High Resolution Data Analysis: Plans and Prospects

Shriram Jois, Leanne Duffy, Neil Sullivan, David Tanner, and William Wester

Abstract A report on the progress on the high resolution data analysis of the ADMX experimental results is presented. In this paper, tools are developed and tested on a blind injection mimicking a Maxwellian like signal in the frequency domain. This blind injection will be used as a test bed which can be later implemented on all the high resolution data. The high resolution data is stored in the Fermilab server. In this analysis a PostgreSQL query was made to ensure the blind injection is in the middle of the frequency spectrum and 19 such files were found. The time series data is read using a c++ program. An apodization function is applied on the time series data and zero filled to reduce the frequency spacing in order to achieve a better interpolation. A FFTW header is used to compute the Fourier transform of the time series data. A Savitzky–Golay filter is applied on the unnormalized power which then can be used to remove the spectral shape. Each frequency spectrum has a bandwidth of 50 kHz.

Keywords ADMX · Axion · Dark matter · Sine wave injection · Synthetic axion injection · High resolution data analysis · Apodization functions · Spectral shape · Savitzky–Golay filter · Exponential distribution

S. Jois (✉) · N. Sullivan · D. Tanner
Department of Physics, University of Florida, Gainesville, FL, USA
e-mail: ramjois@ufl.edu; sullivan@phys.ufl.edu; tanner@phys.ufl.edu

L. Duffy
Los Alamos National Laboratory, Los Alamos, NM, USA
e-mail: ldd@lanl.gov

W. Wester
Fermi National Accelerator Laboratory, Batavia, IL, USA
e-mail: wester@fnal.gov

G. Carosi, G. Rybka (eds.), *Microwave Cavities and Detectors for Axion Research*, Springer Proceedings in Physics 245, https://doi.org/10.1007/978-3-030-43761-9_11

1 Introduction

The Axion Dark Matter eXperiment (ADMX) uses a cylindrical RF cavity in a high magnetic field and a superconducting amplifier to search for a hypothetical particle called the axion, which if it exists is a solution to the strong CP problem in QCD [1] and a possible cold dark matter candidate [2]. During the experiment the axion decays into photons in the presence of a high magnetic field in a process called the Primakoff conversion. Since the mass of axion is unknown, to detect these photons the RF cavity is tuned to a frequency that matches the total energy of the axion and resonantly enhances the conversion process ($f_a = \frac{m_a c^2}{h}$). Since the galactic halo axions are non-relativistic, the total energy of axion is given by $E_a = m_a c^2 + \frac{1}{2} m_a v^2$. The signal undergoes a diurnal and annual modulation due to the Earth's rotation and revolution around the Sun. The frequency modulation is given by $\frac{\delta f}{f_a} = \frac{v_a |\delta v|}{c^2}$. The frequency shift has been calculated by Ling et al. in 2004 [3]. The shift in a 100s time interval due to the Earth's orbital motion Δf_o is $\mathcal{O}(10^{-12} f_0)$ and for the Earth's rotation Δf_r is $\mathcal{O}(10^{-11} f_0)$. The high resolution channel has a frequency resolution of 0.01 s and will be able to detect these frequency modulation when the axions are found, whereas the velocity dispersion of axions in the medium resolution channel is $\mathcal{O}(10^{-4} c)$ [4]. To test and develop a search pattern, a sine wave injection or blind injection can be used. This has been demonstrated by the advanced LIGO group to validate the discovery of gravitational waves [5].

2 Fast Fourier Transform Techniques

The time series gets truncated due to the finite number of points. The instrument function $I(\omega)$ of the time series data $f(t)$ is defined by the exponential Fourier transformation

$$I(\omega) = \frac{1}{2\pi} \int_{\frac{T}{2}}^{\frac{T}{2}} dt \;\; f(t) a(t) e^{-i\omega t}. \tag{1}$$

If the time series data, $f(t)$ contains a periodic signal, e.g. a sine wave, the Fourier transform of $f(t)\,a(t)$ will be a convolution of $F(\omega)$ and $A(\omega)$. In the absence of an apodization function, $a(t) = 1$. The Fourier transform of the rectangular apodization is, $A(\omega) = \frac{1}{2\pi} \frac{\sin\left(\omega \frac{T}{2}\right)}{\omega \frac{T}{2}}$. Because the $F(\omega)$ is a delta function, the convolution of $F(\omega)$ and $A(\omega)$ will be a sinc function. To minimize the sidelobes in the sinc function, an apodization function is often used in the spectral analysis. In a high resolution data analysis, the spectrum is dominated by noise, hence minimizing the side lobes is not required. The frequency resolution of the spectrum is set by the measurement time which for this case is 0.01 Hz. However a better interpolation is achieved by adding zeros to the time series [6]. If f(n) is the time series and f(k) its Fourier transform

$$f(k) = \sum_{n=0}^{N-1} f(n)e^{-i\frac{2\pi kn}{N+M}} + \cancelto{0}{\sum_{n=N}^{M+N-1} f(n)e^{-i\frac{2\pi kn}{N+M}}} = \sum_{n=0}^{N-1} f(n)e^{-i\frac{2\pi kn}{N+M}}. \tag{2}$$

