

## Research Article

# EMU/GAMA: A technique for detecting active galactic nuclei in low mass systems

Jahang Prathap<sup>1,2,3</sup> , Andrew M. Hopkins<sup>1,2</sup>, Aaron S.G. Robotham<sup>4,3</sup>, Sabine Bellstedt<sup>4</sup>, José Afonso<sup>5,6</sup>, Ummee T. Ahmed<sup>7,8</sup> , Maciej Bilicki<sup>9</sup>, Malcolm N. Bremer<sup>10</sup>, Sarah Brough<sup>11</sup>, Michael J.I. Brown<sup>12</sup> , Yjan Gordon<sup>13</sup>, Benne W. Holwerda<sup>14</sup>, Denis Leahy<sup>15</sup> , Ángel R. López-Sánchez<sup>1,2,3</sup>, Joshua R. Marvil<sup>16</sup>, Tamal Mukherjee<sup>1,2,3</sup> , Isabella Prandoni<sup>17</sup>, Stanislav S. Shabala<sup>18</sup> , Tessa Vernstrom<sup>4</sup> and Tayyaba Zafar<sup>1,2,3</sup> 

<sup>1</sup>School of Mathematical and Physical Sciences, Macquarie University, Sydney, NSW, Australia, <sup>2</sup>Macquarie University Astrophysics and Space Technologies Research Centre, Sydney, NSW, Australia, <sup>3</sup>ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO-3D), Canberra, Australia, <sup>4</sup>International Centre for Radio Astronomy Research (ICRAR), The University of Western Australia, Crawley, WA, Australia, <sup>5</sup>Instituto de Astrofísica e Ciências do Espaço, Universidade de Lisboa, OAL, Tapada da Ajuda, Lisbon, Portugal, <sup>6</sup>Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, Lisbon, Portugal, <sup>7</sup>Australian Astronomical Optics, Macquarie University, North Ryde, NSW, Australia, <sup>8</sup>Centre for Astrophysics, University of Southern Queensland, Springfield Central, QLD, Australia, <sup>9</sup>Centre for Theoretical Physics, Polish Academy of Sciences, Warsaw, Poland, <sup>10</sup>H.H. Wills Physics Laboratory, University of Bristol, Bristol, UK, <sup>11</sup>School of Physics, University of New South Wales, Kensington, NSW, Australia, <sup>12</sup>School of Physics, Monash University, Clayton, VIC, Australia, <sup>13</sup>Physics Department, University of Wisconsin-Madison, Madison, WI, USA, <sup>14</sup>Department of Physics and Astronomy, University of Louisville, Louisville, KY, USA, <sup>15</sup>Department of Physics and Astronomy, University of Calgary, Calgary, AB, Canada, <sup>16</sup>National Radio Astronomy Observatory, Socorro, NM, USA, <sup>17</sup>INAF – Istituto di Radioastronomia, Bologna, Italy and <sup>18</sup>School of Natural Sciences, University of Tasmania, Hobart, TAS, Australia

### Abstract

We propose a new method for identifying active galactic nuclei (AGN) in low mass ( $M_* \leq 10^{10} M_\odot$ ) galaxies. This method relies on spectral energy distribution (SED) fitting to identify galaxies whose radio flux density has an excess over that expected from star formation alone. Combining data in the Galaxy and Mass Assembly (GAMA) G23 region from GAMA, Evolutionary Map of the Universe (EMU) early science observations, and Wide-field Infrared Survey Explorer (WISE), we compare this technique with a selection of different AGN diagnostics to explore the similarities and differences in AGN classification. We find that diagnostics based on optical and near-infrared criteria (the standard BPT diagram, the WISE colour criterion, and the mass-excitation, or MEx diagram) tend to favour detection of AGN in high mass, high luminosity systems, while the “PROSPECT” SED fitting tool can identify AGN efficiently in low mass systems. We investigate an explanation for this result in the context of proportionally lower mass black holes in lower mass galaxies compared to higher mass galaxies and differing proportions of emission from AGN and star formation dominating the light at optical and infrared wavelengths as a function of galaxy stellar mass. We conclude that SED-derived AGN classification is an efficient approach to identify low mass hosts with low radio luminosity AGN.

**Keywords:** Galaxies: active; galaxies: evolution; galaxies: nuclei; galaxies: spectral energy distribution; galaxies: radio AGN

(Received 23 October 2023; revised 29 January 2024; accepted 15 February 2024)

## 1. Introduction

One of the primary objectives in extragalactic astronomy is the identification and differentiation of active galactic nuclei (AGN) from star forming galaxies (SFGs). A significant number of galaxies host a supermassive black hole (SMBH), which acts as their central engine, and identification of these active nuclei is crucial in understanding their role in galaxy evolution (Magorrian et al. 1998; Kormendy & Kennicutt 2004; Kormendy & Ho 2013).

Luminous radio sources ( $L_{1.4\text{ GHz}} > 10^{24} \text{ W Hz}^{-1}$ ) are typically AGN, whereas only a few percent, typically around 10%, are

considered radio-loud (White et al. 2000). This fraction increases with optical and X-ray luminosities (Urry & Padovani 1995). In the mid-infrared (mid-IR) regime, strong ionisation lines and polycyclic aromatic hydrocarbon (PAH) features serve as excellent diagnostics for identifying AGN (Laurent et al. 2000). Luminous, centrally located compact X-ray and  $\gamma$ -ray sources are classified as AGN and do not require confirmation from other methods (Mushotzky 2004). However, depending on its accretion state or mode and duty cycle, the presence of an active nucleus might not always be simultaneously visible in all spectral regions, and consequently, not all diagnostic methods are necessarily consistent (Hine & Longair 1979; Hardcastle, Evans, & Croston 2007).

A significant fraction of radio sources do not show broad optical emission lines or even any emission lines at all (BL Lacertae objects; Padovani & Giommi 1995, low-excitation radio galaxies (LERGs); Best & Heckman 2012), adding complexity to the classification process. Infrared observations of both X-ray selected

**Corresponding author:** Jahang Prathap; Email: jahangprathap12@gmail.com

**Cite this article:** Prathap J, Hopkins AM, Robotham ASG, Bellstedt S, Afonso J, Ahmed UT, Bilicki M, Bremer MN, Brough S, Brown MJI, Gordon Y, Holwerda BW, Leahy D, López-Sánchez ÁR, Marvil JR, Mukherjee T, Prandoni I, Shabala SS, Vernstrom T and Zafar T. (2024) EMU/GAMA: A technique for detecting active galactic nuclei in low mass systems. *Publications of the Astronomical Society of Australia* 41, e016, 1–14. <https://doi.org/10.1017/pasa.2024.9>

and optically selected AGN samples reveal a wide range of IR to X-ray SEDs and significant variation in the IR colours of AGN (Mushotzky 2004; Padovani *et al.* 2017). Consequently, relying solely on IR colours can pose challenges in conducting an AGN survey. The IR spectral parameters do not exhibit a correlation with the optical spectral slope, IR luminosity, mid-IR-to-far-IR luminosity ratio, or the effects of inclination-dependent extinction in the context of a dusty torus model (Mushotzky 2004). Additionally, certain AGN classes display relatively low soft X-ray-to-optical ratios, indicating the existence of an IR-selected AGN group with broad optical emission lines that are considered X-ray “weak” (Wilkes *et al.* 2002). The key takeaway is that a multi-wavelength analysis is essential for accurately identifying galaxies hosting AGN.

### 1.1. Classifying AGN

Currently, a wide array of diagnostics is employed for AGN identification. A traditional optical diagnostic, the Baldwin-Phillips-Terlevich (BPT; Baldwin *et al.* 1981) diagram, uses the intensity ratios of [OIII]/H $\beta$  and [NII]/H $\alpha$  to classify galaxies as AGN or SFGs. It was later revised by Veilleux & Osterbrock (1987) to include [SII] and [OI] lines, Kewley *et al.* (2001) to provide a theoretical maximum for SF, and Kauffmann *et al.* (2003) for a semi-empirical demarcation between SFGs and AGN making the optical selection more comprehensive. For higher redshifts ( $z > 0.4$ ), where [NII] and H $\alpha$  lines may not be available, alternative diagnostic diagrams have been developed. The mass-excitation (MEx; Juneau *et al.* 2011) diagnostic diagram uses the [OIII]/H $\beta$  line ratio and stellar mass as a substitute for [NII]/H $\alpha$ , while the colour-excitation (CEX; Yan *et al.* 2011) diagnostic diagram employs the rest-frame  $U - B$  colour along with the [OIII]/H $\beta$  line ratio for classification. In addition to the BPT diagram, other diagrams use different line ratios to classify galaxies. The blue diagram (Lamareille *et al.* 2004) incorporates the [OIII]/H $\beta$  and [OII]/H $\beta$  line ratios, while the equivalent width (EW) of H $\alpha$  ( $W_{H\alpha}$ ) versus [NII]/H $\alpha$  (WHAN; Cid Fernandes *et al.* 2010, 2011) diagram uses the [NII]/H $\alpha$  line ratio combined with EW of H $\alpha$ . The Trouille-Barger-Tremonti (TBT; Trouille *et al.* 2011) diagram uses the rest-frame ( $g - z$ ) colour and the line ratio [NeIII]/[OII]. The OHNO diagram (Zeimann *et al.* 2015; Backhaus *et al.* 2022) uses the line ratios [OIII]/H $\beta$  and [NeIII]/[OII] for classification.

The mid-IR region can be used to produce AGN diagnostics that primarily uses colours and line-flux intensities originating from dust re-emission. Several methods (the Lacy wedge; Lacy *et al.* 2004, 2007; Sajina *et al.* 2005, the Stern wedge; Stern *et al.* 2005, the Infrared Array Camera (IRAC) power-law; Alonso-Herrero *et al.* 2006; Donley *et al.* 2007, 2012, and the KI & KIM methods; Messias *et al.* 2012, 2014) have been developed, each with its specific approach and advantages. To mitigate degeneracies, incorporating redshift information along with colour parameters has been suggested (Juneau *et al.* 2013). Wide-field Infrared Survey Explorer (WISE; Wright *et al.* 2010; Jarrett *et al.* 2011) photometry has also been used to derive criteria for distinguishing AGN from SFGs (Mateos *et al.* 2012; Stern *et al.* 2012; Assef *et al.* 2013, 2018). Mid-IR line ratios have been explored for AGN identification as well (Genzel *et al.* 1998; Sturm *et al.* 2002). No AGN selection is 100% complete and IR AGN selection is no exception. Several studies (e.g., Barmby *et al.* 2006; Gürkan *et al.* 2014; Messias *et al.* 2012) have highlighted the

instances of contamination caused by high- $z$  galaxies or inactive galaxies.

### 1.2. Radio AGN

One of the key properties of AGN is how their emission is distributed across the entire electromagnetic spectrum. The signature is not uniform throughout the spectrum as noted above, resulting in numerous “types” or classes of AGN depending on how they are diagnosed. In terms of radio properties, the AGN population can be broadly classified into radio-loud and radio-quiet depending on the radio power (e.g., radio-to-optical flux ratio; Kellermann *et al.* 1989, radio luminosity; Willott *et al.* 1999). Since the accretion rate of AGN can still vary within these classes, based on the spectral intensities, AGN are further classified into radiative-mode and jet-mode (Hine & Longair 1979; Hardcastle *et al.* 2007; Heckman & Best 2014). Radiative-mode AGN, where the energy is transferred radiatively from the accretion disc, consist of high-excitation radio galaxies (HERGs; Best & Heckman 2012; Pracy *et al.* 2016) or strong-line radio galaxies (SLRGs; Tadhunter 2016), Seyferts, and radio-quiet quasi-stellar objects (QSOs). Jet-mode AGN, whose primary energy output is in two-sided collimated jets, consist of LERGs (Best & Heckman 2012; Pracy *et al.* 2016), or weak-line radio galaxies (WLRGs; Tadhunter 2016) and low-ionisation nuclear emission line regions (LINERs) when they are AGN (Belfiore *et al.* 2016).

