Fast Radio Bursts with Apertif

Oostrum, L.C.

Link to publication

Creative Commons License (see https://creativecommons.org/use-remix/cc-licenses):
Other

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 4

Technical and scientific commissioning of the Apertif Radio Transient System


To be published as part of

Abstract

Fast Radio Bursts (FRBs) must be powered by uniquely energetic emission mechanisms. This requirement has eliminated a number of possible source types, but several remain. Identifying the physical nature of FRB emitters arguably requires good localisation of more detections, and broadband studies enabled by real-time alerting. Increasing detection rates depends strongly on enlarging telescope Field of View (FoV). APERture Tile In Focus (Apertif) is a new phased array feed system that delivers a \( \sim 30 \)-fold increase of this kind for the Westerbork Synthesis Radio Telescope. We here present the Apertif Radio Transient System (ARTS), a supercomputing radio-telescope instrument for Apertif that performs real-time FRB detection and localisation on this interferometer. It reaches coherent-addition sensitivity over the entire FoV of the primary-dish beam. Using a high-performance GPU pipeline, AMBER, for the transient detection and a real-time processing pipeline, DARC, the system is capable of sending out triggers of interesting events to the outside world within seconds of the light arriving at the telescope receivers. Our commissioning results include the detection of known pulsars, and of terrestrial signals such as perytons. We also determine the system sensitivity, and any effects of radio frequency interference. Together these commissioning steps verified that the system performs as planned, and is ready for scientific operation at design sensitivity.
4.1 Introduction

Many transient and time variable phenomena are observable in the radio sky, including fading radio afterglows of supernovae and gamma-ray bursts on many-month timescales (Dubner & Giacani 2015; van der Horst et al. 2007), solar radio bursts lasting milliseconds to minutes (Barrow et al. 1984; Zucca et al. 2018), and radio pulses from Galactic pulsars on timescales of milliseconds (Lorimer & Kramer 2005). However, the most recent and tantalising addition to this phase space was the discovery of bright (∼Jy), short (∼ms) radio pulses of extragalactic origin by Lorimer et al. (2007), now known as Fast Radio Bursts (FRBs).

Since their discovery in 2007, the study of FRBs has become a well-developed field in its own right, with more than 100 sources now reported (see Petroff et al. 2019a, for a review). The origins of these radio bursts is not yet known, but theories invoking young, highly magnetised neutron stars, black holes, and the explosions involved in their birth have gained traction (for a theory review, see Platts et al. 2019). A subset of the published FRB sources has been observed to repeat, supporting stable or non-cataclysmic progenitor theories (Spitler et al. 2016; CHIME/FRB Collaboration et al. 2019c; Fonseca et al. 2020).

The first FRBs were serendipitously discovered in all-sky surveys for radio pulsars. The systems necessary to find FRBs require high time resolution (∼1ms) and at least modest fractional bandwidth (∆ν/ν ≳ 0.25). The Berkeley Parkes Swinburne Recorder (BPSR) employing single pulse search software on Graphics Processing Units (GPUs) at the Parkes radio telescope proved particularly effective in early FRB searches (Keith et al. 2010; Thornton et al. 2013; Champion et al. 2016).

Increasingly, dedicated search hardware and software have been developed with the primary goal of discovering (and in some cases precisely localising) large numbers of new FRB sources; many with the stated goal of discovering new FRB pulses in real time. The majority of these systems employ GPUs, Field-Programmable Gate Arrays (FPGAs) and other specialised hardware on dedicated compute clusters to perform the most computationally expensive tasks such as dedispersion (Caleb et al. 2016; CHIME/FRB Collaboration et al. 2018; Law et al. 2018).

Despite the rapid advance of the field in recent years, many outstanding challenges remain in understanding the FRB phenomenon. Despite the high all-sky rate of FRBs of several thousand per day (Lawrence et al. 2017), many early instruments (such as Parkes) reported a new event only every few months due to the limited Field of View (FoV) inherent to single dishes or single receivers. Some of the most basic properties of the FRB population remain unknown; the underlying distributions of the population in pulse duration, dispersion measure, distance, energy, and spectral structure have all been difficult to pin down. This is partly due to a small sample size, but also in part due to insufficient instrumental resolution in the case of the pulse duration, dispersion measure, and spectral structure distributions.

Recent research was able to preserve data which capture the Stokes parameters of new FRBs and interpret the polarisation properties of individual bursts for a growing fraction of the population. Many of these bursts have been seen to be highly linearly polarised (Michilli et al. 2019b).
and analysis of the linear polarisation in some cases has revealed the presence of a strong magnetic medium local to the FRB source (Michilli et al. 2018). Polarisation may be critically important for understanding the emission and local environment of FRBs. However, preserving polarisation information requires a survey to either keep full-polarisation data for all survey observations, to be analysed post-detection; or to search data in real-time while the full-polarisation data remains in a memory buffer, to be excised and saved for a burst. Each choice comes with its own set of challenges for data storage and real-time data processing power, respectively.

A subset of the FRB population has been observed to repeat, producing multiple bursts at the same (or very similar) dispersion measure, with some sources now observed to repeat over many years (Spitler et al. 2016; Chapter 5). Sustained follow up of an FRB source is needed to eventually detect repeats. However, at present only 20 repeating sources have been published and the overall fraction of repeating sources in the FRB population is unknown. Whether all FRBs eventually repeat is an open question, one that is being tackled currently both with observational efforts (Fonseca et al. 2020) and modelling (Gardenier et al. 2019).

In addition to the challenges of understanding the underlying population(s) of FRBs, their physical properties, and their progenitors, there are also technical challenges involved in their discovery at scale. With next generation telescopes such as the Square Kilometre Array (SKA) it will no longer be possible to preserve the raw survey data for offline searches (Macquart et al. 2015). Instead, new sources will need to be identified in real time to capture the telescope data for later analysis. New automated FRB search techniques and pipelines taking advantage of classification and machine learning tools are being developed to prepare for this future reality (Connor & van Leeuwen 2018).

To address all these challenges, new FRB search efforts are increasingly employing interferometers to survey the sky (Caleb et al. 2017; Bannister et al. 2017; Maan & van Leeuwen 2017; Law et al. 2018; CHIME/FRB Collaboration et al. 2018). Interferometers, coherently or incoherently combining signals from many smaller elements or dishes, have the advantage of a large instantaneous FoV. Recent technological advances have resulted in new receivers such as Phased Array Feeds (PAFs), which place many dipoles at the focus of each dish of a telescope array.

One of the largest challenges of interferometric radio astronomy has always been computation. Beamforming within the telescope FoV requires a powerful correlator to combine the signals from all elements in phase. This is more difficult still when combining the multi-element PAF systems to form beams on the sky. Forming coherent beams and searching the time stream of each for impulsive radio signals such as FRBs provides an added technical challenge.

Faster and more agile processing units available in recent years have made it possible to form more beams and search them quickly, in some cases in real time. These searches still require large compute clusters to deal with the massive amounts of data streaming from the telescope and distribute it over many processing nodes. Many FRB search efforts, including
the searches described here with the APERture Tile In Focus (Apertif), now house dedicated computing clusters on-site to search the data in real time for pulses (Sect. 4.3.3).

By combining different elements of the array and the feed, the larger FoV of an interferometric array can be sampled by many smaller beams, enabling a much more precise localisation of any new source. The raw localisation ability of an interferometer depends on the length of the longest baseline, but even more precise localisation is possible for brighter signals that appear in several beams (Sect. 6.2), or where the raw voltage streams from the telescope are still available for offline correlation and beamforming.

For FRBs, arcsecond or better localisation is needed to identify a host galaxy unambiguously (Eftekhari et al. 2018), and interferometers provide the only path to such an association. More than 10 FRBs have been precisely localised and traced back to their hosts using interferometric arrays such as ASKAP, DSA-10, the EVN, and the VLA (Chatterjee et al. 2017; Bannister et al. 2019; Ravi et al. 2019; Marcote et al. 2020).

Since it is unknown whether all FRBs repeat, and on what timescale an individual source will produce repeating pulses, the goal in most cases is to localise an FRB from the discovery pulse. Identifying an FRB in real time in the data stream can aid localisation efforts, particularly in the case where the data are preserved at lower resolution or the raw voltages are not stored. If an FRB is found in the incoming data, the higher resolution data can be preserved, also including full-Stokes data in some cases (e.g., Petroff et al. 2015a; Fonseca et al. 2020).

Real-time searches are also critical in searching for prompt emission from the FRB source in other wavelength regimes, but at other radio frequencies as well. Previous multi-wavelength searches for related emission following real-time FRB detections were unsuccessful. However, these early efforts triggered follow-up several hours after the initial FRB (Petroff et al. 2015a).

It is still unknown how broadband FRB pulses can be, and down to what radio frequencies they are detectable. FRB emission has been seen down to the bottom of the CHIME band at 400 MHz (CHIME/FRB Collaboration et al. 2019b), but previous low frequency searches with LOw Frequency ARray (LOFAR) and the Murchison Widefield Array (MWA) have been unsuccessful (Coenen et al. 2014; Karastergiou et al. 2015; Rowlinson et al. 2016). However, a triggered search at low frequencies from a detection at 1 GHz may yield interesting results. Such coincident searches require real-time classification and triggering.

An ideal observing setup to tackle all of these challenges at once would combine large FoV with high spatial resolution on the sky for a high rate of localised FRBs. To address as many of the population unknowns as possible, such a system should also be able to resolve FRBs in time and frequency and capture polarisation information. The Apertif Radio Transient System (ARTS) is designed with these considerations in mind; The system operates on an interferometer (the Westerbork Synthesis Radio Telescope) equipped with PAFs, called APERture Tile In Focus (Apertif), to provide a large instantaneous FoV but much more precise spatial resolution (see Sect. 4.3.2 for more details). ARTS processes a bandwidth of 300 MHz centred on 1370 MHz.

