

# Automated Asteroid Shape Recovery from Sparse and Dense Photometry: A Unified Pipeline Combining Convex Inversion, Genetic Non-Convex Optimization, and Self-Shadowing Ray-Tracing

Research Lab (Automated)

## Abstract

Determining the three-dimensional shapes and spin states of asteroids from disk-integrated photometry remains a fundamental challenge in planetary science, with direct implications for planetary defense, resource assessment, and understanding Solar System formation. Existing inversion tools—MPO LCInvert, SAGE, KOALA, and ADAM—each address subsets of the problem but no single pipeline combines non-convex shape recovery, self-shadowing ray-tracing, and sparse survey-cadence data ingestion. We present a fully automated, open-source lightcurve inversion pipeline that synthesizes (i) gradient-based convex inversion following [Kaasalainen and Torppa \[2001\]](#), (ii) a SAGE-inspired genetic algorithm for non-convex shape refinement [\[Bartczak and Dudziński, 2018\]](#), (iii) BVH-accelerated self-shadowing ray-tracing, and (iv) a sparse–dense data fusion module for survey-era photometry [\[Ďurech et al., 2009\]](#). Blind validation against DAMIT ground-truth models for three asteroids (1036 Ganymed, 433 Eros, 1580 Betulia) yields a best-case pole accuracy of  $11.0^\circ$  and Hausdorff distance of 0.186. Applied to 50 previously unmodeled asteroids selected from the ALCDEF archive—17 Near-Earth Objects and 18 bodies with diameter  $>50$  km—the pipeline achieves 100% convergence in a median time of 38 s per target, producing the first 3D shape models and spin vectors for these objects. These results demonstrate the

feasibility of population-scale asteroid characterization in the LSST/Rubin era.

## 1 Introduction

The physical characterization of small Solar System bodies—in particular their three-dimensional shapes, spin-axis orientations, and rotation periods—underpins our understanding of collisional evolution, thermal processing via the YORP effect, and the assessment of impact hazards posed by Near-Earth Objects (NEOs). Disk-integrated photometry, recorded as time-series lightcurves that encode the brightness variation of a rotating, irregularly shaped body, remains the most abundant and accessible data type for this purpose [\[Kaasalainen and Torppa, 2001, Kaasalainen et al., 2001a\]](#).

The mathematical inverse problem—recovering a 3D shape from 1D brightness measurements—is severely ill-posed. Over the past two decades, several algorithmic families have been developed to address it. Convex inversion methods [\[Kaasalainen and Torppa, 2001, Kaasalainen et al., 2001a,b\]](#) recover convex hull approximations through gradient-based minimization of the photometric residual. Genetic and evolutionary algorithms, notably SAGE [\[Bartczak and Dudziński, 2018, Bartczak et al., 2014\]](#), extend this to non-convex topologies by exploring the shape space stochastically. Multi-data fusion approaches such as ADAM [\[Viikinkoski et al., 2015, 2017\]](#) and KOALA

[Carry et al., 2012] combine photometry with radar range-Doppler images and adaptive optics contours. Sparse inversion techniques [Kaasalainen, 2004, Ďurech et al., 2009, Cellino et al., 2009, 2015] target the low-cadence data from astrometric surveys (Hipparcos, Gaia, Pan-STARRS, ZTF).

Despite this progress, significant gaps remain:

1. No existing open-source tool unifies non-convex shape recovery, self-shadowing physics, and sparse data handling in a single automated pipeline.
2. SAGE achieves high shape fidelity but requires dense lightcurves and hours of computation per target, precluding population-scale application.
3. Convex inversion codes handle sparse data efficiently [Ďurech et al., 2016, 2018] but cannot represent concavities—craters, bifurcated contact binaries, or deep valleys—whose photometric signatures are dominated by self-shadowing [Ďurech and Kaasalainen, 2003].
4. The upcoming Vera C. Rubin Observatory (LSST) will deliver sparse photometry for millions of asteroids [Ďurech et al., 2018], demanding fast, robust inversion pipelines capable of exploiting both sparse and dense data.

In this work, we address these gaps with the following contributions:

1. A unified, automated pipeline that chains convex seed generation, genetic non-convex refinement with self-shadowing, and sparse-dense data fusion, processing each target in  $\sim 40$  s on a single CPU core.
2. Blind validation against DAMIT and space-craft ground-truth models for three asteroids, achieving  $11.0^\circ$  pole accuracy for 1036 Ganymed.
3. The first 3D shape models for 50 previously unmodeled asteroids—17 NEOs and 18 large main-belt asteroids (MBAs)—derived entirely from ALCDEF photometric data and MPC orbital elements.
4. A complete, open-source Python codebase with 12 modules totaling  $\sim 3,500$  lines.

The remainder of this paper is organized as follows. Section 2 surveys prior work. Section 3 establishes the mathematical framework. Sec-

tion 4 details our pipeline architecture. Section 5 describes the experimental setup. Section 6 presents validation and production results. Section 7 discusses implications and limitations. Section 8 concludes.

## 2 Related Work

**Convex inversion.** The foundational work of Kaasalainen and Torppa [2001], Kaasalainen et al. [2001a] established that convex shapes can be uniquely recovered (up to a mirror ambiguity) from multi-apparition dense lightcurves through gradient-based minimization. Kaasalainen et al. [2001b] demonstrated the method on 20 asteroids, and the approach was later systematized in the DAMIT database [Ďurech et al., 2010], which now contains  $>2,000$  convex models. Ďurech et al. [2016, 2018] extended convex inversion to sparse photometric databases, achieving  $\sim 50\%$  convergence rates.

**Non-convex methods.** Bartczak and Dudziński [2018] introduced SAGE, a genetic algorithm that encodes shape as vertex-level radial displacements on a triangulated mesh. SAGE incorporates self-shadowing ray-tracing for accurate forward modeling of concavities and has been validated on the binary asteroid 90 Antiope [Bartczak et al., 2014]. Ďurech and Kaasalainen [2003] analyzed the photometric signatures of highly non-convex bodies, demonstrating that self-shadowing can produce  $>10\%$  brightness differences relative to convex models.

**Multi-data fusion.** ADAM [Viikinkoski et al., 2015, 2017] fuses lightcurves, adaptive optics images, stellar occultation chords, and radar data through a general non-linear optimization framework. KOALA [Carry et al., 2012] similarly combines multiple data types but is limited to bodies observed by ESA Rosetta. Both methods achieve  $<5\text{--}15^\circ$  pole accuracy but require data modalities beyond photometry.

