Astronomy & Astrophysics

Shape and spin distributions of asteroid populations from brightness variation estimates and large databases

H. Nortunen[1], M. Kaasalainen[1], J. Ďurech[2], H. Cibulková[2], V. Ali-Lagoa[3,4], and J. Hanuš[2]

[1] Tampere University of Technology, Department of Mathematics, PO Box 553, 33101 Tampere, Finland
e-mail: hari.nortunen@tut.fi
[2] Charles University, Faculty of Mathematics and Physics, Astronomical Institute, V Holešovičkách 2, 18000 Praha 8, Czech Republic
[3] Observatoire de la Côte d'Azur, Boulevard de l'Observatoire, CS 34229, 06304 Nice Cedex 4, France
[4] Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse 1, 85748 Garching, Germany

ABSTRACT

Context. Many databases on asteroid brightnesses (e.g. ALCDEF, WISE) are potential sources for extensive asteroid shape and spin modelling. Individual lightcurve inversion models require several apparitions and hundreds of data points per target. However, we can analyse the coarse shape and spin distributions over populations of at least thousands of targets even if there are only a few points and one apparition per asteroid. This is done by examining the distribution of the brightness variations observed within the chosen population.
Aims. Brightness variation has been proposed as a population-scale rather than individual-target observable in two studies so far. We aim to examine this approach rigorously to establish its theoretical validity, degree of ill-posedness, and practical applicability.
Methods. We model the observed brightness variation of a target population by considering its cumulative distribution function (CDF) caused by the joint distribution function of two fundamental shape and spin indicators. These are the shape elongation and the spin latitude of a simple ellipsoidal model. The main advantage of the model is that we can derive analytical basis functions that yield the observed CDF as a function of the shape and spin distribution. The inverse problem can be treated linearly. Even though the inaccuracy of the model is considerable, databases of thousands of targets should yield some information on the distribution. We employ numerical simulations to establish this and analyse photometric databases that provide sufficiently large numbers of data points for reliable brightness variation estimates.
Results. We establish the theoretical soundness and the typical accuracy limits of the approach both analytically and numerically. We propose a robust brightness variation observable η based on at least five brightness points per target. We also discuss the weaker reliability and information content of the case of only two points per object. Using simulations, we derive a practical estimate of the model distribution in the (shape, spin)-plane. We show that databases such as Wide-field Infrared Survey Explorer (WISE) yield coarse but robust estimates of this distribution, and as an example compare various asteroid families with each other.

Key words. methods: statistical – methods: numerical – techniques: photometric – minor planets, asteroids: general – methods: analytical

1. Introduction

Most of the current roughly one thousand asteroid shape and spin models, such as the ones given in the Database of Asteroid Models from Inversion Techniques (DAMIT[1]), are based on photometry (Kaasalainen & Lamberg 2006; Ďurech et al. 2015). Databases from large sky surveys and the inversion methods of sparse lightcurves (Kaasalainen 2004; Ďurech et al. 2009, 2016a) will greatly expand the list of models of individual asteroids. However, the databases also contain measurements that are not sufficient for individual models, but nevertheless can be expected to provide information on the statistical shape and spin distributions of the observed asteroid populations. Such measurements are, for example, brightness sequences ranging from a few points to full lightcurves. These can be transformed into statistical data by examining the population-level distribution of the brightness variation within each observed sequence.

The variation among each target's brightnesses, sampled over a wide range of rotation phases, can mostly be attributed to the shape elongation and the sub-Earth aspect angle of the object. The more detailed shape of the body, and especially its irregularity, is another important factor, but this cannot be included in a population-level model due to its complexity. The elongation and aspect have simply describable effects on the brightness variation, with monotonous dependencies. The detailed illumination and viewing geometry also have a somewhat complicated effect. Fortunately, if we include these factors as a part of the modelling error by simply using opposition geometry in the model, they will not make a large contribution to the total error budget as we will discuss below.

A realization of the statistical approach was presented by Szabo & Kiss (2008). They included over 10^4 pieces of pairwise brightness differences and, while their study did not contain analytical or numerical inspection of the generic inverse problem, they concluded that a statistical analysis is possible. The asteroid populations were characterized by shape elongation

---

[1] http://astro.troja.mff.cuni.cz/projects/asteroids3D/web.php

distributions. A similar type of observable and method was used by McNeill et al. (2016). We aim to establish the usefulness of the statistical approach by investigating the inverse problem both analytically and numerically by simulations, including the role of the insufficient model and other assumptions that do not necessarily hold in practice. We also seek to define a good observable of the brightness variation such that its information content is as high as possible.

We generate cumulative distribution functions (CDFs) of the brightness variation levels observed within large asteroid populations, aiming to study what the CDF reveals about the properties of the population. We choose the CDF since it is the most direct, well-defined, and stable data product describing the distribution statistics of a one-dimensional observable. Morever, the analytical study of the inverse problem requires the CDF integral of the model in the first place. To keep our model CDF simple and solvable, we choose to utilize as few parameters as possible, namely the shape elongation p and the spin β. We define p as the ratio of the equatorial widths of the asteroid, and β is the ecliptic polar angle of the spin axis. The main principle is to derive analytical basis functions that describe the contribution of the proportion of targets in a given β and p-bin to the observed CDF. These functions allow both the inspection of the information content of the data and the use of robust inversion methods. In particular, we show that it is possible to obtain information about the β distribution in addition to p.

In addition to the thorough analysis of the inverse problem, one of our main goals is to introduce an especially useful observable, η, that is a measure of the variation of the squared intensities of a sequence. The estimate η can be employed in a variety of contexts. Cibulková et al. (2016) used η to investigate brightness data that were not sufficient for sparse lightcurve inversion but suitable for creating a number of most probable simple asteroid models. These were used especially to demonstrate the slight periodic anisotropy of the distribution of rotation longitudes.

As the shape and spin distributions of asteroid populations are complex to interpret by themselves, we aim to introduce a tool for comparing the distributions of different populations. We do not make astronomical interpretations of the populations or their differences, but our objective is to show that the CDF-based method is a useful tool for the statistical investigation of populations.

We use asteroid databases that provide a number (usually at least five) of points for effectively one rotation: that is, obtained essentially randomly within a few nights such that the aspect angles of the Earth and the Sun are effectively constant. This allows the use of analytical basis functions for the CDFs as well as a rigorous study of the inverse problem. Other scenarios can be used as well, but these require additional assumptions and/or purely numerical treatment, further increasing the model noise.

In Sect. 2 we formulate the observables and the forward problem of the derivation of a CDF from the population model. In Sect. 3 we discuss the solution methods of the inverse problem and prove its fundamental uniqueness and stability properties. In order to verify the applicability of our method and obtain information on the level of error, we perform realistic simulations in Sect. 4 to assess the information content in practice. In Sect. 5 we use observations from databases to analyse asteroid families. We do not consider observational biases here: we simply take the available data at face value and analyse them as such. Additionally, we introduce a tool for a statistical comparison of distinct families. Conclusions are presented in Sect. 6, and mathematical details are given in Appendices A to C.

2. Observables and forward problem

In this section, we define our model and formulate the forward problem. We introduce our main observable, denoted as η, based on the variation of squared brightness intensities, and show why it is useful both analytically and by its information content. We also briefly consider the applicability of measurements with only two observed points. Mathematical details are given in Appendices A and B.

Our model shape is the triaxial ellipsoid, since it has a particularly simple analytical expression for the area of its projection in any given viewing direction (Connelly & Ostro 1984). In this paper, we use the terms brightness and projection area interchangeably, because they are physically almost the same (up to a scaling factor) for dark targets when the viewing and illumination directions coincide (Kaasalainen & Lamberg 2006). We further simplify the model (semiaxes a, b, c) with $b = c = 1$ and use $p := b/a$ for describing the shape elongation (the smaller the p, the more elongated the body). This is a coarse shape approximation for individual targets of general shape, but even if our model is actually not very realistic in practice, it should portray some coarse-scale population tendencies correctly when we have many observations. Indeed, as we will show by simulations, it suffices to have a model that represents the effects of shape elongation and spin direction in a roughly correct manner.

2.1. Amplitude A and its CDF C(A)

Let the polar aspect angle of the viewer be given by θ: $\cos\theta = \boldsymbol{v} \cdot \boldsymbol{e}$, where $\boldsymbol{v}$ is the spin direction (given by the polar coordinates (β, λ) in the inertial frame) and $\boldsymbol{e}$ the line of sight (unit vectors). Due to model symmetry, we only need to consider the interval $0 \le \theta \le \pi/2$. With ϕ for the longitudinal angle in a coordinate frame fixed to the ellipsoid, the area I of the ellipsoid's projection in the direction $\boldsymbol{e}$ is (Connelly & Ostro 1984)

$$I = \pi abc \sqrt{\frac{\sin^2\theta\cos^2\phi}{a^2} + \frac{\sin^2\theta\sin^2\phi}{b^2} + \frac{\cos^2\theta}{c^2}}.$$

In terms of our model definitions, the brightness L scaled against the maximal possible value πa is

$$L = \sqrt{p^2\sin^2\theta\cos^2\phi + \sin^2\theta\sin^2\phi + \cos^2\theta}$$
$$= \sqrt{1 + (p^2 - 1)\sin^2\theta\cos^2\phi}. \tag{1}$$

The statistical observable can be anything that describes the variation of the brightness as the target rotates (at a fixed θ). A simple version is the peak-to-peak amplitude; here we consider the ratio $A = L_{\min}/L_{\max} = L|_{\phi=0}/L|_{\phi=\pi/2}$ (i.e. an "inverse amplitude": the smaller the A, the larger the variation). Thus we have chosen the convenient $0 < p \le 1$ and $0 < A \le 1$ (rather than either of these extending to infinity). We would like to note that the amplitude A is based on intensity; we do not use magnitudes anywhere. The assumption is that all objects rotate about an axis (the ellipsoid's c-axis), which produces the observed projections random in ϕ. At first, we consider the randomness of θ to be due to the uniform distribution of rotation axis directions on the unit sphere S^2; later, we take the θ-distribution to be caused by a shifting viewing position.

The amplitude A is given by

$$A = \sqrt{\cos^2\theta + p^2\sin^2\theta} = \sqrt{1 + (p^2 - 1)\sin^2\theta}. \tag{2}$$

Using the amplitude, we can derive analytical basis functions, the linear combination of which yields the CDF $C(A)$ of a population with a given distribution of p (and β); see Appendix A for details.

In an approximation consistent with the coarseness of the model, it is practical to divide the population under study into a moderate number n of bins in each of which all members have the same p (and β). Then, if we have only p-bins and isotropic θ, the CDF of the values of A observed in the model population is

$$C(A) = \sum_{i=1}^{n} w_i \, F_i(A), \tag{3}$$

where the basis functions $F_i(A)$ are, from Eq. (A.2),

$$F_i(A) = \begin{cases} 0, & A \le p_i \\ \sqrt{\dfrac{A^2 - p_i^2}{1 - p_i^2}}, & A > p_i. \end{cases} \tag{4}$$

The range of the monotonously increasing F_i is $[0, 1]$, and $F_i = 1$ at $A = 1$ (Fig. 1). The occupation numbers of the bins are given by w_i.

