



## **Master Thesis**

Nanna Bach-Møller

# Atmospheric Disequilibrium Model

Model Development for Exoplanet Atmosphere Simulations

Advisor: Uffe Gråe Jørgensen

Handed in: October 14, 2020

## Abstract

Detecting biosignatures in the spectra of exoplanet atmospheres might be one of our strongest tools in the search for life elsewhere in the universe. To date, we have only been able to do very few detailed spectroscopic observations of exoplanet atmospheres, but in the coming years, missions such as JWST, ELT and ARIEL have exactly this as one of their main goals. In order to be able to analyse the observations from these missions, we need to have an in-depth understanding of the atmospheres we are observing, and here atmosphere modelling plays a critical role.

Many atmosphere models calculate the atmospheric chemistry based on an assumption of chemical equilibrium. This can be problematic since many atmospheres are expected to be out of equilibrium, making them difficult to reproduce with equilibrium models. In addition to this disequilibrium has been suggested as a potential biosignature, and might be used as a means to study the physical structure of an atmosphere, solely based on its spectrum. It is therefore important to allow for chemical disequilibrium in atmosphere models, which is not currently the case for the models, MARCS and GGchem used by *StarPlan*.

The aim of this project has been to develop a chemistry model that can be coupled to MARCS, and will allow for chemical disequilibrium in the atmosphere. This new model estimates the concentration of atmospheric species based on the reaction rate of a specified set of reactions. From the reaction rates, the changes in concentration are expressed as a set of chemical ordinary differential equations (ODEs) and solved numerically. In this study I test the stability of two different methods for solving ODEs for a simple and more complex set of reactions respectively. I find that the explicit Forward Euler's method remains stable with relatively few iterations for the very simple set of reactions. For more complex reaction chains, with a larger diversity in the reaction rates, Forward Euler's method loses stability, and it might be necessary to use a semi-implicit or implicit method. The implicit Backward Euler's method remains stable for longer periods of time, and for larger time steps for both the simplified and the complex reaction set. However, problems with overflowing arise, that I have not been able to solve within the scope of the project. The model is intended to be further developed, and throughout this thesis I explain the choices I have made to make the model user friendly and easy to build upon, and I reflect on some of the potential improvements that might be implemented in the future.

## **Acknowledgement**

First and foremost, I would like to thank Uffe Gråe Jørgensen for his help and supervision on this project as well as on several other aspects of my scientific career. It is a true pleasure working with you, and I thank you so much for all our inspiring conversations and the many opportunities you have given me! I would also like to thank Kristian Holten Møller and Henrik Grum Kjærgaard for providing me help and guidance in the wondrous world of chemistry, and for not losing patience with me despite my complete lack of knowledge about chemical reactions. Then I would like to thank the Center for Star and Planet Formation for giving me access to their HPC network, and Rune Kildetoft and Troels Haugbølle for helping me in this regard. Finally, I would like to truly thank Frederik Pless Johannesson, Sidse Lærke Lolk and my family, for always being there for me and giving me all the help and mental support I could possibly wish for. Thank you.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Theory</b>	<b>2</b>
2.1	Atmospheres . . . . .	2
2.2	Physics . . . . .	4
2.3	Chemistry . . . . .	7
2.4	Solving chemical differential equations numerically . . . . .	10
2.5	Chemical disequilibrium . . . . .	14
<b>3</b>	<b>Models</b>	<b>16</b>
3.1	Atmosphere modelling . . . . .	16
3.2	Physics models and MARCS . . . . .	17
3.3	Chemistry models and GGchem . . . . .	19
3.4	Disequilibrium models . . . . .	20
<b>4</b>	<b>Methods</b>	<b>22</b>
4.1	Structure of disequilibrium model (DISEQ) . . . . .	22
4.2	Solving chemical ODEs . . . . .	26
4.3	Setting parameter values . . . . .	28
4.4	Ease of use and coding etiquette . . . . .	29
<b>5</b>	<b>Results</b>	<b>30</b>
5.1	Simplified reaction for oxidation of methane . . . . .	30
5.2	Reaction chain for oxidation of methane . . . . .	34
5.3	Comparing stability of ODE solvers . . . . .	38
5.4	How does it fit with actual methane production . . . . .	40
<b>6</b>	<b>Discussion and conclusions</b>	<b>41</b>
<b>7</b>	<b>Additional Projects</b>	<b>44</b>
7.1	Article on eccentricity-multiplicity correlation . . . . .	44
7.2	Observations at La Silla Observatory . . . . .	46
	<b>References</b>	<b>47</b>



<b>A</b>	<b>Disequilibrium model, DISEQ</b>	<b>51</b>
<b>B</b>	<b>Paper by Bach-Møller and Jørgensen, 2020</b>	<b>59</b>
<b>C</b>	<b>Paper by Kelley et al, 2020</b>	<b>70</b>

## 1 Introduction

One of the most promising aspects of exoplanet research in the near future is the observations of exoplanet atmospheres. Despite having observed exoplanets for almost 30 years, we have only just started getting detailed observations of their atmospheres. One of the most exciting discoveries in the field of exoplanets in recent years has been the detection of water in the atmosphere around the exoplanet K2-18b, which was reported by two individual studies based on observations from the Hubble Space Telescope done by Benneke et al. [2019] (also reported by Tsiaras et al. [2019]). We can expect to get many more of these observations in the coming years, when missions such as James Webb Space Telescope (JWST), the Extremely Large Telescope (ELT), and Atmospheric Remote-sensing Infrared Exoplanet Large-survey (ARIEL) begin their observations [Beichman et al., 2014, Greene et al., 2016, Tinetti et al., 2018, Udry et al., 2014]. The two space missions focus on transmission spectroscopy of exoplanet atmospheres, while ELT will be able to obtain direct images and spectra of exoplanets. All the facilities will have the search for biosignatures as one of their main goals.

In order to be able to properly analyse these observations, it is critical that we have a thorough understanding of exoplanet atmospheres, and an essential part of this understanding comes from atmosphere models. Through atmosphere modelling we can simulate the evolution and final state of an atmosphere. These simulations not only allow us to follow and study complex atmospheric processes, but also to recreate spectra that can then be used to analyse actual observations.

As we become able to study exoplanet atmospheres it will also be possible to search for atmospheric biosignatures on the distant planets. A very clear indication of the presence of life here on Earth, is the chemical disequilibrium found in our atmosphere (e.g. Seager and Dotson [2010], Schwieterman et al. [2018], Robinson and Reinhard [2019]). Atmospheric disequilibrium has long been suggested as a potential biosignature [Lovelock, 1965], and since then a number of studies have supported especially redox-disequilibrium as a promising biosignature (e.g. Krissansen-Totton et al. [2018], Wogan and Catling [2020]).

The atmosphere models that are currently used at *StarPlan* are the stellar atmosphere model MARCS [Gustafsson et al., 2008] in cooperation with the equilibrium chemistry model GGchem [Woitke et al., 2018]. As an equilibrium model, GGchem is not designed to simulate atmospheres in disequilibrium, and as such might not be suitable for the analysis of disequilibrium-biosignatures. The aim of this project is to develop an atmospheric chemistry model that allows for chemical disequilibrium. This model is meant to work in cooperation with the MARCS atmosphere model, by calculating the molecular concentration of the atmospheric species of each iteration of the MARCS model.

Throughout this report I will present the work I have done in developing my atmospheric disequilibrium model DISEQ. In Sec. 2 I will go through some of the basic theory regarding the physics and chemistry that takes place in an atmosphere, and how this is treated numerically. In Sec. 3 I will explain the basics of atmosphere modelling, and describe some of the models used by *StarPlan* (MARCS and GGchem). In Sec. 4 I will go into detail with the structure of my new disequilibrium

model, (DISEQ), as well as some of the choices I have made in the development of the model. In Sec. 5 I show the results of several test runs of the model. Finally in Sec. 6 I summarize my findings, and further reflect on the subject of atmospheric equilibrium models.

In addition to the main research of my Master's thesis project I have been working on two different projects, with the aim of expanding my knowledge in the field of exoplanet research. These projects are presented shortly in Sec. 7.

## 2 Theory

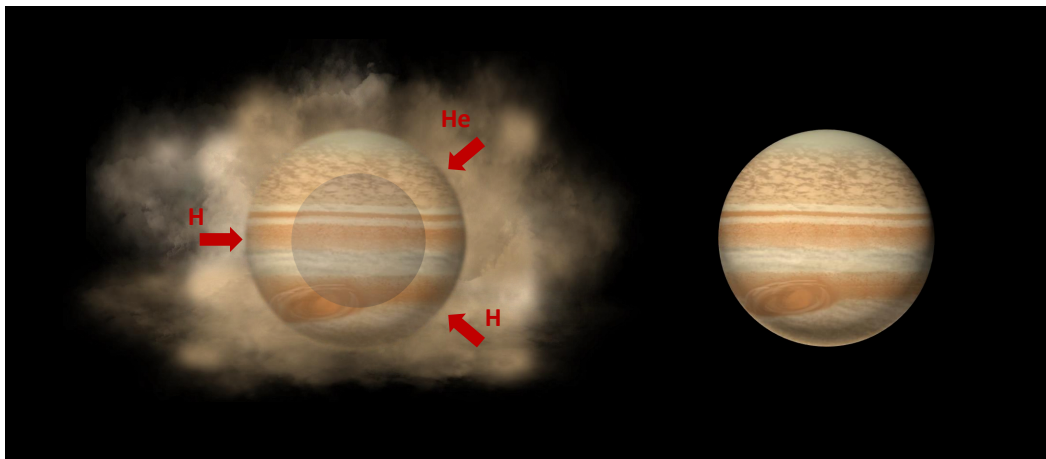
In this section I will go through some of the basic theory behind atmosphere models. I will briefly explain some of the defining characteristics of an atmosphere and the aspects of the physics and chemistry that is incorporated in most atmosphere models. The chemistry of atmosphere models is often treated through ordinary differential equations, and I will describe a few of the methods that can be used to solve these equations numerically. Finally I will present some of the research that has been made in chemical disequilibrium in atmospheres.

### 2.1 Atmospheres

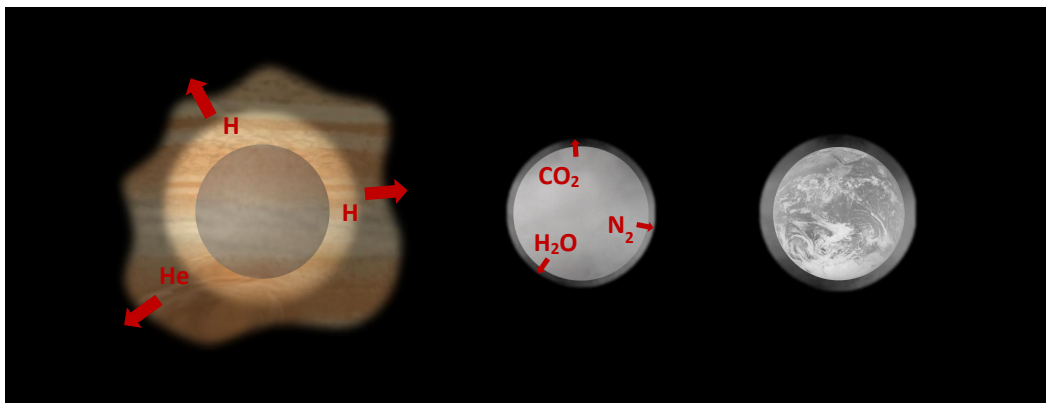
An atmosphere is an outer gaseous layer at the boundary between an object such as a planet, a moon, or a star, and its surroundings. On Earth, as well as on many other planets, the atmosphere only constitute a minimal fraction of the entire planet. But despite its modest proportions, understanding an atmosphere is essential for us to understand the object it belongs to for several reasons:

**The atmosphere itself** First of all, the atmosphere is in and of itself an interesting and very dynamical system. The entire energy balance of the planet is highly affected by the atmosphere, since the energy exchange between a planet and its surroundings happens through energy transfer in the atmosphere. In addition to this the planets main influx of energy from the host star is highly affected by the planet's albedo, which in turn depends on scattering and absorption due to different atmospheric molecular species and cloud formation [Fortney, 2018]. Atmospheric research involves studying how radiative transfer and convective dynamics drives wind systems and climate patterns. These mechanisms are complex in themselves, and are further complicated by processes such as day-and-night temperature differences caused by stellar radiation, the formation of clouds, atmospheric escape, and interactions with the surface. Understanding atmospheric processes is both important for understanding weather and climate patters here on Earth, but also for research on the other Solar System planets and exoplanets. When observing planets from afar our observations will often be highly affected, and in some cases completely obscured, by the atmosphere, and as such understanding the atmospheric system is critical for us to be able to interpret our observations.

**Interactions with the surface** The atmosphere itself can reveal a lot about the object it belongs to. The presence of e.g. oceans or active volcanoes on the surface will affect the chemical composition of the atmosphere, and these can therefore be indirectly studied through the atmosphere. The atmo-



(a) Accretion of gas from protoplanetary disk (left), to form the primary atmosphere of a gas giant (right).



(b) Escape of primary atmosphere (left), followed by outgassing from the surface (middle), to form secondary atmosphere of a terrestrial planet (right).

**Figure 1:** The formation of primary and secondary atmospheres.

sphere of Earth is also highly affected by its biosphere, and the presence of life on distant planets might similarly be detected through biosignatures in the atmosphere.

**Evolution of planet and planetary system** Atmospheres are not closed systems, and they will interact both with the surface of the planet and the surrounding space. For this reason an atmosphere can reveal details about the evolution of the object it belongs to as well as the system it is found in, depending on the type of atmosphere.

There are two major categories of atmospheres; *Primary* and *secondary* atmospheres [Seager and Deming, 2010]. **Primary atmospheres** are atmospheres that were captured or accreted from the protoplanetary disk during the late stages of planet formation, as illustrated in Fig. 1a. These atmospheres are often dominated by  $H$  and  $He$ , and are the type of atmospheres found among e.g. the gas giants of the Solar Systems. By studying the composition of these atmospheres we get an indication of the composition of the protoplanetary disk, and the early evolution of the planetary systems. **Secondary atmospheres** are atmospheres that have been outgassed from the interior of the planet as illustrated in Fig. 1b. These planets have often lost their primary atmospheres of  $H$  and  $He$  due to atmospheric escape, and have later formed new atmospheres through an outgassing from the surface

of heavier volatiles and in some cases silicates. These atmospheres can in some cases evolve further due to the re-absorption of some of the atmospheric molecules into the surface or potential oceans. The terrestrial planets of the Solar System all have secondary atmospheres, and by comparing the outgassing and re-absorption that has taken place on e.g. Earth, Mars and Venus we can increase our understand of planetary evolution [Fortney, 2018].

### 2.1.1 Observing exoplanet atmospheres

Atmospheres are found around all the planets of the Solar System, as well as some of the moons, and within the past 20 years we have also started to observe atmospheres around exoplanets. In very rare cases, with giant planets that are either very bright or located far from their host stars, exoplanet atmospheres can be observed through direct imaging, but more often the atmospheres are observed in relation to a transit. During transit observations the radiation from an exoplanet is observed together with its host star, and as such the radiation from the host star must be subtracted to evaluate the spectrum coming from the atmosphere. Two different spectra can be retrieved for transiting exoplanets [Kreidberg, 2017]: The **transmission spectrum** can be observed during the transit of the planet in front of the host star, as illustrated at position *a* in Fig. 2. As the planet transits, some of the radiation from the star moves through the planetary atmosphere, and absorption features from the atmosphere can be observed in the combined spectrum. The **flux spectrum** can be observed when the planet approaches the secondary transit, before it moves behind the star, as seen in position *b* in Fig. 2. At this point the planets day-side faces the Earth, and the thermal emission can be observed together with light from the star reflected from the planetary atmosphere.

Together these observations can give us information on the depth of the atmosphere, molecular composition, albedo, and temperature profile of the planet. In order to interpret the observation, however, we need an thorough understanding of atmospheric structures and chemistry - an understanding we largely get through atmosphere modelling. In the following I will go through some of the physical and chemical theory behind these models, and describe how it has been used in previous models as well as in the model that was developed as a part of this study.

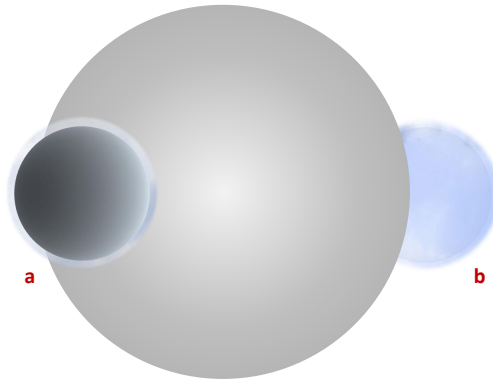
## 2.2 Physics

The physical aspect of atmosphere models mainly deals with energy balance and transfer within the atmosphere, in order to investigate the pressure-temperature profile.

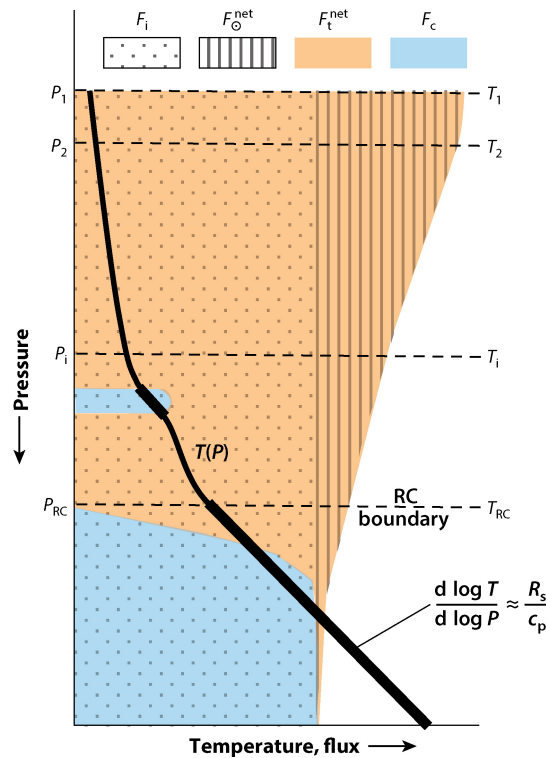
The dependence of pressure ( $P$ ) on altitude ( $z$ ) is described through the *equation of hydrostatic equilibrium* that balances the pressure and density ( $\rho$ ) of an atmosphere with gravity ( $g$ ):

$$\begin{aligned} \frac{dP}{dz} &= -\rho g \\ P(z) &= P_0 \cdot e^{-\frac{z-z_0}{H}} \end{aligned} \quad (1)$$

Where  $P_0$  and  $z_0$  are the initial pressure and height, and  $H$  is the pressure scale height. Based on this, the vertical position of an atmospheric layer can be directly expressed in terms of pressure rather than



**Figure 2:** A transmission spectrum can be observed during primary transit (a), when radiation from the star passes through the planet atmosphere. A flux spectrum can be observed prior to secondary transit (b), when thermal radiation from the planet can be observed together with stellar radiation reflected from the planetary atmosphere. Based on Fig. 2 from Robinson [2017] and Fig. 1 from Kreidberg [2017].



**Figure 3:** Pressure-temperature profile for a modelled atmosphere of a Jupiter-like planet [Marley and Robinson, 2015]. Black line indicates the temperature profile while the shaded areas indicate flux distribution and dominating energy transportation mechanism. The total flux of the atmosphere will be the sum of the internal energy flux ( $F_i$ ), and the irradiation from the host star ( $F_{\odot}$ ). The energy transport is dominated by thermal radiation ( $F_t$ ) at the upper layers, and convection ( $F_c$ ) below the radiative-convective (RC) boundary.

altitude, and as such the pressure-temperature profile will describe the vertical temperature profile in the atmosphere.

The temperature of a planetary atmosphere is determined by energy from the radiation from its host star ( $F_{\odot}$ ) as well as intrinsic sources of energy ( $F_i$ ), such as secular cooling of the interior or radiative decay. So the total energy flux ( $F_{total}$ ) of the atmosphere can be described as:

$$F_{total} = F_i + F_{\odot} \quad (2)$$

For gas giants and young terrestrial planets the energy from the intrinsic sources can be significant, but for older terrestrial planets, such as the Earth, the radiation from the host star dominates the energy balance. The thermal structure of a gas giant similar to Jupiter can be seen in Fig. 3, as modeled by the study of Marley and Robinson [2015]. Here we see the total temperature, expressed as the energy flux, along the primary axis, with a decreasing pressure along the secondary axis to indicate the altitude. As can be seen, a consistent energy flux due to intrinsic energy sources is found throughout the atmospheric layers (dotted), while an increasing flux due to radiation from the star (striped) can be seen in the upper layers.

The transfer of energy in the atmosphere occurs either through radiation or, if the temperature gradient becomes too large, through convection. Radiative energy transfer can be expressed through the *radiative transfer equation*:

$$\frac{dI_{\nu}}{dz} = \kappa_{\nu}\rho(S_{\nu} - I_{\nu}) \quad (3)$$

Where  $I_{\nu}$  is the specific intensity at frequency  $\nu$ , that travels along direction  $\mathbf{z}$ ,  $\kappa$  is the absorption coefficient,  $\rho$  is the density, and  $S_{\nu}$  is the source function. The radiative transfer equation in short terms states that the change in intensity along a distance,  $z$ , must be caused by either emission or absorption by the medium, and it will be highly dependent on the atmospheric composition.

Whether radiation or convection dominates in a specific atmospheric layer depends on the density and opacity of the gas [Fortney, 2018]. In Fig 3 it can be seen that the upper layers are dominated by radiative or thermal energy transfer ( $F_t$ , orange) while the dense, opaque layers in the lower atmospheres is dominated by convection ( $F_c$ , blue). In accordance with energy conservation, the energy received by the atmosphere from the interior and stellar irradiation is transferred to, and re-emitted from, the upper atmosphere, and the atmospheric system remains in thermal equilibrium [Fortney, 2018]. In this way each of the atmospheric layers will be in *radiative-convective equilibrium* such that:

$$F_t + F_c - (F_i + F_{\odot}) = 0 \quad (4)$$

The temperature profile of the atmosphere will depend largely on the dominating mechanism of energy transfer in the different layers. In convective layers the energy transfer is so efficient, that the temperature gradient will start following, so called, convective adiabats, which is indicated by thick black lines in Fig. 3. Atmosphere models dealing with radiative-convective energy transfer work by finding a temperature profile that will satisfy the transfer of energy, both from intrinsic sources and the host star, in each of the atmospheric layers. This is done by solving the three equations that were introduced in this sections: the *equation of hydrostatic equilibrium*, *radiative-convective equilibrium*,

and the *radiative transfer equation*.

## 2.3 Chemistry

In addition to the pressure and temperature, an atmosphere is also highly defined by its chemical composition. Due to the chemical interaction between the surface and the atmosphere of the planet, investigating the chemical composition of the atmosphere can tell us a lot about the surface and the general climate of the planet, both present and, to some degree, past. Even if one is only interested in the physical aspects of the atmosphere, the atmospheric composition is still very important. Since different atoms and molecules absorb radiation at different wavelengths, the temperature of the atmosphere will be highly dependent on the chemical composition. Furthermore, the formation of clouds will alter the albedo of the atmosphere further affecting the temperature, and since cloud formation is also highly dependent on the chemical composition, there is a high dependency of temperature on the chemistry. In the following section I will explain some of the basic theory behind atmospheric chemistry, and describe some of the methods used to describe chemical systems through differential equations.

### 2.3.1 Chemical reactions and equilibrium

Most chemical reactions will have the following structure:



Where the reactants  $A$  and  $B$  react to form the products  $C$  and  $D$ . The number ratio at which they react  $a$ ,  $b$ ,  $c$  and  $d$  are called the stoichiometric coefficients. The double arrow between the two sides of the reaction indicates that the reaction can run both ways. Which direction the reaction will run depends on the state of the system, such that the direction that will cause the greatest decrease of the total energy of the system will dominate. As the energy of the system is minimized, the reaction reaches an equilibrium at which the reaction will run equally fast in both directions.

The energy state of the system can be expressed through the free Gibbs energy, which is the energy available to do work. For a closed system the Gibbs free energy will be the difference between the enthalpy ( $H$ ) and the product of the entropy and the temperature ( $ST$ ):

$$G = H - ST \quad (6)$$

Where the enthalpy can be expressed through the internal energy of the system ( $E_{int}$ ), the pressure ( $P$ ), and the volume ( $V$ ).

$$G = E_{int} + PV - ST \quad (7)$$

In a closed system, where the temperature and volume remains constant, the change in Gibbs free



energy can be expressed as:

$$dG = dP \cdot V = dP \cdot \frac{nRT}{P} = nRT \cdot \ln\left(\frac{P}{P_0}\right) \quad (8)$$

Where  $n$  is the number of moles of gas and  $R$  is the ideal gas constant and  $P_0$  is the initial pressure. By defining the standard Gibbs free energy  $G^\ominus$  at the standard state of  $P_0 = 1 \text{ atm}$ , the Gibbs free energy can be expressed as:

$$G = G^\ominus + nRT \cdot \ln(P) \quad (9)$$

To look at the change in Gibbs free energy caused by the reaction in Eq. 5, one finds the difference between the energy of the reactants and products.

$$\Delta G = \sum(G_{\text{Products}}) - \sum(G_{\text{Reactants}}) = \Delta G^\ominus + RT \cdot \ln\left(\frac{P_C^c P_D^d}{P_A^a P_B^b}\right) \quad (10)$$

Where the pressure  $P_A$  are the partial pressures of species  $A$ , and similarly with the other participating species. The expression in the parenthesis is the reaction quotient,  $Q = \frac{P_C^c P_D^d}{P_A^a P_B^b}$ , and as the the Gibbs free energy of the system is minimized, and  $\Delta G \rightarrow 0$ , the reaction quotient will become the equilibrium constant  $K$  such that.

$$0 = \Delta G^\ominus + RT \cdot \ln(K) \quad \Rightarrow \quad (11a)$$

$$K = \exp\left(-\frac{\Delta G^\ominus}{RT}\right) = \frac{P_C^c P_D^d}{P_A^a P_B^b} \quad (11b)$$

Since the equilibrium constant, like the reaction quotient, states the ratio between the partial pressures of the products and reactants, Eq. 11b describes the dependence between the temperature, the change in standard Gibbs free energy and the partial pressures of the molecules when the system is at equilibrium. This correlation is used by many atmosphere models to estimate the relative abundances of the molecules in an atmosphere at equilibrium. I will elaborate on this in Sec. 3.3.

### 2.3.2 Reaction rates

Before the chemical system reaches equilibrium it should be described by its reaction rate rather than the equilibrium constant. The *reaction rate* is the rate at which a reaction takes place, and describes the change in concentration over time of any reactant. The reaction rate ( $r$ ) of the reaction in Eq. 5 is defined as [McNaught and Wilkinson]:

$$r = -\frac{1}{a} \frac{d[A]}{dt} = -\frac{1}{b} \frac{d[B]}{dt} = \frac{1}{c} \frac{d[C]}{dt} = \frac{1}{d} \frac{d[D]}{dt} \quad (12)$$

Where  $[A]$ ,  $[B]$ ,  $[C]$  and  $[D]$  are the concentrations of the reactants and products respectively, and  $a$ ,  $b$ ,  $c$  and  $d$  are their stoichiometric coefficients.

The reaction rate can be expressed based on the concentration of the reactants, depending on the

*reaction order* with respect to each of them.

$$r = k \cdot [A]^p [B]^q \quad (13)$$

Here  $[A]$  and  $[B]$  are again the concentrations of the reactants,  $k$  is the rate coefficient, and  $p$  and  $q$  are the partial reaction orders of each reactant. The reaction orders depend on the exact steps the reaction follows, and which temporary products, *intermediates*, occur along the way. For single-step reactions, or so-called *elementary* reactions, where no reaction intermediates have been detected, the reaction order for each of the reactants will be the same as their stoichiometric coefficients. In all other cases the reaction orders must be determined experimentally, by looking at how the reaction rate depends on the concentration of each reactant. The overall order of the reaction will be the sum of the partial reaction orders for each of the reactants. An elementary reaction between two molecules (a bimolecular reaction) will therefore be second order.

The rate coefficient,  $k$ , relates the reaction rate to the concentrations of the reactants, and is itself independent of these concentrations. The rate coefficient is determined by the number of collisions that leads to successful reactions, and it therefore depends on the temperature and pressure of the medium. The temperature dependence of the rate coefficient has been expressed by the *Arrhenius Equation*:

$$k = A \cdot \exp\left(-\frac{E_A}{RT}\right) \quad (14)$$

Where the activation energy,  $E_A$ , is the energy necessary for a reaction to take place, and  $R$  is the gas constant.  $A$  is often called the *pre-exponential factor*, or the *frequency factor*, and it is proportional to the frequency of collisions where the molecules are oriented in a such way that the reaction is possible. The activation energy and frequency factor are both determined experimentally by studying how the reaction rate, and hence the reaction coefficient, depends on temperature.

### 2.3.3 Chemical differential equations

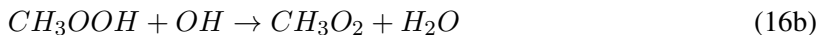
Both the physical and chemical processes in an atmosphere can often be expressed through differential equations. Looking at the expression for the reaction rate, the rate of change for any molecule due to a specific reaction can be expressed as an ordinary differential equation (ODE):

$$\begin{aligned} -\frac{1}{a} \frac{d[A]}{dt} &= r = k \cdot [A]^p [B]^q \\ \frac{d[A]}{dt} &= -k \cdot a \cdot [A]^p [B]^q \end{aligned} \quad (15)$$

Eq. 15 illustrates how the rate of change in the concentration of a molecule depends on the current concentration of said molecular species. The dynamics of the entire chemistry of the atmosphere, including all present molecules, can be expressed as a system of coupled ODEs, where the rate of change for each molecule will depend not only on the present concentration of the molecular species itself, but also on the concentration of other species it reacts with.

To illustrate a set of coupled ODEs, we can look at some of the reactions that participate in the

oxidation of methane in the atmosphere (a more complete reaction chain can be seen in Table. 2 in the results section, but will not be needed now):



Since these reaction are all elementary reactions (i.e. single-step reactions, with no known intermediates), the orders of the reactions are easily found from the stoichiometric coefficients of the reactants: All reactions are first order in relation to each of their reactants, and whereas 16a and 16b are both bimolecular elementary reactions with a total order of two, 16c is a unimolecular, first order reaction.

The rates of the three reactions will therefore be:

$$r_{CH_4+OH} = k_1 \cdot [CH_4]^1 [OH]^1 \quad (17a)$$

$$r_{CH_3OOH+OH} = k_2 \cdot [CH_3OOH]^1 [OH]^1 \quad (17b)$$

$$r_{CH_3OOH} = k_3 \cdot [CH_3OOH]^1 \quad (17c)$$

From the reaction rates we can write up the ODEs to describe the change in concentration for each of the molecules. The change in concentration for a specific species will be the difference between the production and loss of that species, calculated from the reaction rates of reactions where the species act as reactant and product respectively. Taking  $CH_4$  and  $OH$  as examples:

$$\frac{d[CH_4]}{dt} = -k_1 \cdot [CH_4][OH] \quad (18a)$$

$$\frac{d[OH]}{dt} = -k_1 \cdot [CH_4][OH] - k_2 \cdot [CH_3OOH][OH] + k_3 \cdot [CH_3OOH] \quad (18b)$$

Since  $OH$  participate in all three reactions, either as a reactant or product, the rate of all three reactions will affect the final change in concentration of  $OH$ . As such the final concentration of  $OH$  will also depend on the concentration of all the reactants, and the complete set of differential equations will be coupled to one another.

## 2.4 Solving chemical differential equations numerically

In order to study the evolution of the reactions shown in Eq. 16, one would need to solve the ODEs listed in Eq. 18 as well as similar equations for all other molecules participating in the reactions.

Many individual ODEs can easily be solved analytically by integrating both sides of the expression.

Looking at Eq. 18a as an example, the solution would be:

$$\begin{aligned} \int \frac{1}{[CH_4]} d[CH_4] &= - \int k_1 \cdot [OH] dt \quad \Rightarrow \\ \ln([CH_4]) &= -k_1 \cdot [OH]t + C \quad \Rightarrow \\ [CH_4]_t &= [CH_4]_0 \cdot e^{-k_1[OH]_t t} \end{aligned} \quad (19)$$

From this solution the concentration of  $CH_4$  at time  $t$  can be found from the rate coefficient  $k_1$ , the current concentration of  $OH$ , and the initial concentration of  $CH_4$ .

Not all ODEs can be solved analytically, and as the number of reactions and molecules in a model increase, the complexity of the set of coupled differential equations increases as well, and the system must be solved numerically.

Solving a problem numerically is a way to approximate the analytical solution often by iterating over smaller steps from an initial value [Jacobson, 2005]. There are five properties that must, especially, be taken into account when choosing a method to solve a chemical ODE for an atmosphere model [Cariolle et al., 2017, Jacobson, 2005]:

- **Stability:** The model must remain stable and produce reasonable results during the entire run.
- **Mass-conservation:** There should be no net production or loss of atoms and thereby mass.
- **Accuracy:** It is desirable for a model to be as accurate and close to a potential analytical solution as possible, even though it will always be an approximation.
- **Positivity:** The model should never produce concentrations below zero.
- **Speed:** It is desirable for a model to run quickly and with as little computation as possible, especially for models investigating complex processes over long periods of time.

Numerical methods can either be explicit, semiimplicit, or implicit, and any numerical solution will always offer a compromise between the properties mentioned above. In the following sections I will discuss a few different ODE solvers based on the properties above.

#### 2.4.1 Explicit methods:

The simplest methods, and often the ones that require the least computation, are the explicit methods. During each iteration explicit methods step closer to the true solution based on values, that are already known from previous iterations, and as such the computation lies mainly in the iterations themselves, often making explicit methods very quick.

**Forward Euler's method** One of the most basic numerical ODE solvers is the explicit Euler's method, also known as the *Forward Euler's method* (hereinafter referred to as FEM). Using a step size of  $h$ , FEM estimates the function value,  $N_t$ , of the present time step,  $t$ , based on the value at

the previous time step  $N_{t-h}$ . This method assumes that the function follows a linear trend within the step, and that the slope of the trend is the first derivative at the previous time step  $\frac{dN_{t-h}}{dt}$ , such that:

$$N_t = N_{t-h} + \frac{dN_{t-h}}{dt} \cdot h \quad (20)$$

Taking  $CH_4$  from the reactions in Eq. 16 as an example and substituting 18a into the equation, FEM solves the ODE as follows:

$$\begin{aligned} [CH_4]_t &= [CH_4]_{t-h} + \frac{[CH_4]_{t-h}}{dt} \cdot h \\ &= [CH_4]_{t-h} - k_1 \cdot [CH_4]_{t-h} [OH]_{t-h} \cdot h \end{aligned} \quad (21)$$

Where  $k_1$  is still the rate constant of the reaction in Eq. 16a and  $h$  is the time step. FEM is mass conserving, and for small enough time steps it is both highly stable and accurate, in accordance with the requirements on page 11. The weakness of this method, and other explicit methods, is that their stability is dependent on the size of the time steps. Explicit methods can only be used if the time step is significantly smaller than the smallest life time of the participating molecular species [Cariolle et al., 2017]. If the time step is too big, the model will become unstable, and since it has no demand of positivity, the concentrations can become negative. Most atmosphere models contain radicals with life times of a few milliseconds, and the demand for time steps lower than this value makes explicit methods unsuitable for atmospheric chemistry, due to the high number of required computations.

#### 2.4.2 Implicit and semiimplicit methods:

Where the explicit methods base their calculations on values known from previous time steps, implicit methods use values from the same time step as the value investigated, and semiimplicit methods use both previous and "current" values. Since the values needed are unknown at the beginning of the step, they must be found simultaneously with the desired solution. This often requires more computation, but in return the methods can have much larger time steps without losing stability, and some implicit methods are unconditionally stable for all step sizes.

**Backward Euler's method** An example of a semiimplicit method is the implicit Euler's method, or *Backward Euler's method* (hereinafter referred to as BEM). This method is similar to FEM, but instead of evaluating the first derivative solely from the previous time step,  $\frac{dN_{t-h}}{dt}$ , the derivative is evaluated for the present time step,  $t$ , in regard to the molecule that is being investigated:

$$N_t = N_{t-h} + \frac{dN_{t,t-h}}{dt} \cdot h \quad (22)$$

Again taking  $CH_4$  from the reactions in Eq. 16 as an example, we get:

$$\begin{aligned} [CH_4]_t &= [CH_4]_{t-h} - k_1 \cdot [CH_4]_t [OH]_{t-h} \cdot h && \Rightarrow \\ [CH_4]_t &= \frac{[CH_4]_{t-h}}{1 + k_1 \cdot [OH]_{t-h} \cdot h} \end{aligned} \quad (23)$$

Backward Euler's method has no time step instability, and it will not generate negative concentrations. The weakness of the Backward Euler's method is that it is not mass-conserving, due to the fact that each reaction will get different rates, depending on which participating species we are observing. The errors related to this uncertainty can accumulate significantly throughout the run.

**Gear's method** One of the few ODE solvers that meets all the criteria listed on page 11 is the semiimplicit *Gear's method*, especially when combined with sparse-matrix and computer optimization techniques, that increase the speed and allow the method to be used in 3D models [Jacobson, 2005]. The main strength of Gear's method is that it has been specifically designed for so-called *stiff* systems. Stiff systems are systems for which the time derivatives among the species differ greatly, and therefore systems where the species have very varying demands for the time steps [Jacobson, 2005]. For stiff systems the entire stability of the model is therefore dependent on the step size, whereas for other systems the step size will mostly just affect the accuracy of the model [Shampine and Gear, 1979]. As mentioned previously, the chemical species in chemistry models often have very different life times, and chemical systems are therefore, generally, stiff systems.

In order to account for this, Gear's method adapts the time step, so that the model can remain stable throughout the entire run. Gear's method is a backward differentiation formula, and as such it uses both values from the current and previous time steps. For a set of species  $\hat{N}_t$ , Gear's method solves the following ODE:

$$\hat{N}_t = \sum_{j=1}^s (a_{s,j} \hat{N}_{t-jh}) + h\beta_s \frac{d\hat{N}_t}{dt} \quad (24)$$

Where  $\alpha$  and  $\beta$  are scalar multipliers,  $s$  is the order of approximation, and  $h$  is the time step. The strength of Gear's method is that the size of the time step can be adapted for each iteration and atmospheric species, such that a minimum number of time steps are required. A weakness of the original method is that for each iteration of the Gear's method large matrices of partial derivatives must be solved using decomposition and backsubstitution, which requires a lot of computation. However, since many of the matrix entries will be zero, the matrix is *sparse*, and by using sparse-matrix techniques together with computer optimization techniques the computation required for each iteration can be drastically reduced. As a result, Gear's method is an effective and highly accurate solver of stiff chemical ODEs, but it is complicated to implement, and has not been in the scope of this project.

