









The *ViSta* Method for Stacking in the Fourier Domain and its Application to the Dusty Star-forming Galaxies in the ALMA Science Archive

Martina Torsello^{1,4} , Marcella Massardi^{1,2} , Elisabetta Liuzzo² , Gayathri Gururajan^{1,3} , Francesca Perrotta¹ , and Andrea Lapi^{1,3,4,5} 

¹ Scuola Internazionale Superiore di Studi Avanzati, Via Bonomea 265, 34136 Trieste, Italy; mtorsell@sissa.it, massardi@ira.inaf.it, ggururaj@sissa.it, perrotta@sissa.it, lapi@sissa.it

² INAF—Istituto di Radioastronomia - Italian ALMA Regional Centre, Via Gobetti 101, 40129 Bologna, Italy; e.liuzzo@ira.inaf.it

³ Institute for Fundamental Physics of the Universe (IFPU), Via Beirut 2, 34014 Trieste, Italy

⁴ INAF—Istituto di Radioastronomia, Via Gobetti 101, 40129 Bologna, Italy

⁵ Istituto Nazionale Fisica Nucleare (INFN), Sezione di Trieste, Via Valerio 2, 34127 Trieste, Italy

Received 2025 August 18; revised 2025 October 29; accepted 2025 November 14; published 2025 December 8

Abstract

We present *ViSta*, a Visibility Stacking method to combine interferometric observations in the Fourier domain at radio to sub-millimeter wavelengths for galaxies. The goal of our method is to maximize the exploitation of available archival interferometric data. By stacking visibilities of galaxies with secure spectroscopic redshifts directly in the Fourier domain and transforming them into the rest-frame, we can enhance the stacked signal, suppress noise, and improve image reconstruction thanks to an extended coverage of the visibility domain. The *ViSta* method is highly flexible, allowing stacking of visibilities regardless of the array configuration or spectral setup. It is effective for both targeted sources and spurious detections offset from the phase center, whether unresolved or extended, within the field of view of the telescope. We validated the method using simulated interferometric datasets. For point-like sources, we can reconstruct the true emission with approximately 90% accuracy, obtaining similar results to classical image-plane stacking. In contrast, for faint and extended sources below the noise level, our method can provide a more accurate estimate of the signal compared to traditional image-based approaches. Finally, we applied *ViSta* to a sample of dusty star-forming galaxies observed with the Atacama Large Millimeter/sub-millimeter Array to detect the CO(3–2) emission line. As for the simulated case, we demonstrated that our tool performs better than image-plane stacking when the signal from individual sources is no longer easily detectable, achieving higher SNR. Finally, we outline potential future applications of this stacking approach.

Unified Astronomy Thesaurus concepts: [Astrochemistry \(75\)](#); [Radio spectroscopy \(1359\)](#); [Astronomy data reduction \(1861\)](#); [Galaxies \(573\)](#)

1. Introduction

The radio to sub-millimetric emission of a galaxy is key to investigating the fundamental stages of galaxy evolution as it traces various components of the interstellar medium (ISM), the star formation, and the feedback mechanisms in a galaxy. At cosmic noon (redshift, $z \sim 2$), we see the peak of the cosmic star formation rate density (SFRD), nuclear activity, and the molecular gas content of galaxies. Among the most remarkable objects observed at high redshift at these wavelengths are the dusty star-forming galaxies (DSFGs). These massive ($M_* \gtrsim 10^{10} M_\odot$), dust-rich, and highly star-forming systems host extreme star formation rates (SFRs) in excess of

$500 M_\odot \text{ yr}^{-1}$, contributing approximately $\sim 20\%$ to the total SFRD and $\sim 30\%$ – 50% to the stellar mass density of the Universe at $z \sim 2$ – 4 (Casey et al. 2014; Madau & Dickinson 2014; Talia et al. 2021; Behiri et al. 2023; Bosi et al. 2025). In addition to their continuum emission at FIR-radio wavelengths, characterized by a combination of synchrotron radiation (associated with nuclear activity and stellar processes) and thermal dust emission (heated by young stars in star-forming regions), DSFGs present rich chemistry associated with gaseous and molecular phases of the ISM. Spectral lines such as CO and [C II] trace different phases of the ISM in DSFGs, providing insights into molecular gas reservoirs, kinematics, and excitation conditions.

However, even for intrinsically bright DSFGs, the signal in individual spectral channels can be very weak (often of the order of only a few tens of μJy), making the detection of line emission challenging at high redshift. To overcome this



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

limitation, strong gravitational lensing is sometimes exploited to increase both the observed flux and the apparent angular resolution of these galaxies (Negrello et al. 2010; Vieira et al. 2013; Gururajan et al. 2022; Giulietti et al. 2024), enabling the study of intrinsically faint dusty star-forming galaxies that would otherwise be inaccessible. In addition, spatial and/or spectral stacking techniques are employed to boost the signal-to-noise ratio (SNR). Image-domain stacking has been broadly and successfully used at these wavelengths to recover higher line flux and retrieve the average properties of the ISM of selected samples of high- z extragalactic sources, which would otherwise remain inaccessible in low-sensitivity images (Knudsen et al. 2005; Carilli et al. 2008; Pannella et al. 2009; Karim et al. 2011; Delhaize et al. 2013; Decarli et al. 2014; Spilker et al. 2014; Lindroos et al. 2015; Jolly et al. 2020; Vernstrom et al. 2021; Neumann et al. 2023).

Even for a sensitive telescope such as the Atacama Large Millimeter Array (ALMA), the observation of most of the spectral emission from fainter targets remains extremely time-consuming, and frequently, only the brightest features (typically CO transitions between $J = 3$ and $J = 6$, [C II], and the dust continuum) can be detected. As a consequence, other potentially important spectral lines are often not searched for or remain undetected in observations.

For instance, HCN and HCO⁺ are effective tracers of dense molecular gas, as they exhibit strong correlations with star formation rates across various populations of galaxies (Imanishi et al. 2023). H₂O lines serve as excellent tracers of warm, dense gas, offering insight into star-forming regions (Perrotta et al. 2023). SiO and CS are commonly used as tracers of shocks in the interstellar medium, as their enhanced emission is typically associated with shocked gas produced by processes such as outflows, cloud–cloud collisions, or galaxy interactions (Martín et al. 2015; Imanishi et al. 2018). Methanol (CH₃OH) can be released from icy grain mantles under similar conditions, indicating dynamic processes such as outflows or cloud–cloud collisions (Huang et al. 2024).

Stacking of archival datasets could be a gold mine for investigating the DSFG population and improving the detectability of the spectral lines discussed above. It is fundamental to note that archival resources are often inhomogeneous, as a result of many projects with different scientific motivations and varying telescope settings (i.e., sensitivity, spatial and spectral resolution); therefore, a standard approach to stacking has usually been to align the images (produced within the same project or even the same mosaic image Decarli et al. 2014; Spilker et al. 2014) before stacking.

Radio astronomical images obtained with an interferometer are the result of an inverse Fourier transform of the observed visibility distribution, namely the map of correlated signals from each antenna couple projected over a plane (usually referred to as “ uv ” or “Fourier domain”) centered on the target position (tagged as “phase center”) (Thompson et al. 2017,

e.g.,). The sampling of the uv -domain is inherently incomplete due to the finite number of antennas and limited observing time, leading to gaps in the Fourier space. In addition, converting visibility data into an image involves several steps: resampling the uv -plane, applying convolution, and performing an inverse Fourier transform, all while accounting for instrumental effects. The resulting image is then used to predict visibilities, which are subtracted from the observed data to generate residuals. Deconvolution is performed to remove the blurring caused by incomplete uv -plane sampling and to refine the image by distinguishing true sky features from distortions. This iterative process gradually improves the sky model, addressing incomplete sampling, and enhancing the image reconstruction.

This technique allows for different possibilities in the weighting scheme, in the deconvolution process, and in the pixelization scheme of the resulting image. Therefore, multiple images can be generated from the same dataset, with different approximations and features appearing differently in the various cases. Furthermore, the noise no longer remains Gaussian upon imaging the visibilities and it corresponds to the telescope response function due to the deconvolution process involved in the imaging. Therefore, uncertainties in the stacking process should be carefully considered.

To overcome these issues, stacking the interferometric data before imaging (i.e., matching the visibilities in the uv -domain) is demonstrated to be more efficient than stacking in the image domain (see Hancock et al. 2011; Knudsen et al. 2015; Lindroos et al. 2015; Hill et al. 2024). However, such methods have been restricted to local galaxies or rely on outdated techniques. To effectively perform visibility plane stacking on sources with different redshift ranges and at high redshifts, it is essential to align the data in the rest frame by applying appropriate frequency scaling and redshift corrections. This process must account for several factors: the effective increase of bandwidth in the observed frequency range, the influence of luminosity distance on the observed flux densities, and the required k -corrections to ensure a consistent visibility-domain stacking across data from different redshifts. To achieve this, we developed the Visibility Stacking tool `ViSta`, a new method to investigate sources at high- z by exploiting the availability of radio and sub-mm archival interferometric data observed using differing telescope settings and for sources at different redshifts, by combining the data in the visibility domain. Thanks to this approach, we can manipulate interferometric datasets directly in the visibility plane, shifting them to the rest-frame while preserving the original file format. This maintains compatibility with the standard calibration and imaging routines of each telescope’s data processing software, ensuring seamless integration of the stacked visibilities into existing analysis pipelines.

The method is presented in Section 2, and the details of its application to the data collected in the ALMA Science Archive

(ASA) are discussed in Section 3. Processing strategies for different source morphologies are described in Section 4. To verify the functionality and the efficiency of our method in retrieving spectral features from ALMA observations, we have generated simulated data and compared the `ViSta` results with a more traditional stacking in the image domain (see Section 5). Applications to selected real ALMA observations to address specific science cases will be detailed in future papers from our collaboration as they go beyond the scope of this paper. However, as an example of application to real data, we report in Section 6 the results of stacking to investigate the CO(3–2) line in a small sample of DSFG observations stored in the ASA. Finally, in Section 7 we summarize our approach and findings and discuss its applications.

Throughout this paper, we adopt a flat Λ CDM cosmology (Planck Collaboration 2020) with round parameter values $h \approx 0.67$, $\Omega_m \approx 0.3$ and $\Omega_\Lambda \approx 0.7$.

2. The `ViSta` Method

The basic principle of stacking is to align and combine multiple observations of faint or undetected sources, such as galaxies, to enhance their collective signal. As we are interested in spectral features, we match sources for which we know the position and spectroscopic redshift with high confidence. Indeed, during data combination, even small errors of a few percent can cause a flux loss of nearly 50% and lead to an increasingly significant broadening of the emission line (Jolly et al. 2020). Successful stacking requires verifying that each dataset’s spatial filtering and phasing are compatible, ensuring the emission is sampled consistently across all observations. In a radio or sub-mm interferometer, the coverage of the uv plane for a given antenna configuration depends on the observing wavelength λ , total observation time, and the range of baselines, from the shortest distance between any pair of antennas b_{\min} to the longest b_{\max} . As the Earth rotates, the projected baselines between antenna pairs trace different points in the uv plane, thereby improving the sampling over time.

Therefore, the array is able to recover emissions from the source components with angular sizes between the observed resolution $\Theta_{\min} \sim \lambda/b_{\max}$ and the maximum recoverable scale $\Theta_{\max} \sim \lambda/b_{\min}$: all emission from regions with size $\Theta > \Theta_{\max}$ is filtered out, while those $\Theta < \Theta_{\min}$ are smoothed to the resolution scale. Different array configurations, spectral set-up and observation times sample different angular scales that should be properly matched to reconstruct average properties.

If a source is observed in a configuration with its angular scale significantly smaller than Θ_{\min} , it could be approximated to a Dirac δ in the phase center: its Fourier transform is a constant and therefore the visibility measured amplitude value will be the same over all baselines (i.e., whatever is the array configuration).

We can also include resolved sources with angular size Θ_s such that $\Theta_{\min} < \Theta_s < \Theta_{\max}$, provided that the uv plane is uniformly sampled across the dataset. Moreover, any differences in angular scale sampling across the dataset (e.g., due to different array configurations or observation durations) must be carefully characterized. One strategy to mitigate this is to apply tapering during imaging to match the effective resolution and sensitivity to extended emission. Alternatively, if we want to be even more strict, we can apply tapering that downweights baselines longer than $b > \lambda/\Theta_s$, effectively smoothing the observations so that the source appears point-like at the tapered resolution.

The presence of other signals in the same field of view (FoV) might contaminate the visibilities, thus in the case of bright contaminating sources, the measured signal is no longer consistent with a constant. Similar cases must be discarded to avoid undesired contamination.

In addition, we aim to exclude high-resolution observations of strongly lensed sources, as well as sources that are significantly more resolved than the rest of the sample. The possible approaches we may adopt for handling these cases, which are still under investigation, are described in the subsequent sections. Therefore, we select only unlensed and moderately resolved sources that meet the above criteria. Once we have selected a well-defined, homogeneous sample of sources, we can perform stacking directly in the visibility domain. To combine the visibilities, before performing the Fourier transform, it is essential that the uv -plane sampling remains consistent across all sources, ensuring a uniform angular reference scale. Since each visibility intrinsically encodes its observing wavelength, coherence can only be achieved by reprojecting every visibility to a single, common reference, i.e., the rest-frame wavelength λ_{RF} . Spectral matching is performed by shifting each dataset to the rest-frame wavelength $\lambda_{\text{RF}} = \lambda/(1+z)$. Therefore, the input dataset at wavelength λ is equivalent to being observed at λ_{RF} by the same array but with a more compact configuration with antennas at distances $b/(1+z)$. Furthermore, the spectral resolution is also increased to $\Delta\lambda_{\text{RF}} = \Delta\lambda/(1+z)$. Once coherence has been established across all sources, their visibilities can be concatenated to achieve consistent coverage of the same uv -plane.

The core of the data pre-processing in the `ViSta` method is to adjust the telescope settings and manipulate the interferometric data to simulate an observation of the same identical source as if it were at redshift $z = 0$, but emitting at its rest-frame wavelength with a rest-frame luminosity density $L_{\text{RF}} = S_{\text{obs}} 4\pi D_L^2/(1+z)$, where S_{obs} is the observed flux density and D_L the luminosity distance. Once every dataset to be stacked has been shifted to the rest frame and properly centered, we also apply a correction for the primary beam attenuation that the source would have experienced at its

original position in the field. This ensures that the flux density is accurately recovered after the spatial shift to the phase center. Visibilities must then be spectrally regridded to a finer common spectral grid. We adopt the spectral sampling of the dataset with the highest native resolution in order to prevent undersampling and preserve the fidelity of narrow spectral features.

Once the pre-processing steps (centering with primary beam correction, rest-framing, and spectral alignment) are complete, we can concatenate all datasets and treat them as if they were multiple epochs or configurations observing the same target. Note that in this process, there is no loss of information about the spectral profile and each dataset is intrinsically weighted for its thermal noise limits.

After concatenation, the resulting dataset is fully processed and ready for scientific analysis. Additional post-processing steps can be applied depending on the specific scientific goals and the data’s characteristics. It is important to note that these steps are optional, as the core stacking procedure formally concludes with the concatenation, while any further operations are left to the user’s discretion.

For example, once the stacking is complete, we can perform continuum subtraction to isolate the spectral line signals and simultaneously estimate the average continuum flux of the sample. In this way, we also estimate the average continuum flux of the sample. However, continuum removal may not be recommended when the line emission does not significantly stand out from the continuum, as this could lead to unintended suppression of the signal.

Similarly, the weighting step is another optional operation used to correct for uneven spectral coverage. To avoid statistical biases, which can arise when spectral windows do not perfectly overlap and some channels are averaged more heavily than others, each source can be weighted in each channel to account for possible repetitions in the concatenation, and the resulting visibilities are re-normalized channel by channel to achieve a more uniform combination of the datasets. This weighting may not be advantageous in cases of marginal line detections, where it could inadvertently suppress the signal, particularly if nearby lines fall within the stacking window.