Here N is the number of points in the time series and M is the zeros added to the time series. The frequency resolution of the original time series was $\Omega = \frac{2\pi}{N}$. This got reduced to $\Omega_e = \frac{2\pi}{N+M}$.

The data analysis can proceed in two steps. The first is to find the candidates. Once the candidates are found, it can be zero filled to reduce the frequency spacing which results in a better interpolation. A peak detection algorithm is likely to find the center frequency of the axion signal more accurately in a zero filled spectrum than the original frequency spectrum due to smaller frequency spacing. A compromise needs to made since a higher order zero will fill up the disk space faster than a lower order zero fill (Fig. 1).

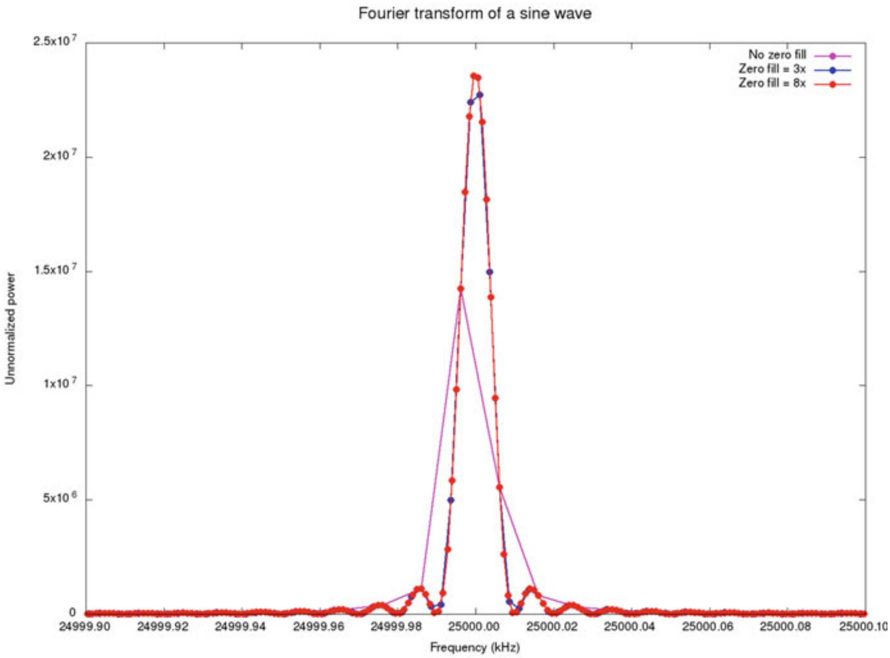

Fig. 1 This figure shows the Fourier transform of a sine wave with a 25 kHz frequency with various degrees of zero fill. The peak of the signal is closer to 25 kHz on a zero filled spectrum than the original spectrum

3 Power Spectrum and Distribution

The time series data has a time spacing of $\Delta t = 10.24\,\mu s$. Each data file contains a total of 9 million data points. The zero fill is done as a multiple of the total number of data points. Making the total number of points a power of two will make the program run faster. Once the apodization function is chosen and the amount of zero fill is decided, a Fourier transform of the time series is computed (Fig. 2). In this paper, a FFTW package was used on a c++ program to compute the Fourier transform [7]. The Fourier frequency is computed as $f_F = \frac{1}{k\Delta t}$ and k can have values from 0 to $\frac{n-1}{2}$, where n is the total number of data points in the time series. The start frequency of the spectrum is stored in the time series file which can be read and added to the Fourier frequency to get the actual frequency of the spectrum. The complex Fourier transform can be expressed as sine and cosine transforms. The sine and cosine transform of the noise in the time series follows a Gaussian distribution with two random variables in the frequency domain. The distribution function is given by

$$\frac{dP}{dA} = \frac{1}{2\pi\sigma^2} e^{\frac{-x^2}{2\sigma^2}} e^{\frac{-y^2}{2\sigma^2}}, \tag{3}$$

