Speech Filtering for Improving Intelligibility in Noisy Transients

by Andrew Lewine

B.S. M.I.T., 2010

Submitted to the Department of Electrical Engineering and Computer Science in Partial Fulfillment of the Requirements for the Degree of Master of Engineering in Electrical Engineering and Computer Science at the Massachusetts Institute of Technology

ARCHIVES

[June 2011]

May 2011

Author _____

Department of Electrical Engineering and Computer Science

May 20, 2010

Certified by _____

Dr. Rahul Sarpeshkar

Associate Professor, Department of Electrical Engineering and Computer Science

Thesis Supervisor

Accepted by _____

Dr. Christopher J. Terman

Chairman, Masters of Engineering Thesis Committee

1

Speech Filtering for Improving Intelligibility in Noisy Transients

by Andrew Lewine

B.S. M.I.T., 2010

Submitted to the Department of Electrical Engineering and Computer Science

May 19, 2011

ABSTRACT

Hearing impairment is a problem that affects a large percentage of the population. Cochlear implants allow those with profound or total hearing loss to regain some hearing by stimulating auditory nerve fibers with implanted electrodes, in response to sound picked up by an external microphone. The signal processing chain from microphone input to stimulation output is an important factor in the overall speech intelligibility of the implant system. This thesis work improves on an existing ultra-low-power cochlear implant system by utilizing an improved noise and power efficient bandpass filter bank to implement a novel frequency-selective gain control algorithm capable of reducing, and in some cases removing, loud transient noises, thereby improving speech intelligibility. This gain control algorithm takes advantage of the inherent frequency-specific gain control afforded by the improved bandpass filter topology. This contribution makes an improvement to the existing state-of-the-art system in both power efficiency and performance.

Thesis Supervisor: Dr. Rahul Sarpeshkar

Title: Associate Professor, Department of Electrical Engineering and Computer Science

Acknowledgments

My sincerest thanks go to Prof. Rahul Sarpeshkar, the AVBS group, my family, and my dear Emily.

Contents

List of Figures

1 Introduction

This document serves as the Masters of Engineering thesis project writeup. The aim of this project was to explore ultra-low-power architectures for speech filtering for use in cochlear implants. This project was two-pronged. Part of it focused on utilizing a low-power sub-threshold complementary metal oxide semiconductor (CMOS) bandpass filter to implement speech filtering algorithms to improve the intelligibility of speech in noisy environments. The other part focused on creating a proof-of-concept cochlear implant system utilizing several integrated circuits (ICs) designed by other members of the Analog and Biological VLSI Systems group.

2 Cochlear Implants

Cochlear implants, or bionic ears as they are sometimes referred to, are electronic devices that restore hearing to people with severe sensorineural hearing loss. Candidates for cochlear implants span from young children who were born deaf, to elderly people who have lost hearing as an effect of illness or disease. A complete cochlear implant system consists of several parts. First, a microphone and external processing unit pick up sound and speech from the environment and perform signal processing on the data in order to prepare it for transmission into to skull of the patient. Next, this data is wirelessly streamed into the implanted unit using some form of telemetry; radio frequency (RF) transmission using a pair of wire coils is a popular choice. Finally, the sound and speech data is used to trigger electrodes that are in contact with the patient's auditory nerve fibers. The electrical stimulation of the auditory nerve fibers send data directly to the brain via the auditory nerve system where it can be processed and interpreted by the recipient as meaningful sound. In order for a cochlear implant to be successful, the recipient must still have functioning auditory nerve fibers. This requirement excludes some hearing impaired people from using cochlear implants.

Cochlear implants have been successfully used for decades to restore hearing to the hearing impaired[1]. However, they typically employ digital speech analyzers that are only moderately energy efficient. This all-digital approach maximizes flexibility and processing capability (as the external unit is in essence a small computer which can be reprogrammed at will) at the expense of power consumption. Typical

---

[1] "History of Cochlear Implants." n.d. Web. 4 Mar. 2010.
<http://biomed.brown.edu/Courses/BI108/BI108_2001_Groups/Cochlear_Implants/history.html>.

battery lives of these systems are only on the order of several days. This requires the implant user to recharge the system daily in order to ensure continuous operation. While daily recharges are not in and of themselves a death knell for these systems, maximum battery lifespan is. Batteries degrade in capacity with each recharge cycle; even the most advanced batteries currently in production are only rated to around one thousand charge-discharge cycles[2]. Sanyo currently produces lithium polymer batteries whose cycle discharge curve indicates that on average, one of these batteries has dropped to 80% of its initial capacity after only 500 charge-discharge cycles[3]. These best-of-the-best batteries are not suitable for medical applications; they are much too large and do not have the required FDA certifications. Even if they did, however, they are still not good enough to enable multi-month recharge-free operation of non-power-efficient cochlear implants. Reducing the power consumption of cochlear implant system allows room for making trade-offs in other realms. For example, with lower power consumption, a cochlear implant system could use a smaller capacity (and therefore smaller size) battery while retaining the same time between recharges. Or, with the same capacity battery, the time between recharges and thus the maximum battery lifespan could be increased. Looking to the future, as cochlear implant systems become more advanced and employ more sophisticated processing algorithms, a lower-power system could pack more total effective processing power into the same area before crossing the biological limits on the maximum safe temperature increase of the tissue inside the brain. Needless to say, lowering the power consumption of cochlear implant systems opens the doors to an array of improvements in other related metrics.

This project leveraged an existing cochlear implant sound processing integrated circuit (IC). This chip, henceforth referred to as the Analog Bionic Ear Processor[4] (ABEP), incorporates a microphone preamp, sound processing circuits, and an analog to digital converter. In a complete system, the digital bits output from the ABEP would be fed to a wireless telemetry unit to stream the sound and speech data to the internal stimulation unit. The ABEP takes several design cues from biological systems and is partly analog rather than fully digital. The analog-digital hybrid architecture allows for an ultra-low-power design and increased efficiency.

---

[2]"Core Technology." A123 Systems. n.d. Web. 13 March 2011. <http://www.a123systems.com/technology-core.htm>.

[3]"Sanyo Lithium Polymer Rechargeable Batteries." Sanyo. n.d. Web. 24 Feb. 2011.<www.sanyo.com.hk/battery/en/shared/file/lipolymerE.pdf>.

[4]R. Sarpeshkar, C. Salthouse, J.J. Sit, M. Baker, S. Zhak, T. Lu, L. Turicchia, and S. Balster, "An Ultra-Low-Power Programmable Analog Bionic Ear Processor," IEEE Transactions on Biomedical Engineering, Vol. 52, No. 4, pp. 711-727, April 2005.
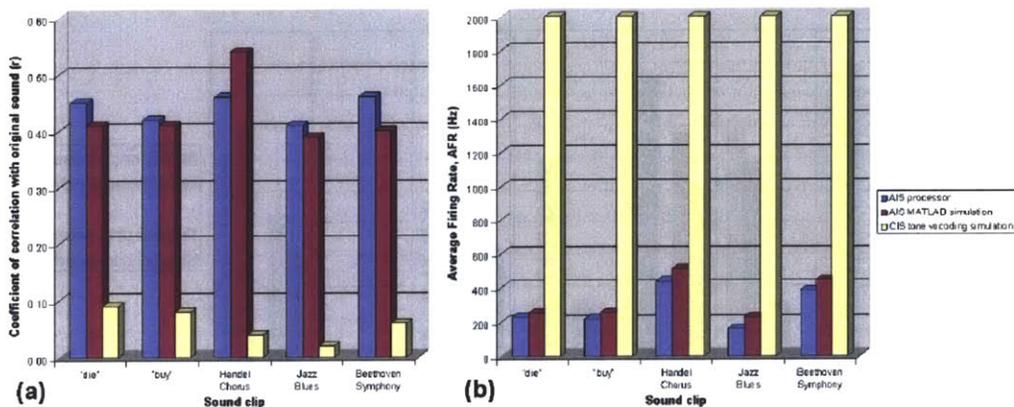
Figure 3.1 – Comparison of AIS and CIS performance

3 Analog Bionic Ear Processor

The current state-of-the-art in ultra-low-power cochlear implant speech processing units is the Analog Bionic Ear Processor developed and tested by the Analog VLSI and Biological Systems group at MIT. This speech processor uses two novel techniques to reduce power consumption and improve performance over typical cochlear implant speech processors. The first is the Asynchronous Interleaved Sampling[5] (AIS) algorithm. Compared to the standard Continuous Interleaved Sampling (CIS), AIS reduces stimulation power consumption by performing a winner-takes-all contest between the stimulation channels. This technique allows high-intensity stimulation channels to fire frequently, while reducing the firing rate of low-intensity channels. This mimics the cochlea's behavior which exhibits intensity masking; that is, a loud sound will mask the presence of a simultaneous soft sound. In order to allow low-intensity channel to fire at all, recent winners are handicapped from winning in the near future. Figure 3.1 shows both the correlation coefficients between the simulation pattern and the original sound as well as the average stimulation firing rate. Compared to CIS, AIS both produces more highly correlated stimulation pulses as well as reduces the average firing rate, reducing stimulation power consumption.

The ABEP also employs a biology-inspired companding algorithm in order to improve speech intelligibility in noisy environments. As compared to standard companding algorithms[6], this novel companding algorithm mimics the behavior of the cochlea by modeling spectral masking. Spectral masking is a phe-

---

[5]J. Sit, A. M. Simonson, A. J. Oxenham, M. A. Faltys, and R. Sarpeshkar, "A low-power asynchronous interleaved sampling algorithm for cochlear implants that encodes envelope and phase information", IEEE Transactions on Biomedical Engineering, Vol. 54, pp. 138-149, 2007.

[6]Yang, J., Luo, F., and Nehorai A. (2003). "Spectral contrast enhancement: Algorithms and comparisons," Speech Comm. vol. 39, 33-46.
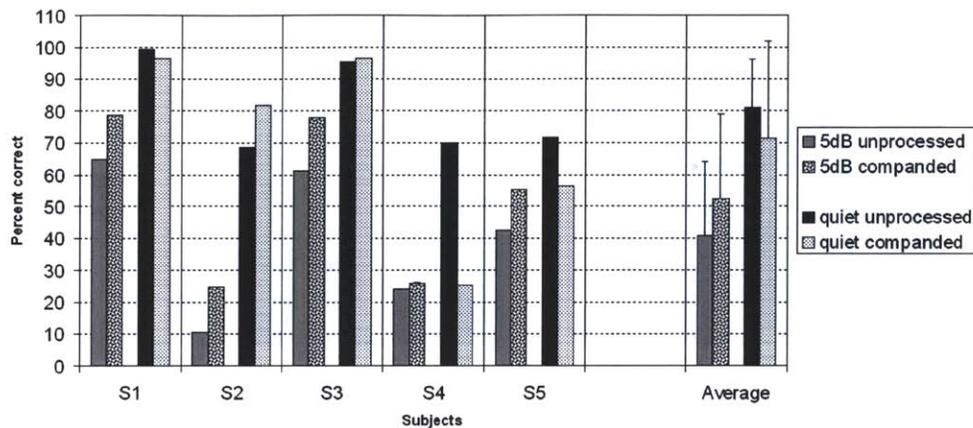
Figure 3.2 – The performance of hearing-impaired users in a speech intelligibility study using the novel companding algorithm. The sentences were either preprocessed (dotted bars) by the companding algorithm or left unprocessed (solid bars). Error bars indicate standard errors of the mean.

nomenon by which a slightly higher-intensity sound masks the presence of another sound that is close in frequency to the first. By modeling this process and mimicking it, the ABEP is able to improve speech intelligibility in noisy environments over standard speech processors. This speech processor was tested by cochlea implant patients[7] for a day and was compared to a standard vocoding-based speech processor[8]. Figure 3.2 shows the results of the study. Average correct scores for hearing-impaired patients rose from 41% to 53%.

This thesis work improves on the existing ABEP chip in two important domains: power consumption and speech intelligibility. The power consumption has been reduced through the implementation of a new ultra-low-power sub-threshold CMOS bandpass filter proposed by Arfin et. al.[9] The speech intelligibility has been improved through the development of new filtering algorithms that are conducive to translation into low-power circuit architectures.

---

[7]A. Bhattacharya and F. G. Zeng, "Companding to improve cochlear-implant speech recognition in speech-shaped noise", Journal of the Acoustical Society of America, Vol. 122, No. 2, pp. 1079—1089, 2007.

[8]Dudley, H. (1939). "The Automatic Synthesis of Speech," Proc Natl Acad Sci USA. vol. 25, 377–383.

[9]Adaptively Biased Bandpass Filters: S. K. Arfin, S. Mandal and R. Sarpeshkar, "Dynamic-Range Analysis and Maximization of Micropower Gm–C Bandpass Filters by Adaptive Biasing," Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Taipei, Taiwan, May 2009, pp. 2954-2957

4 Power- and Noise-efficient Bandpass Filter

The sound processing that must take place in order to translate sounds and speech picked up by the microphone of a cochlear implant in to electrical stimulation pulses of the auditory nerve fibers is in its simplest form a Fourier Transform. In a hearing-capable person, the human cochlea performs this function itself. However, when hearing damage exists, the cochlea can longer perform this essential operation. Shaped like a coil, the cochlea has hair fibers lining its inside. These hair fibers react to different frequency bands depending on their location inside the cochlea. The further within the spiral a hair occurs, the lower the frequencies to which it responds. In this way, the cochlea performs a simple frequency-to-spacial mapping. In order to create a successful cochlear implant, the sound processing that is performed must replicate this function. The standard procedure for all-digital systems is to perform a computation-intensive Fourier Transform. This power spectral data is then used to stimulate electrodes placed around the length of the cochlea in order to mimic the frequency-to-spacial mapping of a working cochlea. For systems that employ digital processors, the Fourier Transform method is adequate. However, for ultra-low-power systems, this method is not power efficient enough to meet demanding power budgets. Instead, the state-of-the-art cochlear implant sound processor uses hybrid analog-digital systems in order to perform these same operations in whichever domain (analog or digital) is most efficient. By allowing some operations to take place in the analog domain (amplification, filtering, envelope detection) and some operations in the digital domain (chip programming, data transmission), this cochlear implant speech processor achieves extremely low-power operation.

In the current Analog Bionic Ear Processor, a series of bandpass filters, followed by envelop detectors and finally a logarithmic analog-to-digital converter performs the same function as the Fourier Transform, but in a much more power-efficient manner. In order to get proper frequency separation of each channel, each bandpass filter must be high-Q. In the "ideal" bandpass filter bank, each bandpass filter would have a brick-wall cutoff so that only frequency information in that particular filter's pass band would be translated into spectral energy information. However, this is not possible, and even good approximations of brick-wall filters are too power-inefficient to be useful in this application. Instead, we use high-Q second-order bandpass filters to replicate the function of the ideal brick-wall filter. In addition to being high-Q, we must also ensure that each filter has low noise and is power efficient. In its current iteration, the bandpass filters employed in the Analog Bionic Ear Processor use a two-transconductor architecture
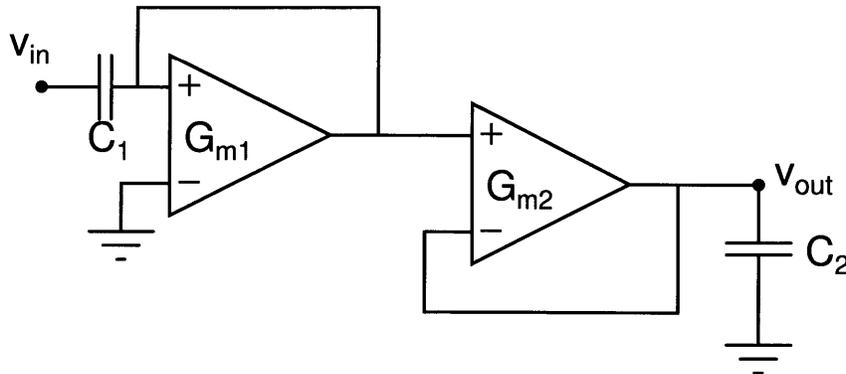
Figure 4.1 – Original Analog Bionic Ear Processor two-transconductor bandpass filter architecture

shown in figure 4.1[10].

It can be shown from the characterizing equations that this bandpass filter architecture has a noise performance characteristic that scales as Q^2. That is, if we increase the Q of the filter by a factor of two, the input-referred noise will quadruple. The power consumption of the filter, however, scales as Q^{-1}. Thus, if we evaluate the performance of these filters by looking at the product of the input-referred noise times the power consumption, we see that this figure of merit scales as Q:

$$FOM = P * \overline{v_i^2} = O(Q)$$

In this work, we implement a new design for the bandpass filter that achieves a better figure of merit. This filter was originally proposed by Arfin et. al[11]. The improvement of this filter comes by the fact that the input referred noise scales linearly with Q (better by a factor of Q), while the power consumption still scales inversely with Q. Thus, our input-referred noise times power figure of merit scales independently of Q:

$$FOM = P * \overline{v_i^2} = O(1)$$

The performance of this new filter architecture was verified through simulation, and a chip implementing a filter suitable for use in a cochlear implant speech processing unit is currently in fabrication.

---

[10]R. Sarpeshkar, C. Salthouse, J.J. Sit, M. Baker, S. Zhak, T. Lu, L. Turicchia, and S. Balster, "An Ultra-Low-Power Programmable Analog Bionic Ear Processor," IEEE Transactions on Biomedical Engineering, Vol. 52, No. 4, pp. 711-727, April 2005.

[11]ADAPTIVELY BIASED BANDPASS FILTERS: S. K. Arfin, S. Mandal and R. Sarpeshkar, "Dynamic-Range Analysis and Maximization of Micropower Gm–C Bandpass Filters by Adaptive Biasing," Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Taipei, Taiwan, May 2009, pp. 2954-2957.
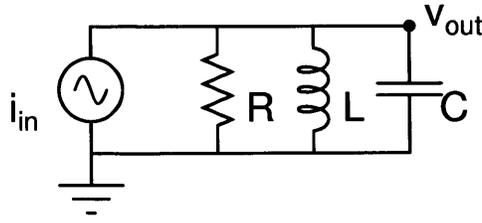
Figure 4.2 – Simplified bandpass filter architecture
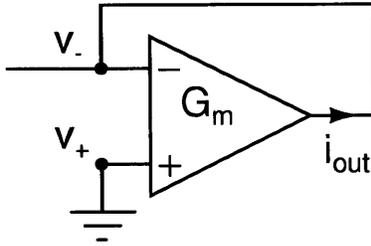
Figure 4.3 – Transconductor-capacitor equivalent of a grounded resistor

4.1 $O(1)$ bandpass filter architecture

This transconductor (G_m)-C filter topology is in its simplest form a parallel R-L-C resonant circuit (figure 4.2). We use an input current and resonate it with the parallel R, L, and C elements, taking the output voltage as the voltage across the capacitor. In order to translate this simplified passive element architecture into a G_m-C topology, we can use element substitution. This method directly maps passive elements to transconductor-capacitor building blocks, allowing simple replacement of passives to construct complex filter topologies. Since we are using a G_m-C topology, we do not need to replace the parallel capacitor. We will begin by replacing the resistor. We can do this by using a transconductor in a negative feedback loop as shown in figure 4.3.

The device input-output relationship for the transconductor is given by

$$i_{out} = G_m(V_+ - V_-)$$

With the positive terminal of the transconductor grounded, we get

$$i_{out} = -G_m * V_-$$

That is, the transconductor sinks a current from the V_- terminal to ground equal to $G_m * V_-$.

19
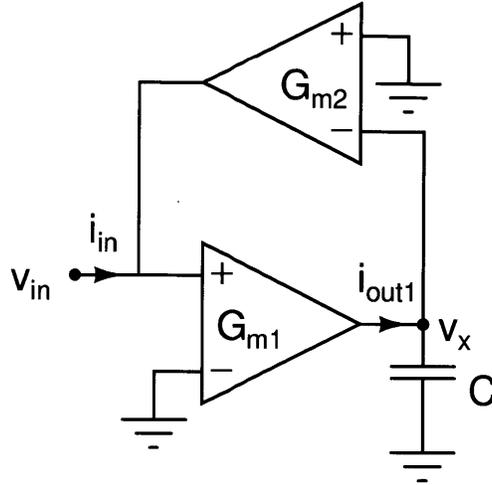
Figure 4.4 – Transconductor-capacitor equivalent of a grounded inductor

Now consider a resistor between two terminals: V_- and ground. The current through this resistor, from the V_- terminal to ground is $i = V_-/R$. Thus, our transconductor has replicated the effect of a resistor with one side tied to ground, with resistance given by:

$$R_{effective} = \frac{1}{G_m}$$

We can also recreate the effect of an inductor by using two transconductors and a capacitor, as shown in figure 4.4.

