

Generating all-sky radio continuum clustering simulations with GHOST

Brandon Venville^{1*}, Anna Bonaldi², David Parkinson³, Natasha Hurley-Walker¹, Tim Galvin⁴, Nick Seymour¹,

¹ICRAR-Curtin, Curtin University, Bentley, 6102, Western Australia, Australia

²SKA Organization, Jodrell Bank, Lower Whittington, Macclesfield SK11 9DL, UK

³Korea Astronomy and Space Science Institute, Daejeon 34055, Republic of Korea

⁴ATNF, CSIRO Space & Astronomy, Bentley, WA, Australia

Abstract

Techniques using multiple tracers of the large scale structure of the universe show great promise for examining the fundamentals of our Universe’s cosmology. Such techniques rely on the different relationship between the overdensity of tracers and the broader matter overdensity, enabling cosmic-variance-free tests of primordial non-Gaussianity in the initial curvature perturbations. There is a great opportunity for current and future all-sky extra-galactic radio surveys to make use of this technique to test for non-Gaussianity at a precision greater than existing all-sky constraints from the cosmic microwave background. To realize this goal there is a need for accurate simulations. Previous radio galaxy simulations have either been realistic but covering only a small area (and so unhelpful for cosmological forecasts), or all-sky dark matter only cosmological simulations but having no connection to a real radio galaxy population. In this study, we use the FLAMINGO suite of cosmological surveys, as well as the matching of dark matter halos to radio galaxy population, to create an accurate sky simulation in order to examine the feasibility of multi-tracer techniques. We present an analysis of the clustering (with a bias model for the simulation), as well as redshift distributions, source counts and radio luminosity functions, and discuss future work on non-Gaussianity detection.

1 INTRODUCTION

Multi-tracer techniques show great promise in constraining primordial non-gaussianity (PNG) when applied to next generation large-scale extra-galactic surveys. PNG encodes the dynamics of the early Universe and offers a sharp test of inflationary physics. In the widely used ‘local’ model (Bartolo et al., 2004; Chen, 2010), its amplitude is captured by f_{NL} , the fractional non-linearity, which imprints a distinctive, scale-dependent modulation of galaxy bias b on ultra-large scales k and at redshift z , $\Delta b(k, z) \propto f_{\text{NL}} k^{-2}$ (Dalal et al., 2008; Gomes et al., 2020a), enhancing power on the largest scales.

While attractor single-field slow-roll inflation predicts a nearly Gaussian field with $f_{\text{NL}} \ll 1$ (Maldacena, 2003), non-attractor, or other non-standard dynamics can naturally yield f_{NL} of order unity and distinctive bispectrum shapes (Bartolo et al., 2004; Chen, 2010). Current CMB bispectrum limits show no detection but still is consistent with local type non-Gaussianity $f_{\text{NL}}^{\text{local}} = -0.9 \pm 5.1$ (Planck Collaboration et al., 2020), leaving room for further tests. Forecasts for upcoming surveys, for exam-

ple with the SKA Observatory (SKAO) telescopes - the SKAO low-frequency radio telescope (SKA-Low) and the SKAO mid-frequency radio telescope (SKA-Mid) - and with *Euclid* (Amendola et al., 2018), indicate sub-unity precision on $f_{\text{NL}}^{\text{local}}$ under realistic assumptions, motivating a multi-tracer approach (Yamauchi et al., 2014; Gomes et al., 2020a)¹.

Multi-tracer methods are uniquely suited to this regime: by combining galaxy populations with different linear biases and redshift kernels, one can cancel cosmic variance in shared modes and convert relative clustering into information on f_{NL} . This typically involves identifying tracers of large scale structure which trace the underlying total matter distribution differently, and using the common distribution to negate the effects of cosmic variance. Large sky area surveys provide access to large angular scales, where the signals of non-Gaussianity are most likely to be found. Theoretical

¹Here $f_{\text{NL}}^{\text{local}}$ denotes the amplitude of the “local” bispectrum shape; in general f_{NL} is shape-dependent (e.g. local, equilateral, orthogonal), but large-scale structure is maximally sensitive to the local case because it induces a scale-dependent galaxy bias $\Delta b(k) \propto k^{-2}$. For brevity, unless stated otherwise we use $f_{\text{NL}} \equiv f_{\text{NL}}^{\text{local}}$ hereafter.

*email:brandon.venville@postgrad.curtin.edu.au

studies indicate we may be able to perform such techniques (Gomes et al., 2020b; Ferramacho et al., 2014), where previously a lack of appropriately observed source populations, with identification of different tracers, has hampered efforts.

Practical sensitivity depends on population mixing, relative ratios in clustering between active galactic nuclei (AGN) and star forming galaxies (SFG), flux- and frequency-dependent selection, photometric- z errors, survey masks and depth gradients, resolution/morphology effects, and calibration systematics that can mimic large-scale power (Abramo & Leonard, 2013; Ferramacho et al., 2014; Gomes et al., 2020b). A new, end-to-end forward simulation that ties halo populations, luminosity functions, and spectral assumptions to survey-specific observing effects is therefore essential to quantify biases, propagate uncertainties, and validate multi-tracer estimators in conditions that match real data.

Prior studies such as Bonato et al. (2017) and the Tiered Radio Extragalactic Continuum Simulatio (T-RECS) (Bonaldi et al., 2019) underpin our approach. Bonato et al. (2017) matched redshift-dependent star formation rate (SFR) functions to 1.4-GHz radio luminosity functions (RLFs) to calibrate the radio-SFR mapping, explored linear vs. mildly non-linear relations (including scatter), updated the Massardi et al. (2010) evolutionary model for radio-loud AGN, and used number counts to motivate a mild redshift evolution in the synchrotron/SFR ratio. T-RECS provides an end-to-end population synthesis for flat-spectrum radio quasars (FSRQ), BL Lac, and steep-spectrum AGN and for SFGs, including empirical/evolutionary RLFs, spectral-index prescriptions, size and subclass proxies, and mock catalogues designed to reproduce source counts and clustering. In this work we adopt those physically motivated ingredients (RLF parameterizations, SFR to radio calibrations, and soft high excitation radio galaxies (HERG)/low excitation radio galaxies (LERG) mapping with mass weighting) but embed them in a halo-based framework tied to a cosmological lightcone (FLAMINGO; Schaye et al., 2023; Kugel et al., 2023) and a forward survey emulator. This lets us enforce consistency between RLFs and the halo supply via abundance matching, and propagate realistic selection effects into counts, redshift distributions, and clustering—bridging population models and large-volume cosmology.

This study seeks to use the aforementioned N-body simulations and analytic halo population to verify the possibilities of using multi-tracer techniques in cosmology, with a realistic sky and experimental systematics. There will be two papers in this series: In this first paper we build GHOST, an all-sky, clustered, survey-ready radio mock catalogue tied to a cosmological lightcone, expressly to validate multi-tracer pipelines for f_{NL} . We

construct population- and survey-aware catalogues that reproduce wide-area measurements, quantify how selections reshape redshift kernels and deliver the calibrated inputs (galaxy positions, fluxes and halo mass mapping) required for end-to-end PNG tests. By supplying tracer families with distinct linear biases and redshift kernels, the catalogue enables controlled tests of cosmic-variance cancellation as it concerns non-Gaussianity constraints, estimator choices in configuration and harmonic space, and the impact of realistic masks, photo- z tails, and population misclassification (Seljak, 2009; Abramo & Leonard, 2013; Ferramacho et al., 2014; Yamauchi et al., 2014; Dalal et al., 2008; Gomes et al., 2020b; Planck Collaboration et al., 2020).

Paper II will use these products to perform the f_{NL} analysis on the simulated sky. The surveys of current instruments, such as the LOFAR Two-metre Sky Survey (LoTSS; Shimwell et al. 2017) using the LOw-Frequency ARray (LOFAR; van Haarlem et al. 2013), as well as the planned surveys of future instruments like *Euclid* and the SKAO telescopes shall be used to inform source selection, identification and cosmological analysis, to predict the constraints that may be obtained on cosmological parameters.

GHOST also will provide a simulation in the vein of T-RECS (Bonaldi et al., 2019) and European SKA Design Study Simulated Skies (SKADS) (Wilman et al., 2008), but with all-sky coverage as well as large-scale clustering. These improvements will support numerous cosmological applications, for example late-time CMB cross-correlation. Full-sky density fields with known selection allow pipeline development for the integrated Sachs–Wolfe effect and cross-correlation with CMB lensing convergence, including null tests under injected large-scale systematics (dipoles, depth gradients, bright-source masks) and validation of mitigation strategies using pseudo- C_ℓ machinery matched to survey footprints (e.g. Bahr-Kalus et al., 2022).

Weak-lensing and magnification applications for GHOST are supported via add-on ray-tracing. Coupling the sky to a lensing engine (e.g. Borzyszkowski et al., 2017, LIGER) turns the catalogue into a forward model for magnification and shear of radio sources, enabling tests of number-count and flux-weighted magnification estimators and of their sensitivity to the population mix and selection (Borzyszkowski et al., 2017).

The GHOST simulation is also suited to studies of the cosmic radio dipole and survey-induced dipoles. With contiguous sky coverage and explicit selection functions, users can inject a kinematic dipole and survey systematics to quantify amplitude/direction bias, or measure the density dipole of flux-limited samples under controlled population mixtures. This connects directly to the radio-dipole literature (e.g. Böhme et al., 2025; Wa-

genveld et al., 2025). For survey design, the large sky area of GHOST provides a sandbox for confusion, flux boosting, and spatially varying completeness. Because classes are labelled, the catalogue supports training and stress-testing of cross-identification and population-classification models under controllable class imbalance, and exploration of how selection affects downstream cosmology.

Finally, masking GHOST to medium/small fields enables direct forecasts of cosmic-variance floors and assessment of when multi-tracer or cross-correlation approaches deliver meaningful gains for deep tiers, while maintaining consistency with the wide-area selection and clustering that drive ultra-large-scale signals.

The subsequent Sections include a discussion of the FLAMINGO simulations (Schaye et al., 2023; Kugel et al., 2023) and how they were used here (Section 2), the model descriptions (Section 3.2, 3.3 and 4) and results from the simulation in Section 5, including clustering (Section 5.3), redshift distributions (Section 5.6), source counts (Section 5.7), and RLF results (Section 5.8). The luminosity distribution with redshift is discussed in Section 6, and future improvements to GHOST in Section 7. We then finish with conclusions of the GHOST simulation in Section 9.

2 FLAMINGO SIMULATION BASICS

The FLAMINGO simulations (Schaye et al., 2023; Kugel et al., 2023) comprise a suite of 28 runs — 16 hydrodynamical and 12 dark matter only (DMO) — spanning a range of particle numbers, particle masses, cosmological models, and box sizes. Developed by the Virgo Consortium, FLAMINGO aims to support upcoming large-scale structure surveys through high-fidelity predictions. The simulations use the SWIFT hydrodynamics code (Schaller et al., 2024), solving the fluid equations with the SPHENIX flavour of smoothed particle hydrodynamics (SPH) (Borrow et al., 2022). FLAMINGO features a novel approach to calibrating subgrid physics — including AGN feedback — by applying machine learning techniques trained on observed local datasets. Halo identification, whether in DMO or hydrodynamic simulations, was performed using the VELOCIRAPTOR halo finder (Elahi et al., 2019). Subhalo properties were calculated using the Spherical Overdensity and Aperture Processor (SOAP) (McGibbon et al., 2025) code, developed specifically for FLAMINGO. Available data products include SOAP catalogues for multiple redshift snapshots, lightcones for select runs, and HEALPIX maps of derived lightcone properties. The full FLAMINGO simulation specifications are detailed by Schaye et al. (2023).

In this work, we use a FLAMINGO DMO simula-

tion based on the best-fit cosmology from the Dark Energy Survey Year 3 analysis, specifically the “ $3 \times 2pt + \text{All Ext } \Lambda\text{CDM}$ ” model. The corresponding cosmological parameters are listed in Table 1.

We specifically use the L1000N3600 run, with a 1 Gpc^3 box and 3600^3 particles of mass $8.40 \times 10^8 M_\odot$, with the intention of providing sufficient mass resolution to resolve the halos of less massive, low SFR SFG. $\sim 2.29 \times 10^9$ halos were selected from a total of $\sim 1.14 \times 10^{11}$ available in redshift $z = 0$ to $z = 10$ and assigned celestial coordinates (RA and Dec) based on their 3D positions in the simulation box at each redshift snapshot. The observer is located at $(0, 0, 0)$, the center of the coordinate system.

The FLAMINGO lightcone provides the two ingredients our f_{NL} tests are most sensitive to: ultra-large angular modes from a $\sim \text{Gpc}$ volume so that the lowest multipoles and cross-population covariance are realized in a single, contiguous sky; and a well-sampled halo mass function down to the scales that host low-SFR SFG. We intentionally anchor the abundance matching and host assignment to this specific halo field, because the PNG signal enters large-scale clustering primarily through the tracers’ (mass-dependent) bias. Using the same halo supply for catalog construction and for clustering predictions ensures that $n(z)$, effective bias, and survey selections are mutually consistent when we calculate angular clustering statistics. While this paper employs a DMO run, baryonic effects are subdominant at the very large scales driving the scale dependent bias from non-Gaussianity; their residual impact is absorbed in the empirical validation against counts and $n(z)$, and in the bias calibration we carry forward to Paper II.

3 RELATION TO T-RECS AND RATIONALE FOR A NEW IMPLEMENTATION

Our goal is an all-sky, clustered, survey-ready radio mock catalogue built to validate multi-tracer pipelines on ultra-large scales and, ultimately, to constrain f_{NL} . This demands contiguous full-sky light-cones so that the lowest multipoles and cross-population covariances are realized, a halo assignment re-derived on our N -body field so that large-scale bias reflects the actual halo supply, and end-to-end, instrument-aware selection functions. T-RECS provides excellent population prescriptions and public clustered mocks, but its clustered realizations are limited in area and tied to a different halo pipeline, which is sub-optimal for the variance-cancellation regime targeted here (Bonaldi et al., 2019). For f_{NL} , where the signal appears as a k^{-2} modulation of bias at the largest angular scales, these constraints are decisive.

We adopt the same high-level taxonomy and broad-

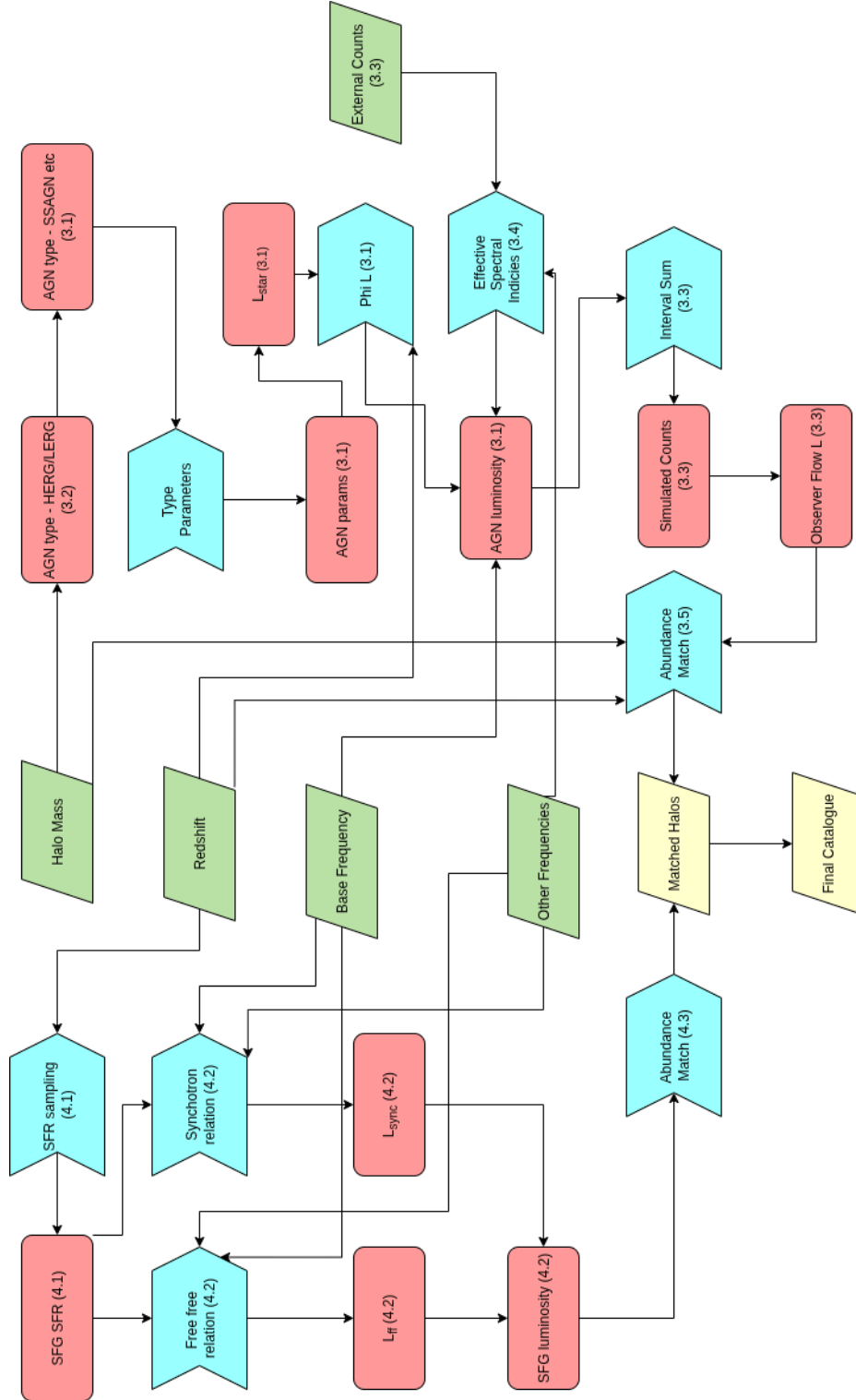


Figure 1. Population workflow used in *ghost*. The diagram summarizes how radio sources are generated for the two branches, with relevant subsections of the paper indicated by the bracketed numbers: **SFG** (left) and **AGNs** (right). *Green slanted boxes* are external inputs (halo mass, redshift, base/other frequencies, external counts); *cyan chevrons* are operations/relations (e.g. SFR sampling, synchrotron/free-free relations, evolving RLF $\Phi(L)$, spectral-index mapping, abundance matching); *red rounded boxes* are intermediate products or outputs (component luminosities $L_{\text{ff}}/L_{\text{sync}}$, SFG/AGN luminosities, simulated counts, observer-frame fluxes). On the SFG side, sampled SFRs are converted to radio luminosity via synchrotron and free-free relations, then propagated to other frequencies. On the AGN side, type parameters (HERG/LERG, steep-spectrum AGN (SSAGN), etc.) and the evolving RLF (via L_{star} and $\Phi(L)$) set the AGN luminosity, with spectral indices mapping to other bands. Both branches feed an *abundance-matching* step (using halo mass and redshift) to assign hosts and form the final matched-halo catalogue; the counts/interval-sum nodes provide predicted number counts used for verification. Arrows indicate data flow from inputs to catalogue outputs.

Table 1 Cosmological parameters adopted for the fiducial flat Λ CDM model used in this work. Columns list the symbol, parameter name, and adopted value (all quoted at $z = 0$). We use the dimensionless Hubble constant h ; the total matter, dark-energy, and baryon density parameters Ω_m , Ω_Λ , and Ω_b ; the summed neutrino mass $\sum m_\nu$; the amplitude and spectral index of the primordial scalar power spectrum, A_s and n_s ; and the present-day r.m.s. density fluctuation in $8 h^{-1}$ Mpc spheres, σ_8 . We also quote two derived parameters: $S_8 \equiv \sigma_8 \sqrt{\Omega_m}/0.3$ and the neutrino density parameter Ω_ν defined via $\Omega_\nu h^2 = \sum m_\nu / (93.14 \text{ eV})$.