Among the post-processing operations, we also include two methods for extracting the final flux. We can measure the spectrum directly by averaging the visibility values (over the whole uv -domain) in each channel: the signal would be amplified and the Gaussian noise would be minimized. It is also possible to generate an image out of the concatenated visibilities and measure the luminosity of the stacked source. The resulting image will benefit from an improved uv -coverage compared to those obtained from any individual input dataset. While these two approaches are provided as practical tools to recover the bulk of the emitted flux from the

stacked signal, the user is free to adopt any alternative method that best suits their scientific objectives.

3. Applications to the ALMA Science Archive

In this section we describe how each step of the `ViSta` procedure is implemented to deal with ALMA data stored in the ASA. While the described method could, in principle, be applied to any sample of sources with known position, redshift and size, we mainly focus on DSFGs. We selected a sample of ~ 200 DSFGs with the criteria that the galaxy falls within the redshift range $z = 2\text{--}4.5$ and has been observed in at least one ALMA band. Subsamples can be defined depending on specific scientific goals (which we will describe in Section 7 for our sample of galaxies).

After calibrating each dataset,⁶ visibilities appear in the Measurement Set (Dijkema et al. 2019, MS) format: a collection of inter-dependent tables, containing information on the visibilities, telescope settings, applied calibration gains, metadata summarizing the dataset history, among other relevant parameters.

To analyze the ALMA Measurement Sets within the `ViSta` framework, we mainly employed version 6.6 of the Common Astronomy Software Applications (CASA) package, which is the primary data processing software for ALMA (Team et al. 2022). It supports a wide range of operations, including calibration, imaging, and data analysis, making it well-suited for handling our datasets. In addition, we employ several Python libraries to facilitate data manipulation and analysis, including `Astropy` for managing astronomical data structures and units, `SciPy` and `NumPy` for numerical computations, and `Matplotlib` for data visualization.

Figure 1 presents a comprehensive overview of the `ViSta` workflow. At each step, we indicate the specific packages involved and the functions used to modify the data, along with the MS tables that we edited via Python scripts. This figure serves as a visual guide to the data processing pipeline for stacking ALMA data described in the following paragraphs. In the case of other interferometric telescopes, the pipeline will remain conceptually the same, although different software tools will be employed and the data will be adjusted accordingly.

3.1. Centering

The centering process ensures that all images are aligned to a common reference position, allowing coherent stacking of multiple observations. First, we need to determine the precise position of the source. This position can be derived either from the continuum emission of the observation itself, for instance, by fitting a Gaussian to the source emission, or by exploiting a

⁶ We mostly retrieved calibrated datasets by exploiting the EU ARC CalMS service <https://almascience.eso.org/tools/eu-arc-network/the-european-arc-calms-service> both for archival products and for the ARI-L project products (Massardi et al. 2021).

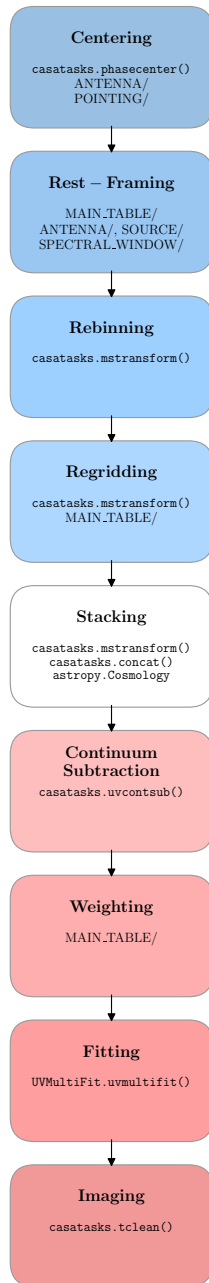


Figure 1. The ViSta method general workflow. We also provide specifics for ALMA Measurement Sets, listing the functions from CASA or other astronomical libraries and the tables that have been manually modified. The blue color cells refers to the input data editing phase, while the red color cells represents the output extraction steps.

different observation of the same source retrieved from the ALMA archive. Once the accurate coordinates of the source are identified, the phase center of the observation must be shifted to align with this reference position.

This phase center shift is performed using the `casatasks.phaseshift()` function, which is applied to each input MS of the sample, hereafter referred to as *file.ms*, and

applies a phase correction based on the new coordinates (the specified “phasecenter”) and produces a corrected MS *file.ms.center* where the source is now at the center of the field. This approach ensures that the original data remain untouched.

After the phase center has been shifted, we modify the metadata stored in the FIELD, SOURCE, and POINTING tables of the MS. In particular, we simulate that all sources are located at coordinate zero, which is essential for correctly concatenating multiple datasets. This is done by setting to zero the entries corresponding to the phase direction PHASE_DIR, reference direction REFERENCE_DIR, and delay direction DELAY_DIR columns in the FIELD folder. Additionally, we set the source direction entries in the POINTING table to zero, ensuring full consistency across the dataset. These steps guarantee that all sources in the sample are treated as if they were observed at the same position, allowing accurate combination and stacking in the following analysis steps.

Once the source has been shifted to the center of the field, we divide each visibility by the correction for primary beam attenuation that it would have experienced at its original position in the field of view. This correction is crucial to recover the true flux density of the source prior to stacking. The primary beam (PB) response of an ALMA 12 m antenna can be approximated by a normalized symmetric Gaussian function that depends on the observing frequency ν and the angular offset θ from the pointing center:

$$\text{PB}(\theta) = \exp \left[-4 \ln(2) \left(\frac{\theta \nu d}{1.2 c} \right)^2 \right], \quad (1)$$

where d is the antenna diameter and c is the speed of light. This ensures that all sources are placed on a consistent flux scale regardless of their original location in the field, enabling accurate stacking across the sample.

3.2. Restframing

The process of rest-framing the visibilities begins by shifting the observed frequencies to the rest frame. The process of rest-framing the visibilities begins by shifting the observed frequencies to the rest frame. This transformation requires modifying the frequencies, along with all dependent parameters (for example, the channel spectral resolution and the total bandwidth) by applying the corrective factor $(1 + z)$. For a typical MS from a given ALMA observation, majority of the parameters that depend directly on the frequency can be found in the SPECTRAL_WINDOW folder. These parameters include the central frequencies for each channel, CHAN_FREQUENCY, the reference frequency for the spectral window, REF_FREQUENCY, the spectral width of each channel, CHAN_WIDTH, the effective noise bandwidth of each spectral channel, EFFECTIVE_BW, the effective spectral resolution of each channel, RESOLUTION, and the total bandwidth for this

window, TOTAL_BANDWIDTH. We create a copy file called (*file.ms.center.rest*) where we apply all these modifications.

Furthermore, the baseline vectors and antenna positions stored in the MS must also be corrected, since they are expressed in units of wavelength λ and thus depend directly on the observed frequency. In particular, the visibility coordinates in the main table UVW, the antenna positions in the ANTENNA table POSITION, and the dish diameter DISH_DIAMETER must be divided by the corrective factor $1 + z$. In addition, the dish diameter must also be rescaled to preserve the angular resolution as the beam size $\theta \approx \lambda/d$, so adjusting d by $(1 + z)$ thus preserves the consistency of θ between the observed and the rest-frame.

3.3. Rebinning

Before merging different datasets, it is essential to ensure consistency along the spectral axis across all MS, meaning that all datasets must share the same spectral resolution and channel width so that spectral features are properly aligned and can be reliably combined. Initially, the spectral resolution of each observation is adjusted using the `casatasks.mstransform()` tool to match the widest spectral resolution among the sample and returns a rebinned measurement set *file.ms.center.rest.rebin*. This step prevents the introduction of artifacts during subsequent data combination. To improve spectral smoothness, one can first apply a Hanning smoothing filter with `casatasks.hanningsmooth()`, which performs a weighted running average of each channel with its two neighbors (0.5 for the central channel, 0.25 for each adjacent channel) and flags the first and last channels, producing *file.ms.center.rest.smooth*. This smoothed MS can be further rebinned with `casatasks.mstransform()`, yielding *file.ms.center.rest.smooth.rebin*. Note that if the correlator has already pre-averaged the data during calibration, `hanningsmooth` will detect this and the smoothing step for those spectral windows can be skipped.

3.4. Regridding

The next step is to standardize all observations so that they have the same number of channels. This is a prerequisite for using the CASA `casatasks.concat()` function subsequently, which requires consistent spectral window shapes across the datasets to merge them. In the future, however, the plan is to avoid combining the data with this function, allowing for more flexibility in the choice of spectral regions. Therefore, we adopt a number appropriate for the dataset, constrained by the MS with the fewest, and for each MS we select the spectral window that best overlaps with the others and matches our scientific interests. As a result, some regions of the spectrum will be covered by multiple observations, while others will have fewer contributions. We can decide to limit ourselves exclusively to the spectral region where the entire dataset

overlaps, although this would discard a substantial amount of data, or, as we will explain in the following steps, carefully weight each channel according to the number of observations that have a detection in that channel. Another possibility is to add “dummy” channels to avoid being limited by the size of the final spectral window. For example, suppose we want the largest possible spectral window that covers all frequencies observed in our dataset. To define the final frequency range, we select the minimum frequency from the lowest observed range and the maximum frequency from the highest observed range across all datasets. To standardize the number of channels, we identify the observation with the largest number of channels and use it as a reference. The spectral coverage of the other datasets is then expanded by adding flagged channels at the edges of their spectral axes, ensuring that all observations have the same number of channels without exceeding the overall minimum and maximum frequency limits. To add the missing channels, we first created multiple copies of the original MS, to preserve its full table structure, including all spectral windows and metadata. For each copy, we then run `casatasks.flagdata()`, the all-purpose CASA flagging task that marks data as invalid, thereby turning every visibility into a “placeholder” flagged channel. Next, we merge these flagged MS copies back into the original dataset using `casatasks.concat()`, which preserves the main structure of its spectral features by adding the new channels on the old one. Finally, we apply `casatasks.mstransform()` to cut the measurement set and extract our final spectral cube. We trim any excess flagged channels outside the desired overall frequency bounds and regrid the data onto the target channelization, producing a uniformly sampled cube ready for stacking called *file.ms.center.rest.rebin.regrid*.

3.5. Stacking

Once the data are prepared, we proceed with the actual stacking operation. Before combining the measurement sets, we rescale the visibilities of each observation by multiplying them by a correction factor to adjust the flux densities to a common reference redshift. Specifically, each visibility is multiplied by

$$w = \frac{D_L^2(z_{\text{source}})}{D_L^2(z_{\text{common}})} \times \frac{1 + z_{\text{common}}}{1 + z_{\text{source}}} \quad (2)$$

where z_{source} is the source redshift, z_{common} is the reference redshift chosen for the stacking analysis, and $D_L(z)$ is the luminosity distance at redshift z . This correction accounts for the difference in luminosity distance and redshift effects, ensuring that flux densities are consistent and comparable across the sample. This rescaling step is optional and can be skipped or adjusted according to the specific scientific goals of the analysis. After this rescaling, the measurement sets are

concatenated using the CASA `concat()` function to produce a single combined dataset, *concat.ms*.

Furthermore, if a source is gravitationally lensed and remains unresolved or partially resolved in our data, its flux is corrected by dividing the corresponding visibilities by the lensing magnification factor before stacking. This removes the artificial flux boost due to lensing, allowing the stacked dataset to reflect the intrinsic (unlensed) brightness of all sources.

If the data have been properly pre-processed and aligned, the `concat()` function treats all the observed sources as a single field centered at zero with a unified spectral window. Although the sample is homogeneous, each source is added to *concat.ms* as a separate observation. This structure can complicate operations such as time averaging using `mstransform`, which cannot average across scans belonging to different observations. Time averaging is also necessary to reduce computational load and to visualize the stacked compact spectra in the *uv*-plane using plotting tools like `casaplotms`.

To address these issues, we modify the OBSERVATION table so that all scans appear as part of a single observation. First, we determine the global time span of the dataset by extracting the minimum and maximum values from the TIME_RANGE column. We then update the first row (observation ID 0) of the OBSERVATION table to cover this full time range. Subsequently, we overwrite the OBSERVATION_ID column in the main table, setting all entries to 0. This effectively unifies the dataset under a single observation ID. At this point, we can safely use `casatasks.mstransform()` to extract only observation ID 0 and discard the others in a new *concat.ms.unified* file, on which we can perform operations like time averaging.

As discussed previously, the stacking procedure formally concludes with the concatenation step, and any further processing is optional and depends on the specific scientific goals.

3.6. Continuum Subtraction and Analysis

After stacking the data, we may proceed to separate the continuum emission and isolate the spectral line signals. This is achieved using `casatasks.uvcontsub`, which performs continuum fitting and subtraction in the *uv*-domain. The task estimates the continuum emission by fitting polynomials to the real and imaginary parts of the visibility data over line-free channels specified by the `fitspw` parameter. Typically, a zero-order polynomial (i.e., a constant) is used to avoid overfitting and to preserve the line flux. The fitted continuum model is then subtracted from the data, resulting in a new MS *concat.ms.unified.contsub*, which contains only the line emission. Optionally, if the parameter `write_model = True` is set, the task also writes the fitted model in the MODEL column of the main table of the output MS. In fact, this process provides the

continuum term concatenated over the whole sample, corresponding to the stacked visibility of the continuum signal. By producing a multi-frequency-synthesis (mfs) image, for example with the TCLEAN task, we obtain the stacked continuum map, ready for any population continuum analysis.

3.7. Weighting

During the imaging process, the data are weighted according to the noise in each of the source datasets (i.e., noisier data, coming from shorter integration times, receive less weight). This inverse-variance weighting is usually optimal for maximizing the overall signal-to-noise ratio and is applied automatically during the cleaning step. Moreover, as we have seen in the regridding step, some channels may be covered by fewer observations than others, and we can account for this in the subsequent processing. To standardize the frequency grid across all observations while regridding, we add flagged data, which leads to regions that could be averaged with inhomogeneous numbers of sources. Therefore, to ensure statistical consistency across the spectral window, we account for the number of observations that cover each frequency and apply corresponding statistical weights to each channel. The weight assigned to each channel is defined as:

$$w(\nu) = \sqrt{\frac{N_{\text{coverage}}(\nu)}{N}}, \quad (3)$$

where $N_{\text{coverage}}(\nu)$ is the number of observations that include the frequency ν , and N is the total number of sources in the sample.

To compute $N_{\text{coverage}}(\nu)$, we first construct a uniform frequency array with the same shape as the unified spectral window. For each observation in the dataset, we determine the indices of the grid corresponding to its original frequency range (from its minimum to maximum frequency) and increment the entries in this range by one, accumulating the counts of sources contributing to each frequency bin. This results in an array that encodes the number of overlapping observations per grid frequency. As the centers of the unprocessed spectral channels may not align exactly with the grid points, we interpolate this coverage array onto the precise channel frequencies to obtain a continuous estimate of $N_{\text{coverage}}(\nu)$.

Moreover, if two observations of the same source overlap in some channels, we must reduce the weight of each observation on those channels (by dividing by the number of times the source appears) to avoid multiple counting. This ensures that, statistically, each source contributes only once in a single channel to the final stacked dataset, regardless of how many times it is observed at that frequency across different datasets.

The choice of weighting is flexible and can be adapted according to the specific scientific goals of the analysis. Users may choose to skip this step entirely or modify individual

observations, allowing them to apply the weighting scheme that best suits their needs.

3.8. Spectral Line Fitting in the Visibility Domain

To fit the stacked data from ALMA observations directly on the visibility plane, we utilize the `UVMultiFit` package, a versatile tool designed to fit models directly on interferometric visibility data within the CASA environment. This approach bypasses the imaging process, enabling more accurate parameter estimation, in particular for unresolved or partially resolved sources.

In our analysis, we employ the `UVMultifit.uvmulti-fit()` function, selecting either the `delta` model (corresponding to a point source) or the `Gaussian` model, depending on the source morphology.

