Skepthical review: *Mapping Thermophysical Diversity in Asteroid Families via Spin-Orbit V-Shape Morphology*

Summary

This manuscript proposes an extension of classical asteroid-family Yarkovsky “V-shape” analyses from the usual $(a,\ 1/D)$ space to a spin–size space using semimajor axis a versus a transformed variable based on the product $P \times D$ (spin period times diameter). Using merged catalog data (Sec. 2.1–2.2), the authors assemble a working sample of family members with complete a, D, P information, define a subset of “well-populated” families, and quantify each family’s left/right V-shape wings by fitting 95th-percentile linear quantile-regression envelopes (Sec. 2.3.2). From the fitted wing slopes $(m_L,\ m_R)$ they derive a steepness metric k and a consistency metric C (Sec. 2.3.3), and they explore correlations between family age and both orbital spread (IQR in a) and k (Sec. 2.5, Sec. 3.3). The extensive figures suggest that for some families the $P \times D$-based representation visually yields a sharper envelope than D or P alone (Sec. 3.1), while many families exhibit “inverted” or otherwise atypical k values (Sec. 3.2). However, substantial problems in internal accounting of sample sizes/family sets, incomplete or mis-referenced result tables, under-specified data provenance and uncertainty treatment, and ambiguities in the physical motivation and sign conventions currently limit reproducibility and complicate interpretation—especially for the prevalence/meaning of negative k and for the age–k conclusions (Sec. 1, Sec. 2.2–2.5, Sec. 3.2–3.4, Sec. 4).

Strengths

- Clear, conceptually interesting attempt to incorporate spin information into family V-shape morphology via a simple combined proxy $(P \times D)$, motivated by Yarkovsky-driven evolution (Sec. 1, Sec. 2.3.1).
- Systematic pipeline for merging multiple catalogs and enforcing completeness in key parameters (a, D, P, family membership), enabling a uniform analysis set (Sec. 2.1–2.2).
- Appropriate use of quantile regression to target distribution envelopes rather than central trends, aligned with the goal of tracing V-shape boundaries (Sec. 2.3.2).
- Large set of diagnostic multi-panel figures comparing D-only, P-only, and $P \times D$ parameterizations that helps readers assess the empirical behavior family-by-family (Sec. 3.1).
- Connection to broader evolutionary questions via age–dispersion and age–morphology analyses, with an explicit discussion of limitations and selection effects (Sec. 2.5, Sec. 3.3–3.4).

Major issues

1. **Internal inconsistencies in dataset sizes, family counts, and which subsets are used for which analyses undermine clarity and reproducibility (Abstract; Sec. 2.2; Sec. 3.1–3.3; Sec. 4).** The manuscript variously cites (i) an aggregated dataset of 16,774 asteroids and 37 well-populated families (Abstract), (ii) 3,872 asteroids in 89 families with 15 well-populated families (Sec. 2.2), and (iii) 32 families used in age correlations (Sec. 3.3), while figures/text discuss additional families beyond the listed "well-populated" set (Sec. 3.1–3.2). It is not currently possible to tell which families enter the quantile fits, the k/C distributions, and the correlation tests.

   Recommendation: In Sec. 2.2 and at the start of Sec. 3, define and use a consistent set of named samples (e.g., $N_{\rm raw}$, $N_{\rm merged}$, $N_{\rm core} = a, D, P, \mathrm{fam}$, $N_{\rm fit}$ where V-fits attempted, $N_{\rm robust}$ where fits pass QA, $N_{\rm age}$ where ages exist). Add a single flow table/diagram with counts at each merge/filter step and a table listing every family that appears in any figure or enters any statistic, including: family name, N used, whether it meets $N \geq 50$, whether both wings pass fit-quality criteria, and whether an age is available. Ensure the Abstract and Sec. 4 summarize the same final analysis sample (or explicitly state multiple tiers, e.g., "37 families plotted; 15 primary families used for quantitative summaries; 32 used for age correlations").

2. **Key quantitative outputs are missing or mis-referenced, preventing verification and reuse (Sec. 2.2; Sec. 3.2; Sec. 3.4).** Table numbering appears inconsistent (Table 1 used as a template/summary but later referenced as containing full fit results), and the manuscript does not provide a complete per-family results table with (a_c, m_L, m_R, k, C) and fit status/flags. As a result, claims about ranges of k/C, which families have NaN slopes (e.g., Karin), and which families drive "inverted" behavior cannot be checked.

