ECE4634 - Digital Communications
September 26, 2020
Nima Mohammadi
nimamo@vt.edu

Sampling and Quantization

Design Project 1
Technical Memorandum

I. Introduction

In this technical report we will investigate the impact of quantization and sampling procedures on a sample voice recording. To mimic the analog audio signal, the voice recording data (of 5.86 seconds) has been provided with a relatively high sampling rate of $f_s = 65536$ and large number of quantization levels. Upon inspection of the voice recording, we can observe that a nonlinear quantization scheme (Fig. 1) with 268858 quantization levels have been employed

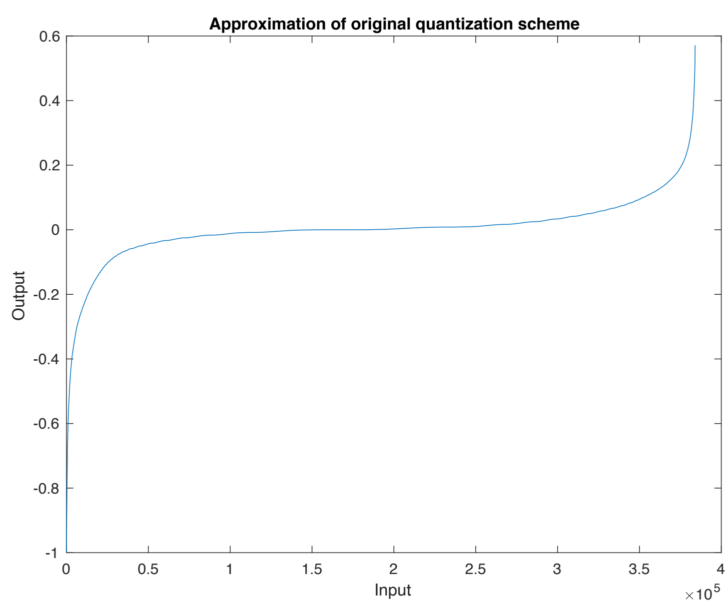

Fig 1. Approximation of quantization scheme of original audio recording

ranging from -1 to 0.5714; needless to say that original number of quantization level have probably been higher, but possibly not all those levels have been realized in this recording.

The sound recording contains the voice of a male subject reading a sentence. Fig 2. visualizes the audio clip in time domain. Power Spectral Density (PSD), as the name implies, is

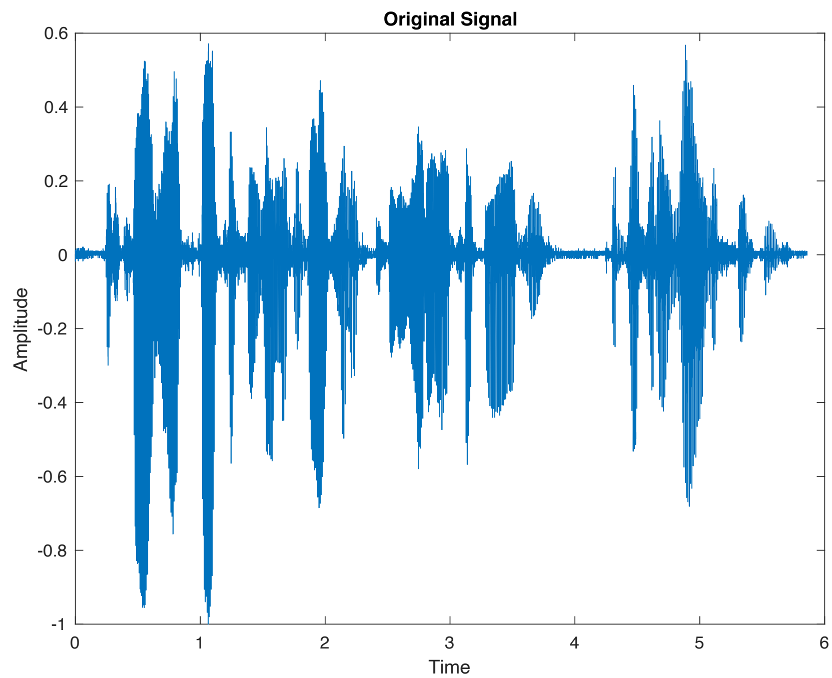

Fig 2. Plot of the original audio signal in time domain

defined based on power rather than on energy. Energy Spectral Density (ESD) is a better measurement tool here since our signal has finite energy over time. As the energy of our signal is limited to a finite time interval, we have used ESD, depicted in Fig. 3 with both ordinary and logarithmic scales (note that in the ordinary-scale plot it has been zoomed in). The bandwidth holding 99.9% of the signal energy is 3.685kHz.

