FAST MUSIC – AN EFFICIENT IMPLEMENTATION OF THE MUSIC ALGORITHM FOR FREQUENCY ESTIMATION OF APPROXIMATELY PERIODIC SIGNALS

Orchisama Das, Jonathan S. Abel, Julius O. Smith III

Center for Computer Research in Music and Acoustics,
Stanford University
Stanford, USA
[orchi|abel|jos]@ccrma.stanford.edu

ABSTRACT

Noise subspace methods are popular for estimating the parameters of complex sinusoids in the presence of uncorrelated noise and has applications in musical instrument modeling and microphone array processing. One such algorithm, MUSIC (Multiple Signal Classification) has been popular for its ability to resolve closely spaced sinusoids. However, the computational efficiency of MUSIC is relatively low, since it requires an explicit eigenvalue decomposition of an autocorrelation matrix, followed by a linear search over a large space. In this paper, we discuss methods for and the benefits of converting the Toeplitz structure of the autocorrelation matrix to *circulant* form, so that the eigenvalue decomposition can be replaced by a Fast Fourier Transform (FFT) of one row of the matrix. This transformation requires modeling the signal as at least approximately periodic over some duration. Moreover, we derive a closed-form expression for finding spectral peaks, yielding large savings in computation time. We test our algorithm to resolve closely spaced piano partials and observe that it can work with a much shorter data window than a regular FFT.

1. INTRODUCTION

Sinusoidal parameter estimation is a classical problem with applications in radar, sonar, music, and speech, among others. When the frequencies of sinusoids are well resolved, looking for spectral peaks is adequate. It is shown in [1] that the maximum likelihood (ML) frequency estimate for a single sinusoid in Gaussian white noise is given by the frequency of the magnitude peak in the periodogram. The ML approach is extended and Cramer-Rao bounds are derived for multiple sinusoids in noise in a follow-on paper by the same authors [2]. Some other estimators are covered in [3].

For closely spaced sinusoidal frequencies, however, other approaches have been developed. Noise subspace methods are a class of sinusoidal parameter estimators that utilize the fact that the noise subspace of the measured signal is orthogonal to the signal subspace. Pisarenko Harmonic Decomposition [4] makes use of the eigenvector associated with the minimum eigenvalue of the estimated autocorrelation matrix to find frequencies. However, it has been found to exhibit relatively poor accuracy [3]. Schmidt [5] improved over Pisarenko with the MUSIC (MUltiple SIgnal Classification) algorithm which could estimate the frequencies of multiple closely spaced signals more accurately in the presence of noise. In MUSIC, a pseudospectrum is generated by projecting a complex sinusoid onto all of the noise subspace eigenvectors, defining peaks where this projection magnitude is minimum. This method is shown to be asymptotically unbiased. An enhancement to MUSIC, root-MUSIC, was proposed in [6]. It uses the properties of the signal-space eigenvectors to define a rational spectrum with poles and zeros. It is said to have better resolution than MUSIC at low SNRs. Similarly, another popular algorithm, ESPRIT [7], was invented by Roy et al. which makes use of the underlying rotational invariance of the signal subspace. The generalized eigenvalues of the matrix pencil formed by an auto-covariance matrix and a cross-covariance matrix gives the unknown frequencies. ESPRIT performs better than MUSIC, especially when the signal is sampled nonuniformly. More recently, another enhancement to MUSIC, gold-MUSIC [8] has been proposed which uses two stages for coarse and fine search, respectively.

One of the disadvantages of the MUSIC algorithm is its computational complexity. Typical eigenvalue decomposition algorithms are of the order $O(N^3)$ [9]. In this paper, we use the fact that, for periodic signals, the autocorrelation matrix is *circulant* when it spans an integer multiple of the signal's period. In this case, looking for the eigenvalues with largest magnitude is equivalent to looking for magnitude peaks in the FFT. We know in the circulant case that our noise eigenvectors are DFT sinusoids, and hence we can derive a closed-form solution when we project our search space onto the noise subspace, thereby reducing further the calculations required to find the pseudospectrum. Replacing eigenvalue decomposition in MUSIC with efficient Fourier transform based methods has been previously studied in [10] where the eigenvectors are derived to be some linear combinations of the data vectors, while maintaining the orthonormality constraint. In this paper, we take a different approach by utilizing the circulancy of autocorrelation matrices.

