Collective Gravity of Minor Planets in the Outer Solar System

by

Alexander Zderic

B.A., University of Washington, 2015

M.S., University of Colorado, 2018

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Astrophysical and Planetary Science

2021

Committee Members:

Ann-Marie Madigan, Chair

Mitch Begelman

Meredith MacGregor

David Nesvorný

Mark Rast

Daniel Scheeres

Zderic, Alexander (Ph.D., Astrophysical and Planetary Science)

Collective Gravity of Minor Planets in the Outer Solar System

Thesis directed by Prof. Ann-Marie Madigan

Axisymmetric disks of eccentric Kepler orbits are susceptible to the inclination instability, an out-of-plane buckling characterized by exponentially growing inclinations, decreasing eccentricities, and clustering in argument of periapsis. Observations of trans-Neptunian Objects (TNOs), minor planets orbiting around and beyond the orbit of Neptune, reveal a similar structure. Here we explore if the inclination instability could explain these observations if it occurred in the young Solar System within Neptune's primordial Scattered Disk. We describe the mechanism governing the instability using a two-orbit mutual torque model, and we derive an expression for the timescale of the instability as a function of disk mass for realistic astrophysical disks. We find that the inclination instability can occur in disks modelled after the Scattered Disk in the outer Solar System, but we find that precession induced in the outer Solar System by the quadrupole potential of the giant planets can suppress the instability if the total disk mass is not large enough. We find that the primordial Scattered Disk would need to host $\gtrsim 20$ Earth masses of material in order to undergo the inclination instability. Finally, we find that the post-inclination instability disk evolves into a lopsided, ϖ-clustered state which can explain similar observations of the TNOs.

Dedication

To my parents, without whom I would have never made it this far.

Acknowledgements

First and foremost, I would like to thank my advisor, Ann-Marie Madigan. Without her support and guidance this thesis would not have been possible. I'm grateful for the group members past and present that helped contribute to this thesis directly and indirectly. In no particular order, Heather Wernke, Jacob Fleisig, Maria Tiongco, Aleksey Generozov, and Angela Collier all made substantial contributions to this work. I would also like to thank NASA and the NESSF/FINNEST Fellowship for their monetary support of this research.

I would like to thank my parents, James and Libi, for their support throughout my life. Without their help, this path in life would've been closed to me. I'd like to thank Momo and Brian for many great times spent climbing and skiing. Finally, I'd like to thank my girlfriend Lauren and my dog Cat whose love and company has been a great solace during my PhD.

Contents

Chapter

Appendix

Tables

Table

Figures

Figure

Chapter 1

Introduction

This thesis is an anthology of papers exploring the idea that a collective gravitational instability, the inclination instability, could have happened in the young solar system shortly after planet formation. We invoke the inclination instability to explain some current peculiarities observed in the trans-Neptunian region, the area around and beyond the orbit of Neptune.

Here in the Introduction, I motivate the thesis and provide some essential celestial mechanics background. In Section 1.1 I provide secular celestial mechanics background and define important coordinates, in Section 1.2 I give a background covering the dynamical history of the outer Solar System, in Section 1.3 I outline some unexplained observations motivating the thesis research and review some competing theories, and in Section 1.4 I introduce the inclination instability and explain how it might resolve the issues outlined in Section 1.3.

1.1 Keplerian Orbital Elements and Some Celestial Mechanics

All of the dynamics presented in this thesis are near-Keplerian, and I will regularly use the Kepler elements to describe the orbits of bodies. Near-Keplerian systems are systems where, to lowest order, the dynamics are described by the Newtonian two-body problem. All interesting dynamics are a perturbation on top of this background. Near-Keplerian dynamics describe the dynamics in any system where the potential in that system is dominated by a single body, e.g. the Solar System, single star stellar systems, and galactic centers[1] .

---

[1] Where the potential is dominated by a supermassive black hole.

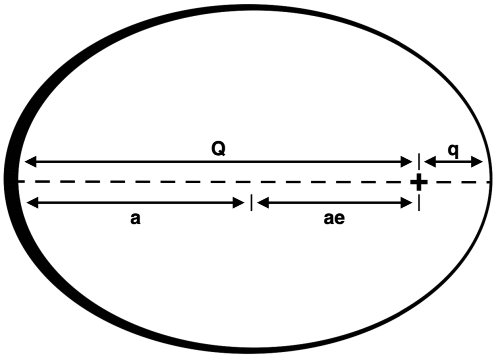

Figure 1.1: Elliptical Kepler orbit with important elements marked. The orbit thickness shows Gauss' approximation (the thickness is inversely proportional to the orbital speed).

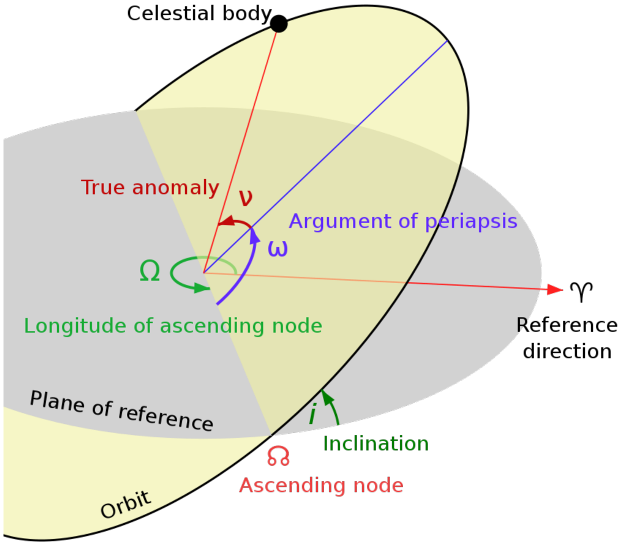

Figure 1.2: A diagram showing the orientation angles for a Kepler orbit in 3D, inclination, i, argument of periapsis, ω, and longitude of the ascending node, Ω, and the true anomaly, ν. Diagram by Lasunncty, distributed under CC-BY-SA-3.0 license.

The standard Keplerian orbital elements are semi-major axis (a), eccentricity (e), inclination (i), argument of periapsis (ω), longitude of the ascending node (Ω), and true anomaly (ν) (see Figure 1.1 and 1.2). These elements fully describe the relative position and velocity of a body w.r.t. a primary (the other body in the two-body problem). Because these elements describe the orbit in the two-body problem, they require two bodies to be specified. Here, unless specified otherwise, the primary is always the central, massive body. Semi-major axis and eccentricity describe the size and shape of the orbit, inclination, argument of periapsis, and longitude of the ascending node describe the orbit's position in 3D space, and true anomaly defines the body's position on the orbit. The three angles i, ω, and Ω are defined w.r.t. a reference plane and direction (see Figure 1.2). In the Solar System this is usually the ecliptic plane and vernal point. Also useful are periapsis and apoapsis[2] , $q = a(1-e)$ and $Q = a(1+e)$ respectively, mean anomaly ($\mathcal{M}$), and the longitude of periapsis, $\varpi = \omega + \Omega$. Periapsis and apoapsis are the closest and further approach of a body on a Keplerian orbit to the primary (see Figure 1.1), $\mathcal{M}$ is an alternative anomaly that specifies the bodies position evenly in time, and ϖ is a compound angle that gives the angular position of the orbits periapsis in the x-y plane if the inclination is small.

In the pure two-body problem, all Kepler elements are constant, except the true anomaly, which circulates as the body orbits. In the near-Keplerian case, the Kepler elements are generally all non-constant, e.g. $a(t)$. In this case, they are called osculating orbital elements. Osculate is an old verb meaning touch, and the osculating orbital elements define the Kepler orbit that touches the body's current position. As the body moves in the near-Keplerian potential, the osculating orbit moves and changes shape to follow the body as it moves along it's a not-quite Keplerian trajectory (see Figure 1.3). The osculating orbital elements give the instantaneous, unperturbed trajectory of the particle.

The osculating orbital elements experience cyclical changes per orbit and slow drifts over many orbits. For example, imagine there is an excess of mass near the periapsis of the osculating orbit, e.g. from a planet close to the sun. Every orbit, the body will pick up excess speed at periapsis due to the additional gravitational attraction of the planet which changes the eccentricity and semi-major axis in a cyclic fashion. On longer timescales, the excess speed at periapsis causes the osculating orbit to slowly precession around the central body. In general, we are interested in the slow, secular evolution of the osculating orbital elements that occurs over many orbits. To do this, we use Gauss' approximation: we average over the body's motion over it's osculating orbit. Formally, this is an

---

[2] Perihelion and aphelion can be used when the Sun is the primary; pericenter and apocenter are also commonly used.

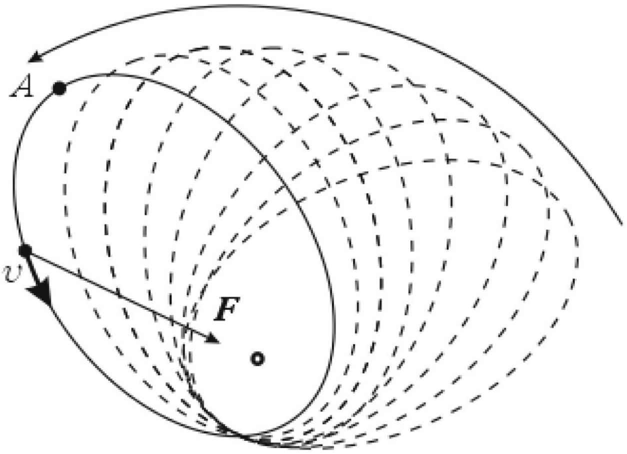

Figure 1.3: An diagram portraying apsidal precession in a osculating Kepler orbit. Reproduced from Kondratyev (2014).

integral over mean anomaly (a.k.a. time), and it's only accurate when the body's osculating orbit is evolving slowly. Geometrically, Gauss' approximation can be pictured as spreading the body's mass over it's orbit-averaged osculating orbit with linear mass density inversely proportional to the body's velocity. The mass density of the resulting elliptic ring is smallest at periapsis and largest at apoapsis (see Figure 1.1). The dynamics of orbit rings like this are governed by orbit-averaged torques which can change the shape and orientation of the ring. In this thesis, we use N-body simulations, but we think of the dynamics as being N-ring.[3]

The orbital angles ω and ϖ are important to the following discussion as apparent clustering in these angles in the outer Solar System are what originally motivated this research. Unfortunately, these angles are somewhat unintuitive, at least geometrically, and a short explanation of why these angles are used is needed. First, the angles i, Ω, and ω are proper rotation angles (the Euler angles). This is useful from a mathematics point a view, the 3D orientation of the orbit is defined by three consecutive applications of the rotation matrix. However, this makes the angles difficult to picture as the axes that the angles are rotated about are themselves rotated by later rotations (i.e. rotations are not commutative). Second, $\mathcal{M}$, ω, and Ω are the angles in the canonical Delaunay action-angle coordinates for the two-body problem (the conjugate momenta/actions for the angles

---

[3] N-ring codes exist, but they suffer from singularities associated with spreading the body out into a ring, i.e. intersecting rings will have nonphysical strong interactions.

are $I = \sqrt{GMa}$, $H = I\sqrt{1-e^2}$, and $H_z = H\cos i$). The longitude of periapsis, ϖ, is one of *modified* Delaunay angles (also canonical) which are used when i and/or e are zero (ω and/or Ω are not well defined in this case). See Morbidelli (2002) for a derivation of the Delaunay coordinates. As canonical coordinates, the Delaunay action-angle coordinates are extremely useful when studying quasi-integrable, near-Keplerian systems from Hamiltonian perspective, and the modified Delaunay coordinates are useful when studying systems with low/no mutual inclination and/or negligible eccentricity (e.g. the planets, and inner Solar System in general). They are not quite as useful for studying near-Keplerian systems from a Newtonian perspective, i.e. our secular torque approach, or when studying a highly 3D systems like the outer Solar System. Still these angles have great staying power in celestial mechanics and are the default coordinates used to describe orbits of bodies in the Solar System.

1.2 Dynamical History of the Outer Solar System

Here we give an overview contextualizing the thesis work. We outline a "Standard Model" for the post-gas dynamical processes that are believe to have shaped the Solar System shortly after the giant planets formed with a focus on the outer Solar System, especially the Trans-Neptunian Objects (TNOs). In Section 1.2.1, we review planet formation theories for the Solar System and how these shaped the orbital configuration of the outer Solar System, in Section 1.2.2 we introduce planetesimal-driven migration of the giant planets, and in Sections 1.2.3 and 1.2.4 we describe how this migration could have instigated significant orbital instability in the giant planets and shaped the small body structures in the outer Solar System.

1.2.1 Solar System Dynamics Initial Conditions

The post-gas dynamics of the Solar System are determined by the initial conditions set by planet formation within the protoplanetary disk. Observations suggest that protoplanetary disk lifetimes are 2-10 Myr (Haisch et al., 2001). In that time, gas within the disk is either accreted onto the star or photo-evaporated by it (Williams & Cieza, 2011). Assuming planet formation in our

Solar System was typical, the gas planets had about that much time to form. Efforts to calculate the age of Jupiter agree with this estimate, suggesting that it reached about 50 Earth masses ($M_\oplus$) in 3-4 Myr (Kruijer et al., 2017). Our giant planets, Jupiter, Saturn, Uranus, and Neptune, probably formed via core accretion; they first formed a solid core of $\mathcal{O}(10M_\oplus)$ then they accreted gas from the surrounding disk (Mizuno et al., 1978; Pollack et al., 1996).

The solid cores of the gas giants started as micron-sized dust within the disk. The growth process from micron-sized dust to $10\,M_\oplus$ planetary cores covers a dizzying orders of magnitude and is driven by different processes at different scales (Armitage, 2020; Youdin & Kenyon, 2013). The first stage in this process is the formation of planetesimals, $\gtrsim$10 km bodies that form the building blocks for the planets. From micron to about meter sizes, bodies rapidly (10^3-10^4 yr at $\sim$1 AU) grow through sticky collisions (Weidenschilling & Cuzzi, 1993; Dominik et al., 2007). Once the bodies are approximately meter-sized, collisions will no longer cause growth (relative velocities are higher and collisions aren't as sticky). In addition, there exists the "meter-sized barrier" where a gas headwind causes rapid ($\sim 10^2$ yr) inward radial drift of solids (Weidenschilling, 1977; Takeuchi & Lin, 2002). Continued growth and the circumvention of the "meter-sized barrier" could be accomplished by the streaming instability, a turbulent, gas-driven process that concentrates solids in cores that gravitationally collapse into planetesimal-sized bodies (Youdin & Goodman, 2005; Johansen et al., 2007). Overall, the growth of micron-sized bodies to planetesimal-sized, is expected to take $\sim 10^5$ yr (Bitsch et al., 2015; Carrera et al., 2015).

Once planetesimals are formed, further growth is driven by the gravitational interaction of the planestimals. Two-body collisions between planetesimals will generally result in a gravitationally bound remnant with an increased mass. For giant planet formation, collisions like this proceed until a solid core of $\sim 10\,M_\oplus$ is formed at which point a hydrodynamic instability is triggered where gas is rapidly accreted in large amounts from the disk (Perri & Cameron, 1974; Mizuno, 1980). The timescale for the growth of the core is a function of orbital period. Jupiter probably formed a gas-accreting core first then Saturn did and so on. The density of solids in the disk also decreases with radius limiting the absolute size of cores formed further from the Sun (Williams & Cieza, 2011).

Because of this, it's thought that Jupiter and Saturn formed large solid cores before Uranus and Neptune, and were able to accrete significant amounts of gas from disk before it dissipated. In fact, it's difficult to get Uranus and Neptune to form at all at their current orbits before the disk was entirely dissipated. One possibility is that they formed closer to the Sun, and migrated outward to their current orbits (Ford & Chiang, 2007). It's also possible for the giant planets to migrate radially inward while still embedded in the gas disk (Kley, 2000). Overall, its likely that the giant planets were in a more compact configuration when they emerged from the protoplanetary disk.

Outside of the orbit of the ice giants (Uranus and Neptune), planet formation proceeded at even slower pace, so that by the time the gas in the disk was dissipated, no large cores had formed. This left a disk of planetesimals outside of Neptune's orbit, a primordial Kuiper Belt (Nesvorný, 2018). This disk likely had a total mass of a few tens of Earth masses, slightly greater than the individual masses of Uranus and Neptune, and the bodies in the disk likely had modest eccentricities and inclinations. This gives us the post-formation, post-gas picture of the Solar System – the gas giants, possibly in a more compact configuration, with an exterior disk of small bodies of total mass of a few tens of Earth masses.

1.2.2 Planetesimal-driven Migration

This configuration, the giant planets interior to a massive exterior planetesimal disk, is not static. Studies show that planetesimals within this disk will scatter with the giant planets causing them to migrate radially. If a planet of mass M ejects a mass m_p of planetesimals at a radius of r, the orbital radius of the planet decreases by a factor ${}^{\delta r}/_{r} \approx -C\,{}^{m_p}/_{M}$, where $C \sim 1$ to conserve angular momentum (Fernandez & Ip, 1984). If the planetesimal is nearly ejected (e.g. placed in the Oort Cloud), the same will happen with a slightly smaller factor C, and if the planetesimal is scattered inward the planetary orbital radius will increase. Simulations show that the outer giant planets, Saturn, Uranus, and Neptune tend to migrate outward, while Jupiter tends to migrate inward (Fernandez & Ip, 1984). The key to this migration pattern is Jupiter's ability to eject planetesimals so efficiently. It takes many scattering encounters for the smaller planets (Saturn, Uranus, and Neptune) to eject a planetesimal. Jupiter, on the other hand, will almost always eject the body on the first encounter. Saturn, Uranus, and Neptune scatter bodies both inward and outward. Some of the outwardly scattered planetesimals will end up in the Oort Cloud. The inwardly scattered bodies interact with the planets interior to the scattering planet. Because Jupiter quickly ejects most planetesimals that interact with it, some of the planetesimals scattered inward by the planets are quickly ejected. This results in an irreversible transfer of angular momentum from the planetesimals to the outer three giant planets moving them outward. Because Jupiter typically ejects particles, it's angular momentum decreases, moving it inward. Jupiter is so massive and on such a tight orbit, that it doesn't migrate far (~0.3 AU for 20 Earth masses of planetesimals ejected), but the lighter and less tightly bound planets, Uranus and Neptune, can migrate as much as 10 AU.

This radial migration is dangerous; it increases the likelihood of a mean motion resonance being set up between planets. Mean motion resonances (MMRs) occur when two bodies in a near-Keplerian potential have orbital periods that form a low order ratio, e.g. if the inner body has an period $^1/_3\,P$ and the outer body has a period $^2/_3\,P$ they are in a 3:2 MMR. Bodies in an MMR have correlated motions. This leads to long-term, coherent gravitational interactions which can cause significant secular orbital evolution[4]. In an MMR, the two body's motions can be correlated such that close-encounters are avoided (phase-protected), or the relative motions can are correlated such that catastrophic close-encounters repeatedly occur. During radial migration, the planets sweep through a range of orbital periods, which increases the likelihood that a resonance condition will be met at some point during migration. This is especially likely for planets that migrate in opposite radial directions (e.g. Jupiter and Saturn) as their period ratio sweeps through a wider range of values.

---

[4] MMRs are the fraternal twin of Gaussian secular torques: they're both mechanisms that can drive long-term orbital evolution, but the position of the bodies on the orbits is critical for MMRs and completely irrelevant for Gaussian torques.

1.2.3 Giant Planet Instability

It's possible that planetesimal-driven migration in the young Solar System did result in an MMR between Jupiter and Saturn. This was originally proposed by Tsiganis et al. (2005) and dubbed the Nice Model. Using a suite of N-body simulations, the researchers tried to explain the relatively large eccentricities and inclinations of the orbits of the giant planets[5] by instigating a period of instability and close encounters between the giant planets. In their model, Tsiganis et al. (2005) placed the giant planets in a more compact configuration with the outer ice giant at $\lesssim 20$ AU and Saturn was placed just inside of Jupiter's 2:1 resonance. They also added a disk of planetesimals of total mass 30-50 $M_\oplus$ extending from just outside the outermost giant planet out to 30-35 AU. The giant planets were initialized on very circular and coplanar orbits.

During the simulations, the planetesimals slowly migrate Jupiter inward and Saturn, Uranus, and Neptune outward (see Figure 1.4). Saturn inevitably hits Jupiter's 2:1 mean motion resonance, and their eccentricities increase as they "jump" across the resonance. Now with larger eccentricities, Jupiter and Saturn exert significant secular torques on the Uranus and Neptune. These torques dramatically increase Uranus and Neptune's eccentricities to the point that their orbits cross Saturn's orbit (orbit crossing is also facilitated by the initial compactness of the system). A period of instability involving strong planet-planet scattering occurs, wherein the inclinations of the giant planets are increased, and the ice giants are scattered outward[6], deep into the disk of remaining planetesimals. The planetesimal disk is rapidly dispersed, which dampens the eccentricities of the ice giants, and starts a period of fast radial migration for all planets.

The model was successful in explaining the current eccentricities and inclinations of the giant planets. It was retroactively able to describe a variety of other Solar System features such as the Late Heavy Bombardment[7] and the orbital structure of the Jupiter's Trojans (Gomes et al., 2005a; Morbidelli et al., 2005; Nesvorný et al., 2007; Levison et al., 2011). However, the success rate of the

---

[5] Planet formation models predict low eccentricities and inclinations due to interactions with the disk (Pollack et al., 1996; Lubow et al., 1999).

[6] About half the time the two ice giants swap positions.

[7] Interestingly, evidence for the Late Heavy Bombardment has recently weakened (Michael et al., 2018; Hartmann, 2019).

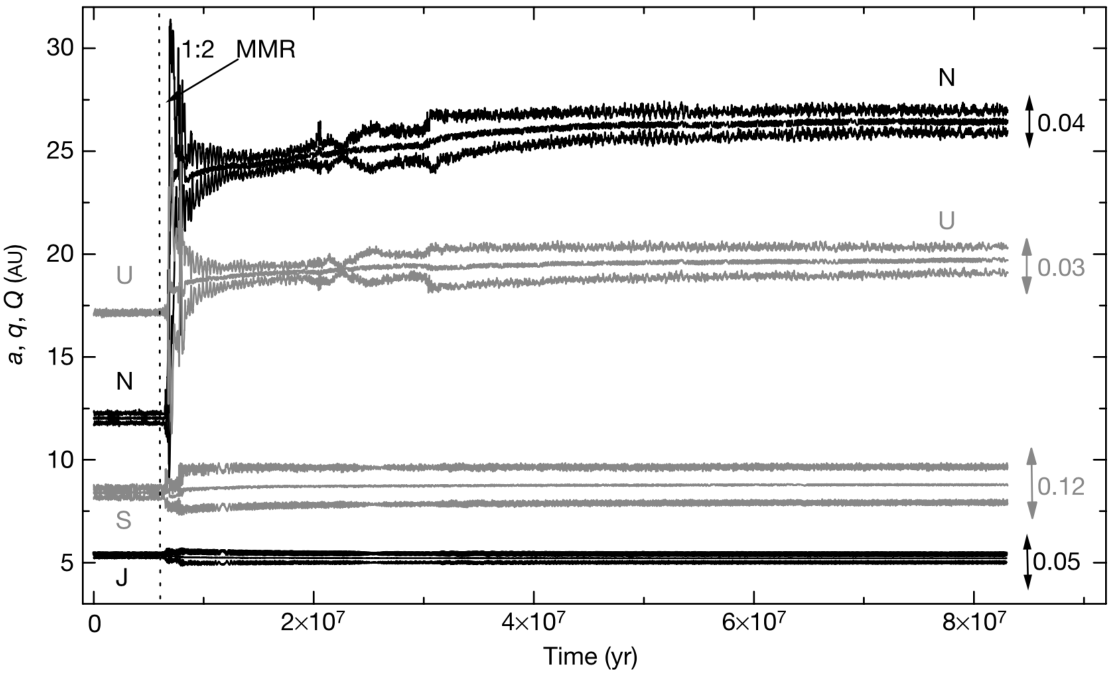

Figure 1.4: Reproduced from Tsiganis et al. (2005). Semi-major axis (a), perihelion (q), and aphelion (Q) of the Jupiter, Saturn, Uranus, and Neptune in one iteration/simulation of the Nice model. The vertical dashed line marks the point where Jupiter and Saturn enter a 1:2 mean motion resonance due to planetesimal-drive migration. This increases Jupiter and Saturns eccentricities which then torque the ice giants, Uranus and Neptune, to high eccentricities and crossing orbits. This destabilizes the system resulting in planetary encounters that scatter Neptune and Uranus into the planetesimal disk at larger semi-major axes. Planetesimal scattering damps the eccentricities of the ice giants, re-circularizing their orbits.

model per simulation was low (33% of simulations). The remaining 67% of simulations either eject one of the ice giants or don't produce high enough proper eccentricities and inclinations. The Nice model had a few other problems too. Notably, it did not properly produce the current structure of the terrestrial planets (Brasser et al., 2009). In addition, studies of the earlier, gas-driven migration of Jupiter and Saturn showed it was unlikely that they emerged from the protoplanetary disk near the 2:1 resonance. Instead these studies found that Jupiter and Saturn probably skipped over the 2:1 resonance became locked in a 3:2 resonance while gas was still present (Masset & Snellgrove, 2001; Morbidelli & Crida, 2007; Pierens & Nelson, 2008). One specific and influential model of this is the Grand Tack Model (Walsh et al., 2011). In this model, Jupiter migrates inward to 1.5 AU through interactions the the gas in the protoplanetary disk. Saturn also migrates inwards, passes over the

2:1 resonance with Jupiter and locks in the 3:2 resonance. Jupiter and Saturn then clear a shared gap in the disk, which reverses the migration direction (tacking), and they then migrate outward to ∼5 AU. This model helps explain terrestrial planet formation by truncating the protoplanetary disk at 1.5 AU.

Other giant planet instability models have been proposed since the original Nice model, fixing some of the problems with that model. One refinement was the introduction of a jumping Jupiter, i.e. Jupiter undergoes a discontinuous jump in semi-major axis. There are a variety of reasons that a jumping Jupiter is desired (see Morbidelli et al. (2009); Brasser et al. (2009); Morbidelli et al. (2010)). Jupiter's "jump" is attained by a strong scattering interaction with one of the ice giants during instability. However, a problem with jumping Jupiter models is the ice giant is often ejected as a result of the interaction. One solution to this is to include an extra ice giant, a 5th giant planet. From a planet formation perspective, this is possible (Ford & Chiang, 2007; Izidoro et al., 2015). This extra ice giant, closest to Saturn, is excited by resonance with Saturn and ultimately ejected by Jupiter. Nesvorný & Morbidelli (2012) looked at a large suite of giant planet instability simulations with 4, 5, and 6 planets. They found that four planet models did not accurately reproduce the current Solar System structure, while the five planet models did, suggesting that our Solar System lost a planet shortly after it's formation.

1.2.4 Details of Neptune Migration

Of all the planets, Neptune is believed to have migrated the furthest from planetesimal interactions. Neptune's current orbit is at ∼30 AU, but it's orbit just after the departure of gas is believed to have been at ∼20 AU. By scattering planetesimals, it migrated out 10 AU through the bulk of the disk. While it migrated, it dispersed much of the disk, sending some of the planetesimals to the Oort cloud, ejecting some itself, and scattering others inward to ultimately be ejected by Jupiter. Many of the planetesimals that did not meet one of these fates ended up in the Kuiper Belt where they remain today. In fact, it is the current orbital structure of the Kuiper Belt, or, more generally, the Trans-Neptunian Objects (TNOs) that provides the best evidence for Neptune's

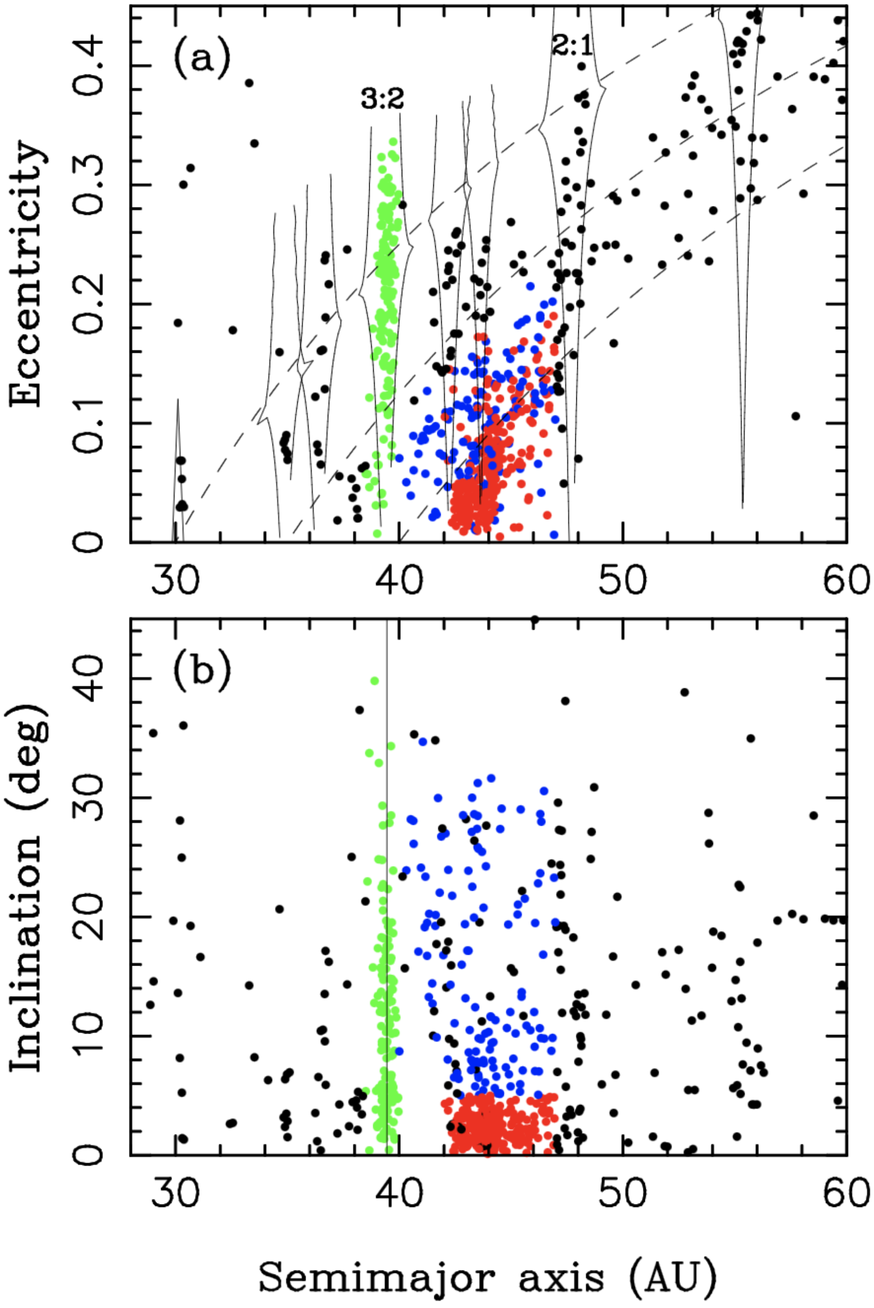

Figure 1.5: Inclination and eccentricity vs. semi-major axis of observed bodies (3 or more oppositions) in the outer Solar System. Mean motion resonances are shown with brackets in the top panel. The green bodies are "plutinos", bodies in Neptune's 3:2 resonance. The blue and red bodies are the Hot and Cold Classicals respectively. Bodies not inside a resonance or part of the classical Kuiper Belt are Scattered Disk objects. The dashed grey lines in the top panel mark perihelion of 30, 35, and 40 AU. Like Pluto, there are numerous other resonant bodies that cross Neptune's orbit (i.e., bodies that lie above the first grey dashed curve). Reproduced from Nesvorný (2018).

long-range migration (see Figure 1.5, reproduced from Nesvorný (2018)). This structure is impressive, and within it are variety of distinct dynamical populations. However, when Neptune migration was first proposed by Malhotra (1993), there were only a handful of discovered TNOs. In fact, Malhotra (1993) was trying to explain the orbit of just one TNO, Pluto.

Since it's discovery in 1930[8] , Pluto has been delightfully confusing to astronomers. Pluto was observed by Clyde Tombaugh at the Lowell Observatory in Arizona (Tombaugh, 1946; Levy, 1991). Clyde was actually looking for Planet X, a ∼7 Earth mass planet proposed by Percival Lowell to explain the apparent (but erroneous) unexplained deviations in the motion of Neptune and Uranus (Standish, 1993). The Lowell Observatory was founded by Percival with the specific intent to search for Planet X. Unfortunately, Percival died in 1916 before he could complete his search, and Clyde was hired to continue the search (see Batygin et al. (2019) for a review of planet searches beyond Uranus). It was immediately obvious that Pluto was not Planet X; it was unresolved and far to dim to be 7 Earth masses[9] (Slipher, 1930; Nicholson & Mayall, 1931). In addition, Pluto had a strangely inclined ($i = 17^\circ$) and eccentric ($e = 0.25$) orbit different than the other planets (Nicholson & Mayall, 1931). Pluto's orbit is so eccentric, that it actually crosses the orbit of Neptune. Typically, crossing orbits are unstable; assuming the motion of the bodies on their two orbits are uncorrelated, crossing orbits necessitate close encounters and strong scattering events. However, Pluto and Neptune's motion is correlated. Pluto is in a 2:3 mean motion resonance with Neptune (Cohen & Hubbard, 1965; Milani et al., 1989). This resonance is phase protected: the relative phase of Pluto and Neptune on their orbits is such that close encounters are avoided even though their orbits overlap. This resonance and Pluto's large eccentricity are not expected from in-situ formation. It is this resonance that Malhotra (1993) sought to explain.

Malhotra's idea was to have Neptune start 5-6 AU closer to the Sun at 24-25 AU and to place Pluto on a nearly circular orbit slightly outside of Neptune's 3:2 resonance (at ∼32-33 AU). Using a

---

[8] There was a ∼60 year gap between the discovery of the first Kuiper Belt object, Pluto (1930), and the second, Albion (1992) (Jewitt et al., 1992).

[9] The mass today is known to be 0.002 Earth masses, almost 4 orders of magnitude smaller than was expected (Christy & Harrington, 1978; Buie et al., 2006).

prescribed migration formula (smooth e-folding migration), she migrated Neptune outward to it's current semi-major axis. As Neptune migrated, it's resonances also migrated, sweeping through the outer Solar System ($P \propto a^{3/2}$). Using linear perturbation theory and numerical simulations, she found that Pluto could be swept up into Neptune's 3:2 resonance and it's eccentricity subsequently pumped up by Neptune. Renu recognized that the resonance capture method should work for other planetesimals besides Pluto, and, in Malhotra (1995), she predicted that all of the low-order resonances of Neptune should be populated.

Some of the details of Malhotra's original proposal, notably the migration distance and timescale, were off, but the concept as a whole has been extremely successful at explaining the observed TNOs. In Figure 1.5, there are three major TNO populations, the Resonant Bodies, the Classical Delt, and the Scattered Disk. Each of these populations has a formation story related to the Neptune's migration (Gladman et al., 2008). The resonant bodies are TNOs that are in one of Neptune's mean motion resonances. The 3:2 resonance is the largest resonant population (green points in Figure 1.5), but the 2:1 and 5:2 are also quite large (Malhotra, 2019). As mentioned, Pluto is in the 3:2 resonance, and objects within this resonance are referred to as plutinos. Like Pluto, all resonant bodies were swept up and trapped in Neptune's resonances while Neptune migrated outward through the planetesimal disk (Levison & Morbidelli, 2003; Gomes, 2003; Hahn & Malhotra, 2005). The resonant populations have relatively high eccentricities, a result of resonant trapping. In Malhotra (1993), Neptune is migrated $\sim$5 AU with a prescribed e-folding timescale of 1.5 Myr. However, more recent studies have found that the migration distance needs to be further ($\sim$ 10 AU) and the timescale longer ($\gtrsim$ 10 Myr) (Nesvorný, 2015). In addition, the smooth migration in Malhotra (1993) tends to overestimated the resonant population size compared to the classical belt. Adopting a grainier migration scheme, consistent with Neptune scattering with 1000-4000 Pluto-sized planetesimals while it migrated, fixes this problem (Nesvorný & Vokrouhlický, 2016). Also, in Malhotra (1993), Neptune is given an initially circular and co-planar orbit, and more recent, post-Nice model studies of Neptune migration have looked at resonant population formation with an eccentric and inclined migrating Neptune (Levison et al., 2008; Batygin et al., 2011).

The classical belt is composed of non-resonant bodies with $a \in [40, 50]\,\mathrm{AU}$ and perihelia greater than ~ 35 AU. The classical belt is the largest Kuiper Belt population (blue and red in Figure 1.5). The high perihelia of these bodies guarantees that close encounters with Neptune are avoided. The inclination distribution of the classical belt is bimodal, and objects with $i > 5^\circ$ are Hot Classicals (HCs, blue points in Figure 1.5) while objects with $i < 5^\circ$ are Cold Classicals (CCs, red points in Figure 1.5). The CCs appear different than other Kuiper belt objects. They have higher albedos (Brucker et al., 2009) and a steeper mass distribution (Bernstein et al., 2004), and they're redder (Tegler & Romanishin, 2000) and often in wide binaries (Noll et al., 2008). It's believed that the CCs formed in situ, unlike rest of the Kuiper Belt. Although the HCs fill a similar orbital space as the CCs, they are actually closely related to the resonant population. Gomes (2003) and Gomes et al. (2005b) describe a mechanism to form the HCs. First, bodies in the planetesimal disk are scattered during Neptune's migration attaining large eccentricities and inclinations. Then, they can enter one of Neptune resonances where they undergo secular dynamics that can decrease their eccentricity (e.g. Kozai-Lidov cycles). Finally, either these cycles reverse and the body eventually stabilizes in the resonance, or, because Neptune is migrating, the body drops out of Neptune's resonance with a low e, HC orbit (Nesvorný & Vokrouhlický, 2016).

Finally, the Scattered Disk Objects (SDOs) are TNOs that are actively scattering off of Neptune, and, unlike the resonant population and the classical belt, these bodies have short dynamical lifetimes (Duncan & Levison, 1997). The SDOs orbits have perihelia near Neptune's orbit, but they are not protected from strong interactions with Neptune by a resonance. They have large semi-major axes, up to hundreds of AU, and they're very dynamically hot, often having extreme eccentricities and inclinations. The scattered disk objects are presumably the remnants of a continually eroding population of planetesimals not in one of Neptune's resonances that were scattered out of the primordial Kuiper Belt when Neptune migrated.

Neptune migration models can also be used to constrain features of the original planetesimal belt. As already mentioned, the planetesimal belt needed to have 1000-4000 Pluto-sized bodies in order to explain the relative sizes of the HCs and the plutinos (Nesvorný & Vokrouhlický, 2016).

In addition, the total mass and density distribution of the planetesimal belt can be inferred from the distance and timescale of migration (Hahn & Malhotra, 1999). Gomes et al. (2004) found two migration regimes determined by the mass density of the disk. Below a critical value, Neptune migration will stall shortly after it starts and fail to clear the disk, and, above that value, Neptune will migrate through the entirety of the disk. Getting Neptune to migrate $\sim 10\,\mathrm{AU}$ requires a disk mass > 15 AU (above the critical mass density) and getting the migration timescale $\gtrsim 10\,\mathrm{Myr}$ requires a disk mass $\lesssim 20\,M_{\oplus}$ (Nesvorný, 2018). A disk mass of 15-20 $M_{\oplus}$ is consistent with current theories. An awkward problem here is actually getting Neptune to *stop* migrating at 30 AU in a disk with a super-critical mass density. This requires some mechanism to truncate the planetesimal disk at 30 AU.

1.2.5 Summary

Overall, it's likely that the giant planets in the Solar System formed in a more compact and resonant configuration (possibly with an extra ice giant) with a massive exterior belt of planetesimals. An orbital resonance between Jupiter and Saturn likely occurred which resulted in significant orbital restructuring of the giant planets. During this time, Neptune migrated ∼10 AU through the disk of planetesimals, dispersing most of it, and sculpting the rest into the impressive structure we observe today. A more comprehensive review of this picture is available in Nesvorný (2018).

1.3 Outstanding Issues

The giant planet instability and Neptune migration models have done wonders for understanding the current structure of the Solar System, especially the TNO populations. However, some phenomena remain unexplained, and the central focus of this thesis is to invoke the inclination instability (Section 1.4) to explain them. Some important observed anomalies for this theory are ω- and ϖ-clustering (§1.3.1), (highly)-detached bodies (§1.3.2), extreme high inclination bodies (§1.3.3), and the perihelion gap (also §1.3.3). In addition, in §1.3.4 I will discuss two theories competing with and parallel to our inclination instability theory, the Planet Nine theory and an apsidally-aligned disk theory.

Unlike the Classical Belt and Resonant Populations which live at low semi-major axes (30-50 AU), the Scattered Disk extends out to hundreds of AU. The observed peculiarities discussed here occur in bodies in the Scattered Disk in bodies with $a \gtrsim 150\,\mathrm{AU}$. In fact there is a special designation, the Extreme Trans-Neptunian Objects (ETNOs), which separates these bodies from the other TNOs. The ETNOs are defined to have $a \gtrsim$ 150-250 AU and $q \gtrsim 37\,\mathrm{AU}$ (Kavelaars et al., 2020). Within this population there is a small but extremely unusual subpopulation, the Sednoids, which are named after Sedna, the first body discovered in the population (Brown et al., 2004a). The Sednoids are ETNOs that are significantly detached from Neptune ($q > 50\,\mathrm{AU}$).

1.3.1 Orbital Angle Clustering

The orbital angles of the ETNOs may be clustered. While reporting the discovery of the second Sednoid, 2012 VP_{113} (nicknamed Biden), Trujillo & Sheppard (2014) noticed that the argument of perihelion (ω) of the ETNOs appeared clustered around $\omega \approx -20^\circ$ (see Figure 1.6). In the discovery paper for the inclination instability, Madigan & McCourt (2016) reported that they could explain this clustering via the collective gravity of planetesimals. Indeed, it was the results of Trujillo & Sheppard (2014) that started the line of inquiry followed in this thesis. While proposing the existence of a 9th, Neptune-like planet, Batygin & Brown (2016a) noticed that the longitude of perihelion (ϖ) of the ETNOs also appeared clustered.

These angles, ω and ϖ, aren't particularly intuitive and the significance of their clustering isn't immediately clear. Argument of periapsis (ω) is a Kepler element, and it's the angle from the periapsis of the orbit to the ascending node (see Figure 1.2 for a refresher). Because ω is a rotation angle, it's meaning with regards to the actual physical orientation of the orbit is generally dependent on the other rotation angles i and Ω[10]. However, we are lucky that the clustering in ω occurs around 0° in this case. Trujillo & Sheppard (2014)'s observation says that the ETNOs preferentially pass from below the ecliptic to above the ecliptic at pericenter. Said another way, all the ETNO orbits

---

[10] Two bodies with the same ω could have very different orbital orientations if they have different i and Ω.

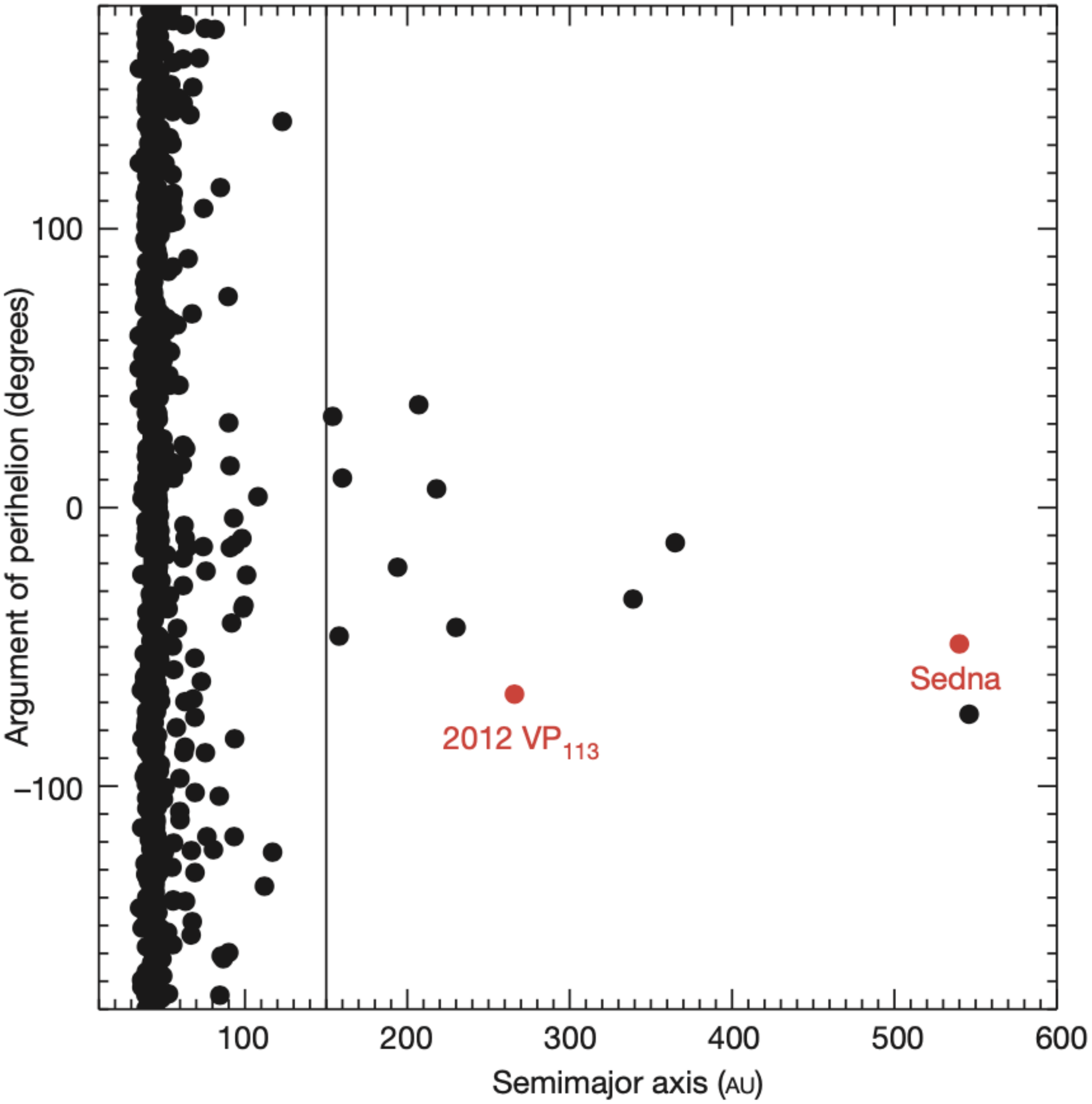

Figure 1.6: Reproduced from Trujillo & Sheppard (2014), ω-clustering in the ETNOs. The Sedna and Biden are marked in red. The clustering has blurred some with more recent observations, but there is still a puzzling lack of orbits at ±180° (Kavelaars et al., 2020)

roll around their major axes in the same direction. Longitude of periapsis (ϖ) clustering is a little easier to understand. If $i \sim 0^{\circ}$[11], ϖ is the angle between periapsis and some reference direction. Therefore, ϖ-clustering is the physical clustering of periapses; ϖ-clustered orbits all "point" in the same direction. Angular clustering like this is not expected from the standard Neptune migration history. Scattering is a random process, and it should not be able to preferentially roll orbits or cluster them in physical space. Thus, ω and ϖ-clustering demand additional theories.