### 1.3. Spectral Energy Distributions and AGN

The SED of a galaxy serves as a comprehensive blueprint, reflecting the intrinsic characteristics of a galaxy, including its stellar and gas content, age and enrichment, dust content, and AGN activity. SED fitting is, hence, a powerful technique that allows for derivation of the physical properties of a galaxy by fitting a model to the observed data (e.g., Walcher *et al.* 2011; Conroy 2013). SED fitting has also been used as an AGN diagnostic tool (Thorne *et al.* 2022; Best *et al.* 2023; Yang *et al.* 2023).

There are many state-of-the-art SED fitting tools that are commonly used (see Pacifici *et al.* 2023 for a short description of some of these tools and their application to datasets). We choose PROSPECT (Robotham *et al.* 2020; Bellstedt *et al.* 2020a; Thorne *et al.* 2022) as a generative galaxy SED package to perform SED fitting. Our choice is motivated by the capability of PROSPECT to extend the SED to the radio regime. Also, the dust modelling used in PROSPECT uses the two-component form (see Section 3.4). Few SED fitting tools combine these two features. Here we explore the AGN diagnostic capabilities of the SED fitting technique through radio selection and find that this technique selects an AGN population residing in host galaxies of low stellar mass, which also show low radio luminosities.

The structure of the paper is as follows: We describe the photometric and spectroscopic data in Section 2. The diagnostic tools are discussed in Section 3, where we also introduce a binary classification scheme to identify which criteria classify each object as star forming or AGN. We compare different AGN diagnostic methods in Section 4. We demonstrate the capability of PROSPECT as a new tool for identifying AGN using SED fitting in low mass galaxies in Section 5. We summarise our main findings and conclude in Section 6. Throughout the paper we assume the Planck Collaboration *et al.* (2016) cosmology with  $H_0 = 67.8 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_M = 0.308$  and  $\Omega_\Lambda = 0.692$ . The PROSPECT tool adopts the Chabrier (2003) initial mass function (IMF).

## 2. Data

SED fitting requires panchromatic data. We fit SEDs to observed photometry spanning the far-ultraviolet (FUV) to the far-infrared (FIR) and predict the radio output. We then compare it with the observed radio emission. We use data from two main surveys: the Galaxy and Mass Assembly (GAMA; Driver 2009; Driver et al. 2011, 2022; Liske et al. 2015) survey (FUV-FIR bands) and early science data from the Evolutionary Map of the Universe (EMU; Norris et al. 2011, 2021; Gürkan et al. 2022) survey (radio, at 888 MHz). Additionally, we incorporate data from the AllWISE catalogue (Cutri et al. 2012, 2021), which is derived from WISE for IR photometry.

### 2.1. GAMA data

The GAMA Survey has emerged as a natural extension of legacy surveys such as the Sloan Digital Sky Survey (SDSS; York et al. 2000) and the Two Micron All-Sky Survey (2MASS; Skrutskie et al. 2006). This comprehensive survey covers five separate fields, with a limiting magnitude of  $r < 19.8$  mag ( $i < 19.2$  mag for the G23 field), collectively spanning 286 square degrees. Four of the GAMA fields (G09, G12, G15, and G23) are strategically located within the footprints of the European Southern Observatory (ESO) VST Kilo Degree Survey (KiDS; de Jong et al. 2015) and ESO Visible and Infrared Survey Telescope for Astronomy (VISTA) Kilo-degree Infrared Galaxy Public Survey (VIKING; Edge et al. 2013) surveys. These, along with data from Galaxy Evolution Explorer (GALEX; Martin et al. 2005), WISE, and Herschel (Pilbratt et al. 2010), ensure that the GAMA survey benefits from extensive multiwavelength coverage from FUV to FIR.

GAMA Data Release 4 (Driver et al. 2022) combines photometric data from these surveys together with the spectroscopy from the Anglo-Australian Telescope (AAT) AAOmega spectrograph. We are interested in the G23 field ( $338.1^\circ < \text{RA} < 351.9^\circ$ ,  $-35^\circ < \delta < -30^\circ$ ) as it has radio measurements available from the EMU early science observations (Gürkan et al. 2022).

The GAMA database is stored as tables in data management units (DMUs). The DR4 contains the `gkvScienceCatv02` table (Bellstedt et al. 2020b) in the `gkvInputCatv02` DMU, with photometry for 458 844 sources and the `gkvFarIRv03` table (Bellstedt et al. 2020b) in the `gkvFarIRv03` DMU with photometry for 458 783 sources. Spectra are taken from `GaussFitSimplev05` (Gordon et al. 2017, 49 623 sources) in the `SpecLineSFRv05` DMU and stellar mass estimates from `StellarMassesGKvV24` (Taylor et al. 2011, 67 006 sources) in the `StellarMassesv24` DMU. We use Topcat (Taylor 2005) to cross-match between these tables, forming the ‘GAMA’ sample of 47 774 sources with the spectrophotometric data.

### 2.2. EMU data

The Australian Square Kilometre Array Pathfinder (ASKAP; Johnston et al. 2007; McConnell et al. 2016; Hotan et al. 2021) is a precursor telescope for the Square Kilometre Array (SKA). ASKAP is designed to observe in the frequency range of 800–1 800 MHz with an instantaneous bandwidth of 288 MHz and offers rapid observations with good resolution and sensitivity.

EMU (Norris et al. 2011, 2021) is an ongoing wide-field radio continuum survey conducted using ASKAP. Its primary goal is to provide a deep radio continuum map of the entire southern sky, extending up to  $+30^\circ$  in declination. The survey achieves a

sensitivity of approximately  $20 \mu\text{Jy beam}^{-1}$  and a resolution of  $15''$ . As the survey progresses, it is anticipated that EMU will detect and catalogue approximately 40 million galaxies, including SFGs up to redshift  $z \sim 1$ , powerful starbursts reaching even greater redshifts, and AGN to the edge of the visible Universe.

The G23 field has been observed as part of the ASKAP commissioning (Leahy et al. 2019) and later as part of the EMU early science programme. In this study, we use the most recent ASKAP observation of the G23 field, carried out as part of EMU early science (Gürkan et al. 2022). These data cover an area of  $82.7 \text{ deg}^2$  and achieve a sensitivity of  $0.038 \text{ mJy beam}^{-1}$  at a frequency of 887.5 MHz. 55 247 sources were detected in this field. Of these, 39 812 have signal-to-noise (S/N) ratios  $\geq 5$ .

In our analysis, we focus on the common sources between GAMA and EMU early science radio data in G23. A cross-match between the two employing a matching radius of  $5''$  (Norris et al. 2021; Ahmed et al. 2023) results in 10 991 common sources with FUV-FIR photometry and 888 MHz radio measurements. Of these, there are 6 425 sources with spectroscopic measurements and stellar masses. We refer to this sample as ‘GAMA-EMU.’ We adopt the integrated radio flux densities from Gürkan et al. (2022) for our analysis throughout, to ensure we are using the total radio flux densities for each source.

### 2.3. WISE data

WISE (Wright et al. 2010; Jarrett et al. 2011) is a NASA medium-class explorer mission consisting of a 40 cm space-based infrared telescope that conducted a comprehensive digital imaging survey of the entire sky. WISE observations were performed in four mid-infrared bandpasses, namely 3.4, 4.6, 12, and 22 micrometres, denoted as  $W_1$ ,  $W_2$ ,  $W_3$ , and  $W_4$  with minimum  $5\sigma$  sensitivities of 0.08, 0.11, 0.8, and 6 mJy, respectively. During its mission, WISE produced and released a digital image atlas, providing coverage of the sky in four survey bands.

The AllWISE catalogue (Cutri et al. 2012, 2021) is an extensive database containing  $W_1$  through  $W_4$  magnitudes of 747 634 026 sources. Cross-matching the AllWISE catalogue with GAMA produced 44 244 sources within a matching radius of  $3''$ , which we refer to as ‘GAMA-WISE’ throughout the paper. The cross-match radius is motivated by the fact that the point spread function (PSF) fitting of the AllWISE pipeline can affect the resulting accuracy of the position particularly for resolved sources (Cluver et al. 2014). The latest data release of WISE (the CatWISE catalogue; Marocco et al. 2021) provides  $W_1$  and  $W_2$  magnitudes of more than a billion objects with improved photometry. The choice of AllWISE over CatWISE in this analysis is based on the diagnostic in Section 3.2, which is defined based on AllWISE photometry. Tests using the CatWISE measurements show no significant differences in our main results, and we adopt the AllWISE measurement for consistency throughout. In order to link this to the galaxies with radio measurements, we cross-match the GAMA-WISE sample with the GAMA-EMU sample. This resulted in 6 266 sources with FUV-FIR photometry, WISE magnitudes, and radio measurements, which we refer to as ‘GAMA-WISE-EMU.’

### 2.4. Final sample

We now have 6 266 galaxies with FUV-radio photometry, mid-IR magnitudes, relevant spectroscopic data and stellar mass estimates. We selectively remove galaxies under the following categories:

**Table 1.** The number of galaxies cross-matched between GAMA G23, WISE, and EMU.

Catalogue	Number of sources
GAMA	47 774
GAMA-WISE	44 244
GAMA-EMU	6 425
GAMA-WISE-EMU	6 266
GEM	2 956

- those with redshifts below the quality index (denoted as  $nQ$  in the GaussFitSimplev05 catalogue),  $nQ < 3$  (Gordon *et al.* 2017); 123 sources
- those with spectral lines in absorption, including  $H\alpha$  and  $H\beta$  EWs; 2 330 sources
- those with unphysical stellar masses; 2 sources
- those with a starmask (potential contamination of the photometry of a galaxy by a background star; Bellstedt *et al.* 2020b): 149 sources.

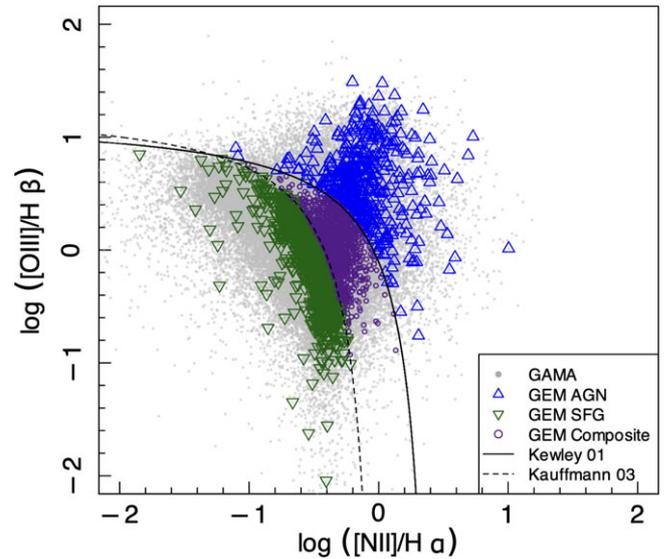
We are now left with 3 662 sources. We do not exclude galaxies further on the basis of the Balmer line S/N ratio so as not to further reduce the sample. If we impose a  $S/N > 2$  limit, we are left with only 2 253 galaxies, and our main results remain unchanged. The requirement to have  $H\alpha$  lines for the BPT classification limits our sample to  $z < 0.34$ .