The fitting procedure is performed on a per-channel basis by setting `OneFitPerChannel = True`, providing an estimate of the flux density and its associated uncertainty for each spectral channel. Additionally, using the `write = "residuals"` parameter, the residual visibilities are saved in a column of the measurement set. This enables the assessment of the fit quality through residual analysis.

Finally, we plot the fitted flux densities as a function of frequency, providing the spectrum in the visibility domain. We can model this spectrum by fitting a Gaussian profile, facilitating the characterization of spectral features directly from the visibility data.

3.9. Imaging

The final step involves transforming the stacked visibilities into the image plane using the `casatasks.tclean()` task from the CASA software package, which is a versatile imaging tool that reconstructs images from interferometric data by deconvolving the point-spread function (PSF) from the observed visibilities.

To image our stacked spectra, we adopt the following parameter settings:

- (1) *Deconvolution Algorithm.* We use the *Hogbom* algorithm, a classic cleaning method suitable for fields dominated by point sources or compact emission. This algorithm iteratively identifies the brightest pixel in the residual image, subtracts a scaled PSF, and updates the model image accordingly.
- (2) *Weighting Scheme.* We select *natural* weighting to maximize sensitivity. This weighting scheme gives more weight to densely sampled regions in the uv -plane, resulting in lower noise levels at the expense of angular resolution. Since our primary goal is to recover the total flux of the stacked source rather than to resolve its detailed spatial structure, the loss in resolution is acceptable for our scientific goals.

- (3) *Cell Size.* We define the cell size to be 1/5 of the major axis of the restoring beam. This choice ensures adequate sampling of the PSF.

First, we generate a dirty image ($niter=0$) to check for possible artifacts introduced by the stacking process and to evaluate whether the source is already detected above the noise. We determine the cleaning threshold for the deconvolution based on the average sensitivity expected for the sample, which is derived from the estimated sensitivity σ_k of each individual observation k as:

$$\sigma^2 = \frac{1}{\sum_{k=1}^N 1/\sigma_k^2}. \quad (4)$$

The threshold is typically set at a level of 3σ . With these parameters, we perform the deconvolution using `tclean`, which produces the image cube, along with the residuals, model image, point-spread function (PSF), and primary beam. These products can then be visualized with an image viewer such as `CARTA`, to analyze the spatial and spectral properties of the detected emission.

4. Processing Strategies for Different Source Morphologies

The visibility–stacking procedure described in the previous sections is fundamentally independent of the intrinsic angular size or complexity of the target sources, provided all the datasets uniformly sample the uv -plane between a common minimum baseline (Θ_{\min}) and maximum baseline (Θ_{\max}). However, source morphology must be taken into account when combining the ensemble of visibilities. Point-like, partially resolved, and extended (or lensed) sources require different treatments, and combining them together requires specific post-processing considerations. Therefore, in Sections 4.1, 4.2 and 4.3, we outline how to handle different source morphologies and describe the corresponding signal extraction strategies for unresolved sources, partially resolved sources, and resolved or lensed sources.

4.1. Unresolved Sources

In the simplest case, where the entire signal is contained within the synthesized beam, the source can be treated as unresolved. If we choose to extract the flux after imaging, we can simply take the peak intensity within the synthesized beam. Alternatively, for the analysis in the visibility domain, we use `uvmultifit` with a Gaussian profile: the size parameters are allowed to vary within physically motivated priors, while the position is fixed at the center. This approach provides a more robust estimate than a delta-function model, particularly in the presence of noise or slight deviations from ideal point-like structure.

4.2. Partially Resolved Sources

In cases where the sample includes partially resolved sources, it is important to evaluate whether their intrinsic angular sizes are smaller or larger than the common restoring beam determined by the combined uv -coverage. When the beam is at least three times larger than the intrinsic source size, individual emissions are smoothed in the stacking process, permitting treatment analogous to that of point-like sources within the same framework. However, if the signal-to-noise ratio is low or the data contain a substantial amount of noise, those noisy fluctuations will also be convolved into the beam and propagated into the final stack. This effectively raises the noise floor of the stacked image and can bias any measurement of faint emission.

In contrast, when the angular extent of the source approaches or surpasses the beam size, its peak image intensity no longer captures the total flux density. In such cases, the appropriate procedure is to explicitly model the spatial emission profile. If the angular size of the source in the band of interest is known, one could measure the flux density directly using `casatasks.imstat()` within an aperture tailored to the emission; however, since the intrinsic extent typically varies with frequency and the target line may exhibit velocity dispersion across the band, it is preferable to perform a two-dimensional Gaussian fit with `casatasks.imfit()`. Alternatively, a visibility-domain Gaussian fit with `uvmultifit` with a Gaussian model and appropriately broad size priors can be used to estimate the total flux density of the stacked source, accounting for the intrinsic source size variation across frequency.

Another strategy is to enforce a common resolution across the sample directly during imaging. This can be achieved using the `restoringbeam` parameter in the `tclean` task in CASA, which forces the final image to adopt a beam size corresponding to the worst (i.e., largest) resolution in the dataset. This ensures uniform angular resolution across all sources, allowing a consistent and unbiased comparison within the stacked sample. Finally, the flux of the stacked source can then be estimated with `casatasks.imfit()` similar to the previous case.

4.3. Resolved and High-resolution Sources

When sources are well resolved (typically at angular resolutions of $0''.2$ – $0''.1$) and exhibit clearly extended structures, as in the case of high-resolution observations of lensed galaxies, they require additional care to be included in a stacked sample alongside unresolved or partially resolved sources. An effective strategy to address this is the application of uv -tapering. Tapering is a weighting technique applied during imaging that downweights visibilities at longer baselines, effectively reducing the angular resolution of the final image. By applying a tapering function (typically Gaussian in

the uv -plane), one can smooth high-resolution data to match the angular scale probed by lower-resolution observations. This ensures that all sources in the sample are treated with a consistent effective beam size, thereby enabling a meaningful combination in the stacking process. Once the appropriate tapering has been applied to smooth the more highly resolved sources to a resolution compatible with the rest of the sample, these sources can be modeled following the same procedure described for partially resolved sources, with an image-plane modeling or tailored visibility-domain fitting with flexible priors. However, these strategies are yet to be further validated and could be refined by future studies.

5. Example on Simulated Data

5.1. Building Simulated ALMA dataset

To verify the proper functioning of `ViSta`, we test the validity and functioning of `Vista` on a sample of 50 simulated galaxies. Each of the mock galaxies is assigned a random value of redshift between 2 and 4, where we simulate both the continuum and the line emission. The emission line is modeled as a Gaussian bump superimposed on the continuum, and it is located at the same rest-frame wavelength for all sources.

Then, we use the task `simobserve`⁷ in the Common Application Software for Astronomy (CASA) v.6 to simulate ALMA observation for each source. The task accepts as input several parameters like the sky model, the telescope configuration (thus the angular resolution of the observation), the observation date and the total observation time and returns a simulated MS with the same construct of an actual ALMA observation.

`simobserve` allows for adding thermal noise according to realistic models of the atmospheric profile for the ALMA site, tabulated in dedicated libraries,⁸ and described as:

$$S_\nu(\text{noise}) = \frac{4\sqrt{2}k}{\eta_a \eta_c \pi d^2 \sqrt{\Delta\nu} \Delta t} T_{\text{sys}} \quad (5)$$

where k is the Boltzmann constant, η_a and η_c are the antenna and correlator efficiency, d is the diameter of the antennas, $\Delta\nu$ is the observed bandwidth and Δt is the integration time. The system temperature, T_{sys} , comprises contributions from various components and is defined as:

$$T_{\text{sys}} = T_{\text{CMB}} + \eta_s T_{\text{atm}}(e^{\tau_\nu} - 1) + (1 - \eta_s) T_{\text{amb}} e^{\tau_\nu} + T_{\text{RX}} e^{\tau_\nu} \quad (6)$$

where T_{CMB} is the temperature of the cosmic microwave background, η_s indicates the spillover efficiency of the antenna, T_{atm} is the atmospheric temperature, τ_ν denotes the atmospheric opacity at the observed frequency, T_{amb} is

⁷ https://casaguides.nrao.edu/index.php/Simulating_Observations_in_CASA_6

⁸ https://cab.inta-csic.es/users/jrpardo/class_atm.html

the ambient temperature and T_{RX} represents the receiver temperature. Together, these factors account for the total thermal noise introduced during the observation process.

However, the corruption procedure in the `simobserve` task only calculates a single average T_{sys} for all visibilities and integration times, without considering the phase delay for each antenna introduced by the atmosphere. As a result, the simulator generates individual noises with an identical standard deviation and the weights and sigma associated with each individual antenna are identical and set to unity. This not only fails to accurately represent a realistic observation, where each antenna is affected differently by the atmospheric contribution, but also does not allow us to exploit the full potential of stacking in the visibility plane (as it does not include random fluctuations of the noise).

For these reasons, we choose to generate the MS for each mock source without noise and then corrupt them using custom Python functions. These functions add thermal noise, with the functional description given by Equation (5), by exploiting the same libraries as `simobserve`, but introducing phase delay and fluctuations. The atmospheric temperature and thermal opacity are individually calculated for each antenna, taking into account factors such as antenna elevation and the air mass above each one. Other parameters, such as humidity, ambient temperature, etc., are set based on the values suggested by the ALMA Sensitivity Calculator,⁹ with random variations of up to 10% applied to key observational parameters, including atmospheric opacity, ground temperature, and receiver temperature, to ensure each simulated observation is unique.

We generate simulations for two types of sources: a point-like source and an extended Gaussian source. For each case, we produce 50 synthetic observations of sources with redshifts randomly distributed between 2 and 4, using different antenna configurations from CASA’s predefined `alma.cycle.11.n.cfg` files, with n ranging from 2 (corresponding to configurations with baselines up to 314 m) to 5 (with baselines up to 1400 m).¹⁰

The input sky models consist of spectral cubes with 100 channels, each 1 MHz wide. The reference frequency, corresponding to the central channel of all cubes, is set to 345.7598 GHz. This value is chosen to provide consistency with the real-case application of the method to the stacking of the CO(3–2) emission line, discussed later in the paper. However, it does not represent the line itself, and no specific molecular transition is modeled at this stage. All cubes are constructed such that their central channel corresponds to the observed frequency at which the redshifted line falls, ensuring that the same rest-frame frequency is aligned across sources, even though the observed frequencies differ due to their individual redshifts. Each channel is represented by a

512×512 pixel image, with a pixel scale of $0.05''$, ensuring oversampling relative to the synthesized beam of all configurations. The sources are not placed exactly at the center of the image but are instead offset by a fixed amount from the central pixel, while still being fully contained within the sky model field of view.

In the point-source case, the cube includes a spatially and spectrally centered point-like source with both continuum and line emission. The continuum is modeled as a flat component with an intensity of 0.6 mJy per channel across the full bandwidth. The line emission is added on top of the continuum and modeled as a Gaussian profile centered on the source position, confined to the 20 central spectral channels, with a peak flux of 2 mJy. This observed frequency is chosen so that, at the considered redshifts, the emission line falls within ALMA Band 3.

In the extended-source case, we model a spatially resolved Gaussian source with a FWHM of $5''$, also centered in space and frequency. This source includes the same flat continuum level (~ 0.6 mJy per channel) and the Gaussian profile centered on the middle channel and limited to the 20 central spectral channels, with a peak flux of ~ 2 mJy.

In both cases, we then corrupt the mock sources by adding the realistic thermal noise described above, following the functional form given by Equation (5); for the redshift range considered in this study, the observed emission lines fall within ALMA Band 3, where the thermal noise level at a channel width of 1 MHz, under the adopting observing setup, is about ~ 1 mJy (Cortes et al. 2025).

These conditions ensure that the sources (both continuum and line emission) are not detectable in individual observations, enabling an exploration of detection and reconstruction strategies under low-SNR regimes.

For all sources, we avoided post-processing steps such as continuum subtraction to prevent bias in the characterization of the emission. Additionally, weighting was not applied since we considered only the frequency range common to all sources in the final analysis. We did not perform any correction for flux scaling due to redshift differences, as this is a simulated case where all fluxes were set equal from the beginning without any redshift-dependent adjustment.

Figure 2 provides a direct view of the effect of the `ViSta` rest-framing step described in Section 3.2 on the uv -plane of a single source. The uv -plane coverage appears different, with the uv -distances scaled down by a factor of $(1+z)$. Nevertheless, when the `ViSta` procedure is correctly applied and the MSs are properly adjusted, the resulting image is effectively identical to the original, differing only by the shift to the rest-frame frequency.

In Figure 3, we instead illustrate the uv -plane coverage of the stacked mock observations after applying the `ViSta` procedure. The left panel displays the different array configurations adopted in the observing setups, while the right panel highlights the visibilities color-coded by redshift. The

⁹ <https://almascience.eso.org/proposing/sensitivity-calculator>

¹⁰ <https://almascience.eso.org/observing/observing-configuration-schedule>

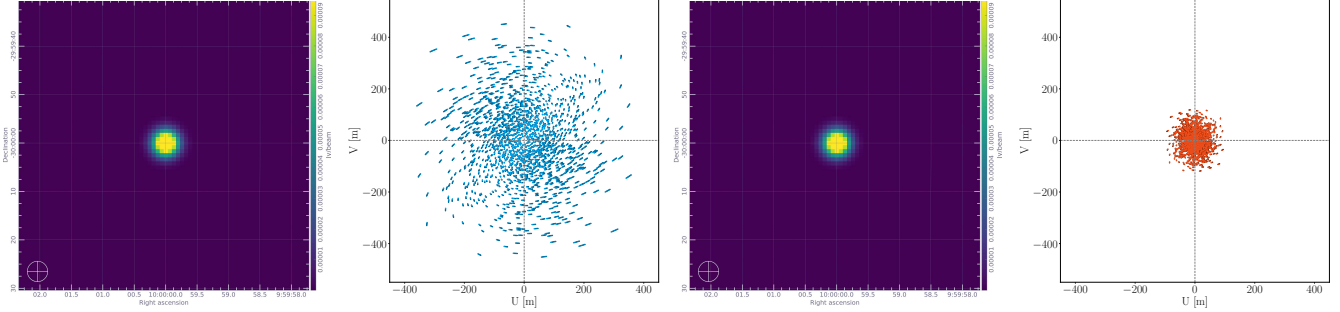


Figure 2. Example of the image and uv -plane coverage in units of meters for one of the uncorrupted simulated sources at redshift 2.9, observed with ALMA configuration 4. The left panels (1 and 2) shows the image and uv -plane coverage in the observed frame, while the right panels (3 and 4) shows the resulting image after rest-framing.

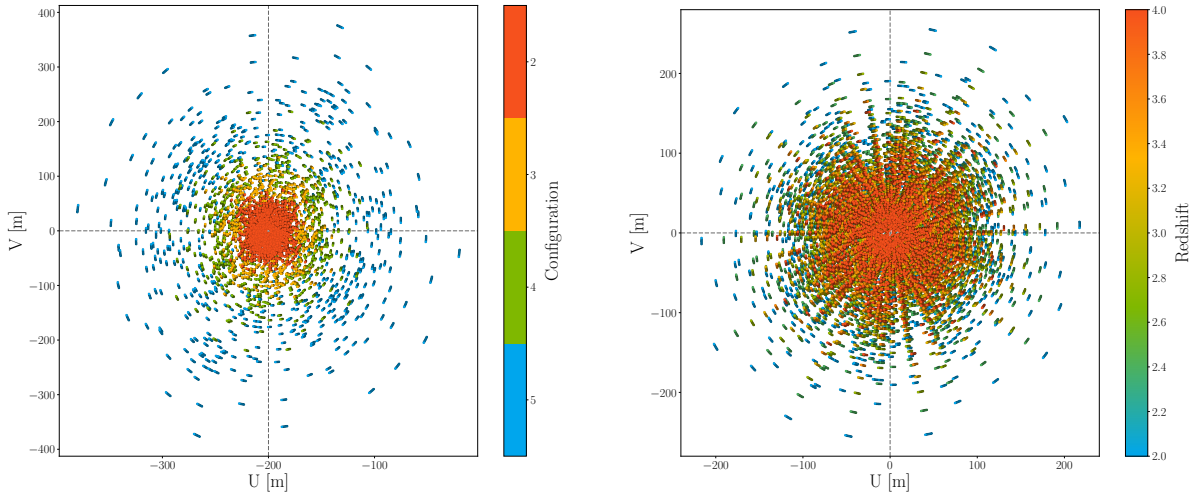


Figure 3. uv -plane coverage of the stacked simulated source. Visibilities from different array configurations have been centered on the phase center and shifted to the rest frame using the `ViSta` method applied to the ALMA data. The left panel shows the rest-framed visibilities position in the uv -plane for all the 50 mock observation, color-coded according to the observing configuration. The right panel shows the difference in uv -plane coverage for the same configuration (5 sources at configuration 4, with max baseline at 780 m) observed at different redshifts.

combination of a wide range of redshifts and multiple array configurations contributes to a denser and more homogeneous uv -plane coverage.