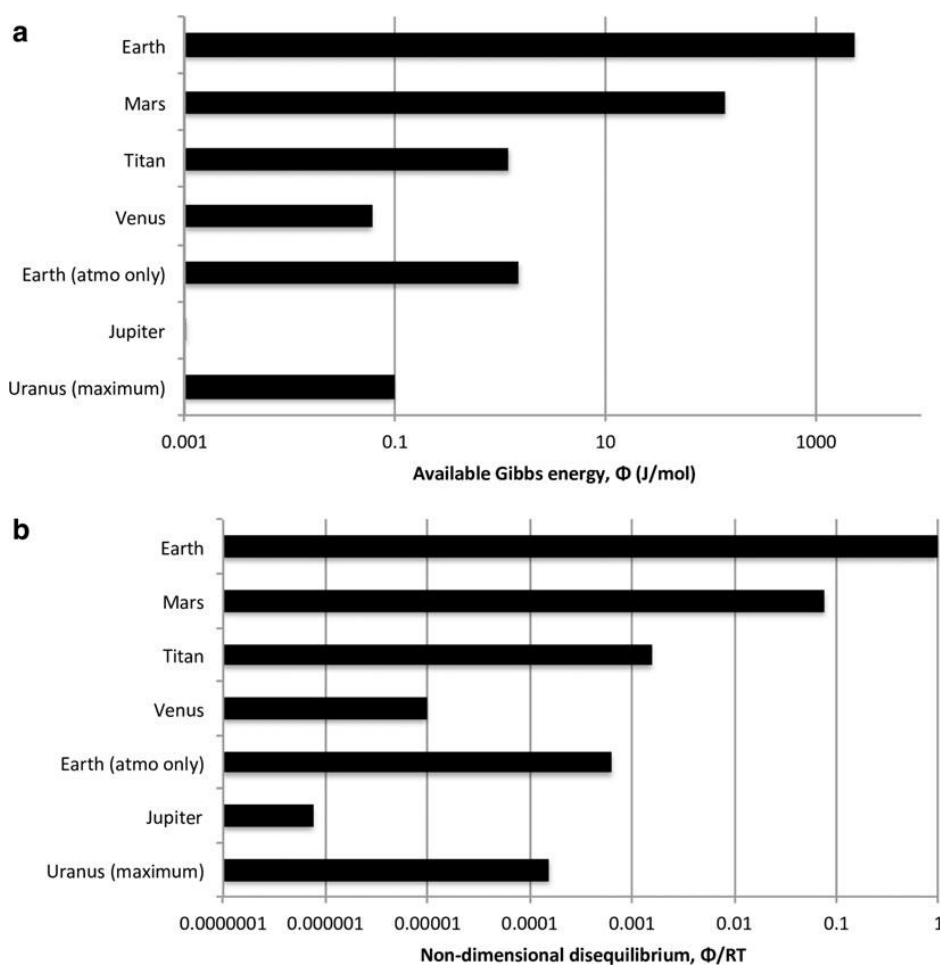
## 2.5 Chemical disequilibrium

As mentioned in Sec. 2.3.1 a closed system will reach chemical equilibrium when the free Gibbs energy is at a minimum, and as such the relative chemical abundances in an atmosphere can be found by minimizing the Gibbs free energy of the system. However, planetary atmospheres are not closed systems, and as such they are rarely in chemical equilibrium. For stellar atmospheres and high-temperature gas giants ( $T \geq 1200K$ , [Woitke et al. \[2018\]](#)) the temperatures are so high, and the chemical timescales correspondingly short, that the atmosphere in many cases can be assumed to be in local thermodynamic equilibrium. At lower temperatures and in secondary atmospheres of terrestrial planets the chemistry in the atmosphere becomes more complicated. Secondary atmospheres generally have a wider range of components, and lower temperatures allow for more complex molecules to be present and for clouds to be formed. These atmospheres will generally be driven out of thermodynamic equilibrium by mechanisms, such as UV photolysis and transport-induced quenching [[Moses, 2014](#)]. *UV photolysis* is especially dominant in the upper atmosphere and is caused by radiation from the host star. UV photolysis is one of the main causes of chemical disequilibrium in many atmospheres [[Venot and Agúndez, 2015](#)]. *Transport-induced quenching* takes place when the mass transport in the atmospheres due to e.g. convection and eddy diffusion, happens on shorter timescales than the chemical reactions [[Moses, 2014](#)]. In these situations the gas will be mixed before a chemical equilibrium can be reached, which will lock the concentrations at disequilibrium. Since the chemical reaction rate is highly dependent on temperature, transport-induced quenching is especially seen at lower temperatures.

Terrestrial atmospheres can be brought further into disequilibrium by their interactions with the surface, e.g. through chemical exchange with potential surface oceans and volcanic outgassing [[Wogan and Catling \[2020\]](#)]. These interactions act like sources or sinks for specific molecular species, which will prevent specific species from reaching a concentration appropriate for equilibrium.

When looking at the atmospheres found among the planets and moons of the Solar System, Earth has one of the largest thermodynamic disequilibria. A study by [Krissansen-Totton et al. \[2016\]](#) quantified the chemical thermodynamic equilibrium of Solar System atmospheres by comparing the Gibbs free energy of the observable atmosphere with the expected free energy of the same atmosphere in equilibrium. The results of their study are shown in Fig. 4, where it can be seen that Earth's atmosphere has one of the highest amounts of free Gibbs energy, only surpassed by Mars and possibly Titan. When the entire atmosphere-ocean system of Earth is included, rather than just the atmosphere, the disequilibrium of Earth is orders of magnitudes higher than any other object in the study. In order to account for how the irradiation from the Sun drives the thermodynamic disequilibrium, [Krissansen-Totton et al. \[2016\]](#) has shown both the available Gibbs energy (a), and the non-dimensional disequilibrium (b), where the latter has been calculated as the free Gibbs energy divided by the temperature and gas constant. The effect of using the non-dimensional disequilibrium can especially be seen when looking at Titan, that shows a significantly larger disequilibrium than Earth atmosphere in (b) but slightly smaller in (a).

In their study, [Krissansen-Totton et al. \[2016\]](#) found that the excess of available Gibbs energy on Earth was mainly due to a redox-disequilibrium between  $O_2$  and  $CH_4$ . Oxygen is a highly oxidizing



**Figure 4:** Fig. 8 from the study by Krissansen-Totton et al. [2016]. A comparison of the Gibbs free energy (a) and dimensionless free energy (b) among Solar System atmospheres. The dimensionless free energy is found as the Gibbs free energy divided by temperature and the gas constant. "Earth" refers to the entire atmosphere-ocean system of Earth, whereas "Earth (atmo only)" only includes the atmosphere.

species, whereas methane is reducing, and as such they should not be able to coexist over longer periods of time [Seager and Deming, 2010, Schwieterman et al., 2018]. The  $O_2 - CH_4$  redox couple has an expected lifetime of  $\sim 10yr$ , so in order for the concentration of both to remain high longer than this, there must be a constant production of especially  $CH_4$ . Krissansen-Totton et al. [2016] found abiotic causes for the disequilibrium found among other planets and moons, but the disequilibrium of Earth can not be explained solely by abiotic mechanisms, but must also have a biological source.

For this reason the disequilibrium caused by  $O_2 - CH_4$  and similar redox couples has been suggested as a potential biosignature. All life, as we know it, has a metabolism, and if the biomass or activity is great enough, the by-products of this metabolism can accumulate in the atmosphere causing chemical disequilibrium as observed on Earth. Especially redox chemistry is used in metabolism to produce energy, and as such redox disequilibrium is thought to be especially relevant as a biosignature, since abiotic processes rarely lead to redox disequilibrium [Seager and Deming, 2010].

A study by Wogan and Catling [2020] investigated the chemical disequilibrium through the evolu-



tion of Earth, and found that disequilibrium might be able to work both as a biosignature and an anti-biosignature, depending on the couples in disequilibrium. They concluded that the disequilibrium between  $H_2$ ,  $CO_2$ ,  $CO$  and  $H_2O$  found in the prebiotic Earth's atmosphere can work as a anti-biosignature, since these species can participate in redox reactions with very low activation energies, and as such easily could be utilized in metabolism if life is present on the planet. However, redox couples with very high activation energies, such as the disequilibrium between  $O_2$ ,  $N_2$ ,  $H_2O$  and  $CH_4$  found on modern Earth, are biosignatures, since they show that all the redox couples suitable for metabolism have been utilized.

A study by Krissansen-Totton et al. [2018] found the oxidation of methane as expressed through the reaction:



to be one of the main reasons for the chemical disequilibrium on Earth, especially during the Proterozoic era, where the concentration of methane in the atmosphere was higher. This reaction will be used later in this study to test the disequilibrium model and will be discussed further in Sec. 5.1. Krissansen-Totton et al. [2018] have similarly suggested the disequilibrium of the redox couple  $CH_4-CO_2$  as potential biosignatures. They found that in anoxic atmospheres, mixing ratios of more than  $10^{-2}$  for  $CH_4$  must be expected to be biological, since it is unlikely for abiotic mechanisms to maintain this concentration.

## 3 Models

### 3.1 Atmosphere modelling

An atmosphere model is meant to reproduce the evolution or final state of an atmosphere, based on a series of numerical calculations of the known dynamics, physics and chemistry that takes place in the atmosphere. The atmospheric processes are mostly expressed as ordinary and partial differential equations, and are solved numerically through approximations such as finite-difference [Jacobson, 2005] or the numerical methods described in Sec. 2.4. One of the greatest strengths of atmosphere models is the concept of *forward modelling*, where an atmosphere is simulated from the ground up solely based on some input parameters and computations of atmospheric processes [Fortney, 2018], which allows the models to be self-consistent. By forward modelling an atmosphere we are able to predict what observations of this atmosphere would look like and compare them to actual observations. This not only gives us insight into atmospheric processes, but can also be used to identify which processes takes place in observed exoplanet atmospheres, based on their spectrum. Constructing an atmosphere model is a constant compromise between the accuracy of the model and the required computer power, and as such the strength and limitations of a model often lies in the approximations used during the computations.

One of the defining qualities of an atmosphere model is its dimensions. Atmosphere models are generally divided into units or boxes, in which the conditions are assumed to be homogeneous, and

these boxes exchange particles with their surroundings. A model consisting of one box is called a zero-dimensional model or a box model. A line of vertically- or horizontally stacked boxes is a one-dimensional model, whereas two- and three-dimensional models have these lines of boxes stacked in one and two more directions respectively. Three-dimensional models that have both altitude-, longitude-, and latitude-dimensions can naturally reproduce more accurate depictions of dynamics and transport processes, as well as global variations from e.g. north to south or from the day to the night side. However, 3D-models are also very demanding computationally, and are often better suited for short term simulations of processes with timescales of months to years [Jacobson, 2005]. For the modelling of exoplanet atmospheres we are often interested in the long-term evolution of the atmosphere, and as such a 3D-model is rarely suitable. In addition to this, all observable light from an exoplanet will be combined into one pixel, and as such it is less important to model the global variation, when we wish to compare our models to observations. For these reasons exoplanet atmosphere models are often one-dimensional, and they model the variation among vertically stacked atmospheric boxes or layers.

Many atmosphere models are separated into two individual sections, one responsible for the physical mechanisms of the atmosphere and one responsible for the chemistry. In this section I will briefly describe the idea behind these two types of models, and especially focus on the stellar-atmosphere model MARCS and chemistry model GGchem, that are the main atmosphere models used at the *StarPlan* group at the University of Copenhagen.

### 3.2 Physics models and MARCS

Physics models mainly reproduces the pressure-temperature profile of the atmosphere, by modelling e.g. energy transfer and the absorption of radiation from the host star. In doing so they often account for e.g. radiative transfer, convection, condensation and cloud formation, surface fluxes, scattering, and absorption [Lincowski et al., 2018]. In order to model the pressure-temperature profile of an atmosphere, three different equations need to be solved: The radiative transfer equation, the equation of hydrostatic equilibrium, and the equation of radiative-convective equilibrium describing the conservation of energy [Seager and Deming, 2010].

The main atmosphere model used by StarPlan is a one-dimensional model called MARCS originally developed to model both the physics and chemistry of stellar atmospheres. MARCS is self-consistent and forward model the final equilibrium state of an atmosphere based on an initial guess of a pressure-temperature profile as well as elemental abundances. By being self-consistent the result of the MARCS model will always produce a reasonable solution in compliance with the known physical and chemical processes.

The first step of the MARCS model is to calculate the energy input of the atmosphere due to the absorption of radiation from the host star and the planetary surface/interior, based on the opacity of the involved species and the assumption of local thermodynamic equilibrium. MARCS will then model the effect on this energy input on the atmospheric state. As mentioned previously, the main physical aspects an atmosphere model generally have to deal with are radiative transfer, hydrostatic equi-

librium, and radiative- or radiative-convective equilibrium. The MARCS model deals with radiative transfer by solving the radiative transfer equation, Eq. 3, with the source function described as:

$$S_\nu = \frac{\kappa_\nu}{\kappa_\nu + \sigma_\nu} B_\nu(T) + \frac{\sigma_\nu}{\kappa_\nu + \sigma_\nu} J_\nu \quad (26)$$

Where  $\kappa_\nu$  describes the absorption coefficient,  $\sigma_\nu$  is the scattering coefficient,  $J_\nu$  is the mean intensity, and  $B_\nu$  is the Planck function describing black body radiation [Gustafsson et al., 2008].

Hydrostatic equilibrium is introduced by balancing the pressure with gravity (as explained in Eq. 1), where the total pressure, ( $\nabla P_{tot}$ ), will be the sum of the pressure caused by radiation, gas, and turbulence.

$$\begin{aligned} \nabla P_{tot} &= -\rho g \\ \nabla P_{tot} &= \nabla P_{rad} + \nabla P_{gas} + \nabla P_{tur} \end{aligned} \quad (27)$$

Finally the model is kept in radiative-convective equilibrium by balancing the incoming energy in each box or layer of the atmosphere with the energy transported through radiation or convection, as described in Eq. 4. MARCS introduces convection using mixing-length theory in which "parcels" of gas are moved as units before the contents mixes with the new surroundings. Several atmosphere models approximates convection through mixing-length theory (e.g. VPL Climate, Lincowski et al. [2018]), however, this approximation might introduce some uncertainty to the model, since it has been found that the mixing-length scheme causes a much more rapid variation of the convective velocity with depth, than what was found by more realistic simulations [Gustafsson et al., 2008]. The calculations are generally made under the assumptions that the atmosphere is in local thermodynamic equilibrium, and that the layers are plane-parallel or that there is spherical symmetry, depending on the depth of the atmosphere in relation to the size of the planet. These approximations poses some uncertainty to the MARCS model, but the main limitation of the model is that it is one-dimensional, and that it has difficulties converging for low temperature atmospheres. It is yet uncertain why the model does not converge at lower temperatures, but it might be related to the fact that the electron pressure is used to adjust the structure of the iterations, and that for lower temperatures the electron pressure could get low enough to cause numerical issues. The main strengths of the model is the fact that it is self-consistent. Non-self-consistent models often base their calculations on fitting regression, where the physical parameters are fitted to existing observations. These models will not necessarily provide credible physical results, they might not be reproducible, and for this reason fitting regression models are less trustworthy than self-consistent models. Another strength of MARCS is that it has already been tested against the observations of numerous stars and brown dwarfs, and the paper describing the current version of the model [Gustafsson et al., 2008] has been cited by more than 1500 studies. The fact that MARCS can model both stellar and planetary atmospheres might allow the model to recreate entire self-consistent systems, of exoplanets and their host stars. This is currently being done at *StarPlan*, where they are attempting to model the planetary- and stellar atmosphere using the same code, such that the irradiation from the host star into the planetary atmosphere can be modelled as accurately as possible.

In addition to the atmospheric processes treated by MARCS itself, the model can also be coupled

to several other individual models. Doing this, MARCS can deal with diffusion, dust- and cloud formation, through a coupling with the model DRIFT, and the processing of chemical reactions can be improved through the coupling with one of several different chemistry models. One of these chemistry models is the equilibrium model GGchem, that I will introduce next.

### 3.3 Chemistry models and GGchem

The chemistry taking place within each of the atmospheric boxes or layers is often modelled by individual chemistry models. These models must account for all chemical reactions as well as the photochemistry that takes place in the upper atmosphere due to the irradiation from the host star, and in some cases aerosol condensation and cloud formation. Some models also allow for atmospheric escape, an exchange of species between surface and atmosphere, and the chemistry of oceans in their models [Wogan and Catling, 2020]. The chemical reactions of the atmosphere will always strive to bring the atmosphere into chemical equilibrium, and many models are based on the assumption that this equilibrium has already occurred, which allows the model to estimate the concentration of each atmospheric species based on equilibrium constants.

The chemistry model GGchem, which can be coupled to MARCS, is one of these equilibrium models. For each of the atmospheric layers MARCS calls GGchem to find the concentration of the atmospheric species at equilibrium by minimizing the Gibbs free energy. As described in Sec. 2.3.1, a closed system will be at equilibrium when the Gibbs free energy is at a minimum, such that  $\Delta G = 0$ . At this point, the equilibrium constant can be expressed as:

$$K = \exp\left(-\frac{\Delta G^\ominus}{RT}\right) \quad (11b, \text{copy})$$

GGchem solves this equation for each of the chemical reactions, based on the pressure and temperature of that specific atmospheric layer, by calculating the standardized Gibb's free energy for each reaction. This is done by also accounting for element and charge conservation, and together they form a system of non-linear equations. For higher temperatures ( $T > 1000K$ ) these systems can be solved numerically using the Newton-Raphson method. However, as the temperature is lowered to adapt the model to exoplanet atmospheres, the chemistry becomes more complex, as molecules can become larger and the particles can begin to condensate. To account for this complexity [Woitke et al. \[2018\]](#) have introduced several pre-iterations to find a better initial guess before the use of Newton-Raphson, making the model stable down to 100K. GGchem now has more than 500 molecules and excels compared to many other chemistry models by maintaining stability even at lower temperatures.

The main limitation of the GGchem model is the assumption of chemical equilibrium. As mentioned in Sec. 2.5, most planetary atmospheres, especially those of lower temperatures, have been brought out of chemical equilibrium [Moses, 2014, Venot and Agúndez, 2015, Wogan and Catling, 2020], by e.g. photolysis in the upper atmospheres or interactions with the surface. In addition to this, disequilibrium might be an important biosignature, and as such it might be critical for us to be able to understand and model atmospheres in chemical disequilibrium.

### 3.4 Disequilibrium models

When modelling atmospheres out of chemical equilibrium it is not enough to look at the equilibrium constants, the model also has to account for the processes bringing the atmosphere out of equilibrium and allow the final state to be in disequilibrium.

One way to allow disequilibrium in a model atmosphere is by introducing "quenching" into an equilibrium chemistry model. As described in Sec. 2.5, *transport induced quenching* takes place when an atmospheric species is transported to an area faster than it can be removed through chemical reactions, in other words when the transportation timescale is shorter than the chemical timescale. This will cause the concentration of the species to be locked (or quenched) at a non-equilibrium value. This *quench approximation* can be applied by finding the *quench point*, where the transportation timescale becomes shorter than the chemical timescale for specific molecules, and then locking the concentration of these molecules at this point (as reviewed by Marley and Robinson [2015] and implemented by Zahnle and Marley [2014], Miles et al. [2020], Fortney et al. [2020]). For all other species and for layers where the chemical is shorter than the transportation timescale, the atmosphere is assumed to be in chemical equilibrium, and the concentrations can be found e.g. by minimizing Gibbs free energy. The main strength of the quench approximation is that it can be introduced to already existing chemical equilibrium models, without having to change the model drastically. The main weakness is the assumption that transport induced quenching should be the main cause of disequilibrium in the atmosphere. In addition to this the models are not necessarily self-consistent, since the disequilibrium is often not accounted for when making the pressure-temperature profile, as argued by Fortney et al. [2020].

A different approach to modelling an atmosphere in disequilibrium is to calculate the concentrations of the atmospheric species using the reaction rates rather than the equilibrium constants, as described in Sec. 2.3.2. The reaction rate for many atmospheric reactions have been determined experimentally for specific temperature ranges, and the results are listed on databases such as NIST [2020]. Many atmospheric chemistry models calculate the concentrations based on these rates, by passing the rates as a function of temperature from the databases as an input to the model (e.g. Zahnle et al. [2009], Zahnle and Marley [2014], Arney et al. [2016], Lincowski et al. [2018]). Sources of disequilibrium can then be introduced as e.g. externally induced reactions, such as photolysis in the upper atmosphere, or fixed abundances of species at the surface, and the evolution of the atmosphere can be followed through time. An example of this, can be seen in a study by Arney et al. [2016], who set a fixed abundance of  $CH_4$  at the surface layer (corresponding to a constant release of  $CH_4$  at the surface), which caused disequilibrium throughout the atmospheric layers, as  $CH_4$  was transported upward from the surface. The strength of using the reaction rates is that it will produce self-consistent results, since it is a very direct way of forward modelling the chemistry. In addition to this, a time dependency is introduced through the reaction rate, which allows the model to follow the evolution of the atmosphere through time, and not just model the final equilibrium state. A limitation of this method is, that the reaction rate and all values related to the rate must be determined experimentally. As explained in Sec. 2.3.2 the reaction rate is found based on a temperature dependent rate constant, and as such the constant for each reaction must either be determined experimentally at all

relevant temperatures, or the temperature dependency must be known. These values have not been determined for all reactions that take place in an atmosphere, and as such, estimations must be made [Zahnle et al., 2009].

Most of the studies mentioned above have calculated their reaction rates based on temperature-dependencies that were provided on NIST [2020]. In the development of my own model I have taken a similar approach, as I will explain in Sec. 4, except that I have calculated the temperature dependency myself. The calculations are done based on the activation energy and frequency factor for each reaction using Arrhenius equation, as explained in Sec. 2.3.2. I believe that this provides more flexibility to the model, and allows it to be easily updated, as future experiments will produce more exact values for the activation energies and frequency factors.

## 4 Methods

The main purpose of this study is to build a chemical model that allows the atmosphere to be in chemical disequilibrium, and to test this model on different cases.

In order to allow for disequilibrium, the change in concentration for the atmospheric species is found based on the reaction rate, rather than the equilibrium constants, based on the same arguments as presented in 3.4. In this section I will describe the overall structure of `DISEQ`, how the reaction rates are found using the Arrhenius equation, and how the chemical ordinary differential equations are solved. I also explain some of the choices I have made in regard to layout, methods, user-friendliness, and parameter values.

### 4.1 Structure of disequilibrium model (`DISEQ`)

The model currently consists of a single file with scripts for the main program and subroutines. This file also contains a section specifying the species, reactions and parameters the user wish to use for the model. The entire model file has been attached as appendix A. The only other file needed for the model is a simple input file containing all the initial concentrations of the participating species. The model is written in Fortran in order to stay consistent with the other atmosphere models used at *StarPlan* (`MARCS` and `GGchem`), and to make it easier for a potential coupling to `MARCS`.

The main file has the following structure:

- Text introduction
- `parameters` module
- Main program
- Subroutines; `calc_conc_euler` and `calc_rates`

After a text introduction explaining the structure of the script, the code is initiated by a module named `Parameters` where all universal parameters used in both the main program and subroutines are introduced and defined. The `Parameters` module is followed by the main program, that performs the computations of the model by calling different subroutines.

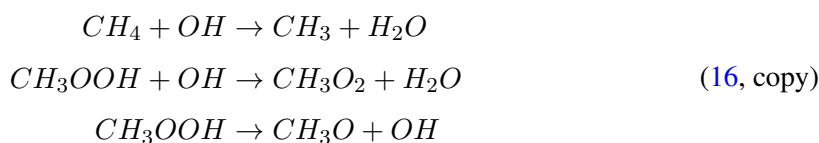
#### 4.1.1 Main program

**Reactions:** The first section of the main program lists the different species and reactions included in the model. In this section the user can add reactions to the model by providing the stoichiometric coefficients and exponents of each of the participating species, as well as frequency factor and activation energy for the reaction.

A *reaction matrix* is formed to indicate the participation of each species in each of the reactions. The entries of the matrix are given as the stoichiometric coefficients of the species in each reaction, with



the coefficient of the reactants being negative. As such the reaction matrix indicates the production and loss of each species in the reactions. Taking the reactions from 16 on page 10 as an example:



The reaction matrix would be:

$$\begin{array}{l} 16a \\ 16b \\ 16c \end{array} \begin{pmatrix} H_2O & CH_4 & CH_3 & OH & CH_3O & CH_3O_2 & CH_3COOH \\ 1 & -1 & 1 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & 1 & 0 & -1 \end{pmatrix} \quad (28)$$

In addition to the reaction matrix, this section also lists in the frequency factor, activation energy, and exponents for each of the reactions, all values that are needed to calculate the reaction rates using Arrhenius equation, as explained in Sec. 2.3.2. For elementary reactions the exponents equal the stoichiometric coefficients, and are set automatically by the model, but for reactions that are not elementary the exponents must be specified manually by the user.

Both the forward and the reverse reactions must be included in the model. In the current version of the model the reaction matrix and exponents of the reverse reactions are set automatically for all elementary reactions, whereas non-elementary reactions must be set manually. The frequency factor and activation energy must be set individually for each direction of the reaction. Currently the reactions are listed one at a time, in order to make it easy for users to add or remove reactions to suit their investigations. Taking reaction 16a as an example the code would look as follows:

---

```
! REACTION 1 and 2: CH4 + OH <-> CH3 + H2O
reactions(:,1) = (/&
  ! H2O CH4 CH3 OH CH2O CH3O2 CH3OOH
  & 1 , -1 , 1 , -1 , 0 , 0 , 0 , 0 , 0 &
  & /)
reactions(:,2) = -reactions(:,1)
! Both reactions are first order in regard to all reactants:
where (reactions(:,[1,2])<0) exponents(:,[1,2]) = 1 ! Setting exponents
act_energy(1) = 10.24 ! kJ/mol
fre_factor(1) = 4.16E-13 ! cm^3/(molecules s)
act_energy(2) = 6.22E01 ! kJ/mol
fre_factor(2) = 1.20E-14 ! cm^3/(molecules s)
```

---

Each participating reaction will have a section as the one above with the indices (1 and 2) changing consecutively as more reactions are added. As can be seen from the code above, the coefficients for each of the species are listed in the reaction matrix, and since the reaction is a elementary reaction, the exponents for both directions are easily found from the reaction matrix. Comments are added to make the code easier to follow, as will be further explained in Sec. 4.4



**Read Input Values** The next section reads in the input variables, and has been intended to read concentrations and temperature from e.g. MARCS output files. Currently it reads a simple input file with initial concentrations given in units of *molecules cm<sup>-3</sup>*, and store these concentrations as the variable `concentrations`. During a coupling with MARCS also the temperature of the specific atmospheric layer must be passed to DISEQ, but for the isolated test runs performed in this project the temperature is set at 300 K.

**Calculate new concentrations** The final part of the main program uses the reaction matrix and input values to calculate the new concentrations. This is done by solving the ordinary differential equation that describes the dependence between the reaction rates and the current concentrations of the reactants, as explained in Sec. 2.3.3. Solving the ODE is done using subroutines in the model, such that ODE solver can rather easily be changed, without having to adapt the rest of the code. The basic version of the model calls the subroutine `calc_conc_euler` that calculates the new concentrations using the forward Euler's method (FEM), but a subroutine for the Backward Euler's method (BEM) has also been made, and will be explained later.

#### 4.1.2 Subroutines

The main program is followed by several subroutines, and the transition is clearly marked with a heading. I have chosen to use subroutines for all major calculations that are replaceable, such that it is easy for the user to adjust or replace the calculations. Currently there are subroutines for the ODE solvers FEM and BEM, and for the calculation of the reaction rates using Arrhenius equation.

**calc\_conc\_euler** The Forward Euler subroutine is given the time and time step for the run, the reaction matrix and the current concentrations for all participating species, in addition to the parameters needed to calculate the reaction rate using Arrhenius equation.

The model loops over the different time steps, and for each it solves the differential equation using FEM as explained in Eq. 20 in Sec. 2.4.

$$N_{i,t} = N_{i,t-h} + \frac{dN_{i,t-h}}{dt} \cdot h \quad (20, \text{copy})$$

Where  $N_{i,t}$  and  $N_{i,t-h}$  are the concentration of species  $i$ , at time step  $t$  and  $(t - h)$  respectively,  $h$  is the size of the time steps, and  $\frac{dN_{i,t-h}}{dt}$  is the rate of change in concentration of species  $i$  at  $(t - h)$ . In order to find the change in concentration of species  $i$ , the reaction rates of the reactions it participates in must be calculated, which is done using the subroutine `calc_rates`, as will be explained further down. As described in Sec. 2.4 FEM finds the total change in concentration of a species as the sum of the rates of the reactions the species participates in, multiplied by the stoichiometric coefficient of the species in those reactions. Since the stoichiometric coefficients are listed in the reaction matrix,

( $R$ ), the total change in concentration per time can be found as:

$$\frac{dN_i}{dt} = \sum_j R_{i,j} \cdot r_j \quad (29)$$

Where  $R_{i,j}$  is the reaction matrix that indicates the stoichiometric coefficient of species  $i$  in reaction  $j$ , and  $r_j$  is the rate of reaction  $j$ . The change in concentration is multiplied by the size of the time step and added to the concentration from the previous time step.

At each time step, new concentrations are found, before the model moves on to a new time step. The iteration is continued until the desired time period is reached. The code for this part of the model can be seen below.

---

```
do ii_time = 0, nint(time/time_step)
  ! Calculate rate for all reactions
  call calc_rates(concentrations, fre_factor, act_energy, exponents, reactions, rate
    , k_const)

  ! Calculate change in concentration for each species
  do ii_spec = 1, n_species
    concentrations_change(ii_spec) = sum(reactions(ii_spec, :)*rate)
  end do

  ! Calculate new concentrations
  concentrations = concentrations + concentrations_change * ii_time * time_step
    * s_per_year
end do
```

---

**calc\_rates** As a part of the forward Euler's model shown above the reaction rates must be calculated, using the subroutine `calc_rates`. This subroutine calculates the rates based on Arrhenius equation using the frequency factors, activation energies, and exponents (indicating the order of the reactions in regard to each reactant), as well as the temperature of the atmospheric layer being studied. Rewriting Eq. 13 and 14 to look at the rate,  $r$  of reaction  $j$ , we get:

$$r_j = k_j \cdot \prod_{i \in reac} N_i^{p_i} \quad (13, \text{rewritten})$$

$$k_j = A_j \cdot \exp\left(-\frac{E_{A,j}}{RT}\right) \quad (14, \text{rewritten})$$

Where  $k_j$  is the rate constant,  $N_i$  the current concentration of species  $i$ ,  $reac$  is the set of indices for reactants in reaction  $j$ ,  $p_i$  is the exponent or order of the reaction in regard to the reactant,  $A_j$  and  $E_{A,j}$  are the frequency factor and activation energy for reaction  $j$ , and  $R$  and  $T$  are the gas constant and temperature.

The rate constants are easily found for all reactions from the listed constants and the temperature. The model then finds the rate of each reaction, by looping over each species in the reaction matrix, to see if the species acts as a reactant, i.e. has a negative entry in the reaction matrix. If that is the case, the index of the species,  $i$ , is a part of the reactant set,  $reac$ , indicated in Eq. (13, rewritten) and the

concentration of the species participates in the product. The code for the subroutine, `calc_rates` looks as follows:

---

```

! CALCULATE RATE CONSTANTS, k_const
R_ideal_gas = 8.31                               ! Ideal gas constant, in J/(mol K)
R_ideal_gas = R_ideal_gas/(6.022E23)            ! Converting to J/(molecules K)
k_const = fre_factor*exp(-act_energy/(R_ideal_gas*temp_layer)) ! Rate constant

! CALCULATE RATE FOR EACH REACTION, rate
do ii_reac = 1,n_reactions
  conc_product = 1

  ! Find which reactants participate in reaction
  ! If entry in reaction matrix is negative, species is a reactant
do ii_spec = 1,n_species
  if (reactions(ii_spec,ii_reac) < 0) then
    ! Calculate product of concentrations, with exponents
    conc_product = conc_product*concentrations(ii_spec)**exponents(ii_spec,
      ii_reac)
  end if
end do

  ! Calculate rate from rate constant and concentration product
  rate(ii_reac) = k_const(ii_reac)*conc_product
end do

```

---

The `calc_rates` subroutine is kept as a individual subroutine, separated from `calc_conc_euler`, in order to be able to change to a different ODE-solver, while still being able to use Arrhenius equation to calculate the reaction rates through `calc_rates`.

## 4.2 Solving chemical ODEs

As mentioned in Sec. 2.4 the stability of a chemical model will depend highly on which numerical method is used to solve the chemical ODEs, and the size of the time step. In order to test the stability of the model, I have chosen to implement two different methods for solving ODEs; the explicit FEM and the implicit BEM. I have already explained the subroutine implementing the FEM. This method requires very little computation for each of the iterations, but as an explicit method the stability of the model requires that the time step is significantly shorter than shortest life time among the atmospheric species. When the time step of FEM becomes too big, the model will loose its positivity, which can cause some issues as I will show you in the results section, Sec. 5.

### 4.2.1 Implementing Backward Euler's method, `calc_conc_euler_back`

The BEM, on the other hand, has guaranteed positivity, and as such the stability of the model is less dependent of the size of the time step. In return, BEM is a bit more complex and it requires more

computation for each of the iterations.

As shown in Sec. 2.4 BEM solves the following ODE:

$$N_t = N_{t-h} + \frac{dN_{t,t-h}}{dt} \cdot h \quad (22, \text{copy})$$

Where the derivative of the concentration is evaluated at the current time step,  $t$ , for the species that is currently being studied, and at the previous time step,  $t - h$ , for all other species. Since the derivative depends on unknown values for time  $t$ , I will rewrite the expression to isolate all concentrations evaluated at  $t$ . The derivative can be expressed in terms of the loss rate ( $L$ ) and the production rate ( $P$ ).

$$N_t = N_{t-h} + (P_{t-h} - L_{t,t-h}) \cdot h \quad (30)$$

The production of a species,  $i$ , will depend on the rate of the reactions where the species,  $i$ , acts as a product. The loss will depend on the rates of the reactions where the species,  $i$ , act as a reactant. Since the reaction rates only depend on the concentrations of the reactants (see Eq. 13), the concentration of species,  $i$ , will only influence the loss rate. The production rate can therefore be evaluated at  $t - h$ , whereas the loss must be evaluated at  $t$ . For all elementary reactions, the loss rate of a reactant will be directly proportional to the reactant itself. For we can therefore define an implicit loss coefficient for species  $i$ ,  $IL_i$ , that is the loss rate divided by the concentration of species,  $i$ .

$$IL_{i,t-h} = \frac{L_{i,t,t-h}}{N_{i,t}} \quad (31)$$

Since all other species are evaluated at  $t - h$ , the implicit loss coefficient will also be evaluated at  $t - h$ . By introducing this to the ODE, we get:

$$\begin{aligned} N_{i,t} &= N_{i,t-h} + (P_{i,t-h} - IL_{i,t-h} \cdot N_{i,t}) \cdot h \\ N_{i,t} &= \frac{N_{i,t-h} + P_{i,t-h} \cdot h}{1 + IL_{i,t-h} \cdot h} \end{aligned} \quad (32)$$

We can now express the production rate and the implicit loss coefficient as:

$$\begin{aligned} P_{i,t-h} &= \sum_j R_{i,j}(+) \cdot r_{j,t-h} \\ IL_{i,t-h} &= \sum_j R_{i,j}(-) \cdot k_j \cdot \prod_{k \in \text{react}, \forall \neq i} N_k^{p_k} \end{aligned} \quad (33)$$

Where  $j$  indicates the reactions,  $R_{i,j}(\pm)$  is the value for reaction matrix for species  $i$  in reaction  $j$  only including the positive or negative entries (i.e. entries where the species is a product and reactant respectively),  $r_{j,t-h}$  is the reaction rates evaluated at  $t - h$ ,  $k_j$  is the rate constant, and the product for the implicit loss coefficient is over the concentrations of all reactants (*react*) for each reaction, not including species  $i$ .

The code for the BEM subroutine can be seen as a part of the complete model in Appendix A.

The user can easily change between the desired ODE solvers by changing which solver is called in the "calculate new concentrations" part of the main program. Choosing FEM and BEM respectively

would look as follows:

---

```
! Calculate new concentrations , call solver
  call calc_conc_euler(concentrations , reactions , fre_factor , &
    & act_energy , exponents , rate )
```

---

```
! Calculate new concentrations , call solver
  call calc_conc_euler_back(concentrations , reactions , fre_factor , &
    & act_energy , exponents , rate )
```

---

### 4.3 Setting parameter values

In order for the model to calculate the new molecular concentrations five input parameters are needed: 1) The temperature of the atmospheric layer. 2) The initial molecular concentrations. 3) The activation energy of each reaction. 4) The frequency factor of each reaction. 5) The exponents of each reactant, decided by the order of each reaction in regard to that reactant.

**Temperature** The temperature of the atmospheric layer is decided before each run of the model, and will remain constant throughout the run. When the model is coupled to the MARCS climate model, the temperature for each atmospheric layer will be found by the climate model and read as an input to the chemical disequilibrium model. For the test runs the temperature is set to 300K, since most of the other parameter values used in the test run have been decided experimentally at this temperature.

**Initial molecular concentrations** The initial concentration of each of the atmospheric species are used in the first iteration of the ODE solver. During a coupling with the MARCS model, these concentrations should be read as outputs from MARCS. For the test runs done in this project the concentrations are set to values similar to those found in the modern Earth atmosphere. I have currently not been able to find any universal database listing the average concentration of all atmospheric species in Earth's atmosphere, and as such the concentrations are found from a number of different studies. All concentrations used in the test runs can be seen in Table 3. When looking at the more complex reaction chains, e.g. in Sec. 5.2, short lived radicals such as  $OH$  and  $HO_2$  play important roles as catalysts. These species have very low concentrations and are so short lived that it is extremely difficult to measure their actual concentration in the atmosphere. For this reason most of the radicals used in the model have been given initial concentrations similar to that of  $OH$ , which should pose no great uncertainty to the model, since the radicals will quickly be used and reproduced. In general, it is in no way necessary to know the exact concentration of each species in order to use the model, and since the actual atmospheric concentrations varies significantly, we can still get a valuable insight into the atmospheric chemistry even from rough initial estimates.

**Activation energy, frequency factor and reaction order** In order to find the reaction rate using Arrhenius equation both the activation energy, frequency factor and the exponents indicating the

reaction order of the reactants for each of the reactions must be known. These values are usually determined experimentally at different temperatures, and results from many of the experiments have been listed on the database NIST [2020]. For most of the reactions used in this study, NIST [2020] has several different studies reporting their results regarding each specific reaction. In order to be consistent in my choice of which reported values to use in this study, I generally pick the first study on the list (preferably a review, and preferably of newer date), that report both a frequency factor and an activation energy within a wide temperature range (generally from  $\sim 300K$  to  $\sim 2500K$ ). In doing so, I hope to get more reliable results from recent studies that have gone through peer review, and values that are relevant within a temperature range suitable for a warm atmosphere. In cases where the reactions are not listed in the database, I estimate the values from similar reactions, as will be explained in the later sections.

#### 4.4 Ease of use and coding etiquette

In order for others to be able to use and continue to develop the model, user-friendliness and coding etiquette has been taken into consideration when building the model. The following points have been taken into account.

- **Text sections** that explain the contents and structure of the following code are found at the beginning of the entire model, and before each section of the main program and each subroutine, to make it easier for the user to follow and potentially alter the process.
- **Comments** are added throughout the code to clarify the purpose of each command and to specify e.g. units.
- **The names of parameters, subroutines and files are mostly self-explanatory**, and are further introduced through comments at their first appearance in the code, where also units are specified if relevant. The variable names are never shorter than three characters to make it easier to search for them. **Only letters and underscores are used in the names** to make it easier to separate from code, parameter values, and format specifications. Subroutines are named such that their purpose should be clear from the name, e.g. `calc_conc_euler` that calculates the new concentrations based on Euler's method.
- Fortran is not case sensitive, but **all code is written in lower case letters** to ease the reading and stay consistent. Comments are written in both capital and lower case letters, and headings are written in all capital letters, to make them easy to identify.
- **Indentations** are used for all nested loops to clearly confine them.
- **The unit numbers for in- and output files are listed** in the text introducing the model at the beginning.
- **Subroutines are used** for specific computational methods, e.g. ODE solvers, to make it easy to replace without having to do major changes to the main program.

## 5 Results

In the following section I will go through the results of a few test runs, for different sets of reactions and species, and for different settings in regard to the method of solving the chemical ODEs. For each test run I will describe the background for the run, the choice of specific parameter settings, and finally present the results.

### 5.1 Simplified reaction for oxidation of methane, Forward Euler's method

In order to test the model with a relatively simple, yet essential, reaction, I first look at the oxidation of methane. The oxidation, of methane is the main reason methane and oxygen is not found together in the atmosphere, and as such metabolism is needed in order for both to be present. The true reaction chain for the oxidation of methane is described in Sec. 5.2, but to test the reaction I will look at this simplified version:



As explained in Sec. 2.5 the oxidation of methane is thought to be one of the main sources of chemical disequilibrium on Earth, making it an interesting reaction to study.

To test the model with the reaction from Eq. 25, a REACTIONS section must be made in the model, as explained in Sec. 4.1. The reaction matrix is easily made from the stoichiometric coefficients, but because the reaction is not a primary reaction, the exponents can not be read from the coefficients. Finding the exponents is further complicated, since they can not be decided experimentally, due to the fact that the reaction is a simplified version of a larger reaction chain, and as such is not a "real" reaction at all. However, since the concentration of  $O_2$  is significantly higher than the concentration of  $CH_4$ , we can to a first approximation assume that it is constant, especially during the small periods of time relevant for each time step. This allows us to treat the reaction as a pseudo-first order reaction, and thereby set the exponents to one. The reaction matrix and exponents will look as follows:

---

```
! REACTION 1 and 2: CH4 + 2O2 <-> 2H2O + CO2
reactions(:,1) = (/&
    ! O2  H2O  CO2  CH4
    & -2 , 2 , 1 , -1 &
    & /)
reactions(:,2) = -reactions(:,1)
! Pseudo-first order, so assuming first order in regard to all reactants:
where (reactions(:,[1,2])<0) exponents(:,[1,2]) = 1 ! Setting exponents
```

---

As can be seen above, I am only looking at the reaction from Eq. 25, and have only included the four species that participates in the reaction, ignoring all others.

**Table 1:** Input values for the model test run using the simplified reaction for oxidation of methane shown in Eq. 25.

Reaction	A [ $cm^3 molecules^{-1} s^{-1}$ ]	$E_a$ [ $kJ/mol$ ]	$E_a$ [ $kJ/molecules$ ]
$CH_4 + 2O_2 \rightarrow 2H_2O + CO_2$	$\sim 10^{-25}$	$\sim 10$	$\sim 10^{-23}$
$2H_2O + CO_2 \rightarrow CH_4 + 2O_2$	$\sim 10^{-24}$	$\sim 60$	$\sim 10^{-22}$

(a) Values for the frequency factor and activation energy used to find the reaction rate for each of the two directions of the reaction.

