EEE 6504 Adaptive Signal Processing, Project 1

Mark Rosenberg

Abstract—**An interference canceling filter is adapted with the NLMS algorithm to remove noise from speech. The effect of the two free parameters, filter order and step size, on performance is evaluated. The result is a signal with a high SNR.**

Index Terms—**Adaptive Interference Canceling, NLMS**

I. INTRODUCTION

A speech signal is recorded at the primary microphone in the presence of interference from a noise source n, which is recorded both at the reference microphone and at the primary microphone. An interference cancelation circuit is constructed with the reference signal as the adaptive filter input and the primary as the desired signal. Consider a delay t_0 until the primary microphone receives the n_0 interference signal and a delay t_1 until the reference microphone receives the n_1 interference signal, both from the source n. Then, ideally, the adaptive filter would represent a delay system from the reference microphone to the primary microphone characterized by the time difference t_2, which satisfies a triangle inequality with t_0 and t_1 and in general is not an integer multiple of the sampling period. If the delay were an integer multiple of the sampling period, the impulse response from the ideal weights of the adaptive filter would isolate from the tap-input vector the sample corresponding to the appropriate delay if the filter order were large enough. Otherwise, for an ideal delay system, the impulse response resembles a shifted sinc function for interpolating the noise signal. If it were an ideal delay system, the distance between the primary and reference nodes could be calculated using the sampling frequency and the speed of sound in air. In this problem, however, $n_1 \neq n_2$, so the purpose of the adaptive filter is to use the correlation between the filter input and the desired signal so that the error signal is close to the speech signal

$$e(k) = s(k) + (n_0(k) - y(k)) \approx s(k), \tag{1}$$

where y is the filter output. If the reference and primary microphones are moving, then the ideal weights of the adaptive filter will be time varying, as the delay and distance between the two nodes varies. Since the reference microphone is closer to the noise source than is the primary microphone, it will receive the noise signal first, and the adaptive filter representing the recording delay from the reference to the primary will be causal since the delay is positive.

II. METHODS

The goal is to minimize the noise power in the error signal. First, Equation 1 is squared and the expected value of both sides is taken. The cross term vanishes because the signal and noise are assumed to be uncorrelated, which is a fair assumption for this problem, though there is some signal leakage in the reference microphone. The squared signal term is independent of the weights, so minimizing the cost of squared error is equivalent to minimizing the $(n_0 - n_1)^2$ term, our goal.

The learning curve is the MSE over time, which is the value of the performance surface over time as the weights track on the abscissa dimensions. For the transient section of time, the weights evolve from the initial condition to the optimal weights that characterize the system, and they do so by an update equation of the NLMS algorithm, which is based on gradient descent. The step size μ characterizes the speed of convergence, and is responsible for how quickly the weights reach the optimal weight neighborhood from the initial condition, and for how high the average error is between the approximation of the system and the true system, once the weights are in the optimal neighborhood. The learning curve is best accompanied by a plot of the weight tracking to show how they reach a steady state at the same time the learning curve values reflect the minimum neighborhood of the MSE performance surface. In this problem, the error approximates the signal, so the learning curve approximates the squared signal.

III. RESULTS

A. Filter Order Two

The simplest way to visualize the performance surface is as contours in 2 dimensions, which corresponds to a filter with 2 taps. In order to plot the contours of the performance surface, an array of 2-dimensional tap-weight vectors are used to filter the reference, compute the error signal, and the MSE over time is the value of the performance surface. Figure 1 shows the effect of the remaining free parameter, step size μ, on the MSE. The optimal solution $\mu = 0.0011$ minimizes the MSE for $M = 2$. For optimal μ, the tap-weight vector evolves as in Figure 2 from the zero initial condition to a time-varying system whose output approximates the interference noise in the primary node. Figure 3 shows the same as Figure 2, except represented as a vector and with the contours of the performance surface. Since NLMS is derived from steepest descent, the path of the tap-weight vector is perpendicular to the contours toward the ellipse center. This is shown both from the initial condition and from large perturbations in the weights due to the weight update equation's dependency on the instantaneous sample for approximating the gradient as well as the change in the noise around $4.7 \times 10^4 s$. Figure 4 shows the learning curve, which approximates the squared speech signal; if the only component of the error signal were the quantity to be minimized, it would approach zero.