In our analysis of the PROSPECT SED outputs described in detail below in Section 3.4, we remove the bad fits by using the log-posterior (LP) parameter, which serves as a goodness-of-fit metric akin to the  $\chi^2$  statistic. The LP values within our sample exhibit a wide range, spanning -15 to -4 000, with the more negative values standing out as outliers. Following a careful examination, we set a threshold of  $LP \geq -50$ . We are left with a sample of 2 958 galaxies. To further validate the quality of this sample, each of these 2 958 galaxies was visually inspected, resulting in the removal of two bad fits. This gives our final sample of 2 956 galaxies, which we call ‘GEM’ (for the final robustly defined GAMA-WISE-EMU sample) throughout the paper. A summary of the catalogues is shown in Table 1.

### 3. AGN diagnostics

We introduce a new AGN diagnostic based on the PROSPECT SED fitting tool. In order to understand how this new diagnostic relates to existing well established methods, we use a selection of these here. We choose to compare against the BPT diagnostic (Baldwin *et al.* 1981), the WISE  $W_1 - W_2$  criterion (Assef *et al.* 2018), and the mass-excitation (MEx) diagram (Juneau *et al.* 2011). There is an extensive array of AGN diagnostics in the literature, but of necessity, we limit our choice here to a selection of the more common in order to keep the analysis tractable. The GEM sample is selected in such a way that all galaxies have the relevant parameters to be classified by this suite of diagnostics.

Balmer lines are affected by stellar absorption. To account for this, we apply a simple constant EW correction of 1.3 Å (Hopkins *et al.* 2003; Gunawardhana *et al.* 2011). It is important to note that the spectra in the GAMA dataset that are drawn from the Two-degree Field Galaxy Redshift Survey (2dFGRS) survey are



**Figure 1.** The BPT diagram: an emission line diagnostic diagram which uses the ratios  $[OIII]/H\beta$  and  $[NII]/H\alpha$  to delineate SFGs from AGN. The diagram shows the data from the full survey (G09, G12, G15, and G23) as GAMA BPT SFGs and AGN (grey dots), GEM BPT SFGs (green triangles), and GEM BPT AGN (blue triangles). Galaxies below the theoretical Kewley line (solid line) and above the empirical Kauffmann line (dashed line) form a sequence of composite galaxies (purple dots). These galaxies are included among the BPT AGN in our analysis.

not flux calibrated. For diagnostic diagrams, however, since we use ratios of lines close in wavelength, the lack of flux calibration is not an issue. Whenever Balmer line luminosities are used, we address both the stellar absorption and the flux calibration issues. To achieve this, we use the measured EWs and the r-band absolute magnitude to derive the respective luminosities, following the method outlined in Hopkins *et al.* (2003).

#### 3.1. The BPT diagram

In the seminal work by Searle (1971), it was demonstrated that emission line spectra from different types of HII regions can be classified based on line-intensity ratios, such as  $[OIII]\lambda 5007/H\beta$  (excitation) or  $[OII]\lambda 3727/[OIII]\lambda 5007$ . However, relying solely on the excitation threshold, though powerful for identifying Seyfert 2 galaxies, has some drawbacks. It is prone to contamination by sufficiently ionised starbursts and often excludes many LINERs (Baldwin *et al.* 1981; Osterbrock & De Robertis 1985). It was also found that the strength of low-ionisation lines such as  $[NII]$ ,  $[SII]$ , and  $[OI]$  can be used to distinguish HII regions from Seyfert 2s (Baldwin *et al.* 1981; Osterbrock & De Robertis 1985).

To address these limitations, Baldwin *et al.* (1981) extended the one-parameter model to a more comprehensive two dimensional classifier of HII regions resulting from power-law photoionisation and shock wave heating. Further studies on a more accurate classification of HII regions, including a broader range of sources like Seyferts, LINERs, and narrow-line radio galaxies (NLRGs) followed.

Notably, Osterbrock & De Robertis (1985) and Veilleux & Osterbrock (1987) provided a semi-empirical boundary line that effectively separates starbursts from Seyfert galaxies. In our study, we employ the BPT diagram with the line ratios  $[OIII]/H\beta$  versus  $[NII]/H\alpha$ , and we use the theoretical boundary line proposed by Kewley *et al.* (2001) and the empirical line proposed by Kauffmann

et al. (2003) to distinguish between AGN and SFGs. Fig. 1 illustrates the BPT classifier showing the GAMA data and our GEM sample. The data points in blue and green represent the GEM AGN and SFGs, respectively, overlaid on the GAMA sample shown in grey. Following the prescription in Kauffmann et al. (2003), we select all those galaxies above the dashed line as AGN. Composite galaxies, shown here in purple dots, are incorporated into our AGN classification for all our subsequent analyses unless otherwise noted in order that we are as complete as possible in identifying contributions from an AGN component.

### 3.2. The WISE criterion

AGN derive their power from the accretion disc formed by matter falling into the SMBH. The dusty torus surrounding most of the AGN absorbs a part of this power and re-emits this in the IR. This signature is evident across the IR and contributes measurably to the mid-IR spectrum. This process gives AGN characteristic red mid-IR colours. As a result, mid-IR colour criteria are a good way to select AGN (Stern et al. 2012). Consequently, WISE data are now extensively used for AGN selection.

Stern et al. (2012) put forward a simple magnitude independent colour criterion,  $W_1 - W_2 \geq 0.8$ , providing a reliable AGN sample for  $W_2 < 15.05$  mag. As an extension to fainter magnitudes, Assef et al. (2013) developed a new mid-IR magnitude-dependent selection criterion. There are four flavours of the criterion depending on the level of reliability and completeness of AGN selection for a limiting magnitude,  $W_2 = 17.11$  mag. Here, we adopt the magnitude-dependent AGN selection criterion described in Assef et al. (2018) developed for AllWISE data:

$$W_1 - W_2 > \begin{cases} \alpha_R \exp^{\beta_R(W_2 - \gamma_R)^2} & , W_2 > \gamma_R \\ \alpha_R & , W_2 \leq \gamma_R, \end{cases}$$

optimised for reliability, where the coefficients,  $(\alpha_R, \beta_R, \gamma_R) = (0.650, 0.153, 13.86)$ , correspond to 90% reliable AGN selection. Any other choice of WISE colour diagnostic, for instance, the  $W_1 - W_2$  classifier (Stern et al. 2012), does not change our main conclusions. The distribution of GEM and GAMA WISE galaxies in the  $(W_1 - W_2)$  vs  $W_2$  parameter space and the selection criterion is shown in Fig. 2.

SED fitting (in principle) should recover each IR-dominated AGN because the SED fits are equally sensitive to the same flux ratios that lead to an IR colour excess. In our analysis, however, as detailed below (Section 3.4), we do not use PROSPECT’s AGN mode in the SED fits.

### 3.3. The MEx diagram

Traditional nebular-line based diagnostic diagrams suffer from two main issues: the absence of near-infrared (NIR) lines at higher redshifts and the non-zero probability of galaxies near dividing lines being in either of the classes. Juneau et al. (2011) introduced the mass-excitation (MEx) diagnostic diagram, a novel probabilistic approach splitting galaxies into sub-categories more confidently than alternative diagnostic diagrams developed to identify AGN at similar redshifts.

The line ratios used in the BPT diagram are a combination that probes two important aspects of galaxies: the hardness of their ionising radiation and their gas-phase metallicity. However, at higher redshifts, the lines [NII] and  $H\alpha$  become redshifted out of the optical window, excluding their use unless NIR spectra are available.

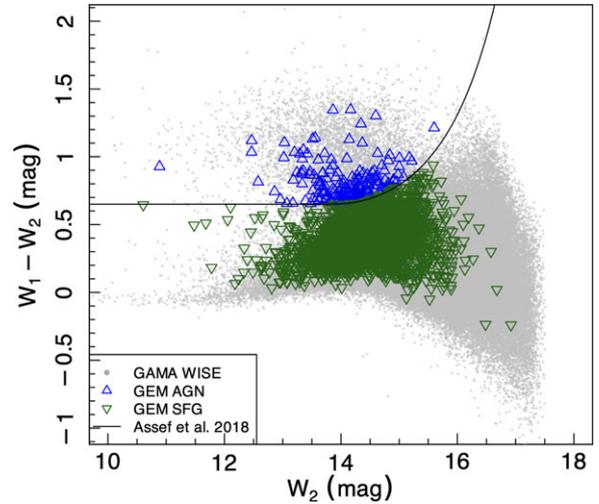


Figure 2. WISE colour diagram. The WISE AGN selection criterion described in Section 3.2, is shown as the solid line. The diagram shows the data from the full survey (G09, G12, G15, and G23) as GAMA WISE SFGs and AGN (grey dots), GEM WISE SFGs (green triangles), and GEM WISE AGN (blue triangles).

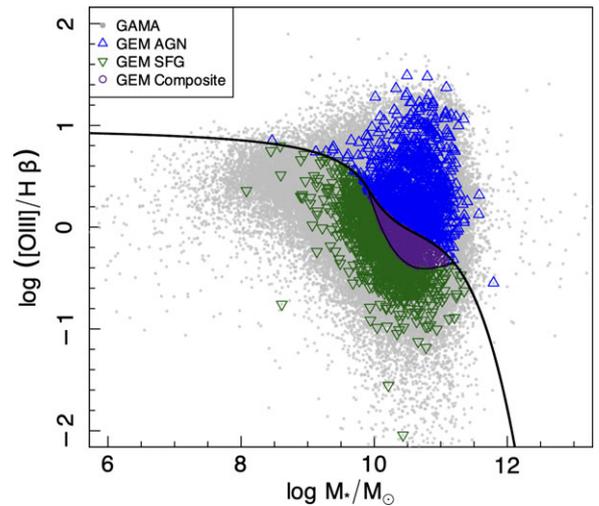


Figure 3. The mass-excitation (MEx) diagnostic. This diagnostic compares the excitation ( $[OIII]/H\beta$ ) and stellar mass. The diagram shows the data from the full survey (G09, G12, G15, and G23) as GAMA MEx SFGs and AGN (grey dots), GEM MEx SFG (green triangles), GEM MEx composite (purple dots), and GEM MEx AGN (blue triangles). The regions occupied by SFGs, composites, and AGN in the MEx parameter space are analogous to those in the BPT parameter space. The MEx composites are included among MEx AGN in our analysis.

We can leverage other empirical relations and properties of galaxies to mimic this approach to AGN identification. The empirical mass-metallicity relation (Tremonti et al. 2004) of normal galaxies suggests that mass is an indicator of metallicity ( $[NII]/H\alpha$ ). The value of the ratio  $[NII]/H\alpha$  is a direct measure of the energy output, which is in turn suggestive of AGN activity (Kauffmann et al. 2003; Kewley et al. 2006). Given these relationships, stellar mass can serve as an excellent proxy for the  $[NII]/H\alpha$  ratio at both lower and higher redshifts when it is not available or is challenging to measure accurately.

The MEx diagram classifies galaxies into three groups: AGN, composite, and SFG. Fig. 3 shows the GAMA and GEM AGN, SFGs and composites identified using the MEx diagram. As with

**Table 2.** The free parameters of the PROSPECT fitting code and their characteristics. Different parameters corresponding to modelling the SFH, dust absorption and re-emission, and metallicity are shown in the first column. A short description of each parameter is given in column 2, and the scale (log or linear) adopted while fitting in column 3. Columns 4 and 5 give the units of the free parameters and the range that was chosen to constrain the fitting, respectively. Column 6 shows the prior function used, applied only for the alpha parameters. PROSPECT is an interface to the simple stellar population (SSP) models, dust attenuation and re-emission libraries, which then uses the Highlander package (a combination of parameter optimisation and Bayesian inferential techniques) to optimise the free parameters and generate the SED corresponding to the observations.