Evaluating the input-output relationship for the first transconductor, we can write that

$$i_{out1} = V_{in} * G_{m1}$$

This current is then integrated on capacitor C:

$$V_x = \frac{i_{out1}}{sC} = \frac{V_{in}G_{m1}}{sC}$$

Using the input-output relationship for the second transconductor, we find that

$$i_{in} = -(-V_x * G_{m2}) = \frac{V_{in}G_{m1}G_{m2}}{sC}$$
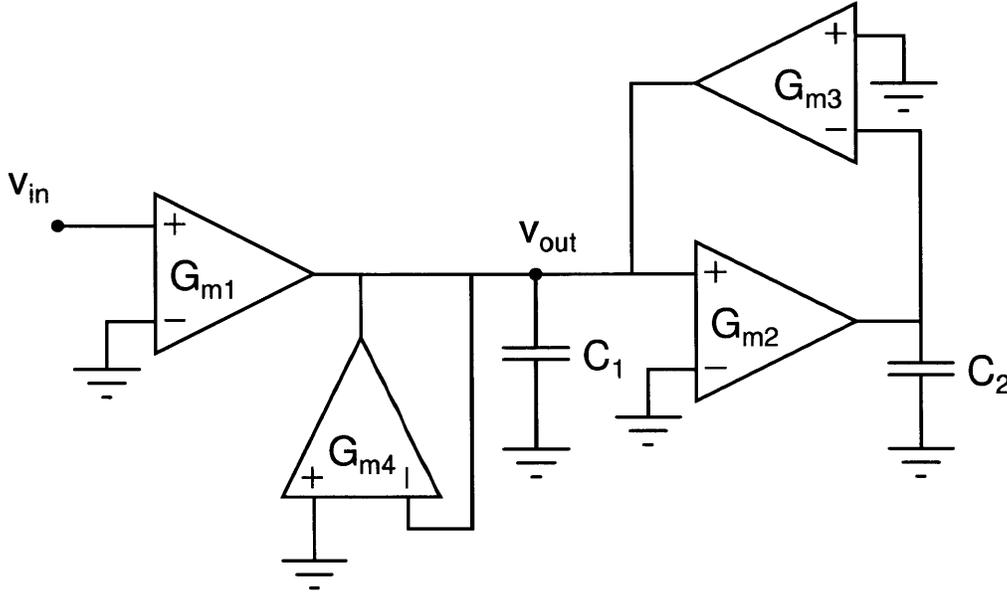
Figure 4.5 – O(1) transconductor-capacitor bandpass filter

Finally, we can write the effective impedance of this block as

$$Z_{effective} = \frac{V_{in}}{i_{in}} = \frac{sC}{G_{m1}G_{m2}}$$

Thus, this block acts as an inductor, with effective inductance

$$L_{effective} = \frac{C}{G_{m1}G_{m2}}$$

Our last building block is simply a single transconductor, used to convert a driving voltage to a driving current, necessary for our parallel RLC bandpass architecture. In total, the transconductor-capacitor bandpass filter architecture is shown in figure 4.5.

We can evaluate the ideal transfer function of this bandpass filter by first evaluating the transfer function of the parallel RLC circuit and translating each passive component into its transconductor-capacitor equivalent. For the current-driven bandpass filter of figure 4.2, the transfer function is given by

$$\frac{v_{out}}{i_{in}} = R\frac{s\frac{L}{R}}{s^2LC + s\frac{L}{R} + 1}$$

We can make the substitutions

$$R = \frac{1}{G_{m4}}$$

$$C = C_1$$

$$L = \frac{C_2}{G_{m2}G_{m3}}$$

An additionally include the fact that $\frac{i_{in}}{v_{in}} = G_{m1}$ to write the transconductor-capacitor bandpass filter

transfer function as

$$\frac{v_{out}}{v_{in}} = \frac{G_{m1}}{G_{m4}}\left(\frac{s\frac{G_{m4}C_2}{G_{m2}G_{m3}}}{s^2\frac{C_1C_2}{G_{m2}G_{m3}} + s\frac{G_{m4}C_2}{G_{m2}G_{m3}} + 1}\right)$$

Or, in canonical, second-order system notation

$$\frac{v_{out}}{v_{in}} = A\frac{\frac{\tau s}{Q}}{\tau^2 s^2 + \frac{\tau s}{Q} + 1}$$

with

$$A = \frac{G_{m1}}{G_{m2}}$$

$$\tau = \sqrt{\frac{C_1 C_2}{G_{m2}G_{m3}}}$$

$$Q = \frac{1}{G_{m4}}\sqrt{\frac{C_1 G_{m2}G_{m3}}{C_2}}$$

Thus, this transconductor-capacitor filter implements a second-order bandpass filter, with tunable gain

(A), center frequency (τ), and quality factor (Q). It is important to note that this topology allows

independent manipulation of the three filter variables. The product of G_{m2} and G_{m3} sets the filter center

frequency. Once this is set, G_{m4} sets the filter Q. Finally, G_{m1} sets the filter gain.

Two things are important about this filter topology. The first is the power and noise scaling; the input-

referred noise of this filter scales linearly with Q while the power consumption scales inversely with Q.

Thus our figure of merit is independent of Q. Compared to the bandpass filter currently in use in the

Analog Bionic Ear processor, for high-Q operation (which is precisely what we want for cochlear implant

speech processing), this filter is more power and noise efficient by a factor of Q. With quality values

typically programmed at $6 < Q < 12$, this savings can be quite large. The other important aspect of

this filter topology is that it allows independent control of filter gain. For a given cochlear implant user,

the center frequency and quality factor of each filter in the bank will be individually tuned to give that

particular user the best hearing response possible. While these may need to be adjusted as the user

ages, this adjustment is never made on-the-fly. Thus, with the center frequency and Q of each filter fixed, we have the ability to adjust the gain of each filter by modulating the transconductance G_{m1}. The frequency-selective gain control algorithm discussed later takes advantage of this control handle and uses it to improve filtering capabilities in noisy transients.

4.2 $O(1)$ bandpass filter implementation

The transconductor-capacitor bandpass filter was implemented in a $0.5\mu m$ AMI CMOS process. The capacitors were implemented on-chip, using a poly-insulator-poly structure. Each transconductor was implemented using a wide-linear-range amplifier as proposed by Sarpeshkar et. al.[12] This novel transconductor topology minimizes the transconductance and therefor maximizes the linear range of the amplifier by using the wells of two p-channel devices as the inputs, as well as gate degeneration, source degeneration transistors, and bump linearization. All of these techniques combine to give a linear range of nearly $V_L = 1.2V$ even though the transconductance amplifier operates in the subthreshold regime. The linear range of the transconductance amplifiers is important because they will be used downstream from a microphone pickup and preamp, which has an output range of approximately one volt.

The wide linear range transconductance amplifier is shown in figure 4.6. On order to tune the parameters of the bandpass filter, we must be able to change the transconductances of the amplifiers. This can be done by changing the bias current (I_B) of the amplifier. It is shown in Sarpeshkar et. al.[13], that the approximate transconductance relationship is given by

$$i_{out} = I_B tanh(\frac{v_+ - v_-}{V_L})$$

where V_L is the linear range of the amplifier. As long as we operate well within the linear range, we can approximate the transconductance by

$$G_m = \frac{I_B}{V_L}$$

Thus, for a fixed transconductance amplifier topology (fixed linear range V_L), we can vary the transconductance linearly with bias current.

---

[12]Low-power Transconductance Amplifier: R. Sarpeshkar, R.F. Lyon, and C.A. Mead, "A Low-Power Wide-Linear-Range Transconductance Amplifier," Analog Integrated Circuits and Signal Processing, Vol. 13, pp. 123-151, 1997.

[13]Ibid.
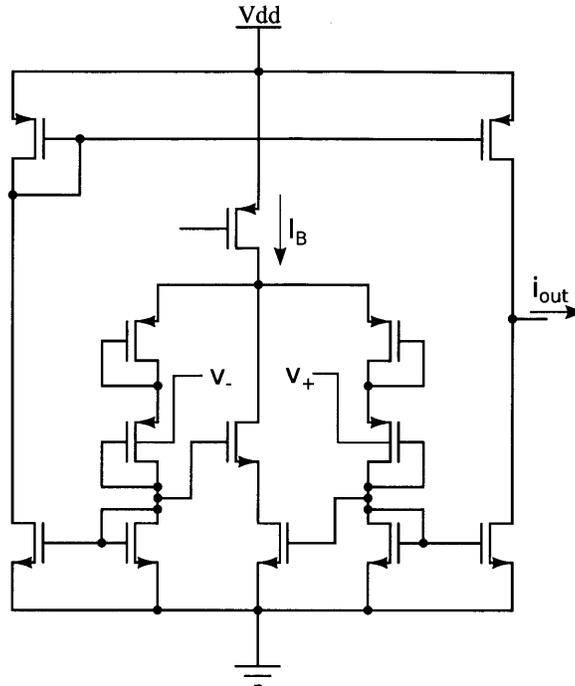
Figure 4.6 – Wide linear range transconductance amplifier

This filter topology was also simulated in Cadence to verify its operation and performance. Figure 4.7 shows a tunable center frequency range of $12Hz < f_c < 30kHz$. Figure 4.8 shows a tunable Q range of $1 < Q < 12$. These simulations show that this power- and noise-efficient bandpass filter will work in a filter bank for a cochlear implant processing unit.

This filter is currently in fabrication and is expected back in June of 2011. Figure 4.9 shows a picture of the final layout of the filter chip. Once back, it will be tested to verify performance, then will be used to implement the frequency-selective gain control algorithm discussed below.

5 Speech in Noisy Environments

One of the major existing hurdles for cochlear implant users (and more generally for hearing aid users) is the ability to correctly parse speech in noisy environments. In quiet environments, hearing aid and cochlear implant users can meet or even exceed the speech recognition accuracy of normal, hearing, people. However, in environments with significant background noise, the speech recognition accuracy of hearing aid and cochlear implant users drops dramatically. Both transient noises (such as a door slam or car engine backfire) as well as stationary noises (such as a computer fan or road noise while driving inside
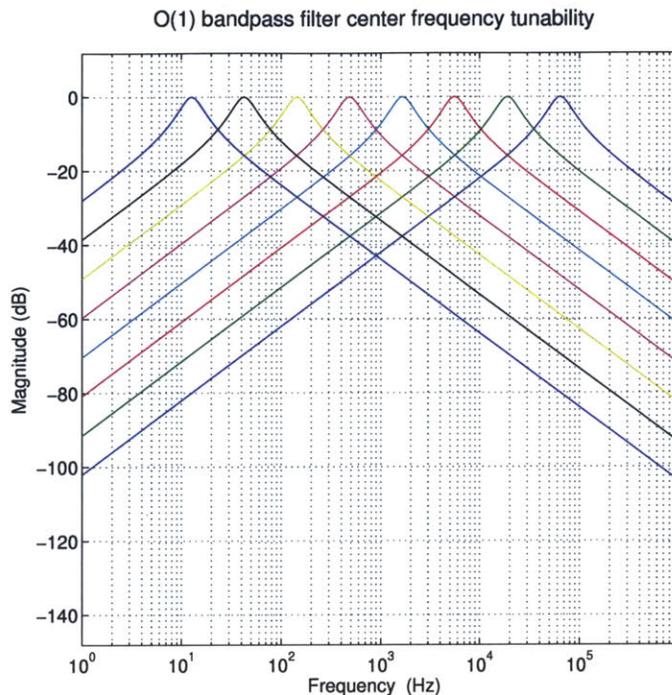
Figure 4.7 – O(1) bandpass filter simulated center frequency tuning range: 12Hz to 5.8kHz
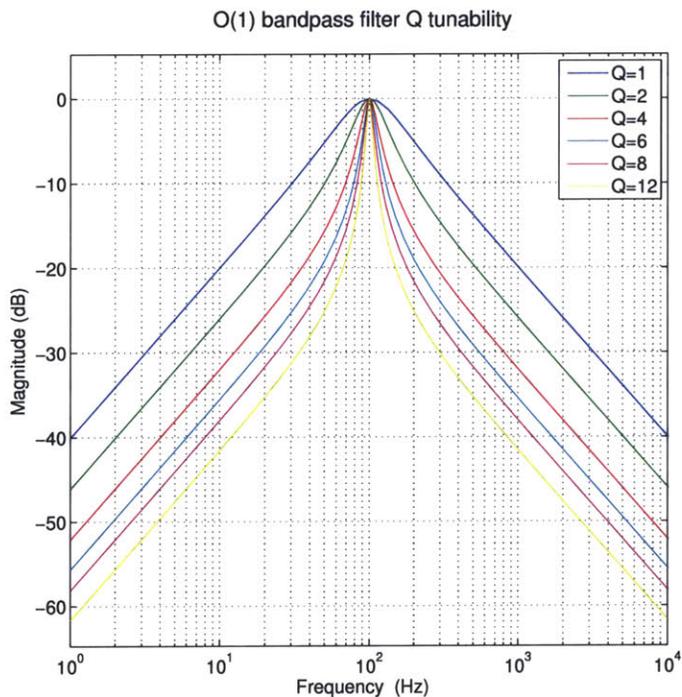
Figure 4.8 – O(1) bandpass filter simulated Q tuning range: 1 to 12

25
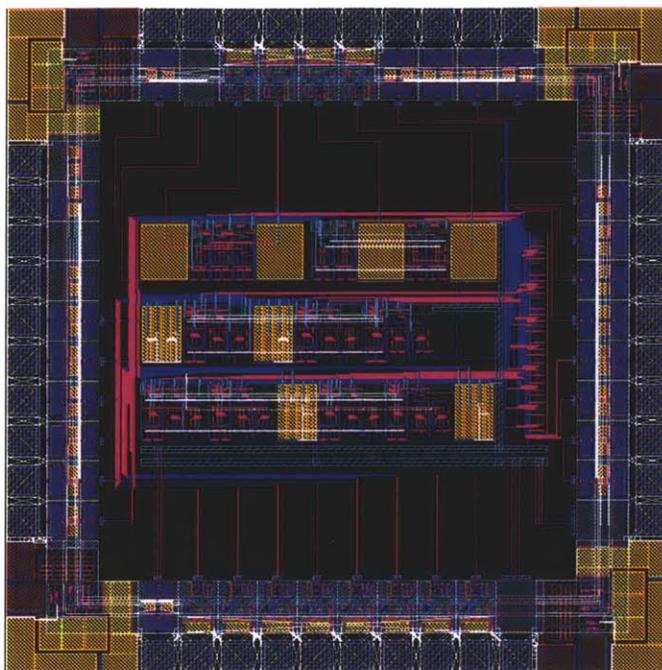
Figure 4.9 – Bandpass filter integrated chip layout (currently in fabrication)

a car) can reduce the effectiveness of cochlear implants. Many methods exist to try to improve hearing in noise, but none are able to return hearing aid and cochlear implant users back to the level of speech recognition accuracy of hearing people. We will focus here on a novel gain control algorithm for improving speech intelligibility in noisy transients.

5.1 Existing Solutions

Although the algorithm discussed below and the system on which it is based is a cochlear implant system, the same algorithm could be used to improve the performance of hearing aids. Unlike cochlear implants, hearing aids do not stimulate nerve fibers, but rather amplify sounds so that people who have partial hearing loss can hear more easily. While the method of actuation (stimulation in the case of the cochlear implant and speaker output in the case of the hearing aid) differ between the two systems, the fundamental problem of filtering speech in noisy environments is the same in both cases. There are currently hundreds of different types of hearing aids on the market. We will look at two existing solutions first, and how they attempt to solve this problem.

One current solution avoids the problem of speech filtering in noisy environments by wirelessly transmit-

ting the speech signal of interest directly to a user's hearing aid. This system, called Hearing Loop[14], employs a large RF coil (the "loop") installed in a public space. Popular installations have been in churches, public transportation buses, or anywhere where there is a large hard-of-hearing audience and difficult acoustics. This RF coil transmits audio (in our examples, perhaps speech from the pastor of the church or the bus driver) directly to a small coil embedded in the hearing aid. The hearing aid then amplifies this received signal instead of amplifying a signal picked up by its own microphone. The result is improved intelligibility. However, these systems are not in widespread use. While the Hearing Loop technology could be ported to a cochlear implant, the transmission loops themselves must be installed in the locations of interest. Furthermore, wireless transmission is susceptible to interference from an increasingly vast array of electronic devices.

Another unique solution comes from the realm of hearing protection. This device, from CENS Digital[15], is an electronic noise suppressor designed for hunters and sport shooters. The aim of this device is to protect the user's hearing by muffling the sound of a gunshot. While this is not a cochlear implant or hearing aid, and does not even purport to improve speech intelligibility, it is an example of how automatic gain control algorithms can be used to suppress noise. The CENS Digital suppressor is a custom-molded in-ear device that has a large amount of passive noise attenuation. With a good fit, custom-molded in-ear plugs can provide up to $50 dB$ of broadband attenuation. Much like a hearing aid, the device also has an outward-facing microphone and a small inward-pointing speaker. When no gunshots are detected, the speaker receives input from the microphone and allows the user to hear as normal. However, when a gunshot is detected, the speaker stops as quickly as possible, and the passive attenuation provided by the in-ear plug is used to muffle the gunshot. The loud transient suppressive behavior of this product is similar to the aims of the noise-suppressive frequency-specific automatic gain control algorithm presented below, but it does not act in a frequency-dependent way. We will see that having independent gain control for different frequencies can increase the noise rejection capabilities while not eliminating speech that may be occurring at non-noisy frequencies.

5.2   Automatic Gain Control

Automatic gain control systems as used in cochlear implants and hearing aids attempt to modulate the gain of the sound-to-stimulation path so as to allow users to both hear soft sounds (by increasing the

---

[14]"Hearing Loop". n.d. Web. 3 April 2011.<http://www.hearingloop.org/>

[15]"CENS Digital". n.d. Web 13 Jan. 2011 <http://www.censdigital.com/home.html>.
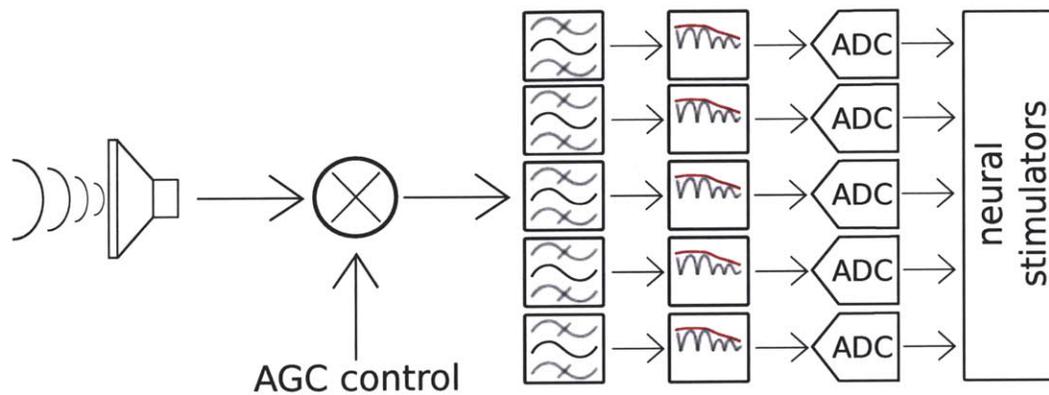
Figure 5.1 – Standard AGC control architecture

gain) as well as not be overwhelmed by loud sounds (by decreasing the gain). These systems work by monitoring the magnitude of the microphone input and adjusting the gain up if the input is low and adjusting the gain down if the input is high. In most existing AGC systems, the gain control modifies the front-end gain of the entire system, as can be seen in figure 5.1. This gain control mechanism provides a coarse control over the total sound level. Much work has already been done to try to improve the performance of AGC systems in transients. Work in both digital and analog systems has further built on this conceptually simple gain control architecture[16] [17] [18]. A more complicated technique requiring multiple feedback loops has also been demonstrated in an ultra-low-power analog design[19]. All these techniques share a fundamental design architecture: the gain control is wideband and modulates total system signal level.

This work proposes a new automatic gain control system to improve speech intelligibility in noisy transients by employing a frequency-selective gain control mechanism.

5.3 Frequency-selective gain control architecture

The standard signal processing chain in an analog cochlear implant is as follows. For each channel (frequency band), the wideband input signal is first bandpass filtered to remove all content in neighboring

---

[16]W. A. Serdijn, A. C. van der Woerd, J. Davidse, and H. M. van Roermund, "A low-voltage low-power fully-integratable front-end for hearing instruments," IEEE Trans. Circuits Syst. I, Fundam. Theory Applicat., vol. 42, no. 11, pp. 920–932, Nov. 1995.

[17]D. G. Gata et al., "A 1.1-V 270-μA mixed-signal hearing aid chip," IEEE J. Solid-State Circuits, vol. 37, no. 12, pp. 1670–1678, Dec. 2002.

[18]S. Kim, J.-Y. Lee, S.-J. Song, N. Cho, and H.-J. Yoo, "An energy- efficient analog front-end circuit for a sub-1V digital hearing aid chip," in Symp. VLSI Circuits Dig. Tech. Papers, 2005, pp. 176–179.