We test our algorithm to resolve closely spaced partials of the A3 note played on a piano. It is a well known fact that the strings corresponding to a particular piano key are slightly mistuned. Coupled motion of piano strings has been studied in detail by Weinreich in [11, 12]. There is a slight difference in frequency of the individual strings, giving rise to closely spaced peaks in the spectra. A very long window of data is needed to resolve these peaks with the FFT. We show that FAST MUSIC can resolve two closely spaced peaks with a much shorter window of data than a regular FFT, and much faster than MUSIC.

The rest of this paper is organized as follows : Section 2 gives an outline of the MUSIC algorithm, Section 3 derives the FAST MUSIC algorithm, Section 4 describes the experimental results on a) an artificially synthesized signal containing two sinusoids with additive white noise and b) a partial of the A3 note played on the piano which contains beating frequencies. We conclude the paper in Section 5 and delineate the scope for future work. Estimation of real-valued sine wave frequencies with MUSIC has been studied in [13]. For all derivations in this paper, we work with real signals, because we are interested in audio applications. This saves some time in computing the pseudospectrum since we know it will be symmetric. Our derivations can be easily extended to complex signals.

2. MUSIC

2.1. Model

We wish to estimate the parameters of a signal composed of additive sinusoids from noisy observations. Let $y(n)$ be the noisy signal, composed of a deterministic part, $x(n)$, made of r real sinusoids and random noise, $w(n)$. We assume that $w(n) \sim N(0, \sigma^2)$, and that $w(n)$ and $x(n)$ are uncorrelated. The sinusoidal phases ϕ_i's are assumed to be i.i.d. and uniformly distributed $\phi_i \sim U(-\pi, \pi)$.

$$\begin{aligned} y(n) &= \sum_{i=1}^{r} A_i \cos(\omega_i n + \phi_i) + w(n) \\ y(n) &= x(n) + w(n) \end{aligned}$$

In vector notation, the signal $\mathbf{y} \in \mathbb{R}^M$ can be characterized by the $M \times M$ autocorrelation matrix $K_y = \mathbb{E}(\mathbf{y}\mathbf{y}^T)$. For a zero-mean signal, the autocorrelation matrix coincides with the covariance matrix. Since this matrix is Toeplitz and symmetric positive-definite, its eigenvalues are real and nonnegative (and positive when $\sigma > 0$). We can perform an eigenvalue decomposition on this matrix to get a diagonal matrix Λ consisting of the eigenvalues, and an eigenvector matrix Q. The $2r$ eigenvectors corresponding to the $2r$ largest eigenvalues, Q_x, contain signal plus noise information, whereas the remaining *M-2r* eigenvectors, Q_w, only represent the noise subspace. Thus, we have the following relationships:

$$\begin{aligned} &K_y = K_x + K_w = K_x + \sigma^2 I \\ &K_y = Q \Lambda Q^H \\ &K_y = \begin{bmatrix} Q_x & Q_w \end{bmatrix} \begin{bmatrix} \Lambda' & 0 \\ 0 & \sigma^2 I_{M-2r} \end{bmatrix} \begin{bmatrix} Q_x^H \\ Q_w^H \end{bmatrix} \end{aligned} \tag{1}$$

2.2. Pseudospectrum Estimation

Let a vector of M harmonic frequencies be denoted as $\mathbf{b}(\omega) = [1, e^{j\omega}, e^{2j\omega} \cdots e^{(M-1)j\omega}]^T$. We project this vector onto Q_w, i.e., the subspace occupied by the noise (where there is no signal component). We find the following spectrum as a function of a set of ω's:

$$\begin{aligned} P(\omega) &= \frac{1}{\mathbf{b}(\omega)^H Q_w Q_w^H \mathbf{b}(\omega)} \\ P(\omega) &= \frac{1}{||Q_w^H \mathbf{b}(\omega)||^2} \end{aligned} \tag{2}$$

For a particular value of ω that is actually present in the signal, the sum of projections of **b** onto the eigenvectors spanning the noise subspace will be zero. This is because the subspace occupied by the signal is orthogonal to that occupied by noise since they are uncorrelated. Thus, we see that $P(\omega)$ will take on a very high value in such cases. In conclusion, we can find peaks in the function $P(\omega)$ and those will correspond to our estimated frequencies. Since the search space can consist of any number of densely packed frequencies, very closely spaced peaks can show up in the pseudospectrum, which a simple FFT might not be able to resolve in the presence of noise. However, as the search-space grows, so does computational complexity.

3. FAST MUSIC

Under conditions to be specified, it is possible to replace the eigenvalue decomposition required in MUSIC by a Fast Fourier Transform (FFT). In this section, we derive the structure and order of the autocorrelation matrix for which it is circulant instead of only Toeplitz. We also derive a closed-form expression for finding the pseudospectrum.

