A Dedicated Pulsar Timing Array Telescope (a.k.a. "Pulsar Town") ngVLA Memo #34

Robert Selina, Paul Demorest, Wes Grammer

May 15, 2018

Abstract

As part of the transition plan from the VLA to the ngVLA, we describe re-purposing the VLA antennas for a dedicated Pulsar Timing Array Telescope (PTAT). The VLA antennas and supporting systems would be retrofitted with ngVLA technology with two purposes: (1) reduce system complexity to allow the operation of the array for a fraction of current costs, and (2) improve the instantaneous bandwidth and sensitivity of the instrument for this use case. Proposed changes could allow the PTAT to be supported for a fraction (goal of 1/5 to 1/10) of the current VLA operations cost, while providing the U.S. pulsar community with a dedicated instrument with ∼1.5x GBT continuum sensitivity at f_c of 1.5 GHz. Depending on the deployment schedule, such an instrument could bridge the pulsar community to the ngVLA. It could also complement pulsar timing observations on the ngVLA by providing lower frequency coverage, and increased telescope time on bright sources.

1 Introduction

Timing observations of radio millisecond pulsars (MSPs) provide us with access to a set of high-precision astronomical “clocks” that may be used for a variety of physics and astrophysics experiments. For example, pulsar timing is able to accomplish some of the most constraining tests of theories of gravity, investigate the properties of matter at super-nuclear density and provide a unique probe of the intervening ionized interstellar medium. High-precision timing observations of a set of MSPs can also act as a probe of very low frequency gravitational waves (GW); this concept is known as a pulsar timing array (PTA). PTAs are sensitive to GW with frequency on the order of the inverse experiment duration, typically in the nHz range, and as such are complementary to higher-frequency (kHz) GW experiments like LIGO.

Pulsar timing array sensitivity to GW is dramatically improved by increasing the number of pulsars that are monitored [SEJR13]. However, as telescopes become larger, timing results tend to become dominated by intrinsic noise processes in the pulsars themselves [L+16]; even with a high-gain instrument, long observations continue to be required to reduce this noise. Because of this, there is growing realization that large shared-use facilities such as the ngVLA and SKA may not be able to efficiently accomplish a next-generation PTA experiment, and that a lower-cost, dedicated facility (even one with lower sensitivity) is needed to complement the large telescopes.

The VLA antennas are not reused within the ngVLA array. Absent another use, they may be decommissioned. We propose to re-purpose the antennas and significant portions of the existing electronics and infrastructure to turn the VLA into a dedicated Pulsar Timing Array Telescope (PTAT). The array would operate with ∼80% of clock hours spent timing PTAs, and the remaining ∼20% allocated to maintenance/downtime and test time. The instrument would be operated as a finite duration experiment (rather than a multi-user observatory) and operations would be simplified by supporting pulsar timing (PT) modes only.

In our proposed concept, the antennas would be permanently left in D or C-configuration, and the electronics simplified to reduce the maintenance burden. A single prime focus ultra wide band (UWB) receiver system would replace the focus rotation mechanism (FRM) and all existing receivers. The backend would be replaced in its entirety with a beamformer/correlator and pulsar timing back-end to free up space in the control building for ngVLA systems. Such a system could be housed in modified shipping containers similar to the CHIME back-end. Using ngVLA technology would reduce both NRE and operational/maintenance overheads.

2 The Need for a Dedicated Pulsar Timing Telescope

Precision timing of pulsars depends on state of the art sensitivity and significant time on sky. The ability to detect perturbations in the pulse period caused by gravitational waves is dependent on the number and distribution of the pulsars within the pulsar timing array (PTA), requiring many observations of discrete sources across the sky. The optimal observing strategy differs for various GW signal types. Detection and characterization of a stochastic GW background requires large numbers of pulsars monitored. The NANOGrav project currently monitors ∼70 pulsars using roughly 10% of the total available time at the GBT and Arecibo; a future timing program will consist of at least ∼200 pulsars. In contrast, sensitivity to GW signals from discrete sources (individual SMBH binaries) is improved by maximizing timing sensitivity on a smaller set of high-quality pulsars; for future work timing improvements of at least a factor of a few versus the current state of the art are necessary. A general-purpose GW detection program will incorporate both strategies.