[19]Low-power Automatic Gain Control (AGC) Circuits: M. Baker and R. Sarpeshkar, "Low-Power Single Loop and Dual-Loop AGCs for Bionic Ears," IEEE Journal of Solid-State Circuits, Vol. 41, No. 9, pp 1983—1996, September 2006.
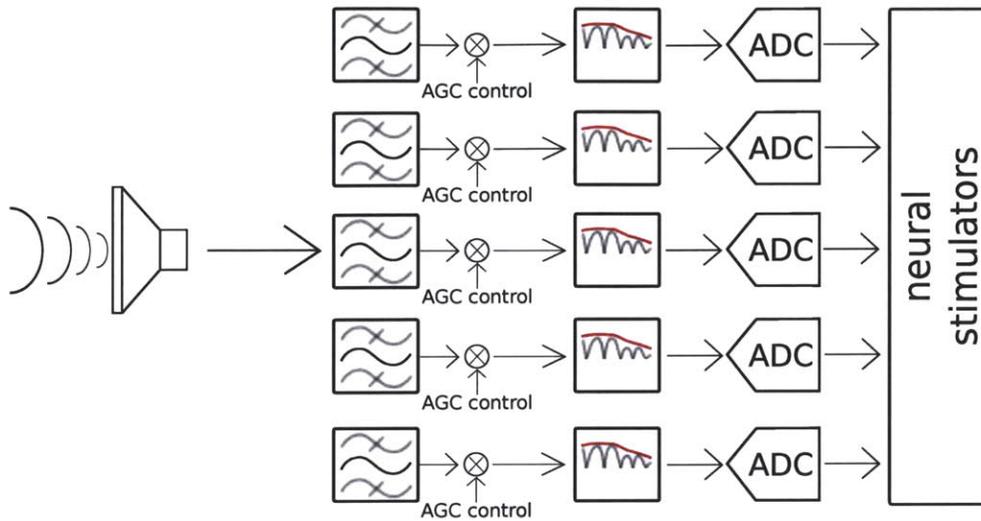
Figure 5.2 – Proposed AGC control architecture

frequency bands other than the one of interest. Next, the filtered signal is rectified and the envelope is detected. The slower-moving envelope signal is then digitized and used to actuate downstream neural stimulators. The gain control architecture proposed here takes advantage of the bank of bandpass filters already present in a low-power analog cochlear implant system to allow frequency-selective gain control. By modifying the gains of the signal post-filtering, we can achieve noise-suppressive behavior from our AGC system. This architecture is shown in figure 5.2 . It should be made clear that from a system behavior standpoint, we have not diminished the realm of potential behaviors of the system. If we were to command the frequency selective gain controls to always have the same value as each other, we would replicate the behavior of the previous architecture. By shifting the gain control post-filter, we have only expanded the gain control and noise reduction possibilities. The improved bandpass filter architecture discussed above allows us an easy mechanism by which we can modify the gain of each filter without changing its center frequency or Q.

5.4 The Acoustics of Transients

Before proceeding further, it is important to understand the acoustical phenomena at work when a transient sound propagates through space. Transient sounds are defined here to be sounds that are very short in time. More formally, if we consider the environment in figure 5.3 , we require that the duration over which the source is creating sound pressure waves be much less than the propagation time from the
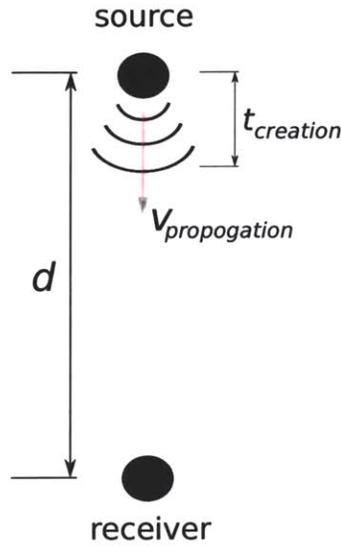
Figure 5.3 – Transient sound environment

source to the receiver:

$$t_{creation} \ll \frac{d}{v_{propogation}}$$

We will assume that as long as this condition is met, our noise source can be considered a noise event generator. That is, our noise source generates single pulses of sounds, each one considered a unique event. We do not restrict the spectral signature of these impulses. In a strict sense, a true impulse must have a uniform spectral distribution. However, we allow for simply short pulses of noise, thus the spectral signature may (and often does) vary from one noise source to another. We will concentrate our discussion here on gunshots as our noise source. Gunshots have a very uniform power spectral density; that is, they create an impulse of sound with nearly equal power at all frequencies. This makes them a good noise source for testing our noise reduction algorithm, ensuring that this gain control architecture can reduce the presence of noise at all frequencies.

In order to have a good understanding of the spectral signatures likely to be encountered at the receiver from loud noise transients, it is important to understand the propagation of noise through an environment. When a sound is created from a source with no directivity, sound pressure waves radiate outwards from the source in a spherical shape. Thus, the energy of the sound falls off as the inverse of the distance squared. This assumes a point source in free space, with no environment interactions. In reality, the transient noises picked up at the receiver will have been affected by a complex environment. The acoustics associated with sound propagation in complex environments is an entire field of science unto itself. Thus, for our analysis

here, we will defer to the ISO 9613 standard for the attenuation of sound during propagation outdoors[20]. This standard provides a detailed method for calculated sound attenuation as it travels through outdoor environments, perfect for our analysis of gunshots.

The ISO 9613 standard gives the following equation to determine the sound pressure level at a receiver:

$$L_{fT}(DW) = L_W + D_C - A$$

where L_W is the sound power level (dB) output of the noise source, D_C is the directivity correction (dB) of the source (0dB for an omnidirectional source), and A is the attenuation (dB) due to propagation from source to receiver. The important term in this equation is the attenuation. This term tells us how pure tones at different frequencies are attenuated as a function of the acoustic environment. The ISO standard outlines several important sources of attenuation: geometrical divergence, atmospheric absorption, environmental absorption, ground effect, barrier attenuation, and miscellaneous attenuation. We will only consider the first three, as they are the most important factor in understanding the spectral signature of gunshots in outdoor environments.

Sound produced from a omnidirectional point source radiates outward spherically, giving rise to an inverse square law sound pressure level equation. This inverse square law dictates that every doubling in distance results in a 6dB attenuation. This is a frequency-invariant attenuation that simply arises out of the geometry of sound propagation.

Atmospheric attenuation refers to the energy that is lost due to the interaction between the sound waves and the propagation medium–the atmosphere. This attenuation is due to two major factors: friction between neighboring air molecules, and the absorption/re-radiation process. The friction loss is simply due to the interaction between the molecules of the atmosphere transmission medium when disturbed by a sound pressure wave. The absorption/re-radiation loss is due to the fact that air particles in the presence of a sound pressure wave will very briefly absorb the energy, translating it into increased movement of the particle, before re-radiating that energy and returning to its original energy state. This process is not perfectly efficient, so some energy is lost, thereby resulting in an attenuation of the sound pressure wave. The extent to which the atmosphere absorbs a particular sound is dependent on the frequency of that sound. While the ISO 9613 standard provides numerical models for estimating the level of absorption, for

---

[20]ISO IS 9613-2:1996: Acoustics – Attenuation of sound during propagation outdoors – Part 2: General method of calculation.
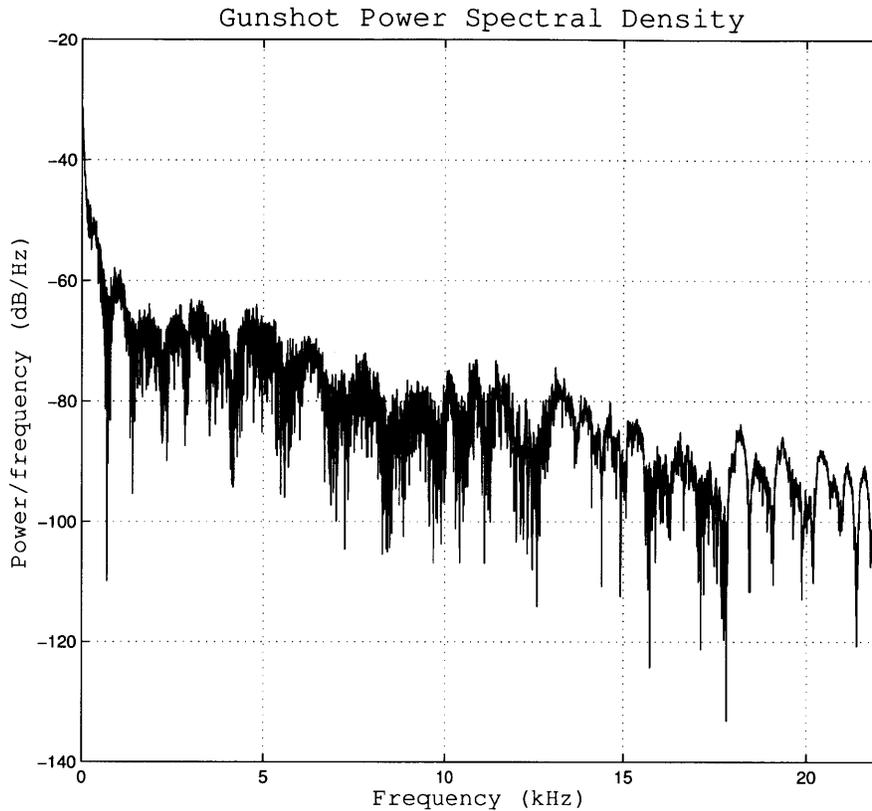
Figure 5.4 – Power spectral density of M4 rifle recorded at 75 ft

our purposes it is sufficient to understand that the relative absorption increases with increasing frequency. Thus, ignoring all other sources of attenuation, a low-frequency tone will travel farther through the atmosphere than a high-frequency tone.

These first two factors contribute to an important conclusion: a white noise transient at a source will produce at a receiver, a noise transient whose wave front has spectral energy that falls off as frequency increases. Since the wave front (the very first moment of sound) will contain the sound that traveled directly from the source to the receiver, it will have been attenuated only by the geometrical divergence (frequency-independent) and atmospheric absorption (frequency-dependent) processes. This can be seen in figure 5.4 , which shows the power spectral density of the wave front of a gunshot (typically considered a good approximation of a white noise source in the audio band) from a recording of an M4 rifle at 75 ft. The spectral energy shows that the high frequencies have been attenuated by nearly $30dB$ more than the low frequencies.

In addition to the wave front, the sound at a receiver due to a transient at a source contains a decay due to reverberations and reflections of the sound in the given environment. In order to understand the entire noise transient likely to be encountered at a receiver, we must also understand the acoustics of the reverberations and reflections.

Environmental absorption refers to the attenuation of sound as it reflects off of elements in the environment. In outdoor environments, this may include buildings, trees, grass, or any other surface. Every surface can be characterized by the extent to which it absorbs sound at a particular frequency. The noise reduction coefficient (NRC) measures this attenuation on a zero to one scale. An NRC of zero at a particular frequency indicates perfect reflection whereas an NRC of one indicates perfect absorption. In general, most materials have NRC values that increase with increasing frequency. That is, they are better at absorbing sounds at high frequency than they are at absorbing sounds at low frequency. Just as was the case with atmospheric absorption, environmental absorption tends to reduce the intensity of high frequencies in relation to low frequencies. Thus, for a white noise transient, the receiver will detect a transient whose wave-front contains more low frequency energy than high frequency energy, and whose decay will take longer for low frequency energy than for high frequency energy.

This theoretical absorption model was confirmed by analyzing recordings of gunshots at distances from 75 feet to 150 feet. In all cases, we see more wave-front energy the lower the frequency band, and longer decay times the lower the frequency band. Plots of both spectrograms as well as frequency-specific decay characteristics for different gun barrel lengths, ammo type, and distances can be seen in appendices A and B. The information gleaned by analyzing acoustic propagation of gunshots is important because it provides a sense of the room for improvement with frequency-specific gain control. The difference in wave-front spectral energy between the high frequencies ($10kHz$ and up) and low frequencies ($500Hz$ and below) can be dramatic. In the power spectral density plot of figure5.4, this difference is nearly $30dB$. In a wideband gain control system, no matter what our policy is for modulating gain, we perform sub-optimally. In such a system, we can only detect total broadband energy, thus we change the gain in response to roughly the average of the power spectral density. We attenuate high frequencies more than we should, so any speech that is also taking place in this frequency range will be needlessly attenuated. We also attenuate low frequencies less than we should, so the long decay time of the low frequency noise needlessly contaminates our speech signal. If instead, we respond on a per-frequency-band basis, we can do a much better job of attenuating noise in the bands in which it occurs, while leaving speech in noise-free bands un-attenuated. This observation is the basis on which the frequency-selective gain control
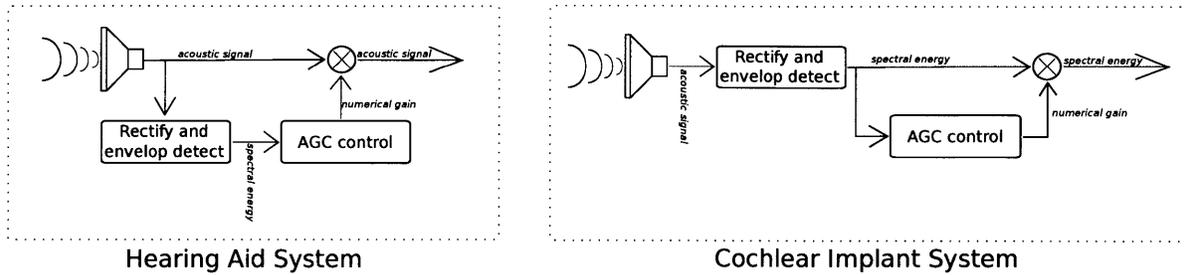
Figure 5.5 – Comparison of hearing aid and cochlear implant gain control architectures

algorithm is based.

5.5 Frequency-selective Gain Control Algorithm

We now describe the frequency-selective gain control algorithm. Although the algorithm applies to applications of both cochlear implants as well as hearing aids, there is a slight architectural difference between these two systems. Cochlear implant systems use spectral energy information to stimulate auditory nerve fiber endings in order to recreate the sensation of hearing. Hearing aid systems use acoustic signals to power a small speaker which amplifies the sound and delivers it to the inner ear. Thus, the actuation step of the two systems rely on different data; the cochlear implant uses spectral energy information while the hearing aid uses acoustic signals. Thus, the architecture of these two systems is different, as is shown in figure 5.5. It is important to realize that although the architecture is different, the gain control block is identical in the two systems. In both cases, the gain control block receives spectral energy information in each frequency channel and uses this information to modulate the gain of each channel independently. In the cochlear implant system, the gain modulation acts on a spectral energy signal, whereas in the hearing aid system, the gain modulation acts on an acoustic signal.

The gain control algorithm itself is simple. We use a transient threshold detector followed by an expansive non-linearity in order to translate the spectral energy information into a gain signal. Under normal operation (with no loud transients), the threshold detector inactivates the non-linear gain control, and the gain is fixed at unity. However, when the threshold detector trips, the non-linear gain is applied to modulate the output signal. This gain control scheme is shown in figure 5.6. When activated, by employing an expansive non-linearity, we create a total input-to-output relationship which is compressive. The expansive non-linearity is given by
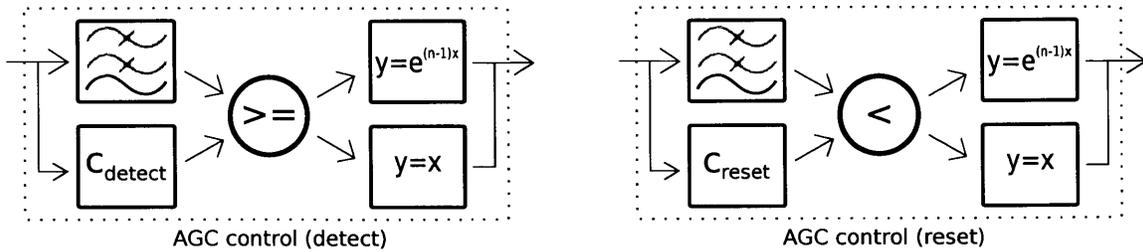
$$y = e^{(n-1)x}$$

Figure 5.6 – Automatic gain control algorithm

where n is the compression coefficient, between 0 and 1. When this gain control term is multiplied by the original signal (spectral energy for cochlear implants or acoustic for hearing aids), the overall input-output relationship is approximated by

$$y = e^{nx}$$

Thus, we have achieved a compressive behavior. This compressive function is what performs the noise reduction. Signals which are relatively loud are attenuated more than signals which are relatively soft. We can control the extent of this compression by varying n, the compression ratio. A low n will give us aggressive compression, whereas a high n will give us moderate compression. In our transient suppression scheme, we only apply this non-linear gain when a transient has been detected; once it has been detected, we apply aggressive compression to attenuate the noise.

The threshold detector senses the presence or absence of a loud transient, and switches the gain control between unity (when no transients are detected), and the non-linear compression scheme (when a transient is detected). The threshold detector simply compares the current spectral energy information to a low-pass filtered version. When the current spectral energy level times some constant (c_{detect}) is greater than the filtered version, the threshold trips. In essence, we are comparing the scaled instantaneous value to the averaged value. When the instantaneous value is enough more than the average value, we declare that a loud transient is occurring. Once the instantaneous times some constant (c_{reset}) drops below the averaged value, we declare that the transient is over and we resume normal unity gain operation.

The low-pass filter allows us to set the effective attack and release times for the transient condition detection. Lowering the cutoff frequency of the low-pass filter increases the attack and release time constants. More formally, if the cutoff frequency of the bandpass filter is f_c, then the attack and release time constants are given by

$$t_{attack/release} = \frac{1}{2\pi f_c}$$

The attack and release time constants control how quickly a change in the total power in a given channel is reflected by the "averaged" value we are using to set the transient detection threshold. The attack corresponds to increases in total power and the release corresponds to decreases in total power. The inclusion of the low-pass filter in this gain control architecture is important. One could argue that the same function is performed by the envelop detection step performed when translating the acoustic signal to a power spectral signal. However, since a full cochlear implant system contains an envelope detector per channel, we must set the attack and release time constants of the envelope detector to appropriate values for each frequency band. By including a low-pass filter we have decoupled the attack and release time constants associated with the envelope detector (which in large part are dictated by the frequencies at which the envelope detector must work), and the attack and release time constants associated with the gain control algorithm (which are dictated by how quickly we would like the gain to adapt to changing noise levels). While this decoupling is still possible with a standard broadband AGC system, it cannot be done at the per-channel level.

5.6   Results

The frequency-selective gain control algorithm presented above was simulated using MATLAB and Simulink. A preexisting Simulink model of the Analog Bionic Ear Processor, with which this gain control architecture would work, was used to model the rectification and envelope detection already designed and implemented in the ABEP. This preexisting model was combined with a model of the improved band-pass filter (discussed above) derived from Cadence simulations. These models were used to simulate the behavior of the gain control system under a variety of inputs.

Single gunshot recordings were used to test the ability of the gain control algorithm to suppress noisy transients. Figure 5.7 shows spectrograms comparing an unfiltered gunshot recording to one filtered using a MATLAB and Simulink simulation of the frequency-selective gain control algorithm. More simulations with speech and gunshot recordings with different barrel lengths, ammo type, and distance can be seen in appendix C. As can be seen in all these plots, the gain control algorithm quickly reduces the gain of each channel in response to the loud noise transient.

An important point must be made here. It can be seen from the spectrogram of the processed gunshot that the algorithm cannot immediately reduce the gain of each frequency band, thus we cannot completely eliminate a gunshot. This has to do with the inherent delay associated with the signal processing of each
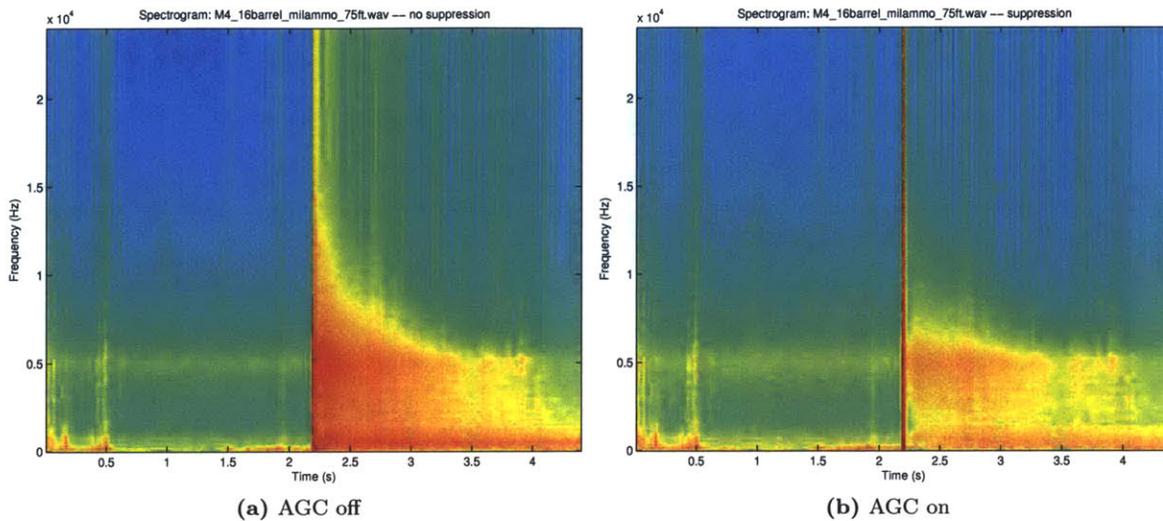
Figure 5.7 – Comparison of gunshot (M4, 16 inch barrel, military ammo, at 75 ft.) with AGC system on and off

frequency band channel. It takes a certain amount of time to perform band-pass filtering, rectification, envelop detection, low-pass filtering, and apply the expansive non-linearity. By the time we can detect that the noise power in a given channel has changed, the original sound associated with that change has passed. In essence, we are always acting on old data. Our response time to a gunshot is limited by the delay of the processing. In certain circumstances we may be able to remove the effect of this delay. This will be discussed below in section 5.7.1.

