UNIVERSITY OF CALIFORNIA
SANTA CRUZ

GAS GIANTS IN EXTREME ENVIRONMENTS

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

ASTRONOMY AND ASTROPHYSICS

by

Renata Frelikh

September 2023

The Dissertation of Renata Frelikh
is approved:

Professor Ruth-Murray-Clay, Chair

Professor Jonathan Fortney

Professor Enrico Ramirez-Ruiz

Professor Leslie Rogers

Peter Biehl
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

Gas Giants in Extreme Environments

by

Renata Frelikh

In this thesis, I will explore three ideas that relate to giant planet formation, an incredibly dynamic process that affects not only the orbital configurations, but also the atmospheric structure and processes driven by the often extreme environments in which these planets are embedded: (1) The formation of our solar system's ice giants may have involved significant migration due to a disk dynamical instability, as well as concurrent accretion of solids and gas from the protoplanetary disk, that may have consequences for their current atmospheric composition and evolution. (2) In exoplanetary systems, gas giants have been discovered on orbits that are extremely eccentric, and in the inner regions corresponding to the traditionally rocky planet region of our own solar system. These exotic systems may have undergone a phase of giant impacts, analogous to the phase that shaped our solar system's inner planet region and resulted in the Moon-forming impact. Such a phase produces outcomes that reproduce interesting features of the population: the correlation of mass with eccentricity, and eccentricity with stellar metallicity. (3) Finally, the atmospheres of the close-in planets are bathed in stellar irradiation. Photochemical reactions occur as the upper layers of the atmospheres are irradiated, and the massive amounts of deposited energy drive hydrodynamic winds from the planets. This results in mass loss that can not only be observed in the cases

of giants, but can shape the global structure of smaller planets, leaving an imprint on entire populations of these planets.

To Matilda Sophie and James Ian.

Acknowledgments

I thank my advisor Ruth for her mentorship and unwavering support. I want to thank her for her infinite patience and wisdom. Her curiosity and insight into any subject is inspiring. She is truly a great scientist and person. I want to thank Enrico for believing in me. In one of our conversations he told me I could have it all, and those words kept me going through it all. Thank you to everyone in my research group, Mickey, John, Tyler, Diana, Maddy, Arcelia. Thank you to my collaborators over the years, and James Owen and his research group for being so helpful in understanding atmospheric escape. Thank you Matilda for being my sunshine. Thank you James for keeping me company while I write this dissertation, your smiles always make me so happy. I love you both more than I could ever describe. And my parents, Lidia and Eduard, thank you for all of your help.

The text of this dissertation includes reprints of the following previously published material, with permission of the listed coauthors, led by myself.

Chapter 2 was previously published in the literature as Frelikh & Murray-Clay [25]. All the figures and text are my own, with substantial feedback from Ruth Murray-Clay, who supervised the research and edited the paper.

Chapter 3 was previously published in the literature as Frelikh et al. [26]. All the figures and text are my own. I ran the set of simulations presented in the paper and analyzed the results. Ruth Murray-Clay supervised the project and edited the paper. Hyerin Jang, then an undergraduate at UC Santa Cruz, worked with me

to produce excellent initial results that helped conceptualize the project. Cristobal Petrovich provided very helpful feedback and advice on the project.

Thank you Ruth Murray-Clay for supporting my work through NSF CAREER grant AST-1555385, and through Heising-Simons award number 2021-3126. Thank you Enrico Ramirez-Ruiz and the Heising-Simons foundation for your support of my research.

Chapter 1

Introduction

Giant planet atmospheres are a significant component of their overall structure, in contrast to the relatively tenuous atmospheres of the rocky planets. The solar system's gas giants, Jupiter and Saturn, as well as the Jupiter-like exoplanets, such as the hot and warm Jupiters discovered by radial-velocity surveys in exoplanetary systems, are mostly atmospheres of H and He accreted from the primordial protoplanetary disk in the process of runaway gas accretion. As such, their composition will roughly match solar. The ice giants, such as Uranus and Neptune in our Solar System, and Neptune analogs in exoplanetary systems, possess of order 10 % atmospheres by mass. Chapter 2 focuses on how these planets formed, specifically the question: are the intermediate-sized atmospheres of the ice giants consistent with the classical planet formation scenario via planetesimal accretion, while pebbles drift in with the protoplanetary disk gas from the outer disk? Can the ice giants avoid runaway gas accretion into gas giants - is there a way to limit their growth without appealing to a fine-tuned argument?

In Chapter 3, we focus on stage of formation of the gas giants in exoplanetary systems that occurs after the gas disk has dissipated. An interesting feature in the population of radial-velocity warm Jupiters is that mass and eccentricity are correlated. This is counterintuitive from what we would expect as a traditional outcome of planet-planet scattering. When we view the sample as a population coming from a distribution of disks with varying initial total planet mass, our n-body simulations produce the highest-eccentricity planets in the highest-mass disks, reproducing the trend from the observations. Furthermore, as heavy element mass and stellar metallicity may be correllated, using the disk mass as a proxy for stellar metallicity in our simulations allows us to reproduce the observed correlation between stellar metallicity and eccentricity. If the warm Jupiter population is an outcome of such a giant impacts phase, it will leave in imprint on the distribution of giant planets with semimajor axis, with the collisional growth products being more common in the region of collisional growth, $\sim 1 - 8$ au, and this will be tested with direct imaging surveys.

Chapter 4 focuses on the processes involved in shaping the atmospheres of the giant planets on close-in orbits, specifically extreme-ultraviolet-driven atmospheric escape. Bathed in ionizing radiation, the gas in the thermospheres of these planets ionizes, and the residual energy is deposited as heat. This drives winds from the planets, reaching velocities of km s^{-1} to $\sim$ tens of km s^{-1}. However, the ionizing radiation also drives photochemical processes that determine the chemical composition of the atmosphere at different altitudes. The molecular hydrogen layer at the base of the thermospheric region also contains H_3^+, an important infrared coolant for the gas. The atomic hydrogen in the

thermally dissociated layer above is a source of significant Lyman-α cooling, an often even stronger radiative coolant than the H_3^+. The energy radiated away is not available to power escape, reducing the mass-loss rate substantially. How the energy budget of the planet is allocated between radiative processes and driving a wind determines the efficiency of atmospheric escape, and allows us to predict under which conditions escape can occur.

Chapter 2

Formation of Uranus and Neptune:

Fine-tuning in Core Accretion

Abstract

Gas Giants in Extreme Environments

by

Renata Frelikh

Uranus and Neptune are ice giants with $\sim 15\%$ atmospheres by mass, placing them in an intermediate category between rocky planets and gas giants. These atmospheres are too massive to have been primarily outgassed, yet they never underwent runaway gas accretion. The ice giants never reached critical core mass ($M_{\rm crit}$) in a full gas disk, yet their cores are $\gtrsim M_{\rm crit}$, suggesting that their envelopes were mainly accreted at the end of the disk lifetime. Pebble accretion calls into question traditional slow atmospheric growth during this phase. We show that the full-sized ice giants predominantly accreted gas from a disk depleted by at least a factor of ~ 100. Such a disk dissipates in $\lesssim 10^5$ years. Why would both cores stay sub-critical for the entire $\sim$ Myr disk lifetime, only to reach $M_{\rm crit}$ in the final 10^5 years? This is fine tuned. Ice giants in the outer disk have atmospheric mass fractions comparable to the disk gas-to-solid ratio during the bulk of their gas accretion. This point in disk evolution coincides with a dynamical upheaval: the gas loses its ability to efficiently damp the cores' random velocities, allowing them to be gravitationally excited by Jupiter and Saturn. We suggest that the ice giants' cores began growing on closer-in orbits (staying sub-critical), and migrated out during this dynamical instability. There, their orbits circularized after accreting much of their mass in solids. Finally, they accreted their envelopes from a depleted nebula, where the

sparseness of feeding zone gas prevented runaway.

2.1 Introduction

Uranus and Neptune, ice giants with masses 14.5 $M_{\rm E}$ and 17 $M_{\rm E}$, respectively, consist of cores of rock and ice, surrounded by hydrogen and helium envelopes. Models of the planets' interiors indicate that these atmospheres comprise about 12-14% of Uranus and 14-16% of Neptune by mass [123]. This places the ice giants in an intermediate category between the rocky planets, with atmospheres likely dominated by outgassing, and the gas giants, whose atmospheres are a result of runaway gas accretion from the nebula.

A simple scaling argument suggests that the atmospheres of Uranus and Neptune were primarily accreted from the nebula, rather than outgassed from the cores. Suppose that the hydrogen was originally bound in the core as H_2O. Releasing 2 $M_{\rm E}$ in hydrogen would require a starting mass of 18 $M_{\rm E}$ of water, which is comparable to the entire mass of the core. This is unfeasible, since after outgassing the core would almost entirely consist of oxygen. (D. Stevenson - private communication). In principle, 2 $M_{\rm E}$ of hydrogen can be released from 11 $M_{\rm E}$ of NH_3 or 8 $M_{\rm E}$ of CH_4. However, observations show that the carbon in protoplanetary disks is primarily bound in CO and CO_2 [124, 69], and that only 10-12% of the nitrogen is in ammonia ices [107]. This can be understood by a theoretical argument: the disk abundances are inherited from the interstellar medium, where the carbon is in CO_2, and CO [110], and the nitrogen is in N_2. Though protoplanetary disk conditions favor CH_4 and NH_3, the reaction timescales for the conversion of nitrogen to ammonia and carbon monoxide to methane are too

long compared to the disk lifetime [119].

For our argument, we are interested in maximizing the potential contributions of CH_4 and NH_3 to the core composition at the ice giants' disk locations, which are comparable to or beyond the snowline locations of CH_4 ($\simeq$ 4 AU) and NH_3 ($\simeq$ 23 AU) [127]. Since the presence of large amounts of pure CH_4 and NH_3 in the disk is unlikely, we can consider an H_2O planetary core with CH_4 and NH_3 mixed in at abundances corresponding to the maximum observed abundances in a disk. We define $n_\times$ as the number density of species X. Using the maximum observed abundances of $n_{CH_4,max} = 0.13 n_{H_2O}$ [124], and $n_{NH_3,max} = 0.15 n_{H_2O}$ [107], outgassing 2 M_E of H_2 requires a starting mass of $\sim$ 17 M_E of a mixture of H_2O, CH_4, and NH_3. Forming a $\sim$ 2 M_E atmosphere would require releasing all of the hydrogen from the core, leaving behind a $\sim$ 15 M_E C/N/O core. Such an outcome is implausible: the ice giants' cores almost certainly contain hydrogen. A 100% dissociation efficiency is highly unlikely; moreover, most of the hydrogen would remain mixed in the core, as it would not phase separate [e.g. 135]. We conclude that the hydrogen envelopes of Uranus and Neptune cannot be outgassed.

The standard scenario for the formation of planets with a significant gas component is core growth by planetesimal accretion, followed by gas accretion [130] once a critical core mass is reached. For the purposes of this paper, we define the critical core mass (M_{crit}) to be the mass required for the protoplanetary core to be able to accrete a gas envelope with a mass M_{gas} comparable to the core mass by a given time (generally, the gas disk lifetime). Once $M_{core} = M_{gas}$, the atmosphere starts to contribute substan-

tially to the gravitational potential of the planet, triggering runaway gas accretion by nucleated instability [e.g. 130, 131]. We note that $M_{\rm crit}$ is defined in two ways in the literature: the steady state solution described by Rafikov [70], and the time evolving solution [130]. The Rafikov [70] solution considers a core with an accretion luminosity that is high enough that its gas envelope is able to reach equilibrium on a short timescale compared to the disk lifetime. In this case, the envelope achieves a pseudo-steady state, with radiative losses balancing heating by accreting planetesimals. So, $M_{\rm crit}$ is the core mass that is able to hold onto its own mass in gas, while continually being injected with energy from accretion. The planet undergoes runaway gas accretion if the core is able to grow to $M_{\rm crit}$ in the lifetime of the disk. On the other hand, the Pollack et al. [130] solution requires accretion to halt (or slow substantially) once the supply of planetesimals is depleted, allowing the envelope to accumulate gas as it gradually cools and undergoes Kelvin-Helmholtz (KH) contraction. If the core is able to accumulate its own mass in gas within the lifetime of the full disk, it undergoes runaway gas accretion and becomes a gas giant, with $M_{\rm crit}$ being the core mass at the point when runaway is triggered. When not otherwise specified, below we use $M_{\rm crit}$ generically to mean the mass at which a core is able to accrete $M_{\rm gas} = M_{\rm core}$ by any mechanism.

What separates the ice giants from the gas giants is their formation history: during core growth, the former never exceeded $M_{\rm crit}$ in a full gas disk. While sub-critical, the relatively small cores were unable to accrete much gas, with the core mass as a conservative upper limit to the mass of the atmosphere at any given time before the planets reached critical core mass. On the other hand, if given sufficient time to

accumulate gas, a super-critical core is able to accrete most of the gas in its feeding zone, which in a full disk is typically sufficient to produce a gas giant.

The cores of Uranus and Neptune exceed reasonable estimates of the critical core mass. One can imagine several scenarios to prevent runaway. A high enough planetesimal accretion rate can cause the critical core mass to exceed the current core masses of the ice giants (Section 2). However, the high accretion rate must be sustained for the bulk of the gas disk lifetime, which is unlikely. As an alternative, the standard explanation of Pollack et al. [130] avoids fine tuning by proposing the existence of a stage of slow gas and planetesimal accretion that persists for a couple Myr. In this stage, the composition of the ice giants remains similar to their present-day makeup for a significant fraction of the gas disk lifetime. However, pebble accretion [117, 125] makes this explanation newly problematic, as we will show in section 2.2. Finally, our ice giants could have reached their final core masses in a largely depleted gas disk, when there was just enough gas left in their feeding zones to accrete their intermediate-sized atmospheres. We show in Section 3 that this corresponds to a disk mass of at most $\lesssim 0.01$ times its original mass, and show in Section 4 that this depleted disk gas dissipates fairly quickly ($\lesssim 10^5$ years).

Once the gas is depleted to $\simeq 0.01$ of the full gas disk, the mass of gas in the disk becomes comparable to the mass of solids. A solid-to-gas ratio of order unity is a known factor for dynamical instability. This dynamical instability is triggered when the gas becomes so depleted that it can no longer effectively damp the protoplanetary cores' random velocities. In Section 5, we discuss a possible solution for the fine tuning

problem: there is a dynamical reason for the ice giants to accrete most of their mass in gas at the end of the gas disk lifetime.

2.2 Avoiding runaway

We now discuss the standard ways to avoid runaway. In 2.1, we discuss the possibility that the growing protoplanetary core could have accreted planetesimals so quickly that the entropy deposited into the envelope was enough to prevent collapse. In other words, $M_{\rm crit}$ was larger than the current mass of the ice giants throughout the gas disk lifetime. In this fast accretion regime, the atmosphere satisfies a static solution, where the core and the envelope have reached a state of thermal equilibrium, with the energy in balancing the energy out [70]. In 2.2, we consider the classical argument of Pollack et al. [130], in which the ice giants' core growth effectively halts at a mass lower than would be necessary to quickly accrete an envelope. In this dynamic solution, which differs from the scenario discussed in 2.1 in the sense that the atmosphere never reaches a steady state, the atmosphere instead progresses through a period of slow gas accretion limited by the cooling rate.

2.2.1 Can the cores stay sub-critical due to fast planetesimal accretion?

As planetesimals are accreted onto a growing core, their gravitational potential energy is released near the surface with an accretion luminosity $L = GM_p\dot{M}/R$. This increases the temperature at the base of the atmosphere and puffs it up, thus decreasing

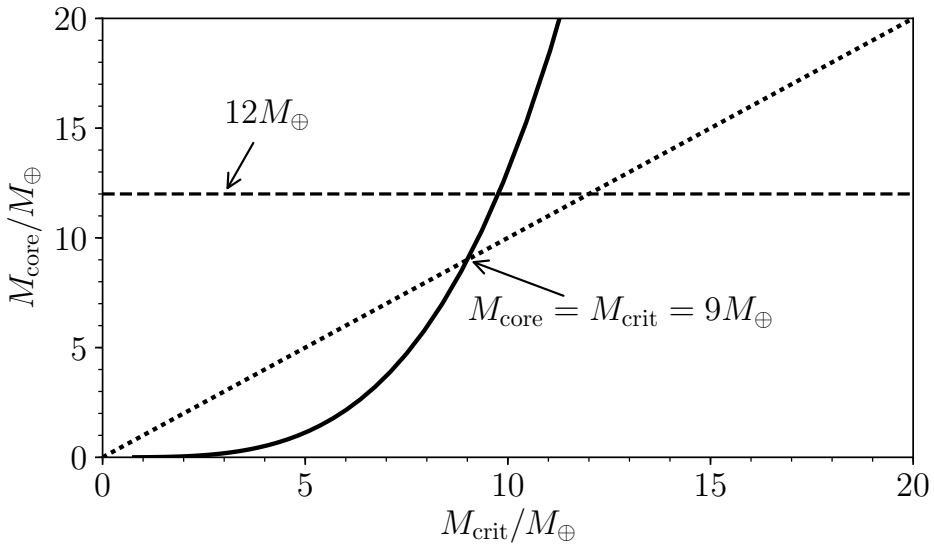

Figure 2.1 Is it possible to keep the core sub-critical by accreting at a constant $\dot{M}$ for the ~ 3 Myr gas disk lifetime? We define $M_{\rm core} = \dot{M}\tau_{\rm disk}$ as the size the core can attain in the lifetime of the disk, assuming it is accreting at the rate $\dot{M}$ associated with $M_{\rm crit}$ (Equation 2.1) for the entire disk lifetime (solid line). The cutoff mass for accretion to keep the core sub-critical is $M_{\rm core} = M_{\rm crit}$ (dotted line). At a higher accretion rate, the cores cannot stay sub-critical because the accretion rate required to prevent runaway would cause the core to grow larger than its assumed size.

The values of $M_{\rm core}(M_{\rm crit})$ are calculated for $\rho_c = 3$ g cm^{-3} (the typical core mean density for Uranus and Neptune from Nettelmann et al. [123]), $\kappa_0 = 0.1$ cm^{2}g^{-1}, and $\mu = 2.3\, m_p$ (mean molecular mass of H_2/He gas).

the total atmospheric mass compared to that of an otherwise identical non-accreting core, and increasing the amount of energy that must be radiated away to bind additional

gas.

In Equation 2.1, we express the functional dependence of the critical core mass on the planetesimal accretion rate ($\dot{M}$) and the gas opacity in the outer radiative layer of the atmosphere (κ_0) from Rafikov [131]. We note that this is the applicable value for $M_{\rm crit}$ in the high accretion rate limit, for which energy in balances energy out.

$$M_{\rm crit} = \left(\left(\frac{k_B}{\mu} \right)^4 \left(\frac{4\pi\rho_c}{3} \right)^{1/3} \frac{\kappa_0 \dot{M}}{\zeta_0 \sigma G^3} \right)^s , \tag{2.1}$$

with constants of order unity $s = \frac{3}{7+3\delta}$ and $\delta = 1.2$, a constant $\zeta_0 = 4 \times 10^{-32}$ g$^{-\delta}$ that depends on δ, the mean molecular mass μ, the Boltzmann constant $k_B = 1.38 \times 10^{-16}$ cm^2 g s^{-2} K^{-1}, the gravitational constant $G = 6.67 \times 10^{-8}$ cm^3 g^{-1} s^{-2}, the Stefan-Boltzmann constant $\sigma = 5.67 \times 10^{-5}$ g s^{-3} K^{-4}, and the mean density of the core ρ_c. In this context, the critical core mass is calculated assuming steady state can be achieved. We note that a high accretion luminosity (e.g. comparable to that of a core in the fast or intermediate accretion regime of Rafikov [70]) increases the total luminosity of the core, causing it to evolve toward this steady state on timescales typically shorter than the disk lifetime ($\tau \sim E/L$, where E is the total energy of the envelope).

How massive can the core grow in the lifetime of the gas disk (using $\tau_{\rm disk} \simeq 3$ Myr as a conservative estimate) and stay sub-critical? The maximum accretion rate that the core can sustain for the full disk lifetime without growing beyond its final size $M_{\rm core}$ is: $\dot{M} = M_{\rm core}/\tau_{\rm disk}$.

In Figure 1, we plot Equation 2.1 for $\kappa = 0.1$ cm^2 g^{-1}, a commonly assumed opacity in protoplanetary disks [e.g 106, 129]. If $M_{\rm core} = \dot{M}\tau_{\rm disk} < M_{\rm crit}$, where $\dot{M}$

is the accretion rate associated with $M_{\rm crit}$, then accretion can in principle keep the core sub-critical. If accretion is not sustained at this value throughout the final stages of core growth, which in standard models is comparable to the full disk lifetime, the effective value of $M_{\rm crit}$ is reduced. We may therefore calculate the maximum core mass consistent with runaway prevented by sustained accretion at the rate $\dot{M}$. We conclude from Equation 2.1 that $\sim 10 M_{\rm E}$ is the greatest mass the core can grow to in $\tau_{\rm disk}$ and remain just sub-critical. In Figure 1, the accretion rate that gives a critical core mass of $10 M_{\rm E}$ allows the core to grow to $10 M_{\rm E}$ in the lifetime of the disk. Increasing the accretion rate to above this value is not enough to prevent runaway; to the contrary, it gives rise to a super-critical core. We conclude that for our assumed values of opacity, a high accretion rate could not have kept the $\gtrsim 12 M_{\rm E}$ cores of Uranus and Neptune from reaching critical core mass.

The dependence of the critical core mass on opacity introduces a degree of uncertainty. For example, increasing the grain opacity by an order of magnitude ($\simeq$1 $\rm cm^2 g^{-1}$) doubles the critical core mass, allowing the cores to stay sub-critical for the lifetime of the disk for a limited range of accretion rates. However, this still requires fine tuning: the cores would have had to accrete at a specific rate for the entire disk lifetime.

In the discussion above, we have considered a scenario of core accretion in which a steady-state solution for the envelope exists, given a high planetesimal accretion rate $\dot{M}$. In this scenario, planetesimal accretion acts as a continuous heating source for the envelope that balances its outgoing luminosity, which prevents the envelope from

contracting and sustains it in a steady state. It is plausible that the protoplanets were in this high-accretion regime at some point in their evolution, but we have shown that they could not have sustained it for their entire formation time, hence this is not a way to prevent collapse. In 2.2, we discuss an alternate way the ice giants could have avoided runaway by considering growth with a low planetesimal accretion rate, which can be sustained for a significant fraction of the gas disk lifetime. If $\dot{M}$ is slow enough, in principle the cores can be limited to small enough masses that they wouldn't be able to quickly accrete the amounts of gas necessary to trigger runaway gas accretion.

2.2.2 Discussion of the Classical Core Accretion Scenario

We now discuss the classical work of Pollack et al. [130], which models the formation of the giant planets with a 3-phase process that includes the concurrent accretion of gas and solids. We focus on the key assumptions of this model, and refer to the original work for a detailed description of the model.

During Phase 1, the core grows by planetesimal accretion with a rate $\dot{M}_{\rm solids}$ until it depletes the bulk of the material in its feeding zone with no replenishment from the outside, reaching isolation mass at a time $\tau_{\rm iso}$. This model assumes that the protoplanetary core carves a gap in the planetesimal disk, causing core growth to halt at roughly the local isolation mass. The core starts accreting an envelope once the core is massive enough that its escape speed is greater than the local sound speed of the gas. The envelope is required to cool and contract for it to grow, and its growth rate can be stalled by external heating. For example, the envelope cannot cool as efficiently via KH

contraction if planetesimal accretion deposits extra energy into the envelope, compared to the case with no planetesimal heating: the planetesimal accretion luminosity will be part (if not the main component) of the planet's total outgoing luminosity. This model relies on the planetesimal accretion luminosity to be initially large, and then to drop to less than the KH luminosity of the envelope. This can be achieved by cutting off the supply of planetesimals by growing the core to isolation mass, assuming no radial drift of solids through the gas. The key point is that accretion cuts off while the core remains low enough mass to take extra time to accrete gas.

However, this is problematic: observations have shown the existence of pebbles in the outer disk [105], which are mm/cm-sized particles that are well-coupled to the gas. These particles may drift into the feeding zone with the gas and rapidly accrete onto the core via pebble accretion [125, 117, 132], which has been shown to be extremely efficient in forming cores even in the outer disk.

Rapid accretion of mm-cm sized planetesimals (known as pebbles) by gas-assisted growth can reduce the growth timescales sufficiently to allow giant planets to form at wide orbital distances. We define the Hill radius to be the distance from the planet at which the gravity of the planet dominates over the tidal gravity of the Sun:

$$r_H = a \left(\frac{M_c}{3M_\odot} \right)^{1/3}, \tag{2.2}$$

where M_c is the mass of the core, $M_\odot$ is the mass of the Sun, and a is the semimajor axis of the planetary orbit. Under the most favorable conditions to gas-assisted growth, pebbles are able to accrete once they enter the Hill sphere, increasing the cross-section

for accretion significantly from the classical one derived for gravitational focusing. For a review of classical growth rates, see Goldreich et al. [114].

The most favorable velocity at which the accreting particles approach the core is given by the Hill velocity, $v_H = \Omega r_H$ ("shear-dominated regime" Goldreich et al. [114], section 2.2). This is an appropriate value for low dispersion velocities, i.e. $\lesssim v_H$, for which accretion is 2-dimensional and particles approach the core due to Keplerian shear. In the absence of damping, the cores themselves excite nearby particles to $\sim v_H$. We note for reference that gas can both damp the dispersion velocities of particles and - when turbulent - kick them to higher velocities. Which effect dominates depends on particle size. According to pebble accretion, in the most favorable case, the collision rate is equal to the Hill entry rate. Therefore, the cross-section for accretion from a thin disk of particles is $\sigma_{\rm acc} \sim H r_H$, where H is the particle scale height. For large ($\gtrsim M_{\rm E}$) cores accreting pebbles under typical disk conditions, the maximum growth rate (see Rosenthal et al. [132]), assuming the existence of pebbles coupled to the gas in the outer disk, is then:

$$\dot{M} = \rho \sigma_{\rm acc} v_{\rm rel} = \frac{\Sigma}{H} r_H H r_H \Omega = \Sigma {r_H}^2 \Omega, \tag{2.3}$$

where $v_{\rm rel}$ is the relative velocity between the pebbles and the core, ρ and Σ are the respective volume and surface mass densities of solids in the disk, and Ω is the local Keplerian frequency. The cores cannot stay at 12 $M_{\rm E}$ for very long: the timescale for such a core to double its mass via pebble accretion can be as short as 10^5 years (Equation 2.3) at the current distance of Neptune in a disk with a Minimum Mass Solar Nebula (MMSN) solid surface density profile [116]. We note that the minimum timescale for gas-

assisted growth does not depend on the surface density of gas, which primarily affects the sizes of the pebbles accreted [132], and that increasing the surface density in solids can significantly decrease this timescale. The fast growth rates predicted by pebble accretion for large cores and the fact that radial drift through the disk replenishes the planetesimal supply are problematic for the original core accretion interpretation for ice giants, which relies on core growth halting at the isolation mass. In other words, having a roughly fixed-mass core over timescales comparable to the disk lifetime is needed for slow gas accumulation to explain the properties of the ice giants' envelopes without fine tuning, and gas-assisted growth suggests that $\sim 12 M_{\rm E}$ cores should grow rapidly.

Pebble isolation [118] can possibly save the original interpretation: pebble accretion can be halted once the core reaches a mass at which the core can gravitationally perturb the gas disk such that entrainment in gas flowing around the core stops the drift of pebbles onto the core. We now estimate the pebble isolation mass (see Rosenthal et al. [132] for a detailed discussion). We define the Bondi radius as the radius at which the gravitational potential energy of the core equals the thermal energy of the gas: $R_B = 2GM_c/c_s^2$, where c_s is the sound speed of the gas and M_c is the core mass. The pebble isolation mass is the core mass required for the Bondi radius to surpass the Hill radius:

$$M_{\rm iso,peb} \simeq \left(\frac{a^3}{24 M_\odot G^3}\right)^{1/2} c_{\rm s}^3 \simeq 27 M_{\rm E} \left(\frac{a}{20{\rm AU}}\right)^{3/4}, \tag{2.4}$$

using a temperature profile of $T = 300(a/{\rm AU})^{-1/2}$ K appropriate for the MMSN in estimating the sound speed. Our estimate of $M_{\rm iso,peb}$ at ice-giant distances is close enough (to within an order of magnitude) to the current core masses of Uranus and

Neptune that halting accretion of pebble-sized objects via this mechanism cannot be ruled out with certainty. Alternatively, pebble accretion can halt if the supply of pebbles in the outer disk is cut off, which, based on the observations of Andrews et al. [105], we suppose is a less likely explanation.

Assuming core growth can be halted at a mass low enough that the planet cannot quickly accumulate a mass in gas large enough to trigger runaway gas accretion, the planet enters a long-lived phase of simultaneous, slow accretion of solids and gas (Phase 2), which ends when the core accumulates a mass in gas comparable to its own mass. We note that the length of Phase 2 depends strongly on the opacity, which is uncertain. Work including realistic opacities by Movshovitz et al. [122] and Mordasini [121] suggests that the values for the opacities used in the classical Pollack et al. [130] simulations were too high, suggesting that it is even worth considering grain-free models in giant planet formation. Additional models including protoplanetary disk opacities by Piso et al. [128] showed that taking into account grain growth lowered the envelope opacity, yielding significantly faster runaway accretion timescales.

Finally, for this 3-phase scenario to work, the timescale to accrete the $\sim 2M_{\rm E}$ atmospheres of Uranus and Neptune would have to be comparable to or greater than $\tau_{\rm disk} - \tau_{\rm iso}$, which is a couple Myr. Furthermore, Jupiter and Saturn must have reached runaway in the $\sim$ 3 Myr lifetime of the gas disk, which does not happen for the high opacities used in Pollack et al. [130]. We note that the timescale for KH contraction (and hence the critical core mass for low planetesimal accretion rates) varies as a function of distance from the star primariy due to changes in opacity. Opacities in the interstellar

medium scale with temperature as $\kappa \propto T^2$, and, as shown in Piso et al. [128], the critical core mass declines with distance from the star. More realistically, once grains have grown, the opacity typically loses its temperature dependence, and the resulting critical core mass is a weakly declining function of stellocentric distance. So, for this scenario to be correct, Jupiter and Saturn must have had only marginally enough time to achieve runaway.

Given the difficulty in constraining the timescales of envelope growth due to the uncertainties in opacities, the specific opacities required to grow the atmospheres of Jupiter and Saturn within reasonable estimates for gas disk lifetimes without turning Uranus and Neptune into gas giants, and the challenges in halting the growth of a core experiencing pebble accretion, we find it valuable to consider our proposed dynamical alternative.

2.3 How Depleted is the Gas at the Onset of Gas Accretion?

We have suggested that the fully grown cores of Uranus and Neptune could not have accreted their atmospheres from a full gas disk. Instead, the cores accreted the bulk of their gas from a depleted disk late in its lifetime. Here we calculate the maximum gas mass available for accretion in the depleted disk. We note that the atmospheric growth timescale is dominated by the last doubling timescale for the atmosphere [e.g. 127], and that the fully grown cores were massive enough to accrete all of the gas in

their respective feeding zones on short timescales, unless limited by the solid accretion luminosity (see Section 5).

To calculate how depleted the disk has to be at the onset of gas accretion to produce the observed envelope masses of the ice giants, we set the mass of gas in each feeding zone equal to the mass of each planet's atmosphere, respectively. We approximate a planetary feeding zone with an annulus at a distance a (the planet's semi-major axis) from the Sun, with a width of Δa of $\simeq 5$ Hill radii (see Schlichting [133]). The Hill radius, r_H, is the distance from the planet at which the gravity of the planet dominates the gravity of the Sun (i.e. the distance scale relevant to the planet's ability to accrete material):

$$r_H = a \left(\frac{M_c}{3M_\odot} \right)^{1/3}, \tag{2.5}$$

where M_c is the mass of the core, $M_\odot$ is the mass of the Sun. The atmosphere of an ice giant is not massive enough to contribute substantially to the planet's gravity, so we use the core mass instead of the total planet mass in the Hill radius definition. Thus, though not entirely self-consistent, the following is reasonable estimate of the gas feeding zone mass, given that the core dominates the gravitational potential of the planet:

$$M_g \simeq 2\pi a \Delta a \Sigma_g(a), \text{ where} \tag{2.6}$$

$\Sigma_g(a)$ is the mass surface density of gas in the depleted disk at a distance a from the Sun. We assume that $\Sigma_g(a)$ stays roughly constant across the width of the feeding zone. Taking $\Delta a \simeq 5 r_H$, this gives us:

$$M_g \simeq 10\pi a^2 \left(\frac{M_c}{3M_\odot} \right)^{1/3} \Sigma_g(a). \tag{2.7}$$

We denote the surface density of the full gas disk as $\Sigma_{g,0}$. We define a disk depletion factor: $f \equiv \Sigma_g/\Sigma_{g,0}$. Setting $M_g = M_{\rm atm}$ in Equation 2.7 gives us the depletion factor at the onset of gas accretion:

$$f = \frac{\Sigma_g}{\Sigma_{g,0}} \simeq \frac{M_{\rm atm}}{M_{g,0}} \simeq \frac{M_{\rm atm}}{10\pi a^2 \Sigma_{g,0}} \left(\frac{3M_\odot}{M_c}\right)^{1/3}. \tag{2.8}$$

This depletion factor can also be understood as the fraction of the original, full gas disk feeding zone mass, $M_{g,0}$, that ended up in the atmosphere of the planet.

We use the MMSN model's disk surface density profile [116],

$$\Sigma_{g,0} = 1700 \left(\frac{a}{\rm AU}\right)^{-3/2} \text{g cm}^{-2}, \tag{2.9}$$

the current values of a, and total planetary gas masses from Nettelmann et al. [123] to calculate that $f \simeq 0.01$ when Uranus and Neptune entered the stage of gas accretion during which their atmospheres acquired most of their mass. This calculated f is an upper limit because we assumed the planets accreted a minimum amount of gas, and have neglected viscous feeding zone replenishment and any significant early-on gas accretion.

We rewrite this criterion in a simple form. We define the solid-to-gas ratio in the full gas disk: $f_s \equiv \Sigma_s/\Sigma_{g,0}$. For the planet's atmospheric mass to be comparable to its gas feeding zone mass, we again set $M_g = M_{\rm atm}$ in Equation 2.7:

$$M_{\rm atm} \simeq 10\pi a^2 \left(\frac{M_c}{3M_\odot}\right)^{1/3} \Sigma_g. \tag{2.10}$$

So,

$$\frac{M_{\rm atm}}{M_c} \simeq \frac{10\pi a^2 \Sigma_g}{(3M_\odot)^{1/3} M_c^{2/3}} = \frac{10\pi a^2}{(3M_\odot)^{1/3}} M_c^{-2/3} \Sigma_{g,0} f. \tag{2.11}$$

The cores of Uranus and Neptune are comparable to their local isolation masses [e.g. 113]. Though we favor pebble accretion as the dominant mode of core growth at ice giant distances, we nonetheless find it useful to express planetary masses in units of isolation mass because it allows us to write the scaling relations in a simple form.

Classically, the isolation mass is the mass of a core that has consumed all of the solid material in its feeding zone:

$$M_{\rm iso} \simeq 2\pi a \Delta a \Sigma_s(a), \tag{2.12}$$

where $\Sigma_s(a)$ is the disk surface mass density in solids at a distance a from the Sun. Taking the width of the feeding zone in Equation 2.12 to be $\Delta a \simeq 5 r_H = 5a(M_{\rm iso}/(3M_\odot))^{1/3}$, we solve for $M_{\rm iso}$:

$$M_{\rm iso} \simeq \frac{(10\pi a^2 \Sigma_s)^{3/2}}{(3M_\odot)^{1/2}} \tag{2.13}$$

Then,

$$\frac{10\pi a^2}{(3M_\odot)^{1/3}} \simeq \frac{M_{\rm iso}^{2/3}}{\Sigma_s}. \tag{2.14}$$

Thus, from Equations 2.11 and 2.14, we have:

$$\frac{M_{\rm atm}}{M_c} \simeq f \frac{\Sigma_{g,0}}{\Sigma_s} \left(\frac{M_{\rm iso}}{M_c}\right)^{2/3} = \frac{f}{f_s} \left(\frac{M_c}{M_{\rm iso}}\right)^{-2/3}, \tag{2.15}$$

and for $M_c \simeq M_{\rm iso}$, $M_{\rm atm}/M_c \simeq f/f_s$. When forming an isolation-mass scale ice giant, for the atmospheric mass to be comparable to the total planet mass, we need the surface

density of gas to be comparable to the surface density of solids at the onset of gas accretion (see Figure 2).

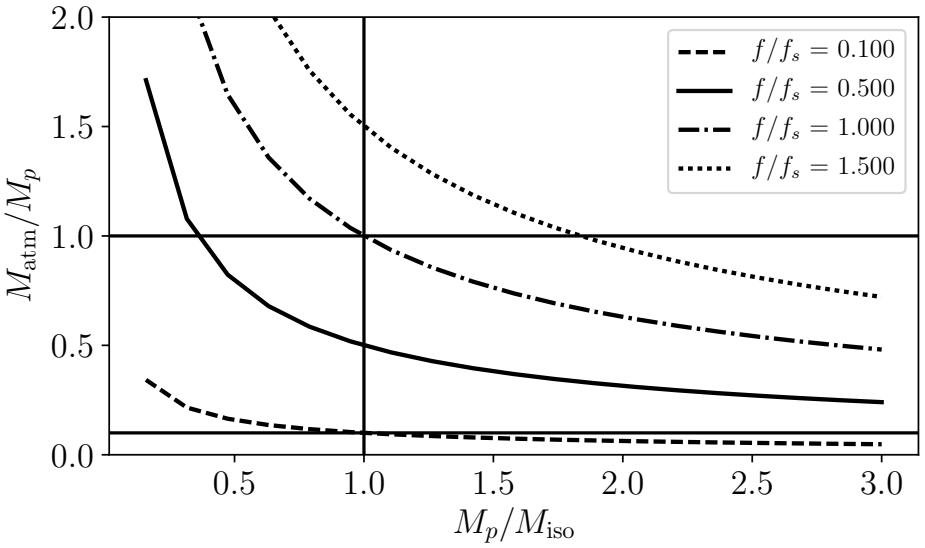

Figure 2.2 $M_p/M_{\rm iso}$ gives the planet mass in terms of the isolation mass. $M_{\rm atm}/M_p$ is the atmospheric mass fraction. Here we vary the gas-to-solid ratio, illustrating that for $\sim$ isolation-mass cores the gas-to-solid ratio at the onset of gas accretion is comparable to the final atmospheric mass fraction. The horizontal lines mark ratios of 0.1 and 1, which are reasonable for the onset of dynamical instability (see text).