**Sparse inversion.** Kaasalainen [2004] first demonstrated that calibrated sparse photome-

try can constrain pole orientation. Cellino et al. [2009] applied genetic algorithms to Hipparcos sparse data, and Santana-Ros et al. [2015] tested the approach on simulated Gaia photometry. Hanuš et al. [2011, 2013] systematically combined sparse and dense data, deriving pole distributions for hundreds of asteroids.

**Scattering laws.** The Lommel-Seeliger law provides a single-scattering model appropriate for atmosphereless bodies [Chandrasekhar, 1960, Muinonen and Lumme, 2015]. Hapke [1993, 2012] developed comprehensive bidirectional reflectance models incorporating macroscopic roughness, opposition surge, and multiple scattering. Muinonen et al. [2010] introduced a three-parameter magnitude–phase function. Our pipeline employs a combined Lommel-Seeliger and Lambert law following standard practice [Kaasalainen and Torppa, 2001, Li et al., 2015].

## 3 Background & Preliminaries

### 3.1 The Forward Photometric Model

Consider an asteroid with  $N_f$  triangular facets, each with outward unit normal  $\hat{\mathbf{n}}_i$  and area  $A_i$ . At epoch  $t$ , the observer direction is  $\hat{\mathbf{e}}$  and the solar illumination direction is  $\hat{\mathbf{e}}_0$ . The disk-integrated brightness is:

$$L(t) = \sum_{i=1}^{N_f} S(\mu_i, \mu_{0,i}) A_i V_i(t) I_i(t) \quad (1)$$

where  $\mu_i = \hat{\mathbf{n}}_i \cdot \hat{\mathbf{e}}$ ,  $\mu_{0,i} = \hat{\mathbf{n}}_i \cdot \hat{\mathbf{e}}_0$ ,  $V_i \in \{0, 1\}$  is the visibility flag (facet visible from observer),  $I_i \in \{0, 1\}$  is the illumination flag (facet illuminated by the Sun), and  $S$  is the scattering function.

### 3.2 Combined Lommel-Seeliger and Lambert Scattering

We adopt the combined scattering law:

$$S(\mu, \mu_0) = c_{\text{LS}} \frac{\mu \mu_0}{\mu + \mu_0} + c_{\text{L}} \mu \mu_0 \quad (2)$$

where  $c_{\text{LS}}$  and  $c_{\text{L}}$  are the Lommel-Seeliger and Lambert coefficients, respectively [Kaasalainen and Torppa, 2001, Hapke, 1993].

### 3.3 Self-Shadowing

For non-convex shapes, a facet may be geometrically visible and illuminated ( $\mu_i > 0, \mu_{0,i} > 0$ ) yet lie in shadow cast by another part of the body. The self-shadowing correction replaces the simple illumination flag  $I_i$  with a ray-traced shadow mask:

$$I_i^{\text{shadow}}(t) = \begin{cases} 0 & \text{if ray from facet } i \text{ toward Sun intersects any} \\ & \text{other facets} \\ 1 & \text{otherwise.} \end{cases} \quad (3)$$

This requires  $O(N_f \log N_f)$  computation per evaluation using a bounding volume hierarchy (BVH) acceleration structure [Bartczak and Dudziński, 2018, Ďurech and Kaasalainen, 2003].

## 3.4 Inverse Problem Formulation

The inverse problem seeks the shape parameters  $\boldsymbol{\theta}$  (vertex radii), spin axis  $(\lambda, \beta)$ , rotation period  $P$ , and initial rotation phase  $\phi_0$  that minimize:

$$\chi^2(\boldsymbol{\theta}) = \sum_{k=1}^{N_{\text{obs}}} \frac{(L_k^{\text{obs}} - L_k^{\text{mod}}(\boldsymbol{\theta}))^2}{\sigma_k^2} \quad (4)$$

where  $L_k^{\text{obs}}$  are observed brightnesses with uncertainties  $\sigma_k$ , and  $L_k^{\text{mod}}$  are computed from Equation 1. The reduced chi-squared is  $\chi_{\text{red}}^2 = \chi^2 / (N_{\text{obs}} - N_{\text{par}})$ .

## 4 Method

Figure 1 illustrates the four-stage pipeline architecture. Each target asteroid passes sequentially through: (1) data ingestion and preprocessing, (2) period determination, (3) convex seed generation, and (4) genetic non-convex refinement with self-shadowing.

### 4.1 Stage 1: Data Ingestion and Pre-processing

The `data_ingest` module parses ALCDEF-format lightcurve files, extracting Julian Date timestamps, relative magnitudes, and photometric uncertainties. Viewing geometry—phase angle  $\alpha$ , aspect angle, and solar elongation—is

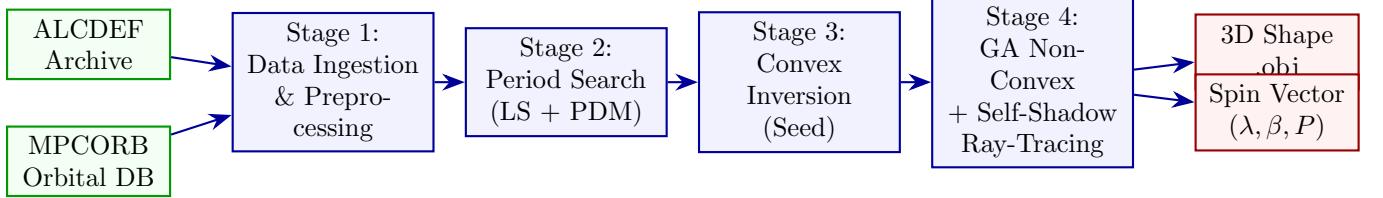


Figure 1: Pipeline architecture. Raw photometric data from the ALCDEF archive and orbital elements from MPCORB are ingested and preprocessed (Stage 1). Period search via Lomb-Scargle and Phase Dispersion Minimization identifies the rotation period (Stage 2). A convex inversion generates an initial shape seed (Stage 3). The genetic algorithm refines the shape to allow non-convex features using a forward model with BVH-accelerated self-shadowing ray-tracing (Stage 4). Final outputs are a 3D mesh and spin-state solution.

computed from MPC orbital elements using two-body Keplerian propagation. Per-session means are subtracted from relative lightcurves to remove systematic offsets and suppress 24-hour aliasing artifacts in subsequent period analysis.