In the next section, we will present the results separately for unresolved (point-like) and resolved Gaussian sources. We will apply the `ViSta` method to fit the spectral profile of the sources, performing fits both in the visibility domain, which we will refer to as `ViSta-uvFit`, and in the image domain after imaging the stacked data, referred to as `ViSta-ImFit`. These results will then be compared with those obtained from the classical image-plane stacking procedure, which we will call `ImSta-ImFit`.

5.2. Results

5.2.1. Unresolved Source

We analyse a sample of 50 simulated, perfectly point-like sources with the `ViSta` method. We begin with centering and the rest-framing of these sources (as discussed above) and then

we proceed with the spectral rebinning and regridding onto a common frequency axis. The new channel width is determined by the source in the sample with the largest intrinsic frequency dispersion, which, since all mock sources initially share the same rest-frame properties, corresponds to the source at the highest redshift. We have performed a fit of the stacked data both in the uv -plane and in the image plane.

For the latter case, i.e., `ViSta-ImFit`, we follow the procedure described in 2: we do not specify a restoring beam explicitly but we adopt the default one, i.e., a Gaussian fitted to the main lobe of the PSF resulting from the relatively uniform uv -coverage of the combined dataset. Furthermore, we select natural weighting to enhance sensitivity at the expense of resolution; this choice is particularly suitable for sources that are essentially point-like at the observed resolution, similar to our case. Like we explained in 4.1, for the case of point-like sources, we considered the peak value in the image as a measure of the flux density source, without integrating over an extended region.

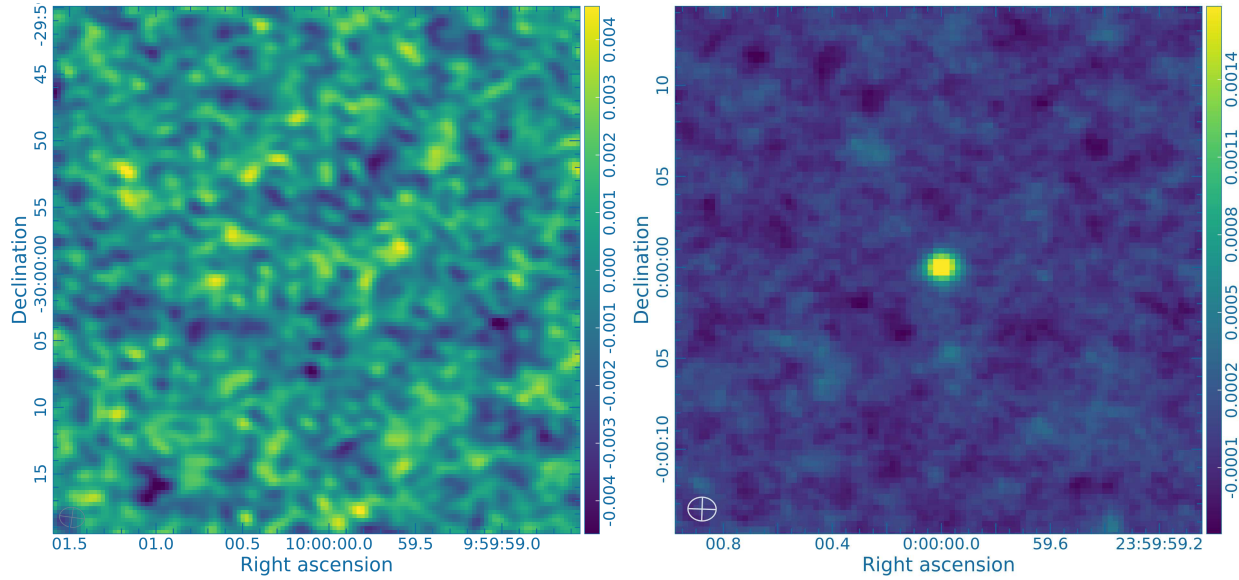


Figure 4. Comparison between the peak channel image of a single mock source before stacking, where the source signal is below the noise level (left panel), and the result after stacking 50 mock sources using the ViSta method (right panel).

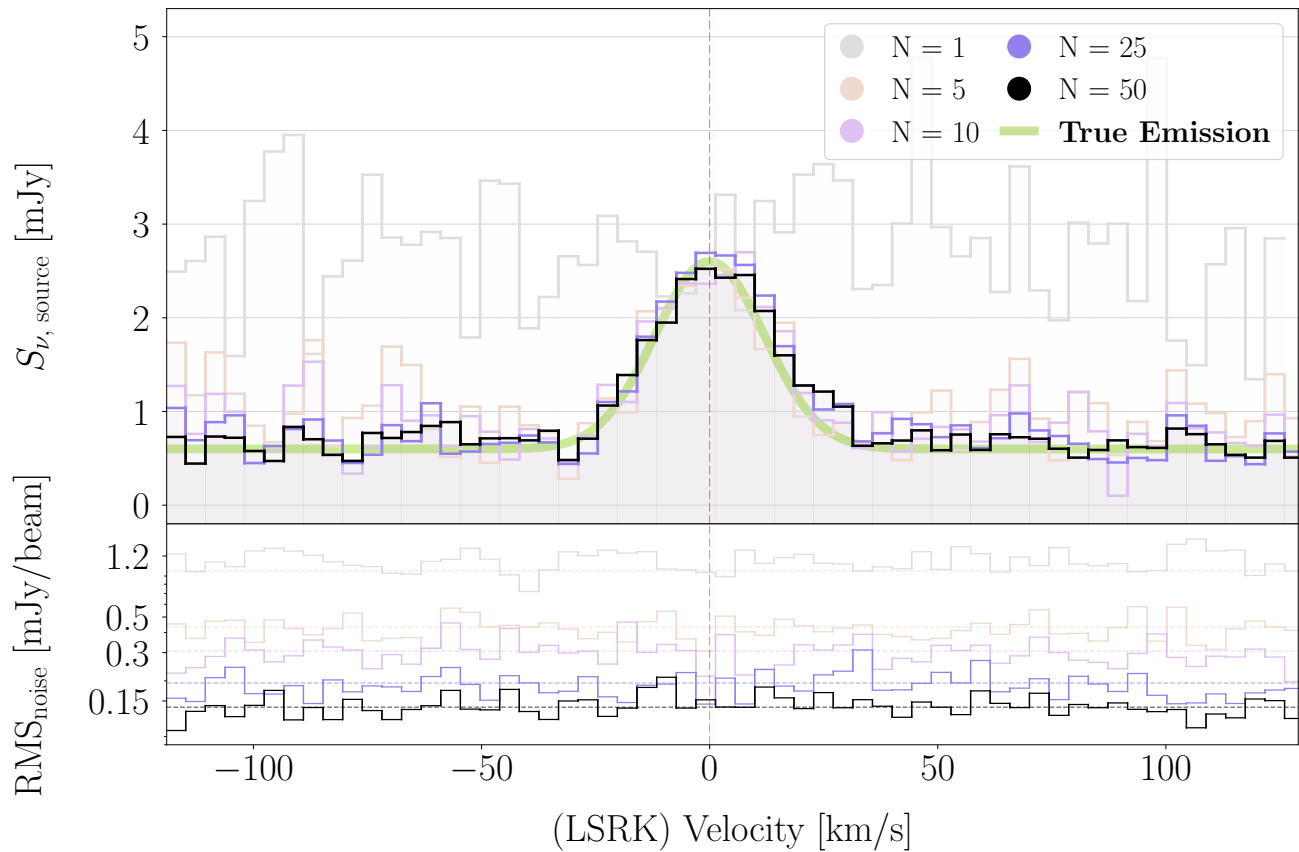


Figure 5. Top panel: Spectral line profiles recovered using the ViSta-ImFit method for different numbers of stacked point-like mock sources ($N = 1, 5, 10, 25, 50$), compared to the true emission profile (green solid line). Bottom panel: measured rms noise of the corresponding stacked spectra as a function of N . The dashed line represents the expected $1/\sqrt{N}$ scaling, corresponding to the theoretical noise reduction with ideal stacking. The agreement between the measured and expected values confirms the statistical robustness of the method.

Table 1
Peak SNR Results for ViSta-ImFit

N_{Source}	S_{peak} (mJy beam ⁻¹)	σ_{chan} (mJy beam ⁻¹)	Peak SNR
1	3.3125	1.1191	2.9600
5	2.5168	0.4240	5.9358
10	2.4579	0.3009	8.1695
25	2.6655	0.1876	14.2058
50	2.4283	0.1404	17.3007

Note. Signal-to-noise ratio (SNR) results for the ViSta-ImFit procedure for the central channel. S_{peak} is the peak flux, σ_{chan} is the average channel rms.

In Figure 4, we show the comparison between the image of a single mock source, where the signal is below the noise level, and the result obtained after stacking all 50 mock sources using the ViSta method. As can be seen, while the individual source is not detectable, the stacking procedure clearly amplifies the signal. Figure 5 further illustrates how the signal-to-noise ratio improves as a function of the number of stacked sources, and the corresponding quantitative measurements for the central channel are reported in Table 1. The advantage of the ViSta method increases with the number of stacked sources. As expected, the rms noise decreases approximately as $1/\sqrt{N}$ (the trend illustrated by the dashed line in Figure 5). In our case, stacking 50 datasets reduces the average channel rms σ_{chan}^* (initially ~ 1.2 mJy for the reference image) by a factor ~ 8 consistent with this scaling, giving a final peak SNR of approximately 17.3 on stacking.

We then performed a spectral fitting directly in the visibility domain using `uvm.uvmultifit`, adopting a Gaussian model to provide greater flexibility in the source representation. To assess the consistency of this method, we also compared it with the standard image-stacking approach `ImSta-ImFit`. For each target in the sample, we reconstructed the image using its original observational parameters, i.e., synthesized beam, sensitivity and field of view. Following a procedure analogous to the ViSta-ImFit, we extracted the flux density from the central pixel in each spectral channel. The final stacked spectrum was then obtained by averaging the flux values across all sources for each channel.

Figure 6 presents the stacked spectra obtained using the various methods described above. The green line represents the expected average emission, derived from averaging all uncorrupted observations. The black diamonds indicate the best-fit flux densities from the ViSta-uvFit procedure, with corresponding uncertainties, while the black dashed line represents the Gaussian model fitted to these points. The fit was obtained through a Monte Carlo approach, by perturbing the data according to their uncertainties and averaging the resulting fits. The magenta and blue lines correspond to the spectral fits obtained using the ViSta-ImFit and ImSta-ImFit

methods, respectively, with the dashed lines in matching colors indicating the corresponding Gaussian fits. All three methods successfully reproduce the Gaussian line profile, recovering the true emission with about 90% accuracy and showing only minor differences among them. However, they appear to slightly overestimate the continuum level. This discrepancy may arise because the continuum is significantly weaker and comparable to the average noise level, making it more difficult to constrain reliably. These results indicate that all three approaches are robust and fully reliable for the analysis of unresolved sources.

5.2.2. Resolved Source

For resolved or partially resolved sources, the stacking procedure remains conceptually identical to that adopted for point-like sources; the main difference lies in how the flux density is measured. In this regime, the peak pixel value no longer provides a reliable estimate of the total emission, and the flux must instead be integrated over a spatially extended region. To explore this, we present an analysis based on three different scenarios, always comparing our method to the standard ViSta-ImFit stacking approach. In the first case, we assume that the intrinsic size of the sources is known, allowing us to measure the flux density directly within the corresponding area. In the second case, we assume no prior knowledge of the source size, and instead fit each detection with a 2D Gaussian model using `castasks.imfit()`, from which we derive both the morphology and the integrated flux. In the third case, where the individual sources are not suitable for reliable fitting, we convolve all images to a common (i.e., lowest) spatial resolution to ensure consistency. The stacking is then performed on the convolved images, followed by a fit on the stacked result.

Native Source Size at Native Resolution. The first test focuses on the ideal case in which the intrinsic size of the sources is well known. This option suits our simulated case well, where all sources share the same spatial extent, characterized by a FWHM of $5''$. This allows us to measure the flux density directly within the appropriate region, using a circular aperture with a radius of 3σ (where σ is the standard deviation of the Gaussian profile) to encompass the majority of the source emission while minimizing contamination from the background noise. For the ViSta-ImFit, after the classic pre-processing and stacking procedure, we performed the imaging step as per standard procedure and we measured the flux density for every channel of the stacked cube within the known source extension. For the ViSta-uvFit, we use the Gaussian model within the fixed region, letting only the embedded flux change. Finally, for the image stacking ImSta-ImFit, we generate every image for every mock observation like for the point-like scenario, we extract for each source the flux density embedded in the same region and we then average the obtained values along the spectral axis.

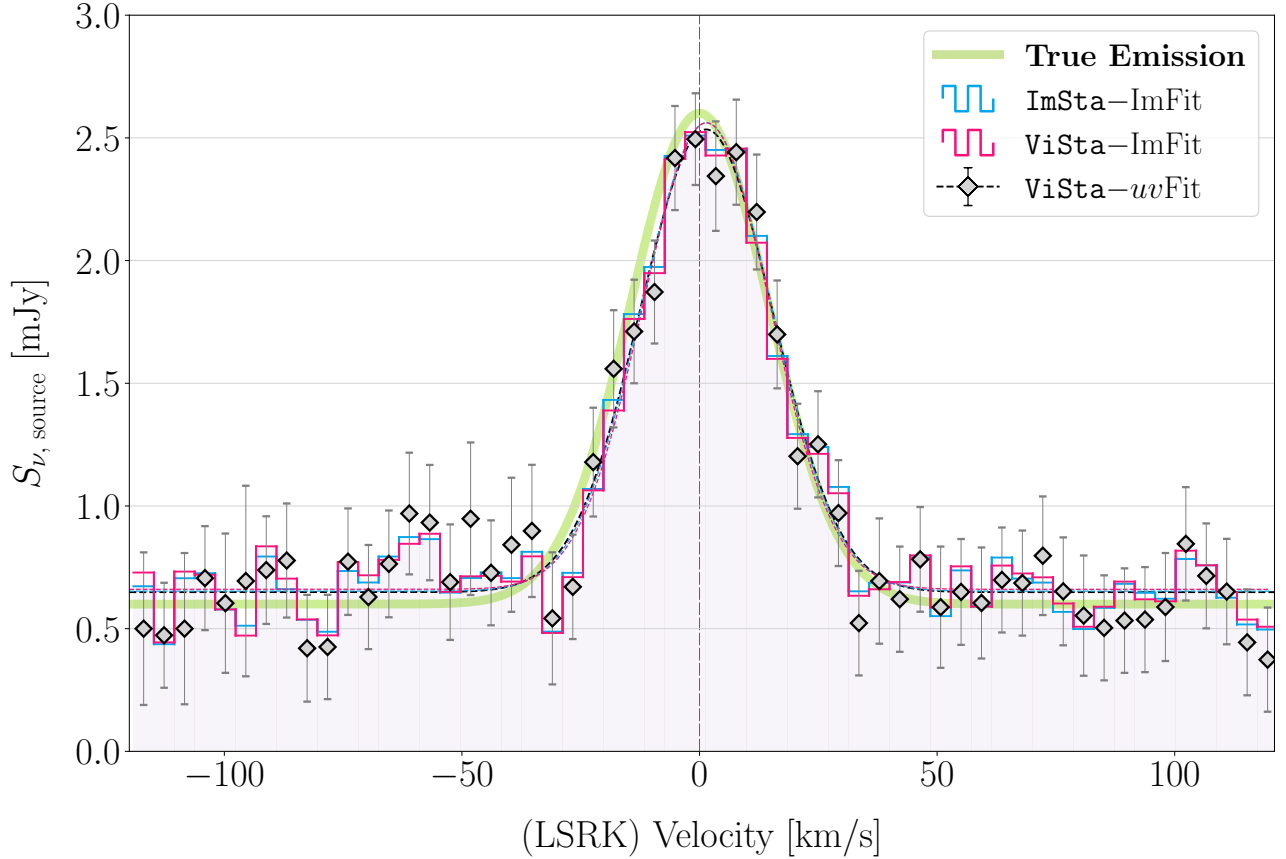


Figure 6. Comparison of the recovered spectral line profiles obtained using different stacking and fitting methods on simulated data for unresolved sources. The green solid line represents the true emission model used in the simulation. The blue and the magenta lines show respectively the results from the `ImSta-ImFit` and `ViSta-uvFit` procedures, and the dashed lines represent the Gaussian fits to the data points. The diamonds represent the data points fitted in the `uvFit` method with 1σ uncertainties; for the `ViSta-uvFit` case, where error bars were available, the Gaussian fit was performed via Monte Carlo sampling: synthetic realizations of the data were generated within the uncertainties, and the final fit (black dashed line) represents the average result. The plot shows that the three methods yield consistent results and recover the signal reasonably well.

