An Ultra-High Time Resolution Cosmic Ray Detection Mode for the Murchison Widefield Array

A. Williamson\(^1\), C.W. James\(^1\), S.J. Tingay\(^1\), S.J. McSweeney\(^1\), S.M. Ord\(^1\)

\(^1\)International Centre for Radio Astronomy Research, Curtin University, Bentley, WA 6102, Australia, alexander.williamson1@postgrad.curtin.edu.au

Received (to be inserted by publisher); Revised (to be inserted by publisher); Accepted (to be inserted by publisher);

The radio-wavelength detection of extensive air showers (EAS) initiated by cosmic-ray interactions in the Earth’s atmosphere is a promising technique for investigating the origin of these particles and the physics of their interactions. The Low Frequency Array (LOFAR) and the Owens Valley Long Wavelength Array (OVRO-LWA) have both demonstrated that the dense cores of low frequency radio telescope arrays yield detailed information on the radiation ground pattern, which can be used to reconstruct key EAS properties and infer the primary cosmic-ray composition. Here, we demonstrate a new observation mode of the Murchison Widefield Array (MWA), tailored to the observation of the sub-microsecond coherent bursts of radiation produced by EAS. We first show how an aggregate 30.72 MHz bandwidth (3072 × 10 kHz frequency channels) recorded at 0.1 ms resolution with the MWA’s voltage capture system (VCS) can be synthesised back to the full bandwidth Nyquist resolution of 16.3 ns. This process, which involves ‘inverting’ two sets of polyphase filterbanks, retains 90.5% of the signal-to-noise of a cosmic ray signal. We then demonstrate the timing and positional accuracy of this mode by resolving the location of a calibrator pulse to within 5 m. Finally, preliminary observations show that the rate of nanosecond radio-frequency interference (RFI) events is 0.1 Hz, much lower than that found at the sites of other radio telescopes that study cosmic rays. We conclude that the identification of cosmic rays at the MWA, and hence with the low-frequency component of the Square Kilometre Array, is feasible with minimal loss of efficiency due to RFI.

Keywords: cosmic rays; radio astronomy; astronomical methods

1. Introduction

The digital era of radio astronomy has led to a revival of the radio-detection of cosmic rays (Huege, 2016; Schröder, 2017). Extensive air showers (EAS) initiated by cosmic-ray interactions in the atmosphere produce coherent bursts of low frequency (\(\sim 100\) MHz) radio-wave radiation over a localised ground region of a few hundred metres diameter. The dense cores of modern low-frequency astronomical telescopes are well-suited for observing this ground pattern from cosmic rays in the \(10^{16}-10^{18}\) eV energy range, and provide complementary information to dedicated cosmic ray facilities such as the Pierre Auger Observatory (Aab \textit{et al.}, 2016a) and Tunka-Rex (Bezyazeekov \textit{et al.}, 2015). LOFAR, the Low Frequency Array (van Haarlem \textit{et al.}, 2013), has demonstrated that studying the ground pattern yields precise information on the primary cosmic rays, and can be used to study their spectrum and, hence, their origin (Schellart \textit{et al.}, 2013; Buitink \textit{et al.}, 2014, 2016). The Owens Valley Long Wavelength Array (OVRO-LWA) has also recently developed a cosmic-ray detection mode (Monroe \textit{et al.}, 2020), and cosmic-ray observations with the low-frequency component of the Square Kilometre Array, SKA1-Low, have been proposed (Huege \textit{et al.}, 2015). However, cosmic ray signals appear similar to a bandwidth-limited impulse, so that data analysis ideally requires access to baseband data sampled at the Nyquist rate (direct measurements of the voltage induced at the antennas by the radio waves). This makes it necessary for telescopes aiming to detect cosmic rays to develop a dedicated, high-time-resolution observation mode.

