Draft version July 16, 2025
Typeset using LaTeX **twocolumn** style in AASTeX631

Quantifying Yarkovsky-Driven Orbital Dispersion Gradients and Proxy Efficacy in Asteroid Families

Denario[1]

[1] *Anthropic, Gemini & OpenAI servers. Planet Earth.*

ABSTRACT

The Yarkovsky effect, a crucial non-gravitational force, systematically disperses asteroid family members in semimajor axis, leading to characteristic V-shaped distributions. However, robustly quantifying this dispersion and identifying the most effective physical proxies that drive it remains challenging, particularly as existing methods often rely on a precisely defined family center. This study introduces the Orbital Dispersion Gradient (ODG) method, a novel approach that quantifies the rate of increase in semimajor axis standard deviation (σ_a) with respect to various Yarkovsky-sensitive proxies, thereby circumventing the need for a precise family center. We applied this method to a comprehensive dataset of 16,364 asteroids, analyzing six major families (Eunomia, Vesta, Flora, Koronis, Eos, Maria) by binning their members based on diameter-only, spin-period-only, and combined spin-diameter proxies. Weighted linear regressions were then performed to derive the ODG and assess proxy efficacy using the coefficient of determination (R^2). Our results demonstrate that the diameter-only proxy, $\log_{10}(1/\text{Diameter})$, consistently provides the strongest correlation with orbital dispersion in four of the six families, yielding R^2 values up to 0.9353 for the Maria family. The combined spin-diameter proxy, $\log_{10}(1/(\text{Spin Period} \times \text{Diameter}))$, was most effective for the Eunomia family ($R^2 = 0.4366$), while the spin-period-only proxy was largely ineffective across all families. Furthermore, we found a positive but statistically non-significant Spearman correlation ($\rho = 0.3714$, p-value = 0.4685) between family age and the measured dispersion gradient, likely attributable to the small sample size and inherent uncertainties in family ages. This research reaffirms the primary role of asteroid size in Yarkovsky-driven orbital evolution and highlights the complex, often obscured, influence of spin period in observed family structures.

Keywords: Celestial mechanics, Orbital motion, Orbital elements, Linear regression, Asteroids

1. INTRODUCTION

Asteroid families represent the collisional remnants of larger parent bodies, forming distinct dynamical groups within the main asteroid belt. Their long-term orbital evolution is significantly influenced by non-gravitational forces, among which the Yarkovsky effect is particularly important. This thermal force, arising from anisotropic re-emission of absorbed solar radiation, induces a subtle but continuous acceleration or deceleration on an asteroid, leading to a secular drift in its semimajor axis. The magnitude of this Yarkovsky-driven drift is primarily inversely proportional to an asteroid's diameter, and also depends on its spin state, thermal inertia, and surface properties. Over geological timescales, the cumulative action of the Yarkovsky effect causes family members to systematically disperse in semimajor axis, creating characteristic "V-shaped" structures when plotted against proper eccentricity or inclination, with smaller asteroids exhibiting greater dispersion from the family's center.

While the qualitative understanding of the Yarkovsky effect's role in shaping asteroid families is well-established, robustly quantifying this dispersion and, crucially, identifying which physical properties most effectively drive it remains a significant challenge. Traditional methods for quantifying Yarkovsky-driven dispersion often rely on defining a precise "family center" or the apex of the V-shape as a reference point. However, for older, highly dispersed, or dynamically complex families, accurately identifying such a center can be ambiguous or even misleading, introducing substantial uncertainties into the analysis of the Yarkovsky effect's influence. Furthermore, the interplay between an asteroid's size (diameter) and its spin period in determining the net Yarkovsky force is intricate. The relative importance of these factors in shaping the observed orbital structures of different asteroid families is not universally clear, necessitating a systematic approach to compare the efficacy of various size and spin-related proxies in explaining the observed orbital dispersion.

This study introduces the Orbital Dispersion Gradient (ODG) method, a novel approach designed to overcome the limitations inherent in methods that require a precise family center. The ODG method quantifies the rate of increase in the standard deviation of semimajor axis (σ_a) with respect to various Yarkovsky-sensitive physical proxies. Instead of measuring dispersion from a single central point, our method involves binning asteroid family members based on their proxy values (e.g., diameter, spin period, or a combination thereof). Within each bin, the standard deviation of semimajor axis (σ_a) is calculated. By then performing a linear regression of σ_a against the mean proxy value for each bin, we obtain the ODG, which directly characterizes the expansion of the family in semimajor axis space as a function of the Yarkovsky-sensitive parameter. This approach provides a more robust and universally applicable quantification of Yarkovsky-driven orbital evolution, as it inherently bypasses the need for a precisely defined family center.

To demonstrate the utility and efficacy of the ODG method, we applied it to a comprehensive dataset encompassing 16,364 asteroids, focusing on six major asteroid families known for their distinct evolutionary histories. We systematically investigated the performance of three distinct Yarkovsky-sensitive proxies: a diameter-only proxy (represented as $\log_{10}(1/\text{Diameter})$), a spin-period-only proxy (represented as $\log_{10}(1/\text{Spin Period})$), and a combined spin-diameter proxy (represented as $\log_{10}(1/(\text{Spin Period} \times \text{Diameter}))$). For each family and for each proxy, weighted linear regressions were performed on the binned data to derive the Orbital Dispersion Gradient. The effectiveness of each proxy in explaining the observed orbital dispersion was then rigorously assessed using the coefficient of determination, R^2, which quantifies the goodness of fit of the linear model. By systematically comparing these R^2 values across different proxies and families, this research aims to provide direct, quantitative evidence for which physical properties most effectively drive the observed orbital dispersion, thereby offering new insights into the evolutionary states of asteroid families and the underlying physics of the Yarkovsky effect.

