

# Comparing metallicity correlations in nearby non-AGN and AGN-host galaxies

Song-Lin Li,<sup>1,2★</sup> Zefeng Li<sup>1,2,3★</sup>, Emily Wisnioski<sup>1,2</sup>, Mark R. Krumholz<sup>1,2</sup>  
and Sebastián F. Sánchez<sup>4,5</sup>

<sup>1</sup>Research School of Astronomy & Astrophysics, Australian National University, Canberra ACT 2611, Australia

<sup>2</sup>ARC Centre of Excellence for All-Sky Astrophysics in 3 Dimensions (ASTRO 3D), Canberra ACT 2611, Australia

<sup>3</sup>Department of Physics, Centre for Extragalactic Astronomy, Durham University, South Road, Durham DH1 3LE, UK

<sup>4</sup>Instituto de Astronomía, Universidad Nacional Autónoma de México, AP 106, Ensenada 22800, BC, México

<sup>5</sup>Instituto de Astrofísica de Canarias, Vía Láctea s/n, 38205 La Laguna, Tenerife, Spain

Accepted 2024 November 19. Received 2024 November 19; in original form 2024 July 4

## ABSTRACT

The gas-phase metallicity distribution within galaxies records critical information about galactic evolution. In this work, we investigate how active galactic nuclei (AGNs) influence this distribution by measuring the two-point correlation functions of gas-phase metallicity in 95 non-AGN and 37 AGN-host galaxies from the Calar Alto Legacy Integral Field spectroscopy Area integral field spectrographic survey. We measure metallicity using a novel Bayesian method that properly includes both stellar and AGN contributions to emission line fluxes and allows us to measure metallicities in both AGN-host and non-AGN galaxies in a single, consistent framework. We find that the two-point correlation functions of both AGN-host and non-AGN galaxies are well fit by a simple injection-diffusion model, and that the correlation lengths  $l_{\text{corr}}$  we derive for the non-AGN galaxies are reasonably consistent with those obtained in earlier work. The AGN-host galaxies generally have smaller  $l_{\text{corr}}$  than non-AGN galaxies at fixed stellar mass, but similar  $l_{\text{corr}}$  at fixed star formation rate (SFR), suggesting that the primary effect of hosting an AGN in this sample is a reduction in SFR at fixed stellar mass, and that this in turn suppresses the correlation length. Our findings further indicate that, while both SFR and stellar mass are positively correlated with metallicity correlation length  $l_{\text{corr}}$ , the former is more fundamental, implying that fluctuations in the metallicity distribution within galaxies are driven more by short-term responses to physical processes such as star formation that can change much faster than a Hubble time.

**Key words:** galaxies: abundances – galaxies: ISM – galaxies: Seyfert.

## 1 INTRODUCTION

Gas-phase metallicity, regulated by galactic inflows, outflows, star formation, and stellar nucleosynthesis, records key processes of galaxy formation and evolution (e.g. Sharda et al. 2021, 2024). Thanks to large spectroscopic sky surveys such as Sloan Digital Sky Survey (SDSS), a number of empirical scaling relations for metallicity have been uncovered that can help us understand these processes. One of the most famous empirical relations is the mass–metallicity relation (MZR), a positive correlation between a galaxy’s stellar mass and its gas-phase metallicity (e.g. Tremonti et al. 2004; Kewley & Ellison 2008; Andrews & Martini 2013; Pérez-Montero et al. 2013). Beyond this primary dependence, a number of authors have also identified secondary dependencies on parameters other than stellar mass. For example, metallicity anticorrelates with star formation rate (SFR) at fixed stellar mass, as first demonstrated by Lara-López et al. (2010) and Mannucci et al. (2010). The correlation between stellar mass, SFR, and metallicity is likely

more fundamental than the MZR, and is therefore known as the fundamental metallicity relation (FZR); however, for a contrasting perspective see Sánchez et al. (2013, 2017, 2019) and Barrera-Ballesteros et al. (2017). A secondary dependence of metallicity on atomic or molecular hydrogen mass has also been explored (Bothwell et al. 2013, 2016; Brown et al. 2018). These scaling relations are also found at high redshifts, showing a clear evolution through cosmic time (e.g. Erb et al. 2006; Maier et al. 2014; Zahid et al. 2014), and have recently extended to lower stellar masses and to redshifts up to 10 by *JWST* observations (Nakajima et al. 2023; Li et al. 2023b; He et al. 2024).

The prevalence of the integral field spectroscopy (IFS) surveys of nearby galaxies, such as MaNGA (Mapping Nearby Galaxies at Apache Point Observatory; Bundy et al. 2015), SAMI (Sydney-AAO Multi-object Integral-field spectrograph; Bryant et al. 2015), and CALIFA (Calar Alto Legacy Integral Field Area; Sánchez et al. 2012), makes it possible to explore spatially resolved metallicity scaling relations. For example, there is a local correlation between stellar mass surface density, star formation surface density, and metallicity, which has led some authors to suggest that the MZR and FZR are simply the cumulative effects of a more fundamental local scaling relation, an interpretation that emphasizes the importance of

\* E-mail: [songlin.li@anu.edu.au](mailto:songlin.li@anu.edu.au) (S-LL); [zefeng.li@durham.ac.uk](mailto:zefeng.li@durham.ac.uk) (ZL)

\*

local properties such as gas fraction in regulating the metallicity (e.g. Rosales-Ortega et al. 2012; Barrera-Ballesteros et al. 2016; Teklu et al. 2020; Baker et al. 2023). An azimuthally averaged metallicity gradient is another property that can be measured from IFS data (Sánchez et al. 2014). Belfiore et al. (2017) and Poetrodjojo et al. (2018) find that the metallicity gradients steepen with the stellar mass in local star-forming (SF) galaxy populations. If we envision increasing stellar mass as an evolutionary sequence, this steepening requires metal redistribution by gas flows toward outskirts or metal dilution by pristine gas inflow (e.g. Mott, Spitoni & Matteucci 2013; Maiolino & Mannucci 2019).

In addition to mean metallicities and radial gradients, studies of IFS metallicity maps have begun to explore more complex statistical characterizations of 2D metallicity fields. For example, Sánchez-Menguiano et al. (2020) explore the metallicity variations between spiral-arm and interarm regions. However, perhaps the simplest non-trivial statistic one can compute from IFS data is the two-point correlation function, which reveals the interplay between metal production and metal mixing. Krumholz & Ting (2018) propose a theoretical model for this statistic based on competition between supernova production and metal diffusion, and several authors have measured it, or closely related statistics, in small samples of very nearby galaxies (e.g. Kreckel et al. 2019, 2020; Metha, Trenti & Chu 2021; Metha et al. 2022). By using the CALIFA and AMUSING++ IFS surveys (López-Cobá et al. 2020), Li et al. (2021, hereafter L21) and Li et al. (2023a, hereafter L23) successfully achieve larger sample sizes ( $> 100$  galaxies) and a significantly wider range of galaxy mass, SFR, and morphological type than the earlier studies. They found that Krumholz & Ting’s model provides good fits to the observed two-point correlation functions across the full sample. The key fitting parameter derived from this analysis is the correlation length ( $l_{\text{corr}}$ ) of the metallicity field. Observed  $l_{\text{corr}}$  are typically  $\sim \text{kpc}$ , and are positively correlated with stellar mass ( $M_*$ ), SFR, and galaxy effective radius ( $R_e$ ). Which of these is the most fundamental is not clear, as these three parameters are all correlated with each other as well.

Most studies of gas-phase metallicities to date have excluded galaxies where active galactic nuclei (AGNs) makes significant contributions to the luminosity. This exclusion has been necessary because most metallicity calibrators have been developed for H II regions ionized by stellar sources, and are not reliable for gas ionized by non-stellar sources that typically produce much harder spectra (Kewley, Nicholls & Sutherland 2019a). There are a few exceptions where authors have attempted to derive the metallicity for AGNs (e.g. Carvalho et al. 2020; Dors et al. 2020; Dors 2021) using metallicity calibrators developed specifically for AGN-like ionizing spectra. However, these calibrators are limited in pure AGN-dominant galaxies and not applicable to galaxies with significant contributions from H II regions. Moreover, because different metallicity diagnostics have substantial systematic differences (Kewley et al. 2019a, and the references therein), it is not straightforward to compare metallicities derived with these AGN-specific methods to those derived for non-AGN galaxies using different methods. When studying how the presence of an AGN influences galaxy metallicities, we require a more general approach that uses a uniform set of line diagnostics to disentangle the relative contribution from the SF-driven and AGN-driven ionisation in a self-consistent way.

The need for such a method was a primary motivation for the development of NEBULABAYES, a Bayesian code that estimates the gas properties including metallicity by comparing a set of observed emission-line fluxes to photoionization model grids (Thomas et al. 2018). By adopting a mixture of HII-region grids and narrow-line

region (NLR) grids where the gas is ionized by nuclear activity, NEBULABAYES is able to measure metallicity in galaxies ranging from purely SF to Seyferts where AGN is the dominant excitation source. Thomas et al. (2019) use NEBULABAYES to investigate the MZR for local Seyfert galaxies with SDSS single-fibre spectra. Li et al. (2024, hereafter L24) extend this work to the FZR by applying NEBULABAYES to the MaNGA IFS survey. They find that AGN hosts generally have higher metallicities than non-AGN galaxies at fixed stellar masses, but that this is primarily due to their lower SFR compared to non-AGN galaxies, i.e. the difference is a natural result of the existence of FZR.

Our primary goal in this work is to combine the Bayesian AGN-SF decomposition method developed in L24 to the study of full 2D metallicity distributions and their two-point correlations using the statistical techniques pioneered by L21 and L23. Our motivation is to investigate how the presence of an AGN influences the distribution of metals *within* galaxies, in addition to their overall mean metallicities. For example, one might anticipate that AGN could influence the metallicity fields by triggering galactic-scale gas outflows, or that AGN activity might be correlated with merging events that are expected to increase the correlation lengths (e.g. Harrison et al. 2014; Fischer et al. 2015; King & Pounds 2015; Koss et al. 2018; Comerford et al. 2024). We aim to determine whether any such influence is visible in the data.

Our plan for the rest of this paper is as follows. In Section 2, we introduce our sample selection procedures, the implementation of NEBULABAYES, and how to measure correlation lengths from metallicity fluctuation maps. We present our results and discuss the implication of the results in Sections 3 and 4, respectively. Finally, we summarize our findings in Section 5.

Throughout the work, we adopt a Chabrier (2003) initial mass function (IMF) and a flat WMAP7 cosmology:  $H_0 = 70.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_M = 0.27$ , and  $\Omega_\Lambda = 0.73$  (Komatsu et al. 2011).

## 2 METHODS

In this section, we start with a brief introduction to the CALIFA survey from which we draw our sample, and to our sample selection criteria, in Section 2.1. We then introduce our implementation of NEBULABAYES and describe how we use it to obtain metallicity maps in Section 2.2. In Sections 2.3 and 2.4, we explain our procedures to calculate galactic two-point correlation functions and fit them by Markov chain Monte Carlo (MCMC) sampling to acquire the correlation lengths  $l_{\text{corr}}$ , respectively. Finally, in Section 2.5, we discuss ancillary data on galaxy masses and SFRs that we use in our analysis.

### 2.1 Sample selection

CALIFA is an IFU survey using the 3.5 m telescope at the Calar Alto Observatory to obtain three-dimensional spectra of nearby galaxies ( $0.0005 < z < 0.05$ ) covering a wide range of stellar mass, morphology, and environment (Sánchez et al. 2012; Walcher et al. 2014). The observations are performed with the Potsdam Multi Aperture Spectrograph (PMAS) in the PPak mode (Roth et al. 2005; Kelz et al. 2006). In this work, we use the V500 grating that covers 3750–7500 Å with a typical spectral resolution  $R \sim 850$  around 5000 Å ( $\sim 360 \text{ km s}^{-1}$ ). The PMAS instrument covers at least two effective radii ( $R_e$ ) of each target. The data reduction pipeline before DR3 (v2.2) recovers a typical spatial resolution  $\sim 2''.50$ , corresponding to 760 pc at  $z \sim 0.015$  (Sánchez et al. 2016).

The latest extended data release (eDR3) contains 895 galaxies and develops a new reduction algorithm (v2.3) that enhances the spatial resolution to  $\sim 1''.50$  (Sánchez et al. 2023). Our parent sample is from eDR3. However, we still use the v2.2 pipeline to acquire spatially resolved properties to allow direct comparison with the work of L21, who used this pipeline, and who demonstrate that the spatial resolution  $\sim 2''.50$  available from it is sufficient to obtain accurate measurements of  $l_{\text{corr}}$ . The three-dimensional spectral cubes are fed into the spectroscopy analysis package `pyPipe3D` to analyse the stellar populations and fit emission lines (Lacerda et al. 2022).

Sánchez et al. (2023) provide classifications for galaxies in the CALIFA eDR3 sample based on the locations on Baldwin, Phillips, & Terlevich (BPT) diagram (Baldwin, Phillips & Terlevich 1981; Kewley et al. 2001) applied to the emission-line flux ratios extracted from central 1.50 arcsec of each target galaxy. In this work, we only use three out of six groups provided. The galaxies in these three categories are required to have a signal-to-noise ratio (SNR) greater than 3 in the  $H\alpha$  emission line and greater than unity in the  $H\beta$ ,  $[N\text{ II}] \lambda 6583$ , and  $[O\text{ III}] \lambda 5007$  lines. With this restriction, the three groups we use are

- (i) star forming (BPT-SF): galaxies in this class are characterized by an equivalent width in the  $H\alpha$  emission line,  $\text{EW}(H\alpha)$ , above  $3\text{ \AA}$ , and by line ratios  $\log([O\text{ III}]/H\beta)$  and  $\log([N\text{ II}]/H\alpha)$  that place them below the maximum starburst line proposed by Kewley et al. (2001) on the BPT diagram.
- (ii) weak AGN (wAGN): galaxies in this category have line ratios above the maximum starburst line, and  $\text{EW}(H\alpha)$  from 3 to  $6\text{ \AA}$ .
- (iii) strong AGN (sAGN): These are galaxies whose line ratios place them above the maximum starburst line and that have  $\text{EW}(H\alpha)$  greater than  $6\text{ \AA}$ .