\(^1\)Corresponding author.
The Murchison Widefield Array (MWA) is a radio interferometer located at the Murchison Radio-astronomy Observatory (MRO) in the Mid West of Western Australia (Tingay et al., 2013). Following an upgrade, the MWA consists of 256 tiles, of which 128 can be used concurrently, spread across an area greater than 20 km² (Wayth et al., 2018). Each tile consists of 16 dual polarisation antennas arranged in a grid, which are analog beamformed prior to digitisation. The MWA operates between 80 and 300 MHz, with a bandwidth of 30.72 MHz. Its location in a pristine radio-quiet environment, layout, and frequency range make it well-placed to study the radio pulses from EAS initiated by cosmic-ray interactions in the Earth’s atmosphere (Beardsley et al., 2019). To this end, a prototype particle detector has already been installed at the MRO to be used as a trigger (Bray et al., 2020), similar to LOFAR’s LORA detector (Thoudam et al., 2014).

The MWA can return 30.72 MHz of processed bandwidth, comparable to the effective bandwidth of LOFAR. This makes cosmic-ray detection with the MWA an attractive experimental prospect, especially since it has been estimated that the 100-200 MHz range provides the greatest sensitivity to cosmic-ray signals (Aswathi et al., 2020). Key to implementing cosmic ray detection with the MWA is the reconstruction of a time series equivalent to a Nyquist sampled 30.72 MHz bandwidth, starting from the 3072×10 kHz channels produced by the default MWA analog/digital signal chain.

This paper describes a new processing pipeline for the MWA, focused on the detection of cosmic ray signals, which performs this reconstruction process. In Section 2, we briefly describe the properties of the polyphase filterbanks (PFBs) that perform the coarse and fine channelisation steps in the MWA signal chain, and the previously published process for inverting the fine PFB. We then describe our methods for inverting the coarse PFB, and present the expected response of this process to cosmic ray signals. In particular, we focus on the challenges faced when only 24 of 256 channels are available for inversion, and our solution. Section 3 describes an on-site experimental validation of this mode using a calibration pulse, which demonstrates that we have obtained our desired inverse-bandwidth resolution. Section 4 then provides a characterisation of the radio-frequency interference (RFI) environment at nanosecond timescales at the MRO. We present our conclusions in Section 5.

2. Polyphase Filterbank (PFB) and inversion implementation

PFBs are commonly used in radio astronomy to convert time series voltages measured from an antenna into their conjugate frequency space in a fast and efficient manner, while minimising spectral leakage between the resulting frequency channels. The process can be generalised to a decimator that first separates the signal into multiple phases that each undergo a finite impulse response (FIR) filter and a Fourier transform (FT), and are then summed across the phases (Crochiere & Rabiner, 1983). The result is a time series for each frequency channel. Since the FT and FIR filter operations commute, the FIR can be represented in the frequency domain and applied after the FT. This representation of a PFB is shown in Figure 1; the time-domain representation of the coarse channel FIR filter is shown in Figure 2.

MWA tile-beamformed voltages for each X and Y polarisation are sampled at 655.36 Msamples s⁻¹ (time resolution of 1.53 ns), and channelised using a PFB into 256x1.28 MHz ‘coarse’ channels (Prabu et al., 2015). This gives a time resolution of 781.3 ns, with 5 real bits and 5 imaginary bits. This first PFB stage uses eight taps of 512 samples each. Of these coarse channels, 24 are returned for central processing, and are passed through a fine channeliser, producing a total of 3072×10kHz channels with a time resolution of 100 µs (Ord et al., 2015). This second PFB stage uses 12 taps of 128 samples each.

The MWA voltage capture system (VCS) allows these fine channels to be recorded directly to disk, which is the observation mode commonly used for pulsar observations (Tremblay et al., 2015). One second of VCS observation produces 3072 frequency channels with 10⁴ time samples, for each of 128 MWA tiles for 2 polarisations, resulting in 7.3 GB of data. This means that a one hour observation produces a non-trivial 25.7 TB of data.

In order to search for and reconstruct cosmic ray signals with inverse bandwidth resolution, this large quantity of data must be returned to inverse-bandwidth time resolution by inverting both fine and coarse PFBs, and then searched for cosmic-ray signals.