A simple expectation is that, all else being equal, a telescope with an improvement in point source sensitivity (SEFD) of a factor of x would require a factor of x^2 less telescope time to achieve results comparable to those of a less sensitive system. However, at high sensitivity pulsar timing results eventually become dominated by noise processes intrinsic to the pulsar radio emission, rather than by radiometer noise. This intrinsic noise is known as "pulse jitter," and will add in quadrature to the usual radiometer noise. A simplified expression for pulse time of arrival (TOA) uncertainty (σ), taking both jitter (σ_J) and radiometer noise (σ_R) into account is (see for example [L+16]):

$$\sigma^2 = \sigma_R^2 + \sigma_J^2 = \frac{w^3}{P}\left(\frac{S_{sys}}{S_{psr}}\right)^2 \frac{1}{BT} + c^2 w^2 \frac{P}{T} \tag{1}$$

Here, w is the pulse width, P is the pulse period, S_{sys} and S_{psr} are the SEFD and pulsar flux respectively, B is the total bandwidth, T is the total integration time, and c is a factor of order unity related to the jitter properties. For faint pulsars (or less sensitive telescopes) the σ_R term dominates, and results can be improved by reducing SEFD or increasing bandwidth. For bright pulsars (or more sensitive telescopes), the σ_J term dominates and only increased integration time helps.

This effect is illustrated graphically in Figure 1, which uses Eqn. 1 with typical values for pulse width, period, bandwidth, etc, to compare timing results from two hypothetical timing programs: One uses a sensitive telescope (SEFD of 1.5 Jy, comparable to ngVLA or SKA1-mid), whereas the other uses a less-sensitive telescope (10 Jy, comparable to GBT) but with a factor of 10 more observing time. The "less-sensitive" program produces better results for sources brighter than about 0.5 mJy. It seems reasonable to expect that, even if very highly rated, any single project is unlikely to obtain more than ∼5–10% of the total available time on a competitive shared-use facility like ngVLA. Therefore improving PTA results significantly beyond the current state of the art is likely to require a combination of an order of magnitude increase in observing time (i.e., a dedicated instrument) with current (∼GBT-level) sensitivity, as well as smaller amounts of time on future, more sensitive instruments (ngVLA/SKA) for monitoring fainter pulsars.

3 Scientific and Technical Requirements

Pulsar timing observations require monitoring pulsars of known sky position, pulse period and dispersion measure. The array is phased and coherently summed to form a single beam on the target source. The signal is processed into specified frequency resolution, coherently dedispersed, detected, folded (averaged modulo the known pulse period) into pulse profiles consisting of a specified number of pulse phase bins, and recorded once every few seconds. The relevant measurement eventually extracted is a pulse time of arrival, i.e. the average pulse phase shift observed between the measured profiles and a model for the specific pulsar. In addition to the sensitivity and integration time requirements discussed in the previous section, high-precision pulsar timing also requires accurate timekeeping, and careful polarimetry. While timing is not intrinsically a polarization-based measurement, calibration errors will affect the observed pulse shapes, leading to systematic errors in the timing measurements.

Therefore, the basic, high-level scientific requirements for the PTAT are:

- Continuum Sensitivity comparable to GBT. Desirable to have sensitivity of order 2x GBT.

