Skepthical review: *Constraining Asteroid Thermal Properties Through Analysis of Spin-Orbital Correlations Within Families*

Summary

This manuscript uses asteroid families as ensemble laboratories to probe long-term coupled Yarkovsky–YORP evolution and, ultimately, constrain thermophysical properties. The authors compile a merged dataset (12 CSV sources) of orbital elements, diameters, spin periods, obliquities, spectral types, and family ages; select 19 families with sufficient membership; define each asteroid's semimajor-axis offset from a family center (Δa relative to a_c); and measure intra-family correlations/regressions between Δa (and sometimes $|\Delta a|$), diameter, spin period, and obliquity, including some taxonomy-based subsamples (Sec. 2.1–2.5; Sec. 3.1–3.2). They then forward-model coupled Yarkovsky drift and YORP spin/obliquity evolution over each family's age using a grid over two "efficiency" parameters (C_{yark}, C_{yorp}) and compare simulated vs observed distributions using regression-slope comparisons and goodness-of-fit statistics (χ^2 / KS-style tests) to infer best-fit parameter ranges (Sec. 2.6–2.7; Sec. 3.3–3.4). The overall approach is promising and the qualitative narrative (obliquity linked to Δa sign; size dependence of drift magnitude; family-to-family differences with age/taxonomy) is plausible, but the paper currently reads more like an exploratory workflow and figure set than a fully reproducible physical inference. Key elements are under-specified (exact dynamical equations, parameter definitions/units, objective function for fitting), several analysis choices may bias or blur the inferred trends (definition of $a_c/\Delta a$, mixing signed vs absolute drift, selection effects in obliquity availability, multiple comparisons), and the mapping from fitted efficiency coefficients to concrete thermal properties remains largely qualitative. Addressing these points with explicit model equations, clearer data provenance/selection-function discussion, a unified and defensible inference metric with uncertainty quantification, and comprehensive per-family summary tables would substantially strengthen the robustness and interpretability of the claimed constraints.

Strengths

- Timely scientific goal: leveraging asteroid-family spin–orbit structure to constrain long-term Yarkovsky/YORP evolution and relate it to thermophysical properties (Sec. 1; Sec. 4).
- Clear end-to-end workflow: data compilation → intra-family correlation/regression → forward simulations → model–data comparison, which is conceptually coherent and easy to follow at a high level (Sec. 2.1–2.7).
- Empirical trends generally align with expected qualitative physics: obliquity linked to drift direction (Δa sign) and size dependence affecting drift magnitude, providing a reasonable starting point for forward modeling (Sec. 3.1–3.2).

- Forward-modeling approach is appropriate in principle: scanning (C_{yark}, C_{yorp}) grids per family and comparing simulated to observed distributions is a natural route to constrain efficiencies (Sec. 2.6–2.7; Sec. 3.3).
- Use of robust/non-parametric measures (e.g., Spearman/robust regression) is a sensible choice given outliers and non-Gaussian scatter typically present in family data (Sec. 2.4–2.5).
- Figures appear designed to support side-by-side observation/simulation comparison (density overlays/contours), which is helpful for diagnosing model behavior family-by-family (Figures 1–21; Sec. 3.3).

Major issues

1. **The coupled Yarkovsky–YORP evolution model is not specified with sufficient mathematical and algorithmic detail to be assessed or reproduced (Sec. 2.6–2.7; Sec. 3.3).** In particular, the manuscript does not provide the explicit forms of $\mathrm{d}a/\mathrm{d}t$, $\mathrm{d}\omega/\mathrm{d}t$, and $\mathrm{d}\epsilon/\mathrm{d}t$ (or equivalent), whether diurnal/seasonal components are included, how heliocentric-distance scaling enters, how diameter/density/albedo/thermal inertia enter, and how numerical integration is performed (time step, scheme, random seeds, number of realizations). Closely related: C_{yark} and C_{yorp} are not defined mathematically (units/dimensions, baseline normalization), so reported best-fit magnitudes cannot be audited or compared to prior work.

   Recommendation: In Sec. 2.6, add an explicit "Model equations and parameters" subsection that (i) writes the governing evolution equations used for $a(t)$, $\omega(t)$, and $\epsilon(t)$, with citations; (ii) defines C_{yark} and C_{yorp} precisely (including units or explicit statement that they are dimensionless scalings relative to a stated baseline model), and lists all symbols with dimensions; (iii) states whether diurnal and seasonal Yarkovsky are both modeled and how thermal parameters are represented (explicitly or absorbed into C_{yark}); and (iv) documents the integration method, timestep, duration (= family age), random seeding, and number of Monte Carlo realizations per grid point. Provide the full (C_{yark}, C_{yorp}) grid ranges/spacing in a table and ensure Sec. 3.3 and figure captions reference these exact definitions.