ARTS also addresses some of the challenges above by combining a range of innovations: first, delivering high time and frequency resolution over a frequency range of 1220 MHz
to 1520 MHz; second, the ability to capture full-Stokes polarisation data with a new FRB search pipeline and third, a machine learning classifier to better identify and trigger on FRB candidates (Sect. 4.5.5.3). The ARTS project also investigates the feasibility of detecting FRBs at low frequencies through a targeted triggering effort with the LOFAR telescope in the Netherlands (Sect. 4.3.1).

4.2 Science motivation for ARTS

4.2.1 Fast Radio Bursts

FRBs exhibit dispersion curves and frequency-dependent scattering that are characteristic for sources located far outside our Galaxy (Lorimer et al. 2007); and a number have been localised to host galaxies at redshifts ranging from 0.12 to 0.7 (Marcote et al. 2020; Ravi et al. 2019). Their bright and brief (∼10 Jy, ∼1 ms) radio emission, over the derived luminosity distances of several Gpc, implies an extraordinarily energetic and compact cosmological origin: The brightness temperature exceeds $10^{32}$ K (Petroff et al. 2019a), while the light travel distance over the duration of the pulse substructure limits the source size to ∼10 km (see e.g. Farah et al. 2018).

In many respects the radio pulses we observe from FRBs are very similar to those from radio pulsars. Both pulsars and FRBs emit in the same, broad frequency bands, spanning of 100s of MHz to 10s of GHz (comparing e.g. Camilo et al. 2006 with Gajjar et al. 2018). Pulses for both are around 1 ms in duration and show some intriguing, similar subpulse behaviour (Hessels et al. 2019). In stark contrast, the pseudo luminosity between the two is different by over ten orders of magnitude. Even though a significant number of the theories have been put forward (Platts et al. 2019), the origin of the bursts remains unexplained.

Determining the nature of FRBs and next using them as tools in, for example, cosmology requires progress on a number of observational fronts. Larger numbers of detected FRBs will allow for more significant statistical studies, using methods that are currently being first demonstrated (e.g. Gardenier et al. 2019), to indicate whether multiple populations of FRB-emitting sources exist. Detection of rare, bright bursts at high Signal-to-Noise ratio (S/N) will provide the details required to help study the radio emission characteristics. Reaching these goals requires a large number of detections – thus, a survey with good sensitivity, FoV, and time on sky.

These three characteristics are provided by ARTS. Given the estimated $10^3 - 10^4$ bursts per day above a fluence of 1 Jy ms (e.g. Thornton et al. 2013), many will occur in the Apertif 8 sq. deg. FoV. As Apertif is a full-time survey machine, ARTS should detect an FRB roughly every week of observing (see Maan & van Leeuwen 2017).

A factor that potentially is more important than rates, in determining the formation of FRBs, is the localisation of the bursts. If FRBs are formed by young neutron stars, the galaxy in which they reside will need to have recently been forming massive stars. To identify the host galaxy with high confidence, the FRB position error box must be small enough to hold only a single candidate host. A young neutron star emitting FRBs may still be surrounded by a nebula,
that could be detected in follow-up observations if the FRB is well enough localised. Theories in which Active Galactic Nuclei (AGN) are related to the formation of FRBs could be falsified if localisation regions never contain these; or continue to be possible if they do. Finally, for nearby bursts, sub-arcsec localisation, using Very Long Baseline Interferometry (VLBI), can connect the (repeating) FRB emitters with features within the host galaxy. ARTS contains the hardware to connect the Westerbork Synthesis Radio Telescope (WSRT) to VLBI; but more importantly, the addition of Apertif onto an interferometer with a baseline of over 1 km provides very good instantaneous FRB localisation.

### 4.2.2 Neutron Stars

The surface gravity of these extremely compact stars, about $10^9$ times the gravity on Earth, is the largest of any object visible in the Universe. The internal densities of ten times nuclear density have not existed elsewhere since the Universe was about 1 ms old.

The combination of this high density and the millisecond rotation periods turns neutron stars into near-perfect cosmic time keepers (e.g. Hulse & Taylor 1975). Performing high precision timing on individual binary stars, such as Double Neutron Star (DNS) systems, informs us of the underlying binary evolution (e.g. van Leeuwen et al. 2015), and enables tests of general relativity (Kramer et al. 2006b; Desvignes et al. 2019).

Statistical analysis of the ages and distributions of large numbers of radio pulsars can provide neutron-star birth rates (cf. Hartman et al. 1997). Three groups of neutron stars are only very sporadically active in radio: rotating radio transients (RRATs), intermittent pulsars and radio-transient magnetars. Given the odds against their detection, the number of such transient neutron stars must be comparable to that of radio pulsars (Keane et al. 2011). Either supernovae make more neutron stars than we thought, or these sub-populations evolve into one another.

In survey mode, ARTS performs full-FoV searches for neutron-star single pulses. In timing mode, it provides a high-time resolution data stream, with real time coherent-dedispersion, and online folding. Together these allow for both searches for, and studies of radio-emitting neutron stars.

### 4.2.3 Prompt emission from slow transients

Some nearby DNS systems can be studied through pulsar timing from which we can derive that their orbits shrink. That orbital decay is due to the emission of gravitational waves. DNS systems that were formed $\sim10^8-9$ years ago, have currently reached the point that the two stars will merge. The energy reservoir in this coalescence is so large that multiple stages could produce radio emission, on different timescales (see, e.g. Chu et al. 2016). Prompt emission could possibly be generated at the merger, and possibly at the slightly later collapse of an intermediary massive neutron star to a black hole. Incoherent radio emission is expected from the reverse shock and the afterglow and produces an image-domain, slow
radio transient (Hallinan et al. 2017). To enable studies of both kinds of emission, ARTS is capable of simultaneously observing in the time domain as well as the image domain.

4.3 System overview

4.3.1 High-level overview of the Apertif Radio Transient System

Time-domain observations with Apertif are defined by the science teams and scheduled by the observatory operators. A hierarchical series of beams are formed (see Sect. 4.3.2) and used for automated and manual scientific processing and analysis. Transient detections in ARTS produce triggers to allow for follow-up in close to real time. Finally, processed data products are stored in the Apertif Long-Term Archive (ALTA), where they are publicly available.

ARTS, the Apertif time-domain system, comprises the following major subsystems, each described in more detail in the subsections below. These work together as illustrated in Fig. 4.1.

The first subsystem is the hardware platform (Sect. 4.3.3) that consists of the Apertif PAFs and the dishes of the WSRT east-west interferometer. Front-end beamformers provide dish processing. Tied-array beamformers built using FPGAs on high-performance processing boards (UniBoard and UniBoard\(^2\)), are connected through fast networking to a GPU cluster.

Second is the firmware and software subsystem (Sect. 4.3.4) that controls and produces one or multiple Tied-Array Beams (TABs), or ‘pencil beams’. These can be in Nyquist sampled, complex-voltage format for pulsar timing (and VLBI). The data in these TABs can also be ‘detected’, that is, converted to the four Stokes parameters, allowing subsequent partial integration to reduce data rates. That way, many hundreds of beams can be streamed out.
The third subsystem comprises the ARTS pipelines (Sects. 4.4 and 4.5). These perform transient searching and pulsar timing. For pulsar timing, the central single TAB is coherently dedispersed and folded in real time, on a single multi-GPU node. For the transient search, all Stokes-I TABs are cleaned of Radio Frequency Interference (RFI), dedispersed over a number of trial Dispersion Measures (DMs), corrected for chromatic effects, and searched for transient events. Good candidates immediately trigger data dumps from a ring buffer of Stokes-IQUV data. A deep learning implementation further classifies all candidates.

The fourth and final subsystem provides the integration with facilities external to Apertif, through the real-time broadcast of events, directly, for low-frequency follow-up, to LOFAR, and generally to the outside world through public VOEvents and the archive. Information about high-confidence candidates is sent within seconds of detection to LOFAR. Its Transient Buffer Boards (TBBs) freeze at the time of the expected burst arrival, and data is read out, and imaged offline, for localisation of the burst.

4.3.2 Overview of hierarchical beamforming

One of the innovative aspects of ARTS is its use of hierarchical beamforming to allow for searches throughout the entire primary-beam FoV, at coherent-addition sensitivity. That is a challenge in many wide-field fast-transient instruments (for example the upcoming Square Kilometer Array; Backer 2000, and LOFAR; van Leeuwen & Stappers 2010). Here we give a short overview of these beamforming steps. More details are available in Appendix 4.A.

Figure 4.2 provides an overview of all constituents of the beamforming hierarchy. The base is formed by the 40 Compound Beams (CBs) formed by the PAF in each Apertif dish. Each CB has a FoV diameter of ~ 0.5° and produces a full-bandwidth, Nyquist sampled output data stream. Together, the CBs cover the compound FoV of the Apertif system as shown in panel a of Fig. 4.2.
The data streams of corresponding CBs of multiple dishes are combined using either an incoherent beamformer or a coherent beamformer. Using the incoherent beamformer, 40 Incoherent-Array Beams (IABs) are formed that together cover the compound FoV of Apertif with a sensitivity improvement over a single dish that scales with the square root of the number of dishes used. When using coherent beamforming, the sensitivity scales linearly with the number of dishes involved, but more data streams need to be analysed as up to 12 TABs are formed for each CB. Together, these TABs fill the FoV of their CB completely owing to the grating response of the regularly-spaced WSRT array. This high TAB filling factor is illustrated in panel b of Fig. 4.2. This requires that we only use those dishes that are spaced equidistantly. Generally these are the 8 dishes RT2–RT9. The movable dishes RTA and B can be included if they are located on the common baseline grid.

