

Full length article

A novel greedy approach to harmonic summing using GPUs

K. Adámek^a, J. Roy^b, W. Armour^{c,*}^a Faculty of Information Technology, Czech Technical University, Thákurova 9, 160 00, Prague, Czech Republic^b National Centre for Radio Astrophysics, Tata Institute of Fundamental Research, Pune 411 007, India^c Oxford e-Research Centre, Department of Engineering Sciences, University of Oxford, 7 Keble road, OX1 3QG, Oxford, United Kingdom

ARTICLE INFO

Article history:

Received 13 October 2021

Accepted 15 July 2022

Available online 19 July 2022

Keywords:

GPUs

Pulsar search

Fourier transform

Radio astronomy

ABSTRACT

One-dimensional incoherent harmonic summing is a technique used to improve the sensitivity of Fourier domain search methods. In time-domain radio astronomy, the harmonic sum is a part of the Fourier domain periodicity search, which aims to detect isolated single pulsars. We present a new harmonic sum algorithm based on the greedy approach and an implementation of this on NVIDIA GPUs using the CUDA programming language. The new algorithm determines which samples to add according to short-term gains. We present an evaluation of the sensitivity of the new approach and its performance against the PRESTO harmonic sum a standard used in radio astronomy. The new Greedy harmonic sum has, on average, 20% higher sensitivity whilst it has the same performance as the standard algorithm. This work forms part of the AstroAccelerate project which is a GPU accelerated software package for processing time-domain radio astronomy data.

© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A convenient and computationally efficient way of detecting pulsars in time-domain radio astronomy is to use techniques based on the Fourier transform. The Fourier transform utilises the periodic nature of the pulsar's signal, but the use of the Fourier transformation is not without pitfalls. For example, variations in the pulsar's, otherwise periodic, emissions lead to degradation of the signal in the Fourier domain. This can occur when a pulsar is present in a binary system with another compact object (e.g. white dwarf, neutron star, black hole). In such a case, the Doppler effect changes the pulsar's period over the observing span resulting in a smearing of the response in the Fourier plane. The standard Fourier transformation is then inadequate for detecting such pulsar, and other methods, such as FDAS (Ransom et al., 2002; Dimoudi et al., 2018), must be used.

The pulsar's signal can be described as a pulse train, where pulsar emissions are visible only in a portion of its period. The proportion between the part of the period when we see the pulsar's signal and the period duration is called the *duty cycle*. For such signals, the Fourier transform distributes the power contained in the pulsar's signal into the fundamental frequency bin and multiple higher harmonic bins. The source of these harmonics is the combined effects of sampling and aliasing introduced by the discrete Fourier transform. In effect, some of the power contained in the pulse train is also added to the higher harmonics, which

can be used to boost the pulsar's signal in the Fourier domain and increase the probability of detection. Moreover, this spread of power in higher harmonics increases for narrow duty cycle pulsars.

The incoherent harmonic sum (HS) aims to increase the pulsar's signal-to-noise ratio (SNR) by accounting for power (either fully or partially) from an increasing number of harmonics. Thus, the harmonic sum adds the signal present in the data and averages out the noise, increasing the pulsar's SNR and consequently the probability of detection.

Taylor and Huguenin (1969) first used the harmonic sum in 1969, only a few years after the discovery of pulsars by Antony Hewish and Jocelyn Bell Burnell (Hewish et al., 1968). More recently, Ransom et al. (2002) described how the duty cycle of the pulsar affects the number of significant harmonics. In the same work, Ransom et al. (2002) also compared data from 600 pulsars and showed that a pulsar with a duty cycle below 0.1 could have more than ten significant harmonics. This shows the importance of the harmonic sum algorithm and its role in pulsar detection. Furthermore, Yu (2018) performed a theoretical probabilistic analysis of the harmonic sum technique and proved the usefulness of this technique.

The one-dimensional harmonic sum algorithm is a standard function present in many software packages which process time-domain radio astronomy data, such as SIGPROC (Lorimer, 1999) or the PRESTO software package (Ransom, 2002). However, none of these existing packages have the harmonic sum in a computationally accelerated form. The higher-dimensional harmonic

* Corresponding author.

E-mail address: wes.armour@oerc.ox.ac.uk (W. Armour).

sum is used in the Fourier domain acceleration search technique (FDAS) (Ransom et al., 2002), or JERK search (Andersen and Ransom, 2018), which searches for accelerated pulsars.

In this work, we present a new algorithm to perform the harmonic sum and GPU implementations of this algorithm, evaluating its sensitivity and performance. Like the existing CPU implementations this algorithm is also approximate and does not perform an exhaustive harmonic sum. For example, the Lyne–Ashworth algorithm (described in Lorimer and Kramer (2004)) used in the SIGPROC package, widely used in pulsar detection, sums only harmonics which are powers-of-two. The harmonic sum algorithm used in the PRESTO software package (Ransom, 2002) also uses a subset of the harmonics.

We propose a novel harmonic sum algorithm that sums all harmonics, not only powers-of-two, and is based on a greedy approach. Our algorithm, called Greedy harmonic sum, does not perform an exhaustive search in the Fourier domain. Instead, we have used a greedy method, where the algorithm determines which time samples to sum according to the short-term gains rather than finding the optimal sum of harmonics. The Greedy harmonic sum achieves higher sensitivity with a performance similar to standard harmonic sum algorithms.

To compare our proposed algorithm in both sensitivity and performance, we have implemented a PRESTO harmonic sum algorithm on NVIDIA GPUs using CUDA. Furthermore, we have implemented and used for comparison a modified version of PRESTO harmonic sum that has higher sensitivity and sums all harmonics, but it is substantially slower.

We have integrated the algorithms discussed in this work into AstroAccelerate¹ (Armour et al., 2020), a GPU accelerated real-time data processing software package for time-domain radio astronomy. The AstroAccelerate package contains GPU accelerated versions of commonly used algorithms for time-domain radio astronomy. These are de-dispersion transform (Armour et al., 2012), single pulse detection (Adámek and Armour, 2020), Fourier domain acceleration search (Dimoudi et al., 2018) and accelerated JERK search (Adámek et al., 2020). AstroAccelerate is actively used as part of scientific pipelines like MeerTRAP (Morello et al., 2021; Rajwade et al., 2021), Greenburst (Agarwal et al., 2020) and GHRSS survey.²

This paper is organised as follows: in Section 2 we describe the harmonic sum, its purpose and problems cause by sampling; in Section 3 we introduce our new greedy harmonic sum algorithm and also describe existing algorithms; in Section 4 we present signal loss, recovered SNR, and compute performance of selected harmonic sum algorithms; in Section 5 we discuss achieved results, and lastly we conclude the paper in Section 6.

