ALMA North America Cycle 3 Study Project Final Report: Extensions and Enhancements to the ALMA Phasing System

L. D. Matthews[1], G. B. Crew[1], & M. H. Hecht[1]

ABSTRACT

The Atacama Millimeter/submillimeter Array (ALMA) Phasing Project (APP) has successfully brought Very Long Baseline Interferometry (VLBI) to ALMA. Nine VLBI science projects were observed in 2017 during ALMA's inaugural VLBI campaign as part of Cycle 4. This marked the culmination of an international 5-year effort that involved both hardware and software contributions from the APP Team to the ALMA Observatory. A Cycle 3 ALMA North America (NA) Study was proposed to enable ongoing support of VLBI at ALMA and the investigation of enhancements to the ALMA Phasing System (APS) that were not within the scope of the original APP project. These included: (1) an extension of phasing capabilities to the submilleter (Band 7); (2) an exploration of correlation techniques to compensate for the mismatch in sampling rates between ALMA and other VLBI stations; (3) prescriptions for optimization of ALMA baseband delay application; (4) defining and documenting data calibration and analysis pathways for experiments utilizing phased ALMA data.

This report summarizes outcomes from the Cycle 3 Study. Work on the APP remains ongoing under a Cycle 4 Study award and will continue under a pending ALMA NA Cycle 5 Development Project that is expected to enable full implementation of the capabilities explored under the Cycle 3 and Cycle 4 Studies.

1. Overview and Context

The ALMA Phasing Project (APP) was conceived to exploit the extraordinary sensitivity of ALMA for VLBI science at millimeter (mm) and sub-mm wavelengths (Doeleman et al. 2010; Matthews et al. 2018; hereafter M18). In addition to operating as a connected element interferometer, ALMA can function as the equivalent of a single very large aperture

[1]Massachusetts Institute of Technology Haystack Observatory, 99 Millstone Road, Westford, MA 01886 USA

antenna if the data from its individual antennas are phase-corrected and coherently added. The goals of the APP were to provide the hardware and software necessary to perform these functions, to record the resulting data in a format suitable for VLBI, and to provide the infrastructure needed for ALMA to function as a VLBI station.

An optimally phased array provides a collecting area equivalent to the combined effective area of the individual antennas. Thus the addition of a phased ALMA to existing mm VLBI networks, including the Global mm-VLBI Array (GMVA) and the Event Horizon Telescope (EHT) network, offers up to an order of magnitude boost in sensitivity. The geographical location of ALMA also provides a significant improvement in u-v coverage for many experiments, thereby enhancing the ability to reconstruct images of sources. In the case of the GMVA, the addition of ALMA improves the north-south angular resolution by a factor of two compared with their present array.

The potential science applications of a beamformed ALMA are wide-ranging (Fish et al. 2013; Tilanus et al. 2014; Asada et al. 2017). For example, by providing angular resolution of $\sim$30 μs, VLBI observations of nearby supermassive black holes at mm wavelengths are able to probe event-horizon scales and can be used to study accretion onto black holes in unprecedented detail and to offer stringent new tests of the predictions of General Relativity. High-resolution imaging of jets in larger samples of active galactic nuclei are expected to illuminate jet formation and collimation processes, and observations of high brightness temperature maser lines at (sub)mm wavelengths with VLBI resolution can be used as diagnostics of the physical conditions in star-forming regions, the atmospheres of evolved stars, and the circumnuclear environment of other galaxies. Furthermore, observations of maser sources over multiple epochs have the potential to provide higher astrometric precision than has been possible to date, enabling valuable new measurements of mass and distance of sources.

Using phased ALMA as a "single-dish" rather than as part of a global VLBI network is also expected to enable new discoveries. For example, observations of pulsars at ALMA frequencies are poised to aid our understanding of pulsar emission processes (e.g., Torne et al. 2015; Mignani et al. 2017), and pulsar searches near the Galactic Center may uncover one or more systems that can be used for relativistic tests at very high precision (Cordes et al. 2004, 2017; Eatough et al. 2015; Psaltis et al. 2016). In addition, sensitive observations of absorption lines at cosmological distances are expected to provide information on the physical conditions of the early universe and chemical evolution over time and can also be used to constrain possible variability of fundamental physical constants.

The initial implementation of the ALMA Phasing System (APS) focused on the simplest use case: VLBI on continuum sources that are bright enough to permit phasing up the array

on the science target (see M18). The APS "Phase 1" has now been fully commissioned and offered to the community for VLBI science observations in ALMA Bands 3 (3 mm) and 6 (1.3 mm), where the GMVA and the EHT, respectively, are available to serve as partner networks. The first science observations that included the APS in this capacity were conducted in 2017 April as part of ALMA Cycle 4.

The present ALMA NA Cycle 3 Study ("Extensions and Enhancements to the ALMA Phasing System") was launched to begin laying the necessary groundwork for an "APP Phase 2" that would bring additional capabilities to the APS beyond those that could be completed within the scope of the initial APP funding awards [a Major Research Instrumentation (MRI) Award from the National Science Foundation (NSF), which included contributions from international cost-sharing partners, and an ALMA NA Cycle 1 Augmentation Award]. These new capabilities are designed to broaden and diversify the scientific applications of the APS, allowing it to reach its full scientific potential as described by Fish et al. (2013) and Tilanus et al. (2014). The period of performance for the Cycle 3 Study was nominally 2016 January 1 to 2017 September 30.

At the time of this report, an ALMA NA Cycle 4 Study program titled "Diversifying the Scientific Applications of the ALMA Phasing System"—a follow-on to the current Cycle 3 award—is ongoing and slated to extend through 2018 January 31. Additionally, an ALMA NA Cycle 5 Development award has been approved to enable implementation of the new APS capabilities investigated in the Cycle 3 and 4 Studies. The latter program (hereafter "APP-2") will enable additional ALMA VLBI capabilities to be offered in ALMA Cycle 7 (or later).

In the sections that follow, we summarize the work and results related to each of the four topics proposed under our Cycle 3 Study: (1) extension of phasing capabilities to the sub-mm (Band 7; Section 2); (2) development of correlation techniques to compensate for the mismatch in sampling rates between ALMA and other VLBI stations (Section 3); (3) optimization of ALMA baseband delay application to minimize decorrelation losses (Section 4); (4) exploration of data reduction and analysis pathways for experiments utilizing phased ALMA data (Section 5).

2. Investigations of Phasing in Band 7

2.1. Background

An extension of VLBI observations to the sub-mm (i.e., Band 7, or $\lambda \sim 0.8$ mm) was discussed in the original MRI proposal for the APP. There is significant interest in this capa-

bibility because of the superior angular resolution compared with Band 6, and in particular because of the diminished impact of interstellar scattering at these shorter wavelengths. This in turn is expected to significantly enhance the ability to reconstruct images of sources (Johnson & Gwinn 2015; Johnson 2016).

The lower-level APS software is agnostic with respect to observing band: the machinery solves for phases without any explicit knowledge of the frequency. However Band 7 has not yet been offered by ALMA for VLBI use for a combination of reasons. First, at the time APP Phase 1 was being commissioned, few of the EHT observatories had 345 GHz receivers. Secondly, the requirement to commission an additional band would have delayed the ability to offer ALMA VLBI capabilities in time for ALMA Cycle 4. Thus Band 7 was deferred, and its practical applications first began to be explored under the current Cycle 3 Study.

Additionally, there are practical differences in operating the phasing system in Band 7 because of the shorter coherence timescales compared with Bands 3 and 6. Precipitable water vapor (PWV) has a more significant impact on phase stability at higher frequencies, making Band 7 observations more susceptible to coherence loss during sub-optimal weather and/or on longer ALMA array baselines. One option for mitigating this is to optimize the array size for phase-up in Band 7; excluding outlying stations decreases the effective collecting area, but may also improve overall phasing efficiency. A second potentially useful capability for Band 7 is the so-called "fast" phasing loop. To help to mitigate the corrupting effects of variable tropospheric water vapor on phase stability during standard (connected element) interferometry, ALMA's 12-m antennas are equipped with water vapor radiometers (WVRs) to provide data that can be used to correct for phase fluctuations, either in real time or during post-processing. A mode of phasing adjustments referred to as the "fast loop" was designed and tested by the APP for the purpose of applying these corrections in real time (see M18). The general idea is that the WVR delays can be taken out on a rapid timescale which then allows the slower timescale phase solution loop ("slow loop") to keep the antennas phased. In very good weather the fast loop corrections were found to add noise to Band 6 data (M18), but these corrections are expected to become increasingly useful at higher frequencies.

2.2. Results and Findings

When conditions warranted, and as authorized by the designated ALMA Site Lead, our team utilized time intervals when the ALMA array was not scheduled for approved science projects to attempt Band 7 phasing experiments. A summary of these tests is provided in Table 1. Although in several instances these data were taken in conditions that were sub-

optimal for Band 7 observing, these data sets have allowed us to demonstrate the feasibility of phasing at sub-mm wavelengths and to begin to explore the impact of weather conditions on Band 7 phasing, as well as the parameter space for optimal phasing operations at higher frequencies. This work remains ongoing as additional data are obtained. Below we describe some results to date.

Table 1: **Band 7 Test Schedule Blocks**

Date	Archive UID	Duration(s)†	Comments
2015 Mar 30	uid://A002/X9cdda2/X42c	225.759	PWV∼0.6 mm
2015 Aug 1	uid://A002/Xa72fea/X1191	1958.222	VLBI w/APEX
2015 Aug 1	uid://A002/Xa72fea/X12a6	256.790	VLBI w/APEX
2015 Aug 1	uid://A002/Xa72fea/X1144	....	...
2015 Aug 1	uid://A002/Xa72fea/X136e	2365.310	VLBI w/APEX
2015 Aug 1	uid://A002/Xa72fea/X132c	....	...
2015 Aug 2	uid://A002/Xa73e10/X28dc	486.572	
2015 Aug 2	uid://A002/Xa73e10/X28cf	216.235	
2015 Aug 2	uid://A002/Xa73e10/X2926	486.644	PWV∼0.6 mm
2016 Jul 9	uid://A002/Xb53e10/X883	4000.286	No delay correction
2016 Jul 9	uid://A002/Xb53e10/Xa7a	3037.934	No delay correction
2016 Jul 9	uid://A002/Xb53e10/X983	3032.294	No delay correction
2017 Jan 29	uid://A002/Xbd27fa/Xd41	2476.982	No delay correction
2017 Jan 30	uid://A002/Xbd3836/X87c	869.892	
2017 Jan 30	uid://A002/Xbd3836/X4ba	485.004	PWV∼7 mm
2017 Jan 30	uid://A002/Xbd3836/X739	978.948	
2017 Jan 30	uid://A002/Xbd3836/X579	901.932	
2017 Feb 1	uid://A002/Xbd3836/X4363	3283.644	PWV∼1.6 mm
2017 Apr 8	uid://A002/Xbec3cb/X6734	....	...
2017 Apr 13	uid://A002/Xbedd69/Xc61	....	...

†Columns with ellipses designate corrupted Archive data and/or data sets with a missing `CalAppPhase.xml` file.

