Gaia, the universe in 3D: an overview of the mission

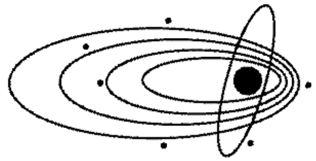
X. Luri, ICCUB/IEEC

Outline of the talk

- Gaia's context in ESA

- The Gaia mission
  - Scientific goals
  - Spacecraft and instruments
  - Data reduction: the DPAC

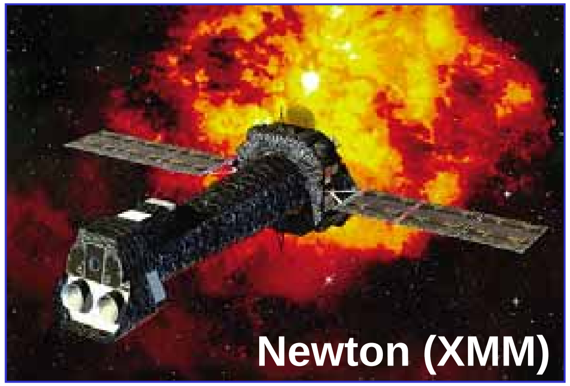
- The Gaia simulator
  - Simulations in DPAC
  - Simulator structure
  - The Universe Model and the Instrument Model
  - Data generators
  - Lessons learnt
  - Examples

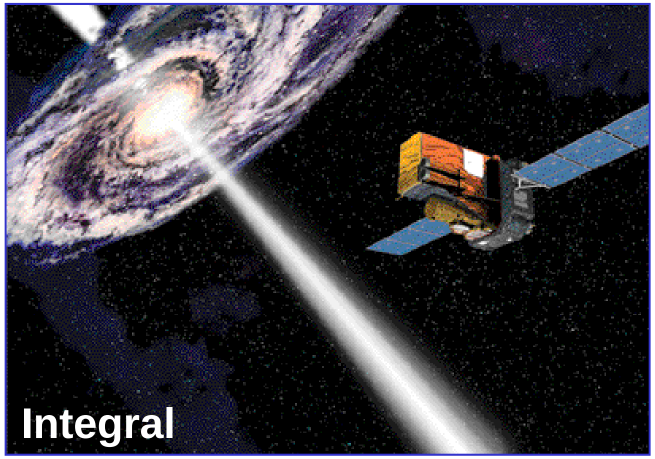


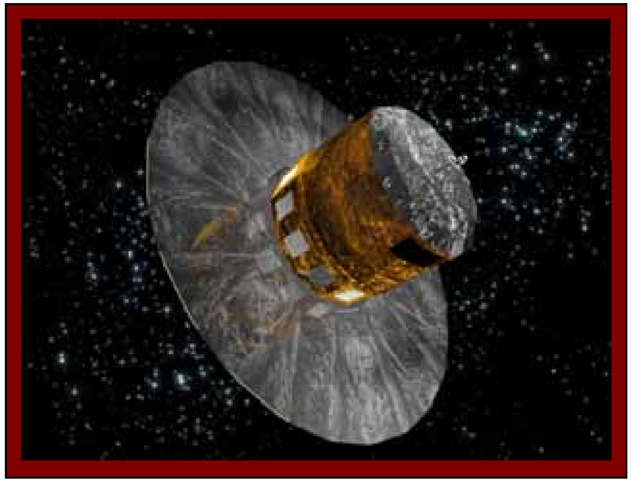
- All ESA’s member states participate in its activities and in a series of common space science programs (mandatory programs)

- ESA’s areas of work:
  - Earth’s space environment
  - Sun-Earth interaction
  - Interplanetary medium
  - The Moon and the planets
  - The stars and the universe

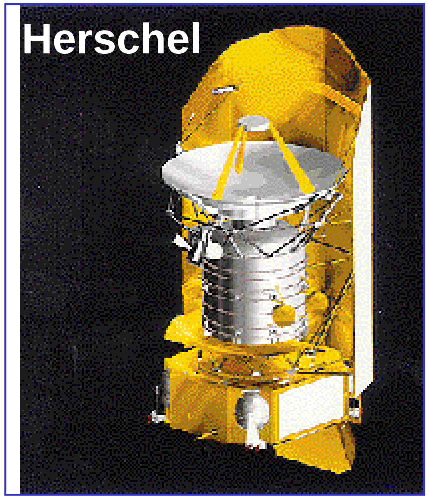
Cosmic vision 2002-2012

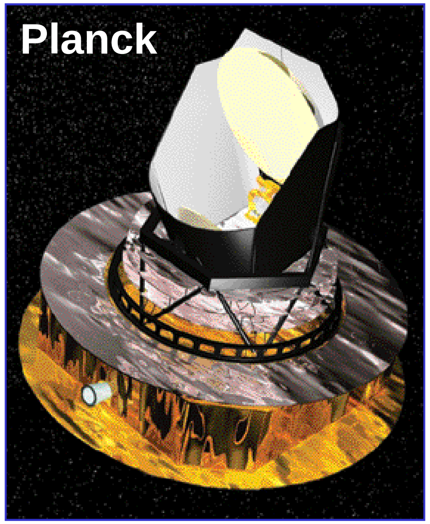






Gaia





The Gaia mission

Gaia history

- Gaia is the successor of the Hipparcos satellite, the first space astrometry mission. The Hipparcos catalogue is today an essential reference in astronomy and has led to more than 1600 refereed publications since 1996

  http://www.rssd.esa.int/index.php?project=HIPPARCOS&page=science_results

- Gaia is the *Cornerstone 6* in the frame of ESA’s “Horizon 2000+” program. It was approved in 2001 and its launch is scheduled for 2013.

- **1998-2000:** proposal to ESA
- **2001:** ESA selection
- **2001-03:** preparatory studies
  - Mission feasibility
  - Definition of the scientific goals
- **2005-2006:** Phase B (detailed design)
- **2006-2012:** Phase C/D (construction)
- **Mid-2013: launch**
- **2013-2018:** nominal operation (+1yr extended)
- **2013-2021:** data reduction

Main scientific goals

Structure and kinematics of our galaxy:

- Form and rotation of the bulge, disc and halo
- Internal motion of the regions of stellar formation, clusters, etc.
- Nature of the spiral arms and the "warp"
- Spatial motion of all the satellite systems satellite of the Galaxy

Stellar populations:

- Physical characteristics of all the galactic components
- Initial mass function, binary masses, chemical evolution
- History of the stellar formation

Tests of the galactic formation:

- Dynamic determination of the dark mass distribution
- Reconstruction of "mergers" and history of accretion

$\Rightarrow$ **Origin, Formation and evolution of the galaxy**

Additional goals

- **Stellar astrophysics**
- **Multiple stellar systems**
- **Solar system objects**
- **Extrasolar planets**
- **Fundamental physics (test of general relativity)**
- **Galaxies & QSOs**

Gaia in a nutshell

- ESA-only mission (prime EADS Astrium SAS)
- Launch date: June 2013 (schedule without margin)
- Launcher: Soyuz–Fregat from CSG (Guiana)
- Orbit: L2 Lissajous orbit (1.5 million km from Earth)
- Lifetime: 5 years (1 year potential extension)
- Ground stations: Cebreros + New Norcia + Malargüe
- Downlink rate: 4 – 8 Mbps (~50 GB day^{-1})

- Science data processing by DPAC (400+ scientists)
- Science exploitation by GREAT (ESF + EU-FP7)
- Final catalogue ~2021 (archive preparation by GAP)
- Intermediate catalogues under definition (~2016 + ...)
- Science alerts released immediately
- No proprietary data rights

Figure courtesy EADS Astrium and Starsem

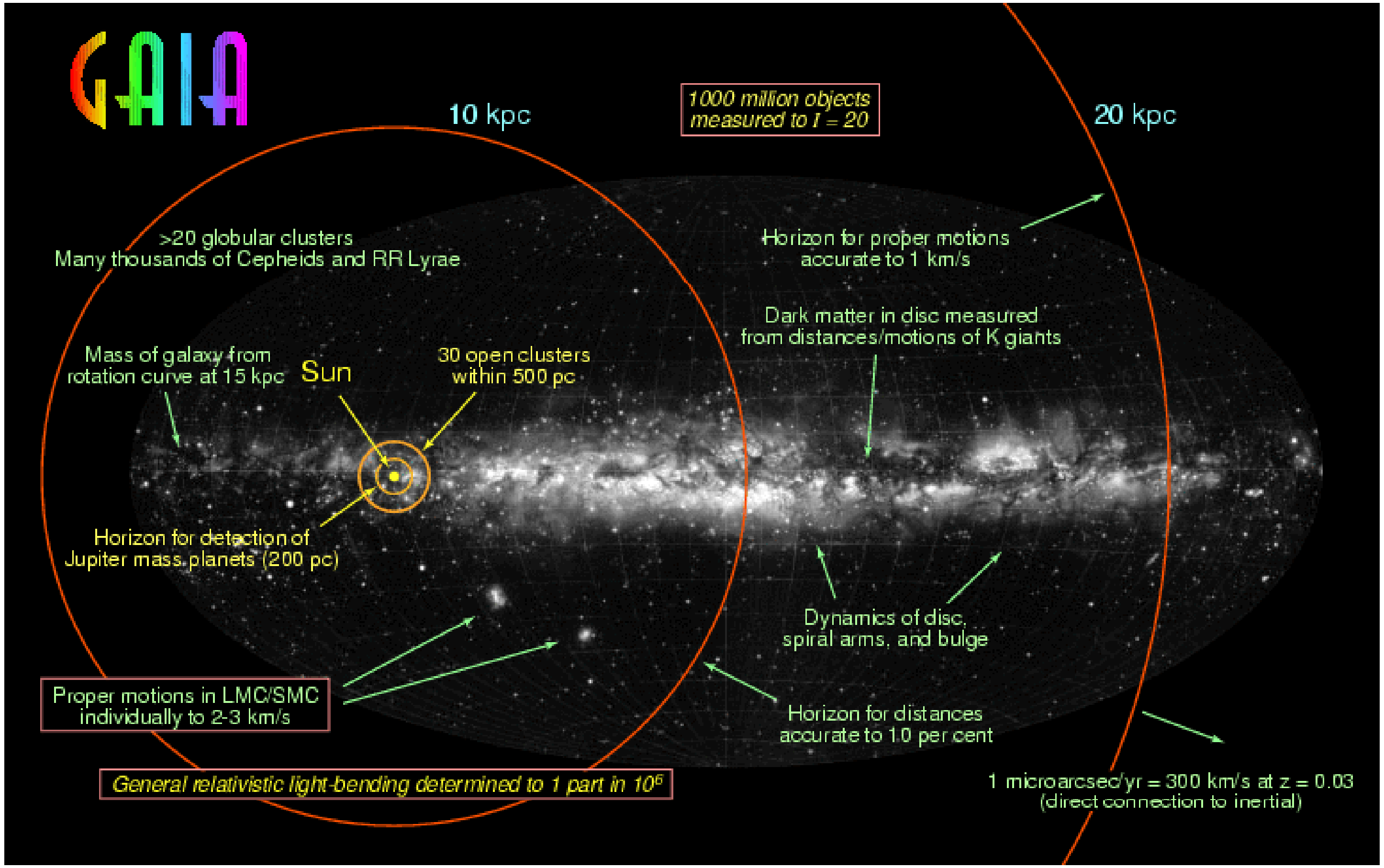
Gaia: an astrometric mission

Will provide the most complete 3D survey of objects in our Galaxy (and beyond)

- **>10^9 objects**
- **Complete up to 20th magnitude**
- Positions, velocities and parallaxes
  - Nominal precision (15^{th} mag): ~25μas
- Spectrophotometry
- Spectroscopy and radial velocities (G<16)
- No input catalogue → unbiased survey

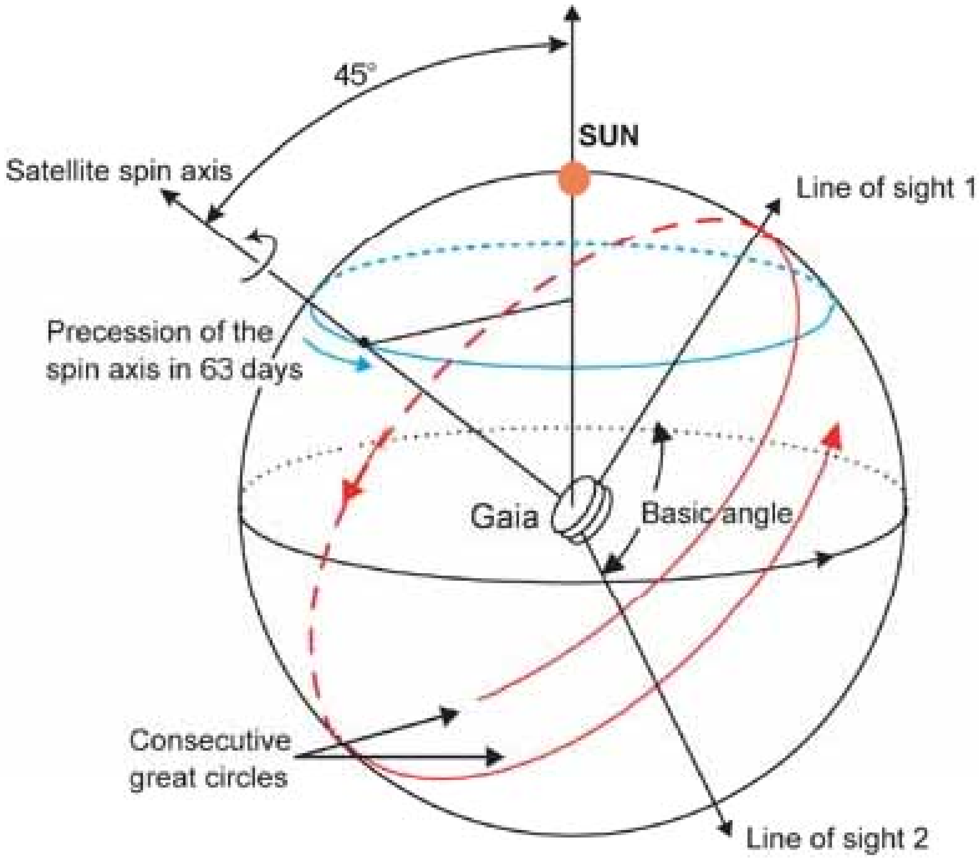
>10^9 objects

About 1% of the Milky Way! → statistically significant galactic census

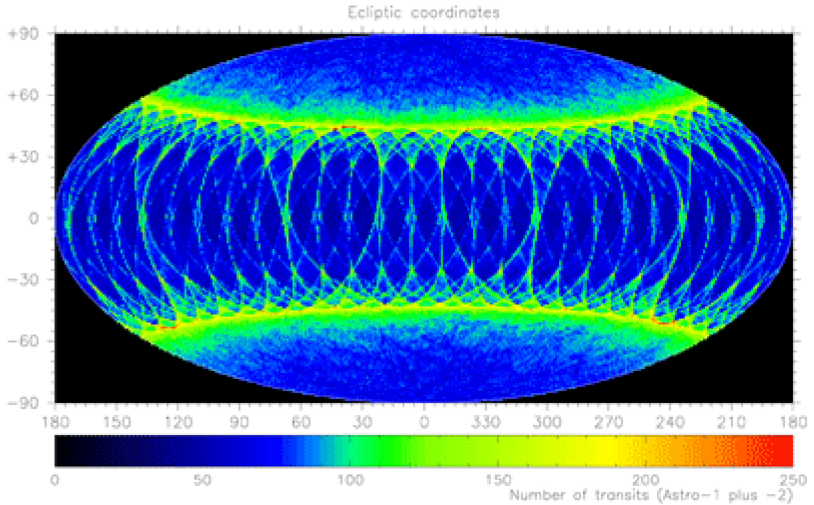


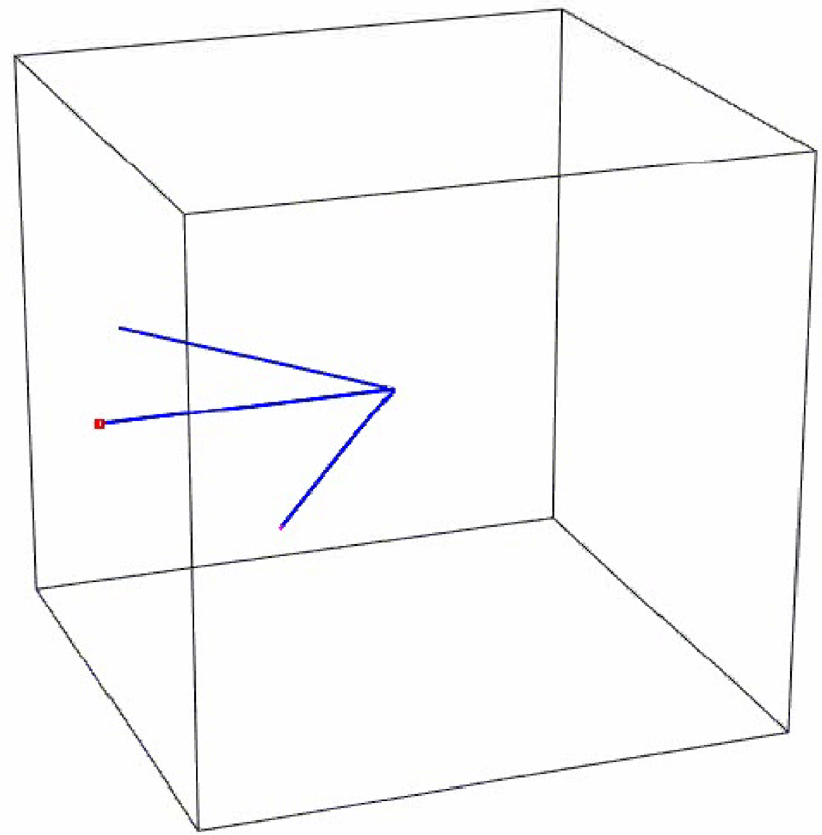
Gaia principles

Two telescopes scanning the sky



Scanning law; provides full sky coverage





Multiple pairs of instantaneous observations provide an all-sky grid of angular measurements. A global reduction process converts them into global astrometry: positions, parallaxes and proper motions.

Spectrophotometric measurements provide physical parameters for all the observed objects.

Spectroscopic measurements provide radial velocities.