II. Impact of Sampling Rate

Various sampling rates have been tested in this experiment, namely, f_s = 2048, 4096, 8192, 16386. Fig. 4 depicts the results of performing sampling with these rates. The left column depicts the time-domain signal for a short period of time (i.e. 15ms). We can vividly see how choosing a

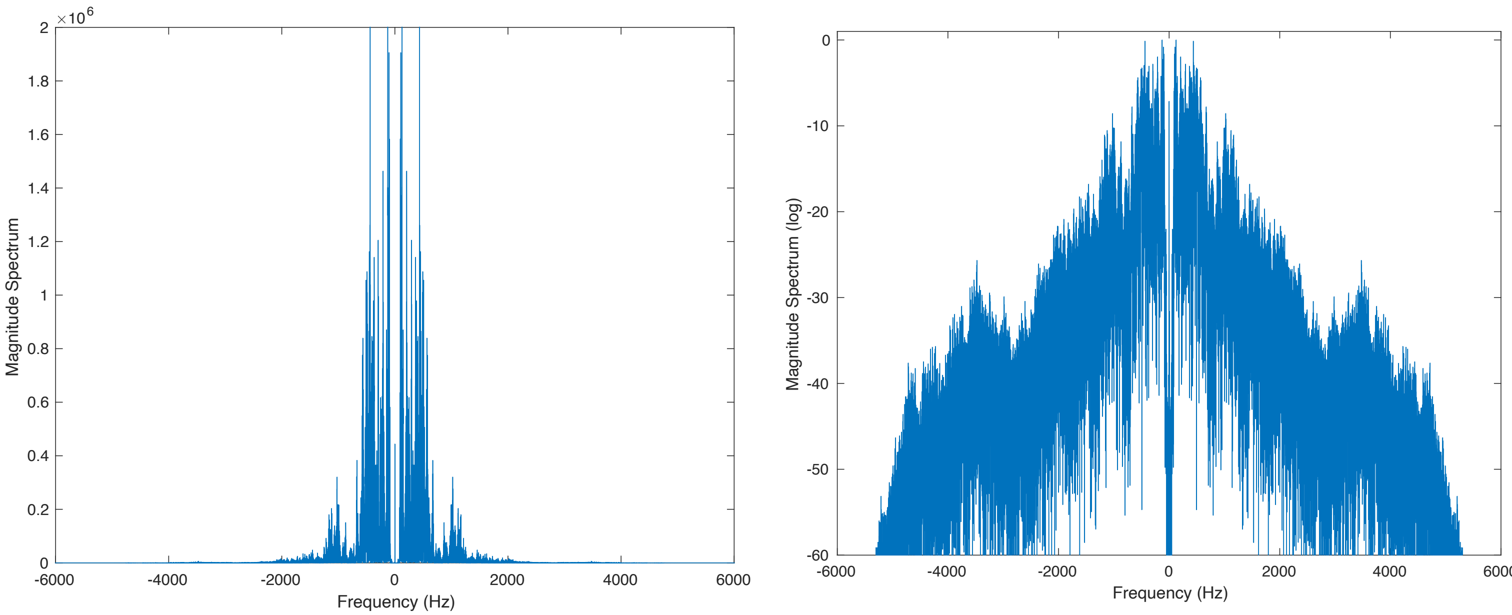

Fig 3. Energy Spectral Density of the original signal a) ordinary scale b) logarithmic scale

higher sampling rates would preserve more information and better reconstruct the signal. The column in the middle shows the log-scale ESD plots for each sampling rate. The horizontal axis of the plot has been purposefully chosen to be identical across different sampling rates to better depict the frequency components that have been inhibited as a result of down-sampling. The right column on the other hand holds the plots for ordinary scale ESDs. Notice that for these set of plots, however, I have not chosen identical boundaries for the horizontal axes to get a higher resolution view of the spectrum.

Playing each sampled signal would enable us to get a qualitative assessment of the impact of sampling at different rates on the voice quality. Needless to say, this is a subjective measure and the perception of the sound might vary for different people based on age and damage to people's hearing as a result of abuse/misuse of earbuds and headphones. The signal sampled at 2048Hz is almost unintelligible and the words can not be distinguished. The 4096Hz sampled data significantly improves the quality of the sampled data, however, it is still very unnatural. Next, for 8192Hz we can hear a much clearer voice, but it still sounds a bit huskier and deeper than the original signal. Moving up to our highest sampling rate, I can not distinguish the original recording with 16384Hz-sampled signal. Given these observations, I would go with either 8kHz or 16kHz sampling rates if the system is only transmitting or storing human voice recordings.