M. Rosenberg is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA e-mail: markrosenberg@ufl.edu

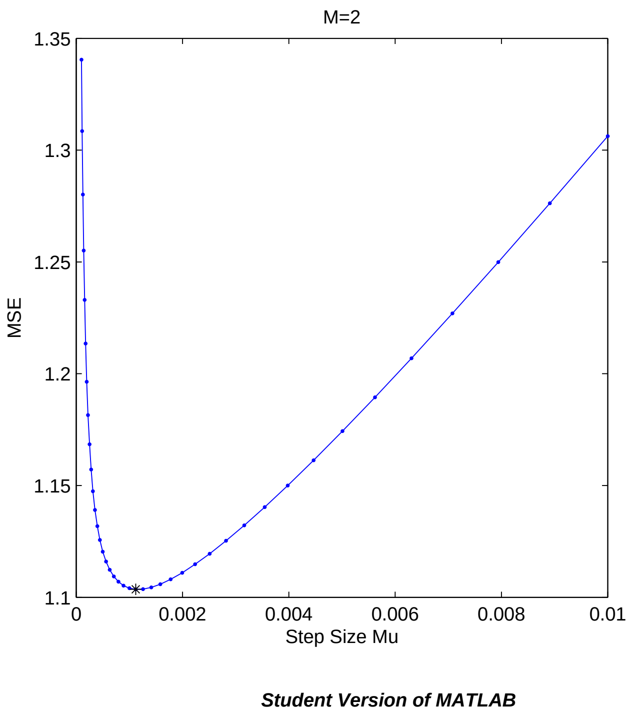


Fig. 1. MSE vs Step Size μ for Filter Order $M = 2$

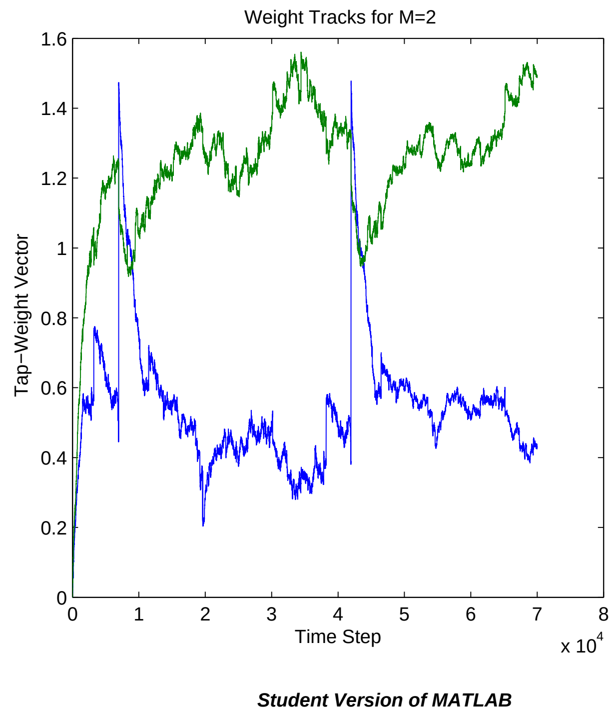


Fig. 2. Tap-Weight Vector over Time for $M = 2$

B. Higher Order Filter

Two taps may not be sufficient to characterize the ideal interference canceling filter. For $M = 3$ through $M = 15$ and for $\mu = 10^{0:-0.3:-3}$, the NLMS algorithm was executed, and the MSE vs $(M, log(\mu))$ was plotted in Figure 5. This plot shows that there are diminishing returns for $M > 10$. For higher filter orders, the adapted filter's impulse response for the tap weights greater than 10 is approximately zero. For this reason, $M = 10$ was selected to show how increased filter order improves performance. Figure 6 shows MSE vs μ with optimal step size $\mu = 0.2818$ for $M = 10$. With that optimal step size, NLMS was executed and the tap-weight vector tracks are shown in Figure 7. The figure shows that the weights adapt to disturbances and initial conditions quickly and then fluctuate about those values. Higher fluctuations arise from larger step sizes, yet reducing the step size slows adaptation. The fluctuations are related to the misadjustment, discussed in Section III-C. Figure 7 implies that the ideal filter is relatively fixed with different values before and after $k = 4.7 \times 10^4 s$. From this observation, in order to improve the speech signal at the expense of longer adaptation speed, which is not detrimental during the regions with relatively fixed ideal weights, we can reduce the step size. Figures 8-9 show this effect for $\mu = 0.01$ and $\mu = 0.001$, respectively. The clarity of the speech in the error signal is improved for majority of the time; however, as shown in Figure 6, this does still raise the MSE over the entire duration. Furthermore, using $\mu = 0.001$ produces an error signal where the interference noise has a longer transient period, and is audibly noticeable.