This point coincides with the onset of a dynamical instability in the disk: the gas can no longer damp the eccentricities and inclinations of the growing cores once the surface densities of solids and gas are comparable.

The required disk gas-to-solid ratio for the onset of instability may be less than 1. For instance, Dawson et al. [108] found in simulations of planetary embryos embedded

in a gas disk that gas stopped being able to efficiently damp the embryos' eccentricities at $f = 10^{-3}$, i.e. $f/f_s \sim 0.1$. In another simulation by Ford & Chiang [112] of planetary cores in a planetesimal disk, a dynamical instability was triggered once the density of planetesimals dropped to roughly 0.1 times the density of big bodies. As discussed in section 5.2 of Goldreich et al. [113], dynamical friction due to gas can be treated in the same way as planetesimal dynamical friction, replacing the small bodies' surface density with that of the gas, and the random velocity of the small bodies with the local sound speed. If the instability starts at $f/f_s \sim 0.1$, the eventual atmospheric masses will be approximately 0.1 M_p, which matches the ice giants' atmospheric masses quite well.

2.4 How Long Can the Gas Disk Stay in a Depleted State?

Observations of protoplanetary disks suggest that gas disks dissipate on a $\lesssim$ 3-5 Myr timescale, with most young stars stars losing their gas disks by 10 Myr [115, 104, 111]. Disk dispersal mechanisms include: viscous accretion onto the host star, viscous spreading, and photoevaporation driven by irradiation from the host star, as well as the local interstellar radiation field. Transitional disks, if representative of disks in the process of clearing, have lifetimes of $\sim$ 10% of the full disk lifetime (e.g. Alexander et al. [104], Espaillat et al. [111]). We note that the $\sim$ 10% transition disk fraction could instead reflect the $\sim$ 10% frequency of planets massive enough to carve holes in the gas disk.

The predominant mode of disk dispersal can vary throughout the extent and

lifetime of the disk, with the inner regions clearing due to accretion onto the star, and photoevaporation driving mass loss from the outer ones. We are interested in the outer disk, where photoevaporation can be important. The protoplanetary disk undergoes photoevaporation as it is heated by far-ultraviolet (FUV: 6 eV to 13.6 eV), extreme-ultraviolet (EUV: $>$ 13.6 eV), and X-ray photons [101]. FUV photons strike the disk and heat the gas to temperatures of $\approx 10^3$ K, EUV photons ionize the gas and heat it to $\approx 10^4$ K, and X-ray flux heats the gas surface layer to $\approx 10^3 - 10^4$ K. The critical radius is the relevant length scale to determine how far in the disk photoevaporative mass loss starts to be efficient (see Equation 1, Adams et al. [101]:

$$r_g = \frac{GM_\odot \mu_{\rm gas} m_p}{kT}, \tag{2.16}$$

where $\mu_{\rm gas}$ is the mean molecular weight of the gas, and $m_p = 1.67 \times 10^{-24}$ g is the proton mass. Beyond r_g, the local sound speed of the externally heated gas is greater than the local escape velocity from the star, allowing heated gas to escape. For a solar-mass star and atomic hydrogen gas, the critical radius for FUV is $r_g \sim 100$AU, for EUV is $r_g \sim 10$AU, and for X-ray the range of critical radii is $r_g \sim 10 - 100$AU [126].

We estimate the timescale for disk dispersal due to solar EUV-driven mass loss, where the stellar irradiation drives thermal escape of gas with outflow velocities of $v_{\rm flow} \simeq 10\,{\rm km\ s^{-1}}$. We check that the hydrogen gas is mostly ionized at the $\tau = 1$ layer of the disk, where τ is the optical depth to the EUV radiation. Since $\tau = N_{\rm H}\sigma_{\rm ion}$, the neutral column depth to $\tau = 1$ is $N_{\rm H} = 1/\sigma_{\rm ion}$, where $\sigma_{\rm ion} = 6.3 \times 10^{-18}$ cm^2 is the photoionization cross-section for hydrogen by Lyman limit photons. The

gas outflows at the critical radius $r_{\rm g}$, where the heating puffs the disk up to a scale height comparable to the stellocentric distance a: $H \sim c_s/\Omega \sim v_{\rm esc}/\Omega \sim a$. Thus, the volumetric number density of neutral hydrogen at r_g is $n_{\rm H} \sim N_{\rm H}/H \sim 10^3$ cm^{-3}. By balancing photoionization with recombination, we obtain the number density of ionized hydrogen at the $\tau = 1$ surface: $n_{\rm H}\sigma\Phi_i/(4\pi a^2) = \alpha {n_p}^2$, where the coefficient for radiative recombination for hydrogen at 10^4 K is $\alpha = 2.6 \times 10^{-13}$ cm^3 s^{-1}, n_p is the number density of ionized hydrogen/electrons (assuming charge balance), a is the stellocentric distance, and Φ_i is the ionizing stellar flux. For a solar-mass star,

$$n_p \simeq \left(\frac{\Phi_i}{4\pi a^3 \alpha}\right)^{1/2} \simeq 10^5 \left(\frac{\Phi_i}{10^{41}{\rm s}^{-1}}\right)^{1/2} {\rm cm}^{-3}, \tag{2.17}$$

so the assumption of ionization is correct. The the recombination timescale, $\tau_{\rm rec} \simeq 1/(\alpha n_p^2) \simeq 4 \times 10^2$ s, is much less than the outflow timescale $\tau_{\rm flow} \simeq a/v_{\rm flow} \simeq 10^8$ s, justifying the assumption of ionization equilibrium. The mass-loss rate is therefore: $\dot{M} \simeq 4\pi a^2 m_{\rm p} n_p v_{\rm flow} \simeq (4\pi a \Phi_i/\alpha)^{1/2} m_{\rm p} v_{\rm flow} \simeq 7 \times 10^{-10} M_\odot$ yr^{-1}, assuming an ionizing luminosity of 10^{41} photons s^{-1} appropriate for T Tauri stars [102]. Using the MMSN initial gas disk mass of $M = 0.01 M_\odot$ [109], the gas disk already depleted by a factor of 100 disperses on a timescale of: $\tau_{\rm disp} = M/\dot{M} = 10^{-4} M_\odot/(7$ x $10^{-10} M_\odot$ yr$^{-1}) \approx 10^5$ yr. We note that this calculation is an order-of-magnitude estimate. However, it agrees with transitional disk lifetime estimates both from statistical arguments based on observations [120], and on theoretical work [103].

Photoevaporation of circumstellar disk gas may also be driven by far-ultraviolet (FUV) photons from the largest stars in the local star forming region. Classically, it is

thought that FUV can only result in substantial mass loss if the outer disk boundary r_d extends beyond the critical radius. We note that this radius is beyond the region of the solar system's ice giants. However, Adams et al. [101] found that FUV may be important at disk distances of $\sim 5-10$ times less than r_g. Conceptually, a sub-sonic gas outflow originates at r_d, which then flows through a sonic point (being accelerated to the sound speed), eventually reaching r_g and escaping. Assuming external irradiation, Adams et al. [101] found that FUV may drive outflows at the current locations of the ice giants, up to stellocentric distances of $\simeq$ 15 AU. For example, the mass-loss rates due to photoevaporation by FUV photons are 10^{-9} to $10^{-8} M_{\odot}$ yr^{-1} for a strongly FUV-irradiated disk at respective disk radii of 20 to 40 AU (see Fig. 6 of Adams et al. [101]). In this case, the gas disk depleted by a factor of 100 disperses on a timescale of: $\tau_{\rm disp} \equiv M/\dot{M} \simeq 10^4$ to 10^5 years.

X-rays may play an important role in determining the gas disk lifetime. X-rays heat the gas to $T \sim 10^3 - 10^4$ K, which corresponds to a critical radius for a solar-mass star of 10-100 AU. Owen et al. (2011) analytically estimated the mass-loss rate to be $\sim 8 \times 10^{-9} M_{\odot}$ yr^{-1} for a Solar-mass star with an X-ray luminosity of 10^{30} erg s^{-1} which would give a lifetime of $\sim 10^4$ yr for our depleted disk.

For completeness, we consider the timescale for viscous accretion of gas onto the Sun. Considering the model of viscous disk accretion in a disk with turbulence parameterized by the α parameter [134], the timescale for gas accretion in the nebula

is:

$$\tau_{\rm acc} = \frac{1}{\alpha\Omega}\left(\frac{H}{r}\right)^{-2}. \tag{2.18}$$

We note that this timescale is independent of disk surface density. Using a passive disk temperature profile, and an α value[1] of 10^{-3}, this timescale is $\sim$ 1-2 Myr on orbital separations of 20-40 AU.

2.5 Proposed Solution

We suggest that the fine tuning problem may be resolved by proposing a dynamical reason for the ice giants to accrete their gas envelopes in the last $\lesssim 10^5$ years of the gas disk lifetime. Why did they start accreting their atmospheres precisely at that time? In Section 3, we calculated that the ice giants have to accrete the bulk of their masses in gas from a disk depleted to $10^{-3} - 10^{-2}$ times its original surface density, for their current atmospheric masses to be equal to the mass of gas in their respective feeding zones. Since the dust-to-gas ratio is ~ 0.01 in a full gas disk, in a depleted disk it's order unity: $\Sigma_s/\Sigma_g = \Sigma_s/\Sigma_{g,0}$ x $(\Sigma_g/\Sigma_{g,0})^{-1} \simeq 1$.

A full gas disk efficiently damps the random velocities of protoplanetary cores, forcing them to stay on roughly circular orbits. However, once the gas is depleted to a point where the surface density of the solids becomes comparable to the surface density of the gas, gas drag can no longer balance dynamical excitations by neighboring bodies, causing the growing cores' random velocities to increase. The resulting dynamical

[1]Values of 10^{-3} to 10^{-2} are often used partly because they give the approximate observed disk lifetimes

instability may cause the cores to travel to a new location, continue their core growth there, and accrete a gas envelope from the remaining gas.

Why might the ice giants have remained small until this late stage? We suppose that the ice giants' cores started growing on closer-in orbits, where they were prevented from reaching a critical core mass. There are several possibilities that could have halted core growth on closer-in orbits. The differing turbulence structure in the disk could have affected pebble accretion. Alternatively, the ice giants could have formed in between Jupiter and Saturn, with Saturn consuming the infalling pebbles. Proximity to the gas giants, which excite the random velocities of neighboring bodies, can slow growth by accretion of large planetesimals. Finally, as the pebble isolation mass scales with stellocentric distance, the cores could have reached a sub-critical pebble isolation mass on closer-in initial orbits, if pebble accretion was the dominant mode of core growth at their initial orbital locations. We note that the core growth at smaller stellocentric distances (such as that of Jupiter and Saturn) is not limited to the pebble isolation mass, as significant accretion of km-sized planetesimals via the standard Goldreich et al. [114] regimes can occur within the gas disk lifetime.

Plotted in Figure 3 are estimates on limits of core growth as a function of stellocentric distance, where the dominant accretion regime can vary depending on the orbital timescale. The mass a core can grow to within the disk lifetime $\tau_{\rm disk}$ by accreting planetesimals at a rate $\dot{M}_s$ is: $M_{\rm core,max} = \dot{M}_s \tau_{\rm disk}$. Using the Goldreich et al. [114] growth rate for gravitational focusing at random planetesimal velocities of order the

Hill velocity,

$$\dot{M}_s \simeq \Sigma_s(a)\Omega R_p^2 \left(\frac{v_{\rm esc}}{v_H}\right)^2, \tag{2.19}$$

where $\Sigma_s(a)$ is the disk surface density in solids at a stellocentric distance a, $\Omega = (GM_\odot/a^3)^{1/2}$ is the local Keplerian frequency, R_p is the protoplanetary radius, and $v_{\rm esc} = (2GM_{\rm core}/R_p)^{1/2}$ is the escape velocity at its surface, we solve for the maximum core size that a protoplanet of a mean density ρ_c can achieve within the disk lifetime by accreting planetesimals:

$$M_{\rm core,\ max} \simeq \frac{54}{\pi}\frac{G^{3/2}M_\odot^{1/2}\tau_{\rm disk}^3\Sigma_s(a)^3}{\rho_c a^{3/2}}. \tag{2.20}$$

As the normalization of the pebble isolation mass is uncertain, we plot Equation 2.4 scaled by a factor of 1/2 for the masses to be comparable to the current ice giants' masses. To make this plot, we chose disk surface density and temperature profiles appropriate for the MMSN, a gas disk lifetime of 3 Myr, and a mean density $\rho_c = 3$ g cm^{-3} for the core.

At distances close to the star, cores can grow efficiently via standard planetesimal accretion: there is no timescale problem as there is in the outer disk, and the core mass is limited instead by the supply of planetesimals. In the outer disk, the dynamical timescales are too long to form giant planets with standard growth mechanisms without appealing to accretion from a very thin planetesimal disk. However, as long as a core can grow to a fraction of an $M_{\rm E}$ required for pebble accretion to begin, it can grow to its pebble isolation mass very quickly. We observe in Figure 3 a portion of the plot where the maximum core masses falls below the nominal value for the critical core mass of 10

$M_{\rm E}$. This dip in the plot illustrates that there can be a region in the disk where gas giant cores cannot form in situ, surrounded by an area where they can.

We have discussed above several means by which the cores could have stayed sub-critical until onset of the dynamical instability in the depleted disk, bearing in mind that any model of solar system formation must keep Uranus and Neptune sub-critical during the full disk lifetime.

Once the gas was depleted enough to allow the cores' orbits to be gravitationally excited and not damped, the gas giants flung them outward, either to their current locations, or to a sufficient distance that they began migrating outward due to planetesimal-driven migration. Accreting the pebbles at their new locations allowed the cores to at least double in size to their final masses. Each core's orbit was circularized as a result of accreting a mass of roughly circular-orbit solids (the pebbles are coupled to the gas) comparable to its own mass.

Under the most favorable conditions for pebble accretion, from Equation 2.3 the timescale for last doubling is $t_{\rm growth} = M/\dot{M}$, or approximately 10^5 years for a Neptune-sized core at its current location, with a surface density in solids given by the MMSN. We note that pebble accretion continues to operate in a depleted gas disk, though the sizes of pebbles that are effectively accreted change [132]. The depleted disk dissipates on a $\lesssim 10^5$ year timescale due to EUV- and FUV- driven mass loss. Our depleted disk may dissipate on a shorter timescale of 10^4 years due to photoevaporation driven by stellar X-ray radiation. If this shorter timescale applies, the cores of Uranus and Neputune may nonetheless undergo their last doubling in a timescale comparable

to the remaining disk lifetime if they accrete from a disk enhanced in solids by a factor of only a few compared to the MMSN. Whether pebble accretion can decrease core growth timescales sufficiently to allow Neptune-sized cores to double within $10^4 - 10^5$ yr depends on the size distribution of planetesimals and the level of turbulence in the depleted gas disk (see Rosenthal et al. [132] for a detailed discussion). Here, we merely note that fast enough growth is possible.

Such a high accretion rate limits the atmospheric mass that a core can sustain in a steady state [131], possibly to a fraction of an Earth mass for the ice giants. For our scenario to work, the core accretion luminosity cannot be so high during the entire phase of last doubling that it prevents the accumulation of a couple $M_{\rm E}$ of gas.

Because the last doubling and disk dispersal timescales are comparable, it is reasonable to suggest that the ice giants' cores could have finished their growth as they accreted gas envelopes from the remaining gas disk. If the ice giants' cores doubled their mass while finishing their growth at their final orbital locations, their growth rate during this final doubling phase is approximately: $\dot{M} \simeq 6M_{\rm E}/(10^5{\rm yr}) \simeq 10^{16}{\rm g\ s^{-1}}$. Rafikov [131] estimates the critical core mass in the fast planetesimal accretion regime as the mass of the core that can hold onto its own mass in gas. With masses of around $2M_{\rm E}$, the atmospheres of Uranus and Neptune are not massive enough to undergo nucleated instability. We claim that the cores need not be super-critical to maintain $\sim 2M_{\rm E}$ envelopes in a steady-state during rapid planetesimal accretion: a lower core mass is sufficient. Assuming that the planetesimal accretion luminosity is transported out via radiative diffusion through the atmosphere with a constant gas opacity κ_0, Rafikov [131]

finds that:

$$M_{\rm atm} \simeq \zeta \left(\frac{GM_p\mu}{k}\right)^4 \frac{\sigma}{\kappa_0 L}, \tag{2.21}$$

where L is the accretion luminosity, R_p is the radius of the core, and $\zeta = \zeta_0 M_p^\delta$, with ζ_0 and δ as defined for Equation 2.1. Therefore, a core accreting at a rate $\dot{M} \simeq 10^{16}$g s^{-1} at a distance $a = 20$AU can hold onto a $2M_{\rm E}$ atmosphere as long as $M_{\rm core} \gtrsim 12M_{\rm E}$. Thus, the cores of Uranus and Neptune in the final stage of growth were able to capture all of the gas in their feeding zones, forming steady state envelopes heated by fast planetesimal accretion. We note that the evolution timescale (given by the KH timescale) to such a steady state can be short, as a consequence of the high planetesimal accretion luminosity. If we suppose that the dominant source of the luminosity is the accretion luminosity ($L = GM_p\dot{M}/R$), then the KH timescale can be approximated as follows:

$$\tau_{KH} \simeq \frac{E}{L} \simeq \frac{GM_{\rm p}M_{\rm atm}}{RL} \simeq \frac{M_{\rm atm}}{\dot{M}}. \tag{2.22}$$

Therefore, for a $2M_{\rm E}$ atmosphere surrounding a core accreting at a rate $\dot{M} \simeq 10^{16}$g s^{-1}, the evolution timescale is approximately $\simeq 3 \times 10^4$ yr. Additional effects such as the pollution of the envelope from the disruption of icy accreting solids have been shown to decrease this timescale substantially [136]. If the appropriate timescale for the disappearance of the depleted disk is 10^4 years, then the rate of pebble accretion (and, therefore, the luminosity) would have to be higher by a factor of 10 from the scenario considered above. In that case, according to Equation 19 the cores would only be able to hold onto ~ 0.2 Earth masses in gas, and we would require fast accretion to be halted before the dispersal of all of the disk gas.

We've proposed a dynamical scenario in which the ice giants would complete their core growth and accrete their gas envelopes in the time that the depleted disk dissipates. We emphasize that even if this scenario is incorrect, pebble accretion introduces a fine tuning problem in the standard scenario for the formation of Uranus and Neptune.

2.6 Summary and Discussion

Recent work on pebble accretion successfully resolves the timescale problem in growing gas giant cores at large orbital distances, but it gives rise to a fine tuning problem in the formation of the ice giants. In the standard scenario, the majority of solid accretion halts once the cores reach a certain mass, at which point the cores slowly begin accreting significant amounts of gas. However, pebbles, which are mm- to cm-sized particles observed to exist in the outer regions of protoplanetary disks, accrete extremely efficiently onto protoplanetary cores at the mass scale of the ice giants. They accrete so quickly that it is difficult for ice giant cores to exist at roughly their current core masses for a significant fraction of the disk lifetime, while accreting a gas envelope. To avoid runaway, they must therefore complete their growth at a specific time, when the gas disk is partially, but not entirely, depleted. Halting core growth at a specific point in time is fine tuned, and in this paper we propose a scenario for resolving this problem of fine tuning in the growth of Uranus and Neptune.

We suggest that the onset of the majority of gas accretion onto the planetary

cores coincides with a dynamical instability, where the gas was no longer able to damp the growing cores' random velocities. We suppose that the cores of Uranus and Neptune could have begun their growth closer to the gas giants. Then, at the point where $\Sigma_{\rm solid} \sim \Sigma_{\rm gas}$, their orbits were gravitationally excited and they may have traveled outward to their present locations. There, the ice giants' orbits circularized, and they accreted their $\sim 15\%$ atmospheres from the now-depleted nebula, avoiding atmospheric runaway.

We emphasize that the fine tuning problem exists regardless of the explanation. Since they never underwent runaway, the ice giant cores could not have grown to their current masses in a full gas disk, in which they can grow rapidly by accreting pebbles. As an alternative to our scenario, the cores may halt their growth if:

- there is no radial drift of pebbles into the ice giants' respective feeding zones. This is improbable, given protoplanetary disk observations of pebbles;
- the supply of pebbles in the outer disk is somehow cut off, which would require a cutoff mechanism;
- core masses are limited by pebble isolation *and* the grain opacity is higher than expected (such that $M_{\rm iso,peb} < M_{\rm crit}$ for the ice giants). This high opacity is unlikely, given the existence of the solar system's gas giants.

We predict that the occurrence of isolation-mass scale ice giants beyond 10 AU with 0.1 atmospheres by mass may be strongly correlated with the presence of gas giants in the system. If our proposed mechanism is correct, then gas giants—by

facilitating substantial migration during the epoch of instability—may play a key role in forming planets with intermediate-size atmospheres in the outer disk. Without gas giants, planets likely won't move around as much at our preferred time, so they are less likely to experience enhanced growth while the gas disk is dissipating. They could still grow to the scale of the ice giants after the gas disk is gone, but in that case, they would have small rather than intermediate-sized atmospheres. Therefore, we predict that when gas giants are not present, outer planets will host smaller atmospheres.

We are thankful to Jonathan Fortney, Jack Lissauer, Michael Rosenthal, and Yanqin Wu for helpful discussions. We thank the referee for comments that improved this paper. This work was supported by NSF Career grant number AST-1555385.

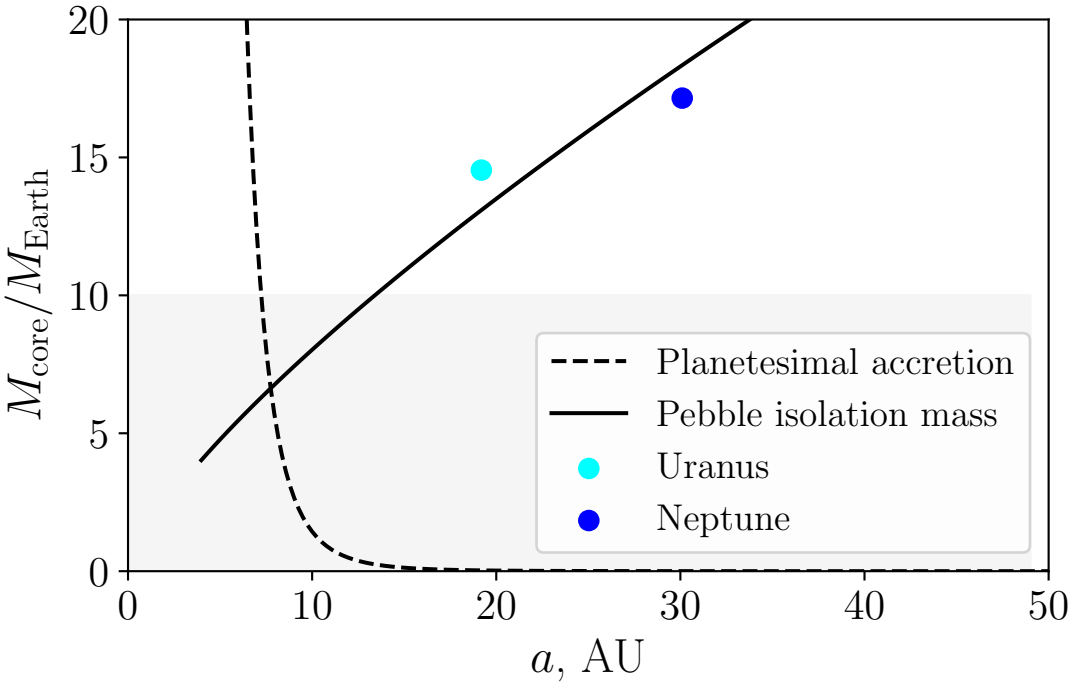

Figure 2.3 How massive can a core grow as a function of stellocentric distance (outside the terrestrial planet region)? Here, we speculate that pebble isolation sets the maximum core mass at distances comparable to those of the ice giants. Closer in, standard planetesimal accretion may set the core size. The maximum core mass that can be achieved by accreting planetesimals at random velocities of order the Hill velocity is shown (dashed line, Equation 2.20). This mass is limited by the supply of planetesimals, so plotted here is the (extreme) upper limit for standard growth. The solid line marks the pebble isolation mass (Equation 2.4, scaled by a factor of 1/2), which is an upper limit for core growth via pebble accretion. The shaded region below 10 $M_{\rm Earth}$ represents (roughly) the scale at which cores are sub-critical.

Chapter 3

Signatures of a planet-planet impacts phase in exoplanetary systems hosting giant planets

Abstract

Gas Giants in Extreme Environments

by

Renata Frelikh

Exoplanetary systems host giant planets on substantially non-circular, close-in orbits. We propose that these eccentricities arise in a phase of giant impacts, analogous to the final stage of Solar System assembly that formed Earth's Moon. In this scenario, the planets scatter each other and collide, with corresponding mass growth as they merge. We numerically integrate an ensemble of systems with varying total planet mass, allowing for collisional growth, to show that (1) the high-eccentricity giants observed today may have formed preferentially in systems of higher initial total planet mass, and (2) the upper bound on the observed giant planet eccentricity distribution is consistent with planet-planet scattering. We predict that mergers will produce a population of high-mass giant planets between 1 and 8 au from their stars.

3.1 Introduction

Observations of exoplanetary systems have found many gas giants with orbital distances less than 5 au from their host stars. The orbits of these planets are often eccentric, deviating from the roughly circular orbits in our Solar System. Several mechanisms have been proposed to account for these eccentricities, including: secular chaos [94], the Kozai-Lidov mechanism [e.g. 17, 44, 49, 59, 82], secular oscillations due to an outer planetary companion [67], resonant interactions with a gas disk [13], and planet-planet scattering [e.g. 12, 41, 68, 72, 23, 21]. However, these studies have not yet explained an important clue to these systems' dynamical histories: most observed planets with eccentricities $e > 0.5$ are more massive than Jupiter, while lower-mass planets are confined to lower eccentricities [92]. This observation is potentially surprising, since lower-mass planets are excited to higher eccentricities in typical dynamical simulations [e.g. 12, Figures 14 and 20].

We explore the possibility that some stars initially host multiple hydrogen-rich planets in their inner systems, which go through a giant-impacts phase, analogous to the final stage of inner Solar System assembly that resulted in Earth's moon-forming impact. Collisions cause these planets to grow in mass. Though in a given system, the lower-mass planets are more likely to have higher eccentricities, when viewed as a population, the trend reverses. Stars hosting more total mass in planets produce more high-mass giants, which in turn are able to excite each other to high orbital eccentricities. We use the observed mass distribution for planets at semimajor axes $a < 5$ au to construct a

distribution of initial total planet masses in that region. We demonstrate, using stellar metallicity as an observational proxy for initial mass in planets, that our ensemble produces giant-impact phase outcomes consistent with observations.

3.2 Observational Sample

We compare our simulations to a sample of observed exoplanets obtained from the Exoplanet Orbit Database on April 3, 2019, hosted on exoplanets.org [93]. We include 311 planets from 243 stellar systems, discovered via the radial-velocity method, orbiting FGK stars (0.5 to 1.4 $M_{\odot}$). Following Dawson & Murray-Clay [18], we remove 70 evolved stars with log g < 4. We exclude 11 planets for which no data for log g is reported. We further exclude 27 planets with no reported eccentricity. We note that all of the masses in the observational data set are reported as $M \sin i$, which henceforth we refer to as the mass. The planets in this sample are subject to several biases, including radial-velocity selection biases toward higher-mass, close-in planets [32, 28]. It is not appropriate to make a statistical comparison between the simulated and the observed data sets, as there is a risk that the observational sample used is non-uniform. We therefore focus on the qualitative features of the distribution.

3.3 The Highest-Eccentricity Planets are Found Around Metal-rich Stars

Figure 3.1A displays the distribution of eccentricity as a function of planetary mass for the observational sample. The highest-eccentricity planets tend to be more massive, contradicting our intuition that the lower masses are more readily excited. Figure 3.1B provides a hint to the solution. After dividing the data into planets orbiting metal-rich stars ([Fe/H]>0, blue) and metal-poor stars ([Fe/H]<0, red), we find that planets with the highest observed eccentricities orbit stars that are preferentially metal-rich[1]. Therefore, if metal-rich stars tend to produce a greater quantity of high-mass planets, these giants would scatter their planetary companions to higher eccentricities. This is a key point of our Letter.

Shifting our focus to the orbital-separation distribution, several further observational features become apparent. In Figure 3.2A, there is an upper limit to the eccentricities in the inner ~ 1 au, which we call the eccentricity envelope (black solid line). A mass segregation in semimajor axis is seen when we separate the planets into higher-mass ($M_{\rm p} > 1.17\ M_{\rm J}$, orange) and lower-mass ($M_{\rm p} < 1.17\ M_{\rm J}$, purple), when compared to the median planet mass of $1.17 M_{\rm J}$ found in the observational sample. Higher-mass planets are preferentially found with $a > 0.5$ au. Furthermore, dividing the planets by metallicity in Figure 3.2B reveals that the planets approaching the ec-

[1]The Solar [Fe/H]=0 cutoff is employed because the occurrence rate of giant planets rises steeply for stellar metallicities above [Fe/H]=0 [73, 19]. 76 planets are observed in systems with [Fe/H]<0, and 235 - in systems with [Fe/H]>0. The median metallicity in our observational sample is [Fe/H]=0.14.

centricity envelope (blue band) are more likely to come from higher-metallicity systems, following the trend from Figure 3.1. In summary, the shape of the eccentricity distribution, together with its mass- and metallicity-dependent features, are important clues to the formation of these systems.

3.4 Eccentricities as a Result of a Giant-impacts Phase

The observational features that we attempt to match with our model are: (1) the eccentricity envelope, (2) the mass-eccentricity relation, and (3) the correlation between eccentricity and stellar metallicity. We do this by making a non-standard assumption that the planets are the result of collisional growth in systems initially consisting of multiple giant planets that underwent a giant-impacts phase, at least in systems where eccentricity excitation occurs. If one is willing to make that leap, all three features can be recovered. We plot our simulation results in Figure 3.1C, Figure 3.1D, Figure 3.2C, and Figure 3.2D.

Our proposed giant-impacts phase is analogous to a dynamical instability that occurred in Solar System formation, where collisions assembled the inner, rocky planets. In general, a gravitational encounter between two protoplanets leads to either a physical collision or a scattering that results in a velocity deflection. Which is statistically more likely to occur depends on the relative velocities of the bodies. Recall that the collision cross section goes as $\sigma_{\rm coll} \sim \pi R^2(1 + v_{\rm esc,p}^2/v^2)$, while the strong scattering cross section goes as $\sigma_{\rm scat} \sim \pi R^2(v_{\rm esc,p}/v)^4$. Here, R is the physical radius, $v_{\rm esc,p}$ is the planet's

escape velocity, and v is the protoplanets' relative velocity. Thus, for small velocities $v < v_{\rm esc,p}$, $\sigma_{\rm coll} < \sigma_{\rm scat}$ and scattering dominates. As v increases and approaches $v_{\rm esc,p}$, the cross section for strong scattering approaches the physical cross section, πR^2, which is always less than or approximately equal to $\sigma_{\rm coll}$. Therefore, for sufficiently large v, $\sigma_{\rm coll} > \sigma_{\rm scat}$, and the planets are more likely to collide than be deflected further [e.g. 68]. This limits the v that can arise from scattering to, at most, $\sim v_{\rm esc,p}$. The Solar System's inner planets have $v_{\rm esc,p}$ ($\sim$ 11 km/s for Earth) less than the orbital escape velocity, $v_{\rm esc,\ star}$ ($\sim$ 42 km/s at 1 au). In general, they cannot readily eject each other through gravitational encounters, allowing for the existence of a giant-impacts phase. Conversely, the Solar System's giant planets have $v_{\rm esc,p}$ ($\sim$ 24 km/s for Neptune, $\sim$ 60 km/s for Jupiter) larger than $v_{\rm esc,\ star}$ ($\sim$ 8 km/s at Neptune, $\sim$ 19 km/s at Jupiter), making a giant-impacts phase unlikely. However, exoplanetary systems hosting close-in giants are in the regime where $v_{\rm esc,\ star} > v_{\rm esc,p}$, allowing giant impacts to leave a detectable imprint on the giant exoplanet mass distribution (Figure 3.3).

3.5 Methods

To generate our numerical results, we perform N-body simulations using Mercury6 [10], including growth by planet-planet collisions with mass and momentum conservation. Our planet radii follow nominal mass-radius relations constructed from the literature [24]. After each collision, we calculate the planet's mass and reassign its radius using these expressions. The reason we do this is because the escape velocity is

critical for the dynamical evolution. The radii for planets below 10 Earth masses ($M_{\rm E}$) were calculated using the mass-radius relation for a rocky core in the absence of a gas envelope, Fortney et al. [24, Equation 7]. Planets with masses between 10 and 200 $M_{\rm E}$ were assigned radii based on the theoretical models of Fortney et al. [24, Table 2] for a 10 $M_{\rm E}$ core surrounded by a H/He envelope at 300 Myr. Planets above 200 $M_{\rm E}$ were assigned 1.1 Jupiter radii ($R_{\rm J}$), which is consistent with the models, which show that the radius of Jupiter-sized planets stays roughly fixed above a Jupiter mass. The computed radii R, for planets of mass $M_{\rm p}$, in $M_{\rm E}$, are as follows:

$$R = \begin{cases} (0.16(\log_{10}M_{\rm p})^2 + 0.74\log_{10}M_{\rm p} + 1.12)R_{\oplus} & \text{if } M_{\rm p} < 10M_{\oplus} \\ (-1.0\times10^{-6}M_{\rm p}^4 + 1.8\times10^{-4}M_{\rm p}^3 - 1.1\times10^{-2}M_{\rm p}^2 & \\ +0.28M_{\rm p} - 1.7)R_{\rm J} & \text{if } 10 < M_{\rm p} < 28M_{\oplus} \\ (-7.2\times10^{-11}M_{\rm p}^4 + 8.1\times10^{-8}M_{\rm p}^3 - 3.2\times10^{-5}M_{\rm p}^2 & \\ +5.5\times10^{-3}M_{\rm p} + 0.8)R_{\rm J} & \text{if } 28 < M_{\rm p} < 200M_{\oplus} \\ 1.1R_{\rm J} & \text{if } M_{\rm p} > 200M_{\oplus}. \end{cases} \tag{3.1}$$

The planet densities ρ in the simulations were then set to be: $\rho = 3M/(4\pi R^3)$. We note that our mass-radius relation is only an approximation. Collisions cause inflation and they may not result in a perfect merger. For example, Hwang et al. [30] have shown that collisions between sub-Neptune-sized bodies may result in significant atmospheric mass loss. Given these results, collision outcomes between gas giant planets

merit future study.

For all of the simulations presented in this Letter, the initial eccentricities were set to 0, the mean anomalies and longitudes of the ascending node were drawn randomly from a uniform distribution between 0 and 360 degrees. We define the limits of the simulation to be at 1000 au from the host star, beyond which planets are removed from the simulation. The initial timestep was set to be the minimum of 3 days or 1/15 of the orbital period of the closest-in planet, and the accuracy parameter was set to 10^{-12}. We set the radius of the central body to 0.005 AU, the mass to 1 Solar mass, and the Hybrid integrator changeover distance to 3 Hill radii. All simulations were run for 2×10^7 years, a reasonable time beyond which instabilities become much less frequent [12, 41]. The eccentricities reach a statistical equilibrium after this timescale [41]. We found that our features of interest were already imprinted at 1 Myr, and persisted at 5, 10, and 15 Myr snapshots. We expect that integration for the several-Gyr ages, typical of observed systems, would reduce the final number of planets, but only by a modest amount.

The main results of this work are presented for a set of 694 numerical integrations, the initial inclinations i for which were drawn randomly from a uniform distribution between 0 to 1 degree.

We call the initial total mass in planets in each system the "disk mass". In constructing our simulations, we first considered a single disk mass. Each disk mass has a corresponding eccentricity envelope, which is higher for greater disk masses because they typically form higher-mass planets. Thus, the disk mass is what determines how high the

eccentricities can get. As a result, higher-mass planets have higher eccentricities across systems, contrary to our expectation for a single system. The correlation of eccentricity and metallicity in the observational data suggests that the observed distribution reflects a collection of varied disk masses. Therefore, to reproduce the observational data, we need to construct a set of simulations with initial disk masses drawn from a distribution. We do not know a priori what this distribution is. However, if we choose a distribution that is a few times the observed planet mass distribution, our (Figure 3.3) set of systems evolves to match the observed distribution of planet masses, after ejections and mergers. We choose to draw the disk masses $M_{\rm disk}$ from the exponential distribution,

$$f(M_{\rm disk}/M_{\rm J}, 1/\beta) = 1/\beta {\rm e}^{-M_{\rm disk}/(\beta M_{\rm J})}, \tag{3.2}$$

with the scale parameter $\beta = 20$ set to produce the planet mass distribution. We re-draw any disks with masses $< 0.1 M_{\rm J}$, as our disks need to be able to produce giant planets. An upper disk mass cutoff of 50 $M_{\rm J}$ is employed. Our simulations are not otherwise tuned. Each disk mass is allocated into planets in the following way: 10 planet masses are first drawn from a uniform distribution from 0 to 1, then each planet mass is scaled, such that the sum of the planet masses in each system is equal to the disk mass. The planets are uniformly distributed in $\log(a)$ between 0.03 and 10 au. To justify the logarithmic spacing in semimajor axis, we recall that the number of planetary Hill radii, $R_{\rm H}$, that can dynamically fit into a given semimajor axis range is: $a/R_{\rm H} = (M_{\rm p}/(3M_{\rm star}))^{-1/3}$. Here, the Hill radius, $R_{\rm H} = a(M_{\rm p}/(3M_{\rm star}))^{1/3}$, sets the distance scale from the planet at which gravitational interactions with other planets become strong. Since this ratio

is independent of a, each successive bin in $\log(a)$ from the star can fit the same number of planetary $R_{\rm H}$ as the one prior, and our uniform spacing in $\log(a)$ is appropriate.