   Recommendation: Populate the data-summary table in Sec. 2.2 with actual numbers and create a separate, correctly numbered results table (main text, appendix, or supplement) listing for each fitted family: N, a_c definition used, wing Ns, m_L and m_R with uncertainties, k (and any asymmetry metric), C, and a fit-quality flag (pass/fail with reason). Fix all in-text references (Sec. 3.2, Sec. 3.4) to point to the correct table(s).

3. **Physical justification for the specific choice of $P \times D$ (and the assumed linear envelope in a–log space) is currently ambiguous and at points logically inconsistent with the stated drift scaling (Sec. 1; Sec. 2.3.1; Sec. 3.1–3.2).** The Introduction cites an idealized scaling resembling $\mathrm{d}a/\mathrm{d}t \propto 1/(P \cdot D^2)$, then concludes "Consequently" that $P \times D$ is a comprehensive proxy—this does not follow algebraically. More broadly, Yarkovsky's dependence on rotation period is regime-de-

pendent (thermal parameter, diurnal/seasonal components, obliquity), so the expected mapping from Yarkovsky drift limits to an envelope in $(a, \log(P \times D))$ needs either a derivation or a clear statement that the metric is empirical/heuristic.

Recommendation: Revise Sec. 1 and Sec. 2.3.1 to either (a) provide a short, explicit scaling argument linking a boundary in $|a - a_c|$ to a boundary in a chosen spin–size variable (and justify why D^2 is reduced to D, if that is intended), including assumptions/regime; or (b) explicitly frame $P \times D$ as an empirical proxy selected for exploratory morphology, and soften causal language ("Consequently"). In either case, state what physical behavior would produce an upright vs inverted V-shape under your plotting convention, and what deviations might plausibly indicate (e.g., selection bias, resonance truncation, heterogeneous thermophysics, spin-state asymmetries).

4. **Sign conventions and axis definitions are inconsistent across Methods and Results, making the meaning of "upright/inverted" V-shapes and negative k unclear (Sec. 2.3.1; Sec. 3.1–3.2; figure axes).** The Methods define $Y' = \log_{10}(P \times D)$, while many figures/Results use $\log_{10}(1/(P \times D)) = -Y'$ to obtain a visually upright V. This interacts with the choice of fitting an upper (95th) quantile and with the expected signs of m_L and m_R; without a single fixed convention, some negative-k cases may be artifacts of plotting/definition rather than physical inversions.

   Recommendation: Choose one dependent variable definition for fitting and interpretation and enforce it everywhere: either $Y = \log_{10}(1/(P \times D))$ or $Y = \log_{10}(P \times D)$. Then: (i) explicitly state whether you fit the upper or lower envelope ($\tau = 0.95$ vs $\tau = 0.05$) and why; (ii) derive the expected sign pattern for m_L and m_R under an "ideal" family; (iii) update the definition/interpretation of k accordingly. Add a short note in Sec. 3.2 clarifying which negative-k cases remain negative under the unified convention.

5. **The definitions of k and C, and their interpretation as "steepness" and "consistency/sharpness," are not statistically robust as currently used (Sec. 2.3.3; Sec. 3.2).** $k = (m_R - m_L)/2$ is described as an "average magnitude" in places, which is only true under strict sign assumptions; when slopes change sign (a core result here), k conflates opening, asymmetry, and sign convention. Meanwhile, C (fraction of points below a fitted 95th-quantile line) is close to 0.95 largely by construction and can deviate due to finite-sample effects or model misfit, so treating small departures as morphological "filled-in vs sharp" is risky.

   Recommendation: In Sec. 2.3.3, redefine or supplement the metrics so they remain meaningful under sign violations: e.g., report an opening metric $k_{\text{open}} = (|m_L| + |m_R|)/2$ and an asymmetry metric $k_{\text{asym}} = |m_R| - |m_L|$ (or a ratio), while separately tracking sign anomalies as flags. For "sharpness/consistency," either (i) treat C as a diagnostic of quantile-regression calibration with uncertainty (bootstrap CI for $C - 0.95$), or (ii) replace/supplement C with a dispersion-based envelope metric (e.g., median distance to the fitted envelope; density contrast in a thin band below the envelope; or quantile spacing between $\tau = 0.95$ and $\tau = 0.80$). Update Sec. 3.2 interpretations accordingly.