Parameter	Description	Scale	Units	Range	Prior
mSFR	peak SFR	Log	$M_{\odot}/\text{yr}$	[−3,6]	–
mpeak	lookback time when the peak SFR occurred	Linear	Gyr	[−2,13.38]	–
mperiod	width of the SFH	Log	Gyr	[0.3,3]	–
mskew	skewness of the SFH	Linear	–	[−0.5,3]	–
Zfinal	final gas-phase metallicity	Log	–	[−4,−1]	–
alpha_SF_birth	power-law of the radiation heating the birth cloud	Linear	–	[0.01,4]	$\exp\left(\frac{-1}{2}\left(\frac{\alpha_{\text{birth}} - 2}{1}\right)^2\right)$
alpha_SF_screen	power-law of the radiation heating the ISM	Linear	–	[0.01,4]	$\exp\left(\frac{-1}{2}\left(\frac{\alpha_{\text{screen}} - 2}{1}\right)^2\right)$
tau_birth	optical depth of the birth cloud	Log	–	[−2.5,1]	–
tau_screen	optical depth of the ISM screen	Log	–	[−5,1]	–

our approach to composite galaxies in the BPT classifier, here we likewise include the MEx composites among AGN.

### 3.4. The PROSPECT selection

PROSPECT (Robotham *et al.* 2020; Bellstedt *et al.* 2020a; Thorne *et al.* 2022) is a generative galaxy SED fitting package. It is equipped with the stellar population libraries BC03 (Bruzual & Charlot 2003) and EMILES (Vazdekis *et al.* 2016), both using a fixed Chabrier IMF (Chabrier 2003). The package incorporates the Charlot & Fall (2000) dust attenuation model, as well as the dust re-emission library of FIR templates presented by Dale *et al.* (2014).

Initially, PROSPECT allowed fitting of SEDs only extending to the FIR region. In a recent extension by Thorne *et al.* (2023), the package now supports fitting up to radio wavelengths. The radio flux densities in these models are estimated from the FIR fluxes through the link between FIR-derived star formation rates (SFRs) and thermal and non-thermal radio emission. They adopt power-laws with  $\alpha = -0.1$  and  $\alpha = -0.7$  (where  $S_{\nu} \propto \nu^{\alpha}$ ) for these free-free and synchrotron components respectively (Thorne *et al.* 2023; Marvil *et al.* 2015). This forward-modelled radio flux solely originates from star formation processes. The simulated radio flux is then compared with the EMU radio data to distinguish between AGN and SFGs. Any excess radio emission above a defined threshold can be considered as an indication of AGN activity.

To model the star formation history (SFH) we use the `massfunc_snorm_trunc` parameterisation, where `mpeak`, `mSFR`, `mperiod`, and `mskew` are set as free parameters. For the metallicity history, we use the `Zfunc_massmap_lin` keeping `Z_final` as a free parameter. These parameters play crucial roles in determining the stellar mass and SFR of the galaxy at different epochs, taking into account any variations in the SFH and metallicity over time.

For modelling dust, we include four additional free parameters: `alpha_SF_birth`, `alpha_SF_screen`, `tau_birth`, and `tau_screen`. The `alpha_SF_birth` and `alpha_SF_screen` parameters represent the intensity of the dust radiation field for the birth and ISM clouds, respectively, whereas `tau_birth` and `tau_screen` parameters are measures of dust opacity for the same. To facilitate the fitting process, a prior function is included for the alpha parameters (Bellstedt *et al.* 2020a). This prior

function helps guide the fitting process and enables PROSPECT to better model the SED and derive more accurate physical properties faster. Table 2 provides a summary of the different parameters used in our SED fitting process and their characteristics.

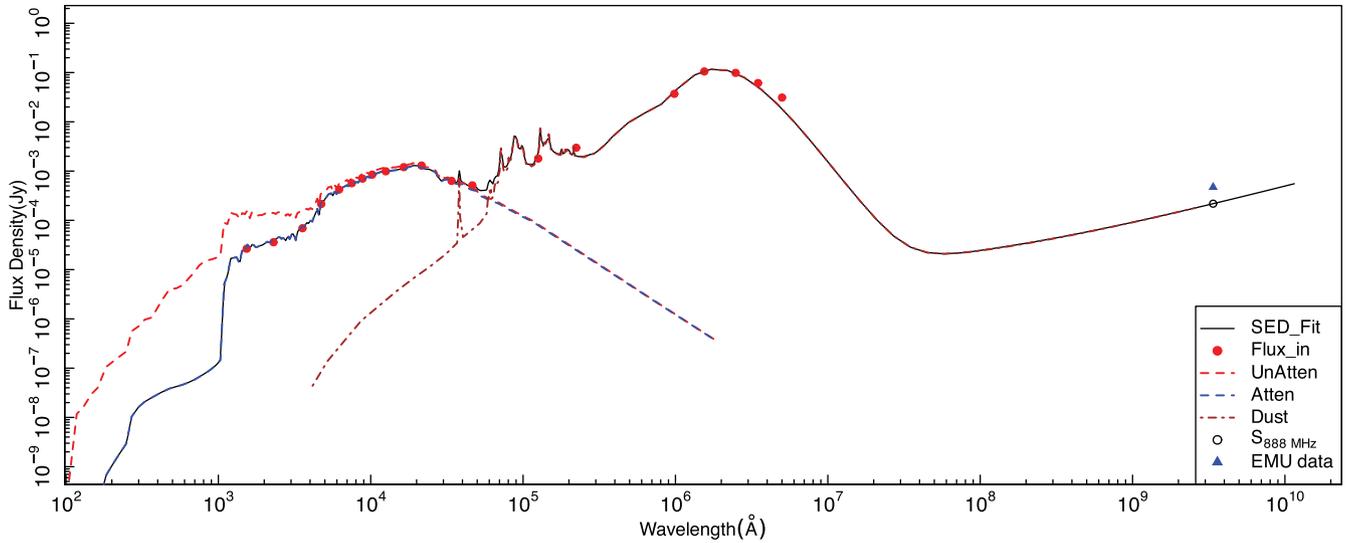
The FUV-FIR photometry gives the flux densities in twenty different bands. We note here that some objects do not have photometry in certain bands. In such cases, the SED was fitted using only the available photometry. We use floor values to supplement the uncertainties associated with each band to ensure reliable fits even when reported uncertainties are too optimistic (FUV = 0.2, NUV = 0.3, u = 0.2, g = 0.2, r = 0.2, i = 0.2, Z = 0.2, Y = 0.2, J = 0.2, H = 0.2, Ks = 0.2, W<sub>1</sub> = 0.05, W<sub>2</sub> = 0.14, W<sub>3</sub> = 0.145, W<sub>4</sub> = 0.16, P100 = 0.001, P160 = 0.001, S250 = 0.001, S350 = 0.001, and S500 = 0.001) and the final fitting uncertainties are calculated as described in Bellstedt *et al.* (2020a).

In the fitting process, we use the Highlander<sup>a</sup> package to model the SED of a galaxy. We implement this in a Bayesian manner using the Componentwise Hit and Run Metropolis (CHARM) algorithm, which is a Markov Chain Monte Carlo (MCMC) method provided by the LaplacesDemon (LD) package in R (Statisticat & LLC. 2021). For each galaxy, MCMC is run in two iterations, with each iteration optimising the parameters in 1 000 steps and then feeding the optimised parameters into LD (500 steps in the first iteration and 1 000 steps in the second iteration).

To estimate the radio emission from the SED, we enable the `addradio_SF=TRUE` option. The underlying concept is to model the SED of the galaxy assuming no AGN contribution, which results in an estimate of the radio emission purely from star formation processes. An example SED of a galaxy with the contributing factors (stellar emission, dust absorption, and re-emission) is shown in Fig. 4. To distinguish between AGN and SFGs, we employ a threshold for the ratio of the observed EMU radio data to the radio flux estimate obtained from the SED fitting without AGN.

For this analysis, and to be conservative, we choose a threshold of 3. This choice is somewhat arbitrary and could be lower or higher depending on how the diagnostic is to be used. We did

<sup>a</sup><https://github.com/asgr/Highlander>.

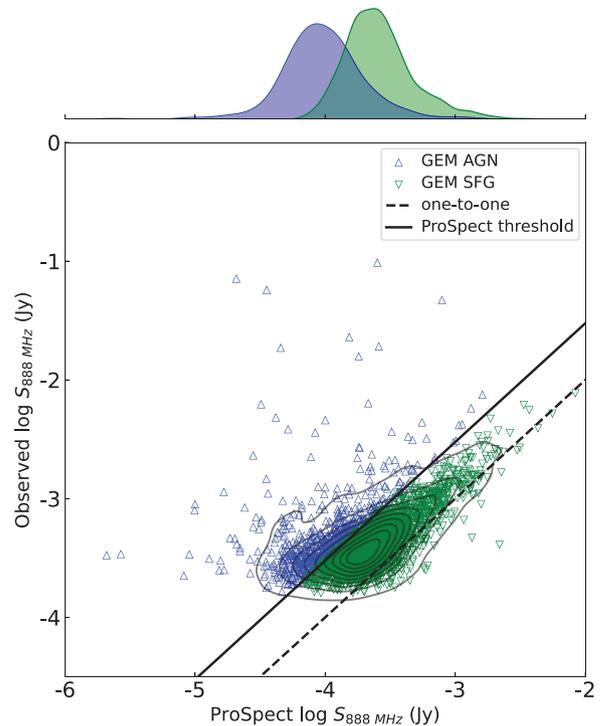


**Figure 4.** An example SED fit using PROSPECT (black solid line). The red data are the photometry from the GAMA catalogue. The EMU observation is shown as the blue triangle and the PROSPECT predicted radio flux is shown as the black open circle. The unattenuated stellar emission (red dashed line) is subjected to dust absorption and the resulting attenuated stellar emission is shown as the blue dashed line. The corresponding energy difference is re-emitted in the infrared spectrum (brown dot-dashed line) taking energy balance into account. The PROSPECT AGN component is not used here, thus the radio prediction arises from star formation processes alone. Based on our criterion, this example is classified as a PROSPECT SFG.

check the impact on our AGN selection of varying the threshold value, and, as expected, lowering the threshold increases the sample of potential PROSPECT AGN, but also increases the number classified as SF galaxies by the other classifiers, implying greater contamination. For the current process, we are aiming to demonstrate that this is a reliable AGN diagnostic, and accordingly, we choose a relatively high threshold so as not to bias our result through contamination by star formation from fits with large uncertainties.

The FIR-radio correlation (Condon 1992) exhibits a mass dependency (Gürkan et al. 2018; Smith et al. 2021; Delvecchio et al. 2021). Low mass galaxies exhibit a higher value of  $q_{IR}$  ( $L_{IR}/L_{1.4GHz}$ ; Delvecchio et al. 2021) than the high mass galaxies (see figure 14 of Delvecchio et al. 2021). This can be attributed to the escape of cosmic ray electrons (CRE) from low mass galaxies before they radiate away their energies. The consequence of this is a reduction in the radio luminosity of the low mass galaxies from that predicted by the linear FIR-radio correlation. This may result in some low mass galaxies which would be AGN but not be selected by the PROSPECT method. The FIR-radio correlation is tight with a scatter of only 0.26 (e.g., Yun, Reddy, & Condon 2001), corresponding to a factor of 1.9. With the PROSPECT selection threshold of 3, this would not result in any SFG being misclassified as AGN. If anything, this conservative selection might exclude some AGN.