2. **Definition of the family center a_c and the derived variable Δa may be biased and is not aligned with standard family analyses unless carefully justified (Sec. 2.3; Sec. 3.1).** The manuscript sets a_c as the mean of member semimajor axes and uses $\Delta a_i = a_i - a_c$, but the mean can be shifted by Yarkovsky-evolved tails, truncation by resonances, and interlopers; and it is unclear whether proper or osculating elements are used. Since Δa is central to every correlation and to the simulation matching, an inconsistent anchor can systematically distort both the observed slopes and inferred efficiencies.

Recommendation: In Sec. 2.3, explicitly state whether a is proper or osculating and justify the choice. Replace or benchmark the mean-based a_c against at least one robust alternative (median/sigma-clipped mean; using only large D objects less affected by drift; or adopting published family centers from a standard family catalog). Perform a sensitivity test for a subset of key families (including one with apparent anomalies, e.g., Gefion) showing how (i) ϵ–Δa correlations, (ii) D–$|\Delta a|$ relations, and (iii) best-fit (C_{yark}, C_{yorp}) regions change under alternative a_c definitions and/or outlier clipping. Report the outcome explicitly in Sec. 3.1/3.3.

3. **Signed Δa and $|\Delta a|$ are used inconsistently in interpreting size-dependent drift (Sec. 3.1; related figures across Sec. 3).** The physical expectation from Yarkovsky scaling is primarily a relationship between size and drift magnitude ($|\Delta a|$ or V-shape envelope), while the sign of Δa depends on spin-axis orientation (e.g., $\cos\epsilon$). Several places motivate a diameter–Δa correlation with arguments that actually correspond to diameter–$|\Delta a|$, which risks confusion and can generate misleading null/positive correlations if both drift directions are present.

   Recommendation: Audit Sec. 3.1–3.2 and all relevant plots to state explicitly, for each size–drift test, whether Δa is signed or absolute. For the core "size dependence of drift" claim, prioritize D vs $|\Delta a|$ (and/or a V-shape boundary method) and report those results consistently. If any signed D–Δa analysis is retained, provide a clear physical rationale (e.g., asymmetric obliquity distribution, resonance removal on one side, selection effects) and demonstrate that the inference does not hinge on this choice.

4. **The statistical inference and goodness-of-fit framework used to select best-fit (C_{yark}, C_{yorp}) is not defined in a unified, reproducible way (Sec. 2.7; Sec. 3.3).** It is unclear exactly which 1D distributions are compared with which tests, how χ^2 is computed (binning/normalization), whether any multi-dimensional comparison is attempted, how regression-slope comparisons enter, and what scalar objective function determines "best". Use of KS p-values as an optimization target (especially a min-p rule) is ad hoc and sensitive to sample size; additionally, near-flat contour maps in some families could indicate weak identifiability, plotting-range issues, or an implementation bug.

   Recommendation: In Sec. 2.7, define a single objective function used to rank grid points (or a clearly specified multi-criteria decision rule), and justify it. Specify: (i) which observables are compared (Δa, spin period, obliquity; possibly transformed variables like $\cos\epsilon$); (ii) the exact test statistics used, including binning/smoothing choices for any χ^2; and (iii) how statistics are combined into one score. Strongly consider adding or substituting a distance metric better suited for optimization (e.g., Wasserstein distance for 1D; energy distance/MMD for 2D like $(\Delta a, \cos\epsilon)$). For each family, quantify identifiability by showing example simulated distributions at several grid points and report whether the objective varies meaningfully; if contours are flat, state that constraints are weak and explain why (data sparsity, degeneracy, or model insensitivity).

5. **Measurement uncertainties and observational selection effects—especially for obliquity—are not incorporated, risking biased correlations and biased (C_{yark}, C_{yorp}) inference (Sec. 2.1–2.5; Sec. 3.1–3.3).** The requirement of complete property sets likely selects larger/brighter objects with dense photometric coverage, and the availability of spin-axis solutions is strongly non-random. Family ages also have non-negligible uncertainties that directly trade off against inferred drift efficiencies.

   Recommendation: In Sec. 2.1–2.2, provide data provenance and typical uncertainties for D, P, ϵ, and ages, with citations. For each family, report: total catalog members vs those used after requiring (a, D, P, ϵ, taxonomy), and summarize how the used subset differs in D (or H), inclination, and a. In Sec. 2.4–2.5 and Sec. 2.7, propagate uncertainties via bootstrap/Monte Carlo resampling (including perturbing ages within quoted uncertainties) to produce confidence intervals on correlations, regression slopes, and best-fit parameter regions. At minimum, repeat the key ϵ–Δa and model-fitting steps on a diameter-limited subset where spin-axis completeness is closer to uniform, and discuss any changes in Sec. 4.