Let us now include the β-distribution[2] by assuming that there is a concentration of viewing geometries towards the ecliptic plane (see Appendix A). If we assume a (p_i, β_j)-grid, $i = 1, \ldots, l$ and $j = 1, \ldots, m$, then we have $n = lm$ bins in the grid. We can write the CDF $C(A)$ as

$$C(A) = \sum_{ij} w_{ij} \, F_{ij}(A), \tag{5}$$

where, from Eq. (A.4), the monotonously increasing basis functions $F_{ij}(A)$ with the range $[0, \pi/2]$ are,

$$F_{ij}(A) = \begin{cases} 0, & A \le p_i \\ \dfrac{\pi}{2} - \arccos \dfrac{\sqrt{A^2 - p_i^2}}{\sin \beta_j \sqrt{1 - p_i^2}}, & p_i < A < \mathcal{F}(p_i, \beta_j) \\ \dfrac{\pi}{2}, & A \ge \mathcal{F}(p_i, \beta_j), \end{cases} \tag{6}$$

where $\mathcal{F}(p_i, \beta_j) = \sqrt{\sin^2 \beta_j + p_i^2 \cos^2 \beta_j}$. The $F_{ij}(A)$ are sigmoidal functions (Fig. 2), approaching the step function when $p_i \to 1$ (step at $A = 1$) or $\beta_j \to 0$ (step at $A = p_i$). Because of our choice of scale of p and A, parts of the F_{ij} tend to pack together at the low end of A, making them less well distinguishable than those with the slope in the higher end of A, but on the other hand, p-values less than 0.4 are not likely for real celestial bodies.

The occupation numbers w_{ij} are assigned to each bin. It should be noted that occupation levels proportional to $\sin \beta$ mean a uniform density on the direction sphere: that is, a constant $f(\beta)$. For applications, we adopt the convention of reporting the actual (relative) target numbers w_{ij} for a given β-slot (absorbing the factor $\sin \beta$), and we plot these as the density functions (DF; number densities in β rather than on the sphere) in the following sections.

---

[2] We note that our β is measured from the pole: $0 \le \beta \le \pi$. The ellipsoidal model, however, folds here all solutions of β into the interval $0 \le \beta \le \pi/2$; that is, the model cannot distinguish between pole latitudes above and below the ecliptic plane.
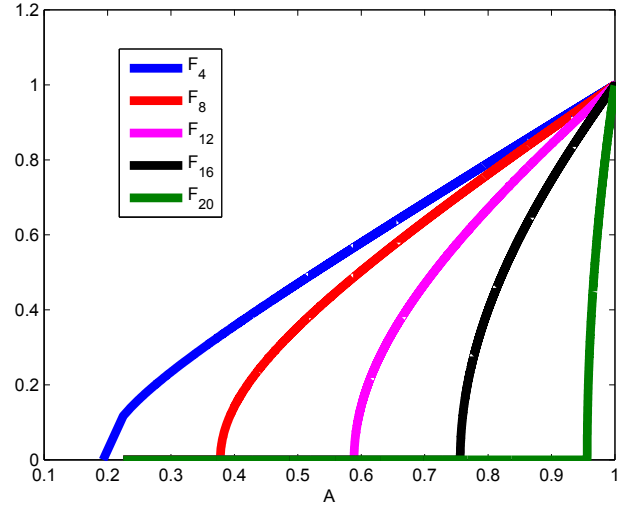

Fig. 1. Sample basis functions F_i on a set of bins p_i, where $i = 1, \ldots, 20$.
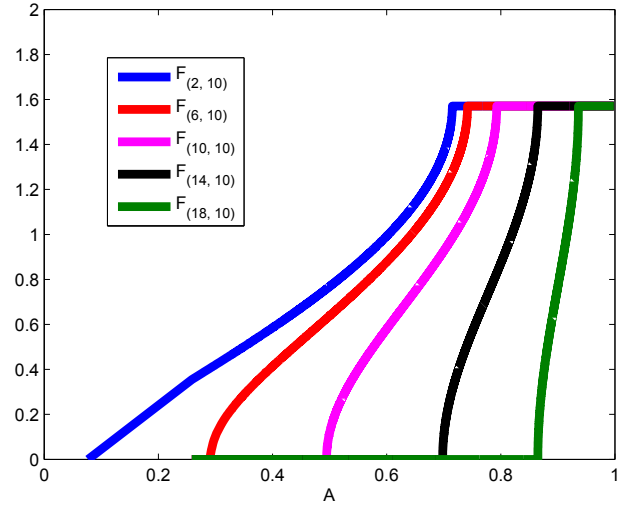

Fig. 2. Sample basis functions F_{ij} on a set of bins (p_i, β_j), where $i = 1, \ldots, 20$ and $j = 1, \ldots, 19$. The shape of the basis functions shows that they are linearly independent.

2.2. Brightness variation η

If the amplitude cannot be measured directly, a practical observable is the brightness variation around some mean value, requiring fewer points. Using intensity squared, L^2, to get rid of the square root in integrands, we obtain from Eq. (1) a simple average quantity over model rotation at a constant θ:

$$\langle L^2 \rangle = \frac{1}{2\pi} \int_0^{2\pi} \left(1 + \sin^2 \theta (p^2 - 1) \cos^2 \phi \right) d\phi$$
$$= 1 + \frac{1}{2} \sin^2 \theta (p^2 - 1).$$

Now, a measure of variation[3] for L^2 over a rotation is

$$\Delta(L^2) = \sqrt{\langle (L^2 - \langle L^2 \rangle)^2 \rangle} = \sqrt{\langle [\sin^2 \theta (p^2 - 1)(\cos^2 \phi - 1/2)]^2 \rangle}$$
$$= \sin^2 \theta (1 - p^2) \left[\frac{1}{2\pi} \int_0^{2\pi} (\cos^4 \phi - \cos^2 \phi) \, d\phi + \frac{1}{4} \right]^{1/2}$$
$$= \sin^2 \theta (1 - p^2)/\sqrt{8},$$

---

[3] Other definitions could be used as well, but this form leads to simple closed-form formulae.

and normalizing this with $\langle L^2 \rangle$ yields

$$\eta(\theta, p) := \Delta(L^2)/\langle L^2 \rangle = \sqrt{\left\langle \left(\frac{L^2}{\langle L^2 \rangle} - 1 \right)^2 \right\rangle}$$

$$= \frac{1}{2\sqrt{2}} \left[\frac{1}{\sin^2 \theta (1 - p^2)} - \frac{1}{2} \right]^{-1}. \tag{7}$$

We note that $0 \leq \eta \leq 1/\sqrt{2}$. Thus, by Eq. (2), our brightness variation η is directly related to the amplitude A:

$$\eta = \frac{1}{\sqrt{8}} \left(\frac{1}{1 - A^2} - \frac{1}{2} \right)^{-1}, \quad A = \sqrt{1 - \left(\frac{1}{\sqrt{8}\eta} + \frac{1}{2} \right)^{-1}}. \tag{8}$$

This is a particular advantage of the biaxial model: we can use all available estimates of A (available for dense lightcurves) and η together to form a $C(A)$. One can also directly compute a $C_\eta(\eta)$ with a procedure similar to that of Appendix A, resulting in similar types of integrals, but we choose the A-based formulation as it is more intuitive and leads to simpler equations. The end result is naturally the same in both cases. For the triaxial ellipsoid, a similar simple conversion between A and η is not possible since η would depend on θ (and b) in addition to A (Cibulková et al. 2016).

The condition $0 \leq \eta \leq 1/\sqrt{2}$ from Eq. (7) may be violated at some measurements of $L(\phi, \theta, p)$ when the parameter p is low ($\lesssim 0.4$). The maximal theoretical value of η for all lightcurve shapes (not just those from ellipsoids) approaches one, given by a boxcar-shaped lightcurve with half of the values at a constant level and half approaching zero. Real lightcurves have lower values of η because the lightcurve is smoother than the step-function type. If $\eta > 1/\sqrt{2}$ (this may happen due to an irregular shape, outliers, and/or particular spacing of the sample points), it follows that the amplitude A becomes purely imaginary according to Eq. (8). For computational purposes, we have omitted complex amplitudes in our study (these are rarely encountered).

2.3. Two-point brightness variation

The accuracy of the η estimate depends on the number of data points (and their coverage of the rotational phase) used to approximate $\Delta L^2 / \langle L^2 \rangle$. To analyse the information content of the minimal case of two points per rotation, we briefly consider simple pairwise brightness differences. For a group of N points for one target, the number of such values is $N(N-1)/2$, ordered such that the difference $0 < q \leq 1$ is $q = L_{\text{dimmer}}/L_{\text{brighter}}$. We do not need to have more than one such pair for one target, so one object does not have to cover the rotational phases well. This is the observable used in Szabo & Kiss (2008). We examine its properties from the inversion point of view in Appendix B. Since they turn out to be considerably inferior to those of η (above all, no information can be obtained on the distribution of β), we do not consider the two-point data further in the main text or database analysis.

McNeill et al. (2016) used a similar type of observable, with the two points connected by a short time interval (effectively yielding the slope of a lightcurve). This problem is even more complicated as it necessarily introduces the rotation period of the target into the forward model, with overlapping effects of p, spin, and period distributions (see Appendix B). Thus, in practice, this observable necessitates the heavy use of a priori assumptions in the inverse problem. Again, this is outside our aim of minimal use of parameters and prior functions, so we do not consider the slope version of two-point data further.

3. Inverse problem

In this section, we consider the fundamental properties of the inverse problem version of the forward model above before moving to realistic shapes and numerical results in the following sections. In particular, we present and prove a uniqueness result that shows how the distributions of both p and β can be uniquely obtained from the CDF $C(A)$. That is, we show why η-data contain unambiguous information on $f(p, \beta)$. This may seem counterintuitive at first glance, since the effects of p and β are certainly mixed for a single observation of η (i.e. a lightcurve). The point is that, under the ecliptic-plane assumption of Appendix A, the distribution of η in a large population separates the effects from each other.

For the p-only case of isotropic θ, the inverse problem can be cast linearly in matrix form. From Eq. (3), we write

$$Mw = C, \tag{9}$$

where $C \in \mathbb{R}^k$, $w \in \mathbb{R}^n$, $M_{ji} = F_i(A_j)$, and $F_i(A_j)$ are given by Eq. (4). The k observed values of A (derived from η) are sorted in ascending order, and the vector C contains the observed CDF: each element $C_j = j/k$ is the value of $C(A_j)$. In Appendix C, we discuss the analytical stability results of the distribution function $f(p)$ obtained from $C(A)$ (i.e. η-scatter data). In particular, we show that the inverse problem is not strongly ill-posed: the errors in the details of the observed CDF do not amplify fast in the error of the recovered $f(p)$.

In the inverse problem of full $f(p, \beta)$, we can write Eq. (5) in the form of Eq. (9) as well; now $C \in \mathbb{R}^k$, $w \in \mathbb{R}^n$ and M is a $k \times n$-matrix ($n = lm$),

$$M = \begin{pmatrix} F_{11}(A) & \dots & F_{1m}(A) & \dots & F_{l1}(A) & \dots & F_{lm}(A) \end{pmatrix},$$

and the occupation numbers w_{ij} are given in w with indexing similar to M above.

Uniqueness result. Perhaps surprisingly, the distribution of both p and β is unambiguously recoverable; that is, the bin model coefficients w_{ij} are uniquely determined by the $C(A)$. To show this, we notice that the pairs of end points (i.e. the values of A between which F_{ij} changes: A_- at $F_{ij} = 0$ and A_+ at $F_{ij} = \pi/2$), are unique for each F_{ij}. Any combination of F_{ij} will start to deviate from zero at the lowest A_- of the set, and stop changing at the highest A_+ of the set. Thus both end points of an F_{ij} cannot be matched by a superposition of other F_{rs}, so the F_{ij} are linearly independent (see Fig. 2 for illustration). Since the model $C(A)$ is a linear combination of the F_{ij}, the w_{ij} are unique for the observed $C(A)$. As w_{ij} are the occupation numbers of each (p_i, β_j) bin, this proves that the full p and β distribution is uniquely obtained for the ecliptic-orbit model.