As long as the planets were dynamically spaced close enough to interact, the final outcome was not significantly dependent on the way we distributed the disk mass into planets. For example, a 20 $M_{\rm J}$ disk, allocated into 7 planets, each with mass 2.9 $M_{\rm J}$, produced similar results to a 20 $M_{\rm J}$ disk allocated into 50 planets with mass 0.4 $M_{\rm J}$. Both typically produce planets with masses up to $\sim$ 10 $M_{\rm J}$ after collisions. We chose 10 planets to allow the planets to have multiple close encounters within the first $\sim$ 10 thousand years of the simulation, after which the number of planets decreased. We have found, in agreement with previous work [e.g. 41, Figures 1 and 11], that although the planet multiplicity generally dropped further as we extended the time to 10^8 years, and the individual planet orbital properties varied with time, our overall features of interest in the distributions of eccentricity with mass and semimajor axis persisted. The final planet mass distribution from the simulations (Figure 3.3B) was consistent with the mass distribution in the observational sample (Figure 3.3A), suggesting that our starting guess for the disk masses was reasonable. We note that planet migration before the gas disk dissipates may mean that fewer planets per system may be required, if it causes planets to migrate in and experience dynamical instability.

For our simulation results, plotted in Figures 3.1C, 3.1D, 3.2C, 3.2D, 3.3B, and 3.3C, we consider planets observable if they have $a < 6$ au and radial velocities >2 m/s, which roughly matches the typical precision of radial-velocity surveys contributing to our observational sample [8]. Radial velocities, K, are estimated by $28.4(M_{\rm p}/{\rm M_J})(T/{\rm yr})^{1/3}$

m/s, with $M_{\rm p}$ - the planet mass and T - the orbital period [52]. We comment that including the eccentricity dependence $(1-e^2)^{-1/2}$ in the expression for K does not significantly change our results. To determine the likelihood of observing systems with multiple planets, we "observe" each resulting system of our set of simulations from 1000 random directions. We evaluate the observability of each planet by computing the radial velocity component along the line of sight to our random "observer", i.e. we multiply the mass of each planet by sin$i$ in the radial-velocity formula. We find that $\sim 70\%$ of systems have an observable multiplicity of 1 or 2, $\sim 20\%$ have no observable planets per system, and $\sim 10\%$ have 3 or more. On average, ~ 1.6 planets can be observed per system. Our simulated distribution of Msin$i$, generated by observing each system from random angles as described above, is shown in Figure 3.3C.

Trends with stellar metallicity that arose in the observed populations were of particular interest (Figures 3.1 and 3.2). Stellar metallicity is not a parameter in our simulations. However, we can make a simple assumption that a higher total planet mass is correlated with stellar metallicity, which is reasonable given that high-metallicity stars are more likely to host close-in planets [19], and eccentric warm Jupiters are preferentially found in high-metallicity systems [18]. To map the synthetic distribution of disk masses containing observable planets to the observed distribution of stellar metallicities, we first assume a linear relationship between $\log_{10}(M_{\rm disk}/M_{\odot})$ and [Fe/H] of the form $[{\rm Fe/H}] = a\log_{10}(M_{\rm disk}/M_{\odot}) + b$, where $M_{\odot}$ is the Solar mass. We then find that $a = 0.4$ and $b = 0.9$ in the formula give us the best fit, such that the range of observed metallicities matches the range of assigned metallicities from our simulations. Therefore, we

convert the disk mass from our simulations to metallicity using the following relation: $[\mathrm{Fe/H}] = 0.4 \log_{10}(M_{\mathrm{disk}}/M_{\odot}) + 0.9$. We note that we do not expect there to exist a perfect one-to-one correspondence between stellar metallicity and disk mass, and our nominal relation is only suggestive.

3.6 Discussion

Figure 3.1B displays the distribution of eccentricities as a function of planet mass in our simulations. Contours correspond to enclosed probabilities. The upper contour encloses 90% of the points, and each successive contour below is a 10% decrement in enclosed probability from the one above. The contours are defined the same throughout all of the figures in this Letter, to allow one-to-one comparison of the simulations and the observations. We remind the reader that the only tuning required was the choice of an initial disk mass distribution to match the observed planet mass distribution, and a sufficient number of planets to ensure dynamical interaction. Our simulations in Figure 3.1D reproduce the observational feature seen in Figure 3.1C: $e > 0.6$ planets have $M_{\mathrm{p}} > 0.5\ M_{\mathrm{J}}$ and [Fe/H]>0. Higher-metallicity systems have more mass available for planet formation, which leads to more and larger planets that can excite each other to higher eccentricities.

In Figure 3.2, both the observations (A) and the simulations (C) are bounded by the eccentricity envelope. This is the upper limit to the eccentricities produced by planet-planet scattering, which deflects the planetary velocities in random directions.

The orbit deviation from circular is often called the "random" velocity, and can be approximated as $v_{\rm r} = ev_{\rm K}$, where $v_{\rm K} = ({\rm G}M_{\rm star}/a)^{1/2}$ is the orbital Keplerian velocity, G - the gravitational constant, and $M_{\rm star}$ - the stellar mass [114]. Since scattering can only excite $v_{\rm r}$ up to $v_{\rm esc,p}$ (Section 4), the eccentricity is limited to: $v_{\rm r} = ev_{\rm K} \approx ev_{\rm esc,p}$. This yields a maximum eccentricity of $e = v_{\rm esc,p}/v_{\rm K} \approx v_{\rm esc,p}/v_{\rm esc,star}$. Here, $v_{\rm esc,star} = (2{\rm G}M_{\rm star}/a)^{1/2}$ is the escape velocity from the star at the planet's a, which is comparable to $v_{\rm K}$. The eccentricity envelope curves in Figure 3.2 are

$$e = v_{\rm esc,p}/v_{\rm esc,star} = (M_{\rm p}a/(RM_{\rm star}))^{1/2}, \tag{3.3}$$

shown for $M_{\rm p} = 1.17M_{\rm J}$, the median observed planet mass. Planets in higher-metallicity systems are scattered to a greater range of eccentricities than those with lower metallicity (blue band in Figure 3.2B and D).

Because we are interested in the implications of planet-planet scattering for planets at all semimajor axes, not just those that are currently observable, we perform two additional sets of simulations with 20 planets uniformly distributed in $\log(a)$ between 0.03 and 100 au. The outcome for a single disk mass of 5 $M_{\rm J}$ is presented in Figures 3.4A and B for a set of 192 simulations. For each realization of the 20-planet systems, we re-drew the semimajor axes uniformly in $\log(a)$, the inclinations uniformly from 0 to 1 degree, and the mean anomalies and longitudes of the ascending node uniformly from 0 to 360 degrees. In Figure 3.4A, the higher-mass products of planet-planet collisions are seen to the left of the teal solid line, which marks $v_{\rm esc,p} = v_{\rm esc,star}$. Lower-mass planets are seen to the right, where scattering is a more likely outcome of

planet-planet interactions. The detailed mass distribution is shown in (B). Green, red, and blue dots show the maximum, median, and mean masses, respectively. The vertical blue lines display one standard deviation about the mean. The highest-mass planets typically form interior to the distance where $v_{\rm esc,p} = v_{\rm esc,star}$.

The result for a distribution of disk masses is presented in Figure 3.4C and D for a set of 175 simulations. We draw the disk masses from the exponential distribution (Equation 3.2), scaled by a factor of 1.4 to account for the additional bin in $\log(a)$. We employed an upper disk mass cutoff of 70 $M_{\rm J}$. The 1-8 au region hosts planets of higher mass. Interestingly, adding mass to the outer disk reduces the number of planets in the inner regions of higher-mass disks. The mass needed beyond 10 au to match the mass segregation feature from Figure 3.2A, while reproducing the metallicity feature from Figure 3.2B, merits future study.

We find it illustrative to repeat our main simulations in two dimensions, to avoid potential complications from Kozai oscillations. We start out with 10 planets using the same set of initial disk masses as described in Section 3.5, except we initialize the inclinations to 0. The starting semimajor axes of the planets are randomly distributed in $\log(a)$ from 0.03-1 au, and from 1-10 au, with the density reduced by a factor of 5 in the inner region. This selection is appropriate to match the observed mass-segregation feature in Figure 3.2A. We plot the results for this set of 180 simulations in Figure 3.5, for direct comparison to Figures 3.1 and 3.2. The features of interest are recovered in these two-dimensional simulations, suggesting that they are a result of planet-planet scattering. In the absence of mutually inclined orbits, the planets experience more

collisional growth, especially in higher-mass disks, where the planets are dynamically spaced close enough to be more likely to interact. The smaller planets in the higher-mass disks get absorbed by collisions more frequently (Figure 3.5A) than the mutually inclined planets in Figure 3.1C. The eccentricity vs. semimajor axis distribution in Figure 3.5D reproduces the observational feature seen in Figure 3.2B: the planets from higher-metallicity systems extend above the lower-metallicity region in a blue band on the plot. We plot the same distribution, instead dividing the planets by mass, in Figure 3.5C. The initial reduction in planet number density in the inner region allowed us to reproduce the mass segregation feature seen in the observations (Figure 3.2A). This mass reduction may be appropriate if giant planets do not form as easily in the inner 1 au of protoplanetary disks. On average, ∼1.8 observable planets are left per system in our 2D simulations.

3.7 Conclusion

We have suggested a giant-impacts phase in the evolution of giant exoplanetary systems, which creates a population of higher-mass planets in the collisional growth region (peaking at ~ 3 au for a Solar-mass star; see Figure 3.4D). Beyond, we predict a population of lower-mass planets that avoided mergers, some of which were scattered out on high-eccentricity orbits. This is consistent with the results of the Gemini Planet Imager Exoplanet Survey (GPIES) [60], who found that planets more massive than $\sim 3M_{\rm J}$ around Solar-type stars are mostly found in the radial-velocity regime ($a < 5$

au), rather than in the direct-imaging regime (a >10 au). Future microlensing and direct imaging surveys, sensitive to finding planets outside of the distance limits of radial-velocity surveys, will enable further observational tests of this predicted mass separation in the outer giant exoplanet population.

As evidence of mergers could correlate with orbital eccentricities, we speculate that inflated Jupiters would exist on preferentially eccentric orbits. Though collisions can happen on any timescale, most occur early, so this correlation would be most likely observable for stars with ages less than the $\sim$ 100 Myr cooling time for giants [24]. In addition, since some warm Jupiters have substantial cores [85], it is reasonable to suggest that early-on in their evolution planetary systems could have consisted of multiple planets, which experienced collisional growth.

We thank Jonathan Fortney, Yanqin Wu, and Kassandra Anderson for helpful discussions. We thank Bruce Macintosh for pointing us to the GPIES results of Nielsen et al. [60]. This work made use of the Exoplanet Orbit Database at exoplanets.org. RMC and RF are supported by NSF CAREER grant AST-1555385.

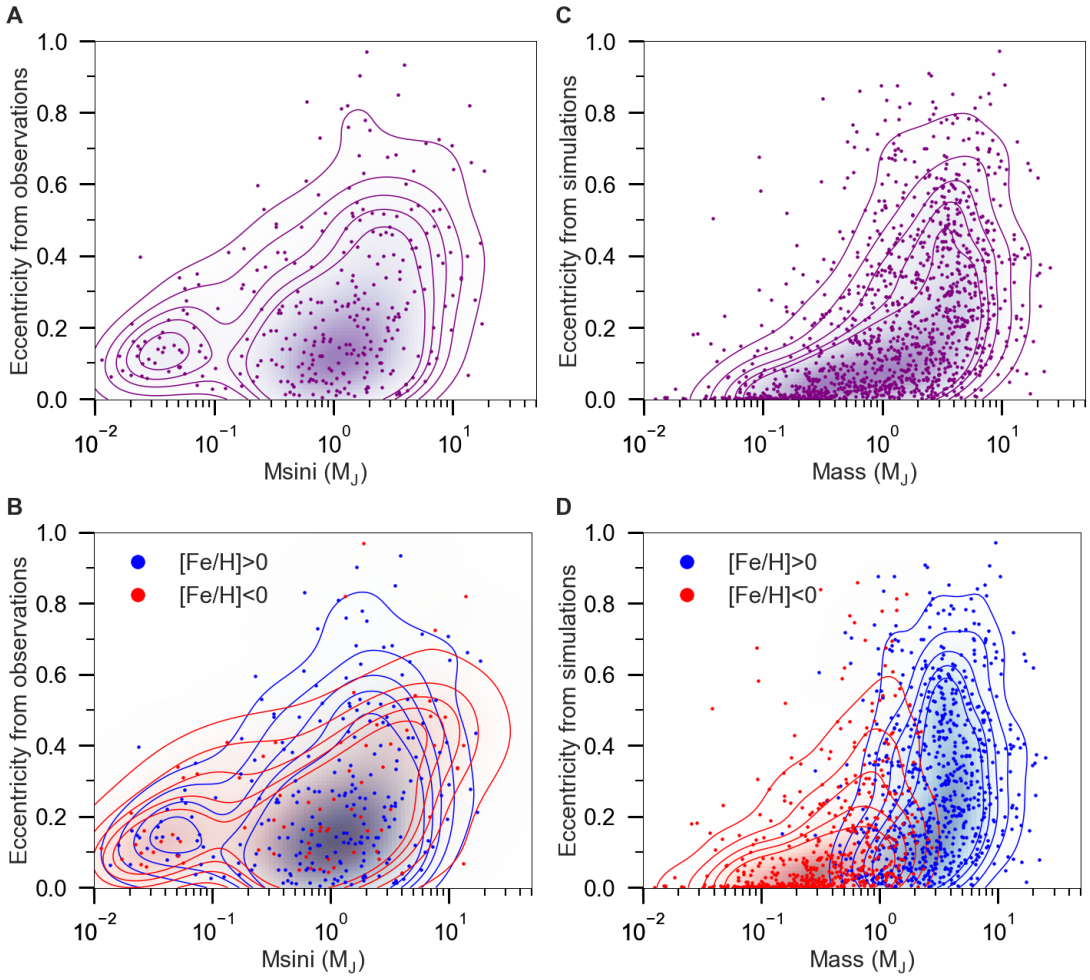

Figure 3.1 Distribution of eccentricities as a function of planet mass for (A) & (B) observations, and (C) & (D) simulations. The eccentricity of observed exoplanets is correlated with planet mass (A), which is recovered in our simulations (C). In (B), planets orbiting metal-rich ([Fe/H]>0, blue, 235 points) stars exhibit a range of eccentricities. Planets orbiting metal-poor stars ([Fe/H]<0, red, 76 points) are confined to $e < 0.6$. The highest-eccentricity planets ($e > 0.6$) are giant planets with $M_{\rm p} > 0.5$ Jupiter masses ($M_{\rm J}$).

Figure 3.1 This is matched in (D), in which we use total initial mass in planets as a proxy for metallicity (see Section section 3.6). On all panels, the upper contour denotes an enclosed probability of 90%, and each successive contour below is a 10% decrement to the one above.

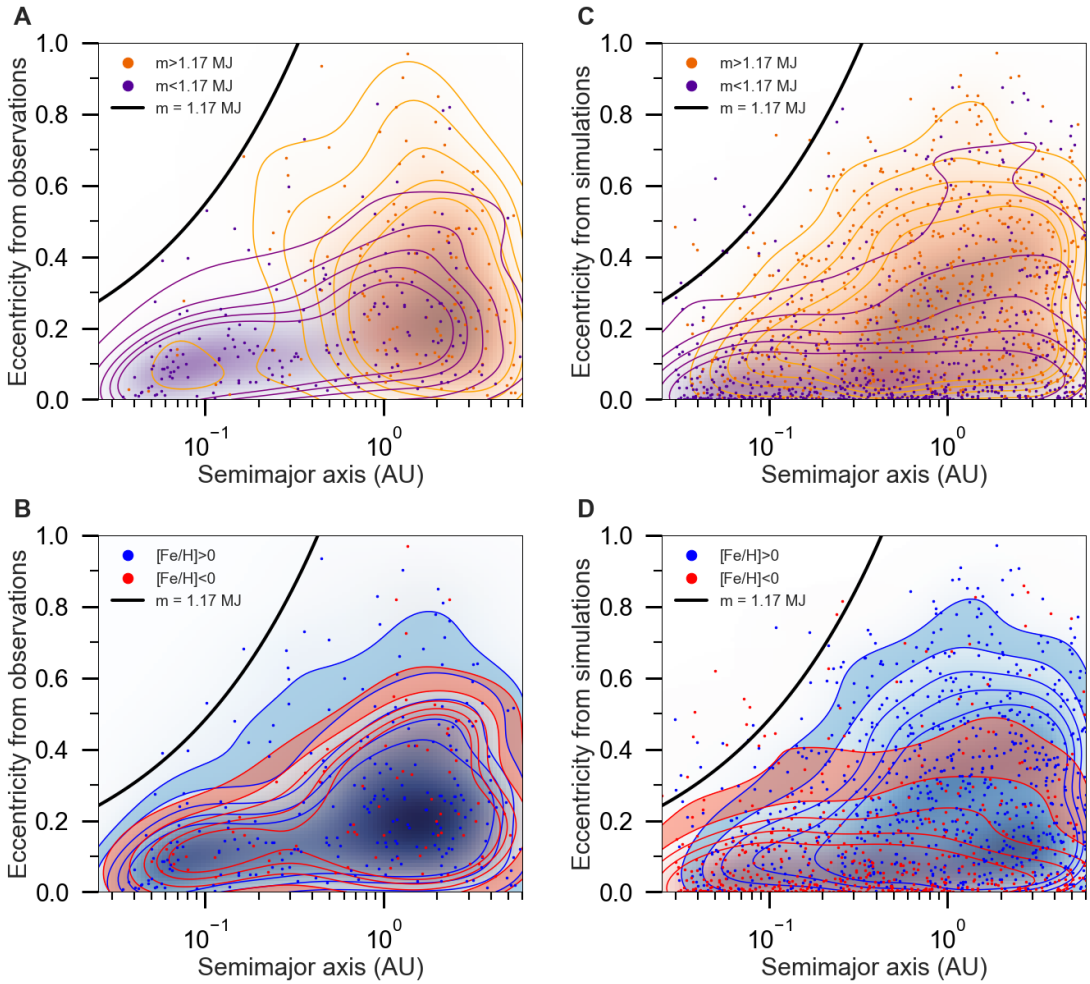

Figure 3.2 Eccentricity vs. semimajor axis. (A) Observational eccentricity distribution. The high-mass giant planets ($m > 1.17\ M_{\rm J}$, orange) exhibit eccentricities $0 < e \lesssim 1$, and semimajor axes generally beyond 0.3 au, while low-mass giants (purple) span 0.03-6 au in a. This is matched by the simulations in (C), though the mass segregation in a is weaker. The eccentricity envelope (black solid line) denotes the planet-planet scattering limit to the eccentricity (Equation 3.3).

Figure 3.2 In (B) and (D), we divide the planets into those belonging to systems orbiting metal-rich ([Fe/H]>0, blue) and metal-poor ([Fe/H]<0, red) stars. Note that observed planets in (B) with eccentricities approaching the eccentricity envelope (blue band on plot) preferentially come from systems with super-solar metallicity ([Fe/H]>0), which is matched by our simulations (D). The contours are defined as in Figure 3.1.

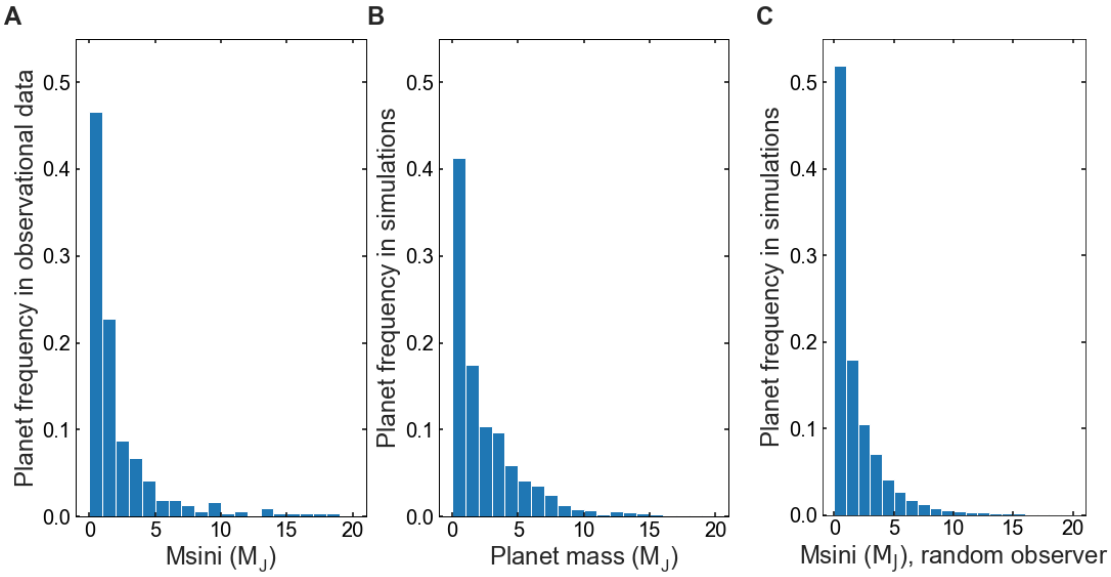

Figure 3.3 Distribution of planet masses corresponding to Figures 3.1 and 3.2 for (A) observations and (B) & (C) simulations. (C) displays the distribution of Msin$i$ in our simulations that would be measured by the radial-velocity method, as seen from a random "observer" (see Section 3.5)

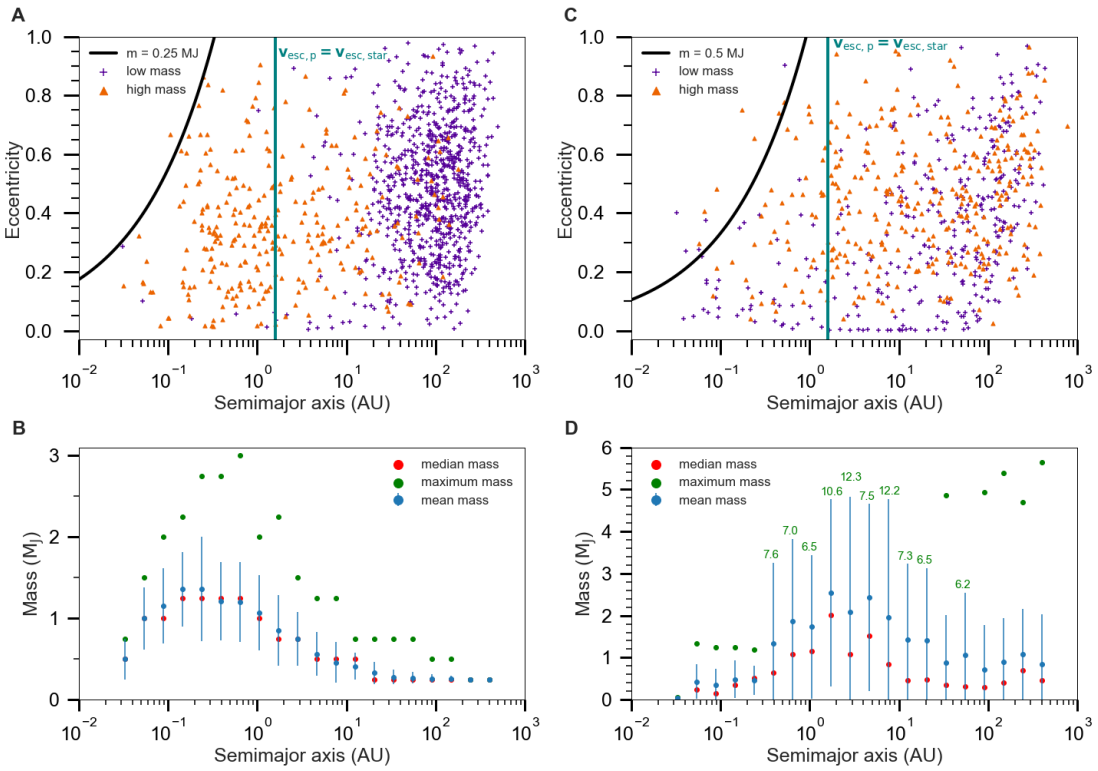

Figure 3.4 Eccentricities and masses vs. semimajor axis for simulations starting with 20 planets in the semimajor axis range 0.03 to 100 au. (A) Initial planet masses are all 0.25 M_J. The inner, higher-mass planets (orange triangles) grew via planet-planet collisions. The lower-mass (purple pluses), outermost planets never collided. The vertical teal line at 1.6 au marks where the escape velocity of a 0.25 M_J Saturn-radius planet is equal to the orbital escape velocity of the planet from the star. (B) shows the corresponding distribution of planet masses, including the median (red dot), mean (blue dot), and maximum (green dot) masses. The vertical blue bars mark one standard deviation about the mean.

Figure 3.4 Most high-mass planets form interior to the distance where $v_{\rm esc,p} = v_{\rm esc,star}$. (C) and (D) provide corresponding plots to (A) and (B), respectively, for a set of initial disk masses designed to match the distribution used in Figures 3.1-3.3 (see text), but adjusted to incorporate the additional bin in $\log_{10}(a)$ compared to the results shown in Section 5. In (D), the values for the maximum planet masses that extend beyond the axis range are printed in green by the error bars corresponding to each bin in semimajor axis. The highest-mass products of collisional growth have $1 < a < 8$ au.

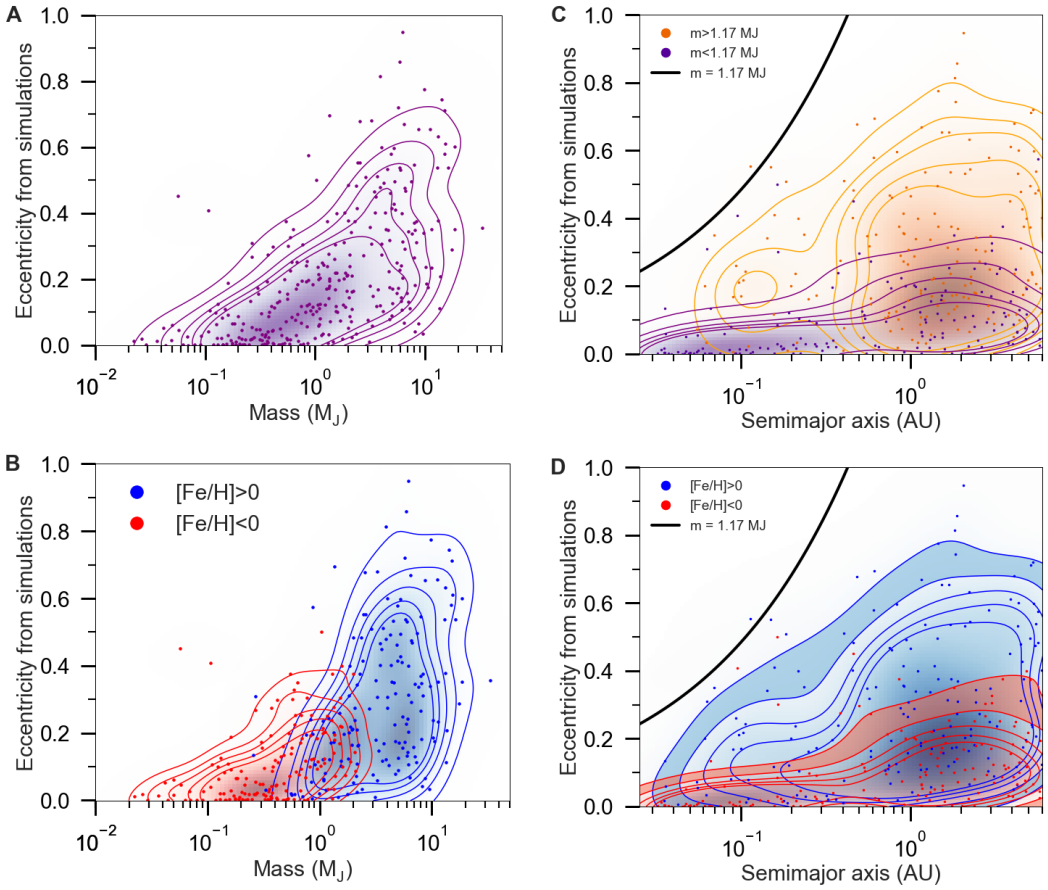

Figure 3.5 (A) Our 2D simulations reproduce the correlation between mass and eccentricity in the distribution. (B) Planets orbiting high-metallicity ([Fe/H]>0, blue) stars are able to be excited to a higher eccentricities than planets orbiting low-metallicity stars ([Fe/H]<0, red). Metallicities are used as a proxy for the initial disk masses (Section 3.6). (C) and (D) Distribution of eccentricities as a function of semimajor axis in our 2D simulations. In (C), the planets of higher mass (orange) occur more frequently beyond ∼0.5 au, and can reach higher eccentricities than the lower-mass (purple) planets, which are confined to $e < 0.4$.

Figure 3.5 In (D), planets orbiting high-metallicity stars are excited to higher eccentricities than planets orbiting low-metallicity stars throughout the entire semimajor axis range where they occur. The contours are defined as in Figure 3.1.

Chapter 4

Efficiency of Hydrodynamic Escape

Abstract

Gas Giants in Extreme Environments

by

Renata Frelikh

We develop a flexible one-dimensional code to model the escape of hydrogen and helium from a hot Jupiter as a result of photoionization from extreme-ultraviolet (EUV) radiation. We include stellar spectrum heating and ionization, radiative cooling by Lyman-α and H_3^+, tidal gravity, and a H-He reaction network. For our fiducial hot Jupiter, we uncover a three-layer structure: an H_3^+-cooled layer of molecular hydrogen at the base, enveloped by a Lyman-α-cooled layer of neutral hydrogen, which transitions into an ionized wind layer that is cooled by adiabatic expansion. The highest spectral energy photons are deposited in the molecular layer - they are cooled the most via H_3^+. We run a grid of models, varying the distance of our fiducial planet from the star. We find that Lyman-α cooling increases in significance with increasing stellocentric distance, diverging the mass-loss rate from the energy-limited mass-loss regime substantially. At increasing stellocentric distances, the outflow becomes increasingly more neutral. The neutral hydrogen starts to decouple from the ionized outflow, free-streaming out and potentially decreasing the mass-loss rate from our estimates. This two-component structure to the outflow is interesting, both in terms of evaluating where the wind is cut off, and in terms of observational signatures.

4.1 Introduction

Atmospheric escape is known to affect the atmospheres of hot Jupiters, gas giants on extremely close-in orbits from their host stars. Winds reaching velocities of order km s^{-1} are driven from these strongly irradiated, heated planets, resulting in atmospheric mass loss as their hydrogen-dominated atmospheres are ionized. However, H_3^+ and Lyman-α are known to radiatively cool atmospheres of such planets. We aim here to investigate in depth how radiative cooling affects the photoevaporative outflow process, and how the effect varies with distance from the star. For that reason, we need to resolve the transition from the molecular hydrogen that forms deeper in the atmosphere, shielded from extreme ultraviolet (EUV) radiation, to the ionized hydrogen that forms the wind.

EUV radiation drives hydrodynamic escape from planets with hydrogen- dominated thermospheres by photoionizing hydrogen and heating the upper layer of the atmosphere [91]. This escaping gas has not only been inferred from population analyses (e.g., Owen & Wu 66), but it has been seen directly for larger planets. Atmospheric escape has been measured for hot Jupiter HD 209458b (e.g., Vidal-Madjar et al. 88, Vidal-Madjar et al. 89). Vidal-Madjar et al. 88 observed that, as the planet transits its host star, the Ly-α emission line from the star is reduced significantly more than expected purely from the planet obscuring part of the stellar surface. This suggests that a cloud of neutral hydrogen extending out to $\sim$4 planetary radii surrounds the planet. There is evidence for metals in the upper atmosphere extending beyond the

Roche lobe (i.e. the region surrounding a planet within which material is gravitationally bound to it) of the planet in transit (e.g. Vidal-Madjar et al. 89, Linsky et al. 50). HD 189733b (Des Etangs et al. [20] and Bourrier et al. [7]) is another hot Jupiter for which an escaping atmosphere has been observed in Lyman-α transit, with a possible detection of O [4]. Tentative detections of metals have been reported for WASP-121b (Sing et al. 77, Cubillos et al. 15). Escaping neutral H has also been observed in the H-α line (e.g. Cabot et al. 9 in Wasp-121b, Jensen et al. 38 in Wasp-12b, Yan et al. 98 in Wasp-33b, and Wyttenbach et al. 95 in Kelt-9b - with additional Balmer lines). He has been detected in the escaping atmosphere of WASP-107b [78], as well as for a warm Neptune around GJ 3470 [61]. This is strong evidence for neutral hydrogen and possibly metals to be undergoing hydrodynamic escape.

Theoretical work on atmospherical escape include models of escaping hydrogen and helium atmospheres bathed in EUV radiation, in addition to possible metals entrained in the wind. Two important classes of hydrodynamic models are core-powered mass loss (e.g., Ginzburg et al. [33], Ginzburg et al. [34]) and photoevaporative mass loss driven by stellar EUV and/or X-ray radiation (e.g. Lammer et al. [46], Baraffe et al. [2], Yelle [99], García-Muñoz [27]). We focus our work on photoevaporative mass loss. Prior theoretical studies of stellar-irradiation driven escape from Jupiter-sized extrasolar planets close to their host stars can be categorized into steady-state models (e.g., Murray-Clay et al. [58], Oklopčić & Hirata [63]) or time-dependent models, either one-dimensional (e.g., García-Muñoz [27], Yelle [99]), two-dimensional (e.g., Owen & Adams [65], Trammell et al. [86]), or three-dimensional (e.g., Koskinen et al. [43],

Tripathi et al. [87], McCann et al. [53]).

The goal of our work is to understand what fraction of a planet's incoming stellar flux is radiatively cooled, and what fraction is available to power the hydrodynamic outflow. It is necessary to understand the interface layer from molecular hydrogen gas (that is efficiently radiatively cooled) to the photodissociated and/or thermally dissociated atomic layer that comprises that outflow. To this end, we couple photochemistry to an outflow model. While this transition from the radiatively cooled molecular layer to the atomic outflow is key to understanding hot Jupiter atmospheric escape, it may be even more important for smaller planets, whose atmospheres are significantly shaped by these outflows, which are likely to have higher metallicities and stronger molecular coolants. We built a self-consistent hydrodynamic photochemical model that explicitly addresses the transition from the radiatively cooled molecular layer to the outflow, and how that feeds back into the efficiency of incoming stellar EUV radiation powering atmospheric escape.

We now discuss how our model relates to a few of the aforementioned models, to understand where it fits in the context of the broader set of models. For example, García-Muñoz [27] and Yelle [99] created one-dimensional time-dependent models that both introduce the general structure of hot Jupiter thermospheres: a molecular layer of hydrogen with interesting photochemistry that transitions into an atomic outflow. Bathed in EUV radiation, the hot Jupiters retained a molecular layer of hydrogen that dissociated into atomic hydrogen. García-Muñoz [27] found that the escape can be supersonic. As also discusssed in Yelle [99], the H_3^+ ion was found at the base of the

modeled atmosphere, which is an effective radiative coolant for energy deposited in that region. They found that the escape approaches the energy-limited (Watson et al. 91, Lammer et al. 46) regime for strongly irradiated hot Jupiter planets. However, Lyman-α cooling was not included as part of the radiative coolants in this study. Nevertheless, our model qualitatively agrees with both models in terms of the structure: a molecular layer at the base cooled by H_3^+, which is thermally disssociated into atomic hydrogen, which outflows as an ionized wind. Both models found outflows reaching of order 10^4 K, which were cooled via adiabatic expansion above the H_3^+ layer, however, Yelle [99] did not model the region where the outflow reaches supersonic velocities. A one-dimensional steady-state atmospheric escape model has been developed by Murray-Clay et al. [58], in which the equations of hydrodynamics are coupled with the ionization equation for atomic hydrogen and solved for via numerical relaxation. This study focused on the region above the molecular layer, specifically on the outflow structure. Murray-Clay et al. [58] found that Lyman-α cooling can be significant, particularly for higher stellar ionizing fluxes. In addition, Koskinen et al. [43] found that H_3^+ radiative cooling can significantly limit outflows. In summary, the time-dependent models we mentioned here provide a base in understanding the general thermospheric structure, while the steady-state model provides a detailed understanding of the outflow, in particular, including the Lyman-α cooling in the interface region. Our model incorporates both frameworks: to understand the efficiency of hydrodynamic escape driven by EUV radiation, it is necessary to have a detailed model of the molecular-atomic interface, as well as the outflow region. Run to steady state, our time-dependent model produces atmospheric

profiles as a function of height. We analyze the results of our model in terms of where in the atmosphere the EUV photons of different energies are absorbed, as well as which cooling terms dominate. Deposited near the base of our atmosphere, the highest-energy photons in our spectrum are radiatively cooled by H_3^+. Lyman-α cooling is important in the atomic-hydrogen-dominated layer just above the H_3^+-cooled region near the base. As we move the planet out in distance from the star, Lyman-α cooling becomes a more significant cooling source. We find that hydrodynamic outflows can be driven out to 0.9 au, as long as they're not shut off by magnetic field confinement from the planet. We model the outflow along the substellar ray. As long as the wind is not magnetically confined, the energy will be redistributed around to the nightside of the planet.

We discuss the methods in Section 4.2, where we describe our model. Additional details of the implementation are in the Appendix. Section 4.3 contains the results for our fiducial hot Jupiter at 0.05 au, as well as a grid of models with varying distance from the star. There we discuss the structure of the atmosphere, including the molecular layer and the outflow, and which parts of the spectrum dominate in the atmosphere at different heights. For the grid, we discuss the mass-loss rates and compare them to the energy limited regime. We conclude with Section 4.4.

4.2 Methods

To address the question of efficiency of atmospheric escape, we developed a photochemical-hydrodynamic model that allows us to simulate a planetary outflow from

a hydrogen- and helium-dominated atmosphere. The atmospheric composition is initialized to a gas consisting of molecular hydrogen and atomic helium in ratios corresponding to Solar abundance (the mass fraction X=0.75 for H, and Y=0.25 for He [51]). We used a modified the low-flux solution of Murray-Clay et al. [58] as the starting point for our fiducial planet. For Jupiter-sized planets with deep potential wells, the profiles need to be initialized with an initial velocity greater than zero for stability purposes. The choice of initial conditions can be relatively arbitrary - the initial profiles do not have to be self-consistent to start, just enough to be able to evolve and avoid situations that code might have difficulty with, such as a snowplow phase [53]. For smaller-sized planets such as Neptune and Earth analogs, one may simply initialize the profiles to isothermal hydrostatic.

