*URSI GASS 2021, Rome, Italy, 28 August - 4 September 2021*

# An End-to-End Verification of Pulsar Search Signal Chain for Square Kilometre Array Mid Central Signal Processor

Thushara Gunaratne(1), and Mitchell Mickaliger(2)
(1) Herzberg Astronomy and Astrophysics Research Centre, National Research Council Canada, 717 White Lake Road, Penticton, British Columbia, CANADA, V0H 1K0.
(2) Jodrell Bank Observatory, Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, The University of Manchester, Alan Turing Building, Oxford Road, Manchester, United Kingdom, M13 9PL.

## Abstract

An end-to-end simulation has been conducted in order to verify the performance of the pulsar search (PSS) signal chain for the Square Kilometre Array Phase-1 (SKA1) Central Signal Processor (CSP) for the Mid Telescope. This simulation considered combined signal chains in the Mid Correlator Beamformer (CBF) and the Pulsar Search Engine (PSSE). A frequency-domain method has been developed to generate the test vectors containing the simulated pulses with given spin periods, dispersion characteristics and power. The generated test vectors are first processed by the model for the Pulsar Search (PSS) Beamforming for the Mid CBF. The resultant PSS power beams are then processed by the model for Mid.PSSE, which determines the parameters of the pulsars, if any such signals are present. The simulation results confirm that the proposed signal chain for Mid.CBF and Mid.PSSE meets the requirements.

## 1 Introduction

The Square Kilometre Array (skatelescope.org) telescope will be the largest 'radio' telescope in the World, observing at frequencies in the range 50 MHz to 15.4 GHz. In phase 1 (SKA1) of the Mid-frequency telescope there will be 197 dishes distributed in an area of diameter 160 km in the Karoo region of South Africa. Some of the major science goals of the SKA include searches for gravitational waves and strong-field tests of gravity, in order to further test general relativity. Strong-field tests of gravitational physics can be accomplished by timing pulsars that are in binary systems with massive companions, such as black holes [1]. These *pulsar-black hole binary systems* have never been observed before, and hence are thought to be very rare. Finding these systems will not merely require the sensitivity of the SKA, but also an extensive search for pulsars. In the roughly 50 years since pulsars were discovered [2], less than 3000[1]. have been found [3] Given the projected sensitivity of SKA1 Mid, we expect to discover between 20000 to 30000 pulsars, a factor of 10 more than the current population in only a few years. Therefore, the pulsar search required for the SKA needs to achieve an unprecedented level of speed, automation, and sophistication.

In order to verify that the proposed signal chain configurations meet the requirements, models of the signal chains have been developed for all the signal processing subsystems of SKA1, including the Mid.CBF [4] and the Mid.PSSE [5]. These models for the subsystems use the test vectors generated specifically to test the particular requirements of each subsystem [6, 7]. For example, in the Mid CBF, the sensitivity of PSS and PST beams were determined by cross-correlating the channelized beams against the channelized ideal signal [4]. On the other hand, in Mid.PSSE, pulses are generated using pulsar-simulation software[2], based on an open-source pulsar-search software suite [8], in order to verify the accuracy of algorithms used to determine the spin period, dispersion characteristics, and signal-to-noise ratio (SNR) of detected pulsar candidates.

As a part of the system critical design review of the SKA Telescope, an end-to-end model has been developed by combining signal processing models for subsystems such as the DISH, Mid.CBF and Mid.PSSE, etc. Here, we describe the combined models for Mid.CBF PSS beamforming and Mid.PSSE and the simulations conducted to verify that the requirements are met.

## 2 End-to-End Signal Chain Model

The main signal processing blocks of the end-to-end signal chain for PSS are shown in Figure 1. One of the key features of this model is the signal generator that generates test sequences containing dispersed pulses that are delayed accordingly, as if these are emanating from an off-boresight pulsar and received by dishes at specific locations. The generation of dispersed pulses for the test sequences is explained in detail in Section 3.