After selecting galaxies in these three categories, we apply a series of additional cuts to the sample. We remove galaxies with axis ratios smaller than 0.4; we calculate the axis ratio as  $b/a = \sqrt{1 - e^2}$ , where  $e$  is the eccentricity provided in the property summary catalogue for eDR3. This cut excludes galaxies that are close enough to edge-on to suffer from significant dust attenuation and mixing of light from regions with different physical conditions.

In addition to removing galaxies from the sample based on axis ratios, within each galaxy we follow L24 by also masking spaxels where  $\text{EW}(H\alpha)$  is below  $3\text{ \AA}$ , or wherever the  $H\alpha$  or  $H\beta$  SNR is smaller than 3. The  $\text{EW}(H\alpha)$  cut is intended to remove possible contamination from post-asymptotic giant branch (pAGB) stars (Cid Fernandes et al. 2011; Belfiore et al. 2016; Lacerda et al. 2018), whereas the Balmer-line only SNR criterion is to avoid a possible sample bias to lower metallicity when applying SNR cuts independently to multiple metal lines (Salim et al. 2014; Thomas et al. 2019). L24 demonstrate that applying a higher SNR cut to Balmer lines does not change the results. Finally, after applying these pixel-by-pixel masks, we remove from the sample any galaxies where we have fewer than 500 non-masked spaxels, a number that L21 find is the minimum required for accurate measurements of two-point correlation functions. The final sample we use after applying these filters consists of 97 BPT-SF galaxies, 3 wAGN galaxies, and 35 sAGN galaxies.

Finally, we note that by using the BPT diagram to identify galaxies as AGN hosts, we introduce a potential bias: we might be more likely to mis-classify AGN-host galaxies with higher SFRs as non-hosts because their bright stellar-driven emission masks the AGN signature. This is a potential concern since, as we show below, we find that there is a systematic difference in the SFRs of AGN-host and non-host galaxies. To mitigate this concern, in Appendix A we follow

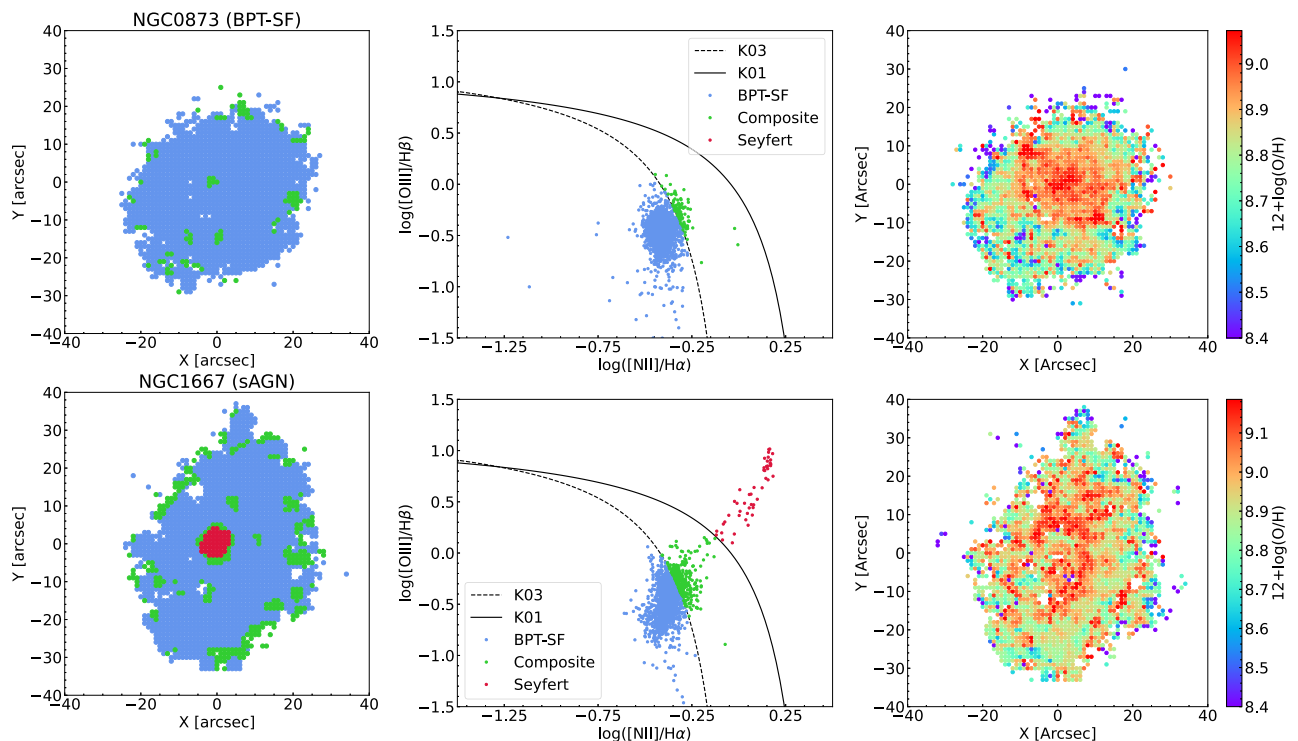
Tozzi et al. (2023) by searching for signatures of  $\text{He II}$  emission as an additional AGN diagnostic, one less subject to this bias. In the appendix we show that, while this check does cause us to reclassify two of our BPT-SF galaxies as AGN hosts, this is too small an effect to materially alter our conclusions. For consistency with prior work, we therefore continue to use the classification based on the BPT diagram explained above for the main body of this paper.

## 2.2 Application of NEBULABAYES

We use NEBULABAYES (Thomas et al. 2018) to calculate metallicity maps for our sample. NEBULABAYES estimates a series of Bayesian posterior parameters by comparing the observed emission-line fluxes with model grids produced by a photoionization code such as MAPPINGS (Sutherland & Dopita 2017). Instead of using the Solar-scaled scheme, we adopt the more realistic Galactic Concordance (Nicholls et al. 2017) scheme to set the relative abundance of metal elements as a function of gas-phase metallicity. In particular, the nitrogen-to-oxygen ratio as a function of  $12 + \log(\text{O}/\text{H})$ , which determines the  $[N\text{ II}]/[O\text{ II}]$  ratio we use as a prior to estimate the metallicity in NEBULABAYES, is obtained by fitting nearby stars and nebulae with a combination of two linear functions that address both the primary and secondary nitrogen. In this work, we use the same procedure and set of photoionization model grids as in L24. Here, we briefly summarize the procedures and refer readers to L24 for more details.

To apply NEBULABAYES, we first classify spaxels as BPT-SF, Composite, or Seyfert based on their position in the BPT diagram; this classification is an efficiency measure, since it lets us avoid carrying out computationally expensive fits using arbitrary combinations of stellar and AGN ionization models in spaxels where there is clearly no AGN contribution. Note that, these spaxel-by-spaxel classifications are not identical to the overall galaxy classifications discussed in Section 2.1, which are based on integrated central spectra, so that a galaxy that is classified as BPT-SF may still contain spaxels with an AGN contribution, and vice versa. Fig. 1 shows the spaxel-by-spaxel classifications for one example galaxy classified as BPT-SF and one as sAGN. The left panels show the spaxel classifications, while the middle panels show the locations of each spaxel in the BPT diagram. Spaxels that fall below the demarcation line proposed by Kauffmann et al. (2003, dashed lines) are classified as BPT-SF, and are assumed to be ionized by radiation from massive stars. Spaxels above the theoretical maximum starburst line proposed by Kewley et al. (2001, solid lines) are classified as Seyfert, and are presumably dominated by AGN ionization. Spaxels lying between the two lines are classified as Composite, and likely mix both sources of ionisation.

For BPT-SF spaxels, we assume emission is solely from  $\text{H II}$  regions, and we therefore fit to  $\text{H II}$ -region-only models, which have three free parameters: the oxygen abundance  $12 + \log(\text{O}/\text{H})$  (i.e. the gas-phase metallicity), the ionization parameter at the inner edge of the modeled nebula  $\log U_{\text{H II}}$ , and the gas pressure  $\log(P/k)$ . We use the ‘Line-ratio prior mode’ in NEBULABAYES for these spaxels; in this mode, we run NEBULABAYES three times, with each run using only one emission-line ratio that is sensitive to one parameter:  $[N\text{ II}] \lambda 6583/[O\text{ II}] \lambda \lambda 3626$ , 29 for  $12 + \log(\text{O}/\text{H})$  (Kewley & Dopita 2002),  $[O\text{ III}] \lambda 5007/[O\text{ II}] \lambda \lambda 3626$ , 29 for  $\log U_{\text{H II}}$  (Kewley & Dopita 2002; Kobulnicky & Kewley 2004), and  $[S\text{ II}] \lambda 6716/[S\text{ II}] \lambda 6731$  for  $\log(P/k)$  (Kewley et al. 2019b). We correct dust reddening by using the difference between the intrinsic  $H\alpha/H\beta$  ratio of each point of the model grid and the observed  $H\alpha/H\beta$  ratio to deredden all lines before comparing the data to that model (see L24 for details). This yields three sets of posterior



**Figure 1.** A walkthrough of our analysis procedure for two sample galaxies: the upper row shows NGC0873, which is classified as a BPT-SF galaxy, while the lower row shows NGC1667, which is classified as sAGN. The left panels show the classifications we assign to each pixel; different colours correspond to different locations in the BPT diagrams shown in the middle panels, as indicated in the legend. The dashed lines in the middle panels are the demarcations proposed by Kauffmann et al. (2003, K03) to distinguish pure SF galaxies from those with AGN contributions, while the solid lines show the theoretical maximum starburst locus proposed by Kewley et al. (2001, K01); pixels below the dashed lines are classified as BPT-SF, those above the solid lines are classified as Seyfert, and those between the two as Composite. The right panels show the final metallicity maps obtained from NEBULABAYES.

likelihoods for each point of our model grid, which we then multiply together to get the final posterior, with the  $[\text{N II}]/[\text{O II}]$  weighted three times as much as other line ratios since metallicity is the key parameter of interest in this work. As discussed in L24, this simplified procedure for BPT-SF spaxels yields results for the metallicities that are very close to those obtained by carrying out a full run of NEBULABAYES using all available emission lines, but at a small fraction of the computational cost.

For Composite and Seyfert spaxels, we fit a full model whereby we assume that the emission in that spaxel represents a linear combination of an H II-region component and a narrow-line region (NLR) component ionized by a theoretical AGN spectrum generated using the OXAF code (Thomas et al. 2016). We adopt a power-law index  $\Gamma = -2.0$  for the non-thermal component of the AGN spectrum and assume that this component contributes a fraction  $p_{\text{NT}} = 0.15$  of the total AGN luminosity. We fix the ionization parameter of H II regions in these spaxels to  $\log U_{\text{HII}} = -3.25$ , since this is the median  $\log U_{\text{HII}}$  of BPT-SF spaxels found in L23, and we assume that each spaxel is characterized by a single oxygen abundance and gas pressure. Under these assumptions, we have five free parameters:  $\log(P/k)$ ,  $12 + \log(\text{O}/\text{H})$ , the ionization parameter of the NLR  $\log U_{\text{NLR}}$ , the energy at which the Big Blue Bump component from the AGN accretion disc has its peak  $\log(E_{\text{peak}}/\text{keV})$ , and the fractional contribution of H II regions to the  $\text{H}\alpha$  luminosity  $f_{\text{HII}}$ . In addition, since all spaxels within the same galaxy should see the same AGN spectrum, we require  $\log(E_{\text{peak}}/\text{keV})$  to be the same for all spaxels in a single galaxy; we adopt a value of  $-1.35$  for this parameter for spaxels in BPT-SF galaxies (Thomas et al. 2019), while for wAGN and sAGN

