

Searching for HI around MHONGOOSE galaxies via spectral stacking

S. Veronese^{1,2,*}, W. J. G. de Blok^{1,2,3}, J. Healy¹, D. Kleiner^{1,4}, A. Marasco⁵, F. M. Maccagni⁴, P. Kamphuis⁶, E. Brinks⁷, B. W. Holwerda⁸, N. Zabel³, L. Chemin⁹, E. A. K. Adams^{1,2}, S. Kurapati^{3,1}, A. Sorgho¹⁰, K. Spekkens¹¹, F. Combes¹², D. J. Pisano³, F. Walter¹³, P. Amram¹⁴, F. Bigiel¹⁵, O. I. Wong^{16,17}, and E. Athanassoula¹⁴

¹ Netherlands Institute for Radio Astronomy (ASTRON), Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands

² Kapteyn Astronomical Institute, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands

³ Department of Astronomy, University of Cape Town, Private Bag X3, 7701 Rondebosch, South Africa

⁴ INAF – Osservatorio Astronomico di Cagliari, Via della Scienza 5, 09047, Selargius, CA, Italy

⁵ INAF – Padova Astronomical Observatory, Vicolo dell’Osservatorio 5, I-35122 Padova, Italy

⁶ Ruhr University Bochum, Faculty of Physics and Astronomy, Astronomical Institute (AIRUB), 44780 Bochum, Germany

⁷ Centre for Astrophysics Research, University of Hertfordshire, College Lane, Hatfield AL10 9AB, United Kingdom

⁸ Department of Physics and Astronomy, University of Louisville, Natural Science Building 102, Louisville, KY 40292, USA

⁹ Universidad Andrés Bello, Facultad de Ciencias Exactas, Departamento de Ciencias Físicas – Instituto de Astrofísica, Fernández Concha 700, Las Condes, Santiago, Chile

¹⁰ Instituto de Astrofísica de Andalucía (IAA-CSIC), Glorieta de la Astronomía s/n, E-18008 Granada, Spain

¹¹ Department of Physics, Engineering Physics and Astronomy, Queen’s University, Kingston, ON K7L 3N6, Canada

¹² Observatoire de Paris, LERMA, Collège de France, CNRS, PSL University, Sorbonne University, 75014 Paris, France

¹³ Max Planck Institute für Astronomie, Heidelberg, Germany

¹⁴ Aix-Marseille Univ., CNRS, CNES, LAM, 38 rue Frédéric Joliot Curie, 13338 Marseille, France

¹⁵ Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany

¹⁶ Australia Telescope National Facility, CSIRO, Space and Astronomy, PO Box 1130, Bentley, WA 6102, Australia

¹⁷ International Centre for Radio Astronomy Research, The University of Western Australia, Crawley, WA 6009, Australia

Received 2 September 2024 / Accepted 15 November 2024

ABSTRACT

The observed star formation rates of galaxies in the Local Universe suggests that they are replenishing their gas reservoir across cosmic time. Cosmological simulations predict that this accretion of fresh gas can occur in a hot or a cold mode, yet the existence of low column density ($\sim 10^{17} \text{ cm}^{-2}$) neutral atomic hydrogen (HI) tracing the cold mode has not been unambiguously confirmed by observations. We present the application of unconstrained spectral stacking to attempt to detect the emission from this HI in the circumgalactic medium (CGM) and intergalactic medium (IGM) of six nearby star-forming galaxies from the MHONGOOSE sample for which full-depth observations are available. Our stacking procedure consists of a standard spectral stacking algorithm coupled with a one-dimensional spectral line finder designed to extract a reliable signal close to the noise level. In agreement with previous studies, we find that the amount of signal detected outside the HI disk is much smaller than implied by simulations. Furthermore, the column density limit that we achieve via stacking ($\sim 10^{17} \text{ cm}^{-2}$) suggests that direct detection of the neutral CGM and IGM component might be challenging in the future, even with the next generation of radio telescopes.

Key words. methods: statistical – galaxies: evolution – intergalactic medium

1. Introduction

The process of cold gas accretion, whereby pristine gas from the intergalactic medium (IGM; Dekel & Birnboim 2006; Somerville & Davé 2015; Danovich et al. 2015) or processed gas from the circumgalactic medium (CGM; Putman et al. 2012; Afruni et al. 2021; Li et al. 2021; Marasco et al. 2022) or from galaxy mergers (Sancisi et al. 2008; Putman et al. 2012; de Blok et al. 2020; Porter et al. 2022) fuels the growth of galaxies, is a crucial mechanism in galaxy formation and evolution.

It is believed that neutral atomic hydrogen (HI) clouds in the IGM and CGM serve as a tracer for the reservoirs of cool

($T < 10^5 \text{ K}$) gas that can be funnelled onto galaxies, providing the necessary fuel for star formation and driving the growth of stellar mass over cosmic time (Sancisi et al. 2008; Madau & Dickinson 2014; Davé et al. 2017; Tumlinson et al. 2017; Chen et al. 2021; Faucher-Giguère & Oh 2023). The detection and characterisation of these HI clouds is essential for understanding the efficiency of their accretion onto galaxies (Dekel et al. 2009; Putman et al. 2012; Kamphuis et al. 2022; Afruni et al. 2023; Lochhaas et al. 2023). By studying their spatial distribution, abundance, and physical properties, we can gain insights into the mechanisms responsible for the transport and deposition of gas from the cosmic web onto galaxies (Popping et al. 2009; Somerville & Davé 2015; Schaye et al. 2015; Hopkins et al. 2018; Nelson et al. 2019; van de Voort et al. 2019; Ramesh & Nelson 2024). Additionally,

* Corresponding author; veronese@astron.nl

the kinematics and gas dynamics of the HI clouds can provide valuable information about the accretion processes and the interplay between galaxies and their surrounding environment.

The presence of approximate kiloparsec-scale, low column density ($N_{\text{HI}} \sim 10^{17} \text{ cm}^{-2}$), HI clouds in the CGM and IGM has been predicted by simulations (Popping et al. 2009; van de Voort et al. 2019; Ramesh & Nelson 2024), yet not unambiguously confirmed observationally (Braun & Thilker 2004; Wolfe et al. 2013, 2016; Tumlinson et al. 2017; Das et al. 2020; Kamphuis et al. 2022; Xu et al. 2022; Liu et al. 2023; Das et al. 2024). The detection of CGM and IGM clouds has been a challenging endeavour. On the one hand, by looking at the HI absorption against background quasars, we can reach column densities well below 10^{17} cm^{-2} but at the cost of having few single pencil-beam measurements per galaxy, given the sparse distribution of sufficiently bright background quasars (Tumlinson et al. 2013; Werk et al. 2014; Turner et al. 2014; Rubin et al. 2015; Lehner et al. 2015; Tumlinson et al. 2017; Tchernyshyov et al. 2023; Afruni et al. 2023). On the other hand, the detection of the HI clouds in emission has always been limited by either spatial resolution or column density sensitivity.

Single-dish radio telescopes are able to reach limiting column densities of 10^{17} cm^{-2} in HI but with poor spatial resolution ($>200''$). Outside the Local Group, this size corresponds to $>3 \text{ kpc}$ and this is insufficient to constrain the nature of an isolated detection, that is, it is not possible to distinguish between a primordial HI cloud or a tiny dwarf galaxy, or a tidal feature (Fraternali et al. 2001; Thilker et al. 2004; Westmeier et al. 2007; de Blok et al. 2014; Kerp et al. 2016; Veronese et al. 2023). Beam dilution can also wash out the signal from a gas cloud if it is very small with respect to the beam.

Previous generations of interferometric observatories achieved the necessary spatial resolution ($<100''$) but their column density sensitivity was usually not better than $\sim 10^{19} \text{ cm}^{-2}$. Previous surveys (e.g. HIPASS, Barnes et al. 2001; ALFALFA, Giovanelli et al. 2007; THINGS, Walter et al. 2008; HALOGAS, Heald et al. 2011) have made significant contributions to the understanding of the HI content in galaxies. However, for the reasons just described, they were limited in their ability to probe the diffuse atomic gas in the CGM and IGM, which remains largely unexplored. The MeerKAT HI Observations of Nearby Galactic Objects: Observing Southern Emitters (MHONGOOSE, de Blok et al. 2024) survey represents a recent advancement in observational capabilities, leveraging high sensitivity ($<10^{18} \text{ cm}^{-2}$) coupled with the excellent spatial resolution ($\sim 30''$) and wide-field ($\sim 1.5^\circ$) coverage of the MeerKAT radio telescope¹ (Jonas & MeerKAT Team 2016). The MHONGOOSE survey aims to detect and study the HI gas in nearby star-forming galaxies, exploring their kinematics, gas dynamics, and environmental effects. The rich dataset collected by MHONGOOSE also provides an opportunity to investigate the presence of faint signatures of pristine clouds embedded within the CGM and IGM. While this is usually done by attempting to find direct detections in data cubes with source finders, or even by eye, we can also use alternative techniques that can push the sensitivity limit for the detection of faint HI structures around nearby galaxies even further down.

The most well-known of these is the spectral stacking, which consists of co-adding a large number of individual spectra with known redshifts to improve the signal-to-noise ratio (S/N) of the final spectrum. HI spectral stacking has been employed mainly to characterise the average HI properties of samples of galaxies (Fabello et al. 2011; Delhaize et al. 2013; Geréb et al. 2014, 2015a,b; Kanekar et al. 2016; Maccagni et al. 2017; Ianjamasimanana et al. 2018; Healy et al. 2019, 2021; Chowdhury et al. 2022; Amiri et al. 2023; Bera et al. 2023), but to date two studies used it to systematically look for HI emission in the CGM and IGM around galaxies (Das et al. 2020, 2024).

For this paper we used this technique to search for HI clouds around the MHONGOOSE galaxies but with a new approach. What differs from previous stacking experiments is that in this case we know the redshift of each galaxy, but there is no information about how the gas around them (i.e. outside the disk) is moving with respect to the disk, where it is located and what its morphology is. In other words, we performed a kind of “unconstrained” stacking. For this reason, we used state-of-the-art cosmological simulations to find the best priors for our experiment, such as the size of the HI clouds and their velocity distribution, prior to applying the technique to the MHONGOOSE data. We also employed a one-dimensional spectral-line finder, based on the Source Finding Application (SoFIA-2, Serra et al. 2015; Westmeier et al. 2021), designed to deal with low S/N lines at unknown spectral and spatial locations.

The rest of the paper is as follows. In Sect. 2 we describe the two datasets used in this paper: cosmological simulations that we used to test our methods and the observational data from the MHONGOOSE survey that we used for the actual stacking. The employed stacking and source finder algorithm is illustrated in Sect. 3 and its testing and calibration with simulated data is provided in Sect. 4. We present its application to the MHONGOOSE galaxies and the results in Sect. 5, discussing also the statistical significance of the detected signals in the real data. A summary of our conclusions is provided in Sect. 6, while some future prospects are briefly outlined in Sect. 7.

2. Dataset

To test our algorithm we used mock HI data based on the TNG50 simulation (Nelson et al. 2019; Ramesh et al. 2023), which predicts the presence of low-column density HI clouds around galaxies. This enabled us to calibrate our algorithm and understand its limitations. Once the performance of our algorithm was assessed using these mock data, we applied it to the HI cubes from the MHONGOOSE survey. In the following we briefly describe some of the properties of both the observed and simulated data.

2.1. The MHONGOOSE sample

The MHONGOOSE survey (de Blok et al. 2024) using MeerKAT has performed ultra-deep HI observations of 30 nearby gas-rich dwarf and spiral galaxies spanning a wide range of morphology, stellar mass ($4.7 \leq \log M_* \leq 10.7 M_\odot$) and star formation rate ($-2.52 \leq \log \text{SFR} \leq 0.75 M_\odot \text{ yr}^{-1}$). One of the main goals of the project is to identify and characterise the distribution of low column density HI, which could potentially be inflowing towards the galaxies. These observations also

¹ The MeerKAT telescope is operated by the South African Radio Astronomy Observatory, which is a facility of the National Research Foundation, an agency of the Department of Science and Innovation.

Table 1. Main properties of the simulated galaxies used in this work.

ID	M_{halo} [$10^{12} M_{\odot}$]	M_{*} [$10^{10} M_{\odot}$]	M_{HI} [$10^{10} M_{\odot}$]	SFR [$M_{\odot} \text{ yr}^{-1}$]
520885	1.17	5.47	3.17	2.76
555013	0.89	3.66	1.85	1.48

Notes. The galaxies are part of the TNG50 Milky Way-like sample provided by Ramesh et al. (2023) and the HI properties has been derived by Marasco et al. (in prep). The ID refers to the subhalo index in the group catalogues as determined by the subfind algorithm (Dolag et al. 2009). M_{halo} , M_{*} and M_{HI} are the dark matter halo, stellar and HI masses, respectively. SFR is the current star formation rate.

seek to establish connections between this HI and ongoing star formation processes within these galaxies.

Each of the 30 targets is observed for a total of 55 h, spread across 10 observing sessions lasting 5.5 h each. All ten observation tracks underwent the same calibration procedure and were combined to produce the “full-depth” cube. For every galaxy we also have a “single-track” cube, obtained from data reduction of a single 5.5 h observation.