Parameter	Value
h	0.681
Ω_m	0.306
Ω_Λ	0.694
Ω_b	0.0486
$\sum m_\nu$	0.06 eV
A_s	2.099×10^{-9}
n_s	0.967
σ_8	0.807
S_8	0.815
Ω_ν	1.39×10^{-3}

band philosophy as T-RECS, and inherit some of the modeling: radio AGN split into FSRQ, BL Lacertae (BLLac), and SSAGN, and SFG with synchrotron and free-free contributions to their radio spectral energy distributions (SEDs). We retain smoothly evolving RLFs, class-dependent spectral-index distributions, and a minimalist Fanaroff–Riley (FR) proxy for internal labeling. GHOST differs from T-RECS in three ways most relevant to f_{NL} . First, we re-derive the luminosity–halo mapping against the FLAMINGO light-cone and use a rank-preserving, no-replacement match within each redshift slice; this enforces monotonicity, respects the available halo counts, and stabilizes large-scale bias. Second, we generate contiguous, full-sky clustered realizations so the ultra-large angular modes are present. Third, our validation focuses on the observables that feed PNG analyses—Euclidean-normalized counts, redshift kernels $n(z)$, and angular clustering $w(\theta)/C_\ell$ —for total and per-population samples under survey-like cuts.

For AGN, T-RECS draws host halo masses from HERG/LENG-based probabilities (inferred via stellar-to halo-mass relations) and associates each source to the nearest-mass halo in the same redshift slice, excluding already-assigned halos. For SFG, it fits an $L(M_h)$ relation by abundance matching and inverts it to map radio luminosities to halo masses, assigning those to halos per slice, with very faint sources below the halo-mass floor placed randomly on the sky. In GHOST we instead perform a global, rank-order match within each slice (monotonic $L \leftrightarrow M_h$, no replacement), which prevents Eddington-style leakage across the steep end of the mass function, guarantees that the integral bias of each subsample matches the simulated halo supply, and reduces slice-to-slice bias fluctuations that would otherwise propagate into the k^{-2} PNG signature.

3.1 Types of galaxies in GHOST

GHOST includes four observational families that dominate the extragalactic radio sky at metre–centimetre wavelengths. Radio AGN appear in three classes—FSRQ, BLLac, and SSAGN — which differ in beaming/orientation and spectral behaviour but are often unified as the same central engine viewed at different angles (Urry & Padovani, 1995; Heckman & Best, 2014). Star-forming galaxies (SFG) form a separate branch whose radio output tracks recent star formation. Below we summarize the properties and the labels used in this work.

- **FSRQ:** Flat-spectrum radio quasars show core-dominated, compact radio emission with flat spectra (assumed mean $\alpha = -0.1^2$), bright optical continua, and broad emission lines (Urry & Padovani, 1995). In orientation-based unification they are relativistic jets viewed close to the line of sight, hence they have Doppler-boosted luminosities and are compact.
- **BLLac:** Blazars with weak/absent emission lines and flat radio spectra, likewise interpreted as beamed (Doppler-boosted) jets at small inclination; compared to FSRQ they typically have lower line equivalent widths and similarly compact morphologies (Urry & Padovani, 1995).
- **SSAGN:** Steep-spectrum radio AGN, representing the bulk of radio-loud AGN, with mean spectral index $\zeta = -0.8$. This class includes spectroscopic and morphological sub-taxonomies:
 - **LERG:** Jet-dominated systems with radiatively inefficient, low-Eddington accretion and

²We define the radio spectral index via $S_\nu \propto (\nu/\nu_0)^\alpha$; unless stated otherwise we take $\nu_0 = 1.4 \text{ GHz}$. When quoted as $\alpha_0 \pm \sigma_\alpha$ this denotes the population mean and its 1σ scatter; by this convention synchrotron sources have $\alpha < 0$.

weak/absent optical lines.

- HERG: Radiatively efficient (“radiative-mode”) accretors with strong optical emission lines (Heckman & Best, 2014; Bonaldi et al., 2019).
- FRI: Edge-darkened morphology (brightness peaks nearer the core; jets fade outward), more common at lower radio power (Fanaroff & Riley, 1974).
- FRII: Edge-brightened lobes with terminal hotspots, more common at higher radio power (Fanaroff & Riley, 1974).

These schemes probe different physics and are not one-to-one (e.g. FRII LERG exist). In our catalogue the FR tag is assigned via the size-ratio parameter R_s (Section 3.3), defined as the separation of peak-brightness regions divided by the total source extent; $R_s > 0.5$ is labelled FRII and $R_s < 0.5$ FRI.

- SFG: Radio emission powered predominantly by recent star formation: non-thermal synchrotron from cosmic-ray electrons accelerated in supernova remnants plus thermal free-free from H II regions (Condon, 1992; Murphy et al., 2011). At GHz frequencies the ensemble spectral index is steep (typ. $\zeta \approx -0.7$ to -0.8) with a mild flattening from free-free toward higher frequencies. In GHOST the SFG branch is modelled with a calibrated radio luminosity function at 1.4 GHz and class-appropriate K -corrections; it is treated separately from AGN because its radio output traces star-formation physics rather than black-hole accretion.

Radio AGN typically inhabit more massive halos than SFG and therefore have a higher large-scale bias; within AGN, radiatively inefficient/jet-dominated systems tend to occupy the most massive environments, while radiatively efficient (quasar-like) systems are somewhat less clustered. Conversely, radio-selected SFG are associated with lower-mass halos and exhibit a lower bias. These trends are well-established in wide/deep radio clustering studies (e.g. Lindsay et al. 2014; Magliocchetti et al. 2017; Hale et al. 2018) and are reproduced in our Section 5.3. The bias contrast between SFG and AGN is precisely the lever arm exploited by multi-tracer estimators; as such the presence of the trend in our simulation is of the upmost importance.

3.2 Star-forming galaxy radio luminosity function (SFG RLF)

We require an SFG description that is dust-insensitive and straightforward to propagate through survey selections across metre-centimetre bands. Rather than constructing an SFR function from UV/IR tracers and mapping $\text{SFR} \rightarrow L_\nu$, we calibrate the *radio* luminos-

ity function of SFG directly at 1.4 GHz and model its smooth redshift evolution. This keeps the baseline independent of assumptions about obscured star formation and the FIR-radio correlation, letting the radio data set the space density as a function of L_ν and z .

Relative to T-RECS, we omit the IR/SFRF intermediate (no L_{IR} term, no explicit FIR-radio coupling); GHOST treats SFG emission radio-first. This choice reduces dependence on dust corrections and the redshift evolution of the FIR-radio correlation. The trade-off is that we do not separately track an explicitly IR-luminous subpopulation; any such behaviour is absorbed into the smooth evolution of the radio RLF. As discussed in Section 5.6, this tends to produce a smoother $n(z)$ at fixed flux (less pronounced features around $z \sim 1-2$) while preserving number counts and large-scale clustering.

For the functional form we adopt a smoothly broken (double-power-law) RLF with redshift-dependent normalization and knee luminosity. All parameters are obtained by fitting to 1.4 GHz SFG RLF measurements in redshift slices, then validated against differential counts at 150 MHz (Franzen et al., 2016; Mandal et al., 2021), 1.4 GHz (Bondi et al., 2008; Bridle et al., 1972; Ciardi & Loeb, 2000; Fomalont et al., 2006; Gruppioni et al., 1999; Hopkins et al., 2003; Ibar et al., 2009; Kellermann et al., 2008; Mitchell & Condon, 1985; Owen & Morrison, 2008; Richards, 2000; Seymour et al., 2008; White et al., 1997; de Zotti et al., 2010) and 3 GHz (Smolčić et al., 2017) using the SFG spectral prescription applied elsewhere in this work. Fluxes at other frequencies are generated via standard K -corrections with that spectral model; no infrared (IR) quantities enter the catalogue construction. The explicit evolution law and fitted coefficients are listed in Appendix A.

In the multi-tracer context the SFG branch provides the lower-bias tracer with a broadly lower-redshift kernel set by the radio RLF and survey flux limit. Together with the higher-bias AGN branch this delivers the bias contrast and distinct $P(z)$ kernels needed for cosmic-variance cancellation on ultra-large scales. Calibrating the SFG RLF directly in radio reduces sensitivity to dust-related systematics that could otherwise imprint spurious large-scale structure in $n(z)$ and $w(\theta)$, improving robustness for f_{NL} forecasts.

3.3 AGN population

For each class we adopt the T-RECS double-power-law radio luminosity function with smooth redshift evolution in normalisation and knee luminosity; the functional form is listed in Appendix A. Coefficients are refitted where required to remain consistent with the host assignment on our simulation and with the class mix used in this work; the resulting values and uncertainties are reported in Table 7. Observed flux densities at 150 MHz,

1.4 GHz, and 3 GHz are obtained from L_ν via standard K -corrections using class-dependent spectral-index distributions (flat for FSRQ/BLLac; steep for SSAGN), as in T-RECS. We do not model variability.

Sizes and a soft edge-brightening proxy are assigned to SSAGN to provide a morphology label used only for host weighting (Section A.4); this proxy is converted to a probabilistic HERG/LERG tag to allow overlap between morphology and excitation classes. Photometry is unaffected by these labels.

Hosts are assigned using the abundance-matching scheme defined in Section 4. In brief, AGN are matched to halos within redshift shells using a monotonic, no-replacement map, with class-dependent host probabilities and stellar-mass-weighted excitation fractions inspired by Janssen et al. (2012). Tying the placement to our halo supply stabilises the large-scale bias of each class and ensures that the predicted $w(\theta)$ reflects the simulation’s actual halo population.

Beamed subclasses are handled by drawing jet orientations isotropically and enforcing a maximum viewing angle per class and redshift to reproduce the observed beamed fractions. This yields flat spectra and compact sizes for the beamed subset while leaving the large-scale clustering unchanged.

In summary, we inherit the T-RECS class taxonomy, RLF functional form, and spectral prescriptions, while refitting the RLF coefficients where needed and replacing the original halo mapping with our abundance-matching scheme on the FLAMINGO light-cone.

4 HOST ASSIGNMENT AND ABUNDANCE MATCHING

All clustering results that follow use a host-assigned catalogue: every mock source is attached to a specific halo in the FLAMINGO DMO light-cone. We place this section before the clustering analysis to make explicit what catalogue underlies all $w(\theta)$ and C_ℓ measurements.

Given a survey selection, the requested number of sources in a $(\log L, z)$ bin is

$$\Delta N_{i,j} = \Omega \Phi(L|z_i) \left. \frac{dV}{dz} \right|_{z_i} \Delta \log L_{i,j} \Delta z_i, \quad (1)$$

with Φ the class radio luminosity function (double-power-law with smooth evolution; see Appendix A). For SFG we realise these targets by fitting, per redshift slice i and flux density interval j , a monotone halo-to-luminosity map $L_\nu(M_h; z)$ (Equation 2) with optional log-normal scatter and then rank-ordering halos to luminosities without replacement. For AGN we first draw hosts with mass-weighted HERG/LERG properties (Appendix A), then assign L_ν monotonically within the selected subset to reproduce the class RLFs. When

halos are scarce, we apportion the available hosts across populations in proportion to their requested counts (largest-remainder rule; Section 4.5) so the hosted mixture tracks the selection-driven requested mixture. This alignment of mixture, bias hierarchy, and selection is the key ingredient for robust multi-tracer f_{NL} forecasts.

4.1 Assumptions and limitations

We assign a single central radio source per halo (no explicit satellites). The scatter $\sigma_{\log L}$ is taken to be constant within a shell. HERG/LERG labels are proxies for host-mass weighting rather than strict spectroscopic classes. At the highest redshifts, finite halo supply can saturate the brightest SFG targets; we quantify hosted versus requested counts in Section 5.8.

4.2 Inputs and redshift shells

We partition redshift into disjoint shells: $\mathcal{Z}_i \equiv [z_i, z_{i+1})$ ($i = 0, \dots, N-1$), so that $\bigcup_{i=0}^{N-1} \mathcal{Z}_i = [z_0, z_N)$ and $\mathcal{Z}_i \cap \mathcal{Z}_j = \emptyset$ for $i \neq j$. (the same grid used for $n(z)$ and RLFs). In each shell we compute the comoving shell volume per steradian and the target numbers per luminosity bin and population from the relevant RLF under the survey selection (cf. Equation 1). Candidate hosts are all halos in the shell, each with mass M_h and light-cone position. We assign a stellar-mass proxy via the redshift-dependent $M_\star(M_h, z)$ relation of Aversa et al. (2015).

4.3 SFG branch: monotonic L – M_h abundance matching

For SFG we fit a smooth, monotonic mapping so that the host-assigned SFG luminosity histogram reproduces the radio SFG RLF $\Phi_{\text{SFG}}(L_\nu, z)$ (Section 3.2). We adopt

$$L_\nu(M_h; z) = N(z) \left[\left(\frac{M_h}{M_\star(z)} \right)^\gamma + \left(\frac{M_h}{M_\star(z)} \right)^\beta \right]^{-1}, \quad (2)$$

, where M_h is the halo mass, and $M_\star(z)$, γ, β and $N(z)$ are free parameters. with log-normal scatter $\epsilon \sim \mathcal{N}(0, \sigma_{\log L}^2)$. The parameters $\{N, M_\star, \gamma, \beta, \sigma_{\log L}\}$ are inferred independently in each shell by maximising a Poisson likelihood comparing the target counts (from the RLF) to the counts produced by rank-ordering halos in M_h and mapping to L_ν via Equation 2. We explore the posterior with an ensemble Markov-Chain Monte-Carlo (MCMC) (emcee, Foreman-Mackey et al., 2013) and adopt median parameters for assignment.

4.4 AGN branch: mass-weighted host selection and monotone assignment

For radio AGN we draw hosts with a stellar-mass-dependent probability that encodes HERG/LERG tendencies and then assign luminosities monotonically within the selected subset so that the class RLFs are reproduced.

Each SSAGN is given a stochastic size from DiPompeo et al. (2013) and edge-brightening proxy R_s from Lin et al. (2010); we label $R_s > 0.5$ as FR II and $R_s < 0.5$ as FR I and use this as a soft proxy for excitation class. Following Janssen et al. (2012); Bonaldi et al. (2019), we weight hosts by stellar mass, with a high-mass cap (“saturation”) for the LERG branch:

$$P_{\text{LERG}}(M_\star|z) \propto \begin{cases} (M_\star/M_0)^{2.5}, & M_\star \leq M_{\text{sat}}, \\ (M_{\text{sat}}/M_0)^{2.5}, & M_\star > M_{\text{sat}}, \end{cases} \quad (3)$$

$$P_{\text{HERG}}(M_\star|z) \propto \left(\frac{M_\star}{M_0}\right)^{1.5}, \quad (4)$$

with $M_{\text{sat}} \simeq 10^{11.6} M_\odot$. By “saturating” we mean the LERG triggering weight stops increasing above M_{sat} . We renormalise $P_t(M_\star|z)$ over candidate hosts in each redshift shell and draw hosts so that HERG/LERG totals match the RLF targets for that shell (FSRQ $\rightarrow$ HERG-like; BLLac $\rightarrow$ LERG-like; SSAGN split by R_s). Hosts are sampled without replacement. Within each class and shell we sort hosts by M_h , sort luminosities by L_ν , and map by rank (with optional log-normal L scatter) to reproduce the class RLF while anchoring large-scale bias to the actual halo field. FSRQ/BL Lac orientations are drawn isotropically (uniform in $\cos\theta$).

4.5 Population-weighted host apportionment

Let $\mathcal{L}_t(s)$ be the set of luminosities requested in slice s for population $t \in \{\text{FSRQ}, \text{BL Lac}, \text{SSAGN}, \text{SFG}\}$, with size $R_t(s) = |\mathcal{L}_t(s)|$. Let $K(s)$ be the number of available hosts (unassigned halos) in the slice. If $K(s) \geq R_{\text{tot}}(s) \equiv \sum_t R_t(s)$, all requests are hosted.

When $K(s) < R_{\text{tot}}(s)$, we apportion the $K(s)$ hosts across populations in proportion to requests using a largest-remainder rule:

$$q_t = \frac{R_t(s)}{R_{\text{tot}}(s)} K(s), \quad K_t = \min(R_t(s), \lfloor q_t \rfloor), \quad (5)$$

then distribute the remaining $K(s) - \sum_t K_t$ one by one to populations with the largest fractional remainders $q_t - \lfloor q_t \rfloor$ (never exceeding R_t). This guarantees $\sum_t K_t = \min(K(s), R_{\text{tot}}(s))$ and $K_t \leq R_t$ for all t .

Within each population, sort luminosities descending, $L_{t,(1)} \geq \dots \geq L_{t,(R_t)}$, and map rank to a host-mass

index inside the quota via

$$i(r) = \left\lceil r \frac{K_t}{R_t} \right\rceil \in \{1, \dots, K_t\}. \quad (6)$$

Across all populations we collect these (predicted-mass, population) pairs, sort by predicted mass, and assign the top $K(s)$ to the actual top $K(s)$ halos (strictly preserving mass rank). Leftover requests $R_t - K_t$ are written as “no-halo” rows (*NH) at the slice mid-redshift, preserving per-population totals and making host scarcity explicit.

4.6 Algorithmic summary

In each shell we compute target counts per class and luminosity bin from the relevant RLF (Equation 1) under the survey selection, gather all halos and assign $M_\star(M_h, z)$ from Aversa et al. (2015), and then perform assignment by branch. For SFG we fit $L_\nu(M_h)$ via Equation 2 using an ensemble MCMC on a Poisson likelihood and assign by rank without replacement, adding log-normal scatter. For AGN we draw HERG/LERG hosts using $P_{\text{HERG/LERG}}(M_\star)$ and, within each class, assign L_ν monotonically to match the class RLF. We then attach sky positions from hosts, apply spectral prescriptions and K -corrections, and finally apply the survey selection (flux/SNR survey selection functions and completeness surrogates).

5 INITIAL RESULTS AND CATALOGUE VALIDATION

In this section, we present representative results from the simulated sky catalogue described above. We perform consistency checks against the input luminosity and redshift distributions, assess population statistics, and provide example sky visualizations. We will discuss applications to multi-tracer techniques. The simulated catalogue contains a total of $\sim 2.29 \times 10^9$ galaxies, comprising both AGN and SFG across 4π steradians to a sensitivity of $1 \mu\text{Jy}$ at 1.4 GHz. The AGN population includes $\sim 2.32 \times 10^7$ FSRQ, $\sim 2.08 \times 10^7$ BLLacs, and $\sim 6.60 \times 10^7$ SSAGN, while the SFG population contributes $\sim 2.18 \times 10^9$ sources. The redshift distribution spans the range $z \in [0, \sim 6]$, with the majority of sources concentrated below $z \sim 4$.

We use the same survey selections for both the redshift-distribution $n(z)$ and clustering analyses to enable like-for-like comparisons with the literature. We re-classify for comparison with Magliocchetti et al. (2004) because the rule is a single, explicit luminosity threshold that can be mirrored exactly. We do not re-classify to compare with Hale et al. (2018) because those clustering measurements treat the sample as a mixed population without a uniform AGN/SFG split. For a different reason, we do not re-classify for comparison with Hale

et al. (2018) because their labels are derived from multi-band diagnostics that are not available uniformly in our mock; we therefore emulate only the detection flux density threshold and separate population-mix effects (our classes) from selection effects (their S/N cut).

5.1 Selections used for validation

This Section defines the cuts and clarifies how each survey separated AGN from SFG (where they did), and how we mirrored those choices in GHOST. The three emulations are: the VLA-COSMOS 1.4 GHz luminosity threshold used by Magliocchetti et al. (2017); the LoTSS-DR2 (Hale et al., 2024; Shimwell et al., 2022a) wide-area 150 MHz flux and peak-S/N survey selection functions; and the VLA-COSMOS 3 GHz peak-S/N threshold and multi-wavelength classes of Hale et al. (2018). Table 2 summarises the survey selection functions and our emulation strategy. As emphasised in Section 5.3, these selections compress or broaden the redshift kernel $n(z)$ and thus reshape the projected clustering amplitude/slope without altering the intrinsic bias ordering, so we present raw and selected catalogues side by side in what follows.