However, the statistical significance of ω- and ϖ-clustering is hotly debated because TNO observational surveys are deeply biased (Shankman et al., 2017; Lawler et al., 2017; Brown, 2017; Brown & Batygin, 2019; Kavelaars et al., 2020; Napier et al., 2021). It's extremely difficult to observe TNOs. TNOs are observed with reflected light so their apparent magnitude drops as r^{-4} instead of r^{-2}, they're very small and often unresolved, and, finally, due to their distance from the Sun, they move so slowly ($v^2 \propto {}^{1}\!/_{r}$) it can be difficult to differentiate them from background sources. Most TNOs and all ETNOs, are observed at or very near their perihelion, when they have the highest apparent magnitude and are moving the fastest. Observational surveys have to dig very deep to find TNOs, typically looking at small sections of the sky during periods of the year where seeing is the best. In addition, most surveys bias their observations to look at the ecliptic because this is where they are most likely to observe new bodies. Because of this, the announcement of ω- and ϖ-clustering was met with some skepticism by observers. Recent surveys like the outer Solar System Origins Survey (OSSOS) have carefully characterized their observational biases, and the results from these surveys have cast the significance of ω- and ϖ-clustering into doubt. In Shankman et al. (2017), OSSOS found that the ETNOs discovered by the survey had ω- and ϖ-that were consistent with a uniform distribution if the observational biases of the survey were properly accounted for. However, OSSOS's conclusions only hold for the handful of ETNOs they discovered during their survey, and, using an unorthodox method, Brown (2017) and Brown & Batygin (2019) proposed that ϖ-clustering in the ETNOs was statistically significant even if it wasn't in the OSSOS

---

[11] This is not true in the outer Solar System, however, in this case, ϖ-clustering does still reflect a physical clustering of periapses.

sample. Recently, Napier et al. (2021) combined the ETNOs found in OSSOS, the Dark Energy Survey (DES), and a survey by Sheppard and Trujillo into a single sample, and, accounting for the biases in those surveys, found that the ω and ϖ of the ETNOs found in those surveys is statistically consistent with a uniform distribution. However, Brown & Batygin (2021) recently rebuffed this, saying that clustering in the ETNOs is still significant at the 99.6% level.

1.3.2 Detached bodies

The scattering history of bodies in the scattered disk requires that those bodies have perihelia near the orbit of Neptune. In a near-Keplerian potential, planetary scattering can change the size (a) and shape (e) of an orbit, but the pre-scattered orbit and the post-scattered orbit must intersect[12] . However, the ETNOs have perihelia detached from the orbit of Neptune $q \gtrsim 37\,\mathrm{AU}$, with some, like the Sednoids, being extremely detached. If ETNOs are believed to have been scattered out to their current orbits when Neptune migrated through the left-over planetesimal disks, how did they detach their orbits from Neptune?

Two existing explanations can explain many of the detached TNOs. The first has already been mentioned: resonant dropout during grainy Neptune migration. The mechanism described in Gomes (2003) that forms the Hot Classicals combined with the grainy migration from Nesvorný & Vokrouhlický (2016) can create large a and large q bodies (Nesvorný et al., 2016; Kaib & Sheppard, 2016). Neptune has resonances within the scattered disk extending out to large a. These resonances are weaker than the lower-order resonances closer to Neptune like the 3:2 and the 2:1, but they are still expected to be populated. As mentioned, bodies within a phase-protected resonance aren't strongly scattering, and they can undergo secular Kozai-Lidov oscillations. If the initial inclination of the orbit before it entered the resonance is sufficiently high[13] , it can undergo large eccentricity oscillations that detach it from Neptune. If the resonance moves while this happens (e.g. Neptune migrates to a different a), the orbit can be left "fossilized" on a detached orbit. The observation of

---

[12] Gravitational scattering cannot discontinuously transfer a body from one orbit to another

[13] Neptune scattering is sufficient to raise i to high enough values.

some detached objects near resonances supports this (Pike et al., 2015; Lawler et al., 2019).

A second explanation is diffusion inward from the Inner Oort Cloud. The Inner Oort Cloud is a fairly mysterious population of bodies that spans from roughly 10^3 to $10^4\,\mathrm{AU}$. The bulk of the Oort Cloud is isotropic, but the inner portion is thought be be more disk-like. The Oort Cloud is in essence a huge population of detached Scattered Disk objects. In their case, they were scattered out onto very long-period orbits where interactions with passing stars and galactic tides were sufficient to raise their perihelia and detach them from Neptune. In the case of the detached, high a ETNO 2013 SY_{99} ($q = 50\,\mathrm{AU}$, $a \approx 730\,\mathrm{AU}$) long-term simulations show that weak scattering interactions with Neptune are sufficient to diffuse 2013 SY_{99} to and from the Inner Oort Cloud (Bannister et al., 2017). Diffusion out to large a in the Inner Oort Cloud then allows galactic tides to detach/reattach SY_{99} from/to Neptune's orbit.

These mechanisms cannot explain the most detached ETNOs, the Sednoids (see Figure 1.7). The Sedna and Biden have such high perihelia and low semi-major axes that it's unlikely they diffused inward from the Inner Oort Cloud though interactions with Neptune. In addition, they actually have relatively low inclinations ($i < 25°$) making resonant dropout also unlikely. Long-term orbit integrations of the third Sednoid, Leleākūhonua, including the giant planets and galactic tides show that it's orbit is also stably detached from Neptune (Sheppard et al., 2019).

1.3.3 Extremely High Inclination Bodies, and the Perihelion Gap

A handful of TNOs have extremely high inclinations ($i > 45°$) (Gladman et al., 2009; Chen et al., 2016; Becker et al., 2018). These inclinations are too high to be explained purely by planetary scattering (Kaib et al., 2019). The Oort Cloud could be one source for these high inclination TNOs, but it can't account for all of them. A distant $\sim 5\,M_\oplus$ planet can reproduce the number of observed high inclination bodies, but it produces too broad of an inclination distribution.

Recently, Kavelaars et al. (2020) showed and Oldroyd & Trujillo (2021) proved that there is a statistically significant perihelion gap between Sednoids and the rest of the ETNOs (see Figure 1.7). The Sednoids are so detached and difficult to observe that observational surveys should've found

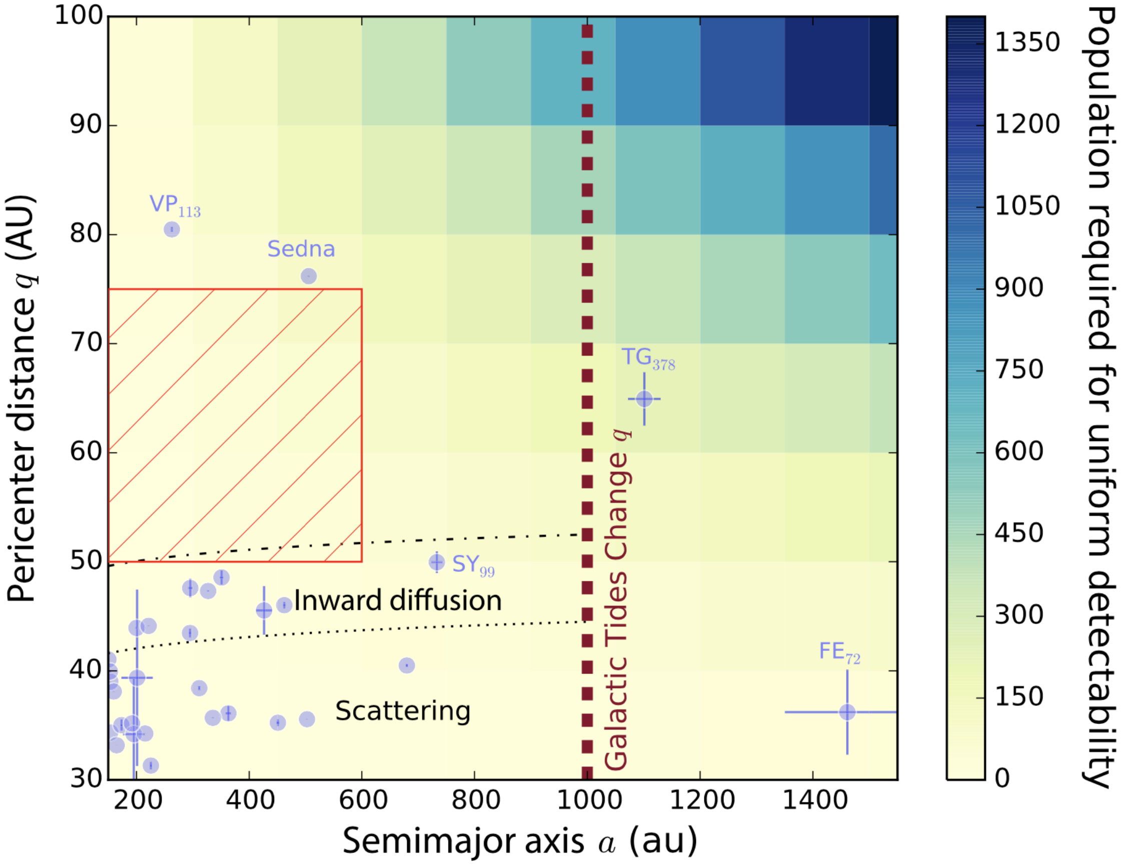

Figure 1.7: Reproduced from Kavelaars et al. (2020), pericenter vs. semi-major axis for the ETNOs. Bodies with pericenters greater than ∼ 40 AU are detached from the orbit of Neptune. There is a gap in pericenter between the rest of the TNOs and the Sednoids. This gap is statistically significant, i.e. there is a paucity of bodies in this region.

orbits with slightly lower perihelia if those orbits were there. Kavelaars et al. (2020) found that an distant ice giant could not account for this gap, but Oldroyd & Trujillo (2021) found that, in some cases, a distant giant planet could explain the gap.

1.3.4 Competing Theories

A few theories have been put forward to resolve some of these unexplained observations. The most well known is the Planet Nine (P9) theory. This theory was originally put forward on the basis of N-body simulations by Batygin & Brown (2016a) to explain ω- and ϖ-clustering and detachment. The original iteration of this theory proposed that a $\gtrsim 10 M_\oplus$ planet on a distant, inclined, and eccentric orbit ($a \sim 700\,\mathrm{AU}$, $e \sim 0.6$, $i \sim 30^\circ$) could explain the observed ω- and ϖ-clustering and detachment in the ETNOs. It was found that a distant planet like this could explain a few other Solar System features as well, such as the solar obliquity and the highly inclined TNOs (Bailey et al., 2016; Batygin & Brown, 2016b). A combination of secular dynamics and mean motion resonances is responsible for the clustering and detachment of the TNOs by P9 (Batygin & Morbidelli, 2017). Extensive observational searches have thus far failed to find P9, and the preferred orbital elements have changed over time with current estimates suggesting P9 is smaller, closer, more circular, and less inclined ($M_{P9} \sim 6\,M_\oplus$, $a \sim 380\,\mathrm{AU}$, $e \sim 0.2$, and $i \sim 16^\circ$), further exacerbating the tension with failed observational surveys[14] (Brown & Batygin, 2021). In addition, the origin of P9's orbit is difficult to explain (Batygin et al., 2019). At that distance, in-situ formation is extremely unlikely (not enough mass, not enough time, and too dynamically excited). Other formation theories involve either capturing P9 in the birth cluster or forming P9 much closer to the Sun, scattering it outward (e.g. the 5 planet giant planet instability models), and detaching it with perturbations from passing stars or from a planetesimal disk. These formation pathways are possible but statistically unlikely (Li & Adams, 2016; Mustill et al., 2016; Parker et al., 2017).

Sefilian & Touma (2019) proposed an alternative theory. They found that self-gravity in a distant, large ($\sim$10 $M_\oplus$), and apsidally-clustered disk of planetesimals could also explain the

---

[14] In addition, the updated P9 orbit can no longer explain the solar obliquity.

observed angular clustering and detachment. This disk is expected to have a moderate eccentricity, $e \sim 0.2$, and, like Planet Nine, to be ϖ-clustered *opposite* to the observed clustered ETNOs. Similar to a P9, they found that an eccentric lopsided disk of small bodies like this shepherd a smaller population of more eccentric anti-aligned orbits, the observed ETNOs. They found that the disk was self-consistent, stabling trapping aligned, moderately eccentric orbits in the main part of the disk, and shepherding anti-aligned, higher eccentricity orbits. Interestingly, this theory has some similarities to the Planet Nine theory, if the mass of Planet Nine were spread out into a disk of smaller bodies. This theory suffers from the same formation problems as Planet Nine. A disk of the mass described in Sefilian & Touma (2019) couldn't have formed in situ at it's current radius. It's moderate eccentricity and detachment from the inner Solar System requires a dynamical explanation.

1.4 The Inclination Instability: A Collective Gravity Solution

Our theory explaining the observational peculiarities outlined in Section 1.3 is that the disk of scattering planetesimals could've undergone a self-gravitational instability, the inclination instability, during Neptune's migration. This theory was originally put forward in Madigan & McCourt (2016) and subsequently developed in Madigan et al. (2018b) (Chapter 2), Zderic & Madigan (2020) (Chapter 3), and Zderic et al. (2020, 2021) (Chapter 4). The inclination instability is a collective instability that occurs in disks of eccentric near-Keplerian orbits, and it's driven by the mutual secular torques between orbits within the disk (Madigan et al., 2018b). It's named after the characteristic exponential growth in orbital inclination seen during the instability.

In Figure 1.8, we show the inclination, eccentricity, and argument of perihelion of bodies in an inclination-unstable, idealized disk of 400 particles with a total mass equal to a Jupiter mass. The bodies in the disk all have approximately the same initial semi-major axis, the same eccentricity ($e = 0.7$) and inclination ($i = 10^{-3}\,\mathrm{rad}$) and the disk is axisymmetric. Time units are in orbital periods, $P \propto a^{-3/2}$. In Figure 1.9, we show an orbit diagram of the initial conditions of the disk and the mid-instability, buckled disk. Figure 1.8 shows the exponential growth in inclination of all the bodies in the disk during the inclination instability. There is a corresponding drop in eccentricity

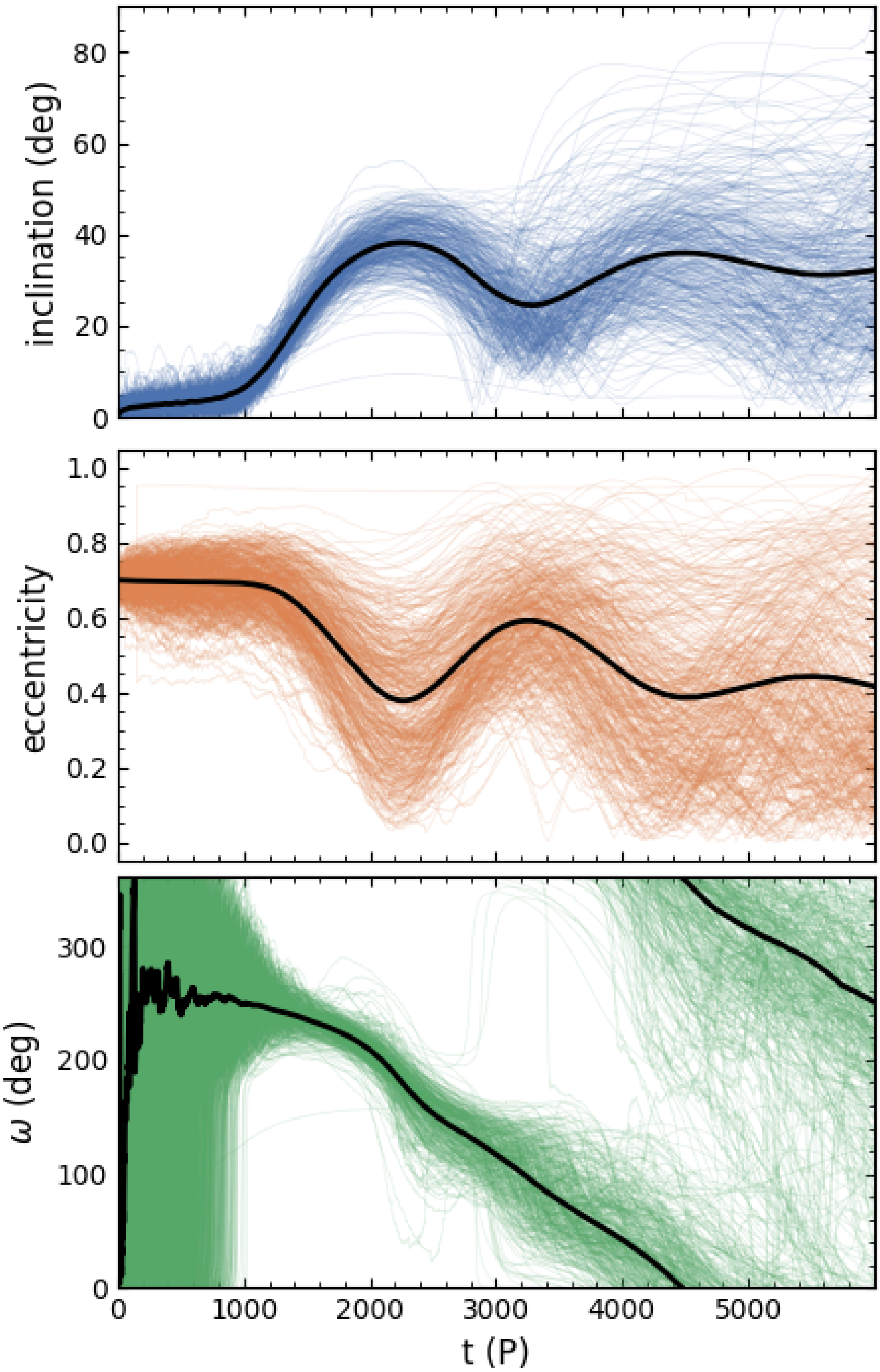

Figure 1.8: The inclination instability for the same disk as in the next figure shown with orbital elements. Mean values shown with a black line, and time in orbital periods shown on the x-axis. Inclination exponentially increases for all bodies and eccentricity decreases to conserve angular momentum. Significant ω-clustering occurs during the inclination instability, producing a bowl-shape in the orbits.

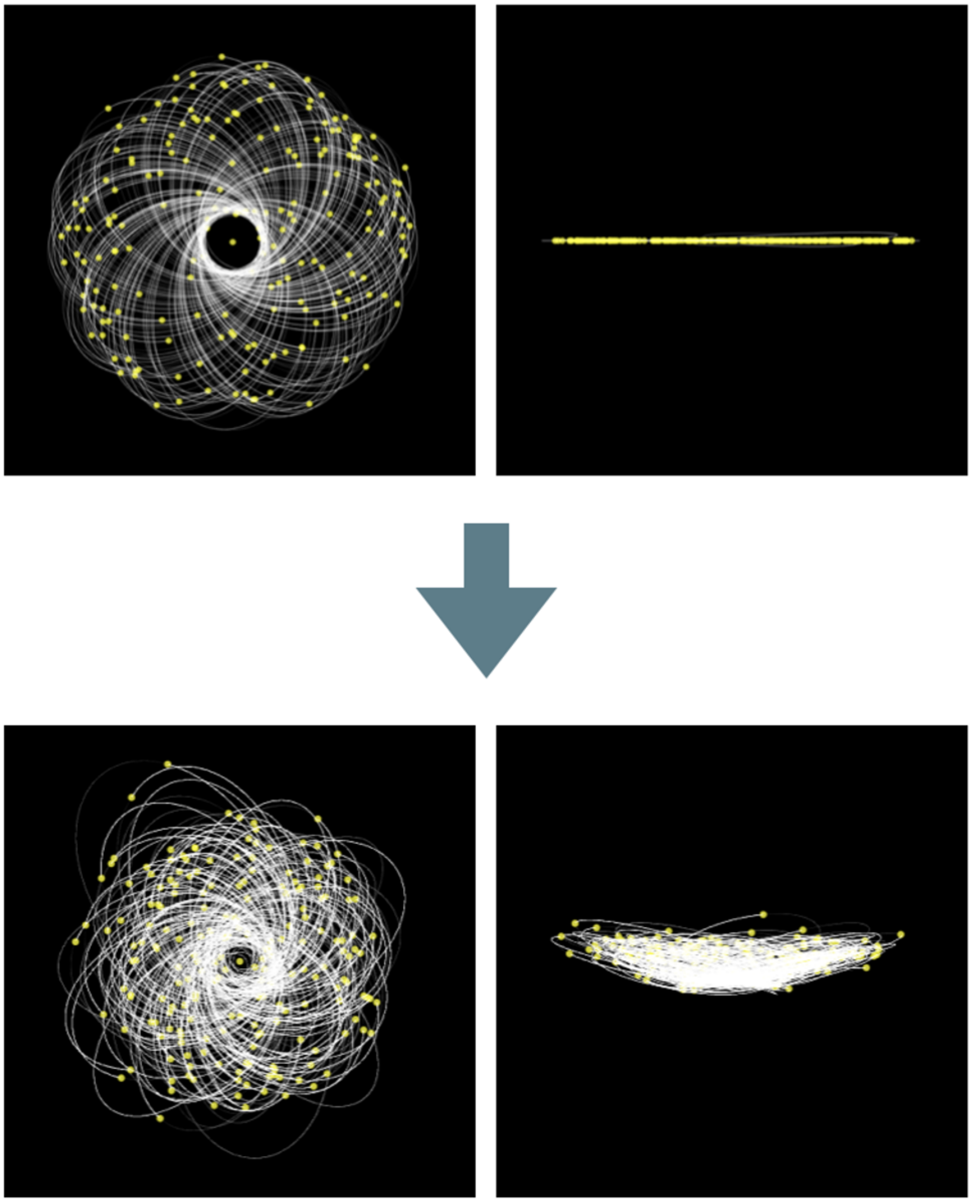

Figure 1.9: The inclination instability in physical space (taken from our N-body simulations). N-body particles are shown in dots, and their osculating orbits are shown in thin lines. Initial conditions shown on the left: an axisymmetric, nearly-flat disk of eccentric orbits ($e = 0.7$). When integrated, the disk buckles into a bowl shape with exponentially growing inclinations and decreasing eccentricities. The buckling direction is random; the disk is equally likely to buckle in the $+z$ or $-z$ direction.

to conserve vector angular momentum. An interesting feature of the instability is it's ability to increase the individual angular momentum of all the orbits in the disk while still conserving total angular momentum. This unique ability is facilitated by the asymmetric buckling of the disk (see Figure 1.9). Orbits increase their angular momentum by adding angular momentum parallel to the disk (in the x-y plane). Axisymmetry of the disk ensures that these x-y components cancel, maintaining net vector angular momentum. During the instability, all bodies within the disk roll and pitch in the same direction by the same amount. This amounts to a very tight clustering in the ω of the bodies in the disk. Recently, we've discovered that a weak apsidal clustering (ϖ-clustering) appears in the post-instability disk (Chapter 4)

The orbital distribution shown undergoing the inclination instability in Figures 1.8 and 1.9 does not reflect any real astrophysical disk. However, we have explored a variety of orbital configurations to determine some general constraints that disks have to satisfy to undergo the inclination instability:

(1) The disks have to be quite axisymmetric. Without axisymmetry, the buckling motion and the corresponding increase in individual orbital angular momentum cannot occur.

(2) Most of the orbits in the disk have to have eccentricities greater than ∼0.5. The linear stability of the disk is directly related to the eccentricity of the orbits (Chapter 2).

(3) The disk has to be relatively flat. We've found that disks with mean inclinations greater than about 10° don't undergo the inclination instability.

(4) *Typically*, the semi-major axis distribution of the orbits needs to be narrow. However, there is one important exception, the Scattered Disk.

This last point is interesting, and related to the angular momentum distribution of the disk. By a Scattered Disk structure, we mean a disk of orbits with roughly equal values of pericenter. The semi-major axis distribution can be broad (e.g. a^{-1} in $[100, 1000]\,\mathrm{AU}$), provided that eccentricity increases such that $q = a(1-e)$ is roughly constant. It isn't fully understood why this is the case. It probably has partly to do with the fact that higher eccentricity orbits are more unstable to this instability (Chapter 2). It may also have something to do with the angular momentum distribution of the Scattered Disk distribution. The angular momentum of a Keplerian orbit is given by $j = \sqrt{a(1-e^2)}$. If a and e are uncorrelated, $j \propto \sqrt{a}$. However, in the Scattered Disk, $e(a) = 1 - q/a$, and $j \propto \sqrt{2 - q/a}$. As a result, the Scattered Disk has an extremely flat angular momentum distribution. It seems the inclination instability is intricately related to the individual angular momentum of the orbits in the disk, and, perhaps, this flat angular momentum distribution specifically lends itself to instability. Regardless, it is fortunate for this theory that a Scattered Disk orbital configuration is unstable.

At face value, the inclination instability has the potential to explain all of the observed peculiarities mentioned in Section 1.3. It produces ω- and ϖ-clustering (see Chapter 4 for ϖ-clustering), it has the ability to reduce the mean e of orbits in the disk and would therefore be capable of detaching scattering orbits from Neptune, it's capable of producing very high inclination bodies (up to $i \sim 90°$), and if, the influences of the giant planets are included, it can produce a perihelion gap (Chapter 3). Because the inclination instability can produce a lopsided disk post-instability, it is a plausible formation mechanism for Sefilian & Touma (2019)'s disk theory. Unfortunately, we are faced with numerical limitations that make a direct, fully-interacting Neptune migration simulation impossible. Therefore, we have to approach the inclination instability's application to the Solar System indirectly.

The chapters in this thesis are composed of four published papers. Chapter 2 is Madigan et al. (2018b) wherein we describe the mechanism driving the inclination instability using a two-orbit torque model. In addition, we calculate the timescale for the inclination instability as a function of total disk mass (M_d) and number of particles (N), deriving an expression for the inclination instability in a realistic disk where $N \to \infty$. This result will prove critical for the application of the inclination instability to the Solar System. Chapter 3 is Zderic & Madigan (2020), wherein we study how the influence of the giant planets (to the quadrupole level) affect the inclination instability. Here we leverage the timescale derived in Chapter 2 to estimate the minimum Scattered

Disk mass required for the inclination instability to occur within it. Chapter 4 combines Zderic et al. (2020) and Zderic et al. (2021), wherein we find and explain the ϖ-clustering that appears post-instability within the disk. Finally, in Chapter 5, we summarize the results of these papers to paint a semi-complete picture of the viability of the inclination instability in the Solar System, and I outline relevant future work.

Chapter 2

On the Dynamics of the Inclination Instability

This article has been published the Astronomical Journal as: Madigan, A.-M., Zderic, A., McCourt, M., & Fleisig, J. 2018b, AJ, 156, 141, doi: 10.3847/1538-3881/aad95c

Abstract Axisymmetric disks of eccentric Kepler orbits are vulnerable to an instability that causes orbits to exponentially grow in inclination, decrease in eccentricity, and cluster in their angle of pericenter. Geometrically, the disk expands to a cone shape that is asymmetric about the mid-plane. In this paper, we describe how secular gravitational torques between individual orbits drive this "inclination instability". We derive growth timescales for a simple two-orbit model using a Gauss N-ring code, and generalize our result to larger N systems with N-body simulations. We find that two-body relaxation slows the growth of the instability at low N and that angular phase coverage of orbits in the disk is important at higher N. As $N \rightarrow \infty$, the e-folding timescale converges to that expected from secular theory.

2.1 Introduction

In Madigan & McCourt (2016), we introduced a new dynamical instability in axisymmetric Kepler disks of eccentric orbits. We showed that secular (orbit-averaged) gravitational torques between orbits in the disk drive exponential growth of their inclinations. As the orbits' inclinations grow, they tilt in the same way with respect to the disk plane. This leads to clustering in their angles of pericenter and the initially thin disk expands into a cone shape. Concurrently, the orbital eccentricities decrease and perihelion distances increase. We proposed that this could be at work in the outer solar system between minor planets scattered to large orbital eccentricities by the giant planets. It can explain the observed clustering in the angle of pericenter and detached, high perihelia objects (Trujillo & Sheppard, 2014).

This paper serves to explain the physics behind the inclination instability. In Section 2.2 we describe the mechanism of the instability through a simple, yet predictive, two-orbit toy model. We show analytically that inclinations grow exponentially in disks of eccentric orbits, and numerically derive the associated growth timescale. We discuss our results and their application to the outer solar system in Section 2.6 .

2.2 Mechanism of Instability

We consider a thin, axisymmetric disk of particles on eccentric orbits ($e \gtrsim 0.6$) around a much more massive central object. All orbits take part in the inclination instability which changes their orientation. However, to illustrate the instability's physical mechanism, we consider just two orbits within the disk and see how this system responds to a small perturbing force. The response of the system to the small perturbing force is best described with a set of angles different than the traditionally used Kepler elements (inclination i, longitude of ascending node Ω, and argument of pericenter ω). These new angles are i_{a}, i_{b}, and i_{e}, and we originally introduced them in Madigan & McCourt (2016). The angles represent rotations of the orbit about its semi-major ($\hat{\boldsymbol{a}}$) axis, semi-minor ($\hat{\boldsymbol{b}} \equiv \hat{\boldsymbol{j}} \times \hat{\boldsymbol{a}}$) axis, and angular momentum vector ($\hat{\boldsymbol{j}}$), respectively,[1] , i.e.,

$$i_{\mathrm{a}} = \arctan\left[\frac{\hat{b}_{\mathrm{z}}}{\sqrt{1-\hat{b}_{\mathrm{z}}^2}}\right], \tag{2.1a}$$

$$i_{\mathrm{b}} = \arctan\left[-\frac{\hat{a}_{\mathrm{z}}}{\sqrt{1-\hat{a}_{\mathrm{z}}^2}}\right], \tag{2.1b}$$

$$i_{\mathrm{e}} = \arctan\left[\hat{a}_{\mathrm{y}}, \hat{a}_{\mathrm{x}}\right]. \tag{2.1c}$$

---

[1] Note that these are not true rotation angles. See Appendix A for more information.

The subscripts x, y, and z denote an inertial Cartesian reference frame with unit vectors, $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$. The angles i_{a}, i_{b}, and i_{e} are equivalent to that of an aircraft's roll, pitch and yaw. We give a graphical representation of i_{a} and i_{b} in figure 2.1 and convert between these variables and Kepler elements in Appendix A.

2.3 Two-orbit toy model

We consider how two initially coplanar, anti-aligned eccentric orbits respond to a perturbing force, as depicted in figure 2.1.

(1) We perturb orbit 1 with a uniform force in the $+\hat{\boldsymbol{z}}$-direction. Such a force could arise from, e. g., a temporary fluctuation in the center of mass of the disk. Because of the orbit's symmetry around the $\hat{\boldsymbol{a}}$ axis, all torques in the $\hat{\boldsymbol{a}}$-direction cancel. However, because of the orbit's asymmetry around the $\hat{\boldsymbol{b}}$ axis, the vertical force perturbation creates a net torque in the $\hat{\boldsymbol{b}}$-direction. This asymmetry between the torques is further exacerbated by the time-averaged torque, $\overline{\boldsymbol{\tau}} = \overline{\boldsymbol{r} \times \boldsymbol{f}}$, at apocenter being greater than at pericenter because the particle spends a greater fraction of its time at apocenter, ($\sim [(1+e)/(1-e)]^{3/2}$). Thus, the net time-averaged torque points in the $+\hat{\boldsymbol{b}}$-direction. Consequently, the angular momentum vector $\boldsymbol{j}$ of orbit 1 rotates counter clockwise about its semi-major axis, rolling the orbital plane such that $i_{\mathrm{a}} < 0$.

(2) The roll of orbit 1 induces out-of-plane forces along the sides of orbit 2. These forces are asymmetric around the $\hat{\boldsymbol{a}}$ axis and imply a time-averaged torque on orbit 2 in its $+\hat{\boldsymbol{a}}$-direction. This torque rotates orbit 2 about its semi-minor axis, pitching the orbital plane such that $i_{\mathrm{b}} > 0$.

(3) This rotation raises orbit 2's apocenter, and thus its center of mass, above the original orbital plane. The raising of orbit 2's apocenter raises the center of mass of the disk, which

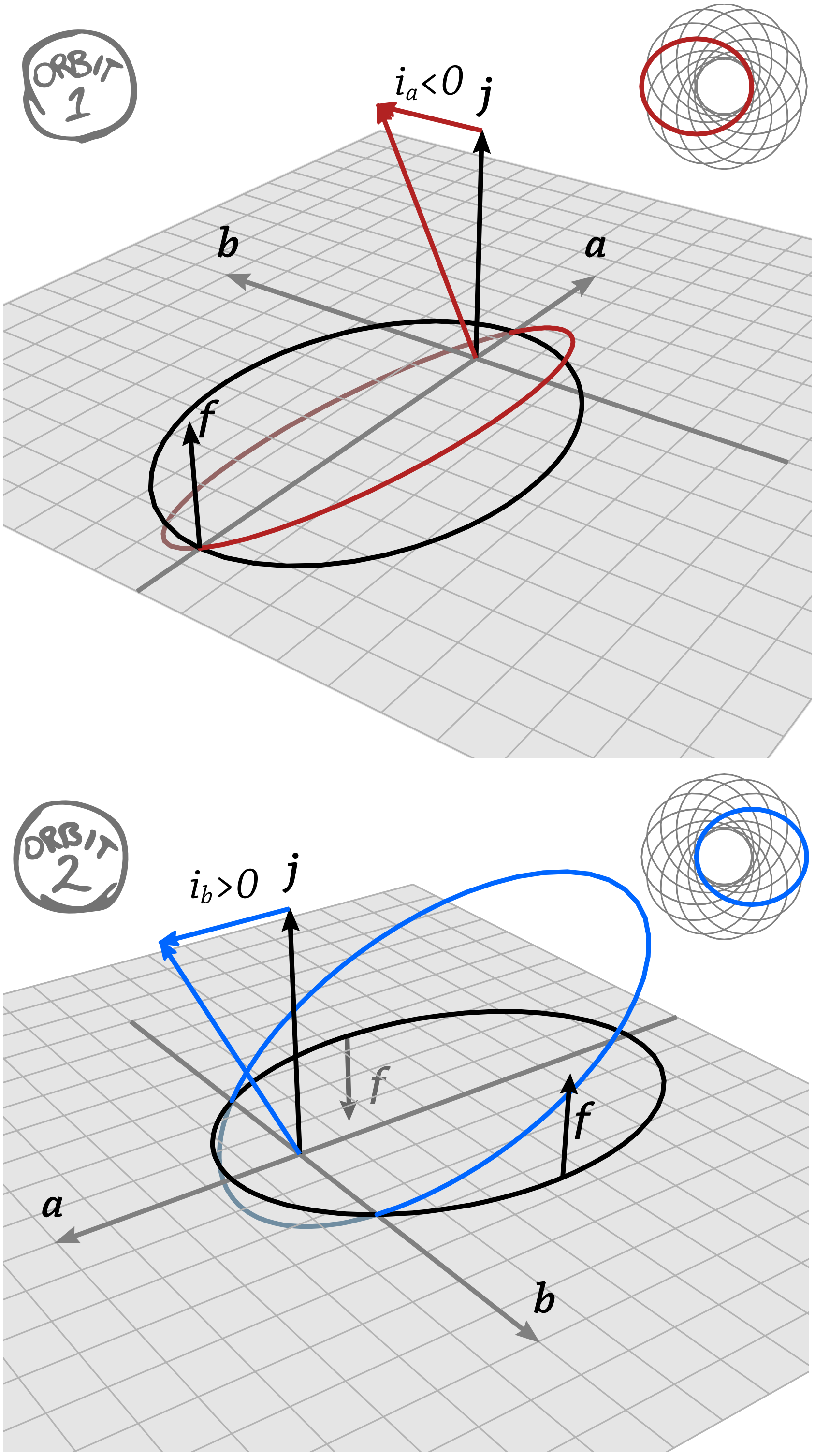

Figure 2.1: Idealized, “two-orbit” toy model for the inclination instability. The top right of each panel shows the location of the orbit in the disk from a face-on perspective. (*top*): Orbit 1 experiences a net upward force This force produces a torque along the axis, rotating the orbital plane such that (*bottom*): Diametrically opposed orbit 2 feels a force due to the rotation of orbit 1; this force produces a torque along the direction of orbit 2, rotating its orbital plane such that

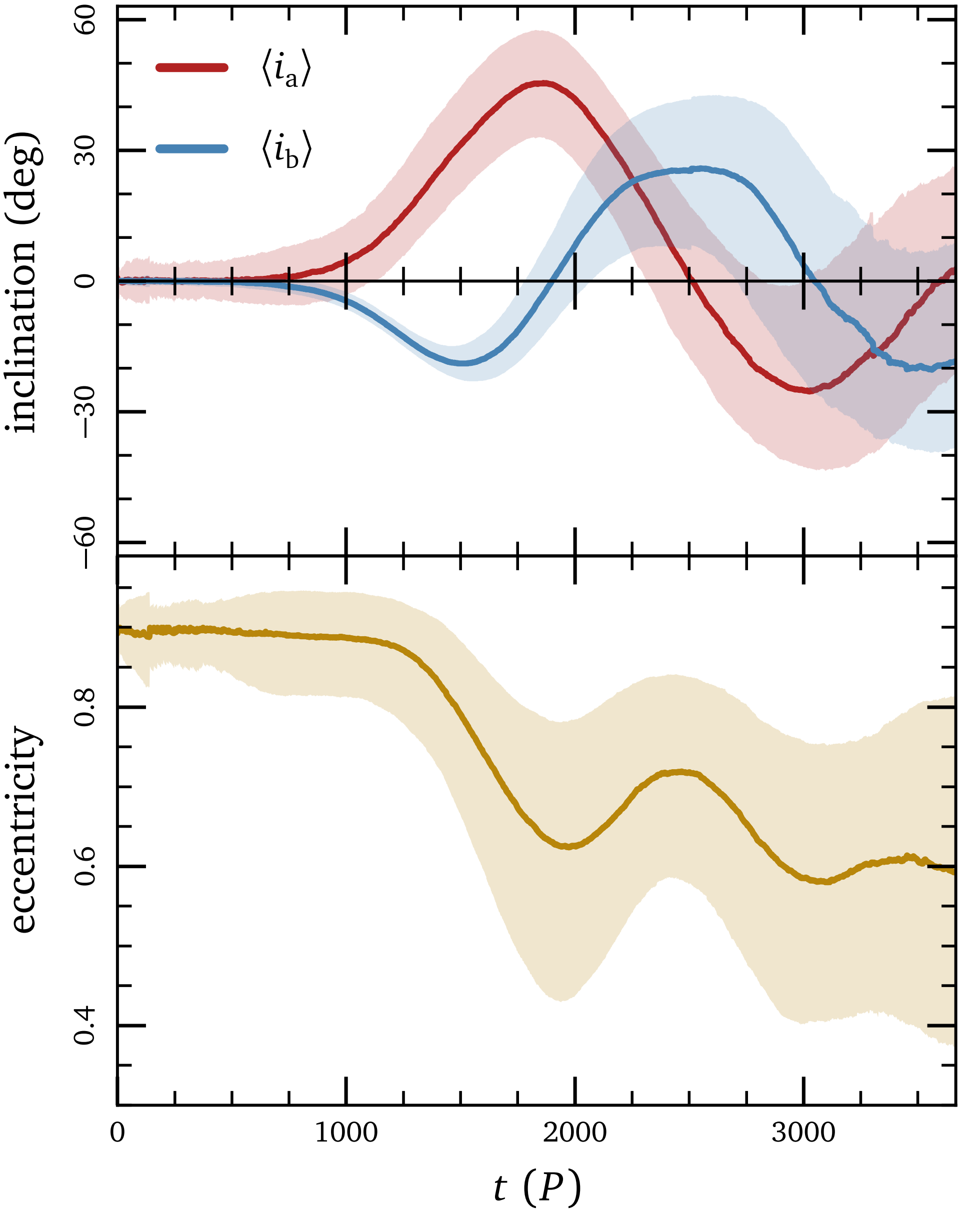

Figure 2.2: (*top*): Median values of angles i_a and i_b for all particles in a disk as a function of time in units of orbital period. Shaded regions indicate one-sigma quantile contours. Both i_a and i_b rise exponentially at early times. As predicted by the two-orbit toy model, i_a and i_b increase with opposite sign. (*bottom*): Median value of orbital eccentricity for the same particles. As inclinations increase, eccentricities decrease.

reinforces the initial perturbation. Thus, the disk is unstable to the initial perturbation, and the inclination will run away to large values.

As the orbits incline off the disk plane, conservation of angular momentum requires that they decrease in eccentricity. They do so via the same secular gravitational torques that drive growth in inclination. Let us take orbit 1 as an example. After orbit 1 begins to roll over its major axis, its angular momentum vector is no longer parallel to the initial perturbing force. A component of the force acts therefore in its orbital plane. This component causes a positive time-averaged torque resulting in a decrease in eccentricity. A similar analysis shows that orbit 2 must also decrease in eccentricity.

Though we have only considered two orbits in the disk, all orbits take part in the instability. We show this in figure 2.2 in which we plot the median values of i_{a} and i_{b} (top), and eccentricity (bottom) for a representative N-body simulation. We refer the reader to figure 2.4 to clearly see exponential growth of i_{b}. Here we use the `REBOUND` code (Rein & Liu, 2012), with the IAS15 integrator (Rein & Spiegel, 2015). We distribute $N = 100$ massive particles in a thin axisymmetric disk in the potential of a central body of mass M. The orbit eccentricities are initialized to $e = 0.9$, semi-major axes are uniformly distributed in the narrow range $[0.9, 1.1]$. Mean anomaly, argument of pericenter, and angle of ascending node are distributed uniformly in $(0, 2\pi]$.

The two-orbit toy model makes several testable predictions confirmed by N-body experiments (see also figure 2 of Madigan & McCourt (2016)):

(1) If the magnitudes of the torques are linear in i_{a} and i_{b}, as one might expect for small deviations from an equilibrium, this instability should grow exponentially in time.

(2) The orbits should grow their $|i_{\mathrm{a}}|$ and $|i_{\mathrm{b}}|$ values with opposite signs and $i_{\mathrm{b}}/i_{\mathrm{a}}$ in constant ratio.

(3) A constant ratio $i_{\mathrm{b}}/i_{\mathrm{a}}$ implies a constant angle of pericenter, as for small inclinations $\omega(i_{\mathrm{a}}, i_{\mathrm{b}}) \sim \arctan|i_{\mathrm{b}}/i_{\mathrm{a}}|$ ($+\pi$ if $i_{\mathrm{a}} < 0$). As i_{a} and i_{b} depend on the eccentricities of

orbits, $\omega(i_{\rm a}, i_{\rm b})$ should cluster to one of two eccentricity-dependent values, depending on the direction of the initial perturbing force.

2.4 Growth timescale of instability

We estimate the e-folding timescale of the instability between two orbits as follows. For a given orbit with its $\hat{\boldsymbol{a}}$ and $\hat{\boldsymbol{b}}$ axes, two components of the torque that rotate its orbital plane are:

$$\tau_{\rm a} \equiv \boldsymbol{\tau} \cdot \hat{\boldsymbol{a}}, \tag{2.2a}$$

$$\tau_{\rm b} \equiv \boldsymbol{\tau} \cdot \hat{\boldsymbol{b}}. \tag{2.2b}$$

We denote the torque on orbit 2 due to orbit 1 as $\tau^{(2,1)}$ and the torque on orbit 1 due to orbit 2 as $\tau^{(1,2)}$. Torque acting on orbit 2 can be expressed as

$$\tau_{\rm a}^{(2,1)}(i_{\rm a}^{(1)}) = \frac{dj^{(2)}(i_{\rm b}^{(2)})}{dt} = j^{(2)}\frac{di_{\rm b}^{(2)}}{dt}, \tag{2.3}$$

(where the last step is simply vector rotation), and expanded via Taylor series and linearized

$$\tau_{\rm a}^{(2,1)}(i_{\rm a}^{(1)}) = \left(\frac{d\tau_{\rm a}^{(2,1)}}{di_{\rm a}^{(1)}}\right) i_{\rm a}^{(1)}. \tag{2.4}$$

Since $i_{\rm a}$ represents a rotation of angular momentum in the $(-\hat{\boldsymbol{b}})$ direction and $i_{\rm b}$ represents a rotation of angular momentum in the $\hat{\boldsymbol{a}}$ direction, we have the linearized equations of motion:

$$j^{(2)}\frac{di_{\rm b}^{(2)}}{dt} = \tau_{\rm a}^{(2,1)}(i_{\rm a}^{(1)}) = \left(\frac{d\tau_{\rm a}^{(2,1)}}{di_{\rm a}^{(1)}}\right) i_{\rm a}^{(1)} = \alpha\, j^{(2)}\, i_{\rm a}^{(1)}, \quad \text{and} \tag{2.5a}$$

$$j^{(1)}\frac{di_{\rm a}^{(1)}}{dt} = -\tau_{\rm b}^{(1,2)}(i_{\rm b}^{(2)}) = -\left(\frac{d\tau_{\rm b}^{(1,2)}}{di_{\rm b}^{(2)}}\right) i_{\rm b}^{(2)} = -\beta\, j^{(1)}\, i_{\rm b}^{(2)}, \tag{2.5b}$$

where in the last step we have defined the coefficients

$$\alpha \equiv \frac{1}{j^{(2)}}\left(\frac{d\tau_{\rm a}^{(2,1)}}{di_{\rm a}^{(1)}}\right) = \frac{1}{j^{(1)}}\left(\frac{d\tau_{\rm a}^{(1,2)}}{di_{\rm a}^{(2)}}\right), \tag{2.6a}$$

$$\beta \equiv \frac{1}{j^{(1)}}\left(\frac{d\tau_{\rm b}^{(1,2)}}{di_{\rm b}^{(2)}}\right) = \frac{1}{j^{(2)}}\left(\frac{d\tau_{\rm b}^{(2,1)}}{di_{\rm b}^{(1)}}\right). \tag{2.6b}$$

There are similar expressions for $i_{\rm a}^{(2)\prime}(t)$ and $i_{\rm b}^{(1)\prime}(t)$. In the following, we assume the parameters α and β are constants, which is consistent with N-body simulations of the instability. We then differentiate equation 2.5a and combine it with equation 2.5b to yield an equation of motion for i_{b}:

$$i_{\mathrm{b}}^{(2)}{}''(t) = \alpha\, i_{\mathrm{a}}^{(1)}{}'(t) = -\alpha\beta\, i_{\mathrm{b}}^{(2)}. \tag{2.7}$$

For small inclinations, the inclination angles evolve in time as

$$\{i_{\mathrm{a}}, i_{\mathrm{b}}\} \propto \exp(\gamma t), \tag{2.8}$$

$$\gamma \equiv \sqrt{-\alpha\beta}, \tag{2.9}$$

where γ is the growth rate of the instability. Equation 2.7 shows that the instability exponentially grows orbit inclinations through mutual secular gravitational torques between eccentric orbits.

