Publication Year	2018
Acceptance in OA@INAF	2021-02-24T10:33:46Z
Title	KAFE: the Key-analysis Automated FITS-images Explorer
Authors	BURKUTEAN, SANDRA; GIANNETTI, ANDREA; LIUZZO, Elisabetta Teodorina; MASSARDI, MARCELLA; RYGL, Kazi Lucie Jessica; et al.
DOI	10.1117/1.JATIS.4.2.028001
Handle	http://hdl.handle.net/20.500.12386/30568
Journal	JOURNAL OF ASTRONOMICAL TELESCOPES, INSTRUMENTS, AND SYSTEMS
Number	4

Journal of Astronomical Telescopes, Instruments, and Systems

AstronomicalTelescopes.SPIEDigitalLibrary.org

KAFE: the Key-analysis Automated FITS-images Explorer

Sandra Burkutean
Andrea Giannetti
Elisabetta Liuzzo
Marcella Massardi
Kazi Rygl
Jan Brand
Francesco Bedosti
Matteo Bonato
Quirino D'Amato
Vincenzo Galluzzi
Claudia Mancuso
Felix Stoehr
Cristina Knapic
Riccardo Smareglia

SPIE.

Sandra Burkutean, Andrea Giannetti, Elisabetta Liuzzo, Marcella Massardi, Kazi Rygl, Jan Brand, Francesco Bedosti, Matteo Bonato, Quirino D'Amato, Vincenzo Galluzzi, Claudia Mancuso, Felix Stoehr, Cristina Knapic, Riccardo Smareglia, "KAFE: the Key-analysis Automated FITS-images Explorer," *J. Astron. Telesc. Instrum. Syst.* **4**(2), 028001 (2018), doi: 10.1117/1.JATIS.4.2.028001.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes,-Instruments,-and-Systems on 03 Dec 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Journal of Astronomical Telescopes, Instruments, and Systems 4(2), 028001 (Apr–Jun 2018)

KAFE: the Key-analysis Automated FITS-images Explorer

Sandra Burkutean,[a,*] Andrea Giannetti,[a] Elisabetta Liuzzo,[a] Marcella Massardi,[a] Kazi Rygl,[a] Jan Brand,[a] Francesco Bedosti,[a] Matteo Bonato,[a] Quirino D'Amato,[b] Vincenzo Galluzzi,[b,c] Claudia Mancuso,[a] Felix Stoehr,[d] Cristina Knapic,[c] and Riccardo Smareglia[c]
[a]INAF, Istituto di Radioastronomia—Italian ARC, Bologna, Italy
[b]Università di Bologna, Dipartimento di Fisica e Astronomia, Bologna, Italy
[c]INAF, Osservatorio Astronomico di Trieste, Trieste, Italy
[d]European Southern Observatory, Garching, Germany

Abstract. We present KAFE—the Key-analysis Automated FITS-images Explorer. KAFE is a web-based FITS image postprocessing analysis tool designed to be applicable in the radio to sub-mm wavelength domain. KAFE was developed to complement selected FITS files with metadata based on a uniform image analysis approach as well as to provide advanced image diagnostic plots. It is ideally suited for data mining purposes and multi-wavelength/multi-instrument data samples that require uniform data diagnostic criteria. We present the code structure and interface, the keyword definitions, the products generated for selected users' science cases, and application examples. © *2018 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JATIS.4.2.028001]

Keywords: methods; data analysis; techniques; interferometric; astronomical data bases; miscellaneous.

Paper 18008 received Feb. 26, 2018; accepted for publication Apr. 30, 2018; published online May 22, 2018.

1 Introduction

Interferometric data reduction in the mm to sub-mm regime is increasingly being automated by instrument-specific pipelines both at the raw data reduction stage as well as during the process of image generation. This is often done for the purpose of generating a uniform data archive as well as to provide a data reduction framework. Given the increasing size of data archives, the role of data mining in astronomy will increase in the next few years.[1] In fact, with the Square Kilometre Array[2] —a radio-interferometer that is expected to produce huge amounts of data (of the order of hundreds of PB/year) to be processed and visualized—astronomy is approaching an era in which data mining will rely on image analysis tools as well as automated data mining cross-match strategies of data originating from different instruments. Data archive contents increasingly comprise not only calibrated datasets but often also include imaging products. The Atacama Large Millimeter/submillimeter Array[3] (ALMA) archive provides image cubes as part of its output package for a subselection of all data while the NRAO Very Large Array (VLA) archive survey images[4] repository contains images for selected VLA projects. Archival image products are typically in flexible image transport system (FITS)[5] format recognized as the standard archival data format for astronomical data sets. The strength of the FITS format lies in its ability to encode not only the data themselves but also a selection of the data characteristics in a machine-independent form, nowadays often exploited for data mining purposes in astronomy.[6] The presence of FITS images in interferometric data archives with their intrinsic metadata content facilitates their use in postprocessing analyses, accessible not only to the expert community in the particular wavelength regime but also to a wide range of astronomers from across the fields. One of the most convenient ways to supply such information is via keyword-value pairs in the FITS header structure that is easily machine readable and can be incorporated into archival databases. In the case of archived images, the FITS header metadata should subsequently be the primary location for the repository of this information. When dealing with archival images from a range of different instruments, observing bands or indeed simply a large sample of sources extracted from a single data archive, easily accessible and uniformly generated image diagnostics are essential to assess and thus exploit the full data potential.

We therefore present KAFE,[7] the Key-analysis Automated FITS-images Explorer, a tool we developed to specifically address these issues in a single framework. Working directly on the FITS images, KAFE offers a user-friendly image processing platform suitable for several of the major radio to sub-mm astronomical facilities [ALMA,[3] Australia Telescope Compact Array[8] (ATCA), Karl G. Jansky Very Large Array[9] (VLA), Combined Array for Research in Millimeter-wave Astronomy[10] (CARMA), Very Long Baseline Array[11] (VLBA), etc.], as it is free from any telescope-specific data format constraint. It is comprised of a suite of tools to quickly perform (in "a few clicks") all the most commonly used image analysis steps to give an overall picture of the target radio-sub-mm data.

KAFE works on three levels, namely, the image header contents, the FITS image data itself as well as catalog cross-matching. The first focuses on the production of additional and uniformly defined keyword-value pairs allowing for a quick observing setup and data quality comparison across instruments. The KAFE image diagnostic suite produces advanced analysis plots that allow in-depth exploitation of large data samples as well as quick insights into each dataset content and scientific

*Address all correspondence to: Sandra Burkutean, E-mail: burkutean@ira.inaf.it

2329-4124/2018/$25.00 © 2018 SPIE

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes,-Instruments,-and-Systems on 03 Dec 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

exploitation capabilities in the spatial, frequency, and time domain. The presented suite goes beyond the common preview tools [SAOImage DS9,[12] Karma,[13] Cube Analysis and Rendering Tool for Astronomy[14] (CARTA)] that require significant "by-hand" user input as KAFE defines settings and thresholds from the data in a fully automated way while also offering advanced processing products (e.g., instantaneous three-dimensional (3-D) views, spectral analysis features, and component detection). The large suite of diagnostic plots and analysis tools makes it easily accessible to respond to a broad range of user cases. The third part of KAFE allows users to query astronomical catalogs automatically based on signal detection algorithms in the imaging suite.

The KAFE interface layout with its extended keyword suite and image analysis options are provided in Sec. 2. Section 3 outlines the FITS header keywords that KAFE extracts from the input images and optionally adds to the FITS headers. Sections 4 and 5 summarize the postprocessing analysis, image visualization tools, and catalog cross-match possibilities illustrated via the application of KAFE to interferometric data. In Sec. 6, we report a selection of scientific example cases that exploit the functionalities of KAFE. Finally, in Sec. 7, we give a summary of the KAFE suite and provide details on how to access the KAFE web interface.[7]

2 KAFE Suite

KAFE is a fully automated image diagnostics suite that explores interferometric FITS images in the spatial, spectral, and/or temporal parameter space, thus bridging the gap between data visualization and image science potential assessment. In the light of data archive mining projects, KAFE facilitates image sample property classifications, cross-matching, and comparison across radio to sub-mm instruments.

The code is fully Python-based and exploits the CASA[15] task capabilities as well as the packages Scipy,[16–19] APLpy,[20] Astropy,[21,22] and Astroquery.[23] Its functionalities are accessible through a web interface,[7] hosted by one of the web-servers of the Italian center for Astronomical Archive. In addition, KAFE incorporates the image-based part of the ALMA Keyword Filler (AKF[24]) tool package. The AKF package aims at improving the ALMA Science Archive miner experiences by supplementing the keywords already available in the FITS headers with additional keyword-value pairs. The AKF toolkit includes parameters that describe the observational setup (e.g., time ranges and array conditions), the instrumental setup (e.g., coordinates, frequency ranges, and angular and spectral resolutions) as well as image characteristics (e.g., dynamic range, noise levels, and peak flux) extracted from the calibrated measurement sets as well as the associated images.