Payload Module (PLM)

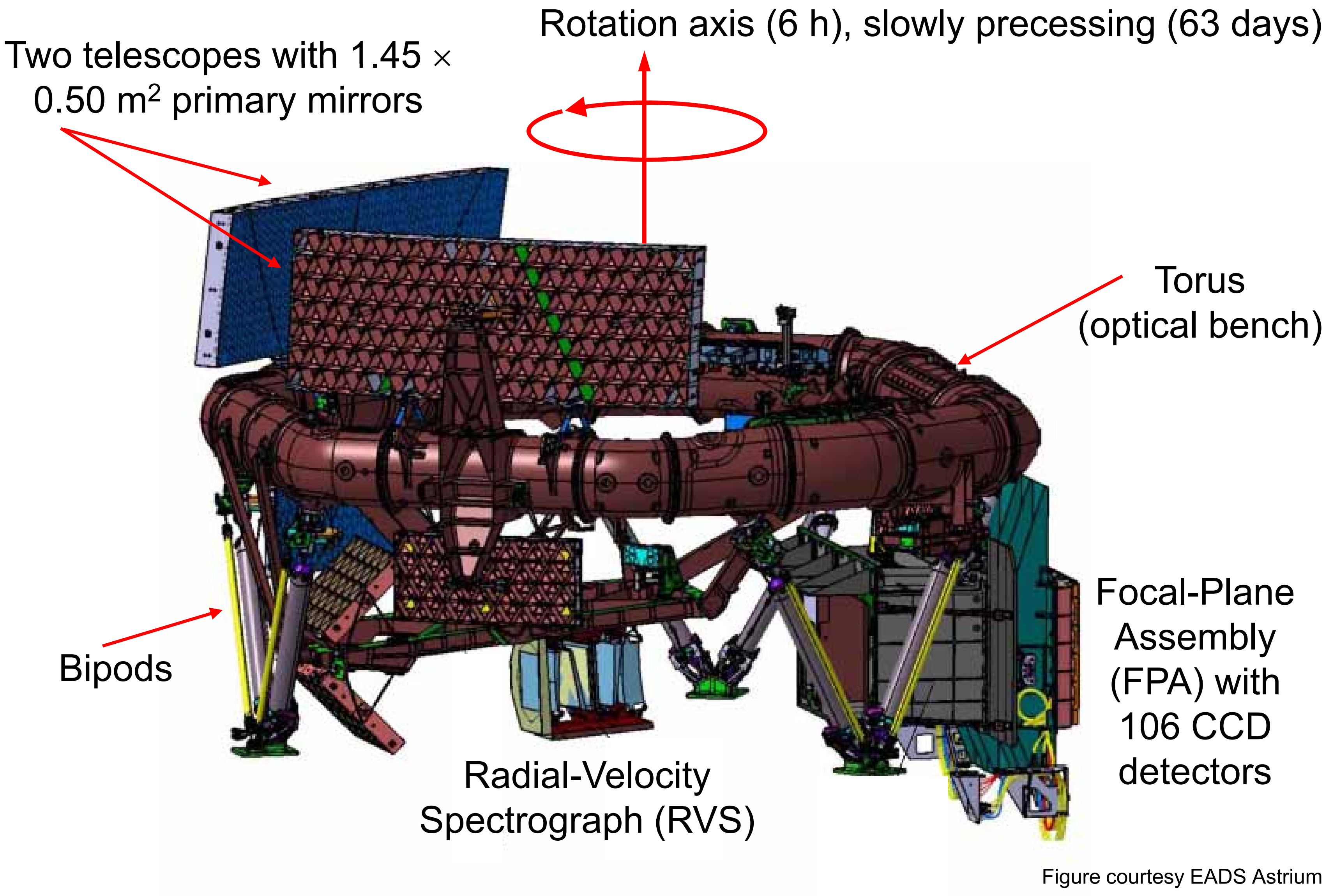
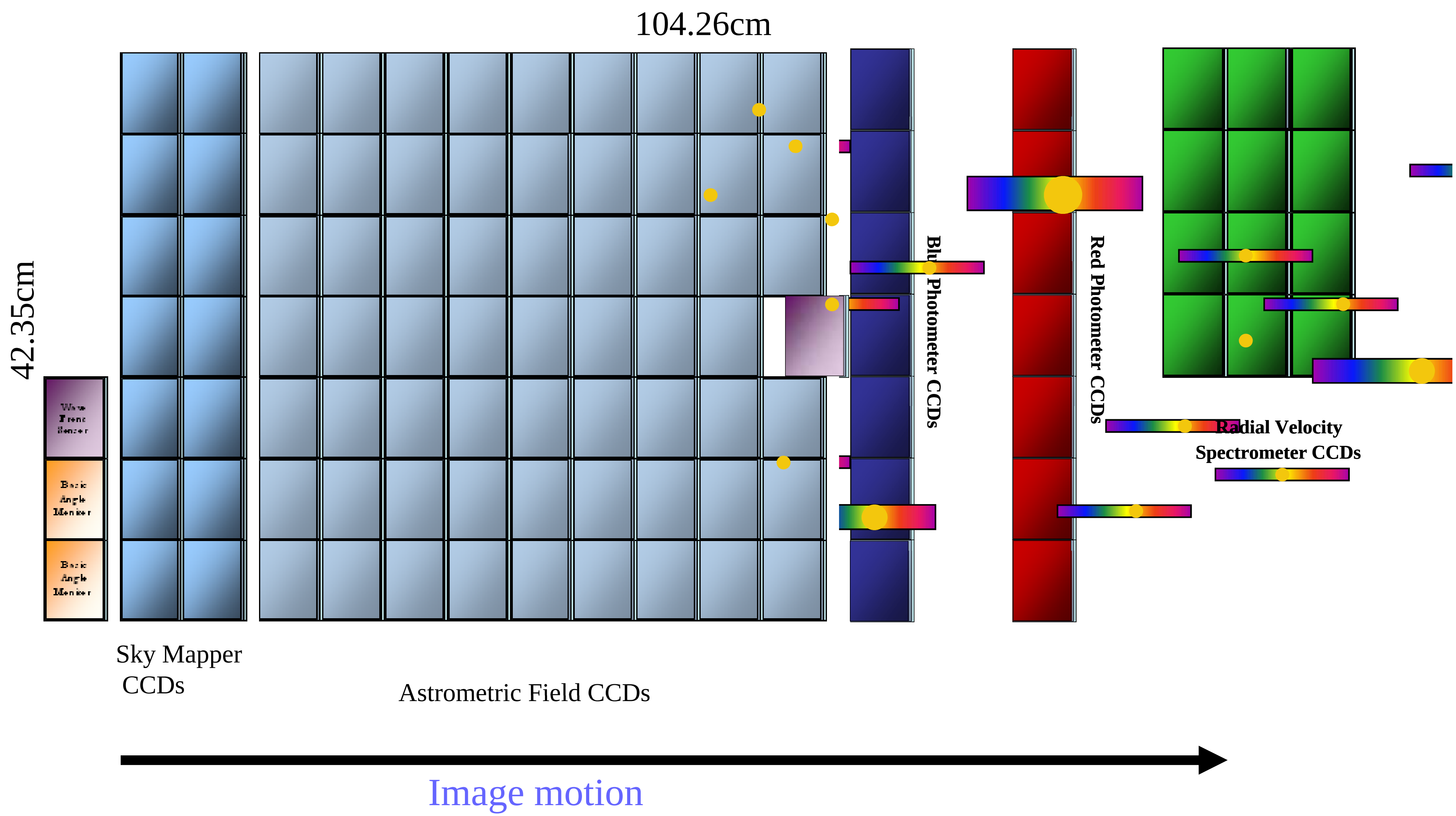


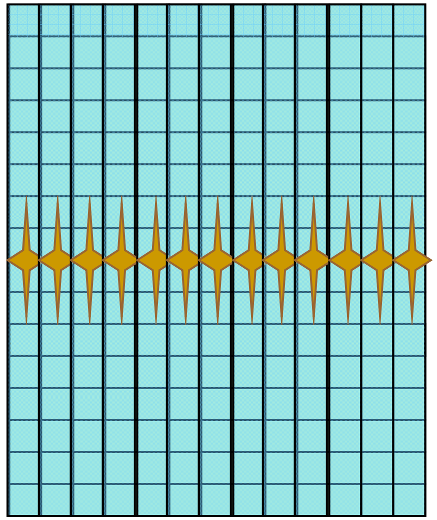
Figure courtesy EADS Astrium

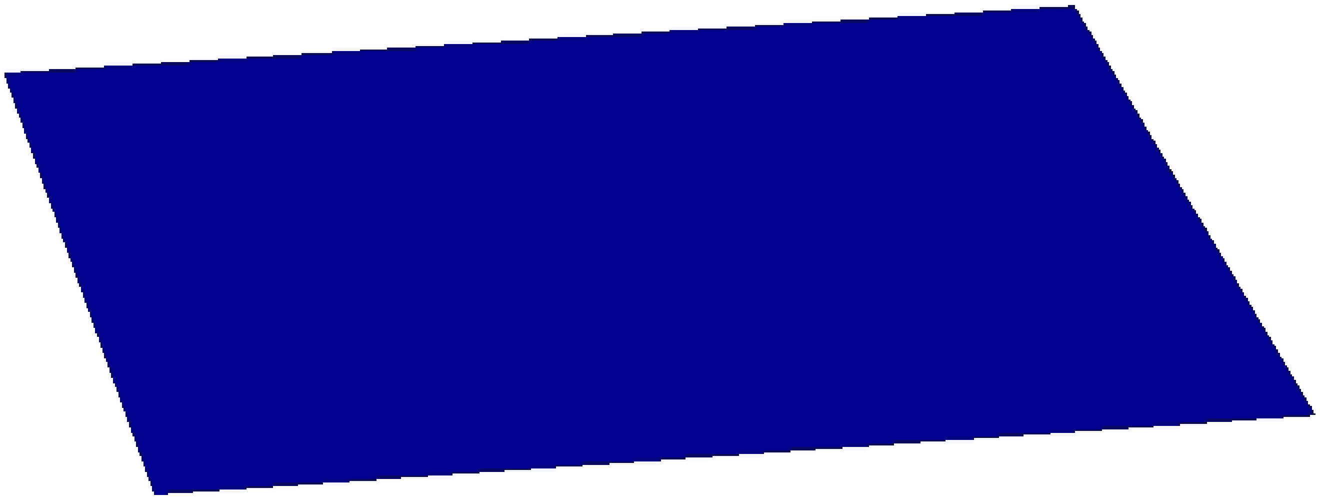
Focal plane

106 CCDs , 938 million pixels, 2800 cm^2

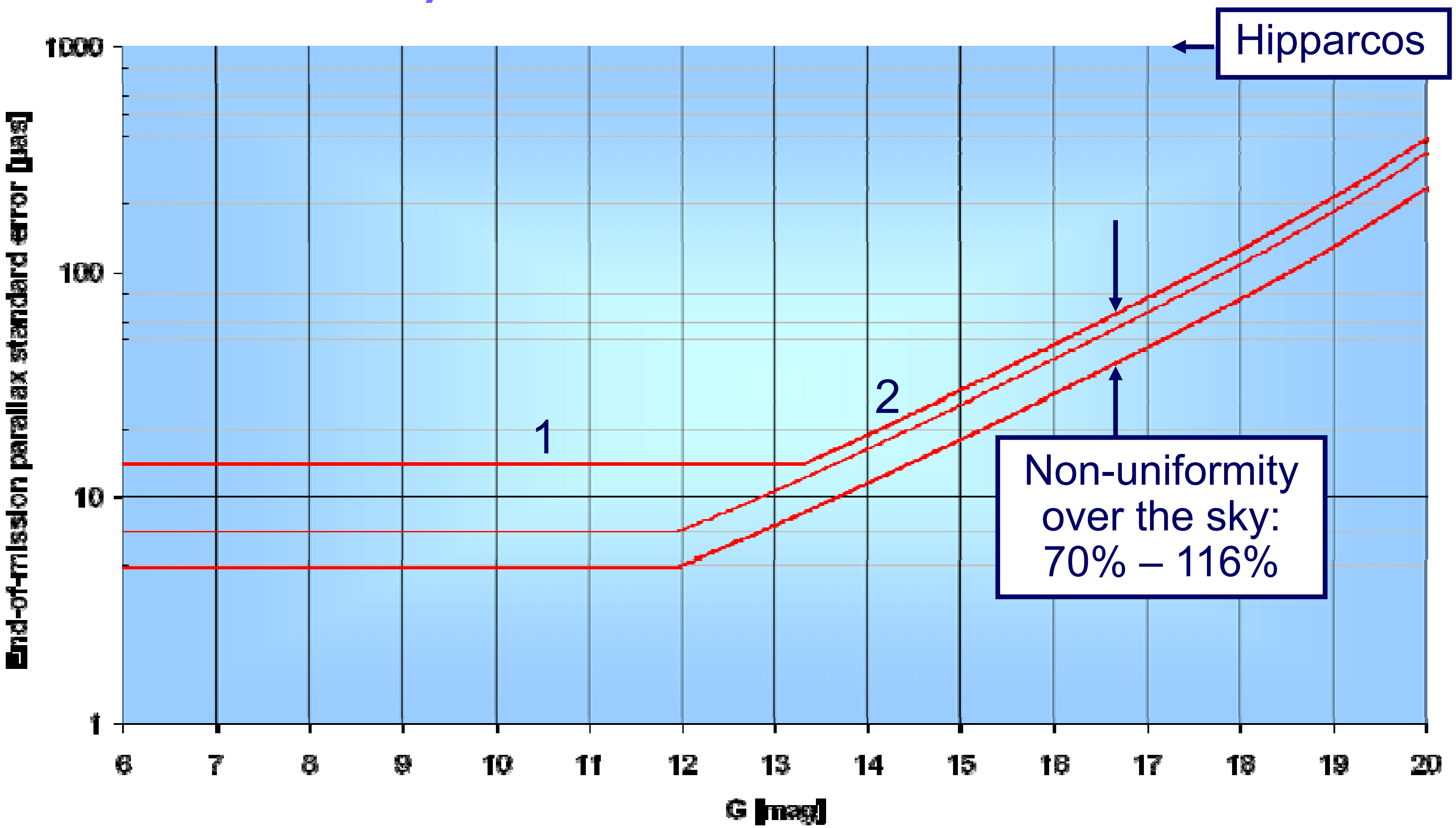


Time-delayed integration





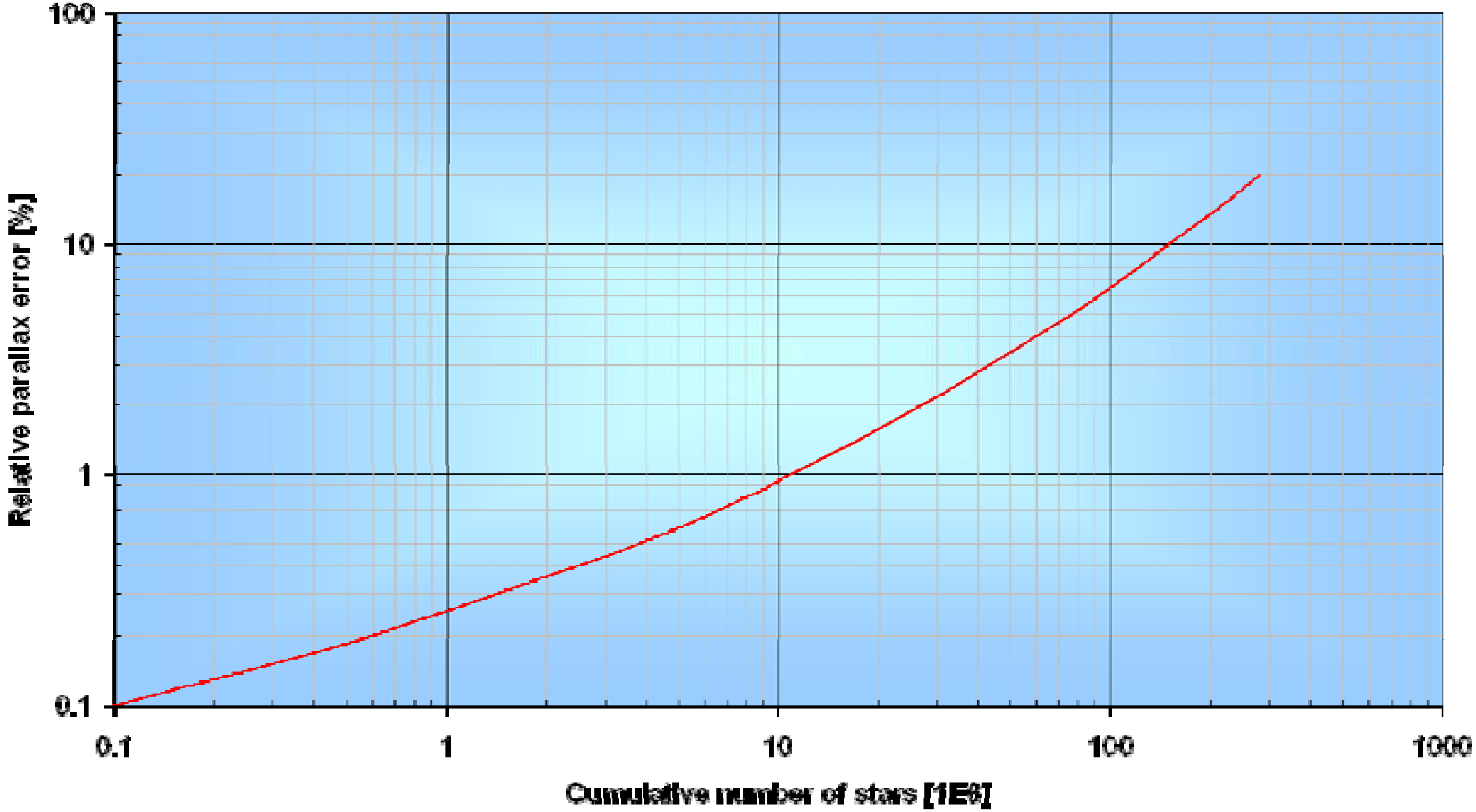
End-of-life parallax errors



1. 6 < G < 12: bright-star regime (calibration errors, CCD saturation)
2. 12 < G < 20: photon-noise regime, with sky-background noise and electronic noise setting in around G ~ 20 mag