Molecule	$O_2$	$H_2O$	$CO_2$	$CH_4$
Concentration [ $molecules\ cm^{-3}$ ]	$5.3 \cdot 10^{18}$	$1.0 \cdot 10^{18}$	$8.9 \cdot 10^{15}$	$4.3 \cdot 10^{13}$

(b) Initial concentration for each of the participating atmospheric species.

### 5.1.1 Choosing parameter values

Finding the activation energy and frequency factor needed to calculate the reaction rate is also complicated by the fact that the reaction is a simplification. Since the values cannot be decided experimentally, I estimate them based on a reaction similar to the one in question, namely:



If we assume that the change in concentration of  $CH_4$  is the same for the two reactions we know that:

$$\frac{d[CH_4]}{dt} = k_{O_2}[CH_4][O_2] = k_{OH}[CH_4][OH] \quad (35)$$

Where the rate constant for the reaction in 34 would be:

$$k_{OH} = A_{OH} \cdot \exp\left(\frac{-E_{a,OH}}{RT}\right) = 4.16 \cdot 10^{-13} cm^3 molecules^{-1} s^{-1} \cdot \exp\left(\frac{-10.24 kJ/mol}{RT}\right) \quad (36)$$

The frequency factor ( $A_{OH}$ ) and activation energy ( $E_{a,OH}$ ) used above were listed in the database NIST [2020]. As an approximation I assume that the activation energy is the same for the reactions from Eq. 25 and 34, ( $E_{a,OH} = E_{a,O_2}$ ) and by inserting the rate constants into Eq. 35, and isolating the frequency factor for Eq. 25, I get:

$$A_{O_2} = \frac{[OH]}{[O_2]} \cdot 4.16 \cdot 10^{-13} cm^3 molecules^{-1} s^{-1} = 7.85 \cdot 10^{-26} cm^3 molecules^{-1} s^{-1} \quad (37)$$

Where the above is based on the concentrations of  $OH$  and  $O_2$  respectively from Earth's atmosphere, as they are listed in Lodders et al. [1998]:

$$[OH] \approx 10^6 molecules\ cm^{-3} \quad [O_2] \approx 5.3 \cdot 10^{18} molecules\ cm^{-3} \quad (38)$$

I am doing the same for the reverse reaction, by using the values for the following reaction as a first approximation:





The values for the activation energies and frequency factors for both directions of the reaction in Eq. 25 can be seen in Table 1a. Since the concentration of the radical  $CH_3$  is very small, and its lifetime is in general short, it is difficult to measure its concentration in the atmosphere, and I have not been able to find a study reporting it. For the sake of this test, I will simply just assume that the concentration of  $CH_3$  is similar to that of the  $OH$  radical.

$$A_{CO_2} = \frac{[CH_3]}{[CO_2]} \cdot 1.20 \cdot 10^{-14} \text{ cm}^3 \text{ molecules}^{-1} \text{ s}^{-1} = 1.35 \cdot 10^{-24} \text{ cm}^3 \text{ molecules}^{-1} \text{ s}^{-1} \quad (40)$$

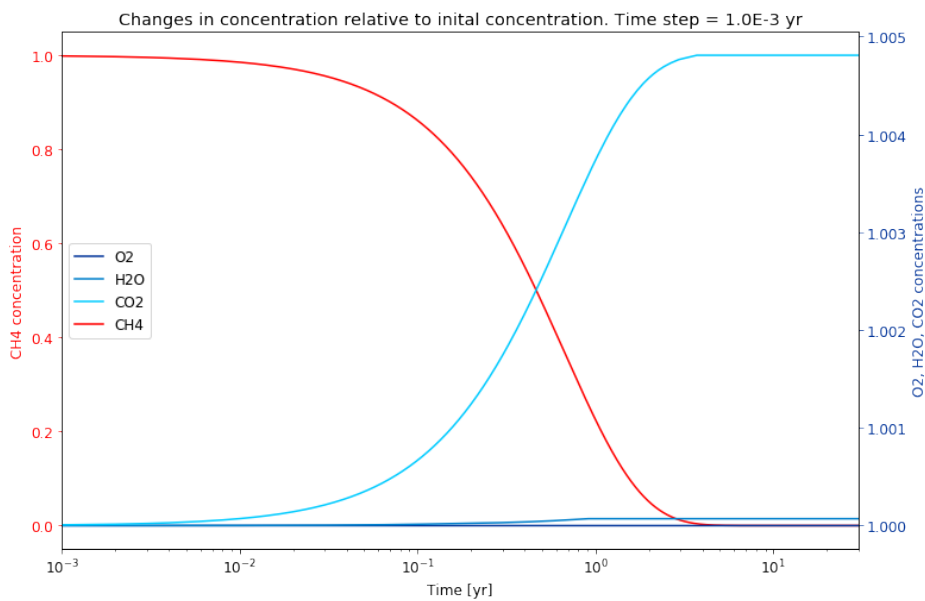
In addition to the values used for the reaction rate calculations, Table 1 shows the initial concentrations for each of the participating molecular species. These values are the approximate concentrations found in the modern Earths atmosphere, as they are listed in Lodders et al. [1998].

### 5.1.2 Concentrations over time

The results from this initial model run, based on the inputs listed in Table 1, can be seen in Fig. 5. Fig. 5 shows the change in concentration relative to the initial concentration of the four participating species ( $O_2$ ,  $H_2O$ ,  $CO_2$  and  $CH_4$ ) as a function of time. These relative concentrations are found by dividing the concentration at each time step with the initial concentration

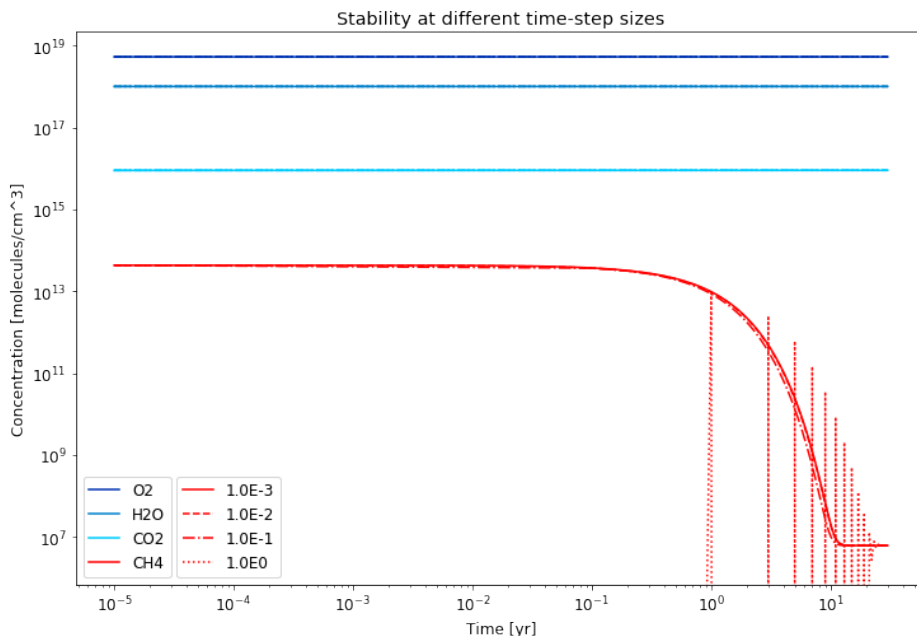
$$N_{t,relative} = \frac{N_t}{N_0} \quad (41)$$

and hence have no unit. Since we are especially interested in  $CH_4$ , and this species has the most drastic relative change during the run, the concentration of  $CH_4$  has been isolated on the left vertical axis and shown in red. whereas the other three species are shown in blue on the right vertical axis. Notice that the time axis is logarithmic. From Fig. 5 we can see that almost all the

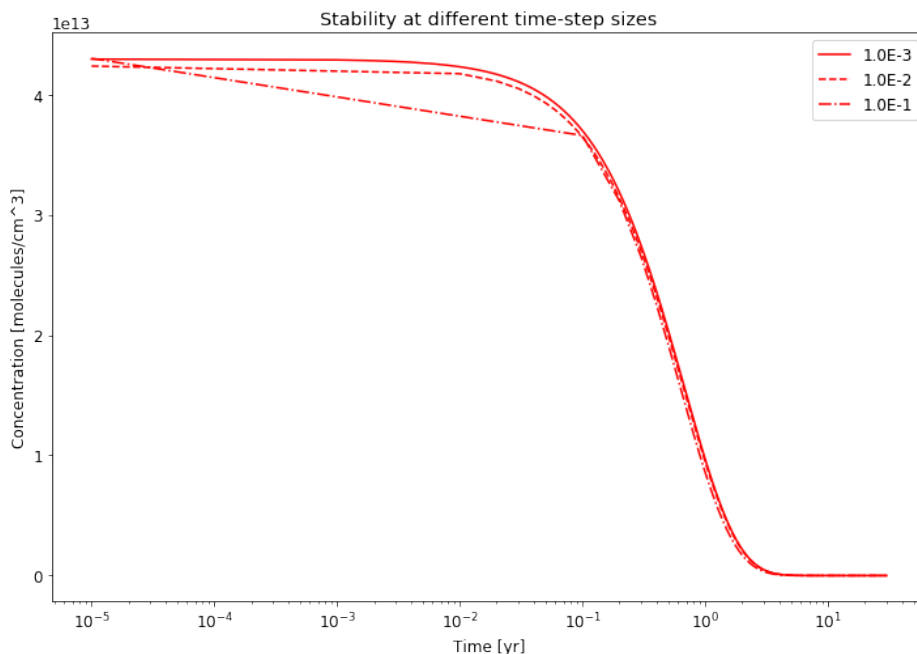


**Figure 5:** Relative change in concentration for test run of simplified reaction of the oxidation methane as shown in Eq. 25. ODEs are solved using FEM.

$CH_4$  is exhausted early on, and an equilibrium is reached as the curve flattens, at a concentration of  $[CH_4] \approx 6 \cdot 10^6 \text{ molecules cm}^{-3}$ . During the same period of time there is a slight relative increase in the concentration of  $CO_2$  ( $\sim 0.5\%$ ), whereas  $O_2$  and  $H_2O$  remain largely the same.



(a) Concentration of all participating species over time.



(b) Concentration of  $CH_4$  over time for three different time steps. Notice that unlike 6a, this figure has a linear secondary axis.

**Figure 6:** Concentrations of molecular species over time, for model runs with different time steps of the simplified oxidation of methane. The different line types indicate different sizes of the time step ( $h$ ): Full line:  $h = 10^{-3} \text{ yr}$ . Dashed line:  $h = 10^{-2} \text{ yr}$ . Dot-dashed line:  $h = 10^{-1} \text{ yr}$ . And dotted line:  $h = 10^0 \text{ yr}$ . ODEs are solved using FEM.

The time step used for the model in Fig. 5 is  $10^{-3} \text{ yrs} \approx 8 \text{ hr}$ , and the model runs for a year in total. The values for the time step and total time were chosen, such that the model could remain stable and an equilibrium could be reached. As explained in Sec. 2.3.3 and 4.2 the stability of a model can depend highly of the size of the time step, especially if the chemical ODEs are solved using an explicit method.

Fig. 6a shows the concentration of the different species as a function of time for four different time steps. Here we can see that the model becomes unstable at larger time steps, e.g. 1 year. From Fig. 6a we can see that the time steps of  $10^{-3}$ ,  $10^{-2}$ , and  $10^{-1}$  all follow similar trends of stable concentrations of  $CH_4$  with a gradual decrease until the concentration stabilizes at  $\sim 6 \cdot 10^6 \text{ molecules cm}^{-3}$ . For the time step of  $h = 1 \text{ year}$ , on the other hand, we see that the concentration of  $CH_4$  fluctuates rapidly between positive and negative values, indicating that the model is not stable at this time step. This is caused by the fact that the stability of FEM for solving ODEs depend on the size of the time step. Since FEM has no guarantee for positivity (which was one of the model demands listed on page 11), the estimated concentration might become negative when the time steps are too large, as explained in Sec. 2.4. For FEM to remain stable, the time steps should be significantly smaller than the shortest lifetime among the species present. In the case of this model the shortest lifetime is found for  $CH_4$ , and from Fig. 5 we can estimate this to  $\sim$  a few years, in accordance with the model loosing stability for time steps of  $\sim 1 \text{ year}$ .

If we look a bit closer at the short time steps in Fig. 6b we can see that there is a slight difference between the trends, mostly due to the data points, but that they are almost identical, indicating that these three time steps all yield the same results. If the code should be run for long periods of time, or if it should be coupled to MARCS and run for several iterations and atmospheric layers, it could therefore be beneficial to use the time step of  $10^{-1}$  to minimize the number of iterations, while still keeping the model stable. This simplified model reaches equilibrium after  $t \approx 17.5 \text{ years}$ .

## 5.2 Reaction chain for oxidation of methane, Forward Euler's method

To further test the model on a more complex network of reactions I will now look at the actual reaction chain for the oxidation of methane. There are a great number of reactions that might participate and dominate in this reaction chain, all depending on the exact conditions present during the reactions, such as temperature, pressure and which other molecular species are present in addition to the ones participating in the reaction chain. In this study, I will follow the reaction chain suggested in chapter 11 of Jacob [1999], which discusses the oxidizing power of the troposphere, to make sure that the reaction chain is appropriate for atmospheric conditions. The reactions participating in the reaction chain can be seen in Table 2. The first 13 reactions all participate directly in the conversion of  $CH_4$  and  $O_2$  to  $H_2O$  and  $CO_2$ , whereas the last three reactions are related to the production of the radicals  $OH$  and  $HO_2$ , both of which play critical roles in the main reaction chain.

**Table 2:** Reaction chain for oxidation of methane according to Jacob [1999]. Values for the frequency factor (A) and activation energy ( $E_a$ ) for both directions of the reaction have been found on the database NIST [2020]. Some reactions were not listed on the database, and have been given the values of similar reactions. These reactions are marked by a superscript in the table, and the reaction their values are based on are listed below.

Reaction	Forward		Reverse		Order
	A	$E_a$ [kJ/mol]	A	$E_a$ [kJ/mol]	
$CH_4 + OH \leftrightarrow CH_3 + H_2O$	4.16 E-13	10.24	1.2E-14	6.22E01	2
$CH_3 + O_2 \leftrightarrow CH_3O_2$	9.86E-09	22.37	2.03	139	2
$CH_3O_2 + HO_2 \leftrightarrow CH_3OOH + O_2$	7.69E-13	-10.81	1.20E-12 <sup>1</sup>	-1.08 <sup>1</sup>	2
$CH_3O_2 + NO \leftrightarrow CH_3O + NO_2$	2.80E-12	-2.37	2.80E-14 <sup>2</sup>	23.7 <sup>2</sup>	2
$CH_3OOH + OH \leftrightarrow CH_2O + OH + H_2O$	1.68E-12	-1.08 <sup>3</sup>	4.09E-15 <sup>4</sup>	429 <sup>4</sup>	2
$CH_3OOH + OH \leftrightarrow CH_3O_2 + H_2O$	1.20E-12	-1.08	4.00E-12 <sup>5</sup>	41.57 <sup>5</sup>	2
$CH_3OOH \leftrightarrow CH_3O + OH$	4.00E15	180	9.04E-16 <sup>6</sup>	7.89 <sup>6</sup>	1
$CH_3O + O_2 \leftrightarrow CH_2O + HO_2$	7.82E-14	9.56	5.63E-12	8.00E01	2
$CH_2O + OH \leftrightarrow HCO + H_2O$	8.20E-12	-0.33	8.54E-13	1.09E2	2
$CH_2O + O_2 \leftrightarrow HCO + HO_2$	3.40E-11	163	3.21E-30 <sup>7</sup>	1.79 <sup>7</sup>	2
$CH_2O \leftrightarrow CO + H_2$	3.49E-09	146	2.20E-10 <sup>8</sup>	43.56 <sup>8</sup>	2
$HCO + O_2 \leftrightarrow CO + HO_2$	8.60E-12	2.04	3.60E-13 <sup>9</sup>	75.99 <sup>9</sup>	2
$CO + OH \leftrightarrow CO_2 + H$	5.40E-14	-2.08	2.51E-10	1.11E02	2
$H + O_2 \leftrightarrow HO_2$	2.37E-33	-5.35	2.41E-08	2.03E02	3
$2HO_2 \leftrightarrow H_2O_2 + O_2$	2.20E-13	-4.99	9.00E-11	1.66E02	2
$H_2O_2 \leftrightarrow 2OH$	2.01E-07	1.90E02	6.71E-33	-1.79E01	2

<sup>1</sup>  $CH_3OOH + OH \leftrightarrow CH_3O_2 + H_2O$

<sup>2</sup> Both A and  $E_a$  are rough estimates from the forward reaction.

<sup>3</sup> Same as  $CH_3OOH + OH \leftrightarrow CH_3O_2 + H_2O$ .

<sup>4</sup>  $CH_2 + H_2O \leftrightarrow CH_3 + HO_2$

<sup>5</sup>  $CH_3O_2 + H_2O_2 \leftrightarrow CH_3OOH + HO_2$

<sup>6</sup>  $CH_3O + OH \leftrightarrow CH_2O + H_2O$

<sup>7</sup>  $HCO + H \leftrightarrow CH_2O$

<sup>8</sup>  $CO + H_2 \leftrightarrow HOCO$

<sup>9</sup>  $CO + HO_2 \rightarrow CO_2 + OH$

### 5.2.1 Choosing parameter values

The frequency factors and activation energies in Table 2 are found on the database NIST [2020], and for the reactions not listed on the database, I have used values reported for similar reactions. In the rightmost column the order of each of the reactions are listed, and as can be seen most of them are second order reactions. This is caused by the fact that the reactions are mainly bi-molecular elementary reactions, and as such the order of the reactions in relation to each of the two reactants will equal the stoichiometric coefficient for that reactant. A few of the reactions had both listings as second and third order reactions on the database, but for these I chose the values appropriate for second order reactions, in order to keep it simple for the model. The only reaction that was solely listed as a third order reaction is  $H + O_2 \rightarrow HO_2$ , but since the concentration of  $O_2$  must be expected to be much higher than that of  $H$ , I will list this reaction as a pseudo-first order in the model.

The concentrations for each of the species from Table 2 can be seen in Table 3. The middle column shows the abundances as they have been listed in different sources, dominated by Lodders et al. [1998]. Since  $N_2$  does not participate in the reaction chain, I have not included this species in the

Species	Abundances	Concentrations [ <i>molecules cm<sup>-3</sup></i> ]
$O_2$	20.9 %	$5.3 \cdot 10^{18}$
$H_2O$	< 4 %	$1.0 \cdot 10^{18}$
$CO_2$	350 ppm	$8.9 \cdot 10^{15}$
$CH_4$	1.7 ppm	$4.3 \cdot 10^{13}$
$CH_3$		$1.0 \cdot 10^6$
$OH$	$10^6$ molecules per $cm^{-3}$	$1.0 \cdot 10^6$
$HO_2$	10 ppt <sup>1</sup>	$2.6 \cdot 10^8$
$H_2O_2$	1 ppb	$2.6 \cdot 10^{10}$
$H$		$1.0 \cdot 10^6$
$H_2$	0.6 ppm <sup>2</sup>	$1.5 \cdot 10^{13}$
$HCO$		$1.0 \cdot 10^6$
$CO$	125 ppb	$3.2 \cdot 10^{12}$
$CH_2O$	$0.5 \mu g m^{-3}$ <sup>3</sup>	$1.0 \cdot 10^{10}$
$CH_3O$		$1.0 \cdot 10^6$
$CH_3O_2$	$\sim 10^8$ molecules per $cm^{-3}$ <sup>4</sup>	$1.0 \cdot 10^8$
$CH_3OOH$	0.1 ppb <sup>5</sup>	$2.6 \cdot 10^9$
$NO$	100 ppt	$2.6 \cdot 10^9$
$NO_2$	100 ppt	$2.6 \cdot 10^9$

**Table 3:** The abundances in this table are found in a Planetary Scientists Companion, with a few exceptions marked by superscripts in the table and listed below. The concentrations are found assuming an atmospheric density of  $2.55 \cdot 10^{19}$  *molecules cm<sup>-3</sup>* [Jacobson, 2005]. I have not been able to find reports on the abundances for some trace species, and the concentration of these are marked in grey and have been set to the lowest value among the known species [ $OH$ ] =  $1.0 \cdot 10^6$  *molecules cm<sup>-3</sup>*

<sup>1</sup> Concentration estimated from Bianchi et al. [2019]

<sup>2</sup> Concentration from Glueckauf and Kitt [1957]

<sup>3</sup> Concentration from sec. 1.3.3 in for Research on Cancer et al. [1995]

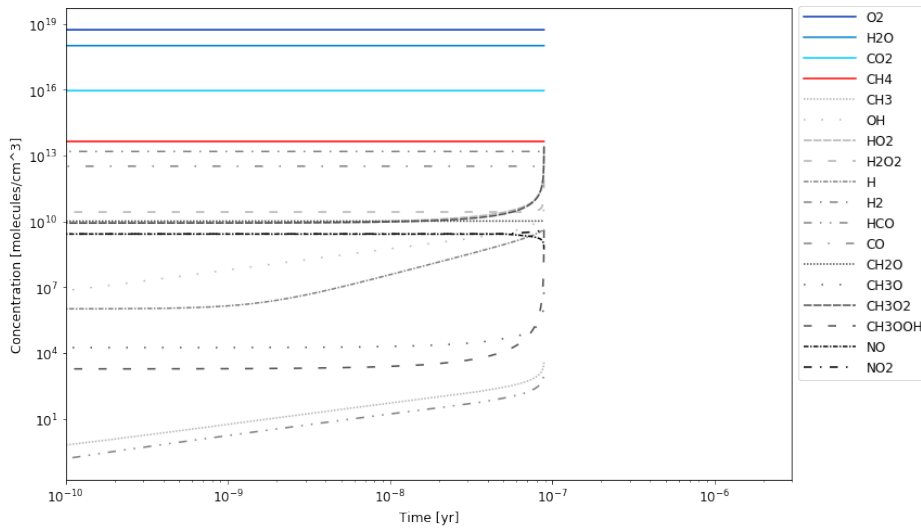
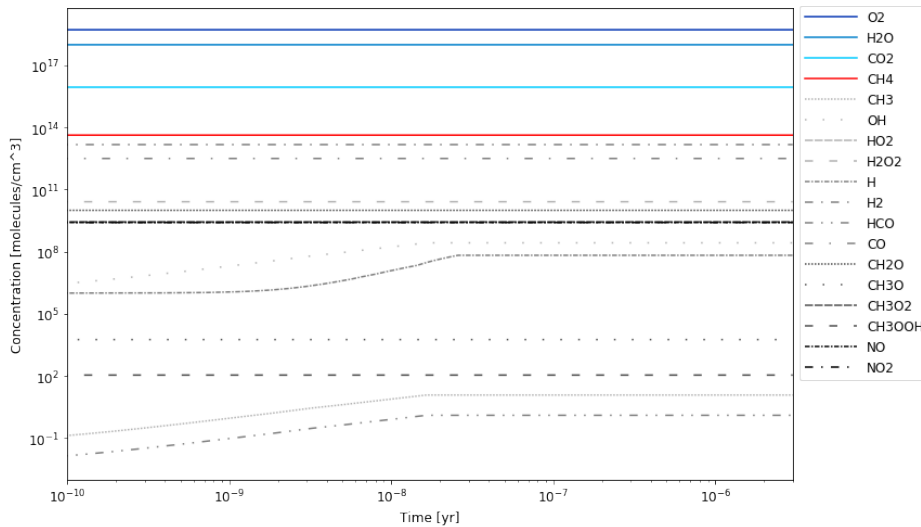
<sup>4</sup> Concentration from Onel et al. [2017]

<sup>5</sup> Concentration from Zhang et al. [2012]

model. The abundances in the middle column will therefore not sum up to 100%. This should in no way affect the model. The rightmost column shows the concentrations in *molecules cm<sup>-3</sup>*, as I have listed them in the input file for the model. Some of the species that were not listed in Lodders et al. [1998], have been found in various other papers, as indicated in the table footnotes. The few trace elements that I have not been able to find references for, have been listed to have the same initial concentrations as the radical  $OH$ , since this has the lowest documented concentration among the species in the table.

### 5.2.2 Concentrations over time

In Fig. 7 we see how the concentration of the different molecules from Table 3 changes over time for two different time steps. These models were difficult to stabilise, as I will explain later, and in order for them to remain stable for any period of time, I had to introduce some measure of positivity. This was done by fixing all concentrations below  $10^{-2}$  *molecules cm<sup>-3</sup>* at  $10^{-2}$  for each time step. Since it is not physically possible for an atmospheric concentration to be negative, I find this to be

(a) Concentrations of all species over time, for  $h = 10^{-14}$  years(b) Concentrations of all species over time, for  $h = 10^{-15}$  years

**Figure 7:** Concentrations of all species over time for the reaction chain of the oxidation of methane. The four main species  $O_2$ ,  $H_2O$ ,  $CO_2$ , and  $CH_4$ , that were also studied in the simplified reaction are plotted in blue and red respectively, in compliance with their color code in Fig. 5. The remaining species are plotted in shades of grey with different line types. ODEs are solved using FEM.

a reasonable approximation, and it was introduced to all following models of the methane reaction chain. The total time period observed is  $3.0 \cdot 10^{-6}$  years and the time steps are  $10^{-14}$  and  $10^{-15}$  years. The system in Fig. 7a does not reach an equilibrium before the model becomes unstable at  $\sim 10^{-7}$  years. The reason for this is, that the complexity of the reaction chain puts very high demands on the size of the time step. In addition to the higher number of molecules and reactions introduced in this model, the model also contains several radicals, that generally have very short lifetimes. Since the stability of FEM is highly dependent on the lifetime of the most short-lived species, and some radicals can have lifetimes of less than milliseconds ( $\sim 10^{10}$ ), the time step must be significantly below this threshold [Cariolle et al., 2017].

Fig. 7b, on the other hand, does seem to reach a plateau, but this happens without any change in

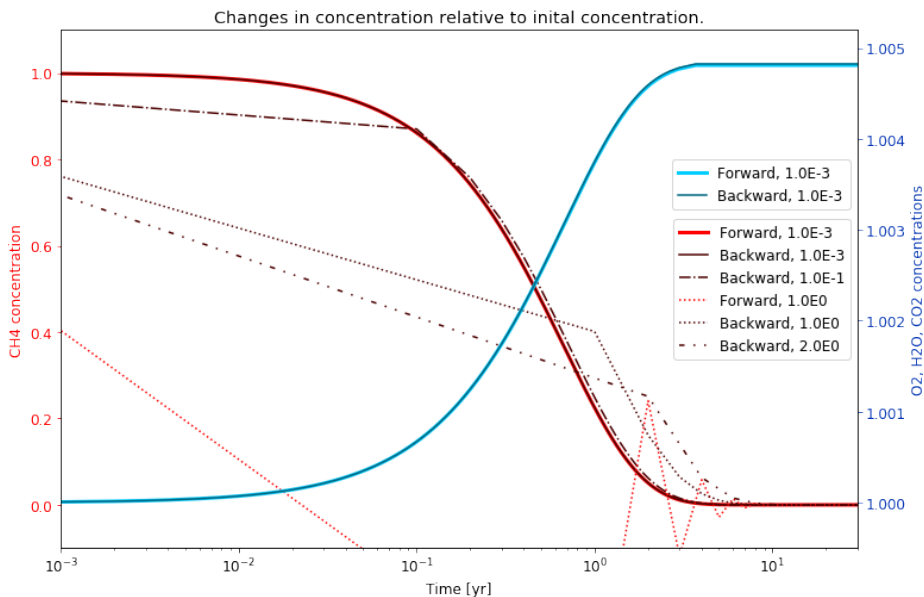
concentration of the four main species we are observing,  $O_2$ ,  $H_2O$ ,  $CO_2$ , and  $CH_4$ . The lack of change in concentration of some species might be caused by the fact that some reactions have much lower rates than others, and that the species participating in these reactions therefore need more time for any significant change to occur. The combination of very short time scales for some species (such as some of the radicals) and very long time scales for others (such as  $O_2$ ) is what makes chemical systems *stiff*, as explained in Sec. 2.4. Stiff systems are very difficult to treat using explicit ODE solvers, like FEM, since the short lived species require very short time steps for the model to remain stable, and the long lived species require a very long total time for the model to run. This leads in a very high number of iterations and a proportionately great number of computations which results in a very long run time, which is not suitable for a model we would like to follow for several years.

### 5.3 Comparing stability of ODE solvers; Forward- and Backward Euler's methods

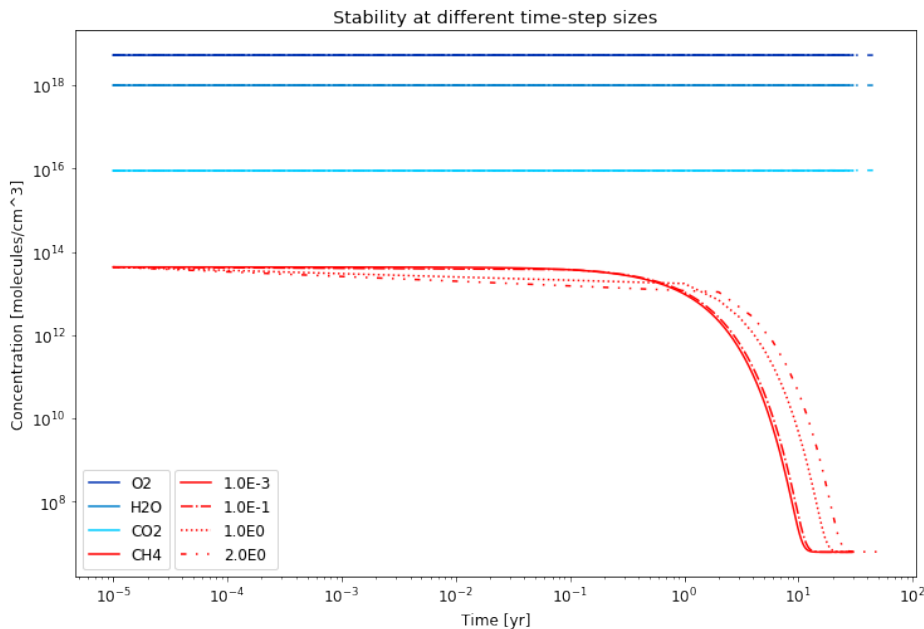
In order to better process the stiff chemical systems I have tried to run the models using the implicit Backward Euler's method (BEM), rather than the explicit Forward Euler's method, as described in Sec. 4.2.

Running the model run for the simplified reaction studied in Sec. 5.1, where the BEM was used to solve the ODEs, gave the results that can be seen in Fig. 8. Fig. 8a shows the relative change in concentrations of  $CH_4$  and  $CO_2$  with different time steps for the Forward and Backward Euler's method respectively. The full lines illustrates the model runs for  $h = 10^{-3}$  years, for each of the models, corresponding to the results in Fig. 5. Here we see that the concentrations for both  $CH_4$  and  $CO_2$  are identical for the two Euler's methods, indicating that the models are equally accurate at this time step size. At  $h = 10^0$  years (the dotted lines), however, FEM becomes unstable, whereas BEM maintains stability for both  $h = 10^0$  years,  $h = 2.0 \cdot 10^0$  years, and even higher time steps. The stability of BEM is even more visible in Fig. 8b, where we can see the concentrations of each of the four species at time steps from  $10^{-3}$  to 2 years. Here we can see that all models remain stable and follow the same trend. Since the greatest change in concentration happens within the first year, however, the time steps should be significantly shorter than this making both Euler's methods viable.

To test the stability of BEM for the more complex reaction chain studied in Sec. 5.2, I have plotted the change in concentration for all species for a model run of  $10^{-12}$  years in Fig. 9. In Fig. 9 we see that the concentrations vary a lot more than what we saw in Fig. 7. We also see a drastic decrease in the concentration of methane before the curve mostly flattens at  $t \sim 3 \cdot 10^{-6}$  years. The concentrations of some species keeps changing throughout the entire time period, and the model does not reach a final equilibrium before it becomes unstable at  $\sim 10^{-5}$  years. BEM has a requirement for positivity, and as such the instability does not occur due to fluctuations between positive and negative values (as was the case for  $h = 10^0$  years in Fig. 6). Instead the instability occurs due to overflow and/or possibly precision errors during the calculations. I have not yet been able to locate the exact source of the errors and fix it. All relevant variables at the model run for Fig. 9 are at double precision in order to minimize this effect slightly.



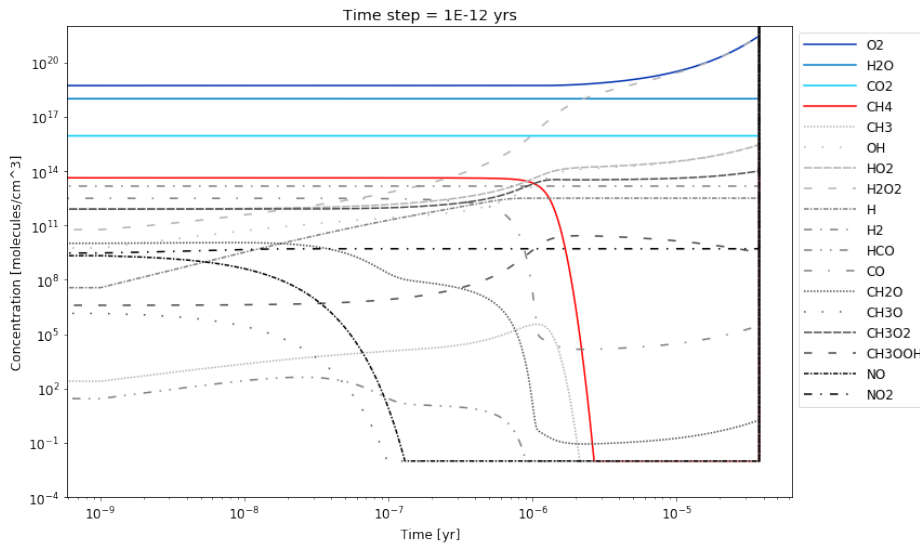
(a) Comparing the stability of the Forward and Backward Euler’s method at different time step sizes. The blue lines indicate the relative change in concentration for CO<sub>2</sub>, as is marked by the right secondary axis. The red lines indicate the relative change in concentration for CH<sub>4</sub>, as is marked by the left secondary axis.



(b) Concentration of the four participating species over time.

**Figure 8:** Results of the model runs for the simplified oxidation of methane, where the Backward Euler’s method was used to solve the chemical ODEs. The different line types indicate model runs for different sizes of the time step ( $h$ ) as listed in the legends where the values have the unit [years].





**Figure 9:** Concentrations plot for all species in the reaction chain for the oxidation of methane. The set of chemical ODEs is solved using the Backward Euler's method. The color code is consistent with previous plots.

#### 5.4 How does it fit with actual methane production

None of the results from the cases I have tested so far have any real scientific value, since the first reaction was extremely simplified, and the reaction chain was both simplified never had the chance to properly stabilize. To illustrate this, I will try comparing the results from my test runs with the actual methane production of Earth.

According to my test runs the entire atmospheric abundance of methane was exhausted after 17.5 years for my simplified case and  $3 \cdot 10^{-6}$  years for the reaction chain using BEM. In order for the concentration of methane to remain constant in the two cases, this corresponds to an annual release of roughly  $2.5 \cdot 10^{12}$  and  $1.4 \cdot 10^{19}$  molecules  $CH_4$  per  $cm^3$  atmosphere per year, respectively. If we assume that the atmosphere has a volume of  $\sim 10^{12} km^3$ , and the molar mass of  $CH_4$  is  $16.04 g mol^{-1} = 2.7 \cdot 10^{-23} g molecules^{-1}$ , the total annual emission in the two test cases would be:

$$\begin{aligned} 2.5 \cdot 10^{12} \text{ molecules } cm^{-3} \cdot 10^{27} cm^3 \cdot 2.7 \cdot 10^{-23} g \text{ molecules}^{-1} &= 6.7 \cdot 10^{16} g = 6.7 \cdot 10^{13} kg \\ 1.4 \cdot 10^{19} \text{ molecules } cm^{-3} \cdot 10^{27} cm^3 \cdot 2.7 \cdot 10^{-23} g \text{ molecules}^{-1} &= 3.8 \cdot 10^{23} g = 3.8 \cdot 10^{20} kg \end{aligned} \quad (42)$$

The actual global annual emission of  $CH_4$  into the atmosphere by both natural sources and human activity is estimated to be  $570 Mt = 5.7 \cdot 10^{11} kg$  [IEA, 2020]. As can be seen the test results differ from the actual emission with at least two to nine orders of magnitude, indicating that we are far from being able to reproduce the reactions taking place in Earth's atmosphere. It should be noted that these calculations are very rough estimates and that the model result might vary significantly, if a  $CH_4$  source was actually introduced to the model and implemented for each iteration. However, the comparison does illustrate that the DISEQ model still needs a lot of work both in regard to stability and reaction network complexity before we can expect to get trustworthy results.

## 6 Discussion and conclusions

The next big step in exoplanet exploration is going to be the investigation of their atmospheres. During the next ten years missions such as JWST and ARIEL will make detailed observations of the spectra of exoplanet atmospheres, and in order to be able to analyse these spectra, we need to understand the mechanisms taking place in atmospheres. This understanding is gained through atmosphere modelling, that can help us recreate not only the final state and spectrum of an atmosphere, but also its evolution.

The models that are currently used at the research group StarPlan at the University of Copenhagen are the atmosphere model MARCS and the equilibrium chemistry model GGchem. Basing an atmospheric chemistry model on equilibrium constants, such as it has been done for GGchem, might be problematic, since it prevents the model from simulating atmospheres in disequilibrium. It has been found that low temperature atmospheres rarely are in equilibrium due to e.g. photolysis in the upper atmosphere, interactions with the surface, and transport induced quenching [Moses, 2014, Venot and Agúndez, 2015, Wogan and Catling, 2020]. Metabolism caused by the presence of life on a planet might also be a great source of chemical disequilibrium, and as such atmospheric disequilibrium, especially between redox-couples, has been suggested as a potential biosignature (e.g. Lovelock [1965], Krissansen-Totton et al. [2016, 2018], Wogan and Catling [2020]). That it might not be easy to interpret the upcoming exoplanet spectra is well illustrated by the fact of how difficult and yet inconclusive it has been to understand the physics and potential biology behind the identified  $CH_4$  on Mars and  $PH_3$  on Venus [Greaves et al., 2020, Gloesener et al., 2020]. In this thesis I have introduced the beginning of one necessary factor to facilitate these interpretations within the tools we are working with at StarPlan.

During this project I have designed a chemistry disequilibrium model, DISEQ, that can calculate the change in concentration of chosen atmospheric species based on the reaction rates of the reactions they participate in. The reaction rates are calculated using the Arrhenius equation based on values provided by the user for the frequency factor and activation energy for each of the reactions. The change in concentration for each of the atmospheric species are expressed as a set of chemical ordinary differential equations (ODEs) that are solved numerically. During test runs for DISEQ I tested the stability of the model for two different ODE solvers for the oxidation of methane expressed as a simplified reaction and through a reaction chain respectively.

*The explicit Forward Euler's method* remained stable for the simplified reaction for time steps of  $h = 10^{-1}$  years and smaller. The model ran for a total of 30 years and reached equilibrium after  $\sim 17.5$  years. For the more extended reaction chain the time step had to be at  $10^{-15}$  years or smaller for the model to remain stable for any period of time. Even then the model could not run for longer than  $10^{-5}$  years without losing stability by getting negative concentrations.

*The implicit Backward Euler's method* had guaranteed positivity and remained stable for all time step sizes tested for the simplified model ( $10^{-3} \leq h \leq 2 \cdot 10^0$  years). For the reaction chain the model remained stable and yielded results for time steps up to  $h = 10^{-11}$  years. However, for longer runs I kept experiencing problems with overflowing and possibly rounding errors. I did not manage to find the cause for these problems, but I believe they might be related to the combination of very large val-

ues, e.g. for the molecular concentrations ( $10^{20}$ ), and very small values, e.g. for the frequency factor ( $10^{-30}$ ). The issue might be solved by re-scaling to more manageable, possibly unit-less, variables. In general the explicit method excels by having very little computation for each iteration, and as such might be the best suited to very simple cases that require few iterations, such as the simplified oxidation of methane. The implicit method takes more computation for each of the iterations, but excels by remaining stable at lower time step sizes, and thereby require few iterations. This might be advantageous for more complex reaction chains. A beneficial future prospect for DISEQ could be to implement a semiimplicit ODE solver with adaptable time steps, such as Gear's method [Shampine and Gear, 1979, Jacobson, 2005]. This method is especially well suited to stiff systems, such as chemical systems, where the reaction rates might vary significantly from reaction to reaction, as it adapts the size of the time step to minimize the number of iterations while still remaining stable.

The model has been designed to be user friendly and flexible, and many decisions have been made with this focus. All major calculations (such as the calculation of reaction rates using Arrhenius equation and the two Euler's method) are done through subroutines, such that they can easily be changed or replaced if desired. Currently changing between methods has to be done in the script itself, but a future prospect could be to add a "switch" in the form of an input variable specifying which method to use. At the moment all reactions have individual sections at the beginning of the script, such that additional reactions can be added by the user. The plan was to have all reactions and input parameters (i.e. time, time step, ODE solver setting, choice of input file, and output file name) in a separate file, such that the user only had to deal with that file to change any of the inputs.

