Sub-Nyquist Sampling and Reconstruction Schemes

- Novel techniques to minimize data consumption for variable-bandwidth signals -

Project Report

Liam Mulshine & Casey Meehan

Harvard University
AM205 | Numerical Methods
December 18, 2017

Contents

Chapter 1

Introduction and Background

Sampling is a widely employed technique for storing information about a continuous time signal. It is used in nearly every electronic system that interacts with the outside world. Consider, for example, systems that perform digital audio recording. In such systems, samples from the analog signal generated by a microphone are collected and stored sequentially in memory. Assuming the samples are taken at a sufficient rate, the original continuous signal is roughly represented by its discrete samples. However, if, during the audio recording, useful signals grow sparse, it is unnecessary to continue sampling rapidly for accurate reconstruction. In this paper, we describe a novel sampling and reconstruction technique that addresses this issue exactly. Explicitly, we develop a method to determine the optimal sampling rate of a sparse, band-limited signal in real-time. Since most sampling techniques use a constant sampling rate, they are constrained to sample above what is known as the Nyquist rate. The method that we propose varies the sampling rate based on the signal's instantaneous bandwidth, effectively reducing the number of samples required to accurately reconstruct the original signal.

This project was inspired by our class discussions on interpolation, and the effects of both the distribution and density of samples on interpolation accuracy, efficiency and stability. While we discuss how the interpolation methods that we covered in class apply to the exact problem that we are trying to solve, we predominantly pull from mathematical concepts in a separate research area. Specifically, we utilize foundational concepts and mathematical techniques from the field of Digital Signal Processing (DSP). This field is focused on the relationship between real-time signals and their frequency domain equivalents. In the following section we will provide a brief background on a few important concepts in digital signal processing. Then we will enter a discussion on previous work related to the fields of sampling and reconstruction. Finally, we will provide an outline of the sampling and reconstruction technique that we employ.

1.1 Frequency Domain and Bandlimited Signals

Nyquist, Shannon, and Gabor, pioneers in the field of communication theory, showed that a signal can be represented as a linear combination of shifted, variable frequency sinusoids. A real signal can be represented in the frequency domain, by the magnitude of the component frequencies that make it up. For example, consider a widely used signal in electronics, the square wave. The square wave can be constructed as a sum of weighted sinusoids at odd multiples of the cycle frequency, f.

$$x_{square}(t) = \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{sin(2\pi(2k-1)ft)}{2k-1} \tag{1.1}$$

The signal can be represented in the frequency domain as

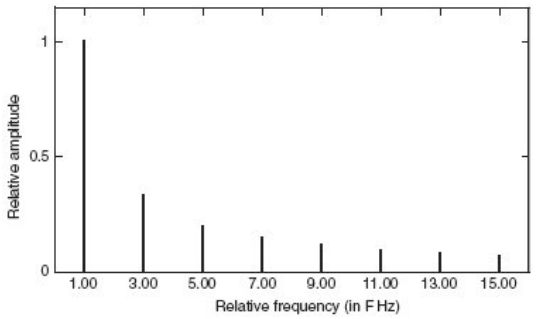

Figure 1.1: Frequency Domain Representation of Square Wave

This plot depicts the magnitude of each frequency component within the square wave signal. Since the square wave is composed only of odd multiples of the fundamental, all other frequencies have zeros magnitude.

The frequency domain representation of a signal is useful for determining the signal's bandwidth. Bandwidth is defined as the maximum frequency for which the spectral density (units W / Hz), exceeds some threshold. To put this explicitly, suppose the spectral distribution, H, of a signal is known, providing information about the signal's power as a function of frequency. We can define the relative spectrum as the original spectral distribution, divided by the value of the maximum spectral component. In signal processing, it is common to represent this value in decibels

$$\text{Relative Power (dB)} = 20 \log_{10} \left(\frac{H}{\max(H)} \right) \tag{1.2}$$

Suppose that we set the threshold for Relative Power to -20 dB, corresponding to a relative amplitude of $\frac{1}{10}$. In the case of the square wave, the maximum frequency component that exceeds this relative power is at nine times the fundamental frequency, $9f$, corresponding to the the fifth component of the series square wave representation. Since all frequency components above $9f$ have a lower spectral density,

$9f$ is, by definition, the signal bandwidth. In a real system, this bandwidth condition would be enforced by filtering out all frequency components above the specified bandwidth.

1.2 Reconstruction from Uniform Samples

In classical sampling theory, reconstruction of a bandlimited signal from its uniform samples is achieved through discrete-time convolution. Consider a signal sampled on a uniform grid with spacing, T. Provided the sample rate exceeds the Nyquist rate, (i.e. $\frac{1}{T} > 2\Omega$, where Ω is the signal bandwidth), the original signal, $x(t)$ can be reconstructed perfectly from its samples, x_n, using [[7]]

$$\sum_{n=-\infty}^{+\infty} x_n \frac{\Omega T}{\pi} \frac{sin(\Omega(t-nT))}{\Omega(t-nT)} \tag{1.3}$$

This expression originates from the earlier work of Whittaker as a general interpolation method. Nyquist and Gabor then demonstrate that $2\Omega T$ samples from a signal of length T and bandwidth W are sufficient for perfect reconstruction using (1.1). In [[7]], Shannon relates the independent work of these individuals to communication theory, proving (1.1) and presenting the "Sampling Theorem," laying the foundation for much of classical sampling theory.

1.3 Reconstruction from Nonuniform Samples

After presenting the Sampling Theorem, Shannon briefly mentions that the $2\Omega T$ required samples need not be uniformly spaced to guarantee perfect reconstruction, so long as at least $2\Omega T$ samples are taken. However, he states that under conditions of non-uniformity, (1.1) no longer applies. In fact, Shannon claims that as the sampling grid grows increasingly non-uniform, bandlimited reconstruction becomes both more complex, and less numerically stable. While Shannon does not explicitly define a method for signal reconstruction from non-uniform samples, after only a short time, J. L. Yen filled this void with his publication, "On Nonuniform Sampling of Bandwidth-Limited Signals" [[8]].

Most research on reconstruction from non-uniform samples stems from Yen's work. In his paper, Yen considers signal reconstruction from four distinct classes of non-uniform sampling: (1) a finite number of samples deviate from a uniform sampling grid, (2) the uniform sampling grid shifts at one instance in time, (3) a recurrent non-uniform sampling scheme, and (4) arbitrary sampling. Each class has practical significance. However, each must also satisfy the constraint that the average sampling rate over the entire signal is equal to or exceeds the Nyquist rate. As an example of what this entails, suppose you have a band-limited signal of length, t, with bandwidth, Ω. Assuming uniform sampling at the Nyquist rate, a total of $2\Omega t$ samples will be collected. While each of Yen's non-uniform sampling classes permit

deviation from the uniform sampling grid, each ultimately requires $2\Omega t$ samples to be collected within the time interval, t.