Figure courtesy J. de Bruijne ESA

Parallax statistics



Data, based on GUMS, courtesy Xavier Luri and Lennart Lindegren

Photometric passbands: G, BP & RP

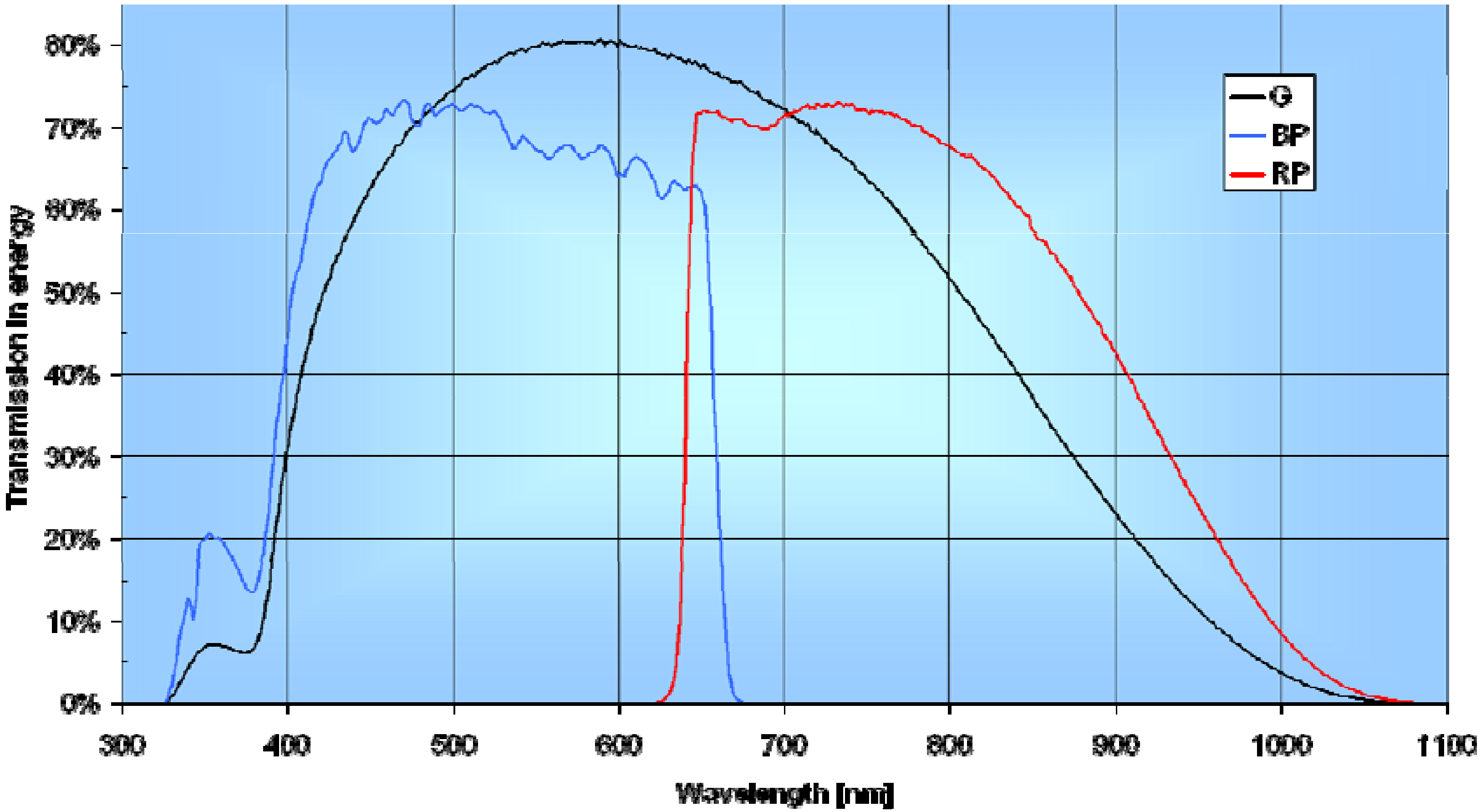
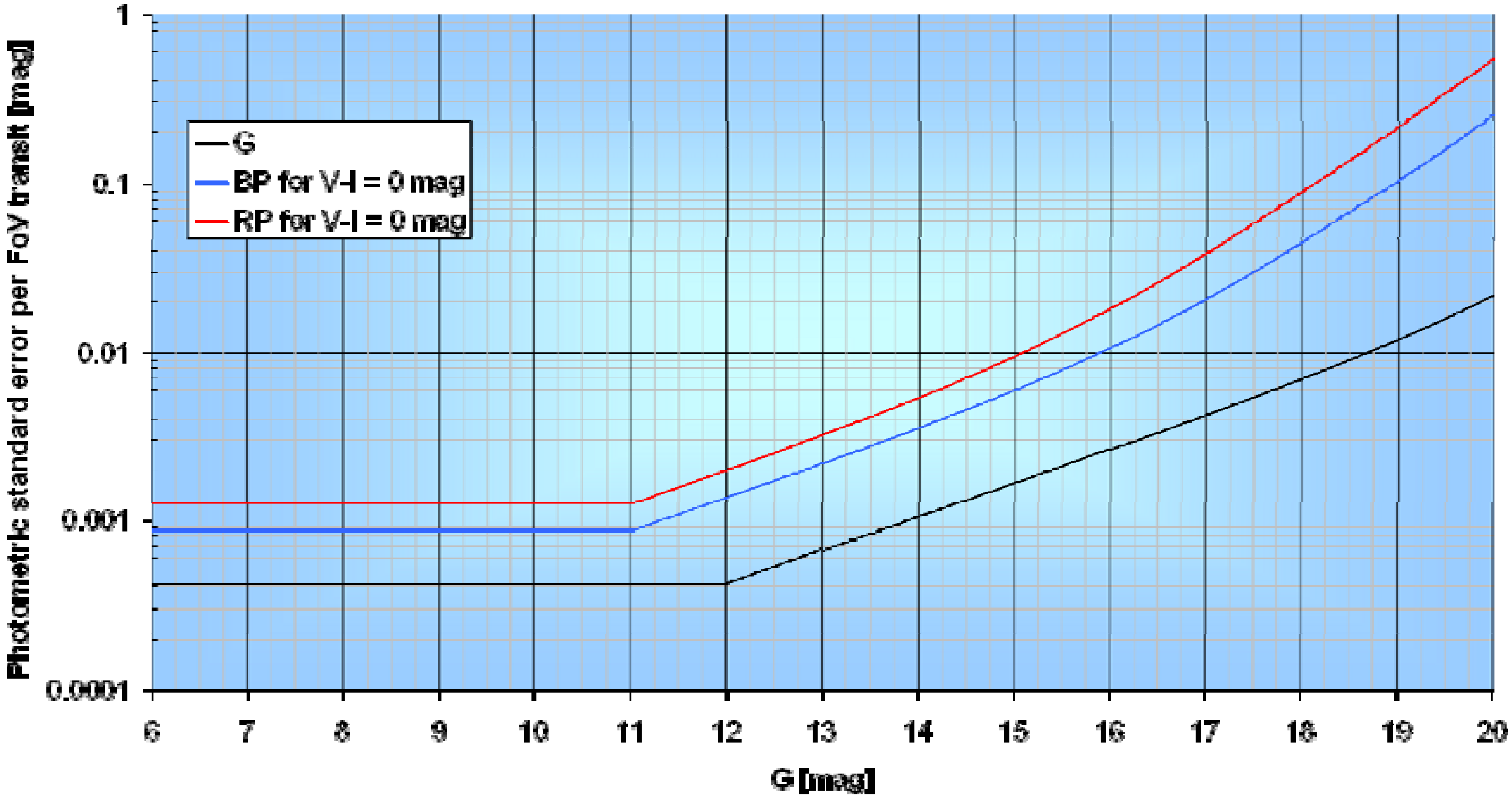


Figure courtesy J. de Bruijne ESA

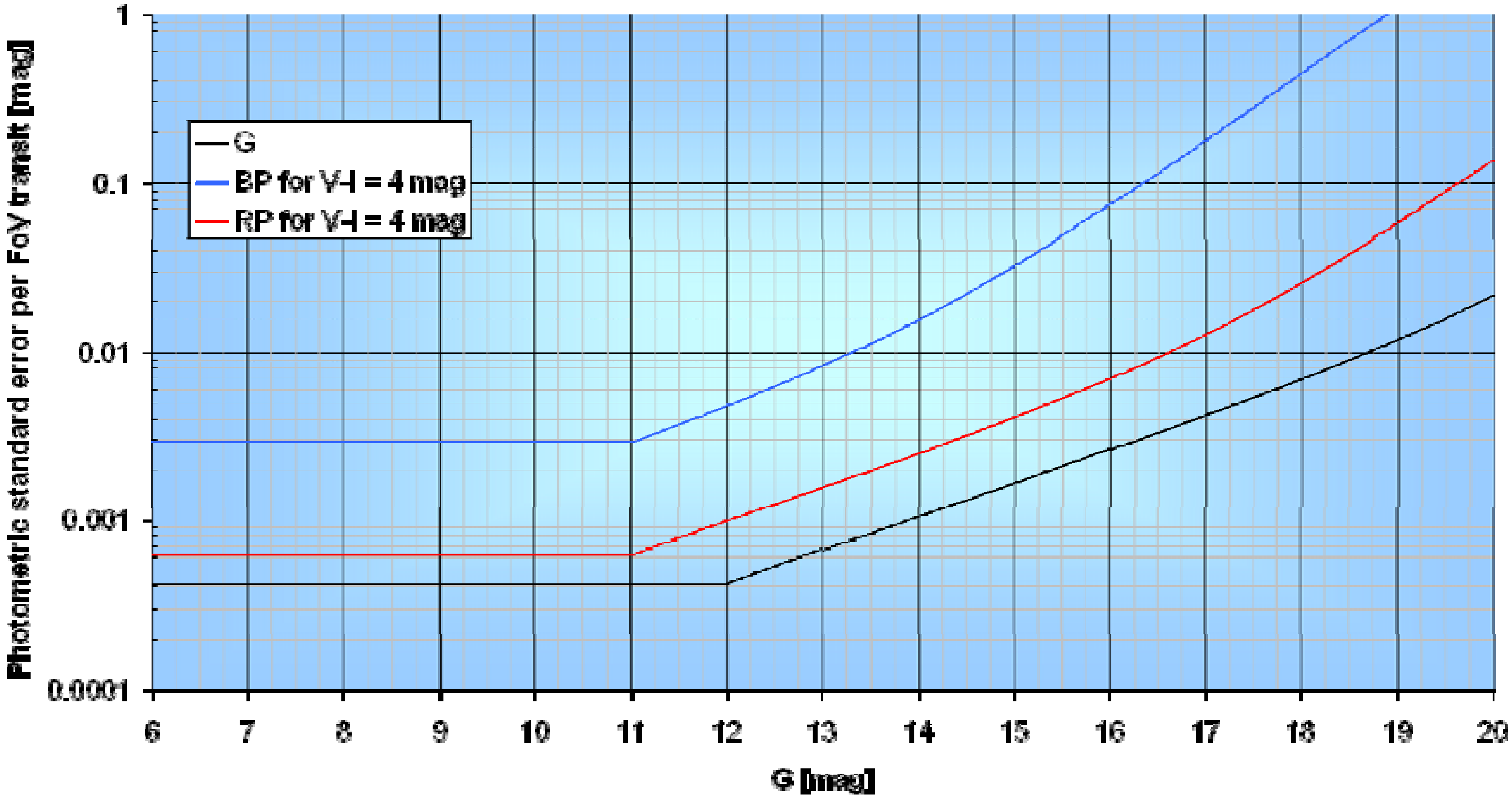
Transit-level integrated photometry



Sky-average end-of-mission number of transits ~ 70

Figure courtesy J. de Bruijne ESA

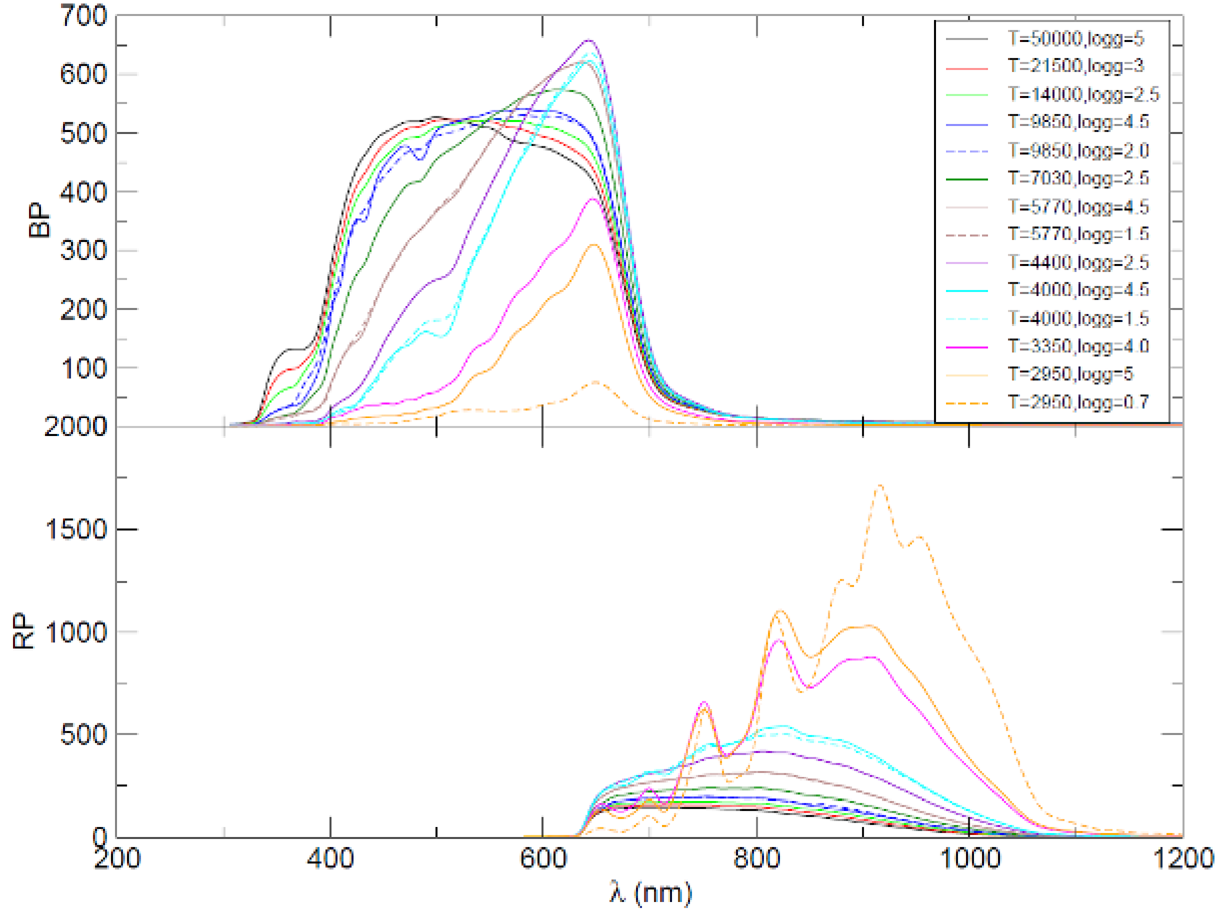
Transit-level integrated photometry



Sky-average end-of-mission number of transits ~ 70

Figure courtesy J. de Bruijne ESA

BP/RP spectrophotometric data



Illustrative BP/RP spectra for 14 solar-metallicity stars at G = 15 mag (2010A&A...-523A..48J)

Figure courtesy J. de Bruijne ESA

Goals: astrophysical parameters, for instance extinction within ~0.1 mag, surface gravity within ~0.2 dex, metallicity within ~0.2 dex, and effective temperature within ~200 K (2010MNRAS.403...96B)

End-of-life radial velocity errors

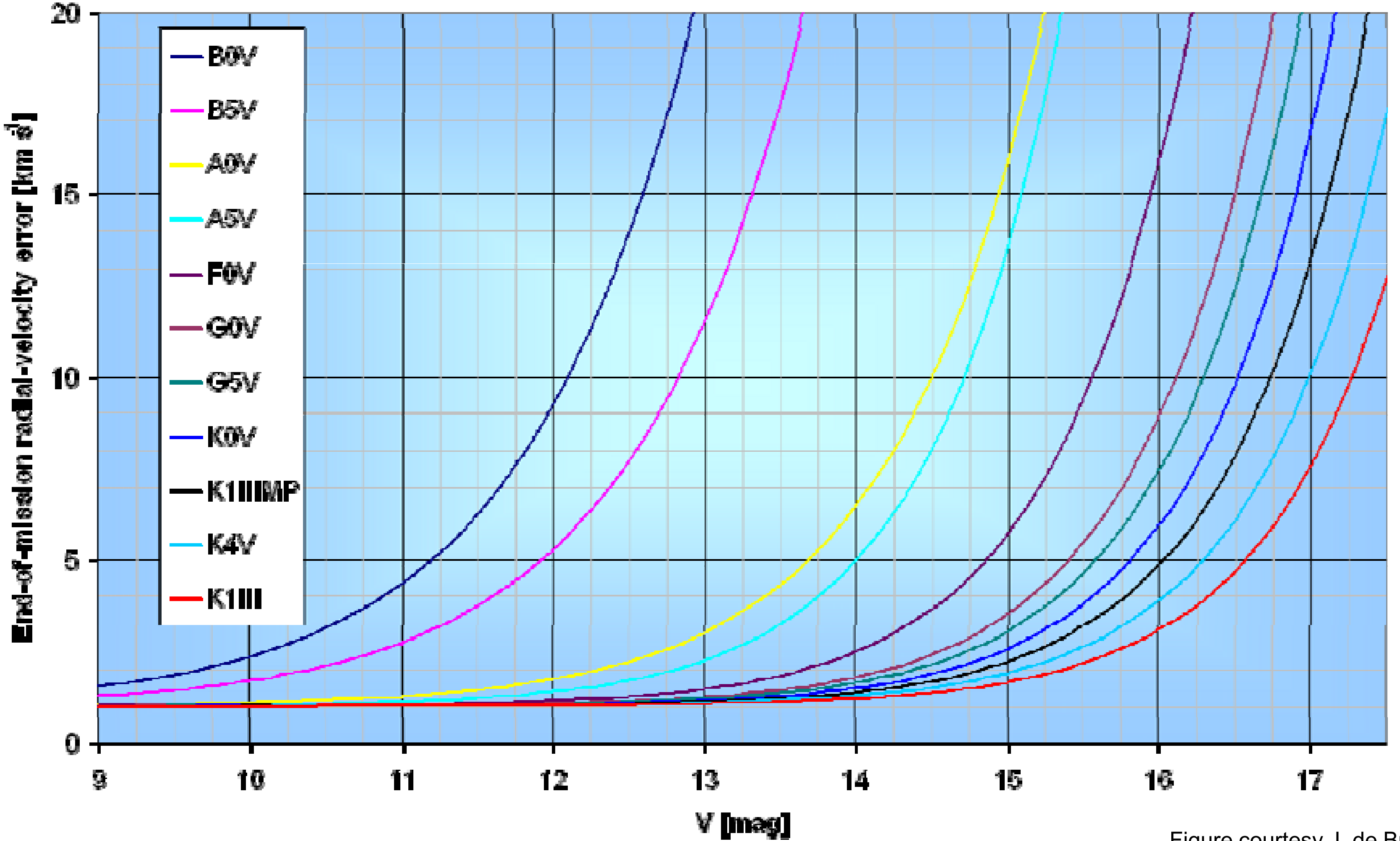


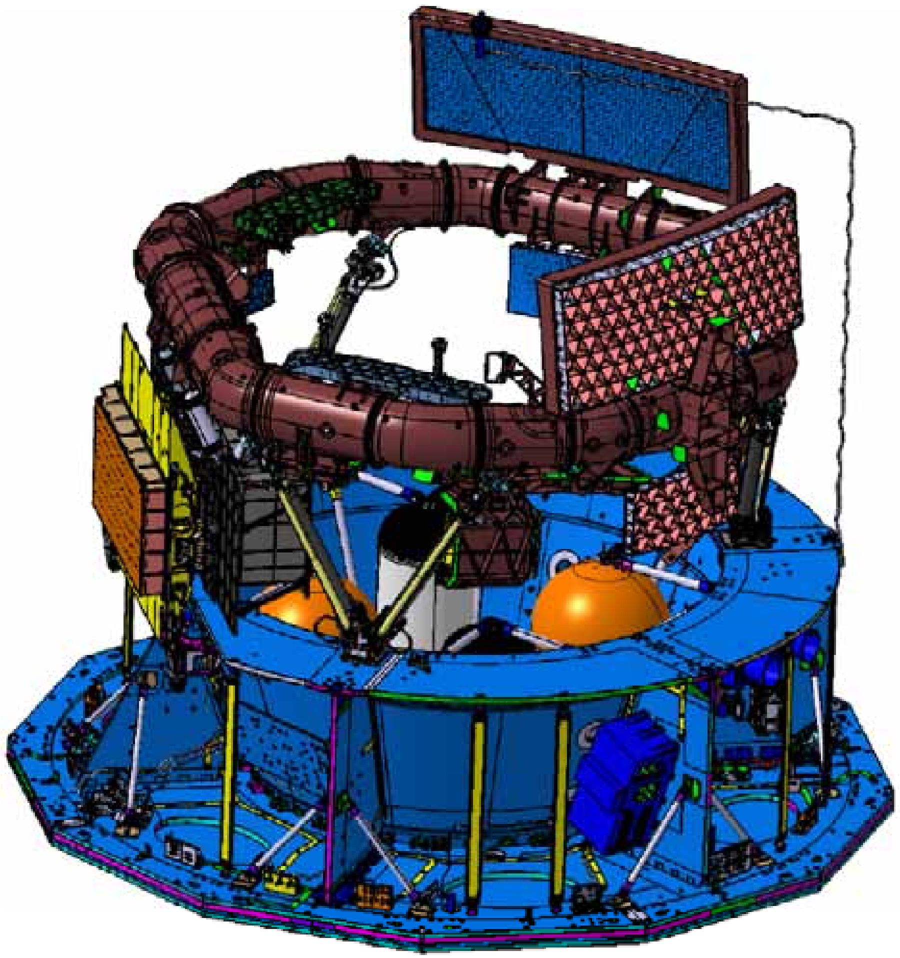
Figure courtesy J. de Bruijne ESA

Launch pad is ready

Figure courtesy ESA and Arianespace

Spacecraft is getting there ...