The calibration was performed using the Containerized Automated Radio Astronomy Calibration (CARAcal) pipeline (Józsa et al. 2020) as described in de Blok et al. (2024). CARAcal provides a unified platform encompassing the standard calibration and reduction steps, including data flagging, cross-calibration, target separation, self-calibration, as well as continuum subtraction, spectral line imaging, and deconvolution.

The resultant HI cubes were generated via WSClean (Offringa et al. 2014; Offringa & Smirnov 2017) within the CARAcal pipeline. Cleaning of the HI data followed a three-step iterative strategy which uses SoFIA-2 to produce the cleaning masks. The process exploited the full MeerKAT resolution and sensitivity capabilities by producing six distinct HI cubes for each galaxy with a 1.5° field of view (FoV). These cubes span a range of resolutions, from $7''$ to $90''$, achieved through variations in weighting and tapering.

As in this paper we aim to look for the HI at the $\sim 10^{17} \text{ cm}^{-2}$ column density level, we refrain using the high resolution cubes (r00_t00 and r05_t00 in de Blok et al. 2024) as this is not computationally efficient given their limiting column density being more than two order of magnitudes higher. The lowest resolution cubes (r05_t60 and r10_t90) were also excluded given their sidelobe level being potentially able to affect the stacking. As the mock cubes described in Sect. 2.2 are only used for testing purposes and a thorough comparison between TNG50 and MHONGOOSE will be provided in Marasco et al. (in prep.), between the intermediate resolution cubes r10_t00 and r15_t00 we decided for the latter because of its lower limiting column density (3σ , 16 km s^{-1} sensitivity of $2.8 \times 10^{18} \text{ cm}^{-2}$). The beam size of $34.4'' \times 25.4''$ is slightly larger than for the mock cubes ($26'' \times 18''$), however, as described in Sects. 3.1 and 4.1 the effect of the beam will be mitigated via re-gridding and the stacking will be performed on scales much larger than the beam.

For the r15_t00 cubes we also used their derived SoFIA-2 products obtained as described in de Blok et al. (2024): zeroth moment (intensity), first moment (intensity-weighted line-of-sight velocity), and second moment (line-of-sight dispersion) maps, as well as the detection masks used to produce these maps. For this work we used the single-track and full-depth cubes of

the 18 MHONGOOSE galaxies having inclination angle $\leq 60^{\circ}$. This selection is described in more detail in Sect. 4.2.

2.2. The TNG50 galaxies

The mock data used to test our algorithm were taken from IllustrisTNG (Nelson et al. 2019), a collection of extensive, cosmological simulations carried out using the AREPO moving-mesh code (Springel 2010). The project comprises of three distinct simulation volumes (comoving box sizes: approximately 50, 100, and 300 Mpc on each side) and incorporates a comprehensive framework of galaxy formation physics, enabling a realistic representation of galaxy evolution across cosmic epochs (Weinberger et al. 2017; Pillepich et al. 2018).

For this work, we selected from the TNG50 Milky Way-like sample provided by Ramesh et al. (2023) two galaxies, 520885 and 555013² whose main properties are listed in Table 1. In Figs. 1 and 2 we show their high-resolution HI intensity and intensity-weighted line-of-sight velocity map³, respectively, and as observed with an inclination angle of 60 degrees. We chose these two galaxies because of their diverse HI morphology, consisting of a well-defined HI disk surrounded by clumpy and filamentary faint HI features located up to ~ 100 kpc of distance from the centre. These features are more evident for system 520885 (left panel of Fig. 1). The variety of HI structures shown by these two galaxies make them optimal testing targets for the stacking algorithm we use in this paper.

The derivation of these HI distributions is presented in Marasco et al. (in prep.). The goal of that paper is to derive the HI distributions of a sample of simulated Milky Way-like galaxies, and to infer what they would have looked like if these systems were “observed” as a part of the MHONGOOSE survey. To achieve this, Marasco et al. (in prep.) produce synthetic observed HI cubes for each simulated system matching the resolution, noise level and field of view of a real MHONGOOSE dataset. Here we use the synthetic data cubes for our two simulated test galaxies, derived with a robustness parameter of 1.0 (r10_t00, cf. de Blok et al. 2024), giving a beam size of $26'' \times 18''$ and a 3σ column density sensitivity of $4.9 \times 10^{18} \text{ cm}^{-2}$ over 16 km s^{-1} . The simulated galaxies are all assumed to be at a distance of 20 Mpc, which is the average distance of the MHONGOOSE galaxies that Marasco et al. (in prep.) use as a comparison sample.

3. The algorithm

The algorithm we use here to search for HI signatures consists of a spectral stacking routine combined with a one-dimensional spectral-line finder. These two functions are wrapped into a publicly-available Python package⁴. Throughout the paper we refer to the spectral stacking task as STACKER and to the line finding task as FINDER. In the following sections we describe each task in detail and we refer to Sect. 4.1 for details on how the values of the input parameters were determined.

² The ID refers to the subhalo index in the group catalogues as determined by the subfind algorithm (Dolag et al. 2009).

³ The velocity field colourmap used throughout this paper is taken from English et al. (2024).

⁴ <https://github.com/SJVeronese/nicci-package>.

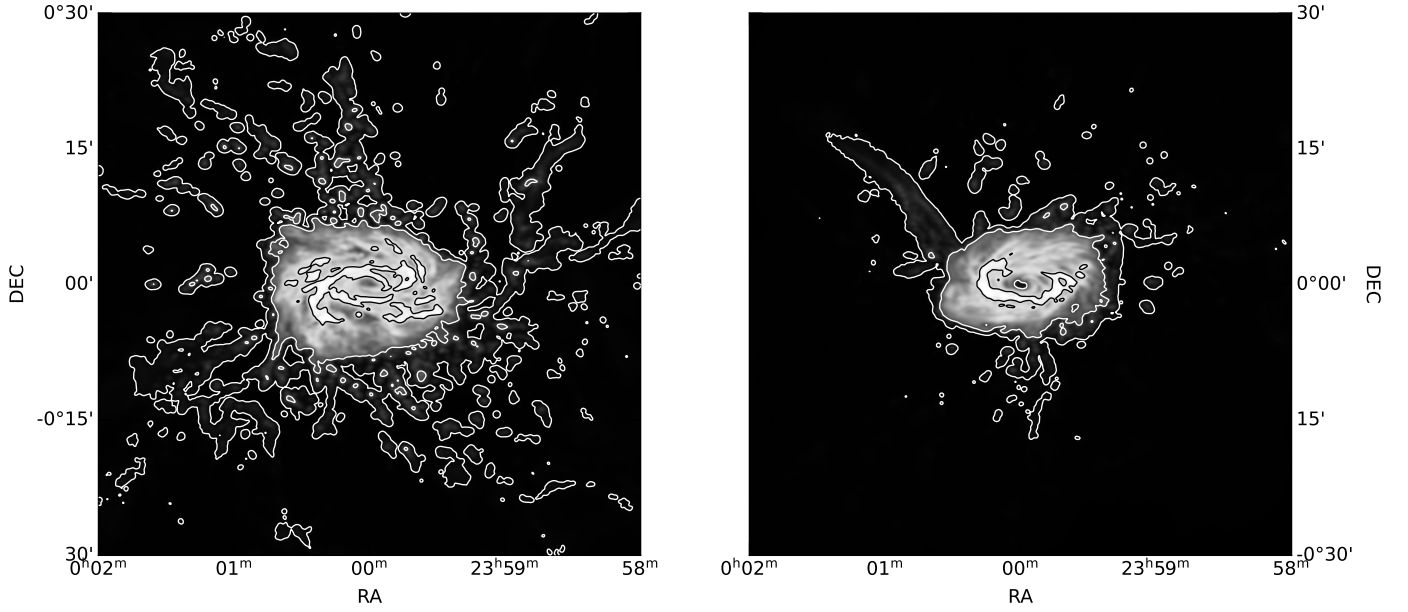


Fig. 1. Moment 0 map of the two TNG50 galaxies presented in this paper: 520855 is on the left and 555013 is on the right. Contours are denoting the $\text{Log}(N_{\text{HII}}) = (16, 19, 21) \text{ cm}^{-2}$ column density in both panels. The field of view has a side-length of 350 kpc, sufficient to encompass the virial radius of both galaxies.

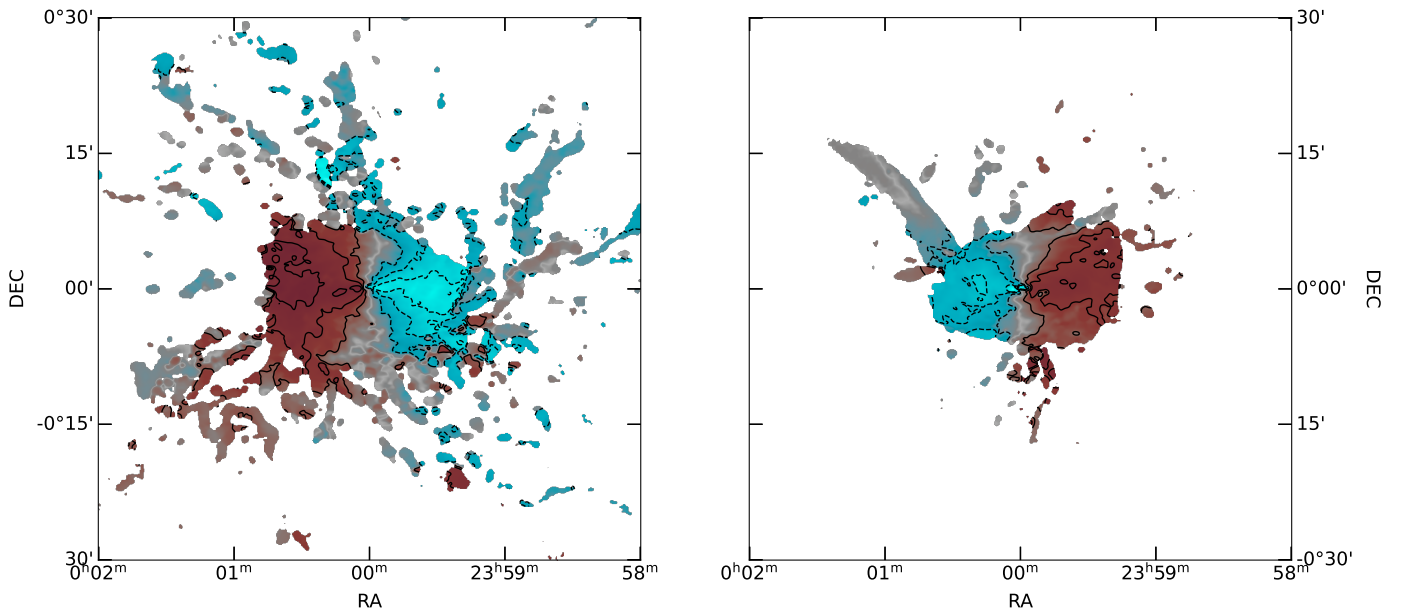


Fig. 2. Moment 1 map of the two TNG50 galaxies presented in this paper: 520855 is on the left and 555013 is on the right. Black solid-contours denote the receding (50, 100, 150) km s^{-1} velocities with respect to the systemic velocity. Black dashed-contours instead refer to the approaching (-150, -100, -50) km s^{-1} velocities. The maps are clipped at the 10^{16} cm^{-2} column density level, as this is the noise level we expected to achieve via stacking.

3.1. Spectral stacking

STACKER begins with several preliminary manipulations on the input data cube:

- it blanks specific lines-of-sight (LoS) based on a supplied mask to avoid stacking the emission from the main galaxy and other already-known sources. For this we used the SoFiA-2 detection mask that was also used to create the moment maps as described in [de Blok et al. \(2024\)](#);
- it aligns the spectra according to a velocity field, a process called “shuffle”. It consists of shifting each spaxel along the spectral axis so that its central channel corresponds to the value of the same LoS in the velocity field. To preserve the

noise properties of the data, the shifting wraps the spectrum around⁵, meaning it inserts the low frequency channels at the high frequency end. For the TNG50 galaxies we initially used the noiseless moment 1 map, as it contains the kinematical information of the gas outside the SoFiA-2 detection mask (see Sect. 4). As this information is missing when dealing with real data, we tested several velocity field assumptions as described in Sect. 4.2;

⁵ The spectral alignment is performed using our implementation of the GIPSY ([van der Hulst et al. 1992](#); [Vogelaar & Terlouw 2001](#)) task SHUFFLE.

- it re-grids the resulting shuffled cube to mitigate the spatial correlation of the voxels, so that the noise decreases as the square root of the number of co-added spectra (under the assumption of Gaussian noise). To achieve this, the pixel size is increased to match the beam size⁶.

We do not apply any primary beam correction at this stage because stacking requires the noise to be as uniform as possible. A primary beam correction increases the σ level towards the outskirts of the FoV.

