

Near-Earth-Object identification over apparitions using n -body ranging

Mikael Granvik and Karri Muinonen

Observatory, P.O. Box 14, 00014 University of Helsinki, Finland
email: mikael.granvik@helsinki.fi

Abstract. Under ideal conditions, Earth-based telescopes can observe near-Earth objects (NEOs) continuously from a few days to months during each apparition. Due to the usually complicated dynamics of the Sun-Earth-NEO triplet, the time interval between consecutive apparitions typically ranges from months to several years. On these time scales, exiguous single-apparition sets (SASs) of observations having short observational time-intervals lead to substantial orbital uncertainties. Linking of SASs over apparitions thus becomes a nontrivial task. For example, of a total of roughly 4,100 NEO observation sets, or orbits, currently known, some 2,300 are SASs, for which the observational time interval is less than 180 days. Either these SASs have not been observed at an apparition following the discovery apparition or the linkage of SASs has failed, an option which should preferably be eliminated. As a continuation to our work on the short-arc linking problem at the discovery moment (Granvik & Muinonen, 2005, *Icarus* 179, 109), we have investigated the possibility of using a similar method for linking exiguous SASs over apparitions. Assuming that the observational time-interval for SASs of NEOs is typically at least one day (minimum requirement set by the Minor Planet Center), the orbital-element probability-density function is constrained as compared to the typical short-arc case with an observational time interval of only a few tens of minutes. Because of the smaller orbital-element uncertainty, we can use the short-arc method (comparison in ephemeris space) for longer time spans, or even do the comparison directly in the orbital-element space (Cartesian, Keplerian, equinoctial, etc.), thus allowing us to assess the problem of linking SASs of NEOs. Due to possible close approaches with the Earth and other planets, and substantial propagation intervals, we have developed new n -body techniques for the orbit computation.

Keywords. Identification; statistical methods; data analysis

1. Introduction

An observation set of an asteroid is typically classified as belonging to a certain asteroid group based on a single point estimate of the orbital-element probability-density function (p.d.f.), which is usually obtained via either the least-squares solution assuming Gaussian statistics, or a single Väisälä solution based on human judgement. For short observation time intervals leading to wide, clearly non-Gaussian, and strongly nonlinear orbital-element p.d.f.'s, the use of point estimates for classification is prone to errors. When searching for near-Earth objects (NEOs), an observation set may, for instance, seem a probable main-belt object (MBO; which in this particular case is uninteresting), because the point estimate places it within the main asteroid belt at the observation date. However, in reality, the observation set may belong to a near-Earth object that only spends part of its time in the main asteroid belt. Currently, some 50,000 provisionally designated observation sets in the Minor Planet Center's (MPC) observation database span less than 48 hours. Most of these single-apparition sets (SASs) of observations have been observed for at least two nights, as the MPC guidelines require to get a provisional designation. The overwhelming majority of these so-called two-nighters are

currently classified as MBOs. For the NEO suspects in the two-nighter data, the linkage to other observation sets may have failed with the current methods. This could, for instance, happen for fast-moving objects on high-eccentricity orbits as these objects also spend a substantial fraction of their time in the main asteroid belt around aphelion, that is, masquerading themselves as MBOs.

For the current long-term linking problem, we will use the same scheme that has successfully been applied by Granvik & Muinonen (2005) in the short-term case. However, in the long-term NEO case, some of the assumptions are no longer valid. The main obstacle is the two-body approximation, which can be acceptable even for NEOs in the short-term case assuming current orbital uncertainties, but not in the long-term case. Even though the differences between the two-body and the n -body p.d.f.'s are negligible around the inversion epoch (assuming, in turn, that the inversion epoch is close to the observation dates), the nonlinear propagation of the p.d.f.'s to a comparison epoch, often several years from the inversion epoch, will generally lead to notable differences between the two cases. The methods that are used to solve the two-point boundary-value problem (the problem of solving an orbit from two heliocentric Cartesian positions) in Ranging Virtanen, Muinonen & Bowell (2001) have so far been either the p -iteration method or the continued-fraction method for details, see, e.g., Danby (1992), both of which are based on the two-body approximation. Until now, the use of two-body solutions in Ranging has been acceptable because, for example, the observational data sets have spanned time intervals short enough for perturbations to be negligible Virtanen & Muinonen (2006) e.g.,. However, n -body perturbations have been accounted for in the propagations for the computation of collision probabilities.

