Astro2020 Science White Paper

Planet formation – The case for large efforts on the computational side

Thematic Areas:  ✔ Star and Planet Formation

Principal Author:
Name: Wladimir Lyra
Institution: California State University, Northridge
Email: wlyra@csun.edu
Phone: +1 818-677-7464

Co-authors: Thomas Haworth (Imperial College London), Bertram Bitsch (MPIA), Simon Casassus (Universidad de Chile), Nicolás Cuello (Pontificia Universidad Católica de Chile), Thayne Currie (NASA Ames), Andras Gáspár (University of Arizona), Hannah Jang-Condell (University of Wyoming), Hubert Klahr (MPIA), Nathan Leigh (AMNH/Universidad de Concepcion), Giuseppe Lodato (Università di Milano), Mordecai-Mark Mac Low (AMNH/CCA), Sarah Maddison (Swinburne University of Technology), George Mamatsashvili (Niels Bohr Institute), Colin McNally (Queen Mary University of London), Andrea Isella (Rice University), Sebastián Pérez (Universidad de Santiago de Chile), Luca Ricci (CSUN), Debanjan Sengupta (NASA Ames), Dimitris Stamatellos (University of Central Lancashire), Judit Szulágyi (University of Zürich), Richard Teague (University of Michigan), Neal Turner (JPL), Orkan Umurhan (NASA Ames), Jacob White (Konkoly Observatory), Al Wootten (University of Virginia).

Co-signers:
Felipe Alarcon (University of Michigan), Daniel Apai (University of Arizona), Amelia Bayo (Universidad de Valparaiso), Edwin Bergin (University of Michigan), Daniel Carrera (Penn State University), Ilse Cleeves (University of Virginia), Asantha Cooray (UC Irvine), Gregor Golabek (University of Bayreuth), Oliver Gressel (Niels Bohr Institute), Mark Gurwell (Harvard-Smithsonian Center for Astrophysics), Sebastiaan Krijt (University of Arizona), Cassandra Hall (University of Leicester), Ruobing Dong (University of Victoria), Fujun Du (University of Michigan), Ilaria Pascucci (LPL), John Ilee (University of Leeds), Andre Izidoro (Universidade Estadual Paulista), Jes Jorgensen (Niels Bohr Institute), Mihkel Kama (University of Cambridge), Dimitri Mawet (Caltech) Jinyoung Serena Kim (University of Arizona), David Leisawitz (NASA Goddard), Tim Lichtenberg (University of Oxford), Nienke van der Marel (NRC Herzberg), Margaret Meixner (STScI), John Monnier (University of Michigan), Giovanni Picogna (Ludwig-Maximilians-Universität München), Klaus Pontoppidan (STScI), Hsien Shang (ASIAA), Jane Simon (SwRI), David Wilner (Harvard-Smithsonian Center for Astrophysics),...
1 Introduction

Planet formation is simultaneously one of the oldest and one of the newest concerns of human inquiry. “How did the Earth come to be?” is a question that almost invariably appears in the cosmogonies of the ancients. In the late modern period it was understood that planets must form in disks of gas and dust around young stars, a process that modern astronomy has finally been able to observe in reasonable resolution and detail.

Developments in technical capabilities at optical/near-infrared and sub-millimeter wavelengths in the 2010–2019 decade, in particular the advent of the Atacama Large Millimeter Array (ALMA) and high-contrast adaptive optics have produced spatially resolved observations of the structures within protoplanetary disks (see reviews by Casassus 2016 and Sicilia-Aguilar et al. 2016). This has revealed a plethora of sub-structure, including rings and gaps (ALMA Partnership et al., 2015), spirals (Muto et al., 2012; Garufi et al., 2013; Benisty et al., 2015; Currie et al., 2014, 2015), warps (Casassus et al., 2015; Cuello et al., 2019), shadows (Stolker et al., 2016; Kama et al., 2016; Cuello et al., 2019), cavities (Andrews et al., 2011), and dust traps (van der Marel et al., 2013; Casassus et al., 2018). For a recent survey, see Garufi et al. (2018).

These processes are understood under the framework of disk-planet interaction, a process studied analytically and modeled numerically for over 40 years (Goldreich & Tremaine, 1979; Lin & Papaloizou, 1993; Kley & Nelson, 2012). Long a theoreticians’ game, the wealth of observational data has been allowing for increasingly stringent tests of the theoretical models. Although the observed structures qualitatively matched the general predictions from these models, one of the highlights of the decade of 2010–2019 in planet formation was the attempt to bring the hydrodynamical models to the level of quantitative agreement with the new detailed observations, a task that has been unexpectedly challenging. Modeling efforts are crucial to support the interpretation of direct imaging analyses, not just for potential detections but also to put meaningful upper limits on mass accretion rates and other physical quantities in current and future large–scale surveys.