Therefore, while allowing for a more flexible sampling grid, the non-uniform sampling and reconstruction methods that Shannon hints at and Yen explores in depth still enforce the $2\Omega T$ lower bound on the number of samples required for perfect reconstruction. Since our intention is to reduce the number of samples required to reconstruct a real-time, band-limited signal below the average Nyquist threshold, we will not satisfy Yen's constraints. To achieve this sampling reduction while preserving reconstruction integrity, a new sampling and reconstruction technique must be devised. We propose that by varying the sampling rate based on the instantaneous signal bandwidth, near-perfect reconstruction of the original signal is possible using a piece-wise sinc-interpolation, assuming the sampling rate at any point in time exceeds $2W_i t$, where W_i is the instantaneous bandwidth. This follows directly from Shannon's sampling theorem.

1.4 Choosing Nonuniform Samples

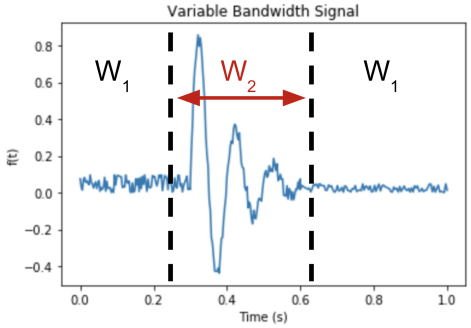

Figure 1.2: Example of Variable Bandwidth Signal

Consider a time-limited signal like that shown in **Figure 1.2** with time-dependent bandwidth. After anti-alias filtering, the bandwidth of the first and last segments will be significantly lower than that of the middle segment. By Shannon's theory, the density of samples required to reconstruct the full signal is twice the maximum bandwidth. As such, reconstruction of this signal would require a sample density of $2W_2$ in all three segments. However, if one were to sample and reconstruct the three segments independently, the sample density of the first and third segments would only need to be $2W_1$ by Shannon's theory. This piecewise reconstruction strategy would be ideal by Shannon's standards within each of the three segments, and minimize the amount of sampling required. Simply stated, signals with temporally changing bandwidths may be sampled more economically if the local bandwidth can be resolved.

The practice of describing a signal by its local frequency characteristics has a rich

history in time-frequency distributions [[1]]. In this field, a time signal, $s(t)$ is transformed from one to two dimensions, $p(\omega, t)$, in which every time value corresponds to a unique local frequency distribution. A time-frequency distribution is most simply captured by taking a *short-time Fourier transform* (STFT):

$$p(\omega, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} w(\tau - t) s(t + \tau) e^{-j\omega\tau} d\tau$$

where $w(t)$ is a 'window function'. There are a variety of window functions used in different applications, the details of which are beyond the scope of this discussion. Most window functions are even to avoid adding group delay to the transformed signal. In the STFT, the widow maximally weights the signal value at time, t and supresses content far away from t. In practice, the window function is considered 0 when below a certain threshold, w_{th}. This allows the STFT to be a finite integral:

$$p(\omega, t) = \frac{1}{2\pi} \int_{-\text{tlim}}^{\text{tlim}} w(\tau - t) s(t + \tau) e^{-j\omega\tau} d\tau$$

Where $\text{tlim} = w^{-1}(w_{th})$.

There is a wealth of research behind time-frequency distributions ([[1]], [[6]]) and behind reconstruction from nonuniform samples ([[2]], [[4]]). However, we have not found investigations leveraging time-frequency distributions to optimally *select* nonuniform samples in real-time. The objective of this project is to – in real-time – localize the required sample rate of variable bandwidth signals. As the example in **Figure 1.2** samples according the segment's bandwidth, we consider sampling according to the signal's instantaneous bandwidth. We approximate the instantaneous bandwidth by applying a STFT to the signal's Taylor expansion. We restrict our available mathematical operations such that instantaneous bandwidth may be determined with analog electronic computation. We then apply a variety of reconstruction schemes to the nonuniform sampled signal to consider the reconstruction error relative to a Nyquist sampled version, and discuss the trade-off between sample density (data volume) and error (signal fidelity).

Chapter 2

Sampling Techniques

2.1 Instantaneous Bandwidth

To resolve the instantaneous bandwidth of the signal, we apply a STFT to the Taylor expansion of the signal in real time. Fortunately, differentiation of arbitrary order and addition are trivial for analog circuity, so the Taylor series expansion, which consists of a linear combination of the the signal's derivatives, may be implemented using analog hardware, allowing our technique to be run in real-time.

The Gabor Transform [[6]] is a classic STFT. By using a Gaussian window (maximal entropy), assumptions made about the signal being processed are minimized. To avoid infinite integration, we truncate the Gaussian window for all values less than $w_{th} = 0.001$, such that our transform becomes:

$$G(\omega, t) = \int_{-\text{tlim}}^{\text{tlim}} \sqrt{\frac{\alpha}{\pi}} e^{-\alpha(\tau - t)^2} f(t + \tau) e^{-j\omega\tau} d\tau$$

To generalize this formula for signals of arbitrary bandwidth, we normalize the integrated time domain such that the frequency in radians, ω is π at the signal bandwidth, W. In other words, provided a signal with bandwidth, W, the proposed normalization can be carried out by scaling the time axis by $\beta = \frac{W}{\pi}$. A visual depiction of this transformation is provided in **Figure 2.1**, **Figure 2.2** and **Figure 2.3**. This technique, of course, requires that the signal's maximum bandwidth is known. Fortunately, the bandwidth can be set and easily enforced with simple analog circuitry, a common practice in applications of sampling and reconstruction. We then choose our window to span two periods of the signal's would-be Nyquist grid. This normalization can be performed with a simple change of variables:

Consider the signal centered about the current time, t_0:

$$q(\tau) = f(t_0 + \tau)$$

With a window spanning a Nyquist period, $\frac{1}{2W}$, above and below t_0, the Gabor

transform is:

$$G(\omega, t_0) = \int_{-\frac{1}{2W}}^{+\frac{1}{2W}} \sqrt{\frac{\alpha}{\pi}} e^{-\alpha\tau^2} q(\tau) e^{-j\omega\tau} d\tau$$

The window size, in combination with our truncation restraint: $w_{th} = 0.001$ determines the α parameter of our window:

$$e^{-\alpha\tau^2} = w_{th}$$

$$\alpha = \frac{-\log(w_{th})}{\tau^2}|_{\tau=\frac{1}{2W}} = -4W^2 \log(w_{th})$$

We can rescale the time domain such that the Nyquist period is normalized from $\frac{1}{2W}$ to $\frac{1}{2\pi}$ by performing a change of variables from τ to $\lambda = \frac{W}{\pi}\tau = \beta\tau$:

Let $y(\lambda) = q(\beta^{-1}\lambda) = q(\tau)$

$$G(\omega, t_0) = \frac{\pi}{W} \int_{-\frac{1}{2\pi}}^{+\frac{1}{2\pi}} \sqrt{\frac{\alpha}{\pi}} e^{-\alpha\beta^2\lambda^2} y(\lambda) e^{-j\Omega\lambda} d\lambda$$