In this context, we also investigated how the performance of the `ViSta-ImFit` method evolves with the number of individually extracted sources included in the stack and we present the results in Figure 7. Similar to the point-source case, the noise decreases as expected, scaling approximately with the square root of the number of sources. However, we find that recovering a well-defined spectral line profile requires a significantly larger sample. This is because, when integrating over an extended area, a greater amount of noise is also accumulated, reducing the efficiency with which the spectral signal emerges. As a result, although the method remains effective, the convergence toward the intrinsic line profile is slower than in the unresolved case, and a larger number of sources is needed to achieve comparable spectral quality.

The stacked spectra resulting from the different methods are shown in Figure 8, using the same color map and linestyles used for the point-source case. All three approaches recover the line emission with good fidelity. The direct visibility-plane extraction via `ViSta-uvFit` provides the closest match to the

expected profile, whereas `ViSta-ImFit` tends to slightly overestimate the Gaussian peak intensity. This problem can be attributed to the heterogeneous beam configurations of the individual datasets. When these observations are concatenated, the resulting effective beam differs subtly from each input synthesized beam, which can concentrate flux into a narrower core and artificially boost the measured peak. Consequently, slight variations in beam size and shape across the sample propagate into the stacked spectrum, leading to the observed bias of the stacked source size. Finally, the spectra derived from `ImSta-ImFit` exhibit larger point-to-point scatter, likely a consequence of including more noise in the integration region.

Estimated Source Size at Native Resolution. We repeat the analysis presented above without assuming prior knowledge of the source size. In the `ViSta-uvFit` approach, to obtain a more accurate estimate of the source geometry, we first performed a fit on the data averaged over all channels. The best-fit size parameters (major, minor axes and orientation of

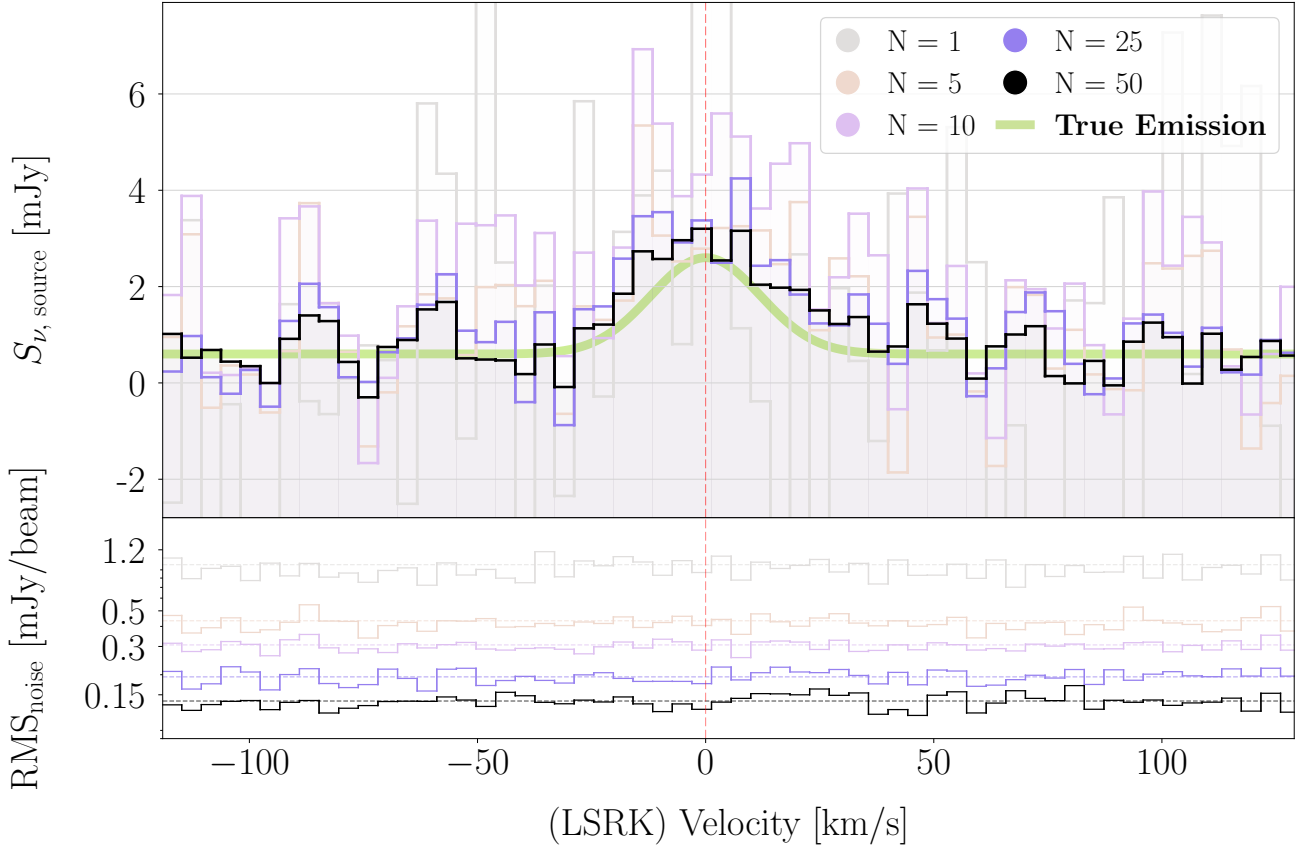


Figure 7. Top panel: Spectral line profiles recovered using the `ViSta-ImFit` method for different numbers of stacked mock extended sources ($N = 1, 5, 10, 25, 50$), compared to the true emission profile (green solid line). Bottom panel: measured rms noise of the corresponding stacked spectra as a function of N . The dashed line represents the expected $1/\sqrt{N}$ scaling, corresponding to the theoretical noise reduction with ideal stacking. The agreement between the measured and expected values confirms the statistical robustness of the method.

the spatial Gaussian) were then kept fixed and used to estimate the flux density independently in each individual channel. While fitting in the image plane, on the other hand, we use CASA’s `imfit` task: it performs a two-dimensional Gaussian fit to a user-defined region of emission, deconvolving the trial model from the synthesized beam to recover the intrinsic source morphology. `imfit` returns several best-fit parameters: integrated flux density, centroid coordinates, deconvolved major and minor axes, and position angle. For the `ViSta-ImFit` workflow, we first perform the `uv`-stacking and then compute the zero-moment map of the stacked visibilities. We then fitted the 2D Gaussian model to the moment-0 image, yielding the deconvolved source size. In contrast, in the (`ImSta-ImFit`) approach, we begin by generating a moment-0 map for each individual cube. We then fit each map with a 2D Gaussian, extract the integrated flux density within the fitted aperture for every source, and finally average these per-source fluxes across the entire sample on a channel-by-channel basis to assemble the stacked spectrum. For both methods, the flux density was extracted within a radius of 3σ from the fitted Gaussian. The results are presented in Figure 9. They highlight

the robustness of the `ViSta` method in recovering both continuum and line emission, whether the fitting is performed in the `uv` plane or in the image plane. This is largely due to its ability to accurately resolve the source size, even when the emission is faint and spatially extended, allowing for a clean flux measurement that is minimally affected by noise. On the contrary, the `ImSta-ImFit` approach struggles to precisely determine the intrinsic source size of single sources, even in the continuum map. This leads to inaccurate flux estimates for the line emission and results in a stacked spectrum that shows a lower peak emission and significantly larger fluctuations.

Estimated Source Size at Lowest Resolution. Finally, we repeated the analysis on the full sample after degrading all data to the lowest common angular resolution. In this configuration, both the integrated flux density and the noise level (now uniform across all images in the dataset) were estimated using the (`ViSta-ImFit`) and (`ImSta-ImFit`) approaches. Figure 10 shows the results of both image-plane fitting methods applied to the two stacking strategies at the lowest common resolution. In this case, stacking directly in the `uv` domain results in a more uniform filling of the Fourier plane

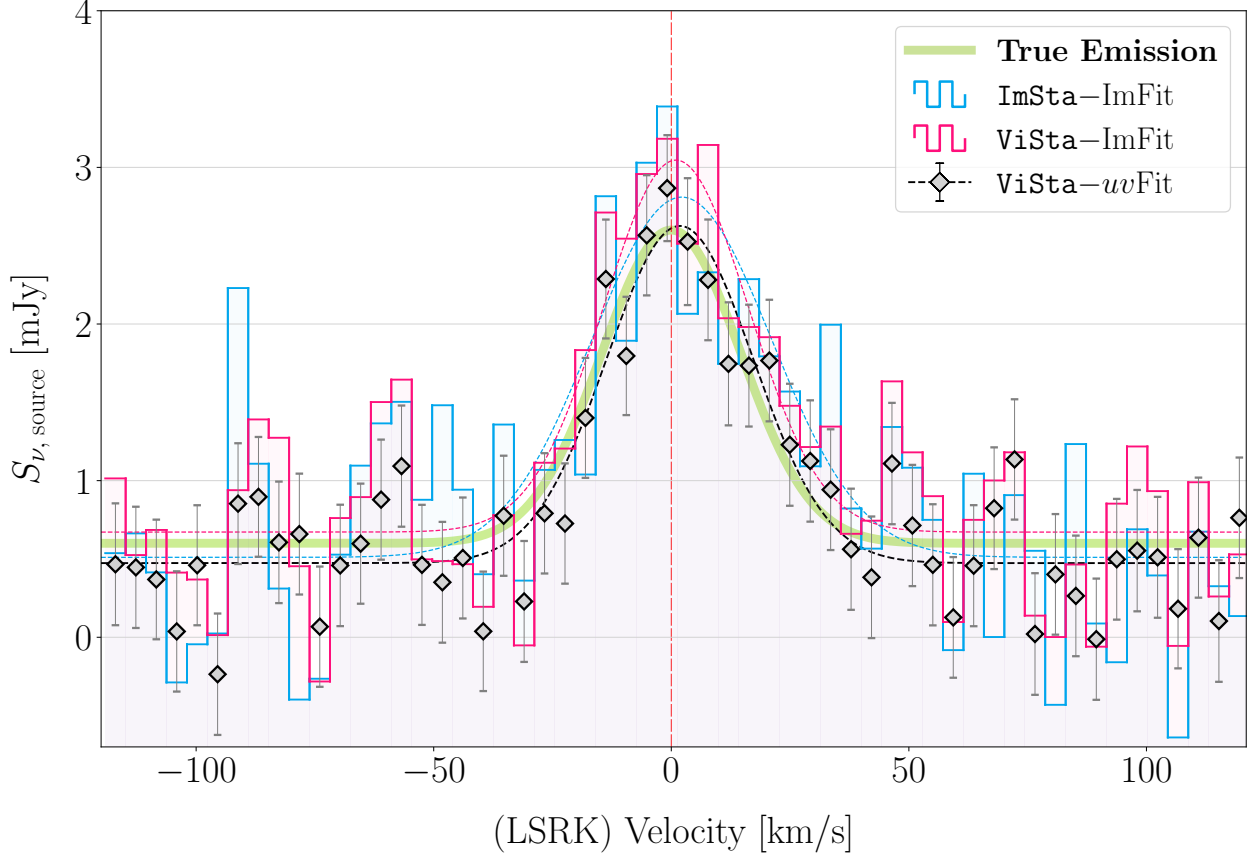


Figure 8. Comparison of the recovered spectral line profiles obtained using different stacking and fitting methods on simulated data for a resolved source, for known source sizes. The green solid line represents the true emission model used in the simulation. The blue and magenta lines show respectively the results from the ImSta-ImFit and ViSta-ImFit procedures, and the dashed lines represent the Gaussian fits to the data points. The diamonds represent the data points fitted in the ViSta-uvFit method with 1σ uncertainties and the dashed lines is the Gaussian fit performed via Monte Carlo sampling.

and naturally down-weights the noisiest baselines, yielding a higher-fidelity reconstruction of the source core. In contrast, convolving each individual image to the lowest resolution, each with relatively low SNR, introduces additional noise into stacking, leading to a broader scatter in the measured spectrum, a reduced peak amplitude, and a slightly elevated noise. Thus, after resolution matching, when the SNR is low in the single observation of the sample, the ViSta method consistently better captures the true source profile. Figure 11 shows the convolved stacked image with both methods for the central channel of the cube. In addition, it also demonstrates that the deconvolved source size recovered in both the ViSta and ImSta methods closely matches the true input value. However, the image-plane extraction exhibits a more pronounced noise pattern, with an average rms for the central channel of approximately $\sim 0.285 \text{ mJy beam}^{-1}$, whereas the one stacked in the visibility plane yields a smoother background, with an rms of $\sim 0.22 \text{ mJy beam}^{-1}$. Considering the flux densities obtained with these two methods, respectively 2.070 and 2.566 mJy, the ViSta method yields a gain of $\sim 54\%$ in the central channel compared to the image-plane stacking.

Continuum extraction at lowest resolution. We also tested the use of the ViSta method to extract only the continuum emission and to quantify the potential gain provided by this approach. In this case, we produced the continuum image from the ViSta-stacked measurement set, using only the channels identified as free from line emission. This approach was feasible because, in our simulated dataset, the line-free channels were known a priori. If applied to a realistic dataset, the continuum should instead be obtained by subtracting the line emission following the procedure described in Section 3.6.