We numerically evaluate α and β from equation 2.6 using Gauss' approximation, as described in Gürkan & Hopman (2007).[2] We estimate the time-averaged torques between two orbits by smearing the bodies out into "wires" with a linear mass density along the orbit inversely proportional to the instantaneous Kepler velocity; we then split each wire into n segments equally spaced in eccentric anomaly, and we directly sum the n^2 gravitational torques of all pairs of segments. We repeatedly double the number of segments until the total torque converges. Though our numeric estimate for the time-averaged torques between orbits does not require gravitational softening, we include it in our calculation as arbitrarily large torques can arise for certain orientations of the pair of orbits. This is unsatisfactory from a theoretical perspective because the large torques arising over small spatial scales destroy the linearity presumed in equation 2.5. These large torques are also unphysical, because they would drive scattering or diffusion of the orbits on a timescale shorter than the instability growth timescale. They therefore should not be included in our calculation, and gravitational softening is one way to approximate this behavior. We adopt a fiducial force-softening length $s = (3 \times 10^{-2})\, a$, chosen to exclude close encounters that would drive diffusion on a timescale shorter than $1/\gamma$. With this choice, the linear theory we present applies over the majority of the parameter space. The boundaries between stable and unstable regions do not depend significantly on our chosen value of s. However, the maximum growth rate scales with the softening length as $\gamma_{\max} \sim s^{-3/4}$.

---

[2] Our source code is publicly available at `https://github.com/mkmcc/orbit-torque`.

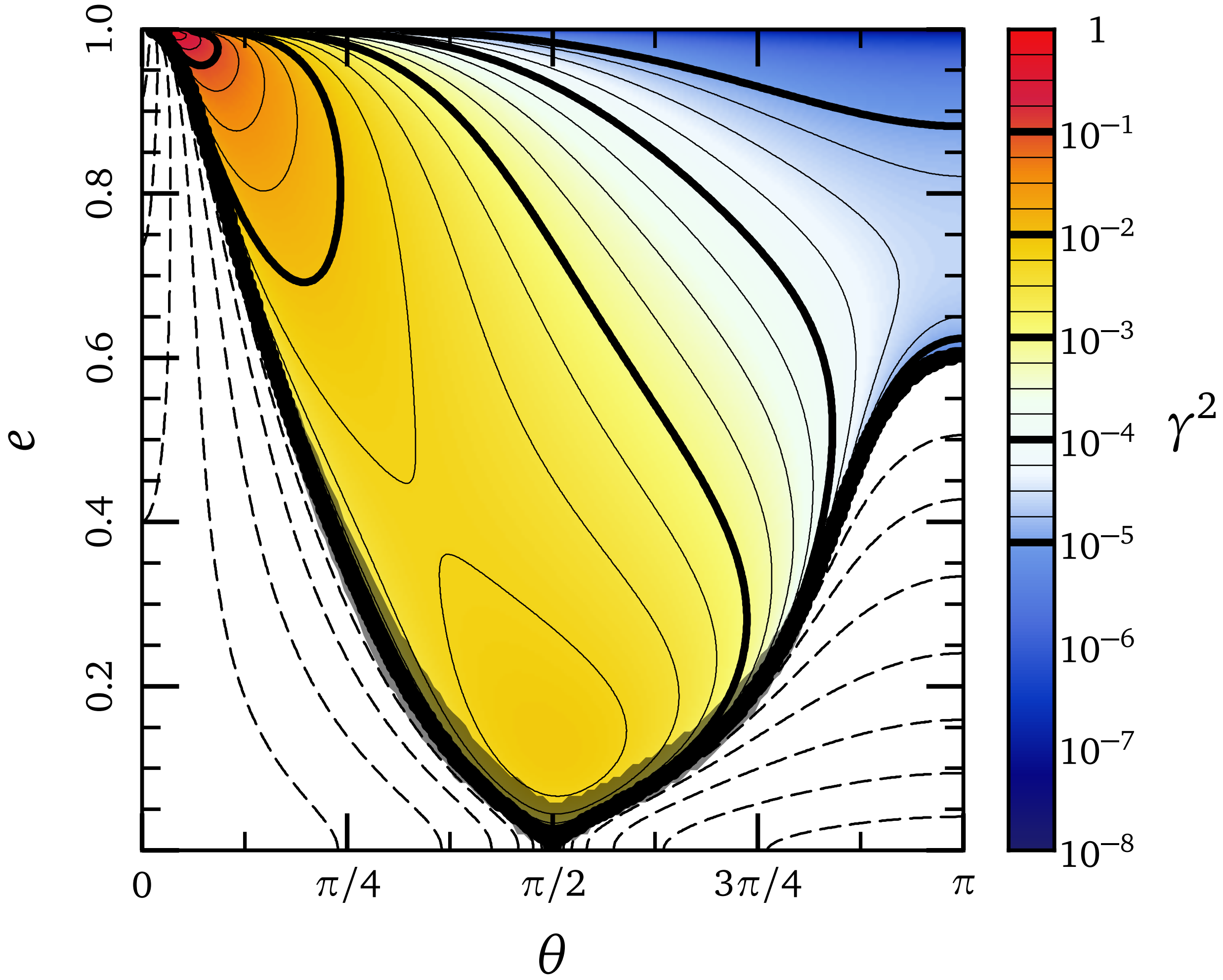

Figure 2.3: Squared growth rate γ^2 for two orbits gravitationally torquing each other. This is estimated from the linearized equations of motion for the system (equation 2.9), as a function of orbital eccentricity and angular separation between orbits in the plane, θ. Only the unstable region with $\gamma^2 > 0$ is colored. The white, uncolored region of the plot with dashed contours is stable. The thick, gray band near the $\gamma^2 = 0$ boundary marks the region in which the linear approximation in equation 2.5 breaks down, causing an order-unity uncertainty in our estimate of the timescale. This region also depends on the softening length we adopt.

Using this numerical approach, we calculate the squared growth rate, γ^2, as a function of orbital eccentricity and angular separation, $\theta = \arccos(\hat{\boldsymbol{e}}_1 \cdot \hat{\boldsymbol{e}}_2)$. We initialize the equal-mass orbits with similar semi-major axes. The growth rate is normalized to the orbital period, so that γ is the number of e-foldings per orbit. We show γ^2 as a function of θ and e in figure 2.3. At every eccentricity, the orbits are unstable in some range of angular separations. The largest growth rates occur at the highest orbital eccentricities and small-to-intermediate angular separations. In general, as eccentricity increases, the range of stable angular configurations shrinks. At very small angular separations, the two orbits are unstable only for very high eccentricities. Interestingly, our results show that the angular separation ($\theta = \pi$) used by the toy model in section 2.3 is the least-unstable unstable region of figure 2.3. The inclination instability only occurs in high eccentricity disks ($e \gtrsim 0.6$) because below this boundary most orbits are stable with respect to one another; tests with N-body simulations confirm this.

There is a marked boundary in $e - \theta$ space between stable and unstable configurations. The location of this boundary can be understood from the two-orbit model: For low θ, orbits exert stabilizing vertical forces on each other because their sides overlap. As θ increases, the distance between their sides grows until the force acting on each side reverses and becomes unstable. This reversal happens at smaller separations for high eccentricity orbits as the width of an orbit scales as $\sqrt{1-e^2}$. Orbits of nonzero eccentricity are unstable at $\theta = \pi/2$. As θ increases towards π, the growth rate of the instability decreases due to the increasing distance and decreasing gravitational force between the orbits. Low-eccentricity orbits are the most stable as forces at pericenter become stronger than those at apocenter. The squared growth rate γ^2 is negative in the stable regions which implies that the two orbits may be undergoing oscillatory behavior. We will explore the dynamical behavior of the orbits in the regions of stability in a future study.

2.5 Generalization to N bodies

To apply the instability to real astrophysical systems with more than two bodies, we need to know the growth rate for large N systems. We start with the hypothesis that the inclination instability acts on a secular timescale,

$$t_{\text{sec}} \sim \frac{1}{2\pi}\frac{M}{M_{\text{d}}}P, \tag{2.10}$$

where P is the orbital period. Here t_{sec} is derived as the time it takes to change an orbit's angular momentum by order of its circular angular momentum using a specific torque over one orbital period of $\tau \sim Gm/a$. To test this hypothesis, we use numerical N-body simulations with the `REBOUND` code. We simulate the instability for a range of disk masses ($M_{\text{d}} = [10^{-2}, 10^{-3}, 10^{-4}]$, where the central mass $M = 1$), and number of particles ($N =$ [100, 200, 500, 1000]), averaging many tens of simulations for each parameter set to reduce the noise. Each particle has $m = M_{\text{d}}/N$, $e = 0.7$, and $i = 10^{-4}$ rad. Semi-major axes are distributed uniformly in $[0.9, 1.1]$. Mean anomaly, argument of pericenter, and angle of ascending node are randomly distributed in $(0, 2\pi]$. We do not use gravitational softening.

We show the results of our simulations in figure 2.4. The top plot shows the mean inclination of particles as a function of time for simulations with a fixed total mass $M_{\text{d}} = 10^{-2}M$. The various lines correspond to simulations with differing numbers of particles. A surprising result is how dependent the inclination growth rate is on the number of particles per simulation. The instability proceeds faster with increasing N. The bottom plot extends this result for different disk masses, showing the growth rate γ as a function of N for $M_{\text{d}} = [10^{-2}, 10^{-3}, 10^{-4}]$. We derive the growth rate in each simulation from the linear parameter of a quadratic model fit to the median $\log i_{\text{b}}$ value[3] . We optimize the fit parameters with a linear least-squares method and find the error using a bootstrapping method. Our data is Gaussian distributed so a least-squares optimization is equivalent to a maximum likelihood estimator (MLE) approach.

If the e-folding time of the instability matched the secular timescale given by equation 2.10, the data in figure 2.4 would align with the dashed gray lines. This is clearly not the case however, particularly at low N and large M_{d}. We identify three reasons for the deviation of the instability growth rate from a secular one.

---

[3] We use i_{b} instead of i_{a} to calculate the exponential growth rate as it is the less noisy parameter (it requires more force for an orbit to change its i_{b} value as it must lift its apocenter).

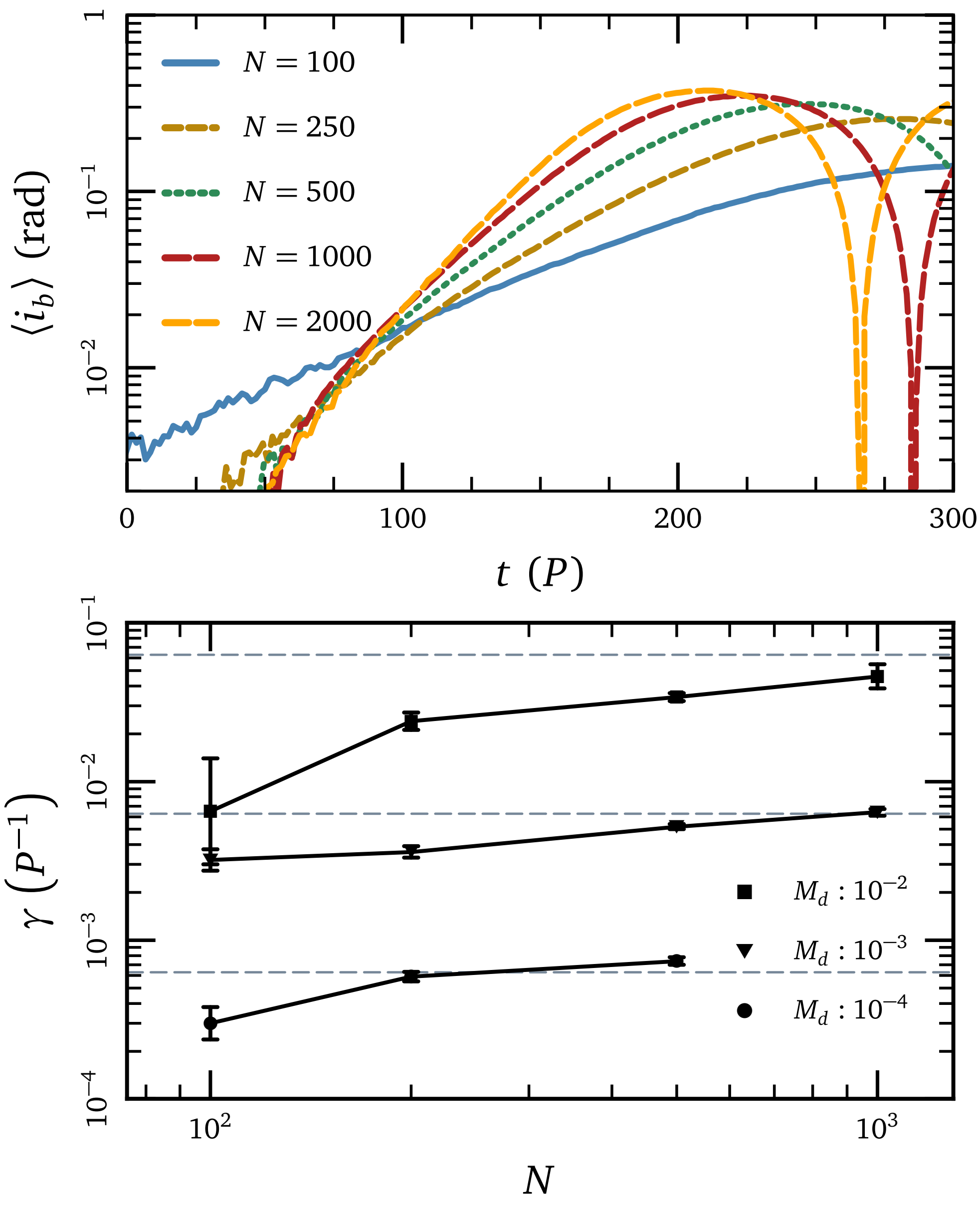

Figure 2.4: (*top*): Exponential growth of the inclination angle i_{b} for different numbers of disk particles (N) and fixed total mass $M_{\mathrm{d}} = 10^{-2}M$. Time is in units of orbital period at semimajor axis $a = 1$. Note that the slopes (growth rate) of the lines show a strong dependence on N. (*bottom*): Instability growth rate (γ) for different values of N and M_{d}. The expected secular growth rate (t_{sec}^{-1}) is plotted as dashed gray lines (corresponding from top to bottom to disk masses of $M_{\mathrm{d}} = [10^{-2}, 10^{-3}, 10^{-4}]$). Note that the growth rate's dependence on N weakens as M_{d} is reduced, and ultimately converges to the secular growth rate.

(1) Angular phase coverage of orbits in disk

In the two-orbit model, the instability growth rate strongly depends on the angular separation of orbits (see figure 2.3). In fact, large sections of phase space are stable ($\gamma^2 < 0$). Generalizing to N bodies, the angular phase coverage of orbits in a disk greatly impacts the overall dynamics, including whether or not the disk is stable or unstable. Non-axisymmetric disks are stable; a good example being the eccentric nuclear disk of stars in the Andromeda nucleus (Madigan et al., 2018a). Furthermore, in the limit of small N, few angular separations are realized and axisymmetric disks can also be stable; tests show that at $N \lesssim 20$ the instability doesn't occur within the predicted timescale. As we randomly select the angular orientation of the orbits in our simulations, the unstable region of phase space gets filled at a rate $\propto N^{1/2}$. Assuming the most unstable angular separation dominates the growth rate, we would then expect the growth rate to follow an $N^{1/2}$ dependence.

(2) Apsidal precession

In limit of $M_{\mathrm{d}} \rightarrow M$, the Kepler orbit approximation breaks down. Orbits precess rapidly due to the gravitational potential of the disk, and we expect orbit-averaged torques to weaken significantly. Hence the timescale for the instability should strongly deviate from the secular timescale at large M_{d}/M values, as seen in figure 2.4.

(3) Two-body relaxation

In a near-Keplerian system, the two-body relaxation time is given by

$$t_{\mathrm{rel}} \equiv \frac{M^2}{m^2 N \ln(\Lambda)} \tag{2.11}$$

(Rauch & Tremaine, 1996), where $\ln(\Lambda)$ is the Coulomb logarithm. Therefore, the ratio of the two-body timescale to the secular timescale is

$$\frac{t_{\mathrm{rel}}}{t_{\mathrm{sec}}} \sim \frac{M}{m \ln(\Lambda)} \propto m^{-1}. \tag{2.12}$$

As m decreases, the two-body relaxation timescale increases relative to the secular timescale. Conversely, for larger values of m, strong two-body scattering perturbs particles off their orbits and slows the rate of the secular instability. Thus, we expect the growth rate to have a m^{-1} dependence from two-body relaxation. For a fixed M_d, this will emerge as an N^{-1} dependence.

All three of these physical effects coexist simultaneously in our simulations. To explore their relative importance, we parameterize deviation from the secular growth rate with an arbitrary function α,

$$\gamma^{-1} = \alpha(N, M_\mathrm{d})\ t_\mathrm{sec}. \tag{2.13}$$

We fit our simulation data using two different models, $\alpha = aN^{-1/2} + b$ and $\alpha = aN^{-1} + b$, for disk phase-coverage and two-body relaxation, respectively. The phase coverage model most accurately fits the large N evolution of the instability timescale ($N = 200$–1000). Additionally, it predicts that α converges to the same value for all M_d ($b \approx 0.4$). However, the two-body scattering model more accurately fits the low N data points. Therefore, we see a mixture of phase coverage and two-body scattering for the N we simulated. In both models, the value of the a parameter decreases with M_d. That is, the growth rate's dependence on N weakens as M_d decreases. Apsidal precession is the likely source of this additional M_d dependence. In summary, our exploration into the growth rate of the instability for large N, low M_d systems finds that the e-folding timescale is

$$t_\mathrm{e-fold} \sim \frac{0.2}{\pi}\frac{M}{M_\mathrm{d}}P. \tag{2.14}$$

2.6 Discussion

In this paper, we describe the mechanism behind the inclination instability in axisymmetric, eccentric Kepler disks as introduced by Madigan & McCourt (2016). The instability results in an exponential growth of orbital inclinations, a decrease in orbital eccentricities, and clustering in the argument of pericenter during its linear phase. We explain how the instability works with a simple toy model, examining the secular torques between just two orbits (section 2.3). Using linearized equations of motion, we analytically demonstrate the exponential growth of the instability and numerically derive the e-folding rate (section 2.4). We expand to large N systems (section 2.5) with a suite of N-body simulations. Our results show that two-body scattering, angular phase coverage of orbits within the disk, and apsidal precession reduce the growth rate of the instability. At low N ($\approx 10^2$), two-body scattering is the dominant physical process decreasing the growth rate. At large N, the angular phase coverage of orbits is important. Apsidal precession further decreases the growth rate as $M_\mathrm{d} \to M$. This result is general to all N-body simulations involving secular dynamical behavior. Thus, we conclude that N-body simulations (which for computational efficiency typically use low N and high M_d/M values) generally underestimate the strength of secular dynamics.

2.7 Outer Solar System

The results of section 2.5 can be used to apply the inclination instability to real astrophysical systems that are currently impossible to directly simulate. Studies of inner Oort cloud formation in a stellar cluster suggest that 1-10 Earth masses of cometary material may exist at hundreds of au (Brasser et al., 2012). The comets are scattered outward by Jupiter and Saturn, and their pericenters are later decoupled from the planets by perturbations from cluster gas and nearby stars. We compare important dynamical timescales for bodies in this region, taking the orbital parameters of the minor planet Sedna as an example (Brown et al., 2004b, $e \approx 0.85$, $a \approx 507$ au, $i \approx 12°$).

The induced motion of Sedna's orbit due to the four giant planets can be approximately calculated by assigning to the Sun an artificial quadrupole moment,

$$J_2 = \frac{1}{2} \sum_i^4 \frac{m_i}{M_\odot} \left(\frac{a_i}{R_\odot} \right)^2 . \tag{2.15}$$

The apsidal and nodal precession frequencies are then given by

$$\dot{\omega} = \frac{3 J_2}{4} n \left(\frac{R_\odot}{a} \right)^2 \left(\frac{5 \cos^2 i - 1}{(1-e^2)^2} \right) , \tag{2.16}$$

and

$$\dot{\Omega} = -\frac{3 J_2}{2} n \left(\frac{R_\odot}{a} \right)^2 \left(\frac{\cos i}{(1-e^2)^2} \right) , \tag{2.17}$$

where $n = 2\pi/P$ is the mean motion of Sedna's orbit (Brasser et al., 2006). With $J_2 = 2.6149 \times 10^3$, the corresponding precession timescales are $t_{\mathrm{apsidal}} = 2\pi/\dot{\omega} \approx 1.4\,\mathrm{Gyr}$, $t_{\mathrm{nodal}} = 2\pi/\dot{\Omega} \approx 2.6\,\mathrm{Gyr}$.

The Kozai-Lidov timescale for Sedna due to the galactic potential is,

$$t_{\mathrm{KL}} \sim 4.0 \left(\frac{0.1 M_{\odot}\mathrm{pc}^{-3}}{\rho_G}\right) \left(\frac{10^4\ \mathrm{au}}{a}\right)^{3/2}\ \mathrm{Gyr} = 360\ \mathrm{Gyr}, \tag{2.18}$$

where ρ_G is the mass density of the Galaxy in the solar neighborhood, and the angular momentum diffusion timescale for Sedna from stellar passages is,

$$t_{\mathrm{diff}} \sim 25 \left(\frac{q}{30\ \mathrm{au}}\right) \left(\frac{10^4\ \mathrm{au}}{a}\right)^{2}\ \mathrm{Myr} = 25\ \mathrm{Gyr}, \tag{2.19}$$

with the timescales taken from Brasser et al. (2008)[4]. The timescale for apsidal precession induced by the giant planets is the shortest of these three. The inclination e-folding timescale for Sedna is,

$$t_{\mathrm{inst}} \sim 0.4\, \frac{1}{2\pi} \left(\frac{M_{\odot}}{M_{\mathrm{d}}}\right) P = 725 \left(\frac{M_{\odot}}{M_{\mathrm{d}}}\right)\ \mathrm{yr}. \tag{2.20}$$

If $M_{\mathrm{d}} \approx 0.18 M_{\oplus}$, the instability timescale is of the same order of magnitude as the apsidal precession timescale. Thus, we find that in this scenario the self-gravity of small bodies in the Sedna region is the dominant dynamical driver.[5]

If we instead hypothesize that the minor planets in the Sedna region came from the scattered disk (a natural starting point for the instability, which requires only eccentric orbits in an axisymmetric disk), and detached themselves via self-gravitational torques, the more relevant comparison is that of the instability e-folding timescale to the energy diffusion time. The energy diffusion time due to planetary perturbations is given by Duncan et al. (1987),

$$t_{D(x)} \approx 1 \times 10^6 \left(\frac{10^4 \mathrm{au}}{a}\right)^{1/2} \left(\frac{10^{-4} \mathrm{au}^{-1}}{D(x)}\right)^{2}\ \mathrm{yr}, \tag{2.21}$$

where $D(x)$ is the typical energy perturbation per perihelion passage. For a body originating in the scattered disk with perihelion distance of $p \approx 35\mathrm{au}$ and a low orbital inclination, $D(x) \approx 10^{-5}\mathrm{au}^{-1}$ (figure 1 of Duncan et al., 1987). This yields an energy diffusion time at Sedna's semi-major axis

---

[4] The angular momentum diffusion timescale and Kozai-Lidov timescale were derived for $e \sim 1$ orbits in Brasser et al. (2008).

[5] Of course this statement does not hold if a massive planet 9 (Batygin & Brown, 2016a) exists in this region.

of $t_{D(x)} \approx 450$ Myr, consistent with results from Bannister et al. (2017). If $M_{\rm d} \approx 0.5 M_\oplus$, the instability timescale equals the energy diffusion timescale. Thus in this scattered disk scenario, the self-gravity of small bodies in the Sedna region is the dominant dynamical driver if there exists about half an Earth mass of material at hundreds of au. In general, self-gravity in the inner Oort Cloud acts to increase the clearing timescale due to planetary perturbations. Gravitational torques increase the orbital inclination and the perihelion distance of minor planets, thus weakening the strength of scattering events. Minor planets can be more effectively retained.

Fan & Batygin (2017) recently presented simulations of self-gravitating planetesimal disks within the context of the Nice Model. They found no evidence of the inclination instability, attributing its absence to mass depletion via ejections, and rapid precession of orbits due to the presence of the giant planets. We note however that the simulations ran for only 20 Myr during the most violent phase of solar system evolution. We do not expect the inclination instability to dominate while the giant planets are migrating tens of au. Furthermore, there is the important issue of particle number. If the original disk mass of $30 M_\oplus$ depletes to $\approx 0.5 M_\oplus$, the number of particles reduces from $N = 10^3$ to $N \approx 16$. The instability cannot be captured at this low particle number, as discussed in section 2.5. Thus, we cannot conclude from the Fan & Batygin (2017) results that the Nice model is incompatible with the inclination instability.[6]

If the observed clustering of arguments of pericenter (ω) of minor planets is a real effect (see Shankman et al., 2017), we need to either (i) be catching the instability during its linear phase, or (ii) be in the nonlinear, saturated phase and have something maintaining the clustering. It is interesting to note that the cone-shaped gravitational potential produced by the instability induces precession of eccentric orbits such that $\dot{\omega} < 0$, while the presence of the giant planets induces the opposite, $\dot{\omega} > 0$ (for $\cos i < 1/\sqrt{5}$; see Equation 2.16). There may well be an interesting region of phase space where the two cancel and clustering is maintained (see Sefilian & Touma (2019) for a similar idea involving longitude of perihelion (ϖ) clustering with a massive, apse-aligned disk).

---

[6] It's interesting that the success ratio of the Nice model in reproducing the architecture of the current day solar system is a factor of ~10 lower than in simulations without self-gravity. Self-gravity of minor planets would appear to be a crucial component in the dynamical evolution of the solar system.

We have, so far, focused on the linear phase of the inclination instability in an idealized set-up. In future studies, we will directly apply the inclination instability to the solar system and examine its nonlinear evolution (see Chapter 3). We will include the gravitational influence of the giant planets, and vary our initial conditions to reflect the scattered disk population and the inner Oort cloud formation in a stellar cluster. Finally, instead of presenting results averaged for many particles, we will focus on individual minor planets, particularly those that flip to retrograde orbits.

Chapter 3

Giant Planet Influence on the Collective Gravity of a Primordial Scattered Disk

This article has been published in the Astronomical Journal as: Zderic, A., & Madigan, A.-M. 2020, AJ, 160, 50, doi:10.3847/1538-3881/ab962f

Abstract Axisymmetric disks of high eccentricity, low mass bodies on near-Keplerian orbits are unstable to an out-of-plane buckling. This "inclination instability" exponentially grows the orbital inclinations, raises perihelion distances and clusters in argument of perihelion. Here we examine the instability in a massive primordial scattered disk including the orbit-averaged gravitational influence of the giant planets. We show that differential apsidal precession induced by the giant planets will suppress the inclination instability unless the primordial mass is $\gtrsim 20$ Earth masses. We also show that the instability should produce a "perihelion gap" at semi-major axes of hundreds of AU, as the orbits of the remnant population are more likely to have extremely large perihelion distances ($\mathcal{O}(100$ AU$)$) than intermediate values

3.1 Introduction

Structures formed by the collective gravity of numerous low-mass bodies are well-studied on many astrophysical scales, for example, stellar bar formation in galaxies (Sellwood & Wilkinson, 1993) and apsidally-aligned disks of stars orbiting supermassive black holes (Kazandjian & Touma, 2013; Madigan et al., 2018a). The driver of these dynamics are long-term (secular) gravitational torques between orbits.

The corresponding structures in planetary systems are relatively under-explored. This may be due to the presence of massive perturbers (planets) that are assumed to dominate the dynamics. While this is often the case on small scales close to the host star, there may be significant regions of phase space in which the influence of massive planets is small and the orbital period of bodies is short enough for collective gravitational torques to be important.

In Madigan & McCourt (2016) we presented the discovery of a gravo-dynamical instability driven by the collective gravity of low mass, high eccentricity bodies in a near-Keplerian disk. This "inclination instability" exponentially grows the orbital inclination of bodies while decreasing their orbital eccentricities and clustering their arguments of perihelion (ω).

In Madigan et al. (2018b) (Chapter 2) we explained the mechanism behind the instability: secular torques acting between the high eccentricity orbits. We also showed how the instability timescale scaled as a function of disk parameters. One important result is that the growth timescale is sensitive to the number of bodies used in N-body simulations. A low number of particles suppresses the instability due to two-body scattering and incomplete angular phase coverage of orbits in the disk.

We showed that the amount of mass needed for the collective gravity of extreme trans-Neptunian objects (eTNOs) to be the dominant dynamical driver in the outer Solar System ($\sim 100 - 1000$ AU) was about half an Earth mass. However, to observe significant clustering in ω within the age of the Solar System we required a mass closer to a few Earth masses. We note that this is very similar to the predicted mass of Planet 9 (Batygin & Brown, 2016a; Batygin et al., 2019). It is perhaps no coincidence that the mass requirements are the same as dynamics are driven by gravitational torques in both (Batygin & Morbidelli, 2017), but a disk of individually low mass bodies with high perihelion and inclinations will be harder to observe than a single massive body at the same distance. In Fleisig et al. (2020) we moved from simulations of a single mass population to a mass spectrum. In this paper we add two more additional complexities to the system: a more realistic orbital configuration and the gravitational influences of the giant planets. Our goal is to determine the parameters under which the presence of giant planets completely suppresses the inclination instability in a orbital configuration modeled on a primordial scattered disk (Luu et al., 1997; Duncan & Levison, 1997). We note that this was first addressed by Fan & Batygin (2017), who found that the inclination instability did not occur in their simulations of the Nice Model containing 30 Earth masses of self-gravitating planetesimals. These simulations, however, lacked a sufficient number of particles between 100 - 1000 AU ($N < 16$) for the inclination instability to occur. We find that differential apsidal precession induced by the giant planets can suppress the inclination instability in the scattered disk. However, if the mass of the primordial scattered disk is large enough ($\gtrsim$ 20 Earth masses) then the instability will occur.

In Section 3.2, we describe our N-body simulations including how we emulate the influence of the giant planets with a quadrupole (J_2) potential. In Section 3.3, we discuss how the instability is changed by the J_2 potential and show results for a primordial scattered disk configuration. We also discuss the generation of a "perihelion gap" at hundreds of AU. In Section 3.4, we scale our results to the solar system, obtaining an estimate for the required primordial mass of the scattered disk for the inclination instability to have occurred within it. Finally, in Section 3.5, we summarize our results and discuss the implications of our work.

3.2 Numerical Methods

3.2.1 N-body Simulations

To study the collective gravitational effects of minor bodies in the outer Solar System we run simulations using `REBOUND`, an open-source N-body integration framework available in C with a Python wrapper. `REBOUND` offers a few different integration methods and gravity algorithms (Rein & Liu, 2012). For this work, we use the direct gravity algorithm (N^2 scaling) and the IAS15 adaptive time-step integrator. We also use the additional package `REBOUNDx` which provides a framework for adding additional physics (e.g. general relativity, radiation forces, user-defined forces) (Tamayo et al., 2020).

3.2.2 JSUN as $J_{2,\mathrm{Sun}}$

The most straight-forward way to incorporate the giant planets would be to simulate them directly as N-bodies. However, this is much harder to do than it might seem. Out of computational necessity we simulate implausibly large particle masses and scale our results to realistic values. If we wanted to simulate the correct mass ratio between giant planets and the disk, the mass ratio between the Sun and the giant planets would be too small, in which case the potential would no longer be near-Keplerian. If we were to simulate a more realistic disk mass, the simulations would take proportionally longer and we would need to use fewer particles. Scattering interactions between disk particles and the planets would naturally depopulate the disk, further reducing numerical resolution of the simulation. The instability cannot be captured at low particle numbers (Madigan et al., 2018b, Chapter 2).

Our solution is to model the Sun and the giant planets with a multipole expansion keeping only the two largest terms, the monopole and quadrupole term (the dipole term is zero in the center of mass frame). We ignore the contributions of the planets to the monopole term because this results in a negligible (1 part in thousand) change in the Sun to disk mass ratio. In spherical coordinates (r, θ, ϕ), the multipole expansion potential is,

$$\Phi(r,\theta) = -\frac{GM}{r}\left(1 - \frac{J_2 R^2}{r^2}\, P_2\left(\cos\theta\right)\right) \tag{3.1}$$

where J_2 is a weighting factor for the quadrupole moment, R is the mean radius of the mass distribution, and $P_2\left(\cos\theta\right)$ is the $n = 2$ Legendre polynomial. The first term in the parentheses is the monopole term and the second is the quadrupole term. For the giant planets, the orbit-averaged quadrupole moment is given by,

$$J_2 = \frac{1}{2M_\odot R^2}\sum_{i=1}^{4} m_i a_i^2 \tag{3.2}$$

where i iterates over the giant planets (Batygin & Brown, 2016a). We further assume that the Sun's inherent J_2 moment is negligible compared to the contributions of the giant planets.

Equation 3.1 is not a general multipole expansion; we have already implicitly assumed there is no longitudinal (ϕ) dependence in Φ. Thus, this expansion assumes the giant planet's orbits have no inclination or eccentricity and their mass is spread out along their orbit. Formally, a multipole expansion only converges to the actual potential for $r > d$ where d is the size of the system. Thus, the mathematically correct method would be to set $R = a_N$, the semi-major axis of Neptune, and remove any particles that went inside R.

Due to the artificially strong self-stirring in our low-N, large-mass disks, most bodies in our simulations violate this requirement during integration, and, if we removed them, we would end up with a depopulated disk that is numerically unable to undergo the instability. Therefore, we ignore this convergence requirement.

We do not use the actual J_2 value of the giant planets because of the unrealistic disk mass and N used in our simulations. For a given set of simulations, we fix the number of particles, N and the mass of the disk, M_d, and vary the J_2 value until we find the instability is suppressed. To extrapolate our results to the solar system, we determine how these J_2 values scale with N, M_d, and orbital configuration of the disk.

3.2.3 Initial Orbit configurations

In this paper we simulate two distinct systems, a compact configuration and a scattered disk configuration. The compact configuration is a thin, mono-energetic disk of orbits (nearly identical semi-major axes), and the scattered disk configuration models a population of bodies with equal perihelion and an order-of-magnitude range in semi-major axis.

In the compact configuration, the disk of orbits is initialized to have a semi-major axis a distribution drawn uniformly in $[0.9, 1.1]$, eccentricity $e = 0.7$, and inclination $i = 10^{-4}\,\mathrm{rad}$. The disk is initially axisymmetric (ω and Ω and mean anomaly, $\mathcal{M}$, drawn from a uniform distribution in $[0, 2\pi]$). The total mass of the disk is $10^{-3}\,M$ and the number of disk particles, $N = 400$. This configuration is ideal for physical analysis.

In the scattered disk configuration, orbits are initialized with an order-of-magnitude range in semi-major axes and identical perihelion distances. Specifically, we draw the orbit's semi-major axis from an a^{-1} distribution in the range $[1, 10]$, define eccentricity e from the relation $e = 1 - {}^{p}\!/_{a}$ for a chosen perihelion p, and draw inclination i from a Rayleigh distribution with a mean inclination of $5°$.[1] The disk is initially axisymmetric with $\mathcal{M}$ drawn uniformly in the range $[0, 2\pi]$.

We look at two different scattered disk configurations: 'sd100' and 'sd250'. The 'sd100' configuration represents a scattered disk with inner-most semi-major axis of 100 AU and a perihelion of 30 AU, while the 'sd250' configuration represents a scattered disk with the same perihelion but an inner-most semi-major axis of 250 AU.

We run these two simulations to explore the effect of distributing the peak of the mass density of the scattered disk in a different location. Apsidal precession due to the J_2 moment is a steep function of semi-major axis, $a^{-7/2}$; perhaps the gravitational torques between orbits in a scattered disk with peak mass density at larger radius can better resist the differential precession from the giant planets?

The Newtonian N-body problem is scale-free. Simulation times are presented in units of the secular timescale,

$$t_{\rm sec} \sim \frac{1}{2\pi}\frac{M}{M_{\rm d}}P, \tag{3.3}$$

where P is the orbital period at $a = 1$. In this paper, $M_d = 10^{-3}\,M$ such that $t_{\rm sec} \approx 160\,P$. $a = 1$ may be scaled to, for example, $a = 100$ AU with the conversion $P = 1000$ yr. The J_2 potential is not scale-free however. In simulations with added J_2 we appropriately scale the semi-major axes, a_i, in equation 3.2, and present our results in solar system units (distances in AU etc.).

3.3 J2 and the Inclination Instability

The inclination instability timescale, $t_{\rm e-fold}$, scales linearly with the secular time. It also depends non-trivially on N and orbital configuration. We use the orbital angle coordinates defined in Madigan & McCourt (2016) to describe the instability and quantify its timescale. The angles represent rotations of the orbit about its semi-major ($\hat{\boldsymbol{a}}$) axis, semi-minor ($\hat{\boldsymbol{b}} \equiv \hat{\boldsymbol{j}} \times \hat{\boldsymbol{a}}$) axis and

---

[1] The instability can occur in scattered disk simulations with initial inclinations drawn from Rayleigh distributions with means up to $\sim 15°$, but it's hard to measure the instability growth rate in these systems because the instability is linear for a short time.

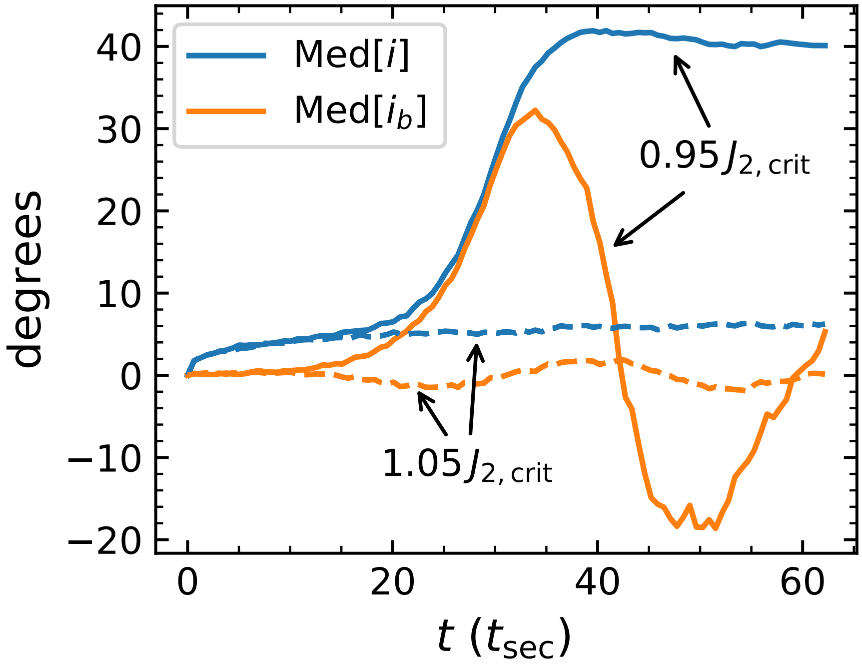

Figure 3.1: Median inclination (blue) and i_b (orange) of disk orbits for two different simulations, one with J_2 less than $J_{2,\mathrm{crit}}$ (solid) and one with J_2 greater than $J_{2,\mathrm{crit}}$ (dashed), vs. time measured in secular times ($t_{\mathrm{sec}} \approx 160\,P$). The simulation with added J_2 less than $J_{2,\mathrm{crit}}$ is susceptible to the inclination instability while the other is not. The two simulations have very similar J_2 values, only different by about 20%, showing the abruptness of the transition from unstable to stable.

angular momentum vector ($\hat{\boldsymbol{j}}$), respectively,

$$i_{\mathrm{a}} = \arctan\left[\frac{\hat{b}_{\mathrm{z}}}{\sqrt{1-\hat{b}_{\mathrm{z}}^2}}\right], \tag{3.4a}$$

$$i_{\mathrm{b}} = \arctan\left[-\frac{\hat{a}_{\mathrm{z}}}{\sqrt{1-\hat{a}_{\mathrm{z}}^2}}\right], \tag{3.4b}$$

$$i_{\mathrm{e}} = \arctan\left[\hat{\boldsymbol{a}}_{\mathrm{y}}, \hat{\boldsymbol{a}}_{\mathrm{x}}\right]. \tag{3.4c}$$

The subscripts x, y, and z denote an inertial Cartesian reference frame with unit vectors, $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$. These angular coordinates are useful for understanding the effect of torques on orbits.

The inclination instability is characterized by exponential growth in median i_{a} and i_{b} with opposite sign (i.e. if i_{a} increases to positive values, i_{b} increases to negative values). A constant ratio $i_{\mathrm{b}}/i_{\mathrm{a}}$ implies a constant angle of perihelion, as for small inclinations $\omega(i_{\mathrm{a}}, i_{\mathrm{b}}) \sim \arctan|i_{\mathrm{b}}/i_{\mathrm{a}}|$ ($+\pi$ if $i_{\mathrm{a}} < 0$). We use the exponential growth of median i_{b} as a diagnostic for the instability and define the inverse of its growth rate, γ, as $t_{\mathrm{e-fold}}$. As orbits incline, their eccentricities decrease. This means that the magnitude of the angular momentum vectors of all orbits increase (semi-major axes remain constant apart from scatterings due to two-body relaxation). This may seem counter-intuitive at first, but the *vector* sum of all the angular momenta is conserved.

3.3.1 Compact Configuration

A sufficiently large J2 value suppresses the inclination instability. We call the threshold value above which the disk does not undergo the instability $J_{2,\mathrm{crit}}$. For $J_2 < J_{2,\mathrm{crit}}$, we find two different regimes:

(1) The 'instability-dominated region' defined by $J_2 \leq 0.1\, J_{2,\mathrm{crit}}$. Here the system is unaffected by the additional J_2.

(2) The "transition region" defined by $(0.1 - 1)J_{2,\mathrm{crit}}$. Here the dynamics of the instability are altered by the presence of the J_2, but the instability still occurs.

In Figure 3.1, we plot the median inclination i and i_{b} of a disk of particles in two simulations,

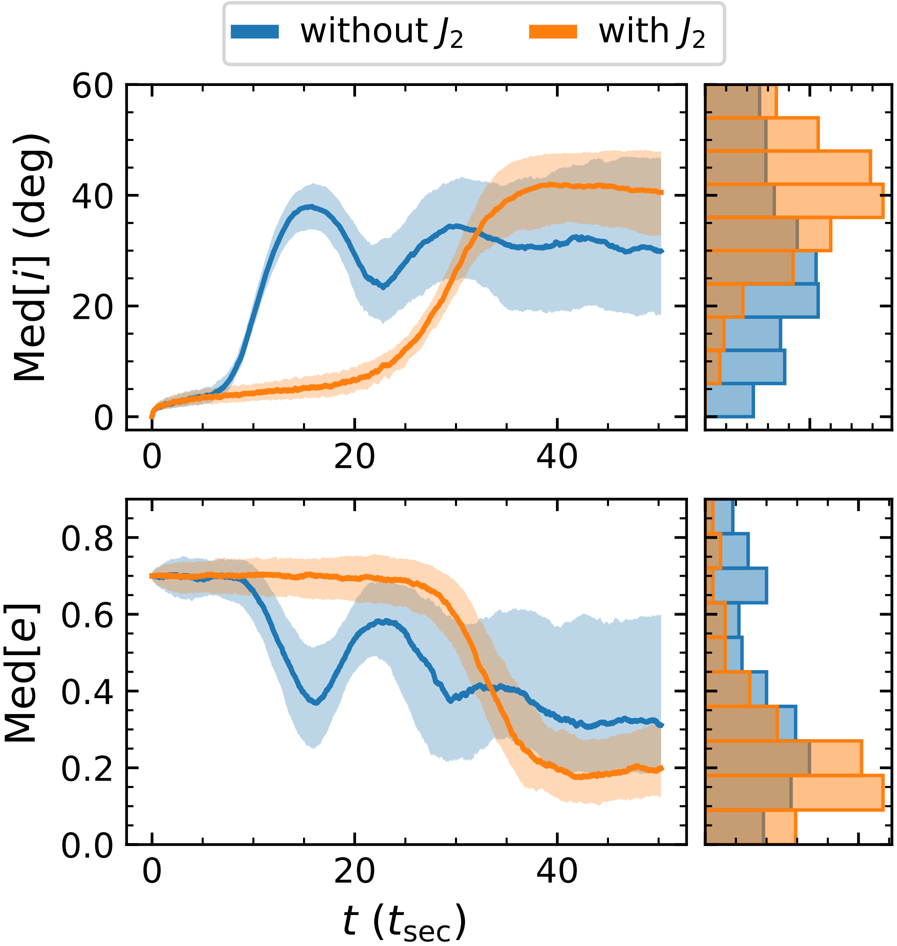

Figure 3.2: Median inclination (deg) and eccentricity for two simulations, one with added J_2 and one without J_2. The "with J_2" simulation has an added J_2 in the transition region ($J_2 = 0.9\,J_{2,\rm crit}$). The left column shows the time evolution of the median orbital elements with their upper and lower quartiles. The right column is a histogram showing the distribution of the particle's orbital elements in the simulation at the end of the plotted time evolution ($\sim 50\,t_{\rm sec}$). The "with J_2" simulation has a smaller growth rate than the "without J_2" simulation, but it reaches a higher median inclination and lower median eccentricity post-instability.

one with $J_2 = 0.95\ J_{2,\rm crit}$ and another with $J_2 = 1.05\ J_{2,\rm crit}$. This figure shows that the inclination instability is suppressed for $J_2 > J_{2,\rm crit}$ and that the transition around $J_{2,\rm crit}$ is rapid, with the inclination behavior of the disk changing dramatically for only slight changes ($\sim 20\%$) in the value of added J_2.

In Figure 3.2, we show that the average post-instability orbital elements of the disk are different in the transition region ($J_2 = 0.9\, J_{2,\rm crit}$). The orbits attain higher (lower) post-instability inclinations (eccentricities) on average than systems with no/low J_2, despite the fact that the instability growth rate is reduced by the added J_2. The right columns show histograms of the orbital elements at 50 secular times. The histograms are limited in range for clarity; two out of eight hundred bodies have reached polar orientations of $i \gtrsim 90^\circ$.

In Figure 3.3, we show that the growth rate of the instability decreases across the transition region. At $\sim 0.1\, J_{2,\rm crit}$, the growth rate of the instability is identical to the instability with no J_2 moment, and at $J_{2,\rm crit}$, the instability has a growth rate of zero. Above $J_{2,\rm crit}$, we find that the median i_b of the disk oscillates rather than grows exponentially; the growth rate is imaginary.