If we want to use regularization to smooth the solutions for either p or β, we may apply, for example, the following $(n-1) \times n$ regularization matrix in the p-only case:

$$(R_p)_{ij} = \begin{cases} -1/(p_{i+1} - p_i), & i = j \\ 1/(p_{i+1} - p_i), & j = i + 1 \\ 0, & \text{elsewhere} \end{cases}$$

and its generalization for the (p, β)-grid, as well as similarly R_β with β. These approximate the gradients at each w_{ij} in the p- and β-directions only; one can construct more general matrices, but we found these to suffice for our problem. The occupation numbers can be obtained as a solution to an optimization problem:

$$\hat{w} = \arg\min_w \left(\|C - Mw\|^2 + \delta_p \|R_p w\|^2 + \delta_\beta \|R_\beta w\|^2 \right), \quad w \in \mathbb{R}_+^n. \tag{10}$$

To obtain the solution $\hat{w}$, we create an extended matrix $\tilde{M}$:

$$\tilde{M} = \begin{pmatrix} M \\ \sqrt{\delta_p} R_p \\ \sqrt{\delta_\beta} R_\beta \end{pmatrix}, \quad \tilde{C} = \begin{pmatrix} C \\ 0_{(l-1)m} \\ 0_{l(m-1)} \end{pmatrix}, \tag{11}$$

assuming a (p, β)-grid of the size $n = lm$ with, respectively, l and m equally spaced p- and β-values, and we find the least-squares solution of $\tilde{M}w = \tilde{C}$ with the constraint that each element of w be larger than or equal to zero. The extended vector $\tilde{C}$ due to the regularization ensures that there are always more equations than unknowns regardless of the number of original data points in C. Due to the instability of the problem, the direct unconstrained matrix solution would lead to negative values, but in for example the Matlab environment, the positivity constraint is simple to enforce with a standard function. We found that this is more practical than nonlinear optimization with, for example, $w_i = \exp(z_i)$.

We emphasize here that, despite the similar fitting procedures, CDFs are quite different from lightcurves as data. First of all, noise does not show as signal deviations because CDFs are monotone curves. The error in the observed CDF curve is essentially due to convolution (the distribution function is multiplied by the error probability function under the CDF integral), causing the smoothing of the curve (very noisy data would produce a featureless CDF resembling a step function). In our analysis, we do not attempt to deconvolve the original CDF, since the convolution (i.e. error) function is not known: in addition to the random and systematic brightness errors, it depends on the number and temporal distribution of the data points. Also, the addition of more measurements does not fill the gaps between points as in lightcurves: it alters the whole shape of the CDF. A visually good density of points in a CDF does not imply that its shape is near the correct theoretical one from infinitely many points. Computationally, one can use very high densities for interpolating between the actual CDF values to construct the values in the data vector. Then one does not have to use the full high number of observations, which may be helpful for software dealing with the positivity constraint of the solution based on Eq. (11).

Obviously p-values lower than about 0.4 start to become unrealistic, so one could also use an additional regularization function and a lower limit on p-values. We have, however, used the whole scale since the ostensibly unrealistic p-values are usually not heavily occupied and may carry information. For example, especially for smaller asteroids, the small p-values may also indicate irregular shapes (and the shadowing effects of nonzero solar phase angles) or an otherwise increased "noise level" due to systematic and modelling errors. In any case, the CDF method is meant to give a quick overview of a population instead of a detailed portrait, so trying to extract information via prior constraints is not a key concept here. An abundance of small p-values may also indicate that the data are simply not sufficient or otherwise suitable for a reliable result, so such a warning should not be suppressed by regularization.

4. Simulations

To assess the performance of our method with actual data and to check the effect of the simplified model, we perform several simulations. In Sect. 4.1, we explain the setup for our simulations and experiment with the synthetic data, giving graphical presentations of both actual and computed (p, β) solutions. In Sect. 4.2, we discuss ways to apply a post-solution "deconvolution" in order to correct the systematic errors in the solution; the corrected solutions are graphically presented as well.

4.1. Synthetic data

Since the modelling errors and sampling effects in CDF construction dominate over the noise of original brightness data and the CDF errors do not show as signal noise, no standard error estimates are available. Therefore, the only way to test the reliability of our method is via simulations. In our setup, we utilize synthetic data for brightness measurements and attempt to reconstruct the (p, β) distributions. In the simulations with the synthetic data, we use the same geometries (i.e. the direction vectors e_{sun} and e_{earth} as seen from the asteroid-fixed frame) and measurement time information that are used in the real asteroid databases, such as the Lowell Observatory database (Bowell et al. 2014), the Asteroid Lightcurve Data Exchange Format (ALCDEF; Warner et al. 2011), the Uppsala Asteroid Photometric Catalogue (UAPC; Lagerkvist et al. 1987; Piironen et al. 2001) and the Wide-Field Infrared Survey Explorer (WISE; Mainzer et al. 2011).

In our forward model, we take asteroid models from DAMIT and apply basic transformations such as stretching on them to obtain a desired shape distribution with a large number of objects. For DAMIT shapes, the concept of elongation is no longer as well defined as for ellipsoids, but we estimate $p = b/a$ simply by choosing a to be the longest diameter in the equatorial xy-plane, and b the width in the corresponding orthogonal direction. In this way, we generate a (p, β) distribution with one peak by choosing suitable values of p and β for the objects.

For computing the brightnesses of the synthetic asteroids, we use a combination of the Lommel–Seeliger and Lambert scattering laws as in Kaasalainen & Lamberg (2006). We also add random perturbations to L to simulate noise. When the brightness function has been computed, η can be obtained using the discrete approximation of $\Delta(L^2)/\langle L^2 \rangle$ from the available synthetic data points close enough in time to depict one rotation in a fixed geometry. Then, we get A from Eq. (8), and thus, the CDF of A: that is, $C(A)$. Other scattering models such as Hapke's could be used as well, but this represents only small brightness changes to separate objects and is thus not relevant to the collective results here. In fact, asteroid models in DAMIT are mostly constructed using the combined Lommel–Seeliger and Lambert law, so this choice reproduces the typical observed asteroid brightnesses best in this simulation.

In the inverse problem, we attempt to reconstruct the original distribution. First of all, the competence of our method depends on how close our obtained distribution is to the original one. We can check numerically how well the function series $\sum_{ij} w_{ij} F_{ij}(A)$ of Eq. (5) converges to the CDF of A, $C(A)$, by computing the relative error:

$$\frac{\left\| C(A) - \sum_{ij} w_{ij} F_{ij}(A) \right\|}{\|C(A)\|} \overset{\text{Eq. (9)}}{=} \frac{\|C - Mw\|}{\|C\|}. \tag{12}$$

Figure 3 depicts a typical fit of the analytical basis functions F_{ij} to the data created with complex shapes and sampled more sparsely than implicitly assumed by the CDF integrals. We can see that the model usually fits CDF data perfectly, so the analytical basis functions provide a very good set despite the crudeness of the model approximations. The main question is thus the accuracy of the result rather than the explainability of the model.

In our simulations with synthetic data, we generate populations with a single (p, β) peak in their joint distribution, and we attempt to reconstruct this peak. Each asteroid can have multiple brightness measurements, and we require at least five observation points for a valid estimate of the variation observable η. From the results of the simulations, we have found this to be
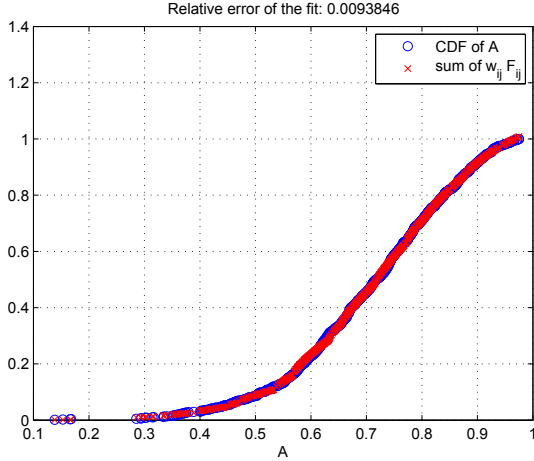
Fig. 3. Function series $\sum_{ij} w_{ij}F_{ij}$ from Eq. (5) plotted in the same figure with the CDF $C(A)$ from data. The minimal error in the fitting should be noted.

the typical minimum number of data points for sampling one rotation of the target. This is also simple to estimate analytically by considering the possible permutations of random samples of a boxcar-shaped sinusoidal signal: then the average error of η drops fast from the 50% of two sample points to close to 10% with five or six points. Such error levels already fit well in the total error budget. Numerical examples of lightcurves give similar results.

In order to obtain enough observations of η and thus accurate distributions, we use populations of 1000 asteroids. This is a realistic population size, as real major databases contain $\propto 10^3$–10^4 objects. Smaller populations start to suffer from too sparse sampling of geometries. Also, the systematic errors caused by the inaccurate assumptions of the model and the sparsity of rotational phases in η estimation are best counteracted by the averaging effect of a large number of samples. For the bins, $p_i \in [0, 1]$ and $\beta_j \in [0, \pi/2]$, where we have selected $i = 1, \ldots, 20$ and $j = 1, \ldots, 29$, with a random point near the centre of each equally spaced bin to represent its β or p value. This way, every bin is about 0.05×0.05 units in size, and the computation of the inverse problem is fast enough.

