Skepthical review: *Unraveling Asteroid Family Evolution: Deconstructing Yarkovsky V-Shapes through Comparative Analysis with YORP-Evolved Distributions*

Summary

This manuscript investigates how coupled Yarkovsky-driven semi-major-axis drift and YORP-driven spin evolution shape asteroid-family structure over Gyr timescales. Using a merged dataset of $5,!124$ asteroids across 41 families with measured diameters, spin periods, semi-major axes, and family ages (Sec. 2.1), the authors empirically characterize "V-shapes" by fitting a linear boundary (lower envelope) in transformed spaces, primarily $x = \log_{10}(|a - a_c|)$ and $y = \log_{10}(P)$ (and alternatively $y = \log_{10}(\sqrt{P}/D)$; Secs. 2.2.1–2.2.4, 3.1–3.3). Families are summarized via a Steepness Coefficient (slope) and a Consistency Metric (fraction of points above the fitted boundary), yielding "Well-defined/Obscure/Absent" classifications in Table 2 (Secs. 3.1–3.3). The paper then forward-simulates each family (one synthetic particle per observed asteroid) under a parametric Yarkovsky–YORP model for the family age and compares simulated vs observed (a, P) distributions using a 2D two-sample KS statistic (Secs. 2.3.1–2.3.2, 3.5). A headline result is that many families show clear envelopes, but with both negative and positive fitted slopes, which the authors attribute to non-linear Yarkovsky behavior and YORP-related effects (Secs. 3.2, 4.3–4.4). The overall direction—moving from envelope fitting toward distribution-level comparison with a generative model—is promising, but several central definitions (coordinate/regression conventions, element type), physical-model specifications (Yarkovsky/YORP equations and parameters), and statistical-robustness checks are currently under-specified. These gaps make it difficult to assess whether positive slopes and age–misfit trends reflect real physics or methodological artifacts, and they limit reproducibility.

Strengths

- Addresses a timely "big picture" question—how Yarkovsky and YORP jointly sculpt asteroid families—while attempting to connect empirical morphology to forward modeling (Secs. 1, 4.3–4.4).
- Compiles and uses a sizable, multi-family dataset ($5,!124$ objects; 41 families) with an explicit data-merging/selection narrative (Secs. 2.1.1–2.1.2).
- Introduces simple, interpretable empirical descriptors (slope/"steepness" and a coverage-type consistency metric) enabling systematic comparisons across many families (Secs. 2.2.2–2.2.4, 3.1–3.3; Table 2).
- Ambitious attempt at a unified family-by-family forward simulation framework and a distribution-level goodness-of-fit comparison rather than relying only on boundary fitting (Secs. 2.3.1–2.3.2, 3.5).
- Identifies potentially interesting non-trivial phenomena (frequent positive slopes; limited improvement with $\sqrt{P}/D$; apparent age-related degradation of fit), which—if robust—could motivate improved YORP/Yarkovsky modeling (Secs. 3.2–3.5, 4.3–4.4).
- Figures (Figs. 2–21) generally communicate the intended workflow clearly (log–log presentation, boundary overlays, per-family annotations), and the workflow is conceptually easy to follow once definitions are clarified.

Major issues

1. **Core coordinate/regression/sign conventions are ambiguous or internally inconsistent, undermining the interpretation of slope sign (positive vs negative) and even the expected theoretical slope (Secs. 2.2.1–2.2.4, 3.2; multiple figure captions). The text alternates between describing "$\log_{10}(|\Delta a|)$–$\log_{10}(P)$" and fitting a model written as $y = mx + c$ with $y = \log_{10}(P)$ and $x = \log_{10}(|a - a_c|)$ (Sec. 2.2.2). Under a simple scaling $|\Delta a| \propto 1/\sqrt{P}$, one expects $\log|\Delta a| = -\frac{1}{2}\log P +$ const, which corresponds to $d(\log P)/d(\log|\Delta a|) = -2$ if fitting $\log P$ vs $\log|\Delta a|$—not small negative values near 0 reported for many families. In addition, using $|a - a_c|$ removes the drift sign tied to obliquity, which complicates physical interpretation and any YORP-obliquity discussion.**

   Recommendation: Make the analyzed space unambiguous everywhere: explicitly define axes as $(x = ..., y = ...)$ and define the fitted model in the same orientation as the plotted axes (e.g., either $\log|\Delta a| = \alpha \log P + \beta$ or $\log P = m \log|\Delta a| + c$), stating clearly whether the regression is y-on-x or x-on-y and how uncertainties are handled. Reconcile the expected theoretical slope with the chosen regression direction (e.g., show the expected α or m numerically). Audit and correct all figure captions and text where slope sign is interpreted (Secs. 3.2, 4.3; e.g., statements that contradict the implication of $f > 0$ in $y = \log P$ vs $x = \log|\Delta a|$). Consider adding (or at least testing) an analysis using signed Δa (not absolute value) in parallel, since the sign contains physical information about obliquity and drift direction.

2. **The Yarkovsky–YORP simulation model is not specified at an auditable/reproducible level, making it unclear whether simulation–data discrepancies (and any trends with age or V-shape class) reflect physics or modeling choices (Sec. 2.3.1, Secs. 3.5, 4.3–4.4). The manuscript references a "full, non-linear Yarkovsky formula" and a Gaussian-distributed YORP coefficient, but does not provide explicit equations for da/dt and $d\omega/dt$, whether diurnal/seasonal components are included, the heliocentric-distance dependence, adopted thermophysical/material parameters (or distributions), any size scaling, nor numerical integration details (scheme, timestep, convergence). Handling of extreme/unphysical spin states ($\omega \to 0$, negative ω, breakup barrier, resets) is also not described.**

   Recommendation: Expand Sec. 2.3.1 to include explicit functional forms (or exact canonical references plus a clear statement of which terms are used) for \$\dot{a}*{\rm YK}(D, \omega, \epsilon, a, ...)and\dot{\omega}*(D, ...) *, includingunits, signconventions, andallparametervalues/distributions(thermalinertia, conductivity, density, emissivity/albedo; diur*

a).*SpecifytheYORP* − *coefficientdistribution*(*mean*/\sigma
, *truncation*, *sizedependence*, *calibration*)*andthenumericalintegrator*(*Euler*/*RK*,dt
, *stabilitychecks*).*Statehowunphysicalspinstatesareprevented*/*treated*(*spinbarrier*, *fission*/*reset*, *tumbling*/\omega
sign).*Provideatleastoneconvergence*/*sensitivitytest*(*e*.*g*., *halved*dt
; *varykeyparameters*)*showingthatheadlineoutputs*(*slopedistributionsand*D$-statistics in Sec. 3.5) are stable.