Fig. 5 presents a comparison between the observed and the predicted radio fluxes based on SED fitting. This is effectively an AGN selection criterion based on a radio emission threshold. Galaxies that fall above this threshold are identified as AGN, while those below it are classified as SFGs. An apparent asymmetry is evident in the distribution of AGN and SFGs around the one-to-one line. This is associated with our sample selection, as we require the objects to be radio-detected. SFGs that would otherwise contribute to a symmetrical distribution may fall below the radio sensitivity limit of the survey and, hence, are not represented here.



**Figure 5.** PROSPECT AGN selection. The new SED-based AGN diagnostic scheme showing the relation between observed EMU radio flux and the predicted radio flux from SF processes, both at 888 MHz. The blue triangles are the GEM PROSPECT AGN and the green triangles are the GEM PROSPECT SFGs. The demarcation (solid line) between the two corresponds to the observed radio flux being at least 3 times the predicted radio flux from SF. The dashed line shows the one-to-one relation.

Given the challenge of tracking which classification each galaxy has from each diagnostic, we introduce a binary classification scheme. Galaxies are labelled 0 and 1, meaning they are SFG or

**Table 3.** The binary classification scheme.

Identifier	BPT	WISE	PROSPECT	MEx
0000	SFG	SFG	SFG	SFG
0001	SFG	SFG	SFG	AGN
0010	SFG	SFG	AGN	SFG
0011	SFG	SFG	AGN	AGN
0100	SFG	AGN	SFG	SFG
0101	SFG	AGN	SFG	AGN
0110	SFG	AGN	AGN	SFG
0111	SFG	AGN	AGN	AGN
1000	AGN	SFG	SFG	SFG
1001	AGN	SFG	SFG	AGN
1010	AGN	SFG	AGN	SFG
1011	AGN	SFG	AGN	AGN
1100	AGN	AGN	SFG	SFG
1101	AGN	AGN	SFG	AGN
1110	AGN	AGN	AGN	SFG
1111	AGN	AGN	AGN	AGN

AGN dominated, respectively, in a particular diagnostic and the scheme is detailed in Table 3.

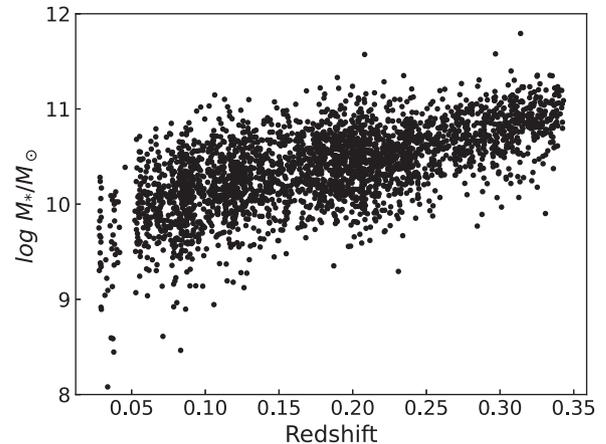
#### 4. Results

As PROSPECT identification is based on radio excess, we can have a high level of confidence in the AGN detected by this method. In this section, we explore and compare various physical properties of the AGN identified by the classifiers. We investigate whether PROSPECT provides unique information beyond what other methods offer. The analyses in the upcoming subsections use the binary identifier scheme introduced in Section 3. Table 3 is a guide to the identification scheme.

##### 4.1. Luminosities: $H\alpha$ – radio

Fig. 6 shows how the host stellar masses of our sample are distributed over redshift. The deficit of low mass hosts at higher redshifts is a consequence of the flux limit of the survey whereas that of massive galaxies in the local universe is due to the small volume sampled in that redshift range. Fig. 7 shows the relation between  $H\alpha$  and 888 MHz luminosities in four different stellar mass bins. We choose to represent the data by looking at the relationship between these luminosities, as both luminosities are sensitive to SFR. Excesses above a one-to-one correlation in this diagram would imply an excess radio emission that could be attributed to a radio AGN, although low (radio) luminosity AGN will not necessarily show such an excess.

We chose the division in mass to have a roughly similar number of objects in the middle two bins, which contain the most galaxies, while still separately sampling the highest and lowest mass ranges. Table 4 presents the number of galaxies in these mass bins arising from the different combinations of AGN indicators. The colours chosen in the key have been carefully selected to help guide the eye in interpreting this very busy figure (and those to follow). In broad terms, and neglecting the MEx classifier (for simplicity in the first instance), the green colours are galaxies classified as star forming, blue are those classified as AGN by PROSPECT (or PROSPECT and



**Figure 6.** Variation of stellar masses of the GEM galaxies with redshift. The requirement of reliable spectral lines for the emission line diagnostics, especially the BPT diagram, confines the redshift of the GEM sample to a maximum of  $z = 0.34$ . At higher redshifts, the [NII] and  $H\alpha$  lines are redshifted out of the optical window. The typical trend seen in flux-limited samples is visible, with an apparent trend of increasing mass with redshift. This is a consequence of the flux limit excluding low mass galaxies at higher redshift, while at low redshift, the sampled volume is not large enough to capture rare massive galaxies.

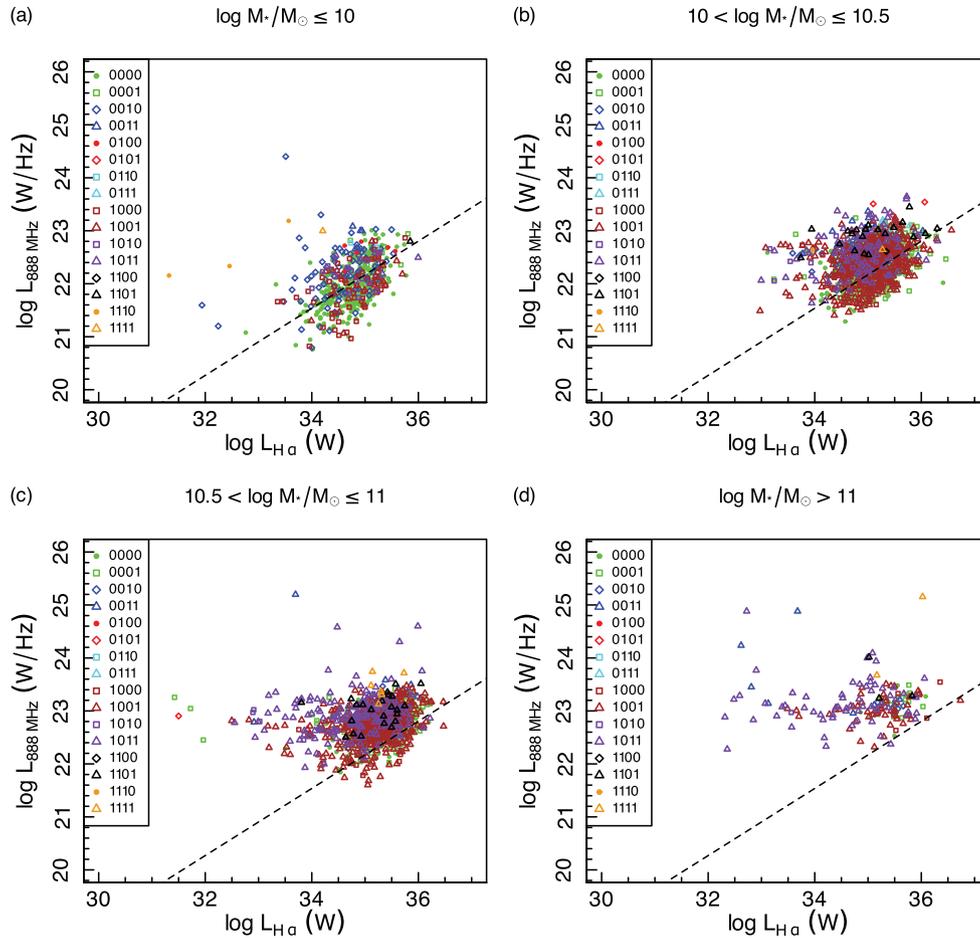
MEx), and brown are those classified as AGN by BPT (or BPT and MEx). The red symbols (WISE, or WISE and MEx) are very few in number, and the remaining colours, purple, black, and yellow, refer to galaxies classified as AGN by multiple indicators.

Most AGN, especially those identified by BPT and MEx, populate the two middle mass bins, and it is noticeable that very few WISE and relatively less MEx-identified AGN exist in the lowest mass bin. The lowest masses, however, do contain a significant number of AGN, and these are primarily identified through the PROSPECT and BPT classifiers. This is reflected in the numbers in Table 4. From this initial interpretation, it appears that PROSPECT is effective at identifying AGN in galaxies of low mass, some of which demonstrate low radio and  $H\alpha$  luminosities. The distribution of AGN by indicator as a function of host galaxy stellar mass and as a function of radio luminosity is shown in Fig. 8, which reinforces this result.

In the lowest mass bin (Fig. 7), there is a trend between the radio and  $H\alpha$  luminosities for most galaxies (indicated by the dashed line). Those AGN identified by the BPT follow this trend, but those identified by PROSPECT fall above or on the upper edge of this trend, as expected from radio excess systems. In the higher mass bins, the populations move to higher luminosities and have an increased scatter of radio excess objects. This is a consequence of the variation of host galaxy mass with redshift.

It is worth noting that even in the highest mass bin ( $\log M_*/M_\odot > 11$ ) we still find galaxies classified as star forming only by all four classifiers. We have visually inspected the spectra of all 13 of these potential SFGs and confirmed that 6 of them are genuine star formers with no apparent AGN signature in the indicators we have utilised. The rest were assigned manually to the BPT AGN class. These galaxies lie in the upper range of our redshift space, with  $z > 0.1$  for all massive star formers.

A close look at Figs. 7, 8, and Table 4 clearly demonstrates that the low mass end of the AGN population in our GEM sample is roughly equally populated by BPT and PROSPECT AGN. With the mean mass of the PROSPECT AGN population being  $10^{9.9} M_\odot$ ,



**Figure 7.** 888 MHz luminosity as a function of H $\alpha$  luminosity for four different stellar mass bins. The colours of the data points are selected to aid in distinguishing the 16 classes. The emission line detected AGN are dominant in the three highest mass bins, whereas a large number of AGN in the lowest mass bin are identified by PROSPECT. The dashed line is a linear fit to the population in the lowest mass bin and is shown for reference in other mass bins. This emphasises the upward shift of galaxies in luminosity as mass increases, along with an increased scatter in the radio excesses.

PROSPECT is efficient at detecting low mass AGN systems which would not have been identified with narrow line and mid-IR diagnostics alone. Though BPT, in combination with other methods, detects a fairly large number of AGN in the lowest mass bin, the PROSPECT method detects the largest number of AGN here.

#### 4.2. Radio luminosity fractional AGN contribution

We define the parameter  $J$  as the ratio of excess radio luminosity compared to that predicted by PROSPECT ( $L_{\text{Obs}} - L_{\text{PROSPECT}}$ ) to the observed radio luminosity:

$$J = \frac{L_{\text{Obs}} - L_{\text{PROSPECT}}}{L_{\text{Obs}}}$$

The quantity  $J$  is the ratio of the AGN’s contribution to the total radio luminosity of the galaxy. It answers the question, “What fraction of a galaxy’s radio emission is from AGN as opposed to star formation?” Since there are galaxies that lie below the one-to-one line in Fig. 5, there will be systems with a negative value of  $J$ . These by definition are classified as star forming by the PROSPECT diagnostic, but could still be classified as AGN by other methods.

