`rPICARD`: A CASA-based Calibration Pipeline for VLBI Data

I. Calibration and imaging of 7 mm VLBA observations of the AGN jet in M87

M. Janssen[1], C. Goddi[1,2], I. M. van Bemmel[3], M. Kettenis[3], D. Small[3], E. Liuzzo[4], K. Rygl[4], I. Martí-Vidal[5], L. Blackburn[6], M. Wielgus[6,7], and H. Falcke[1]

[1] Department of Astrophysics/IMAPP, Radboud University Nijmegen, PO Box 9010, 6500 GL Nijmegen, The Netherlands
e-mail: M.Janssen@astro.ru.nl
[2] ALLEGRO/Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, The Netherlands
[3] Joint Institute for VLBI ERIC, Dwingeloo, Postbus 2, 7990 AA, The Netherlands
[4] Istituto di Radioastronomia (INAF - IRA), Via P. Gobetti 101, I-40129 Bologna, Italy
[5] Observatorio de Yebes - IGN, Cerro de la Palera S/N, 19141 Yebes (Guadalajara), Spain
[6] Harvard-Smithsonian Center for Astrophysics, 60 Garden St., Cambridge, MA 02138, USA
[7] Black Hole Initiative at Harvard University, 20 Garden St., Cambridge, MA 02138, USA

ABSTRACT

Context. With the recent addition of a fringe-fitter, the Common Astronomy Software Application (CASA) software suite, the current state of the art package for radio astronomy can now reduce Very Long Baseline Interferometry (VLBI) data.
Aims. Here, we present the Radboud PIpeline for the Calibration of high Angular Resolution Data (`rPICARD`) – an open-source VLBI calibration and imaging pipeline built on top of the CASA framework. The pipeline is capable of reducing data from different VLBI arrays. It can be run non-interactively after only a few non-default input parameters are set and deliver high-quality calibrated data. CPU scalability based on a Message Passing Interface (MPI) implementation ensures that large bandwidth data from future arrays will be processable within reasonable computing times.
Methods. Phase calibration is done with a Schwab-Cotton fringe-fit algorithm. For the calibration of residual atmospheric effects, optimal solution intervals are determined based on the signal-to-noise ratio (SNR) of the data for each scan. Different solution intervals can be set for different antennas in the same scan to increase the number of detections in the low SNR regime. These novel techniques allow `rPICARD` to calibrate data from different arrays, including high-frequency and low-sensitivity arrays. The amplitude calibration is based on standard telescope metadata and a robust algorithm can solve for atmospheric opacity attenuation in the high frequency regime. Standard CASA tasks are used for CLEAN imaging and self-calibration.
Results. In this work we demonstrate `rPICARD`'s capabilities by calibrating and imaging 7 mm Very Long Baseline Array (VLBA) data of the central radio source in the M87 galaxy. The reconstructed jet image reveals a complex collimation profile and edge-brightened structure, in accordance with previous results. A potential counter-jet is detected, having 10 % of the approaching jet's brightness, which constrains jet speeds close to the radio core to about half the speed of light for small inclination angles.

Key words. Atmospheric effects – Techniques: high angular resolution – Techniques: interferometric – Methods: data analysis

1. Introduction

Very Long Baseline Interferometry (VLBI) is an astronomical observing technique used to study radio sources at very high angular resolution, down to milli- and micro-arcsecond scales. Global VLBI uses a network of radio telescopes as an interferometer to form a virtually Earth-sized telescope. The large distances between telescope sites impose the need of recording the radio wave signals and the use of independent and very precise atomic clocks at individual VLBI stations which allows to cross-correlate the signals between all pairs of antennas post-facto. The lack of real-time synchronization, the typical sparsity of VLBI arrays, and the fact that the signals received by each ground station are distorted by unique local atmospheric conditions make the process of VLBI data calibration especially challenging. At the correlation stage, these effects are partially corrected for with a model of station locations, source positions, the earth orientation and atmosphere, tides, ocean loading, and relativistic corrections for signal propagation (see Thompson et al. 2017; Taylor et al. 1999, for example). But these models are never perfect and the data must be calibrated in a post-correlation stage to correct for residual errors. All further references to calibration procedures in this manuscript implicitly refer to post-correlation calibration.

For decades, the Astronomical Image Processing System (AIPS) (e.g., Greisen 2003) has been the standard package to calibrate radio-interferometric and VLBI datasets. However, official support for AIPS has now been discontinued, and its successor, the Common Astronomy Software Application (CASA) (McMullin et al. 2007) package has become the main tool for the calibration and analysis of radio-interferometric data in the recent years. Nevertheless, AIPS continued to be the standard package for VLBI data reduction, since CASA was missing a few key VLBI calibration functions – most notably a fringe-fitting task. The missing functionalities have now been added to CASA and the calibration framework has been augmented by a global fringe-fitting task, thanks to an initiative from the BlackHoleCam (BHC) (Goddi et al. 2017) project in collabora-

tion with the Joint Institute for VLBI ERIC (JIVE) (van Bemmel 2019). This task is based on the Schwab & Cotton (1983) algorithm and is similar to the FRING task in AIPS. CASA presents some clear advantages over AIPS: a) an intuitive IPython interface implies a low learning curve for the new 'python generation' of radio astronomers (Momcheva & Tollerud 2015); b) software and data structure are designed to facilitate batch processing, providing much more control and flexibility for pipeline-based data processing compared to AIPS, even when combined with the ParselTongue python framework (Kettenis et al. 2006), and c) strong community support, largely by the Atacama Large Millimeter/submillimeter Array (ALMA) (Wootten & Thompson 2009) plus Karl G. Jansky Very Large Array (VLA) (Thompson et al. 1980) userbases, ensures the development and maintenance of a healthy software, with quick bug detection and adjustments to the most recent needs of the community. Under these conditions, it is natural to expect that CASA will soon become the standard package also for VLBI data processing.

While traditionally only raw data taken by a radio interferometer was delivered to the principle investigator (PI), new-generation facilities like ALMA provide raw data along with calibration tables obtained by running automated calibration pipelines. Next-generation facilities like the Square Kilometre Array (SKA) (e.g., Dewdney et al. 2009) will generate much larger raw data volumes and only fully-pipeline-calibrated datasets and images will be delivered to the PIs. For VLBI experiments, the number of participating stations is typically much smaller than for connected interferometers. This leads to much smaller datasets, so it is feasible to pass all the data on the PI for post-correlation calibration. A notable exception is the European VLBI Network (EVN) (e.g., Porcas 2010), which provides users with a set of pipeline-generated calibration files and diagnostics (Reynolds et al. 2002). The main purpose of this pipeline is a quick assessment of the quality and characteristics of a dataset. Advances in data recording rates and wide-field VLBI capabilities will make it increasingly difficult to reduce VLBI data interactively on single personal machines in the near future. This necessitates data handling methods, where computing power scales with available hardware. Data reduction pipelines are an attractive solution to this problem, as they promote reproducibility of scientific results and circumvent difficulties of VLBI data reduction. As a by-product, this will also attract more astronomers to the field of VLBI.

We have developed a highly modular, Message Passing Interface (MPI)-parallelized, fully-automated VLBI calibration pipeline based on CASA, called the Radboud PIpeline for the Calibration of high Angular Resolution Data (`rPICARD`). The purpose of the pipeline is to provide science-ready data and thereby make VLBI more accessible to non-experts in the community. It should be noted that VLBI data is prone to a large variety of data corruption effects. Some of these effects can not be remedied with calibration techniques and may escape the flagging (removal of corrupted data) methods of `rPICARD`. Depending on the severity of these effects and the required quality of the data for scientific analysis, user interaction to address the data issues may be inevitable. For these cases, `rPICARD`'s verbose diagnostics, tuneability, and interactive capabilities can be used to obtain the required data quality. `rPICARD` v1.0.0 is able to handle

data from any VLBI array given that the raw data is in FITS-IDI[1] or Measurement Set (MS)[2] format.

So far, EVN, Very Long Baseline Array (VLBA) (Napier et al. 1994), Global Millimeter VLBI Array (GMVA) (Krichbaum et al. 2006), and Event Horizon Telescope (EHT) (Doeleman et al. 2009) datasets have successfully been calibrated and imaged. Phase-referencing and simple polarization calibration are supported. Future releases will be able to also handle spectral line observations. CASA-based pipelines are already used for the reduction of ALMA and VLA data[3] and with `rPICARD`, a CASA-based VLBI pipeline is now available as well.

We describe the general CASA calibration scheme in Section 2 and the structure of the `rPICARD` framework in Section 3. A description of the pipeline's calibration strategies follows in Section 4. The implementation of CASA-based automated imaging and self-calibration routines in `rPICARD` are specified in Section 5. A verification of the pipeline based on VLBA data is presented in Section 6. Section 7 gives an overview of future features and a concluding summary is given in Section 8. The necessary information on how to use the code are given in Appendix G.

2. The CASA Calibration Framework

CASA employs Jones matrix algebra (Jones 1941) based on the Hamaker-Bregman-Sault Measurement Equation (Hamaker et al. 1996) to calibrate full Stokes raw complex VLBI visibilities formed at the correlator. In this framework, the ideal, uncorrupted visibilities $\mathcal{V}_{mn}^{\text{true}}$ on baseline m-n, which would be obtained from a perfect measurement device, are related to the measured visibilities $\mathcal{V}_{mn}^{\text{obs}}$ via a 4×4 Jones matrix J_{mn}, which contains all accumulated measurement corruptions on baseline m-n:

$$\mathcal{V}_{mn}^{\text{obs}} = J_{mn}\mathcal{V}_{mn}^{\text{true}} \; . \tag{1}$$

Note that Equation 1 assumes that telescopes are linear measurement devices, so no higher order terms of J_{mn} are considered. Examples of corruptions that factor into J_{mn} are antenna gain errors, antenna bandpasses, and atmospheric phase distortions. We can denote the individual constituents of J_{mn} by g_{mn}^k, where each index k represents a different corruption effect. The order of k should be equal to the reverse order in which the corruptions occur along the signal path, i.e. first the instrumental effects introduced by the signal recording, then effects from the receiving elements, and finally atmospheric signal corruptions. The combined effect of all corruption effects can be represented as

$$J_{mn} = \prod_k g_{mn}^k \; . \tag{2}$$

Lastly, only antenna-based corruption effects are typically removed in the calibration process, as baseline-based effects have a much smaller magnitude and are harder to determine. This means that g_{mn}^k can be rewritten as $g_m^k \otimes \left(g_n^k\right)^*$, where the star denotes a complex conjugate and the $\otimes$ operator represents the tensor product. Therefore, Equation 1 can be written as

$$\mathcal{V}_{mn}^{\text{obs}} = \prod_k \left[g_m^k \otimes \left(g_n^k\right)^*\right]\mathcal{V}_{mn}^{\text{true}} \; . \tag{3}$$

---

[1] See https://fits.gsfc.nasa.gov/registry/fitsidi/AIPSMEM114.PDF for a description of the fits-idi data format.
[2] See https://casa.nrao.edu/Memos/229.html for a description of the current Measurement Set format.
[3] See https://casa.nrao.edu/casadocs-devel/stable/pipeline.

CASA keeps the complex visibility data, together with auxiliary metadata (antennas, frequency setup, system temperatures, etc.), stored locally in a contained form – as binary tables which make up a Measurement Set. The calibration philosophy is to perform incremental calibration based on the inverse of Equation 3 with separate tables for each corruption effect, containing the calibration solutions. These calibration tables have the same structure as the MS and are locally contained binary tables as well. When solving for a new calibration, previous calibration solutions for other corruption effects can be applied 'on-the-fly', meaning that the data are calibrated while they are passed to the solver. The new calibration solutions are therefore relative to the previous ones. Typically, the dominant data corruption effects are calibrated out first. This can be an iterative process if different corruption effects are not sufficiently independent. If no good calibration solution can be obtained, a flagged solution will be written in the calibration table. This generally happens when the signal-to-noise ratio (SNR) of the data is not sufficient to obtain a reliable calibration. Applying flagged solutions to the MS will cause the corresponding data to be flagged. The final goal of the calibration process is to obtain a close representation of $\mathcal{V}_{mn}^{\mathrm{true}}$, by applying all calibration tables to the measured $\mathcal{V}_{mn}^{\mathrm{obs}}$.

It should be noted that station-based calibration is generally an overdetermined problem: For N antennas, there are $N(N-1)/2$ baseline-based visibility measurements.

CASA keeps track of *SIGMA* and *WEIGHT* columns corresponding to the visibility data. *SIGMA* represents the noise within a frequency channel of width $\Delta\nu$ in a time bin Δt. For a single complex cross-correlation visibility data point, this noise is given by

$$\sigma = \frac{1}{\sqrt{2\Delta\nu\Delta t}} \ . \tag{4}$$

CASA will take into account correlator weights when estimating σ. E.g., for data from the DiFX correlator (Deller et al. 2007), weights are determined based on the amount of valid data present in each integration bin Δt when initializing the *SIGMA* column. After that, *SIGMA* will only be modified when data is averaged according to the changes in $\Delta\nu$ and Δt. The *WEIGHT* column is used to weight data according to their quality when averaging within certain bins. The column is initialized based on the initial *SIGMA* column as σ^{-2} for each visibility. After initialization, the product of all applied station-based gains will modify the weights of the baseline-based visibilities. E.g., high system temperature values and channels that roll-off at the edge of the bandpass will be down-weighted in that process. For frequency-dependant weights, the CASA *SIGMA_SPECTRUM* and *WEIGHT_SPECTRUM* columns are used.

The spectral setup of the MS data format, consists of 'spectral windows' (spws), which are sub-divided into frequency channels. Channels are formed at the correlator and determine the frequency resolution of the data. Spectral windows correspond to distinct frequency bands, equivalent to AIPS intermediate frequencies (IFs). These frequency bands are usually formed by telescopes' heterodyne receivers when the high frequency sky signal is mixed with a local oscillator signal. The resulting frequency down-conversion enables analogue signal processing (Thompson et al. 1980). Data from different spectral windows have usually passed through different electronics, so instrumental effects must often be calibrated for each spw separately before combing the data from the full observing bandwidth.

3. `rPICARD`'s Code Structure

`rPICARD` is a software package for the calibration and imaging of VLBI data based on CASA. This section describes the most important features of the pipeline's source code, in particular code philosophy (§ 3.1), input and output (§ 3.2), handling of metadata (§ 3.3), interactive capabilities (§ 3.4), handling of data flag versions (§ 3.5), and the MPI implementation (§ 3.6). `rPICARD`'s license, version control, and a description on how to install the code are given in Appendix G.

3.1. Code Philosophy

The source code is written based on a few maxims:

1. Every parameter can be set in input files; nothing is hard-coded.

2. No parameter needs to be set by hand thanks to sensible- or self-tuning default values.

3. By default, every run of the pipeline is tracked closely with very verbose diagnostic output.

4. Every step of the pipeline can always be repeated effortlessly.

This allows experienced users to tune the pipeline to their needs, while new users will be able to use `rPICARD` to learn about the intricacies of VLBI data reduction. Similarly, the pipeline can be used either for a quick and dirty analysis or to obtain high-quality calibrated data ready for scientific analysis.

`rPICARD` is a highly modular code, with a standardized wrapper for each calibration task. The wrapper automatically takes care of potential smoothing and plotting of calibration solutions as set by the user via a set of standardized input parameters, on-the-fly calibration, and the passing of any required complementary metadata from the Measurement Set (e.g., frequency, antenna, and scan information), in single objects. This makes it straightforward to adjust any calibration strategy or to add new functionalities.

3.2. Input and Output

Input parameters are read in from simple configuration text files. The raw input visibility data from the correlator can either be a MS or a set of FITS-IDI files, which will be imported as MS into CASA with the *importfitsidi* task. Optional metadata files are read in when available as described in Section 3.3.

Additional command line arguments can be used to enable the interactive mode, and to select which pipeline steps are to be redone when experimenting with different, non-default calibration parameters (e.g., SNR cut-offs, interpolation methods of solutions, and selection of calibrator sources).

Finally, a calibrated MS with user-defined spectral- and time-averaging is produced. For backwards comparability with older radio astronomy software packages, UVFITS files will be created from the calibrated MS as well.[4] [5]

---

[4] See https://fits.gsfc.nasa.gov/fits_home.html for a description of the fits file data format.

