



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/214273/>

Version: Published Version

---

**Article:**

Blaylock-Squibbs, G.A., Parker, R.J. and Daffern-Powell, E.C. (2024) No signature of the birth environment of exoplanets from their host stars' Mahalanobis phase space. *The Astrophysical Journal*, 968 (2). 108. ISSN: 0004-637X

<https://doi.org/10.3847/1538-4357/ad4be0>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



# No Signature of the Birth Environment of Exoplanets from Their Host Stars' Mahalanobis Phase Space

George A. Blaylock-Squibbs<sup>1</sup>, Richard J. Parker<sup>1</sup>, and Emma C. Daffern-Powell

Department of Physics and Astronomy, The University of Sheffield, Hounsfield Road, Sheffield S3 7RH, UK; [gablaylock-squibbs1@sheffield.ac.uk](mailto:gablaylock-squibbs1@sheffield.ac.uk)

Received 2024 March 6; revised 2024 May 13; accepted 2024 May 13; published 2024 June 18

## Abstract

The architectures of extrasolar planetary systems often deviate considerably from the “standard” model for planet formation, which is largely based on our own solar system. In particular, gas giants on close orbits are not predicted by planet formation theory and so some processes are thought to move the planets closer to their host stars. Recent research has suggested that hot-Jupiter host stars display a different phase space compared to stars that do not host hot Jupiters. This has been attributed to these stars forming in star-forming regions of high stellar density, where dynamical interactions with passing stars have perturbed the planets. We test this hypothesis by quantifying the phase space of planet-hosting stars in dynamical  $N$ -body simulations of star-forming regions. We find that stars that retain their planets have a higher phase space than nonhosts, regardless of their initial physical density. This is because an imprint of the kinematic substructure from the regions birth is retained, as these stars have experienced fewer and less disruptive encounters than stars whose planets have been liberated and become free-floating. However, host stars whose planets remain bound but have had their orbits significantly altered by dynamical encounters are also primarily found in high phase space regimes. We therefore corroborate other research in this area that has suggested the high phase space of hot-Jupiter host stars is not caused by dynamical encounters or stellar clustering, but rather reflects an age bias in that these stars are (kinematically) younger than other exoplanet host stars.

*Unified Astronomy Thesaurus concepts:* Star forming regions (1565); Exoplanet migration (2205); Stellar dynamics (1596); N-body simulations (1083)

## 1. Introduction

Observations and theory have shown that stars form in groups with tens, to tens of thousands, of stellar siblings, with a wide range of stellar densities that exceed those in the Galactic disk by several orders of magnitude (Lada & Lada 2003; Bressert et al. 2010).

Star-forming regions exhibit spatial and kinematic substructure (Gomez et al. 1993; Larson 1995; Cartwright & Whitworth 2004; Sánchez & Alfaro 2009; André et al. 2014), which is erased over time through dynamical evolution (Klessen & Kroupa 2001; Parker & Meyer 2012), sometimes solely two-body relaxation, or a combination of this and violent relaxation.

Furthermore, observations show that planet formation occurs contemporaneously with star formation (Haisch et al. 2001; Alves et al. 2020), such that the star formation environment may strongly influence the planet formation process, depending on the density of the star-forming region and how long the planets' host stars spend in the region (Parker 2020).

In the lowest-density star-forming regions ( $\bar{\rho} \sim 10 \text{ s } M_{\odot} \text{ pc}^{-3}$ ), if massive stars are present (e.g., in OB associations) then photoionizing radiation will lead to the evaporation of gas (and to a lesser extent, dust) from the disks (Sclay & Clarke 2001; Haworth et al. 2018a, 2018b), possibly hindering the formation of gas giant planets (Clarke & Pringle 1991; Nicholson et al. 2019;

Concha-Ramírez et al. 2019) and altering the structure of the dust in the disks (Sellek et al. 2020).

In moderately dense star-forming regions ( $\bar{\rho} \sim 100 \text{ s } M_{\odot} \text{ pc}^{-3}$ , which most nearby star-forming regions are at least as dense, if not more so at birth, Parker & Alves de Oliveira 2017; Parker & Schoettler 2022) planetary orbits can be disrupted. This can be either a relatively minor change in their semimajor axis and/or eccentricity, or a more energetic interaction that leads to the planets becoming unbound from their original host stars and becoming a free-floating within the star-forming region (Kobayashi & Ida 2001; Daffern-Powell et al. 2022).

In the most dense star-forming regions ( $\bar{\rho} > 1000 \text{ s } M_{\odot} \text{ pc}^{-3}$ ) protoplanetary disks can even be truncated by interactions with passing stars (Sclay & Clarke 2001; Olczak et al. 2006; Steinhausen et al. 2012; Vincke & Pfalzner 2016; Winter et al. 2018), potentially limiting the material available for planet formation in the outer disks.

A significant unanswered question in astrophysics is to what degree does the star-forming environment shape the architectures of planetary systems? For instance, are hot Jupiters (Dawson & Johnson 2018) formed from the migration of planets in disks (Masset & Papaloizou 2003), or are they produced later via dynamical instabilities in the planetary systems caused by interactions with other stars (Wu & Murray 2003; Benkendorff et al. 2024; Li et al. 2024)?

However, most exoplanet host stars (and indeed our own Sun) are main-sequence stars in the Galactic field, and are likely to have migrated from their birth star-forming regions. There appears to be little prospect of identifying the dynamical signatures of planet host stars' birth environments.

However, Winter et al. (2020) used Gaia observations to build up a 6D phase space (i.e., position and velocity)

<sup>1</sup> Royal Society Dorothy Hodgkin fellow.

distribution for exoplanet host stars, and quantified the phase space around stars using the dimensionless Mahalanobis (1936) distance. Winter et al. (2020) then use the Mahalanobis distance to define a phase space “density.”<sup>2</sup>

After finding the relative Mahalanobis “densities” of exoplanet host stars in their data, Winter et al. (2020) categorize each exoplanet host star according to its local Mahalanobis phase space; either low, ambiguous, or high. In a low Mahalanobis phase space nearby stars will have very different velocities and Winter et al. (2020) call these “field stars;” though technically all of the stars in their samples are Galactic field stars. In a high Mahalanobis phase space stars will exhibit clustering in their position–velocity phase space, and Winter et al. (2020) suggest that the high Mahalanobis phase space regime indicates the imprint of a dense, clustered birth environment.