Where $\Omega = \beta\omega$ is the normalized frequency, so any signal has a maximum frequency of $\Omega = \pi$.
Since the signal's maximum frequency is now normalized from W to π, the Nyquist sampling period is $1/2\pi$, and our Gabor transform is integrated from $\lambda = -1/2\pi$ to $\lambda = 1/2\pi$.
Resubstitute α:

$$G(\omega, t_0) = \int_{-\frac{1}{2\pi}}^{+\frac{1}{2\pi}} \sqrt{\frac{-4\pi^2 \log(w_{th})}{\pi}} e^{-(-4\pi^2 \log(w_{th}))\lambda^2} y(\lambda) e^{-j\Omega\lambda} d\lambda$$

Notice that the $\frac{\pi}{W}$ term outside absorbs into the $\sqrt{\frac{\alpha}{\pi}}$ term inside. Effectively, $-4\pi^2 \log(w_{th})$, becomes the normalized bandwidth, α_0

$$G(\omega, t_0) = \int_{-\frac{1}{2\pi}}^{+\frac{1}{2\pi}} \sqrt{\frac{\alpha_0}{\pi}} e^{-\alpha_0\lambda^2} y(\lambda) e^{-j\Omega\lambda} d\lambda$$

The above formula captures the time-frequency representation of $f(t)$ if the full signal is available. But in the application of of real-time sampling, only the signal's instantaneous value and, possibly, some of its past values are available. As mentioned above, differentiation of an arbitrary signal can be implemented with analog circuitry, so we can estimate the signal's instantaneous form in real time with a Taylor expansion centered at $\lambda = 0$:

$$\begin{aligned} y(\lambda) &= y(0) + y'(0)\lambda + \frac{1}{2!}q''(0)\lambda^2 + \frac{1}{3!}y'''(0)\lambda^3 + O(\lambda^4) \\ &= \sum_{m=0}^{K} \frac{1}{m!} y^{(m)}(0)\lambda^m + O(\lambda^4) \end{aligned}$$

Where $O(\lambda^4)$ is residual error caused by truncating all terms fourth-order and higher. It is important to note that the m^{th} order derivative is a constant, which, in application, will be generated by a real-time analog differentiator. As such, the available derivative is of the signal $f^{(m)}(t)|_{t=t_0} = \frac{d^m q}{d\tau^m}|_{tau=0}$, not with respect to the normalized λ. To account for this, we define the m^{th} order derivative of y as:

$$\begin{aligned} y^{(m)}(0) &= \frac{d^{(m)}y}{d\lambda^m}\Big|_{\lambda=0} \\ &= (\beta^{-m})\frac{d^{(m)}q}{d\tau}\Big|_{\tau=0} \\ &= (\beta^{-m})\frac{df}{dt}\Big|_{t=t_0} = \Big(\frac{\pi}{W}\Big)^m D_m \end{aligned}$$

where we have replaced the m^{th} time derivative of our signal by the variable D_m. Substituting the time normalized Taylor approximation into our Gabor STFT:

$$\begin{aligned} G(\omega, t_0) &= \int_{-\frac{1}{2\pi}}^{+\frac{1}{2\pi}} \sqrt{\frac{\alpha_0}{\pi}} e^{-\alpha_0\lambda^2} y(\lambda) e^{-j\Omega\lambda} d\lambda \\ &= \int_{-\frac{1}{2\pi}}^{\frac{1}{2\pi}} \Big[\sum_{m=0}^{K} \frac{1}{m!}\Big(\frac{\pi}{W}\Big)^m D_m\lambda^m + O(\tau^{K+1})\Big] \sqrt{\frac{\alpha_0}{\pi}} e^{-\alpha_0\lambda^2} e^{-j\Omega\lambda} d\lambda \\ &\approx \int_{-\frac{1}{2\pi}}^{\frac{1}{2\pi}} \Big[\sum_{m=0}^{K} \frac{1}{m!}\Big(\frac{\pi}{W}\Big)^m D_m\lambda^m\Big] \sqrt{\frac{\alpha_0}{\pi}} e^{-\alpha_0\lambda^2} e^{-j\Omega\lambda} d\lambda \\ &= \sum_{m=0}^{K} \frac{1}{m!}\Big(\frac{\pi}{W}\Big)^m D_m \sqrt{\frac{\alpha_0}{\pi}} \int_{-\frac{1}{2\pi}}^{\frac{1}{2\pi}} \lambda^m e^{-\alpha_0\lambda^2} e^{-j\Omega\lambda} d\lambda \end{aligned}$$

This expression can be simplified by grouping terms as follows

$$
\begin{aligned}
G(\omega, t_0) &= \sum_{m=0}^{K} \frac{1}{m!}\Big(\frac{\pi}{W}\Big)^m \frac{d^{(m)} f(t_0)}{dt} \sqrt{\frac{\alpha_0}{\pi}} \int_{-\frac{1}{2\pi}}^{\frac{1}{2\pi}} \lambda^m e^{-\alpha_0 \lambda^2} e^{-j\Omega\lambda} d\lambda \\
&= \sum_{m=0}^{K} C_m\Big(\frac{d^m f(t_0)}{dt^m}\Big) \text{ where,} \\
C_m &= \frac{1}{m!}\Big(\frac{\pi}{W}\Big)^m \sqrt{\frac{\alpha_0}{\pi}} \int_{-\frac{1}{2\pi}}^{\frac{1}{2\pi}} \lambda^m e^{-\alpha_0 \lambda^2} e^{-j\Omega\lambda} d\lambda
\end{aligned}
$$

Notice that our estimation of instantaneous spectrum is simply a linear combination of the instantaneous time derivatives of the signal. Not unlike the gauss quadrature method that we discussed this semester, this STFT technique allows for *precomputation* of rather complex integrals for band-limited signals.

2.2 Determining Instantaneous Bandwidth

The previous section demonstrates how the signal's instantaneous spectrum can be calculated for a given normalized frequency $0 \leq \Omega \leq \pi$. To determine instantaneous bandwidth we sample this instantaneous spectrum at H discrete, equally spaced frequencies. If we approximate the signal from its $K+1$ different orders (0 to K_{th} derivative), we require H new coefficients for each of the $K+1$ orders. For a time t_0 the spectrum magnitude at each frequency Ω_i is represented by:

$$
\begin{aligned}
||G(\omega_i, t_0)||_2 &= \Big|\Big| \sum_{m=0}^{K} C_m(\Omega_i)\Big(\frac{d^m f(t_0)}{dt^m}\Big)\Big|\Big|_2 \text{ , where} \\
C_m(\Omega_i) &= \frac{1}{m!}\Big(\frac{\pi}{W}\Big)^m \sqrt{\frac{\alpha_0}{\pi}} \int_{-\frac{1}{2\pi}}^{\frac{1}{2\pi}} \lambda^m e^{-\alpha_0 \lambda^2} e^{-j\Omega\lambda} d\lambda \text{, and} \\
\Omega_i &= \Big(\frac{\pi}{W}\Big)\omega_i
\end{aligned}
$$

