

# Global astrometric sphere reconstruction in Gaia: challenges and first results of the Verification Unit

Alberto Vecchiato<sup>1</sup>, Ummi Abbas<sup>1</sup>, Beatrice Bucciarelli<sup>1</sup>, Mario G. Lattanzi<sup>1</sup>, Roberto Morbidelli<sup>1</sup>

<sup>1</sup>Osservatorio Astronomico di Torino,  
via Osservatorio 20, 10025 Pino Torinese (TO), Italy  
email: [vecchiato@oato.inaf.it](mailto:vecchiato@oato.inaf.it)

**Abstract.** Gaia will estimate the astrometric and physical data of approximately one billion objects. The core of this process, the global sphere reconstruction, is represented by the reduction of a subset of these objects, which will constitute the largest and most precise catalog of absolute astrometry in the history of Astronomy, and will put General Relativity to test by estimating the PPN parameter  $\gamma$  with unprecedented accuracy. As the Hipparcos mission showed, and as it is natural for all kind of absolute measurements, possible errors in the data reduction can hardly be identified at the end of the processing, and can lead to systematic errors in all the works which will use these results. In order to avoid such kind of problems, a Verification Unit was established by the Gaia Data Processing and Analysis Consortium (DPAC). One of its jobs is to implement and perform an independent global sphere reconstruction, parallel to the baseline one, to compare the two results, and to report any significant difference.

**Keywords.** Astrometry, catalogs, relativity

---

## 1. Introduction

The Gaia astrometric catalog, with its  $\sim 10^9$  objects at  $\sim 10$  to  $\sim 100$  microarcsecond accuracy, will be the richest and most precise ever produced. It will be a milestone of paramount importance for several science topics, going from almost every subject of astrophysics and astronomy to fundamental physics (Turon *et al.*, 2005).

The most evident example comes from the possibility of having parallaxes at the 10% accuracy level at galactic distances, that would result in a complete revision of the cosmic distance ladder since its first step (Webb, 1999).

At the same time, the kind of measurements performed by Gaia belongs to the area of *absolute astrometry*, and the main result of the mission is a *catalog* realizing an astrometric reference frame. A reference frame, as commonly said, is the materialization of a reference system; therefore, it is very difficult to identify possible errors in the measurements or in the data reduction process that brings to the definition of the final catalog, given its nature.

Gaia inherits many ideas from its “parent” mission HIPPARCOS, such as the concept of getting absolute parallaxes from simultaneous observations of two different fields of view separated by a large angle, or that of a scanning law which makes it possible for the satellite instruments to observe the whole celestial sphere. The main goal of HIPPARCOS was also to produce a catalog of absolute positions. Therefore, HIPPARCOS has faced the same kind of problems as above, but at a much smaller scale because of its lower precision and the much smaller size compared to that of the future Gaia catalog. It is then worth learning from that mission in order to tackle these difficulties.

The data reduction process in HIPPARCOS was carried out by two consortia, FAST and NDAC, which operated independently on the same data. Their two results were then compared and appropriately merged in order to obtain the final catalog.

This ideal solution cannot be applied to the case of Gaia. Due to the size of the problem, the data reduction task is much demanding both in terms of the needed resources and manpower. To retain as much as possible the HIPPARCOS approach, without requiring excessive resources, the Gaia Data Processing and Analysis Consortium (DPAC) adopted a strategy that foresees an Astrometric Verification Unit (AVU) within a single data processing pipeline. Before detailing the function of this unit, a brief summary of the structure of the Gaia data reduction pipeline is needed.

## 2. Overview of the Gaia data processing design

The DPAC is organized in *Coordination Units* (CUs), each of which is in charge of some parts of the whole task. CU3 takes care of the so-called *core processing*, i.e. it will consider a subset of “well-behaved” stars (e.g. single stars, photometrically and astrometrically stable, not too faint, etc.) and will reconstruct very precisely their five astrometric parameters. The number of stars processed by CU3 will be approximately some tens of million, up to  $10^8$ , i.e.  $\lesssim 10\%$  of the size of the final Gaia catalog. These will constitute a network to which the position of the other objects will refer. The CU3 reconstruction of the astrometric celestial sphere differs from that involving the other sources because it is *global*, i.e. without any reference to other objects, at least until this global reference frame will be linked to the ICRF.†

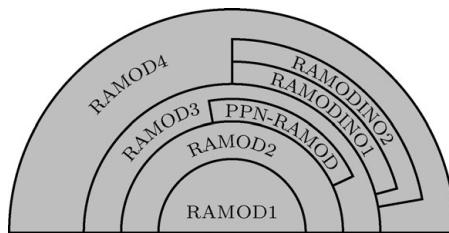
The Core Processing is a complex procedure that includes several steps. Since it is here that the “bootstrap” of the Gaia Reference Frame takes place, the DPAC has chosen some of its critical parts to be processed by two independent sub-systems in order to reproduce, for these specific steps, the structure of the two HIPPARCOS consortia. CU3 is designed in such a way that, for each of those steps, there will exist a verification counterpart operating independently from the main data reduction chain. All of these verification sub-systems have been gathered in the so-called AVU, each of them having the task to compare its results with those of its counterpart.