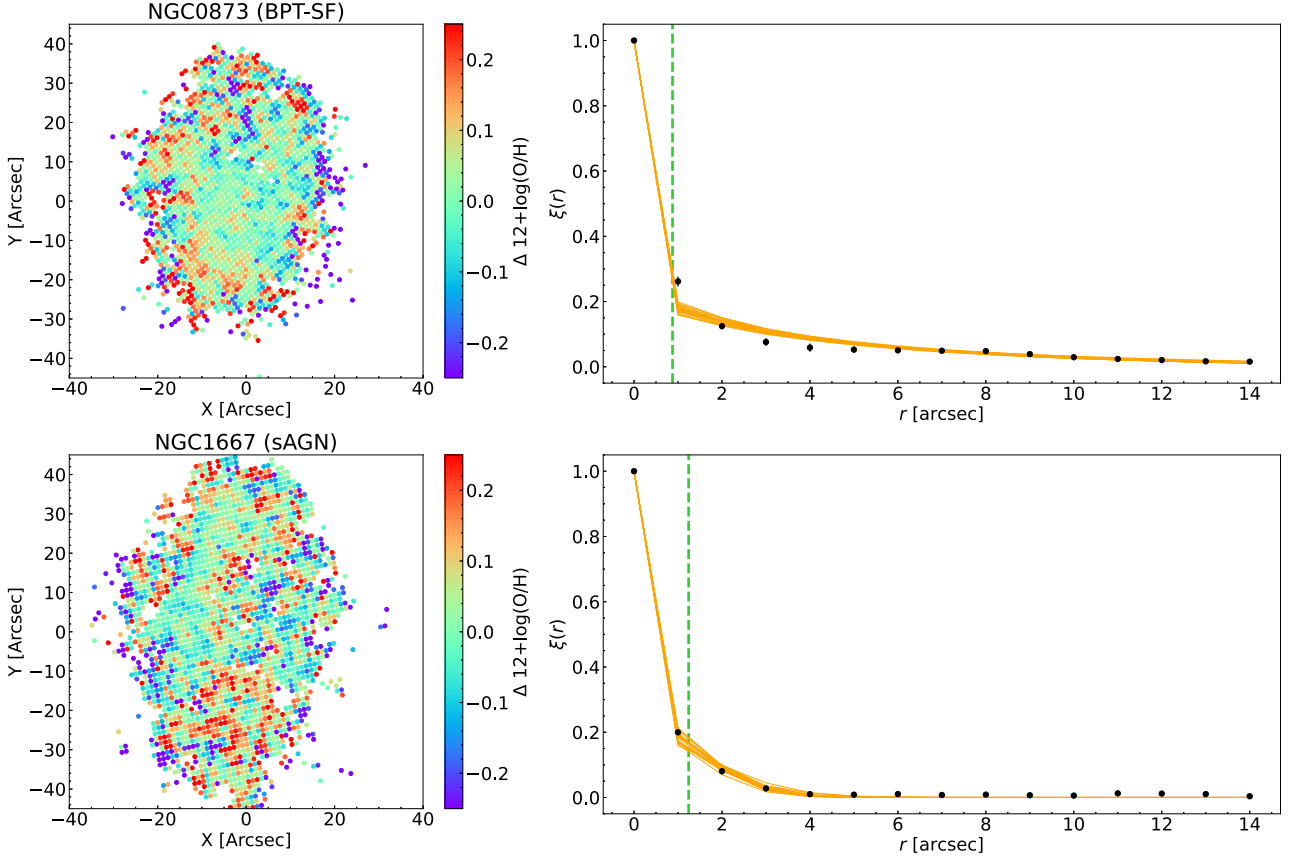
galaxies, we choose a value by selecting Seyfert spaxels within central  $3''$  and running NEBULABAYES with  $\log(E_{\text{peak}}/\text{keV})$  left as a free parameter, and fix the value of  $\log(E_{\text{peak}}/\text{keV})$  for that galaxy to the modal outcome over these spaxels. We then run NEBULABAYES three times to determine the remaining four parameters. The first two runs are in ‘Line-ratio prior mode’ with  $[\text{N II}]/[\text{O II}]$  and  $[\text{S II}] \lambda 6716/[\text{S II}] \lambda 6731$ , while the final run uses the ‘full-line likelihood mode’ on the 10 emission lines:  $[\text{O II}] \lambda 3726$ , 29,  $[\text{Ne III}] \lambda 3869$ ,  $\text{H}\beta$ ,  $[\text{O III}] \lambda 5007$ ,  $\text{He I} \lambda 5876$ ,  $[\text{O I}] \lambda 6300$ ,  $\text{H}\alpha$ ,  $[\text{N II}] \lambda 6583$ ,  $[\text{S II}] \lambda 6716$ , and  $[\text{S II}] \lambda 6731$ . We obtain the final posterior by multiplying the posteriors produced by these three runs together. Dust reddening is handled exactly as for BPT-SF spaxels.

The right panels of Fig. 1 show the metallicity maps we derive by this procedure for our two example galaxies. Both the two galaxies show a clear negative metallicity gradient. We apply this procedure to the full sample to generate a metallicity map for each galaxy.

### 2.3 Measuring the correlation functions

We next measure the two-point correlation functions for our metallicity maps following the procedure outlined by L21; as with Section 2.2, we only summarize here, and refer readers to the original paper for full details of the procedure. Our first step is to generate metallicity fluctuation maps from the metallicity maps output by NEBULABAYES. To do so, we deproject and rotate the metallicity maps to acquire a circular profile with the major axis of the original map along the  $x$  axis. The required transformation from the original





**Figure 2.** Left: the metallicity fluctuation maps of the same two example galaxies shown in Fig. 1. Right: the dots and the error bars (which are too small to be seen for most points) show the observed two-point correlation functions and the corresponding uncertainties, comparing with 20 MCMC-fit models (orange lines) randomly selected from the converged sampling chains. The dashed lines indicate the the median values of the posterior PDF of  $\sigma_{\text{beam}}$ . See Section 2.4 for details.

coordinate  $(x, y)$  of each spaxel to the new one  $(x', y')$  is given by

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta / \cos i & \cos \theta / \cos i \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad (1)$$

where  $\theta$  is the position angle adopted from the eDR3 catalogue,  $i$  is the inclination angle, determined by

$$\cos^2 i = \frac{(b/a)^2 - q_0^2}{1 - q_0^2}, \quad (2)$$

where the  $b/a$  is the axis ratio, and  $q_0 = 0.13$ . L23 show that adopting a larger  $q_0$  does not change the results.

We then calculate the metallicity fluctuation of the  $i$ th spaxel  $Z'_i$  by subtracting the mean metallicity  $\bar{Z}_r$ , calculated within an annular bin with the same distance  $r$  to the galactic center as the  $i$ th spaxel and with a bin width of 1 arcsec, which is the spaxel size of CALIFA datacubes. The left panels of Fig. 2 show the rotated and deprojected metallicity fluctuation maps of the same examples from Fig. 1.

The two-point correlation function of the metallicity fluctuation map is

$$\xi(\mathbf{r}) = \frac{\langle Z'(\mathbf{r} + \mathbf{r}')Z'(\mathbf{r}') \rangle}{\langle Z'(\mathbf{r}')^2 \rangle}, \quad (3)$$

where  $Z'(\mathbf{x})$  is the metallicity fluctuation we obtained above at position  $\mathbf{x}$  in the map, and the angle brackets  $\langle \cdot \rangle$  denote averaging over the dummy position variable  $\mathbf{r}'$ . In practice, we integrate out the direction of the vector  $\mathbf{r}$  since we are only interested in the

scalar separation  $r$  between two spaxels. We compute the two-point correlation function averaged over bins of separation  $r$  that are  $1''$  wide, equal to the spaxel size of CALIFA data cubes. We compute the two-point correlation function of the  $n$ th bin as

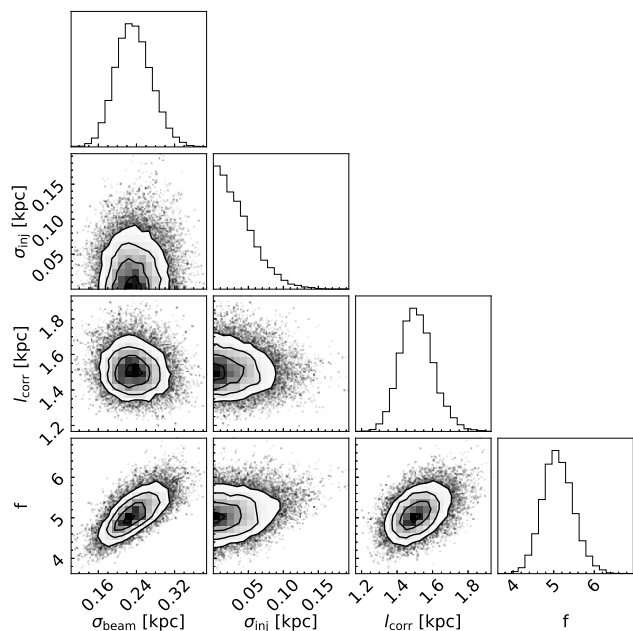
$$\xi_n = \frac{\sigma_{Z'}^{-2}}{N_n} \sum_{r_n < r_{ij} \leq r_{n+1}} Z'_i Z'_j, \quad (4)$$

where  $r_n = n \times (1 \text{ arcsec})$ , the sum runs over the  $N_n$  spaxel pairs  $(i, j)$  whose separations lie in the range  $r_n < r_{ij} \leq r_{n+1}$ , and

$$\sigma_{Z'}^2 = \frac{1}{N_p} \sum_{i=1}^{N_p} Z_i'^2 \quad (5)$$

is the sum of the variances of all  $N_p$  spaxels in the fluctuation map. We evaluate the two-point correlation to a maximum separation of  $15''$ , corresponding to 4 kpc at  $z \sim 0.015$ , the mean redshift of the CALIFA sample.

To estimate the uncertainty of the two-point correlation function, we randomly draw a value for the flux of each emission line in each pixel of the observed map from a Gaussian distribution with a mean and standard deviation set equal to the central estimate and estimated error for that spaxel in the observed map. This yields a new flux map. We mask spaxels in this map for which the flux is negative in the



**Figure 3.** The posterior distribution of four parameters describing the two-point correlation function for NGC0873, a BPT-SF example. The corresponding BPT map, BPT diagram, metallicity map, metallicity fluctuation map, and two-point correlation function are shown in Figs 1 and 2. The contours represent the 0.5, 1, 1.5, and 2 $\sigma$  confidence levels, i.e. the 1 $\sigma$  contour encloses 39.3 per cent of the total probability, etc.

H  $\alpha$ , H  $\beta$ , [O III], [N II], [O II], [S II]  $\lambda$ 6716, or [S II]  $\lambda$ 6731 lines.<sup>1</sup> We then run NEBULABAYES on the resulting synthetic map of line fluxes exactly as described in Section 2.2 to generate a synthetic metallicity map, and compute the two-point correlation of this metallicity map as described above. We repeat this procedure 20 times, yielding 20 two-point correlation functions. We take the standard deviation of these 20 realizations at each separation bin as our estimate of the uncertainty for that bin.

The dots in the right panels of Fig. 2 show the two-point correlation functions of the same example galaxies used in Fig. 1. Note that,  $\xi(0)$  is always unity, since every spaxel is perfectly correlated with itself. Error bars are shown, but for most points they are invisibly small because each bin represents a sum over a large number of spaxel pairs, and this summation averages down the final uncertainties.

## 2.4 Estimation of the correlation length

L21 and L23 find that the two-point correlation functions of non-AGN host galaxies in the CALIFA and AMUSING++ samples are well fit by the injection-diffusion model proposed by Krumholz & Ting (2018) with additional terms to account for the effects of beam smearing and observational noise. The functional form for this model

is (see Appendix D of L21 for details)

$$\xi_{\text{model}}(r) = \frac{2}{\ln \left( 1 + \frac{2l_{\text{corr}}^2}{\sigma_0^2} \right)} \left[ \frac{\Theta(r - \ell_{\text{pix}})}{f} + \Theta(\ell_{\text{pix}} - r) \right] \times \int_0^\infty e^{-\sigma_0^2 a^2 / 2} \left( 1 - e^{-2l_{\text{corr}}^2 a^2} \right) \frac{J_0(ar)}{a} da, \quad (6)$$

where  $\sigma_0^2 = \sigma_{\text{beam}}^2 + 2\sigma_{\text{inj}}^2 \cdot \sigma_{\text{beam}}$  is the observational beam size,  $\sigma_{\text{inj}}$  is the width over which supernovae inject metals,  $l_{\text{corr}}$  is the correlation length,  $f$  is the ratio of the variance of the metallicity fluctuation map including observational uncertainties to the real variance,  $\ell_{\text{pix}}$  is the pixel scale of the fluctuation map,  $J_0(x)$  is the modified Bessel function of the first kind of order 0, and  $\Theta(x)$  is the Heaviside step function. The terms involving  $\Theta(x)$  account for the fact that the errors within each spaxel of the map are perfectly correlated with themselves, but are totally uncorrelated with the errors in different spaxels.

The free parameters in this model are  $\sigma_{\text{beam}}$ ,  $\sigma_{\text{inj}}$ ,  $l_{\text{corr}}$ , and  $f$ , and we fit these to our measured correlation functions using MCMC sampling with the python package EMCEE (Foreman-Mackey et al. 2013). We set several priors for these four parameters. All parameters are required to be positive, so we set priors for non-positive values to zero. For  $\sigma_{\text{beam}}$ , we cross-match our sample with the point spread function (PSF) catalogue from Sánchez et al. (2016), in which they measure the PSF by comparing the reconstructed  $g$ -band images from CALIFA data cubes to SDSS images; this catalogue provides beam sizes for 96 of the 135 galaxies in our sample. Because we deproject the metallicity maps when generating the metallicity fluctuation maps in Section 2.3, the beam sizes in our science maps are stretched by a factor of  $1/\cos i$  along the galactic minor axis, where  $i$  is the inclination angle calculated from equation (2). For the 96 galaxies with PSF measurements  $\sigma_{\text{PSF}}$  in the Sánchez et al. catalogue, we assign a Gaussian prior with a mean of  $\sigma_{\text{PSF}}(1 + \cos i)/(2 \cos i)$ , the arithmetic mean of the major and minor axis PSF sizes, and a standard deviation of  $\sigma_{\text{PSF}}(1 - \cos i)/(2 \cos i)$ , which is half of the difference between the major and minor axis PSF sizes. This prior covers the plausible range of a circular beam size due to the beam-stretching because of the deprojection. For galaxies without a PSF measurement, we assign a Gaussian prior with a mean of  $1.06 \text{ arcsec} \times (1 + \cos i)/(2 \cos i)$  and a standard deviation of  $0.14 \text{ arcsec} \times (1 + \cos i)/(2 \cos i)$ , where 1.06 and 0.14 arcsec are the median and the 1 $\sigma$  scatter of the distribution of the CALIFA beam sizes. In Appendix B, we test the sensitivity of our results to this treatment of beam smearing by testing the alternative approaches of fixing  $\sigma_{\text{beam}}$  to either  $\sigma_{\text{PSF}}$  or  $\sigma_{\text{PSF}}/\cos i$ , i.e. the the largest and smallest plausible values. We find that doing so does not alter the qualitative results.

For  $f$ , we assign a flat distribution in logarithmic space up to a maximum  $f = 40$ . We set this upper limit to be large enough so that for values of  $f$  near the upper limit any real metallicity fluctuations are completely hidden in the noise and therefore cannot be estimated. A value of  $f$  near the upper limit effectively counts as a non-detection of any correlation in the metallicity fluctuation map. Finally, we adopt a flat prior on  $l_{\text{corr}}$  from zero to a physical distance corresponding to an angular size of 80 arcsec, the same as the bundle size of CALIFA. Values of  $l_{\text{corr}}$  larger than this are unmeasurable given CALIFA's field of view.