The path towards a complete theory of planet formation remains elusive. While some processes have been definitely observed (such as rings, gaps, spirals, and vortices), and some indirectly observed, such as turbulence (Flaherty et al., 2015; Teague et al., 2018) other processes are less likely to be directly observed (such as planet formation via streaming instability and pebble accretion).

As even more detail is expected with ground based interferometers, extremely large telescopes (ELTs) and the James Webb Space Telescope (JWST) in the next decade (Ricci et al. 2018 and white papers by Isella, Currie, and Jang-Condell), a burning question, and the central point of this white paper, is what efforts on the computational side are required in the next decade to advance our theoretical understanding, explain the observational data, and guide new observations?

2 Overview of computational planet formation

The dynamical state of the protoplanetary disk is the fundamental canvas on which the planet formation narrative is etched. Understanding the evolution of disks, the structures that we are observing within them and the planet formation process presents a formidable challenge. Disks are composed of material spanning conditions ranging from cold, dense, and molecular, through to diffuse, hot, and ionized. Densities and temperatures vary by about 10 and 3 orders of magnitude, respectively. The gravitational potential from the parent star, self-gravity of the disk, hydrodynamic
torques in the disk, radiation from the parent star or other nearby stars, dust, cosmic rays, and non-ideal magnetohydrodynamics (MHD) all play important roles (see reviews by Dullemond & Monnier, 2010; Turner et al., 2014; Haworth et al., 2016). Furthermore, the dynamical evolution of dust grains of moderate size must be solved in addition to the gas dynamics (see reviews by Testi et al., 2014; Johansen et al., 2014). Disks are also not necessarily in a steady state, and can be subject to a range of instabilities, such as gravitational fragmentation (Durisen et al., 2007; Young & Clarke, 2015; Forgan et al., 2015; Meru, 2015), the streaming instability (Youdin & Goodman, 2005), Rossby wave instability (e.g. Lovelace et al., 1999; Tagger, 2001; Lyra et al., 2008, 2009), convective and vertical shear instabilities, which can form and grow vortex structures (see review by Lyra & Umurhan, 2018), the magnetorotational instability (MRI, Balbus & Hawley, 1991, 1998), as well as instabilities driven by non-ideal MHD (Kunz, 2008; Lesur et al., 2014).

Given the importance of these links, ultimately one wishes to identify which physical processes affect each other in a non-negligible fashion, and to model all of them simultaneously. The modelling of protoplanetary disks is therefore a daunting task. Each physical mechanism requires sufficient rigor and detail that modelling them constitutes an active field of protoplanetary disk research in its own right (for reviews of physical processes in protoplanetary disks, see e.g. Hartmann et al., 1998; Armitage, 2011; Williams & Cieza, 2011; Armitage, 2015). In practice, we have neither the numerical tools nor computational resources to achieve complete multi-physics modelling of protoplanetary disks in the immediate future. However, we can set out a roadmap towards this goal whilst outlining achievable milestones along the way.

2.1 Unsolved questions in computational planet formation

2.1.1 The nature of accretion

Despite major efforts, we still cannot answer a central question about accretion disks: how do they accrete? Since the landmark work of Shakura & Sunyaev (1973) the mechanism has been thought to be turbulence, with the MRI identified in the early 1990s as a plausible candidate to drive it. For disks hot enough to be fully ionized the MRI seems to be the most promising mechanism. However, even these disks have not conclusively been shown to be appropriately magnetized. The situation is even more dire in protoplanetary disks, as these disks are cold and poorly ionized, so the MRI is not a plausible mechanism, except for the hot inner disk (≤0.1 AU), and perhaps the low density atmosphere in the outer disk, if ionization by cosmic rays and stellar ultraviolet and X-rays is sufficiently high. The rest of the disk is poorly ionized, and thus non-ideal MHD has to be taken into account. This divides the disk into a small region in which MRI is active, and a large dead zone, itself split into regions dominated by Ohmic, Hall, and ambipolar diffusion. The current decade has seen the realization that angular momentum actually may be primarily transferred through magnetocentrifugal winds launched from upper atmosphere regions dominated by ambipolar diffusion (Bai & Stone, 2013; Bai, 2013; Gressel et al., 2015; Simon et al., 2018). Some disks are seen to have free-free emission (GM Aur, Ricci et al., 2010; Owen et al., 2013; Macías et al., 2016) but it is not clear whether the wind is photoevaporative (Owen et al., 2013; Canovas et al., 2018) or magnetocentrifugal (Banzatti et al., 2019). The Hall-dominated region is prone to a Hall-shear instability (Kunz, 2008), that would drive laminar accretion Lesur et al. (2014). Finally, the Ohmic region has three possible instabilities: the Zombie Vortex Instability (Marcus et al., 2015; Barranco et al., 2018), the Vertical Shear Instability (Nelson et al., 2013;
Flock et al., 2017), and the Convective Overstability (Klahr & Hubbard, 2014; Lyra, 2014), all generating large scale vortices and moderate levels of turbulence.