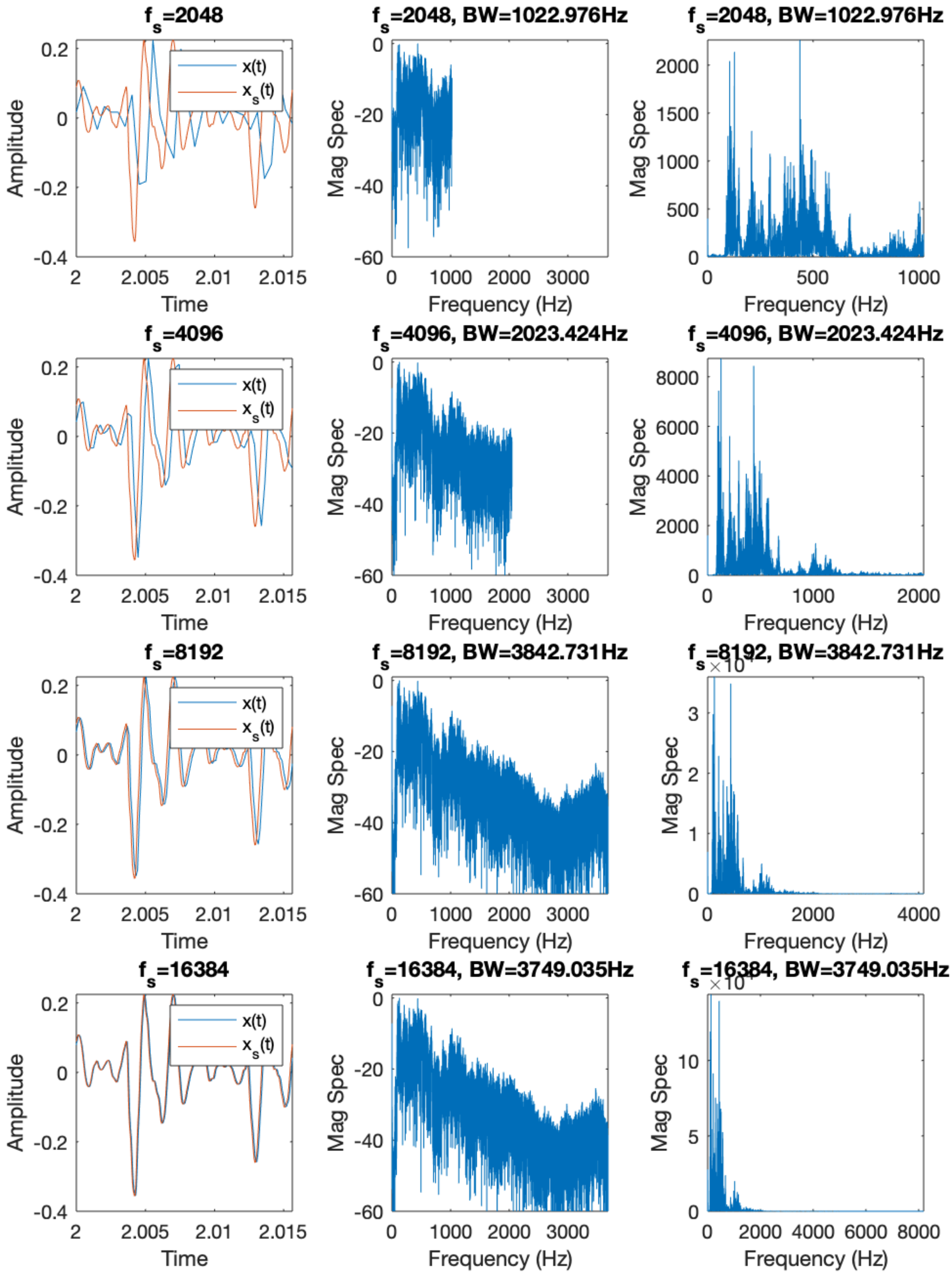

Fig. 4. Different Sampling Rates (left) Time domain (middle) ESD log scale (right) ESD ordinary scale

Compared with the 3.685kHz and taking into account the Nyquist sampling rate, it does seem that choosing 8kHz is justified and it also confirms our observation of lower sampling rates being inadequate.

Having interpolated the sampled signals by a proper factor, we can calculate the SNR with respect to the original signal. The result is shown in the table below. The table shows an almost monotonic relationship between the sampling rate and SNR. However, surprisingly the signal sampled at 8kHz has higher SNR compared to 16kHz signal. My prediction was that given the bandwidth corresponding to 99.9% energy of the signal falls in the area covered by both sampling rates, we should not observe much improvement for 16kHz sampling rate, but I did not expect observing a decline. Upon further pondering, I have formed a conjecture about this. We know that both 8kHz and 16kHz sampling can capture the "bulk" of the original signal, and that the remaining 0.1% energy lying above 3.685kHz are high frequency components that can be framed as noise. Now, I am making the assumption that the noise is centered around zero and what I am claiming is that the loss os SNR for 16kHz sampling is a consequence of interpolation technique we used. The interpolation, interp() function in MATLAB, connects consecutive points, effectively averaging out and canceling the noise in 8kHz which was an adequate sampling rate for our case. But for 16kHz, the linear interpolation, given the erratic behavior of high frequency noise, tries to fill in the gaps, attempting to mimic the noise, which is unsuccessful due to the inherent irregularity of this noise. We might have achieved better SNR for higher sampling rates if we tweaked the interpolation scheme (e.g. quadratic, cubic or higher orders interpolations), but since those high frequency components are irrelevant to us, we refrain from doing so.

Sampling Rate	SNR
2048	8.7185
4096	17.8478
8192	238.1315
16384	48.7612

III. Impact of Quantization

Next, we are going to examine different quantization techniques. Quantization is the process of mapping input values from a large set to output values in a (discrete) smaller set. Here we investigate uniform quantization, the effect of amplifier gain control, and nonuniform quantization.

III.I. Uniform Quantization

Starting with uniform quantization, six different number of quantization levels, namely 2, 8, 32, 128, 512 and 1024, have been tested. These quantization transformations are plotted in Fig 5. Applying these transformation, we get the signals depicted in Fig. 6, converging more and more to the original signal as we increase the quantization levels.