2. Harmonic sum

The incoherent harmonic sum is given by equation

$$X(f)_H = \sum_{h=1}^H x(hf_p), \quad (1)$$

where H is the total number of harmonics summed, f is frequency, and x is the continuous power spectra. The integer multiple h will be called the *harmonic order* throughout this work. The harmonic sum algorithm produces a series of partial sums. For each sum it calculates its SNR and then selects the maximum. That is

$$\text{SNR}(X(f)_H) = \max_H \left(\frac{X(f)_H - \mu_H}{\sigma_H} \right), \quad (2)$$

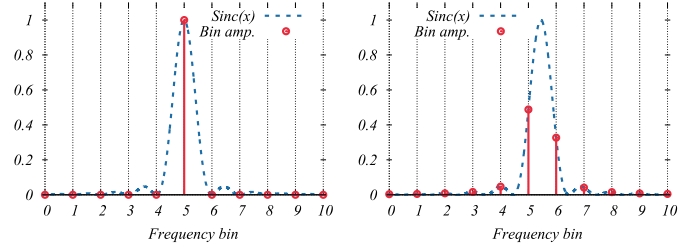


Fig. 1. Scalloping loss due to a discrete representation of frequencies.

where μ_H and σ_H are the mean and standard deviation respectively for partial sum of length H . The number of harmonics we need to sum is governed by the duty cycle of the pulsar we are looking for.

When searching for a pulsar, where the frequency of the pulsar is unknown, the harmonic sum algorithm needs to perform partial sums up to the highest harmonic for every frequency bin (up to a limit, for example, dictated by observation or a theoretical model).

The harmonic sum algorithm is performed after Fast Fourier Transformation (FFT) and further preprocessing, which includes spectrum whitening or power spectrum calculation which may include inter-binning as introduced by Ransom et al. (2002).

2.1. Discrete harmonic sum

When we have continuous frequencies, it is easy to pick the correct frequency of the higher harmonics as it is just an integer multiple of the pulsar's frequency $f_h = hf_p$. However, when dealing with discrete data choosing the correct frequency for summation is more complicated, and it is at the core of any harmonic sum algorithm.

In the discrete data, the true frequency of the pulsar f_p is represented by a *fundamental frequency* f_0 which could be different from the pulsar's true frequency due to the finite resolution of the frequency representation of the observed data. Each frequency represented in the discrete Fourier domain has a *frequency bin*, an integer index of that frequency. Therefore a bin that corresponds to the fundamental frequency is called the *fundamental frequency bin*, and frequency bins that represent higher harmonics will be called *harmonic bins*. We will use the terms bin and frequency interchangeably since there is a direct link between them.

When the frequency we want to represent in the discrete spectrum does not coincide with the centre of any of the frequency bins, then the signal's power is distributed across all frequency bins – the distribution follows a Sinc function. The consequences of the redistribution of power across all frequency bins is shown in Fig. 1, where we see a noticeable decrease in amplitude for the frequency outside the centre of the frequency bin compared to the represented frequency at the centre of the frequency bin. The effect of decreased amplitude for off-centre frequencies is called a *scalloping loss*. The scalloping loss can cause up to 36% loss (Lyons, 2011) of the amplitude, assuming we know which bin to pick.

2.2. Peak drift

The problem of picking the correct harmonic bin for summation is that the representation of the higher harmonic frequencies is, in the discrete spectrum, based on the integer multiple of the pulsar's true frequency $f_h = hf_p$ and not on the fundamental frequency f_0 . Therefore, there might be a frequency bin that more closely represents the higher harmonic f_h . The frequency of this

¹ <https://github.com/AstroAccelerateOrg/>.

² <http://www.ncra.tifr.res.in/~bhaswati/GHRSS.html>.

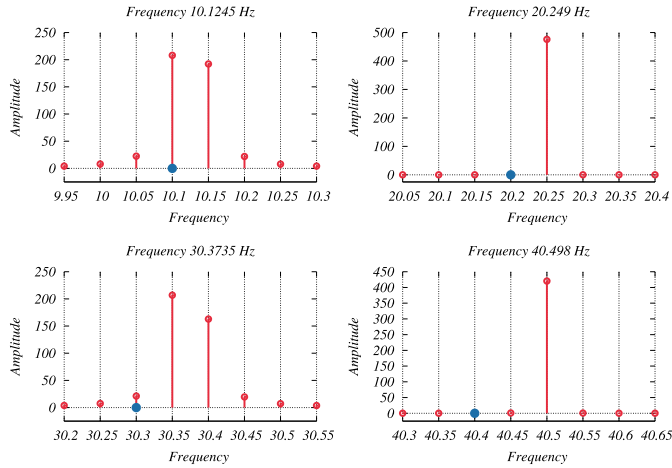


Fig. 2. Signal of the pulsar with fundamental frequency $f = 10.1245$ Hz after Fourier transformation. Individual plots shown are the fundamental frequency and the first three harmonics. The integer multiple of the fundamental frequency bin ($f_b = 10.1$ Hz) is emphasised by the filled blue point.

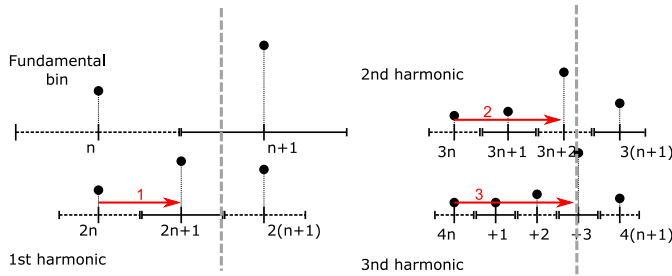


Fig. 3. Frequency bin drift, when interpreted from one bin centre to another, is an increasing function. The drift and its value are in red. The grey dashed vertical line represents the true frequency value and its integer multiples at higher harmonics.

bin might be different from the integer multiple of the fundamental frequency bin hf_0 , thus introducing *drift* in the peak position at higher harmonics. The drift is illustrated in Fig. 2, where the frequency $f = 10.1245$ is chosen to be almost exactly in between two frequency bins. A filled blue point depicts the fundamental frequency bin and its integer multiples. We see that at higher harmonics, a mere multiple of the fundamental frequency bin would miss the peak and would result in no increase to the signal-to-noise ratio of the pulsar.

In this work, we have adopted an interpretation that the peak can drift from one centre of the frequency bin to another, rather than being a peak's distance from the centre of one frequency bin. Therefore, we define peak drift as the distance from the fundamental frequency bin (or its integer multiple) to the peak at the same harmonic order. In our interpretation, the drift is an increasing function, while the value of deviation from the frequency bin $n + 1$ centre could also decrease. Our chosen interpretation is depicted in Fig. 3, where the true frequency is outside the fundamental frequency bin n and lies in the first half of the frequency bin $n + 1$. As we sum higher harmonics, the peak (true pulsar frequency) position is more precisely represented. This is because, in our interpretation, higher harmonics will have a higher resolution in frequency because the interval between bins n and $n + 1$ is mapped into more and more bins, which lay between bins hn and $h(n + 1)$, as the harmonic order h increases.