6. **Quantile-regression methodology, fit-quality criteria, and uncertainty estimates are under-specified, which jeopardizes claims about extreme/inverted k and any downstream correlations (Sec. 2.3.2–2.3.3; Sec. 3.2–3.4).** The manuscript does not specify the software/solver/options used, how NaNs/duplicates are handled, whether any robustness options are used, or explicit criteria for declaring a wing/family "fitable" (minimum N per wing, minimum a-span). Uncertainties on m_L, m_R, k, and derived statistics are not reported.

   Recommendation: Expand Sec. 2.3.2–2.3.3 to specify the quantile-regression implementation (package, solver, convergence criteria), preprocessing for each wing, and explicit fit-quality thresholds (e.g., N_{left} and N_{right} minima; minimum Δa per wing; handling of leverage points). Use bootstrap/resampling within each family to compute uncertainty intervals for m_L, m_R, and derived metrics, and propagate these into any family-level comparisons. In Sec. 3.2, clearly mark and exclude failed/unstable fits from summary histograms and from age correlations, and report results for both the "all attempted" and "robust-only" sets.

7. **Age-correlation analyses (age–IQR(a) and especially age–k) lack clear sample definition, treatment of missing/uncertain values, and robustness checks (Sec. 2.5; Sec. 3.3).** It is unclear whether semimajor axis is osculating or proper (family work typically uses proper elements), which age compilation is used, how age uncertainties are handled, and whether outliers or families with unstable k dominate the correlation results. The mismatch between counts with age data (e.g., 41 vs 32) further clouds interpretation.

   Recommendation: In Sec. 2.5 and Sec. 3.3: (i) specify whether a is proper or osculating and justify; if possible, redo using proper semimajor axis for the family analysis; (ii) provide a table of the exact families used in each correlation, with age values and references (plus uncertainties/ranges when available), N, IQR(a), and k (with uncertainty); (iii) state how NaN/failed fits are handled; (iv) add robustness checks (Spearman with/without extreme $|k|$, Kendall's τ, and sensitivity to plausible age uncertainties). Report n used in each test and interpret "non-correlation" cautiously when uncertainty is large.

8. **Selection effects from requiring measured spin periods (and the joint availability of D and P) are acknowledged but not quantified; they may materially shape the observed envelopes and the apparent "improvement" of $P \times D$ (Sec. 3.4; implications for Sec. 3.1–3.2).** Because spin-period measurements are highly incomplete and biased toward larger/brighter objects and certain observing campaigns, the sampled boundary at small D (where drift is largest) may be truncated or distorted, potentially creating spurious sharpness or inversions.

Recommendation: Quantify selection bias per family: show, for a few representative families and/or in an appendix, (i) the D distribution of all cataloged family members versus the subset with P, and (ii) how the fitted envelope changes when restricting to size ranges with more uniform completeness. Where possible, compare against the full family distribution in a standard family catalog (even if P is missing) to assess how representative the P-available subset is. Temper Sec. 3.1 claims about systematic improvement of $P \times D$ unless supported under these checks.

9. **Reproducibility is limited by incomplete data provenance and processing documentation (Sec. 2.1; Sec. 2.2).** Listing local CSV filenames is insufficient: sources, versions/dates, units, and conflict-resolution rules (multiple P solutions, diameter choices, family membership definitions) are not specified. This prevents independent replication.

   Recommendation: In Sec. 2.1–2.2, provide full provenance for each input (catalog/survey name, citation, version/date, access link/DOI if available), the units used (P in hours/days; D in km; a in au), and rules for resolving duplicates/conflicting measurements. Add a short reproducibility statement indicating whether code and processed tables will be archived (recommended), and at minimum include enough detail that a reader can reconstruct the merged dataset.

Minor issues

1. Dimensional inconsistency: the manuscript applies $\log_{10}$ to dimensionful quantities (e.g., $\log_{10}(P \times D)$, $\log_{10}(1/P)$, $\log_{10}(1/D)$, $\log_{10}(1/(P \times D))$), so slopes/intercepts (and thus k comparisons) depend on arbitrary unit choices (Sec. 2.3.1; figure axes).

   Recommendation: Nondimensionalize before taking logarithms, e.g., $\log_{10}\left[\left(\frac{P}{P_0}\right) \cdot \left(\frac{D}{D_0}\right)\right]$ with stated reference values (P_0, D_0), and ensure all figure labels reflect the normalization. Briefly note how changing reference units would (or would not) affect k and comparisons across families.

2. Choice of family center a_c as the median semimajor axis is not tested for sensitivity; asymmetric/resonance-truncated families may yield biased wing partitioning and slopes (Sec. 2.3.2; Sec. 3.2).