C. Misadjustment

The misadjustment is defined as the ratio of the steady state excess MSE over the minimum MSE. The minimum MSE is the value of the performance surface at the optimal Wiener solution, which occurs at the center of the ellipse contours. Steady state excess MSE is the expected value of the learning curve at convergence minus the minimum MSE. There is an ideal time-varying system that the adaptive filter approximates, so the optimal weights are time-varying. With time-varying ideal weights and a non-stationary environment, the idea of convergence is with respect to the ideal time-varying weights, and the idea of steady state may in this problem refer to sections of time where the ideal weights are relatively constant and the filter weights are in that neighborhood. Higher step size gives a higher misadjustment because in the neighborhood of the minimum, larger steps are more likely to overshoot the minimum, resulting in a higher value in the performance surface. In fact, executing a second trial where the initial condition is a tap-weight vector from a first NLMS trial with zero initial condition, as well as using a small step size, resulted in smaller fluctuations around the optimal weights (until the ideal weights changed since they're time-varying). Then the tradeoff between adaptation speed and average fluctuations around the ideal weights was observed. This process can be repeated to achieve more accurate initial conditions. Since the ideal weights are only relatively constant over a subset of the signal duration, the misadjustment in that region can be approximated by scanning the tap-weight space until obtaining the tap-weight vector with the minimum MSE. The learning curve is then averaged over time and divided by the minimum