Since the drift is either increasing by one or stays the same, we can picture it as a tree of nodes, as shown in Fig. 4, where there is an ideal path and two different trials. With this model,

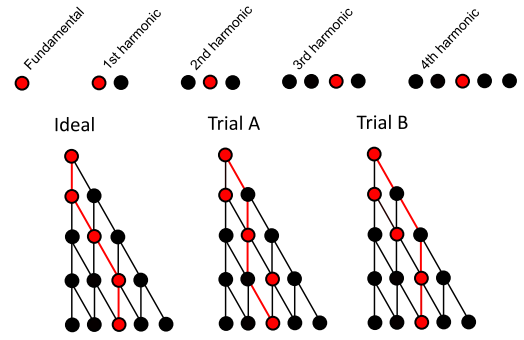


Fig. 4. One potential partial sum of the harmonic sum algorithm for a fixed fundamental frequency bin could be interpreted as a path in a tree where the layers are made out of frequency bins at higher harmonics. Each path (shown in red) would then be defined only by the drift value at harmonic order h .

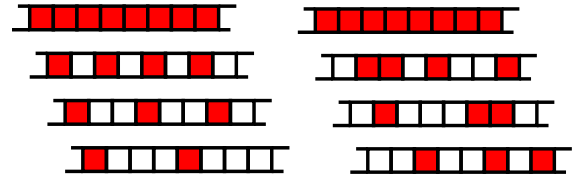


Fig. 5. Stride in the memory access pattern for the harmonic sum without drift on the left and including drift on the right.

each harmonic sum for a fixed fundamental frequency is one path through the tree.

The *maximum drift* a peak undergoes depends on the harmonic order which is being summed, and it is given as $d_m(h) = h - 1$. This is the number of bins between two consecutive fundamental frequency bins at the same harmonic order. The actual drift a peak acquires depends on the pulsar's frequency and increases with the difference between the fundamental frequency f_0 and the true frequency of the pulsar f_p .

3. Algorithm and implementation

We have developed a new harmonic sum algorithm based on the greedy approach. The algorithm prefers short term SNR gain rather than the optimal partial sum, which could give the higher SNR gain for a given fundamental bin. This algorithm is then compared with the established harmonic sum algorithm, which is used in the PRESTO software package (Ransom, 2002). The second harmonic sum algorithm used for comparison is the simple harmonic sum which ignores the drift and sums only integer multiples of the fundamental bin. We have implemented GPU versions of all of these harmonic sum algorithms, and we compared these algorithms in performance and sensitivity.

Any harmonic sum algorithm must solve two main problems. The first is an unfavourable memory access pattern, and the second is the number of possible partial sums explored per a single fundamental bin.

The poor memory access pattern is due to the distance between integer multiples of two consecutive fundamental bins. This distance increases with the harmonic order, which is being added to the partial sum starting with one up to the maximum harmonic order. In other words, we access the data with an increasing stride. Fig. 5 (left) illustrates such a memory access pattern that, in addition, ignores the peak's drift. Taking peak drift into account, we add randomness into the memory access pattern, as shown in Fig. 5 (right). The stride in any memory accesses represents a problem because it significantly reduces data

utilisation per cacheline³ unless this is mitigated by CPU/GPU cache.

The second problem faced by the harmonic sum algorithm is how to limit the number of partial sums performed. As the peak drift is unknown, to get the highest SNR, we need to explore a number of partial sums containing different elements to estimate the value of the drift. However, the calculation of these various partial sums can be computationally expensive. Therefore the harmonic sum algorithm must prioritise which partial sums to use.

In the following sections, we will describe our new Greedy harmonic sum algorithm as well as other harmonic sum algorithms that are used for comparison.

3.1. Simple harmonic sum

The simple way to perform the harmonic sum is to ignore the peak's drift and sum only the values located at the integer multiple of the fundamental frequency bin. This approach has the potential to be fast, but it sacrifices SNR gain for performance. The simple harmonic sum can be summarised by the equation

$$X[n] = \sum_{h=1}^H x[hn] , \quad (3)$$

where n is the fundamental frequency bin and h is the harmonic order. The simple harmonic sum serves as a bottom line for other harmonic sum algorithms regarding SNR gain.

3.1.1. Implementation

In our GPU implementation of the simple harmonic sum, each GPU thread calculates the harmonic sum for one fundamental bin. The threads that are next to each other process fundamental bins in sequence. Since the data for higher harmonics are accessed with an increasing stride, as discussed above and shown in Fig. 5, this decreases cacheline utilisation. When reading fundamental bins, a cacheline is fully utilised. In contrast, for higher harmonics, it is utilised to $1/h$, where h is the harmonic order used. Cacheline utilisation has the potential to decrease performance if not mitigated by caches. In the case of the simple harmonic sum, which does not suffer from random access to memory introduced by the drift, we can improve the access pattern if we transpose the data. Instead of having fundamental bins from a single time-series spatially close to each other, we change the layout of the data to put the same bin from different time-series next to each other. The advantage of this transformation is that when reading data for higher harmonic orders, we read data that are contiguous in memory. That way, threads process the same fundamental bin but from different time-series. It also means we have to pay the price of transposing data, however.

During our investigation, we discovered that the transpose of the data is not beneficial on the newer generation of NVIDIA GPUs since they have a write-through cache. The older generations of NVIDIA GPU benefit more from such pre-processing.

3.2. PRESTO harmonic sum

The harmonic sum algorithm used in the PRESTO software package is based on the following formula

$$X[n] = \sum_{h=1}^H x \left[\frac{h}{H} n \right] , \quad (4)$$

³ A cacheline refers to a block of data that is spatially next to each other and transferred together in one memory transfer. This is due to memory architecture.

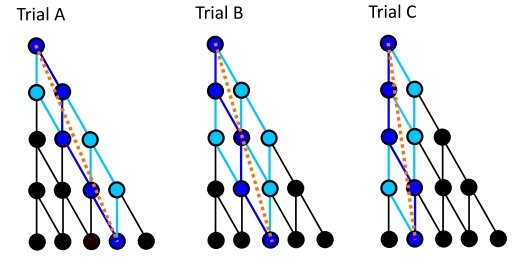


Fig. 6. PRESTO chooses a path based on the last element of the sum.

where as before H is the number of harmonics to sum, and x is the input time-series. The index n in the case of the PRESTO algorithm is not the fundamental bin but the last bin to be summed in a harmonic series.