[5] It should be noted that the UVFITS files produced by CASA can currently not be read into AIPS. This was tested with CASA version 5.4 and AIPS version 31DEC18. The issue is related to different formatting conventions of the 'AIPS AN' extension of the UVFITS file. Difmap (Shepherd et al. 1994) is able to read in the CASA UVFITS files without problems.

3.3. Metadata

A-priori knowledge is crucial for the calibration of VLBI data. `rPICARD` looks for the following files and reads them in automatically as metadata if they exist:

– Standard ANTAB[6] files containing system temperatures (T_{sys}) and station gains which are used for the amplitude calibration (Section 4.3). Some correlators will already attach a T_{sys} table to the raw visibility data itself. Otherwise, custom scripts are used to attach the ANTAB data to the visibilities.

– Files containing stations' receiver temperature information, which increases the accuracy of the correction for the signal attenuation caused by the opacity of the Earth's troposphere. This additional opacity correction is recommended for high frequency observations which do not measure opacity corrected T_{sys} values (Section 4.3). This is the case for the switched-noise calibration method of the VLBA for example.

– Weather tables from antennas' weather stations, which are only needed if an additional opacity correction is to be performed. Usually, these tables are already attached to the visibility data. If the weather information is only present in ASCII format it will be attached to the visibilities with custom scripts.

– Files containing flagging instructions either in the native CASA format or the AIPS-compliant UVFLG format. Files with flagging instructions are typically compiled from stations' observer logs and are handed to PIs together with the correlated data.

– Files containing models of the observed sources. When available, these source models can improve some calibration steps (e.g., fringe-fitting). Usually a-priori source models not needed (Section 4).

– A file with stations' mount-type corrections in case they are not correctly specified in the MS or FITS-IDI files. Station's mount-types are important for the polarization calibration, specifically the feed rotation angle correction (Section 4.5). The feed rotation describes the rotation of the orthogonal polarization receiving elements of a telescope as seen from the sky.

The naming conventions for each of the files mentioned above are described in the code manual.

Additionally, `rPICARD` determines a set of ancillary or 'internal' metadata derived from the MS data itself. This data is stored in a single object for quick access from the different calibration modules. The stored information is about the polarization, frequency setup, data integration time, stations, scans, and observed sources.

3.4. Interactivity Capabilities

For a typical VLBI dataset, the basic flagging mechanisms (Section 4.1) and fringe non-detections (Appendix D) should take care of bad data. For severely corrupted datasets, e.g., with phase instabilities on very short timescales, data dropouts, or unrecoverable correlation errors, user-interaction is inevitable if a high-quality calibration is to be achieved. In `rPICARD`, an interactive mode is enabled with a simple command line flag. In this mode, the software waits for user input to advance to the next step.

This allows for a careful inspection and refinement via manual flagging and/or smoothing of calibration tables. Moreover, it is always possible to flag bad visibilities responsible for erroneous calibration solutions and quickly redo the calibration step. The pipeline provides an external module containing functions for convenient post-processing of calibration tables. It can be loaded into an interactive CASA session and be used in between interactive steps.

3.5. Flag Versions

`rPICARD` keeps track of different flag versions of the visibility data in the MS using the default lightweight CASA `.flagversions` extensions of the MS. Three flagging journals are stored:

1. An initial blank flag version created immediately after the data has been loaded for the first time.
2. A version with all a-priori flags applied (Section 4.1).
3. A final version after all calibration tables have been applied. This version will include the a-priori flags and flags based on failed calibration solutions.

These different flag versions are relevant when re-running parts of the pipeline for an optimization of calibration steps. E.g., the first version can be used when starting from scratch with a new strategy, and the second version can be used when re-running only certain steps with a different set of parameters. The flag version can be specified at the start of a pipeline run.

3.6. MPI Parallelization

With the steadily increasing data recording rates of VLBI arrays, it is necessary for downstream calibration and analysis software packages to scale up with the available computing power, and in particular to fully exploit the parallelism of modern multi-node, multi-core machines. Within CASA, central processing unit (CPU) scalability is implemented via an MPI infrastructure. The MS is partitioned into several sub-MSs, which are virtually concatenated.

In `rPICARD`, the MS is sub-divided across scans. In that way, multiple workers can calibrate multiple scans simultaneously. The most significant speed-up is achieved for the fringe-fitting steps (Section 4.5), where the least-squares globalization steps to go from baseline-based fringes to station-based solutions require significant CPU time, especially for large bandwidths. For other CASA tasks, which are internally parallelized, shorter computing times are achieved as well. The CASA MPI implementation is still being developed, so more and more CASA tasks will be upgraded with an MPI functionality in the future. In Appendix H, we present a test case which demonstrates the significance of the CPU scalability.

An optional input parameter can be set to control the memory usage of the MPI servers – jobs will only be dispatched when enough memory is available. This option is disabled by default because the memory monitoring and resource scheduling will slightly decrease the performance of the pipeline. Moreover, there will be no memory limitations for most VLBI datasets[7] on systems with a reasonable ratio of CPU cores to random-access memory (RAM).

---

[6] The ANTAB format is the current standard used for VLBI data calibration. See http://www.aips.nrao.edu/cgi-bin/ZXHLP2.PL?ANTAB for more information.

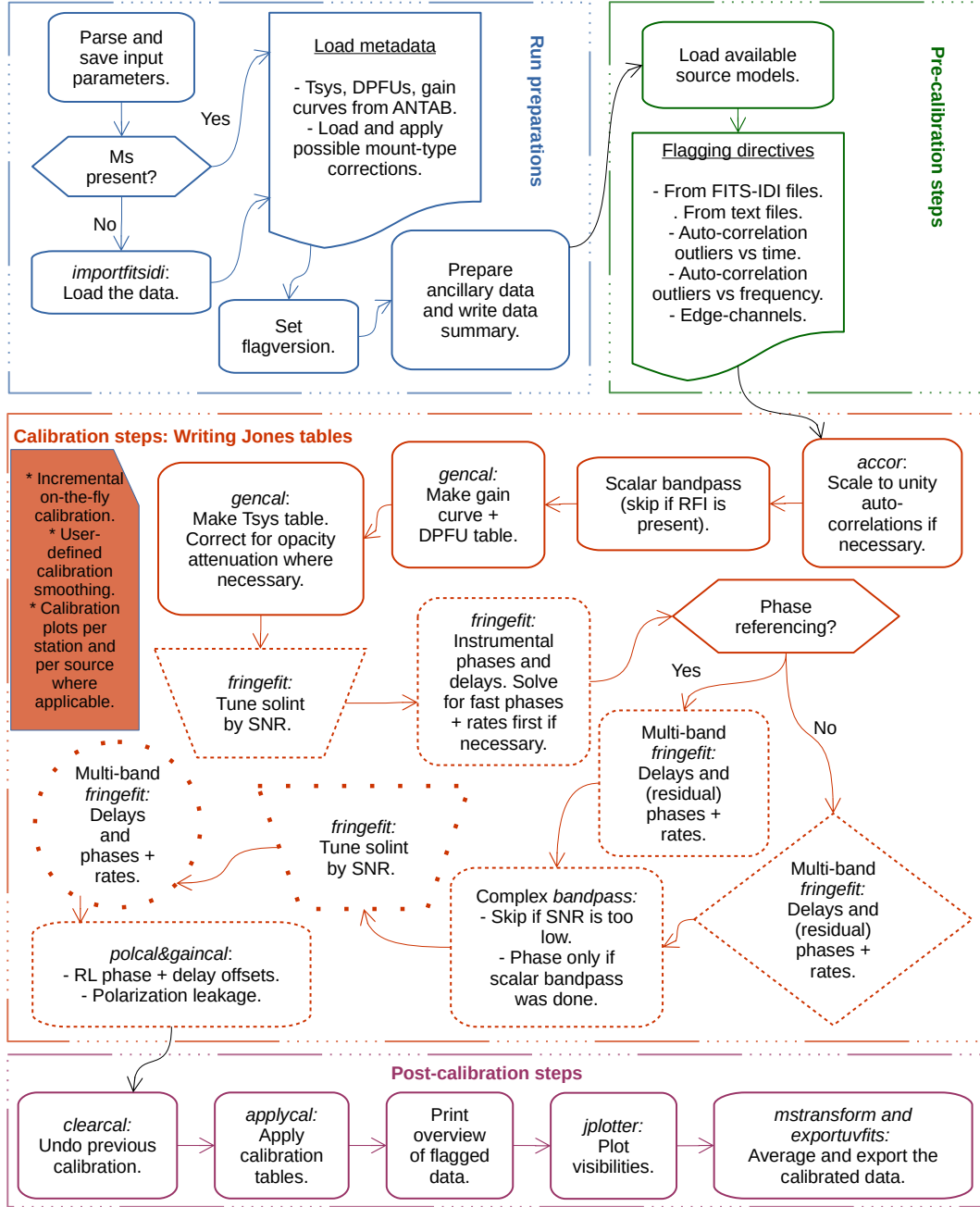[7] Wide-field and very wide-bandwidth VLBI observations may be exceptions.

Fig. 1. Overview of the `rPICARD`'s calibration scheme. CASA tasks used by the pipeline are written in italics. In the orange block in the middle, distinctions are made based on which sources are used to solve for the specific calibration tables and to which sources these tables are applied. Boxes with a solid border belong to calibration steps where all sources were used to obtain solutions. Dashed borders mean that only calibrators were used and dotted borders correspond to solutions obtained from the science targets only. Note that the line-style of the borders only describes the source selection, an averaging of solutions from different sources is not implied. Rectangular shaped boxes indicate solutions that are applied to all sources. Diamonds represent solutions applied to calibrators only, circles are used when the solutions are applied to science targets only, and trapezoids indicate intermediate calibration solutions which are not applied to the data.

4. `rPICARD`'s VLBI Calibration Procedures

This section describes the calibration procedures employed by `rPICARD`. As per standard VLBI data reduction, the user has to specify suitable calibrator sources for the different calibration steps in the pipeline input files. For the calibration of the phase bandpass and other instrumental phase corruption ef-

fects, bright, compact, flat-spectrum sources should be used. For phase-referencing, a calibrator close to the science target on the sky is required. A linearly polarized source tracked over a wide range of feed rotation angles is needed for proper polarization calibration.

An overview of the overall calibration scheme is given in Fig. 1. The standard VLBI calibration techniques introduced in this section, using ANTAB data and fringe-fitting techniques, will be referred to as 'first-order' calibration to distinguish it from the model-dependant additional self-calibration introduced in Section 5. This section presents flagging methods utilized by the pipeline to remove bad visibilities in § 4.1. § 4.2 describes the calibration for digitization effects, § 4.3 outlines the amplitude calibration, § 4.4 describes the amplitude bandpass correction, § 4.5 presents the phase calibration framework, and § 4.6 outlines the methods used for the polarization calibration. Amplitude calibration steps are done first, as they will adjust the data weights and therefore guide the phase calibration solutions (Appendix B).

4.1. Flagging Methods

Bad data can severely inhibit the science return of VLBI data. Some corruptions cannot be calibrated and the affected data should be removed (flagged). Typically this is low SNR data for which no calibration can be obtained by bootstrapping from neighbouring calibration solutions. Examples are data recorder failures and edge channels without signal. `rPICARD` has several independent steps to deal with data flagging:

1. Flags from the correlator are applied, which are typically complied from station's log files, which describe when telescopes are off-source. For the pipeline, correlator flags can come attached to the FITS-IDI input files, or as separate files either as AIPS-compliant UVFLG flags or as a file with flag commands in the CASA format. UVFLG flags are first converted to the CASA flag format with custom scripts.

2. Edge-channels can be flagged if the spectral windows in the data are affected by a roll-off at the edges. The user can specify the number of channels to be flagged. The default is to flag the first and last 5% of channels.

3. Users can add their own flagging statements based on data examinations as flagging files for easy bookkeeping.

4. Finally, experimental flagging algorithms have been implemented in `rPICARD`. These try to find bad data based on outliers in auto-correlation spectra as a functon of time and frequency with respect to median amplitudes.

   In the frequency space, an auto-correlation spectrum is formed by averaging in time over scan durations. Outliers are identified if the running difference in amplitude across channels is larger than the median difference across all channels by a user defined fraction. This analysis is done on a per scan, per antenna, per spectral window, and per parallel-hand correlation basis. The purpose is to identify bad data due to dead frequency channels and Radio-Frequency Interference (RFI).

   Along the time axis, the data is averaged across all channels and all spectral windows. Then, the median amplitude is found and outliers are identified if the amplitude of a data point is lower than the median by a user defined fraction. This analysis is done for each parallel-hand correlation and each station separately. The purpose is to identify bad data due to recorder issues or antennas arriving late on source.

   The derived flagging commands are compiled into flag tables and they are applied to both the auto- and cross-correlations. These experimental flagging algorithms should be used with care as they have only been tested with (and successfully applied to) EHT data.
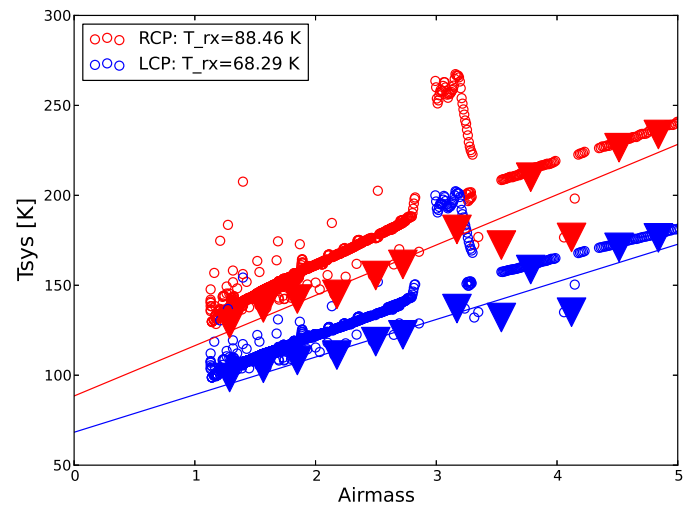

Fig. 2. System temperature from the Brewster VLBA station at 7mm on 04 June 2013 as a function of airmass (given by $1/\sin(\text{elevation})$) for the RCP and LCP receivers (red and blue open circles, respectively). Overplotted are solid triangles from the smallest T_{sys} values in each bin of 0.3 airmass size, which are used to find the receiver temperature by fitting a straight line to the lower T_{sys} envelope (formed by the binning). Small T_{sys} variations induced by weather (e.g., the jump seen at an airmass of ~ 3) do not affect the fit.

The flagging instructions from each step are written as ASCII CASA flag tables and applied to the data with the CASA *flagdata* task prior to each calibration step.

4.2. Sampler Corrections

The analogue signals measured by each station are digitized and recorded for later cross-correlation to form visibilities. Erroneous sampler thresholds from the signal digitization stage are determined with the *accor* CASA task. Correction factors g^{accor} are derived based on how much each station's auto-correlation spectrum deviates from unity.[8]

4.3. Amplitude Calibration

The digital sampling of the received signals at each station causes a loss of information about the amplitude of the electromagnetic waves. The lack of unresolved sources at VLBI resolutions prohibits the use of calibrator sources with 'known' brightness to recover the correct flux densities on all baselines. Instead, the system equivalent flux densities (SEFDs) of all antennas can be used to perform the amplitude calibration. The SEFD of an antenna is defined as the total noise power – i.e. a source with a flux density equal to the SEFD would have a SNR of unity. It can be written as

$$\text{SEFD} = \frac{T_{\text{sys}}e^{\tau}}{\text{DPFU} \cdot \text{gc}} . \tag{5}$$

Here, T_{sys} is the system noise temperature in Kelvin and the DPFU is the 'degrees per flux unit' factor in Kelvin per Jansky (Jy).[9] The DPFU describes a telescope's gain at an optimal elevation, relating a measured temperature to a flux.

---

[8] For some datasets, this correction is not necessary: The SFXC correlator already applies this correction for EVN data for example.

[9] The Jansky unit is defined as 10^{-26} watts per square meter per hertz.