Notice that the integral term is a complex integral. Each of the C_m coefficients are complex-valued, and the spectrum magnitude at any Ω_i is the root sum of the real and imaginary components of the $K+1$ complex terms squared. Unfortunately, this operation is not simply performed by analog circuitry. Instead, we calculate an upper bound of the signal magnitude at each frequency:

$$
||G(\omega_i, t_0)||_2^2 = \Big|\Big| \sum_{m=0}^{K} C_m(\Omega_i)\Big(\frac{d^m f(t_0)}{dt^m}\Big)\Big|\Big|_2 \leq \sum_{m=0}^{K} \Big(\frac{d^m f(t_0)}{dt^m}\Big)\Big|\Big|C_m(\Omega_i)\Big|\Big|_2
$$

Since we can precompute the $\Big|\Big|C_m(\Omega_i)\Big|\Big|_2$ terms, the remaining operations of calculating the derivatives and taking a linear combination of the coefficients can be calculated with analog circuitry.

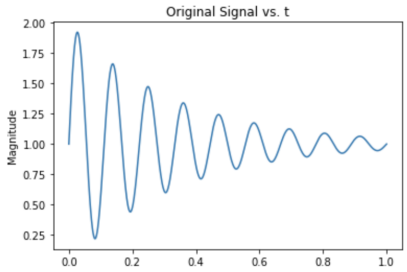

Figure 2.1: Original signal in time: $f(t)$

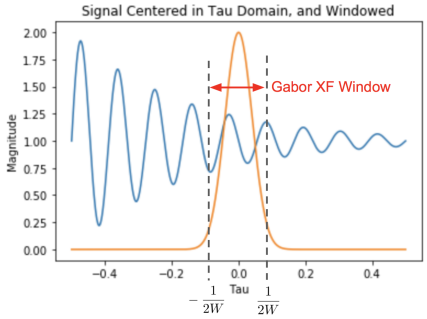

Figure 2.2: Center the signal, apply Gabor window: $q(\tau) = f(\tau + t_0)$

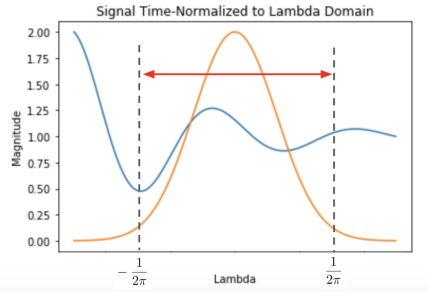

Figure 2.3: Normalize the time axis to : $y(\lambda) = q(\beta^{-1}\lambda)$

Figure 2.4: Demonstration of normalizing the time axis for the Gabor transform. 2.2 is normalized to the same domain for all signals, so the corresponding Taylor series may be precomputed. (Note: signal and Gaussian to scale temporally, but not in magnitude for visibility)

For a system approximating the signal from its K derivatives, the derivatives at time t_0 may be represented as a vector:

$$\mathbf{v}(t_0) = \left[f(t_0), \frac{df(t_0)}{dt}, \frac{d^2f(t_0)}{dt^2}, ..., \frac{d^mf(t_0)}{dt^m}\right]$$

The $(K+1) \times H$ coefficient magnitudes, $\left|\left|C_m(\Omega_i)\right|\right|_2^2$, can be represented in a matrix, $\mathbf{C}$, where

$$\mathbf{C}_{kh} = \left|\left|C_k(\Omega_h)\right|\right|_2$$

In this matrix, each row holds coefficients for the $K+1$ derivatives and each column holds coefficients for the H discrete frequencies. The inner product of the derivative vector, $\mathbf{v}$, with any column of $\mathbf{C}$, $\mathbf{C}_h$, is an upper bound of the signal magnitude at frequency h. An upper bound of the spectrum magnitude at each of the H discrete frequencies can be calculated by simply taking the inner product of the derivative vector and the coefficient matrix:

$$\mathbf{z} = \mathbf{v}^T\mathbf{C}$$

We then consider the signal bandwidth to be the highest frequency component of $\mathbf{z}$ greater than a threshold, $s_{\min}$. Since $\mathbf{z}$ represents *upper bounds* of the spectrum magnitude, we can be confident that the bandwidth is equal to or lower than this frequency. As such, we can treat this as the 'instantaneous Nyquist frequency' of the signal and sample at twice its value as specified by Shannon. More formally, our sampling frequency at a time, t_0 is:

$$F_s(t_0) = 2\max\{\mathbf{z}_u : \mathbf{z}_u \geq s_{\min}\}$$

In this way, our sampling algorithm samples at the locally maximum frequency, as opposed to the globally maximum frequency, conserving the number of samples needed.

2.3 Hardware Implementation

A major advantage of our bandwidth calculation is that it can be calculated with analog circuitry instantaneously. The three operations that our formula uses to calculate bandwidth are differentiation, weighted summation, and thresholding which can be accomplished by high-pass filters, summing amplifiers, and comparators, respectively.

2.3.1 Differentiation

High-pass filters effectively take the derivative of a signal. If $f(t_0)$ is a voltage in time, the current passing through the first series capacitor is

$$i_{C_1} = C\frac{df(t_0)}{dt}$$

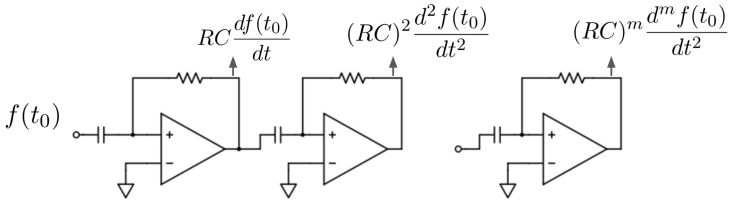

Figure 2.5: Arbitrary order differentiation with high-pass filters

The amplifier uses high open-loop gain to pass a nearly identical current through the resistor. By ohms law, the voltage at the output of the resistor, and thereby the amplifier, is:

$$\begin{aligned} v_{out} &= i_{res}R \\ &\approx i_{C_1}R \\ &= (RC)\frac{df(t_0)}{dt} \end{aligned}$$

By chaining m of these amplifiers together and using the same values of RC, we can achieve a derivative of order m scaled by $(RC)^m$. By choosing the time constant of the amplifiers, RC to be $\beta^{-1} = \frac{\pi}{W}$, the output of the n_{th} amplifier is:

$$v_n(t_0) = \left(\frac{\pi}{W}\right)^n \frac{d^n f(t_0)}{dt} = \mathbf{v}^T\boldsymbol{\beta}$$

Where $\boldsymbol{\beta}^{-1}$ is a vector with the n_{th} element being $(\frac{\pi}{W})^n = \beta^{-n}$

2.3.2 Weighted Summation

The operation, $\mathbf{v}^T\mathbf{C}$, where $\mathbf{C}$ is a $(K+1) \times H$ matrix, is effectively a weighted summation of $\mathbf{v}$ using a column of $\mathbf{C}$ as weights. This operation can be performed using a summing amplifier.