6. **Key quantitative results are not comprehensively reported across all 19 families, limiting robustness assessment and reproducibility (Sec. 3.1–3.4).** The text highlights examples, but there is no single table listing, per family, N used per analysis, correlation coefficients (with uncertainty), regression parameters, and the inferred best-fit (C_{yark}, C_{yorp}) with uncertainty/acceptable ranges and fit-quality metrics.

   Recommendation: Add summary tables (Sec. 3.1–3.3): (i) per-family sample sizes for each observable (N with D; with P; with ϵ), and Spearman/Pearson (or preferred) correlations for ϵ–Δa (or $\cos\epsilon$–Δa), D–$|\Delta a|$, and any spin-related relations, with p-values and bootstrap confidence intervals; (ii) per-family best-fit (C_{yark}, C_{yorp}) with uncertainty regions (e.g., Δobjective threshold), alongside the associated goodness-of-fit values. Ensure Sec. 3.3 and Sec. 4 base comparative claims (C-type vs S-type; young vs old) on these tables rather than selective examples.

7. **The physical interpretation "constraints on thermal properties" is not quantitatively supported because the mapping from (C_{yark}, C_{yorp}) to thermal inertia/conductivity/albedo/roughness/density is left qualitative and is likely degenerate (Sec. 3.3–3.4; Sec. 4).** Without an explicit thermophysical model and stated priors/assumptions, a constrained efficiency coefficient does not uniquely imply thermal inertia; similarly, C_{yorp} is strongly shape/roughness dependent.

Recommendation: Either (a) explicitly map inferred (C_{yark}, C_{yorp}) ranges to thermal inertia (and/or conductivity) ranges using stated formulae and assumed parameters (density, albedo, beaming/roughness, shape distribution), including a degeneracy discussion; or (b) narrow claims to "relative efficiencies" and avoid direct thermal-property language. For a few benchmark families (e.g., Themis, Koronis, Vesta or other well-studied cases), provide an order-of-magnitude conversion with uncertainty bands and compare to published thermophysical constraints. Clearly flag where inference is model-dependent (Sec. 4).

8. **Several key modeling/analysis choices that strongly affect long-term spin-state distributions are missing or un-justified: collisional/stochastic reorientation, YORP self-limitation/cycles, spin-barrier/fission and any reset rule, and the handling of resonances or dynamical depletion that can reshape Δa distributions (Sec. 2.6–2.7; Sec. 3.3–3.4).** Without these, apparent fits may be non-unique, and outliers (e.g., anomalous trends such as Gefion's reported size–drift behavior) cannot be interpreted confidently.

   Recommendation: In Sec. 2.6–2.7, explicitly list included vs neglected physical processes and justify omissions with timescale arguments (size- and heliocentric-distance dependent). If collisional reorientation and/or YORP cycling is omitted, discuss how this might bias obliquity clustering and spin-period distributions and thus inferred C_{yorp}. In Sec. 3.4 or Sec. 4, add a limitations/alternatives subsection addressing resonance truncation, initial ejection-velocity fields, and dynamical diffusion; for families with anomalous behavior (e.g., Gefion noted in Sec. 3.1), add a short quantitative check (resonance proximity, asymmetry in Δa distribution, interloper contamination, or age uncertainty sensitivity).

Minor issues

1. Obliquity is treated as a linear variable in regressions despite being bounded/angular (0–180°), and the theoretically natural dependence is often on $\cos\epsilon$ (sign and magnitude of diurnal Yarkovsky component) rather than ϵ itself (Sec. 2.5; Sec. 3.2). Linear fits in ϵ can be hard to interpret and can be sensitive to boundary effects.

   Recommendation: In Sec. 2.5 and Sec. 3.2, either (i) switch primary analyses to $\cos\epsilon$ (or an axial/circular-statistics treatment) when relating spin axis to Δa, or (ii) justify why ϵ-linear regression is adequate as a purely descriptive fit over the observed range. Report robustness of key results under $\epsilon \to \cos\epsilon$ transformation.

2. Multiple comparisons are not handled explicitly despite testing many correlations across 19 families, multiple variable pairs, and sometimes taxonomy subsamples (Sec. 2.4; Sec. 3.1). This increases false-discovery risk and can inflate the apparent number of "significant" trends.

Recommendation: State the number of tests performed and apply a multiple-testing correction (e.g., Benjamini–Hochberg FDR) or justify a different approach. In Sec. 3.1, mark which correlations remain significant after correction and adjust language for marginal cases.

3. Data provenance is described in implementation terms (merging CSVs) rather than as a scientific description of catalogs, versions, and quality control (Sec. 2.1). The specific sources for diameters, spin periods, obliquities, taxonomy, family membership, and ages are not clearly listed, nor is the handling of conflicting entries or quality flags.

