nature astronomy

Article

https://doi.org/10.1038/s41550-026-02990-8

Two families of multi-planet systems with distinct architectures and stellar hosts

Received: 1 July 2022

Accepted: 1 September 2026

Published online: 13 October 2026

Check for updates

Diego Turrini[1,13] ✉, Camilla Danielski[2,3,13] ✉, Danae Polychroni[1], Robert Przyłuski[4], Hans Rickman[4,5], Diego Bossini[6,7], J. M. Diederik Kruijssen[8], Angelo Zinzi[9], Simon Goodwin[10], Steven N. Longmore[8,11], Mélanie Chevance[8,12] & Lise Ramambason[12]

The rapidly growing census of exoplanets now includes over a thousand multi-planet systems, revealing architectures that span more than four orders of magnitude in mass and five in orbital scale. Explaining the origin of this striking diversity remains a central challenge for planet formation theory, compounded by observational biases that limit access to Solar System-like regimes. Here we introduce architecture-agnostic metrics to uniformly analyse four independently defined samples of multi-planet systems, uncovering unifying patterns across their diversity. We identify a clear bimodality in system masses, dividing known architectures into two distinct families. The lower-mass family, predominantly hosted by subsolar-to-solar stars, exhibits tightly packed configurations with mass distributions consistent with the 'peas-in-a-pod' pattern and modal dynamical excitation comparable with that of the Solar System. By contrast, the high-mass family forms in metal-rich environments around solar-to-supersolar stars, displays gravitationally compact architectures and is an order of magnitude dynamically hotter than both the Solar System and the lower-mass counterparts. Their high level of dynamical excitation points to violent evolutionary pathways driven by global instabilities and planet–planet scattering. Together, these results show that the architectures of planetary systems are born from the combined influence of formation processes and host star properties. Signatures of these roots are preserved across the post-formation dynamical evolution that reshaped most known multi-planet systems.

The known 1,000+ multi-planet systems display a notable diversity of architectures, both as a population and with respect to the Solar System. Despite substantial progress in classifying this architectural diversity[1–8], especially for transiting planets thanks to the NASA mission Kepler[9–14], its roots and full extent remain elusive. Owing to observational biases and instrumental limitations, the planets and planetary systems we observe are not necessarily the most frequent products of planet formation, just the ones most accessible to our current capabilities[15]. As well as this incomplete observational picture, planetary population synthesis models face several challenges in simultaneously reproducing the Solar System planets and the observed exoplanetary demographics[16].

This situation is analogous to the parable of the blind men attempting to describe an elephant by touching different body parts. Each observer produces a different description due to their limited perspectives. Different observational approaches probing specific aspects of planetary systems lead to partial interpretations, a more complete picture emerging only when these perspectives are combined. In this population study, we, thus, investigate the global properties of known multi-planet systems, without pre-selecting by architecture or stellar

A full list of affiliations appears at the end of the paper. ✉e-mail: diego.turrini@inaf.it; camilla.danielski@uv.es

https://doi.org/10.1038/s41550-026-02990-8

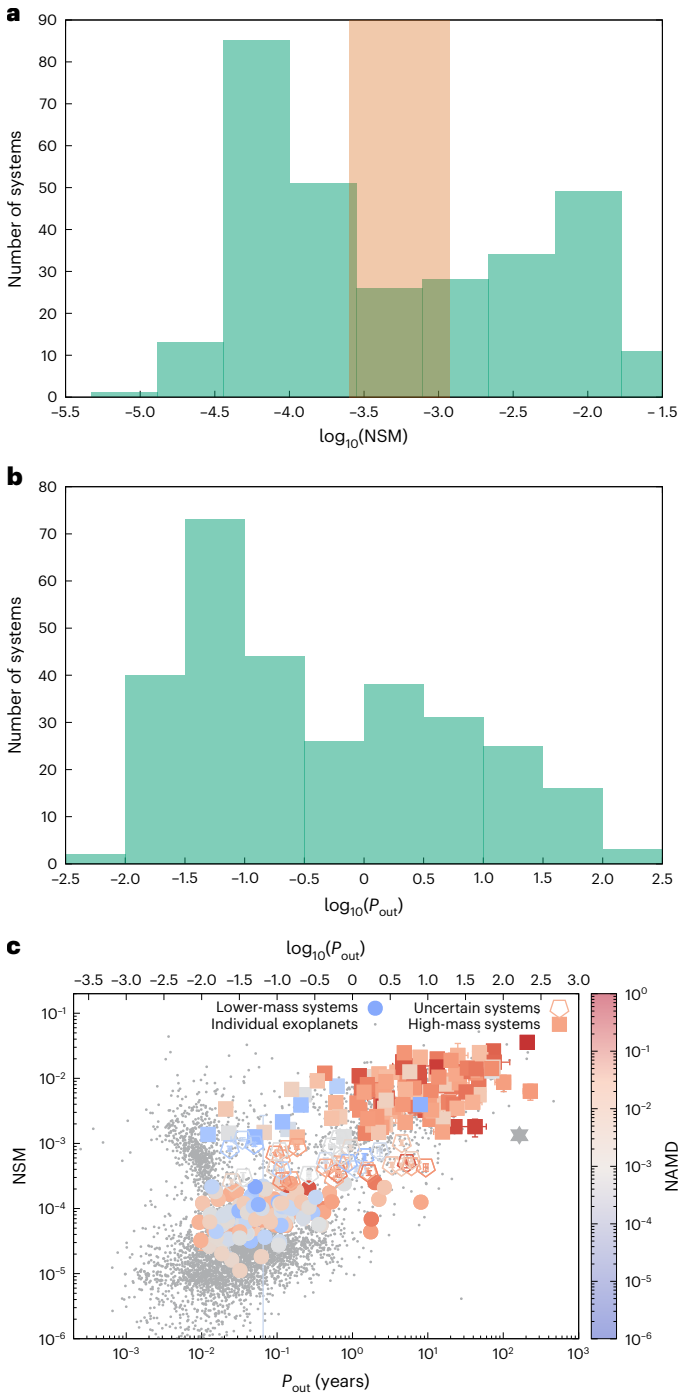


Fig. 1 | Physical properties of the multi-planet systems in sample L. a, Bimodal distribution of the NSMs. The region highlighted in red marks the mass range where systems can belong to the lower- or high-mass families with comparable probabilities (Methods). **b**, Bimodal distribution of the orbital periods of the outer planets P_{out}. **c**, Comparison between the known exoplanet population and the high- and lower-mass families identified in **a**. Data are presented as modal values with the error bars identifying the 16th and 84th percentiles (most error bars are smaller than the markers and not visible). Systems with uncertain attribution are identified by empty symbols. The Solar System is shown for comparison as the grey star.

type, to uncover the underlying drivers of their diversity. To address the observational biases affecting the exoplanetary demography, we perform four parallel investigations on samples selected through different criteria that we consistently analyse by means of architecture-agnostic metrics and extensive statistical testing.

Multi-criteria selection and characterization of the planetary samples

The first sample, labelled L for large selection, is composed of all multi-planet systems for which we know both the semimajor axis a_p and mass m_p of each planet ('The sample of multi-planet systems and their host stars' in Methods). Following ref. 2, we consider as valid masses both the absolute mass M_p and the minimum mass $M_p \sin i_p$ where i_p is the orbital inclination of the planet. To investigate the dynamical excitation of these systems, we retrieved the orbital eccentricity e_p of the planets whenever available. For those systems where the eccentricities are not known for all planets, we filled the gaps by using the correlation between the number of planets N and their average eccentricity derived from Kepler observations[12] ('Treatment of the missing eccentricities and inclinations in samples L, E and S' in Methods). Sample L has $n_L = 298$ multi-planet systems, that is 28% of their known population, and is our reference sample.

The second sample, labelled H for homogeneous stars, was created by cross-matching sample L with the SWEET-Cat catalogue[17,18], which provides homogeneous stellar atmospheric parameters (effective temperature, surface gravity and [Fe/H]). We then estimated the homogeneous stellar masses by fitting isochrones ('The sample of multi-planet systems and their host stars' in Methods and Supplementary Table 1). Sample H contains $n_H = 189$ multi-planet systems and was used to validate the observed trends for the properties of the host star.

The third sample, labelled E for excitation, is the subset of sample L for which the orbital eccentricity e_p is known for all planets ('The sample of multi-planet systems and their host stars' in Methods). Sample E contains $n_E = 89$ systems and was used to validate any trend in the dynamical excitation[4,19] emerging from samples L and H.

Finally, the fourth sample, labelled S for strict selection, includes only those systems in sample E for which the absolute planetary mass M_p and its relative uncertainty are known for all planets ('The sample of multi-planet systems and their host stars' in Methods). Sample S is composed of $n_S = 68$ systems and was used to validate any trend of parameters involving the planetary mass against possible biases introduced by the inclusion of the minimum mass $M_p \sin i_p$.

The complete architectural and dynamical characterization of the samples requires knowing the relative inclinations of the planetary orbits[2,4,20], yet this information is available for only a handful of systems. For the other systems, we assumed the average equipartition of dynamical excitation among the secular degrees of freedom[2,12] to statistically constrain the relative inclinations of the planets. We adopted different approaches for the four samples ('Treatment of the missing eccentricities and inclinations in samples L, E and S' in Methods). For samples L and H, for which the orbital eccentricities are not always known, we adopted the multiplicity–inclination correlation from ref. 12. For samples E and S, where eccentricities are known by definition, we used the eccentricity–inclination correlation from refs. 2,4. Further details of the samples, their selection criteria and the queries used to retrieve them are provided in 'The sample of multi-planet systems and their host stars' and 'Treatment of the missing eccentricities and inclinations in samples L, E and S' in Methods.

We characterized the global properties of the samples in a uniform way with nine normalized parameters. First, the basic physical properties of each system are quantified by (1) the normalized system mass (NSM), that is the ratio between the total planetary mass and the stellar mass[11,15], and (2) the orbital period of the outermost planet P_{out}, quantifying how wide the system is the global architecture of each system is characterized by (3) planetary multiplicity N, that is, the number of planets; (4) the fraction of the total planetary mass contained in the most massive planet S_m (ref. 20); (5) the orbital spacing S_s (refs. 7,20), quantifying how spatially and gravitationally compact the system is; and (6) the normalized angular momentum deficit (NAMD)[4,20], quantifying the dynamical excitation and intuitively interpreted as its 'dynamical temperature'[19].

https://doi.org/10.1038/s41550-026-02990-8

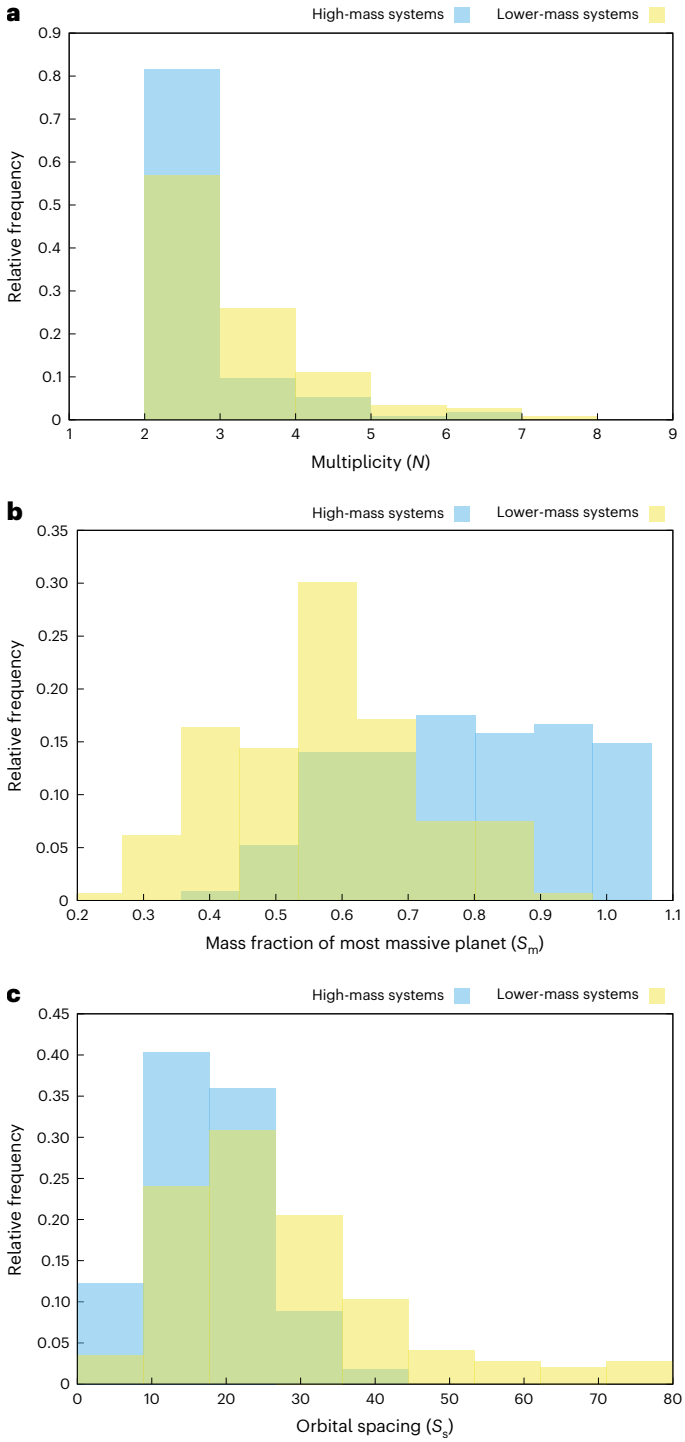


Fig. 2 | Global architectures of the multi-planet systems in sample L. **a**, Distributions of the multiplicity N in the lower- and high-mass families of multi-planet systems. **b**, Distributions of the mass fraction of the most massive planets S_m. **c**, Distributions of the orbital spacing metric S_s.

The properties of the host star are characterized by (7) metallicity ([Fe/H]), (8) stellar mass M_* in units of solar masses and (9) total planet-building material (PBM), which quantifies the amount of solids (in dex as $\log_{10}(M_{dust}/M_\oplus)$) in its, now-lost, class II protoplanetary disk, by combining the information provided by stellar mass and metallicity with the star mass M_* versus disk dust mass M_{dust} correlation[21,22]. Detailed descriptions of the normalized metrics and the computation of their uncertainties is provided in 'Normalized parameters and uncertainty propagation' in Methods. The stellar properties of sample H are provided in Supplementary Table 1.