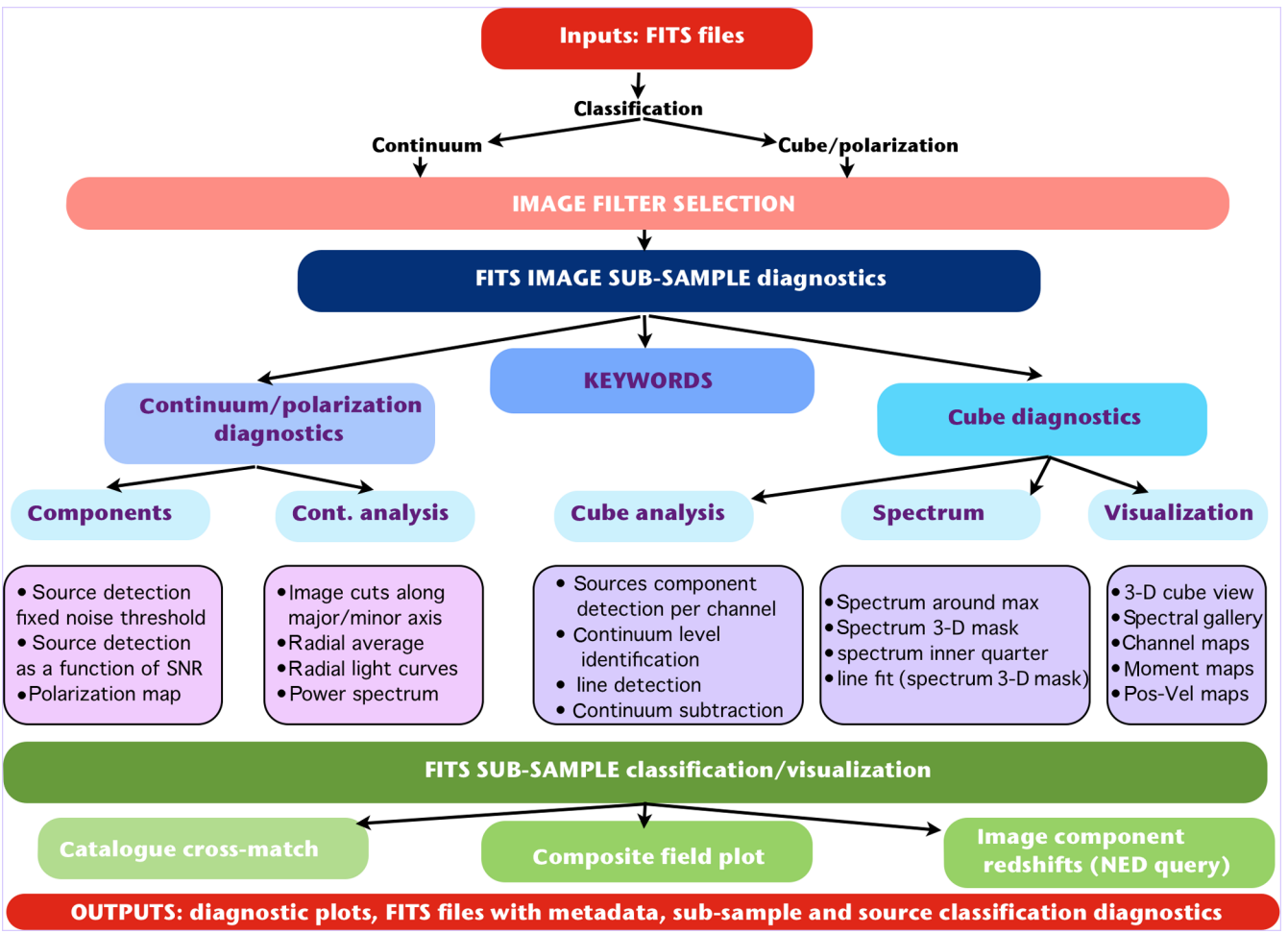

Fig. 1 The KAFE code structure. KAFE is structured into four sections: data input, image selection filter application, image postprocessing (divided into metadata computation, cube and single-plane image diagnostics), and catalog cross-matching.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes,-Instruments,-and-Systems on 03 Dec 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

KAFE's code structure is shown in Fig. 1. The input can be a single or a collection of radio to sub-mm FITS images provided by the user (see an image of the input interface in Fig. 2). These are either previously downloaded directly from the respective archives by the user or produced by the user's imaging routines.

Since KAFE is strongly dependent and limited on the quality of the input FITS images, we caution that users should be well

Fig. 2 Snapshot of version 1.0 of the KAFE interface whose layout mirrors the code structure in Fig. 1. The interface was designed to be able to incorporate further future capabilities as KAFE is devised as a long-term image analysis toolkit flexible in its incorporation of additional features. Note that the filter option is an optional feature and not mandatory for the execution of KAFE runs.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes,-Instruments,-and-Systems on 03 Dec 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

```
############################## test1.fits#################################
KRACENTER:          334.376917917  deg
KDECCENTER:         0.0173538888889  deg
KSPATRES:           0.52591746722  arcsec
KBNDCTR:            271812516941.0  Hz
KBNDRES:            18628936524.2  Hz
KBNDWID:            18628936524.2  Hz
KCHANRMS:           [0.000249888063990511]  Jy/beam
KDYNRANGE:          [10.079535518414469]
KFLUXTOT:           [0.0140585788445162217]  Jy
KDATAMAX:           [0.0025187556166201 83]  Jy/beam
KDATAMIN:           [-0.0016350860241800547]  Jy/beam
KSTOKES:            ['I']
############################## test2.fits#################################
KRACENTER:          334.21485  deg
KDECCENTER:         0.304776944445  deg
```

Fig. 3 Example of the KAFE keyword catalog output for a list of images. See the Appendix for the AKF keyword definitions.

aware of the image content, quality, and production history. For archives whose image contents are not denoted as "science-ready," we recommend that data are recalibrated and reimaged according to the users' needs before applying KAFE to extract values to be used in publications. Once uploaded, each input image is initially classified as being a continuum (i.e., a single layer map) or a cube (i.e., a collection of maps at varying frequency channels). After the input classification stage, KAFE provides image sample filter options based on a single or multiple keyword-value pairs that KAFE calculates from the images or extracts from their respective header information. If a filter option has been selected, the subsequent diagnostic plots will only be made for the resultant subsample.

The image analysis features are divided into three parts: keyword computation, image cube spectrum analysis, and single-plane image diagnostics as well as catalog cross-matching and composite plot features. In the next sections, we will give an outline of the AKF keyword metadata currently implemented into KAFE, illustrate the image postprocessing tools, and summarize the catalog cross-match and KAFE image visualization capabilities.