After these preliminary manipulations, STACKER assigns each position on the sky to a stacking region, that we choose to be square cells (see Fig. 3). The stacked spectrum for each region is then computed as follow:

$$F_S = \frac{\sum_i F_i \cdot w_i}{\sum_i w_i}, \quad (1)$$

where F_S is the stacked flux in a given channel, calculated by summing the fluxes (F_i) of the spectra in that channel, weighted by their respective weights (w_i). The common weighting scheme used in the literature is the so-called “ σ^2 -weight” (e.g. Fabello et al. 2011; Kanekar et al. 2016; Ianjamasimanana et al. 2018; Healy et al. 2019, 2021; Chowdhury et al. 2022; Bera et al. 2023): $w_i = \frac{1}{\sigma_i^2}$, where σ_i represents the i th-spectrum’s noise level. As the presence of signal hidden in the noise potentially increases σ_i , leading to the down weighting of the corresponding spectrum and thus reducing the contribution of any hidden signal, we also considered the “equal-weight” scheme ($w_i = 1 \forall i$) and tested against the σ^2 -weight as described in Sect. 4.1.

The stacked spectra, one for each region, are passed to FINDER for source finding.

3.2. Source finding

FINDER employs a simplified one-dimensional version of the SOFIA-2 smooth-and-clip (S+C) algorithm, a procedure involving smoothing the data on multiple scales and apply for each scale a σ -clip (Serra et al. 2012b). The output is binary masks, one for each stacked spectrum, which are subsequently used to create the source catalogue by combining the detected channels in each mask, that is, assigning consecutive detections to a single source. Positive and negative sources spanning fewer than a given number of channels are removed from the catalogue.

Following the method described in detail by Serra et al. (2012a), FINDER computes the reliability of the remaining sources to discriminate genuine lines from noise peaks. Each source is placed in a three-dimensional space ($\frac{F_{\max}}{\sigma}$, $\frac{F_{\text{sum}}}{\sigma}$ and $\frac{F_{\text{mean}}}{\sigma}$), where F_{\max} , F_{sum} and F_{mean} are the peak, total and mean flux (see Fig. 4 for an example of the parameter space plot). For each positive source $P_i = (F_{i,\max}, F_{i,\text{sum}}, F_{i,\text{mean}})$ a volume density of positive (n_{pos}) and negative (n_{neg}) sources, as determined on the collection of stacked spectra, is calculated through the Gaussian kernel density estimation (Hastie et al. 2001):

$$n = \frac{1}{h} \sum_i \mathcal{G}\left(\frac{x - x_i}{h}\right), \quad (2)$$

⁶ The re-gridding is done via the `reproject_interp` function from the `Astropy Reproject` package (Astropy Collaboration 2013, 2018, 2022). The interpolation is done using a nearest-neighbour method to preserve the underlying noise properties. Using different interpolation orders or `Reproject` functions leads to negligible improvement, but a significant cost in terms of computing time.

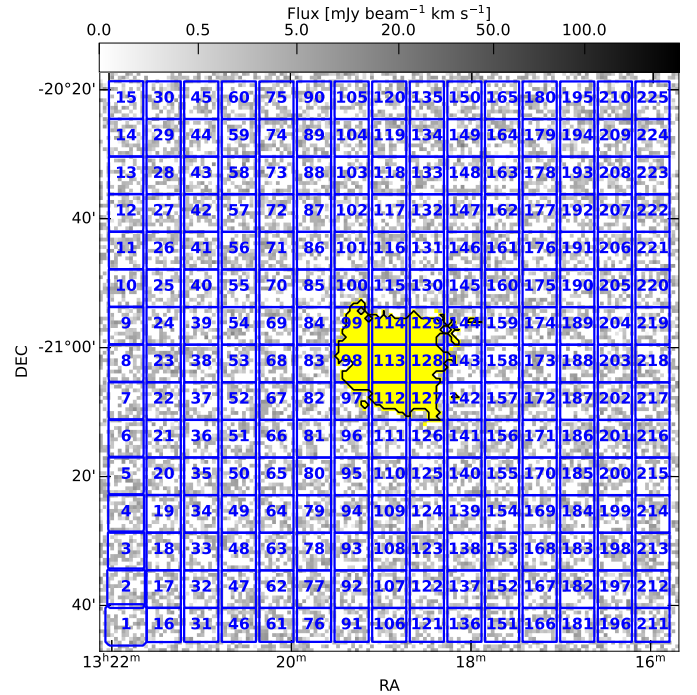


Fig. 3. Stacking regions for J1318-21 reported as blue squares. For the cell size displayed in the figure, STACKER will provide 225 stacked spectra, one for each region. The numbering helps for the bookkeeping. The background grey-scale image is the cube resulting from the preliminary manipulation steps (i.e. masking, shuffling and re-gridding) collapsed along the spectral axis and the black contours enclose the SoFIA-2 mask denote with a yellow colour.

where $x = (x_1, x_2, \dots, x_n)$ are the position of the sources, \mathcal{G} is a Gaussian kernel and h is a smoothing factor. The reliability (ϱ) of P_i is then given by

$$\varrho = \begin{cases} \frac{n_{\text{pos}} - n_{\text{neg}}}{n_{\text{pos}}}, & \text{for } n_{\text{pos}} \geq n_{\text{neg}} \\ 0, & \text{otherwise} \end{cases}. \quad (3)$$

Only sources for which ϱ is above a given threshold are retained in the final source catalogue.

3.3. Gaussianity tests

MeerKAT noise is known to be Gaussian to a high degree, as illustrated in Sect. 5.2. To quantify this further we incorporated two Gaussianity tests in FINDER. The first is the Anderson-Darling test (Stephens 1974, 1976, 1977, 1979), which evaluates the hypothesis that a given sample follows a specific distribution, in our case, a Gaussian distribution. This statistical test quantifies the disparity between the observed data and the expected values by computing the Anderson-Darling statistic A

$$A^2 = -n - \sum_{i=1}^n \frac{2i-1}{n} [\ln(F(x_i)) + \ln(1 - F(x_{n+1-i}))], \quad (4)$$

where n is the number of elements in the sample, F is the hypothesised cumulative distribution function (in our case, a Gaussian) and x_i is the i th-datapoint in progressive order ($x_1 < \dots < x_n$).

A lower Anderson-Darling statistic indicates a better fit to the Gaussian distribution, suggesting that the data can reasonably be assumed to follow it. Conversely, a higher statistic indicates a poorer fit, meaning non-Gaussian behaviour. A

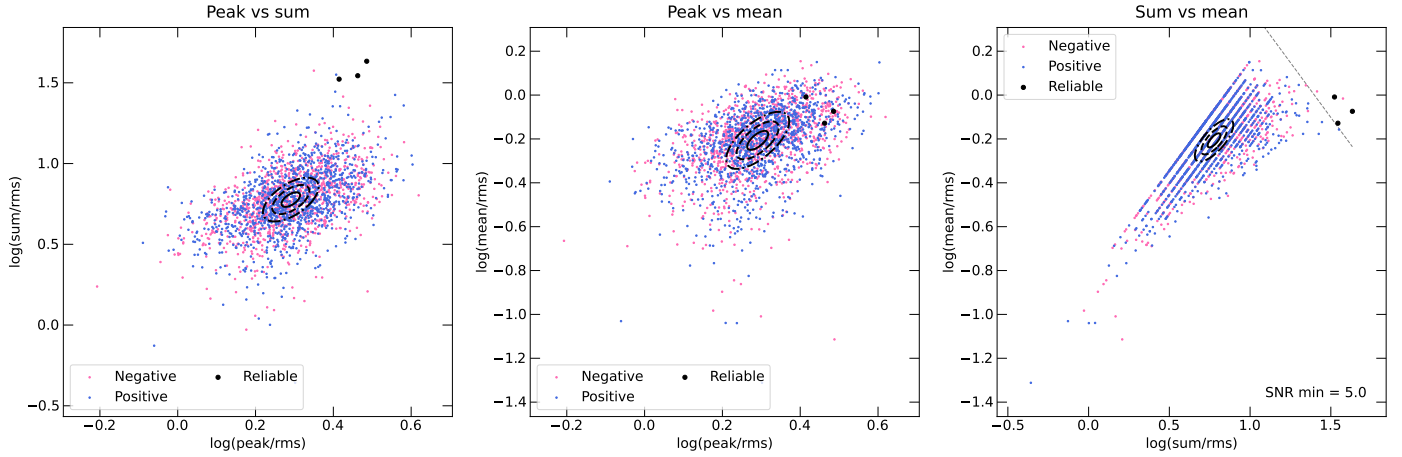


Fig. 4. Reliability plot of the sources detected in the J1318-21 full-depth cube (intermediate cell size). *Left panel:* $\frac{F_{\text{sum}}}{\sigma}$ vs. $\frac{F_{\text{peak}}}{\sigma}$ projection of the $(\frac{F_{\text{max}}}{\sigma}, \frac{F_{\text{sum}}}{\sigma}$ and $\frac{F_{\text{mean}}}{\sigma})$ parameter space. Blue, pink and black points are the positive, negative and reliable, detected in the 225 stacked spectra computed for the J1318-21 full-depth cube when using the intermediate cell size. The solid, dashed and dash-dotted black ellipses represent, respectively, the two-dimensional 1, 2 and 3 times the standard deviation of the source distribution. *Centre panel:* $\frac{F_{\text{mean}}}{\sigma}$ vs. $\frac{F_{\text{peak}}}{\sigma}$ projection. *Right panel:* $\frac{F_{\text{mean}}}{\sigma}$ vs. $\frac{F_{\text{sum}}}{\sigma}$ projection. The dashed dark-grey line is the minimum S/N of a source in order to be considered reliable.

p -value

$$p = \begin{cases} e^{1.2937-5.709A+0.0186A^2}, & \text{if } A \geq 0.6 \\ e^{0.9177-4.279A-1.38A^2}, & \text{if } 0.34 \leq A < 0.6 \\ e^{-8.318+42.796A-59.938A^2}, & \text{if } 0.2 \leq A < 0.34 \\ e^{13.436+101.14A-223.73A^2}, & \text{otherwise} \end{cases} \quad (5)$$

is associated with the outcome of the test. Since $p = 0.05$ is a “standard” value for rejecting the alternative hypothesis, in our case spectra for which $p < 0.05$ are considered to have non-Gaussian noise.

We complement the Anderson-Darling with another Gaussianity test which involves smoothing the spectrum using a boxcar filter with a width of 10 channels (N. Kanekar, priv. comm.). The σ before and after the smoothing is then compared, with the expectation that it should decrease by a factor of $\sqrt{10}$ if the noise adheres to a Gaussian distribution. Detections from spectra that do not satisfy either of these conditions are rejected from the final source catalogue.

4. Testing the method

Both STACKER and FINDER depend on a number of parameters which need to be optimised, mainly, the weighting scheme, the size of the square cells, the smooth-and-clip scales and the reliability threshold. We tested and optimised our procedure using simulated galaxies, which provide precise information about the priors, that is, location, column density, and kinematics of the gas. Therefore, we performed the stacking on the two TNG50 cubes described in Sect. 2.2.

The emission from the galaxy was blanked based on the SoFIA-2 mask and the cubes were shuffled using the noiseless moment 1 map, which also contains kinematic information on the low column density gas otherwise hidden in the noise. The latter choice is to maximise the S/N of each stacked spectrum, as we are assuming to know precisely the kinematics of the gas. In practice, these velocities are not precisely known, thus we test the effect of these uncertainties in Sect. 4.2.

We then re-grid the shuffled cube to the beam FWHM. As already mentioned in Sect. 3.1, this will make voxels (almost)

independent and will allow the noise to decrease with the square root of the number of co-added LoS.

4.1. Parameters optimisation

We first optimised the size of the stacking regions, aiming to balance the number of stacked spectra with the filling factor of a source. The noise (σ_s) in the stacked spectrum decreases as \sqrt{N} , where N is the number of co-added LoS. Increasing the cell size, that is, increasing N leads on the one hand to a lower σ_s , but on the other hand it might potentially lead to the inclusion of more noise-only spectra which wash out the signal. Conversely, reducing the cell size, that is, decreasing N , will increase σ_s .

Given that the noise decreases as \sqrt{N} , we set the smallest cell size to be 5×5 beams, which corresponds to an improvement in the limiting column density of up to a factor of 5. It is not worth opting for smaller cell sizes because of the otherwise negligible improvement in sensitivity. We then progressively increased the cell size until the number of detections is maximised. This results in choosing the largest size to be 14×14 beams, equivalent to an improvement in the limiting column density by a factor of 14. We also employed an intermediate size of 9×9 beams.

In a general cosmology with $H_0 = 69.6 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_m = 0.286$ and $\Omega_b = 0.714$ (Bennett et al. 2014), for a galaxy at 22.7 Mpc (the furthest in the MHONGOOSE sample) observed with the r15_t00 beam (see Sect. 2.1) the cell sizes correspond to physical scales of 52×52 , 34×34 and $19 \times 19 \text{ kpc}^2$, respectively, while for a galaxy at 3.3 Mpc (the closest in the MHONGOOSE sample) to 8×8 , 5×5 and $3 \times 3 \text{ kpc}^2$, respectively. These scales overlap with the predicted physical size of HI clouds in the CGM (Ramesh & Nelson 2024). Given the wide range in distances, keeping the cell size fixed to a set of physical scales is not feasible for our stacking experiment, thus, we fixed them instead to the aforementioned values of 14×14 , 9×9 and 5×5 beams.