   Recommendation: Add a sensitivity check for a_c: compare median-a_c to an alternative (literature family center / parent-body a / optimized a_c minimizing envelope residuals). Report how much m_L, m_R, and k change for a few key families (especially those with negative k).

3. Figures generally do not overlay the fitted quantile-regression boundary lines, do not mark a_c, and do not report per-panel N, making it hard to verify the fitting and compare parameterizations (Sec. 3.1; figures throughout).

Recommendation: Overlay fitted envelope lines (with uncertainty bands if available) and mark a_c in each panel, at least for a representative subset in the main text. Include N per panel and a short caption note stating whether the plot shows the robust-fit sample or all points.

4. The claim that $P \times D$ "systematically" yields clearer V-shapes is mostly qualitative (Sec. 3.1).

   Recommendation: Add a compact quantitative comparison across the three parameterizations ($1/D$, $1/P$, $1/(P \times D)$) using a consistent metric (e.g., cross-validated quantile loss, median absolute distance to the fitted envelope, or quantile-spacing). Summarize across families (median and IQR) to support or qualify the visual impression.

5. Interpretations of anomalous/inverted V-shapes are largely speculative without basic diagnostics (Sec. 3.2–3.4).

   Recommendation: For a small set of highlighted families (e.g., Vesta, Eos, Agnia, Massalia), add simple diagnostics: wing sample sizes, spin-period distributions by wing, and proximity to major resonances/overlapping families. Clearly label proposed mechanisms as hypotheses unless supported by these checks.

6. Notation/labeling inconsistencies ($\log$ vs $\log_{10}$, $P \cdot D$ vs $P \times D$, units omitted on axes, varying capitalization "V-shape/V-Shape") reduce readability and can create sign/definition confusion (Sec. 2.3.1; figures and captions).

   Recommendation: Standardize notation throughout: use $\log_{10}$ consistently, specify units on all axes (and normalization if nondimensionalized), and use a single multiplication symbol and term ("$P \times D$"). Ensure captions explicitly state which transformed variable is plotted ($\log_{10}(P \times D)$ vs $\log_{10}(1/(P \times D))$).

Very minor issues

1. Cross-references and numbering issues (e.g., Sec. 2.4 citing the wrong section for the V-shape procedure; re-use/mislabeling of "Table 1") impede verification (Sec. 2.1–2.4; Sec. 3.2).

   Recommendation: Do a full cross-reference pass: ensure each section/table/figure is uniquely numbered and every in-text reference points to the correct item.

2. Minor typography/formatting problems (broken hyphenations like "ex-pected", inconsistent heading styles such as hash-prefixed headings, small caption/axis fonts in some figures) slightly hinder reading (Sec. 1–4).

   Recommendation: Proofread and standardize formatting: fix hyphenation/line breaks, harmonize heading styles with the journal template, and increase figure font sizes/contrast; export figures at print-quality resolution or as vector graphics.

3. Quantile notation like $Y'_{0.95}$ can be misread as an exponent rather than a conditional quantile (Sec. 2.3.2).

   Recommendation: Use unambiguous quantile notation, e.g., $\hat{Q}*{0.95}(Y|a)=m a + c*(*or \hat{q}*(a)$), and define it once.

Key statements and references

- ✔ **For numerous asteroid families, plots of semi-major axis versus the combined parameter $\log_{10}(1/(P \times D))$ reveal a significantly more defined and clearer V-shape than plots using diameter or spin period alone, indicating that the product $P \times D$ more effectively captures the thermophysical properties governing Yarkovsky-induced orbital evolution and serves as a more robust tracer of this evolution across families such as Eos, Gefion, Massalia, Juno, Hertha, Emma, Beagle, Ursula, Veritas, Hungaria, Hoffmeister, Flora, Euphrosyne, Adeona, Tirela, Erigone, Barcelona, Chloris, Maria, and Dora.**
- *Reference(s):* Denario [11]
- *Justification:* No valid PDFs found; assumed supported.

- ✔ **Across the analyzed asteroid families, the steepness coefficient k of the V-shape in $\log_{10}(P \times D)$ versus semi-major axis space spans a broad range from -52.2 (Agnia) to 152.3 (Karin), with a mean of -0.16 and a standard deviation of 32.6, demonstrating substantial diversity in V-shape morphology, including both theoretically expected upright V-shapes (positive k, e.g., Koronis with $k = 5.1$ and Veritas with $k = 23.6$) and unexpected inverted V-shapes (negative k, e.g., Vesta with $k = -2.0$, Eos with $k = -8.0$, and Eunomia with $k = -2.1$).**

- *Reference(s):* Denario [11]
- *Justification:* No valid PDFs found; assumed supported.