We solve the coupled equations of hydrodynamics and chemical network for a planet being irradiated by extreme ultraviolet (EUV) photons in the 165.3-12 eV range by time evolving the species' number density, bulk velocity, and specific internal energy of the gas to steady state. The input spectrum used for ionization and subsequent heating is from Richards et al. [74], a solar EUV flux model that we scale for the corresponding distance from the star. This spectrum is appropriate for modeling planets around evolved stars, as X-rays do not dominate the outflow. Our code is flexible: an alternate spectrum may be used to model planets around stars early in their evolution, as well as expanded into the far-ultraviolet range to include molecular photochemistry that may be especially important in smaller-sized planet atmospheres.

A planet atmosphere can be modeled as a hydrodynamic outflow only if the

gas remains collisional up to the sonic point. In other words, the gas can be treated as a fluid. For the gas to be collisional, its mean free path $\lambda = 1/(n\sigma)$ must be less than the scale height of the atmosphere $H = k_B T/(\mu g)$, with μ being the mean molecular weight and $g = GM_p/{R_\mathrm{P}}^2$ - the gravitational acceleration. The collisional cross section σ is either the Coulomb cross section $\sigma_C = 3.899 \times 10^{-14}(T/10^4)^{-2}$ cm^2 (for ion-ion scattering) or the hard body cross section $\sigma_{hb} = 3.519 \times 10^{-16}$ cm^2 (for neutral-neutral scattering), depending on whether the gas is mostly ionized (in the wind, for a high stellar flux) or mostly neutral (at the base). We confirm that our fiducial case presented in Section 4.3.1 is collisional throughout. We discuss collisionality in detail in Section 4.3.8.

The methods section is structured as follows: the equations of hydrodynamics are desribed in Section 4.2.1, the numerical method used for solving them in Section 4.2.5, and the chemical reactions and solver in Section 4.2.8 and Appendix A.3. The spectrum and a discussion of the photoionization processes are given in Section 4.2.9.

4.2.1 Hydrodynamics

The hydrodynamic equations are the conservation of mass, momentum, and energy. Written together, these comprise the Euler equations. In conservation form, the mass, momentum, and energy conservation equations are (e.g., Clarke & Carswell 14):

$$\frac{d\rho}{dt} + \vec{\nabla} \cdot (\rho \vec{u}) = 0, \tag{4.1}$$

$$\frac{d(\rho\vec{u})}{dt} + \vec{\nabla}\cdot(\rho\vec{u}\,\vec{u} + P\mathbf{I}) = -\rho\vec{\nabla}\Psi, \tag{4.2}$$

$$\frac{d(\rho e)}{dt} + \vec{\nabla}\cdot(\vec{u}(\rho e + P)) = -\rho\vec{u}\cdot\vec{\nabla}\Psi + \rho(\Gamma - \Lambda). \tag{4.3}$$

Here, ρ is the mass density of all of the species of the atmosphere, and $\vec{u}$ is their bulk velocity. We note that the $\vec{\nabla}\cdot$ operator in Equation (4.2) is the tensor divergence, which in one-dimensional spherical coordinates is just $\hat{\mathrm{r}}\,\partial/\partial r$. Here, $\vec{u}\,\vec{u}$ is the velocity outer product, $\mathbf{I}$ is the identity matrix, and P is the scalar pressure that is isotropic. The term e in Equation 4.3 is known as the specific internal energy - the internal energy per unit mass (ρe is the internal energy density, i.e. the internal energy per unit volume). Finally, Γ is a heating source term, and Λ is a radiative cooling source term. The heating comes from the EUV spectrum (Section 4.2.9), and the cooling included is Lyman-α cooling (Section 4.2.4) and infrared cooling due to H_3^+ emission (Section 4.2.3)[1]. Bolometric heating and cooling forces the temperature in the lower boundary region, which we interpret to be the temperature set by bolometric heating and cooling in Section 4.2.2.

Equation (4.1) states that the mass in a fixed grid cell changes only as a result of fluid flowing in or out, in the absence of sources or sinks. Equation (4.2) states that the momentum in a cell changes as a result of momentum flowing in or out, the pressure gradient, and external forces which are written in terms of the gradient of a

[1]Recombination cooling contributes to radiative energy loss, with the rate $\Lambda_{rr} \simeq 6.11 \times 10^{-10} k_B T^{0.11} n_{H^+} n_e/\rho$ erg s^{-1} [64]. We confirm in post-processing that it is negligible for the parameter space of our models, so it is turned off in our models.

scalar potential, $-\rho\vec{\nabla}\Psi$. If the force is gravity, then $\Psi = -GM_P/r$, and the force term is $-\rho GM_p/r^2\,\hat{\mathrm{r}}$. It is included on the right-hand side of the momentum equation as an external force, and on the right-hand side of the energy equation as a potential energy source term. Equation 4.3 states that the internal energy of a cell changes as a result of energy flux through the cell boundary, PdV work, heating and cooling source terms, and advection through a potential field. The total energy density E of the gas incorporates a kinetic energy term due to the bulk velocity, in addition to the internal energy term: $E = 1/2\rho u^2 + \rho e$. The ideal gas law relates the pressure to the density:

$$P = Nk_{\mathrm{B}}T = \rho\frac{k_{\mathrm{B}}T}{\mu} = \rho\frac{e}{\gamma - 1}, \tag{4.4}$$

where k_{B} is the Boltzmann constant, μ is the mean molecular weight of the gas, and $N = \sum_s n_s$ is the total number density of the gas. The mean molecular weight $\mu = \sum_s n_s m_s/N$, where m_s and n_s are the molecular weight and number density of species s, respectively. Thus, the specific internal energy is related to the pressure via: $P/\rho = e/(\gamma - 1)$. The adiabatic index $\gamma = 5/3$ for a monatomic ideal gas is used throughout.

As we are working in one dimensional spherical coordinates, Equation 4.1 can be written

$$\frac{\partial\rho}{\partial t} + \frac{1}{r^2}\frac{\partial}{\partial r}(\rho u r^2) = 0 \tag{4.5}$$

An individual mass conservation equation can be written for each species s of our atmosphere. We add a chemical source term S on the right-hand side of the mass conservation equation to denote the creation and destruction of species in our

chemical network, which we solve for in a separate substep of our routine. The diffusive velocities u_s and v_s correspond to molecular and eddy diffusion, respectively. A detailed discussion of diffusion is presented in Section A.2.

$$\frac{\partial \rho_s}{\partial t} + \frac{1}{r^2}\frac{\partial}{\partial r}(\rho_s u r^2) + \frac{1}{r^2}\frac{\partial}{\partial r}(\rho_s (u_s + v_s) r^2) = S \tag{4.6}$$

The bulk density $\rho = \sum_s \rho_s$ is conserved.

Equation (4.5) can be rewritten as:

$$\frac{\partial \rho}{\partial t} + u\frac{\partial \rho}{\partial r} = -\frac{\rho}{r^2}\frac{\partial (u r^2)}{\partial r} \tag{4.7}$$

Similarly, we write Equation 4.2. The last term on the right-hand side is the tidal gravity term, derived in Appendix A.1:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial r} = -\frac{GM}{r^2} - \frac{1}{\rho}\frac{\partial P}{\partial r} + \frac{3GM_\star a}{r^3} \tag{4.8}$$

Finally, Equation 4.3 can be written (see Appendix A.5):

$$\frac{\partial e}{\partial t} + u\frac{\partial e}{\partial r} = -\frac{P}{\rho r^2}\frac{\partial (u r^2)}{\partial r} + \Gamma - \Lambda \tag{4.9}$$

4.2.2 Bolometric heating and cooling

The base of our atmosphere is optically thick to incoming EUV radiation. We implement a bolometric heating and cooling term in this region, as the EUV heating is unable to balance the cooling processes to keep this region isothermal. Our atmosphere is least opaque to the highest-energy photons in our spectrum (165.3 eV). Thus, our

calculation is only strictly physical in the region above approximately $\tau = 1$ for the 165.3 eV photons.

In the bolometric region, the temperature is set by the balance between heating by the entire spectrum of the star and blackbody radiation at the effective temperature of the planet (in the infrared (IR)). We do not model the IR emitters in detail and instead parameterize the opacity by $\kappa = 1$, which is an aritficially high value for the opacity. A more accurate model would include dust sinking as well and contributions to the opacity from molecular lines and atoms. Thus, a detailed calculation of the transition between the bolometric region and the optically thin region, implementing appropriate opacities for the optical and IR photons, is left for future work.

The bolometric heating term heats the layer below $\tau = 1$ to the equilibrium temperature at a given distance from the star. The effective temperature is set by a balance between stellar heating and thermal infrared cooling. We neglect dust heating by re-absorption of IR from the emitters, either dust or molecules. Per mass, the effective heating rate is then: $Q_{\rm bol} = F n_{\rm absorbers} \sigma_{\rm absorbers} / \rho_{\rm absorbers} = F \kappa_{\rm absorbers} = \kappa_{\rm absorbers} \sigma_{\rm SB} T_{eq}^4$. The cooling term radiates at the effective temperature as a blackbody with a cooling flux per unit area $F_{\rm cool} = \sigma_{\rm SB} T_{eff}^4$, which per mass can be written as: $\Gamma_{IR} = F_{\rm cool} n_{\rm absorbers} \sigma_{\rm IR} / \rho_{\rm absorbers} = \kappa_{\rm IR} \sigma_{\rm SB} T_{eff}^4$. The net bolometric heating and infrared cooling term is then written as:

$$Q_{\rm bol} - \Gamma_{IR} = \kappa \sigma_{\rm SB} (T_{eq}^4 - T_{eff}^4). \tag{4.10}$$

4.2.3 H3+

The H_3^+ ion, also known as trihydrium, is a significant coolant in gas giant planet thermospheres. In the astronomical context, it was first detected in Jupiter's auroral atmosphere by infrared emission. It has also been observed to be present in the atmospheres of the Solar System giants Saturn and Uranus. However, it has not yet been detected in an exoplanet.

It is primarily produced from the hydrogen molecule by a chain of two chemical reactions. First, molecular hydrogen is ionized by a photon with energy above the ionization threshold to produce the dihydrogen cation: $H_2 + h\nu \rightarrow H_2^+ + e^-$. This ionized molecular hydrogen then reacts with neutral molecular hydrogen to form H_3^+: $H_2^+ + H_2 \rightarrow H_3^+ + H$. The main destruction pathway is dissociative recombination, the outcomes of which are twofold: $H_3^+ + e^- \rightarrow H_2 + H$ and $H_3^+ + e^- \rightarrow H + H + H$.

Due to its structure, the H_3^+ molecule is an efficient emitter in the infrared. For a detailed review of the vibrational modes contributing to the infrared spectrum of H_3^+, see Miller et al. [57] and references therein. We implement the H_3^+ cooling function calculated by Miller et al. [56]:

$$\begin{aligned} E(30 < T < 300K) = &- 81.9599 + 0.886768T - 0.0264611T^2 + 0.000462693T^3 \\ &- 4.70108 \times 10^{-6}T^4 + 2.84979 \times 10^{-8}T^5 \\ &- 1.03090 \times 10^{-10}T^6 + 2.13794 \times 10^{-13}T^7 \\ &- 2.26029 \times 10^{-16}T^8 + 8.66357 \times 10^{-20}T^9 \end{aligned} \tag{4.11}$$

$$E(300 < T < 800K) = 92.2048 + 0.298920T - 0.000962580T^2 + 1.82712 \times 10^{-6}T^3 - 2.04420 \times 10^{-9}T^4 + 1.24970 \times 10^{-12}T^5 - 3.22212 \times 10^{-16}T^6 \quad (4.12)$$

$$E(800 < T < 1800K) = -62.7016 + 0.0526104T - 7.22431 \times 10^{-5}T^2 + 5.93118 \times 10^{-8}T^3 - 2.83755 \times 10^{-11}T^4 + 7.35415 \times 10^{-15}T^5 - 8.01994 \times 10^{-19}T^6 \quad (4.13)$$

$$E(1800 < T < 500K) = -55.7672 + 0.0162530T - 7.68583 \times 10^{-6}T^2 + 1.98412 \times 10^{-9}T^3 - 2.68044 \times 10^{-13}T^4 + 1.47026 \times 10^{-17}T^5 \quad (4.14)$$

The temperature-dependent cooling function allows us to express the H_3^+ contribution to the cooling term in our code as: $\Gamma_{H_3^+} = 4\pi e^{E(T)} n_{H_3^+} \times 10^7/\rho$. We note that the cooling term is zero for temperatures above the range provided in Equations (4.11) through (4.14), as we do not expect H_3^+ to exist there in any significant amounts due to its source H_2 being thermally and/or photo-dissociated in the regions of the atmosphere where these higher temperatures occur.

4.2.4 Lyman-α cooling

The other significant radiative cooling mechanism in our model is Lyman-α cooling, which occurs after neutral hydrogen atoms that are collisionally excited by

electrons de-excite. This cooling converts thermal energy into radiation via electron impacts. As our gas is partially ionized - neutral hydrogen and ionized hydrogen (the main contributor to the electrons) coexist at least near the base of the outflow, with the gas still being mostly neutral - and the gas is sufficiently heated via EUV radiation ionizing the gas - Lyman-α cooling is expected to be significant. We use the expression used by Murray-Clay et al. [58] from Black [5]:

$$\Lambda_{\mathrm{Ly}\alpha} = -7.5 \times 10^{-19} n_{\mathrm{H}} n_{\mathrm{e}} e^{-118348/T} \mathrm{erg\,cm^{-3}\,s^{-1}}, \tag{4.15}$$

where the density of electrons $n_{\mathrm{e}} = n_{\mathrm{H^+}} + n_{\mathrm{H_2^+}} + n_{\mathrm{H_3^+}} + n_{\mathrm{He^+}} + n_{\mathrm{HeH^+}}$ is the sum of the number densities of the ionized species (we enforce charge neutrality). We note that the Lyman-α photons must be able to escape in order for the cooling to be efficient. The gas is not optically thin towards the radiation in this energy; however, Murray-Clay et al. [58] show that these photons can freely escape after scattering into the Lyman-α line wings.

4.2.5 Hydrodynamic solver

We adopt the CIP (constrained interpolation profile) scheme [96] for solving the hydrodynamics equations. It minimizes numerical diffusion by keeping track of not only the cell-centered quantities, but also of their spatial gradients. As suggested by [96], we also adapt a staggered mesh to avoid the problem of damping oscillations caused by the decoupling of odd and even grid cells. Avoiding unnecessary dissipation improves mass and energy conservation, which is important especially for low scale heights at the

base of the atmosphere, where spatial concentration gradients are steep. This scheme was tested in the context of atmospheric escape by Kuramoto et al. [45], who found that it reproduces hydrodynamic outflow solutions with relatively high accuracy.

In Section 4.2.1, we split the fluid equations (Equations 4.7, 4.8, and 4.9) into an advective part on the left, and source terms on the right. The advective part will be solved via the CIP method. Then, the source terms on the right are explicitly marched forward in time with a finite differencing method. The diffusion terms are dealt with separately via the implicit Crank-Nicolson method, and the chemical network is solved for via the semi-implicit method from Numerical Recipes [71].

We solve the mass conservation equation of each of the 8 component species separately. We assume that the gas is thermalized, i.e. the temperatures of the species are the same. All of the species travel with the bulk velocity u, and diffusion is dealt with implicitly in the mass conservation equation.

We keep track of the variables ρ, u, and e, as well as ρ', u', and e', where the primes denote the gradient of the variables n, i.e. $n' = \partial n/\partial r$. The initial n' are calculated using a central difference formula, ${n_i}' = (n_{i+1} - n_{i-1})/(r_{i+1} - r_{i-1})$. A forward difference is used for the cell at the lower boundary, ${n_0}' = (n_1 - n_0)/(r_1 - r_0)$, and, similarly, a backward difference is used for the cell at the upper boundary, $n'_{-1} = (n_{-1} - n_{-2})/(r_{-1} - r_{-2})$.

The method of solving the hydrodynamic equations by splitting them into an advective and non-advective substep is known as operator splitting, which involves solving two separate sub-parts of the equation which are subsequently combined to advance

the solution one full timestep. In the following, we discuss the advective substep.

4.2.5.1 Advective sub-step

The advection equation is any equation of the form

$$\frac{dn}{dt} + u\frac{\partial n}{\partial r} = 0. \tag{4.16}$$

Here, n is the quantity that can be advected with the velocity profile u; in our case, the quantity can be ρ, u, or e.

In the CIP method, we wish to advance the solution from timestep $m-1$ to timestep m. The profile is known at the location of the cell centers of the grid at timestep $m-1$. To obtain the solution at timestep m, we need to map the cell-centered values back to their locations at the previous timestep $m-1$: $x_i^m - u\Delta t$. However, as the locations do not coincide with the cell centers, we need to reconstruct the quantity in each cell. To do that, we use a third-degree polynomial. Recall as we are also advecting the gradient of the profiles, we reconstruct its solution with a second-degree polynomial. The remapping process is what makes the code semi-Lagrangian, and the preservation of the gradient is what makes the code less dissipative than other methods. Appendix A.4.2 describes in detail how we reconstruct the profiles and their gradients within each cell, as well as how the solutions are mapped from the previous timestep.

4.2.5.2 Non-advective sub-step

The right-hand side of Equations 4.7, 4.8, and 4.9 can be expressed as a central difference. The staggered velocity grid makes it easy to express the divergence terms

in the mass and energy conservation equations. The mass conservation term can be written on a grid as:

$$\frac{\partial \rho_i}{\partial t} = -\frac{\rho_i}{r_i^2} \frac{u_{i+1} r_{h,i+1}^2 - u_i r_{h,i}^2}{r_{h,i+1} - r_{h,i}}. \tag{4.17}$$

The density is then updated via $\rho_i^{m+1} = \rho_i^m + (\partial \rho_i / \partial t)^m \Delta t$. Similarly, the non-advective derivative in the energy equation is discretized as:

$$\frac{\partial e_i}{\partial t} = -\frac{P_i}{\rho_i r_i^2} \frac{u_{i+1} r_{h,i+1}^2 - u_i r_{h,i}^2}{r_{h,i+1} - r_{h,i}} + \Gamma_i - \Lambda_i. \tag{4.18}$$

The first term on the right hand size is the work done by expansion (or contraction) of the gas. The second and terms, Γ_i and Λ_i are the heating and cooling terms per mass. The heating term encompasses the heating from EUV (extreme ultraviolet) photons from the star, which will be described below. The cooling term is radiative cooling: Lyman-α cooling and molecular cooling (H_3^+ infrared radiation).

4.2.6 Artificial viscosity

For stability, we add a numerical viscosity term Q_i to the pressure term, analagous to Yabe & Aoki [96] for the shock-tube problem. First, we calculate the divergence of the velocity: $\vec{\nabla} \cdot \vec{u} = ((r^2 u)_{i+1/2} - (r^2 u)_{i-1/2})/r_i^2$, as suggested by Yabe & Aoki [96]. Then, the artificial viscosity is given by a combination of two terms [48], linear and quadratic in the divergence of velocity:

$$\begin{aligned} Q_i &= 0.75 \left(-\rho_i \left(\frac{\gamma P_i}{\rho_i} \right)^{1/2} (\vec{\nabla} \cdot \vec{u})_i + \frac{\gamma + 1}{2} \rho (\vec{\nabla} \cdot \vec{u})_i^2 \right) \qquad \text{if } (\vec{\nabla} \cdot \vec{u}_i) < 0 \\ &= 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if } (\vec{\nabla} \cdot \vec{u})_i \geq 0. \end{aligned} \tag{4.19}$$

The first term, linear in $(\vec{\nabla} \cdot \vec{u})$, is required to stabilize shock waves that might arise, since the scheme reduces numerical dissipation [62]. It is the product of the density, adiabatic sound speed, and the divergence of the velocity. This term reduces short-wavelength numerical oscillations. The second term is a von-Neumann viscosity term [90] - it supresses longer-wavelength numerical oscillations. The tunable constant of order 1 in front of the viscosity term controls the strength of the artificial viscosity, 0.75 was chosen following Yabe & Aoki [96]. We note that our implementation is similar to that of the tensor artificial viscosity of Stone & Norman [79], where instead of the standard gradient implementation the divergence of velocity is used.

4.2.6.1 Grid

We solve Equations 4.7, 4.8, and 4.9 by time marching the density, velocity, and specific internal energy to a steady state on a one-dimensional non-uniform grid in spherical coordinates. We use a non-uniform grid to allow for a higher uniform resolution at the base, increasing the cell size with increasing altitude. The planetary radius $R_{\rm P}$ is set to 0.7 $R_{\rm J}$. The base cell is set to 10^5 cm, and the grid cell size is further set by a geometric progression, $\Delta r_i = 1.014\,\Delta r_{i-1}$. The grid spans 1 $R_{\rm P}$ to ~ 8.8 $R_{\rm P}$. This grid is used for the density and energy profiles - the variables are calculated at the cell centers. The velocity grid is staggered from the regular grid such that the cell-centered velocity values coincide with the cell boundaries on the regular grid. Staggering the grid this way allows for the exact computation of $\partial(ur^2)/\partial r$ on the right-hand side of the mass and energy conservation equations using adjacent cells. There are 2 ghost cells at

the base, and 2 ghost cells at the top, for a total of 504 cells.

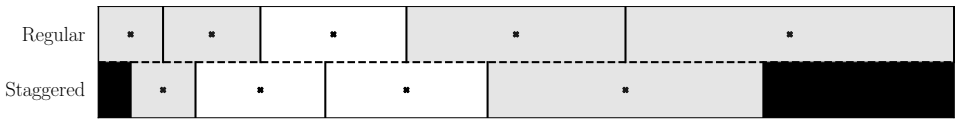

Figure 4.1 Illustration of a staggered grid with uneven spacing. The staggered grid cells are centered at the regular grid cell boundaries. The ghost cells are shaded in gray.

Using the same grid for density, velocity, and energy may result in high-frequency oscillations in the solution. The velocity is coupled to the pressure via the momentum equation. Ignoring the advective term on the left and the force terms on the right, the finite-difference form of Equation (4.8) for cell i on the grid is:

$$\frac{\partial u_i}{\partial t} = -\frac{1}{\rho_i}\frac{P_{i+1} - P_{i-1}}{r_{i+1} - r_{i-1}}, \tag{4.20}$$

where we have used the central difference method for calculating gradients. Note that P_i does not appear in the calculation of $\partial u_i/\partial t$ - even and odd cells are decoupled, which results in two separate modes in the solution that may appear as a zig-zag pattern (also called odd-even grid oscillation or odd-even decoupling). This is a numerical artifact that may impact the stability of the solution [83]. To mitigate this issue, the resolution may be increased (at a computational cost), a filter (which may impact accuracy) or artificial viscosity (which introduces numerical dissipation) may be used; another way is to use a staggered grid. That way, instead of Equation (4.20) we have:

$$\frac{\partial u_{i+1/2}}{\partial t} = -\frac{2}{\rho_i + \rho_{i+1}}\frac{P_{i+1} - P_i}{r_{i+1} - r_i} \tag{4.21}$$

with no decoupling, where we use half-indeces such as $i + 1/2$ to denote cells centered on the grid staggered by half a cell size from the cells i on the regular grid.

We stagger our velocity grid with respect to the regular grid, with the cell centers of the velocity grid matching up with the cell boundaries of the regular grid. There is 1 ghost cell at the base, and 1 ghost cell at the top, for a total of 503 cells. Figure 4.1 illustrates our setup for a single cell on the regular grid. The size of the first cell is set by $\Delta r_{h,0} = \Delta r_0$. The following cells are calculated via: $\Delta r_{h,i} = 2\,\Delta r_i - \Delta r_{h,i-1}$.

4.2.7 Boundary conditions

As the flow is supersonic, we use outflow boundary conditions for the number densities, velocity, and pressure at the outer boundary. For example, where x is ρ or e and N is the number of grid cells, we have $x_{N-2} = x_{N-3} + (x_{N-3} - x_{N-4})/(r_{N-3} - r_{N-4})(r_{N-2} - r_{N-3})$, and similarly for x_{N-1} and the velocity boundary condition. No boundary conditions have to be specified at the outer boundary if the flow is supersonic, as waves cannot enter the domain from the outer boundary. At the inner boundary, the flow is subsonic, so we specify the boundary conditions for the number densities of each species and the pressure. We extrapolate the velocity from the grid: $u_0 = u_1 - (u_2 - u_1)/(r_2 - r_1)(r_1 - r_0)$: a wave can travel from the domain to the inner boundary, as the flow is subsonic, so the value of the velocity is part of the solution and cannot be imposed as a boundary condition.

4.2.8 Chemical network

The chemical network of reactions (Table 4.1) includes the ionization processes that heat the gas and drive a wind. They also include the reactions that generate H_3^+, an important infrared coolant in the molecular layer. In addition, they determine how the wind transitions from the molecular to the atomic layer. The reactions were compiled by Yelle [99] and García-Muñoz [27]. The rate constants for the photoionization reactions are determined as in Section 4.2.9. For the 7 atomic/molecular species (+electrons) undergoing 22 reactions in our chemical network, we can write a set of differential equations governing the evolution of their concentrations with time. As the rates of the reactions vary by several orders of magnitude, the reactions network cannot be solved via a straightforward explicit method - the set of equations is said to be "stiff". We implement the semi-implicit method of solving stiff systems of equations from Press et al. [71], the details of which are given in Appendix A.3.

4.2.9 Spectrum

Here we describe the EUV spectrum of the star. We use the binned spectrum of Richards et al. [74], which provides the solar flux in 37 wavelength bins corresponding to a high level of stellar activity.

We calculate a wavelength-dependent optical depth, $\tau_\nu(r)$, for each spectral bin:

$$\tau_\nu(r) = \sum_s \int_{r_{\rm out}}^{r} n_s(r)\sigma_{\nu,s}, \tag{4.22}$$

Table 4.1. Chemical network

Label	Reaction	Rate	Reference
k1	$H + h\nu \rightarrow H^+$		Schunk & Nagy [76]
k2	$H^+ + e^- \rightarrow H$	$4.0 \times 10^{-12}(300/T)^{0.64}$	Storey & Hummer [80]
k3	$H_2 + h\nu \rightarrow H_2^+ + e^-$		Schunk & Nagy [76]
k4	$H_2 + h\nu \rightarrow H + H^+ + e^-$		Schunk & Nagy [76]
k5	$H_2^+ + e^- \rightarrow H + H$	$2.3 \times 10^{-8}(300/T)^{0.4}$	Auerbach et al. [1]
k6	$H_3^+ + e^- \rightarrow H_2 + H$	$2.9 \times 10^{-8}(300/T)^{0.65}$	Sundström et al. [81]
k7	$H_3^+ + e^- \rightarrow H + H + H$	$8.6 \times 10^{-8}(300/T)^{0.65}$	Datz et al. [16]
k8	$H_2^+ + H_2 \rightarrow H_3^+ + H$	2.0×10^{-9}	Theard & Huntress Jr [84]
k9	$H_2^+ + H \rightarrow H^+ + H_2$	6.4×10^{-10}	Karpas et al. [42]
k10	$H^+ + H_2 \rightarrow H_2^+ + H$	$1.0 \times 10^{-9} e^{-2.19\times10^4/T}$	Yelle [99]
k11	$H_3^+ + H \rightarrow H_2^+ + H_2$	2.0×10^{-9}	Yelle [99]
k12	$H_2 + M \rightarrow H + H + M$	$1.5 \times 10^{-9} e^{-4.8\times10^4/T}$	Baulch et al. [3]
k13	$H + H + M \rightarrow H_2 + M$	$8.0 \times 10^{-33}(300/T)^{0.6}$	Ham et al. [35]
k14	$He + h\nu \rightarrow He^+$		Schunk & Nagy [76]
k15	$He^+ + H_2 \rightarrow HeH^+ + H$	4.2×10^{-13}	Schauer et al. [75]
k16	$He^+ + H_2 \rightarrow H^+ + H + He$	8.8×10^{-14}	Schauer et al. [75]
k17	$HeH^+ + H_2 \rightarrow H_3^+ + He$	1.5×10^{-9}	Bohme et al. [6]
k18	$HeH^+ + H \rightarrow H_2^+ + He$	9.1×10^{-10}	Karpas et al. [42]
k19	$He^+ + e^- \rightarrow He$	$4.6 \times 10^{-12}(300/T)^{0.64}$	Storey & Hummer [80]
k20	$HeH^+ + e^- \rightarrow He + H$	$1.0 \times 10^{-8}(300/T)^{0.6}$	Yousif & Mitchell [100]
k21	$H^+ + H_2 + M \rightarrow H_3^+ + M$	3.2×10^{-29}	Miller et al. [55]
k22	$H_2 + h\nu \rightarrow H + H$		Schunk & Nagy [76]

where $\sigma_{\nu,s}$ is the total photoabsorption cross section of species s per wavelength ν, r is the radial coordinate, $n_s(r)$ is the number density of species s at a radial distance r, and r_{out} is the outer radius of the grid, where we take the optical depth to be zero.

Here we calculate the photoionization rate for species s by photons of wavelength ν and energy E_ν. The photons may be absorbed by species other than species s if the cross sections for absorption $\sigma_{abs,\nu}$ for these other species are non-zero. Therefore, we must allocate the ionizing flux in each wavelength bin between the species. Suppose the absorbing species in our gas are H_2, H, and He. Then, the fraction x_{H} absorbed by H can be written as [64]:

$$x_{\text{H}} = \frac{n_{\text{H}}\sigma_{abs,\nu,\text{H}}}{n_{\text{H}_2}\sigma_{abs,\nu,\text{H}_2} + n_{\text{H}}\sigma_{abs,\nu,\text{H}} + n_{\text{He}}\sigma_{abs,\nu,\text{He}}}, \tag{4.23}$$

and similarly for H_2 and He. Here, $n_s\sigma_{abs,\nu,s}$ is the number density of species s, and $\sigma_{abs,\nu,s}$ is the cross section for absorption of species s for photons of wavelength ν. To calculate the photoionization rate for species s in each cell, consider a cell of width Δr and cross-sectional area A (Figure 4.2) located at a depth r from the surface of the atmosphere of the planet.

Suppose that the incoming flux Φ_0 from the star is coming in from the left. This incoming flux $\Phi_0 = \Phi_\nu e^{-\tau_\nu(r)}$, where $\tau_\nu(r)$ is the optical depth at radius r given by Equation 4.22. The total number of photons that gets absorbed per volume in the cell is $\Phi_0(1 - e^{-\tau_{\nu,cell}})/\Delta r$, where $\tau_{\nu,cell} = (n_{\text{H}_2}\sigma_{abs,\nu,\text{H}_2} + n_{\text{H}}\sigma_{abs,\nu,\text{H}} + n_{\text{He}}\sigma_{abs,\nu,\text{He}})\Delta r$ for small Δr such that the number densities of all species are approximately constant within each cell. Then, of the photons that get absorbed, atomic hydrogen absorbs a fraction

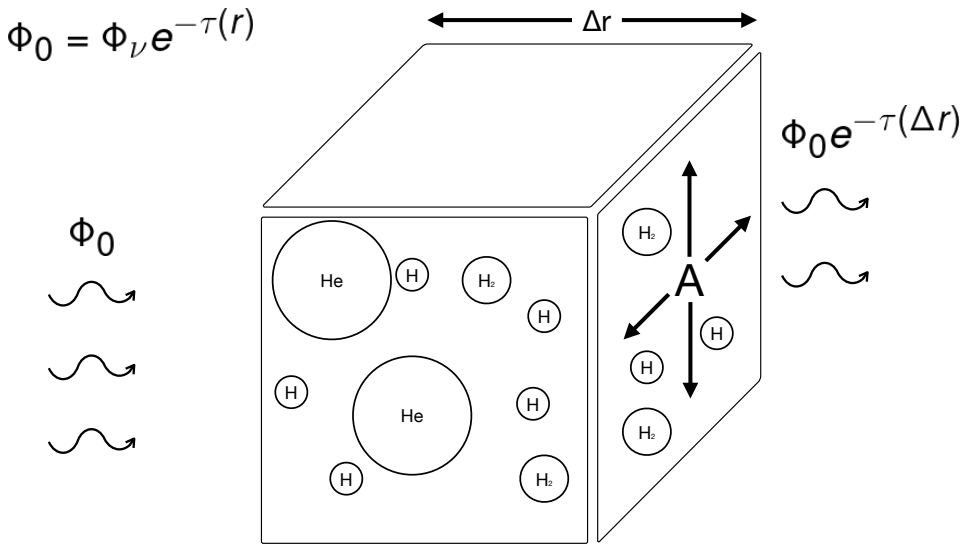

Figure 4.2 Grid cell of length Δr and cross-sectional area A containing absorbing species H, H_2, and He. Ionizing stellar flux of wavelength ν impacts the cell on the left. After being absorbed in the cell, the flux is reduced to $e^{-\tau}$ times the original flux, Φ_0, where τ is the optical depth through the cell for the particular wavelength of interest.

$x_{\rm H} = \tau_{\nu,{\rm H,cell}}/\tau_{\nu,cell}$ (Equation 4.23), where $\tau_{\nu,{\rm H,cell}} = n_{\rm H}\sigma_{abs,\nu,{\rm H}}\Delta r$ is the contribution to the optical depth in the cell due to atomic hydrogen. The rate at which photons are absorbed by atomic hydrogen is therefore:

$$\text{rate} = \frac{\mathrm{x_H}\Phi_0(1-\mathrm{e}^{-\tau_{\nu,\mathrm{cell}}})}{\Delta \mathrm{r}}. \tag{4.24}$$

For either absorbing species s, the reation rate is then:

$$\text{rate} = \mathrm{n_s}\sigma_{\mathrm{ion},\nu,\mathrm{s}}\frac{\mathrm{F}_\nu \mathrm{e}^{-\tau_\nu(\mathrm{r})}}{\mathrm{E}_\nu}\frac{(1-\mathrm{e}^{-\tau_{\nu,\mathrm{cell}}})}{\tau_{\nu,\mathrm{cell}}}. \tag{4.25}$$

Here, F_ν is the energy flux per bin of wavelength ν, E_ν is the average photon energy

in the bin (we assume all photons in a bin have the same energy), and $\sigma_{ion,\nu,s}$ is the photoionization cross section for the specific photoionization reaction, as for example H_2 photoionization can proceed via Reaction 3 or 4.

4.2.9.1 Optically thin limit

For computational purposes, we implement the following expression for the optically thin limit ($\tau_{\nu,cell} < 10^{-9}$): as $\tau_{\nu,cell} << 1$, $1 - e^{-\tau_{\nu,cell}} \approx \tau_{\nu,cell}$, reducing the rate expression to:

$$\text{rate} = \frac{\tau_{\nu,\text{H,cell}}}{\tau_{\nu,\text{cell}}}\Phi_0\frac{\tau_{\nu,\text{cell}}}{\Delta\text{r}} = \Phi_0 \text{n}_\text{H}\sigma_{\text{abs},\nu,\text{H}}, \tag{4.26}$$

which allows us to calculate the photoionization rate for either species s via:

$$\text{rate} = \text{n}_\text{s}\sigma_{\text{ion},\nu,\text{s}}\frac{\text{F}_\nu \text{e}^{-\tau_\nu(\text{r})}}{\text{E}_\nu} \tag{4.27}$$

Following the discussion in Section 4.2.9 on photoionization rates, we can calculate the heating rate due to the photoionization of H, H_2, and He, as well as the dissociative photoionization and dissociation of H_2. The optical depth due to photon absorption by H, H_2, and He in a grid cell is $\tau_{\nu,cell} = (n_{\text{H}_2}\sigma_{abs,\nu,\text{H}_2} + n_\text{H}\sigma_{abs,\nu,\text{H}} + n_\text{He}\sigma_{abs,\nu,\text{He}})\Delta r$. The incoming flux of photons of energy $h\nu$ at the cell boundary is $\Phi_\nu(r) = F_\nu e^{-\tau_\nu(r)}$. Per volume, rate of photoreactions j for species s is:

$$r_j = n_s\sigma_j\frac{F_\nu e^{-\tau_\nu(r)}}{E_\nu}\frac{(1-\tau_{\nu,cell})}{\tau_{\nu,cell}} \tag{4.28}$$

We then calculate the heating rate per volume:

$$\Gamma_\nu = \sum_j n_s\sigma_j\epsilon_j F_\nu e^{-\tau_\nu(r)}\frac{(1-\tau_{\nu,cell})}{\tau_{\nu,cell}}, \tag{4.29}$$

where the heating efficiency ϵ_j is given by:

$$\epsilon_j = \frac{(h\nu - E_{lim})}{h\nu}, \tag{4.30}$$

where E_{lim} is the threshold energy for the ionization of H (13.6 eV), He (24.6 eV), H_2 (15.4 eV), the disssociative ionization of H_2 (18.08 eV, Ford et al. [22]), and the dissociation of H_2 (4.74 eV, Herzberg & Monfils [36]).

4.3 Results

We discuss the results for our fiducial planet in Section 4.3.1. We show the density, velocity, pressure, and temperature profiles, as well as the composition by species with altitude. We discuss our model in detail, including the deposition of photons of varying spectral energy with altitude, and show how it produces three layers. Reaction rates involving H, H_2, and He are shown, both to demonstrate the convergence of the model and to highlight the key reactions involved in the molecular layer and the outflow. Heating and cooling rates are discussed, as well as the photoreactions that contribute to the heating. In Section 4.3.2, we discuss how the cooling terms are assigned to the heating terms, both by altitude and by wavelength. In Section 4.3.3 we introduce the heating efficiency parameter, ϵ, that we will use to compute the energy-limited mass-loss rate. Section 4.3.4 discusses the differences between various radiative modes of cooling turned on and off for our fiducial hot Jupiter at 0.05 au, and Section 4.3.5 contains the results of our grid of models, varying the semimajor axis of the planet from 0.015 to 0.9 au.

4.3.1 HD209458b

We model the upper atmosphere of our fiducial planet with parameters similar to that of the hot Jupiter HD209458b (Table 4.2). The steady-state profiles of our model are shown in Figure 4.3.