We show how  $J$  varies in Fig. 9, which presents radio SFR as a function of H $\alpha$  SFR, colour-coded by  $J$ . The radio SFR here is

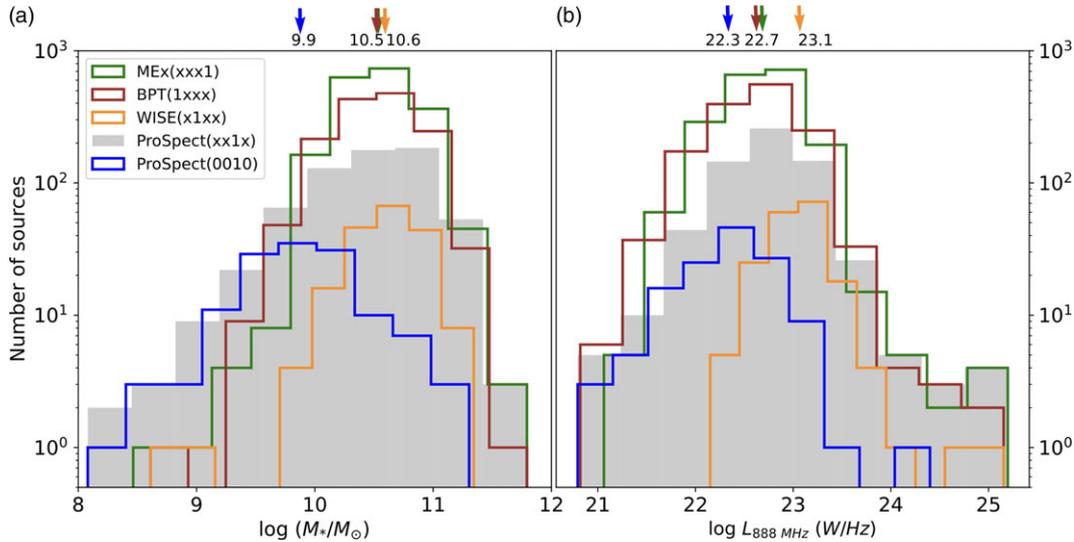
calculated after converting our 888 MHz luminosities to 1.4 GHz, assuming a radio spectral index of  $\alpha = -0.7$ , and the SFR calibration of Hopkins et al. (2003). This confirms that sources with an excess of radio emission over the otherwise broad correlation have higher values of  $J$ , as expected.

Fig. 10 presents this parameter as a function of mass and radio luminosity. Here we only show positive values of  $J$ . The PROSPECT detected sources, by definition, lie at  $J > 0.67$  because of our threshold value of 3 (defined in Section 3.4), which we call the “PROSPECT threshold”. The transition to different AGN diagnostics being dominant above  $\log (M_{\bullet}/M_{\odot}) = 10$ , reinforces the preferential selection of AGN from low mass hosts by PROSPECT.

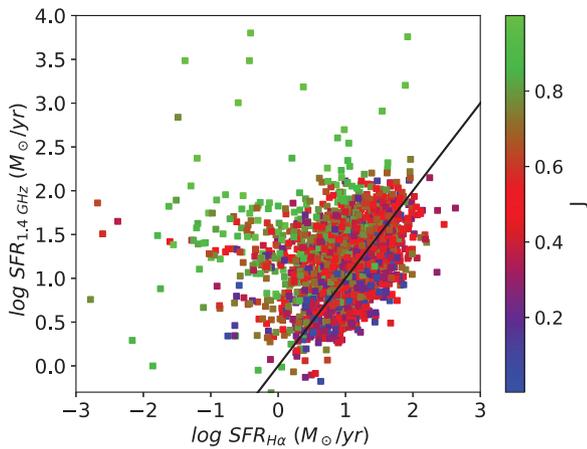
Above the PROSPECT threshold, AGN show a clear demarcation of diagnostics in the region of parameter space they occupy in Fig. 10. PROSPECT AGN, hosted by low mass galaxies, extend towards the low radio luminosity end to a degree not seen with other indicators. Fully one-fifth (22%) of the PROSPECT AGN lie at  $L_{888 \text{ MHz}} < 10^{22} \text{ W Hz}^{-1}$ . This compares to a fraction of one-tenth (10%) for the BPT AGN. The population exhibits a mixing at mid-radio luminosity giving way to AGN detected by multiple diagnostics at high radio luminosity end. Below the PROSPECT threshold, AGN lie predominantly at higher radio luminosities,

**Table 4.** The numbers of AGN in different mass bins identified by different classifiers. The first column shows the mass bins and the total AGN numbers in each mass bin are shown in the second column. Here, for clarity in the display, we neglect the WISE AGN due to their small numbers. Columns three to five list the AGN detected by BPT, MEX, and PROSPECT regardless of the identification by other classifiers. Columns six to eight show the AGN detected by BPT, MEX, and PROSPECT alone.

mass_bin	AGN/Total	BPT AGN (1xxx)	MEx AGN (xxx1)	PROSPECT AGN (xx1x)	BPT alone (1000)	MEx alone (0001)	PROSPECT alone (0010)
$\log(M_*/M_\odot) \leq 10$	206/481	95	57	113	51	17	81
$10 < \log(M_*/M_\odot) \leq 10.5$	970/1 269	583	857	217	56	261	38
$10.5 < \log(M_*/M_\odot) \leq 11$	948/1 051	659	903	240	21	201	11
$\log(M_*/M_\odot) > 11$	148/155	122	137	77	6	11	3



**Figure 8.** Panel (a): Distribution of host galaxy stellar mass for AGN identified by different classifiers. AGN detected by narrow-line diagnostics are shown in green and brown, and WISE detected AGN in orange. The grey filled distribution shows the AGN identified by multiple classifiers including PROSPECT and the PROSPECT alone AGN are represented by the blue histogram. The arrows coloured according to the histograms show the mean (the medians are almost identical to the mean) of these distributions in the respective colours (BPT and MEX have the same mean values). Panel (b): Distribution of host galaxy radio luminosity for AGN identified by different classifiers. The colours of the distributions and the arrows are as in panel (a). The distribution of PROSPECT AGN clearly favours the low mass, low radio luminosity systems, while the distributions of AGN detected by multiple classifiers favour higher host galaxy mass and radio luminosities.



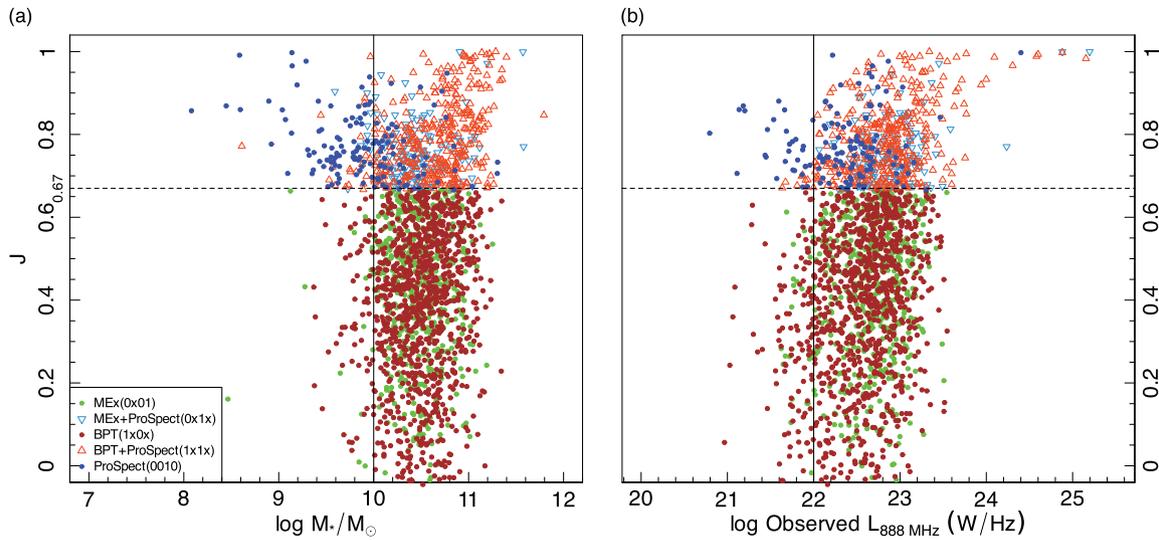
**Figure 9.** 1.4 GHz radio SFR as a function of H $\alpha$  SFR. The 1.4 GHz radio SFRs are calculated after converting the 888 MHz radio luminosities to 1.4 GHz. The solid black line represents the one-to-one correlation between both SFRs. The points are colour-coded according to  $J$ . The sources with higher values of  $J$  ( $J \geq 0.67$ ) exhibit excess radio emission.

with very few extending to the faintest levels. This is true also for those non-PROSPECT AGN with negative  $J$  values.

## 5. Discussion

### 5.1. Comparison with a low mass AGN sample

Salehirad *et al.* (2022) identified 388 low mass AGN hosts in the GAMA DR4 equatorial and G23 regions consistent with our definition of low mass ( $\log M_*/M_\odot \leq 10$ ). Of this sample, 53 lie in G23. Their search was primarily using narrow-line diagnostic techniques, including the BPT diagram. A cross-match between the 53 low mass AGN hosts and the EMU early science data with a 5" radius resulted in 11 common sources. Thus the radio selection is the primary reason for the large difference in their sample and ours. A cross-match between these 11 low mass AGN hosts and our GEM sample resulted in 9 common sources whose details are shown in Table 5. One of these is classified by all four classifiers as an SFG in our sample. The two missing sources in the GEM sample are removed by the LP selection described in Section 2.4.



**Figure 10.** The radio luminosity fractional AGN contribution ( $J$ ) as a function of (a) stellar mass and (b) radio luminosity. For conciseness in the key, we use an “x” to avoid repetition of both 1 and 0 where classifications are consolidated. So, for example, 0x01 means both 0001 and 0101. That is, these are objects classified as AGN by MEx and SF by PROSPECT and BPT, independent of the WISE classification. To be clear we spell out here what each colour represents in terms of explicitly which binary classifications are jointly included. The colour scheme is identical for both panels. Green (MEx AGN: 0001, 0101); Cyan (MEx+PROSPECT AGN: 0011, 0110, 0111); Brown (BPT AGN: 1000, 1001, 1100, 1101); Orange (BPT+PROSPECT AGN: 1010, 1011, 1110, 1111); Blue (PROSPECT AGN only: 0010). Any datapoint below  $J = 0.67$  (the horizontal dashed line) will, by construction, not be identified by PROSPECT. Consequently, the green and brown dots are AGN identified by all classifiers except PROSPECT. Above the PROSPECT threshold, the blue dots are AGN detected by PROSPECT alone, whereas the orange and cyan triangles are AGN identified by multiple identifiers, including PROSPECT. The figure clearly demonstrates that PROSPECT is sensitive to AGN hosted by low mass galaxies with low radio luminosities. The solid line represents the low mass threshold in (a) and the low radio luminosity threshold in (b).

While PROSPECT identifies 2 out of the 8 remaining AGN in common, none of the AGN detected by PROSPECT alone in our GEM sample are present in their low mass AGN sample. This is consistent with the above results demonstrating that narrow-line techniques are not complete in identifying AGN in low mass galaxies. While narrow-line diagnostic tools can identify bona fide AGN, such identification of AGN from low mass hosts is likely to be more incomplete than at high masses (Reines et al. 2013), emphasising the need for the use of a suite of diagnostics to optimise the completeness of AGN samples.