Because Band 7 is more susceptible to the time-variable effects of PWV, as noted above, judicious use of the fast phasing loop is expected to become increasingly important in this band. However, during the study period our experimentation with the fast phasing loop was restricted to Band 6. We note, however, that to date, our experience with weather at ALMA suggests that it is not possible to predict conditions sufficiently in advance to permit reprogramming the Schedule Block. Thus any routine use of the fast mode should

be dynamic in the sense that either the `PhasingController` or the TelCal phasing engine can make real-time choices. Alternatively, an Astronomer-on-Duty (AoD) could be allowed to make this choice through a script or GUI.

2.2.1. ALMA-Only Band 7 Testing

Our first attempt to operate the APS in Band 7 was on 2015 March 30 during an APP Commissioning and Science Verification (CSV) session. This short test was performed as part of the check-out of the current implementation of the "delay fix" (see Section 4), to insure that phasing was working nominally across multiple frequency bands (see Matthews & Crew 2015b for additional details). Testing in Band 7 had initially been included in the plan of work of the original APP NSF MRI award before it was descoped for the reasons described above.

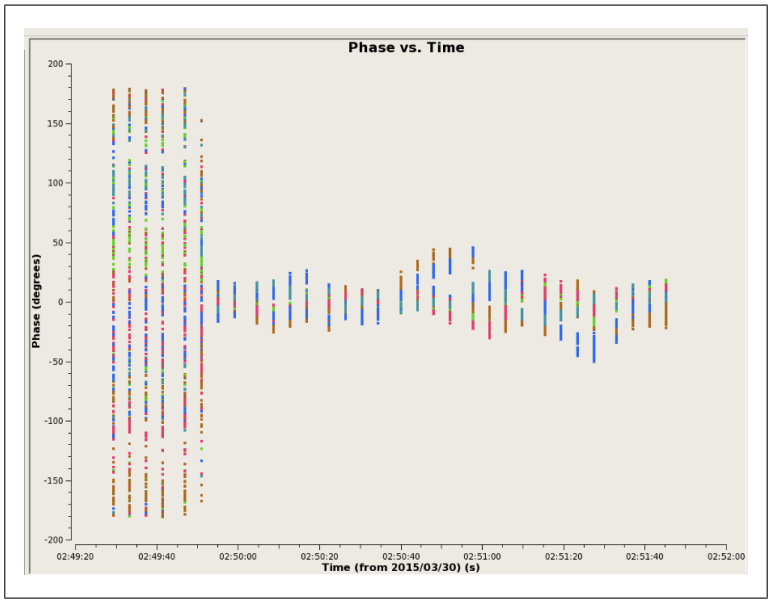

Fig. 1.— Results from the first test of the APS in Band 7 on 2015 March 30. Phase is plotted as function of time for a scan on the 4 Jy source 3C273. Baselines between the phasing reference antenna and antennas that are part of the phased sum are shown, delineated by different colors. Data from all four correlator quadrants (Stokes YY) are overplotted. Once phase-up is achieved, the RMS phase fluctuations on all phased baselines are $\sim 14°$, implying a phasing efficiency of $\sim$97%.

During the 2015 March test there were 9 phased antennas in the array, with baseline lengths ranging from $\sim$15–180 m and PWV$\sim$0.6 mm. Figure 1 plots phases as function of

time for the single Band 7 scan that was obtained. Baselines between the phasing reference antenna and antennas that are part of the phased sum are shown, delineated by different colors. All four correlator quadrants in Stokes YY are overplotted. Over the full range of baselines between antennas in the phased sum, the RMS phases after phase-up is achieved is $\sim 14°$, implying a phasing efficiency of $\sim$97%[1] Figure 2 shows phase as a function of u-v distance (in kλ) for the same data set. As expected, dispersion in the phases increases with baseline length, but remains $< 18°$ RMS on the longest baselines.

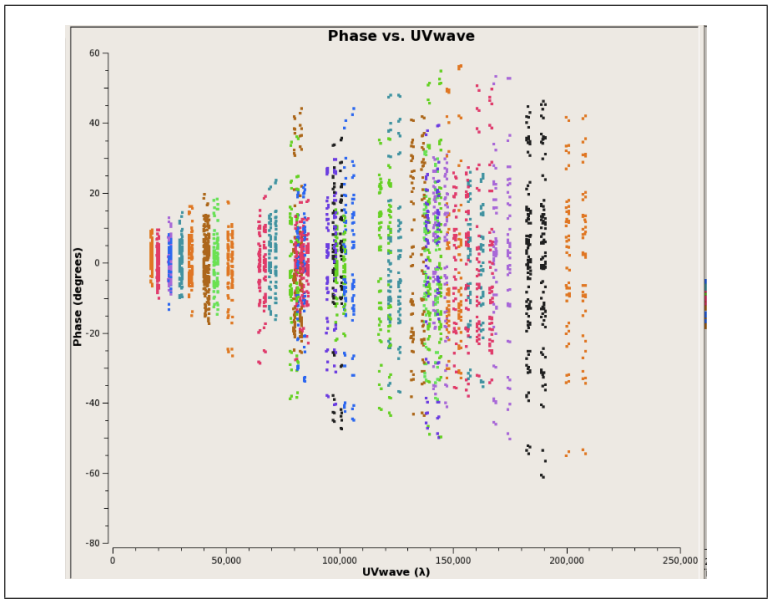

Fig. 2.— Phase as a function of baseline length for the same data set shown in Figure 1.

Another Band 7 data set was obtained on 2015 August 2 with similar PWV conditions to the March 2015 data (PWV$\sim$0.6 mm), but this time with a phased array of 35 antennas and baselines ranging from 15 m to 1470 m. In this case, the RMS dispersion in the phases over all baselines between antennas in the phased sum (after phase-up is achieved) is $\sim 45°$—considerably higher than in the 2015 March data (Figure 3). Limiting the comparison only to baseline lengths that overlap between the two data sets (<200 kλ), phase RMS is $\sim$28°—still twice as high as the March data. While factors in addition to the PWV are expected to affect phase coherence (e.g., wind speed and the overall stability of the atmosphere), this nonetheless provides preliminary evidence that *the presence of baselines longer than a few*

[1]RMS phase fluctuations reduce correlated amplitude by a factor $\epsilon \approx \exp\left(\frac{-\sigma^2_{\rm RMS}}{2}\right)$ where $\sigma_{\rm RMS}$ is the RMS phase scatter in radians (e.g., Carilli et al. 1999). For the present discussion we use ϵ as a measure of phasing efficiency achieved by the APS phasing software, although it does not take into account quantization losses and other inefficiencies present across the entire APS signal path (see M18 for a detailed discussion).

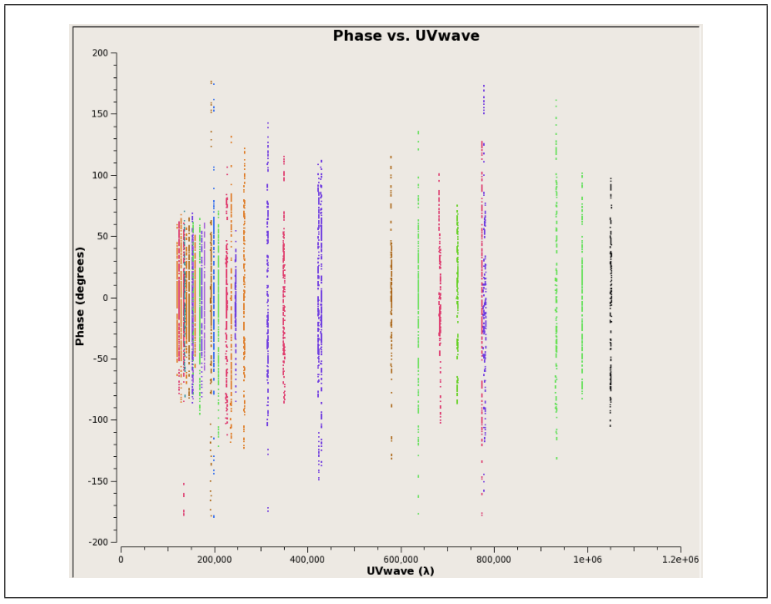

Fig. 3.— Phase as a function of baseline length for a Band 7 data set obtained in 2015 August. The target was the ∼4.5-Jy source B0521−365. Weather conditions were similar to the 2015 March test shown in Figure 2, but longer baselines (up to 1.4 km) were included in the phased array. The mean RMS dispersion in the phases is significantly larger than in the March data, even on the shorter baselines (note the different in the y-axis scale range of the two plots).

hundred meters may detrimental to the overall phasing efficiency in Band 7 even in relatively good weather conditions.

On 2017 February 1 we were able to obtain test data in both Bands 6 and 7 using identical arrays and set-ups. There were 41 antennas in the phased array with baseline lengths ranging from ∼15 m to 326 m. Unfortunately, the weather conditions were rapidly varying, which complicated comparisons between the performance of the two bands. Nonetheless, the results suggest that over this range of baseline lengths, phasing efficiency does not degrade significantly between Bands 6 and 7. As an example, in Figure 4 we compare results from 5-minute scans on the source B0521-365 in both bands. During the Band 6 observation the PWV was ∼2.1 mm and during the Band 7 observation (approximately 40 minutes later) it was ∼1.6 mm. The RMS phase fluctuations for all phased baselines were 49° in Band 6 and 41° in Band 7, implying phasing efficiencies in the two bands of 69% and 77%, respectively.

While these modest efficiencies reflect the sub-optimal weather conditions for observing in these bands, as illustrated in Figure 4, phase RMS is nonetheless comparable between the two bands over the full range of baseline lengths present in the array. While comparison

between additional data Band 6 and 7 sets is still needed, *our initial results suggest that for compact array sizes (baselines $\lesssim$300 m) and moderately good or better observing conditions (PWV $\lesssim$2 mm), phasing efficiency in Band 7 with the current APS will not be significantly lower compared with Band 6.*

In Figure 5 we illustrate phasing efficiency in another way for the 2017 February Band 7 data by plotting the correlated amplitude as a function of time on baselines between the phased sum and each of two unphased reference antenna (blue and turquoise points) and on baselines between these same two comparison antennas and the phasing reference antenna (purple and gold points). For a perfect phased array, the correlated amplitude should be higher on the baselines to the phased sum compared with baselines to a single antenna by a factor equal to the square root of the number of phased antennas. However, in the actual data, the amplitude is reduced by decorrelation losses (in this case a factor of ϵ =0.77, as described above), 2-bit quantization of the sum signal ($\times$0.88), and an additional efficiency loss of $\sim \times 0.8$ whose origin within the APS remains a topic of ongoing investigation (see M18 for discussion). These latter two factors affect only the *phase sum signal*, not the data from the individual baselines, such as those plotted in Figure 4. After accounting for these known losses, the jump in correlated amplitude between standard baselines and baselines to the phased sum is predicted to be $\sqrt{41} \times 0.77 \times 0.88 \times 0.8 \approx 3.5$, in agreement with the observations. This is consistent with the expectation that the phasing system is frequency-independent, aside from efficiency losses resulting from shorter coherence timescales at higher frequencies.

2.2.2. Band 7 VLBI Tests

Our first attempts to coordinate VLBI tests in Band 7 were in 2015 July/August. Initially, intercontinental tests were planned in Band 7 involving ALMA, the Submillimeter Array (SMA) and the Atacama Pathfinder EXperiment (APEX) telescope. However, because of marginal weather conditions in Hawaii, as well as concerns over data quality issues from a Band 6 VLBI test done earlier in the campaign, the Band 7 testing was abandoned, and instead additional VLBI testing was performed in Band 6 involving these sites (see Matthews & Crew 2015c for details).