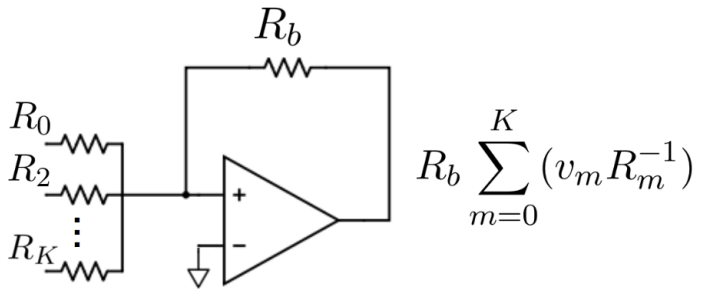

Figure 2.6: Weighted summation of the K different derivative voltages

We include H different amplifiers, one for each frequency bin. Each amplifier has $K+1$ different inputs, one for each of the different derivative orders. We choose the resistor values such that the output voltage is our upper bound on frequency content:

Knowing that the n_{th} input voltage is:

$$v_n(t_0) = \Big(\frac{\pi}{W}\Big)^n \frac{d^n f(t_0)}{dt}$$

We choose R_n and R_b of the i_{th} frequency amplifier (i_{th} frequency bin, Ω_i) such that:

$$\begin{aligned}
\sum_{n=0}^{K} R_b R_n^{-1} v_n &= \sum_{n=0}^{K} R_b R_n^{-1} \Big(\Big(\frac{\pi}{W}\Big)^n \frac{d^n f(t_0)}{dt^n} \Big) \\
&= \sum_{m=0}^{K} \Big(\frac{d^m f(t_0)}{dt^m} \Big) \Big|\Big| C_m(\Omega_i) \Big|\Big|_2 \geq ||G(\omega_i, t_0)||_2 \text{ so,} \\
R_b R_n^{-1} &= \frac{\Big|\Big| C_n(\Omega_i) \Big|\Big|_2}{\Big(\frac{\pi}{W}\Big)^n} \\
&= \frac{1}{n!} \sqrt{\frac{\alpha_0}{\pi}} \Big|\Big| \int_{-\frac{1}{2\pi}}^{\frac{1}{2\pi}} \lambda^n e^{-\alpha_0 \lambda^2} e^{-j\Omega\lambda} d\lambda \Big|\Big|_2
\end{aligned}$$

So, our amplifier resistor values scale the derivatives by the precomputed Gabor transform of the normalized Taylor components. With this configuration, the output voltage of the i_{th} amplifier represents an upper bound on the magnitude of the i_{th} frequency bin, $\omega_i = \Big(\frac{W}{\pi}\Big)\Omega_i$

2.3.3 Thresholding

Each of the H weighted summation amplifiers have an output voltage proportional to an upper bound on its associated frequency bin. Each of these voltages can be thresholded by a comparator, that receives an analog input (the magnitude bound) and drive a digital output:

$$\text{Comp}_i = \textbf{Indicator}(v_m > v_{th})$$

Where v_{th} is a threshold voltage set by a reference circuit, which are common in most converters. The highest frequency bin to trigger its comparator sets the sampling frequency of the converter.

Chapter 3

Reconstruction Techniques

Signal reconstruction is a widely researched field. In Chapter 1, we provided a brief overview of this field, discussing a few different classes of reconstruction. In particular, we focused on reconstruction of a bandlimited signal from both uniform and nonuniform samples. In Chapter 2, we discussed a non-uniform sampling scheme designed to reduce the number of samples needed for perfect reconstruction below the conventional Nyquist lower bound. In this Chapter, we will discuss three different reconstruction methods for bandlimited signals sampled on a non-uniform grid.

First we will relate the problem at hand to a few of the interpolation methods that we've discussed in class this semester, specifically Langrange and cubic-spline interpolation. Then we will discuss an interpolation method that we devised in order to reconstruct signals from non-uniformly sampled data collected using the methods described in Chapter 2. This reconstruction technique applies Shannon's Sampling Theorem to a piecewise uniformly sampled data-set. Assuming that the instantaneous sampling rate across all time exceeds the Nyquist rate, we expect this piecewise Shannon reconstruction method to perform as well as it would assuming uniform sampling at the true Nyquist rate.

3.1 Lagrange Interpolation

Lagrange interpolation is a form of polynomial interpolation based on the Lagrange basis functions

$$l_j(t) = \frac{\prod_{k=1,k\neq j}^{n}(t - t_k)}{\prod_{k=1,k\neq j}^{n}(t_j - t_k} \tag{3.1}$$

From this basis, a polynomial of degree $n-1$ interpolating the data-set, (x_i, y_i), $i = 1, 2, ..., n$, can be determined as a linear combination of the Lagrange basis functions, weighted by the sample values

$$p_{n-1}(t) = y_1 l_1(t) - y_2 l_2(t) + ... + y_n l_n(t) \tag{3.2}$$

The resulting polynomial is guaranteed to pass through each sample point, regardless of the spacing of each sample in the time domain. In fact, through an analysis of Lagrange interpolation on Runge's function, we discussed in class that non-uniform sampling can *improve* interpolation accuracy. Unfortunately, Lagrange interpolation becomes unwieldy as the data set size grows large, since for every additional data point, the interpolating polynomial's order increases by one. The sampling method that we've proposed requires signal reconstruction/interpolation from a set of samples of arbitrary length. As the number of samples grows from tens to hundreds to thousands, the effectiveness and accuracy of a single polynomial fit diminishes rapidly. This realization quickly turned us away from applying a single polynomial fit to our data. Instead, we began to consider piece-wise reconstruction techniques.

Piecewise reconstruction can take on a variety of forms. The general idea is to segment the original data set, and find an interpolating function for each segment. The entire interpolated data set is then defined piecewise in terms of each segment's interpolating function. In the simplest case, a linear piecewise reconstruction can be employed, interpolating neighboring samples with a constant slope line. However, this imposes high reconstruction error when the sampling rate is low. A better fit is made through a piecewise Lagrange interpolation. This scheme, segments a data set sampled on an arbitrarily spaced grid into segments containing 5 data points each. Within each segment, a polynomial of order 4 is fit to the 5 data points, using Lagrange's method, and the resulting interpolation is defined piecewise over the entire sampling grid.

While this method ultimately produces reasonable results, large errors can arise near the boundaries of each segment, since no constraints are placed on how neighboring piecewise functions are combined. This naive piecewise polynomial interpolation method can be improved significantly using the cubic spline.