The inversion process to reconstruct coarse channels from fine filterbank data is described in McSweeney
Fig. 1. Schematic of a PFB process, using the frequency domain representation of the filters. 1) An initial time-domain data stream at rate $dt$ is decimated into blocks of length equal to the number of desired channels, $N_{\text{chan}}$. 2) A discrete Fourier transform (DFT) is then applied along the channel axis. 3) The data are convolved with a forward filter to produce a channelised form of the input data with the time resolution reduced by a factor equal to half the number of channels formed. The inverse process is similar, with the exception that the convolution and (inverse) DFT stages are swapped.

et al. (2020), and has already been validated on pulsar observations (Kaur et al., 2019). The inverse PFB process is qualitatively similar to the operation of the PFB itself: the data are Fourier transformed back to the time domain, convolved with a FIR filter, and unfolded. Since the PFBs present in the MWA are critically sampled, it is not generally possible to perfectly reconstruct the input time series from the fine channel voltages, due to significant leakage of power between fine channels near the coarse channel edges. Thus we do not investigate the inversion method described by Morrison et al. (2020) and used for fast radio bursts detected by the Australian Square Kilometre Array Pathfinder (Cho et al., 2020), which is based on an oversampled PFB. Therefore, our aim is to choose an inverse filter to reconstruct the input data as accurately as possible.

McSweeney et al. (2020) investigate several possible inverse filters, finding that the reconstruction accuracy of the coarse channels improves with the number of taps used for the inverse filter. In our application, however, reconstruction of the time-domain data will be limited by the small number of coarse channels available for the inversion process. We therefore use the ‘mirror filter’ described by McSweeney et al. (2020) for the inversion of both fine and coarse channel data. That is, the forward and inverse filters are identical.

2.1. Effect of finite channels

The accuracy of reconstructing a cosmic ray signal with MWA data is primarily limited by the availability of only 24 of 256 coarse channels for further processing. The simplest treatment is to set all unavailable coarse channels to zero prior to the coarse PFB inversion. However, this will cause significant oversampling in the output and will cause the processing to be dominated by calculations on empty elements. This effect can be negated by choosing the 24 coarse channels to fit into the same Nyquist zone. Fourier transforming only the coarse channels in that zone — including complex conjugation if it is an odd zone — will produce a time-domain signal with minimal oversampling. For dedicated observations, the 24 channels can be chosen to lie exactly within a 30.72 MHz Nyquist zone, resulting in time-domain data at 61.44 Msamples s$^{-1}$. In general, a larger band must be chosen, e.g. a Nyquist zone 32 coarse channels wide corresponds to a 1/16 downsampling of the original time-domain data to a rate of 81.2 Msamples s$^{-1}$, with 75% band occupancy. Only the FIR filter coefficients corresponding to the used channels are applied prior to the
Fig. 2. The FIR filter used for both the coarse PFB and its inversion. It is a sinc function convolved with a Hanning window and is split into eight taps of 512 samples each. The amplitude is unitless and normalised to unity.

Fourier transform.

We reproduce real-valued time-domain data using a complex-to-real inverse FFT from FFTW (Frigo & Johnson, 2005). The data are then aligned into the time series and written to disk. The output consists of a file for each second, tile, and polarisation.

2.2. Performance

Figure 3 presents the effect of our PFB inversion on a test impulse passed through software that emulates the MWA coarse and fine filterbanks. The signal is then reconstructed as described above using all 256 coarse channels (i.e. no downsampling), and compared to the input. The same signal is then band-limited to 24 coarse channels (96–119), and inverted as described in the previous section.

The performance of this process varies with the location of the initial impulse within the 512-sample tap length. When the signal is located near the edge of the tap, reconstruction using 256 channels is almost perfect, since PFB coefficients are dominated by the central branch. In the middle of the tap however, the performance is worst, since the PFB coefficients are spread out over the branches. The best and worst cases shown in Figure 3, respectively, correspond to impulses at the edge and centre of the tap. The performance as a function of tap position is shown in Figure 4.