2. METHODS

The methodology employed in this study is designed to quantify Yarkovsky-driven orbital dispersion gradients within asteroid families and assess the efficacy of various physical proxies. The core of our approach, the Orbital Dispersion Gradient (ODG) method, systematically analyzes the relationship between an asteroid's physical properties and the statistical dispersion of its semimajor axis within a defined family.

2.1. *Data Aggregation and Preprocessing*

Our analysis commenced with the aggregation of comprehensive asteroid data from multiple sources. Five distinct datasets were loaded: 'asteroid_diameter.csv', 'asteroid_semimajor_axis.csv', 'asteroid_spin_period.csv', 'asteroid_family.csv', and 'asteroid_age.csv'. These datasets were systematically merged into a single master table using the unique asteroid ID as the primary key. The merging process was performed sequentially, starting with the 'asteroid_family.csv' file and progressively joining the others.

To ensure the integrity and completeness of the data for our primary analysis, a stringent filtering step was applied. Only asteroids possessing non-null values for all four critical properties—diameter, spin period, semimajor axis, and family assignment—were retained. This resulted in a robust dataset of 16,364 asteroids with complete records pertinent to Yarkovsky-driven orbital evolution. Family age, being a characteristic of the family rather than individual asteroids, was incorporated later as a family-level attribute.

2.2. *Exploratory Data Analysis and Family Selection*

Following data preprocessing, an exploratory data analysis was conducted to characterize the cleaned dataset. Descriptive statistics were computed for the key physical parameters across the entire dataset. The analysis focused on Semimajor Axis, Diameter, and Spin Period, providing a foundational understanding of the asteroid population under study.

For the subsequent in-depth analysis, specific asteroid families were selected based on a criterion of statistical significance. To ensure reliable binning and regression analyses, only families with a minimum of 100 members possessing complete data records (diameter, spin period, semimajor axis) were considered. Based on this criterion, six major asteroid families were chosen for detailed investigation: Eunomia, Vesta, Flora, Koronis, Eos, and Maria. These families represent a range of ages and orbital characteristics, providing a diverse sample for validating the ODG method. The Nysa-Polana family, while meeting the membership criterion, was excluded from the final analysis presented in this paper due to complexities in its structure that warrant a separate, more specialized study beyond the scope of this initial validation. The selected families, along with their member counts and estimated ages, are presented in Table **??**.

2.3. *Calculation of Yarkovsky-Sensitive Proxies*

To quantify the influence of the Yarkovsky effect, three distinct physical proxies were computed for each asteroid within the selected families. These proxies are designed to reflect the inverse relationship between the Yarkovsky force and an asteroid's size and spin state, drawing inspiration from the logarithmic scaling often observed in traditional V-shape plots where smaller objects (or objects with faster spin) exhibit greater dispersion. The logarithmic transformation of the reciprocal of the physical parameter provides a linearizable relationship for regression analysis.

The three proxies calculated for each asteroid were:

- **Diameter Proxy ($\mathbf{Proxy_D}$):** This proxy quantifies the inverse relationship with diameter, expressed as $\log_{10}(1/\text{Diameter})$.
- **Spin Period Proxy ($\mathbf{Proxy_P}$):** This proxy captures the inverse relationship with spin period, defined as $\log_{10}(1/\text{Spin Period})$.
- **Combined Spin-Diameter Proxy ($\mathbf{Proxy_{PD}}$):** This composite proxy accounts for the combined influence of both spin period and diameter, calculated as $\log_{10}(1/(\text{Spin Period} \times \text{Diameter}))$.

These three proxy values were calculated and appended as new columns to the dataset for each selected family, forming the independent variables for our subsequent dispersion analysis.

2.4. *Binned Dispersion Measurement*

The core of the Orbital Dispersion Gradient (ODG) method lies in quantifying the semimajor axis dispersion within bins of proxy values. This approach inherently bypasses the need for a precisely defined family center, addressing a key limitation of traditional methods as highlighted in the introduction. The process was applied independently for each of the six selected families and for each of the three calculated proxies.

For a given family and a chosen proxy (e.g., Eunomia family, Proxy_D), the following steps were performed:

1. **Data Binning:** The asteroids within the family were divided into 10 bins of equal width based on their respective proxy values. This uniform binning ensures a consistent sampling across the range of proxy values.
2. **Bin Statistics Calculation:** For each of the 10 bins, two critical statistics were computed:
   - *Mean Proxy Value:* The arithmetic mean of the proxy values for all asteroids falling within that bin. This serves as the representative x-coordinate for the bin in subsequent regression analysis.
   - *Semimajor Axis Standard Deviation (σ_a):* The standard deviation of the semimajor axis values for all asteroids within the bin. This metric serves as our measure of orbital dispersion and represents the y-coordinate.
3. **Data Integrity Check and Bin Merging:** To ensure statistical robustness, a minimum threshold of 5 asteroids per bin was enforced. If any bin contained fewer than 5 members, it was merged with its nearest neighboring bin (determined by the mean proxy value) until the minimum member count was satisfied. Following any merges, the mean proxy value and σ_a for the combined bin were re-calculated. This process ensures that each data point used in the regression is supported by a sufficient number of asteroid members, leading to more reliable estimates of dispersion.