Winter et al. (2020) find that hot-Jupiter host stars are predominantly found in high Mahalanobis phase space, which they interpret as being due to them forming in high-density star-forming regions, where interactions with passing stars led to the formation of hot Jupiters. However, several authors (Adibekyan et al. 2021; Mustill et al. 2022) have contended that the high Mahalanobis phase space is due to an age bias. Hot Jupiters tend not to survive due to orbital decay and photoevaporation of their atmospheres (Hamer & Schlaufman 2019; Chen et al. 2023). Younger stars tend to be kinematically “cooler,” which results in the stars being closer to each other in the 6D Mahalanobis phase space.

In this work we investigate whether stars whose planetary systems are perturbed in their birth star-forming regions retain an imprint of the initial stellar clustering. We follow the evolution of the 6D Mahalanobis phase space for sets of  $N$ -body simulations where the stars host Jupiter-mass planets. We then look for correlations between stars in high Mahalanobis phase space and the orbital architectures of their planetary systems. The paper is organized as follows. We outline our methods in Section 2, we present our results in Section 3, and we provide a discussion in Section 4. We conclude in Section 5.

## 2. Methods

In this section we summarize the setup of the  $N$ -body simulations and then discuss how we quantify the 6D Mahalanobis phase space metric.

### 2.1. Simulation Setup

We use four sets of 20 subvirial (virial ratio  $\alpha_{\text{vir}} = 0.3$ , where  $\alpha_{\text{vir}} = 0.5$  is virial equilibrium) simulations from Daffern-Powell et al. (2022) containing 1000 stars drawn from a Maschberger (2013) IMF. 500 of these stars are randomly assigned a single Jupiter-mass planet at either 5 or 30 au. The orbital eccentricities and inclinations are all set to zero. The average mass of host stars in the simulations is  $\sim 0.5 M_{\odot}$ , which is lower than the observed mean hot-Jupiter host star mass of  $1 M_{\odot}$  (Schneider et al. 2011). There is a small dependence of planet perturbation on stellar mass in these simulations (Parker 2023, who find that planets are more likely to be disrupted around M dwarfs than G dwarfs). However, this

<sup>2</sup> This phase space “density” is not the physical volume or surface density of the stars, but the two may be related as postulated by Winter et al. (2020).

**Table 1**  
Table Summarizing the Simulations Used in This Work

$D_f$	$\bar{\rho}$ ( $M_{\odot} \text{pc}^{-3}$ )	$\alpha_{\text{vir}}$	$N_*$	$a_i$ (au)
1.6	$10^2$	0.3	1000	5
1.6	$10^2$	0.3	1000	30
1.6	$10^4$	0.3	1000	5
1.6	$10^4$	0.3	1000	30

**Note.** From left to right the columns are the fractal dimension  $D_f$ , the median local stellar density of the simulations  $\bar{\rho}$ , the virial ratio  $\alpha_{\text{vir}}$ , the number of stars  $N_*$  and the initial semimajor axis of the Jupiter-mass planets’ initial orbits,  $a_i$ .

subtle difference (of order 10%) would be difficult to observe in reality.

To test if different densities of the simulations results affects the numbers of planet-hosting stars in high Mahalanobis phase space regimes, we use low and high-density simulations with local median volume densities of  $\sim 10^2 M_{\odot} \text{pc}^{-3}$  and  $\sim 10^4 M_{\odot} \text{pc}^{-3}$ , respectively.

For additional information on the simulations see Daffern-Powell et al. (2022). We summarize the simulations in Table 1.

#### 2.1.1. Dynamical Evolution

The simulations are evolved using the *kira* integrator, a part of the *starlab* software package (Portegies Zwart et al. 1999, 2001). The *kira* integrator uses a fourth-order Hermite scheme with individual time steps. We limit our planetary systems to one Jupiter-mass planet and therefore treat the star-planet systems as a binary (Portegies Zwart et al. 1999, 2001).

#### 2.1.2. Dynamical Timescales

The simulations are run for 10 Myr which allows for sufficient dynamical evolution and relaxation to occur. The two stellar density regimes in the simulations result in different crossing times, which in turn results in different relaxation times. The high-density ( $\bar{\rho} = 10^4 M_{\odot} \text{pc}^{-3}$ ) regions have shorter crossing times of around 0.1 Myr, whereas the low-density ( $\bar{\rho} = 10^2 M_{\odot} \text{pc}^{-3}$ ) regions have crossing times on the order of 2 Myr. Since both the crossing times for the simulations are well below 10 Myr we are able to simulate the dynamical interactions between stars and planets to suitable degree. The longer crossing time for the low-density regions results in the spatial and kinematic substructure being present for longer, although it is still erased within 5 Myr.

Our choice of 10 Myr for the end time of simulations is also informed by observations that highlight the paucity of star-forming regions remaining bound beyond 10 Myr (Lada & Lada 2003; Kruijssen 2012).

## 2.2. Mahalanobis Phase Space Metric

The Mahalanobis phase space “density” metric uses the Mahalanobis distance introduced by Mahalanobis (1936) to measure the distance of a point from a distribution. The Mahalanobis distance allows for the comparison of high-dimensional data sets (e.g., position–velocity phase space is 6D) in which each parameter can be very different in physical scale from one another (i.e., positions measured in parsecs and velocities in kilometers per second). The Mahalanobis distance

between two points (in this case stars) is

$$m_d(\vec{x}, \vec{y}) = \sqrt{(\vec{x} - \vec{y})^T \mathbf{S}^{-1} (\vec{x} - \vec{y})}, \quad (1)$$

where  $\vec{x}$  and  $\vec{y}$  are vectors containing the position and velocity information for two stars and  $\mathbf{S}^{-1}$  is the inverse of the covariance matrix. By multiplying the data set by  $\mathbf{S}^{-1}$  it is variance normalized and the units are removed. This rescaling means that moving one unit along on one axis in 6D Mahalanobis space is the same as moving one unit along any other axis in the 6D Mahalanobis space.