The model is written in Fortran, and meant to work in cooperation with MARCS, such that MARCS can pass on the temperature and initial molecular abundances of a specific atmospheric layer, and DISEQ can pass back new molecular concentrations. This coupling has not yet been done.

For this project I have chosen to calculate the concentrations using the reaction rates by solving a series of chemical ODEs through iterations. I thereby avoid equilibrium chemistry, equilibrium constants, and the equilibrium model GGchem completely in my calculations. In contrast, other disequilibrium models base their calculation on equilibrium states, by e.g. adding measures of disequilibrium to specific species in otherwise equilibrium systems [Miles et al., 2020]. One way to do this could be to base the model on the *quench approximation*, where the atmosphere is assumed to generally be in equilibrium, except for species in areas where the transport timescale exceeds the chemical timescales, and where the concentration of that species is locked at a disequilibrium value [Marley and Robinson, 2015, Zahnle and Marley, 2014, Fortney et al., 2020].

The main benefit of calculating the concentration based on the reaction rates is that it is a way of forward modelling, and that it will result in self-consistent solutions. This is not necessarily the case for the disequilibrium models based on equilibrium states [Fortney et al., 2020]. In addition to this a time dependency is introduced through the reaction rate, and the evolution of the atmosphere model can be followed in detail.

One of the benefits of basing the calculations on equilibrium states rather than calculating the concentrations directly from reaction rates, is that many of the values needed to calculate the reaction rates must be determined experimentally. Values such as the activation energy and frequency factor

must be determined experimentally for each of the reactions included in the model, and for many atmospheric reactions this has not been done yet. During the test runs done in this project there were several reactions for which the values had not been determined experimentally, and for these reactions the values were estimated based on similar reactions. This will of course introduce a significant uncertainty, since there is no guarantee that similar reaction will have similar activation energies and frequency factors, and it should be avoided for non-test runs. For reactions where at least one direction has been investigated experimentally (which is the case for all in this study), this issue can be solved by determining the rate for one direction based on the rate of the other, as done by Zahnle et al. [2009], Zahnle and Marley [2014]. This can be done by determining the equilibrium constant ( $K$ ) through a minimization of Gibb's free energy, and then utilizing that the equilibrium constant will equal the rate between the two directions of the reaction  $K = \frac{r_1}{r_2}$ . One can then find the rate constant of the unknown reaction, and thereby avoid the need of knowing the frequency factor and activation energy beforehand.

As mentioned previously atmospheric disequilibrium between redox-couples has been suggested as a potential biosignature, or even as a potential anti-biosignature depending on the species in disequilibrium [Wogan and Catling, 2020]. This could, however, be complicated by the fact that spectral lines caused by redox-couples are often widely separated on the spectrum, and as such are difficult to observe simultaneously [Seager and Deming, 2010]. Even then, it is still highly advantageous to model disequilibrium, since it will both allow us to more accurately reproduce the atmospheres that are out of equilibrium, and it can give us an increased insight into the structure of an atmosphere. Since disequilibrium can change the overall structure of an atmosphere, studies have used disequilibrium tracers to probe the pressure-temperature profiles of brown dwarfs Miles et al. [2020], which has been suggested to not only help us understand the dynamics and chemistry taking place, but also to some degree the history and evolution of the atmosphere [Fortney et al., 2020].

Chemical disequilibrium models are important tools for understanding atmospheres, and if we are to be able to properly analyse the exoplanet atmosphere spectra we can soon expect to get from telescopes such as JWST, ELT, and ARIEL, it is highly beneficial to include disequilibrium in our atmospheric chemistry models.

## 7 Additional Projects

In order to expand my general understanding of the field of exoplanets I have done a few additional "projects" parallel to the main research of my Master's thesis. In the following sections I will give a short overview of these projects, and how I believe they have benefitted me professionally.

### 7.1 Article on the correlation between orbital eccentricity and multiplicity of exoplanet systems

During the past few years I have been doing research on the dynamics of exoplanet systems under the supervision of Uffe Gråe Jørgensen. When we look at our known sample of exoplanetary systems, we can see that they are often very different from the Solar System, which could potentially be worrying if we are searching for Earth-twins and biosignatures. Two of the characteristics that seem to be unusual about the Solar System compared to most observed exoplanet systems, are the very circular orbits of the planets in the Solar Systems and the high number of planets in the system. Previous studies have found that there might be a correlation between the orbital eccentricity of exoplanets and the number of planets in their systems (multiplicity). We wanted to test this correlation based on observational data from the database *exoplanet.eu*, and to use our findings to estimate how unusual a system such as the Solar System might be. Our study differed from previous studies by having a much larger data sample, by focusing on minimising bias related to detection method and planet type, and by expanding on our results to hopefully get an insight into the underlying planet population.

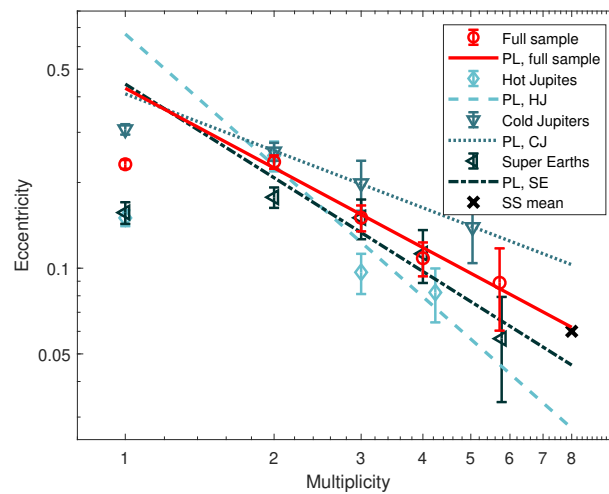
Similarly to previous reports, we found that the correlation between the eccentricity and multiplicity could be described as a power law for multiplicities of two and above. One of the figures showing the eccentricity-multiplicity correlation from my study can be seen in Fig. 10. Fig. 10 shows the correlation for the full sample (red) as well as three subsamples consisting of systems with different planet types. The figure indicates that the full sample follows a power law correlation (fitted as a full, red line) almost perfectly for multiplicities above one. The Solar System also follows this trend without having been included in the fit for the power law, indicating that the low eccentricities of our orbits are not unusual, when the multiplicity of our system is taken into account. The power law for the full sample can be described by:  $e(M) = 0.429 \cdot M^{-0.93}$

We can also see that the subsamples of systems with hot-Jupiters, cold-Jupiters, and super-Earths (i.e. systems with no giant planets) all follow similar trends. This indicates that the power law correlation is not dependent on the planet type, which is comforting since some planet types are very difficult to observe, and are therefore highly underrepresented in our observation samples.

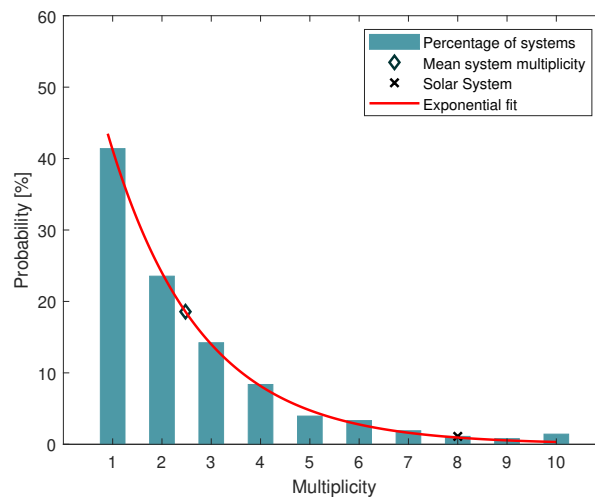
The only outliers from the power law correlations, for any of the subsamples, were the one-planet systems. Other studies had suggested that many one-planet systems can be expected to contain additional undiscovered planets, which could explain why they deviate from the power law correlation. Based on this assumption we estimated how many planets each one-planet system should contain, according to their eccentricities and the power law correlation written earlier. We redistribute the systems based on the estimated multiplicities, to obtain a probability distribution that might more

accurately reflect of the underlying multiplicity distribution of exoplanet systems. This distribution can be seen in Fig. 11. Based on the new probability distribution we estimated that the probability of a systems having eight planets or more, like the Solar System, was  $\sim 1\%$  which would mark the Solar System as an unusual system but far from exclusive.

Parallel to the main research of my Master’s thesis project, this study was written up as a paper and submitted to Monthly Notices of the Royal Astronomical Society in March with me as the main author. The third revision has recently been handed in after only minor changes were requested before the article would be ready to be reconsidered for publication. The (probably) final version of the paper can be seen in appendix B.



**Figure 10:** Eccentricity-multiplicity correlation plotted for the full sample and three subsamples sorted for planet types. Dashed: Systems containing at least one hot-Jupiter planet (HJ). Dotted: Systems containing at least one cold-Jupiter planet (CJ). Dot-dashed: Systems with super-Earths and no giant planets. The Solar System (SS) is plotted as a black cross. Power law correlations have been fitted to all samples and are referred to as PL in the legend.



**Figure 11:** Probability distribution of the planet sample after planets that might have been circularized through planet-star interactions have been excluded, and systems with potentially undiscovered planets have been redistributed.

## 7.2 Observations at La Silla Observatory

So far, all my research in physics has been based on either programming, data processing, or theoretical work. In order to get a better understanding of the observational aspects of astrophysics I volunteered to act as the on-site observer at the Danish 1.54-metre telescope at La Silla Observatory for two weeks at the beginning of my Master's thesis project.

During this time I did observations as a part of the MiNDSTEp (Microlensing Network for the Detection of Small Terrestrial Exoplanets) collaboration. MiNDSTEp especially focuses on microlensing observations of exoplanets towards the center of the galaxy, but also participates in smaller side projects including instrument development, globular clusters, transits, and asteroid and comet observations. The Danish 1.54-metre telescope has two main instruments, Lucky Imaging (LI) and Danish Faint Object Spectrograph and Camera (DFOSC). I spent most of my time operating and calibrating LI to do microlensing observations and globular cluster observations, and had a few longer sessions of transit observations using DFOSC. I was the only observer at the telescope during most of my stay and did the observations alone or in cooperation with a remote observer. This has given me a valuable experience in observational astronomy and the retrieval of data (and has additionally given me a huge respect for observers).

I have not processed any of the data from the MiNDSTEp consortium myself, nor participated in the analysis of the results, but MiNDSTEp data has recently been used for a study of cometary activity by Kelley et al. [2020], in which I am a co-author due to my contribution to the MiNDSTEp observations. Based partly on data taken by DFOSC Kelley et al. [2020] study mini-outbursts of the comet 46P/Wirtanen by comparing their frequency and magnitude to those seen from other comets, in order to analyse the nuclear surface of the comet and its evolution. They find that the surface of 46P/Wirtanen seem to be less rough (have fewer cliffs per area) than two out of the three comets it is compared to, which is in agreement with a predicted correlation between mini-outbursts and the collapse of surface features such as cliffs. The article by Kelley et al. [2020] is attached as appendix [C](#).



## References

- Giada Arney, Shawn D Domagal-Goldman, Victoria S Meadows, Eric T Wolf, Edward Schwieterman, Benjamin Charnay, Mark Claire, Eric Hébrard, and Melissa G Trainer. The pale orange dot: the spectrum and habitability of hazy archean earth. *Astrobiology*, 16(11):873–899, 2016.
- Charles Beichman, Bjoern Benneke, Heather Knutson, Roger Smith, Pierre-Olivier Lagage, Courtney Dressing, David Latham, Jonathan Lunine, Stephan Birkmann, Pierre Ferruit, et al. Observations of transiting exoplanets with the james webb space telescope (jwst). *Publications of the Astronomical Society of the Pacific*, 126(946):1134, 2014.
- Björn Benneke, Ian Wong, Caroline Piaulet, Heather A Knutson, Joshua Lothringer, Caroline V Morley, Ian JM Crossfield, Peter Gao, Thomas P Greene, Courtney Dressing, et al. Water vapor and clouds on the habitable-zone sub-neptune exoplanet k2-18b. *The Astrophysical Journal Letters*, 887(1):L14, 2019.
- Federico Bianchi, Theo Kurten, Matthieu Riva, Claudia Mohr, Matti P Rissanen, Pontus Roldin, Torsten Berndt, John D Crouse, Paul O Wennberg, Thomas F Mentel, et al. Highly oxygenated organic molecules (hom) from gas-phase autoxidation involving peroxy radicals: A key contributor to atmospheric aerosol. *Chemical reviews*, 119(6):3472–3509, 2019.
- Daniel Cariolle, Philippe Moinat, Hubert Teyssèdre, Luc Giraud, Béatrice Josse, and Franck Lefèvre. Asis v1. 0: an adaptive solver for the simulation of atmospheric chemistry. 2017.
- International Agency for Research on Cancer et al. Volume 62: Wood dust and formaldehyde. *IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans*, 62, 1995.
- Jonathan J Fortney. Modeling exoplanetary atmospheres: An overview. In *Astrophysics of Exoplanetary Atmospheres*, pages 51–88. Springer, 2018.
- Jonathan J Fortney, Channon Visscher, Mark S Marley, Callie E Hood, Michael R Line, Daniel P Thorngren, Richard S Freedman, and Roxana Lupu. Beyond equilibrium temperature: How the atmosphere/interior connection affects the onset of methane, ammonia, and clouds in warm transiting giant planets. *arXiv preprint arXiv:2010.00146*, 2020.
- Elodie Gloesener, Özgür Karatekinb, and Véronique Dehant. Stability and composition of ch<sub>4</sub>-rich clathrate hydrates in the present martian subsurface. *Icarus*, page 114099, 2020.
- E Glueckauf and GP Kitt. The hydrogen content of atmospheric air at ground level. *Quarterly Journal of the Royal Meteorological Society*, 83(358):522–528, 1957.
- Jane S Greaves, Anita MS Richards, William Bains, Paul B Rimmer, Hideo Sagawa, David L Clements, Sara Seager, Janusz J Petkowski, Clara Sousa-Silva, Sukrit Ranjan, et al. Phosphine gas in the cloud decks of venus. *Nature Astronomy*, pages 1–10, 2020.
- Thomas P Greene, Michael R Line, Cezar Montero, Jonathan J Fortney, Jacob Lustig-Yaeger, and Kyle Luther. Characterizing transiting exoplanet atmospheres with jwst. *The Astrophysical Journal*, 817(1):17, 2016.



- Bengt Gustafsson, Bengt Edvardsson, Kjell Eriksson, Uffe Gråe Jørgensen, Åke Nordlund, and Bertrand Plez. A grid of marcs model atmospheres for late-type stars-i. methods and general properties. *Astronomy & Astrophysics*, 486(3):951–970, 2008.
- IEA. Methane Tracker, 2020. Information retrieved from International Energy Agency (IEA), (accessed: 13.10.2020), <https://www.iea.org/reports/methane-tracker-2020>.
- Daniel J Jacob. *Introduction to atmospheric chemistry*. Princeton University Press, 1999.
- Mark Z Jacobson. *Fundamentals of atmospheric modeling*. Cambridge University Press, 2005.
- Michael S. P. Kelley, Tony L. Farnham, Jian-Yang Li, et al. Quintet of outbursts of comet 46p/wirtanen. unpublished, submitted to *The Planetary Science Journal*, 2020.
- Laura Kreidberg. Exoplanet atmosphere measurements from transmission spectroscopy and other planet-star combined light observations. *arXiv preprint arXiv:1709.05941*, 2017.
- Joshua Krissansen-Totton, David S Bergsman, and David C Catling. On detecting biospheres from chemical thermodynamic disequilibrium in planetary atmospheres. *Astrobiology*, 16(1):39–67, 2016.
- Joshua Krissansen-Totton, Stephanie Olson, and David C Catling. Disequilibrium biosignatures over earth history and implications for detecting exoplanet life. *Science advances*, 4(1):eaao5747, 2018.
- Andrew P Lincowski, Victoria S Meadows, David Crisp, Tyler D Robinson, Rodrigo Luger, Jacob Lustig-Yaeger, and Giada N Arney. Evolved climates and observational discriminants for the trappist-1 planetary system. *The Astrophysical Journal*, 867(1):76, 2018.
- Katharina Lodders, Bruce Fegley, et al. *The planetary scientist’s companion*. Oxford University Press on Demand, 1998.
- James E Lovelock. A physical basis for life detection experiments. *Nature*, 207(4997):568–570, 1965.
- Mark S Marley and Tyler D Robinson. On the cool side: Modeling the atmospheres of brown dwarfs and giant planets. *Annual Review of Astronomy and Astrophysics*, 53:279–323, 2015.
- Alan D McNaught and Andrew Wilkinson. IUPAC, Compendium of chemical terminology, 2nd ed. (the "Gold Book"). URL <https://doi.org/10.1351/goldbook>. Scientific Publications, Oxford (1997). Online version (2019-) created by S. J. Chalk. ISBN 0-9678550-9-8.
- Brittany E Miles, Andrew JI Skemer, Caroline V Morley, Mark S Marley, Jonathan J Fortney, Katelyn N Allers, Jacqueline K Faherty, Thomas R Geballe, Channon Visscher, Adam C Schneider, et al. Observations of disequilibrium co chemistry in the coldest brown dwarfs. *The Astronomical Journal*, 160(2):63, 2020.
- Julianne I Moses. Chemical kinetics on extrasolar planets. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2014):20130073, 2014.

- NIST. National Institute of Standards and Technology, 2020. Data retrieved from NIST Chemical Kinetics Database, (accessed: 05.09.2020), <https://kinetics.nist.gov/kinetics/index.jsp>.
- Lavinia Onel, Alexander Brennan, Paul W Seakins, Lisa Whalley, and Dwayne E Heard. A new method for atmospheric detection of the  $\text{CH}_3\text{O}_2$  radical. *Atmospheric Measurement Techniques*, 10 (10):3985–4000, 2017.
- Tyler D Robinson. A theory of exoplanet transits with light scattering. *The Astrophysical Journal*, 836(2):236, 2017.
- Tyler D. Robinson and Christopher T. Reinhard. Earth as an exoplanet, 2019.
- Edward W Schwieterman, Nancy Y Kiang, Mary N Parenteau, Chester E Harman, Shiladitya Das-Sarma, Theresa M Fisher, Giada N Arney, Hilairy E Hartnett, Christopher T Reinhard, Stephanie L Olson, et al. Exoplanet biosignatures: a review of remotely detectable signs of life. *Astrobiology*, 18(6):663–708, 2018.
- Sara Seager and Drake Deming. Exoplanet atmospheres. *Annual Review of Astronomy and Astrophysics*, 48:631–672, 2010.
- Sara Seager and Renéé Dotson. Exoplanets. space science series, 2010.
- Lawrence F Shampine and Charles William Gear. A user’s view of solving stiff ordinary differential equations. *SIAM review*, 21(1):1–17, 1979.
- Giovanna Tinetti, Pierre Drossart, Paul Eccleston, Paul Hartogh, Astrid Heske, Jérémy Leconte, Giusi Micela, Marc Ollivier, Göran Pilbratt, Ludovic Puig, et al. A chemical survey of exoplanets with ariel. *Experimental Astronomy*, 46(1):135–209, 2018.
- Angelos Tsaras, Ingo P Waldmann, Giovanna Tinetti, Jonathan Tennyson, and Sergey N Yurchenko. Water vapour in the atmosphere of the habitable-zone eight-earth-mass planet k2-18 b. *Nature Astronomy*, 3(12):1086–1091, 2019.
- S Udry, C Lovis, F Bouchy, A Collier Cameron, T Henning, M Mayor, F Pepe, N Piskunov, D Polacco, Didier Queloz, et al. Exoplanet science with the european extremely large telescope. the case for visible and near-ir spectroscopy at high resolution. *arXiv preprint arXiv:1412.1048*, 2014.
- Olivia Venot and Marcelino Agúndez. Chemical modeling of exoplanet atmospheres. *Experimental Astronomy*, 40(2-3):469–480, 2015.
- Nicholas F Wogan and David C Catling. When is chemical disequilibrium in earth-like planetary atmospheres a biosignature versus an anti-biosignature? disequilibria from dead to living worlds. *The Astrophysical Journal*, 892(2):127, 2020.
- Peter Woitke, Ch Helling, GH Hunter, JD Millard, GE Turner, M Worters, J Blečić, and JW Stock. Equilibrium chemistry down to 100 k-bar impact of silicates and phyllosilicates on the carbon to oxygen ratio. *Astronomy & Astrophysics*, 614:A1, 2018.

- K Zahnle, MS Marley, R St Freedman, K Lodders, and JJ Fortney. Atmospheric sulfur photochemistry on hot jupiters. *The Astrophysical Journal Letters*, 701(1):L20, 2009.
- Kevin J Zahnle and Mark S Marley. Methane, carbon monoxide, and ammonia in brown dwarfs and self-luminous giant planets. *The Astrophysical Journal*, 797(1):41, 2014.
- X Zhang, SZ He, ZM Chen, Y Zhao, and W Hua. Methyl hydroperoxide (ch<sub>3</sub>ooh) in urban, suburban and rural atmosphere: ambient concentration, budget, and contribution to the atmospheric oxidizing capacity. *Atmospheric Chemistry and Physics*, 12(19):8951, 2012.

## A Disequilibrium model, DISEQ

The next 7 pages contain the model for the simplified reaction for the oxidation of methane, with subroutines for both the Forward and Backward Euler's method.

---

```

!
! This program will calculate the new concentrations of atmospheric
! species based on an input of temperature, activation energy and
! previous concentrations
!
! The program is build up by the following structure:
!
! - Module "parameters":
!   - Defines universal parameters.
!   - USER must specify number of species and reactions
!   - USER must specify total time and time step
!
! - Main program "testdiseqbacknetwork":
!   - Set up reaction matrix
!   - Coefficient of each participating species in the reactions
!   - Parameters for reactions (activation energy, frequency factor)
!   - USER can add reactions and values at will
!   - Read input files
!     (- Temperature from MARCS output)
!     -Initial abundances (e.g. from MARCS output)
!   - Calculate new concentrations using OD solver in subroutine
!   - Set time and time step
!   - Call subroutine calc_conc_xxx, where xxx specifies ODE solver
!   - USER must specify ODE solver
!
! - Subroutines
!
! Files in use
! (100 - Output from MARCS)
! 101 - Initial abundances
! 102 - Output
!

```

---

```

module parameters
  implicit none

  ! Defining parameters
  integer(kind=8) :: ii_spec, ii_reac, ii_time      ! Loop indicies
  integer, parameter :: n_species = 18           ! Number of species
  integer, parameter :: n_reactions = 32         ! Number of reactions
  real, parameter :: temp_layer = 300           ! Temperature, REDEFINE if using MARCS input
  real, parameter :: time = 3E-5                ! Total time period, [yrs]
  real, parameter :: time_step = 1E-12         ! Length of each time step [yrs]
  real, parameter :: s_per_year = 3.15e7       ! Seconds per year
  real, parameter :: time_step_s = time_step*s_per_year ! Time step [s]
  real, parameter :: mols_per_mol = 6.022e23   ! Molecules per mole
end module parameters

program testdiseqbacknetwork
  use parameters
  implicit none

  ! Defining integers and real values for use in script
  real :: act_energy(n_reactions), fre_factor(n_reactions) ! Activation energy, frequency
  factor
  real :: concentrations(n_species) ! Concentrations
  integer :: reactions(n_species, n_reactions) ! Reaction matrix
  integer :: exponents(n_species, n_reactions) ! Exponents

  !
  ! REACTION MATRIX
  ! Matrix listing the atmospheric species and their participation in the chosen reactions.
  ! The reactions and along the columns, and the species are along the rows, but due to
  ! Fortran it will seem opposite.
  !

```

---

```

! PARTICIPATING SPECIES AND REACTIONS

```

```

! Order of species in input:
! O2, H2O, CO2, CH4, CH3, OH, HO2, H2O2, H, H2, HCO, CO, CH2O, CH3O, CH3O2, CH3OOH, NO,
NO2

! REACTION 1 and 2: CH4 + OH <=> CH3 + H2O
reactions(:,1) = (/&
! O2 H2O CO2 CH4 CH3 OH HO2 H2O2 H H2 HCO CO CH2O CH3O CH3O2 CH3OOH
NO NO2
& 0 , 1 , 0 , 0 , -1 , 1 , -1 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 ,
0 , 0 &
& /)
reactions(:,2) = -reactions(:,1)
! Both reactions are first order in regard to all reactants:
where (reactions(:,1,2])<0) exponents(:,1,2]) = 1 ! Setting exponents

act_energy(1) = 10.24 ! kJ/mol
fre_factor(1) = 4.16E-13 ! cm /((molecules s))
act_energy(2) = 6.22E01 ! kJ/mol
fre_factor(2) = 1.20E-14 ! cm /((molecules s))

! REACTION 3 and 4: CH3 + O2 <=> CH3O2
reactions(:,3) = (/&
! O2 H2O CO2 CH4 CH3 OH HO2 H2O2 H H2 HCO CO CH2O CH3O CH3O2 CH3OOH
NO NO2
& -1 , 0 , 0 , 0 , 0 , -1 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 1 , 0 ,
0 , 0 &
& /)
reactions(:,4) = -reactions(:,3)
! Both reactions are first order in regard to all reactants:
where (reactions(:,3,4])<0) exponents(:,3,4]) = 1 ! Setting exponents
act_energy(3) = 22.37 ! kJ/mol
fre_factor(3) = 9.86E-09 ! cm /((molecules s))
act_energy(4) = 139 ! kJ/mol FROM SIMILAR REACTION
fre_factor(4) = 2.03 ! cm /((molecules s)) FROM SIMILAR REACTION

! REACTION 5 and 6: CH3O2 + HO2 <=> CH3OOH + O2
reactions(:,5) = (/&
! O2 H2O CO2 CH4 CH3 OH HO2 H2O2 H H2 HCO CO CH2O CH3O CH3O2 CH3OOH
NO NO2
& 1 , 0 , 0 , 0 , 0 , 0 , -1 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , -1 , 1 ,
0 , 0 &
& /)
reactions(:,6) = -reactions(:,5)
! Both reactions are first order in regard to all reactants:
where (reactions(:,5,6])<0) exponents(:,5,6]) = 1 ! Setting exponents
act_energy(5) = -10.81 ! kJ/mol
fre_factor(5) = 7.69E-13 ! cm /((molecules s))
act_energy(6) = -1.08 ! kJ/mol FROM SIMILAR REACTION
fre_factor(6) = 1.20E-12 ! cm /((molecules s)) FROM SIMILAR REACTION

! REACTION 7 and 8: CH3O2 + NO <=> CH3O + NO2
reactions(:,7) = (/&
! O2 H2O CO2 CH4 CH3 OH HO2 H2O2 H H2 HCO CO CH2O CH3O CH3O2 CH3OOH
NO NO2
& 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 1 , -1 , 0 ,
-1 , 1 &
& /)
reactions(:,8) = -reactions(:,7)
! Both reactions are first order in regard to all reactants:
where (reactions(:,7,8])<0) exponents(:,7,8]) = 1 ! Setting exponents
act_energy(7) = -2.37 ! kJ/mol
fre_factor(7) = 2.80E-12 ! cm /((molecules s))
act_energy(8) = 23.7 ! kJ/mol FROM SIMILAR REACTION
fre_factor(8) = 2.80E-14 ! cm /((molecules s)) FROM SIMILAR REACTION

! REACTION 9 and 10: CH3OOH <=> CH2O + H2O
reactions(:,9) = (/&
! O2 H2O CO2 CH4 CH3 OH HO2 H2O2 H H2 HCO CO CH2O CH3O CH3O2 CH3OOH
NO NO2
& 0 , 1 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 1 , 0 , 0 , -1 ,
0 , 0 &
& /)
reactions(:,10) = -reactions(:,9)
! Both reactions are first order in regard to all reactants:
where (reactions(:,9,10])<0) exponents(:,9,10]) = 1 ! Setting exponents
act_energy(9) = -1.08 ! kJ/mol
fre_factor(9) = 1.68E-12 ! cm /((molecules s))
act_energy(10) = 429 ! kJ/mol FROM SIMILAR REACTION
fre_factor(10) = 4.09E-15 ! cm /((molecules s)) FROM SIMILAR REACTION

```

```

! REACTION 11 and 12: CH3OOH + OH <=> CH3O2 + H2O
reactions(:,11) = (/&
! O2  H2O  CO2  CH4  CH3  OH  HO2  H2O2  H  H2  HCO  CO  CH2O  CH3O  CH3O2  CH3OOH
NO  NO2
& 0  , 1  , 0  , 0  , 0  , -1  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 1  , -1  ,
0  , 0  &
& /)
reactions(:,12) = -reactions(:,11)
! Both reactions are first order in regard to all reactants:
where (reactions(:,[11,12])<0) exponents(:,[11,12]) = 1 ! Setting exponents
act_energy(11) = -1.08 ! kJ/mol
fre_factor(11) = 1.20E-12 ! cm /(molecules s)
act_energy(12) = 41.57 ! kJ/mol FROM SIMILAR REACTION
fre_factor(12) = 4.00E-12 ! cm /(molecules s) FROM SIMILAR REACTION

! REACTION 13 and 14: CH3OOH <=> CH3O + OH
reactions(:,13) = (/&
! O2  H2O  CO2  CH4  CH3  OH  HO2  H2O2  H  H2  HCO  CO  CH2O  CH3O  CH3O2  CH3OOH
NO  NO2
& 0  , 0  , 0  , 0  , 0  , 1  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 1  , 0  , -1  ,
0  , 0  &
& /)
reactions(:,14) = -reactions(:,13)
! Both reactions are first order in regard to all reactants:
where (reactions(:,[13,14])<0) exponents(:,[13,14]) = 1 ! Setting exponents
act_energy(13) = 180 ! kJ/mol
fre_factor(13) = 4.00E15 ! cm /(molecules s)
act_energy(14) = 7.89 ! kJ/mol FROM SIMILAR REACTION
fre_factor(14) = 9.04E-16 ! cm /(molecules s) FROM SIMILAR REACTION

! REACTION 15 and 16: CH3O + O2 <=> CH2O + HO2
reactions(:,15) = (/&
! O2  H2O  CO2  CH4  CH3  OH  HO2  H2O2  H  H2  HCO  CO  CH2O  CH3O  CH3O2  CH3OOH
NO  NO2
& -1  , 0  , 0  , 0  , 0  , 0  , 1  , 0  , 0  , 0  , 0  , 0  , 1  , -1  , 0  , 0  ,
0  , 0  &
& /)
reactions(:,16) = -reactions(:,15)
! Both reactions are first order in regard to all reactants:
where (reactions(:,[15,16])<0) exponents(:,[15,16]) = 1 ! Setting exponents
act_energy(15) = 9.56 ! kJ/mol
fre_factor(15) = 7.82E-14 ! cm /(molecules s)
act_energy(16) = 8.00E01 ! kJ/mol
fre_factor(16) = 5.63E-12 ! cm /(molecules s)

! REACTION 17 and 18: CH2O + OH <=> HCO + H2O
reactions(:,17) = (/&
! O2  H2O  CO2  CH4  CH3  OH  HO2  H2O2  H  H2  HCO  CO  CH2O  CH3O  CH3O2  CH3OOH
NO  NO2
& 0  , 1  , 0  , 0  , 0  , -1  , 0  , 0  , 0  , 0  , 0  , 1  , 0  , -1  , 0  , 0  , 0  ,
0  , 0  &
& /)
reactions(:,18) = -reactions(:,17)
! Both reactions are first order in regard to all reactants:
where (reactions(:,[17,18])<0) exponents(:,[17,18]) = 1 ! Setting exponents
act_energy(17) = -0.33 ! kJ/mol
fre_factor(17) = 8.20E-12 ! cm /(molecules s)
act_energy(18) = 1.09E02 ! kJ/mol
fre_factor(18) = 8.54E-13 ! cm /(molecules s)

! REACTION 19 and 20: CH2O + O2 <=> HCO + HO2
reactions(:,19) = (/&
! O2  H2O  CO2  CH4  CH3  OH  HO2  H2O2  H  H2  HCO  CO  CH2O  CH3O  CH3O2  CH3OOH
NO  NO2
& -1  , 0  , 0  , 0  , 0  , 0  , 1  , 0  , 0  , 0  , 0  , 1  , 0  , -1  , 0  , 0  , 0  ,
0  , 0  &
& /)
reactions(:,20) = -reactions(:,19)
! Both reactions are first order in regard to all reactants:
where (reactions(:,[19,20])<0) exponents(:,[19,20]) = 1 ! Setting exponents
act_energy(19) = 163 ! kJ/mol
fre_factor(19) = 3.40E-11 ! cm /(molecules s)
act_energy(20) = 1.79 ! kJ/mol FROM SIMILAR REACTION
fre_factor(20) = 3.21E-30 ! cm /(molecules s) FROM SIMILAR REACTION

! REACTION 21 and 22: CH2O <=> CO + H2
reactions(:,21) = (/&
! O2  H2O  CO2  CH4  CH3  OH  HO2  H2O2  H  H2  HCO  CO  CH2O  CH3O  CH3O2  CH3OOH
NO  NO2
& 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 1  , 0  , 1  , -1  , 0  , 0  , 0  ,
0  , 0  &

```

```

      0 , 0 &
    & /)
  reactions(:,22) = -reactions(:,21)
  ! Both reactions are first order in regard to all reactants:
  where (reactions(:,[21,22])<0) exponents(:,[21,22]) = 1 ! Setting exponents
  act_energy(21) = 146 ! kJ/mol
  fre_factor(21) = 3.49E-09 ! cm /(molecules s)
  act_energy(22) = 43.56 ! kJ/mol FROM SIMILAR REACTION
  fre_factor(22) = 2.20E-10 ! cm /(molecules s) FROM SIMILAR REACTION

  ! REACTION 23 and 24: HCO + O2 <=> CO + HO2
  reactions(:,23) = (/&
    ! O2 H2O CO2 CH4 CH3 OH HO2 H2O2 H H2 HCO CO CH2O CH3O CH3O2 CH3OOH
    NO NO2
    & -1 , 0 , 0 , 0 , 0 , 0 , 1 , 0 , 0 , 0 , -1 , 1 , 0 , 0 , 0 , 0 ,
    0 , 0 &
    & /)
  reactions(:,24) = -reactions(:,23)
  ! Both reactions are first order in regard to all reactants:
  where (reactions(:,[23,24])<0) exponents(:,[23,24]) = 1 ! Setting exponents
  act_energy(23) = 2.04 ! kJ/mol
  fre_factor(23) = 8.60E-12 ! cm /(molecules s)
  act_energy(24) = 75.99 ! kJ/mol FROM SIMILAR REACTION
  fre_factor(24) = 3.60E-13 ! cm /(molecules s) FROM SIMILAR REACTION

  ! REACTION 25 and 26: CO + OH <=> CO2 + H
  reactions(:,25) = (/&
    ! O2 H2O CO2 CH4 CH3 OH HO2 H2O2 H H2 HCO CO CH2O CH3O CH3O2 CH3OOH
    NO NO2
    & 0 , 0 , 1 , 0 , 0 , -1 , 0 , 0 , 1 , 0 , 0 , -1 , 0 , 0 , 0 , 0 ,
    0 , 0 &
    & /)
  reactions(:,26) = -reactions(:,25)
  ! Both reactions are first order in regard to all reactants:
  where (reactions(:,[25,26])<0) exponents(:,[25,26]) = 1 ! Setting exponents
  act_energy(25) = -2.08 ! kJ/mol
  fre_factor(25) = 5.40E-14 ! cm /(molecules s)
  act_energy(26) = 1.11E02 ! kJ/mol
  fre_factor(26) = 2.51E-10 ! cm /(molecules s)

  ! REACTION 27 and 28: H + O2 <=> HO2
  reactions(:,27) = (/&
    ! O2 H2O CO2 CH4 CH3 OH HO2 H2O2 H H2 HCO CO CH2O CH3O CH3O2 CH3OOH
    NO NO2
    & -1 , 0 , 0 , 0 , 0 , 0 , 1 , 0 , -1 , 0 , 0 , 0 , 0 , 0 , 0 , 0 ,
    0 , 0 &
    & /)
  reactions(:,28) = -reactions(:,27)
  ! REACTION SHOULD ACTUALLY BE THIRD ORDER (???) but here I assume
  ! first order in regard to all reactants:
  where (reactions(:,[27,28])<0) exponents(:,[27,28]) = 1 ! Setting exponents
  act_energy(27) = -5.35 ! kJ/mol
  fre_factor(27) = 2.37E-33 ! cm /(molecules s)
  act_energy(28) = 2.03E02 ! kJ/mol
  fre_factor(28) = 2.41E-08 ! cm /(molecules s)

  ! REACTION 29 and 30: 2HO2 <=> H2O2 + O2
  reactions(:,29) = (/&
    ! O2 H2O CO2 CH4 CH3 OH HO2 H2O2 H H2 HCO CO CH2O CH3O CH3O2 CH3OOH
    NO NO2
    & 1 , 0 , 0 , 0 , 0 , 0 , -2 , 1 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 ,
    0 , 0 &
    & /)
  reactions(:,30) = -reactions(:,29)
  ! Reaction is first order in regard to all reactants, but since coefficient is two:
  exponents(:,29) = (/&
    & 0 , 0 , 0 , 0 , 0 , 0 , 2 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 0 ,
    0 , 0 &
    & /)
  exponents(:,30) = (/&
    & 1 , 0 , 0 , 0 , 0 , 0 , 0 , 0 , 1 , 0 , 0 , 0 , 0 , 0 , 0 , 0 ,
    0 , 0 &
    & /)
  act_energy(29) = -4.99 ! kJ/mol
  fre_factor(29) = 2.20E-13 ! cm /(molecules s)
  act_energy(30) = 1.66E02 ! kJ/mol
  fre_factor(30) = 9.00E-11 ! cm /(molecules s)

  ! REACTION 31 and 32: H2O2 <=> 2OH
  reactions(:,31) = (/&

```

```

! O2  H2O  CO2  CH4  CH3  OH  HO2  H2O2  H  H2  HCO  CO  CH2O  CH3O  CH3O2  CH3OOH
& 0  , 0  , 0  , 0  , 0  , 2  , 0  , -1  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  ,
& 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 2  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  ,
& /)
reactions(:,32) = -reactions(:,31)
!! Reaction is second order in total:
exponents(:,31) = (/&
& 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 2  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  ,
& 0  , 0  , 0  , 0  , 0  , 2  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  ,
& /)
exponents(:,32) = (/&
& 0  , 0  , 0  , 0  , 0  , 2  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  ,
& 0  , 0  , 0  , 0  , 0  , 2  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  , 0  ,
& /)
act_energy(31) = 1.90E02 ! kJ/mol
fre_factor(31) = 2.01E-07 ! cm /((molecules s)
act_energy(32) = -1.79E01 ! kJ/mol
fre_factor(32) = 6.71E-33 ! cm /((molecules s)

!-----
! READ PRESSURE, TEMPERATURE AND ABUNDANCES
!-----

! Open file with input temperature and save as temp_layer
! open(unit=100, file='...')
! read(100,*) temp_layer

! Open files with abundances and save as concentrations
open(unit=101, file='./nbm_test/test_abun_network.dat')
read(101,*) concentrations

print *, "Initial concentrations"
print *, concentrations

!-----
! CALCULATE NEW CONCENTRATIONS
! Calculate new species concentrations, using different ODE-solvers
! - Change settings, e.g. temperature
! - Call solver
! - Write to output file
!-----

! Calculate new concentrations, call solver
call calc_conc_euler_back(concentrations, reactions, fre_factor, &
& act_energy, exponents)

end program testdiseqbacknetwork

!-----
! SUBROUTINES
!-----

!-----
! CALC_CONC_EULER
! Solving chemical ODE's using the Forward Euler method.
!
! [N](t) = [N](t-dt) + dt * d[N](t-dt)/dt
! d[N](t)/dt = sum( r * reac )
!
! [N](t) - The concentration of a specific species at time t
! dt - Time step
! r - The reaction rate found using Arrhenius equation
! reac - The reaction matrix (indicating the participation
! and reaction coefficient of each species in each reaction)
!
! Steps in reaction
! Loop over entire time period (time)
! 1) Calculate reaction rates for all reactions using calc_rates
! 2) Find change in concentration for each species, d[N](t-dt)/dt
! 3) Calculate new concentration
! 4) Write concentrations at this specific timestep to file