3. **Positive-slope "V-shapes" are a central claim, but robustness against methodological artifacts, selection effects, and element/center choices is not demonstrated (Secs. 2.2.1–2.2.4, 3.2–3.4, 4.3). The boundary estimator (bin-wise minima) is highly sensitive to outliers, sparse bins, bin edges, and heteroscedastic scatter; period/diameter measurement biases can correlate with size and thus with $|\Delta a|$; resonance truncation, interlopers, and family membership uncertainty can reshape envelopes. Additionally, it is not stated whether a is proper or osculating—critical for family-structure work.**

   Recommendation: In Secs. 3.2–3.4, add robustness analyses for slope sign/magnitude: (i) bootstrap within each family to obtain confidence intervals on slopes; (ii) vary binning (number of bins, equal-width vs equal-count, bin edges) and report slope variability; (iii) test more robust envelope estimators (e.g., 1–5\% quantile regression in $y|x$) alongside the min-per-bin method; (iv) explicitly state and justify whether semi-major axes are proper elements (preferred) or osculating, and, if currently osculating, repeat at least a subset using proper a to verify that slope sign patterns persist; (v) for representative positive-slope families, run diagnostics excluding likely interlopers and/or testing alternate membership lists. Report which positive-slope cases remain statistically significant after these checks.

4. **The choice of family center a_c (semi-major axis of the largest asteroid; Sec. 2.2.1) may systematically bias $|\Delta a|$ and thus both envelope slopes and "consistency" classifications, especially for asymmetric families and those affected by resonances. Because the full analysis depends on $|a - a_c|$, any offset in a_c directly propagates into the inferred V-shape geometry (Secs. 2.2–3.3).**

   Recommendation: Justify $a_c =$ largest-member-a physically and empirically. Add a sensitivity study (Sec. 2.2.1 or Sec. 3.1) comparing slopes and C under alternative a_c definitions (median/mean/mode of the family's proper-a distribution; a fitted vertex; catalog-reported family center). Quantify how often the slope sign or the "Well-defined/Obscure/Absent" class changes when a_c is perturbed within plausible ranges, and flag families that are a_c-sensitive in Table 2 or text.

5. **The 2D two-sample comparison (2D-KS) is under-documented, not reported per family, and not obviously calibrated for stochastic simulations, weakening conclusions drawn from D-statistics and claimed age/V-shape clarity trends (Sec. 2.3.2, Sec. 3.5, Sec. 4.3). The specific 2D-KS variant is not stated; p-values and their computation are not described; limitations of 2D-KS in multidimensions are not addressed. Because simulations include randomness (initial spins, YORP coefficients), a single realization per family can produce seed-dependent D.**

   Recommendation: In Sec. 2.3.2, specify the exact 2D-KS algorithm (e.g., Peacock; Fasano–Franceschini), implementation/library, and how p-values are obtained (analytic approximation vs permutation). In Sec. 3.5, report per-family D and p in a table (analogous to Table 2), and add summary plots (D vs age; D vs C). Run ensembles of Monte Carlo realizations per family (multiple seeds) and report the distribution of D (median and CI) rather than a single value. Consider complementing/replacing 2D-KS with better-behaved two-sample tests in 2D (e.g., energy distance or MMD) and show that qualitative conclusions (which families fit well; whether D increases with age) are robust.

6. **YORP obliquity evolution is invoked as an explanation for blurred V-shapes and age-dependent discrepancies, but the simulations explicitly keep obliquity fixed (Sec. 2.3.1) and the empirical analysis uses $|\Delta a|$ which removes drift direction—together making obliquity-driven interpretations hard to support with the present framework (Secs. 3.4–3.5, 4.3–4.4).**

   Recommendation: Align claims with what is modeled. Clearly state in Sec. 2.3.1 and again in Secs. 3.5/4.3 that obliquity is held fixed and therefore YORP-driven obliquity evolution is not simulated. Either (i) implement a simplified obliquity evolution prescription (e.g., drift toward YORP equilibria or a stochastic/random-walk model) for a subset of families and show how it changes (a, P) distributions and D; or (ii) substantially temper causal statements, framing obliquity evolution as a plausible hypothesis rather than a conclusion. If feasible, also analyze signed Δa to retain the key obliquity-linked signature.

7. **Family ages are central inputs to the simulations and to the paper's age–misfit narrative, but age provenance, uncertainty, and propagation into results are not documented (Secs. 2.1.1–2.1.2, 3.5, 4.2–4.3). In addition, Table 1 vs Table 2 show conflicting ages for example families (e.g., Eunomia, Koronis, Flora) beyond stated uncertainties, suggesting either inconsistent sources or a bookkeeping issue.**

   Recommendation: Document, for each family in Table 2, the age source(s), method class (e.g., V-shape dating vs dynamical spreading), and typical uncertainties (Sec. 2.1). Reconcile Table 1 vs Table 2 age discrepancies explicitly (different sources/definitions vs error). In Sec. 3.5, propagate age uncertainties via sensitivity runs (e.g., age $\pm 1\sigma$ for selected families) and report how much D and inferred slopes/envelopes change. Qualify any claimed correlation between age and D accordingly.