Within a TAB, the WSRT grating lobes consist of multiple strong sidelobes. We exploit this fact to increase the sky coverage of each TAB. It does, however, also imply that there is degeneracy on the location of a detected event, as we do not know through which grating sidelobe the signal is coming in. We disambiguate between grating responses through their frequency dependence. Any grating response that is not in the phase centre has a frequency-dependent position. This is illustrated in panel c of Fig. 4.2. As shown in that panel, a source slightly further from the phase centre (measured parallel to the WSRT array) that is slightly further than a specific grating response at 1500 MHz may be exactly in that grating response at 1350 MHz and slightly closer to the phase centre than that grating response at 1200 MHz. This also implies that most grating responses only make a detection over a limited frequency range. For each point in the FoV of the CB and for each frequency band, we should thus select the TAB whose grating response is best positioned. To make a detection over the full bandwidth, that is, at full sensitivity, different TABs are selected over different frequency ranges and their responses combined to form a Synthesised Beam (SB) as illustrated in panel d of Fig. 4.2. In ARTS, up to 71 SBs are formed in each CB. As it is integrated into the subband dedispersion (Sect. 4.4), this frequency re-organisation itself comes with no additional computational cost. However, the increased number of beams (12 TABs to 71 SBs) does require more compute power down stream. The SBs provide instantaneous localisation with a resolution in one direction determined by the extent of the array of dishes used for the observation and a resolution in the orthogonal direction determined by the size of the CB.

The orientation of the grating responses on the sky changes with time due to Earth rotation. Over time, a specific locus in the CB therefore moves through different SBs with the phase centre as notable exception. This is illustrated in panel e of Fig. 4.2. To track a specific locus in the FoV of each CB, we should thus pick the best positioned SB for each instant in time during the observation and combine the corresponding data. This results in the last constituent of our hierarchy of beams: the Tracking Beam (TB). Over a 12-hour observation, ~3000 unique loci (and hence TBs), are defined within each CB.

In this section, we have introduced the hierarchy of beams in a conceptual and qualitative way. A more quantitative description is given in Appendix 4.A.
4.3.3 Hardware

4.3.3.1 Uniboard/Uniboard\(^2\) FPGA system

At the dishes, 9.3 Tb/s is digitised (Eq. 1.11); these are processed by a rack of eight Uniboards (UNBs). Each UNB contains eight FPGAs, divided into two sections of four nodes: the Front Node FPGAs (FNs) and Back Node FPGAs (BNs) on the UniBoards. The BNs receive the data from the Analogue-to-Digital Converters (ADCs) and run the filterbank. The data is then transported to the FNs, which run the beamforming and send out the data to a central building. X and Y polarisation are processed independently by four UNBs each. Each of the 16 nodes that is thus available for the filterbank and beamformer per polarisation, processes \(1/16\) of the subbands. In total, the twelve Apertif-equipped dishes produce \(\sim 3.5\) Tb/s of compound-beam data.

The central ARTS beamformer consists of 16 UNBs, for a total of 128 FPGAs. Each central UNB processes \(1/16\) of the bandwidth of all dishes. The FNs receive the X-polarisation data from half of the dish FNs, while the BNs receive the Y-polarisation data from the other half of the dish FNs. Each node receives the data from three dishes. For the dedicated FRB survey, the data are reordered such that each FPGA processes the data from five CBs. Each of the 128 FPGAs thus outputs \(1/16\) of the bandwidth for five CBs.

A set of four Uniboards\(^2\) (UNB2s) is installed to be able to run imaging and time-domain at the same time. The UNB2s receive the data from the 16 UNBs. When the dedicated FRB survey is running, the UNB2s simply forward the data to network switches. For the commensal setup, they receive intermediate-stage output from the UNBs, which are then operating as a correlator. Each UNB2 contains four Processing Nodes (PNs). A single UNB2 PN processes the full bandwidth for five CBs.

During dedicated FRB survey observations, the Uniboards generate 480 TABs with a time resolution of 81.92 \(\mu\)s and 1536 frequency channels over a bandwidth of 300 MHz. For each TAB, both a Stokes I and a Stokes-IQUV data stream are created, for a total output data rate of \(\sim 360\) Gb/s.

4.3.3.2 ARTS GPU cluster

The high data volume produced by the ARTS beamformer is processed by a 41-node GPU cluster. Each node has identical hardware, which is listed in Table 4.1. One node serves as master/login node; it does not receive any data from the beamformer. The other 40 nodes provide a total of 160 GPUs, 1600 CPU cores, 5 TB RAM, and \(\sim 1.3\) PB storage, for a theoretical 32-bit peak performance of \(\sim 2\) PFLOP/s (Peta FLoating-point OPerations per second).

Each of the 40 worker nodes processes the data from one CB. The incoming data rates per CB are 1.8 Gbps of Stokes-I data, and 7.3 Gbps of Stokes-IQUV data. Half of the storage is available for the incoming Stokes-I data, which, when writing continuously, fills up in \(\sim 20\) hrs. Hence any processing that requires access to these data should take at most a few
4.3 System overview

Table 4.1: Hardware overview of a GPU cluster node. The cluster consists of 41 nodes with identical hardware.

<table>
<thead>
<tr>
<th>Hardware</th>
<th>Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>2 × Intel Xeon E5-2640 v4 (40 cores total)</td>
</tr>
<tr>
<td>GPU</td>
<td>4 × Nvidia 1080 Ti</td>
</tr>
<tr>
<td>RAM</td>
<td>128 GB</td>
</tr>
<tr>
<td>Network</td>
<td>40 Gbps, full-duplex</td>
</tr>
<tr>
<td>Storage</td>
<td>32 TB</td>
</tr>
</tbody>
</table>

hours to run. Stokes-IQUV data is only stored upon a trigger from the transient pipeline, so even though the Stokes-IQUV data rate is four times that of Stokes I, its storage will fill up very slowly. Both Stokes I and IQUV data are buffered in RAM, for 10 and 15 s, respectively, and made available to the data writers and pipelines.

4.3.4 Apertif/ARTS monitoring and control system

The WSRT is controlled by the Apertif Monitoring And Control system (MAC). The MAC is split into several parts, called controllers, that each control a specific part of the system. Observations are scheduled through the Apertif Task DataBase (ATDB). When the system has to start an observation, ATDB generates a parset file, and sends it to all relevant controllers. The controllers set up the PAFs, dish pointing, and back-ends. Most of these controllers are shared between imaging and time-domain. Two controllers were added to be able to control the ARTS GPU cluster: ARTSSURVEYCONTROL and ARTSSURVEYNODECONTROL. For an overview, see Fig. 4.3.

ARTSSURVEYCONTROL runs on the master node of the GPU cluster. This controller receives the parset when an observation needs to be started. The beamformer produces data with timestamps in increments of 1.024 s. The current timestamp is read from the FPGAs, and the start time and end time of the observation are aligned accordingly. The parset contains a list of CBs that need to be recorded. For each of these CBs, ARTSSURVEYCONTROL creates a header file to be used by PSRDADA\(^1\) memory buffers and a settings file. The settings file contains all necessary output paths and settings such as the S/N threshold used in the transient search, the central frequency, observing mode, and pointing coordinates of the CB. Each worker node runs an instance of ARTSSURVEYNODECONTROL and receives the header and settings file from ARTSSURVEYCONTROL.

ARTSSURVEYNODECONTROL is responsible for starting up all software required to run an observation. Central to all processing software are the data memory buffers. Based on experience in ARTS predecessor PuMa-II (Karuppusamy et al. 2008), we use tools from PSRDADA to control these buffers. In normal operations, four buffers are created: two buffers for Stokes-I data and two buffers for Stokes-IQUV data. The two buffers of a single Stokes mode are connected with DADA_DBEVENT. One of the buffers is the main buffer, which holds all incoming data. The other is the triggered buffer, which is empty until an external trigger is

\(^1\) http://psrdada.sourceforge.net/
Figure 4.3: Overview of the ARTS software. The Apertif MAC monitors the ARTS GPU cluster and schedules observations. The ARTS MAC consists of ARTSSurveyControl and ARTSSurveyNodeControl. Observations are started through ARTSSurveyControl, which sets up the observation on the GPU cluster. ARTSSurveyNodeControl starts the memory buffers for the incoming data, disk writers, and the FRB search pipeline AMBER. Low-resolution FITS data are stored in the Apertif Long-Term Archive (ALTA). Additionally, the DARC processing pipeline runs during the observations. It performs real-time analysis of candidate metadata and triggers Stokes-IQUV data dumps and/or LOFAR TBB observations. Additionally, it performs an offline in-depth analysis of FRB candidates, which are then sent to the astronomers for visual verification.
4.4 Apertif Monitor for Bursts Encountered in Real time

The Apertif Monitor for Bursts Encountered in Real Time (AMBER; Sclocco et al. 2016) is the real-time software pipeline for single pulse detection used within ARTS. The goal of AMBER is to produce a list of dedispersed signals with high peak signal-to-noise-ratios (i.e. FRB candidates) from the beamformed TAB or IAB data stream described in Sect. 4.3.2. AMBER is optimised for execution on highly parallel architectures, such as the GPUs installed in the ARTS cluster.

One of the key points in the design of AMBER is modularity. Each component of the pipeline is a separate module, and is developed and maintained on its own. This design allows developers to work on new modules, or make changes to existing ones, with minimal impact on the AMBER code base. Moreover, modules can be reused in other projects such as SKA without requiring a hard dependency on AMBER.

