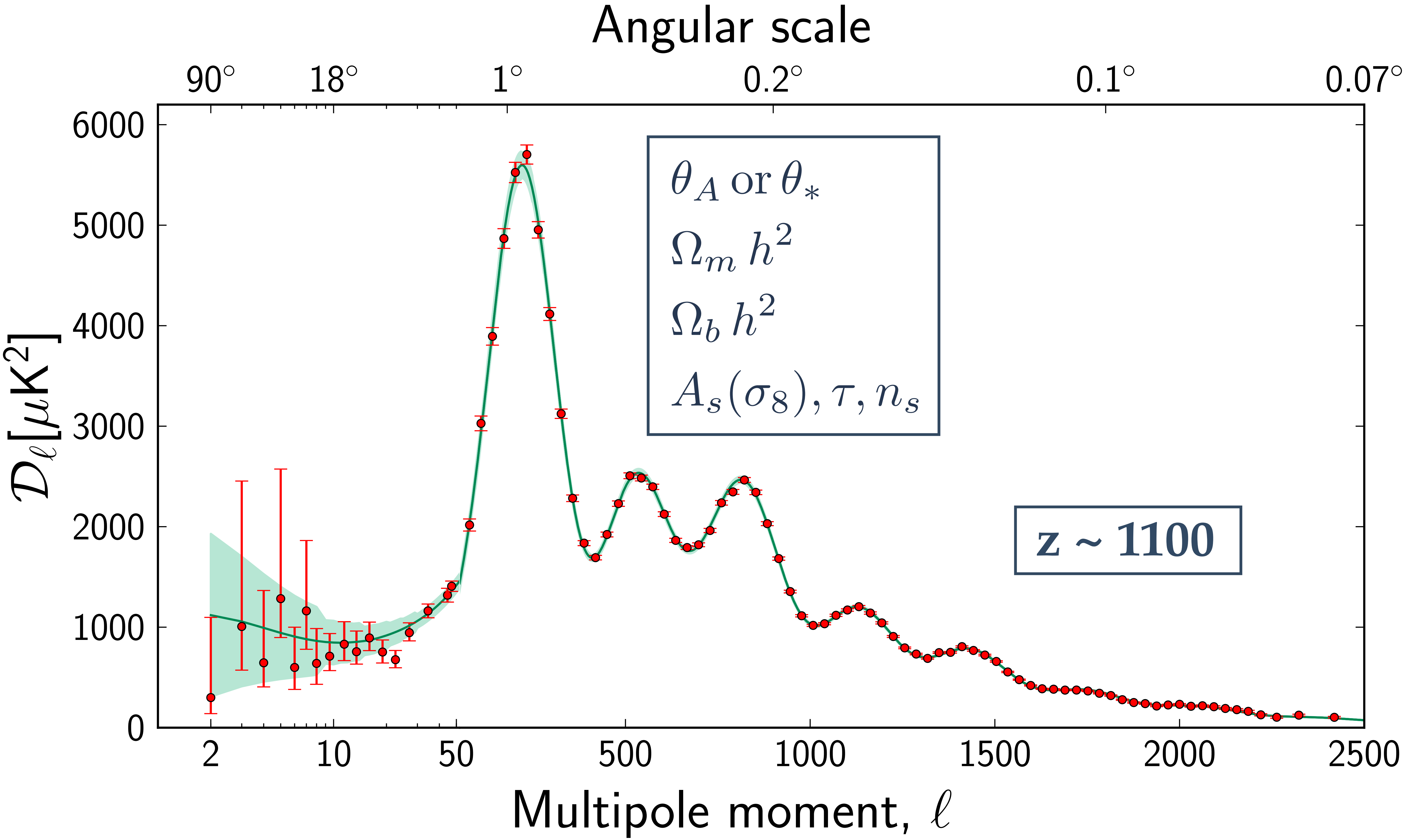
- ✤Recent cosmological observations imply a flat universe
- ✤Matter (including dark matter) is about a quarter of the critical density. Most of the energy density in the universe seems to be made of a mysterious component behaving like Λ
- ✤Λ : Repulsive gravity !
- ✤Vaccum energy (quantum fluctuations) ➠ Dark Energy ?
- ✤Determination of state equation of this cosmic fluid:

$$p = w(z)\,\varrho$$

- ✤w(z) = -1 for the cosmological constant (Λ)

http://www.nsf.gov/mps/ast/detf.jsp

A



Planck TT power spectrum , Planck collaboration arXiv 1303.5075

8. Constraints on the basic six-parameter ΛCDM model using *Planck* data. The top section contains constraints on the si[x]
[primar]y parameters included directly in the estimation process, and the bottom section contains constraints on derived parameters.

	Planck		*Planck*+WP	
Parameter	Best fit	68% limits	Best fit	68% limits
$\Omega_b h^2$	0.022068	0.02207 ± 0.00033	0.022032	0.02205 ± 0.00028
$\Omega_c h^2$	0.12029	0.1196 ± 0.0031	0.12038	0.1199 ± 0.0027
$100\theta_{MC}$	1.04122	1.04132 ± 0.00068	1.04119	1.04131 ± 0.00063
τ	0.0925	0.097 ± 0.038	0.0925	$0.089^{+0.012}_{-0.014}$
n_s	0.9624	0.9616 ± 0.0094	0.9619	0.9603 ± 0.0073
$\ln(10^{10} A_s)$	3.098	3.103 ± 0.072	3.0980	$3.089^{+0.024}_{-0.027}$
Ω_Λ	0.6825	0.686 ± 0.020	0.6817	$0.685^{+0.018}_{-0.016}$
Ω_m	0.3175	0.314 ± 0.020	0.3183	$0.315^{+0.016}_{-0.018}$
σ_8	0.8344	0.834 ± 0.027	0.8347	0.829 ± 0.012
z_{re}	11.35	$11.4^{+4.0}_{-2.8}$	11.37	11.1 ± 1.1
H_0	67.11	67.4 ± 1.4	67.04	67.3 ± 1.2
$10^9 A_s$	2.215	2.23 ± 0.16	2.215	$2.196^{+0.051}_{-0.060}$
$\Omega_m h^2$	0.14300	0.1423 ± 0.0029	0.14305	0.1426 ± 0.0025
Age/Gyr	13.819	13.813 ± 0.058	13.8242	13.817 ± 0.048
z_*	1090.43	1090.37 ± 0.65	1090.48	1090.43 ± 0.54
$100\theta_*$	1.04139	1.04148 ± 0.00066	1.04136	1.04147 ± 0.00062
z_{eq}	3402	3386 ± 69	3403	3391 ± 60

Planck cosmological parameters , Planck collaboration arXiv 1303.5075

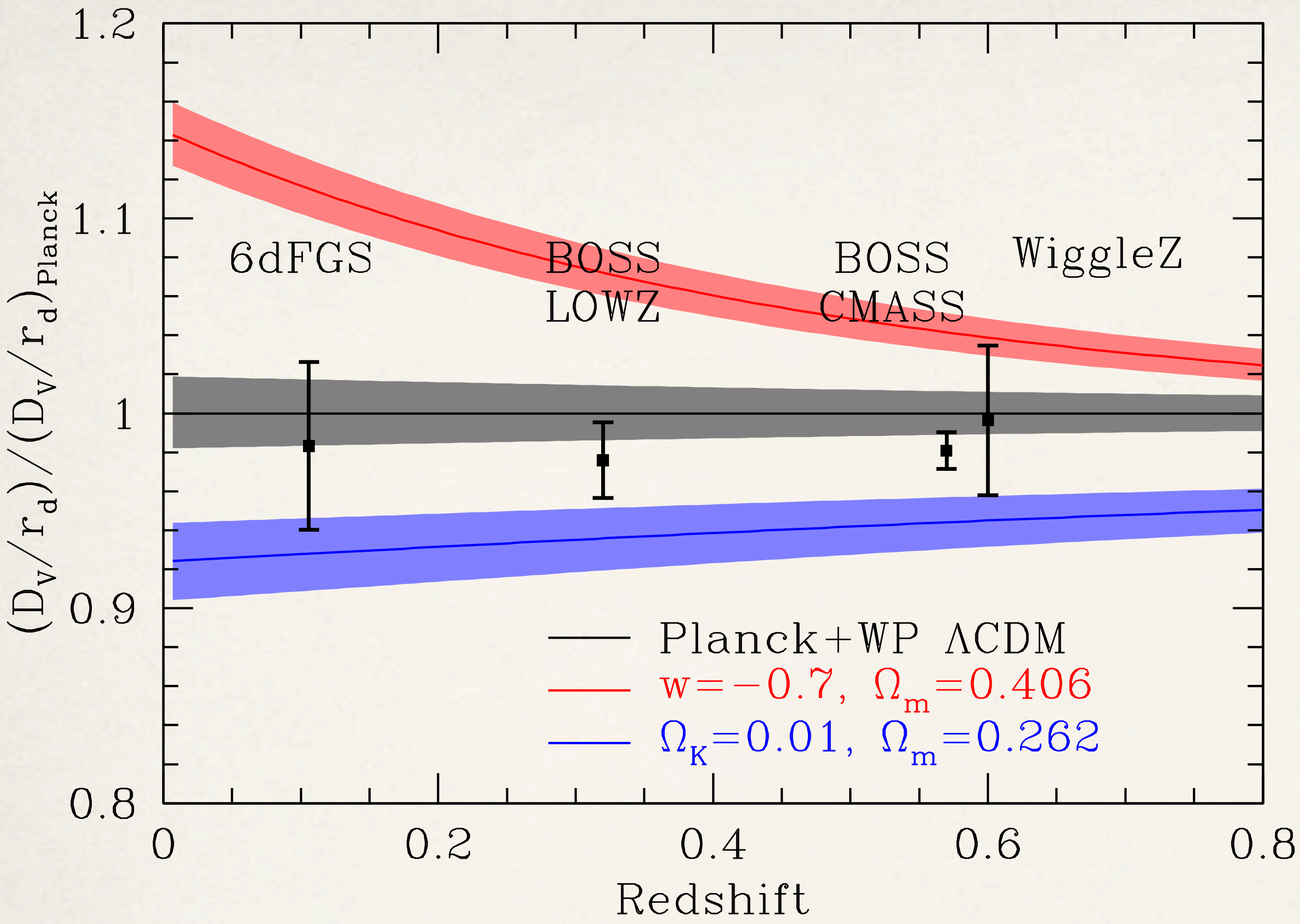


Figure 24. The $D_V(z)/r_d$ measured from galaxy surveys, divided by the best-fit flat ΛCDM prediction from the Planck data. All error bars are 1 σ. We now vary the cosmological model for the Planck prediction. Red shows the prediction assuming a flat Universe with $w = -0.7$; blue shows the prediction assuming a closed Universe with $\Omega_K = -0.01$ and a cosmological constant.

⬆ SDSS-III BOSS ⬆

Anderson et al, arXiv 1312.4877

Planck (Cosmo. Param)

⬇ arXiv 1303.5076 ⬇

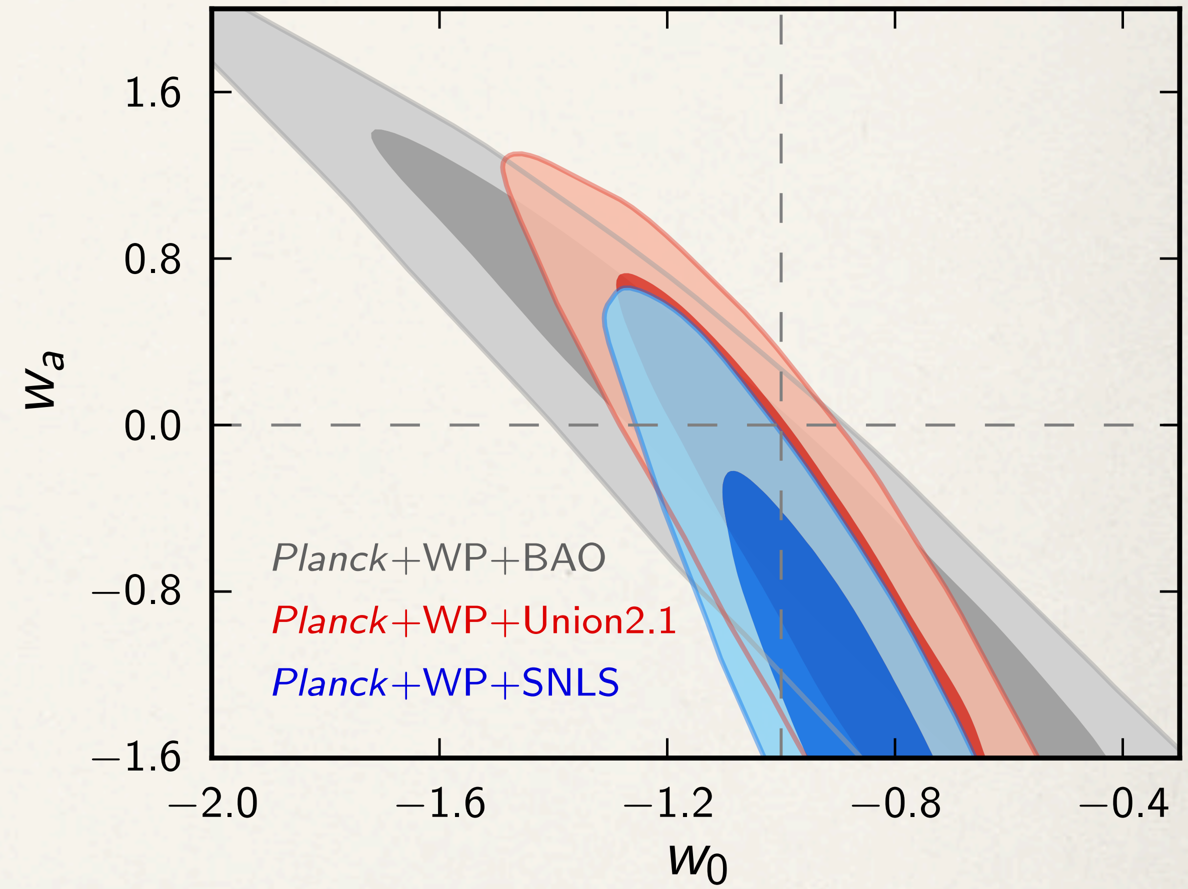
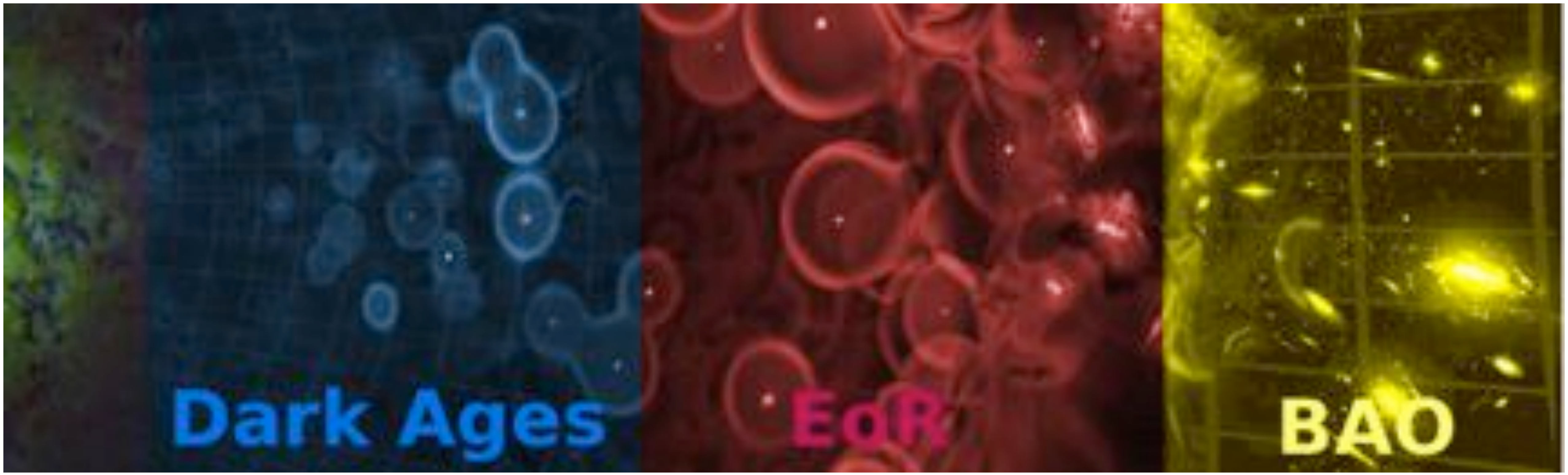


Fig. 36. 2D marginalized posterior distributions for w_0 and w_a, for the data combinations *Planck*+WP+BAO (grey), *Planck*+WP+Union2.1 (red) and *Planck*+WP+SNLS (blue). The contours are 68% and 95%, and dashed grey lines show the cosmological constant solution.