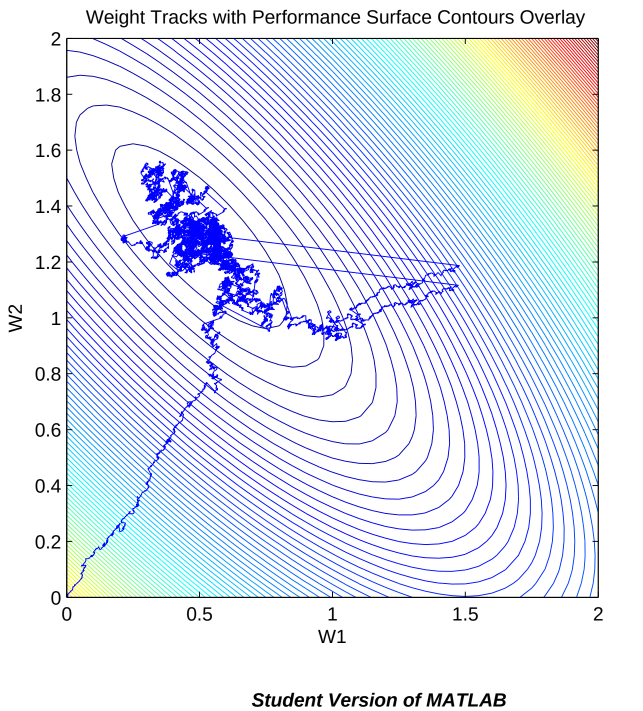


Fig. 3. Weight Tracks with Performance Surface Contours Overlay for $M = 2$

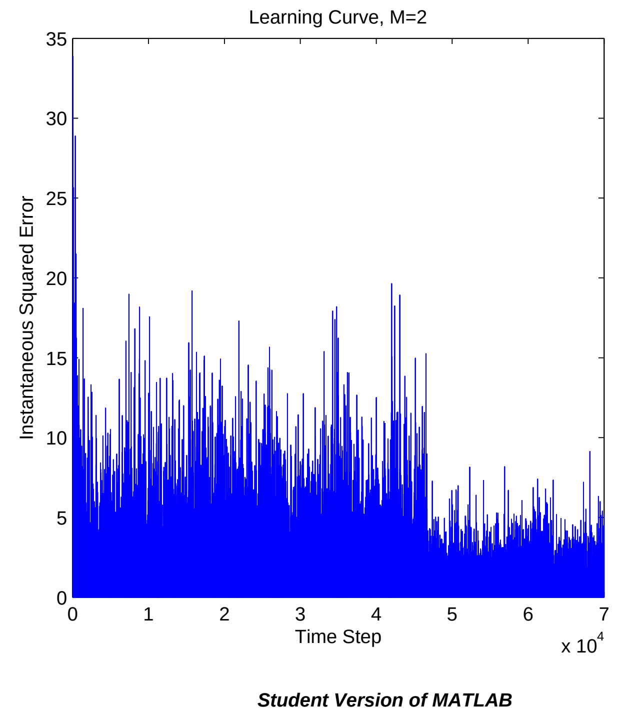


Fig. 4. Learning curve for $M = 2$

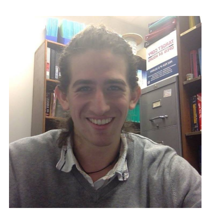
MSE. For higher dimensions, the computational complexity of scanning the tap-weight vectors grows. However, from the results of NLMS, we know the neighborhood of the ideal tap-weight vector. One approach would be to average the tap-weight vectors over the time the ideal tap-weight vector is assumed to be constant and NLMS has converged. This would approximate the minimum MSE. From this approach over time $[1000, 40000]$, the misadjustment was calculated to be 0.5212 for the optimal $mu = 0.2818$ at $M = 10$.

D. Frequency Response

The ratio of the spectra of two time series is the frequency response from one to the other. If the time series are input and output of a filter, the transfer function of the filter evaluated at $e^{j\omega}$ is the frequency response. The adaptive filter transforms the interference noise n_1 to the output y, which is an approximation of n_0; however, this system does not relate the error and desired signals to each other. The FFT is used to estimate the spectra of the error and desired signals. The magnitude of the frequency response from the desired to the error for $M = 2$ is shown in Figure 10 and for $M = 10$ is shown in Figure 11. The SNR improvement in dB is given by the echo return loss enhancement, $ERLE = 10log(E(d^2)/E(e^2))$. For $M = 2$, $ERLE = 5.8$ and for $M = 10$, $ERLE = 19.0$. Recall from Figure 5 that for larger filter order M, the MSE is smaller, so it follows that the ERLE will be greater, and thus the results validate the theory.

IV. Conclusion

The result of NLMS is a signal with a high SNR, and is recognizable as a line from the movie, "The Phantom Menace." It is important to consider that although it's well-defined to solve the optimization problem for the optimal step size μ, higher quality speech in the error signal was obtained for a smaller step size. In this problem domain, it was more important to consider that brief intervals of low SNR from either initial transience or rapid change in the ideal tap-weight vector were preferable to a minimal MSE over the entire signal duration because of the idea that below a threshold SNR, inaudible speech can't become further inaudible. Perhaps minimizing a cost function equal to the minimum of the squared error and some upper threshold would result in a smaller optimal step size; however since NLMS is derived from the structure of the cost function, future work would investigate whether this adjustment leads to a similar algorithm.