One day later (2015 August 1), two hours were made available at the APEX telescope for a short baseline VLBI test with ALMA in Band 7 (see Matthews & Crew 2015c). The sessions were scheduled between 19:00-20:00 UT and 21:30-22:30 UT, respectively. The separation between the sessions was inserted to improve parallactic angle coverage of the target sources and thereby enable more robust testing of the `PolConvert` software used by

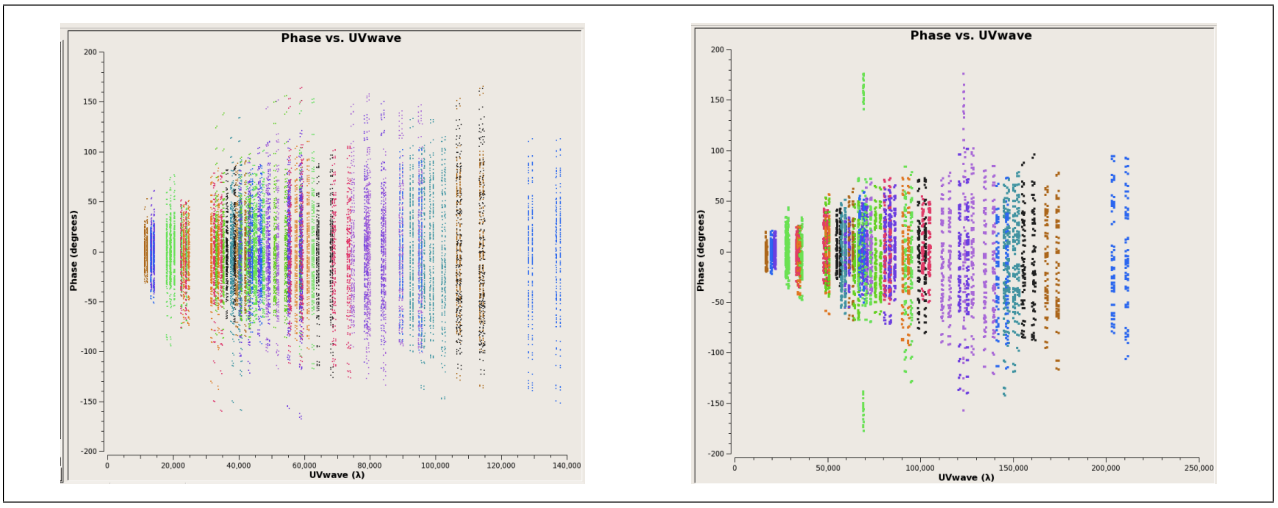

Fig. 4.— Phase as a function of baseline length during scans of a few minutes duration on the source B0521−365 using the APS on 2017 February 1. *Left:* Band 6 scan. *Right:* Band 7 scan. Although the absolute spread in the phase values is slightly higher in the left-hand plot, RMS phase fluctuations are comparable between the two bands (see Text for details).

the APP convert ALMA's linearly polarized data products to circular.[2]) These Band 7 tests were performed in a manual observing mode using a VLBI test observing script that was modified to create a spectral window with Band 7 tunings that APEX could match. Since the conditions were dry and stable, we did not attempt to use the fast loop to apply WVR-based phasing corrections.

Unfortunately, this test was plagued with problems. APEX experienced pointing problems as well as clock offset issues during its first four scans. In addition, the APEX receiver exhibited serious problems with phase stability. On the ALMA side, we experienced issues with the correlator (which were later resolved), but which made the ALMA data difficult to analyze. These included a number of instances of scans that exhibited autocorrelation data comprising purely zeros. While the VLBI recordings from both sides appeared to be robust, we did not attempt a VLBI correlation given the multitude of known problems with the data. *As of this writing, no VLBI fringes have been obtained with ALMA in Band 7 and no further opportunities have been available for our team to attempt VLBI testing with ALMA in Band 7.*

[2]`PolConvert` is a specialized software tool developed by the APP Team to convert ALMA's linearly polarized data products to a circularly polarized basis in order to match what is currently standard at other VLBI sites. See Martí-Vidal et al. (2016b) and M18 for a full description.

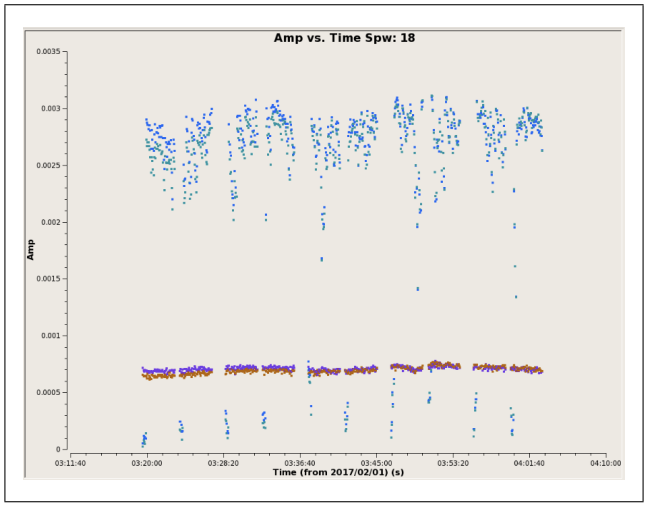

Fig. 5.— Correlated amplitude as a function of time on four baselines from the 2017 February Band 7 observation shown in Figure 4. Stokes XX data for a single baseband are plotted. The blue and turquoise points are baselines between the phased sum signal and each of two unphased comparison antennas. The purple and brown points are baselines between the same two comparison antennas and the phasing reference antenna. The difference in amplitude between the baselines that include the phased sum and those that include only a single antenna from the phased sum (i.e., the phasing reference antenna) provides a measure of the overall phasing efficiency of the entire APS, including atmospheric effects, quantization losses, and other inefficiencies (see Text for discussion).

3. VLBI Correlation Techniques

3.1. Background

VLBI backends at ALMA's peer observatories within the GMVA and EHT networks generally channelize data into subbands whose frequency widths are based on powers of two, i.e., 2^n MHz where n is an integer [but note the exception of the SMA, discussed below]. This usage dates from earlier generations of analog VLBI equipment where such subbanding was necessary in order to cover wider frequency ranges. In contrast, ALMA's correlation configuration partitions each 2-GHz baseband into thirty-two 62.5-MHz channels, each corresponding to one digital tunable filter band (TFB) channel. The innermost 15/16ths of each TFB channel has an approximately flat response with roll-off at the edges. Therefore the TFB channels are slightly overlapped in frequency, yielding a total usable bandwidth of 1.875 GHz per polarization in each of the four basebands. The consequences for VLBI are that (1) the sampling rates of the recorded channels are 125 MHz (which is not a power of 2), and (2) the centers of the (slightly overlapping) channels have a 58.59375 MHz spacing.

Originally, the APP proposed to accommodate the different sampling rates of ALMA and its peer VLBI observatories through the use of multiple ($\geq$16) spectral windows across each ALMA baseband, with each spectral window tuned so as to align the DC edge with those of the 32 MHz VLBI channels at other sites, and with 2.5 MHz gaps between the spectral windows. A so-called "zoom" mode would then be used to handle this type of configuration during correlation. The software correlation package DiFX (Deller et al 2007, 2011) already supported a zoom mode that was designed to carve out spectral regions of interest (typically regions containing spectral lines), and a modification of this zoom mode also offers a solution to the problem of correlating ALMA VLBI data, since the non-overlapping portions of each channel can be assigned to zoom bands and then correlated and analyzed in a straightforward manner (i.e., as if this had been the original channelization).

At the time the APP was proposed, polyphase filter bank (PFB) implementations were the state-of-the-art, and a scheme such as shown in Figure 6 was contemplated to allow ALMA to match a 512 MHz band covered by 32 MHz subbands (channels). Variations on that theme would allow 64 MHz or even 512 MHz channels at ALMA's VLBI peer observatories.

Extensive work by APP team members at the Academia Sinica Institute of Astronomy and Astrophysics (ASIAA) investigated this and similar schemes using software developed to simulate observations, as well as software to automate the configuration of DiFX (see https://github.com/zhengmeyer/autozoom and Pradel 2014). However, one of the drawbacks of this approach is that large fractions of the total bandwidth available cannot be appropriately correlated or used without significant post-correlation software to put the pieces back together into bands usable for analysis.[3]If this software is eventually developed to be flexible, it should allow more options for handling the "mis-matched channel problem" which was identified at the outset of the APP, including the use of narrower channels. We have largely resolved the problem through a combination of zoom bands and the very wide channels. In addition, one is confined to 32 or 16 MHz channels in order to get suitable overlaps. To avoid gaps in the VLBI correlation, one has to give up ALMA bandwidth (see Figure 7).

---

[3]During the recent Cycle 4 VLBI campaign with the GMVA there was a commanding error (PFB mode rather than DDC mode was used in the backends) which led to a significantly mismatched band and resulting correlation problems. As of this writing, an effort is currently under way at the Max-Planck-Institut für Radioastronomie (MPIfR) in Bonn, Germany to develop suitable repair software.

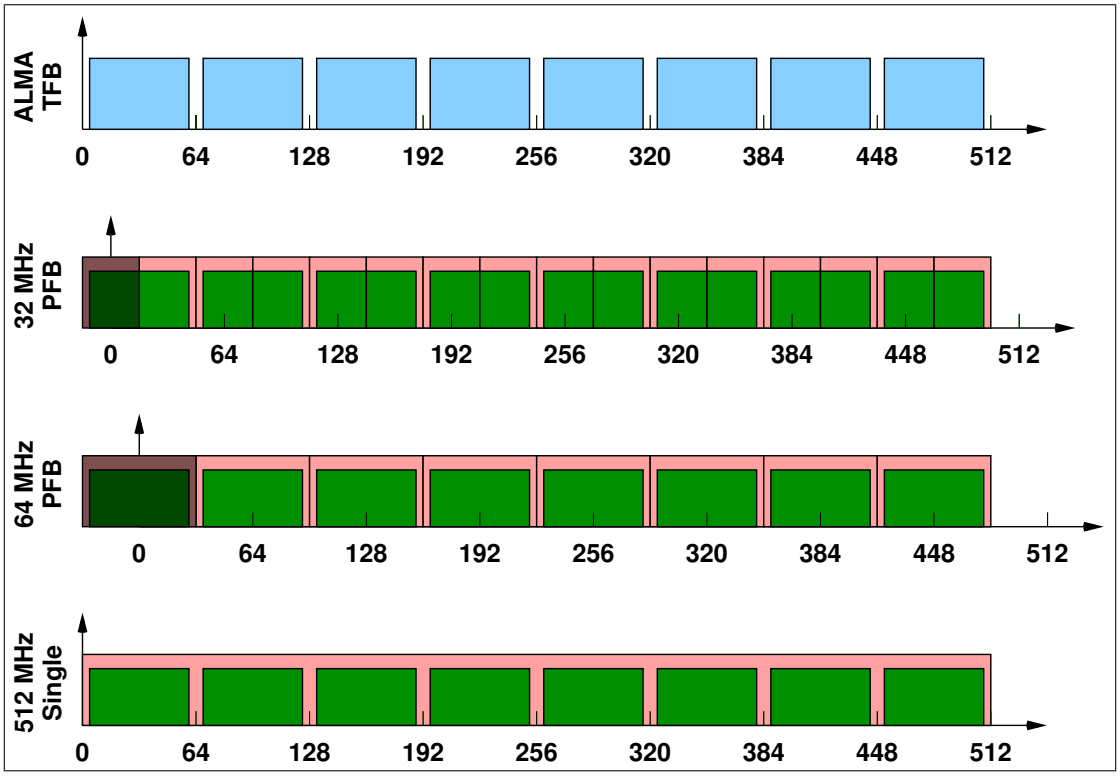

Fig. 6.— The original plan for using zoom bands to match 32 MHz polyphase filterbank (PFB) channels with ALMA 62.5 MHz TFB subbands. This scenario assumed that the ALMA bands could be placed at 64 MHz frequency intervals.