3.1. Deriving the autocorrelation matrix

In vector form, $\mathbf{y} \in \mathbb{R}^M$ can be written as:

$$\begin{aligned} \mathbf{y} &= S\mathbf{a} + \mathbf{w} \\ \mathbf{w} &\sim N(0, \sigma^2 I) \end{aligned} \tag{3}$$

$$\begin{bmatrix} y(n) \\ y(n-1) \\ \vdots \\ y(n-M+1) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots \\ \cos(\omega_1) & \sin(\omega_1) & \cdots \\ \vdots & \vdots & \cdots \\ \cos[(M-1)\omega_1] & \sin[(M-1)\omega_1] & \cdots \end{bmatrix} \times \begin{bmatrix} A_1 \cos(\omega_1 n + \phi_1) \\ A_1 \sin(\omega_1 n + \phi_2) \\ \vdots \\ A_r \sin(\omega_r n + \phi_r) \end{bmatrix} + \begin{bmatrix} w(n) \\ w(n-1) \\ \vdots \\ w(n-M+1) \end{bmatrix} \tag{4}$$

Since $\mathbf{y}$ is zero-mean, its covariance matrix is

$$K_y = \mathbb{E}(\mathbf{y}\mathbf{y}^T) = S K_a S^T + \sigma^2 I. \tag{5}$$

We now want to get K_y in terms of K_a. We have assumed $\phi_i \sim U(-\pi, \pi)$ (uniformly identically distributed random phase). We observe that every term of K_a is of the form $K_a(i,j) = \mathbb{E}[A_i \cos(\omega_i n + \phi_i)\, A_j \cos(\omega_j n + \phi_j)]$, or $\mathbb{E}[A_i \sin(\omega_i n + \phi_i) A_j \sin(\omega_j n + \phi_j)]$, or $\mathbb{E}[A_i \sin(\omega_i n + \phi_i) A_j \cos(\omega_j n + \phi_j)]$, or $\mathbb{E}[A_i \cos(\omega_i n + \phi_i) A_j \sin(\omega_j n + \phi_j)]$. All of these terms are zero, except the first two when $i = j$, i.e. $\mathbb{E}[A_i^2 \cos(\omega_i n + \phi_i)^2] = \mathbb{E}[A_i^2 \sin(\omega_i n + \phi_i)^2] = \frac{A_i^2}{2}$, which makes it a diagonal matrix:

$$K_a = \begin{bmatrix} \frac{A_1^2}{2} & 0 & \cdots & 0 & 0 \\ 0 & \frac{A_1^2}{2} & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \frac{A_r^2}{2} \end{bmatrix} \tag{6}$$

The autocorrelation matrix of the observed signal is given in (7). This is an $M \times M$ real, symmetric Toeplitz matrix, but with the right order M it can become circulant.

3.2. Circulancy of the autocorrelation matrix

We have seen that the autocorrelation matrix K_y is symmetric Toeplitz. However it is to be noted that for $k = 1, 2, ...$, if M is an integer such that $M = 2\pi n / \omega_i, n \in \mathbb{Z}^+$, then

$$\sum_{i=1}^{r} \frac{A_i^2}{2} \cos(M-k)\omega_i = \sum_{i=1}^{r} \frac{A_i^2}{2} \cos k\omega_i \tag{8}$$

$$SK_aS^T + \sigma^2 I = \begin{bmatrix} \sum_i \frac{A_i^2}{2} + \sigma^2 & \sum_i \frac{A_i^2}{2}\cos\omega_i & \sum_i \frac{A_i^2}{2}\cos 2\omega_i & \cdots & \sum_i \frac{A_i^2}{2}\cos(M-1)\omega_i \\ \sum_i \frac{A_i^2}{2}\cos\omega_i & \sum_i \frac{A_i^2}{2} + \sigma^2 & \sum_i \frac{A_i^2}{2}\cos\omega_i & \cdots & \sum_i \frac{A_i^2}{2}\cos(M-2)\omega_i \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \sum_i \frac{A_i^2}{2}\cos(M-1)\omega_i & \sum_i \frac{A_i^2}{2}\cos(M-2)\omega_i & \cdots & \cdots & \sum_i \frac{A_i^2}{2} + \sigma^2 \end{bmatrix} \tag{7}$$