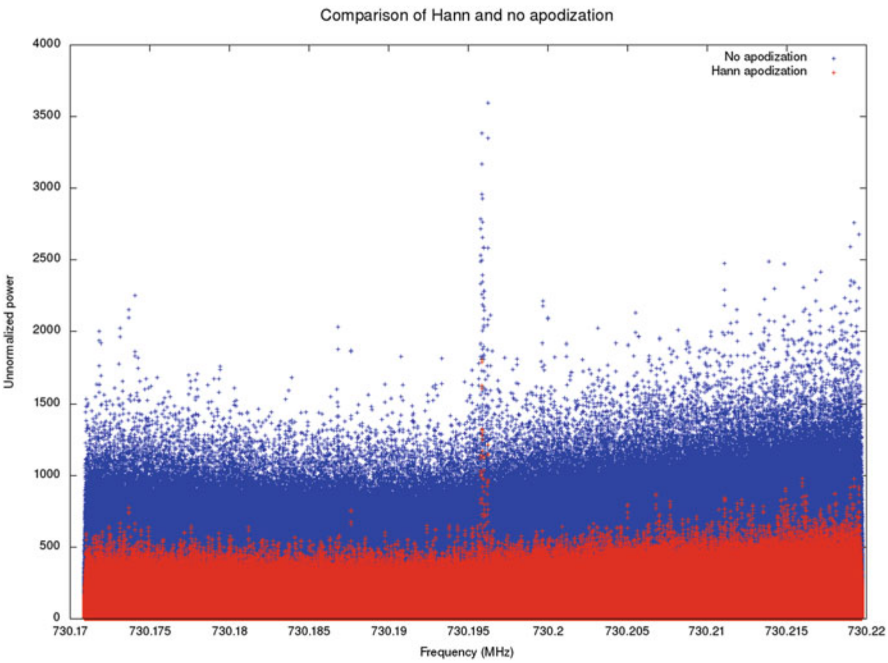

Fig. 2 Comparison of a Hann apodization and a no-apodization frequency spectrum with background noise and a blind axion injection

converting x and y to polar coordinates r and θ. The θ integral can be integrated to get 2π. This gives us the Rayleigh distribution

$$P = \frac{1}{2\pi\sigma^2}\int_0^{\infty} r\,dr \int_0^{2\pi} d\theta\, e^{\frac{-r^2}{2\sigma^2}} = \frac{1}{\sigma^2}\int_0^{\infty} r\,dr\, e^{\frac{-r^2}{2\sigma^2}}. \tag{4}$$

Since we are interested in the power due to Primakoff conversion, the amplitude can be squared to find the rms power. The resulting probability distribution function is an exponential distribution

$$\frac{dP}{dp} = \frac{1}{\sigma^2} e^{\frac{-p}{\sigma^2}}. \tag{5}$$

Taking the logarithm of the above equation, we find

$$\log(P(x)) = -\frac{p}{\sigma^2} + c. \tag{6}$$

This is of the form $y = mx + c$, a straight line with a negative slope (Fig. 3). A generalized distribution function for n-bin data has been derived by L. Duffy et. al [8].

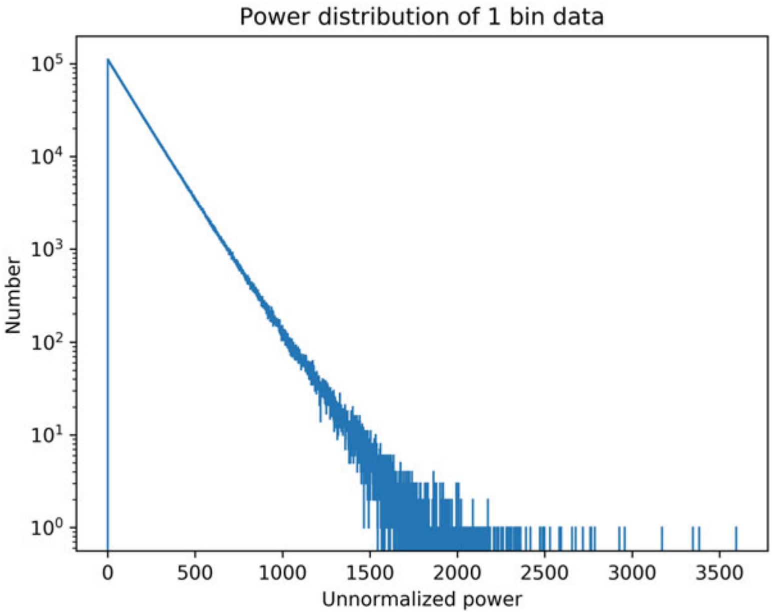

Fig. 3 Histogram of unfiltered noise spectrum on a semilog scale which follows Eq. (7). The scattering at the end of the line is due to the blind axion injection. A spectrum with a real axion signal is expected to have a similar distribution