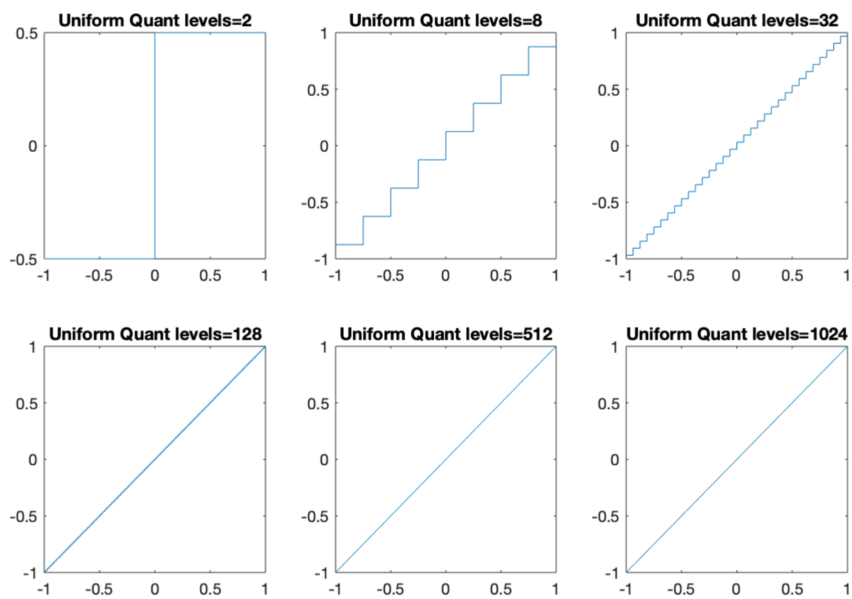

Fig. 5. Uniform Quantization

As for the perception of the sound, the 2-level quantization was much better than I initially expected. The 2-level quantization is completely unintelligible, but one can still figure out that it is a voice recording and not some random sound. The quality of 8-level quantization is still very bad, but moving to 32-level we can actually understand the statement being read on the voice recording. The associated plot in Fig. 6 also indicates that this is where the quantized signal starts to approach the shape of waves in the original signal. The 128-level quantization can be easily comprehended, but some irritating noise-like sound can be heard, making it unpleasant to hear. The 512 and 1024 levels of quantizations are both quite acceptable, and I can not easily distinguish them with my commodity computer speakers. Given my observations, I would go for 512-level quantization, that is 9 bits.

The table below summarizes the results for uniform quantization. The theoretical maximum signal-to-quantization-noise ratios are calculated based on $1.8 + 6R$, where R is the number of quantization bits.

Number of bits	Measured SNR	Theoretical SNR	Sound Quality
1	-13.3539	7.8	Gibberish
3	-0.0701	19.8	Very bad
5	13.3540	31.8	Bad, understandable but w/ artifacts
7	26.1555	43.8	Good, with noise
9	38.1261	55.8	Almost perfect
10	44.3737	61.8	Almost perfect

Calculating the difference between the original signal and the quantized versions of the signal we get the error signals, depicted in Fig. 7. The error signals have been plotted with the same scale on the horizontal axis, to better rank the efficacy of different quantization levels. The plotted error signals are only limited to 10 milliseconds for the sake of better resolution. We can see that increasing the number of quantization levels introduces less error, with 2-levels of quantizations naturally having the higher magnitude of error.