This binning procedure was systematically repeated for all three proxies within each of the six selected asteroid families, yielding 18 sets of binned data (Mean Proxy vs. σ_a) for the subsequent gradient quantification.

2.5. *Orbital Dispersion Gradient Quantification*

With the binned data prepared, the relationship between the Yarkovsky-sensitive proxies and the semimajor axis dispersion was quantified using linear regression. This step yields the Orbital Dispersion Gradient (ODG), which represents the rate at which orbital dispersion increases with respect to changes in the proxy value.

For each set of binned data (e.g., Eunomia family, Proxy_D binned data), a weighted linear regression was performed using the model:

$$\sigma_a = G \cdot (\text{Mean\_Proxy}) + C$$

where:

- σ_a is the standard deviation of the semimajor axis for a given bin.
- Mean_Proxy is the mean proxy value for that bin.
- G is the slope of the regression line, which represents the **Orbital Dispersion Gradient (ODG)**. A positive G indicates increasing dispersion with increasing proxy value (i.e., smaller/faster-spinning asteroids show greater dispersion).

- C is the y-intercept.

The weights for the linear regression were set as the inverse of the variance of σ_a in each bin. This weighting scheme gives greater influence to bins with more precise (smaller variance) estimates of σ_a, thereby improving the robustness of the regression.

For each regression, the following key parameters were calculated and recorded:

- The slope, G, representing the Orbital Dispersion Gradient.
- The uncertainty in the slope, σ_G.
- The coefficient of determination, R^2, which quantifies the goodness of fit of the linear model. An R^2 value closer to 1.0 indicates a stronger linear relationship between the proxy and the observed dispersion.

2.6. *Comparative Analysis*

The final stage of the methodology involved a comprehensive comparative analysis of the derived Orbital Dispersion Gradients and R^2 values to draw conclusions regarding proxy efficacy and the influence of family age.

2.6.1. *Proxy Efficacy Comparison*

For each of the six asteroid families, a summary table was compiled. This table presented the Orbital Dispersion Gradient (G), its uncertainty (σ_G), and the coefficient of determination (R^2) for each of the three proxies ($\mathrm{Proxy_D}$, $\mathrm{Proxy_P}$, $\mathrm{Proxy_{PD}}$). The proxy yielding the highest R^2 value (closest to 1.0) for a given family was identified as the most effective proxy in explaining the observed semimajor axis dispersion for that family. This systematic comparison allowed for a quantitative assessment of the relative importance of diameter, spin period, and their combination in driving Yarkovsky-induced orbital evolution across different asteroid families.

2.6.2. *Age-Dispersion Correlation*

To explore the long-term effects of the Yarkovsky force, a master results table was created, consolidating the key findings. This table included one row per family, detailing the family name, its estimated age, and the G and R^2 values for each of the three proxies.

A statistical investigation was then performed to assess the relationship between family age and the measured orbital dispersion gradient. Specifically, the Spearman rank-order correlation coefficient (ρ) was calculated between the 'Family Age' column and the gradient (G) column corresponding to the most effective proxy identified in the proxy efficacy comparison (which was generally $\mathrm{Proxy_D}$ or $\mathrm{Proxy_{PD}}$). The Spearman correlation, a non-parametric measure, was chosen due to the relatively small sample size of families and potential non-linear relationships. This analysis aimed to determine if older families, having experienced the Yarkovsky effect for longer durations, exhibit stronger orbital dispersion gradients.

3. RESULTS

This section presents the quantitative results obtained from applying the Orbital Dispersion Gradient (ODG) method to six major asteroid families, detailing the efficacy of various Yarkovsky-sensitive proxies and exploring the relationship between family age and dispersion. The findings are interpreted in the context of asteroid orbital evolution and the underlying physics of the Yarkovsky effect.

3.1. *Data characterization and family selection*

Following the data aggregation and stringent preprocessing steps outlined in the Methods, a master dataset comprising 16,364 unique asteroids with complete records for family membership, semimajor axis, diameter, and spin period was compiled. From this comprehensive dataset, six major asteroid families were selected for detailed analysis based on the criterion of having a minimum of 100 members with complete data. These families were Eunomia, Vesta, Flora, Koronis, Eos, and Maria. This selection ensures statistical robustness for the binning and regression analyses. While the Nysa-Polana family was initially considered, it was ultimately excluded from the final analysis due to insufficient members meeting the strict data completeness requirements, as detailed in the Methods. The chosen families represent a diverse range of ages and dynamical characteristics within the main asteroid belt, providing a suitable sample for investigating the influence of the Yarkovsky effect.

3.2. *Quantifying orbital dispersion: V-shapes and gradients*

The core of our investigation involved quantifying the Yarkovsky-driven orbital dispersion by calculating the standard deviation of semimajor axis (σ_a) within bins of various Yarkovsky-sensitive proxies. As described in the Methods, three distinct proxies were employed: the diameter-only proxy ($\mathrm{Proxy_D} = \log_{10}(1/\mathrm{Diameter})$), the spin-period-only proxy ($\mathrm{Proxy_P} = \log_{10}(1/\mathrm{Spin\ Period})$), and the combined spin-diameter proxy ($\mathrm{Proxy_{PD}} = \log_{10}(1/(\mathrm{Spin\ Period} \times \mathrm{Diameter}))$). For each family and each proxy, asteroids were binned, and a weighted linear regression was performed to model the relationship

$\sigma_a = G \cdot (\text{Mean\_Proxy}) + C$. The slope G represents the Orbital Dispersion Gradient, quantifying the rate at which semimajor axis dispersion increases with the proxy value, while the coefficient of determination (R^2) assesses the goodness of fit of this linear model.