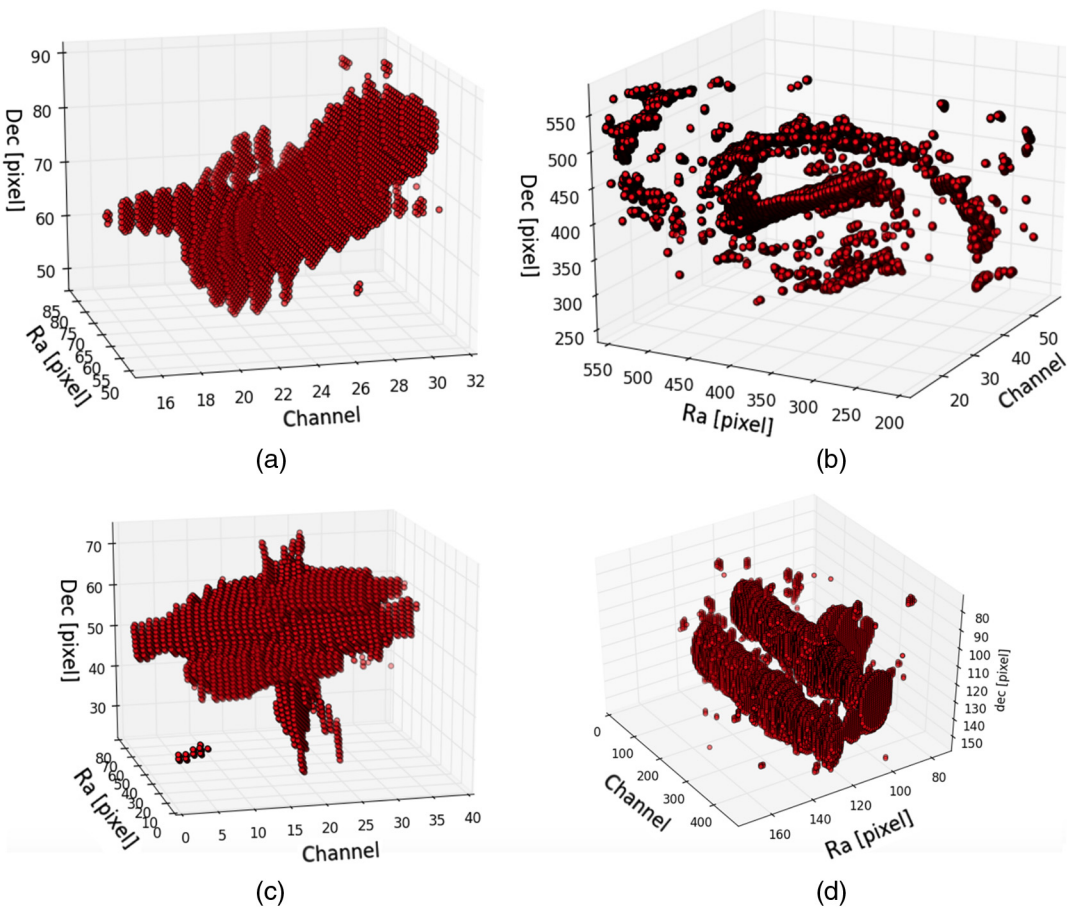

Fig. 4 Illustration of the 3-D emission/absorption structures within the cubes in an automatically generated single 3-D view. (a) ALMA NGC3256 CO (1-0) data, (b) ALMA 12-m data of M100 in CO (1-0) clearly illustrating the spiral structure in 3-D, (c) ALMA observations of the OrionKL region in band 6, and (d) ALMA IRAS16293 band 6 observations with two separate source regions clearly visible.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes,-Instruments,-and-Systems on 03 Dec 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

3 FITS Image Keywords Definitions

When examining images from a range of instruments, data miners might encounter difficulties if the keywords present in the headers lack supplementary documentation on their definitions. Even though the standard FITS 4.0 keywords (at least for the most recent datasets) should be available in the headers of current FITS formats, headers specific to the data reduction might not follow the standard 4.0 conventions or might have nonuniform definitions across instruments. KAFE does not edit or remove any of the keywords already present in the header, but adds a few more via the AKF[24] routines on demand. The new metadata are flagged with an initial "*K*" to be clearly recognizable within the header structure and to minimize the risk of confusion with existing keywords, see Fig. 3. The keyword value definitions are outlined in the Appendix and are entirely based on information gained from the image content in conjunction with existing image headers. Their purpose is not only to complement image FITS headers but also to facilitate a uniform comparison of images from the same or multiple instruments (e.g., data of different sources, data on the same source at different observing setups or epochs). Though some keyword values are also noted in the standard FITS headers, we include them in the KAFE metadata so that the KAFE metadata catalog can easily be reused for future source catalog cross-matching.

4 FITS Image Postprocessing Image Diagnostics

KAFE provides a range of image diagnostic tool suites: image cube diagnostics, single image plane component classification as well as image sample comparisons across the time and frequency domains. Below, we summarize and illustrate the main diagnostic features currently available within KAFE via its application to ALMA archival data. We stress that their purpose in this case is not to show scientific results, which in many cases are already presented in dedicated publications but to illustrate the KAFE suite capabilities.

After continuum/cube classification, optional selection filter application, and AKF keyword value production within KAFE, the images are checked for previous application of primary beam correction, in which case, KAFE excises the outer boundary of the images to avoid noisy edge features to be picked up in the analysis.

4.1 *Image Cube Analysis Plots*

The image cube analysis diagnostic plots are divided up into 3-D cube representation, spectral analysis, and spatial component detection. Using a dynamic-range-sensitive noise threshold, the most significant image cube components are accessible in 3-D image cube views (see Fig. 4).

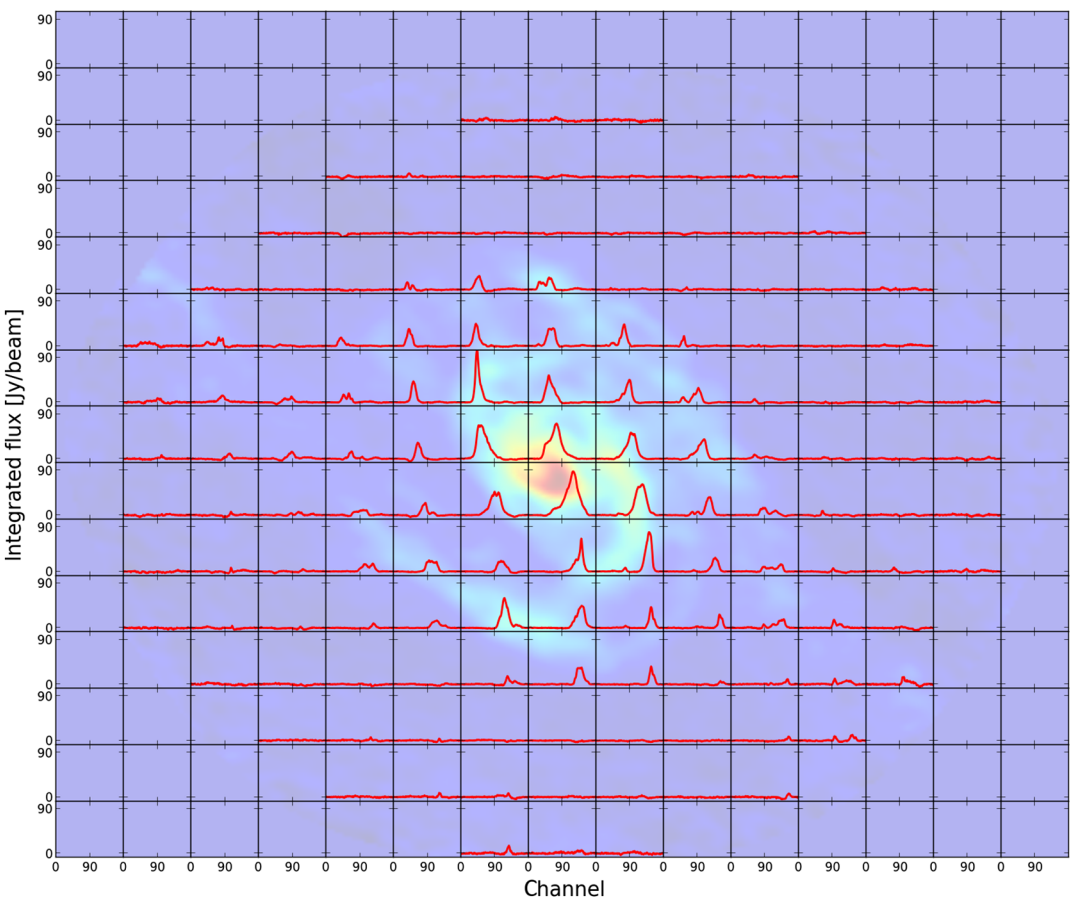

Fig. 5 Spectral gallery of ALMA CO (1-0) data of the Circinus galaxy. Note that these spectra are made before continuum subtraction and that we define the spectrum to be the sum of the flux in each channel within the selected region (according to CASA's spectral sum definition).

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes,-Instruments,-and-Systems on 03 Dec 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

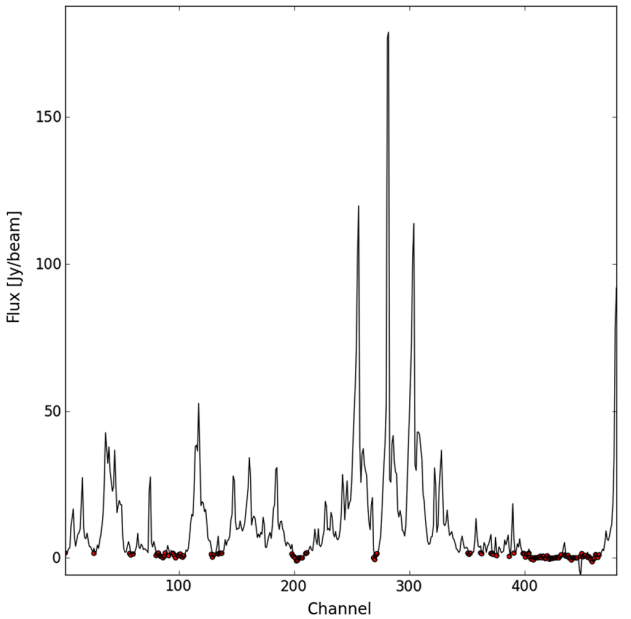

Fig. 6 Continuum channel identification in the spectrum of ALMA IRAS16293 band 6 observations (red points are continuum channels). These channels are used for continuum subtraction in the image plane to generate the line-only cube. The identified continuum level is used as an initial value for the continuum level fit during the line fitting process.