The continuum analysis was carried out following the procedure described in the preceding paragraph, homogenizing all data to the lowest angular resolution of the dataset. This ensures a realistic and coherent noise estimation for both methods, facilitating a fair comparison and enabling the quantification of the resulting improvement in SNR. While for the line stacking we generated the spectral cubes with TCLEAN using the `spemode='cube'` parameter, for the continuum we used `spemode='mfs'` (*multi-frequency synthesis*) and selected only the channels containing line-free continuum emission. The results are presented in Figure 12. For both

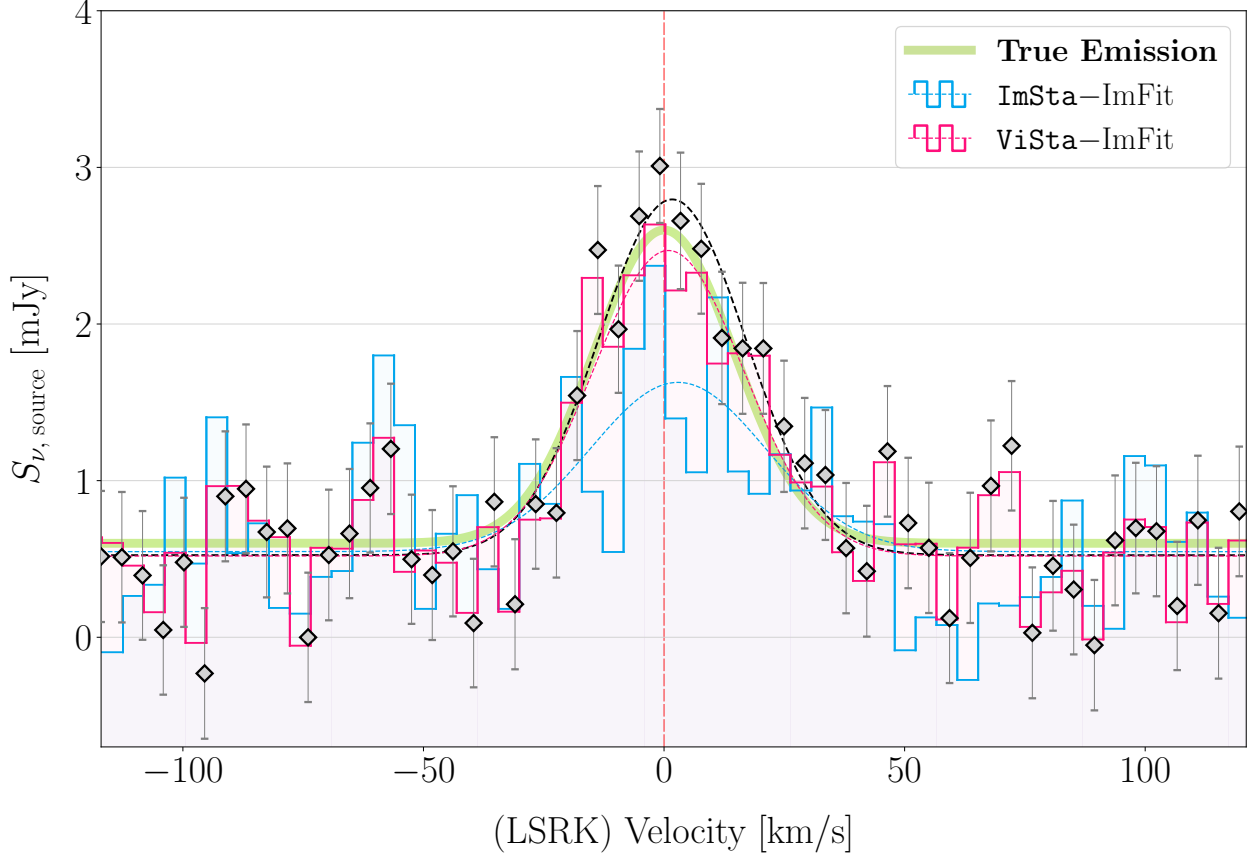


Figure 9. Comparison of the recovered spectral line profiles obtained using different stacking and fitting methods on simulated data for a resolved source, derived after fitting the size of the source. The green solid line represents the true emission model used in the simulation. The blue and the magenta lines show respectively the results from the `ImSta-ImFit` and `ViSta-ImFit` procedures, and the dashed lines represent the Gaussian fits to the data points. The diamonds represent the data points fitted in the `ViSta-uvFit` method with 1σ uncertainties and the dashed lines is the Gaussian fit performed via Monte Carlo sampling.

approaches, we measured the peak flux within a beam size region centered on the source. We note, however, that the peak flux does not represent the total flux density of the source, as the new restoring beam is comparable in size to the source itself and not significantly larger, as required for a point-like case. Nevertheless, it ensures that the SNR difference is driven solely by the gain of the two methods rather than by uncertainties in size estimation. Using the `ImSta-ImFit` method, the peak flux is $0.404 \text{ mJy beam}^{-1}$ with an rms noise of $0.058 \text{ mJy beam}^{-1}$, giving a SNR of 6.966. With `ViSta-ImFit`, the peak flux is $0.419 \text{ mJy beam}^{-1}$ and the rms noise $0.051 \text{ mJy beam}^{-1}$, resulting in a SNR of 8.216. This represents a $\sim 18\%$ increase in peak SNR, consistent with the 1.2 gain reported in recent ALMA archival combination analysis which compares uv -plane and image-plane stacking (Hill et al. 2024).

6. Application to a Real Case

After validating the method on the simulated case, we now apply the same analysis to a sample of galaxies observed with

ALMA, targeting the CO(3–2) line emission. The sample consists of 21 partially resolved or unresolved galaxies, four of which had two available observations, in the redshift range of $z \sim 1.407\text{--}2.73$ and angular resolutions between $0''.25$ and $2''.8$. For each source, we verified the accuracy of its coordinates and ensured that every observation was free from instrumental issues or artifacts.

The selected sources are part of the sample analyzed in the BEARS I, II, and III studies (Bendo et al. 2022; Urquhart et al. 2022; Hagimoto et al. 2023), which provided spectroscopic redshifts through ALMA spectral scans, multi-frequency continuum observations to constrain the source spectral energy distributions (SEDs), and molecular line detections to characterize the ISM properties of these galaxies. Detailed information for each source, including coordinates, resolution, and spectroscopic redshift, is provided in Table 2, along with the peak SNR value in the case of a spectral line detection.

As we can estimate an approximate source size from the zero order moment maps for our sample, we used the same approach as the simulations (see Section 5.2.2.). However, for

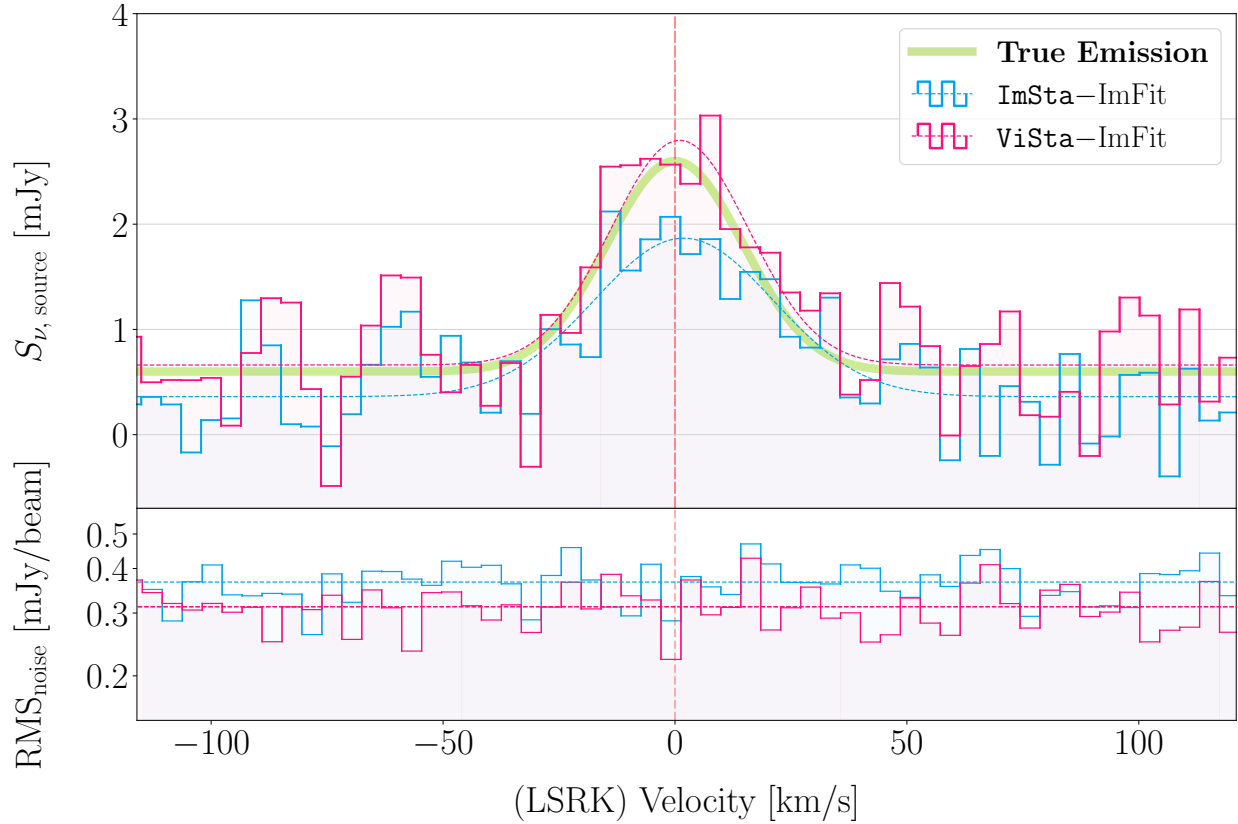


Figure 10. Comparison of the recovered spectral line profiles obtained using different stacking and fitting methods on simulated data for resolved source, after convolving all sample to the lowest resolution. The green solid line represents the true emission model used in the simulation. The blue and the magenta lines show respectively the results from the `ImSta-ImFit` and `ViSta-ImFit` procedures, and the dashed lines represent the Gaussian fits to the data points.

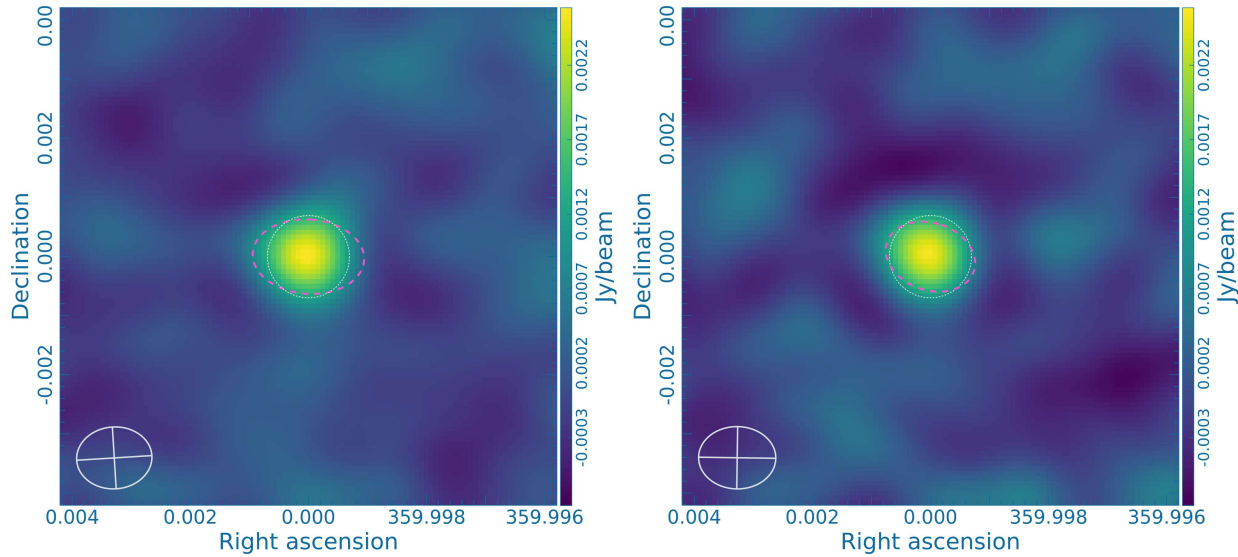


Figure 11. Stacked images of the central channel of the cube using two different convolution methods. Left: Image stacked in the image plane using the `ViSta` method, then convolved with a largest beam of the sample. Right: Image stacked in the image plane after convolving all input images to the worst resolution. The dashed purple circle indicates the source size as fitted with `casatasks.imfit()`, while the white dotted circle shows the effective size of the stacked source. The intrinsic source size is reliably recovered by both methods, but visibility-plane stacking yields a smoother background.

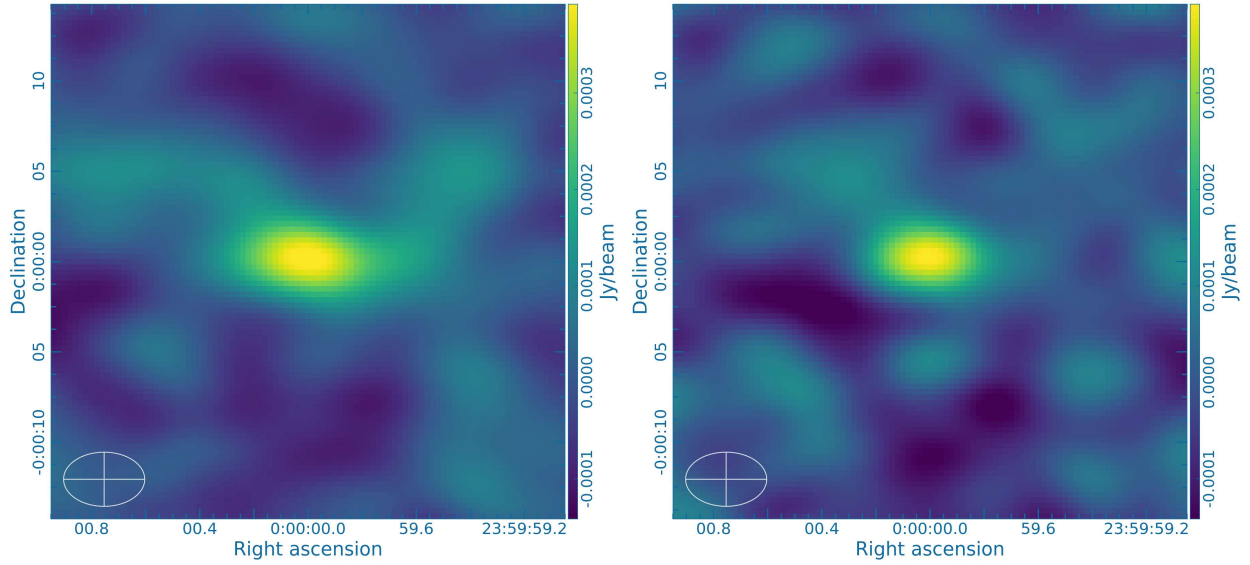


Figure 12. Stacked images of the continuum emission using two different convolution methods. Left: Continuum image stacked in the uv -plane using the `ViSTA` method, then convolved with a largest beam of the sample. Right: Continuum image stacked in the image plane after convolving all input images to the worst resolution. The visibility-plane stacking yields a smoother background and captures a larger fraction of the source emission.

Table 2
Source Properties

Source	z_{Spec}	R.A.	Decl.	D_L factor	Ang Res ($''$)	Peak SNR
HerBS-11	2.631	1:24:07.518	-28:14:34.84	0.460	0.317	1.500
HerBS-24	2.198	0:47:36.096	-27:29:51.943	0.370	2.445	1.106
HerBS-37	2.619	23:26:23.137	-34:26:43.993	0.457	0.361	...
HerBS-45	2.434	0:51:32.943	-30:18:49.566	0.418	2.366	...
HerBS-63	2.432	0:51:31.700	-30:20:20.383	0.418	2.367	0.626
HerBS-77	2.228	0:56:29.253	-31:12:07.431	0.376	2.459	0.741
HerBS-80	2.231	23:00:02.546	-31:50:08.808	0.376	2.505	...
HerBS-86	2.564	23:53:24.571	-33:11:11.766	0.446	2.661	1.008
HerBS-93	2.400	23:47:50.447	-35:29:29.932	0.411	0.250	...
HerBS-107	2.553	1:45:20.060	-31:38:32.954	0.443	2.731	...
HerBS-135	2.401	22:56:11.799	-32:56:51.802	0.412	0.348	...
HerBS-138	1.407	1:17:30.439	-32:07:22.445	0.214	1.663	1.262
HerBS-145	2.730	1:23:34.643	-31:46:23.483	0.481	0.319	...
...	2.744	...
HerBS-159	2.236	23:51:21.740	-33:29:00.400	0.377	2.669	1.756
HerBS-178A	2.658	1:18:50.254	-28:36:43.941	0.466	0.318	...
HerBS-178B	2.655	1:18:50.127	-28:36:40.863	0.465	0.318	...
HerBS-182	2.227	23:05:38.804	-31:22:05.351	0.376	2.471	...
HerBS-184	2.507	23:49:55.665	-33:08:34.370	0.434	0.253	...
...	2.544	1.298
HerBS-200	2.151	1:43:13.303	-33:26:33.125	0.360	2.725	...
HerBS-208	2.478	22:57:44.560	-32:42:33.566	0.428	1.527	...
—	0.350	...
HerBS-209	2.272	22:49:20.518	-33:29:41.137	0.385	2.508	...
—	2.802	...