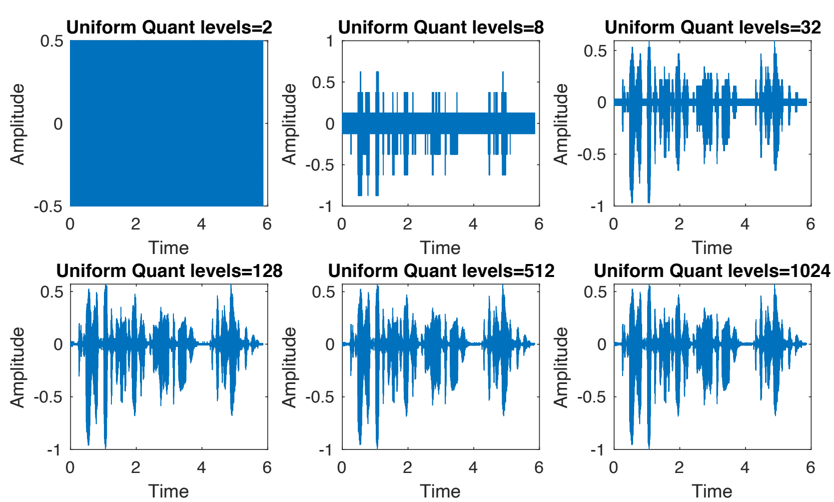

Fig. 6. Uniform Quantization of the audio signal

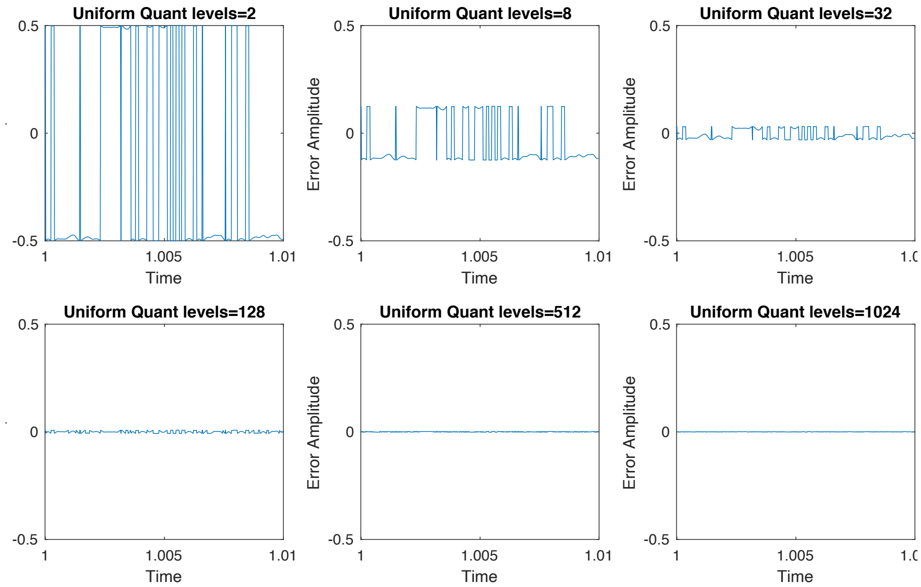

Fig. 7. Uniform Quantization Error for 10ms

III.II. Uniform Quantization with Amplifier Gain Control

In this section, we examine the impact of tweaking the amplifier gain in our uniform quantization procedure. We start by sampling the signal with a sampling rate of 8192Hz. The sampled signal is then quantized with 4-bits (16 quantization levels). The voice message in the resulting signals is understandable but we can hear undesired artifacts. Then we shrink the gain of the signal by a factor of 2 (multiplying the sampled signal by $\frac{1}{2}$) and doubling the gain of the quantized signal. This will result in a worse-quality sound. We repeat the 4-bit quantization, but this time we increase the input gain by a factor of 5 and halving the output gain of quantized signal. The resulting sound has less "volume", which indicates that the input and output gains are not simply canceling out each other. The latter form of gain control brings a better sounding signal with less artifacts and more pleasing quality.

4-bit Quant	**Measured SNR wrt x**	**Sound Quality**
$Q(x,4)$	7.0098	Good
$2Q(\frac{1}{2}x,4)$	0.2008	Very bad
$\frac{1}{5}Q(5x,4)$	6.7731	Better than no gain control

This experiment shows us the importance of proper gain control. With decreasing the input gain (multiplication of signal by 0.5) we are effectively constraining the input range of the quantization procedure, which is basically squashing quantization levels, and as we have seen this can not be recovered by amplifying the quantized signal by the reciprocal of this value. On the other hand, amplifying the input signal whose values are distributed around the mean, increases the resolution for amplitudes around zero which are more prevalent, resulting in better-quality signal. The effectiveness of this procedure, however, depends on the characteristics of the signal, that is if the person whose voice we are hearing in the audio clip was shouting, the reverse procedure might have been preferable.

The Fig. 8 depicts the error signal of 10ms for these three cases. As we can see from the plots, the third case (amplification of input and damping output of quantization) results in less error. This is despite its SNR being negligibly smaller than the quantized signal with no gain control. My understanding is that SNR is a statistic summarizing the whole signal, which is flawed by its dependence on possibly few but potentially significant "outliers" that can greatly affect its value. Overall, in my opinion, looking at the plot of an error signal is a better tool of judgement.

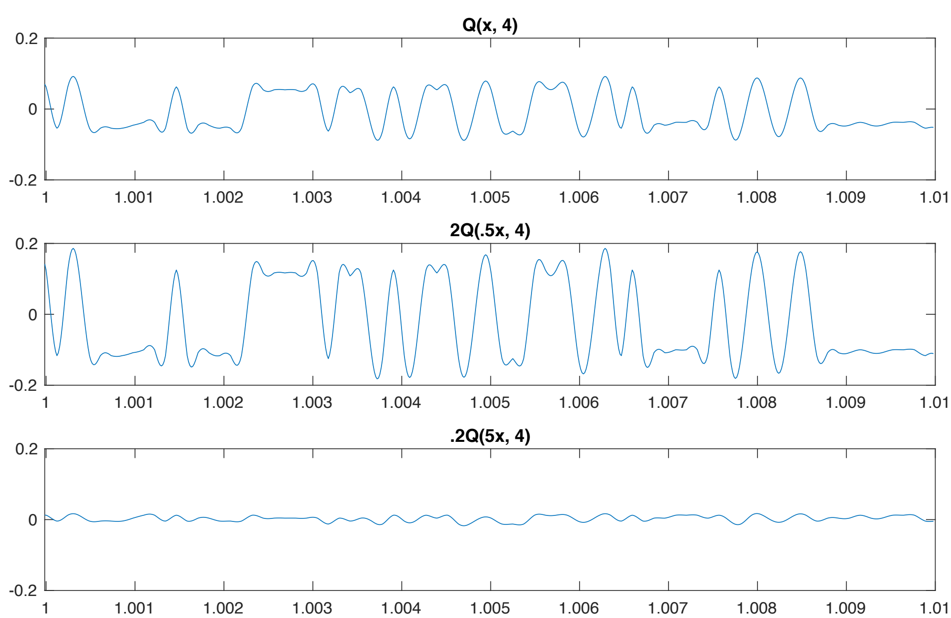

Fig. 8. Error noise for Amplification Gain Control

III.III. Nonlinear Quantization

This section investigates another technique of quantization technique, moving away from the uniform quantization to realm of nonlinearity. To this end, we will augment uniform quantization with a "compander" with $\mu-$law characteristics. The nonuniform quantization using $\mu-$law with $\mu = 255$ and a 4-bit uniform quantization is depicted in Fig. 9.