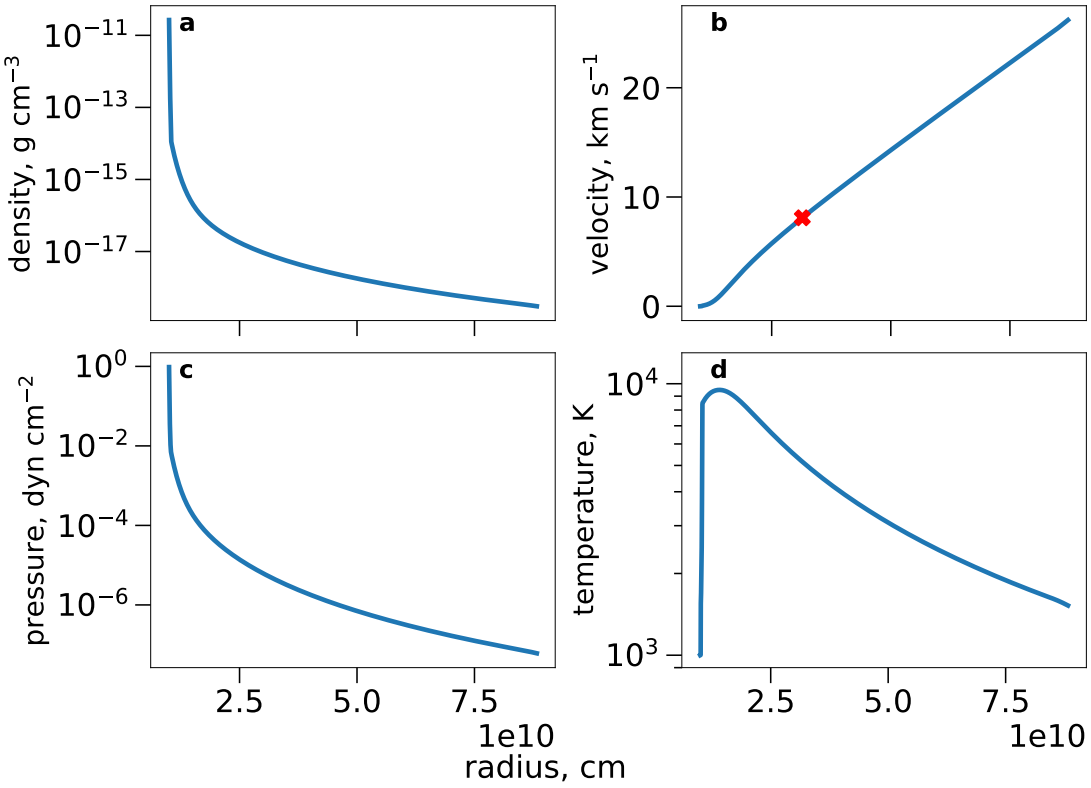

Figure 4.3 Panel (A) shows the density profile for fiducial hot Jupiter at 0.05 au. Panel (B) - its velocity profile. The red "X" marks the sonic point of the wind. Panel (C) displays the pressure profile, and panel (D) - the temperature profile.

The wind reaches a velocity of ~ 8 km s^{-1} at the sonic point distance of 3.1 $R_{\rm P}$. The maximum temperature reached is $\sim 10^4$ K, as the Lyman-α cooling prevents

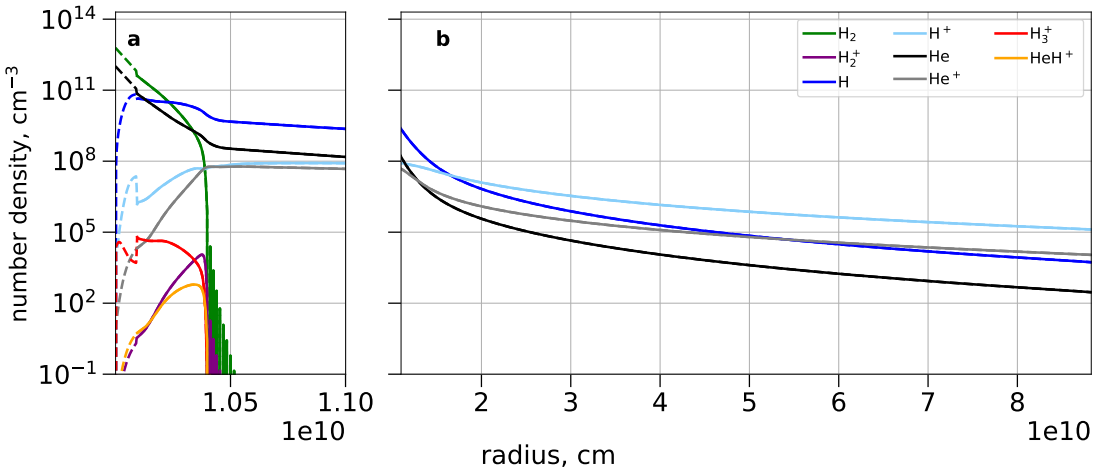

Figure 4.4 The composition of the atomosphere of our fiducial (0.05 au hot Jupiter) planet by species. The dashed lines on Panel A denote the bolometric region. The transition from: $H_2 \rightarrow H$ occurs at $1.02R_P$, $H \rightarrow H^+$ - $1.64R_P$, and $He \rightarrow He^+$ - at $1.29R_P$. The molecular layer drops off sharply at $1.04R_P$.

further heating. The mass-loss rate, $\dot{M} \sim \rho u r^2 = 6.4 \times 10^9$ g s^{-1}, is evaluated per steradian along the ray connecting the planet to the star on the dayside. Integrating over 4π steradians would lead to a mass-loss overestimate of the true mass-loss rate.

Figure 4.4 shows the composition of the atmosphere, zoomed in on the base (panel a) and depicting the outflow (panel b). H_2 and He are the major constituents of the atmosphere at the base of our simulation, which we denote the molecular hydrogen layer. This layer also contains the important coolant H_3^+. The composition transitions to H and He in the intermediate layer between the base and the outflow layer: the H_2 is thermally dissociated and the He is not yet ionized. Finally, the outflow is dominated by H^+, which is created by ionizing H and advected away by the outflow. The ionization

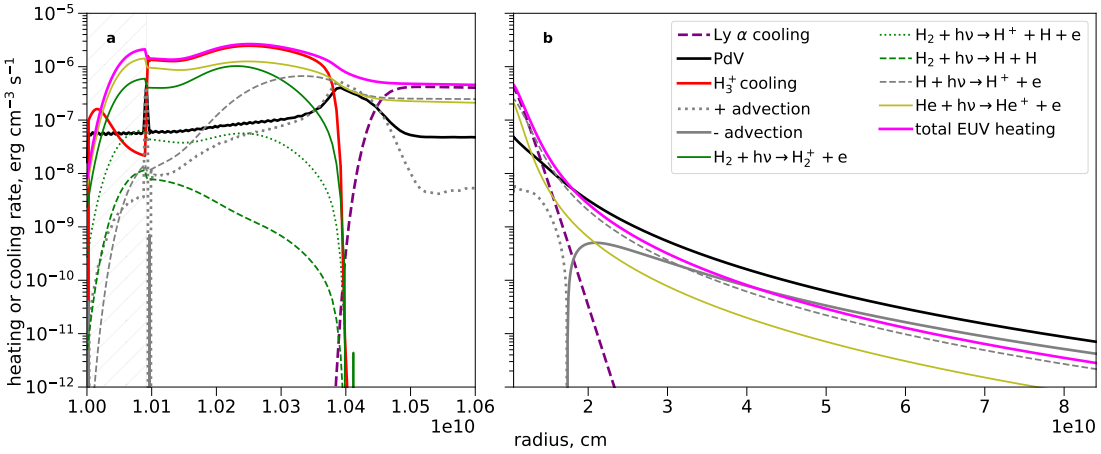

Figure 4.5 Panel (A): The heating and cooling rates as a result of the photoreactions in the wind at the base near the molecular hydrogen layer ($r < 1.04 \times 10^{10}$ cm). The H_2 layer is primarily heated by the photoioization of He, with contributions from the photoionization of H_2 deeper in the layer, and the photoionization of H higher up. This is radiated away by the H_3^+. In the atomic hydrogen layer ($1.04 \times 10^{10} < r < 1.4 \times 10^{10}$ cm), Lyman-α is the primary coolant, balancing equal contributions of the photoionization of H and He. Panel (B): The heating and cooling rates as a result of the photoreactions in the wind. The EUV heating is balanced by the PdV cooling in the outflow ($r > 1.4 \times 10^{10}$ cm). At the base, Lyman-α contributes significantly. The EUV heating is predominately due to the ionization of atomic hydrogen in the wind. We note that the PdV work term here balances the sum of the EUV heating and advection terms - as temperature decreases with distance from the planet, thermal energy is added to power the wind.

processes contribute to heating the gas.

Specifically, the energy left over from ionization heats the gas. This energy is

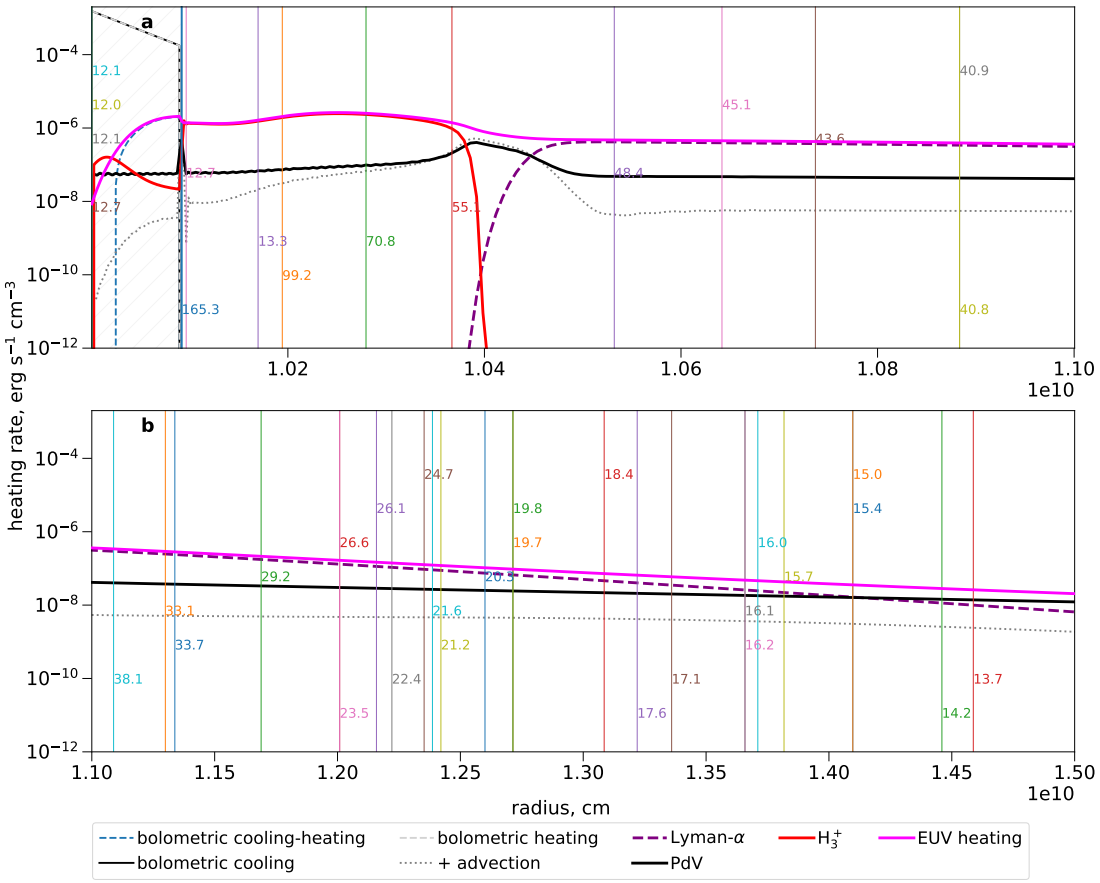

Figure 4.6 The balance of energy deposition and loss at the base of our fiducial hot Jupiter atmosphere. Panel a is zoomed in to the H_2 layer at the base, which is H_3^+-cooled. Panel b shows the region just above, dominated by H, which is primarily Lyman-α-cooled. The EUV heating rate (after overcoming the ionization potentials of the ionizing species) is shown in magenta. The cooling processes that balance the heating are: H_3^+ in the molecular layer (red), Lyman-α in the neutral atomic hydrogen layer (purple dashed), transitioning to a PdV-dominated outflow layer beyond ~ 1.45 R_P (black).

Figure 4.6 The highest-energy photons (165.3-55.1 eV) $\tau_\nu = 1$ surfaces lie within the molecular layer - a significant part of the heating attributed to those energies is H_3^+-cooled. The photons in the 55.1-15.0 eV energy range have $\tau_\nu = 1$ within the layer cooled by Lyman-α - heating due to especially the higher end of this energy range is affected significantly by Lyman-α cooling. Finally, photons in the 15.0-13.6 eV energy range ionize neutral hydrogen and are mostly all available to power the outflow.

available to either power an outflow (cool it by adiabatic expansion), or to be radiated away. Either way, the heating received by a steady-state atmosphere must exactly balance the cooling. Two important sources of radiative cooling are included in our model: Lyman-α cooling and H_3^+ cooling. We wish to understand how these cooling processes affect the outflow. To do this, we identify which of the cooling processes are dominant as a function of altitude, as well as heating from what spectral ranges are affected by the radiative cooling processes the most. For example, the photons driving the wind have to be absorbed close to the H_3^+ for the H_3^+ to matter when cooling the wind. To form H_3^+, H_2 must be present, and it also must be ionized. In Figure 4.4, we observe the presence of both H_2 and H_2^+ at the base, satisfying this criterion for forming H_3^+. Furthermore, if we look just above this molecular hydrogen layer on the plot, we observe partially ionized (but still mostly neutral) hydrogen. The temperature in Figure 4.3 rises steeply just outside of the molecular hydrogen layer. The H_3^+ cooling no longer effectively cools the gas here, yet in order to escape the gravitational well of the hot Jupiter, the gas must reach a temperature that allows the thermal pressure to

accelerate it to a significant fraction of the escape velocity, allowing adiabatic cooling to set in. Thus, with no cooling mechanism yet available, the gas heats up enough for Lyman-α cooling to ultimately begin: the electrons[2] have high enough energies to excite the neutral H, and there is enough of them for the cooling to be significant. Further up still in the model, we observe the outflow being accelerated to velocities of km s^{-1} to tens of km s^{-1} (Figure 4.3). The outflow region corresponds to the H^+-dominated region in Figure 4.4. Connecting this to the energy processes, Figure 4.5 displays the rate of EUV heating (after subtracting the contribution to overcoming the ionization potentials) as a function of altitude. It also displays the cooling processes: H_3^+-dominated at the base, Lyman-α-dominated higher up, and PdV-dominated in the outflow, suggesting a three-layer structure for our model. The H_3^+ cooling term (red solid line) drops as soon as the H_2 in the layer is thermally dissociated. The temperature increases as H_3^+ is no longer available to cool in the layer above, leading to an onset of Lyman-α cooling (purple dashed line). This term drops off as the wind acquires enough energy to be able to start climbing out of the gravitational potential of the hot Jupiter, allowing the PdV term to cool the outflow. Note that in Panel b of Figure 4.5, the PdV term is balanced by the sum of the EUV heating and advection terms. The advection term is the left-hand side of the energy equation in steady-state form (Equation 4.34). When it is negative (i.e., $\partial e/\partial r < 0$, since $\rho u > 0$ in the outflow), this term acts as an additional source of energy, advecting energy from lower in the atmosphere. For this reason, we denote it "- advection" in the outflow. In contrast, at the base as the temperature

[2]We note that Lyman-α cooling occurs via collisional excitation, and electron impacts are most efficient due to their high velocities relative to the other gas species in energy equipartition.

rises with distance, $\partial e/\partial r > 0$, and the advective term actually acts as an energy sink, transporting energy up (labeled "+ advection" in Panel a).

Zooming in to the base of the atmosphere, we have Figure 4.6. In this figure, we mark the $\tau_\nu = 1$ surfaces on the plot, with the EUV heating rate and cooling rates as described in Figure 4.5. Here, we get an idea of where photons of a given energy get deposited, and therefore which mode of cooling balances their resultant heating.

Going back to our three-layer model, its structure is a result of the interplay between photoionization reactions and the mutual reactions of the photochemical products. These reactions both form the constituents in the chemical network and determine the temperature structure, which in turn determines both the reaction rates and feeds back into the radiative cooling rates of the gas. To understand the composition in Figure 4.4, resolving the structure of the outflow, we must understand the chemical network. We focus our discussion of the chemical network on the following three processes:

- reactions forming and destroying H_3^+,
- reactions destroying H_2 to understand the transition between molecular and atomic hydrogen,
- ionization reactions of the neutral species H, H_2, and He and which reactions dominate the ionization.

Physically, what controls the ionization processes is the wavelength-dependent cross section to ionization and the incoming photon flux of that wavelength. The flux declines as $e^{-\tau_\nu}$ as we traverse the atmosphere from the top to the base. Furthermore,

the presence of multiple ionizing species and, for the case of molecular hydrogen, ionization outcomes, means that the incoming photon flux must be allocated according to each process' respective cross section. The photoabsorption cross sections H_2, H and He are shown in Figure 4.7. A key constituent of the base layer of our atmosphere is H_3^+. H_2^+ must be able to form within the layer of H_2 (Table 4.1, k3 and k8) in order for H_3^+ to form. H_2^+ forms through ionization of H_2 (Table 4.1, k3). However, the ionization of helium (k14) and atomic hydrogen (k1) are the main competitors with this process. Panel c of 4.7 displays these competing processes as a function of wavelength: photons with energies less than the 24.6 eV ionization threshold for He are all used to ionize atomic hydrogen. The longer-wavelength photons of our spectrum are almost completely absorbed by neutral hydrogen, leaving only the highest-energy photons in our spectrum to ionize molecular hydrogen. As we will discuss in Section 4.3.2, these photons' (in the 13.6-24.6 eV range) $\tau = 1$ surfaces all lie outside of the molecular hydrogen layer: the heating from this energy range of photons will be unaffected by cooling by H_3^+, as they are absorbed in a layer above.

Figure 4.8 depicts the reaction rates that are relevant to the production and destruction of molecular hydrogen at the base of our atmosphere. H_2 can be ionized by photons with energies exceeding the photoionization threshold $h\nu = 15.4$ eV via $H_2 + h\nu \rightarrow H_2^+ + e^-$ (blue line). However, in our model the layer of atomic hydrogen absorbs most of the ionizing photons up to the 24.6 eV ionization threshold of He (Figure 4.7). Photodissociation can occur followed by the ionization of atomic hydrogen: $H_2 + h\nu \rightarrow H + H^+ + e^-$. In addition, thermal dissociation can occur: $H_2 + M \rightarrow H + H + M$,

where M is a species that collisionally dissociates H_2 (orange line).

Molecular hydrogen is primarily ionized by photons in the 165.7-50 eV range (Figure 4.7), the $\tau = 1$ surfaces of which lie within this molecular layer (Figure 4.6). Figure 4.5, panel a displays the heating and cooling rates at base of our atmosphere in the layer of molecular hydrogen. The cooling done by H_3^+ is dominant in this layer - a large fraction of the EUV heating here is radiated away. Going back to Figure 4.8, the transition from molecular to atomic hydrogen occurs within the molecular hydrogen layer. Here, atomic hydrogen is able to be ionized (purple line) by higher-energy photons, but similarly that ionization heating is efficiently radiated away by H_3^+. In summary, the ionization of molecular hydrogen ultimately does not drive hydrodynamic escape of the atmosphere, as the energy deposited in this layer is efficiently radiated away by infrared cooling from the H_3^+.

At the upper edge of the molecular layer, ionizing radiation heats up the gas enough for the molecular hydrogen to dissociate (reaction k12) after collisions with the dominant species, H_2 and H. We find that photodissociation of H_2 is not the main pathway for the dissociation of H_2. We also include the dissociation of molecular hydrogen by photons in the far-ultraviolet (FUV) range of the spectrum (k22) [37, 39], with wavelengths $> 912\,\mathring{A}$ (Figure 4.7). Shortward, ionization of H dominates over this process. These photons resonantly excite H_2, which then decays with a $\sim 10\%$ probability of dissociating into two hydrogen atoms (and the remaining $\sim 90\%$ de-excites via fluorescent emission). The FUV photons probe deeper into the molecular hydrogen layer, and this process is the dominant dissociation pathway of H2 (and therefore the main source

of atomic hydrogen) further in (as opposed to thermal dissociation, which happens at higher temperatures). The heating from this process is predominately radiated away by H_3^+ (Figure 4.10, $\lambda > 12.4$ eV), even though these are the lowest-energy photons in our spectrum. The destruction of H_2, the source of H_2^+ which is the precursor of H_3^+, means that H_3^+ cooling is shut off in this layer.

In Figure 4.9 we show the main reactions involving H. The thermal dissociation of molecular hydrogen increases as the temperature of the gas starts rising to $\sim 10^4$ K. As molecular hydrogen gives way to atomic hydrogen, some of the atomic hydrogen is ionized (producing electrons capable of exciting the atomic hydrogen) and the gas is heated. Now, the predominant mode of cooling is Lyman-α cooling. Lyman-α cooling keeps the temperature here to 10^4 K. Thus, atomic H can be produced here at a high rate - which further allows for the heating of the gas and the destruction of molecular hydrogen. We note that most of the energy leftover from ionization here in this neutral atomic hydrogen layer above the molecular layer (Figures 4.4 and 4.5) is radiated away by Lyman-α cooling. Some of the heating here is converted to PdV work, partially powering a hydrodynamic outflow. Further up in the wind, the gas becomes mostly ionized. The ionizations (k1) here are not balanced by recombinations (k2), but by advection. Figure 4.9, panel b shows the reactions involving H in the outflow: advection of H and a minor contribution from recombination exactly balance ionization. Furthermore, adiabatic expansion is what mostly cools the gas, as is shown in Figure 4.5, panel b. The PdV term is the dominant cooling term in the outflow: it balances EUV heating (magenta line) as well as heat advected from below (solid gray line). We will show in

Section 4.3.2 that what drives the wind is photons in the 13.6-70 eV energy range.

The advective term $Nu\,\partial f_s/\partial r$ in Figures 4.8 and 4.9 is derived from the continuity equation. The advective term for the total number density N is:

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 N u\right) = u\frac{\partial N}{\partial r} + \frac{N}{r^2}\frac{\partial}{\partial r}\left(u r^2\right). \tag{4.31}$$

Therefore, in steady state we have:

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(u r^2\right) = -\frac{u}{N}\frac{\partial N}{\partial r} \tag{4.32}$$

Similarly, we write the advection term for an individual species s:

$$\begin{aligned}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 n_s u\right) =& u\frac{\partial n_s}{\partial r} + n_s\frac{1}{r^2}\frac{\partial}{\partial r}\left(u r^2\right) = \\ & u\frac{\partial n_s}{\partial r} - u\frac{n_s}{N}\frac{\partial N}{\partial r} = Nu\left(N\frac{\partial n_s}{\partial r} - n_s\frac{\partial N}{\partial r}\right)/N^2 = \\ & Nu\frac{\partial (n_s/N)}{\partial r} = Nu\frac{\partial f_s}{\partial r},\end{aligned} \tag{4.33}$$

where we define $f_s = n_s/N$ as the fraction of the total density in species s.

Figure 4.10 is a summary of the three-layer structure that is the outcome of our model. The H_2-dominated base layer is cooled by H_3^+, with the highest energies of our spectrum (165.3-55.1 eV) deposited there. The Lyman-α-cooled layer containing a significant fraction of neutral hydrogen absorbs photons of energies 48.4-15 eV. Finally, the H in the PdV-dominated layer is quickly converted to H^+, with 14.2-13.6 eV photons' $\tau = 1$ surfaces within this layer. As an exception, the 12.7 eV photons contributing to the far-ultraviolet resonant dissociation of H_2 are deposited in the H_3^+-cooled layer, not contributing to the outflow. The distance from the planet where the wind velocity

Table 4.2. Simulation data

Parameter	Value
Planet/star parameters	
Planet mass, M_P (M_J)	0.7
Stellar mass, $M_\star$ ($M_\odot$)	1
Planet radius, R_P (cm)	10^{10}
Semimajor axis, a (au)	0.05
Temperature at base, T_0 (K)	1000
Pressure at base, P_0 (dyn cm^{-2})	0.96
Adiabatic index, γ	5/3
Constants	
Jupiter mass, M_J (g)	1.898×10^{30}
Solar mass, $M_\odot$ (g)	1.989×10^{33}
Astronomical unit, au (cm)	1.496×10^{13}
Boltzmann constant, k_B (cm^2 g s^{-2} K^{-1})	1.380658×10^{-16}
Gravitational constant, g (dyn cm^2 g^{-2})	6.67259×10^{-8}
Stefan-Boltzmann constant, σ_SB (g cm^{-3} K^{-4})	5.6704×10^{-5}
Photoreaction thresholds (eV)	
$\mathrm{H} + \mathrm{h}\nu \rightarrow \mathrm{H}^+$	13.6
$\mathrm{H_2} + \mathrm{h}\nu \rightarrow \mathrm{H_2^+}$	15.4
$\mathrm{H_2} + \mathrm{h}\nu \rightarrow \mathrm{H} + \mathrm{H}^+ + \mathrm{e}$	18.08
$\mathrm{H_2} + \mathrm{h}\nu \rightarrow \mathrm{H} + \mathrm{H}$	4.74
$\mathrm{He} + \mathrm{h}\nu \rightarrow \mathrm{He}^+$	24.6

reaches the sound speed is the sonic point, located at 2.9 planetary radii for our fiducial model. The Hill radius, where the tidal gravity from the star starts exceeding the gravity of the planet, i.e. where the atmosphere becomes unbound from the planet, is shown at 4.5 planetary radii. The exobase is the location where the mean-free path to collisions is comparable to the scale height of the atmosphere. Beyond, the particles are no longer collisional, and the atmosphere can no longer be modeled as a hydrodynamic outflow. The neutral species' cross section is smaller than the Coulomb cross section appropriate for ionized particles: they decouple from the outflow before the gas becomes collisionless, starting to free-stream out. Kinetic modeling is appropriate for the decoupled particles, which is beyond the scope of our work. Nevertheless, for our fiducial planet at 0.05 au, the gas remains collisional out to the sonic point, allowing us to model the outflow hydrodynamically. Collisionality is further discussed in Section 4.3.8.

4.3.2 Energy allocation

In order to evaluate which parts of the spectrum are generating heating, and which parts are cooled (and which location that corresponds to in the atmosphere), we will use the steady-state version of Equation (4.9),

$$\rho u \frac{\partial e}{\partial r} = \frac{P}{\rho} u \frac{\partial \rho}{\partial r} + \rho(\Gamma - \Lambda), \tag{4.34}$$

which we derive in Appendix A.5. The left-hand side is denoted the advective term, and the right-hand side is the PdV work term, as well as the heating and cooling source terms. We calculate the contribution of each heating term from each physical region

to understand how much of the heating goes into PdV work, Lyman-α cooling, and H_3^+ cooling, and also its spectral dependence. We're going to analyze the output of our simulation by evaluating these terms.

To get the total heating rate at a given radius from the planet, we sum over the spectrum: $\Gamma(r) = \sum_\nu \Gamma_\nu(r)$. We define a cumulative heating rate $\Gamma_C(r) = \int_{R_\mathrm{P}}^{r} \rho(r) r^2 \Gamma_\nu(r) dr$ to get the total energy deposited after ionization heating from the base of the atmosphere to a height r, for example, the outer radius of our grid $r = r_\mathrm{out}$. Then, the heating rate is $\Gamma_C(r_\mathrm{out})$. Similarly, if the radiative coolants are H_3^+ cooling $\Lambda_{H_3^+}$ and Lyman-α cooling $\Lambda_{\mathrm{Ly}\alpha}$, such that $\Lambda = \Lambda_{H_3^+} + \Lambda_{\mathrm{Ly}\alpha}$, the cumulative cooling rate to the outer radius is $\Lambda_C(r_\mathrm{out}) = \Lambda_{C,H_3^+}(r_\mathrm{out}) + \Lambda_{C,\mathrm{Ly}\alpha}(r_\mathrm{out}) = \int_{R_\mathrm{P}}^{r_\mathrm{out}} \rho(r) r^2 \Lambda_{H_3^+}(r) dr + \int_{R_\mathrm{P}}^{r_\mathrm{out}} \rho(r) r^2 \Lambda_{\mathrm{Ly}\alpha}(r) dr$.

As shown in Figure 4.6, the highest-energy photons from our spectrum tend to be deposited in the molecular layer of the atmosphere at the base (the 55.1-165.3 eV range). This layer is primarily radiatively cooled via H_3^+ emission - most of the heating in this layer is balanced by this radiative cooling. Although mostly radiated away, a fraction of the energy deposited in this layer still contributes to PdV work and therefore the outflow (black solid line). We are interested in the contribution of heating in this layer to the total PdV work term. To calculate this, first we calculate the total integrated PdV work out to a radius r. This is per unit solid angle, as we are working in one-dimensional spherical coordinates:

$$PdV_{tot}(r) = \int_{R_\mathrm{P}}^{r} \left(\frac{P}{\rho} u \frac{\partial \rho}{\partial r} \right) r^2 dr, \tag{4.35}$$

where the term in parentheses is the PdV term from the steady-state energy equation. In Figure 4.6, we point out the region between the edge of the bolometric region acting as a boundary condition (hatched area) and the edge of the layer with significant H_3^+ cooling ($r = 1.039 \times 10^{10}$ cm) - where the H_3^+ cooling curve crosses the PdV work curve. The contribution to the total PdV work from this region is 7.2%. As we go up in the atmosphere, the next primary cooling mechanism is briefly PdV and advection. The contribution of this region ($1.039 \times 10^{10} < r < 1.045 \times 10^{10}$ cm) to the PdV work is 3.6%, after which the primary cooling mechanism transitions to Lyman-α cooling. The primary cooling mechanism transitions back to PdV work at $r \approx 1.039 \times 10^{10}$ cm, and the contribution to the total PdV from this region is 31.3%. The total contribution to the PdV work from the radiatively-limited region is significant. Finally, 56.5% of the PdV work comes from the cooling in the PdV-dominated upper layer, which transitions at $r = 1.42 \times 10^{10}$ cm.

We can look at the energy balance in this upper layer. The contribution from PdV as a fraction of the heating and advection is 89.8%, placing the region above approximately 1.4 planetary radii in the energy-limited regime, suggesting that the effective radius for launching the wind is 1.4 planetary radii - higher up in the atmosphere than the molecular layer that we have chosen as the base. We note that the sonic point for this planet is at 3.14 planetary radii.

To calculate the percentages shown in Figure 4.11, we first define the ratio of the heating from a specific energy bin to the total heating at a given radius: $f_\nu = \Gamma_\nu(r) / \sum_\nu \Gamma_\nu(r)$. Then, we multiply each cooling rate ($\Lambda_{H_3^+}, \Lambda_{Ly\alpha}$, and PdV) at

that radius by that fraction. This way, we estimate how much of each of the cooling processes balance each heating energy bin: $\Lambda_{H_3^+,\nu}(r) = f_\nu \Lambda_{\mathrm{H_3^+}}(r)$, $\Lambda_{\mathrm{Ly}\alpha,\nu}(r) = f_\nu \Lambda_{\mathrm{Ly}\alpha}(r)$, and $PdV_\nu(r) = f_\nu PdV(r)$. We can obtain the total contribution to the wavelength bin by integrating over radius[3]: $\Lambda_{H_3^+,\nu} = \int_{R_\mathrm{P}}^{r_{out}} \Lambda_{H_3^+,\nu}(r)\rho(r)r^2dr$, $\Lambda_{\mathrm{Ly}\alpha,\nu} = \int_{R_\mathrm{P}}^{r_{out}} \Lambda_{\mathrm{Ly}\alpha,\nu}(r)\rho(r)r^2dr$, and $PdV_\nu = \int_{R_\mathrm{P}}^{r_{out}} PdV_\nu(r)r^2dr$. Finally, we can divide this by the total heating, $\int_{R_\mathrm{P}}^{r_{out}} (\sum_\nu \Gamma_\nu(r))\rho(r)r^2dr$, to find out what fraction of the heating goes into what cooling process for each respective energy bin.

Figure 4.12 depicts the energy budget broken up into wavelength bins. The highest-energy photons are primarily balanced by $\mathrm{H_3^+}$ cooling, then Lyman-α in the 64-28 eV range. Then, for energies below around 28 eV around 0.5 of the heating goes into PdV. This may suggest that the outflow is energy-limited from the ionization threshold of H (13.6 ev) to 28 eV. However, the $\tau = 1$ surface for photons from 28 eV to 15 eV lies at around 1.17 planetary radii, which is where Lyman-α dominates over the PdV work term (Figure 4.6). This is because the higher-energy photons also contribute significantly to the heating in that region, and their contribution to the heating rate is mostly radiated away. Thus, to determine where the base of the wind is we make Figure 4.12 for radii above 1.4 planetary radii, where the PdV term starts to dominate.

If we restrict the radius range to $r > 1.4$ planetary radii, we can see that 15% of the heating goes into Lyman-α, and the rest goes into PdV or is advected away over the entire range of the spectrum.

As in Figure 4.12, Figure 4.13 displays the cooling rates normalized by the

[3]After multiplying by ρ in the integrals the quantities are in units of energy, not energy density.

available heating from the spectrum. However, we plot the PdV term as a cumulative sum from the lower-energy bins of the spectrum (which are entirely energy-limited) to the higher-energy bins (which are almost entirely radiation-limited). This allows us to get a quantitative estimate for which spectral energies are driving the wind: most of the spectral energy used to power the outflow is from energies less than ~ 70 eV.

4.3.3 Stellar heating efficiency parameter, ϵ

A regime frequently used in the literature to evaluate the mass-loss rate is the energy-limited regime. It gives the maximum mass-loss rate for a planetary outflow, assuming that all of the incoming stellar UV flux goes into powering the outflow:

$$\dot{M}_{\rm elim} = \frac{\epsilon R_{\rm P}^3 F_{uv}}{G M_p}. \tag{4.36}$$

In the literature, this efficiency factor ϵ is sometimes used to mean the overall efficiency, including the ionization and the radiative cooling. We only consider the ionization in this section, such that the efficiency factor $\epsilon = 0.5$ is used to account for the energy required to ionize the species in the atmosphere, which is unavailable for heating. For comparison with this expression, we evaluate our results for what the effective ϵ would be. Recall that we initialized the simulations to mass fractions X=0.75 for H and Y=0.25 for He. If we assume the hydrogen is atomic, the ratio of number densities H:He is 4:3. Then we can approximate the efficiency factor

$$\epsilon = \frac{4}{7}\sum_{\nu}\left(\frac{E_\nu - I_{\rm H}}{E_\nu}\frac{F_\nu}{F}\right) + \frac{3}{7}\sum_{\nu}\left(\frac{E_\nu - I_{\rm He}}{E_\nu}\frac{F_\nu}{F}\right), \tag{4.37}$$

where $F = \sum_\nu F_\nu$ is the total EUV flux. This gives an estimate of $\epsilon \simeq 0.51$, which we will use in Section 4.3.5 to discuss the energy-limited mass-loss rate for a grid of models with varying semimajor axis.

We may compute the actual heating efficiency by defining it as a fraction of the total input EUV flux that goes into heating, after accounting for overcoming the ionization potentials of the species. It may be computed for a given distance from the planet (Figure 4.14, panel a), or for a given energy bin (Figure 4.14, panel b). The total efficiency computed over the entire spectral range, integrated over the planetary radius range, is $\epsilon = 0.54$, showing that our estimate from Equation (4.37) is reasonable. We note here that although the photons of higher energy are more efficient at heating the gas, they do not necessarily power the outflow more, as they are deposited at altidudes that have significant radiative heating, both H_3^+ for the highest-energy photons in our spectrum ($\sim 165.3 - 70$ eV), and Lyman-α ($\sim 70 - 20$ eV).

From our simulations including radiative cooling, we do not find mass-loss rates equal to this energy-limited rate. We provide a calculation of the output estimate in Section 4.3.5. The radiative cooling provides an extra factor of loss of efficiency, which depends on distance from the star (and therefore on the EUV flux).

4.3.4 4 radiative cases for fiducial 0.05au hot Jupiter

In this section, we compare 4 cases for our fiducial hot Jupiter at a 0.05 au orbital separation: one with full radiative cooling, a case with either just H_3^+ cooling or just Lyman-α cooing turned on, and one - with all radiative cooling turned off. We

contrast the density, pressure, velocity in the outflow, and temperature structure in Figure 4.15. The composition for all 4 cases is plotted in Figure 4.16, both zoomed-in to the base (left-hand plots), and depicting the larger-scale outflow (right-hand plots). Figure 4.17 shows a zoomed-in temperature plot on the left, along with the larger scale shown on the right. Finally, Figure 4.18 shows the full reaction rate network for reactions involving H_2 - the molecular layer at the base - to contrast how the temperature structure affects the dominant reaction rates, as well as the resulting composition in Figure 4.16.

In Figure 4.15, the density, pressure, and velocity in the wind drops as we turn on radiative cooling, with H_3^+ and Lyman-α having similar contributions in magnitude to the general outflow. Both cooling mechansims have a similar (in order of magnitude) impact on the mass-loss rate. However, they affect the temperature structure differenty, as they operate in different regions of the outflow (as seen in the three-layer structure revealed by our model, e.g. Figure 4.10). Both cooling mechanisms decrease the total internal energy that the pressure-driven wind is able to draw from to power itself, thereby decreasing both the velocity and temperature thoughout the outflow. The temperature plot is shown in Figure 4.17 with a focus on the base of our simulated region in panel a. While H_3^+ lowers the the temperature at the base in the molecular region, Lyman-α thermostats the peak temperature to $\sim 10^4$ K (panel b).

In Figure 4.16, in the H_3^+-cooled base layers in panels a and e the transition from molecular to atomic hydrogen occurs at a slightly higher altitude, as H_2 is less thermally dissociated when temperatures are low.