These test sequences are processed by the model for the PSS beamforming signal chain for Mid.CBF shown in Figure 1. Note that with Sample Clock Frequency Offset (SCFO) [9] sampling each of these sequences is having a different sampling rate. As shown there, the wideband

[1] ATNF Pulsar Catalogue: www.atnf.csiro.au/people/pulsar/psrcat/

[2] PSS Test Vector Generator: gitlab.com/ska-telescope/pss-test-vector-generator

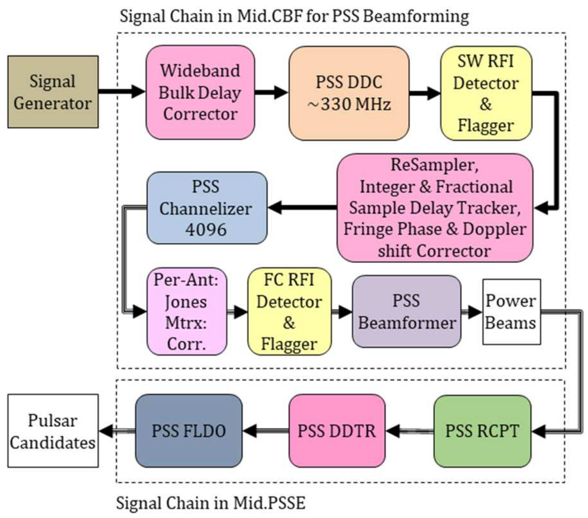


**Figure 1.** The signal chain model for PSS. This is consisted of the cascading signal chain models of the Mid.CBF PSS beamforming and the PSSE.

signal is first processed by the wideband bulk delay corrector. In this module most of the geometric propagation delay [10] is taken out. Following bulk delay correction, the wideband signal is processed by a Digital Down Convertor (DDC) [11] extracting ~330 MHz band from the wideband signal. The central ~300 MHz of this band is 'spectrally pure', that is the leaked spectral components through the stopband are attenuated at least by 60 dB. The extracted band is first analyzed for RFI and then processed by the ReSampler that consists of fractional-delay filter-bank [11] and phase modulator. The ReSampler corrects the remaining geometric propagation delay and resamples the signal sequence into the common sample rate of 330.30144 Msps. The resampled sequence is then segmented into 4096 channels by the PSS channelizer [11]. The central 3687 channels containing the spectrally pure 300 MHz band are first analyzed for RFI and processed by the PSS Beamformer. For each channel, the PSS Beamformer applies the phase-rotations required to steer the beams at the specific directions and adds the corresponding channels from multiple antennas forming 'voltage' beams. The magnitudes of these 'voltage' beams are squared to form 'power' beams.

These power beams are first read into the Mid.PSSE by the receptor module (RCPT), which unpacks the data from Mid.CBF[3]. Once unpacked, the power beams are dedispersed by the dedispersion transform module (DDTR) at many trial *dispersion measures*[4] (DMs), as the test sequences processed by the Mid.CBF are dispersed at an a priori unknown DM. For a given trial DM, this module shifts each channel in time by $D \cdot DM \cdot F_{Ch}^{-2}$, where $F_{Ch}$ is the center frequency of the channel and $D$ is the *dispersion constant* $\approx 4.15\,x\,10^{15}\ \mathrm{Hz^2cm^3pc}$ [12]. Once the channels have been appropriately shifted, all of the channels are added together, producing a time series representing power as a function of time; there is one time series produced for each DM searched. These dedispersed times series are sent on to the folding module (FLDO). Here, each dedispersed time series is first Fourier-transformed. As the data in the dedispersed time series are real-valued, the resulting spectra are symmetric, and only half of the spectra are saved. These frequency spectra are searched for peaks, by looking for frequency bins where the power exceeds some threshold. In the full PSS pipeline, each Fourier bin is added to many of its harmonically related frequency bins, in order to detect signals that have too little power in their fundamental to be detectable above the noise. This is because, in general, each pulsar pulse has a low duty cycle within a spin period, so the power at the spin frequency of the pulsar is spread over many harmonics. In the Mid.PSSE, once the Fourier spectra have been searched, the dedispersed time series, from which a particular Fourier spectrum was created, is folded at the period relating to the frequency of the strongest bin in the Fourier spectrum. The SNR of the peak of each folded pulse for each dedispersed time series are compared, and the DM resulting in a folded pulse with the highest SNR is considered the real DM.