In Figure 3.4, we show the effects of J_2 on the clustering of argument of perihelion, ω, in the compact configuration using the Kuiper test, a variation of the Kolmogorov-Smirnov test that is applicable to circular quantities (Kuiper, 1960). We use the test to compare the simulation ω distribution to a uniform distribution for two different simulations, one with J_2 in the transition region and one with no added J_2. We show the test statistic value for a p-value of 0.05. Larger test statistic values correspond a greater likelihood that the simulation ω distribution is not uniform. In both simulations, the test statistic is initially consistent with a uniform distribution. Within a single orbit, the system develops a bi-modal distribution in ω with peaks at 0° and 180° due to small oscillations in i_a. Later, the test statistic increases to a large peak as the instability clusters the orbit's ω. Post-instability, the ω-clustering is not maintained, and differential precession washes out the clustering.

Surprisingly, the duration of ω-clustering isn't significantly changed in the transition region. The J_2 potential term causes prograde ($\dot{\omega} > 0$) precession, and the *post-instability* disk potential

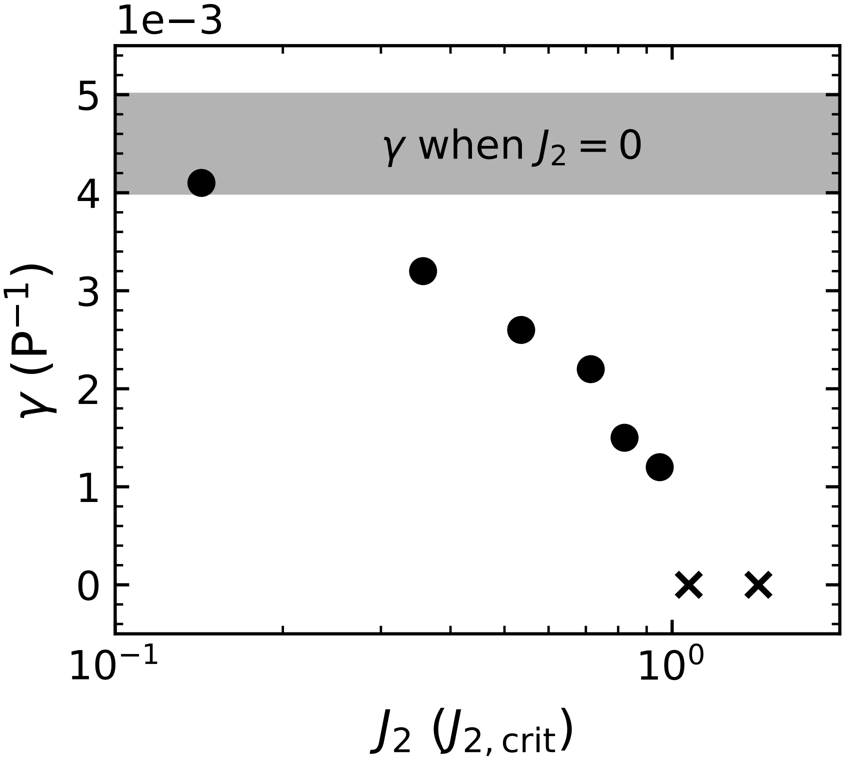

Figure 3.3: Growth rate of the inclination instability (γ) as a function of added J_2 moment. For $J_2 \lesssim 0.1 J_{2,\text{crit}}$ the growth rate of the instability is the same as if there were no added J_2. In the region between 0.1 to 1.0 $J_{2,\text{crit}}$ the growth rate steadily drops until the instability disappears for $J_2 > J_{2,\text{crit}}$. Above this, the disk is stable and the growth rate becomes imaginary as signified by the change in marker.

causes retrograde precession. One might expect that the two competing sources of precession would reduce the overall precession rate, and increase the duration of ω-clustering. However, this is not what we see. When J_2 is added to the system, the growth rate slows, and the rise time to peak clustering increases. The mean ω precession rate decreases, but the *differential* precession rate increases. Thus, the ω-clustering is washed out faster. Overall, the duration of clustering is relatively unchanged in the compact configuration.

In summary, we find that the addition of a J_2 term to the Keplerian potential suppresses the inclination instability above a critical value, $J_{2,\text{crit}}$. As this critical value is approached from below, the post-instability orbital elements and growth rate of the instability are changed in a transition

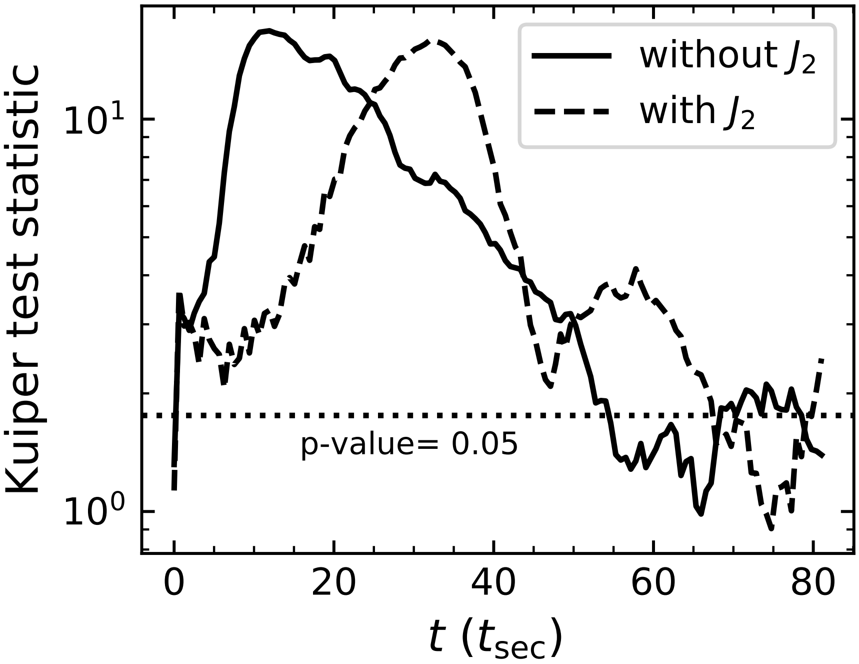

Figure 3.4: Kuiper's test statistic comparing the ω distribution of N-body simulations to a uniform distribution as a function of time. Two simulations are shown, one without J_2 and another with added J_2 in the transition region ($J_2 = 0.8 J_{2,\mathrm{crit}}$). A horizontal line marks the test statistic value for a p-value of 0.05. The test statistic reflects the dynamical behavior of ω over the course of the simulation with the obvious peak corresponding to the peak clustering during the instability.

region, $0.1\,J_{2,\mathrm{crit}}$ to $J_{2,\mathrm{crit}}$. Finally, we find that the duration of ω-clustering is unchanged in this transition region.

3.3.2 Scattered Disk Configuration

In previous publications, we focused on the compact configuration for ease of analysis. However, we have explored the inclination instability in a range of different orbital initial conditions. Our findings can succinctly be summarized: compact systems with mean eccentricity $\gtrsim 0.5$ and/or mean inclinations $\lesssim 20°$ are unstable, systems with an order of magnitude spread in semi-major axis *with constant eccentricity* are either stable or have very small growth rates, and systems with an order of magnitude spread in a, but ${}^{de}/_{da} > 0$, i.e. the scattered disk, are unstable.

In Figure 3.5, we show the inclination instability in a system with scattered disk initial

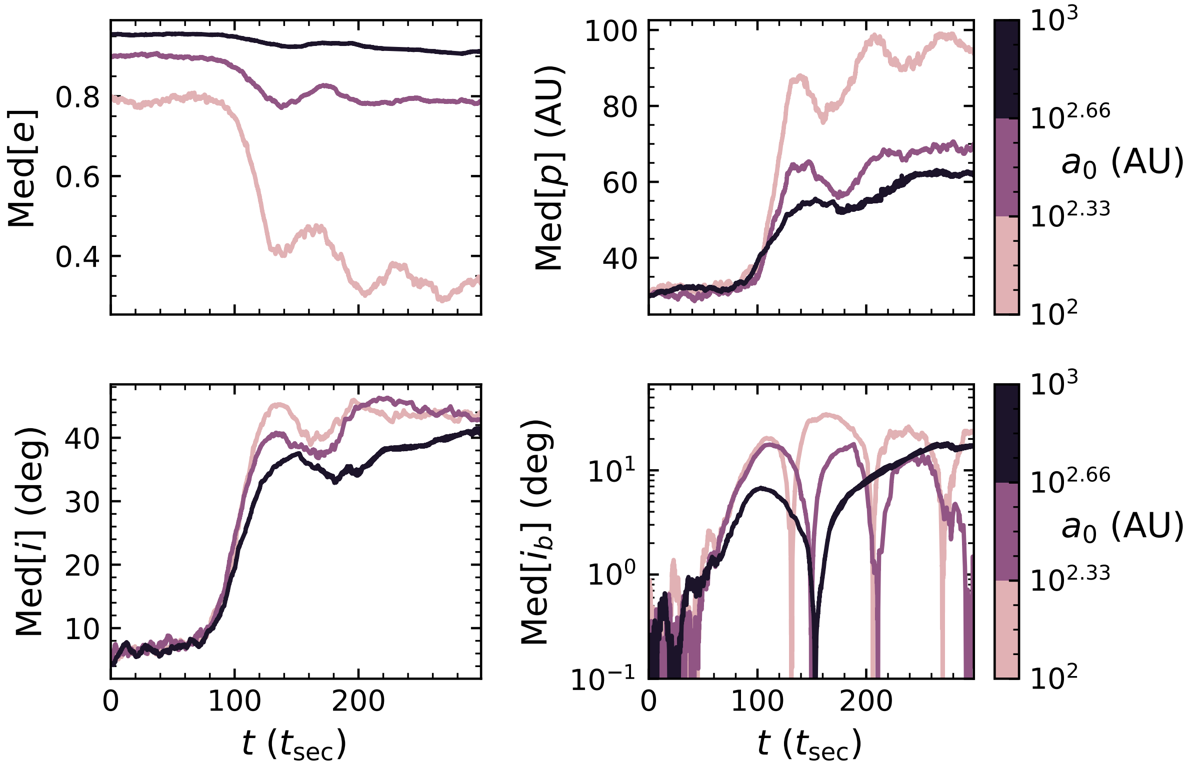

Figure 3.5: Median eccentricity (top left), perihelion distance (top right), inclination (bottom left) and $i_{\rm b}$ (bottom right) as a function of time in units of secular times for a scattered disk ('sd100') simulation with no J_2 moment. Orbits have been binned by their initial semi-major axis. The lower the initial eccentricity/semi-major axis, larger the change in eccentricity/perihelion during the instability. All bodies attain a similar post-instability inclination, with the larger semi-major axis bodies attaining slightly lower final inclinations. The larger semi-major axis orbits have slightly lower instability growth rates (slope of Med[i_b] during the instability) than smaller semi-major axis orbits, but the instability begins at the same time at all radii.

conditions ('sd100') and no J_2 moment. The disk undergoes the instability simultaneously at all radii, though the orbits at larger semi-major axis have a slightly smaller growth rate. In general, the lower the initial eccentricity (and semi-major axis) of the orbit, the larger the change in eccentricity during the instability and the larger the final perihelion distance. The final median inclination is similar for all semi-major axis bins ($i \approx 40^\circ$).

Figure 3.6 shows the same information as Figure 3.5, but for a simulation with added J_2 in the transition region (in this case $J_2 \sim 0.9\, J_{2,\mathrm{crit}}$). Again, the instability occurs simultaneously throughout the disk. The smallest semi-major axis bin is barely unstable, however, and has a lower post-instability inclination and a higher eccentricity. This is due to the significant differential apsis precession caused by the added J_2. The larger semi-major axis bins have larger post-instability inclinations ($i \approx 60^\circ$), lower eccentricities ($0.5 \lesssim e \lesssim 0.85$), and larger perihelia ($100\,\mathrm{AU} \lesssim p \lesssim 150\,\mathrm{AU}$) than they do in simulations without J_2.

Overall, the addition of the J_2 moment to simulations has a similar effect on the scattered disk orbital configuration as it has on the compact configuration, i.e., increased (decreased) post-instability inclination (eccentricity) and reduced instability growth rate. One significant difference is the inner-most part of the disk barely undergoes the instability. Indeed, if we simulate the inner portion of the disk ($a \in [100, 200]\,\mathrm{AU}$) without the outer portion it does not undergo the instability at all. The inner portion of the disk is being pulled along by the outer portion as the outer portion undergoes the instability. We can think of this as the disk having two components, a stable component and an unstable component. The inner-most part is stabilized by differential precession from the J_2 moment while the outer portion is still unstable (J_2 precession has a steep $a^{-7/2}$ dependence). Below $J_{2,\mathrm{crit}}$, the inner-most component is small enough that it can be coerced into instability by the outer-most portion. At the critical J_2, the stable, inner-most component of the disk is massive enough that the outer portion of the disk is held back from lifting out of the plane. The inclination instability is a global phenomenon, and we find that the disk as a whole is stabilized if $\sim 30\%$ of the mass is in the stable component.

We find that the duration of ω-clustering in the scattered disk simulations is enhanced by the

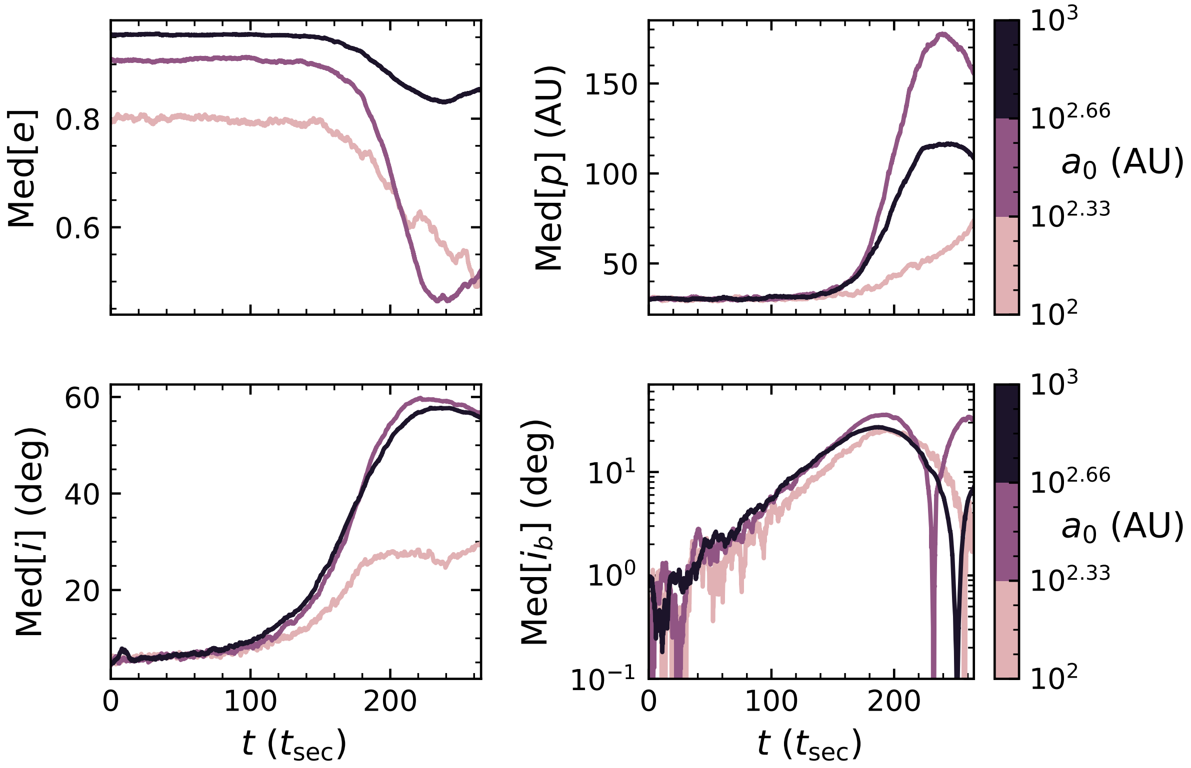

Figure 3.6: Median eccentricity (top left), perihelion distance (top right), inclination (bottom left) and $i_{\rm b}$ (bottom right) as a function of time for a scattered disk ('sd100') orbital configuration with J_2 in the transition region. The orbits have been binned by their initial semi-major axis. Compared to Figure 3.5, the growth rate of the instability is smaller and all bins have the same growth rate. The inner-most semi-major axis bin is barely unstable (with a reduced post-instability inclination), and the outer bins have larger post-instability perihelia and larger post-instability inclinations.

addition of J_2 in the transition region, in contrast to our findings for the compact configuration. In Figure 3.7, we plot the argument of perihelion, ω, as a function of semi-major axis, a, for each disk orbit at three different times corresponding to the simulations shown in Figures 3.5 and 3.6. At the beginning of each simulation, the distribution of ω values is uniform. Later, the instability causes ω values to cluster. After $\sim 260\, t_{\rm sec}$, the ω values for $a \gtrsim 200\,\mathrm{AU}$ are significantly less clustered in the simulation without J_2 while the simulation with J_2 still retains significant ω-clustering. This difference is due to both the reduced growth rate (and delayed instability saturation time) of the instability due to the added J_2 potential and the reduced differential precession rate in the $a \gtrsim 200\,\mathrm{AU}$ portion of the disk due to competition between the disk and J_2 precession.

The global nature of the instability has an interesting consequence on the perihelion distribution of the post-instability orbits. As we see in Figure 3.5 in which all the orbits are unstable, orbits of different semi-major axes end up with similar mean inclinations. Specific orbital angular momentum increases with semi-major axis across the scattered disk ($\sim 10\%$ change from 100 to 1000 AU). This means that as orbits incline, those at lower semi-major axis will gain a larger fractional increase in orbital angular momentum than those at higher semi-major axis. This results in orbits at lower semi-major axis decreasing their eccentricities and increasing their perihelia more so than those at higher semi-major axis. This naturally generates a perihelion gap at the inner edge of the disk that has undergone the instability.

In Figure 3.8, we show a 2D histogram of time-averaged post-instability values of semi-major axis vs perihelion for the two simulations shown in Figures 3.5 and 3.6 (scattered disk configurations without J_2 and with J_2). We use a time-average to get sufficient numerical resolution to make this plot. We've added a red box to the figures to show the observed perihelion gap between VP_{113} and Sedna and the rest of the minor bodies (see Figures 1 and 2 in Kavelaars et al. (2020)). In the simulation with J_2, the inclination instability empties the region corresponding the observed perihelion gap. The size of the region vacated by the inclination instability is related to the magnitude of J_2, a larger J_2 vacates a larger region of a-p space.

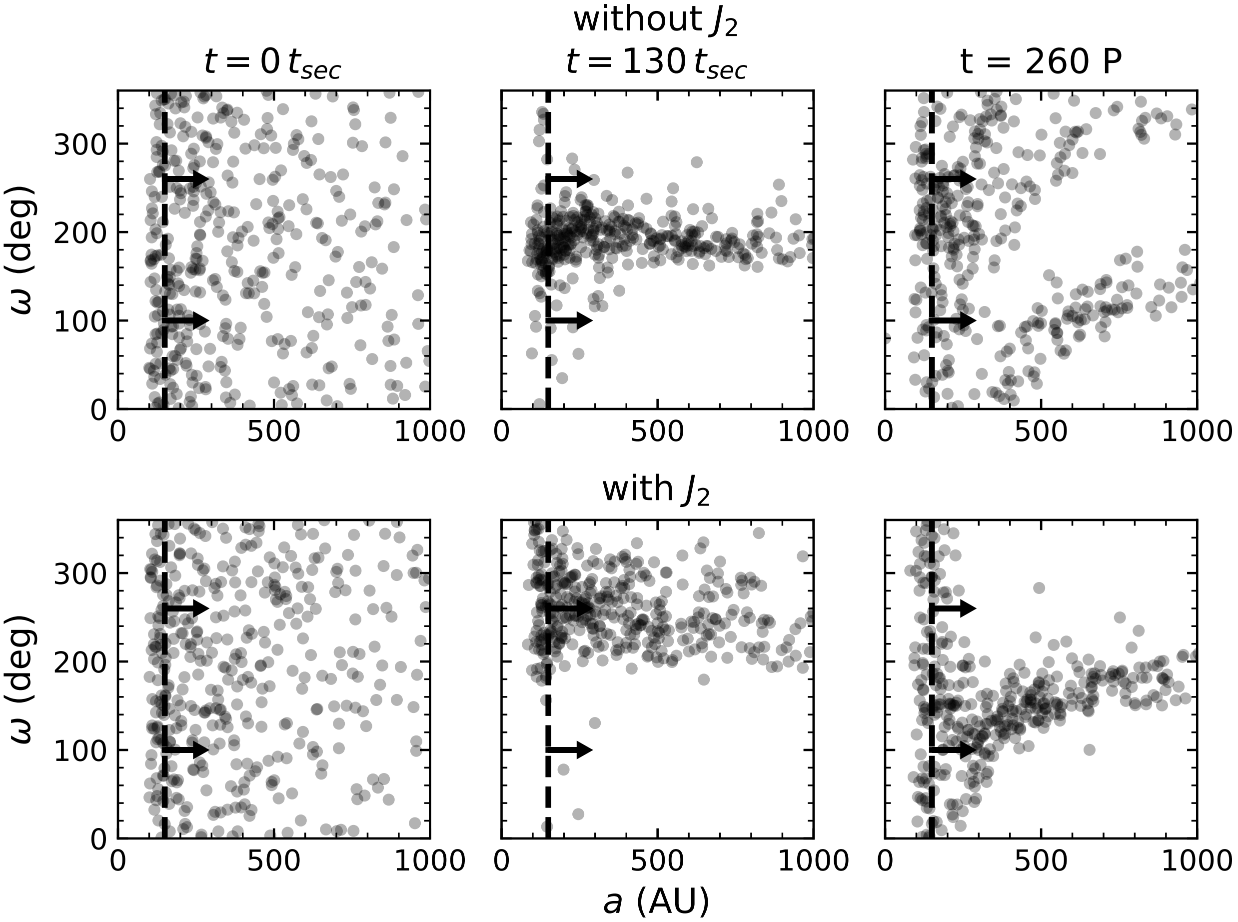

Figure 3.7: Argument of perihelion, ω, as a function of semi-major axis, a, for a simulation without J_2 (top) and one with added J_2 in the transition region (bottom), corresponding to Figures 3.5 and 3.6. The vertical dashed line marks $a = 150\,\mathrm{AU}$. (Left column) orbits initially have a uniform random ω distribution. (Middle column) ω is clustered while orbits undergo the inclination instability. (Right column) ω-clustering is lost in the simulation without J_2 due to differential precession, but it is maintained for $a \gtrsim 200\,\mathrm{AU}$ in the simulation with J_2 due to the reduced differential precession and instability growth rate.

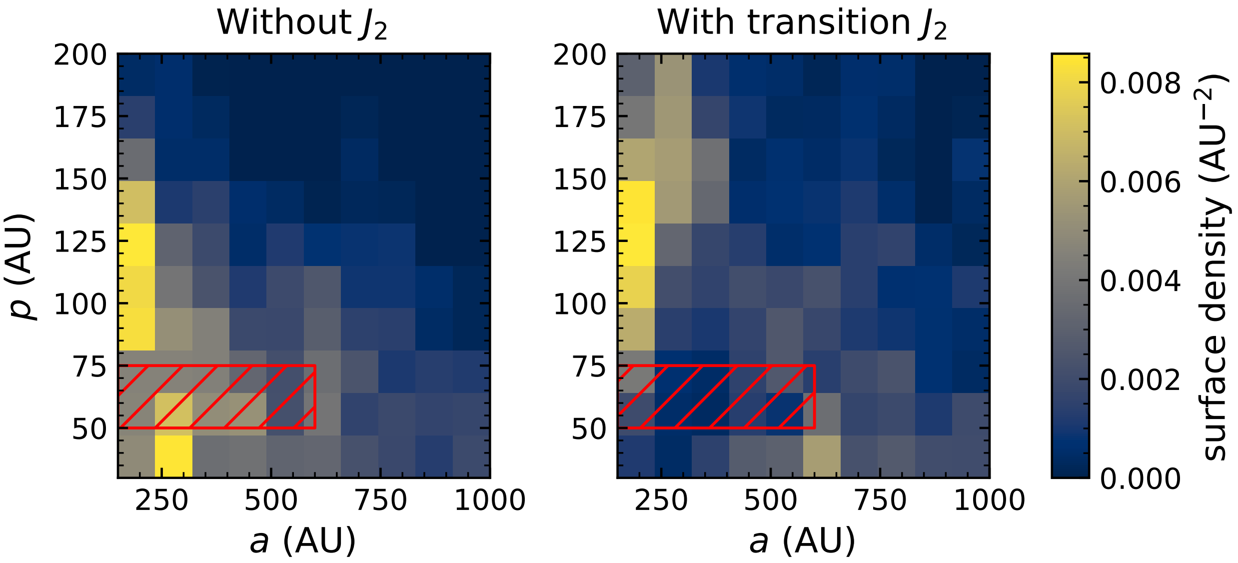

Figure 3.8: Time-averaged surface density in the semi-major axis (a) - perihelion (p) plane for two scattered disk ('sd100') simulations, one without J_2 and one with J_2 in the transition region ($J_2 \sim 0.9\, J_{2,\mathrm{crit}}$). For this plot, we histogram the a and p of each particle in the simulation at all times post-instability. Areas of high density will have more observable bodies than areas of low density. Also drawn is a red box covering the observed "perihelion gap" in the solar system. When J_2 is added to the simulations in the transition region, the region of a-p space corresponding to the observed perihelion gap in the solar system is vacated by the inclination instability. The size of the gap changes with the added J_2; larger J_2 values produce a larger gap.

3.4 Scaling to the Solar System

Our goal in this section is to explain how $J_{2,\mathrm{crit}}$ depends on system parameters such as number of particles, N, mass of disk, M_d, and initial orbital configuration, which allows us to then extrapolate our simulation results to the solar system.

The instability mechanism relies on a secular average where the individual bodies can be approximated as rings in the shape of the body's osculating Keplerian orbit with a linear mass density inversely proportional to its instantaneous velocity. The validity of this average depends on how quickly the body's osculating Keplerian orbital elements change. If the osculating orbit changes rapidly the approximation fails and the mutual secular torques responsible for the instability weaken to the point that the instability can no longer occur. As the instability relies on inter-orbit secular torques, it is inter-orbit or differential apsidal precession that is responsible for the weakening of the secular torques. The addition of the quadrupole term to the potential increases differential apsidal precession within the disk. However, the disk itself also causes apsidal precession in its constituent orbits, and this source of apsidal precession must be considered in combination with that from the J_2.

The addition of the quadrupole term causes secular changes in the ω and Ω of the orbits in the disk. Assuming the added quadrupole term is a small perturbation on the potential of the central body, the evolution of the osculating Keplerian elements of orbits in the potential can be determined with the disturbing function formalism and the Lagrange planetary equations,

$$\dot{\omega} = \frac{3J_2}{4}\, n\, \frac{R^2}{a^2}\, \frac{5\cos^2 i - 1}{(1-e^2)^2}, \tag{3.5a}$$

$$\dot{\Omega} = -\frac{3J_2}{2}\, n\, \frac{R^2}{a^2}\, \frac{\cos i}{(1-e^2)^2}, \tag{3.5b}$$

where n is the mean motion of the body, $n^2 a^3 = \mu = GM$. $\varpi = \omega + \Omega$ gives the apsidal angle of the orbit and $\dot{\varpi}$ the apsidal precession rate. Therefore, the apsidal precession rate from the added J_2 is,

$$\dot{\varpi}_{J2} = \frac{3J_2}{4}\, n\, \frac{R^2}{a^2}\, \frac{5\cos^2 i - 2\cos i - 1}{(1-e^2)^2}. \tag{3.6}$$

For $i \lesssim 46°$, which holds for our initial disk configurations, apsis precession due to J_2 is prograde (with respect to orbital motion). In the compact configuration, $e \approx 0.7$ and $i \approx 0$. In Figure 3.9, we show Equation 3.6 with these approximations as a dotted line. For reference, $J_{2,\rm crit} = 0.26$ for this simulation (length scaled to 100 AU).

The disk potential induces retrograde precession (see Appendix B for differences between orbital configurations) whereas added J_2 potential induces prograde precession. The two precession sources compete and the mean precession rate is reduced. However, secular torques are not always strengthened by the added J_2 potential. In the compact configuration, $|\dot{\varpi}_{J2}|$ decreases with semi-major axis and the precession rate due to the disk, $|\dot{\varpi}_d|$, increases with semi-major axis. The result, as shown in Figure 3.9 with a solid line, is an amplified differential precession rate (slope). Thus, in this orbital configuration, the added J_2 weakens the gravitational torques between orbits which hinders the growth of the instability. On the other hand, in the scattered disk configurations

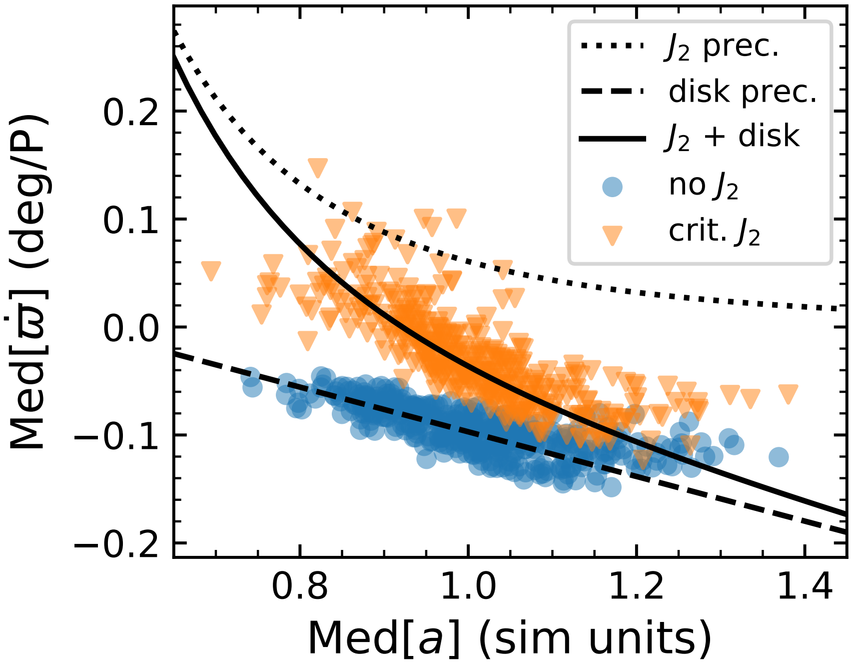

Figure 3.9: Median apsidal precession rate as a function of median semi-major axis (calculated from the first ~ 6 secular times) of a simulation with a compact configuration. Orange triangles show apsidal precession rates from a simulation with $J_2 \approx J_{2,\mathrm{crit}}$, and blue dots show one without J_2. The lines on the plot show models for the different precession sources. The dotted line shows the J_2 contribution to $\dot{\varpi}$, the dashed shows the disk contribution, and the solid line shows the sum of the dashed and dotted lines. The differential precession rate (slope of solid line) is enhanced by the presence of the added J_2 although the average precession rate is reduced.

$|\dot{\varpi}_d|$ decreases with semi-major axis. Therefore, the scattered disk configuration does a better job of resisting the added J_2. For example, with $N = 400$ and $M_\mathrm{d} = 10^{-3}\,M$, the scattered disk configuration has $J_{2,\mathrm{crit}} = 0.325 \pm 0.025$ while the compact configuration has $J_{2,\mathrm{crit}} = 0.258 \pm 0.008$ (length scaled to 100 AU). Despite having a much lower mass density (and growth rate), the scattered disk configuration handles the added J_2 better than the compact configuration.

We find that $J_{2,\mathrm{crit}}$ corresponds to the point where the differential apsidal precession and the inclination instability timescales are comparable. The timescale for the instability is defined as the inverse of the exponential growth rate, $t_{\mathrm{e-fold}} = \gamma^{-1}$, which we obtain from simulations by fitting the exponential growth of the median i_b of the disk orbits (Madigan & McCourt, 2016; Madigan

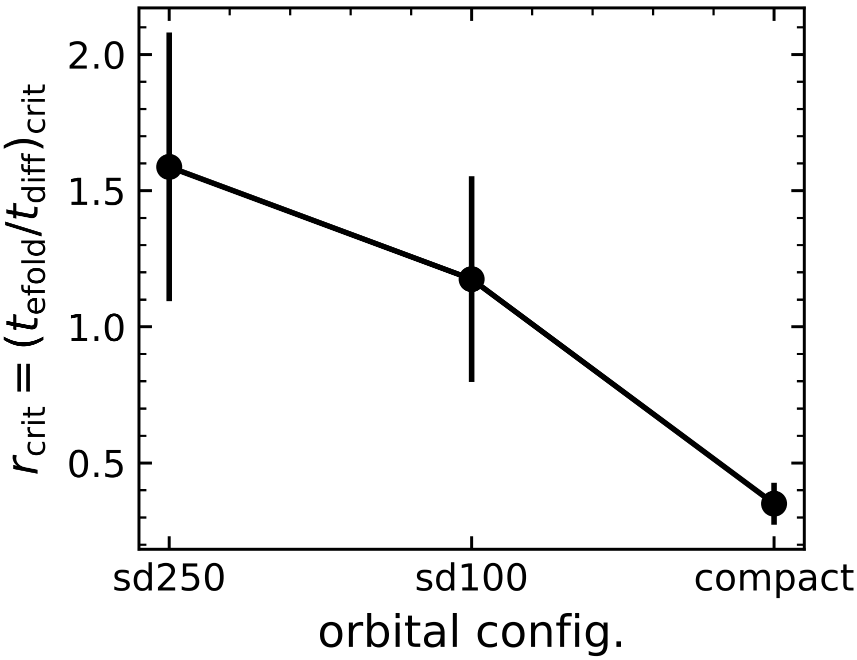

Figure 3.10: Ratio of e-folding to differential apsidal precession timescale at the critical J_2 where the instability is suppressed for different simulation initial conditions, r_{crit}. $t_{\mathrm{e-fold}}$ is measured when $J_2 = 0$ and t_{diff} is measured when $J_2 \approx J_{2,\mathrm{crit}}$. The timescales are roughly equal ($r_{\mathrm{crit}} \sim 1$) at the transition from instability to stability. The scattered disk simulations have a larger r_{crit} indicating that these orbital configurations can resist more differential precession relative to their instability growth rate than the compact configuration.

et al., 2018b, Chapter 2). We define the differential precession timescale as,

$$t_{\mathrm{diff}} = \frac{1\ \mathrm{rad}}{\dot{\varpi}_{\mathrm{diff}}} \tag{3.7}$$

where $\dot{\varpi}_{\mathrm{diff}}$ is the total apsidal differential precession rate of the disk. We calculate $\dot{\varpi}_{\mathrm{diff}}$ from our simulations by calculating the difference in precession rate between the fastest precessing quartile of the disk and the slowest precessing quartile, $\dot{\varpi}_{\mathrm{diff}} = \dot{\varpi}_{\mathrm{uqrt}} - \dot{\varpi}_{\mathrm{lqrt}}$. The choice to look at quartiles comes from our observation that if about $\sim 30\%$ of the orbits in the disk can't undergo the instability for any reason, the whole disk will fail to undergo the instability.

We define the ratio of these two timescales as

$$r = \frac{t_{\mathrm{e-fold}}}{t_{\mathrm{diff}}}. \tag{3.8}$$

The instability should occur for $r \ll 1$ and should be suppressed for $r \gg 1$. With numerous runs, we locate the $J_{2,\mathrm{crit}}$ in simulations with different N, M_d, and orbital configurations and calculate $r(J_{2,\mathrm{crit}}) = r_\mathrm{crit}$.

We find that r_crit is roughly constant with N provided that $N > 100$. For low N, self-stirring within the disk causes a wide spread in semi-major axes and eccentricity which artificially amplifies differential precession. In addition, we find that r_crit is constant with mass of the disk. Therefore, the r_crit measured at $N = 400, M_\mathrm{d} = 10^{-3}M$ in the scattered disk simulations should be consistent with r_crit as $N \to \infty$ for all M_d.

We find that r_crit does change with the orbital configuration. This is shown in Figure 3.10. Here 'compact' refers to the compact configuration while 'sd100' and 'sd250' refers to the scattered disk orbital configurations discussed in section 3.3.2. For each configuration, $N = 400$ and $M_\mathrm{d} = 10^{-3}\,M$. From this figure, we see that $r_\mathrm{crit} \approx 1$ as expected. Large r_crit values mean that in one e-folding time the disk orbits have differentially precessed by more than a radian with respect to one another, meaning that this configuration is more resistant than expected to added J_2. Small values of r_crit mean that the system is less resistant, the disk orbits having precessed less than a radian in one e-folding time. Notably, the compact configuration is worse at resisting added J_2 ($r_\mathrm{crit} \sim 0.3$) than the scattered disk configurations ($r_\mathrm{crit} \sim 1$).

To scale our results to the solar system, we find the timescale ratio $r = {}^{t_{\mathrm{e-fold}}}\!/_{t_\mathrm{diff}}$ in the solar system for different disk masses, and compare it to the r_crit calculated for 'sd100' and 'sd250' (i.e. the points shown in Figure 3.10). The e-folding timescale for the instability in large N, low M_d, *compact* systems is calculated in Madigan et al. (2018b) (Chapter 2),

$$t_{\mathrm{e-fold}} \sim \chi \frac{0.2}{\pi} \frac{M}{M_\mathrm{d}} P, \tag{3.9}$$

where we have included a scaling factor, χ, to extrapolate this result to the scattered disk configurations. From simulations, we find that is $\chi \sim 12$ for the 100 AU scattered disk and $\chi \sim 11$ for the 250 AU scattered disk. That is, the e-folding timescale increases by $\mathcal{O}(10)$ accounting for the drop in mass density as particles are spread across a broad range of semi-major axes.

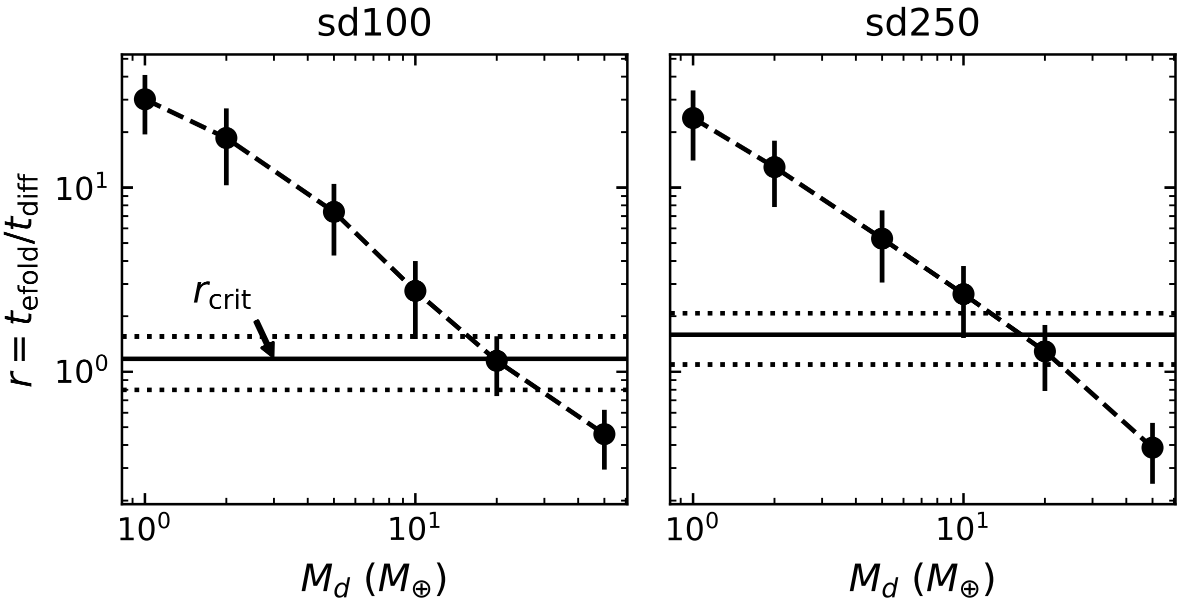

Figure 3.11: Ratio of e-folding to differential apsidal precession timescale, $r = t_{\rm e-fold}/t_{\rm diff}$, in the solar system as a function of disk mass $M_{\rm d}$ for two different versions of a scattered disk. The horizontal lines mark the critical timescale ratio and its error for the indicated orbital configuration as calculated from simulations (see Figure 3.10). Where $r \lesssim r_{\rm crit}$ the instability will occur. This corresponds to ~ 20 earth masses between 100-1000 AU.

We calculate the differential precession timescale directly from simulations of scattered disk configurations at the correct disk masses, $M_{\rm d} = 1, 2, 5, 10, 20, 50\ M_\oplus$. We include the J_2 value for the giant planets in the solar system using the appropriate semi-major axis conversion. Our simulation particles are fully interacting so the contribution to the precession rate from the disk potential is accounted for. These simulations are integrated for $\sim 6\, t_{\rm sec}$ (1000 orbits), far too short to see the instability, but more than enough time to calculate the differential precession rate. This rate is then used to calculate $t_{\rm diff}$ which in combination with the $t_{\rm e-fold}$ above gives us the timescale ratio for the solar system.

We present our results in Figure 3.11. We find that for both scattered disk configurations, $r \approx r_{\rm crit}$ at $M_{\rm d} \approx 20\, M_\oplus$. Thus, the disk mass required for the inclination instability to occur in a primordial scattered disk in the solar system under the gravitational influence of the giant planets is $\mathcal{O}(20\,M_\oplus)$. For smaller disk masses, the differential precession due to giant planets suppresses the instability. Our previous estimate for the total mass required in the outer solar system with $a > 100$ AU for the instability to occur was about an Earth mass. This estimate was based on a compact configuration, we which adopted to clearly demonstrate the discovery of a new instability in near-Keplerian disks. This new estimate is $\mathcal{O}(10)$ times greater than our previous estimate demonstrating the importance of both disk mass density and differential apsidal precession in this global instability. Sources of error in this estimate include the unknown mass and inclination distribution as a function of radius in the primordial scattered disk. We have taken reasonable best estimates and a full exploration of parameter space is beyond the scope of the paper.

The orbital evolution of a $\sim 20\,M_\oplus$ mass primordial scattered disk in the solar system is modelled by our 'sd100' simulations in the transition region. Figure 3.6 shows the expected orbital evolution of a massive primordial scattered disk. The instability saturates at $\sim 250\,t_{\rm sec}$. With $P = 1000\,{\rm yr}$, this is a saturation time of 40 Myr. Scaling using the secular timescale, a 20 Earth mass primordial scattered disk with the correct solar system value of J_2 will have a saturation timescale $\lesssim 660$ Myr. Post-instability, the intermediate to large semi-major axis population ($a \in [200, 1000]$ AU) is extremely detached from the inner solar system with perihelia of $\approx 100 - 150$ AU. The $a \in [200, 400]$ AU range actually have the largest perhelia values. Bodies with $a \lesssim 200$ AU will have inclinations of $i \approx 30^\circ$ while bodies with $a \gtrsim 200$ AU will have inclinations twice as large, about 60°. This latter population will be very difficult to detect due to their extreme detachment.

3.5 Summary and Conclusions

In this paper, we continue our exploration on the collective gravity of high eccentricity orbits in a near-Keplerian disk. We simulate the "inclination instability", a dynamical instability akin to buckling in barred disk galaxies which comes about from orbit-averaged torques between the individually low mass, but collectively massive, population. The disk orbits incline exponentially off the mid-plane, drop in eccentricity and tilt over their axes in a coherent way which leads to clustering in arguments of perihelion ω. Starting from an unrealistic (but tractable) compact configuration of orbits, we build up to simulating a massive primordial scattered disk in the outer solar system. We include the orbit-averaged gravitational influence of the giant planets using a quadrupole J_2 moment of the central body. This causes the scattered disk orbits to deferentially precess with respect to one another, weakening the strength of inter-orbit torques. We summarize our findings as follows:

(1) The e-folding timescale of the instability increases by $\mathcal{O}(10)$ when we simulate orbits in a scattered disk rather than in a compact configuration (Equation 3.9). This is due to the drop in mass density as particles are spread across an order-of-magnitude range of semi-major axes.

(2) We identify a critical J_2 moment in each simulation configuration, $J_{2,\mathrm{crit}}$, beyond which the instability is suppressed. The growth rate of the instability decreases (by a factor of a few) with added J_2 moment across a transition region $(0.1 - 1)\, J_{2,\mathrm{crit}}$, becomes zero at $J_{2,\mathrm{crit}}$, and is imaginary above $J_{2,\mathrm{crit}}$ (Figure 3.3).

(3) The median post-instability inclination/eccentricity increases/decreases with the addition of $J_2 < J_{2,\mathrm{crit}}$ (Figure 3.2).

(4) The time over which ω-clustering is maintained (about 60 secular times) is not strongly affected by the addition of a J_2 moment for a disk in the compact configuration (Figure 3.4). In a scattered disk configuration however, the addition of J2 increases the timescale over which ω-clustering is maintained (Figure 3.7). The clustering persists for orbits with semi-major axes $a \gtrsim 200$ AU until the end of the simulation, about 300 secular times.

(5) Physically, $J_2 = J_{2,\mathrm{crit}}$ in a given simulation is reached when orbits in the disk precess ≈ 1 radian apart from each other within an e-folding time (Figure 3.10). Above this value, orbits differentially precess too rapidly for long-term coherent torques to sustain the instability.

(6) The instability is a global phenomenon. If enough mass (and angular momentum) remains pinned to the mid-plane of the disk, the remainder of the disk is preventing from lifting off.

(7) The mass required for the inclination instability to occur in a primordial scattered disk between 100 $\sim$ 1000 AU in the solar system under the gravitational influence of the giant planets is $\mathcal{O}(20\,M_\oplus)$. We look at two different scattered disk configurations to explore the effect of distributing the peak of the mass density in a different location. Figure 3.11 shows they yield the same result.

(8) Unstable orbits at different semi-major axes end up with similar mean inclinations post-instability (Figures 3.5 and 3.6). Hence, those at lower semi-major axis gain a larger fractional increase in orbital angular momentum than those at higher semi-major axis. These orbits decrease their eccentricities and increase their perihelia more so than those at higher semi-major axis. This naturally generates a perihelion gap at the innermost radius of the disk that has undergone the instability. This gap appears as an under-density of orbits at perihelia of ≈ 50 AU at semi-major axes of $\approx 200 - 600$ AU. In Figure 3.8, we show the time-averaged surface density of perihelion p vs semi-major axis a (for all times post-instability). We need to time-average the simulation to get sufficient resolution to make this plot. This means we cannot attempt to precisely match the observed perihelion gap in the solar system (Trujillo & Sheppard, 2014; Bannister et al., 2018; Kavelaars et al., 2020). The parameters of the region depleted by the instability are related to the magnitude of J_2; a larger J_2 depletes a larger region of a-p space.

(9) Orbits with semi-major axes $\approx 200 - 1000$ AU will, on average, obtain extremely large perihelion distances (p $\gtrsim$ 100 AU) and inclinations ($i \sim 60°$). Figure 3.2 shows however that there is a broad range of final inclination and eccentricity values. It is possible to produce high inclination, high eccentricity eTNOs such as 2015 BP519 (Becker et al., 2018).