3.2. Results and Findings

Despite this extensive preparatory work, the above correlation scheme ultimately proved to be untenable. It was discovered during APP commissioning that the ALMA baseline (BL) correlator is presently unable to support the use of more than two spectral windows within a single baseband. An additional problem is that the finite frequency resolution of the digital local oscillator (LO) within each TFB limits tuning to 31 kHz resolution. While small LO offsets are commonly used at some VLBI sites, traditionally these have been at most a few Hz. Some simulations done by the ASIAA group established that 31 kHz LO offsets are probably usable, though they have never been tested in actual VLBI experiments. Furthermore, with 16 MHz channels expected at peer VLBI observatories in the future, the above scheme results in a very large number of subbands to process.

In a parallel development, the EHT (for programmatic reasons) did not develop channelized backends for its build-out to support ALMA-sized basebands. Instead, the Reconfigurable Open Architecture Computing Hardware 2 (ROACH2) Digital Backend (R2DBE;

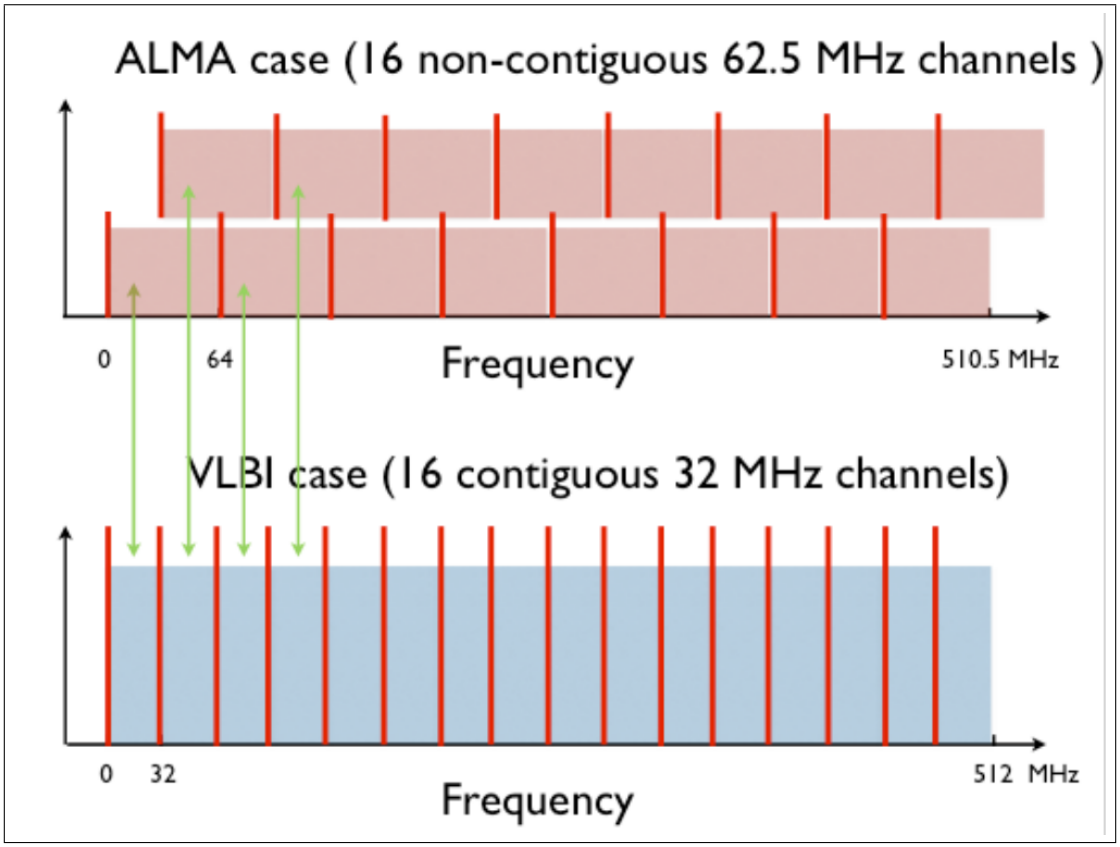

Fig. 7.— Sample of zoom band matching where 16 ALMA TFB channels can be placed to cover a 512 MHz VLBI peer site with 16× 32 MHz channels. In this case all bandwidth is covered, but 75% of the ALMA band is wasted. Adapted from Pradel (2014).

Vertatschitsch et al. 2015) was deployed, which utilizes a single 2048 MHz channel. This makes it very simple to carve out zoom bands for data from EHT sites, as there are no channel boundaries to cross. No additional complications are expected as the EHT moves to wider bandwidths for Cycle 5. For the GMVA, a digital downconvertor or "DDC" mode is available that provides two 128 MHz bands; in this case, four zoom bands can easily match four ALMA TFB bands.

A third related development was related to the Haystack Observatory Postprocessing System (HOPS)[4] software commonly used for mm VLBI data reduction. The Astronomical Image Processing System (AIPS) has for many years been able to work with channels ("IFs") comprising arbitrary frequency widths and arbitrary numbers of spectral channels. To enable

[4]https://www.haystack.mit.edu/tech/vlbi/hops.html

similar capabilities in HOPS, this software was recently enhanced by R. Cappallo (MIT Haystack Observatory) to use the Fastest Fourier Transform in the West (FFTW) library for its Fourier transforms. This library includes support for Discrete Fourier Transforms (DFT) rather than only Fast Fourier Transforms (FFT). Finalization and testing of these new capabilities were done under the current award. This enhancement meant that the zoom channels could have an arbitrary number of spectral channels—not simply a power of two as was previously required. Figure 8 shows one of the first tests done using a non-power-of-two number of lags (540 in this example).

Because of these developments, we have now devised a correlation configuration with 58-MHz zoom channels, 1-sec integrations and 0.5-MHz spectral resolution which is comparable to previously used VLBI correlation configurations at 1 mm and 3 mm. The only residual issue is that there is still some small loss of bandwidth ($\sim$1%) because of samples excluded from the correlation, and some delay issues that result from the number of samples that must go into each DFT while the Earth rotates. In the latter case, the loss is 1-2% and is quantifiable and acceptable.

One other technology change that became relevant to the problem of VLBI correlation for phased ALMA is the Adaptive Phased-array Heterogeneous Interpolating Downsampler for SWARM (APHIDS), where "SWARM" stands for the SMA Wideband Astronomical ROACH2 Machine system. This was developed for handling the SMA's VLBI data for the EHT (Primiani et al. 2016). While ALMA introduced 125 MHz sampling into the zoom band numerology, the comparable sampling rate at the SMA is 143 MHz. The latter zoom band scheme depends on performing DFTs that are sufficiently large that spectral channels can be matched-up between disparate recorded bands. While the scheme is just barely feasible for ALMA ($2^5 \times 5^2$) and the R2DBE (2^{14}), or equally for the SMA and an R2DBE ($2^5 \times 11 \times 13$), it becomes totally unworkable for all three together ($2^{14} \times 5^2 \times 11 \times 13$) is more than 14 million samples per FFT. Even assuming the processors could perform the calculation, the delay losses induced by such a large FFT size would be unsupportable. (The underlying DiFX algorithm assumes that the FFTs are not too large.)

The APHIDS solution was to record raw SWARM data with greater bit depth (effectively 4-bits per sample, but with the samples in the spectral domain) and build a GPU processor capable of ingesting this data stream and synthesizing the 2-bit single channel that would have been generated by, e.g. an R2DBE. The existing implementation, as customized for SWARM, is able to perform the reprocessing at a speed several times slower than the experiment duration. While having to wait for this extra processing step is an inconvenience, it is preferable to complicating the DiFX correlation machinery to allow doing the equivalent processing as part of the DiFX correlation stage.

During the course of this Study, our team has been in regular contact with Walter Brisken (Director Long Baseline Observatory) regarding DiFX needs and requirements for ALMA. In addition to the activities described above, we note some other developments related to DiFX under our current Study:

- Our team uncovered a bug in DIFX2FITS that led to a problem populating the POLCALA and POLCALB arrays in the case where NOPCAL=0 and which caused confusion in certain AIPS tasks. This was fixed in DiFX revision 7349.
- `PolConvert` and its required support scripts have been added to the DiFX trunk.
- The Quality Assurance Level 2 (QA2) calibration script(s) required for processing ALMA data have been added to a subdirectory (QA2) for reference on the DiFX trunk. The current versions are fully functional, although further testing and revision of these scripts remains ongoing as of this writing.
- A subdirectory (PP) was created for `PolConvert` post-processing at the correlator using the QA2 inputs. As the QA2 scripts change, there may be further changes to these scripts.