21 cm Cosmology projects



z = 100 to 20
15 to 70 MHz

z = 20 to 5
70 to 240 MHz

z = 5 to 0
240 to 1400 MHz

20 — 5 — z

- LOFAR
- GMRT
- MWA

- SKA-LOW
- HERA

- **CHIME**
- **Tianlai**
- GBT
- BAOBab
- SKA-mid

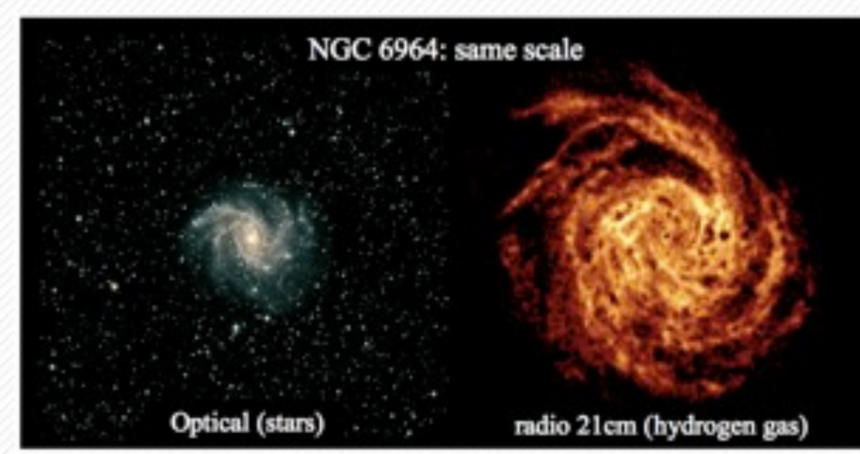
SKA
SQUARE KILOMETRE ARRAY

The Square Kilometre Array
Exploring the Universe with the world's largest radio telescope

Five Key Science projects

- **Galaxy evolution, Cosmology and Dark Energy**
- Strong-field tests of gravity using pulsars and black holes
- Origin and evolution of cosmic magnetism
- Probing the dark ages
- Cradle of life

What is mysterious dark energy?



How are galaxies born and how do they evolve?

Great Observatories for the coming decades

Slide by R. Braun
SKA Science director

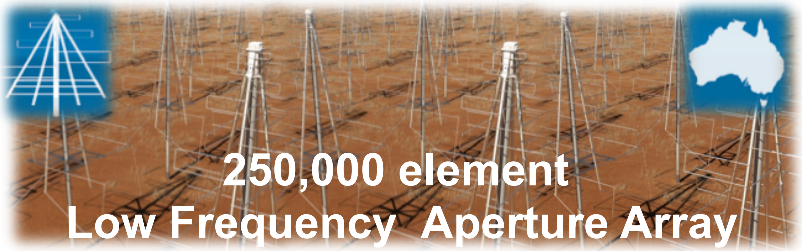
Square Kilometre Array: radio
Construction start 2017/18

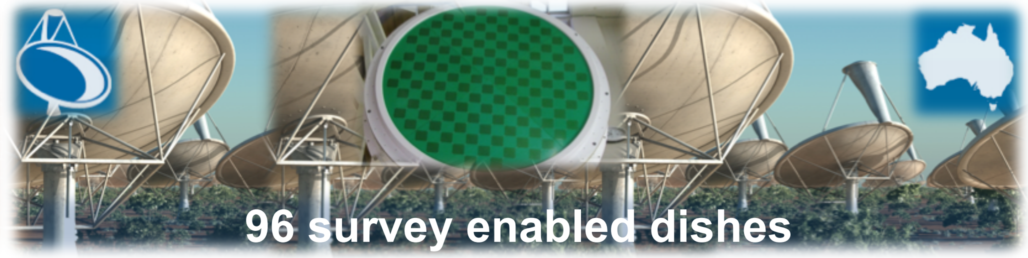
Exploring the Universe with the world's largest radio telescope

SQUARE KILOMETRE ARRAY

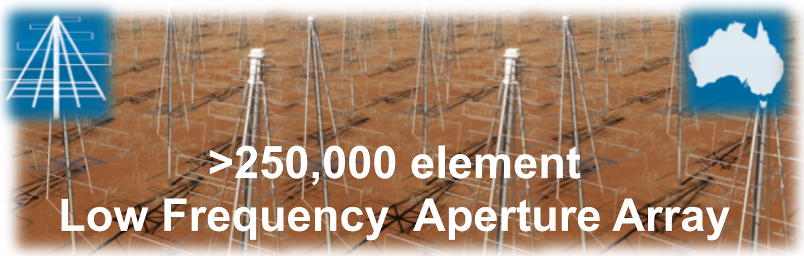
Slide by R. Braun
SKA Science director

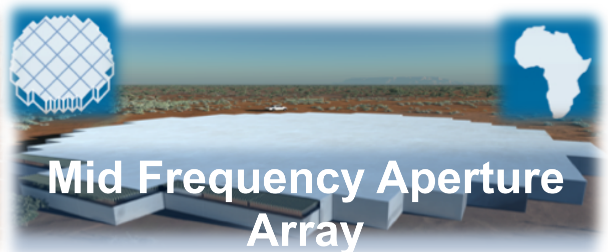
Phase I : 2020



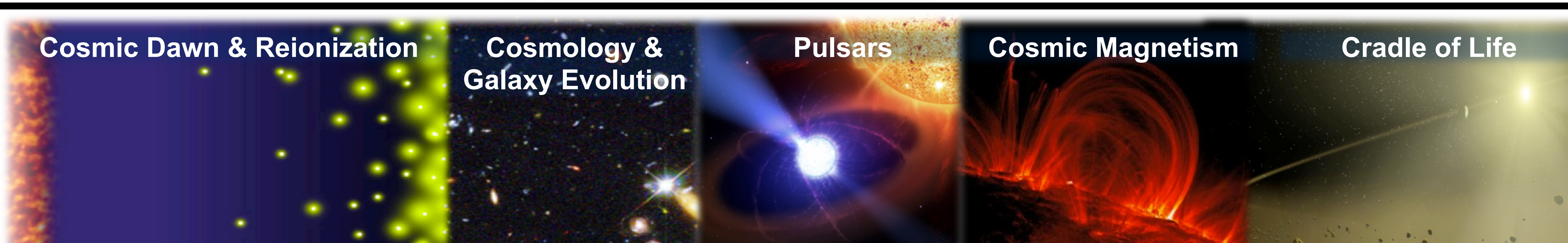


Phase II : 2024





Science



50 MHz **100 MHz** **1 GHz** **10 GHz**

How does SKA1 baseline redefine state-of-art?

		JVLA	MeerKAT	SKA1-mid	ASKAP	SKA1-survey	LOFAR-NL	SKA1-low
Aeff/Tsys	m^2/K	265	321	1630	65	391	61	1000
Survey FoV	deg^2	0.14	0.48	0.39	30	18	6	6
Survey Speed FoM	$deg^2\ m^4\ K^{-2}$	0.98×10^4	5.0×10^4	1.0×10^6	1.3×10^5	2.8×10^6	2.2×10^4	6.0×10^6
Resolution	arcsec	1.4	11	0.22	7	0.9	5	11

A_{eff}/T_{sys}: 6xJVLA 6xASKAP 16xLOFAR

Survey Speed: 100x 22xASKAP 270x

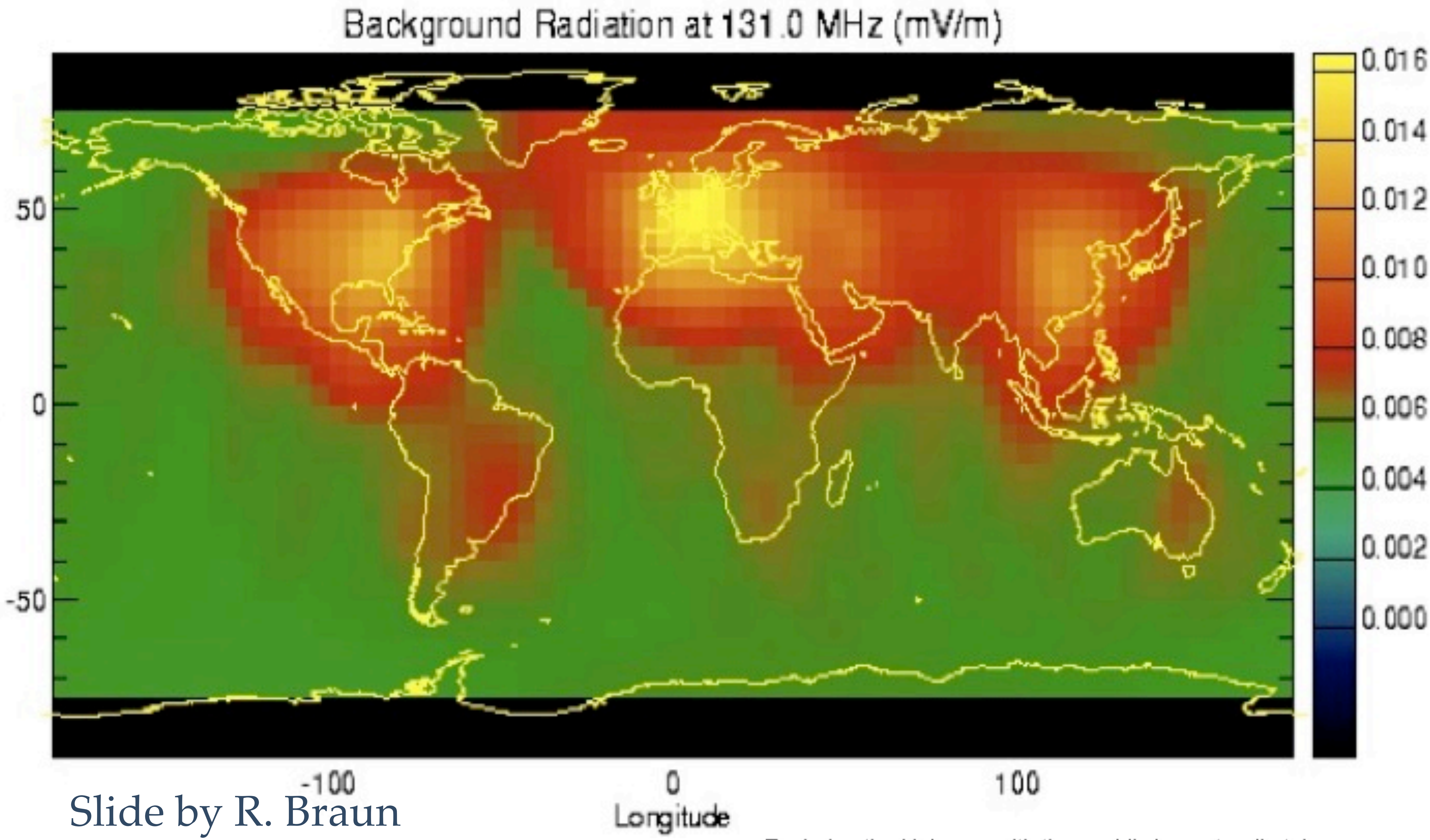
280xJVLA

Slide by R. Braun

SKA Science director

Exploring the Universe with the world's largest radio telescope

How did we choose the site?