The standard pipeline is composed of seven processing stages sequentially applied to data chunks: (1) RFI mitigation, (2) downsampling, (3) dedispersion, (4) integration, (5) S/N evaluation, (6) candidate selection, and (7) clustering. Among these seven stages, four are optional and can be enabled by the user when starting AMBER, while the others are always executed. We briefly describe each stage below.

**RFI mitigation** (optional). Two distinct filters are applied to the input data to identify and remove bright wide-band low-DM RFI and bright narrowband RFI, respectively; this module is described in more detail in Sclocco et al. (2020).

---

 Downsampling (optional). Here we reduce the time resolution of the input data. Downsampling data is particularly useful when searching for transients at high DMs, as high DM steps require larger chunks of data in time, which can be prohibitive due to memory limitations of the hardware. Additionally, a very high time resolution is not required at high DMs due to intra-channel DM smearing.

Dedispersion. This module constitutes the base for the searching algorithm. It implements dedispersion using an algorithm derived from brute force approach. Recognising that, in GPU dedispersion, optimizing memory limitations is much more fruitful than computing bounds, the module is highly tuned for data reuse. It can dedisperse both in a single step, or using a more efficient two step process including subbanding. During subbanding, the module creates the Synthesised Beams (SBs) described in Sect. 4.3.2. An overview of the module design and performance is available in Sclocco et al. (2016).

Integration (optional). FRBs pulses span a range of time durations. The S/N is highest if the time series are down sampled such that all emission is collected in a single time bin. We thus search over a range of down sampling factors. During integration, the dedispersed time series are downsampled in the time dimension according to this user-defined discrete set of trial pulse widths. This integration thus acts as a convolution kernel that smooths the signal and approaches the maximum intrinsic S/N of transients of various pulse widths.

S/N evaluation. This module provides a way to compute the S/N of all peaks in the dedispersed, and optionally integrated, timeseries. Different ways of computing the S/N of a time series are implemented in this module, and users can select the method that best fits their data.

Candidate selection. Candidates with S/N higher than a user-defined threshold are stored and made ready to be included in the output.

Clustering (optional). At this stage, selected candidates with similar DM or pulse width are clustered together, thereby reducing the total number of candidates reported in the output. Each cluster is represented by the candidate with the highest S/N.

Eventually, all candidates found in the current input chunk are stored in a text file. After the output is saved, AMBER continues processing the next chunk of input data.

AMBER is distributed under version 2.0 of the Apache License, and the source code is available on GitHub\(^1\). Source code portability, i.e. the ability to compile and run AMBER on different hardware platforms, is provided by using standardised and open languages such as C++ and OpenCL. However, AMBER is not just portable at the source code level: As a result of the combination of run-time code generation, user configurability, and auto-tuning, AMBER can also provide performance portability, and can be automatically adapted to achieve high performance on different hardware platforms (see e.g. Mikhailov & Sclocco 2018), and for different observational parameters and search strategies.

\(^1\) https://github.com/AA-ALERT/AMBER
4.5 Data Analysis of Real-time Candidates

The output of AMBER consists of only the metadata of FRB candidates. Based on these metadata, the system has to decide whether or not to store the Stokes-IQUV data buffer to disk, and/or trigger a LOFAR observation. In order to do this, we have designed a pipeline that processes the AMBER FRB candidates in real-time: Data Analysis of Real-time Candidates (DARC). Additionally, DARC provides offline processing of the FRB candidates. This involves reading back the raw data for good candidates and classifying them with a deep neural network. Finally, plots of the FRB candidates are sent to the ARTS team by e-mail for visual inspection. DARC is distributed under the GPL-3.0 license and available on GitHub\(^1\).

Speed is of importance for both the online and offline modes. The Stokes-IQUV data buffer is 15 s in size, and thus a dump to disk has to be triggered well within that time. Typically, a 7-s data block is written for each trigger, so the pipeline should send the trigger within 8 s of the FRB arrival time. LOFAR should be triggered within 10 s of the FRB arrival time, so the Stokes-IQUV trigger provides the most stringent constraint on the online pipeline performance.

The offline processing is allowed to be much slower than the real-time system, but still has the significant constraint that the raw data can only be stored for \(\sim\)15 hrs. The astronomer responsible for inspecting the e-mailed candidates has to delete the raw data, and as ARTS is designed to run continuously, the processing should not take more than a few hours at most.

In the following subsections, we first describe the global design of the pipeline, followed by a detailed description of each DARC module.

4.5.1 Design constraints

The design of DARC is constrained by the layout of ARTS. As each GPU cluster node stores the data products locally, DARC has to run on all nodes as well.

Offline data processing should read the AMBER output, and verify whether or not each candidate is in fact a real transient using the filterbank data. The results of all 40 CBs have to be combined and sent to the astronomers. This can be achieved by storing the results on shared storage. The master node of the cluster can then read the results and generate an overview of the observation.

For the real-time system, DARC has to read the AMBER output and send any Stokes-IQUV trigger to the \texttt{psrdada} event listener that can trigger the Stokes-IQUV data dump. The Stokes-IQUV triggering runs on each node independently so does not require cross-node communication. High-significance events should trigger the LOFAR TBBs. Communications to LOFAR happen through a VOEvent broker that is external to ARTS. As an ARTS FRB may be detected in several CBs simultaneously, the possible LOFAR triggers of multiple CBs should be sent to the master node, which then decides which trigger to send to the VOEvent broker.

\(^{1}\) https://www.github.com/loostrum/darc
Table 4.2: Overview of DARC modules and their tasks, see also Fig. 4.3.

<table>
<thead>
<tr>
<th>Type</th>
<th>Module</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monitoring and control</td>
<td>DARCMaster</td>
<td>Manages all other modules</td>
</tr>
<tr>
<td></td>
<td>StatusWebsite</td>
<td>Generates a webpage with the status of each module</td>
</tr>
<tr>
<td>Offline</td>
<td>OfflineProcessing</td>
<td>Full offline processing pipeline</td>
</tr>
<tr>
<td>Real-time</td>
<td>AMBERListener</td>
<td>Reads AMBER candidates from disk</td>
</tr>
<tr>
<td></td>
<td>AMBERClustering</td>
<td>Determines for which candidates to trigger a Stokes-IQUV dump and/or</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LOFAR TBB observation</td>
</tr>
<tr>
<td></td>
<td>DADATrigger</td>
<td>Executes Stokes-IQUV triggers through psrdada</td>
</tr>
<tr>
<td></td>
<td>VOEventGenerator</td>
<td>Executes LOFAR triggers through VOEvent system</td>
</tr>
</tbody>
</table>

4.5.2 Global design

To ease development and improve readability of the source code, DARC is split into several modules that each perform a specific task. The modules are listed in Table 4.2. They can be split in roughly three categories: monitoring and control, offline processing, and real-time processing. An overview of the DARC modules and how DARC connects to the MAC is given in Fig. 4.3.

Communication between the modules is provided by Python queues. Queues follow the first-in; first-out principle. A module can put any type of Python data on a queue, which is then read in the same order by another module. This setup allows for passing information (for example AMBER candidates) from one module to the next. It is also possible to use queues in a server-client fashion. This is used for LOFAR triggers: The server queue runs on the cluster master node and each worker node is a client that puts LOFAR triggers on the server queue.

4.5.3 Monitoring and control

4.5.3.1 Master module and command line interface

All other DARC modules are controlled from the master module, DARCMaster. When DARC is started, DARCMaster sets up all relevant modules and connects them with queue objects. DARCMaster listens for commands on a network socket, allowing for communication between nodes of the GPU cluster. Commands can be sent to DARCMaster through the darc executable. The executable can be used to start and stop observations, as well as to control the modules and change global settings.
When the MAC issues a `start_observation` command, it starts the data writers and AMBER, but also provides DARC with the path to a parset file containing the observation settings. DARCMASTER parses this file and instructs each module to start an observation with the specified settings through their respective queues.

### 4.5.3.2 Status website

The `StatusWebsite` module on the master node monitors the status of DARC across the GPU cluster. At periodic intervals, it queries the status of each DARC module on each node. This is then converted into an HTML webpage showing offline module in red, and online modules in green. The data are sorted by node, so an offline node is easily spotted.

### 4.5.4 Real-time system

#### 4.5.4.1 Reading FRB candidate metadata

AMBER writes a list of FRB candidates to disk every second. The real-time AMBERListener module reads the AMBER output files and makes them available to the rest of DARC. As there are three instances of AMBER running for every observation, there are three files to read candidates from.

At the start of an observation, AMBERListener starts three threads that each wait for one of AMBER's output files to appear and read new lines from the file every second. If a candidate file does not exist yet, AMBERListener retries reading it until a configurable timeout has passed.

AMBERListener is purposely agnostic to the exact content of each line of the candidate files. Instead of parsing the lines, they are put on the output queue as a single string per line. This ensures that when, for example, the number of columns in the candidate file is ever changed, the AMBERListener module will not need to be updated.

#### 4.5.4.2 Candidate clustering

Each candidate produced by AMBER has an associated beam number, downsampling factor, arrival time, width, DM, and S/N. A single transient may be detected in multiple points in this 6-dimensional parameter space. AMBER already clusters candidates that are detected at the same time in neighbouring DMs or downsampling steps by only saving the candidate with the highest S/N. However, a transient may still be detected at slightly different times or DMs and hence be written to the candidate file several times. The AMBERClustering module clusters the candidates further and produces triggers for a Stokes-IQUV data dump through DADATrigger, or a LOFAR trigger through VOEventGenerator.