Two families of multi-planet systems in the mass–period space

First, we characterized the reference sample (L) and validation samples (H, E and S) in terms of their normalized mass (NSM) and their radial extent (P_{out}) by analogy with the classical exoplanet mass–period distribution. Although the four samples of multi-planet systems show continuous distributions in the mass–period parameter space, both NSM and P_{out} reveal bimodal distributions, as illustrated in Fig. 1a,b for sample L and in Extended Data Figs. 1–3 for samples H, E and S, respectively ('Binning methodology' in Methods). We verified these bimodalities for all four samples using a one-sample Kolmogorov–Smirnov (KS) test ('Statistical characterization of the two families and their properties' in Methods). Nonetheless, given that the NSM and P_{out} bimodalities are probably connected through observational biases, we conservatively focused on the NSM bimodality to identify the two families associated with the modal peaks and the uncertain systems.

Using the Gaussian mixture method ('Statistical characterization of the two families and their properties' in Methods), we defined the NSM parameter space boundaries associated with each modal peak. For sample L, the family of lower-mass systems has $\log_{10}(\mathrm{NSM}) < -3.602$ (corresponding to $\lesssim 80\ M_\oplus$ around a solar mass star), and the family of high-mass systems has $\log_{10}(\mathrm{NSM}) > -2.927$ (corresponding to $\gtrsim 400\ M_\oplus$ around a solar mass star). The boundaries of the two families for samples H, E and S are reported in Supplementary Table 2. Uncertain membership regions are highlighted in red in Fig. 1a and by empty markers in Fig. 1c for sample L and in Extended Data Figs. 1–3 for samples H, E and S. After removing these uncertain memberships from all four samples, we verified the bimodality of our metrics (P_{out}, S_s, S_m, NAMD, PBM, N, M_* and [Fe/H]) using two-sample KS tests[23] corrected for multiple comparisons to account for false positives within the data ('Statistical characterization of the two families and their properties' in Methods). We confirmed that the results are not artefacts of the uncertainties of the parameters (especially samples E and S, which are smaller) using Monte Carlo-resampled Anderson–Darling (AD) tests with Benjamini–Hochberg (B-H) false discovery rate control ('Statistical characterization of the two families and their properties' in Methods). All repeated tests consistently confirmed the bimodality for all samples.

Overlaying the samples onto the known population of exoplanets (source: NASA Exoplanet Archive[24], https://exoplanetarchive.ipac.caltech.edu), the lower- and high-mass systems (Fig. 1c and Extended Data Figs. 1–3) map into the main groups of exoplanets discovered through the transit and radial velocity methods (source: ExoArchive Scatter, https://exoplanetarchive.ipac.caltech.edu/ScatterPlot/), justifying our conservative choice of focusing on the NSM bimodality. The transit method is sensitive to planets in short-period orbits ($P_{out} \lesssim 1$ year), whereas planets in longer-period orbits are identified through radial velocities, direct imaging and astrometry[15,16]. However, at large orbital distances, these methods presently detect only high-mass planets[15,16], meaning that the only extended multi-planet systems we can discover are those with high total masses (Fig. 1c and Extended Data Figs. 1–3).

The following analysis shows that, notwithstanding these observational biases, the NSM bimodality identifies two physically distinct families of multi-planet system architectures with host stars characterized by different properties. The comparison with the known exoplanetary demography also highlights that the parameter space region where hot Jupiters reside is visibly underpopulated (Fig. 1c). We investigate this in Supplementary Information, section 1.

Different architectures, dynamical excitations and host star properties

The multiplicity N of both lower- and high-mass systems is dominated by two-planet architectures but the families have different abundances of systems with $N > 2$ (Fig. 2a): up to 45% for lower-mass systems and 25% for high-mass systems. In most lower-mass systems, the most massive

https://doi.org/10.1038/s41550-026-02990-8

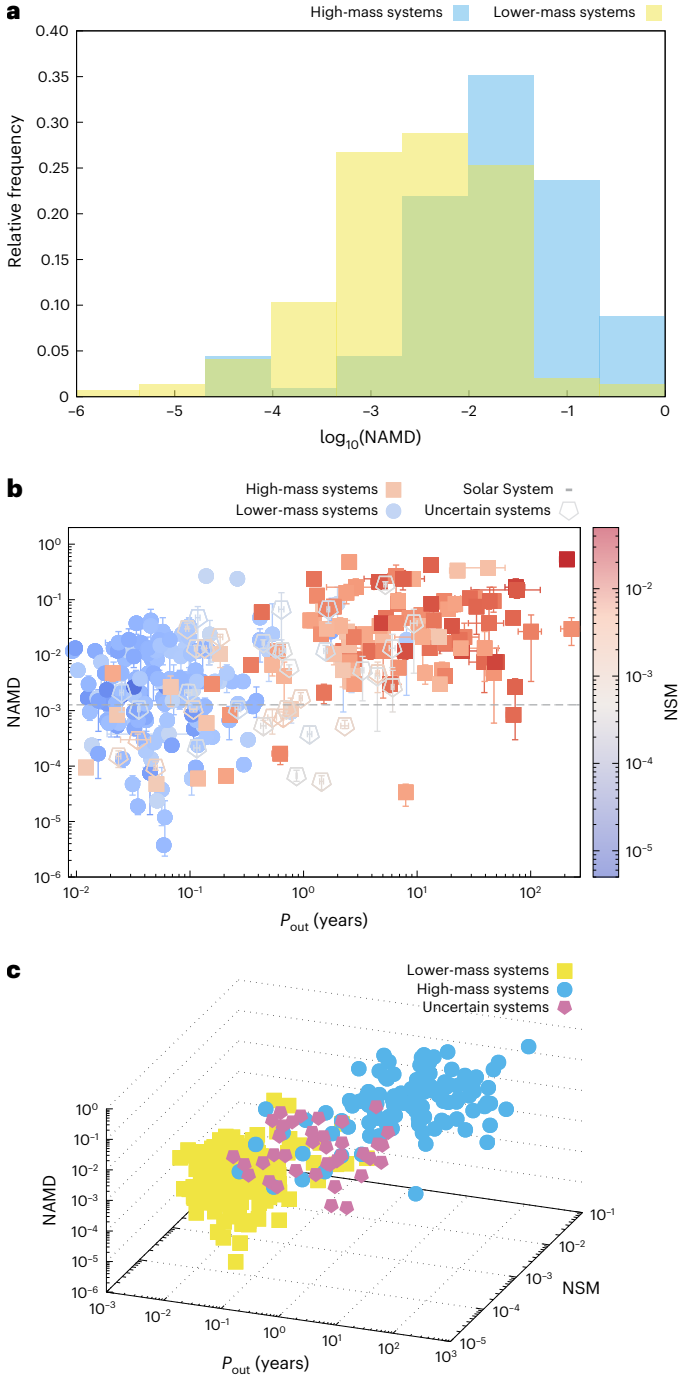


Fig. 3 | Dynamical excitation of the multi-planet systems in sample L. **a**, Distributions of NAMD for lower- and high-mass systems. **b**, NAMD versus P_{out} scatter plot. Data are presented as modal values, and the error bars are the 16th and 84th percentiles; the Solar System $NAMD_{ss}$ is marked with the dashed grey line. The colour scale shows NSM. **c**, Scatter plot in the NAMD–P_{out}–NSM parameter space (see Supplementary Video 1 for an animated version).

planet contains between one- and two-thirds of the total planetary mass (Fig. 2b and Extended Data Figs. 1–3). As more than half of this family is composed of two-planet systems, the S_m distribution of lower-mass systems is compatible with the peas-in-a-pod scenario, which refers to the observational finding that neighbouring transiting planets observed by the NASA mission Kepler within the same multi-planet system generally have similar radii[8,10,11]. High-mass systems show, instead, a flat S_m distribution between 0.5 and 1 (Fig. 2b and Extended Data Figs. 1–3), meaning that systems composed of similar-mass planets are as probable as systems where almost all the mass is concentrated in one single object. Given that the family of high-mass systems has a larger fraction of $N = 2$ systems in all four samples (Fig. 2a and Extended Data Figs. 1–3), this difference does not arise from higher planetary multiplicities. By using statistical samples an order of magnitude larger, these results extend the findings of ref. 25, which studied how the architectures of multi-planet systems in the candidate sample of ESA's Ariel space mission depend on the homogeneously characterized masses and metallicities of their host stars. The Solar System ($S_{m,ss} = 0.7$) falls in the S_m range common to both families.

The orbital spacing metric S_s reveals that high-mass systems, despite their wider orbital extension traced by P_{out}, are globally packed the same as or even more packed than lower-mass systems (Fig. 2c and Extended Data Figs. 1–3). By contrast, the lower-mass systems are largely dominated by spatially packed configurations spanning only a fraction of an astronomical unit. As S_s accounts for both the radial and the gravitational spacing between planets, this suggests the high-mass systems are more gravitationally packed. The bulk of high-mass systems have S_s comparable with or larger than that of the outer Solar System[20] ($S_{s,oss} = 11.6$; Fig. 2c and Extended Data Fig. 1; this is not apparent in Extended Data Figs. 2 and 3 due to fewer and wider bins), whereas the bulk of lower-mass systems have S_s markedly smaller than that in the inner Solar System[20] ($S_{s,iss} = 37.7$; Fig. 2c and Extended Data Figs. 1–3). Both families have S_s values larger than that of the whole Solar System[20] ($S_{s,ss} = 8.1$).

The NAMD metric shows that both families contain large fractions of systems more dynamically excited than the Solar System[4] ($NAMD_{ss} = 1.27 \times 10^{-3}$; Fig. 3a,b and Extended Data Figs. 1–3). In sample L, where most eccentricity values are estimated through the multiplicity–eccentricity correlation found among Kepler planets[12], the NAMD modal peak of lower-mass systems is overall comparable to that of the Solar System. In sample E, for which measured eccentricities are available, the peak is shifted to higher values consistently with the proposed role of dynamical instabilities in sculpting Kepler multi-planet systems[12,13]. High-mass systems are always about an order of magnitude more excited than both lower-mass ones and the Solar System (Fig. 3a,b and Extended Data Fig. 2), suggesting violent dynamical histories shaped by instabilities and planet–planet scattering[4,19,20]. This behaviour is consistent with their more massive and gravitationally packed architectures, which lead to stronger gravitational perturbations during close planetary encounters. The higher NAMD of high-mass systems and their larger abundance of $N = 2$ systems are in agreement with the dynamical excitation–multiplicity anticorrelation found in previous studies[4,12,26,27].

When we consider the high- and lower-mass families together, the general trend in all four samples is that the larger the NSM, the higher the NAMD (Fig. 3b and Extended Data Figs. 1–3), that is, the system is more dynamically excited. This global NSM–NAMD correlation is consistent with the picture from the orbital spacing S_s. Lower-mass systems have a similar or larger spacing than their high-mass counterparts but are subject to stronger stellar gravitational bounds and weaker planetary perturbations, resulting in overall less violent instabilities. In the P_{out}–NSM–NAMD space, the correlation between these parameters extends across the three dimensions (Fig. 3c, Extended Data Figs. 1–3 and Supplementary Video 1), with the family of uncertain systems falling in the intermediate region between lower- and high-mass systems (Supplementary Video 1). The real trend is probably more complex due to the presence of undiscovered wide, intermediate-mass systems.

The differences between high- and lower-mass systems extend to the properties of their host stars (Fig. 4) and provide insight into the planet-forming environments that gave birth to the two families. Lower-mass systems mainly orbit stars with a solar-to-subsolar mass whose modal peak of metallicity falls between −0.1 dex and 0.1 dex, whereas the metallicity distribution of the solar-to-supersolar mass stellar hosts of high-mass systems peaks around 0.2 dex, that is about ×1.5 larger (Fig. 4a,b and Extended Data Figs. 1–3). Although the

https://doi.org/10.1038/s41550-026-02990-8

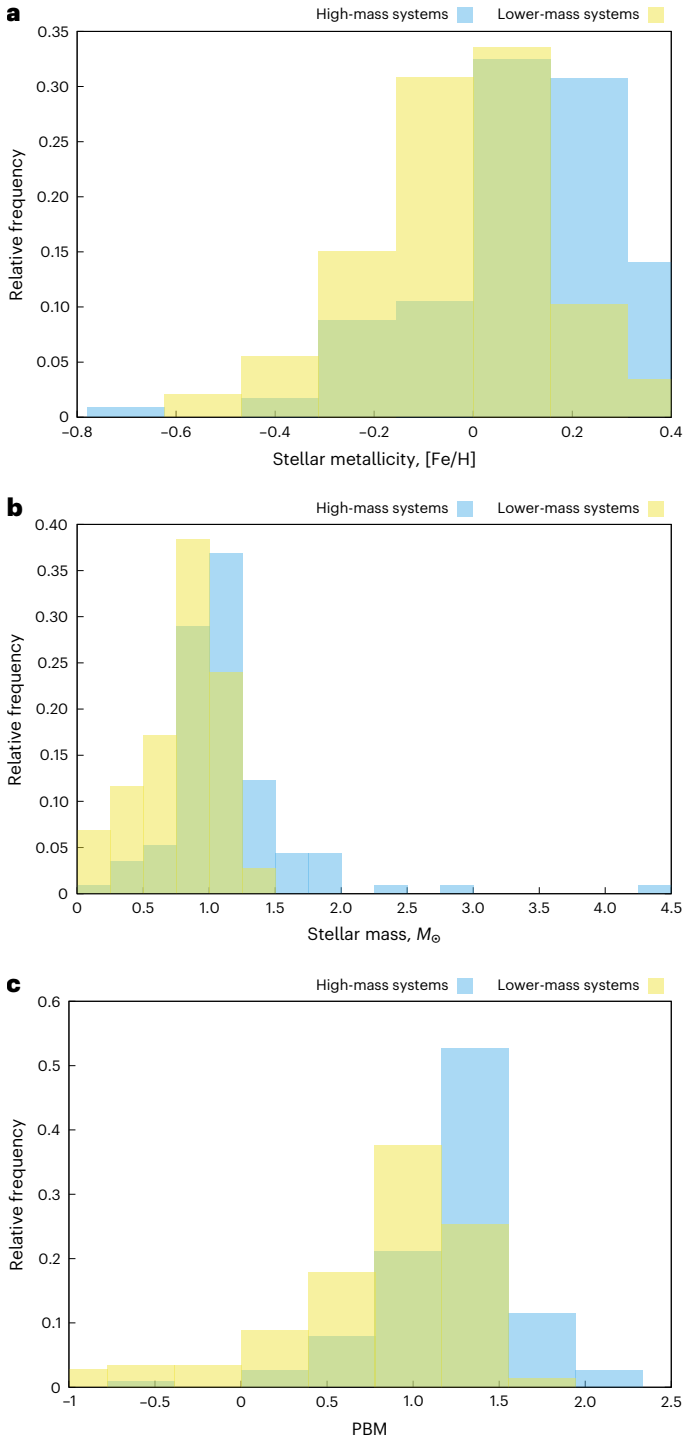


Fig. 4 | Stellar properties of the host stars of sample L. a, Distributions of the stellar metallicity [Fe/H]. **b**, Distributions of the stellar masses $M_\star$. **c**, Distributions of PBM.

differences in stellar mass and metallicity may seem moderate individually, they interact constructively due to the proportionality between stellar and disk masses revealed by star–disk correlations[21,22]. As a result, the planet-forming environments of high-mass systems are characterized by PBM values about 0.5 dex higher than those of low-mass systems (Fig. 4c and Extended Data Figs. 1–3). This shift in PBM means that their natal disks contained roughly ×3 more heavy elements, hence dust and ices, than the disks of their lower-mass counterparts.