Winter et al. (2020) define the Mahalanobis phase space “density” (we hereon refer to this as a “metric” rather than a “density” to avoid confusion with the volume density of the stars) as

$$\rho_{m,N} = N m_{d,N}^{-D_p}, \quad (2)$$

where  $N$  is the nearest neighbor number,  $m_{d,N}$  is the Mahalanobis distance to the  $N$ th nearest neighbor and  $D_p$  is the number of dimensions (parameters; Winter et al. 2020).

We calculate this Mahalanobis phase space metric in two different ways in this work. First we calculate the Mahalanobis phase space metric for all stars with respect to all other stars and then we calculate the relative Mahalanobis phase space metric of host stars using the neighborhood method from Winter et al. (2020).

The neighborhood method proceeds as follows. For each host star in turn we define a large neighborhood (80 pc) and a local neighborhood (40 pc). We check that the host star has at least 400 other stars within 40 pc of it; if there is not the host star is excluded from the analysis. If there are more than 600 stars within 40 pc we randomly pick 600 stars within the local neighborhood of the host star. Then we calculate the Mahalanobis phase space metric for the host star with respect to all other stars within the local neighborhood (not just the 400/600 found in the previous step). Then for each of the chosen 400/600 stars in the host stars local neighborhood we calculate their Mahalanobis phase space metrics with respect to all other stars within 40 pc of them. Once the Mahalanobis phase space metric is calculated for all the chosen stars around the host star the results are median normalized to enable direct comparisons between different neighborhoods.

### 2.2.1. Mahalanobis Phase Space Regimes

For each type of neighborhood (i.e., population of stars as defined above) we split them into a low and high phase space regime following Winter et al. (2020).

To calculate if a exoplanet host star is in a low, ambiguous, or high Mahalanobis phase space regime we fit two lognormals using Gaussian mixture modeling, as done by Winter et al. (2020). We then calculate the probability,  $P_{\text{high}}$ , that the host belongs to the high phase space regime. We use the same thresholds as Winter et al. (2020) where a  $P_{\text{high}} < 0.16$  corresponds to a low phase space regime,  $P_{\text{high}} > 0.86$  corresponds to a high phase space regime and anything between these values is said to belong to an ambiguous phase space regime. The threshold values are chosen to be one standard deviation apart.

Unlike the Winter et al. (2020) work we do not test if the simulations can be described using a single lognormal, due to the fact that our regimes are being defined for single star-forming regions. We have tried creating a Galactic disk population model by combining the outputs of different

simulations. However, the results were not significantly different to a single stellar population from each individual  $N$ -body simulation.

We show four examples of single host stars taken from simulations with different initial conditions in Figure 1. Figure 1 shows the distribution of Mahalanobis phase space metrics calculated for host stars and 600 randomly picked stars within 40 pc. These plots show examples of host stars in low (left-hand column) and high (right-hand column) phase space regimes for simulations with planets at 30 au (top row) and 5 au (bottom row). The gray dashed-dotted line shows the Mahalanobis phase space metric for the host star. The black solid line shows the Gaussian mixture model consisting of two lognormals, with one corresponding to a low phase space regime and the other a high phase space regime.

## 3. Results

In this section we present the Mahalanobis phase space evolution of  $N$ -body simulations of 1000 stars with  $\sim 500$  planets with initial semimajor axes of 5 or 30 au. We then show the number of times host stars are found in either the low, ambiguous, or high Mahalanobis phase space regimes at  $t = 0$  Myr and  $t = 10$  Myr. We then show the number of perturbed host stars in each of the phase space regimes at  $t = 10$  Myr.