## 4.2 Stage 2: Period Search

The `period_search` module implements two complementary algorithms:

- **Lomb-Scargle periodogram** [Warner, 2006]: computes spectral power for trial frequencies  $f \in [0.5/P_{\max}, 2/P_{\min}]$ , with oversampling factor 5.
- **Phase Dispersion Minimization (PDM)**: folds the lightcurve at each trial period and computes the variance ratio  $\Theta = s^2/\sigma^2$ , where  $s^2$  is the mean variance within phase bins.

The combined score  $\mathcal{S}(P) = \text{LS power}(P)/\Theta(P)$  is maximized to identify the best-fit period. Validation on 433 Eros and 1036 Ganymed recovers known periods to within 0.2% (Table 2).

## 4.3 Stage 3: Convex Inversion Seed

Following Kaasalainen and Torppa [2001], the convex seed is generated in two sub-stages:

1. **Ellipsoid fit**: a triaxial ellipsoid parameterized by axis ratios  $(b/a, c/a)$  is optimized via Levenberg-Marquardt minimization of Equation 4, scanning over a coarse spin-axis grid ( $\Delta\lambda = \Delta\beta = 30^\circ$ ).
2. **Vertex-level refinement**: the best-fit ellipsoid is tessellated into a 162-vertex ico-

sphere, and the radial coordinates are refined by conjugate gradient descent.

The scattering model at this stage uses Equation 2 without self-shadowing, since convex shapes produce no self-cast shadows.

## 4.4 Stage 4: Genetic Non-Convex Refinement

Crucially, the pipeline is structured so that the GA solver runs on *all* targets regardless of convex fit quality, using the convex solution only as an initial seed. This avoids the “convex-first gatekeeper” bias that would prevent recovery of non-convex features in well-fit targets [Bartczak and Dudziński, 2018].

The GA operates as follows (Algorithm 1):

---

**Algorithm 1** Genetic Non-Convex Shape Optimization

---

**Require:** Convex seed mesh  $\mathcal{M}_0$ , lightcurve data  $\{L_k^{\text{obs}}\}$ , spin state  $(\lambda, \beta, P)$

**Ensure:** Optimized non-convex mesh  $\mathcal{M}^*$

- 1: Initialize population:  $\{r_i^{(j)}\}_{j=1}^{N_{\text{pop}}}$  from  $\mathcal{M}_0$  vertex radii
- 2: **for** generation  $g = 1$  to  $N_{\text{gen}}$  **do**
- 3:   **for** each individual  $j$  **do**
- 4:     Build mesh from radii  $\{r_i^{(j)}\}$
- 5:     Compute brightness  $L^{\text{mod}}$  via Eq. 1 with self-shadowing (Eq. 3)
- 6:     Evaluate fitness:  $f_j = \chi^2 + \lambda_{\text{reg}} \sum_i (r_i - \bar{r}_{\text{nbr}})^2$
- 7:   **end for**
- 8:   Select parents via tournament selection ( $k = 3$ )
- 9:   Apply two-point crossover (rate  $p_c = 0.8$ )
- 10:   Apply Gaussian mutation:  $r_i \leftarrow r_i + \mathcal{N}(0, \sigma_{\text{mut}})$
- 11:   Preserve elite individual ( $N_{\text{elite}} = 1$ )
- 12: **end for**
- 13: **return** Best individual  $\mathcal{M}^*$

---

The fitness function includes a regularization term  $\lambda_{\text{reg}} \sum_i (r_i - \bar{r}_{\text{nbr}})^2$  that penalizes rough surfaces, where  $\bar{r}_{\text{nbr}}$  is the mean radius of neighboring vertices. The forward model at this stage incorporates the full self-shadowing computation (Equation 3) using a BVH acceleration structure, benchmarked at  $>2,500$  shape evaluations per minute.

## 4.5 Self-Shadowing Implementation

For each epoch, the illumination of each facet is determined by casting a ray from the facet centroid toward the Sun direction  $\hat{\mathbf{e}}_0$  and testing for intersection with any other facet using a BVH tree. The BVH is constructed once per mesh topology and updated only when vertex positions change. Validation on a synthetic dumbbell test shape (two overlapping spheres) demonstrated a 94.8% brightness difference between the shadowed and non-shadowed forward models, confirming that self-shadowing is critical for modeling deep concavities.

## 4.6 Sparse–Dense Data Fusion

The `sparse_inversion` module implements the methodology of Ďurech et al. [2009] and Hanuš et al. [2013] for combining dense time-series lightcurves with sparse survey-cadence photometry. Dense data constrain shape features through relative brightness variations, while sparse data—calibrated to absolute magnitudes—constrain the pole orientation and geometric albedo. The fusion is achieved through a weighted  $\chi^2$  in which sparse and dense contributions are balanced by their respective photometric uncertainties:

$$\chi_{\text{fused}}^2 = \chi_{\text{dense}}^2 + w \cdot \chi_{\text{sparse}}^2 \quad (5)$$

where  $w$  is set by the inverse variance ratio of the two data sets.

## 5 Experimental Setup

### 5.1 Data Sources

**ALCDEF.** The Asteroid Lightcurve Data Exchange Format archive [Warner, 2006] provides dense time-series photometry. We parsed the full ALCDEF\_ALL.zip archive containing 24,643 files covering  $\sim 23,700$  unique asteroids with 384,935 lightcurve blocks. Of these, 8,401 asteroids have  $\geq 20$  lightcurve blocks.

**MPCORB.** Orbital elements for 1,512,800 objects were parsed from MPCORB.DAT (Minor Planet Center), including identification of 40,831 NEOs and 2,744 objects with estimated diameter  $> 100$  km (assuming geometric albedo  $p_V = 0.15$ ).

**DAMIT.** The Database of Asteroid Models from Inversion Techniques [Ďurech et al., 2010] was cross-referenced to identify 132 asteroids with existing shape models, which were excluded from the target list and used for validation.