4 Spectral Shape of the Noise

The spectral shape of the noise (Fig. 2) refers to the amplitude dependence of the noise as a function of the frequency and is due to the systematic effects that are introduced in the receiver chain. These systematic errors should be corrected before establishing a noise baseline. In the previous high resolution searches [4, 8], the spectral shape was removed by dividing the frequency data by a polynomial of order n as a function of frequency offset using the equivalent circuit model which estimates the rms value of noise power at the NRAO output. The spectral shape can be removed even without the equivalent circuit model by a textbook implementation of Savitzky–Golay digital filter [9] (Fig. 4). Due to the number of points in the time series, the window size of such a filter should be large to achieve a smaller standard deviation in the filtered data. With the use of high performance computing techniques, the Savitzky–Golay filters can be employed for the removal of spectral shape. The Savitzky–Golay filter has two variables, window size and order. The filtered data is a convolution of Savitzky–Golay coefficients and the unnormalized data and can be written as [10]

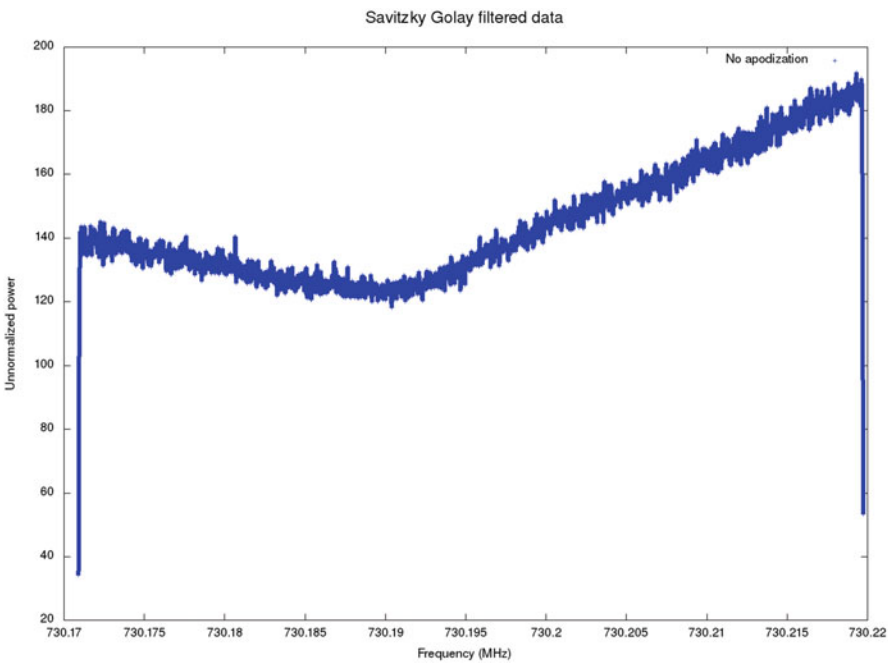

Fig. 4 Plot of the frequency spectrum after the application of a Savitzky–Golay filter on the noise spectrum. The Savitzky–Golay has preserved the spectral shape of the noise distribution which can be used to flatten the noise spectrum

$$P_j^{SG}(f) = \sum_{i=-\frac{M-1}{2}}^{\frac{M-1}{2}} C_i P_{i+j}(f). \qquad (7)$$

Here, the window size of the filter is M and C_i are the normalized Savitzky–Golay convolution coefficients. The coefficients can be calculated by forming a least square polynomial of order n and using the Moore–Penrose inverse to solve a set of M+1 linear equations. The sum of all coefficients is one. The sharp dip at the ends of the Savitzky–Golay filtered spectrum is due to the Gibbs phenomenon. Since the Savitzky–Golay filter requires convolution, the data needs to be extended at the start and the end. This determines the length of the dip at the ends of the spectrum. One way to remove the dip is to use Gegenbauer polynomials [11]. However, in this paper it has been removed by chopping off the ends. The data that has been lost can be covered in the next spectrum.