These differences in the properties of the host stars are confirmed by the homogeneously characterized stars of sample H (Extended Data Figs. 1–3). All the architectural and stellar differences discussed above were confirmed by statistical two-sample KS tests, incorporating the Holm–Bonferroni and the B-H procedures, as well as by Monte Carlo-resampled AD tests with B-H false discovery rate control (Extended Data Figs. 4–7 and 'Statistical characterization of the two families and their properties' in Methods).

Exploring the roots of the excited planetary architectures

We complemented the architectural study with an exploratory campaign of instability simulations to assess which dynamical scenarios can result in excitations like those of the two families. For both lower- and high-mass architectures, we simulated the dynamical evolution of systems starting in (1) unstable non-resonant configurations, (2) incomplete resonant chains, (3) stable non-resonant configurations and (4) stable resonant chains. The two initially stable configurations (3 and 4) are perturbed by stellar passages that may (or may not) destabilize their systems ('Dynamical simulations of the possible origins of the excited architectures of the sample' in Methods). We adopted equal-mass planets as a reference set-up for all initial architectures as this peas-in-a-pod-like initial state is compatible with both lower- and high-mass systems. For scenario 4, we considered an alternative architecture where the mass partition is that of the outer Solar System (4-ss), as this architecture is also common to both families.

The dynamical excitations of both high- and lower-mass systems in sample L seem to be intermediate between those resulting from the two initially unstable scenarios (1 + 2), with naturally unstable architectures providing the closest match to those of sample E (Fig. 5a,c,d). Initially stable scenarios, independently of whether they are resonant or not (3 + 4 + 4-ss), do not reproduce the fractions of high- and lower-mass systems more excited than the Solar System[4] (Fig. 5a,c,d). Most instability simulations producing excited architectures like those in samples L and E resulted in the loss of at least one planet (Fig. 5a), consistently with the dynamical excitation–multiplicity anticorrelation[4,12,19,26] and the results of stability investigations[13,27,28]. Among these studies, ref. 13 suggests a central role of planet loss by giant impacts in shaping the architectures of Kepler multi-planet systems. This scenario is consistent with most lower-mass systems having S_m values between the curves where all planets share the same mass, like the peas-in-a-pod scenario[10], and where the most massive planet formed from the collisional merge of two planets from an initially $N + 1$ higher multiplicity peas-in-a-pod system (Fig. 5b).

Only a small fraction of high-mass systems falls within the same two curves, with the family showing a more continuous distribution that includes many systems with higher S_m than the Solar System (Fig. 5b). This distribution is, plausibly, a joint outcome of planet formation and dynamical evolution. Although giant planets forming at different times within the same protoplanetary disk naturally have a wide mass dispersion[29–32], the loss of planets by impacts and ejections is a widespread outcome of the strong dynamical instabilities that shaped these massive systems[4,19] (Fig. 5). As such events increase the mass concentration S_m in the largest planet by removal or accretion, their effects magnify any initial difference in the S_m distribution.

Insights into planet formation and the connection to stellar properties

The distinct architectural properties of the high- and lower-mass families of multi-planet systems can be explained by the interplay between planet formation and the different characteristics of their stellar hosts. Lower-mass systems emerge from Solar System-like planet-forming environments (as mapped by the [Fe/H] and $M_\star$ of their host stars), and both their spatially compact architectures (S_s and P_{out}) and mass partitioning (S_m) can be explained by formation tracks in which the growth rates of planets are lower than their migration rates[33]. These formation tracks lead to large displacements during planet formation

https://doi.org/10.1038/s41550-026-02990-8

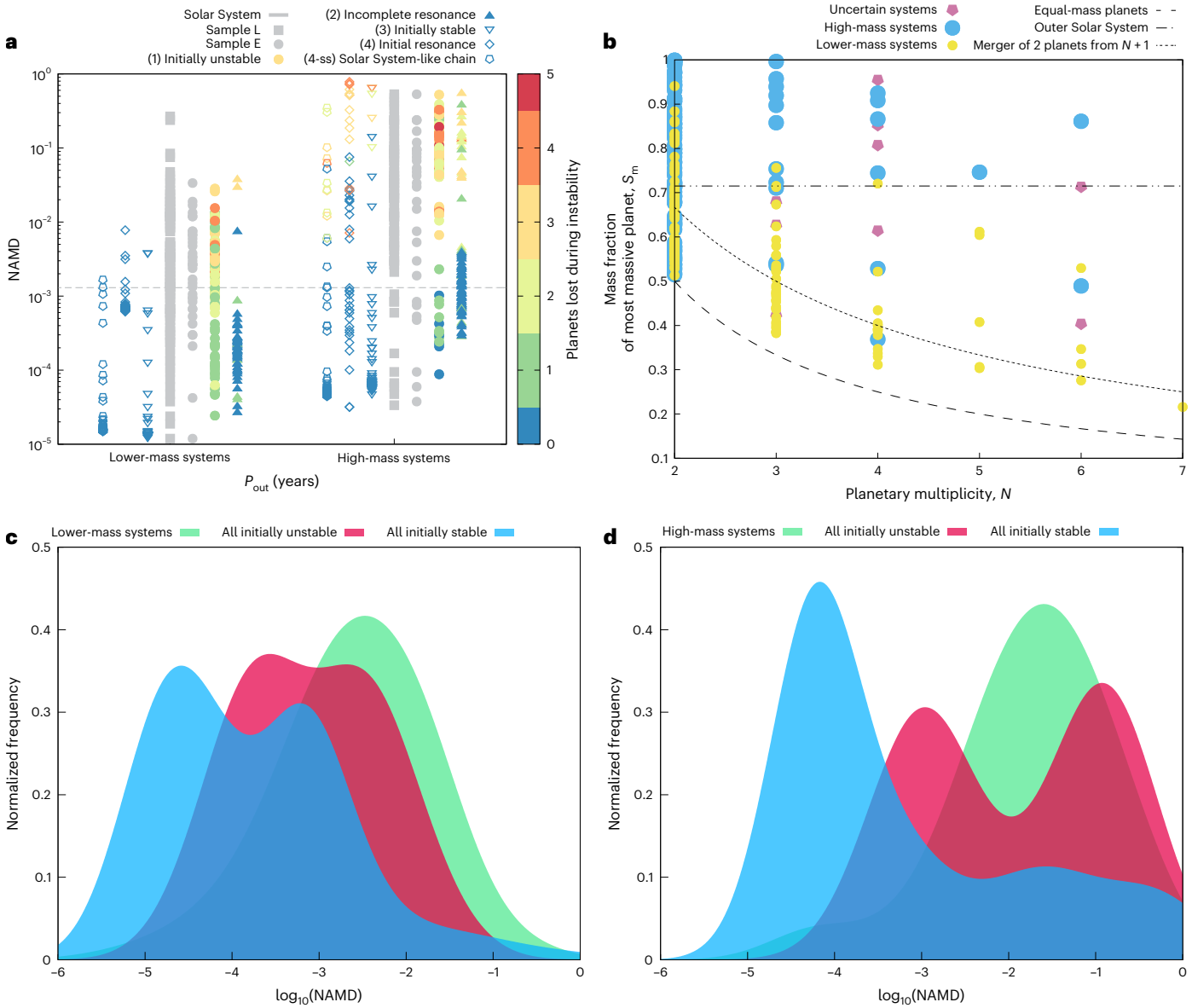


Fig. 5 | Constraints on the dynamical histories of samples L and E. a, Comparison of the dynamical excitation of sample L (grey squares) and sample E (grey circles) with synthetic analogues undergoing dynamical evolution starting from stable, unstable and resonant initial architectures. The colour scale shows the number of initial planets lost during the dynamical evolution of the synthetic systems. **b**, Comparison of the mass fraction in the most massive planets as function of the planetary multiplicity N of lower-mass, high-mass and uncertain systems. The horizontal line represents Jupiter in the Solar System (S_m = 70% of the mass of the planetary system). The two curves identify the mass-wise peas-in-a-pod scenario (the planets initially have the same mass) and the mass fraction resulting from the collisional merge of two equal-mass planets in a system with initial multiplicity N + 1. **c**, Overlapping kernel density plot showing the NAMD distributions of lower-mass systems in sample L and the aggregated NAMD distributions of all initially unstable (1 + 2) and all initially stable (3 + 4 + 4-ss) synthetic lower-mass systems. **d**, Same overlapping kernel density plot for high-mass systems.

and, when shaped by pebble accretion, produce comparatively similar planetary masses due to the limited variation of the pebble isolation mass across the inner protoplanetary disk[33].

High-mass system are, instead, born from planet-forming environments that are, on average, three times richer in gas and solids (as mapped by their PBM), so that planetary growth becomes more efficient than orbital migration[33] and promotes the formation of massive planets at larger orbital distances (as recorded by their NSM and P_{out}). The diversity of planets within individual high-mass systems (S_m) is consistent with their gas- and dust-rich disk environments, which can sustain planet formation over longer timescales and allows planets to form at different times under markedly different solid and gas accretion rates[25,29–32]. However, high-mass systems have highly dynamically excited architectures, suggesting that their S_m distribution is not exclusively primordial but was sculpted, at least partially, by dynamical instabilities.

The two families of multi-planet systems share common properties that set them apart from the Solar System. Notwithstanding their different spatial extensions (P_{out}), both families have architectures that are more compact (S_s) than the inner Solar System and less compact than the outer Solar System. Furthermore, both families show large fractions of systems more dynamically excited than our Solar System. In high-mass systems, this fraction becomes dominant and, together with their larger population of N = 2 systems, points to more violent dynamical and collisional histories. This is consistent with the multiplicity–excitation anticorrelation reported in previous studies[4,12,19,26,27]. These violent dynamical histories can produce hot Jupiters through high-eccentricity migration, a scenario compatible with most hot

Jupiters inhabiting high-NAMD systems and not being their most massive planet (Supplementary Fig. 1).

The distinct architectural properties of the high- and lower-mass families of multi-planet systems and their connections to different stellar properties show the importance of accounting for the star–planet genetic link instead of adopting the Sun and the Solar System as universal templates. Both high- and lower-mass systems are better described as the outcome of 'initially unstable architectures' or 'incomplete resonance chains' (Fig. 5), supporting previous population studies showing that dynamical instabilities played a large role in shaping multi-planet systems[2,5,13]. When combined with the observed paucity of multi-planet systems in resonant chains[34,35], this suggests that planet formation is only moderately efficient at producing stable resonant architectures.

Methods

The sample of multi-planet systems and their host stars

We collected planetary data from the NASA Exoplanet Archive[24] (https://exoplanetarchive.ipac.caltech.edu/) on 14 October 2025 using `ExoplAn3T` (Exoplanet Analysis and 3D Visualization Tool[36], https://tools.ssdc.asi.it/exoplanet/) developed by the Space Science Data Center of the Italian Space Agency. `ExoplAn3T` allows searches based on the characteristics of planetary systems instead of individual exoplanets and has been used in previous studies focused on global investigations of exoplanetary systems[3,4]. We selected the sample of multi-planet systems through a two-step procedure. The first step involves a planetary query to identify exoplanets matching a specified set of characteristics, while the second step performs a system-wide query to retrieve data about all their planetary companions and their host stars. The data collected for sample L through the planetary and system-wide queries were filtered to preserve only multi-planet systems for which measurements and uncertainties on planetary mass (here both M_p and $M_p \sin i_p$) and semimajor axis were available for all planets. Sample H was derived by cross-matching sample L with the catalogue of stars for which we have homogeneous stellar parameters, as discussed in the following. For samples E and S, we included further filters on the orbital eccentricities and the planetary mass (here M_p exclusively), respectively, and their uncertainties. All queries filtered out systems with controversial planets, binary stellar systems and circumbinary planets to avoid biases due to the effects of the stellar companion.

In the format of the NASA Exoplanet Archive[24], the combined planetary and system queries for sample L are as follows:

```
default_flag=1
sy_snum=1
sy_pnum>1
cb_flag=0
pl_controv_flag=0
pl_orbsmax not null
pl_orbsmaxlim=0
pl_orbsmaxerr1 not null
pl_orbsmaxerr2 not null
pl_bmassj not null
pl_bmassjlim=0
pl_bmassjerr1 not null
pl_bmassjerr2 not null
```

with the query for sample E adding:

```
pl_orbeccen not null
pl_orbeccenlim=0
```

and that for sample S including also:

```
pl_massj not null
pl_massjlim=0
pl_massjerr1 not null
pl_massjerr2 not null
```

During the computation of the global parameters of samples E and S, we also filtered out those systems containing planets for which the orbit was reported as non-circular but for which no uncertainty estimate was provided for the orbital eccentricity (only planets on circular orbits can have no uncertainty reported).

For each exoplanet, we retrieved the following parameters:

- Host star name
- Planet letter
- Planet name
- System multiplicity
- Planetary orbital eccentricity and associated uncertainties
- Planetary mass (in Jupiter masses) and associated uncertainties
- Orbital semimajor axis of the planet and associated uncertainties
- Planetary orbital inclination and associated uncertainties
- HD name of the star
- HIP name of the star
- Right ascension of the star
- Declination of the star
- Gaia IDs of the star

The creation of sample H from sample L required that we derive homogeneous stellar data, which allowed us to more robustly test correlations and trends between planets and their host stars (for example, refs. 25,29,31,32,37–40). We cross-matched the multi-planet systems in sample L with the SWEET-Cat catalogue[17,18] (http://sweetcat.iastro.pt), and we selected those stars that have been homogeneously characterized. This means that the effective stellar temperature $T_{\rm eff}$, metallicity [Fe/H] and surface gravity $\log g$ have been determined with the same technique from high-resolution, high signal-to-noise spectra. The planet and host stars of sample H have temperatures 3,200 K $\leq T_{\rm eff} \leq$ 6,600 K, metallicity −0.64 dex $\leq$ [Fe/H] $\leq$ 0.48 dex and surface gravity 2.6 dex $\leq \log g \leq$ 5.1 dex.