It should be noted here that this comparison does not check the completeness of the AGN population detected by PROSPECT. Our intention is to compare the low mass AGN samples and check whether the PROSPECT AGN are unique, which seems to be the case thereby reinforcing the importance of multiwavelength analysis.

### 5.2. Deciphering the 0010s

We have established that PROSPECT is efficient at picking up AGN from low mass systems, which are potentially weak radio emitters. Radio AGN are classified as weak-line radio galaxies (WLRGs) or strong-line radio galaxies (SLRGs) based on their optical emission line intensities (Tadhunter et al. 1998). With respect to their properties, WLRGs and SLRGs are almost similar to LERGs and HERGs, respectively (Tadhunter 2016). PROSPECT clearly identifies both strong-line and weak-line species. It detects HERGs as it is sensitive to some of the higher mass, higher radio luminosity AGN that are also identified by the emission line techniques. It also identifies low mass and low radio luminosity AGN that are most likely to be LERGs, as they are not identified by the emission line techniques.

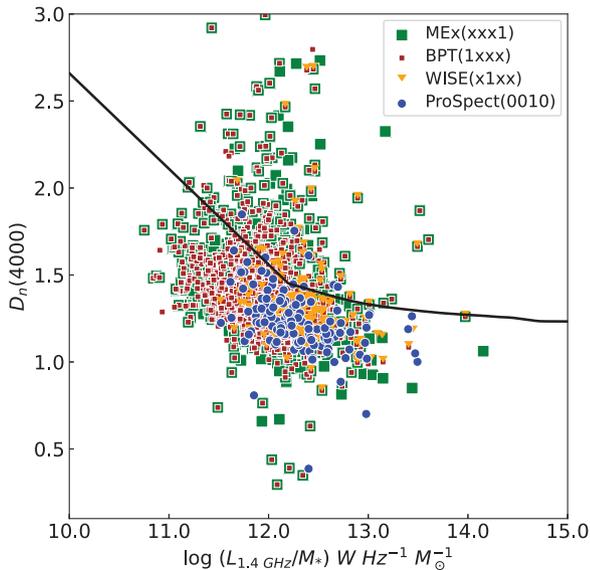
Fig. 11 shows the 1.4 GHz radio luminosity over galaxy stellar mass as a function of the  $D_n(4000)$  index. This method was introduced by Best et al. (2005), and they derived a demarcation between radio-loud AGN and SF galaxies using BC03 SSP models. The bulk of the AGN identified by PROSPECT alone (blue) lie in the region of this diagram identified by Best & Heckman (2012) as LERGs. Accordingly, those AGN identified by PROSPECT alone are most likely to be AGN in this low-excitation state.

We explain this result in terms of proportionally lower mass black holes in low mass galaxies as opposed to massive systems. Black holes are assumed to exist in the centres of almost all galaxies. The influence that these central engines have on their host galaxies can significantly affect their evolution, which in turn can have a feedback effect on the black holes. The proportion of black hole mass has been found to vary as a function of the mass of the host galaxy (see, e.g., figure 3 from Graham & Sahu 2023) such that lower mass galaxies host proportionally smaller mass black holes when compared to massive galaxies. Quantitatively, while a  $M_* = 10^{12} M_\odot$  galaxy typically hosts a  $10^{10} M_\odot$  black hole, a  $M_* = 10^9 M_\odot$  galaxy is found to typically host a black hole of mass of only  $10^5 M_\odot$ .

This could potentially explain why narrow-line diagnostics are relatively less efficient at detecting AGN in low mass host galaxies. Massive galaxies, where the black hole is a larger proportion by mass, can produce ionisation lines strong enough that then dominate and are easily detected by BPT and MEx. However, when the black holes are proportionally less massive, which is the case with low mass galaxies, the stellar contribution may outweigh the AGN contribution in the optical spectra, consequently minimising the AGN optical signature. This argument is reinforced by Fig. 11 since the AGN identified by PROSPECT alone lie in the LERG region, consistent with having weak optical emission features. A

**Table 5.** The cross-match between the GEM and low mass galaxy samples. There are 9 galaxies common to both datasets, with 8 identified by multiple classifiers and 1 SFG. The absence of AGN unique to PROSPECT means that the SED-based AGN hosted by low mass galaxies in our GEM sample is a novel set. We can be confident that the PROSPECT AGN are true AGN since the detection is rooted in radio excess.

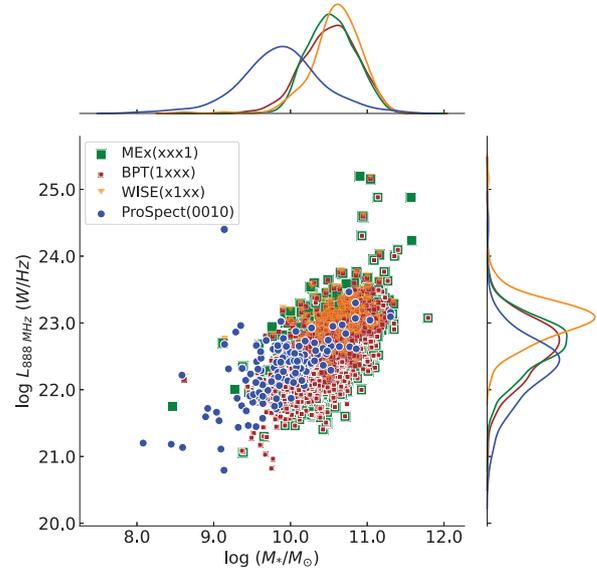
CATAID	Identifier	$\log(M_*/M_\odot)$	$\log L_{888 \text{ MHz}}$
5154526	0000	9.8	21.8
5364072	1000	9.7	21.7
5280475	1000	9.8	21.4
5241739	1000	9.8	22.9
5266552	1000	9.9	22.6
5197149	1000	9.9	22.5
5240995	1001	9.9	21.8
5368772	1010	9.7	22.4
5230865	1011	9.9	22.5



**Figure 11.**  $D_n(4000)$  as a function of 1.4 GHz radio luminosity per galaxy stellar mass. The blue dots are the AGN identified by PROSPECT alone, most of them which lie below the selection line. Brown and green squares represent the BPT and MEX AGN, respectively. The orange triangles are WISE AGN. The solid line delineates the optically active radio AGN (above) from the optically weak radio AGN (below). It corresponds to the 3 Gyr exponentially decaying SF track as prescribed in Best & Heckman (2012). By definition, SFGs with radio emission occupy the region below the demarcation line. The presence of PROSPECT alone AGN in this region indicate that these are low-excitation species with a radio excess.

more detailed investigation of this relationship is explored by Shabala (2018). That analysis links radio jets (and hence radio luminosity), accretion discs (optical emission line luminosity) and black hole mass. It is feasible for lower mass black holes to produce a radio jet that is more powerful (per unit black hole mass) than their higher mass counterparts, especially if they are accreting close to the Eddington limit (Turner & Shabala 2015).

Fig. 12 shows radio luminosity as a function of stellar mass colour-coded by the AGN diagnostics. This is a summary of the key outcomes regarding the capabilities of different classifiers. The figure reiterates the capabilities of PROSPECT in picking up



**Figure 12.** The variation of observed 888 MHz radio luminosities with stellar mass. The blue dots are the AGN identified by PROSPECT alone, which preferentially correspond to lower host stellar mass and radio luminosity. Other data points are AGN identified by multiple classifiers. Green squares are AGN by MEX, brown squares are AGN by BPT, and yellow triangles are AGN by WISE regardless of the classification by the other methods. The marginal plots on the X and Y axes, colour-coded using the same key, represent the density distribution of the points. AGN detected by PROSPECT alone (blue) are distributed towards low host galaxy stellar mass. These AGN are also distributed towards low radio luminosity when compared to AGN detected by other diagnostics.

potentially low radio luminosity AGN in low mass hosts over and above those picked up by other diagnostics.

Many AGN diagnostics exist, each with its own capabilities and parameter space in which they maximise their probability of identifying authentic AGN. The binary classifier scheme that we use makes it easy to see AGN uniquely identified by different diagnostic schemes. Not all diagnostics are perfectly coupled and the presence of AGN may not be simultaneously perceptible by different techniques. Since almost all galaxies are believed to host SMBH, it seems likely that identifying their presence as an AGN becomes more probable as more AGN diagnostics are brought to bear. Thus we suggest that the better discussion to have is about the fraction of AGN and star formation contribution from a galaxy rather than as an artificial binary distinction into AGN or SFG. The probabilistic approach suggested by Juneau *et al.* (2011) is an important step in this direction. The mixing between emission from star-formers and AGN in the BPT diagnostic have been studied as well (e.g., Thomas *et al.* 2018a, 2018b) and now SED tools like PROSPECT can add a new dimension to further extend this more detailed approach.

## 6. Conclusion

We have presented a spectral energy distribution fitting technique as a tool to identify nuclear activity. We have used the GAMA DR4 G23 data (spectrophotometry and stellar masses), the EMU early science data, and the AllWISE catalogue to perform a panchromatic SED fitting using PROSPECT, a generative SED fitting package. We used four AGN diagnostic tools: the BPT diagram and the mass-excitation (MEX) diagnostic diagram (two narrow-line based methods), the WISE colour criterion, together with the SED-based PROSPECT AGN identification tool.

The combined sample of 2956 G23 field galaxies were classified by each of these four diagnostics. For ease of illustration, we introduce a binary-based identification scheme (Table 3) where 0 corresponds to SFGs and 1 corresponds to AGN. The sample was divided into four mass bins, and we found 81 low mass galaxies hosting AGN unique to PROSPECT at masses below  $10^{10} M_{\odot}$ . These low mass hosts are weaker in their radio emission as well compared to those identified by other methods. A cross-match with the low mass AGN hosts in GAMA G23 compiled by Salehirad et al. (2022) also confirms that the SED-based low mass AGN hosts are unique to PROSPECT. Narrow-line diagnostics appear to struggle in identifying AGN in low mass galaxies. We present a possible cause for this selection difference in terms of the proportion of black hole mass in low and high mass galaxies. Low mass galaxies host proportionally lower mass black holes (Graham & Sahu 2023) thereby enabling their star forming signature to dominate and resulting in these galaxies being identified as star formers by narrow-line diagnostics. This effect is likely to be more prominent for radio galaxies in their low-excitation mode.

**Acknowledgement.** We thank Philip Best for providing the necessary data points for the demarcation line in Fig. 11. JP is supported by the Australian Government International Research Training Program (iRTP) Scholarship. IP acknowledges support from CSIRO under its Distinguished Research Visitor Programme to work on the GAMA23 EMU Early Science project. JA acknowledges financial support from the Science and Technology Foundation (FCT, Portugal) through research grants PTDC/FISAST/29245/2017, UIDB/04434/2020 (DOI:10.54499/UIDB/04434/2020) and UIDP/04434/2020 (DOI:10.54499/UIDP/04434/2020). MBI is supported by the Polish National Science Center through grants no. 2020/38/E/ST9/00395, 2018/30/E/ST9/00698, 2018/31/G/ST9/03388 and 2020/39/B/ST9/03494, and by the Polish Ministry of Science and Higher Education through grant DIR/WK/2018/12.