If we choose M carefully, then the autocorrelation matrix may be written as :

$$\begin{bmatrix} \sum_i \frac{A_i^2}{2} + \sigma^2 & \sum_i \frac{A_i^2}{2}\cos\omega_i & \cdots & \sum_i \frac{A_i^2}{2}\cos\omega_i \\ \sum_i \frac{A_i^2}{2}\cos\omega_i & \sum_i \frac{A_i^2}{2} + \sigma^2 & \cdots & \sum_i \frac{A_i^2}{2}\cos 2\omega_i \\ \vdots & \vdots & \cdots & \vdots \\ \sum_i \frac{A_i^2}{2}\cos\omega_i & \sum_i \frac{A_i^2}{2}\cos 2\omega_i & \cdots & \sum_i \frac{A_i^2}{2} + \sigma^2 \end{bmatrix} \tag{9}$$

This matrix is circulant! Hence, its eigenvectors are given by the DFT sinusoids and its eigenvalues are the DFT coefficients of the first row [14]. The eigenvalues can be computed using an FFT algorithm when M is a power of 2 or highly composite. The relationship between eigenvalues of Toeplitz matrices and those of asymptotically equivalent circulant matrices have been studied in [15].

For example, if the signal consists of 3 sinusoids with frequencies $\frac{\pi}{2}$, $\frac{\pi}{4}$ and $\frac{\pi}{5}$ then the minimum order of M which will make the autocorrelation matrix circulant is given by $2\times LCM(2,4,5) = 40$. However, if any of the denominators is irrational, then the LCM does not exist, and hence no value of M will make the autocorrelation matrix circulant. Of course, in reality we do not know the frequencies and cannot determine M this way. However, we can instead detect when the autocorrelation corresponds to a signal that is periodic. If we can find the periodicity of the estimated autocorrelation function and set M to be that period, then the resulting autocorrelation matrix will be circulant. Since no signal is truly precisely periodic, this procedure can be viewed as introducing an approximation based on assuming the signal is periodic. Such a periodic/harmonic approximation is common when the underlying signal source is known to be a quasi periodic oscillator such in voiced speech, bowed strings, woodwinds, flutes, brasses, organs, and so on. In this paper, we use the Average Magnitude Difference Function [16] to detect the period. We find all local minima in the AMDF and pick the period as the lowest minimum index which is smaller than its adjacent neighbors.

3.3. Searching over a large range of frequencies

Suppose we want to calculate the pseudospectrum for N distinct frequencies $\omega_k = 2\pi\frac{k}{N}$ for $k = -\frac{N}{2}, \ldots, \frac{N}{2} - 1$ covering the range $[-\pi, \pi)$. In that case, $\mathbf{b}(k) = [1, e^{\frac{2\pi jk}{N}}, e^{\frac{4\pi jk}{N}} \ldots e^{\frac{2\pi(M-1)jk}{N}}]^T$. Since the eigenvectors spanning the noise subspace are also complex exponentials, the projection function can be simplified. The matrix Q_w is composed of columns of noise eigenvectors, such that each column is denoted by $\mathbf{q}_{w_{n_i}} = \frac{1}{\sqrt{M}}[1, e^{\frac{2\pi jn_i}{M}}, e^{\frac{4\pi jn_i}{M}}, \cdots, e^{\frac{2\pi(M-1)jn_i}{M}}]^T$ where n_i are the indices of the noise eigenvectors for $i = 1, \ldots, L$ and $L = M - 2r$.

$$\begin{aligned} Q_w^H\mathbf{b}(k) &= \frac{1}{\sqrt{M}} \begin{bmatrix} \sum_{m=0}^{M-1} \exp[2\pi jm(\frac{k}{N} - \frac{n_1}{M})] \\ \vdots \\ \sum_{m=0}^{M-1} \exp[2\pi jm(\frac{k}{N} - \frac{n_L}{M})] \end{bmatrix} \\ &= \frac{1}{\sqrt{M}} \begin{bmatrix} e^{-\pi j(\frac{k}{N} - \frac{n_1}{M})(M-1)} \frac{\sin[\pi(\frac{k}{N} - \frac{n_1}{M})M]}{\sin[\pi(\frac{k}{N} - \frac{n_1}{M})]} \\ \vdots \\ e^{-\pi j(\frac{k}{N} - \frac{n_L}{M})(M-1)} \frac{\sin[\pi(\frac{k}{N} - \frac{n_L}{M})M]}{\sin[\pi(\frac{k}{N} - \frac{n_L}{M})]} \end{bmatrix} \end{aligned} \tag{10}$$

From the pseudospectrum estimate, it can be seen that, whenever $\frac{k}{N} = \frac{n_i}{M}$, one term in the denominator sum dominates, and we can evaluate $P(k)$ using L'Hospital's rule.