5.1.1 LoTSS-DR2 wide 150 MHz (Shimwell et al., 2022b; Hale et al., 2024).

We emulate the LoTSS selection by combining an integrated-flux cut $S_{150} \geq 1.5$ mJy with a peak-S/N threshold $S/N_{\text{peak}} \geq 7.5$. The image rms varies across the mosaic due to multiplicative effects, such as primary beam sensitivity, mosaicking overlap, and calibration/dynamic range, so we model the local rms σ as log-normal in an effort to replicate this, $\ln \sigma \sim \mathcal{N}(\mu_{\ln}, \sigma_{\ln}^2)$. Anchoring the distribution to the median $\tilde{\sigma} = 83 \mu\text{Jy beam}^{-1}$ and 95th percentile $171 \mu\text{Jy beam}^{-1}$ gives

$$\mu_{\ln} = \ln \tilde{\sigma}, \quad \sigma_{\ln} = \frac{\ln(171 \mu\text{Jy}) - \mu_{\ln}}{1.645}, \quad (7)$$

which reproduces the observed quantiles by construction. For a source of integrated flux S (we take $S_{\text{peak}} \approx S$ at 1.5 mJy; for resolved sources this is conservative), the S/N cut is equivalent to $\sigma \leq \sigma_{\text{thr}} = S/7.5$, yielding a detection probability

$$p_{\text{keep}}(S) = \Phi\left(\frac{\ln \sigma_{\text{thr}} - \mu_{\ln}}{\sigma_{\ln}}\right), \quad (8)$$

with Φ the standard normal CDF. Using the class counts dN_X/dS , we define the area-weighted acceptances

$$\text{acc}_X = \frac{\int_{S_{\min}}^{\infty} \frac{dN_X}{dS}(S) p_{\text{keep}}(S) dS}{\int_{S_{\min}}^{\infty} \frac{dN_X}{dS}(S) dS}, \quad X \in \{\text{SFG}, \text{AGN}\}, \quad (9)$$

and form the effective mixture entering the total $n(z)$,

$$w_{\text{SFG}} = \frac{\text{acc}_{\text{SFG}} N_{\text{SFG}}}{\text{acc}_{\text{SFG}} N_{\text{SFG}} + \text{acc}_{\text{AGN}} N_{\text{AGN}}}, \quad (10)$$

where $S_{\min} = 1.5$ mJy and N_X are class totals from the un-gated $n(z)$. At $S_{150} = 1.5$ mJy the flux threshold removes only a few percent of sources in both classes, so the resulting $n(z)$ is dominated by the intrinsic SFG/AGN mix (Figure 9). Hale et al. (2024) does not supply a homogeneous AGN/SFG split for this clustering sample, so we emulate the selection only and compare the *total* $n(z)$ and clustering; by-class predictions shown elsewhere use our internal labels to isolate population-mix effects from detection selection.

5.1.2 VLA-COSMOS 1.4 GHz (Magliocchetti et al., 2017).

We mirror the flux density limit $S_{1.4\text{GHz}} \geq 0.15$ mJy and re-classify sources using their redshift-dependent radio-luminosity boundary. Radio powers are computed as $P_{1.4} = F_{1.4} D_A^2 (1+z)^{3+\alpha}$ (W Hz⁻¹ sr⁻¹) with $\alpha = 0.7$ and D_A the angular diameter distance, and the AGN/SFG division is

$$\log_{10} P_{\text{cross}}(z) = \begin{cases} 21.7 + z, & z \leq 1.8 \\ 23.5, & z > 1.8 \end{cases} \quad (11)$$

(we convert to isotropic luminosity via $L_{\text{cross}} = 4\pi P_{\text{cross}}$). Objects with $L_{1.4} \geq L_{\text{cross}}(z)$ are labelled “AGN-like”, the rest “SFG-like”. This reproduces their threshold classifier exactly and makes the comparison sensitive to population mix and redshift tails rather than to differences in detection survey selection functions; it also emphasises the overlapping mJy regime where both classes contribute (see Figure 10).

5.1.3 VLA-COSMOS 3 GHz (Hale et al., 2018).

We emulate the catalogue’s peak-flux threshold $S/N \geq 5.5$ with a field median rms $\tilde{\sigma} \simeq 2.3 \mu\text{Jy beam}^{-1}$, modelling field-to-field rms as lognormal with

$$\mu_{\ln} = \ln \tilde{\sigma}, \quad \sigma_{\ln} = 0.30. \quad (12)$$

For a source of flux density S , the rms required to pass is $\sigma_{\text{thr}} = S/5.5$; the area fraction that admits detection is

$$p_{\text{keep}}(S) = \Phi\left(\frac{\ln(S/5.5) - \mu_{\ln}}{\sigma_{\ln}}\right). \quad (13)$$

We then compute a class-wise, area-weighted acceptance by averaging p_{keep} over the per-population 3 GHz counts histogram $H_X(S_i)$ (from the simulation),

$$\text{acc}_X = \frac{\sum_i H_X(S_i) p_{\text{keep}}(S_i)}{\sum_i H_X(S_i)}, \quad X \in \{\text{SFG}, \text{AGN}\}, \quad (14)$$

and form the effective mixture

$$w_{\text{SFG}} = \frac{\text{acc}_{\text{SFG}} N_{\text{SFG}}}{\text{acc}_{\text{SFG}} N_{\text{SFG}} + \text{acc}_{\text{AGN}} N_{\text{AGN}}}. \quad (15)$$

At the Hale depth the selection function is weak; in our run the acceptances are close to unity for both classes (numerically $\text{acc}_{\text{SFG}} \approx 0.89$, $\text{acc}_{\text{AGN}} \approx 0.94$), so the predicted $n(z)$ is governed mainly by the intrinsic SFG/AGN mix. Because the Hale classification relies on deep multi-wavelength diagnostics (X-ray, MIR colour, UV-IR SED, rest-frame NUV colour, and radio excess relative to SFR), we do not attempt to reproduce their labels inside GHOST; we emulate only the S/N selection with a final flux cut of $12.5 \mu\text{Jy}$ at 3 GHz and compare the *total* $n(z)$, using our internal classes where a by-class illustration is helpful.

5.2 Abundance-matching outcomes

All results that follow use the host-assigned catalogue from Section 4: every mock source is either attached to a specific FLAMINGO halo or, if the slice's host supply is exhausted, written as a "no-halo" (*NH) row. When halos are scarce we apportion the available hosts across populations in proportion to their requested counts (largest-remainder rule; Section 4.5) and then assign luminosities monotonically within each population so that the class RLFs are reproduced.

At fixed flux limit the accessible radio luminosity rises with redshift,

$$L_{\text{lim}}(z, \zeta) = 4\pi D_L^2(z) S_{\text{lim}} (1+z)^{\alpha-1}, \quad (16)$$

so the observed (L_ν, z) domain is bounded jointly by Equation 16 and the evolving halo mass function. With the corrected slice projection, both the abundance-matching targets and the host-assigned catalogue populate this domain smoothly across all slices.

Figure 2 summarizes the delivered mapping. Star-forming galaxies occupy lower-mass halos and lower L_ν ; steep-spectrum AGN extend to higher M_h with a steeper L - M trend; beamed subclasses (FSRQ, BLLac) preferentially inhabit the most massive halos and dominate the bright tail. With increasing redshift, median ridge-lines shift as expected from K -corrections and the evolving halo supply, and the per-population marginals narrow modestly.

We track requested versus hosted counts per slice and population (Section 4.5). The hosted fraction is high across all panels; occasional *NH rows appear only at very low z and L_ν and are negligible for the selections used. In the RLF grids we show the model projected on the same (L, z) grid alongside the host-assigned estimate; the two agree within uncertainties over the full dynamic range.

The host-assigned catalogue embeds a clear bias hierarchy—SFG < SSAGN < FSRQ/BL Lac—set by

the actual halo field under the survey selection. Because our population-weighted apportionment keeps the hosted mixture close to the requested mixture for any flux/SNR cut, changes in $w(\theta)$ across selections reflect projection kernels rather than a reversal of intrinsic bias ordering. This alignment of mixture, bias hierarchy, and selection is exactly what multi-tracer f_{NL} estimators require.

5.3 The angular correlation function

Our goal in this section is to verify that the catalogue places each population in environments consistent with their expected large-scale bias, and to produce clustering summaries that feed the PNG analysis in Paper 2. We therefore measure the two-point angular correlation of the simulated sky, both in configuration space, $w(\theta)$, and in harmonic space, C_ℓ . Given a redshift kernel $\phi(z)$ (from the $n(z)$ of each population), these observables probe the projection of the 3D matter power spectrum modulated by the tracers' bias, so their amplitudes and slopes are sensitive to the typical host halo masses of SFGs versus AGN subclasses and how survey selections reshape $\phi(z)$.

Concretely, we compute $w(\theta)$ for AGN and SFG samples (and for specific survey-like cuts), fit a compact power-law model $w(\theta) = A(\theta/\theta_0)^{1-\gamma}$ over 0.02° – 1° to quote comparable amplitudes and slopes, and cross-check in harmonic space by fitting C_ℓ with fixed $\phi(z)$ templates to infer effective linear bias (both constant and gently evolving with redshift).

5.3.1 Definitions and modelling

We quantify angular clustering with the two-point correlation function $w(\theta)$, related to the angular power spectrum C_ℓ by

$$w(\theta) = \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} C_\ell P_\ell(\cos \theta) \simeq \int \frac{\ell d\ell}{2\pi} J_0(\ell\theta) C_\ell, \quad (17)$$

where the Bessel form is accurate for small angles. In the Limber regime, the auto-spectrum for a population $X \in \{\text{AGN}, \text{SFG}\}$ is

$$C_\ell^{XX} = \int dz \underbrace{\frac{H(z)}{\chi^2(z)} \phi_X^2(z)}_{\text{projection}} \underbrace{b_X^2(k, z) D^2(z)}_{\text{bias/growth}} \times P_m\left(k = \frac{\ell+1/2}{\chi(z)}, z\right), \quad (18)$$

with comoving distance $\chi(z)$, Hubble rate $H(z)$, linear growth $D(z)$, matter spectrum P_m , and a normalized redshift kernel $\phi_X(z)$ (unit area).

Across a finite angular decade the observed slope of $w(\theta)$ reflects a weighted average of the 3D logarithmic

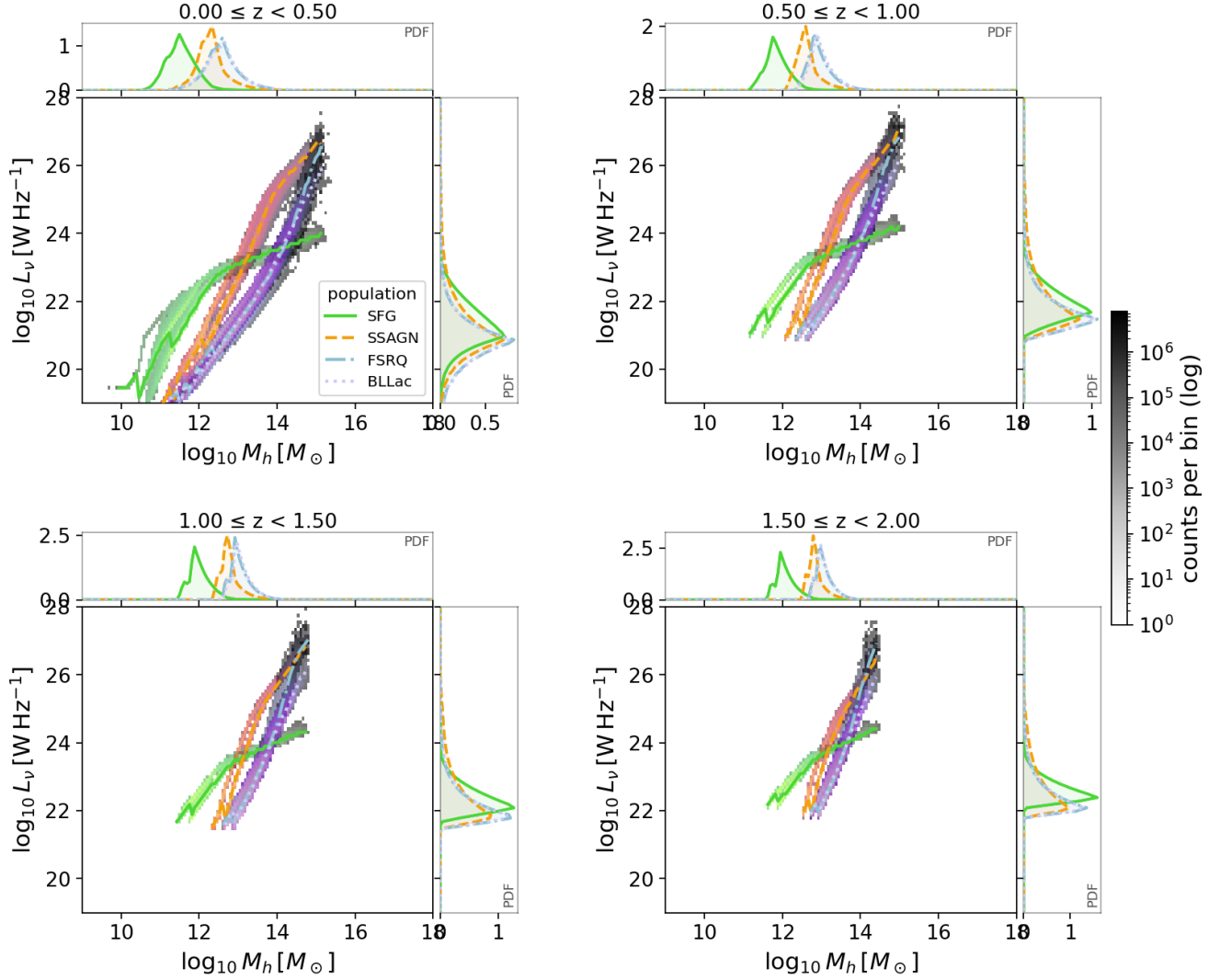


Figure 2. All populations — radio luminosity vs. halo mass with per-population marginals (shared density scale). Each panel shows 2D binned abundances in a fixed redshift window; colored semi-transparent maps overlay populations (SFG *solid*, SSAGN *dashed*, FSRQ *dash-dot*, BL Lac *dotted*). All density layers share the same logarithmic colorbar, so intensity is directly comparable across populations and panels. Solid lines trace the median (“ridgeline”) of $\log_{10} L_\nu$ at fixed $\log_{10} M_h$. The small top (right) axes show, for each population, the *normalized* marginal PDFs $p(\log_{10} M_h)$ and $p(\log_{10} L_\nu)$ —with $\int p d \log M_h = 1$ and $\int p d \log L = 1$ —using the same color/linestyle as the ridgelines. *Reading the figure:* SFG peak at lower host masses and luminosities; SSAGN extend to higher M_h with a steeper L_ν – M_h trend; FSRQ and BL Lac preferentially inhabit the most massive halos and dominate the bright radio tail. With increasing redshift, ridgelines shift to higher M_h and the PDFs narrow, reflecting both selection effects and the evolving halo-occupation mix.

Table 2 Classification schemes of comparison surveys and the emulation used in this work.

Dataset	Selection	Their classification	Our emulation
VLA-COSMOS, 1.4 GHz (Magliocchetti et al., 2017)	$F_{1.4} \geq 0.15$ mJy	Luminosity split: $L_{1.4} \geq L_{\text{cross}}(z)$ with $\log L_{\text{cross}}(z) = 21.7 + z$ (to $z \leq 1.8$), else 23.5	We apply the same $L_{1.4}$ vs $L_{\text{cross}}(z)$ split to GHOST
LoTSS DR2, 150 MHz (Shimwell et al., 2022b)	$S_{150} \geq 1.5$ mJy + peak S/N threshold	LoTSS DR2 per-catalogue classes (data-driven mix)	We emulate the survey selection function
VLA-COSMOS 3 GHz (Hale+18)	$S/N_{\text{peak}} \geq 5.5$	Multi-wavelength (SED/line) AGN and SFG classes	We emulate the S/N threshold and compare total and by-class using our internal classes

slope $d \ln[P_m b_X^2]/d \ln k$ evaluated at $k \simeq \ell/\chi(z)$, convolved with ϕ_X^2 . For direct comparison to the literature and to quote compact summary numbers we fit a power law

$$w(\theta) = A \left(\frac{\theta}{\theta_0} \right)^{1-\gamma}, \quad (19)$$

using full covariances where available and a pivot θ_0 taken to be the geometric mean of the fitted θ to minimise the A - γ covariance. Unless stated otherwise we fit over $\theta \in [0.02^\circ, 1.0^\circ]$ ($\ell \simeq 50$ –400).

To interpret slope changes induced by selections, we also quote *effective* redshifts and distances computed with Limber-like weights

$$W_X(z) \propto \frac{H(z)}{\chi^2(z)} \phi_X^2(z) D^2(z), \quad (20)$$

$$z_{\text{eff},X} = \frac{\int W_X(z) z dz}{\int W_X(z) dz}, \quad (21)$$

$$\chi_{\text{eff},X} = \frac{\int W_X(z) \chi(z) dz}{\int W_X(z) dz}. \quad (22)$$

A fixed angular separation then corresponds to a characteristic projected comoving scale

$$r_\perp(\theta) = \theta \chi_{\text{eff},X}. \quad (23)$$

5.3.2 Raw GHOST sample.

In Figure 3, we present the angular correlation function of the full GHOST sample showing AGN lying above SFG in amplitude across all θ , consistent with their higher bias and host-halo masses. Power-law fits give $\gamma_{\text{AGN}} = 1.886 \pm 0.028$ and $\gamma_{\text{SFG}} = 1.647 \pm 0.076$. The effective depths are $(z_{\text{eff}}, \chi_{\text{eff}}) = (0.318, 1.22 \text{ Gpc})$ for AGN and $(0.602, 2.12 \text{ Gpc})$ for SFG, so $r_\perp(0.05^\circ) = (1.06, 1.85) \text{ Mpc}$, respectively. The shallower SFG slope in the raw GHOST sample is naturally explained by their broader, higher- z kernel $\phi(z)$, which mixes a wider range of $k = \ell/\chi$ in projection (Equations 18–22).

5.3.3 VLA-COSMOS 1.4 GHz

Applying the $L_{1.4} \geq L_{\text{cross}}(z)$ selection, as detailed in Section 5.1.2, compresses the SFG kernel to very low z . We find $\gamma_{\text{AGN}} = 1.858 \pm 0.041$ and $\gamma_{\text{SFG}} = 1.667 \pm 0.010$, with $(z_{\text{eff}}, \chi_{\text{eff}}) = (0.339, 1.30 \text{ Gpc})$ for AGN and $(0.094, 0.41 \text{ Gpc})$ for SFG, implying $r_\perp(0.05^\circ) = (1.14, 0.35) \text{ Mpc}$. In Fig. 4, the SFG correlation lies above the AGN over the fitted range. This does not imply a larger intrinsic SFG bias; it follows from the projection weights in the Limber integral, which scale as $\phi^2(z)/\chi^2(z)$. Our selection compresses the SFG kernel to very low redshift ($z_{\text{eff}} \simeq 0.094$, $\chi_{\text{eff}} \simeq 0.41 \text{ Gpc}$) while the AGN kernel sits higher ($z_{\text{eff}} \simeq 0.339$, $\chi_{\text{eff}} \simeq 1.30 \text{ Gpc}$). The resulting geometric factor $(\chi_{\text{AGN}}/\chi_{\text{SFG}})^2 \simeq (1.30/0.41)^2 \approx 10$ boosts the projected SFG amplitude and can outweigh the larger AGN bias, thereby producing the observed inversion in $w(\theta)$. The residual shape mismatch — most evident in the flatter SFG slope — likely reflects blending of one- and two-halo terms when projecting such a low- z kernel over 0.02° – 1° ($r_\perp \sim 0.14$ – 7 Mpc), and modelling choices that suppress small-scale power (e.g. our current radio-SFG occupation lacks an explicit satellite component). These effects account for the difference in gradient while preserving the relative bias hierarchy that underpins the multi-tracer analysis.