Following L23, for each galaxy, we use 100 walkers in the MCMC and run for 1000 steps. We treat the first 500 steps as the burn-in period for the walkers to converge, and only keep the last 500 steps. The posterior PDF of each parameter thus consists of 50 000 sampling points. Fig. 3 shows the posterior distribution of four

<sup>1</sup>For other weak lines with negative flux used for the Composite and Seyfert spaxels, we assign a flux of 5 per cent of the error. This assignment has little weight in our likelihood calculation, but satisfies the formal requirement that all fluxes input to NEBULABAYES be positive.

parameters describing the two-point correlation function for the BPT-SF example, NGC0873, shown in Figs 1 and 2. The values of  $\sigma_{\text{beam}}$ ,  $l_{\text{corr}}$ , and  $f$  are well constrained, but we are only able to obtain an upper limit on  $\sigma_{\text{inj}}$ . As discussed in L21, this is expected, since  $\sigma_{\text{inj}}$  is expected to be  $< 100$  pc, always smaller than the CALIFA beam size. Given that  $\sigma_{\text{beam}}$  and  $\sigma_{\text{inj}}$  are combined into  $\sigma_0$ , at our spatial resolution we only expect to be able to obtain an upper limit on  $\sigma_{\text{inj}}$ . In the right panels of Fig. 2, we randomly select 20 MCMC-fit models from the converged chains and plot them on top of the observed two-point correlation functions (orange lines). They fit the observed data quite well for both the BPT-SF and AGN-host galaxies. We also show the median values of the posterior PDF of  $\sigma_{\text{beam}}$  as dashed lines for reference. In what follows, we take the median value of the posterior PDF of the  $l_{\text{corr}}$  as our central estimate for each galaxy and the half of the 16th to 84th percentile range as the corresponding uncertainty.

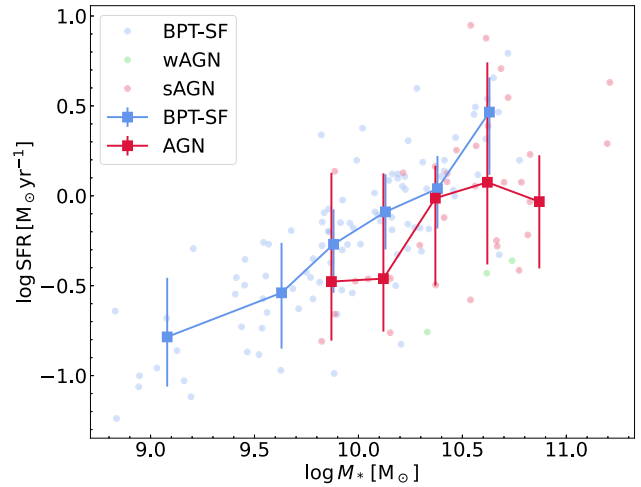
While this procedure yields well-behaved and narrowly peaked posterior PDFs for 132 of our sample of 135 galaxies, we find that three galaxies show long tails in their posterior distributions of  $l_{\text{corr}}$  and  $f$ . Examination of the two-point correlation functions for these galaxies shows that they are close to 0, with large errors, except for the bin at  $\xi(0)$  that is always unity. We discuss these galaxies further in Appendix C, where we show that our analysis fails for these galaxies because we measure essentially no meaningful correlation, and our fitting procedure cannot distinguish between two possible explanations: either these galaxies have  $l_{\text{corr}}$  significantly smaller than the beam size or observational errors so large that any real correlations are unmeasurable against the noise. Given the degeneracy of these two possibilities, and the fact that it affects only 2 per cent of our sample, we simply remove these three galaxies from our analysis, leaving a final sample consisting of 95 BPT-SF galaxies, 3 wAGN galaxies, and 34 sAGN galaxies with well-measured correlation lengths. Because we have so few wAGN in our sample, we merge the wAGN and the sAGN categories into a single AGN-host category in the following analysis.

## 2.5 Ancillary data: stellar mass and SFR

In addition to the correlation length, we require ancillary data on other galaxy properties, in particular, the stellar mass and star formation rate. We take these directly from the eDR3 catalogue. The stellar mass is estimated by fitting the integrated spectra with simple stellar population (SSP) templates, and the SFR is converted from the dust-corrected integrated  $H\alpha$  luminosity (Sánchez et al. 2023). The stellar mass and SFR in the eDR3 catalogue are derived assuming a Salpeter IMF, and we correct both of them by 0.25 dex to the Chabrier IMF that we use throughout this work (Bernardi et al. 2010).

Sánchez et al. (2023) point out that the conversion from the  $H\alpha$  luminosity to SFR in the catalogue does not account for potential contributions to the ionizing luminosity from sources other than massive stars, and thus might overestimate the true SFRs in galaxies with significant AGN contributions. However, we show in Appendix D that, while AGN contributions to individual pixels can be significant, contributions to the integrated ionizing luminosity of the full galaxy, and thus potential errors in the inferred SFR, are small. Consequently, making a correction to the SFR for a potential AGN contribution does not produce any qualitative difference to our results. We therefore omit such corrections for simplicity.

Fig. 4 shows SFR versus stellar mass for both the BPT-SF (blue) and AGN-host galaxies (red, combining wAGN and sAGN groups) in our sample. The squares show the median SFRs in each stellar mass bin. To estimate the  $1\sigma$  levels of the scatter, for each galaxy lying in one stellar mass bin, we randomly draw 1000 realizations



**Figure 4.** SFR versus stellar mass for BPT-SF galaxies (blue dots), wAGN galaxies (green dots), and sAGN galaxies (red dots). The squares are the median SFR in each stellar mass bin for BPT-SF galaxies (blue), and AGN-host galaxies (red, combining the wAGN and sAGN groups). The error bars show the  $1\sigma$  scatter estimated via the procedure described in Section 2.5. The stellar mass bins are the same for both BPT-SF and AGN-host galaxies, but the points have been offset slightly for readability.

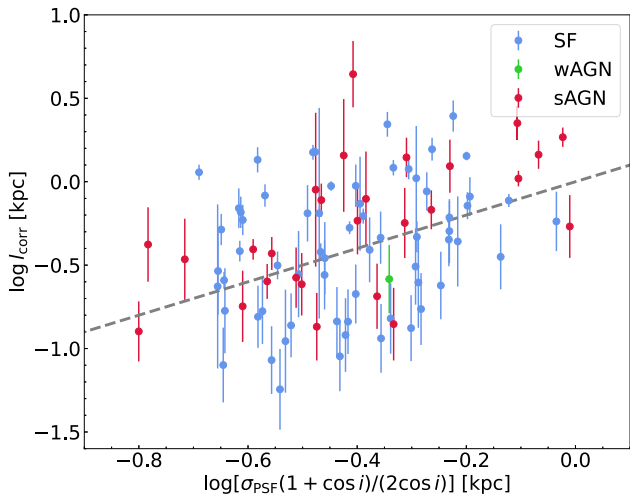
from a Gaussian distribution with a standard deviation equal to the stated uncertainty on the SFR from the eDR3 catalogue. We then take the 84th and the 16th percentiles over all realizations for all galaxies lying in this stellar mass bin as the upper and the lower uncertainties, respectively. It is clear that AGN-host galaxies generally have lower SFR than BPT-SF (non-AGN) galaxies at fixed stellar mass, consistent with earlier studies of optically selected AGN samples (e.g. Schawinski et al. 2014; Ellison et al. 2016; Lacerda et al. 2020).

## 3 RESULTS

Here, we first carry out two tests of the robustness of our results against possible errors and biases, and then present the main findings of this study.

### 3.1 Robustness tests

Before we examine how galaxy correlation lengths differ between AGN-host and non-AGN galaxies, and depend on other galactic properties in both populations, we first verify that our correlation lengths measurements are reliable and not subject to the influences of the spatial resolution of CALIFA. Following L21, Fig. 5 shows the estimated correlation lengths,  $l_{\text{corr}}$ , as a function of the beam size  $\sigma_{\text{PSF}}(1 + \cos i)/(2 \cos i)$  in physical scales (kpc) for the 96 galaxies with PSF measurements in CALIFA. The Pearson correlation coefficient for the BPT-SF galaxies is 0.27, similar to what found in L21 and indicating that our estimates of  $l_{\text{corr}}$  are at most weakly dependent on spatial resolution. One might be concerned that the measured  $l_{\text{corr}}$  are comparable to the beam size, but note that the two-point correlation function contains significant structure even at sizes of several times  $l_{\text{corr}}$  (cf. Fig. 1 of Krumholz & Ting 2018), and thus that observations can constrain  $l_{\text{corr}}$  even if the inner structure of the two-point correlation function on scales  $< l_{\text{corr}}$  is masked by beam smearing.



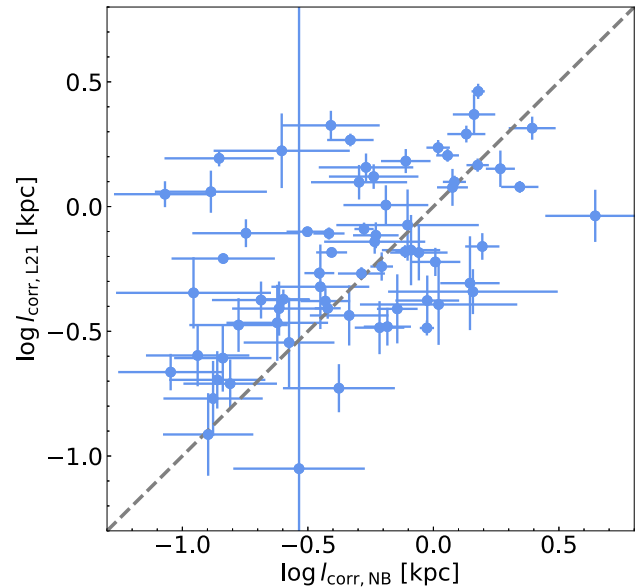
**Figure 5.** The relation between estimated correlation lengths and the beam size  $\sigma_{\text{PSF}}(1 + \cos i)/(2 \cos i)$  for 96 galaxies with PSF measurements in CALIFA, where  $\sigma_{\text{PSF}}$  is the beam size in physical scale (kpc), and  $i$  is the inclination angle calculated by equation (2). The dashed grey line shows the one-to-one relation.

While the BPT-SF sample shows only a weak correlation between beam size and  $l_{\text{corr}}$ , the Pearson correlation coefficient of AGN-host galaxies (sAGN and wAGN) is 0.55, showing a stronger correlation that is potentially of concern. Before concluding that this is a beam-smearing effect, however, we must consider an alternative explanation: due to CALIFA’s fixed bundle size, the survey selects galaxies of roughly fixed angular size, and as a result the galaxies it includes at larger distances tend to have larger physical radii and higher masses. Since L21 show that more massive galaxies in turn have larger  $l_{\text{corr}}$ , we would therefore expect galaxies with larger PSFs – which are on average further away and more massive – to have larger  $l_{\text{corr}}$  even in the absence of beam effects.

To understand whether the correlation we find in Fig. 5 between beam size and  $l_{\text{corr}}$  is due to this innocuous effect or to beam smearing, we carry out a partial-correlation test between stellar mass, beam size, and  $l_{\text{corr}}$  for AGN-host galaxies. When controlling for beam size, we find that the partial correlation coefficient between stellar mass and  $l_{\text{corr}}$  is 0.46 with a  $p$ -value of 0.02; by contrast, when controlling for stellar mass the partial correlation coefficient between beam size and  $l_{\text{corr}}$  is 0.22 with a  $p$ -value of 0.26, strongly suggesting that the correlation between  $l_{\text{corr}}$  and beam size in the AGN sample is a byproduct of CALIFA’s sample selection, not an indication of beam-smearing effects.

For completeness, we also repeat the partial-correlation test for BPT-SF galaxies. When controlling for beam size in this sample the partial correlation coefficient between stellar mass and  $l_{\text{corr}}$  is 0.26 with a  $p$ -value of 0.04, while the partial correlation coefficient between beam size and  $l_{\text{corr}}$  controlling for stellar mass is 0.12 with a  $p$ -value of 0.33. This suggests that the minor correlation between  $l_{\text{corr}}$  and beam size that we find in the BPT-SF sample is also primarily due to the correlation between  $l_{\text{corr}}$  and stellar mass interacting with CALIFA’s selection function, and not to beam-smearing.

Having shown that our results of  $l_{\text{corr}}$  are not strongly influenced by CALIFA’s resolution, we next verify that our correlation lengths are in reasonable agreement with those derived by earlier authors. Fig. 6 compares the  $l_{\text{corr}}$  values we derive here with those obtained by L21 for the 68 galaxies common to both samples (which in



**Figure 6.** Comparison between the correlation lengths calculated by L21 from metallicity maps produced using the  $[\text{N II}]/[\text{O II}]$  metallicity diagnostic from Kewley et al. (2019b) (vertical axis) and the correlation lengths estimated in this work using NEBULABAYES (horizontal axis). Each point corresponds to one of the 68 galaxies in common to both samples; points show median values, while error bars show the 16th to 84th percentile range. The dashed grey line indicates the 1 to 1 relation.