$$\begin{aligned} P(k) &= \frac{M}{\sum_{i=1}^{M-2r} \left| \frac{\sin[\pi(\frac{k}{N} - \frac{n_i}{M})M]}{\sin[\pi(\frac{k}{N} - \frac{n_i}{M})]} \right|^2} \\ &\approx \frac{M}{\sum_{i=1}^{M-2r} M^2} \quad if \quad \frac{k}{N} = \frac{n_i}{M} \\ &\approx \frac{1}{M(M-2r)} \quad if \quad \frac{k}{N} = \frac{n_i}{M} \end{aligned} \tag{11}$$

4. EXPERIMENTS AND RESULTS

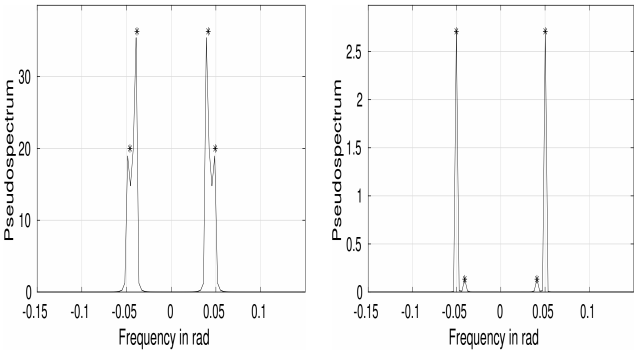
4.1. Synthesized signal

To compare FAST MUSIC with MUSIC, we tested a cosine signal composed of frequencies 0.04rad and 0.05rad at $f_s = 1$ Hz.

$$y(n) = \cos(0.05n) + 0.5\cos(0.04n + \phi) + w(n) \tag{12}$$

To detect periodicity of the autocorrelation function, we need at least two periods of the signal. For closely spaced sinusoids, the period is likely a large number, and hence more samples of the signal are needed to estimate the autocorrelation function. In this example, the length of the signal is 2000. The pseudospectra of MUSIC and FAST MUSIC are given in Figure 1.

To measure computation time, we compared various eigenvalue decomposition algorithms with Fast Fourier Transform algorithms for increasing orders of the autocorrelation matrix. The results can be seen in Figure 2. QR factorization is used to find eigenvalues and eigenvectors. QR factorization with Gram Schmidt orthogonalization is slow, symmetric tridiagonal QR with implicit Wilkinson shift is slightly faster whereas reduction to the Hessenberg form is fastest. More details about these algorithms can be found in [9]. The Fourier transform algorithms are orders of magnitude faster, with the DFT dominating at lower orders and self-sorting mixed radix FFT [17] and resampled split radix FFT [18] giving faster speeds at higher orders. It is to be noted that the order of the autocorrelation matrix for which circulancy is achieved



(a) MUSIC (b) FAST MUSIC

Figure 1: Pseudospectrum of signal with frequencies 0.04 rad and 0.05 rad.

is not likely to be a power of 2, and hence we cannot use the well-known radix-2 FFT algorithm. However, we can first resample the data to a power of 2 [19] and then apply a radix-2/split radix FFT algorithm or used a mixed-radix FFT algorithm on any composite order. All of these functions have been implemented by the authors in MATLAB. More efficient implementations can be done in C, where the FFT functions should overtake the DFT at much lower orders.

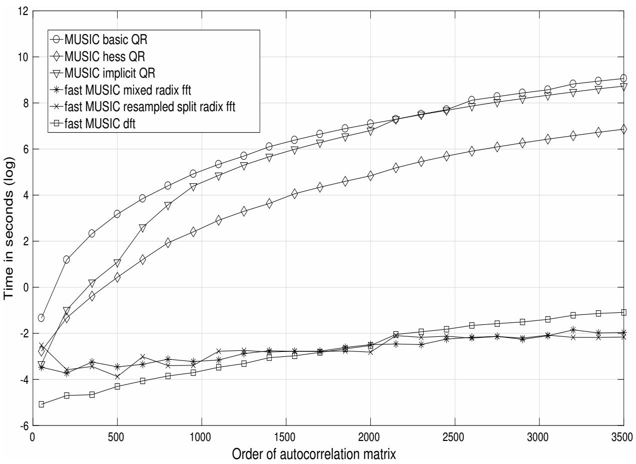
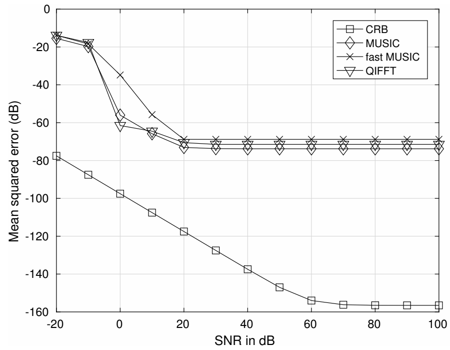
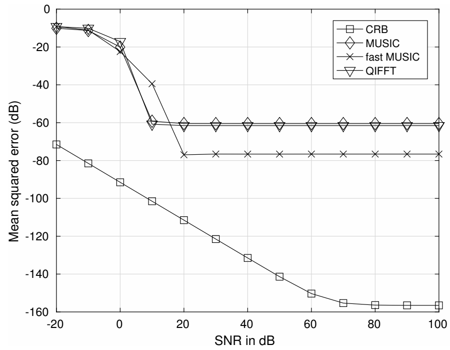