5.3.4 VLA-COSMOS 3 GHz

The same pattern repeats at 3 GHz, with the selection described in Section 5.1.3: $\gamma_{\text{AGN}} = 1.903 \pm 0.037$ and $\gamma_{\text{SFG}} = 1.662 \pm 0.010$, with $(z_{\text{eff}}, \chi_{\text{eff}}) = (0.313, 1.21 \text{ Gpc})$ for AGN and $(0.161, 0.67 \text{ Gpc})$ for SFG, hence $r_\perp(0.05^\circ) = (1.05, 0.59) \text{ Mpc}$. Relative to the Hale et al. 3 GHz bands our SFG points lie higher and show a slightly shallower slope, eventually overtaking the AGN at large angles. This behavior is expected from the different redshift kernels: the SFG selection sits

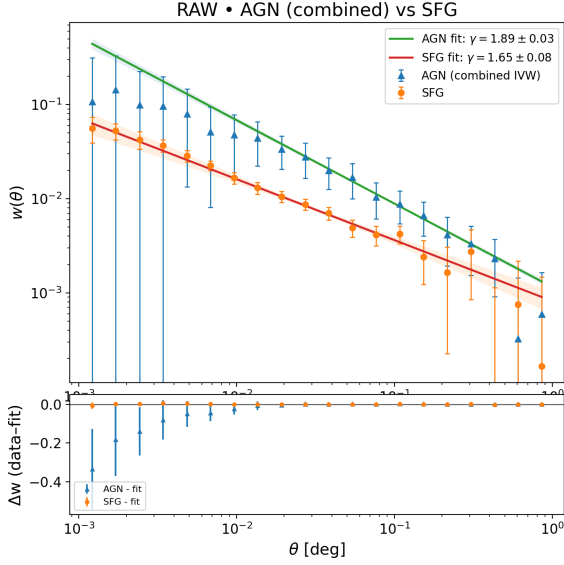


Figure 3. raw GHOST sample: Angular correlation function $w(\theta)$ for AGN (blue triangles) and SFG (orange circles). Solid lines show inverse-variance weighted power-law fits $w(\theta) = A(\theta/\theta_0)^{1-\gamma}$ over $\theta \in [0.02^\circ, 1.0^\circ]$; shaded bands are the propagated 1σ model uncertainties from the (A, γ) covariance. Best-fit slopes are $\gamma_{\text{AGN}} = 1.886 \pm 0.028$ and $\gamma_{\text{SFG}} = 1.647 \pm 0.076$. The lower panel shows data–fit residuals. The amplitude ordering AGN>SFG reflects the intrinsic bias hierarchy, while the slightly shallower SFG slope is consistent with a broader, higher- z kernel $\phi(z)$ that mixes more $k = \ell/\chi(z)$ in projection (Equations 17–18).

at lower χ_{eff} , so the projection integral schematically $w(\theta) \propto \int dz \phi^2(z) b^2(z) D^2(z)/\chi^2(z)$ weights it more strongly, and the same angular range probes smaller $r_\perp$, mixing some one-halo power into the fitted scales and flattening the apparent slope. By contrast, the AGN kernel projects to larger $r_\perp$ where the two-halo term dominates and the slope is steeper. Residual differences with the Hale et al. (2018) bands are therefore plausibly explained by modest kernel mismatches (e.g. a slight SFG overweight at $z \lesssim 0.3$ in the mocks), survey-specific treatments of integral-constraint and close pairs/multi-component sources that suppress small-angle power, and minor AGN/SFG cross-contamination near the luminosity split. The cross-frequency consistency (compare 1.4 GHz) supports the interpretation that these are selection-driven projection effects rather than frequency-dependent systematics.

5.3.5 Picture across frequencies

The differences we see across the raw, Magliocchetti and Hale-like samples are driven primarily by how each selection reshapes the redshift kernel $\phi(z)$, not by any change in the intrinsic bias ordering. In Equation 18 the amplitude scales as $b^2 \phi^2/\chi^2$ and the slope reflects which 3D wavenumbers are projected ($k \simeq \ell/\chi$) given

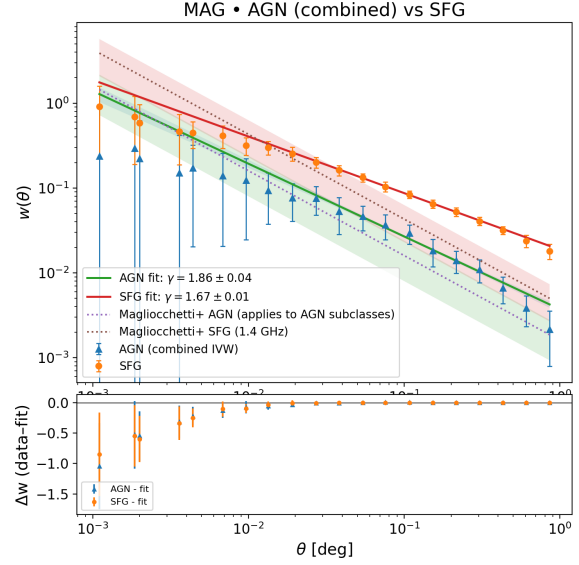


Figure 4. VLA-COSMOS selection (1.4 GHz): $w(\theta)$ for AGN and SFG after the $L_{1.4} \geq L_{\text{cross}}(z)$ selection function detailed in Section 5.1.2. Solid curves are our power-law fits over $\theta \in [0.02^\circ, 1.0^\circ]$ (shaded 1σ bands); dotted lines and translucent bands show the literature reference for AGN and SFG, respectively. We obtain $\gamma_{\text{AGN}} = 1.858 \pm 0.041$ and $\gamma_{\text{SFG}} = 1.667 \pm 0.010$. The SFG amplitude lies above the reference at large scales and the slope is slightly flatter. From the effective-depth summary (Table 3) the SFG kernel compresses to $z_{\text{eff}} \simeq 0.094$ and $\chi_{\text{eff}} \simeq 0.41$ Gpc, so a fixed θ probes small $r_\perp$ (e.g. $r_\perp(0.05^\circ) \approx 0.35$ Mpc), flattening the angular scaling (Equations 22–23).

the kernel. Cuts that compress the SFG kernel to lower redshift (e.g. luminosity splits or high S/N survey selection functions) increase ϕ^2/χ^2 and shift the map toward larger k , yielding the observed “lift and flattening” of SFG $w(\theta)$. AGN kernels move much less under the same cuts, so the intrinsic hierarchy $b_{\text{AGN}} > b_{\text{SFG}}$ remains intact.

This explains the apparent discrepancy between the Magliocchetti and Hale measurements: once their selection functions are emulated, both behaviours arise naturally from projection effects with a fixed bias ordering. The cross-frequency consistency (1.4 vs 3 GHz) further shows the effect is selection-driven rather than band-dependent. For multi-tracer f_{NL} applications, the lesson is operational: clustering inferences must be made with tracer-specific $\phi(z)$ that match the actual survey selection functions; otherwise slope and amplitude shifts can be misread as bias changes. Here we carry forward bias parameters together with the matched kernels to ensure the large-scale modes relevant for f_{NL} are interpreted consistently. The statistics of the angular correlation function (ACF) are summarized in Table 3.

Table 3 Projection summary for AGN and SFG by selection. Errors are 1σ from power-law fits over $\theta \in [0.02^\circ, 0.5^\circ]$. Effective depths use the Limber-like weights of Equation (22). We quote the characteristic projected comoving separation $r_\perp(\theta) = \theta \chi_{\text{eff}}$ at three angles.

Cut	Group	γ	z_{eff}	χ_{eff} [Gpc]	$r_\perp(0.02^\circ)$ [Mpc]	$r_\perp(0.05^\circ)$ [Mpc]	$r_\perp(0.10^\circ)$ [Mpc]
raw GHOST sample	AGN	1.886 ± 0.028	0.318	1.220	0.426	1.065	2.130
	SFG	1.647 ± 0.076	0.602	2.117	0.739	1.847	3.695
Magliocchetti (1.4 GHz)	AGN	1.858 ± 0.041	0.339	1.303	0.455	1.137	2.275
	SFG	1.667 ± 0.010	0.094	0.406	0.142	0.354	0.708
Hale-like (3 GHz)	AGN	1.903 ± 0.037	0.313	1.208	0.422	1.054	2.108
	SFG	1.662 ± 0.010	0.161	0.674	0.235	0.588	1.176

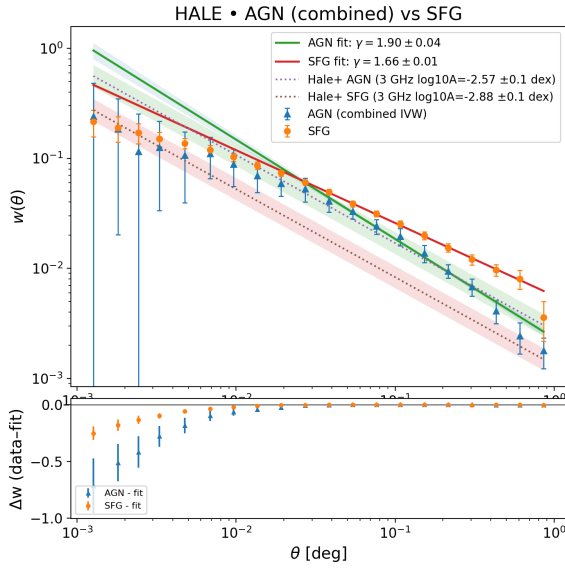


Figure 5. VLA-COSMOS selection (3 GHz): Same as Fig. 4 but for the 3 GHz cut (Section 5.1.3). Our fits give $\gamma_{\text{AGN}} = 1.903 \pm 0.037$ and $\gamma_{\text{SFG}} = 1.662 \pm 0.010$. AGN closely follow the band from Hale et al. (2018), while SFG sit high and slightly flatter. The projection summary (Table 3) shows $z_{\text{eff}, \text{SFG}} \simeq 0.161$ and $\chi_{\text{eff}, \text{SFG}} \simeq 0.67$ Gpc (so $r_\perp(0.05^\circ) \approx 0.59$ Mpc), again explaining the slope change as a selection-driven projection effect rather than frequency-dependent systematics. Residuals (lower panel) are shown relative to the fitted power laws.

5.4 Angular power spectra and bias fits

We model the auto-spectra of each radio population $X \in \{\text{SFG}, \text{SSAGN}, \text{BLLac}, \text{FSRQ}\}$ in the Limber regime using Equation 18. We construct, for each tracer, a *template* T_ℓ^X by evaluating equation 18 with a unit, redshift-independent bias (i.e. $b(z) \equiv 1$). The measured spectrum in a given ℓ -band then determines an overall amplitude related to the bias.

We fit two bias models over $\ell \in [\ell_{\min}, \ell_{\max}]$ (default 40–250), after sanitizing $n(z)$ (Section 5.6) and optionally subtracting a shot component from the auto-correlations when available.

5.4.1 Constant bias

We assume $b_X(z) = b_X$ and fit an amplitude $a_X = b_X^2$ against the fixed template T_ℓ^X . Using (optionally) Knox weights $w_\ell = [2/((2\ell + 1)f_{\text{sky}})]/\sigma_\ell^2$ or per-band errors we solve the weighted least-squares problem

$$\hat{a}_X = \frac{\sum_\ell w_\ell T_\ell^X C_\ell^{\text{obs}}}{\sum_\ell w_\ell (T_\ell^X)^2}, \quad \hat{b}_X = \sqrt{\hat{a}_X}, \quad (24)$$

with error propagation $\sigma_a^2 \simeq 1/\sum_\ell w_\ell (T_\ell^X)^2$ and $\sigma_b \simeq \sigma_a/(2\hat{b})$. Goodness of fit is summarized by $\chi^2 = \sum_\ell w_\ell (C_\ell^{\text{obs}} - \hat{a} T_\ell^X)^2$.

5.4.2 Evolving bias

We test a minimal two-parameter form

$$b_X(z) = b_{0,X} (1+z)^{\zeta_X}. \quad (25)$$

Within our pipeline this redshift dependence is folded into the *shape* of the projection kernel by replacing $\phi_X(z) \rightarrow \phi_X(z) (1+z)^{\zeta_X}$ (and re-normalizing to unit area), while keeping the overall amplitude as b_0^2 . For each tracer we form the profile likelihood in ζ by optimizing b_0 at fixed ζ using Equation 24, i.e. $\hat{b}_0(\zeta) = \arg \min_{b_0} \chi^2(b_0, \zeta)$. We then select $(\hat{b}_0(\hat{\zeta}), \hat{\zeta})$ with $\hat{\zeta} = \arg \min_\zeta \chi^2(\hat{b}_0(\zeta), \zeta)$.

5.4.3 Results

Table 4 summarises the fits. In the constant model we find the expected hierarchy $b_{\text{SFG}} < b_{\text{SSAGN}} \lesssim b_{\text{FSRQ}} \simeq b_{\text{BLLac}}$. The evolving model always yields a slightly lower χ^2 statistic, indicating better fit, with (ζ, b_0) increasing from SFG to radio-loud AGN, consistent with rarer, more strongly biased hosts at higher redshift. Figure 6 shows the measured bandpowers with the best-fit constant (dashed) and evolving (solid) curves.

5.4.4 Implications

The bias ordering mirrors the halo-mass distributions inferred from the catalogue: SFG peak at lower masses ($\sim 10^{12} M_\odot$) while AGN subclasses peak at higher masses (few $\times 10^{12.5-13} M_\odot$), naturally producing larger large-scale clustering amplitudes. The mild preference

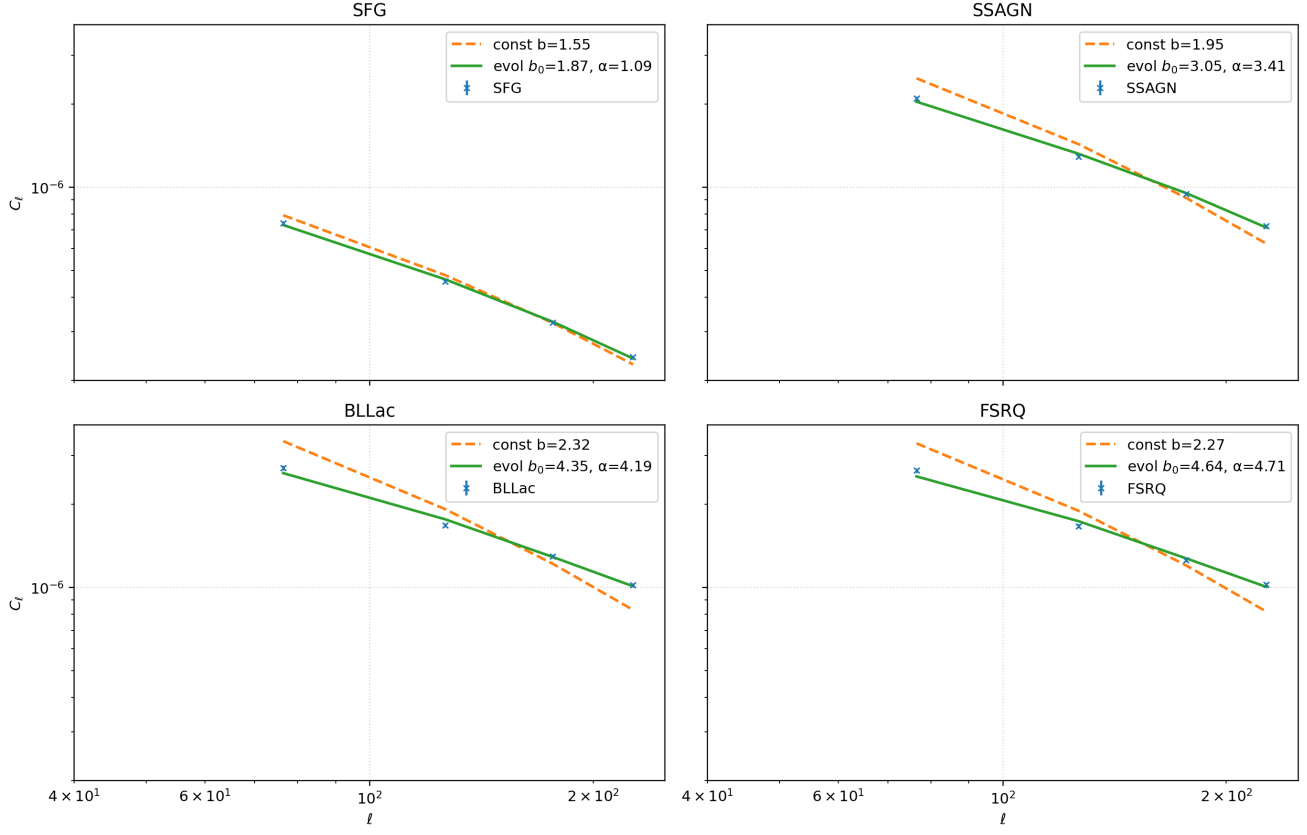


Figure 6. C_ℓ bias fits by population. Points show measured bandpowers; dashed lines are best-fit constant-bias models ($b = \text{const}$) and solid lines are evolving models $b(z) = b_0(1+z)^\zeta$, both fit over $\ell \in [40, 250]$. Numerical values are listed in Table 4. The amplitude ordering $\text{SFG} < \text{SSAGN} \lesssim \text{FSRQ} \simeq \text{BLLac}$ and the positive ζ for AGN are consistent with the halo-mass distributions and the configuration-space results (Section 5.3).

for positive ζ in AGN ($\zeta \sim 3\text{--}5$) indicates bias growth with redshift, aligning with the picture of AGN inhabiting rarer halos at earlier times. Overall, the C_ℓ fits and the configuration-space results (Section 5.3) are mutually consistent: amplitudes trace b^2 while the $w(\theta)$ slope variations are explained by projection effects rather than changes in the underlying bias hierarchy.

5.5 Host halo mass distributions

To connect clustering amplitudes with halo occupation, we examine the distribution of host halo masses M_h for each population. Figure 7 shows per-population PDFs of $\log_{10}(M_h/M_\odot)$ (normalized per dex). Table 5 lists the median, the central 68% interval, and the sample size.

The ordering is clear: SFG inhabit lower-mass halos (median $\log_{10} M_h \simeq 11.97$), while all AGN subclasses peak near $\sim 10^{12.6\text{--}13.0} M_\odot$ (SSAGN $\simeq 12.65$, BLLac $\simeq 12.95$, FSRQ $\simeq 12.92$). The interquartile widths (IQR) are similar ($\sim 0.63\text{--}0.68$ dex), indicating that the dominant difference is a shift in typical mass rather than a change in spread.

Because the large-scale bias $b(M, z)$ is a monotonic

function of halo mass (e.g. in peak-background split models),

$$b(M, z) = 1 + \frac{a\nu^2 - 1}{\delta_c} + \frac{2p}{\delta_c [1 + (a\nu^2)^p]}, \quad (26)$$

$$\nu \equiv \frac{\delta_c}{\sigma(M, z)},$$

(with $\delta_c \simeq 1.686$ and $(a, p) \approx (0.707, 0.3)$ as a common choice), the $\sim 0.7\text{--}1.0$ dex upward shift from SFG to radio-loud AGN naturally implies larger b at the redshifts that dominate our projection. This is precisely what we measure from C_ℓ in Section 5.4: the constant-bias fits obey $b_{\text{SFG}} < b_{\text{SSAGN}} \lesssim b_{\text{FSRQ}} \simeq b_{\text{BLLac}}$, and the evolving model prefers positive ζ for AGN, consistent with rarer halos at earlier times.

Overall, the mass PDFs corroborate the bias/clustering hierarchy found in both configuration and harmonic space: AGN populate more massive halos than SFG, hence exhibit higher large-scale clustering amplitudes, while the small changes in $w(\theta)$ slope across selections are driven by projection (Section 5.3) rather than a reversal of the intrinsic bias-mass ordering.

Table 4 Bias fits from C_ℓ over $\ell \in [40, 250]$. We fit a constant-bias model ($b = \text{const}$) and an evolving form $b(z) = b_0(1+z)^\zeta$. Quoted χ^2 values are for the best fit of each model. The last column gives the constant-bias ratio b/b_{SFG} .

Tracer	b (const)	χ^2_{const}	b_0 (evol)	ζ	χ^2_{evol}	b/b_{SFG}
SFG	1.55 ± 0.04	0.7	1.87 ± 0.05	1.09	0.0	1.00
SSAGN	1.95 ± 0.06	4.2	3.05 ± 0.09	3.41	0.1	1.26
BLLac	2.32 ± 0.07	8.1	4.35 ± 0.12	4.19	0.2	1.50
FSRQ	2.27 ± 0.07	8.5	4.64 ± 0.13	4.71	0.3	1.46

Table 5 Summary of host halo mass distributions per population. We report the sample size N , the median $\text{med}[\log_{10} M_h/M_\odot]$, the central 68% interval (16th–84th percentiles), the IQR width, and the median offset relative to SFG.