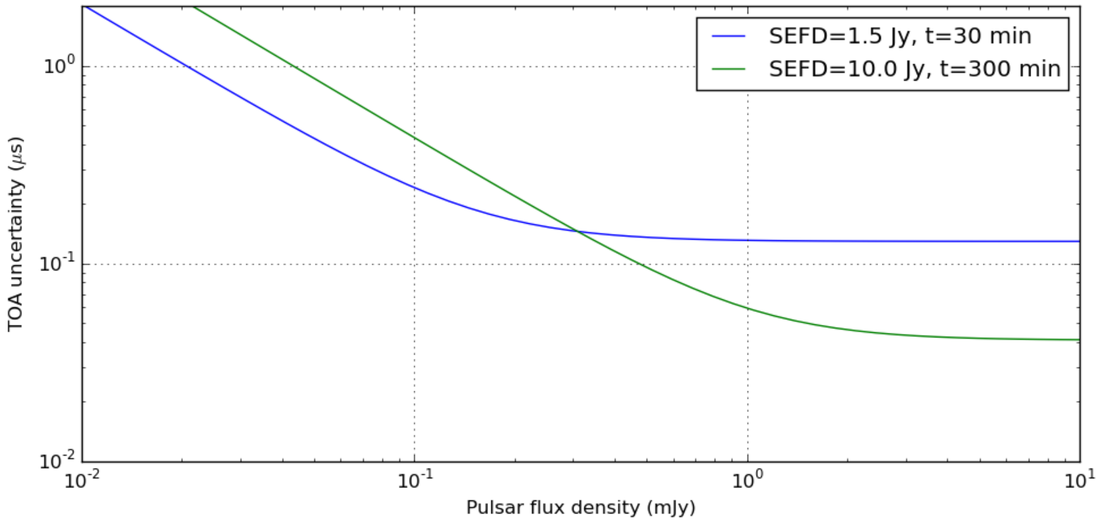

Figure 1: Typical expected time-of-arrival uncertainty versus pulsar flux density, incorporating both jitter and radiometer noise effects, for two hypothetical observing programs. More time on the less-sensitive telescope produces better results for pulsars brighter than ∼0.5 mJy.

- Frequency span overlapping with current instruments and programs, largely 300 MHz to 3 GHz. Desirable to provide frequencies up to 4 GHz.
- Beamforming, coherent dedispersion, folding and recording for a single beam on sky. Desirable to support multiple subarrays and beams (3 or less).
- Operational availability, slew rates and supporting specifications to monitor of order 200 MSPs on a weekly cadence with a minimum of 30 minutes on source within that epoch.
- Ability to correct data timestamps to a known time standard (e.g., TAI or GPS) with an error of at most 10 ns. This correction can be retroactively applied.
- Ability to perform polarimetric calibration with an accuracy of ∼3% on boresight. Stability is more important than absolute calibration.

The technical requirements that guide this concept are shown in Table 1.

Description	Requirement	Goal
Minimum Frequency	1 GHz	300 MHz
Maximum Frequency	2 GHz	4 GHz
Bandwidth	1 GHz	3.7 GHz
SEFD	10 Jy	5 Jy
Frequency Resolution	1 MHz	100 kHz
Profile Bins	2048	
Polarization accuracy	3%	1%
Pulse Period	1 ms - 1 s	
Sub-arrays	1	3

Table 1: Key Technical Requirements of the PTAT.

4 Technical Concept

In order to both reduce the maintenance burden and improve the sensitivity of the VLA antennas for the dedicated PTAT use case, we propose the following broad changes to the system:

- All 28 antennas are placed in D-config (or C-config) permanently. Antennas would receive any major maintenance and overhaul before being placed in position, so the reconfiguration track can be abandoned in place or removed. (No future service in the antenna barn).
- Altitude and Azimuth gear drives are tuned to reduce wear/tear at the expense of pointing performance and slew speed. Performance tuned for observations below 5 GHz.
- A new single 500 MHz to 2.5 GHz wide-band receiver is installed at the prime focus. The focus-rotation mechanism (FRM) and subreflector are permanently removed. Amplified RF signals are transmitted over coax to the vertex room.
- The 2 GHz of bandwidth is directly sampled by a new 8-bit integrated receiver module located in the vertex room. The integrated receiver transmits the unformatted serial link to the central beamformer over existing fiber optic lines.
- The central beamformer (CBF) will produce up to 3 beams of 2 GHz bandwidth that are transmitted to pulsar timing engines. The CBF is based on the existing NRC frequency slice processor architecture [PZC+17].
- The pulsar timing engine (PTE) coherently dedisperses, detects and folds the data in to a specified number of phase bins, which are recorded on 1 to 30 second periods. The PTE would likely be a GPU-based architecture using commercial off-the-shelf hardware.
- Existing VLA time and frequency distribution electronics are retained and reused. Sampler clocks are synthesized at the antenna by modifying the L300 to produce the 5120 MHz clock by installing a new bandpass filter. All other LO/IF equipment in the vertex room is removed.