For quick identification of spatial variations in the spectral image properties, KAFE offers a spectral gallery view (see Fig. 5). Furthermore, the user can choose between a spectral extraction around a dynamic-range-sensitive noise threshold within the whole cube, a spectrum within the inner quarter of the image cube, as well as a spectrum of the region with maximum SNR using an excision radius of three times the beam size. In the future, we will also offer more sophisticated image cube region selections for the spectra computations based on our cube component feature detection. Throughout KAFE, we define the spectrum to be the sum of the flux in each channel within the selected region (according to CASA's spectral sum definition). Note that line + continuum detection as well as line fitting features are currently solely implemented for the spectrum obtained via 3-D dynamic range-sensitive masks (see Fig. 4). In addition, continuum identification is currently based on excision regions depending on whether the input image is primary beam corrected or not. In case of primary beam correction, the same mask as in the 3-D cube plot is used, being dependent on the dynamic range in the image. Otherwise, a three-sigma threshold is used on the whole cube to include all possible emission/absorption regions. We note that faint features may be averaged using this approach and more sophisticated automated spectra excision algorithms are in development. The continuum level and line features in the image cubes are detected using a combination of statistical thresholding in combination with

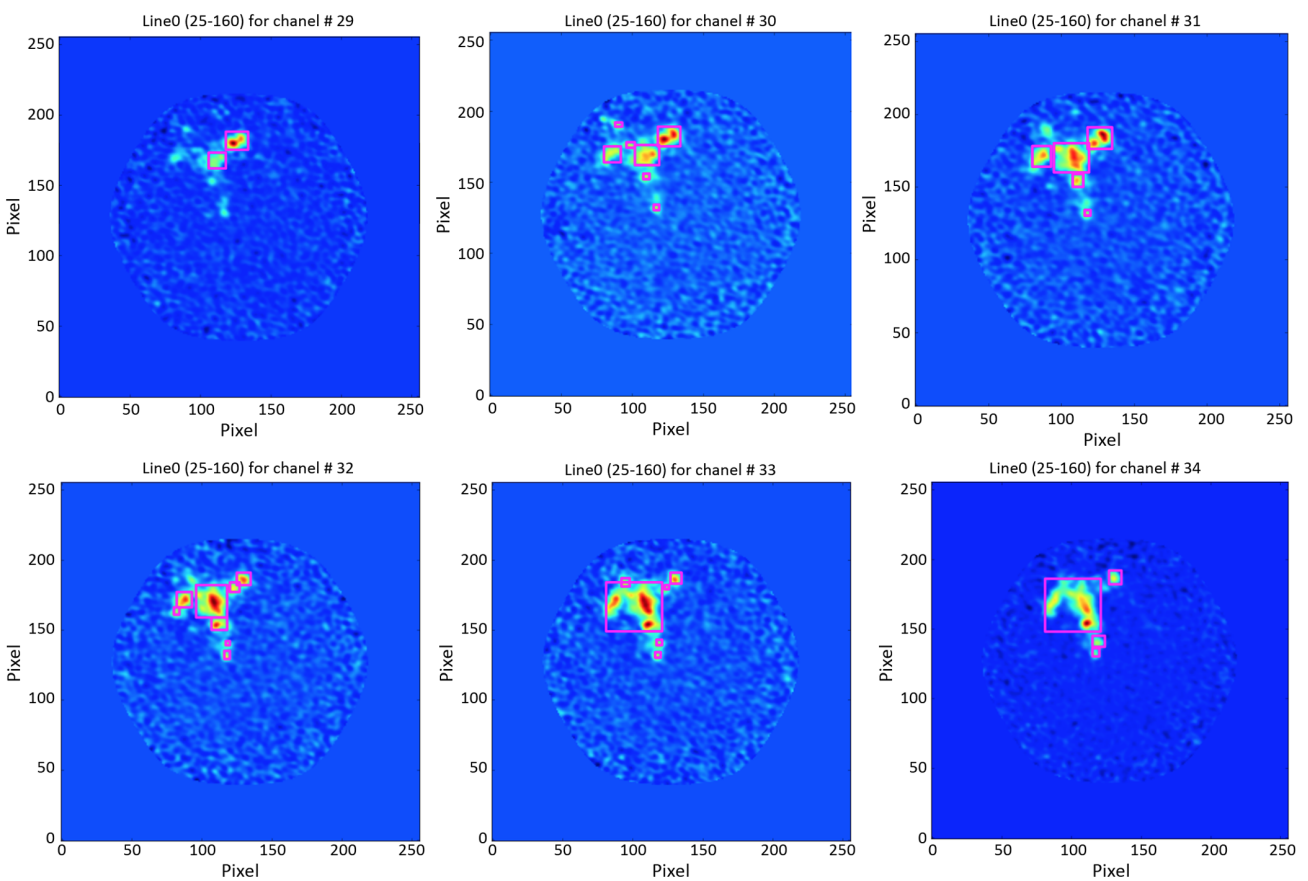

Fig. 7 Spectral image gallery of the line region in the Circinus galaxy in CO (1-0) after continuum subtraction in the image plane with the component detection algorithm, described at the end of this section, applied at an SNR threshold of 5.0 in each channel, and illustrated via the magenta rectangles. Image components are defined as independent nonconnected islets in the corresponding SNR threshold channel map. Displayed channel maps do not have the noise threshold applied.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes,-Instruments,-and-Systems on 03 Dec 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

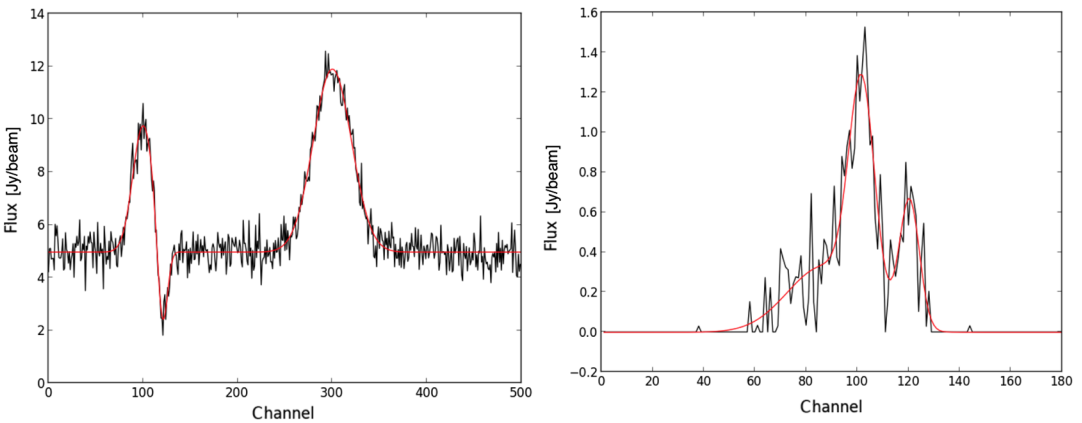

Fig. 8 The line fitting algorithm illustrated on a simulated spectrum with both emission and absorption line features (left) and CN ALMA data of the Circinus galaxy (fit shown as red line).

a least-squares cross-validation for a KDE approach[25] and a wavelet-based peak detection algorithm using SciPy[16] routines (see Fig. 6). Continuum subtracted channel maps are subsequently made for inspection (Fig. 7). Detected lines are fit on the input FITS-file cube using a nonlinear least-squares Gaussian fitting approach assuming a zeroth-order continuum level whose starting value for the fit is taken from the continuum finding code and subsequently fit jointly with the identified lines (see Fig. 8).

A 3-D position-velocity diagram extracted from rectangular spatial slabs allows the user to diagnose the spectral features as a function of two orthogonal directions along the RA and DEC directions, thus giving an averaged view of spectral variations as a function of 2-D image position (see Fig. 9). In addition, moment 0, 1, and 2 maps are produced for each detected line using an optimized dynamic range-dependent noise threshold defined by the 3-D mask from the continuum and line channel identification stage (see Fig. 10). The object orientation is identified from the moment 0 map and position-velocity diagrams along the major and minor axes within a width of four times the synthesized beam are made (see Fig. 11).

For polarization data, polarization vector plots are provided with vector lengths scaled according to the intensity (see Fig. 12). The length scaling can be linear or logarithmic depending on the dynamic range of the image. Morphological feature detection, using SciPy[16] routines, as a function of signal-to-noise threshold ensure polarization vector displays centered on the main structural features in the image.