Population	N	$\text{med}[\log_{10} M_h]$	[16, 84] perc.	IQR width	Δ med vs SFG
SFG	2184221793	11.97	11.64–12.27	0.63	0.00
SSAGN	65986633	12.65	12.31–12.96	0.65	0.69
BLLac	20794032	12.95	12.59–13.27	0.68	0.8
FSRQ	23240345	12.92	12.56–13.24	0.68	0.95

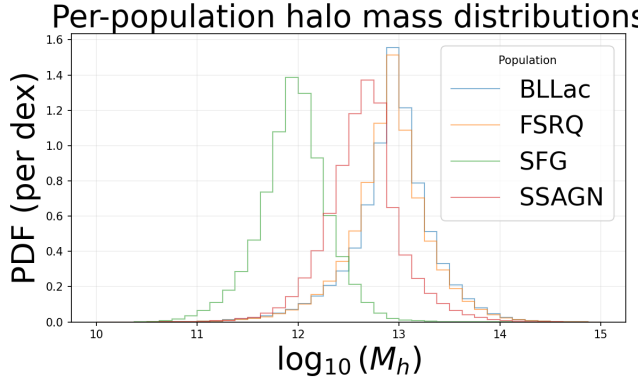


Figure 7. Host halo mass PDFs by population. Normalized distributions of $\log_{10}(M_h/M_\odot)$ (per dex) for SFG, SSAGN, BLLac, and FSRQ. Distributions for the AGN subclasses peak near $10^{12.8-13} M_\odot$, while SFG peak near $10^{12} M_\odot$. The mass ordering maps onto the bias hierarchy measured in Section 5.4.

5.6 Redshift distributions $n(z)$

Figure 8 shows normalized redshift distributions for the full GHOST catalogue and for AGN and SFG subsamples (area under each curve equals unity). Before any survey selection, the total population peaks near $z \sim 0.8$ with a long tail to $z > 3$ (median $z = 1.21$, IQR 0.74–1.81). The AGN distribution is narrower (median $z = 0.85$, IQR 0.50–1.31), while SFG peak at slightly higher redshift (median $z = 1.23$, IQR 0.75–1.83) and maintain the broader high- z tail. This is consistent with radio-loud AGN dominating the bright counts at low z and SFG contributing increasingly toward the sub-mJy regime.

At our pre-selection stage with the raw GHOST catalogue, the $n(z)$ for SFG and AGN look consistent with the behavior from deep μJy radio surveys: SFG dominate at faint fluxes and exhibit a broad distribution that

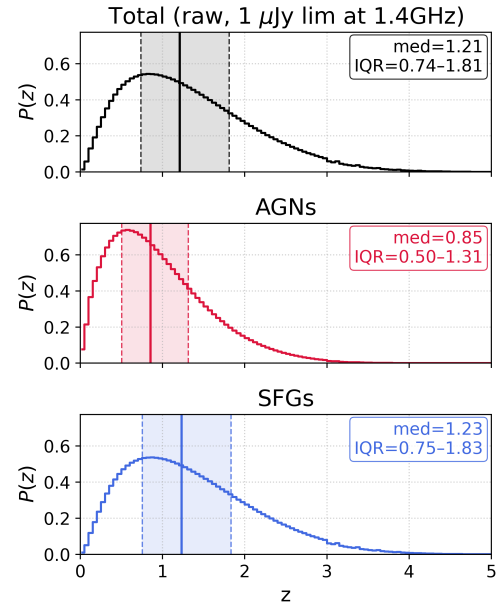


Figure 8. Normalized PDFs $P(z)$ (unit area) for the full GHOST catalogue (top) and by class (AGN, SFG; middle/bottom). Vertical lines mark medians (solid) and IQRs (dashed). The total peaks near $z \sim 0.8$ with a long high- z tail, establishing the baseline cosmic distribution before survey selection.

extends to $z \approx 4$, while AGN show a more concentrated, lower z peak. Similar behavior is seen in the LOFAR Two-metre Sky Survey (LoTSS) Deep fields (Best et al., 2023; Bonzini et al., 2013; Smolčić et al., 2017). Disagreements that emerge once survey cuts are applied should therefore be read chiefly as mixture/evolution issues rather than a failure of the overall redshift density.

Our SFG $p(z)$ shape—a broad peak around $z \sim 1$ with a long, shallow high- z tail—is qualitatively consistent with the component in empirical SKAO telescope forecasting work (e.g. Mancuso et al. (2015), their Fig. 10). Two differences help explain why our curves are smoother and show no pronounced ‘bump’ near $z \sim 1$ at $\sim \mu\text{Jy}$ depths:

1. *Population partitioning.* In those forecasts, star-forming galaxies are split into at least two SFG families with distinct evolution: “late-type” discs (normal+starburst) that dominate at $z \lesssim 1-1.5$, and high- z dusty star-forming systems (often tagged as “proto-spheroids”). The sum of two or more evolving families can imprint a mild feature near the transition redshift (Mancuso et al., 2015). In GHOST we treat SFG as a single, smoothly evolving family, which naturally washes out such structure.
2. *How dust enters.* Although radio SFR tracers are intrinsically insensitive to dust attenuation, many forecasting frameworks first construct an SFR function from UV, $\text{H}\alpha$, and IR data (the IR term capturing the obscured star formation) and then map $\text{SFR} \rightarrow L_\nu$ via radio–SFR calibrations (Madau & Dickinson, 2014; Bonato et al., 2017; Mancuso et al., 2015). In such models, a distinctly evolving, dust-rich SFG population that grows rapidly near the cosmic SFR peak ($z \sim 1-2$) can imprint mild features in $n(z)$ at fixed flux (e.g. late-type versus proto-spheroidal contributions) (Mancuso et al., 2015). By contrast, our implementation fits the SFG radio luminosity function directly at 1.4 GHz with smooth redshift evolution, so a given flux limit projects to a correspondingly smooth redshift distribution.

5.6.1 Redshift distributions $n(z)$: rationale for the comparisons

Our goal in this Section is verification of the GHOST mock catalogue. There is no all-sky radio survey to $\sim 1 \mu\text{Jy}$, so we cannot validate the pre-selection $n(z)$ directly on the full sky. Instead, we forward-apply the selection functions of existing surveys (flux/SNR thresholds, peak-vs-integrated survey selection functions, band-passes) to the same underlying catalogue and compare the resulting $n(z)$ to the corresponding data products. This matters because $n(z)$ sets the projection kernels $\phi(z)$ that weight our angular statistics and biases (e.g. the Limber integrals for C_ℓ and the response to the k^{-2}

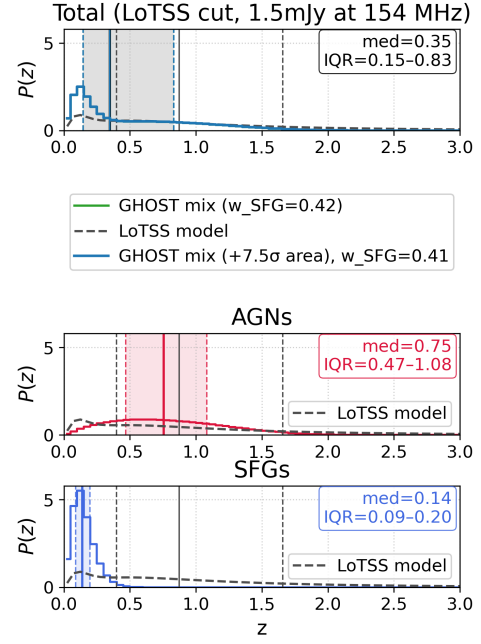


Figure 9. Redshift distributions after applying the flux threshold $S_{150} \geq 1.5 \text{ mJy}$ (Section 5.1). Top: total $P(z)$ from GHOST (black). The green curve reconstructs the total as $w_{\text{SFG}} P_{\text{SFG}}(z) + w_{\text{AGN}} P_{\text{AGN}}(z)$ using the intrinsic flux-limited class fractions ($w_{\text{SFG}}, w_{\text{AGN}}$); the blue curve uses the same $P_{\text{SFG}}(z), P_{\text{AGN}}(z)$ but reweights the class fractions by the LoTSS peak-S/N ≥ 7.5 area acceptance from a log-normal rms model ($\bar{\sigma} = 83 \mu\text{Jy beam}^{-1}$, 95th percentile $171 \mu\text{Jy beam}^{-1}$). The dashed line shows the LoTSS-DR2 parametric model. The green and blue reconstructions are essentially identical, indicating that the peak-S/N cut only weakly perturbs the class mix at 1.5 mJy . Middle/bottom: AGN and SFG PDFs (unit area) with medians and IQRs indicated. Relative to LoTSS, the simulated total places excess weight at $z \lesssim 0.4$ and underpowers the $z \gtrsim 0.6$ shoulder, suggesting a slightly SFG-heavy mix.

PNG term); counts alone cannot diagnose class mixture or redshift tails. Agreement across multiple, independent selections implies that the global LF+SED calibration produces a plausible cosmic history for each class, and our survey emulation is realistic. Conversely, coherent offsets across selections isolate population-level issues (e.g. SFG/AGN mix, high- z tails, spectral/K-correction choices) rather than artefacts of any one data set. The comparisons below are therefore tests designed to stress our population model and selection operators in regimes where data exist, not predictions for a hypothetical all-sky μJy survey.

5.6.2 LoTSS-like selection ($S_{150} \geq 1.5 \text{ mJy}$)

We emulate the LoTSS DR2 (Hale et al., 2024) wide-area selection with an integrated-flux threshold $S_{150\text{MHz}} \geq 1.5 \text{ mJy}$ and a peak-S/N threshold, as described in Section 5.1.1. After replicating the 1.5 mJy sensitivity the selection function removes only a few percent of sources,

so the resulting $n(z)$ is driven by the intrinsic class mixture, shown in Figure 9. Where shown, the “LoTSS model” is the parametric $p(z)$ derived from Deep Fields. The total $p(z)$ is low- z dominated (median $z = 0.35$, IQR 0.15–0.83), with a conspicuous SFG contribution at $z \lesssim 0.3$ (SFG median $z = 0.14$, IQR 0.09–0.20). The AGN component peaks broadly around $z \sim 0.5$ –1 (median $z = 0.75$, IQR 0.47–1.08). Using the measured 150 MHz counts-by-population to set the mixture yields $w_{\text{SFG}} \approx 0.42$. Figure 9 compares the resulting PDFs to the LoTSS model. Relative to the LoTSS model, the total median is lower and the high- z shoulder is underpowered, indicating a SFG-heavy mix at the mJy level and an underrepresentation of moderate- z AGN in the total.

Increasing the flux of the cut shifts the prediction as expected (more AGN-dominated, higher median z), yet even at 10 mJy the median remains below the LoTSS curve (Table 6). The direction and size of the offset are consistent with the RLF diagnostics (Section 5.8; possible population-level origins are deferred to sections 5.6).

S_{min} (mJy)	w_{SFG}	median z
1.5	0.41	0.35
2.0	0.37	0.41
3.0	0.29	0.51
5.0	0.22	0.59
10.0	0.14	0.66

Table 6 Effect of increasing the 150 MHz flux threshold on the mixture weight (w_{SFG}) and the total-sample median redshift.

5.6.3 VLA-COSMOS split at 1.4 GHz (Magliocchetti et al., 2017)

To mirror VLA-COSMOS (Magliocchetti et al., 2017), we apply a flux cut $F_{1.4} \geq 0.15$ mJy and partition sources using a redshift-dependent luminosity boundary, as detailed in Section 5.1.

The total resulting distribution, as shown in Figure 10 remains low- z dominated (median $z = 0.37$, IQR 0.17–0.87). The AGN-like PDF peaks at $z \sim 0.6$ –1 with a tail to $z \sim 2$, while the SFG-like PDF is tightly concentrated at $z \lesssim 0.3$. Integrating above/below $L_{\text{cross}}(z)$ gives an SFG fraction for this selection of

$$f_{\text{SFG}} \equiv \frac{N_{\text{SFG}}}{N_{\text{SFG}} + N_{\text{AGN}}} = 0.49. \quad (27)$$

The selection function split recovers the expected qualitative separation in $n(z)$ prevalent in Magliocchetti et al. (2017), with SFG tightly clustered at low redshift and a broader AGN distribution, but the integrated SFG share is high for this selection, again pointing to

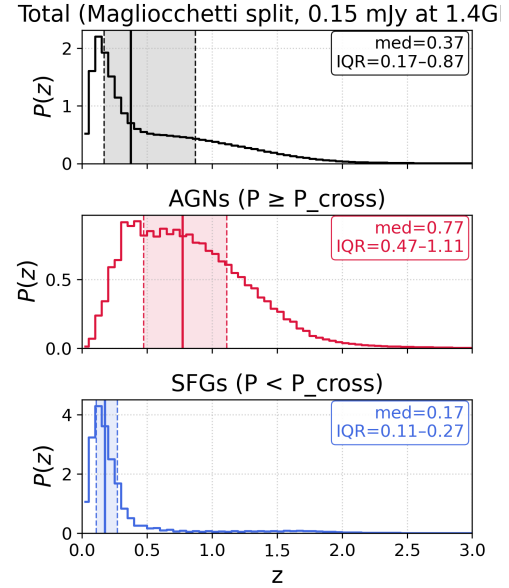


Figure 10. VLA-COSMOS-like $n(z)$ with 1.4 GHz selection. We apply $F_{1.4} \geq 0.15$ mJy and classify sources via $L_{1.4} \geq L_{\text{cross}}(z)$ with $L_{\text{cross}} = 4\pi P_{\text{cross}}$ and $\log_{10} P_{\text{cross}}(z) = 21.7 + z$ ($z \leq 1.8$) or 23.5 otherwise (Section 5.1.2). Top: total $P(z)$ with median and IQR; middle/bottom: AGN-like and SFG-like PDFs. The expected qualitative in redshift - with SFG tightly clustered and AGN more broad - is recovered, but the integrated SFG-like fraction is high, again pointing to a slightly SFG-heavy mix at the mJy level.

a SFG-heavy mix around the mJy level. Because this diagnostic does not depend on peak-vs-integrated flux (unlike LoTSS), it isolates the population modelling; too many low- z SFG and/or too few moderate- z radio-loud AGN.

Relative to the COSMOS-based fractions of (Magliocchetti et al., 2017), our f_{SFG} is larger in the same sense indicated by the LoTSS test. A hard L -threshold is an imperfect classifier (e.g. it can move luminous SFG at $z \sim 1$ into the AGN-like bin), so exact numerical agreement is not expected; the direction of the difference is robust and consistent across selections.

5.6.4 VLA-COSMOS-like selection at 3 GHz (Hale et al., 2018)

We emulate the $S/N \geq 5.5$ 3 GHz selection (median rms $\simeq 2.3 \mu\text{Jy beam}^{-1}$) and compare to the multi-wavelength SFG/AGN classes of Hale et al. (2018). At the Hale et al. (2018) depth the applied flux threshold and selection function detailed in Section 5.1 is weak: numerically the acceptances are close to unity for both classes (in our run $\text{acc}_{\text{SFG}} \approx 0.89$, $\text{acc}_{\text{AGN}} \approx 0.94$), so the predicted $n(z)$ is governed mainly by the intrinsic SFG/AGN mix. The total has median $z = 0.38$. Class medians are $z_{\text{med}}^{\text{SFG}} = 0.30$ (IQR 0.19–0.45) and $z_{\text{med}}^{\text{AGN}} = 0.80$ (IQR 0.48–1.22). The effective SFG fraction is $w_{\text{SFG}} = 0.69$.

At this depth and frequency the selection is intrinsi-

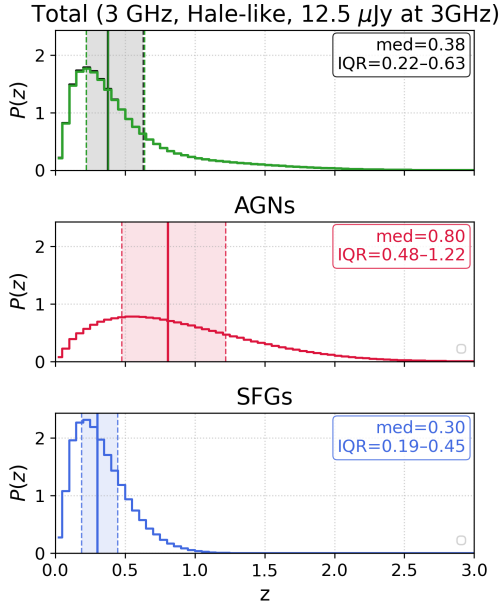


Figure 11. VLA-COSMOS-like $n(z)$ at 3 GHz. GHOST total (top) and class PDFs (middle/bottom) for a $S/N \geq 5.5$ selection function (Section 5.1.3); the top panel also shows the effective counts-weighted mixture (green). Medians and IQRs are indicated. The total median is low and the SFG fraction is high compared to Hale et al. (2018), consistent with the trend that GHOST overweights nearby SFG relative to moderate- z radio-loud AGN.

cally SFG-rich, but the simulated mixture is still somewhat SFG-heavy relative to Hale et al. (2018), again implicating the population balance rather than the survey selection function. Hale et al. (2018) classify $\sim 79\%$ of sources via multi-wavelength diagnostics and report a SFG share of $\simeq 3704/3704 + 2937 \approx 0.56$ in the masked field, together with a higher total median z (see their Fig. 3). Our Hale-like selection therefore echoes the LoTSS and Magliocchetti results: GHOST places relatively more weight in nearby SFG and underweights moderate- z radio-loud AGN. The sign and scale of the difference are consistent with the RLF and clustering comparisons.

5.7 Differential and Euclidean-normalized number counts

This section tests the total source density in GHOST against the literature via Euclidean-normalized differential counts, $S^{2.5} dN/dS$, at 150 MHz, 1.4 GHz, and 3 GHz. Because number counts integrate over luminosity and redshift, they primarily probe the combination of luminosity-function normalisations and radio SEDs, largely independent of the host-halo assignment used for clustering.

Figure 12 shows that over nearly seven decades in flux density the simulation follows the observed trends at all three frequencies. At the bright end ($S \gtrsim 10$ mJy),

the counts are close to Euclidean ($S^{2.5} dN/dS \approx \text{const}$), consistent with a population dominated by radio-loud AGN.

A tension appears around the classical sub-mJy upturn, where the data rise due to the emergence of SFG. In our model this rise is weaker at $S \sim 0.1\text{--}0.3$ mJy. This behaviour is fully consistent with the adopted SFG RLF: the radio-first, smoothly evolving SFG luminosity function underpredicts the luminous tail of SFG at $z \gtrsim 1$, which suppresses the upturn when projected to flux. Since total counts depend on the RLFs and SEDs rather than on host assignment, abundance matching does not drive this feature.

5.7.1 Residuals

The residual panels plot $(\text{model}/\text{data}) - 1$ at the locations of the observational points. Away from the very faint and very bright limits, the residual clouds at 150 MHz, 1.4 GHz, and 3 GHz are centred on zero within the quoted errors, with a typical scatter of $\pm(10\text{--}30)\%$ across ~ 4 dex in S , indicating no frequency-dependent bias in the combination of RLFs and SEDs. Where residuals dip near the sub-mJy regime (most visibly at 1.4 GHz, with corresponding ranges at 150 MHz and 3 GHz), the shortfall follows from the SFG RLF’s high- L tail rather than from halo scarcity or assignment choices. At the brightest fluxes, excursions of order $\gtrsim 30\%$ reflect the rarity of luminous AGN and the small effective survey volumes of the anchoring datasets.

5.7.2 Implications

The close match to multi-frequency counts implies that the global normalisations of the luminosity functions, the adopted radio SEDs, and the applied survey selections are mutually consistent: GHOST produces the correct number of sources as a function of flux density without frequency-dependent bias, apart from the discussed SFG underprediction at the sub-mJy upturn. Because counts constrain the sum over populations, different SFG/AGN partitions can yield similar totals; discrepancies seen later in $n(z)$, the RLFs, and clustering should therefore be interpreted as redistributions across populations and redshifts rather than a failure in the overall source density. The 3 GHz panel of Figure 12 shows the same behaviour, indicating that the spectral extrapolation from 1.4 GHz to 3 GHz is consistent at the precision of current counts.