```



---

```

subroutine calc_conc_euler(concentrations , reactions , fre_factor ,&
& act_energy , exponents)
  use parameters
  implicit none

  ! Input variables
  real :: fre_factor(n_reactions) , act_energy(n_reactions)
  real :: concentrations(n_species)
  integer :: reactions(n_species , n_reactions) , exponents(n_species , n_reactions)
  ! Calculated variables
  real :: rate(n_reactions) , k_const(n_reactions) ! Reaction rate and rate constant
  real :: concentrations_change(n_species) ! Change in concentration

  ! Output file containing the concentrations at each time step
  open(unit=102, file='./nbm_test/new_conc_network.dat')

  do ii_time = 0, nint(time/time_step , 8)

    ! Write concentrations to output file
    if (mod(ii_time,1000)==0) write(102,*) ii_time , ii_time*time_step , &
& concentrations

    ! Calculate rate for all reactions
    call calc_rates(concentrations , fre_factor , &
& act_energy , exponents , reactions , rate , k_const)

    ! Calculate change in concentration for each species
    do ii_spec = 1, n_species
      concentrations_change(ii_spec) = sum(reactions(ii_spec , :)*rate)
    end do

    ! Calculate new concentrations
    concentrations = concentrations + concentrations_change * time_step_s

    ! Locking trace elements to stabilize
    where (concentrations <= 1E-2) concentrations=1E-2

  end do

  301 format(i5 , f8.2 , 4e21.6)
end subroutine calc_conc_euler

```

---

```

! CALC_CONC_EULER_BACK
! Solving chemical ODE's using the Backward Euler method.
!
!  $[N](t) = ([N](t-dt) + P(t-dt)*dt) / (1 + IL(t-dt)*dt)$ 
!  $P(t-dt) = \text{sum}(r * \text{reac}(+))$ 
!  $IL(t-dt) = \text{sum}(r(=/N) * \text{reac}(-))$ 
!
!  $[N](t)$  – The concentration of a specific species at time t
! dt – Time step
! P(t) – Is the production of N during time t
! IL(t) – Is the implicit loss coefficient of n during time t
! r – The reaction rate found using Arrhenius equation
! reac – The reaction matrix (indicating the participation
! and reaction coefficient of each species in each reaction)
! (+),(-) – Only including entries that are positive , negative
! (=N) – Excluding the concentration of N in the calculation of the reaction rate
!
! Steps in reaction
! Loop over entire time period (time)
! 1) Set initial concentration for time step to current concentration
! 2) Calculate rated based on initial concentration
! 3) Loop over species
!   – Find reactions where species is reactant , reac(-)
!     – Calculate implicit loss coefficient
!   – Find reactions where species is product , reac(+)
!     – Calculate production
!   – Calculate new concentration for species
!     (Stored in array separate from initial concentration)

```

---

```

subroutine calc_conc_euler_back(concentrations , reactions , fre_factor ,&

```

```

& act_energy , exponents)
use parameters
implicit none

! Input variables
real :: fre_factor(n_reactions), act_energy(n_reactions)
real :: concentrations(n_species), concentrations_init(n_species)
integer :: reactions(n_species , n_reactions), exponents(n_species , n_reactions)
! Temporary variables used in calculations
integer :: ii_spec2          ! Second species index
real :: loss_product        ! Loss product used to calculate im_loss
! Calculated variables
real :: im_loss, prod       ! Implicit loss coefficient, Production,
real :: rate(n_reactions), k_const(n_reactions)

! Output file containing the concentrations at each time step
open(unit=102, file='./nbm_test/new_conc_backward_network.dat')

do ii_time = 0, nint(time/time_step)

! Write pressures to output file
! The if statements provides extra detail at the timesteps where there is overflow
if (any(concentrations > huge(concentrations))) then
write(102,*) ii_time-1, (ii_time-1)*time_step, concentrations_init
write(102,*) ii_time, ii_time*time_step, concentrations
end if
if (mod(ii_time,1000)==0) write(102,*) ii_time, ii_time*time_step, concentrations

concentrations_init = concentrations ! Set initial concentration for time step

! Calculate rate for all reactions
call calc_rates(concentrations , fre_factor , &
& act_energy , exponents , reactions , rate , k_const)

! Calculate new concentration for each species
do ii_spec = 1, n_species

! Resetting variables
im_loss = 0 ! Implicit loss coefficient
prod = 0    ! Production

! Loop over reactions
do ii_reac = 1, n_reactions

loss_product = 1 ! Used to calculate im_loss

! If species is a reactant, calculate loss
if (reactions(ii_spec , ii_reac) < 0) then
do ii_spec2 = 1, n_species
if (ii_spec2 == ii_spec) cycle ! Exclude species from im_loss
if (reactions(ii_spec2 , ii_reac) < 0) then ! Find other reactants
loss_product = loss_product * concentrations_init(ii_spec2) &
& **exponents(ii_spec2 , ii_reac)
end if
end do
im_loss = im_loss + loss_product * k_const(ii_reac) &
& *abs(reactions(ii_spec , ii_reac)) ! Implicit loss coefficient

! If species is a product, calculate production
else if (reactions(ii_spec , ii_reac) > 0) then
prod = prod + rate(ii_reac) * reactions(ii_spec , ii_reac)
end if
end do
concentrations(ii_spec) = (concentrations_init(ii_spec) + prod * time_step_s) &
& / (1 + im_loss * time_step_s)
end do

! Locking trace elements to stabilize
where (concentrations <= 1E-2) concentrations = 1E-2
end do
end subroutine calc_conc_euler_back

!-----
! CALC_RATES
! Calculating reaction rates based on concentrations ([]), temperature (T),
! activations energies (E), and frequency factors (A), using Arrhenius Law.
! r = k * [A]^n * [B]^m

```

```

!      k = A * exp(-E/(RT))
!
! R - Ideal gas constant = 8.314 J/(K*mol) = 1.38e23 J/(K*molecules)
!
! 1) Calculate the rate constants for all reactions , k
! 2) Calculate rate for each reaction
!   Loop over reactions
!   - Find out which species participate in reaction ii_reac
!   - Calculate product of concentrations , [A]^n * [B]^m
!   - Calculate rate , r
!-----

subroutine calc_rates(concentrations , fre_factor , &
& act_energy , exponents , reactions , rate , k_const)
  use parameters
  implicit none

  ! Input variables
  real :: R_ideal_gas      ! Ideal gas constant
  real :: fre_factor(n_reactions) , act_energy(n_reactions)
  real :: concentrations(n_species)
  integer :: reactions(n_species , n_reactions) , exponents(n_species , n_reactions)
  ! Calculated variables
  real :: k_const(n_reactions) , conc_product , rate(n_reactions)

! CALCULATE RATE CONSTANTS, k_const
  R_ideal_gas = 8.31E-3      ! Ideal gas constant , in kJ/(mol K)
  k_const = fre_factor*exp(-act_energy/(R_ideal_gas*temp_layer)) ! Rate constant

! CALCULATE RATE FOR EACH REACTION, rate

  do ii_reac = 1 , n_reactions
    conc_product = 1

    ! Find which reactants participate in reaction
    ! If entry in reaction matrix is negative, species is a reactant
    do ii_spec = 1 , n_species
      if (reactions(ii_spec , ii_reac) < 0) then
        ! Calculate product of concentrations , with exponents
        conc_product = conc_product*concentrations(ii_spec)**exponents(ii_spec , ii_reac)
      end if
    end do

    ! Calculate rate from rate constant and concentration product
    ! currently with unit molecules/cm /s
    rate(ii_reac) = k_const(ii_reac)*conc_product
  end do

end subroutine calc_rates

```

---

## **B Bach-Møller and Jørgensen, 2020: Eccentricity-multiplicity correlation**

The following 10 pages contain the third (and likely final) revision of the article by Uffe and I as it has been resubmitted to *MNRAS*.

# Orbital Eccentricity - Multiplicity Correlation for Planetary Systems and Comparison to the Solar System

Nanna Bach-Møller,<sup>1</sup>★ Uffe G. Jørgensen,<sup>1</sup>

<sup>1</sup>Centre for Star and Planet Formation, Niels Bohr Institute, University of Copenhagen, Østervoldgade 5, 1350 Copenhagen, Denmark

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

The orbit eccentricities of the Solar System planets are unusually low compared to the average of known exoplanetary systems. A power law correlation has previously been found between the multiplicity of a planetary system and the orbital eccentricities of its components, for systems with multiplicities above two. In this study we investigate the correlation for an expanded data sample, by focusing on planetary systems as units (unlike previous studies that have focused on individual planets). Our full data sample contains 1171 exoplanets, in 895 systems, and the correlation between eccentricity and multiplicity is found to follow a clear power law for all multiplicities above one. We discuss the correlation for several individual subsamples, and find that all samples consistently follow the same basic trend regardless of e.g. planet types and detection methods. We find that the eccentricities of the Solar System fit the general trend and suggest that the Solar System might not show uncommonly low eccentricities (as often speculated) but rather uncommonly many planets compared to a "standard" planetary system. The only outlier from the power law correlation is, consistently in all the samples, the one-planet systems. It has previously been suggested that this may be due to additional unseen exoplanets in the observed one-planet systems. Based on this assumption and the power law correlation, we estimate that the probability of a system having 8 planets or more is of the order of 1%, in good agreement with recent predictions from analyses based on independent arguments.

**Key words:** planetary systems – planets and satellites: general – planets and satellites: fundamental parameters – methods: data analysis – astrobiology

## 1 INTRODUCTION

Extrasolar planets reveal orbital eccentricities much higher than those found among the planets of the Solar System, a deviation that in the beginning was considered so strange that it even led some people to doubt whether the radial velocity exoplanet measurements actually showed real planets. In the present study we will show that the eccentricity of the Solar System planets actually follow the same trend as all other known planetary systems, but belong to the tail of a continuous distribution. When searching for extraterrestrial life we often focus on Earth-like planets and Solar System-like systems, and so low eccentricities are included in our search criteria. But exactly how the habitability of a planet might be affected by the eccentricity of its orbit is yet unknown. A planet on a high-eccentricity orbit can undergo drastic seasonal changes in surface temperature due to the difference in stellar radiation from perihelion to aphelion. These seasonal changes could lead to periods of time without liquid water on the surface, which would greatly limit the habitability of the planet (Bolmont et al. 2016). However, a series of studies (reviewed

in Kopparapu et al. (2019)) have found that often the atmosphere and oceans of a planet can act like a buffer to the temperature variations, in which case the surface climate will be determined by the average stellar radiation rather than the seasonal extremes. In other cases large seasonal variability was found to expand the habitable zone for the planet, by allowing water to remain liquid at larger semi major axes (Linsenmeier et al. 2015). Since it is still uncertain how orbit eccentricities affect the habitability of a planet, it is critical for us to study and understand the eccentricities in the existing exoplanet sample and how they might deviate from those in the Solar System.

From previous investigations (Chatterjee et al. 2008; Ford & Rasio 2008; Jurić & Tremaine 2008b,a; Carrera et al. 2019), planet-planet interaction has been suggested as the dominating mechanism determining orbital eccentricities of planets, either through dynamical relaxation or planet-planet scattering. The dynamical interactions of planetary systems is reviewed in Davies et al. (2014). As a conclusion of this, a correlation between orbital eccentricity and multiplicity (number of planets) is predicted. This prediction has been tested empirically by Limbach & Turner (2015) based on 403 exoplanets detected by the radial velocity method (RV)

★ E-mail: nanna@bachmoeller.dk

and listed in *exoplanets.org*. A strong anti-correlation between eccentricity ( $e$ ) and multiplicity ( $M$ ) was found, and for multiplicities above two the correlation could be described by a power law:  $e(M) \approx 0.584 \cdot M^{-1.20}$ . The eccentricity-multiplicity correlation has later been investigated by [Zinzi & Turrini \(2017\)](#), who found a similar correlation for multiplicities above one based on 258 selected RV and transit planets from NASA Exoplanet Archive. Both of the previous investigations have based their analyses on individual planets rather than treating the systems as units.

The main motivation for this article is to further the investigations by [Limbach & Turner \(2015\)](#) and [Zinzi & Turrini \(2017\)](#) using the expanded planet sample known to date, comparing search methods, population groups, and databases, and aiming to set the results in perspective to our own Solar System and habitability. Our planet sample contains planets found by several detection methods including RV, transiting planet (transit), microlensing (ML) and others. By including all planets, regardless of detection method, we will be able to comment on whether there is an observational bias related to the specific methods, and the large dataset available today makes it possible to exclude more planets that might potentially introduce unwanted bias into the correlation. Unlike the previous investigations we will treat each system as a unit by conducting the analysis based on the average orbital eccentricities in the systems rather than the eccentricity of each individual planet. This is done since both the multiplicity and potential planet-planet interactions are properties of the planetary system as a whole rather than the individual planets.

From the resulting eccentricity-multiplicity correlation an estimate of the mean multiplicity of a planetary system can be obtained in addition to a probability distribution of the multiplicity of planetary systems. From this we wish to set our Solar System in perspective against a "standard" planetary system. We envision that planetesimals are formed in relatively circular orbits, then gravitationally scatter one another into higher eccentricity, before they over longer timescales collide to build up solid planets or the planetary cores of giants. After the evaporation of the gas disk, planet-planet interaction would be the dominating mechanism determining the final eccentricities, in such a way that the more planets there end up being in the system the more circular the orbits become. This is a logic scenario to provide an image of the physical process behind the correlation we investigate in the present paper, but we stress that this is only an image that helps us (and hopefully the reader, too) to imagine the process. Our study is empirical, and hence have no apriori assumption about which exact mechanisms cause the correlation. In order to further the development of the theoretical understanding, we take advantage of the large sample now available to also analyze whether different populations of exoplanets show different correlations.

A major concern when investigating extrasolar planets is that we are highly constrained by limitations in our detection methods. When using RV the detection probability of a planet is biased towards large masses, and when using transit it is biased towards ultra short periods. That leaves a large parameter space where planets go mainly undetected, and thereby bias conclusions about standard planetary systems drawn from the limited sample. Today the two most abundant detection methods (RV and transit) basically have shown us that exoplanetary systems very different from our own Solar System are abundant. Direct observational estimates of how abundant exoplanetary systems resembling our own Solar System are, may most likely come from future extensive microlensing surveys from space (perhaps from a dedicated mi-

cro-lensing satellite ([Bennett & Rhie 2002](#)) or from WFIRST ([Penny et al. 2019](#)) or from the ground (perhaps from GravityCam-like instruments ([Mackay et al. 2018](#))), and they will give us the full set of orbital parameters of solar-system-like exoplanets ([Gaudi 2012; Ryu et al. 2018](#)), as opposed to today where orbital eccentricity has been obtained for only one microlensing exoplanet ([Gaudi et al. 2008](#)). Until then it can be useful to look at indirect evidences for what a standard exoplanetary system looks like. A motivation for this article is to go beyond the data sample by finding a general theory for all systems (including those with planets yet undetected), and from this estimate the characteristics of standard planetary systems. This may give us some insight into the standard formation mechanism of planetary systems and how they develop into the most common configurations of planets, give hints about what to look for and thereby which instruments to develop, and maybe contribute to give us a more realistic view on how abundant truly Earth-like exoplanets might be. One such indirect method is the study of the eccentricity distribution among known exoplanets, as presented here.

In Sect. 2 the dataset is discussed. In Sect. 3 the correlation between eccentricity and multiplicity is examined, both for the full data samples from two different databases, for subsamples sorted for detection methods and for population groups, and for a high-eccentricity subsample in which we attempt to exclude most systems containing undiscovered planets. Based on the correlation a power law is found. In Sect. 4 some of the potential implications of the power law correlation are explored. A probability distribution of the multiplicity is found, and from this a mean multiplicity of planetary systems is estimated. In Sect. 5 the results and theories are discussed. Finally in Sect. 6 the conclusions are summarized.

## 2 THE DATASET

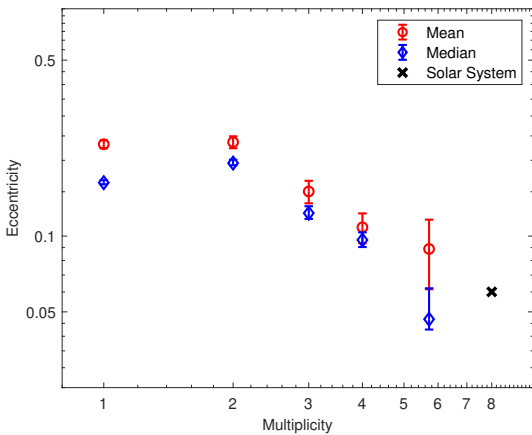
Our data from *exoplanet.eu* were retrieved in August 2019. All confirmed planets regardless of detection method are included. We are aware that *exoplanet.eu*, like most other databases, might be subject to errors in their data listing. For the sake of this study we mostly try not to question the validity of the data found on the website. Planets without listed eccentricities or where the eccentricity is falsely listed as zero (i.e. without listed uncertainties) are excluded from the sample. Of the 4103 planets listed on *exoplanet.eu* a total of 1171 planets remain in the sample, 2932 are excluded due to unknown eccentricities and 60 of these have eccentricities listed as zero with unknown uncertainties. In Table 1<sup>1</sup> the number of planets sorted by multiplicity can be seen for each of the included detection methods.

Because no multiplicities are listed on *exoplanet.eu* each planet has been given a multiplicity based on the number of confirmed planets orbiting the same star listed in the database. Since some of the systems might contain yet undiscovered planets the known companions in these systems will initially be sorted into the wrong multiplicity bins, and the actual distribution might differ from Table 1. Due to the small number of systems with high multiplicities, all systems with more than 5 known planets have been combined in one bin. The multiplicity of this bin is calculated as the mean multiplicity of the included systems. Note that the number of planets in each bin is not necessarily a multiple of the multiplicity. This is caused

<sup>1</sup> All planets and systems with a multiplicity of X will henceforth be referred to as MX-planets or MX-systems

**Table 1.** Planets included in data samples. Retrieved from *exoplanet.eu*. Planets are sorted for detection method. Rightmost column show the number of systems present in each multiplicity bin, whereas columns 2-5 show number of individual planets.

Multiplicity	Total	RV	Transit	Other	Systems
M1	667	408	234	23	667
M2	274	215	52	5	151
M3	121	65	50	6	45
M4	63	43	17	3	20
$\geq M5$	46	34	10	2	12
Total	1171	765	363	39	895



**Figure 1.** Mean and median values of the eccentricity for each multiplicity. The mean eccentricity of the Solar System is plotted with a black  $\times$ . The multiplicity of the  $\geq M5$  multiplicity-bin is plotted as  $M = 5.7$ .

by the fact that not all planets in each system are included, mainly because their eccentricities are unknown. Our dataset is three to four times larger than any of the previous analyses (1171 in this study, compared to 403 in [Limbach & Turner \(2015\)](#) and 258 in [Zinzi & Turrini \(2017\)](#)). We have not accounted for the uncertainties listed for each of the eccentricities in the database in this analysis, which will be discussed further in Sec. 5.

### 3 ECCENTRICITY AND MULTIPLICITY

Each system is assigned an eccentricity found as the mean eccentricity of the planets in the system. This differs from previous studies, where the planets were not sorted into systems, and the authors looked at the eccentricities of the individual planets. The final results from the two methods do not differ greatly, but we find that sorting the planets into systems is more meaningful, since the effects we observe might be caused by planet-planet interactions within the systems and will change the system as a whole. These assigned system eccentricities are then used to calculate overall mean and median eccentricities within each multiplicity bin. In Fig. 1 mean and median values of the system eccentricities are plotted for each of the multiplicity bins, together with our Solar System with a multiplicity of eight.

The errors are calculated using the following methods: Mean; As the standard deviation of system means found by the Bootstrap method. Median; As the one-third and two-thirds quantiles from a Cumulative Distribution Function divided by  $\sqrt{N-1}$ , where  $N$  is the number of systems in the multiplicity bin. Notice that the errors indicate the uncertainties of the mean and median eccentricities of each multiplicity bin, and not the spread of the eccentricities among the individual planets, which is significantly larger than the errors shown. Fig. 1 suggests a trend of decreasing eccentricity for increasing multiplicity. As can be seen the Solar System too follows this trend indicating that our system does not deviate from the norm. An exception for this trend, is the M1 systems. Whereas the other data points seems to approximately follow a power law (seemingly linear because of the logarithmic axes), the eccentricities for M1 deviate from the trend by being too low to follow the power law. This deviation will be discussed later.

#### 3.1 Planet populations

A potential uncertainty related to the study of an eccentricity-multiplicity correlation is the dependence of the correlation on factors such as planet mass and semi major axis. [Turrini et al. \(2020\)](#) and [Laskar & Petit \(2017\)](#) therefore looked at the correlation of multiplicity and angular momentum deficit (AMD), rather than multiplicity and eccentricity. The AMD does depend on the eccentricity, but also on the semi major axis and the mass of the planets, and [Turrini et al. \(2020\)](#) found an anticorrelation between the normalized angular momentum deficit (NAMD) and the multiplicity. [Turrini et al. \(2020\)](#) argues that the eccentricity-multiplicity correlation found by other studies is a reflection of the underlying NAMD-multiplicity correlation. The study of the NAMD-multiplicity is complicated by the fact that few planets have both their masses, eccentricity and semi-major axis well-known, and as such the dataset is smaller. The larger sample in our data set compared to previous data sets, allows us to study directly the correlation of eccentricity and multiplicity for a number of different subsamples, in order to test how the planet mass ( $m_p$ ) and semi major axis (or period,  $P$ ) might affect the eccentricity-multiplicity correlation.

To test the impact of mass and period, we have divided the systems into three different populations: 1) Systems containing a hot-Jupiter ( $m_p > 0.1M_J$  and  $P < 100$  days). 2) Systems containing a cold-Jupiter ( $m_p > 0.1M_J$  and  $P > 100$  days), and no hot-Jupiters. 3) Systems dominated by super-Earths ( $m_p < 0.1M_J$ ) with no giant planets. In order to increase the data sample, planets with no listed mass in the database, have been sorted based on their  $mass \cdot \sin(i)$  value, when this is known, and a total of 849 systems are sorted into the population categories. The distribution of systems in each population category can be seen in Table 2. It should be noted, that the observed planet sample does not represent the true planet population since some planet types are more easily observed than others, but the differences between the populations, as shown here, might still give us an insight into the uncertainties of the eccentricity-multiplicity correlation. Research in the actual occurrence rate of different planet types is reviewed in e.g. [Winn & Fabrycky \(2015\)](#).

Table 2 shows that different multiplicities are dominated by different populations of planets, such that most of the M1 systems are giant-planet systems, whereas the larger multiplicity systems are dominated by super-Earths. A priori one could expect that since the cold-Jupiters dominate the M1 systems, we could seek the explana-



**Table 2.** Distribution of the systems in Table. 1 where in addition to the eccentricity, also the mass or  $msin(i)$  is known, such that they can be divided into the groups: hot-Jupiters (HJ), cold-Jupiters (CJ), super-Earths (SE), and plotted in Fig. 2. Last column shows the number of systems (Number). A total of 849 systems are included.

Multiplicity	HJ	CJ	SE	Number
M1	39.2%	51.3%	9.4%	637
M2	22.9%	57.6%	19.4%	144
M3	25.6%	20.5%	53.8%	39
M4	21.1%	26.3%	52.6%	19
$\geq M5$	10.0%	30.0%	60.0%	10

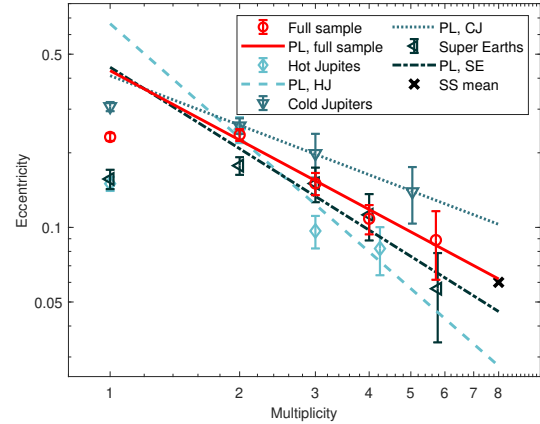
tion for the deviation from the power law followed by the  $M > 1$  systems in the cold-Jupiter population. However, we find that this is not the case, when we look at the mean eccentricities plotted as a function of multiplicity in Fig. 2.

Fig. 2 shows the mean eccentricities plotted for the full sample (equivalent to the mean values from Fig. 1) together with the three different populations introduced above. A power law has been fitted to all samples for multiplicities above one, not including the Solar System, i.e.  $1 < M < 8$ . The power law has been fitted to the overall mean eccentricities for all systems in each multiplicity bin, corresponding to the data points seen in the figure. Due to the small sample of Jupiter-systems with four or more planets, the  $M4$  and  $\geq M5$  bins have been combined for the hot-Jupiter and cold-Jupiter systems. The multiplicity for these bins are the mean multiplicities among the systems combined in the bins. The main conclusion from Fig. 2 is that all three populations follow similar power law trends to the one for the full sample (although of course with larger scatter of the individual points due to the smaller data sample). We notice that the cold-Jupiter population is not the cause of the low eccentricities of the M1 systems, but on the contrary displays the highest eccentricities of the M1 systems among all populations.

### 3.2 The undiscovered planets in the systems

To get further understanding of the uncertainties of the power law correlation, Fig. 3 shows the mean eccentricities plotted as a function of multiplicities for three additional subsamples: Beside the full system sample from *exoplanet.eu*, are shown a high-eccentricity subsample consisting of only the 75% systems with highest eccentricities, a subsample consisting of RV planets listed on *exoplanets.org* before 2014 (L&T) equivalent to the sample used by Limbach & Turner (2015), and a full sample of the 704 planets with known eccentricities from the database *exoplanets.org*. Power laws have been fitted to all samples for multiplicities above one.

The high-eccentricity subsample has been created to exclude systems containing undiscovered planets. According to the trend visible in Fig. 1 larger systems have lower eccentricities, and systems with additional, undiscovered, planets should therefore have eccentricities below what is appropriate for their given multiplicity. We might therefore expect, that the systems showing the lowest orbital eccentricities, could have extra undiscovered planets. Removing these systems from the fit does change the relation a bit (obviously shifting the line to somewhat higher eccentricities), but do keep the same trend of a fine linear fit to the systems with  $M > 1$



**Figure 2.** Mean values of eccentricities for each multiplicity for four subsamples. Full red line: The full sample from *exoplanet.eu* identical to the mean values from Fig. 1. Dashed: Subsample of systems containing a hot-Jupiter (HJ). Dotted: Subsample of systems containing a cold-Jupiter (CJ). Dot-dashed: Subsample of systems only containing smaller planets. Mean value of the Solar System (SS) is plotted in black. Power laws (PL) have been fitted to all four samples for multiplicities above one; this is discussed in Sect. 3.1.

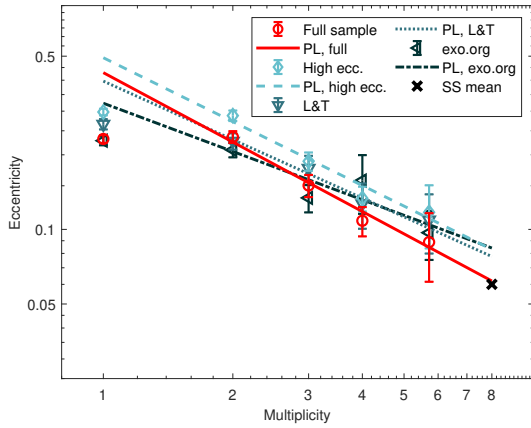
and a substantially lower average eccentricity for the M1 systems than expected from the power law.

Since both of the dominating detection methods (the radial velocity method and the transit method) depend on the size of the planets, smaller planets are more difficult to detect, and only few planets with a size comparable to Mercury or Mars have been found. Mars and Mercury represent one fourth of the (known) planets in the Solar System, and following this line of argument a first attempt of a qualified guess on a typical number of undetected planets could be, that a minimum of 25% of the planets in exoplanet systems remain undiscovered. By removing the 25% systems with the lowest eccentricities in each multiplicity-bin we hope to lower the bias in the correlations by "contamination" due to systems with unknown planets. No systems are removed from the M8 bin, since this only consist of the Solar System. We see from Fig. 3 that the high-multiplicity systems are less affected than the low-multiplicity systems when removing the 25% lowest eccentricity systems, indicating that high-multiplicity systems could be more completely surveyed.

The L&T subsample has been plotted to compare the power law correlation found in this study with one found using a data sample similar to the one used in the original study by Limbach & Turner (2015). Notice that whereas the mean eccentricities for the full, high-eccentricity, and *exoplanets.org* subsamples are found as the mean of the system eccentricities for each multiplicity, the mean eccentricities of the L&T subsample are found as the mean of all planets in each multiplicity-bin (to stay consistent with the analysis methods used by Limbach & Turner (2015) as explained previously).

In order to further constrain potential uncertainties related to our data, we repeated the entire analysis using data from the database *exoplanets.org*. It should be remembered that our main database, *exoplanet.eu*, is more complete and up to date than *exoplanets.org*, but that the planets listed on *exoplanets.org* have





**Figure 3.** Mean values of eccentricities for each multiplicity for four subsamples. Full red line: The full sample from *exoplanet.eu* identical to the mean values from Fig. 1. Dashed: High-eccentricity subsample consisting of 75% systems with highest eccentricities. Dotted: Subsample of RV planets detected before 2014 equivalent to the sample used by *Limbach & Turner (2015)*. Dot-dashed: Full sample from *exoplanets.org*. Mean value of the Solar System (SS) is plotted in black. Power laws (PL) have been fitted to all samples for multiplicities above one; this will be discussed in Sect. 4.

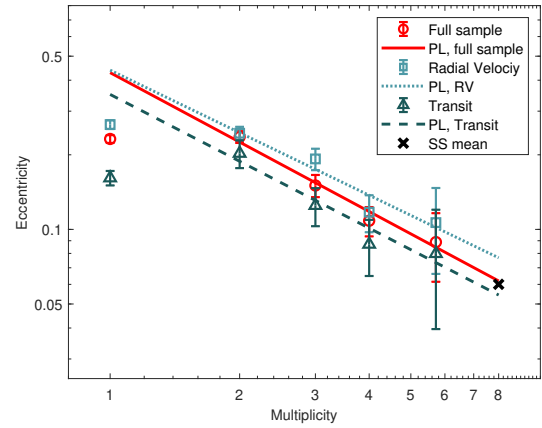
undergone a more strict selection process in regard to peer-review (*Han et al. (2014); Schneider et al. (2011)*), and personal communication with Jason Wright and Françoise Roques). Although the two databases therefore will not contain the exact same data sample, comparison of the results based on both databases gives more clear impression of the uncertainties.

Fig. 3 shows that all the subsamples, display the same general tendency of a power law correlation between eccentricity and multiplicity for  $M > 1$  as the full sample, and a lower eccentricity of the M1 systems not following the power law trend of the higher multiplicity systems. The slopes, however, vary for the different samples.

### 3.3 Detection methods

Whereas the L&T subsample consists only of RV planets our sample contains planets found by all detection methods. To test how this difference might affect the eccentricity-multiplicity correlation, and to better understand whether the behaviour of the correlation could be dominated by a bias effect related to the detection method, a plot for the transit and RV subsamples together with the full sample can be seen in Fig. 4. It should be noted that the eccentricities listed for planets discovered with the transit method are often determined from followup RV observations, so the two populations are not completely separated. Fig. 4 shows that both the transit and the RV subsamples have eccentricity-multiplicity correlations similar to that of the full sample, and the trend of the M1 systems falling below the  $M > 1$  relation is identical.

We also see that the transit systems show lower eccentricities at all multiplicities compared to the RV systems. This bias, that transit planets generally have lower eccentricities, is in correspondence with a study by *Van Eylen & Albrecht (2015)* who found high-multiplicity Kepler planets to generally have lower eccentricities than the RV planet sample. This tendency might be caused by the



**Figure 4.** Mean values of eccentricities for each multiplicity for three subsamples. Full red line: The full sample. Dashed: Subsample consisting of planets discovered by the transit method. Dotted: Subsample consisting of planets discovered by RV. Mean value of the Solar System (SS) is plotted in black. Power laws (PL) have been fitted to all samples for multiplicities above one.

bias, that there are more low-mass planets in the transit subsample than in the RV sample, and that lower mass planets are more easily circularised by planet-planet interaction (*Kane et al. 2012*). We see a hint of the same tendency in Fig. 2 where the super-Earth subsample shows lower eccentricities than the full sample, and the important conclusion is that independent of the shift and its potential explanation in an observational bias, the same tendencies discussed above applies to both of the subsamples.

It is also possible that planet-planet scattering could cause a spread in the orbital inclinations (*Chatterjee et al. 2008*) in addition to lowering the multiplicity of the system. The spread in inclination could lead to a higher number of undiscovered planets in the transit systems and thereby a higher number systems with eccentricities too low to fit their assigned multiplicity. This trend would be strongest for low-multiplicity systems, as seen in Fig. 4, if these are formed due to severe planet-planet scattering. It can be seen from the errorbars given in Fig. 4 that the listed eccentricities of the transit planets have a greater variation than the RV planets, possibly caused by a larger uncertainty in their determination (*Kane et al. 2012; Van Eylen & Albrecht 2015*).

### 3.4 Kolmogorov-Smirnov test

To statistically test the correlation between multiplicity and eccentricity, a two-sample Kolmogorov-Smirnov is conducted on the full system sample. The test compares the multiplicity-bins one and one to test the difference in the eccentricity distributions of the systems. The test results can be seen in Table 3. Notice that the distribution of eccentricities for the individual *planets* is used for the Solar System, whereas the distributions of the *systems* are used for the rest.

It can be seen that the eccentricities of most of the multiplicity-combinations show significant differences, on a 5% significance level. This indicates that the difference in eccentricity for systems of different multiplicity is caused by a connection between the two factors and not by coincidence. The higher p-values seen for high-multiplicity combinations might be caused by the small num-

**Table 3.** Test result for Kolmogorov-Smirnov test.

	<i>M</i> 1	<i>M</i> 2	<i>M</i> 3	<i>M</i> 4	$\geq M$ 5	<i>M</i> 8(SS)
<i>M</i> 1	1					
<i>M</i> 2	< 0.01	1				
<i>M</i> 3	0.04	0.01	1			
<i>M</i> 4	0.01	< 0.01	0.15	1		
$\geq M$ 5	0.04	< 0.01	0.06	0.31	1	
<i>M</i> 8 (SS)	0.01	< 0.01	0.05	0.38	0.65	1

ber of systems in these multiplicity-bins. Altogether the statistical test supports, that there is a correlation between multiplicity and eccentricity.

### 3.5 Quantification of the multiplicity-eccentricity correlation

In the standard core-accretion model for the formation of planetary systems, the dust component of the disk relatively quickly clumps together (via simple condensation or even faster via streaming instability) to form many objects of planetesimal sizes (Johansen & Lambrechts 2017). Over a longer timescale the planetesimals then excite one another's orbits by gravitational interaction, leading to collisions and hence growth to planet size. After the dissipation of the protoplanetary disk the orbits of the planets are largely determined by planet-planet interactions, indicating a correlation between the orbital eccentricity and the number of interactions and hence planets. The numerical simulations by Chatterjee et al. (2008) and Jurić & Tremaine (2008a) confirms that this expectation is correct, by showing that the final architecture of a system is almost independent of the assumed starting conditions of planetesimals, and suggesting that planet-planet interaction is the dominating mechanism for changing the average orbital eccentricity. The simulations do not in themselves predict a specific analytical correspondence between eccentricity and multiplicity, which, however, can be done by fitting the corresponding observational data. In Fig. 3 it was indicated that the high-multiplicity systems seemed to have fewer undiscovered planets, and in both Fig. 2, 3 and 4 we quantified the relation by fitting the mean eccentricities for  $M > 1$  to a power law. Our best fit to the full set of data (as shown in red in the figures) can be expressed as:

$$e(M) = 0.429 \cdot M^{-0.93} \quad (1)$$

where  $e$  is the eccentricity and  $M$  is the multiplicity. Fig. 2-3 and 4 further demonstrates that this fit also agrees with the Solar System despite the fact that the  $M = 8$  was not included in the fit. This adds extra confidence in believing that the quantification is universal, and two fits, with and without the Solar System, showed the following correlation coefficient;  $R^2 = 0.98$  for  $M = [2; 7]$  and  $R^2 = 0.99$  for  $M = [2; 8]$ .

Since the physical cause behind the relation is thought to be planet-planet gravitational interaction, one should expect the decreasing tendency to range all the way from *M*1 systems to a maximum number of planets,  $M_{\max}$ , for which the systems can still remain stable, (Papaloizou & Terquem 2001; Jurić & Tremaine 2008b), with the *M*1 systems having the largest average eccentricity. Observationally, the *M*1 planets, obviously, do not show the high eccentricity expected from the correlation, and therefore the observed

*M*1 systems must be affected differently from the multi-planet populations. In the following section, Sect. 4, we will elaborate on one potential explanation for the deviation of the *M*1 systems from the trend, namely the idea that the low *M*1 eccentricity is caused by a combination of mechanisms other than the general planet-planet interaction, lowering the eccentricities, plus an observational bias. When correcting for these two effects, the remaining *M*1 systems are made to follow the same trend as the rest of the systems, and potential implications for the trend are explored.

An alternative explanation for the discrepancy between the *M*1 and multi-planet systems could be that they are dominated by different planet populations. To analyze if any specific population dominates the lowering of the *M*1 eccentricities, we investigated, in Sect. 3.1, whether the population of large planets (which observationally dominates the *M*1 and *M*2 systems) and the population of smaller planets (that have a more dominating role in the higher multiplicity systems), show different observational trends. We concluded that all of the populations follow the same general trend between eccentricity and multiplicity, indicating that the same general mechanism is responsible for all the observed populations of exoplanets from *M*1 to *M*8 (and is likely to be planet-planet interaction with some correction for the *M*1 systems).

In all cases, it is obvious from Fig. 1-4 that the observed *M*1 systems do not follow the trend expressed in Eq. 1. If a reasonable transformation from the observed abundance of *M*1 systems to intrinsic *M*1 system abundances can be obtained, it will be possible from Eq. 1 to give an estimate of the true probability distribution of multiplicities among the observed systems.

## 4 PERSPECTIVE AND IMPLICATIONS: CONVERSION OF OBSERVED MULTIPLICITY DISTRIBUTION TO ACTUAL DISTRIBUTION

Fig. 1-4 demonstrates that the observed average eccentricity of one-planet systems (*M*1) falls below the relation for multi-planet systems. The main assumption in this further analysis is that the *M*1 systems intrinsically follow the same eccentricity correlation as the other multiplicities. This assumption is supported by a series of studies by He et al. (2019, 2020), who recreated the multiplicity distribution of the Kepler observations, by forward-modelling multi-planet systems at the AMD-stability limit (introduced in Laskar & Petit (2017); Petit et al. (2017)). He et al. (2020), found that all multiplicities from one to ten followed the same eccentricity-multiplicity power law correlation, with the intrinsic *M*1 systems having higher eccentricities than the multi-planet systems, and they found that most observed *M*1 systems contain yet undiscovered planets. In this section will try to identify these systems with undiscovered planets, and redistribute them to the multiplicity bin appropriate to their multiplicities.

We will first investigate whether some of the low eccentricity *M*1 planets can have got their low eccentricity due to other mechanisms than the general planet-planet interaction assumed to be responsible for Eq. 1.

Exoplanets in ultra small orbits are often tidally locked to the host star, which could lead to circularisation of the planetary orbit (Jackson et al. 2008). By looking at the eccentricity damping timescale (Ogilvie 2014), the eccentricity damping from these planet-star interactions can be approximated by:

$$\dot{e} \propto \frac{m_*}{m_p} \frac{1}{a^5} \quad (2)$$

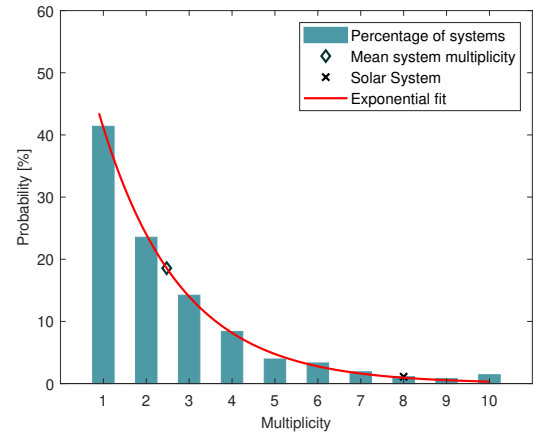
where  $\dot{e}$  is the change in eccentricity,  $a$  is the semi major axis of the planet, and  $m_p$  and  $m_*$  are the masses of the planet and the star respectively.

In order to distinguish systems that have low eccentricities due to planet-star interactions from those that may have low eccentricities for other reasons, all planets for which the value from Eq. 2 exceeds a certain threshold are excluded. The threshold was chosen to be  $6.77 \times 10^5$ , and 191 M1 planets, and 100 planets among the other multiplicities, were excluded on this basis. These planets will be excluded in the following probability analysis, but were not excluded in the making of Eq. 1 (which would have very small effect as described below). The chosen threshold is the value of Mercury, and even though Mercury is far from being circularised (it holds the highest eccentricity in the Solar System), it is "almost" tidally locked (in a 2/3 orbital/rotational resonance), and is the planet in the Solar System that has the highest potential for tidal circularisation. In an analysis of hot-Jupiters with known obliquities, Hjort (2019) was able to divide the planets into two distinct groups, with 15% of the planets having extremely low obliquity (and hence low eccentricity) and 85% having a continuous obliquity distribution. Hjort (2019) ascribed the former group to planet migration in the disk and the latter to migration due to planet-planet interaction (scattering). It is therefore likely that also a fraction of the M1 systems will have much lower eccentricities than expected from Eq. 1 due to disk-migration.