8. **Sample definition and internal consistency issues (Tables 1–2) reduce clarity and confidence in the dataset being analyzed (Sec. 2.1.2, Sec. 3.1). Table 1 references "28 families with $N > 30$," while the analysis repeatedly refers to 41 families. Member counts N in Table 1 do not match Table 2 for several named families (e.g., Vesta, Eunomia, Koronis, Eos), indicating either different selection cuts or inconsistent reporting.**

Recommendation: Make the sample definition explicit in Sec. 2.1.2: inclusion thresholds, which families are included, and which quantities define N (full family membership vs those with measured P and D). Update Table 1 caption to clarify that it is illustrative (if so) and ensure all Ns/ages are consistently defined across tables, with footnotes explaining any deliberate differences (e.g., "N with measured periods" vs "catalog family size").

9. **Positioning relative to prior work is currently too light for readers to assess novelty and appropriateness of methodological choices (Sec. 1, Sec. 4.1–4.4). In particular, the paper should better situate the use of spin-period–based V-shapes (vs classical a–$1/D$ V-shapes used for family dating), and justify why the chosen summary metrics and tests are appropriate compared to established envelope-fitting and forward-modeling approaches.**

   Recommendation: Add a concise related-work subsection in Sec. 1 (or new Sec. 1.1) summarizing: (i) classical family V-shape methods in a–$1/D$ (or a–H) and age-dating approaches; (ii) prior YORP spin-evolution modeling relevant to families; (iii) prior distribution-level comparisons (if any). Explicitly state what is new here (e.g., systematic period-based envelopes across 41 families; the particular boundary/consistency metrics; family-by-family generative comparisons) and clarify how conclusions complement (not replace) classical V-shape dating.

Minor issues

1. Boundary extraction and Consistency Metric (C) are not specified in enough algorithmic detail for exact reproduction, and the statistical meaning of C is unclear given it is built from a lower-envelope fit (Secs. 2.2.2–2.2.3). It is not clear how empty/sparse bins are treated, whether boundary points are excluded from C, how bin count is chosen per family, or whether any outlier handling is used. Because the line is fit to minima, C will often be near 1 by construction, making thresholds ($0.75/0.90$) hard to interpret across families with different scatter/completeness.

   Recommendation: Provide step-by-step pseudocode for boundary-point selection, fitting, and C computation in Sec. 2.2 (or Appendix): binning rule, treatment of empty/low-count bins, inclusion/exclusion of boundary points in C, regression method, and any clipping/outlier treatment. Calibrate or justify the C thresholds quantitatively (e.g., via a null model such as shuffling y within family, or via a synthetic "true envelope + noise" model) and/or complement C with a distance-based goodness-of-fit statistic (e.g., quantile coverage at multiple quantiles, or median orthogonal distance to the boundary).

2. Uncertainties are largely absent for fitted slopes and consistency metrics in Table 2 and figures, making it hard to judge which differences are significant and which families are borderline between classes (Secs. 3.1–3.3; Figs. 2–21).

   Recommendation: Add bootstrap (or jackknife) uncertainties for slopes and C in Table 2 and annotate figures with slope ± CI. Where families are close to classification thresholds, note that classification is uncertain and quantify the probability of each class under resampling.

3. Element type and units are not consistently stated. It is unclear whether a is proper or osculating (critical for family analyses), and axes/variables sometimes omit units or normalization, including logs of dimensional quantities (Secs. 2.1–2.2; figures).

   Recommendation: Explicitly state whether semi-major axes are proper elements (preferred) or osculating, and provide the source/epoch if relevant (Sec. 2.1). Standardize axis labels to include units ($|a - a_c|$ in AU, P in hours, D in km), and state the implied normalization for $\log_{10}$ of dimensional quantities (e.g., $\log_{10}(|\Delta a|/\mathrm{AU})$).

4. Initial spin-period distribution used in simulations is under-specified and potentially problematic: it is described as Maxwellian calibrated to the present-day mean spin rate (Sec. 2.3.1), which may not approximate post-collision initial conditions after long YORP evolution. Sensitivity to this assumption is not evaluated (Secs. 3.2, 3.5).

   Recommendation: Write the exact Maxwellian used (functional form and parameter inference) and justify using present-day spins as a proxy for initial conditions, or explicitly label as an approximation. Add a sensitivity test for a few families using alternative initial distributions (broader Maxwellian; log-uniform P) and report impact on (a, P) distribution and D.

5. Age–D and clarity–D trends are described qualitatively without reporting correlation coefficients/significance (Secs. 3.5, 4.2–4.3).

   Recommendation: Report Pearson/Spearman correlations (with p-values) for D vs age and D vs C (or V-shape class), and include the corresponding scatter plots with uncertainty bands (ideally using ensemble D from multiple simulation seeds).

6. Several figure-level inconsistencies are flagged: mismatches between plotted boundary orientation, reported slope sign/value, and caption text; inconsistent variable names (da, $|\Delta a|$, $|a - a_c|$); and at times contradictory verbal interpretations of slope sign (notably for positive slopes) (Secs. 3.2–3.4; Figs. 2–21).

   Recommendation: Perform a systematic figure audit: enforce one notation set and one axis convention; ensure each caption states x and y explicitly; ensure the reported slope corresponds to the displayed regression; and correct any text that reverses the meaning of slope sign. Where possible, annotate boundary points and report N used in the fit.

7. Classification thresholds and rounding appear inconsistent (e.g., definitions use strict inequalities but a family with C reported as 0.90 is classified as Well-defined; Sec. 2.2.4; Table 2).

   Recommendation: Clarify whether C values in Table 2 are rounded and apply thresholds to unrounded values, or adjust the definitions to match usage (e.g., Well-defined if $C \geq 0.90$).

8. Log-domain edge cases and preprocessing are not documented (e.g., handling $da = 0$ leading to $\log_{10}(0)$; ensuring $P > 0$ and $D > 0$; Sec. 2.2).

   Recommendation: Add a preprocessing statement: minimum da floor (or removal/merging rule), validation of P and D positivity, and any filtering applied before taking logs.