We have plotted some all-round cases of the (p, β) distribution of the forward model and the solution distribution of the inverse problem (from data mimicking WISE) in Figs. 4–8. In the forward model, the peak of the (p, β) distribution has been placed in the middle, bottom left, top left, bottom right, and top right positions in the (p, β) plane, respectively. In Fig. 4, we notice that the approximate location of the (p, β)-peak is correct, but the solution spreads when moving away from the peak, particularly when moving towards $\pi/2$ (spin direction in the ecliptic plane). In addition, the peak is too much to the left in p-axis (towards more elongated bodies). The same phenomenon can be observed in the other figures as well. In Fig. 6, when the peak was located in the top left corner of the (p, β) plane (elongated bodies in the ecliptic plane), the contour looked visually messy every time. Hence, we included an additional plot of the solution of the inverse problem in the $(p, \beta, \mathrm{DF}(p, \beta))$ coordinates. From the three-dimensional perspective, we notice that the peak of the shape elongation is once again too far in the left in p-axis, and the solution spreads when moving away from the peak. Indeed, these errors are systematic, and they occur in a solution every time. The p-shift is inevitable: even in the absence of noise, near-spherical targets with $p = 1$ do not portray a completely
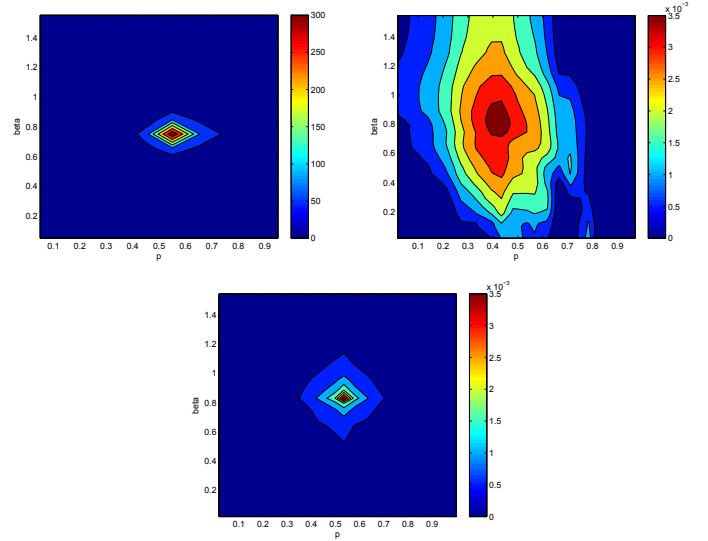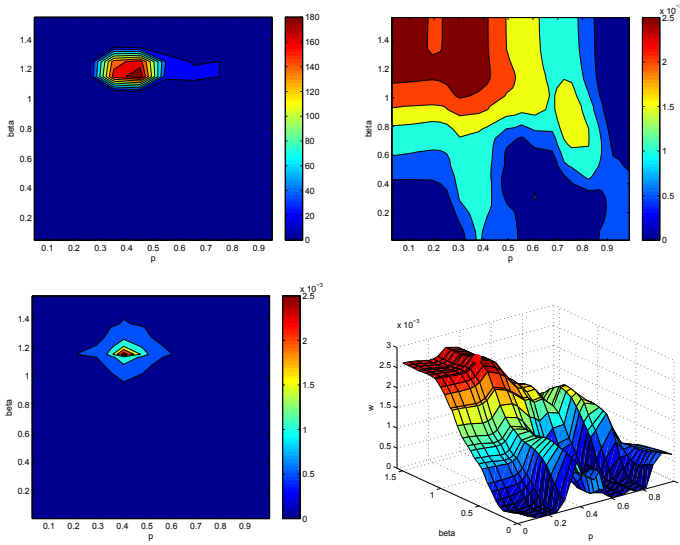
Fig. 4. Actual joint distribution of (p, β) (*top left*) of synthetic asteroids compared to the solution of the WISE-based inverse problem (*top right*). The colours depict the occupation number of each (p, β)-cell (on arbitrary scales). The absolute value of the ecliptic latitude of the spin axis decreases from bottom (perpendicular to the ecliptic plane) to top (in the ecliptic plane), and the shape elongation decreases from left (thin cigar) to right (sphere). On the *bottom* is the solution of the inverse problem after applying deconvolution for correction.
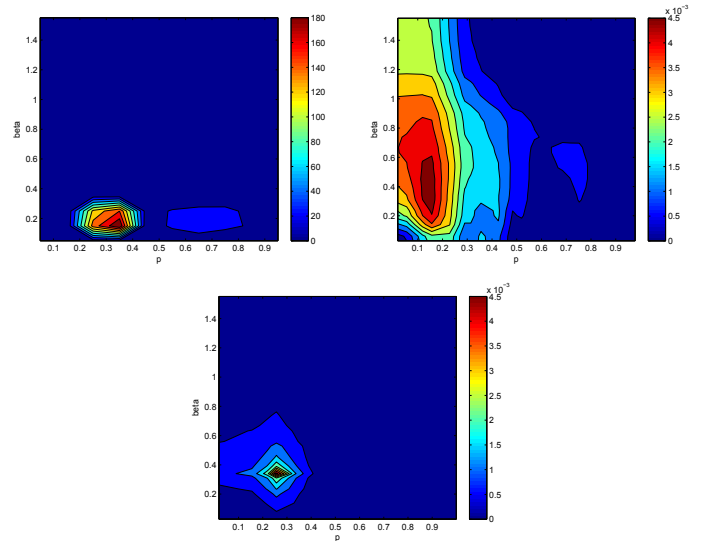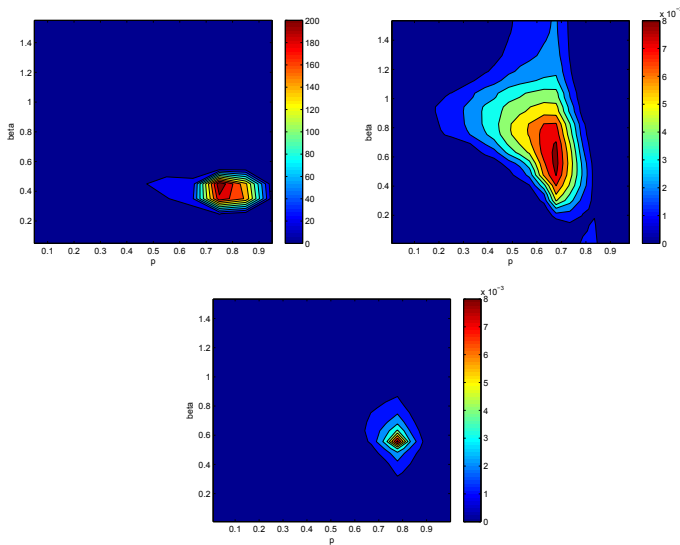

Fig. 5. Actual joint distribution of (p, β) (*top left*) of synthetic asteroids compared to the solution of the WISE-based inverse problem (*top right*). Here we have tested how accurately the solution is obtained if the peak of both p and β distributions is low. On the *bottom* is the solution with deconvolution added.

flat lightcurve because of local shape irregularities. In addition, Figs. 5, 7, and 8 show a trend of the peak of the β solution to move slightly towards the middle (away from the $\beta = 0$ and $\beta = \pi/2$ ends). The error is common but does not occur every time. In general, the errors are encountered because of both modelling errors and noisy measurements. They are rather regular and predictable, and therefore, it is possible to formulate a post-solution correction in order to revise the solution distribution.
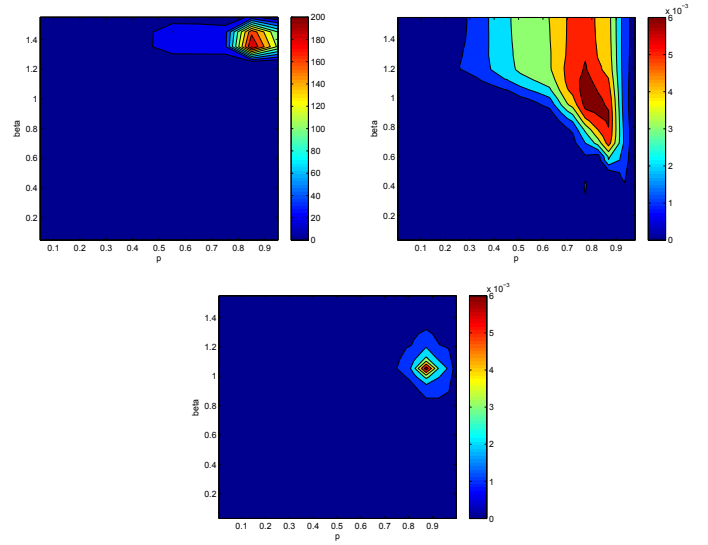
Fig. 6. Actual joint distribution of (p, β) (*top left*) of synthetic asteroids compared to the solution of the WISE-based inverse problem (*top right*). To make the solution plot more easily readable, we have plotted it from another perspective on the *bottom right*, with the *z*-axis depicting the weights w of each (p, β) bin. Here we have tested how accurately the solution is obtained if the peak of the distributions is low for p and high for β. On the *bottom left* is the solution with deconvolution.

Fig. 7. Actual joint distribution of (p, β) (*top left*) of synthetic asteroids compared to the solution of the WISE-based inverse problem (*top right*). Here we have tested how accurately the solution is obtained if the peak of the distributions is high for p and low for β. On the *bottom* is the solution with deconvolution.

4.2. Correction in (p, β)-plane

The "deconvolution" of the noisy solution "image" in the (p, β)-plane is a visual aid based on experiments performed on the synthetic data. With the help of simulations, we were able to acquire a good understanding of how much our computational solution typically differs from the actual distribution when one assumes that there is a dominant peak in the latter. This way, we could deduce the typical point-spread function of the solution in the plane. We introduce damping on bins further away from the peak of the centre of the solution. Then we move the values of p a constant (fixed) step to the right: $p_i \rightarrow p_i + \Delta\mathcal{P}$. According to our simulations with the WISE database, the solution is typically

Fig. 8. Actual joint distribution of (p, β) (*top left*) of synthetic asteroids compared to the solution of the WISE-based inverse problem (*top right*). Here we have tested how accurately the solution is obtained if the peak of both p and β distributions is high. On the *bottom* is the solution with deconvolution.

shifted about 0.1 p-units to the left due to noise. Therefore, we choose $\Delta\mathcal{P} = 0.1$ when we use WISE data. The systematic error in β direction is irregular, and there is no way to know whether the obtained β is too small or too large. For β, we observed that the obtained distribution is usually accurate if the actual β peak is somewhere near $\pi/4$ (usually when $\beta \in [0.5, 1]$), but if the actual peak is near the extreme end values 0 or $\pi/2$, then the solution tends to shift the peak away from the extremes, towards the middle values.

We show the deconvoluted solution in the bottom figures of Figs. 4–8 (bottom left picture in Fig. 6). The deconvolution has been used to correct the solution presented in the top right picture of the same figure. The corrected solution is close to the distribution shape of the forward model in the top left picture. We will apply deconvolution solely on the joint (p, β) distribution. In order to reduce errors and loss of information, the marginal p and β distributions are presented without corrections. In their cases, the main point is that the peak of the p distribution is usually slightly more to the right in the p-axis than in the obtained solution.

5. Results from astronomical databases

In this section, we plot distributions of different asteroid families and introduce a method for comparing such distributions. The setup we use is very similar to the one used for synthetic data. We receive our data (geometries, brightness values, measurement times) from the WISE database. We downloaded the data from the Infrared Science Archive: infrared Processing and Analysis Center (IRSA/IPAC archive[4]) and used the same selection criteria as Ali-Lagoa et al. (2014). The combined ALCDEF & UAPC lightcurve database (hereafter called simply ALCDEF) is also useful for various analyses, but its denser lightcurves (yielding improved η estimates) do not really compensate for the larger number of objects in the WISE data that is crucial to the robustness of the statistical CDF approach, as discussed earlier. Moreover, the lightcurves sample well only asteroids with short

---

[4] http://irsa.ipac.caltech.edu/Missions/wise.html

rotation periods, because observations from different nights usually cannot be combined together because of poor or completely missing calibration. The large number of brightness variation samples, rather than the accuracy of the observable, is the main reason why the method can tolerate the crude underlying shape model. Even though the WISE data are in mid-infrared wavelengths (we used measurements at 12 and 22 μm), the η derived for them is essentially the same as from the projected area since the infrared regime mainly causes a lag in the lightcurves (compared to visual data) that does not affect the brightness variation (Ďurech et al. 2016a).

Another possible rich source of asteroid photometry is the Lowell Observatory photometric database Bowell et al. (2014), which was used by Oszkiewicz et al. (2011) and Cibulková et al. (2016). However, the large errors of photometric points of ∼0.1–0.2 mag would bring another source of systematic error into our model.

We consider one η-estimate to consist of measurements done within a three-day time window to keep the observing geometry sufficiently constant. We only accept estimates based on at least five measured values. In principle, our model requires the phase angle between the Sun and Earth to be close to zero degrees. However, adding this restriction would greatly reduce the number of brightness measurements we could use, and would eventually lead to a considerably lower number of η values. According to simulations performed on synthetic data, the error caused by a non-zero phase angle is so small compared to other error sources that its effect is negligible. Therefore, we set a very liberal requirement that

$$\arccos(\boldsymbol{e}_{\text{sun}} \cdot \boldsymbol{e}_{\text{earth}}) \leq 30°. \tag{13}$$

5.1. Discussion about bias

Before we move on to plotting asteroid families, we discuss some possible sources for biases. The number of possible η-estimates varies between asteroids; if an asteroid yields n estimates of η, we can formally give each of these estimates the weight $1/n$. However, if there are many estimates associated with some asteroids, there could be a bias, as the solution could be favouring such targets. We checked if the solution was affected if we only took one estimate for each target. There was no noticeable change from the situation when all estimates were considered, so we can conclude that there is no significant bias from the weights of individual asteroids when using large databases for a large number of objects. This underlines the safety in large numbers, so the method can be used even if we do not know which observation is from which asteroid and use all η estimates "blindly".