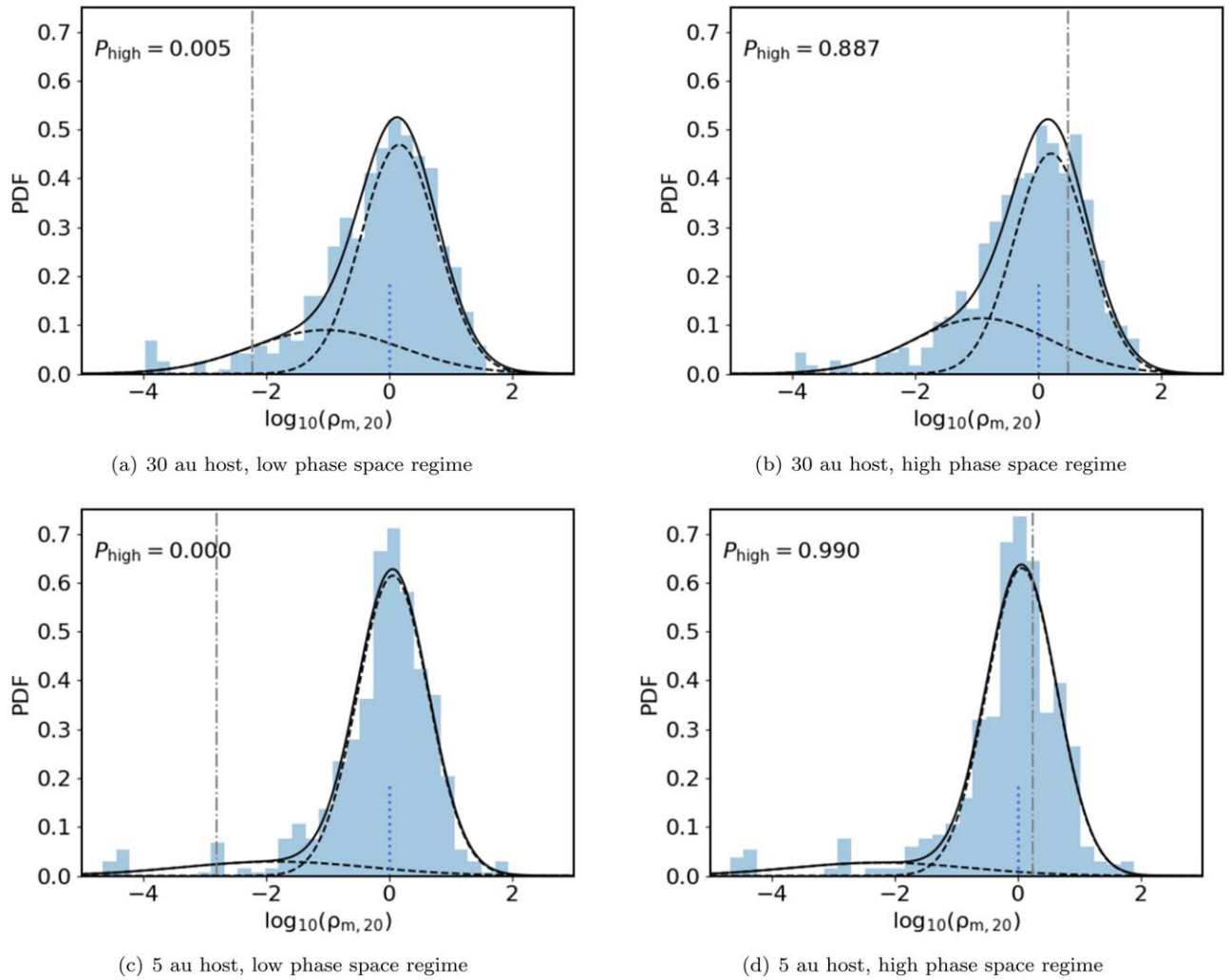
### 3.1. Definitions

In this work we define planet “host” stars as being stars that are mutual nearest neighbors with a planet and that the total binding energy of the star–planet system is negative, meaning they are gravitationally bound. We define “former hosts” as stars that originally hosted a planet, but no longer do so in the current snapshot. We define “perturbed” hosts as stars whose planet’s semimajor axis has changed by more than 10% of its initial value.

### 3.2. Mahalanobis Phase Space Evolution

In Figure 2 we show the evolution of the mean 6D (position and velocity information) Mahalanobis phase space metric (as defined by Equation (2)) of nonhost stars (stars with no gravitationally bound planet), host stars (stars with at least one gravitationally bound planet), and former host stars (stars that have lost their planet). Figure 2(a) shows the evolution of the mean Mahalanobis phase space metric for high initial density ( $\bar{\rho} \sim 10^4 M_{\odot} \text{ pc}^{-3}$ ) simulations with planets initially on 30 au orbits. The host stars (blue dashed line) have a higher Mahalanobis phase space metric, whereas the nonhosts and former hosts have a slightly lower phase space metric. However, the uncertainties between different realizations of the same initial conditions overlap to such an extent that observing one snapshot in time would make differentiating between host, nonhost, and former host stars difficult.

In Figure 2(b) we show the evolution of the Mahalanobis phase space metric for the initially low density ( $\bar{\rho} \sim 10^2 M_{\odot} \text{ pc}^{-3}$ ), but where the planets are again on 30 au orbits. In these simulations the former host stars initially have much lower mean Mahalanobis phase space metrics. We attribute this to the effects of low number statistics as early on in the simulations there are no former hosts, because interactions that disrupt planetary systems take longer to occur in these low-density simulations.



**Figure 1.** Mahalanobis phase space metric distributions around host stars in different simulations. The gray dashed–dotted line shows the Mahalanobis phase space metric for the host star. The black solid line shows the Gaussian mixture model consisting of two lognormals, with one corresponding to a low phase space regime and the other a high phase space regime. The dashed black lines show the component Gaussians and the blue dotted line which is always centered on zero highlights the median of the distribution after normalization.

We now repeat the above analyzes but for simulations where the planets are initially all on 5 au orbits. Figure 2(c) shows the evolution of the Mahalanobis phase space metric for the high-density simulations. There are slight differences in the evolution of the metric depending on whether we are looking at planet hosts, nonhosts, or former hosts, with the former hosts having the lowest phase space metric values. However, there is significant overlap in the uncertainties in the metrics between the host, nonhost, and former host stars.

In Figure 2(d) we show the evolution of the mean Mahalanobis phase space metric for the low-density simulations. There is no significant difference between the evolution of the Mahalanobis phase space metric for nonhost and host stars, but there is a difference in the evolution of the metric for the former hosts, in that the former host stars consistently have lower minimum mean Mahalanobis phase space metrics.

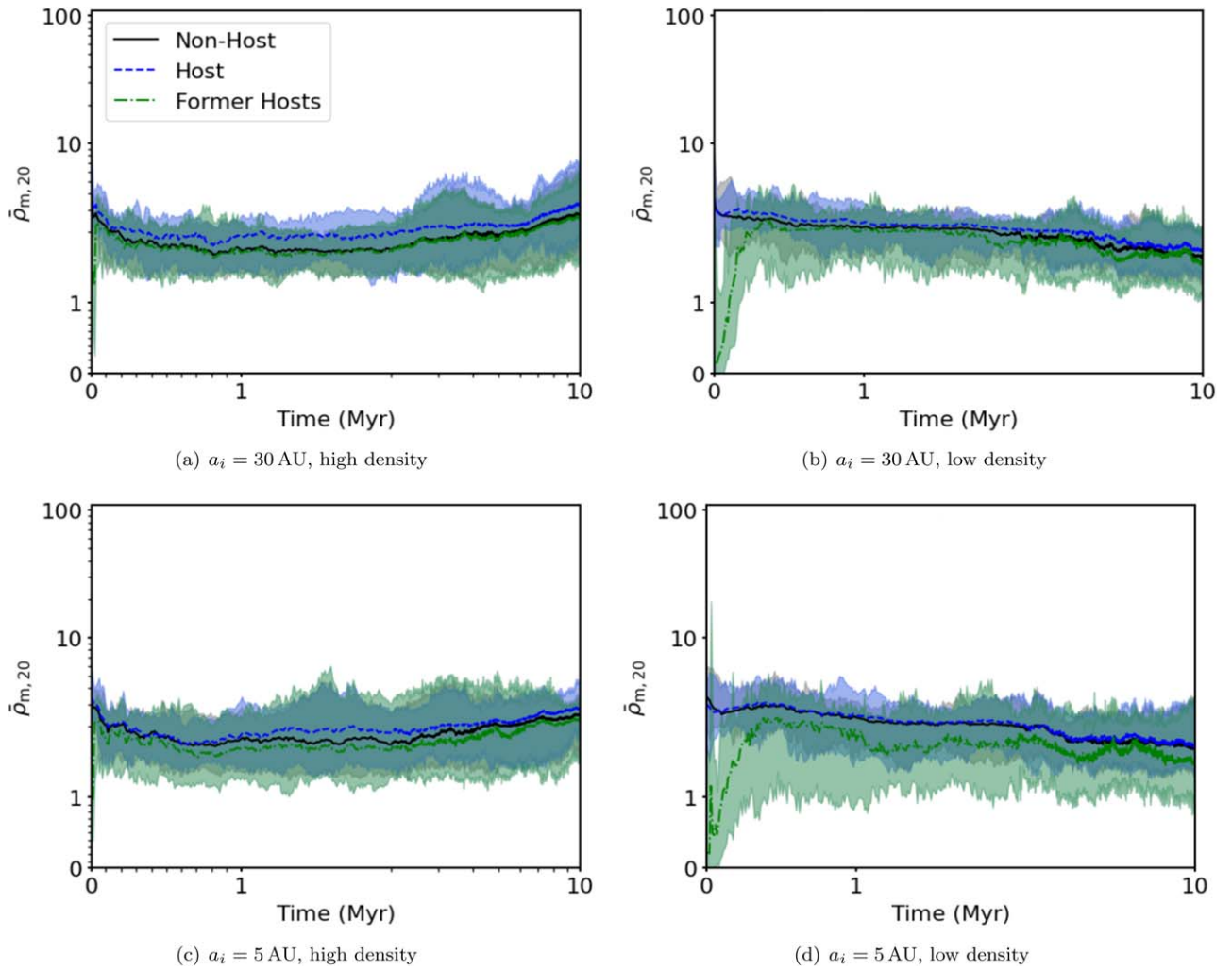
We interpret the lower Mahalanobis phase space metric for the former host stars as being due to is due to the kinematics of the former host stars.