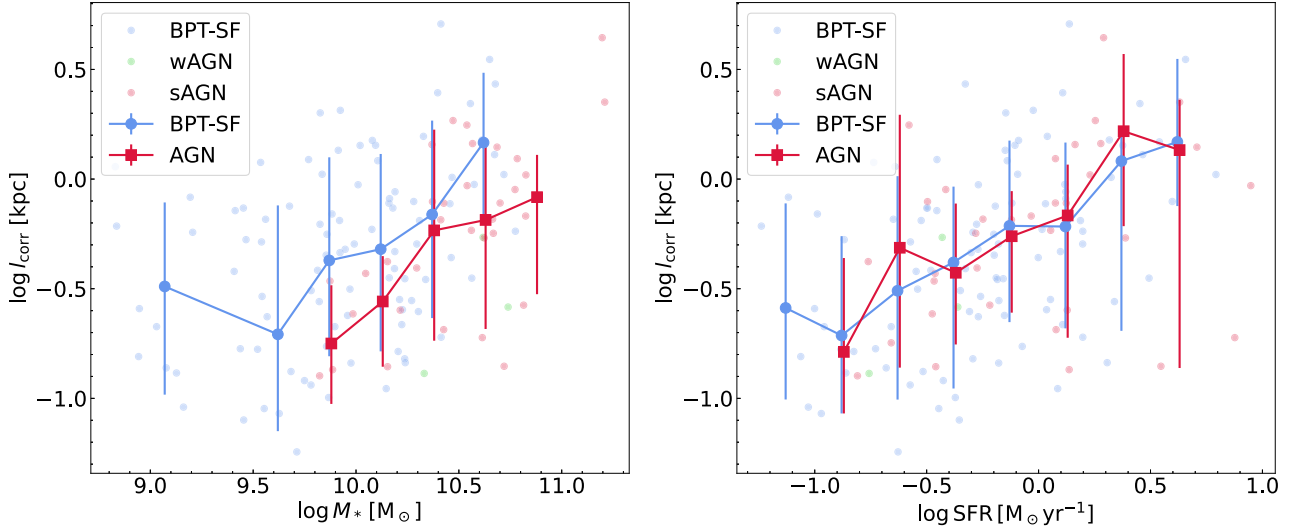
the case of L21 includes only BPT-SF galaxies, since they were required to exclude AGN). L21 obtained their metallicities using the  $[\text{N II}]/[\text{O II}]$  metallicity diagnostic from Kewley et al. (2019b). As the figure shows, the results from the two studies generally follow the 1 to 1 relation with a  $1\sigma$  scatter of  $\approx 0.34$  dex; the Pearson correlation coefficient between the two data sets is 0.49. L21 compared the correlation lengths of maps derived using three different line-ratio metallicity diagnostics and found similar scatter and Pearson correlations when comparing them. We thus confirm that our  $l_{\text{corr}}$  estimates are consistent with those obtained by L21 to the extent that we expect given that we are using different metallicity diagnostics.

### 3.2 Correlation lengths in AGN-host versus BPT-SF galaxies

Fig. 7 shows  $l_{\text{corr}}$  versus stellar mass (left panel) and SFR (right panel) for both BPT-SF (non-AGN, blue) and AGN-host (combining the wAGN and sAGN groups, red) galaxies. Points show the median over all MCMC samples for  $l_{\text{corr}}$  (excluding the burn-in period) for all galaxies falling into a given bin of stellar mass or SFR, while error bars indicate the 16th to 84th percentile range over these samples; for both stellar mass and SFR, bins are 0.25 dex wide. For both BPT-SF and AGN-host galaxies, we find  $l_{\text{corr}}$  increases with both stellar mass and SFR. These positive correlations are similar to those found in L21 and L23.

However, while BPT-SF and AGN-host galaxies show trends in the same direction, we also find that there is a systematic offset between the two galaxy groups at fixed stellar mass – AGN-host galaxies have systematically lower  $l_{\text{corr}}$  than BPT-SF galaxies of similar stellar mass. There is no corresponding systematic difference at fixed SFR. To verify that this visual impression is statistically robust, we carry out  $t$ -tests comparing the distributions of  $l_{\text{corr}}$  values for BPT-SF and AGN-host galaxies over the four bins of stellar mass where the two





**Figure 7.** Left:  $\log l_{\text{corr}}$  versus  $\log M_*$  for BPT-SF (blue), wAGN (green), and sAGN (red) galaxies. The blue circles with error bars represent the median and 16th to 84th percentile range of  $l_{\text{corr}}$  for BPT-SF galaxies in each stellar mass bin, computed via the procedure described in Section 3; red squares with error bars show the same quantities for AGN-host galaxies, combining the wAGN and sAGN groups. The stellar mass bins are the same for both BPT-SF and AGN-host galaxies, but the points have been offset slightly for readability. Right: same as the left panel, but showing  $\log l_{\text{corr}}$  versus  $\log \text{SFR}$  rather than versus  $\log M_*$ .

samples overlap (the left panel of Fig. 7). The null hypothesis of this test is that the mean  $l_{\text{corr}}$  for AGN-host galaxies is greater than or equal to that of the BPT-SF galaxies in the same mass bin. The  $p$ -values returned by this test are, from lowest to highest stellar mass bin, 0.020, 0.111, 0.327, and 0.007. If we take  $p = 0.05$  to mark significance then the results rule out the null hypothesis in two of the four bins.<sup>2</sup> Repeating this exercise for the two galaxy classes binned by SFR with the null hypothesis that the mean  $l_{\text{corr}}$  for AGN-host galaxies is equal to that of BPT-SF galaxies in the same SFR bin (right panel of Fig. 7), on the other hand, yields  $p$ -values of 0.80, 0.43, 0.85, 0.76, 0.82, 0.22, and 0.48 from the lowest to the highest  $\log \text{SFR}$  bin, consistent with there being no statistically significant difference in the distributions.

Comparing Fig. 7 to 4 yields the interesting conclusion that the offset between BPT-SF and AGN-host galaxies in the  $l_{\text{corr}}$  versus  $M_*$  relation is qualitatively very similar to the offset between the two samples in SFR versus  $M_*$ . That is, at fixed stellar mass, galaxies that host AGNs have both systematically smaller SFRs and metallicity  $l_{\text{corr}}$ , and the direction and size of the offset are similar in both diagrams. Consequently, there is no offset between the two galaxy classes at fixed SFR, and thus we can attribute the full difference in  $l_{\text{corr}}$  values between BPT-SF and AGN-host galaxies at fixed stellar mass to differences in SFR.

## 4 DISCUSSION

In this section, we explore the implications of our results. We first discuss the implications of our findings for how AGN affect galaxy properties in Section 4.1, and then discuss the implications for what processes regulate galactic metallicity distributions in Section 4.2.

<sup>2</sup>Note that the apparent difference between the BPT-SF and AGN-host samples in the bin  $\log(M_*/M_\odot) \in [10.25, 10.5]$  does not reach statistical significance because there are only four AGN galaxies in this stellar mass bin.

### 4.1 Implications for how AGN influence galaxy properties

We have found that AGN-host galaxies have smaller  $l_{\text{corr}}$  than BPT-SF galaxies (non-AGN galaxies) at fixed stellar mass (the left panel of Fig. 7), but that this result is fully accounted for by the smaller SFR of AGN hosts at fixed stellar mass (Fig. 4), such that there is no difference in  $l_{\text{corr}}$  between BPT-SF and AGN-host galaxies of similar SFR (right panel of Fig. 7). This suggests that AGN activity does not directly influence galactic metallicity distributions, but instead only affects them indirectly by suppressing star formation or at least by correlating with low SFR.

At first, this result is surprising. In the context of the Krumholz & Ting (2018) model, the correlation length is expected to be  $l_{\text{corr}} = \sqrt{\kappa t_*}$ , where  $\kappa$  is a diffusion coefficient describing diffusion of metals in the ISM and  $t_*$  is a characteristic time-scale over which star formation has taken place. The diffusion coefficient in turn should be related to the velocity dispersion  $\sigma_g$  and scale height  $h$  of the ISM as  $\kappa \approx h\sigma_g/3$  (Karlsson, Bromm & Bland-Hawthorn 2013, L23). Any physical processes that enhance the strength of the turbulence should produce an increase in  $\sigma_g$ , and thus an increase in the  $l_{\text{corr}}$ . Powerful AGN like quasars are able to trigger strong outflows and shocks and produce large velocity dispersions (e.g. Veilleux, Cecil & Bland-Hawthorn 2005; Rich, Kewley & Dopita 2015; Davies et al. 2017; Oh et al. 2022), and thus, we might expect AGN activity to increase  $l_{\text{corr}}$ . Moreover, AGN activity has long been thought to be correlated with galactic mergers (e.g. Hopkins et al. 2008; Koss et al. 2018; Comerford et al. 2024), and merging systems are also observed to have large  $l_{\text{corr}}$  (L21, L23). Considering these effects, one might expect that AGN-host galaxies possess larger  $l_{\text{corr}}$  than non-AGN galaxies. However, as discussed above, we find the opposite.

We suspect this is because in local universe, most AGN are not powerful enough to trigger strong galactic-scale winds or shocks. As the example sAGN shown in Fig. 1 demonstrates, even sAGN in our sample have AGN excitation restricted to the central regions of the galaxy, while the bulk of the disc is still ionised by massive stars. Quantitatively, in Appendix D we show that even in AGN-host galaxies the AGN typically provides only  $\sim 1$  per cent of the

total ionization. López-Cobá et al. (2019, 2020) find that in both the CALIFA and AMUSING++ samples only a small fraction of AGN hosts harbour detectable outflows. Together all this evidence suggests that most AGN in the local Universe are not capable of substantially altering the bulk properties of the ISM. Similarly, while mergers may be important triggers of quasars and similar very bright AGN activity, external perturbations may contribute only marginally to the sorts of low-luminosity AGN present in our sample; these may instead be triggered predominantly by local secular processes occurring in the galactic nucleus, with no substantial contribution from the types of mergers and interactions that can raise  $l_{\text{corr}}$  (Orban de Xivry et al. 2011; Kocevski et al. 2012).

To the extent that this hypothesis is correct and our AGN are too dim to affect ISM properties directly via shocks or to be correlated with mergers, an alternative effect arising from AGN can be dominant. This is that optically selected AGN have lower SFR than galaxies on the star-forming main sequence (Schawinski et al. 2014; Ellison et al. 2016; Lacerda et al. 2020). This correlation in turn can arise through two possible channels. One is negative AGN feedback that takes the form not of violent ejection of mass from the ISM, which presumably would be sufficient to affect metallicity distributions directly (e.g. Hopkins et al. 2008; Hopkins & Elvis 2010; Bing et al. 2019), but via a more gentle path of preventing ongoing gas accretion, namely ‘starvation’ (Kumari et al. 2021). The other is that AGN activity is associated with bulge growth (Heckman et al. 2004; Kormendy & Ho 2013; Sánchez et al. 2018), which inhibits star formation (so-called morphological quenching) due to increased shear stabilizing the gaseous components of galactic discs (Martig et al. 2009; Fang et al. 2013). In either of these scenarios, AGN would affect galaxy metallicity distributions only indirectly, via the resulting reduction in the SFR, consistent with what we see in the observations.

Our finding that the offset between AGN-hosts and SF galaxies at fixed stellar mass is due to differences in SFR is also consistent with the results of L24, who find that AGN hosts generally have higher metallicity than BPT-SF galaxies at fixed stellar mass, but that this difference is a natural result of the existence of the FZR, whereby the SFR anti-correlates with the metallicity at fixed stellar mass. Since AGN hosts have smaller SFRs at fixed stellar mass, the FZR implies that they will have larger metallicities, an indirect effect directly analogous to the one we have found here. Taken together, the two results suggest that, at least for low-luminosity AGN, the primary way that AGN affect the gas-phase metallicities of galaxies – both the total metallicity and the metallicity distribution – is that hosting an AGN is associated with galaxies having lower SFRs, rather than via any direct effects of the AGN itself.

#### 4.2 Which galactic parameter is most important to regulating metallicity distributions?

Both L21 and L23 find that galaxies’  $l_{\text{corr}}$  is positively correlated with their stellar mass, SFR, and size. However, because each of these quantities correlates with the other two, they are not able to determine which of these three parameters is the fundamental one when it comes to setting  $l_{\text{corr}}$ , and which are secondary correlations. Our current sample offers the opportunity to break this degeneracy: the lower  $l_{\text{corr}}$  of AGN hosts compared to non-AGN at fixed stellar mass but similar  $l_{\text{corr}}$  at fixed SFR strongly suggests that SFR is more fundamental than stellar mass when it comes to regulating  $l_{\text{corr}}$ , and thus galaxies’ internal metallicity distributions more broadly. To quantify this intuition, we carry out a partial-correlation test among  $l_{\text{corr}}$ , stellar mass, and SFR for BPT-SF galaxies with the

PINGOUINPYTHON package. When controlling the stellar mass, the partial correlation coefficient between  $l_{\text{corr}}$  and SFR is 0.21 with a  $p$ -value of 0.04; by contrast, controlling for SFR the partial correlation coefficient between  $l_{\text{corr}}$  and stellar mass is 0.08 with the  $p$ -value of 0.42. This test is consistent with the result suggested by the comparison of the BPT-SF and AGN-host samples, which is that SFR is the more important parameter. This is also consistent with theoretical expectation: there is a well-known and theoretically expected correlation between galaxies’ SFR and velocity dispersions Krumholz et al. (2018, and references therein), and to the extent that larger velocity dispersion yields larger diffusion coefficients, this provides a natural pathway by which an increase in SFR could increase  $l_{\text{corr}}$ .

As in Section 4.1, we can make a useful connection to the results of L24 for total metallicity. The gas-phase metallicity primarily records the cumulative effects of metal enrichment via stellar nucleosynthesis in a galaxy through cosmic time (Vale Asari et al. 2009). If we take stellar mass as representing an evolutionary sequence, the mass-metallicity relation then reflects the cumulative buildup of metals over cosmological time-scales (e.g. Maiolino et al. 2008; Zahid, Kewley & Bresolin 2011). However, on top of this secular trend, the metallicity also responds to processes like star formation that can fluctuate on much shorter time-scales, and the anti-correlations between SFR (or  $\Sigma_{\text{SFR}}$ ) and metallicity at fixed stellar mass (or  $\Sigma_*$ ) are the signature of this response to short timescale phenomena (e.g. Lara-López et al. 2010; Mannucci et al. 2010; Baker et al. 2023). In this context, our findings here present another example of this short time-scale response, this time in the spatial structure of fluctuations in metallicity fields rather than in the mean metallicity.

## 5 CONCLUSION

In this work, we measure the two-point correlation functions of the metallicity fields of 95 galaxies that do not contain optical AGN and 37 AGN-host galaxies drawn from CALIFA IFS survey. The main methodological innovation in our work is that we measure metallicities using a Bayesian method (NEBULABAYES) that self-consistently accounts for the AGN contribution to emission-line fluxes spaxel-by-spaxel, which allows us to derive metallicities for the AGN-host and non-AGN galaxies using the same method and without incurring the large systematic uncertainties that are inevitable when comparing metallicities derived using different diagnostics. This in turn makes a direct comparison between AGN-host and non-host galaxies possible, allowing us to achieve our primary science aim: understanding how AGN influence galactic metallicity distributions.