- ✔ **For the 32 asteroid families with published age estimates, there is a moderate, statistically significant positive Spearman rank correlation between family age and the interquartile range of semi-major axis (IQR_a), with $\rho = 0.57$, p-value $= 0.00067$, and a 95% confidence interval of $(0.26, 0.83)$, supporting the interpretation that older families are more dynamically dispersed in semi-major axis due to cumulative Yarkovsky-induced orbital drift over gigayear timescales.**

- *Reference(s):* Denario [11]
- *Justification:* No valid PDFs found; assumed supported.

- ✔ **For the same set of families, the correlation between V-shape steepness coefficient k (derived from $\log_{10}(P \times D)$ versus semi-major axis) and family age is weak and statistically non-significant, with Spearman $\rho = 0.34$, p-value $= 0.055$, and a 95% confidence interval of $(-0.01, 0.65)$, indicating that the thermophysical characteristics encoded in k do not exhibit a simple monotonic evolution with age and are likely dominated by intrinsic family properties and complex dynamical histories.**
- *Reference(s):* Denario [11]
- *Justification:* No valid PDFs found; assumed supported.
- ✔ **The consistency metric C, defined as the fraction of family members lying on or below the fitted 95th-percentile quantile-regression boundaries in $\log_{10}(P \times D)$ versus semi-major axis space, takes values between 0.909 and 0.949 across the analyzed families (e.g., Agnia with $C = 0.9353$), confirming that the 95% quantile regression generally succeeds in capturing the intended upper V-shape envelope while revealing that some families have more diffuse or "filled-in" boundaries than others.**
- *Reference(s):* Denario [11]
- *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: substantial

The paper's analytic content centers on defining a transformed dependent variable based on spin period and diameter, fitting left/right linear 95th-percentile quantile-regression boundaries in $(a, \text{transformed variable})$ space, and summarizing each family by slope-derived metrics k and C. The main internal-consistency risks are sign/orientation mismatches between $Y' = \log_{10}(P \cdot D)$ versus $\log_{10}(1/(P \cdot D))$, dimensional issues with logarithms of quantities with units, and an unproven step from an asserted drift scaling $\propto 1/(P \cdot D^2)$ to using $P \cdot D$ as the proxy.

Checked items

1. ⚠ **Yarkovsky scaling to proxy choice** (Introduction, p.2 (paragraph beginning "Theoretically, for an idealized spherical body…"))
    - **Claim:** States drift rate is inversely proportional to $P \times D^2$ and concludes "Consequently, the product $P \times D$ serves as a more comprehensive proxy…".
    - **Checks:** algebraic implication check, definition consistency
    - **Verdict:** UNCERTAIN; confidence: medium; impact: moderate

- **Assumptions/inputs:** Only the proportionality statement in the text is available; constants and omitted factors are not provided.
- **Notes:** From the written proportionality $\mathrm{d}a/\mathrm{d}t \propto 1/(P \cdot D^2)$, a direct algebraic proxy aligned with it would naturally involve $P \cdot D^2$ (or its inverse). The step to $P \cdot D$ is not justified within the PDF; it may be heuristic, but then "Consequently" overstates the derivation.

2. ✖ **Definition of transformed variable Y'** (Sec. 2.3.1 Variable Transformation, p.3)

   - **Claim:** Defines $Y' = \log_{10}(P \times D)$ as the dependent variable for V-shape characterization.
   - **Checks:** dimensional/units consistency, notation consistency
   - **Verdict:** FAIL; confidence: high; impact: critical
   - **Assumptions/inputs:** P and D are physical quantities with units (not specified).
   - **Notes:** The argument of $\log_{10}$ is not dimensionless unless $P \times D$ is normalized by reference units. Without normalization, changing time/length units adds constants to Y' and changes fitted slopes/intercepts, undermining cross-family comparability of k.

3. ✖ **Expected V-shape orientation with Y'** (Sec. 2.3.1 Variable Transformation, p.3 (sentences "Plotting Y' versus a is theoretically expected to produce an upright V-shape... more pronounced for bodies with smaller $P \times D$ values."))