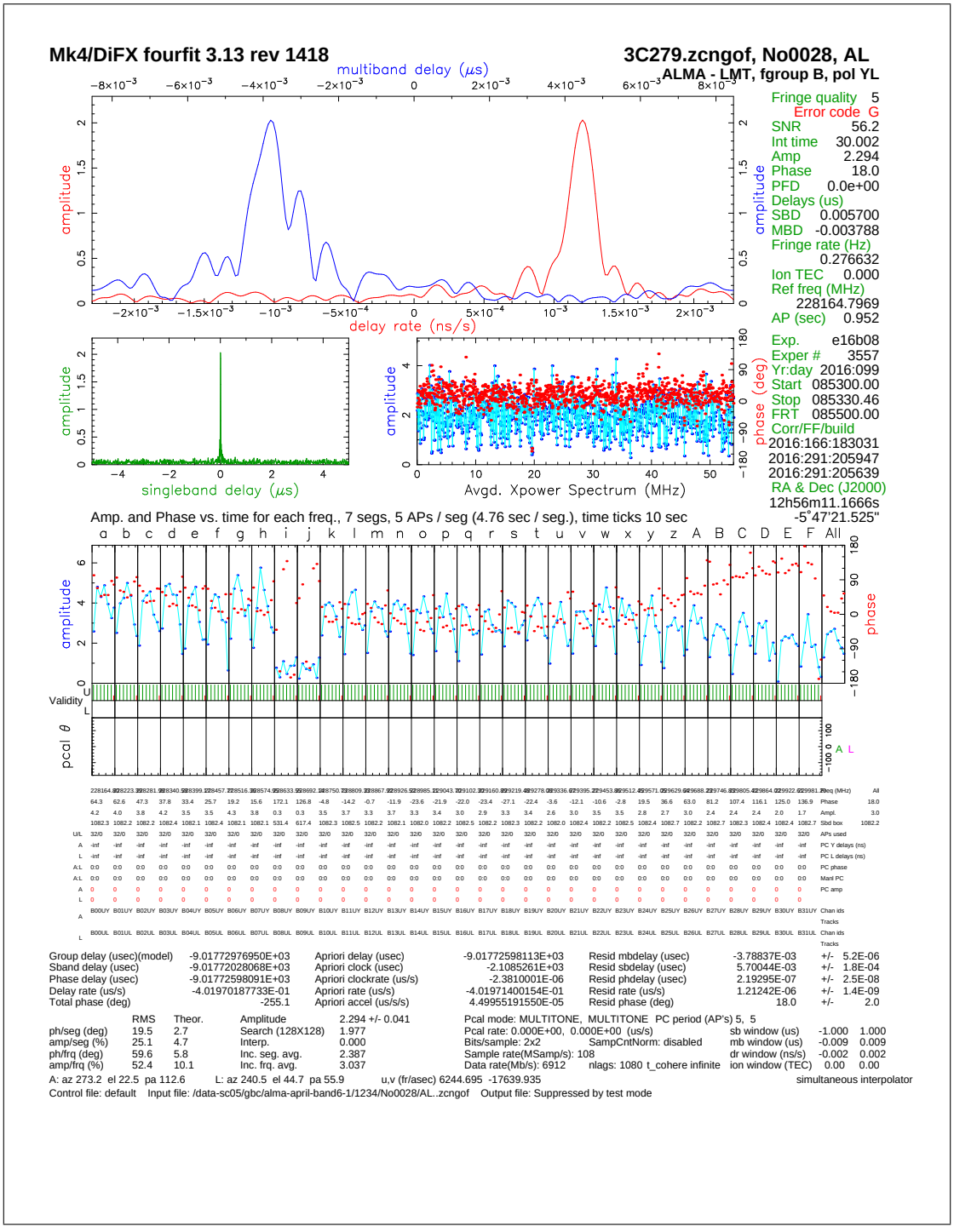

Fig. 8.— Sample fringe plots illustrating an enhancement to HOPS to allow handling arbitrary numbers of frequency channels (i.e., numbers that are not a power of two). This example has 520 lags and a total bandwidth of 1728 MHz. The data are from a Band 6 APP CSV observation of the quasar 3C279 on a baseline between ALMA and the Large Millimeter Telescope (LMT) taken in 2016 April.

4. Optimization of Baseband Delays

4.1. Background

During the initial phases of APP CSV it was recognized that while the phasing system was working properly in baseband 1 (BB_1), polarization X in Band 3, it was not operating effectively in the other polarization (Y) or in the other basebands (BB_2, BB_3, BB_4), and phasing efficiency was further reduced in Band 6 (Matthews & Crew 2015a). After extensive discussions with ALMA-knowledgeable developers and astronomers, we ultimately traced the problem to a poorly documented aspect of the delay system. Ultimately, a change in the manner that the APS handles baseband delays (BBDs) was required in order to achieve acceptable phasing performance in all basesbands, polarizations, and observing bands.

Because of time pressures to reach an operational capacity within the funded duration of the APP project, and in time to offer ALMA VLBI to the community in Cycle 4, the APP team devised and implemented a work-around for the underlying issue (hereafter the "delay problem") that involved turning off the BBD correction applied in the Correlator Data Processor (CDP) computers. We provide a more detailed explanation of what this entails below, but for context, we first briefly review how the ALMA delay system works.

As with any connected element interferometric array, the signals from the ALMA antennas need compensation for the geometric delays prior to correlation. In the ALMA system, there is a software component (`DelayServer`) that understands the geometry of the array and periodically publishes a set of delay adjustments on the real-time Correlator Area Network (CAN) bus. These adjustments are assigned specific times of application (at a resolution of 1 ms) and are in the range 0 - 100 μs for an array of ALMA's size.(There is a so-called "causality" offset added to the delays to ensure that these are all positive, which simplifies the low-level implementation details.) At ALMA the delays are decomposed into several pieces which can be optimally applied in various hardware components within the antenna plus correlator system. This partitioning is somewhat arbitrary; but ultimately traces to the relative convenience of making the corrections in the various hardware components. The terminology varies somewhat within the ALMA documentation, but the slide shown in Figure 9 (from R. Amestica, 2013) documents the delay system as viewed from the correlator software perspective, as the correlator software is ultimately responsible for passing the correlated visibility data into the ALMA Archive without any sizable delay on any baselines.

One portion of the delay (hereafter the "fine" delay) is applied in the digitizer at the antenna, and is implemented by rotating the phase of the 4000 MHz digitizer clock (DGCK). This is a digital implementation which allows 16 rotation steps per sample for a resolution

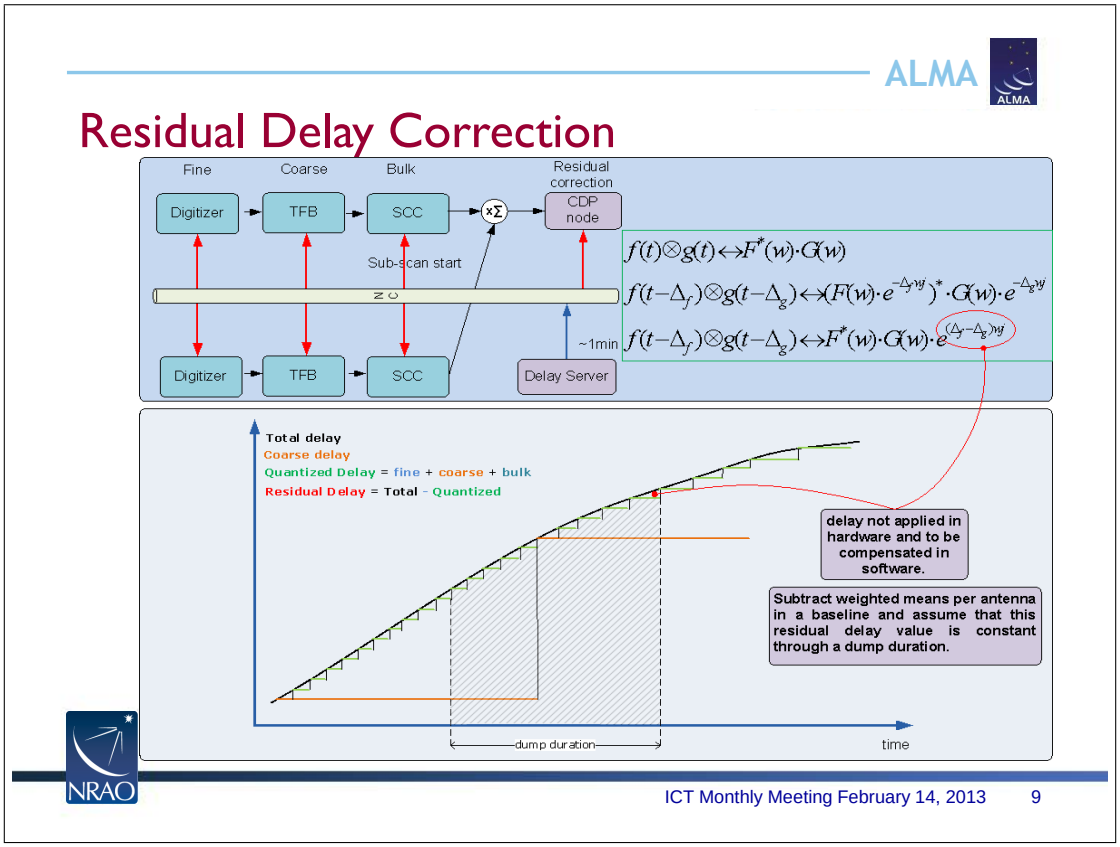

Fig. 9.— This graphic (from R. Amestica) shows how the delay system is decomposed for use within the ALMA BL correlator. See text for full discussion.

of 250 ps / 16 = 15.625 ps and is the smallest delay adjustment possible in the system.

Every complete turn (16 steps) corresponds to a sample shift, so the remaining adjustments are at the sample level and are managed by queuing the samples in the station cards and releasing them for correlation by adjusting the read position in the queue. These adjustments are subdivided into "bulk" delays (applied in the station cards) and "coarse" delays (applied in the TFBs within the station cards), amounting to some integral number of (250 ps) samples, again on a per-antenna/polarization data stream basis. Once the correlator is programmed, these data streams are passed to the correlator cards for cross-correlation and accumulation ($\times\Sigma$ in the diagram), and then dumped to the CDP nodes for final processing. This CDP spectral processing includes a number of corrections to the data, but most importantly, it applies a final correction piece, the so-called residual delay correction (RDC), which is 0–15.625 ps and is the remaining part of the correction which cannot be applied in hardware. (Note that this correction is imperfect since the adjustments are made on a 1 ms cadence; however the error is, by design, tolerable.) The algebra in Figure 9

shows that CDP RDC correction is simply a complex multiplication by $\exp[(\Delta_f - \Delta_g)\omega j]$ for residual delays of Δ_f and Δ_g at frequency ω on two streams labelled f and g (and where $j = \sqrt{-1}$).

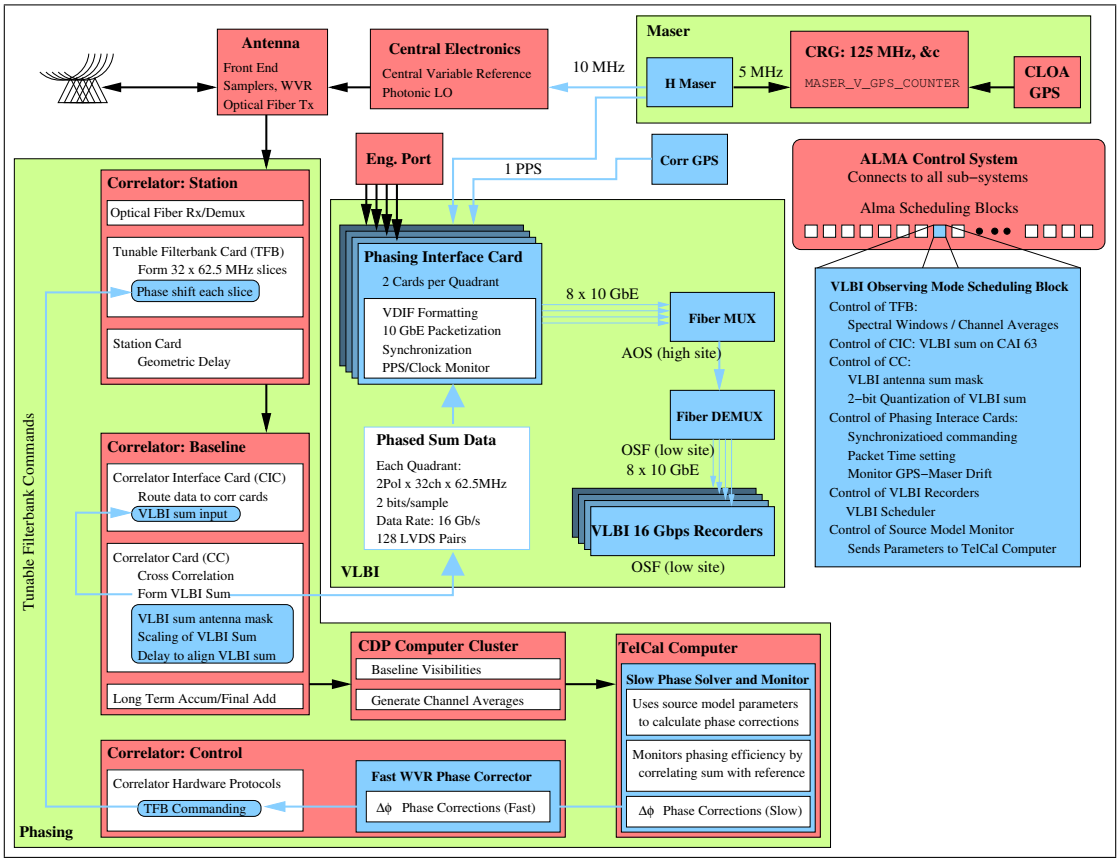

Fig. 10.— Diagram of the APS, illustrating how existing ALMA systems were modified to support VLBI. Pre-existing objects are shown in pink and objects in blue are new components added by the APP. The green backgrounds group related components.