Another assumption used in the short-term case, but proven problematic in the long-term case, is that finding a common orbital solution for two observation sets separated in time is straightforward and can be done efficiently with, for instance, Ranging. For the long-term case, it is clear that the least-squares method Danby (1992) see, e.g., is both a valid and an optimal tool for the verification of trial linkages. But finding an initial orbit for the method is cumbersome; one could think of using sample orbits computed with, e.g., Ranging for the separate observation sets, but it turns out that these orbits are often too far off at either end to allow the least-squares method to converge, even if the linkage would be correct. Another possibility would be to use a semi-analytical method like the one by Kristensen (1995) to find an initial orbit by using observations from both sets simultaneously. However, the semi-analytical methods are problematic because they, again, use the two-body approximation which is not valid over long time intervals.

New n -body methods are thus needed both for the solution of the two-point boundary-value problem, and for the generation of initial orbits to be used by the least-squares method. Our aim is to (i) present new methods for these two specific tasks, (ii) present the overall structure of a new sampling method which can link exiguous observation sets over several apparitions, and (iii) apply the linking method to simulated data to prove its feasibility. The paper is organized as follows. In Sect. 2, we present the methods and techniques, while Sect. 3 explains the generation of the simulated data. We present and discuss the results in Sect. 4, and finally give our conclusions in Sect. 5.

2. Methods

The current long-term linking method uses two filters; the first one is used to find a substantially reduced set of trial linkages (as compared to all possible trial linkages) worth to be analyzed in detail. The second filter tries to find an orbital solution which ties two separate observation sets together assuming realistic observation uncertainties.

If successful, the second step thus implies a linkage between the two sets. Before getting into details of the linking method, we need to go back to the basics of orbit computation, and develop a robust method that solves the two-point boundary-value problem using an n -body dynamical model.

2.1. Robust n -body solution to the two-point boundary-value problem

Starting with two Cartesian heliocentric positions $\mathbf{r}_0(t_0)$ and $\mathbf{r}_1(t_1)$ (where t_0 and t_1 indicate the corresponding epochs), the solution to the two-point boundary-value problem is a velocity $\mathbf{v}_0(t_0)$ such that once an orbit $\mathbf{P}(t_0) = (\mathbf{r}_0(t_0), \mathbf{v}_0(t_0))$ at epoch t_0 is propagated to the epoch t_1 , the distance

$$d_{\text{Car}} = |\mathbf{r}_1(t_1) - \text{pos}[\mathbf{P}(t_1)]| < \epsilon_{\text{acc}}, \quad (2.1)$$

where ϵ_{acc} is an adjustable parameter that defines the accuracy of the solution and the operator $\text{pos}[\mathbf{P}]$ gives the Cartesian position vector of orbit \mathbf{P} . A suitable fixed value for ϵ_{acc} could be, for instance, 10 m or 100 m. Optimally, ϵ_{acc} should be adjusted according to both the observation accuracy and the topocentric distance of the observed object. Several solutions to the inverse problem can be found in the literature (see, e.g., Danby (1992)) but we are not aware of any that would not require the validity of the two-body approximation. Note, however, that a problem resembling the current two-point boundary-value problem has been solved for collision orbits by Muinonen (1999) and Muinonen, Virtanen & Bowell (2001). In essence, the two-point boundary-value problem is an optimization problem where the three components of the velocity vector $\mathbf{v}_0(t_0)$ are free parameters, and the distance d_{Car} has to be minimized. By assuming that the difference between the two-body solution and the n -body solution is fairly small, we initialize a simplex optimization routine (for a description of a simplex method, see, e.g., Press, Teukolsky, Vetterling, *et al.* (1992), Press, Teukolsky, Vetterling, *et al.* (1999)) with the two-body solution. The simplex routine then uses the full n -body approach to make a correction to the initial two-body solution. On one hand, the n -body solution will be found substantially slower for close-approaching orbits as compared to the two-body solution. On the other hand, only one n -body propagation is needed if the object is far from perturbers, or over very short time intervals, because the n -body correction turns out to be negligible.