   - **Claim:** Claims $Y' = \log_{10}(P \cdot D)$ vs a should yield an upright V where drift is more pronounced for smaller $P \cdot D$.
   - **Checks:** limiting/sanity case, sign consistency
   - **Verdict:** FAIL; confidence: high; impact: critical
   - **Assumptions/inputs:** As stated elsewhere in the paper, greater drift corresponds to smaller $P \cdot D$ (or smaller drift denominator)., Family center at a_c; farther from center corresponds to greater cumulative drift.
   - **Notes:** If smaller $P \cdot D$ implies greater drift, then objects farther from a_c should have smaller $P \cdot D$, hence smaller Y'. That makes Y' large near the center and small at the extremes (an inverted V/$\wedge$), not an upright V/$\vee$. An upright V would correspond to using $-Y' = \log_{10}(1/(P \cdot D))$ instead.

4. ✔ **Equivalence of plotted variable to $-Y'$** (Sec. 3.1, p.5 (sentence "$\log_{10}(1/(P \times D))$ ... equivalent to $-Y'$ as defined in the Methods section"))

   - **Claim:** States $\log_{10}(1/(P \cdot D))$ equals $-\log_{10}(P \cdot D)$.
   - **Checks:** algebra
   - **Verdict:** PASS; confidence: high; impact: minor
   - **Notes:** Identity holds: $\log_{10}(1/x) = -\log_{10}(x)$ for $x > 0$. (Dimensionality issue still applies if x has units.)

5. ⚠ **Quantile-regression boundary model equations** (Sec. 2.3.2 V-Shape Boundary Fitting, p.3 (bullet equations for left/right wings))

    - **Claim:** Models the 95th-percentile boundary by $Y'{0.95} = m_L a + c_L (left) and Y' = m_R a + c_R$ (right).
    - **Checks:** notation consistency, model specification sanity
    - **Verdict:** UNCERTAIN; confidence: medium; impact: moderate
    - **Assumptions/inputs:** $Y'_{0.95}$ denotes the conditional $\tau = 0.95$ quantile of Y' given a for each wing.
    - **Notes:** As linear quantile-regression forms, these are syntactically consistent. However, whether the 95th percentile corresponds to the intended "boundary" depends on whether the dependent variable is Y' or $-Y'$ (the paper uses both conventions in different places).

6. ✖ **Expected signs of wing slopes (m_L negative, m_R positive)** (Sec. 2.3.2 V-Shape Boundary Fitting, p.3 (sentence describing expected signs))

    - **Claim:** Claims m_L is expected negative and m_R positive for the upright V-shape in Y' vs a.
    - **Checks:** sign consistency, consistency with variable definition
    - **Verdict:** FAIL; confidence: high; impact: critical
    - **Assumptions/inputs:** Upright V means larger y values at larger $|a - a_c|$.
    - **Notes:** The sign expectation matches using $y = \log_{10}(1/(P \cdot D)) = -Y'$ (since y increases as a decreases on the left wing and increases as a increases on the right wing). It contradicts the explicit definition $Y' = \log_{10}(P \cdot D)$ if smaller $P \cdot D$ corresponds to larger drift.

7. ⚠ **Steepness coefficient definition** (Sec. 2.3.3(1) and also Introduction p.2 ($k = (m_R - m_L)/2$))

    - **Claim:** Defines $k = (m_R - m_L)/2$ and describes it as the average magnitude of left and right wing slopes.
    - **Checks:** algebra, definition-text consistency
    - **Verdict:** UNCERTAIN; confidence: high; impact: moderate
    - **Assumptions/inputs:** Expected upright V has $m_L \leq 0 \leq m_R$.
    - **Notes:** Algebraically, if m_L is negative and m_R positive, then $(m_R - m_L)/2 = (|m_R| + |m_L|)/2$, matching "average magnitude." But when slopes violate expected signs (which the paper reports), the formula is not an average magnitude; the text and formula diverge in those cases.

8. ✔ **Negative k interpretation conditions** (Sec. 3.2, p.6 (paragraph beginning "A negative k value implies that $m_R - m_L < 0$…"))

    - **Claim:** Explains negative k occurs if m_R is negative with sufficient magnitude or m_L is positive.

- **Checks:** algebra
- **Verdict:** PASS; confidence: high; impact: minor
- **Assumptions/inputs:** $k = (m_R - m_L)/2$ as defined.
- **Notes:** Correct: $k < 0 \Leftrightarrow m_R < m_L$, which can happen via m_R negative and/or m_L positive (or both), or even both negative with m_R more negative than m_L.