Figure 4.18, Panel (A) plots the reaction rates involving H_2. H_2^+ and H ex-

change an electron and almost balance out to the edge of the molecular layer (yellow solid line and dotted grey line) - these don't involve the descruction of H_2, though their rates are quite high. The main destruction pathways for H_2 are the production of H_3^+ throughout the layer, and the thermal destruction of H_2 at the edge of the layer, which causes the level of H_2 to drop sharply. Note that the photodissociation and ionization reaction ($H_2 + h\nu \rightarrow H + H^+ + e$, purple solid line) is not the dominant destruction pathway of H_2, and neither is the dissociation of H_2 by far-ultraviolet photons deeper in the molecular layer (red solid line). In panels b and d of Figure 4.18, molecular hydrogen is dissociated thermally throughout the layer after being advected up from the lower boundary region - as H_3^+ cooling is turned off, the temperatures get quite high. Only when the temperature reaches $\sim 2 \times 10^3$ K is thermal destruction significant. In contrast, in the presence of H_3^+ cooling, the temperatures in the molecular layer (Figure 4.17, around 1.02 au) do not get high enough for H_2 to thermally dissociate (solid green line) throughout - instead, H_3^+ formation (solid pink line) is the dominant pathway for H_2 destruction throughout the layer. With advection unable to replenish H_2 at a high enough rate from the inner boundary region, H_2 falls off rapidly (Figure 4.16, left panels), transitioning into atomic hydrogen.

Our case with Lyman-α cooling turned off is qualitatively similar to the compositions described in Yelle [99] and García-Muñoz [27].

4.3.5 Grid of models

We address the question of how the efficiency of atmospheric escape varies with stellocentric distance a. For that purpose, we run a grid of models, placing our fiducial hot Jupiter at 0.015-0.9 au from the host star. We scale the EUV flux by a^{-2}, effectively varying the distance. Tidal gravity is present in our model (Appendix A.1), enhancing the mass-loss rate for the closest-in planets. As the term scales as a^{-3}, beyond ~ 0.1 au the effect is negligible. We plot the density (panel a), velocity (panel b), pressure (panel c), and temperature (panel d) profiles in Figure 4.19. The sonic point moves outward as the stellar EUV flux decreases with semimajor axis and less energy is available to lift the gas out of the potential well of the planet. The molecular hydrogen transitions to atomic at a slightly increasing height in the atmosphere for a decreasing semimajor axis. This is because the temperature for the close-in planets is actually lower at the base in the molecular region, allowing a slighly more extended molecular region. This is due to H_3^+ cooling: more H_3^+ is able to form close-in to the star, allowing for more cooling.

Figure 4.20, panel a displays the fraction of the total EUV heating (after accounting for the energy required to overcome the ionization potentials) that goes into H_3^+ cooling, Lyman-α cooling, and PdV as a function of semimajor axis. We plot the total heating rate and the heating that goes into directly heating the gas in Figure 4.20, panel b. The heating efficiency is approximately constant over the semimajor axis range, suggesting that our value of $\epsilon = 0.51$ is reasonable for making an estimate of the

energy-limited escape rate. We will use this expression for the energy-limited escape rate to evaluate how energy-limited our wind is, i.e. what fraction of the total available energy that is used to heat the gas goes into powering the wind. Panel b also shows the power output in both H_3^+ and Lyman-α cooling, as compared to the effective EUV heating rate (blue line). For the set of models including both H_3^+ and Lyman-α cooling (black dashed and soild lines), the Lyman-α power output exceeds that of the H_3^+, comprising a greater fraction of the incoming flux further out from the star. Lyman-α cooling drops in importance once tidal gravity becomes strong for the 0.015au case.

The mass-loss rates for our grid of models are plotted in Figure 4.21. The silver solid line marks the mass-loss rates for our series of runs with all radiative cooling turned off: all of the EUV heating (minus the heating required for ionization) is available to power the outflow. The blue dashed line is the energy-limited approximation to the mass-loss rate, with efficiency $\epsilon = 0.51$ to account for the energy required to overcome the ionization potentials of the ionizing species. This expresssion matches the mass-loss rate well. The mass-loss rate for this case goes as $F_{\rm EUV}^1$, which is not necessarily true in cases where radiative cooling is significant. For example, the model of Murray-Clay et al. [58] that includes Lyman-α cooling found that the energy-limited escape rate trends with flux as $F_{\rm EUV}^{0.9}$, reflecting possibly the increased importance of Lyman-α cooling with increasing flux in the UV flux range $10^2 - 10^4$ erg cm^{-2} s^{-1} of their 20 eV grey stellar irradiation model. In our simulations, for the case where all radiative cooling is turned on (black solid line), the mass-loss rate is approximately 0.3 times the energy-limited rate at 0.05 au. However, the mass-loss rate does not scale as $F_{\rm EUV}^1 \sim a^{-2}$ - it is

a steeper function of semimajor axis (and stellar flux). The mass-loss rate was found to go as $a^{-2.4}$, or as the energy-limited rate times $a^{-0.4}$ (neglecting the increased mass-loss rate for the 0.015au case due to tidal gravity), and is shown on the plot as a light gray dashed line. Escape becomes less energy-limited with increasing semimajor axis, with a greater fraction of the deposited stellar energy being radiated away by Lyman-α cooling (Figure 4.20). Interestingly, H_3^+ cooling was found to be a roughly constant fraction of the deposited energy heating the gas for the parameter range of our grid of models. In addition, for distances $a < 0.1$ au, tidal gravity enhances the mass-loss rate - note the diverging black solid and the light gray dashed lines. The increased importance of Lyman-α coolling can be observed in the case where H_3^+ cooling is turned off (grey dotted line) - the line trends toward the case with all radiative cooling turned on with increasing distance, suggesting that the departure from the energy-limited regime in the low-flux case is predominately due to Lyman-α cooling. On the other hand, the H_3^+-only case (grey dashed line) follows the energy-limited case with semimajor axis: it radiates away a roughly constant fraction of energy deposited as heat.

4.3.6 Stellar wind, magnetic field, and radiation pressure

The plantary outflow may be significantly redirected by the stellar wind and/or the stellar magnetic field if the stellar wind ram pressure, $P_\star$, the magnetic field pressure, $P_{B,\star}$, and the thermal pressure, P_k exceeds the sum of the thermal and ram pressures from the planetary outflow, $P_{k,p} + P_p$ (for example, forming a cometary tail, as in McCann et al. 53). The stellar wind ram pressure is $P_{\rm ram} = \rho_\star v_\star^2$. If we

take the stellar wind velocity as the approximately constant $v \sim 400$ km s^{-1} for our semimajor axis range $a < 0.9$, and the number density of protons as approximately $n_p \simeq 2.4 \times 10^3 (0.05\text{au}/a)^2$ cm^{-3} [58], the ram component of the stellar wind pressure can be estimated: $P_{\text{ram}} \simeq m_{\text{H}} n_p v^2 \simeq 6.4 \times 10^{-6} (0.05\text{au}/a)^2$ dyn cm^{-2}, or 6.4 picobars at 0.05 au. Taking the proton temperature to be 10^6 K, the thermal pressure due to protons in the stellar wind is $P_k = n_p k T_p \simeq 3.3 \times 10^{-7} (0.05\text{au}/a)^2$ dyn cm^{-2}, or 0.3 picobars at 0.05 au. The pressure due to the solar magnetic field is $B_r^2/(8\pi)$. The radial component of the "monopole" magnetic field B_r is 5×10^{-5} G $(\text{au}/a)^2$, so we estimate $P_B \sim 4 \times 10^{-8} (0.05\,\text{au}/a)^4$ dyn cm^{-2}, or 0.04 picobars at 0.05 au. As the semimajor axis range of our simulations is outside the Alfven radius, the pressure is not dominated by the magnetic field pressure - this component is small compared to the ram and thermal pressures. Therefore, we estimate the pressure to be approximately 7 picobars at 0.05 au. At the sonic point distance from the planet, the thermal and ram pressures of the outflow are equal. At 0.05 au, both components are approximately 5 picobars, so $P_{k,p} + P_p \simeq 10$ picobars. As this is comparable to the stellar wind pressure, especially at close-in distances we conclude that the outflow shape may be affected by the stellar wind.

If the pressure of the stellar wind exceeds the pressure of the planetary wind upstream of the sonic point, the wind may be significantly redirected and possibly confined. If the pressure of the stellar wind exceeds the planetary wind pressure downstream of the sonic point, the supersonic solution is valid, as the boundary conditions will not travel upstream.

Radiation pressure is typically considered for interactions with dust grains. However, radiation pressure can act on hydrogen atoms. The cross section for scattering by ionized hydrogen is negligible compared to the scattering by neutral hydrogen of Lyman-α photons, which can be significant [54]. With increasing distance from the star, our outflow becomes increasingly neutral. While radiation pressure will not affect the strongly ionized outflows close to the star, we expect radiation pressure could play a significant role for the neutral hydrogen farther out, especially as the neutral hydrogen starts to decouple from the ionized outflow (see also Section 4.3.8).

4.3.7 Planetary magnetic field confinement

A planetary outflow may be confined in the presence of a strong planetary magnetic field. We estimate strength of the magnetic field at the surface of the planet, such that the magnetic field pressure P_B becomes equal to the wind pressure $P = P_k + \rho v^2$ - the sum of the thermal and the ram pressures - at the sonic point. The magnetic field pressure of a dipole field goes as:

$$P_B(r) = P\left(\frac{r}{r_{sonic}}\right)^{-6}. \tag{4.38}$$

Then, the magnetic field at the surface is: $B = (8\pi P_B(R_{\rm P}))^{1/2}$, which for our fiducial hot Jupiter is 0.5 G. For comparison, the magnitude of Jupiter's magnetic field is 4.17 G, which is considered quite strong. However, the magnitude of the magnetic fields on hot Jupiters is currently unknown.

4.3.8 Collisionality

An outflow may be modeled as a fluid if it is collisional out to the sonic point. In other words, the mean-free path to collisions, $1/(n_{col}\sigma)$ is less than the pressure scale height, $H = k_{\rm B}T/(\mu g)$, with the gravitational acceleration $g = GM_P/r^2$. There are two cross sections for collision relevant to our problem: the hard-body cross section that is used to approximate the collision cross section for neutral hydrogen, σ_{hb}, and the Coulomb scattering cross section used for neutrals, σ_C. The gas is mostly ionized in our models close to the host star, $a < 0.2$ au. Beyond that, for the semimajor axis range $0.2 - 0.9$ au considered for our grid of models, the gas is a mix of ions and neutrals, with the ionization fraction approaching 0.5 for the 0.9 au case. The ionized hydrogen is collisional throughout our semimajor axis range out to the sonic point. We note that at the base, even the neutral hydrogen remains collisional. However, the neutral hydrogen starts to decouple from the ionized hydrogen in the wind as we go out in semimajor axis: the neutral hydrogen mean free path to collisions generally exceeds the scale height for our models. We will not attempt to model the cross section for a mixture of neutral and ionized hydrogen. As the neutral hydrogen decouples from the pressure-driven hydrodynamic outflow, it starts to free-stream out. However, if the neutral hydrogen in the outflow actually spends most of its time in the ionized state, where it has plenty of time to thermalize with the surrounding ions, the neutral hydrogen can still be treated as collisional. To evaluate this, we can compare the free-streaming timescale $\tau_{fs} \simeq H/u$ (where u is the outflow velocity) to the ionization timescale per neutral hydrogen atom

(τ_{ion}). While the free-streaming timescale is comparable to the ionization timescale close-in ($a < 0.1$ au), further out from the star the neutral hydrogen is able to travel a scale height before ionizing. Therefore, only the ionized hydrogen can be represented as a hydrodynamic outflow, and the mass-loss rates from our model further out in semimajor axis will be reduced, especially as the neutral fraction goes up. The 0.9 au case is marginal: the wind is 50% ionized hydrogen and 50% neutral, so we use that as the maximum distance from the star we attempt to simulate (keeping in mind that the neutral hydrogen is likely decoupled from the ionized wind at that point, reducing the true mass-loss rate) - beyond, a kinetic model would be needed.

4.4 Discussion

Our goal was to evaluate the efficiency of atmospheric escape as a result of EUV heating, while being radiatively cooled by H_3^+ and Lyman-α. For this, we created a one-dimensional hydrodynamic model that solves the fluid equations to steady state, coupled to a chemical network solver that allowed us to include reactions between the various species of H and He in our atmosphere.

As a result of our simulations, we uncovered a three-layer structure to the atmosphere:

- a molecular hydrogen layer at the base (cooled by H_3^+),
- an intermediate layer of predominately neutral atomic hydrogen (cooled by Lyman-α), and

- an outflow of H^+, in which PdV was the main cooling mode.

We then ran a grid of models with increasing distance from the star. We found that the proportion of total EUV flux that is radiatively cooled by H_3^+ stayed roughly constant within the parameter range of our study. In contrast, as the outflow became increasingly neutral with distance, the Lyman-α cooling had an increasingly important effect as a fraction of the total heating. This decreased the mass-loss rate even faster than the naïve estimate of a^{-2}: the mass-loss rate declined as $a^{-0.4}$ times the energy-limited estimate. Tidal forces enhanced the mass-loss rate close-in to the star.

We found that H_3^+ cooling preferentially affects the higher-energy bins of our spectrum. For the purposes of theoretical modeling, ~ 115 eV is the energy cutoff above which $\sim$70% of the energy is radiated away as H_3^+ cooling.

We have developed a flexible code that can be expanded to include the effects of higher-energy photons, which are important for planets around younger stars. In addition, expanding the spectrum to the far-ultraviolet range would allow us to include molecular dissociation, which is especially important for smaller planets. Diffusion has been implemented in the code, but not turned on, as its effects are negligible for the purposes of this study. However, it is important for planets with metals in the atmosphere.

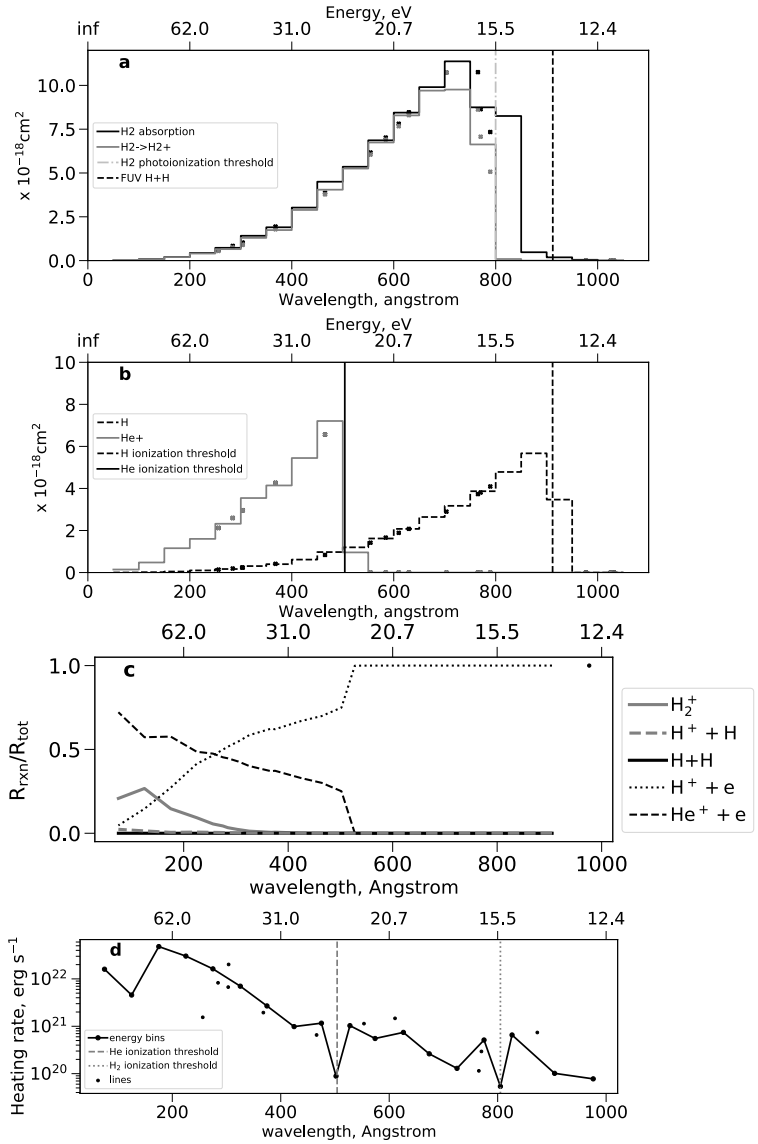

Figure 4.7 Cross sections for absorption and ionization of H_2, H and He. The bins are wavelength bins of width 50$\mathring{A}$, and the dots are emission lines. Panel a: total photoabsorption cross section of H_2 (black). Plotted in black is the minimum wavelength for dissociation of H_2 into atomic hydrogen via absorption of a far-ultraviolet photon.

Figure 4.7 Wavelengths shortward are absorbed by atomic hydrogen. 1110 Å is the beginning of the line-dominated spectrum: longward absorption is weakened. The H_2 photoionization threshold is the gray dotted-dashed line. Panel c: rate of given photoreaction as a fraction of the total photoreaction rate rate for different spectral energies. H absorbs all photons below the ionization threshold of He and > 13.6 eV. Panel d: the heating rate of the gas as a function of energy, with the bins connected by the black solid line, and the emission lines plotted as dots. Close to the ionization thresholds of H and He, a significant fraction of the energy goes toward ionization, corresponding to the dips in the plot. We note that this heating rate is the energy available to heat the gas after taking ionization into account, not accounting for radiative cooling.

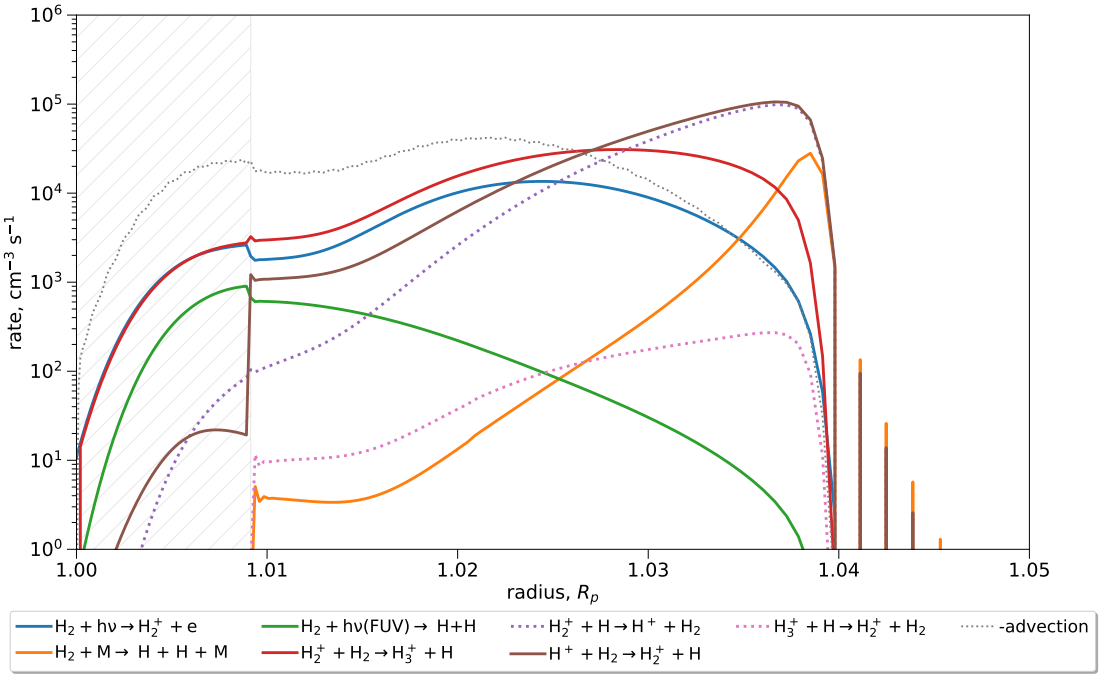

Figure 4.8 Selected reaction rates involving H_2. The advective term is plotted as derived in Equation (4.33). The brown solid and purple dashed lines represent the process of H_2 and H exchanging an electron: this reaction pair does not lead to the destruction of H_2. Instead, H_2 is thermally dissociated (orange line) at $\sim$ 1.04 planetary radii. The key reaction of H_2^+ with H_2 to form H_3^+, the most important coolant in this layer, is plotted in red. The photoionization of H2 is plotted in blue; the formation of H_2^+ is a crucial step in forming H_3^+. Interestingly, the far-ultraviolet dissociation of H_2 (green line) increases as we go deeper in the layer - the 12.7 eV photons responsible for this process are also some of the lowest-energy in our spectrum. Advection is plotted as a gray dotted line; its sign follows the convention established in Figure 4.5: as the sign is negative, the advective term represents an addition of H_2 from the boundary.

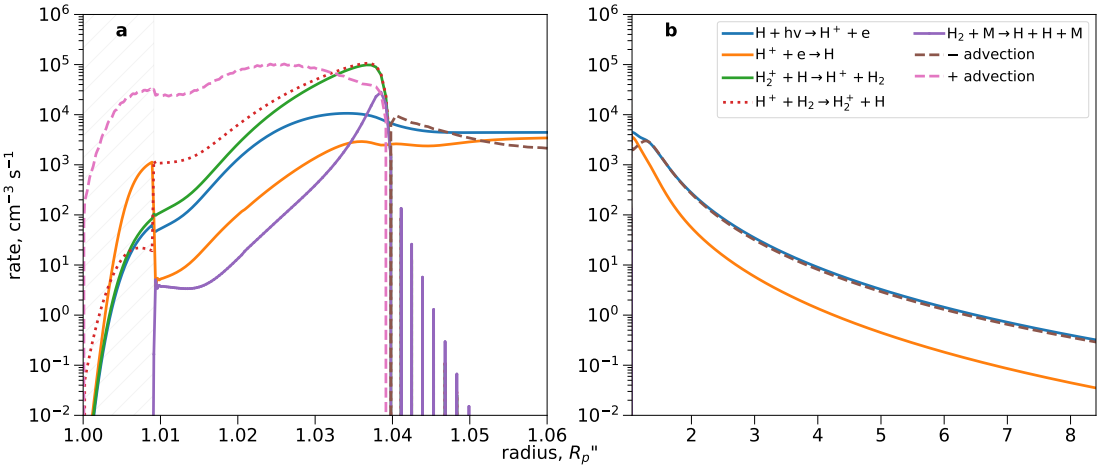

Figure 4.9 Reaction rates involving H. Panel a: zoomed in to molcular hydrogen layer. Panel b: 1.06 to $\sim 8.8R_{\rm P}$. As in Figure 4.8, the green solid and red dotted lines represent the electron exchange between H_2 and H. The purple solid line marks the production of H as a result of the thermal dissociation of H_2 at the boundary of the molecular layer. The blue solid line is the photoionization of H. In the outflow (Panel b), the ionization of H is exactly balanced by the advection of H (brown dashed line), with a small contribution from recombination (orange). The sign convention for advection is "−" for the advection in of H, and "+" for the advection of H away, as in Figures 4.5 and 4.8.

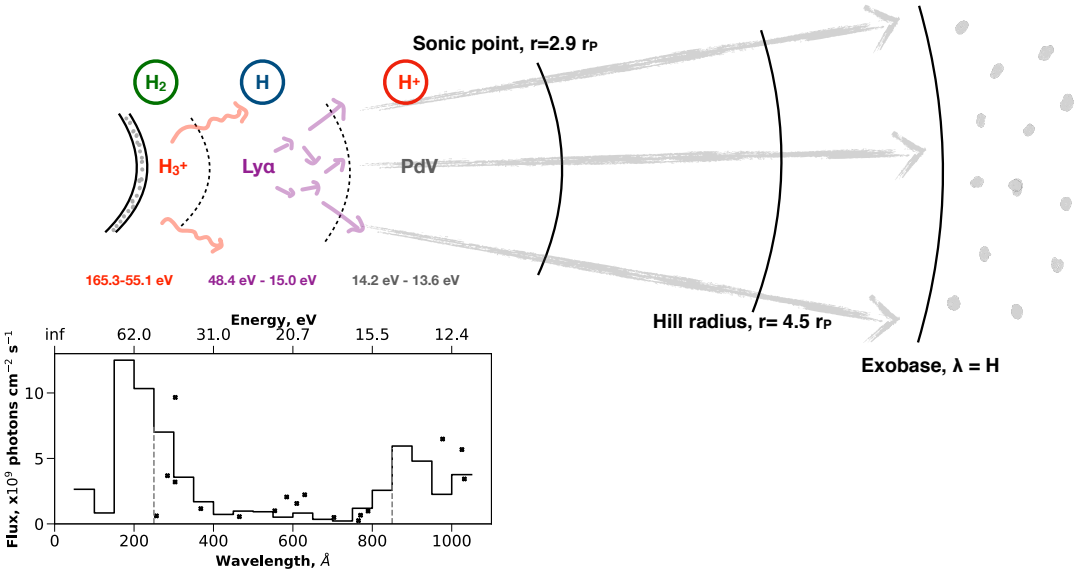

Figure 4.10 The three-layer diagram of the outcome of our model. The inner layer is the H_2 layer cooled by H_3^+, where the highest-energy photons are deposited. The neutral atomic hydrogen layer further up is predominately cooled by Lyman-α, with a contribution from PdV as well. Intermediate-energy photons are depostied there. The outflow layer is cooled via PdV. Its structure is shown as a hydrodynamic outflow out to the sonic point, extending out to the Hill radius of the planet, where the gas becomes unbound from the planet. The exobase, where the outflow can is no longer collisional, is shown. The spectrum below the diagram corresponds to the P200 model from Richards et al. [74] with a high level of steller activity. The dots on the spectrum represent emission lines, and the bins are 20 wavelength bins with widths of 50 Å.

Figure 4.10 The 165.3-55.1 eV photons' $\tau = 1$ surfaces lie within the H_3^+-cooled layer, the 48.4-15.0 eV - within the Lyman-α-cooled layer, and the 14.2-13.6 - within the outflow layer. The 12.7 eV photons responsible for the resonant dissociation of H_2 reach deeper in the H_3^+ - cooled layer.

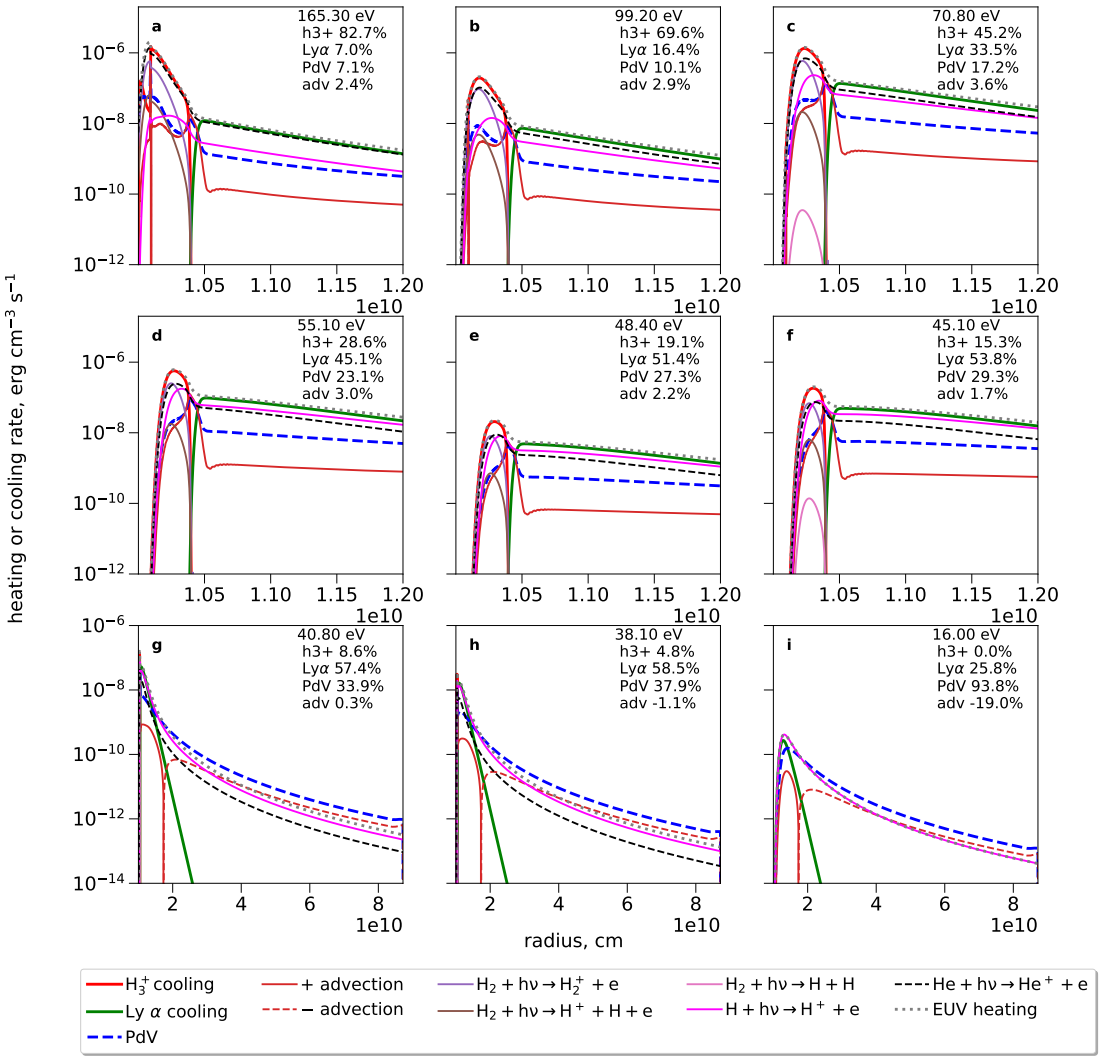

Figure 4.11 Heating and cooling rates for selected energy bins. The energy is allocated according to the expressions from Section 4.3.2. The panels display the breakdown of the heating rate as a result of all of the photoreactions with radius. Panels a-f are zoomed in at the base, where radiative cooling is important. H_3^+ cooling dominates panels a-c, whilst Lyman-α is the main cooling source in panels d-f. He ionization is an important heating source in panels a-c, while H ionization becomes more important for heating in panels d-i.

Figure 4.11 Although dominated by Lyman-α, panels g and h still display significant contributions to the PdV term from the intermediate energy ranges in our spectrum. Finally, panel i displays an energy bin whose heating is primarily goes into driving the PdV outflow (though the advective term here pulls extra energy from the internal energy of the gas in addition to the heating). The heating in this bin is all due to H ionization. The advective term sign convention follows that of Figure 4.5.

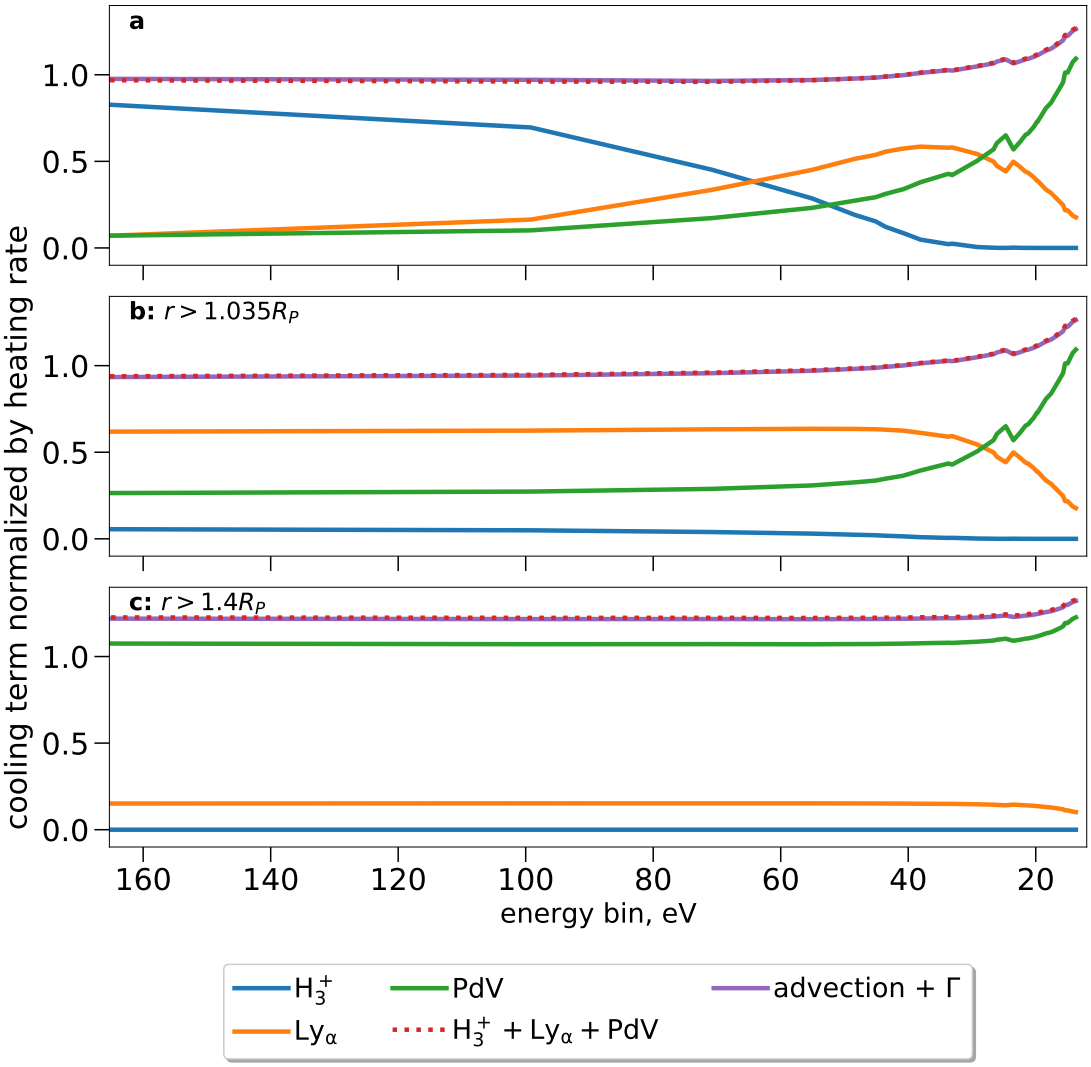

Figure 4.12 PdV_ν, $\Lambda_{H_3^+,\nu}$, $\Lambda_{\mathrm{Ly}\alpha,\nu}$ normalized by the stellar heating rate Γ_ν. The cooling terms balance the sum of advection and stellar heating (normalized by the stellar heating) in steady state (Equation (4.34)). Panel a is shown for the entire radius range (beyond the bolometric region acting as a boundary region). Panel b is shown for $r > 1.035R_\mathrm{P}$, the transition between H_3^+- and Lyman-α-dominated cooling (Figure 4.6). Panel c is shown for $r > 1.4R_\mathrm{P}$, beyond which the outflow becomes energy-limited.

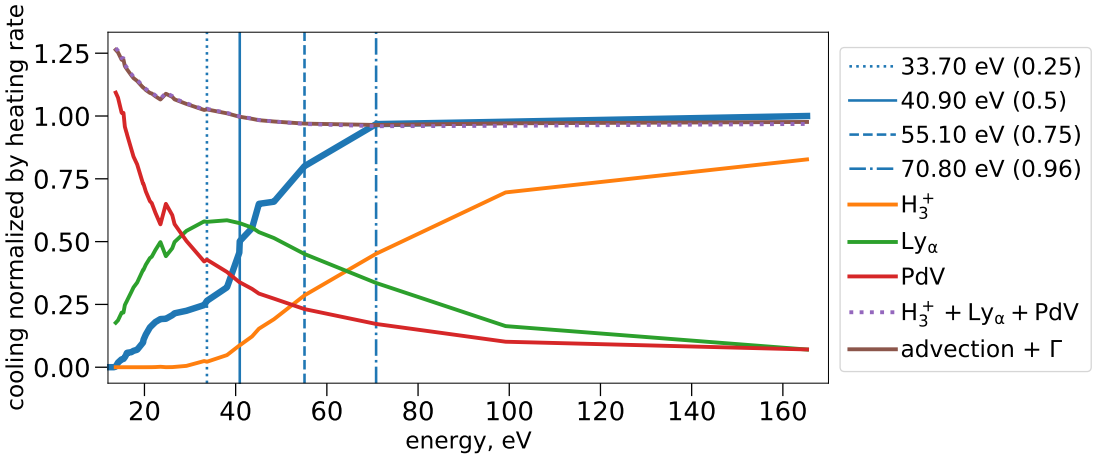

Figure 4.13 As in Figure 4.12, $\Lambda_{H_3^+,\nu}$ and $\Lambda_{\mathrm{Ly}\alpha,\nu}$ are normalized by the stellar heating rate Γ_ν. The PdV term is shown as a cumulative fraction from the low- to the high-energy bins. As the low-energy bins are dominated by PdV-cooling (red line), the cumulative term gives a sense of where in the spectrum most of the PdV outflow power is derived. 0.25 of the PdV comes from bins left of the vertical dotted line, 0.5 - of the solid line, 0.75 - of the dashed line, and 0.96 - of the dashed-dotted line. Beyond $\sim$ 70.8 eV, almost all of the energy is radiated away, predominately by H_3^+ from the base molecular layer in the atmosphere.

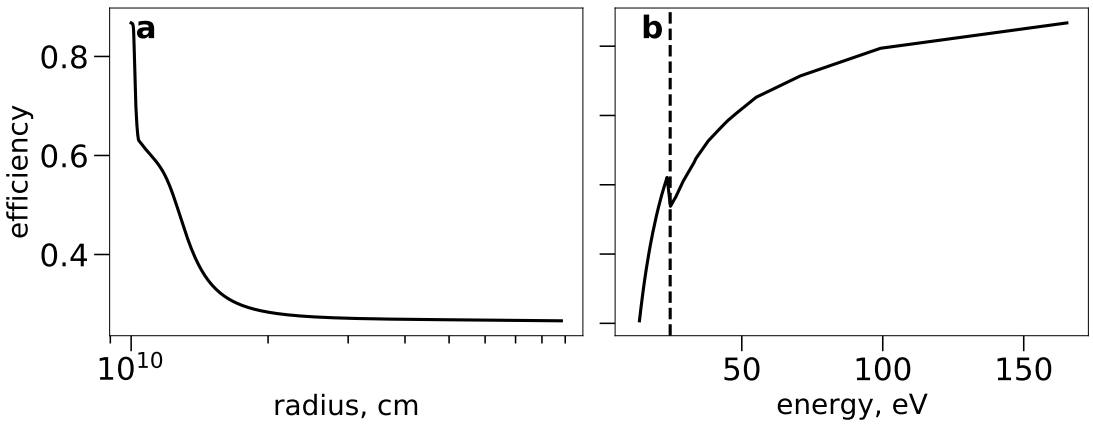

Figure 4.14 Efficiency of EUV heating, defined as the fraction of total EUV heating that is available to heat the gas by ionization processes, after overcoming the ionization potential of the species. Panel a shows the efficiency with radius - heating is more efficient in the base molecular layer, as the photons deposited there are of higher energy. Panel b shows the efficiency with energy. The vertical dashed line marks the ionization threshold of He: as the photons rightward are available to ionize, the efficiency drops slightly. We emphasize that this is a deposition efficiency of heating, after accounting only for the ionization processes. This does not yet include any losses due to radiative cooling, either H_3^+ or Lyman-α.