Algorithm 1: Pseudo-code for the PRESTO harmonic sum for a bin n . A series of harmonic sums of increasing harmonic order H is performed on the input data x up to a maximum harmonic order H_{\max} . Output is the maximum SNR found SNR_{\max} and related harmonic order h_m . To calculate SNR, mean $\mu[h]$ and standard deviation $\sigma[h]$ is used.

```

array  $\mu[H_{\max}], \sigma[H_{\max}]$ ;
 $\{\mu, \sigma\} = \text{GetMeanStdev}()$ ;
Sum = 0;
 $\text{SNR}_{\max} = 0$ ;
 $h_m = 0$ ;
for  $H = 1$  to  $H_{\max}$  by  $H = 2H$  do
    Fundamental frequency bin;
     $P = \text{Round}(1/Hn)$ ;
    Sum =  $x[P]$ ;
    for  $h = 2$  to  $H$  by  $h = h + 2$  do
         $P = \text{Round}(h/Hn)$ ;
        Sum = Sum +  $x[P]$ ;
        SNR = CalculateSNR(Sum,  $\mu[h], \sigma[h]$ );
         $\{\text{SNR}_{\max}, h_m\} = \text{CompareSNR}(\text{SNR}, h, \text{SNR}_{\max}, h_m)$ ;
    end
     $\text{SNR}_o = \text{SNR}_{\max}$ ;
     $h_o = h_m$ 
end

```

By starting from the last element, some paths are eliminated, the remaining paths are further reduced by selecting only one path by rounding bin location $(h/H)n$ to the nearest integer. Paths are chosen for the same fundamental, but different end elements of the harmonic sum are shown in Fig. 6, where the chosen path is shown in dark blue, and all the possible paths leading to that element are shown in light blue. The pseudo-code for the PRESTO harmonic sum algorithm can be found in Algorithm 1.

The advantage of this algorithm is that it covers most of the drift possibilities and it is deterministic. The disadvantage of this approach is that the last element from which the algorithm begins, acts as the last element for multiple fundamental bins but at different harmonic orders. For example, bin 800 is the last element for the fundamental frequency bin 100 at $H = 8$ or for fundamental frequency bin 200 at $H = 4$. Furthermore, the set of elements that need to be summed differs by the fundamental bin. Therefore, we have to perform an additional loop (the first loop in Algorithm 1) to cover these fundamental bins and thus increase the complexity of the algorithm to $\mathcal{O}(H^2)$. These harmonic sums may include different elements depending on rounding, making any data reuse difficult.

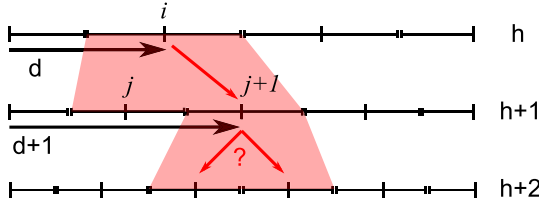


Fig. 7. Localisation of the peak using a binary search like algorithm.

The PRESTO algorithm for each bin n performs all partial sums that are power-of-two long $H_p = 2^i$, that is, for $H = 32$ partial sums performed would be $H_p = \{2, 4, 8, 16, 32\}$. Further, PRESTO sums only even harmonics. These choices lower the complexity of the algorithm to $\mathcal{O}(H \log_2(H))$, but at the cost of sensitivity.

There are two ways to decrease sensitivity loss and thus improve the SNR of the PRESTO harmonic sum. Improvements however, come at the expense of algorithm performance. The first improvement is to sum elements of all higher harmonics, which we will call PRESTO_{all}. The second is to perform the harmonic sum for all harmonic orders in addition to summing all higher harmonics. We will call this algorithm PRESTO+.

We have used the GPU implementation of the PRESTO harmonic sum and PRESTO+ harmonic sum as a standard to which we compare the greedy harmonic sum algorithm. Our implementation of the PRESTO harmonic sum algorithm was tested against the PRESTO package. Both algorithm implementations agree on the result: candidates frequency bin, detected number of harmonics, period and summed power. The performance was not considered.

3.3. Greedy harmonic sum

The greedy algorithm tracks the peak position and updates the drift value if the peak shifts. The greedy algorithm is listed in Algorithm 2. At each new harmonic order, the algorithm loads values of two consecutive frequency bins. The first frequency bin index m is

$$m = hn + d, \quad (5)$$

where n is the fundamental bin index, h is the harmonic order, and d is the value of the drift. Values of both bins are compared, and if the value of the second frequency bin is higher, the peak has shifted and drift is increased by one.

The algorithm for following the peak position as harmonic order increases is inspired by the binary search algorithm. The binary search algorithm finds a number in a sorted array of numbers by splitting the array into two sub-arrays. Then (for an ascending array), the number in the middle is compared to the number we are searching for, and if the number we are searching for is smaller, we continue the search in the first half of the array. If not, the second half of the array is used for further searches. The total number of sub-arrays is doubled with each iteration, as each sub-array is split in two. This allows us to contain the number we are searching for in ever-smaller sub-arrays until it is found.

In contrast to the binary search method, the number of sub-intervals that divide the interval between two consecutive fundamental frequency bins (fundamental interval) is not doubled at each harmonic order but only increased by one to $n + 1$ dividing sub-intervals. The situation is depicted in Fig. 7.

Let us suppose that we want to use a similar algorithm to the binary search for peak position at higher harmonic orders. In this case, we must show that a frequency bin i with drift d and width w at harmonic order h is contained in two consecutive frequency bins j and $j + 1$ at harmonic order $h + 1$. That is, the

Algorithm 2: Pseudo-code for the greedy harmonic sum. Where H is maximum harmonic order summed, x is input data in the frequency domain, SNR_{\max} is maximum SNR found, h_m is related harmonic order, and $\mu[h]$, $\sigma[h]$ is mean and standard deviation respectively for partial sum of h elements. The algorithm's output is the maximum SNR found $\text{SNR}_o[n]$ for each relevant fundamental frequency bin and related harmonic order h_o . The variable drift keeps track of the peak drift in the number of frequency bins.

```

array  $\mu[H], \sigma[H]$ ;
 $\{\mu, \sigma\} = \text{GetMeanStdev}()$ ;
HarmonicSum begin
  for  $n = 1$  to  $N$  do
     $\text{sum} = 0$ ;
     $d = 0$ ;
     $\text{SNR}_{\max} = 0$ ;
     $h_m = 0$ ;
    fundamental bin;
     $x_d = x[n]$ ;
     $x_{(d+1)} = x[n + 1]$ ;
    if  $x_{(d+1)} > x_d$  then
       $d++$ ;
       $\text{sum} = \text{sum} + x_{(d+1)}$ ;
    else
       $\text{sum} = \text{sum} + x_d$ ;
    end
     $\text{SNR}_{\max} = \text{CalculateSNR}(\text{sum}, \mu[0], \sigma[0])$ ;
    higher harmonics;
    for  $h = 1$  to  $H$  do
       $x_d = x[hn + d]$ ;
       $x_{(d+1)} = x[hn + d + 1]$ ;
      if  $x_{(d+1)} > x_d$  then
         $d++$ ;
         $\text{sum} = \text{sum} + x_{(d+1)}$ ;
      else
         $\text{sum} = \text{sum} + x_d$ ;
      end
       $\text{SNR} = \text{CalculateSNR}(\text{sum}, \mu[h], \sigma[h])$ ;
       $\{\text{SNR}_{\max}, h_m\} = \text{CompareSNR}(\text{SNR}, h, \text{SNR}_{\max}, h_m)$ ;
    end
     $\text{SNR}_o[n] = \text{SNR}_{\max}$ ;
     $h_o = h_m$ 
  end
end