We find that the two-point correlation functions of both AGN host and non-host galaxies are well fit by the simple metal injection and diffusion model proposed by Krumholz & Ting (2018). The key fitting parameter we can derive from this model is correlation length  $l_{\text{corr}}$ , which describes the characteristic length scale over which metallicity fluctuations are correlated. Our results for non-AGN host galaxies show reasonable consistency with earlier results from Li et al. (2021) derived from metallicity maps obtained from traditional emission-line flux ratio diagnostics, validating our new Bayesian methods.

Comparing  $l_{\text{corr}}$  between non-AGN and AGN-host galaxies, we find that AGN hosts generally have smaller  $l_{\text{corr}}$  than non-host galaxies at fixed stellar mass (the left panel of Fig. 7), but similar  $l_{\text{corr}}$  at fixed SFR (the right panel of Fig. 7). We show that the difference in  $l_{\text{corr}}$  at fixed stellar mass is fully explained by the fact that AGN-host galaxies have reduced SFR at fixed stellar mass

compared to non-hosts (Fig. 4). This finding suggests that AGN, at least over the luminosity range we are able to probe, influence the metallicity distributions within galaxies only indirectly, either by suppressing star formation through starvation or because AGN growth is correlated with bulge growth, which suppresses star formation through morphological quenching. We find no evidence for other, more direct effects. The fact that the two galaxy-subgroups have similar  $l_{\text{corr}}$  at fixed SFR, but not at fixed stellar mass, further suggests that SFR is more fundamental than stellar mass when it comes to regulating galaxy metallicity distributions. Partial-correlation tests among the  $l_{\text{corr}}$ , stellar mass, and SFR confirm this impression. This suggests that, while galaxy mean metallicities are the result of accumulation over the full star formation history of galaxies, fluctuations in the metallicity distribution within galaxies are more driven by short-term responses to physical processes such as star formation that can change much faster than a Hubble time.

## ACKNOWLEDGEMENTS

We thank the referee Prof. Roberto Maiolino for a detailed report that helped significantly in improving the presentation of our work. This study uses data provided by the Calar Alto Legacy Integral Field Area (CALIFA) survey (<http://califa.caha.es/>). This study is based on observations collected at the Centro Astronómico Hispano Alemán (CAHA) at Calar Alto, operated jointly by the Max-Planck-Institut für Astronomie and the Instituto de Astrofísica de Andalucía (CSIC). This study also used computational resources provided by the National Computational Infrastructure, which is supported by the Australian Government, through award jh2. MRK acknowledges support from the Australian Research Council through its Laureate Fellowship funding scheme, award FL220100020. SFS thanks the PAPIIT-DGAPA AG100622 project and CONACYT grant CF19-39578. We thank Trevor Mendel for comments and ideas for this work.

## DATA AVAILABILITY

The data underlying this article will be shared on reasonable request with the corresponding author.

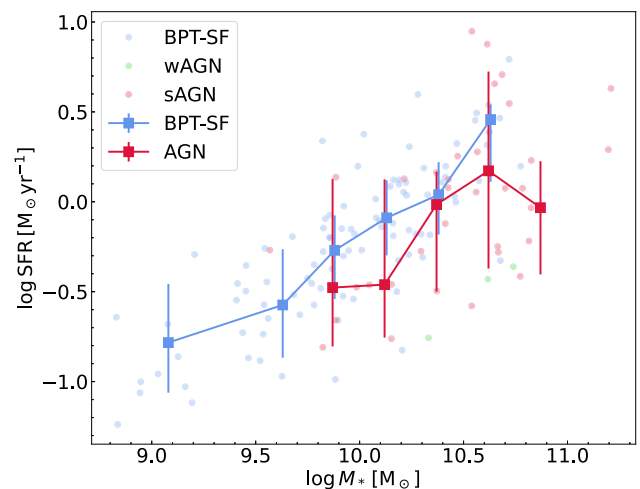
## REFERENCES

- Andrews B. H., Martini P., 2013, *ApJ*, 765, 140  
 Baker W. M. et al., 2023, *MNRAS*, 519, 1149  
 Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, 93, 5  
 Barrera-Ballesteros J. K. et al., 2016, *MNRAS*, 463, 2513  
 Barrera-Ballesteros J. K., Sánchez S. F., Heckman T., Blanc G. A., MaNGA Team, 2017, *ApJ*, 844, 80  
 Belfiore F. et al., 2016, *MNRAS*, 461, 3111  
 Belfiore F. et al., 2017, *MNRAS*, 469, 151  
 Bernardi M., Shankar F., Hyde J. B., Mei S., Marulli F., Sheth R. K., 2010, *MNRAS*, 404, 2087  
 Bing L. et al., 2019, *MNRAS*, 482, 194  
 Bothwell M. S., Maiolino R., Kennicutt R., Cresci G., Mannucci F., Marconi A., Ciccone C., 2013, *MNRAS*, 433, 1425  
 Bothwell M. S., Maiolino R., Ciccone C., Peng Y., Wagg J., 2016, *A&A*, 595, A48  
 Brown T., Cortese L., Catinella B., Kilborn V., 2018, *MNRAS*, 473, 1868  
 Bryant J. J. et al., 2015, *MNRAS*, 447, 2857  
 Bundy K. et al., 2015, *ApJ*, 798, 7  
 Carvalho S. P. et al., 2020, *MNRAS*, 492, 5675  
 Chabrier G., 2003, *PASP*, 115, 763  
 Cid Fernandes R., Stasińska G., Mateus A., Vale Asari N., 2011, *MNRAS*, 413, 1687  
 Comerford J. M. et al., 2024, *ApJ*, 963, 53  
 Davies R. L. et al., 2017, *MNRAS*, 470, 4974  
 Dors O. L., 2021, *MNRAS*, 507, 466  
 Dors O. L. et al., 2020, *MNRAS*, 492, 468  
 Ellison S. L., Teimoorinia H., Rosario D. J., Mendel J. T., 2016, *MNRAS*, 458, L34  
 Erb D. K., Shapley A. E., Pettini M., Steidel C. C., Reddy N. A., Adelberger K. L., 2006, *ApJ*, 644, 813  
 Fang J. J., Faber S. M., Koo D. C., Dekel A., 2013, *ApJ*, 776, 63  
 Fischer T. C., Crenshaw D. M., Kraemer S. B., Schmitt H. R., Storch-Bergmann T., Riffel R. A., 2015, *ApJ*, 799, 234  
 Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306  
 Harrison C. M., Alexander D. M., Mullaney J. R., Swinbank A. M., 2014, *MNRAS*, 441, 3306  
 He X. et al., 2024, *ApJ*, 960, L13  
 Heckman T. M., Kauffmann G., Brinchmann J., Charlot S., Tremonti C., White S. D. M., 2004, *ApJ*, 613, 109  
 Hopkins P. F., Elvis M., 2010, *MNRAS*, 401, 7  
 Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, *ApJS*, 175, 356  
 Karlsson T., Bromm V., Bland-Hawthorn J., 2013, *Rev. Mod. Phys.*, 85, 809  
 Kauffmann G. et al., 2003, *MNRAS*, 346, 1055  
 Kelz A. et al., 2006, *PASP*, 118, 129  
 Kewley L. J., Dopita M. A., 2002, *ApJS*, 142, 35  
 Kewley L. J., Ellison S. L., 2008, *ApJ*, 681, 1183  
 Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, *ApJ*, 556, 121  
 Kewley L. J., Nicholls D. C., Sutherland R. S., 2019a, *ARA&A*, 57, 511  
 Kewley L. J., Nicholls D. C., Sutherland R., Rigby J. R., Acharya A., Dopita M. A., Bayliss M. B., 2019b, *ApJ*, 880, 16  
 King A., Pounds K., 2015, *ARA&A*, 53, 115  
 Kobulnicky H. A., Kewley L. J., 2004, *ApJ*, 617, 240  
 Kocevski D. D. et al., 2012, *ApJ*, 744, 148  
 Komatsu E. et al., 2011, *ApJS*, 192, 18  
 Kormendy J., Ho L. C., 2013, *ARA&A*, 51, 511  
 Koss M. J. et al., 2018, *Nature*, 563, 214  
 Kreckel K. et al., 2019, *ApJ*, 887, 80  
 Kreckel K. et al., 2020, *MNRAS*, 499, 193  
 Krumholz M. R., Ting Y.-S., 2018, *MNRAS*, 475, 2236  
 Krumholz M. R., Burkhardt B., Forbes J. C., Crocker R. M., 2018, *MNRAS*, 477, 2716  
 Kumari N., Maiolino R., Trussler J., Mannucci F., Cresci G., Curti M., Marconi A., Belfiore F., 2021, *A&A*, 656, A140  
 Lacerda E. A. D. et al., 2018, *MNRAS*, 474, 3727  
 Lacerda E. A. D., Sánchez S. F., Cid Fernandes R., López-Cobá C., Espinosa-Ponce C., Galbany L., 2020, *MNRAS*, 492, 3073  
 Lacerda E. A. D., Sánchez S. F., Mejía-Narváez A., Camps-Fariña A., Espinosa-Ponce C., Barrera-Ballesteros J. K., Ibarra-Medel H., Lugo-Aranda A. Z., 2022, *New A*, 97, 101895  
 Lara-López M. A. et al., 2010, *A&A*, 521, L53  
 Li Z., Krumholz M. R., Wisnioski E., Mendel J. T., Kewley L. J., Sánchez S. F., Galbany L., 2021, *MNRAS*, 504, 5496  
 Li Z. et al., 2023a, *MNRAS*, 518, 286  
 Li M. et al., 2023b, *ApJ*, 955, L18  
 Li S.-L., Grasha K., Krumholz M. R., Wisnioski E., Sutherland R. S., Kewley L. J., Chen Y.-M., Li Z., 2024, *MNRAS*, 529, 4993  
 López-Cobá C., Sánchez S. F., Bland-Hawthorn J., Moiseev A. V., Cruz-González I., García-Benito R., Barrera-Ballesteros J. K., Galbany L., 2019, *MNRAS*, 482, 4032  
 López-Cobá C. et al., 2020, *AJ*, 159, 167  
 Maier C., Lilly S. J., Ziegler B. L., Contini T., Pérez Montero E., Peng Y., Balestra I., 2014, *ApJ*, 792, 3  
 Maiolino R., Mannucci F., 2019, *A&A Rev.*, 27, 3  
 Maiolino R. et al., 2008, *A&A*, 488, 463  
 Mannucci F., Cresci G., Maiolino R., Marconi A., Gnerucci A., 2010, *MNRAS*, 408, 2115  
 Martig M., Bournaud F., Teyssier R., Dekel A., 2009, *ApJ*, 707, 250

- Metha B., Trenti M., Chu T., 2021, *MNRAS*, 508, 489
- Metha B., Trenti M., Chu T., Battisti A., 2022, *MNRAS*, 514, 4465
- Mott A., Spitoni E., Matteucci F., 2013, *MNRAS*, 435, 2918
- Nakajima K., Ouchi M., Isobe Y., Harikane Y., Zhang Y., Ono Y., Umeda H., Oguri M., 2023, *ApJS*, 269, 33
- Nicholls D. C., Sutherland R. S., Dopita M. A., Kewley L. J., Groves B. A., 2017, *MNRAS*, 466, 4403
- Oh S. et al., 2022, *MNRAS*, 512, 1765
- Orban de Xivry G., Davies R., Schartmann M., Komossa S., Marconi A., Hicks E., Engel H., Tacconi L., 2011, *MNRAS*, 417, 2721
- Pérez-Montero E. et al., 2013, *A&A*, 549, A25
- Poetrodjojo H. et al., 2018, *MNRAS*, 479, 5235
- Rich J. A., Kewley L. J., Dopita M. A., 2015, *ApJS*, 221, 28
- Rosales-Ortega F. F., Sánchez S. F., Iglesias-Páramo J., Díaz A. I., Vílchez J. M., Bland-Hawthorn J., Husemann B., Mast D., 2012, *ApJ*, 756, L31
- Roth M. M. et al., 2005, *PASP*, 117, 620
- Salim S., Lee J. C., Ly C., Brinchmann J., Davé R., Dickinson M., Salzer J. J., Charlot S., 2014, *ApJ*, 797, 126
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Sánchez-Menguiano L., Sánchez S. F., Pérez I., Ruiz-Lara T., Galbany L., Anderson J. P., Kuncarayakti H., 2020, *MNRAS*, 492, 4149
- Sánchez S. F. et al., 2012, *A&A*, 538, A8
- Sánchez S. F. et al., 2013, *A&A*, 554, A58
- Sánchez S. F. et al., 2014, *A&A*, 563, A49
- Sánchez S. F. et al., 2016, *A&A*, 594, A36
- Sánchez S. F. et al., 2017, *MNRAS*, 469, 2121
- Sánchez S. F. et al., 2018, *RMxAA*, 54, 217
- Sánchez S. F. et al., 2019, *MNRAS*, 484, 3042
- Sánchez S. F., Galbany L., Walcher C. J., García-Benito R., Barrera-Ballesteros J. K., 2023, *MNRAS*, 526, 5555
- Schawinski K. et al., 2014, *MNRAS*, 440, 889
- Sharda P., Krumholz M. R., Wisnioski E., Forbes J. C., Federrath C., Acharyya A., 2021, *MNRAS*, 502, 5935
- Sharda P., Ginzburg O., Krumholz M. R., Forbes J. C., Wisnioski E., Mingozi M., Zovaro H. R. M., Dekel A., 2024, *MNRAS*, 528, 2232
- Sutherland R. S., Dopita M. A., 2017, *ApJS*, 229, 34
- Teklu B. B., Gao Y., Kong X., Lin Z., Liang Z., 2020, *ApJ*, 897, 61
- Thomas A. D., Groves B. A., Sutherland R. S., Dopita M. A., Kewley L. J., Jin C., 2016, *ApJ*, 833, 266
- Thomas A. D., Dopita M. A., Kewley L. J., Groves B. A., Sutherland R. S., Hopkins A. M., Blanc G. A., 2018, *ApJ*, 856, 89
- Thomas A. D., Kewley L. J., Dopita M. A., Groves B. A., Hopkins A. M., Sutherland R. S., 2019, *ApJ*, 874, 100
- Tozzi G., Maiolino R., Cresci G., Piotrowska J. M., Belfiore F., Curti M., Mannucci F., Marconi A., 2023, *MNRAS*, 521, 1264
- Tremonti C. A. et al., 2004, *ApJ*, 613, 898
- Vale Asari N., Stasińska G., Cid Fernandes R., Gomes J. M., Schlickmann M., Mateus A., Schoenell W., 2009, *MNRAS*, 396, L71
- Veilleux S., Cecil G., Bland-Hawthorn J., 2005, *ARA&A*, 43, 769
- Walcher C. J. et al., 2014, *A&A*, 569, A1
- Zahid H. J., Kewley L. J., Bresolin F., 2011, *ApJ*, 730, 137
- Zahid H. J., Dima G. I., Kudritzki R.-P., Kewley L. J., Geller M. J., Hwang H. S., Silverman J. D., Kashino D., 2014, *ApJ*, 791, 130

## APPENDIX A: POTENTIAL BIASES DUE TO AGN-HOST IDENTIFICATION METHOD

As discussed in Section 2.1, we introduce a potential bias by selecting our AGN-host sample based on the BPT diagram, since this may selectively miss AGN-hosts that have higher SFRs and thus stellar-driven line emission bright enough to mask weak AGN signatures.