Component classification for each channel in a cube is also offered for a predefined dynamic-range-dependent noise threshold clipping level, Fig. 7. The algorithm is based on a structuring element characterized by the beam dimensions that identifies connected regions above different signal-to-noise thresholds. For single-plane (continuum or single input channel) FITS images, morphological structure detection is furthermore offered as a function of different signal-to-noise thresholds (see Fig. 13). Further developments are currently underway to quantitatively characterize the detected components in terms of their characteristic scales (consult the KAFE cookbook[26] for updates).

4.2 *Single Plane Image Analysis Plots*

Besides the image component detection features described above, we offer further diagnostic features based on temporal and image plane structural classifications for continuum input FITS images. In particular, users can plot the light-curve of the peak emission if the input FITS file list consists of the same target observed in different observing runs or at different times on the same day (i.e., images made from different scans), see Fig. 14.

Note that we provide a CASA-based script in the KAFE cookbook[26] that writes a scan selection-based time

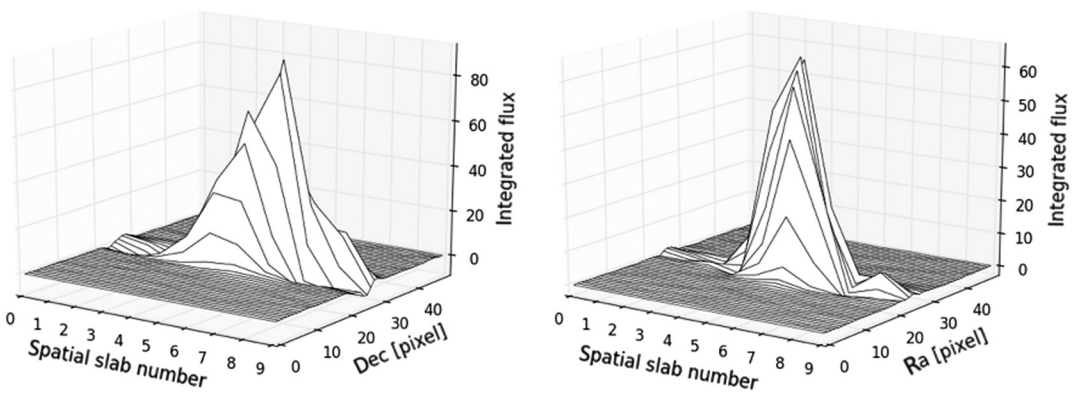

Fig. 9 3-D position-velocity views using spatial slabs in the RA and DEC direction respectively of NGC 3256 in CO.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes,-Instruments,-and-Systems on 03 Dec 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

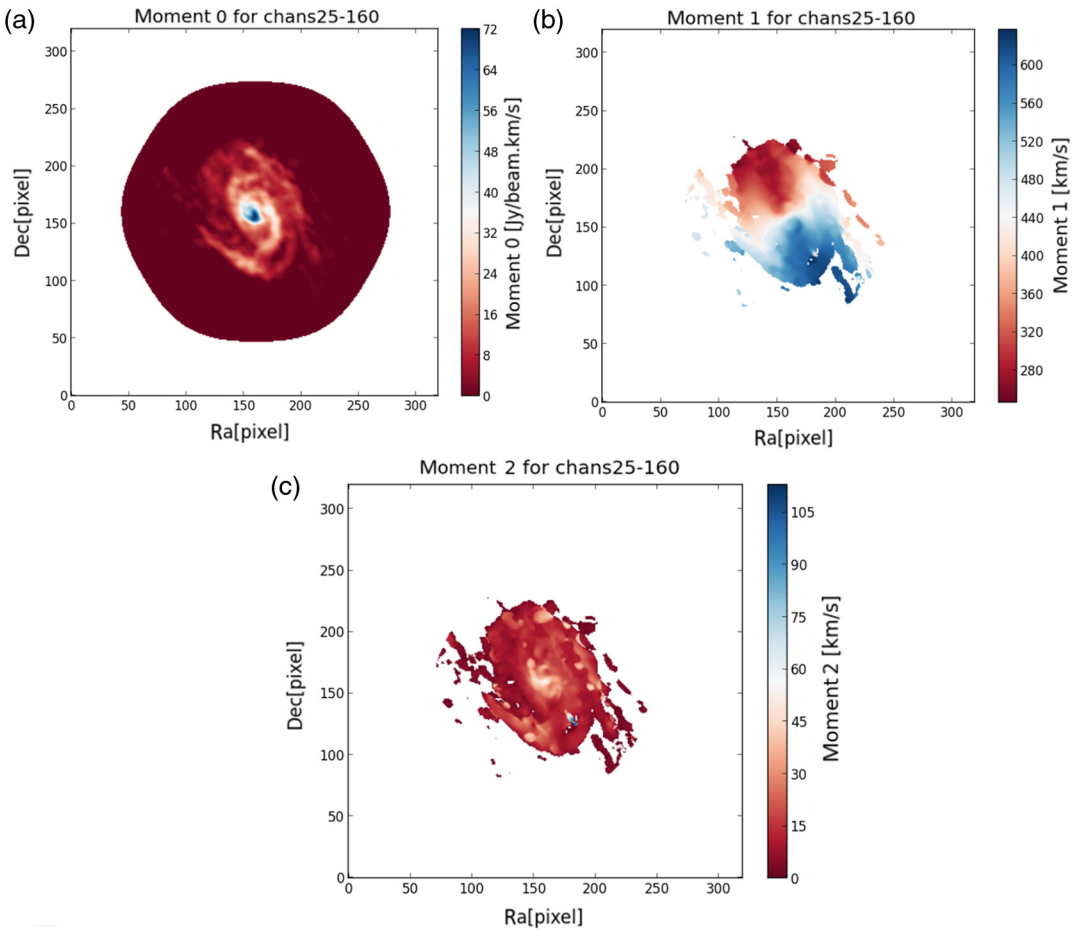

Fig. 10 Moment 0, 1, and 2 maps of ALMA 12 m Circinus galaxy CO (1-0) data made using the automatic mask generation. (a) Moment 0 corresponds to the map of the integrated line flux density, (b) moment 1 corresponds to the map of the mean of the velocity distribution, and (c) moment 2 corresponds to the dispersion of the velocity distribution.

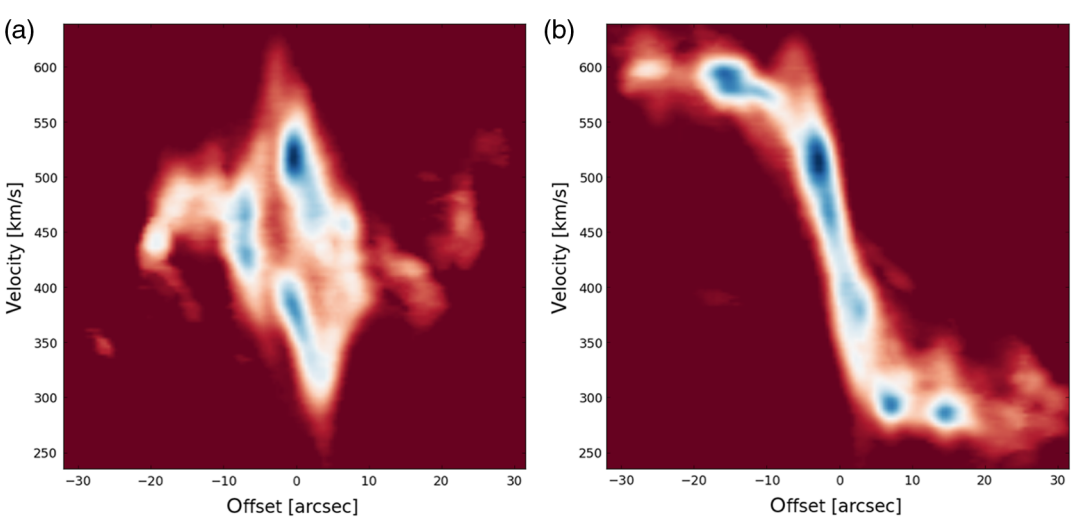

Fig. 11 Fully automatically computed position-velocity diagram of ALMA 12 m Circinus galaxy CO (1-0) data along the (a) minor and (b) major axes. An extraction width of four times the beam size is used.

keyword to the header of the cleaned image since the correct time stamp is required for the light-curve plot (and not the default time keyword that denotes the start of the observation).