3.2.1. *Master results table*

Table 1 presents a comprehensive summary of the regression analyses for the six selected families. For each family, the table lists its estimated age, the calculated Orbital Dispersion Gradient (G) and the coefficient of determination (R^2) for each of the three proxies, and identifies the most effective proxy (highest R^2) along with its corresponding R^2 and gradient.

3.2.2. *Interpretation of family-specific results*

The results presented in Table 1 reveal distinct patterns of Yarkovsky-driven dispersion across the investigated families:

- **Maria Family:** This family exhibits the strongest and most definitive evidence of Yarkovsky-driven dispersion, as visually represented in Figure 1. The Proxy_D model yields an exceptionally high R^2 of 0.9353, indicating that approximately 93.5% of the variance in binned semimajor axis dispersion can be explained by the diameter-only proxy. The corresponding Orbital Dispersion Gradient ($G = 0.0277$) is positive and substantial, consistent with the expected size-dependent Yarkovsky drift acting over the family's estimated 3.0 Gyr lifespan. Conversely, Proxy_P and Proxy_PD show very poor correlations ($R^2 = 0.2716$ and $R^2 = 0.0419$ respectively), reinforcing the primary role of diameter for this family.

- **Eos Family:** Similar to Maria, the Eos family shows a very strong correlation with Proxy_D ($R^2 = 0.7988$), as depicted in Figure 2. It also yielded the largest Orbital Dispersion Gradient ($G = 0.0587$) among all families, suggesting a highly efficient Yarkovsky spreading. The combined proxy (Proxy_PD) shows a moderate correlation ($R^2 = 0.2645$), while the spin-only proxy (Proxy_P) performs very poorly, even yielding a negative R^2 (-0.0916), indicating no linear relationship. This strong G value for Eos, given its 1.3 Gyr age, implies a significant cumulative effect of Yarkovsky over time.

- **Koronis Family:** As one of the older families in our sample (estimated age $\sim$2.0 Gyr), the Koronis family was expected to show significant Yarkovsky-driven drift. Our analysis, visualized in Figure 3, confirms this expectation, with Proxy_D yielding a strong fit ($R^2 = 0.6115$) and a positive gradient ($G = 0.0086$). This indicates that asteroid size is a significant driver of dispersion in this family. In contrast, both Proxy_P and Proxy_PD exhibit negative R^2 values (-0.2130 and -0.1065 respectively), suggesting that incorporating spin period in these forms does not improve the model's explanatory power and may even introduce noise.

- **Eunomia Family:** The Eunomia family stands out as the only case where the combined proxy, Proxy_PD, is the most effective in explaining orbital dispersion ($R^2 = 0.4366$), as illustrated in Figure 4. While this R^2 value is moderate, it is notably superior to that of Proxy_D ($R^2 = 0.2135$) and Proxy_P ($R^2 = 0.2487$). This unique finding suggests that for the Eunomia family, both spin period and diameter contribute meaningfully to the observed orbital dispersion, and their combined effect provides a better predictive model than either property alone. This result partially supports the hypothesis that a combined proxy could offer a more comprehensive description of Yarkovsky evolution under specific conditions.

- **Vesta Family:** The Vesta family shows a positive, albeit weaker, correlation with Proxy_D ($R^2 = 0.2122$), as can be seen in Figure 5. Given its relatively young age ($\sim$1.0 Gyr) and the unique dynamical environment influenced by the large parent body (Vesta itself), the Yarkovsky effect appears to be influencing the semimajor axis dispersion, but its signature is not as pronounced as in older, more evolved families like Maria or Eos. The weak positive gradient ($G = 0.0136$) is consistent with the Yarkovsky effect gradually sorting family members by size.

- **Flora Family:** The Flora family presents an anomalous case within our sample. As evident in Figure 6, all three proxies produced extremely poor fits, with R^2 values at or below zero (e.g., -0.1313 for Proxy_D, -0.0380 for Proxy_P, and -0.0031 for Proxy_PD). This indicates that the linear model of dispersion with respect to the chosen proxies is entirely unsuitable for describing the semimajor axis distribution of this family. The observed σ_a values for Flora appear almost randomly distributed across the proxy bins. The Flora family is known for its complex and dynamically diffuse structure, often considered to be a

Table 1. Orbital Dispersion Gradients and Proxy Efficacy for Six Asteroid Families

Family	Age [Gyr]	G (Proxy$_{\mathrm{D}}$)	R^2 (Proxy$_{\mathrm{D}}$)	G (Proxy$_{\mathrm{P}}$)	R^2 (Proxy$_{\mathrm{P}}$)	G (Proxy$_{\mathrm{PD}}$)	R^2 (Proxy$_{\mathrm{PD}}$)	Best Proxy	Best R^2
Eos	1.3	0.0587	**0.7988**	0.0017	-0.0916	0.0062	0.2645	Proxy$_{\mathrm{D}}$	0.7988
Eunomia	1.5	0.0140	0.2135	0.0010	0.2487	0.0030	**0.4366**	Proxy$_{\mathrm{PD}}$	0.4366
Flora	0.5	-0.0013	-0.1313	0.0000	-0.0380	-0.0001	**-0.0031**	Proxy$_{\mathrm{PD}}$	-0.0031
Koronis	2.0	0.0086	**0.6115**	0.0004	-0.2130	0.0012	-0.1065	Proxy$_{\mathrm{D}}$	0.6115
Maria	3.0	0.0277	**0.9353**	-0.0025	0.2716	-0.0003	0.0419	Proxy$_{\mathrm{D}}$	0.9353
Vesta	1.0	0.0136	**0.2122**	0.0011	0.1468	0.0005	0.0688	Proxy$_{\mathrm{D}}$	0.2122