The gc = gc(elevation) variable describes a telescope's gain-elevation curve – a function normalized to unity, that takes into account the changing gain of the telescope due to non-atmospheric (e.g., gravitational deformation), elevation effects. The τ factor is the atmospheric opacity which attenuates the received signal. A e^τ correction enters in the SEFD to represent the signal from 'above the atmosphere' (before atmospheric attenuation). In this way, source attenuation will be accounted for in the amplitude calibration.

In the following, we describe how `rPICARD` can solve for the e^τ atmospheric opacity correction factor. The system noise temperature can generally be written as

$$T_{\rm sys} = T_{\rm rx} + T_{\rm sky} + T_{\rm spill} + T_{\rm bg} + T_{\rm loss} + T_{\rm source} , \qquad (6)$$

with $T_{\rm rx}$ the receiver temperature, $T_{\rm sky}$ the sky brightness temperature, $T_{\rm spill}$ the contribution from stray radiation of the telescope's surroundings, $T_{\rm bg}$ the cosmic microwave background and galactic background emissions, $T_{\rm loss}$ the noise contribution from losses in the signal path, and $T_{\rm source}$ the small contribution from the source. Rewriting the sky brightness temperature in terms of the attenuated actual temperature of the atmosphere $T_{\rm atm}$, yields

$$T_{\rm sky} = \left(1 - e^{-\tau}\right) T_{\rm atm} , \qquad (7)$$

with $\tau = \tau_0 / \sin{\text{(elevation)}}$ the opacity at zenith τ_0, corrected for the airmass to the horizon, which is given as $1 / \sin{\text{(elevation)}}$. Using Equation 7, the dominant terms of Equation 6 can be written as

$$T_{\rm sys} \simeq T_{\rm rx} + \left(1 - e^{-\tau}\right) T_{\rm atm} . \qquad (8)$$

The standard hot-load calibration method of high frequency observatories will measure an 'opacity-corrected' system temperature $T_{\rm sys}^* \equiv T_{\rm sys} e^\tau$ directly (Ulich & Haas 1976), while many other telescopes, e.g., the VLBA stations, will use a switched-noise method (e.g. O'Neil 2002), which does not correct for the atmospheric attenuation. For mm observations, where the signal attenuation is significant, `rPICARD` will solve for e^τ for each opacity-uncorrected $T_{\rm sys}$ measurement, using Equation 8. The receiver temperatures can either be inserted beforehand from a-priori knowledge about a station's frontend or the pipeline will estimate $T_{\rm rx}$ based on the fact that $T_{\rm sys}$ should converge to $T_{\rm rx}$ towards zero airmass. This is done by fitting the lower envelope (to be robust against τ variations) of system temperature versus airmass, exemplified in Fig. 2. This method was adopted from (Martí-Vidal et al. 2012). The other unknown in Equation 8 – the atmospheric temperature corresponding to a system temperature measurement – is estimated based on the Pardo et al. (2001) atmospheric model implementation in CASA with input from local weather stations.

With these assumptions, `rPICARD` can determine $T_{\rm sys}^*$ for the SEFD calculations from each $T_{\rm sys}$ measurement. Outlier values are removed by smoothed interpolation to make the fitting robust even in bad weather conditions. It is possible to perform the additional opacity correction only for a subset of stations in the VLBI array.

Within CASA, polynomial gain curves from the ANTAB files are first multiplied by the DPFUs, then the square root is taken, and finally, the polynomial coefficients are re-fitted with respect to 90 degrees elevation instead of zero. By that 'VLA-type' gain curves are formed. These gain curves and the system temperatures are converted into standard CASA amplitude calibration tables with the *gencal* task.

On a baseline between stations m and n, the amplitude calibration and sampler correction $g^{\rm accor}$ will adjust visibility amplitude $\mathcal{V}_{mn}^{\rm amp}$ as

$$\mathcal{V}_{mn}^{\rm true,\,amp} = g_m^{\rm accor}\, g_n^{\rm accor}\, \sqrt{\text{SEFD}_m \text{SEFD}_n}\, \mathcal{V}_{mn}^{\rm obs,\,amp} . \qquad (9)$$

4.4. Scalar Bandpass Calibration

The data from each spectral windows passes through a bandpass filter. These filters never have a perfectly rectangular passband, which `rPICARD` corrects for by determining an accurate amplitude (scalar) bandpass calibration solutions from the auto-correlations within each spectral window. The auto-correlations are formed by correlating a station's signal with itself, i.e. they are the Fourier Transform of the power spectrum and do not contain phase information. Due to their high SNR, a scalar bandpass solution can be computed for each scan by averaging the data over the whole scan duration and taking the square root of normalized per-channel amplitudes. This is done with a custom python script using basic CASA tools. The user should disable this step if the auto-correlations are affected by RFI. In that case the complex bandpass calibration (Section 4.5.5) can be used to correct the amplitudes.

4.5. Phase Calibration

The large distances between telescopes in VLBI arrays imposes the use of independent local oscillators at the stations and causes the signals to be corrupted by independent atmospheres. When visibilities are formed at correlators, sophisticated models to compute station- and source positions are used to align the wavefronts from pairs of stations (e.g., Taylor et al. 1999). However, the correlator models are never perfect and residual phase corruptions will still be present in the data, These are phase offsets, phase slopes with frequency (delays) and phase slopes with time (rates or 'fringe-rates'). The principal task for any VLBI calibration software is to calibrate out these errors with a fringe-fitting process (e.g., Thompson et al. 2017), so that the data can be averaged in time and frequency to facilitate imaging and model-fitting in the downstream analysis.

In practice, fringe-fitting is used to correct for both instrumental and atmospheric effects. Instrumental effects occur as data from different spectral windows passes through different electronics, causing phase and delay offsets between the spws. This can be modelled as a constant or slowly varying effect over time. The atmosphere causes time-varying phases, delays, and rates. Delays are typically stable over scans, unless the signal path changes, e.g., when clouds pass by or when the source's elevation changes significantly over the scan duration. Phases and rates, as they describe the change of phases with time, can vary on timescales given by the atmospheric coherence time. These coherence times are very short, $O(\text{seconds})$, for mm observations. In the centimetre (cm) regime, the coherence times are $O(\text{minutes})$.

`rPICARD` employs multiple fringe-fitting steps, using the recently added *fringefit* CASA task, which is based on the Schwab & Cotton (1983) algorithm. First, optimal solution intervals for atmospheric effects are determined for the calibrator sources (Section 4.5.1). For mm observations, these solution intervals are used to perform a 'coherence calibration', where intra-scan phase and rate offsets are corrected to increase the coherence times for the calibrator sources (Section 4.5.2). The next step is to correct for instrumental phase and delay offsets (Section 4.5.3). Then, residual atmospheric effects are corrected for
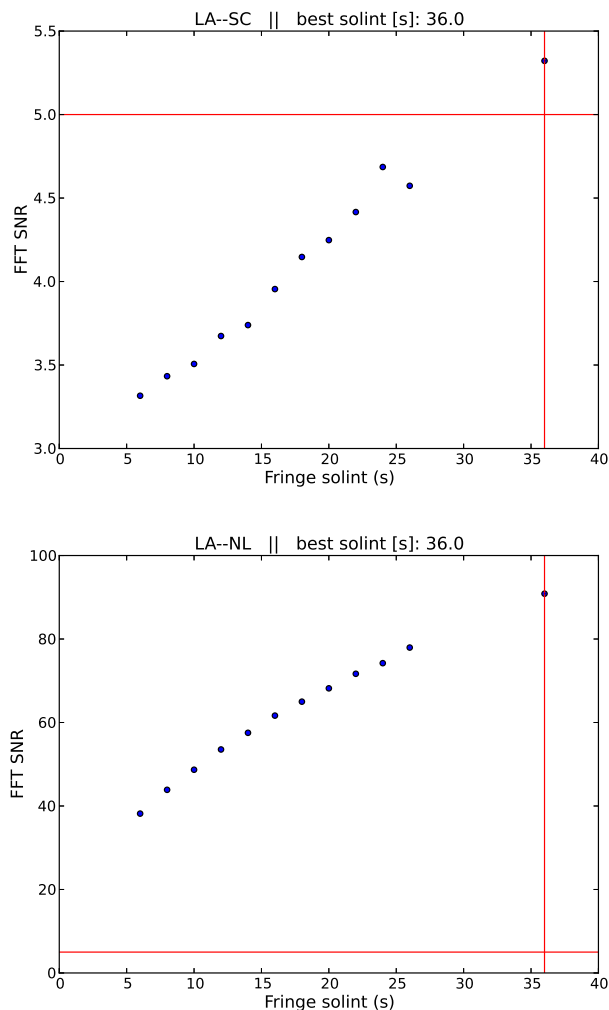
Fig. 3. Example plots of a fringe-fit solution interval optimization search based on 7mm VLBA data (Section 6). Los Alamos (LA) is chosen as the reference station for the scan shown. The blue points show the FFT SNR as a function of solution interval, the horizontal red line indicate the SNR cut-off of 5 on each baseline, and the vertical red line shows the smallest solution interval of 36 seconds needed to reach the SNR cut-off on each baseline for this scan. The top plot is showing the baseline to the Saint Croix (SC) station which has driven the solution interval to larger values until a detection was made when integrating for 36 seconds. The bottom plot shows data from the same scan for the baseline to the North Liberty (NL) station, where the source was detected for each solution interval, yielding the expected increase in SNR with the square root of the solution interval.

the calibrator sources (Section 4.5.4), which finalizes the fringe calibration for the calibrators. If the SNR of the calibrator scans is good enough, corrections for the antenna phase bandpass can be obtained (Section 4.5.5). With all instrumental effects taken out, the phases of the weaker science targets are calibrated (Section 4.5.6). More details about fringe-fitting in CASA are given in Appendix D. It should be that all fringe-fit solutions are obtained after the geometric feed rotation angle phase evolution has been corrected (Appendix C).

4.5.1. Finding the Optimal Solution Intervals for the Calibrator Sources

The solution interval on which the antenna-based corrections are performed is the essential parameter for the phase calibration of VLBI data. The shortest timescales on which baseline phases vary are driven by the coherence times of the atmospheres above the two stations. However, the SNR is often not sufficiently high for the fringe-fitter to find solutions on these short timescales. Therefore, the optimal fringe-fit solution interval should be as close as possible to the phase variation timescales, while being high enough to allow for reliable fringe detections for all stations. It follows, that these intervals strongly depend on the observed source (flux density and structure), antenna sensitivity, weather, and the observing frequency.

`rPICARD` determines phase calibration solution intervals for each scan within an open search range, based on the observing frequency and array sensitivity. The default search ranges are given in Appendix F. The search algorithm uses the SNR from the FFT stage of the fringe-fit process as a metric (Appendix D). The FFTs are determined quickly[10] and an SNR value of five can be taken as indicator for a solid detection. The search stops at the smallest solution interval, where the median SNR on each baseline is above the detection threshold, as shown in Fig. 3. Baselines where the SNR is below the detection threshold are discarded from the search.

In the mm regime, short solutions intervals are used to calibrate for the atmospherically-induced phase fluctuations. Here, it may be desirable to obtain fringe solutions for different antennas on different timescales. Typically, this becomes necessary when some baselines are much more sensitive than others. In that case, fringe solutions within the coherence time can be obtained for some stations, while much longer integration times may be necessary to obtain detections for others. `rPICARD` solves this problem with a two-steps search on two different ranges of fringe-fit solution intervals. The first search is done on short timescales and for each scan, any antenna which does not detect the source on a baseline to the reference station is recorded. In a subsequent search for detections on longer solution intervals, all scans with failed solutions are fringe-fitted again. Detections obtained on the smallest successful solution interval from this second search are used to replace the previously failed solutions. The default search intervals are given in Appendix F.

The *fringefit* task is used to perform the FFT over the full observing bandwidth for the solution interval search. Even when instrumental phase and delay offsets are present in the data, the FFT will still have a decent sensitivity. The optimal solution intervals are determined for the bright calibrator sources first. These sources are always easily detected and for high frequency observations, it is beneficial to calibrate for atmospheric effects before using the calibrator data to solve for instrumental effects. For the science targets, the solution interval search is done after all instrumental effects have been solved, to get more detections in the low SNR regime.

4.5.2. Coherence Calibration for High Frequency Observations

The first step of fringe-fitting typically solves for instrumental phase and delay offsets. This single-band fringe-fitting is done over scan durations to maximize the SNR (Section 4.5.3). For

---

[10] FFTs are limited only by disk I/O, while the least-squares algorithm for a full global fringe-fit requires significant CPU time.

high frequency observations above 50 GHz (e.g., for the EHT and GMVA), the short coherence times degrade the SNR significantly over scan durations. To overcome this problem, `rPICARD` will first solve for phase and rate offsets on optimized atmospheric calibration timescales determined with the method introduced in Sect 4.5.1, before doing the instrumental phase and delay calibration. This 'coherence calibration' fringe search is done for the calibrators with the full bandwidth of the observation, i.e. using data from all spectral windows to maximize the SNR. As instrumental phase and delay offsets have not been solved for yet, potentially spurious solutions for the multi-band delay may come out of the least squares solver. Therefore, no delay solutions are applied from this fringe search; the atmospherically induced multi-band delays are instead corrected in a later fringe search after the instrumental phase and delay calibration. In the cm regime, where the coherence times are much longer, the coherence calibration procedure is not necessary.

4.5.3. Instrumental Phase and Delay Calibration

Depending on telescopes' front- and back-ends, different spectral windows can have different phase and delay offsets. These data corruptions are mostly constant or vary on very long timescales, meaning they can be taken out with a few observations of bright calibrator sources by fringe-fitting the data from each spectral window separately (single-band fringe-fit).

`rPICARD` will integrate over each scan of the brightest calibrators for this calibration step. No distinct time-dependent phase drifts should occur for different spectral windows, as rates are frequency independent, and no single rate solution should be applied to an entire observing track, as an extrapolation of the solution would introduce growing phase slopes with time. Consequently, rate solutions are obtained from the scan-based multi-band fringe-fitting step (Section 4.5.4) and only phase and delay solutions from this instrumental calibration procedure are applied to the data; the single-band rate solutions are zeroed. The instrumental phase and delay solutions are clipped with a high SNR cut-off of 10 by default, to ensure a very small false detection probability and the remaining corrections for delay and phase offsets of each spectral window are interpolated across scans. This allows to correct for small drifts in the electronics for long observations. If the instrumental phase and delay offsets are sufficiently stable, it suffices to use a single scan for the calibration or to pick the single solution with the highest SNR per station and receiver out of all solutions from the calibrator scans. These different options are implemented as a simple input parameter switch in `rPICARD`.

4.5.4. Multi-Band Phase Calibration of Calibrators

After the instrumental phase and delay calibration, the visibility phases are coherent across spectral windows. This allows for an integration over the whole frequency band for an increased SNR to calibrate atmospheric effects. The solutions intervals are determined following the method introduced in Section 4.5.1. A slightly smaller SNR cut-off (between 4 and 5) can be employed for this global fringe-fitting step compared to the threshold on which the optimal solution intervals are based. This ensures that solutions on short timescales are not flagged within scans when the SNR fluctuates around the chosen SNR cut-off for detections. Delay and rate windows (not necessarily centred on zero) can be used for the FFT fringe search to reduce the probability of false detection in the low SNR regime. No windows are put in place

by default, because the sizes of reasonable search ranges are very different for different observations. The user should inspect the fringe solutions obtained by `rPICARD` and, if there are outliers, re-run the fringe-fitting step with windows that exclude the mis-detections or flag the outliers in the calibration table. Since the atmospherically induced multi-band delays are slowly varying, the solutions will be smoothed in time to remove imperfect detections. First, delay solutions outside of the specified search windows are replaced by the median delay solution value of the whole scan.[11] Then, a median sliding window with a 60 seconds width by default is applied to the delay solutions, leaving the fast phase and rate solutions untouched. Depending on the stability of the multi-band delays within scans, different smoothing times can be set in the input files. To ensure constant phase offsets between the parallel-hand correlations of the different polarization signal paths, the same rate solutions are applied to all polarizations. They are formed from the average rates of the solutions from the individual polarizations, weighted by SNR.

