



A Continuous-time Speech Enhancement Front-end for Microphone Inputs

Heejong Yoo¹, Rich Ellis¹, David V. Anderson¹, Paul Hasler¹

David W. Graham¹, Mat Hans

Mobile and Media Systems Laboratory

HP Laboratories Palo Alto

HPL-2002-311

November 7th, 2002*

noise
suppression,
speech
enhancement,
analog VLSI,
low-power

In this paper, we present a real-time noise suppression system implemented with analog VLSI. The algorithm implemented is designed to reduce stationary background noise in single-microphone signals while preserving the non-stationary signal component. Because the system relies on analog computation rather than digital, it has benefits such as extremely low power consumption and real-time computation. The algorithm is based on an adaptive Wiener filter algorithm, adapted to take advantage of analog processing capabilities provided by floating-gate analog VLSI circuits. Noise suppression processing is performed on continuous-time signals in one-third octave subbands. The analog components described as parts of this system include a C^4 second-order section band-pass filter, peak and minimum detectors, a translinear division circuit, and a differential multiplier. Floating-gate circuits are used to set bias points and adjust offsets.

* Internal Accession Date Only

Approved for External Publication

¹ School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA

© Copyright Hewlett-Packard Company 2002

A Continuous-time Speech Enhancement Front-end for Microphone Inputs

Heejong Yoo, Rich Ellis, David V. Anderson,
Paul Hasler, David W. Graham, and Mat Hans *

Abstract

In this paper, we present a real-time noise suppression system implemented with analog VLSI. The algorithm implemented is designed to reduce stationary background noise in single-microphone signals while preserving the non-stationary signal component. Because the system relies on analog computation rather than digital, it has benefits such as extremely low power consumption and real-time computation. The algorithm is based on an adaptive Wiener filter algorithm, adapted to take advantage of analog processing capabilities provided by floating-gate analog VLSI circuits. Noise suppression processing is performed on continuous-time signals in one-third octave subbands. The analog components described as parts of this system include a C^4 second-order section band-pass filter, peak and minimum detectors, a translinear division circuit, and a differential multiplier. Floating-gate circuits are used to set bias points and adjust offsets.

1 Introduction

Audio signal enhancement by removing additive background noise from corrupted noisy signal is not a new idea; but, with the prosperity of the portable communication devices, it has received renewed attention recently. Traditional methods of noise suppression include spectral subtraction, Wiener filtering, and a number of modifications on these methods that increase the intelligibility of the processed audio signal and/or reduce adverse artifacts (for example see [1, 2, 3, 4, 5, 6, 7]).

While most noise suppression methods focus on processing discrete-time audio signals, we propose a continuous-time noise suppression method in which the processing is performed on the microphone signal prior to the A/D conversion, as illustrated in Fig. 1. Analog audio signal enhancement received some attention in the 1960's, but interest in this has also revived recently due to its applications in telecommunications. Schroeder and Diethorn have presented methods of noise suppression using subband processing of analog and digital signals in [8] and [9], respectively. We are building on these and other concepts, using advances in technology and the new tools that are available, including the computational

*Heejong Yoo, Rich Ellis, David V. Anderson, Paul Hasler, and David W. Graham are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA. Mat Hans is with the Appliance Platforms Department, Hewlett-Packard Laboratories, Palo Alto, CA.

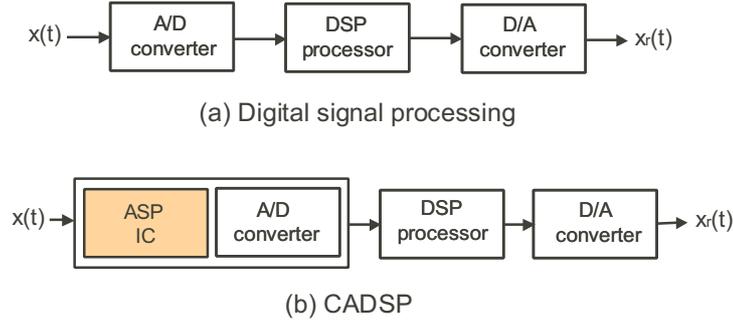


Figure 1: (a) Typical signal processing system where processing is done between the A/D and D/A converters. (b) Cooperative analog and digital signal processing (CADSP). Continuous-time noise suppression is done before the A/D converter.

potential of analog VLSI systems that use floating gate circuits, to design a system that operates in real time and uses extremely low amounts of power. By performing this portion of the processing in low-power analog circuits and by utilizing analog/digital computation in mutually beneficial way [10], we hope to enhance the overall functionality of an entire system. A review of the cooperative analog/digital computing paradigm with its benefits is presented in [11, 12].