Next, this gain control system was tested with single transients in the presence of speech. A comparison of an unfiltered gunshot with speech to a filtered gunshot with speech is shown in figure 5.8 . Further simulations with speech and gunshot recordings with different barrel lengths, ammo type, and distance can be seen in appendix D. It can be seen that with the frequency-selective automatic gain control algorithm in place, the speech becomes much more pronounced in comparison to the gunshot. Since there are many frequency bands for which the noise transient from the gunshot does not overlap with the speech, our frequency-selective algorithm performs well by attenuating the gunshot noise more in bands with no speech. In these examples, during the gunshot transients, the speech-to-noise level was improved by $10dB$ to $16dB$.
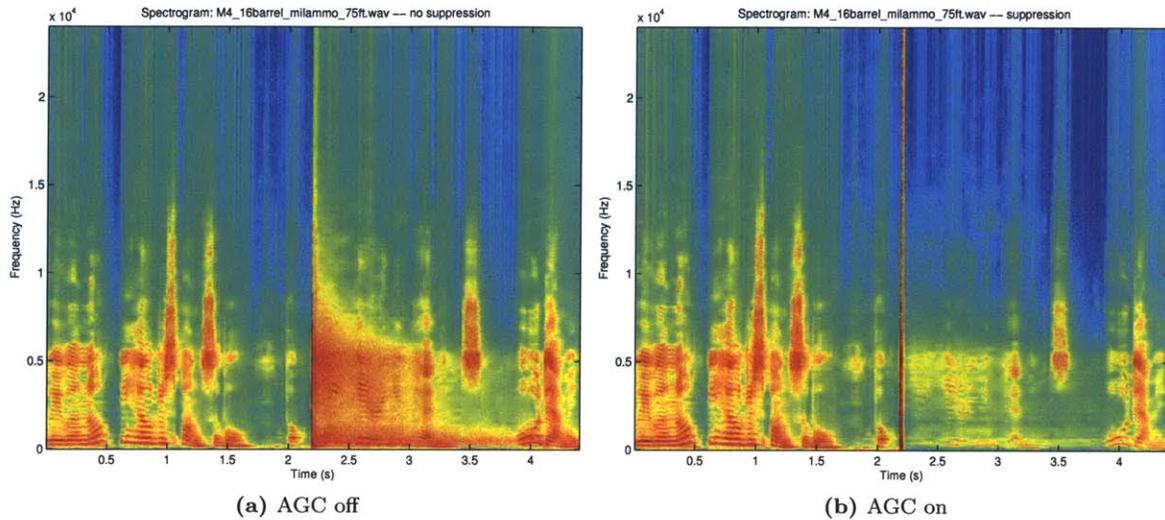
(a) AGC off   **(b)** AGC on

Figure 5.8 – Comparison of gunshot (M4, 16 inch barrel, military ammo, at 75 ft.) plus speech with AGC system on and off

5.7   Future Work

In this section, we discuss further improvements that could be made on the frequency-selective automatic gain control algorithm discussed previously.

5.7.1   Delay Compensation

While we cannot reduce the delay of the processing, we can introduce a delay to the un-processed sound such that we re-align the processed data on which we are basing our gain modulation and the original data on which we are performing that modulation. As long as the introduced delay is not significant (less than several tens of milliseconds), this would be acceptable to do in cochlear implants, where the user cannot hear the original sound, only the perceived sound based on stimulations controlled by the implant. If this delay is too great, however, there would be a perceptible lag between the visual cues of speech and the resulting cochlear stimulation. If we simply delay the triggering of stimulation by the length of time it takes to process the sound, we can completely "remove" gunshots; that is, we can completely prevent stimulation in the presence of gunshots. In order for this same technique to work in devices used by hearing-capable people, the device would need to first reduce or eliminate the ability of the user to hear sounds not passed deliberately through the device. Then, any noise deemed acceptable could be reproduced by the device. Without the noise-insulating property of the device, users would be able to

hear gunshots even though the device itself is not reproducing them.

5.7.2   Independent Attack and Release Times

The transient noise detection employs a simple low-pass filter as a means to find the local "average" spectral power. This average value is then compared to the instantaneous value in order to determine if a loud transient has occurred. Instead of a low-pass filter, a true averaging block could be used instead. This would allow independent effective attack and release time constants. If the average is increasing, we would use a local average window that corresponds to the desired attack time constant. If the average is decreasing, we would use a local average window that corresponds to the desired release time constant. It may be the case that independent control of attack and release time constants would allow a more robust detection of noise transients.

5.7.3   Applications to Stationary Noise Reduction

Although this algorithm was developed with noisy transients in mind, it can also be applied to stationary noise as well. Stationary noise is noise whose spectral signature does not change over time. This could be noise such as road noise while traveling in a car, or noise from a ceiling fan in a room. Since our gain control algorithm has independent control over the gains of each frequency channel, it is well-suited to stationary noise reduction. Instead of detecting loud transients and using this to trigger the use of the compressive input-to-output gain control, we could instead simply leave the compressive behavior in place all the time (still with a low-pass filter in place), but with a much less aggressive compression ratio. Thus, channels that tend to have continuous high spectral energy will be attenuated more than channels with low spectral energy. The result of this is that strong noise tones at a particular frequency will be greatly attenuated, while neighboring channels will be unaffected.

This noise reduction scheme was tested using the same MATLAB and Simulink framework developed for the tests of the transient reduction algorithm. Appendix E shows comparisons of unfiltered and filtered stationary noises. Appendix F shows comparisons of unfiltered and filtered stationary noises with speech. These noise recordings (all military in nature) come from the Signal Processing Information Base, sponsored by the Signal Processing Society and the National Science Foundation[21]. All of these

---

[21]"Signal Processing Information Base." Web. n.d. 10 Nov. 2011. <http://spib.rice.edu/spib/data/signals/noise/>.

figures show a reduction in noise when the frequency-specific automatic gain control algorithm is used. The spectrogram plots with speech show that the speech becomes more prominent as the background stationary noise is attenuated. Further work might be done to develop a joint algorithm which implements both the aggressive compression on detection of a loud transient as well as weak compression to suppress stationary noise.

6 Cochlear Implant with Stimulation

We will now shift to the other aspect of this thesis work: a proof-of-concept cochlear implant system with stimulation, leveraging existing speech processing and stimulation ICs. These two chips were combined on a single printed circuit board (PCB), along with an FPGA and other discrete parts to create a working implant system. This implant system represents an end-user product; once miniaturized, this system would be a functioning cochlear implant for use in patients with profound hearing loss.

6.1 Implant System Overview

The cochlear implant system is composed of three main parts. The first is the Analog Bionic Ear Processor, which is a state-of-the-art cochlear implant speech processor. This IC receives sound through a microphone, processes it, and outputs digital spectral energy information. This spectral energy information represents how much energy there is in any one of the processor's 16 frequency bands. This information is read by an FPGA and packetized so that it can be sent to a neural stimulator IC[22]. The stimulator then sends stimulation pulses to one of four electrodes in response to the data it receives. At system start-up, the FPGA also sends programming commands to the ABEP to set the center frequencies and Qs of the bandpass filters in its filter bank, as well as several other programmable options. This general architecture is shown in figure 6.1.

Neural stimulation is an essential last step in a cochlear implant system. The neural stimulation chip that has already been designed and tested is a wireless stimulator receiver transmitter pair that can be operated at up to a meter apart. For this project, only the receiver and electrode drivers were be used. The neural stimulator has four controllable stimulation channels, 32 selectable current levels from

---

[22]S. K. Arfin, M. A. Long, M. S. Fee, and R. Sarpeshkar, "Wireless Neural Stimulation in Freely Behaving Small Animals," Journal of Neurophysiology, Vol. 102, No. 1, pp. 598-605, July 2009.
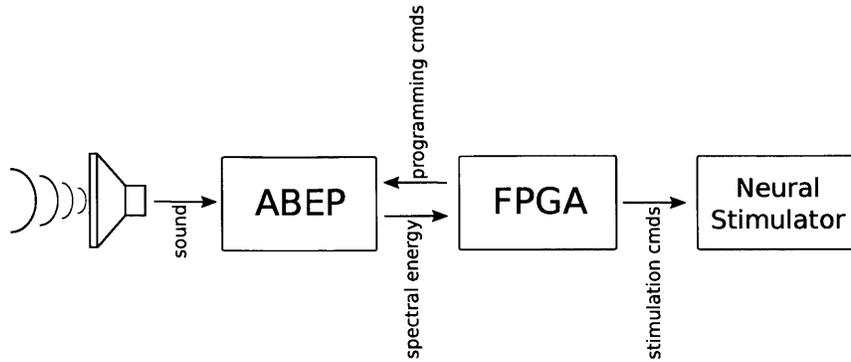
Figure 6.1 – Cochlear implant system architecture

$10\mu A$ to $1mA$, and a positive or negative stimulation phase selection. The neural stimulator chip delivers stimulation through a novel blocking-capacitor-free charge balanced system[23].

The neural stimulator has been tested in freely behaving birds. In zebra finches, when random neural stimulation was delivered during singing, premature termination of bird songs was observed[24].

6.2 ABEP and FPGA interaction

The FGPA and ABEP interact in two distinct ways. First, on power-up, the FPGA is used to send programming commands to the ABEP to set all of its user-programmable settings. These settings include the center frequencies and Qs of each filter and the bandpass filter bank, current reference levels, sampling frequency, and many others. This start-up routine is programmed in the ROM that the FPGA reads from and is editable by uploading a modified version of the software code running on the FPGA. The ABEP chip uses a serial interface for many of its programming options, so the frequency and ordering of the bits is important during the programming phase.

The second interaction happens during normal operation. Once the FPGA has programmed the ABEP, the ABEP begins outputting digital data to the FPGA. The ABEP digital outputs are a 7-bit bus that is shared among all 16 frequency channels. Thus, the ABEP is only outputting a single channel's data at any given time. The FPGA commands the ABEP to continue to output the next channel's data

---

[23] J. Sit and R. Sarpeshkar, "A low-power, blocking-capacitor-free, charge-balanced electrode-stimulator chip with less than 6nA DC error for 1mA full-scale stimulation," IEEE Transactions on Biomedical Circuits and Systems, Vol. 1, No. 3, pp. 172-183, September 2007.

[24] S. K. Arfin, M. A. Long, M. S. Fee, and R. Sarpeshkar, "Wireless Neural Stimulation in Freely Behaving Small Animals," Journal of Neurophysiology, Vol. 102, No. 1, pp. 598-605, July 2009.

by providing a scan clock to the ABEP. The FPGA uses a crystal as its clock source, and supplies a divided down version of this clock as the scan clock to the ABEP. During normal operation, the FPGA is constantly commanding the ABEP to scan through all 16 channels, and is updating the spectral energy data it is storing for any given channel each time it cycles back to that channel. Thus, the FPGA always has an up-to-date store of all 16 channel's worth of spectral energy data from the ABEP.

6.3 FPGA data translation

The FPGA must translate the spectral energy data output from the ABEP into a format usable by the stimulator chip. There are three important aspects to this translation. First, the ABEP digital outputs are seven bits wide, but the stimulator chip only has five bits of stimulation intensity precision. Naturally, we would like to command the stimulator to output a stimulation pulse whose intensity is proportional to the spectral energy information from the ABEP. Thus, the FPGA truncates the ABEP seven-bit precise data down to five bits. Second, the stimulator chip has only four output stimulation channels. While a real, implanted system would need to use four stimulator chips in order to match the sixteen channels in the ABEP, this prototype system uses only a single stimulator chip for simplicity. The FPGA must select four of the sixteen available channels from the ABEP and map them to the four channels in the stimulator. For this demo system, the choice is somewhat arbitrary, but the selected ABEP channels should have a good spread across the sixteen available channels. Lastly, the FPGA must packetize the data appropriately in order to be recognized by the stimulator. The packetization requires that a special recognition sequence (four ones) be prepended to the data, and that the bits representing the stimulation channel, stimulation intensity, and stimulation phase be organized in a particular way.

6.4 FPGA and neural stimulator interaction

After the FPGA as mapped the ABEP data into a format appropriate for the neural stimulator, the FPGA must encode the digital bits as a pulse-width-modulated (PWM) signal, itself modulated up to $13MHz$. The initial PWM translation is performed by the FPGA utilizing the clock provided by the crystal. This PWM data is then used to modulate the base of a BJT in a standard Colpitts oscillator. The output from the oscillator is the desire PWM signal modulated to $13MHz$ and is sent to the neural stimulator. The Colpitts oscillator is implemented on the custom PCB using discrete parts.
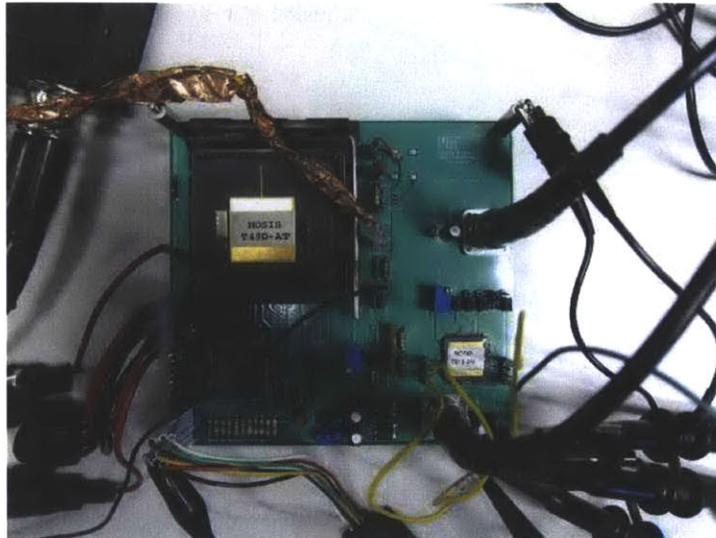
Figure 6.2 – Experimental setup of cochlear implant system with stimulator

6.5   Prototype system

This cochlear implant system was fully implemented and tested using a custom PCB. A Xilinx Spartan 3E and outboard ROM were chosen as the FPGA. The experimental setup is shown in figure 6.2. This prototype system has the ability to display the spectral energy information output by the ABEP to a computer running MATLAB. This data is used to shade a real-time spectrogram display. The stimulation pulses can also be viewed on an oscilloscope. Four stimulation channels with stimulation pulses are shown in figure 6.3.

7   Conclusions

This thesis writeup demonstrated an improvement on an existing ultra-low-power cochlear implant system by utilizing an improved noise and power efficient bandpass filter bank to implement a novel frequency-selective gain control algorithm capable of reducing, and in some cases removing, loud transient noises, thereby improving speech intelligibility. This gain control algorithm was able to improve the speech to noise ratio in the presence of both loud noise transients as well as stationary noise. We also showed a prototype full cochlear implant system, from speech processor all the way to neural stimulation.