The multi-band fringe-fit is done for the calibrator sources first. This can be done in a single pass as the calibrators are easily detected, even on short timescales. The weaker science targets are calibrated after the complex bandpass has been solved for.

4.5.5. Complex Bandpass Calibration

Once the calibrator sources' phase, rate, and delay offsets are calibrated, yielding coherent visibility phases in frequency and time, the SNR should be high enough to solve for the complex bandpass. While the scalar bandpass calibration (Section 4.4) can only solve for amplitudes, the complex bandpass calibration makes use of the cross-correlations and can therefore solve for amplitude and phase variations introduced by each stations' passband filter within each spectral window. `rPICARD` uses the CASA *bandpass* task to solve for the complex bandpass. Single solutions for each antenna, receiver, and spectral window over entire experiments are obtained based on the combined data of all scans on the specified bandpass calibrators. If the SNR is good enough, per-channel solutions can be obtained. Otherwise, polynomials should be fitted. If a prior scalar bandpass calibration was done, the amplitude solutions from the complex bandpass calibration are set to unity. If the amplitudes are to be solved for here, flat spectrum sources should be chosen as bandpass calibrators (this is especially important for wide bandwidths) or an a-priori source model encompassing the frequency structure must be supplied to rPICARD.

If the SNR is too low or not enough data on bright calibrators is available, the complex bandpass calibration should be skipped.

4.5.6. Phase Calibration of Science Targets

After all instrumental phase corruptions are removed with the instrumental phase and delay calibration and complex bandpass solutions, the phases of the weaker science targets can be calibrated. In this step, all previous calibration tables are applied. These are the sampler corrections, the amplitude calibration, the amplitude bandpass, the phase bandpass, the instrumental phase and delay calibration, and calibrator multi-band fringe solutions in the case of phase-referencing (Appendix E).

Three different science target phase calibration paths are implemented in `rPICARD`:

---

[11] The starting point for the fringe-fit globalization step is constrained to lie within the search widows. However, solutions can wander outside of that range during the unconstrained least-squares refinement.

1. If phase-referencing is disabled, a first multi-band fringe search is done over long integration times to solve for the bulk rate and delay offsets with open search windows for each scan. This will determine if the source can be detected or not; fringe non-detections will flag data. Then, intra-scan atmospheric effects are corrected in a second multi-band fringe-fit with narrow search windows following the method presented in Section 4.5.4. This is the default option for high frequency observations.

2. If phase-referencing is enabled and no fringe search on the science targets for residuals is to be done, phases, delays, and rates are calibrated using only the phase-referencing calibrator solutions. All valid data of the science targets will be kept. This is the default option for low frequency observations if the science targets are very weak or for astrometry experiments.

3. If phase-referencing and a search for residuals on the science targets are enabled, the calibration is done with the same two-step fringe-fitting approach as in option 1, while also applying phase-referencing calibrator solutions on-the-fly. Here, very narrow search windows can be used. The goal is to solve for residual phase, delay, and rate offsets not captured by the phase-referencing. This is the default option for non-astrometry experiments in the cm regime with strong enough science targets.

4.6. Polarization Calibration

All calibration solutions described so far are obtained independently for the different polarization signal paths (RCP and LCP for circular feeds). To calibrate the cross-hand correlations, strong, polarized sources must be observed over a wide range of feed rotation angles. This enables imaging of all four Stokes parameters. For circular feeds, the Stokes parameters for total intensity I, linear polarization Q and U, and circular polarization V, are formed as

$$I = \frac{1}{2}(RR + LL) \tag{10}$$

$$Q = \frac{1}{2}(RL + LR) \tag{11}$$

$$U = \frac{i}{2}(LR - RL) \tag{12}$$

$$V = \frac{1}{2}(RR - LL) , \tag{13}$$

with $i = \sqrt{-1}$.

Feed impurities are typically constant over experiments. Therefore, constant solutions are obtained for the polarization calibration by combining the data from all polarization calibrator scans. The CASA *gaincal* task with the type 'KCROSS' setting is used to solve for cross-hand delays and the *polcal* task with the type 'Xf' is used to solve for cross-hand phases. Both tasks determine solutions per spectral window under the assumption of a point source model with 100% Stokes Q flux. This assumption will affect the absolute cross-hand phase which determines the Electric Vector Position Angle (EVPA) orientation across the source – if the source is imaged, adding an arbitrary phase will rotate the EVPA Φ of each linear polarization vector by the same amount. The absolute EVPA can be calibrated using connected-element interferometric or single-dish observations under the assumption that the field orientation remains the same on VLBI

scales. The EVPA follows from the Stokes Q and U fluxes as

$$\Phi = \frac{1}{2}\tan^{-1}\left(\frac{U}{Q}\right) \tag{14}$$

Secondly, the leakage – or 'D-terms' – (Conway & Kronberg 1969) between the two polarizations of each receiver are corrected. Some power received for one polarization will leak into the signal path of the other one and vice versa, causing additional amplitude and phase errors in the cross-hand visibilities. For common VLBI circular feeds, the quarter-wave plates are often the major cause of this effect. Depending on the SNR, `rPICARD` can obtain separate leakage solutions for each spectral window or a single solution for the whole frequency band to correct for the cross-hand visibility errors. It can be assumed that extragalactic synchrotron sources have a negligible fraction of circular polarization[12] and that the parallel-hand correlations are perfectly calibrated, including corrections for stations' feed rotation angle (Appendix C). In this case, small D-terms that can be modelled by a first order approximation for circular feeds on the *m-n* baseline can be written as (Leppänen et al. 1995):

$$RL_{mn}^{obs} = RL_{mn}^{true} + \left(D_m^R e^{2i\chi_m} + \left(D_n^L\right)^* e^{2i\chi_n}\right)I \tag{15}$$

$$LR_{mn}^{obs} = LR_{mn}^{true} + \left(D_m^L e^{-2i\chi_m} + \left(D_n^R\right)^* e^{-2i\chi_n}\right)I . \tag{16}$$

Here, D are the leakage terms, with a superscript indicating the polarization, and subscript indicating the antenna. It can be seen that the effect of the D-terms is a rotation in the complex plane by twice the feed rotation angle, which makes this instrumental effect discernible from the constant true source polarization. The complex D-terms are estimated with the CASA *polcal* task. A SNR-weighted average of the individual calibration solutions will be formed when more than one polarization calibrator is used. This approach is similar to the AIPS PCAL task and it requires a good total intensity source model. However, for strongly resolved sources with varying polarization structure along the different spatial components, *polcal* will fail to determine accurate D-terms. To address this issue, a task similar to AIPS LPCAL will be added to the pipeline in the future (Section 7).

4.7. Post-Processing and Application of Calibration Tables

The solution tables obtained form each calibration steps described in this section can be post-processed with user-defined median or mean filter smoothing, and different options for time and frequency interpolation schemes offered by CASA (nearest, linear, cubic, or spline) can be used for the application of the solutions. The default linear interpolation should be the best option in most cases.

4.8. Diagnostics

By default, `rPICARD` stores many diagnostics from every run in a dedicated folder within the working directory:

– Plots of calibration solutions from each calibration table. Where applicable, plots will be made on a per-source basis.
– Solution interval searches as shown in Fig. 3.
– Receiver temperature fits as shown in Fig. 2.

---

[12] For a typical observation which includes multiple calibrator sources, this assumption is easily tested.

- Plots of visibilities made with the jplotter[13] software. These plots can be made for the raw visibilities and once again after applying the calibration solutions – a comparison of these plots serves as a good metric for the performance of the pipeline. The visibility plots are made for a number of scans which are selected from the start, middle, and end of the experiment based on the number of baselines present. The plots show visibility phases and amplitudes, as a function of time (frequency-averaged) and as a function of frequency (time-averaged).
- Lists of all applied flags and an overview of flagged data percentages per station.
- CASA log files containing detailed information about every step of the pipeline.
- Copies of the used command line arguments and configuration files.

These logging procedures, together with the use of a single set of input files, which determine the whole pipeline calibration process, makes the results fully reproducible. Examples of diagnostic plots from a full calibration run are presented in Appendix A.

5. Imaging and Self-Calibration

The 'first order' calibration described in Section 4 is typically done without a-priori knowledge about the source model. This limits the amplitude calibration to the measured SEFDs, which can have large uncertainties. Values for the DPFUs and gain curves are affected by measurement errors, system temperature variations during scans are often not captured, phasing efficiencies for phased interferometers have uncertainties, and off-focus as well as off-pointing amplitude losses are not captured in most cases. With the default assumption of a point source in the first order calibration, the phase calibration will naturally steer baseline phases to the reference station towards zero (Appendix. D). With the correct source model known, the phases can be calibrated towards the true values and on shorter (atmospheric) timescales.

The basic principle of self-calibration is to image the data to obtain a source model and to use that model to derive station-based amplitude and phase gain solutions, correcting for the aforementioned shortcomings in the first order calibration (Readhead & Wilkinson 1978; Pearson & Readhead 1984). For CLEAN-based imaging algorithms (Högbom 1974), this is often an iterative process. As the source model improves by recovering weaker emission features, better self-calibration solutions are obtained, which will in turn improve the imaging. This process should converge to a final image after a finite number of self-calibration iterations on increasingly shorter solution intervals.

Within rPICARD, imaging and self-calibration capabilities are added as an interactive feature. It is a single module, independent from the rest of the pipeline, that can be used in an interactive CASA session to image sources from any MS or UVFITS file. The CASA *tclean* function is used as MPI-parallelized, multi-scale, multi-frequency synthesis state-of-the-art CLEAN imaging algorithm. The robustness scheme can be used to weight the data (Briggs 1995). The rPICARD imager will perform iterations of imaging and incremental phase plus amplitude self-calibration steps. The CASA *gaincal* task is used to do the self-calibration. Fig. 4 provides an overview of the algorithm steps and Table F.1 in Appendix F describes the most important input parameters.

To avoid cleaning too deeply and thereby picking up source components from the noise, the maximum number of clean iterations is set to a low number for the first few iterations. As the data improves with every self-calibration iteration, the maximum number of clean iterations is gradually increased. Additionally, the cleaning will stop earlier if the peak flux of the residual image is smaller than the theoretically achievable point-source sensitivity of the array or (optionally) when a $3\sigma_n$ stopping threshold has been reached, with σ_n given as 1.4826 times the median absolute deviation in the residual image.

For the self-calibration, SNR cut-offs of 3 for the phases and 5 for the amplitudes are employed to avoid corrupting the data by an erroneous source model. Failed solutions are replaced by interpolating over the good solutions, except for the final self-calibration iteration. This ensures that data without sufficient quality for self-calibration will be excluded from the final image. First, the self-calibration is done for visibility phases, normally starting with a a point source model, which allows for a subsequent time-averaging with reduced coherence losses and results in a better starting point of a bright component at the phase center for the imaging. Next, as the CLEAN model is constructed, a few phase-only self-calibration steps are preformed on increasingly shorter solution intervals to first recover the basic source structure. Then, the self-calibration is used to adjust amplitude gains on long timescales at first to avoid freezing-in uncertain source components. These solution times are then lowered by a factor of 2 for every iteration as the model improves and definite source components are obtained. If necessary, data from certain u-v ranges can be excluded from the amplitude self-calibration, e.g., long baselines with low SNR.

The imaging function can be run entirely interactively, where CLEAN boxes can be placed and updated for each major cycle of each imaging iteration. Once a final set of CLEAN boxes is obtained, the algorithm can be set to run automatically for all remaining iterations. Alternatively, the imaging can be fully automated when a set of CLEAN masks is supplied as input or when the *tclean* auto-masking options are used. This is done with the CASA 'auto-multithresh' algorithm. The auto-masking algorithm will try determine regions with real emission primarily based on noise statistics, and sidelobes in the image. A detailed description of the algorithm can be found online in the official CASA documentation.[14] The default auto-masking parameters set in rPICARD should produce reasonable masks for compact sources, but if faint extended emission has to be recovered, the conservative auto-masking approach may not be sufficient. In that case parameter tweaking or interactive user interaction on top of the auto-masking is required.

The image convergence is easily tracked, as plots of all self-calibration solutions, the model, the data after each self-calibration, and the images (beam-convolved model plus residuals) of each cycle are made by default.

6. A Science Test Case: 7mm VLBA Observations of the AGN Jet in M87

As a demonstration of the rPICARD calibration and imaging capabilities, we present here the results from an end-to-end processing of a representative VLBI dataset. For this purpose, we have selected continuum observations of the central AGN in M87 conducted with the VLBA at 7 mm (43 GHz). The chosen dataset is an eight hour long track on M87, including 3C279 and

---

[13] https://github.com/haavee/jiveplot.

[14] https://casa.nrao.edu/casadocs-devel/stable/imaging/synthesis-imaging/masks-for-deconvolution
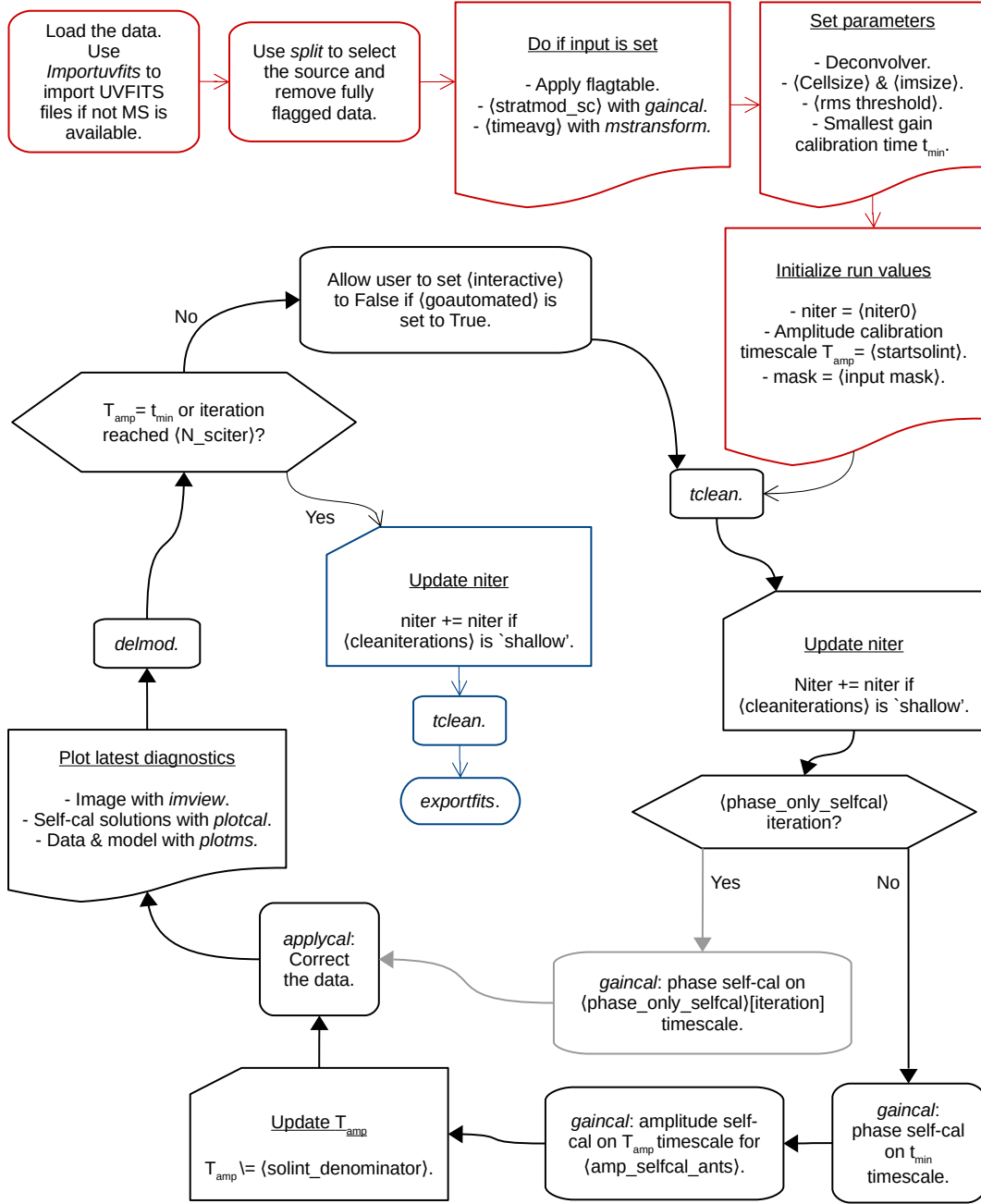
Fig. 4. Overview of the `rPICARD`'s imager performing *tclean* and self-calibration iterations. CASA tasks used by the pipeline are written in italics and input parameters are written inside angle brackets. The solutions from each self-calibration table are incremental with respect to all previous solutions. Interpolated values from good self-calibration results are used instead of failed (flagged) solutions unless an antenna is fully flagged. Optionally, only phase self-calibration can be performed or a single image can be made without any self-calibration. Within each *tclean* operation, the mask is updated between CLEAN cycles – by CASA if 'auto-multithresh' is enabled and/or by the user in interactive mode.