Now we come to the question, just how unreasonable is it to expect the primordial scattered disk to contain $\mathcal{O}(20\,M_\oplus)$? Current theories of planet formation suggest that the giant planets migrated significantly in a massive planetesimal disk. Scattering planetesimals fled to more stable regions of the solar system including into the various populations we observe today (hot classical resonant Kuiper belt, the scattered disk, the Trojans, irregular satellites, etc. see Nesvorný (2018) for a recent review). Comets in the Oort Cloud were scattered outward, until the gravitational influence of the Galaxy could torque their orbits and detach them from the inner solar system (for review see Dones et al., 2015). On their way out, they would have passed through the region of space that we are most interested in, $\approx 100 - 1000$ AU. If enough mass existed on scattered, high eccentricity orbits in this region at any given time, the orbits would collectively have gone unstable.

The Nice Model of giant planet migration supposes some $30 - 50$ earth masses of planetesimals existing from the orbit of the outermost giant planet to ~ 35 AU (Gomes et al., 2005a; Tsiganis et al., 2005; Morbidelli et al., 2005). Nesvorný & Vokrouhlický (2016) show that 1000-4000 Pluto mass bodies are needed in a primordial outer planetesimal disk to match the observed current population of Neptune's resonant bodies. Their estimate is obtained by modelling Neptune's migration through this disk. The authors find that a primordial disk of ~ 20 Earth masses is consistent with the observed resonant populations (see also Nesvorný & Morbidelli, 2012).

More recently, Shannon & Dawson (2018) estimate the mass of the primordial scattered disk using the survival of ultra-wide binaries in the Cold Classical Kuiper belt. At a 95% upper limit they find the disk could have contained 9 Earths, 40 Mars, 280 Lunas, and 2600 Plutos (at the 68% upper limit the numbers are 3, 16, 100, 1000). The combined mass of just these bodies is $\sim$23 Earth masses ($\sim$8 Earth masses at 68%). The total mass of the disk would be significantly more than this.

If at any point 20 Earth masses of material scattered onto high eccentricity orbits between $\approx 100 - 1000$ AU, it would have undergone the inclination instability. Now we turn to the question of how much of this mass would remain today. The total mass remaining in this region depends critically on the outgoing flux from this population, or how many bodies have been lost over the age of the solar system. Post-instability, the population is relatively isolated as the orbits drop in eccentricity at the same time they incline off the ecliptic. This raises their perihelia and lowers their aphelia, reducing the influence of the inner solar system planets on one side, and galactic tides and passing stars on the other. The population fossilizes at high inclinations and extraordinary values of perihelion distances. The outgoing flux should therefore depend on secular gravitational interactions between the orbits themselves rather than outside influences. These secular torques in turn depend on the distribution of the eTNOs today, particularly on whether or not they align in physical space (Sefilian & Touma, 2019) (see Chapter 4). The secular gravitational torques will cause long-term angular momentum changes transferring some bodies from the large detached population into and around the current-day scattered disk.

Hills (1981) provides an early estimate for the current amount of mass in the region spanning the orbit of Neptune to 10^4 AU. Using a variety of heuristic arguments involving the Oort cloud, Hills suggests there could be anywhere from a few to a few thousand Earth masses of material (an average yields tens of Earth masses).

Hogg et al. (1991) derives a current estimate by considering the dynamical influence of a massive ecliptic disk with $a \gg 30$ AU on the ephemerides of the giant planets and Halley's comet. They find that there could be hundreds of Earth masses of material in the disk based on the giant planets ephemerides or a few Earth masses of material based on the ephemerides of Halley's comet. Hogg's latter limit is the more constraining estimate, suggesting that the primordial scattered disk population, if it exists, must be whittled down to a few Earth masses. It's worth noting however that they assume a flat disk with the potential modelled as an outer quadrupole moment.

Gladman et al. (2009) report the discovery of a TNO on a retrograde orbit and find the object is unlikely to be primordial. They suggest a supply mechanism from a long-lived source, for example a population of large-inclination orbits beyond Neptune. This could also be a source for Halley-type comets (Levison et al., 2006). The inclination instability in a primordial scattered disk produces such a reservoir. As posited in Madigan & McCourt (2016), there may be a massive (1 - 10 $M_\oplus$) reservoir of icy bodies at large orbital inclinations beyond the Kuiper Belt. The Vera C. Rubin Observatory (Ivezić et al., 2019) will be instrumental in the discovery and orbital classification of this population should it exist.

Finally, the results we present here focus on the instability during or soon after its linear phase. However, a 20 Earth mass primordial scattered disk has a saturation timescale $\lesssim$ 660 Myr. If the outer solar system did undergo this instability, it must be in the non-linear, saturated state.

The long-term behavior of the instability is therefore crucial to understand in order to compare with current-day observations.

Chapter 4

Long-term Evolution

4.1 Apsidal Clustering following the Inclination Instability

This article has been published in the Astrophysical Journal Letters as: Zderic, A., Collier, A., Tiongco, M., & Madigan, A.-M. 2020, ApJ, 895, L27, doi:10.3847/2041-8213/ab91a0

Abstract Disks of low-mass bodies on high-eccentricity orbits in near-Keplerian potentials can be dynamically unstable to buckling out of the plane. In this letter, we present N-body simulations of the long-term behavior of such a system, finding apsidal clustering of the orbits in the disk plane. The timescale over which the clustering is maintained increases with number of particles, suggesting that lopsided configurations are stable at large N. This discovery may explain the observed apsidal (ϖ) clustering of extreme trans-Neptunian Objects in the outer solar system.

4.1.1 Introduction

Collective gravity is responsible for most large-scale structure in disk galaxies, e.g. buckling, bars, and spiral arms. This letter is a continuation of a series exploring analogous effects in the near-Keplerian potential of the Solar System. In Madigan & McCourt (2016) we presented a new dynamical instability driven by the collective gravity of low mass bodies in an axisymmetric near-Keplerian disk, and applied our results to the outer Solar System ($\sim 100 - 1000$ AU). This "inclination instability" exponentially grows the orbital inclinations of bodies while decreasing their orbital eccentricities, raising their perihelia and clustering their arguments of perihelion (ω). It appears in many ways similar to the out-of-plane buckling instability of barred disk galaxies (Raha et al., 1991; Friedli & Pfenniger, 1990). In Madigan et al. (2018b) (Chapter 2) we explained the mechanism behind the instability: long-term (secular) torques acting between high-eccentricity orbits, and showed it scaled with the number of particles in the simulation. In Fleisig et al. (2020), we moved from simulations of a single mass population to a mass spectrum. In Zderic & Madigan (2020), we showed that $\mathcal{O}(20\,M_\oplus)$ is required for the instability to occur in a primordial scattered disk between $\sim 100 - 1000$ AU in the solar system under the gravitational influence of the giant planets (Chapter 2). We also demonstrated how the instability naturally generates a gap in perihelion at a few hundred AU. The saturation timescale for the instability in a 20 Earth mass disk is $\lesssim 660$ Myr. Hence the non-linear, saturated state of the instability is important to understand. In this letter we look at the long-term behavior of the system and discover a new effect: apsidal clustering of orbits in the disk plane.

4.1.2 Long-term evolution of the Inclination Instability

We perform N-body simulations using the IAS15 adaptive time-step integrator in `REBOUND` (Rein & Liu, 2012). We use `REBOUNDx` (Tamayo et al., 2020) to add a zonal quadrupole (J_2) term to the potential of the central object to approximate the influences of the giant planets (Jupiter, Saturn, Uranus, and Neptune).

In these simulations, we return to our simplified, approximately mono-energetic setup as in Madigan & McCourt (2016); Madigan et al. (2018b) (Chapter 2). This idealized setup is chosen in place of a scattered disk configuration as higher particle density is needed to see the apsidal clustering (see section 4.1.3). The disk of orbits is initialized with a semi-major axis a distribution drawn uniformly in $[0.9, 1.1]$, eccentricity $e = 0.7$, and inclination $i = 10^{-4}\,\text{rad}$, and is initially axisymmetric (argument of perihelion, ω, longitude of ascending node, Ω, and mean anomaly, $\mathcal{M}$, drawn from a uniform distribution in $[0, 2\pi)$). The large initial eccentricities are motivated by the solar system's scattered disk. The Newtonian N-body problem is scale-free. This means we can apply our results to different semi-major axes by scaling the timescale. For example, if we choose

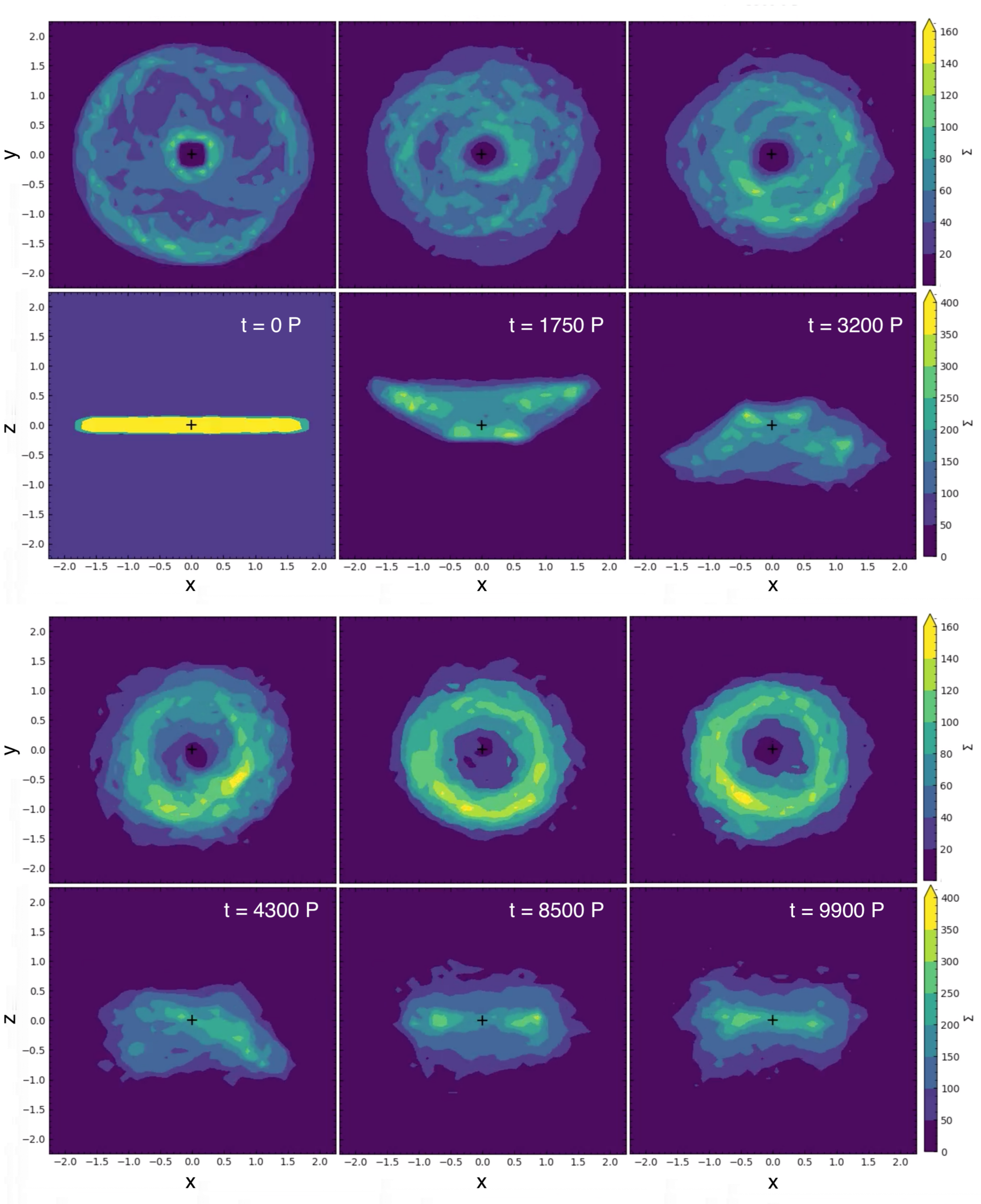

Figure 4.1: Surface density of the disk at different times showing mode development in x-y plane (top rows) and x-z plane (bottom rows). Orbital motion is CCW in the x-y plane, and individual disk orbits precess CW. Time progresses from top left to bottom right. The initially flat disk undergoes the inclination instability, buckling out above the x-y plane and dropping in orbital eccentricity ($t \sim 1750\,P$). The orbits precess back through the plane, moving the 'cone' of orbits below the x-y plane ($t \sim 3200\,P$). The $m = 1$ mode grows while the orbit cone disperses due to differential precession (from $t \sim 4300\,P$). This whole process takes $\sim 60\,t_{\rm sec}$. The evolution of mean normalized eccentricity vector $\boldsymbol{\mu}_{\hat{\boldsymbol{e}}}$ for this simulation is plotted as the stable model in Figure 4.2.

$a = 1 = 100\,AU$, one orbital period corresponds to $P = 10^3$ years. The total mass of the disk is $M_{\rm d} = 10^{-3}\,M$ and the number of disk particles, $N = 400$. The inclination instability scales with the secular timescale,

$$t_{\rm sec} \sim \frac{M}{M_{\rm d}} \frac{P}{2\pi}. \tag{4.1}$$

With this definition, $t_{\rm sec} \approx 170\,P$. As in our previous publications, we deliberately simulate an unrealistically large mass ratio, $M/M_{\rm d}$, to reduce the secular timescale and hence the simulation wall-time. To apply our results to the solar system, we rescale using the secular timescale. For example, the timescale for the instability to occur in a ~20 Earth mass disk will be ~16 times longer than in the simulations presented here.

In Figure 4.1 we show surface density evolution of the disk with face-on and edge-on lines of sight. The orbits incline out of the plane, collectively pitching over their semi-latus rectum and rolling over their major axis. Collectively, the orbits describe a cone shape. They drop in eccentricity as they incline, visibly contracting the surface area of the disk. The orbits reach peak mean inclination at $\sim 2000\ P$, and we observe the formation of a prograde-precessing $m = 1$ mode in the disk soon after. The mode starts in the inner disk as a single spiral arm (top right panel) and then moves to larger radii forming a banana-shaped over-density. This is a slow mode (Tremaine, 2001, 2005), with a pattern speed of $\sim 7 \times 10^{-4}\,\mathrm{rad\,P^{-1}}$. In the bottom left panel we see asymmetry both in and out of the disk plane.

We quantify alignment of the orbits using the mean normed eccentricity vector, $\boldsymbol{\mu}_{\hat{\boldsymbol{e}}} = \sum_{i=1}^{N} \hat{e}_i/N$, where $\hat{\boldsymbol{e}}_i$ is a unit vector pointing from orbit i's focus to pericenter. The inclination instability reveals itself as a rapid increase in $\mu_{\hat{\mathbf{e}},z}$. The apsidal clustering that follows occurs in the x-y plane, which we quantify using $\mu_{\hat{\mathbf{e}},R} = \sqrt{\mu^2_{\hat{\mathbf{e}},x} + \mu^2_{\hat{\mathbf{e}},y}}$.

The time series evolution of $\mu_{\hat{\mathbf{e}},z}$ and $\mu_{\hat{\mathbf{e}},R}$ for two simulations is shown in Figure 4.2. The first simulation is unstable to the inclination instability. This creates an out-of-plane asymmetry quantified by $\mu_{\hat{\mathbf{e}},z}$ which then appears to seed an in-plane asymmetry quantified by $\mu_{\hat{\mathbf{e}},R}$ as the orbits precess back through the plane. This in-plane over-density attracts more orbits, increasing the

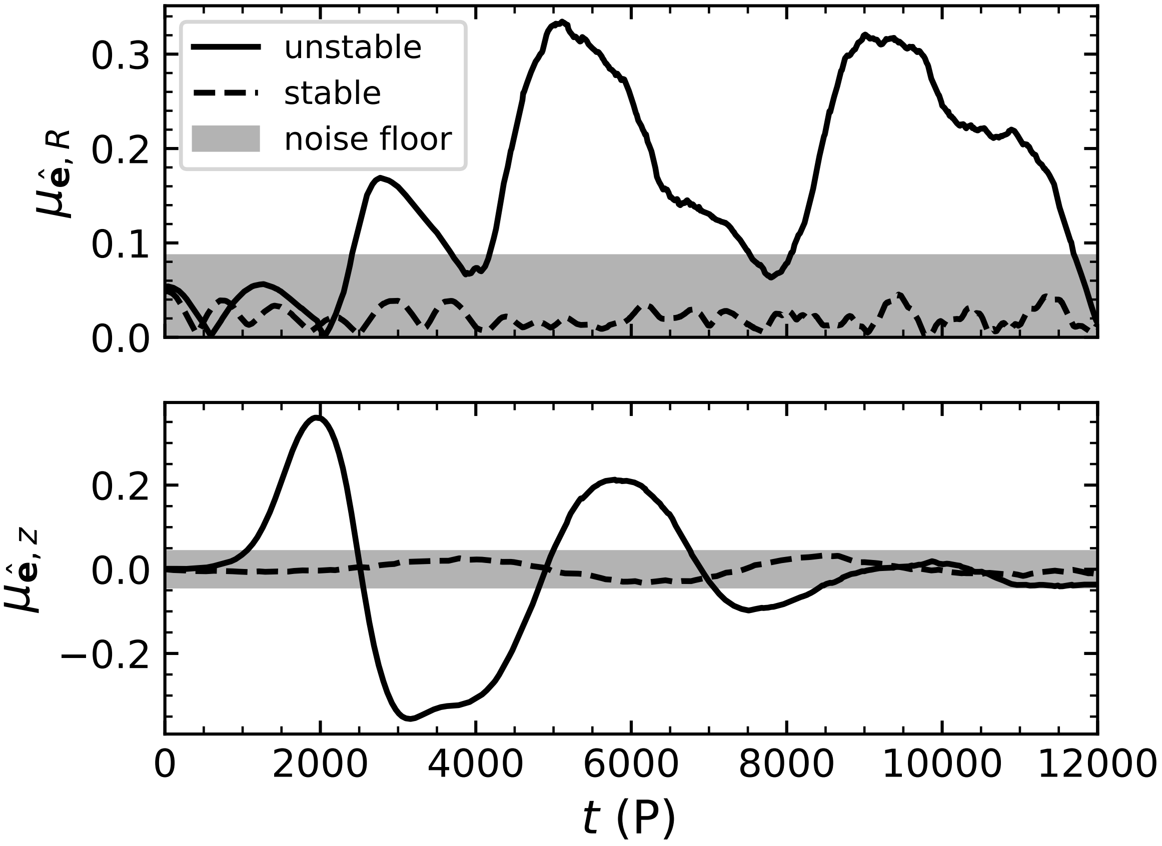

Figure 4.2: Length of the projection of the normalized mean eccentricity vector in the x-y plane, $\mu_{\hat{\mathbf{e}},R}$, and z-component of the normalized mean eccentricity vector, $\mu_{\hat{\mathbf{e}},z}$, as function of time for two simulations. The grey band shows the noise floor. Both simulations have identical initial orbital distributions, but one has a strong J_2 moment added to the central body that suppresses the instability (stable) and the other has no added J_2 (unstable).

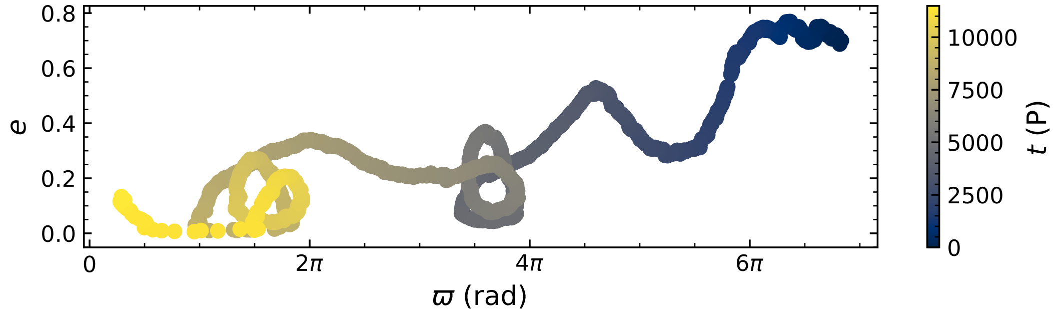

Figure 4.3: Evolution of the longitude of perihelion ϖ, and orbital eccentricity, e, as a function of time demonstrating the transient trapping of an orbit in an m=1 mode. The orbit precesses with retrograde motion (moving right to left) over 10^4 orbital periods. At $\sim 5000\ P$ and $\sim 9000\ P$ the orbit becomes temporarily trapped in the m=1 mode. It librates within the mode, being secularly torqued by the mode to higher and lower eccentricity.

strength of the perturbation. The mode disperses, recurring some time later. Differential precession causes the out-of-plane orbital clustering to disappear after $\sim 10^4\,P$. We do not anticipate this clustering to reappear beyond this time as the conditions which drove the instability in the first place (low orbital inclinations, high orbital eccentricities) will no longer be met. Artificially strong two-body scattering causes the in-plane clustering to disappear after this. We expect in-plane clustering to be sustained in simulations with larger particle numbers (see Section 4.1.3.2 and Figure 4.4).

In the second simulation, the disk is made stable against the inclination instability by the addition of a zonal quadrupole ($J_2 = 3 \times 10^{-5}$) moment of the central body, just strong enough to suppress the inclination instability with differential apsidal precession of disk orbits (see Zderic & Madigan, 2020, Chapter 3). The eccentricity vector components remain below the noise floor for the entire length of the simulation.[1] Smaller J_2 moments (i.e. low enough for the instability to still occur) actually increase the longevity of the apsidal clustering.

In Figure 4.3, we show the time evolution of the longitude of perihelion, ϖ, and orbital eccentricity, e, of a single particle in the unstable simulation. The particle's orbit precesses with

---

[1] The noise floor for $\mu_{\hat{e},R}$ is calculated as the 95th percentile of 100 iterations of the initial, axisymmetric $\mu_{\hat{e},R}$. The noise floor for $\mu_{\hat{e},z}$ is calculated by bootstrapping the 2σ error of $\mu_{\hat{e},z}$ at late times ($t \sim 10^4$ P).

retrograde motion (right to left) and decreases in eccentricity during the instability. At $\sim 5000\ P$, the orbit becomes transiently trapped in the $m = 1$ mode, and librates in ϖ-e space. After a few cycles in the mode, the particle escapes, circulates retrograde for a single cycle, and becomes transiently trapped in the mode again at $\sim 9000\ P$. Secular gravitational torques exerted on the orbit by the mode are responsible for the ϖ-e oscillations. This same mechanism stabilizes eccentric nuclear disks of stars around supermassive black holes (Madigan et al., 2018a).

4.1.3 Discussion

4.1.3.1 Why does an $m = 1$ mode develop in the plane of the disk?

Lopsided modes in near-Keplerian disks can develop spontaneously if the disk contains a large fraction of retrograde orbits (e.g., Touma 2002; Touma et al. 2009; Kazandjian & Touma 2013). In the simulations presented here, the inclination instability produces retrograde orbits, but only a few ($\lesssim 1\%$) and only after apsidal alignment has already occurred. Thus, retrograde orbits are not responsible for the clustering observed here, distinguishing these results from previous work.

A recent series of papers (Touma et al., 2019; Tremaine, 2020a,b) show that spherical near-Keplerian potentials tend toward an ordered, lopsided state when the system is cooled below a critical temperature (directly related to RMS eccentricity of the orbits). The ordered lopsided state results from a phase transition, rather than dynamical instability, driven by resonant relaxation (Rauch & Tremaine, 1996).

Our simulations show the development of a spontaneous lopsided mode in a three-dimensional near-Keplerian potential. This is done without seeding asymmetry (beyond that from numerical noise), forcing mode development with gas dynamics, or superimposing retrograde orbits. The appearance of the mode appears to be contingent on the inclination instability altering the initial orbital configuration, requiring lower eccentricities, higher inclinations and clustering in arguments of perihelion.[2] . Indeed, in the 'stable' simulation of Figure 4.2 in which the instability is suppressed,

---

[2] Low eccentricity ($e < 0.5$) disks that don't undergo the instability don't show apsidal clustering, nor do post-instability disks in which we deliberately randomize the arguments of perihelion

there is no apsidal clustering. We hypothesize that the inclination instability allows the system to undergo a phase transition to a lopsided state similar to the transition reported in Touma et al. (2019) for spherically symmetric systems. This hypothesis will be explored in future work.

4.1.3.2 Lifetime of the mode

While the inclination instability appears in our compact simulations with as few as $N = 100$ particles, more particles are needed to resolve and stabilize the $m = 1$ mode. This is reminiscent of bar development in galactic disks. The bar instability in an N-body disk occurs on nearly identical timescales, initial mode density, and growth rate for similar disks of increasing N. However, if N is too low, the bar will dissolve soon after formation (Dubinski et al., 2009). Similarly, N-body galactic disks are known to easily form recurrent short-lived, transient spirals (James & Sellwood, 1978; Sellwood, 2012, 2020). When particle number is increased in these simulations, the over-density reaches some minimum value and results in exponential growth of the mode (e.g., Toomre & Kalnajs (1991); Weinberg (1998); Sellwood (2012)). In disk galaxies this presents as long-lasting $m = 2$ spirals or an $m = 2$ bar mode. In our system, the low number of particles results in a coarse, under-populated mode. Artificially strong two-body interactions perturb the osculating orbits of bodies that stream through the mode and weaken the secular torques that would trap them. Therefore, while our simulations show robust results, increasing particle number is well motivated.

In Figure 4.4, we show that the strength and duration of the mode increases with N, the number of particles in the simulation. We do so by calculating A_{e_R}, the integral of $\mu_{\hat{e},R}$ above the noise floor over a time period of $10{,}000\,P$. These initial tests indicate that the disk produces longer-lasting and stronger apsidal alignment with increasing N.

At $N = 400$, the mode is already quite long-lived. Rescaling our results for a disk mass of $M_\mathrm{d} \sim 20\,M_\oplus$, using $a = 1 = 100\,\mathrm{AU}$, $P = 1000\,\mathrm{yr}$, the mode lasts ~ 160 Myr. For realistic numbers of particles, many orders of magnitude larger than what we use here, it may be reasonable to expect that the mode would be stable for the age of the solar system.

4.1.3.3 Application to the Solar System

Observations reveal many unusual orbital features in the population of extreme trans-Neptunian Objects (eTNOs). This includes detached orbits (perihelia well beyond the orbit of Neptune), high inclinations and even retrograde orbits, clustering in arguments of perihelion ω, and clustering in longitudes of perihelion ϖ. In particular, the clustering in ϖ has been an important motivator of the Planet 9 hypothesis (Batygin & Brown, 2016a; Batygin et al., 2019). For recent reviews, including in-depth discussions of observational biases, see Trujillo (2020); Kavelaars et al. (2020).

Sefilian & Touma (2019) show that test particles interacting with the potential of a thick, apsidally-aligned eccentric disk in the outer Solar System will cluster in ϖ and reproduce other key orbital features of the eTNO population. However, they don't directly simulate this eccentric disk or its formation[3] . Here we have shown that the late-time evolution of the inclination instability can produce such a structure from an axisymmetric disk, and that this structure is likely stable at large N. The drop in eccentricity during the instability isolates the system, both from the influence of giant planets at the inner edge and external perturbations (galactic tides, passing stars) at the outer edge. The isolated nature of the system, and the increasing stability of the mode with N, suggests that the structure should be stable over a long timescale. The observations of Sednoids ($a \gtrsim 150 AU$, $p \gtrsim 50 AU$; Brown et al., 2004b; Trujillo & Sheppard, 2014) support this picture.

4.2 Conclusion

In a recent series of papers we have shown that the inclination instability in high-eccentricity, near-Keplerian disks results in high orbital inclinations, raised perihelia and ω-clustering. Here we show that the system's long-term behavior results in ϖ-clustering. The strength and duration of the apsidal clustering increases with increasing N. We find that both ω-clustering and ϖ-clustering can occur at the same time. In the context of the solar system, the collective gravity of eTNOs could explain the observed ω-clustering, ϖ-clustering, detached objects, and even a perihelion gap

---

[3] On this issue, Sefilian & Touma (2019) point to a paper by Kazandjian et al. in prep.

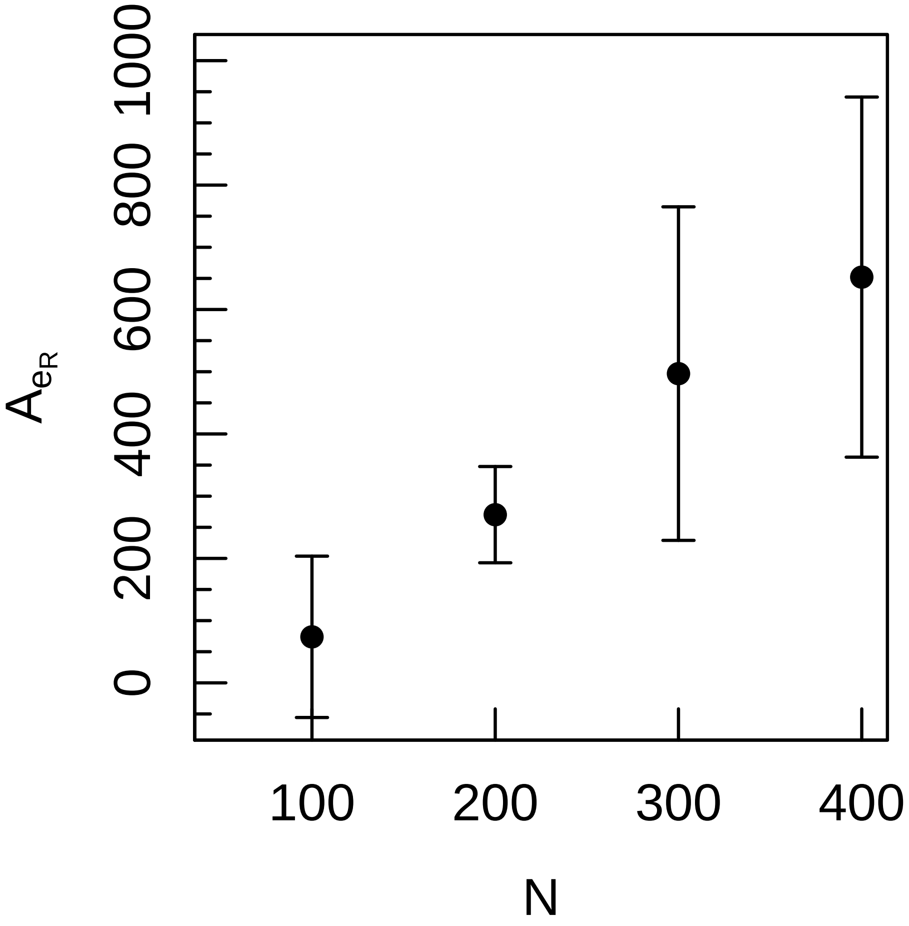

Figure 4.4: A_{e_R}, the integral of $\mu_{\hat{\mathbf{e}},R}$ above the noise floor, as a function of particle number, N, for a simulation length of 10,000 P . The points and error bars show the median and standard deviation of 6 simulations in each group.

(see Zderic & Madigan, 2020, Chapter 3).

In this letter, we present results from simulations with highly idealized initial conditions for both the simplicity of analysis and tractable computational expense. In future, we plan to simulate the long-term evolution of a high-mass primordial scattered disk including the presence of the giant planets at high-N. We are working on the modification of existing codes which will allow us to advance in this direction.

4.3 A Lopsided Outer Solar System?

This article has been accepted (Oct. 2021) by the Astronomical Journal as: Zderic, A., Tiongco, M., Collier, A., Wernke, H., Generozov, A., & Madigan, A.-M., AJ (accepted)

Abstract Axisymmetric disks of eccentric orbits in near-Keplerian potentials are unstable and undergo exponential growth in inclination. Recently, Zderic et al. (2020) showed that an idealized disk then saturates to a lopsided mode (Section 4.1). Here we show, using N-body simulations, that this apsidal clustering also occurs in a primordial scattered disk in the outer solar system which includes the orbit-averaged gravitational influence of the giant planets. We explain the dynamics using Lynden-Bell (1979)'s mechanism for bar formation in galaxies. We also show surface density and line of sight velocity plots at different times during the instability, highlighting the formation of concentric circles and spiral arms in velocity space.

4.3.1 Introduction

Something odd is going on in the outer solar system: distant bodies in orbit beyond Neptune appear clustered in argument of perihelion (ω; Trujillo & Sheppard, 2014) and longitude of perihelion (ϖ; Batygin & Brown, 2016a). Some have extreme inclinations that cannot be generated in the standard model of solar system evolution (Gladman et al., 2009; Chen et al., 2016; Becker et al., 2018; Kaib et al., 2019), and others are "detached" in the sense that they have perhelia that lie far beyond the gravitational reach of the giant planets (e.g., Brown et al., 2004a). Observational biases have been carefully demonstrated in outer solar system surveys (Shankman et al., 2017; Lawler et al., 2017; Kavelaars et al., 2020; Napier et al., 2021) but whether they can fully explain the anomalous orbital structure of Trans-Neptunian objects (TNOs) remains a contentious issue (Brown, 2017; Brown & Batygin, 2019). If they do not, the outer solar system requires a new source of gravitational perturbation. One such source could be a planet far beyond the orbit of Neptune (for reviews see Batygin et al. (2019) and Trujillo (2020)). We propose a different, internal source: the self-gravity of the bodies themselves.

The collective gravity of bodies on eccentric orbits in an axisymmetric near-Keplerian disk drives a dynamical instability (Madigan & McCourt, 2016; Madigan et al., 2018b, Chapter 2).

This “inclination instability” exponentially grows the inclinations of orbits while decreasing their eccentricities, raising their perihelia and clustering their arguments of perihelion (ω). In a recent paper, Zderic & Madigan (2020), we showed that $\mathcal{O}(20)$ Earth masses are required for the instability to occur in a primordial scattered disk between $\sim 10^2 - 10^3$ AU in the solar system under the orbit-averaged, gravitational influence of the giant planets at their current locations (Chapter 3). The instability can also generate a gap in perihelion at $\sim 50 - 75$ AU, as observed in the outer solar system (Kavelaars et al., 2020; Oldroyd & Trujillo, 2021). The saturation timescale, that is, the time at which inclinations cease exponential growth, for the instability in a 20 Earth mass disk is far less than the age of the solar system. Therefore, to connect to the present-day outer solar system we need to understand the non-linear, saturated state of the instability. We are further motivated by the results of Zderic et al. (2020) where we discovered late-time apsidal clustering of orbits in the disk plane, albeit in simulations with highly idealized initial conditions (Section 4.1). Here we show that the same late-time clustering occurs in a primordial scattered disk between $\sim 10^2 - 10^3$ AU in the solar system under the gravitational influence of the giant planets. We essentially take the more realistic simulation conditions of Zderic & Madigan (2020) (Chapter 3) and extend them past saturation to look for in-plane clustering. We show that the apsidal clustering can be explained using Lynden-Bell (1979)'s mechanism for bar formation in disk galaxies.

Our paper proceeds as follows: in §4.3.2 we describe the Lynden-Bell mechanism for bar formation and show how it may be applied to near-Keplerian systems. In §4.3.3 we describe our numerical methods and in §4.3.4 present our results. In §4.3.5 we show surface density and line of sight velocity plots of our simulations at different times, and we conclude in §4.3.6.

4.3.2 The Lynden-Bell Mechanism in near-Keplerian Systems

In 1979, Donald Lynden-Bell described a mechanism by which bars may be formed in the centers of galaxies. We reproduce the basic argument here.

In a general galactic potential, a typical orbit is a rosette with an angle between π and 2π linking consecutive apocenters. If we view an orbit from rotating axes, we may choose the rotation speed ν_i such that the angle between apocenters will be π. The orbit will then be bisymmetric, like a centered oval or ellipse. If ν is the mean angular speed of a star about the galaxy and κ is its radial angular frequency, then we should choose $\nu_i = \nu - \kappa/2 > 0$. For near-circular orbits, ν_i will not vary much over a large region of a galaxy (Binney, 2008).

We now introduce a weak, bar-like potential, rotating with pattern speed $\nu_p \approx \nu_i$, and consider its interaction with an orbit. In the frame co-rotating with ν_p, the star's orbit is an almost closed oval which rotates at a slow rate $\nu_i - \nu_p << \nu$. There is no time for the weak perturbing potential to affect the star's fast motion around the oval, so the orbit has an adiabatic invariant, $1/2\pi \oint \boldsymbol{p} \cdot d\boldsymbol{q} = 2J_f \sim$ const, where $\boldsymbol{q}$ is a vector of polar coordinates (R, ϕ), $\boldsymbol{p}$ is the polar conjugate momentum, and the integral is taken over one closed, bi-symmetric orbit. However, the potential will exert a persistent weak torque on the oval as a whole because they move slowly with respect to one another. Hence the oval will change to another oval with the same J_f but different angular momentum j.

If the orbit is ahead of the bar in its rotation, its angular momentum will decrease due to the gravitational torque from the bar. Normally, ν_i will increase in response such that the orbit is repelled by the bar. In other words, the bar repels the orbit because $\partial\nu_i/\partial j|_{J_f}$ (the Lynden-Bell derivative; Polyachenko 2004) is negative. In the abnormal case in which the Lynden-Bell derivative is positive however, the orbit will oscillate about the bar-like potential. In such cases, the orbit adds to the strength of the potential which will then be able to capture more and more orbits.

To discover what regions of a galaxy lead to the barring of near-resonant orbits, Lynden-Bell calculated $\nu_i(J_f, j)$ for an isochrone galactic potential which permits analytic expressions for the angular frequencies κ and ν. He showed that an abnormal region is associated with central regions in this model where circular velocity rises with radius. Polyachenko & Shukhman (2020) recently expanded upon Lynden-Bell's work by mapping the equilibria of orbits as a function of ν_i, the Lynden-Bell derivative, and the orbit's responsiveness to the bar potential.

We now extend this argument to a near-Keplerian system, where the gravitational potential is dominated by a central mass. A typical orbit is an almost-closed ellipse. As in Lynden-Bell (1979) we focus on the idealized planar problem, though we note that our simulations in the next sections are three-dimensional. The orientation of the ellipse in the orbital plane is given by the longitude of pericenter, ϖ, and its rate of change $\dot{\varpi} = d\varpi/dt = \nu - \kappa$ indicates its precession rate. If we view the orbit from rotating axes, we may choose the rotation speed such that the angle between apocenters is zero, $\nu_i = \dot{\varpi} = 0$. The orbit will then be a closed ellipse with the central body occupying one focus.

Following Lynden-Bell's argument, we now introduce a weak, *lopsided* potential rotating with pattern speed $\nu_p \approx \dot{\varpi}$ and consider its gravitational influence on an orbit. Here the precession rate of the orbit is by definition much less than the orbital period even in an inertial frame. In this case, $\dot{\varpi} \ll \nu$ and $\nu_p \ll \nu$, thus $\dot{\varpi} - \nu_p \ll \nu$. Therefore, the secular average over mean anomaly is equivalent to Lynden-Bell's average over fast orbital motion, and $J_f \to I$ where $I = \sqrt{GMa}$, M is the central mass and a is the semi-major axis (see Merritt 2013; Fouvry et al. 2020).

The lopsided potential exerts a persistent torque on the orbit, changing the orbit's angular momentum at fixed semi-major axis. The specific angular momentum of a Kepler orbit is given by $j = \sqrt{GMa(1-e^2)}$, where e is the magnitude of the orbital eccentricity. At fixed semi-major axis, angular momentum is a monotonically decreasing function of eccentricity. In Kepler elements, the Lynden-Bell derivative $\left({}^{\partial\nu_i}\!/_{\partial j}|_{J_f}\right)$ is $\propto -{}^{\partial\dot{\varpi}}\!/_{\partial e}|_a$.

Lynden-Bell's 'abnormal region' specifically refers to prograde precession with magnitude increasing with increasing angular momentum. In Kepler elements, this corresponds to a region where precession is prograde with magnitude decreasing with increasing eccentricity. The interpretation of 'normal' and 'abnormal' regions changes with context, e.g. the abnormal region described above is actually typical in lopsided eccentric disks (Madigan et al., 2018a). Therefore, we will refer to regions where apsidal clustering is supported as *clustering regions*, and regions where clustering is not supported as *anti-clustering regions.*

We note that it is also possible to have a clustering region with retrograde precessing orbits: if precession is retrograde and the magnitude of the precession rate increases with increasing eccentricity then orbits will be attracted to a perturbing potential. Orbits can be trapped in modes

Model Id	J_2	N
N400	No	400
N800	No	800
J2N400	Yes	400
J2N800	Yes	800

Table 4.1: Model names and parameters.

in near-Keplerian systems provided that $\partial\dot{\varpi}/\partial e|_a < 0$ regardless of the sign of ϖ; we define the clustering region to be any region in the disk where $\partial\dot{\varpi}/\partial e|_a < 0$.

4.3.3 N-Body Simulations

Our N-body simulations use the open-source framework REBOUND with the IAS15 adaptive timestep integrator (Rein & Liu, 2012; Rein & Spiegel, 2015)[4]. Additionally, we use REBOUNDx (Tamayo et al., 2020) to add a zonal harmonic, J_2, to the central body to emulate the orbit-averaged effects of the giant planets. All particles in our simulations are massive and fully-interacting. In this paper, the Kepler elements semi-major axis (a), eccentricity (e), inclination (i), argument of pericenter (ω), longitude of the ascending node (Ω), and mean anomaly ($\mathcal{M}$) are used to describe the orbits.

The total disk mass used in the simulations is $M_d = 10^{-3}\,M_\odot$ and the number of particles is 400 or 800 (see Table 1). This unrealistically large disk mass is chosen to accelerate secular dynamics (see Equation 4.2) within the disk reducing the number of orbits we need to simulate. In addition, the low N is required to reduce the simulation walltime per orbit. The orbital distribution of our disks are initialized to approximately model a primordial scattered disk in the outer solar system (Duncan et al., 1987). The model is axisymmetric with an order of magnitude spread in semi-major axis, a_0, the values of which are drawn from a 1D log-uniform distribution between $[10^2, 10^3]$ AU (this is equivalent to a surface density distribution of a^{-2})[5]. All bodies have the same

---

[4] The fixed timestep WHFast integrator, while faster than IAS15, doesn't conserve energy and angular momentum well in this high eccentricity problem (see also Rauch & Holman, 1999). The performance of the MERCURIUS integrator is similar to IAS15's due to frequent close encounters between particles.

[5] We have simulated other 1D semi-major axis distributions, for example $a^{-0.7}$ (Napier et al., 2021) and $a^{-2.5}$ Duncan et al. (1987). The instability proceeds similarly but its timescale decreases with increasing distribution

initial pericenter distance, $p_0 = 30$ AU. Inclination i_0 is drawn from a Rayleigh distribution with a mean of 5°, and ω, Ω, and $\mathcal{M}$ are chosen uniformly from 0 to 2π radians[6] .

We add a J_2 potential to the central body in half of our simulations (see Table 1), and pick the J_2 moment to lie in the "transition region" where the J_2 potential alters the inclination instability without suppressing it (Zderic & Madigan, 2020, Chapter 3). Our chosen disk mass and number of particles, choices forced by numerical limitations, are unrealistic. The instability timescale and the max J_2 that the disk can resist (that is, still undergo the instability) both depend on these key parameters. The low N in our simulations leads to artificially strong self-stirring that weakens the secular torques that cause the inclination instability and increases differential precession by excessively spreading out the disk (Madigan et al., 2018b, Chapter 2). For the same total mass, a disk with more particles will be able to resist a larger J_2 (that is, still undergo the instability). We determined how the inclination instability timescale scales with M_d and N in Madigan et al. (2018b) (Chapter 2). Then in Zderic & Madigan (2020) (Chapter 3), we used that timescale scaling along with simulations of these disks with added J_2 to find that a $\sim 20\,M_\oplus$ primordial scattered disk could resist the J_2 of the giant planets. We found that this realistic system would be in the transition region. The J2N400 and J2N800 models in this paper are in the transition region too. Therefore, these models, which have unrealistic J_2, M_d, and N, are dynamically similar to a 20 Earth mass primordial scattered disk, at least with regards to J_2.

For the sake of reproducibility, the J_2R^2 used in these simulations is $0.3\,\mathrm{AU}^2$. We use the same J_2R^2 value for the J2N400 and J2N800 even though these simulations have different critical J_2 because this J_2R^2 is sufficient to put both models in the transition region. The J_2R^2 for the solar system is 0.06 AU^2 and using solely a secular scaling ($10^{-3}/20M_\oplus \approx 16$) J_2R^2 would be 0.96 AU^2 for a $10^{-3}M_\odot$ mass disk. This J_2R^2 is about 3 times larger than the actual J_2R^2 used in our simulations. A $N \to \infty$ disk can resist about 3 times more J_2R^2 than a $N = 400$ disk.

---

steepness.

[6] Disks with larger initial inclinations also undergo the instability provided that mean i_0 is less than $\sim 20°$.

Simulation times are given in units of the secular timescale:

$$t_{\rm sec} = \frac{1}{2\pi}\frac{M_\odot}{M_{\rm d}}P \tag{4.2}$$

where P is the orbital period at the innermost part of the disk. For $M_d = 10^{-3}\,M_\oplus$, $t_{\rm sec} \approx 160\,P \approx 0.16\,{\rm Myr}$, where $P(a = 100\,{\rm AU}) = 10^3\,{\rm yr}$. We give timescales for a more realistic 20 Earth mass primordial scattered disk with in Section 4.3.6.

4.3.4 Results

We measure apsidal clustering using the mean, normed eccentricity vector,

$$\boldsymbol{\mu}_{\hat{\mathbf{e}}} = \sum_{i=1}^{N} \frac{\hat{\boldsymbol{e}}_i}{N}, \tag{4.3}$$

where

$$\boldsymbol{e}_i = \frac{(\boldsymbol{v}_i \times \boldsymbol{j}_i)}{GM_\odot} - \hat{\boldsymbol{r}}_i \tag{4.4}$$

is the eccentricity vector of the ith orbit, and $\boldsymbol{r}_i$, $\boldsymbol{v}_i$, and $\boldsymbol{j}_i$ are the position, velocity, and specific angular momentum of the ith particle. The eccentricity vector points from the apocenter to the pericenter of the orbit. We use the cylindrical coordinates of $\boldsymbol{\mu}_{\hat{\mathbf{e}}}$ to look for apsidal clustering,

$$e_R = \sqrt{\mu_{\hat{\mathbf{e}},x}^2 + \mu_{\hat{\mathbf{e}},y}^2}, \tag{4.5a}$$

$$e_\theta = \arctan\left[\frac{\mu_{\hat{\mathbf{e}},y}}{\mu_{\hat{\mathbf{e}},x}}\right], \tag{4.5b}$$

$$e_z = \mu_{\hat{\mathbf{e}},z}. \tag{4.5c}$$

The radial component, e_R, quantifies in-plane apsidal clustering, the azimuthal component, e_θ, is used to calculate the pattern speed and direction of in-plane apsidal clustering, and the z component, e_z, quantifies out-of-plane apsidal clustering. See Appendix D for a comparison to standard measures of apsidal clustering.