2.2. Phase-space address comparison

In practice, the first filter requires that the p.d.f.'s of the orbital elements (or any quantities derived from them) computed from the two separate observation sets have to overlap at one or more epochs. see Granvik & Muinonen (2005). For exiguous observation sets, a rigorous sampling of the orbital-element p.d.f. is critical to be able to link observation sets over long time intervals. For the inversion of the SASs, we use either stepwise Ranging Granvik & Muinonen (2005), or the Volume-of-Variation method (VoV; Muinonen, Virtanen, Granvik *et al.* (2006)), depending on the number of observations and the length of the observational time interval. Cartesian orbital elements, and a complete n -body approach, including the n -body solution of the two-point boundary-value problem, is used during the inversions. The inversion epoch of each observation set is the midnight (TT) closest to the mid-date of the observations.

All orbital-element p.d.f.'s are then propagated to the comparison epoch by using a full n -body dynamical model. We tried several different comparison variables (Cartesian orbital elements, Keplerian orbital elements, equinoctial orbital elements, Poincaré variables, the angular momentum vector, heliocentric spherical coordinates, etc.) and a few of their combinations. The number of comparison epochs was also altered when spherical

Table 1. Current phase-space discretization parameters for the address comparison algorithm. The parameters and their bin sizes will be further optimized in the future.

	a [AU]	e	i [°]	Ω [°]	ω [°]
Lower limit	0.0	0.0	0.0	0.0	0.0
Upper limit	500.0	1.0	180.0	360.0	360.0
Bin size	0.1	0.05	0.2	5.0	5.0

coordinates were tested. Since most of these choices performed more or less equally well, we decided to use Keplerian orbital elements at one comparison epoch and without the mean anomaly for the time being. We chose to use a comparison epoch after the last observations, which, in the future, will allow us to efficiently compare “new” discoveries with old SASs. However, the comparison epoch can easily be changed were there more optimal choices found later.

In practice, we do not compare smooth volumes in the phase space, but use a large amount, say 5,000, discrete orbits that sample the true orbital-element p.d.f.. For the comparison phase, the five-dimensional comparison vector is squeezed into a single scalar, which permits fast comparison of different orbits (for details, see Granvik & Muinonen (2005), Muinonen, Virtanen, Granvik *et al.* (2005)). Values that we currently use for the discretization of the phase space are shown in Table 1. Once identical scalars (integers) are found for orbital-element p.d.f.’s originating from two separate observation sets, we conclude that the observation sets give rise to similar orbits thus implying a potential linkage. The maximum deviation of the two orbits in comparison space is explicitly given by the bin sizes. The next step is to try to find a single orbital solution tying both of the corresponding observation sets together assuming realistic observational uncertainties.

2.3. Initial orbits through optimization

As stated earlier, the optimal way to scrutinize linkages between SASs in the second filter is to use the least-squares method. The reason why the least-squares method is the best choice is also its weak point. While the inversion of exiguous SASs of observations typically produces wide orbital-element p.d.f.’s (Ranging or VoV applicable), the inversion of a combination of two SASs separated by several years produces very constrained p.d.f.’s (least squares applicable). Typically, the difference in the orbital uncertainties are several orders of magnitude. It is thus clear that using a sample orbit from either one of the SASs as an initial orbit for the least-squares method will usually fail to converge, simply because the initial orbit is too far from the final solution and the linearity assumption does not apply.

The orbits computed for the SASs naturally provide a good first approximation and should hence be used. The apparent thing to do is to find the two orbits corresponding to the two SASs that are closest to one another in the orbital-element space at some common epoch, here, naturally, the comparison epoch. After tests on different orbital elements, we decided to use the Keplerian ones. We then tested several metrics, all of which are based on the work by Southworth & Hawkins (1963), and found the (a,e,i) -version used by Nesvorný, Bottke, Dones, *et al.* (2002) to be adequate for our use:

$$d_{Kep} = na\sqrt{C_a(\delta a/a)^2 + C_e(\delta e)^2 + C_i(\delta \sin i)^2}, \quad (2.2)$$

where na is the heliocentric velocity of an asteroid on a circular orbit with semimajor axis a , $\delta a = |a_1 - a_2|$, $\delta e = |e_1 - e_2|$, $\delta \sin i = |\sin i_1 - \sin i_2|$. For the constants C_a , C_e , and C_i we use 1.25, 2, and 2, respectively. Note that Nesvorný, Bottke, Dones, *et al.* (2002) used proper elements while trying to find fragments from a single break-up event, whereas we

are comparing osculating elements (at a common epoch) that have been derived from two different observation sets and (might) refer to the same object.