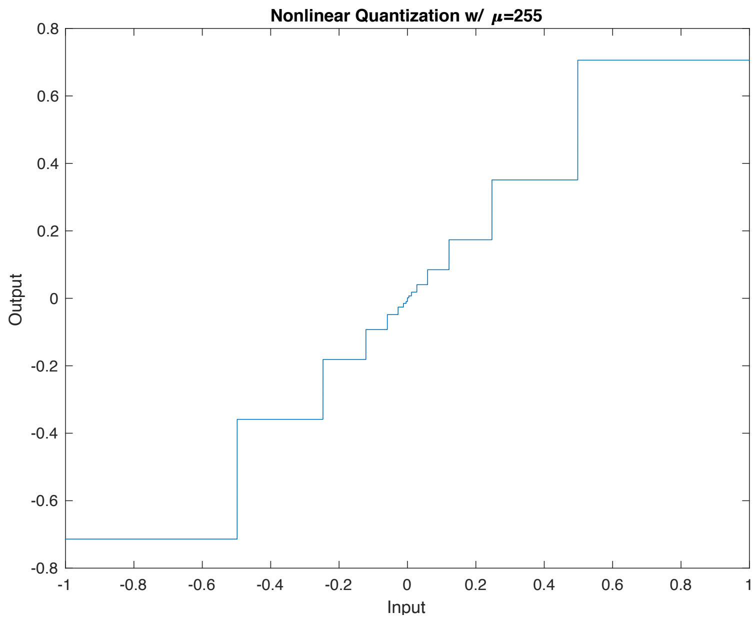

Fig. 9. Nonuniform Quantization using $\mu-$law

We can wee in Fig. 9 that more quantization levels are allocated for amplitudes around zero, and before applying this scheme we can predict that it would improve the performance similar to gain control experiment where we amplified input signal.

Applying the procedure, we compare the the signal (sampled at 8192Hz) with 4-bit uniform quantization and 4-bit nonuniform quantization using $\mu-$law.

Quantization	**Measured SNR wrt x**	**Sound Quality**
$Q(x,4)$	7.0098	Good
$NQ(x, 4, 255)$	13.7196	Very good

Again, as before, we plot the resulting quantization errors. Fig. 10 shows that the nonuniform quantization scheme, as suggested by its SNR, is indeed achieving errors of less magnitude. As we mentioned this is due to the fact that the value of the signal is centered most of the time around the mean value of the signal. That higher resolution around the mean, as achieved with the nonuniform schemes, results in better sounding quantization. From this perspective, a nonuniform quantization scheme outshines a uniform one!