```

frequency $f_{i,b} = f_i - w/2$ at the beginning of the bin i , mapped to a higher harmonic order $h + 1$, must be higher than the frequency $f_{j,b} = f_j - w/2$ at the beginning of the frequency bin j :

$$f_{j,b} \geq \frac{h+1}{h} f_{i,b}, \quad (6)$$

where $(h + 1)/h$ maps frequencies from harmonic order h to the frequency range of the harmonic order $h + 1$. Substituting $f_i = hf_f + dw - w/2$ and $f_j = (h + 1)f_f + dw - w/2$ into Eq. (6) gives condition

$$2d - 1 \geq 0, \quad (7)$$

which holds for $d > 0$. The bin at $d = 0$ is at the beginning of the fundamental interval where we have only the second half of the bin therefore this rule cannot be applied. However all bins with $d = 0$ begin at integer multiple of the fundamental bin therefore the condition is met.

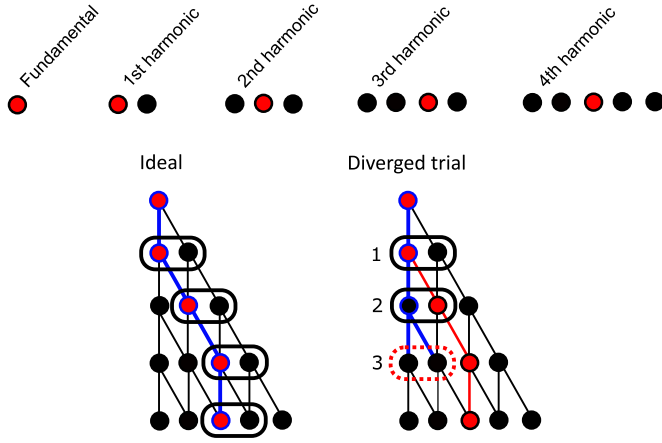


Fig. 8. Greedy algorithm. Decision points are shown as rounded boxes.

The second condition is that the end of the bin i must be less than, or equal to, the end of the second bin $j + 1$. That is

$$\frac{h+1}{h} f_{i,e} \leq f_{(j+1),e}, \quad (8)$$

where $f_{i,e} = hf_f + dw + w/2$ is the frequency at the end of the frequency bin i and $f_{(j+1),e} = (h+1)f_f + (d+1)w + w/2$ is the frequency at the end of the bin $j + 1$. After substitution we arrive at the condition

$$2d + 1 \leq 2h. \quad (9)$$

As the maximum value of the drift is $d = h - 1$, the inequality (9) holds for any value $d > 0$.

Hence, we can track the position of the peak as the drift d changes with increasing harmonic order h .

The greedy algorithm evaluates which element to sum based on short term SNR gain. Each rounded box in Fig. 8 shows a decision that chooses the next direction of the summation. However, a single element containing a high noise value can potentially divert the greedy algorithm to a path which, in the long term, sums to a lower SNR than it otherwise would have if the algorithm followed the correct path. The divergent path is shown in Fig. 8 on the right.

To increase the recovered SNR further, we can apply the same principle of selecting the bin with the highest value for the fundamental frequency bin. Since drift is defined from one bin centre to the next, it is possible that the true frequency is closer to the frequency bin $n + 1$, which is next to the fundamental frequency bin n . Bin $n + 1$ would then have a higher amplitude than the fundamental frequency bin n . However, as the true frequency is lower in the centre of the bin $n + 1$, the higher harmonics are contained between integer multiples of bins n and $n + 1$. This increases recovered SNR.

4. Results

The comparison between our Greedy Harmonic sum algorithm and both PRESTO harmonic sum algorithms is presented in this section. We will compare the implemented harmonic sum algorithms ability to detect pulsar signals and their compute performance. To compare the ability to detect pulsars, we use the recovered SNR (RSNR) that is the SNR returned by the harmonic sum algorithm. We distinguish recovered SNR from the injected SNR of the signal as the RSNR depends on other steps, chiefly on calculating the mean and standard deviation. Therefore the RSNR may deviate from the injected SNR of the signal. However,

all algorithms use the same pre-processing steps (mean and standard deviation calculation, power spectrum calculation, and interbinning). Thus the presented RSNR values are comparable between algorithms. The SNR is calculated using Eq. (2). Only the highest SNR value is returned for each element n .

Besides comparing absolute values of RSNR, we will also compare the sensitivity loss incurred when the pulsar frequency is between the centre of frequency bins and the dependency of the RSNR on pulsar frequency. For these comparisons, we will use artificial data produced using the von Mises distribution. Following this, we will compare all algorithms using synthetic data with white noise included. Lastly, we will test both algorithms integrated into the AstroAccelerate software package, where we will compare the result from real pulsar data from the GHRSS survey (Bhattacharyya et al., 2016) using the Giant Metrewave Radio Telescope (GMRT).

The performance of the harmonic sum algorithm will be measured in the number of processed fundamental frequency bins per second. Processing a fundamental frequency bin means constructing a partial sum, or sums, and returning the SNR value.

4.1. Artificial data without noise

To create an artificial pulsar, we have followed (Ransom et al., 2002) which used a modified von Mises distribution. The modified von Mises distribution is given by

$$f(\kappa, t) = a \frac{e^{\kappa \cos(2\pi f_r t + \varphi)} - e^{-\kappa}}{I_0(\kappa) - e^{-\kappa}}, \quad (10)$$

where a is amplitude, f_r is the pulsar's frequency, $0 \leq t \leq T$ is time, T is the time of the observation, φ is the phase, and I_0 is the modified Bessel function of zeroth order. The sampling time used was $t_s = 64 \mu\text{s}$. The shape parameter κ can be used to change the full width at half-maximum (FWHM), therefore the duty cycle of the pulsar which affects the number of significant harmonics present in the power spectra. More details can be found in Ransom et al. (2002). This way, we can create artificial pulsar data that can be used to test and evaluate the harmonic sum algorithm.

4.1.1. Sensitivity loss

The sensitivity loss comes from two sources that are connected. The first is the drift of the peak. The second is the scalloping loss. In order to have a non-zero drift, the true frequency of the pulsar must be outside of the centre of the frequency bin, which also means that the signal will suffer from scalloping loss.