Figure 6.3 – Stimulation pulses from prototype cochlear implant system

A Gunshot Spectrogram Plots

These are spectrogram plots (figures A.1, A.2, A.3, A.4, A.5, A.6, A.7, and A.8) generated by analyzing real gunshots recorded and provided by the United States military. Each plots shows a spectrogram of a gunshot. We can see two phenomena in these plots: 1)in general, the wave-fronts (at the impulse in the spectrogram) contain more energy as the frequency gets lower, and 2)the decay times become longer as the frequency gets lower.
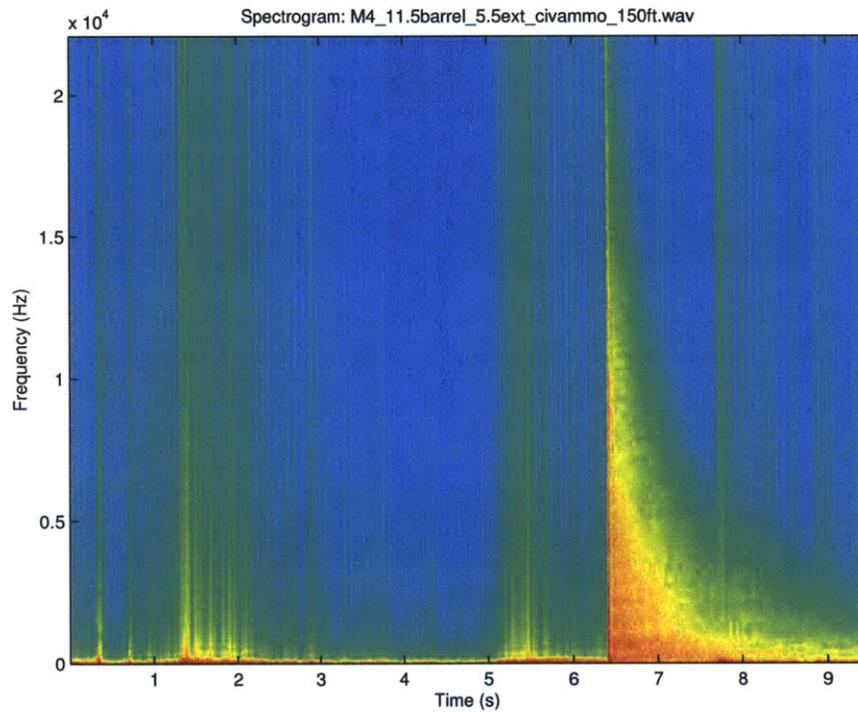
Figure A.1 – Spectrogram plot for M4 rifle with 11.5 inch barrel and 5.5 inch extension, civilian ammo, at 150 feet
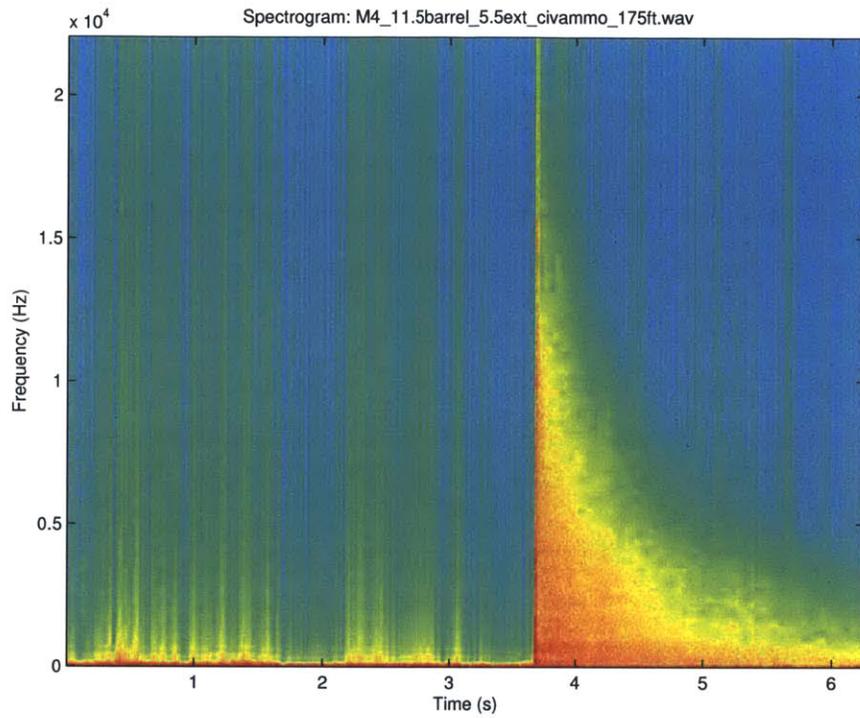
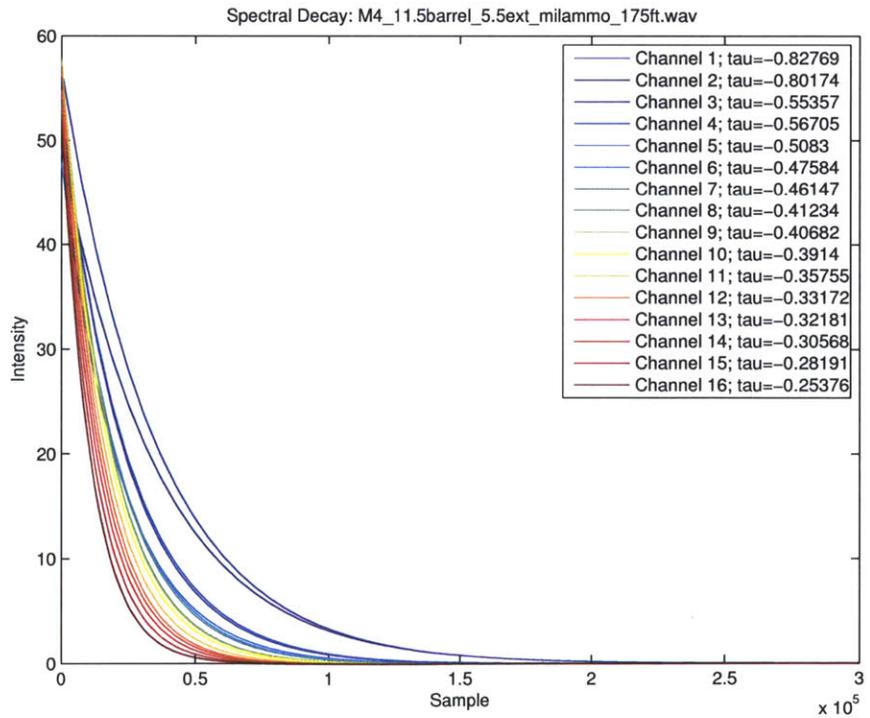**Figure A.2** – Spectrogram plot for M4 rifle with 11.5 inch barrel and 5.5 inch extension, civilian ammo, at 175 feet
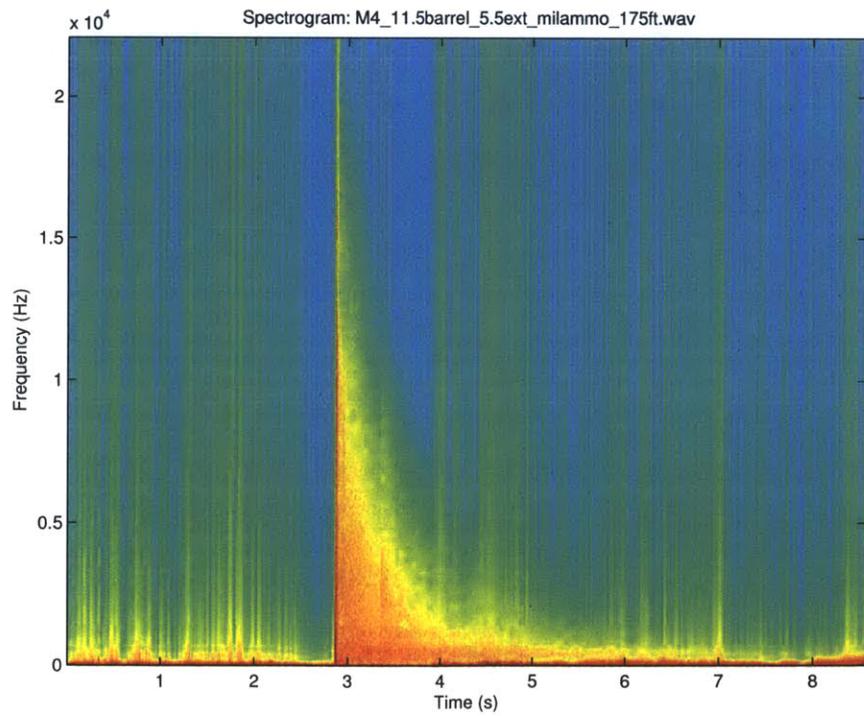
**Figure A.3** – Spectrogram plot for M4 rifle with 11.5 inch barrel and 5.5 inch extension, military ammo, at 175 feet
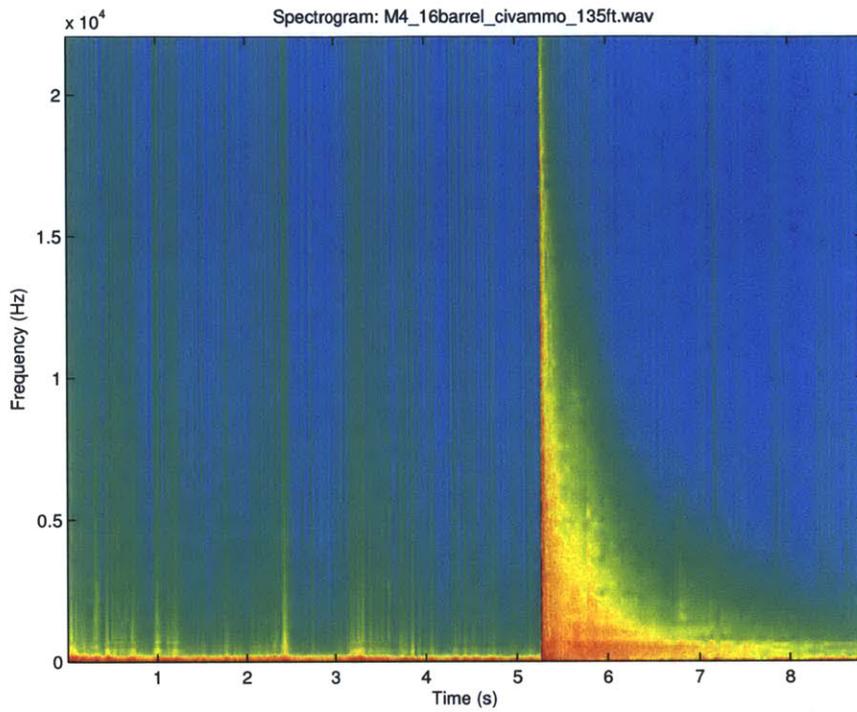
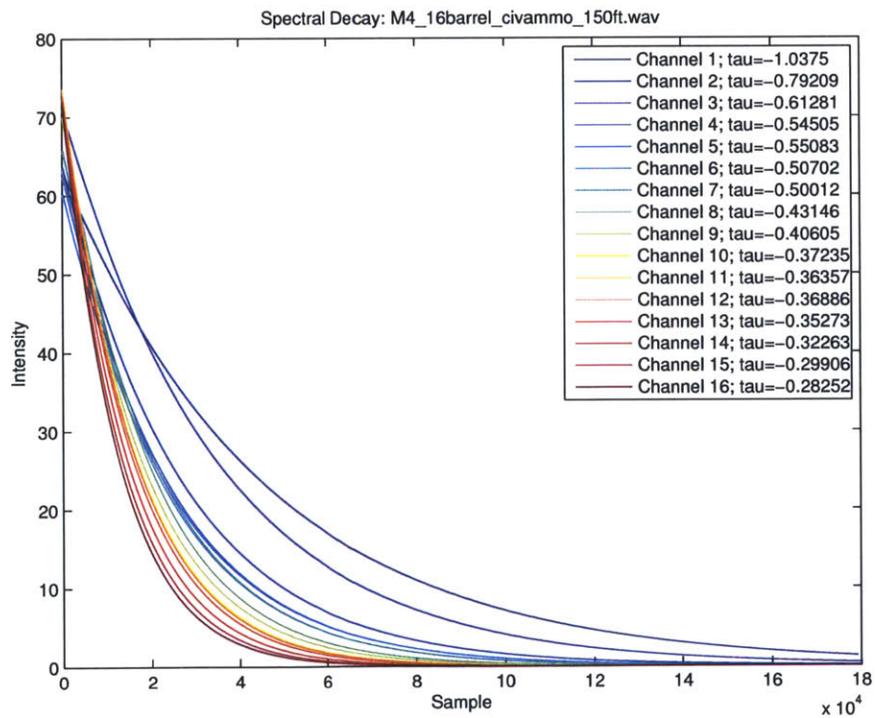**Figure A.4** – Spectrogram plot for M4 rifle with 16 inch barrel, civilian ammo, at 135 feet
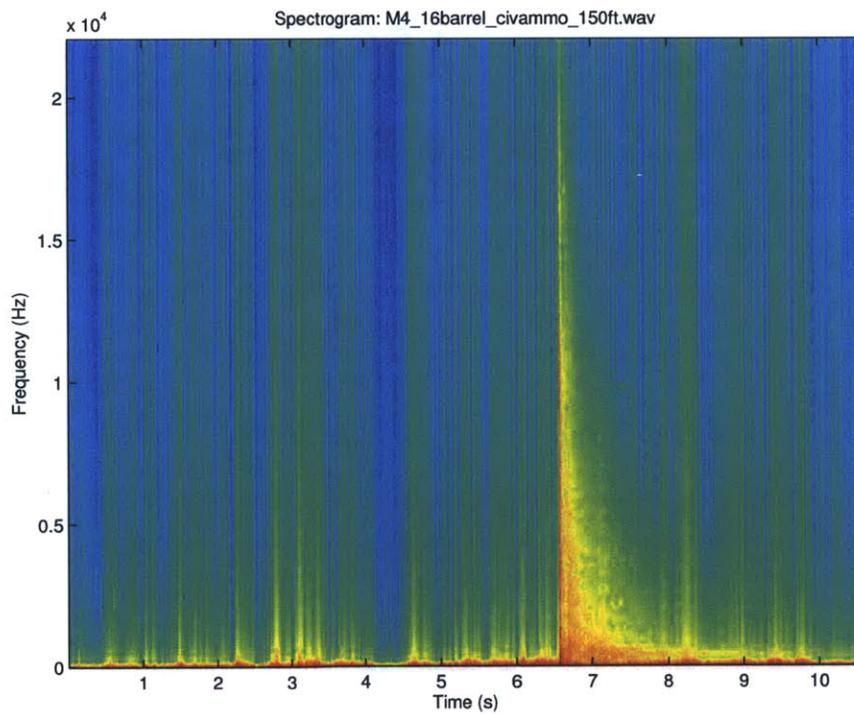
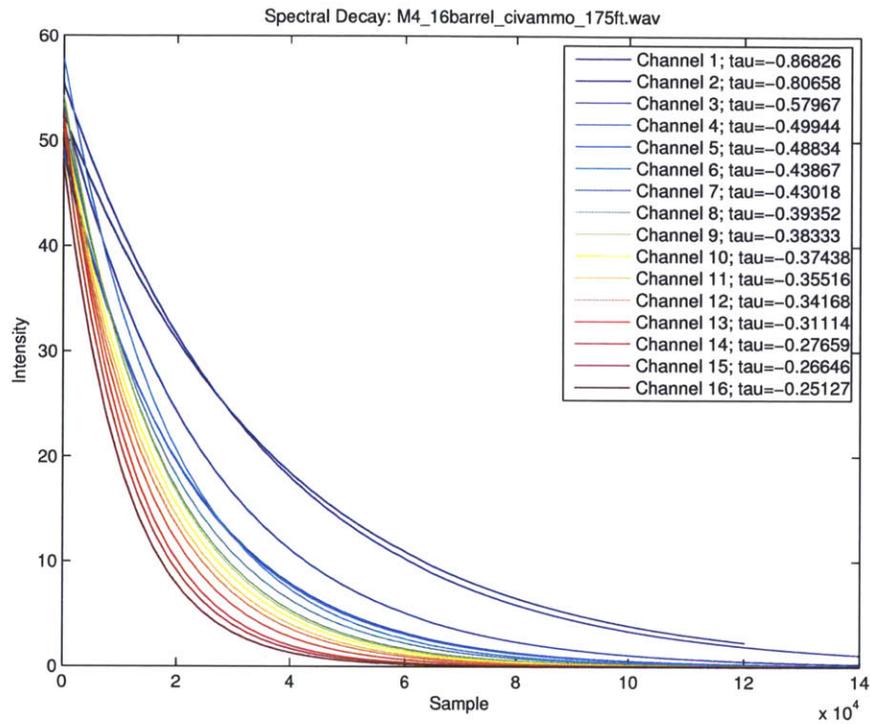**Figure A.5** – Spectrogram plot for M4 rifle with 16 inch barrel, civilian ammo, at 150 feet
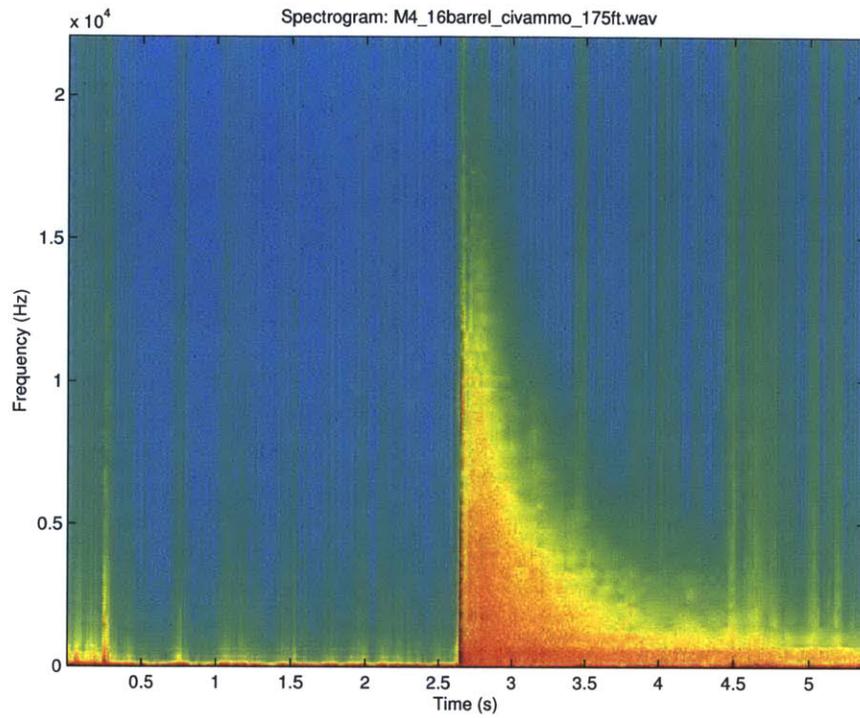
Figure A.6 – Spectrogram plot for M4 rifle with 16 inch barrel, civilian ammo, at 175 feet
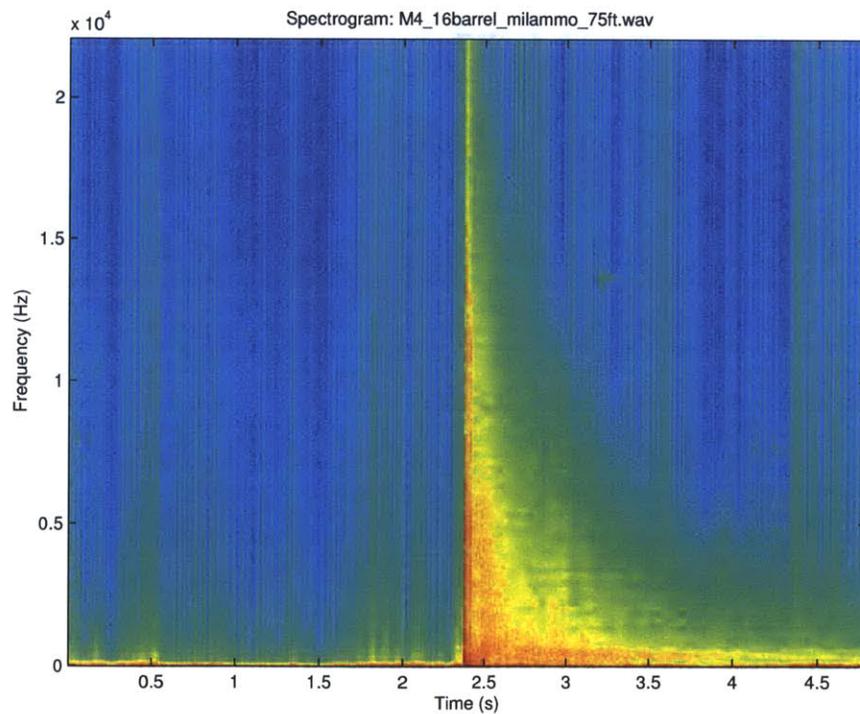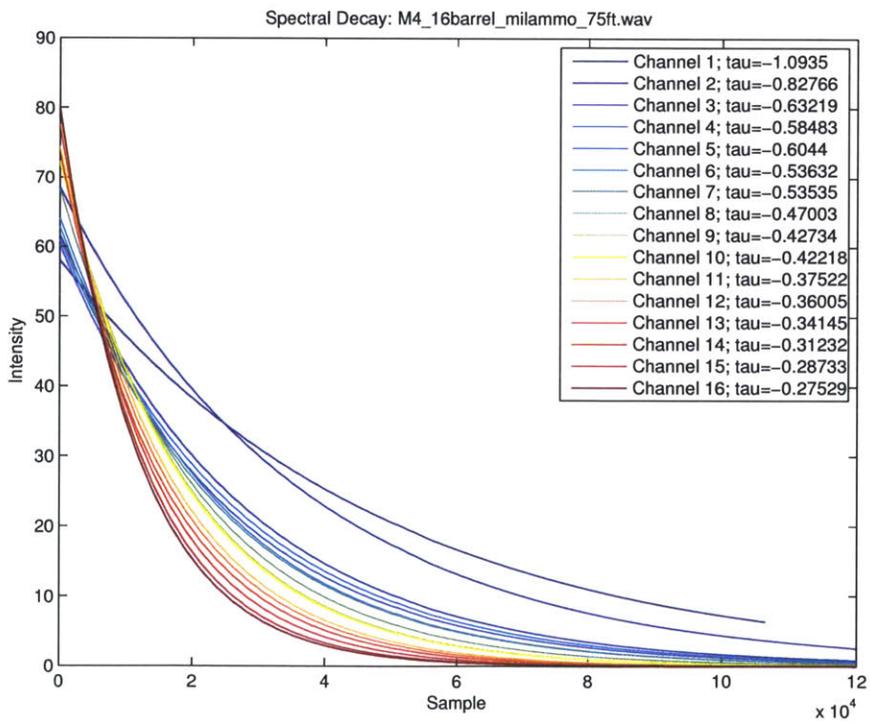

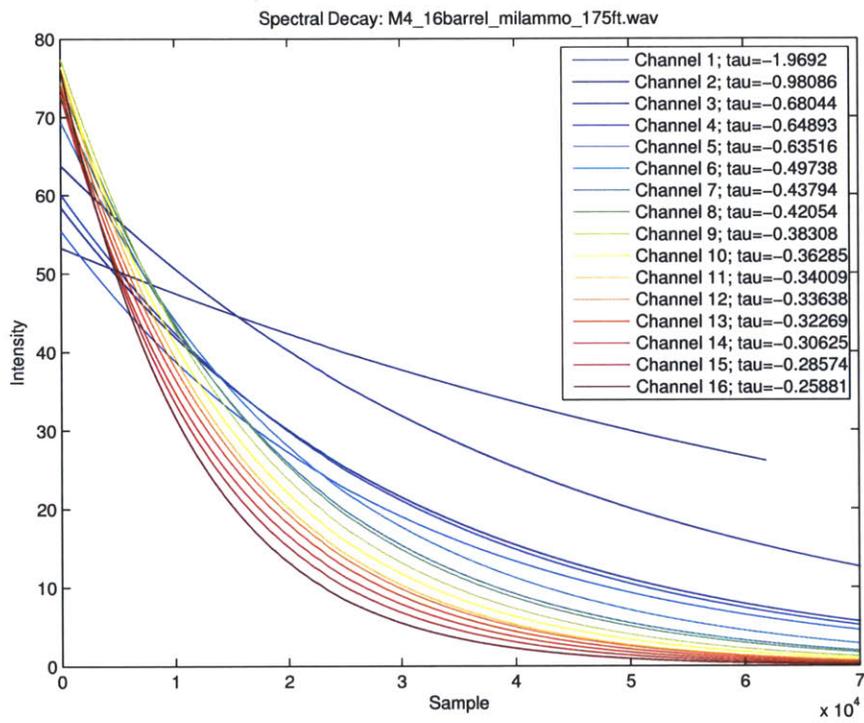**Figure A.7** – Spectrogram plot for M4 rifle with 16 inch barrel, military ammo, at 75 feet
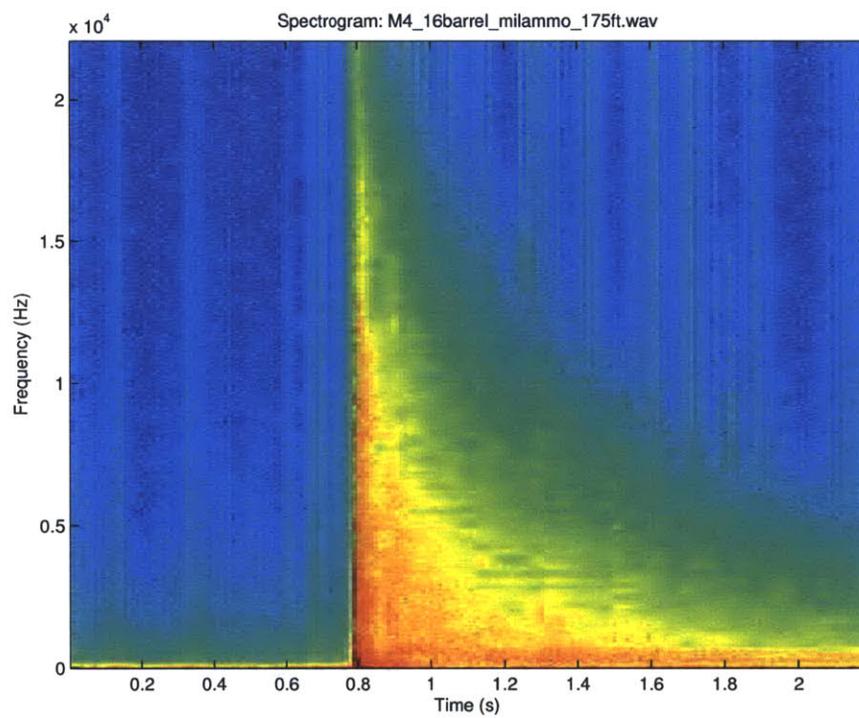
Figure A.8 – Spectrogram plot for M4 rifle with 16 inch barrel, military ammo, at 175 feet