As we know from asteroid lightcurves, large asteroids are generally more spherical than small asteroids. This was also shown by Cibulková et al. (2016), for example. We checked this result by comparing WISE subpopulations of different sizes. We divided the WISE asteroids into four subpopulations: ones with diameter $D < 10$ km (≈65 000 bodies), 10 km $\leq D < 25$ km (≈6000 bodies), 25 km $\leq D < 50$ km (≈1000 bodies), and $D \geq 50$ km (≈1000 bodies). While the population sizes are noticeably different (the number of small asteroids clearly surpasses the number of large ones), the subpopulations were selected so that each of them would have a sufficiently large sample size in order to acquire reliable results. The obtained distributions confirmed that large asteroids indeed tend to be more spherical (see Fig. 9 for comparison). Our result is considerably different from the one obtained by McNeill et al. (2016), where the peak of the
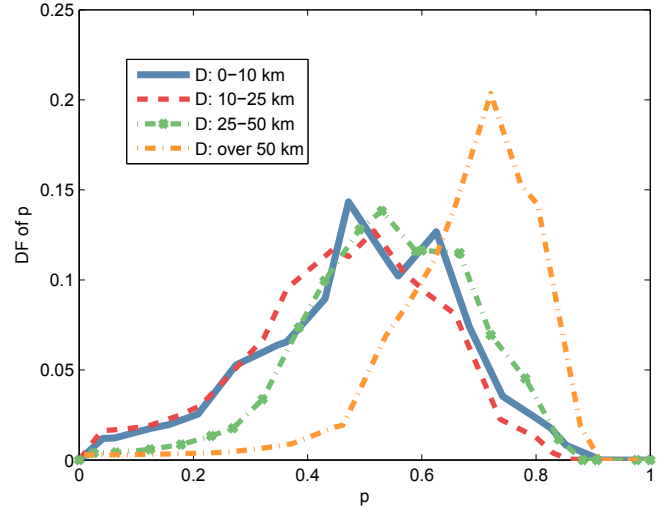
Fig. 9. Comparison of the marginal DFs of the shape elongation p for asteroids of different sizes from WISE data. The vertical axis depicts the occupation number of each p-slot (on an arbitrary scale). The shape elongation decreases from left (thin cigar) to right (sphere).
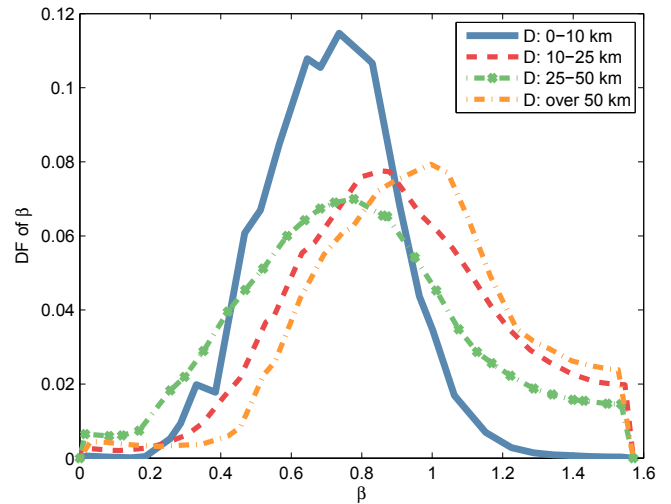

Fig. 10. Comparison of the number densities of the spin β for asteroids of different sizes from WISE data. The absolute value of the ecliptic latitude of the spin axis decreases from left (perpendicular to the ecliptic plane) to right (in the ecliptic plane).

p distribution for $D < 8$ km was located at a near-spherical value of $b/a = 0.85$. Of course, the distribution tail of small p-values of our result for $D < 10$ km is more due to systematic and model errors (especially irregular shapes) than to actually very elongated bodies.

Similarly, large asteroids tend to have their spin axes closer to the ecliptic plane than the small ones (see Fig. 10), which qualitatively agrees with the results of Hanuš et al. (2011) and Ďurech et al. (2016b), who studied the distribution of spins of asteroids on a sample of several hundred individual models. Contrary to the results based on individual models where the spins are clustered towards poles of ecliptic ($\beta = 0$ in our notation), our analysis shows that small asteroids have a remarkably sharp peak extending some 15° on both sides of the ecliptic spin latitude of 50°, but the lack of values $\beta \approx 0$ might be caused by some systematic effects of our simple model (but we note the effect of $\sin\beta$ in the plot as discussed in Sect. 2.1).
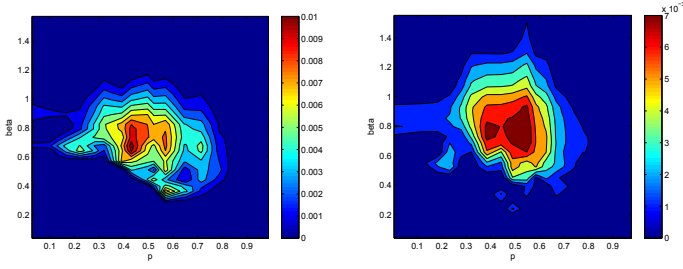
Fig. 11. Contour solution of the joint (p, β) distribution based on WISE (*left*) and WISE+ALCDEF data (*right*). Deconvolution is not used to preserve the double-peak information.
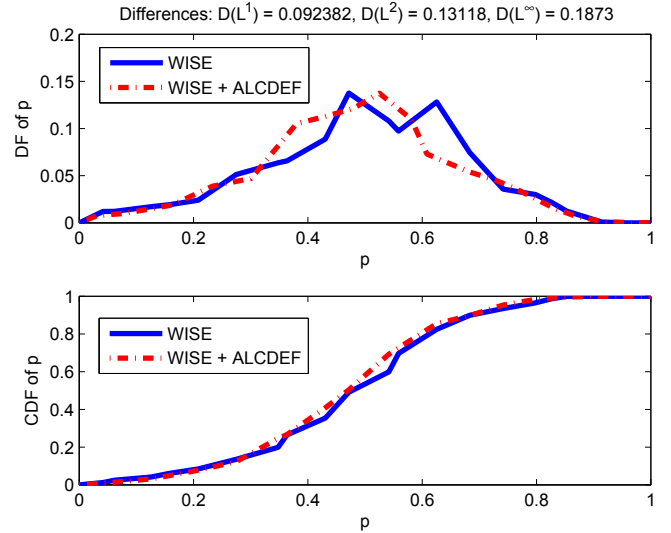

Fig. 12. Comparison of the marginal DFs (*top*) of the shape elongation p for the entire WISE population and merged ηs from WISE and ALCDEF databases, and of their marginal CDFs (*bottom*).

When comparing asteroid families, the size of the targets may thus be a factor. The shape and/or spin difference between two families may be partly driven by the difference in their size distributions. On the other hand, the best statistical material is acquired without adding the size as another dimension in the distribution function since there usually are just not enough targets to split a family into size bins. Thus we report the family p and β distributions here as such, using all family members without considering the size distribution a bias factor, even though a closer analysis between families may require taking at least the division into small and intermediate sizes into account. The proportion of large asteroids is small, so their contribution to the population distributions is usually small as well. As an example, we investigated the Eos family of well over 3000 η estimates. We performed some comparisons of p and β distributions for the whole family and its subset of smaller bodies with a diameter less than 20 km. For the p distribution, the inclusion of large asteroids mainly affected the width of the peak and there was no noticeable difference in the distributions. The differences were even smaller for the β distribution. These results suggest that the biases caused by large objects are insignificant and we will include full populations in our examples. For this and other comparison purposes, we propose a measure of difference tailored to our case (instead of the Kolmogorov–Smirnov test that typically produces indecisive statistics).

Let S_1 and S_2 be two sets of population samples. For S_i, let $F_p(S_i)$ and $F_\beta(S_i)$ be the CDFs for the marginal distributions of DF solutions for p and β, respectively (normalized to the interval $[0, 1]$). We define the statistical difference measure between S_1 and S_2 as

$$\begin{cases} D_p(S_1, S_2) = \alpha_k \left\| F_p(S_1) - F_p(S_2) \right\|_k \\ D_\beta(S_1, S_2) = \alpha_k \left\| F_\beta(S_1) - F_\beta(S_2) \right\|_k, \end{cases} \quad (14)$$

where usually $k = 1$, $k = 2$ or $k = \infty$, and α_k is a norm-based scaling factor to fix the statistical difference to the same magnitude for all norms; typically, $\alpha_1 = 1/4$, $\alpha_2 = 1$ and $\alpha_\infty = 2$. The case $k = \infty$ is used in the Kolmogorov–Smirnov test. In addition to the L^∞ norm, we will also compute the L^1 and L^2 norms, as this way we will have a better understanding of the type of statistical difference, for example, do the distributions differ in terms of the maximum or mean difference. Generally, our simulations suggest that in this context two distributions can be considered statistically different if $D \gtrsim 0.2$, although one number does not tell the whole story, and it is more instructive to perform a visual inspection on the marginal DF and CDF plots.

In Eq. (14), we chose to compare the CDFs rather than the DFs since the latter is the derivative of the former and CDFs are monotone functions, so computing the norm of their differences is more stable when one aims at one number depicting the

difference. The CDF difference tells whether the distributions are really different in the first place, and the DFs give additional details of the potential differences.
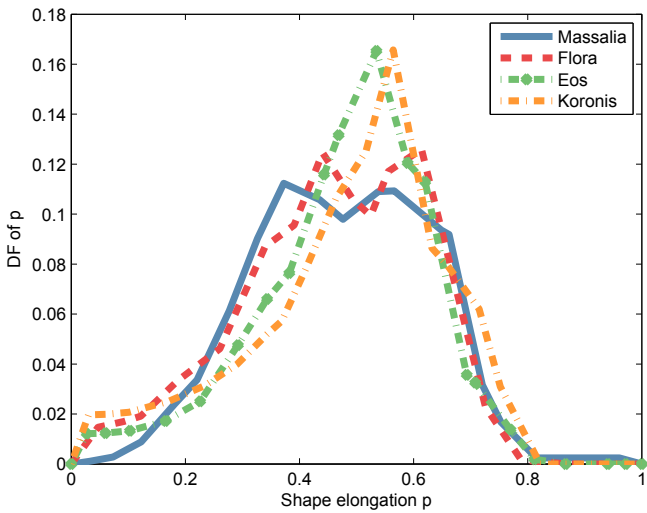
Finally, we ran tests to see if there is a bias associated with the selected database. WISE is one of the largest databases, with our method being able to cover about 85 000 asteroids. We can get about 86 000 values of η from WISE, which means we can get approximately one value for our observable from each asteroid. Indeed, WISE is one of the biggest asteroid databases available. The average number of brightness measurements available for one η is $\langle n_L \rangle \approx 9$. The ALCDEF lightcurve database contains about 14 000 asteroids suitable for our method, and yields about 39 000 values of η, resulting in less than three values of η from each asteroid. For ALCDEF, $\langle n_L \rangle \approx 18$. Despite its smaller sample of asteroids, the ALCDEF's η-per-asteroid ratio is better than WISE's. The drawback of the ALCDEF is the weak or nonexistent calibration as well as selection effects.

We perform a consistency check by comparing two distributions. First, we compute the solution of the inverse problem using all ηs obtained from the WISE population. Then we do the same, but using all the ηs from both WISE and ALCDEF. The (p, β)-plane plots are shown in Fig. 11.