To evaluate the sensitivity loss of a harmonic sum algorithm, we have created a set of time-series with increasing true pulsar frequency f_p from one centre of the frequency bin n to the next centre of the frequency bin $n + 1$. We have set the observation length to $T = 20$ s with sampling time $t_s = 64 \mu\text{s}$, thus the width of the frequency bin is $W_f = 0.05$ Hz. The frequency bin width was then divided into 50 steps. To determine the sensitivity loss of an algorithm, we have normalised the RSNR value returned by an algorithm, by the value of RSNR returned by the same algorithm at the centre of the frequency bin. That is, the sensitivity loss at the pulsar frequency f is

$$l(f) = \frac{\text{RSNR}(f)}{\text{RSNR}(f_0)}. \quad (11)$$

Two algorithms may have similar sensitivity loss but different RSNR and thus different probabilities of detecting a pulsar.

The value of the drift increases with pulsar frequency, and scalloping loss increases until the border of two bins is reached, where it starts to diminish up to the point of having no effect at the centre of the next frequency bin. If the tracking of the drift in a

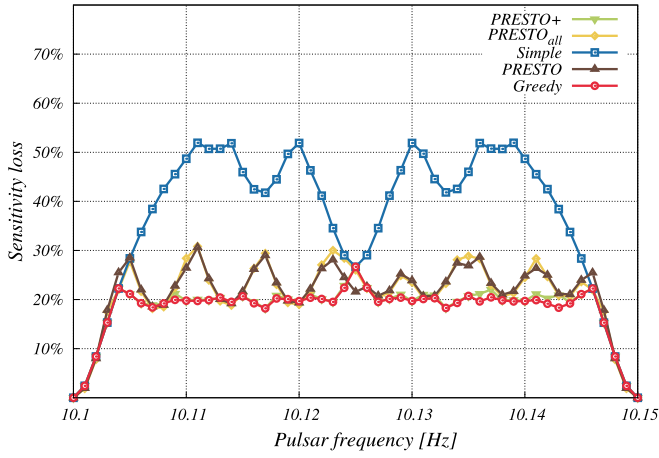


Fig. 9. Sensitivity loss of different harmonic sum algorithms.

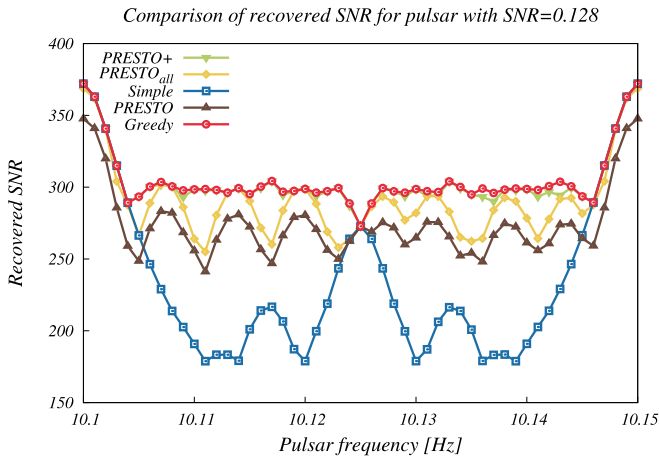


Fig. 10. Recovered SNR by different harmonic sum algorithms. Note that recovered SNR of the PRESTO algorithm is lower even at the centre of the frequency bin because PRESTO does not sum all higher harmonics.

harmonic sum is correct, we should see a symmetrical sensitivity loss around the border of the two bins. The comparison of the two versions of the PRESTO algorithms, simple harmonic sum and Greedy harmonic sum algorithm in sensitivity loss is shown in Fig. 9.

4.1.2. Recovered SNR

The RSNR returned by a harmonic sum algorithm depends also on the number of harmonic orders summed. One of the two differences between PRESTO, PRESTO_{all}, PRESTO+ and the Greedy harmonic sum algorithm is the number of harmonic orders summed. The recovered SNR for different algorithms is shown in Fig. 10. This is for $H = 32$.

To show the dependence of RSNR on the different fundamental frequencies and the maximum number of harmonics summed, we can use an average RSNR. To calculate the average RSNR, we calculate a mean RSNR from all RSNR returned by a harmonic sum algorithm as the pulsar frequency changes from one centre of the frequency bin to the next. That is, a mean RSNR of all RSNRs shown in Fig. 10 for a given algorithm represents one average RSNR value. The average RSNR for different fundamental frequency bins and maximum harmonics summed is shown in Fig. 11.

The average RSNR decreases from the fundamental frequency $f = 400$ Hz because of a poor representation of the pulsar

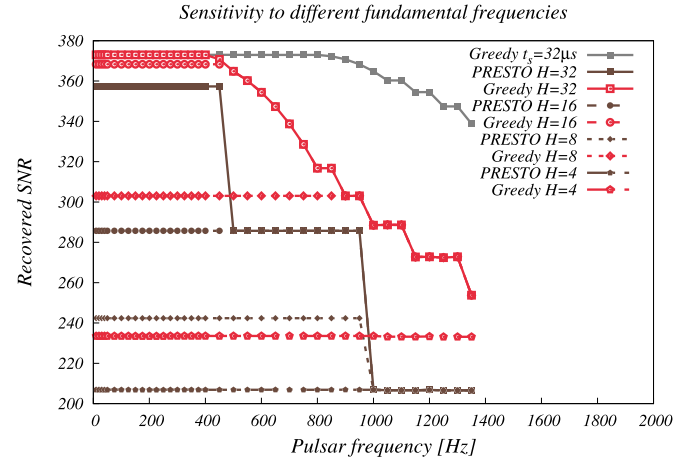


Fig. 11. Sensitivity loss of different fundamental frequencies and different maximum harmonic orders H .

Table 1

Recovered SNR by the AstroAccelerate for in-beam pulsars in GHRSS survey data.

Pulsar	Period (ms)	Harmonic sum		
		Greedy	PRESTO	PRESTO+
J0835–4510	89.39	12k	11k (–10%)	12k ()
J1328–4357	532.66	25	24 (–7%)	29 (+14%)
J1337–4441	1257.41	36	26 (–28%)	36 ()
J1517–4356	650.82	134	97 (–28%)	163 (+21%)
J1517–4356	650.78	53	42 (–27%)	55 (+4%)
J1542–5034	599.27	37	34 (–9%)	36 (–4%)
J1613–4714	382.34	209	163 (–22%)	209 ()
J1708–3426	692.16	25	21 (–15%)	24 (–8%)
J1823–3106	284.08	105	79 (–25%)	104 ()
J1817–3618	387.04	105	79 (–25%)	104 ()
J0418–4154	757.15	14	12 (–8%)	15 (+14%)

signal by the data. For sampling time $t_s = 64 \mu\text{s}$, the pulse width is below two samples wide. Thus, in addition to the other effects discussed above, we have introduced more signal loss by undersampling the pulsar signal. For comparison, we have added a time-series with sampling time $t_s = 32 \mu\text{s}$ that extends beyond $f = 400$ Hz and starts to decline at $f = 800$ Hz as expected. The artificial pulsar had about 30 significant harmonics and duty cycle of 4%.