Every second, all candidates received from AMBERListener are parsed by AMBERClustering and fed to a clustering algorithm. The clustering algorithm is simple: We take the highest S/N event in a small box in the DM/arrival time plane, across all SBs of the CB. The
clustering runs in two different modes simultaneously: one for known sources, and one for new sources. The modes use different thresholds as listed in Table 4.3.

For any remaining candidates, a Python dictionary with the arrival time, DM, beam number, width, and S/N is created. This is done separately for Stokes-IQUV triggers, which are sent to the DADATrigger module, and for LOFAR triggers, which are sent to the VOEventGenerator module that runs on the master node of the cluster.

### 4.5.4.3 Stokes-IQUV triggers

The DADATrigger module converts an incoming set of FRB parameters to a Stokes I or IQUV trigger. The Stokes-I trigger is not yet used in the current pipeline (March 2020). As it provides a way of accessing Stokes-I time-frequency data of candidates in real-time real time, it enables future efforts to move also the neural-net FRB classification to real time.

The input to DADATrigger is a list of dictionaries with FRB parameters, as provided by AMBERClustering. For each FRB candidate, the UTC start and end time of the Stokes-IQUV trigger need to be calculated. Because the arrival time as given by AMBER corresponds to the top of the frequency band, this time is also the latest possible start time of the Stokes-IQUV trigger. The length of the trigger is the dispersion delay across the band with a minimum of 2s, plus an extra 2 s to ensure there are data without an FRB to determine noise statistics from. The start and end time of the Stokes-IQUV trigger have to be aligned to the incoming data, which come in chunks of 1.024 s. The start time is therefore rounded down to the nearest multiple of 1.024 s, while the end time is rounded up in the same way. This ensures the full burst is captured in the Stokes-IQUV data.

The Stokes-IQUV trigger itself is a multi-line string, with one trigger per line. The triggers are processed in the MAC by dada_dbevent, a tool from psrdada. dada_dbevent receives the Stokes-IQUV trigger from DADATrigger through a network socket. It automatically combines overlapping triggers into one longer trigger, so in principle there is no dead time. Even the ARTS hard-disk pool, however, cannot continuously write Stokes-IQUV data because the data rate is too high. Stokes-IQUV triggers are therefore limited to one per minute.

### Table 4.3: Thresholds used during candidate clustering. LOFAR triggering is disabled for known pulsars.

<table>
<thead>
<tr>
<th>Trigger type</th>
<th>Source type</th>
<th>S/Nmin</th>
<th>DM range (pc cm$^{-3}$)</th>
<th>Downsampling$_{max}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stokes IQUV</td>
<td>Known pulsar</td>
<td>10</td>
<td>DM$_{src}$ ± 10</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Known FRB</td>
<td>10</td>
<td>DM$_{src}$ ± 10</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>New source</td>
<td>10</td>
<td>&gt; 1.2 × DM$_{YMW16}$</td>
<td>100</td>
</tr>
<tr>
<td>LOFAR</td>
<td>Known FRB</td>
<td>12</td>
<td>DM$_{src}$ ± 10</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>New source</td>
<td>12</td>
<td>&gt; 2.0 × DM$_{YMW16}$</td>
<td>80</td>
</tr>
</tbody>
</table>
4.5 Data Analysis of Real-time Candidates

4.5.4.4 LOFAR triggers

While the worker nodes decide when LOFAR should be triggered, the VOEventGenerator module on the master node is responsible for executing these triggers. Once a potential LOFAR trigger is received from one of the nodes, the system waits for one second in case any additional triggers arrive. This could happen if an FRB is detected in multiple CBs. The highest S/N trigger is then selected to actually be sent to LOFAR.

For communications with LOFAR we use the VOEvent standard for FRBs (Petroff et al. 2017). An xml-format VOEvent is generated from the trigger parameters. This is then sent to the VOEvent broker that LOFAR listens to.

After sending a VOEvent, further triggering of LOFAR is disabled. Once the LOFAR data has been read out, triggering can be enabled again manually through the DARC command line interface.

4.5.5 Offline processing

The full offline processing pipeline is executed by OfflineProcessing. On worker nodes, this module produces an overview of FRB candidates of the observation. The master node waits for all workers to finish, then gathers the results and sends them to the astronomers. Several tools used for system verification and calibration are also run by this module. Below we describe each step in the offline processing pipeline in more detail.

4.5.5.1 Initialisation

When an observation is started, the OfflineProcessing module receives the observation settings. Processing should start only after the data recording has finished. The system thus idles until the end time of the observation as defined by the MAC. Then, OfflineProcessing creates the necessary output directories and starts the processing. Depending on the observation type and source name, it also starts calibration tools (see Sect. 4.5.5.5).

4.5.5.2 Candidate clustering and extraction

The offline candidate clustering uses the same method as the real-time system (Sect. 4.5.4.2). All candidates from the observation are fed to the clustering algorithm at the same time. Any candidates with a DM $< 20$ pc cm$^{-3}$ or S/N $< 10$ are discarded.

For each post-clustering candidate, the few seconds of data around it are extracted from the filterbank files and cleaned of RFI. One complication is that the data on disk are not SBs, but TABs. The SBs thus need to be regenerated from the TABs. DARC includes the SBGenerator tool which can be used to determine out of which TABs a specific SB is made, and to convert TAB data into SB data. Both reading the data from disk and generating the SBs are intensive tasks. Therefore, the data extraction is run using two threads. While one thread is waiting for Input/Output (I/O), the other can run the SB synthesis.
We then dedisperse to the DM found by AMBER and downsample in time to the pulse width that maximises the S/N of the frequency-averaged time stream. While the S/N is already reported by AMBER, its value may not be accurate due to AMBER having limited data (1.024 s) available to determine the noise level. This is especially important for wide pulses that would be detected at the highest downsampling factor of 250. One batch of data then consists of only 50 samples. The offline processing loads several seconds worth of data from disk, and thus does not have this limitation. Therefore, the initial AMBER S/N calculation is refined here. This is done using a matched filter that tries many different box-car widths, which is different from how it is calculated in AMBER. If the S/N is below 5, the candidate is treated as a spurious event and discarded. For all candidates with sufficiently high post-processing significance, the dynamic spectra, DM-time arrays, and metadata are saved to an HDF5 file that can be classified by our machine learning classifier.

4.5.5.3 Candidate classification

Due to the real-time nature of the ARTS pipeline and the large number of false-positives relative to true astrophysical transients, our candidate classification had to be automated. To this end, we built a binary classifier using deep neural networks (DNNs) to select true FRBs and discard false positives generated by RFI and noise fluctuations (Connor & van Leeuwen 2018). The publicly-available package is called single_pulse_ml and uses Keras with a TensorFlow backend for the construction, training, and execution of its convolutional neural networks (Connor 2018). Our machine learning classifier was trained on tens of thousands of false positive triggers from Apertif, as well as an equal number of ‘true positives’ that were generated either by injecting simulated FRBs into real telescope data or by detecting single pulses from Galactic pulsars.

The classifier assigns each candidate a probability of it being a real astrophysical transient. If the probability is above a set threshold (currently 50%), a diagnostic plot is generated showing the frequency-time intensity array (i.e. dynamic spectrum), DM-time intensity array, and the pulse profile, as well as metadata such as the beam number, S/N, classifier probability, and width. These plots, along with a summary file containing the number of candidates at different steps in the pipeline, are stored to shared storage so they can be picked up by the master node emailing system.

4.5.5.4 E-mailing system

Once each worker node finishes processing an observation, the results are loaded by an emailing system running on the master node. First, an overview of the observation settings is generated that includes the pointing on sky of each CB, the set of dishes used, and the maximum expected Galactic DM as determined with YMW16 (Yao et al. 2017). Secondly, a table is generated with a summary per CB of the number of raw, post-clustering, and post-classifier candidates. Finally, an overview of the FRB candidate metadata is generated,
sorted by the frequency-time classifier probability of being an astrophysical transient. This information is converted into an e-mail, and sent to the astronomers with the diagnostic plots of the candidates as attachments. Additionally, the same data is made available on a local website, where the team can go through all observations and check the candidates.

4.5.5.5 Calibration tools

As noted in Sect. 4.5.5.1, the offline processing can automatically run tools used for verification and calibration. Currently, two tools are used.

The first tool is the folding of test pulsar data. DARC recognises observations of several test pulsars by the source name defined in the observation settings. If the test pulsar is located in the CB a specific worker node is processing, the filterbank data of the central TAB of the CB are folded with prepfold from PRESTO\(^1\) (Ransom 2011).

The second tool is the calibration of drift scan data. We regularly perform drift scans of calibrator sources in order to determine the sensitivity of the system and to be able to do bandpass calibration. Typically, a drift scan is done over several CBs. This information is encoded in the source name. For example, 3C147drift2126 would be a drift scan of 3C147 through CBs 21 to 26. DARC extracts this information from the source name, and determines whether or not the CB that is processed by the worker node is part of the drift scan. If it is, an external script is called which analyses the drift scan data.

4.6 Commissioning results

4.6.1 Pulsars

Several pulsars have single pulses that are bright enough to be detectable with ARTS. Additionally, the filterbank data can be folded at known pulsar parameters to increase the S/N and detect the integrated pulse profile. As part of commissioning, we regularly observe the four pulsars listed in Table 4.4. Single pulses are detectable from each, although the DM of B0950+08 is so low that the pulses are below the DM thresholds of our real-time search pipeline; and are potentially clipped by the RFI mitigation algorithm that is optimised for FRBs (See Sect. 4.4).