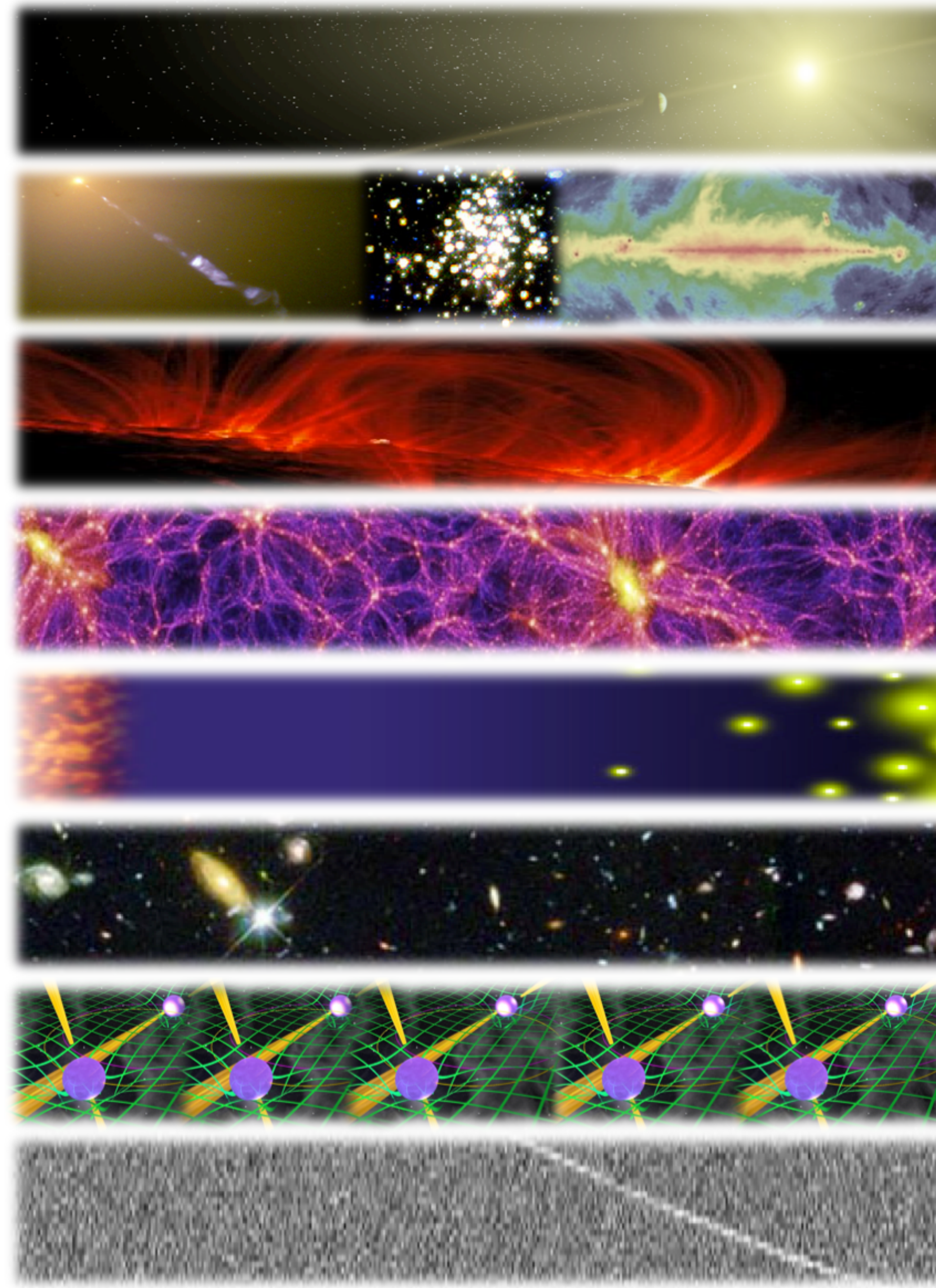


Slide by R. Braun
SKA Science director

Exploring the Universe with the world's largest radio telescope

The Science Working Groups

- **Astrobiology (“The Cradle of Life”)**
  – *Project Scientist:* Tyler Bourke
  – *Working Group Chair:* Melvin Hoare
- **Galaxy Evolution – Continuum**
  – *Project Scientist*: Jeff Wagg
  – *Working Group Chairs:* Nick Seymour & Isabella Prandoni
- **Cosmic Magnetism**
  – *Project Scientist:* Jimi Green
  – *Working Group Chairs:* Melanie Johnston-Hollitt & Federica Govoni
- **Cosmology**
  – *Project Scientist:* Jeff Wagg
  – *Working Group Chair:* Roy Maartens
- **Epoch of Reionisation & the Cosmic Dawn**
  – *Project Scientist:* Jeff Wagg
  – *Working Group Chair:* Leon Koopmans
- **Galaxy Evolution – HI**
  – *Project Scientist:* Jimi Green
  – *Working Group Chairs:* Lister Staveley-Smith & Tom Osterloo
- **Pulsars (“Strong field tests of gravity”)**
  – *Project Scientist:* Jimi Green
  – *Working Group Chairs:* Ben Stappers & Michael Kramer
- **Transients**
  – *Project Scientist:* Tyler Bourke
  – *Working Group Chair:* Rob Fender



Exploring the Universe with the world’s largest radio telescope

Slide by R. Braun
SKA Science director

LSST


LSST (Large Synoptic Survey Telescope)
Wide … Fast … Deep

LSST "MISSION"

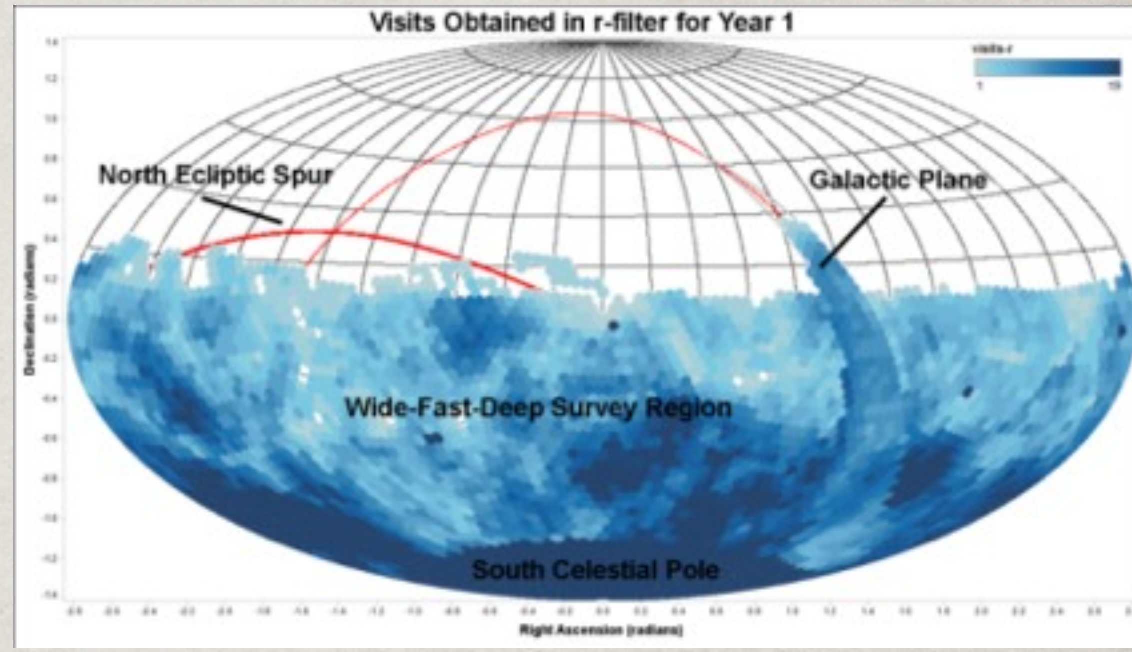
- Photometric survey of half of the sky (~ 20 000 sq. deg.) during 10 years
- Complete coverage every 4 nights
- One 10 sq. deg. field every 40 seconds
- Fast alert system (60s) for detection of violent phenomenon

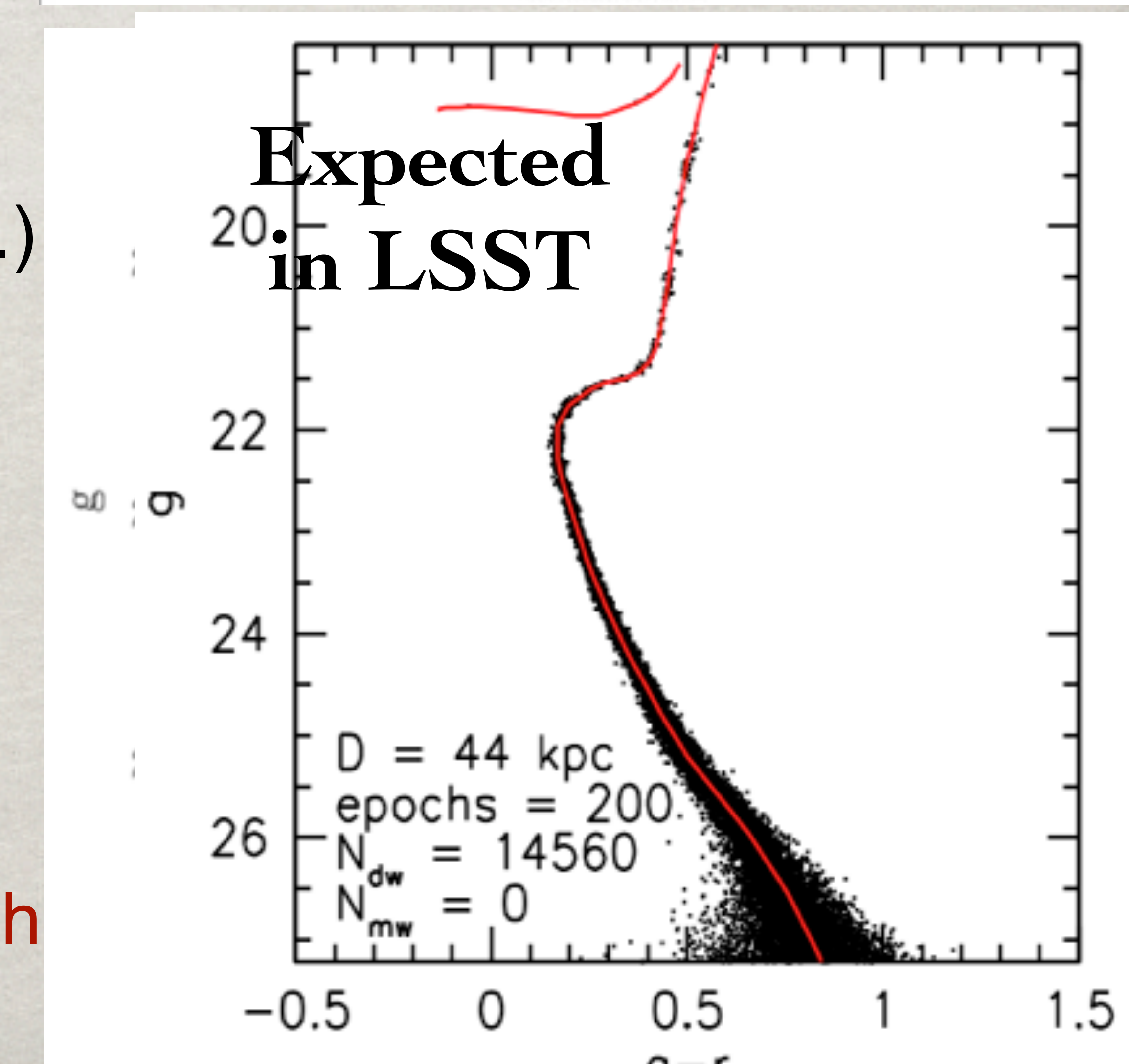
<u>Deliverables</u>

- « 4D » object mapping (stars, galaxies...)
  (α,δ) positions on the sky
  Redshifts z
  Time variations (SN, lensing, AGN…)

<u>Cosmology</u>

- Archive more than 3×10^9 galaxies with photometric redshifts up to z=3
- Detection of 250 000 SN Ia per year (with photo-z < 0.8).