B   Gunshot Decay Plots

These are simulation plots ((figures B.1, B.2, B.3, B.4, B.5, B.6, B.8, and B.7) generated by analyzing real gunshot data recorded and provided by the United States military. Each plot shows sixteen individual decay traces. Each trace corresponds to the output from one channel of the Analog Bionic Ear Processor. Channel one is the lowest frequency and channel sixteen is the highest frequency. We can see two phenomena in these plots: 1)in general, the wave-fronts ($t = 0$) contain more energy as the frequency gets lower (channel number reduces), and 2)the decay times become longer as the frequency gets lower.
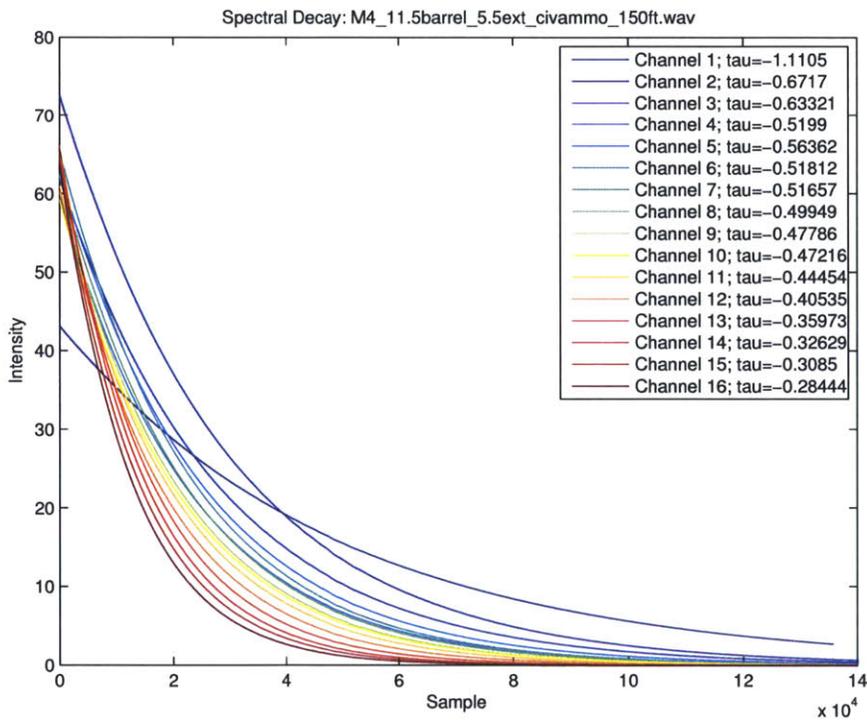
Figure B.1 – Decay plot for M4 rifle with 11.5 inch barrel and 5.5 inch extension, civilian ammo, at 150 feet

51
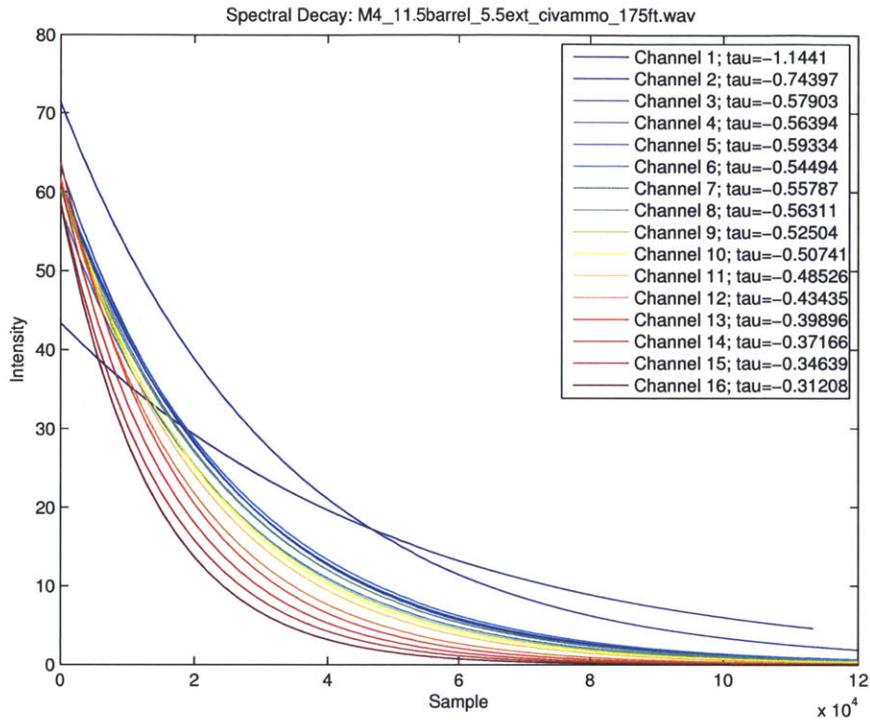
Figure B.2 – Decay plot for M4 rifle with 11.5 inch barrel and 5.5 inch extension, civilian ammo, at 175 feet

Figure B.3 – Decay plot for M4 rifle with 11.5 inch barrel and 5.5 inch extension, military ammo, at 175 feet
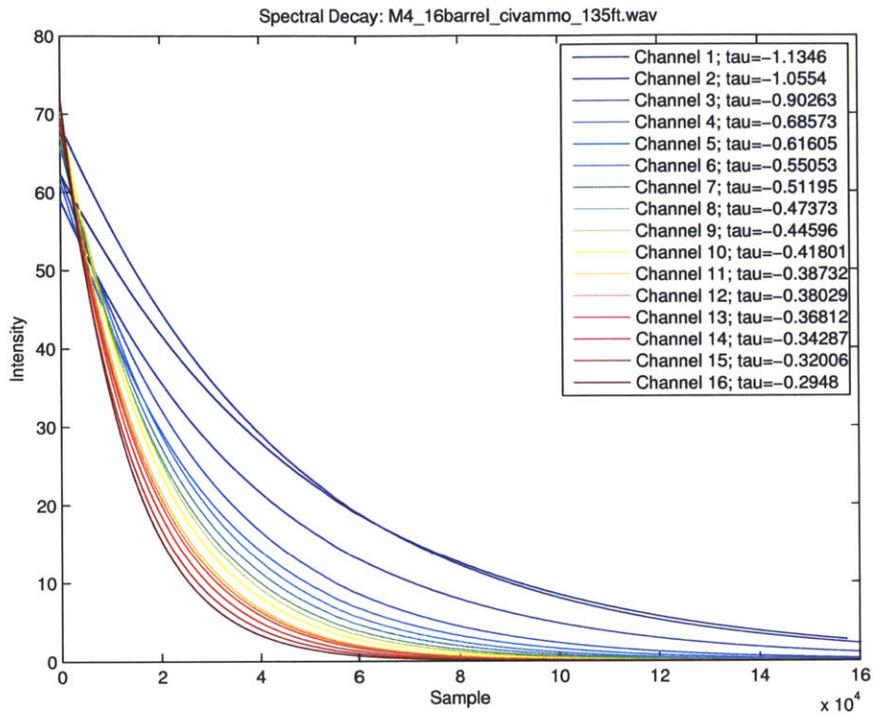
Figure B.4 – Decay plot for M4 rifle with 16 inch barrel, civilian ammo, at 135 feet

Figure B.5 – Decay plot for M4 rifle with 16 inch barrel, civilian ammo, at 150 feet

53

Spectral Decay: M4_16barrel_civammo_175ft.wav

Channel 1; tau=-0.86826
Channel 2; tau=-0.80658
Channel 3; tau=-0.57967
Channel 4; tau=-0.49944
Channel 5; tau=-0.48834
Channel 6; tau=-0.43867
Channel 7; tau=-0.43018
Channel 8; tau=-0.39352
Channel 9; tau=-0.38333
Channel 10; tau=-0.37438
Channel 11; tau=-0.35516
Channel 12; tau=-0.34168
Channel 13; tau=-0.31114
Channel 14; tau=-0.27659
Channel 15; tau=-0.26646
Channel 16; tau=-0.25127

Figure B.6 – Decay plot for M4 rifle with 16 inch barrel, civilian ammo, at 175 feet

Spectral Decay: M4_16barrel_milammo_75ft.wav

Channel 1; tau=-1.0935
Channel 2; tau=-0.82766
Channel 3; tau=-0.63219
Channel 4; tau=-0.58483
Channel 5; tau=-0.6044
Channel 6; tau=-0.53632
Channel 7; tau=-0.53535
Channel 8; tau=-0.47003
Channel 9; tau=-0.42734
Channel 10; tau=-0.42218
Channel 11; tau=-0.37522
Channel 12; tau=-0.36005
Channel 13; tau=-0.34145
Channel 14; tau=-0.31232
Channel 15; tau=-0.28733
Channel 16; tau=-0.27529

Figure B.7 – Decay plot for M4 rifle with 16 inch barrel, military ammo, at 75 feet

Figure B.8 – Decay plot for M4 rifle with 16 inch barrel, military ammo, at 175 feet

C Gunshot Noise Suppression Plots

These plots compare single gunshots with the transient noise suppression AGC system on and off. With the AGC on, the gunshot transient is nearly entirely eliminated.

(a) AGC off   (b) AGC on
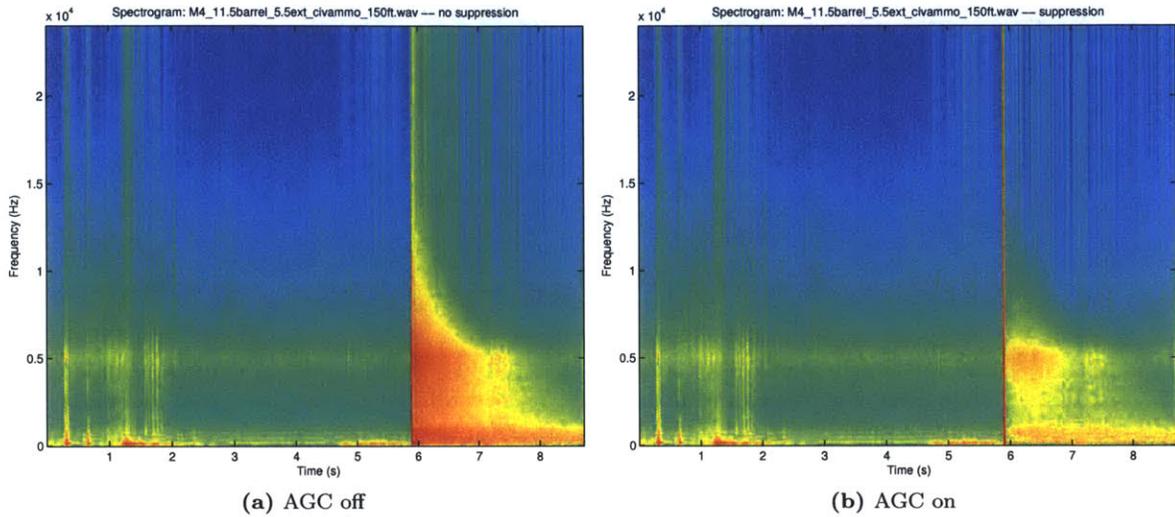
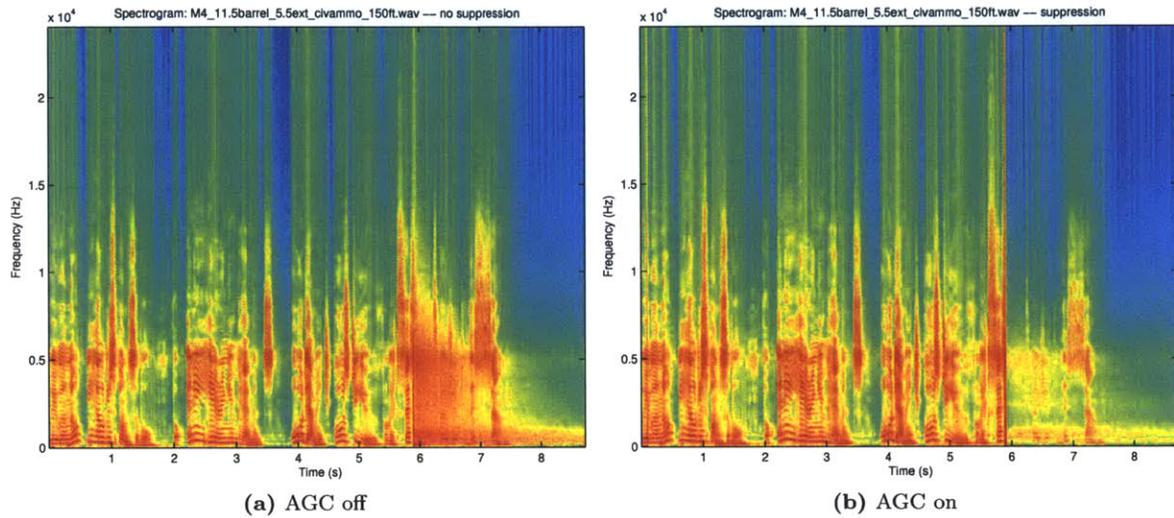**Figure C.1** – Comparison of gunshot (M4, 11.5 inch barrel with 5.5 inch extension, civilian ammo, at 150 ft.) with AGC system on and off

(a) AGC off   **(b)** AGC on
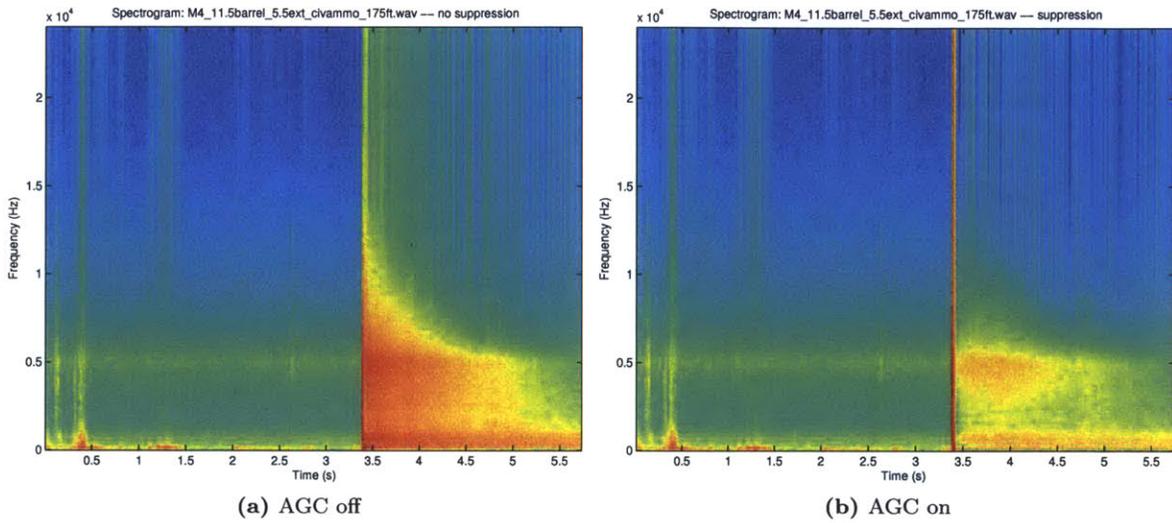
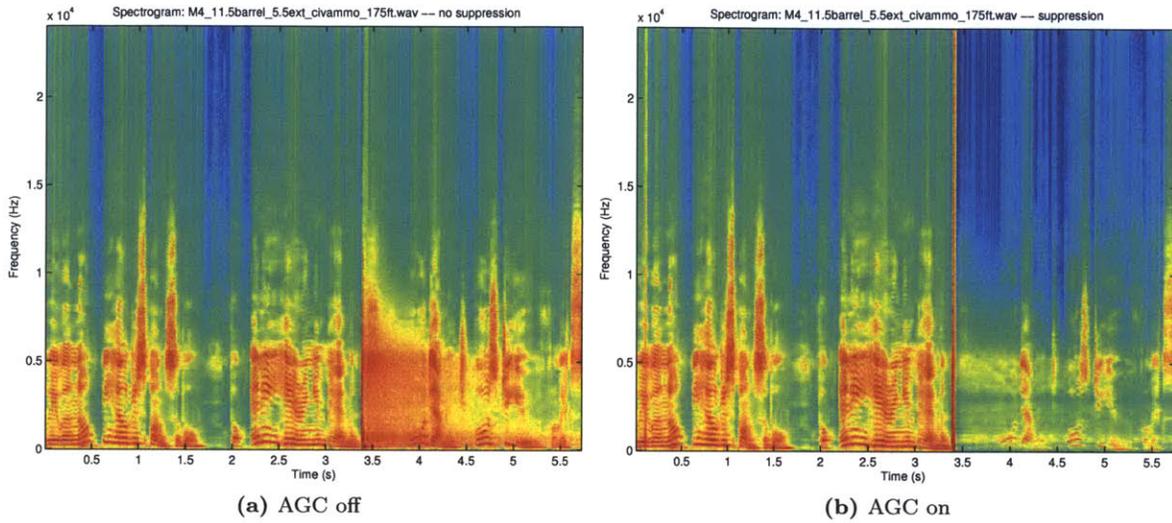**Figure C.2** – Comparison of gunshot (M4, 11.5 inch barrel with 5.5 inch extension, civilian ammo, at 175 ft.) with AGC system on and off

(a) AGC off   **(b)** AGC on
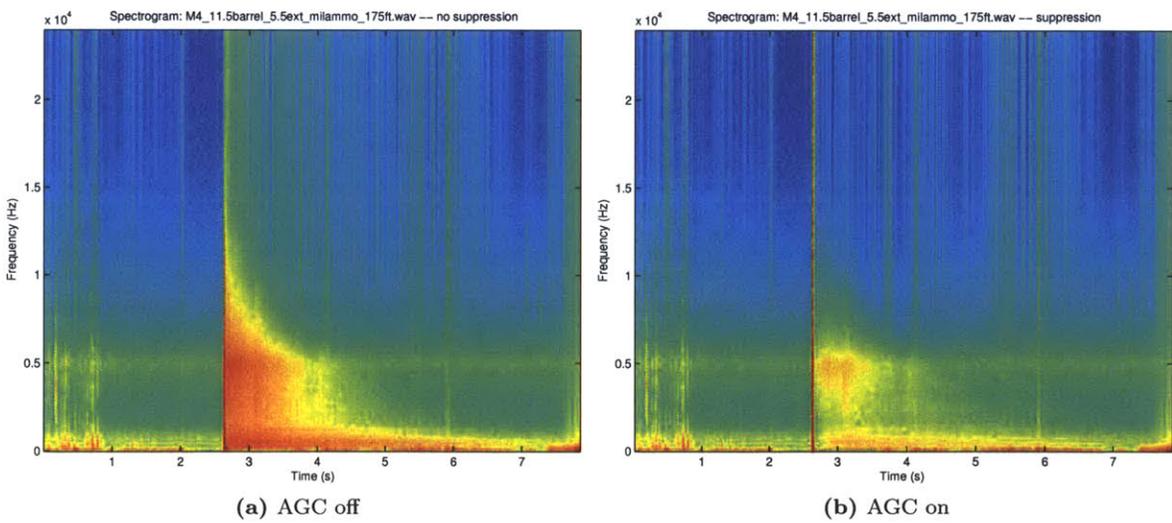
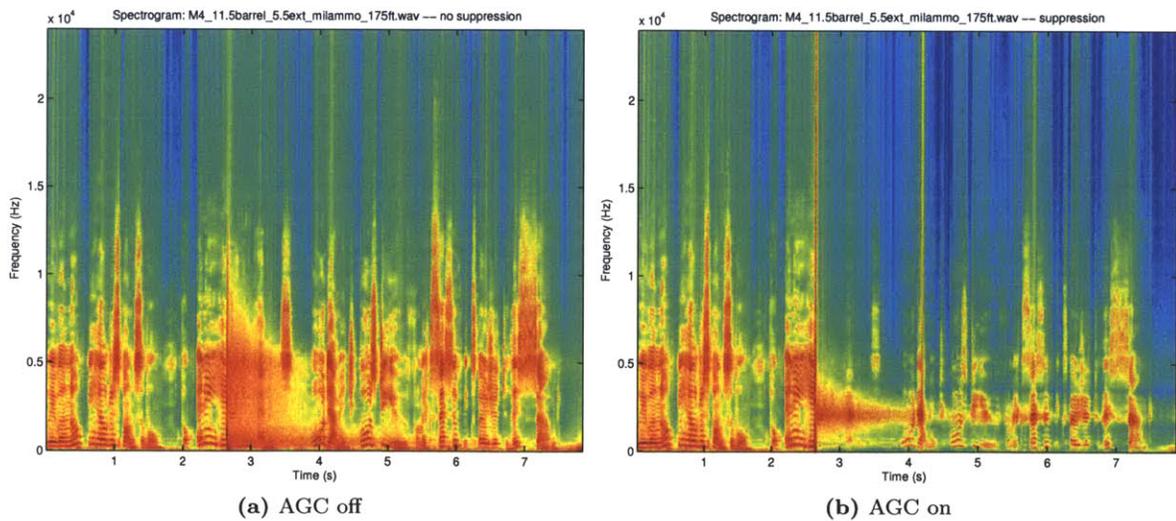**Figure C.3** – Comparison of gunshot (M4, 11.5 inch barrel with 5.5 inch extension, military ammo, at 175 ft.) with AGC system on and off