We calculate the noise floor for e_R and e_z by creating one thousand $N = 400$ or $N = 800$ axisymmetric disks. For each disk, we draw argument of perihelia and longitude of the ascending node from a uniform distribution and inclination from a Rayleigh distribution with mean inclination

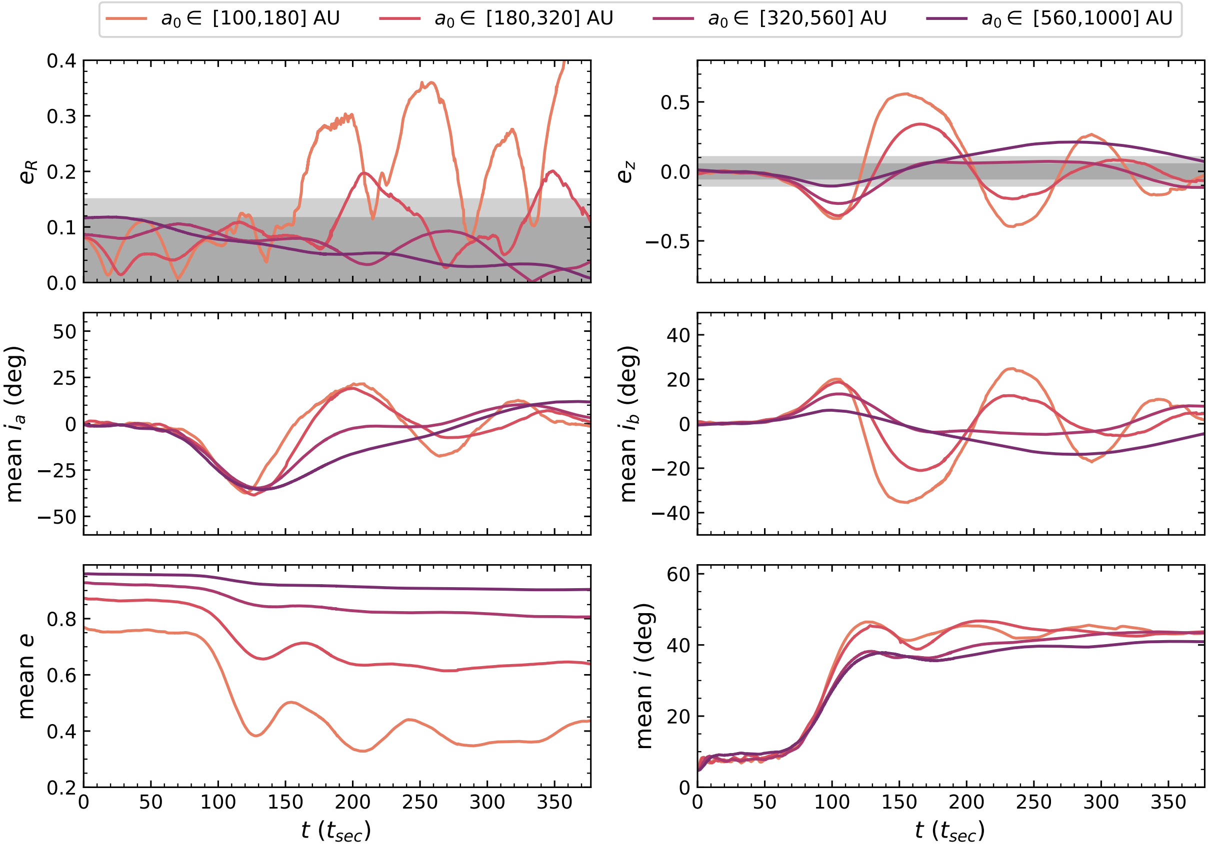

Figure 4.5: Model N400: Apsidal clustering occurring in the inner edge of a scattered disk model without added J_2 after the inclination instability has saturated. The plotted quantities are binned by initial semi-major axis. Two noise floors are shown for e_R and e_z (see Section 4.3.4 for noise floor definition). The inclination instability is shown by exponential growth in i, i_a, and i_b and a corresponding decrease in e to conserve total angular momentum of the disk. The instability saturates at $t \sim 125\,t_{\rm sec}$. About $25\,t_{\rm sec}$ later, e_R for the inner bin ($a_0 \in [100, 180]$ AU) increases above the noise floor indicating in-plane apsidal clustering. About $50\,t_{\rm sec}$ later, slight in-plane apsidal clustering appears in the next bin ($a_0 \in [180, 320]$ AU).

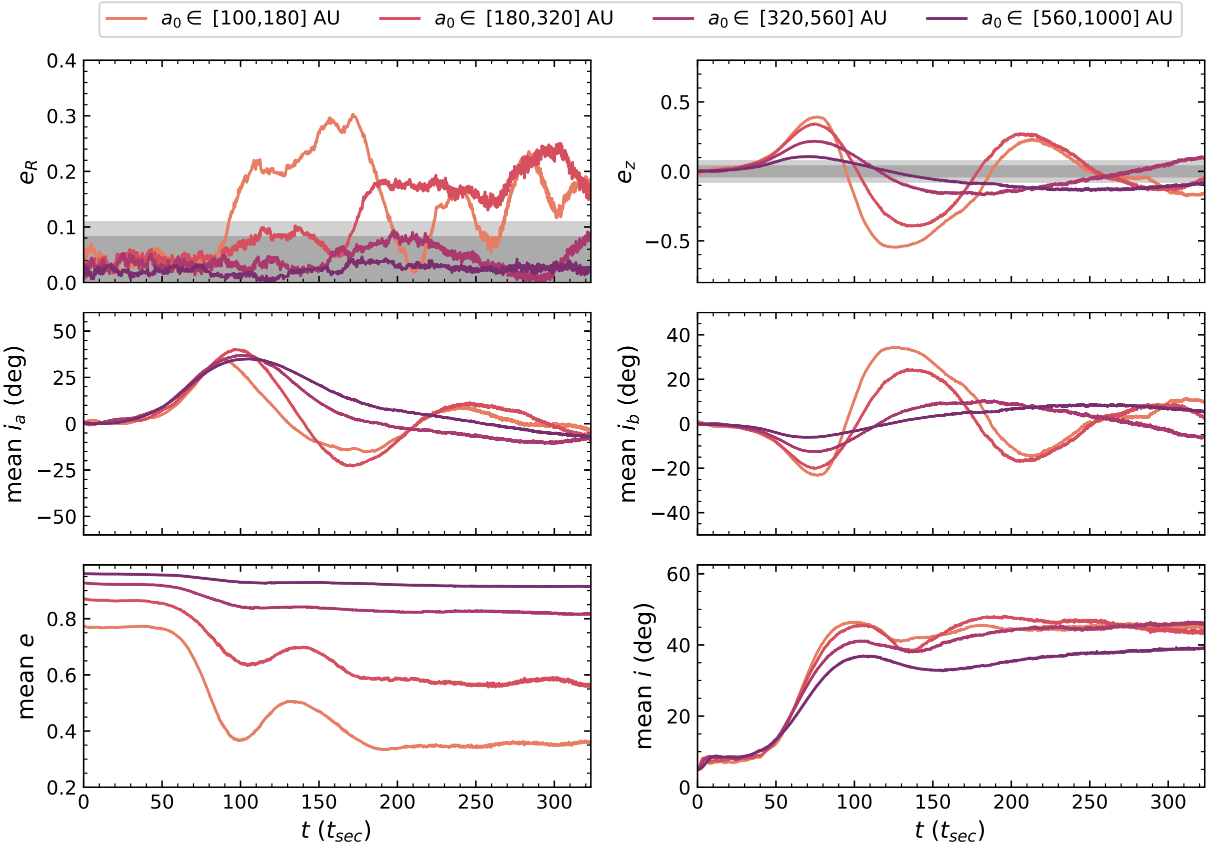

Figure 4.6: Model N800: Apsidal clustering occurring in the inner edge of an 800 particle scattered disk model without added J_2 after the inclination instability has saturated. Same panels as in Figure 4.5. The plotted quantities are binned by initial semi-major axis, and e_R and e_z noise floors are shown. Compared to N400, the instability saturates at an earlier time, $t \sim 100\, t_{\rm sec}$, and, in-plane apsidal clustering (shown by e_R) in the inner a_0 bin begins immediately after the instability saturates. Like N400, in-plane apsidal clustering propagates into the next a_0 bin ($a_0 \in [180, 320]$ AU) about $75\, t_{\rm sec}$ later. Apsidal clustering in the $a_0 \in [180, 320]$ AU bin is stronger and more consistent in N800 than in N400.

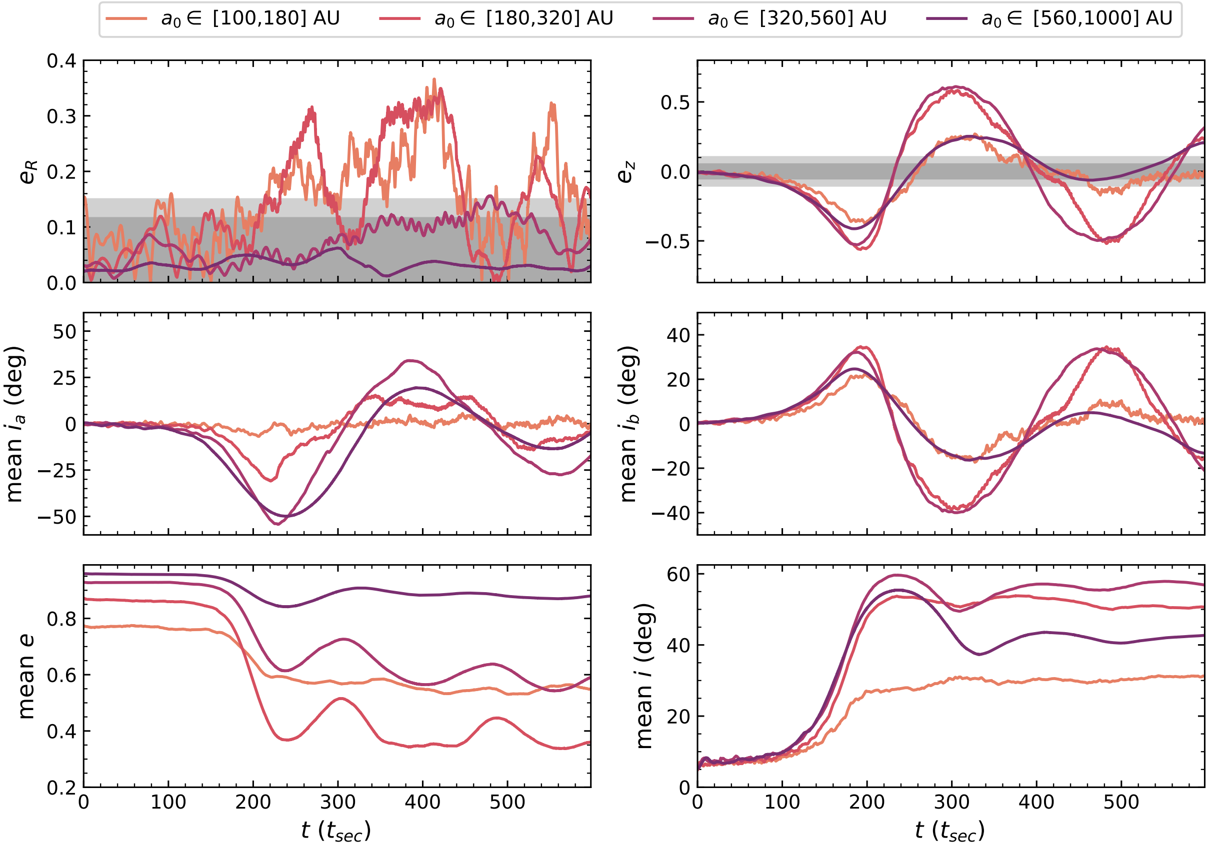

Figure 4.7: Model J2N400: Apsidal clustering occurring in the inner edge of a scattered disk model with added J_2 after the inclination instability has saturated. Same panels as in Figure 4.5. Compared to N400 and N800, the inclination instability is delayed by the added J_2, saturating at $\sim 200\, t_{\text{sec}}$, and the innermost part of the disk barely undergoes the instability. Like in N400 and N800, a lopsided mode develops on the inner edge of the disk shortly after the inclination instability has saturated. Unlike N400 and N800, apsidal clustering appears in both inner disk bins $a_0 < 320$ AU simultaneously. In addition, apsidal clustering in the $a_0 \in [180, 320]$ bin is as strong as in the $a_0 \in [100, 180]$ bin.

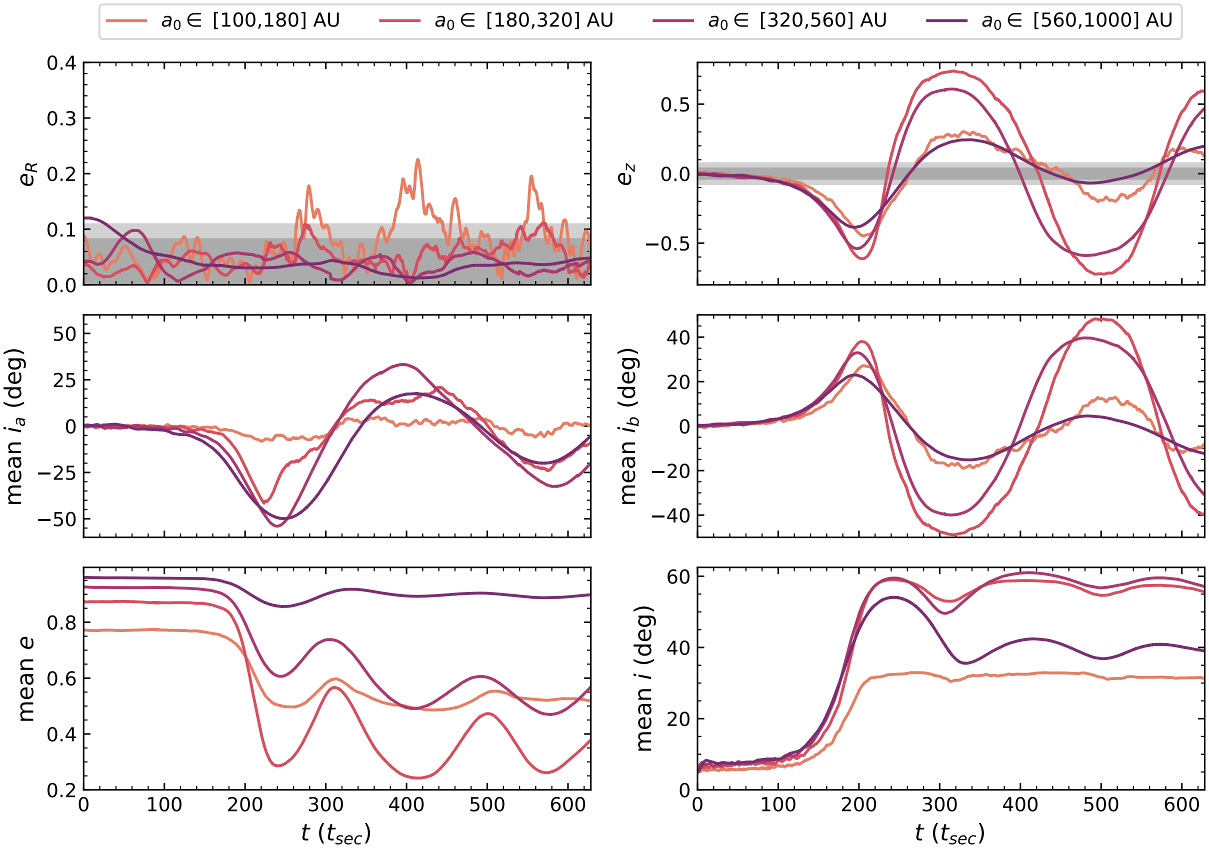

Figure 4.8: Model J2N800: Apsidal clustering occurring in the inner edge of an 800 particle scattered disk model with added J_2 after the inclination instability has saturated. Same panels as in Figure 4.5. Compared to J2N400, apsidal clustering is weakened in this simulation. Statistically significant clustering only occurs in the $a_0 \in [100, 180]$ AU bin and this clustering is weaker than in the J2N400. This is different than the simulations without added J_2 where we saw similar to slightly more in-plane apsidal clustering as we increased the in particle number.

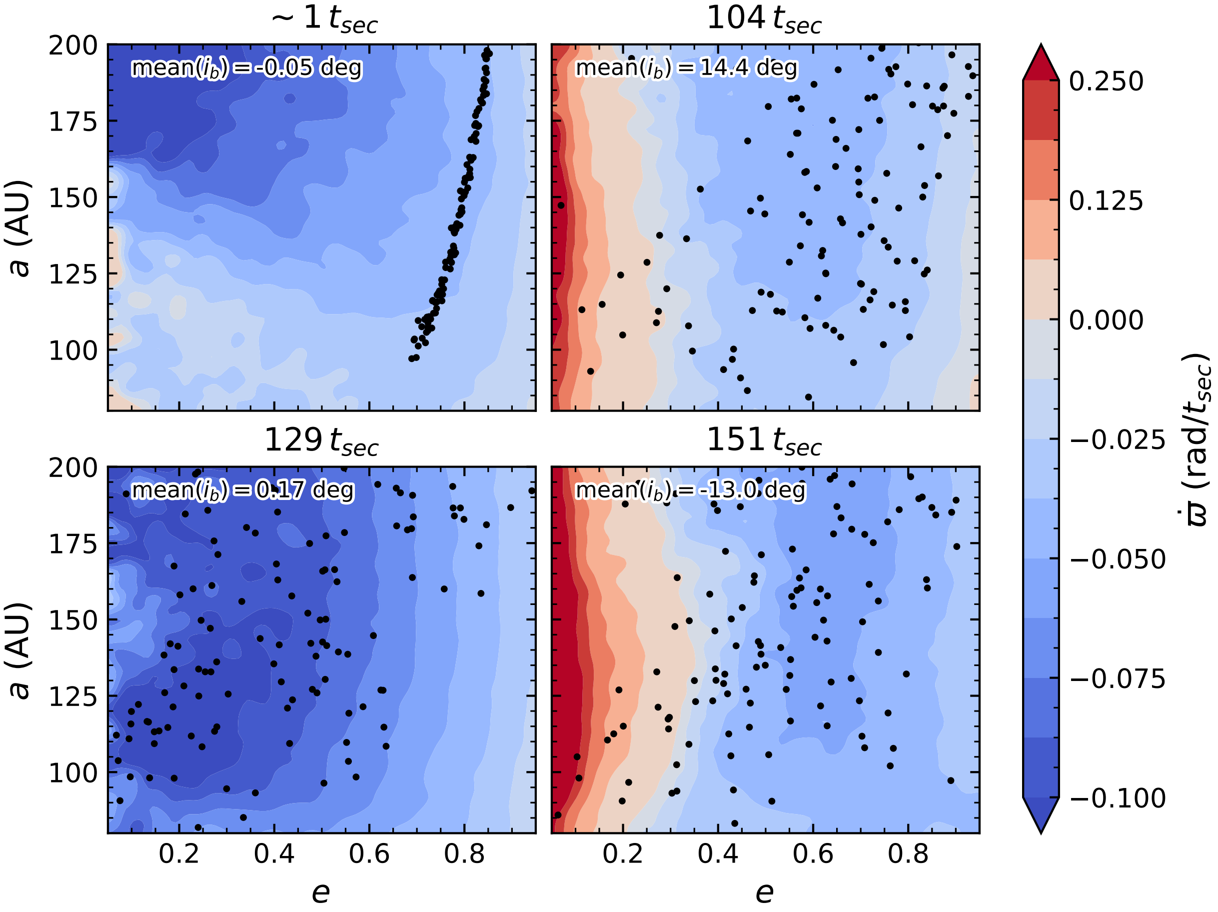

Figure 4.9: Model N400: Contours of the time derivative of the longitude of pericenter, $\dot{\varpi}$, for the inner edge ($a \in [80, 200]$ AU) of the scattered disk simulation without J_2 from Figure 4.5 at four times. The e and a of the disk particles are shown with black points and the mean i_b of the disk at the time is shown in the top left. There is a clustering region at $t = 104\, t_{\rm sec}$ and $t = 151\, t_{\rm sec}$ for $e \in [0.25, 0.60]$. This clustering region is associated with the bowl-shaped orbital configuration created by the inclination instability. Note that the region of prograde precession, $e \in [0.05, 0.25]$, also supports apsidal clustering as described in Section 4.3.2. Apsidal clustering doesn't appear until after $151\, t_{\rm sec}$. At $104\, t_{\rm sec}$ the region $e \in [0.05, 0.60]$ is only occupied by ~ 20 particles, by $151\, t_{\rm sec}$ this has increased to ~ 65 particles. There aren't enough particles in $e \in [0.05, 0.60]$ at $104\, t_{\rm sec}$ for apsidal clustering to begin earlier.

equal to the mean inclination of the post-instability disk (e.g. $\sim 50\,\mathrm{deg}$). We calculate e_R and e_z for each disk to obtain an empirical distribution for e_R and for e_z with $N = 1000$ samples. We calculate the noise floors from these distributions (68th and 95th percentile centered on the mean, corresponding to one and two standard deviations of the Gaussian distribution). The noise floor is a function of N with lower N simulations having higher noise floors. We show the noise floors in Figures 4.5, 4.6, 4.7 and 4.8 with grey bands. e_R and e_z values above the noise floor indicate statistically significant apsidal clustering.

As first described in Madigan & McCourt (2016), the spatial orientation of orbits can be quantified with the angles, i_{a}, i_{b}, and i_{e} representing rotations of an orbit about its semi-major ($\hat{\boldsymbol{a}}$) axis, semi-minor ($\hat{\boldsymbol{b}} \equiv \hat{\boldsymbol{j}} \times \hat{\boldsymbol{a}}$) axis, and angular momentum vector ($\hat{\boldsymbol{j}}$), respectively, such that

$$i_{\mathrm{a}} = \arctan\left[\frac{\hat{b}_{\mathrm{z}}}{\sqrt{1-\hat{b}_{\mathrm{z}}^2}}\right], \tag{4.6a}$$

$$i_{\mathrm{b}} = \arctan\left[-\frac{\hat{a}_{\mathrm{z}}}{\sqrt{1-\hat{a}_{\mathrm{z}}^2}}\right], \tag{4.6b}$$

$$i_{\mathrm{e}} = \arctan\left[\hat{a}_{\mathrm{y}}, \hat{a}_{\mathrm{x}}\right]. \tag{4.6c}$$

The subscripts x, y, and z denote an inertial Cartesian reference frame with unit vectors, $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$. The angles i_{a}, i_{b}, and i_{e} are equivalent to the roll, pitch and yaw of a boat or plane, and are useful for understanding the net gravitational torque acting on an orbit. The inclination instability is characterized by the mean i_a (roll) and i_b (pitch) of all the orbits in the disk growing exponentially with opposite signs. We use these angles in upcoming plots to see how the inclination instability proceeds in simulations with different parameters and how that affects the subsequent growth of a lopsided mode.

4.3.4.1 Inclination Instability

The axisymmetric disks of eccentric orbits in our simulations undergo a dynamical instability called the inclination instability due to the secular gravitational torques between orbits. The instability is characterized by exponential growth in inclination and a corresponding decrease in eccentricity. The initially thin disk expands into a cone or bowl shape[7]. As the orbits' inclinations grow, they tilt in the same way with respect to the disk plane and oscillate coherently in i_a and i_b. We describe the physical mechanism behind the inclination instability in Madigan et al. (2018b) (Chapter 2).

In Figures 4.5, 4.6, 4.7, and 4.8, we show the inclination instability and its aftermath for models N400, N800, J2N400, and J2N800, respectively. Particles are binned by their initial semi-major axis, with the bin boundaries chosen such that the number of particles per bin is approximately equal.[8] Note that the figures have different x-axes (time) but identical y-axes.

In Figures 4.5 and 4.6 (models N400 and N800), the largest growth in inclination occurs in the two innermost semi-major axis bins. In Figures 4.7 and 4.8 (models J2N400 and J2N800) however, the innermost bin ($a < 180$ AU) flattens in inclination after a shorter exponential phase. This difference is seen again in the eccentricity evolution; the innermost semi-major axis bin drops to the lowest eccentricity values in the simulations without added J_2 whereas the drop is suppressed in the simulations with added J_2. In addition, the inclination instability saturates at later times in the J2N400 and J2N800 models than in N400 and N800 models ($\sim 200\,t_{\rm sec}$ vs. $\sim 100\,t_{\rm sec}$). We attribute the difference between the two models to the strong differential apsidal precession in the innermost bin induced by the gravitational influence of the giant planets. This effect decreases the coherence time over which inter-orbit torques can act.

The inclination instability produces *out-of-plane* apsidal clustering, captured by both e_z and i_b. We note that the longitude of pericenter, $\varpi = \omega + \Omega$, fails to find this clustering because it is sensitive only to in-plane clustering.

---

[7] For a visualization of the bowl shape, the reader can jump ahead to the second row, third panel from the left of Figure 4.13.

[8] We've verified that particles do not drift far from their initial semi-major axis during integration.

4.3.4.2 Apsidal Clustering in the Scattered Disk

In Zderic et al. (2020) (Section 4.1), we found *in-plane* apsidal clustering after the inclination instability had saturated in a simple, unrealistic orbital configuration. This “compact configuration” is characterized by an axisymmetric, nearly-flat disk of Keplerian orbits in which all bodies have identical eccentricities and nearly identical semi-major axes. Here, we report the same findings for our scattered disk model with and without added J_2.

In Figures 4.5 and 4.6, e_R traces the development of apsidal clustering in the x/y-plane at the inner edge ($a_0 \in [100, 180]$ AU) of Models N400 and N800—a massive scattered disk without giant planets. Values of e_R above the noise floor indicate statistically significant in-plane apsidal clustering. As in Zderic et al. (2020) (Section 4.1), apsidal clustering appears in the disk after the inclination instability has saturated (Section 4.1). Note that apsidal clustering only appears for bodies with $a \lesssim 320$ AU with clustering first appearing in the $a_0 \in [100, 180]$ AU bin and then travelling out into the $a_0 \in [180, 320]$ AU bin at later times. Comparing the two models, apsidal clustering begins earlier, is more consistent (fewer oscillations), and is stronger in the $a_0 \in [180, 320]$ AU bin in the higher N model, N800, than in the N400 model. Finally, note that the mode strength regularly oscillates below the noise floor, particularly in N400.

Figures 4.7 and 4.8 show the development of apsidal clustering in the inner edge of our J_2 models, J2N400 and J2N800—a massive scattered disk with giant planets. We get apsidal clustering in both models (starting after $\sim 200\, t_{\text{sec}}$), though this clustering is weaker than it is for the models without added J_2. Apsidal clustering appears at later times in the J_2 models because the inclination instability is slowed by the added J_2, and clustering does not appear until after the instability has saturated. In J2N400, apsidal clustering in the $a_0 \in [180, 320]$ AU bin is stronger than it is in the N400 and N800 models. This reflects a general trend of our J_2 results. The J_2 potential disrupts the instability for the lowest a bodies (compare the $a_0 \in [100, 180]$ AU bin in Figures 4.5 and 4.7), but it strengthens the instability in the outer a_0 bins. Bodies with $a_0 \gtrsim 180$ AU attain higher mean i, lower mean e, and stronger apsidal clustering post-instability in the J2N400 model than in N400 and N800 models. In J2N800, statistically significant apsidal clustering only appears in the $a_0 \in [100, 180]$ AU bin, and it's weaker and shorter-lived than all the other models. This is unexpected—clustering was generally stronger in N800 than in N400 so we expect J2N800 to show apsidal clustering similar to or stronger than J2N400. We have multiple simulations of the $N = 400$ models, all showing apsidal clustering. The N800 and J2N800 simulations we show here are the only ones of that particle number that were run long enough to show apsidal clustering. In the $N = 400$ simulations, we find that the strength of clustering varies from simulation to simulation (being as weak as $e_R \approx 0.20$ at peak). Thus, it is possible that the weak apsidal clustering seen in J2N800 is just a peculiarity of that simulation's specific initial conditions.

In all models, in-plane apsidal clustering appears in the inner semi-major axis bin(s) after the instability has saturated. The occurrence of apsidal clustering shortly after the inclination instability in both models suggests that this instability is responsible for the in-plane apsidal clustering.

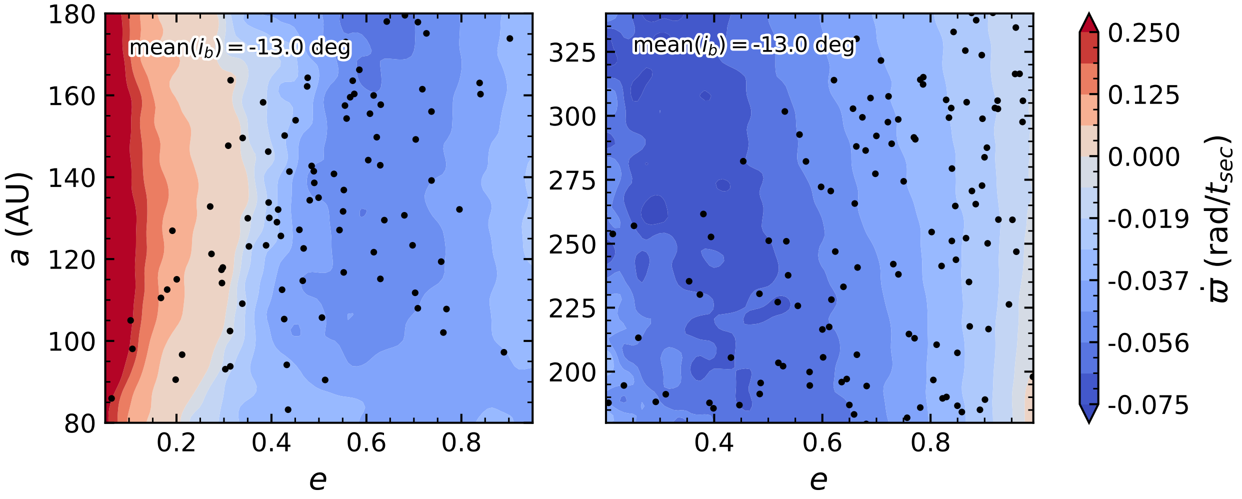

Figure 4.10: Model N400: $\dot{\varpi}$ contour plots for the innermost two a_0 bins at $t = 151\, t_{\rm sec}$. The left panel is the same as the bottom right panel in Figure 4.9, and it shows a large clustering region ($e \lesssim 0.6$). The right panel shows a small, underpopulated clustering region at the lowest a (note that the e axes are different in the two panels). This explains why we only see strong apsidal clustering for $a \lesssim 180\,$AU. The contour plots differ at their boundary, 180 AU. This is due to the different mean test particle ω in these two bins and it demonstrates the importance of ω in forming the clustering region.

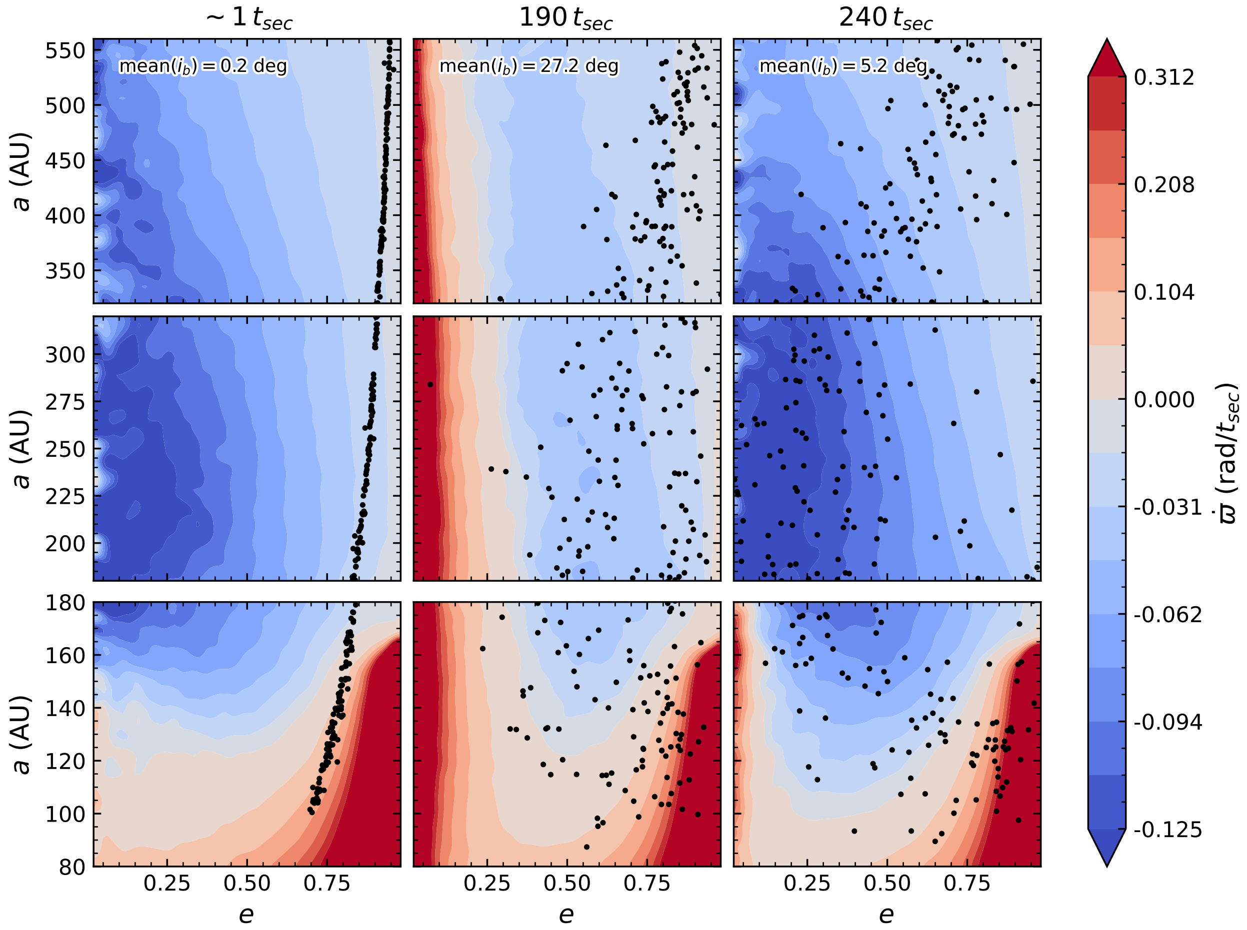

Figure 4.11: Model J2N400: The emergence of a clustering region post-instability in a scattered disk model including J_2 (the same simulation as shown in Figure 4.7). We show contours of $\dot{\varpi}$ from three different a bins ($a_0 \in [100, 180]$ AU, $a_0 \in [180, 320]$ AU and $a_0 \in [320, 560]$ AU) at three different times. Mean i_b for the disk is shown in the top left of the top panel. Peak mean i_b is attained at $190\,t_{\text{sec}}$, and a corresponding, populated clustering region is shown for $e \in [0.05, 0.55]$ and $a \in [125, 300]$ AU. The deep trough of prograde precession at high e and low a in the bottom row of panels is due to the J_2 potential.

4.3.4.3 Emergence of the clustering region

Apsidal clustering in our simulations occurs after a Lynden-Bell clustering region appears. Here we show that the clustering region appears during the saturation of the inclination instability and is associated with the unique, 'bowl'-shape of the mass distribution post-instability.

We simulate test particles in the frozen gravitational potential of the disk to find $\dot{\varpi}$ as a function of a and e. The disk particles from a fully interacting simulation are frozen onto their orbits at a specific time and a test particle is integrated in this background disk potential. We use these test particles to make contour plots of $\dot{\varpi}$ in e-a space at specific instances to find clustering regions, if any exist. In these test particle simulations we have a test particle, a background disk, and a central body (with/without added J_2). The background disk particles are given the REBOUND/MERCURY 'small particle' designation meaning they do not interact with each other; they only interact with the test particle and the central body (they do actually interact with each other indirectly through the central body as mentioned in Peng & Batygin (2020); this effect is small). Paradoxically, we must give the test particle mass in order for it to interact with the background disk. The mass of the test particle is set to be so small that it negligibly affects the background disk bodies. The test particle simulations are integrated for 10 orbits, and the test particles are initialized with a and e drawn from a grid (96×96), ω and i calculated from the mean values of the local disk (same a bin), and an Ω of 0, $\pi/4$, $\pi/2$, or $3\pi/4$. The median $\dot{\varpi}$ is calculated from each set of four test particle simulations. This is the method used to create the contour plots shown in Figures 4.9, 4.10, and 4.11. We have checked the accuracy of our test particle simulations with an alternative method which calculates the instantaneous precession rate of the test orbit directly from the torques and forces it experiences; the results are are consistent, as shown in Appendix C. We only show $\dot{\varpi}$ for the $N = 400$ simulations as the $N = 800$ results are the same.

In Keplerian elements, a clustering region (see Section 4.3.2) is defined by $\partial\dot{\varpi}/\partial e|_a < 0$. Apsidal precession within our disks is initially retrograde ($\dot{\varpi} < 0$) and with *magnitude* decreasing with increasing eccentricity ($\partial\dot{\varpi}/\partial e|_a > 0$). When clustering regions appear within our disks, we see regions where the precession is retrograde with increasing magnitude and regions where precession is prograde with decreasing magnitude. In the contour plots, clustering regions are regions where, at fixed semi-major axis (a), the contours go from warmer to cooler colors with increasing eccentricity (e).

In Figure 4.9, we show the development of a clustering region in the inner edge of N400, the scattered disk model without added J_2. This figure shows the time derivative of the longitude of pericenter, $\dot{\varpi}$, in the scattered disk as a function of semi-major axis a and orbital eccentricity e for the inner edge of the disk. We show mean i_b for the disk in the top left of the panels. Initially, all bodies in the scattered disk are on the line $a\,(1-e) = 30\,\mathrm{AU}$ (top left). Later, the inclination instability reduces the disk orbits' eccentricity, e, at roughly fixed semi-major axis, a (top right), and causes the disk to buckle in to a bowl-shape. Notably, the $\dot{\varpi}$ contours have changed to admit a clustering region ($\partial\dot{\varpi}/\partial e|_a < 0$) covering $e \in [0.25, 0.60]$ and $a \in [80, 200]\,\mathrm{AU}$. This retrograde clustering region smoothly blends into a prograde region ($e \in [0.05, 0.25]$) which also facilitates apsidal clustering. Thus the whole region $e \in [0.05, 0.60]$ supports apsidal clustering. Immediately after the instability saturates, the apsidal clustering region is lightly populated. The eccentricity continues to drop after the inclination instability leaves the linear regime (bottom left). However, the clustering region has disappeared. This is because the disk has precessed out of the bowl-shape (mean $i_b \sim 0°$). Finally, the orbits at the inner edge have precessed through the ecliptic and inverted the bowl-shaped mass distribution (mean $i_b > 0$) (bottom right). Again, the clustering region appears, but now it is sufficiently populated for in-plane apsidal clustering to take hold.

Two things are apparent from this sequence. First, the clustering region coincides with the bowl-shaped orbital distribution (large mean i_b). Second, in-plane apsidal clustering only appears once the clustering region is sufficiently populated.

Once apsidal clustering has been established and the lopsided mode has grown, it is no longer reliant on the clustering region produced by the bowl-shape to exist. The bowl-shape is not actively maintained after the inclination instability saturates. In Figure 4.5, differential precession slowly erodes mean i_b in the disk, and will eventually erase the bowl-shape altogether. However, e_R appears unaffected by this, and apsidal clustering actually reaches peak strength by the end of the simulation even though the mean i_b has dropped quite low. The bowl-shape seeds the lopsided mode, but, once seeded, the mode is self-sustaining even though the strength of the mode oscillates.

The clustering region appears towards the inner edge of the disk in N400. In Figure 4.10, we show the inner two semi-major axis bins of the disk ($a \in [100, 320]\,\mathrm{AU}$) at $151\,t_{\mathrm{sec}}$. The clustering region extends to $a > 200$ AU, but it is largest at lower a and it's unpopulated for $a \gtrsim 250$ AU. This explains why apsidal clustering is primarily only seen in the inner two bins of these simulations, and why apsidal clustering is slightly weaker for $a_0 > 180$ AU. The precession rates are discontinuous at 180 AU (top of the left panel and bottom of the right panel), and the two panels have different x axes, exacerbating the apparent discontinuity. The discontinuity is due to the test particles in these two panels having different bin-averaged ω_0 and i_0.

The general features found in the N400 model are repeated in the J2N400 model: the clustering region appears around peak mean i_b (in the 'bowl'-shape), the semi-major axis location of the clustering region traces apsidal clustering, and the clustering region precedes apsidal clustering. In Figure 4.11, we show contours of $\dot{\varpi}$ across the disk at three distinct times for model J2N400, the same simulation depicted in Figure 4.7. A deep trough of prograde ϖ precession from the added J_2 is seen at large eccentricity and small semi-major axis. Bodies near this trough do not undergo the inclination instability and their eccentricities are stable. The clustering region still forms (middle column), but at slightly larger semi-major axis because the lower semi-major axis portion of the disk is too disrupted by the J_2 potential. This is reflected in Figure 4.7 by apsidal clustering in $a_0 \in [100, 180]\,\mathrm{AU}$ and $a_0 \in [180, 320]\,\mathrm{AU}$. The mean i_a of the inner edge of the disk is ~ 0 (due to the added J_2 precession), but mean $i_b > 0$.

In our simulations, the formation of the clustering region coincides with peak mean i_b, signifying that the unique out-of-plane orbital distribution resulting from the inclination instability is responsible for the in-plane apsidal clustering. The post-instability bowl-shaped potential creates a clustering region at low eccentricity, and this region is simultaneously populated by the circularizing effect of the instability.

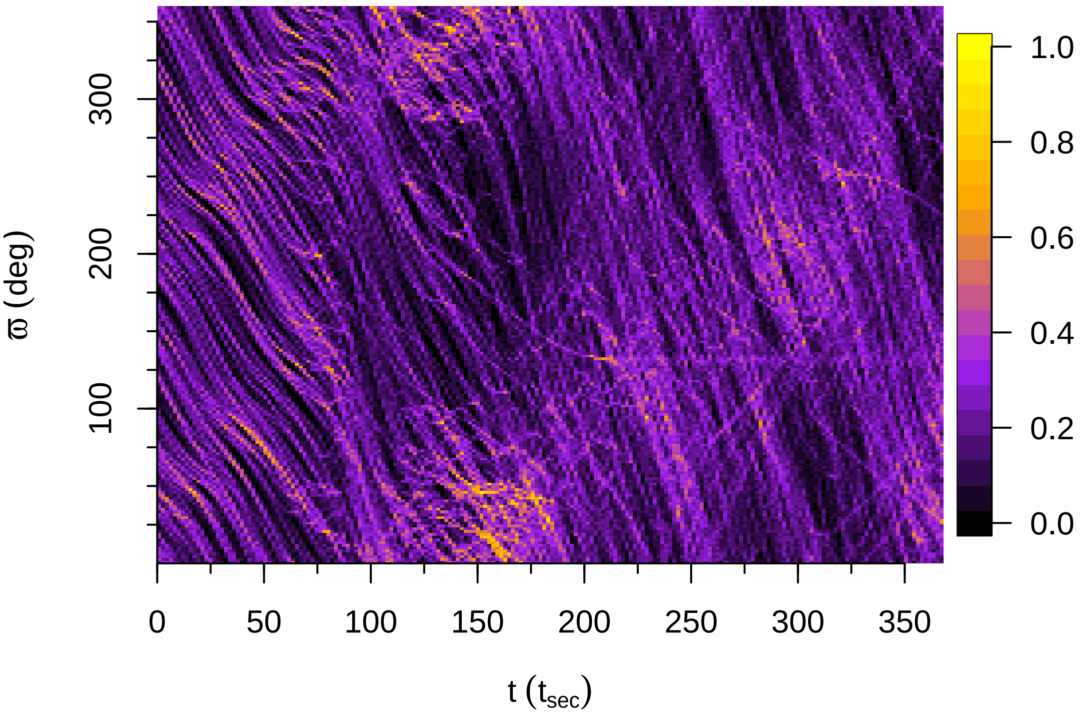

Figure 4.12: Model N800: Clustering in ϖ as it evolves over time in the inner a_0 bin ($a_0 \in [100, 180]$ AU). This figure shows a 2d-histogram of ϖ as a function of time for all particles in this bin, and the colors represent the density of points (normalized to the densest area). We identify the mode as the highly dense region in this figure that generally precesses prograde, opposite to the individual orbits that precess retrograde. Retrograde precessing orbits can be captured in the prograde precessing mode, and orbits actually caught in the mode librate within the mode. The pattern speed of the mode is $\sim 1.5° \, t_{\rm sec}^{-1}$ which is of the same order of magnitude as the precession rate of the orbits in the disk.

A particle's precession rate depends on the forces it experiences throughout its orbit. From equation (C.1) the orbital precession rate is

$$\dot{\varpi} \approx \langle f_r(r) \rangle \, \frac{\sqrt{1-e^2}}{e} \sqrt{\frac{a}{GM}}, \tag{4.7}$$

where $\langle f_r(r) \rangle$ is the orbit-averaged specific, radial force. This is negative if the force is radially inwards, such that precession will be retrograde. For retrograde precession, the magnitude of $\dot{\varpi}$ must increase with eccentricity in a clustering region. In contrast, the second term on the right-hand-side of equation (4.7) decreases monotonically with eccentricity. Therefore, the orbit-averaged force must increase with eccentricity in a retrograde clustering region. This is somewhat unusual considering that typically in a Keplerian potential (i) precession is dominated by the forces near apocenter and (ii) these forces will be smaller at larger apocenters. However, this is not the case in the post-instability bowl-shaped orbital distribution.

In Appendix C, we measure the precession rates of orbits in the clustering region, by calculating the forces and torques at many points along the orbit. We find that changes in the precession rate (with eccentricity) can be dominated by points near pericenter. Additionally, forces can increase with apocenter. This behavior allows a clustering region to appear.

For prograde precession, clustering will occur if the magnitude of $\dot{\varpi}$ decreases with eccentricity, which is typical for a Keplerian potential. In fact this will occur for any external force that decreases with radius (see equation 4.7).

4.3.4.4 Mode direction

In galaxies, the slowness condition, $\nu_i - \nu_p \ll \nu$, will only be satisfied if the mode and the orbits precess in the same direction because $\nu_i \lesssim \nu$ (galactic orbits are generally rosettes in the inertial frame). If $\text{sgn}(\nu_i) = -\text{sgn}(\nu_p)$ then $\nu_i - \nu_p \sim \nu_i$ and the orbits are not nearly closed in the rotating frame of the bar perturbation. However, in near-Keplerian systems, both $\dot{\varpi}_i$ and ν_p are much less than ν and $\dot{\varpi}_i - \nu_p \ll \nu$ is true even if $\text{sgn}(\dot{\varpi}_i) = -\text{sgn}(\nu_p)$. If the mode and the orbits precess in opposing directions, the relative orbital precession rate in the frame rotating with the mode will be greater than if the mode and orbit precess in the same direction. However, the relative orbital precession rate will still be $\mathcal{O}(t_{\text{sec}}^{-1})$, and secular torques between the orbit and the mode can still be dynamically important. Therefore, in near Keplerian systems, it is not dynamically forbidden for a mode to form via the Lynden-Bell mechanism with a precession direction opposite the orbital precession direction. Indeed, we generally see the mode precess opposite to that of the orbits in our simulations.