9. ⚠ **Consistency metric C definition vs quantile level** (Sec. 2.3.3(2), p.4 and Sec. 3.2, p.7-8)
   - **Claim:** Defines C as fraction of points on/below the fitted 95th-percentile boundary, expecting $C \approx 0.95$ for a sharp boundary.
   - **Checks:** definition consistency, sanity check (conceptual)
   - **Verdict:** UNCERTAIN; confidence: medium; impact: minor
   - **Assumptions/inputs:** Quantile regression is fit separately per wing at $\tau = 0.95$.
   - **Notes:** C being near 0.95 is largely implied by the choice $\tau = 0.95$ (subject to model misfit and finite-sample effects). The PDF does not provide enough analytic detail to verify whether C is computed wing-wise consistently with the fitting, or whether it meaningfully measures boundary sharpness.

10. ✖ **Results section statement of modeled space** (Sec. 3.2, p.5 (first sentence: "To quantitatively characterize the V-shape in the $\log_{10}(P \times D)$ versus a space…"))
    - **Claim:** States quantile regression was applied in $\log_{10}(P \cdot D)$ vs a space.
    - **Checks:** notation consistency, sign convention consistency
    - **Verdict:** FAIL; confidence: high; impact: critical
    - **Assumptions/inputs:** Figures and Sec. 3.1 emphasize $\log_{10}(1/(P \cdot D))$ as the clearer V-shape variable.
    - **Notes:** This statement conflicts with the V-shape variable emphasized in Sec. 3.1 ($\log_{10}(1/(P \cdot D)) = -Y'$) and with the slope-sign expectations in Methods. If the regression is actually run on $\log_{10}(P \cdot D)$, then the expected slope signs and the interpretation of positive/negative k are inverted relative to the figures.

Limitations

- The PDF does not provide the explicit quantile-regression objective function, residual definitions, or exact algorithm for computing C, so some verification is necessarily limited to consistency of stated definitions.
- No derivations are provided for the asserted Yarkovsky scaling with P and D, so only logical/algebraic consistency of the stated implications can be audited (not the underlying physical derivation).
- The main V-shape results table referenced in Results ("Table 1") is not present in the provided text, limiting cross-checking of symbol usage against tabulated definitions.

Numerical results audit

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

10 numeric checks were executed: 9 PASS and 1 FAIL. Most internal consistency checks (dataset sizes, family counts, table thresholds, quantile-bound logic, and correlation CI/p-value coherence) passed. One cross-section mismatch was identified for the number of families with age data used in correlation analysis (41 vs 32).

Checked items

1. ✔ **CAND-001** (Abstract (page 1) vs Methods/EDA (page 2) vs Results (page 4))
    - **Claim:** Three different total-asteroid counts appear: Abstract says an aggregated dataset of 16,774 asteroids; EDA says refined working dataset of 3,872 asteroids after filtering; Results reiterate 3,872 asteroids. Check internal consistency/relationship (e.g., 16,774 vs 3,872).
    - **Checks:** cross_section_count_consistency
    - **Verdict:** PASS
    - **Notes:** Results dataset count matches filtered-core dataset exactly (3,872). Abstract total (16,774) is $\geq 3{,}872$, but differs in scope; filtered/abstract ratio reported as 0.2308, so wording should clarify which count was analyzed.
2. ✔ **CAND-002** (EDA text (page 2) vs Table 1 (page 3))
    - **Claim:** Initial merged dataset size is $789{,}012$ unique asteroid entries; filtered complete-core dataset is $3{,}872$ asteroids. Table 1 repeats these. Verify Table 1 equals the narrative values.
    - **Checks:** table_vs_text_exact_match
    - **Verdict:** PASS
    - **Notes:** Narrative values match Table 1 exactly for both $789{,}012$ and $3{,}872$.
3. ✔ **CAND-003** (Table 1 (page 3) vs EDA text (page 2))
    - **Claim:** Table 1: Number of families represented = 89; Families with age data available = 41. EDA text: 89 distinct asteroid families. Verify matching and that $41 \leq 89$.
    - **Checks:** count_consistency_and_bounds
    - **Verdict:** PASS
    - **Notes:** The 89-family count matches between text and Table 1; the 41-with-age count is ≤ 89.
4. ✔ **CAND-004** (Table 2 (page 3))

- **Claim:** Table 2 lists the 'Top 15 most populated asteroid families' with member counts. Verify there are exactly 15 families and each listed member count satisfies the stated criterion $N > 50$.
- **Checks:** table_row_count_and_threshold_check
- **Verdict:** PASS
- **Notes:** Parsed 15 family member counts; minimum count is 52, satisfying the strict threshold $N > 50$.