We choose the weighting scheme (see Sect. 3.1) by comparing the outcomes when employing the equal- and the σ^2 -weight. The difference was negligible, both in terms of σ_s and source detection, so we ended up choosing the equal-weight for the reasons provided at the end of Sect. 3.1. Because the noise injected in the mock cubes is uniform by construction (see Sect. 2.2), it

Table 2. Source finder parameters used on the TNG50 cubes.

Parameters		Cell size		
		14 × 14 beams	9 × 9 beams	5 × 5 beams
Flux threshold	[σ]	2.5	2.5	2.5
Smoothing kernels	[channels]	1, 5, 10	1, 5, 11	1, 7, 12
Minimum linewidth	[channels]	5	5	7
Minimum source S/N		5	5	5
Reliability threshold		0.85	0.85	0.85

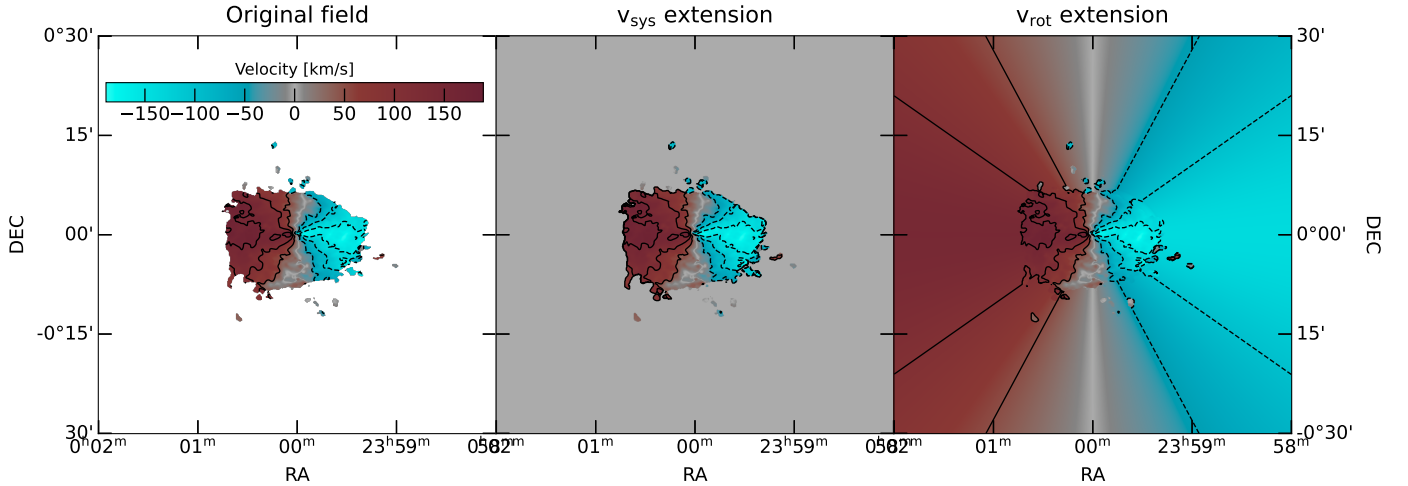


Fig. 5. Comparison between differently extended velocity fields. *Left panel:* SoFia-2 moment 1 map of the TNG50 galaxy 520855. Black solid contours denote receding (50, 100, 150) km s⁻¹ velocities with respect to the systemic velocity. Black dashed contours instead refer to approaching (-150, -100, -50) km s⁻¹ velocities. *Centre panel:* Velocity field extended with the systemic velocity. *Right panel:* Velocity field extended with the rotation velocity of the HI disk.

is not surprising that the two weighting schemes lead to comparable results.

Lastly, we optimised the FINDER specifications, namely, the σ -threshold, the smoothing kernels, the linker size and the reliability parameters. We found that imposing σ -threshold = 2.5 allows for the detection of a sufficient number of noise peaks to provide a robust reliability calculation. The optimal smoothing kernels, linker size and reliability parameters were selected by changing each of them until the difference between the observed and expected cumulative distribution of the Skellam variable

$$K = \frac{n_{\text{pos}} - n_{\text{neg}}}{\sqrt{n_{\text{pos}} + n_{\text{neg}}}} \quad (6)$$

is minimised. A thorough description of K is provided by Serra et al. (2012a). In Table 2 we list the values employed in FINDER to detect signal in the stacked spectra around the two simulated galaxies.

4.2. Assumption on the gas kinematics

So far we have aligned the spectra with the noiseless moment 1 map, that is, assuming to know precisely the kinematics of the gas. As this will not be the case when stacking the MHONGOOSE data, we proceeded by evaluating two possible assumptions about the gas kinematics that are used as a prior: systemic velocity, which is equivalent to no alignment, or flat rotation curve. The former represents the scenario where the gas outside the galaxy is assumed to be unaffected by the rotation of

the galaxy, the latter assumes that the CGM and IGM co-rotates with the HI disk.

Ideally, to test the second hypothesis one wants the rotation curve to be derived and correctly extrapolated to large radii, that is, beyond the “edge” of the disk as identified in the moment maps. This prior, namely the geometry of the gaseous disk as a function of radius, is generally unknown. Exploring the outcome of stacking when the spectral alignment is done with the geometry and the rotation velocity as a free parameter is computationally expensive and beyond the scope of this paper. It would, however, be a natural follow-up of this work (see Sect. 7).

To evaluate the two hypothesis we extrapolated to larger radii the moment 1 map of the two TNG50 galaxies as produced by SoFia-2. For the case where the gas is assumed to move at the systemic velocity, this is done by simply replacing any empty value in the map with the systemic velocity of the galaxy. For the assumption that the gas is co-rotating with the galaxy, the extension of the moment 1 map is less trivial. We retrieved the values of the moment 1 map along the major axis (v_{rot}) and, together with the azimuthal angle (θ) and inclination (i) of the galaxy, we computed the velocity field across the entire FoV:

$$v = v_{\text{sys}} + v_{\text{rot}} \cos \theta \sin i, \quad (7)$$

for which we provide in Fig. 5 the result for the galaxy 520885⁷.

We then calculated for each galaxy the residual maps $R = |\text{mom}_1 - v|$, where mom_1 is the noiseless moment 1 and v is the

⁷ The extension of the velocity field is performed using our implementation of the GIPSY task VELFI.

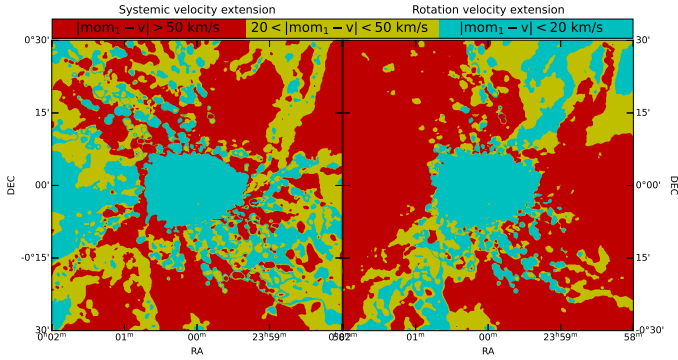


Fig. 6. Error in the spectral alignment for the TNG50 galaxy 520885 under two different CGM and IGM kinematics assumptions. *Left panel:* the residuals in terms of $|\text{mom}_1 - v|$, where mom_1 is the noiseless moment 1 map and v the assumed kinematics, when $v = v_{\text{sys}}$. Light-blue corresponds to $|\text{mom}_1 - v| < 20 \text{ km s}^{-1}$, yellow to $20 < |\text{mom}_1 - v| < 50 \text{ km s}^{-1}$ and red is $|\text{mom}_1 - v| > 50 \text{ km s}^{-1}$. *Right panel:* same as left panel but when v is given by assuming the gas is co-rotating with the galaxy with a flat rotation curve.

extended velocity field under one of the two aforementioned assumptions. The residual maps for the galaxy 520885 are provided in Fig. 6, while the residuals for the galaxy 555013 are given in Appendix B. Since R corresponds to the alignment error for each LoS, by looking at Fig. 6 one can see where each assumption will produce too large an error ($>50 \text{ km s}^{-1}$, Khandai et al. 2011; Maddox et al. 2013; Neumann et al. 2023) in the spectral alignment. In this case using the systemic velocity generally provides a better result.

It will be clear that one should be careful not to generalise this result. However, we can use the inclination value to our advantage when applying the stacking. If i is the inclination angle and v_{rot} the rotation velocity of the gas, the maximum velocities measured along each LoS are related approximately by Eq. (7). For a face-on galaxy ($i = 0^\circ$), $v \sim v_{\text{sys}}$ no matter the rotation velocity. If the gas accretion happens along the plane of the disk, assuming the HI clouds are moving at the systemic velocity is a reasonable choice. For an edge-on galaxy, v_{rot} is the dominant factor in Eq. (7), hence, any uncertainty in the estimate of the kinematics is likely to produce a larger error in the alignment. Again, a study of the most suitable assumption on the gas kinematics as a function of the geometrical properties of the galaxy is beyond the scope of this work and will be analysed in a forthcoming paper. In this paper we employ the systemic velocity as the assumed kinematics of the gas in the CGM and IGM for stacking the MHONGOOSE data and we apply our algorithm to the 18 galaxies with $i \leq 60^\circ$.

A last point to consider for the quality of the alignment is the coherency of the assumed kinematics with respect to the real one on the scales probed by the chosen cell sizes. For instance, if in a given cell size the assumed kinematics cover only a small velocity range, then even if the spectra in the cell are assigned erroneous velocity shifts, this will still result in a coherently aligned stacked profile, just with an extra velocity shift with respect to the reference velocity. However, measuring this coherency in the absence of the prior kinematics, as when we stack the MHONGOOSE cubes, is not possible.

4.3. The limitation of our method: The effect of the S/N cut

Figure 7 shows the results of the stacking experiment for TNG50 520855 galaxy where the alignment has been done using

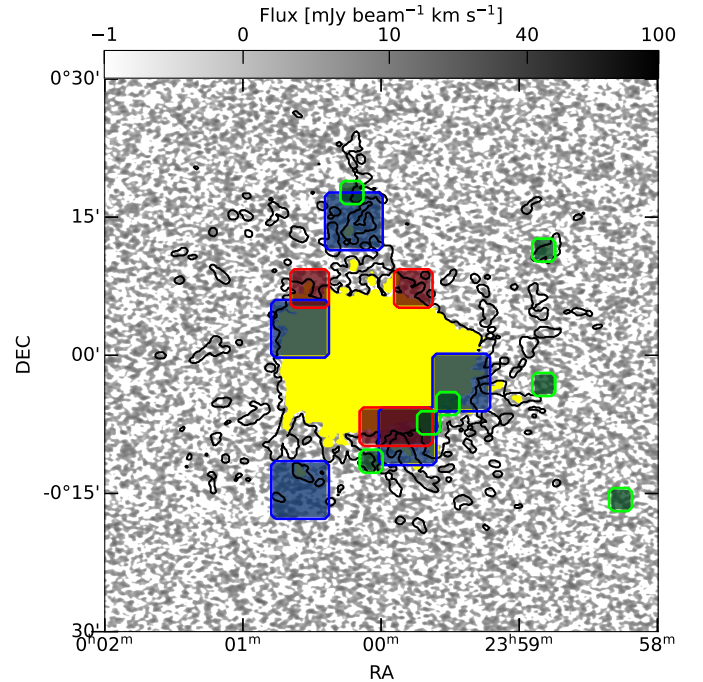


Fig. 7. Stacking detection map. The background grey-scale image is the cube of the TNG50 520855 galaxy collapsed along the spectral axis and blanked from the galaxy emission. Black contours denote the noiseless moment 0 map clipped at the column density value of $3.6 \times 10^{17} \text{ cm}^{-2}$. The coloured squares indicate the stacking regions where their stacked spectrum contains a reliable detection. Different colours represent different cell size: blue for 14×14 , red for 9×9 and green for 5×5 beams.

the systemic velocity. Regions where a reliable detection has been found are denoted with the blue, red and green boxes, corresponding to cell sizes of 14×14 , 9×9 and 5×5 beams, respectively. The boxes are overlaid with the background grey-scale image resulting from the collapse of the cube along the spectral axis. The galaxy emission is blanked according to the SoFIA-2 mask marked in yellow. Since this emission was blanked prior to the stacking, the sources in the coloured boxes represent signal below the SoFIA-2 detection threshold of 4σ .

The 2.5σ - 16 km s^{-1} column density limit that we were able to achieve with stacking is $1.3 \times 10^{17} \text{ cm}^{-2}$ for the largest cell size, $2.1 \times 10^{17} \text{ cm}^{-2}$ for the intermediate size and $3.6 \times 10^{17} \text{ cm}^{-2}$ for the smallest size. However, only $\sim 35\%$ of the emission above the $3.6 \times 10^{17} \text{ cm}^{-2}$ column density level, shown as black contours in Fig. 7, was detected. The reason why not all emission was detected is that stacking gives a weighted mean, hence, a line intrinsically bright enough to potentially fall above the detection threshold, will become too faint to be picked up if averaged with only noise. Furthermore, through the reliability procedure we introduce a filtering based on the S/N of a source: detections with S/N lower than a threshold are automatically rejected and do not enter the detection map (see the grey dashed line of the top-right panel in Fig. 4).