Note. List of sources with redshift, coordinates, luminosity distance factor D_L , angular resolution of the observation and Peak SNR (in case of detection) included in the real sample described in Section 6. When multiple observations exist for the same source, only the angular resolution is shown for secondary entries.

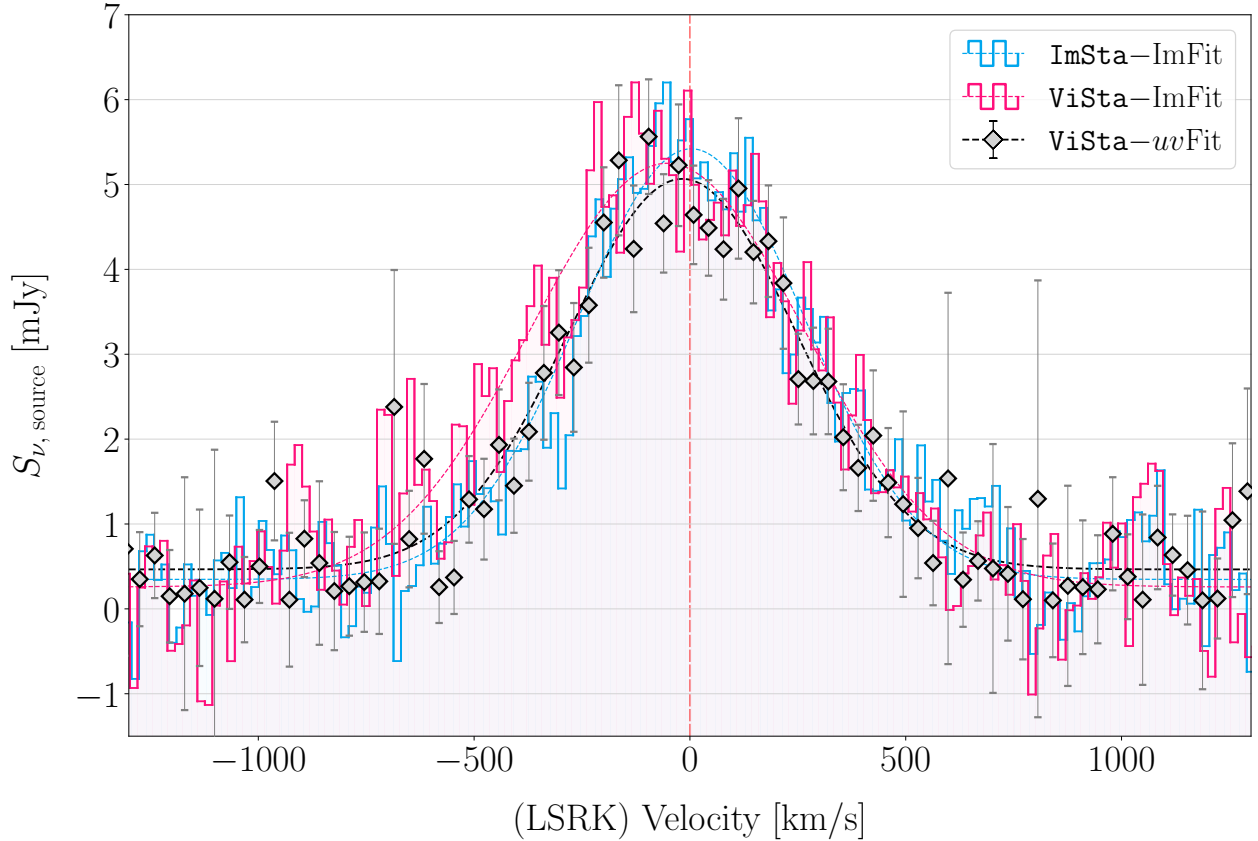


Figure 13. Stacked CO (3–2) spectra obtained through the different methods on the real sample data. The black diamonds show the best-fit results from the `UVMultiFit` analysis with 1σ uncertainties, and the blue line indicates the overall spectral fit to those data points. The blue and the magenta lines show respectively the results from the `ImSta-ImFit` and `ViSta-ImFit` procedures, and the dashed lines represent the Gaussian fits to the data points.

Table 3
Gaussian Fit Parameters for the Real Sample

Case	μ (km s^{-1})	σ (km s^{-1})	A (mJy)	C (mJy)
<code>ViSta-uvFit</code>	-16.8184	264.9165	4.5835	0.4636
<code>ViSta-ImFit</code>	-52.6187	317.8469	4.9921	0.2579
<code>ImSta-ImFit</code>	3.3452	262.5233	5.0738	0.3467

Note. Summary of Gaussian fit parameters (mean, sigma, amplitude, shift) with uncertainties omitted for clarity. The values correspond to three different stacking methods applied to the real data sample.

the `ViSta-uvFit` case, we deliberately chose not to estimate the source size from the moment-zero map. We instead kept the spatial parameters of the Gaussian model free in each channel. This choice allows for greater flexibility, as line properties, such as velocity dispersion and extent, are not known a priori and may vary across the profile. Moreover, we did not subtract the continuum to avoid biasing the line emission. We applied a weighting scheme to account for varying frequency coverage across observations and to avoid multiple counting sources with duplicate observations, ensuring that each source contributes

appropriately to the final stacked dataset. However, we did apply a correction for luminosity distance differences, adopting a reference redshift of 2.5 for consistent comparison with previous stacking analyses of high- z DSFGs (Spilker et al. 2014; Hagimoto et al. 2023).

The results are presented in Figure 13. The various fitting methods appear to be reasonably consistent, as summarized in Table 3. The velocity dispersion and the amplitude of the spectral line show good agreement across methods, with differences of approximately 10% in both cases. The continuum level shows the largest relative variation, which is expected, given that the continuum is generally detected at a lower SNR than the line itself in most of our sample. In contrast, the position of the peak is consistent within 3 spectral channels across methods, which is a reasonable level of agreement considering the typical line widths and the possibility of minor asymmetries or extended emission that may affect the line differently. The resulting line properties are broadly consistent with other estimates in the literature (Spilker et al. 2014; Hagimoto et al. 2023). However, the result presented in this paper should not be interpreted as a precise measurement of the line emission, as we only applied a

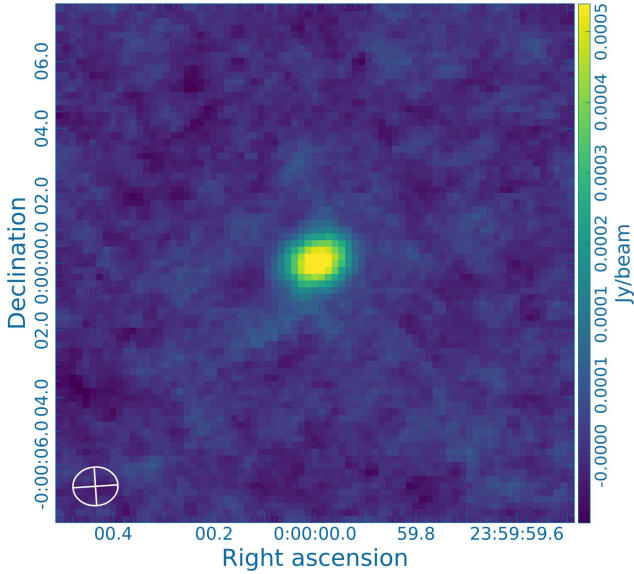


Figure 14. Stacked CO (3–2) continuum extracted as a post-processing step of the `ViSta` method after the concatenation of the visibilities. We estimate a continuum average flux density of 0.832 ± 0.064 mJy.

correction for the luminosity distance but did not account for other factors such as flux normalization or lensing magnification. This application to the real dataset is intended primarily as a test and comparison of the `ViSta` method with respect to both the simulated case and the classical stacking methods, rather than as a definitive estimate of the total line flux, which will be addressed in a future study.

As an additional test, we explored estimating the continuum as a post-processing step of the `ViSta` method. From the resulting continuum image, shown in Figure 14, we estimate a continuum average flux density of 0.832 ± 0.064 mJy.

In the case presented, all sources were detected from the first integration, and for nearly all targets the spectral line was clearly detected. This allowed for a robust characterization of the line parameters, which naturally explains why both stacking approaches produce consistent results.

A more stringent observational test of the capability of the `ViSta` method can be performed by artificially degrading the sensitivity of individual observations. This can be achieved by reducing the number of integration intervals, and thus the uv -coverage, which lowers the overall data quality. When a source is no longer individually detectable, we expect the `ViSta` method to recover a larger fraction of the intrinsic signal and achieve a higher SNR compared to conventional image stacking, as we demonstrated in the simulated case. To explore this, we selected one fourth of the available data for each observation and repeated the analysis, thereby testing the method under conditions of reduced sensitivity and sparser uv -sampling. In the first phase, we compared the recovered line flux between the classical stacking method and `ViSta` at native

resolution. To determine the area over which to measure the flux density, we referred to the same moment-zero maps generated from the higher-sensitivity observations, as done in the previous case. The results, shown in Figure 15, indicate that when the sensitivity of the observations is reduced and the line signal is increasingly embedded in the noise, stacking in the image plane produces a source that appears significantly more scattered and diffuse compared to one in the uv -plane. Thus, under conditions approaching the detection limit, `ViSta` seems to be more effective in capturing the line emission.

However, to obtain a consistent estimate of the actual SNR improvement achieved by stacking in the visibility domain, it is necessary to evaluate the noise coherently for both the visibility and image-plane stacking. To do so, we refer to the procedure described in Section 5.2.2. and convolve all datasets to the worst angular resolution available, corresponding to a synthesized beam of $5.063'' \times 2.820''$. Figure 16 illustrates the obtained results: the top panels show the spectra extracted using the `ViSta-ImFit` and `ImSta-ImFit` methods for both the high- and low-sensitivity datasets, while the bottom panels display the corresponding rms noise.

When considering the datasets at their original sensitivity, stacking in the uv -plane yields a more effective suppression of noise, consistent with the behaviour observed in the simulated tests. Quantitatively, the SNR increases from 7.837 for `ImSta-ImFit` to 10.924 for `ViSta-ImFit`, corresponding to an improvement of approximately 39%. When the data sensitivity is reduced, the SNR increases from 3.41 to 5.24, resulting in a relative gain of about 53%. These results confirm our earlier hypothesis that visibility-based stacking becomes progressively more advantageous as the data approach the noise-dominated regime, owing to its ability to preserve the phase coherence of the astronomical signal while mitigating the incoherent addition of noise inherent to image-plane stacking.

7. Summary and Outlook

In this paper, we have introduced `ViSta`, a novel technique for stacking radio-to-submm data directly in the uv -domain. This method enables the combination of observations of galaxies at different redshifts, obtained with various telescopes and array configurations, into a unified analysis framework.

We have outlined the advantages of performing stacking in the visibility plane according to the `ViSta` method, presented the related workflow pipeline, and demonstrated its efficiency on simulated and real ALMA observations for the study of the CO(3–2) average line emission on a sample of dusty star-forming galaxies.

The main results of our analysis are listed as follows:

1. Overall, the `ViSta` method performs correctly and reliably when tested with the simulated dataset. While for point-like sources its performance is comparable to

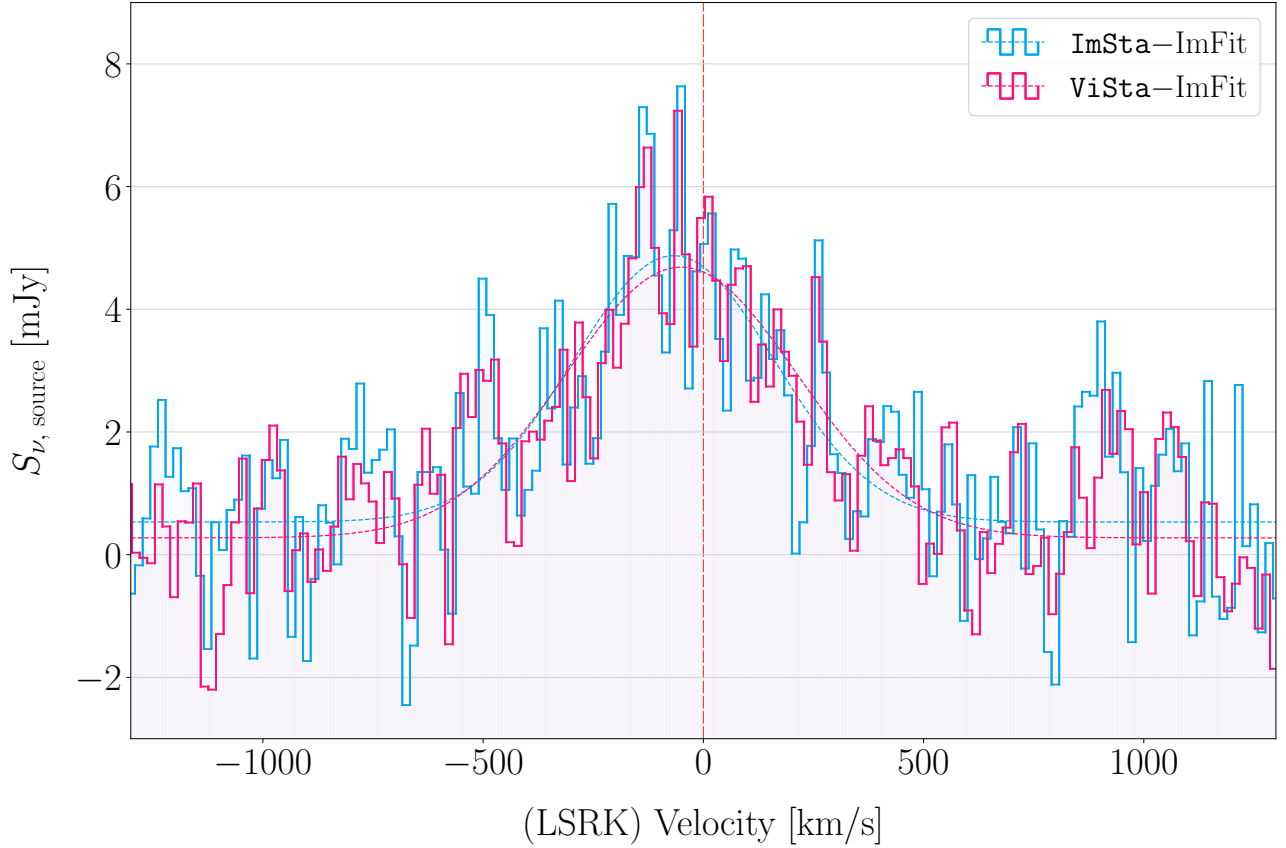


Figure 15. Stacked CO (3–2) spectra obtained through *ImSta-ImFit* and *ViSta-ImFit* procedures after convolving to the worst resolution of the dataset. The blue and the magenta lines show respectively the results from the *ImSta-ImFit* and *ViSta-ImFit* procedures, and the dashed lines represent the Gaussian fits to the data point. The spectrum derived from the *ImSta-ImFit* procedure appears noticeably more scattered and exhibits a higher level of noise fluctuations, whereas the *ViSta-ImFit* result shows a smoother and more coherent line profile, indicative of a more efficient suppression of random noise in the visibility-based stacking.

classical image-plane stacking, the true advantage becomes evident in the case of extended sources, where the intrinsic flux and size cannot be reliably recovered in individual images because of the low SNR. In such cases, performing the fit directly in the visibility domain after stacking yields a more accurate estimate of the total flux and a lower noise level compared to the approach of fitting each individual image and then averaging the recovered flux densities. When applied to extract only the continuum emission, the *ViSta* method provides a gain of about 18% over classical image-plane stacking in our analysis. This highlights the strength of visibility-domain stacking in preserving faint, spatially extended emission that would otherwise be underestimated.