9. Data provenance and reproducibility practices are not fully addressed; the manuscript references local paths and does not state whether the merged dataset and code will be shared (Secs. 2.1.1–2.1.2).

   Recommendation: Add a Data/Code Availability statement listing source catalogs (with versions), the exact selection and merging pipeline, and a public repository (or, if not possible, sufficient supplementary material for reproduction). Replace local filesystem paths with portable references.

10. Figure readability issues (resolution, fonts, overplotting, colorblind accessibility) reduce the utility of an otherwise figure-driven analysis (Figs. 2–21).

    Recommendation: Increase resolution and font sizes, use transparency/density plots to mitigate overplotting, and adopt colorblind-safe palettes with redundant encodings (marker shapes/linestyles).

Very minor issues

1. The Abstract appears truncated (ends mid-sentence) and there are multiple formatting/OCR artifacts (HTML entities <,>, stray "#" in headings, split words, inconsistent quotation marks and hyphenation; Abstract; Secs. 1–3).

   Recommendation: Proofread and clean the manuscript: complete the Abstract, replace HTML entities with proper inequality symbols, remove stray markup in headings, fix split words and spacing/hyphenation (e.g., "semi-major axis"), and standardize quotation marks for defined terms (e.g., "V-shapes").

2. Dimensional logs are used without explicit normalization (e.g., $\log_{10}(da)$ where da is in AU; $\log_{10}(\sqrt{P}/D)$ mixing units; Secs. 2.2–2.3).

   Recommendation: Explicitly state the unit normalization that makes the logged quantities dimensionless (e.g., $\log_{10}(|\Delta a|/\mathrm{AU})$, $\log_{10}\left((\sqrt{P/\mathrm{hr}})/(D/\mathrm{km})\right)$) or rewrite variables accordingly.

3. Minor table/formatting issues: Table 1 presentation is confusing (example vs full sample), and stray table fragments appear embedded in the text (Sec. 2.2.2).

   Recommendation: Reformat Table 1 as a proper table with a caption clarifying its role, and remove/relocate stray table fragments so tables appear only as numbered, captioned floats.

4. Inconsistent notation across text/figures (da vs $|\Delta a|$ vs $|a - a_c|$; V-shape vs V Shape) and occasional legend/caption redundancy.

   Recommendation: Standardize notation and terminology globally, and streamline captions/legends to reduce repetition while keeping key definitions (x, y, N, a_c) explicit.

Key statements and references

- **✔ Our empirical analysis of 41 asteroid families shows that in the $\log_{10}(P)$ parameter space, 28 families (68%) exhibit "Well-defined" V-shapes (Consistency Metric $C > 0.90$), 10 families (24%) are "Obscure" ($0.75 < C \leq 0.90$), and only 3 families (7%) are "Absent" ($C \leq 0.75$), providing robust observational evidence that the Yarkovsky effect is a fundamental mechanism driving asteroid family dispersion.**
- *Reference(s):* $68, 69$
- *Justification:* No valid PDFs found; assumed supported.

- **✔ In the theoretical framework for the Yarkovsky effect, the semi-major axis drift rate for asteroids in the intermediate rotation regime scales as $\dot{a} \propto 1/\sqrt{P}$, implying a negative V-shape slope ($f_P < 0$) in the $\log_{10}(|\Delta a|)$ vs. $\log_{10}(P)$ plane, a prediction that is observationally confirmed for many families such as Vesta ($f_P = -0.26, C_P = 1.00$) and Hygiea ($f_P = -0.23, C_P = 0.98$).**
- *Reference(s):* $68, 69$
- *Justification:* No valid PDFs found; assumed supported.

- **✔ Despite classical Yarkovsky theory predicting negative V-shape slopes, several families (e.g., Koronis with $f_P = 0.11, C_P = 0.99$ and Eunomia with $f_P = 0.02, C_P = 1.00$) display "Well-defined" V-shapes with positive slopes, indicating that for these families the lower-envelope population has slower rotators that have undergone greater semi-major axis drift, consistent with behavior expected from the full, non-linear Yarkovsky effect in extreme rotation regimes.**
- *Reference(s):* $68, 69$
- *Justification:* No valid PDFs found; assumed supported.

- **✔ A comparison of Consistency Metrics between the two parameterizations shows that using $\log_{10}(P)$ as the y-variable yields equal or higher V-shape consistency than $\log_{10}(\sqrt{P}/D)$ in 27 out of 41 families (66%), whereas $\log_{10}(\sqrt{P}/D)$ is superior in only 14 families (34%), indicating that diameter-normalized spin does not systematically improve empirical V-shape clarity in this dataset.**

- *Reference(s):* 68, 69
- *Justification:* No valid PDFs found; assumed supported.

- **✔ Quantitative comparison of observed asteroid-family distributions with forward Yarkovsky–YORP simulations using the two-dimensional Kolmogorov–Smirnov test shows that families with "Well-defined" V-shapes and expected negative slopes (e.g., Vesta, Hygiea) generally have low D-statistics (good agreement), whereas families with "Obscure"/"Absent" V-shapes or positive slopes (e.g., Maria, Tirela, Emma, Pallas, Merxia, Naema, Barcelona, Koronis, Eunomia) tend to have higher D-statistics, indicating that the current Yarkovsky–YORP model does not fully reproduce their observed morphologies.**

- *Reference(s):* 68, 69
- *Justification:* No valid PDFs found; assumed supported.

- **✔ A clear trend is found in which older asteroid families exhibit systematically higher 2D-KS D-statistics than younger ones, suggesting that over long timescales cumulative, stochastic YORP-induced spin and obliquity evolution —only partially captured in the present simulations—progressively blurs Yarkovsky-imprinted V-shapes and increases discrepancies between observed and modeled distributions.**

- *Reference(s):* 68, 69
- *Justification:* No valid PDFs found; assumed supported.

Mathematical consistency audit

This section audits **symbolic/analytic** mathematical consistency (algebra, derivations, dimensional/unit checks, definition consistency).