Therefore, we tested, given our set of filtering parameters, what our stacking algorithm should be able to pick up. We impose an integrated S/N cut of 5 (see Table 2), that with our spectral resolution of 1.4 km s^{-1} (see Sect. 2.2) corresponds to a per-channel S/N of 1.5 for a linewidth of 16 km s^{-1} .

We proceeded to quantify the amount of gas that will fall below our chosen S/N threshold, as illustrated in Fig. 8. The integrated S/N for a source spanning n channels in a stacked

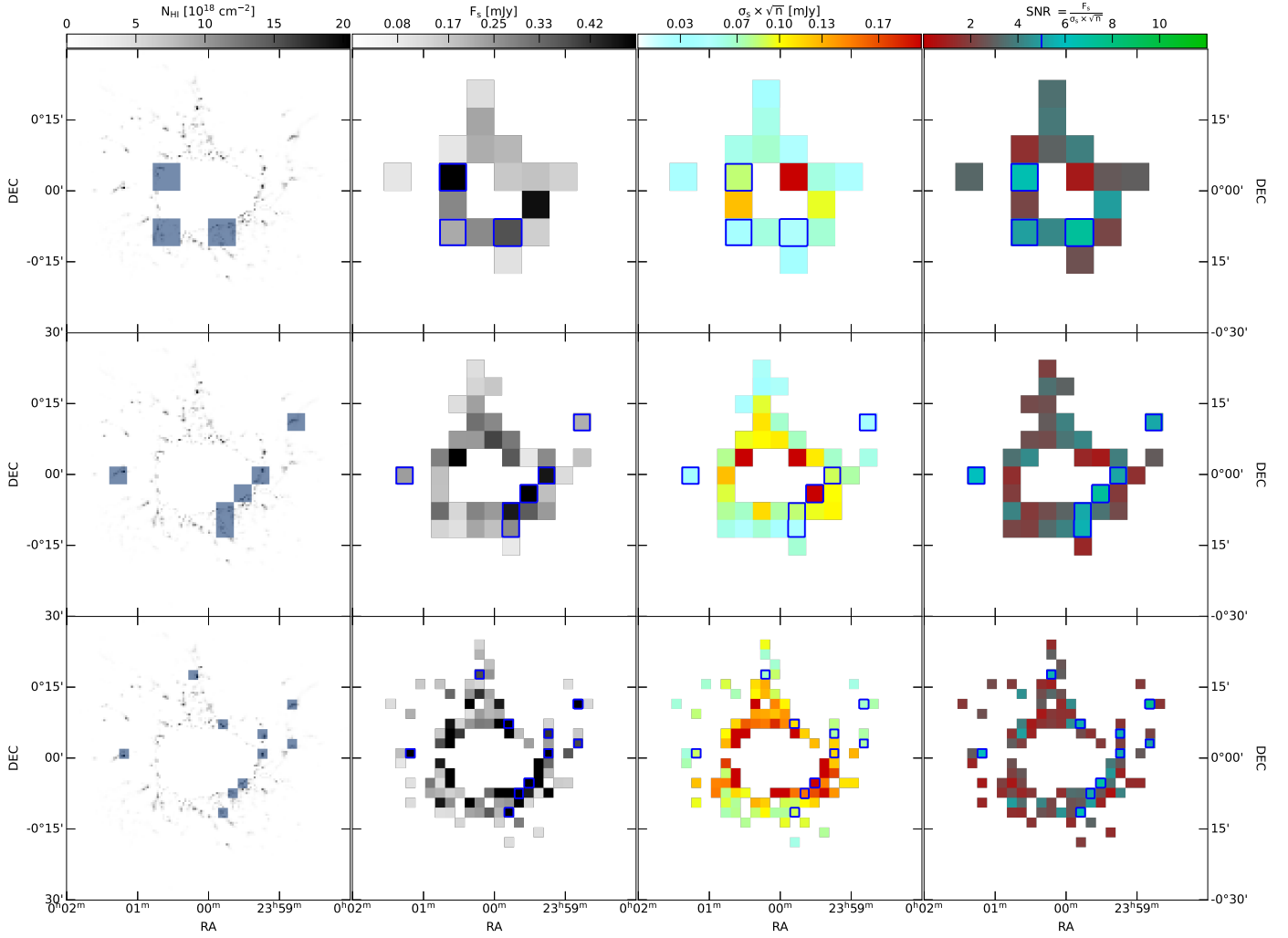


Fig. 8. Stacking S/N maps for the TNG50 galaxy 520885. *First column:* stacking regions where $S/N > 5$ (blue squares) overlaid with the grey-scale noiseless moment 0 map, blanked from the emission already detected by SoFIA-2. Top to bottom is for cell sizes of 14×14 , 9×9 and 5×5 beams, respectively. *Second column:* the integrated flux of the signal for stacking regions where the HI column density is $> 3.6 \times 10^{17} \text{ cm}^{-2}$. The blue contours enclose the cells where $S/N > 5$. *Third column:* the noise of the stacked spectrum multiplied by the square root of the number of channels covered by the source. *Fourth column:* integrated S/N of the signal in the stacked spectrum for the regions. The $S/N = 5$ level is indicated also on the colourbar.

spectrum having noise σ_s is defined as:

$$S/N = \frac{\sum F_i}{\sigma_s \sqrt{n}}, \quad (8)$$

where F_i is the flux in the i -th channel covered by the source. We computed Eq. (8) for each stacking region by running STACKER on the noiseless cube. We retrieved first the dispersion of the line in every stacked spectrum, which gives us n , and then by integrating each noiseless stacked spectrum within a window of size n around its peak we calculated $\sum F_i$. Finally, σ_s is retrieved from the “noisy” stacked spectrum. The result is a map with the S/N for each stacking region.

Clipping this map at $S/N > 5$ provides the regions where the signal in their stacked spectrum is bright enough to be detected given the S/N cut we have imposed. This is illustrated in Fig. 8, where we show the outcome for the TNG50 galaxy 520885 (the result is analogous for the other mock galaxy as shown in Appendix B). The blue squares are the regions where $S/N > 5$.

Consequently, the reason why in Fig. 7 not all the emission above the $3.6 \times 10^{17} \text{ cm}^{-2}$ level (black contours) has been detected with stacking is the following: if the mean flux within

each region is not enough to survive the imposed S/N cut, then that flux will not be detected, no matter how it is distributed within the region.

5. Applying the method

We proceeded with the application of STACKER and FINDER to the eighteen MHONGOOSE galaxies with $i \leq 60^\circ$ listed in Table A.1. For each galaxy, we have a “single-track” cube obtained from the data reduction of a single 5.5 h track, and for ten of them we also have a “full-depth” cube obtained by combining all 10 observing sessions.

We first applied our stacking on the single-track dataset. The reason was to test our algorithm in an “uncontrolled” environment, which means that, unlike when working with simulated data, we are now truly unconstrained, having no information about the priors. Any detection present in the stacked spectra of the single-track cubes was then checked against the corresponding full-depth dataset (when available) to provide an indication on the reliability of each detection and the robustness of any source that will be found in the data. Indeed, any source

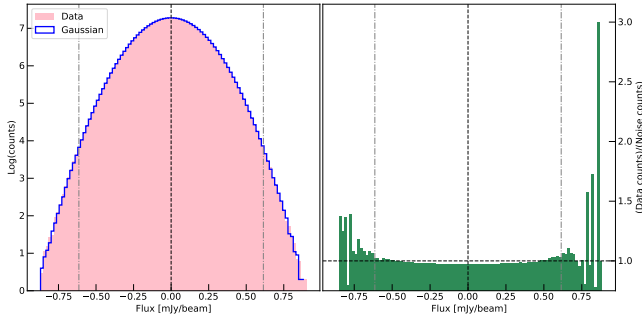


Fig. 9. Comparison between the MeerKAT and Gaussian noise distribution. *Left panel:* the pink linear-log histogram represent the number of voxels with a given flux value in a MeerKAT cube, known to contain no bright galaxies. The blue linear-log histogram is instead computed from a Gaussian noise cube. The black dashed vertical line is the reference 0-flux level, whereas the two grey dash-dotted lines are denoting the $\pm 4\sigma$ level. *Right panel:* ratio between the pink and blue histograms of the left panel. If the MeerKAT noise is truly Gaussian, the ratio should be ~ 1 for every flux bin. The black dashed horizontal line marks the 1-to-1 ratio.

which is present in both the single-track and full-depth stacked spectrum of a given region is likely genuine, while any excess of emission picked up in the single-track cubes but not in the full-depth might provide insight on how to further discriminate plausible emission line from artefacts.

5.1. Stacking the MHONGOOSE cubes

As was done for the TNG50 galaxies, we used the SoFiA-2 moment map to blank the already detected emission of the target galaxies and their satellites. As discussed in Sect. 4.2, we employed the systemic velocity as the reference redshift. Again, the re-gridding has been chosen so that 1 pixel equals 1 beam and the weighting scheme was set to be the equal-weight. This procedure was repeated in the same way for both the single-track and the full-depth data.

We opt for a common set of parameters for the line finder. To determine this common set we optimised the FINDER parameters for each galaxy following the procedure presented at the end of Sect. 4.1, that is by minimising the difference between the observed and expected cumulative distribution of the variable defined by Eq. (6), and then we took the mean value of each parameter⁸. The full list of parameters used for the source finding are given in Table 3.

5.2. The Gaussianity of MHONGOOSE noise

The reliability of any detection extracted by our algorithm is meaningful only if the MHONGOOSE noise distribution is Gaussian. We therefore created a “MHONGOOSE noise cube” by reimagining a full-depth cube in a $\sim 1000 \text{ km s}^{-1}$ spectral range away from the target galaxy and free of known emission. The noise level in this cube is $\sigma = 0.15 \text{ mJy beam}^{-1}$. To compare with, we also created a cube with the same size but containing values extracted from a Gaussian noise distribution with dispersion $\sigma_0 = 0.15 \text{ mJy beam}^{-1}$.

In the left-side panel of Fig. 9 we plotted the histogram of the flux values (F) in the MHONGOOSE noise cube (pink) and compare this with the Gaussian histogram (in blue). Taking the

⁸ Optimising the values for each galaxy led to negligible differences in the number of detected lines.

ratio, we directly measured the presence of any non-Gaussianity. In the right-side panel of Fig. 9 it is shown that the ratio diverges from 1 beyond a level of $\sim 4\sigma$ (grey dash-dotted vertical lines).

For a Gaussian distribution, 0.000063% of the data have $|F| > 4\sigma$, which corresponds to 25606 voxels given the size of the cube. In the MHONGOOSE noise cube we found that instead 0.000069% of the voxels (27864) have $|F| > 4\sigma$. This is equivalent to an excess of 8.8%.

Of these 27864 voxels having $|F| > 4\sigma$, 14053 are positive (P) and 13811 are negative (N). The almost symmetric distribution ($\frac{P}{N} = 1.02$) suggests that this excess is due to low-level artefacts such as residual continuum emission or RFI. A further confirmation of the above is provided by checking the spatial and spectral location of the voxels with $|F| > 4\sigma$. We found that they are spatially randomly distributed and spectrally non-contiguous.

We also repeated the aforementioned analysis for the Stokes- Q ⁹ component of the MHONGOOSE noise cube, as well as for the single-track and full-depth cubes used for stacking, finding always similar results. Thus, we concluded that the MHONGOOSE noise distribution is mostly Gaussian, with a small excess at the highest absolute flux values. The symmetry of this excess argues that it is inherent to the MHONGOOSE noise distribution, caused by low-level artefacts, rather than real sources.

5.3. Detections in the single-track data

The stacking on the single-track data cubes of the MHONGOOSE sample results in 25 reliable sources out of 82479 detections, for which the stacked spectra are available at Zenodo and the main properties are listed in Table A.2. Three of the authors independently visually inspected the single-track cubes at the spatial and spectral location of each detection and all agreed with the following classifications:

- eleven sources were not visually identified. For five we also have the corresponding full-depth data and there was no counterpart there as well. Therefore, these five are likely noise peaks;
- ten sources were identified in their single-track cube. However, we were not able to check them also in their full depth cube because it was not available, as the survey is still to be completed at the time of writing;
- three sources have been associated with faint HI emission from the outer disk of the target galaxies or their companions;
- one source is galaxy MKTJ125225.4-124304.1, which is listed in de Blok et al. (2024). In that paper the galaxy is listed with uncertain parameters, as not the entire velocity range was imaged. Also in the stacked spectrum the velocity range is incomplete, however, based on the full-depth line profile the systemic velocity of the galaxy can now be updated to 1279 km s^{-1} .

These last four sources were detected by SoFiA-2 in the corresponding full-depth cube. This detection of HI in emission via an unconstrained spectral stacking thus confirms the capability of our procedure.

We plot the peak, total and mean flux of the 25 sources as a function of their line width and of the cell sizes on the background of Fig. 11. We found no preferred size for the stacking regions and no clear threshold in terms of flux properties or line width to discriminate between genuine detections and false

⁹ Stokes- Q is given by the difference between the two linear polarisations, thus, the signal cancels out and only the noise remains.

Table 3. Source finder parameters for each cell size and dataset.