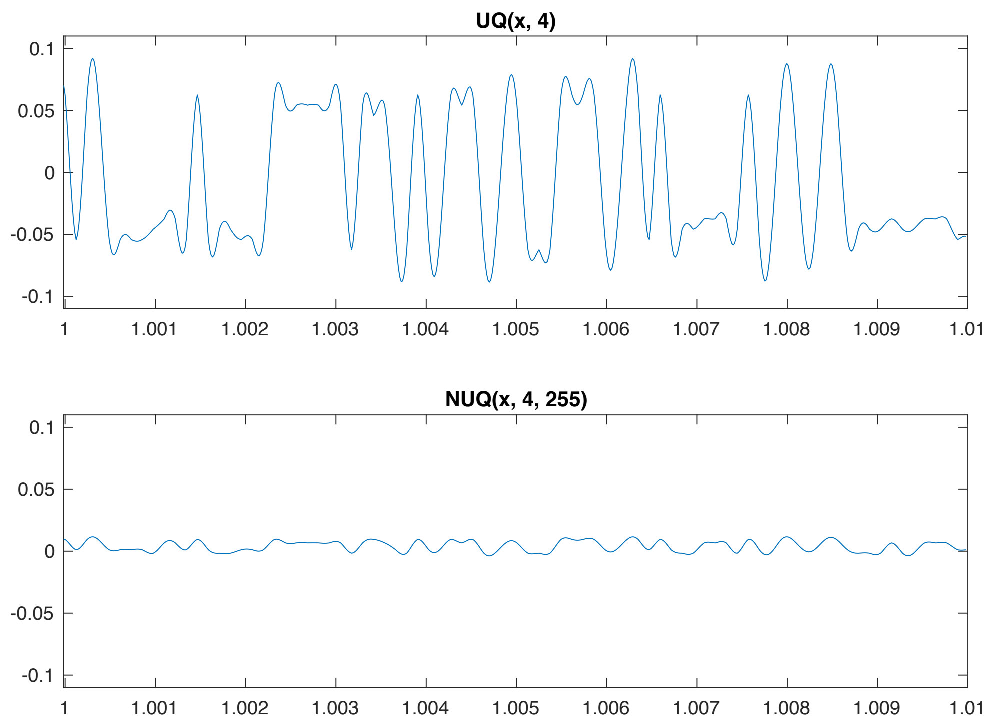

Fig. 10. Nonunform and Uniform Quantization Errors

IV. An Economical System

Assume we want to design a system that would cost 0.25 cents for 1kbps data rate. This is very important to understand the nature of the signal we want to work with. Here, it is assumed that the signal is speech voice recordings. That is, the system may not scale well for instrumental music for example. Note how fixing the output bitrate introduces a trade-off between the number of bits used for quantization and the sampling rate. Although increasing sampling rate decreases the in-band quantization noise, it also reduces number of bits we can use for quantization. However, here we have not been given a cap on the cost per minute. I can complain about this lack of information in the question, as with no constraint and no scarcity, one may come up with a design by willy-nilly choosing very high sampling rate and number of bits for quantization! Based on the characteristic of the signal, one may still argue that regardless of the cost, some configurations still do not make sense. For example, in this article (https://bit.ly/2HAc8Hn), it is claimed that "there is no point to distributing music in 24-bit/192kHz format. Its playback fidelity is slightly inferior to 16/44.1 or 16/48, and it takes up 6 times the space."

We first start with the sampling rate. The human hearing range spans 20Hz to 20kHz. This suggests that according to Nyquist theorem 40kHz sampling rate should be enough. In compact disks they have extended this to 44.1kHz, mainly for practical reasons. But we actually can limit this more, as our own results show. For human voice recordings, it has been characterized to peak around 2.1 KHz and roll off below 300 and over 3000 Hz. The 3kHz bandwidth is enough for voice and 5kHz would give us some margin. Nyquist-Shannon theorem say that we would need twice this amount to use digital signals. As we have seen in our data, a bandwidth of 3.685kHz holds 99.9% of signal energy. Rounding it up, we can settle for 8kHz sampling rate. As we observed during our experiment, higher sampling rates would not bring any more value to the table.

As for the quantization part, we saw that naïvely using uniform sampling could be inferior to a case where we perform proper gain control or nonuniform quantization. A nonlinear relationship might need more sophisticated hardware, but the improvement is indeed noticeable. The decision of using $\mu-$law algorithm for companding in the digital telecommunication systems in North America and Japan also confirms our choice for our system. I choose 4-bit (16 levels) nonuniform quantization using $\mu-$law for my proposed system.

With a sampling rate of 8kHz and 4-bit quantization, we will have a bitrate of 32kbps. This system would cost $0.08 per minute, or $4.8 per hour.

V. Final Thoughts

In this technical report, we examined the impacts of various sampling and quantization schemes. We saw that we could gain higher fidelity with increasing the sampling rate and quantization levels. We empirically observed how the Nyquist sampling rate, using the 99.9% energy bandwidth, can guide us in choosing the proper sampling rate. Moreover, we saw that we can improve a "naïve" uniform quantization technique, with the same number of quantization bits, by either proper amplifier gain control, or even better, with nonuniform quantization using companding as we examined the $\mu-$law. Upon performing different transformations to the signal, we learned that calculating the SNR and plotting the noise signal are indispensable tools. Assessing the transformation based on subject perception is also important, and we saw how a theoretical measure can be an indicator of empirical ones (as with SNR). At last we proposed a proper combination if sampling rate and quantization configuration that could give rise to a cost-effective solution.