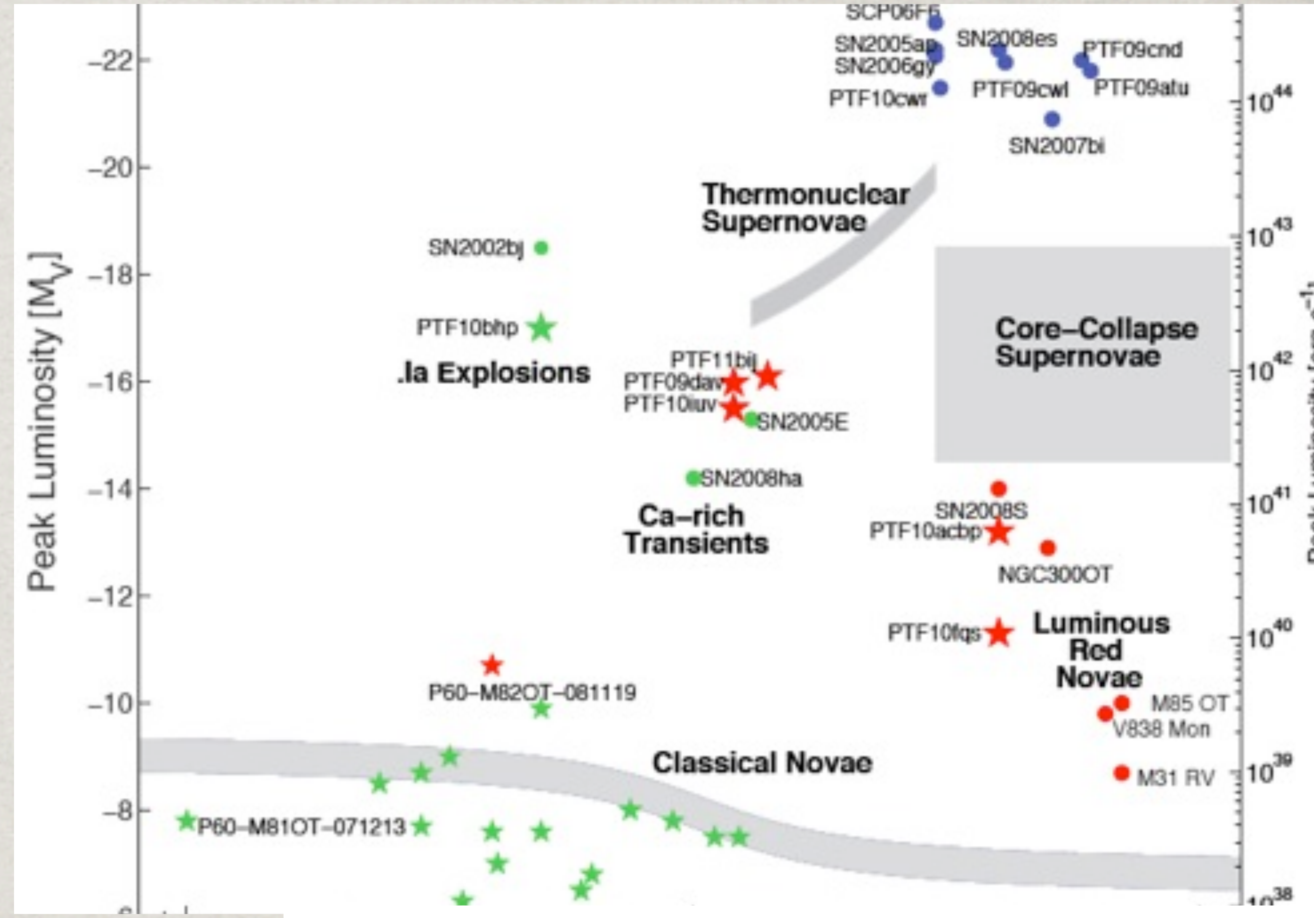


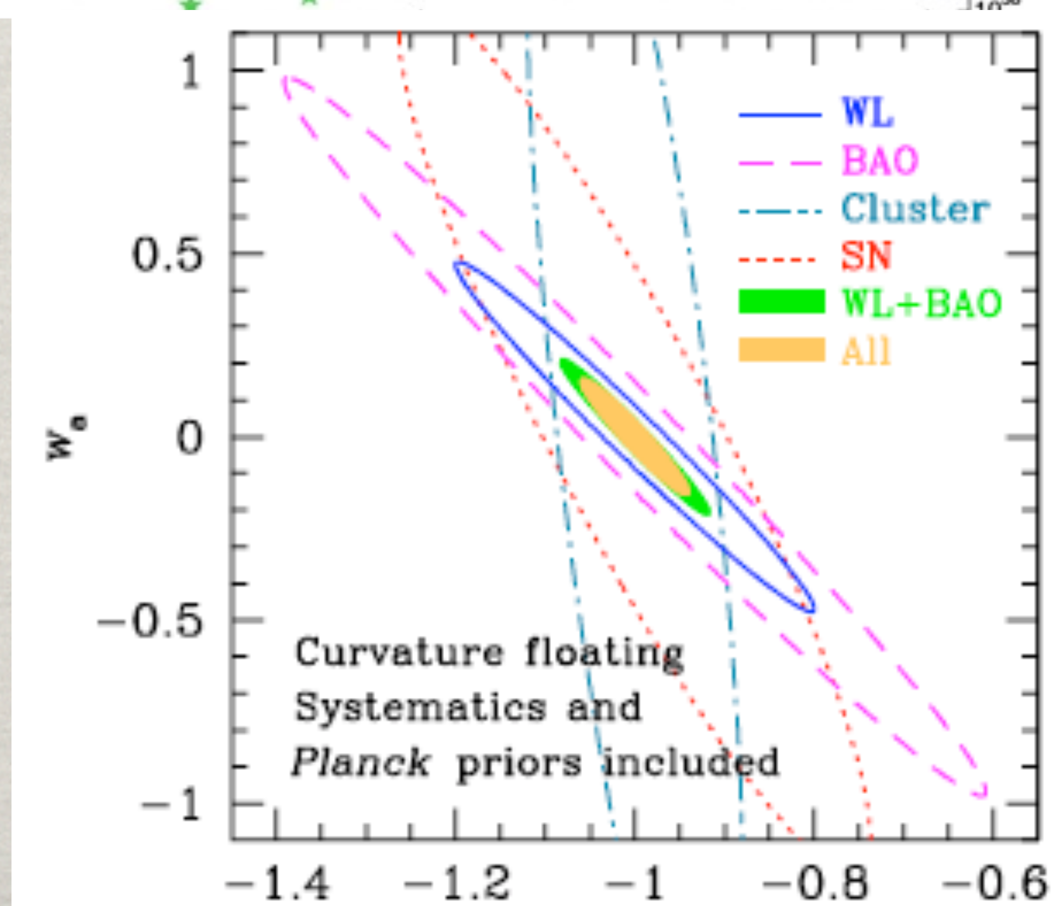


R. Ansari , CC-IN2P3, Jan 2014

The Science Enabled by LSST

- Time domain science
  - Nova, supernova, GRBs
  - Source characterization
  - Gravitational microlensing
  - Interstellar scintillation
- Finding moving sources
  - Asteroids and comets
  - Proper motions of stars
- Mapping the Milky Way
  - Tidal streams
  - Galactic structure
- Dark energy and dark matter
  - Gravitational lensing
  - Slight distortion in shape
  - Trace the nature of dark energy