   Recommendation: Rewrite Sec. 2.1 to list each input catalog with citation and access date/version, specify merge/priority rules, and document any quality cuts (e.g., LCDB quality codes; spin-axis solution reliability). Move pandas/file-format details to an appendix or repository documentation.

4. Taxonomy usage is under-specified: the adopted taxonomic system, treatment of ambiguous/multiple classifications, and the criteria for labeling a family as C/S/X/B (or mixed) are not clearly defined (Sec. 2.4; Sec. 3.3–3.4). This matters because taxonomy informs assumed densities/albedos and the interpretation of "thermal property" differences.

   Recommendation: In Sec. 2.4, state the taxonomy system(s) used (e.g., Bus–DeMeo/SDSS/SMASS), how conflicts are resolved, and the threshold for assigning a dominant type to a family. If densities/albedos differ by type in simulations, list the adopted values and sources.

5. Figures 1–21 are difficult to evaluate quantitatively due to readability/annotation gaps (small fonts, crowded panels) and missing key numeric context (N, fitted slopes, ρ, p-values, best-fit parameters, KS/χ^2 values). Some axis labels/units and Δa definitions appear inconsistent across panels.

   Recommendation: Increase resolution and font sizes; standardize axis labels with units (e.g., Δa in AU; spin period in h); add panel labels referenced in captions; and include per-panel/caption annotations: sample sizes (N_{obs}, N_{sim}), correlation/regression statistics (ρ, slope ± CI), and best-fit (C_{yark}, C_{yorp}) with fit metrics. Where relevant, add reference lines ($\Delta a = 0$, $\epsilon = 90^\circ$).

6. The Conclusions emphasize qualitative trends but do not clearly separate robust results from tentative inferences, and do not provide a concise "deliverables" summary (Sec. 4).

   Recommendation: Revise Sec. 4 to list the most robust quantitative findings (by family, referencing the new summary tables), then explicitly list limitations/degeneracies (selection effects, age uncertainties, model simplifications), and end with specific next steps (e.g., proper-element anchoring, selection-function forward modeling, more realistic YORP cycling).

7. Keywords in the Abstract include unrelated topics (e.g., solar flares/galaxy abundances/comets), suggesting template or quality-control issues and potentially harming indexing.

   Recommendation: Replace keywords with manuscript-relevant terms (asteroid families; Yarkovsky; YORP; radiative torques; spin–orbit evolution; main-belt asteroids; thermal inertia).

8. Reproducibility artifacts: the manuscript does not clearly state code/data availability, random seeds, run counts per grid point, or how to regenerate the master dataset and simulation outputs (Sec. 2.1; Sec. 2.6–2.7).

   Recommendation: Provide a repository link (or supplementary materials) with: data-ingest scripts and catalog references, simulation code or pseudocode, parameter grids per family, random seeds, and instructions to reproduce each figure/table. If code cannot be released, provide enough pseudocode and parameter tables for independent reimplementation.

Very minor issues

1. Notation and formatting are inconsistent (e.g., Δa vs Delta_a_AU; a_c vs a_{c}; ϵ/ϵ vs Obliquity_deg), and ϵ can be visually confused with “efficiency” near $C_{\mathrm{yark}}/C_{\mathrm{yorp}}$ (Sec. 2.3–2.7; Sec. 3.1–3.3).

   Recommendation: Standardize manuscript notation (use Δa, a_c, and ϵ consistently in math text; reserve code-style names for tables). Add a symbol table and avoid using ϵ if ϵ is used elsewhere.

2. At least one regression formula-like statement is malformed/ambiguous (e.g., missing operator in a model specification such as “SpinPeriod_hr Delta_a_AU”) (Sec. 2.5).

   Recommendation: Rewrite all model specifications using standard notation (e.g., $y \sim x$ or $y = \beta_0 + \beta_1 x$) and ensure consistency across sections and captions.

3. Several sentences are overly long and there are minor typographical/hyphenation artifacts (e.g., broken words) (Sec. 1; Sec. 4).

   Recommendation: Proofread and tighten sentence structure, especially in Sec. 1 and Sec. 4, to improve readability without changing content.

4. Edge-case conventions are not stated (e.g., how $\epsilon = 90^\circ$ is classified in prograde/retrograde statements; consistent inequality usage) (Sec. 3.1).

   Recommendation: State explicitly how $\epsilon = 90^\circ$ is treated and keep inequality conventions consistent throughout.

5. Non-standard affiliation text (if present in the manuscript) appears unlikely to meet typical journal formatting requirements.

Recommendation: Ensure author affiliations and acknowledgements conform to the target journal's policies and standard scientific style.