From Figure 3, the 256 channel inversion in the best case produces a perfect impulse. In general however, the peak of the signal is reduced in amplitude, with secondary peaks showing up as artefacts separated by the tap length of 0.78125 µs. This results in a reduction in pulse height (which is the signal-to-noise ratio, S/N, when using a peak detection method) of up to 19%. The reduction is similar when limited to a 24 channel band. The average of the 256-channel performance when varying the location of the impulse in the initial data stream (as per Figure 4) is 90.5%. Since electric field peak height is almost proportional to primary cosmic ray energy (Aab et al., 2016b), the cosmic ray spectrum above the knee of \( dN_{CR}/dE_{CR} \sim E_{CR}^{-3.1} \) (Apel et al., 2012) implies a reduction of 18% in event rate compared to an idealised experiment.

3. Demonstration with a pulse calibrator

In order to test the effectiveness of the inverse PFB, an experiment was conducted at the MRO during scheduled maintenance. A hand-held barbeque gas lighter, which generates a spark via a piezoelectric, was
used to produce radio pulses that were both strong enough to be detected by an MWA tile, and with sufficient structure on nanosecond scales to test the cosmic-ray mode.

The gas lighter was triggered a number of times at multiple locations near the core of the MWA during a ten-minute period. Dual polarisation data from coarse channels 96–119 (122.88–153.6 MHz) were collected from all 128 tiles in the MWA’s compact configuration using the MWA VCS and processed through the inverse PFB. This produced a data set with which to validate the high-time-resolution mode.

An example of a measured gas-lighter signal is shown in Figure 5. In order to identify the signal by eye, the unprocessed VCS (fine channel) data were integrated over all tiles and polarisations in 1 ms intervals. This showed a typical total signal duration of a few ms. Following the inverse PFB, the signal is prominent in a single tile’s time series, and it is possible to discern multiple sub-pulses with microsecond structure.
An automated pulse detection algorithm was implemented to identify gas lighter signals that are coincident in time between multiple tiles. Initially, a simple $5.5 \sigma$ threshold search is used on each tile/polarisation independently. The time of arrival of each signal detected at each tile is compared to the signals from all other tiles, and signals that are coincident in time are grouped. The coincidence condition requires that the signals exceed the $5.5 \sigma$ threshold within the light travel time between tiles, with an additional $17 \mu s$ margin (corresponding to the light travel time across the array), after adjusting for system delays. Groups with four or more coincident tiles are noted for further analysis.

A key difference between cosmic-ray and gas-lighter signals is the total signal duration. The time at which a signal with an intrinsically long duration — such as that from the gas lighter — passes above threshold can vary according to signal shape and the influence of random noise. Any signals noted in the coincidence test are therefore correlated to extract a more robust time delay between the signals, resulting in a time resolution comparable to the sampling time. In the case of cosmic-ray signals, which have a duration of a few samples at most, any above-threshold signals will always have an accurate time-of-arrival.

It is possible to calculate the origin of a nearfield signal via an analytic approach provided that at least 4 different tiles detect the signal — three tiles yield two degenerate solutions on the horizontal plane. Since individual tiles can yield large timing errors — the structure of the gas-lighter signal is such that an incorrect peak in the correlation function is sometimes used to calculate time delays — candidate source positions for all combinations of four tiles are calculated. The median location of all of these calculated points then gives a stable solution to the source location.

Figure 6 shows the final source locations for multiple pulses that occurred within the same second, i.e. with the same source location. The scatter of the location solutions is about $5m$, consistent with the expected uncertainty due to the resolution of the time series data, and the known location of the gas lighter at the time.
A cosmic ray detection Mode for the Murchison Widefield Array

4. High Time Resolution RFI at the MRO

One of the major challenges encountered in identifying cosmic ray signals is eliminating impulsive RFI events. The nanosecond-scale RFI environment at the MRO to which this observation mode is sensitive is largely unknown, since it would be averaged out during normal MWA operating modes and would only contribute to a slightly higher noise level. We therefore present first results for low-frequency RFI detected at the MRO, during a one-hour observation.