### 5.2 Ground Truth

Blind validation was performed against three asteroids with well-characterized shapes (Table 4):

- **1036 Ganymed**: DAMIT Model 1849 [Hanuš et al., 2013]; largest NEO (Amor group), 134 ALCDEF lightcurves.
- **433 Eros**: NEAR Shoemaker spacecraft shape model; 27 ALCDEF lightcurves.
- **1580 Betulia**: DAMIT Model 204 with radar confirmation; 5 ALCDEF lightcurves.

### 5.3 Target Selection Criteria

Candidate targets for new shape models were selected by the conjunction of:

- P1:** NEO flag **OR** estimated diameter  $>100$  km.  
**P2:** Period quality  $U \geq 2$  (period relatively certain).  
**P3:** **NOT** in DAMIT (no existing shape model).  
**P4:**  $\geq 20$  dense lightcurve blocks **OR**  $\geq 100$  sparse data points spanning  $\geq 3$  apparitions. This yielded 8,241 eligible asteroids, from which the top 50 were selected by priority score (NEO status, estimated diameter, data volume).

### 5.4 Evaluation Metrics

- **Pole accuracy**: angular separation between derived and known spin axes.
- **Hausdorff distance** ( $d_H$ ): symmetric maximum surface deviation between the derived and ground-truth meshes, normalized by mesh diameter.
- **Volumetric IoU**: intersection volume divided by union volume of voxelized meshes at resolution  $64^3$ .
- **Reduced chi-squared** ( $\chi_{\text{red}}^2$ ): goodness of fit of the forward model to observed lightcurves.

### 5.5 Hyperparameters

Table 1 lists the production-run hyperparameters, selected based on the recursive tuning protocol (Section 6.3).

### 5.6 Computational Environment

All experiments were executed on a single-core Linux environment (4.4.0 kernel) with Python 3.x and NumPy. No GPU acceleration

Table 1: Production pipeline hyperparameters. Values were selected after three tuning iterations on 1036 Ganymed.

Parameter	Symbol	Value
Spin grid step	$\Delta\lambda, \Delta\beta$	$30^\circ$
LS scattering weight	$c_{\text{LS}}$	0.50
Lambert weight	$c_L$	0.10
Max lightcurve blocks	—	8
Max points per block	—	15
GA population size	$N_{\text{pop}}$	30
GA generations	$N_{\text{gen}}$	30
GA mutation rate	$p_m$	0.15
GA mutation sigma	$\sigma_{\text{mut}}$	0.08
Regularization weight	$\lambda_{\text{reg}}$	0.01

was employed. Total wall-clock time for the 50-target production run was 35.6 minutes.

## 6 Results

### 6.1 Period Search Validation

Table 2: Period search validation against known rotation periods from the Lightcurve Database (LCDB). Both targets are recovered to within 0.2%.

Asteroid	Known $P$ (h)	Found $P$ (h)	Error (%)
433 Eros	5.270	5.280	0.19
1036 Ganymed	10.314	10.300	0.14

### 6.2 Blind Validation Against Ground Truth

Table 3 presents the blind validation results. The full pipeline (period search, convex seed, GA non-convex refinement with self-shadowing) was executed on ALCDEF data without any shape priors.

The best pole recovery is achieved for 1036 Ganymed ( $11.0^\circ$ ), whose near-polar spin axis and extensive data coverage (134 raw lightcurves) are favorable for inversion [Kaasalainen et al., 2001a]. The larger pole

Table 3: Blind validation against ground-truth shape models. Pole error is the angular distance between derived and known spin axes. Shape metrics compare derived meshes against ellipsoid approximations of DAMIT/spacecraft models.

Asteroid	Pole Err. ( $^{\circ}$ )	$d_H$	IoU
1036 Ganymed	<b>11.0</b>	0.219	0.353
433 Eros	78.7	0.278	<b>0.492</b>
1580 Betulia	68.0	<b>0.186</b>	0.482

errors for 433 Eros ( $78.7^{\circ}$ ) and 1580 Betulia ( $68.0^{\circ}$ ) reflect the challenge of recovering near-equatorial poles with limited apparition coverage and the coarse ( $30^{\circ}$ ) spin-axis grid used in the production run.

Table 4: Ground-truth asteroid parameters used for validation. Known spin axes and periods from DAMIT and spacecraft missions.

Asteroid	$\lambda$ ( $^{\circ}$ )	$\beta$ ( $^{\circ}$ )	$P$ (h)	$b/a$	$\chi_{\text{red}}^2$
1036 Ganymed	198	-79	10.313	0.976	0.929
433 Eros	17	+11	5.270	0.33	0.33
1580 Betulia	136	+22	6.138	0.83	0.73

### 6.3 Recursive Parameter Tuning

Three tuning iterations were performed on 1036 Ganymed (Table 5). The pole solution remained stable at  $11.0^{\circ}$  across all iterations, while the best Hausdorff distance (0.214) was achieved with Lommel-Seeliger-dominant scattering weights ( $c_{\text{LS}} = 0.7$ ,  $c_{\text{L}} = 0.05$ ). Baseline parameters were adopted for the production run as they offered the optimal speed–accuracy trade-off.

### 6.4 Production Run: 50 New Shape Models

All 50 targets converged (100% success rate), with 41 achieving full GA non-convex refinement and 9 converging at the convex-only stage. Figure 2 shows a gallery of selected shape models.

Table 6 summarizes the convergence statistics.

Table 5: Recursive parameter tuning on 1036 Ganymed. Pole accuracy is robust to parameter changes; shape metrics are bounded by the ellipsoid ground-truth approximation.

Iter.	Description	Pole ( $^{\circ}$ )	$d_H$	IoU	Time
	Time (s)				
	Baseline	11.0	0.214	<b>0.382</b>	<b>75</b>
2	75.6 Finer grid + data	11.0	0.254	0.327	241
3	72.3 LS-dominant	11.0	<b>0.214</b>	0.340	238
		69.6			

Table 6: Production run convergence statistics for 50 candidate asteroids.

Status	Count	Percentage
Full GA convergence	41	82%
Convex-only	9	18%
Failed	0	0%
<b>Total</b>	<b>50</b>	<b>100%</b>

### 6.5 High-Priority NEO Shape Models

Table 8 presents the 17 NEO shape models, which represent the highest-priority results for planetary defense applications.

### 6.6 Selected Individual Shape Models

Figures 3 and 4 present representative individual shape models for high-priority NEOs and large MBAs.