Even though the resulting orbits are superficially very similar, the least-squares method does not usually yet converge. Before improving the orbits with least squares, they thus need to be slightly optimized with a robust nonlinear method for which partial derivatives are not needed. To solve the optimization problem, we, again, use a simplex method. The six Cartesian orbital elements are now the free parameters, and the resulting χ^2 with respect to the observations is minimized. To initialize the simplex we use the three closest pairs of orbits between the sets, while the seventh orbital-element set is the arithmetic mean of the elements of the closest pair of orbits. The simplex naturally uses the full n -body model. When one of the seven orbits reaches some predefined χ^2 -value, it is used as the initial orbit for the least-squares method, which now converges (at least for correct linkages). The final acceptance for the linkage is given if the residuals of the least-squares fit conform to the assumed observational uncertainties. The verification of the correctness of an accepted linkage can only be made by additional, archive or new, observations.

3. Simulated observations

Simulated observations of NEOs were generated using the ASurv software (see, e.g., Granvik & Muinonen (2005)), which randomly draws uncorrelated orbital elements and absolute magnitudes from specified p.d.f.'s. Here we used NEO p.d.f.'s based on the work by Jedicke, Larsen & Spahr (2002) and an upper limit of $H = 18$ mag for the absolute magnitude. Positions and apparent magnitudes for these random objects are then computed for specified observation dates. If the position falls inside the observation window and the apparent magnitude is lower than a given threshold (here, $V_{\text{lim}} = 19$ mag), the observation is accepted. Here we allowed the target to be anywhere on the sky with the only limit being a minimum solar elongation of 45° . The dynamical model used in propagations between observation dates took into account perturbations induced by all planets as well as the dwarf planet Pluto. Finally, random Gaussian noise with standard deviation of $\sigma = 0.5''$ was added to the observations.

We used a cadence of two observations separated by one hour on two consecutive nights, which was repeated after a time interval of 2800 days, or roughly seven years and eight months. For a single object, we thus got a maximum of $2 \times 2 \times 2 = 8$ observations, which was split up into two two-nighters.

The previous parameters resulted in 211 two-night observation sets and 188 different objects. The maximum number of trial linkages between the observation sets is thus $211 \times (211 - 1) = 44310$, while the number of correct linkages is 23. The number of observation sets that can be linked assuming realistic observation uncertainties, is somewhere between these two extremes, usually closer to the lower number.

4. Results and Discussion

To show how the method performs for a correct linkage, we randomly chose the following two simulated geocentric observation sets which we know belong to the same object (designation, date, R.A., Dec., magnitude):

00107	2004 01 20.24925	16 27 20.254	-08 59 15.78	18.88V
00107	2004 01 20.28925	16 27 43.694	-09 00 01.92	18.89V
00107	2004 01 21.24925	16 36 58.287	-09 18 15.96	18.97V
00107	2004 01 21.28925	16 37 21.076	-09 19 00.34	18.97V
00111	2011 09 20.24924	03 45 55.689	-59 03 24.40	16.32V
00111	2011 09 20.28924	03 45 08.261	-59 02 59.67	16.32V
00111	2011 09 21.24924	03 26 44.342	-58 47 45.73	16.31V
00111	2011 09 21.28924	03 25 59.625	-58 46 55.75	16.31V

Orbital-element p.d.f.'s were then obtained by solving the inverse problem for both observation sets at local inversion epochs, that is, 2004/1/21.0 TT for 00107 and 2011/9/21.0 TT for 00111. The resulting marginal distributions are shown in Figs. 1 and 2. After