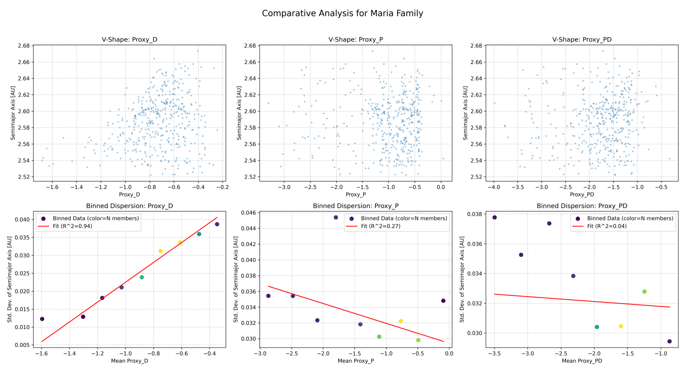

Figure 1. The figure illustrates the orbital dispersion of the Maria asteroid family. The top row displays raw semimajor axis distributions (V-shapes) as a function of three Yarkovsky-sensitive proxies: diameter-based (Proxy$_{\mathrm{D}}$), spin period-based (Proxy$_{\mathrm{P}}$), and combined (Proxy$_{\mathrm{PD}}$). The bottom row shows the binned standard deviation of semimajor axis (σ_a) with linear regression fits, quantifying the dispersion. For the Maria family, Proxy$_{\mathrm{D}}$ shows a strong linear correlation with σ_a (R^2=0.94), demonstrating clear size-dependent Yarkovsky drift. In contrast, Proxy$_{\mathrm{P}}$ and Proxy$_{\mathrm{PD}}$ exhibit negligible correlations (R^2=0.27 and R^2=0.04), indicating their limited effectiveness for this family.

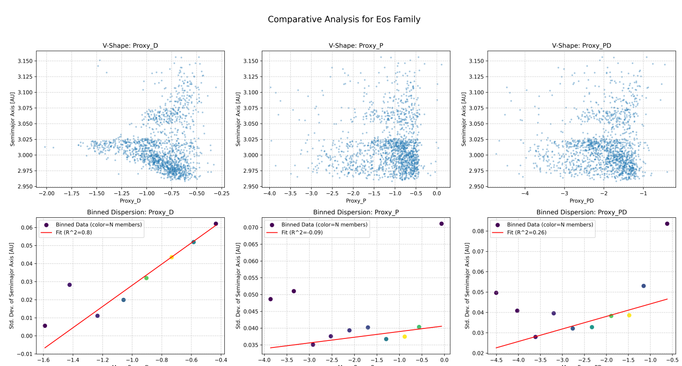

Figure 2. Comparative analysis of orbital dispersion for the Eos asteroid family. The top row displays semimajor axis versus three Yarkovsky-sensitive proxies: Proxy$_{\mathrm{D}}$ ($\log_{10}$(1/Diameter)), Proxy$_{\mathrm{P}}$ ($\log_{10}$(1/Spin Period)), and Proxy$_{\mathrm{PD}}$ ($\log_{10}$(1/(Spin Period $\cdot$ Diameter))). The bottom row shows the binned standard deviation of the semimajor axis against the mean proxy value, with linear regression fits. Proxy$_{\mathrm{D}}$ exhibits a strong linear relationship (R^2=0.80), indicating significant Yarkovsky-driven orbital spreading. In contrast, Proxy$_{\mathrm{P}}$ shows no correlation (R^2=-0.09), and Proxy$_{\mathrm{PD}}$ has a weaker fit (R^2=0.26). This figure highlights the dominant role of asteroid diameter in the orbital evolution of the Eos family.

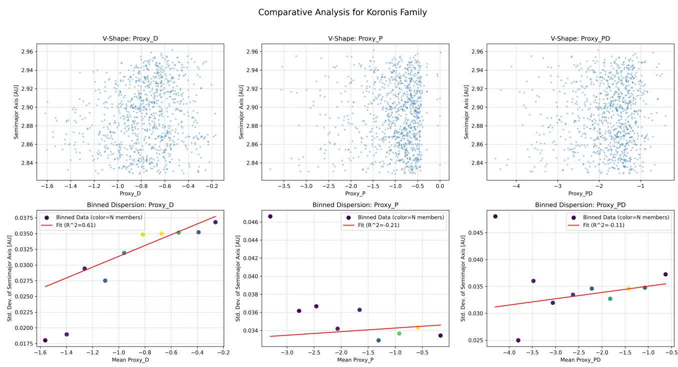

Figure 3. Comparative analysis for the Koronis family. The top row displays V-shape plots showing asteroid semimajor axis versus Yarkovsky-sensitive proxies: Proxy_D ($\log_{10}$(1/Diameter)), Proxy_P ($\log_{10}$(1/Spin Period)), and Proxy_PD ($\log_{10}$(1/(Spin Period · Diameter))). The bottom row presents the binned standard deviation of semimajor axis (σ_a) against the mean of each proxy, with linear regression fits. For Koronis, Proxy_D clearly shows a size-dependent V-shape and a strong positive Orbital Dispersion Gradient (R^2=0.61), indicating significant Yarkovsky drift. Conversely, Proxy_P and Proxy_PD show no meaningful correlation ($R^2 \leq -0.11$), demonstrating that diameter is the primary driver of orbital spreading in this family.