Figure 2: Order of matrix vs time in seconds (log)

We also conducted 1000 Monte-Carlo simulations on the above example, with uniformly distributed random phase and plotted the mean-squared errors vs SNR for MUSIC, FAST MUSIC and QIFFT [20], along with the Cramer-Rao bounds (CRB) as given in [21]. The plots can be seen in Figure 3. The performances are nearly equivalent in Figure 3a, with QIFFT performing the best and FAST MUSIC performing the worst, but FAST MUSIC performs significantly better than the others in Figure 3b for high SNRs. FAST MUSIC performs poorly for SNRs below 20 dB.



(a) Frequency = 0.04rad



(b) Frequency = 0.05rad

Figure 3: MSE vs SNR plots

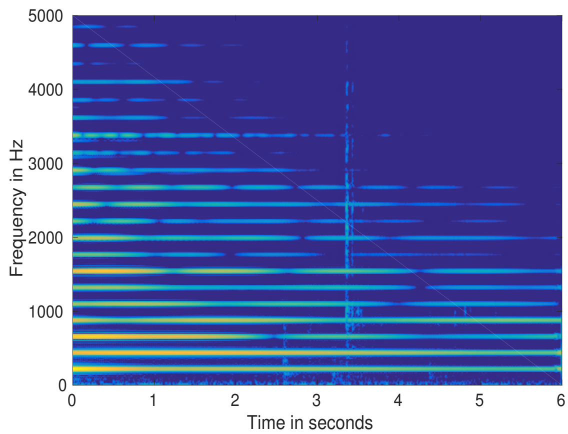
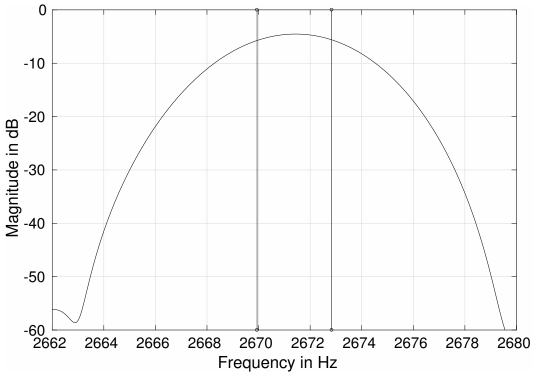


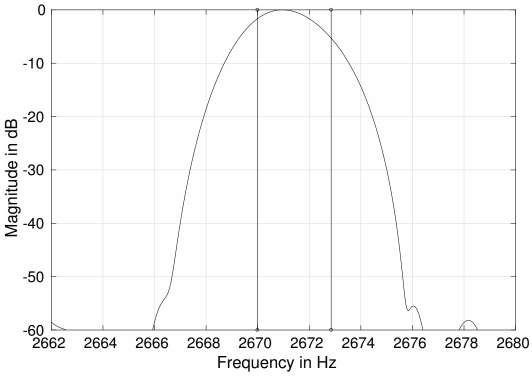
Figure 4: Spectrogram of piano note

4.2. Piano data

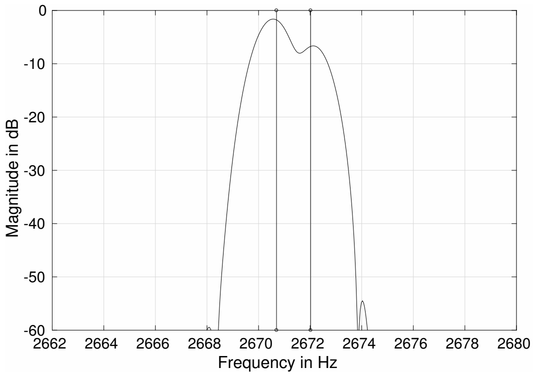
We tested our algorithm on the A3 note played on the piano. The spectrogram of the steady state portion of the note is given in Figure 4. We can observe beating in some of the partials. We decided to work with the 11th partial, located close to 2600 Hz, where a beating of roughly 1 Hz is observed. We bandpass-filtered the signal using a 4th-order Butterworth filter with cut-off frequencies at 2400 Hz and 2900 Hz. We ran FFTs with the Blackman