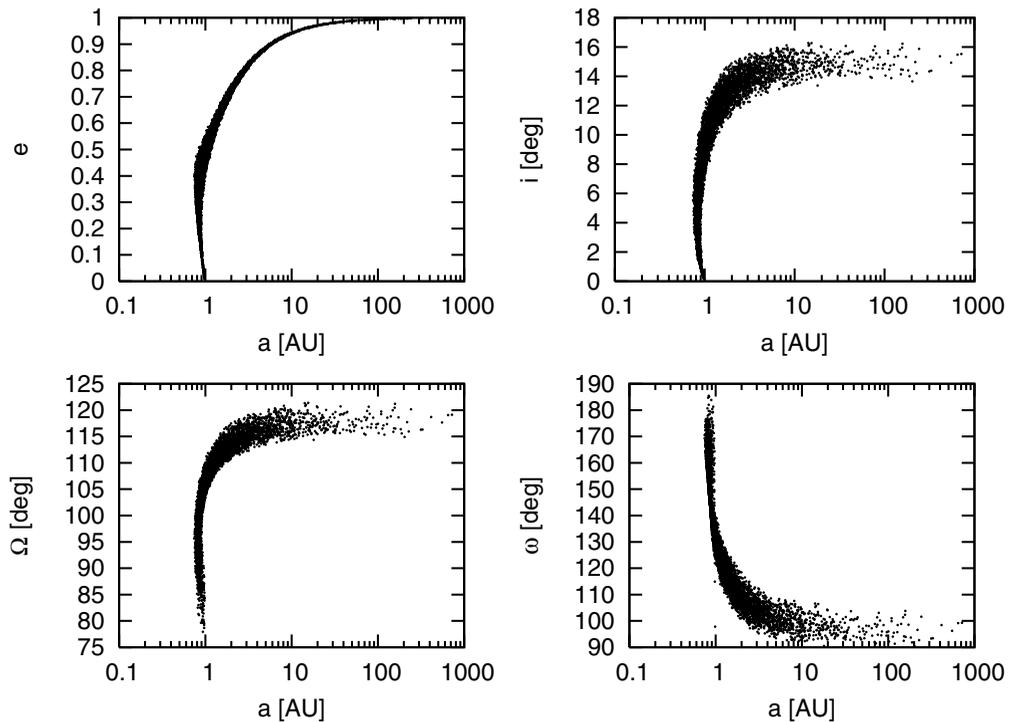


Figure 1. Selected marginal distributions of the orbital-element p.d.f. of 00107 at the inversion epoch 2004/1/21.0 TT.

the inversion, the p.d.f.'s were propagated to the comparison epoch 2012/1/1.0 TT (see Fig. 3). Recalling that the uncertainty in Cartesian orbital elements at the inversion epoch lies mainly along the line of sight (thus producing a long, narrow cone in the configuration space), the nonlinearity induced by the propagation to the comparison epoch is clearly visible in Fig. 4.

The comparison algorithm finds similar orbits between the sets, which correctly implies a possible linkage. A more elaborate inspection reveals astonishing similarity of the orbits considering the exiguous data and the substantial length of the propagation (see Table 2). In the current example, the least-squares method would actually converge, if the fourth orbit in Table 2 is fed into the least-squares routine. However, to make sure that the