“Naked view” of



Payload Module
(instrument)

Service Module
(support
functions)

Figure courtesy EADS Astrium

Service Module: testing of units

Figure courtesy EADS Astrium

Deployable Sunshield Assembly (DSA)

Figure courtesy ESA, SENER, RUAG, and EADS Astrium

Torus and bipods are ready

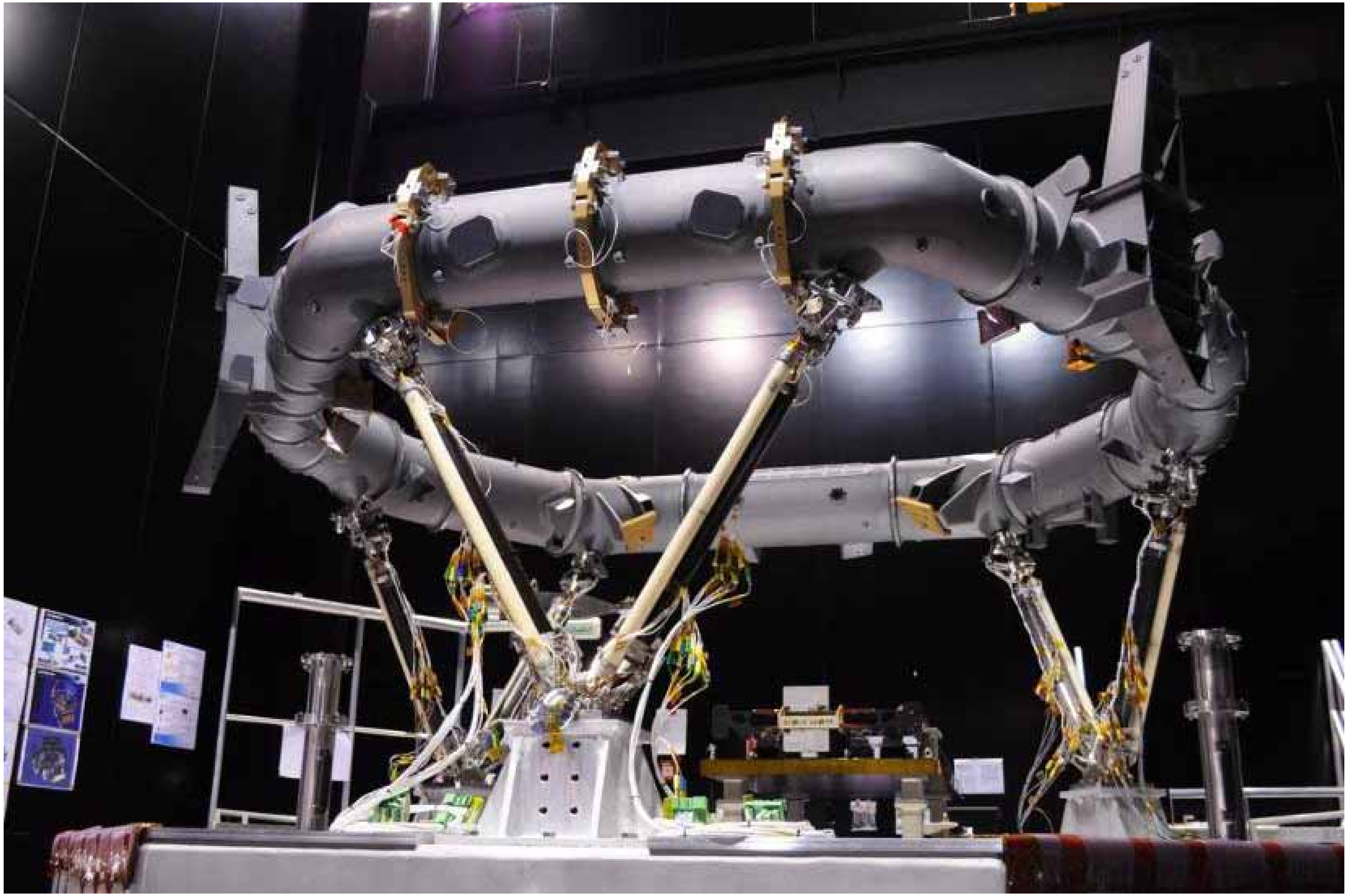


Figure courtesy RUAG, Boostec, and EADS Astrium

Nine of the ten mirrors are ready

Figure courtesy EADS Astrium, Sagem, and Boostec

Seven mirrors have been integrated

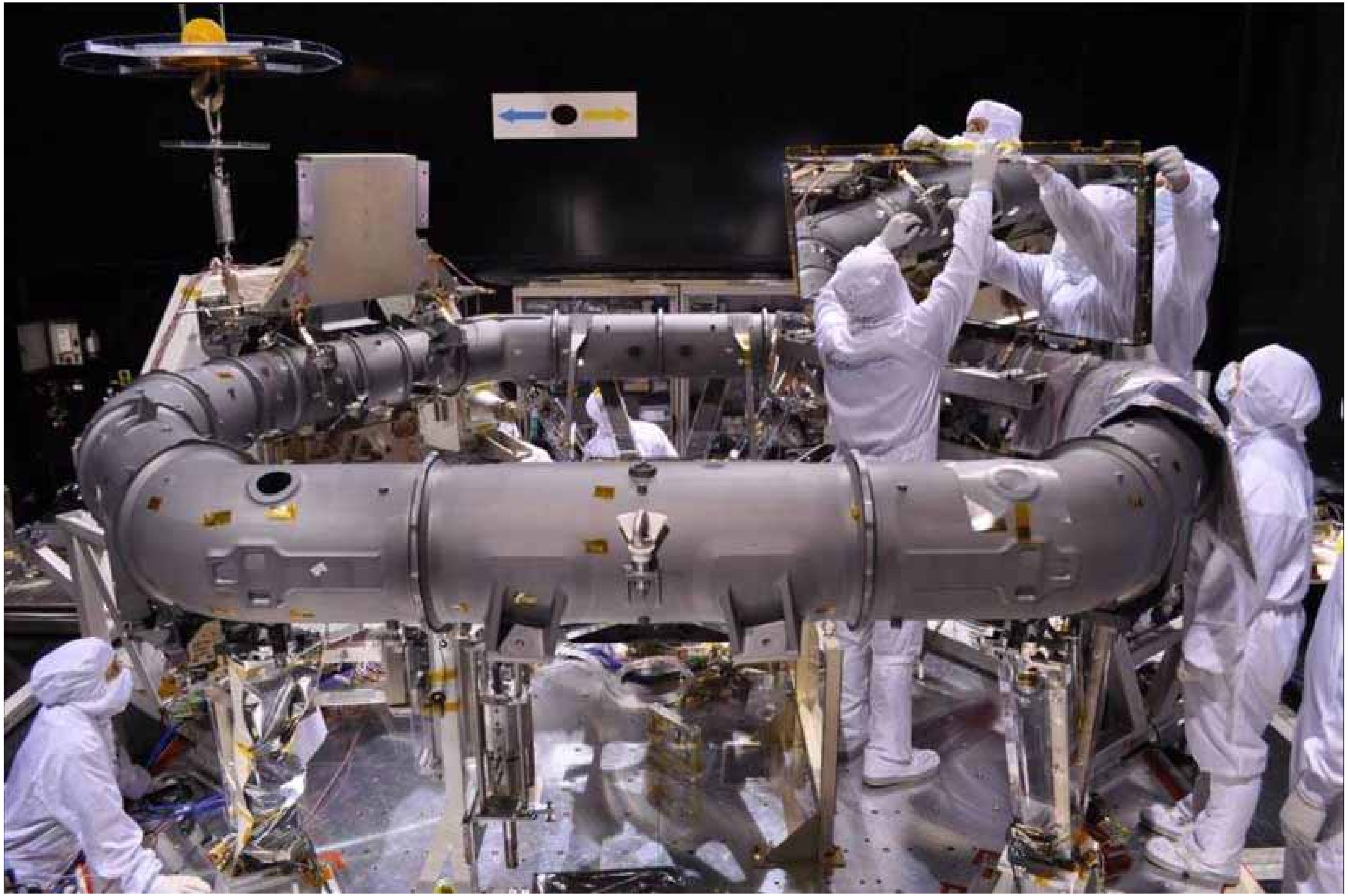
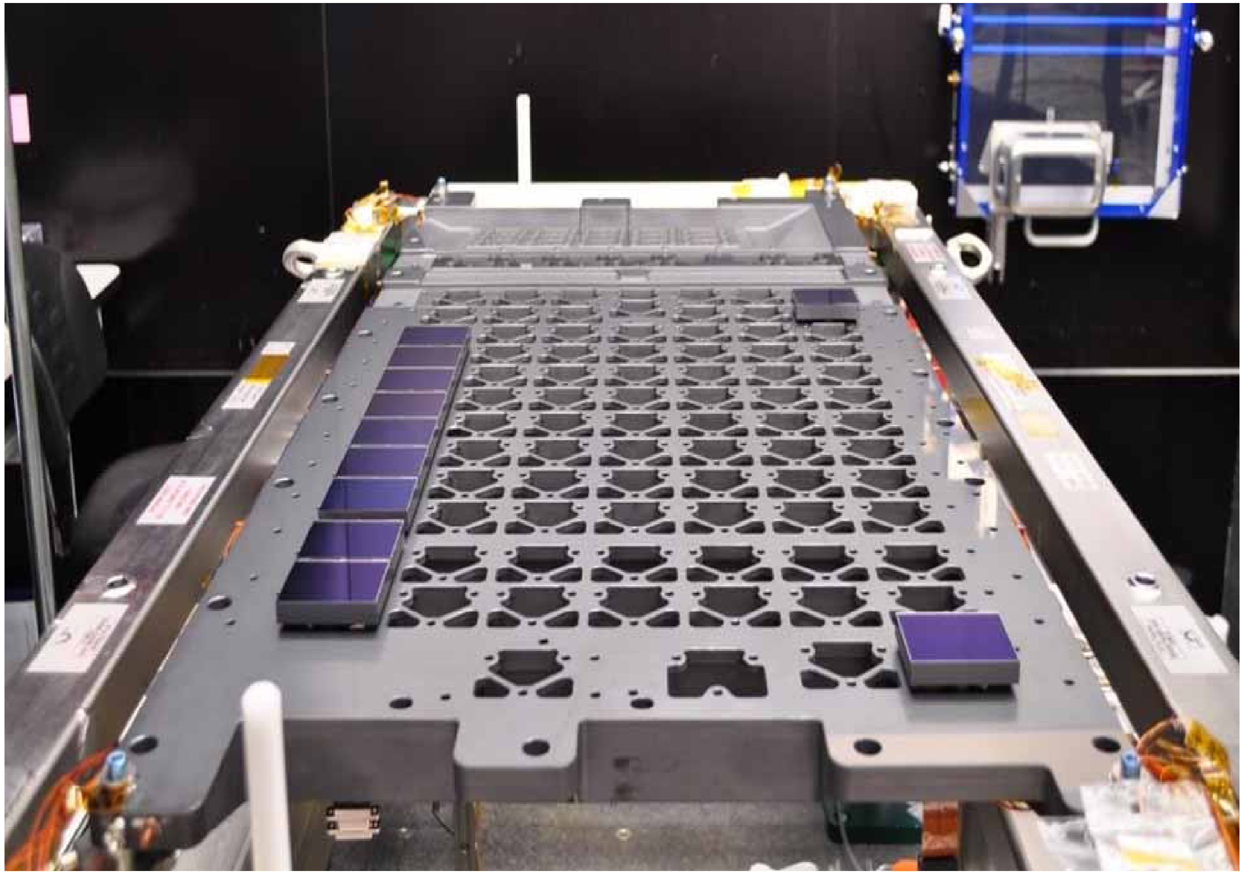


Figure courtesy EADS Astrium

Focal-Plane-Assembly (FPA) integration



All CCDs are ready and 80/106 have been mounted

Figure courtesy EADS Astrium

BP and RP prisms are ready

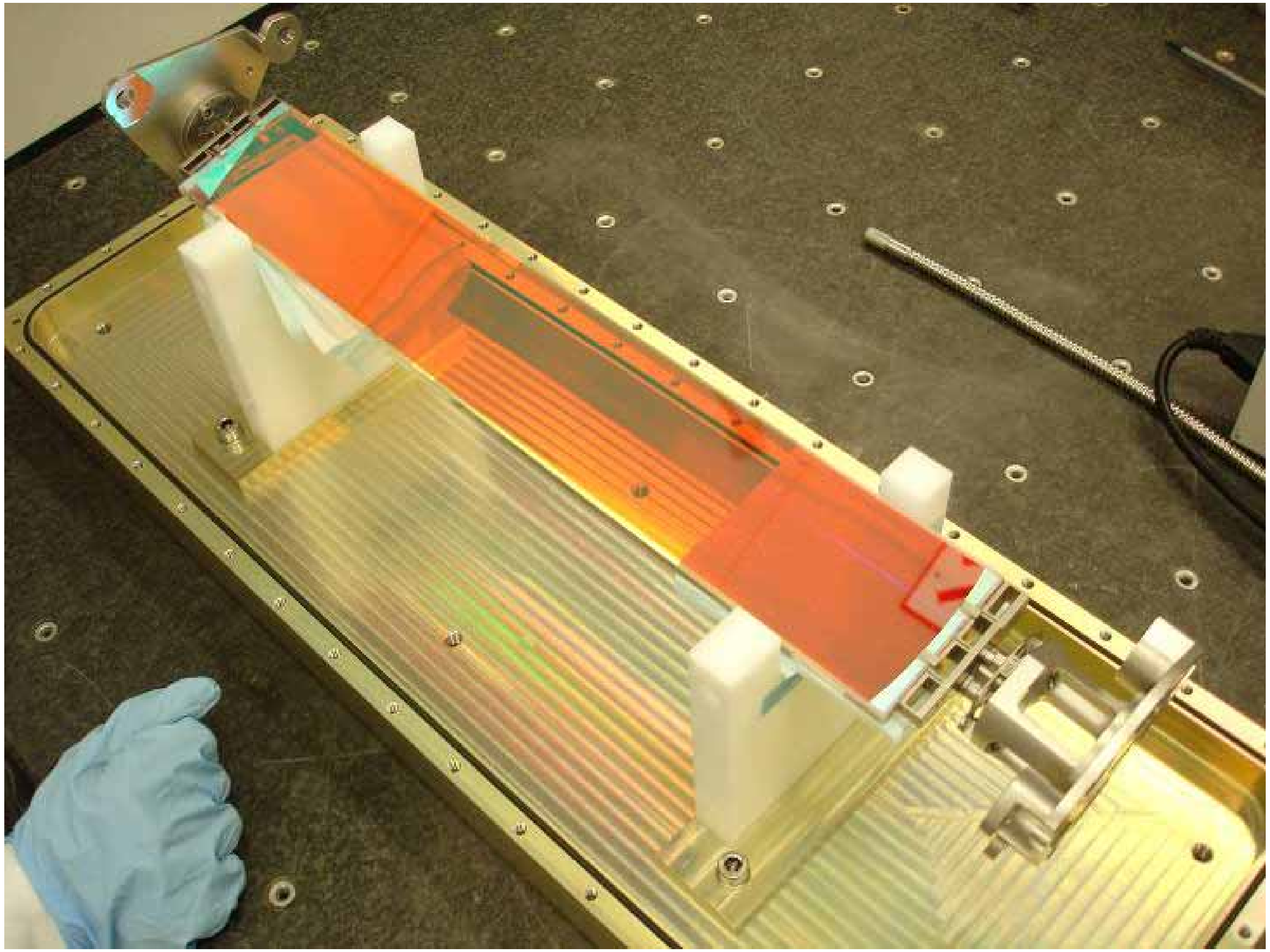


Figure courtesy SESO, Selex Galileo, and EADS Astrium

RVS spectrograph is being integrated …

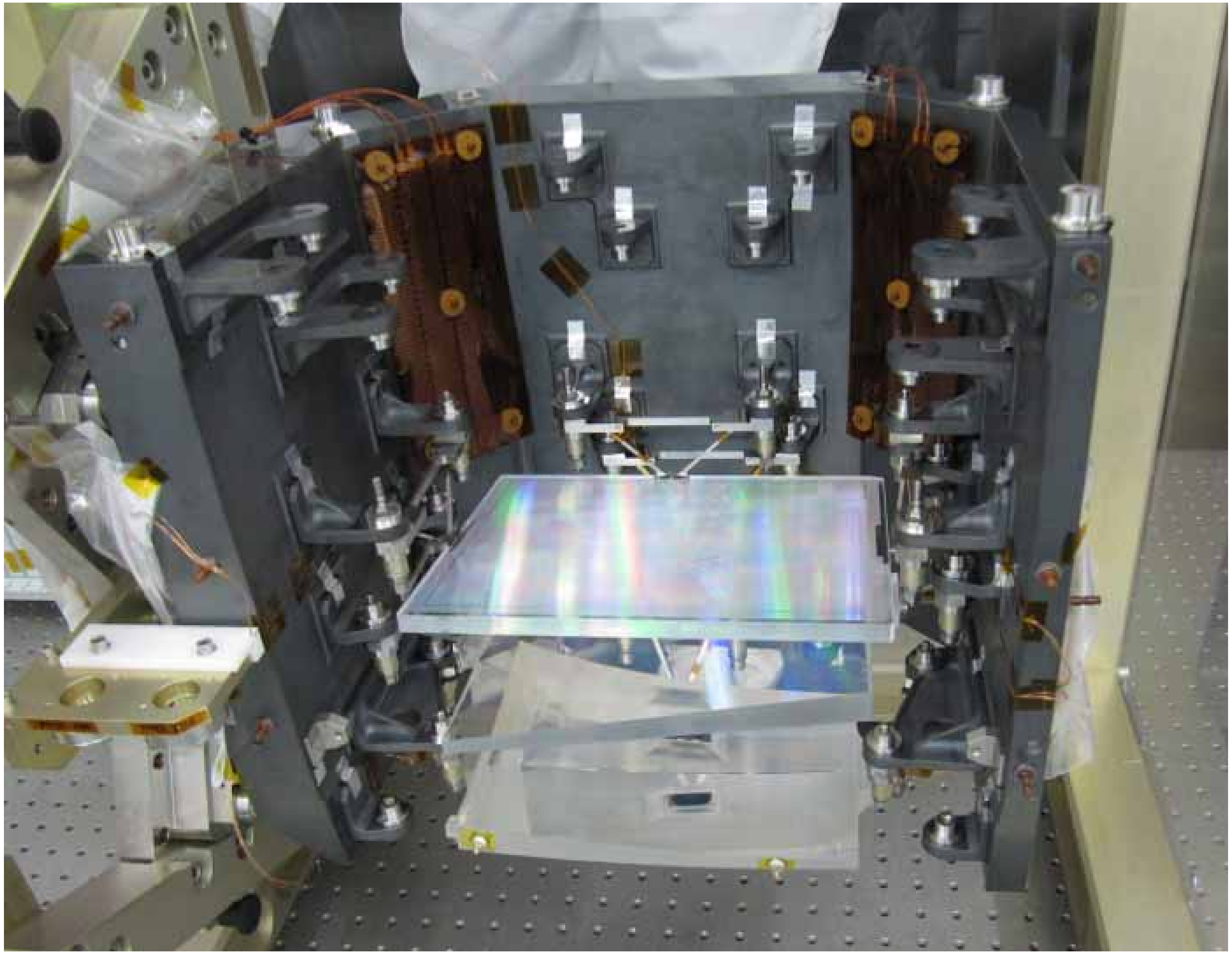
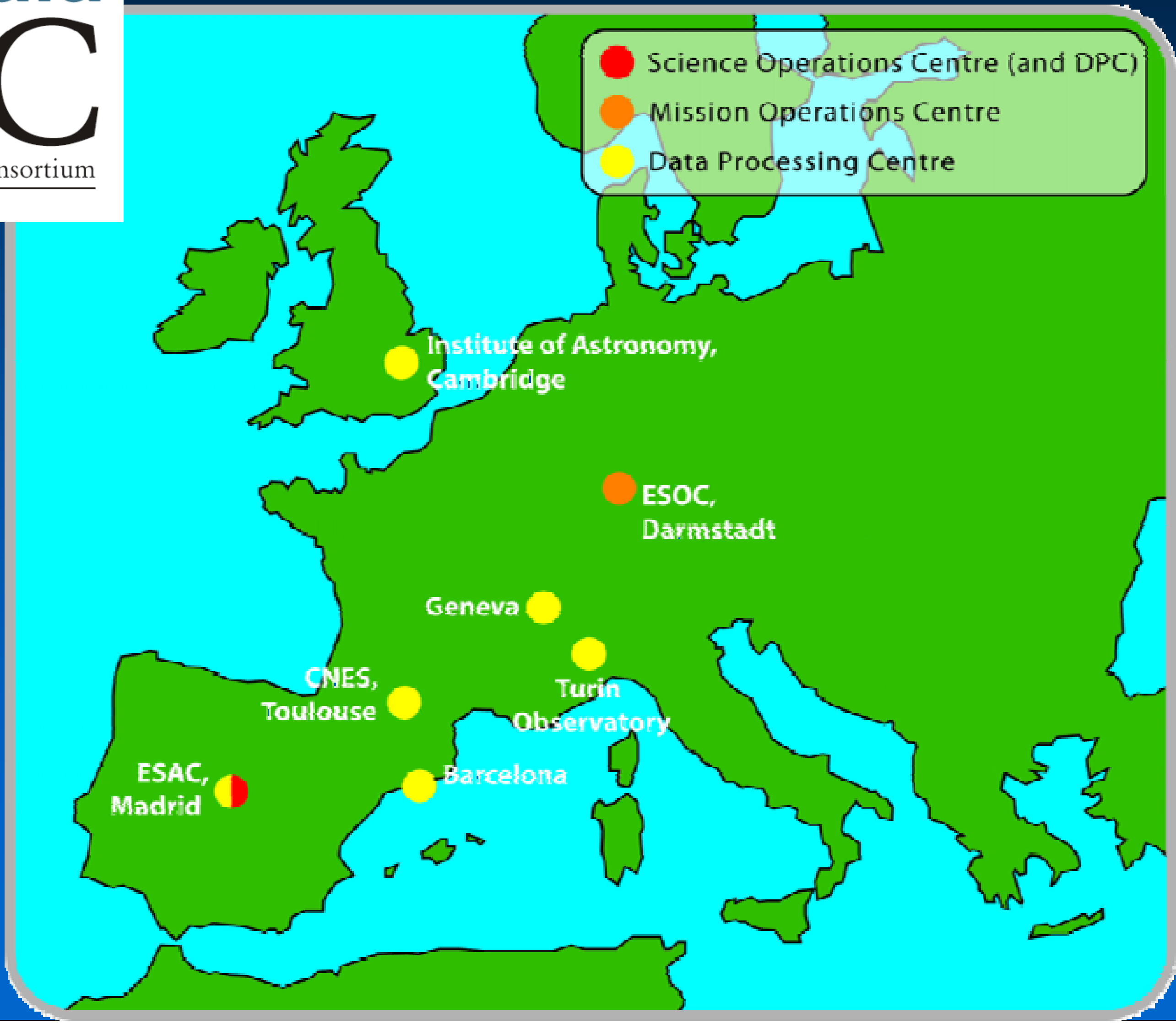