Note that the technical solution proposed is just one of many alternatives, and is provided as a straw-man for the evaluation of the broader VLA-PTAT concept.

4.1 Antenna Optics

The VLA antenna is a 25m aperture Cassegrain system with a focus-rotation mechanism (FRM) for band selection and focus adjustment. In order to minimize blockage, a relatively small subreflector

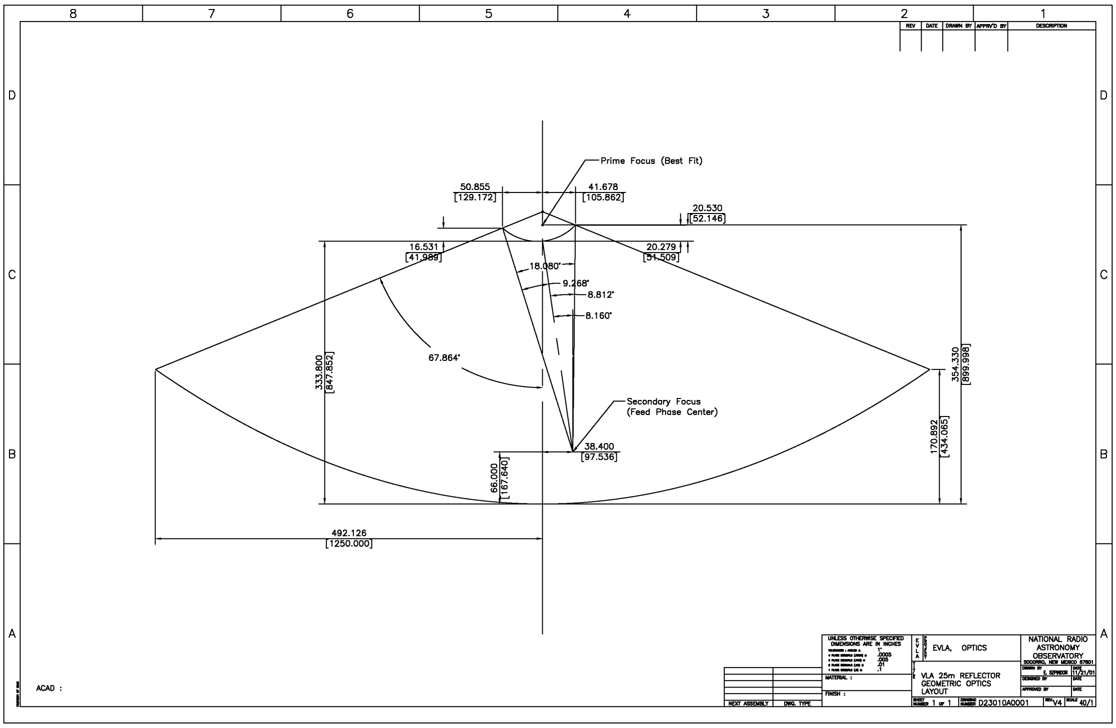

Figure 2: Optical sketch of the VLA antenna.

of 2.3m in diameter is used (Figure 2). This geometry is an appropriate compromise for high frequency observation, but is suboptimal for 1 GHz operation, leading to diffraction effects and increased T_{spill} below 2 GHz.

In addition, the FRM is a source of complexity and requires frequent repair and maintenance. The elevation and azimuth drive systems, by comparison, are relatively trouble free. The prospect of removing the FRM makes a prime focus instrument attractive. However, the VLA optics are a shaped pair, so the main reflector and subreflector both have a deviation from the canonical parabola/hyperbola. The main reflector deviates by of order 25cm from its canonical shape, leading to destructive interference over the aperture when sampled at the prime focus. This term, referred to here as η_{phase}, reduces aperture efficiency as we increase in frequency (Table 2). The total aperture efficiency at the prime focus is still competitive and useful, but this feature of the antenna restricts the operating frequency of prime focus to frequencies below 2.5GHz.