Furthermore, in order to derive the orientation of extended sources, moments of inertia and the associated axes are computed. The orientation of the image is then used to derive image cuts along the respective major and minor axes, see Fig. 15.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes,-Instruments,-and-Systems on 03 Dec 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

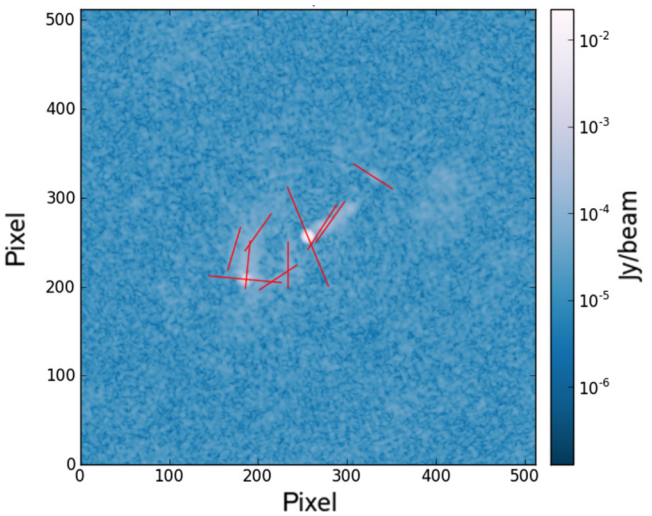

Fig. 12 A polarization vector plot of PKS0521-365. The length of the polarization vectors is logarithmically proportional to intensity (for images with less dynamic range the respective vector length scaling is linear).

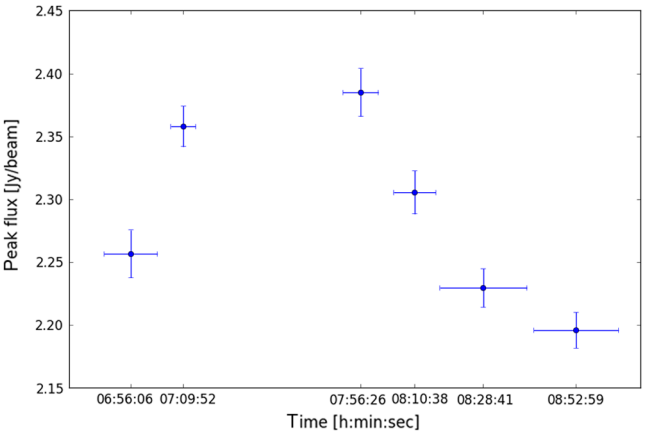

Fig. 14 KAFE-generated light curve of ALMA band 3 Sagittarius A* observations.[27] The horizontal error bars indicate the time range over which the peak flux was computed.

5 Catalog Cross-Matching and Multifield Visualization

Having classified the images in terms of their spatial, spectral, and/or temporal features, KAFE then offers the possibility to do catalog cross-matching for redshift and multifrequency flux analysis. We note that we do not aim to do aperture photometry or redshift identification in this tool at the current stage as these require careful consideration of the different beam sizes. Instead, we merely focus on a position-query-based cross-match between the catalogs to offer the user a larger database to draw scientific conclusions from.

In addition, we provide composite field plots (for a maximum composite field size of 1 deg), Mollweide projection plots for large image samples, right-ascension-redshift pie plots, and

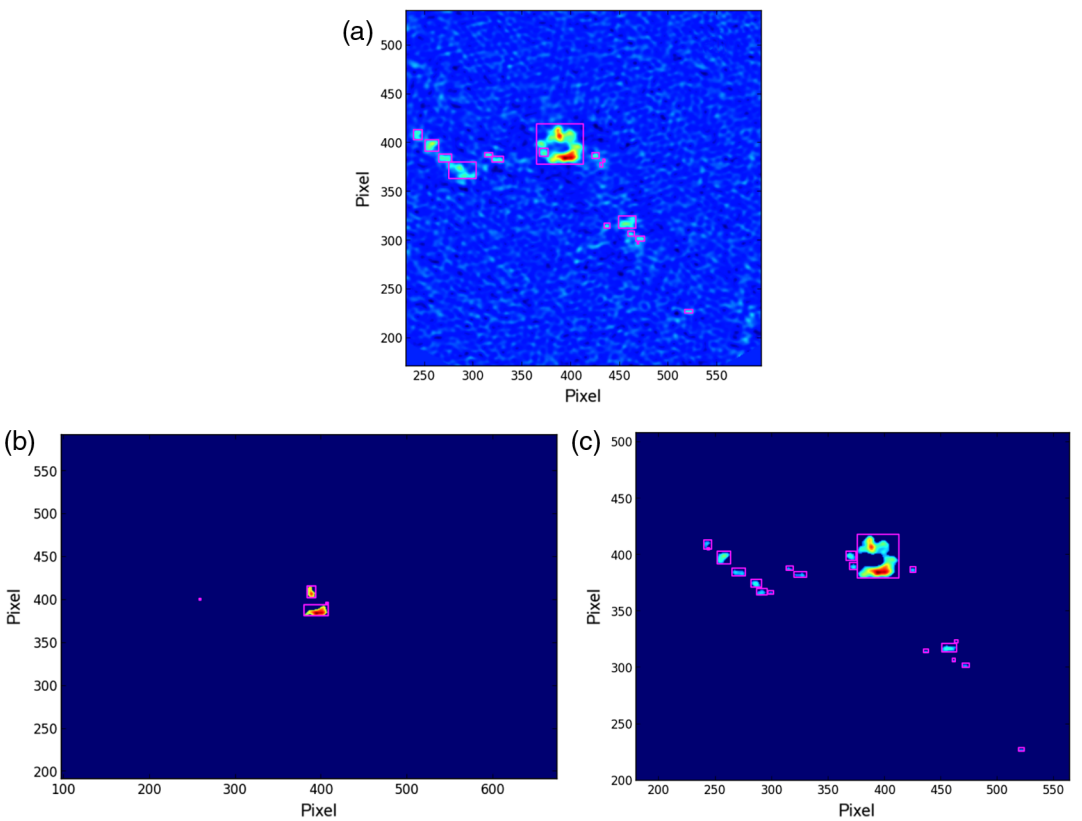

Fig. 13 The automated component detection algorithm applied to a single channel FITS file in ALMA M100 data. (a) Component detection using a dynamic range-dependent threshold. (b) and (c) Component detection as a function of SNR thresholds [(b) SNR $\approx$ 14, (c) SNR $\approx$ 9] illustrating the characteristic emission scales as a function of emission/absorption feature strength.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes,-Instruments,-and-Systems on 03 Dec 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

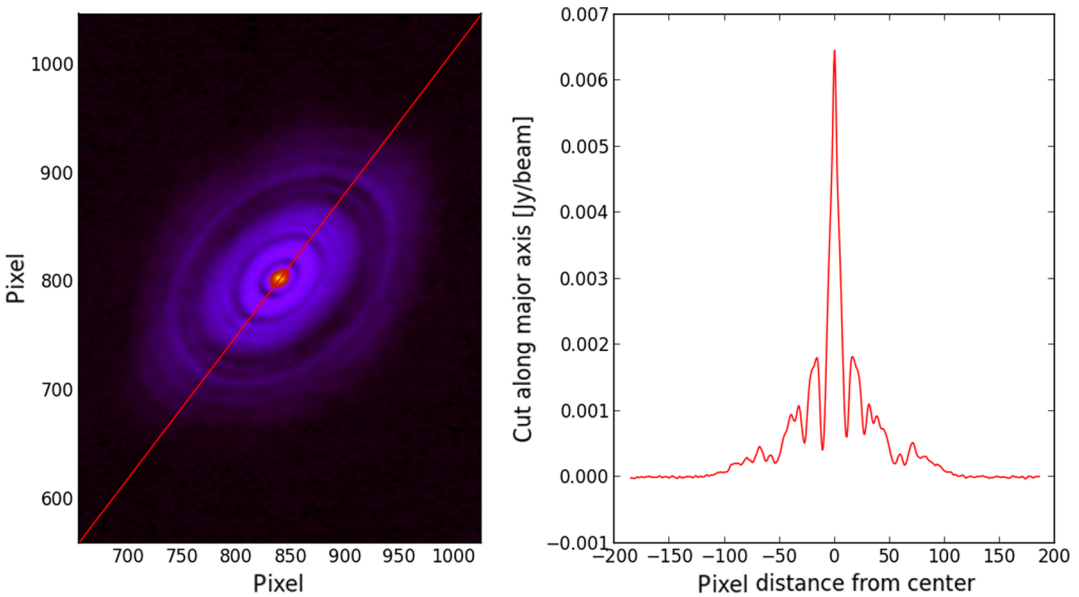

Fig. 15 High angular resolution HL Tau band 6 ALMA data and the associated cut along the major axis whose orientation is computed automatically.

three-color images (see Fig. 16). Detected source positions are cross-checked with the NASA/IPAC Extragalactic Database[29] (NED) for redshift information and multiband flux values. We also offer cross-match queries with deep-field catalogs (HUDF and COSMOS for example) as well as several object classification sample catalogs (consult the KAFE cookbook[26] for further details). In addition, a cross-match with DustPedia[30] is currently in planning. KAFE outputs could also be used with Virtual Observatory tools such as the Tool for OPerations on Catalogues And Tables[31] to cross-match the sample properties with other catalogs and databases not already incorporated into KAFE to investigate property relations (e.g., source counts). Having run the metadata, diagnostic plot features, and catalog cross-match options according to the user input, KAFE output PNG files are displayed on the KAFE website for each FITS input image and a download button allows the resultant KAFE products to be retrieved.