We computed the stellar masses using the optimization tool PARAM (ref. 41–43), which we used to fit isochrones by comparing a grid of theoretical stellar isochrones with a set of observables, such as metallicity, effective temperature and photometric luminosity. The input luminosity L was computed separately by converting the bolometric magnitude $M_{\rm bol}$ through the standard formula:

$$L = 10^{M_{\rm bol}} = 10^{-0.4(m+{\rm BC}_m-(-5+5\log(d))-M_{\rm bol,\odot}-A_{\rm m})}, \quad (1)$$

where m is the generic magnitude, ${\rm BC}_m$ is its bolometric correction, d is the distance (in parsecs), $M_{\rm bol,\odot}$ is the bolometric magnitude of the Sun (4.74) and $A_{\rm m}$ is the extinction for the m-band. $M_{\rm bol}$ was derived from a Markov chain Monte Carlo fitting procedure, which took as input a combination of several magnitudes and their respective uncertainties. The magnitudes were taken from 2MASS[44] (J-, H- and K_s-bands), Gaia[45] ($G_{\rm BP}$ and $G_{\rm RP}$) and when available, Johnson B and V magnitudes. The Johnson magnitudes were taken either from the synthetic photometry generated from Gaia BP/RP mean spectra[46] or from the Tycho-2 catalogue (derived from the Tycho-2 $B_{\rm T}$ and $V_{\rm T}$ magnitudes)[47], and we prioritized the first source. The distance was derived by a simple parallax inversion, with the parallax given by Gaia eDR3. The bolometric corrections were calculated using YBC[48], which interpolates a series of precomputed tables for the bolometric correction with respect to $T_{\rm eff}$, [Fe/H], $\log g$ and $A_{\rm v}$. For the theoretical models (the isochrones), we relied on PARSEC v2.1 (Padova and Trieste Stellar Evolution Code[49], http://stev.oapd.inaf.it/cgi-bin/cmd).

Treatment of the missing eccentricities and inclinations in samples L, E and S

The query to create sample L (and, consequently, H) returns systems where the orbital eccentricity and inclination values are not known for all planets. To fill these gaps and compute the NAMD of these systems,

we took advantage of the trends derived from the population of transiting planets characterized by the Kepler mission to fill the gaps with physically justified values. Specifically, we adopted the trends based on the criterion for the stability of the angular momentum deficit[12]:

$$\mu_{\mathrm{i,n}} = 1.10^{+0.15}_{-0.11} \times \left(\frac{N}{5}\right)^{-1.73^{+0.09}_{-0.08}} {}^{\circ}, \quad (2)$$

for the average multiplicity-dependent inclination and

$$\mu_{e,n} = 0.031^{+0.004}_{-0.003} \times \left(\frac{N}{5}\right)^{-1.74^{+0.11}_{-0.07}} \quad (3)$$

for the average multiplicity-dependent eccentricity. Rewriting these trends in the general form $y = a \times C^b$ where $C = N/5$ is a constant factor for each planetary system, the 1σ relative variability is $\sigma_{\mathrm{rel},y} = \sqrt{(\sigma_a/a)^2 + (\ln(C)\sigma_b)^2}$. The eccentricity and inclination of the individual planets in each system were computed assuming Gaussian distributions truncated to 0, centred on the average values set by $\mu_{\mathrm{e,n}}$ and $\mu_{\mathrm{i,n}}$, and with the standard deviation associated with the relevant relative uncertainty $\sigma_{\mathrm{rel},y}$. The queries to create samples E and S include, instead, systems where the inclination values are not known for all planets. To compute the NAMD of these systems, we took advantage of the knowledge of the orbital eccentricity, as it is known by definition for these samples, and assumed the Gaussian distributions of inclination centred on 0 with standard deviations equal to the modulus of the arcsin of the eccentricity $\sigma_{\mathrm{i}} = |\arcsin e|$ (ref. 4). This relation arises from the equipartition of the dynamical excitation $\sqrt{e^2 + \sin^2(i)}$ (ref. 50) along all secular degrees of freedom of the system[2].

Normalized parameters and uncertainty propagation

To generate the Monte Carlo samples and quantify the uncertainties of the normalized parameters used in the analysis, we followed the Monte Carlo approach introduced by ref. 2 and updated in ref. 4. For each planetary system, we generated 10^5 synthetic copies where we randomly extracted the planetary and stellar parameters. Unless otherwise stated, we assumed for all planetary and stellar quantities Gaussian distributions truncated to 0 and centred on the nominal values, with standard deviations equal to half the full 1σ, 68% confidence interval. We quantified the nominal values and 1σ uncertainties of all normalized parameters from the modal values and the 16th and 84th percentiles of their posterior Monte Carlo distributions.

The normalized parameters adopted to quantify the basic physical properties of the multi-planet systems in the sample are the normalized planetary system mass NSM[11,15] and the orbital period of the outermost planet P_{out}, by analogy with the classical mass–period classification. The NSM metric is defined as

$$\mathrm{NSM} = \Sigma_i m_i / M_*, \quad (4)$$

where m_i are the planetary masses, M_* is the mass of the host star and the index i ranges from 1 to the planetary multiplicity N of the system. The nominal mass of the host star is that resulting from fitting the isochrone. The P_{out} metric is defined as

$$P_{\mathrm{out}} = \sqrt{\max(a_i)^3 / M_*} \text{ years}, \quad (5)$$

where the planetary semimajor axes a_i are in astronomical units and M_* is in solar masses. The normalized parameters used to characterize the global architecture of the planetary systems in the sample are the planetary multiplicity N (ref. 20), the fraction of the total planetary mass contained in the most massive planet S_{m} (ref. 20), the orbital spacing metric S_{s} (ref. 20) and the normalized angular momentum deficit NAMD (refs. 4,19,20) quantifying the dynamical excitation. In this study, we consider the planetary multiplicity N to be the number of known planets in each system, meaning that $N > 1$ by definition, as we focus on multi-planet systems. The mass partition S_{m} metric is defined as

$$S_{\mathrm{m}} = m_{\mathrm{mp}} / \Sigma_i m_i, \quad (6)$$

where m_{mp} is the mass of the most massive planet. The orbital spacing metric S_{s} is defined as

$$S_{\mathrm{s}} = \frac{6}{N-1}\left(\frac{a_{\max} - a_{\min}}{a_{\max} + a_{\min}}\right)\left(\frac{3M_*}{2\overline{m}}\right)^{1/4}, \quad (7)$$

where $\overline{m} = \Sigma_i m_i / N$ is the mean mass of the planets, and $a_{\max}$ and $a_{\min}$ are the largest and shortest semimajor axes of the planets in the system. As discussed by ref. 20, this metric is also connected to the stability of multi-planet systems. Finally, NAMD is defined as

$$\mathrm{NAMD} = \frac{\sum_i m_i \sqrt{a_i}\left(1 - \sqrt{1 - e_i^2}\cos i_i\right)}{\sum_i m_i \sqrt{a_i}}, \quad (8)$$

where e_i and i_i are the orbital eccentricity and inclination values of the planets, respectively. Following refs. 4,19,26, we adopted the NAMD of the Solar System[4] (1.27×10^{-3}) as the reference value for the transition between dynamical order and chaotic evolution (see refs. 2,51,52 and references therein).

For the masses of planets and stars we reused the values generated in the computation of NSM and S_{m}. The shortest semimajor axis $a_{\min}$ is the semimajor axis of the innermost planet for each synthetic planetary system. The values of $a_{\max}$ were computed from the 10^5 Monte Carlo values of P_{out} by inverting equation (5).

For the NAMD, we adopted the methodology from ref. 2 as expanded by ref. 4. We reused the 10^5 Monte Carlo extractions of the planetary masses m_i and the semimajor axes a_i described above and integrated them with 10^5 Monte Carlo extractions of the eccentricity and inclination values of all planets in the system. For those systems for which we possess inclinations for all planets, we assumed Gaussian distributions centred on the relative inclinations with respect to the most massive planet in the system and standard deviations equal to half the reported 68% confidence intervals[2,3,26]. For the systems for which we lack the eccentricity or inclination values of some planets, we followed the different approaches described in 'Treatment of the missing eccentricities and inclinations in samples L, E and S' for samples L and H and for samples E and S. For the systems in samples E and S for which we know the eccentricities of all planets, we followed the approach described by ref. 2 and assumed Gaussian distributions for the rectangular eccentricity coordinates ($e\cos\omega$, $e\sin\omega$) where ω is the argument of periapsis. As the average value of ω did not influence the computation of the eccentricity distribution, its value was set to $\omega = 0$ (ref. 2). The distribution of the first rectangular eccentricity coordinate $\tilde{e} = e\cos\omega$ is therefore centred on the mean value $\langle e\rangle$ of the eccentricity, and the distribution of second rectangular eccentricity coordinate $e\sin\omega$ is centred on 0. The standard deviations of both distributions were set to half the reported 68% eccentricity confidence interval. The distribution of e was derived from that of $\tilde{e}$ through the relation $e = \sqrt{\tilde{e}^2 + (\tilde{e} - \langle e\rangle)^2}$ (ref. 2). In rare cases where the resulting value of e was outside the valid range [0:1), we repeated the Monte Carlo extraction of $\tilde{e}$.

The normalized parameters used to characterize the host stars are the stellar metallicity [Fe/H], the stellar mass M_* and the total PBM. For M_*, we reused the 10^5 Monte Carlo values extracted for the computation of the NSM. For [Fe/H], we used the nominal values and uncertainties reported by SWEET-Cat. We used the synthetic values of stellar mass and metallicity to quantify PBM:

$$\mathrm{PBM} = \alpha + \beta \log_{10}(M_*) + [\mathrm{Fe/H}]\ \mathrm{dex}. \quad (9)$$

This dependence of PBM on the parameters of the host star is based on the joint insights provided by the study of protoplanetary disks and of star–planet correlations. The first two terms on the right-hand side of equation (2) arise from the relation between the stellar mass and the dust mass in protoplanetary disks found by recent surveys[21,22], with α and β being the fit parameters of this empirical correlation. In our analysis, we adopted the values $\alpha = 1.1 \pm 0.2$ and $\beta = 1.9 \pm 0.4$ computed for the 1–2-Myr-old star-forming region L1688 by ref. 22. We extracted the 10^5 synthetic values of α and β from Gaussian distributions centred on the nominal values and with standard deviations equal to the uncertainties reported by ref. 22. We verified with numerical experiments that the specific choice of the star-forming region adopted as a reference did not affect the analysis, as its effect was simply to shift the absolute values of the PBM of the multi-planet systems. The second term in the right-hand side of equation (2) links the abundance of solid material in protoplanetary disks to the metallicity of their stars. The stellar metallicity is a proxy for the total abundance of heavy elements that can condense as solids in protoplanetary disks, and recent studies of the impact of stellar properties on planet formation highlight its important role in shaping the frequency and mass of planets (for example, refs. 25,29–31 and references therein).

Binning methodology

Owing to the large differences in the populations of the four samples, to plot the distributions of the normalized parameters in a uniform way we adopted for the optimal number of bins k in each plot the minimum between the values provided by Sturge's formula ($k = \log_2(n) + 1$), the Rice rule ($k = 2\sqrt[3]{n}$) and the Terrell–Scott rule ($k = \sqrt[3]{2n}$), after rounding to their least greater integers, where n is the total number of systems in the plot. Depending on the specific plot, therefore, n is either the total population of the sample or the combined population of lower-mass and high-mass systems, i.e. excluding the systems with uncertain attribution.

Statistical characterization of the two families and their properties

We identified the lower- and high-mass families by investigating the sample of multi-planet systems in the P_{out}–NSM parameter space. We first analysed the distribution of each parameter individually using one-sample KS tests with the `ks.test` algorithm in the `dgof` package (https://www.rdocumentation.org/packages/dgof/versions/1.5.1/topics/ks.test) to test whether both NSM and P_{out} are either normally or uniformly distributed. The comparison of both distributions provided P values much lower than 0.05 (Supplementary Table 3, top section) and rejected the hypothesis that either P_{out} or NSM are normally or uniformly distributed. We then focused exclusively on the NSM metric and fitted a bimodal distribution to its values using a Gaussian mixture model implemented via `GaussianMixture` in the Python `sklearn.mixture` package[53]. We used a model with two components and max_iteration = 1,000 points, and we filtered out the systems from the sample that fell in the region where the two Gaussian components had comparable probabilities. We report in Supplementary Table 2 the number of systems U_{sys} falling within this range, which we label as the 'uncertain membership region'.

We then performed a statistical comparison between the two families of systems: the one with lower-mass architectures and the one with high-mass architectures (Fig. 1 and main text). These families were defined by the limits given by the Gaussian mixture fit of the NSM. We tested the hypothesis that the two families come from the same underlying distribution in terms of planetary architectures and stellar hosts by performing a two-sample KS test on all metrics: S_m, S_s, P_{out}, N, NAMD and stellar properties: M_*, [Fe/H], and PBM. We applied multiple testing corrections to the two-sample KS test results using the Holm–Bonferroni procedure, which controls the family-wise error rate in a stepwise manner while retaining the diagnostic power of the stricter Bonferroni method, yet without the same high rejection rate of true results. We also used the B-H procedure, which controls the false discovery rate under independence or certain forms of positive dependence among tests. For sample L, all results reported $P < 0.05$ (Supplementary Table 3), thus robustly supporting that the two families have intrinsically different properties. We repeated the same analysis on the H, E and S samples, whose P values always proved significant. The largest P values among all the metrics were 1.330×10^{-3} for the H sample (multiplicity N), 1.325×10^{-2} for the S sample (PBM metric) and finally 2.061×10^{-2} for the E sample (multiplicity N). When any of the presented metrics contained ties, we avoided the KS test, as it can yield inaccurate P values in such cases. Instead, we applied the non-parametric two-sample Wilcoxon–Mann–Whitney test by introducing an infinitesimal jitter (10^{-10}) to both distributions to break the ties. The test was implemented using `wilcox.test` from the `stats` package in R[54,55] (https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/wilcox.test).

Next, we tested whether the bimodality is a function of the stellar evolutionary phase using the H sample. Among the 189 multi-planet systems, 87.3% have main sequence stars. In this sample, only one lower-mass system has an evolved host star, whereas the high-mass family contains 23 evolved stars (Supplementary Fig. 2). Given that most of our sample are main sequence stars, we performed our tests on the main sequence subsample alone by conservatively removing all stars with $\log g < 4$ from both lower- and high-mass families. This subsample includes 90 high-mass systems and 67 lower-mass systems around main sequence host stars. We note that both subsamples cover stellar spectral types from M to A. We performed the two-sample statistical test (either KS or Wilcoxon–Mann–Whitney) for S_s, S_m, [Fe/H], P_{out}, PBM, NAMD, M_* and N on these two populations. Each test (Holm–Bonferroni or B-H) resulted in $P < 0.05$ (the largest P value being 4.125×10^{-3} for N from the B-H correction), showing that the bimodality did not depend on the evolutionary stage of the host stars.