	14 × 14 beams		9 × 9 beams		5 × 5 beams	
	Single-track	Deep	Single-track	Deep	Single-track	Deep
Flux threshold [σ]	2.5	2.5	2.5	2.5	2.5	2.5
Smoothing kernels [channels]	1, 5, 9	1, 5, 10	1, 7, 10	1, 7, 10	1, 7, 12	1, 7, 12
Minimum linewidth [channels]	5	5	7	7	7	7
Minimum S/N	5	5	5	5	5	5
Reliability threshold	0.85	0.85	0.85	0.85	0.85	0.85

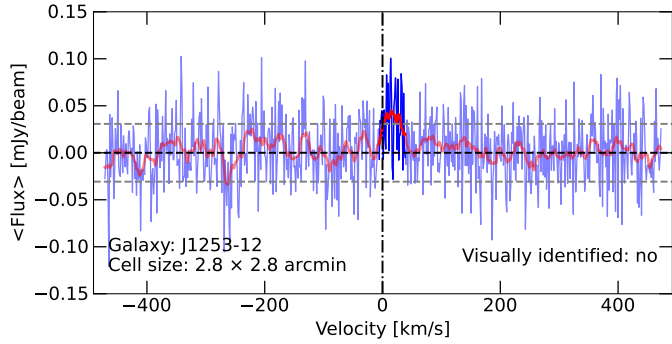


Fig. 10. Reliable source around the galaxy J1253-12. The stacked spectrum and its 9-channel boxcar smoothed version are given in blue and red, respectively. The detected source is highlighted. The horizontal grey dashed lines are the $\pm\sigma$ level for the unsmoothed spectrum, while the black dashed line is the 0-flux level. The vertical black dashed-dotted line is, instead, the 0 km s^{-1} reference velocity. The galaxy name and the cell size are provided in the bottom-left corner, while at the bottom-right is reported if the detection was also visually identified in its corresponding cube.

positives, although low flux detections are generally not visible by eye in the cube.

5.4. Detections in the full-depth data

The stacking on the full-depth data leads to a total of thirteen reliable sources out of 34839 detections corresponding to a detection rate of ~ 2.1 sources per cube, slightly higher than for the single-track data (~ 1.8). The full source list is provided in Table A.3 and a representative stacked spectrum for a reliable detection is given in Fig. 10. The full sample of reliable detections is available online at [Zenodo](#). By visually inspecting the cubes we found that four sources are identifiable by eye, while nine do not have a clear counterpart in the cube. These numbers are equivalent to $\sim 15\%$ of the reliable detections being visible by eye, $\sim 15\%$ not unambiguously distinguishable and $\sim 70\%$ not visible. For the single-track data, these percentages are $\sim 20\%$, $\sim 36\%$ and $\sim 44\%$, respectively.

The larger number of non-visible detections is due to the flux properties of the sources found in the stacked spectra of the full-depth cubes. The comparison with single-track detections made in Fig. 11 shows that the full-depth detections lie in the region of parameter space where also for the single-track observations we have not found the corresponding source in the HI cube. Given the analysis of the MeerKAT noise presented in Sect. 5.2, we suspect the non-identified sources are not low-level artefacts still present in the cubes but rather genuine HI excess, although only through a direct detection via deeper observations we will be able to confirm this.

5.5. Explaining the missing gas

In Sect. 4.1 we have presented the application of our algorithm to two TNG50 mock galaxies detecting emission outside the SoFiA-2 mask. However, we find very few plausible HI detections when stacking the MHONGOOSE observations. We suspect that the reason why little excess of HI is found in the full-depth cubes is that the radial profile of the HI disks exhibits a sharp decrease in the column density, because the gas likely becomes ionised, thus requiring a detection limit well below 10^{17} cm^{-2} in order to be detected. This has been suggested previously to occur at higher column densities as well (Maloney 1993; Dove & Shull 1994; Ianjamasimanana et al. 2018), and a radial profile study of 18 MHONGOOSE galaxies made with the Green Bank Telescope (GBT) seem to support the view of this decrease happening at the $\sim 10^{17} \text{ cm}^{-2}$ column density level (Sardone et al. 2021).

Recent observations with the GBT of NGC 891 and NGC 4565 did also reveal that the HI column density in the CGM and IGM is $< 10^{17} \text{ cm}^{-2}$ (Das et al. 2024). Similar results were achieved with the Five-Hundred-meter Aperture Spherical radio Telescope by Xu et al. (2022), observing of Stephan’s Quintet, and Liu et al. (2023) pointing at NGC 4490/85. It is important to note that, as discussed in Sect. 1, single-dish observations are limited by their low spatial resolution ($> 200''$). Reaching the $< 10^{17} \text{ cm}^{-2}$ sensitivity at a higher angular resolution with interferometers will require significant observing time even with the next generation of interferometers (such as SKA-MID, see Fig. 2 of de Blok et al. 2024).

6. Conclusion

We present a method to detect emission from the HI in the CGM and IGM around resolved nearby galaxies with no prior information about the spatial and spectral distribution of the gas. The method consists of a standard stacking procedure, described in Sect. 3.1, combined with a one-dimensional line finder that uses the reliability calculation presented in Serra et al. (2012a) to discriminate noise peaks from potentially genuine emission (see Sect. 3.2).

We applied our algorithm to two TNG50 mock galaxies taken from the Ramesh et al. (2023) sample and optimised its free parameters, mainly, the size of the stacking regions, the weighting scheme, and the source finder parameters (Sect. 4.1). As simulated data provide us with exact knowledge of the gas kinematics, we tested whether using the systemic velocity or the rotation velocity to align the spectra provides the best result (Sect. 4.2) and find that for averaged-inclined galaxies using the systemic velocity is preferable. The TNG50 data have also been used to study the limitations of the stacking, as discussed in Sect. 4.3, finding that imposing a minimum S/N in the stacked spectrum inevitably leads to the loss of the majority of the

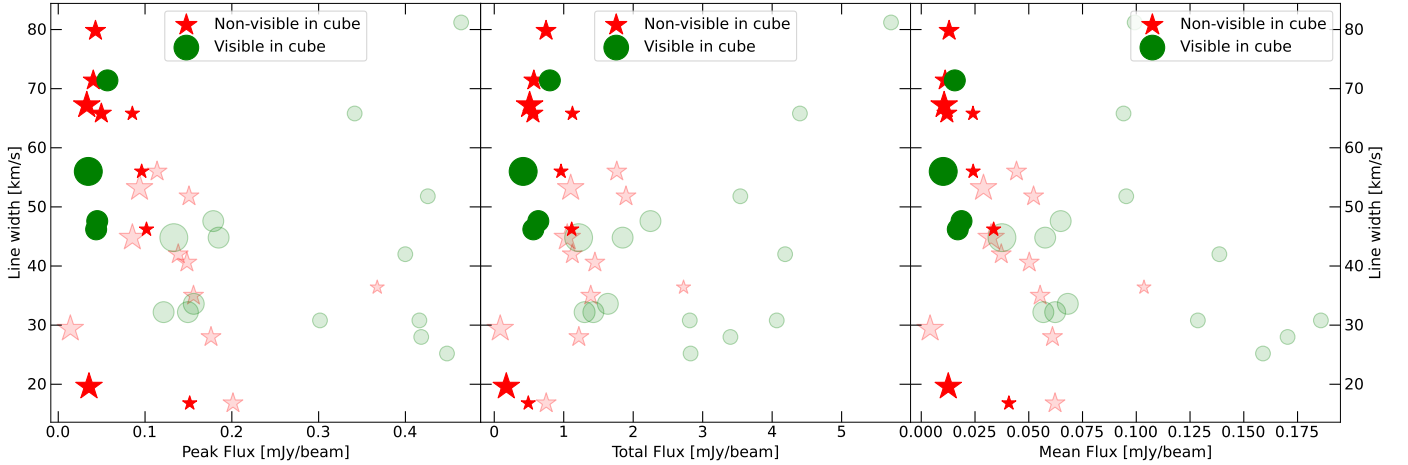


Fig. 11. Peak flux (left panel), total flux (central panel), and mean flux (right panel) of the stacking detections as a function of their width. Detections in single-track stacked spectra are provided in the background, while detections in full-depth stacked spectra are given in the foreground. Detections with no counterpart visible in the cube are labelled with red stars. Green circles refer to visually identified HI emission. The size of the markers is indicative of the size of the square regions used for the stacking.

detections whose peak flux falls above the detection limit (see Fig. 8). However, lowering the S/N threshold to include more signals is not a viable option, as this would also increase the number of false detections. TNG50 data are, by construction, filled with Gaussian noise. In Sect. 5.2 we explain how we quantified the level of Gaussianity of the MHONGOOSE noise, finding a $\sim 10\%$ excess of voxels with an absolute flux $>4\sigma$ compared to a Gaussian.

When stacking the MHONGOOSE data (Sect. 5.4), we found very little excess HI emission. As discussed in Sect. 5.5, assuming our stacking method works, the lack of HI at the $\sim 10^{17} \text{ cm}^{-2}$ level beyond the periphery of the HI disks may be the result of a sudden drop in the HI column density of the gaseous disk, as radial profile studies using GBT observations suggest (Sardone et al. 2021). If validated by future analysis of the HI radial profiles of the MHONGOOSE galaxies with the higher resolution MeerKAT data, this will confirm the need for further comparisons with simulations where extended HI distributions are a more general feature. Stacking also has its limitations. Increasing the number of stacked spectra does not necessarily mean that the S/N of the emission we are trying to detect will increase. However, if our conclusions hold, then longer integration times can help with the direct detection of the HI. One would need many tens of hours even on SKA-MID to reach the necessary column density limit and be able to resolve features even at an arc-minute resolution.

7. Future prospects

This work represents a first step in the application of unconstrained spectral stacking to detect HI in emission around nearby resolved objects. There is still space for further improvements. For example, future studies can focus on designing a more sophisticated way to minimise the discrepancy between the assumed kinematics and the real velocity of the gas. This can be done, for instance, by repeating the analysis presented in Sect. 4.2 on a much larger sample of simulated galaxies and study which is the best assumption as a function of the inclination angle. Alternatively, one can derive the rotation curve and model the velocity field leaving the kinematics and geometry of the outer disk as free parameters, to check which configuration maximises the number of detections. The work by

Das et al. (2024) is a step in this direction, as they stacked along the major and minor axis where it is simpler to determine the kinematics of the gas, but ultimately one wants to detect the HI all around the galaxies. Another limitation of our method is the use of regions with fixed shape and size. Implementing a dynamical determination of the stacking regions where the shape and size of the stacking regions is continuously varied until a line is detected and its S/N maximised, like employing Voronoi tessellation (Cappellari & Copin 2003; Cappellari 2009), may lead to even more reliable identification of any excess of HI embedded into the noise.

Another way forward to understand the properties of the HI in the CGM of nearby isolated star-forming galaxies may consist in the application of this stacking procedure on the whole MHONGOOSE sample normalised in size and redshift. This experiment will resemble the “typical” stacking analysis carried out to study the mean properties of sample of galaxies (e.g. Fabello et al. 2011; Delhaize et al. 2013; Geréb et al. 2014, 2015a,b; Kanekar et al. 2016; Maccagni et al. 2017; Ianjamasimanana et al. 2018; Healy et al. 2019, 2021; Chowdhury et al. 2022; Amiri et al. 2023; Bera et al. 2023). As mentioned in the previous section, it is imperative to do a detailed radial profile study of the HI in the MHONGOOSE galaxies which will improve the results of Sardone et al. (2021), further constraining the shape and extension of the gaseous disk. Finally, it is also worth noting that as the limit on the HI column density outside the gaseous disk keeps getting lower, it will be possible to better distinguish between various modes of accretion (for example, cold vs. hot accretion). Further work on the MHONGOOSE data, but also future deep surveys should be able to provide more observational constraints on these scenarios.

Acknowledgements. We thank the anonymous referee for the constructive comments. This work has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 882793 “MeerGas”). SV thanks N. Kanekar for the insightful discussion about the stacking technique. KS acknowledges support from the Natural Sciences and Engineering Research Council of Canada (NSERC). NZ and DJP gratefully acknowledge support by the South African Research Chairs Initiative of the Department of Science and Technology and the National Research Foundation. PK is partially supported by the BMBF project 05A23PC1 for D-MeerKAT. L.C. acknowledges financial support from the Chilean Agencia Nacional de Investigación y Desarrollo (ANID) through

Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT) Regular Project 1210992.