The prime focus feed package can be easily accommodated within the support ring connecting the quad-legs. The reduced mass of the receiver system (compared to the FRM) should also reduce gravitational deformation with elevation. At the frequencies of interest, it is expected that the system can be collimated mechanically/manually and left in a fixed position with no adjustment required for thermal deformation or other effects.

4.2 Receiver Concept

The proposed receiver concept is an ultra-wideband (5:1, 0.5-2.5 GHz) prime focus receiver based on the Parkes UWB 0.7 - 4.2GHz receiver [D+17] (Figure 3), but with a frequency-scale of the feed by a factor of ∼0.8 to move the lower frequency cutoff to below 500 MHz. This design employs a novel dielectrically-loaded quad ridge feed horn and has been both constructed and measured. Projected performance of the front end is summarized in Table 2.

$f_L(GHz)$	η_{illum}	η_{focus}	η_A	η_{spill}	T_{spill}, K	T_{rx}, K	T_{sky}, K	T_{sys}, K	$T_{sys}/\eta_A, K$
0.50	0.72	0.96	0.59	0.95	15	15	12	47	79
1.00	0.65	0.89	0.50	0.98	6	6	4.5	21.5	43
1.50	0.68	0.77	0.45	0.98	6	7	4.5	22.5	50
2.00	0.66	0.64	0.36	0.98	6	7	4.5	22.5	62
2.50	0.65	0.50	0.28	0.99	3	8	4.5	20.5	73

Table 2: Estimated Ultra-Wideband (UWB) Prime-Focus Receiver Performance on a VLA Antenna.

- η_{illum} and η_{spill} from [D+15] A. Dunning et. al, "An Ultra-Wideband Dielectrically Loaded Quad-Ridged Feed Horn for Radio Astronomy" (2015); Figure 5, frequency-scaled by 0.8. Assumes f/d=0.41 case, for minimum T_{sys}/η and spillover at prime focus.
- η_{focus} from [Sri17] S. Srikanth, NRAO, private communication. Estimated phase efficiency of the shaped VLA antenna primary reflector, at the prime focus.

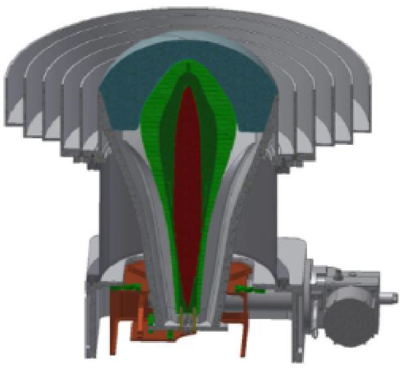

Figure 3: CSIRO developed 6:1 UWB Quad-Ridge Feed design [D+15] [D+17].

- η_A assumes negligible surface efficiency degradation, i.e. $\eta_{ruze} = 1.0; \eta_{block} = 0.86$
- T_{rx} from [T+17] Tzioumis, "Technologies for Radio Astronomy", CSIRO, Oct. 2017, p.7.
- T_{sky} from [BS18] B. Butler and F. Schinzel, NRAO, private communication. Assumes antenna elevation angle of 45 degrees. Sky temperature at 500 MHz can vary from 12-39 K, depending on how much of the galaxy is in the beam.
- T_{sys} is the sum of T_{spill}, T_{rx}, and T_{sky} and T_{misc}, where $T_{misc} = 5K$. This accounts for an additional ~5K of noise of unknown origin measured during on-antenna tests at Parkes. [T+17] VLA performance may vary.
- The full-array SEFD in Jy can be computed as $0.2T_{sys}/\eta_A$.

The overall T_{sys}/η_A is ~33% better than the VLA L-band receiver [Gra17] at the center of the band as can be seen in Figure 4. If pursued, the design should be investigated in detail to best match the f/D of the VLA optics and the mechanical attachments available.