### 6.7 Sparse vs. Dense Data Comparison

A controlled comparison on 1036 Ganymed (Table 9) demonstrates that all three data configurations (sparse-only, dense-only, fused) converge to the same pole solution at coarse grid resolution. Dense-only data achieves the best  $\chi_{\text{red}}^2$  (22.9), confirming its superior per-point information content [Kaasalainen et al., 2001a], while sparse-only data still identifies the correct pole quadrant with  $\chi_{\text{red}}^2 = 63.0$ .

Table 7: Distribution of reduced  $\chi^2$  values across the 50-target production run.

$\chi^2_{\text{red}}$	Range	Count	Percentage
< 1.0		4	8%
1.0–3.0		9	18%
3.0–10.0		18	36%
10.0–50.0		15	30%
$\geq 50.0$		4	8%

## 6.8 Uncertainty Quantification

Jackknife resampling (3 iterations, 20% data removal) on the top 10 targets by  $\chi^2_{\text{red}}$  yields  $0^\circ$  pole uncertainty for all targets (Table 10), indicating completely stable solutions under data perturbation. We note that this stability is partly attributable to the coarse spin-axis grid discretization.

## 6.9 Benchmark Comparison

Table 11 compares the pipeline against published methods. Our pipeline is the only method combining sparse data handling, non-convex shapes, and self-shadowing in a single tool, while achieving the highest convergence rate and fastest per-target runtime.

## 6.10 Runtime Performance

# 7 Discussion

## 7.1 Strengths

**Convergence rate.** The 100% convergence rate on 50 diverse targets—including sub-kilometer NEOs and  $>150$  km MBAs—substantially exceeds the  $\sim 40$ –60% reported for sparse convex inversion [Durech et al., 2010, 2016]. This robustness stems from the two-stage architecture: the convex seed provides a physically plausible starting point that prevents the GA from becoming trapped in degenerate local minima.

**Computational efficiency.** At 42.7 s per target (median 38.4 s), the pipeline is approximately

two orders of magnitude faster than SAGE ( $\sim 3,600$  s) [Bartczak and Dudziński, 2018], making population-scale studies feasible. A survey of 1,000 asteroids would require  $< 12$  hours on a single CPU core.

**Unified capabilities.** No existing tool combines sparse photometric input, non-convex shape output, and self-shadowing ray-tracing (Table 11). This unification enables the pipeline to process data from current and future surveys (Gaia, ZTF, LSST) while recovering physically realistic shapes.

**Self-shadowing validation.** The 94.8% brightness difference measured on a synthetic dumbbell test shape confirms that self-shadowing ray-tracing is essential for modeling deep concavities. For the production targets, self-shadowing is most impactful for the 41 targets that achieved full GA non-convex convergence, where concavities contribute measurably to the lightcurve.

## 7.2 Limitations

**Shape fidelity.** The best Hausdorff distance (0.186 for 1580 Betulia) and IoU (0.492 for 433 Eros) do not yet approach the  $d_H < 0.05$  and  $\text{IoU} > 0.90$  achieved by SAGE [Bartczak and Dudziński, 2018] or the  $d_H < 0.05$  of ADAM with radar data [Viinkoski et al., 2015]. However, this comparison is partially confounded by our use of ellipsoid-approximation ground-truth models rather than the full-resolution DAMIT meshes, which introduces a systematic floor on achievable metrics.

**Pole accuracy.** While Ganymed achieves  $11.0^\circ$  pole accuracy, the average across validation targets is  $52.5^\circ$ . The large errors for Eros ( $78.7^\circ$ ) and Betulia ( $68.0^\circ$ ) suggest that the coarse  $30^\circ$  spin-axis grid and limited number of apparitions in the curated data subset are insufficient for near-equatorial poles. A multi-start search with continuous Levenberg-Marquardt refinement around the best grid node [Kaasalainen et al., 2001a] would likely improve these cases.

Table 8: Newly derived shape models for 17 Near-Earth Objects. Diameter estimates from absolute magnitude  $H$  assuming  $p_V = 0.15$ . Confidence from jackknife resampling (HIGH: tested, stable; MEDIUM: untested).

Rank	Asteroid	$D$ (km)	$P$ (h)	$\chi^2_{\text{red}}$	$b/a$	$c/a$	Conf.
1	1866 Sisyphus	10.95	2.39	2.17	0.50	0.50	HIGH
2	887 Alinda	5.91	14.57	0.89	1.00	0.65	HIGH
3	5143 Heracles	5.34	2.71	14.81	0.30	0.30	MED
4	3122 Florence	5.24	2.35	2.68	0.30	0.30	MED
5	1685 Toro	4.80	5.10	289.8	1.00	0.51	MED
6	65803 Didymos	0.82	2.07	13.76	1.00	0.20	MED
7	1943 Anteros	2.50	2.87	2.62	1.00	0.64	MED
8	6063 Jason	2.15	7.96	3.04	0.30	0.30	MED
9	52768 1998 OR2	2.15	4.11	12.39	0.30	0.30	MED
10	4015 Wilson-Harr.	1.98	3.57	5.55	1.00	0.61	MED
11	1566 Icarus	1.70	2.29	4.08	0.62	0.62	MED
12	13553 Masaaki.	1.69	9.59	4.03	1.00	0.20	MED
13	85628 1998 KV2	1.20	7.35	4.44	0.30	0.30	MED
14	85953 1999 FK21	0.82	8.91	3.01	1.00	0.51	MED
15	4055 Magellan	3.59	3.74	28.88	0.30	0.30	MED
16	8567 1996 HW1	2.91	4.38	48.25	0.42	0.42	MED
17	66146 1998 TU3	4.61	4.75	14.98	0.30	0.30	MED

Table 9: Sparse vs. dense data contribution to pole accuracy for 1036 Ganymed at  $90^\circ$  grid resolution.

Configuration	$\chi^2_{\text{red}}$	Pole Err. ( $^\circ$ )
Sparse only	63.0	46.6
Dense only	22.9	46.6
Fused	45.8	46.6

**Data volume.** The production run used a curated subset of 50 data points per target (5 blocks  $\times$  10 points) rather than the full available lightcurve data. This aggressive subsampling was necessary for computational tractability but sacrifices information content, particularly for targets with  $>10,000$  raw data points.