This was all understood in the design of the APS, which is shown schematically in Figure 10. In this figure, the L-shaped greenish domain groups the correlator components and TelCal software component where the phasing adjustments are calculated. The signals from the antennas flow down through the correlator station cards, are correlated in the correlator baseline cards, and pass (horizontally at the bottom of the Figure) through the CDP Computer Cluster where channel averages are formed and published to TelCal (and also to the Archive, not shown). The APP TelCal phasing engine computes the phases on all of the channel averages and passes them back to the Correlator Control Computer (CCC) which issues the commands to actually make the corrections. Specifically, it inserts the negative of the measured phase into a special phase offset register (see Figure 11) in each TFB (in the

correlator station cards) which when applied makes for an additive phase correction. After the correction is made, the signals TelCal sees on the next correlator subscan dump should reflect the phase change and be far less in need of correction. (If the atmosphere is perfectly stable, one correction is all that is required; however, under normal conditions, adjustments every subscan are required.)

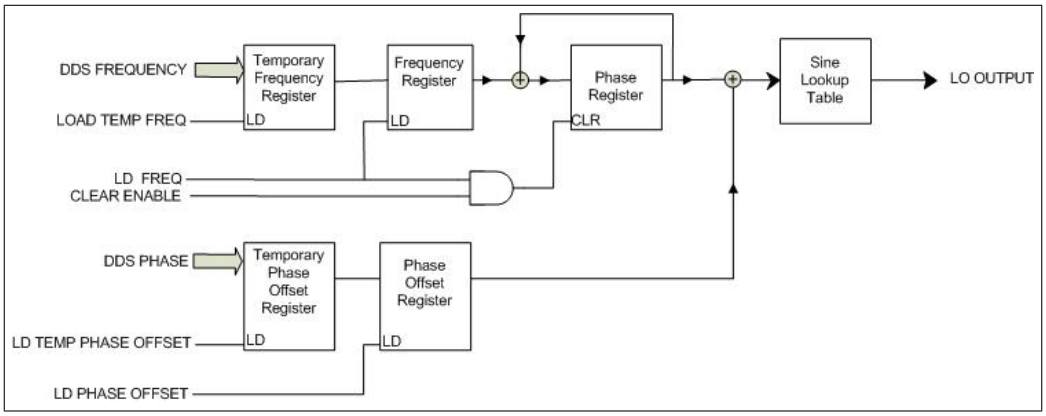

Fig. 11.— In normal operation of each TFB, the logic in the upper half of this figure is active: the desired frequency is loaded and the contents of the phase register are cleared at the start of the observation. Thereafter the LO output is driven from a sine look-up table. When the APS is active, an additional 16-bit phase offset register is loaded with a value for the appropriate fraction of a turn of phase at every correlator subscan in order to correct the phase of this TFB. This phase offset may be cleared (or not) at the start of a scan. The capability of leaving the phase offset register contents unchanged will enable future use of a "passive" phasing mode.

The above discussion is missing one crucial component, namely the stable circuit delays for the individual analog circuit paths through the receivers, switches, and filters on their way to the digitizers. Although similar, these paths have relative delays on the order of a few hundred ps. The collective sum of these various delays is termed the "baseband delay" (BBD) and it can be measured through calibration exercises performed periodically at ALMA and stored in a database. If the various digitizers within each antenna had separate DGCKs, it would be possible to read these delays from the database and adjust the geometric delays to exactly compensate for them in the `DelayServer` commands. However, separate LOs were not provided, so the "fine" delay portion (0–250 ps) of the BBD can only be fixed for one baseband polarization data stream, or alternatively minimized for all of them. On the other hand, since ALMA does not require real-time correctness for its normal observations, it is sufficient to make the correction within the CDP spectral processors. Indeed, the RDC correction module is already coded for the necessary algebra. Thus Δ_f and Δ_g are

merely adjusted by the BBD (i.e., a delay correction of RDC + BBD is made, but all the documentation and code references are for a "residual delay correction". This is why it was missed in our initial review of the correlator design and code.)

The root of the "delay problem" for the APP can be seen in Figure 10: the CDP processing lies between the TFB phase registers and TelCal. Hence the BBD correction provides TelCal with a considerably different signal from the one it can adjust in the TFB registers. The solution currently adopted by the APP is to turn off the RDC correction. Or more precisely, since on longer baselines the RDC correction is useful to stabilize the relative delay variations (as the the DGCK is stepped), we added commanding to separately control the application of the BBD correction. Thus operating with BBD correction turned off allows the phasing loop to function normally, but at a cost of signatures of the uncorrected BBD being present in the ALMA interferometry data that are archived in parallel with the VLBI recordings whenever the APS is operational.

Owing to software modifications introduced after the delay problem was discovered, the delay is partially compensated by a frequency-varying phase correction. As part of this fix, the allowed number of channel averages across each baseband was increased to 32, and if each of the 32 TFBs is programmed to correspond to a channel average, the delay can be removed via a phase ramp across the 32 TFBs. In practice, there is a trade-off between the number of channel averages used and the SNR of the phasing solutions computed by TelCal for each one. As a compromise, recent operation of the APS, including the Cycle 4 science campaigns, has employed 8 channel averages per baseband.

An example of the resulting bandpasses for a case where 8 channel averages were used per baseband is shown in Figure 12. The plot shows visibility phase as a function of frequency for a baseline between an unphased comparison antenna and the phasing reference antenna. In this plot, different colors represent a time series of subscans. At the start of the sequence, before the array is phased up (red and orange points), the visibility phase on the plotted baseline shows a significant phase ramp across the entire band. At later times, after phase-up is achieved, this phase slope (delay) is removed, but a "sawtooth" pattern is present. In this experiment, 8 channel averages were used, with 4 TFBs per channel average, so the "saw" has 8 teeth. Thus we see that a consequence of the current solution to the delay problem implemented by the APP team is that antennas excluded from the phased array (comparison antennas) will have a large residual delay. Since these comparison antennas are used for characterization of the phasing efficiency, this statistic becomes compromised. Another drawback of this approach is the need to break the baseband into several pieces reduces the available bandwidth for each phasing calculation. This reduces the SNR for the phasing calculations, leading to an increase in flux density threshold for sources on which

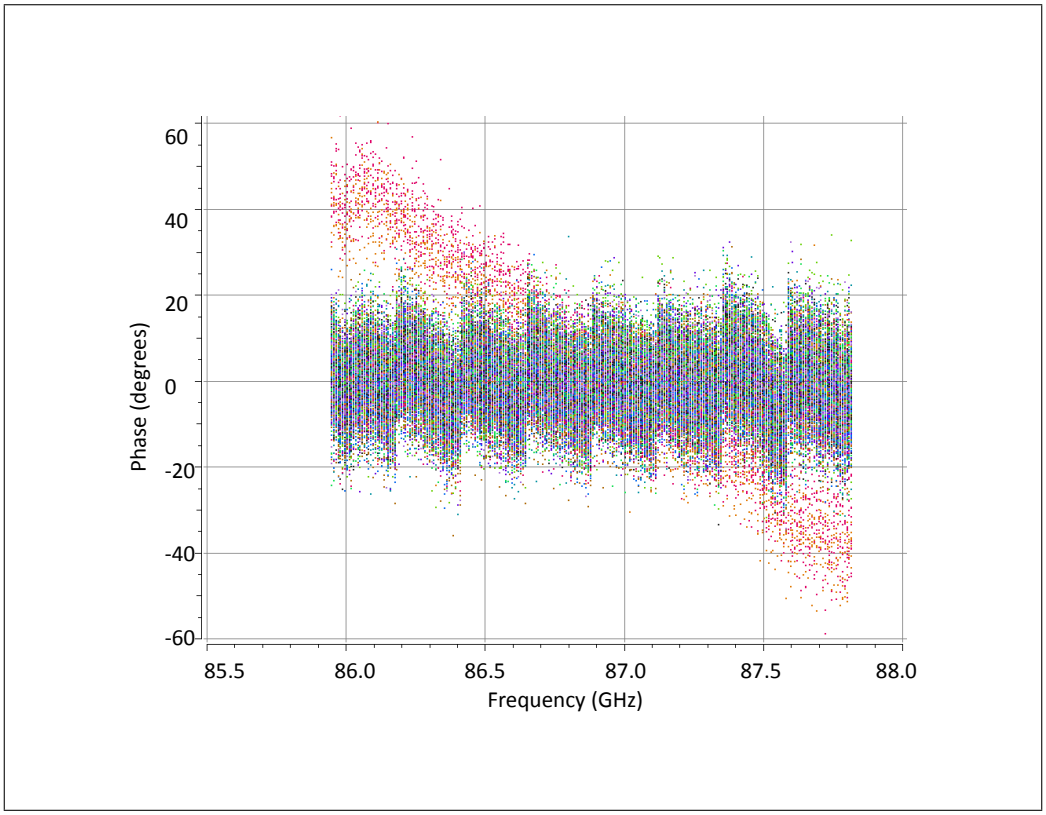

Fig. 12.— Example of the impact of the current "delay fix" on data taken with the APS. Visibility phase is plotted as a function of frequency on a baseline between the phasing reference antenna and an antenna that was part of the phased sum. Different colors correspond to a time series of subscans. Early in the sequence, before the array is phased up (red and orange points), a phase ramp is present across the entire 1.8 GHz band, reflecting the fact that the data are uncorrected for the BBD. As phase-up is achieved in subsequent subscans, the points become clustered near zero phase. The phasing corrections remove most of the delay error (i.e., the mean phase slope across the full band is now zero), but a residual "sawtooth" pattern remains as an artifact of the application of phasing adjustments on a per-channel-average basis. In this example, there are 8 spectral averages across the band, with 4 TFBs per spectral average; the phasing adjustments are unable to remove these delay-like phase slope within each subset of TFBs.

phase-up can be achieved. Additionally, the phased antennas (and the phased sum) suffer modest decorrelation losses due to the phase ramp ("tooth") across each chunk of bandwidth used to compute the 8 channel averages from which the phasing solutions are determined. Finally, the resulting bandpass effects require special handling during analysis of the ALMA single field interferometry (SFI) data archived with each ALMA VLBI experiment.

4.2. Results and Findings

The above described imperfections in the current delay fix were deemed acceptable for the initial offering of the APS to the community for VLBI, but ALMA Management requested that the APP team investigate alternative solutions to the delay problem that would mitigate or minimize these issues. That problem was considered as part of this Cycle 3 Study, leading to the following four possible options to fixing the delay problem. Below we summarize and the merits and drawbacks of each approach and summarize our recommended path for implementation of a new, improved delay fix under our pending Cycle 5 Development award.