In the low-density regions stars will undergo fewer dynamical interactions, thus the velocities of the majority of stars (including the planet hosts) are unlikely to be significantly

altered. However, planets initially on (say) 5 au orbits will be strongly bound to their stars, and therefore any interaction that is energetic enough to remove that planet will significantly alter the velocity of that star. We believe this is the reason for the former hosts displaying slightly lower Mahalanobis phase space metrics in the low-density simulations (panels (b) and (d) of Figure 2), and the high-density simulations where the planets are initially tightly bound (Figure 2(c)).

In the high-density simulations where the planets are on more weakly bound orbits (Figure 2(a)), more stars undergo energetic interactions, and these interactions are more likely to disrupt planets on 30 au orbits. For this reason, the Mahalanobis phase space metric is not significantly different for the former host stars.

The difference in the kinematics for planets in different density regions is further highlighted by the mean number of former hosts in each set of simulations. For the simulation sets with planets at 30 au we find that by the end of the simulations ( $t = 10$  Myr) there are 169 and 308 former hosts for low- and high-stellar-density regions, respectively. We see a similar trend for the simulation sets with planets initially at 5 au, with 76 and 163 former hosts at the end of the simulations for low



**Figure 2.** The evolution of the Mahalanobis phase space metric against time for the four sets of simulations. The shaded gray, blue, and green areas show the range of mean Mahalanobis phase space metrics calculated across 20 simulations for nonhost stars, host stars, and former host stars, and the solid black, blue dashed, and green dotted–dashed lines show the median values for the 20 simulations. Simulations with planets with initial semimajor axes of 30 au are shown in the top row and simulations with planets with initial semimajor axes of 5 au are shown on the bottom row. The left-hand column shows the high initial stellar density ( $\bar{\rho} \sim 10^4 M_{\odot} \text{pc}^{-3}$ ) simulations and the right-hand column shows the low initial stellar density ( $\bar{\rho} \sim 10^2 M_{\odot} \text{pc}^{-3}$ ) simulations.

and high stellar densities, respectively. The lower number of former hosts for planets initially at 5 au in both low and high stellar densities is due to planets at 5 au being more strongly gravitationally bound to their host stars, making them harder to remove via external interactions with other stars (Heggie 1975; Hills 1975; Fregeau et al. 2006; Parker & Reggiani 2013).

This result is only discernible because we know everything about our simulations, i.e., we know exactly which stars have lost planets in dynamical interactions. This is not possible observationally; however a statistical method could be developed to estimate the number of stars likely to have lost planets, but this method would require assumptions about the typical planetary system architecture and the initial density of the star-forming region.

### 3.3. Host Star Mahalanobis Phase Space Regimes

We now examine the Mahalanobis phase space metric distribution in more detail at  $t = 0$  Myr (before any dynamical evolution has taken place) and  $t = 10$  Myr (the end of our simulations). Based on their Mahalanobis phase space metric,

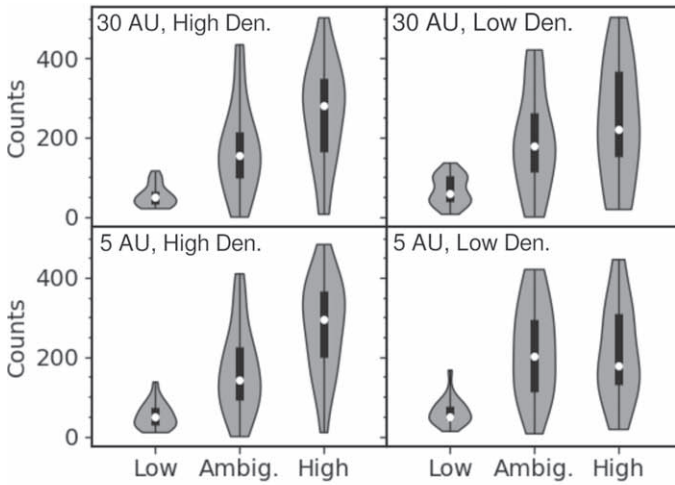
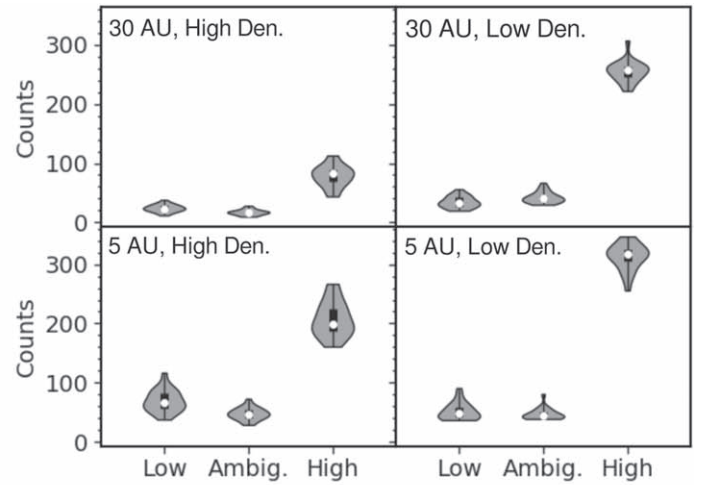
we divide the stars into three regimes (as detailed in Section 2.2.1): low, ambiguous, and high (phase space metric).