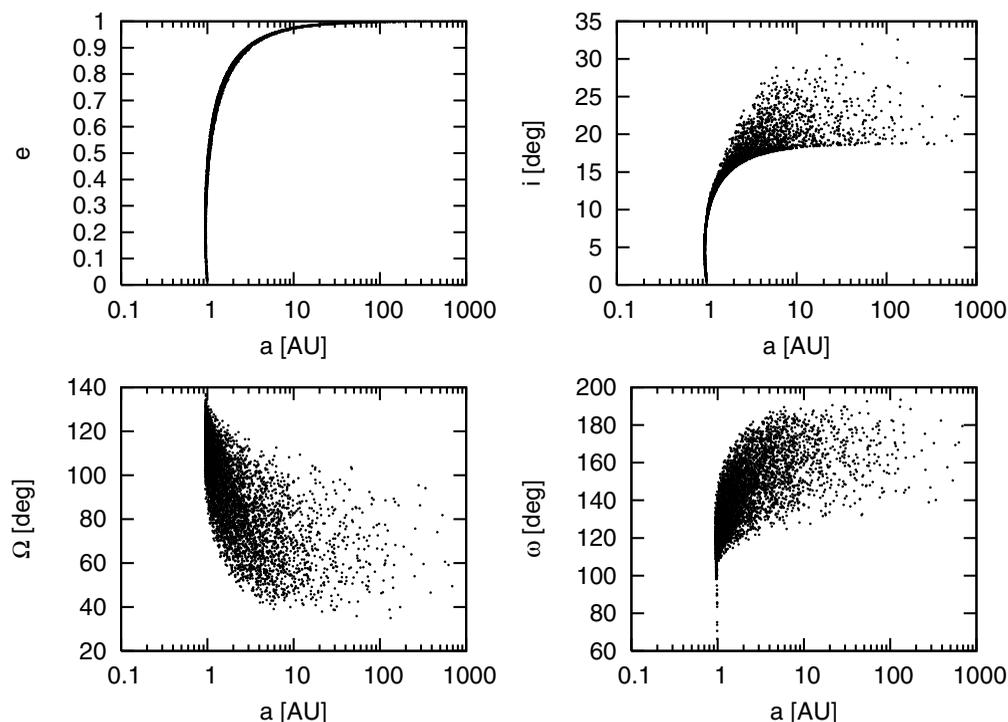


Figure 2. Selected marginal distributions of the orbital-element p.d.f. of 00111 at the inversion epoch 2011/9/21.0 TT.

Table 2. The three closest pairs of orbits at the comparison epoch. Note that there are only two different orbits shown for 00107.

SAS	a [AU]	e	i [°]	Ω [°]	ω [°]	M [°]
00107	1.17397	0.597785	12.5345	111.997	130.005	60.178
00111	1.17365	0.597630	12.5302	120.544	117.272	128.100
00107	1.03496	0.495752	9.5710	106.713	128.354	155.926
00111	1.03520	0.495986	9.5595	106.815	128.113	161.302
00107	1.03496	0.495752	9.5710	106.713	128.354	155.926
00111	1.03521	0.495296	9.5622	107.184	127.803	161.326

method converges for all correct linkages, the six orbital elements can also be further optimized with the simplex method described in Sect. 2.3. To induce as much stability as possible into the system, the inversion epoch is, again, the midnight (TT) closest to the mid-date of the observations, which in this particular case was 2007/11/21.0 TT. When the optimization reaches a certain level, the orbits are fed into the least-squares routine. The convergence is fast due to the optimized initial orbit, and the results (with a comparison to the ground truth) are given in Table 3. The orbital uncertainties for the combined SASs are several orders of magnitude lower than for the SASs separately (for example, for the semimajor axis roughly by a factor of 10^{-6} !). The small uncertainty of the resulting orbit makes verification by additional observations fairly easy.