We first establish the fidelity of our data by analysing the reconstructed voltage amplitudes. Figure 7 shows two examples of the sample amplitude distributions at each tile. These have been split in X and Y polarisation (orientated East-West and North-South, respectively) and have been fit with a Gaussian curve. The distribution on most tiles/polarisations (e.g. Tile 51) appears broadly Gaussian over the full time range analysed. However, all tiles begin to show an excess of samples above about 60–90 digital sample units. RFI events induce a long tail in the distribution, as is evident from the data for Tile 52.

To study these RFI events in more detail, we first identified all $6\sigma$ events, i.e. sample excursions with magnitude greater than six times the root-mean-square (RMS) sample value in each tile/polarisation. We then grouped all samples within a coincidence window of 17 $\mu$s, corresponding approximately to the light travel time across the MWA. Each event was characterised primarily by the number of tiles, $N_{\text{coinc}}$, with excursions within the coincidence window.

Figure 8 shows the distribution of $N_{\text{coinc}}$ as well as the predicted distribution from Gaussian noise at a $6\sigma$ threshold. As expected from Figure 7, the excess of high sample values leads to $N_{\text{coinc}}$ significantly exceeding predictions. There is an inflection point near $N_{\text{coinc}}=4$–5, indicating that detections with five or more coincident tiles are likely to be dominated by signals that cannot be attributed to thermal noise. These signals will either be RFI or radio emission from cosmic ray air showers. Following an extensive search, the majority of the signals that contribute to the excess of $6\sigma$ detections in the $N_{\text{coinc}}=1$–3 range appear to be random and uncorrelated, rather than due to locally generated RFI.

The total number of RFI events will be less than the total number of events with $N_{\text{coinc}} \geq 4$ — 363 events in this sample, i.e. one event every 10 seconds. Due to the prevalence of RFI, most other experiments that search for cosmic ray events use a particle detector array to act as a trigger (Schellart et al., 2013;
Fig. 7. A example of the sample distribution of two core tiles for both polarisation (E-W and N-S) measurements for 442 seconds. The dashed lines represent Gaussian fits to the individual datasets. This figure demonstrates the amount of RFI (i.e. deviations from the fit) detected by the array.

Fig. 8. A histogram of the number of coincident tiles (window of 17.7 µs) per detected signal for an hour observation. The dashed line is the expected rate of coincidence from thermal noise.

Aab et al., 2016a). Those without such a trigger typically suffer from very high rates of coincident RFI. For example, Monroe et al. (2020) used radio data only to identify 10 cosmic ray events in a 40 hr window. However, the raw rate of candidate events was 500 Hz, with the resulting RFI rejection cuts leading to a loss of 55% of the total observation time. In a related experiment to use the Parkes radio telescope to look
for nanosecond radio pulses produced by ultra-high-energy particles hitting the Moon, Bray et al. (2015) recorded a raw rate of 250 Hz RFI events, which was reduced to 1.6 Hz using an anti-coincidence filter. That the MWA sees a much lower rate of RFI events will make the identification of cosmic ray events significantly easier, and is testament to the low-RFI environment of the MRO.

The cosmic ray event rate over the MWA core should be similar to that of LOFAR and OVRO-LWA, with a rate of approximately one per hour (Schellart et al., 2013; Monroe et al., 2020). A unique characteristic of cosmic ray signals is their localised ground pattern, with emission concentrated in a ring of a few hundred metres diameter, which gets projected onto the ground plane according to the cosmic ray arrival direction (Huege, 2016). Far-field RFI, on the other hand, should be detected across the array, and be readily distinguishable from cosmic ray events. Any locally generated RFI may appear as a localised transient signal originating from the ground, and should be distinguishable using our proven ability to localise RFI sources. The expected East-West polarisation of cosmic ray events is a further discriminant that can be used to identify these signals. We therefore plan to use these techniques to search for cosmic ray events in a much larger data sample in the near future.