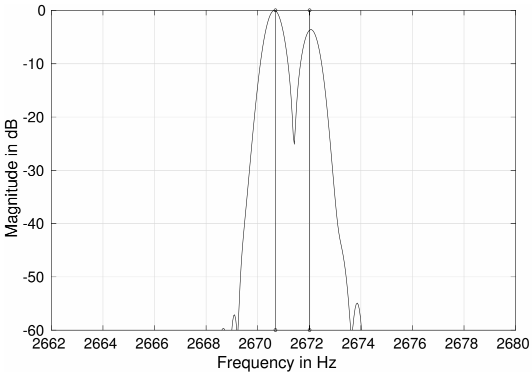
(a) Data length= 2^{14}



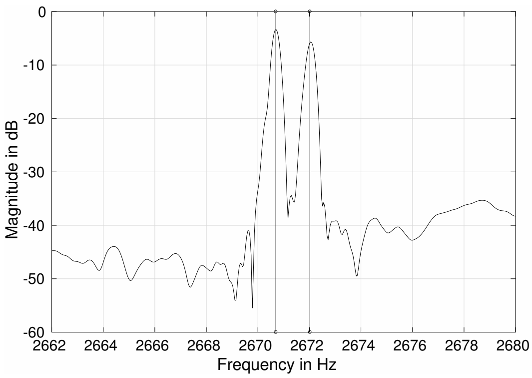
(b) Data length= 2^{15}



(c) Data length= 2^{16}



(d) Data length= 2^{17}



(e) Data length= 2^{18}

Figure 5: FFT magnitude plots and FAST MUSIC frequency estimates (vertical lines)

window and FAST MUSIC on different data lengths, as shown in Figure 5, where the vertical lines indicate the frequencies detected by FAST MUSIC. We see that for window sizes 2^{14} and 2^{15}, the FFT magnitude does not exhibit two separately discernible peaks at all, whereas FAST MUSIC provides two peak frequencies with some error. For larger window sizes, both FFT and FAST MUSIC are able to resolve the two peaks with greater accuracy. FAST MUSIC may be preferred over the FFT for higher resolution when the available data size is limited. One potential application is in piano tuning, where FAST MUSIC could be used to quickly resolve closely spaced peaks caused by the coupled motion of the piano strings.

5. DISCUSSION AND FUTURE WORK

In this paper, we have proposed a computationally efficient implementation of the MUSIC algorithm. The autocorrelation matrix was derived and approximated by a circulant matrix. This approximation allowed us to replace computationally intensive eigenvalue decomposition algorithms with an FFT. We subsequently derived a closed-form expression for searching over a range of frequencies. These modifications yielded a significant improvement in computational speed.

A key factor in the accuracy of FAST MUSIC is the precision in periodicity detection. If the periodicity is off by a significant number of samples, the autocorrelation matrix is no longer circulant and FAST MUSIC falls apart. AMDF based periodicity detector is time consuming, not foolproof and often yields wrong results if the number of lags in the autocorrelation function is very high. Ideally, a better method for periodicity detection should be used. However, the aim of this paper is not to come up with a novel solution for detecting the period of a signal, so we leave this problem open for future work.

Another issue is finding the number of sinusoids present in a given signal, if it is not known apriori. To do so, one can look at the relative magnitude of the eigenvalues (FFT peak values in our case). This works well if the signal to noise ratio is high. More robust partitioning schemes have been used in [8]. Once the signal frequencies are known, the estimation of amplitudes is trivial and can be done using linear least squares.

We found that the estimator variances for both FAST-MUSIC and the QIFFT method were dominated by bias at high SNRs. Bias can be reduced arbitrarily in the QIFFT method by increasing the amount of zero-padding, as well as by other methods [22], and we expect it to reduce similarly in FAST-MUSIC when the upsampling factor is increased to higher powers of 2 and when the number of points in the search space is increased. Future work should reduce or eliminate the bias so that the relative performance can be observed at high SNRs.