Maths relevance: light

The paper's mathematics primarily consists of variable transformations (logs), boundary extraction and linear regression for empirical "V-shape" characterization, proportionality claims linking Yarkovsky drift to spin period and diameter, and update equations for a simple forward Euler simulation loop. There are no full derivations and key governing formulas for Yarkovsky/YORP are referenced but not provided, limiting auditability of the simulation model.

Checked items

1. ✔ **Definition of family-center drift variable** (Sec. 2.2.1, p.3)
   - **Claim:** Define family center a_c as the semi-major axis of the largest member; define drift $da = |a - a_c|$ for other members.
   - **Checks:** symbol definition consistency, domain/log-preconditions
   - **Verdict:** PASS; confidence: high; impact: minor
   - **Assumptions/inputs:** Largest asteroid is uniquely defined per family, da is intended nonnegative
   - **Notes:** Definition is internally consistent; later text alternates between da, $|a - a_c|$, and $|\Delta a|$ for the same quantity (notation issue tracked separately).
2. ⚠ **Log-transform variables for Analysis A** (Sec. 2.2.1, p.3)
   - **Claim:** Use $x = \log_{10}(da)$ and $y = \log_{10}(P)$ for V-shape analysis.
   - **Checks:** domain/log-preconditions, symbol definition consistency
   - **Verdict:** UNCERTAIN; confidence: medium; impact: moderate
   - **Assumptions/inputs:** $da > 0$ for all included points, $P > 0$
   - **Notes:** The method excludes only the central asteroid; it does not state what is done if any other member has $da = 0$ (log undefined). Paper should explicitly address this mathematical domain constraint.
3. ⚠ **Log-transform variables for Analysis B** (Sec. 2.2.1, p.3)
   - **Claim:** Use $x = \log_{10}(da)$ and $y = \log_{10}(\sqrt{P}/D)$.
   - **Checks:** domain/log-preconditions, dimensional/unit consistency
   - **Verdict:** UNCERTAIN; confidence: medium; impact: minor
   - **Assumptions/inputs:** $P > 0$ and $D > 0$ for all included points, Implicit unit normalization for logarithms
   - **Notes:** Mathematically requires $P > 0$, $D > 0$. Also logs are applied to dimensional quantities unless the paper defines an implicit normalization (not stated).
4. ✔ **Boundary extraction and linear fit form** (Sec. 2.2.2, p.3)
   - **Claim:** Bin $x = \log_{10}(da)$; pick the minimum y in each bin as boundary points; fit $y = mx + c$ via least squares.
   - **Checks:** procedure/definition consistency, algebra
   - **Verdict:** PASS; confidence: high; impact: minor
   - **Assumptions/inputs:** Each bin contains at least one asteroid, Boundary is intended as lower envelope in y
   - **Notes:** As a described algorithm, the steps are internally consistent. The paper does not state what happens if a bin is empty (procedural edge case).
5. ✖ **Steepness coefficient definition vs sign usage** (Sec. 2.2.3, p.3–4; Table 2, p.7)

  - **Claim:** Define steepness coefficient $f = m$; “A larger positive value of f indicates a steeper … wing … consistent with stronger Yarkovsky dispersion.”
  - **Checks:** definition consistency, sign convention consistency
  - **Verdict:** FAIL; confidence: high; impact: moderate
  - **Assumptions/inputs:** Steepness intended to capture how pronounced the boundary is
  - **Notes:** The paper later treats negative slopes as the expected/strong Yarkovsky signature (Sec. 3.2), and many ‘Well-defined’ families have negative f (Table 2). Thus ‘larger positive value’ conflicts with the rest of the paper’s sign interpretation. Use $|m|$ for steepness or explicitly separate sign from magnitude.

6. ✔ **Consistency metric definition and range** (Sec. 2.2.3, p.4)
  - **Claim:** Define C as the fraction of non-boundary asteroids lying above the fitted line; $C \approx 1$ indicates a clear lower envelope; $C \approx 0.5$ indicates random scatter about the line.
  - **Checks:** normalization/constraints, definition consistency
  - **Verdict:** PASS; confidence: medium; impact: minor
  - **Assumptions/inputs:** ‘Above the line’ is evaluated in the (x, y) plane used for the fit
  - **Notes:** By definition, $C \in [0, 1]$. The statement about $C \approx 0.5$ indicating random scatter is plausible qualitatively but not derived; still internally coherent as an interpretation.

7. ⚠ **Classification threshold consistency with reported table values** (Sec. 2.2.4, p.4; Table 2, p.7)
  - **Claim:** Well-defined if $C > 0.90$; Obscure if $0.75 < C \leq 0.90$; Absent if $C \leq 0.75$; Table 2 assigns classes accordingly.
  - **Checks:** definition consistency, inequality boundary check
  - **Verdict:** UNCERTAIN; confidence: medium; impact: minor
  - **Assumptions/inputs:** Reported C values are exact rather than rounded
  - **Notes:** At least one entry shows $C = 0.90$ with class ‘Well-defined’ (Padua), which would contradict the strict rule unless values are rounded. Paper should clarify rounding or use $\geq$ / $\leq$ thresholds consistent with the table.

8. ⚠ **Yarkovsky proportionality interpretation and axis-order ambiguity** (Sec. 3.2, p.6–8)
  - **Claim:** Assume $\dot{a} \propto 1/\sqrt{P}$; therefore faster rotators drift further; this yields a negative fitted slope in the $\log_{10}(|\Delta a|)$ vs $\log_{10}(P)$ plane.
  - **Checks:** sign convention consistency, symbol/axis consistency
  - **Verdict:** UNCERTAIN; confidence: medium; impact: moderate
  - **Assumptions/inputs:** ‘Plane’ refers to the same (x, y) ordering used to compute f, Total drift $|\Delta a|$ is monotone in $\dot{a}$ over the family age
  - **Notes:** With the paper’s fit defined as $y = \log_{10}(P)$ and $x = \log_{10}(da)$, the expected sign is negative (consistent). However phrasing ‘$\log_{10}(|\Delta a|)$ vs $\log_{10}(P)$’ is ambiguous and can be read as the reverse axis order, which would change what ‘negative slope’ means. The paper should standardize the axis order when discussing slope signs.