3.2 Piecewise Spline Interpolation

The cubic spline is a piecewise defined function, $F(t)$ of degree 3 that interpolates a dataset, (x_i, y_i), $i = 1, 2, ..., n$. This function can be defined as follows:

$$F(t) = \begin{cases} f_1(t) = a_1x^3 + b_1x^2 + c_1x + d_1 & t_1 \leq x \leq t_2 \\ \quad \vdots & \\ f_{n-1}(t) = a_{n-1}x^3 + b_{n-1}x^2 + c_{n-1}x + d_{n-1} & t_{n-1} \leq x \leq t_n \end{cases} \tag{3.3}$$

Notice that (3.3) is defined in terms of a set of unknown constant scaling factors for each cubic polynomial. In order to fully define the interpolant, this set of coefficients $\{a_i, b_i, c_i, d_i\}$, $i = 1, ..., n-1$, where n is the number of points in the data set, must be found. With $4(n-1)$ unknowns, $4(n-1)$ constraints are required for a unique solution. Interpolating the data points exactly provides $2(n-1)$ constraints, since each cubic polynomial must equal the two sampled values at the endpoints of

its interval. Furthermore, requiring continuous first and second derivatives at each sample point imposes and additional $2n - 4$ constraints. Therefore, 2 additional constraints are required to fully constrain the system. We decided to impose the constraints of a "natural cubic spline", requiring the second derivatives at the two endpoints of the reconstruction to equal zero.

Clearly, as the number of samples increases, the size of the linear system necessary to solve for the coefficients in F grows rapidly. Fortunately, a solution to this large linear system has been defined using a specially constructed tradiagonal matrix that captures the constraints posed above. Since cubic spline interpolation was not the focus of this project, we will not provide a derivation of the exact form of this solution, and will instead point the reader to [[3]] for more information..

We implemented this interpolation technique numerically using a function from python's scipy package, InterpolatedUnivariateSpline. We've used spline reconstruction as a comparison against our windowed sinc reconstruction technique, which we define in the next section.

3.3 Piecewise Sinc Interpolation

From Shannon's Sampling Theorem, we known that signal reconstruction from a set of uniform samples taken at or above the Nyquist rate can be achieved perfectly using equation (1.1). Unfortunately, the reconstruction formula spans an infinite time domain, and, thus, is not numerically realizable. Considering this, we turn to a classical reconstruction method from uniform samples known as upsampling. Upsampling interpolates between samples of a bandlimited signal. Unlike Langrange and Spline interpolation, upsampling does not produce an analytic function that can be evaluated at any point along the signal's domain. However, it can be used to accurately reconstruct a bandlimited signal on a very tight grid, effectively recovering the signal's original continuous form.

As mentioned, upsampling interpolates between points of a bandlimited signal. Explicitly, given an array of uniformly spaced samples of a bandlimited signal, $S = [s_1, s_2, ..., s_N]$, upsampling by a factor of L increases the number of samples in S from N to $N(L-1)$. Theoretically, the new samples are exact, in that they match samples of the original signal exactly, had that signal been sampled L times faster. Upsampling is a two step process. In the first step, known as expansion, $L-1$ zeros are inserted between each sample in S. In the frequency domain, this is equivalent to compressing the frequency axis, bringing each spectrum image closer by a factor of L. In the ideal case, the original signal can be perfectly reconstructed from the expanded signal by applying an ideal low pass filter with bandwidth $\frac{\pi}{2L}$. Unfortunately, an ideal lowpass filter spans an infinite time domain, so a windowed filter is constructed, and convolved with the signal instead.

This process is only effective for reconstructing signals sampled on a uniform grid. The samples that we've collected, however, are "piecewise uniform". In other words, within finite windows, the collected samples are uniformly spaced; however,

across windows, the sampling grid varies widely. To account for this, we augmented the upsampling method described above to vary its upsampling rate across windows. Additionally, we introduced a simple technique to smoothly merge each individual upsampled window.

The reconstruction algorithm that we devised proceeds as follows. The method assumes that a set of samples and the index in time at which each sample was taken has been provided. First, each sampling window is identified from the array of time indices by determining the points in time at which a sample rate transition occurred. Next, reconstruction is carried out within each window individually, using the uniform upsampling method described above, with one point of overlap between neighboring windows. The upsampling rate employed is equivalent to the time difference between samples within the window so that the reconstructed signal has the same number of samples as the original signal had the original signal been sampled at our system's maximum sampling rate.

Chapter 4

Experimental Results

To assess the effectiveness of our sampling and reconstruction scheme, we compared its performance against three baseline sampling and reconstruction techniques, considering the following metrics: reconstruction mean-squared error and the number of samples needed for reconstruction. We first tested against the classical sampling technique, in which samples are taken on a uniform grid at the twice the bandwidth (i.e. the Nyquist rate). We then compared our method's performance against a technique that samples on a uniform, *sub-Nyquist* sampling grid. The sub-Nyquist rate employed is directly related to the reduced number of samples taken by our variable rate sampling technique. Finally, we compared our reconstruction results to those achieved with piecewise spline interpolation on a non-uniformly sampled grid. In the following sections, we discuss the simulation process in a bit more detail and present the results that we achieved on our simulated data set.

4.1 Experimental Setup

As discussed in Chapter 2, an important requirement of the signals that our sampling scheme processes is that they are band-limited. In 1.1 we discuss the concept of a band-limited signal, so it is clear that we need to ensure that the signals that we process have negligible frequency content above a known frequency (i.e. the bandwidth). It is easy to ensure that synthetic data (i.e. data generated using a known analytic function) has no frequency content above a certain threshold. However, we wanted to process real physical signals, not merely synthetic data.

Specifically, we wanted to process a well known human biological signal: a heart's pulse, as recorded by an echo-cardiogram. Fortunately, we were able to gain access to an online database of such signals [NIGMS]. After identifying a appropriate signal, we needed to devise a way to ensure that this signal was band-limited. At a high level, this boils down to three steps. First, the signal bandwidth must be identified through an analysis of the signal's frequency domain. Next, any component frequencies above the specified bandwidth must be filtered out to ensure that only the frequencies below the specified bandwidth remain in tact. Finally, an analytic expression for the signal

as a function of time must be constructed. For this final step, a piecewise cubic spline is used to fit the filtered data, because of the ease with which it can be constructed and the accuracy that it provides in regard to both the signal value and its derivatives up to order 4.

When computing the signal bandwidth, care needs to be taken when transforming the signal into the frequency domain. An accurate representation of a signal's frequency spectrum can be determined from its uniform samples using an FFT. Note, however, that Fourier analysis typically assumes an infinite time domain. To account for this, when analyzing finitely sampled signals, a technique called "windowing" is employed to minimize frequency domain distortion during the time to frequency transformation. In our implementation, we applied a kaiser window to the finite sample, and then took the FFT of this windowed signal data to resolve its frequency content. Next we identified the signal bandwidth using the technique described in section 1.1, by finding the maximum frequency whose relative power (calculated using (1.2)) exceeds a pre-specified threshold.