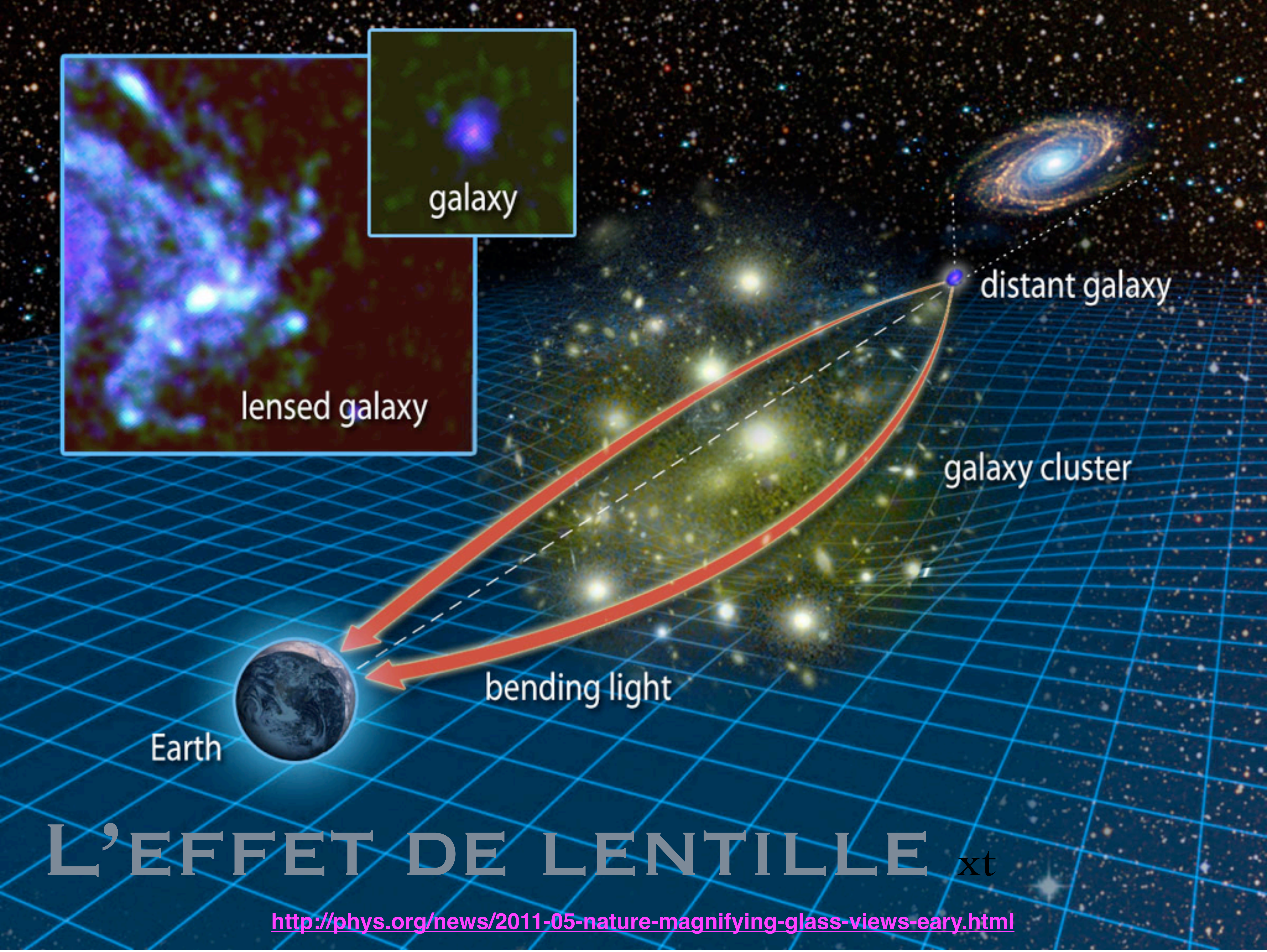




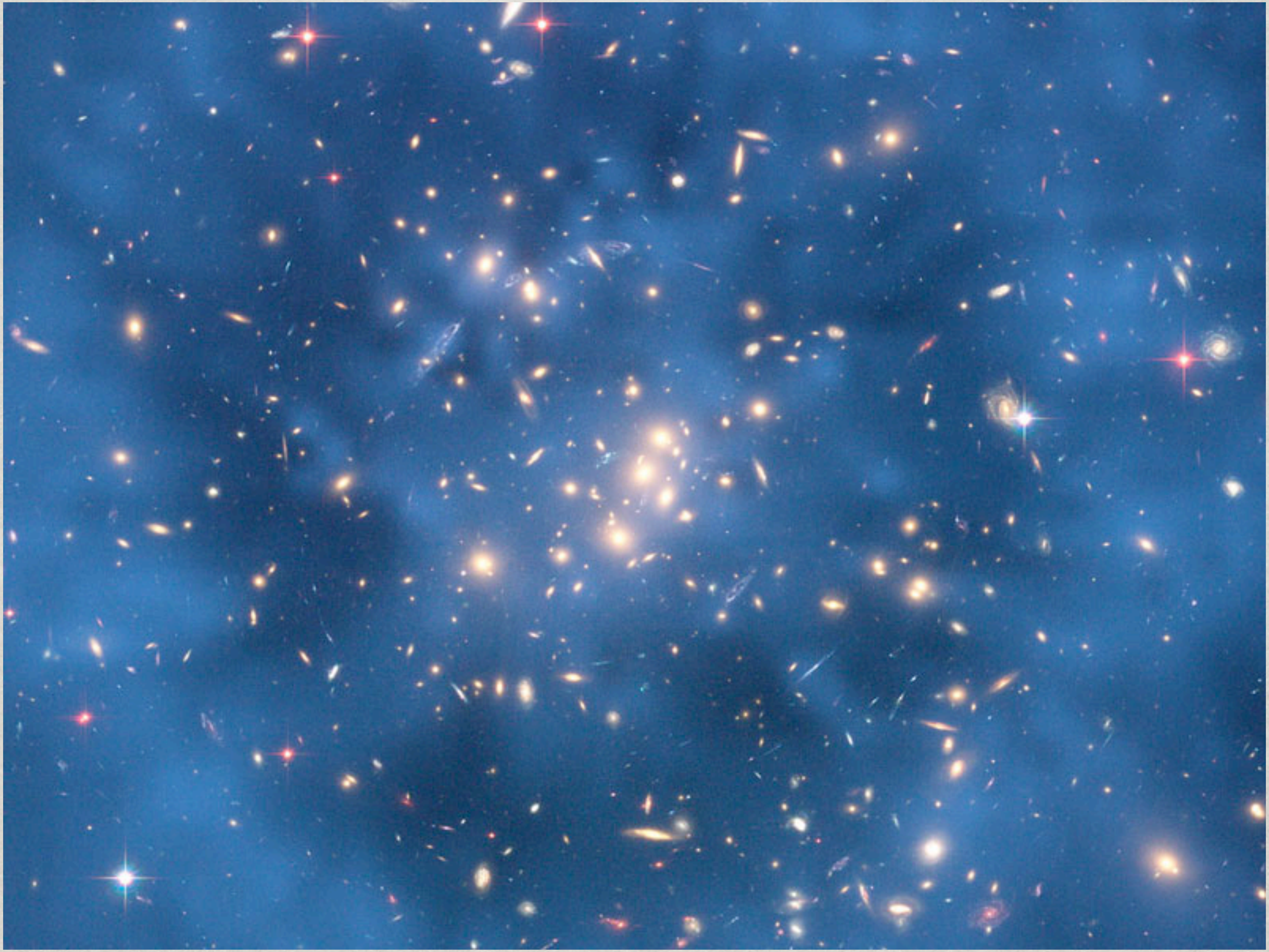
R. Ansari , CC-IN2P3, Jan 2014

Large Synoptic Survey Telescope

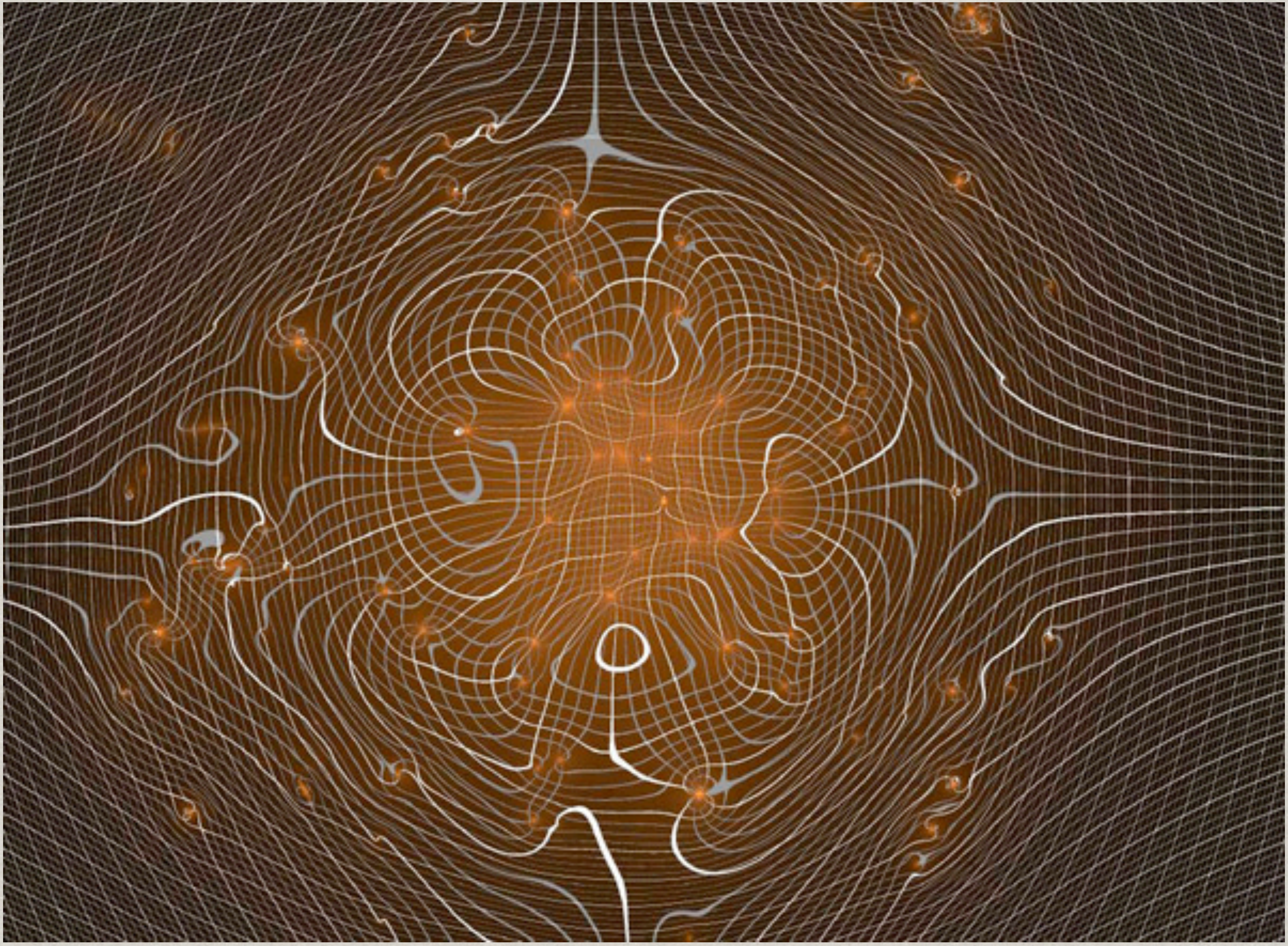
L'EFFET DE LENTILLE xt



http://phys.org/news/2011-05-nature-magnifying-glass-views-eary.html

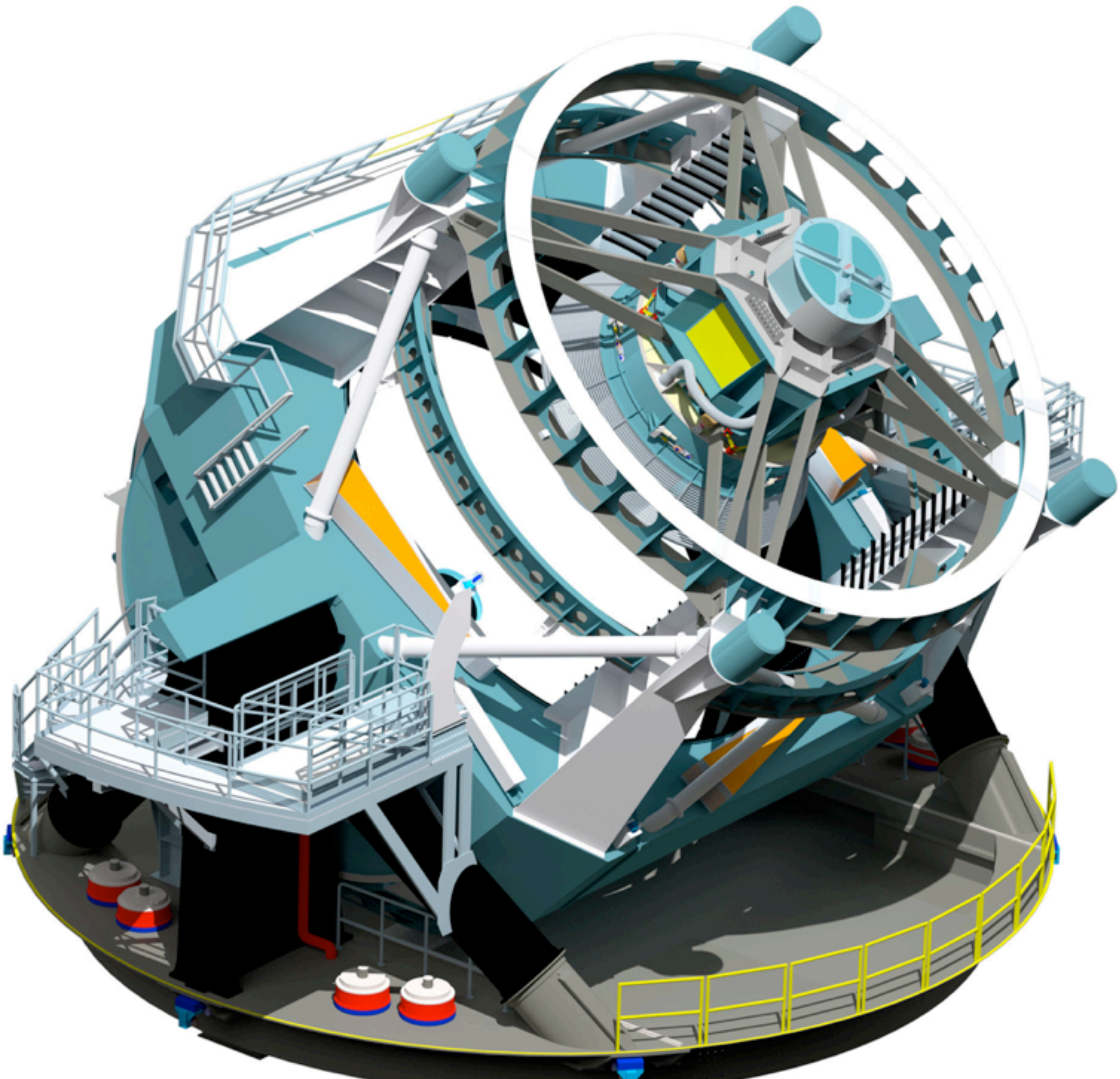


Amas de galaxies CL0024+17 (z ~ 0.39 , ~ 5. 10^9 AL) vu par HST



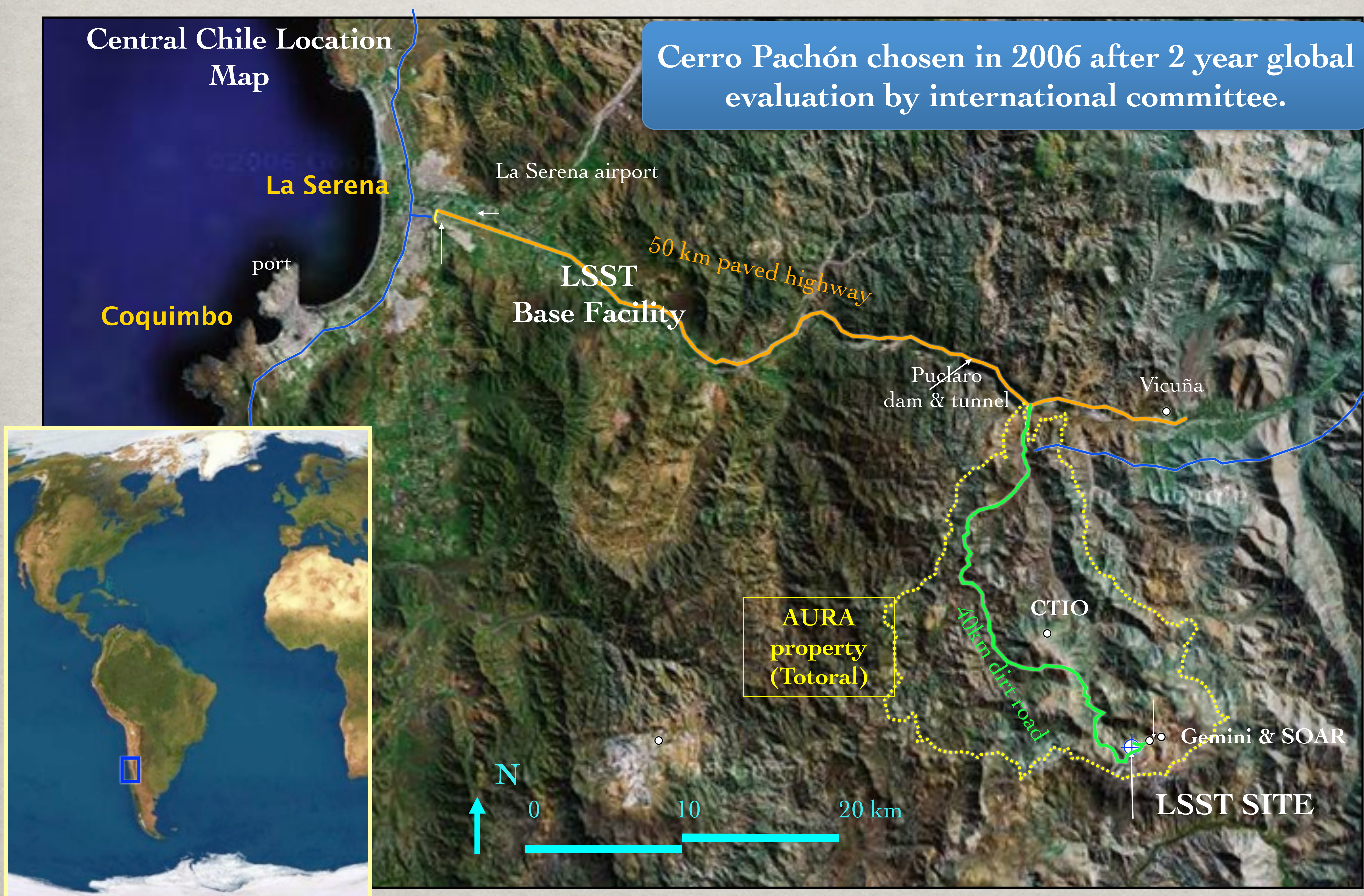
Reconstitution du potentiel gravitationnel (matière noire) et distortion de l'espace-temps

LSST : l'instrument

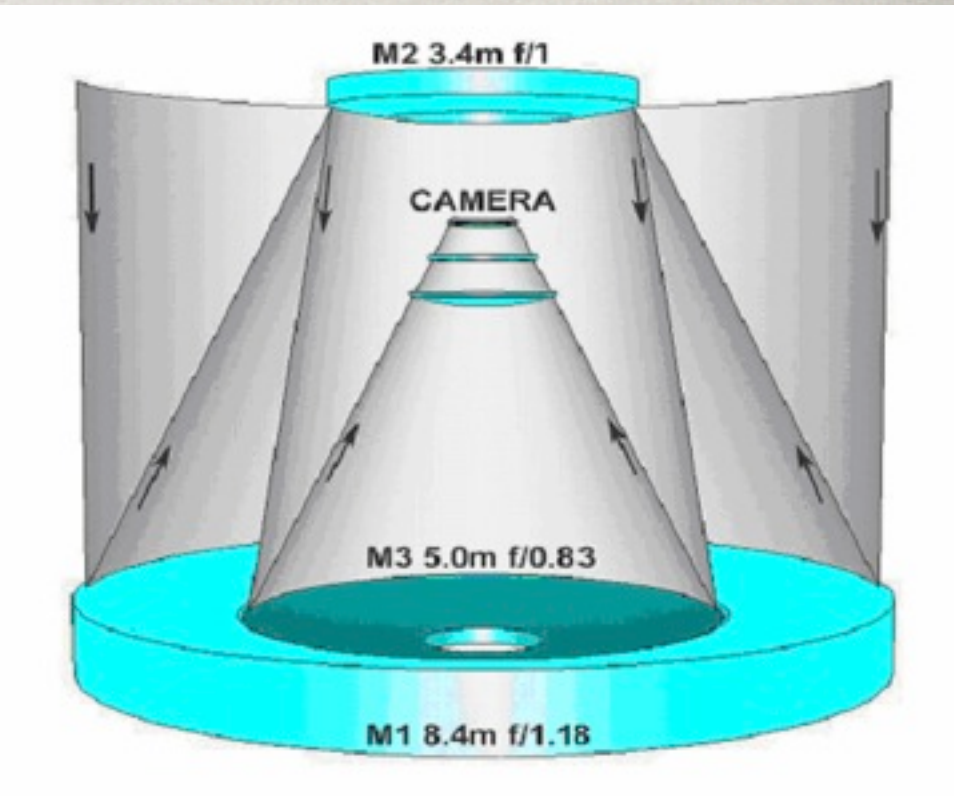


R. Ansari , CC-IN2P3, Jan 2014

The LSST observatory facility will be located on Cerro Pachón, Chile



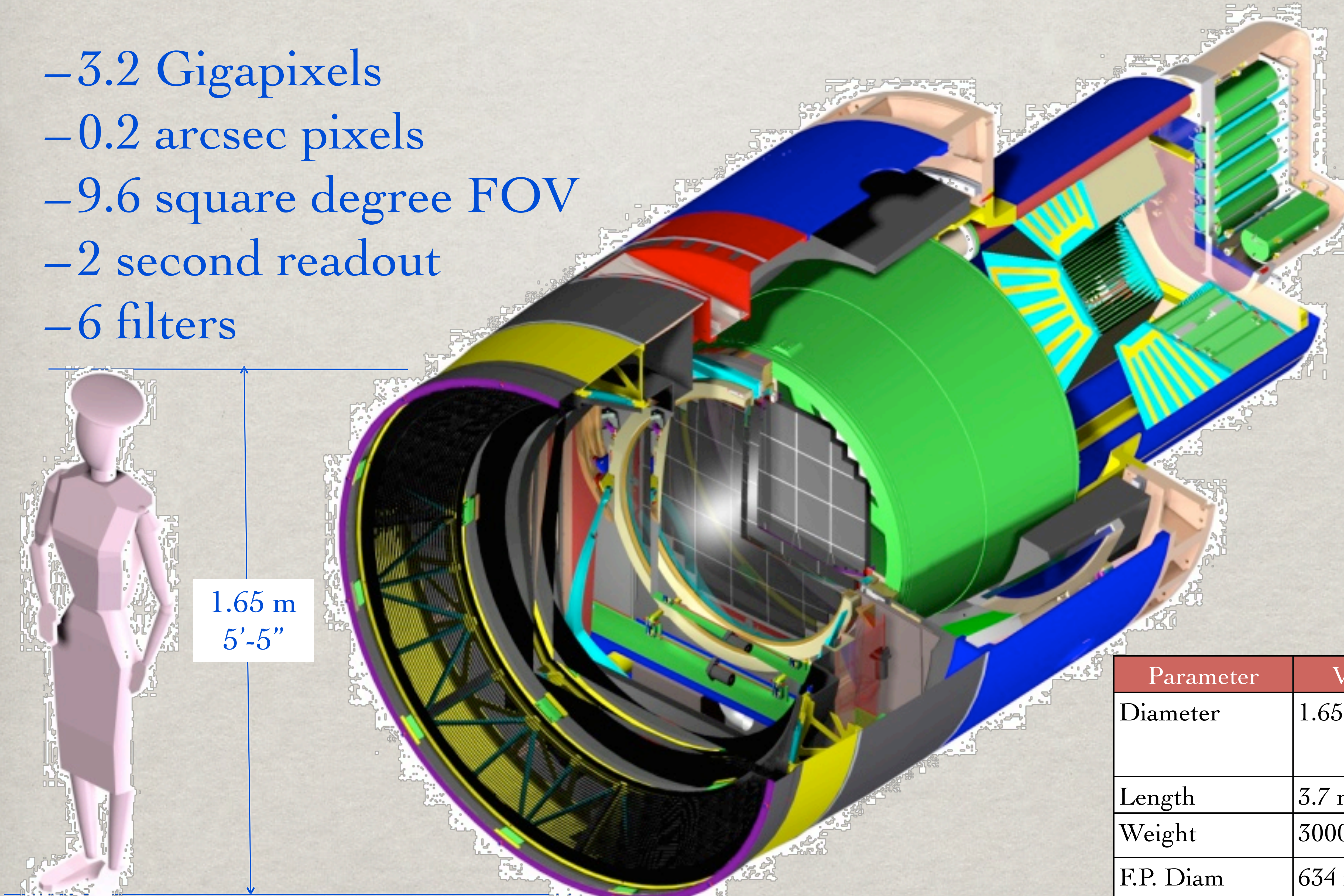
Blank of the first/tertiary mirror
(sept. 08)



Diameter: 8.4m
f/D = 1.23
field: 3.5°
Image: 0.2''

Camera Overview

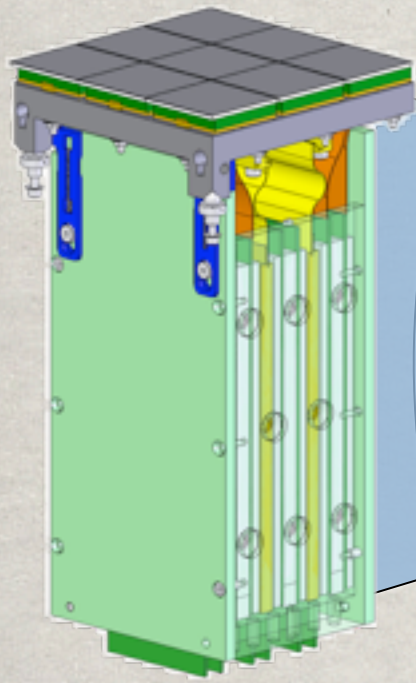
– 3.2 Gigapixels
– 0.2 arcsec pixels
– 9.6 square degree FOV
– 2 second readout
– 6 filters