## 3 Generating Dispersed Pulses

A frequency-domain method [13] has been used to generate coherent wideband test sequences that are used to verify the performance of Mid.CBF sub-element [4]. These sequences contain sampled aperiodic Gaussian-distributed pseudo-random signals that correspond to a celestial point source. Due to the limitations in the computational resources, the maximum length of the wideband test sequence that can be generated at a time is approximately 65 million samples. Given that the sampling rate is ~4 Gsps, this corresponds to a signal duration of 16.3 ms. However, the output samples corresponding to partially filled signal processing pipelines are considered 'invalid' and therefore, have to be removed before the integration or beamforming. This takes out another few microseconds worth of integration time. Therefore, in order to conduct an integration or form a beam for a longer duration, multiple test sequences containing ~65 million samples are processed through the signal chains of Mid.CBF and the invalid outputs are discarded before the integration or the beam-sum.

A similar method has been used to generate the dispersed pulses corresponding to a given DM, spin period ($T_{SP}$) and occupying a certain frequency range [$F_l$, $F_u$]. Following [12], the delay between the arrival of spectral components

[3] Note that the Mid.PSSE signal chain differs from the PSS pipeline to be implemented in Mid CSP in that it lacks some processing modules; these missing modules mainly deal with interference and are not required in the end-to-end test to verify the performance requirements

[4]Dispersion Measure (DM) is the integrated column density of electrons along the path of propagation.

corresponding to frequencies $F_l$ and $F_u$ of a pulse that has been propagating through the interstellar medium (ISM) is given by

$$\tau = D \cdot DM \left( \frac{1}{F_u^2} - \frac{1}{F_l^2} \right). \qquad (1)$$

An example of the distribution of spectral components of several dispersed pulses containing spectral components in the band $f \in [F_l, F_u]$ during the observed time interval $t \in [T_S, T_E]$ is shown in Figure 2.

Given that the group delay $\tau(f)$, and phase delay $\phi(f)$, are related such that $\tau(f) \triangleq \frac{-1}{2\pi} \cdot \frac{d\phi(f)}{df}$, it can be shown that the phase delay of a dispersed pulse is given by

$$\phi(f) = \frac{2\pi \cdot D \cdot DM}{f}. \qquad (2)$$

Hence, the spectrum of a dispersed pulse that is shifted by $t = T_D$ is

$$X(f) = H(f) \cdot e^{i2\pi \cdot \left( \frac{D \cdot DM}{f} + f \cdot T_D \right)}, \qquad (3)$$

where $H(f)$ is the magnitude spectrum of the pulse. The inverse-Fourier transform of $X(f)$ results in a time series of the 'dispersed pulse'.

The linearity of Fourier and inverse-Fourier transforms [14] is exploited in generating 'dispersed pulse trains', modeling the pulsar signals received by the dishes in the SKA1 Mid. For efficient evaluation of the sampled sequences of the dispersed pulses, the discrete Fourier transform (DFT) methods [13, 14] are used. However, one of the key issues to consider here is that the inverse discrete Fourier transform (IDFT) produces a periodic sequence and thereby the dispersed-pulse can 'wrap around'. Considering the example shown in Figure 2, using the expression for a delayed dispersed pulse given in (3), the combined spectrum, $X_T(f)$, of the dispersed pulse-train for the duration $t \in [T_S, T_E]$ can be expressed as

$$X_T(f) = X_L(f) + X_R(f) + H(f) \cdot e^{i2\pi \cdot \left( \frac{D \cdot DM}{f} + f \cdot (T_0 - T_S) \right)} \cdot \sum_{k=0}^{2} e^{i2\pi k f T_{SP}}, \qquad (4)$$

where $X_L(f)$ and $X_R(f)$ are the spectra of the left-most and right-most dispersed pulses in Figure 2, respectively. Note that $X_L(f)$ can be expressed as

$$X_L(f) = H_L(f) \cdot e^{i2\pi \cdot \left( \frac{D \cdot DM}{f} + f (T_0 - T_{SP} - T_S) \right)}, \qquad (5)$$

where $H_L(f) = 0$ for $f \in [\acute{F}_u, F_u]$. This prevents the IDFT of the dispersed pulse to appear in the region $t \in [T_E - (T_0 - T_{SP}), T_E]$. Similarly, $X_R(f)$ can be expressed as

$$X_R(f) = H_R(f) \cdot e^{i2\pi \cdot \left( \frac{D \cdot DM}{f} + f (T_0 + 3T_{SP} - T_S) \right)}, \qquad (6)$$

where $H_R(f) = 0$ for $f \in [F_l, \acute{F}_l]$. This prevents the IDFT of the dispersed pulse to appear in the region of $t \in [T_S, T_0]$.