Figure courtesy IOF, Selex Galileo, Boostec, and EADS Astrium

The Gaia data reduction

Gaia
DPAC
Data Processing & Analysis Consortium



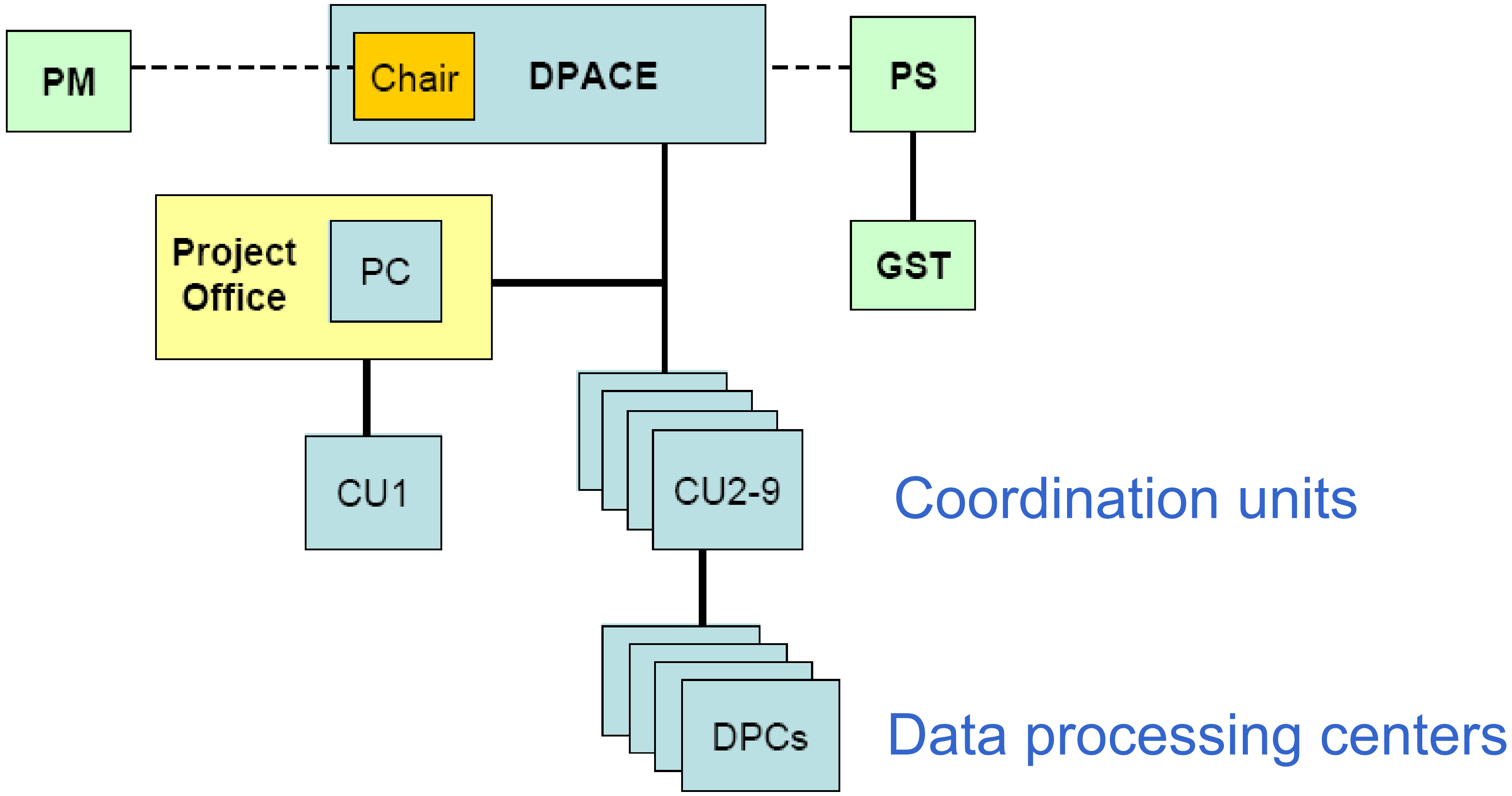
Data Processing and Analysis Consortium

- Formed to answer the Announcement of Opportunity (AO) for Gaia data processing

- Involves large number of European institutes and observatories (>400 people)

- The science community must fund the majority of the Gaia processing (not ESA)

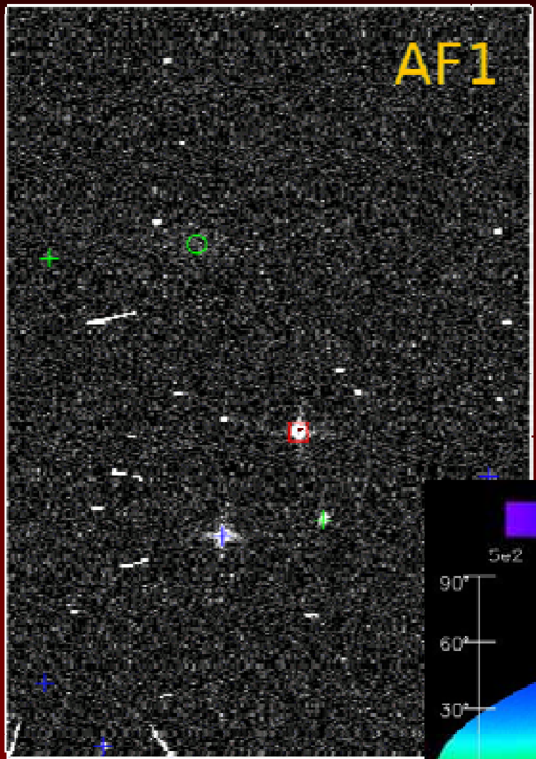
DPAC organigram

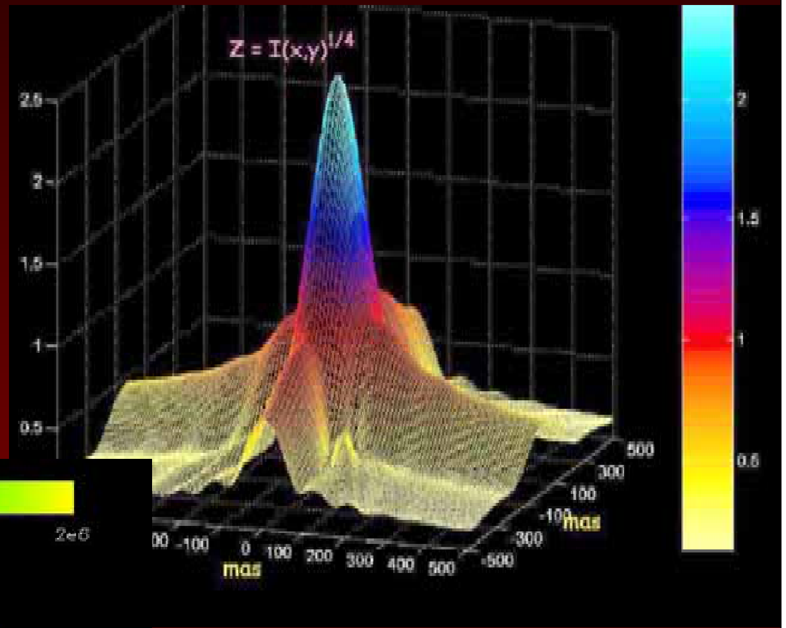


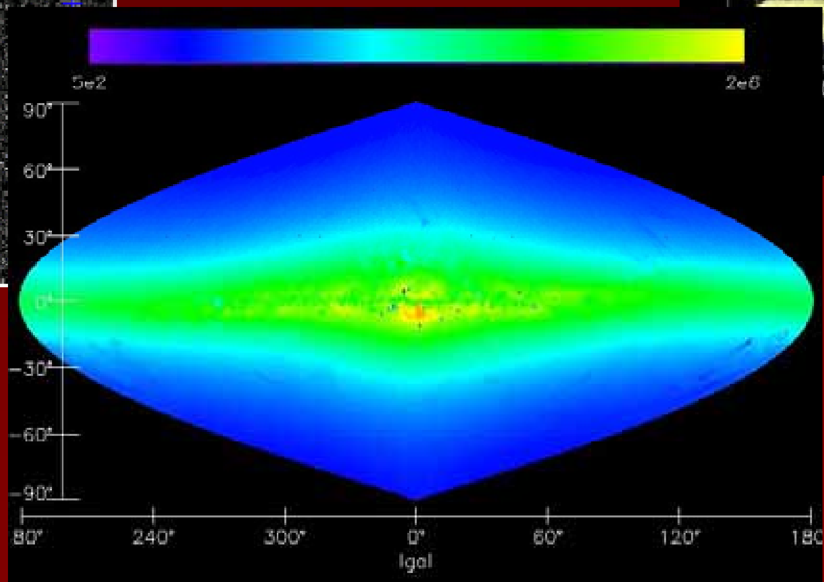
DPAC coordination units

- CU1: System Architecture
- CU2: Data Simulations X. Luri (manager)
- CU3: Core Processing
- CU4: Object Processing
- CU5: Photometric Processing
- CU6: Spectroscopic Processing
- CU7: Variability Processing
- CU8: Astrophysical Parameters
- CU9: Catalogue Access

The Gaia simulator







The CU2

The CU2 task is to cover the simulation needs for the work of other CUs, ensuring that reliable data simulations are available for the development and testing of the various stages of the data processing development.

Note: before the creation of the CU2 the UB and ObsPM had already been working for 5 years on the generation of Gaia simulations for instrument design and data reduction feasibility demonstration.

The CU2 operates in close coordination with the rest of the DPAC:

- Requests for simulated data are sent in each cycle by the different CUs in DPAC
- The development of the different modules of the simulator is jointly agreed, taking into account each CU's needs, their global priorities within the overall consortium, the availability of the simulation models (which can depend on industrial tests or other CU accuracy models) and the available CU2 manpower.

Key point:

we are presently serving the **simulation needs of a data reduction consortium**. We are not doing simulations for instrument design (although we did in the past) and we are just now starting to provide simulations to prepare the science exploitation (simulated Gaia catalogue to be released in 2012).

This has set our priorities and design choices up to now.

We are part of a larger development effort:

- We follow common coding standards (document in LL), use common libraries (GaiaTools, GPDB) and interfaces (gbin format)

- We are subject to the 6-month development cycle agreed in the consortium (an updated version of the simulator every 6 months)

We have not been able to cover absolutely all the needs:

- CTI simulator
- CU7 using real catalogues for variable studies

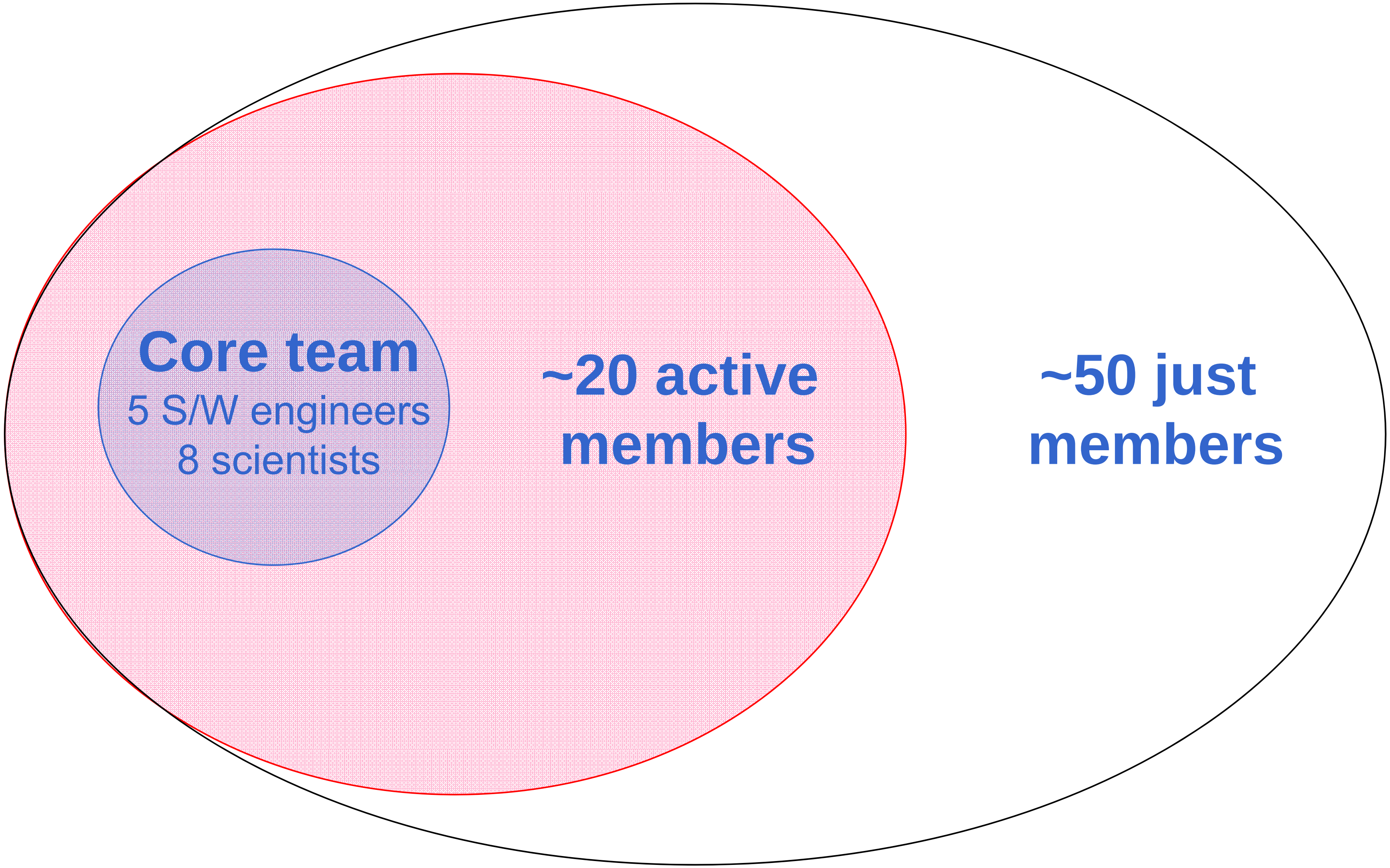
Complexity:

The Gaia simulator is (five) four simulators in one:

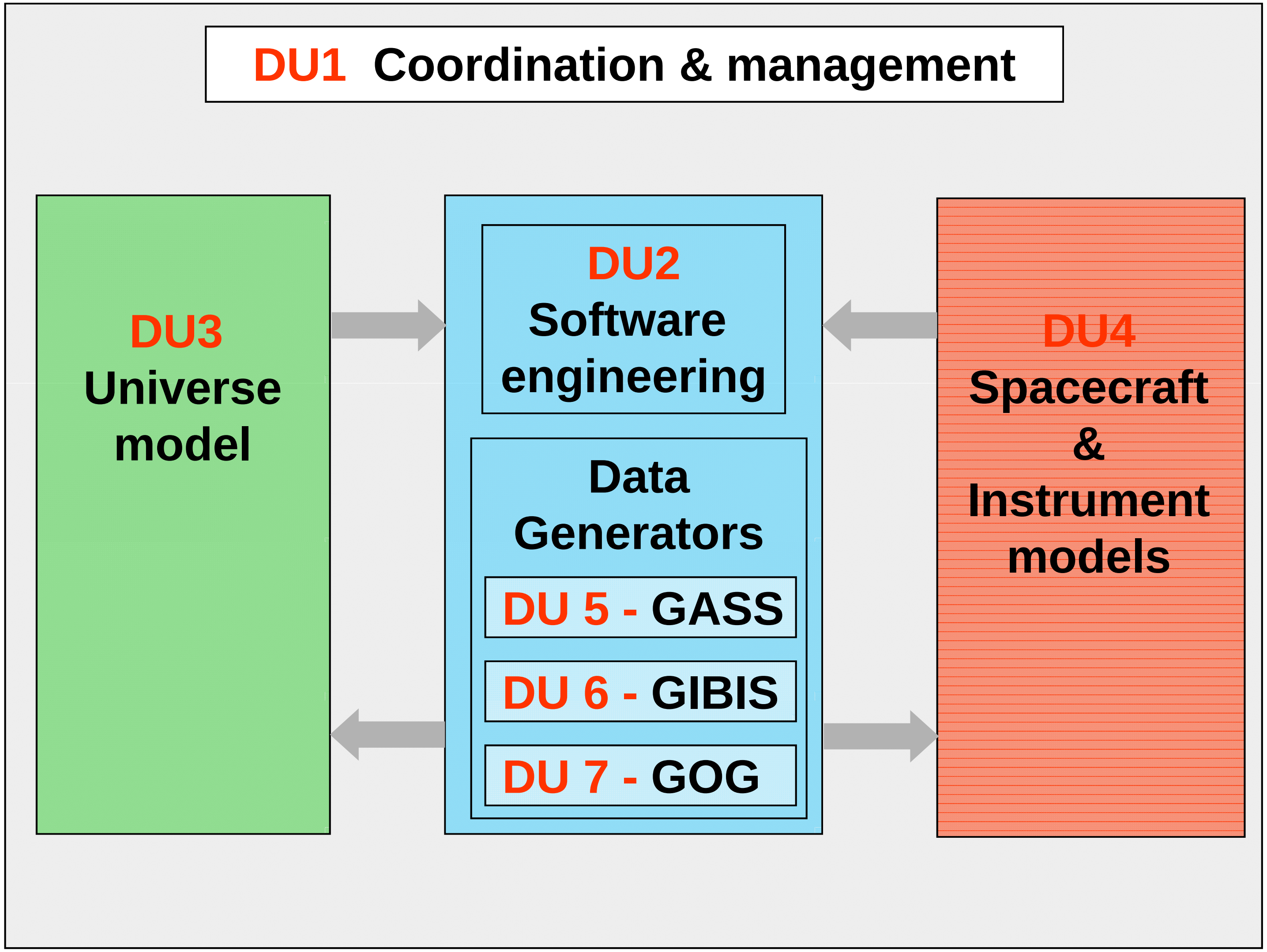
- (The spacecraft)
- The Sky mapper
- The astrometric field
- The spectrophotometer BP/RP
- The Radial Velocity Spectroscope