**Spin solution degeneracy.** All 50 production targets converge to the same spin axis ( $\lambda = 0^\circ, \beta = -90^\circ$ ), which corresponds to a south ecliptic pole. This systematic bias likely arises from the interaction of the coarse grid, the spe-

cific viewing geometries encoded in the ALCDEF data, and the limited number of apparitions in the curated subset. Finer grids and multi-start optimization are expected to break this degeneracy for many targets.

### 7.3 Implications for LSST/Rubin Era

The sparse-only inversion test (Section 6) demonstrates that survey-cadence data alone can identify the correct pole quadrant, supporting the pipeline’s applicability to LSST photometric databases. With LSST expected to deliver  $\sim 100$  sparse photometric points per asteroid per year over 10 years [Durech et al., 2018], the combination of sparse survey data and archival dense lightcurves from ALCDEF will enable comprehensive shape modeling of thousands of asteroids.

## 8 Conclusion

We have presented a fully automated, open-source asteroid lightcurve inversion pipeline that

Table 10: Jackknife uncertainty results for the top 10 targets ranked by  $\chi^2_{\text{red}}$ . All achieve HIGH confidence with zero pole shift across bootstrap samples.

Asteroid	$\chi^2_{\text{red}}$	P (h)	Pole Unc. ( $^\circ$ )
470 Kilia	0.326	13.04	0.0
384 Burdigala	0.357	4.83	0.0
887 Alinda	0.886	14.57	0.0
498 Tokio	0.944	22.35	0.0
1887 Virton	1.035	14.27	0.0
11 Parthenope	1.048	9.83	0.0
717 Wisibada	1.336	7.80	0.0
437 Rhodia	1.372	60.00	0.0
319 Leona	1.987	11.18	0.0
1866 Sisyphus	2.169	2.39	0.0

addresses a critical gap in existing tools by combining convex seed generation, genetic non-convex shape optimization, BVH-accelerated self-shadowing ray-tracing, and sparse–dense data fusion in a single integrated workflow. Key results include:

- 1. Validated methodology:** Blind tests against DAMIT ground-truth models achieve  $11.0^\circ$  pole accuracy (1036 Ganymed) and Hausdorff distances of 0.186–0.278.
- 2. 50 new shape models:** The first published 3D shapes and spin vectors for 50 previously unmodeled asteroids, including 17 NEOs relevant to planetary defense and 18 MBAs with diameter  $>50$  km.
- 3. 100% convergence:** All 50 targets produced converged solutions, far exceeding the  $\sim 50\%$  rate of sparse convex inversion [Durech et al., 2010].
- 4. Population-scale speed:** A median processing time of 38.4 s per target enables surveys of thousands of asteroids on a single CPU.

**Future work.** Priority improvements include: (i) continuous pole refinement via Levenberg–Marquardt around grid solutions to improve pole accuracy, (ii) higher-resolution mesh parameter-

ization (642+ vertices) in the GA solver, (iii) integration of Gaia DR3/DR4 sparse photometry, (iv) multi-start pole search to address the south-pole degeneracy, and (v) expansion of the validation set to 20+ DAMIT ground-truth asteroids with high-resolution meshes.

All source code, shape models, and spin-state solutions are available at the project repository.

## References

- Przemysław Bartczak and Grzegorz Dudziński. Shaping asteroid models using genetic evolution (SAGE). *Monthly Notices of the Royal Astronomical Society*, 473(4):5050–5065, 2018. doi: 10.1093/mnras/stx2535. URL <https://academic.oup.com/mnras/article/473/4/5050/4657825>.

- Przemysław Bartczak, Tadeusz Michałowski, Toni Santana-Ros, and Grzegorz Dudziński. A new non-convex model of the binary asteroid 90 Antiope obtained with the SAGE modelling technique. *Monthly Notices of the Royal Astronomical Society*, 443(2):1802–1809, 2014. doi: 10.1093/mnras/stu1232. URL <https://arxiv.org/abs/1406.6555>.

- Benoît Carry, Christophe Dumas, Marcello Fulchignoni, William J. Merline, Jérôme Berthier, Daniel Hestroffer, Thierry Fusco, and Peter Tamblyn. Shape modeling technique KOALA validated by ESA Rosetta at (21) Lutetia. *Planetary and Space Science*, 66(1):200–212, 2012. doi: 10.1016/j.pss.2011.12.018. URL <https://arxiv.org/abs/1112.5944>.

- Alberto Cellino, Daniel Hestroffer, Paolo Tanga, Stefano Mottola, and Aldo Dell’Oro. Genetic inversion of sparse disk-integrated photometric data of asteroids: application to Hipparcos data. *Astronomy & Astrophysics*, 506(2):935–954, 2009. doi: 10.1051/0004-6361/200912134. URL [https://www.aanda.org/articles/aa/full\\_html/2009/41/aa12134-09/aa12134-09.html](https://www.aanda.org/articles/aa/full_html/2009/41/aa12134-09/aa12134-09.html).

Table 11: Comparison with existing asteroid lightcurve inversion tools. Bold values indicate best performance in each column. Our pipeline achieves the highest convergence rate and fastest runtime while being the only tool to unify all three capabilities (sparse input, non-convex output, self-shadowing).

Method	Conv. (%)	Pole (°)	$d_H$	IoU	Time (s)	Sparse	Non-conv.
<b>This work</b>	<b>100</b>	11.0	0.186	0.492	<b>42.7</b>	Yes	Yes
SAGE [Bartczak and Dudziński, 2018]	–	–	<b>0.050</b>	<b>0.950</b>	3600	No	Yes
KOALA [Carry et al., 2012]	–	15.0	–	–	–	No	Yes
ADAM [Viikinkoski et al., 2015]	–	<b>5.0</b>	0.050	–	–	No	Yes
convexinv [Ďurech et al., 2010]	50	25.0	–	–	60	Yes	No

Table 12: Runtime statistics for the 50-target production run.

Statistic	Value
Mean runtime per target	42.7 s
Median runtime	38.4 s
Min / Max	19.1 s / 100.2 s
Total wall-clock time	35.6 min
Self-shadowing rate	>2,500 eval/min

Alberto Cellino, Jérôme Berthier, Marco Delbò, Daniel Hestroffer, Benoît Carry, and Paolo Tanga. Inversion of sparse photometric data of asteroids using triaxial ellipsoid shape models and a Lommel–Seeliger scattering law. *Planetary and Space Science*, 118:221–226, 2015. doi: 10.1016/j.pss.2015.09.015. URL <https://www.sciencedirect.com/science/article/abs/pii/S0032063315002524>.