## 3. The sphere reconstruction at a glance

The main pipeline process which will reconstruct the global sphere is called Astrometric Global Iterative Solution (AGIS). In its bare bones, the sphere reconstruction consists in the solution, in the least-squares sense, of a large and sparse system of linearized equations.

Each equation corresponds to a Gaia observation whose known term contains the measurement, while the other ones are functions of the unknowns to be estimated. Since Gaia is a self-calibrating instrument, the function describing the satellite measurements does not depend just on the (relativistic) *astrometric* model, but also on the *attitude* and the instrument *calibration* parameters for the time of observation. Finally, since one of the expected by-products of the core processing is the estimation of the  $\gamma$  parameter of the Parametrized Post-Newtonian formulation, there is at least one unknown in a last set. Its elements appear in every equation of the system, and therefore these unknowns are called *global parameters*.

† We emphasize that the Gaia reference system will be materialized by all of the  $\sim 10^9$  objects of the Gaia catalog, since the stars not processed by the CU3 will be linked to the same reference frame.



**Figure 1.** RAMOD identifies a family of astrometric models with increasing accuracies. The attitude models belonging to the project are called RAMODINO1 and RAMODINO2. The present relativistic model implemented in GSR is an adaptation of PPN-RAMOD to the Gaia-type of measurement.

#### 4. The Gaia Sphere Reconstruction in AVU

AGIS has a “duplication” in one of the sub-systems of the AVU called Global Sphere Reconstruction (GSR). The input of both AGIS and GSR is a set of pre-processed data from the Gaia telemetry.

As said in section 2, AGIS will process the data for up to  $10^8$  well-behaved stars. It is presently foreseen that GSR will rather use a subset of up to 10 million of stars chosen from the AGIS dataset.

To keep the two reductions as independent as possible, GSR will differ from AGIS both from the point of view of the astrometric model and for the algorithm adopted for the sphere reconstruction, i.e. for solving the system of the linearized observation equations.

##### 4.1. GSR astrometric model

The astrometric model of AGIS is GREM (Gaia RELativistic Model) (Klioner, 2003), which is an extension of a seminal study (Klioner & Kopeikin, 1992) conducted in the framework of the post-Newtonian (pN) approximation of General Relativity. In GREM this model has been formulated according to a Parametrized Post Newtonian (PPN) scheme accurate to 1 micro-arcsecond.

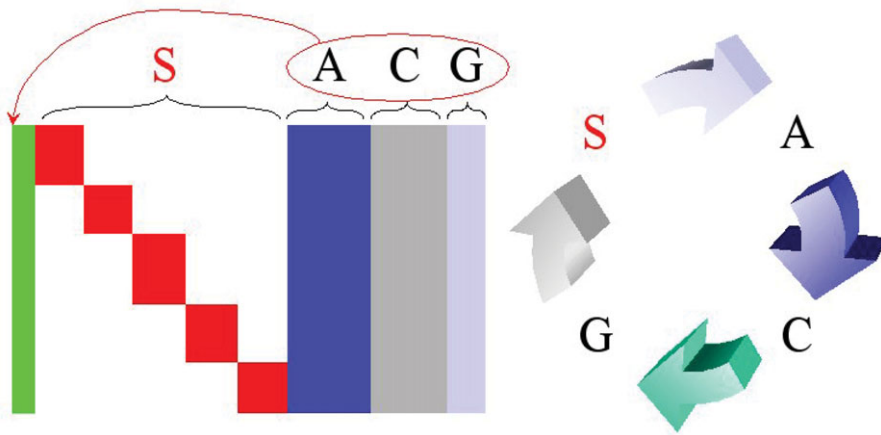
The astrometric model of GSR is taken from the RAMOD project, which identifies a family of astrometric models with increasing accuracies (see, e.g. Vecchiato *et al.*, 2003, de Felice *et al.*, 2006 and references therein) conceived to solve the inverse ray-tracing problem in a general relativistic framework and to use the tetrad formalism for the description of the observer’s reference system.

##### 4.2. GSR algorithm for the sphere reconstruction

AGIS takes its name after the method used for solving the system of equations, i.e. the *Global Iterative Solution*. The adopted strategy mainly consists in considering separately each type of parameter: astrometric, attitude, calibration, and global. When, e.g., the astrometric parameters are solved, all of the others are not computed and their present approximate values are used to calculate the known terms. Then the attitude parameters are solved, and the latest estimation for the astrometric ones is used for the known terms, and similarly for the calibration and the global parameters. A complete cycle over all of the parameter types is called *external iteration*, and the process is iterated until convergence is reached (Fig. 2).

This approach allows to easily parallelize the mathematical problem and is probably mandatory when the size of the system of equations to be solved is that of the AGIS.