The p distributions of WISE and WISE+ALCDEF are plotted in Fig. 12, while the β distributions of WISE and WISE+ALCDEF are plotted in Fig. 13. We have not used deconvolution procedures since, being designed for one-peak distributions, they would smooth out the two-peaked result. For p distributions, a visual inspection shows some differences between the DFs, such as the unrealistic boost of p-values around 0.4 by the ALCDEF addition. The CDFs are similar enough to suggest that the added ALCDEF data do not greatly distort the p distribution. For β, on the other hand, the addition of the ALCDEF data shifts the distribution to the right, and the β distribution obtained from the hybrid data is obviously different from the one with WISE data only. Therefore, we conclude that there may be a database-related bias included, especially with the β solutions, and it is advised to use caution in the selection of a database. Indeed, the ALCDEF data are distorted by a number of selection effects due to the visibility and popularity of the targets. WISE targets are more evenly and comprehensively spread and
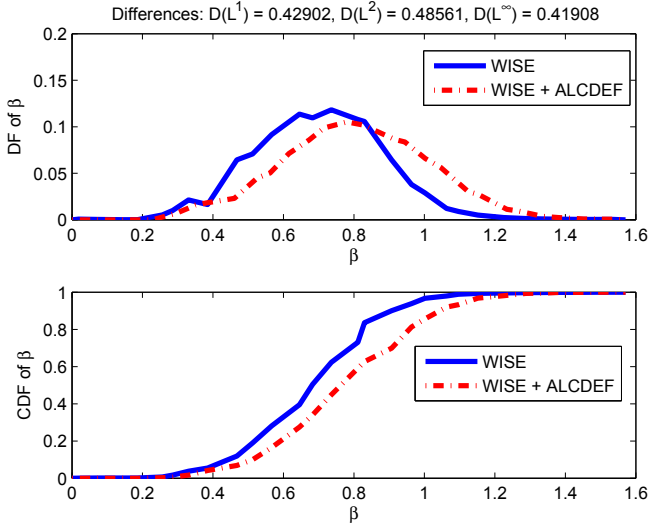
Fig. 13. Comparison of the marginal DFs (*top*) of the spin β for the entire WISE population and merged ηs from WISE and ALCDEF databases, and of their marginal CDFs (*bottom*).
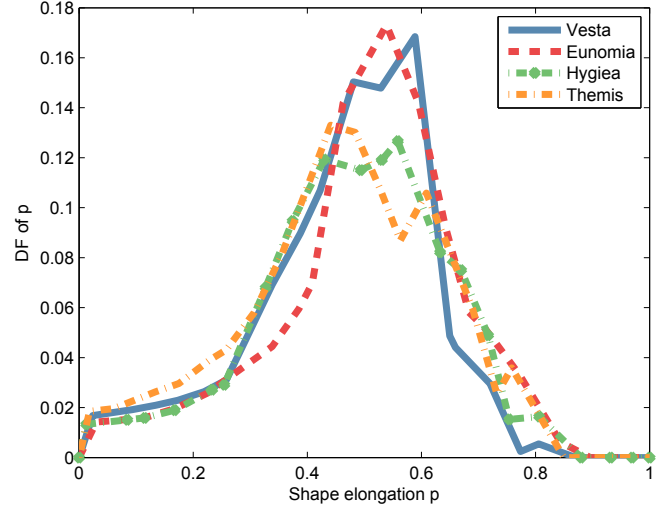

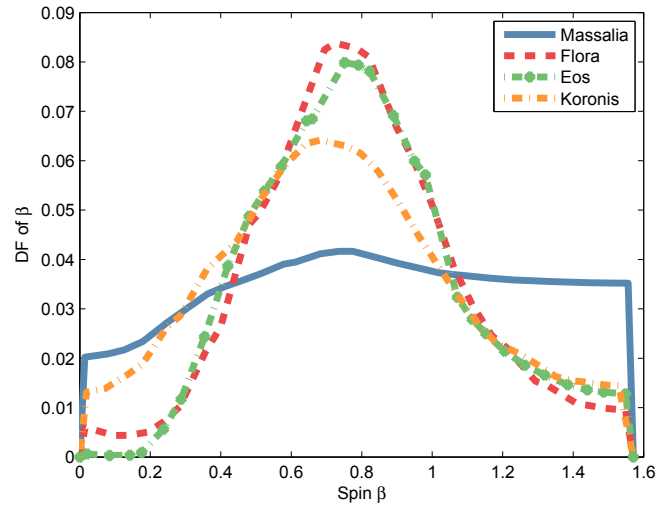**Fig. 15.** Obtained p distributions for the asteroid families Vesta, Eunomia, Hygiea, and Themis.

Fig. 14. Obtained p distributions for the asteroid families Massalia, Flora, Eos, and Koronis.

Fig. 16. Obtained β distributions for the asteroid families Massalia, Flora, Eos, and Koronis.

observed from a satellite, so the biases are smaller. Due to this and the sufficiently large number of database targets, we consider WISE data more reliable for distribution analysis and use them in the studies below.

5.2. Examples of distributions and their comparison

Below we list some results from family distribution reconstruction and comparison by their inferred marginal distributions of p and β. We chose families that were interesting from a statistical point of view, mainly to demonstrate the features, differences, and similarities the method can discover. Obviously, in addition to the interpretation and analysis of the results, there are many other families as well as populations other than families to consider in further work from astronomical points of view. We note that the computed distributions for asteroid families vary slightly in the figures, just to illustrate that randomized inverse grids lead to slightly different details in distribution solutions.
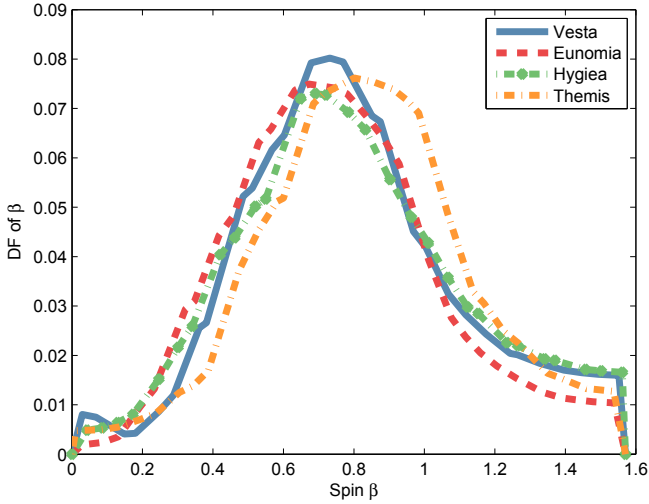
The shape elongation and spin distributions for eight different asteroid families are shown in Figs. 14–17. The obtained shape elongation distributions are different from those obtained by Szabo & Kiss (2008), who assumed a uniform distribution of spin axes and utilized the less reliable two-point brightness scatter observable. The number of available asteroid samples per family in the WISE database varies. Flora and Eos have about 3000 WISE samples, while Vesta, Eunomia, Hygiea, and Themis have 1000–2000. Koronis has 642 WISE samples, while Massalia is limited to only 154 samples. According to our simulations, solving the inverse problem several times for Koronis leads to fairly good regular solutions, so the sample size can be trusted to be large enough. Massalia, on the other hand, has stability problems with the small sample size, so its results cannot be considered to be as reliable as the others, but we include it for completeness. Typically, a sample of at least 500 objects is required in order to obtain stable solutions that can successfully recover from the model errors and noise.
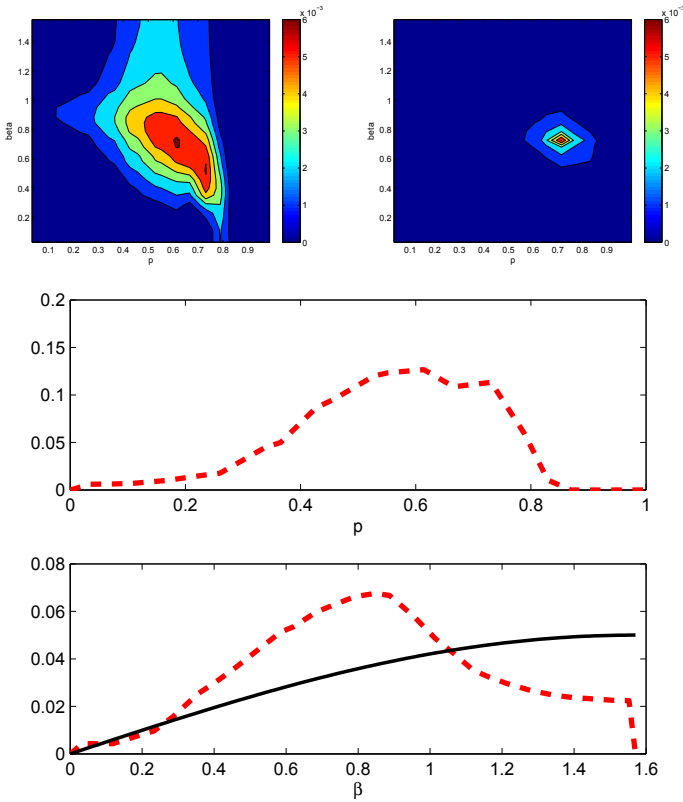
We observed that the Alauda family (some 800 samples) contains a somewhat higher ratio of near-spherical bodies than other families. This is likely to be an intrinsic quality of the family, as the same result holds when the large asteroids ($D > 20$ km) have

Fig. 17. Obtained β distributions for the asteroid families Vesta, Eunomia, Hygiea, and Themis.

Fig. 18. Contour solution of the joint (p, β) distribution of the Alauda family (*top left*), the deconvoluted smoothing of the same solution (*top right*), and the normalized marginal distributions for p (*middle*) and β (*bottom*). The black solid curve ($\sin \beta$) depicts the curve shape of a constant level of spin distribution on the sphere. Deconvolution is not used for the marginal distributions in order to avoid the loss of information in the smoothing.

been filtered out. The contour solution of the joint (p, β) distribution, both with and without deconvolution, as well as the marginal p and β distributions (without deconvolution) for the Alauda family are plotted in Fig. 18. We plotted an additional $\sin \beta$ curve in the β plot to illustrate what the spin DF would look like if it was uniformly distributed on the sphere.

To show examples of difference classes between families, we consider cases of small, borderline, and large values of the

Fig. 19. Comparison of the marginal DFs (*top*) of the spin β for Gefion and Koronis families, and of their marginal CDFs (*bottom*).

difference measures. The very similar β distributions of the Gefion and Koronis families are shown in Fig. 19, while the β distributions of the Phocaea (some 1000 samples) and Alauda families are plotted in Fig. 20. Phocaea has slightly more asteroids with spins closer to perpendicular to the ecliptic plane, while both families have heavy tails close to the plane. Generally, our solutions appear to give less weight to the values of β close to 0 or $\pi/2$. In the case of $\beta = 0$, this is partly due to the factor of $\sin \beta$ in the occupation numbers. For $\beta = \pi/2$, this can be due to, for example, size distribution, noise, and orbit positions away from the ecliptic plane. In our simulations, we did not find any particular mechanism or tendency for the scarcity of solutions close to the ecliptic plane. Finally, in Fig. 21, we give an example of the clearly different shape distributions of the Themis and Alauda families.

To confirm the reliability of distributions of β and p for individual families, it would be ideal to compare our CDFs with those constructed from individual models derived by lightcurve inversion. Unfortunately, this is not possible at this stage, because the number of known models for a typical family is a few tens at most (Hanuš et al. 2013). Another possibility would be to compare CDFs reconstructed from different and independent data sets. For example, the difference between Themis and Alauda families is also significant when we do the same analysis with the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS) data (Cibulková et al., in prep.).

6. Discussion and conclusions

The statistical CDF approach is a fast way of testing hypotheses about shape and spin distributions of asteroid populations without constructing models of separate objects. It is applicable for discovering the existence of peaks in distributions of parameters and for comparing the distributions of different populations.