Observational tests to determine whether disks are magnetized, the prevalence of winds, and the level of turbulence throughout the whole disk column should provide the necessary information to decide between these models. Computational models with the necessary physics and resolution to solve for all these dynamical instabilities together is sorely missing but possible sometime early in the next decade. Such models will inform the observations of the main dynamical processes expected in disks and where to expect them.

### 2.2 Ab initio planet formation

Planetesimals and planetary embryos are $\leq 10^3$ km in size, suggesting that direct observations of their formation will remain difficult in the next decade. On this front, the best data comes from Solar System missions to primitive asteroids and comets such as provided by Rosetta, New Horizons, Hayabusa 2, OSIRIS-REx, and Lucy. Planetesimal formation by streaming instability (SI, Youdin & Goodman, 2005; Johansen et al., 2007), followed by pebble accretion to planetary embryos (Ormel & Klahr, 2010; Lambrechts & Johansen, 2012) stands as the best candidate for the formation of rocky and icy planets. Pebble trapping in vortices (Barge & Sommeria, 1995; Lyra et al., 2008, 2009) is also a plausible candidate, and can be tested observationally, as the spatial scales are resolved in some disks (Oph IRS 48, van der Marel et al. 2013, and MWC 758, Casassus et al. 2018). These hypotheses should be seen as complementary rather than mutually exclusive. Models still have problems making the SI work with solar metallicity (Carrera et al., 2015), needing at least twice the amount of metals (a small but numerically robust factor). In real disks, it is likely the streaming instability feeds off dust concentration in local pressure maxima such as rings or vortices (e.g. Lyra et al., 2008; Raettig et al., 2015). A detailed global model of streaming instability should be able to address these questions. What is needed is a model with sufficient dynamical range (1–100 AU), and enough resolution (at least 20 grid zones per SI wavelength), with full 360° azimuthal coverage in the MRI-dead zone in the presence of hydrodynamical instabilities. Other questions pertain to the angular momentum distribution of clumps formed by SI during gravitational collapse, as initially investigated by Nesvorný et al. (2010). In particular, the bizarre shape of Ultima Thule, as well as the frequency of contact binaries among small Solar System bodies are problems that have been emphasized to the community after the Rosetta and New Horizons encounters, and will take effort on the computational side in the next decade to explain.

### 2.3 Disk formation and early evolution

There is growing evidence for planet formation happening early. How then, are disks assembled and what progress towards dust processing/planet formation is made at this early stage? How does the star forming environment sculpt the disk parameters and can this imprint upon planetary populations? Unfortunately, there are too many free parameters to clearly connect the environment of molecular clouds to the initial conditions for circumplanetary disks. Although disk formation has been an active area of research (e.g. Inutsuka, 2012; Tsukamoto, 2016; Wurster & Li, 2018) we lack observational data on the first $10^4$ years of their life when evolution is fast.

A possible relevant process in this early stage is stellar flybys that affect not only disk truncation but also disk warping and morphology (Cuello et al., 2019). The younger the cluster, the more
likely the encounters, which can strongly affect where and when planetesimals form. This is also particularly important for the gravitational instability model of giant planet formation, which would have to happen early, while the disks are massive. While probably not the dominant mode of planet formation, some planetary systems are indeed extremely difficult to explain by core accretion.

### 2.3.1 Circumplanetary disks

As disk observations advance in accuracy and detail, the goal of imaging a planet in the making has been achieved, around the star PDS70 (Keppler et al., 2018). While proving a unequaled laboratory to study planet-disk interaction, it remains unclear whether the emission originates from the planet’s photosphere or from a circumplanetary disk (CPD) surrounding the protoplanet. CPDs are formed in high resolution numerical simulations around giant planets, and are a natural explanation for the regular satellite system of Jupiter. A gas CPD of about half the Hill radius should exist as long as the circumstellar disk (CSD) exists, because the CSD continuously feeds the CPD (Szulágyi et al., 2014; Fung & Chiang, 2016). Modeling observations of CPDs is challenging, requiring 3D radiation-hydrodynamics with adaptive mesh refinement. The post-processing is also challenging, since the temperatures and opacities remain poorly known. Another possible way to detect them is via kinematics: signposts of circumplanetary kinematics have been predicted from 3D hydrodynamic simulations of CPDs as localized deviations from Keplerian velocity in protoplanetary disks (Perez et al., 2015), and have been later detected in ALMA observations of the HD163296 disk (Pinte et al., 2018). Modeling of CPDs is an area that we recommend significant effort in the next decade.