5. ✔ **CAND-005** (Methods: Quantile regression (page 3) and Consistency metric definition (page 4))

   - **Claim:** Quantile regression uses $\tau = 0.95$; consistency metric C is defined as fraction on/below fitted 95th percentile boundary; thus C should be close to 0.95. Later, C range reported 0.909 to 0.949. Verify range values are ≤ 0.95 and plausibly close (compute deviations).
   - **Checks:** range_vs_expected_quantile_check
   - **Verdict:** PASS
   - **Notes:** Hard bound check passed: $C_{\text{max}} = 0.949$ and $C_{\text{min}} = 0.909$ are both $\leq \tau = 0.95$. Deviations $\tau - C$ are 0.0010 (at max) and 0.0410 (at min).

6. ✔ **CAND-006** (Results: V-shape metrics (page 5))

   - **Claim:** Steepness coefficient k range and summary: min $k = -52.2$ (Agnia), max $k = 152.3$ (Karin), mean $k = -0.16$, standard deviation $= 32.6$. Check basic plausibility constraints: min $\leq$ mean $\leq$ max; sd ≥ 0; and (max-min) ≥ 0.
   - **Checks:** summary_stat_sanity_check
   - **Verdict:** PASS
   - **Notes:** Sanity inequalities pass: $-52.2 \leq -0.16 \leq 152.3$; $sd = 32.6$ is non-negative; range computed as 204.5.

7. ✔ **CAND-007** (Results: definition of k (page 4) vs example narrative (page 6))

   - **Claim:** k is defined as $(m_R - m_L)/2$. For Eos, narrative states m_L was positive (14.7031) and Eos has $k = -8.0$ (page 5). Derive implied m_R from k and m_L: $m_R = 2k + m_L$, then check sign/magnitude matches the 'inverted V' explanation (m_R negative and magnitude large enough).
   - **Checks:** algebraic_backsolve_from_definition
   - **Verdict:** PASS
   - **Notes:** Implied $m_R = 2 \times (-8.0) + 14.7031 = -1.2969$, which is negative and consistent with the stated inverted-V interpretation (noting k is rounded).

8. ✔ **CAND-008** (Correlation analysis: Orbital spread vs Age (page 8))

- **Claim:** Spearman correlation reported: $\rho = 0.57$ with p-value $= 0.00067$; 95% CI $(0.26, 0.83)$ does not include zero. Verify CI excludes 0 and that sign of ρ matches CI bounds.
- **Checks:** ci_contains_zero_and_sign_consistency
- **Verdict:** PASS
- **Notes:** $\rho = 0.57$ lies within CI $(0.26, 0.83)$; CI excludes 0; p-value $0.00067 < 0.05$.

9. ✔ **CAND-009** (Correlation analysis: k vs Age (page 9))

   - **Claim:** Spearman correlation reported: $\rho = 0.34$, p-value $= 0.055$; 95% CI $(-0.01, 0.65)$ includes zero. Verify CI includes 0; verify p-value slightly above 0.05; verify ρ is within CI.
   - **Checks:** ci_includes_zero_and_pvalue_threshold
   - **Verdict:** PASS
   - **Notes:** CI includes 0 $(-0.01$ to $0.65)$; $\rho = 0.34$ is within CI; p-value $0.055 > 0.05$.

10. ✖ **CAND-010** (Age-data counts: Table 1 (page 3) vs Correlation section (page 8))

    - **Claim:** Table 1 says 'Families with age data available' $= 41$. Correlation section says analysis used 'the 32 families where such data were available.' Verify mismatch (41 vs 32) and flag for reconciliation.
    - **Checks:** cross_section_count_mismatch_detection
    - **Verdict:** FAIL
    - **Notes:** Counts differ by 9 (41 in Table 1 minus 32 in correlation section). While $32 \leq 41$ holds, the discrepancy requires an explicit explanation of differing scopes/exclusions.

Limitations

- Only parsed text provided; any numeric values present solely inside figures/plots cannot be extracted reliably without pixel-based reading (excluded by scope).
- Per-family k, C, m_L, m_R tables appear missing or mis-referenced (Table 1 is a data summary), preventing verification of reported extrema/means/SD from underlying lists.
- Several numeric statements involve statistical outputs (p-values, confidence intervals) that cannot be recomputed without the underlying dataset; only logical cross-checks among the reported numbers are feasible.
- Inconsistent counts across sections (notably 41 vs 32 families with age data) indicate potential scope differences or missing exclusions that limit confidence in cross-section comparability.