We show the initial and final Mahalanobis phase space metric distributions in Figure 3 of host stars across the 20 simulations for each set of stars that are in either low, ambiguous, or high Mahalanobis phase space regimes.

We find that initially most host stars are either in the ambiguous or high phase space regime (Figure 3(a)), with a similar number of hosts in both of these regimes. This is expected as these simulated star-forming regions are initially highly spatially and kinematically substructured. This means that stars close to each other in physical space have small velocity dispersions, which leads to a high Mahalanobis phase space metric.

We show the final ( $t = 10$  Myr) distributions in Figure 3(b). Most host stars across all four sets of simulations are in the high Mahalanobis phase space regime.

We interpret the far higher numbers of host stars in the high Mahalanobis phase space regime at 10 Myr as being due to stars that remain hosts not having undergone as much dynamical processing as former hosts. If they had undergone dynamical encounters, they would likely have lost their planet, becoming former hosts. Conversely, the general population of stars will


 (a) Host star phase space regimes at  $t = 0$  Myr

 (b) Host star phase space regimes at  $t = 10$  Myr

**Figure 3.** Violin plots of the number of host stars in either low, ambiguous, or high Mahalanobis phase space regimes at 0 Myr (panel (a)) or 10 Myr (panel (b)). These plots show the counts across all 20 simulations in the four different sets we use. The white dot is the median count, the black bar is the interquartile range, the thin black line shows the full range of values and the thickness of the plot corresponds to the probability of finding a particular value. The top left plot shows the counts for hosts with planets initially at 30 au in high-stellar-density regions ( $\bar{\rho} \sim 10^4 M_{\odot} \text{pc}^{-3}$ ), the top right plot shows the same but for low-stellar-density regions ( $\bar{\rho} \sim 10^2 M_{\odot} \text{pc}^{-3}$ ). The bottom left plot shows the counts for hosts with planets initially at 5 au in high-stellar-density regions, and the bottom right plot shows the same but for low-stellar-density regions.

have velocities that have been changed due to dynamical interactions which will manifest as an overall decrease in the Mahalanobis phase space metric. This results in more host stars residing in the high Mahalanobis phase space regimes.

### 3.3.1. Host Stars with Perturbed Planets

We now plot the Mahalanobis phase space metric distributions for host stars whose planets have been perturbed (i.e., the semimajor axis has changed by  $\pm 10\%$  of its initial value, after 10 Myr of dynamical evolution. Figure 4 shows the number of perturbed hosts in either the low, ambiguous, or high Mahalanobis phase space regime.

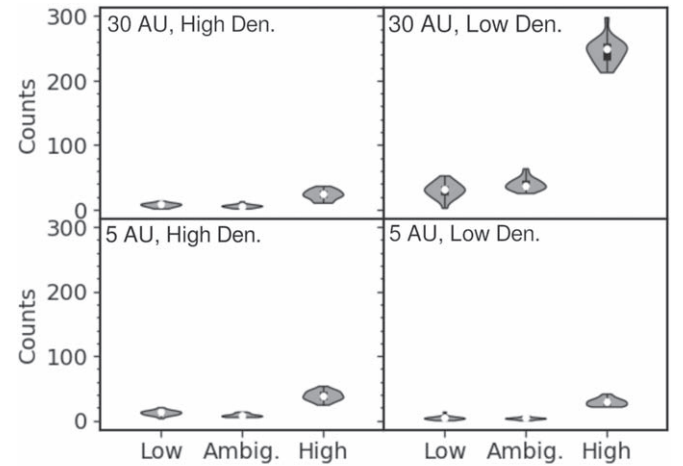
In all sets of simulations, the trends are similar to the distributions for all planet host stars we present in Figure 3(b). This is to be expected, as the perturbed hosts represent a subset of all of the host stars.

However, we see far more perturbed hosts in the high Mahalanobis phase space regimes in low-density ( $\bar{\rho} \sim 10^2 M_{\odot} \text{pc}^{-3}$ ) simulations where the planets have initial semimajor axes of 30 au, than in any other simulations. There are fewer perturbed hosts in the high-density simulations due to the interactions in the simulations being more likely to produce free-floating planets, whereas the low-density simulations are less likely to remove planets and instead just perturb them (this can be seen in the difference between distributions in the top left and top right of Figure 4).

Conversely, there is very little difference in the Mahalanobis phase space distributions when the host stars' planets are at 5 au; this is due to very few planets being perturbed at low stellar densities.

## 4. Discussion

Our main result is that the Mahalanobis phase space “density”—a metric that incorporates information on the position and velocity of each star and provides a dimensionless distance from the average in a distribution—of planet host stars only weakly depends on the dynamical history of the star-


 (a) Perturbed host star phase space regimes  $t = 10$  Myr

**Figure 4.** Violin plots of the number of perturbed host stars (stars whose planets' semimajor axes have changed by  $\pm 10\%$ ) after 10 Myr of dynamical evolution in either low, ambiguous, or high Mahalanobis phase space regimes across all 20 simulations. In all panels, the white dot is the median count, the black bar is the interquartile range, the thin black line shows the full range of values and the thickness of the plot corresponds to the probability of finding a particular value. The top left plot shows the counts for hosts with planets initially at 30 au in high-stellar-density ( $\bar{\rho} \sim 10^4 M_{\odot} \text{pc}^{-3}$ ) regions, top right shows the same but for low-stellar-density ( $\bar{\rho} \sim 10^2 M_{\odot} \text{pc}^{-3}$ ) regions. The bottom left plot shows the counts for hosts with planets initially at 5 au in high-stellar-density regions, and bottom right shows the same but for low-stellar-density regions.

planet system. When we plot the evolution of the Mahalanobis phase space metric over time, stars whose planets have been liberated by dynamical interactions have a slightly lower Mahalanobis phase space metric, but one that is within the spread of the metric for stars who still retain their planets, and stars that have never hosted planets.