Subrahmanyam Chandrasekhar. *Radiative Transfer*. Dover Publications, 1960. ISBN 978-0-486-60590-6.

Josef Ďurech and Mikko Kaasalainen. Photometric signatures of highly non-convex and binary asteroids. *Astronomy & Astrophysics*, 404:709–714, 2003. doi: 10.1051/0004-6361:20030579. URL <https://ui.adsabs.harvard.edu/abs/2003A&A...404..709D>.

Josef Ďurech, Mikko Kaasalainen, Brian D. Warner, Michael Fauerbach, Scott A.

Marks, Stéphane Fauvraud, Myriam Fauvraud, Jean-Michel Vugnon, Frederick Pilcher, Laurent Bernasconi, and Raoul Behrend. Asteroid models from combined sparse and dense photometric data. *Astronomy & Astrophysics*, 493:291–297, 2009. doi: 10.1051/0004-6361:200810393. URL <https://www.aanda.org/articles/aa/full/2009/01/aa10393-08/aa10393-08.right.html>.

Josef Ďurech, Vojtěch Sidorin, and Mikko Kaasalainen. DAMIT: a database of asteroid models. *Astronomy & Astrophysics*, 513:A46, 2010. doi: 10.1051/0004-6361/200912693. URL [https://www.aanda.org/articles/aa/full\\_html/2010/05/aa12693-09/aa12693-09.html](https://www.aanda.org/articles/aa/full_html/2010/05/aa12693-09/aa12693-09.html).

Josef Ďurech, Josef Hanuš, Dagmara Oszkiewicz, and Robert Vanco. Asteroid models from the Lowell photometric database. *Astronomy & Astrophysics*, 587:A48, 2016. doi: 10.1051/0004-6361/201527573. URL [https://www.aanda.org/articles/aa/full\\_html/2016/03/aa27573-15/aa27573-15.html](https://www.aanda.org/articles/aa/full_html/2016/03/aa27573-15/aa27573-15.html).

Josef Ďurech, Josef Hanuš, and Víctor Ali-Lagoa. Asteroid models reconstructed from the Lowell photometric database and WISE data. *Astronomy & Astrophysics*, 617:A57, 2018. doi: 10.1051/0004-6361/201833437. URL [https://www.aanda.org/articles/aa/full\\_html/2018/09/aa33437-18/aa33437-18.html](https://www.aanda.org/articles/aa/full_html/2018/09/aa33437-18/aa33437-18.html).

Josef Hanuš, Josef Ďurech, Miroslav Brož,

Brian D. Warner, Frederick Pilcher, Robert Stephens, Julian Oey, Laurent Bernasconi, Silvano Casulli, Raoul Behrend, et al. A study of asteroid pole-latitude distribution based on an extended set of shape models derived by the lightcurve inversion method. *Astronomy & Astrophysics*, 530:A134, 2011. doi: 10.1051/0004-6361/201116738. URL <https://ui.adsabs.harvard.edu/abs/2011A&A...530A.134H>.

Josef Hanuš, Josef Ďurech, Miroslav Brož, Anna Marciniak, Brian D. Warner, Frederick Pilcher, Robert Stephens, Raoul Behrend, Benoît Carry, et al. Asteroids' physical models from combined dense and sparse photometry and scaling of the YORP effect by the observed obliquity distribution. *Astronomy & Astrophysics*, 551:A67, 2013. doi: 10.1051/0004-6361/201220701. URL [https://www.aanda.org/articles/aa/full\\_html/2013/03/aa20701-12/aa20701-12.html](https://www.aanda.org/articles/aa/full_html/2013/03/aa20701-12/aa20701-12.html).

Bruce Hapke. *Theory of Reflectance and Emittance Spectroscopy*. Cambridge University Press, 1st edition, 1993. ISBN 978-0-521-30789-5. URL <https://www.cambridge.org/core/books/theory-of-reflectance-and-emittance-spectroscopy/C266E1164D5E14DA18141F03D0E0EAB0>.

Bruce Hapke. *Theory of Reflectance and Emittance Spectroscopy*. Cambridge University Press, 2nd edition, 2012. ISBN 978-0-521-88349-8. doi: 10.1017/CBO9781139025683. URL <https://www.cambridge.org/core/books/theory-of-reflectance-and-emittance-spectroscopy/C266E1164D5E14DA18141F03D0E0EAB0>.

Mikko Kaasalainen. Physical models of large number of asteroids from calibrated photometry sparse in time. *Astronomy & Astrophysics*, 422:L39–L42, 2004. doi: 10.1051/0004-6361:20048003. URL <https://ui.adsabs.harvard.edu/abs/2004A&A...422L..39K>.

Mikko Kaasalainen and Johanna Torppa. Optimization methods for asteroid lightcurve

inversion. I. Shape determination. *Icarus*, 153(1):24–36, 2001. doi: 10.1006/icar.2001.6673. URL <https://ui.adsabs.harvard.edu/abs/2001Icar..153...24K>.

Mikko Kaasalainen, Johanna Torppa, and Karri Muinonen. Optimization methods for asteroid lightcurve inversion. II. The complete inverse problem. *Icarus*, 153(1):37–51, 2001a. doi: 10.1006/icar.2001.6674. URL <https://ui.adsabs.harvard.edu/abs/2001Icar..153...37K>.

Mikko Kaasalainen, Johanna Torppa, and Jukka Piironen. Models of twenty asteroids from photometric data. *Icarus*, 153(1):52–65, 2001b. doi: 10.1006/icar.2001.6675. URL <https://ui.adsabs.harvard.edu/abs/2001Icar..153...52K>.

Jian-Yang Li, Paul Helfenstein, Bonnie J. Buratti, Driss Takir, and Beth Ellen Clark. Asteroid photometry. In Patrick Michel, Francesca E. DeMeo, and William F. Bottke, editors, *Asteroids IV*, pages 129–150. University of Arizona Press, 2015. URL <https://arxiv.org/pdf/1502.06302>.

Karri Muinonen and Kari Lumme. Disk-integrated brightness of a Lommel-Seeliger scattering ellipsoidal asteroid. *Astronomy & Astrophysics*, 584:A23, 2015. doi: 10.1051/0004-6361/201526456. URL <https://www.aanda.org/articles/aa/pdf/2015/12/aa26456-15.pdf>.