Key statements and references

- **• Old, C-type asteroid families such as Themis, Alauda, and Ursula are best reproduced by YORP–Yarkovsky simulations when the Yarkovsky efficiency coefficient is in the range $C_{\text{yark}} \approx 10^{-9}$–$10^{-8}$ and the YORP efficiency coefficient is $C_{\text{yorp}} \approx 10^{-14}$, implying relatively high thermal inertia and comparatively inefficient YORP torques for these populations.**
- *Reference(s):* (none)
- **• S-type asteroid families, exemplified by Koronis and Eunomia, are matched by simulations with slightly higher Yarkovsky efficiencies of $C_{\text{yark}} \approx 10^{-8}$, which is interpreted as evidence for lower thermal inertia compared to C-type families.**
- *Reference(s):* (none)
- **• In the Maria family (S-type, 3.0 Gyr), there is a strong negative correlation between asteroid diameter and the absolute value of semimajor axis dispersion Δa, with a Spearman correlation coefficient $\rho = -0.388$ and $p < 0.0001$, indicating that smaller asteroids are located at larger $|\Delta a|$ from the family center, consistent with size-dependent Yarkovsky drift.**
- *Reference(s):* (none)

Mathematical consistency audit

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

Maths relevance: light

The paper is primarily descriptive/statistical and simulation-driven. Explicit mathematics in the provided PDF text is limited to the definition of semimajor-axis dispersion ($\Delta a_i = a_i - a_c$), qualitative sign relationships between obliquity and drift direction, and example linear regression fits (obliquity as a linear function of Δa). The core simulation model is referenced but not written as explicit equations; the key fitted parameters C_{yark} and C_{yorp} are not defined mathematically, which blocks a full internal analytic consistency check of the main physical inference.

Checked items

1. ✔ **Family center semimajor axis definition** (Sec. 2.3, p.2)
    - **Claim:** The family's central semimajor axis is defined as $a_c = \text{mean}(a_i)$ over family members.
    - **Checks:** definition consistency, dimensional consistency

   - **Verdict:** PASS; confidence: high; impact: minor
   - **Assumptions/inputs:** Family membership set is fixed for computing the mean, a_i are in consistent units (AU per column name SemimajorAxis_AU)
   - **Notes:** Definition is clear and dimensionally coherent (a_c has same units as a_i).

2. ✔ **Semimajor-axis dispersion formula** (Sec. 2.3, p.2)

   - **Claim:** For each asteroid i, $\Delta a_i = a_i - a_c$ is computed and stored as Delta_a_AU.
   - **Checks:** algebra, definition consistency
   - **Verdict:** PASS; confidence: high; impact: minor
   - **Assumptions/inputs:** a_c computed as the mean semimajor axis of the same set
   - **Notes:** Algebra is correct; also implies $\mathrm{mean}(\Delta a) = 0$ by construction, which is internally consistent with using Δa as a centered measure of dispersion.

3. ✔ **Prograde/retrograde drift sign logic** (Sec. 3.1, p.4)

   - **Claim:** Prograde rotation ($\epsilon < 90°$) tends to yield outward drift (positive Δa), retrograde ($\epsilon > 90°$) inward drift (negative Δa).
   - **Checks:** logical consistency, notation consistency
   - **Verdict:** PASS; confidence: medium; impact: moderate
   - **Assumptions/inputs:** ϵ is obliquity measured from orbit normal, ranging 0 –180°, Δa is signed relative to family center
   - **Notes:** Within the paper's own definitions, the mapping between ϵ regimes and sign of Δa is stated consistently and supports the subsequent correlation interpretation.

4. ✔ **Negative obliquity–Δa correlation inference** (Sec. 3.1, p.4)

   - **Claim:** Given the drift sign behavior, the observable manifests as a negative correlation between obliquity and Δa.
   - **Checks:** sanity/limiting-case check
   - **Verdict:** PASS; confidence: medium; impact: moderate
   - **Assumptions/inputs:** Data contain a mix of prograde (low ϵ, $+\Delta a$) and retrograde (high ϵ, $-\Delta a$) rotators
   - **Notes:** If points cluster at (low ϵ, $+\Delta a$) and (high ϵ, $-\Delta a$), rank/linear correlation is indeed negative. No algebraic inconsistency detected.

5. ⚠ **Yarkovsky size dependence statement** (Sec. 3.1, p.4)

   - **Claim:** The Yarkovsky drift rate is inversely proportional to asteroid diameter, so smaller asteroids drift further from the family center.

   - **Checks:** model-to-observable mapping (symbolic)
   - **Verdict:** UNCERTAIN; confidence: low; impact: moderate
   - **Assumptions/inputs:** Drift accumulates roughly linearly with time for fixed parameters, Other dependencies are held fixed or average out
   - **Notes:** The statement is plausible but the paper provides no explicit analytic expression for $\mathrm{d}a/\mathrm{d}t$ as a function of diameter and thermal parameters, so the proportionality cannot be internally verified from the PDF alone.