There are numerous possibilities of analysing and comparing asteroid families, and a comprehensive analysis and interpretation of results is not the aim of this paper. Our main goal was to discuss the usefulness and various aspects of the CDF approach, and build mathematical tools for its efficient use. As we have seen, the data and inversion procedures of the problem are, in fact, very simple and fast to generate and apply as such. The main burden lies in the judicious interpretation of the
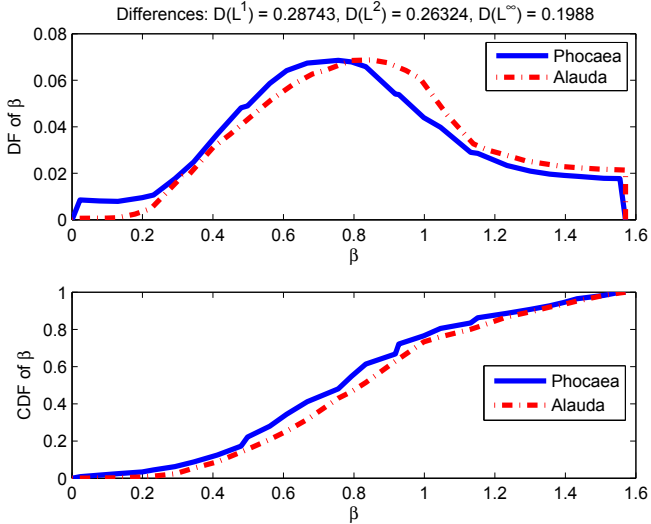
Fig. 20. Comparison of the marginal DFs (*top*) of the spin β for Phocaea and Alauda families, and of their marginal CDFs (*bottom*).
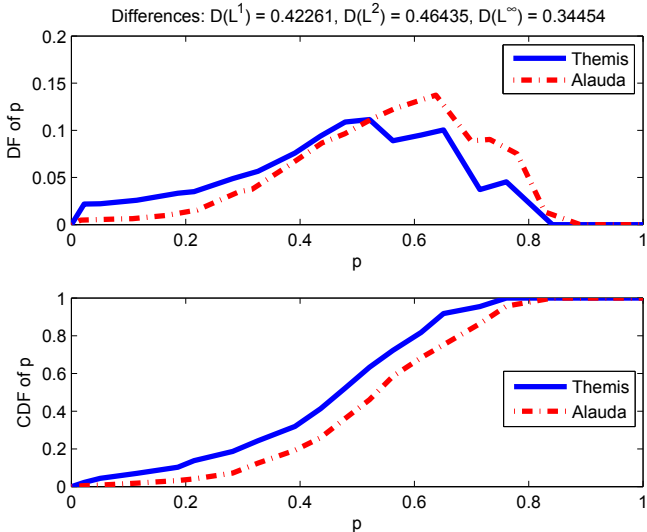
Fig. 21. Comparison of the marginal DFs (*top*) of the shape elongation p for Themis and Alauda families, and of their marginal CDFs (*bottom*).

results. A scrutiny of the usable databases may well be necessary, as some data sources may cause skewed results due to biases and/or noise.

We introduced a robust observable that can provide information on both the shape elongation and spin properties of asteroid populations, and performed an analysis on the theoretical background, providing also some examples of the obtained distributions, inspected from the statistical point of view. In our analysis, we proved that unique solutions can be obtained for both shape elongation and spin distributions, and we performed numerical simulations in order to verify that they coincide with the analytical results. It is interesting to note that, while the abundance of orbits close to the ecliptic plane means that many individual asteroid models necessarily have an ambiguity of 180 degrees in the spin longitude (Kaasalainen & Lamberg 2006), the same ecliptic orbital configuration makes possible the population-level information on spin latitudes.

Due to the model noise and the assumptions made, we cannot expect to obtain detailed, high-resolution solutions of the distributions, but we get the overall picture when the observational noise is sufficiently low (at most 0.05 mag or so). If the data noise is large, the whole point of using the brightness variations as the observable is challenged, and the prior information needed for regularization would dominate the solution. Typically, high noise means that the real p-information on near-spherical bodies disappears, and the β-information is similarly severely diluted. For low observational noise, we can additionally use a deconvolution filter to correct the systematic errors caused by modelling errors and noise. The deconvolution is a visual tool which attempts to illustrate what the distribution of the parameters actually looks like, based on prior information obtained from simulations performed for synthetic data.

Trans-Neptunian objects (TNOs) are especially interesting from the statistical point of view due to their slow orbital motion. Since their observing geometries change very slowly, proper individual models of TNOs cannot be made with the current ground-based instruments in less than tens of years. Also, for TNOs the solar phase angle is essentially zero, leading to ambiguous shape solutions from photometry. The statistical approach, however, applies to TNOs just as well as to other populations so long as there are sufficient targets in the observed set. This is, in fact, the only way to model TNO populations with data from large-scale surveys. A bonus with TNOs is that, because of the essentially fixed geometry and near-zero solar phase angle, all calibrated survey data points are usable for η-estimation even if they are separated by long time intervals. With main-belt asteroids, sparse data points from one apparition may be usable together for η estimates if one uses a solar phase correction as in Kaasalainen (2004), Ďurech et al. (2009), or Cibulková et al. (2016). The additional systematic error from this is not necessarily very large considering the total error budget.

We emphasize that the statistical use of brightness variation, while a promising approach, should always be treated with caution. Above all, any proposed type of observable and implementation should be checked with realistic simulations where the synthetic data are created with a model different from the one used in inversion. The information potential of each dataset should be assessed by using its actual observing geometries in the simulations. These simulations yield insight into the uniqueness and stability properties and accuracy expectations. Based on analytical considerations and simulations mimicking real databases, we advocate the use of the η observable and the corresponding analytical basis functions in the inverse problem.

We plan to offer a software application as a statistical (and simulation) tool that can be used for experimenting with different populations that are defined by the user. It may also be useful to construct solution procedures tailor-made for input populations. For example, one can create basis functions numerically by making synthetic CDFs for each (p, β)-bin with DAMIT-based shapes placed in the orbits of the populations.

References

Ali-Lagoa, V., Lionni, L., Delbo, M., et al. 2014, A&A, 561, A45

Bowell, E., Oszkiewicz, D. A., Wasserman, L. H., et al. 2014, Meteoritics and Planetary Science, 49, 95

Cibulková, H., Ďurech, J., Vokrouhlicky, D., Kaasalainen, M., & Oszkiewicz, D. 2016, A&A, 596, A57

Connelly, R., & Ostro, S. 1984, Geometriae Dedicata, 17, 87

Ďurech, J., Kaasalainen, M., Warner, B., et al. 2009, A&A, 493, 291

Ďurech, J., Carry, B., Delbo, M., Kaasalainen, M., & Viikinkoski, M. 2015, in Asteroid Models From Multiple Data Sources, in Asteroids IV, eds. P. Michel, et al. (Tucson: University of Arizona Press), 183

Ďurech, J., Hanuš, J., Ali-Lagoa, V., Delbo, M., & Oszkiewicz, D. 2016a, WISE data and sparse photometry used for shape reconstruction of asteroids, in Asteroids: New Observations, New Models, Proceedings of the International Astronomical Union, IAU Symp., 318, 170

Ďurech, J., Hanuš, J., Oszkiewicz, D., & Vančo, R. 2016b, A&A, 587, A48

Hanuš, J., Ďurech, J., Brož, M., et al. 2011, A&A, 530, A134

Hanuš, J., Brož, M., Ďurech, J., et al. 2013, A&A, 559, A134

Kaasalainen, M. 2004, A&A, 422, L39

Kaasalainen, M., & Lamberg, L. 2006, Inverse Problems. 22, 749

Lagerkvist, C., Barucci, M. A, Capria, M. T., et al. 1987, Asteroid photometric catalogue, Istituto di Astrofisica Spaziale, Consiglio Nazionale delle Ricerche

Mainzer, A., Bauer, J., Grav, T. et al. 2011, ApJ, 731, 53

McNeill, A., Fitzsimmons, A., Jedicke, R., et al. 2016, MNRAS, 459, 2964

Oszkiewicz, D. A., Muinonen, K., Bowell, E., et al. 2011, AAPP, 89, C1V89S1P072

Piironen, J., Lagerkvist, C., Torppa, J., Kaasalainen, M., & Warner, B. 2001, in BAAS, 33, 1562

Szabó, G., & Kiss, L. 2008, Icarus, 196, 135

Warner, B. D., Stephens, R. D., & Harris, A. W. 2011, Minor Planet Bulletin, 38, 172

Appendix A: CDF integrals and analytical basis functions

We recall the expression of the amplitude A, given by Eq. (2):

$$A = \sqrt{1 + (p^2 - 1)\sin^2 \theta}.$$

From this expression, the curves of constant A in the (p, θ)-plane are given by

$$\cos^2 \theta_A(p) = \frac{A^2 - p^2}{1 - p^2} := g_A(p). \tag{A.1}$$

The solutions for θ_A are convex "ripples" starting from the point $(p = 0, \theta = \pi/2)$ (upper left corner) for $A = 0$ and continuing to the lines $\theta = 0$ and $p = 1$ for $A = 1$ (lower right corner). Denoting the model DF of elongation by $f(p)$, we write the unnormalized CDF $C(A)$ as

$$C(A) = \int_0^{p_{\max}(A)} f(p) \int_{\theta_A(p)}^{\pi/2} \sin \theta \, \mathrm{d}\theta \, \mathrm{d}p,$$

where the minimal shape elongation needed to produce amplitude A, obtained at $\theta = \pi/2$, is $p_{\max}(A) = A$. With a change of variable $x = \cos \theta$, we get

$$C(A) = \int_0^A f(p) \int_0^{\sqrt{g_A(p)}} \mathrm{d}x \, \mathrm{d}p = \int_0^A f(p) \sqrt{g_A(p)} \, \mathrm{d}p. \tag{A.2}$$

We can also include the effect of spin distribution. Assuming λ to be isotropic and the observation directions to be in the xy-plane of the inertial frame (as they approximately are for the majority of asteroids, when this plane is that of the Earth's orbit), we study the DF $f_\beta(\beta)$ (or the joint DF $f(p, \beta)$ with p). The minimal aspect angle is $\theta_{\min} = \pi/2 - \beta$. Now, substituting $\boldsymbol{e} = (\cos \lambda_e, \sin \lambda_e, 0)$ into $\cos \theta = e_1 \sin \beta \cos \lambda + e_2 \sin \beta \sin \lambda + e_3 \cos \beta$, we have

$$\cos \theta = \sin \beta \cos \Lambda,$$

where $\Lambda := \lambda - \lambda_e$ is assumed isotropic (evenly distributed longitudes of spins and observing directions). It is sufficient to explore the region $\Lambda \in [0, \pi/2]$ as other quadrants are just symmetric multiples.

The curves of constant θ

$$\Lambda_\theta(\beta) = \arccos \frac{\cos \theta}{\sin \beta} \tag{A.3}$$

in the (β, Λ)-plane are now expanding "ripples" of increasing θ starting from the point $(\beta = \pi/2, \Lambda = 0)$ for $\theta = 0$. The CDF for θ is, with $x = \cos \beta$ (but retaining the argument β in the DF for convenience),

$$C_\theta(\theta) = \int_{\pi/2 - \theta}^{\pi/2} f_\beta(\beta) \sin \beta \int_0^{\Lambda_\theta(\beta)} \mathrm{d}\Lambda \, \mathrm{d}\beta$$

$$= \int_{\pi/2 - \theta}^{\pi/2} f_\beta(\beta) \sin \beta \Lambda_\theta(\beta) \, \mathrm{d}\beta$$

$$\overset{\text{Eq. (A.3)}}{=} \int_0^{\sin \theta} f_\beta(\beta) \arccos \frac{\cos \theta}{\sqrt{1 - x^2}} \, \mathrm{d}x.$$

(Differentiating $dC_\theta(\theta)/d\theta$ yields $\sin \theta$ when $f_\beta = 1$ as expected for isotropic spins.)