The integrated profile of a 5-minute observation of each test pulsar is shown in Fig. 4.4. From their S/N, we can estimate the System-Equivalent Flux Density (SEFD) of the 8-dish system used for these observation using the radiometer equation (Dewey et al. 1985; Lorimer & Kramer 2005):

\[
S = \frac{S/N \text{ SEFD}}{\sqrt{N_{\text{pol}} BW T_{\text{obs}}}} \sqrt{\frac{W}{P - W}},
\]

where \(S\) is the peak flux density, \(N_{\text{pol}}\) is the number of polarisations (two, for Apertif), \(BW\) is the observing bandwidth of 300 MHz, \(W\) is the pulse width, and \(P\) is the pulse period.

\(^1\) https://www.github.com/scottransom/presto
Table 4.4: Overview of test pulsars used by ARTS. All four have detectable single pulses. Pulsar parameters were taken from the ATNF pulsar catalogue (Manchester et al. 2005). We also list the S/N as measured with ARTS in a 5-minute observation, and the derived SEFD.

<table>
<thead>
<tr>
<th>Name</th>
<th>DM (pc cm(^{-3}))</th>
<th>Flux density(^{(a)}) (Jy)</th>
<th>Period (ms)</th>
<th>Pulse width(^{(b)}) (ms)</th>
<th>S/N</th>
<th>SEFD (Jy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B0329+54</td>
<td>26.8</td>
<td>203</td>
<td>715</td>
<td>6.6</td>
<td>3600</td>
<td>250</td>
</tr>
<tr>
<td>B0531+21</td>
<td>56.8</td>
<td>14</td>
<td>33</td>
<td>3.0</td>
<td>24</td>
<td>800</td>
</tr>
<tr>
<td>B0950+08</td>
<td>3.0</td>
<td>100</td>
<td>253</td>
<td>8.9</td>
<td>1400</td>
<td>150</td>
</tr>
<tr>
<td>B1933+16</td>
<td>158.5</td>
<td>58</td>
<td>359</td>
<td>6.0</td>
<td>1200</td>
<td>130</td>
</tr>
</tbody>
</table>

\(^{(a)}\) At 1400 MHz.
\(^{(b)}\) At 50% of peak intensity.

The derived SEFD values are listed in Table 4.4. Due to scintillation, the brightness of a pulsar varies over time and the SEFD values should be taken as rough estimates. For PSR B0531+21 specifically, the Crab Nebula is not resolved out completely by the TAB, so its emission contributes to the noise level and decreases the S/N, thereby increasing the SEFD estimate. Of the four test pulsars, PSR B1933+16 is the most stable and hence expected to give the most accurate SEFD. We note that the SEFD values are valid only for the central CB; The sensitivity of outer CBs is generally expected to be slightly lower (Sect. 4.6.2).

Assuming an aperture efficiency of 70%, a system temperature of 70 K (Oosterloo et al. 2009; van Cappellen et al. 2020), and perfectly coherent beamforming, the theoretical SEFD of eight Apertif dishes is expected to be \(\sim 70\) Jy. The values derived here are a factor 2-4 higher, ignoring B0531+21 for the aforementioned reason. Fluctuations in pulsar brightness on that level are not unexpected, so the measured SEFDs are within expectations. A more accurate SEFD can be derived using scans of calibrator sources, as discussed in Sect. 4.6.2.

### 4.6.2 Sensitivity

The sensitivity of ARTS depends on how sensitive the individual CBs are, as well as on the accuracy of the calibration of the delay and phase offsets between the dishes. Because the system is calibrated weekly, we regularly perform drift scans of calibrator sources at the start and end of every observing week. For each calibrator, the flux density as function of frequency is known and can be described by a polynomial in log space (Perley & Butler 2017),

\[
\log_{10} \left( \frac{S}{Jy} \right) = S_0 + \sum_i a_i \log_{10} \left( \frac{\nu}{GHz} \right)^i,
\]

where \(S\) is the flux density at frequency \(\nu\), \(S_0\) is the flux density at 1 GHz, and \(a_i\) are the polynomial coefficients. An overview of the calibrators used, and their polynomial coefficients, is given in Table 4.5.
4.6 Commissioning results

Figure 4.4: Integrated profiles and dedispersed frequency-phase diagrams of each test pulsar that is regularly observed with ARTS. These were generated from 5-minute observations.

Table 4.5: Overview of flux density calibrators used by ARTS. The spectra are described by a polynomial in log space, of which the coefficients are listed here. Flux densities at 1 GHz are listed. Data from Perley & Butler (2017).

<table>
<thead>
<tr>
<th>Name</th>
<th>Flux density (Jy)</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>$a_4$</th>
<th>$a_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3C48</td>
<td>21.1</td>
<td>−0.76</td>
<td>−0.19</td>
<td>0.05</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3C147</td>
<td>28.3</td>
<td>−0.70</td>
<td>−0.20</td>
<td>0.06</td>
<td>−0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>3C286</td>
<td>17.7</td>
<td>−0.45</td>
<td>−0.18</td>
<td>0.04</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
For each TAB of each CB, we measure the power level at the peak of the TAB when it is pointed exactly at the calibrator (the ‘on’ value) and in a part of the observation where the TAB is off-source (the ‘off’ value). The equivalent flux density of the noise level, the SEFD, can then be determined as

\[
\frac{\text{on} - \text{off}}{\text{off}} = \frac{S_{\text{calibrator}}}{\text{SEFD}},
\]

where \(S_{\text{calibrator}}\) is the flux density of the calibrator source. Figure 4.5 shows the derived SEFD for all TABs of all CBs as determined from sets of drift scan observations on seven different days. The SEFD value is taken as the median value over parts of the frequency band that are not affected by strong RFI. While the sensitivity across TABs for a single CB is relatively constant, the CB sensitivity does vary significantly as function of both time and CB number. This is expected: Both the PAFs and phase offsets between dishes are calibrated weekly. The calibration is done using on-sky data of calibrator sources, and can thus be affected by RFI. Additionally, the accuracy of the calibration could depend on the location of the calibrator source in the sky.

We would expect the sensitivity of a CB to depend on its position within the PAF as well: A CB closer to the centre of the PAF is closer to the focus of the parabolic dish and therefore expected to be more sensitive. From Fig. 4.5 we see that the last CB of a row within the CB layout (Fig. 4.2) is typically less sensitive. To see whether these are truly outliers, we show the the median SEFD of each CB and the CB offset from the centre of the PAF in Fig. 4.6.
4.6 Commissioning results

The six points with an SEFD > 120 Jy are the six CBs at the end of each PAF row, which are indeed outliers. The central CB, CB 00, is on average also less sensitive than the surrounding CBs.

Ignoring the outliers, there is still a trend. Beyond ∼1° from the PAF centre, the SEFD increases to ∼30% worse than the median sensitivity for the outermost beams. For our survey, such a 30% sensitivity loss is more than balanced by the large increase in FoV the PAF provides.

The SEFD can be further defined as the ratio of the system temperature ($T_{\text{sys}}$) and gain ($G$),

$$\text{SEFD} = \frac{T_{\text{sys}}}{GN^\beta_{\text{dish}}}, \quad (4.4)$$

where the gain is that of a single dish, $N_{\text{dish}}$ is the number of dishes, and $\beta$ indicates the beamforming coherence. If there are no losses during beamforming, $\beta = 1$ in TAB mode, and $\beta = 1/2$ in IAB mode. For ARTS, the coherence was measured to be consistent with these theoretical values (Straal 2018). For these TAB drift scans, we thus set $\beta = 1$. Finally, the gain can be related to the aperture efficiency ($\eta$) of a dish as

$$G = \eta \frac{A}{2k_B}, \quad (4.5)$$

where $A$ is the illuminated surface area of the dish and $k_B$ is the Boltzmann constant.

The system temperature of Apertif is ∼70 K (Oosterloo et al. 2009; van Cappellen et al. 2020). At 1400 MHz, the sky background contributes minimally to the measured noise level. We thus use this $T_{\text{sys}}$ value to estimate the aperture efficiency using Eqs. 4.4 and 4.5. The
median sensitivity of 85 Jy translates into an aperture efficiency of $\sim 60\%$. While this aperture efficiency is slightly lower than the typically assumed value of 70\%, we note that the derived value also includes any losses due to beamforming and RFI. Therefore, we consider the value of 60\% to be consistent with theoretical expectations.

### 4.6.3 RFI environment

WSRT is located in a radio-quiet zone, making the RFI situation generally good. However, L-band is not entirely RFI-free. The generation of RFI is becoming an increasing part of human activities. Furthermore, our sky is populated by a growing number of satellites for world-wide telecommunications, which inevitably affect our data. Table 4.6 lists harmful interference previously measured at various frequencies at WSRT. ARTS generally observes in the 1220–1520 MHz band to minimise the impact from the strongest sources of RFI.

RFI is generally stronger than astrophysical signals. This is explained by the inverse-square law of propagation: Electromagnetic radiation dissipates at a quadratic rate with respect to distance. Hence, without a mitigation strategy, RFI can cause false-positive detections (i.e. non-astrophysical pulses erroneously classified as FRBs), while also masking real, weak astrophysical signals and reduce the rate of true positives.

False-positive detections have a direct effect on the search pipeline, as they can rapidly increase the size of the single-pulse candidates list, which in turn requires more processing, in addition to the need for visual verification by astronomers. To reduce the impact of RFI on our pipeline, two strategies are currently used on ARTS: RFI mitigation (Sclocco et al. 2020) that modifies outliers in the time series with local statistics, and a deep-learning classifier (Sect. 4.5.5.3).