Mark Rosenberg is pursuing his Ph.D. in Electrical Engineering at the University of Florida in the Computational NeuroEngineering Lab. He received his B.S. degree in Electrical Engineering from Washington University in St. Louis, St. Louis, MO, USA in 2012, and his M.S. degree in Electrical Engineering from the University of Florida, Gainesville, FL, USA in 2014.

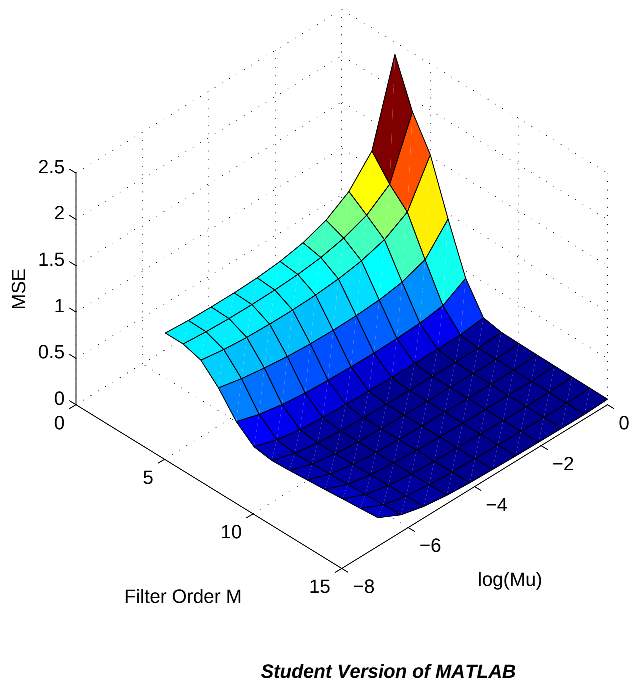


Fig. 5. MSE vs Filter Order M and Log of Step Size μ

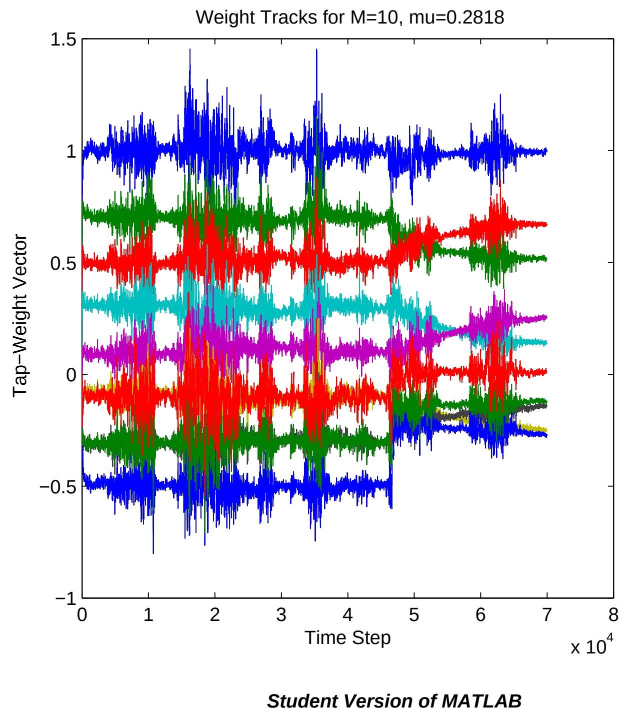


Fig. 7. Weight Tracks for $M = 10$

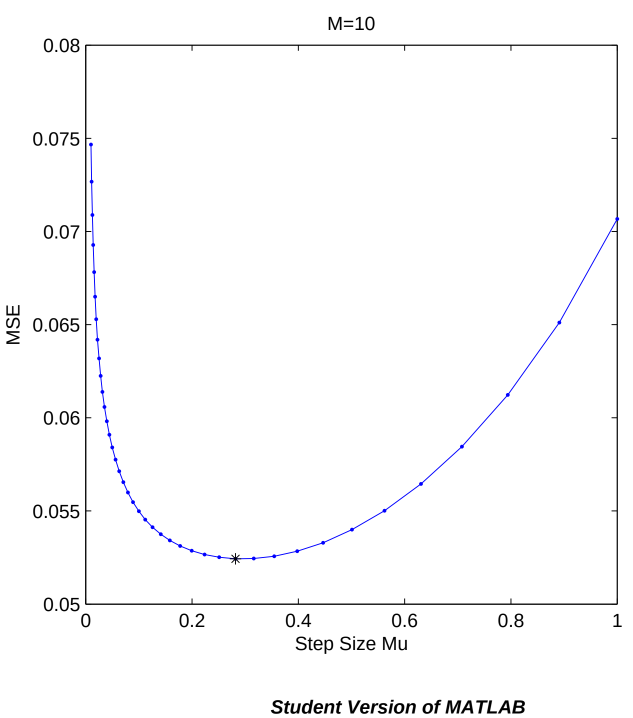


Fig. 6. MSE vs μ for $M = 10$

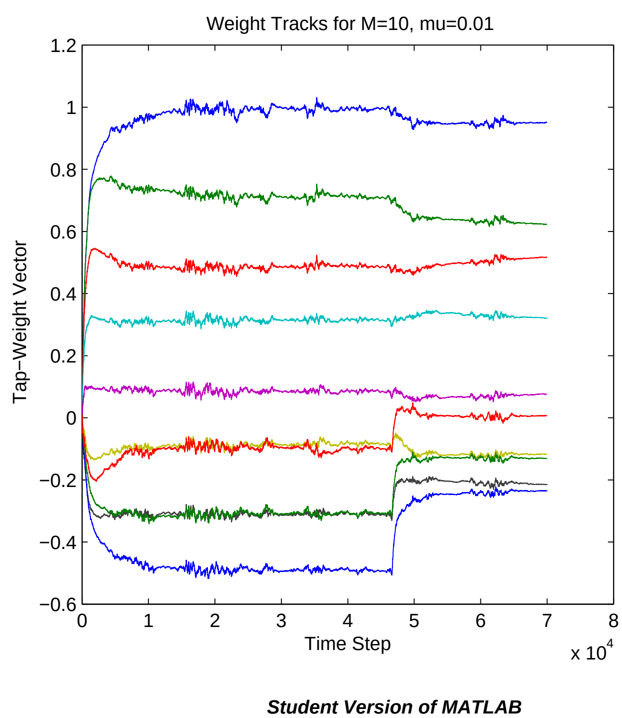