Option 0: As is clear from the discussion above, moving the BBDs into the `DelayServer` would not fully solve the delay problem, so that is not presently a viable solution. However, over the 30-year life of ALMA (especially if front end hardware changes are made), such an approach may become tenable in the future. We therefore make note of this possibility for completeness and future reference.

Option 1: Another option would be to apply the BBD to the spectral data, but leave the channel average data stream uncorrected. This would solve the cosmetic problems (the channel averages are not otherwise used in continuum observing modes) and the spectral data would look "normal". However, the phasing solutions would still have to be computed on a per-channel average basis. That results in a higher flux density limit for active phasing, since the effective bandwidth available for the phasing solution calculation is reduced, i.e., the source must be bright enough to get usable phase solutions in every channel average. When the band is divided up into 8 channel averages, that limit conservatively appears to be $\sim$500 mJy.

Option 2: A better solution (and the one that we have ultimately proposed to lay the groundwork for during our subsequent Cycle 4 Study) is to restore the CDP processing to its normal state (meaning that the Archived data sets would look "normal") and instead fix the delay problem through modifications to TelCal software. The key piece of software development is to provide the TelCal phasing engine with the same look-up methods into the ALMA Telescope Monitor and Configuration DataBase (TMCDB) as are used by the CDP. Since knowledge of the channel frequencies is needed to use these delays, TelCal either needs to retrieve them from the archival science data model (ASDM) file associated with the data set or get them (directly or indirectly) from the correlator software elements. With these enhancements, TelCal could then perform or undo any calculation that the CDP does. Because TelCal works with time-averages of the channel averages, performance requirements are modest and additional processing overheads incurred will not be an issue.

For testing purposes (noting that application of BBD corrections should commute with

phase solutions), one might envision implementing two possible algorithms:

Algorithm A:

A0 CDP does the BBD correction

A1 Undo the BBD correction made by the CDP

A2 Solve for phases

A3 Report phases

Algorithm B:

B0 CDP does the BBD correction

B1 Solve for phases

B2 Undo the BBD correction made by the CDP

B3 Report phases

The Algorithm A pathway is equivalent to what is currently performed (since the work of step A1 reverts the work of step A0, which is equivalent to omitting both steps); steps A2 and A3 are currently coded as part of the phasing algorithm. Algorithm B uses the same code, but inverts the steps B1 and B2 and is closer to what we want, and should also produce the same results as Algorithm A. Finally, step B1 can *average* across the channel averages to increase the SNR of the phasing solutions. This averaging code is present in TelCal, but has not been used since the BBD problem was identified.

We plan to implement both algorithms for comparison purposes. The point of recognizing and implementing these algorithmic steps is to enable some offline testing of the code, which in turn makes it more likely that on-sky testing will have fewer problems. [It turns out to be rather difficult to properly simulate the full system and the Array Operations Site (AOS) is the only standard test environment (STE) that can properly test the phasing system.]

With this general design, the Integrated Computing Team (ICT) task work may be described as follows:

1. TelCal Parameter Tuning Interface

   (a) Provide TelCal with a (per baseband) datum to indicate whether BBDs are being applied or not, and if they are applied, how TelCal should process and encode the phase data.

   (b) Supply TelCal with a datum to specify the channel average frequencies in use (per baseband). These would be the frequency values used by the pack/unpack library to ensure that the phases and delays are consistently handled by both the Control systems and TelCal.

   (c) Supply TelCal with a method to obtain the channel average frequencies directly from the ASDM data.

2. TelCal Phasing Engine

   (a) Develop a method to retrieve BBD data from TMCDB using code similar to what the CDP currently employs.

   (b) Develop a method to use the pack library (or not), as directed.

3. Correlator Control Computer (CCC)

   (a) Revisions to the pack/unpack library (where the code is currently hosted and tested).

   (b) Develop a method to obtain the channel average frequencies so these can be provided to TelCal.

4. SSR Observing Script

   (a) Implement support for the various delay correction pathways in the observing script (for testing and eventual observation).

   (b) Implement `VEX2VOM` support for the testing.

The pack/unpack library contains methods for moving phases between TelCal and the `PhasingController`, which are currently not uniformly used, but might be. Not all of these methods need to be implemented or used, but the underlying issue is that it is difficult to discover errors without exercising the full system on sky. The full set of methods should then allow the phasing system to be tested in a few short on-sky tests, ideally separated by sufficient time for full analysis of each test.

Option 3: Yet another option would be to turn on the BBDs, use a single channel average per baseband, and then have some agent other than TelCal calculate the phase ramp to compensate for the BBD. This agent could be the `PhasingController` or the CCC process.

To summarize, Option 0 above would theoretically be the best one, as it shifts all the work to the `DelayServer`, but it is not currently feasible (though it may be in the future).

Option 1 would simplify some analysis tasks but otherwise keep the *status quo*, including the current flux density threshold for use of the phasing system. Option 3 would also work, but would be more invasive and more difficult to test; nonetheless, it is viable. Our preference and plan is to therefore to implement Option 2 above, as it is the least intrusive and easiest to test both offline and on-sky and it allows relaxing current flux density limits for the phasing system.

5. Data Reduction Explorations

5.1. Background

For VLBI experiments at centimeter (cm) wavelengths, AIPS and other software tools such as Difmap (Shepherd 1997) already offer a variety of well-tested tools for handling post-correlation data calibration, editing, analysis, and imaging. Furthermore, strategies for processing cm wavelength VLBI data sets using these software packages are well-documented (e.g., the VLBI chapters of the AIPS Cookbook[5]). However, data processing strategies appropriate for cm-wave VLBI experiments do not always directly translate to mm VLBI because of additional challenges inherent to observations at these shorter wavelengths. For example, at mm wavelengths, tropospheric fluctuations typically limit atmospheric coherence times to $\sim$10 s, necessitating the use of specialized fringe detection algorithms (Rogers et al. 1995). These algorithms are not presently available in AIPS, but have been implemented the HOPS software and plans are underway by other groups to incorporate them into the Common Astronomy Software Applications (CASA) package.

Including phased ALMA as a station in a VLBI array also introduces other unique challenges. For example, to allow robust amplitude calibration, methods need to be devised to apply the results of regular $T_{\rm sys}$ measurements to phased ALMA data in an optimal manner. Yet another challenge is that unlike other mm VLBI stations throughout the world which record circularly polarized signals (RCP and/or LCP), ALMA records dual linear polarizations (X and Y), requiring the application of specialized software (`PolConvert`; see Footnote 2) to perform a translation to a circularly polarized basis. In particular, the impact of these corrections on sources that have high degrees of intrinsic linear polarization need to be carefully examined to insure that the resulting data products are robust and that the true polarization characteristics of the source can be accurately derived.

In light of these various challenges, we proposed as part of this Study to develop a

[5]http:www.aips.nrao.edu/cook.html

"Cookbook" outlining strategies for post-correlation data processing of (sub)mm VLBI observations involving a phased ALMA. Our work exploring optimal end-to-end data handling procedures for VLBI data that include phased ALMA remains ongoing as of this writing. However, we summarize below progress to date.

5.2. Results and Findings

During the late stages of the APP CSV in 2016, our team carried out two global VLBI sessions of ∼4 hours duration each, one in Band 3, one in Band 6. The resulting data sets were sufficient to validate Cycle 4 software for its eventual use and to demonstrate the feasibility of routine end-to-end VLBI operations at ALMA (including Scheduling Block creation, execution of observations, correlation, verification of VLBI fringes, and `PolConvert` (see Footnote 2). Various outcomes from the analysis of these data sets have been described in Matthews, Crew, & Fish (2017) and Martí-Vidal et al. (2016a). An ALMA Technical Note describing the end-to-end analysis of the Band 3 VLBI data set is currently in preparation. The latter will include a description of the methodology we have developed for fringe-fitting and calibration of Band 3 VLBI data that include phased ALMA, including the approach we have developed for handling phased ALMA $T_{\rm sys}$ measurement within AIPS. With assistance from Eric Greisen (NRAO), work on these data also led to two important updates in AIPS to accommodate ALMA VLBI data. The first was an increase in the buffer size in the TABAN task to accommodate data with a significant number of IFs. The second involved the AT2AN routine that reads calibration from the AT table and is used by tasks such as VBGLU. It was modified to loop over all IFs in the AT table and process the full set of IFs and polarization terms.

Post-correlation analysis of the aforementioned VLBI data sets continues to be a work in progress owing to a combination of factors. First, crucial initial steps in the delivery of ALMA VLBI data products (correlation and QA2, including the application of `PolConvert`) continued to be refined as experience was gained, leading into Cycle 4. In addition, improved techniques and algorithms for analyzing data continue to arise from time spent with significant amounts of science-quality data. Unfortunately, a significant fraction of the data taken in the course of the APP CSV were by default during marginal weather conditions (i.e., data were taken during periods when regular science programs could not be executed) or in the face of correlator issues. In the case of the two global VLBI sessions mentioned above, the Band 3 data set was taken under sub-optimal weather conditions at several of the sites, including ALMA, while the Band 6 test had only two functioning partner sites (a third site was rained out). While the data obtained were sufficient to validate basic capabilities and

functional requirements, undertaking full science-quality data reductions from these data is made more challenging because of these shortcomings. Important new insights therefore continue to be gained since the acquisition of Cycle 4 science data by APP team members and collaborators who are co-investigators on Cycle 4 projects.

As of this writing, members of the wider EHT collaboration are actively engaged in the analysis of data acquired from Cycle 4 VLBI programs. The quality of these data is excellent and the data are enabling further testing analysis of pathways and the comparison of results obtained using different software (e.g., HOPS vs. AIPS). Our plan is to gather up the elements from these exercises into a more useful "cookbook" at some later date.

The first efforts on the Cycle 4 ALMA data were the QA2 processing for correlation; this proceeded over the summer and was largely done by August. A few months of additional effort was required to satisify the ALMA requirements for delivery to the PIs. This resulted in a few more iterations on the QA2 deliverables. Going forward, the same processing scripts should work (with at most minor modifications) for the next few Cycles. The resulting VLBI data, as well as the ALMA SFI data were released to the PIs on in October. The QA2 procedures specific to VLBI are the topic of a memorandum currently in preparation by the VLBI QA2 team, but we briefly outline some key points here.

QA2 for the ALMA Cycle 4 VLBI experiments was carried out primarily by ALMA QA2 personnel who also are EHT Consortium (EHTC) members, Ciriaco Goddi (Radboud University) and Ivan Martí-Vidal (Onsala Observatory), but with regular consultation and input from our Cycle 3 Study team. There were two components to the QA2 work. The first was to reduce the ALMA data for appropriate polarization calibrations that are needed to execute the `PolConvert` process at the GMVA and EHTC correlators in Bonn and Haystack, respectively. With the methodology we have adopted for handling mixed polarizations (see Martí-Vidal et al. 2016b and M18 for details) the DiFX correlator produces mixed polarization products (XR, YL, etc.) which are then rotated into a pure circular polarization basis. The process relies on the generation of measurement set (MS) calibration tables on VLBI scans and on and ALMA-only calibration scans using CASA reduction tasks. It was possible to do this work without inventing new CASA tasks, and as with normal QA2 reductions, there is a driver script that may be run from CASA to do the work.

The second QA2 script addresses the imaging reduction of the VLBI science target. VLBI is a non-standard mode, so the standard reductions were not expected to work as coded. Nevertheless, with some straightforward modifications it was possible to make images from the phased data on the VLBI science targets. These results are still in the proprietary period and will be published elsewhere.