6 Application Examples

In this section, KAFE's versatility is illustrated by examining present and past scientific example cases and their use of KAFE components.

6.1 *Catalog Definition for Statistical Analyses*

Since KAFE accepts multiple FITS images as input, its combined filter functions and image property classifications can provide the necessary parameters to construct catalogs of sources. KAFE was applied to images of I, Q, or U Stokes parameters for a sample of ALMA observation of point-like AGNs to infer their flux densities and construct the catalog.[32,33] For continuum images with single emission/absorption regions, a minimal selection of KRACNTR, KDECCNTR, FREQ, KDATAMAX, KDATAMIN, KFLUXTOT, and KCHANRMS (see Appendix for AKF[24] definitions) is enough to identify sources. In the case of single as opposed to multiple sources in the field of view, one can infer a source detection from the ratio KFLUXTOT/KCHANRMS and specify if the source under investigation has an extended structure within the observed angular scales from the ratio KFLUXTOT/KDATAMAX. KAFE has also been used to assess the image properties of a sample of galaxy cluster Sunyaev-Zel'dovich images. In addition, for spectral line cubes, the KAFE web interface allows one to plot and retrieve spectra and visualize the images of each channel with line emission/absorption features. We are investigating the CO properties in selected samples of starburst galaxies using KAFE to examine the component detection as well as spectral feature classification.

6.2 *Multifrequency Single Source Analyses*

KAFE can be used to compare the spatial and spectral coverage of single sources. KAFE was tested to compare the same spectral line of different star-forming regions in the local galaxy NGC5135.[34] In addition, the properties of CO in different components of the lensed galaxy SDP9[35] were compared to infer hints on the geometry of the lensed system. Also, in these cases of keyword selection and image diagnostics, those suggested in the previous sections are enough to indicate many properties of the investigated portions of the input target and extract statistical analyses or comparison plots for the purpose of first-look diagnostics.

6.3 *Investigation of Light-Curves and Source Variability*

If the input images of the same target have been observed at different epochs, the retrieved catalogs can be used to assess the light-curve or variability of the target itself. Information extracted from continuum images for complete samples of ALMA calibrators were used to reconstruct a catalog with KAFE AKF keyword values of complete samples of AGNs observed at many epochs to investigate blazar variability at ALMA frequencies.[28] A single run of KAFE would be enough to retrieve the whole flux density catalog at all frequencies, speeding up a process that could be time consuming to do

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes,-Instruments,-and-Systems on 03 Dec 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

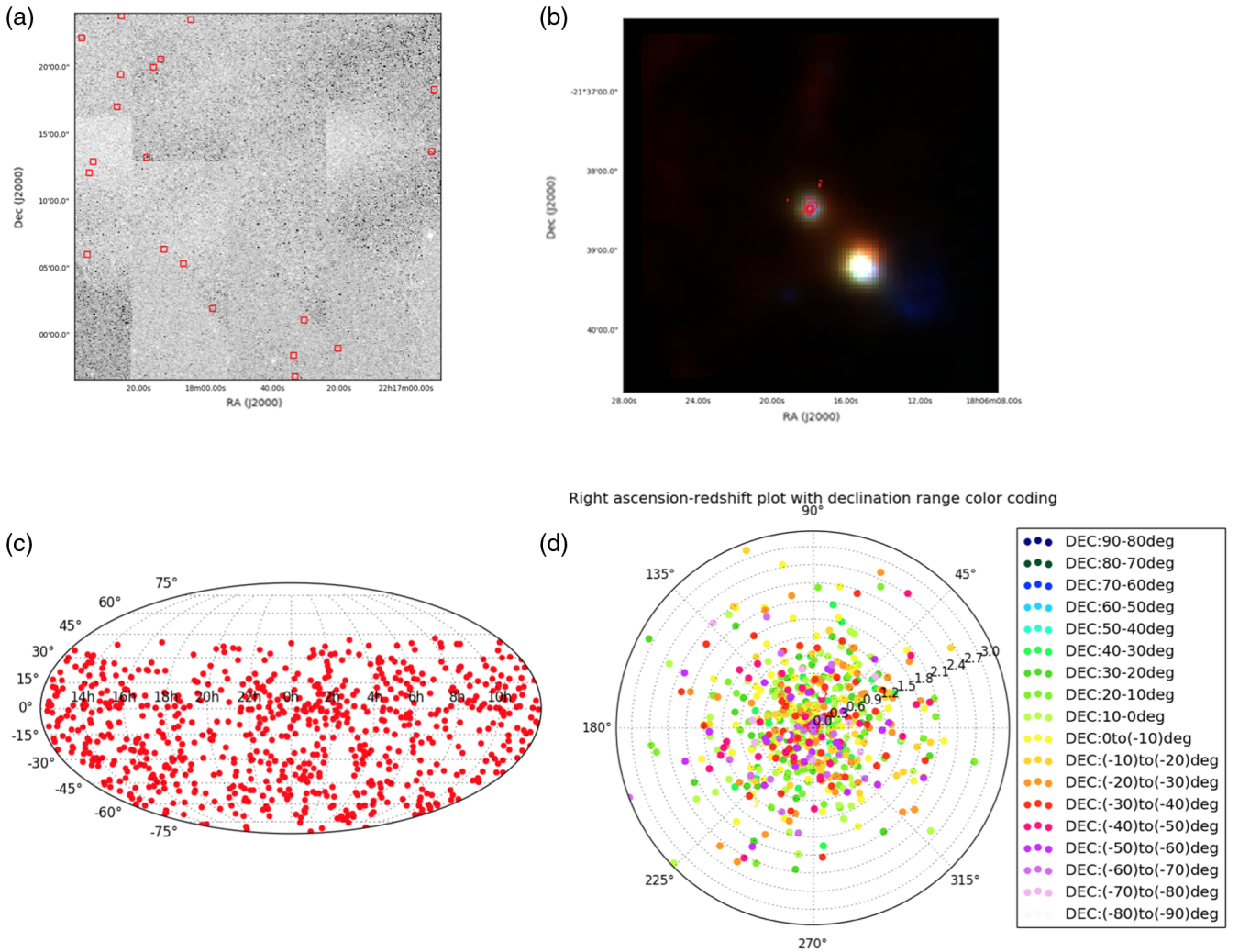

Fig. 16 Image sample diagnostics. (a) KAFE composite field plot for a selection of input FITS files showing the input image FITS extent as red boxes. (b) An example of a three-color image plot as implemented in KAFE using APLpy functionalities.[20] (c) Mollweide projection plot. (d) Right-ascension-redshift pie plot with color-coded declination ranges for a selection of ALMA calibrators[28] constructed via KAFE tasks.

manually as it requires detailed knowledge of ALMA data, CASA tasks, and interferometric image analysis. The catalog cross-matching tools, together with future aperture photometry implementations, will allow one to complement the light-curve analysis with the SED reconstruction over a broad range of flux densities for the whole sample.