6. ⚠ **Signed vs absolute drift in diameter correlations** (Sec. 3.1, p.4 (Maria family uses $|\Delta a|$; surrounding text discusses Δa generally))

   - **Claim:** Diameter is expected to correlate with Δa due to size-dependent drift; an example is given as a negative correlation between diameter and $|\Delta a|$ in Maria.
   - **Checks:** definition consistency, logical consistency
   - **Verdict:** UNCERTAIN; confidence: medium; impact: moderate
   - **Assumptions/inputs:** Physical expectation described is about distance from center, not drift direction
   - **Notes:** The text mixes expectations about drifting 'further from the center' (suggesting $|\Delta a|$) with discussion of correlations involving signed Δa. This is a clarity/consistency gap that can change the sign/interpretation of expected correlations.

7. ✔ **Example robust regression equation (Themis)** (Sec. 3.2, p.4–5)

   - **Claim:** Obliquity $= -360.57 \times \Delta a + 81.04$ (Themis).
   - **Checks:** algebraic form check, dimensional consistency
   - **Verdict:** PASS; confidence: high; impact: minor
   - **Assumptions/inputs:** Obliquity measured in degrees, Δa measured in AU (as defined earlier)
   - **Notes:** Linear form is consistent; implied slope units are deg/AU and intercept deg. No internal contradiction with earlier variable definitions.

8. ✔ **Example robust regression equations (Koronis, Eos)** (Sec. 3.2, p.5)

   - **Claim:** Koronis: Obliquity $= -344.00 \times \Delta a + 85.70$; Eos: Obliquity $= -316.77 \times \Delta a + 94.94$.
   - **Checks:** dimensional consistency, notation consistency
   - **Verdict:** PASS; confidence: high; impact: minor
   - **Assumptions/inputs:** Same units as above
   - **Notes:** Consistent presentation and units by implication.

9. ✖ **Regression-model notation typo/ambiguity** (Sec. 2.5, p.3)

- **Claim:** Linear regression example is written as “SpinPeriod_hr Delta_a_AU”.
- **Checks:** notation consistency
- **Verdict:** FAIL; confidence: high; impact: minor
- **Assumptions/inputs:** Intended to denote a regression of SpinPeriod_hr on Delta_a_AU
- **Notes:** As written it is missing an operator ($\sim$, $=$, vs.). This is a notation error that should be corrected for mathematical clarity.

10. ⚠ **Definition of C_{yark} and C_{yorp} (efficiency coefficients)** (Sec. 3.3, p.5–6 (also referenced Sec. 2.6, p.3))

    - **Claim:** Simulations span a grid of Yarkovsky (C_{yark}) and YORP (C_{yorp}) efficiency coefficients; best-fit values are reported (e.g., $C_{\text{yark}} \approx 10^{-9}$–$10^{-8}$; $C_{\text{yorp}} \approx 10^{-14}$).
    - **Checks:** symbol definition consistency, dimensional/unit consistency
    - **Verdict:** UNCERTAIN; confidence: high; impact: critical
    - **Assumptions/inputs:** C_{yark} and C_{yorp} parameterize the strength of $\mathrm{d}a/\mathrm{d}t$ and YORP-driven spin evolution in the simulator
    - **Notes:** No equations or dimensional definitions are given for $C_{\text{yark}}/C_{\text{yorp}}$. Without the functional form (how they enter $\mathrm{d}a/\mathrm{d}t$, $\mathrm{d}\omega/\mathrm{d}t$, $\mathrm{d}\epsilon/\mathrm{d}t$), their units and meaning cannot be checked, and the mathematical validity/interpretability of the reported parameter ranges cannot be audited from the PDF.

11. ✔ **Best-fit criterion (maximize minimum KS p-value)** (Sec. 3.3, p.5)

    - **Claim:** Best-fit parameters are chosen by maximizing the minimum p-value across KS tests on Δa, spin period, and obliquity distributions.
    - **Checks:** definition consistency
    - **Verdict:** PASS; confidence: medium; impact: minor
    - **Assumptions/inputs:** A set of p-values is computed per parameter pair and then aggregated via $\min(\cdot)$
    - **Notes:** The criterion is mathematically well-defined as stated (argmax over parameter grid of the minimum p-value). Details of which specific KS tests (1D marginals vs conditional slices) are not specified, but no internal contradiction is visible.

12. ✔ **Initial-condition distributions (Maxwellian spin rates; isotropic axes)** (Sec. 2.6, p.3)

    - **Claim:** Synthetic population uses a Maxwellian distribution for initial spin rates and an isotropic distribution for initial spin axes.
    - **Checks:** conceptual/mathematical consistency

  - **Verdict:** PASS; confidence: low; impact: minor
  - **Assumptions/inputs:** Spin rates are nonnegative, Isotropic axes correspond to uniform distribution on the sphere (implying uniform in $\cos$(obliquity))
  - **Notes:** Statements are internally consistent, but the PDF does not provide explicit probability density functions, so only a high-level consistency check is possible.