### 3 Recommendations

#### 3.1 Approximations

Long a popular closure model for turbulence, the $\alpha$ viscosity model (Shakura & Sunyaev, 1973) has outgrown its usefulness. Its use in 2D and 3D modeling has become a hindrance to true progress in the field. Disks are not viscous, they are inviscid. It has been established that $\alpha$ viscosity, as a mean-field theory for disc turbulence, does not capture the spatial distribution and spectral properties now understood, particularly those new mechanisms of Sect. 2.1.1. As such, models with $\alpha$ viscosity result in incorrect small scale properties. This has a major impact on structure formation, the location of planetesimal formation, the shape of planet-induced gaps and vortices. Even the planet-disc interactions and planet migration torques can be impacted. Thus, a major goal for the next decade is to understand the small-scale structure well enough to develop subgrid models suitable for use in large eddy simulations. This will allow the large-scale flows to be modeled without extraordinary computational resources.

Another point (a dismal one to make as late as 2019), is the need to move beyond isothermal models. Already in the decade of 2000–2009 major advances were made in planet formation by relaxing this crippling approximation (e.g., outward migration, Paardekooper & Mellema, 2006). The same happened in 2010–2019. For the observations, especially because the disk scale height controls the scattered intensity in infrared/optical, it is crucial that models that intend to reproduce observations do so both in 3D and without using the isothermal approximation (Fung & Dong,
There is little value in a model, even if it does a reasonable job at reproducing observations, if it is intrinsically flawed. The adoption of $\alpha$ viscosity and isothermal thermodynamics are approximations that we can and should get away from during the next decade.

3.2 Evolutionary models and population synthesis

Evolutionary models must run for the entire lifetime of the disk ($\sim$10 Myr). Such models are currently used for population synthesis, with the goal of comparing results to the observed distribution of exoplanets (e.g. Mordasini et al., 2009a,b; Ida & Lin, 2010; Dittkrist et al., 2014; Mordasini, 2018; Ndugu et al., 2018). These models have also used N-body for the interacting planets (Horn et al., 2012; Coleman & Nelson, 2016; Lambrechts et al., 2019; Izidoro et al., 2019; Bitsch et al., 2019), but with decoupled 1D hydrodynamics for the disk, using formulae for planet-disk interaction derived from hydrodynamical simulations with more degrees of freedom (Paardekooper et al., 2011). The pitfalls of such an approach are many and dangerous.

We identify these models of planetary system evolution as a field would benefit from a significant overhaul. A 2D hydrodynamic model of a disc and planets over 10 Myr timescales is already possible with current computational resources if the inner boundary is at a few AU. However, this has been done with either massively parallel use of CPUs, or only moderately parallel graphical processing units (GPUs) (Regály et al., 2012; Muley et al., 2019). Pérez et al. (2018) report 20 orbits/hr in 3D with 4 GPUs using direct GPU-GPU communication. Extrapolating the speedup in GPU-based computation of a order of magnitude in the last decade is continued into the next, a new generation of evolutionary models directly modelling disc-planet interactions over the full lifetime and spatial scales of the disk may be built.

These models can then be used to make significant advancements such as self-consistently including photoevaporation clearing of the disc gas, or used for the basis of a new generation of high-fidelity population synthesis models. In particular, this second goal, requiring coverage of a vast parameter space, may be enabled by the increasing speed and decreasing cost of GPU-type computing devices in the future.

4 Conclusion

We have identified major fields of interest in computational planet formation: the nature of accretion, ab initio planet formation, early evolution, and circumplanetary disks. We recommend that modelers relax the approximations of $\alpha$ viscosity and isothermal equations of state, on the grounds that these models use flawed assumptions, even if they give good visual qualitative agreement with observations. We similarly recommend that population synthesis move away from 1D hydrodynamics. The computational resources to reach these goals should be developed during the next decade, through improvements in algorithms and the hardware for hybrid CPU/GPU clusters. Advances in computational planet formation, coupled with high angular resolution and great line sensitivity in ground based interferometers, ELTs and JWST, should allow for large strides in the field in the next decade.
References

Armitage, P. J. 2011, ARAA, 49, 195
—. 1998, Reviews of Modern Physics, 70, 1
Casassus, S. 2016, PASA, 33, e013
Dullemond, C. P., & Monnier, J. D. 2010, ARAA, 48, 205
Inutsuka, S.-i. 2012, Progress of Theoretical and Experimental Physics, 2012, 01A307
Kley, W., & Nelson, R. P. 2012, ARAA, 50, 211
Tsukamoto, Y. 2016, PASA, 33, e010