This scientific work uses data obtained from Inyarrimanha Ilgari Bundara/the Murchison Radio-astronomy Observatory. We acknowledge the Wajarri Yamaji People as the Traditional Owners and native title holders of the Observatory site. The Australian SKA Pathfinder is part of the Australia Telescope National Facility (<https://ror.org/05qajvd42>) which is managed by CSIRO. Operation of ASKAP is funded by the Australian Government with support from the National Collaborative Research Infrastructure Strategy. ASKAP uses the resources of the Pawsey Supercomputing Centre. Establishment of ASKAP, the Murchison Radio-astronomy Observatory and the Pawsey Supercomputing Centre are initiatives of the Australian Government, with support from the Government of Western Australia and the Science and Industry Endowment Fund. This paper includes archived data obtained through the CSIRO ASKAP Science Data Archive, CASDA (<http://data.csiro.au>).

GAMA is a joint European-Australasian project based around a spectroscopic campaign using the AAT. The GAMA input catalogue is based on data taken from the SDSS and the UKIRT Infrared Deep Sky Survey. Complementary imaging of the GAMA regions is being obtained by a number of independent survey programmes including GALEX MIS, VST KiDS, VISTA VIKING, WISE, *Herschel*-ATLAS, GMRT and ASKAP providing UV to radio coverage. GAMA is funded by the STFC (UK), the ARC (Australia), the AAO, and the participating institutions. The GAMA website is <http://www.gama-survey.org/>.

**Data availability.** Not applicable.

## References

Ahmed, U. T., et al. 2023, arXiv e-prints, arXiv:2312.11883  
 Alonso-Herrero, A., et al. 2006, ApJ, 640, 167  
 Assef, R. J., et al. 2018, ApJS, 234, 23

Assef, R. J., et al. 2013, ApJ, 772, 26  
 Backhaus, B. E., et al. 2022, ApJ, 926, 161  
 Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5  
 Barmby, P., et al. 2006, ApJ, 642, 126  
 Belfiore, F., et al. 2016, MNRAS, 461, 3111  
 Bellstedt, S., et al. 2020a, MNRAS, 498, 5581  
 Bellstedt, S., et al. 2020b, MNRAS, 496, 3235  
 Best, P. N., & Heckman, T. M. 2012, MNRAS, 421, 1569  
 Best, P. N., Kauffmann, G., Heckman, T. M., & Ivezić, Ž. 2005, MNRAS, 362, 9  
 Best, P. N., et al. 2023, MNRAS, 523, 1729  
 Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000  
 Chabrier, G. 2003, PASP, 115, 763  
 Charlot, S., & Fall, S. M. 2000, ApJ, 539, 718  
 Cid Fernandes, R., Stasińska, G., Mateus, A., & Vale Asari, N. 2011, MNRAS, 413, 1687  
 Cid Fernandes, R., et al. 2010, MNRAS, 403, 1036  
 Cluver, M. E., et al. 2014, ApJ, 782, 90  
 Condon, J. 1992, ARA&A, 30, 575  
 Conroy, C. 2013, ARA&A, 51, 393  
 Cutri, R. M., et al. 2012, VizieR Online Data Catalog, II/311  
 Cutri, R. M., et al. 2021, VizieR Online Data Catalog, II/328  
 Dale, D. A., et al. 2014, ApJ, 784, 83  
 de Jong, J. T. A., et al. 2015, A&A, 582, A62  
 Delvecchio, I., et al. 2021, A&A, 647, A123  
 Donley, J. L., Rieke, G. H., Pérez-González, P. G., Rigby, J. R., & Alonso-Herrero, A. 2007, ApJ, 660, 167  
 Donley, J. L., et al. 2012, ApJ, 748, 142  
 Driver, S. P., 2009, in The Galaxy Disk in Cosmological Context, Vol. 254, ed. J. Andersen, B. m. Nordströara, & J. Bland-Hawthorn, 469  
 Driver, S. P., et al. 2011, MNRAS, 413, 971  
 Driver, S. P., et al. 2022, MNRAS, 513, 439  
 Edge, A., et al. 2013, Msngr, 154, 32  
 Genzel, R., et al. 1998, ApJ, 498, 579  
 Gordon, Y. A., et al. 2017, MNRAS, 465, 2671  
 Graham, A. W., & Sahu, N. 2023, MNRAS, 520, 1975  
 Gunawardhana, M. L. P., et al. 2011, MNRAS, 415, 1647  
 Gürkan, G., et al. 2022, MNRAS, 512, 6104  
 Gürkan, G., Hardcastle, M. J., & Jarvis, M. J. 2014, MNRAS, 438, 1149  
 Gürkan, G., et al. 2018, MNRAS, 475, 3010  
 Hardcastle, M. J., Evans, D. A., & Croston, J. H. 2007, MNRAS, 376, 1849  
 Heckman, T. M., & Best, P. N. 2014, ARA&A, 52, 589  
 Hine, R. G., & Longair, M. S. 1979, MNRAS, 188, 111  
 Hopkins, A. M., et al. 2003, ApJ, 599, 971  
 Hotan, A., et al. 2021, PASA, 38, e009  
 Jarrett, T. H., et al. 2011, ApJ, 735, 112  
 Johnston, S., et al. 2007, PASA, 24, 174  
 Juneau, S., Dickinson, M., Alexander, D. M., & Salim, S. 2011, ApJ, 736, 104  
 Juneau, S., et al. 2013, ApJ, 764, 176  
 Kauffmann, G., et al. 2003, MNRAS, 346, 1055  
 Kellermann, K. I., Sramek, R., Schmidt, M., Shaffer, D. B., & Green, R. 1989, AJ, 98, 1195  
 Kewley, L., Heisler, C., Dopita, M., & Lumsden, S. 2001, ApJS, 132, 37  
 Kewley, L. J., Groves, B., Kauffmann, G., & Heckman, T. 2006, MNRAS, 372, 961  
 Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511  
 Kormendy, J., & Kennicutt, R. C., Jr. 2004, ARA&A, 42, 603  
 Lacy, M., et al. 2007, AJ, 133, 186  
 Lacy, M., et al. 2004, ApJS, 154, 166  
 Lamareille, F., Mouhcine, M., Contini, T., Lewis, I., & Maddox, S. 2004, MNRAS, 350, 396  
 Laurent, O., et al. 2000, A&A, 359, 887  
 Leahy, D. A., et al. 2019, PASA, 36, e024  
 Liske, J., et al. 2015, MNRAS, 452, 2087  
 Magorrian, J., et al. 1998, AJ, 115, 2285  
 Marocco, F., et al. 2021, ApJS, 253, 8  
 Martin, D. C., et al. 2005, ApJ, 619, L1  
 Marvil, J., Owen, F., & Eilek, J. 2015, AJ, 149, 32

- Mateos, S., *et al.* 2012, *MNRAS*, **426**, 3271
- McConnell, D., *et al.* 2016, *PASA*, **33**, e042
- Messias, H., Afonso, J., Salvato, M., Mobasher, B., & Hopkins, A. M. 2012, *ApJ*, **754**, 120
- Messias, H., Afonso, J. M., Salvato, M., Mobasher, B., & Hopkins, A. M. 2014, *A&A*, **562**, A144
- Mushotzky, R. 2004, in *Astrophysics and Space Science Library*, Vol. 308, Supermassive Black Holes in the Distant Universe, ed. A. J. Barger, **53**
- Norris, R. P., *et al.* 2011, *PASA*, **28**, 215
- Norris, R. P., *et al.* 2021, *PASA*, **38**, e046
- Osterbrock, D. E., & De Robertis, M. 1985, *PASP*, **97**, 1129
- Pacifici, C., *et al.* 2023, *ApJ*, **944**, 141
- Padovani, P., & Giommi, P. 1995, *MNRAS*, **277**, 1477
- Padovani, P., *et al.* 2017, *A&A Rev.*, **25**, 2
- Pilbratt, G. L., *et al.* 2010, *A&A*, **518**, L1
- PLANCK Collaboration, *et al.* 2016, *A&A*, **594**, A13
- Pracy, M. B., *et al.* 2016, *MNRAS*, **460**, 2
- Reines, A. E., Greene, J. E., & Geha, M. 2013, *ApJ*, **775**, 116
- Robotham, A. S. G., *et al.* 2020, *MNRAS*, **495**, 905
- Sajina, A., Lacy, M., & Scott, D. 2005, *ApJ*, **621**, 256
- Salehirad, S., Reines, A. E., & Molina, M. 2022, *ApJ*, **937**, 7
- Searle, L. 1971, *ApJ*, **168**, 327
- Shabala, S. S. 2018, *MNRAS*, **478**, 5074
- Skrutskie, M. F., *et al.* 2006, *AJ*, **131**, 1163
- Smith, D. J. B., *et al.* 2021, *A&A*, **648**, A6
- Statisticat, & LLC. 2021, LaplacesDemon: Complete Environment for Bayesian Inference, r package version 16.1.6
- Stern, D., *et al.* 2005, *ApJ*, **631**, 163
- Stern, D., *et al.* 2012, *ApJ*, **753**, 30
- Sturm, E., *et al.* 2002, *A&A*, **393**, 821
- Tadhunter, C. 2016, *A&A Rev.*, **24**, 10
- Tadhunter, C. N., *et al.* 1998, *MNRAS*, **298**, 1035
- Taylor, E. N., *et al.* 2011, *MNRAS*, **418**, 1587
- Taylor, M. B. 2005, in *Astronomical Society of the Pacific Conference Series*, Vol. 347, *Astronomical Data Analysis Software and Systems XIV*, ed. P. Shopbell, M. Britton, & R. Ebert, **29**
- Thomas, A. D., *et al.* 2018a, *ApJ*, **856**, 89
- Thomas, A. D., *et al.* 2018b, *ApJ*, **861**, L2
- Thorne, J. E., Robotham, A. S. G., Bellstedt, S., & Davies, L. J. M. 2023, *MNRAS*, **522**, 6354
- Thorne, J. E., *et al.* 2022, *MNRAS*, **509**, 4940
- Tremonti, C. A., *et al.* 2004, *ApJ*, **613**, 898
- Trouille, L., Barger, A. J., & Tremonti, C. 2011, *ApJ*, **742**, 46
- Turner, R. J., & Shabala, S. S. 2015, *ApJ*, **806**, 59
- Urry, C. M., & Padovani, P. 1995, *PASP*, **107**, 803
- Vazdekis, A., Koleva, M., Ricciardelli, E., Röck, B., & Falcón-Barroso, J. 2016, *MNRAS*, **463**, 3409
- Veilleux, S., & Osterbrock, D. E. 1987, *ApJS* (ISSN 0067-0049), **63**, 295. NSERC-supported research., **63**, 295
- Walcher, J., Groves, B., Budavári, T., & Dale, D. 2011, *Ap&SS*, **331**, 1
- White, R. L., *et al.* 2000, *ApJS*, **126**, 133
- Wilkes, B. J., *et al.* 2002, *ApJ*, **564**, L65
- Willott, C. J., Rawlings, S., Blundell, K. M., & Lacy, M. 1999, *MNRAS*, **309**, 1017
- Wright, E. L., *et al.* 2010, *AJ*, **140**, 1868
- Yan, R., *et al.* 2011, *ApJ*, **728**, 38
- Yang, G., *et al.* 2023, *ApJ*, **950**, L5
- York, D. G., *et al.* 2000, *AJ*, **120**, 1579
- Yun, M. S., Reddy, N. A., & Condon, J. J. 2001, *ApJ*, **554**, 803
- Zeimann, G., *et al.* 2015, in *American Astronomical Society Meeting Abstracts*, Vol. 225, *American Astronomical Society Meeting Abstracts #225*, **206.04**