Fig. 8. Weight Tracks for M=10, $\mu = 0.01$

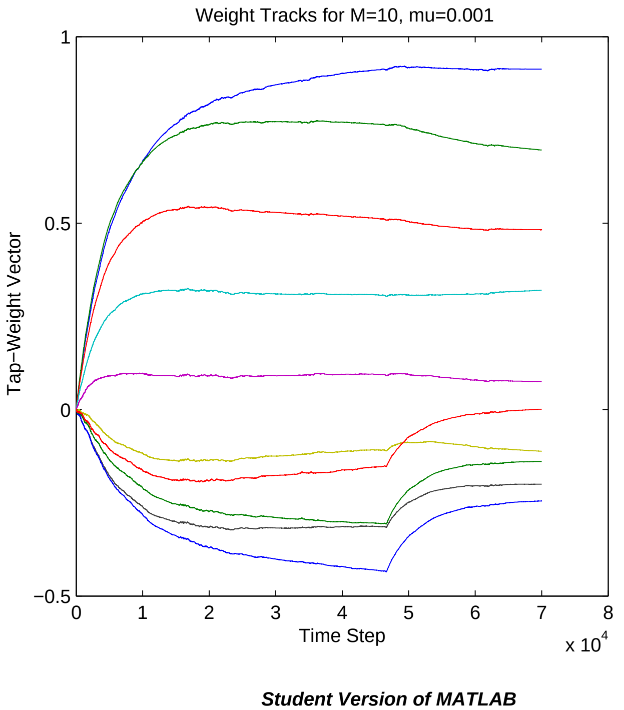


Fig. 9. Weight Tracks for M=10, $\mu = 0.001$

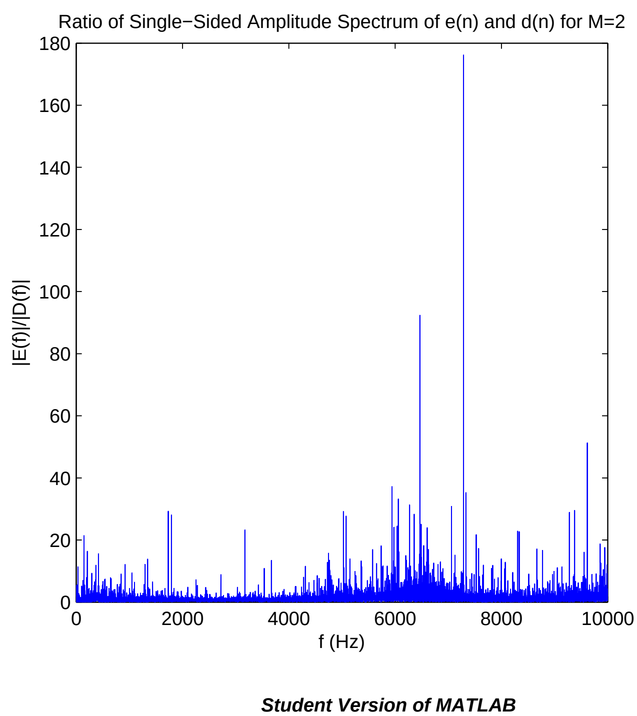


Fig. 10. Frequency response from d to e for $M = 2$

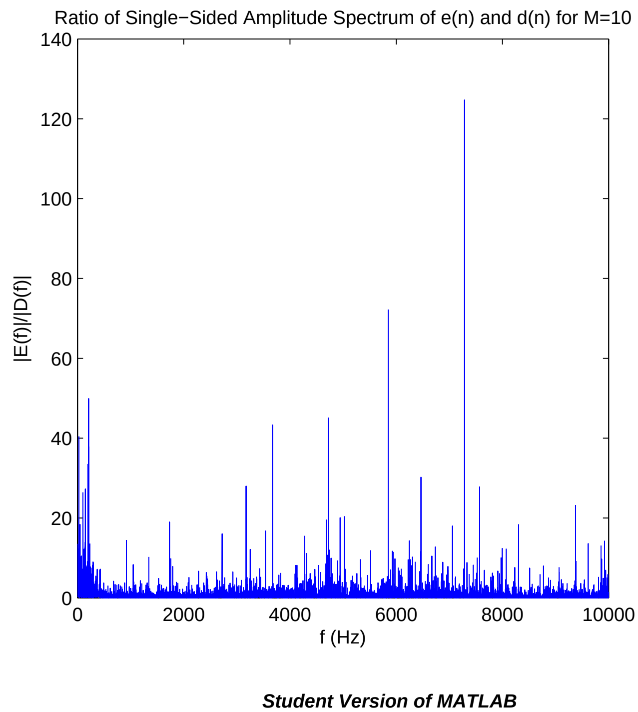


Fig. 11. Frequency response from d to e for $M = 10$