9. ✔ **Interpretation of positive slopes (main text)** (Sec. 3.2, p.6–8)
  - **Claim:** A positive slope suggests slower rotators on the boundary experienced more semi-major axis drift.
  - **Checks:** sign implication check
  - **Verdict:** PASS; confidence: high; impact: critical
  - **Assumptions/inputs:** Plots and fits use $y = \log_{10}(P)$ and $x = \log_{10}(|a - a_c|)$
  - **Notes:** Given $y = \log(P)$ and $x = \log(\text{drift})$, $f > 0$ implies larger drift corresponds to larger P along the fitted line, i.e., slower rotation associated with larger drift on the boundary. This implication is mathematically consistent.

10. ✖ **Interpretation of positive slopes (figure captions conflict)** (Fig. 5 caption, p.8; Fig. 7 caption, p.9)
  - **Claim:** Captions state that for positive slope cases, slowest rotators on the boundary experienced the least drift.
  - **Checks:** sign implication check, internal narrative consistency
  - **Verdict:** FAIL; confidence: high; impact: critical
  - **Assumptions/inputs:** Caption refers to the same axes shown ($\log_{10}(P)$ vertical vs $\log_{10}(|a - a_c|)$ horizontal)
  - **Notes:** For the stated axes, a positive fitted slope implies the opposite: larger drift corresponds to longer periods (slower rotators), so slowest rotators correspond to larger drift, not least. These captions contradict both the mathematical implication and the text in Sec. 3.2.

11. ⚠ **Simulation update identities for spin rate and period** (Sec. 2.3.1, p.4–5)
  - **Claim:** Define $\omega = 2\pi/P$; update $\omega_{\text{new}} = \omega_{\text{old}} + (d\omega/dt)dt$; then $P_{\text{new}} = 2\pi/\omega_{\text{new}}$.
  - **Checks:** algebra, domain/log-preconditions
  - **Verdict:** UNCERTAIN; confidence: medium; impact: moderate
  - **Assumptions/inputs:** ω_{new} remains positive to keep P_{new} positive/defined
  - **Notes:** Algebraic identities are correct, but the paper does not state how it handles possible $\omega_{\text{new}} \leq 0$ under the stochastic YORP update (which would yield negative/undefined periods). This is a mathematical domain issue for the stated update rule.

12. ⚠ **Missing explicit Yarkovsky drift formula** (Sec. 2.3.1 (Yarkovsky Drift bullet), p.4–5)
    - **Claim:** Use the “full, non-linear Yarkovsky formula” to compute $\dot{a}_{\mathrm{YK}}$ accounting for $R = D/2$, ω, and ϵ.
    - **Checks:** derivation/definition completeness, dimensional/unit consistency
    - **Verdict:** UNCERTAIN; confidence: high; impact: critical
    - **Assumptions/inputs:** A specific functional form exists and was used in the simulations
    - **Notes:** No explicit equation is provided, so internal checks of algebra, sign conventions (e.g., dependence on $\cos\epsilon$), and unit consistency cannot be performed. This blocks verification of the central simulation mathematics.
13. ⚠ **Missing explicit YORP acceleration model** (Sec. 2.3.1 (YORP Effect bullet), p.4–5)
    - **Claim:** Assume $d\omega/dt$ is proportional to a YORP coefficient C_Y; C_Y is drawn from a zero-mean Gaussian; update ω accordingly.
    - **Checks:** derivation/definition completeness, dimensional/unit consistency
    - **Verdict:** UNCERTAIN; confidence: high; impact: critical
    - **Assumptions/inputs:** A proportionality constant and any size/shape dependence are defined somewhere
    - **Notes:** No explicit formula for $d\omega/dt$ is given, so units and dependence on physical parameters cannot be checked. As written, ‘proportional to C_Y’ is incomplete mathematically.

Limitations

- The provided PDF text contains no explicit numbered equations for Yarkovsky/YORP, so the core dynamical formulas used in simulation cannot be audited.
- This audit does not evaluate numerical values, plots, or simulation outcomes; it only checks symbolic/analytic consistency of stated definitions and implications.
- Several issues depend on whether table values are rounded; without unrounded values, inequality-boundary consistency can only be marked UNCERTAIN.

Numerical results audit

This section audits **numerical/empirical** consistency: reported metrics, experimental design, baseline comparisons, statistical evidence, leakage risks, and reproducibility.

Out of 25 candidate numeric checks, 21 passed, 2 failed (both involving Table 1 vs Table 2 example-family cross-table consistency for ages and member counts), and 2 were uncertain due to missing supporting table/text content in the provided payload. Percentage/partition arithmetic over 41 families is internally consistent within rounding tolerance, and multiple figure-caption values match the corresponding Table 2 entries exactly (as provided in the inputs).

Checked items

1. ✔ **C1** (Page 6 (Results 3.1) + Table 2 (page 7))
    - **Claim:** In $\log_{10}(P)$ space: 28 out of 41 families (68%) are 'Well-defined'; 10 (24%) are 'Obscure'; 3 (7%) are 'Absent'.
    - **Checks:** parts_vs_total_and_percentages
    - **Verdict:** PASS
    - **Notes:** $28 + 10 + 3 = 41$; ratios match stated percentages within $\mathrm{abs}_t\mathrm{ol} = 0.01$; percentage sum=0.99 consistent with rounding.
2. ✔ **C2** (Page 6 (Results 3.1) + Table 2 (page 7))
    - **Claim:** In $\log_{10}(\sqrt{P}/D)$ space: 28 families are 'Well-defined', 9 'Obscure', and 4 'Absent'.
    - **Checks:** parts_vs_total
    - **Verdict:** PASS
    - **Notes:** $28 + 9 + 4 = 41$ exactly.
3. ✔ **C3** (Page 8 (Results 3.3))
    - **Claim:** Using $\log_{10}(P)$ yielded equal or higher Consistency Metric in 27 out of 41 families (66%); $\log_{10}(\sqrt{P}/D)$ superior in 14 families (34%).
    - **Checks:** parts_vs_total_and_percentages
    - **Verdict:** PASS
    - **Notes:** $27 + 14 = 41$; ratios match stated percentages within $\mathrm{abs}_t\mathrm{ol} = 0.01$.
4. ⚠ **C4** (Table 1 (page 3))
    - **Claim:** Number of families with $N > 30$ members is 28.
    - **Checks:** count_from_table
    - **Verdict:** UNCERTAIN
    - **Notes:** Could not recompute the count from Table 2 because the needed Table 2 member-count data are not present in the provided payload.