The majority of the feed would by necessity be warm given its volume. The LNAs would ideally be cooled to 20K for optimal noise temperature as shown in Table 2. Reusing the existing CTI Gifford-McMahon refrigerator and compressor would be the lowest cost option, and would reduce the existing electrical load to ~1/3 of present (given 1 of 3 compressors running). The preferred approach would be to use dual-stage sterling cycle coolers. Electrical costs could be reduced to less than 1/8 of present and the maintenance cycle could be extended by a factor of three or more. ngVLA will be investigating these designs in the near future, and leveraging ngVLA technology may make such a system affordable. A fall back option would be a commercially available 70K sterling cycle cooler, which would offer the operational benefits (low electrical and maintenance costs) at the expense of ~5K of noise.

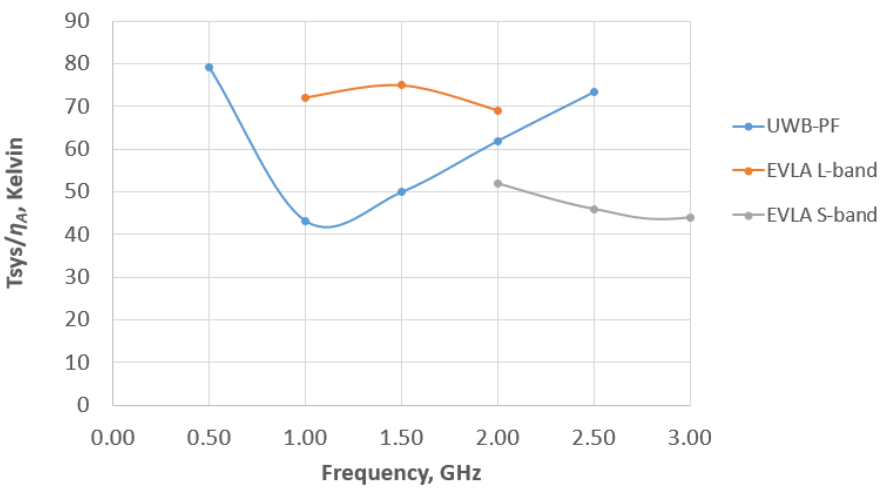

Figure 4: Sensitivity of the UWB PF receiver compared to EVLA receivers.

4.3 Antenna Electronics

The antenna electronics concept would be an integrated receiver module as described in [[MW17]]. The RF from both polarizations would be directly sampled by an IRD module mounted behind the cryostat.An existing EVLA receiver card cage and F318 could be used or M&C of the receiver. An existing vacuum pump would also be relocated to the prime focus.

The existing LO distribution system would provide the 1024 MHz reference from the L305 to a 5x multiplier that would supply the sampler clock.

Data is streamed off the antenna as an unformatted serial digital link and delivered to the central signal processor.

4.4 Central Signal Processor

The Central Signal Processor (CSP) concept is based on the National Research Council of Canada's Processor (FSP) architecture [PZC+17], scaled for the purpose. The CSP will align the bits received from each antenna and recover the sampler clock. Data from each antenna will then be processed by a very coarse channelizer (VCC) to produce delay-corrected 200 MHz subbands. Each 200 MHz subband will be cross-correlated between all antenna pairs, as well as coherently summed to produce the formed beams.

Although cross-correlation or imaging is not directly required for pulsar timing measurements, it is necessary in order to determine per-antenna delay and phase corrections for forming beams. The time/frequency resolution requirements for the cross-correlator are to be determined, but are not necessarily as demanding as they would be for a general-purpose imaging correlator. Simultaneous beamforming and cross-correlation is also not strictly required, but may be desirable to allow (for example) imaging-based commensal science opportunities.

Each coherently summed beam will be transmitted to a new pulsar timing engine. Here the signal is processed into specified frequency resolution, coherently dedispersed, detected, folded into a specified number of pulse phase bins, and recorded to disk once every few seconds. The resulting data rate is modest even by current standards, typically on the order of ∼10 MB/s.