This paper is organized as follows. In Section 2, we present the continuous-time noise suppression algorithm which exploits the characteristics of the sub-threshold operation of floating-gate circuits for real-time and low-power implementation. In Section 3, we present our simulations results. Section 4 discusses some of the circuit details for implementing the noise suppression algorithm and Section 5 contains measured results. We conclude this paper in Section 6.

2 Noise Suppression System

A common model for a noisy signal, $x(t)$, is a signal, $s(t)$, plus additive noise, $n(t)$, that is uncorrelated with the signal

$$x(t) = s(t) + n(t). \quad (1)$$

The goal is to design a real-time system that generates some optimal estimate, $\hat{s}(t)$, of $s(t)$ from $x(t)$. We assume that the additive noise is stationary over a long time period relative to the short-term stationary patterns of normal speech. The signal estimate, $\hat{s}(t)$, may be found in the frequency domain using spectral subtraction or by applying a Wiener filter gain. The basic Wiener gain is found from the power spectral densities (PSD) or from the estimated PSDs as

$$H(\omega) = \frac{\Phi_s(\omega)}{\Phi_n(\omega) + \Phi_s(\omega)}. \quad (2)$$

This can also be expressed as a function of the frequency dependent signal-to-noise ratio (SNR) as

$$H(\omega) = \frac{\Gamma^2(\omega)}{1 + \Gamma^2(\omega)}, \quad (3)$$

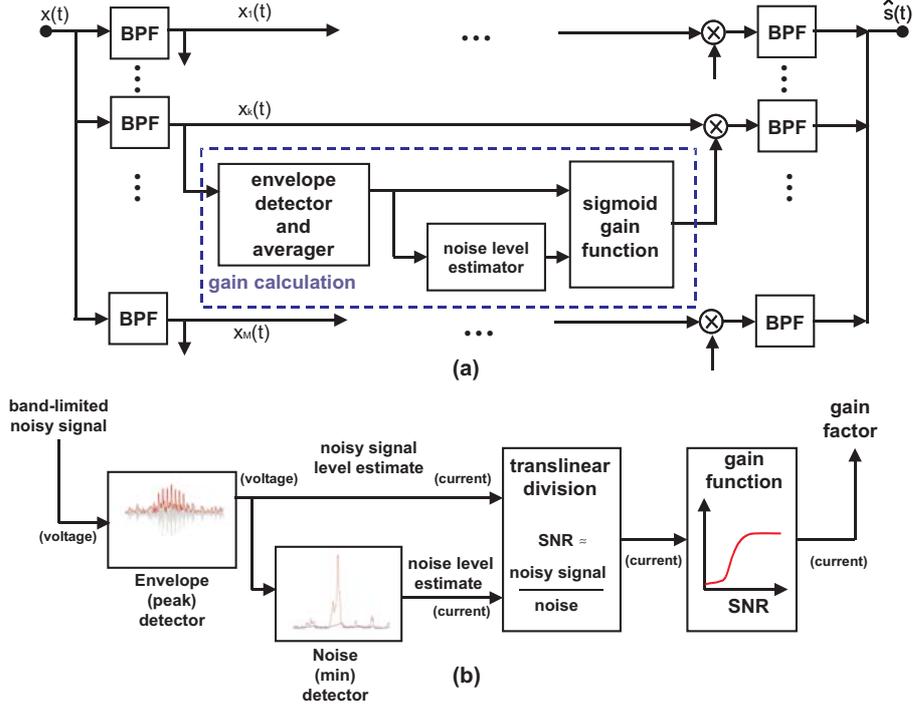


Figure 2: (a) Block diagram of continuous-time noise suppression system. The center frequency of filter bank is spaced exponentially. At each sub-band, gain is calculated from non-linear gain function. To achieve consistent noise suppression for different noise variance, noisy signal estimate, instead of signal estimate, is used as an input of non-linear gain function. The gain is then multiplied with sub-band signal and summed to build full-band signal estimate. (b) Details of the sub-band gain calculation—Within each frequency band, the noisy signal envelope is estimated using a peak detector. Based on the voltage output of the peak detector, the noise level is estimated using a minimum detector. The currents representing the noisy signal and noise levels are input to a translinear division circuit, which outputs a current representing the estimated signal-to-noise ratio. A nonlinear function is applied to the SNR current, and the resulting gain factor is multiplied with the band-limited noisy signal to produce a band-limited “clean” signal. Finally, the output of all of the bands are summed to reconstruct the signal with the noise components significantly reduced.