4.2. Processing of observed data

We have implemented Greedy, PRESTO and PRESTO+ harmonic sum algorithms into a GPU accelerated software package called AstroAccelerate (Armour et al., 2020). AstroAccelerate is used for real-time data processing of time-domain radio astronomy data. We have tested and compared the performance of the greedy harmonic sum against PRESTO algorithms (PRESTO and PRESTO+) using real GHRSS survey observations with the GMRT. The GHRSS (GMRT High Resolution Southern Sky) survey is an off-Galactic plane survey for pulsars and fast transients at 300–500 MHz using the GMRT already having 26 discoveries from CPU (using PRESTO, RIPTIDE Morello et al., 2020) and GPU (using AstroAccelerate) processing. The GHRSS survey produces filterbank data with 4096 spectral channels sampled at $81.92 \mu\text{s}$ time-resolution. We have tested the harmonic sum algorithms using the in-beam pulsars detected in the GHRSS survey distributed over a range of periods and SNR. The results are shown in Table 1.

Table 2

Performance of the implemented harmonic sum algorithms measured in processed fundamental bins per second [Gbin/s].

H	Harmonic sum [Gbin/s]		
	Greedy	PRESTO	PRESTO+
32	8.64	9.64 (+11%)	1.16 (−87%)
16	14.70	15.24 (+ 4%)	3.79 (−74%)
8	26.10	23.38 (−10%)	11.03 (−58%)
4	39.9	37.49 (− 6%)	26.80 (−33%)

4.3. Compute performance

Compute performance was measured using an NVIDIA V100 GPU with 32 GB of GRAM and evaluated as the number of fundamental frequency bins processed per second, which is calculated as

$$[\text{Gbin/s}] = \frac{\text{Total number of frequency bins}}{\text{Execution time}}, \quad (12)$$

where the execution time was measured using the NVIDIA profiler. The tests were performed using 1024 power spectra, each 131 thousand elements long. All implemented harmonic sum algorithms are capable of searching through longer power spectra. The compute performance of all implemented harmonic sum algorithms is listed in Table 2.

5. Discussion

The PRESTO harmonic sum algorithm is deterministic, and it is oblivious to the data it processes. That is, it always sums the same elements (with respect to the fundamental bin) given the same input parameters. On the other hand, the greedy harmonic sum algorithm changes what elements to sum based on the data it is processing. The non-deterministic feature of the greedy harmonic sum may pose a problem.

Before discussing the performance and ability to recover a signal of any harmonic sum algorithm, we should discuss how the result is obtained, as this plays a vital role in signal extraction. Even a deterministic algorithm has to choose a figure of merit on which to judge whenever a particular partial sum of elements is better than any other. In both algorithms (PRESTO and greedy), this figure of merit is the value of SNR, where higher is better. This means that without any other knowledge, even the deterministic algorithm would select a partial sum that gives the highest SNR, even if it is the wrong one, over a partial sum that would contain pulsar's signal. Thus any harmonic sum algorithm could be misled by the noise. The advantage of the deterministic algorithm lies in that it does not abandon the partial sum prematurely.

A greedy approach can be misled by spurious noise into a different path from which it may not recover. As we will see, this may have happened in a few cases as some detections by PRESTO+ are higher than detections by the greedy harmonic sum despite the same signal loss and recovered SNR values in test with artificial data without noise.

The sensitivity loss is shown in Fig. 9. The highest signal loss occurs for the simple harmonic sum where the signal loss can be as high as 50%, which shows that tracking the drift is essential for recovering high RSNR.

The figure also shows three different PRESTO algorithms that differ in sensitivity loss and recovered SNR. Those algorithms are the PRESTO algorithm used in the PRESTO software package, which has the second highest signal loss. The PRESTO_{all} algorithm that is based on the PRESTO algorithm but sums all harmonics for given maximum harmonic order H instead of only even harmonics. And lastly, the PRESTO+ algorithm that adds all harmonics for given maximum harmonic order H and in addition performs the

harmonic sum where the maximum harmonic order is increasing by one (from 1 to H_{\max}) instead of performing the harmonic sum only for H that is a power-of-two.

We see that PRESTO_{all} has almost the same sensitivity loss as the PRESTO algorithm, but it has lower compute performance; hence we have not included it in further comparisons. The PRESTO+ algorithm has the same sensitivity loss as the greedy algorithm.

The recovered SNR is shown in Fig. 10, where we see that it mostly follows the sensitivity loss behaviours. The PRESTO algorithm has lower recovered SNR than the rest of the algorithms because it does not sum all harmonics and does not explore all partial sums in the way PRESTO+ does. We also see that for data without noise, the greedy harmonic sum and PRESTO+ harmonic sum algorithms have the same RSNR.

The results for in-beam pulsars from the GHRSS survey in the presence of real telescope noise, including RFI show that the Greedy harmonic sum always detected higher RSNR than the PRESTO harmonic sum. However, there are cases where PRESTO+ has higher RSNR. In these cases, the greedy algorithm may have followed the wrong path due to noise. However, the greedy algorithm provides up to 28% higher RSNR values than the PRESTO harmonic sum while offering a similar compute performance.

The performance of all implemented harmonic sum algorithms is shown in Table 2, where higher is better. We see that the greedy harmonic sum has a similar compute performance to the PRESTO harmonic sum algorithm. The greedy harmonic sum is slower by 11% at maximum harmonic order search $H = 32$ and 6% faster at $H = 4$. The PRESTO+ algorithm is the slowest. Scaling in the number of power spectra is linear for all implemented harmonic sum algorithms. Algorithms also scale linearly in length of the power spectra. However, each algorithm scales differently with maximum harmonic orders H searched. The greedy harmonic sum scales linearly $\mathcal{O}(H)$ as an increase in the maximum harmonic order by one means increasing the number of steps taken by the algorithm by one. The PRESTO+ algorithm scales as $\mathcal{O}(H^2)$, which is a disadvantage especially prominent in the transition from $H = 16$ to $H = 32$. The PRESTO algorithm scales as $\mathcal{O}(H \log H)$ because the outer loop only goes through H , which are powers of two.

6. Conclusions

We have demonstrated that our new greedy harmonic sum algorithm has lower signal loss and better recovered signal-to-noise ratio (SNR) than the standard algorithm used in the PRESTO software package while achieving similar performance in the number of processed fundamental frequency bins per second.

The signal loss for the pulsar frequencies that fall outside of the centre of the frequency bins is lower by as much as 10 percentage points. The greedy harmonic sum algorithm recovers on average 20% higher SNR than the PRESTO harmonic sum algorithm. The range of recovered SNR was from 7% up to 28% higher for the greedy harmonic sum algorithm compared to the PRESTO harmonic sum algorithm.

In terms of the performance, the greedy harmonic sum algorithm is 6% faster than the PRESTO harmonic sum algorithm for the maximum harmonic order search $H = 4$ and $H = 8$ and 11% slower for harmonic order $H = 32$.