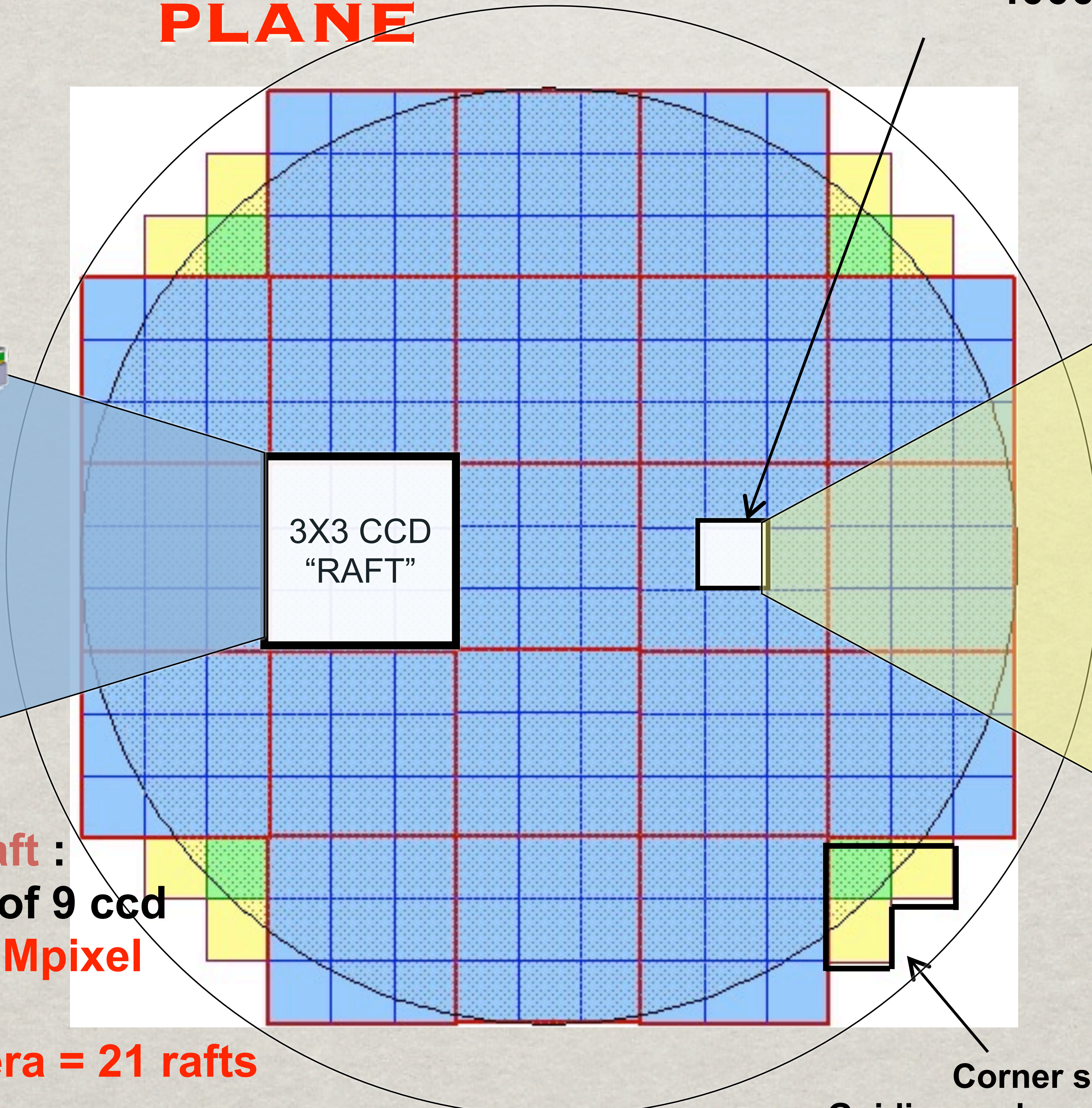


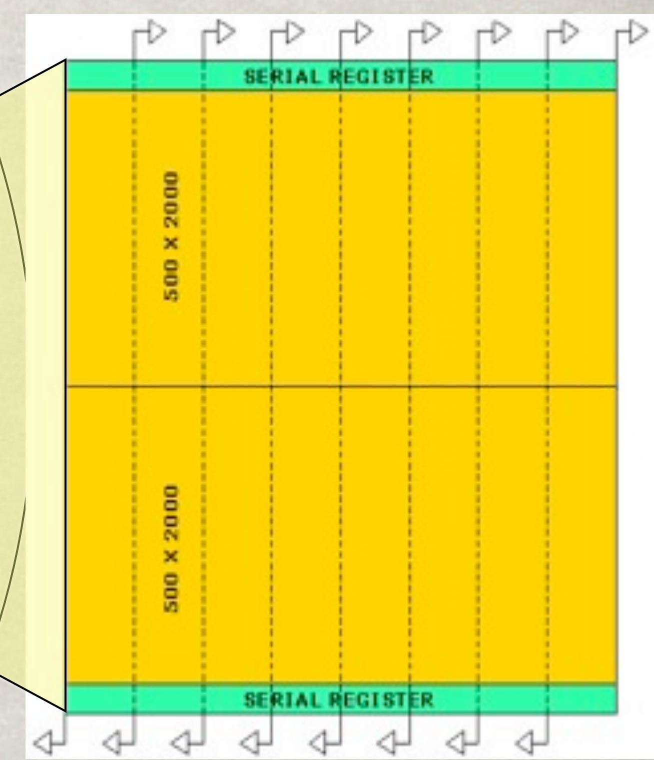
Parameter	Value
Diameter	1.65 m
Length	3.7 m
Weight	3000 kg
F.P. Diam	634 mm

LSST camera : focal plane

189 Science CCD
4000 X 4000 10μm pixels
= 3 x10^9pixels
- 9.6 Deg^2







Segmented CCD
16 segments of 1M pixels

Raft : Matrix of 9 ccd
= 144 Mpixel

LSST camera = 21 rafts

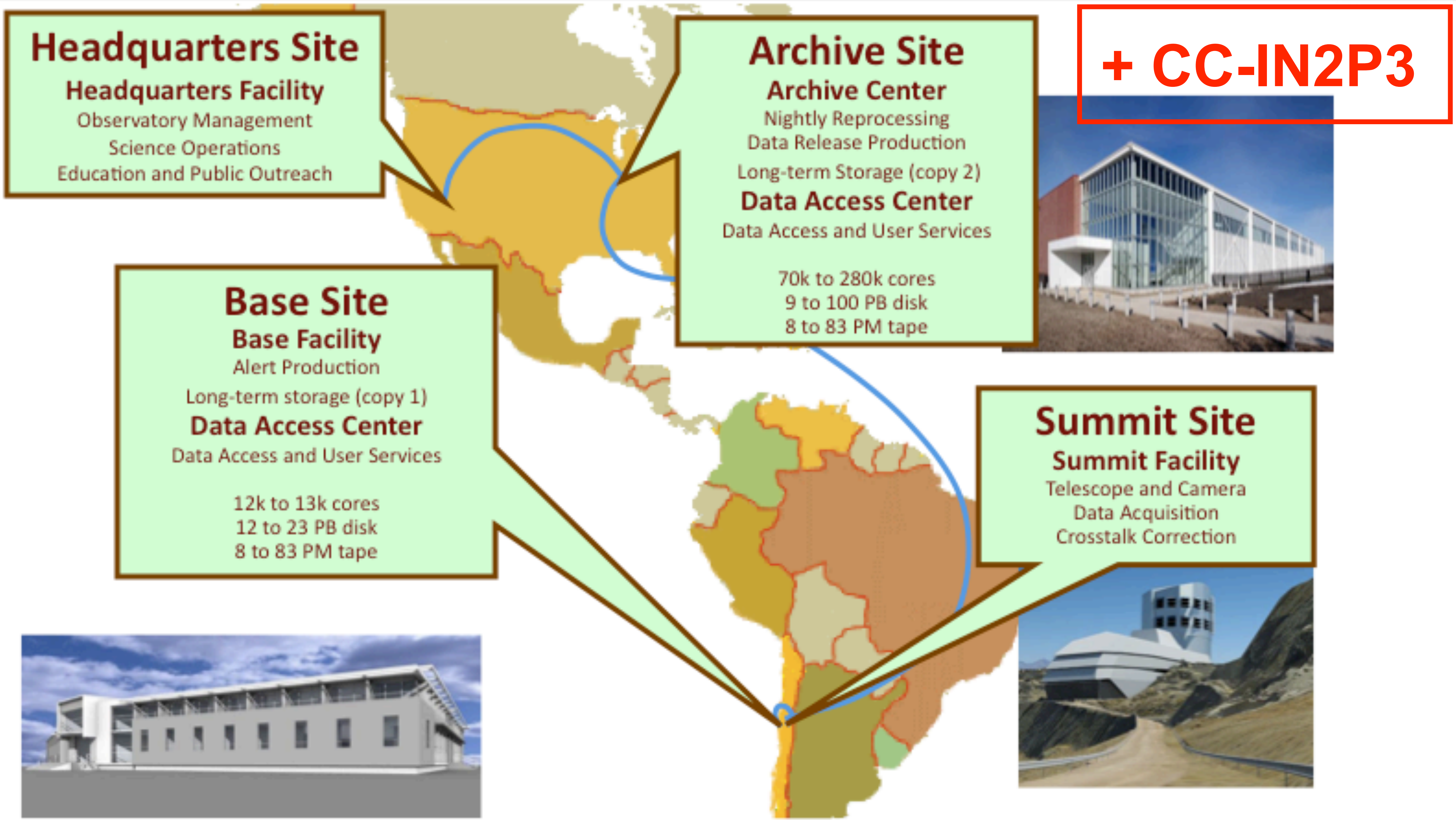
Corner sensors :
Guiding and wave-front analysis

R. Ansari , CC-IN2P3, Jan 2014

Computing challenges

- Data acquisition (volume, on-line processing …)
- Data processing and reduction : image and signal processing methods
- Large volume data storage and management
- Statistical inference
- Large scale modeling and simulation
- Making data usable and available to the scientific community

Cyber infrastructure is defined and capacity has been identified to handle data volume



- Summit-Base network will be installed by the project.
- Working with NSF funded network consortiums on capacity.
- International protected network identified and quoted.

LSST Data Managment

Simulation 1
CCD 4k x 4k

- Flot de données important :
  - Image : 6 Gbyte/17 secondes
  - 15 TB/nuit
  - ~100 PB archives finales images
  - 40. 10^9 objets (table de 100-200 TB)
  - 5 000. 10^9 observations (table de 1-3 PB)
  - Par nuit : alertes sur les transitoires (10^6)
  - Accès « ouvert » aux données : qui ? CC/NCSA ?
- Champ interdisciplinaire avec recherche informatique : BigData
- 2012 : mise en place consortium PetaSky
- 2013 : Test grande échelle au CC

Large Synoptic Survey Telescope

LSST : Sky coverage

~ 3 . 10^6 pointings x 2 x 189 CCD (10 years)
~ 100 PB of raw image data

Simulation Results

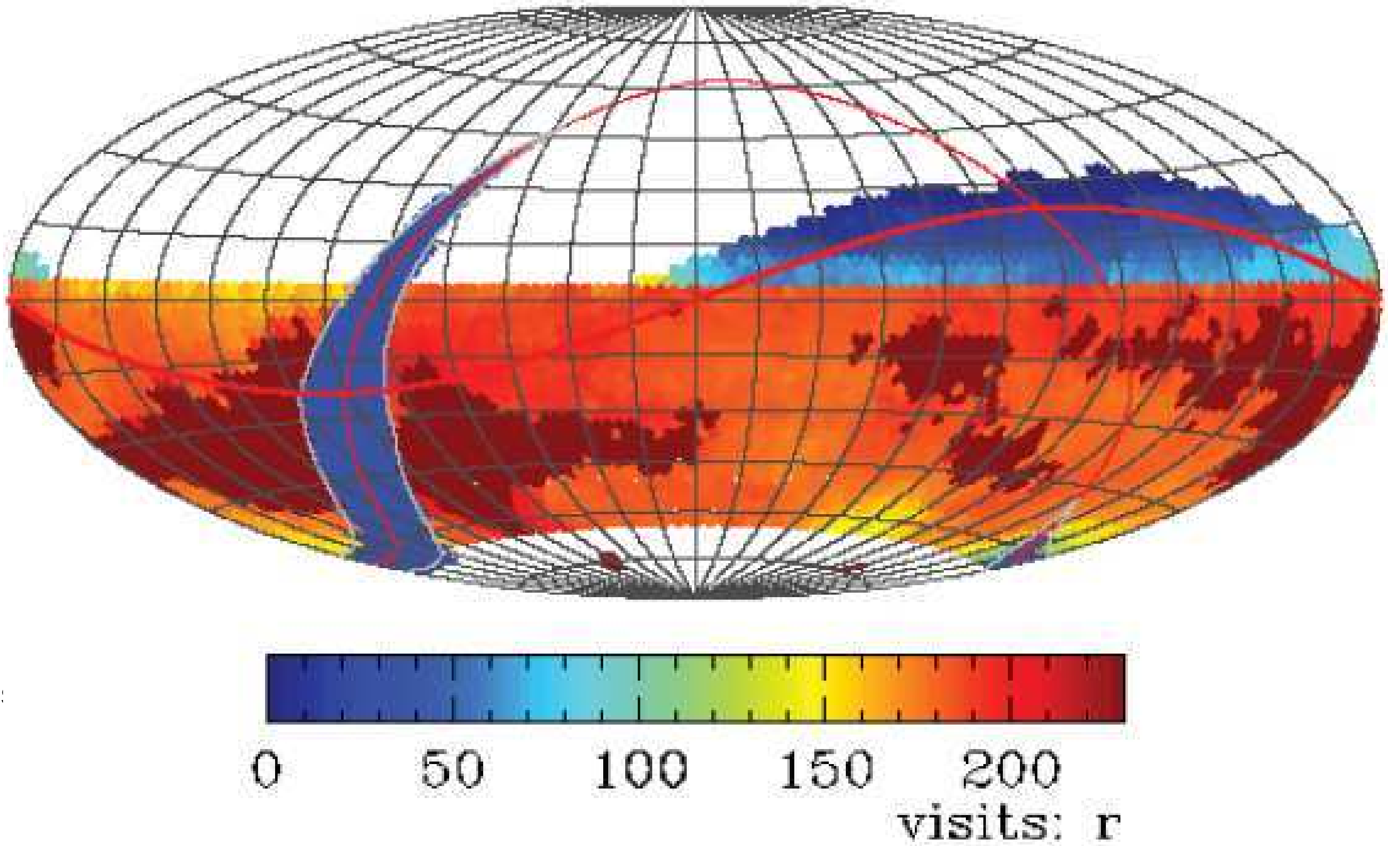
2,767,596 visits (5,535,192 15-second exposures) in 10-year survey

Deep-Wide-Fast: 20,000 square degrees at airmass < 1.4:
per-visit limiting magnitude (5σ, AB):
u: 23.9 g: 25.0 r: 24.7 i: 24.0 z: 23.3 y: 22.1
uniform stacked limiting magnitude (5σ, AB):
u: 26.0 g: 27.4 r: 27.5 i: 27.0 z: 26.2 y: 24.8
average airmass over all observations: 1.2
746,667 pairs of observations separated by 15-60 minutes in griz, an average of 6 per field per lunation
86% recovery of >140m NEA's
excellent period recovery for periods > 0.1 day (see figure 3)

Northern Ecliptic: 4,000 square degrees:
63,497 pairs of observations separated by 15-60 minutes in griz, an average of 2.3 pairs per field per lunation

Deep-Drilling: 150 sequences of 100 day duration with > 85% of observation completed. Dense grizy lightcurves to per-visit limiting magnitude (5σ, AB, for z=1 SN Ia):
g: 26.2 r: 26.3 i: 25.8 z: 25.1 y: 23.7

Galactic Plane: 30 observations in each of ugrizy in each field over 10 years
Uniform stacked limiting magnitude (5σ, AB) :
u: 25.6 g: 26.8 r: 26.5 i: 26.0 z: 25.3 y: 23.9



Large Synoptic Survey Telescope

LSST Survey Strategy: Cadence Design and Simulation

P.A. Pinto (Steward Observatory), K.H. Cook (LLNL), F. Delgado (CTIO), M. Miller (NOAO), L. Denneau (U. Hawaii), A. Saha (NOAO), P.A. Gee (UC Davis), J.A. Tyson (UC Davis), Z. Ivezic (U. Washington) for the LSST Collaboration

R. Ansari , Jan 2014

Paradigme LSST :

Processing Cadence	Image Category (files)	Catalog Category (database)	Alert Category (database)
Nightly	Raw science image Calibrated science image Subtracted science image Noise image Sky image Data quality analysis	Source catalog (from difference images) Object catalog (from difference images) Orbit catalog Data quality analysis	Transient alert Moving object alert Data quality analysis
Data Release (Annual)	Stacked science image Template image Calibration image RGB JPEG Images Data quality analysis	Source catalog (from calibrated science images) Object catalog (optimally measured properties) Data quality analysis	Alert statistics & summaries Data quality analysis

Contribution française au calcul

- L2 : caractérisation → catalogues + stacked images
  - Delivré par projet
- L3 : analyse
  - Hors projet : ex. DESC
  - Quelle relation aux données L2 / reprocessing spécifique ?