Next, we pursue the idea that some of the remaining systems may contain yet undiscovered planets, and that these systems will lower the mean eccentricity of their multiplicity bins, since systems with more planets are expected to have lower eccentricities. Those of the observed systems that have had their eccentricity determined by planet-planet interactions (as opposed to the systems excluded above due to a potential star-planet circularisation) are to first approximation expected to follow the planet-planet eccentricity relation expressed in Eq. 1. We align the mean eccentricities of the multiplicity bins with the power law correlation, by moving the lowest eccentricity systems of the multiplicity bins to an  $M$  corresponding to their observed eccentricity (i.e. assuming undiscovered planets in those systems). During this exercise it was found, that the best alignment occurred when 55% of the M1 systems and all of the  $M > 1$  systems were assumed *not* to contain undiscovered planets, and the rest had new multiplicities estimated based on their eccentricity.

Of the M1 systems 50 (i.e. roughly 10% of the  $667 - 191 = 476$  M1 systems that remained after the exclusion of planets that might have experienced planet-star circularisation) have such low eccentricities that they should be moved to multiplicities that might exceed  $M_{\max}$  (in some cases more than 50 planets). It was therefore assumed that these systems might not contain undiscovered planets, but that other physical mechanisms were responsible for circularizing these 10%. For the proceeding estimates, have in mind that the effect of keeping these 50 planets would be to slightly increase the estimated abundance of M1 systems and decrease correspondingly the abundance of high multiplicity systems like our own solar system. Non-planet-planet interacting mechanisms that could be responsible for circularization of a fraction of this amount of M1 systems could include migration of a single large planet to small orbit while substantial amount of the protoplanetary disk was still in place (Hjort (2019)).

For those of the remaining group of ( $667 - 191 - 50 = 426$ ) M1 systems with eccentricities that potentially could be attributed to yet undiscovered planets, we attempted a redistribution of the systems



**Figure 5.** Percentage of systems with given multiplicity, corresponding to values from Table 4. Probability function found as exponential fit. Mean multiplicity estimated to  $\sim 2.5$ .

by artificially counting them as belonging to higher values of  $M$ . The new multiplicity,  $M_{\text{new}}$ , was determined from the eccentricity of the planet using Eq. 1. A total of 164 M1 systems were redistributed and the new multiplicity distribution can be seen in Table 4.

In addition to the number of systems within each multiplicity bin, Table 4 also shows the percentage- and probability distributions for the redistributed planets. The probability distribution is found by fitting an exponential fit to the percentage distribution as shown in Fig. 2 and explained later.

The redistribution of the M1 systems has been made such that the mean eccentricity of the remaining 262 systems falls on the same relation as the rest of the multiplicity systems described by Eq. 1. For the sake of this experiment, we assume that these remaining M1 systems would be the intrinsic M1 population among the observed systems, with no additional undiscovered planets and whose eccentricity is determined by the same planet-planet interactions as the multi-planet systems. In this sense one can think of the relation given by Eq. 1 applied to all the systems from M1 to  $M_{\max}$  as giving a minimum abundance of M1 systems and corresponding maximum abundance of high multiplicity systems. We stress this fact because it for many might seem intuitively (for example based on the antropoc principle) surprising that our solar system belongs to such a relatively rare type of planetary systems as predicted from Eq. 1 and shown in Fig. 5; without the redistribution suggested above, the Solar System would be predicted to be of an even more rare type of planetary system.

We therefore suggest that the  $M_{\text{new}}$  distribution in Table 3 is a reasonable first qualified guess of the relative distribution of the number of planets in planetary systems, whose average eccentricity distribution is determined by planet-planet interactions. This probability distribution is shown in Fig. 5 and has been fitted to an exponential function described as:

$$P(M) = 0.72 \cdot e^{-0.54M} \quad (3)$$

Where  $P(M)$  indicates the probability of a system having  $M$  planets. This relation has been found by normalizing the exponential fit seen in Fig. 5, such that  $\sum_{M=1}^{10} P(M) = 1$ .

The average number of planets in planetary systems, according to the distribution in Table 4, is  $\langle M \rangle = 2.48$ , and is marked by a diamond in the figure. Based on the discrete probability distribution

**Table 4.** Redistribution of systems. Left; the observed multiplicity distribution of systems from *exoplanet.eu*. Right; the multiplicity distribution of systems after the M1 systems have been redistributed according to their eccentricities as described in the text. The rightmost column indicates the probability of a system having a given multiplicity according to Eq. 3

<i>Multiplicity</i>	<i>Observed distribution</i>		<i>Redistribution</i>		
	<i>Number of Systems</i>	<i>Percentage</i>	<i>Number of Systems</i>	<i>Percentage</i>	<i>Probability</i>
<i>M1</i>	667	75%	262	41%	41%
<i>M2</i>	151	17%	149	24%	24%
<i>M3</i>	45	5%	90	14%	14%
<i>M4</i>	20	2%	53	8%	8%
<i>M5</i>	4	<1%	25	4%	5%
<i>M6</i>	6	<1%	21	3%	3%
<i>M7</i>	1	<1%	12	2%	2%
<i>M8</i>	1	<1%	7	1%	1%
<i>M9</i>			5	<1%	<1%
<i>M10</i>			9	2%	<1%
<i>Total</i>	895		633		

in Eq. 3 the probability of a system having eight planets is  $P(8) \approx 1\%$ , indicating that systems the size of the Solar System are rare but not exceptionally so. In this interpretation the Solar System is in the tail of a distribution of multiplicity, and corresponding orbital eccentricities, near the maximum possible from stability considerations. We have in the above summation assumed that the maximum cut off is near to  $M_{\max} = 10$  planets, but we note that the exact value of  $M_{\max}$  is unimportant for the conclusion, since the integral of Eq. 3 from 8 to infinity is very small. Remark also that the number 1% refers to the fraction of the systems that have had their eccentricities determined by planet-planet interaction, or a similar process that is responsible for Eq. 1 and Eq. 3. If one is counting also the M1 planets that were excluded in deriving Eq.3, then the probability of finding 8 planets would be slightly lower. It should be noted that all results in this analysis rely on the assumption that the power law in Eq. 1 describes the true intrinsic correlation between eccentricity and multiplicity. The redistribution was based on a correlation fitted to the observed multi-planet systems. The fact that some of these multi-planet systems might host yet undiscovered planets could therefore pose a uncertainty to the analysis. However, theoretical studies have found that the observed M1 population is the only one that differ greatly from the theoretical predictions (Johansen et al. 2012). As mentioned previously, some have suggested that this is caused by the fact that the observed M1 systems are especially prone to containing undiscovered planets (He et al. 2020). As such, the analysis should not be affected greatly by undiscovered planets in the multi-planet systems. As mentioned previously planet-star interaction has not been taken into account when making Eq. 1. If the planets from the multi-planet systems that might have experienced planet-star interaction had been excluded in Eq. 1, the mean multiplicity would have been 2.6 rather than 2.5 planets per system.

It is encouraging to note that Jurić & Tremaine (2008b) found the average number of planets to be between 1.8 and 3.0 planets per system from a series of individual simulations with different initial planetesimal conditions, and that Raymond et al. (2018) found the probability of forming planetary systems with a number of planets similar to our own to be  $\sim 1\%$  based on dynamical arguments. Both results are very similar to our result but based on completely different and independent arguments.

## 5 DISCUSSION

We find an anti-correlation between orbital eccentricity and multiplicity of known exoplanet systems similar to the reports by previous studies (Limbach & Turner 2015; Zinzi & Turrini 2017). Our planet sample and method differ from the investigation by Limbach & Turner (2015) by including planets discovered by all detection methods, not just RV, and from both studies by including a much larger dataset and by comparing the results obtained based on different databases with different selection criteria. In addition we have chosen to consider systems as units unlike the previous studies, that treated each planet separately. When comparing our investigation to the previous ones, it should be noted that we, of course, share a great part of our data sample, and although the larger dataset in our analysis has allowed for a more restrictive debias process, all our analyses are biased by the basic limitation by the RV technique (biased towards larger planets), and the transit technique (biased towards ultra small orbits).

The fact that we include all planets regardless of detection methods has shown us that similar eccentricity-multiplicity correlations can be found for the full sample, and RV- and transit subsamples respectively, though with slightly different fits as discussed above. Explicitly, we also studied the eccentricity-multiplicity correlation for subsamples of hot-Jupiters, cold-Jupiters, and super-Earths separately, and found that also these subsamples followed the same general tendency. This shows that the correlation is not solely caused by the giant planets, that currently dominate our observations, or by planets at very short periods, but might also apply to the population of planets that have yet to be observed, with smaller masses and at larger periods.

A correlation between orbit eccentricity and multiplicity is supported by several other studies. Surveys conducted by Howard (2013) and Wright et al. (2009) found lower orbit eccentricities among planets in multi-planet systems compared to single planets. Howard (2013) suggests that the trend could be due to severe planet-planet scattering in the existing single-planet systems where giant planets have excited the eccentricities of its previous companions before ejecting them. Multi-planet systems have had fewer scattering events (otherwise they would no longer be multi-planet), and have thereby been allowed to stay in stable low-eccentricity orbits. Wright et al. (2009) argues that multi-planet systems will



naturally favour low-eccentricity orbits because of the need for high orbital stability in the system. The stability of multi-planet systems was studied further by [Huang et al. \(2017\)](#), who found that a single outer high-eccentricity giant planet would greatly affect the stability of an inner system, by reducing the multiplicity and exciting the eccentricities of the remaining planets. Both [Wright et al. \(2009\)](#) and [Chatterjee et al. \(2008\)](#) support the theory of single high-eccentric planets as a result of ejecting companions. The ejection of planets from planetary systems have been confirmed by [Mróz et al. \(2017\)](#), who from analysis of short duration events in 6 years of microlensing data have found free floating planets of both Jupiter- and Earth-size, although they also conclude that the abundance of free floating planets is small, and can therefore only account for the eccentricity of a small fraction of the M1 systems. A study by [Xie et al. \(2016\)](#) have also reported lower eccentricities in multi-planet systems. This study measured the eccentricity distribution of a sample of transit planets using transit duration statistics, and found that single-planets in general show eccentricities of  $e \approx 0.3$ , whereas their multi-planet counterparts have average eccentricities of  $e \approx 0.04$ . [Xie et al. \(2016\)](#) found all planets from multi-planet systems to follow similar eccentricity distributions, and so, found no general correlation between eccentricity and multiplicity. Several studies have suggested that the correlation between eccentricity and multiplicity originate in an underlying correlation between multiplicity on the stability of the system, or the angular momentum deficit (AMD) ([Laskar & Petit 2017](#); [Turrini et al. 2020](#); [He et al. 2020](#)). In their study, [He et al. \(2020\)](#) recreate the multiplicity distribution observed at the Kepler data, using a forward model, by looking at the AMD-stability limit. They find that the median eccentricities as a function of multiplicity follow a power law correlation for all multiplicities from one to ten. Their model predicts that intrinsic single-planet systems have higher eccentricities than multi-planet systems, whereas most observed single-planet systems contain yet undiscovered planets, similar to our assumptions in Sec. 4. Like previous studies [He et al. \(2020\)](#) argued that the correlation between intrinsic multiplicity and the eccentricity of the systems was caused by the fact that the AMD-stability criteria puts strong demands on the total system AMD and minimum system period ratio, in order for no planet orbits to cross, and thereby destabilizing the system.

The eccentricity-multiplicity anti-correlation is opposed by [Bryan et al. \(2016\)](#) and [Dong et al. \(2014\)](#), who found lower eccentricities among single planets compared to planets with outer companions. Both surveys mainly focus on jovian planets, [Dong et al.](#) solely on warm-Jupiters with jovian companions. [Dong et al. \(2014\)](#) suggest that their results indicate that planet-planet interactions are not the dominating mechanism for creating short-period jovian planets, as opposed to the suggestions by several other studies ([Rasio & Ford 1996](#); [Marzari & Weidenschilling 2002](#); [Chatterjee et al. 2008](#); [Nagasawa et al. 2008](#)).

As argued by [Bryan et al. \(2016\)](#) a significant uncertainty is involved with the investigation by [Limbach & Turner \(2015\)](#) and some of this apply to our study as well. Many planets included have small semi major axes (the majority within 1 AU), and the low eccentricities found in high-multiplicity systems might reflect the fact that systems this closely packed would not be able to remain stable at higher eccentricities. With our larger data sample we have found similar correlations between RV- and transit subsamples which lowers the probability that the correlation is caused by observational biases. [Bryan et al. \(2016\)](#) further emphasize the uncertainty related to the fact that [Limbach & Turner \(2015\)](#) do not account for the in-

dividual errors of each of the listed eccentricities, which could also pose an uncertainty for this study.

Since we have not included the listed uncertainties to the eccentricities of each individual planet, we have not accounted for the uncertainty involved with the estimation of orbit eccentricity of exoplanets. In addition to this, previous studies have found that many eccentricities are systematically overestimated ([Shen & Turner 2008](#)), and that some seemingly high-eccentricity single planets can turn out to have an unknown outer companion that artificially increase their estimated eccentricity ([Fischer et al. 2001](#)). The latter, fits our eccentricity-multiplicity correlation, with a decrease in eccentricity for an increasing number of known planets, it does however represent an uncertainty to our calculated model.

Unlike this study, the study by [Zinzi & Turrini \(2017\)](#) did account for the uncertainties related to the eccentricity measurements. When calculating the mean eccentricities, [Zinzi & Turrini \(2017\)](#) weighted their data with one over the listed uncertainties, which resulted in a steeper curve for the power law correlation compared to unweighted data. Especially the M2 systems seemed to differ between the weighted and unweighted data by having a significantly higher mean eccentricity in the weighted sample. They did not give an explanation as to why the low-eccentricity M2 planets should have generally higher uncertainties. In this study we find that the M2 systems have eccentricities that fit the general eccentricity multiplicity correlation for  $M > 1$  without correction for the uncertainties. In our analysis only the M1 systems falls substantially below the power law fit, but since no M1 systems were included in the analysis by [Zinzi & Turrini \(2017\)](#) we are not able to compare this trend to their results.

## 6 CONCLUSION

During this study we have investigated the correlation between orbital eccentricity and multiplicity for 1171 planets distributed in 895 systems listed in the database *exoplanet.eu*. We found a strong correlation between average eccentricity and multiplicity for all systems with two or more planets, which could be expressed as  $e(M) = 0.429 \cdot M^{-0.93}$  (Eq. 1). The Solar System fits this trend, without being included in the making of the power law, whereas the average eccentricity of the observed M1 systems were markedly lower than predicted from Eq. 1. It is not unexpected from standard core accretion theory that the M2 to  $M_{\max}$  systems fit the same power law distribution, but it is surprising that the M1 systems fall substantially below the correlation. The eccentricity-multiplicity correlation is investigated for at number of different subsamples, in order to explore the stability of the power law correlation, and investigate possible explanations for the deviating M1 average eccentricity. All subsamples show the same general pattern, with all multiplicities fitting a power law correlation well, except the M1 systems having consistently lower eccentricities. The analyzed subsamples include: different planet populations (divided into hot-Jupiter-, cold-Jupiter-, and super-Earth systems), planets detected by the RV- or transit method respectively, etc.

In order to investigate some of the implications of the power law trend, we speculated on the potential consequences, if the trend that was found for  $M > 1$ , in reality applies to all multiplicities. Following the idea that Eq. 1 describes the true eccentricity-multiplicity correlation, we assumed that the seemingly low eccentricities of the M1 systems were caused by a combination of some systems having been circularized through planet-star interactions, and others containing yet undiscovered planets. Correcting for these assumptions,

a probability distribution over the different multiplicities was expressed by Eq. 3, and based on this the mean multiplicity among the observed systems was estimated to  $\langle M \rangle \approx 2.5$ , while the probability of a system having eight planets was  $\sim 1\%$ .

It is not surprising that the probability of finding high-multiplicity systems comes out this low, after all there are very few known exoplanetary systems with more than 6 planets, but it is assuring that the average number of planets in a "standard" exoplanet system in our Galaxy comes out very close to the number predicted independently from numerical simulations of planetesimal collisions (Jurić & Tremaine (2008b)) and that the probability of finding Solar System like multi-planet systems comes out close to recent independent predictions from dynamical simulations (Raymond et al. 2018). This indicates that the orbit eccentricities of the Solar System planets are not unusually low, when the multiplicity of the system is taking into account, but rather that the number of planets in our Solar System is unusually high. The rarity of the large number of planets in our Solar System, and the corresponding low value of the orbital eccentricities, raise the simple and central, but speculative, question "Is there a connection between the high number of planets in our Solar System and the fact that we are here?".

#### ACKNOWLEDGMENTS

This research has made use of The Extrasolar Planets Encyclopaedia at *exoplanet.eu* and the Exoplanet Orbit Database and the Exoplanet Data Explorer at *exoplanets.org*. We are thankful for clarifying discussions with F. Roques about the selection criteria used by *exoplanet.eu* and with J. Wright about the selection criteria used by *exoplanets.org*. We are grateful to an anonymous referee, whose valuable input improved the analyses and argumentation throughout the paper.

#### DATA AVAILABILITY

The data underlying this article are available in The Extrasolar Planets Encyclopaedia, at <http://exoplanet.eu/catalog/>.

#### REFERENCES

Bennett D. P., Rhie S. H., 2002, *The Astrophysical Journal*, 574, 985  
 Bolmont E., Libert A.-S., Leconte J., Selsis F., 2016, *Astronomy & Astrophysics*, 591, A106  
 Bryan M. L., et al., 2016, *ApJ*, 821, 89  
 Carrera D., Raymond S. N., Davies M. B., 2019, *A&A*, 629, L7  
 Chatterjee S., Ford E. B., Matsumura S., Rasio F. A., 2008, *ApJ*, 686, 580  
 Davies M. B., Adams F. C., Armitage P., Chambers J., Ford E., Morbidelli A., Raymond S. N., Veras D., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, *Protostars and Planets VI*. p. 787 ([arXiv:1311.6816](https://arxiv.org/abs/1311.6816)), doi:10.2458/azu\_uapress\_9780816531240-ch034  
 Dong S., Katz B., Socrates A., 2014, *ApJ*, 781, L5  
 Fischer D. A., Marcy G. W., Butler R. P., Vogt S. S., Frink S., Apps K., 2001, *ApJ*, 551, 1107  
 Ford E. B., Rasio F. A., 2008, *ApJ*, 686, 621  
 Gaudi B. S., 2012, *Annual Review of Astronomy and Astrophysics*, 50, 411  
 Gaudi B. S., et al., 2008, *Science*, 319, 927  
 Han E., Wang S. X., Wright J. T., Feng Y. K., Zhao M., Fakhouri O., Brown J. I., Hancock C., 2014, *PASP*, 126, 827

He M. Y., Ford E. B., Ragozzine D., 2019, *Monthly Notices of the Royal Astronomical Society*, 490, 4575  
 He M. Y., Ford E. B., Ragozzine D., Carrera D., 2020, arXiv draft manuscript arXiv:2007.14473  
 Hjort M., 2019, PhD thesis, Department of Physics and Astronomy, Aarhus University, Denmark  
 Howard A. W., 2013, *Science*, 340, 572  
 Huang C. X., Petrovich C., Deibert E., 2017, *AJ*, 153, 210  
 Jackson B., Greenberg R., Barnes R., 2008, *ApJ*, 678, 1396  
 Johansen A., Lambrechts M., 2017, *Annual Review of Earth and Planetary Sciences*, 45, 359  
 Johansen A., Davies M. B., Church R. P., Holmelin V., 2012, *The Astrophysical Journal*, 758, 39  
 Jurić M., Tremaine S., 2008a, in Fischer D., Rasio F. A., Thorsett S. E., Wolszczan A., eds, *Astronomical Society of the Pacific Conference Series Vol. 398, Extreme Solar Systems*. p. 295  
 Jurić M., Tremaine S., 2008b, *ApJ*, 686, 603  
 Kane S. R., Ciardi D. R., Gelino D. M., von Braun K., 2012, *MNRAS*, 425, 757  
 Kopparapu R. k., Wolf E. T., Meadows V. S., 2019, arXiv e-prints, p. arXiv:1911.04441  
 Laskar J., Petit A., 2017, *Astronomy & Astrophysics*, 605, A72  
 Limbach M. A., Turner E. L., 2015, *Proceedings of the National Academy of Science*, 112, 20  
 Linsenmeier M., Pascale S., Lucarini V., 2015, *Planetary and Space Science*, 105, 43  
 Mackay C., et al., 2018, *Publications of the Astronomical Society of Australia*, 35  
 Marzari F., Weidenschilling S. J., 2002, *Icarus*, 156, 570  
 Mróz P., et al., 2017, *Nature*, 548, 183  
 Nagasawa M., Ida S., Bessho T., 2008, *ApJ*, 678, 498  
 Ogilvie G. I., 2014, *ARA&A*, 52, 171  
 Papaloizou J. C. B., Terquem C., 2001, *MNRAS*, 325, 221  
 Penny M. T., Gaudi B. S., Kerins E., Rattenbury N. J., Mao S., Robin A. C., Novati S. C., 2019, *The Astrophysical Journal Supplement Series*, 241, 3  
 Petit A. C., Laskar J., Boué G., 2017, *Astronomy & Astrophysics*, 607, A35  
 Rasio F. A., Ford E. B., 1996, *Science*, 274, 954  
 Raymond S. N., Izidoro A., Morbidelli A., 2018, *Solar System Formation in the Context of Extra-Solar Planets* ([arXiv:1812.01033](https://arxiv.org/abs/1812.01033))  
 Ryu Y. H., et al., 2018, *AJ*, 155, 40  
 Schneider J., Dedieu C., Le Sidaner P., Savalle R., Zolotukhin I., 2011, *Astronomy & Astrophysics*, 532, A79  
 Shen Y., Turner E. L., 2008, *ApJ*, 685, 553  
 Turrini D., Zinzi A., Belinchon J., 2020, *Astronomy & Astrophysics*, 636, A53  
 Van Eylen V., Albrecht S., 2015, *ApJ*, 808, 126  
 Winn J. N., Fabrycky D. C., 2015, *Annual Review of Astronomy and Astrophysics*, 53, 409  
 Wright J. T., Upadhyay S., Marcy G. W., Fischer D. A., Ford E. B., Johnson J. A., 2009, *ApJ*, 693, 1084  
 Xie J.-W., et al., 2016, *Proceedings of the National Academy of Science*, 113, 11431  
 Zinzi A., Turrini D., 2017, *A&A*, 605, L4

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.

---

## **C Kelley et al, 2020: Quintet of Outbursts of Comet 46P/Wirtanen**







The following 27 pages contain the article by Kelley et al. [2020] as it has been submitted to *The Planetary Science Journal*.

## Quintet of Outbursts of Comet 46P/Wirtanen

2 MICHAEL S. P. KELLEY <sup>1</sup>, TONY L. FARNHAM,<sup>1</sup> JIAN-YANG LI (李荐扬) <sup>2</sup>,  
3 DENNIS BODEWITS <sup>3</sup>, COLIN SNODGRASS <sup>4</sup>, JOHANNES ALLEN,<sup>3</sup>

4 ERIC C. BELLM <sup>5</sup>, ANDREW J. DRAKE,<sup>6</sup> DMITRY A. DUEV <sup>7</sup>,  
5 MATTHEW J. GRAHAM <sup>7</sup>, FRANK J. MASCI <sup>8</sup>

6 ZWICKY TRANSIENT FACILITY COLLABORATION

7 M. DOMINIK <sup>9</sup>, U. G. JØRGENSEN,<sup>10</sup> A. ANDREWS,<sup>11</sup> N. BACH-MØLLER,<sup>10</sup>  
8 V. BOZZA,<sup>12,13</sup> M. J. BURGENDORF,<sup>14</sup> J. CAMPBELL-WHITE <sup>15</sup>, S. DIB,<sup>10,16</sup>  
9 Y. I. FUJII <sup>17,18</sup>, T. C. HINSE,<sup>19,20</sup> M. HUNDERTMARK <sup>21</sup>, E. KHALOUEI,<sup>22</sup>  
10 P. LONGA-PEÑA,<sup>23</sup> M. RABUS <sup>24,25</sup>, S. RAHVAR,<sup>22</sup> S. SAJADIAN,<sup>26</sup>  
11 J. SKOTTFELT <sup>11</sup>, J. SOUTHWORTH,<sup>27</sup> J. TREGLOAN-REED,<sup>23</sup> E. UNDA-SANZANA,<sup>23</sup>

12 MINDSTEP COLLABORATION

13  
14 <sup>1</sup>Department of Astronomy, University of Maryland, College Park, MD 20742-0001, USA

15 <sup>2</sup>Planetary Science Institute, 1700 E. Ft. Lowell Rd., Tucson, AZ 85719, USA

16 <sup>3</sup>Physics Department, Leach Science Center, Auburn University, Auburn, AL 36832, USA

17 <sup>4</sup>Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ, UK

18 <sup>5</sup>DIRAC Institute, Department of Astronomy, University of Washington, 3910 15th Avenue NE,  
19 Seattle, WA 98195, USA

20 <sup>6</sup>Department of Astronomy, California Institute of Technology, 1200 E. California Blvd, Pasadena,  
21 CA 91125, USA

22 <sup>7</sup>Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena,  
23 CA 91125, USA

24 <sup>8</sup>IPAC, California Institute of Technology, 1200 E. California Blvd, Pasadena, CA 91125, USA

25 <sup>9</sup>Centre for Exoplanet Science, SUPA, School of Physics & Astronomy, University of St Andrews,  
26 North Haugh, St Andrews KY16 9SS, UK

27 <sup>10</sup>Niels Bohr Institute, Øster Voldgade 5, 1350 Copenhagen, Denmark

28 <sup>11</sup>Department of Physical Sciences, The Open University, Milton Keynes, MK7 6AA, UK

29 <sup>12</sup>Dipartimento di Fisica "E.R. Caianiello", Università di Salerno, Via Giovanni Paolo II 132,  
30 84084, Fisciano, Italy

31 <sup>13</sup>Istituto Nazionale di Fisica Nucleare, Sezione di Napoli, Napoli, Italy

32 <sup>14</sup>Universität Hamburg, Faculty of Mathematics, Informatics and Natural Sciences, Department of  
33 Earth Sciences, Meteorological Institute, Bundesstraße 55, 20146 Hamburg, Germany

34 <sup>15</sup>SUPA, School of Science and Engineering, University of Dundee, Nethergate, Dundee DD1 4HN,  
35 UK

36 <sup>16</sup>Max Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany

37 <sup>17</sup>Institute for Advanced Research, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8601,  
38 Japan



<sup>18</sup>Department of Physics, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8602, Japan

<sup>19</sup>Chungnam National University, Department of Astronomy & Space Science, 34134 Daejeon, South Korea

<sup>20</sup>Centre of Excellence for Astrophysics and Astrochemistry, Nicolaus Copernicus University, Torun, Poland

<sup>21</sup>Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg (ZAH), 69120 Heidelberg, Germany

<sup>22</sup>Department of Physics, Sharif University of Technology, PO Box 11155-9161 Tehran, Iran

<sup>23</sup>Centro de Astronomía (CITEVA), Universidad de Antofagasta, Av. Angamos 601, Antofagasta, Chile

<sup>24</sup>Department of Physics, University of California, Santa Barbara, CA 93106-9530, USA

<sup>25</sup>Las Cumbres Observatory Global Telescope, 6740 Cortona Dr., Suite 102, Goleta, CA 93111, USA

<sup>26</sup>Department of Physics, Isfahan University of Technology, Isfahan 84156-83111, Iran

<sup>27</sup>Astrophysics Group, Keele University, Staffordshire, ST5 5BG, UK

(Received September 30, 2020)

Submitted to The Planetary Science Journal

## ABSTRACT

Cometary activity is a manifestation of sublimation-driven processes at the surface of nuclei. However, cometary outbursts may arise from other processes that are not necessarily driven by volatiles. In order to fully understand nuclear surface and their evolution, we must identify the causes of cometary outbursts. In that context, we present a study of mini-outbursts of comet 46P/Wirtanen. Five events are found in our long-term lightcurve of the comet around its perihelion passage in 2018. The apparent strengths range from  $-0.2$  to  $-1.4$  mag in a  $5''$  radius aperture, and total cross sectional areas correspond to dust masses between  $\sim 10^4$  to  $10^6$  kg, but with large uncertainties due to the unknown grain size distributions. However, the nominal mass estimates are the same order of magnitude as the mini-outbursts at comet 9P/Tempel 1 and 67P/Churyumov-Gerasimenko, events which were notably lacking at comet 103P/Hartley 2. We compare the frequency of outbursts at the four comets, and suggest that the surface of 46P has large-scale ( $\sim 10$ – $100$  m) roughness that is intermediate to that of 67P and 103P, if not similar to the latter. The strength of the outbursts appear to be correlated with time since the last event, but a physical interpretation with respect to solar insolation is lacking. We also examine *Hubble Space Telescope* images taken about 2 days following a near-perihelion outburst. No evidence for macroscopic ejecta was found in the image, with a limiting radius of about 2-m.

## 1. INTRODUCTION

Comet 46P/Wirtanen is a small Jupiter-family comet that has been considered as a potential spacecraft target. The effective radius is 0.6 km (Lamy et al. 1998; Boehnhardt et al. 2002), making it one of the smallest periodic comets (Snodgrass et al. 2011). The comet made an historic flyby of Earth in 2018, passing just 0.0775 au

( $1.16 \times 10^7$  km) on 2018 December 16 (JPL Horizons orbital solution K181/21). The geometry with respect to the Earth and Sun was exceptionally favorable, with long observing opportunities and a total apparent magnitude peaking near  $V \sim 5$  mag (IAU Minor Planet Center Database).

In many respects, comet Wirtanen is considered a near-twin of comet 103P/Hartley 2. They have similar orbits, dust and gas production rates, and nuclear radii (A’Hearn et al. 1995, 2011). As a consequence, both comets are considered to be hyperactive, i.e., their water production rates require a sublimating surface area comparable to the total nuclear surface area, whereas most comets have a ratio  $\lesssim 10\%$  (A’Hearn et al. 1995). Comet Hartley 2 was a flyby target of the *Deep Impact* spacecraft (A’Hearn et al. 2011) and the subject of a large observational campaign in 2010 (Meech et al. 2011). Thus, the 2018 perihelion passage of comet Wirtanen presented an opportunity to apply the knowledge gained from the studies of comet Hartley 2 to comet Wirtanen and the broader comet population.

One important difference between Wirtanen and Hartley 2 is the lack of cometary outbursts in the latter (A’Hearn et al. 2011). Cometary outbursts are brief increases in mass loss (Hughes 1990), instigated by mechanical or thermophysical processes, such as cliff collapse (Pajola et al. 2017), avalanches (Steckloff & Melosh 2016), or nuclear fragmentation (Boehnhardt 2004), and potentially sub-surface pressure release of a gas reservoir (Agarwal et al. 2017) or gas dissolved in a liquid (Miles 2016). Outbursts of many comets have been observed, e.g., comets Kohoutek 1973f, Bowell 1980b, 9P/Tempel 1, and 67P/Churyumov-Gerasimenko (A’Hearn & Cowan 1975; A’Hearn et al. 1984, 2005, 2016), but none have been confirmed for comet Hartley 2. This result is in spite of the 2010 observational campaign, and near-continuous photometry from the *Deep Impact* spacecraft. In contrast, clear outbursts of comet Wirtanen were observed in 1991, 2002, 2008, and 2018 (Yoshida 2013; Kidger 2004, 2008; Kronk et al. 2017; Combi et al. 2020; Farnham et al. 2019).

Dense, long-term photometric and spectroscopic coverage of comets is needed to advance our understanding of cometary activity (A’Hearn 2017). Present-day wide-field time-domain surveys, such as the Zwicky Transient Facility (ZTF; Bellm et al. 2019a; Graham et al. 2019) and the Asteroid Terrestrial-impact Last Alert System (ATLAS; Tonry et al. 2018a), can partially address this challenge with broad-band photometric imaging at a near-daily cadence. In this work, we present a long-term lightcurve of comet Wirtanen and examine it for evidence of outbursts in activity. This paper is a follow-up to the preliminary investigation by Kelley et al. (2019).

## 2. OBSERVATIONS AND DATA

Broad-band photometry of comet Wirtanen was obtained from four observatories in 2018 and 2019: Palomar Observatory, Lowell Observatory, the European Southern Observatory, and the *Hubble Space Telescope*. We first describe the ground-based data, which we use to form a long-term lightcurve of coma, then the *HST* data,

121 which were taken as part of a *Chandra X-Ray Observatory* campaign to study charge  
122 exchange in the cometary coma (Bonamente et al. 2020).

## 123 2.1. Ground-based Observatories

### 124 2.1.1. Palomar Observatory

125 Observations of comet Wirtanen were identified in the ZTF Data Release 3, Part-  
126 nership, and Caltech archives with the ZChecker program (Kelley et al. 2019). ZTF  
127 is a wide-field time-domain survey using the Samuel Oschin 1.2-m telescope at Palo-  
128 mar Mountain with a 16-CCD camera. Each  $6144 \times 6160$  CCD has a  $1''.01$  pixel scale,  
129 yielding a total camera field of view of  $47 \text{ deg}^2$  with an 86% fill factor (Bellm et al.  
130 2019a). The robotic system executes multiple simultaneous surveys, with a range of  
131 science goals (Graham et al. 2019). Comet Wirtanen data was found in 352 images  
132 in total ( $g$ ,  $r$ , and  $i$  bands, 30-s exposure times), taken between 2018 July 13 and  
133 2019 June 06 UTC, observed in the Northern Sky, Galactic Plane, Asteroid Rota-  
134 tion,  $i$ -band, and One-Day Cadence surveys (Bellm et al. 2019b). Most nights have  
135 only one or two images, except during the Asteroid Rotation survey, which observed  
136 46P over 3- to 4-hour periods on 2019 January 24, 25, and 26 UTC with a 255-s  
137 cadence. All data were reduced with the ZTF data pipeline (Masci et al. 2019).  
138 The processing typically includes reference image subtraction, which removes smooth  
139 background and photometrically stable celestial objects, leaving image artifacts and  
140 transients (including solar system objects). We find no significant difference between  
141 small-aperture ( $< 10 \text{ pix}$ ) photometry measured with or without the reference sub-  
142 tracted data, except that the latter are less likely to be affected by background stars.  
143 Therefore, we use reference subtracted data whenever possible for photometry. When  
144 the comet is bright and the angular extent is large, the morphology is best studied  
145 without reference subtraction.

### 146 2.1.2. Lowell Observatory

147 Images of comet Wirtanen were taken with the Lowell Observatory 0.8-m robotic  
148 telescope located at Anderson Mesa (Buie 2010) through an  $R$ -band filter between  
149 2018 September 23 and 2019 February 08 UTC. The camera uses a  $2048 \times 2048$  CCD  
150 with a pixel scale of  $0''.45$ , yielding a  $15'$  field of view. Standard image bias and flat-  
151 field corrections were applied. Typically 3 images were taken per night, with 12- to  
152 300-s exposure times and the telescope tracking at the rate of the comet.

### 153 2.1.3. European Southern Observatory, La Silla

154 After combining the ZTF and Lowell data sets, we identified a gap in temporal  
155 coverage in early August. Select images taken with the Danish 1.54-m telescope at  
156 La Silla Observatory were reduced and examined in order to fill this gap. Observations  
157 utilized the Danish Faint Object Spectrograph and Camera (DFOSC), which has a  
158 field of view of  $13'.7 \times 13'.7$  and a pixel scale of  $0''.39$ , and taken on an approximately  
159 weekly cadence between 2018 June 18 and September 17 UTC, primarily in the  $R$ -

**Table 1.** Comet 46P/Wirtanen geometric, photometric, and derived data.

Column	Name	Unit or scale	Description
(1)	Source	...	Name of telescope
(2)	Date	UTC	Mean time of observations
(3)	$T - T_P$	days	Time offset from perihelion <sup>a</sup>
(4)	$r_h$	au	Comet heliocentric distance
(5)	$\Delta$	au	Comet-observer distance
(6)	$\theta$	deg	Sun-comet-observer (phase) angle
(7)	Filter	...	Filter name
(8)	Exposure	s	Total exposure time
(9)	Airmass	...	Mean airmass of observations
(10)	Seeing	arcsec	FWHM of (potentially trailed) point sources
(11)	$m$	mag	Apparent magnitude in 5'' radius aperture (PS1 system)
(12)	$\sigma_m$	mag	Uncertainty on $m$
(13)	Trail	mag	Trailed-source correction applied to ZTF photometry
(14)	$A(\theta)f\rho$	cm	Comet photometric quantity, based on $m$
(15)	$G$	km <sup>2</sup>	Geometric cross-section, based on $m$
(16)	Trend	mag	$r$ -band magnitude trend from piecewise fit.

<sup>a</sup> $T_P = 2018$  December 12.94146 UTC (Minor Planet Center 2019).

160 band. Additional images were taken in  $UBVI$ -bands later in this period but are not  
 161 included in the work presented here.

#### 162 2.1.4. Photometry

163 All ground-based data are calibrated to the PS1 photometric system using back-  
 164 ground stars in each field. The calibration of the ZTF data are described by Masci  
 165 et al. (2019). The remaining data were calibrated to the  $r_{P1}$ -band (i.e., PS1 system)  
 166 using the ATLAS Refcat2 photometric catalog (Tonry et al. 2018b) and Calviacat  
 167 software (Kelley & Lister 2019). Uncertainties in the absolute calibrations are prop-  
 168 agated into the final measurement errors, but a minimum uncertainty of 0.02 mag is  
 169 assumed. All data are color corrected using the measured coma colors (Section 3)  
 170 and photometric calibration solutions. Photometry within a constant angular aper-  
 171 ture radius of 5'' is given in Table 1, with 372 data points taken on 111 unique nights  
 172 spanning 352 days.

173 Although the comet is bright, it does not saturate the ZTF detectors. In 30-s  
 174 exposures, the saturation limit for point sources is about 13 mag, depending on the  
 175 filter. Since the comet is an extended source, and our photometry is in a 5'' radius  
 176 aperture (whereas seeing is typically around 2'' FWHM), the comet data are not  
 177 saturated despite the bright photometric values reported in this work ( $r \gtrsim 11$  mag).

178 In contrast with the Lowell and Danish telescope observations, the ZTF survey data  
 179 images are tracked in the Celestial reference frame, causing the comet to trail during

the 30-s exposures. With non-sidereal rates up to  $\sim 600'' \text{ hr}^{-1}$ , the comet tailed 0.5–6'' per exposure. Thus, photometry in a 5'' radius aperture may be affected. We attempt to correct for those losses by generating an image of an idealized coma (surface brightness proportional to  $\rho^{-1}$ , where  $\rho$  is the projected distance to the nucleus) and convolving it with a linear kernel. The length of the kernel is equal to the calculated trailed length per exposure, and the correction factor is the ratio of the brightness of the trailed coma to that of the ideal coma, measured in a 5'' radius aperture. The corrections range from  $-0.01$  to  $-0.11$  mag (Table 1), and are applied to all ZTF photometry.

## 2.2. Hubble Space Telescope

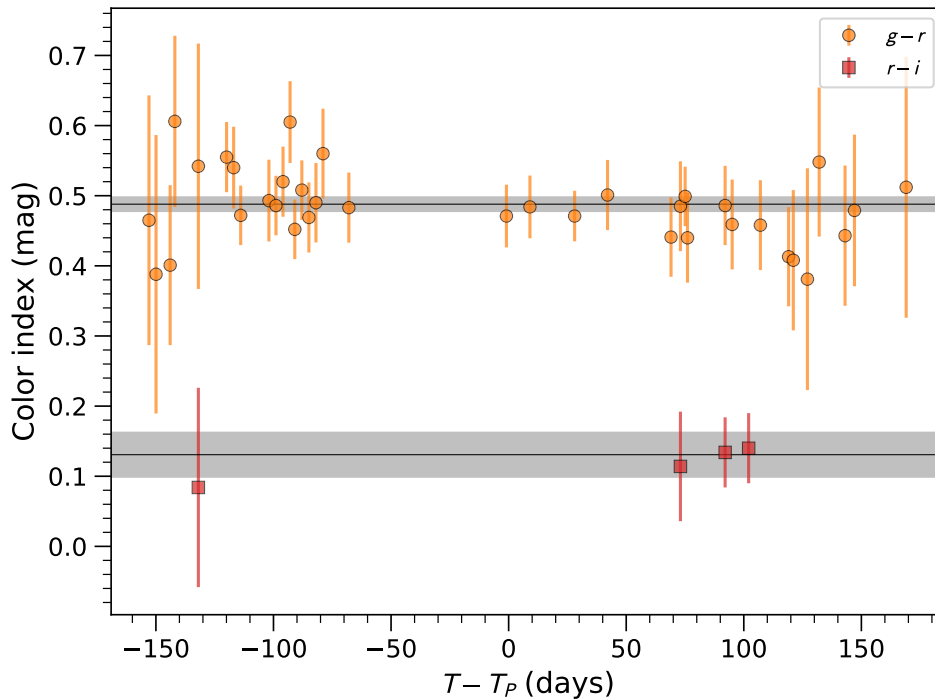
*Hubble Space Telescope* (*HST*) imaged comet Wirtanen with the Wide Field Camera 3 (WFC3) UVIS channel at two epochs. Each epoch contained four *HST* orbits, organized into two two-orbit groups separated by one orbit, covering about 7 hours in duration. The data spanned 2018 December 13 09:15 to 16:18, and December 25 10:30 to 17:33 UTC. The comet was observed through two mid-band filters F689M and F845M (11% wide bandpass) with the  $2\text{k} \times 2\text{k}$  sub-frame, which has a field of view of  $80'' \times 80''$  given the  $0''.04$  pixel scale. Due to the non-linear non-sidereal movement of the comet and the high spatial resolution of the WFC3/UVIS camera, the comet was trailed by up to 4 pixels for all F689M images except one with an 8-pixel trail, and by various amounts up to 9 pixels in the F845M images, despite the short exposure times of 10 and 16 s used for F689M and F845M, respectively. On the other hand, all images are well exposed with the peak brightness up to 24% of the saturation level.