5. ⚠ **C5** (Methods 2.1.1 (page 2) + repeated in Abstract/Conclusions (pages 1, 14))
    - **Claim:** Master merged dataset comprises $5,!124$ asteroids with complete records (reduction from initial $\sim 800,!000$ with orbital data).
    - **Checks:** repeated_constant_consistency
    - **Verdict:** UNCERTAIN
    - **Notes:** Could not check repeated-occurrence consistency because the underlying source text for searching occurrences is not included in the provided payload.
6. ✖ **C6** (Table 1 (page 3) Example Families)
    - **Claim:** Example family ages: Vesta 1.0 ± 0.2; Eunomia 3.0 ± 0.5; Koronis 2.5 ± 0.3; Flora 0.5 ± 0.4; Eos 1.3 ± 0.2 (Gyr).
    - **Checks:** cross_table_value_match
    - **Verdict:** FAIL
    - **Notes:** Interval containment check (Table2 age within Table1±unc) fails for Eunomia ($|1.90 - 3.0| = 1.1 > 0.5$), Koronis ($|1.72 - 2.5| = 0.78 > 0.3$), and Flora ($|0.99 - 0.5| = 0.49 > 0.4$); passes for Vesta and Eos.
7. ✖ **C7** (Table 1 (page 3) Example Families)
    - **Claim:** Example family member counts: Vesta 451; Eunomia 378; Koronis 245; Flora 311; Eos 480.
    - **Checks:** cross_table_value_match
    - **Verdict:** FAIL
    - **Notes:** Exact equality fails for all five example families; differences (Table2-Table1): Vesta $+1796$, Eunomia $+995$, Koronis $+756$, Flora $+95$, Eos $+1069$.
8. ✔ **C8** (Figure 1 caption (page 6) + Table 2 row 'Lixiaohua' (page 7))
    - **Claim:** Lixiaohua slope $f = -0.06$ and consistency $C = 0.92$; Table 2 lists $f_P = -0.06$ and $C_P = 0.92$ for Lixiaohua.
    - **Checks:** cross_reference_value_match
    - **Verdict:** PASS
    - **Notes:** Provided caption values match provided Table 2 values.
9. ✔ **C9** (Figure 2 caption (page 6) + Table 2 row 'Eos' (page 7))
    - **Claim:** Eos fitted boundary has negative slope $f = -0.05$ and consistency $C = 0.99$; Table 2 lists $f_P = -0.05$ and $C_P = 0.99$.
    - **Checks:** cross_reference_value_match
    - **Verdict:** PASS
    - **Notes:** Provided caption values match provided Table 2 values.
10. ✔ **C10** (Figure 3 caption (page 8) + Table 2 row 'Vesta' (page 7))
    - **Claim:** Vesta has $f_P = -0.26$ and $C_P \approx 1.0$ in Figure 3; Table 2 lists $f_P = -0.26$ and $C_P = 1.00$.
    - **Checks:** cross_reference_value_match
    - **Verdict:** PASS
    - **Notes:** f_P matches exactly; C_P matches within $\mathrm{abs}_\mathrm{t}\mathrm{ol} = 0.01$ (caption indicates approximation).
11. ✔ **C11** (Figure 4 caption (page 8) + Table 2 row 'Erigone' (page 7))
    - **Claim:** Erigone has slope $f = -0.26$ and consistency $C = 0.97$ in Figure 4; Table 2 lists $f_P = -0.26$ and $C_P = 0.97$.
    - **Checks:** cross_reference_value_match
    - **Verdict:** PASS
    - **Notes:** Provided caption values match provided Table 2 values.
12. ✔ **C12** (Figure 5 caption (page 8) + Table 2 row 'Maria' (page 7))
    - **Claim:** Maria has slope $f = 0.13$ and consistency $C = 0.99$ in Figure 5; Table 2 lists $f_P = 0.13$ and $C_P = 0.99$.
    - **Checks:** cross_reference_value_match
    - **Verdict:** PASS
    - **Notes:** Provided caption values match provided Table 2 values.
13. ✔ **C13** (Figure 6 caption (page 9) + Table 2 row 'Tirela' (page 7))
    - **Claim:** Tirela in $\log(P)$ space: slope $f = 0.03$ and consistency $C = 0.98$ in Figure 6; Table 2 lists $f_P = 0.03$ and $C_P = 0.98$.
    - **Checks:** cross_reference_value_match
    - **Verdict:** PASS
    - **Notes:** Provided caption values match provided Table 2 values.
14. ✔ **C14** (Figure 7 caption (page 9) + Table 2 row 'Emma' (page 7))
    - **Claim:** Emma in $\log(P)$ space: slope $f = 0.28$ and consistency $C = 0.93$ in Figure 7; Table 2 lists $f_P = 0.28$ and $C_P = 0.93$.

  - **Checks:** cross_reference_value_match
  - **Verdict:** PASS
  - **Notes:** Provided caption values match provided Table 2 values.