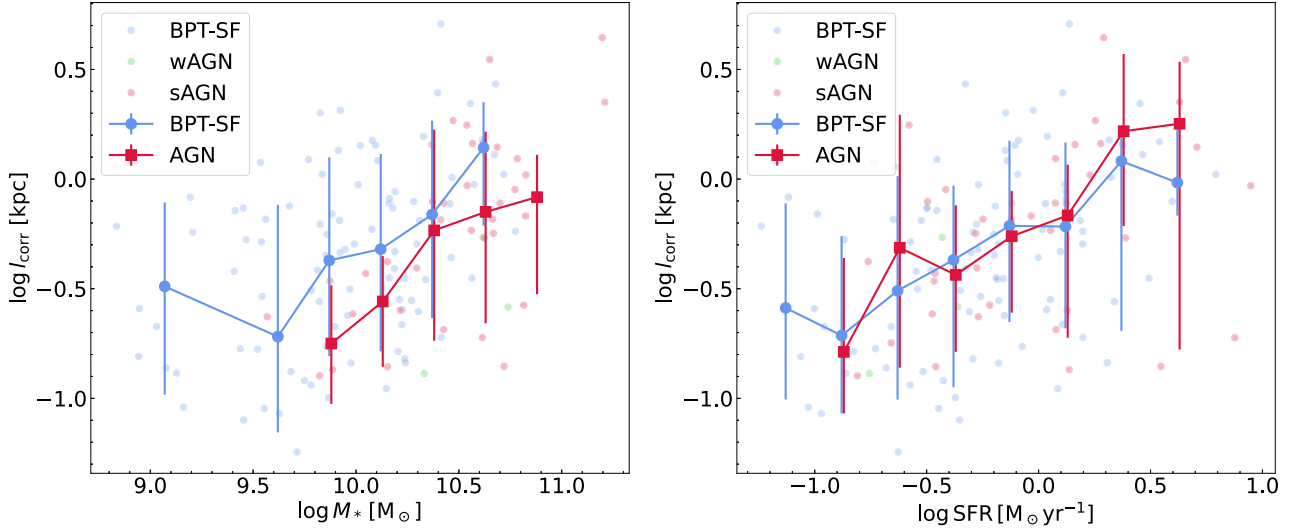
**Figure A1.** Same as Fig. 4, but with two galaxies re-classified from BPT-SF to sAGN based on the He II diagram as discussed in Appendix A.

Tozzi et al. (2023) explore this possibility using the MaNGA sample; they compare AGN identified using the traditional BPT diagram to those identified based on a modified version of the BPT diagram with  $[\text{O III}]/\text{H}\beta$  replaced by  $\text{He II } \lambda 4686/\text{H}\beta$ . The He II line is more sensitive to AGN and much less likely to be masked by bright stellar-driven nebular emission due to its higher excitation energy (54.4 eV versus 35.2 eV for  $[\text{O III}]$ ). Tozzi et al. find that 6 per cent of AGN-host galaxies can be identified only using the He II line, and that these additional AGN-host galaxies all lie on the star-forming main sequence and have SFRs similar to those of BPT-SF galaxies.

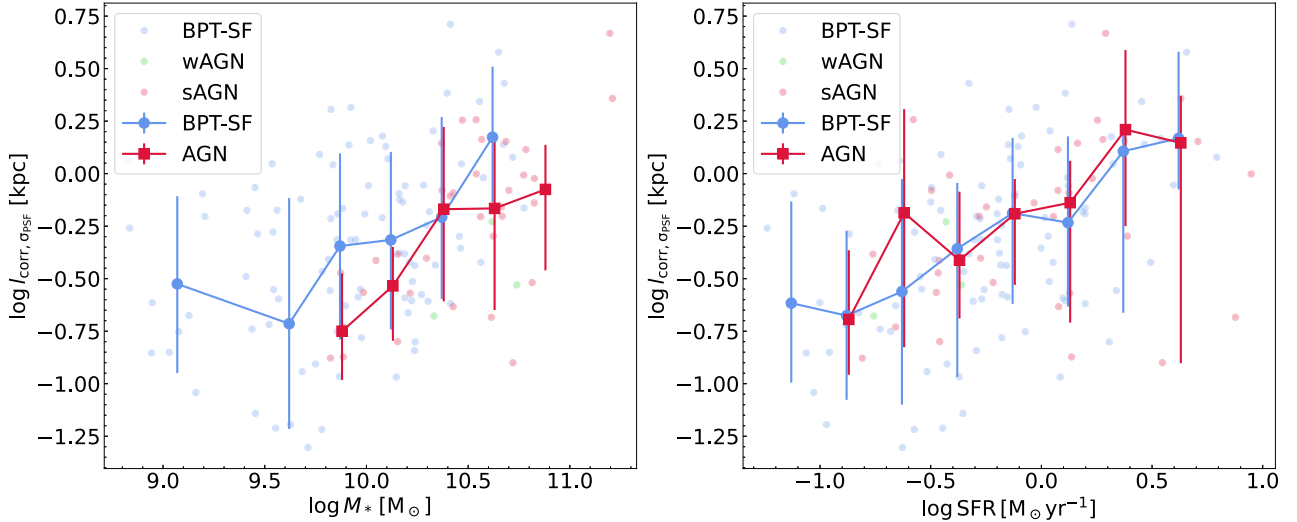
To test for this effect in our sample, we stack the flux of He II, H  $\beta$ , H  $\alpha$ , and  $[\text{N II}]$  within central  $2''.50$  of each galaxy, which is the beam size of CALIFA, and adopt the same SNR cut and the same demarcation line to classify AGN as in Tozzi et al. (2023). We find that seven galaxies out of the 132 in our sample are identified as AGN hosts by this method, but that only two of these were classified as BPT-SF based on the traditional BPT diagram. If we reclassify these extra two galaxies as AGN, our AGN sample grows from 37 to 39, an increase of 6 per cent, completely consistent with Tozzi et al.'s result that 6 per cent of AGN can be found using the He II line only.

We now repeat the analysis in the paper for a modified sample in which we re-classify these two galaxies from BPT-SF to sAGN in order to determine how our results change. Fig. A1 shows the SFR versus stellar mass relation of AGN hosts and BPT-SF galaxies after this reclassification. Comparing with Fig. 4, we see that our qualitative result that AGN hosts have generally smaller SFRs than BPT-SF galaxies at fixed stellar mass remains unchanged. Similarly, Fig. A2 shows the relationship between  $l_{\text{corr}}$ , stellar mass, and SFR we obtain after this reclassification. Comparing this diagram with Fig. 7, the most obvious difference is in the bin  $\log \text{SFR} \in [0.5, 0.75]$ , where now the median  $l_{\text{corr}}$  of AGN hosts is apparently greater than that of BPT-SF galaxies. However, this visual shift is mostly driven by the very small number of AGN in this bin, and is not statistically significant: a  $t$ -test with the null hypothesis that the mean  $l_{\text{corr}}$  of BPT-SF is the same as that of AGN hosts returns a  $p$ -value of 0.98. Our main conclusions therefore again remain unchanged.





**Figure A2.** Same as Fig. 7, but with two galaxies re-classified from BPT-SF to sAGN based on the He II diagram as discussed in Appendix A.

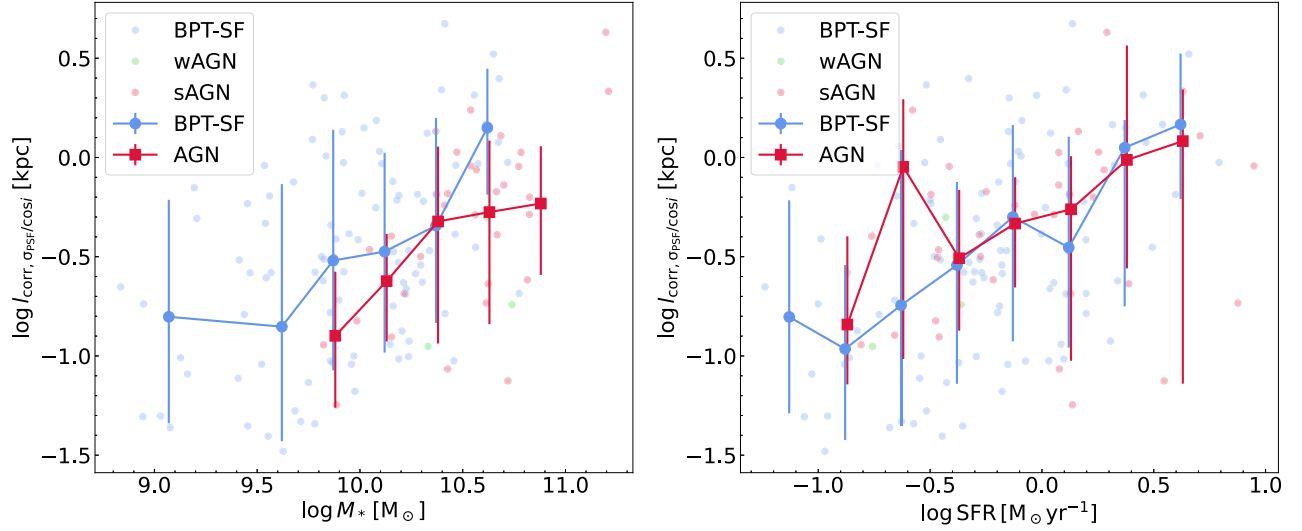


**Figure B1.** Same as Fig. 7, but for  $l_{\text{corr}}$  derived from a fit where we set the prior on  $\sigma_{\text{beam}}$  to a very narrow Gaussian centred on  $\sigma_{\text{PSF}}$ , the measured PSF size from CALIFA; see text for details.

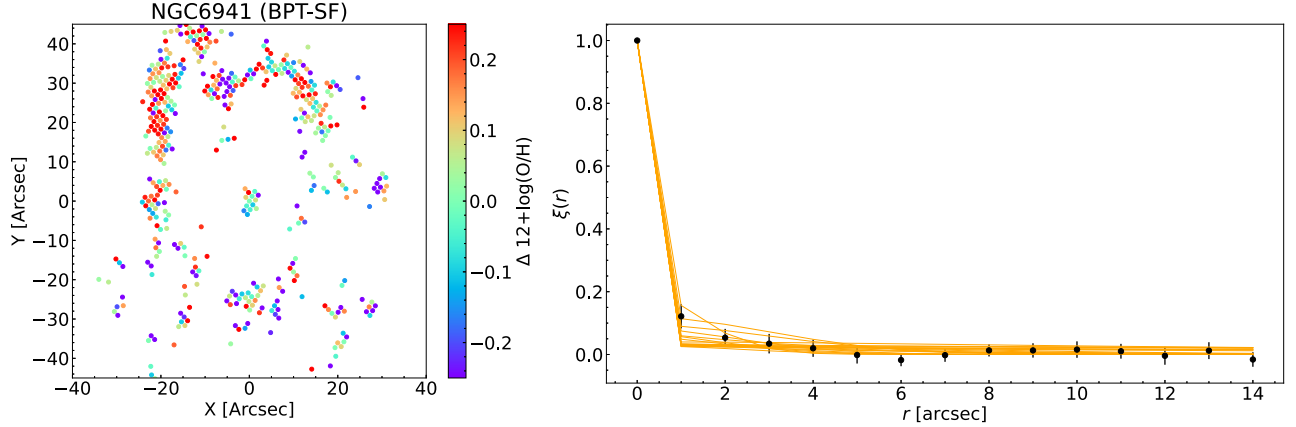
## APPENDIX B: INFLUENCES OF THE PRIORS OF THE BEAM SIZE

As discussed in Section 2.4, due to deprojection of the metallicity maps the observational beam along the minor axis of the galaxies is stretched by a factor of  $1/\cos i$ , where  $i$  is the inclination angle calculated from equation (2). In the main text we handle this issue by setting our prior on the beam size,  $\sigma_{\text{beam}}$ , to a Gaussian whose centre is the arithmetic mean of  $\sigma_{\text{PSF}}$  and  $\sigma_{\text{PSF}}/\cos i$ , the beam sizes along the major and minor axis, and with a width equal to half the difference between the major and minor axis beam sizes. In this appendix we test the sensitivity of our results to this choice by repeating the full analysis pipeline in the main paper using two extreme alternatives for the prior.