GSR will use the well-known LSQR algorithm, instead (Paige & Saunders, 1982). LSQR is an iterative algorithm for solving sparse systems of linear equations based on



**Figure 2.** The right panel shows a complete external iteration for AGIS, where the astrometric (S), attitude (A), calibration (C) and global (G) parameters are solved one by one separately. The left panel represents the step of the solution when the astrometric parameters are solved; it shows that, putting the A, C and G parameters to the left-hand-side, which contains the known terms, the design matrix of the system can be arranged in a block-diagonal scheme. This clearly makes the algorithm parallelization easier. A similar arrangement can be used for the other parameters if the rows of the system are ordered by time.

a conjugate-gradients method. Since the procedure can be optimized for the memory requirements, all of the parameters can be estimated in a single iteration, and therefore, given also the smaller number of stars considered, there is no need to resort to the technique of the external iterations. A parallelized version of LSQR can be implemented to improve on the computing time.

#### 4.3. Algorithms for the comparison

The AVU-GSR sub-system is also in charge of comparing its sphere solution with that coming from the AGIS.

Up to now the comparison process foresees three different algorithms:

- (a) the  $\chi^2$ -test;
- (b) the Kolmogorov-Smirnov test;
- (c) the Infinite Overlapping Circle (IOC) test (Bucciarelli *et al.*, 1993).

Using different algorithms to compare the same sets of data is a further precaution to ensure a controlled and error-free solution to the fullest possible extent. If everything behaves according to the expectations, in fact, all of the algorithms will provide quite similar results. On the other hand, each of them is more sensitive to a particular kind of problem. The IOC method, e.g., is very effective at detecting residual regional errors, if any.

Finally, all of these algorithms will run on user-defined subsets of the parameter space(s). This will allow the sub-system to check and isolate differences possibly caused by a single or a limited set of bad observations.

## 5. Present status of AVU-GSR and future developments

The development path of AVU-GSR includes some successive stages which add refinements to each component of the sub-system. The present version, namely GSR1, is characterized by the first implementation (for RAMOD) of an abscissa-based astrometric model which includes the attitude. The astrometric model uses the PPN-Schwarzschild

metric of PPN-RAMOD (Vecchiato *et al.*, 2003), not including the deflection effects caused by the gravitational pull of the other bodies of the Solar System. Therefore, the main limitation of GSR1 is the need to reject approximately half of the observations, i.e. those too close to the giant planets and to the Earth. The accuracy, however, is enough for the purposes of this version, i.e. a first comparison of the two sphere solutions. Such version could also be used during the first period of the Gaia operational phase, when the overall precision of the sphere reconstruction is of  $\sim 100$  microarcsecond. The attitude model is that of Bini *et al.* (2003) adapted to the PPN-Schwarzschild metric.

Though GSR is able to treat the same input data of AGIS, a complete end-to-end data reduction is still not possible due to the lack of some pieces of software which are still being implemented. Thus, the system is presently running on 1-million-stars self-simulated datasets which allow us to exercise separately the LSQR-algorithm implementation.

The next stage will implement the instrument model (GSR2) while the third and last one will include a fully accurate RAMOD astrometric model (GSR3).

## 6. Conclusions

The absolute nature of the Gaia measurements calls for an HIPPARCOS-like cross-check of the results of the data reduction; however, the size of the problem produced by Gaia is too large and makes it impossible to provide two independent consortia, as in the Gaia predecessor.

The adopted solution is the duplication of some of the most critical parts of the data reduction, which are gathered in an independent Astrometric Verification Unit (AVU). One of the components of the AVU is GSR that replicates to a smaller scale the global sphere reconstruction (AGIS) produced by the main data processing chain. GSR shall provide the results of a comparison with the AGIS solution, obtained using the same datasets.

GSR1, i.e. the first version of GSR which is presently under completion, will exercise the comparison task. Its future versions (GSR2 and GSR3) will complete the sub-system with the implementation of the instrument model and of the final relativistic astrometric model, respectively.

## References

- Bini, D., Crosta, M., & de Felice, F. 2003 *Class. Quantum Grav.*, 20, 4695
- Bucciarelli, B., Taff, L. G., & Lattanzi, M. G. 1993 *J. Statist. Comput. Simul.*, 48, 29
- Klioner, S. A. 2003 *Astron. J.*, 125, 1580
- Klioner, S. A. & Kopeikin, S. M. 1992 *Astron. J.*, 104, 897
- Paige, C. & Saunders, M. A. 1982 *ACM Trans. Math. Software*, 8, 43
- Eds. Turon, C., O'Flaherty, K. S., & Perryman, M. A. C. 2005, *The Three-Dimensional Universe with Gaia*, *Publ. Astron. Soc. Pac.*, 120, 38
- Vecchiato, A., Lattanzi, M. G., Bucciarelli, B., Crosta, M., de Felice, F., & Gai, M. 2003 *Astron. Astrophys.*, 399, 337
- de Felice, F., Vecchiato, A., Crosta, M., Bucciarelli, B., & Lattanzi, M. G. 2006 *ApJ*, 653, 1552
- Webb, S., 1999, *Measuring the Universe*, Springer-Verlag