Finally, we used a Monte Carlo-resampled AD statistical pipeline to verify whether the different properties of the lower- and high-mass systems could have arisen from the uncertainties on the planetary and stellar parameters while simultaneously controlling for the false positive risks inherent in multidimensional parameter spaces. For the parameters spanning several orders of magnitude (NAMD, NSM and P_{out}), we applied logarithmic transformations ($\log_{10}$) to the values, as in the distributions shown in Figs. 1 and 3. To account for measurement uncertainties, we modelled the probability density function of each parameter for each individual system using a Monte Carlo approach. For parameters with symmetric uncertainties, we generated 10,000 synthetic realizations drawn from a standard Gaussian distribution. For parameters exhibiting asymmetric uncertainties (for example, NAMD), we used split-normal distributions, randomly drawing 10,000 samples weighted by the independent upper and lower 1σ error boundaries. All distributions were truncated to 0 whenever necessary to prevent unphysical results. Deterministic parameters with no reported uncertainties (the planetary multiplicity N) were treated as static delta functions.

To determine whether the high- and lower-mass systems were drawn from the same underlying parent population, we employed the non-parametric k-sample AD test[56] using the scipy.stats module within the SciPy module[57] for scientific computing in Python. We independently executed the AD test on all 10,000 synthetic sets of parameters for the two families, yielding a distribution of P values. We extracted the 95th percentile of this P value distribution and considered each parameter to be provisionally distinct if this upper-bound P value fell below the standard significance threshold ($P < 0.05$). We verified the mathematical stability of this threshold by calculating the running 95th percentile across the iterations, ensuring all parameters reached asymptotic convergence well before the 10,000 iterations limit. Because we simultaneously tested nine physically coupled

architectural parameters, we corrected for the probability of obtaining a spurious $P < 0.05$ result due to random chance (type I error) by subjecting the 95th percentile P values to the B-H procedure[58]. We classified a parameter as fundamentally distinct between the two populations only if its B-H-adjusted 95th percentile Monte Carlo P value remained <0.05. We applied this Monte Carlo-resampled AD procedure to all samples and verified that the two families always remain statistically different: the resulting distributions and the associated P values are reported in Extended Data Figs. 4–7. Because the scipy.stats.anderson_ksamp function hard caps its lowest possible reported P value at 0.001, when the output of the AD test corrected for the false discovery rate fell at this lowest bound, it is reported as $p_{FDR} < 0.001$ in the plots.

Dynamical simulations of the possible origins of the excited architectures in the sample

Our exploratory campaign of instability simulations focused on investigating to what extent the dynamical excitation of the sample can inform us about the nature of the dynamical instabilities that caused it, that is whether they originated internally or externally to the planetary systems. Thus, we simulated the dynamical evolution of hundreds of synthetic planetary systems, sampling plausible initial architectures for both the lower- and high-mass families, and compared their final NAMD values with those of the real systems in samples L and E.

The synthetic lower- and high-mass planetary systems were described by the following common properties:

- High-mass systems: P_{out} = 5.25 years and NSM = 560 $M_{\oplus}$
- Lower-mass systems: P_{out} = 0.1 years = 36.5 days and NSM = 33 $M_{\oplus}$
- Host stars: solar mass and radius (the latter quantity being used to test the possibility of stellar engulfments)

These properties position the lower-mass synthetic systems near the abundance peaks of their real counterparts in Fig. 1a,b. For high-mass synthetic systems, as NSM values close to the modal peak of the high-mass family proved extremely prone to becoming immediately unstable even when placed in resonant configurations, we focused the simulations on a lower NSM value. We adopted the following initial orbital architectures for both the lower- and high-mass synthetic systems:

(1) An initially unstable non-resonant configuration (at least one planet in the system is closer to the nearest planetary companion than the critical separation for stability[59])
(2) An incomplete resonant chain, such that the outermost planet is slightly shifted outwards with respect to the orbit that would result in a full resonant chain
(3) Stable non-resonant configurations perturbed by the passage of a stellar intruder (all neighbouring pairs of planets in the system are more distant than the critical separation for stability[59])
(4) A stable resonant chain perturbed by the passage of a stellar intruder

As the instability simulations were solely focused on exploring the most probable instability trigger for lower- and high-mass systems, we adopted equal-mass planets as the reference set-up for all four initial orbital architectures. This choice was motivated by the observation that this initial state is compatible with systems belonging to both families (Fig. 2b and Extended Data Figs. 1–3). For scenarios 2 and 4, we focused on the 2:1 resonance after we verified with numerical experiments that it guarantees the highest initial stability for both lower- and high-mass systems.

To explore the impact of different mass partitionings among the planets, for scenario 4 we also considered the mass partition of the outer Solar System, that is 70% of the planetary mass is in the most massive object. This choice is again motivated by the observation that this initial state is compatible with systems belonging to both families (Fig. 2b and Extended Data Figs. 1–3). This choice also allowed us to take advantage of the extensive literature on the dynamical history of the Solar System and use one of its proposed initial architectures as a template for our synthetic lower- and high-mass systems (scenario 4-ss). Specifically, we followed refs. 28,52 and adopted the initial resonant five-planet configuration from ref. 60, which involves a mixed 2:1 and 3:2 resonant chain with another now-lost ice giant planet, and we scaled the planetary masses and the orbital period of the outermost planet to fit the requirements set for the synthetic lower- and high-mass systems[28]. For each architecture (scenarios 1–4 and 4-ss), we ran 100 instability simulations.

With the exception of scenario 4-ss, for which the initial number of planets was always set to 5, we randomly extracted the initial planetary multiplicity from a uniform distribution between 3 and 6. This choice was motivated by the observed NAMD–multiplicity anticorrelation among multi-planet systems[4,19,26], which suggests that highly dynamically excited systems could have originally possessed more planets, as also illustrated by the dynamical studies by refs. 27,28. In scenarios 2, 4 and 4-ss, as the resonances set the spacing between the involved planets, we randomly varied the orbital angles to start the planets in different initial conditions. The stellar encounters that perturbed the stable systems in scenarios 3 and 4 were taken from the simulations of refs. 28,52. The time step of all simulations was set to 5% of the orbital period of the innermost planet. All instability simulations were run for 10 Myr, and we recorded all planetary collisions, ejections or impacts on the host star. At the end of the simulations, we characterized the outcomes in terms of the NAMD of the final planetary architecture, which we compare with the NAMD values of the real planetary systems in Fig. 3 and Extended Data Fig. 2 and with the number of planets lost through any of the previous processes. The results of this comparison are shown in Fig. 5.

Data availability

The exoplanetary data were obtained from the public database NASA Exoplanet Archive at https://exoplanetarchive.ipac.caltech.edu/ (ref. 24) through the queries provided in Methods. The stellar data for the host stars in sample H were obtained from the public database SWEET-Cat at http://sweetcat.iastro.pt) (refs. 17,18). The values used in the Monte Carlo modelling are reported in Supplementary Information. All data have been processed using methods and software that are publicly available, are described in peer-reviewed publications cited in Methods, or are introduced and fully described in Methods. The four architecturally characterized samples analysed in the study are available via Zenodo at https://doi.org/10.5281/zenodo.21492978 (ref. 61).

Code availability

The codes used for the analysis and described in Methods are available upon reasonable request from the corresponding authors.

References

1. Limbach, M. A. & Turner, E. L. Exoplanet orbital eccentricity: multiplicity relation and the Solar System. *Proc. Natl Acad. Sci. USA* **112**, 20–24 https://doi.org/10.1073/pnas.1406545111 (2015).
2. Laskar, J. & Petit, A. C. AMD-stability and the classification of planetary systems. *Astron. Astrophys.* **605**, A72 https://doi.org/10.1051/0004-6361/201630022 (2017).
3. Zinzi, A. & Turrini, D. Anti-correlation between multiplicity and orbital properties in exoplanetary systems as a possible record of their dynamical histories. *Astron. Astrophys.* **605**, L4 https://doi.org/10.1051/0004-6361/201731595 (2017).
4. Turrini, D., Zinzi, A. & Belinchon, J. A. Normalized angular momentum deficit: a tool for comparing the violence of the dynamical histories of planetary systems. *Astron. Astrophys.* **636**, A53 https://doi.org/10.1051/0004-6361/201936301 (2020).
5. Gajdoš, P. & Vaňko, M. Chaos in multiplanetary extrasolar systems. *Mon. Not. R. Astron. Soc.* **518**, 2068–2075 https://doi.org/10.1093/mnras/stac3200 (2023).

6. Muresan, A., Persson, C. M. & Fridlund, M. Diversities and similarities exhibited by multi-planetary systems and their architectures. I. Orbital spacings. *Astron. Astrophys.* **692**, A122 https://doi.org/10.1051/0004-6361/202451353 (2024).
7. Shariat, C., Hasegawa, Y., Hansen, B. M. S., Yu, T. Y. M. & Hu, R. Predicting the dominant formation mechanism of multiplanetary systems. *Astrophys. J. Lett.* **964**, L13 https://doi.org/10.3847/2041-8213/ad2c8d (2024).
8. Howe, A. R., Becker, J. C., Stark, C. C. & Adams, F. C. Architecture classification for extrasolar planetary systems. *Astron. J.* **169**, 149 https://doi.org/10.3847/1538-3881/adabdb (2025).
9. Xie, J.-W. et al. Exoplanet orbital eccentricities derived from LAMOST-Kepler analysis. *Proc. Natl Acad. Sci. USA* **113**, 11431–11435 https://doi.org/10.1073/pnas.1604692113 (2016).
10. Weiss, L. M. et al. The California-Kepler Survey. V. Peas in a pod: planets in a Kepler multi-planet system are similar in size and regularly spaced. *Astron. J.* **155**, 48 https://doi.org/10.3847/1538-3881/aa9ff6 (2018).
11. Gilbert, G. J. & Fabrycky, D. C. An information theoretic framework for classifying exoplanetary system architectures. *Astron. J.* **159**, 281 https://doi.org/10.3847/1538-3881/ab8e3c (2020).
12. He, M. Y., Ford, E. B., Ragozzine, D. & Carrera, D. Architectures of exoplanetary systems. III. Eccentricity and mutual inclination distributions of AMD-stable planetary systems. *Astron. J.* **160**, 276 https://doi.org/10.3847/1538-3881/abba18 (2020).
13. Ghosh, T. & Chatterjee, S. Orbital architectures of Kepler multis from dynamical instabilities. *Mon. Not. R. Astron. Soc.* **527**, 79–92 https://doi.org/10.1093/mnras/stad2962 (2024).
14. Gilbert, G. J., Petigura, E. A. & Entrican, P. M. No strong associations between eccentricity and orbital architecture in Kepler compact multis. *Astron. J.* **171**, 312 https://doi.org/10.3847/1538-3881/ae579e (2026).
15. Zhu, W. & Dong, S. Exoplanet statistics and theoretical implications. *Annu. Rev. Astron. Astrophys.* **59**, 291–336 https://doi.org/10.1146/annurev-astro-112420-020055 (2021).
16. Burn, R. & Mordasini, C. Planetary population synthesis. In *Handbook of Exoplanets* (eds Deeg, H. J. & Belmonte, J. A.) 1–60 https://doi.org/10.1007/978-3-319-30648-3_143-2 (Springer, 2025).
17. Santos, N. C. et al. SWEET-Cat: a catalogue of parameters for stars with exoplanets. I. New atmospheric parameters and masses for 48 stars with planets. *Astron. Astrophys.* **556**, A150 https://doi.org/10.1051/0004-6361/201321286 (2013).
18. Sousa, S. G. et al. SWEET-Cat 2.0: the cat just got SWEETer. Higher quality spectra and precise parallaxes from Gaia eDR3. *Astron. Astrophys.* **656**, A53 https://doi.org/10.1051/0004-6361/202141584 (2021).
19. Turrini, D. et al. Exploring the link between star and planet formation with Ariel. *Exp. Astron.* **53**, 225–278 https://doi.org/10.1007/s10686-021-09754-4 (2022).
20. Chambers, J. E. Making more terrestrial planets. *Icarus* **152**, 205–224 https://doi.org/10.1006/icar.2001.6639 (2001).
21. Pascucci, I. et al. A steeper than linear disk mass–stellar mass scaling relation. *Astrophys. J.* **831**, 125 https://doi.org/10.3847/0004-637X/831/2/125 (2016).
22. Testi, L. et al. The protoplanetary disk population in the ρ-Ophiuchi region L1688 and the time evolution of class II YSOs. *Astron. Astrophys.* **663**, A98 https://doi.org/10.1051/0004-6361/202141380 (2022).
23. Birnbaum, Z. W. & Tingey, F. H. One-sided confidence contours for probability distribution functions. *Ann. Math. Stat.* **22**, 592–596 https://doi.org/10.1214/aoms/1177729550 (1951).
24. Christiansen, J. L. et al. The NASA exoplanet archive and exoplanet follow-up observing program: data, tools, and usage. *Planet. Sci. J.* **6**, 186 https://doi.org/10.3847/PSJ/ade3c2 (2025).
25. Ramler, H. et al. Ariel stellar characterisation. IV. Fundamental parameters of 18 hot stars in the Ariel mission candidate sample. *Astron. Astrophys.* **710**, A297 https://doi.org/10.1051/0004-6361/202558131 (2026).
26. Bocchieri, A., Zak, J. & Turrini, D. ExoNAMD: leveraging the spin–orbit angle to constrain the dynamics of multi-planet systems. *Astron. Astrophys.* **706**, A222 https://doi.org/10.1051/0004-6361/202557279 (2026).
27. Nigioni, A., Turrini, D., Danielski, C., Polychroni, D. & Chambers, J. E. The quest for Magrathea planets. II. Orbital stability of exoplanets formed around double white dwarfs. *Astron. Astrophys.* **706**, A96 https://doi.org/10.1051/0004-6361/202452890 (2026).
28. Przyłuski, R. et al. Disruption of planetary system architectures by stellar flybys. *Universe* **11**, 240 https://doi.org/10.3390/universe11080240 (2025).
29. Magrini, L. et al. Ariel stellar characterisation. I. Homogeneous stellar parameters of 187 FGK planet host stars: description and validation of the method. *Astron. Astrophys.* **663**, A161 https://doi.org/10.1051/0004-6361/202243405 (2022).
30. Savvidou, S. & Bitsch, B. How to make giant planets via pebble accretion. *Astron. Astrophys.* **679**, A42 https://doi.org/10.1051/0004-6361/202245793 (2023).
31. Filomeno, S. et al. The GAPS programme at TNG. LXI. Atmospheric parameters and elemental abundances of TESS young exoplanet host stars. *Astron. Astrophys.* **690**, A370 https://doi.org/10.1051/0004-6361/202450611 (2024).
32. Tsantaki, M. et al. Ariel stellar characterisation. III. Fast rotators and new FGK stars in the Ariel mission candidate sample. *Astron. Astrophys.* **697**, A102 https://doi.org/10.1051/0004-6361/202453059 (2025).
33. Johansen, A. & Lambrechts, M. Forming planets via pebble accretion. *Annu. Rev. Earth Planet. Sci.* **45**, 359–387 https://doi.org/10.1146/annurev-earth-063016-020226 (2017).
34. Nelson, R. P. Planetary migration in protoplanetary disks. In *Handbook of Exoplanets* (eds Deeg, H. J. & Belmonte, J. A.) 139 https://doi.org/10.1007/978-3-319-55333-7_139 (Springer, 2018).
35. Millholland, S. C., Lara, T. & Toomlaid, J. Spin dynamics of planets in resonant chains. *Astrophys. J.* **961**, 203 https://doi.org/10.3847/1538-4357/ad10a0 (2024).
36. Zinzi, A., Turrini, D., Alei, E. & Verrecchia, F. ExoplAn3T: a new way of exploring large exoplanetary databases and its applications to astrobiology. *Mem. Soc. Astron. Ital.* **92**, 124–127 (2021).
37. Sousa, S. G. et al. Spectroscopic parameters for 451 stars in the HARPS GTO planet search program. Stellar [Fe/H] and the frequency of exo-Neptunes. *Astron. Astrophys.* **487**, 373–381 https://doi.org/10.1051/0004-6361:200809698 (2008).
38. Adibekyan, V. et al. A compositional link between rocky exoplanets and their host stars. *Science* **374**, 330–332 https://doi.org/10.1126/science.abg8794 (2021).
39. Danielski, C. et al. The homogeneous characterisation of Ariel host stars. *Exp. Astron.* **53**, 473–510 https://doi.org/10.1007/s10686-021-09765-1 (2022).
40. da Silva, R. et al. Ariel stellar characterisation. II. Chemical abundances of carbon, nitrogen, and oxygen for 181 planet-host FGK dwarf stars. *Astron. Astrophys.* **688**, A193 https://doi.org/10.1051/0004-6361/202450604 (2024).
41. da Silva, L. et al. Basic physical parameters of a selected sample of evolved stars. *Astron. Astrophys.* **458**, 609–623 https://doi.org/10.1051/0004-6361:20065105 (2006).
42. Rodrigues, T. S. et al. Bayesian distances and extinctions for giants observed by Kepler and APOGEE. *Mon. Not. R. Astron. Soc.* **445**, 2758–2776 https://doi.org/10.1093/mnras/stu1907 (2014).