VI. Apprendix (MATLAB Code)

```
% Nima Mohammadi
% Design Project 1

% Note: You need to comment out command the 'figure'
% command inside `EnergySpectralDensity.m` function

%% Load dataset
load DesignProject1
duration = length(Original) / 65536;
%% Plot of original quantization
plot(sort(Original));
title('Approximation of original quantization scheme');
xlabel('Input');
ylabel('Output');

%% Plot time wave
plot(time, Original);
title('Original Signal');
xlabel('Time');
ylabel('Amplitude');

%% Plot ESD (Energy Spectral Density)
EnergySpectralDensity(Original, 65536, [-6000 6000 0 2000000], 0);
figure;
EnergySpectralDensity(Original, 65536, [-6000 6000 -60 1], 1);
ylabel('Magnitude Spectrum (log)')

%% Calculate bandwidth of the original signal
bw_orig = CalculateBandwidth(Original, 65536, 0.999)

%% Plots for each sampling rate
sampling_freq = {2048, 4096, 8192, 16384};
sampled_signals = cell(1, 4);
sampled_bw = cell(1, 4);
plot_sig_orig = Original(65536*2: 65536*2+65536/64-1);
plot_t_orig = linspace(2, 2 + length(plot_sig_orig)/65536, length(plot_sig_orig));

figure;
set(gcf,'position',[0,0,500,1000])
for i = 1:length(sampling_freq)
    sampled_signals{i} = MyResample(Original, sampling_freq{i}, 65536);
```

```
    sampled_bw{i} = CalculateBandwidth(sampled_signals{i}, sampling_freq{i}, 0.999);
    subplot(length(sampling_freq), 3, 3*i-2);
    plot_sig = sampled_signals{i}(sampling_freq{i}*2: sampling_freq{i}
*2+sampling_freq{i}/64-1);
%     length(plot_sig)
    plot_t = linspace(2, 2 + length(plot_sig)/sampling_freq{i}, length(plot_sig));
    plot(plot_t, plot_sig);
    hold on
    plot(plot_t_orig, plot_sig_orig);
    legend('x(t)', 'x_s(t)');
    xlim([2 plot_t(end)]);
    title(sprintf('f_s=%d', sampling_freq{i}));
    xlabel('Time');
    ylabel('Amplitude');

    subplot(length(sampling_freq), 3, 3*i-1);
    EnergySpectralDensity(sampled_signals{i}, sampling_freq{i}, [0 bw_orig -60 1], 1);
    ylabel('Mag Spec')
    title(sprintf('f_s=%d, BW=%.3fHz', sampling_freq{i}, sampled_bw{i}));

    subplot(length(sampling_freq), 3, 3*i);
    [Y_tmp, f_tmp] = EnergySpectralDensity(sampled_signals{i}, sampling_freq{i});
    ylabel('Mag Spec')
    title(sprintf('f_s=%d, BW=%.3fHz', sampling_freq{i}, sampled_bw{i}));
    ylim([0 max(abs(abs(Y_tmp).^2))]);
    tmp = xlim;
    xlim([0 tmp(2)]);
end

%% Listening to sampled signals
i = 3; % change i from 1 to 4 to hear different SRs
sound(sampled_signals{i}, sampling_freq{i})
% sound(Original, 65536);

%% Calculate SNR for sampled signals
sampled_snr = cell(1, 4);
for i = 1:length(sampling_freq)
    sampled_snr{i} = snr(Original, Original-interp(sampled_signals{i}, 65536/
sampling_freq{i}));
end
sampled_snr

%% Plots different uniform quantizations
```

```
quant_levels = {2, 8, 32, 128, 512, 1024};
x = -1:.0001:1;
figure;
for i = 1:length(quant_levels)
    subplot(length(quant_levels)/3, 3, i)
    plot(x, uniformquantize(x, quant_levels{i}));
    title(sprintf('Uniform Quant levels=%.0f', quant_levels{i}))
end

%% Quantization
quant_signals = cell(1, 6);
figure;
for i = 1:length(quant_levels)
    subplot(length(quant_levels)/3, 3, i)
    quant_signals{i} = uniformquantize(Original, quant_levels{i});
    plot(time, quant_signals{i});
    title(sprintf('Uniform Quant levels=%.0f', quant_levels{i}))
    xlabel('Time');
    ylabel('Amplitude');
end

%% Listening to quantized signals
i = 6; % change i from 1 to 6 to hear different quantizations
sound(quant_signals{i}, 65536)
% sound(Original, 65536);

%% Calculate SNR for quantized signals
quants_snr = cell(1, 6);
theoretical_snr = cell(1, 6);
for i = 1:length(quant_levels)
    quants_snr{i} = snr(Original, Original-quant_signals{i});
    bits = log(quant_levels{i}) / log(2);
    theoretical_snr{i} = 1.8 + 6 * bits;
end
quants_snr
theoretical_snr

%% Plotting Quantization Errors
quant_signals_noise = cell(1, 6);
figure;
for i = 1:length(quant_levels)
    subplot(length(quant_levels)/3, 3, i)
    quant_signals_noise{i} = Original - quant_signals{i};
```

```
    plot(time(65536:65536*1.01), quant_signals_noise{i}(65536:65536*1.01));
    ylim([-.5 .5]);
    title(sprintf('Uniform Quant levels=%.0f', quant_levels{i}))
    xlabel('Time');
    ylabel('Error Amplitude');
end

%% Uniform Quantization with Amplifier Gain Control
figure;
x = MyResample(Original, 8192);
% sound(x, 8192);
y = uniformquantize(x, 16);
y_err = Original-interp(y, 65536/8192);
snr(Original, y_err)
subplot(3, 1, 1);
plot(time(65536:65536*1.01), y_err(65536:65536*1.01))
title('Q(x, 4)');
ylim([-.2 .2])
% sound(y, 8192);
z = 2*uniformquantize(.5*x, 16);
z_err = Original-interp(z, 65536/8192);
snr(Original, z_err)
subplot(3, 1, 2);
plot(time(65536:65536*1.01), z_err(65536:65536*1.01))
title('2Q(.5x, 4)');
ylim([-.2 .2])
% sound(z, 8192);
w = .2*uniformquantize(5*x, 16);
w_err = Original-interp(w, 65536/8192);
snr(Original, w_err)
subplot(3, 1, 3);
plot(time(65536:65536*1.01), w_err(65536:65536*1.01))
title('.2Q(5x, 4)');
ylim([-.2 .2])
% sound(w, 8192);

%% Plotting mu-law Nonuniform Quantization
figure;
plot(-1:.0001:1, expand(uniformquantize(compress(-1:.0001:1, 255), 16), 255));
xlabel('Input');
ylabel('Output');
title('Nonlinear Quantization w/ \mu=255');
```

```
%% Nonuniform quantization of signal using mu-law
figure;
x = MyResample(Original, 8192);
y = compress(x, 255);
x_q = uniformquantize(x, 16);
y_q = uniformquantize(y, 16);
z = expand(y_q, 255);
% sound(x_q, 8192);
% sound(z, 8192);

x_err = Original-interp(x_q, 65536/8192);
snr(Original, x_err)
subplot(2, 1, 1);
plot(time(65536:65536*1.01), x_err(65536:65536*1.01))
title('UQ(x, 4)');
ylim([-.11 .11])

z_err = Original-interp(z, 65536/8192);
snr(Original, z_err)
subplot(2, 1, 2);
plot(time(65536:65536*1.01), z_err(65536:65536*1.01))
title('NUQ(x, 4, 255)');
ylim([-.11 .11])
```