OJ287 as calibrators, with a bandwidth of 256 MHz distributed over two spectral windows with 256 channels each (PI: R. Craig Walker, project code: BW0106). All VLBA stations participated in this experiment.

A description of the data calibration with `rPICARD` and AIPS (to create a benchmark dataset) is given in § 6.1. Imaging and self-calibration steps are descried in § 6.2 and the imaging results are presented in § 6.3.

6.1. First Order Calibration

The data was blindly calibrated with `rPICARD` v1.0.0 in a single pass using the pipeline's default calibration parameters for high frequency VLBA observations. A-priori information was used to flag bad data and to calibrate the amplitudes based on system temperatures, DPFUs, and gain curves. Auto-correlations were used to correct for stations' recorders erroneous sampler
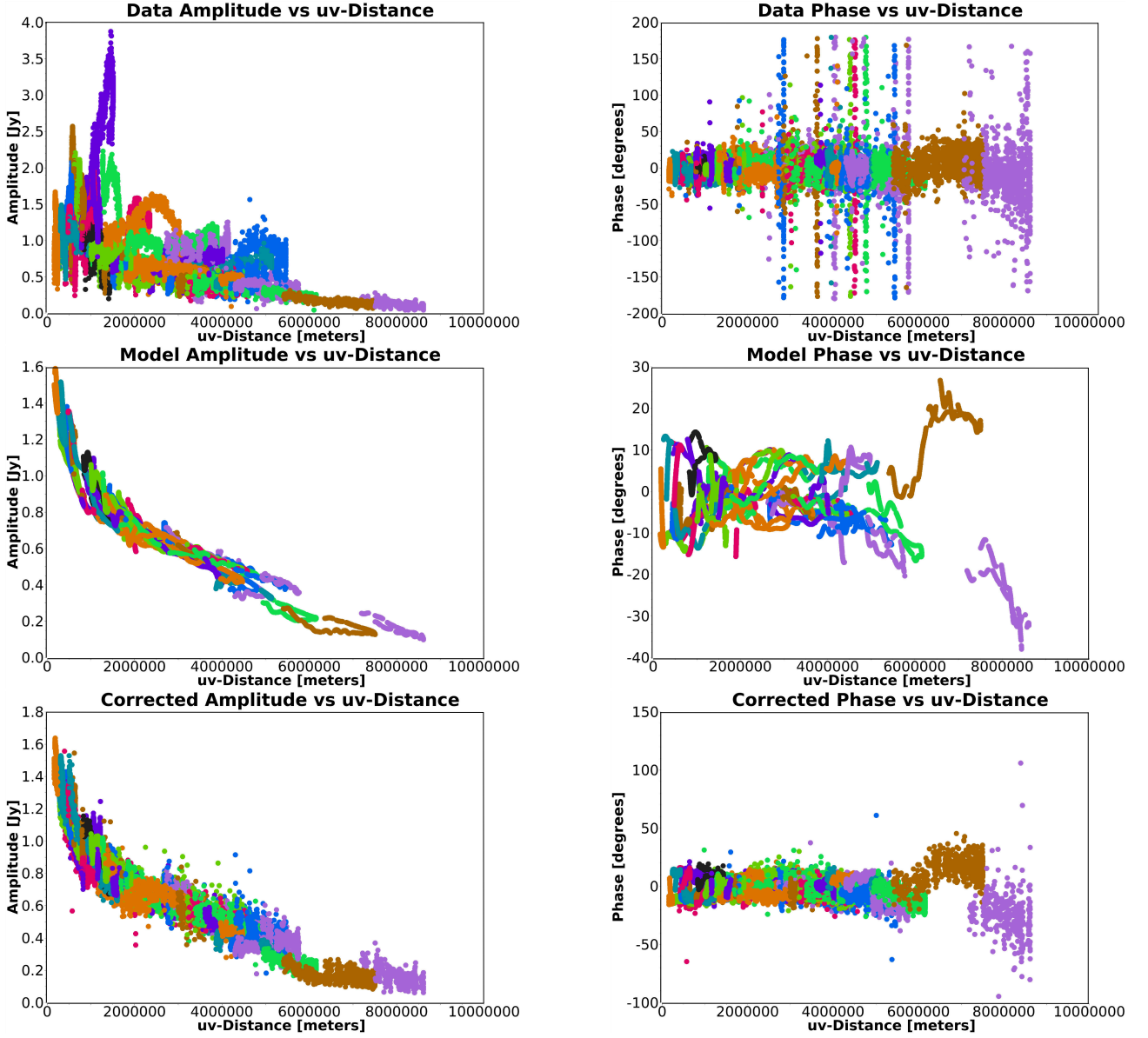
Fig. 5. M87 amplitudes on the left and phases on the right as a function of baseline length (u-v distance), color-coded by baselines. The top panels show visibilities after the first order `rPICARD` calibration (before self-calibration). The middle panel presents the model data from the final image of the imaging plus self-calibration cycles, shown in Fig. 6, The bottom panel depicts the data after the last round of self-calibration. In the first order calibration, anomalously high amplitudes are due to station gain errors; the system temperature measurements of the Fort Davis and Owens Valley stations were doubtful. Large residual phase trends are related to low SNR measurements, where long fringe-fit solution intervals are chosen by `rPICARD`. These amplitude and phase trends are both corrected for with the iterative self-calibration.

thresholds and to calibrate the amplitude bandpasses after edge channels were flagged. Instrumental single-band delay and phase offsets together with phase bandpasses were calibrated using the bright calibrators OJ287 and 3C279. Optimal solution intervals for the fringe-fit phase calibration were determined based on the SNR of each scan.

Independently, the data was manually calibrated in AIPS, following the standard recipe outlined in Appendix C of the AIPS cookbook. In this process, fringe-fitting was done with a 30 s timescale for all scans (the adaptive scan-based solution interval tuning is a unique feature of `rPICARD`). Overall, this integration time yields solid fringe detections, while still capturing the atmospheric phase variations. The AIPS calibrated visibilities serve as a benchmark dataset to cross-compare results between AIPS and CASA/`rPICARD`. Appendix A shows a selec-

tion of `rPICARD` calibration solutions from the different pipeline steps alongside with AIPS calibration solutions for a direct comparison.

6.2. Self-Calibration and Imaging

Both the `rPICARD` and AIPS calibrated datasets are imaged with the `rPICARD` imager using the same default settings, self-calibration steps, and CLEAN windows (Section 5). Initial phase self-calibration steps are done for 300 s and 10 s. After that, the phases are continuously self-calibrated with 10 s solution intervals after each imaging iteration. Only phase gain solutions with SNR > 3 are applied to the data. The first step of amplitude self-calibration used a 2 hour integration, which was lowered by a factor of 2 for each imaging iteration. The final amplitude
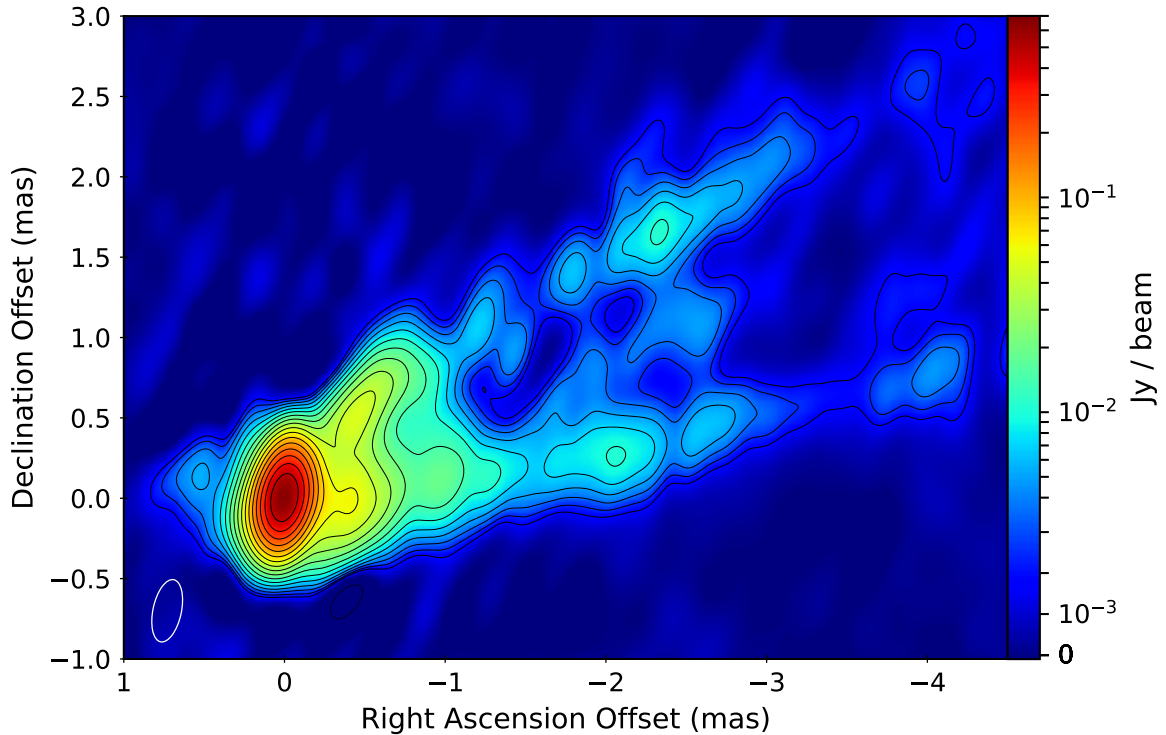
Fig. 6. CLEAN image reconstruction of the 7mm M87 jet from VLBA data using robust weighting of the data (*robust* = 0.5 in *tclean*). The data was calibrated and imaged with `rPICARD`. The restoring beam (0.37 x 0.17 mas at −13.3°), representing the resolving power due to the u-v coverage of the data, is shown in bottom left corner. The color range displayed is −10^{-4} Jy/beam to 0.7 Jy/beam. Contours in mJy per beam are drawn for -1.8, 1.8, 2.5, increasing by factors of $\sqrt{2}$ from there.
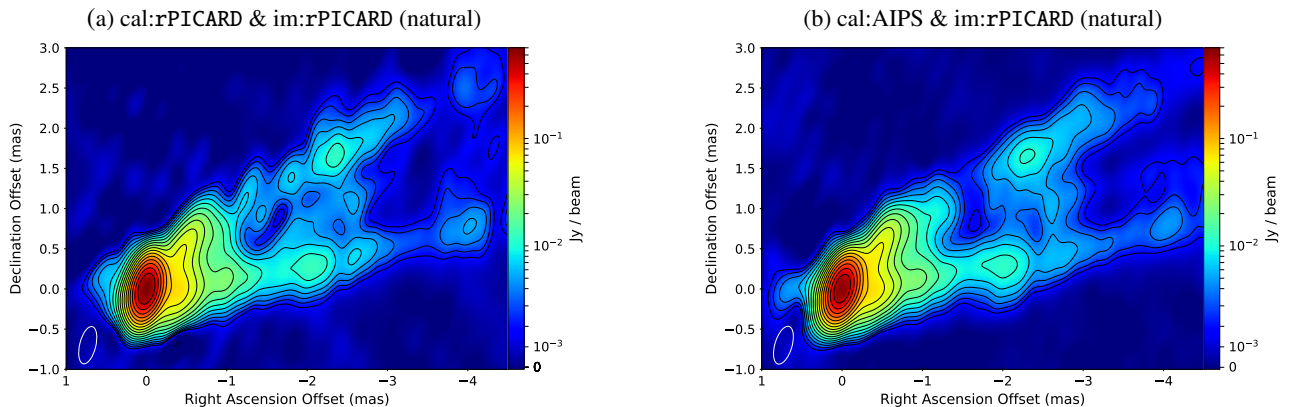
(a) cal:`rPICARD` & im:`rPICARD` (natural)   (b) cal:AIPS & im:`rPICARD` (natural)

Fig. 7. Comparison of CLEAN image reconstructions of the 7mm M87 jet between `rPICARD`-calibrated data in the left panel and AIPS-calibrated data in the right panel. Both images were produced with the `rPICARD` imager using natural weighting. The restoring beam, color rage displayed and contour levels plotted are the same as in Fig. 6 for the two images shown here.

self-calibration step was done on a 10 s solution interval. Only amplitude gain solutions with SNR > 5 are applied to the data.

Fig. 5 shows the effect of the self-calibration after the first order calibration. Before self-calibration, amplitude and phase gain errors are clearly seen (Fig. 5, upper panels) and baseline phases are partially steered towards zero since a point-source model is assumed for the fringe-fitting (Fig. 5, upper right panel). Station-based gain solutions are applied to the data in several self-calibration iterations, which eventually converge to the best source model image (shown in Fig. 6). At the end of the self-calibration process, the corrected visibilities match closely the model visibilities (Fig. 5, lower and middle panels). Note that

at uv-distances > 6$M\lambda$ phases deviate from zero because of the spatial structure of the source.

Fig. 7 shows a direct comparison of image reconstructions for the `rPICARD` and AIPS calibrated data. The overall reconstructed jet structure and aforementioned salient image features are consistent across the two reconstructions. Imaging the same data in Difmap showed no meaningful differences with Fig. 7.

6.3. Results

Fig. 6 shows the fiducial image of the science target obtained after calibration and imaging with `rPICARD`. The bright core of the radio jet in M87 is centred at zero offset in the image. The

western part of the jet is much brighter than the eastern part. This one-sidedness can be explained by Doppler boosting of a jet that is inclined with respect to the line of sight – the approaching jet is brighter than the receding 'counter' jet. Moreover, the edges of the M87 jet are clearly brighter than the central jet 'spine'. This is the result a rotating sheath of material which surrounds the jet spine. The southern jet arm follows an almost straight trajectory, while the track of the northern part of the jet leaves the jet core at a large opening angle of $\sim 50°$ at first. At a right ascension offset of about -1 milli-arcsecond (mas), the upper jet arm bends into a re-collimated trajectory, going parallel to the southern arm as seen on larger scales. The re-collimation at that particular position may be the result of a strong interaction between the jet and the ambient medium surrounding the jet. This structure is also in accordance to earlier work (Walker et al. 2018). A weak emission region is present towards the east of the radio core. The brightness of this feature is about 0.05 Jy – i.e. 10 % of the approaching jet's brightness close to the radio core. When interpreting the weak extension as counter-jet, a symmetrical flow profile would be constrained to about half the speed of light for inclination angles $\lesssim 20°$.

7. Future Features

As rPICARD is built upon CASA, it will benefit from the steady CASA core developments. Notably, these will include enhancements of the Measurement Set data structure, updates of the calibration framework, extensions of the MPI implementation, and imaging improvements.

The CASA fringe-fitter will be able to fit for a dispersive delay for low frequency observations with large fractional bandwidths in the future. This will be important for arrays like LOFAR (van Haarlem et al. 2013) and the SKA.

Additional spectral line specific calibration routines are planned to be included in the next release of rPICARD: A restriction of the bandpass calibration to continuum calibrator sources, a compensation for time-variable Doppler shifts after the bandpass calibration, a delay calibration based only on the continuum calibrators while rate solutions are determined from the spectral line itself, and spectral line amplitude calibration based on the auto-correlations and template spectra of lines.