References

- Afruni, A., Fraternali, F., & Pezzulli, G. 2021, *MNRAS*, **501**, 5575
- Afruni, A., Pezzulli, G., Fraternali, F., & Grønnow, A. 2023, *MNRAS*, **524**, 2351
- Amiri, M., Bandura, K., Chen, T., et al. 2023, *ApJ*, **947**, 16
- Astropy Collaboration (Robitaille, T. P., et al.) 2013, *A&A*, **558**, A33
- Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, *AJ*, **156**, 123
- Astropy Collaboration (Price-Whelan, A. M., et al.) 2022, *ApJ*, **935**, 167
- Barnes, D. G., Staveley-Smith, L., de Blok, W. J. G., et al. 2001, *MNRAS*, **322**, 486
- Bennett, C. L., Larson, D., Weiland, J. L., & Hinshaw, G. 2014, *ApJ*, **794**, 135
- Bera, A., Kanekar, N., Chengalur, J. N., & Bagla, J. S. 2023, *ApJ*, **950**, L18
- Braun, R., & Thilker, D. A. 2004, *A&A*, **417**, 421
- Cappellari, M. 2009, ArXiv e-prints [arXiv:0912.1303]
- Cappellari, M., & Copin, Y. 2003, *MNRAS*, **342**, 345
- Chen, Q., Meyer, M., Popping, A., et al. 2021, *MNRAS*, **508**, 2758
- Chowdhury, A., Kanekar, N., & Chengalur, J. N. 2022, *ApJ*, **937**, 103
- Danovich, M., Dekel, A., Hahn, O., Ceverino, D., & Primack, J. 2015, *MNRAS*, **449**, 2087
- Das, S., Sardone, A., Leroy, A. K., et al. 2020, *ApJ*, **898**, 15
- Das, S., Rickel, M., Leroy, A., et al. 2024, *MNRAS*, **527**, 10358
- Davé, R., Rafieferantsoa, M. H., Thompson, R. J., & Hopkins, P. F. 2017, *MNRAS*, **467**, 115
- de Blok, W. J. G., Keating, K. M., Pisano, D. J., et al. 2014, *A&A*, **569**, A68
- de Blok, W. J. G., Athanassoula, E., Bosma, A., et al. 2020, *A&A*, **643**, A147
- de Blok, W. J. G., Healy, J., Maccagni, F. M., et al. 2024, *A&A*, **688**, A109
- Dekel, A., & Birnboim, Y. 2006, *MNRAS*, **368**, 2
- Dekel, A., Birnboim, Y., Engel, G., et al. 2009, *Nature*, **457**, 451
- Delhaize, J., Meyer, M. J., Staveley-Smith, L., & Boyle, B. J. 2013, *MNRAS*, **433**, 1398
- Dolag, K., Borgani, S., Murante, G., & Springel, V. 2009, *MNRAS*, **399**, 497
- Dove, J. B., & Shull, J. M. 1994, *ApJ*, **423**, 196
- English, J., Richardson, M. L. A., Ferrand, G., & Deg, N. 2024, *Astrophysics Source Code Library* [record ascl:2401.005]
- Fabello, S., Catinella, B., Giovanelli, R., et al. 2011, *MNRAS*, **411**, 993
- Faucher-Giguère, C.-A., & Oh, S. P. 2023, *ARA&A*, **61**, 131
- Fraternali, F., Oosterloo, T., Sancisi, R., & van Moorsel, G. 2001, *ApJ*, **562**, L47
- Geréb, K., Morganti, R., & Oosterloo, T. A. 2014, *A&A*, **569**, A35
- Geréb, K., Maccagni, F. M., Morganti, R., & Oosterloo, T. A. 2015a, *A&A*, **575**, A44
- Geréb, K., Morganti, R., Oosterloo, T. A., Hoppmann, L., & Staveley-Smith, L. 2015b, *A&A*, **580**, A43
- Giovanelli, R., Haynes, M. P., Kent, B. R., et al. 2007, *AJ*, **133**, 2569
- Hastie, T., Friedman, J., & Tibshirani, R. 2001, *Kernel Methods* (New York, NY: Springer), 165
- Heald, G., Józsa, G., Serra, P., et al. 2011, *A&A*, **526**, A118
- Healy, J., Blyth, S. L., Elson, E., et al. 2019, *MNRAS*, **487**, 4901
- Healy, J., Blyth, S. L., Verheijen, M. A. W., et al. 2021, *A&A*, **650**, A76
- Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, *MNRAS*, **480**, 800
- Ianjamasimanana, R., Walter, F., de Blok, W. J. G., Heald, G. H., & Brinks, E. 2018, *AJ*, **155**, 233
- Jonas, J., & MeerKAT Team. 2016, *MeerKAT Science: On the Pathway to the SKA*, 1
- Józsa, G. I. G., White, S. V., Thorat, K., et al. 2020, *Astrophysics Source Code Library* [record ascl:2006.014]
- Kamphuis, P., Jütte, E., Heald, G. H., et al. 2022, *A&A*, **668**, A182
- Kanekar, N., Sethi, S., & Dwarakanath, K. S. 2016, *ApJ*, **818**, L28
- Kerp, J., Kalberla, P. M. W., Ben Bekhti, N., et al. 2016, *A&A*, **589**, A120
- Khandai, N., Sethi, S. K., Di Matteo, T., et al. 2011, *MNRAS*, **415**, 2580
- Lehner, N., Howk, J. C., & Wakker, B. P. 2015, *ApJ*, **804**, 79
- Li, A., Marasco, A., Fraternali, F., Trager, S., & Verheijen, M. A. W. 2021, *MNRAS*, **504**, 3013
- Liu, Y., Zhu, M., Yu, H., et al. 2023, *MNRAS*, **523**, 3905
- Lochhaas, C., Tumlinson, J., Peebles, M. S., et al. 2023, *ApJ*, **948**, 43
- Maccagni, F. M., Morganti, R., Oosterloo, T. A., Geréb, K., & Maddox, N. 2017, *A&A*, **604**, A43
- Madau, P., & Dickinson, M. 2014, *ARA&A*, **52**, 415
- Maddox, N., Hess, K. M., Blyth, S. L., & Jarvis, M. J. 2013, *MNRAS*, **433**, 2613
- Maloney, P. 1993, *ApJ*, **414**, 41
- Marasco, A., Fraternali, F., Lehner, N., & Howk, J. C. 2022, *MNRAS*, **515**, 4176
- Nelson, D., Springel, V., Pillepich, A., et al. 2019, *Comput. Astrophys. Cosmol.*, **6**, 2
- Neumann, L., den Brok, J. S., Bigiel, F., et al. 2023, *A&A*, **675**, A104
- Offringa, A. R., & Smirnov, O. 2017, *MNRAS*, **471**, 301
- Offringa, A. R., McKinley, B., Hurley-Walker, N., et al. 2014, *MNRAS*, **444**, 606
- Pillepich, A., Springel, V., Nelson, D., et al. 2018, *MNRAS*, **473**, 4077
- Popping, A., Davé, R., Braun, R., & Oppenheimer, B. D. 2009, *A&A*, **504**, 15
- Porter, L. E., Orr, M. E., Burkhardt, B., et al. 2022, *MNRAS*, **515**, 3555
- Putman, M. E., Peek, J. E. G., & Jounq, M. R. 2012, *ARA&A*, **50**, 491
- Ramesh, R., & Nelson, D. 2024, *MNRAS*, **528**, 3320
- Ramesh, R., Nelson, D., & Pillepich, A. 2023, *MNRAS*, **518**, 5754
- Rubin, K. H. R., Hennawi, J. F., Prochaska, J. X., et al. 2015, *ApJ*, **808**, 38
- Sancisi, R., Fraternali, F., Oosterloo, T., & van der Hulst, T. 2008, *A&A Rev.*, **15**, 189
- Sardone, A., Pisano, D. J., Pingel, N. M., et al. 2021, *ApJ*, **910**, 69
- Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, *MNRAS*, **446**, 521
- Serra, P., Jurek, R., & Flöer, L. 2012a, *PASA*, **29**, 296
- Serra, P., Oosterloo, T., Morganti, R., et al. 2012b, *MNRAS*, **422**, 1835
- Serra, P., Westmeier, T., Giese, N., et al. 2015, *MNRAS*, **448**, 1922
- Somerville, R. S., & Davé, R. 2015, *ARA&A*, **53**, 51
- Springel, V. 2010, *MNRAS*, **401**, 791
- Stephens, M. A. 1974, *J. Am. Stat. Assoc.*, **69**, 730
- Stephens, M. A. 1976, *Ann. Stat.*, **4**, 357
- Stephens, M. A. 1977, *Biometrika*, **64**, 583
- Stephens, M. A. 1979, *Biometrika*, **66**, 591
- Tchernyshyov, K., Werk, J. K., Wilde, M. C., et al. 2023, *ApJ*, **949**, 41
- Thilker, D. A., Braun, R., Walterbos, R. A. M., et al. 2004, *ApJ*, **601**, L39
- Tumlinson, J., Thom, C., Werk, J. K., et al. 2013, *ApJ*, **777**, 59
- Tumlinson, J., Peebles, M. S., & Werk, J. K. 2017, *ARA&A*, **55**, 389
- Turner, M. L., Schaye, J., Steidel, C. C., Rudie, G. C., & Strom, A. L. 2014, *MNRAS*, **445**, 794
- van de Voort, F., Springel, V., Mandelker, N., van den Bosch, F. C., & Pakmor, R. 2019, *MNRAS*, **482**, L85
- van der Hulst, J. M., Terlouw, J. P., Begeman, K. G., Zwitter, W., & Roelfsema, P. R. 1992, in *Astronomical Data Analysis Software and Systems I*, eds. D. M. Worrall, C. Biemesderfer, & J. Barnes, *ASP Conf. Ser.*, **25**, 131
- Veronese, S., de Blok, W. J. G., & Walter, F. 2023, *A&A*, **672**, A55
- Vogelaar, M. G. R., & Terlouw, J. P. 2001, in *Astronomical Data Analysis Software and Systems X*, eds. J. Harnden, F. A. Primini, & H. E. Payne, *ASP Conf. Ser.*, **238**, 358
- Walter, F., Brinks, E., de Blok, W. J. G., et al. 2008, *AJ*, **136**, 2563
- Weinberger, R., Springel, V., Hernquist, L., et al. 2017, *MNRAS*, **465**, 3291
- Werk, J. K., Prochaska, J. X., Tumlinson, J., et al. 2014, *ApJ*, **792**, 8
- Westmeier, T., Braun, R., Brüns, C., Kerp, J., & Thilker, D. A. 2007, *New A Rev.*, **51**, 108
- Westmeier, T., Kitaëff, S., Pallot, D., et al. 2021, *MNRAS*, **506**, 3962
- Wolfe, S. A., Pisano, D. J., Lockman, F. J., McGaugh, S. S., & Shaya, E. J. 2013, *Nature*, **497**, 224
- Wolfe, S. A., Lockman, F. J., & Pisano, D. J. 2016, *ApJ*, **816**, 81
- Xu, C. K., Cheng, C., Appleton, P. N., et al. 2022, *Nature*, **610**, 461

Appendix A: Stacking detections

Table A.1. MHONGOOSE data cubes used and detections found in the stacked spectra.

Galaxy	Single-track $r = 1.5, t = 0''$	Detected sources Reliable / Non-reliable	Deep $r = 1.5, t = 0''$	Detected sources Reliable / Non-reliable
J0008-34	X ^a	-	N.A.	-
J0031-22	✓	2 / 5099	N.A.	-
J0052-31	✓	1 / 6042	N.A.	-
J0309-41	X ^b	-	X ^b	-
J0310-39	✓	0 / 5806	N.A.	-
J0335-24	✓	2 / 6206	X ^a	-
J0351-38	✓	2 / 6135	N.A.	-
J0419-54	✓	1 / 6011	✓	2 / 6029
J0429-27	✓	1 / 6124	✓	0 / 5957
J0445-59	✓	2 / 6047	N.A.	-
J0454-53	✓	6 / 5774	N.A.	-
J0459-26	✓	2 / 6003	X ^a	-
J0546-52	✓	0 / 6248	✓	2 / 6016
J1253-12	✓	1 / 5905	✓	1 / 5937
J1318-21	✓	2 / 5765	✓	5 / 5651
J1337-28	✓	1 / 5314	✓	3 / 5249
J2257-41	X ^a	-	X ^a	-
J2357-32	X ^c	-	N.A.	-

Notes. N.A. = Not Available.

^(a) Excluded due to residuals as a result of the deconvolution; ^(b) Excluded due to non-optimal noise distribution; ^(c) Excluded due to non-filtered diffuse Magellanic Stream emission.

Table A.2. Stacking detection in the single-track cubes.