#### 4.6.3.1 RFI Bandwidth loss

Two on-line RFI mitigation methods are implemented in AMBER (Sect. 4.4): (1) a ‘time domain mitigation’ method targeting bright low-DM broad-band signals, and (2) a ‘frequency domain mitigation’ method targeting spurious narrowband RFI. Each method is applied in an...
iterative manner, where each consecutive step applies mitigation to a cleaner set of samples than previous steps. Three iterations of each method are applied. To evaluate the impact of RFI on our total bandwidth, we executed these two off-line, on a 3-hour observation that occurred on 2019-08-19 at 05:50:00 UTC (Fig. 4.7). Within this representative sample, we find that at most about 7% of the total bandwidth is affected by RFI for limited time periods (< 0.5 hr). Figure 4.7 highlights that the majority of RFI sources in our band consist of spurious narrow-band emission.

4.6.3.2 RFI direction and time dependence

To assess the variability as function of sky direction and time, we gathered the AMBER results from 448 observations between March 2019 and early September 2019. During this period, RFI mitigation in AMBER was not implemented yet. Any trigger with DM=0 and S/N>10 is assumed to be RFI. Using the observation parameters, the altitude and azimuth of the dishes during each of these triggers was determined. These coordinates were then binned into a 10° × 10° grid. For each grid point, the number of triggers was scaled to the total time the telescopes spent pointing in that direction to get the average number of triggers per minute. A heatmap of the resulting RFI trigger rate as function of azimuth and altitude is shown in the top panel of Fig. 4.8. The RFI trigger rate is uniform over the sky, with one notable exception around an azimuth of 260° and altitude of 65°. The Smilde radio mast is located at this azimuth. Its UHF antenna transmits in the WSRT frequency band. Yet, at 303 m tall it only reaches an altitude of 1.3° as viewed along the direct line of sight from WSRT. Instead, it is likely that WSRT receives signals that were reflected off of the troposphere, which is a known effect of UHF emission. Under the assumption that the light is reflected once, halfway between Smilde and WSRT, and a typical troposphere height of 17 km, the expected altitude
Figure 4.8: Number of RFI triggers per minute (a) as function of telescope pointing and (b) as function of local time. The strong RFI near an azimuth of 260° and altitude of 65° is due to the Smilde radio mast. The RFI environment as function of local time shows an unexplained dip between 16:00 and 21:00 local time.

is 68°, which falls within the bin of strongest RFI in Fig. 4.8. We are thus confident that we are indeed seeing reflected signals from the Smilde mast.

The RFI trigger rate as function of local time is shown in the bottom panel of Fig. 4.8. The trigger rate was expected to be higher during opening hours of the nearby museum Kamp Westerbork, from 10:00 to 17:00 local time. RFI is indeed strong during these hours, but also during the night. The low level of RFI between 16:00 and 21:00 local time is currently unexplained.

4.6.4 Perytons

During the Apertif Science Verification Campaign (SVC) we tested the ARTS real-time transient detection system through the use of perytons. Perytons are a type of RFI first discovered at Parkes, generated when microwave-oven doors are opened while the cavity magnetron still operates (Petroff et al. 2015b). The frequency structure of the interference is very similar to the sweep seen for high-DM sources, which is quite unique for terrestrial RFI. We emitted eight perytons using two different ovens, from outside the WSRT control building. AMBER detected the perytons in real time at a DM of 395 pc cm$^{-3}$ (Fig. 4.9).

4.7 Summary

We have presented ARTS, the new time-domain backend of Apertif at WSRT. The design of ARTS allows for coherent, real-time searches for radio transients over the full 8 sq. deg. FoV provided by the Apertif PAFs. These capabilities are provided by a hybrid FPGA-GPU system. The GPU transient search pipeline AMBER and post-processing pipeline DARC provide real-time triggering, allowing the system to store Stokes-IQUV data for interesting candidates, and follow-up with the LOFAR TBBs.
Figure 4.9: A peryton detected in all 40 compound beams (the circles). In each, the respective time-frequency plot is displayed, dedispersed to 395 pc cm$^{-3}$. The vertical axis spans 300 MHz of bandwidth, the horizontal axis 500 ms of time. The red curve indicates the track a signal with DM = 0 pc cm$^{-3}$ would have followed.
We have commissioned the system using known pulsars and perytons. Drift scans of calibrator sources show that ARTS reaches design sensitivity over most of the FoV. The sensitivity loss due to RFI is limited. We did identify the Smilde radio mast as a significant source of RFI, but only covering a small patch of sky.

Overall we have shown that ARTS is capable of detecting fast radio transients at the designed sensitivity level. As of July 2019, regular survey operations have commenced.

**Acknowledgements**

We thank Joe Callingham, Helga Dénes, Vlad Kondratiev, Alexander Kutkin, Sarrvesh Sridhar, Emma Tiggelaar, and Matthijs van der Wiel for suggesting, providing and operating the microwave oven used in Sect. 4.6.4. This research was supported by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement No. 617199 ('ALERT'), and by Vici research programme 'ARGO' with project number 639.043.815, financed by the Dutch Research Council (NWO). Instrumentation development was supported by NWO (grant 614.061.613 'ARTS') and the Netherlands Research School for Astronomy ('NOVA-ARTS' and 'NOVA-NW3'). This work makes use of data from the Apertif system installed at the Westerbork Synthesis Radio Telescope owned by ASTRON. ASTRON, the Netherlands Institute for Radio Astronomy, is an institute of NWO.
4.A Overview of ARTS hierarchical beamforming

4.A.1 Tied-Array Beams

In this section, we provide a quantitative discussion on the formation of TABs. In following sections, we discuss formation of SBs and Tracking Beams (TBs) in more detail.

To illustrate how coverage of a CB by TABs works, we will assume that the separation between TABs, \( \theta_{\text{sep}} \), is equal to their half power beam width, which depends on the observing wavelength \( \lambda \) and the projected longest baseline \( B_{\text{max}} \) as

\[
\theta_{\text{sep}} = \alpha \frac{\lambda}{B_{\text{max}} \cos(\theta_{\text{proj}})}.
\] (4.6)

where \( \theta_{\text{proj}} \) is the angle between the plane perpendicular to the linear WSRT configuration and the line of sight. Empirically, we found for \( \theta_{\text{proj}} = 0 \) that for the 144-m spaced arrays of eight and ten dishes (Apertif-8, dishes RT2–RT9, \( B_{\text{max}} = 1008 \) m; and Apertif-10, comprising RT2–RTB, \( B_{\text{max}} = 1296 \) m), \( \alpha = 0.78 \) and \( \alpha = 0.80 \) respectively.

If the TABs cover an angular distance equal to the angular distance between the TAB main beam and its first grating response, the TABs will provide full coverage of the CB by a combination of their main beams and first grating response. The grating distance \( \theta_{\text{grat}} \) depends on the observing wavelength and the projected common quotient baseline \( B_{\text{cq}} \) of the regular array as

\[
\theta_{\text{grat}} = \arcsin \left( \frac{\lambda}{B_{\text{cq}} \cos(\theta_{\text{proj}})} \right) \approx \frac{\lambda}{B_{\text{cq}} \cos(\theta_{\text{proj}})},
\] (4.7)

where the approximation holds for small angles.

The number of TABs required per CB, \( N_{\text{TAB}} \), now follows from the ratio of the grating distance and the TAB separation, i.e.,

\[
N_{\text{TAB}} = \frac{\theta_{\text{grat}}}{\theta_{\text{sep}}} = \frac{B_{\text{max}}}{\alpha B_{\text{cq}}}. \tag{4.8}
\]

Note that the wavelength and the projection angle cancel each other out. This implies that the number of TABs required does neither depend on the projection angle nor on frequency. We therefore have to take into account the TAB separation and observing frequency to point the TABs, but we do not need to define a different number of TABs at different frequencies or different observing times. Using Eq. (4.8), we find that 12 and 9 TABs are needed for Apertif-10 and Apertif-8, respectively.

4.A.2 Synthesised Beams

As shown by Eq. (4.7), the grating lobe distance is frequency-dependent. This is illustrated in Fig. 4.10, which shows the grating response of the central TAB at 1500 MHz and the TAB left from the central beam at 1367 MHz for Apertif-10. The first grating response to the right of the centre of these two TABs coincides. This shows that a transient signal received in the first grating response of the central TAB to the right of the centre at 1500 MHz will not be detected in the corresponding grating response at 1367 MHz, but by the corresponding
Grating response of the first TAB left from the central TAB at 1367 MHz. For all but the main beam of the central TAB, we therefore may have to combine chunks of bandwidth from different TABs to form an SB at a given distance from the CB centre.

In principle, there is an optimal combination of chunks of bandwidth from different TABs for each position within the FoV. In practice, the number of SBs that we can form is limited. A practical minimal number of SBs can be set by considering the TAB separation at the highest operating frequency:

$$\theta_{SB} = \theta_{sep}(f_{max}) = \alpha \frac{\lambda_{min}}{B_{max}}.$$  \hspace{1cm} (4.9)

For Apertif-10, we found $\alpha = 0.80$ and we have $B_{max} = 1296$ m. Assuming $f_{max} = 1500$ MHz ($\lambda_{min} = 20$ cm) and a FoV of 30’, we find that the number of SBs required is 71.

At a given frequency, the optimal contribution to the SB denoted by index $n_{SB}$ is coming from a specific grating response denoted with index $n_{gr}$ from TAB with index $n_{TAB}$. In the remainder of this section, we determine which grating response from which TAB gives the optimal contribution to a given SB. Since the position shift of the grating responses over frequency increases with distance from the main beam, it is convenient to choose the reference position to be in the centre of the CB. This intuitively leads to an indexing scheme in which 0 denotes the central position, a negative index implies a position left or westward of the centre and a positive index denotes a position right or eastward of the centre. For example, $n_{gr} = -1$ denotes the first grating response left from the main beam, $n_{TAB} = 0$ denotes the central TAB and $n_{SB} = 5$ denotes the synthesised beam located five grid points to the right of the centre of the CB.
We can now assign a grating index and a TAB index at the highest frequency to each SB index using the following procedure:

1. Determine the closest grating response of the central TAB by

\[ n_{gr} = \left\lfloor \frac{n_{SB}}{N_{TAB}} \right\rfloor, \tag{4.10} \]

where \( \lfloor \cdot \rfloor \) denotes rounding. Many standard rounding routines round half integers away from zero. In our case, we assume that half integers will always be rounded upward towards the next higher integer value.