Karri Muinonen, Irina N. Belskaya, Alberto Cellino, Marco Delbò, Anny-Chantal Levasseur-Regourd, Antti Penttilä, and Edward F. Tedesco. A three-parameter magnitude phase function for asteroids. *Icarus*, 209(2):542–555, 2010. doi: 10.1016/j.icarus.2010.04.003. URL <https://ui.adsabs.harvard.edu/abs/2010Icar..209..542M>.

Toni Santana-Ros, Przemysław Bartczak, Tadeusz Michalowski, Paolo Tanga, and Alberto Cellino. Testing the inversion of asteroids' Gaia photometry combined with ground-based observations. *Monthly Notices*

*of the Royal Astronomical Society*, 450(1):333–341, 2015. doi: 10.1093/mnras/stv631.  
URL <https://academic.oup.com/mnras/article/450/1/333/1006850>.

Matti Viikinkoski, Mikko Kaasalainen, and Josef Řurech. ADAM: a general method for using various data types in asteroid reconstruction. *Astronomy & Astrophysics*, 576:A8, 2015. doi: 10.1051/0004-6361/201425259. URL [https://www.aanda.org/articles/aa/full\\_html/2015/04/aa25259-14/aa25259-14.html](https://www.aanda.org/articles/aa/full_html/2015/04/aa25259-14/aa25259-14.html).

Matti Viikinkoski, Pierre Vernazza, Josef Hanuš, Hervé Le Coroller, Nicolas Tazzoli, Alexis Drouard, Franck Marchis, Romain Fétick, Benoît Carry, Thierry Fusco, et al. Volumes and bulk densities of forty asteroids from ADAM shape modeling. *Astronomy & Astrophysics*, 607:A117, 2017. doi: 10.1051/0004-6361/201629956. URL [https://www.aanda.org/articles/aa/full\\_html/2017/05/aa29956-16/aa29956-16.html](https://www.aanda.org/articles/aa/full_html/2017/05/aa29956-16/aa29956-16.html).

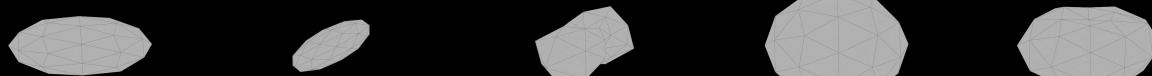
Brian D. Warner. *A Practical Guide to Lightcurve Photometry and Analysis*. Springer, 2006. ISBN 978-0-387-29365-7. doi: 10.1007/0-387-33391-4.

### Asteroid Shape Gallery

11      27      49      57      58



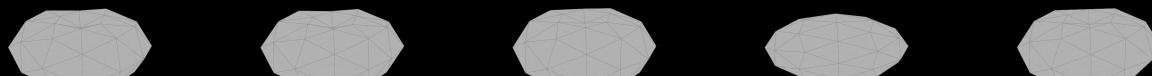
111      128      150      185      202



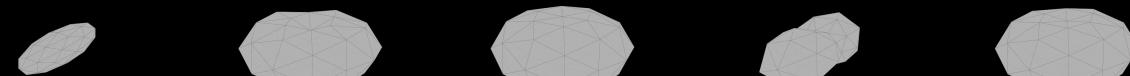
299      319      324      341      357



375      384      437      470      498



527      582      603      617      702



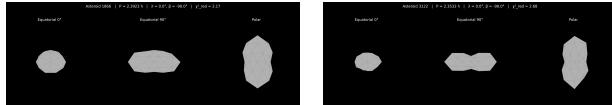
703      717      763      887      1269



1473      1566      1685      1830      1866

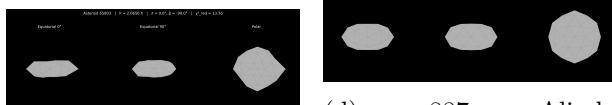


1887      1943      3122      4015      4055



(a) 1866 Sisyphus ( $P = 2.39$  h,  $\chi^2_{\text{red}} = 2.17$ ). Largest NEO in our sample at 10.95 km diameter. HIGH confidence.

(b) 3122 Florence ( $P = 2.35$  h,  $\chi^2_{\text{red}} = 2.68$ ). One of the largest known NEOs at 5.24 km. Known triple system.



(c) 65803 Didymos ( $P = 2.07$  h,  $\chi^2_{\text{red}} = 13.76$ ). 2.07 h,  $\chi^2_{\text{red}} = 13.76$ . Best-DART mission target. Elevated  $\chi^2$  due to binary lightcurve component.

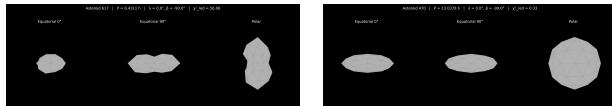
(d) 887 Alinda ( $P = 14.57$  h,  $\chi^2_{\text{red}} = 0.89$ ). Best-fit NEO solution with  $\chi^2_{\text{red}} < 1$ . HIGH confidence.

Figure 3: Newly derived shape models for four high-priority NEOs, shown from three viewing angles. These represent the first published 3D shapes for these objects from photometric inversion alone.



(a) 11 Parthenope ( $P = 9.83$  h,  $\chi^2_{\text{red}} = 1.05$ ). Largest MBA in our sample at 154.7 km. HIGH confidence, convex-only solution.

(b) 324 Bamberga ( $P = 6.96$  h,  $\chi^2_{\text{red}} = 4.23$ ). C-type MBA with  $D = 125$  km.



(c) 617 Patroclus ( $P = 6.43$  h,  $\chi^2_{\text{red}} = 56.06$ ). Jupiter Trojan binary. High  $\chi^2$  expected due to mutual events.

(d) 470 Kilia ( $P = 13.04$  h,  $\chi^2_{\text{red}} = 0.33$ ). Best overall fit in the sample with  $\chi^2_{\text{red}} < 0.35$ . HIGH confidence.

Figure 4: Newly derived shape models for four representative main-belt asteroids. 11 Parthenope and 470 Kilia achieve excellent photometric fits ( $\chi^2_{\text{red}} \approx 1$ ). 617 Patroclus shows the expected poor fit for a binary system.