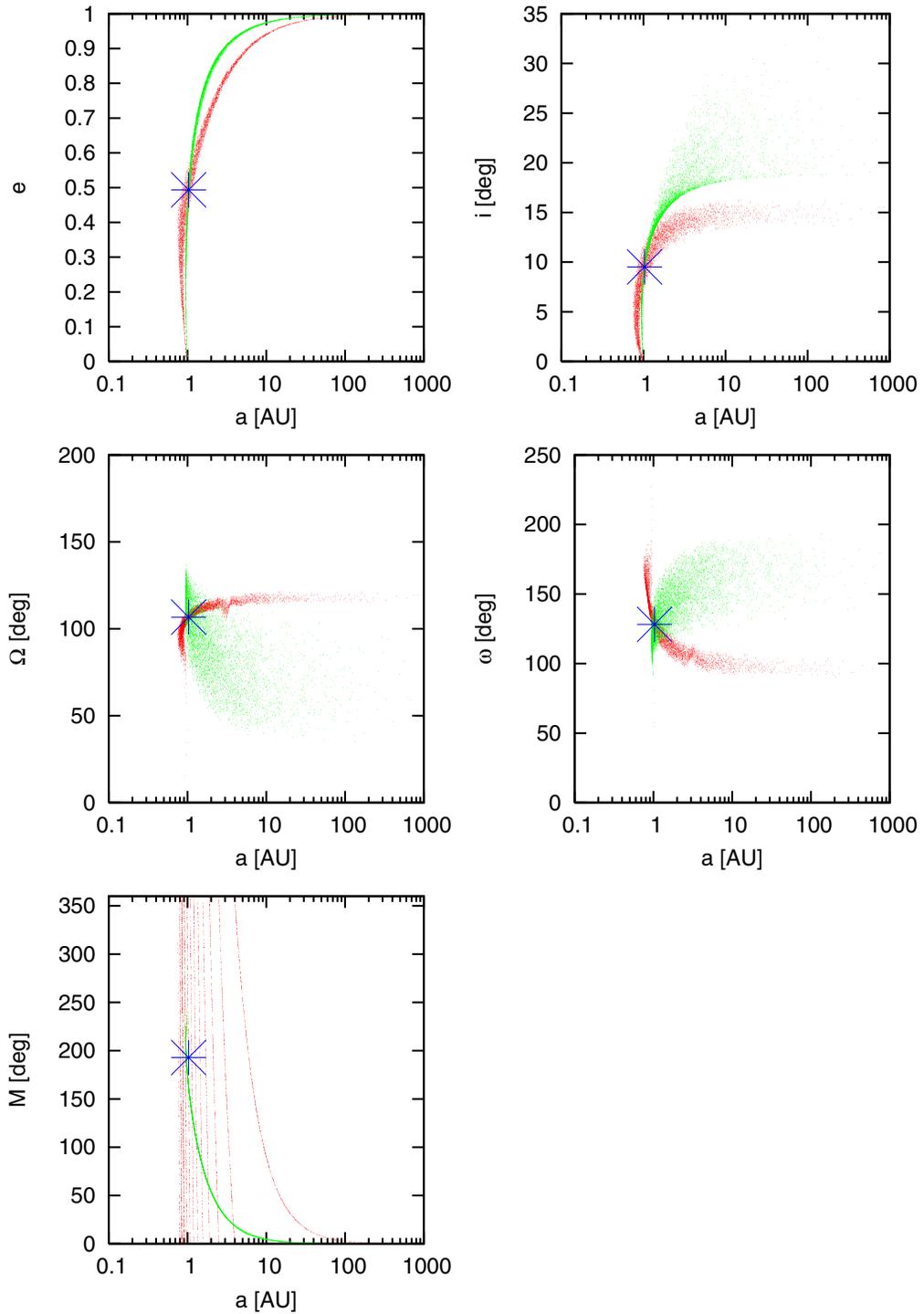


Figure 3. Selected marginal distributions of the Keplerian orbital-element p.d.f.'s of 00107 and 00111 at the comparison epoch 2012/1/1.0 TT. The star shows the ground-truth orbital elements based on which the observation sets were generated.

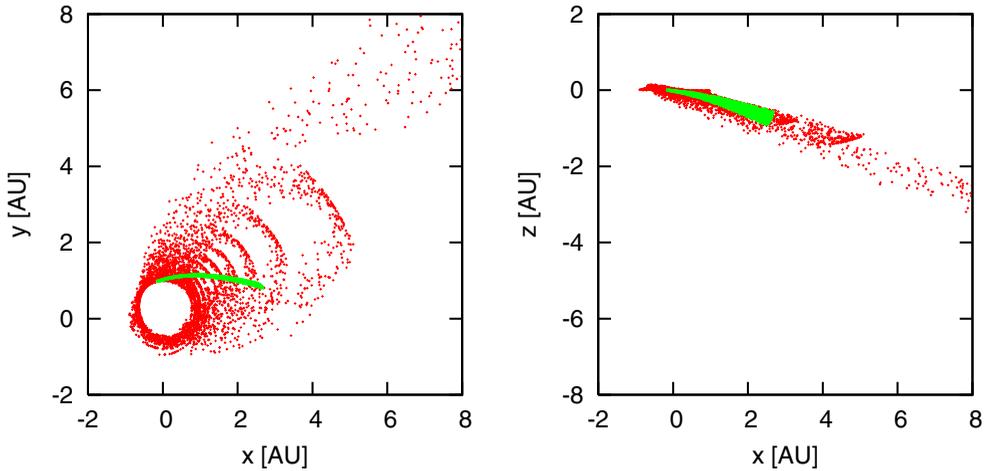


Figure 4. Selected marginal distributions of the Cartesian orbital-element p.d.f.'s of 00107 and 00111 at the comparison epoch 2012/1/1.0 TT. Note that the p.d.f.'s have been cut off at $x = y = 8$.

