The Dispersed Matter Planet Project: Machine Learning applied to space telescope data to find transits and new host stars

**Supervision team:** Professor Carole Haswell, Professor Enrico Motta, Dr Mark H Jones and Dr John Barnes

**Lead contact:** Professor Carole Haswell

*Artist’s impression of the compact multiplanet system DMPP-1. © Mark Garlick*

**Project description:**

The Dispersed Matter Planet Project (DMPP) uses signatures of absorption by circumstellar gas to identify the probable host stars of hot rocky exoplanets for study in the coming era of exoplanetology (Haswell+2020). In our first small supersample of ~6000 stars, we identified 39 bright, nearby target stars. Our state-of-the-art programme of high cadence, high precision radial velocity (RV) observations has an essentially 100% record of detecting rocky exoplanets in short period orbits around them. Our first five planetary systems are DMPP-1: a compact multiplanet system, with multiple super-Earth planets orbiting a bright nearby star (Staab+2020); DMPP-2, the joint first RV discovery of a planet orbiting a strongly pulsating star (Haswell+2020); DMPP-3 which is a ~500d eccentric binary star system hosting one or two circum-primary super-Earths (Barnes+2020, Stevenson+2023); DMPP-4 is a northern hemisphere naked-eye host star of one or more sub-Neptune mass planets; DMPP-5 is a young compact multi-planet system in the Hyades open cluster (Ross+2024). Because the stars were identified by the signatures of absorbing gas ablated from the close-in planets,
these systems are amenable to transmission spectroscopy which can reveal the planet composition. Furthermore, angular momentum considerations suggest these planets have a high probability of transiting. Thus, potentially, our planet discoveries will yield planet masses, radii and compositions, all with small uncertainty ranges. This paves the way for comparative exogeology.

This project will exploit the rapidly growing archival holdings of uniform, high quality data from space telescopes.

- Gaia Data Release 3 provides precise distances and uniform high resolution spectra covering the CaII infrared triplet for a million stars. These data are revolutionising stellar astrophysics, and make it possible to identify stars with absorbing circumstellar gas for much larger samples than we have done so far. Crucially, this will allow us to extend the Dispersed Matter Planet Project to stars at a range of ages. The Neptune desert and the radius valley are prominent features in the demographics of short period exoplanets resulting from the loss of gaseous envelopes / atmospheres from hot planets. It seems likely that they are predominantly carved out early in the main sequence lifetime of the host star. A large-scale analysis of Gaia DR3 data for the signatures of circumstellar absorption should allow us to quantify this planetary mass loss as a function of host star age.

- The Transiting Exoplanet Survey Satellite (TESS) is observing most of the bright stars in the sky, searching for transiting exoplanets. There is a huge international collaboration working through the TESS data in a uniform way. This PhD project will complement the on-going large-scale survey data analysis. Using the proprietary DMPP target selection technique we know a priori the likely host stars of ablating, hot, rocky, transiting planets. Having only a few dozen targets rather than hundreds of thousands means we can sensibly invest significant effort in custom analysis of the TESS data for our targets. We have already demonstrated this works, with the discovery of a transit right at the detection threshold in TESS data on DMPP-1 (Jones+2020). This work drives down the detection threshold for transits in TESS lightcurves: it is among the most shallow transits yet detected in TESS data with a dip of only ~80 parts per million. We have a TESS guest investigator program which ensures the best possible cadence data is collected for our targets. We have secured a very significant award of CHaracterising ExOPlanet Satellite (CHEOPS) space telescope time to follow-up the DMPP-1 transit.

The PhD project will apply data science and machine learning techniques to these archival data. The aims are (i) to identify stars with anomalously low Ca II infrared triplet line cores in the Gaia data and (ii) perform the best possible detrending for targets in the TESS data, informed by the multidimensional time-series data for the targets themselves and other objects observed concurrently.

We expect the Gaia work to identify new bright, nearby target stars for intensive DMPP RV and transit studies along with many fainter potential targets which can be used to draw statistical inferences about planetary mass loss. The TESS work should find further potentially very informative transits, thus providing targets for James Webb Space Telescope (JWST) and
Ariel compositional studies. In addition, techniques for the removal of instrumental systematic effects in TESS data will be developed and assessed. We will produce high quality characterisation of the astrophysical photometric variability of the DMPP target stars. While less exciting than discovering transits, this will be vital for understanding the limits imposed by intrinsic stellar variability on exoplanet spectroscopic studies. We expect the work to open many research avenues; the most promising and immediately tractable of these will be explored in the context of the PhD work.

The work will include opportunities to travel to observatories, including the facilities of the European Southern Observatory in Chile, the South African Astronomical Observatory in the western Roggeveld Mountains in the Karoo, South Africa and Lowell Observatory in Arizona, USA.

The student will be expected to lead-author papers announcing planet discoveries and may present results at national and international exoplanet conferences.

References:


Names in bold are supervisors, underlined names are present and former PhD students or postdocs in Haswell's group.