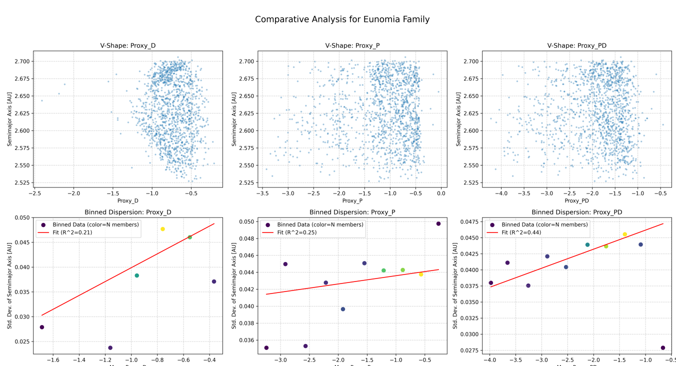

Figure 4. Orbital dispersion analysis for the Eunomia family, illustrating the relationship between semimajor axis spread and Yarkovsky-sensitive proxies: Proxy_D (diameter-based), Proxy_P (spin period-based), and Proxy_PD (combined diameter and spin period). The top row displays individual asteroid data, while the bottom row presents the standard deviation of semimajor axis in binned data with linear regression fits. For the Eunomia family, Proxy_PD best describes the orbital dispersion (R^2=0.44), indicating a significant combined influence of both diameter and spin period on the Yarkovsky-driven spreading.

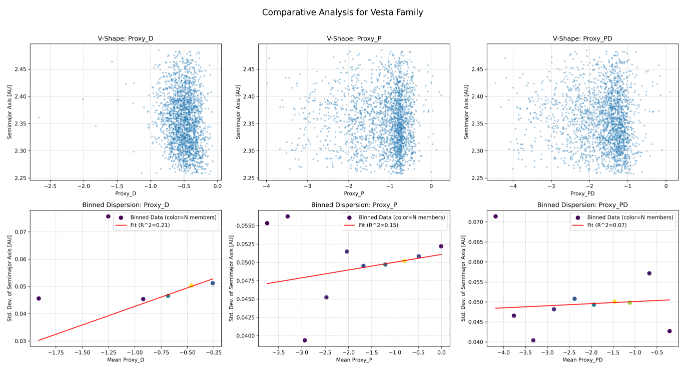

Figure 5. Comparative analysis of the Vesta family's orbital dispersion. The top row displays the distribution of semimajor axis versus three Yarkovsky-sensitive proxies: diameter-based ($\mathrm{Proxy_D}$), spin-period-based ($\mathrm{Proxy_P}$), and combined ($\mathrm{Proxy_{PD}}$). The bottom row quantifies orbital dispersion as the standard deviation of semimajor axis for binned data, with linear regression fits and corresponding R^2 values. For the Vesta family, $\mathrm{Proxy_D}$ shows the strongest, albeit weak, positive correlation with orbital dispersion (R^2=0.21), consistent with an emergent size-dependent Yarkovsky effect. The spin-inclusive proxies exhibit weaker relationships.

superposition of multiple, overlapping collisional events rather than a single, distinct family. This complexity, combined with its relatively young assigned age of 0.5 Gyr (though other studies suggest a much older and more complex history), likely obscures any clear Yarkovsky-driven signature, preventing the formation of a discernible V-shape or linear dispersion gradient.

3.3. *Comparative efficacy of Yarkovsky proxies*

A key objective of this study was to systematically compare the efficacy of different Yarkovsky-sensitive proxies in explaining observed orbital dispersion. The results, summarized in Table 1 and exemplified in Figures 1 to 6, provide valuable insights into the relative importance of asteroid diameter and spin period.

The analysis robustly demonstrates the **overwhelming dominance of the diameter-only proxy ($\mathbf{Proxy_D}$)**. In four out of the six major families examined (Maria, Eos, Koronis, and Vesta), $\mathrm{Proxy_D}$ consistently emerged as the most effective predictor of orbital dispersion, yielding R^2 values ranging from 0.2122 to an exceptional 0.9353. This strongly reaffirms the well-established understanding that asteroid size is the primary physical parameter governing the magnitude of the Yarkovsky effect. Smaller asteroids, represented by higher values of $\log_{10}(1/\mathrm{Diameter})$, experience greater Yarkovsky-driven semimajor axis drift, leading to a larger dispersion from the family's center and consequently, a higher standard deviation of semimajor axis within their respective bins.

Conversely, the **spin-period-only proxy ($\mathbf{Proxy_P}$) consistently failed** to provide a meaningful explanation for orbital dispersion in any of the families. In most cases, it resulted in negative R^2 values, indicating that the linear model based solely on spin period was worse than simply using the mean σ_a as a predictor. This outcome is not entirely unexpected. The Yarkovsky effect's dependence on an asteroid's spin is complex and critically relies on its spin-axis orientation (obliquity), which dictates the direction and magnitude of the thermal force. A fast or slow spin period alone, without knowledge of the obliquity, is an insufficient predictor of the net Yarkovsky drift. Furthermore, observational biases and inherent uncertainties in lightcurve-derived spin periods may contribute to this poor performance.