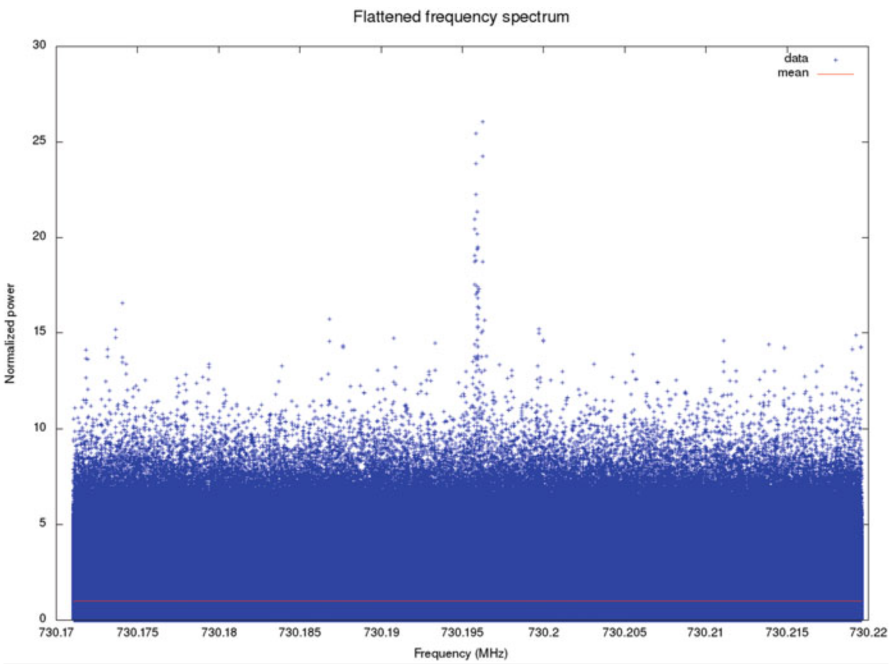

Fig. 5 Plot of the flattened frequency spectrum after the removal of the spectral shape. The Savitzky–Golay has preserved the spectral shape of the noise distribution which was used to flatten the frequency spectrum. The mean is plotted in red and is equal to 1

5 Future Work

Since the window size of the Savitzky–Golay filter is chosen such that the mean of the filtered data is equal to the unnormalized power, the mean of the normalized spectrum is equal to 1. This can be seen in Fig. 5. To improve the S/N, the frequency data can be averaged by sub-dividing the time series data and zero filling the subdivided time series and averaging all the frequency spectra. This is similar to what the spectrum analyzer does. Once the apodization function is chosen and the time series is subdivided and zero filled, the exclusion limits can be set. The search pattern that is developed on the blind injection can be used in the analysis of the data obtained in the experimental search for axions.

Acknowledgments This work was supported by the Department of Energy Grant No. DE-181SC0010296 at the University of Florida.

References

1. R.D. Peccei, H.R. Quinn, CP conservation in the presence of pseudoparticles. Phys. Rev. Lett. **38**(25), 1440 (1977)
2. J.R. Primack, B. Seckel, B. Sadoulet, Detection of cosmic dark matter. Annu. Rev. Nucl. Part. Sci. **38**(1), 757–807 (1988)
3. F.-S. Ling, P. Sikivie, S. Wick, Diurnal and annual modulation of cold dark matter signals. Phys. Rev. D **70**, 123503 (2004)
4. J. Hoskins, N. Crisosto, J. Gleason, P. Sikivie, I. Stern, N.S. Sullivan, D.B. Tanner, C. Boutan, M. Hotz, R. Khatiwada, D. Lyapustin, Modulation sensitive search for nonvirialized dark-matter axions. Phys. Rev. D **94**(8), 082001 (2016)
5. C. Biwer, D. Barker, J.C. Batch, J. Betzwieser, R.P. Fisher, E. Goetz, S. Kandhasamy, S. Karki, J.S. Kissel, A.P. Lundgren, D.M. Macleod, Validating gravitational-wave detections: the advanced LIGO hardware injection system. Phys. Rev. D **95**(6), 062002 (2017)
6. M.L. Forman, Spectral interpolation: zero fill or convolution. Appl. Opt. **16**(11), 2801–2801 (1977)
7. M. Frigo, S.G. Johnson, *The design and implementation of FFTW3*. Proceedings of the IEEE 93.2 216–231 (2005).
8. L.D. Duffy, P. Sikivie, D.B. Tanner, S.J. Asztalos, C. Hagmann, D. Kinion, L.J. Rosenberg, K. Van Bibber, D.B. Yu, R.F. Bradley, High resolution search for dark-matter axions. Phys. Rev. D **74**(1), 012006 (2006)
9. A. Savitzky, M.J. Golay, Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. **36**(8), 1627–1639 (1964). Vancouver
10. S.J. Orfanidis, Introduction to signal processing (Prentice-Hall, Inc, New York, 1995)
11. D. Gottlieb, C.W. Shu, On the Gibbs phenomenon and its resolution. SIAM Rev. 39(4), 644–668 (1997)