After identifying the bandwidth, the original signal data is filtered using a 100-tap finite impulse response filter. This filter is an approximate "ideal" filter. In signal theory, and ideal filter is one that passes all frequencies within a specified frequency band, and attenuates all frequencies outside of this band completely. While such an ideal filter can not be practically realized, a good approximation can be achieved by convolving the sampled data with a windowed sinc with the specified cutoff frequency (i.e. the signal bandwidth that we identified). The following three figures depict the filtering process from start to finish. Take note of the reduction in high frequency content between **Figure 4.1** and **Figure 4.3**.

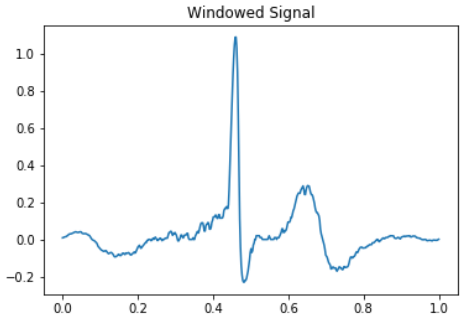

Figure 4.1: Windowed, unfiltered heart-rate signal data

Now that the signal has been properly conditioned, it can be passed through our sampling and reconstruction scheme. First, a tight time grid is defined for the filtered data set, and the original signal value and its derivatives are calculated at each point on this grid to simulate the continuous signal. Then coefficients for the Garbor STFT are calculated for the $m = 0, 1, 2, 3$ order derivatives at $K = 100$ frequencies. Finally, the frequency coefficient magnitudes are computed as a linear

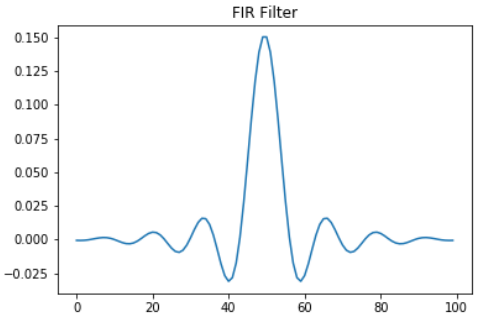

Figure 4.2: Windowed sinc used to filter heart rate data at the specified signal bandwidth via convolution

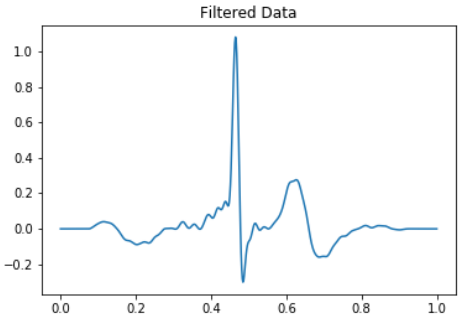

Figure 4.3: Band-limited heart rate signal after convolution with "ideal" low pass filter kernel from **Figure 4.2**

combination of the corresponding derivative and Garbor STFT coefficient terms for each derivative. The instantaneous bandwidth is then computed as explained in section 2.2, from which the instantaneous sampling rate can be directly computed.

4.2 Sampling and Reconstruction Results

Once the signal's instantaneous sampling rate is identified and the signal is sampled, the non-uniformly sampled signal and its time indices are passed to the non-uniform sinc reconstruction method described in section 3.3. Below we've included key figures that demonstrate the functionality and performance of our sampling and reconstruction technique. **Figure 4.4** demonstrates the computed time-frequency spectrum of the input signal. Each curve represents one component frequency magnitude at each point in time. The bandwidth is identified from these curves as the maximum frequency whose component magnitude exceeds a specified threshold. **Figure 4.5** represents the sampling rate as a function of time. Notice that the y-axis denotes the sampling rate as a ratio of the Nyquist rate. As expected, the maximum sampling

rate is required in the interval at which the input signal sees a large amplitude spike, corresponding to the interval in which the highest frequency components are present.

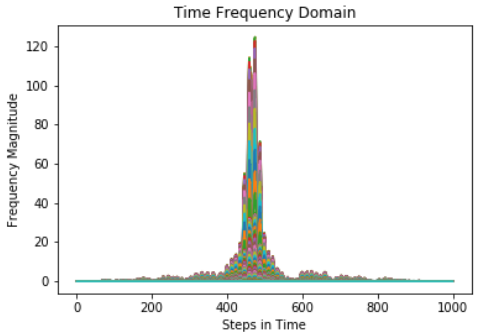

Figure 4.4: Time-Frequency representation of the heart-rate signal. This is a numerical realization of the function, $G(\omega, t)$.

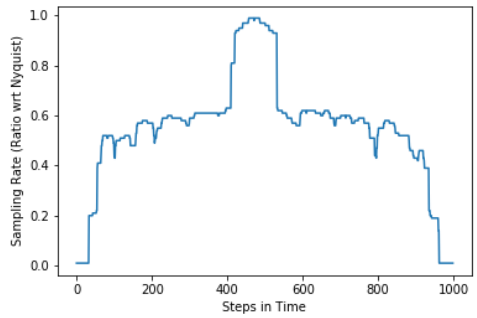

Figure 4.5: Sampling rate as a function of time, defined as a ratio of the Nyquist rate.

Figure 4.6 and **Figure 4.7** demonstrate the reconstruction results achieved from an ideal reconstruction, sampling on a uniform Nyquist grid. This is the classical sampling technique and is used as a baseline for comparison with our reconstruction.

Figure 4.8 and **Figure 4.9** demonstrate the results achieved using our modified sampling and reconstruction scheme, which utilizes instantaneous bandwidth to modulate the sampling rate in real time. Note that, despite some error, the resulting reconstruction is fairly accurate.

Figure 4.10 and **Figure 4.11** demonstrate the reconstruction performance of the original signal had it been sampled uniformly using the same number of points that the variable rate sampling and reconstruction scheme that we've devised uses. Clearly, reconstruction results are very poor within the interval of time with strong high frequency components.

The table below quantitatively demonstrates the performance of each sampling and reconstruction scheme whose results we've presented above.

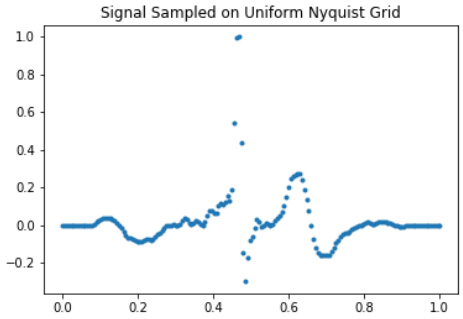

Figure 4.6: Uniform samples of the original signal. The uniform grid spacing was set using the Nyquist Criterion.