In Figure 4.12, we show a 2d-histogram of ϖ as a function of time for all particles in inner a_0 bin ($a_0 \in [100, 180]$ AU) of model N800. We see from this figure that the individual orbits precess retrograde, and cluster together to form a mode starting around 100 t_{sec}. The figure shows that the mode generally precesses prograde with a pattern speed of $\sim 1.5^\circ\, t_{\rm sec}^{-1}$. The mode is capable of capturing orbits that precess counter to it. The captured orbits then librate within the mode, precessing prograde then retrograde within the mode. Eventually, orbits leave the mode and precess retrograde again within the disk.

4.3.5 Simulated Observations

While the inclination instability appears to be promising in explaining the clustered, detached orbits of extreme Trans-Neptunian Objects in the outer Solar System, it should also occur in exoplanet systems with at least one giant planet that can form a massive scattered disk. In particular, it provides a mechanism for creating asymmetric debris disk structures such as the wing-like features in HD 61005 (MacGregor et al., 2019).

In Figure 4.13, we plot snapshots of the J2N400 simulation—a primordial scattered disk with orbit-averaged gravitational influence of the giant planets. We first populate each orbit of the simulation with 100 particles spaced uniformly in mean anomaly to increase the effective resolution. We then make maps of surface density and velocity along the line of sight with a pixel resolution of 20 AU. The surface density, Σ, of the disk in face-on (top frames) and edge-on (bottom frames) orientations are plotted in the left-hand columns. Time is increasing from left to right and down the column. In the x/y-plane the particles orbit in the counter-clockwise direction. Except for the innermost edge of the disk, the orbits precess in the clockwise direction. The initially thin, axisymmetric disk undergoes the inclination instability, buckling above and below the plane ($t \approx$ 196-303 $t_{\rm sec}$). The lopsided mode develops in the x/y-plane as differential precession disperses the asymmetric distribution of orbits in the x/z-plane. At early times, a spiral arm links the inner disk to the most over-dense region of the mode in the outer disk.

To the right of the surface density plots, we plot the corresponding velocity along the line of sight, $v_{\rm los}$. Red and blue colors illustrate red-shifted and blue-shifted velocities with respect to the observer. The initial velocity distribution is dominated by rotation around the Sun, as shown in the x/z-plane. The collective rolling and pitching of the orbits about their major and minor axes

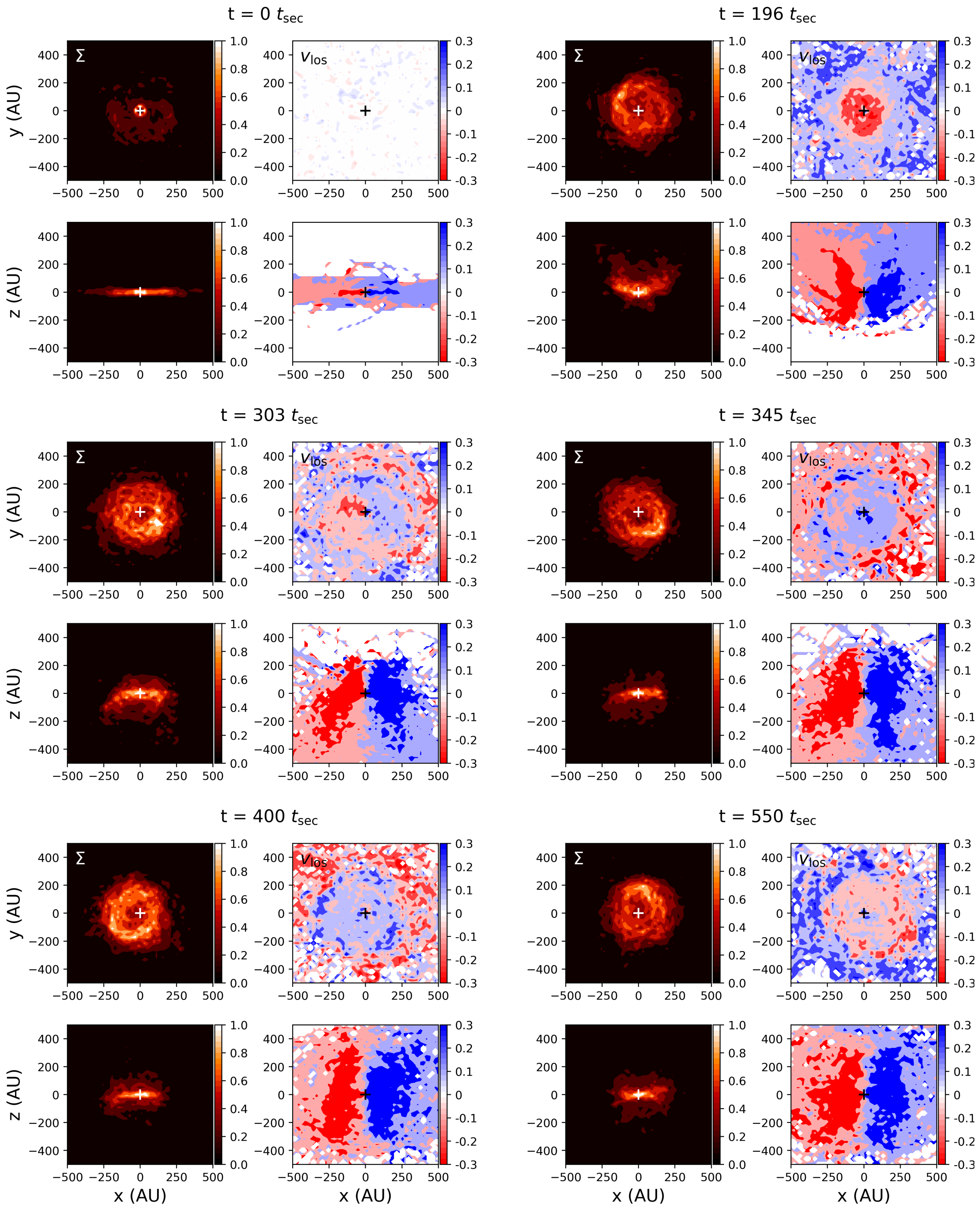

Figure 4.13: Snapshots in time of the J2N400 simulation–a primordial scattered disk with the orbit-averaged gravitational influence of the giant planets. Surface density, Σ, and velocity along the line of sight, $v_{\rm los}$ are plotted for face-on and edge-on orientations. The inclination instability occurs around 196 $t_{\rm sec}$. At 303 $t_{\rm sec}$, we observe a spiral in the line of sight velocity when the disk is viewed face-on.

(captured by the angles i_a and i_b in Figure 4.7) is apparent in the velocity map at 196 $t_{\rm sec}$ which shows resulting concentric circles of red-shifted and blue-shifted velocities. We note that this is equivalent to the clustering of the orbits in argument of pericenter (ω).

At $t \approx 303\, t_{\rm sec}$, a spiral arm in velocity space appears in the x/y-plane. This occurs as the amplitude of i_a for the inner orbits ($a \lesssim 500$ AU) passes through zero but their i_b values are significantly non-zero, as seen in Figure 2 at t $\approx 303\, t_{\rm sec}$. The lopsided mode in the x/y-plane is apparent in surface density before we see the spiral arm in velocity space. The over-dense cluster of orbits leads to positive line of sight velocities on one side of the clump and negative on the other, leading to the appearance of a spiral.

Another spiral arm appears in velocity space when the orbits in the disk pass through $i_a = 0$ again at $t \approx 480\, t_{\rm sec}$ as seen in Figure 2. At all other times we see the concentric circles of red-shifted and blue-shifted velocities, alternating as the orbits coherently precess above and below the mid-plane. Velocities in the x/z-plane continue to show rotation in an increasingly thick disk.

4.3.6 Conclusions

In this paper, we demonstrate the spontaneous apsidal clustering of orbits of low mass bodies in N-body simulations of a primordial scattered disk between ∼100–1000 AU. As in Zderic et al. (2020) (Section 4.1), we find that apsidal clustering begins after the inclination instability has saturated, and that the inclination instability is key to the formation of the lopsided mode. In simulations where the orbit-averaged, gravitational influence of the giant planets is included, we find that apsidal clustering occurs provided that the inclination instability is not suppressed. We also find that apsidal clustering only forms near the inner edge of the disk in the 100–320 AU range with the specific range depending on the model, but we caution that our simulations have low numbers of particles particularly at large semi-major axes. The fast orbital precession caused by the giant planets pushes the location of apsidal clustering out to larger semi-major axis. Finally, we find that the resulting lopsided mode strength oscillates, but appears long-lasting.

Lynden-Bell (1979) proposed a mechanism to explain stellar bar formation in the center of galaxies that we extend here to near-Keplerian systems to explain the apsidal clustering that occurs in our simulations. Orbit-averaged torques from a weak, lopsided mode encourages orbits into precessing towards alignment with the mode.[9] In a Keplerian system, if $\partial\dot{\varpi}/\partial e|_a < 0$ then orbits will tend to align with and reinforce the mode. We call regions of e-a space where $\partial\dot{\varpi}/\partial e|_a < 0$ clustering regions.

We have created contour plots of $\dot{\varpi}$ as a function of eccentricity and semi-major axis within the disk at different times. We find that a clustering region forms during the peak of the inclination instability when the disk has formed a bowl-shape. The clustering region appears just before apsidal clustering begins, and it appears at the inner edges of the disk. In simulations with the orbit-averaged gravitational influence of the giant planets, the added J_2 inhibits circularization of the inner edge of the disk during the instability and amplifies circularization at larger semi-major axis. As a result, the clustering region is populated by bodies with larger semi-major axis in the J_2 models and apsidal clustering correspondingly occurs at larger semi-major axes.

The clustering region is directly correlated with the unique bowl-shaped orbital distribution created by the inclination instability. Due to orbital precession, the bowl-shaped distribution oscillates back and forth across the original plane of the disk, causing the clustering region to repeatedly disappear and reappear, and eventually, the bowl-shape disappears. However, we find that the lopsided mode created by the clustering region persists. We hypothesize that the mode eventually becomes massive enough to trap orbits without the help of the background disk potential.

Surface density plots of our disks during the inclination instability show edge-on wing-like structures reminiscent of some debris disks (e.g. HD61005), and a lopsided mode in face-on views after the instability has saturated. In line-of-sight velocity, we see concentric circles of alternating sign associated with the bowl-shaped orbital distribution post-instability. Later, the lopsided mode creates spiral arms in line-of-sight velocity. Observational signatures like this in exoplanet disks could be caused by the inclination instability provided there is something to pump-up the orbital

---

[9] We assume an initial small mode is seeded by random fluctuations within the disk.

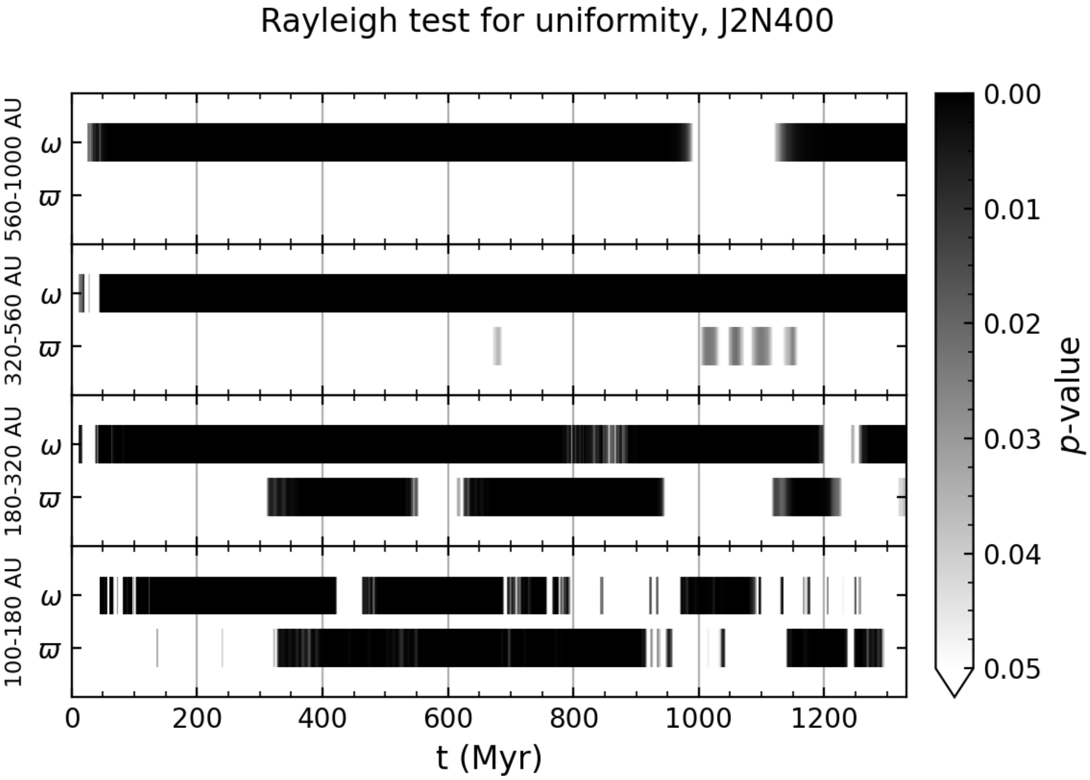

Figure 4.14: Rayleigh test results for ω and ϖ binned by initial semi-major axis as a function of time for the J2N400 model scaled such that the inclination instability saturates (ceases exponential growth) at 250 Myr. Rayleigh test p-values less than 0.05 indicate that the distribution of ω or ϖ is not uniform (i.e. it is clustered) and smaller p-values indicate stronger clustering. ω-clustering begins after ∼10 Myr and persists until the end of the simulation at ∼1300 Myr, except for bodies with $a_0 \in [100, 180]$ AU. ϖ-clustering occurs for bodies with $a_0 \in [100, 320]$ AU after the inclination instability has saturated at ∼250 Myr, and it disappears and reappears intermittently.

eccentricity of the bodies in the disk (e.g. a giant planet).

In Zderic & Madigan (2020) (Chapter 3), we found that ∼20 Earth masses is required for a primordial scattered disk to resist the orbit-averaged quadrupole potential of the giant planets *at their current locations* and undergo the inclination instability[10].

The e-folding timescale for the inclination instability in a scattered disk configuration with

---

[10] Twenty earth masses is extreme for a primordial scattered disk (indeed perhaps too massive for the instability to have occurred in our solar system). While hundreds of earth masses may have been present in the early planetesimal disk beyond ~ 5 AU, only a fraction appears to pass through the scattered disk region (Nesvorný, 2018).

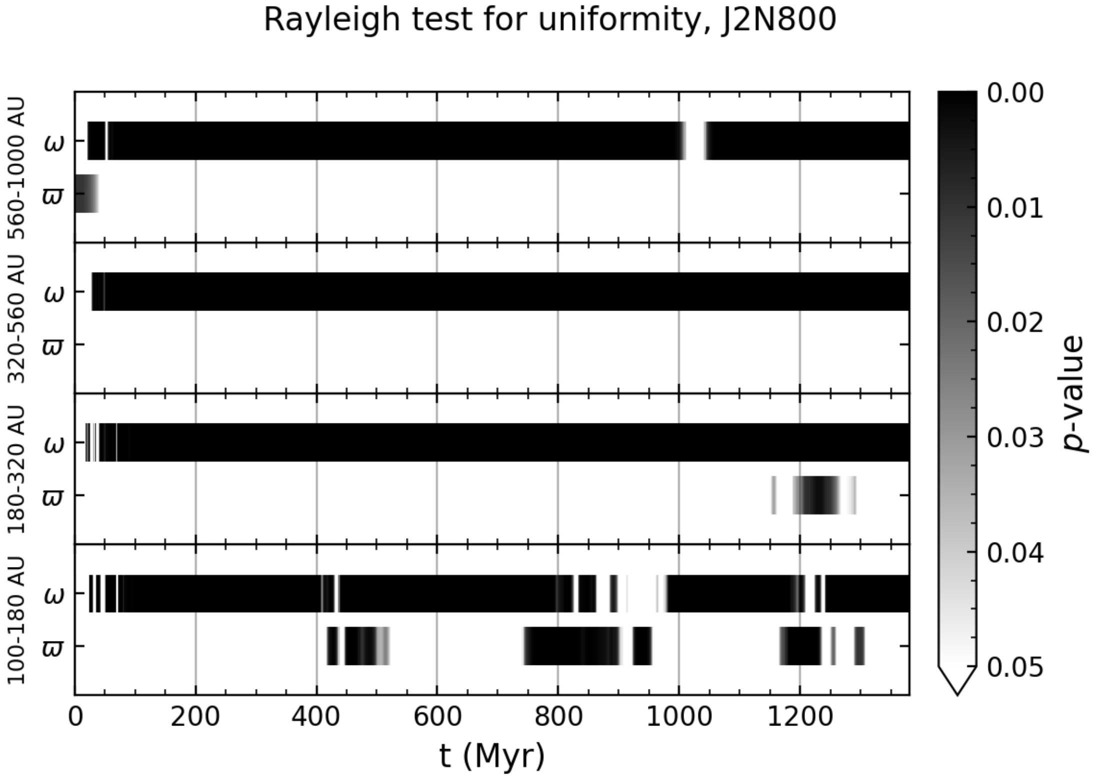

Figure 4.15: Rayleigh test results for ω and ϖ binned by initial semi-major axis as a function of time for the J2N800 model scaled such that the inclination instability saturates (ceases exponential growth) at 250 Myr. Rayleigh test p-values less than 0.05 indicate that the distribution of ω or ϖ is not uniform (i.e. it is clustered) and smaller p-values indicate stronger clustering. ω-clustering begins after ∼10 Myr and persists until the end of the simulation at ∼1300 Myr. ϖ-clustering occurs primarily in bodies with $a_0 \in [100, 180]$ AU after the inclination instability has saturated, and it is more intermittent than in the J2N400 model. The initial ϖ-clustering in the outer a_0 bin is a random quirk of our initial conditions, and it quickly disappears.

$N \to \infty$ and without added J_2 is (Zderic & Madigan, 2020, Chapter 3),

$$t_{\mathrm{e-fold}} \sim \frac{2.4}{\pi} \frac{M_\odot}{M_d} P. \tag{4.8}$$

For a 20 Earth mass disk, $t_{\mathrm{e-fold}} = 1.3 \times 10^4 \, P$. Based on Figure 3.3, we expect this timescale to be increased by a factor of ∼4 in simulations with added $J_2 \lesssim J_{2,\mathrm{crit}}$. With the inner edge semi-major axis being 100 AU, $P = 1000\,\mathrm{yr}$, and the e-folding timescale for the inclination instability in a 20 Earth mass Scattered Disk in the outer solar system is ∼50 Myr. It takes about five e-folding timescales for the inclination instability to saturate. Therefore, the inclination instability in this primordial scattered disk should saturate after $\sim$250 Myr.

We can estimate the duration of ω and ϖ clustering in the 20 Earth mass primordial scattered disk using the J2N400 and J2N800 simulations. We set the saturation time in these simulations to be 250 Myr and scale the subsequent evolution of the disk using the secular timescale. For example, $t_{\rm sec} = 2.6\,{\rm Myr}$ (see Equation 4.2) for a 20 Earth mass disk. The J2N800 simulations runs for $\sim$400 $t_{\rm sec} \approx 1050\,{\rm Myr}$ after the saturation of the instability. We stress that this scaling is approximate and meant to provide a qualitative, order-of-magnitude estimate for the duration of angular clustering in an unstable primordial scattered disk. In Figures 4.14 and 4.15, we show the p-values for the Rayleigh z test for uniformity (Mardia & Jupp, 2000) on ω and ϖ for the J2N400 and J2N800 models as a function of time in Myr using this proposed scaling. Rayleigh z test p-values less than 0.05 signify that the angular distribution is not consistent with a uniform distribution, i.e. that the distribution is clustered. We chose the Rayleigh test as it is sensitive to unimodal deviations from uniformity. In both models, ω-clustering begins after just a few tens of Myr and persists for the duration of the simulation except in the inner semi-major axis bin (100-180 AU) where differential precession is the strongest. Intermittent clustering in ϖ begins after the inclination instability has saturated in the inner two bins in both simulations. Note that binning the results by a_0 partially mitigates the effects of differential precession, prolonging the duration of angular clustering. Overall, we expect a 20 Earth mass primordial scattered disk to be able to sustain ω-clustering for $a \gtrsim 180$ AU (if binned by semi-major axis) and intermittent periods of ϖ-clustering for $a \in [100, 320]$ AU for Gyr timescales.

The inclination instability can raise perihelia and inclinations of bodies in the outer solar system. As such, it can effectively trap planetesimal mass at semi-major axes of hundreds of AU as bodies are isolated from strong scattering encounters with the giant planets. In Zderic & Madigan (2020) (Chapter 3) we show that orbits with semi-major axes between $\sim 200 - 500$ AU end up with a (rather extraordinary) median perihelion distance of 150 AU post-instability; see e.g., Figure 6. The observed Sednoids in this scenario mark the inner edge of a massive reservoir of extremely detached bodies originating from the primordial scattered disk.

The mass remaining at hundreds of AU today is an open question that we are actively exploring. Our simulations show that, post-instability, inter-orbit torques induce eccentricity (but not necessarily inclination) oscillations on particles in this structure. In future work we will calculate the flux of particles back into the inner solar system through these oscillations (which can cause perihelia to drop below the orbit of Neptune) and ultimately the mass loss rate.

Chapter 5

Discussion and Future Work

5.1 Discussion

Madigan & McCourt (2016) originally proposed that the inclination instability, a dynamical instability driven by collective gravity, could explain ω-clustering in the ETNO population. The work presented in this thesis developed this concept further. The most important results from these papers are:

- N-body simulations with artificially low numerical resolution (N) will suppress secular dynamical phenomena like the inclination instability (Chapter 2).
- The mechanism behind the inclination instability can be understood through a two-orbit mutual secular torque model (Chapter 2).
- Differential apsidal precession, induced by the orbit-averaged potential of the giant planets, can suppress the inclination instability within a Scattered Disk orbital distribution if the total mass of the disk is too low. Indeed, this appears to be the primary limiting factor for the instability in the Scattered Disk (Chapter 3).
- The orbital distribution post-instability excites a lopsided $m = 1$ mode in the disk via the Lynden-Bell mechanism for galactic bar formation (Chapter 4).
- About $20\,M_\oplus$ of material are required to have existed in a primordial Scattered Disk orbital structure for a few tens of Myr for the inclination instability to explain the observed

peculiarities in the Outer Solar System.

In light of these results, how viable is the inclination instability the solar system? Recent Neptune migration studies require the primordial Kuiper Belt to have $M_d \gtrsim 15\,M_\oplus$ for Neptune to migrate a significant distance and $M_d \lesssim 20\,M_\oplus$ for the migration to be slow enough to explain the inclination distribution of the Resonant Bodies (Gomes et al., 2004; Nesvorný, 2015, 2018). Therefore, for the inclination instability to explain Outer Solar System observations, essentially all of the mass in the *outer* planetesimal belt would need to end up in the primordial Scattered Disk. This is at odds with our understanding of Neptune migration.

However, the inclination instability may still have a place in the Solar System's history. The Neptune migration simulations do not allow for self-interactions[1] within the planetesimal disk due to numerical limitations. The mass requirements set by these simulations *could* be different in a self-interacting disk. However, it's unlikely this could resolve the mass tension alone. One option is to have the inclination instability occur in the entire population of scattering planetesimals. The reality is that scattering planetesimals were spread throughout and between the giant planets during planetesimal-driven migration which each planet having something like a scattered disk (e.g. see Gomes et al., 2004). This 4-fold scattered disk population would've been more massive than Neptune's Primordial Scattered Disk alone, and possibly able to undergo the inclination instability. A final option is that the inclination instability could possibly occur at an earlier epoch before the gas in the protoplanetary disk disspated. There would've been more mass in planetesimals at this time, however the gas may be able to inhibit the inclination instability in this case. We have not explored the inclination instability in either of these cases[2].

5.2 Future Work

There are several directions in which we could take this research. Here we describe five projects, some already underway: "observing" our simulations using the Vera Rubin Observatory

---

[1] Fan & Batygin (2017) attempted a Nice model with a fully-interacting planetesimal disk, but their results were plagued by numerical resolution problems.

[2] We thank David Nesvorný for pointing out these options to us.

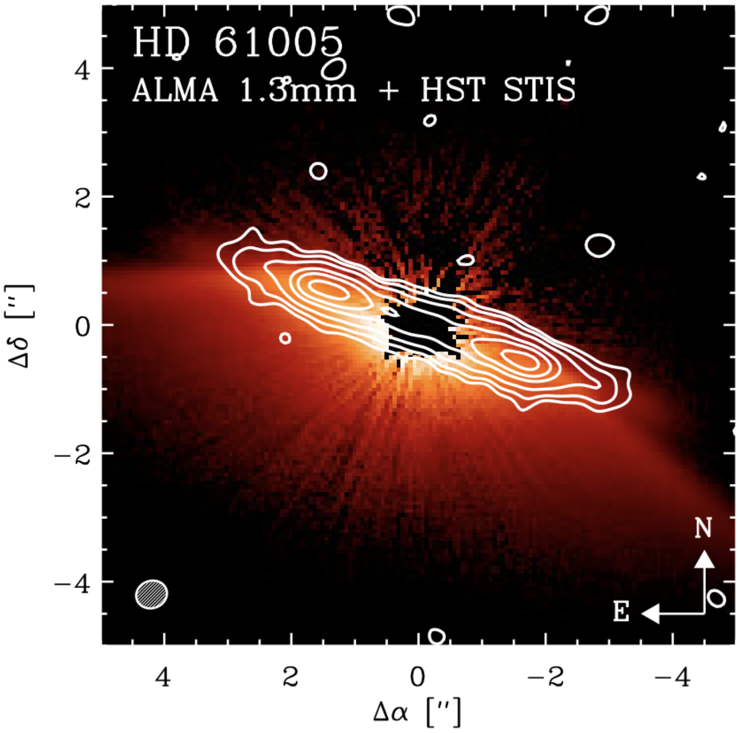

Figure 5.1: Reproduced from MacGregor et al. (2019), scattered light and ALMA (white contours) observations of debris disk HD 61005. An obvious wing- or cone-structure is visible in scattered light. The ALMA observations aren't sufficiently resolved to determine if there is a corresponding structure in the mm.

survey simulator, applying our results to exoplanet debris disks, the generalization of the two-orbit model from Chapter 2 ((Madigan et al., 2018b)) to an N-orbit model, improving an existing code that calculates the equilibria of secular gravitational systems and using it to see if the lopsided mode described in Chapter 4 is a long-term stable structure, and using machine learning to accelerate expensive Neptune migration simulations by having it mediate the interactions between the planetesimals and Neptune.

Vera Rubin Observatory We are working with Professor Mario Jurić simulating Vera Rubin Observatory (VRO) observations with our simulation results. As the numerical resolution of our simulations is low, we're using a Kernel Density Estimation (KDE) method to increase the sampling. The physics in our simulations is also affected by the numerical resolution, e.g. we see significantly more spreading in semi-major axis than we would expect in s system with higher N.

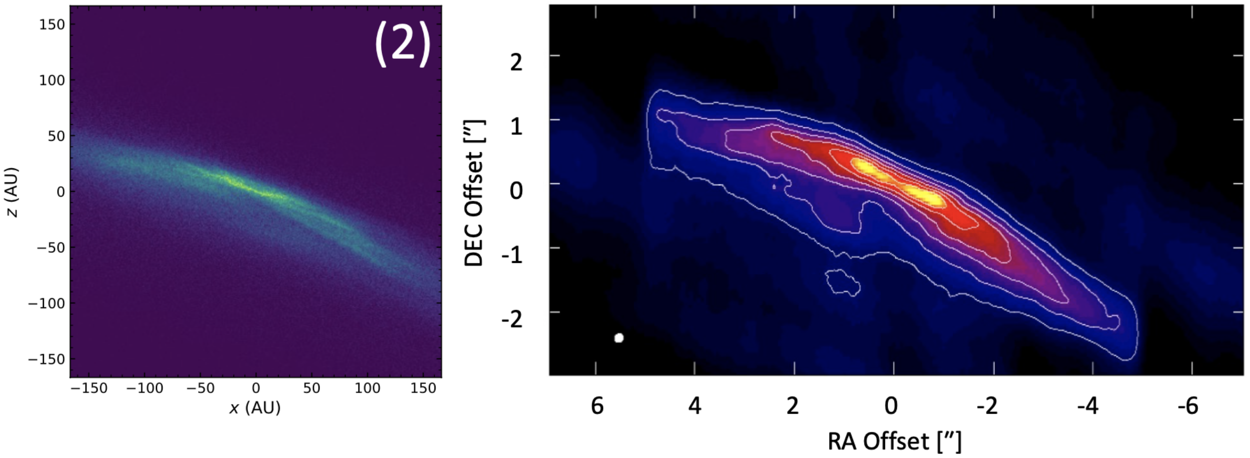

Figure 5.2: (Left) a KDE sampling of a disk mid-inclination instability. (Right) simulated ALMA observations produced by Meredith MacGregor of this disk at a higher resolution than in 5.1 show the wing structure.

This must be corrected in future in order to do more accurate simulated observations.

Exoplanet Debris Disks Even if the inclination instability did not occur in the Solar System, it could occur in other stellar systems. For example, the inclination instability could be used to explain some strange features observed in debris disks. We've started working with Professor Meredith MacGregor looking at a debris disk, HD 61005 ("the Moth"), which shows a remarkable cone shape in scattered light (see Figure 5.1, reproduced from MacGregor et al. (2019)). At early stages of the inclination instability, the disk mass distribution mirrors this structure (see Figure 5.2). Scattered light doesn't always trace the mass distribution and higher resolution ALMA images of HD 61005 are needed to check if the debris disk also has this cone-shaped structure.

***N*-orbit model** The expression for the instability e-folding time as $N \to \infty$ is critical for the application of the inclination instability to any real astrophysical system. We can improve the process detailed in Section 2.5 to get the $N \to \infty$ inclination instability timescale. One method would be to generalize the linear stability analysis used for the two-orbit model to N orbits. The two, coupled, first order linear ODEs defined in Equations 2.5 can be generalized to N orbits. The resulting system of $2N$ linear ODEs is solvable with standard matrix methods (see Chapter 2 in Meiss (2007)). This could provide us with an independent, scattering-free measurement of the growth rate of the instability in a whole disk. It could also elucidate why certain orbital distributions are amenable to the inclination instability (e.g. the Scattered Disk) while others are not.

Statistical Mechanics of the Lopsided Mode The oscillatory nature of apsidal clustering post-instability in our disks throws doubt on the longevity of that apsidal clustering. Some of our results suggest that this is a pathology related to the low N in our simulations (see Figure 4.4). We would like definitive proof that this clustering would be long-lasting in a disk with realistic N. One approach would be to prove that the lopsided mode found post-instability is a secular gravitational equilibrium. Recently, Gruzinov et al. (2020) outlined a new algorithm that can be used to find secular gravitational equilibria for general systems. I reproduced that code, and some of their results. Unfortunately, the code suffers from a shortcoming — it doesn't conserve energy and angular momentum while it converges towards an equilbrium. The equilbria the algorithm finds are valid, but it is impossible to give the algorithm a real system (e.g. one of our disks) and have it find that system's equilibrium. Together with Professor Yuri Levin and Andrei Gruzinov, I have developed an improved version of algorithm that resolves this shortcoming. Once implemented, this code would be able to determine if the lopsided mode found in our simulations is an equilbrium. The updated code can be used to find secular equilbria for other systems too, e.g. it could be used to find the expected distribution of stellar mass black holes in disk at the galactic center.

A quasi-fully-interacting Neptune migration This is the most exploratory of the proposed projects. Recent work with neural networks have found that they can be used to integrate mathematically chaotic systems if they are fed data in Hamiltonian form. Choudhary et al. (2020) showed that a Hamiltonian neural network did significantly better in integrating chaotic trajectories in the Hénon-Heiles potential than a non-Hamiltonian neural network. Perhaps a Hamiltonian neural network could be used to integrate planetary scattering interactions with acceptable accuracy. A neural network like this could be used to mediate the interaction between Neptune and the planetesimals without needing to integrate the scattering interactions directly, possibly providing significant simulation speed-up. This might allow for large N Neptune migration simulations with fully interacting planetesimal disks.

Bibliography

Armitage, P. J. 2020, Astrophysics of Planet Formation, 2nd edn. (Cambridge University Press), doi: `10.1017/9781108344227`

Bailey, E., Batygin, K., & Brown, M. E. 2016, AJ, 152, 126, doi: `10.3847/0004-6256/152/5/126`

Bannister, M. T., Shankman, C., Volk, K., et al. 2017, AJ, 153, 262, doi: `10.3847/1538-3881/aa6db5`

Bannister, M. T., Gladman, B. J., Kavelaars, J. J., et al. 2018, ApJS, 236, 18, doi: `10.3847/1538-4365/aab77a`

Batygin, K., Adams, F. C., Brown, M. E., & Becker, J. C. 2019, Phys. Rep., 805, 1, doi: `10.1016/j.physrep.2019.01.009`

Batygin, K., & Brown, M. E. 2016a, AJ, 151, 22, doi: `10.3847/0004-6256/151/2/22`

—. 2016b, ApJ, 833, L3, doi: `10.3847/2041-8205/833/1/L3`

Batygin, K., Brown, M. E., & Fraser, W. C. 2011, ApJ, 738, 13, doi: `10.1088/0004-637X/738/1/13`

Batygin, K., & Morbidelli, A. 2017, AJ, 154, 229, doi: `10.3847/1538-3881/aa937c`

Becker, J. C., Khain, T., Hamilton, S. J., et al. 2018, AJ, 156, 81, doi: `10.3847/1538-3881/aad042`

Bernstein, G. M., Trilling, D. E., Allen, R. L., et al. 2004, AJ, 128, 1364, doi: `10.1086/422919`

Binney, J. 2008, Galactic dynamics / James Binney and Scott Tremaine., second edition. edn., Princeton series in astrophysics (Princeton, N.J: Princeton University Press)

Bitsch, B., Johansen, A., Lambrechts, M., & Morbidelli, A. 2015, A&A, 575, A28, doi: `10.1051/0004-6361/201424964`

Brasser, R., Duncan, M. J., & Levison, H. F. 2006, Icarus, 184, 59, doi: `10.1016/j.icarus.2006.04.010`

—. 2008, Icarus, 196, 274, doi: `10.1016/j.icarus.2008.02.016`

Brasser, R., Duncan, M. J., Levison, H. F., Schwamb, M. E., & Brown, M. E. 2012, Icarus, 217, 1, doi: `10.1016/j.icarus.2011.10.012`

Brasser, R., Morbidelli, A., Gomes, R., Tsiganis, K., & Levison, H. F. 2009, A&A, 507, 1053, doi: `10.1051/0004-6361/200912878`

Brown, M. E. 2017, AJ, 154, 65, doi: `10.3847/1538-3881/aa79f4`

Brown, M. E., & Batygin, K. 2019, AJ, 157, 62, doi: `10.3847/1538-3881/aaf051`

—. 2021, arXiv e-prints, arXiv:2108.09868. `https://arxiv.org/abs/2108.09868`

Brown, M. E., Trujillo, C., & Rabinowitz, D. 2004a, ApJ, 617, 645, doi: `10.1086/422095`

—. 2004b, ApJ, 617, 645, doi: `10.1086/422095`

Brucker, M. J., Grundy, W. M., Stansberry, J. A., et al. 2009, Icarus, 201, 284, doi: `10.1016/j.icarus.2008.12.040`

Buie, M. W., Grundy, W. M., Young, E. F., Young, L. A., & Stern, S. A. 2006, AJ, 132, 290, doi: `10.1086/504422`

Carrera, D., Johansen, A., & Davies, M. B. 2015, A&A, 579, A43, doi: `10.1051/0004-6361/201425120`

Chen, Y.-T., Lin, H. W., Holman, M. J., et al. 2016, ApJ, 827, L24, doi: `10.3847/2041-8205/827/2/L24`

Choudhary, A., Lindner, J. F., Holliday, E. G., et al. 2020, Phys. Rev. E, 101, 062207, doi: `10.1103/PhysRevE.101.062207`

Christy, J. W., & Harrington, R. S. 1978, AJ, 83, 1005, doi: `10.1086/112284`

Cohen, C. J., & Hubbard, E. C. 1965, AJ, 70, 10, doi: `10.1086/109674`

Dominik, C., Blum, J., Cuzzi, J. N., & Wurm, G. 2007, in Protostars and Planets V, ed. B. Reipurth, D. Jewitt, & K. Keil, 783

Dones, L., Brasser, R., Kaib, N., & Rickman, H. 2015, Space Sci. Rev., 197, 191, doi: `10.1007/s11214-015-0223-2`

Dubinski, J., Berentzen, I., & Shlosman, I. 2009, ApJ, 697, 293, doi: `10.1088/0004-637X/697/1/293`

Duncan, M., Quinn, T., & Tremaine, S. 1987, AJ, 94, 1330, doi: `10.1086/114571`

Duncan, M. J., & Levison, H. F. 1997, Science, 276, 1670, doi: `10.1126/science.276.5319.1670`

Fan, S., & Batygin, K. 2017, ApJ, 851, L37, doi: `10.3847/2041-8213/aa9f0b`

Fernandez, J. A., & Ip, W. H. 1984, Icarus, 58, 109, doi: `10.1016/0019-1035(84)90101-5`

Fleisig, J., Zderic, A., & Madigan, A.-M. 2020, AJ, 159, 20, doi: `10.3847/1538-3881/ab54c0`

Ford, E. B., & Chiang, E. I. 2007, ApJ, 661, 602, doi: `10.1086/513598`

Fouvry, J.-B., Dehnen, W., Tremaine, S., & Bar-Or, B. 2020, arXiv e-prints, arXiv:2011.01673. `https://arxiv.org/abs/2011.01673`

Friedli, D., & Pfenniger, D. 1990, in European Southern Observatory Conference and Workshop Proceedings, Vol. 35, European Southern Observatory Conference and Workshop Proceedings, 265

Gladman, B., Marsden, B. G., & Vanlaerhoven, C. 2008, in The Solar System Beyond Neptune, ed. M. A. Barucci, H. Boehnhardt, D. P. Cruikshank, A. Morbidelli, & R. Dotson (University of Arizona Press), 43

Gladman, B., Kavelaars, J., Petit, J. M., et al. 2009, ApJ, 697, L91, doi: `10.1088/0004-637X/697/2/L91`

Gomes, R., Levison, H. F., Tsiganis, K., & Morbidelli, A. 2005a, Nature, 435, 466, doi: `10.1038/nature03676`

Gomes, R. S. 2003, Icarus, 161, 404, doi: `10.1016/S0019-1035(02)00056-8`

Gomes, R. S., Gallardo, T., Fernández, J. A., & Brunini, A. 2005b, Celestial Mechanics and Dynamical Astronomy, 91, 109, doi: `10.1007/s10569-004-4623-y`

Gomes, R. S., Morbidelli, A., & Levison, H. F. 2004, Icarus, 170, 492, doi: `10.1016/j.icarus.2004.03.011`

Gruzinov, A., Levin, Y., & Zhu, J. 2020, ApJ, 905, 11, doi: `10.3847/1538-4357/abbfaa`

Gürkan, M. A., & Hopman, C. 2007, MNRAS, 379, 1083, doi: `10.1111/j.1365-2966.2007.11982.x`

Hahn, J. M., & Malhotra, R. 1999, AJ, 117, 3041, doi: `10.1086/300891`

—. 2005, AJ, 130, 2392, doi: `10.1086/452638`

Haisch, Karl E., J., Lada, E. A., & Lada, C. J. 2001, ApJ, 553, L153, doi: `10.1086/320685`

Hartmann. 2019, Geosciences, 9, 285, doi: `10.3390/geosciences9070285`

Hills, J. G. 1981, AJ, 86, 1730, doi: `10.1086/113058`

Hogg, D. W., Quinlan, G. D., & Tremaine, S. 1991, AJ, 101, 2274, doi: `10.1086/115849`

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111, doi: `10.3847/1538-4357/ab042c`

Izidoro, A., Morbidelli, A., Raymond, S. N., Hersant, F., & Pierens, A. 2015, A&A, 582, A99, doi: `10.1051/0004-6361/201425525`

James, R. A., & Sellwood, J. A. 1978, MNRAS, 182, 331, doi: `10.1093/mnras/182.3.331`

Jewitt, D., Luu, J., & Marsden, B. G. 1992, IAU Circ., 5611

Johansen, A., Oishi, J. S., Mac Low, M.-M., et al. 2007, Nature, 448, 1022, doi: `10.1038/nature06086`

Kaib, N. A., & Sheppard, S. S. 2016, AJ, 152, 133, doi: `10.3847/0004-6256/152/5/133`

Kaib, N. A., Pike, R., Lawler, S., et al. 2019, AJ, 158, 43, doi: `10.3847/1538-3881/ab2383`

Kavelaars, J. J., Lawler, S. M., Bannister, M. T., & Shankman, C. 2020, in The Trans-Neptunian Solar System, ed. D. Prialnik, M. A. Barucci, & L. Young (Elsevier), 61–77

Kazandjian, M. V., & Touma, J. R. 2013, MNRAS, 784, doi: `10.1093/mnras/stt074`

Kley, W. 2000, MNRAS, 313, L47, doi: `10.1046/j.1365-8711.2000.03495.x`

Kondratyev, B. P. 2014, MNRAS, 442, 1755, doi: `10.1093/mnras/stu841`

Kruijer, T. S., Burkhardt, C., Budde, G., & Kleine, T. 2017, Proceedings of the National Academy of Science, 114, 6712, doi: `10.1073/pnas.1704461114`

Kuiper, N. H. 1960, Indagationes Mathematicae (Proceedings), 63, 38 , doi: `https://doi.org/10.1016/S1385-7258(60)50006-0`

Lawler, S. M., Shankman, C., Kaib, N., et al. 2017, AJ, 153, 33, doi: `10.3847/1538-3881/153/1/33`

Lawler, S. M., Pike, R. E., Kaib, N., et al. 2019, AJ, 157, 253, doi: `10.3847/1538-3881/ab1c4c`

Levison, H. F., Duncan, M. J., Dones, L., & Gladman, B. J. 2006, Icarus, 184, 619, doi: `10.1016/j.icarus.2006.05.008`

Levison, H. F., & Morbidelli, A. 2003, Nature, 426, 419, doi: `10.1038/nature02120`

Levison, H. F., Morbidelli, A., Tsiganis, K., Nesvorný, D., & Gomes, R. 2011, AJ, 142, 152, doi: `10.1088/0004-6256/142/5/152`

Levison, H. F., Morbidelli, A., Van Laerhoven, C., Gomes, R., & Tsiganis, K. 2008, Icarus, 196, 258, doi: `10.1016/j.icarus.2007.11.035`

Levy, D. H. 1991, Clyde Tombaugh: discoverer of planet Pluto (The University of Arizona Press)

Li, G., & Adams, F. C. 2016, ApJ, 823, L3, doi: `10.3847/2041-8205/823/1/L3`

Lubow, S. H., Seibert, M., & Artymowicz, P. 1999, ApJ, 526, 1001, doi: `10.1086/308045`

Luu, J., Marsden, B. G., Jewitt, D., et al. 1997, Nature, 387, 573, doi: `10.1038/42413`

Lynden-Bell, D. 1979, MNRAS, 187, 101, doi: `10.1093/mnras/187.1.101`

MacGregor, M. A., Weinberger, A. J., Hughes, A. M., et al. 2019, ApJL, sub.