LSST DATA

Images: 3. 10^6 x 2 x 189 CCD : ~ 100 PB

Catalogues

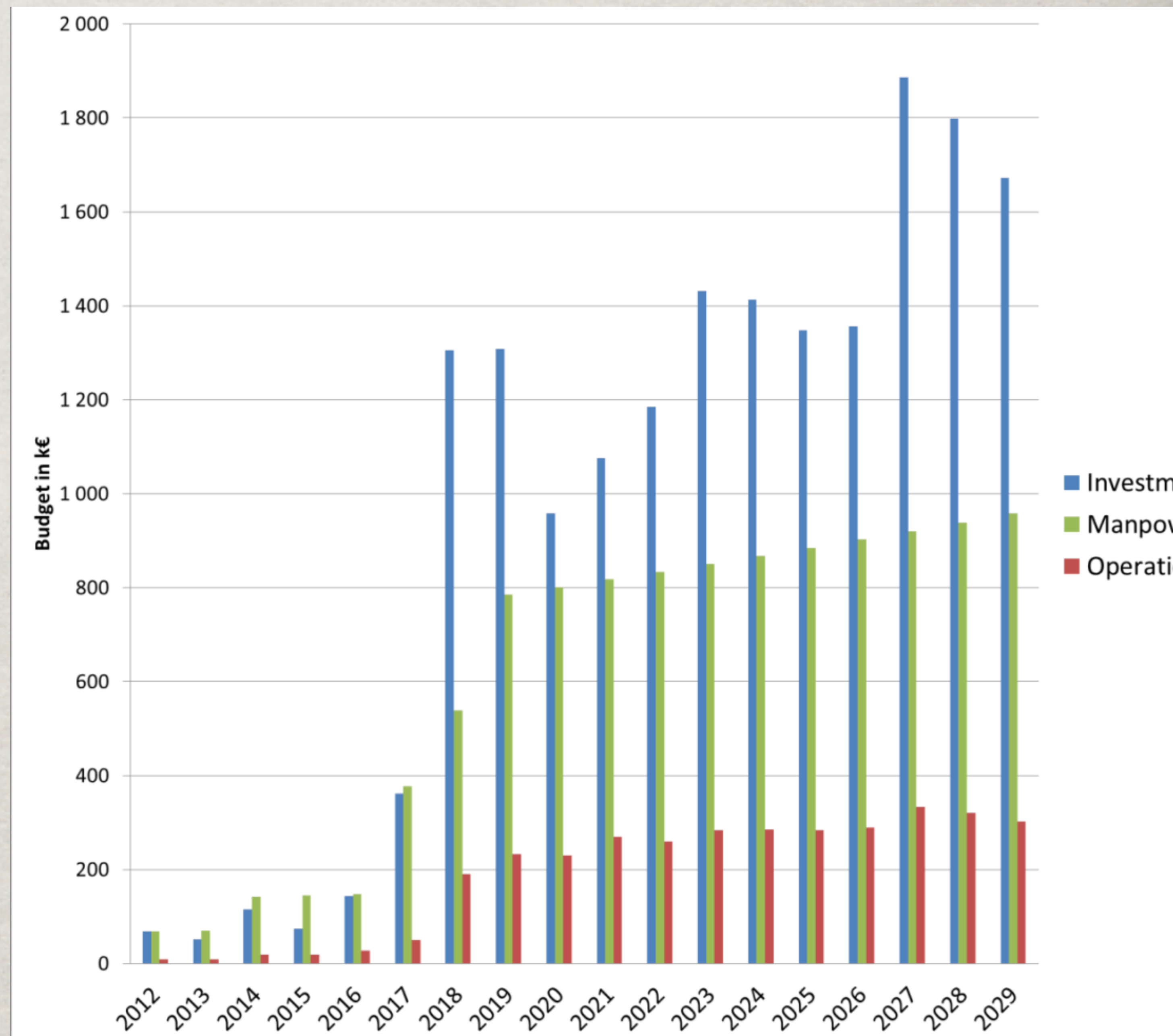
Table	Size [TB]	rows	columns	description
Object	109	~38 billion	~500	Most heavily used, for all common queries on stars/galaxies, including spatial correlations and time series analysis using summarized information
CalibSource	24	~100 billion	~25	Sources used for calibration
DiaSource	71	~200 billion	~50	Alert-related follow up analysis
Source	3,600	~5 trillion	~100	Time series analysis of bright objects and detections
ForcedSource	1,089	~23 trillion	~7	Specialized analysis of faint objects and detections

Data Base table volume : 5 PB !

R. Ansari , CC-IN2P3, Jan 2014

Storage / Computing ressource forecast for LSST at CC-IN2P3

- Disk: 10 PB (1st year) … 25 PB (10th year)
- ~ 10 PB / an (storage space) : 10 (Y1) … 100 PB (Y10)
- Processing power 100 TF (Y1) … 900 TF (Y10)



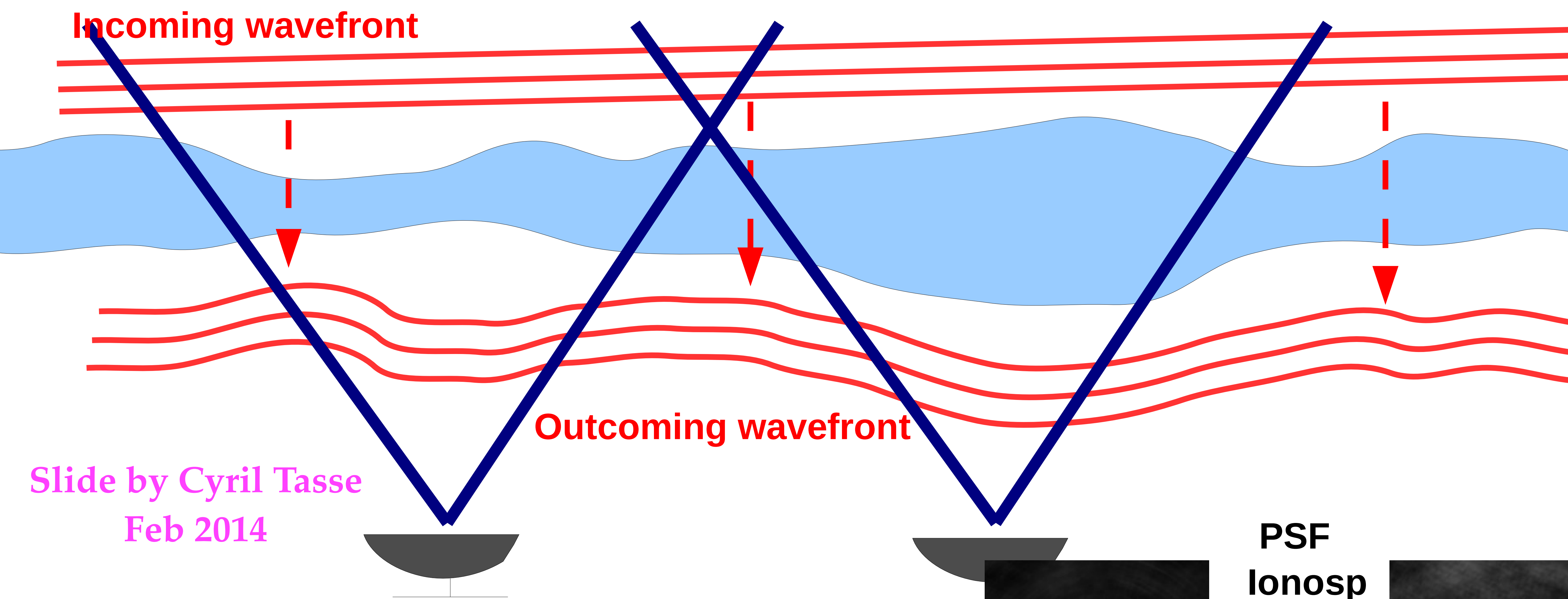
Rapport au CS IN2P3 en décembre 2012

R. Ansari , CC-IN2P3, Jan 2014

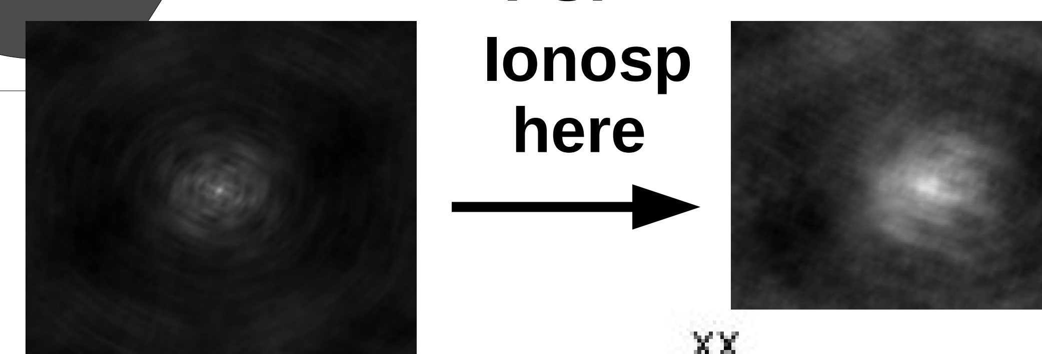
radio interferometry - SKA computing challenges

- Reconstruction of sky maps from interferometric observations (Cyril Tasse slides from LOFAR-SKA radio days, Feb 2014)
- Digital radio interferometry : huge data flow, needs enormous computing and network bandwidth (correlation computation)
- RFI cleaning, gain/phase calibration , beam determination …

… When Direction Dependent Effects (DDE) become a problem : Ionosphere/troposphere



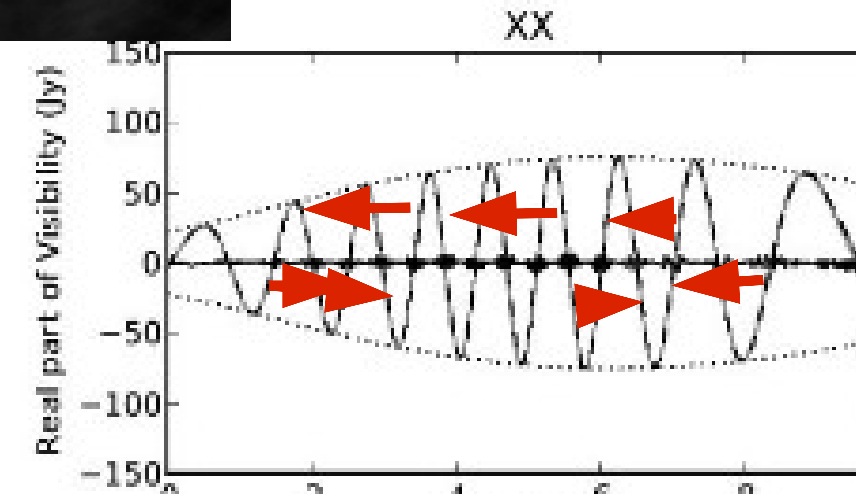
Slide by Cyril Tasse
Feb 2014



Big field of view : station, direction, time and frequency dependent

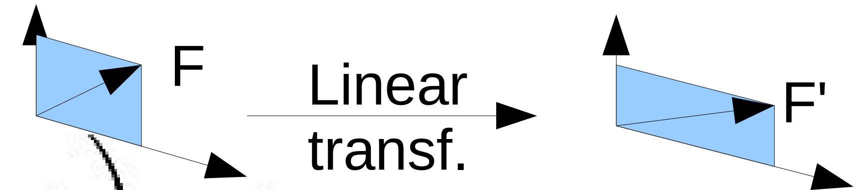
Other direction dependent effects :

- Faraday rotation
- **+ Effect on the polarisation**



The Measurement Equation

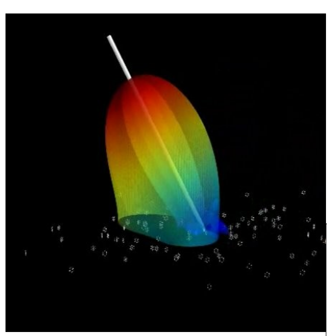
Hamaker 1996



$$V_{pq} = \overbrace{G_p}^{\text{Direction independent}} \left(\sum_{i=1}^{N} \overbrace{\mathrm{B}_{pi}\mathrm{K}_{pi}\mathrm{I}_{pi}}^{\text{Direction dependent}} \overbrace{F_i \, . \, F_i^+}^{\text{Source coherency}} \mathrm{I}_{qi}^+ \mathrm{K}_{qi}^+ \mathrm{B}_{qi}^+ \right) G_q^+$$