When we analyze the Mahalanobis phase space metric for “neighborhoods” of stars (similar to the approach in Winter et al. 2020), we see a clear preference for planet host stars to

reside in areas where the phase space metric (“density”) is high, regardless of the initial *physical* density of the star-forming region. Rather than the physical density, the high phase space metric is actually dominated by the kinematics of the star-forming region.

Star-forming regions form with spatial and kinematic substructure (Larson 1981; Gomez et al. 1993; Cartwright & Whitworth 2004; Sánchez & Alfaro 2009; Alfaro & González 2016), and our simulations are designed to mimic this before any dynamical evolution takes place. For this reason, more stars reside in a high or ambiguous Mahalanobis phase space before dynamical evolution. Once dynamical evolution has occurred, the majority of stars that retain planets reside in high Mahalanobis phase space (Figure 3). This is because stars that do not undergo interactions that would significantly alter the local kinematic substructure are likely to retain their planets.

Winter et al. (2020) interpret the high Mahalanobis phase space densities they find around hot-Jupiter host stars as evidence that these objects formed in a high-physical-density environment, which has caused the formation of the hot Jupiters from dynamical processes, such as a direct encounter with a passing stars (Parker & Quanz 2012; Benkendorff et al. 2024), planet–planet scattering (Raymond et al. 2009) or from von Zeipel–Lidov–Kozai cycles (Wu & Murray 2003; Malmberg et al. 2007; Parker & Goodwin 2009).

However, we find that planet host stars that do form in dense stellar environments will have a high Mahalanobis phase space metric simply because the kinematic signature from their birth environment is retained. To this end, we expect planet host stars in all regions where there is some degree of kinematic coherence (e.g., young moving groups; Zuckerman & Song 2004; Gagné et al. 2015; Schneider et al. 2019) to exhibit a high Mahalanobis phase space metric.

Planets that are perturbed by dynamical encounters with passing stars, but remain in orbit around their parent star, will attain higher-than-average Mahalanobis phase space metric values, but not significantly more so than planets that have not been perturbed. Again, this is due to the Mahalanobis phase space metric being dominated by the velocities of the stars, rather than the spatial density of their birth environment.

We therefore cannot determine the dynamical history of a hot-Jupiter host star using its local Mahalanobis phase space metric, and the question of whether hot Jupiters form due to disk migration processes or external dynamical encounters remains open.

Our results appear complementary to those those of Adibekyan et al. (2021) and Mustill et al. (2022), who posit that a high Mahalanobis phase space metric is indicative of differences in the kinematics of hot-Jupiter host stars compared to stars without hot Jupiters, something these authors attribute to age biases.

Finally, we emphasize that our work is subject to several important caveats and assumptions. First, unlike Winter et al. (2020), who analyze the Galactic field population (which is presumably comprised of many different star-forming regions), our analysis is on individual star-forming regions. In a set of exploratory simulations, we find that our results are similar if we create a composite data set of different snapshots from, e.g., different simulations at different ages, and perform a similar analysis on these stellar populations.

Second, our simulations do not include multiplanet systems, so we cannot quantify whether dynamical instabilities on the

outer planet would induce further instabilities on the inner planets (e.g., Flammini Dotti et al. 2019), potentially leading to more hot Jupiters.

Finally, we do not include the effects of dissolution of our star-forming regions due to feedback from massive stars (Goodwin & Bastian 2006; Shukirgaliyev et al. 2017). If this process is important, it is likely to “freeze in” the velocities of the planet host stars, and reduce the number of perturbed planets, thus not affecting our interpretation.

## 5. Conclusions

We have calculated the 6D Mahalanobis phase space “density” metric for sets of  $N$ -body simulations with different initial physical densities (low density  $\sim 10^2 M_\odot \text{pc}^{-3}$  and high density  $\sim 10^4 M_\odot \text{pc}^{-3}$ ) and different semimajor axes of Jupiter-mass planets (all planets at 5 or 30 au). We then compare the Mahalanobis phase space metric for stars that have lost their planets, stars that have retained their planets, and stars that have retained their planets but whose semimajor axis has changed by more than  $\pm 10\%$ . Our conclusions are as follows.

1. The 6D Mahalanobis phase space evolution for different subsets of stars (planet hosts, nonhosts, and former hosts) is similar for simulations with different initial physical densities.
2. We find that host stars are predominantly found in high Mahalanobis phase space regimes at the end of all the simulations ( $t = 10$  Myr), regardless of the initial physical density. This is due to the host stars retaining some of the kinematic substructure from their formation environment.
3. We find that host star whose planets’ semimajor axes have changed by more than 10% (perturbed host stars) are also found in high Mahalanobis phase space regimes regardless of the initial physical density.


Based on our results, we would not expect dynamical interactions in a planet host star’s birth environment to be responsible for the high Mahalanobis phase space densities, as hypothesized in Winter et al. (2020). Conversely, we find that unperturbed planet host stars are most likely to be found in high Mahalanobis phase space, because they retain some kinematic substructure from their birth star-forming region, which tends to be erased for stars that experience strong or repeated dynamical encounters.

## Acknowledgments

G.A.B.-S. acknowledges a University of Sheffield publication scholarship. R.J.P. acknowledges support from the Royal Society in the form of a Dorothy Hodgkin Fellowship. E.C.D.-P. acknowledges funding from UKRI in the form of an STFC PhD scholarship. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

*Software:* Matplotlib 3.3.4 (Hunter 2007), Numpy (Harris et al. 2020), SciPy (Virtanen et al. 2020).