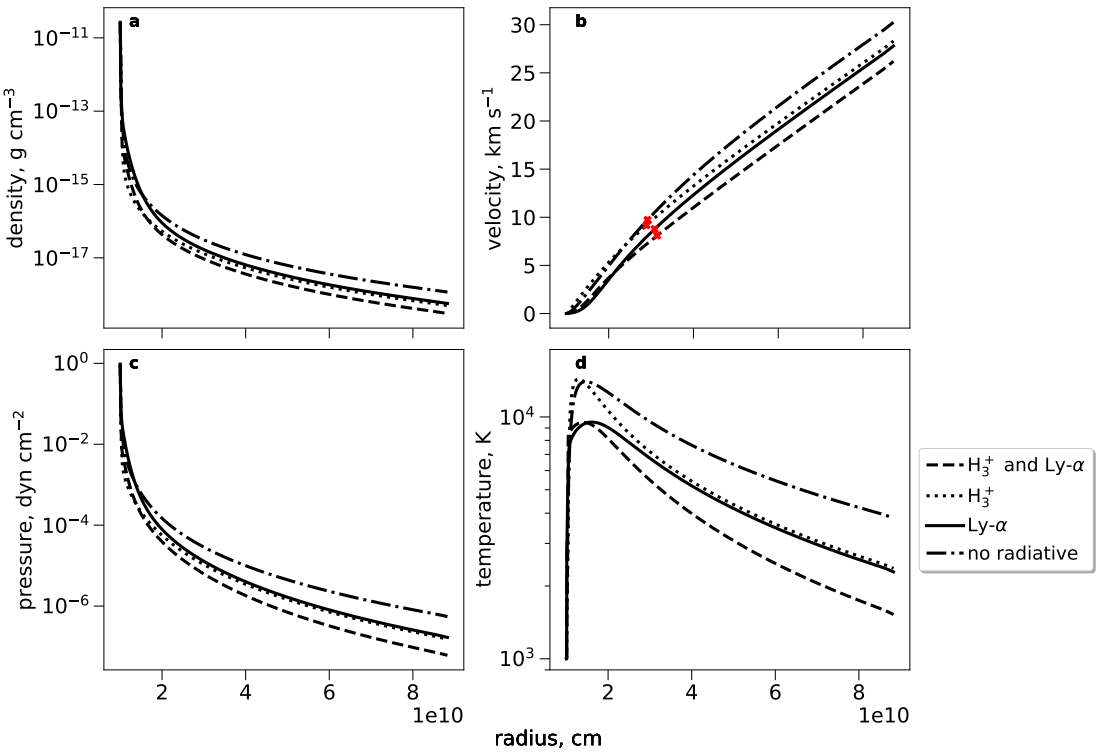

Figure 4.15 Density, velocity, pressure, and temperature profiles for our 4 radiative cases: H_3^+ and Lyman-α cooling turned on (dashed), only H_3^+ - dotted, only Lyman-α - solid, and no radiative cooling - dashed-dotted lines.

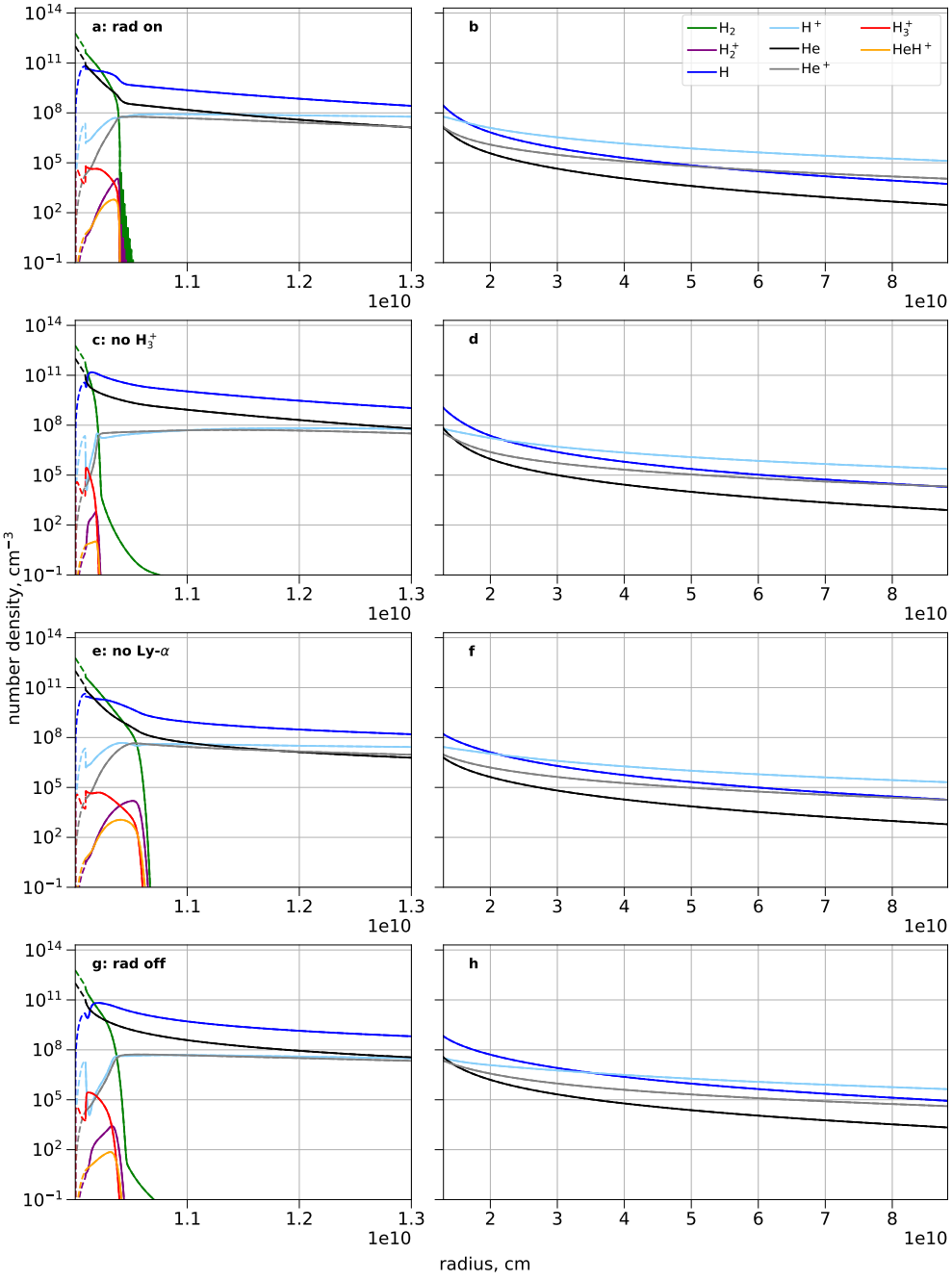

Figure 4.16 The composition of the atomosphere of our fiducial (0.05 au hot Jupiter) planet by species. The two upper panels show our fiducial planet with Lyman-α and H_3^+ cooling turned on.

Figure 4.16 Panels c and d - with only Lyman-α cooling turned on, panels d and f - with only H_3^+ cooling turned on, and the lower panels - with radiative cooling turned off. The dashed lines on the left panels denote the bolometric region. In panel a, the transition from: $H_2 \rightarrow H$ occurs at $1.02\,R_P$, $H \rightarrow H^+$ - $1.64\,R_P$, and $He \rightarrow He^+$ - at $1.29\,R_P$. The molecular layer drops off sharply at $1.04\,R_P$.

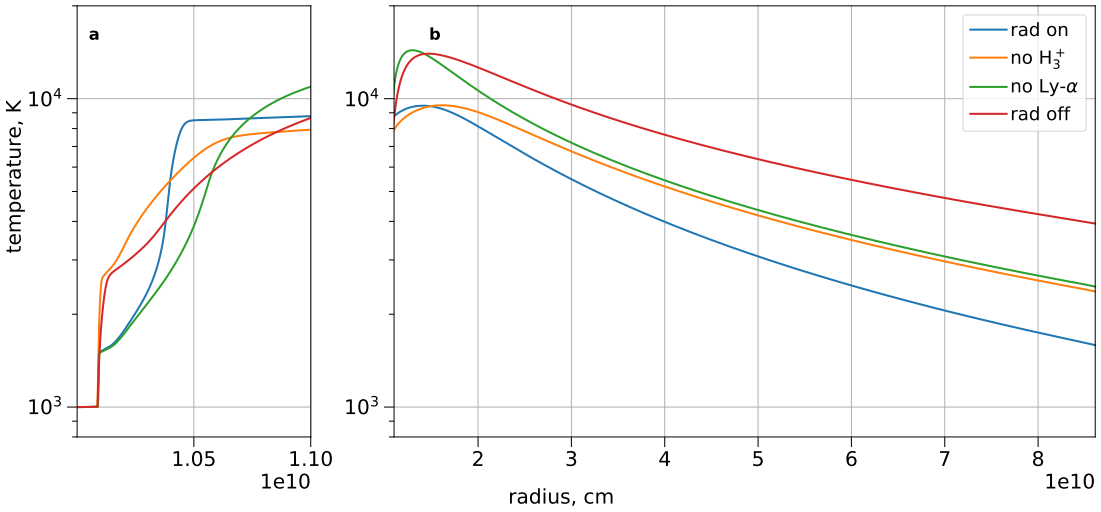

Figure 4.17 Temperature profiles for our fiduial hot Jupiter at 0.05au. The case with radiative cooling turned off (red line) reaches the highest mean outflow temperatures. Including either Lyman-α (orange line) or H_3^+ (green line) cooling produces similar mean outflow temperatures, yet only the case with Lyman-α turned off reaches temperatures exceeding 10^4 K. Lyman-α thermostats the temperature to $\sim 10^4$ K (orange and blue lines) in the region where it is significant ($\sim 1.04-1.4$ au, Figure 4.6). The full radiative heating model is shown in blue. In summary, H_3^+ cools the lower layer, and Lyman-α brings down the peak temperature. We note the temperature in the bolometric region is forced to 10^3 K, acting as a lower boundary condition.

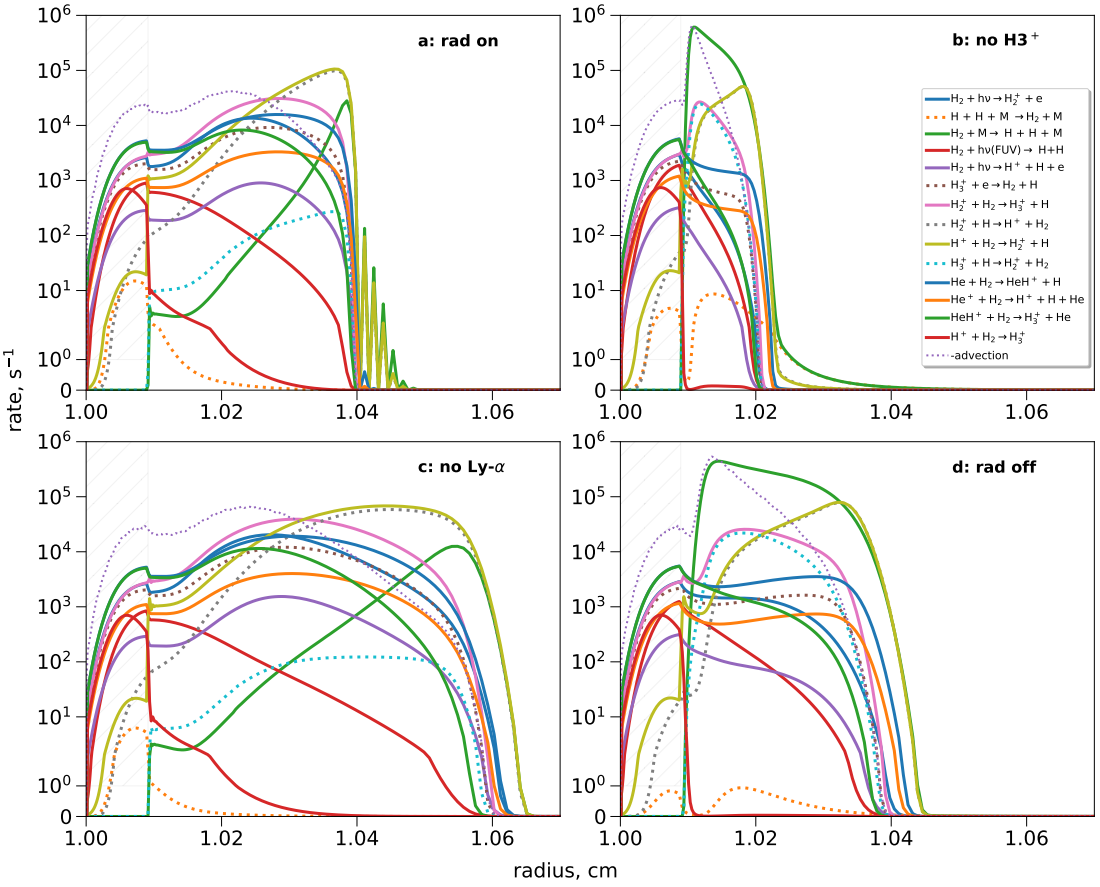

Figure 4.18 Reation rates involving molecular hydrogen. Panel a shows our fiducial hot Jupiter model with full radiative cooling. Panel b - only Lyman-α cooling, panel c - only H_3^+ cooling, and panel d - all radiative cooling turned off. The hatched region at the base denotes the bolometric region.

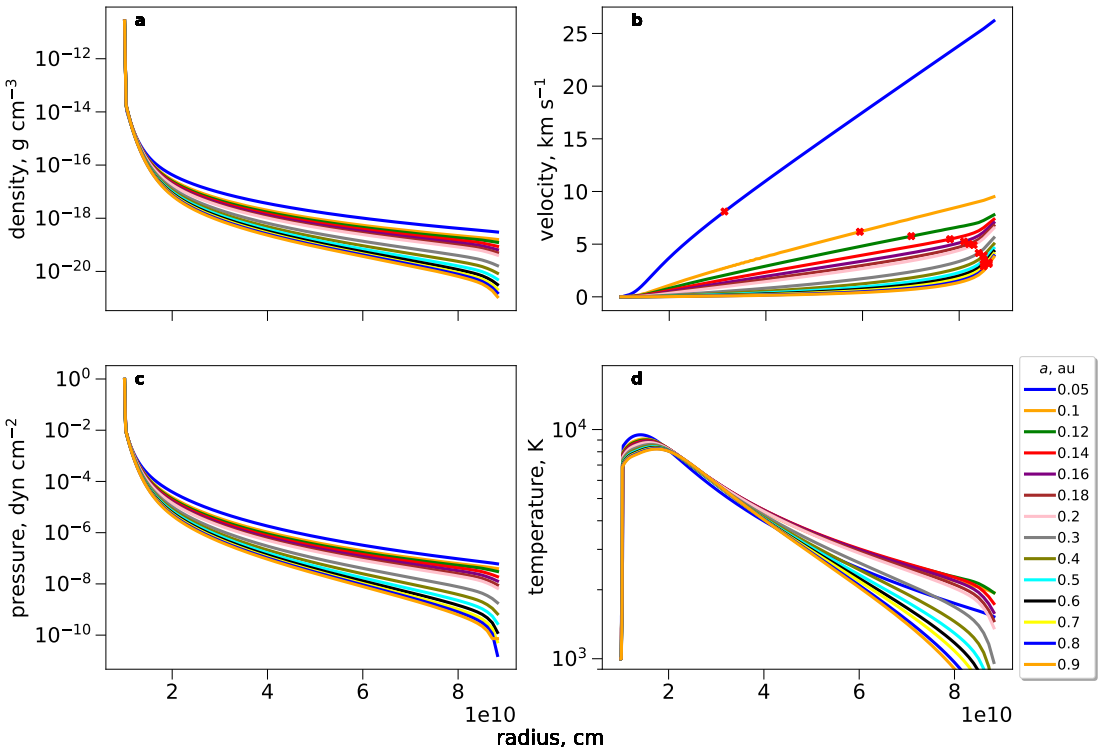

Figure 4.19 Density (panel a), velocity (panel b), pressure (panel c), and temperature (panel d) profiles for varying semimajor axes in the range 0.05-0.9 au.

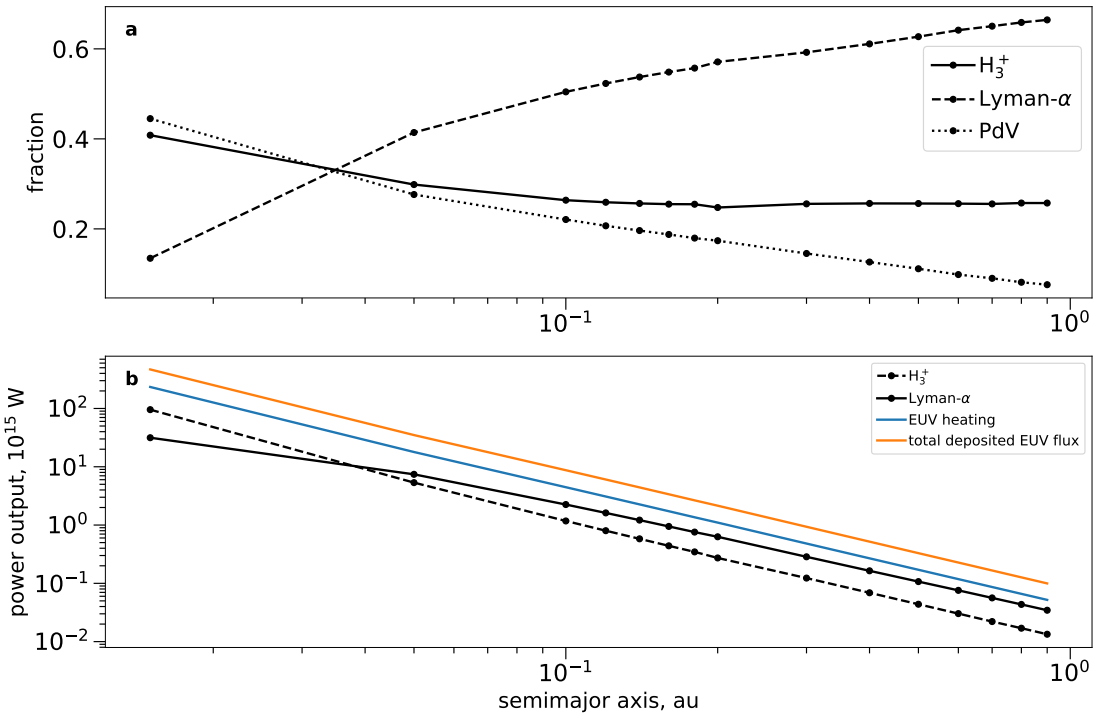

Figure 4.20 Panel a: fraction of the total heating (after accounting for energy required for ionization) that ultimately is balanced by PdV work (dotted), Lyman-α cooling (dashed), and H_3^+ cooling (solid). The PdV fraction decreases with distance, implying that the outflow is increasingly not energy-limited with greater distance from the star. The H_3^+ fraction remains relatively constant: we expect the power output from H_3^+ to go as a^{-2}. The increasing neutral fraction of the outflow causes an increase in the Lyman-α cooling as a fraction of the total avaiable heating. Panel b: the total power output for the model including both H_3^+ and Lyman-α cooling is plotted in black, dashed and solid lines respectively. The dark and light gray lines plot models with either mode of cooling turned on.

Figure 4.20 The total EUV flux from the star incident on the planet is plotted in orange. The heating rate (after ionization) is plotted in blue. The heating efficiency is the fraction of the total EUV flux that goes into heating the gas, as opposed to overcoming the ionization potential energy of the species in the gas. Over the entire semimajor axis range, the heating efficiency is approximately 0.53-0.55, making our initial estimate of 0.51 reasonable. The H_3^+ lines (dashed) scale with the EUV input power, while the Lyman-α decreases more slowly with distance than the input flux.

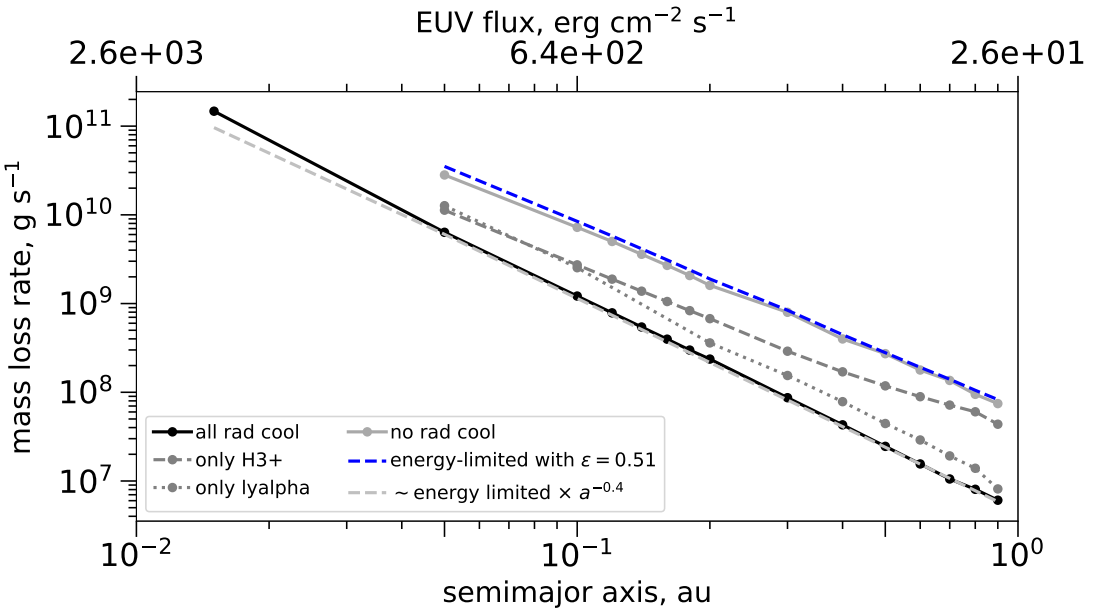

Figure 4.21 Mass-loss rate as a function of distance. Shown are the mass-loss rates $\dot{M} = \rho u r^2$ per solid angle for radiative cooling turned off (silver solid line), only H_3^+ cooling (grey dashed line), only Lyman-α cooling (grey dotted line) and for both H_3^+ and Lyman-α cooling turned on (black line). The mass-loss rates for planets with radiative cooling turned off approach the energy-limited mass-loss rate (blue dashed line, plotted for $\epsilon = 0.51$). The energy-limited mass-loss, scaled by a factor of 0.3 (a/0.05 au)$^{-0.4}$ approximates the mass loss for the full radiative cooling case (light gray dashed line).

Bibliography

[1] Auerbach, D., Cacak, R., Caudano, R., et al. 1977, Journal of Physics B: Atomic and Molecular Physics, 10, 3797

[2] Baraffe, I., Selsis, F., Chabrier, G., et al. 2004, Astronomy & Astrophysics, 419, L13

[3] Baulch, D., Cobos, C., Cox, R., et al. 1992, Journal of Physical and Chemical Reference Data, 21, 411

[4] Ben-Jaffel, L., & Ballester, G. 2013, Astronomy & Astrophysics, 553, A52

[5] Black, J. H. 1981, Monthly Notices of the Royal Astronomical Society, 197, 553

[6] Bohme, D., Mackay, G.-I., & Schiff, H. 1980, The Journal of Chemical Physics, 73, 4976

[7] Bourrier, V., Des Etangs, A. L., Dupuy, H., et al. 2013, Astronomy & Astrophysics, 551, A63

[8] Butler, R. P., Wright, J. T., Marcy, G. W., et al. 2006, Astrophysical Journal, 646, 505

[9] Cabot, S. H., Madhusudhan, N., Welbanks, L., Piette, A., & Gandhi, S. 2020, Monthly Notices of the Royal Astronomical Society, 494, 363

[10] Chambers, J. E. 1999, Monthly Notices of the Royal Astronomical Society, 304, 793

[11] Chapman, S., & Cowling, T. G. 1990, The mathematical theory of non-uniform gases: an account of the kinetic theory of viscosity, thermal conduction and diffusion in gases (Cambridge university press)

[12] Chatterjee, S., Ford, E. B., Matsumura, S., & Rasio, F. A. 2008, Astrophysical Journal, 686, 580

[13] Chiang, E. I., & Murray, N. 2002, Astrophysical Journal, 576, 473

[14] Clarke, C., & Carswell, B. 2007, Principles of astrophysical fluid dynamics (Cambridge University Press)

[15] Cubillos, P. E., Fossati, L., Koskinen, T., et al. 2020, The Astronomical Journal, 159, 111

[16] Datz, S., Sundström, G., Biedermann, C., et al. 1995, Physical Review Letters, 74, 896

[17] Dawson, R. I., & Chiang, E. 2014, Science, 346, 212

[18] Dawson, R. I., & Murray-Clay, R. A. 2013, Astrophysical Journal Letters, 767, L24

[19] Fischer, D. A., & Valenti, J. 2005, Astrophysical Journal, 622, 1102

[20] Des Etangs, A. L., Ehrenreich, D., Vidal-Madjar, A., et al. 2010, Astronomy & Astrophysics, 514, A72

[21] Ford, E. B., & Rasio, F. A. 2008, Astrophysical Journal, 686, 621

[22] Ford, A. L., Docken, K. K., & Dalgarno, A. 1975, Astrophysical Journal, vol. 195, Feb. 1, 1975, pt. 1, p. 819-824. NSF-supported research;, 195, 819

[23] Ford, E. B., Havlickova, M., & Rasio, F. A. 2001, Icarus, 150, 303

[24] Fortney, J. J., Marley, M. S., & Barnes, J. W. 2007, The Astrophysical Journal, 659, 1661

[25] Frelikh, R. & Murray-Clay, R. 2017 Astronomical Journal 154, e98

[26] Frelikh, R., Jang, H., Murray-Clay, R. & Petrovich, C. 2019, Astrophysical Journal, 884, eL47

[27] García-Muñoz, A. 2007, Planetary and Space Science, 55, 1426

[28] Gaudi, B. S., Seager, S., & Mallen-Ornelas, G. 2005, Astrophysical Journal, 623, 472

[114] Goldreich, P., Lithwick, Y., & Sari, R. 2004, Annual Review of Astronomy and Astrophysics, 42, 549

[30] Hwang, J., Chatterjee, S., Lombardi, J., Jr., Steffen, J. H., & Rasio, F. 2018, Astrophysical Journal, 852, 41

[31] Johnson, J. A., Aller, K. M., Howard, A. W., & Crepp, J. R. 2010, Publications of the Astronomical Society of the Pacific, 122, 905

[32] Gaudi, B. S. 2005, Astrophysical Journal, 628, L73

[33] Ginzburg, S., Schlichting, H. E., & Sari, R. 2016, The Astrophysical Journal, 825, 29

[34] —. 2018, Monthly Notices of the Royal Astronomical Society, 476, 759

[35] Ham, D. O., Trainor, D. W., & Kaufman, F. 1970, The Journal of Chemical Physics, 53, 4395

[36] Herzberg, G., & Monfils, A. 1961, Journal of Molecular Spectroscopy, 5, 482

[37] Hoadley, K., France, K., Alexander, R. D., McJunkin, M., & Schneider, P. 2015, The Astrophysical Journal, 812, 41

[38] Jensen, A. G., Cauley, P. W., Redfield, S., Cochran, W. D., & Endl, M. 2018, The Astronomical Journal, 156, 154

[39] Jo, Y.-S., Seon, K.-I., Min, K.-W., Edelstein, J., & Han, W. 2017, The Astrophysical Journal Supplement Series, 231, 21

[40] Johnstone, C. P., Güdel, M., Lammer, H., & Kislyakova, K. G. 2018, Astronomy & Astrophysics, 617, A107

[41] Juric, M., & Tremaine, S. 2008, Astrophysical Journal, 686, 603

[42] Karpas, Z., Anicich, V., & Huntress Jr, W. 1979, The Journal of Chemical Physics, 70, 2877

[43] Koskinen, T. T., Aylward, A. D., & Miller, S. 2007, Nature, 450, 845

[44] Kozai, Y. 1962, The Astronomical Journal, 67, 591

[45] Kuramoto, K., Umemoto, T., & Ishiwatari, M. 2013, Earth and Planetary Science Letters, 375, 312

[46] Lammer, H., Selsis, F., Ribas, I., et al. 2003, The Astrophysical Journal, 598, L121

[47] Landau, L. D., & Lifshitz, E. M. 2013, Fluid Mechanics: Landau and Lifshitz: Course of Theoretical Physics, Volume 6, Vol. 6 (Elsevier)

[48] Landshoff, R. 1955, A numerical method for treating fluid flow in the presence of shocks, Tech. rep., Los Alamos National Lab.(LANL), Los Alamos, NM (United States)

[49] Lidov, M. L. 1962, Planetary and Space Science, 9, 719

[50] Linsky, J. L., Yang, H., France, K., et al. 2010, The Astrophysical Journal, 717, 1291

[51] Lodders, K. 2020, in Oxford Research Encyclopedia of Planetary Science

[52] Lovis, C., & Fischer, D. 2010, Radial Velocity Techniques for Exoplanets, ed. S. Seager, 27–53

[53] McCann, J., Murray-Clay, R. A., Kratter, K., & Krumholz, M. R. 2019, The Astrophysical Journal, 873, 89

[54] McCann, J. R. 2021, Atmospheric Escape in the Hydrodynamic Limit (University of California, Santa Barbara)

[55] Miller, T. M., Moseley, J., Martin, D., & McDaniel, E. 1968, Physical Review, 173, 115

[56] Miller, S., Stallard, T., Tennyson, J., & Melin, H. 2013, The Journal of Physical Chemistry A, 117, 9770–9777, doi: 10.1021/jp312468b

[57] Miller, S., Tennyson, J., Geballe, T. R., & Stallard, T. 2020, Reviews of Modern Physics, 92, doi: 10.1103/revmodphys.92.035003

[58] Murray-Clay, R. A., Chiang, E. I., & Murray, N. 2009, The Astrophysical Journal, 693, 23

[59] Naoz, S., Farr, W. M., Lithwick, Y., Rasio, F. A., & Teyssandier, J. 2011, Nature, 473, 187

[60] Nielsen, E. L., De Rosa, R. J., Macintosh, B., et al. 2019, Astronomical Journal, 158, 13

[61] Ninan, J. P., Stefansson, G., Mahadevan, S., et al. 2020, The Astrophysical Journal, 894, 97

[62] Nishiguchi, A., & Yabe, T. 1983, Journal of Computational Physics, 52, 390

[63] Oklopčić, A., & Hirata, C. M. 2018, The Astrophysical Journal Letters, 855, L11

[64] Osterbrock, D. E., & Ferland, G. J. 2006, Astrophysics of gaseous nebulae and active galactic nuclei, 2nd, University Science Books, Sausalito, CA

[65] Owen, J. E., & Adams, F. C. 2014, Monthly Notices of the Royal Astronomical Society, 444, 3761

[66] Owen, J. E., & Wu, Y. 2017, The Astrophysical Journal, 847, 29

[67] Petrovich, C., & Tremaine, S. 2016, Astrophysical Journal, 829, 132

[68] Petrovich, C., Tremaine, S., & Rafikov, R. 2014, Astrophysical Journal, 786, 101

[69] Pontoppidan, K. M., Boogert, A. C., Fraser, H. J., et al. 2008, The Astrophysical Journal, 678, 1005

[70] Rafikov, R. R. 2006, Astrophysical Journal, 648, 666

[71] Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2007, Numerical recipes 3rd edition: The art of scientific computing (Cambridge university press)

[72] Rasio, F. A., & Ford, E. B. 1996, Science, 274, 954

[73] Santos, N. C., Israelian, G., & Mayor, M. 2004, Astronomy and Astrophysics, 415, 1153

[74] Richards, P., Fennelly, J., & Torr, D. 1994, Journal of Geophysical Research: Space Physics, 99, 8981

[75] Schauer, M., Jefferts, S., Barlow, S., & Dunn, G. 1989, The Journal of chemical physics, 91, 4593

[76] Schunk, R., & Nagy, A. 2009, Ionospheres: physics, plasma physics, and chemistry (Cambridge university press)

[77] Sing, D. K., Lavvas, P., Ballester, G. E., et al. 2019, The Astronomical Journal, 158, 91

[78] Spake, J. J., Sing, D. K., Evans, T. M., et al. 2018, Nature, 557, 68

[79] Stone, J. M., & Norman, M. L. 1992, Astrophysical Journal Supplement Series (ISSN 0067-0049), vol. 80, no. 2, June 1992, p. 753-790. Research supported by University of Illinois., 80, 753

[80] Storey, P., & Hummer, D. 1995, Monthly Notices of the Royal Astronomical Society, 272, 41

[81] Sundström, G., Mowat, J., Danared, H., et al. 1994, Science, 263, 785

[82] Takeda, G., & Rasio, F. A. 2005, Astrophysical Journal, 627, 1001

[83] Tang, H., Dong, W., & Agrawal, A. 2018, Journal of Computational and Applied Mathematics, 333, 404

[84] Theard, L. P., & Huntress Jr, W. T. 1974, The Journal of Chemical Physics, 60, 2840

[85] Thorngren, D. P., Fortney, J. J., Murray-Clay, R. A., & Lopez, E. D. 2016, Astrophysical Journal, 831, 64

[86] Trammell, G. B., Li, Z.-Y., & Arras, P. 2014, The Astrophysical Journal, 788, 161

[87] Tripathi, A., Kratter, K. M., Murray-Clay, R. A., & Krumholz, M. R. 2015, The Astrophysical Journal, 808, 173

[88] Vidal-Madjar, A., Des Etangs, A. L., Désert, J.-M., et al. 2003, Nature, 422, 143

[89] Vidal-Madjar, A., Désert, J.-M., Des Etangs, A. L., et al. 2004, The Astrophysical Journal, 604, L69

[90] VonNeumann, J., & Richtmyer, R. D. 1950, Journal of applied physics, 21, 232

[91] Watson, A. J., Donahue, T. M., & Walker, J. C. 1981, Icarus, 48, 150

[92] Winn, J. N., & Fabrycky, D. C. 2015, Annual Review of Astronomy and Astrophysics, Vol 53, 53, 409

[93] Wright, J. T., Fakhouri, O., Marcy, G. W., et al. 2011, Publications of the Astronomical Society of the Pacific, 123, 412

[94] Wu, Y. Q., & Lithwick, Y. 2011, Astrophysical Journal, 735, 109

[95] Wyttenbach, A., Mollière, P., Ehrenreich, D., et al. 2020, Astronomy & Astrophysics, 638, A87

[96] Yabe, T., & Aoki, T. 1991, Computer physics communications, 66, 219

[97] Yabe, T., Tanaka, R., Nakamura, T., & Xiao, F. 2001, Monthly Weather Review, 129, 332

[98] Yan, F., Wyttenbach, A., Casasayas-Barris, N., et al. 2021, Astronomy & Astrophysics, 645, A22

[99] Yelle, R. V. 2004, Icarus, 170, 167

[100] Yousif, F., & Mitchell, J. 1989, Physical Review A, 40, 4318

[101] Adams, F. C., Hollenbach, D., Laughlin, G., & Gorti, U. 2004, Astrophysical Journal, 611, 360

[102] Alexander, R. D., Clarke, C. J., & Pringle, J. E. 2005, Monthly Notices of the Royal Astronomical Society, 358, 283

[103] —. 2006, Monthly Notices of the Royal Astronomical Society, 369, 229

[104] Alexander, R., Pascucci, I., Andrews, S., Armitage, P., & Cieza, L. 2014, Protostars and Planets VI, 475

[105] Andrews, S. M., Wilner, D. J., Hughes, A. M., Qi, C., & Dullemond, C. P. 2009, Astrophysical Journal, 700, 1502

[106] Beckwith, S. V. W., Sargent, A. I., Chini, R. S., & Guesten, R. 1990, Astronomical Journal, 99, 924

[107] Bottinelli, S., Boogert, A. C. A., Bouwman, J., et al. 2010, Astrophysical Journal, 718, 1100

[108] Dawson, R. I., Lee, E. J., & Chiang, E. 2016, Astrophysical Journal, 822, 54

[109] Desch, S. J. 2007, Astrophysical Journal, 671, 878

[110] Draine, B. T. 2003, Annual Review of Astronomy and Astrophysics, 41, 241

[111] Espaillat, C., Muzerolle, J., Najita, J., et al. 2014, Protostars and Planets VI, 497

[112] Ford, E. B., & Chiang, E. I. 2007, Astrophysical Journal, 661, 602

[113] Goldreich, P., Lithwick, Y., & Sari, R. 2004, Astrophysical Journal, 614, 497

[114] —. 2004, Annual Review of Astronomy and Astrophysics, 42, 549

[115] Haisch, Jr., K. E., Lada, E. A., & Lada, C. J. 2001, Astrophysical Journal, Letters, 553, L153

[116] Hayashi, C. 1981, Progress of Theoretical Physics Supplement, 70, 35

[117] Lambrechts, M., & Johansen, A. 2012, Astronomy and Astrophysics, 544, A32

[118] Lambrechts, M., Johansen, A., & Morbidelli, A. 2014, Astronomy and Astrophysics, 572, A35

[119] Lodders, K. 2003, Astrophysical Journal, 591, 1220

[120] Mamajek, E. E. 2009, in American Institute of Physics Conference Series, Vol. 1158, American Institute of Physics Conference Series, ed. T. Usuda, M. Tamura, & M. Ishii, 3–10

[121] Mordasini, C. 2014, Astronomy and Astrophysics, 572, A118

[122] Movshovitz, N., Bodenheimer, P., Podolak, M., & Lissauer, J. J. 2010, Icarus, 209, 616

[123] Nettelmann, N., Helled, R., Fortney, J. J., & Redmer, R. 2013, Planetary Space Science, 77, 143

[124] Öberg, K. I., Boogert, A. C. A., Pontoppidan, K. M., et al. 2008, Astrophysical Journal, 678, 1032

[125] Ormel, C. W., & Klahr, H. H. 2010, Astronomy and Astrophysics, 520, A43

[126] Owen, J. E., Clarke, C. J., & Ercolano, B. 2012, Monthly Notices of the Royal Astronomical Society, 422, 1880

[127] Piso, A.-M. A., & Youdin, A. N. 2014, Astrophysical Journal, 786, 21

[128] Piso, A.-M. A., Youdin, A. N., & Murray-Clay, R. A. 2015, Astrophysical Journal, 800, 82

[129] Pollack, J. B., McKay, C. P., & Christofferson, B. M. 1985, Icarus, 64, 471

[130] Pollack, J. B., Hubickyj, O., Bodenheimer, P., et al. 1996, Icarus, 124, 62

[131] Rafikov, R. R. 2011, Astrophysical Journal, 727, 86

[132] Rosenthal, M. M. et al. 2017 in prep, Astrophysical Journal

[133] Schlichting, H. E. 2014, Astrophysical Journal, Letters, 795, L15

[134] Shakura, N. I., & Sunyaev, R. A. 1973, Astronomy and Astrophysics, 24, 337

[135] Soubiran, F., & Militzer, B. 2015, Astrophysical Journal, 806, 228

[136] Venturini, J., Alibert, Y., Benz, W., & Ikoma, M. 2015, Astronomy and Astrophysics, 576, A114

Appendix A

Appendix for Efficiency of Hydrodynamic Escape

A.1 Tidal gravity term

Consider the substellar ray of length a from the center of the planet (with mass M_P, radius R_p) to the center of the star (with mass $M_\star$). Consider a test particle that is a distance r from the planet, such that $r << a$. In the frame centered on the planet and co-rotating with the planet, the forces on the particle are: the gravitational force from the planet ($-GM_p/r^2$) and the star ($-GM_\star/(a-r)^2$), as well as the centrifugal force $\Omega \times \Omega \times (r - r_{CM})$, where $r_{CM} = aM_\star/(M_\star + M_P)$. Though present in the co-rotating frame, we neglect the Coriolis term as we are working in one dimensions. Here, $\Omega = G(M_P + M_\star)/a^3$ is the Keplerian orbital frequency. The centrifugal force acting

on the particle is therefore:

$$F_{\text{cent}} = \frac{G(M_P + M_\star)}{a^3}\left(r - a\frac{M_\star}{M_P + M_\star}\right) = \frac{rG(M_P + M_\star)}{a^3} - \frac{GM_\star}{a^2}. \tag{A.1}$$

We make the approximation $M_p << M_\star$ in the first term of the centrifugal force:

$$F_{\text{cent}} \approx \frac{rGM_\star}{a^3} - \frac{GM_\star}{a^2}. \tag{A.2}$$

Liekwise, we make the approximation $r << a$ in the gravitational term due to the star: $GM_\star/(a-r)^2 = GM_\star/(a^2(1-r/a)^2) \approx GM_\star/a^2(1+2r/a) = GM_\star/a^2 + 2rGM_\star/a^3$. Together, after making these approximations the total forces on the particle are:

$$F_{\text{tot}} \approx -\frac{GM_P}{r^2} + \frac{3rGM_\star}{a^3}. \tag{A.3}$$

The last term in Eq. (A.3) is the tidal term: it is the sum of the gravitational term due to the star and the centrifugal term.