2. Determine the TAB index such that \( n_{TAB} = \{-N_{TAB}/2, -N_{TAB}/2 + 1, \ldots, N_{TAB}/2 - 2, N_{TAB}/2 - 1\} \) by calculating

\[ n_{TAB} = n_{SB} \mod N_{TAB} \tag{4.11} \]

and subtracting \( N_{TAB} \) if \( n_{TAB} > N_{TAB}/2 - 1 \).

Figure 4.11 shows the resulting grating indices and TAB indices for all SB indices when \( N_{TAB} = 12 \) and \( N_{SB} = 151 \). Note that the calculation above holds for even \( N_{TAB} \). A similar convention can be defined for odd \( N_{TAB} \). This is not spelled out here to keep this section concise.

The position of the SB with index \( n_{SB} \) is described by

\[ \theta_{n_{SB}} = n_{SB} \theta_{SB} = n_{gr} \theta_{grat}(f_{\max}) + n_{TAB} \theta_{sep}(f_{\max}). \tag{4.12} \]

Since \( \theta_{grat} \) and \( \theta_{sep} \) are frequency-dependent, at a certain (lower) frequency, the grating response of the next TAB will be at the same position and is therefore the optimal choice at that frequency. What constitutes the ‘next TAB’ depends on the sign of the SB index. If the
SB index is positive, the next TAB is the one with index $n_{\text{TAB}} - 1$, if the SB index is negative, the next TAB is the one with index $n_{\text{TAB}} + 1$. To find the frequency at which the position of the grating response of the next TAB at frequency $f_0$ coincides with the grating response of the original TAB at $f_{\text{max}}$, we need to solve $f_0$ from

$$n_{\text{gr}}\theta_{\text{grat}}(f_{\text{max}}) + n_{\text{TAB}}\theta_{\text{sep}}(f_{\text{max}}) = n_{\text{gr}}\theta_{\text{grat}}(f_0) + (n_{\text{TAB}} - \text{sgn}(n_{\text{SB}}))\theta_{\text{grat}}(f_0),$$

(4.13)

where sgn denotes the signum function. Substitution of Eq. (4.6) and Eq. (4.7) while replacing $\lambda$ by $c/f$ and taking $\theta_{\text{proj}} = 0$, we obtain

$$n_{\text{gr}}\frac{c}{f_{\text{max}}B_{\text{eq}}} + n_{\text{TAB}}\frac{\alpha_c}{f_{\text{max}}B_{\text{max}}} = n_{\text{gr}}\frac{c}{f_0B_{\text{eq}}} + (n_{\text{TAB}} - \text{sgn}(n_{\text{SB}}))\frac{\alpha_c}{f_0B_{\text{max}}},$$

(4.14)

which gives

$$f_0 = f_{\text{max}}\frac{n_{\text{gr}}B_{\text{max}} + (n_{\text{TAB}} - \text{sgn}(n_{\text{SB}}))\alpha_cB_{\text{eq}}}{n_{\text{gr}}B_{\text{max}} + n_{\text{TAB}}\alpha_cB_{\text{eq}}}.$$  

(4.15)

Note that a situation in which $n_{\text{TAB}} - \text{sgn}(n_{\text{SB}}) \not\in \{-N_{\text{TAB}}/2, -N_{\text{TAB}}/2 + 1, \ldots, N_{\text{TAB}}/2 - 2, N_{\text{TAB}}/2 - 1\}$ may occur. Since the position of $n_{\text{TAB}} + N_{\text{TAB}}/2$ for $n_{\text{gr}}$ coincides with $n_{\text{TAB}} - N_{\text{TAB}}/2$ for $n_{\text{gr}} + 1$, this can be solved by modifying the indices accordingly. A similar procedure can be followed for $n_{\text{TAB}} - N_{\text{TAB}}/2 - 1$.

From the discussion above, it is clear that for the $n_{\text{SB}}$th SB, for which the chunk of bandwidth close to $f_{\text{max}}$ is coming from the $n_{\text{gr}}$th grating response of the $n_{\text{TAB}}$th TAB, the optimal chunk of bandwidth around $f_0$ is provided by the $n_{\text{gr}}$th grating of the $(n_{\text{TAB}} - \text{sgn}(n_{\text{SB}}))$th TAB. Since the Half-Power Beam Width (HPBW) changes only slowly with frequency and the shift of the grating responses scales linearly with frequency, the optimal switching frequency will be approximately halfway between $f_{\text{max}}$ and $f_0$ if consecutive SBs are separated by the HPBW while the optimal switching frequency will be at approximately a quarter of this frequency interval if consecutive SBs are separated by half the HPBW. The latter applies to ARTS when using 151 SBs, so we can find the frequency at which we should switch to the next TAB by

$$f_{\text{sw}} = 0.75f_{\text{max}} + 0.25f_0$$

(4.16)

for even SBs, and by

$$f_{\text{sw}} = 0.25f_{\text{max}} + 0.75f_0$$

(4.17)

for odd SBs. The resulting switching frequencies are shown in Fig. 4.12.

Figure 4.13 shows the sensitivity, expressed relative to the sensitivity of a single WSRT dish, within the CB FoV as function of frequency when using the switching frequencies calculated above. These sensitivity data indicate that the 151 SBs described above provide an average sensitivity across the full FoV of 81% of the maximum achievable sensitivity of the WSRT array. The latter would require to form a TAB phase centred at each individual point within the CB FoV. That would, strictly speaking, require the formation of an infinite number of TABs, which is practically infeasible.
4.A Overview of ARTS hierarchical beamforming

Figure 4.12: Switching frequencies for all SBs of Apertif-10 covering a FoV of 30' over the frequency range from 1200 to 1500 MHz.

Figure 4.13: Sensitivity within the CB FoV as function of frequency when using the switching frequencies shown in Fig. 4.12. The sensitivity is expressed in terms of the sensitivity of a single WSRT dish.
4. A.3 Tracking Beams

The grating response of each TAB rotates around the centre of the FoV of the CB during an observation. This is illustrated in panel e of Fig. 4.2. A given source may therefore traverse multiple SBs during an observation. To track a specific source (or position within the CB) during an observation, we may thus have to concatenate time domain data from multiple SBs. This section describes a procedure to determine which time intervals from which SBs need to be combined to track a desired position within the CB, that is, to form a TB.

Another way to look at the rotation of the TAB gratings is to consider how a specific point within the CB moves through the TAB gratings in a coordinate system fixed to the TAB grating response. This perspective is sketched in Fig. 4.14. At a specific reference time $t = 0$, a specific locus can be specified by cylindrical coordinates $(\theta_0, \phi_0)$, where $\theta_0$ measures the distance from the field centre and $\phi_0$ measures the angle between the line from the field centre to the locus and the line parallel to the array, i.e., the line orthogonal to the grating responses. During an observation, this locus will follow a circular path through the CB with an angular velocity given by $\omega_E$ as indicated by the red track.

The cross-over points between SBs are indicated in Fig. 4.14 by vertical dashed blue lines. The area between two such lines is associated with a specific SB that can be identified by its SB index as described in Sec. 4.A.2. Figure 4.14 shows that it is very convenient to refer to the central SB with SB index 0 as this reduces the problem of finding the SB associated with
4.A Overview of ARTS hierarchical beamforming

A specific locus at a specific instant \( t \) during an observation by finding out in which SB the point \( \theta_0 \cos(\phi_0 + \omega_E t) \) lies. This is easily done by

\[
  n_{\text{SB}} = \left\lfloor \frac{\theta_0 \cos(\phi_0 + \omega_E t)}{\theta_{\text{SB}}} \right\rfloor, \tag{4.18}
\]

where \( \lfloor \cdot \rfloor \) denotes rounding.

Each point in the CB is covered by the grating response of one of the SBs at \( t = 0 \). Eq. (4.18) provides the SB index at each instant \( t \) during the observation. We can now construct a TB for the full observation for a given point in the CB by combining the responses of individual SBs in consecutive time intervals for SBs that consecutively cover that specific point in the CB. This results in an effective TB focused on that point in the CB. The time series associated with these TBs cover the full length of the observation and therefore allow detection of weaker sources than will be feasible to detect during the drift time for an individual SB.

To cover the full FoV of the CB, we can define a hexagonally close-packed grid of TBs. In such a hexagonal pattern, each hexagon consists of 6 equilateral triangles with height \( \theta_{\text{SB}} / 2 \) giving each unit cell an area of \( 3 \theta_{\text{SB}} / 2 \). The area of each CB can be analysed in a similar way as Apertif’s CBs are also stacked in a hexagonally close-packed pattern. The number of required TBs is then equal to \((\theta_{\text{CB}} / \theta_{\text{SB}})^2\). If the maximum number of TABs is used to fill the space between the grating responses of Apertif-10, \( \theta_{\text{SB}} = 0.398' \). If we assume that \( \theta_{\text{CB}} = 30' \), we will need to synthesise approximately 5700 TBs. If we only like to fill the HPBW of the CB, a circular area with diameter \( \theta_{\text{CB}} \), this number reduces to roughly 3000. For shorter observations, this number may be reduced as the TBs will have an elongated shape in the direction perpendicular to the array at the mid-point of the observation, which would allow for a larger separation between TBs along that direction.