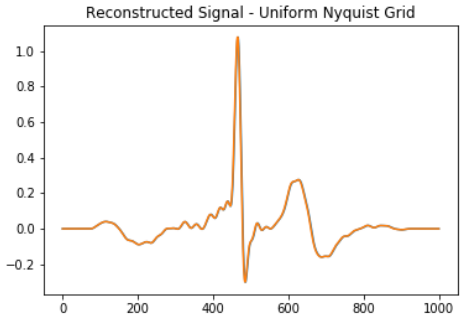

Figure 4.7: Reconstruction of the original signal from its uniform samples taken at the Nyquist rate

Sampling Method	Interpolation Method	MSE	Maximum Error	Number of Samples
Nonuniform	Piecewise Sinc	2.03×10^{-4}	7.2×10^{-2}	82
Uniform	Piecewise Sinc	2.92×10^{-3}	3.66×10^{-1}	82
Nonuniform	Spline	3.84×10^{-6}	8.2×10^{-3}	82
Uniform	Piecewise Sinc	5.47×10^{-7}	4.9×10^{-3}	168 (Nyquist)

Although the best sampling scheme with regard to MSE is uniform Nyquist (row 4), our nonuniform sampling and reconstruction method (row 1) achieves small MSE despite sampling fewer than half as many times as a conventional sampling application would require given the specified signal bandwidth. This sampling reduction would increase considerably for sparse signals. Note that if the original signal is sampled uniformly at the reduced sampling rate (row 2), both the mean squared error and the maximum error increase by more than a factor of 10, hinting at the effectiveness of our variable rate sampler. It is interesting to note that spline reconstruction of the non-uniformly sampled signal (row 3) reconstructs the original signal very well. This is likely due to the fact that the original signal was approximated by

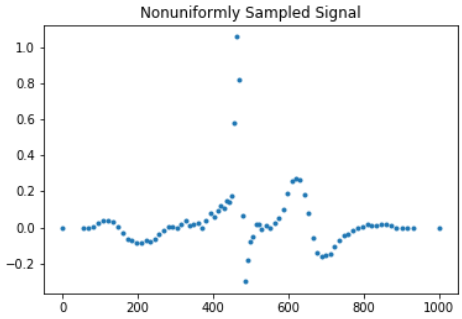

Figure 4.8: Nonuniform samples of the original signal. The nonuniform grid was defined in real time by the sampling scheme that we devised.

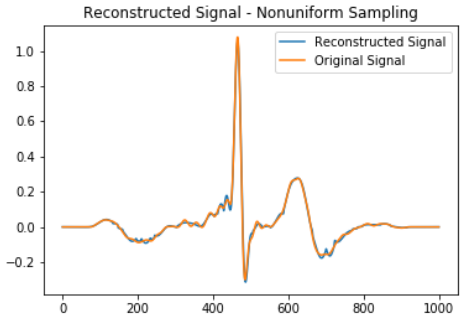

Figure 4.9: Reconstruction of the original signal from its samples taken on the nonuniform grid demonstrated in **Figure 4.8**

a cubic spline.

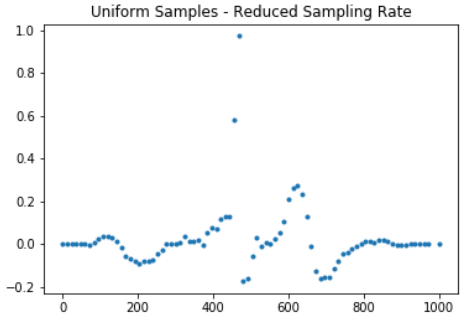

Figure 4.10: Reduced rate, uniformly spaced samples

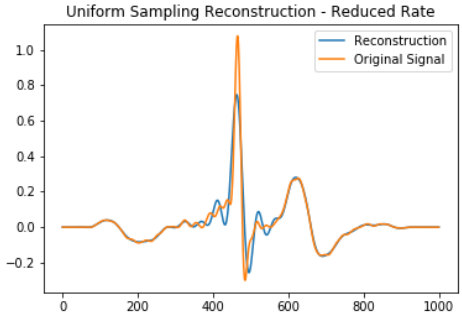

Figure 4.11: Reconstruction of reduced rate uniformly sampled signal.

4.3 Future Work

While the results that we have achieved are strong, there are some areas that we would like to explore and improve upon moving forward. In particular, we would like to identify the cause of the sharp spikes in the both the time-frequency representation of the signal and the resulting instantaneous sampling rate function. Smoothing the sampling rate function reduces reconstruction error, assuming that the instantaneous sample rate exceeds the instantaneous bandwidth.

Additionally, our method avoids performing the nonlinear L2-norm calculation of the complex frequency coefficients and instead calculates an upper bound by summing the magnitude of imaginary and real components. However, RMS calculations can be computed with analog circuitry and the L2-norm is likely similar. It is worth exploring exact resolution of signal bandwidth using this nonlinear operation, and more accurately determining the minimum required sampling rate.

Provided more time, we would also like to test our method on a wider range of signals. In particular, it would be interesting to see how our implementation performs on signals whose frequency content varies more gradually. The chirp signal, commonly used as a baseline test signal in digital signal processing techniques would be a particularly interesting analytic function to use for analysis due to its predictable time-frequency representation.

Chapter 5

Discussion and Conclusion

In this report, we have demonstrated a sampling and reconstruction technique that reduces that number of samples required to reconstruct a band-limited signal with high accuracy below the number of samples required by Nyquist's lower bound. This is achieved by varying the sample rate in real-time based on the instantaneous signal bandwidth. From our research of time-frequency analysis and non-uniform reconstruction, we have not found any work aimed at determining the optimal sampling rate for a given signal in real time. Our method approximates the instantaneous frequency spectrum of a given signal from the Short-Time-Fourier-Transform of the signal's instantaneous Taylor approximation. This allows computation of an upper bound instantaneous bandwidth. As we discussed, since the computation required to identify the instantaneous bandwidth is simply a linear combination of derivatives, our technique could be implemented purely in analog circuitry. This allows for instantaneous determination of required sampling frequency without any digital engine. We have found that – paired with our corresponding interpolation technique – our method is highly adaptive: given a sub-Nyquist limited number of samples, our implementation allocates more effectively than traditional uniform sampling, reducing reconstruction error.

Bibliography

[1] Cohen, L. (1989). Time-Frequency Distributions, A Review. *Proceedings of the IEEE*, 77(7):941–981.

[2] Eldar, Y. C. (2000). Filterbank Reconstruction of Bandlimited Signals from Nonuniform and Generalized Samples. *IEEE Transactions On Signal Processing*, 48(10):2864–2875.

[3] John H. Matthews, K. D. F. (2004). *Numerical Methods Using MATLAB*. Pearson, 4. ed. edition.

[4] Margolis, E. (2008). Nonuniform Sampling of Periodic Bandlimited Signals. *IEEE Transactions On Signal Processing*, 56(7):2728–2745.

[NIGMS] NIGMS. Physionet. https://physionet.org/. Accessed: 2017-12-18.

[6] Qian, S. (1999). Joint Time-Frequency Analysis. *IEEE Signal Processing Magazine*, 16:52–67.

[7] Shannon, C. E. (1949). Communication in the Presence of Noise. *Proceedings of the IEEE*, 37(1):10–21.

[8] Yen, J. L. (1956). On Nonuniform Sampling of Bandwidth-Limited Signals. *IRE Transactons on Circuit Theory*, pages 251–257.