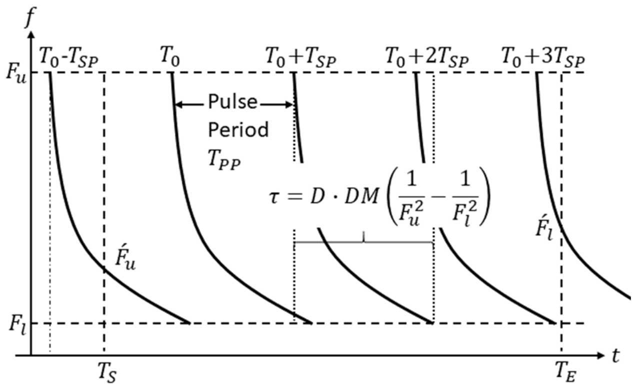


**Figure 2.** An example of the distribution of spectral components of dispersed pulses with respect to observed time.

The spectrogram of a series of dispersed pulses in the SKA1 Band 2 (i.e. $f \in [0.96, 1.76]$ GHz) [6], generated using the aforementioned method with $DM = 1$ cm$^{-3}$pc and $T_{SP} = 0.0015$ s, within $t \in [0.0210, 0.0335]$ s is shown in Figure 3. Note that there is no 'wrapping around' of the dispersed pulses at $t = T_S$ or $t = T_E$.

The spectrum $X_T(f)$, of the dispersed pulse train as specified in (4) is combined with the method proposed in [13] to generate the sample sequences having the corresponding time varying delays as received by the SKA1 Mid dishes at specific locations tracking an off-boresight pulsar.

# 4 Simulation Results

The end-to-end simulation of the PSS signal chain for Mid.CBF consists of two parallel branches of signal chains, performing 'unquantized' and 'quantized' signal processing [4, 5]. At the end, the SNRs of each of the folded pulses evaluated in the two branches are compared. This quantifies the 'loss of sensitivity' [10] due to the quantization at the input as well as the requantization at each signal processing module along the signal chain. These measurements are used in determining whether the proposed hardware architecture for PSS combining the Mid.CBF and the Mid.PSSE is capable of meeting the requirements [6,7].

For the following two simulations, the signal chain model was configured so that the power beams are formed combining the two signals received by two SKA1 Mid dishes that are positioned 20 km apart from each other along an east-west baseline. These two elements were observing the frequency range $f \in [1.4, 1.7]$ GHz in SKA1 Band 2 and tracking a pulsar at the delay-center on the sky at $\delta_S = 0°$ declination and $\alpha_S = 75°$ right ascension at the sidereal rate. It has been shown that this results in the worst-case sensitivity loss [4].

For this initial test, the test sequences were generated as specified in Section 3 with DM = 55 cm$^{-3}$pc and $T_{SP}$ = 0.025 s; both of these parameters were unknown to the

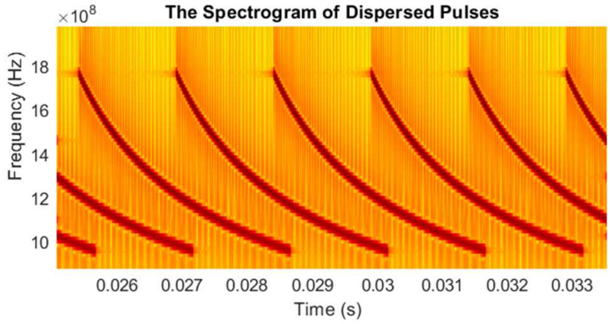


**Figure 3.** The spectrogram of a series dispersed pulses generated using the method outlined in Section 3.

Mid.PSSE during the processing. Note that the generated test sequences did not contain any added receiver noise. This test vector is then processed by the 'unquantized' signal chain, whereas in 'quantized' signal chain, the test vector is optimally quantized to 12-bits before processing. After running these two data sets through the two branches of the signal model, both produced a pulsar candidate with a DM of 55 $cm^{-3}$pc and a spin period of 0.025 s, exactly matching the input parameters. In order to determine the effect of the quantization caused by channelization in Mid.CBF, the SNR of the folded pulse from each data set was evaluated and compared. Here, the SNR of the folded pulse measured by the quantized signal chain was 98% of the SNR achieved in the unquantized signal chain, which is well within the requirements [7].