Limitations

- The provided PDF text contains almost no explicit derivations or numbered equations; most of the work is described procedurally. This limits the audit to checking definitions, stated relationships, and dimensional/notation coherence.
- The numerical model is central to the paper's conclusions, but the governing evolution equations and the definitions (including dimensions/units) of C_{yark} and C_{yorp} are not present in the PDF text, preventing a symbolic verification of the main parameter inference.
- Figures are referenced heavily; however, the audit is restricted to analytic checks and cannot validate whether plotted relationships correspond to the stated formulas beyond basic consistency.
- Where a statement would normally be checked against a standard physical formula (e.g., $\mathrm{d}a/\mathrm{d}t$ scaling), the audit cannot use external sources; if the formula is not in the PDF, the verdict must remain UNCERTAIN.

Numerical results audit

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

17 candidate numeric/logical checks were executed: 14 PASS and 3 UNCERTAIN, with no FAIL outcomes. Most PASS results are bounds/sign/unit sanity checks on reported correlation statistics and regression coefficients; several broader consistency checks could not be completed without full document text or underlying data.

Checked items

1. ✔ **C1** (Page 4, Results §3.1 (Themis family; obliquity–Δa))
   - **Claim:** Themis family (C-type, 2.36 Gyr): Spearman $\rho = -0.288$ with p-value < 0.0001.
   - **Checks:** range_and_sign_check
   - **Verdict:** PASS
   - **Notes:** ρ within $[-1, 1]$. ρ sign negative as stated. Reported p-value is an upper bound $< 1 \times 10^{-4}$ (cannot verify actual p without data). age positive.

2. ✔ **C2** (Page 4, Results §3.1 (Themis B-type subgroup; obliquity–Δa))

  - **Claim:** Within the B-type subgroup of Themis: Spearman $\rho = -0.446$ with p-value < 0.0001.
  - **Checks:** range_and_sign_check
  - **Verdict:** PASS
  - **Notes:** ρ within $[-1, 1]$. ρ sign negative as stated. Reported p-value is an upper bound $< 1 \times 10^{-4}$ (cannot verify actual p without data).

3. ✔ **C3** (Page 4, Results §3.1 (Eos family; obliquity–Δa))

   - **Claim:** Eos family (X-type, 1.3 Gyr): Spearman $\rho = -0.197$ with p-value < 0.0001.
   - **Checks:** range_and_sign_check
   - **Verdict:** PASS
   - **Notes:** ρ within $[-1, 1]$. ρ sign negative as stated. Reported p-value is an upper bound $< 1 \times 10^{-4}$ (cannot verify actual p without data). age positive.

4. ✔ **C4** (Page 4, Results §3.1 (Koronis family; obliquity–Δa))

   - **Claim:** Koronis family (S-type, 1.76 Gyr): Spearman $\rho = -0.189$ with $p = 0.0004$.
   - **Checks:** range_and_sign_check
   - **Verdict:** PASS
   - **Notes:** ρ within $[-1, 1]$. ρ sign negative as stated. p_{value} within $[0, 1]$. age positive.

5. ✔ **C5** (Page 4, Results §3.1 (Maria family; diameter vs $|\Delta a|$))

   - **Claim:** Maria family (S-type, 3.0 Gyr): Spearman $\rho = -0.388$ between diameter and $|\Delta a|$, with p-value < 0.0001.
   - **Checks:** range_and_sign_check
   - **Verdict:** PASS
   - **Notes:** ρ within $[-1, 1]$. ρ sign negative as stated. Reported p-value is an upper bound $< 1 \times 10^{-4}$ (cannot verify actual p without data). age positive.

6. ✔ **C6** (Page 4, Results §3.1 (Koronis family; diameter vs $|\Delta a|$))

   - **Claim:** Koronis family (S-type, 1.76 Gyr): Spearman $\rho = -0.216$ between diameter and $|\Delta a|$, with p-value < 0.0001.
   - **Checks:** range_and_sign_check
   - **Verdict:** PASS
   - **Notes:** ρ within $[-1, 1]$. ρ sign negative as stated. Reported p-value is an upper bound $< 1 \times 10^{-4}$ (cannot verify actual p without data). age positive.

7. ✔ **C7** (Page 4, Results §3.1 (Gefion family overall; diameter vs Δa))

- **Claim:** Gefion family (S-type, 0.48 **Gyr**): positive correlation between diameter and Δa with $\rho = 0.307$, $p = 0.0011$.
  - **Checks:** range_and_sign_check
  - **Verdict:** PASS
  - **Notes:** ρ within $[-1, 1]$. ρ sign positive as stated. p_{value} within $[0, 1]$. age positive.