Note that both FAST-MUSIC and QIFFT are effectively FFT-magnitude based methods. They differ only in how they extract the peak frequency from the FFT magnitude, and there are slight differences in the FFT magnitude computation itself (equivalently, slightly different definitions of sample autocorrelation). Both methods can be regarded as approximate maximum-likelihood methods, so the question becomes which method gives the best average accuracy for a given computational complexity? This remains a topic for future work.

We would also like to add a third method for comparison which is a *true* maximum likelihood estimate, obtainable by finding the optimal window-transform overlap that explains the observed spectrum, under the hypothesis of two sinusoidal peaks being present. Such a method is of course much more expensive than either of the FFT methods considered here, but it should represent the highest accuracy obtainable under our signal modeling assumptions. In the end, we would like a well filled out family of approximate maximum likelihood estimators for closely spaced sinusoidal peaks, providing a range of ways for trading off estimation accuracy versus computational expense.

6. REFERENCES

[1] David C. Rife and Robert R. Boorstyn, "Single-Tone Parameter Estimation from Discrete-Time Observations," *IEEE Transactions on Information Theory*, vol. 20, no. 5, pp. 591–598, 1974.

[2] David C Rife and Robert R Boorstyn, "Multiple tone parameter estimation from discrete-time observations," *Bell Labs Technical Journal*, vol. 55, no. 9, pp. 1389–1410, 1976.

[3] Steven M Kay, "Modern spectral estimation," chapter 13. Prentice Hall, 1988.

[4] Vladilen F Pisarenko, "The retrieval of harmonics from a covariance function," *Geophysical Journal International*, vol. 33, no. 3, pp. 347–366, 1973.

[5] Ralph Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE transactions on antennas and propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[6] Arthur Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'83.* IEEE, 1983, vol. 8, pp. 336–339.

[7] Richard Roy and Thomas Kailath, "Esprit-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on acoustics, speech, and signal processing*, vol. 37, no. 7, pp. 984–995, 1989.

[8] Kaluri V Rangarao and Shridhar Venkatanarasimhan, "gold-music: A variation on music to accurately determine peaks of the spectrum," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 4, pp. 2263–2268, 2013.

[9] Peter Arbenz, "Lecture notes on solving large scale eigenvalue problems," chapter 4. ETH Zurich.

[10] Juha T Karhunen and Jyrki Joutsensalo, "Sinusoidal frequency estimation by signal subspace approximation," *IEEE Transactions on signal processing*, vol. 40, no. 12, pp. 2961–2972, 1992.

[11] Gabriel Weinreich, "Coupled piano strings," *The Journal of the Acoustical Society of America*, vol. 62, no. 6, pp. 1474–1484, 1977.

[12] Gabriel Weinreich, "The coupled motions of piano strings," *Scientific American*, vol. 240, no. 1, pp. 118–127, 1979.

[13] Petre Stoica and Anders Eriksson, "Music estimation of real-valued sine-wave frequencies," *Signal processing*, vol. 42, no. 2, pp. 139–146, 1995.

[14] Robert M Gray et al., "Toeplitz and circulant matrices: A review," *Foundations and Trends® in Communications and Information Theory*, vol. 2, no. 3, pp. 155–239, 2006.

[15] Robert Gray, "On the asymptotic eigenvalue distribution of toeplitz matrices," *IEEE Transactions on Information Theory*, vol. 18, no. 6, pp. 725–730, 1972.

[16] Myron Ross, Harry Shaffer, Andrew Cohen, Richard Freudberg, and Harold Manley, "Average magnitude difference function pitch extractor," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 22, no. 5, pp. 353–362, 1974.

[17] Clive Temperton, "Self-sorting mixed-radix fast Fourier transforms," *Journal of Computational Physics*, vol. 52, no. 1, pp. 1–23, 1983.

[18] Pierre Duhamel, "Implementation of Split-Radix FFT Algorithms for Complex, Real, and Real-Symmetric Data," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, pp. 285–295, 1986.

[19] Julius O. Smith, "Digital Audio Resampling Home Page," Tech. Rep., Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, 2002.

[20] Mototsugu Abe and Julius O. Smith, "Design criteria for simple sinusoidal parameter estimation based on quadratic interpolation of FFT magnitude peaks," in *Audio Engineering Society Conventions and Conferences (AES'04)*, 2004, vol. 117.

[21] Petre Stoica and Arye Nehorai, "Music, maximum likelihood, and cramer-rao bound," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 5, pp. 720–741, 1989.

[22] Kurt James Werner and François Georges Germain, "Sinusoidal parameter estimation using quadratic interpolation around power-scaled magnitude spectrum peaks," *Applied Sciences*, vol. 6, no. 10, pp. 306, 2016.