For an improved polarization calibration, a new task similar to AIPS LPCAL is currently under development. This task will be able to provide more accurate leakage solutions for extended calibrator sources with varying polarization structure. After a source has been imaged in total intensity, it will be possible to decompose a CLEAN model into multiple compact regions with sufficiently constant polarization structure. D-terms can be solved for each region separately and the median leakage will be applied to the data.

The CASA *tfcrop* algorithm will be explored for its capability to automatically identify and flag data corrupted by RFI.

8. Summary

We have presented rPICARD – a CASA-based data reduction pipeline, to calibrate and image VLBI data. rPICARD performs phase calibration using a global Schwab-Cotton fringe-fit algorithm. A robust assessment of the SNR for each scan enables to set optimal fringe-fit solution intervals for different antennas to determine post-correlation phase, delay, and rate corrections. The amplitude calibration is based on standard telescope metadata and a robust algorithm can solve for atmospheric opacity

attenuation in the high frequency regime. rPICARD is agnostic about the details of VLBI data and can run without any user interaction, using self-tuning parameters. Its flexibility allows the user to calibrate VLBI data from different arrays, including high-frequency and low-sensitivity arrays. Standard CASA tasks are used for imaging and self-calibration and fast computing times are achieved by MPI-based CPU scalability.

In order to illustrate the calibration and imaging capabilities of rPICARD, we have selected a 7 mm VLBA observation of the central radio source in the M87 galaxy as a representative VLBI experiment. We have successfully applied the full end-to-end pipeline to this dataset using default parameter settings and produced science-ready results. A qualitative comparison with results obtained from standard techniques with the classic VLBI software suites, AIPS and Difmap, has shown excellent agreement with rPICARD.

References

Briggs, D. S. 1995, in Bulletin of the American Astronomical Society, Vol. 27, American Astronomical Society Meeting Abstracts, 1444
Conway, R. G. & Kronberg, P. P. 1969, MNRAS, 142, 11
Cotton, W. D. 1993, AJ, 106, 1241
Deller, A. T., Tingay, S. J., Bailes, M., & West, C. 2007, PASP, 119, 318
Dewdney, P. E., Hall, P. J., Schilizzi, R. T., & Lazio, T. J. L. W. 2009, IEEE Proceedings, 97, 1482
Doeleman, S., Agol, E., Backer, D., et al. 2009, in Astronomy, Vol. 2010, astro2010: The Astronomy and Astrophysics Decadal Survey
Goddi, C., Falcke, H., Kramer, M., et al. 2017, International Journal of Modern Physics D, 26, 1730001
Greisen, E. W. 2003, in Astrophysics and Space Science Library, Vol. 285, Information Handling in Astronomy - Historical Vistas, ed. A. Heck, 109
Hamaker, J. P., Bregman, J. D., & Sault, R. J. 1996, A&AS, 117, 137
Högbom, J. A. 1974, A&AS, 15, 417
Jennison, R. C. 1958, MNRAS, 118, 276
Jones, R. C. 1941, Journal of the Optical Society of America (1917-1983), 31, 488
Kettenis, M., van Langevelde, H. J., Reynolds, C., & Cotton, B. 2006, in Astronomical Society of the Pacific Conference Series, Vol. 351, Astronomical Data Analysis Software and Systems XV, ed. C. Gabriel, C. Arviset, D. Ponz, & S. Enrique, 497
Krichbaum, T. P., Agudo, I., Bach, U., Witzel, A., & Zensus, J. A. 2006, in Proceedings of the 8th European VLBI Network Symposium, 2
Leppänen, K. J., Zensus, J. A., & Diamond, P. J. 1995, AJ, 110, 2479
Martí-Vidal, I., Krichbaum, T. P., Marscher, A., et al. 2012, A&A, 542, A107
McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 376, Astronomical Data Analysis Software and Systems XVI, ed. R. A. Shaw, F. Hill, & D. J. Bell, 127
Momcheva, I. & Tollerud, E. 2015, ArXiv e-prints [arXiv:1507.03989]
Napier, P. J., Bagri, D. S., Clark, B. G., et al. 1994, IEEE Proceedings, 82, 658
Oliphant, T. E. 2007, Comput. Sci. Eng., 9, 10
O'Neil, K. 2002, in Astronomical Society of the Pacific Conference Series, Vol. 278, Single-Dish Radio Astronomy: Techniques and Applications, ed. S. Stanimirovic, D. Altschuler, P. Goldsmith, & C. Salter, 293–311

Pardo, J. R., Cernicharo, J., & Serabyn, E. 2001, IEEE Transactions on Antennas and Propagation, 49, 1683

Pearson, T. J. & Readhead, A. C. S. 1984, ARA&A, 22, 97

Porcas, R. 2010, in 10th European VLBI Network Symposium and EVN Users Meeting: VLBI and the New Generation of Radio Arrays, 11

Readhead, A. C. S. & Wilkinson, P. N. 1978, ApJ, 223, 25

Reynolds, C., Paragi, Z., & Garrett, M. 2002, Proc. of the XXVIIth General Assembly of the International Union of Radio Science (Gent: URSI), session J8.P.4, paper No. 0924 [`astro-ph/0205118`]

Rogers, A. E. E., Doeleman, S. S., & Moran, J. M. 1995, AJ, 109, 1391

Schwab, F. R. & Cotton, W. D. 1983, AJ, 88, 688

Shepherd, M. C., Pearson, T. J., & Taylor, G. B. 1994, in BAAS, Vol. 26, Bulletin of the American Astronomical Society, 987–989

Taylor, G. B., Carilli, C. L., & Perley, R. A., eds. 1999, Astronomical Society of the Pacific Conference Series, Vol. 180, Synthesis Imaging in Radio Astronomy II

Thompson, A. R., Clark, B. G., Wade, C. M., & Napier, P. J. 1980, ApJS, 44, 151

Thompson, A. R., Moran, J. M., & Swenson, Jr., G. W. 2017, Interferometry and Synthesis in Radio Astronomy, 3rd Edition

Ulich, B. L. & Haas, R. W. 1976, ApJS, 30, 247

van Bemmel, I. 2019, in prep.

Van Der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Comput. Sci. Eng., 13, 22

van Haarlem, M. P., Wise, M. W., Gunst, A. W., et al. 2013, A&A, 556, A2

Walker, R. C., Hardee, P. E., Davies, F. B., Ly, C., & Junor, W. 2018, ApJ, 855, 128

Wootten, A. & Thompson, A. R. 2009, IEEE Proceedings, 97, 1463

Appendix A: Calibration of 7mm M87 BW0106 VLBA Data

Here, we show details about the `rPICARD` and AIPS calibration of the 7mm VLBA data descried in Section 6. Figures A.1 to A.7 were generated with the CASA *plotcal* task (when showing `rPICARD` data) as part of the pipeline's default diagnostics. All plots of calibration solutions shown here are taken from single representative antennas and – except for the multi-band fringe-fit tables presented in Fig. A.6 and Fig. A.7 – show the two spectral windows present in the data color-coded blue and green, respectively, for `rPICARD` data.

A calibration for sampler thresholds (Section 4.2) is shown in Fig. A.1. The corrections are very small and extremely stable in time. A T^*_{sys} calibration table (Section 4.3) is shown in Fig. A.3 for M87. The imprint of the opacity is clearly seen as the source is tracked over a wide range of elevations. Scalar bandpass solutions (Section 4.4) are presented in Fig. A.2 for each scan separately, showing the stability of the bandpass in time. The stability of the instrumental phase and delay calibration (Section 4.5.3) is illustrated by the constant delay offsets between the spectral windows shown in Fig. A.4. Per-channel phase bandpass solutions (Section 4.5.5) are plotted in Fig A.5 for both `rPICARD` and AIPS data. It can be seen how combined solutions from 3C279 and OJ287 have enough SNR to solve for distinct bandpass reposes of a few degrees in a good agreement between AIPS and CASA. Multi-band fringe-fit solutions over full scans (Section 4.5.6) are shown for delays in Fig. A.6 and rates in Fig. A.7 for both `rPICARD` and AIPS. The same fringe solutions are recovered by both calibration packages.
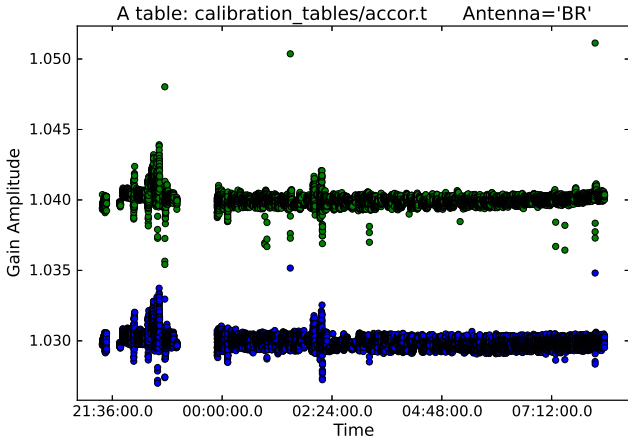

Fig. A.1. *accor* calibration table for the LCP receiver of the Brewster antenna. These solutions correct for sampler threshold offsets per data integration time.

Fig. A.8 shows examples of uncalibrated data from the M87 experiment. Uncorrected post-correlation delays are clearly visible. Fig. A.9 shows the same data after calibration by `rPICARD`. Edge channels are flagged, delays are taken out, bandpasses are flattened, and the phase scatter has been reduced as atmospheric phase variations have been calibrated. There is still some phase scatter present in the Owens Valley – St. Croix baseline due to the SNR-limited timescales on which intra-scan phase variations can be corrected for. The low SNR is also reflected in the amplitudes which are affected by decoherence from the scan-average. The data corresponding to Fig. A.9 after the standard AIPS calibration (Section 6.1) is shown in Figures A.10 and A.11, separately for the RR and LL. The AIPS calibrated data is in good agreement with the `rPICARD` calibrate data.
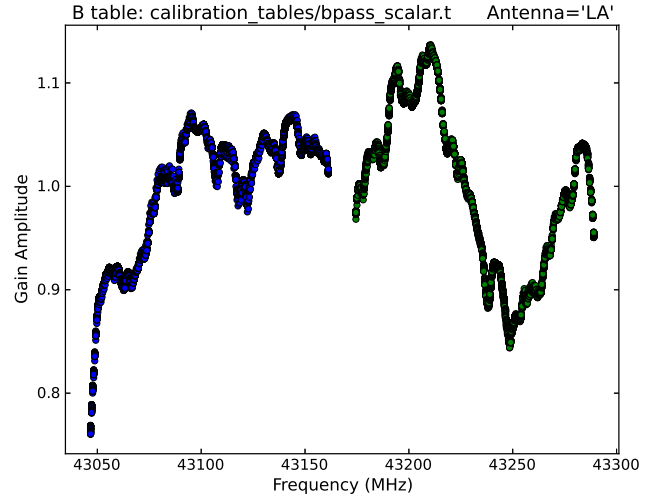

Fig. A.2. Scalar bandpass calibration table for the RCP receiver of the Los Alamos antenna.
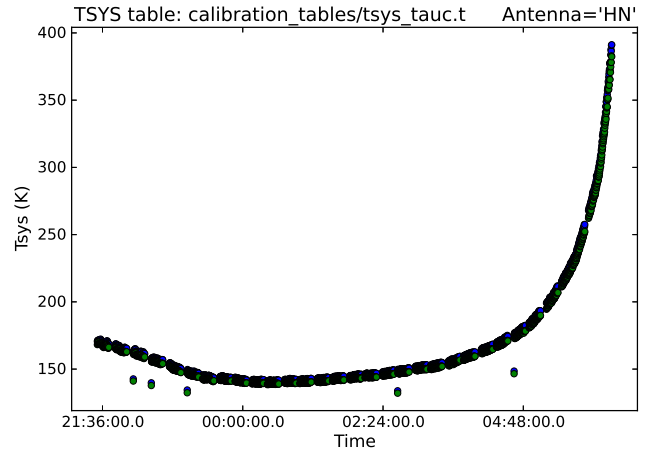

Fig. A.3. Opacity corrected system temperatures for the LCP receiver of the Hancock antenna as M87 is tracked. This calibration table was made by *gencal* using the SYSCAL data from the MS.
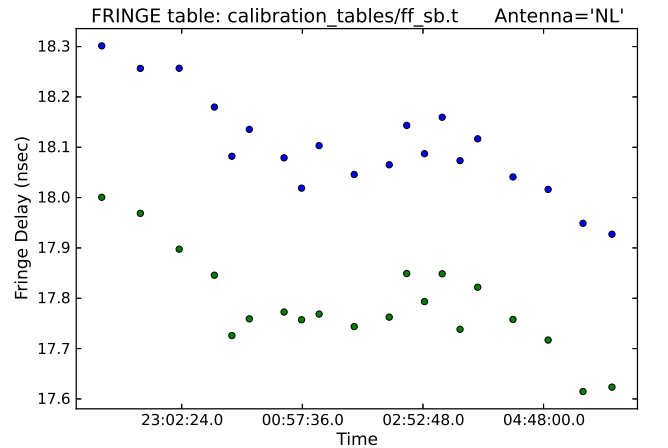

Fig. A.4. *fringefit* calibration table solving for single-band delay offsets for the RCP receiver of the North Liberty station.
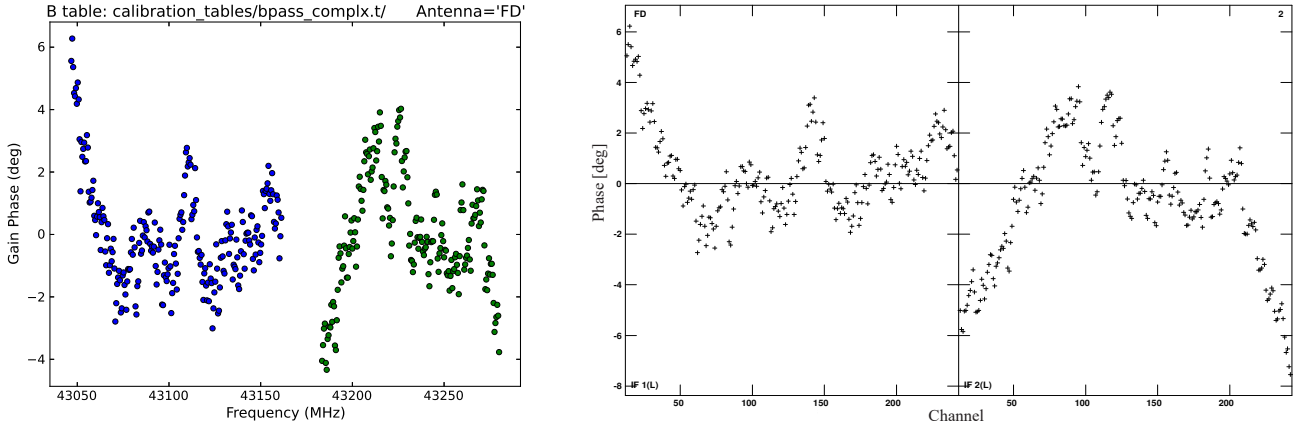
Fig. A.5. Complex *bandpass* calibration tables solving for the phase response within the two spectral windows (IFs) of the Pie Town LCP receiver shown for `rPICARD` calibrated data as a function of frequency (left panel) and for AIPS calibrated data as a function of freqeuncy channel (right panel).
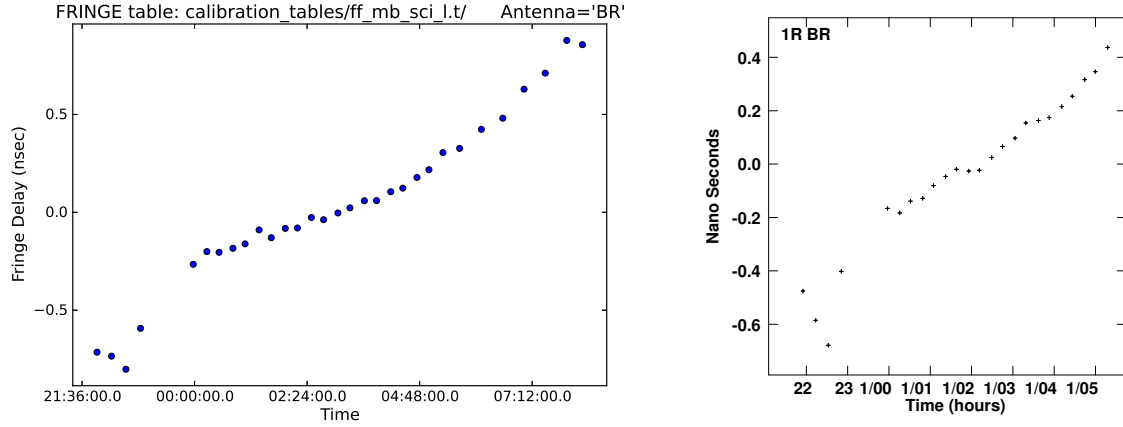