Table 3. The resulting orbital elements for the combined SASs accompanied with the uncertainties (the standard deviation) at the inversion epoch 2007/11/21.0 TT. The third line shows the ground-truth orbit from which the simulated observations were originally generated. The $\Delta\chi^2$ between the the ground-truth orbit and the least-squares solution (that is, the nominal orbit and the corresponding covariance matrix) is approximately 12.1, which, for six degrees of freedom, means that the solution is within the boundaries corresponding to 95.4% of the total probability mass (for a Gaussian distribution this equals to the $2\text{-}\sigma$ boundaries).

	a [AU]	e	i [°]	Ω [°]	ω [°]	M [°]
00107+00111 / elem.	1.03360694	0.4937476	9.51334	106.6827	128.2433	192.90980
00107+00111 / unc.	0.00000021	0.00000040	0.00047	0.0027	0.0030	0.00058
ground truth	1.03360674	0.4937444	9.51414	106.6857	128.2401	192.91025

5. Conclusions

The new methods presented in previous chapters have been successfully applied to case studies, one of which has been presented here. In the very near future we will apply the method simultaneously to the entire set of simulated observations to estimate the accuracy and efficiency of the linking method. In connection with the application of the linking method to larger populations, we will also investigate possibilities of further optimizations or even approximations to reduce the need for CPU time. Our goal, however, is to have a method which will find virtually all correct linkages having time intervals of at least, say, 10–15 years. In the future, we will also apply the method to all known SASs having observational time intervals less than 48 hours.

With the development of the new n -body techniques, we now have the means to completely eliminate the two-body approximation from the Ranging results. Besides allowing long-term linking, this will also refine future collision probability computations.

Acknowledgements

We thank the referee, Giovanni F. Gronchi, for the constructive criticism. Research supported, in part, by the Academy of Finland. MG is grateful for funding from the foundations of Alfred Kordelin and Magnus Ehrnrooth.

References

- Danby, J. M. A. 1992, *Fundamentals of Celestial Mechanics*, 2. edn, Willman-Bell, Inc., Richmond, Virginia, U.S.A.
- Granvik, M. & Muinonen, K. 2005, *Icarus* 179, 109
- Granvik, M., Muinonen, K., Virtanen, J., Delbó, M., Saba, L., De Sanctis, G., Morbidelli, R., Cellino, A., & Tedesco, E. 2005, in: Z. Knežević & A. Milani (eds.), *IAU Colloquium 197: Dynamics of Populations of Planetary Systems*, (Cambridge: Cambridge University Press), p. 231
- Jedicke, R., Larsen, J. & Spahr, T. 2002, in: W.F. Bottke, A. Cellino, P. Paolicchi & R. P. Binzel (eds.), *Asteroids III*, University of Arizona Press, p. 71
- Kristensen, L. K. 1995, *Astron. Nachr.* 316(4), 261
- Muinonen, K. 1999, in: B. A. Steves & A. E. Roy (eds.), *The Dynamics of Small Bodies in the Solar System*, Kluwer Academic, Dordrecht Norwell, MA, p. 127
- Muinonen, K., Virtanen, J. & Bowell, E. 2001, *CMDA* 81, 93
- Muinonen, K., Virtanen, J., Granvik, M., & Laakso, T. 2005, in: M. Perryman (ed.), *Three-Dimensional Universe with Gaia*, ESA Special Publications SP-576, Noordwijk, p. 223
- Muinonen, K., Virtanen, J., Granvik, M., & Laakso, T. 2006, *MNRAS* 368(2), 809
- Nesvorny, D., Bottke Jr., W. F., Dones, L., & Levison, H. F. 2002, *Nature* 417, 720
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, *Numerical Recipes in Fortran The Art of Scientific Computing*, 2. edn, Cambridge University Press
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1999, *Numerical Recipes in Fortran 90 The Art of Parallel Scientific Computing*, 2. edn, Cambridge University Press
- Southworth, R. B. & Hawkins, G. S. 1963, *Smiths. Contr. Astrophys.* 7, 261
- Virtanen, J., Muinonen, K. & Bowell, E. 2001, *Icarus* 154(2), 412
- Virtanen, J. & Muinonen, K. 2006, *Icarus* 184(2), 289