It's like four different interlinked instruments

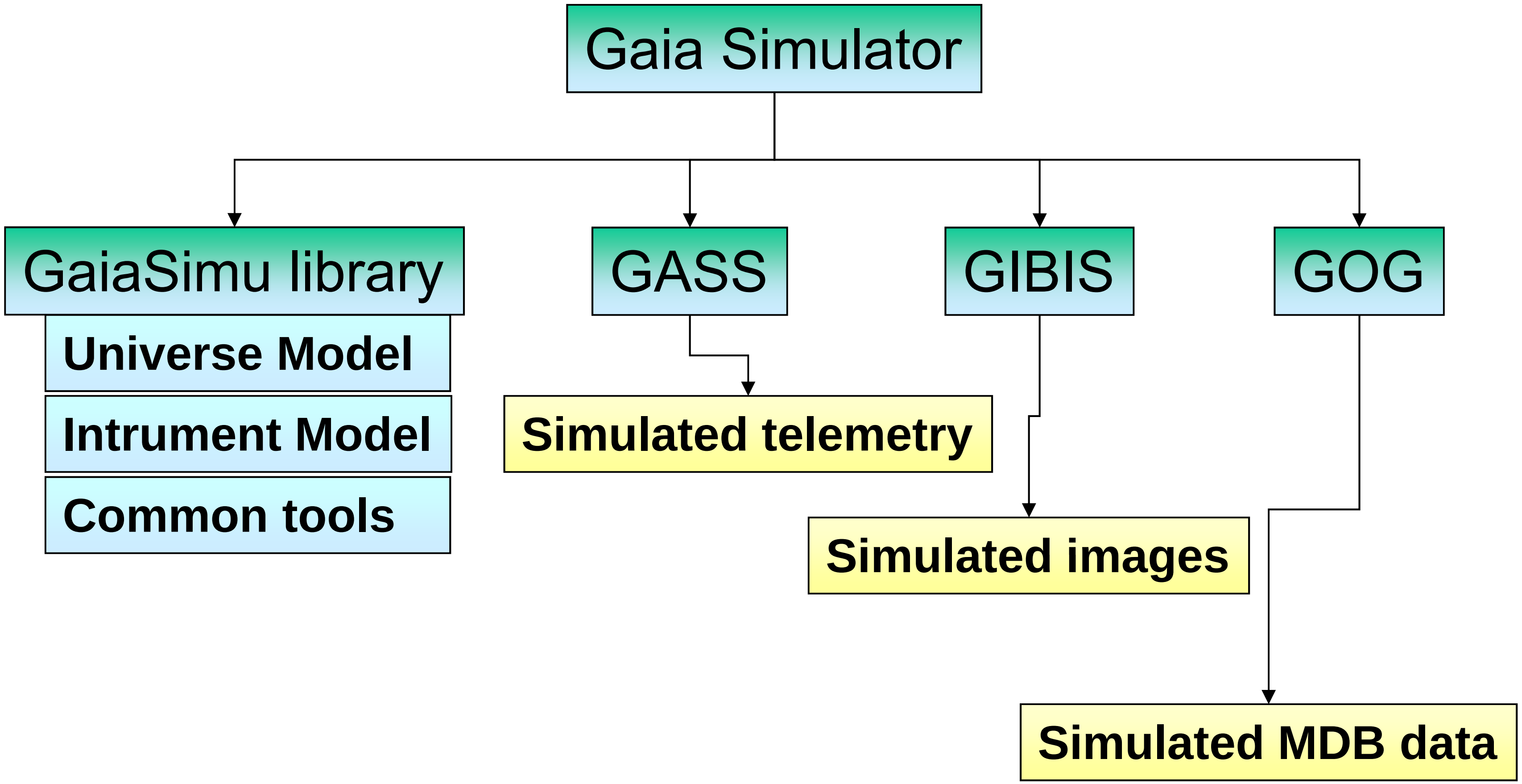
CU2 membership



CU2 structure



CU2 product tree



The GaiaSimu library

GaiaSimu library: this library is the basis for the development of the Gaia simulator. It contains the common models used by the data generators to produce simulations of the Gaia observations. It is composed of three parts:

1. The instrument model: models of the Gaia instruments and elements of the spacecraft
2. The universe model: model of the objects in the sky that Gaia will observe, with all its physical characteristics
3. Common tools: a toolbox for use throughout the simulator

The Universe Model

Solar System

- Sun, Earth, Moon
  (not for observation)
- Planets and satellites
- Minor bodies
  - Asteroids
  - Comets
  - Kuiper belt
- Other components
  - Zodiacal light
  - Solar wind
  - Etc.

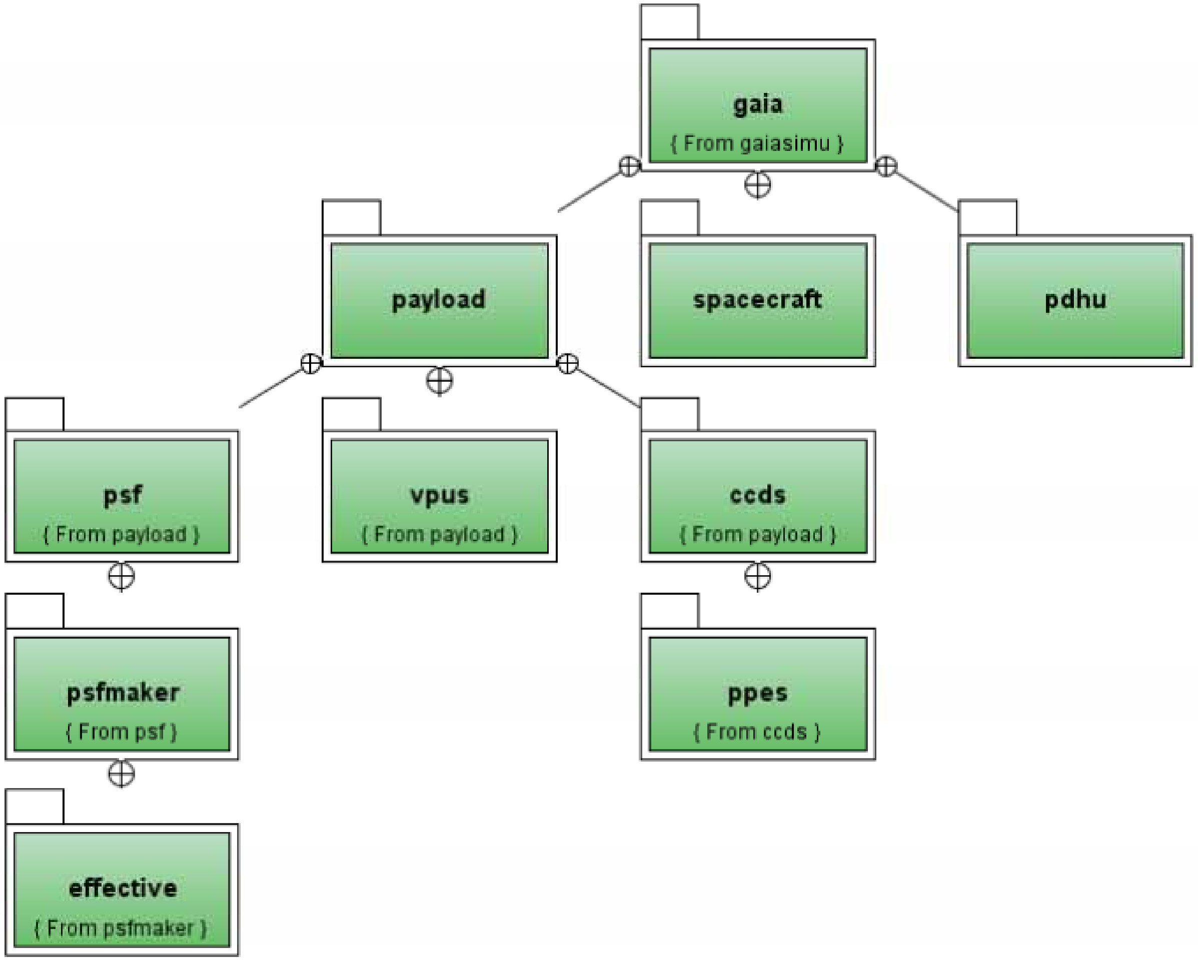
Our Galaxy

- Field stars
  - **“Normal”**
  - **Multiple systems**
  - **Variable stars**
- Stellar clusters
  - Open clusters
  - Globular clusters
  - OB associations
  - Stellar streams
- Extended objects
  - Planetary nebula
  - HII regions
  - Reflection nebula
- Other components
  - Galactic diffuse light
  - Extrasolar planets

Extragalactic objects

- Galaxies with resolved structure (LMC/SMC)
  - Field stars
  - Stellar clusters
  - Surface brightness
  - Supernovae
- Galaxies with unresolved structure
  - Surface brightness
  - Supernovae
- QSO
- Other components
  - Diffuse extragalactic light

The Instrument Model



The Data Generators

The simulator comprises three *data generators*, software components that use the GaiaSimu library to produce specific types of data for the DPAC. The three data generators are:

- The GAia System Simulator (GASS)
- The Gaia Instrument and Basic Image Simulator (GIBIS)
- The Gaia Object Generator (GOG)

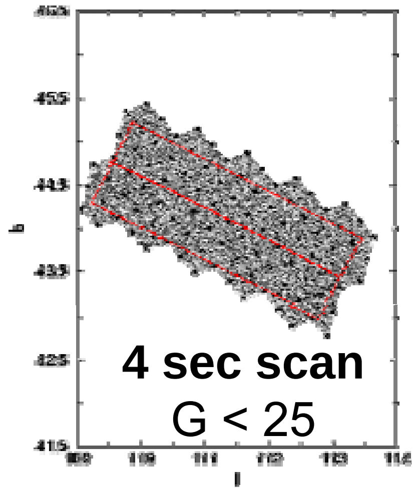
GASS

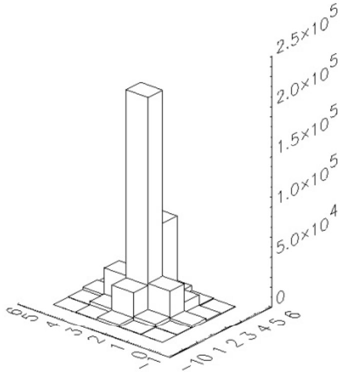
This data generator provides simulations of the telemetry stream of the mission based on some simplications of the instrument and Universe models allowing a large amount of data to be simulated over a signicant period of time.

GASS provides realistic data for:

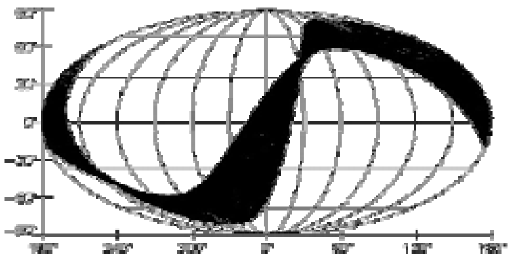
- Predictions to be used for mission design.
- Filling of test databases
- Testing of core reduction algorithms
- Evaluation of mission performances, in particular for peculiar objects (binary stars, NEO's, extrasolar planets,...)

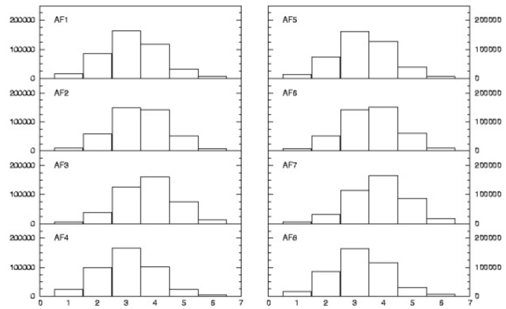
GASS is run in large computer clusters and specially at the **Mare Nostrum supercomputer**.





ASM





ASTRO patches

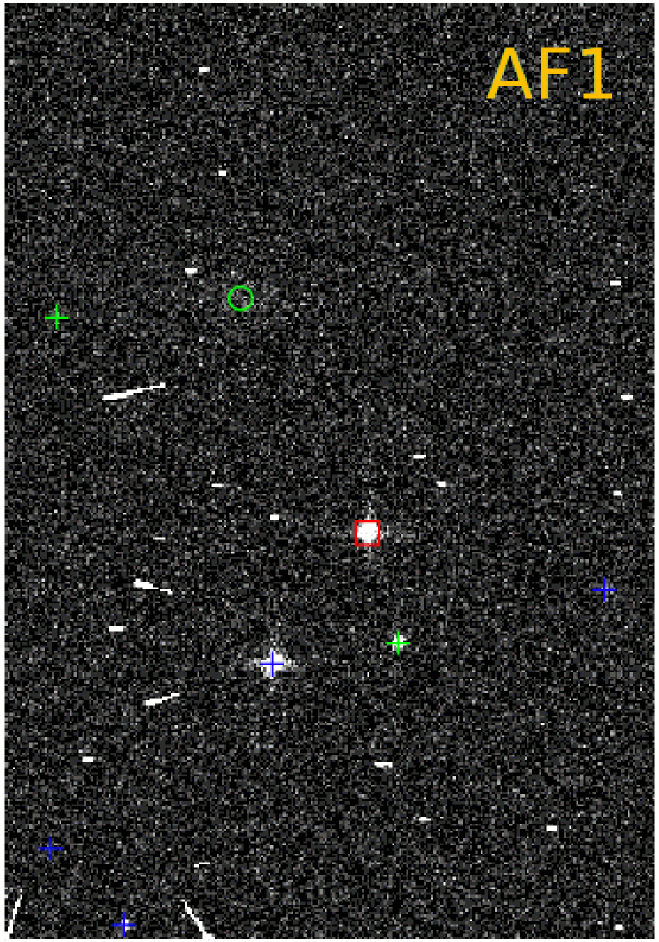
GIBIS

This data generator provides simulations of the data at the pixel level. The resulting simulations are as realistic as possible and restricted, due to computing time limitations, to a region of a sky over a short period of time (~1 hour).

GIBIS provides realistic CCD images for:

- Instrument design
- PDHE design
- Detection & selection algorithm development
- Scientific mission design (e.g. RVS)
- Detailed analysis of reduction algorithms

GIBIS is available as a web service and is also run in batch mode in computer clusters.

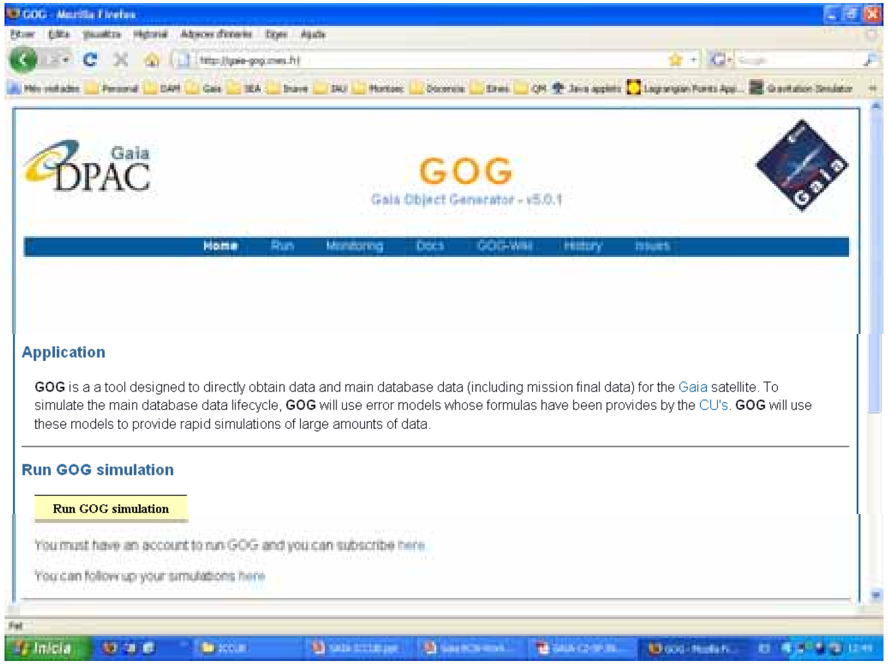


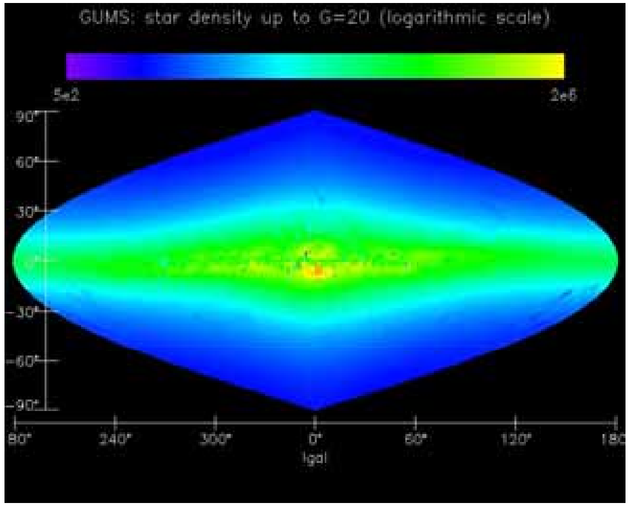
GOG

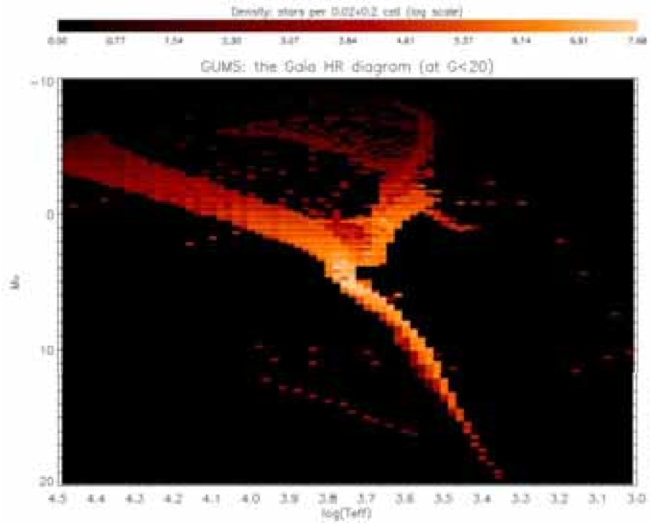
This data generator provides simulations of number counts and lists of observable objects from the Universe model and, for a given source or a collection of sources, simulations of intermediate and end-of-mission Gaia data.

GOG aims to simulating the contents of the MDB at any stage of the data processing.

GOG is available as a web service and is also run in batch mode in computer clusters.







Summary

The Gaia Simulator code amounts today to more than **100.000 lines of code** (+DPAC libraries) and has produced **several terabytes of simulated data** in the last years that have been used for mission design and development and testing of the initial versions of some reduction algorithms.