7 Summary and Future Implementations

We presented KAFE, a web-based FITS image postprocessing analysis tool. It was designed in the framework of the ALMA archive development activities at the Italian node of the European ALMA Regional Center network and developed to be widely applicable to FITS images of the major radio-sub-mm facilities. We outlined its capability to complement the input FITS keyword headers and described the suites of diagnostic postprocessing plots, visualization, and catalog cross-matching tools. To obtain access to the current version of KAFE, send an e-mail to kafe@ira.inaf.it. For further instructions on how to use the web interface,[7] consult the KAFE cookbook.[26]

Appendix: AKF Keyword Definitions

Below, we give a summary of the AKF definitions:[24]

- KRACNTR and KDECCNTR: The center pixel in right ascension and declination of the image.
- KBNDCTR: The center frequency of the image is calculated as

$$\mathrm{KBNDCTR} = 1/2 \times (\mathrm{CRVAL}n + \mathrm{CDELT}n \times \mathrm{NAXIS}n), \quad (1)$$

  where n is such that CTYPEn = FREQ, CRVALn is the reference value for the frequency axis (in the initial channel), CDELTn is the increment per pixel of the frequency axis, and NAXISn is the number of frequency channels. Units are in Hz.
- KBNDRES: The frequency resolution of the image defined as the increment of frequency axis CDELTn where n is such that CTYPEn = FREQ. Units are in Hz.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes,-Instruments,-and-Systems on 03 Dec 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

- KBNDWID: The effective bandwidth of the image is calculated as

$$\mathrm{KBNDWID} = \mathrm{CDELT}n \times (\mathrm{NAXIS}n)\,, \tag{2}$$

  where CDELTn is the increment in the frequency axis, NAXISn is the number of frequency channels where n is such that CTYPEn = FREQ. Units are in Hz.
- KCHANRMS: The interquartile range of the pixel values in the inner third of the image. For spectral line data, it loops over the channel axis, takes the interquartile of the pixel values in each channel, and then calculates the interquartile of the estimated channel rms values. Units are in Jy/beam.
- KDATAMAX and KDATAMIN: The maximum and minimum values in the entire image, respectively. Units are in Jy/beam. In the case of polarization data, one value is given for each Stokes image.
- KDYNRNG: The dynamic range for each Stokes image is defined as KDATAMAX/KCHANRMS. For polarization data, one value is given for each Stokes image.
- KFLUXTOT: The integrated flux of the region (presumably including sources) obtained discarding the image pixels with values below $|3 \times \mathrm{KCHANRMS}|$. In the case of polarization data, one value is given for each Stokes image.
- KSPATRES: The geometric average of the maximum (BMAX) and minimum (BMIN) synthesized beam axes. It exploits the BMIN and BMAX values given in the image headers. For spectral line data, it loops over the frequency axis channels, takes the mean values of all the channel BMIN and BMAX values, and calculates their geometric average.
- KSTOKES: The list of image Stokes parameters. Possible character values are either "*I*" or "*I*, *Q*, *U*, *V*." It cannot handle the case of CRVALn with $n = 2$ or 3.

Acknowledgments

The AKF and KAFE development are part of the activities for the ALMA Re-Imaging Study approved in the framework of the 2016 ESO Call for Development Studies for ALMA Upgrade (PI: Massardi). The study acknowledges partial financial support by the Italian Ministero dell'Istruzione, Università e Ricerca through the grant "Progetti Premiali 2012—iALMA" (CUP C52I13000140001). We thank Massimo Sponza, Giuliano Taffoni, and Gianmarco Maggio for the technical support that allowed the web server machine to be operational for our study activities, and Franco Tinarelli and Sonia Zorba for the future accounting system. We thank Andrea Enia for useful discussions. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2011.0.00002.SV, ADS/JAO.ALMA#2011.0.00004.SV, ADS/JAO.ALMA#2011.0.00015.SV, ADS/JAO.ALA #2013.1.00247.S, ADS/JAO.ALMA#2011.0.00887.S, ADS/JAO.ALMA#2015.1.01522.S as well as Orion KL band 6 and IRAS 16293 band 6 ALMA science verification data. ALMA is a partnership of ESO (representing its member states), NSF (USA), and NINS (Japan), together with NRC (Canada), NSC 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. This research made use of Astropy, a community-developed core Python package for Astronomy (Astropy Collaboration, 2018). This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

References

1. K. Borne, "Scientific data mining in astronomy," ArXiv e-prints arXiv:0911.0505 (2009).
2. "Square Kilometre Array," https://www.skatelescope.org/ (2018).
3. "Atacama Large Millimeter/submillimeter Array," http://www.almaobservatory.org/en/home/ (10 May 2018).
4. "NRAO VLA Archive Survey Images Page," https://archive.nrao.edu/nvas/ (04 September 2013).
5. D. C. Wells, E. W. Greisen, and R. H. Harten, "FITS - a Flexible Image Transport System," *Astron. Astrophys. Suppl. Ser.* **44**, 363 (1981).
6. D. Hanish et al., "SAFIRES: Spitzer archival FIR extragalactic survey," *Bull. Am. Astron. Soc.* **43**, 334.09 (2011).
7. S. Burkutean et al., "The KAFE suite," 2018, https://arc.ia2.inaf.it/kafe.php (08 May 2018).
8. "Australian Telescope Compact Array," http://www.narrabri.atnf.csiro.au/ (15 March 2013).
9. NRAO, "Karl G. Jansky Very Large Array," https://science.nrao.edu/facilities/vla (28 February 2018).
10. CARMA, "Combined Array for Research in Millimeter-wave Astronomy," https://www.mmarray.org/ (10 May 2018).
11. "Very Long Baseline Array," https://science.lbo.us/facilities/vlba.
12. "SAOImage DS9," http://ds9.si.edu/site/Home.html.
13. R. Gooch, "Karma: a visualization test-bed," in *Astronomical Data Analysis Software and Systems V*, G. H. Jacoby and J. Barnes, Eds., Astronomical Society of the Pacific Conference Series, Vol. **101**, p. 80 (1996).
14. "Cube analysis and rendering tool for astronomy," http://cartavis.github.io/.
15. J. P. McMullin et al., "CASA architecture and applications," in *Astronomical Data Analysis Software and Systems XVI*, R. A. Shaw, F. Hill, and D. J. Bell, Eds., Astronomical Society of the Pacific Conference Series, Vol. **376**, p. 127, Tucson, Arizona (2007).
16. E. Jones et al., "SciPy: open source scientific tools for Python," 2001, http://www.scipy.org/ (10 May 2018).
17. J. D. Hunter, "Matplotlib: a 2D graphics environment," *Comput. Sci. Eng.* **9**(3), 90–95 (2007).
18. S. van der Walt, S. C. Colbert, and G. Varoquaux, "The NumPy array: a structure for efficient numerical computation, computing in science and engineering," *Comput. Sci. Eng.* **13**, 22–30 (2011).
19. W. McKinney, "Data structures for statistical computing in python," in *Proc. of the 9th Python in Science Conf.*, pp. 51–56 (2010).
20. T. Robitaille and E. Bressert, "APLpy: Astronomical Plotting Library in Python," Astrophysics Source Code Library (2012).
21. The Astropy Collaboration et al., "The Astropy project: building an inclusive, open-science project and status of the v2.0 core package," arXiv:1801.02634 (2018).
22. The Astropy Collaboration et al., "Astropy: a community Python package for astronomy," *Astron. Astrophys.* **558**, A33 (2013).
23. A. Ginsburg et al., "Astroquery v0.1," (2013).
24. E. Liuzzo et al., "ALMA FITS header keywords: a study from the archive user perspective," (submitted to ALMA Memo Series).
25. M. Brooks and S. Marron, "Asymptotic optimality of the least-squares cross-validation bandwidth in kernel estimates of intensity functions," *Stochastic Processes Appl.* **38**, 157–165 (1991).
26. S. Burkutean, "KAFE v1.0 cookbook," 2018, https://arc.ia2.inaf.it/documents.html (26 April 2018).
27. C. D. Brinkerink et al., "ALMA and VLA measurements of frequency-dependent time lags in Sagittarius A*: evidence for a relativistic outflow," *Astron. Astrophys.* **576**, A41 (2015).
28. M. Bonato et al., "ALMACAL IV: a catalogue of ALMA calibrator continuum observations," (MNRAS in press).

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes,-Instruments,-and-Systems on 03 Dec 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

29. "NASA/IPAC extragalactic database," https://ned.ipac.caltech.edu/ (10 May 2018).
30. C. J. R. Clark et al., "DustPedia: multiwavelength photometry and imagery of 875 nearby galaxies in 42 ultraviolet-microwave bands," *Astron. Astrophys.* **609**, A37 (2018).
31. M. Taylor, "TOPCAT & STIL: Starlink Table/VOTable processing software," in *Astronomical Data Analysis Software and Systems XIV*, P. Shopbell, M. Britton, and R. Ebert, Eds., Astronomical Society of the Pacific Conference Series, Vol. **347**, p. 29, Pasadena, California (2005).
32. V. Galluzzi et al., "ALMA band 3 polarimetric follow-up of a complete sample of faint PACO sources," in preparation.
33. V. Galluzzi et al., "Characterization of polarimetric and total intensity behaviour of a complete sample of PACO radio sources in the radio bands," *Mon. Not. R. Astron. Soc.* **475**, 1306–1322 (2018).
34. G. Sabatini et al., "Unveiling the inner morphology and gas kinematics of NGC 5135 with ALMA," *Mon. Not. R. Astron. Soc.* **476**, 5417–5431 (2018).
35. M. Massardi et al., "Chandra and ALMA observations of the nuclear activity in two strongly lensed star-forming galaxies," *Astron. Astrophys.* **610**, A53 (2018).

Biographies for the authors are not available.

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes,-Instruments,-and-Systems on 03 Dec 2020
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use