The pulsar timing backend system is expected to be built using small compute clusters equipped with GPUs. The GUPPI pulsar backend at Green Bank [FDR10] includes a cluster of eight 2009-era GPU cards (Nvidia GTX 285), each of which can process 100 MHz bandwidth in real time. Current-generation GPUs have an order of magnitude more processing power; recent benchmarking of a GTX 1080 indicates it could handle up to ∼800 MHz BW in a similar operating mode. Even with no assumption of future improvement in technology, three 2-GHz beams for pulsar timing could be handled by a small (∼10-node) computer system. Custom software development is likely to be minimal, and can leverage both the large amount of existing mature software (e.g., GUPPI, dspsr [vB11]) as well as new development done for the ngVLA.

5 Schedule

The deployment schedule of the PTAT concept would be dependent on the ngVLA construction schedule, the transition plan from VLA operations to ngVLA operations, and funding for PTAT design and construction.

The ngVLA design and construction baseline schedule has design activities through 2024 followed by a ten year construction and commissioning phase. The transition from commissioning to operations is gradual, with first science operations scheduled in 2028.

The pulsar timing community need continued, and increasing, access to existing facilities such as Arecibo, GBT and VLA until the PTAT is operational. Constructing the PTAT prior to ngVLA first science would mitigate the loss of an existing facility to PTA projects, but has complex implications for the broader US radio astronomy community and staff availability for the ngVLA.

Deploying the PTAT earlier in the ngVLA design phase could free up critical VLA-support staff to work on ngVLA design, but would of necessity imply that existing VLA operations end prior to ngVLA commissioning. Alternatively, the PTAT project construction could be deferred until after ngVLA first science as part of an operational transition from VLA to ngVLA. This keeps existing VLA capabilities available longer, at increased risk to the PTA projects. If the PTAT project is further developed, the schedule will need to be negotiated amongst multiple stakeholders.

The length of the operations period is flexible and dependent on operations funds as well as the condition and practicalities of continued repair of the VLA antennas. A significant operations period (of order 10 years) would likely be required to justify the capital investment described above.

6 Acknowledgements & Caveats

While this memo has been written by ngVLA project team members, this should not be construed as ngVLA/NRAO support for this option. This memo is provided to document a possible case for this instrument and its technical feasibility. Inclusion of the dedicated PTAT in the VLA-ngVLA transition plan will depend on community support and a demonstrated need, by the radio astronomy community, for the PTAT.

References

[BS18] B. Butler and F. Schinzel. Private communication, 2018.

[D+15] A. Dunning et al. An ultra-wideband dielectrically loaded quad-ridged feed horn for radio astronomy. 2015.

[D+17] A. Dunning et al. An ultra-wideband cryogenic receiver for the parkes radio telescope. 2017.

[FDR10] J. M. Ford, P. Demorest, and S. Ransom. Heterogeneous real-time computing in radio astronomy. In *Software and Cyberinfrastructure for Astronomy*, volume 7740 of *Proc. SPIE*, page 77400A, July 2010.

[Gra17] W. Grammer. Ngvla receiver parameters v4.2, 2017.

[L+16] M. T. Lam et al. The NANOGrav Nine-year Data Set: Noise Budget for Pulsar Arrival Times on Intraday Timescales. *ApJ*, 819:155, March 2016.

[MW17] M. Morgan and S. Wunduke. An integrated receiver concept for the ngvla. *ngVLA Memo Series, Memo #29*, 2017.

[PZC+17] Michael Pleasance, Heng Zhang, Brent Carlson, Ralph Webber, Dean Chalmers, and Thushara Gunaratne. High-performance hardware platform for the square kilomtre array mid correlator & beamformer. August 2017.

[SEJR13] X. Siemens, J. Ellis, F. Jenet, and J. D. Romano. The stochastic background: scaling laws and time to detection for pulsar timing arrays. *Classical and Quantum Gravity*, 30(22):224015, November 2013.

[Sri17] S. Srikanth. Private communication, 2017.

[T+17] Tzioumis et al. Technologies for radio astronomy. 2017.

[vB11] W. van Straten and M. Bailes. DSPSR: Digital Signal Processing Software for Pulsar Astronomy. *PASA*, 28:1–14, January 2011.