2. While testing *ViSta* with real data of high- z DSFGs observed with ALMA, where the spectral line was detectable in some sources or where a reliable zero-moment map enabled fitting of source sizes, both the *ViSta* visibility-plane fitting and classical image-plane stacking methods yield similar flux estimates. This consistency confirms the validity of our approach under

conditions where the source morphology can be reliably constrained, and it is in agreement with simulations and previous stacking experiments on local and Galactic sources (Hancock et al. 2011; Neumann et al. 2023). The fluxes derived from image-plane stacking also agree with prior literature values, further validating our methodology. However, we note that the physically correct total flux would still require additional normalization steps, such as proper flux calibration and correction for lensing magnification or selection biases, which are beyond the scope of this paper.

3. To further test our method, we artificially reduced the sensitivity of the observations by using only one fourth of the data, lowering uv -coverage and overall data quality. Under these conditions, where individual sources are no longer easily detectable, *ViSta* recovers a larger fraction of the intrinsic signal and achieves higher SNR compared to image-plane stacking. At native resolution, image-plane stacking produces more scattered and diffuse sources, while *ViSta* preserves the line emission more effectively. When all datasets are

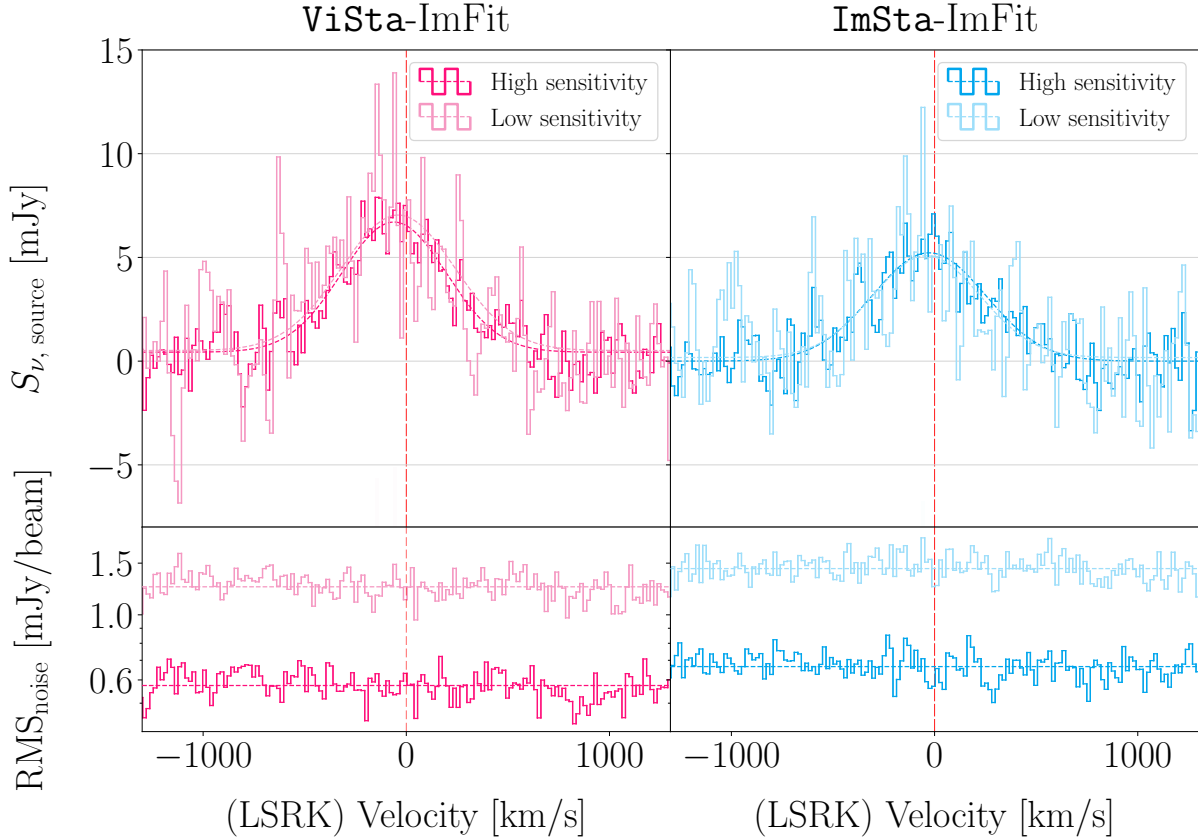


Figure 16. Comparison between stacking methods for a datasets with different sensitivities. Top panels: flux density extracted from the stacked data cubes at both high and low sensitivities. Both are shown with solid lines, with the low-sensitivity data in lighter colors, while the dashed lines represent the corresponding Gaussian fits. The left panel corresponds to the ViSta-ImFit, while the right panel shows the results from conventional ImSta-ImFit. Bottom panels: corresponding rms noise as a function of velocity, derived from line-free channels in each cube. All datasets have been convolved to the same synthesized beam to ensure consistent noise estimation across methods. The spectra demonstrate that, at both sensitivity levels, ViSta yields a higher peak flux density and reduced noise fluctuations relative to ImSta, leading to a systematically higher SNR. The improvement becomes particularly evident in the low-sensitivity case, where the visibility-based stacking more effectively preserves the coherence of the faint line signal.

convolved to the worst angular resolution, this advantage remains evident, with ViSta achieving higher SNR than image-plane stacking increasing from a relative gain of approximately 39% in the high-sensitivity case to about 53% in the reduced-sensitivity case, confirming the method’s robustness under low-SNR and sparse uv -sampling regimes.

In the future, the ViSta method will enable the reconstruction of average molecular line properties, such as SLEDs and line ratios, by stacking ALMA archival data of DSFGs, offering a unique window on the bulk physical conditions of the molecular gas in these systems. For instance, the relative intensities of different CO rotational lines, among the brightest spectral features in these sources, reveal the CO level populations, which depend sensitively on gas temperature and density; this can lead to the inference of the total molecular gas mass under the derived excitation, showing whether these extreme starbursts exhibit nearly thermalized

CO ladders (indicative of very warm, dense gas) or contain large amounts of cooler, sub-thermally excited gas—hence refining our estimates of the ISM conditions and gas content (M. Torsello et al. 2026, in preparation).

By analyzing the low and high excitation level ratios it will be possible to investigate the relative role of nuclear activity (responsible for the highest density and temperature) and star formation. By matching them with other spectral features one could try to select sub-samples and investigate how such relations evolve within the diverse populations across times.

Beyond the brighter emission lines, the power of visibility domain stacking will enable searches for much fainter lines. For example, one could target water (H_2O) lines, which can dominate the cooling of hot, dense molecular gas and intense, dust-enshrouded star formation. We will also employ ViSta to investigate shock tracers (such as HCN or CS) or to search in high-redshift galaxies for intrinsically faint lines that have so far only been detected in the local Universe (i.e., methanol, a tracer of grain-surface chemistry and turbulence). A public

release of the basic `ViSta` code tailored to the ASA will be made available at the time of publication. The code repository is hosted at <https://github.com/martitors/ViSta>.

Furthermore, an advanced version is planned, where particular attention will be given to minimizing the use of computationally expensive operations and integrating high-performance computing (HPC) strategies to handle large datasets efficiently. This is especially crucial given that stacked data products can exceed several hundred gigabytes, far beyond the handling capabilities of standard pipelines or tools designed for individual telescope datasets. In the advanced version of the code, the program will aim to rely on CASA only for calibration, while data extraction and all subsequent operations described in the pipeline will be implemented in a more efficient and optimized way. In this manner, once the *uv*-plane data are extracted, the workflow becomes identical for datasets from any interferometric telescope, such as the VLA, NOEMA, ATCA, ALMA, and MeerKAT. Even if the raw data are initially stored in different formats, they can be extracted using their native software packages and subsequently combined within the same unified framework, enabling the simultaneous stacking of observations from multiple facilities. Looking ahead, tools like `ViSta` will be particularly valuable in the SKA era, where the sheer volume and complexity of interferometric data will demand scalable and efficient analysis solutions.

Acknowledgments

This work was partially funded from the projects: INAF GO-GTO Normal 2023 funding scheme with the project “Serendipitous H-ATLAS-fields Observations of Radio Extragalactic Sources (SHORES)”; INAF Large grant 2022 project “MeerKAT and LOFAR Team up: a Unique Radio Window on Galaxy/AGN co-Evolution”; INAF Large GO 2024 project “MeerKAT and Euclid Team up: Exploring the galaxy-halo connection at cosmic noon”; “Data Science methods for MultiMessenger Astrophysics & Multi-Survey Cosmology” funded by the Italian Ministry of University and Research, Programmazione triennale 2021/2023 (DM n.2503 dd. 2019 December 9), Programma Congiunto Scuole; EU H2020-MSCA-ITN-2019 n. 860744 *BiD4BES*: *Big Data applications for black hole Evolution Studies*; Italian Research Center on High Performance Computing Big Data and Quantum Computing (ICSC), project funded by European Union—NextGenerationEU—and National Recovery and Resilience Plan (NRRP)—Mission 4 Component 2 within the activities of Spoke 3 (Astrophysics and Cosmos Observations); ORP, that is funded by the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101004719 [ORP]. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2019.1.01477.S (P.I.: S. Urquhart), ADS/JAO.ALMA#2021.1.01628.S (P.I.: T. Bakx), and ADS/JAO.ALMA#2022.1.00432.S (P.I.: T. Bakx).

ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), NSTC and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ.

Author Contributions

All authors contributed equally to the planning of the `ViSta` method, to the discussion of the results and to the writing of the present manuscript. M.T. is responsible of the development of the `ViSta` code, and performed the described simulations and analysis.

Facility: ALMA.

Software: `Astropy` (Astropy Collaboration et al. 2013), `Scipy` (Virtanen et al. 2020), `Matplotlib` (Hunter 2007), `CASA` (Team et al. 2022), `CARTA` (Comrie et al. 2018), `UVMultiFit` (Marti-Vidal et al. 2014).

ORCID iDs

Martina Torsello  <https://orcid.org/0000-0001-9301-5209>
 Marcella Massardi  <https://orcid.org/0000-0002-0375-8330>
 Elisabetta Liuzzo  <https://orcid.org/0000-0003-0995-5201>
 Gayathri Gururajan  <https://orcid.org/0000-0002-7472-7697>
 Francesca Perrotta  <https://orcid.org/0000-0001-9808-0843>
 Andrea Lapi  <https://orcid.org/0000-0002-4882-1735>

References

- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, **558**, A33
- Behiri, M., Talia, M., Cimatti, A., et al. 2023, *ApJ*, **957**, 63
- Bendo, G. J., Urquhart, S. A., Serjeant, S., et al. 2022, *MNRAS*, **522**, 2995
- Bosi, M., Lapi, A., Boco, L., et al. 2025, *ApJ*, **984**, 117
- Carilli, C. L., Lee, N., Capak, P., et al. 2008, *ApJ*, **689**, 883
- Casey, C. M., Narayanan, D., & Cooray, A. 2014, *PhR*, **541**, 45
- Comrie, A., Wang, K.-S., Hwang, Y.-H., et al. 2018, *CARTA: The Cube Analysis and Rendering Tool for Astronomy*, Zenodo, doi:10.5281/zenodo.3377984
- Cortes, P., Carpenter, J., Kamenno, S., et al. 2025, ALMA Doc. 12.3, ver. 1.5 ALMA Cycle 12 Technical Handbook, Zenodo, doi:10.5281/zenodo.14933753
- Decarli, R., Smail, I., Walter, F., et al. 2014, *ApJ*, **780**, 115
- Delhaize, J., Meyer, M. J., Staveley-Smith, L., & Boyle, B. J. 2013, *MNRAS*, **433**, 1398
- Dijkema, T. J., Golap, K., Hiriart, R., et al. 2019, MeasurementSet Definition version 3.0.0, NRAO report, July 2019, 42 pages, [E2](#)
- Giulietti, M., Gandolfi, G., Massardi, M., Behiri, M., & Lapi, A. 2024, *Galax*, **12**, 9
- Gururajan, G., Béthermin, M., Theulé, P., et al. 2022, *A&A*, **663**, A22
- Hagimoto, M., Bakx, T. J. L. C., Serjeant, S., et al. 2023, *MNRAS*, **521**, 5508
- Hancock, P. J., Gaensler, B. M., & Murphy, T. 2011, *ApJL*, **735**, L35
- Hill, R., Scott, D., McLeod, D. J., et al. 2024, *MNRAS*, **528**, 5019
- Huang, K. Y., Abbink, D., Viti, S., & García-Burillo, S. 2024, *A&A*, **688**, A130
- Hunter, J. D. 2007, *CSE*, **9**, 90
- Imanishi, M., Baba, S., Nakanishi, K., & Izumi, T. 2023, *ApJ*, **954**, 148
- Imanishi, M., Nakanishi, K., & Izumi, T. 2018, *ApJ*, **856**, 143
- Jolly, J.-B., Knudsen, K. K., & Stanley, F. 2020, *MNRAS*, **499**, 3992
- Karim, A., Schinnerer, E., Martínez-Sansigre, A., et al. 2011, *ApJ*, **730**, 61

- Knudsen, K. K., Lindroos, L., Vlemmings, W. H., Conway, J. E., & Martí-Vidal, I. 2015, *PoS*, **215**, 168
- Knudsen, K. K., van der Werf, P., Franx, M., et al. 2005, *ApJL*, **632**, L9
- Lindroos, L., Knudsen, K. K., Vlemmings, W., Conway, J., & Martí-Vidal, I. 2015, *MNRAS*, **446**, 3502
- Madau, P., & Dickinson, M. 2014, *ARA&A*, **52**, 415
- Martí-Vidal, I., Vlemmings, W. H. T., Muller, S., & Casey, S. 2014, *A&A*, **563**, A136
- Martín, S., Kohno, K., Izumi, T., et al. 2015, *A&A*, **573**, A116
- Massardi, M., Stoehr, F., Bendo, G. J., et al. 2021, *PASP*, **133**, 085001
- Negrello, M., Hopwood, R., De Zotti, G., et al. 2010, *Sci*, **330**, 800
- Neumann, L., den Brok, J. S., Bigiel, F., et al. 2023, *A&A*, **675**, A104
- Pannella, M., Carilli, C. L., Daddi, E., et al. 2009, *ApJL*, **698**, L116
- Perrotta, F., Giulietti, M., Massardi, M., et al. 2023, *ApJ*, **952**, 90
- Planck Collaboration 2020, *A&A*, **641**, A1
- Spilker, J. S., Marrone, D. P., Aguirre, J. E., et al. 2014, *ApJ*, **785**, 149
- Talia, M., Cimatti, A., Giulietti, M., et al. 2021, *ApJ*, **909**, 23
- Team, T. C., Bean, B., Bhatnagar, S., et al. 2022, *PASP*, **134**, 114501
- Thompson, A. R., Moran, J. M., & Swenson, G. W., Jr. 2017, *Interferometry and Synthesis in Radio Astronomy* (3rd ed.; Springer)
- Urquhart, S. A., Bendo, G. J., Serjeant, S., et al. 2022, *MNRAS*, **511**, 3017
- Vernstrom, T., Heald, G., Vazza, F., et al. 2021, *MNRAS*, **505**, 4178
- Vieira, J. D., Marrone, D. P., Chapman, S. C., et al. 2013, *Natur*, **495**, 344
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *NatMe*, **17**, 261