(a) AGC off   **(b)** AGC on
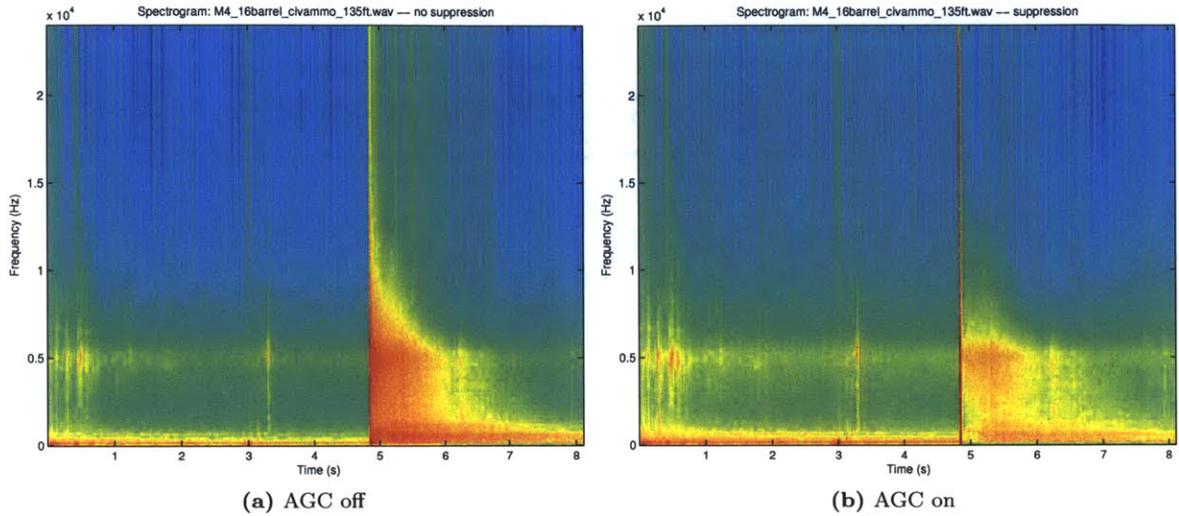
Figure C.4 – Comparison of gunshot (M4, 16 inch barrel, civilian ammo, at 135 ft.) with AGC system on and off

(a) AGC off   **(b)** AGC on
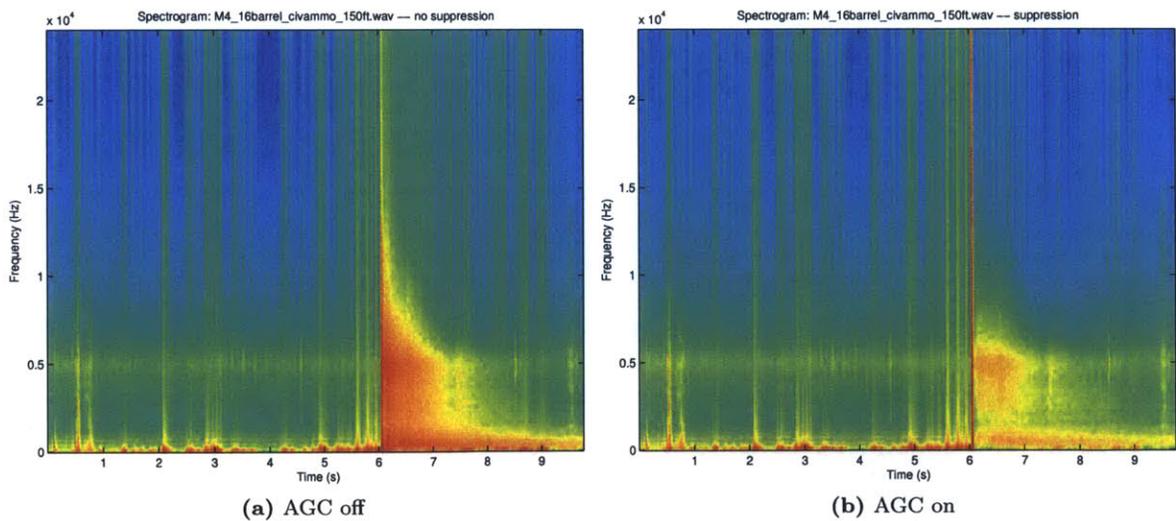
Figure C.5 – Comparison of gunshot (M4, 16 inch barrel, civilian ammo, at 150 ft.) with AGC system on and off

(a) AGC off

(b) AGC on
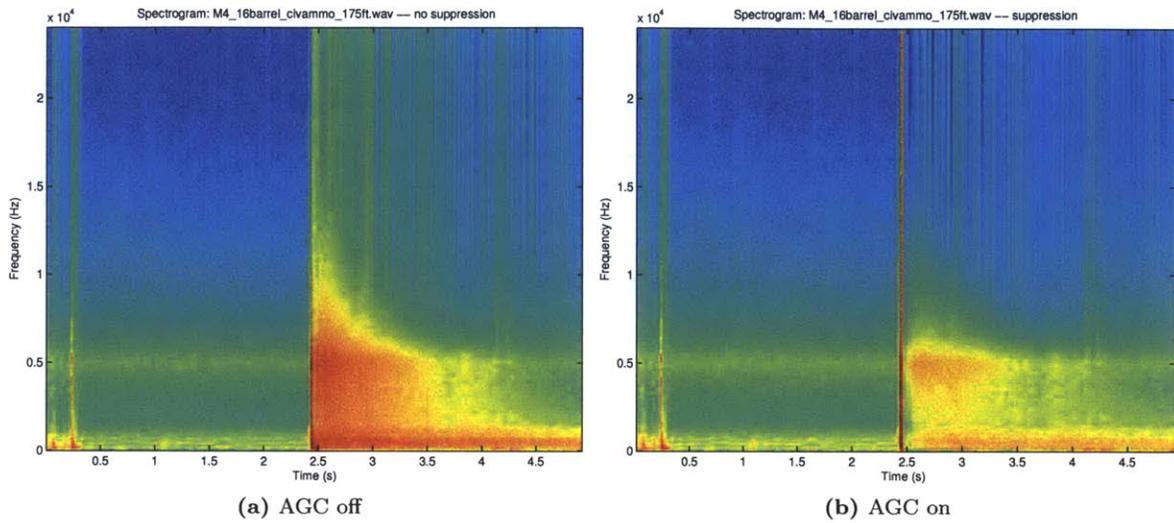
Figure C.6 – Comparison of gunshot (M4, 16 inch barrel, civilian ammo, at 175 ft.) with AGC system on and off

(a) AGC off

(b) AGC on
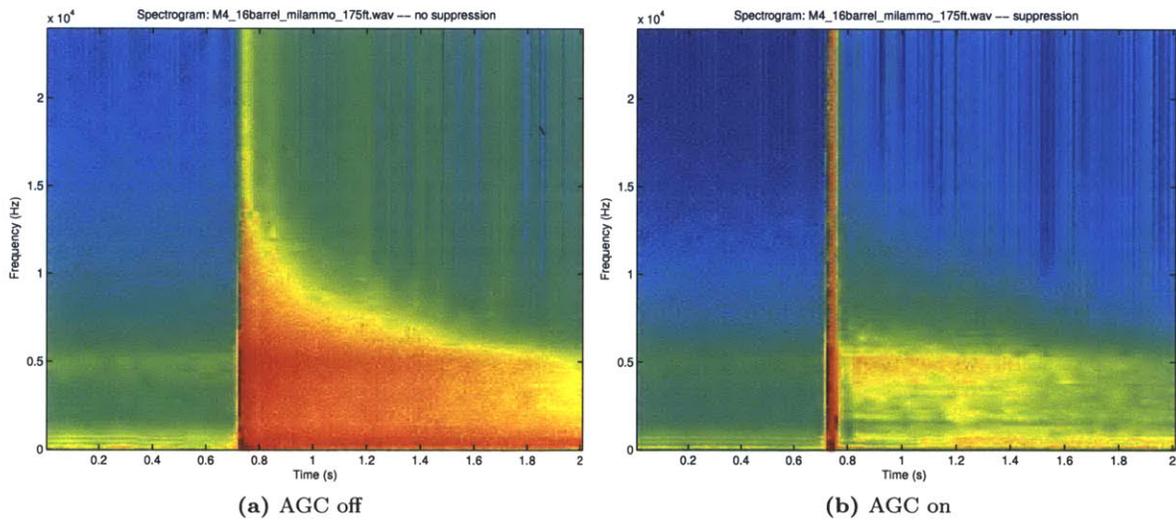
Figure C.7 – Comparison of gunshot (M4, 16 inch barrel, military ammo, at 175 ft.) with AGC system on and off

(a) AGC off   (b) AGC on
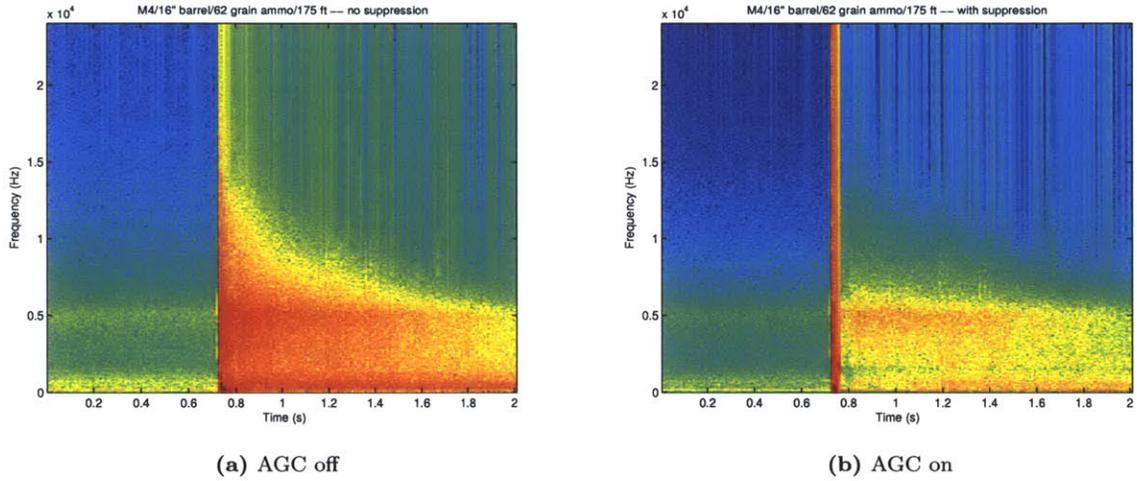
Figure C.8 – Comparison of gunshot (M4, 16 inch barrel, 62 grain ammo, at 175 ft.) with AGC system on and off

D   Gunshot With Speech Noise Suppression Plots

The plots compare single gunshots plus speech with the transient noise suppression AGC system on and off. With the AGC on, speech that occurs just after a gunshot become more pronounced as compared to the noise, thus improving intelligibility.

(a) AGC off   (b) AGC on

Figure D.1 – Comparison of gunshot (M4, 11.5 inch barrel with 5.5 inch extension, civilian ammo, at 150 ft.) plus speech with AGC system on and off

(a) AGC off

(b) AGC on

Figure D.2 – Comparison of gunshot (M4, 11.5 inch barrel with 5.5 inch extension, civilian ammo, at 175 ft.) plus speech with AGC system on and off

(a) AGC off

(b) AGC on

Figure D.3 – Comparison of gunshot (M4, 11.5 inch barrel with 5.5 inch extension, military ammo, at 175 ft.) plus speech with AGC system on and off

(a) AGC off

(b) AGC on
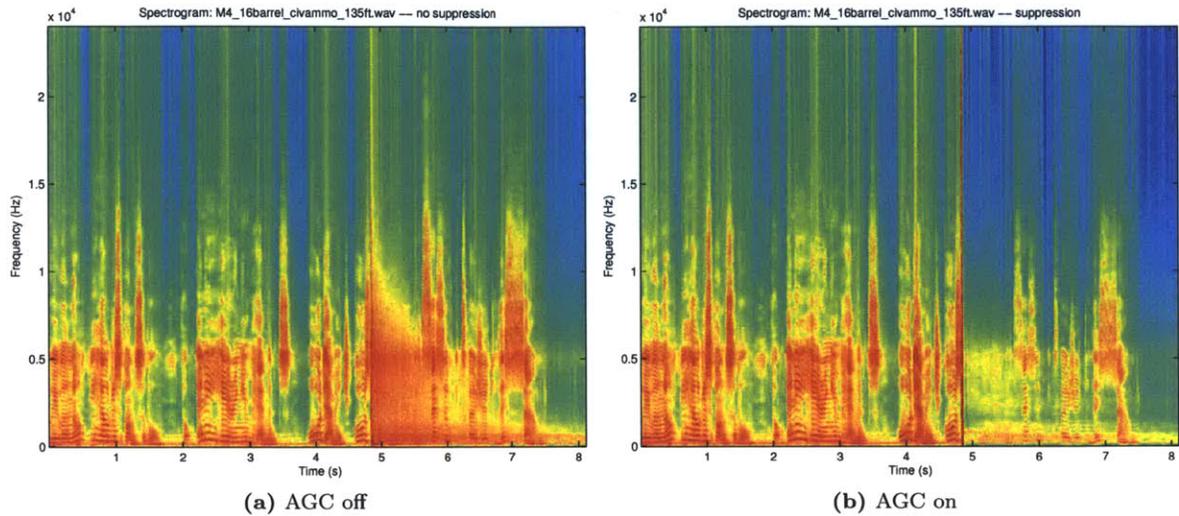
Figure D.4 – Comparison of gunshot (M4, 16 inch barrel, civilian ammo, at 135 ft.) plus speech with AGC system on and off


(a) AGC off

(b) AGC on
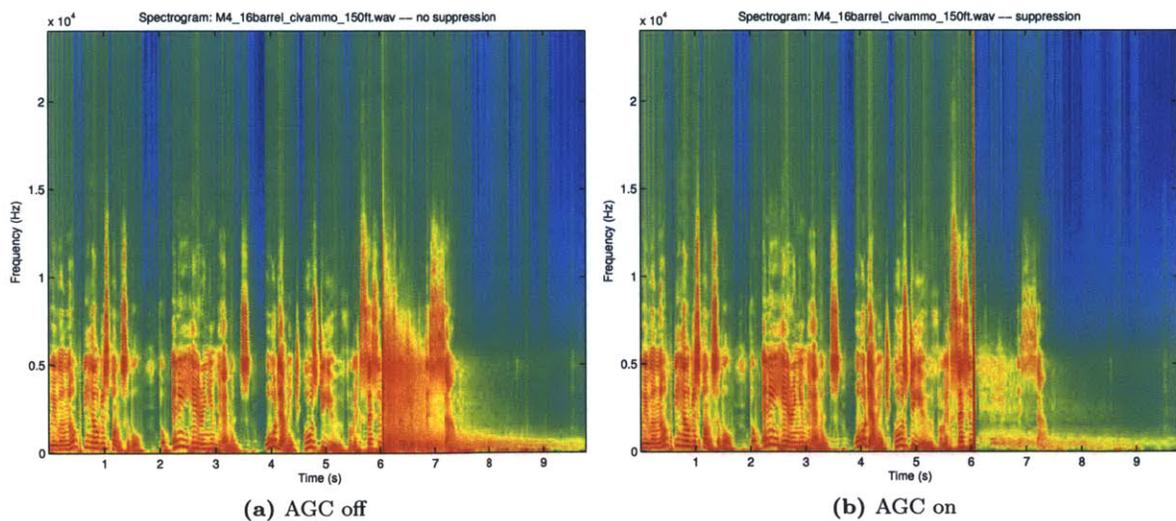
Figure D.5 – Comparison of gunshot (M4, 16 inch barrel, civilian ammo, at 150 ft.) plus speech with AGC system on and off

(a) AGC off

(b) AGC on
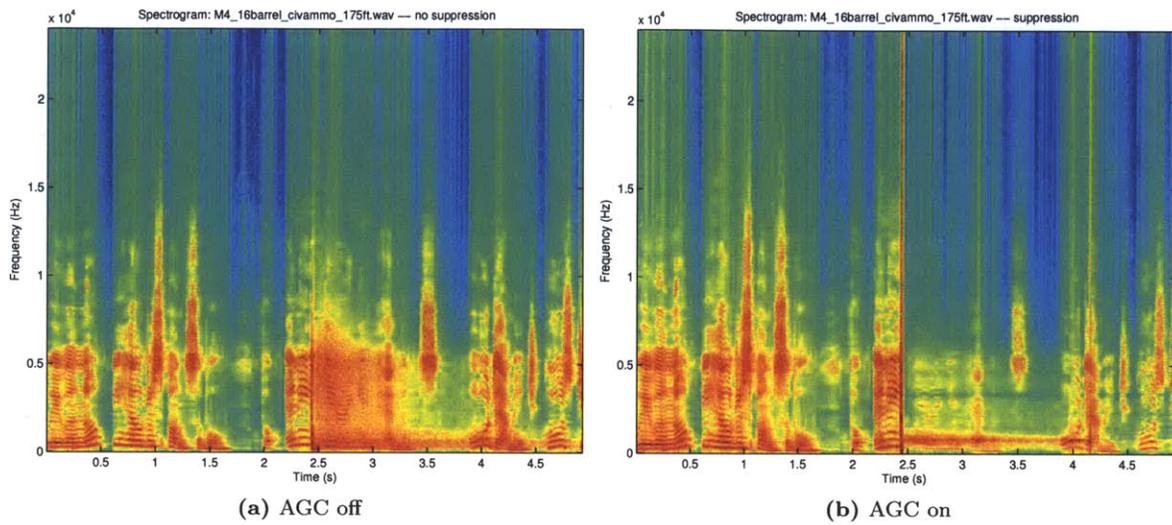
Figure D.6 – Comparison of gunshot (M4, 16 inch barrel, civilian ammo, at 175 ft.) plus speech with AGC system on and off

(a) AGC off

(b) AGC on
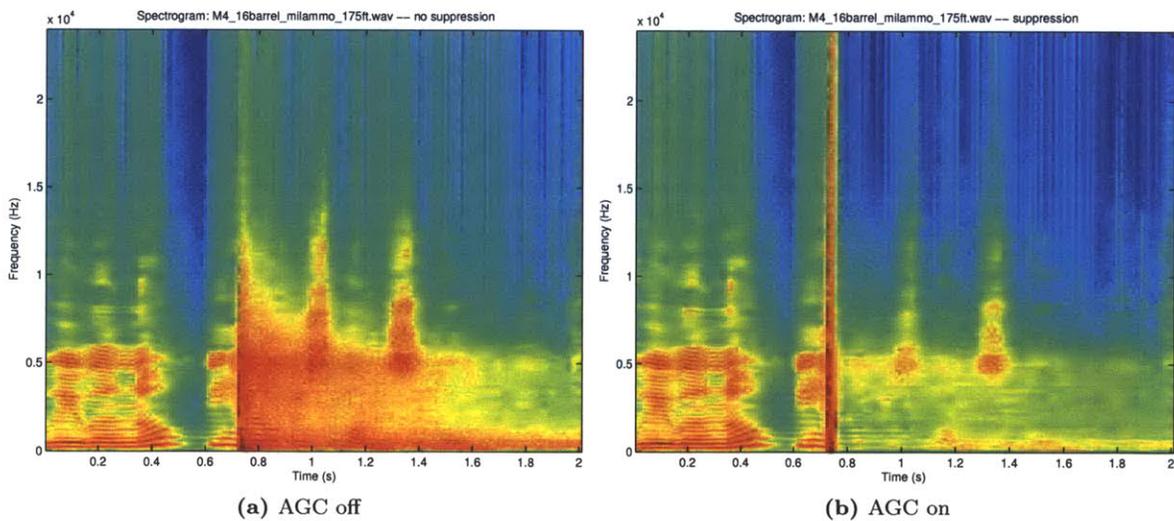
Figure D.7 – Comparison of gunshot (M4, 16 inch barrel, military ammo, at 175 ft.) plus speech with AGC system on and off

(a) AGC off
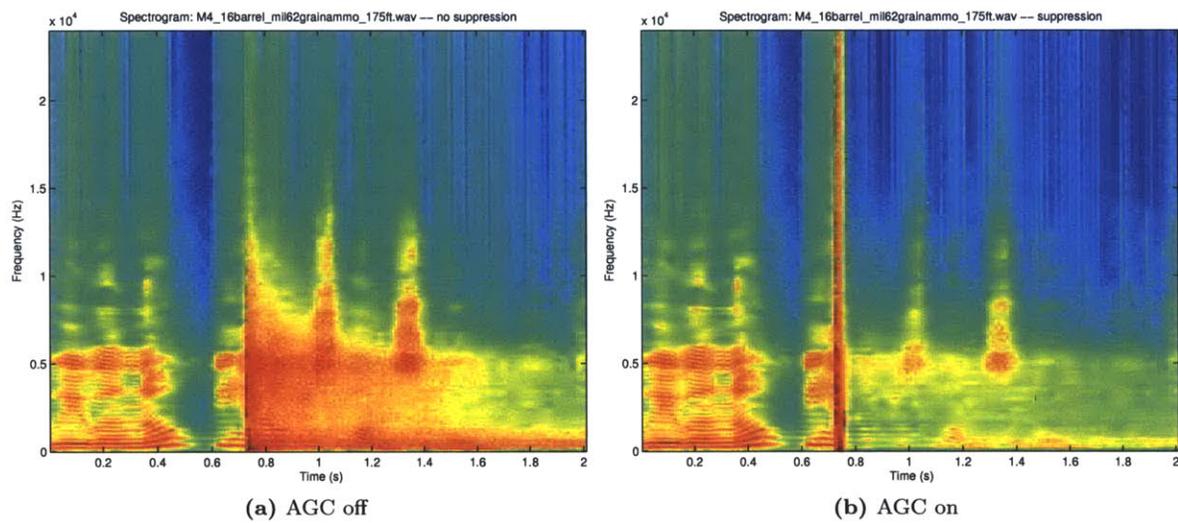
(b) AGC on

Figure D.8 – Comparison of gunshot (M4, 16 inch barrel, 62 grain ammo, at 175 ft.) plus speech with AGC system on and off

E   Stationary Noise Reduction Plots

These plots compare stationary noise with the compressive AGC system on and off. With the AGC on, after time, the frequency bands that contain the most noise are attenuated.