We have also compared our greedy harmonic sum to an improved PRESTO harmonic sum algorithm that achieves the same signal loss values and recovered SNR. However, the greedy harmonic sum algorithm is 7× faster for $H = 32$ and 1.4× faster for $H = 4$.

Due to the nature of greedy algorithms, RFI may pose a major problem if not mitigated. The RFI can cause the greedy harmonic

sum to stray and seek higher harmonics in different bins than their true location. However, as the number of frequency bins searched per fundamental frequency bin is constrained to a limited set, the error caused by the RFI is also constrained to a given fundamental frequency. Moreover, tests on the real data from the GHRSS survey showed that even in the presence of RFI the greedy harmonic sum correctly detects all signals.

CRedit authorship contribution statement

K. Adámek: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing – original draft, Writing – review & editing. **J. Roy:** Data curation, Writing – review & editing. **W. Armour:** Conceptualization, Methodology, Supervision, Project administration, Funding acquisition, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors acknowledge the support of the OP VVV MEYS funded project CZ.02.1.01/0.0/0.0/16_019/0000765 “Research Center for Informatics”. This work has received support from an STFC Grant (ST/T000570/1, ST/W001969/1 and ST/R000557/1). The authors would also like to acknowledge the use of the University of Oxford’s Advanced Research Computing (ARC) (Richards, 2015) facility in carrying out this work. We acknowledge support of GMRT telescope operators for observations. The GMRT is run by the National Centre for Radio Astrophysics of the Tata Institute of Fundamental Research, India.

References

- Adámek, K., Armour, W., 2020. Single-pulse detection algorithms for real-time fast radio burst searches using GPUs. *Astrophys. J.* 247, 56. doi:10.3847/1538-4365/ab7994, arXiv:1910.08324.
- Adámek, K., Novotný, J., Dimoudi, S., Armour, W., 2020. Searching for pulsars in extreme orbits - GPU acceleration of the Fourier domain ‘Jerk’ search. In: Pizzo, R., Deul, E.R., Mol, J.D., de Plaa, J., Verkouter, H. (Eds.), *Astronomical Society of the Pacific Conference Series*. ISBN: 978-1-58381-941-8, p. 671, arXiv:1911.01353.
- Agarwal, D., Lorimer, D.R., Surnis, M.P., Pei, X., Karastergiou, A., Golpayegani, G., Werthimer, D., Cobb, J., McLaughlin, M.A., White, S., Armour, W., MacMahon, D.H.E., Siemion, A.P.V., Foster, G., 2020. Initial results from a real-time FRB search with the GBT. *Mon. Not. R. Astron. Soc.* 497, 352–360, arXiv:2003.14272.
- Andersen, B.C., Ransom, S.M., 2018. A Fourier domain Jerk search for binary pulsars. *Astrophys. J. Lett.* 863, L13, arXiv:1807.07900.
- Armour, W., Adámek, K., Novotný, J., Dimoudi, S., Carels, C., Quannoughi, N., Giles, M., Karastergiou, A., Williams, C., Roberts, S., Brown, A., Chennamangalam, J., Curran, D., Smith, M.S., Nunez-Yanez, J., Mickalliger, M., Stappers, B., Thiagaraj, P., Roy, J., Bassa, C., Hessels, J., Clark, K., Lanfear, T., Sinclair, B., Ling, A., Casselman, S., Sharpe, J., 2020. Astroaccelerate. doi:10.5281/zenodo.4282748.
- Armour, W., Karastergiou, A., Giles, M., Williams, C., Magro, A., Zagkouris, K., Roberts, S., Salvini, S., Dulwich, F., Mort, B., 2012. A GPU-based survey for millisecond radio transients using ARTEMIS. In: Ballester, P., Egret, D., Lorente, N.P.F. (Eds.), *ADASS XXI*. ISBN: 978-1-58381-804-6, p. 33, arXiv:1111.6399.
- Bhattacharyya, B., Cooper, S., Malenta, M., Roy, J., Chengalur, J., Keith, M., Kudale, S., McLaughlin, M., Ransom, S.M., Ray, P.S., Stappers, B.W., 2016. The GMRT high resolution southern sky survey for pulsars and transients. I. Survey description and initial discoveries. *Astron. J.* 817, 130. doi:10.3847/0004-637X/817/2/130, arXiv:1509.07177.
- Dimoudi, S., Adámek, K., Thiagaraj, P., Ransom, S.M., Karastergiou, A., Armour, W., 2018. A GPU implementation of the correlation technique for real-time Fourier domain pulsar acceleration searches. *Astrophys. J.* 239, 28. doi:10.3847/1538-4365/aabe88, arXiv:1804.05335.
- Hewish, A., Bell, S.J., Pilkington, J.D.H., Scott, P.F., Collins, R.A., 1968. Observation of a rapidly pulsating radio source. *Nature* 217, 709–713. doi:10.1038/217709a0.
- Lorimer, D., 1999. Sigproc. <https://sourceforge.net/projects/sigproc>.
- Lorimer, D.R., Kramer, M., 2004. *Handbook of Pulsar Astronomy*, Vol. 4.
- Lyons, R.G., 2011. *Understanding Digital Signal Processing*, third ed. Prentice Hall.
- Morello, V., Barr, E.D., Stappers, B.W., Keane, E.F., Lyne, A.G., 2020. Optimal periodicity searching: revisiting the fast folding algorithm for large-scale pulsar surveys. *Mon. Not. R. Astron. Soc.* 497, 4654–4671, arXiv:2004.03701.
- Morello, V., Rajwade, K.M., Stappers, B.W., 2021. IQRM: real-time adaptive RFI masking for radio transient and pulsar searches. arXiv e-prints, arXiv:2108.12434.
- Rajwade, K., Stappers, B., Williams, C., Barr, E., Christiaan Bezuidenhout, M., Caleb, M., Driessen, L., Jankowski, F., Malenta, M., Morello, V., Sanidas, S., Surnis, M., 2021. MeerTRAP in the era of multi-messenger astrophysics. arXiv e-prints, arXiv:2103.08410.
- Ransom, S.M., 2002. Presto. <https://github.com/scottransom/presto>.
- Ransom, S.M., Eikenberry, S.S., Middleditch, J., 2002. Fourier techniques for very long astrophysical time-series analysis. *Astron. J.* 124, 1788–1809, arXiv:astro-ph/0204349.
- Richards, A., 2015. *Advanced Research Computing*. University of Oxford, doi:10.5281/zenodo.22558.
- Taylor, J.H., Huguenin, G.R., 1969. Two new pulsating radio sources. *Nature* 221, 816–817. doi:10.1038/221816a0.
- Yu, M., 2018. Harmonic summing improves pulsar detection sensitivity: A probability analysis. *Astrophys. J.* 868, 8. doi:10.3847/1538-4357/aae51a.