Madigan, A.-M., Halle, A., Moody, M., et al. 2018a, ApJ, 853, 141, doi: `10.3847/1538-4357/aaa714`

Madigan, A.-M., & McCourt, M. 2016, MNRAS, 457, L89, doi: `10.1093/mnrasl/slv203`

Madigan, A.-M., McCourt, M., & O'Leary, R. M. 2017, MNRAS, 465, 2310, doi: `10.1093/mnras/stw2815`

Madigan, A.-M., Zderic, A., McCourt, M., & Fleisig, J. 2018b, AJ, 156, 141, doi: `10.3847/1538-3881/aad95c`

Malhotra, R. 1993, Nature, 365, 819, doi: `10.1038/365819a0`

—. 1995, AJ, 110, 420, doi: `10.1086/117532`

—. 2019, Geoscience Letters, 6, 12, doi: `10.1186/s40562-019-0142-2`

Mardia, K., & Jupp, P. 2000, Directional Statistics (John Wiley & Sons), doi: `10.1002/9780470316979.ch11`

Masset, F., & Snellgrove, M. 2001, MNRAS, 320, L55, doi: `10.1046/j.1365-8711.2001.04159.x`

Meiss, J. D. 2007, Differential Dynamical Systems (Society for Industrial and Applied Mathematics), doi: `10.1137/1.9780898718232`. `https://epubs.siam.org/doi/abs/10.1137/1.9780898718232`

Merritt, D. 2013, Dynamics and Evolution of Galactic Nuclei (Princeton University Press)

Michael, G., Basilevsky, A., & Neukum, G. 2018, Icarus, 302, 80, doi: `10.1016/j.icarus.2017.10.046`

Milani, A., Nobili, A. M., & Carpino, M. 1989, Icarus, 82, 200, doi: `10.1016/0019-1035(89)90031-6`

Mizuno, H. 1980, Progress of Theoretical Physics, 64, 544, doi: `10.1143/PTP.64.544`

Mizuno, H., Nakazawa, K., & Hayashi, C. 1978, Progress of Theoretical Physics, 60, 699, doi: `10.1143/PTP.60.699`

Morbidelli, A. 2002, Modern celestial mechanics : aspects of solar system dynamics (Taylor & Francis)

Morbidelli, A., Brasser, R., Gomes, R., Levison, H. F., & Tsiganis, K. 2010, AJ, 140, 1391, doi: `10.1088/0004-6256/140/5/1391`

Morbidelli, A., Brasser, R., Tsiganis, K., Gomes, R., & Levison, H. F. 2009, A&A, 507, 1041, doi: `10.1051/0004-6361/200912876`

Morbidelli, A., & Crida, A. 2007, Icarus, 191, 158, doi: `10.1016/j.icarus.2007.04.001`

Morbidelli, A., Levison, H. F., Tsiganis, K., & Gomes, R. 2005, Nature, 435, 462, doi: `10.1038/nature03540`

Mustill, A. J., Raymond, S. N., & Davies, M. B. 2016, MNRAS, 460, L109, doi: `10.1093/mnrasl/slw075`

Napier, K. J., Gerdes, D. W., Lin, H. W., et al. 2021, The Planetary Science Journal, 2, 59, doi: `10.3847/PSJ/abe53e`

Nesvorný, D. 2015, AJ, 150, 73, doi: `10.1088/0004-6256/150/3/73`

—. 2018, ARA&A, 56, 137, doi: `10.1146/annurev-astro-081817-052028`

Nesvorný, D., & Morbidelli, A. 2012, AJ, 144, 117, doi: `10.1088/0004-6256/144/4/117`

Nesvorný, D., & Vokrouhlický, D. 2016, ApJ, 825, 94, doi: `10.3847/0004-637X/825/2/94`

Nesvorný, D., Vokrouhlický, D., & Morbidelli, A. 2007, AJ, 133, 1962, doi: `10.1086/512850`

Nesvorný, D., Vokrouhlický, D., & Roig, F. 2016, ApJ, 827, L35, doi: `10.3847/2041-8205/827/2/L35`

Nicholson, S. B., & Mayall, N. U. 1931, ApJ, 73, 1, doi: `10.1086/143288`

Noll, K. S., Grundy, W. M., Stephens, D. C., Levison, H. F., & Kern, S. D. 2008, Icarus, 194, 758, doi: `10.1016/j.icarus.2007.10.022`

Oldroyd, W. J., & Trujillo, C. A. 2021, AJ, 162, 39, doi: `10.3847/1538-3881/abfb6f`

Parker, R. J., Lichtenberg, T., & Quanz, S. P. 2017, MNRAS, 472, L75, doi: `10.1093/mnrasl/slx141`

Peng, S., & Batygin, K. 2020, ApJ, 898, L46, doi: `10.3847/2041-8213/aba68d`

Perri, F., & Cameron, A. G. W. 1974, Icarus, 22, 416, doi: `10.1016/0019-1035(74)90074-8`

Pierens, A., & Nelson, R. P. 2008, A&A, 482, 333, doi: `10.1051/0004-6361:20079062`

Pike, R. E., Kavelaars, J. J., Petit, J. M., et al. 2015, AJ, 149, 202, doi: `10.1088/0004-6256/149/6/202`

Pollack, J. B., Hubickyj, O., Bodenheimer, P., et al. 1996, Icarus, 124, 62, doi: `10.1006/icar.1996.0190`

Polyachenko, E. V. 2004, MNRAS, 348, 345, doi: `10.1111/j.1365-2966.2004.07390.x`

Polyachenko, E. V., & Shukhman, I. G. 2020, MNRAS, 498, 3368, doi: `10.1093/mnras/staa2573`

Raha, N., Sellwood, J. A., James, R. A., & Kahn, F. D. 1991, Nature, 352, 411, doi: `10.1038/352411a0`

Rauch, K. P., & Holman, M. 1999, AJ, 117, 1087, doi: `10.1086/300720`

Rauch, K. P., & Tremaine, S. 1996, New Astronomy, 1, 149

Rein, H., & Liu, S.-F. 2012, A&A, 537, A128, doi: `10.1051/0004-6361/201118085`

Rein, H., & Spiegel, D. S. 2015, MNRAS, 446, 1424, doi: `10.1093/mnras/stu2164`

Sefilian, A. A., & Touma, J. R. 2019, AJ, 157, 59, doi: `10.3847/1538-3881/aaf0fc`

Sellwood, J. A. 2012, ApJ, 751, 44, doi: `10.1088/0004-637X/751/1/44`

—. 2020, MNRAS, 492, 3103, doi: `10.1093/mnras/staa034`

Sellwood, J. A., & Wilkinson, A. 1993, Reports on Progress in Physics, 56, 173, doi: `10.1088/0034-4885/56/2/001`

Shankman, C., Kavelaars, J. J., Bannister, M. T., et al. 2017, AJ, 154, 50, doi: `10.3847/1538-3881/aa7aed`

Shannon, A., & Dawson, R. 2018, MNRAS, 480, 1870, doi: `10.1093/mnras/sty1930`

Sheppard, S. S., Trujillo, C. A., Tholen, D. J., & Kaib, N. 2019, AJ, 157, 139, doi: `10.3847/1538-3881/ab0895`

Slipher, V. M. 1930, JRASC, 24, 282

Standish, E. M. 1993, AJ, 105, 2000, doi: `10.1086/116575`

Takeuchi, T., & Lin, D. N. C. 2002, ApJ, 581, 1344, doi: `10.1086/344437`

Tamayo, D., Rein, H., Shi, P., & Hernandez, D. M. 2020, MNRAS, 491, 2885, doi: `10.1093/mnras/stz2870`

Tegler, S. C., & Romanishin, W. 2000, Nature, 407, 979, doi: `10.1038/35039572`

Tombaugh, C. W. 1946, Leaflet of the Astronomical Society of the Pacific, 5, 73

Toomre, A., & Kalnajs, A. J. 1991, in Dynamics of Disc Galaxies, ed. B. Sundelius, 341

Touma, J., Tremaine, S., & Kazandjian, M. 2019, Phys. Rev. Lett., 123, 021103, doi: `10.1103/PhysRevLett.123.021103`

Touma, J. R. 2002, MNRAS, 333, 583, doi: `10.1046/j.1365-8711.2002.05437.x`

Touma, J. R., Tremaine, S., & Kazandjian, M. V. 2009, MNRAS, 394, 1085, doi: `10.1111/j.1365-2966.2009.14409.x`

Tremaine, S. 2001, AJ, 121, 1776, doi: `10.1086/319398`

—. 2005, ApJ, 625, 143, doi: `10.1086/429713`

—. 2020a, MNRAS, 491, 1941, doi: `10.1093/mnras/stz3181`

—. 2020b, MNRAS, doi: `10.1093/mnras/staa420`

Trujillo, C. 2020, in The Trans-Neptunian Solar System, ed. D. Prialnik, M. A. Barucci, & L. Young (Elsevier), 79–105

Trujillo, C. A., & Sheppard, S. S. 2014, Nature, 507, 471, doi: `10.1038/nature13156`

Tsiganis, K., Gomes, R., Morbidelli, A., & Levison, H. F. 2005, Nature, 435, 459, doi: `10.1038/nature03539`

Walsh, K. J., Morbidelli, A., Raymond, S. N., O'Brien, D. P., & Mandell, A. M. 2011, Nature, 475, 206, doi: `10.1038/nature10201`

Weidenschilling, S. J. 1977, MNRAS, 180, 57, doi: `10.1093/mnras/180.2.57`

Weidenschilling, S. J., & Cuzzi, J. N. 1993, in Protostars and Planets III, ed. E. H. Levy & J. I. Lunine, 1031

Weinberg, M. D. 1998, MNRAS, 297, 101, doi: `10.1046/j.1365-8711.1998.01456.x`

Williams, J. P., & Cieza, L. A. 2011, ARA&A, 49, 67, doi: `10.1146/annurev-astro-081710-102548`

Youdin, A. N., & Goodman, J. 2005, ApJ, 620, 459, doi: `10.1086/426895`

Youdin, A. N., & Kenyon, S. J. 2013, in Planets, Stars and Stellar Systems. Volume 3: Solar and Stellar Planetary Systems, ed. T. D. Oswalt, L. M. French, & P. Kalas (Springer), 1

Zderic, A., Collier, A., Tiongco, M., & Madigan, A.-M. 2020, ApJ, 895, L27, doi: `10.3847/2041-8213/ab91a0`

Zderic, A., & Madigan, A.-M. 2020, AJ, 160, 50, doi: `10.3847/1538-3881/ab962f`

Zderic, A., Tiongco, M., Collier, A., et al. 2021, arXiv e-prints, arXiv:2106.09739. `https://arxiv.org/abs/2106.09739`

Appendix A

Element Conversions

The inclination instability is due primarily to forces acting *out* of the orbital plane. These forces exert torques on the orbit, $\boldsymbol{\tau} = \boldsymbol{r} \times \boldsymbol{f} = d\boldsymbol{j}/dt$, which result in rotations of the $\boldsymbol{j}$ vector. Whereas none of the standard Kepler elements remain constant during the instability, certain combinations of them do. Hence in Madigan & McCourt (2016) we introduced two new coordinates that represent rotation angles[1] about the semi-major axis of the orbit, $i_{\rm a}$, and about the semi-minor axis[2] of the orbit, $i_{\rm b}$. To complete a transformation from Kepler elements, we also introduced $i_{\rm e}$, the angle between the positive $\hat{\boldsymbol{x}}$-axis and the projection of the semi-major axis (or eccentricity vector) in the xy-plane. In the following conversions, we use $\hat{\boldsymbol{a}}$ and $\hat{\boldsymbol{e}}$ interchangeably as they are equivalent.

---

[1] The inclination angles $i_{\rm a}$ and $i_{\rm b}$ are defined in equation 2.1; with this definition, $i_{\rm a}$ and $i_{\rm b}$ are independent quantities, but they are not formally rotation angles. They are rotation angles to first order, however, so this difference does not affect our analysis in this paper.

[2] Strictly speaking, an orbit does not rotate about its minor axis, rather about a chord parallel to its minor axis which runs through the focus of the ellipse, the latus rectum.

Inclination angles i_a, i_b, and i_e ↔ orbit vectors $\hat{\boldsymbol{a}}$, $\hat{\boldsymbol{j}}$, and $\hat{\boldsymbol{b}}$ ($\equiv \hat{\boldsymbol{j}} \times \hat{\boldsymbol{e}}$):

$$\cos i = \cos i_\mathrm{a} \cos i_\mathrm{b} \left[1 - (\tan i_\mathrm{a} \tan i_\mathrm{b})^2\right]^{1/2} \tag{A.1a}$$

$$\hat{\boldsymbol{a}} = \begin{bmatrix} \cos i_\mathrm{b} \cos i_\mathrm{e} \\ \cos i_\mathrm{b} \sin i_\mathrm{e} \\ -\sin i_\mathrm{b} \end{bmatrix} \tag{A.1b}$$

$$\hat{\boldsymbol{b}} = \begin{bmatrix} \tan i_\mathrm{b} \sin i_\mathrm{a} \cos i_\mathrm{e} - (\cos i / \cos i_\mathrm{b}) \sin i_\mathrm{e} \\ \tan i_\mathrm{b} \sin i_\mathrm{a} \sin i_\mathrm{e} + (\cos i / \cos i_\mathrm{b}) \cos i_\mathrm{e} \\ \sin i_\mathrm{a} \end{bmatrix} \tag{A.1c}$$

$$\hat{\boldsymbol{j}} = \begin{bmatrix} \tan i_\mathrm{b} \cos i \cos i_\mathrm{e} + (\sin i_\mathrm{a} / \cos i_\mathrm{b}) \sin i_\mathrm{e} \\ \tan i_\mathrm{b} \cos i \sin i_\mathrm{e} - (\sin i_\mathrm{a} / \cos i_\mathrm{b}) \cos i_\mathrm{e} \\ \cos i \end{bmatrix} \tag{A.1d}$$

Kepler elements ↔ orbit vectors $\hat{\boldsymbol{e}}$, $\hat{\boldsymbol{j}}$, and $\hat{\boldsymbol{b}}$:

$$\hat{\boldsymbol{e}} = \begin{pmatrix} \cos\omega \cos\Omega - \sin\omega \cos i \sin\Omega \\ \cos\omega \sin\Omega + \sin\omega \cos i \cos\Omega \\ \sin\omega \sin i \end{pmatrix} \tag{A.2a}$$

$$\hat{\boldsymbol{j}} = \begin{pmatrix} \sin i \sin\Omega \\ -\sin i \cos\Omega \\ \cos i \end{pmatrix} \tag{A.2b}$$

$$\hat{\boldsymbol{b}} = \begin{pmatrix} -\sin\omega \cos\Omega - \cos\omega \cos i \sin\Omega \\ -\sin\omega \sin\Omega + \cos\omega \cos i \cos\Omega \\ \sin i \cos\omega \end{pmatrix} \tag{A.2c}$$

$$\boldsymbol{n} = \hat{\boldsymbol{z}} \times \boldsymbol{j} \tag{A.3a}$$

$$i = \arccos\left[\hat{j}_{\mathrm{z}}/j\right] \tag{A.3b}$$

$$\Omega = \arctan\left[\hat{n}_{\mathrm{y}}, \hat{n}_{\mathrm{x}}\right] \tag{A.3c}$$

$$\omega = \arctan\left[\mathrm{sgn}(\hat{e}_{\mathrm{z}})\left[1 - (\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{e}})^2\right]^{1/2}, \hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{e}}\right] \tag{A.3d}$$

Kepler elements ↔ inclination angles i_{a}, i_{b}, and i_{e} (note that these definitions are valid for prograde orbits only):

$$i_{\mathrm{a}} = \arcsin(\sin i \cos\omega) \tag{A.4a}$$

$$i_{\mathrm{b}} = \arcsin(-\sin i \sin\omega) \tag{A.4b}$$

$$i_{\mathrm{e}} = \Omega - \arctan\left[\cos\omega, \cos i \sin\omega\right] + \pi/2 \tag{A.4c}$$

$$i = \arccos\left[\cos i_{\mathrm{a}} \cos i_{\mathrm{b}} \left[1 - (\tan i_{\mathrm{a}} \tan i_{\mathrm{b}})^2\right]^{1/2}\right] \tag{A.5a}$$

$$\omega = \arctan\left[-\mathrm{sgn}(i_{\mathrm{b}})\sqrt{1 - \left(\frac{\sin i_{\mathrm{a}}}{\sin i}\right)^2}, \frac{\sin i_{\mathrm{a}}}{\sin i}\right] \tag{A.5b}$$

$$\Omega = i_{\mathrm{e}} + \arctan\left[\tan i_{\mathrm{b}} \cos i, \frac{\sin i_{\mathrm{a}}}{\cos i_{\mathrm{b}}}\right] \tag{A.5c}$$

Appendix B

Precession due to the disk mass

Here we describe apsidal precession of orbits due to the potentials of the pre-instability (relatively flat) disks presented in this paper. First, we look at numerical results from simulations showing how the rate of change of longitude of perihelion, $\dot{\varpi}$, varies with semi-major axis, eccentricity, and inclination for the compact and scattered disk (sd100) configurations (Figure B.1), and mass of the disk (Figure B.2). Second, we compare the potential of the disk in the compact configuration to an analytic expression derived in Kondratyev (2014). This is shown in Figure B.3. We discuss how this potential can be used to explain some features of the compact and scattered disk apsidal precession profiles.

In Figure B.1, we show median apsidal precession rate vs. median semi-major axis, eccentricity, and inclination for both the compact and 'sd100' orbital configurations with the same disk mass, $M_{\rm d} = 10^{-3}M$, and number of particles $N = 400$. The magnitude of the median precession rate in the compact configuration is approximately 10 times higher than the median precession rate in the scattered disk configuration, reflecting the lower mass density in the latter. In both cases, apsidal precession is retrograde and inclination i has minimal effect on $\dot{\varpi}_d$. In the compact configuration, the magnitude of the precession rate increases with semi-major axis. This relationship is roughly linear. In the scattered disk configuration, the magnitude of the precession rate increases from $a \sim 1$ to $a \sim 2$ after which it decreases out to $a \sim 10$. The sum of $\dot{\varpi}_d$ for the scattered disk and $\dot{\varpi}_{J2}$ results in a flattened precession profile (at least for $a > 2$), while the sum of $\dot{\varpi}_d$ for the compact configuration and $\dot{\varpi}_{J2}$ results in a steeper profile.

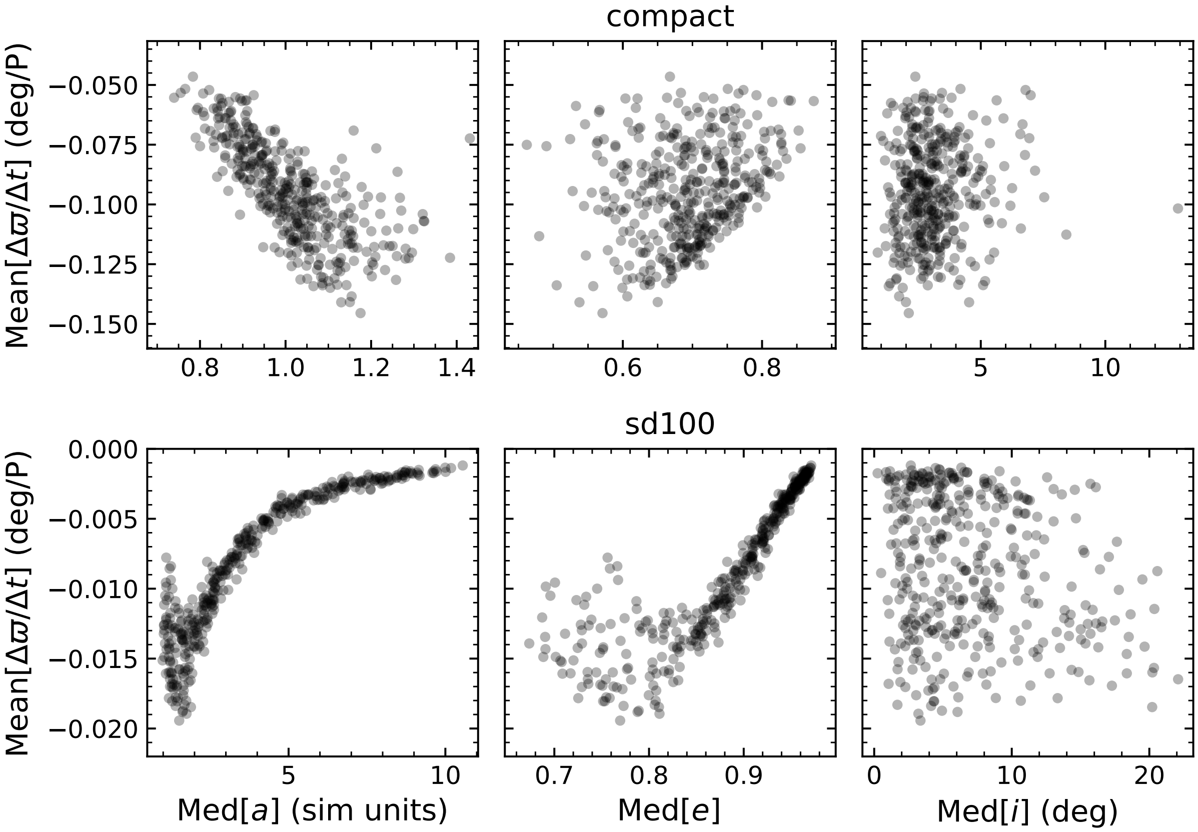

Figure B.1: Median apsidal precession rate for bodies in a compact configuration (top row) and scattered disk (bottom row) simulation. The apsidal precession rate is always retrograde in the disk and is not a strong function of inclination i. (Top Row) Semi-major axis a is the primary source of differential precession. Although the a dependence seems linear, the range is so narrow that the plot reflects a Taylor expansion of the true dependence at $a = 1$. Note that the apsidal precession dependence on a changes for $a \gtrsim 1.1$ from linearly decreasing to increasing. (Bottom Row) The apsidal precession rate is again retrograde and a strong function of semi-major axis. There are two distinct regimes. With a in $[1, 2]$, the retrograde apsidal precession rate increases in magnitude with a then in $[2, 10]$ decreases towards zero.

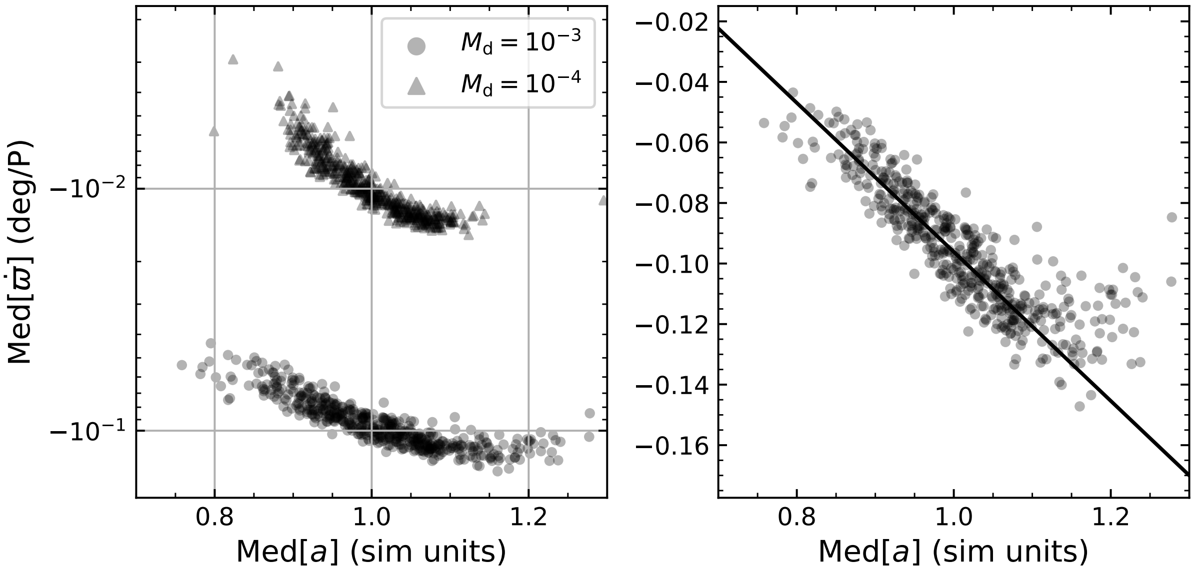

Figure B.2: Apsidal precession rate in the compact orbital configuration (pre-instability) with $N = 500$, as a function of semi-major axis. (Left) We show the precession rate for two different initial disk masses, M_d. The magnitude of the precession rate scales ∼linearly with the mass of the disk and the precession rate is retrograde. (Right) Apsidal precession rate for the case where $M_\mathrm{d} = 10^{-3}\,M$ with a linear y scale and a simple fit. The precession rate varies ∼linearly with a, with the outer edge of the disk precessing faster than the inner edge. On the edges of the disk, the apsidal precession rate doesn't quite follow this linear dependence. Note, only the particles with average semi-major axis in the range [0.9, 1.1] were factored into the shown fit. This linear functional form is conserved over changes in M_d, and the slope of the fit increases as disk mass is decreased.

Figure B.2 shows the disk-only precession rate of orbits in the compact configuration pre-instability with slightly higher N (500 vs. 400 for Figure B.1). In the left panel, we see the secular scaling of the precession rate ($\dot{\varpi}_d \propto M_\mathrm{d}$). In the right panel, we see the linear dependence of $\dot{\varpi}$ on a. This dependence is clearer here due to the increased N.

In the compact orbital configuration, the potential of the disk is well approximated by an expression derived in Kondratyev (2014). Kondratyev found the potential of an infinitely-populated axisymmetric disk of orbits with zero inclination and equal semi-major axis and eccentricity, a_d and e_d. This solid washer mass distribution is characterized by the inner and outer radii, $R_1 = a_d(1 - e_d)$ and $R_2 = a_d(1 + e_d)$. The radial mass density of the washer is the inverse of the radial Kepler velocity. The resulting potential in the plane of the washer is piece-wise, and expressed as integrals

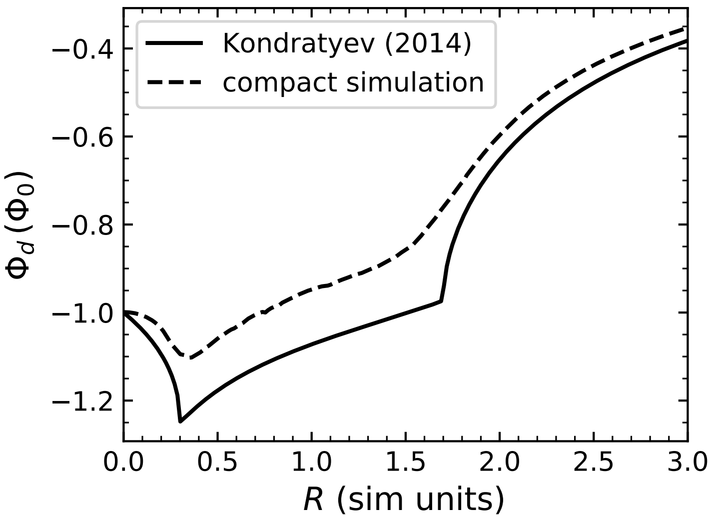

Figure B.3: Disk potential in the xy-plane as a function of cylindrical radius R normalized to the value of the disk potential at $R = 0$. The potential for a disk of eccentric orbits derived in Kondratyev (2014) is shown with $a = 1$ and $e = 0.7$ along with the potential of a simulated disk initialized in the compact configuration at $t = 100\,\mathrm{P}$. The disk potential calculated from the simulation is 'softened' in the sense that each orbit in the disk has been sampled at 40 evenly-spaced mean anomaly points. The general form of the potential is the same for both with the simulation potential lacking the cusp at $R = 0.3$ and $R = 1.7$ of the Kondratyev potential due to the small spread in a and e that the disk has naturally developed from two-body scattering.

over the mass distribution,

$$\Phi_d(R) = \frac{2\phi_0}{\pi^2}\begin{cases}\displaystyle\int_{R_1}^{R_2}\sigma(x)K\left(\frac{R}{x}\right)dx & \text{if } R < R_1\\ \displaystyle\frac{1}{R}\int_{R_1}^{R}x\sigma(x)K\left(\frac{x}{R}\right)dx + \int_{R}^{R_2}\sigma(x)K\left(\frac{R}{x}\right)dx & \text{if } R_1 < R < R_2\\ \displaystyle\frac{1}{R}\int_{R_1}^{R_2}x\sigma(x)K\left(\frac{x}{R}\right)dx & \text{if } R > R_2\end{cases} \tag{B.1}$$

where K is the complete elliptic integral of the first kind, $\phi_0 = -{}^{GM_\mathrm{d}}/_{a_d}$ is the potential at the origin, and $\sigma(x) = ((R_2 - x)(x - R_1))^{-1/2}$.

This potential is shown in Figure B.3 along with the potential of a compact configuration disk with $N = 400$ at $t = 100$ P. Each orbit was sampled at 40 equally-spaced mean anomalies and the potential was averaged along 10 different azimuthal lines in the xy-plane. Note that we expect the simulation potential to differ from the Kondratyev expression because of the (small) initial spread in a, e, and i. Despite the differences, the two potentials share the same bulk characteristics.

A formal expression for the precession rate of the orbits in the disk can be found using a Hamiltonian approach. Restricting ourselves to the xy-plane, the modified Delaunay coordinates in 2D are (Morbidelli, 2002),

$$\lambda = \mathcal{M} + \varpi \qquad I = \sqrt{\mu a}, \tag{B.2}$$

$$\varpi \qquad K = \sqrt{\mu a}(\sqrt{1-e^2} - 1). \tag{B.3}$$

We can then use Hamilton's equation's to get the time evolution of the apsidal angle (see Merritt (2013) for a similar derivation),

$$\dot{\varpi} = \frac{\partial H}{\partial K}, \tag{B.4}$$

where H is the Hamiltonian of the system,

$$H = H_{\text{kep}} + H_d, \tag{B.5}$$

where $H_{\text{kep}} \gg H_d$. The Keplerian Hamiltonian is

$$H_{\text{kep}} = -\frac{1}{2}\left(\frac{GM}{I}\right)^2, \tag{B.6}$$

and we average the disk potential over a Keplerian orbit, such that,

$$H_d = \overline{\Phi}_d, \tag{B.7}$$

where the over-line denotes an average over the unperturbed orbit and Φ_d is the potential of disk. Thus, the apsidal precession rate in the disk is given by,

$$\dot{\varpi}_d = \frac{\partial \overline{\Phi}_d}{\partial K} = -\sqrt{\frac{1-e^2}{\mu a e^2}}\frac{\partial \overline{\Phi}_d}{\partial e}, \tag{B.8}$$

where $\overline{\Phi}_d$ is given by the average of Kondratyev's potential, equation B.1, over the unperturbed orbit. The average over the orbit is

$$\overline{\Phi}_d = \frac{1}{2\pi}\int_0^{2\pi} dE\,(1 - e\cos E)\,\Phi_d(r), \tag{B.9}$$

where E is the eccentric anomaly. r and E are related by $r(E) = a(1 - e\cos E)$. We can pull the partial derivative inside the integral, use integration by parts, and the $r(E)$ expression to obtain,

$$\frac{\partial \overline{\Phi}_d}{\partial e} = \frac{1}{\pi}\int_{a(1-e)}^{a(1+e)} dr \frac{e^2 - (1 - r/a)}{\sqrt{e^2 - (1-r/a)^2}}\frac{d\Phi_d}{dr}, \tag{B.10}$$

$$\dot{\varpi}_d = -\frac{1}{\pi}\sqrt{\frac{1-e^2}{\mu a e^4}}\int_{a(1-e)}^{a(1+e)} dr\, \frac{e^2-(1-r/a)}{\sqrt{e^2-(1-r/a)^2}}\frac{d\Phi_d}{dr}. \tag{B.11}$$

There is no convenient expression for the derivative of Kondratyev's potential, so we will not attempt to find a closed form expression for the apsidal precession rate. However, we can use equation B.11 along with Figure B.3 to understand some basic features of the precession rate shown in figure B.2. Kondratyev's potential has cusps at R_1 and R_2. Further, Kondratyev's potential is concave down for all R ($d^2\Phi_d/dR^2 < 0$). The true disk potential does not have these cusps because the disk orbits have a range in a, e, and i. Instead the disk potential has a region where the potential is concave up ($d^2\Phi_d/dR^2 > 0$) near R_1 and R_2. In these regions $d\Phi_d/dR$ increases with R. The disk orbits are sufficiently eccentric that the orbit-averaged slope of the potential (i.e. the integral in equation B.11) is approximately given by the value of $d\Phi_d/dR$ at apocenter. The orbits with $a \lesssim 1.1$ have apocenters near R_2 in the region where the potential is concave up. The slope of the potential here is positive, and increasing with a (assuming $e \approx 0.7$). Thus, we would expect the apsidal precession rate of the orbits in the disk to be retrograde with magnitude increasing with a until $a \approx 1.1$. This is precisely what we see in Figure B.2.

Appendix C

Finding the Clustering Region

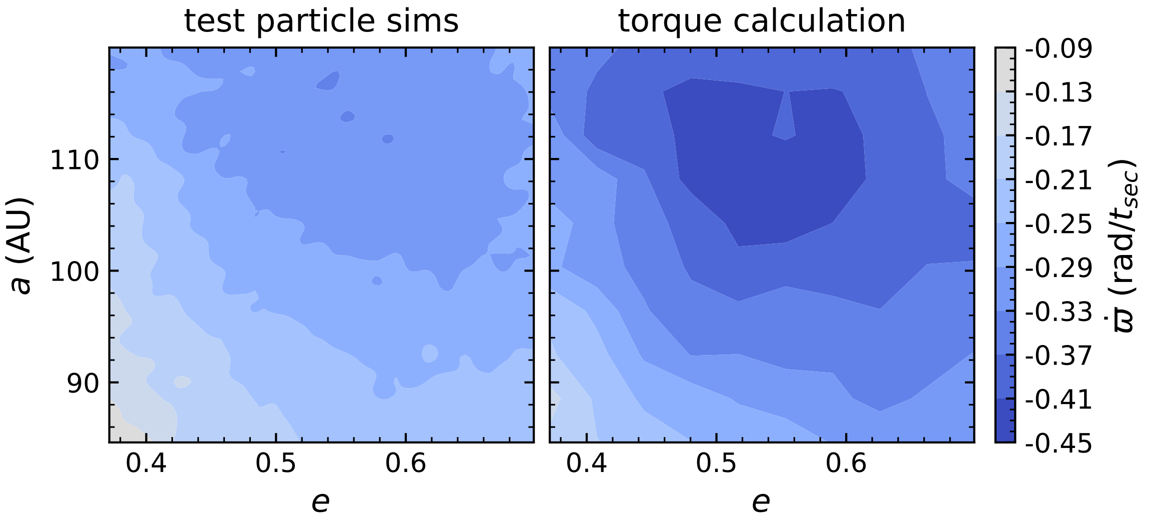

Figure C.1: Comparison of the two methods used to calculate the precession rate in the frozen disk potential. The left panel shows $d\varpi/dt$ calculated from a test particle simulation on a 100 by 100 e-a grid while the right panel shows $d\varpi/dt$ calculated with the torque method from Appendix C on a 10 x 10 grid. The two methods agree qualitatively, though the torque calculation gives faster precession rates than the test particle simulations. Both plots were made using data from an $N = 400$ compact configuration simulation without added J_2.

In Section 4.3.4.3, we show the clustering region within the disk using contour plots of $\dot{\varpi}$ in e-a space. In these plots, $\dot{\varpi}$ is calculated using test particle simulations described in the same section. Alternatively, we can calculate instantaneous precession rate of a test orbit can be directly from the torques and forces it experiences without integrating the orbit. In particular, the orbital precession rate can be computed from the torque and the time derivative of the eccentricity vector, viz.

$$\dot{\boldsymbol{e}} = \frac{\boldsymbol{f} \times \boldsymbol{j}}{GM} + \frac{\boldsymbol{v} \times \boldsymbol{\tau}}{GM}, \tag{C.1}$$

where $\boldsymbol{f}$ and $\boldsymbol{\tau}$ are the specific force and torque on the test orbit; $\boldsymbol{v}$ is the velocity; M is the central mass (see equation 1 in Madigan et al. 2017 and surrounding discussion).

In order to validate the results of § 4.3.4.3, we use equation (C.1) to calculate the precession rates of test orbits injected into N-body simulations. Specifically, we

(1) Discretize each orbit into one thousand, equal mass points, evenly-spaced in mean anomaly.

(2) Compute the total force and torque ($\boldsymbol{\tau}$) on each point along the test orbit from all of the disk orbits.

(3) Use equation (C.1) to determine $\dot{\boldsymbol{e}}$ at each point along the test orbit.

(4) Average $\boldsymbol{\tau}$ and $\dot{\boldsymbol{e}}$ over the test orbit.

(5) Check for convergence by repeating the above steps with a new set of discretrized point (evenly spaced in mean anomaly between existing ones). With one thousand points, the results are always converged within 10%. (And usually it is much better: for 90% of test orbits we obtain convergence within 3%.)

Similar methods have previously been used in studies of resonant relaxation in the Galactic center (e.g. Gürkan & Hopman 2007). To compare to the precession rate in the preceding section it is necessary to convert from $\dot{\boldsymbol{e}}$ and $\boldsymbol{\tau}$ to $\dot{\varpi} = \dot{\Omega} + \dot{\omega}$.[1] The Kepler orbital angles, Ω and ω, are related to the eccentricity and angular momentum vectors as follows

$$\begin{aligned}
&\Omega = \arctan\left(\frac{\hat{\boldsymbol{n}}_y}{\hat{\boldsymbol{n}}_x}\right), \\
&\omega = \arccos\left(\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{e}}\right), \\
&\boldsymbol{n} = \hat{\mathbf{z}} \times \hat{\mathbf{j}}.
\end{aligned} \tag{C.2}$$

---

[1] Note that $\dot{\varpi} = \dot{\Omega} - \dot{\omega}$ for retrograde orbits, which are not considered here.

Then $\dot{\Omega}$ and $\dot{\omega}$ can be approximated as

$$\begin{aligned}
&\dot{\Omega} \approx \frac{\Omega(\boldsymbol{j}(t+\delta t), \boldsymbol{e}(t+\delta t)) - \Omega(\boldsymbol{j}(t), \boldsymbol{e}(t))}{\delta t}, \\
&\dot{\omega} \approx \frac{\omega(\boldsymbol{j}(t+\delta t), \boldsymbol{e}(t+\delta t)) - \omega(\boldsymbol{j}(t), \boldsymbol{e}(t))}{\delta t}, \\
&\boldsymbol{j}(t+\delta t) \approx \boldsymbol{j}(t) + \boldsymbol{\tau}\delta t, \\
&\boldsymbol{e}(t+\delta t) \approx \boldsymbol{e}(t) + \boldsymbol{e'}\delta t,
\end{aligned} \tag{C.3}$$

where δt is a small time interval. Here we use $\delta t = 10^{-6}|\boldsymbol{j}|/|\boldsymbol{\tau}|$. We have verified that the results do not depend on δt.

The precession rate calculated using this torque method is compared to the precession rate from the test particle simulations in Figure C.1. Note that this plot shows results from a compact orbital configuration not a scattered disk orbital configuration. The compact configuration is an axisymmetric, nearly-flat disk of Keplerian orbits in which all bodies have identical eccentricities and nearly identical semi-major axes. This limited radial structure simplifies analysis. The two methods qualitatively agree, however, the test particle simulation method gives slower precession rates than the torque calculation. This is because two-body interactions in the test particle simulations (not accounted for in the torque calculation) weaken secular torques.

Appendix D

Measuring apsidal clustering

Here we demonstrate the connection between the components of the mean, normed eccentricity vector, e_R and e_z (see equation 4.5), and the Kepler elements argument of pericenter, ω, and longitude of pericenter, ϖ . In Figure D.1, we reproduce a result from Zderic et al. (2020) (Section 4.1) in which we demonstrate the appearance of apsidal clustering in a simulation of particles in a compact orbital configuration. Initially, the disk is axisymmetric; e_R is below the noise floor. As the top panel shows, the inclination instability begins at $t \lesssim t_{\rm sec}$ and saturates at $\sim 10\, t_{\rm sec}$. After the instability saturates and orbits apsidally precess back through the mid-plane ($e_z \approx 0$), we begin to see statistically significant e_R, indicating in-plane apsidal clustering. The right panels show histograms of ω and ϖ of all bodies at two times which are marked using colored-matching dashed lines. An e_R above the noise floor corresponds to ϖ-clustering but an e_z below the noise floor can still be ω-clustered.

Using the mean unit eccentricity vector instead of ϖ to measure apsidal clustering has two advantages: the mean unit eccentricity vector is 3D and can capture out-of-plane clustering, and statistical analyses on compound angles like ϖ can be misleading. We demonstrate the first point in Figure D.2. At $t \sim 10\, t_{\rm sec}$, the bodies in the disk have a uniform ϖ distribution suggesting that there is no apsidal clustering. However, the z component of the mean unit eccentricity vector is large (see top left panel of Figure D.1), indicating that the orbits apses are strongly clustered perpendicular to the plane.

Statistics on a compound angle can be misleading. If either ω or Ω is uniformly distributed,

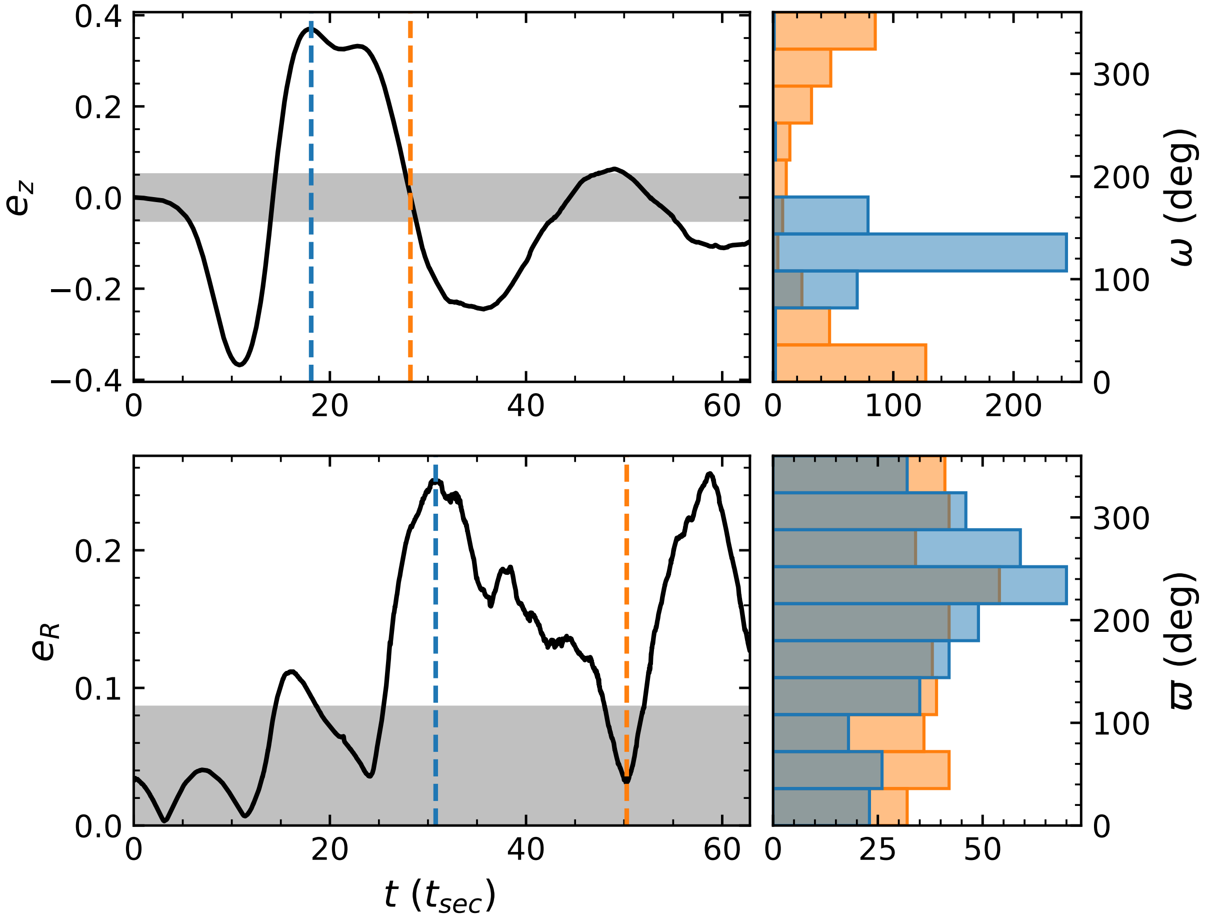

Figure D.1: Post-instability apsidal clustering in a compact orbital configuration. The top row shows e_z on the left and a histogram of each particle's ω on the right at the two times marked by colored vertical lines in the e_z plot. The bottom row shows e_R on the left and a histogram of each particle's ϖ on the right at the two times marked in the e_R plot. This shows that the peaks in e_R correspond to ϖ clustering, and that zero e_z can still be ω-clustered. The inclination instability saturates at around $10\,t_{\rm sec}$ in this simulation. This data comes from a 400 particle compact configuration simulation without added J_2.

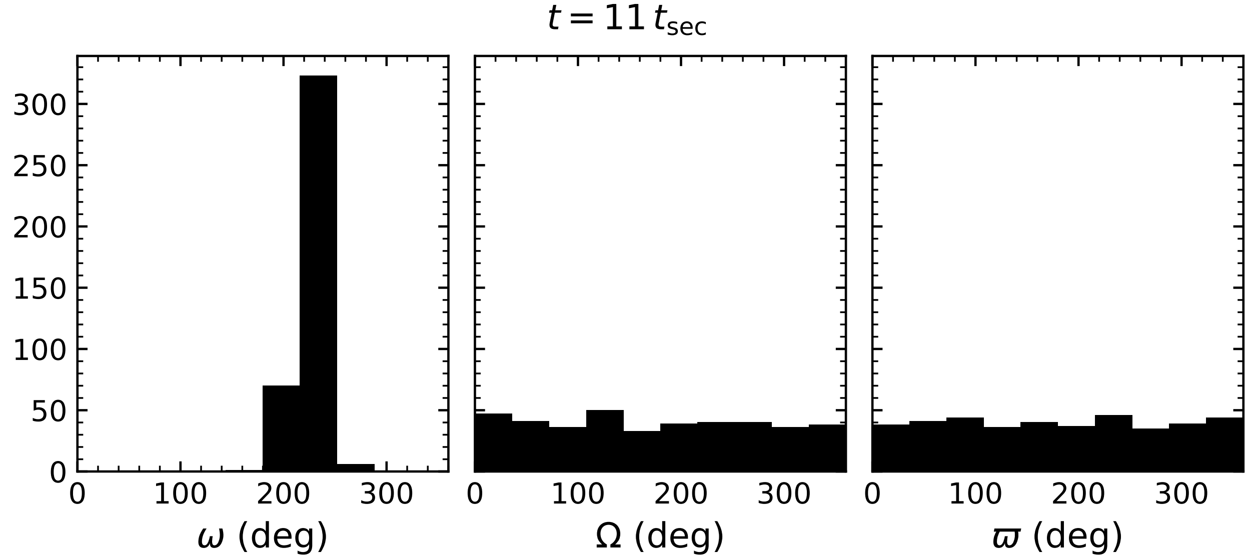

Figure D.2: Orbital angles of bodies from the compact configuration simulation shown in Figure D.1 at $t = 11\, t_{\rm sec}$. Arguments of pericenter ω are highly clustered, while Ω and ϖ show no clustering and are statistically consistent with a uniform distribution (Kuiper's test). This is what the bowl-shape driven by the inclination instability looks like in Kepler angles.

and ω and Ω are independent and have continuous distributions, then ϖ will also be uniformly distributed. In essence, ω or Ω, whichever is uniformly distributed, has the capacity to erase the others distribution in ϖ. This can be seen in Figure D.2. At $11\, t_{\rm sec}$, ω is highly clustered, nearly a delta function, while both Ω and ϖ are uniformly distributed.

We now prove this. We define the normalized distributions of ϖ, ω, and Ω as $f(\varpi)$, $g(\omega)$, and $h(\Omega)$, and recall that $\varpi = \omega + \Omega$. These distributions are periodic, e.g., $f(\varpi) = f(\varpi + 2\pi)$. The distribution of the sum of two continuous, independent random variables is given by the convolution of the two distributions,

$$f(\varpi) = \int_0^{2\pi} g\left(\varpi - \Omega\right) h\left(\Omega\right) d\Omega. \tag{D.1}$$

If the distribution of Ω is uniform, $h(\Omega) = {}^{1}\!/_{2\pi}$, then,

$$f(\varpi) = \frac{1}{2\pi} \int_0^{2\pi} g\left(\varpi - \Omega\right) d\Omega. \tag{D.2}$$

Switching back to ω,

$$
\begin{aligned}
f(\varpi) &= \frac{1}{2\pi}\int_{\varpi-2\pi}^{\varpi} g(\omega)d\omega, && \text{(D.3)}\\
&= \frac{1}{2\pi}\int_{\varpi}^{\varpi+2\pi} g(\omega-2\pi)d\omega, && \text{(D.4)}\\
&= \frac{1}{2\pi}\int_{\varpi}^{\varpi+2\pi} g(\omega)d\omega, && \text{(D.5)}\\
&= \frac{1}{2\pi}\int_{0}^{2\pi} g(\omega)d\omega, && \text{(D.6)}\\
&= \frac{1}{2\pi}, && \text{(D.7)}
\end{aligned}
$$

where we've used the normalization of $g(\omega)$,

$$1 = \int_0^{2\pi} g(\omega)d\omega,$$

and the identity,

$$\int_y^{y+2\pi} F(x)dx = \int_0^{2\pi} F(x)dx,$$

which holds for any $y \in \mathbb{R}$ and function $F(x)$ periodic with period 2π. This proof holds in the case where ω and/or Ω is uniformly distributed.