5. Conclusion

We have successfully implemented a method to reconstruct Murchison Widefield Array data at inverse bandwidth (16.2 ns) resolution. The necessity of inverting two stages of polyphase filterbanks results in the average loss of 9.5% in the peak of an impulse. The method has been tested using a calibrator pulse, allowing the location of the calibrator to be reconstructed at the expected resolution of 5 m. This ultra-high time-resolution mode is necessary and sufficient for detecting cosmic rays, which produce impulse-like radio signals in the MWA frequency range. A preliminary analysis of the nanosecond-scale RFI environment using one hour’s worth of data at the MRO indicates a very low rate of impulsive RFI, of order 0.1 Hz, which should readily allow the characteristic ground pattern of cosmic ray events to be identified. This motivates a search for cosmic ray events with a much longer observation period, and also suggests that this experiment will be feasible with the Square Kilometre Array, the low-frequency component of which will also be deployed at the Murchison Radio-astronomy Observatory.

Acknowledgments

The authors would like to thank Nichole Barry, Aman Chokshi, David Emrich, Jack Line, Andrew McPhail, Adrian Sutinjo, and Andrew Williams for their assistance in performing the gas lighter experiment. This scientific work makes use of the Murchison Radio-astronomy Observatory, operated by CSIRO. We acknowledge the Wajarri Yamatji people as the traditional owners of the Observatory site. Support for the operation of the MWA is provided by the Australian Government (NCRIS), under a contract to Curtin University administered by Astronomy Australia Limited. We acknowledge the Pawsey Supercomputing Centre which is supported by the Western Australian and Australian Governments. This research was supported by the Australian Government through the Australian Research Council’s Discovery Projects funding scheme (project DP200102643). A. Williamson would like to thank Innovation Central Perth, a collaboration of Cisco, Curtin University, Woodside and CSIRO’s Data61, for their PhD scholarship. This research made use of Python libraries Matplotlib (Hunter, 2007), NumPy (Harris et al., 2020), SciPy (Jones et al., 2001–), and Astropy (Astropy Collaboration et al., 2013, 2018), and NASA’s Astrophysics Data System Bibliographic Services.

References

Aab, A., Abreu, P., Aglietta, M., Ahn, E. J., Al Samarai, I., Albuquerque, I. F. M., Allekotte, I., Allison, P., Almela, A., Alvarez Castillo, J. & et. al. [2016a] Phys. Rev. D 93, 122005, doi:10.1103/PhysRevD.93.122005.
Aab, A., Abreu, P., Aglietta, M., Ahn, E. J., Al Samarai, I., Albuquerque, I. F. M., Allekotte, I., Allison, P., Almela, A., Alvarez Castillo, J. & et. al. [2016b] Phys. Rev. Lett. 116, 241101, doi:10.1103/PhysRevLett.116.241101.
Apel, W. D., Arteaga-Velázquez, J. C., Bekk, K., Bertainia, M., Blümer, J., Bozdog, H., Brancus, I. M., Buchholz, P., Cantoni, E., Chiavassa, A. & et. al. [2012] Astroparticle Physics 36, 183, doi:10.1016/j.astropartphys.2012.05.023.
B., Wayth, R. B. & et. al. [2013] PASA 30, e007, doi:10.1017/pasa.2012.007.
Tremblay, S. E., Ord, S. M., Bhat, N. D. R., Tingay, S. J., Crosse, B., Pallot, D., Oronsaye, S. I., Bernardi, G., Bowman, J. D., Briggs, F. & et. al. [2015] PASA 32, e005, doi:10.1017/pasa.2015.6.
van Haarlem, M. P., Wise, M. W., Gunst, A. W., Heald, G., McKean, J. P., Hessels, J. W. T., de Bruyn, A. G., Nijboer, R., Swinbank, J., Fallows, R. & et. al. [2013] A&A 556, A2, doi:10.1051/0004-6361/201220873.
Wayth, R. B., Tingay, S. J., Trott, C. M., Emrich, D., Johnston-Hollitt, M., McKinley, B., Gaensler, B. M., Beardsley, A. P., Booler, T., Crosse, B. & et. al. [2018] PASA 35, 33, doi:10.1017/pasa.2018.37.