43. Rodrigues, T. S. et al. Determining stellar parameters of asteroseismic targets: going beyond the use of scaling relations. *Mon. Not. R. Astron. Soc.* **467**, 1433–1448 https://doi.org/10.1093/mnras/stx120 (2017).
44. Skrutskie, M. F. et al. The Two Micron All Sky Survey (2MASS). *Astron. J.* **131**, 1163–1183 https://doi.org/10.1086/498708 (2006).
45. Gaia Collaboration Gaia Early Data Release 3. Summary of the contents and survey properties. *Astron. Astrophys.* **649**, A1 https://doi.org/10.1051/0004-6361/202039657 (2021).
46. Gaia Collaboration Gaia Data Release 3. The Galaxy in your preferred colours: Synthetic photometry from Gaia low-resolution spectra. *Astron. Astrophys.* **674**, A33 https://doi.org/10.1051/0004-6361/202243709 (2023).
47. Høg, E. et al. The Tycho-2 catalogue of the 2.5 million brightest stars. *Astron. Astrophys.* **355**, L27–L30 (2000).
48. Chen, Y. et al. YBC: a stellar bolometric corrections database with variable extinction coefficients. Application to PARSEC isochrones. *Astron. Astrophys.* **632**, A105 https://doi.org/10.1051/0004-6361/201936612 (2019).
49. Bressan, A. et al. PARSEC: stellar tracks and isochrones with the PAdova and TRieste Stellar Evolution Code. *Mon. Not. R. Astron. Soc.* **427**, 127–145 https://doi.org/10.1111/j.1365-2966.2012.21948.x (2012).
50. Petit, J.-M., Chambers, J., Franklin, F. & Nagasawa, M. Primordial excitation and depletion of the main belt. In *Asteroids III* (eds Bottke Jr, W. F. et al.) 711–723 (Univ. of Arizona Press, 2002).
51. Nesvorný, D. Dynamical evolution of the early Solar System. *Annu. Rev. Astron. Astrophys.* **56**, 137–174 https://doi.org/10.1146/annurev-astro-081817-052028 (2018).
52. Rickman, H. et al. Breakdown of planetary systems in embedded clusters. *Mon. Not. R. Astron. Soc.* **520**, 637–648 https://doi.org/10.1093/mnras/stac3705 (2023).
53. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
54. Bauer, D. F. Constructing confidence sets using rank statistics. *J. Am. Stat. Assoc.* **67**, 687–690 https://doi.org/10.1080/01621459.1972.10481279 (1972).
55. Hollander, M. & Wolfe, D. A. *Nonparametric Statistical Methods* (Wiley, 1973).
56. Scholz, F. W. & Stephens, M. A. *K*-sample Anderson–Darling tests. *J. Am. Stat. Assoc.* **82**, 918–924 https://doi.org/10.1080/01621459.1987.10478517 (1987).
57. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 https://doi.org/10.1038/s41592-019-0686-2 (2020).
58. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 https://doi.org/10.1111/j.2517-6161.1995.tb02031.x (1995).
59. Obertas, A., Van Laerhoven, C. & Tamayo, D. The stability of tightly-packed, evenly-spaced systems of Earth-mass planets orbiting a Sun-like star. *Icarus* **293**, 52–58 https://doi.org/10.1016/j.icarus.2017.04.010 (2017).
60. Nesvorný, D. & Morbidelli, A. Statistical study of the early Solar System's instability with four, five, and six giant planets. *Astron. J.* **144**, 117 https://doi.org/10.1088/0004-6256/144/4/117 (2012).
61. Turrini, D. & Danielski, C. Multi-criteria samples of architecturally-characterized multi-planet systems. *Zenodo* https://doi.org/10.5281/zenodo.21492978 (2026).
62. Chevance, M., Kruijssen, J. M. D. & Longmore, S. N. COOL Research DAO whitepaper—towards community-owned astrophysics for everyone. Preprint at https://doi.org/10.48550/arXiv.2501.13160 (2025).

Acknowledgements

We acknowledge comments and suggestions made by A. Quirrenbach, who provided the seed for the multi-sample design of the study. This work has been developed within the framework of the Ariel Planetary Formation and Stellar Characterization working groups of the ESA Ariel space mission consortium. COOL Research DAO[62] is a decentralized autonomous organization supporting research in astrophysics aimed at uncovering our cosmic origins.

Author contributions

D.T. and C.D. devised the study. D.T., C.D., H.R. and J.M.D.K. devised and selected the normalized parameters used in the analysis. A.Z. prepared the sample of planetary systems. C.D. and D.B. prepared the homogeneous stellar sample, and D.B. computed the homogeneous stellar mass. D.T. and D.P. developed the Monte Carlo code that computed the normalized parameters used in the analysis and their uncertainties. C.D., D.T., J.M.D.K., S.N.L., M.C. and L.R. performed the statistical analysis of the results. H.R., S.G., D.T., D.P. and R.P. designed the dynamical simulations used to study the roots of the dynamical excitation of the sample. D.P. and R.P. performed the dynamical simulations. All authors participated in interpreting the results and writing the article.

Funding

This research was supported by the International Space Science Institute in Bern through the International Teams Project Number 558 (Transformation of Planetary Systems by Environmental Perturbations), PI Kruijssen. D.T., D.P., D.B. and A.Z. acknowledge the support of the Italian Space Agency (ASI-INAF Agreement Number 2021-5-HH.0 and addenda 2021-5-HH.1-2022 and 2021-5-HH.2-2024). D.T. and D.P. acknowledge the support of the Italian National Institute of Astrophysics (INAF) through the INAF Main Stream project Ariel and the Astrochemical Link between Circumstellar Discs and Planets (CUP C54I19000700005) and the PRIN-INAF 2019 Planetary Systems at Young Ages (PLATEA). D.P. acknowledges the support of the Istituto Nazionale di Oceanografia e di Geofisica Sperimentale through the programme HPC-TRES (High Performance Computing Training and Research for Earth Sciences). C.D. acknowledges financial support from the grant RYC2023-044903-I funded by MCIU/AEI/10.13039/501100011033 and by the ESF+, and from the INAF initiative IAF Astronomy Fellowships in Italy (Grant Name GExoLife). D.B. acknowledges funding support by the Italian Ministerial Grant No. PRIN 2022, Radiative Opacities for Astrophysical Applications (2022NEXMP8 and CUP C53D23001220006). M.C. and L.R. gratefully acknowledge funding from the DFG through an Emmy Noether Research Group (CH2137/1-1). Open access funding provided by Istituto Nazionale di Astrofisica within the CRUI-CARE Agreement.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41550-026-02990-8.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41550-026-02990-8.

Correspondence and requests for materials should be addressed to Diego Turrini or Camilla Danielski.

Peer review information *Nature Astronomy* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

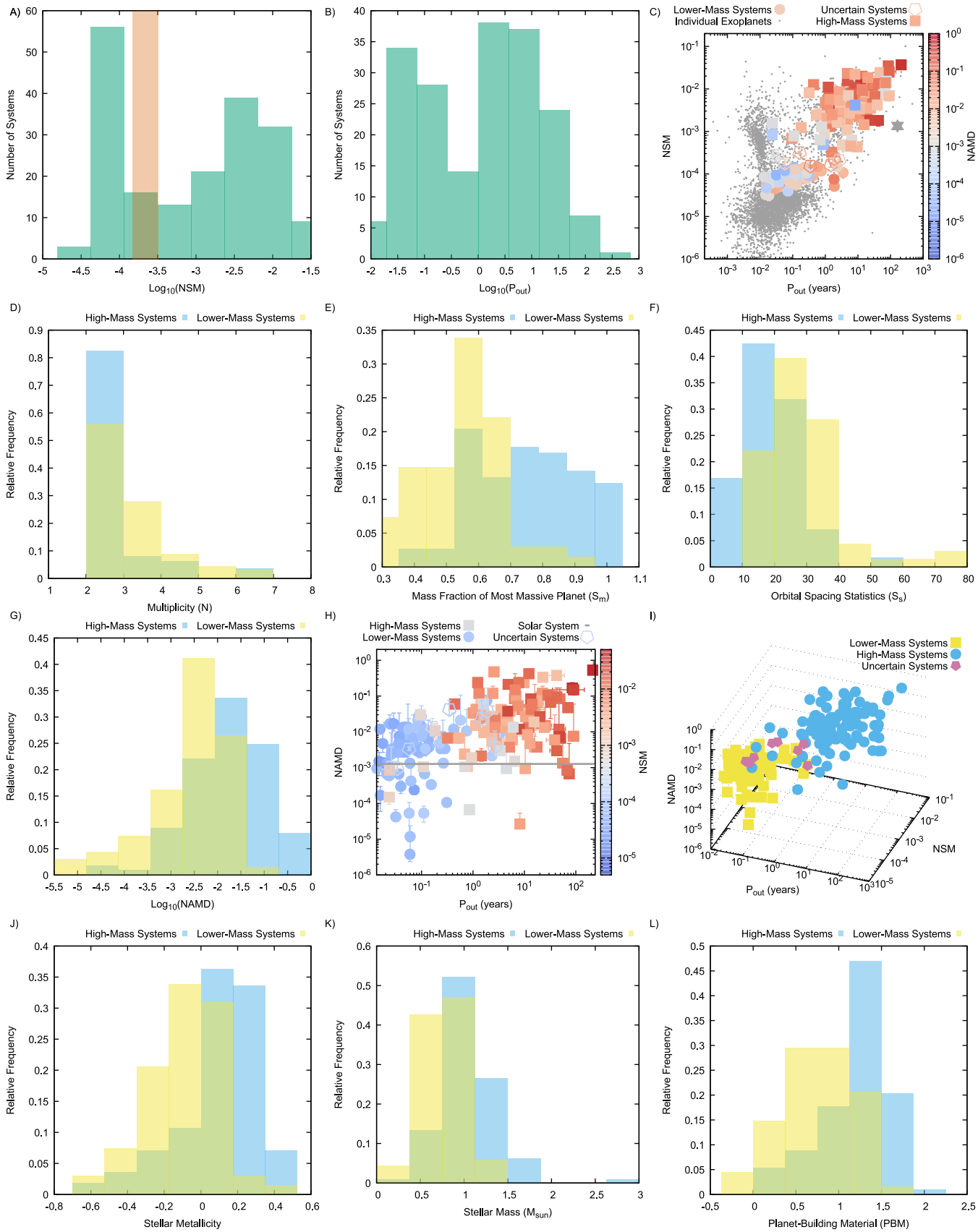
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2026