8. ✔ **C8** (Page 4, Results §3.1 (Gefion S-type members; diameter vs Δa))
   - **Claim:** Within Gefion S-type members: $\rho = 0.581$, $p < 0.0001$.
   - **Checks:** range_and_sign_check
   - **Verdict:** PASS
   - **Notes:** ρ within $[-1, 1]$. ρ sign positive as stated. Reported p-value is an upper bound $< 1 \times 10^{-4}$ (cannot verify actual p without data).
9. ✔ **C9** (Page 4, Results §3.1 (Eunomia family; spin period vs diameter))
   - **Claim:** Eunomia family: Spearman $\rho = -0.170$ between spin period and diameter with $p = 0.0005$.
   - **Checks:** range_and_sign_check
   - **Verdict:** PASS
   - **Notes:** ρ within $[-1, 1]$. ρ sign negative as stated. p_{value} within $[0, 1]$.
10. ⚠ **C10** (Page 4, Results §3.1 (Themis family; spin period vs Δa))
    - **Claim:** Themis family: Spearman $\rho = 0.104$ between spin period and Δa with $p = 0.0453$.
    - **Checks:** range_and_sign_check
    - **Verdict:** UNCERTAIN
    - **Notes:** Insufficient inputs: missing age.
11. ⚠ **C11** (Page 2, Methods §2.3 (family selection criterion))
    - **Claim:** Minimum number of members criterion is $N_{\text{min}} = 40$.
    - **Checks:** constant_reuse_check
    - **Verdict:** UNCERTAIN
    - **Notes:** Cannot regex-search 'full text' because only PAYLOAD is available (no document text provided).
12. ✔ **C12** (Page 2, Methods §2.1 (data files ingested))
    - **Claim:** Data consolidated from twelve CSV files, each containing a specific asteroid property, followed by a list of 12 properties.
    - **Checks:** parts_vs_total_count
    - **Verdict:** PASS
    - **Notes:** Parsed 12 properties from comma-separated list.

13. ✔ **C13** (Page 1 Abstract + Page 5 Results §3.3 + Page 7 Conclusions (number of families))
    - ○ **Claim:** Paper repeatedly states analysis/simulations are for 19 well-characterized/selected families.
    - ○ **Checks:** repeated_constant_consistency
    - ○ **Verdict:** PASS
    - ○ **Notes:** Consistency across provided mentions only; cannot search for additional conflicting mentions without full document text.
14. ✔ **C14** (Page 4, Results §3.2 (Themis regression equation))
    - ○ **Claim:** Fitted model for Themis family: Obliquity $= -360.57 \times \Delta a + 81.04$.
    - ○ **Checks:** unit_reasonableness_check
    - ○ **Verdict:** PASS
    - ○ **Notes:** Slope finite and nonzero. Slope negative. Intercept in $[0, 180]$.
15. ✔ **C15** (Page 5, Results §3.2 (Koronis regression equation))
    - ○ **Claim:** Koronis model: Obliquity $= -344.00 \times \Delta a + 85.70$.
    - ○ **Checks:** unit_reasonableness_check
    - ○ **Verdict:** PASS
    - ○ **Notes:** Slope finite and nonzero. Slope negative. Intercept in $[0, 180]$.
16. ✔ **C16** (Page 5, Results §3.2 (Eos regression equation))
    - ○ **Claim:** Eos model: Obliquity $= -316.77 \times \Delta a + 94.94$.
    - ○ **Checks:** unit_reasonableness_check
    - ○ **Verdict:** PASS
    - ○ **Notes:** Slope finite and nonzero. Slope negative. Intercept in $[0, 180]$.
17. ⚠ **C17** (Page 4, Results §3.1 (obliquity threshold for prograde/retrograde))
    - ○ **Claim:** Prograde rotation: $\epsilon < 90°$; retrograde rotation: $\epsilon > 90°$.
    - ○ **Checks:** logical_partition_check
    - ○ **Verdict:** UNCERTAIN
    - ○ **Notes:** From the stated strict inequalities (ϵ< 90, ϵ> 90), ϵ= 90 is unclassified. Cannot search elsewhere for $\leq/\geq$ usage because full document text is not provided in PAYLOAD.

Limitations

- The provided PDF text contains only a small subset of the computed numerical outputs (most tables of correlations/p-values and fit diagnostics are not included as machine-readable numbers).

- No raw dataset (CSV/Parquet) is included in the PDF text; therefore, correlations, regressions, Δa computations, and simulation-based best fits cannot be recomputed from first principles.
- Figures contain additional numeric information (axes ranges, colorbar scales, best-fit coordinates) but extracting values from images/pixels is excluded by the task constraints.
- Insufficient inputs prevented completion of the Themis spin period vs Δa cross-check (C10) and prevented full-text consistency scans for $N_{\min}$ and the 90° threshold convention (C11, C17).