Using the complement of C_θ (i.e. $\hat{C}_\theta$ in the decreasing direction from $\theta = \pi/2$ to $\theta = 0$) to write the number of states between

$\theta_A(p)$ and $\theta = \pi/2$, our CDF $C(A)$ is, analogously with Eq. (A.2) and using Eq. (A.1),

$$C(A) = \int_0^A \left[\frac{\pi}{2} \int_0^1 f(p, \beta) \, \mathrm{d}x \right.$$
$$\left. - \int_0^{\sqrt{1 - g_A(p)}} f(p, \beta) \arccos \frac{\sqrt{g_A(p)}}{\sqrt{1 - x^2}} \, \mathrm{d}x \right] \mathrm{d}p. \tag{A.4}$$

(The use of x is merely a matter of convenience for the integration limits.)

The basis functions, that is, any CDF $C(A)$ caused by all objects having given fixed p_i and β_j, are now easy to write in closed form. They are obtained by replacing p and β in the integrands by the fixed p_i and β_j, setting $f = 1$, and using the integration limits to describe the inequalities between A and p, β to define the piecewise function $C(A)$. This replaces the integral by a sum of such basis functions each multiplied by the corresponding weight of the p_i and β_j bin. The resulting basis functions are given in Sect. 2.1.

The above assumption of most orbits to be close to the ecliptic plane is only approximate, and one can always define populations (especially those of near-Earth asteroids) for which it is not true even approximately. Thus the validity of this assumption should be checked for the targets used. However, as we show in Sect. 4, the assumption works quite well (given the large model error budget in any case) with typical asteroid populations for which there is some concentration of viewing geometries sufficiently near the ecliptic plane.

Appendix B: Ill-posedness caused by two-point variation observables

As earlier, we consider the case when θ is (approximately) the same for the pair. Now we have, for two rotation phases ϕ_0 and ϕ,

$$\frac{1 + (p^2 - 1)\sin^2 \theta \cos^2 \phi}{1 + (p^2 - 1)\sin^2 \theta \cos^2 \phi_0} = q^2,$$

so, with $0 < q \le 1$, that is, $\phi \le \phi_0$ (due to symmetry, we only need to consider the interval $0 \le \phi \le \pi/2$), we define iso-q contours in the (ϕ, θ) plane (for given p, ϕ_0) by

$$r(q, p, \phi_0, \phi) := \frac{q^2 - 1}{(p^2 - 1)(\cos^2 \phi - q^2 \cos^2 \phi_0)},$$

so, to have viable solutions for θ_q from $\sin^2 \theta_q = r$, we must have $p \le q, \phi \le \phi_0$, and

$$\cos^2 \phi \ge \frac{q^2 - 1}{p^2 - 1} + q^2 \cos^2 \phi_0 := s(q, p, \phi_0) \ge \cos^2 \phi_0,$$

so ϕ exist for given p, q, ϕ_0 only if $s \le 1$; that is,

$$\cos^2 \phi_0 \le \frac{p^2 - q^2}{q^2(p^2 - 1)} := t(q, p).$$

Denoting

$$\tilde{s}(q, p, \phi_0) := \arccos \sqrt{s(q, p, \phi_0)}, \quad \tilde{t}(q, p) := \arccos \sqrt{t(q, p)},$$

our CDF is thus

$$C_q(q) = \int_0^q f(p) \int_{\tilde{t}(q,p)}^{\pi/2} \int_0^{\tilde{s}(q,p,\phi_0)} \int_{\theta(q,p,\phi_0,\phi)}^{\pi/2} \sin \theta' \, \mathrm{d}\theta' \, \mathrm{d}\phi \, \mathrm{d}\phi_0 \, \mathrm{d}p$$

$$= \int_0^q f(p) \int_{\tilde{t}(q,p)}^{\pi/2} \int_0^{\tilde{s}(q,p,\phi_0)} \sqrt{1 - r(q, p, \phi_0, \phi)} \, \mathrm{d}\phi \, \mathrm{d}\phi_0 \, \mathrm{d}p. \tag{B.1}$$

Again, we can include the β-distribution by expanding the integral in the same way as with Eq. (A.4).

The basis function $G_i(q)$ for a given p_i in the two-point brightness scatter case is, from Eq. (B.1),

$$G_i(q) = \begin{cases} 0, & q \le p_i \\ \int_{\tilde{i}(q,p_i)}^{\pi/2} \int_0^{\tilde{s}(q,p_i,\phi_0)} \sqrt{1 - r(q, p_i, \phi_0, \phi)} \, d\phi \, d\phi_0, & q > p_i. \end{cases}$$
(B.2)

Although the ϕ-integral can be given in terms of elliptic functions, this is best computed by evaluating the double integral numerically. The maximum value of $G_i(q)$ is obtained at $q = 1$:

$$G_i(1) = \int_0^{\pi/2} \int_0^{\phi_0} d\phi \, d\phi_0 = \frac{\pi^2}{8}.$$

Our basis functions G_i are closed-form expressions of those computed by Monte-Carlo sampling in Szabo & Kiss (2008). These can be used to determine the p-distribution, although we found the accuracy inferior to the solution based on the variation observable η, which was to expected.

In principle, we can expand G_i to $G_{ij}(q)$ for a (p_i, β_j)-grid in the same way that F_i were expanded to F_{ij}. However, a notable difference between the two-index basis functions of A- or q- data is that the $G_{ij}(q)$ all reach their maxima at the same point $q = 1$ since the two-point comparison can always contain two equal brightnesses for any p and β. Thus the β_j-curves of the $G_{ij}(q)$ of a given p_i form a curve family with the same abscissae for the minimum ($q = p_i$) and maximum ($q = 1$); that is, members of the family can easily be mimicked by a superposition of other members unlike in the case of $F_{ij}(A)$. A number of simulations indeed confirmed that G_{ij} are not usable for solving the inverse problem in practice; that is, β-information is not recoverable from q-data. Adding prior assumptions on the joint distribution did not help either, as it resulted in too heavy regularization, causing the solution to become almost entirely prior-based.

The same problem plagues the lightcurve slope estimate of McNeill et al. (2016) from two neighbouring points, exacerbated by the effect of the rotation period P. For any basis function, the abscissae for the two-point ratio s lie at some minimum $s_m(p, P)$ and maximum $s = 1$. The same abscissae apply not only to all β, but also to infinitely many other combinations of p and P that yield the same s_m. Thus any basis function can be mimicked by numerous different superpositions of basis functions at other p, P, and β, making the solution of the inverse problem ambiguous without heavy prior assumptions.

Appendix C: Stability properties of shape distribution from observed brightness variation

We can analyse the inverse problem of determining $f(p)$ with the same approach as in lightcurve inversion Kaasalainen & Lamberg (2006): we expand both the observed $C(A)$ and $f(p)$ as function series, and examine the relationship between their coefficients. This shows if all coefficients of $f(p)$ can be determined, and also how fast their errors grow as a function of their degree.

Let us expand $f(p)$ as the polynomial

$$f(p) = \sum_{n=1}^{\infty} c_n p^n; \quad p \in [0, 1].$$

For isotropic θ,

$$C(A) = \sum_n c_n \int_0^A p^n \frac{\sqrt{A^2 - p^2}}{\sqrt{1 - p^2}} \, dp,$$

and from tables of integrals we find that this is

$$C(A) = A^2 \sum_n c_n A^n \frac{1}{n+1} F_1\left(\frac{n+1}{2}; \frac{1}{2}, -\frac{1}{2}; \frac{n+1}{2} + 1; A^2, 1\right),$$

where F_1 is the Appell hypergeometric function. This form can be transformed into the usual Gauss hypergeometric function $_2F_1$ so that

$$C(A) = A^2 \sum_n c_n A^n k_n G_n(A),$$

where

$$G_n(x) = {}_2F_1\left(\frac{n+1}{2}, -\frac{1}{2}; \frac{n+4}{2}; x^2\right) = \sum_j^{\infty} b_j^n x^{2j},$$

with

$$b_j^n = \frac{\left(\frac{n+1}{2}\right)_j \left(-\frac{1}{2}\right)_j}{j! \left(\frac{n+4}{2}\right)_j}; \quad (a)_j = \frac{\Gamma(a+j)}{\Gamma(a)}$$

(so $(a)_0 = 1 = b_0^n$), and

$$k_n = \frac{\sqrt{\pi}}{2(n+1)} \frac{\Gamma\left(\frac{n+3}{2}\right)}{\Gamma\left(\frac{n+4}{2}\right)},$$

so $k_n \neq 0$ decreases monotonously as n increases, and $\lim_{n\to\infty} k_n = 0$. The decrease is moderate, approximated by, for example, $\sim(n+1)^{-1}[\log(n/2+3)]^{-3/2}$ for $n < 100$. For the gamma function, $\Gamma(n + 1/2) = \sqrt{\pi}(2n - 1)!!/2^n$ and $\Gamma(n) = (n - 1)!$.

Suppose the observed $C(A)$ is expanded (to hold for $0 \le A \le 1$) as

$$C(A) = A^2 \sum_{n=1}^{\infty} a_n A^n.$$

Then

$$a_1 = c_1 k_1 \Rightarrow c_1 = a_1/k_1; \quad c_2 = a_2/k_2;$$
$$a3 = c_3 k_3 + c_1 k_1 b_1^1 \Rightarrow c_3 = (a_3 - c_1 k_1 b_1^1)/k_3,$$

and so on recursively; that is,

$$c_n = \frac{1}{k_n}\left(a_n - \sum_{i=1}^{[n]-1} c_{n-2i} k_{n-2i} b_i^{n-2i}\right),$$

where $[n]$ is $(n + 1)/2$ or $n/2$ for, respectively, odd or even n. Thus, all coefficients c_n are obtained, and their error grows as $1/k_n$, which is much slower than in, for example, lightcurve inversion.

We note that we can write a formal, more user-friendly one-to-one mapping between the polynomial coefficients determining $f(p)$ and $C(A)$. If $p \in [0, 1[$, and we expand (assuming $f(p)$ to vanish fast enough when $p \to 1$)

$$\frac{f(p)}{\sqrt{1 - p^2}} = \sum_{n=1}^{\infty} d_n p^n,$$

we have

$$C(A) = \sum_n d_n \int_0^A p^n \sqrt{A^2 - p^2} \, dp$$

$$= A^2 \sum_n d_n A^n \frac{{}_2F_1\left(-\frac{1}{2}, \frac{n+1}{2}; \frac{n+3}{2}; 1\right)}{n+1},$$

which is simply

$$C(A) = A^2 \sum_n d_n k_n A^n,$$

so we obtain a simple relationship between data and model:

$$d_n = \frac{a_n}{k_n}.$$

We also note that the above applies to the general triaxial ellipsoid as well. Let us now have a fixed $c \neq 1$, $b = 1$, and $a = 1/p$. Then

$$A^2 = \frac{p^2 \sin^2 \theta + c^{-2} \cos^2 \theta}{\sin^2 \theta + c^{-2} \cos^2 \theta},$$

so the iso-A curves are given by

$$\cos^2 \theta_{A3}(p) := g_{A3}(p) = \frac{A^2 - p^2}{h(A) - p^2},$$

where

$$h(A) := A^2(1 - c^{-2}) + c^{-2}.$$

Now

$$C(A) = \sum_n c_n \int_0^A p^n \sqrt{g_{A3}(p)} \, dp,$$

and this is

$$C(A) = \frac{A^2}{\sqrt{h(A)}} \sum_n c_n A^n \frac{1}{n+1} F_1 \left(\frac{n+1}{2}; \frac{1}{2}, -\frac{1}{2}; \right.$$
$$\left. \frac{n+1}{2} + 1; \frac{A^2}{h(A)}, 1 \right).$$

This can be used to define a series expansion for the observed $C(A)$ with new basis functions instead of polynomials, so we have the same kind of one-to-one correspondence as above.