[1]Osservatorio Astrofisico di Torino, INAF, Turin, Italy. [2]Departament d'Astronomia i Astrofísica, Universitat de València, Burjassot, Spain. [3]Osservatorio Astrofisico di Arcetri, INAF, Florence, Italy. [4]Centrum Badań Kosmicznych Polskiej Akademii Nauk, Warsaw, Poland. [5]Department of Physics and Astronomy, Uppsala University, Uppsala, Sweden. [6]Dipartimento di Fisica e Astronomia Galileo Galilei, Universitá di Padova, Padua, Italy. [7]Osservatorio Astronomico di Padova, INAF, Padua, Italy. [8]Cosmic Origins Of Life (COOL) Research DAO, https://coolresearch.io/. [9]Agenzia Spaziale Italiana, Rome, Italy. [10]School of Mathematical and Physical Sciences, University of Sheffield, Sheffield, UK. [11]Astrophysics Research Institute, Liverpool John Moores University, Liverpool, UK. [12]Institut für Theoretische Astrophysik, Zentrum für Astronomie der Universität Heidelberg, Heidelberg, Germany. [13]These authors contributed equally: Diego Turrini, Camilla Danielski. ✉e-mail: diego.turrini@inaf.it; camilla.danielski@uv.es

https://doi.org/10.1038/s41550-026-02990-8

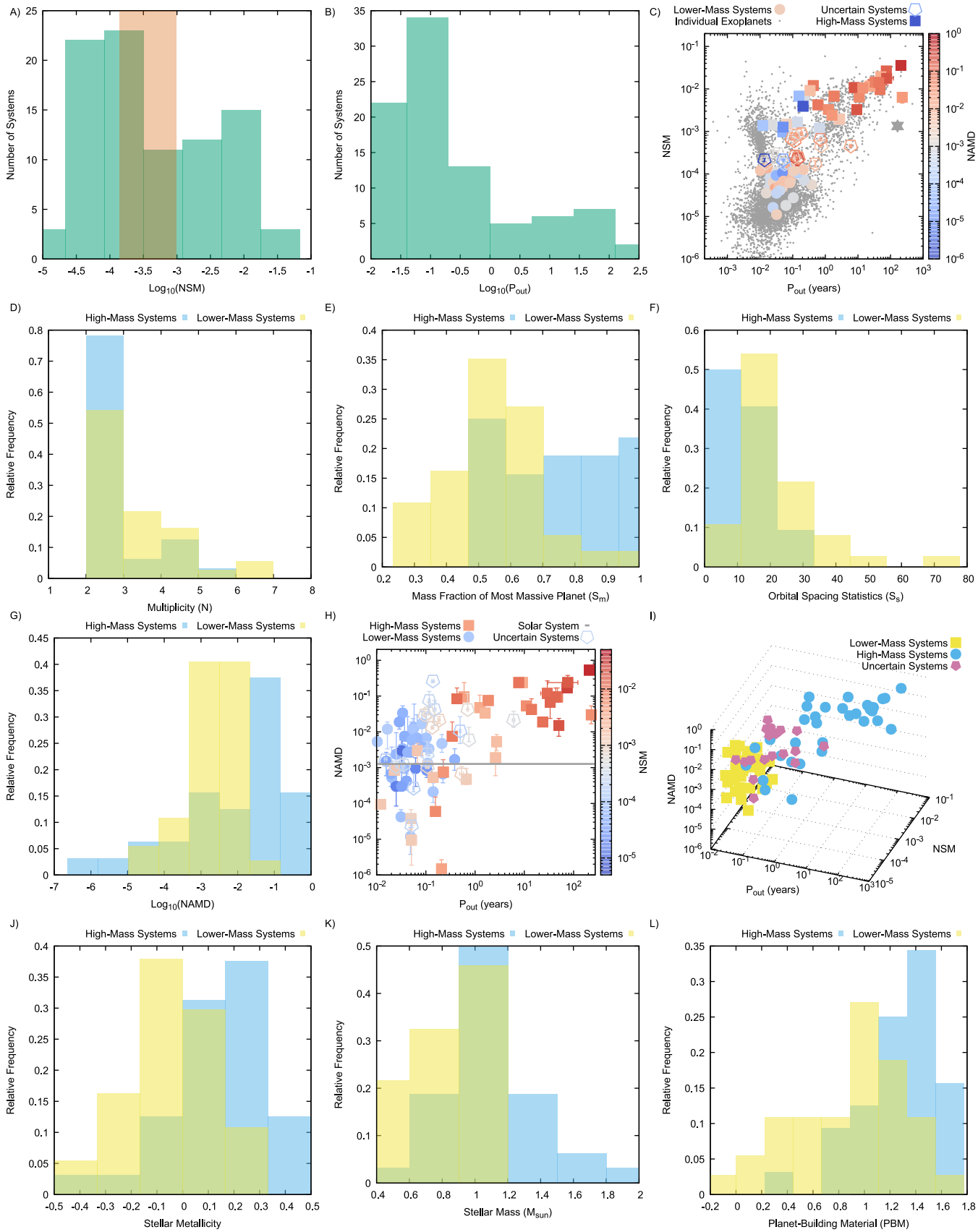


Extended Data Fig. 1 | See next page for caption.

https://doi.org/10.1038/s41550-026-02990-8

Extended Data Fig. 1 | Synoptic view of the analysis of sample H. A) Bimodal distribution of the normalized system masses (NSM). The region highlighted in red marks the mass range where systems can belong to the lower- and high-mass families with comparable probabilities (Methods). **B**) Bimodal distribution of the orbital periods of the outer planets (P_{out}). **C**) Comparison between the known exoplanet population and the high- and lower-mass families. Systems with uncertain attribution are identified by empty symbols, the Solar System is shown for comparison as the gray star. **D**) Distributions of the multiplicity (N) in the lower- and high-mass families of multi-planet systems. **E**) Distributions of the mass fraction of the most massive planets (S_m). **F**) Distributions of the orbital spacing metric (S_s). **G**) Distributions of the normalized angular momentum deficit (NAMD). **H**) NAMD - P_{out} scatter plot, the color scale shows the normalized system mass (NSM), while the dashed line shows the Solar System ($NAMD_{ss}$). **I**) 3D scatter plot in the NAMD-P_{out} -NSM parameter space. **J**) Distributions of the stellar metallicity ([Fe/H]). **K**) Distributions of the stellar masses (M_*). **L**) Distributions of the planet-building material (PBM). Data in panels **C**) and **H**) are presented as modal values with the error bars identifying the 16th and 84th percentiles.

https://doi.org/10.1038/s41550-026-02990-8

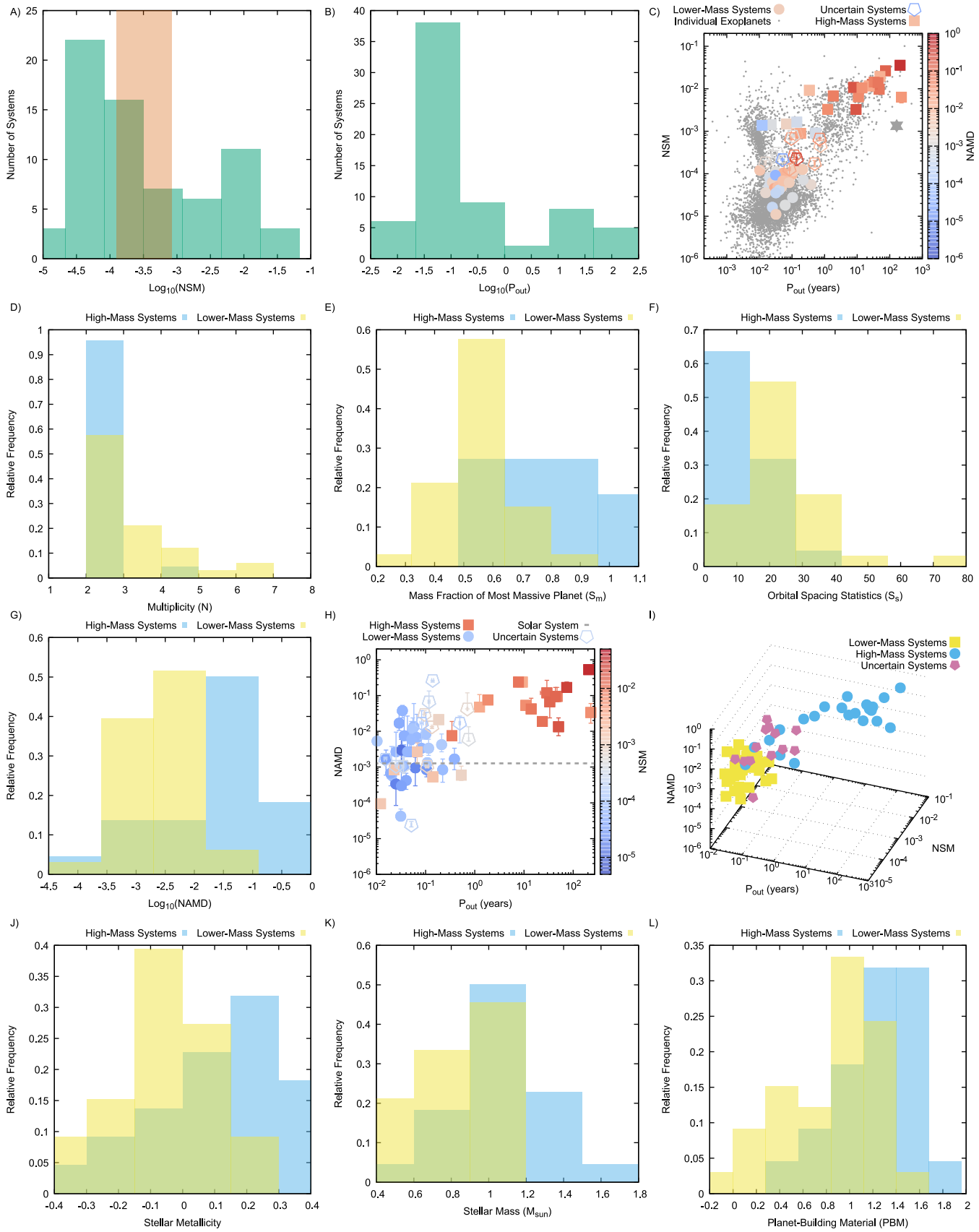


Extended Data Fig. 2 | See next page for caption.

https://doi.org/10.1038/s41550-026-02990-8

Extended Data Fig. 2 | Synoptic view of the analysis of sample E. A) Bimodal distribution of the normalized system masses (NSM). The region highlighted in red marks the mass range where systems can belong to the lower- and high-mass families with comparable probabilities (Methods). **B**) Bimodal distribution of the orbital periods of the outer planets (P_{out}). **C**) Comparison between the known exoplanet population and the high- and lower-mass families. Systems with uncertain attribution are identified by empty symbols, the Solar System is shown for comparison as the gray star. **D**) Distributions of the multiplicity (N) in the lower- and high-mass families of multi-planet systems. **E**) Distributions of the mass fraction of the most massive planets (S_m). **F**) Distributions of the orbital spacing metric (S_s). **G**) Distributions of the normalized angular momentum deficit (NAMD). **H**) NAMD - P_{out} scatter plot, the color scale shows the normalized system mass (NSM), while the dashed line shows the Solar System ($NAMD_{ss}$). **I**) 3D scatter plot in the NAMD-P_{out} -NSM parameter space. **J**) Distributions of the stellar metallicity ([Fe/H]). **K**) Distributions of the stellar masses ($M_\star$). **L**) Distributions of the planet-building material (PBM). Data in panels **C**) and **H**) are presented as modal values with the error bars identifying the 16th and 84th percentiles.

https://doi.org/10.1038/s41550-026-02990-8

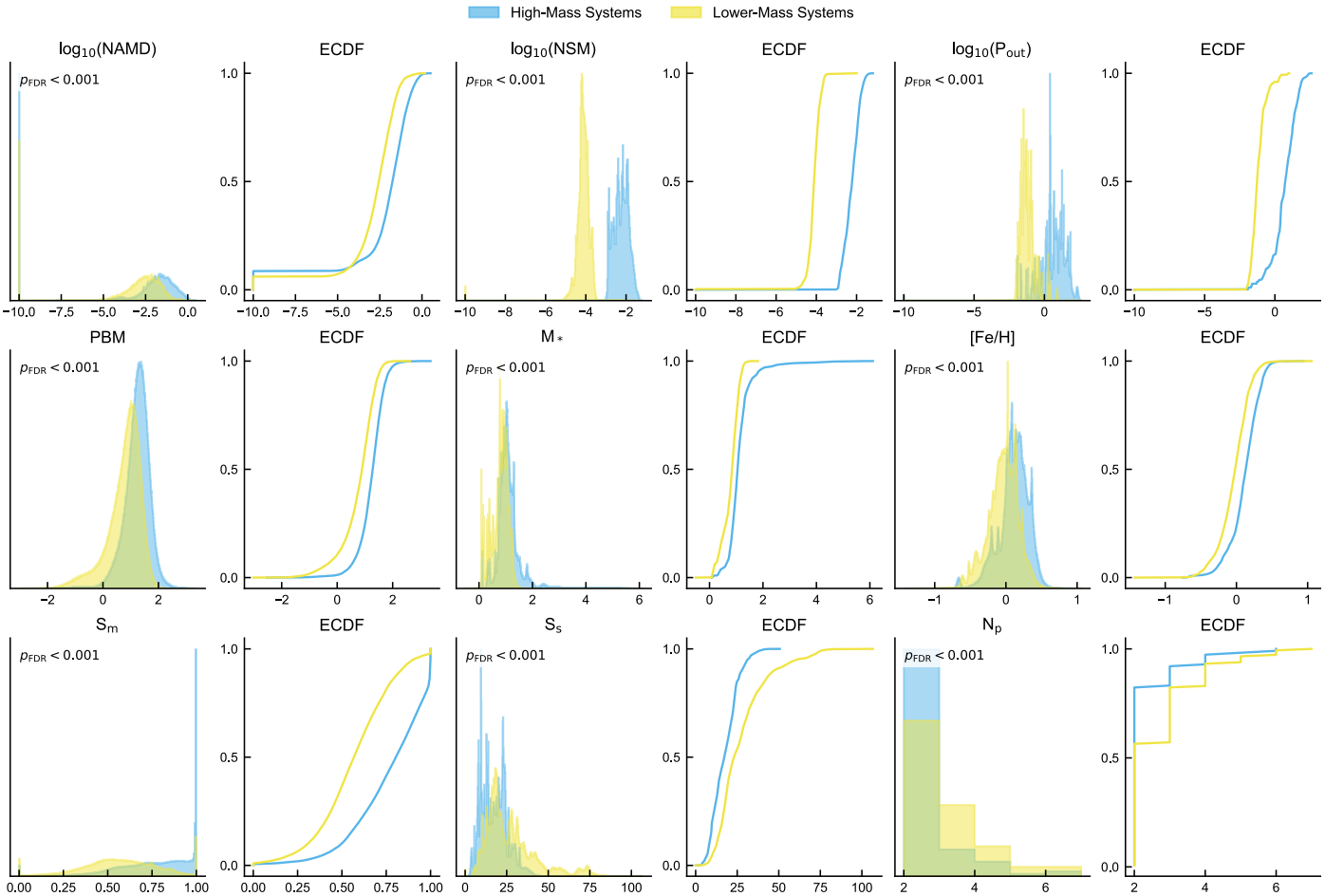


Extended Data Fig. 3 | See next page for caption.