Galaxy	Cell α [hh:mm:ss]	Cell δ [deg:mm:ss]	Cell size [arcmin]	V_0 [km s ⁻¹]	Line width [km s ⁻¹]	F_{tot} [mJy beam ⁻¹]	Visually identified
J0031-22	00:30:32	-23:09:36	5.8	667	16.8	0.7	No
J0031-22	00:29:18	-22:57:02	5.8	722	42.0	1.1	No
J0052-31	00:51:55	-31:13:59	2.8	1571	81.2	5.6	Yes ^a
J0335-24	03:37:42	-24:28:24	2.8	1532	51.8	3.5	Yes ^b
J0335-24	03:37:37	-24:33:02	5.8	1485	47.6	2.2	Yes ^b
J0351-38	03:52:04	-38:14:13	8.7	974	29.4	0.8	No
J0351-38	03:49:44	-38:33:22	5.8	845	56.0	1.8	No
J0419-54	04:17:10	-54:43:08	2.8	1512	36.4	2.7	No
J0429-27	04:32:45	-27:42:04	5.8	1001	35.0	1.4	No
J0445-59	04:47:59	-58:51:39	5.8	1290	44.8	1.8	Yes ^c
J0445-59	04:40:26	-59:02:59	5.8	1341	51.8	1.9	No
J0454-53	04:53:45	-53:26:14	8.7	323	44.8	1.2	Yes ^c
J0454-53	04:53:35	-53:27:49	5.8	328	33.6	1.6	Yes ^c
J0454-53	04:52:18	-53:04:18	5.8	346	30.8	1.3	Yes ^c
J0454-53	04:54:05	-53:34:57	2.8	321	30.8	2.8	Yes ^c
J0454-53	04:53:45	-53:32:02	2.8	326	28.0	3.4	Yes ^c
J0454-53	04:53:45	-53:29:07	2.8	328	25.2	2.8	Yes ^c
J0454-53	04:52:27	-53:05:45	2.8	344	42.0	4.1	Yes ^c
J0459-26	05:01:42	-25:43:59	5.8	851	40.6	1.4	No
J0459-26	04:59:32	-26:30:42	5.8	818	32.2	1.4	Yes ^c
J1253-12	12:55:39	-12:05:05	2.8	916	65.8	4.3	Yes ^b
J1253-12	12:52:28	-12:42:59	2.8	1284	29.4	3.3	Yes ^b
J1318-21	13:20:29	-21:06:55	8.7	690	53.2	1.1	No
J1318-21	13:18:30	-21:31:43	5.8	716	29.4	1.2	No
J1337-28	13:35:01	-28:42:11	8.7	489	44.8	1.0	No

Notes. Cell α and cell δ are the central right ascension and declination of the cell. V_0 is the central velocity of the line (in the non-shuffled cube), and F_{tot} is its total flux.

^(a) Clear detection. Full-depth cube not available; ^(b) Confirmed in the full-depth cube; ^(c) Plausible detection. Full-depth cube not available.

Table A.3. Stacking detection in the full-depth cubes.

Galaxy	Cell α [hh:mm:ss]	Cell δ [deg:mm:ss]	Cell size [arcmin]	V_0 [km s ⁻¹]	Line width [km s ⁻¹]	F_{tot} [mJy beam ⁻¹]	Visually identified
J0419-54	04:23:34	-54:51:57	8.7	1500	56.0	0.4	Yes
J0419-54	04:15:23	-55:35:29	8.7	1475	67.2	0.5	No
J0546-52	05:49:37	-52:28:37	5.8	1158	47.6	0.6	Yes
J0546-52	05:42:02	-51:30:11	5.8	1253	79.8	0.7	No
J1253-12	12:53:28	-12:05:03	2.8	837	46.2	1.1	No
J1318-21	13:19:52	-20:58:10	8.7	523	19.6	0.2	No
J1318-21	13:19:20	-21:08:24	5.8	704	71.4	0.8	Yes ^a
J1318-21	13:18:05	-20:56:43	5.8	634	46.2	0.6	Yes ^a
J1318-21	13:17:40	-20:44:56	5.8	678	65.8	0.6	No
J1318-21	13:16:01	-20:38:56	5.8	681	71.4	0.6	No
J1337-28	13:39:53	-28:13:00	2.8	930	16.8	0.5	No
J1337-28	13:38:53	-27:46:51	2.8	492	56.0	1.0	No
J1337-28	13:34:48	-28:18:52	2.8	573	65.8	1.1	No

Notes. Cell α and cell δ are the central right ascension and declination of the cell. V_0 is the central velocity of the line (in the non-shuffled cube), and F_{tot} is its total flux.

^(a) Plausible detection.

Appendix B: Calibration on TNG50 555013

In Sects. 4.2 and 4.3 we present the results of testing our method on the TNG50 galaxy 520885. For completeness, we report here also the outcomes for the TNG50 galaxy 555013. This mock galaxy is mainly characterised by a regular disk and an extended tail in the north-east. In Fig. B.1 we report the alignment residuals in terms of $|mom_1 - v|$, where v is the assumed kinematics. Also for this galaxy most of the residuals are < 20 km/s when the systemic velocity is used to describe the gas kinematics. Finally, in Fig. B.2 it is shown that the S/N of the stacked spectra is lower than the threshold of 5 for almost all the cells, likely due to the overall lower column density of the gas outside the disk with respect to TNG50 520885.

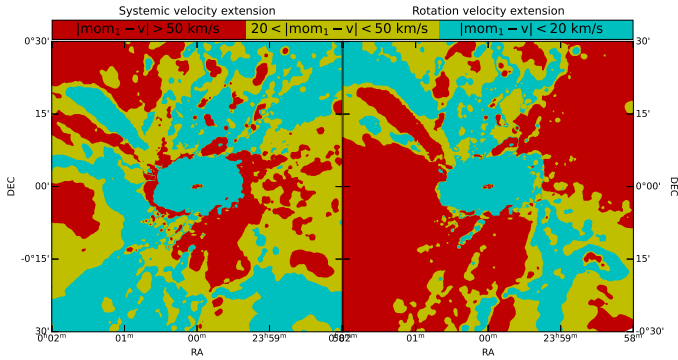


Fig. B.1. Error in the spectral alignment for the TNG50 galaxy 555013 under two different CGM and IGM kinematics assumptions. *Left panel:* the residuals in terms of $mom_1 - v$, where mom_1 is the noiseless moment 1 map and v the assumed kinematics, when $v = v_{sys}$. Light-blue corresponds to $|mom_1 - v| < 20$ km s $^{-1}$, yellow to $20 < |mom_1 - v| < 50$ km s $^{-1}$ and red is $|mom_1 - v| > 50$ km s $^{-1}$. *Right panel:* same as left panel but when v is given by assuming the gas is co-rotating with the galaxy with a flat rotation curve.

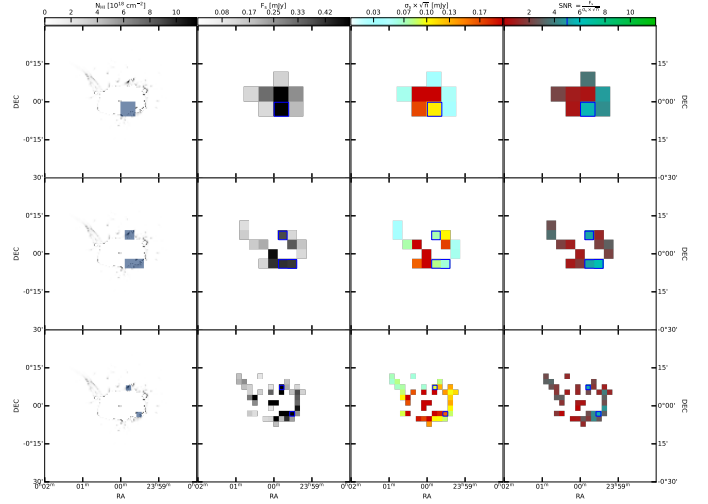


Fig. B.2. Stacking S/N maps for the TNG50 galaxy 555013. *First column:* stacking regions where $S/N > 5$ (blue squares) overlaid with the grey-scale noiseless moment 0 map, blanked from the emission already detected by SoFIA-2. Top to bottom is for cell sizes of 14×14 , 9×9 and 5×5 beams, respectively. *Second column:* the integrated flux of the signal for stacking regions where the HI column density is $> 3.6 \times 10^{17}$ cm $^{-2}$. The blue contours enclose the cells where $S/N > 5$. *Third column:* the noise of the stacked spectrum multiplied by the square root of the number of channels covered by the source. *Fourth column:* integrated S/N of the signal in the stacked spectrum for the regions. The $S/N = 5$ level is indicated also on the colourbar.

Appendix C: The effect of re-gridding

One of the preliminary steps of our stacking procedure is to re-grid the cube such that the pixel size is upscaled to match the beam size. One potential effect of the re-gridding is to lose flux and thus detections. The argument is that after re-gridding the cube violates the Nyquist sampling criterion, as now the pixel size does not match the longest baseline of the interferometer, implying information is lost. To check how important this may be for our stacking experiment, we apply the procedure presented in Sect. 4.1 without re-gridding.

In Fig. C.1 we show the same overlays of Fig. 7. The comparison between the two figures is showing that without re-gridding we pick up less sources. We think that the reason for this counter-intuitive conclusion is that the features we aim to pick up via stacking, that is, HI clouds, are sufficiently large that even after re-gridding they span more than 2.2 beams, satisfying so the Nyquist sampling.

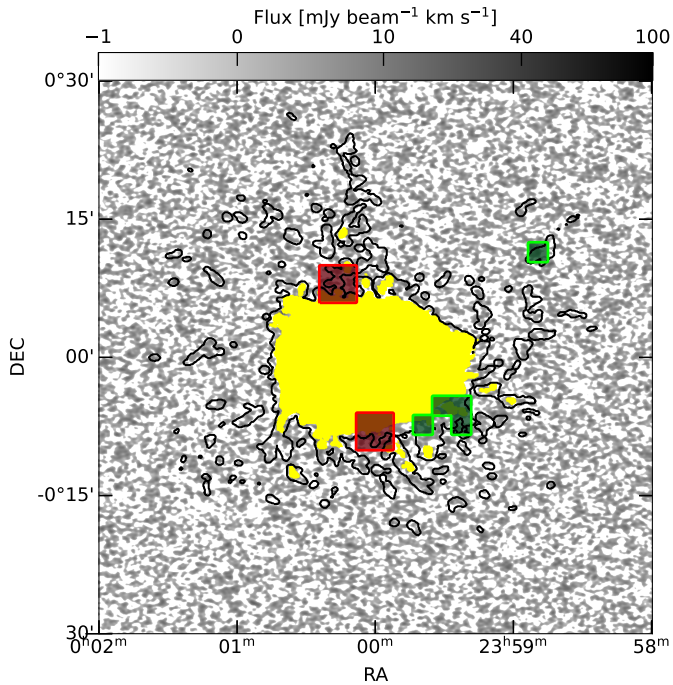


Fig. C.1. Stacking detection map without re-gridding the cube. The background grey-scale image is the cube of the TNG50 520855 galaxy collapsed along the spectral axis and blanked from the galaxy emission. Black contours denote the noiseless moment 0 map clipped at the column density value of $3.6 \times 10^{17} \text{ cm}^{-2}$. The coloured squares indicate the stacking regions where their stacked spectrum contains a reliable detection. Different colours represent different cell size: red for 9×9 and green for 5×5 beams.

Appendix D: Stacking is not equivalent to smoothing

One can argue that the stacking procedure presented in this paper is equivalent to spectrally and spatially smoothing the data at different scales. To check whether this is true we run SoFiA-2 on the masked TNG50 520855 galaxy cube using a spatial smoothing kernel equal to our smallest cell size (i.e. 25×25 pixels, equivalent to 5×5 beams) and spectral smoothing kernels equal to what we have employed with our FINDER (i.e. 1, 7 and 12 channels. cf Table 2). The reliability parameters (minimum S/N of 5 and reliability threshold of 0.85) and the minimum allowed

spatial and spectral sizes for a source (1×1 beam and 7 channels) were set to be equivalent to what we used in our stacking experiment.

In Fig. D.1 we show the moment 0 map as produced by SoFiA-2. Comparing it with Fig. 7, one can see that our stacking procedure is not equivalent to smoothing the cube at different scale. Furthermore, given that the SoFiA-2 detections are outside the $3.6 \times 10^{17} \text{ cm}^{-2}$ column density level, we suspect that what was detected by SoFiA-2 is mainly noise. We want to stress that this comparison is meant to show the differences between our stacking and smoothing and not to compare our procedure with SoFiA-2.

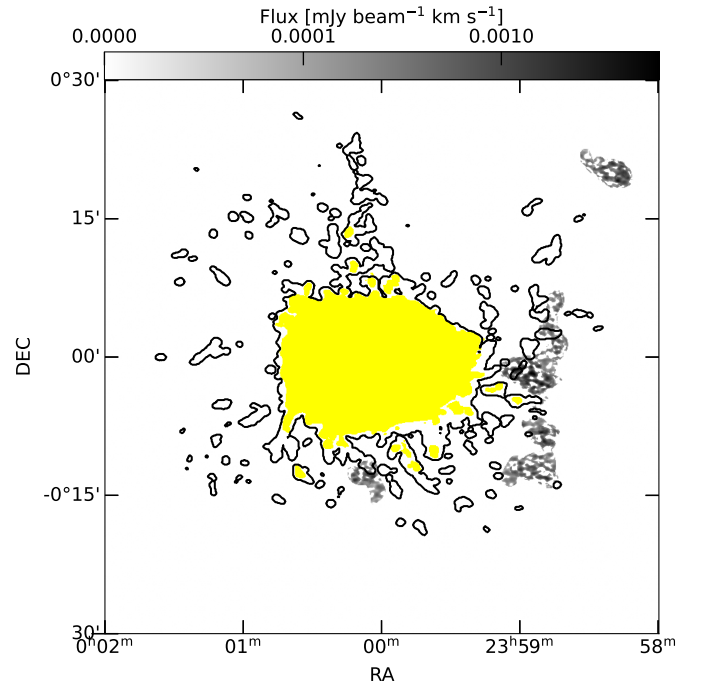


Fig. D.1. Moment 0 map, blanked from the galaxy emission in yellow, produced by SoFiA-2 when emulating our stacking procedure via smoothing. Black contours denote the noiseless moment 0 map clipped at the column density value of $3.6 \times 10^{17} \text{ cm}^{-2}$.