[Voltage antenna p] x [Voltage antenna q]*

Beam

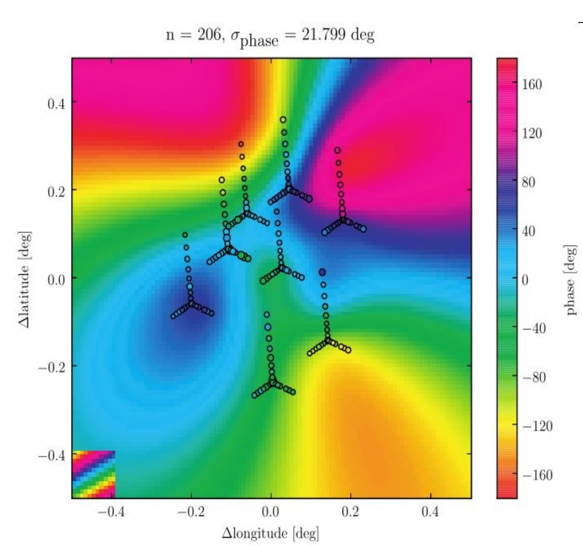


Geometrical delay +Correlator

$$\mathrm{K}_p\mathrm{K}_q^+ = \exp(-2i\pi\phi_{pq}) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

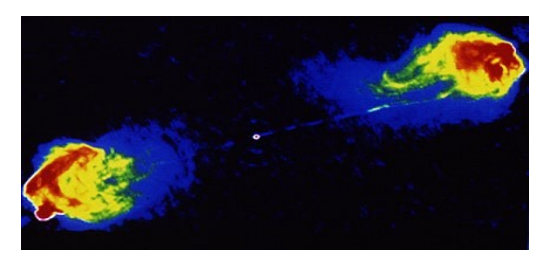
$$\phi_{pq} = u_{pq}l + v_{pq}m + w_{pq}(\sqrt{1-l^2-m^2}-1)$$

Ionosphere



Van der Tol thesis

Electric field

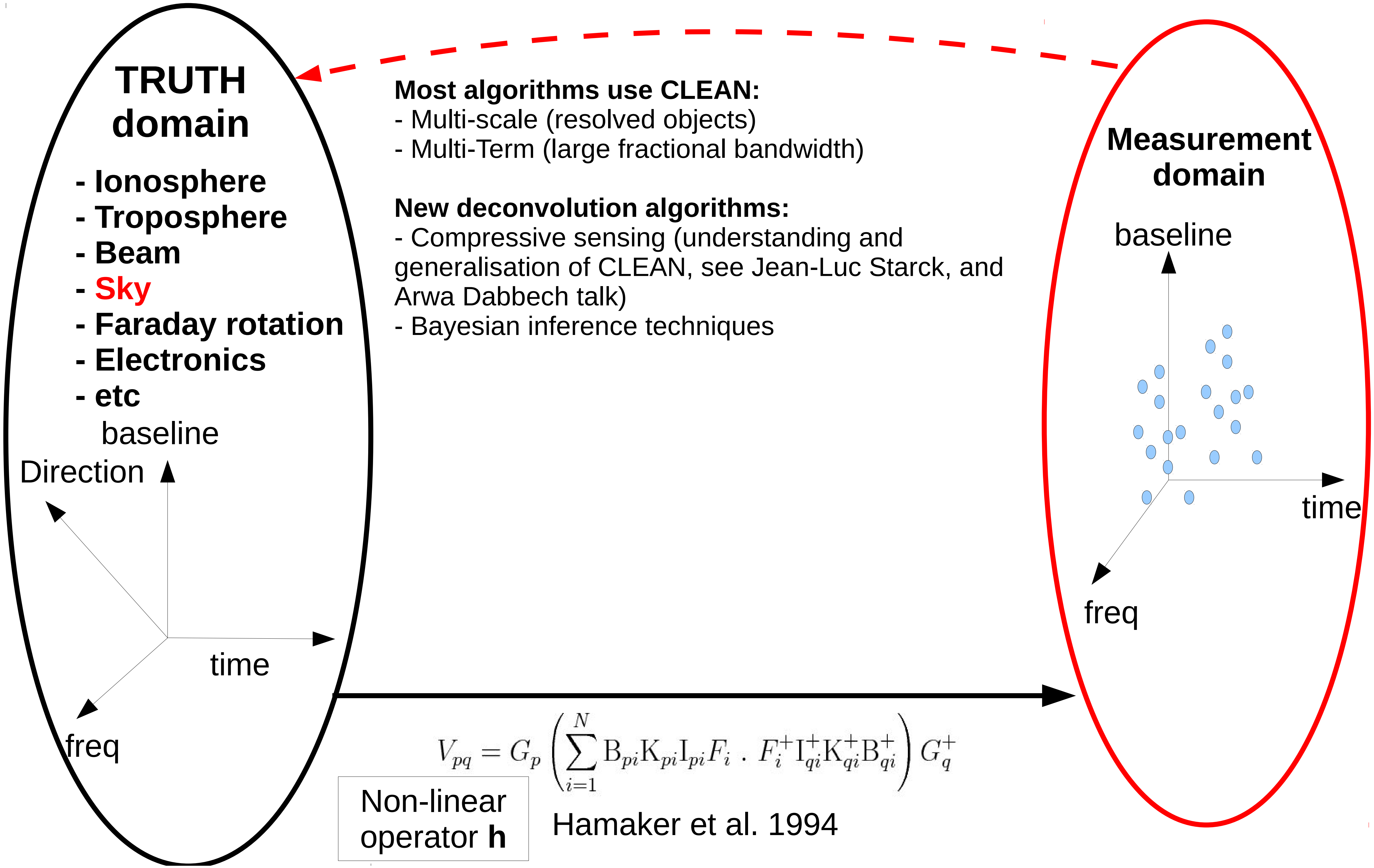


Slide by Cyril Tasse Feb 2014

Imaging and deconvolution

Most algorithms use CLEAN:
- Multi-scale (resolved objects)
- Multi-Term (large fractional bandwidth)

New deconvolution algorithms:
- Compressive sensing (understanding and generalisation of CLEAN, see Jean-Luc Starck, and Arwa Dabbech talk)
- Bayesian inference techniques



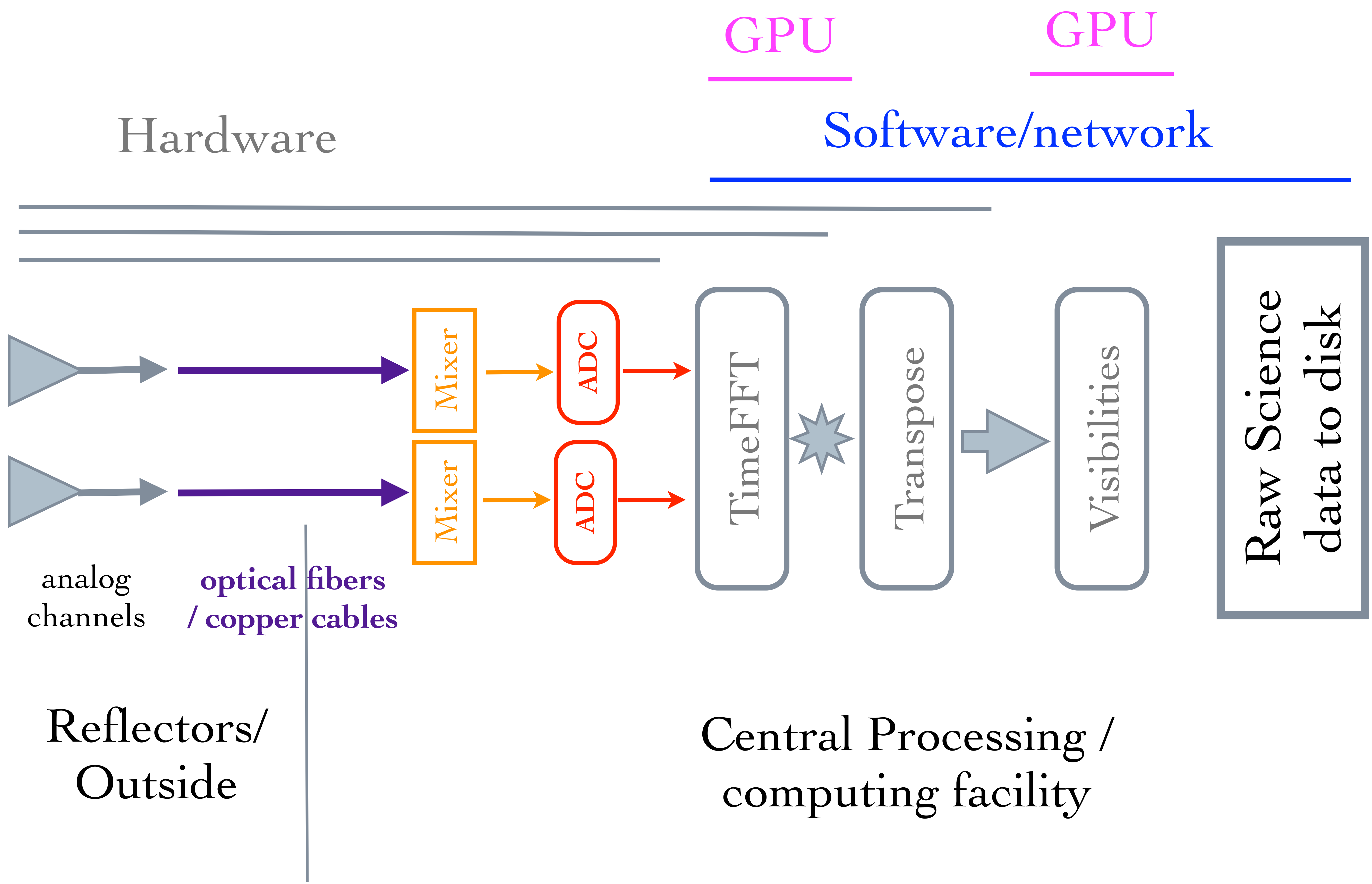
$$V_{pq} = G_p \left(\sum_{i=1}^{N} \mathrm{B}_{pi} \mathrm{K}_{pi} \mathrm{I}_{pi} F_i \ . \ F_i^{+} \mathrm{I}_{qi}^{+} \mathrm{K}_{qi}^{+} \mathrm{B}_{qi}^{+} \right) G_q^{+}$$

Non-linear operator **h**

Hamaker et al. 1994

Slide by Cyril Tasse
Feb 2014

Tianlai Electronic chain overview (1)



R. Ansari - May 2013

Visibilities:

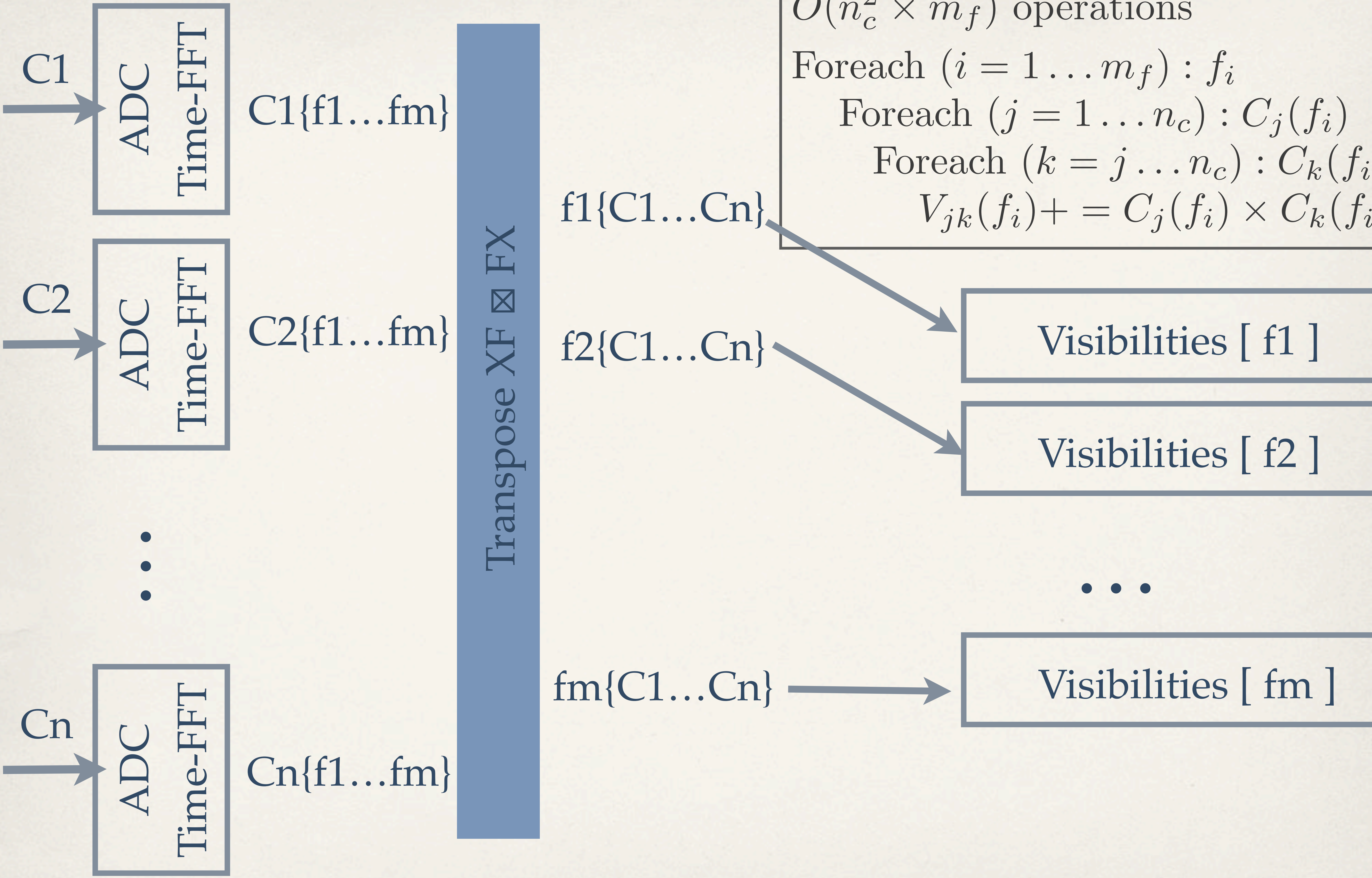
$O(n_c^2 \times m_f)$ operations

Foreach $(i = 1 \ldots m_f) : f_i$

  Foreach $(j = 1 \ldots n_c) : C_j(f_i)$

    Foreach $(k = j \ldots n_c) : C_k(f_i)$

      $V_{jk}(f_i) += C_j(f_i) \times C_k(f_i)$



R. Ansari - May 2013

CPU & bandwidth requirements

	A	B	C
NFeed	32	256	1024
BandWidth	100 MHz	200 MHz	400 MHz
1 ⇢ 2 ⇢ 3	6.4 GBytes / sec	100 GBytes / sec	800 GBytes / sec
M	8	64	256
3 / M ➞	0.8 GBytes / sec	1.6 GBytes / sec	3.2 GBytes / sec
NVis	528	32896	526336
@4 TFlops	~ 1	~ 100	~ 3200
4 ➞ / M	5 MBytes / sec	50 MBytes / sec	400 MBytes / sec

Easy … *Challenging* … *??*

R. Ansari - May 2013