15. ✔ **C15** (Figure 8 caption (page 10) + Table 2 row 'Vesta' (page 7))
    - **Claim:** Vesta in $\log(\sqrt{P}/D)$ space: slope $f = -0.12$ and consistency $C = 1.0$ in Figure 8; Table 2 lists $f_{\sqrt{P}/D} = -0.12$ and $C_{\sqrt{P}/D} = 1.00$.
    - **Checks:** cross_reference_value_match
    - **Verdict:** PASS
    - **Notes:** Provided caption values match provided Table 2 values.
16. ✔ **C16** (Figure 9 caption (page 10) + Table 2 row 'Ursula' (page 7))
    - **Claim:** Ursula in $\log(\sqrt{P}/D)$ space: $f = -0.15$ and $C = 0.96$ in Figure 9; Table 2 lists $f_{\sqrt{P}/D} = -0.15$ and $C_{\sqrt{P}/D} = 0.96$.
    - **Checks:** cross_reference_value_match
    - **Verdict:** PASS
    - **Notes:** Provided caption values match provided Table 2 values.
17. ✔ **C17** (Figure 10 caption (page 10) + Table 2 row 'Beagle' (page 7))
    - **Claim:** Beagle in $\log(\sqrt{P}/D)$ space: $f = 0.03$ and $C = 0.92$ in Figure 10; Table 2 lists $f_{\sqrt{P}/D} = 0.03$ and $C_{\sqrt{P}/D} = 0.92$.
    - **Checks:** cross_reference_value_match
    - **Verdict:** PASS
    - **Notes:** Provided caption values match provided Table 2 values.
18. ✔ **C18** (Figure 11 caption (page 10) + Table 2 row 'Hansa' (page 7))
    - **Claim:** Hansa in $\log(\sqrt{P}/D)$ space: $f = 0.09$ and $C = 0.96$ in Figure 11; Table 2 lists $f_{\sqrt{P}/D} = 0.09$ and $C_{\sqrt{P}/D} = 0.96$.
    - **Checks:** cross_reference_value_match
    - **Verdict:** PASS
    - **Notes:** Provided caption values match provided Table 2 values.
19. ✔ **C19** (Figure 12 caption (page 10) + Table 2 row 'Tirela' (page 7))
    - **Claim:** Tirela in $\log(\sqrt{P}/D)$ space: $f = 0.28$ and $C = 0.98$ in Figure 12; Table 2 lists $f_{\sqrt{P}/D} = 0.28$ and $C_{\sqrt{P}/D} = 0.98$.
    - **Checks:** cross_reference_value_match
    - **Verdict:** PASS
    - **Notes:** Provided caption values match provided Table 2 values.
20. ✔ **C20** (Figure 13 caption (page 11) + Table 2 row 'Baptistina' (page 7))
    - **Claim:** Baptistina in $\log(\sqrt{P}/D)$ space: $f = 0.08$ and $C = 0.97$ in Figure 13; Table 2 lists $f_{\sqrt{P}/D} = 0.08$ and $C_{\sqrt{P}/D} = 0.97$.
    - **Checks:** cross_reference_value_match
    - **Verdict:** PASS
    - **Notes:** Provided caption values match provided Table 2 values.
21. ✔ **C21** (Figure 14 caption (page 11) + Table 2 row 'Adeona' (page 7))
    - **Claim:** Adeona in $\log(\sqrt{P}/D)$ space: $f = 0.28$ and $C = 0.98$ in Figure 14; Table 2 lists $f_{\sqrt{P}/D} = 0.28$ and $C_{\sqrt{P}/D} = 0.98$.
    - **Checks:** cross_reference_value_match
    - **Verdict:** PASS
    - **Notes:** Provided caption values match provided Table 2 values.
22. ✔ **C22** (Figure 15 caption (page 11) + Table 2 row 'Koronis' (page 7))
    - **Claim:** Koronis in $\log(\sqrt{P}/D)$ space: $f = 0.08$ and $C = 0.99$ in Figure 15; Table 2 lists $f_{\sqrt{P}/D} = 0.08$ and $C_{\sqrt{P}/D} = 0.99$.
    - **Checks:** cross_reference_value_match
    - **Verdict:** PASS
    - **Notes:** Provided caption values match provided Table 2 values.
23. ✔ **C23** (Figure 16 caption (page 12) + Table 2 row 'Massalia' (page 7))
    - **Claim:** Massalia in $\log(\sqrt{P}/D)$ space: $f = 0.27$ and $C = 0.97$ in Figure 16; Table 2 lists $f_{\sqrt{P}/D} = 0.27$ and $C_{\sqrt{P}/D} = 0.97$.
    - **Checks:** cross_reference_value_match
    - **Verdict:** PASS
    - **Notes:** Provided caption values match provided Table 2 values.
24. ✔ **C24** (Figure 17 caption (page 12) + Table 2 row 'Pallas' (page 7))
    - **Claim:** Pallas in $\log(P)$ space: $f = 0.07$ and $C = 0.79$ in Figure 17; Table 2 lists $f_P = 0.07$ and $C_P = 0.79$.

    - **Checks:** cross_reference_value_match
    - **Verdict:** PASS
    - **Notes:** Provided caption values match provided Table 2 values.

25. ✔ **C25** (Page 9 (Results 3.4) + Table 2 rows for Theobalda and Veritas (page 7))
    - **Claim:** Theobalda has estimated age of 7 Myr and $C_P = 0.74$; Veritas has age of 8 Myr and $C_P = 0.97$.
    - **Checks:** unit_conversion_and_cross_table_match
    - **Verdict:** PASS
    - **Notes:** Myr→Gyr conversions (0.007 and 0.008) agree with Table 2 age 0.01 within $\mathrm{abs}_\mathrm{t}\mathrm{ol} = 0.003$ (rounding). C_P values match exactly.

Limitations

- Audit uses only the provided PDF parsed text; no access to underlying CSV datasets, code, or simulation outputs.
- Figures are not used for numeric extraction beyond their captions (no pixel/plot reading).
- Some internal checks (e.g., Table 1 vs Table 2) may reflect different subsets/definitions; code can detect mismatches but cannot resolve intent without additional context not in the PDF.
- Some checks could not be completed because required source components were not included in the provided payload (e.g., Table 2 member-count data for recounting $N > 30$; full document text for repeated-occurrence searching).