In working with the APP data, we have found it useful to build a specialized CASA task `aphase_support.py` that can be invoked either from within CASA or run externally on the ASDM `CalAppPhase.xml` table to obtain some information or perform certain reductions. While this is currently an "expert" tool, it could be modified to be of more general use in the future for non-experts. The help text for the tool is as follows:

```
$ ./aphase_support.py --help
Usage: aphase_support.py [options]

App Phase tool

Options:
  --version             show program's version number and exit
  -h, --help            show this help message and exit
  -v MS, --vis=MS       full path to measurement set
  -x FILE, --xml=FILE   full path to CalAppPhase.xml fileor the directory
                        containing it
  -r STRING, --report=STRING
                        report type, use ``help'' for options
  -k, --keep            preserve previous reports
  -a, --perave          do statistics on a per-average basis
  -o FILE, --output=FILE
                        output file name, or empty for stdout
  -b LIST, --bb=LIST    baseband(s) of interest as comma-sep list.  The
                        basebandsmay be specified as BB_x or simply an integer
  -p LIST, --corr=LIST  correlations: XX, YY or XX,YY (default)
  -s LIST, --scan=LIST  a comma-sep list of scan number ranges, empty is all
  -u STRING, --ints=STRING
                        seconds (at end of scan) to use for calculations

This script parses the CalAppPhase table and can produce various reports from
the table itself.  When used in CASA as the aphase task it can do calculations
based using the measurement set.
```

Most of the arguments affect selecting the data of interest for one of several reports (the `-r` option):

```
$ ./aphase_support.py -r help
```

```
Functionality is determined by the ‘‘report’’ argument:

summary     a simple overview of the file
antennas    provides names of antennas used in phasing system
oneline     provides a one-line report on each scan
caves       reports channel average by baseband

phaserms    reports RMS of phase corrections
phases      reports on phase corrections calculated
updates     reports on phase solutions applied

vstats    * visibility statistics for baselines of interest
vsratio   * computes per-antenna visibility ratios
raweffic  * and computes the raw efficiency from that
pheffic   * and finally computes the (corrected) phasing efficiency
mxeffic   * same idea, but max rather than average

leakage   * estimates per-antenna/per-channel leakage terms
leakratio * computes the correction factors from leakage
leakaves  * and average these

summage   * visibility statistics on norms
sumfactor * compute sum scaling factor from summage

Tasks marked with a * require a MS to be present, the
others require only CalAppPhase.xml (from the ASDM).
(MS processing is via CASA and cannot be done standalone.)
```

6. Study Deliverables

Our Study proposal anticipated six deliverables which we address in the following paragraphs.

Deliverable (i): called for a demonstration of a phasing capability in ALMA Band 7, including quantitative measures of phasing efficiency in Band 7 under a range of observing conditions. *This demonstration has been described in Section 2. Our pending Cycle 5 De-*

velopment award calls for us to commission the capability in order to offer it in Cycle 7 (or later).

Deliverable (ii): called for a document specifying recommended upgrades to the DiFX correlator software to handle the correlation of VLBI data that include baselines with ALMA. *Necessary and sufficient enhancements to the DiFX software have been made and are now included in DiFX 2.5.1, released in 25 September 2017.The DiFX 2.5.0 release was used for initial correlations of the GMVA and EHT data from the 2017 Cycle 4 VLBI Campaign. A discussion of the key technology is discussed above in Section 3.*

Deliverable (iii): called for a memorandum detailing pros and cons of various methodologies for handling the application of BBDs that impact APS data products. *This topic is discussed above in Section 4. Conclusions from this exercise were also presented in our Cycle 4 Study proposal. A delay fix is currently being implemented in that award in conjunction with our pending Cycle 5 Development award.*

Deliverable (iv): called for a cookbook describing recommended calibration and analysis procedures for VLBI data that include phased ALMA as a participating station. *This task remains ongoing. Our explorations to date are summarized in Section 5. This intent of this deliverable has been mostly addressed by the QA2 activities for the Cycle 4 VLBI Observations and the QA2 production scripts.*

Deliverable (v): This deliverable called for a prioritized set of recommendations for future software development to maximize the scientific returns on VLBI data produced using the APS. *Our plan to implement the so-called "delay fix" should render the phased scans equivalent to normal ALMA unphased scans and thus reduce the need for special analysis software (see Section 4). As described in Section 3, HOPS has now been enhanced to better handle recent mm VLBI data products, and through consultation with our team, updates have been made in AIPS to accommodate mm VLBI data that include ALMA and employ the latest generation of VLBI backends (Section 5).*While recommendations for changes to `SCHED` were made, the full suite of changes most likely will not be completed until after the upcoming Cycle 5 (2018 April) 64 Gbps campaign is concluded (or later). The VLBI community has a working draft VEX 2.0 which is to be adopted everywhere, but that transition will likely take years.

Deliverable (vi): This deliverable called for a CASA script to apply delay corrections to the standard ALMA interferometry data produced during APS observations. *This has not proved necessary as the QA2 procedures work without it. See Section 5.*

A. Acronym Definitions

ALMA	Atacama Large Millimeter/submillimeter Array
AIPS	Astronomical Image Processing System
AoD	Astronomer-on-Duty
AOS	(ALMA) Array Operations Site
APEX	Atacama Pathfinder EXperiment
APHIDS	Adaptive Phased-array Heterogeneous Interpolating Downsampler for SWARM
APP	ALMA Phasing Project
APS	ALMA Phasing System
ASDM	Archival Science Data Model
ASIAA	Academia Sinica Instiute of Astronomy and Astrophysics
BB	Baseband
BBD	Baseband Delay
CAN	Controller Area Network
CASA	Common Astronomy Software Applications
CC	Correlator Card
CCC	Correlator Control Computer
CDP	Correlator Data Processor
CSV	Commissioning and Science Verification
DBE	Digital Backend
DDC	Digital Downconvertor
DFT	Discrete Fourier Transform
DGCK	Digitizer Clock
EHT	Event Horizon Telescope
EHTC	Event Horizon Telescope Consortium
FFT	Fast Fourier Transform
FFTW	Fastest Fourier Transform in the West
FPGA	Field-Programmable Gate Array
GMVA	Global Millimeter VLBI Array
Gbps	Gigabits per second
GPU	Graphics Processing Unit
GUI	Graphical User Interface
HOPS	Haystack Observatory Processing System
ICT	Integrated Computing Team
IF	Intermediate Frequency
LCP	Left Circularly Polarized
LMT	Large Millimeter Telescope

LO	Local Oscillator
mm	millimeter
MPIfR	Max Planck Institute for Radio Astronomy
MRI	Major Research Instrumentation
NA	North America
NRAO	National Radio Astronomy Observatory
NSF	National Science Foundation
OSF	(ALMA) Operations Support Facility
PFB	Polyphase Filter Bank
PI	Principal Investigator
ps	Picoseconds
PWV	Precipitable Water Vapor
QA2	Quality Assurance level 2
RCP	Right Circularly Polarized
RDC	Residual Delay Correction
R2DBE	ROACH2 Digital Backend
RMS	Root Mean Square
ROACH	Reconfigurable Open Architecture Computing Hardware
SFI	Single Field Interferometry
SMA	Submillimeter Array
SNR	Signal-to-Noise Ratio
SSR	Science Software Requirements
STE	Standard Test Environment
SWARM	SMA Wideband Astronomical ROACH2 System
TFB	Tunable Filter Bank
TMCDB	ALMA Telescope Monitor and Configuration DataBase
VEX	VLBI Experiment file
VOM	(ALMA) VLBI Observing Mode
VDIF	VLBI Data Interchange Format
VLBA	Very Long Baseline Array
VLBI	Very Long Baseline Interferometry
WVR	Water Vapor Radiometer

REFERENCES

Asada, K., Kino, M., Honma, M. et al. 2017, White Paper, arXiv:1705.04776

Carilli, C. L., Carlstrom, J. E., & Holdaway, M. A. 1999, in ASP Conf. Ser., Synthesis Imaging in Radio Astronomy II, ed. G. B. Taylor, C. L. Carilli, & R. A. Perley (San Francisco, CA: ASP), 565

Cordes, J. M. et al. 2017, ALMA NA Cycle 3 Study Final Report, https://science.nrao.edu/facilities/alma/alma-develop-old-022217/

Cordes, J. M., Kramer, M., Lazio, T. J. W., Stappers, B. W., Backer, D. C., & Johnson, S. 2004, NewAR, 48, 1413

Deller, A. T. Tingay, S. J. Bailes, M., & West, C. 2007, PASP, 119, 318

Deller, A. T., Brisken, W. F., Phillips, C. J., Morgan, J., Alef, W., Cappallo, R., Middelberg, E., Romney, J., Rottmann, H., Tingay, S. J., & Wayth, R. 2011, PASP 901, 275

Doeleman, S. 2010, in VLBI and the New Generation of Radio Arrays, PoS, 53

Eatough, R., Lazio, T. J. W., Casanellas, J., et al. 2015, in Proceedings of Advancing Astrophysics with the Square Kiometre Array, PoS, id. 45

Fish, V., Alef, W., Anderson, J., et al. 2013, White Paper, arXiv:1309.3519

Johnson, M. D. 2016, ApJ, 833, 74

Johnson, M. D. & Gwinn, C. R. 2015, ApJ, 805, 180

Martí-Vidal, I., Crew, G., Matthews, L., and the APP Team, 2016a, APP Memo v1.2-r13

Martí-Vidal, I., Roy, A., Conway, J. & Zensus, A. J. 2016b, A&A, 587, A143

Matthews, L. D. & Crew, G. B. 2015a, ALMA Technical Note 16

Matthews, L. D. & Crew, G. B. 2015b, ALMA Technical Note 17

Matthews, L. D. & Crew, G. B. 2015c, ALMA Technical Note 18

Matthews, L. D., Crew, G. B., Doeleman, S. S. et al. 2018, PASP, 130, 015002 (M18)

Matthews, L. D., Crew, G. B. & Fish, V. L. 2017, ALMA Technical Note 19

Mignani, R. P., Paladino, R., Rudak, B., et al. 2017, ApJ, 851, L10

Pradel, N. 2014, "Zoom band test report, beta version", ALMA Phasing Project internal report, https://deki.mpifr-bonn.mpg.de/@api/deki/files/5305/=Zoomband_test_report.pdf

Primiani, R. A., Young, K. H., Young, A., Patel, N., Wilson, R. W., Vertatschitsch, L., Chitwood, B. B., Srinivasan, R., MacMahon, D., Weintroub, J., 2016, J. Astron. Instr. Volume 5, Issue 4, id. 1641006-810

Psaltis, D., Wex, N., & Kramer, M. 2016, ApJ, 818, 121

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

Shepherd, M. C. 1997, "Difmap: an interactive program for synthesis imaging", Astronomical Data Analysis Software and Systems VI, ASP Conference Series, Vol. 125, 77

Tilanus, R. P., Krichbaum, T. P., Zensus, J. A., et al. 2014, White Paper, arXiv:1406.4650

Torne, P., Eatough, R. P., Karuppusamy, R., et al. 2015, MNRAS, 451, L50

Vertatschitsch, L., Primiani, R. Young, A, Weintroub, J., Crew, G. B., McWhirter, S. R., Beaudoin, C., Doeleman, S., & Blackburn, L., 2015, PASP, 127, 1266