Fig. A.6. Scan-based multi-band fringe-fit delay solutions on M87 of the RCP receiver of the Brewster station shown for `rPICARD` (left panel) and AIPS (right panel).
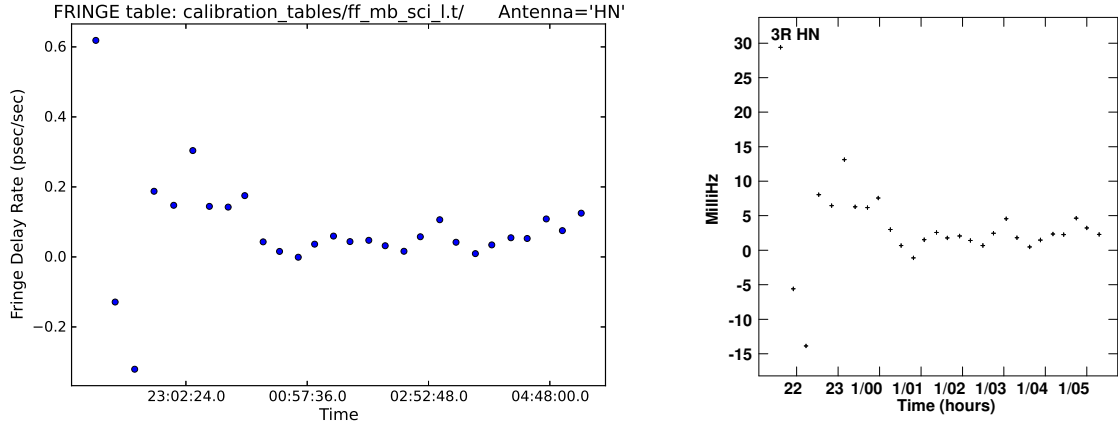

Fig. A.7. Scan-based multi-band fringe-fit rate solutions on M87 of the RCP receiver of the Hancock station shown for `rPICARD` (left panel) and AIPS (right panel). CASA provides dimensionless rates with respect to the reference frequency.
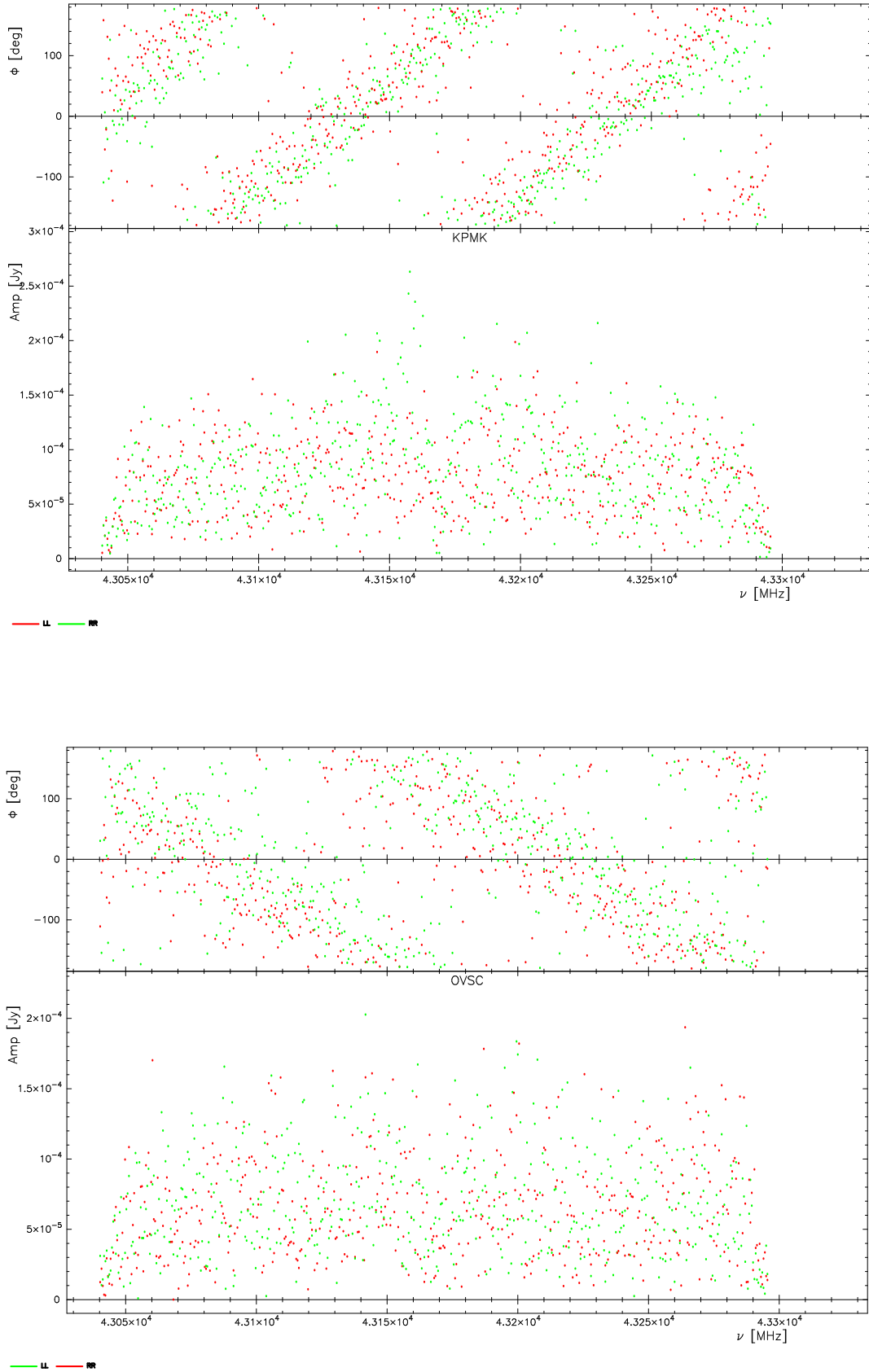
Fig. A.8. Unprocessed raw data of M87. Scan-averaged phases and amplitudes are shown for the two spectral windows in the dataset as a function of frequency for the parallel-hand correlations (color-coded) of two baselines (Kitt Peak – Mauna Kea in the top panel and Owens Valley – St. Croix in the bottom panel). The plots were made with jplotter as part of the diagnostics automatically generated by `rPICARD`.
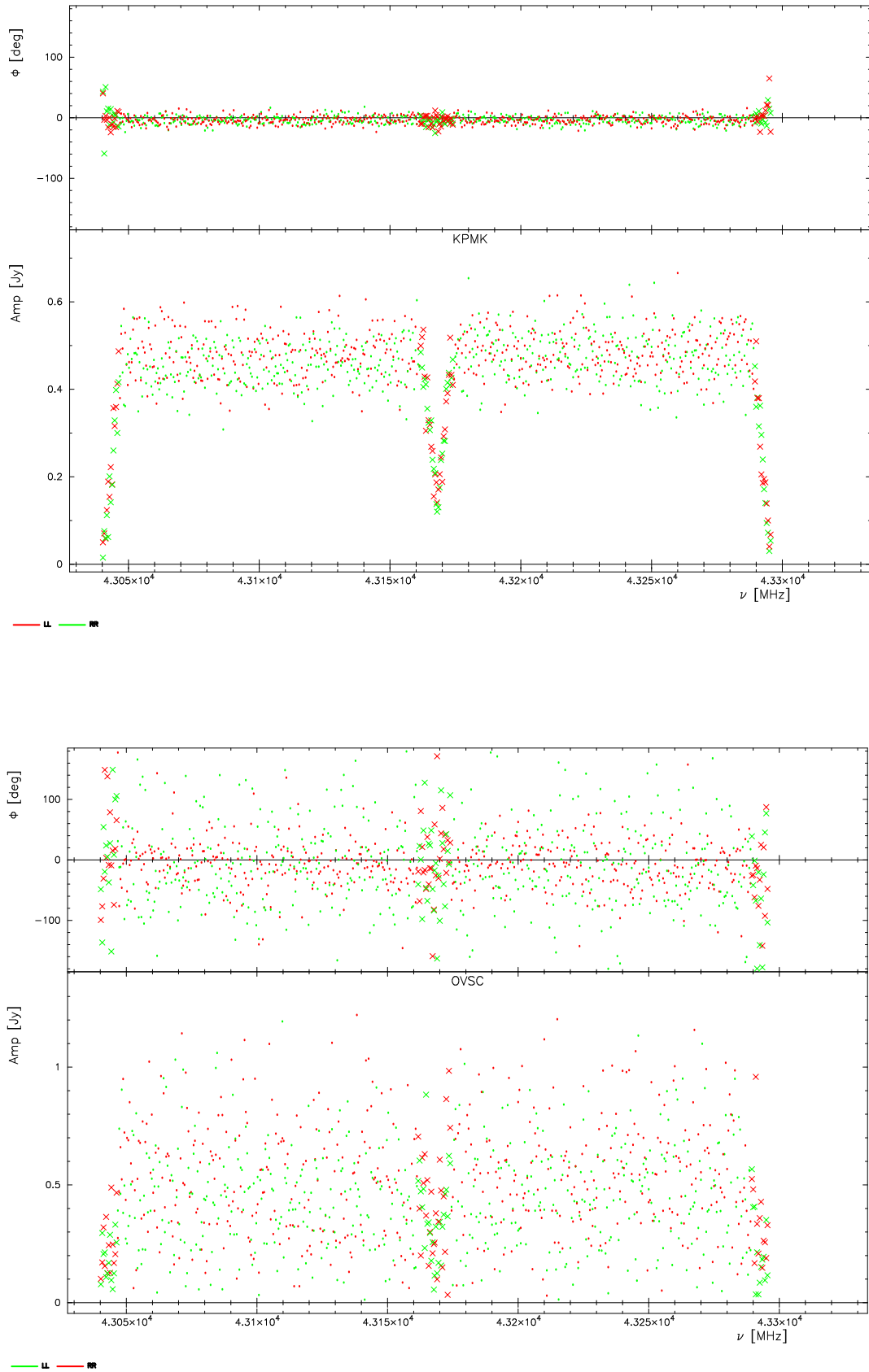
Fig. A.9. rPICARD's pipeline processed data of M87. Scan-averaged phases and amplitudes are shown for the two spectral windows in the dataset as a function of frequency for the parallel-hand correlations (color-coded) of two baselines (Kitt Peak – Mauna Kea in the top panel and Owens Valley – St. Croix in the bottom panel). Crosses indicate flagged data. The plots were made with jplotter as part of the diagnostics automatically generated by rPICARD.
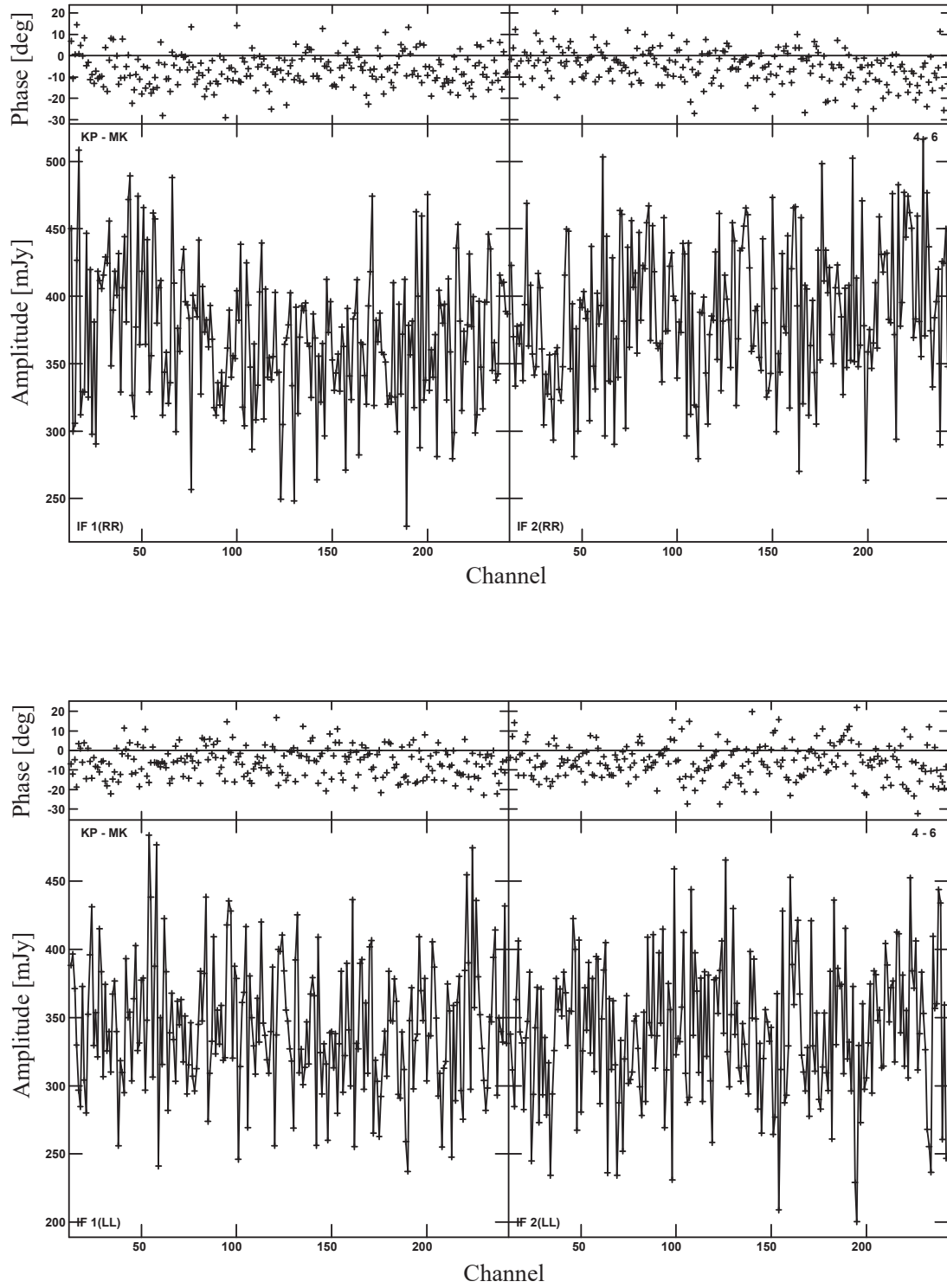
Fig. A.10. AIPS processed data of M87 corresponding to the data shown in the top panel (KP-MK baseline) of Fig. A.9. The AIPS data is shown separately for the RR and LL correlations in the top and bottom panels respectively as a function of frequency channel here.