A.2 Diffusion

The mass conservation equation in one-dimensional spherical coordinates for the bulk gas comprising our atmosphere is:

$$\frac{\partial \rho}{\partial t} + \frac{1}{r^2}\frac{\partial}{\partial r}(\rho u r^2) = 0 \tag{A.4}$$

Here, $\rho = \sum_s \rho_s$ is the total mass density of the gas, and it is defined as the sum of all the individual densities ρ_s of the species comprising the fluid. The bulk velocity u describes the bulk motion of fluid parcels without any individual change in composition.

For each species s of our atmosphere, we can write a separate conservation equation that includes a diffusive term on the left-hand side and a chemical source term on the right-hand side:

$$\frac{\partial \rho_s}{\partial t} + \frac{1}{r^2}\frac{\partial}{\partial r}(\rho_s u r^2) + \frac{1}{r^2}\frac{\partial}{\partial r}(\rho_s (u_s + v_s) r^2) = S \tag{A.5}$$

The conservation equations for each species can be added up together, recovering Eq. (A1): the diffusive terms and chemical source terms all add up to zero after summing over all of the species. Now we focus our discussion on molecular diffusion, so we drop the chemical source term and the advective term due to the bulk gas velocity from Eq. (A2). We keep the diffusive velocity term with the molecular (u_s) component:

$$\frac{\partial \rho_s}{\partial t} + \frac{1}{r^2}\frac{\partial}{\partial r}(\rho_s u_s r^2) = 0. \tag{A.6}$$

We substitute the number density $n_s = \rho_s / m_s$, with m_s the molecular weight of species s in Eq. 4.6 to keep our discussion in terms of concentration rather than mass density:

$$\frac{\partial n_s}{\partial t} = -\frac{1}{r^2}\frac{\partial}{\partial r}(n_s u_s r^2) = -\frac{1}{r^2}\frac{\partial}{\partial r}(\Phi_s r^2). \tag{A.7}$$

The flux $\Phi_s = n_s u_s$ describes the flow of species s across a cell boundary due to diffusion. We will now express the diffusive velocity u_s in terms of a diffusion coefficient and a concentration and temperature gradient[1]. For a binary gas mixture of gases of concentration n_1 and n_2, with the total gas concentration $N = n_1 + n_2$, the diffusion

[1]We neglect the barodiffusion coefficient, as it is present only in fluids with significant pressure gradients (for example, due to external fields) not applicable to our problem.

velocity of gas 1 relative to gas 2 is written as [11, Eq. 14.1.1]:

$$u_s = -\frac{N^2}{n_1 n_2} D_{12} \left(\frac{\partial n_{10}}{\partial r} + \frac{n_1 n_2 (m_2 - m_1)}{N\rho} \frac{\partial \log P}{\partial r} + k_T \frac{1}{T} \frac{\partial T}{\partial r} \right), \tag{A.8}$$

where $n_{10} = n_1/N$. Here, the first two terms on the right-hand side describe the concentration gradient, and the last term describes thermal diffusion with k_T the coefficient of thermal diffusion[2]. This can be simplified to:

$$u_s = -D_{12} \left[\frac{N}{n_2} \left(\frac{1}{n_1} \frac{\partial n_1}{\partial r} - \frac{1}{N} \frac{\partial N}{\partial r} \right) + \left(\frac{m_2 - m_1}{\mu} \frac{1}{P} \frac{\partial P}{\partial r} \right) + \frac{N^2}{n_1 n_2} k_T \frac{1}{T} \frac{\partial T}{\partial r} \right], \tag{A.9}$$

where $\mu = \rho/N$ is the mean molecular weight of the entire gas. We can substitute the thermal diffusion factor α for the coefficient of thermal diffusion $k_T = \alpha n_1 n_2 / N^2$. We are interested in the diffusion of species s against the entire fluid. We can associate gas 1 with species s and gas 2 with the total gas, giving

$$u_s = -D_s \left(\frac{1}{n_s} \frac{\partial n_s}{\partial r} - \frac{1}{N} \frac{\partial N}{\partial r} + \left(1 - \frac{m_s}{\mu} \right) \frac{1}{P} \frac{\partial P}{\partial r} + \frac{\alpha}{T} \frac{\partial T}{\partial r} \right), \tag{A.10}$$

where the diffusion coefficient D_s is for the diffusion of species s against the total gas. This matches Eq. (24) of Johnstone et al. [40], where $\Phi_s = n_s u_s$. We now describe the method of solving this equation, which is adapted from Johnstone et al. [40]. From the ideal gas law $P = NkT$, $\partial P/\partial r = Nk\partial T/\partial r + kT\partial N/\partial r$. Equation A.10 becomes:

$$u_s = -D_s \left(\frac{1}{n_s} \frac{\partial n_s}{\partial r} - \frac{1}{N} \frac{\partial N}{\partial r} + \left(1 - \frac{m_s}{\mu} \right) \frac{1}{P} \left(Nk \frac{\partial T}{\partial r} + kT \frac{\partial N}{\partial r} \right) + \frac{\alpha}{T} \frac{\partial T}{\partial r} \right), \tag{A.11}$$

which, after simplifying, gives us the flux:

$$\Phi_s = -D_s \left(\frac{\partial n_s}{\partial r} - \frac{m_s}{\mu} \frac{n_s}{N} \frac{\partial N}{\partial r} + \left(\alpha + 1 - \frac{m_s}{\mu} \right) \frac{n_s}{T} \frac{\partial T}{\partial r} \right) \tag{A.12}$$

[2]Landau & Lifshitz [47] call it the thermal diffusion ratio, with $k_T D$ the thermal diffusion coefficient.

.

We wish to solve Equation (A.7) by discretizing it on a grid. From now on, we drop the subscript s when referring to species s, and we will use the subscript i when referring to cell i. We use the Crank-Nicolson method by combining the derivatives at timestep n with timestep $n+1$:

$$\frac{n_i^{n+1} - n_i^n}{\Delta t} = \frac{1}{2}\left(\frac{\partial n}{\partial t}\right)_i^{n+1} + \frac{1}{2}\left(\frac{\partial n}{\partial t}\right)_i^{n}. \tag{A.13}$$

Using Equation (A.7), we write the time derivatives of n at timestep n as the differences of the diffusion fluxes at the boundaries of cell i:

$$\left(\frac{\partial n}{\partial t}_i\right)^n = -\frac{1}{r_i^2}\left(\frac{r_{i+1/2}^2\Phi_{i+1/2}^n - r_{i-1/2}^2\Phi_{i-1/2}^n}{r_{i+1/2} - r_{i-1/2}}\right), \tag{A.14}$$

and analogously for timestep $n+1$. Then, we can use Equation (A.14) in Equation (A.13):

$$\frac{n_i^{n+1} - n_i^n}{\Delta t} = -\frac{1}{2r_i^2}\left(\frac{r_{i+1/2}^2\Phi_{i+1/2}^{n+1} - r_{i-1/2}^2\Phi_{i-1/2}^{n+1}}{\Delta r_i} + \frac{r_{i+1/2}^2\Phi_{i+1/2}^{n} - r_{i-1/2}^2\Phi_{i-1/2}^{n}}{\Delta r_i}\right). \tag{A.15}$$

In other words,

$$\begin{aligned} n_i^{n+1} + \frac{\Delta t}{2r_i^2\Delta r_i}\left(r_{i+1/2}^2\Phi_{i+1/2}^{n+1} - r_{i-1/2}^2\Phi_{i-1/2}^{n+1}\right) = \\ n_i^n - \frac{\Delta t}{2r_i^2\Delta r_i}\left(r_{i+1/2}^2\Phi_{i+1/2}^{n} - r_{i-1/2}^2\Phi_{i-1/2}^{n}\right). \end{aligned} \tag{A.16}$$

We are interested the fluxes at the boundaries of cell i, $\Phi_{i+1/2}$ and $\Phi_{i-1/2}$. We write

Equation (A.12) for the right-hand boundary of cell i:

$$\Phi_{i+1/2} = \frac{n_i - n_{i+1}}{r_{i+1} - r_i} D_{i+1/2} + \left[\frac{1}{2} \frac{m}{\mu_{i+1/2}} \frac{1}{N_{i+1/2}} \left(\frac{\partial N}{\partial r} \right)_{i+1/2} D_{i+1/2} - \frac{1}{2} \left(\alpha + 1 - \frac{m}{\mu_{i+1/2}} \right) \frac{1}{T_{i+1/2}} \left(\frac{\partial T}{\partial r} \right)_{i+1/2} D_{i+1/2} \right] \times (n_{i+1} + n_i) \,, \tag{A.17}$$

where we approximate $n_{i+1/2} = (n_{i+1} + n_i)/2$ and $(\partial n/\partial r)_{i+1/2} = (n_{i+1} - n_i)/(r_{i+1} - r_i)$. We can call the factor in square brackets in Equation (A.17) $C_{i+1/2}$, and we group the factors of n_{i+1} and n_i:

$$\Phi_{i+1/2} = n_i \left(\frac{D_{i+1/2}}{r_{i+1} - r_i} + C_{i+1/2} \right) + n_{i+1} \left(-\frac{D_{i+1/2}}{r_{i+1} - r_i} + C_{i+1/2} \right) . \tag{A.18}$$

Similarly, for the flux on the left-hand boundary of cell i we have:

$$\Phi_{i-1/2} = \frac{n_{i-1} - n_i}{r_i - r_{i-1}} D_{i-1/2} + \left[\frac{1}{2} \frac{m}{\mu_{i-1/2}} \frac{1}{N_{i-1/2}} \left(\frac{\partial N}{\partial r} \right)_{i-1/2} D_{i-1/2} - \frac{1}{2} \left(\alpha + 1 - \frac{m}{\mu_{i-1/2}} \right) \frac{1}{T_{i-1/2}} \left(\frac{\partial T}{\partial r} \right)_{i-1/2} D_{i-1/2} \right] \times (n_{i-1} + n_i) \,. \tag{A.19}$$

Then,

$$\Phi_{i-1/2} = n_i \left(-\frac{D_{i-1/2}}{r_i - r_{i-1}} + C_{i-1/2} \right) + n_{i-1} \left(\frac{D_{i-1/2}}{r_i - r_{i-1}} + C_{i-1/2} \right) , \tag{A.20}$$

where the factor in square brackets in Equation (A.19) is $C_{i-1/2}$. We can now use the expressions for the boundary fluxes in Equation (A.16). The diffusion coefficients and C_is are assumed to vary slowly with respect to time, thus they are calculated at the

beginning of a timestep.

$$\underline{n_i^{n+1}} + \frac{\Delta t}{2r_i^2\Delta r_i} r_{i+1/2}^2 \underline{n_i^{n+1}} \left(\frac{D_{i+1/2}}{r_{i+1}-r_i} + C_{i+1/2} \right) +$$
$$\frac{\Delta t}{2r_i^2\Delta r_i} r_{i+1/2}^2 \underline{\underline{n_{i+1}^{n+1}}} \left(-\frac{D_{i+1/2}}{r_{i+1}-r_i} + C_{i+1/2} \right) - \frac{\Delta t}{2r_i^2\Delta r_i} r_{i-1/2}^2 \underline{n_i^{n+1}} \left(-\frac{D_{i-1/2}}{r_i-r_{i-1}} + C_{i-1/2} \right)$$
$$- \frac{\Delta t}{2r_i^2\Delta r_i} r_{i-1/2}^2 \underline{\underline{n_{i-1}^{n+1}}} \left(\frac{D_{i-1/2}}{r_i-r_{i-1}} + C_{i-1/2} \right) = \underset{\sim}{n_i^n}$$
$$- \frac{\Delta t}{2r_i^2\Delta r_i} r_{i+1/2}^2 \underset{\sim}{n_i^n} \left(\frac{D_{i+1/2}}{r_{i+1}-r_i} + C_{i+1/2} \right) - \frac{\Delta t}{2r_i^2\Delta r_i} r_{i+1/2}^2 \underset{\sim\sim}{n_{i+1}^n} \left(-\frac{D_{i+1/2}}{r_{i+1}-r_i} + C_{i+1/2} \right)$$
$$+ \frac{\Delta t}{2r_i^2\Delta r_i} r_{i-1/2}^2 \underset{\sim}{n_i^n} \left(-\frac{D_{i-1/2}}{r_i-r_{i-1}} + C_{i-1/2} \right) + \frac{\Delta t}{2r_i^2\Delta r_i} r_{i-1/2}^2 \underset{\cdots}{n_{i-1}^n} \left(\frac{D_{i-1/2}}{r_i-r_{i-1}} + C_{i-1/2} \right). \tag{A.21}$$

We wish to write Equation (A.21) in the form $a_i n_{i-1}^{n+1} + b_i n_i^{n+1} + c_i n_{i+1}^{n+1} + d_i = 0$. Thus we have:

$$\begin{aligned}
a_i &= -\frac{\Delta t}{2r_i^2\Delta r_i} r_{i-1/2}^2 \left(\frac{D_{i-1/2}}{r_i-r_{i-1}} + C_{i-1/2} \right) \\
b_i &= 1 + \frac{\Delta t}{2r_i^2\Delta r_i} r_{i+1/2}^2 \left(\frac{D_{i+1/2}}{r_{i+1}-r_i} + C_{i+1/2} \right) - \frac{\Delta t}{2r_i^2\Delta r_i} r_{i-1/2}^2 \left(-\frac{D_{i-1/2}}{r_i-r_{i-1}} + C_{i-1/2} \right) \\
c_i &= \frac{\Delta t}{2r_i^2\Delta r_i} r_{i+1/2}^2 \left(-\frac{D_{i+1/2}}{r_{i+1}-r_i} + C_{i+1/2} \right) \\
d_i &= a_i n_{i-1}^n - (2-b_i) n_i^n + c_i n_{i+1}^n.
\end{aligned} \tag{A.22}$$

The polynomial $a_i n_{i-1}^{n+1} + b_i n_i^{n+1} + c_i n_{i+1}^{n+1} + d_i = 0$ represents a system of equations that can be solved via LU decomposition and forward substitution, followed by back substitution. This is a standard way of solving a linear system of equations.

First, we loop from the base of the atmosphere:

$$\begin{aligned} u_i &= c_i \\ l_i &= a_i/d_{i-1} \\ d_i &= b_i - l_i u_{i-1} \\ z_i &= r_i - l_i z_{i-1}, \end{aligned} \tag{A.23}$$

then we back substitute, starting with the outer boundary condition $x_{N-1} = z_{N-1}/d_{N-1}$, and then:

$$n_i = (z_i - u_i n_{i+1})/d_i. \tag{A.24}$$

A.3 Chemical network solver

We label each reaction i in Table 4.1 by its rate constant k_i, which has units of $cm^3 s^{-1}$ for bimolecular reactions and $cm^6 s^{-1}$ for termolecular reactions. The system of equations resulting from the chemical network presented in Table 4.1 is stiff. As the rate constants for the reactions can vary by many orders of magnitude, using a single timestep for explicitly integrating the reactions becomes computationally unfeasible, introducing errors and instabilities. Therefore, implicit methods are needed. We use semi-implicit Euler method, described in Section 17.5.2 of Press et al. [71], to solve the system of equations for the concentrations of the species in our chemical network. We now briefly describe the method.

A.3.1 Scheme

The concentrations of the species in our network evolve under a set of differential equations. We can write the set of concentrations of the species as a vector $\vec{y}$. Consider the set of equations for a vector quantity $\vec{y}$ whose derivative is a function of y that is not necessarily linear:

$$\vec{y'} = \overrightarrow{f(y)} \tag{A.25}$$

We can explicitly advance $\vec{y}$ to the next timestep: $\overrightarrow{y_{n+1}} = \vec{y_n} + dt\,\vec{y'_n}$ - the derivative is computed at timestep n. On the other hand, we can advance the solution implicitly by computing the derivative at timestep $n+1$ instead: $\overrightarrow{y_{n+1}} = \vec{y_n} + dt\,\overrightarrow{y'_{n+1}}$. The advantage of implicit schemes is that they maintain the stability of the solution (generally at a computational cost). We can use the implicit differencing scheme to advance Equation (A.25):

$$\overrightarrow{y_{n+1}} = \vec{y_n} + dt\,\overrightarrow{f(y_{n+1})}. \tag{A.26}$$

As $\overrightarrow{f(y_{n+1})}$ is unknown, we can construct a Jacobian matrix $\partial\vec{f}/\partial\vec{y}$ (evaluated at $\vec{y_n}$) to linearize the equations:

$$\overrightarrow{f(y_{n+1})} = \overrightarrow{f(y_n)} + \frac{\partial\vec{f}}{\partial\vec{y}} \cdot (\overrightarrow{y_{n+1}} - \vec{y_n}). \tag{A.27}$$

Together with Equation (A.26), this gives:

$$\overrightarrow{y_{n+1}} = \vec{y_n} + dt\,\left(\overrightarrow{f(y_n)} + \frac{\partial\vec{f}}{\partial\vec{y}} \cdot (\overrightarrow{y_{n+1}} - \vec{y_n})\right), \tag{A.28}$$

or, after rearranging,

$$\overrightarrow{y_{n+1}} = \vec{y_n} + \left(\mathbf{1} - dt\,\frac{\partial\vec{f}}{\partial\vec{y}}\right)^{-1} dt\,\overrightarrow{f(y_n)}, \tag{A.29}$$

where **1** is the identity matrix. We rewrite this as:

$$\overrightarrow{f(y_n)} = \left(\frac{1}{dt}\mathbf{1} - \frac{\partial \vec{f}}{\partial \vec{y}} \right) (\overrightarrow{y_{n+1}} - \overrightarrow{y_n}). \tag{A.30}$$

This equation is solved via LU decomposition with forward substitution, followed by back substitution, as implemented in section (2.3) of Press et al. [71].

The Jacobian $\partial \vec{f} / \partial \vec{y}$ is calculated analytically for our problem, as are the time derivatives of the concentrations on the left-hand side of Equation A.30. The integration timestep dt must be chosen carefully. Even though we wish to advance the solution by a desired timestep Δt, we need to cross the interval Δt with a series of timesteps dt_i, for each of which we solve Equation A.30. One cannot always use an arbitrarily large timestep in this method - the time interval must be subdivided into sufficiently small timesteps to obtain an accurate solution. The method attempts to increment the solution by dt_i, after which a polynomial extrapolation is performed as described for the Bulirsch-Stoer Method in section (17.3.2) of Press et al. [71] to check for the suitability of the timestep. The timestep is then adjusted as necessary according to the step size sequence defined in (17.5.41) of Press et al. [71].

As an example, consider separately a reaction network formed by reactions 8 and 11, with rate constants k_8 and k_{11}:

$$H_2^+ + H_2 \xrightarrow{k_8} H_3^+ + H \tag{A.31}$$

$$H_3^+ + H \xrightarrow{k_{11}} H_2^+ + H_2. \tag{A.32}$$

We label the time derivatives of the concentrations (n_s) of H_2^+, H_2, H_3^+, and H as $f(n_s)$:

$$\begin{aligned}
f(n_{H_2^+}) &= \frac{dn_{H_2^+}}{dt} = -k_8 n_{H_2^+} n_{H_2} + k_{11} n_{H_3^+} n_H, \\
f(n_{H_2}) &= \frac{dn_{H_2}}{dt} = -k_8 n_{H_2^+} n_{H_2} + k_{11} n_{H_3^+} n_H, \\
f(n_{H_3^+}) &= \frac{dn_{H_3^+}}{dt} = k_8 n_{H_2^+} n_{H_2} - k_{11} n_{H_3^+} n_H, \\
f(n_H) &= \frac{dn_H}{dt} = k_8 n_{H_2^+} n_{H_2} - k_{11} n_{H_3^+} n_H.
\end{aligned} \tag{A.33}$$

As applied to our problem, the Jacobian matrix $\partial \vec{f} / \partial \vec{y}$ is the matrix of partial derivatives of the time rates of change of the concentrations of each species with respect to the concentrations of all of the other species in the network. The Jacobian matrix for this system of equations is:

$$\begin{pmatrix}
\frac{df(n_{H_2^+})}{dn_{H_2^+}} & \frac{df(n_{H_2^+})}{dn_{H_2}} & \frac{df(n_{H_2^+})}{dn_{H_3^+}} & \frac{df(n_{H_2^+})}{dn_H} \\
\frac{df(n_{H_2})}{dn_{H_2^+}} & \frac{df(n_{H_2})}{dn_{H_2}} & \frac{df(n_{H_2})}{dn_{H_3^+}} & \frac{df(n_{H_2})}{dn_H} \\
\frac{df(n_{H_3^+})}{dn_{H_2^+}} & \frac{df(n_{H_3^+})}{dn_{H_2}} & \frac{df(n_{H_3^+})}{dn_{H_3^+}} & \frac{df(n_{H_3^+})}{dn_H} \\
\frac{df(n_H)}{dn_{H_2^+}} & \frac{df(n_H)}{dn_{H_2}} & \frac{df(n_H)}{dn_{H_3^+}} & \frac{df(n_H)}{dn_H}
\end{pmatrix} =
\begin{pmatrix}
-k_8 n_{H_2} & -k_8 n_{H_2^+} & k_{11} n_H & k_{11} n_{H_3^+} \\
-k_8 n_{H_2} & -k_8 n_{H_2^+} & k_{11} n_H & k_{11} n_{H_3^+} \\
k_8 n_{H_2} & k_8 n_{H_2^+} & -k_{11} n_H & k_{11} n_{H_3^+} \\
k_8 n_{H_2} & k_8 n_{H_2^+} & -k_{11} n_H & k_{11} n_{H_3^+}
\end{pmatrix} \tag{A.34}$$

Specifically in the Press et al. [71] implementation, we use the StepperSie method[3] with the parameters as defined in Press et al. [71]: rtol $= 1.0 \times 10^{-7}$, atol $= 1.0 \times 10^{-4}$rtol, $h_1 = 1.0 \times 10^{-6}$, hmin $= 0.0$.

[3] As the grid cells are independent, we are able to run the chemical network in parallel.

A.4 Hydrodynamic solver

The hydrodynamic equations can all be re-cast in the form:

$$\frac{\partial X}{\partial t} + u\frac{\partial X}{\partial r} = G_X. \tag{A.35}$$

Here, the left-hand side describes the advection of a quantity X such as density, momentum, or energy, and u is the bulk velocity of the fluid. The right-hand side contains the non-advective term. For the mass conservation equation it is:

$$G_\rho = -\frac{\rho}{r^2}\frac{\partial(ur^2)}{\partial r}, \tag{A.36}$$

which describes the compression and/or expansion of the gas. For the momentum equation it is:

$$G_u = -\frac{1}{\rho}\frac{\partial P}{\partial r} - \frac{GM}{r^2} + \frac{3GM_\star a}{r^3}, \tag{A.37}$$

where right-hand side describes the force due to a pressure gradient, gravity, and tidal gravity. Finally, for the energy equation it is:

$$G_e = -\frac{P}{\rho r^2}\frac{\partial(ur^2)}{\partial r} + \Gamma - \Lambda, \tag{A.38}$$

which contains the PdV work term, as well as the heating and cooling terms as described in Section 4.2.1. We can solve Equation A.35 by separating the advective part from the non-advective part. This method of solving our differential equation is called operator splitting: as the terms represent distinct physical processes, the numerical methods best suited for solving each of them may be different.

We now apply the CIP (Cubic-Interpolated Propagation) method [96, 97] to solve these equations. Going back to Equation A.35, we can take the spatial derivative

with respect to r of this equation to get:

$$\frac{\partial}{\partial r}\left(\frac{\partial X}{\partial t}\right) + \frac{\partial u}{\partial r}\frac{\partial X}{\partial r} + u\frac{\partial}{\partial r}\left(\frac{\partial X}{\partial r}\right) = \frac{\partial G_X}{\partial r}, \tag{A.39}$$

or written in the form of Equation A.35

$$\frac{\partial X'}{\partial t} + u\frac{\partial X'}{\partial r} = -\frac{\partial u}{\partial r}X' + G'_X, \tag{A.40}$$

where we define $X' = \partial X/\partial r$ and $G'_X = \partial G'_X/\partial r$. Equation A.40 can also be split into an advective part on the right-hand side, and a non-advective part on the left.

A.4.1 Non-advective part

A.4.1.1 Hydrodynamic equations

We write the non-advective substep of Equation A.35 as:

$$\frac{\partial \tilde{X}}{\partial t} = G_{X_n}. \tag{A.41}$$

Given X_n, we solve for $\tilde{X}$ by explicitly integrating with respect to time using a finite difference:

$$\tilde{X} = X + \Delta t G_X. \tag{A.42}$$

On our staggered grid described in Section 4.2.6.1, the mass conservation equation becomes:

$$\tilde{\rho}_i = \rho_i^n - \frac{\rho_i^n}{r_i^2}\frac{u_{i+1/2}^n - u_{i-1/2}^n}{\Delta r}\Delta t. \tag{A.43}$$

The momentum equation becomes:

$$\tilde{u}_{i+1/2} = u_{i+1/2}^n + (-\frac{2}{\rho_{i+1}^n + \rho_i^n}\frac{P_{i+1}^n - P_i^n}{\Delta r} + \frac{GM_P}{r_i^2} + \frac{3GM_\star a}{r_i^3})\Delta t, \tag{A.44}$$

where we've estimated $\rho_{i+1/2}$ as $(\rho_{i+1} + \rho_i)/2$. The energy equation becomes:

$$\tilde{e}_i = e_i^n - \frac{P_i^n}{\rho_i^n r_i^2} \frac{r_{i+1/2}^2 \tilde{u}_{1+1/2} + r_{i+1/2}^2 u_{1+1/2} - r_{i-1/2}^2 \tilde{u}_{1-1/2} - r_{i-1/2}^2 u_{1-1/2}}{2\Delta r} + \Gamma_i - \Lambda_i, \tag{A.45}$$

with the heating and cooling terms Γ_i and Λ_i as described in Section 4.2.

A.4.1.2 Derivative of the hydrodynamic quantities

As in Section A.4.1.1, we can write the non-advective substep of Equation A.40 as:

$$\frac{\partial \tilde{X}'}{\partial t} = -\frac{\partial u}{\partial r} X' + G'_X. \tag{A.46}$$

We can write the finite difference form for the non-advection X'_i associated with G'_i by approximating X' with a central difference of X:

$$\frac{\tilde{X}'_i - X'^n_i}{\Delta t} = \frac{1}{\Delta t} \left(\frac{\tilde{X}_{i+1} - \tilde{X}_{i-1}}{2\Delta r} - \frac{X^n_{i+1} - X^n_{i-1}}{2\Delta r} \right) = \frac{(\tilde{X}_{i+1} - X^n_{i+1}) + (X^n_{i-1} - \tilde{X}_{i-1})}{2\Delta r \Delta t}. \tag{A.47}$$

Using Equation (A.43), we can write the grouped terms of the right-hand side as:

$$\frac{\tilde{X}_{i+1} - X^n_{i+1}}{\Delta t} = G_{X,i+1} \tag{A.48}$$

$$\frac{\tilde{X}_{i-1} - X^n_{i-1}}{\Delta t} = G_{X,i-1} \tag{A.49}$$

So Equation A.47 becomes:

$$\frac{\tilde{X}'_i - X'^n_i}{\Delta t} = \frac{G_{X,i+1} - G_{X,i-1}}{2\Delta t} \tag{A.50}$$

The term on the right-hand side can be complicated to evaluate, so we will use Equation (A.47) instead, as all the terms are known. This allows us to avoid having to evaluate

G. We can write the finite difference form for the first term on the right-hand side of Equation (A.46) as:

$$-\tilde{X}_i' \frac{u_{i+1} - u_{i-1}}{2\Delta t}, \tag{A.51}$$

so the full equation becomes:

$$\frac{\tilde{X}_i' - X_i'^n}{\Delta t} = \frac{(\tilde{X}_{i+1} - X_{i+1}^n) + (X_{i-1}^n - \tilde{X}_{i-1})}{2\Delta r \Delta t} - \tilde{X}_i' \frac{u_{i+1} - u_{i-1}}{2\Delta t}. \tag{A.52}$$

We use this equation to solve for the non-advective part of the time evolution of the gradient of ρ, u, and e.

A.4.2 Advective part

After solving the non-advective part of the fluid equations, we solve the advection equation:

$$\frac{\partial X}{\partial t} + u \frac{\partial X}{\partial r} = 0. \tag{A.53}$$

This equation is solved by the cubic interpolating profile method, in which we keep track of the variables X and also of their gradients X'. We differentiate (A.53) to obtain:

$$\frac{\partial X'}{\partial t} + u \frac{\partial X'}{\partial r} = -\frac{\partial u}{\partial r} X' = \frac{\partial Y}{\partial t} + u \frac{\partial Y}{\partial r} = -\frac{\partial u}{\partial r} Y, \tag{A.54}$$

where we define $Y \equiv X'$. For timestep n, we know X and Y at all grid points. We can write the solution to Equation (A.53) at point i and time t as: $X_i(r_i, t) = X_i(r_i - u\Delta t, t - \Delta t)$: we advect the profiles from the previous timestep at time $t - \Delta t$ to the current time t. However, the exact solution is only given at time $t - \Delta t$ at the center of the gridpoints r_i. We don't know the profile within the cell, so we must interpolate and

we choose a third-order polynomial to do so. The polynomial takes the general form:

$$F_i(\tilde{r}) = a_i\tilde{r}^3 + b_i\tilde{r}^2 + c_i\tilde{r} + d_i \tag{A.55}$$

where $\tilde{r} = r - r_i$. For timestep $n+1$, we shift the profile:

$$X^{n+1} = F(r - u\Delta t), \text{and} \tag{A.56}$$

$$Y^{n+1} = \frac{d}{dr}F(r - u\Delta t). \tag{A.57}$$

$dF(r)/dr = 3ar^2 + 2br + c$. At $r = 0$, $dF/dr = c$, so $c_i = Y_i^n$. Furthermore, at $r = 0$, $F = d$, so $d_i = X_i^n$. Then, we also know F and dF/dr at r_{i+1}. $\tilde{r} = r_{i+1} - r_i = \Delta r$, which is the grid size. We can solve the following two equations for the constants:

$$F(r_{i+1}) = a_i\Delta r^3 + b_i\Delta r^2 + Y_i^n\Delta r + X_i^n = X_{i+1}^n, \text{and} \tag{A.58}$$

$$\frac{\partial F(r_{i+1})}{\partial r} = 3a_i\Delta r^2 + 2b_i\Delta r + Y_i^n = Y_{i+1}^n. \tag{A.59}$$

The constants are:

$$a_i = \frac{Y_{i+1}^n + Y_i^n}{\Delta r^2} - \frac{2(X_{i+1}^n - X_i^n)}{\Delta r^3}, \text{and} \tag{A.60}$$

$$b_i = \frac{3(X_{i+1}^n - X_i^n)}{\Delta r^2} - \frac{Y_{i+1}^n - 2Y_i^n}{\Delta r}. \tag{A.61}$$

A.5 Hydrodynamic equations

García-Muñoz [27] has the hydrodynamic equations in the following form:

$$\frac{\partial r^2\rho}{\partial t} + \frac{\partial}{\partial r}(r^2\rho u) = 0 \tag{A.62}$$

$$\frac{\partial(r^2\rho u)}{\partial t} + \frac{\partial}{\partial r}(r^2(\rho u^2 + P)) = r^2\rho f_{ext} + 2Pr \tag{A.63}$$

$$\frac{\partial(r^2\rho E)}{\partial t} + \frac{\partial}{\partial r}(r^2(\rho E + P)u) = \rho u r^2 f_{ext} + \rho r^2(\Gamma - \Lambda). \tag{A.64}$$

Here, the total energy density $E = 1/2u^2 + e$ is the sum of the kinetic energy density due to the bulk fluid motion and the internal energy density. For the purposes of our numerical method, we wish to write this energy equation as a conservation equation for the internal energy density.

$$\frac{\partial}{\partial t}\left(r^2\rho\left(\underline{\frac{u^2}{2}} + e\right)\right) + \frac{\partial}{\partial r}\left(r^2\left(\rho\left(\underline{\frac{u^2}{2}} + e\right) + \underline{\underline{P}}\right)u\right) = \rho u r^2 f_{ext} + \rho r^2(\Gamma - \Lambda). \tag{A.65}$$

Now we focus on the underlined terms relating to the kinetic energy density:

$$\begin{aligned}
\frac{\partial}{\partial t}\left(\frac{r^2\rho u^2}{2}\right) + \frac{\partial}{\partial r}\left(\frac{r^2\rho u^2 \cdot u}{2}\right) &= \frac{1}{2}\left[u\frac{\partial(ur^2\rho)}{\partial t} + u\frac{\partial(r^2\rho u^2)}{\partial r}\right] \\
&\quad + \frac{1}{2}\left[ur^2\rho\frac{\partial u}{\partial t} + ur^2\rho \cdot u\frac{\partial u}{\partial r}\right] \\
&= \frac{u}{2}\left[r^2\rho f_{ext} - r^2\frac{\partial P}{\partial r}\right] + \frac{\rho u r^2}{2}\left[f_{ext} - \frac{1}{\rho}\frac{\partial P}{\partial r}\right] \\
&= \rho u r^2 f_{ext} - u r^2\frac{\partial P}{\partial r}.
\end{aligned} \tag{A.66}$$

We can also rewrite the double-underlined term:

$$\frac{\partial(r^2 P u)}{\partial r} = P\frac{\partial(ur^2)}{\partial r} + ur^2\frac{\partial P}{\partial r}. \tag{A.67}$$

Thus, we are left with:

$$\frac{\partial(r^2\rho e)}{\partial t} + \frac{\partial}{\partial r}(r^2\rho e u) = -P\frac{\partial(ur^2)}{\partial r} + \rho r^2(\Gamma - \Lambda). \tag{A.68}$$

If we set $\partial(r^2\rho e)/\partial t = 0$, we recognize this as the steady-state energy equation, $\vec{\nabla} \cdot (\rho e \vec{u}) = -P\vec{\nabla} \cdot \vec{u} + \rho(\Gamma - \Lambda)$. Expanding the derivatives, we have:

$$e\left(\frac{\partial(r^2\rho)}{\partial t} + \frac{\partial}{\partial r}(r^2\rho u)\right) + r^2\rho\left(\frac{\partial e}{\partial t} + u\frac{\partial e}{\partial r}\right) = -P\frac{\partial(ur^2)}{\partial r} + \rho r^2(\Gamma - \Lambda). \tag{A.69}$$

We can use the mass conservation equation to eliminate the first two terms on the left, and we are left with:

$$\frac{\partial e}{\partial t} + u\frac{\partial e}{\partial r} = -\frac{P}{r^2\rho}\frac{\partial (ur^2)}{\partial r} + \Gamma - \Lambda. \tag{A.70}$$

From mass conservation, we recognize that $\partial(ur^2)/\partial r = -ur^2/\rho\partial\rho/\partial r$. In steady state, $\partial e/\partial t = 0$, so we have:

$$\rho u\frac{\partial e}{\partial r} = \frac{P}{\rho}u\frac{\partial \rho}{\partial r} + \rho(\Gamma - \Lambda), \tag{A.71}$$

which is a form of Equation (3) from Murray-Clay et al. [58].