Lessons learnt

- **Start very early.** Simulations (specially for instrument design) are needed from the early stages, and can not wait.
- **Start with a good, design, but allow for changes.** Scientific projects can not be as rigid as S/W engineers would like, but can not be redesigned/enlarged/adapted as often as scientists ask for
- **Expect conflicting requirements from users**
- **Defining and sticking to schedules is (very) hard.**
- **Take management very seriously**, specially if the work involves international collaboration
- **Rely on professional S/W development methods for large projects.** Amateur approaches have a limit…

Lessons learnt (ctd)

- **Enforce good programming practices.** Use tracking tools (Mantis, Hudson, SVN, metrics, …), Unit tests, etc.
- **Do not forget to in include documentation in your planning**. Specially if you work for ESA (ECSS standard)
- **In a long-term project take into account personnel mobility.** People may leave, in some cases suddenly.
- **In an international collaboration take into account:**
  - National peculiarities, specially about funding stability
  - Increased complexity of coordination
  - Not all partners are equally committed or reliable

Yes, we code in Java

- Used throughout DPAC. Early decision.

- Reasons:
  - Such a complex project requires an object oriented approach → narrowed to C++ or Java
  - Java offers a much more <u>robust and faster development</u>
  - It's not significantly slower than C, can be even faster

- Quite happy about it after 10 years

See O'Mullane, Luri et al. (2011), Experimental Astronomy (in press)

Some examples of our work

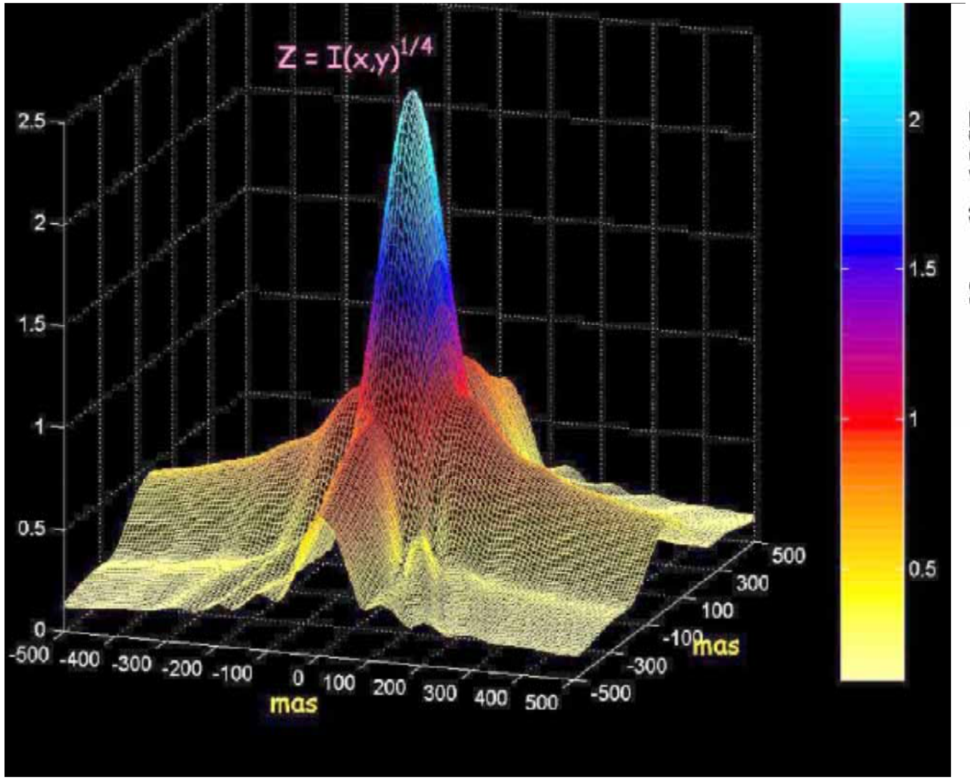
Instrument model: PSF/LSF

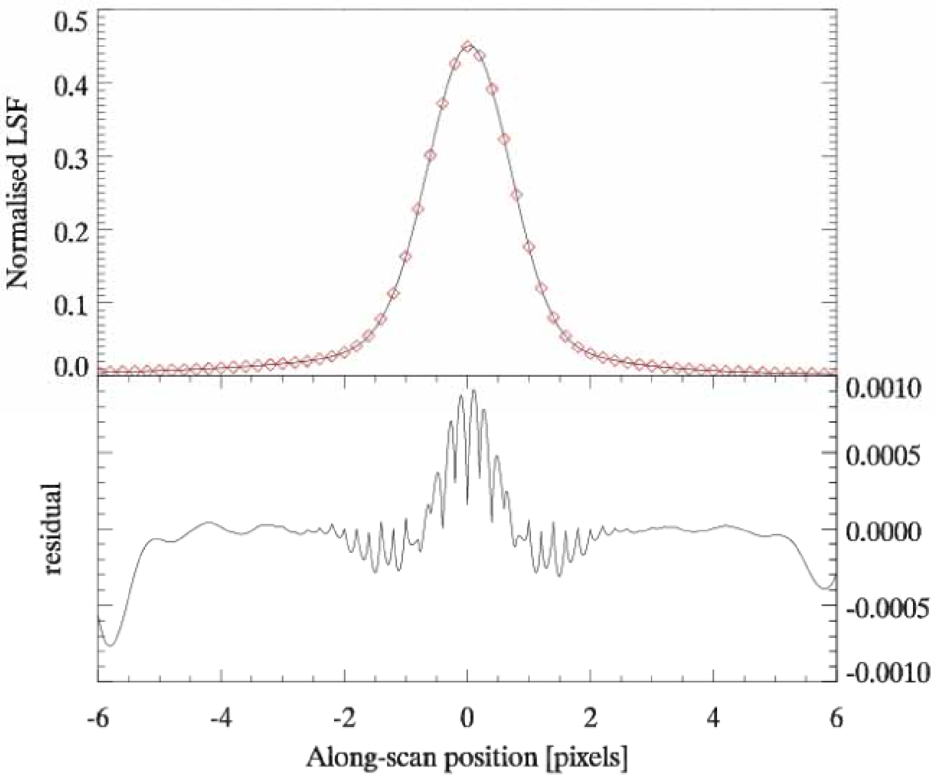
The *Point Spread Function (PSF)* describes how the image of a point source looks like when observed by Gaia. A model of the PSF is needed to produce simulations of the CCD observations.

The PSF varies from point to point of the focal plane, thus in fact a “continous variation” model is needed. Furthermore, the Gaia observations for most of the magnitude range are “compressed” into 1-D windows so what is actually observed are *Line Spread Functions (LSF).* A model for these including cut-out effects is also needed.

PSF model generation process

- Optical model specifications (code V, Astrium)
- Generation of Wave Front Errors (WFE)
- Generation of PSFs at reference focal plane positions and wavelength/color (numerical PSF library) + interpolation between them
- Generation of an LSF library at reference points (several LSFs for each PSF taking into account the AC position) + interpolation between them


Z = I(x,y)^1/4
2.5
2
1.5
1
0.5
0
-500 -400 -300 -200 -100 0 100 200 300 400 500
mas
500
300
100
-100
-300
-500
mas
2
1.5
1
0.5


Normalised LSF
0.5
0.4
0.3
0.2
0.1
0.0
residual
0.0010
0.0005
0.0000
-0.0005
-0.0010
-6 -4 -2 0 2 4 6
Along-scan position [pixels]

Universe model: GUMS

The Gaia Universe Model Snapshot (GUMS) is a full G<20 simulation of the contents of the universe model at T=0 used for validation and testing.

Several versions of GUMS have been generated at various stages of the development. We are currently at GUMS-10.

GUMS-10 Overview

The model has generated 1,000,000,000 galactic objects:

~49% single stars

~51% stellar systems formed by multiple stars and stars with planets

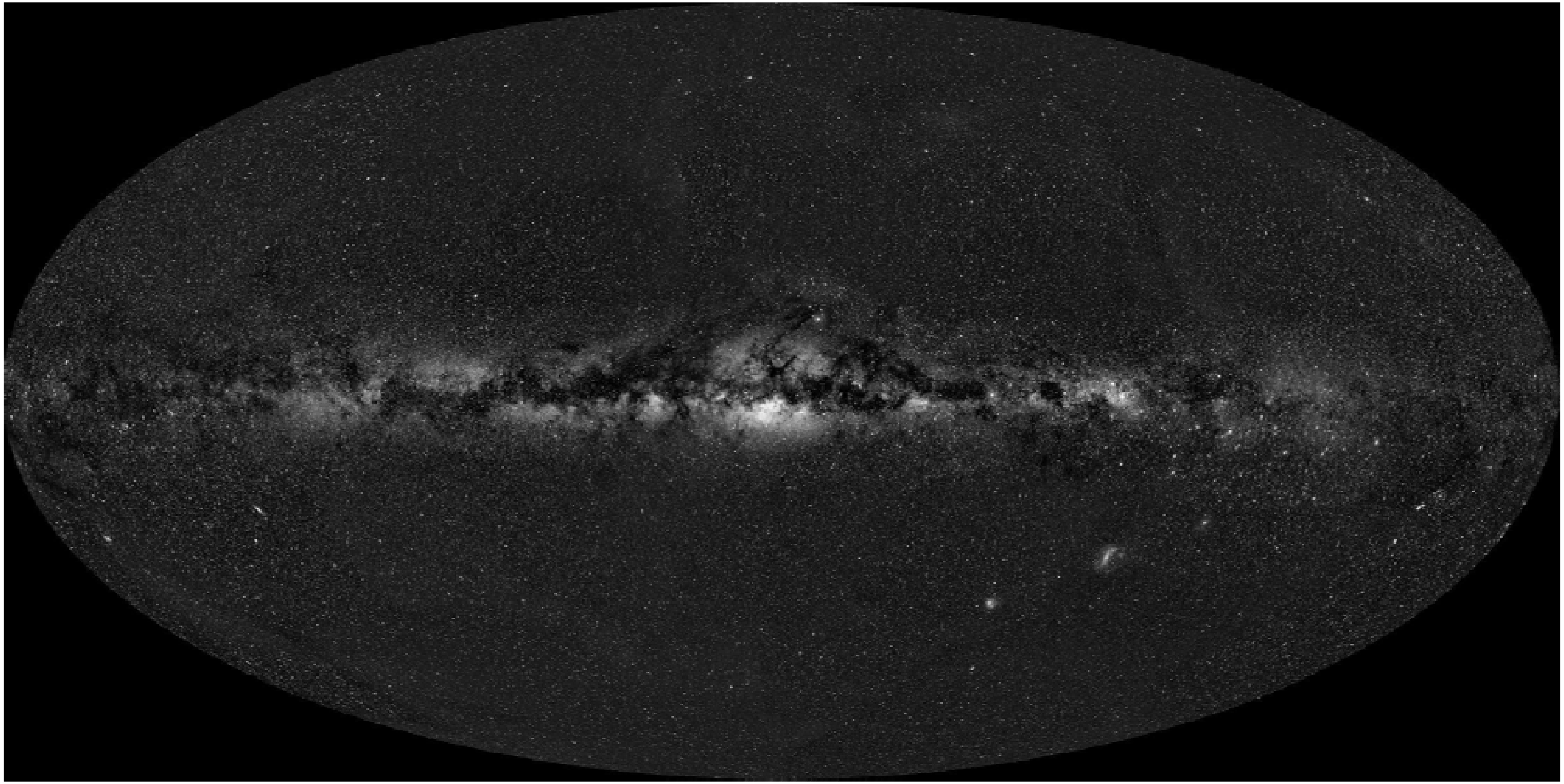
Grand total of 1,600,000,000 individual stars