Photometric measurements were based on the images reduced by the standard WFC3 calibration pipeline (Gennaro et al. 2018). To remove cosmic rays, we divided each image into a grid of  $20 \times 20$ -pixel boxes, then clipped and replaced  $3\sigma$  outliers with the mean in each box. The center  $40 \times 40$  pixel region was excluded from this cosmic ray removal process in order to preserve the inner coma. For the fragment search, we also removed cosmic rays with the LA Cosmic algorithm (van Dokkum 2001). Sky background was estimated by the mean of four  $100 \times 100$  pixel boxes near the corners of the images. The pixel area map of the corresponding detector chip was applied to correct for pixel area change in the spatially distorted (FLT) images before photometric measurement. The total count was then measured in a 5''-radius aperture and converted to flux and apparent magnitude following the photometric calibration constants (Gennaro et al. 2018). Our photometry is limited by the absolute photometric uncertainty for WFC3/UVIS images (2%). The effect of source trailing in our images is negligible for 5''-radius aperture photometry.

## 3. RESULTS

### 3.1. Coma Color

The  $g - r$  color of comet Wirtanen was previously measured from a limited set of ZTF photometry by Kelley et al. (2019) to be  $0.45 \pm 0.02$  mag. We compute  $g - r =$



**Figure 1.** Color index versus time for comet 46P/Wirtanen measured with Zwicky Transient Facility photometry in the  $g$ ,  $r$ , and  $i$  bands. The means and uncertainties are drawn as horizontal lines and shaded areas, respectively.

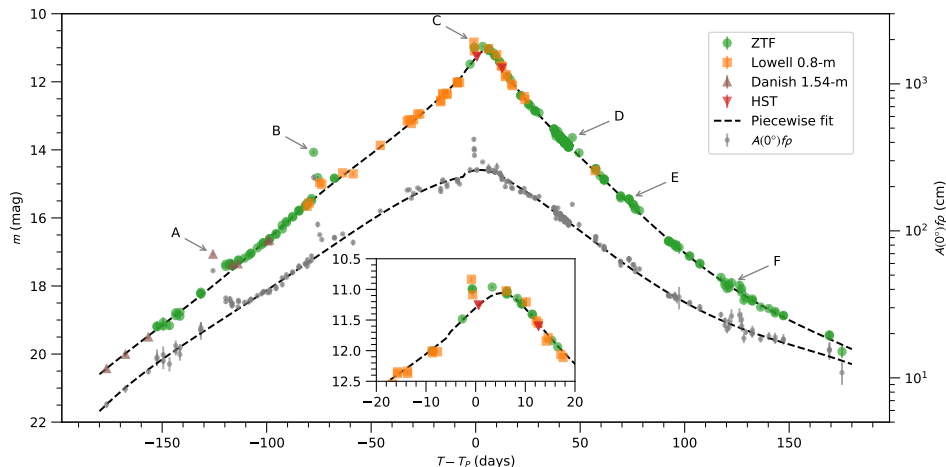
220  $0.49 \pm 0.01$  mag and  $r-i = 0.13 \pm 0.03$  mag from the weighted means of 36 and 4 nightly  
 221 color measurements, respectively. Those colors appear to be consistent throughout  
 222 the data set (Fig. 1), with the largest deviation at the  $2.0\sigma$  level (reduced  $\chi^2$  is 0.5  
 223 for  $g-r$ , 0.1 for  $r-i$ ). The mean color from *HST* is  $m_{689} - m_{845} = 0.15 \pm 0.02$  mag.  
 224 To convert the *HST* photometry into  $r$ -band data, we use the measured *HST* color,  
 225 and extrapolate it to the PS1  $r$ -band with a spectrum of the Sun. Throughout this  
 226 work, we adopt the composite spectrum of the Sun from Haberreiter et al. (2017)  
 227 and Willmer (2018) for filter calibrations (we estimate the apparent magnitude of the  
 228 Sun in the F689M and F845M filters to be  $-27.01$  and  $-27.07$  mag, AB magnitude  
 229 system). Based on the *HST* color, we calculate  $r - m_{689} = 0.04$  mag. Using these  
 230 colors, an effective  $r$ -band lightcurve versus time from perihelion is shown in Fig. 2.

231 The colors of the coma correspond to spectral gradients (A’Hearn et al. 1984) of  
 232  $S_{g,r} = 6.8 \pm 0.7\%$  per 100 nm  $S_{r,i} = 0.7 \pm 2.0\%$  per 100 nm, and  $S_{689,845} = 5.2 \pm 1.2\%$   
 233 per 100 nm, where the subscripts denote the bandpasses used in the calculations.  
 234 The  $S_{r,i}$  and  $S_{689,845}$  are consistent at the  $2\sigma$  level. Note that these colors are not  
 235 necessarily that of the dust coma, as there are gas emission bands present at these  
 236 wavelengths, especially  $C_2$  in  $g$ , but also  $NH_2$  in  $r$ ,  $R$ , and F689M, and  $CN$  in  $i$ .  
 237 See Fink (2009) and Ponomarenko et al. (2018) for relevant optical spectra of comet  
 238 Wirtanen.

239

### 3.2. Quiescent Activity



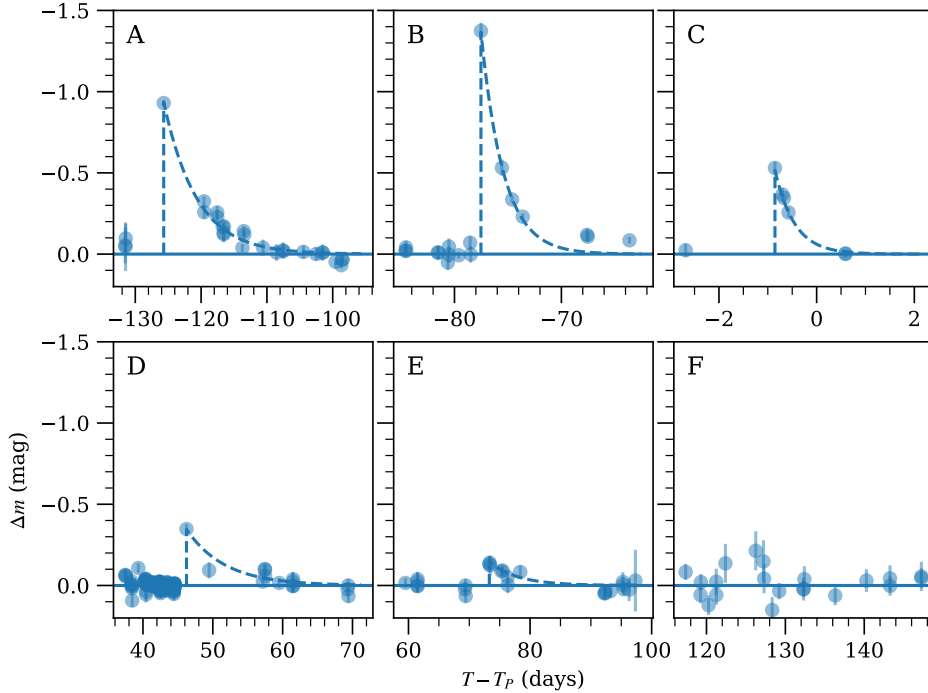


**Figure 2.** Lightcurve of comet 46P/Wirtanen measured within  $5''$  radius apertures. Photometry from the  $g$ - and  $i$ -band have been scaled with the measured coma colors to make an effective  $r$ -band data set. Also shown is the photometry converted to the  $A(0^\circ)f\rho$  quantity. A trend-line based on a piecewise fit to the photometry is shown as a dashed line (see Section 3 for details). Six sets of anomalous data points are labeled A–F.

240 In order to identify outbursts, it helps to define the quiescent activity trend. We use  
 241 the  $Af\rho$  model of A’Hearn et al. (1984). This quantity is based on the brightness of  
 242 the coma within a circular aperture. Formally, it is the product of grain albedo ( $A$ ),  
 243 filling factor within the photometric aperture ( $f$ ), and aperture radius ( $\rho$ , projected  
 244 length at the distance of the comet).  $Af\rho$  carries the units of  $\rho$ , but is proportional  
 245 to dust mass-loss rate under idealized assumptions, e.g., a coma in free expansion  
 246 with a constant production rate, grain size distribution, and composition (i.e.,  $1/\rho$   
 247 surface brightness profile), and photometry free of gas contamination. (See Fink  
 248 & Rubin (2012) for more discussion on the physical interpretation.) The albedo is  
 249 commonly expressed as a function of phase angle,  $\theta$ , in order to explicitly account for  
 250 the phase effect from non-isotropic scattering of sunlight by coma dust. For the phase  
 251 correction, we adopt the Schleicher-Marcus phase function, first used by Schleicher &  
 252 Bair (2011). In Table 1, all photometry is converted to  $A(0^\circ)f\rho$ . In Fig. 2, we plot  
 253 the  $A(0^\circ)f\rho$  values after accounting for the measured color differences.

254 We fit the  $\log A(0^\circ)f\rho$  data with a polynomial as a function of either  $\log r_h$  or time.  
 255 Candidate outbursts were excluded from the fit. The best fit to the entire lightcurve  
 256 is  $(263 \pm 1) r_h^{-4.01 \pm 0.01}$  cm (RMS 0.04 mag). However, we found this trend does not  
 257 have sufficient precision for quantifying outbursts, with local deviations as strong as  
 258 22%. Therefore, we split the lightcurve into three segments with break points based  
 259 on time from perihelion,  $T - T_P = -5$  and  $+15$  days. Each segment is fit with 3rd  
 260 or 4th degree polynomials versus time. The RMS of the residuals are 0.07, 0.05, and  
 261 0.04 mag (excluding possible outbursts).

262 The piecewise approach handles the near-perihelion photometry separately from the  
 263 rest of the data, and allows for short- and long-term asymmetries around perihelion.



**Figure 3.** Lightcurves of six anomalous sets of data points, labeled A–F, identified in the lightcurve of 46P/Wirtanen (Fig. 2). For each set, the baseline photometric trend has been removed, and an illustrative exponential function is shown as a dashed line. Events A–C appear to be outbursts, characterized by a rapid brightening and exponential fading. Event D is sparsely observed, but confirmed as an outburst by image morphology. Event E appears to be a real deviation from the trend, but is not obviously an outburst. Event F was not confirmed in the image morphology (Section 3.3).

264 Near perihelion, the geometrical circumstances vary rapidly. The comet moves  $70^\circ$   
 265 on the sky and through opposition, which occurred 6 days after perihelion. Thus,  
 266 the projection of the potentially non-isotropic coma onto the sky changes substan-  
 267 tially, which affects the small aperture photometry. We find that the  $A(0^\circ)f\rho$  is near  
 268 constant from  $-3$  to  $+9$  days (Table 1), aside from an outburst at  $-1$  day and a  
 269 single-point outlier on day  $+3$  (Fig. 2, inset). Moreover, the near-perihelion  $A(0^\circ)f\rho$   
 270 values are elevated by about 20% with respect to the adjacent pre- and post-perihelion  
 271 trends.

272 Note that our best-fit trends depend on the idealized assumptions of the  $Af\rho$  model  
 273 (especially the assumption of a  $1/\rho$  surface brightness profile), our adopted phase  
 274 curve, and our photometric aperture size. The goal of our investigation is to identify  
 275 and characterize outbursts in the comet’s activity, and the piecewise best-fit trend  
 276 will serve this purpose, but may not be appropriate for other contexts. We plot the  
 277 piecewise trend in Fig. 2 and report the trend values for each observation in Table 1.

278

### 3.3. Outbursts

279 From manual inspection of the lightcurve, we identify six sets of significant photo-  
 280 metric outliers, labeled A–F in Fig. 2. Figure 3 shows each set of data, after removing



the photometric trends. All but event F appear to be brightening events (i.e., outbursts). Event C is followed by a single-point outlier 4 days later on 2018 December 16 UTC. The rapid changes about perihelion, and the fact that the photometry is sparse around this point (it is the only data between December 13 and 19), makes defining the quiescent activity at that time more challenging, therefore we do not interpret this point as an outburst. Event F is also difficult to interpret, due to the weak peak brightness ( $\sim -0.2$  mag), and a possible change in the quiescent trend at the same time. Therefore, we only report F as a possible anomaly.

We visually inspected the candidate outburst image sets for supporting morphological evidence. Because the unresolved nucleus is the ultimate source of any ejecta, the morphology of an outburst is initially that of a point-source, until the ejecta has moved far enough from the nucleus to be detectable as an extended source (as image sensitivity allows). For each event, we defined one or more pre-event images to be used as a baseline model that was scaled and subtracted from the post-event data. By inspection of the residuals, we can help identify the cause of the photometric anomalies. The data were processed with the IPAC Montage software to scale images to a common pixel scale, place the comet at the center of the field, and align the projected Sun direction along the  $+x$ -axis. The images are photometrically scaled to the post-event circumstances using the best-fit lightcurve trend, then median combined and subtracted from a post-event image to reveal the putative outburst ejecta. Events A, B, D, and E are shown in Fig. 4, and event C in Figs. 5 and 6. Details on all sets follow. Comments on the ejecta distributions are based on visual inspection of the images and radial profiles; position angles are measured eastward of Celestial north. Photometry of the residuals are reported in Table 2.

(A) Seven ZTF images taken from 2018 July 22 to 2018 August 03 UTC were combined and subtracted from the three median combined Danish 1.54-m  $R$ -band images taken on August 09. The residuals are extended, but still centrally peaked at the nucleus, and wholly contained within a  $7''.1$ -radius aperture.

(B) Six ZTF images taken from 2018 September 22 to 25 UTC were combined and subtracted from the ZTF  $i$ -band image taken September 26. The ejecta is nearly point-source like, but slightly extended towards  $PA \sim 270^\circ$ . This direction is inconsistent with the proper motion trailing, which is  $0''.4$  along  $PA = 295^\circ$ . A nearby star limits any photometric aperture to  $\leq 7''.1$ , however, this aperture appears to encompass all of the ejecta.

(C) Scaling and subtracting the 2018 December 10  $g$ -band image from the December 12 UTC  $g$ -band image resulted in a halo of negative residuals around the outburst ejecta, perhaps because our photometric scaling is designed for small apertures yet the extended coma at this time is more affected by gas (i.e.,  $C_2$ ). We instead examine the  $r$ -band data from December 04 and 12. Based on these images, the outburst appears to have three components at position angles 36,

72, and 296°. The interpretation of the morphology is affected by the subtraction, which leaves strong negative residuals towards PA $\sim$ 180°, and more subtle residuals towards 55°. We enhanced the *r*-band images by normalizing them with an azimuthally averaged coma (Fig. 5). This confirms that the two components at 36 and 72° are not an artifact caused by over-subtraction of the coma along PA $\sim$ 55°.

The residual emission is distributed as far as 400'' (23,700 km) from the comet. Aside from an ion tail, it is difficult to ascertain how much of this emission beyond 400'' is from the outburst or from residual background. Therefore we only report photometry within this radius.

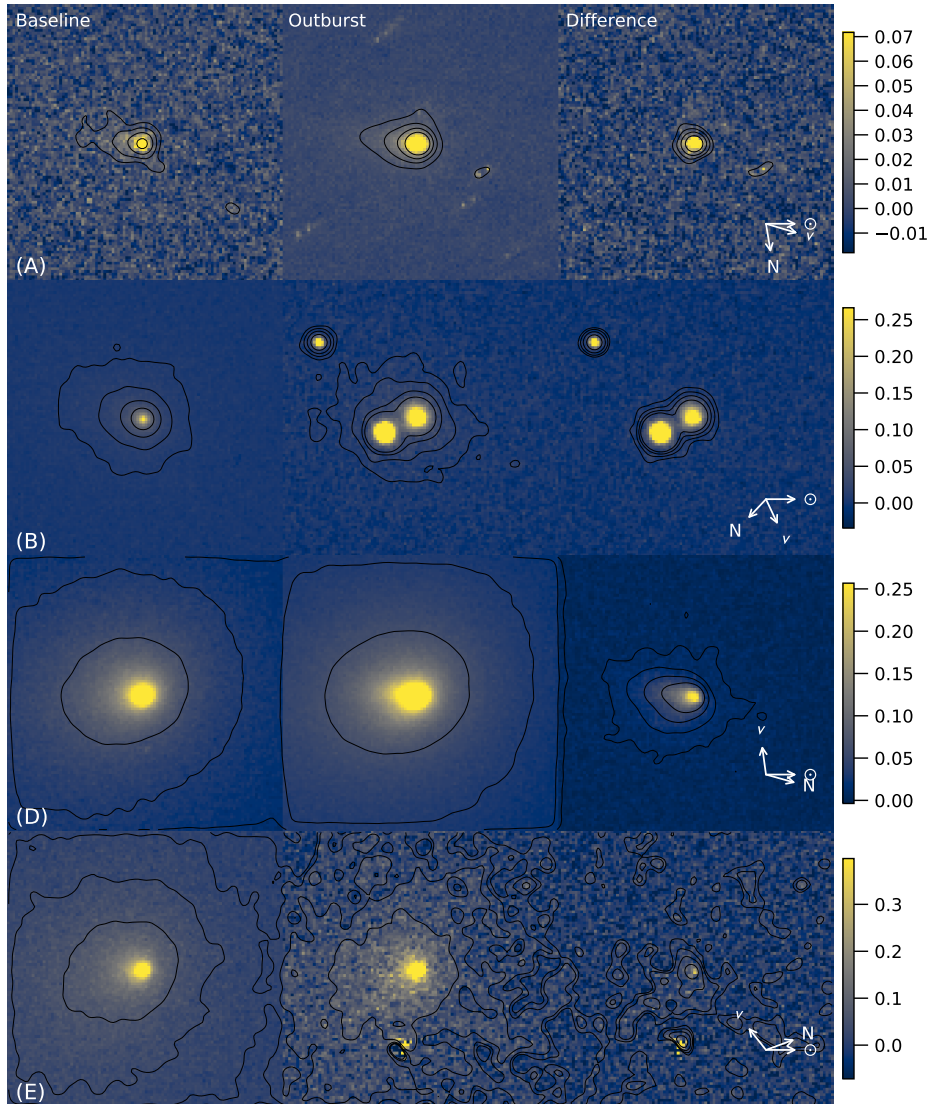
We also inspected the Lowell 0.8-m data taken on 2018 December 05 and on December 12 at 02:07 and 08:46 UTC. Examination of these data reveals ejecta motion over this 6.65-hr period (Fig. 6).

- (D) One ZTF *g*-band image taken 2019 January 24 was subtracted from the *g*-band image taken on January 28 UTC. A small extended source remains in the difference. It has a v-shaped morphology, reminiscent of event C. There is a near linear feature, 27'' long and pointing towards position angle 188°, and a shorter, 21'' long, but broader feature pointing towards 240°. Faint arcminute-scale extended emission is present in the residual image, possibly from C<sub>2</sub> gas.
- (E) After scaling and subtracting three images (1 *g*, 2 *r*) taken on 2019 February 20 from the *r*-band image on 2019 February 24 UTC, a clear residual is detected, no larger than 5'' in radius. However, there is possible extended ejecta towards position angles 180 to 270° in the smoothed contours of the residuals, out to  $\sim$  30''.
- (F) After scaling and subtracting 8 and 11 ZTF images from the data taken on 2019 April 14 and 18 UTC, respectively, we are unable to identify any source in the residuals.

Two of the outbursts have color measurements on the night of the outburst discovery: C and E. The *g* – *r* colors of these events are within 1 $\sigma$  of the mean coma color.

### 3.4. Search for fragments

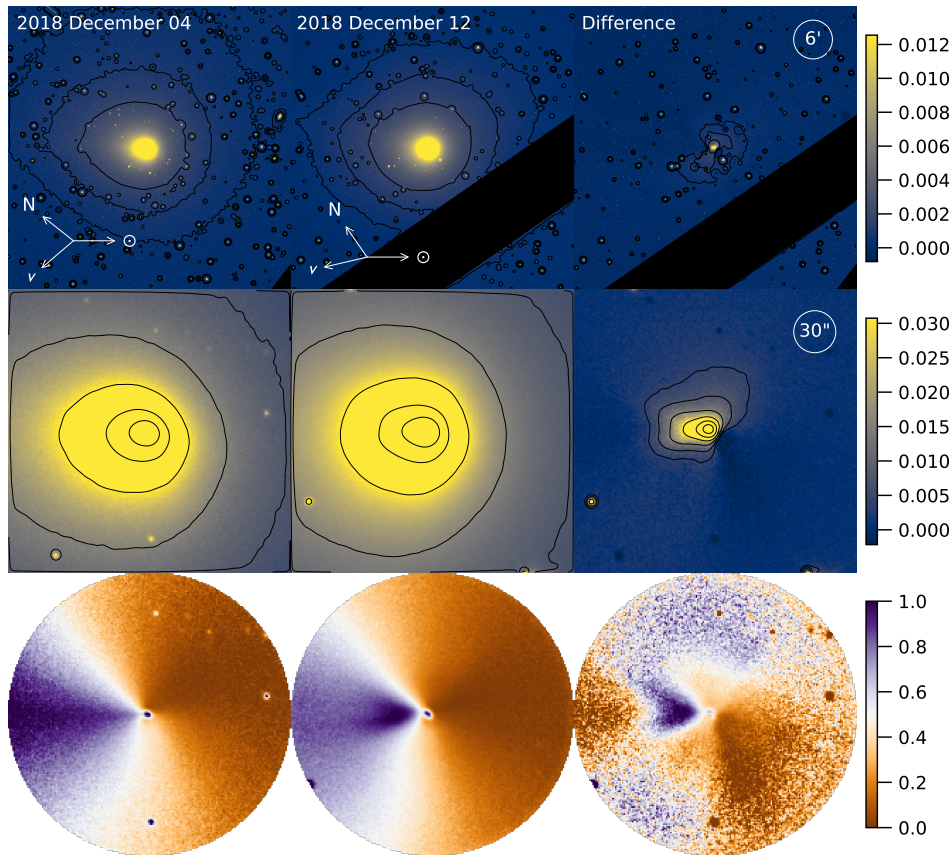
We used the *HST* images of comet Wirtanen obtained on December 13 to look for evidence of any fragments that might have been ejected in the December 11/12 outburst. The close proximity of the comet (0.08 au) and pixel scale of the *HST* WFC3 images (0.04 arcsec/pix) allowed us to investigate the region within a projected distance of around 2300 km of the nucleus for any lingering material. Our observations consist of sequences obtained between 11:32 and 16:18 UTC on December 13 (approximately 35 to 40 hours after the onset of the outburst). Our search



**Figure 4.** (Left and center) Baseline and outburst images for events A, B, D, and E based on ZTF and Danish 1.54-m data. (Right) Difference between outburst and the scaled baseline data. All images are  $1.7' \times 1.7'$ , and scaled with respect to the peak of the comet in the outburst image. Smoothed contours are spaced at factors of two intervals, the brightest of which is at 6.25% of the peak. The projected comet-Sun ( $\odot$ ), comet velocity ( $v$ ), and Celestial north (N) vectors are shown for the outburst image. For outburst B, the comet is the right-most of the two central sources; the other is a background object.

359 utilized four images obtained with the F689M filter and five images with the F845M  
 360 filter, each with exposures short enough for the comet to be untrailed. We used the  
 361 drizzle-processed (DRZ) images, registered on the comet optocenter and rotated so  
 362 that North was up and East to the left.

363 The biggest complication of the search is the large number of cosmic rays that impact  
 364 the *HST* observations, mimicking the types of features that we are looking for. Thus,  
 365 we used cosmic ray cleaned data in addition to using the DRZ images. Although this  
 366 improved the situation somewhat, a significant number of cosmic rays still remained.



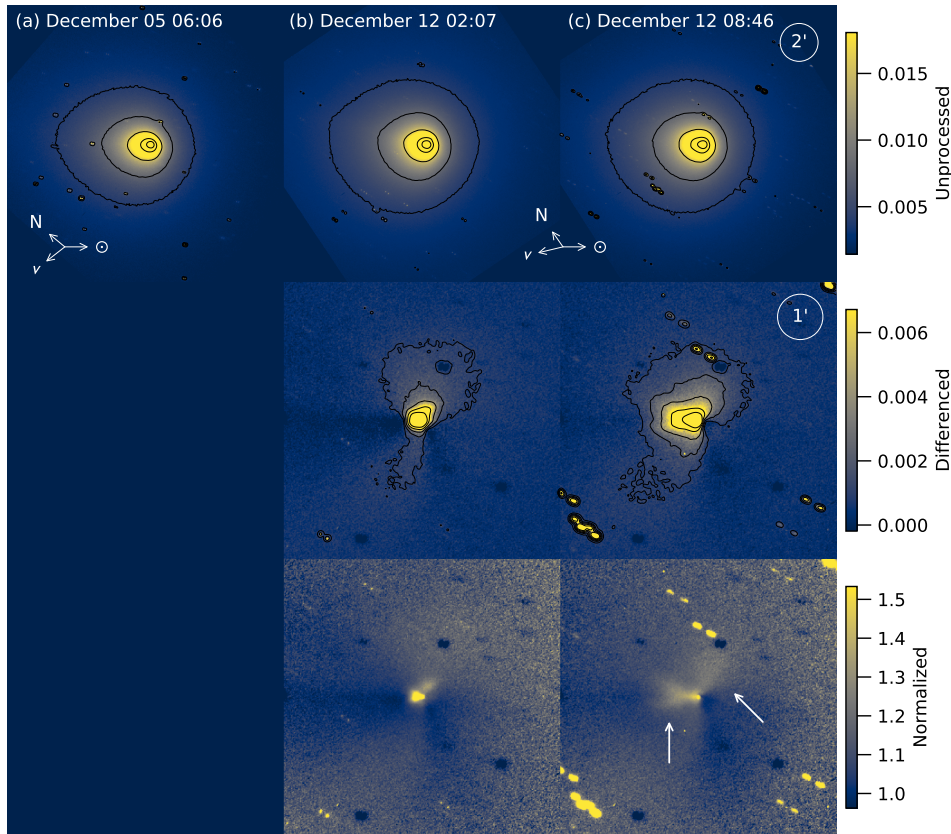
**Figure 5.** Same as Fig. 4 but for event C. (Top)  $44' \times 44'$  field of view with smoothed contours spaced at factors of two intervals, the brightest of which is at 0.20% of the peak. (Center)  $3'4 \times 3'4$  field of view, the brightest contour is 12.5% of the peak. (Bottom) Same as the center, but enhanced by normalizing the data with the azimuthal average, and displayed on a linear scale from the coma minimum to maximum. Note the change in morphology after the outburst with the addition of a v-shaped pattern in the anti-sunward direction. Projected vectors are provided for the baseline and outburst images.

367 Ultimately, we investigated both versions, in case the cosmic ray removal was also  
 368 removing fragments. We also enhanced the images with two different techniques,  
 369 applying an azimuthal average and a (Gaussian) unsharp mask that removes the  
 370 bright central peak of the coma and improves the contrast of any fragments.

371 In order to constrain our search, we assumed that any fragments must be moving  
 372 slowly enough to remain in the field of view for 40 hours (the time from the onset  
 373 of the outburst to the last *HST* observation), setting an upper limit on the proper  
 374 motion of  $25 \text{ pix hr}^{-1}$  (a projected velocity of  $16 \text{ m s}^{-1}$  at the comet). We also  
 375 assumed that particles large enough to be detected will not accelerate significantly  
 376 during the 5-hr window of the *HST* images, and thus any candidates will move along  
 377 a line with spacing proportional to the intervals in the observation times.

378 For each combination of filter/enhancement, we blinked the sequence of images to  
 379 look for candidate particles with acceptable motions. In another approach, we co-  
 380 added the sequences from each filter (and processed as needed), allowing us to look





**Figure 6.** Lowell 0.8-m images taken (a) 2018 December 05, (b) 2018 December 12 at 02:07, and (c) 2018 December 12 at 08:46 UTC. (Top) Unprocessed images and smoothed contours. (Center) Image residuals after subtracting image (a), photometrically scaled according to our baseline photometric trend. (Bottom) Images normalized by the scaled image (a). Arrows indicate two prominent ejecta features in image (c).

381 for linear strings of particles that would represent a moving fragment. In all of our  
 382 searches, we found no convincing evidence for fragments in the *HST* images.

383 Using the cosmic rays as a guide, we estimate that we should have detected any point  
 384 source or central condensation that produces a signal of at least  $2 \times 10^{-18} \text{ W m}^{-2} \mu\text{m}^{-1}$   
 385 (F689M, 0.5 electrons  $\text{s}^{-1}$ ). If we assume an inactive spherical shape with 4% albedo,  
 386 then our detection limit suggests that we should see any fragment larger than  $\sim 2$  m in  
 387 radius, or a mini-comet with a dust cross sectional area of  $\sim 12 \text{ m}^2$ . (These estimates  
 388 ignore issues such as phase effects, but these are small relative to other uncertainties.)

## 389 4. ANALYSIS

### 390 4.1. *Ejecta expansion, grain size*

391 Our general assumption is that all outbursts are brief events, lasting  $\ll 1$  day, and  
 392 that the ejecta can continue to be observed well after the outburst is over. This  
 393 assumption is consistent with the analysis of 30-min cadence observations of out-  
 394 burst B with the *Transiting Exoplanet Survey Spacecraft (TESS)* by Farnham et al.  
 395 (2019), who found coma brightening ceased after 8 hr. Short outburst timescales,

**Table 2.** Summary of outburst circumstances and properties.

Label	Date	$(T - T_P)_0$	$(T - T_P)_1$	$\Delta t$	$\log_{10} H$	$\rho$	$\Delta m_5$	Filter	$m_e$	$G$	$M$
	(UTC)	(days)	(days)	(days)	(J m <sup>-2</sup> )	( $''$ )	(mag)		(mag)	(km <sup>2</sup> )	(kg)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A	2018-08-09	-131.456	-125.683	...	...	7	-0.93	<i>r</i>	$17.49 \pm 0.04$	26.9	$3.6 \times 10^5$
B <sup>a</sup>	2018-09-26	-77.83	-77.81	$51 \pm 3$	9.34	7	-1.4	<i>i</i>	$14.17 \pm 0.04$	73.6	$9.8 \times 10^5$
C	2018-12-12	-2.682	-0.700	$76 \pm 1$	9.80	475	-0.51	<i>r</i>	$9.06 \pm 0.02$	117.9	$1.6 \times 10^6$
D	2019-01-28	44.436	46.234	$47 \pm 2$	9.66	32	-0.15	<i>g</i>	$14.26 \pm 0.03$	16.5	$2.2 \times 10^5$
E	2019-02-24	69.403	73.297	$26 \pm 3$	9.26	5	-0.21	<i>r</i>	$17.80 \pm 0.07$	2.5	$3.3 \times 10^4$

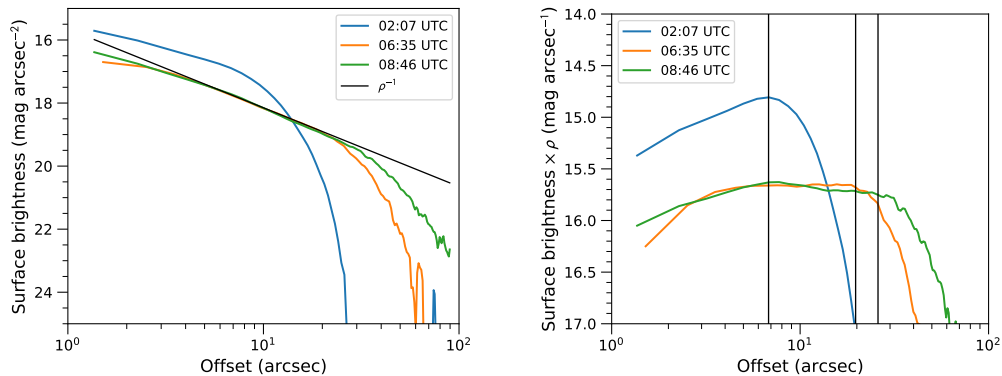
NOTE—Columns: (1) Event label from Fig. 2; (2) Date of first detection; (3) Time of event with respect to perihelion, lower-limit; (4) Time upper-limit; (5) Time since last event and full-range uncertainty; (6) Solar radiant exposure since last event; (7) Photometric aperture radius; (8) Observed peak change in brightness as *r*-band magnitude in 5 $''$  radius aperture; (9) Filter; (10) Total brightness of ejecta in the given filter and 1 $\sigma$  absolute uncertainty; (11) Total geometric cross section; (12) Total mass, assuming  $dn/da \propto a^{-3.5}$  (see Section 4.2 for details).

<sup>a</sup>Outburst timing from Farnham et al. (2019).

396  $\lesssim 1$  hr, are also consistent with the high spatial resolution observations of outbursts  
397 at 9P/Tempel 1 by *Deep Impact* (Farnham et al. 2007) and at 67P/Churyumov-  
398 Gerasimenko by *Rosetta* (Knollenberg et al. 2016; Vincent et al. 2016; Agarwal et al.  
399 2017; Bockelée-Morvan et al. 2017; Rinaldi et al. 2018). Therefore, most outbursts  
400 sampled with a 1 to 3 day cadence and a small aperture will have an observation near  
401 the peak.

402 The fact that event E does not have a distinct photometric peak suggests it occurred  
403 within the 3.9-day gap in data, and that ejecta has moved outside our nominal (5 $''$ )  
404 photometric aperture. In Section 3, we identified faint extended emission in the  
405 outburst residuals, up to  $\sim 30''$  from the nucleus, consistent with this possibility.  
406 An expansion speed of 50 m s<sup>-1</sup> and a projected distance of 30 $''$  corresponds to an  
407 outburst time 2.6 days before the first observation of event E, comfortably within the  
408 3.9-day gap in photometry.

409 For outburst C, we identified motion in the ejecta over a 6.65-hr period. Of the  
410 two features identified in Fig. 6, the anti-sunward feature is brighter and easier to  
411 measure. In Fig. 7, we plot the surface brightness of the ejecta measured in a 5-  
412 pixel-wide box along the anti-sunward direction in the Lowell and ZTF *r*-band data.  
413 Each profile is nearly linear in log-log space closest to the nucleus, then falls with  
414 respect to this line at farther distances. We use the break point (manually estimated  
415 in profiles multiplied by  $\rho$ , Fig. 7, right) to measure the motion of the material. For  
416 break points at 6 $''$ .8, 19 $''$ .7, and 29 $''$ .5, and assuming 1-pix uncertainties, the expansion  
417 speed based on a linear fit is  $55.1 \pm 3.1$  m s<sup>-1</sup>, and an outburst age of  $21.3 \pm 0.9$  hr  
418 in our first image. This places an approximate outburst onset at December 11 04:49  
419 UTC ( $^{+1.9}_{-1.6}$  hr). The reduced- $\chi^2$  statistic is 3.8, but with only 1 degree of freedom



**Figure 7.** (Left) Outburst C ejecta surface brightness based on Lowell 0.8-m and ZTF images. The images are sampled with a 5-pix wide line along the anti-sunward direction, i.e., along the horizontal feature in Fig. 6. (Right) Surface brightness profiles multiplied by distance to the nucleus ( $\rho$ ). Vertical lines mark our estimates of the leading edge.

420 there is a 5% probability of having reduced- $\chi^2 \geq 3.8$  (Bevington & Robinson 1992).  
 421 Thus, we conclude a non-linear expansion is possible, but not strongly supported by  
 422 our data.

423 An upper limit on the outburst C ejecta speed can be estimated from the extent of  
 424 the residuals in the ZTF image ( $400''$ ) and the estimate start time of the outburst.  
 425 Together, they yield an expansion speed of  $250 \text{ m s}^{-1}$ .

426 The lack of outburst ejecta in the *HST* images suggests a lower limit to the expansion  
 427 speed, assuming any slowly moving material is not too diffuse to identify. Given the  
 428 26-hr gap between the last Lowell 0.8-m image and the first *HST* image, and that the  
 429 comet is about  $40''$  from the image edge in the anti-sunward direction, the slowest  
 430 ejecta moved faster than  $\sim 23 \text{ m s}^{-1}$  in projection on the sky.

For the anti-sunward material in outburst C, we can consider the effects of radiation pressure and estimate a lower limit on the grain size assuming the material is in linear expansion. Burns et al. (1979) present the acceleration due to solar radiation pressure,  $a_{\text{rad}}$  on a dust grain as

$$a_{\text{rad}} = \frac{Q_{\text{pr}} L_{\odot} G}{4\pi r_{\text{h}}^2 c m}, \quad (1)$$

431 where  $Q_{\text{pr}}$  is the grain radiation pressure efficiency averaged over the solar spectrum,  
 432  $L_{\odot}$  is the Sun's total luminosity (nominal value  $3.828 \times 10^{26} \text{ W}$ ; Prša et al. 2016),  $G$   
 433 is the grain geometric cross sectional area,  $c$  is the speed of light, and  $m$  is the mass  
 434 of the grain. For simplicity, we take  $Q_{\text{pr}} = 1$ . The projected acceleration on the  
 435 sky is  $a_{\text{rad}}$  attenuated by  $\sin(\theta)$ , where  $\theta$  is the Sun-target-observer angle. Grains  
 436 are accelerated ( $5.3a^{-1}$ )  $\text{m s}^{-1} \text{ hr}^{-1}$  in the anti-sunward direction, projected onto the  
 437 plane of the sky, where  $a$  is the grain radius in  $\mu\text{m}$ . This acceleration corresponds  
 438 to a total displacement of ( $144a^{-1}$ ) km between the ZTF and second Lowell 0.8-m  
 439 epoch (06:35 and 08:46 UT), or at about the level of the seeing ( $59 \text{ km arcsec}^{-1}$ ) for  
 440  $1 \mu\text{m}$  grains. Therefore, the optically dominant grains in this feature are likely at  
 441 least  $1 \mu\text{m}$  in radius.

---

## 4.2. Total geometric cross-sectional area and outburst mass

Converting the observational data into physical quantities allows us to make meaningful comparisons between each outburst and the ambient coma. However, this conversion relies upon several unknown quantities, and therefore will be dependent on our adopted parameters and assumptions. First, we assume a dust  $V$ -band geometric albedo of  $A_p(V) = 4.00\%$ . Given our measured colors, the corresponding albedos are 3.82, 4.19, and 4.22% at  $g$ ,  $r$ , and  $i$ , respectively. Ignoring the dependence of scattering efficiency on grain size, the total geometric cross-sectional area,  $G$ , within a photometric aperture is

$$G = \frac{\pi r_h^2 \Delta^2}{A_p \Phi(\theta)} 10^{-0.4(m-m_\odot)}, \quad (2)$$

where  $\Delta$  is the observer-comet distance in units of length,  $\Phi(\theta)$  is the coma phase function evaluated at phase angle  $\theta$ ,  $m$  is the apparent magnitude of the dust, and  $m_\odot$  is the apparent magnitude of the Sun at 1 au in the same bandpass and magnitude system. For  $r_h$  expressed in units of au,  $G$  will carry the units of  $\Delta^2$ . The coma and outburst photometry are converted to  $G$  and listed in Tables 1 and 2, respectively.

Converting cross-sectional area to dust mass is more uncertain. Here, we require assumptions on the grain density and grain size distribution. For density, we take  $1000 \text{ kg m}^{-3}$ , which allows for some porosity in the grains. Power-law size distributions roughly approximate the grain size distributions observed in situ by spacecraft dust instruments and impacts on the *Stardust* collector (McDonnell et al. 1987; Green et al. 2004; Price et al. 2010; Fulle et al. 2016; Merouane et al. 2017). We assume a differential size distribution,  $dn/da$ , with a power-law slope of  $k = -3.5$ , which is within the estimated time-averaged value of  $-3.3 \pm 0.3$  derived by Fulle (2000) from coma morphology. It is also the cross-over point for mass estimates based on observed brightness, i.e., for values  $< -3.0$  the largest particles dominate the estimated mass, whereas for  $> -4.0$  the smallest particles dominate the mass. Finally, we assume the dust grain radii span from  $0.1 \mu\text{m}$  to  $1 \text{ mm}$ . For these parameters, we convert the outburst geometric cross-sectional area estimates to total mass and provide them in Table 2. The masses range from  $3 \times 10^4$  to  $2 \times 10^6 \text{ kg}$ . For  $k = -3$ , increase the mass estimate by a factor of 10, for  $k = -4$ , decrease the estimate by a factor of 10 (e.g., see Tubiana et al. 2015).