5.8 RLF

We describe the class-dependent radio luminosity function (RLF) by $\Phi_c(L_\nu, z)$ in units of $\text{Mpc}^{-3} \text{dex}^{-1}$, where $c \in \{\text{SFG}, \text{SSAGN}, \text{FSRQ}, \text{BLLac}\}$. Observed flux density at ν_{obs} is related to rest-frame luminosity through

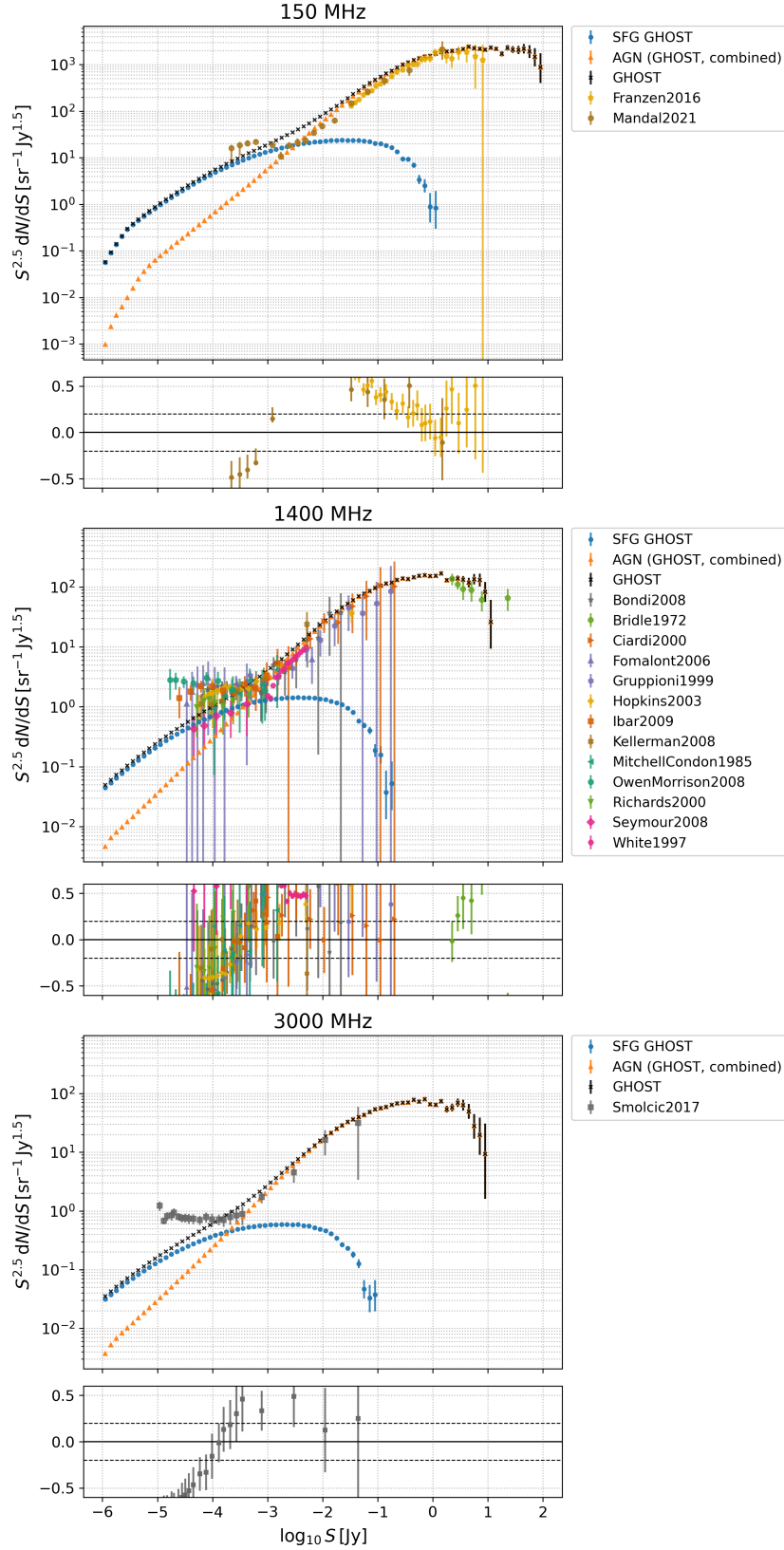


Figure 12. Euclidean-normalized differential counts, $S^{2.5} dN/dS$, for GHOST (black markers) compared to measurements at 150 MHz (Franzen et al., 2016; Mandal et al., 2021), 1.4 GHz (Bondi et al., 2008; Bridle et al., 1972; Ciardi & Loeb, 2000; Fomalont et al., 2006; Gruppioni et al., 1999; Hopkins et al., 2003; Ibar et al., 2009; Kellermann et al., 2008; Mitchell & Condon, 1985; Owen & Morrison, 2008; Richards, 2000; Seymour et al., 2008; White et al., 1997; de Zotti et al., 2010) and 3 GHz (Smolčić et al., 2017). Error bars are 1σ measurement uncertainties, incorporating the estimations of Gehrels (1986) if $N < 20$ for any bin; shaded bands (if shown) indicate the field-to-field variance from GHOST mock realizations with the same area and flux cuts as the data. Residual panels show $(\text{model} - \text{data})/\text{data}$.

the standard K -correction,

$$L_\nu(z, \alpha_c) = 4\pi D_L^2(z) S_{\nu_{\text{obs}}} \left(\frac{\nu_{\text{ref}}}{\nu_{\text{obs}}} \right)^{-\alpha_c} (1+z)^{1+\alpha_c}, \quad (28)$$

so that, for a survey flux limit S_{lim} , the luminosity boundary within a redshift slice is

$$L_{\text{lim}}(z, \alpha_c) = 4\pi D_L^2(z) S_{\text{lim}} \left(\frac{\nu_{\text{ref}}}{\nu_{\text{obs}}} \right)^{-\alpha_c} (1+z)^{1+\alpha_c}, \quad (29)$$

with $L_\nu \propto \nu^{-\alpha}$.

Forward (projected) model. When comparing to data taken over a finite redshift bin $z \in [z_1, z_2]$ we project the intrinsic RLF under the same selection,

$$\frac{\Phi_c^{\text{proj}}(L | z_1:z_2) = \int_{z_1}^{z_2} \Phi_c(L, z) \Theta[L - L_{\text{lim}}(z, \alpha_c)] (dV/dz) dz}{\int_{z_1}^{z_2} (dV/dz) dz}. \quad (30)$$

These orange dashed curves in Figs. 15–16 should be regarded as the baseline “model-to-data” comparison.

Binned estimator from the catalogue (V_{eff}). The blue points are a direct estimate from the catalogue in the $[z_1, z_2] \times [L_1, L_2]$ cell using an effective volume that accounts for partial visibility of the L -bin across the slice. Define the bin width in dex by $\Delta \log_{10} L = \log_{10}(L_2/L_1)$. At any redshift z , the fraction of the luminosity bin that is observable is

$$f_{\text{vis}}(z; L_1, L_2) = \begin{cases} 1, & L_{\text{lim}}(z, \alpha_c) \leq L_1, \\ \frac{\log_{10}(L_2/L_{\text{lim}}(z, \alpha_c))}{\log_{10}(L_2/L_1)}, & L_1 < L_{\text{lim}}(z, \alpha_c) < L_2, \\ 0, & L_{\text{lim}}(z, \alpha_c) \geq L_2. \end{cases} \quad (31)$$

The corresponding effective comoving volume for the cell is

$$V_{\text{eff}}(L_1, L_2; z_1:z_2) = \Omega \int_{z_1}^{z_2} \frac{dV}{dz}(z) f_{\text{vis}}(z; L_1, L_2) dz, \quad (32)$$

with survey solid angle Ω (sr). If N_c catalogue objects of class c fall in the cell, the minimum-variance, Poisson estimator is

$$\hat{\Phi}_c(L | z_1:z_2) = \frac{N_c}{V_{\text{eff}}(L_1, L_2; z_1:z_2) \Delta \log_{10} L}, \quad (33)$$

$$\sigma^2[\hat{\Phi}_c] = \frac{N_c}{[V_{\text{eff}} \Delta \log_{10} L]^2}. \quad (34)$$

If $N < 20$ for any bin, we use the estimators of Gehrels (1986) in constructing the confidence interval. We plot the estimate at the bin centre with vertical error bars (from Equation 34) and omit bins where $V_{\text{eff}} = 0$. Horizontal bars indicate $\Delta \log_{10} L$. For empty bins with $V_{\text{eff}} > 0$ we show one-sided upper limits (Gehrels-style) when useful.

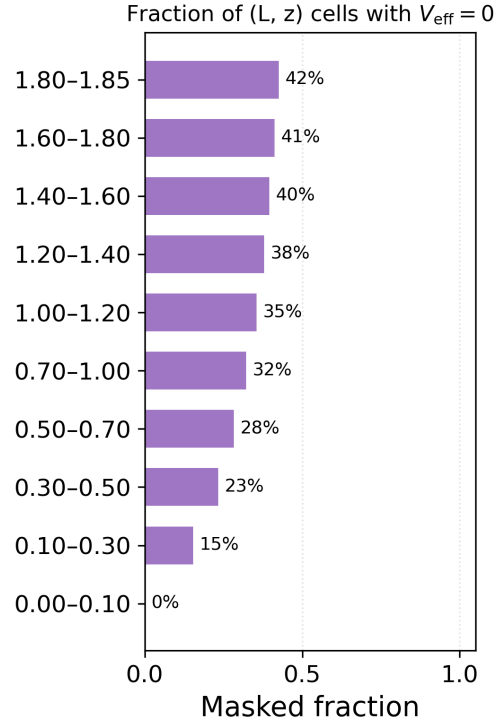


Figure 13. Fraction of (L, z) cells with zero effective volume, $V_{\text{eff}} = 0$, in each redshift bin for the raw selection. Bars show the average masked fraction across AGN and SFG classes; these are indistinguishable within $\lesssim 0.5\%$ in every bin because the mask is set by the survey flux limit (and K -correction) rather than population physics. The fraction rises monotonically from $\lesssim 15\%$ at $z \leq 0.3$ to $\simeq 40\%$ by $z \sim 1.8$, reflecting the shrinking accessible luminosity range in a flux-limited survey. Cells with $V_{\text{eff}} = 0$ are excluded from the panel-averaged RLFs.

Masked fraction of (L, z) cells. Figure 13 shows, for each redshift bin, the fraction of (L, z) cells that are masked because their effective survey volume is zero, $V_{\text{eff}}(z, L) = 0$. A cell is masked precisely when the visibility fraction in Equation 31 vanishes, $f_{\text{vis}}(z; L_1, L_2) = 0$; in those cells we set Φ to NaN and exclude them from the RLF averages. As expected for a flux-limited sample, the masked fraction is $\simeq 0$ in the nearest slice ($0 < z \leq 0.1$) and then rises monotonically with redshift as the flux boundary $L_{\text{lim}}(z, \zeta) \propto D_L^2(z) (1+z)^{\zeta-1}$ moves to higher luminosities. Numerically, the fractions are $\approx 15\%$ at 0.10–0.30, ≈ 23 –28% at 0.30–0.70, ≈ 32 –38% at 0.70–1.40, and ≈ 40 –42% by $z \simeq 1.8$. The AGN and SFG panels are nearly identical because the V_{eff} mask is set by survey geometry and the adopted spectral index used in the visibility correction, not by population physics. Thus the increasing masked fraction simply quantifies the shrinking accessible L -range with redshift and explains the growing white space in the high- z RLF panels; it is a geometric selection effect, not a loss of simulated sources nor a model shortcoming.

Results at 1.4 GHz. Figures 15 and 16 show excellent agreement between the projected model (orange dashed) and the catalogue-based, V_{eff} -corrected estimates (blue points):

- **AGN:** The match is tight across all redshift slices and over the full dynamic range where the survey is complete. This validates both the evolutionary form of the AGN RLF and the class mix encoded in the model.
- **SFG:** At $z \gtrsim 1$ the blue points fall below some literature bright-end measurements. This is a model outcome: our adopted $L_{1.4}$ -SFR calibration is shallower at high SFR (see Figure 14), compressing the bright tail and yielding fewer $L_{1.4} \gtrsim 10^{23-24} \text{ W Hz}^{-1}$ systems when an evolving SFR function is mapped into radio luminosity. The same behaviour is consistent with the divergence of the total counts at the faint flux end where SFG dominate.

A shallow “bump/inflection” occasionally visible in the orange curves near $10^{26} \text{ W Hz}^{-1}$ is naturally produced by the sum of AGN subclasses and their evolution; the double-power-law knee of the steep-spectrum tail, combined with the beamed (FSRQ/BL Lac) component beginning to dominate the bright end with a slightly flatter local slope, and the finite-width slice average in Equation 30. A per-class decomposition of Φ^{proj} confirms that the total slope locally reflects the crossover of SSAGN and beamed contributions. The feature is therefore not required by the data and does not affect the excellent overall agreement.

6 LUMINOSITY-REDSHIFT DISTRIBUTIONS BY CLASS

Figures 17–20 show, for each population, the distribution of rest-frame radio luminosity versus redshift. Shading encodes the relative number density per redshift bin; over-plotted are the median and mean of $\log_{10} L$, central ($p16$ – $p84$) and extended ($p05$ – $p95$) percentile bands, and the minimum/maximum values realised in the catalogue. These panels are intended to summarise, at a glance, how the radio luminosity function (RLF) and the survey selection combine to populate the light-cone for each class.

A few general features are common to all panels. The lower envelope follows the flux-limited luminosity threshold $L_{\text{lim}}(z, \zeta)$ implied by the selection and K -corrections (Section 5.1), which rises rapidly with redshift and produces the Malmquist-like increase of the median at low z . The separation between the mean and the median indicates positive skewness due to a high- L tail; the width of the percentile bands traces the intrinsic luminosity spread convolved with selection. At $z \lesssim 0.5$ the accessible dynamic range in L is largest and the bands

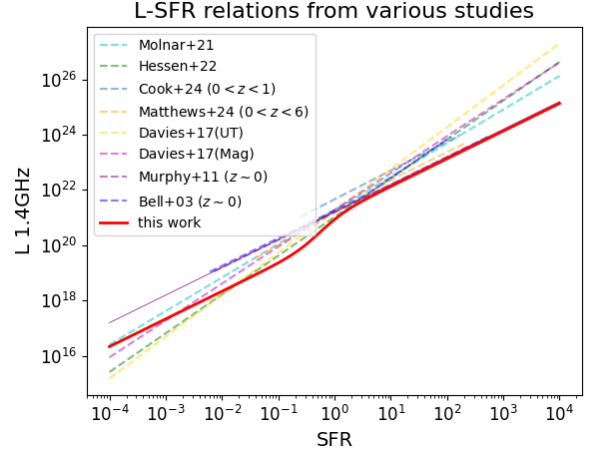


Figure 14. Comparison of 1.4 GHz radio luminosity–SFR calibrations. Dashed curves show literature relations (Molnár et al., 2021; Heesen et al., 2022; Cook et al., 2024; Matthews et al., 2024; Davies et al., 2017; Murphy et al., 2011; Bell et al., 2003). The solid red line (“this work”) is our adopted mapping. At high SFR the slope of our relation is shallower than several published fits, compressing the bright tail when mapping a given $\phi_{\text{SFR}}(z)$ into $L_{1.4}$ and thereby yielding fewer SFG at $L_{1.4} \gtrsim 10^{23-24} \text{ W Hz}^{-1}$. This behaviour contributes to the lower bright-end SFG space densities seen at $z \gtrsim 1$ in Fig. 16. The halo mass range and abundance matching limits the upper SFR range in practice.

are widest. As z increases the faint end is progressively removed by L_{lim} , the median rises, and the bright tail is governed by the evolving knee of the RLF and by the rarity of very luminous sources. Beyond $z \sim 2$ – 3 the high- L tail thins and the percentile bands narrow as space densities decline; small-number statistics appear as occasional spikes in the dotted “max” curves in the highest- z bins.

Star-forming galaxies exhibit the lowest median luminosities and comparatively narrow percentile bands at fixed z . The median increases gently with redshift, while the mean sits only slightly above it, signalling limited skewness. The density shading peaks at intermediate redshift and fades by $z \sim 3$, consistent with the smoothly evolving, radio-first SFG RLF of Section 3.2. This behaviour also explains the modest sub-mJy upturn in the counts (Section 5.7): the high- L SFG tail at $z \gtrsim 1$ is present but not dominant, so the projected rise in $S^{2.5} dN/dS$ is weaker than in datasets whose SFG bright end is steeper or more numerous.

Steep-spectrum AGN span a wide luminosity range. Through $z \sim 1$ – 2 the high- L tail is prominent, lifting the mean well above the median and widening the $p16$ – $p84$ band. The median sits roughly a dex above the SFG median at the same redshift, reflecting the higher typical host masses of radio-loud systems (Section 3.3). At $z \gtrsim 2.5$ the percentile bands narrow and the maximum realised luminosity drops, tracing both the SSAGN RLF evolution and the diminishing supply of massive halos

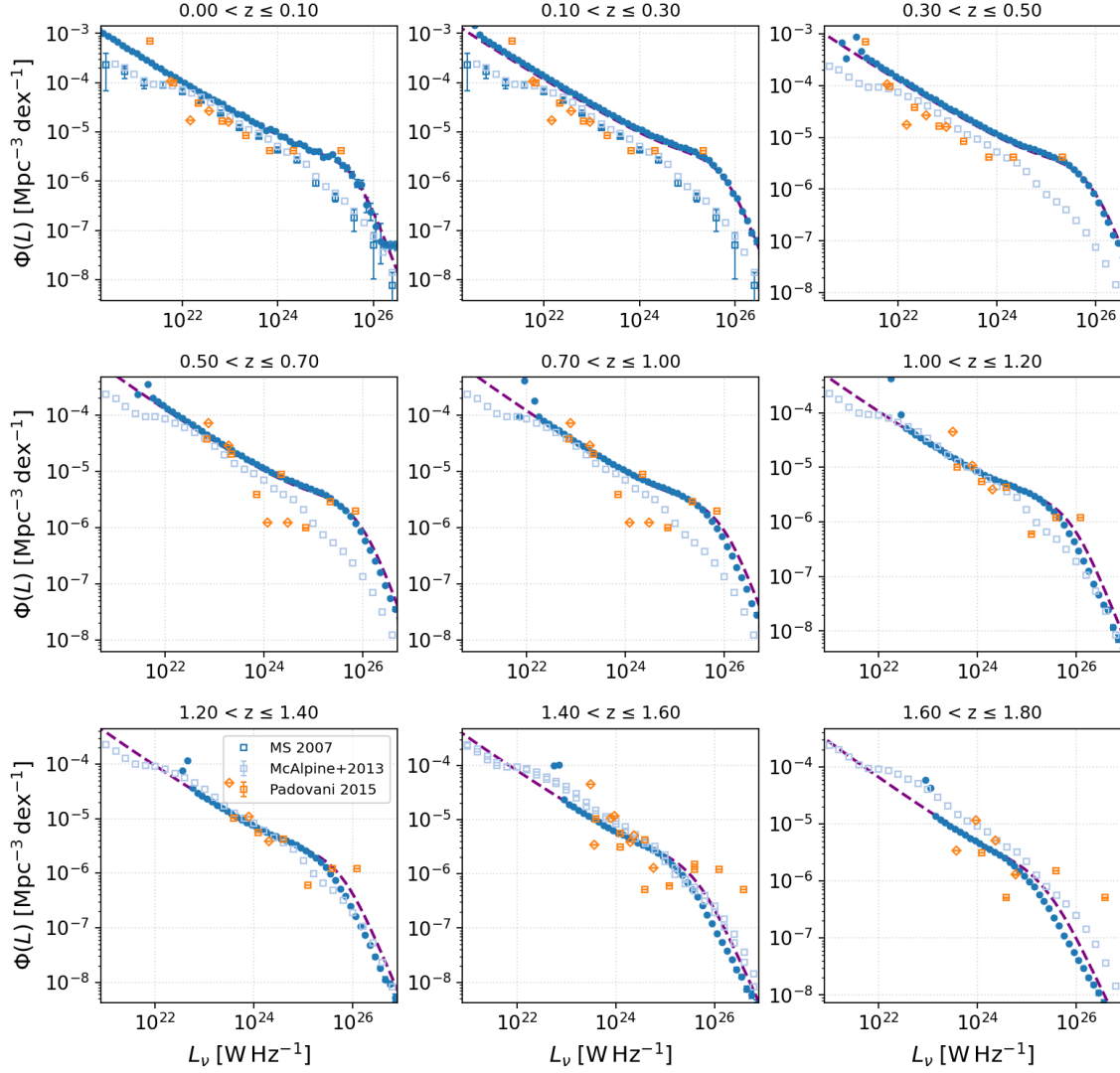


Figure 15. AGN radio luminosity functions at 1.4 GHz in redshift slices (raw cut). Each panel shows $\Phi(L_\nu)$ [$\text{Mpc}^{-3} \text{dex}^{-1}$] for the stated redshift bin. *Blue circles*: RLF measured from the simulation using the effective-volume method, $\Phi = N/(V_{\text{eff}} \Delta \log_{10} L)$, where V_{eff} accounts for the flux limit within the bin (§5.8; Equation 34). Error bars are 1σ Poisson uncertainties propagated through V_{eff} ; cells with $V_{\text{eff}} = 0$ are omitted. *Purple dashed*: forward (projected) model $\Phi^{\text{proj}}(L|z_1:z_2)$ obtained by averaging the intrinsic AGN RLF over the finite Δz of the panel under the same flux selection (Equation 30). *Points in the background*: literature measurements at 1.4 GHz (legend in the top-left panel). Across the ten slices the simulation (blue) and the forward model (purple) show excellent agreement in both normalisation and shape; a mild feature near $L_\nu \sim 10^{26} \text{ W Hz}^{-1}$ reflects the combined contribution of flat-spectrum (FSRQ/BLLac) and steep-spectrum (SSAGN) components with luminosity-dependent evolution (see §5.8). Axis limits in each panel are set by the extent of the observational points to aid readability: Squares indicate RL AGN; diamonds indicate RQ AGN. Filled markers are fully within the z -bin; open markers partially overlap the bin.