Figs B1 and B2 show the relation between  $l_{\text{corr}}$  and stellar mass as well as the relation between  $l_{\text{corr}}$  and SFR for both AGN-host galaxies and BPT-SF galaxies that we derive by setting the mean of the Gaussian prior on the  $\sigma_{\text{beam}}$  to  $\sigma_{\text{PSF}}$  and  $\sigma_{\text{PSF}}/\cos i$ , respectively; in both cases we set the standard deviation on the prior to of 0.01 arcsec, a very small value that essentially fixes  $\sigma_{\text{beam}}$  to the value we assign for the mean. Again, for those without a PSF measurement, we assign the mean of the Gaussian prior on the  $\sigma_{\text{beam}}$  to 1.06 arcsec and 1.06 arcsec/ $\cos i$ , and corresponding standard deviation to 0.14 arcsec and 0.14 arcsec/ $\cos i$ , respectively, where 1.06 arcsec and 0.14 are the mean and the  $1\sigma$  level of scatter of the PSF distribution in CALIFA, respectively. We are therefore testing how our results change if we force the beam sizes to the largest and smallest plausible values.



**Figure B2.** Same as Fig. B1, but now setting the central value of the prior on beam size to  $\sigma_{\text{PSF}} / \cos i$ , where the  $\sigma_{\text{PSF}}$  is the measured PSF size from CALIFA and  $i$  is the inclination angle calculated from equation (2).



**Figure C1.** The same as Fig. 2. However, we show an example of which we fail to fit the two-point correlation function with the procedures elaborated in Section 2.4.

Comparing these two diagrams with Fig. 7, we find a qualitatively similar result that AGN hosts have smaller  $l_{\text{corr}}$  than BPT-SF galaxies at fixed stellar mass, but show no apparent difference at fixed SFR. We thus confirm that our main conclusion remain unchanged even under extreme assumptions about beam sizes.

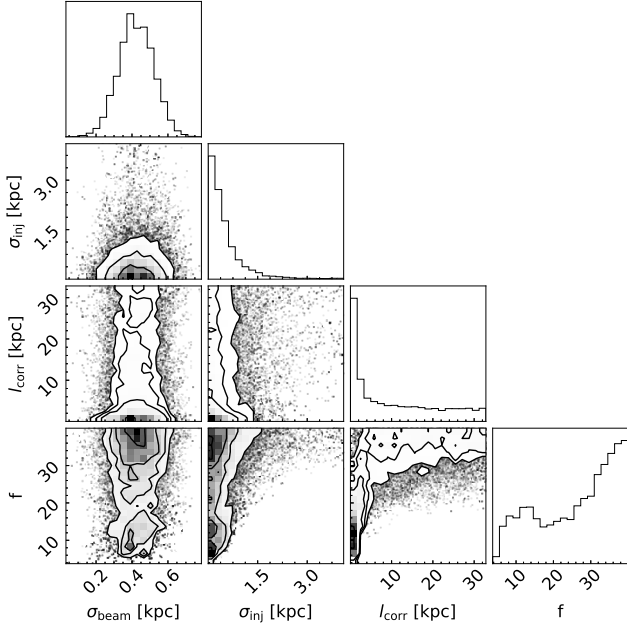
### APPENDIX C: FAILED FITS TO THE TWO-POINT CORRELATION FUNCTION

By inspecting the posterior PDF of the MCMC sampling, we find that we fail to obtain a reasonable  $l_{\text{corr}}$  for three galaxies: NGC4470, NGC1070, and NGC6941. For each of these galaxies, the posterior PDF of  $l_{\text{corr}}$  extends up to the maximum allowed by our prior, and is in general extremely broad, suggesting that our MCMC is not finding an acceptable model. To explore the origin of this failure, in Figs C1 and C2 we show the metallicity fluctuation map, two-point correlation function, and posterior PDF of the parameters describing the two-point correlation function for one of these failed galaxies

NGC6941; the other two are qualitatively similar, and the explanation we present in this Appendix applies equally well to them.

We first see that for this galaxy the metallicity fluctuation map is sparsely sampled due to the SNR cut described in Section 2.1. To understand if this is the primary reason for the failed fit, we calculate the area filling factors within the convex hull following the procedure described in L23. We find filling factors of 0.17 for NGC6941, 0.23 for NGC1070, and 0.35 for NGC4470, all of which are larger than the threshold of 0.04 obtained empirically by L23 for their sample. Moreover, while these filling factors are somewhat low, there are a large number of galaxies in our sample that we do successfully fit despite their similar or even smaller filling factors.

Instead, we can understand the origin of the failure by noting that the joint posterior PDF of  $f$  and  $l_{\text{corr}}$  in Fig. C2 shows two ‘wings’, one with  $l_{\text{corr}}$  smaller than the beam size and smaller  $f$  and another with large  $f$  and essentially a uniform distribution of  $l_{\text{corr}}$ . We can understand the origin of these two branches by examining the two-point correlation function to which we are fitting (right panel of Fig. C1). The essential point to notice is that this galaxy shows essentially no correlations in its metallicity field – excluding the point



**Figure C2.** The same as Fig. 3. But we show the posterior distribution of four parameters for NGC6941, a galaxy for which our MCMC fitting does not produce a good measurement of  $l_{\text{corr}}$ .

at zero lag which is always unity, the most correlated bin only has a central value  $\xi \approx 0.1$ , and given the size of the error bar this point is consistent with  $\xi = 0$  at the  $2\sigma$  level. In terms of the model to which we are fitting, there are two ways to produce a two-point correlation function that looks like the right panel of Fig. C1. One is if the correlation length is significantly smaller than the size of the beam, so that no correlations exist on resolved scales, and this possibility corresponds to the small  $l_{\text{corr}}$  wing of the posterior PDF. The other is if the observational error in the metallicity, parametrized by  $f$ , is so large that any correlations present in the data are buried below the noise; this corresponds to the large  $f$  wing of the posterior PDF, for which the marginal posterior PDF of  $l_{\text{corr}}$  is simply the prior, because the data are too noisy to permit the correlation to be measured.

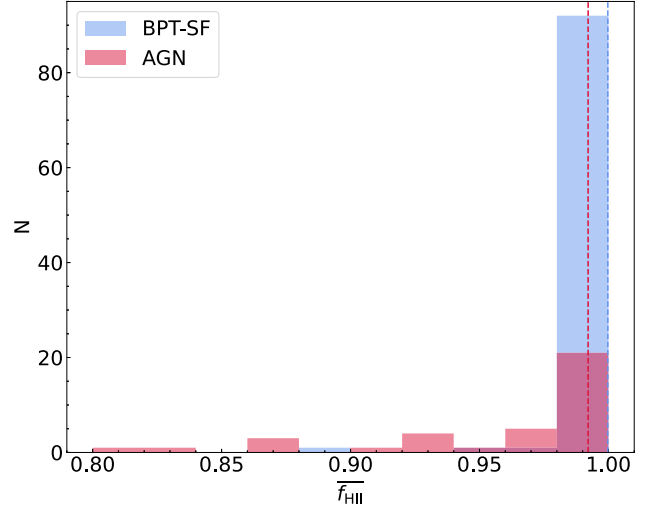
Given this understanding of the fit failures, we conclude that we cannot meaningfully constrain  $l_{\text{corr}}$  for these three galaxies, and we therefore exclude them from our analysis.

#### APPENDIX D: ON AGN CORRECTIONS TO THE SFR

As discussed in Section 2.5, the SFR reported in the eDR3 catalogue is directly converted from  $H\alpha$  luminosity without any attempt to correct for AGN contributions. Accordingly, Sánchez et al. (2023) point out that the SFR for AGN-host galaxies can only be treated as an upper limit. In this appendix, we quantify the level of AGN contamination following the procedure described in L24.

We firstly run NEBULABAYES on our sample with the same spaxel masks and following the same procedure introduced in Sections 2.1 and 2.2. One output of this calculation for every spaxel is  $f_{\text{HII}}$ , which describes the fraction of  $H\alpha$  luminosity attributable to HII regions driven by stars, as opposed to AGN. We then calculate the  $H\alpha$ -luminosity weighted mean  $f_{\text{HII}}$ :

$$\overline{f_{\text{HII}}} = \frac{\sum_i L_{H\alpha,i} f_{\text{HII},i}}{\sum_i L_{H\alpha,i}}, \quad (\text{D1})$$



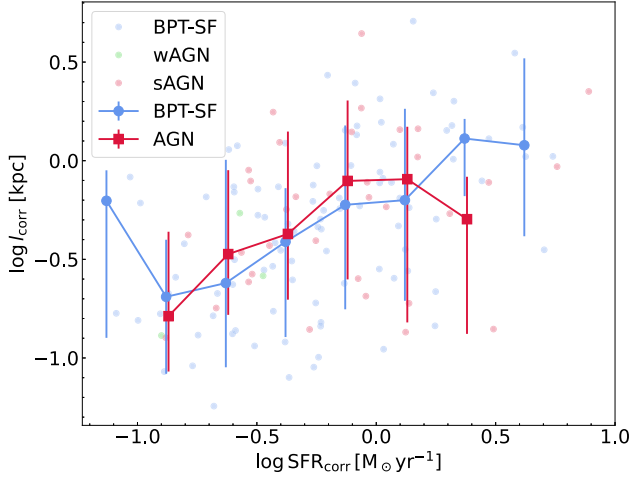
**Figure D1.** The distribution of the  $H\alpha$  luminosity-weighted HII region fraction  $f_{\text{HII}}$  calculated by equation (D1) for BPT-SF galaxies (blue) and AGN-host galaxies (red). The blue (red) dashed line indicates the median  $\overline{f_{\text{HII}}}$  of BPT-SF (AGN-host) galaxies.

where  $L_{H\alpha,i}$  and  $f_{\text{HII},i}$  are the dust-corrected  $H\alpha$  luminosity and value of  $f_{\text{HII}}$  returned by NEBULABAYES for the  $i$ th spaxel; the sum is over all non-masked spaxels. Thus  $\overline{f_{\text{HII}}}$  reflects our best estimate of the fraction of the total  $H\alpha$  luminosity contributed by massive stars, and thus attributable to star formation. Values of  $\overline{f_{\text{HII}}}$  closer to unity indicate less contamination from AGN. Fig. D1 shows the distribution of  $\overline{f_{\text{HII}}}$  values for the BPT-SF (blue) and AGN-host (red) galaxies in our sample. The median value  $\overline{f_{\text{HII}}}$  for BPT-SF galaxies is close to unity, as expected, but even for AGN-host galaxies this falls only to 0.992. It is therefore clear that even for AGN-host galaxies, the AGN contribution to the overall  $H\alpha$  luminosity is marginal. This is because most AGN-host galaxies in our sample are similar to the example shown in Fig. 1, where the AGN contributes significantly only in the centre, and is not the dominant ionizing source throughout the galaxy.

To understand how correcting for AGN contamination might influence our final results, we re-estimate the SFR excluding the AGN contribution using our  $f_{\text{HII}}$  maps:

$$\text{SFR}_{\text{NB}} = \left[ \sum_i L_{H\alpha,i} f_{\text{HII},i} \right] \times \frac{7.9 \times 10^{-42} \text{ M}_{\odot} \text{ yr}^{-1} / (\text{erg s}^{-1})}{10^{0.25}}. \quad (\text{D2})$$

Here  $L_{H\alpha,i}$  and  $f_{\text{HII},i}$  are the same as in equation (D1), and the factor of  $10^{0.25}$  in the denominator is to correct from the Salpeter (1955) IMF to Chabrier (2003) IMF. It is important to note that, because many spaxels in our maps are masked due to low SNR, we cannot directly compare  $\text{SFR}_{\text{NB}}$  to the SFR derived from the catalogue,  $\text{SFR}_{\text{eDR3}}$  – even in galaxies with no AGN contamination the latter is usually significantly larger because it is based on the integral  $H\alpha$  luminosity, whereas  $\text{SFR}_{\text{NB}}$  includes only the  $H\alpha$  flux arising from spaxels bright enough to pass our SNR cuts. Given this situation, our approach to estimating the AGN correction is to fit a linear relation between the  $\log \text{SFR}_{\text{NB}}$  and the  $\log \text{SFR}_{\text{eDR3}}$  for BPT-SF galaxies, where there is no AGN contamination, and use the same relation to calibrate  $\text{SFR}_{\text{NB}}$  to a corrected value  $\text{SFR}_{\text{corr}}$  for AGN-host galaxies. This amounts to assuming that we are missing the same fraction of the flux due to low SNR in both AGN-host and non-AGN galaxies,



**Figure D2.** The same as the right panel of Fig. 7, but showing the relation of  $l_{\text{corr}}$  with the corrected SFR computed following the procedure described in Appendix D.

and if anything this assumption will tend to undercorrect the AGN-host galaxies, since the generally smaller SFR in AGN hosts reduces the number of spaxels with SNR high enough to pass our cuts. We can therefore regard the value of  $\text{SFR}_{\text{corr}}$  we derive from this procedure as

a lower limit on the SFR in AGN-host galaxies, with the uncorrected value  $\text{SFR}_{\text{eDR3}}$  representing an upper limit.

Fig. D2 shows the relation between the  $l_{\text{corr}}$  and the corrected  $\log \text{SFR}_{\text{corr}}$ ; we generate this figure exactly as we generate the right panel of Fig. 7, but using  $\text{SFR}_{\text{corr}}$  instead of  $\text{SFR}_{\text{eDR3}}$ . As in the right panel of Fig. 7, the BPT-SF and the AGN-host galaxies show similar trends. Using the same  $t$ -test procedure as described in Section 3, we find that the  $p$ -values with which we can reject the null hypothesis that the mean  $l_{\text{corr}}$  of the AGN hosts equal to that of BPT-SF galaxies are 0.92, 0.40, 0.22, 0.41, 0.87, and 0.03 from the lowest SFR to the highest bin. Thus, we obtain the same result as we did for  $\text{SFR}_{\text{eDR3}}$ , which is that we cannot reject the null hypothesis. Given the small difference that AGN correction makes to our results, we elect to use the uncorrected SFR from the eDR3 catalogue in the main body of this work.

This paper has been typeset from a  $\text{T}_{\text{E}}\text{X}/\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$  file prepared by the author.