## 4.3. Outburst Frequency

The five outbursts are nearly evenly distributed throughout the observed period. Neglecting the significant gaps in the lightcurve where small events may have taken place (especially near  $-40$ ,  $+80$ , and  $+160$  days), we list the time elapsed between each outburst,  $\Delta t$ , in Table 2 and plot ejecta mass versus  $\Delta t$  in Fig. 8. There is an intriguing correlation between  $\Delta t$  and the amount of material ejected. Pearson's correlation coefficient calculated for  $\Delta t$  and  $\log_{10} M$  is 0.92, indicating a strong significance.



In an attempt to better understand the cause of the apparent correlation, we estimate the solar radiant exposure,  $H$ , based on the comet-Sun distance over the time periods between outbursts, and list them in Table 2 (assuming a solar luminosity of  $3.828 \times 10^{26}$  W; Prša et al. 2016). The correlation is not as good (0.62), as seen in Fig. 8. The main outlier appears to be outburst B, which is unusually strong for the amount of energy received. Our radiant exposure calculation does not consider the source location and pole orientation, nor local topography.

With a pole solution and the assumption of a spherical nucleus, we can explore if a single source illuminated by the Sun could be responsible for all 5 outbursts. The best pole solutions of Knight et al. (in preparation) indicate a high obliquity, with equinox near perihelion. Thus, a near equatorial source could be illuminated during each outburst. For their best pole solution, RA, Dec =  $139^\circ$ ,  $+2^\circ$  (obliquity of  $110^\circ$ ), we find that planetocentric latitudes from  $-30^\circ$  to  $+20^\circ$  are illuminated during outbursts A through E ( $5^\circ$  steps were tested).

We re-calculated the solar radiant energy, this time considering a single source region on a rotating spherical nucleus with the pole orientation of 46P from Knight et al. and latitudes from  $-30^\circ$  to  $+20^\circ$ . We searched for solutions that would improve the mass-energy correlation. More northern latitudes greatly reduced the amount of energy received before outburst E occurred. We show a latitude of  $+20^\circ$  in Fig. 8 as an example. Due to the change in energy for event E, the correlation coefficient between  $\log_{10} M$  and  $\log_{10} H$  increased from 0.59 to 0.89. However, the scatter between events B, C, and D could not be improved. This exercise does not demonstrate that these events are all physically connected, but assuming that they are, it does suggest insolation could be partially responsible for the correlation between ejecta mass and time since the last event.

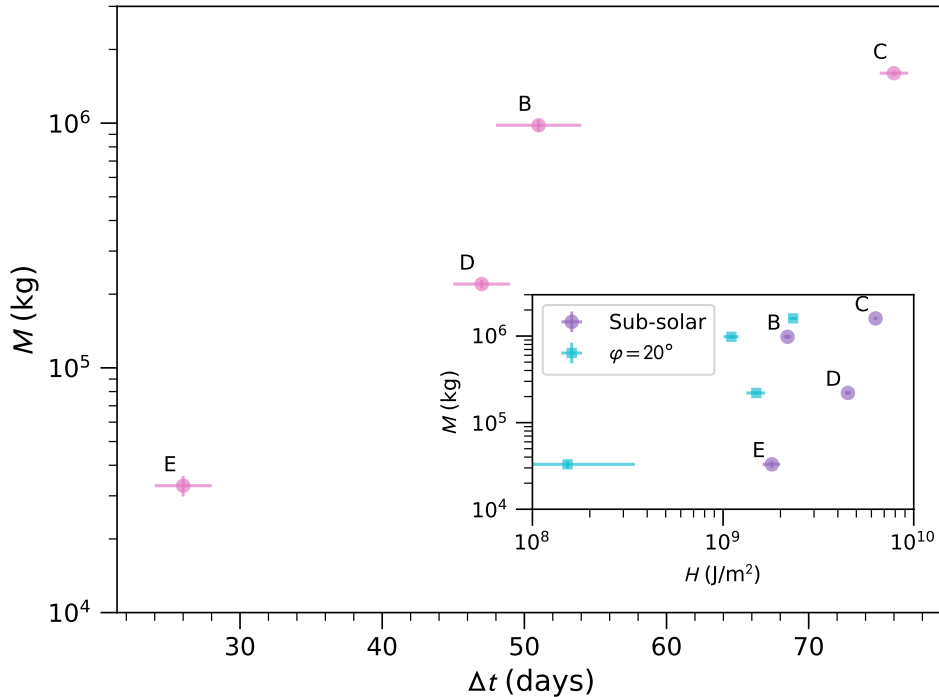
## 5. DISCUSSION

### 5.1. Other Observations of Wirtanen's Outbursts

Farnham et al. (2019) analyzed *TESS* observations of outburst B. They found dust expansion speeds of a few tens of  $\text{m s}^{-1}$ , based on the size of their photometric aperture (25,000 km radius) and the centrally condensed appearance in the images ( $7900 \text{ km pix}^{-1}$ ) that lasted up to 20 days. Farnham et al. (2020) observed outbursts C and D with a near 1-hr cadence, allowing them to estimate dust expansion speeds of  $68 \pm 5$  and  $162 \pm 15 \text{ m s}^{-1}$ , respectively. Our temporal resolution is coarser than that of Farnham et al. (2020), but our estimated expansion speed for the anti-sunward ejecta in outburst C,  $55 \pm 3 \text{ m s}^{-1}$ , is in agreement. The fast moving material in outburst C ( $250 \text{ m s}^{-1}$ ) is less than a factor of two faster than the Farnham et al. (2020) measurement of outburst D.

### 5.2. Mini-outbursts of Wirtanen and Other Comets

In terms of mass, the outbursts of comet Wirtanen are similar to the mini-outbursts of comet 9P/Tempel 1 and 67P/Churyumov-Gerasimenko. At comet 67P, the ejecta



**Figure 8.** Estimated outburst mass (nominal grain parameters) versus time elapsed since last outburst. (Inset) Mass versus solar radiant exposure since last outburst. The exposure is calculated for the sub-solar point and for a source at  $+20^\circ$  latitude.

512 mass estimates are of order  $10^4$  to  $10^5$  kg, based on the analysis of 34 outbursts by  
 513 Vincent et al. (2016). They also re-analyzed the 2005 July 02 outburst of comet  
 514 9P, and, with the same assumptions and techniques, estimated a mass of  $5 \times 10^5$  kg.  
 515 Other mini-outbursts of 9P are the same order of magnitude or smaller (Farnham  
 516 et al. 2007). With the grain parameters of Vincent et al. (2016),  $dn/da = a^{-2.6}$  for 1–  
 517 10 to 1–50  $\mu\text{m}$  in radius, we recalculated the ejecta masses of the Wirtanen outbursts:  
 518  $5.9 \times 10^4$  to  $2.8 \times 10^6$  kg (events E and C, respectively). Thus the Wirtanen outbursts  
 519 are the same order of magnitude to one order larger than the events at 9P and 67P.

520 Vincent et al. (2016) estimated the source locations for the 67P mini-outbursts, and  
 521 found they were correlated with regional boundaries, especially near steep scarps or  
 522 cliffs. Indeed, Grün et al. (2016) correlated an outburst to sunrise on a cliff, Pajola  
 523 et al. (2017) directly connected an outburst to an observed cliff collapse, and Agarwal  
 524 et al. (2017) associated an outburst with the collapse of an overhanging wall.

525 At comet 9P, a correlation with areas of high topographical relief or pits has been  
 526 suggested by Belton et al. (2008). The relationship is intriguing but uncertain. Belton  
 527 et al. (2008) analyzed broad ejecta patterns back to planetocentric coordinates of an  
 528 unresolved nucleus, whereas Vincent et al. (2016) worked with nucleus-resolved data,  
 529 and in some circumstances could visually pinpoint the outburst source to the pixel  
 530 level. The techniques of Belton et al. (2008) inherently assume the nucleus is spherical  
 531 and the outburst ejected normal to the surface. However, 9P is faceted, and many

of these facets face the same direction. Therefore, the projection of planetocentric coordinates to the shape model is multi-valued, and the source regions for the 9P mini-outbursts are uncertain.

No outburst equivalent to those seen at 67P, 9P, and 46P was observed at 103P/Hartley 2. Meech et al. (2011) note an outburst of 103P on 2010 September 16 based on water production rates but without additional details and the event was not seen in *SOHO* observations of the H $\alpha$  coma (Combi et al. 2011). Lin et al. (2013) tentatively associate a relative change in jet brightness in processed data with an outburst, but also consider changes in grain properties as a possibility. We note that the comet’s lightcurve, as observed by *Deep Impact*, has three-peak pattern during this period (A’Hearn et al. 2011; Bodewits et al. 2018), and that the time of the change observed by Lin et al. (2013) corresponds to the brightest of the three peaks.

We take the analyses of the 67P mini-outbursts as a guide, and assume most or all mini-outbursts are related to steep scarps, cliffs, and other features of high topography. If true, then the differences in outburst frequency between 67P, 9P, and 103P are related to differences in terrain. That is, the paucity of cliffs, etc. on the nucleus of 103P results in a lack of mini-outbursts by that comet.

We compare the observed frequency of mini-outbursts at comets 67P, 9P, 46P, and 103P. For 67P and 9P, the frequency is about 0.8 and 1.2 day<sup>-1</sup>, respectively (Belton et al. 2008; Vincent et al. 2016). Based on their measured surface areas, 46.9 and 108 km<sup>2</sup>, respectively (Thomas et al. 2013a; Jorda et al. 2016), the surface area normalized rates are 0.017 and 0.011 day<sup>-1</sup> km<sup>-2</sup>. Comet 46P has an effective radius of 0.56 km (Boehnhardt et al. 2002) and an outburst frequency of 5 in 352 days. Normalized to surface area, this is 0.0036 day<sup>-1</sup> km<sup>-2</sup>, or an order of magnitude smaller than 67P and 9P. Finally, comet 103P has an estimated surface area of 5.24 km<sup>2</sup> (Thomas et al. 2013b) but no outbursts observed over the 180-day lightcurve of Meech et al. (2011). If 103P had the same rate of mini-outbursts as 46P, 2 to 3 events could have been seen, but the details on whether or not they would have been detected depend on observing circumstances and cadence. Perhaps the most sensitive monitoring was executed with the *Deep Impact* spacecraft over a period of approximately 3 months (A’Hearn et al. 2011), but without any reported events (1 to 2 expected). Thus, we estimate 103P’s mini-outburst rate to be no more than that of 46P, or  $\lesssim 0.004$  day<sup>-1</sup> km<sup>-2</sup>.

Vincent et al. (2017) identified a correlation between nuclear surface topography and insolation at comet 67P. Based on an analysis of cliff heights ( $\sim 10$ – $100$ -m scale), they found that regions exposed to more sunlight have fewer large cliffs, and proposed that the erosion of surfaces relaxes their topographies. They continued by analyzing the surfaces of other comets visited by spacecraft, and suggested an evolutionary sequence from comets 81P/Wild 2 and 67P (roughest), to 9P (intermediate), and finally to 103P (smoothest). Kokotanekova et al. (2018) hypothesized a similar sequence, based on a correlation between nuclear phase function and albedo. We build upon these results,

573 adding the correlation between outbursts and cliffs and steep scarps at 67P, and  
 574 propose that the frequency of mini-outbursts is also correlated with topography. With  
 575 respect to surface topography and erosion, comet 46P appears to be in a evolutionary  
 576 state intermediate to 103P and 9P, if not similar to 103P.

What remains to be addressed is the difference between the circumstances of discovery for the mini-outbursts. Comet 67P's and 9P's mini-outbursts were primarily observed by spacecraft. However, two events were observed from the Earth: the 2005 June 14 outburst of 9P observed by [Lara et al. \(2006\)](#) and [Feldman et al. \(2007\)](#) (see also the summary by [Meech et al. 2005](#)), and a tentative outburst on 2015 August 23 at comet 67P identified by [Boehnhardt et al. \(2016\)](#). The lack of events observed at 67P from the Earth, despite the intensive photometric monitoring of that comet ([Snodgrass et al. 2017](#)), can be explained by observing geometry and quiescent activity levels. Setting aside the dependence on observation cadence, the discoverability of an outburst,  $D$ , is inversely proportional to the scattering cross sectional area of dust in an aperture, i.e., the  $Af$  term in  $Af\rho$ . Outbursts are also more readily discovered at high spatial resolution, which reduces the amount of ambient coma in favor of the point-source like outburst ejecta. We let  $\rho$  be inversely proportional to observer-comet distance  $\Delta$  (i.e., fixed angular size). The discoverability of outbursts is

$$D \propto \frac{\rho}{Af\rho} \propto \frac{1}{\Delta Af\rho}. \quad (3)$$

577 For the observational parameters of both comets near perihelion, 1.8 au and 1000 cm  
 578 for 67P ([Snodgrass et al. 2017](#)), and 0.08 au and 300 cm for 46P (this work), the  
 579 ratio is  $D(46P)/D(67P) = 75$ . A  $-1$ -mag outburst of 46P at perihelion in 2018  
 580 (i.e., outburst C) would correspond to a  $-0.03$ -mag outburst of 67P in 2015 at its  
 581 perihelion, assuming the same dust physical parameters and photometric aperture  
 582 angular radius.

### 583 5.3. *Potential for Future Mini-Outburst Studies*

584 The hypothesis that mini-outburst frequency is correlated with surface topography  
 585 could be tested with comet 81P/Wild 2, which has many cliffs, pits, and rough sur-  
 586 face features and a surface area similar to 67P ([Brownlee et al. 2004](#); [Vincent et al.](#)  
 587 [2017](#)). Therefore, this comet may have a mini-outburst every few days. However,  
 588 comet 46P/Wirtanen in 2018/2019 provided favorable circumstances for the study of  
 589 cometary mini-outbursts, and we expect that outbursts of 81P at perihelion in 2022  
 590 would be  $\sim 50$  to  $\sim 70$  times more difficult to detect (based on the  $Af\rho$  measurements  
 591 of 81P by [Farnham & Schleicher 2005](#)), which may require creative solutions in or-  
 592 der to execute such a study. Close approaches to Earth are great opportunities for  
 593 mini-outburst discovery, but 81P will be no closer than 0.65 au from the Earth in the  
 594 next 100 years (JPL Horizons orbit solution K162/9). The next expected cometary  
 595 close approach to Earth with a distance similar to 46P will be 364P/PanSTARRS in

596 April 2023 (0.12 au, via the Center for Near-Earth Object Studies<sup>1</sup>), but low solar  
597 elongations (minimum 45°) will affect the post-approach observability.

598 Outbursts are common events with a wide range of strengths (Ishiguro et al. 2016).  
599 Discoveries of outbursts are becoming more prevalent in recent years due to our in-  
600 creased ability to monitor comets (both in the professional and amateur communities),  
601 and with the increased efficiencies of survey telescopes and precise all-sky photometric  
602 catalogs (e.g., PS1 and SkyMapper; Tonry et al. 2018b; Wolf et al. 2018). Together,  
603 these advances increase our discovery efficiencies, and allow us to identify fainter  
604 events. We expect that current and future cometary outburst surveys will continue  
605 to reveal information about cometary behavior and the evolution of cometary surfaces.

## 606 6. SUMMARY

607 We identified five outbursts in a year-long lightcurve of comet 46P/Wirtanen, with  
608 brightnesses ranging from  $-0.2$  to  $-1.4$  mag with respect to the quiescent trend of  
609 the coma, as measured in  $5''$  radius apertures. The total geometric cross sectional  
610 area of dust in the ejecta ranged from 3 to 120 km<sup>2</sup>, assuming sunlight scattered  
611 according to the Schleicher-Marcus phase function. These areas correspond to  $10^4$   
612 to  $10^6$  kg of dust, but with a factor of 10 uncertainty due to the unknown grain size  
613 distribution. The mass estimates are similar to one order of magnitude larger than the  
614 mini-outbursts observed at comets 9P/Tempel 1 and 67P/Churyumov-Gerasimenko.

615 The expansion speed of material ejected by an outburst near perihelion was at  
616 least  $55 \pm 3$  m s<sup>-1</sup> and up to 250 m s<sup>-1</sup>, projected to the plane of the sky. *Hubble*  
617 *Space Telescope* images taken <2 days after the start of this outburst lack any sign  
618 of macroscopic fragments ( $\sim 2$ -m lower limit radius), or any ejecta at all, indicating  
619 a minimum ejection speed of 23 m s<sup>-1</sup>.

620 The time difference between outbursts ranged from 26 to 75 days, and there appears  
621 to be a correlation between the time elapsed and ejecta mass (or cross sectional  
622 area). We attempted to account for the correlation with the amount of insolation  
623 received at the surface, but could only conclude that this was a potential factor. More  
624 information about the geological or topographic circumstances, and the mechanism(s)  
625 of the outbursts may be needed to further consider this correlation.

626 The mini-outbursts of comet 67P are correlated with steep scarps and cliffs, and  
627 in some circumstances can be directly connected to the collapse of such features  
628 (Vincent et al. 2019, and references therein). Extending this relationship to comets  
629 9P/Tempel 1, 103P/Hartley 2, and 46P, suggests that 46P has fewer cliffs per area  
630 than 67P and 9P, and is more similar to 103P. This comparison is in agreement with  
631 the evolutionary sequence of Vincent et al. (2017), which is based on a correlation  
632 between low topographical relief and insolation on the surface of 67P.

633 Future studies of mini-outbursts and their relationship to topography would help  
634 us understand cometary behavior and nuclear surface evolution. Comet 81P/Wild 2

<sup>1</sup> <https://cneos.jpl.nasa.gov/ca/>

potentially has frequent mini-outbursts, but observational circumstances from the Earth are less favorable for discovery than 46P at this time of our study. However, comet 364P/PanSTARRS may present an opportunity to study mini-outbursts in 2023 ( $\Delta \geq 0.12$  au).

## ACKNOWLEDGMENTS

We thank D. Schleicher and M. Knight for contributing some Lowell 0.8-m telescope data, and B. Skiff and L. Wasserman for assisting with the Lowell observations.

Support for this work was provided through the NASA Solar System Observations program grant number 80NSSC20K0673, and through grant number HST-GO-15372 from the Space Telescope Science Institute. This research also acknowledges financial support from the National Research Foundation (NRF; No. 2019R1I1A1A01059609) and the MINEDUC-UA project, ESR1795.

This work is based on observations obtained with the Samuel Oschin Telescope 48-inch at the Palomar Observatory as part of the Zwicky Transient Facility project. ZTF is supported by the National Science Foundation under Grant No. AST-1440341 and a collaboration including Caltech, IPAC, the Weizmann Institute for Science, the Oskar Klein Center at Stockholm University, the University of Maryland, the University of Washington, Deutsches Elektronen-Synchrotron and Humboldt University, Los Alamos National Laboratories, the TANGO Consortium of Taiwan, the University of Wisconsin at Milwaukee, and Lawrence Berkeley National Laboratories. Operations are conducted by COO, IPAC, and UW.

This research is also based on observations made with the NASA/ESA Hubble Space Telescope obtained from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 526555.

This research made use of Montage. It is funded by the National Science Foundation under Grant Number ACI-1440620, and was previously funded by the National Aeronautics and Space Administration's Earth Science Technology Office, Computation Technologies Project, under Cooperative Agreement Number NCC5-626 between NASA and the California Institute of Technology.

*Facilities:* PO:1.2m (ZTF), LO:0.8m, Danish 1.54m Telescope, HST (WFC3)

*Software:* astropy (Astropy Collaboration et al. 2018), sbpy (Mommert et al. 2014), astroquery (Ginsburg et al. 2019), JPL Horizons (Giorgini et al. 1996), Aperture Photometry Tool (Laher et al. 2012), SEP (Barbary 2016), Ginga (Jeschke et al. 2013), DS9 (Joye & Mandel 2003), imexam (Sosey 2017), astroscrappy (McCully et al. 2018)

## REFERENCES

- |   |     |   |     |
|---|-----|---|-----|
| Agarwal, J., Della Corte, V., Feldman, P. D., et al. 2017, Mon. Not. R. Astron. Soc., 469, s606, doi: 10.1093/mnras/stx2386 | 671 | A'Hearn, M. F. 2017, Philosophical Transactions of the Royal Society of London Series A, 375, 20160261, doi: 10.1098/rsta.2016.0261 | 674 |
|   | 672 |   | 3   |
|   | 673 |   |     |
|   | 674 |   |     |



- 675 A'Hearn, M. F., & Cowan, J. J. 1975, 723  
676 *Astron. J.*, 80, 852, doi: [10.1086/111821](https://doi.org/10.1086/111821) 724  
677 3 725
- 678 A'Hearn, M. F., Millis, R. L., Schleicher, 726  
679 D. G., Osip, D. J., & Birch, P. V. 1995, 727  
680 *Icarus*, 118, 223, 728  
681 doi: [10.1006/icar.1995.1190](https://doi.org/10.1006/icar.1995.1190) 3 729
- 682 A'Hearn, M. F., Schleicher, D. G., Millis, 730  
683 R. L., Feldman, P. D., & Thompson, 731  
684 D. T. 1984, *Astron. J.*, 89, 579, 732  
685 doi: [10.1086/113552](https://doi.org/10.1086/113552) 3, 7, 8 733
- 686 A'Hearn, M. F., Vincent, J. B., & Osiris 734  
687 Team. 2016, in *Lunar and Planetary* 735  
688 *Science Conference, Lunar and* 736  
689 *Planetary Science Conference*, 2678 3 737
- 690 A'Hearn, M. F., Belton, M. J. S., 738  
691 Delamere, W. A., et al. 2005, *Science*, 739  
692 310, 258, doi: [10.1126/science.1118923](https://doi.org/10.1126/science.1118923) 740  
693 3 741
- 694 —. 2011, *Science*, 332, 1396, 742  
695 doi: [10.1126/science.1204054](https://doi.org/10.1126/science.1204054) 3, 20 743
- 696 Astropy Collaboration, Price-Whelan, 744  
697 A. M., Sipócz, B. M., et al. 2018, 745  
698 *Astron. J.*, 156, 123, 746  
699 doi: [10.3847/1538-3881/aabc4f](https://doi.org/10.3847/1538-3881/aabc4f) 23 747
- 700 Barbary, K. 2016, *Journal of Open Source* 748  
701 *Software*, 1(6), 58, 749  
702 doi: [10.21105/joss.00058](https://doi.org/10.21105/joss.00058) 23 750
- 703 Bellm, E. C., Kulkarni, S. R., Graham, 751  
704 M. J., et al. 2019a, *Publ. Astron. Soc.* 752  
705 *Pac.*, 131, 018002, 753  
706 doi: [10.1088/1538-3873/aaecbe](https://doi.org/10.1088/1538-3873/aaecbe) 3, 4 754
- 707 Bellm, E. C., Kulkarni, S. R., Barlow, T., 755  
708 et al. 2019b, *Publ. Astron. Soc. Pac.*, 756  
709 131, 068003, 757  
710 doi: [10.1088/1538-3873/ab0c2a](https://doi.org/10.1088/1538-3873/ab0c2a) 4 758
- 711 Belton, M. J. S., Feldman, P. D., 759  
712 A'Hearn, M. F., & Carcich, B. 2008, 760  
713 *Icarus*, 198, 189, 761  
714 doi: [10.1016/j.icarus.2008.07.009](https://doi.org/10.1016/j.icarus.2008.07.009) 19, 20 762
- 715 Bevington, P. R., & Robinson, D. K. 763  
716 1992, *Data reduction and error analysis*, 764  
717 *for the physical sciences*, 2nd edn. (New 765  
718 York: McGraw-Hill) 16 766
- 719 Bockelée-Morvan, D., Rinaldi, G., Erard, 767  
720 S., et al. 2017, *Mon. Not. R. Astron.* 768  
721 *Soc.*, 469, S443, 769  
722 doi: [10.1093/mnras/stx1950](https://doi.org/10.1093/mnras/stx1950) 15 770
- Bodewits, D., Farnham, T., & A'Hearn, 771  
M. 2018, *NASA Planetary Data System* 772  
20
- Boehnhardt, H. 2004, *Split comets*, ed. 773  
M. C. Festou, H. U. Keller, & H. A. 774  
Weaver, 301 3
- Boehnhardt, H., Riffeser, A., Kluge, M., 775  
et al. 2016, *Mon. Not. R. Astron. Soc.*, 776  
462, S376, doi: [10.1093/mnras/stw2859](https://doi.org/10.1093/mnras/stw2859) 777  
21
- Boehnhardt, H., Delahodde, C., 778  
Sekiguchi, T., et al. 2002, *A&A*, 387, 779  
1107, doi: [10.1051/0004-6361:20020494](https://doi.org/10.1051/0004-6361:20020494) 780  
2, 20
- Bonamente, et al. 2020, this issue 4
- Brownlee, D. E., Horz, F., Newburn, 781  
R. L., et al. 2004, *Science*, 304, 1764, 782  
doi: [10.1126/science.1097899](https://doi.org/10.1126/science.1097899) 21
- Buie, M. W. 2010, *Adv. Astron.*, 2010, 783  
130172, doi: [10.1155/2010/130172](https://doi.org/10.1155/2010/130172) 4
- Burns, J. A., Lamy, P. L., & Soter, S. 784  
1979, *Icarus*, 40, 1, 785  
doi: [10.1016/0019-1035\(79\)90050-2](https://doi.org/10.1016/0019-1035(79)90050-2) 16
- Combi, M. R., Bertaux, J.-L., Quémerais, 786  
E., Ferron, S., & Mäkinen, J. T. T. 787  
2011, *Astrophys. J., Lett.*, 734, L6, 788  
doi: [10.1088/2041-8205/734/1/L6](https://doi.org/10.1088/2041-8205/734/1/L6) 20
- Combi, M. R., Mäkinen, T., Bertaux, 789  
J. L., et al. 2020, this issue, 790  
doi: [10.3847/PSJ/abb026](https://doi.org/10.3847/PSJ/abb026) 3
- Farnham, T. F., et al. 2020, this issue 18
- Farnham, T. L., Kelley, M. S. P., Knight, 791  
M. M., & Feaga, L. M. 2019, 792  
*Astrophys. J., Lett.*, 886, L24, 793  
doi: [10.3847/2041-8213/ab564d](https://doi.org/10.3847/2041-8213/ab564d) 3, 14, 794  
15, 18
- Farnham, T. L., & Schleicher, D. G. 2005, 795  
*Icarus*, 173, 533, 796  
doi: [10.1016/j.icarus.2004.08.021](https://doi.org/10.1016/j.icarus.2004.08.021) 21
- Farnham, T. L., Wellnitz, D. D., 797  
Hampton, D. L., et al. 2007, *Icarus*, 798  
191, 146, 799  
doi: [10.1016/j.icarus.2006.10.038](https://doi.org/10.1016/j.icarus.2006.10.038) 15, 19
- Feldman, P. D., McCandliss, S. R., Route, 800  
M., et al. 2007, *Icarus*, 187, 113, 801  
doi: [10.1016/j.icarus.2006.07.027](https://doi.org/10.1016/j.icarus.2006.07.027) 21
- Fink, U. 2009, *Icarus*, 201, 311, 802  
doi: [10.1016/j.icarus.2008.12.044](https://doi.org/10.1016/j.icarus.2008.12.044) 7
- Fink, U., & Rubin, M. 2012, *Icarus*, 221, 803  
721, doi: [10.1016/j.icarus.2012.09.001](https://doi.org/10.1016/j.icarus.2012.09.001) 8

- 773 Fulle, M. 2000, *Icarus*, 145, 239, 818  
774 doi: [10.1006/icar.1999.6321](https://doi.org/10.1006/icar.1999.6321) 17 819
- 775 Fulle, M., Marzari, F., Della Corte, V., 820  
776 et al. 2016, *ApJ*, 821, 19, 821  
777 doi: [10.3847/0004-637X/821/1/19](https://doi.org/10.3847/0004-637X/821/1/19) 17 822
- 778 Gennaro, M., et al., eds. 2018, *WFC3* 823  
779 *Data Handbook*, v4.0 (Baltimore: Space 824  
780 Telescope Science Institute) 6 825
- 781 Ginsburg, A., Sipócz, B. M., Basseur, 826  
782 C. E., et al. 2019, *Astron. J.*, 157, 98, 827  
783 doi: [10.3847/1538-3881/aafc33](https://doi.org/10.3847/1538-3881/aafc33) 23 828
- 784 Giorgini, J. D., Yeomans, D. K., 829  
785 Chamberlin, A. B., et al. 1996, *Bull.* 830  
786 *Am. Astron. Soc.*, 28, 1158 (abstract) 831  
787 23 832
- 788 Graham, M. J., Kulkarni, S. R., Bellm, 833  
789 E. C., et al. 2019, *Publ. Astron. Soc.* 834  
790 *Pac.*, 131, 078001, 835  
791 doi: [10.1088/1538-3873/ab006c](https://doi.org/10.1088/1538-3873/ab006c) 3, 4 836
- 792 Green, S. F., McDonnell, J. A. M., 837  
793 McBride, N., et al. 2004, *J. Geophys.* 838  
794 *Res. (Planet.)*, 109, 12, 839  
795 doi: [10.1029/2004JE002318](https://doi.org/10.1029/2004JE002318) 17 840
- 796 Grün, E., Agarwal, J., Altobelli, N., et al. 841  
797 2016, *Mon. Not. R. Astron. Soc.*, 462, 842  
798 S220, doi: [10.1093/mnras/stw2088](https://doi.org/10.1093/mnras/stw2088) 19 843
- 799 Haberreiter, M., Schöll, M., Dudok de 844  
800 Wit, T., et al. 2017, *Journal of* 845  
801 *Geophysical Research (Space Physics)*, 846  
802 122, 5910, doi: [10.1002/2016JA023492](https://doi.org/10.1002/2016JA023492) 847  
803 7 848
- 804 Hughes, D. W. 1990, *Q. J. R. Astron.* 849  
805 *Soc.*, 31, 69 3 850
- 806 Ishiguro, M., Kuroda, D., Hanayama, H., 851  
807 et al. 2016, *Astron. J.*, 152, 169, 852  
808 doi: [10.3847/0004-6256/152/6/169](https://doi.org/10.3847/0004-6256/152/6/169) 22 853
- 809 Jeschke, E., Inagaki, T., & Kackley, R. 854  
810 2013, in *Astronomical Society of the* 855  
811 *Pacific Conference Series*, Vol. 475, 856  
812 *Astronomical Data Analysis Software* 857  
813 *and Systems XXII*, ed. D. N. Friedel, 858  
814 319 23 859
- 815 Jorda, L., Gaskell, R., Capanna, C., et al. 860  
816 2016, *Icarus*, 277, 257, 861  
817 doi: [10.1016/j.icarus.2016.05.002](https://doi.org/10.1016/j.icarus.2016.05.002) 20 862
- Joye, W. A., & Mandel, E. 2003, in 863  
*Astronomical Society of the Pacific* 864  
*Conference Series*, Vol. 295, 865  
*Astronomical Data Analysis Software*  
*and Systems XII*, ed. H. E. Payne, R. I.  
Jedrzejewski, & R. N. Hook, 489 23
- Kelley, M. S. P., & Lister, T. 2019,  
*calviacat: Calibrate star photometry by*  
*comparison to a catalog*, 1.0.2, Zenodo,  
doi: [10.5281/zenodo.2635841](https://doi.org/10.5281/zenodo.2635841) 5
- Kelley, M. S. P., Bodewits, D., Ye, Q.,  
et al. 2019, *Research Notes of the*  
*American Astronomical Society*, 3, 126,  
doi: [10.3847/2515-5172/ab3fb4](https://doi.org/10.3847/2515-5172/ab3fb4) 3, 6
- Kelley, M. S. P., Bodewits, D., Ye, Q.,  
et al. 2019, in *ASP Conf. Ser.*, Vol. 523,  
*ADASS XXVIII*, ed. P. J. Teuben,  
M. W. Pound, B. A. Thomas, & E. M.  
Warner, 471.  
[http://www.aspbooks.org/a/volumes/  
article\\_details/?paper\\_id=39288](http://www.aspbooks.org/a/volumes/article_details/?paper_id=39288) 4
- Kidger, M. R. 2004, *Astron. Astrophys.*,  
420, 389,  
doi: [10.1051/0004-6361:20035877](https://doi.org/10.1051/0004-6361:20035877) 3
- Kidger, M. R. 2008, *Comet*  
*46P/Wirtanen: Outburst*,  
[http://www.theastronomer.org/tacirc/  
2008/e2446.txt](http://www.theastronomer.org/tacirc/2008/e2446.txt) 3
- Knollenberg, J., Lin, Z. Y., Hviid, S. F.,  
et al. 2016, *Astron. Astrophys.*, 596,  
A89, doi: [10.1051/0004-6361/201527744](https://doi.org/10.1051/0004-6361/201527744)  
15
- Kokotanekova, R., Snodgrass, C.,  
Lacerda, P., et al. 2018, *MNRAS*, 479,  
4665, doi: [10.1093/mnras/sty1529](https://doi.org/10.1093/mnras/sty1529) 20
- Kronk, G. W., Meyer, M., & Seargent, D.  
A. J. 2017, *Cometography: A Catalog of*  
*Comets, Volume 6, 1983-1993*, Vol. 6 3
- LaHer, R. R., Gorjian, V., Rebull, L. M.,  
et al. 2012, *Publ. Astron. Soc. Pac.*,  
124, 737, doi: [10.1086/666883](https://doi.org/10.1086/666883) 23
- Lamy, P. L., Toth, I., Jorda, L., Weaver,  
H. A., & A'Hearn, M. 1998, *Astron.*  
*Astrophys.*, 335, L25 2
- Lara, L. M., Boehnhardt, H., Gredel, R.,  
et al. 2006, *Astron. Astrophys.*, 445,  
1151, doi: [10.1051/0004-6361:20053833](https://doi.org/10.1051/0004-6361:20053833)  
21



- Lin, Z.-Y., Lara, L. M., & Ip, W.-H. 2013, *Astron. J.*, 146, 4, doi: [10.1088/0004-6256/146/1/4](https://doi.org/10.1088/0004-6256/146/1/4) 20
- Masci, F. J., Laher, R. R., Rusholme, B., et al. 2019, *Publ. Astron. Soc. Pac.*, 131, 018003, doi: [10.1088/1538-3873/aae8ac](https://doi.org/10.1088/1538-3873/aae8ac) 4, 5
- McCully, C., Crawford, S., Kovacs, G., et al. 2018, doi: [10.5281/zenodo.1482019](https://doi.org/10.5281/zenodo.1482019) 23
- McDonnell, J. A. M., Evans, G. C., Evans, S. T., et al. 1987, *Astron. Astrophys.*, 187, 719 17
- Meech, K. J., Ageorges, N., A'Hearn, M. F., et al. 2005, *Science*, 310, 265, doi: [10.1126/science.1118978](https://doi.org/10.1126/science.1118978) 21
- Meech, K. J., A'Hearn, M. F., Adams, J. A., et al. 2011, *Astrophys. J., Lett.*, 734, L1, doi: [10.1088/2041-8205/734/1/L1](https://doi.org/10.1088/2041-8205/734/1/L1) 3, 20
- Merouane, S., Stenzel, O., Hilchenbach, M., et al. 2017, *Mon. Not. R. Astron. Soc.*, 469, S459, doi: [10.1093/mnras/stx2018](https://doi.org/10.1093/mnras/stx2018) 17
- Miles, R. 2016, *Icarus*, 272, 356, doi: [10.1016/j.icarus.2015.12.053](https://doi.org/10.1016/j.icarus.2015.12.053) 3
- Minor Planet Center. 2019, *Minor Planet Circulars*, 114607 5
- Mommert, M., Hora, J. L., Harris, A. W., et al. 2014, *Astrophys. J.*, 781, 25, doi: [10.1088/0004-637X/781/1/25](https://doi.org/10.1088/0004-637X/781/1/25) 23
- Pajola, M., Höfner, S., Vincent, J. B., et al. 2017, *Nature Astronomy*, 1, 0092, doi: [10.1038/s41550-017-0092](https://doi.org/10.1038/s41550-017-0092) 3, 19
- Ponomarenko, V., Chiuryumov, K., & Sergeev, O. 2018, *Bulletin of Taras Shevchenko National University of Kyiv Astronomy*, 58, 6 7
- Price, M. C., Kearsley, A. T., Burchell, M. J., et al. 2010, *Meteoritics and Planetary Science*, 45, 1409, doi: [10.1111/j.1945-5100.2010.01104.x](https://doi.org/10.1111/j.1945-5100.2010.01104.x) 17
- Prša, A., Harmanec, P., Torres, G., et al. 2016, *Astron. J.*, 152, 41, doi: [10.3847/0004-6256/152/2/41](https://doi.org/10.3847/0004-6256/152/2/41) 16, 18
- Rinaldi, G., Bockelée-Morvan, D., Ciarniello, M., et al. 2018, *Mon. Not. R. Astron. Soc.*, 481, 1235, doi: [10.1093/mnras/sty2266](https://doi.org/10.1093/mnras/sty2266) 15
- Schleicher, D. G., & Bair, A. N. 2011, *Astron. J.*, 141, 177, doi: [10.1088/0004-6256/141/6/177](https://doi.org/10.1088/0004-6256/141/6/177) 8
- Snodgrass, C., Fitzsimmons, A., Lowry, S. C., & Weissman, P. 2011, *Mon. Not. R. Astron. Soc.*, 414, 458, doi: [10.1111/j.1365-2966.2011.18406.x](https://doi.org/10.1111/j.1365-2966.2011.18406.x) 2
- Snodgrass, C., A'Hearn, M. F., Aceituno, F., et al. 2017, *Philosophical Transactions of the Royal Society of London Series A*, 375, 20160249, doi: [10.1098/rsta.2016.0249](https://doi.org/10.1098/rsta.2016.0249) 21
- Sosey, M. 2017, *Imexam Version 0.8.0 Release, 0.8.0, Zenodo*, doi: [10.5281/zenodo.1042809](https://doi.org/10.5281/zenodo.1042809) 23
- Steckloff, J., & Melosh, H. J. 2016, in *AAS/Division for Planetary Sciences Meeting Abstracts #48, AAS/Division for Planetary Sciences Meeting Abstracts*, 206.06 3
- Thomas, P., A'Hearn, M., Belton, M. J. S., et al. 2013a, *Icarus*, 222, 453, doi: [10.1016/j.icarus.2012.02.037](https://doi.org/10.1016/j.icarus.2012.02.037) 20
- Thomas, P. C., A'Hearn, M. F., Veverka, J., et al. 2013b, *Icarus*, 222, 550, doi: [10.1016/j.icarus.2012.05.034](https://doi.org/10.1016/j.icarus.2012.05.034) 20
- Tonry, J. L., Denneau, L., Heinze, A. N., et al. 2018a, *Publ. Astron. Soc. Pac.*, 130, 064505, doi: [10.1088/1538-3873/aabadf](https://doi.org/10.1088/1538-3873/aabadf) 3
- Tonry, J. L., Denneau, L., Flewelling, H., et al. 2018b, *Astrophys. J.*, 867, 105, doi: [10.3847/1538-4357/aae386](https://doi.org/10.3847/1538-4357/aae386) 5, 22
- Tubiana, C., Snodgrass, C., Bertini, I., et al. 2015, *Astron. Astrophys.*, 573, A62, doi: [10.1051/0004-6361/201424735](https://doi.org/10.1051/0004-6361/201424735) 17
- van Dokkum, P. G. 2001, *Publ. Astron. Soc. Pac.*, 113, 1420, doi: [10.1086/323894](https://doi.org/10.1086/323894) 6
- Vincent, J.-B., Farnham, T., Kühr, E., et al. 2019, *Space Sci. Rev.*, 215, 30, doi: [10.1007/s11214-019-0596-8](https://doi.org/10.1007/s11214-019-0596-8) 22

- 961 Vincent, J.-B., A’Hearn, M. F., Lin, 972  
962 Z.-Y., et al. 2016, Mon. Not. R. Astron. 973  
963 Soc., 462, S184,  
964 doi: [10.1093/mnras/stw2409](https://doi.org/10.1093/mnras/stw2409) 15, 19, 20<sup>974</sup>
- 965 Vincent, J. B., Hviid, S. F., Mottola, S., 975  
966 et al. 2017, MNRAS, 469, S329,  
967 doi: [10.1093/mnras/stx1691](https://doi.org/10.1093/mnras/stx1691) 20, 21, 22<sup>976</sup>
- 968 Willmer, C. N. A. 2018, The 977  
969 Astrophysical Journal Supplement  
970 Series, 236, 47,  
971 doi: [10.3847/1538-4365/aabfdf](https://doi.org/10.3847/1538-4365/aabfdf) 7
- Wolf, C., Onken, C. A., Luvaul, L. C.,  
et al. 2018, PASA, 35, e010,  
doi: [10.1017/pasa.2018.5](https://doi.org/10.1017/pasa.2018.5) 22
- Yoshida, S. 2013, 46P/Wirtanen (2002),  
[http://www.aerith.net/comet/catalog/  
0046P/2002.html](http://www.aerith.net/comet/catalog/0046P/2002.html) 3