Fig. A.11. AIPS processed data of M87 corresponding to the data shown in the bottom panel (OV-SC baseline) of Fig. A.9. The AIPS data is shown separately for the RR and LL correlations in the top and bottom panels respectively as a function of frequency channel here.
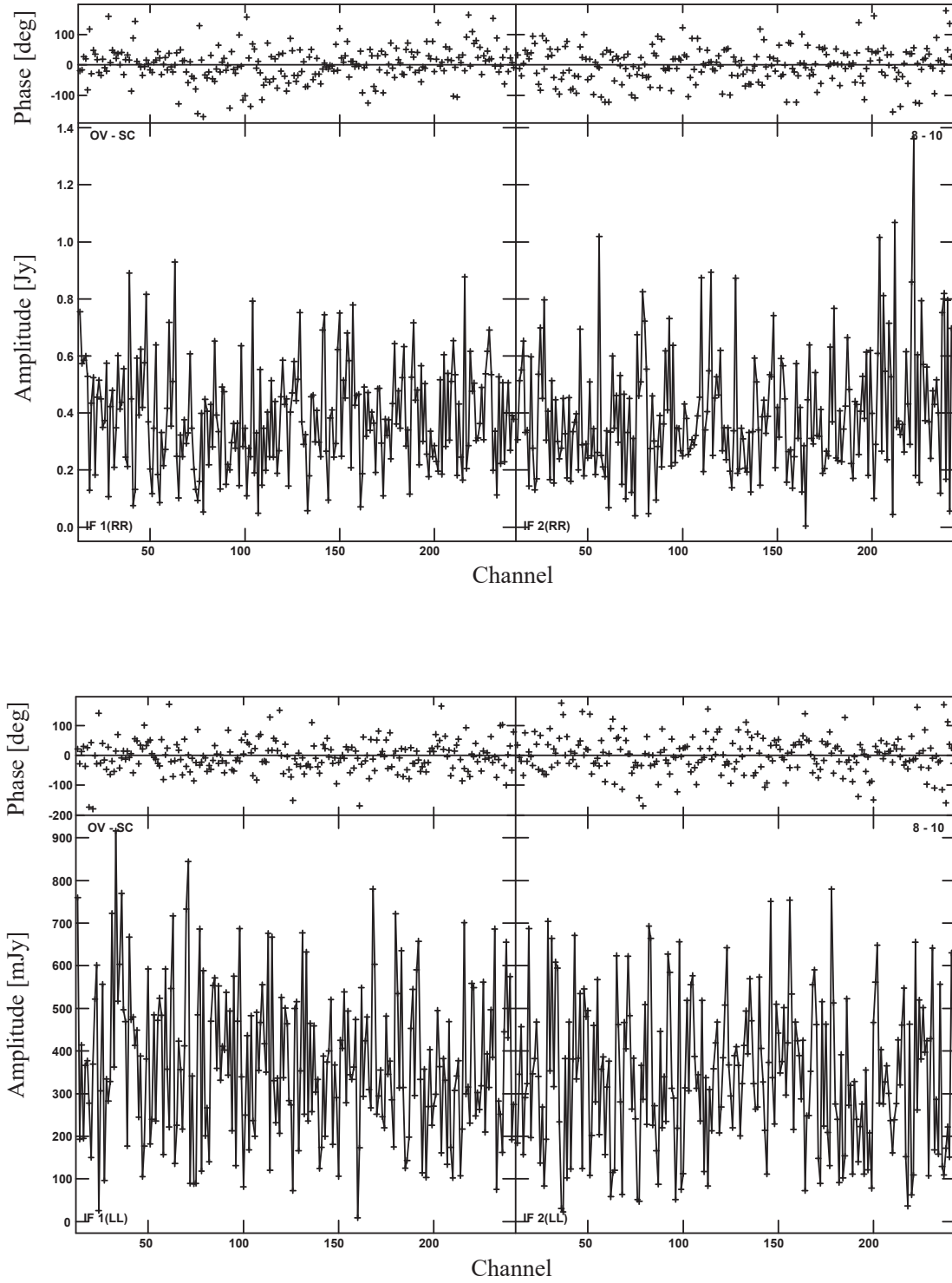
Appendix B: Calibration Weights

The `rPICARD` calibration framework introduced in Section 4 performs incremental calibration steps, based on the CASA calibration scheme. The default choice is to solve for the amplitude calibration solutions first, as these will adjust the WEIGHT column. As follows, the weights of the data within each solution interval will be adjusted, when applying the amplitude solutions on-the-fly for the phase calibration steps. The adjustment of the weights is based on the quality of the data as indicated by the magnitudes of the applied amplitude gains – data which needs large correction factors is considered worse and will be down-weighted. In the time domain, weights are primarily adjusted by system temperatures. For phased arrays, short-timescale phasing efficiencies should enter into the T_{sys} estimates. For single dishes, total power measurements can be used to modulate the system temperature by atmospheric variations within scans. Variations of gain curves have a negligible effect on the weights within the solution intervals. Along the frequency axis, the weights are affected by the amplitude bandpass solution (Section 4.4).

Appendix C: Antenna Mount Types

For the most common telescope mounts, the rotation of the receiver polarization feeds with respect to the source causes a phase evolution χ_f as a function of hour angle. Before fringe-fitting, this simple geometric effect should be corrected, which is done on-the-fly within CASA. This is particularly important for a proper polarization calibration (Section 4.6). Here, an overview of the feed rotation angles for a few common mount types are given. For an altitude-azimuth antenna mount, χ_f evolves with the parallactic angle χ_p as (Cotton 1993):

$$\chi_f^{\text{Alt-Az}} = \chi_p \equiv \tan^{-1}\left(\frac{\cos(\varphi)\sin(h)}{\sin(\varphi)\cos(\delta) - \cos(\varphi)\sin(\delta)\cos(h)}\right),$$

$$(C.1)$$

with φ the station's latitude, δ the declination of the source, and h the source's hour angle. For east-west mounts, the feeds rotate with the co-parallactic angle χ_c as

$$\chi_f^{\text{East-West}} = \chi_c \equiv \tan^{-1}\left(\frac{\cos(h)}{\sin(\delta)\sin(h)}\right).$$

$$(C.2)$$

For Nasmyth mounts, the telescope elevation angle E will enter as an additional term into the phase evolution equation:

$$\chi_f^{\text{Nasmyth}} = \chi_p \pm E.$$

$$(C.3)$$

Here $\chi_p + E$ is the evolution for Nasmyth-right mounts and $\chi_p - E$ is the evolution for Nasmyth-left mounts. For equatorial mounts, the feed rotation angle remains constant:

$$\chi_f^{\text{Equatorial}} = 0.$$

$$(C.4)$$

Appendix D: Fringe-Fitting in CASA

Fringe-fitting is used as the primary phase calibration task by `rPICARD` (Section 4.5). The basic principle is to first Fourier Transform the visibilities as a function of time into the 'rate space' and as a function of frequency into the 'delay space' using a Fast Fourier transform (FFT) (Taylor et al. 1999). If the observed source has been detected on a specific baseline, strong signals will be present at the post-correlation rate and delay offsets. The strength of these signals are characterized by their SNR

as given by their amplitude A_{fringe} in the Fourier domain over the noise σ: $\text{SNR}_{\text{fringe}} = A_{\text{fringe}}/\sigma$. The signal search is done over discrete rate and delay grid cells. Within specified search windows, the cell with the maximum amplitude is picked. Under the assumption that the signal is present in only one of the cells, a false detection occurs if the amplitude in a cell without signal exceeds the signal amplitude (see Thompson et al. 2017, for details). Following Rogers et al. (1995), the probability of false detection P_{error} can be estimated as:

$$P_{\text{error}} \approx n e^{-\text{SNR}_{\text{fringe}}^2/2},$$

$$(D.1)$$

where n denotes the number of discrete rate and delay cells within the search windows. It follows that the FFT $\text{SNR}_{\text{fringe}}$ serves as a good indicator to distinguish detections from non-detections and that a-priori knowledge about the approximate fringe location in the delay-rate search space is important in the low SNR regime.

In practice, the FFTs are obtained for all baselines in a global fringe-fit. `rPICARD` will use a prioritized list of reference stations suitable for the VLBI array to be calibrated, so that different reference antennas can be used for scans where the primary one drops out. The reference station should be sensitive and central in the array to maximize the SNR on each connecting baseline. A SNR cut SNR_{min} is used for $\text{SNR}_{\text{fringe}}$. This determines if an acceptable solution/detection can be obtained for a baseline connecting to a particular station s_i. If $\text{SNR}_{\text{fringe},i} < \text{SNR}_{\text{min}}$, a flagged solution will be written for s_i. All FFT solutions with sufficient SNR are passed as initial estimates to a least least-squares solver, to obtained refined station-based calibration results with a sub-grid resolution (Schwab & Cotton 1983). The least-squares procedure will tune the station gains such that the data matches an assumed source model. In most cases the source model is unknown at the stage of data calibration, so a point source model is assumed. Hence, station gains are adjusted such that baseline phases to the reference station are steered towards zero. It should be noted, that the true source structure is still encoded in the closure phases (Jennison 1958), which are unaffected by the station-based calibration unless the data is averaged in time over long periods. Fringe solutions can be obtained for each receiver polarization channel, over a set integration time, and either per spectral window or for the whole observing bandwidth combined. In the end, all station-based fringe solutions will be re-referenced to a single station with the CASA *rerefant* task to ensure phase continuity. The default choices for reference stations are given in Appendix F.

Appendix E: Phase-Referencing in `rPICARD`

For cm observations, the modest atmospheric influence warrants the use of phase-referencing. With this technique, fringe solutions from a nearby calibrator are transferred to the science target. This enables the calibration of weak science targets, which would be too weak to be detected on their own. Phase referencing is one of the possible phase calibration paths implemented in `rPICARD` (Section 4.5.6). Typically, this is done by interleaving short calibrator scans with science target scans. The length of these scans should not exceed the atmospheric coherence time. Before the fringe transfer, the solutions are first filtered to remove flagged (failed) detections, and smoothed to a single value per scan, antenna, and spectral window.

For astrometry experiments, phase-referencing is essential to obtain accurate positions of the science targets relative to the calibrator sources. If the source position is of no interest and if the

science source is strong enough, it can be fringe-fitted to correct for residual atmospheric effects after the phase transfer. For high frequency VLBI experiments, atmospheric contributions to the phases prohibit phase-referencing unless in-beam calibrators are available.

Appendix F: `rPICARD` Input Parameters

Here, more detailed information is given about some of the input parameters mentioned in the fringe-fitting Section 4.5.1 and imaging Section 5.

The code uses the following prioritized reference station lists for the phase calibration by default:

1. **EHT**: ALMA, LMT, APEX, SMT, IRAM30m, SPT
2. **GMVA**: ALMA, Effelsberg, IRAM30m, Los Alamos, Fort Davis, Pie Town, Kitt Peak
3. **VLBA**: Los Alamos, Fort Davis, Pie Town, Kitt Peak
4. **EVN**: Effelsberg, Yebes, Medicina, Noto

The user could alter these lists based on the actual performance of the stations in an observations. The most important effects that need to be taken into account are weather and technical issues. These will become evident from the diagnostic output of the pipeline; most importantly from plots of system temperatures.

The following default search ranges are used when looking for fringes to calibrate atmospherically induced in-scan phase fluctuations:

1. **EHT**: 2 s – 30 s
2. **GMVA**: 2 s – 180 s
3. **VLBA**: 10 s – 400 s
4. **EVN**: 30 s – 400 s

Depending on the overall weather conditions, exact observing frequency, and source brightness, the user could adjust these search ranges. The best indicator is the number of fringe-non detections within the search windows, which are reported in the diagnostic plots for the fringe-fit solution interval searches. It is possible to specify different search ranges for calibrator sources and science targets. In the mm regime, `rPICARD` will try a few longer solution intervals to get detections for stations in scans outside of the ranges given above. For the EHT, 60 s and 120 s are tried. For the GMVA, 240 s and 300 s are tried.

For both the list of reference stations and phase calibration solution intervals, new lists of solution intervals for other arrays can easily be added.

Table F.1 gives a brief overview of the important parameters for the imager described in Section 5.

Appendix G: The `rPICARD` Source Code

`rPICARD` is hosted on https://bitbucket.org/M_Janssen/Picard. It is an open-source software that can be distributed under the GNU General Public License. Its purpose is to calibrate VLBI data using CASA calibration mechanisms. As it uses the MPI tools within CASA, a native installation is only possible under Linux. Via the Docker[15] software, it is possible to run the pipeline from within system independent containers, allowing the use of `rPICARD` on Windows and MacOS workstations for example. A small how-to guide is provided in the README file of the repository.

---

[15] https://www.docker.com/

The code's version control follows the standard MAJOR.MINOR.PATCH scheme. The current version is the 1.0.0 release. The repository contains a cookbook/manual that describes all features of the pipeline, their implementation, and usage examples. The code itself is further documented with docstrings for all functions in all modules.

A specific CASA branch that contains the latest versions of the VLBI calibration functionalities has to be used for the pipeline. This branch is also open-source; a link to the latest version is given in the README file of the `rPICARD` repository. Eventually, this VLBI branch will be merged into the CASA master branch.

Appendix H: MPI Scalability Test Case

A quick single-core versus multi-core comparison of an eight hour long, 256 MHz bandwidth VLBA experiment (Section 6) demonstrated the significance of the CPU scalability: It took `rPICARD` two hours and three minutes to process the entire dataset sequentially, which was reduced to 34 minutes when using 14 CPU cores in parallel. For the parallelized fringe-fitting steps, a speed-up factor of 6.7 was achieved. This test was done with a Intel(R) Xeon(R) CPU E5-2690 v4 @ 2.60 GHz and a Samsung 960 Pro SSD hard drive connected to four gen-3 PCIe lanes.

However, it should be noted that a general MPI benchmark test has a limited significance, because 1) not all CASA tasks are parallelized, 2) computing speeds are heavily affected by disk I/O (since the datasets are too big to be stored in memory), and 3) large differences in data quality across scans can significantly slow down the multi CPU core speed-up (as a few bad scans can dominate the total computing time).

Table F.1. Description of the most important `rPICARD` imager parameters and how they are updated in each imaging plus self-calibration iteration. For more information about the *tclean* parameter, see https://casa.nrao.edu/casadocs-devel/stable/global-task-list/task_tclean/about.

Parameter Name	Default	Description
cellsize	None	Size of an image pixel. If not set, CASA will determine it based on the longest baseline.
imsize	None	Size of the image in pixels. If not set, CASA will determine it based on the field of view.
niter0	500	Initial value for the maximum number of CLEAN iterations.
cleaniterations	'shallow'	Parameter to update the *tclean niter* parameter, starting from the *niter0* parameter. If set to 'constant', niter will not be updated and for 'shallow', *niter* will be increased by *niter0* after each iteration.
gain	0.1	The CLEAN gain (fraction of source flux subtracted frm residual image).
multiscale	[0,2,6]	Pixel sizes for multi-scale deconvolution.
nterms	1	Number of Taylor coefficients in the spectral model for multi-frequency deconvolution.
robust	0.5	Briggs weighting robustness parameter (Briggs 1995).
threshold	'auto'	Stop cleaning when the residual peak flux reaches this value. The default ('auto') uses the point source sensitivity determined from the data by the CASA *imager* tool.
nsigma	3.0	CLEAN stopping criterion based on the median absolute deviation.
startmod_sc	0	Do an initial phase self-calibration to a point source to align phases. The timescale can be set in seconds. The default (0) is to use the data integration time t_{int}.
timeavg	False	Can average the data in time after *startmod_sc*.
phase_only_selfcal	[300s,10s]	A list of phase-only self-calibration steps to be done before the amplitude self-calibration and after *startmod_sc* and *timeavg*.
startsolint	10 hours	The starting solution interval for the amplitude calibration.
solint_denominator	2	The factor by which the amplitude self-calibration timescale T_{amp} is lowered after each iteration.
N_sciter	0	Maximum number of self-calibration iterations. The default (0) is to stop when T_{amp} reaches t_{int}. If set to -1, only the phases will be self-calibrated. If set to -2, a single image without self-calibration will be made.
minsnr_sc_phase	3.0	The SNR cut for phase self-calibration solutions.
minsnr_sc_amp	5.0	The SNR cut for amplitude self-calibration solutions.
flag_last_sc	True	Flag data for which no self-calibration solutions are obtained in the last iteration.
amp_selfcal_ants	False	Can define a subset of station for which amplitude self-calibration is to be done. The default is to obtain solutions for all stations.
station_constraints	None	Can constrain the range for amplitude self-calibration gains.
station_weights	None	Can modify the data weighting of specific antennas.
uvrange_sc	None	Can constrain the u-v range for which amplitude self-calibration solutions are obtained.
mask	None	Can supply files with CLEAN box masks for the source that are to be used for different iterations.
interactive	True	Build CLEAN boxes interactively in *tclean*.
goautomated	False	Option to set *interactive* to False at runtime for a final set of CLEAN boxes.
usemask	'auto-multithresh'	Enable the *tclean* CLEAN auto-boxing capabilities (see text).