## ORCID iDs

George A. Blaylock-Squibbs  <https://orcid.org/0000-0002-5710-7445>

Richard J. Parker  <https://orcid.org/0000-0002-1474-7848>

Emma C. Daffern-Powell  <https://orcid.org/0000-0002-3767-8187>

## References

- Adibekyan, V., Santos, N. C., Demangeon, O. D. S., et al. 2021, *A&A*, **649**, A111
- Alfaro, E. J., & González, M. 2016, *MNRAS*, **456**, 2900
- Alves, F. O., Cleves, L. I., Girart, J. M., et al. 2020, *ApJL*, **904**, L6
- André, P., Di Francesco, J., Ward-Thompson, D., et al. 2014, in *Protostars and Planets VI*, ed. H. Beuther et al. (Tucson, AZ: Univ. Arizona Press), 27
- Benkendorf, L., Flammini Dotti, F., Stock, K., Cai, M. X., & Spurzem, R. 2024, *MNRAS*, **528**, 2834
- Bressert, E., Bastian, N., Gutermuth, R., et al. 2010, *MNRAS: Letters*, **409**, L54
- Cartwright, A., & Whitworth, A. P. 2004, *MNRAS*, **348**, 589
- Chen, D.-C., Xie, J.-W., Zhou, J.-L., et al. 2023, *PNAS*, **120**, e2304179120
- Clarke, C. J., & Pringle, J. E. 1991, *MNRAS*, **249**, 588
- Concha-Ramírez, F., Wilhelm, M. J. C., Portegies Zwart, S., & Haworth, T. J. 2019, *MNRAS*, **490**, 5678
- Daffern-Powell, E. C., Parker, R. J., & Quanz, S. P. 2022, *MNRAS*, **514**, 920
- Dawson, R. I., & Johnson, J. A. 2018, *ARA&A*, **56**, 175
- Flammini Dotti, F., Kouwenhoven, M. B. N., Cai, M. X., & Spurzem, R. 2019, *MNRAS*, **489**, 2280
- Fregeau, J. M., Chatterjee, S., & Rasio, F. A. 2006, *ApJ*, **640**, 1086
- Gagné, J., Lafrenière, D., Doyon, R., Malo, L., & Artigau, É. 2015, *ApJ*, **798**, 73
- Gomez, M., Hartmann, L., Kenyon, S. J., & Hewett, R. 1993, *AJ*, **105**, 1927
- Goodwin, S. P., & Bastian, N. 2006, *MNRAS*, **373**, 752
- Haisch, K. E. J., Lada, E. A., & Lada, C. J. 2001, *ApJL*, **553**, L153
- Hamer, J. H., & Schlaufman, K. C. 2019, *AJ*, **158**, 190
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Natur*, **585**, 357
- Haworth, T. J., Clarke, C. J., Rahman, W., Winter, A. J., & Facchini, S. 2018a, *MNRAS*, **481**, 452
- Haworth, T. J., Facchini, S., Clarke, C. J., & Mohanty, S. 2018b, *MNRAS*, **475**, 5460
- Heggie, D. C. 1975, *MNRAS*, **173**, 729
- Hills, J. G. 1975, *AJ*, **80**, 809
- Hunter, J. D. 2007, *CSE*, **9**, 90
- Klessen, R. S., & Kroupa, P. 2001, *A&A*, **372**, 105
- Kobayashi, H., & Ida, S. 2001, *Icar*, **153**, 416
- Kruijssen, J. M. D. 2012, *MNRAS*, **426**, 3008
- Lada, C. J., & Lada, E. A. 2003, *ARA&A*, **41**, 57
- Larson, R. B. 1981, *MNRAS*, **194**, 809
- Larson, R. B. 1995, *MNRAS*, **272**, 213
- Li, D., Mustill, A. J., Davies, M. B., & Gong, Y.-X. 2024, *MNRAS*, **527**, 386
- Mahalanobis, P. C. 1936, *Sankhya A*, **80**, 1
- Malmberg, D., Davies, M. B., & Chambers, J. E. 2007, *MNRAS*, **377**, L1
- Maschberger, T. 2013, *MNRAS*, **429**, 1725
- Masset, F. S., & Papaloizou, J. C. B. 2003, *ApJ*, **588**, 494
- Mustill, A. J., Lambrechts, M., & Davies, M. B. 2022, *A&A*, **658**, A199
- Nicholson, R. B., Parker, R. J., Church, R. P., et al. 2019, *MNRAS*, **485**, 4893
- Olczak, C., Pfalzner, S., & Spurzem, R. 2006, *ApJ*, **642**, 1140
- Parker, R. J. 2020, *RSOS*, **7**, 201271
- Parker, R. J. 2023, *MNRAS*, **525**, 2907
- Parker, R. J., & Alves de Oliveira, C. 2017, *MNRAS*, **468**, 4340
- Parker, R. J., & Goodwin, S. P. 2009, *MNRAS*, **397**, 1041
- Parker, R. J., & Meyer, M. R. 2012, *MNRAS*, **427**, 637
- Parker, R. J., & Quanz, S. P. 2012, *MNRAS*, **419**, 2448
- Parker, R. J., & Reggiani, M. M. 2013, *MNRAS*, **432**, 2378
- Parker, R. J., & Schoettler, C. 2022, *MNRAS*, **510**, 1136
- Portegies Zwart, S., Makino, J., McMillan, S. L. W., & Hut, P. 1999, *A&A*, **348**, 117
- Portegies Zwart, S. F., McMillan, S. L. W., Hut, P., & Makino, J. 2001, *MNRAS*, **321**, 199
- Raymond, S. N., Armitage, P. J., & Gorelick, N. 2009, *ApJL*, **699**, L88
- Sánchez, N., & Alfaro, E. J. 2009, *ApJ*, **696**, 2086
- Scally, A., & Clarke, C. 2001, *MNRAS*, **325**, 449
- Schneider, A. C., Shkolnik, E. L., Allers, K. N., et al. 2019, *AJ*, **157**, 234
- Schneider, J., Dedieu, C., Le Sidaner, P., Savalle, R., & Zolotukhin, I. 2011, *A&A*, **532**, A79
- Sellek, A. D., Booth, R. A., & Clarke, C. J. 2020, *MNRAS*, **492**, 1279
- Shukirgaliyev, B., Parmentier, G., Berczik, P., & Just, A. 2017, *A&A*, **605**, A119
- Steinhausen, M., Olczak, C., & Pfalzner, S. 2012, *A&A*, **538**, A10
- Vincke, K., & Pfalzner, S. 2016, *ApJ*, **828**, 48
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *NatMe*, **17**, 261
- Winter, A. J., Clarke, C. J., Rosotti, G., et al. 2018, *MNRAS*, **478**, 2700
- Winter, A. J., Kruijssen, J. M. D., Longmore, S. N., & Chevance, M. 2020, *Natur*, **586**, 528
- Wu, Y., & Murray, N. 2003, *ApJ*, **589**, 605
- Zuckerman, B., & Song, I. 2004, *ARA&A*, **42**, 685