Visibility based on the integrated magnitude G < 20

Ignoring the real spatial resolution of Gaia

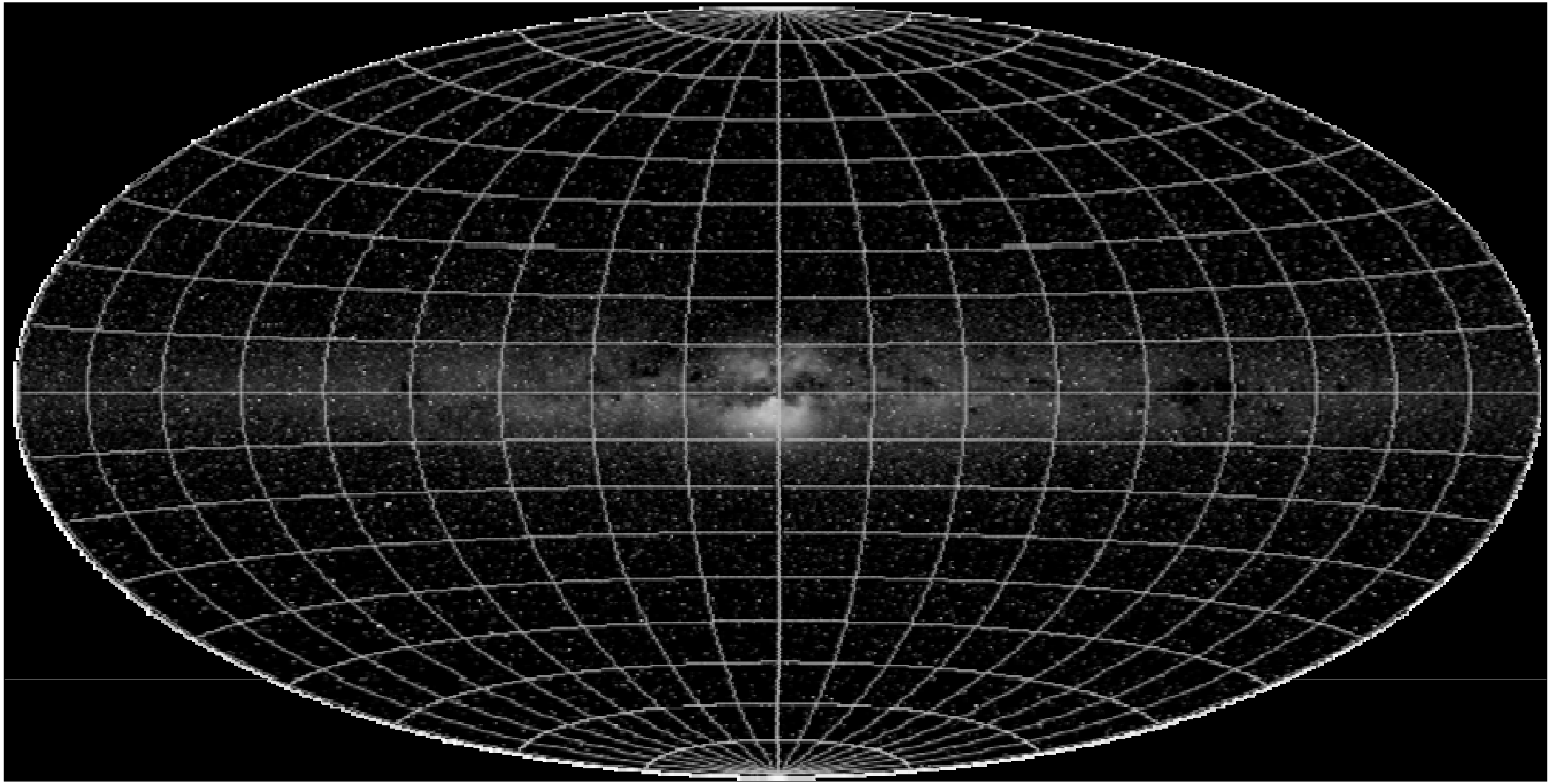
Real sky

https://www.cfa.harvard.edu/~rkirshner/MilkyWay.jpg



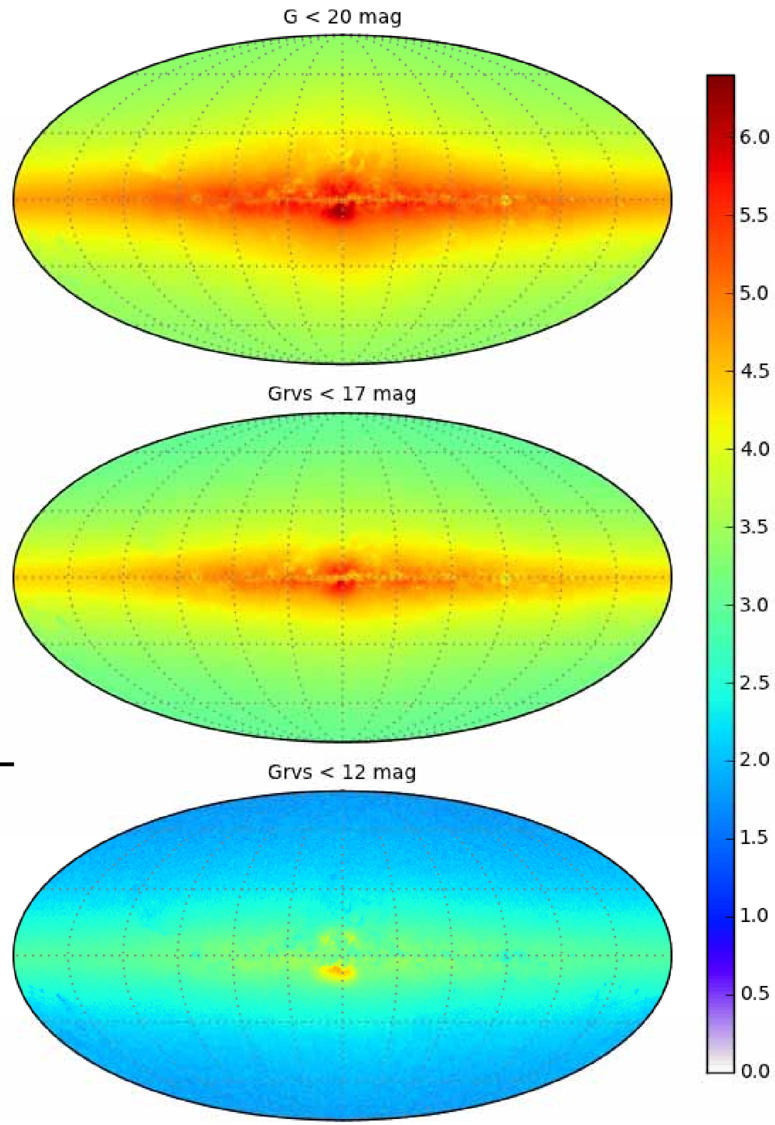
Simulated sky

The sky in G magnitude



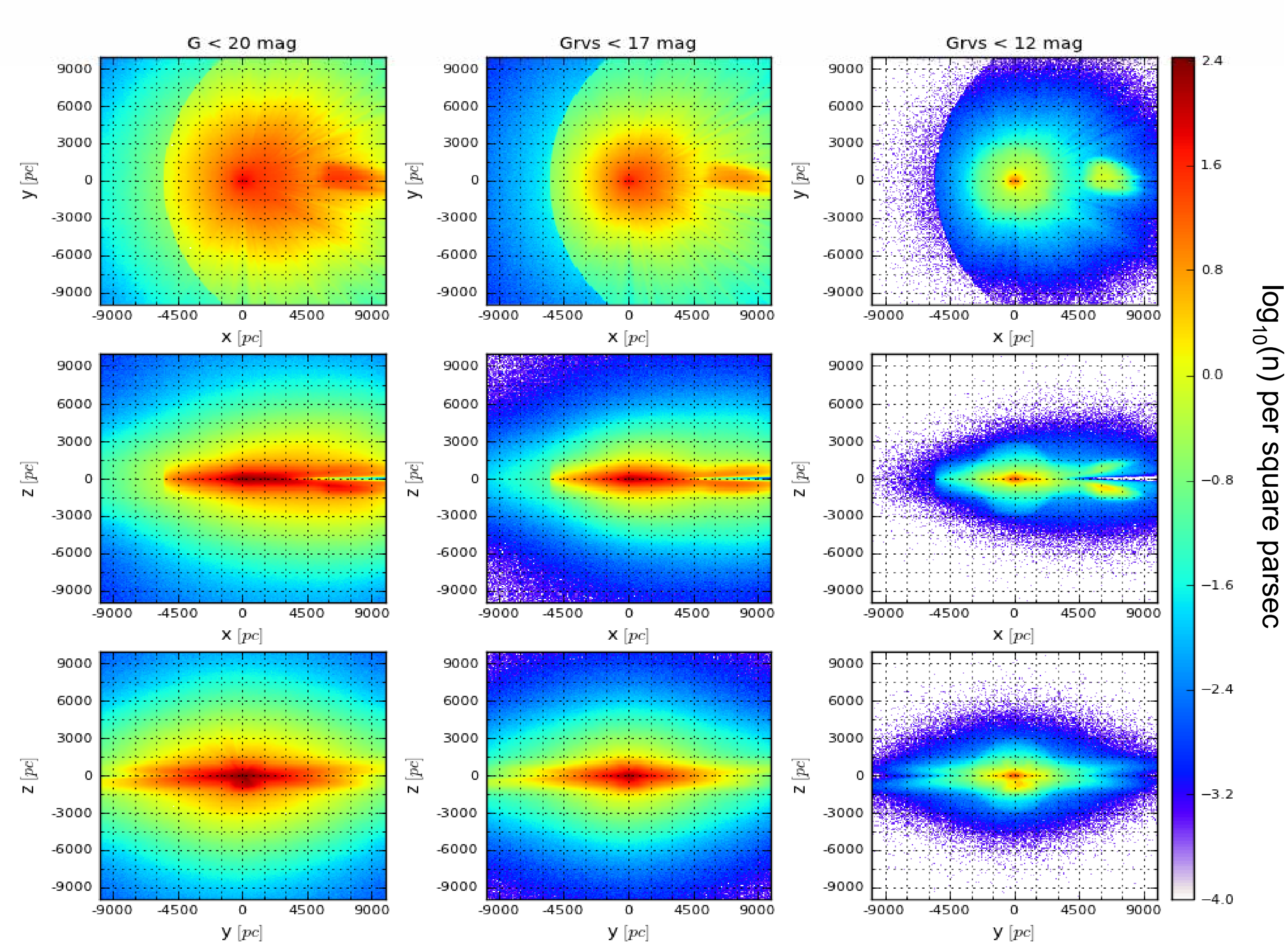
Radial velocities

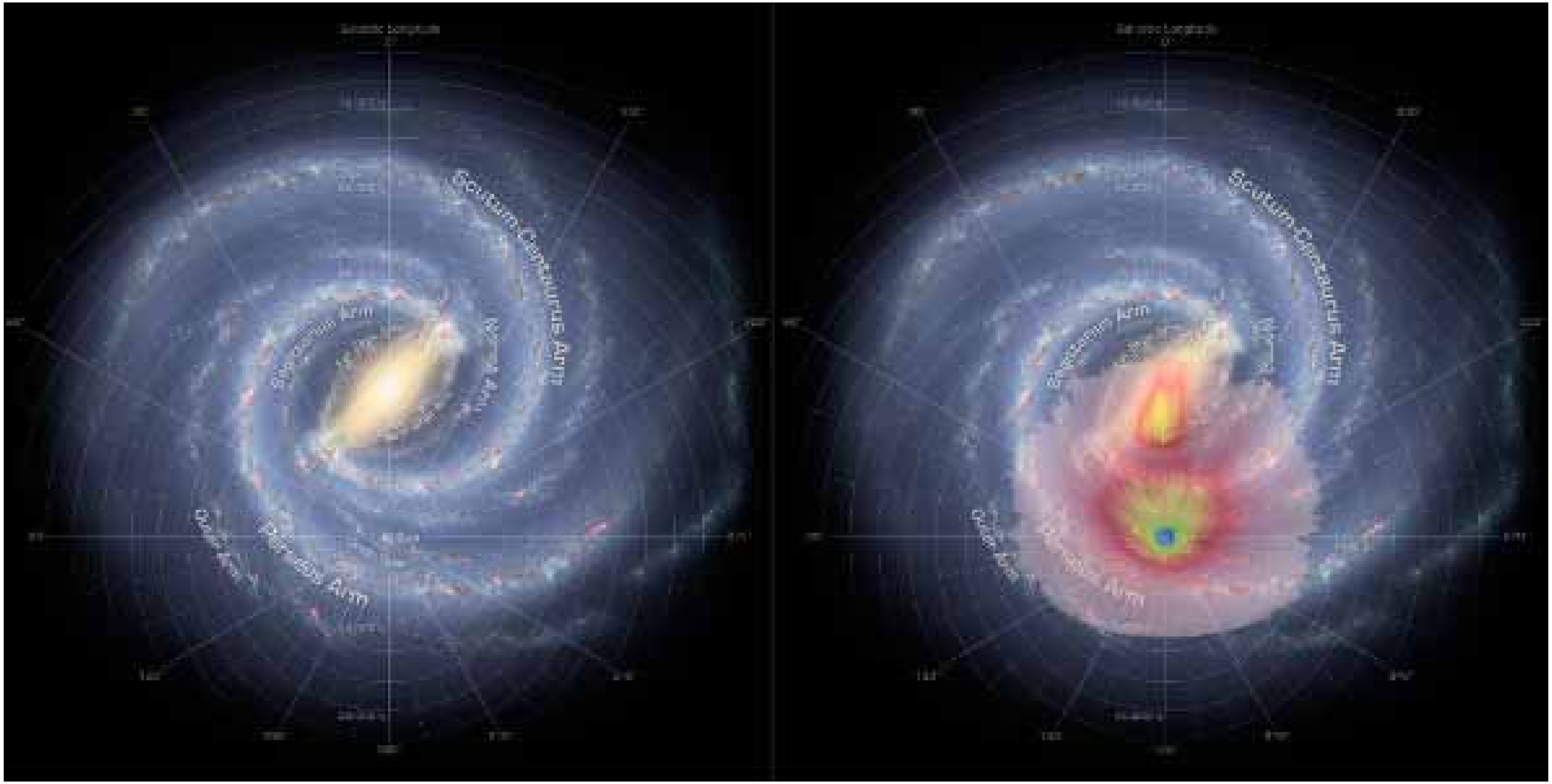
Metallicities



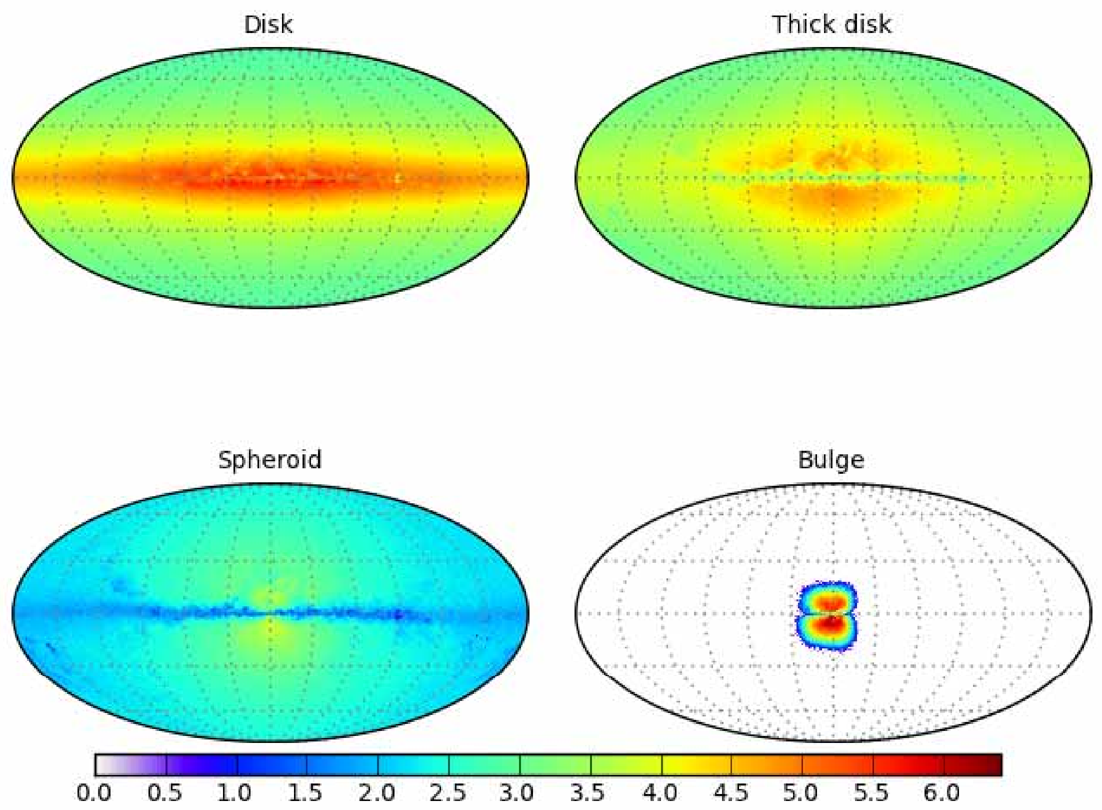
$\log_{10}(n)$ per square degree

Galactic 3D distribution

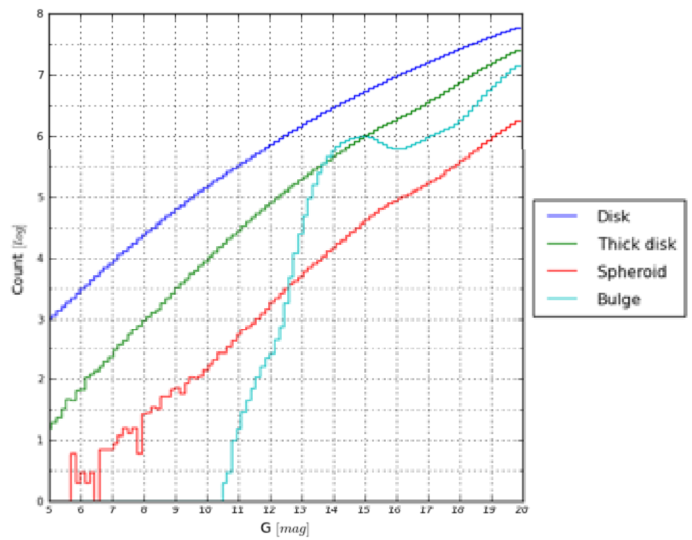



Scutum-Centaurus Arm
Scutum-Centaurus Arm

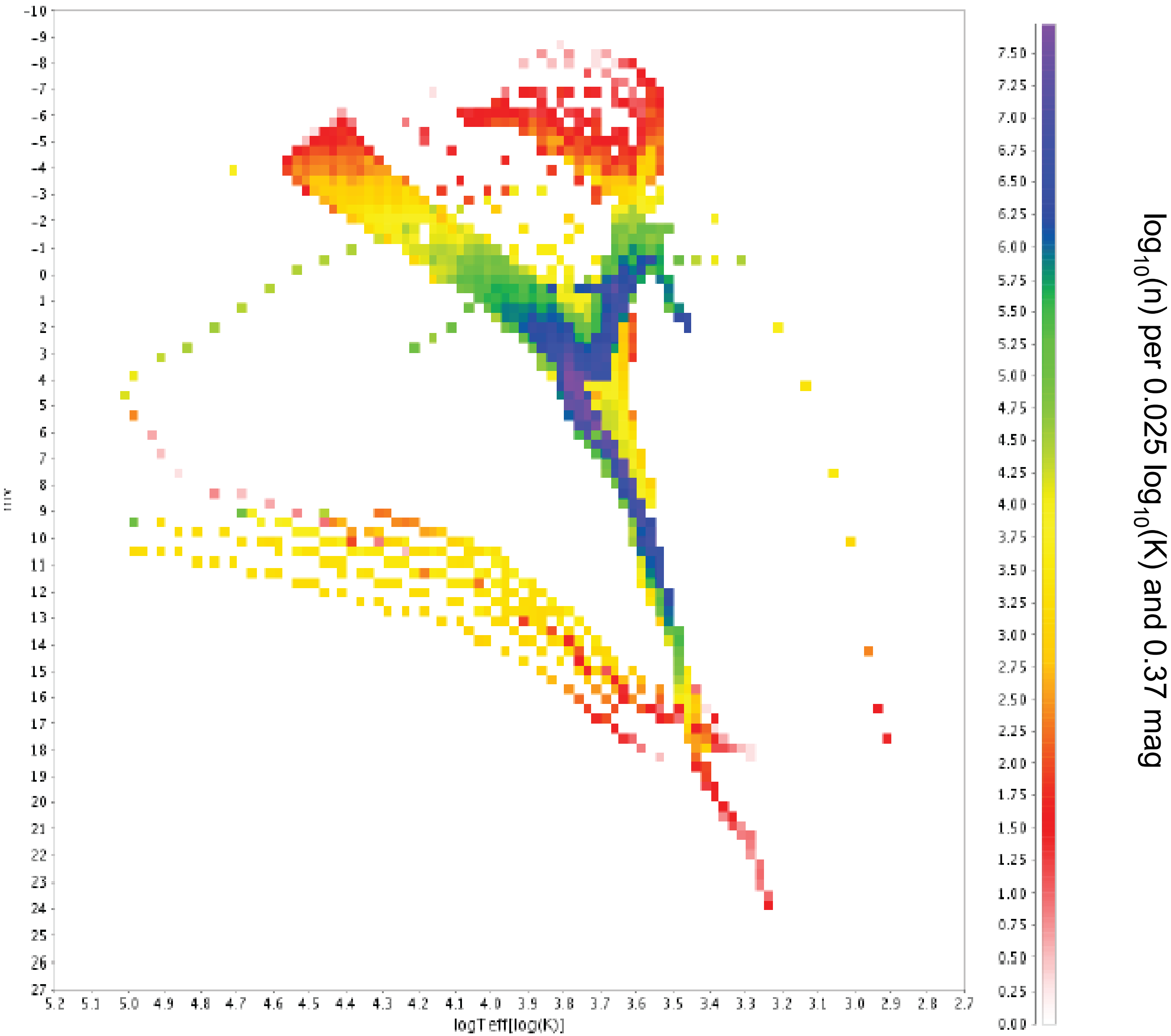
Populations



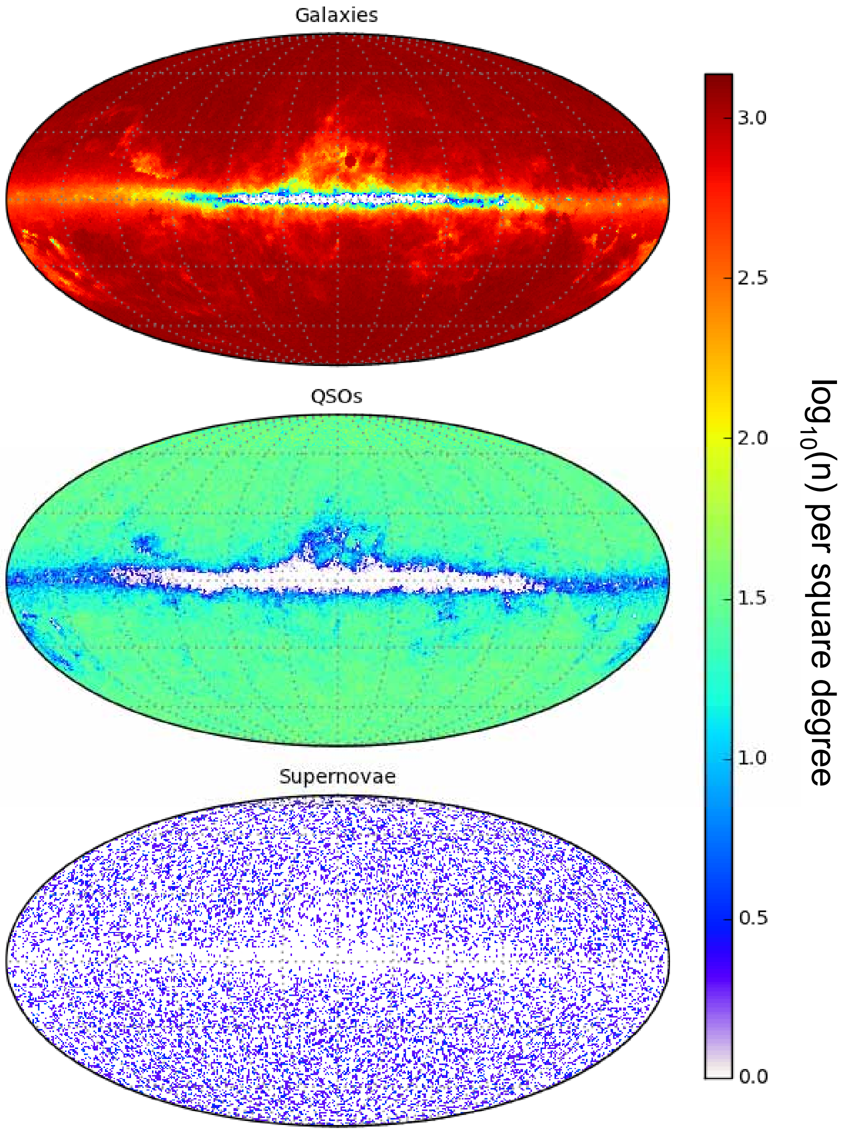
$\log_{10}(n)$ per square degree



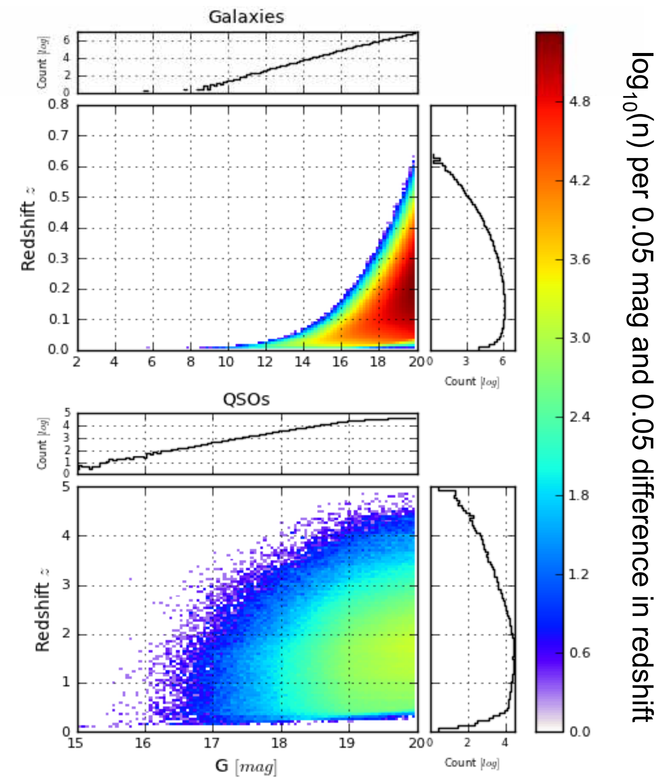
HR diagram


logTeff[log(K)]
log10(n) per 0.025 log10(K) and 0.37 mag
5.2 5.1 5.0 4.9 4.8 4.7 4.6 4.5 4.4 4.3 4.2 4.1 4.0 3.9 3.8 3.7 3.6 3.5 3.4 3.3 3.2 3.1 3.0 2.9 2.8 2.7
-10 -9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
0.00 0.25 0.50 0.75 1.00 1.25 1.50 1.75 2.00 2.25 2.50 2.75 3.00 3.25 3.50 3.75 4.00 4.25 4.50 4.75 5.00 5.25 5.50 5.75 6.00 6.25 6.50 6.75 7.00 7.25 7.50

Extragalactic objects



Stars	G < 20 mag	Grvs < 17 mag	Grvs < 12 mag
Stars in LMC	7,550,000	1,039,000	5,600
Stars in SMC	1,250,000	161,000	950
Unresolved galaxies	38,000,000	3,000,000	4,320
QSO	1,000,000	5,200	11
Supernovae	50,000	-	-



Thank you!