https://doi.org/10.1038/s41550-026-02990-8

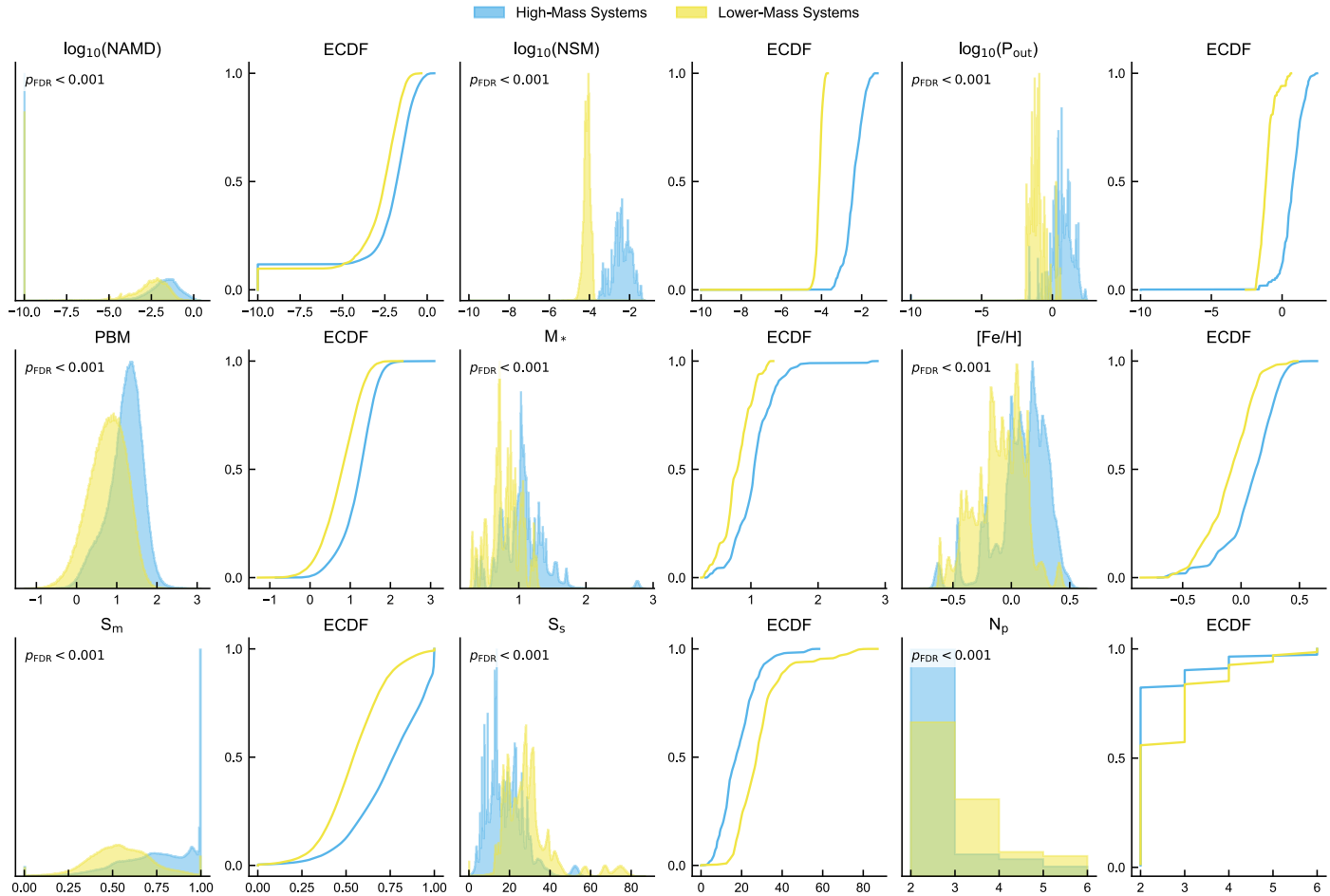
Extended Data Fig. 3 | Synoptic view of the analysis of sample S. A) Bimodal distribution of the normalized system masses (NSM). The region highlighted in red marks the mass range where systems can belong to the lower- and high-mass families with comparable probabilities (Methods). **B**) Bimodal distribution of the orbital periods of the outer planets (P_{out}). **C**) Comparison between the known exoplanet population and the high- and lower-mass families. Systems with uncertain attribution are identified by empty symbols, the Solar System is shown for comparison as the gray star. **D**) Distributions of the multiplicity (N) in the lower- and high-mass families of multi-planet systems. **E**) Distributions of the mass fraction of the most massive planets (S_m). **F**) Distributions of the orbital spacing metric (S_s). **G**) Distributions of the normalized angular momentum deficit (NAMD). **H**) NAMD - P_{out} scatter plot, the color scale shows the normalized system mass (NSM), while the dashed lines shows the Solar System ($NAMD_{ss}$). **I**) 3D scatter plot in the NAMD-P_{out} -NSM parameter space. **J**) Distributions of the stellar metallicity ([Fe/H]). **K**) Distributions of the stellar masses ($M_\cdot$). **L**) Distributions of the planet-building material (PBM). Data in panels **C**) and **H**) are presented as modal values with the error bars identifying the 16th and 84th percentiles.

https://doi.org/10.1038/s41550-026-02990-8



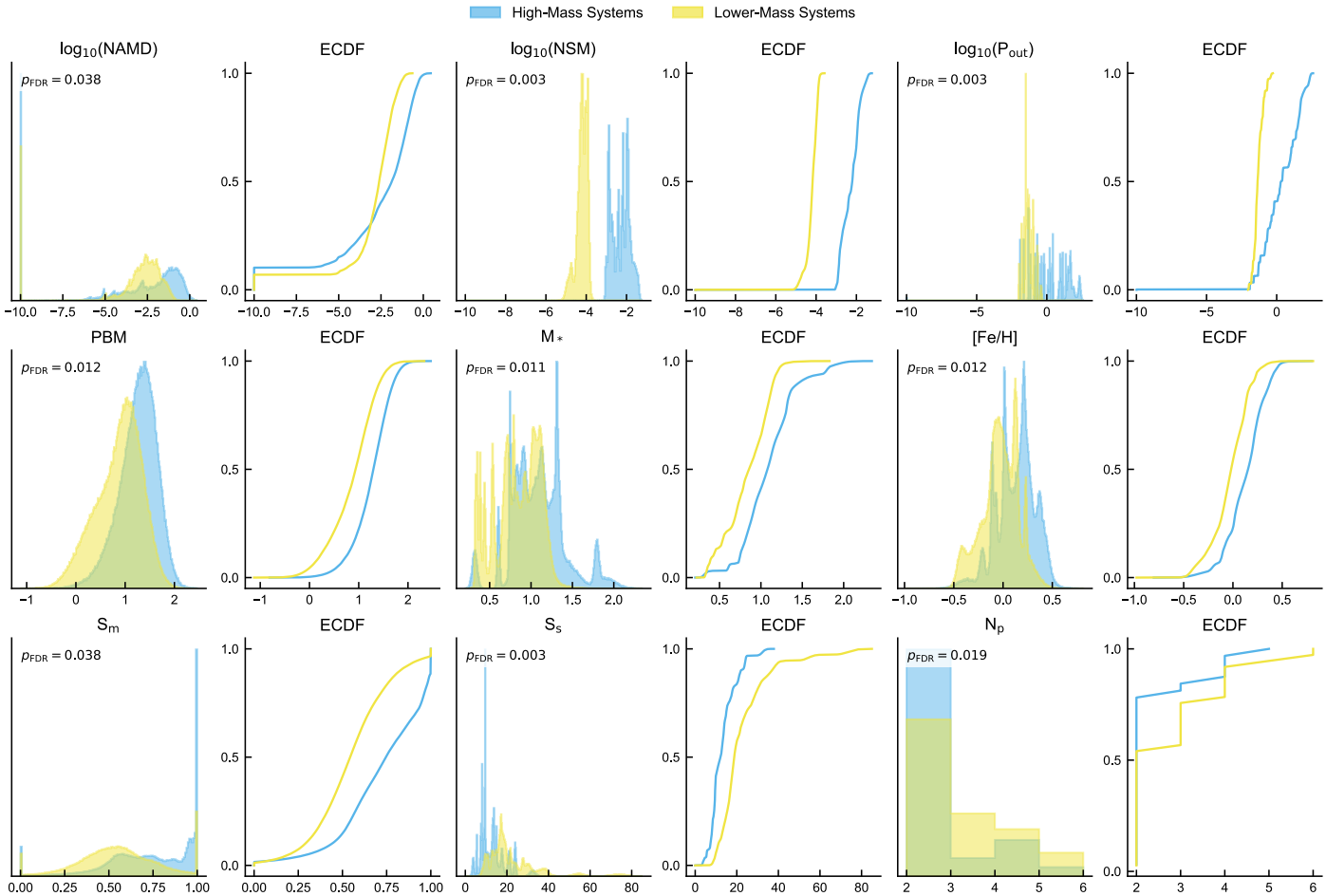
Extended Data Fig. 4 | Monte Carlo-resampled Anderson-Darling test for sample L. The figure displays the probability density histograms and Empirical Cumulative Distribution Functions (ECDFs) for the nine normalized parameters used in the analysis. High-mass systems are represented in light blue, while lower-mass systems are represented in light yellow. The p_{FDR} label located in the top-left corner of each histogram reports the Benjamini-Hochberg False Discovery Rate-corrected *P* value. The distributions and the associated *P* values are generated by iterative k-sample Anderson-Darling testing accounting for Monte Carlo propagation of the uncertainties.

https://doi.org/10.1038/s41550-026-02990-8



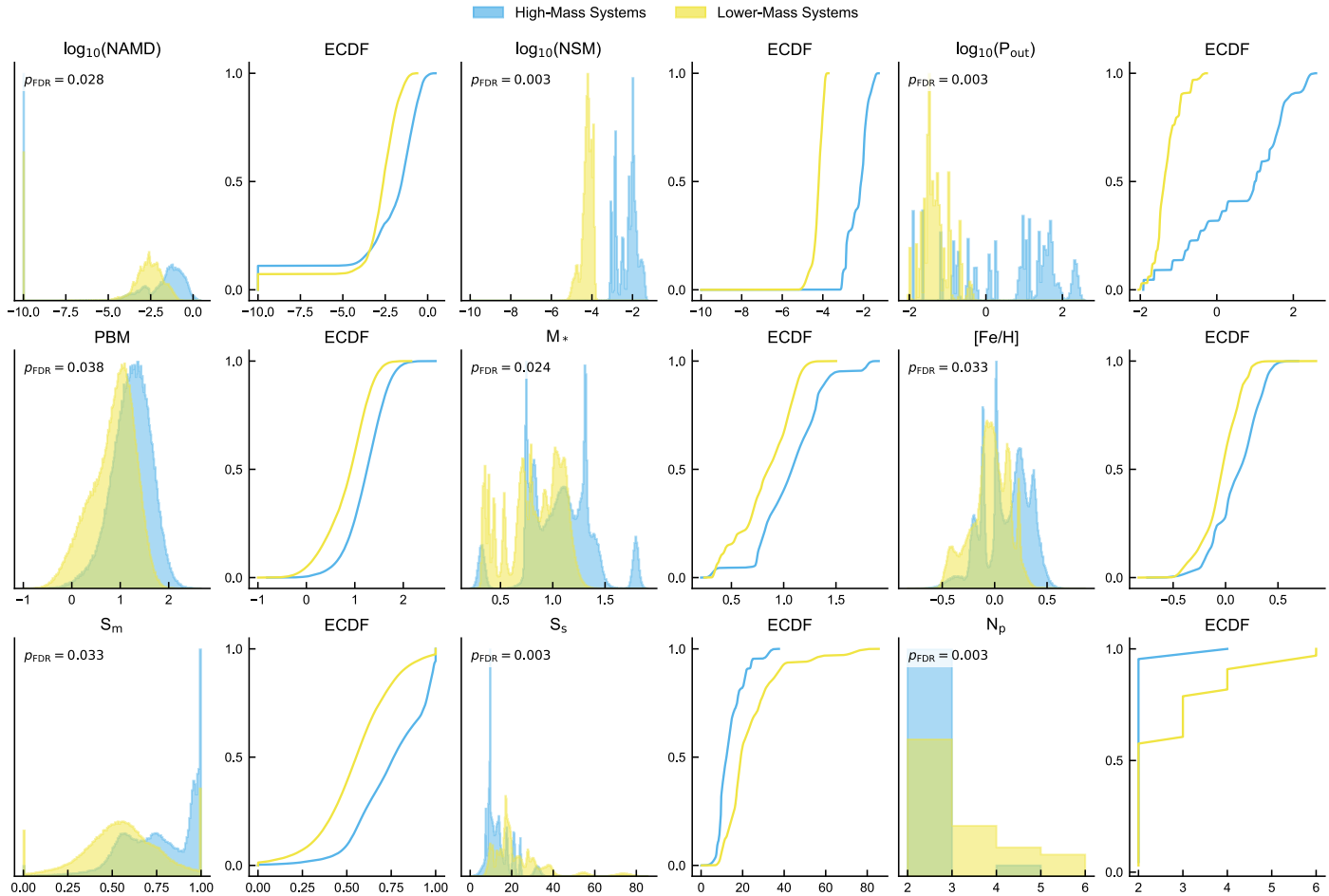
Extended Data Fig. 5 | Monte Carlo-resampled Anderson-Darling test for sample H. The figure displays the probability density histograms and Empirical Cumulative Distribution Functions (ECDFs) for the nine normalized parameters used in the analysis. High-mass systems are represented in light blue, while lower-mass systems are represented in light yellow. The p_{FDR} label located in the top-left corner of each histogram reports the Benjamini-Hochberg False Discovery Rate-corrected *P* value. The distributions and the associated *P* values are generated by iterative k-sample Anderson-Darling testing accounting for Monte Carlo propagation of the uncertainties.

https://doi.org/10.1038/s41550-026-02990-8



Extended Data Fig. 6 | Monte Carlo-resampled Anderson-Darling test for sample E. The figure displays the probability density histograms and Empirical Cumulative Distribution Functions (ECDFs) for the nine normalized parameters used in the analysis. High-mass systems are represented in light blue, while lower-mass systems are represented in light yellow. The p_{FDR} label located in the top-left corner of each histogram reports the Benjamini-Hochberg False Discovery Rate-corrected *P* value. The distributions and the associated *P* values are generated by iterative k-sample Anderson-Darling testing accounting for Monte Carlo propagation of the uncertainties.

https://doi.org/10.1038/s41550-026-02990-8



Extended Data Fig. 7 | Monte Carlo-resampled Anderson Darling test for sample S. The figure displays the probability density histograms and Empirical Cumulative Distribution Functions (ECDFs) for the nine normalized parameters used in the analysis. High-mass systems are represented in light blue, while lower-mass systems are represented in light yellow. The p_{FDR} label located in the top-left corner of each histogram reports the Benjamini-Hochberg False Discovery Rate-corrected *P* value. The distributions and the associated *P* values are generated by iterative k-sample Anderson-Darling testing accounting for Monte Carlo propagation of the uncertainties.