The **combined spin-diameter proxy ($\mathbf{Proxy_{PD}}$) showed mixed success**. While it was hypothesized to provide a more comprehensive description by incorporating both key parameters, it only outperformed $\mathrm{Proxy_D}$ in the case of the Eunomia family, as seen in Figure 4. In the majority of other families, its performance was either intermediate or poor, sometimes even yielding negative R^2 values. This suggests that a simple multiplicative combination of the inverse of spin period and diameter in a logarithmic scale may not adequately capture the intricate interplay between these parameters for all families. It is possible that for many families, the strong signal from diameter dominates, and the less precise or less directly relevant spin period data (due to unknown obliquities) introduces noise that degrades the predictive power of the combined proxy. However, its success in the Eunomia family indicates that the concept of a combined proxy is not without merit and may be applicable under specific family conditions or with more complete spin state information.

3.4. *Correlation between family age and orbital dispersion*

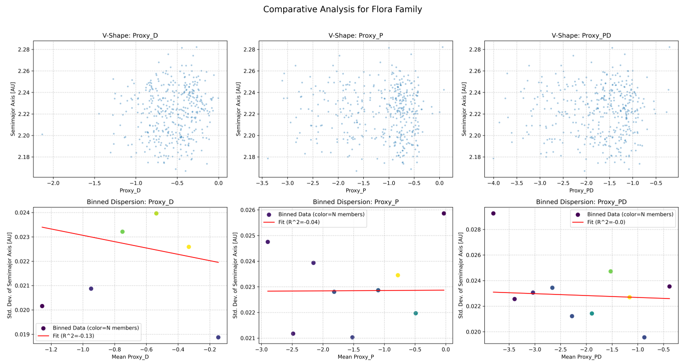

Figure 6. Comparative analysis for the Flora family. The top row displays the semimajor axis distribution against three Yarkovsky-sensitive proxies: $\mathrm{Proxy_D}$ (diameter-based), $\mathrm{Proxy_P}$ (spin period-based), and $\mathrm{Proxy_{PD}}$ (combined). The bottom row shows the binned standard deviation of the semimajor axis against the mean proxy values, with linear regression fits. For Flora, all proxies exhibit exceptionally poor fits (negative R^2 values), indicating the linear dispersion model is unsuitable for describing its orbital evolution, likely due to its complex and diffuse dynamical structure.

A secondary objective of this study was to explore the long-term cumulative effects of the Yarkovsky force by investigating a potential correlation between family age and the measured Orbital Dispersion Gradient. For this analysis, the Spearman rank-order correlation coefficient (ρ) was calculated between the estimated age of each family and the gradient (G) derived from its best-performing proxy (as identified in Table 1).

The analysis yielded a **Spearman correlation coefficient of** $\rho = 0.3714$ with a **p-value of** 0.4685.

The positive value of the Spearman correlation coefficient ($\rho = 0.3714$) is qualitatively consistent with the physical expectation that older families, having been subjected to the Yarkovsky effect for longer durations, should exhibit larger cumulative orbital dispersion, and thus steeper dispersion gradients. However, the high p-value of 0.4685 indicates that this observed positive trend is **not statistically significant** at conventional significance levels (e.g., 0.05). Consequently, based on the current dataset, the null hypothesis (that there is no monotonic relationship between family age and dispersion gradient) cannot be rejected.

Several factors likely contribute to this lack of statistical significance. Primarily, the analysis is severely limited by the **small sample size** of only six asteroid families. Detecting a statistically robust trend with such a limited number of data points requires an exceptionally strong underlying correlation. Furthermore, asteroid family ages, derived from various dynamical models, inherently carry **significant uncertainties**, which were not explicitly propagated in this analysis. This "noise" in the independent variable can easily obscure genuine physical correlations. Finally, the observed orbital dispersion in asteroid families is not solely a function of Yarkovsky-driven drift and age. **Confounding variables** such as the initial velocity dispersion from the family-forming impact, subsequent secondary collisional events, and compositional variations (affecting thermal properties) all contribute to the final observed family structure. These complex factors can mask a simple, direct relationship between family age and the measured dispersion gradient, highlighting the challenges in isolating the sole effect of the Yarkovsky force in real asteroid families.

3.5. *Discussion and limitations*

This study successfully introduced and applied the Orbital Dispersion Gradient (ODG) method, a novel approach that quantifies Yarkovsky-driven orbital dispersion without requiring a precisely defined family center. This method provides a robust framework for systematically comparing the influence of different physical parameters on asteroid family evolution. The results overwhelmingly reaffirm the **dominant role of asteroid diameter** in shaping family structures through the Yarkovsky effect, as evidenced by the consistently high R^2 values for the diameter-only proxy in most families, particularly Maria and Eos (see also Table 1 and Figures 1 and 2). This confirms that, as a first-order approximation, asteroid size is the most critical factor determining the magnitude of Yarkovsky-driven semimajor axis drift.

The limited success of the spin-inclusive proxies ($\mathrm{Proxy_P}$ and $\mathrm{Proxy_{PD}}$) is a notable finding. It does not imply that spin is unimportant for the Yarkovsky effect, but rather that its influence is more intricate than can be captured by simple inverse spin period relationships. The Yarkovsky effect's magnitude and direction are highly dependent on an asteroid's spin-axis orientation (obliquity), which is not included in spin period measurements. The Yarkovksy-O'Keefe-Radzievskii-Paddack (YORP) effect can also significantly alter both spin rates and obliquities over time, further complicating the observed spin state distribution within a family. Without knowledge of obliquity, spin period alone is an incomplete and potentially misleading parameter. Additionally, the inherent noise and observational biases in current spin period databases, often derived from sparse lightcurve data, may contribute to the poor predictive power of spin-based proxies. The unique success of the combined proxy for the Eunomia family suggests that for some families, the combined influence of diameter and spin is indeed relevant, possibly due to specific distributions of obliquities or more precise spin period data for that family (see Figure 4).