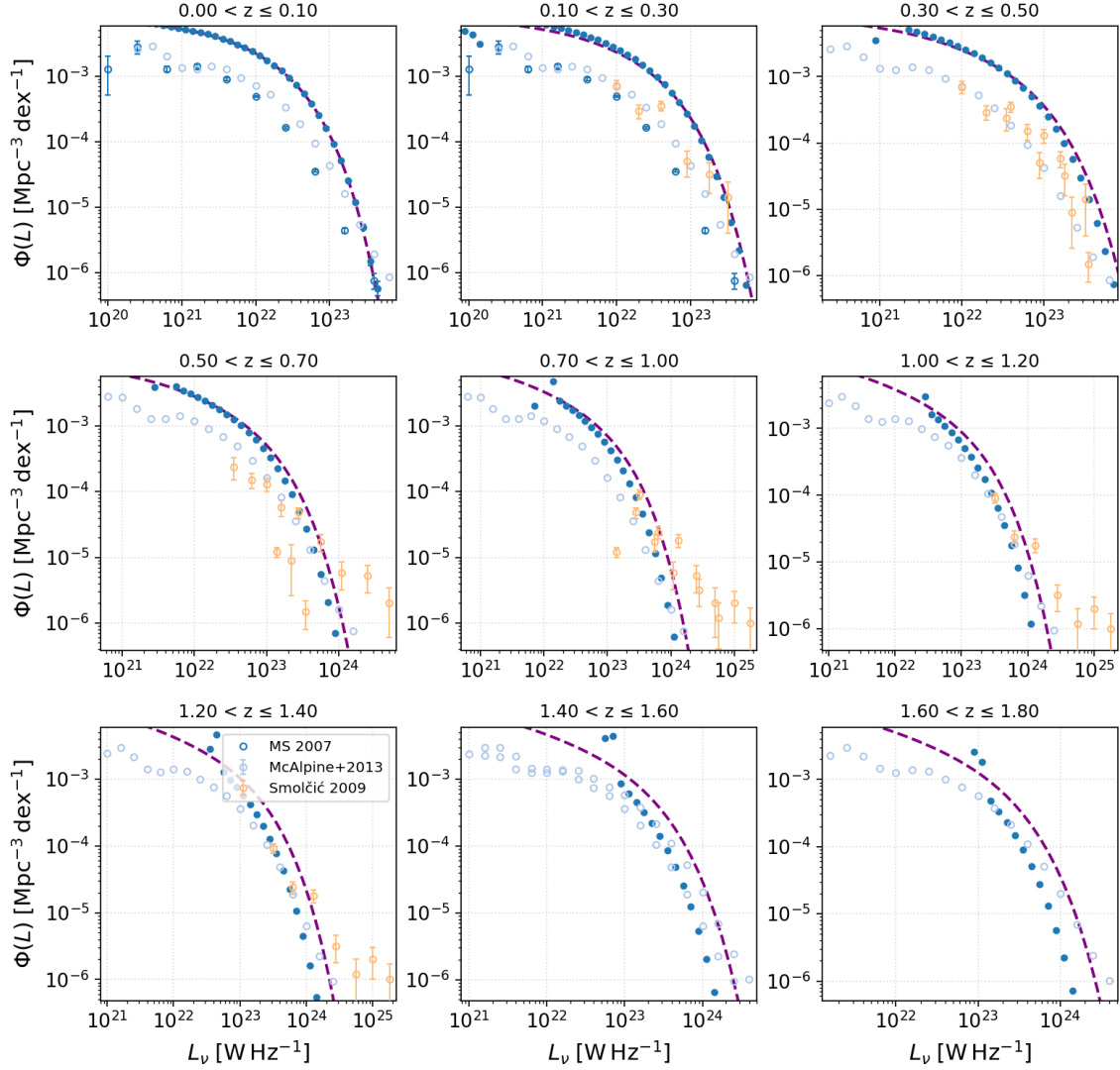


Figure 16. SFG radio luminosity functions at 1.4 GHz in redshift slices (raw cut). As in the AGN figure, *blue circles* show the simulation RLF recovered with the effective-volume estimator $\Phi = N/(V_{\text{eff}} \Delta \log_{10} L)$ (with 1σ Poisson errors), and *orange dashed curves* are the forward (projected) model evaluated over the same Δz and flux selection. *Background points*: literature measurements at 1.4 GHz (legend in the top-left panel). At $z \lesssim 1$ the model and simulation track the observed faint end well. At higher redshift the bright tail in the simulation/model sits below some compilations: this follows naturally from our adopted L -SFR mapping (shallower high-SFR slope), the tightening luminosity threshold imposed by the flux limit within each Δz slice, and the resulting reduction of accessible high-SFR systems (see §5.8). As before, panel limits are set by the extent of the observational points; filled markers are fully within the z -bin; open markers partially overlap the bin.

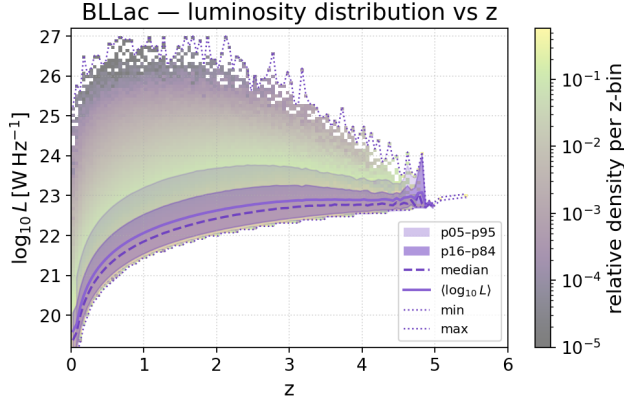


Figure 17. BLLac — luminosity distribution vs. redshift. Rows are normalized by their own maxima; ribbons show p16–p84 and p05–p95 of $\log_{10} L_\nu$. The narrowing and softening of the upper envelope (shown by the dashed purple line) at high z reflect the flux limit and declining BL Lac space density. The lower minimum envelope is shown by the lower dashed line.

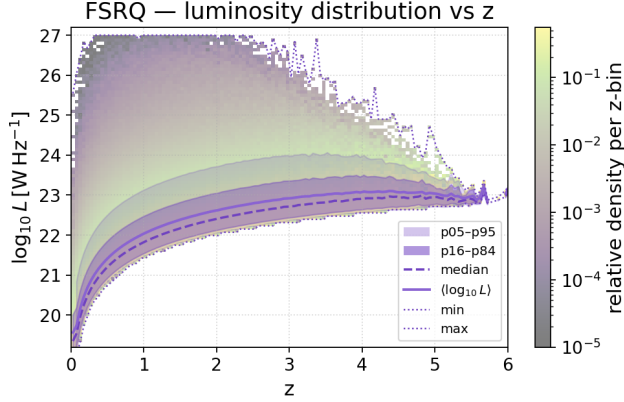


Figure 18. FSRQ — luminosity distribution vs. redshift. Same notation as 17. The broad luminous tail at $z \lesssim 1$ and the subsequent contraction beyond $z \sim 3$ are evident.

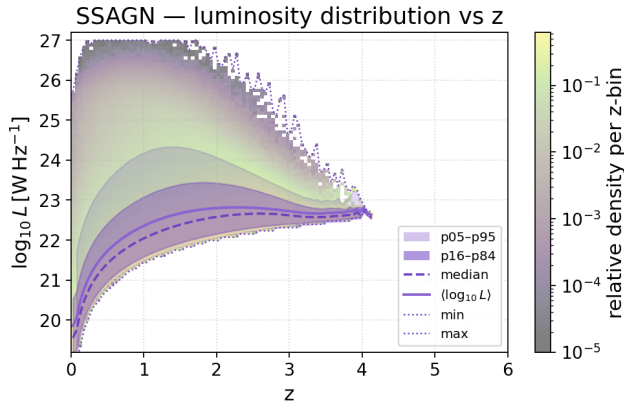


Figure 19. SSAGN — luminosity distribution vs. redshift. Same notation as 17. The median rises to $z \sim 1$ and then turns over, with the distribution tightening as the flux boundary dominates.

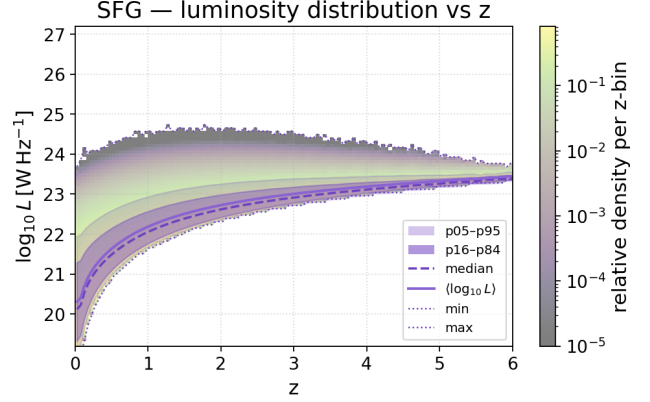


Figure 20. SFG — luminosity distribution vs. redshift. Same notation as 17. The median follows the cosmic growth of typical SFRs to $z \sim 2-3$ and declines gently thereafter; the accessible upper envelope drops beyond $z \sim 3$ as the flux limit and host availability restrict very luminous SFG.

available to host them (Section 4).

Flat-spectrum subclasses show the highest luminosities at fixed redshift. FSRQ display broad percentile bands and a large mean–median gap, as expected for a beamed population with a strong bright tail; the shading peaks around $z \sim 1-2$ and declines rapidly by $z \gtrsim 3$ as the combination of RLF evolution and beamed fractions reduces the space density. BL Lac objects occupy a lower-luminosity locus than FSRQ and a somewhat narrower redshift span; their bands are tighter and the shading is more local-heavy, indicating milder evolution and smaller overall abundance. In both beamed classes the dotted “max” curve shows occasional excursions at high redshift that are consistent with small-number statistics.

These shapes have direct consequences for selection and for the observables compared elsewhere in this paper. Because changes in survey selection functions compress or broaden the effective redshift kernel $\varphi(z)$ (Section 5.1), the rising $L_{\text{lim}}(z)$ cuts different parts of each panel for a given flux limit. SFG are most affected at low to intermediate redshift where their locus lies close to L_{lim} ; SSAGN and flat-spectrum AGN remain visible to higher z but are shot-noise limited in the brightest tail. This selection geometry explains why projected clustering amplitudes and slopes shift with the survey selection functions (Section 5.3) without altering the intrinsic bias ordering inferred from the host-mass distributions: SFG occupy lower-mass halos and populate lower L at a given z ; SSAGN are intermediate; FSRQ/BL Lac trace the luminous, highly biased tail.

For multi-tracer applications the panels make explicit that each class supplies a distinct $P(z)$ kernel and linear-bias evolution. The ordered rise of the medians and the class-to-class offsets stabilise the bias contrast re-

quired for cosmic-variance cancellation on ultra-large scales, while the thinning of the bright tails at $z \gtrsim 3$ sets practical limits on how aggressively very-high- z selections can be pushed before shot noise dominates. Together with the number-count agreement in Section 5.7 and the clustering results in Section 5.3, these luminosity-redshift maps close the loop between population prescriptions, host assignment, and the survey selections used for validation.

7 IMPROVEMENTS

GHOST is an all-sky radio-continuum simulation with population-dependent clustering, redshift distributions, and survey selections designed to emulate realistic wide-area datasets. Below we summarise where the current release is most useful and outline focused directions that will make it more powerful for cosmology and survey work.

- Spectral assumptions and K -corrections. Using a single (or overly narrow) spectral-index model per population tightens the flux-luminosity mapping, especially at high z where rest-frame frequencies shift. *Future direction:* adopt redshift-dependent spectral-index distributions (optionally with curvature), for exact downstream K -corrections, and validate against multi-frequency counts.
- SFG RLF bright end and the sub-mJy upturn. The radio-first SFG RLF underpredicts the high- L SFG tail at $z \gtrsim 1$, weakening the classical upturn near $S \sim 0.1\text{--}0.3$ mJy. *Future direction:* re-fit the SFG RLF with deeper fields and include a cross-check pathway that links to SFR/IR constraints (without making them mandatory) to control the bright tail while preserving the radio-first baseline.
- Satellites and halo occupation. The current catalogue assigns one central per halo and no explicit satellites. This can bias small-scale $w(\theta)$ and environment trends, even if large-scale bias is stable. *Future direction:* add a minimal HOD layer (central + satellites) per population, tuned to small-scale clustering while keeping the ultra-large-scale predictions unchanged.
- Survey realism and detection. Point-like components and simplified morphologies miss PSF/blending effects and spatially varying completeness. *Future direction:* ship an “observe-at-beam” operator (beam, mosaics, noise/rms maps, deblending/thresholding) that produces detection catalogues, randoms, and effective-area curves consistent with pseudo- C_ℓ inputs.
- Documentation and delivered products. *Future direction:* include per-source SED parameters (ζ , curvature), $p(z)$ options, masks/randoms, and $n(z)$ kernels used in the figures so analyses can be repro-

duced exactly.

Taken together, these items indicate where GHOST can grow. They also clarify how to interpret the current results: robust for wide-area cosmology (PNG, ISW, magnification) under the stated assumptions, with clear paths to refine high- z populations, spectral modelling, morphology, and survey realism in future releases.

8 DATA AVAILABILITY

The data and simulated catalogues will be made available on the Data Central service (<https://datacentral.org.au/>).

9 CONCLUSIONS

GHOST is an all-sky, clustered, survey-ready radio mock built on the FLAMINGO light-cone and tailored to ultra-large-scale, multi-tracer cosmology. The catalogue ties population prescriptions to a specific halo field through rank-preserving, no-replacement abundance matching (Section 4), ensuring that large-scale bias follows the halo supply. The SFG branch is calibrated directly in radio via a smoothly evolving 1.4 GHz RLF (Section 3.2), while the AGN branch retains the T-RECS functional form with coefficients refit as needed and a new host assignment on our light-cone (Section 3.3). Together these choices provide transparent control of bias and redshift kernels at the scales relevant for f_{NL} .

Across 150 MHz, 1.4 GHz, and 3 GHz the Euclidean-normalized differential counts are reproduced over nearly seven decades in flux (Section 5.7). The model shows no frequency-dependent bias in the total counts; the subdued sub-mJy upturn is consistent with the high-luminosity tail of the radio-first SFG RLF at $z \gtrsim 1$, rather than with host-assignment choices. This anchors the overall normalisation of the luminosity functions and spectral assumptions used throughout.

Using common selections for both $n(z)$ and clustering, we emulate the survey selection functions used by representative surveys and recover their qualitative behaviour (Section 5.6). Emulating LoTSS-wide, the VLA-COSMOS luminosity split of Magliocchetti et al. (2017), and the 3 GHz S/N threshold of Hale et al. (2018), we find that differences between literature measurements are largely explained by how the survey selection functions compress or broaden the projection kernel $\varphi(z)$. Where a survey defines its own AGN/SFG partition (Magliocchetti et al., 2017), we mimic it explicitly; otherwise we retain our internal classes and emulate only the detection to isolate selection from population-mix effects. This like-for-like procedure keeps $n(z)$ and clustering internally consistent and clarifies the role of classification in published comparisons.

In configuration and harmonic space, the catalogue exhibits a stable bias hierarchy with $b_{\text{SFG}} < b_{\text{SSAGN}} \lesssim b_{\text{FSRQ}} \simeq b_{\text{BLLac}}$ (Section 5.4). Power-law fits to $w(\theta)$ and template fits to C_ℓ agree and map cleanly onto the host–halo mass PDFs: SFG inhabit lower-mass halos; radio-loud AGN occupy more massive halos. Apparent slope and amplitude shifts under different selections are traced to changes in χ_{eff} and $\varphi(z)$, not to a reversal of intrinsic bias ordering. This closes the loop between the population model, host assignment, and the observed large-scale clustering.

For applications, the catalogue supplies the two ingredients required by multi-tracer estimators targeting local-type primordial non-Gaussianity: a robust bias contrast between tracer families and distinct, selection-matched redshift kernels. The same full-sky, selection-aware density fields support late-time CMB cross-correlations (ISW and CMB lensing), magnification studies via add-on ray tracing, radio-dipole tests, and survey design tasks that require controlled injections of depth and masking systematics.

The main caveats and near-term improvements are well circumscribed (Section 7). The SFG bright end at $z \gtrsim 1$ can be revisited with deeper fields and optional SFR/IR cross-checks; spectral-index distributions can be broadened (and made redshift-dependent) to sharpen K -corrections; small-scale clustering and detection realism will benefit from satellites and morphology/PSF pipelines. None of these affect the principal conclusions on large scales.

In summary, GHOST reproduces multi-frequency counts, delivers selection-matched $n(z)$, and yields a physically ordered bias hierarchy tied to the FLAMINGO halo field. These properties make it a practical sandbox for end-to-end ultra-large-scale analyses. Paper II carries the calibrated bias parameters and kernels forward to quantify f_{NL} sensitivity with variance-canceling multi-tracer estimators on the simulated sky.

ACKNOWLEDGMENTS

B.V. would like to thank the members of the FLAMINGO collaboration Joop Schaye, Matthieu Schaller, Roi Kugel and John Helly for the helpful discussions.