(a) AGC off   (b) AGC on
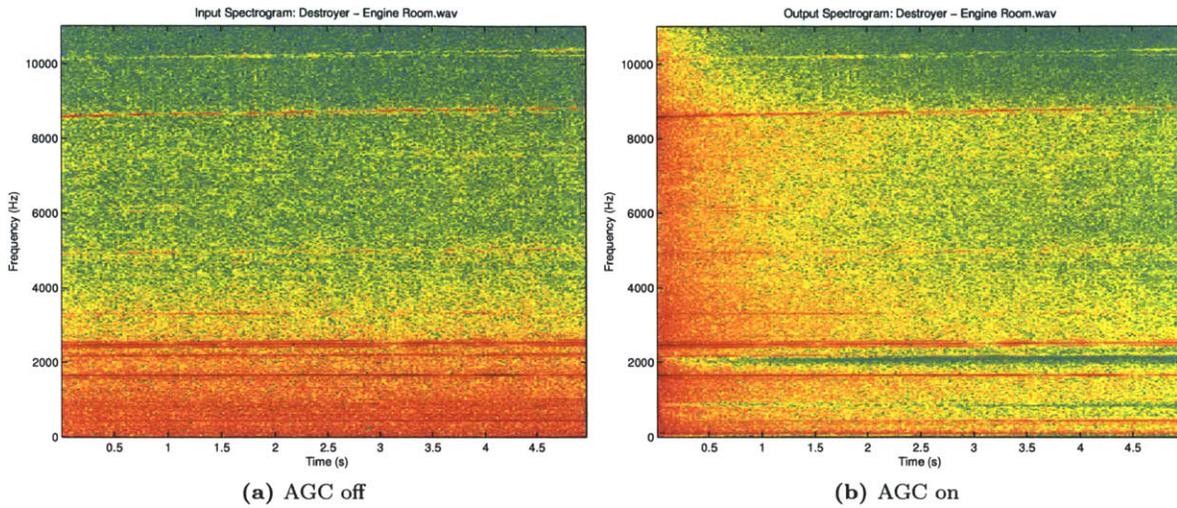
Figure E.1 – Comparison of Destroyer (engine room) ship noise with AGC system on and off

(a) AGC off   (b) AGC on
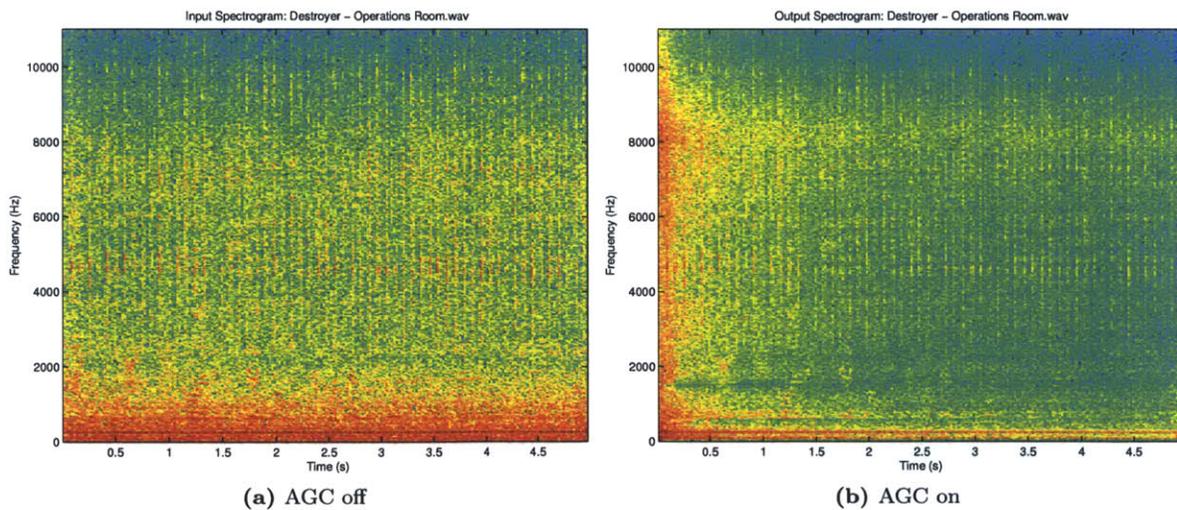
Figure E.2 – Comparison of Destroyer (operations room) ship noise with AGC system on and off
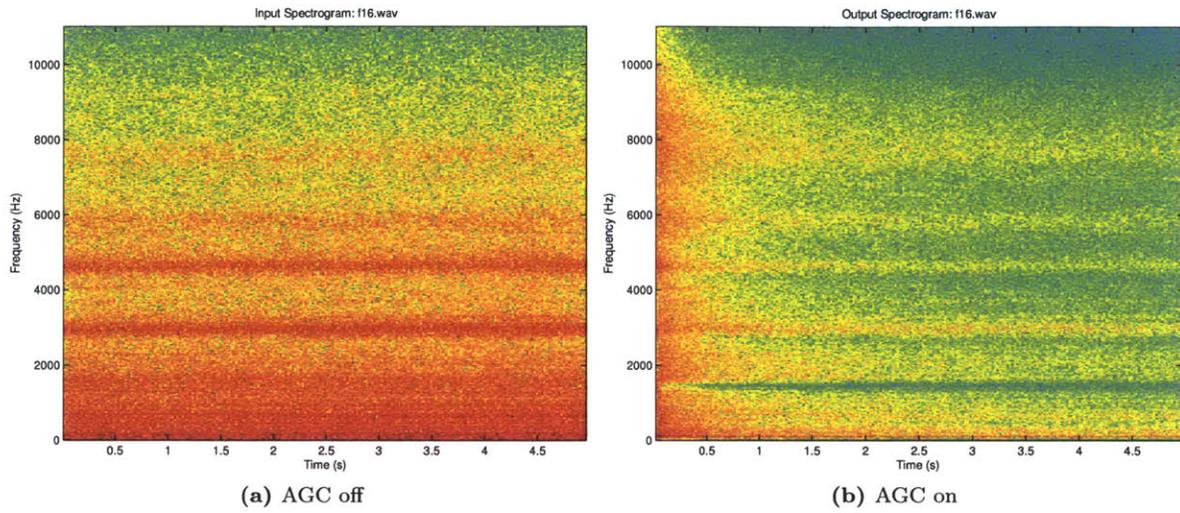
Figure E.3 – Comparison of F16 jet noise with AGC system on and off
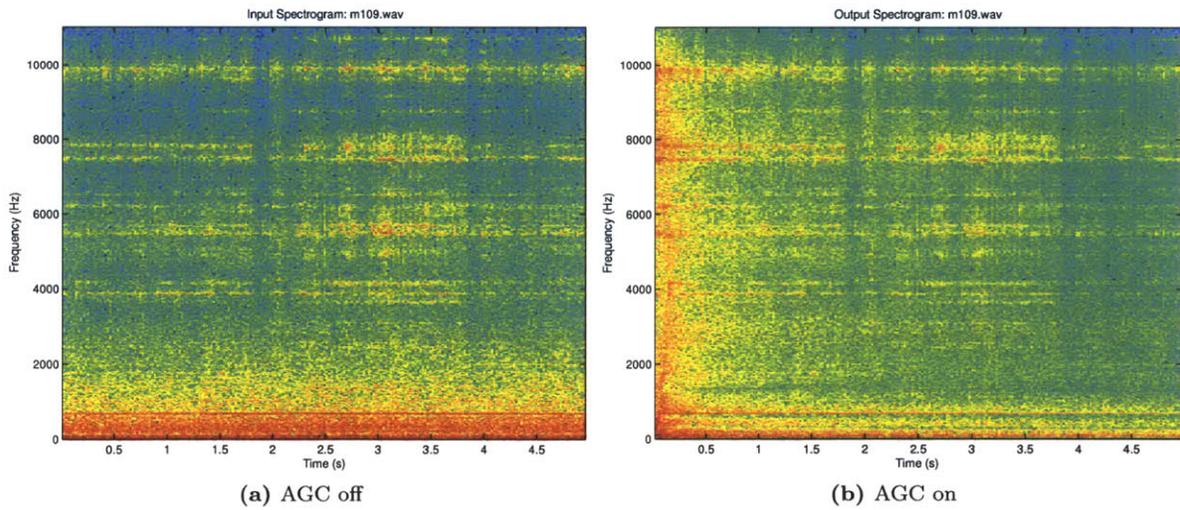
Figure E.4 – Comparison of M109 tank noise with AGC system on and off

70

(a) AGC off
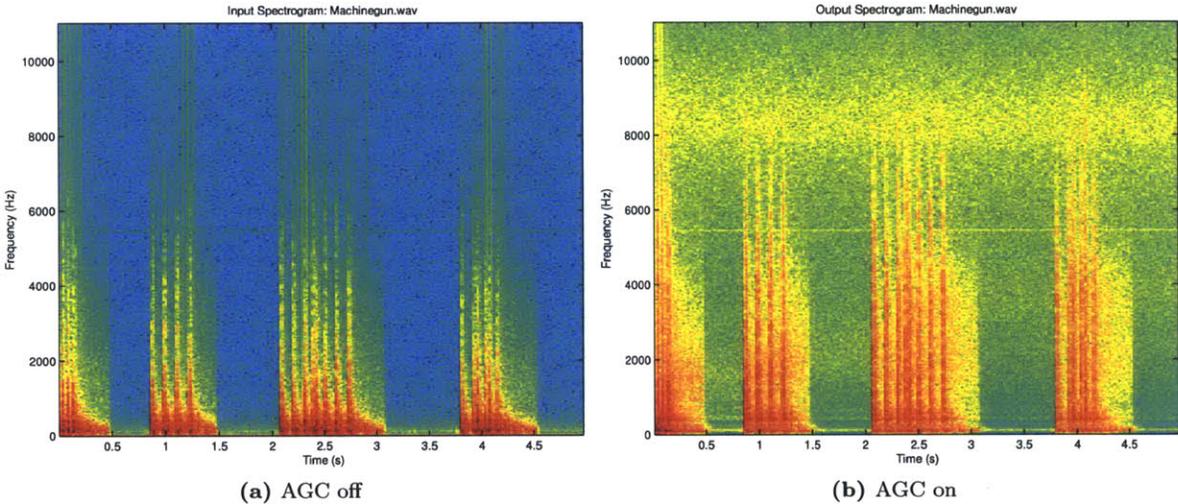
(b) AGC on

Figure E.5 – Comparison of machine gun noise with AGC system on and off

F   Stationary Noise Reduction With Speech Plots

These plots compare stationary noise plus speech with the compressive AGC system on and off. With the AGC on, after time, the noise is attenuated leaving the speech more pronounced, thus improving intelligibility.
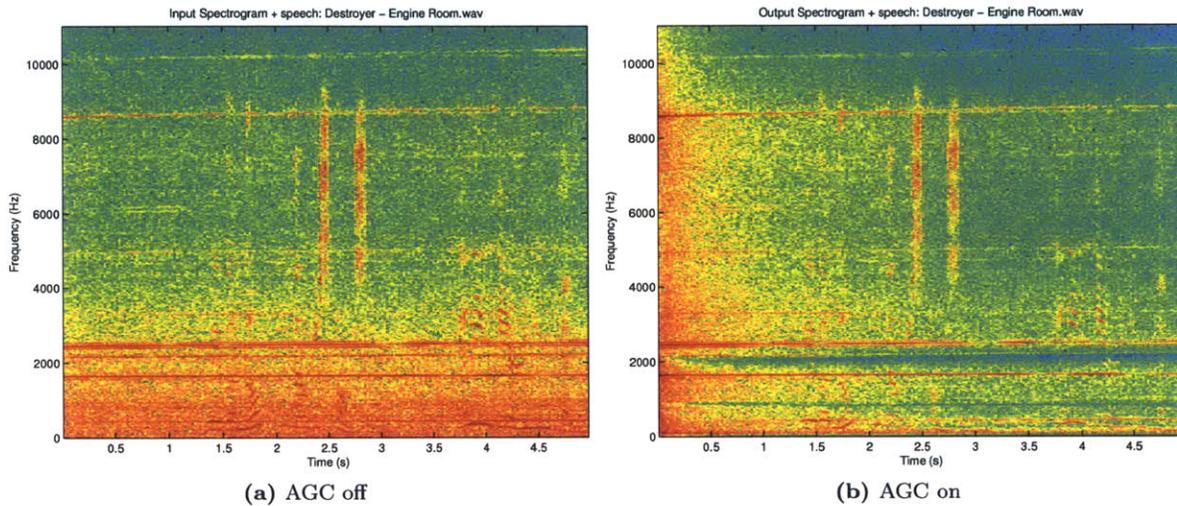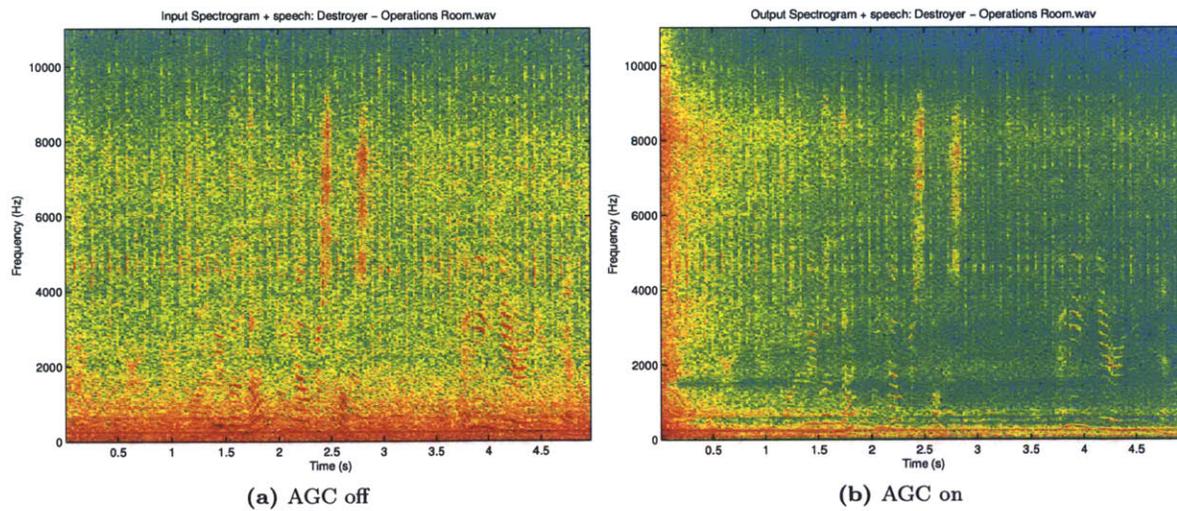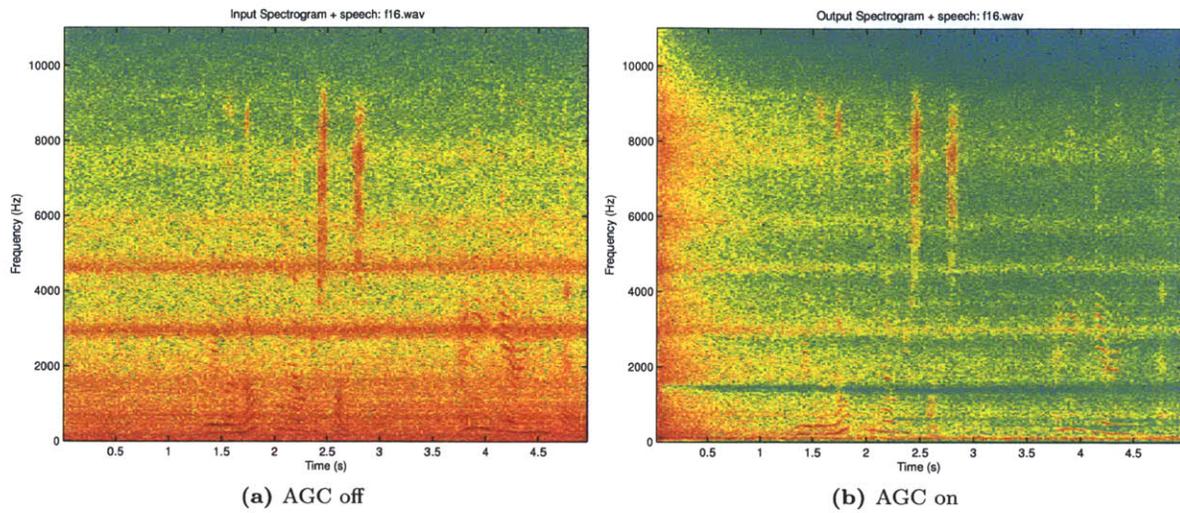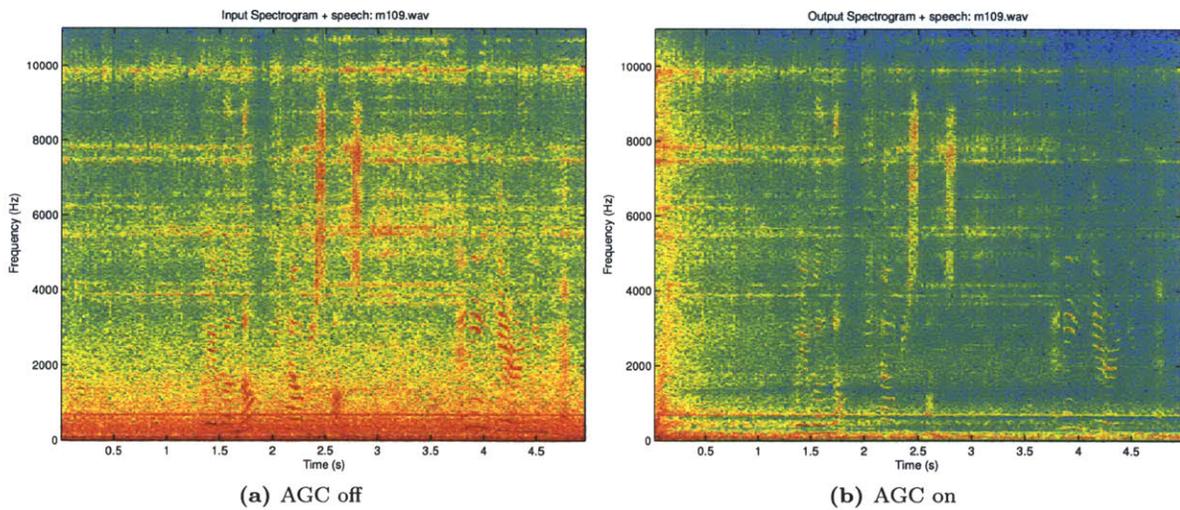

(a) AGC off

(b) AGC on

Figure F.1 – Comparison of Destroyer (engine room) ship noise plus speech with AGC system on and off

(a) AGC off

(b) AGC on

Figure F.2 – Comparison of Destroyer (operations room) ship noise plus speech with AGC system on and off

(a) AGC off

(b) AGC on

Figure F.3 – Comparison of F16 jet noise plus speech with AGC system on and off

(a) AGC off

(b) AGC on

Figure F.4 – Comparison of M109 tank noise plus speech with AGC system on and off
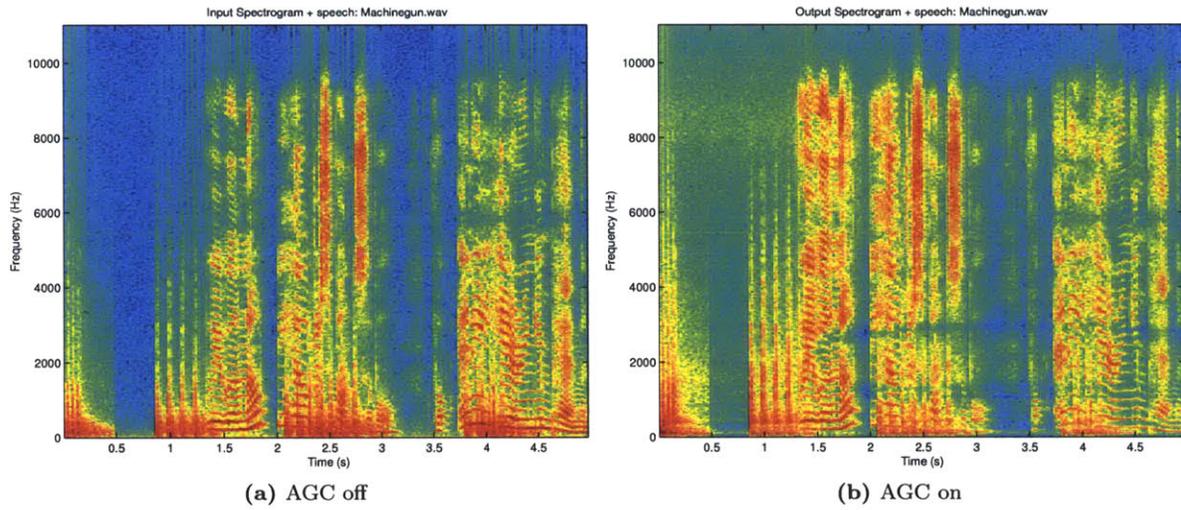
(a) AGC off

(b) AGC on

Figure F.5 – Comparison of machine gun noise plus speech with AGC system on and off