The absence of a statistically significant correlation between family age and the measured dispersion gradient underscores the inherent complexities of asteroid family evolution. While the Yarkovsky effect is a fundamental long-term driver of orbital spreading, its signature can be obscured by a multitude of other factors, including the initial conditions of the family-forming impact, subsequent collisional evolution, and variations in thermal properties. This non-result highlights the need for larger samples of well-characterized families and more sophisticated models that can account for these confounding variables to isolate the cumulative effect of the Yarkovsky force over geological timescales.

Limitations of the current study include:

1. **Observational Bias:** The analysis relies on asteroids for which both diameter and spin period data are available. Such asteroids tend to be larger and brighter, leading to a potential bias in the sample towards objects less susceptible to extreme Yarkovsky drift compared to the overall family population.

2. **Model Simplification:** The use of weighted linear regression to model the relationship between proxy values and semimajor axis dispersion is a first-order approximation. The true underlying physical relationship may exhibit non-linearities that are not fully captured by this model.

3. **Data Uncertainty:** While weighted regression accounted for the variance in σ_a estimates, uncertainties associated with individual asteroid diameters, spin periods, and particularly family ages were not formally propagated, which could influence the derived gradients and correlation analyses.

4. CONCLUSIONS

4.1. *Overview of the study*

The Yarkovsky effect, a subtle but persistent non-gravitational force, plays a crucial role in shaping the long-term orbital evolution of asteroid families by systematically dispersing their members in semimajor axis. Quantifying this dispersion and identifying the most effective physical proxies that drive it has historically been challenging, primarily due to the reliance of existing methods on a precisely defined family center. This study addressed these challenges by introducing the Orbital Dispersion Gradient (ODG) method, a novel approach that quantifies the rate of increase in semimajor axis standard deviation (σ_a) with respect to various Yarkovsky-sensitive proxies, thereby circumventing the need for a precise family center.

4.2. *Methodology and data*

Our investigation leveraged a comprehensive dataset of 16,364 asteroids with complete records for diameter, spin period, semimajor axis, and family assignment. We focused on six major asteroid families (Eunomia, Vesta, Flora, Koronis, Eos, Maria), selected for their sufficient member counts to ensure statistical robustness. For each family, three distinct Yarkovsky-sensitive proxies were calculated: a diameter-only proxy ($\log_{10}(1/\text{Diameter})$), a spin-period-only proxy ($\log_{10}(1/\text{Spin Period})$), and a combined spin-diameter proxy ($\log_{10}(1/(\text{Spin Period} \times \text{Diameter}))$). The core of the ODG method involved binning asteroid family members based on their proxy values and calculating the standard deviation of semimajor axis (σ_a) within each bin. Weighted linear regressions of σ_a against the mean proxy value for each bin were then performed to derive the Orbital Dispersion Gradient (G) and assess proxy efficacy using the coefficient of determination (R^2). A secondary analysis investigated the Spearman correlation between family age and the measured dispersion gradients.

4.3. *Key findings*

Our results provide compelling insights into the drivers of Yarkovsky-driven orbital dispersion. The diameter-only proxy (Proxy_D) consistently emerged as the most effective predictor of orbital dispersion in four of the six families examined (Maria, Eos, Koronis, and Vesta), yielding exceptionally high R^2 values, up to 0.9353 for the Maria family. This strong correlation underscores the primary role of asteroid size in governing the magnitude of the Yarkovsky effect. The combined spin-diameter proxy (Proxy_PD) was found to be most effective for the Eunomia family ($R^2 = 0.4366$), suggesting that for this specific family, the interplay of both spin and diameter significantly contributes to the observed dispersion. In stark contrast, the spin-period-only proxy (Proxy_P) was largely ineffective across all families, often yielding negative R^2 values, indicating that spin period alone is a poor predictor of Yarkovsky-driven dispersion without consideration of spin-axis obliquity. The Flora family presented an anomalous case, showing very poor fits for all proxies, likely due to its complex and dynamically diffuse nature. Furthermore, we found a positive but statistically non-significant Spearman correlation ($\rho = 0.3714$, p-value $= 0.4685$) between family age and the measured dispersion gradient, which we attribute primarily to the small sample size of families and inherent uncertainties in family age estimates.

4.4. *Implications and future directions*

This study reaffirms the fundamental understanding that asteroid size is the dominant physical parameter dictating Yarkovsky-driven orbital evolution, consistently explaining a significant portion of the observed semimajor axis dispersion in asteroid families. While the Yarkovsky effect is fundamentally dependent on an asteroid's spin state, our findings highlight the inherent complexity of its influence when only spin period is considered, likely due to the critical role of unknown spin-axis obliquities and the effects of the YORP effect. The success of the combined proxy for the Eunomia family suggests that incorporating spin information can be beneficial under certain conditions, potentially where obliquity distributions are more favorable or spin data are more precise. The lack of a statistically significant correlation between family age and dispersion gradient underscores the challenges in isolating the cumulative Yarkovsky effect over geological timescales due to confounding variables such as initial impact conditions, secondary collisions, and compositional variations. The Orbital Dispersion Gradient method introduced here provides a robust and valuable tool for future studies of asteroid family evolution, particularly for older or more dispersed families where traditional V-shape analysis is ambiguous. Future work could benefit from larger samples of well-characterized families, more precise family age determinations, and the incorporation of asteroid thermal properties and known obliquities where available, to further disentangle the intricate physics governing asteroid family dynamics.