REFERENCES

Abramo L. R., Leonard K. E., 2013, *MNRAS*, 432, 318
 Amendola L., et al., 2018, *Living Reviews in Relativity*, 21, 2
 Aversa R., Lapi A., de Zotti G., Shankar F., Danese L., 2015, *ApJ*, 810, 74
 Bahr-Kalus B., Parkinson D., Asorey J., Camera S., Hale C., Qin F., 2022, arXiv e-prints, p. arXiv:2204.13436

Bartolo N., Komatsu E., Matarrese S., Riotto A., 2004, *Phys. Rep.*, 402, 103
 Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, *ApJS*, 149, 289
 Best P. N., et al., 2023, *MNRAS*, 523, 1729
 Böhme L., et al., 2025, arXiv e-prints, p. arXiv:2509.16732
 Bonaldi A., Bonato M., Galluzzi V., Harrison I., Mascardi M., Kay S., De Zotti G., Brown M. L., 2019, *MNRAS*, 482, 2
 Bonato M., et al., 2017, *MNRAS*, 469, 1912
 Bondi M., Ciliegi P., Schinnerer E., Smolčić V., Jahnke K., Carilli C., Zamorani G., 2008, *ApJ*, 681, 1129
 Bonzini M., Padovani P., Mainieri V., Kellermann K. I., Miller N., Rosati P., Tozzi P., Vattakunnel S., 2013, *MNRAS*, 436, 3759
 Borrow J., Schaller M., Bower R. G., Schaye J., 2022, *MNRAS*, 511, 2367
 Borzyszkowski M., Bertacca D., Porciani C., 2017, *MNRAS*, 471, 3899
 Bridle A. H., Davis M. M., Fomalont E. B., Lequeux J., 1972, *AJ*, 77, 405
 Chen X., 2010, *Advances in Astronomy*, 2010, 638979
 Ciardi B., Loeb A., 2000, *ApJ*, 540, 687
 Condon J. J., 1992, *ARA&A*, 30, 575
 Cook R. H. W., et al., 2024, *MNRAS*, 531, 708
 Dalal N., Doré O., Huterer D., Shirokov A., 2008, *Phys. Rev. D*, 77, 123514
 Davies L. J. M., et al., 2017, *MNRAS*, 466, 2312
 DiPompeo M. A., Runnoe J. C., Myers A. D., Boroson T. A., 2013, *ApJ*, 774, 24
 Elahi P. J., Cañas R., Poulton R. J. J., Tobar R. J., Willis J. S., Lagos C. d. P., Power C., Robotham A. S. G., 2019, *PASA*, 36, e021
 Fanaroff B. L., Riley J. M., 1974, *MNRAS*, 167, 31P
 Ferramacho L. D., Santos M. G., Jarvis M. J., Camera S., 2014, *MNRAS*, 442, 2511
 Fomalont E. B., Kellermann K. I., Cowie L. L., Capak P., Barger A. J., Partridge R. B., Windhorst R. A., Richards E. A., 2006, *ApJS*, 167, 103
 Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
 Franzen T. M. O., et al., 2016, *MNRAS*, 459, 3314
 Gehrels N., 1986, *ApJ*, 303, 336
 Gomes Z., Camera S., Jarvis M. J., Hale C., Fonseca J., 2020a, *MNRAS*, 492, 1513
 Gomes Z., Camera S., Jarvis M. J., Hale C., Fonseca J., 2020b, *MNRAS*, 492, 1513
 Gruppioni C., Mignoli M., Zamorani G., 1999, *MNRAS*, 304, 199
 Hale C. L., Jarvis M. J., Delvecchio I., Hatfield P. W., Novak M., Smolčić V., Zamorani G., 2018, *Monthly Notices of the Royal Astronomical Society*, 474, 4133
 Hale C. L., et al., 2024, *MNRAS*, 527, 6540

- Heckman T. M., Best P. N., 2014, [ARA&A](#), 52, 589
- Heesen V., et al., 2022, [A&A](#), 664, A83
- Hopkins A. M., Afonso J., Chan B., Cram L. E., Georgakakis A., Mobasher B., 2003, [AJ](#), 125, 465
- Ibar E., Ivison R. J., Biggs A. D., Lal D. V., Best P. N., Green D. A., 2009, [MNRAS](#), 397, 281
- Janssen R. M. J., Röttgering H. J. A., Best P. N., Brinchmann J., 2012, [A&A](#), 541, A62
- Kellermann K. I., Fomalont E. B., Mainieri V., Padovani P., Rosati P., Shaver P., Tozzi P., Miller N., 2008, [ApJS](#), 179, 71
- Kugel R., et al., 2023, [MNRAS](#), 526, 6103
- Lin Y.-T., Shen Y., Strauss M. A., Richards G. T., Lunnan R., 2010, [ApJ](#), 723, 1119
- Lindsay S. N., et al., 2014, [MNRAS](#), 440, 1527
- Madau P., Dickinson M., 2014, [ARA&A](#), 52, 415
- Magliocchetti M., et al., 2004, [MNRAS](#), 350, 1485
- Magliocchetti M., Popesso P., Brusa M., Salvato M., Laigle C., McCracken H. J., Ilbert O., 2017, [MNRAS](#), 464, 3271
- Maldacena J., 2003, [Journal of High Energy Physics](#), 2003, 013
- Mancuso C., et al., 2015, [ApJ](#), 810, 72
- Mandal S., et al., 2021, [A&A](#), 648, A5
- Massardi M., Bonaldi A., Negrello M., Ricciardi S., Raccanelli A., de Zotti G., 2010, [MNRAS](#), 404, 532
- Matthews A. M., et al., 2024, [ApJ](#), 966, 194
- McGibbon R., Helly J., Schaye J., Schaller M., Vandenbroucke B., 2025, [The Journal of Open Source Software](#), 10, 8252
- Mitchell K. J., Condon J. J., 1985, [AJ](#), 90, 1957
- Molnár D. C., et al., 2021, [MNRAS](#), 504, 118
- Murphy E. J., Chary R. R., Dickinson M., Pope A., Frayer D. T., Lin L., 2011, [ApJ](#), 732, 126
- Owen F. N., Morrison G. E., 2008, [AJ](#), 136, 1889
- Planck Collaboration et al., 2020, [A&A](#), 641, A9
- Richards E. A., 2000, [ApJ](#), 533, 611
- Schaller M., et al., 2024, [MNRAS](#), 530, 2378
- Schaye J., et al., 2023, [MNRAS](#), 526, 4978
- Seljak U. c. v., 2009, [Phys. Rev. Lett.](#), 102, 021302
- Seymour N., et al., 2008, [MNRAS](#), 386, 1695
- Shimwell T. W., et al., 2017, [A&A](#), 598, A104
- Shimwell T. W., et al., 2022a, [A&A](#), 659, A1
- Shimwell T. W., et al., 2022b, [A&A](#), 659, A1
- Smolčić V., et al., 2017, [A&A](#), 602, A2
- Urry C. M., Padovani P., 1995, [PASP](#), 107, 803
- Wagenveld J. D., von Hausegger S., Klöckner H. R., Schwarz D. J., 2025, [A&A](#), 697, A112
- White R. L., Becker R. H., Helfand D. J., Gregg M. D., 1997, [ApJ](#), 475, 479
- Wilman R. J., et al., 2008, [MNRAS](#), 388, 1335
- Yamauchi D., Takahashi K., Oguri M., 2014, [Phys. Rev. D](#), 90, 083520
- de Zotti G., Massardi M., Negrello M., Wall J., 2010, [A&A Rev.](#), 18, 1
- van Haarlem M. P., et al., 2013, [A&A](#), 556, A2
- van Hoof P. A. M., Williams R. J. R., Volk K., Chatzikos M., Ferland G. J., Lykins M., Porter R. L., Wang Y., 2014, [MNRAS](#), 444, 420

A WORKING EQUATIONS USED IN GHOST

This appendix collects the definitions and working relations used throughout GHOST and links them to the survey selections and the observables we compare to data.

A.1 Notation, K -corrections, and selection mapping

We define the observed flux density at frequency ν_{obs} by

$$S_{\nu_{\text{obs}}} = \frac{L_{\nu_{\text{rest}}}}{4\pi D_L^2(z)} \frac{1}{1+z}, \quad \nu_{\text{rest}} = (1+z)\nu_{\text{obs}}, \quad (35)$$

with D_L the luminosity distance. For a power-law spectrum $L_\nu \propto \nu^{-\alpha}$ the usual radio K -correction implies

$$L_{\nu_{\text{rest}}} = 4\pi D_L^2(z) S_{\nu_{\text{obs}}} (1+z)^{1+\alpha}. \quad (36)$$

Hence a survey limit in $S_{\nu_{\text{obs}}}$ translates into a luminosity threshold at redshift z for a class with characteristic α_c :

$$L_{\text{lim}}(z, \alpha_c) = 4\pi D_L^2(z) S_{\text{lim}} (1+z)^{1+\alpha_c}. \quad (37)$$

Within any luminosity/redshift bin $[L_1, L_2]$, the fraction of sources visible at z is

$$f_{\text{vis}}(z; L_1, L_2) = \begin{cases} 1, & L_{\text{lim}}(z, \alpha_c) \leq L_1, \\ \frac{\log_{10}(L_2/L_{\text{lim}}(z, \alpha_c))}{\log_{10}(L_2/L_1)}, & L_1 < L_{\text{lim}}(z, \alpha_c) < L_2, \\ 0, & L_{\text{lim}}(z, \alpha_c) \geq L_2. \end{cases} \quad (38)$$

As redshift increases, L_{lim} increases in $\log L$ and progressively trims the faint end of each bin; for a fixed flux limit, this effect is strongest for SFG (larger ζ , lower intrinsic L) and weakest for flat-spectrum AGN. This geometric gate is the lower envelope seen in the L - z panels and drives selection-induced changes in the slope of $w(\theta)$.

The per-steradian redshift distribution for class X under a flux cut is

$$n_X(z) = \frac{dN_X}{dz d\Omega} = \frac{dV}{dz d\Omega} \int d\log_{10} L \Phi_X(L, z) \Theta[L - L_{\text{lim}}(z, \alpha_X)], \quad (39)$$

and the differential counts follow from the change of variables $L \leftrightarrow S$:

$$\frac{dN}{dS}(S) = \int dz \frac{dV}{dz d\Omega} \Phi(L, z) \left. \frac{d\log_{10} L}{dS} \right|_{L=4\pi D_L^2 S(1+z)^{1+\alpha}}, \quad (40)$$

with $d\log_{10} L/dS = [\ln 10]^{-1} S^{-1}$. In practice, total counts constrain the *aggregate* Φ (summing over classes), while $n_X(z)$ and clustering are sensitive to how Φ_X partitions that total across classes and redshifts.

Table 7 Best-fit parameters of the AGN evolutionary RLF adopted from T-RECS. The slopes γ and β correspond to a and b in Bonaldi et al. (2019). Parameters k_{evo} , $z_{\text{top},0}$, δz_{top} , and m_{evo} enter Equations 42–43.

Parameter	FSRQ	BL Lac	SS-AGN
γ (faint-end slope)	0.776	0.723	0.508
β (bright-end slope)	2.669	1.918	2.545
$\log_{10} n_0 [\text{Mpc}^{-3} \text{dex}^{-1}]$	−8.319	−7.165	−5.973
$\log_{10} L_\star(0) [\text{W Hz}^{-1}]$	33.268	32.282	32.560
k_{evo}	1.234	0.206	1.349
$z_{\text{top},0}$	2.062	1.262	1.116
δz_{top}	0.559	—	0.705
m_{evo}	0.136	1.000	0.253

A.2 AGN radio luminosity function and evolution

We adopt the double-power-law form of Bonaldi et al. (2019) with redshift evolution implemented via a de-evolution mapping $L(z) \rightarrow L(0)$:

$$\Phi(L|z) = \frac{n_0}{\left(\frac{L(0)}{L_\star(0)}\right)^\gamma + \left(\frac{L(0)}{L_\star(0)}\right)^\beta} \cdot \frac{d\log_{10} L(0)}{d\log_{10} L(z)}. \quad (41)$$

For a separable evolution $L(0) = L(z)/\mathcal{E}(z)$ the Jacobian equals unity, but we keep it explicit to allow more general mappings.

The characteristic luminosity evolves as

$$L_\star(z) = L_\star(0) \text{dex} \left[k_{\text{evo}} z \left(2z_{\text{top}} - \frac{z^{m_{\text{evo}}} z_{\text{top}}^{1-m_{\text{evo}}}}{1+m_{\text{evo}}} \right) \right], \quad (42)$$

while the peak redshift depends on luminosity through

$$z_{\text{top}}(L) = z_{\text{top},0} + \frac{\delta z_{\text{top}}}{1 + (L(0)/L_\star(0))}. \quad (43)$$

Together, Equations 42–43 produce an early rise and late decline of $L_\star(z)$ and encode “downsizing,” with higher-luminosity AGN peaking earlier. The faint/bright slopes (γ, β) govern the curvature of $S^{2.5} dN/dS$ at the bright end. Given Φ , expected counts in a $(\log L, z)$ bin are

$$\Delta N_{i,j} = \Omega \Phi(L|z_i) \left. \frac{dV}{dz} \right|_{z_i} \Delta \log_{10} L_{i,j} \Delta z_i, \quad (44)$$

the building block for both $n(z)$ and dN/dS via Equations 39–40. The values of the constants in Equations 41 and 43 are given in Table 7.

A.3 HERG/LERG host probabilities and effective bias

Steep-spectrum AGN are split into excitation classes using stellar-mass scalings from Janssen et al. (2012),

$$P_{\text{LERG}}(M_\star|z) \propto \left(\frac{M_\star}{M_0}\right)^{2.5} \quad M_\star \lesssim 10^{11.6} M_\odot, \quad (45)$$

$$P_{\text{HERG}}(M_\star|z) \propto \left(\frac{M_\star}{M_0}\right)^{1.5}, \quad (46)$$

renormalised over candidate hosts. Mapping $M_\star$ to halo mass M via the stellar–halo relation (next subsection) yields the class-effective linear bias

$$b_t(z) = \frac{\int dM n(M, z) b(M, z) W_t(M, z)}{\int dM n(M, z) W_t(M, z)}, \quad (47)$$

$$W_t \equiv P_t(M_\star(M, z), z), \quad (48)$$

with $n(M, z)$ the halo mass function and $b(M, z)$ a standard halo-bias model. Since HERG-like weights favour higher $M_\star$, Equation 48 explains the robust bias hierarchy relative to SFG.

A.4 Sizes

Our catalogue attaches a minimalist morphology proxy to each radio AGN solely to split the SSAGN pool into HERG/LERG-like sub-classes and apply class-dependent, stellar-mass-weighted halo assignment. Sizes do not affect photometry or clustering beyond the sub-class/host selection described here.

Following the approach of Bonaldi et al. (2019), each AGN is given a characteristic linear size D (kpc) via a stochastic draw from the broad, observationally motivated distribution of DiPompeo et al. (2013) with a weak luminosity trend. A random sky orientation is applied (uniform in $\sin \theta$). We do not simulate shapes or images; D is a single scale per source used only as a morphology proxy.

We compute a size-ratio

$$R_s \equiv \frac{d_{\text{peak}}}{D_{\text{tot}}}, \quad (49)$$

the separation of the two peak-brightness regions d_{peak} divided by the end-to-end extent D_{tot} . We adopt the usual FR cut $R_s > 0.5 \Rightarrow \text{FR II}$ and $R_s < 0.5 \Rightarrow \text{FR I}$. The draw of R_s is mildly luminosity-dependent: for $L_\nu \leq 10^{25.4} \text{ W Hz}^{-1}$ we take $R_s \sim \mathcal{N}(\mu = 0.18, \sigma = 0.11)$, and for $L_\nu > 10^{25.4} \text{ W Hz}^{-1}$ we take $R_s \sim \mathcal{N}(\mu = 0.62, \sigma = 0.18)$. This biases high-power sources toward FR II while still allowing overlap in R_s at fixed L .

We then assign the spectroscopic class with a *deterministic* rule at fixed R_s : sources with $R_s > 0.5$ are

tagged as HERG, and those with $R_s \leq 0.5$ as LERG. Because R_s itself is drawn with scatter, the resulting fractions vary smoothly with luminosity; in particular

$$P(\text{HERG} | L) = P(R_s > 0.5 | L) = 1 - \Phi\left(\frac{0.5 - \mu(L)}{\sigma(L)}\right), \quad (50)$$

with Φ the standard normal CDF and $(\mu(L), \sigma(L))$ given above. Hence FR II–LERG and FR I–HERG combinations naturally occur, even though the label is a hard threshold in R_s .

A.5 Stellar–halo mass relation (for host selection)

Host selection uses a redshift-dependent double-power relation (Bonaldi et al., 2019; Aversa et al., 2015),

$$M_\star(M, z) = N_0(z) \left[\left(\frac{M}{M_b(z)}\right)^{\epsilon(z)} + \left(\frac{M}{M_b(z)}\right)^{\omega(z)} \right]^{-1}, \quad (51)$$

interpolated from the fit of Aversa et al. (2015). This sets AGN host probabilities (via $M_\star$) and provides the smooth proxy used to fit the SFG $L(M)$ map. Steepening at high mass boosts HERG bias relative to LERG; shifts in $M_b(z)$ move the characteristic host mass and reshape class-specific $n(z)$ tails.

A.6 SFG radio luminosities: synchrotron and free-free

We model the SFG spectrum as a sum of non-thermal synchrotron and thermal free-free components (Bonaldi et al., 2019; Mancuso et al., 2015). The synchrotron component is

$$L_{\text{sync}}(\nu) = \frac{L_{*,\text{sync}}(\nu)}{\left(\frac{L_{*,\text{sync}}}{\bar{L}_{\text{sync}}}\right)^\beta + \left(\frac{L_{*,\text{sync}}}{\bar{L}_{\text{sync}}}\right)}, \quad (52)$$

with fiducial scaling

$$\bar{L}_{\text{sync}} = 1.9 \times 10^{21} \left(\frac{\text{SFR}}{M_\odot \text{ yr}^{-1}}\right) \left(\frac{\nu}{\text{GHz}}\right)^{-0.85} \left[1 + \left(\frac{\nu}{20 \text{ GHz}}\right)^{0.5}\right]^{-1} \text{ W Hz}^{-1}. \quad (53)$$

The free-free component follows

$$L_{\text{ff}}(\nu) \propto T^{-0.5} g_{\text{ff}}(\nu, T) \nu^{-0.1}, \quad T = 10^4 \text{ K}, \quad (54)$$

with Gaunt factor (van Hoof et al., 2014)

$$g_{\text{ff}}(\nu_{\text{GHz}}, T) = \ln\left(e + \exp\left[5.960 - \frac{\sqrt{3}}{\pi} \ln(\nu_{\text{GHz}} T_4^{-3/2})\right]\right), \quad T_4 \equiv T/10^4 \text{ K}. \quad (55)$$

Below $\sim 10\text{--}20$ GHz, synchrotron with $\zeta \approx 0.8\text{--}0.9$ dominates; free-free adds a flatter $\zeta \approx 0.1$ tail, making extrapolations to a few GHz slightly shallower than pure synchrotron. Together with a smoothly evolving SFG RLF, this SED reproduces multi-frequency counts.

A.7 Abundance matching and the working $L(M)$ map

We assign radio luminosities to halos monotonically in each redshift slice so the realised luminosity histogram reproduces the target RLF. For SFG we fit

$$L(M; z) = N(z) \left[\left(\frac{M}{M_*(z)} \right)^\gamma + \left(\frac{M}{M_*(z)} \right)^\beta \right]^{-1}, \quad (56)$$

apply a log-normal scatter $\sigma_{\log L}$, and perform rank-ordering without replacement. Parameters are inferred per slice with a Poisson likelihood for the RLF-implied counts. Equation 56 behaves as a flexible double-slope map: γ controls the faint-end response to host mass, β the bright-end roll-off, M_* the transition scale, and N the normalisation; the scatter compresses the effective $L\text{--}M$ relation and slightly lowers bias at fixed L . For AGN, we first draw hosts with the mass-weighted HERG/LERG probabilities and then assign L monotonically within the selected subset to match the class RLFs, keeping bias tied to halo mass (Equation 48).

A.8 From populations to observables: effective kernels and bias

Given the above ingredients, the selection-matched redshift kernel for class X is

$$\varphi_X(z) = \frac{n_X(z)}{\int n_X(z') dz'}, \quad (57)$$

where $n_X(z)$ is given by Eq. 39.

For the auto-correlation of tracer X , the effective large-scale bias entering C_ℓ or $w(\theta)$ is

$$b_{\text{eff}, X} = \left[\frac{\int dz \varphi_X^2(z) b_X^2(z) D^2(z)}{\int dz \varphi_X^2(z) D^2(z)} \right]^{1/2}. \quad (58)$$

with $D(z)$ the growth factor, that is, $\delta_m(k, z) = D(z)\delta_m(k, z=0)$, and $b_X(z)$ from Equation 48. Changes to the survey selection alter $\varphi_X(z)$ (projection) and therefore the amplitude/slope of $w(\theta)$ without changing the intrinsic bias ordering set by $b_X(z)$.

A.9 SFR function

We use the SFR density function (constrained by observations) from (Bonaldi et al., 2019; Mancuso et al.,

2015)

$$\phi_{\text{SFR}}(z) = \Phi(z) \left[\frac{\text{SFR}}{\text{SFR}_*(z)} \right]^{1-\zeta(z)} \exp \left\{ - \left[\frac{\text{SFR}}{\text{SFR}_*(z)} \right]^{w(z)} \right\}, \quad (59)$$

with redshift evolution

$$\log \Phi(z) [\text{Mpc}^{-3} \log(\text{SFR})^{-1}] = -2.4 - 2.3\chi + 6.2\chi^2 - 4.9\chi^3, \quad (60)$$

$$\log \text{SFR}_*(z) [M_\odot \text{ yr}^{-1}] = 1.1 + 3.2\chi - 1.4\chi^2 - 2.1\chi^3, \quad (61)$$

$$\zeta(z) = 1.2 + 0.5\chi - 0.5\chi^2 + 0.2\chi^3, \quad (62)$$

$$w(z) = 0.7 - 0.15\chi + 0.16\chi^2 + 0.01\chi^3. \quad (63)$$

where $\chi = \log(1+z)$. In GHOST the SFG branch is radio-first; this SFRF is kept as an optional cross-check to sanity-test the bright tail of Φ_{SFG} at $z \gtrsim 1$ without imposing FIR-radio priors.