A second test was carried out, this time including noise in both the quantized and unquantized input test sequences. The processing of both sequences with the two branches of the signal chain yielded a pulsar candidate with a DM of 45 $cm^{-3}$pc and a spin period of 0.0625 s, again exactly matching the input parameters. In this case, the peak SNR of the folded pulse achieved with the quantized signal chain was 99% the SNR of the folded pulse detected in the unquantized signal chain.

## 5 Conclusions

The combined signals chain of the Mid.CBF and Mid.PSSE is verified through simulations. A frequency-domain method has been used in generating the dispersed pulses corresponding to different DMs and spin periods. The detection of the correct spin periods and DMs, unknown a priori, of two pulsar-like signals by the Mid.PSSE signal chain demonstrates that it performs its intended function. Also, the recovery of over 98% of the SNR of the unquantized signal chain by the quantized signal chain proves that the Mid.CBF PSS beamformer is performing the key signal processing operations including delay tracking well above the specified SNR requirement.

## 6 References


1. R. Smits, M. Kramer, B. Stappers, D. R. Lorimer, J. Cordes, and A. Faulkner, "Pulsar searches and timing with the square kilometre array," *Astronomy and Astrophysics*, vol. 493, no. 3, pp. 1161–1170, 2009, doi: 10.1051/0004-6361:200810383.

2. A. Hewish, S. J. Bell, J. D. H. Pilkington, P. F. Scott, and R. A. Collins, "Observation of a Rapidly Pulsating Radio Source," *Nature*, vol. 217, no. 5130, pp. 709–713, Feb. 1968, doi: 10.1038/217709a0.

3. R. N. Manchester, G. B. Hobbs, A. Teoh, and M. Hobbs, "The Australia Telescope National Facility Pulsar Catalogue," *The Astronomical Journal*, vol. 129, pp. 1993–2006, Apr. 2005, doi: 10.1086/428488.

4. T. Gunaratne, B. Carlson, G. Comoretto, M. Rupen, and M. Pleasance, "An end-to-end model for the correlator and beamformer of the Square Kilometer Array Mid Telescope," in *Modeling, Systems Engineering, and Project Management for Astronomy IX*, Dec. 2020, vol. 11450, p. 1145002, doi: 10.1117/12.2559754.

5. Square Kilometre Array Organization, SKA1 CSP Pulsar Search Sub-element Signal Processing MATLAB Model (ED-7) SKA-TEL-CSP-0000085, Rev. 5, 31 August 2018.

6. Square Kilometre Array Organization, SKA1 CSP Mid Correlator and Beamformer Sub-element Requirement Specification (EB-1) 311-000000-006, Rev. 2, 4 July 2018.

7. Square Kilometre Array Organization, SKA1 CSP Mid Pulsar Search Sub-element Requirement Specifications (ED-1a) 313-000000-001, Rev 5, 20 June 2018.

8. SIGPROC - Pulsar Signal Processing Software (M. Keith's release of D. Lorimer's SIGPROC), http://www.github.com/SixByNine/sigproc.

9. B. Carlson and T. Gunaratne, "Signal processing aspects of the sample clock frequency offset scheme for the SKA1 mid telescope array," in *2017 XXXIInd General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS)*, Aug. 2017, doi: 10.23919/URSIGASS.2017.8105156.

10. A. R. Thompson, J. M. Moran, and G. W. Swenson, *Interferometry and Synthesis in Radio Astronomy*, 2 edition. New York: Wiley-VCH, 2001.

11. F. J. Harris, *Multirate Signal Processing for Communication Systems*, 1 edition. Upper Saddle River, N.J: Prentice Hall, 2004.

12. D. R. Lorimer and M. Kramer, *Handbook of pulsar astronomy*. Cambridge University Press, 2005.

13. T. K. Gunaratne, "Generation of Coherent Signals for the Verification of Signal Processing Algorithms in Radio Astronomy." *2019 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*, 2019, pp. 1–6.

14. J. E. Proakis and D. G. Manolakis, *Digital Signal Processing; Principle, Algorithms and Applications*. Prentice-Hall Inc., Englewood Cliffs, New Jersey 07458, 1996.