where $\Gamma^2(\omega) = \Phi_s(\omega)/\Phi_n(\omega)$.

The proposed noise suppression algorithm is targeted for implementation in a low-power analog computing IC with floating-gate circuits. Therefore, all computation blocks of the algorithm have been developed to reflect the strengths and limitations of the analog VLSI medium. Frequency domain processing is accomplished for the current system by using a one-third octave filter bank. The filter bank separates the noisy signal into narrow-band signals. In each band, the envelope of the noisy signal is detected and smoothed. From the smoothed sub-band signal envelope the noise envelope is estimated in each sub-band. The SNR in each band is estimated from the noisy signal and noise envelopes. A non-linear (sigmoid) gain function is used to approximate the Wiener gain. Finally, the original band-limited signal in each band is multiplied by the respective gain and the result is summed to construct the full-band “clean” signal estimate. The overall structure of the system is shown in Fig. 2(a) with a more detailed view of the gain calculation block of a single band shown in Fig. 2(b).

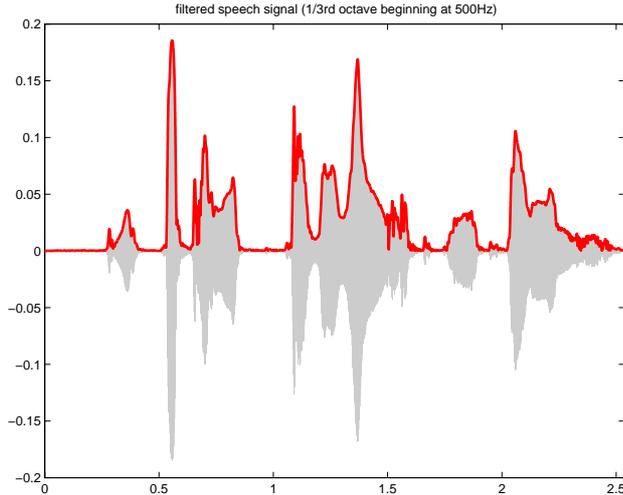


Figure 3: The gray area is $x_k(t) = e_k(t)v_k(t)$. The dark line is $e_k(t)$, the envelope. Since we use $e_k(t)$ as an input of non-linear gain function, we apply averaging to $e_k(t)$ for smoothing gain function.

2.1 Band-pass Filter Bank

The analysis and synthesis filter banks are implemented with Fourier processor IC [13] that has been developed in our previous research. It can easily divide the signal into frequency bands spaced exponentially instead of linearly, as in the typical DFT algorithm. We set the bandwidth of each band-pass filter equal to the bandwidth of the respective critical band in the human ear. This results in approximately one-third octave spacing in the cut-off frequencies of the filter bank. By using one-third octave filters, any frequency distortions—whose bandwidth is on the order of the bandwidth of a band—also lie almost within the same critical band and can be minimized for the perceptual impact [14]. Each continuous-time filter uses a transistor-only version of the autozeroing floating-gate amplifier (AFGA), which is also referred to as a capacitively coupled current conveyer (C^4). Three C^4 's are connected to form a second-order section (C^4 SOS) filter [15]. This filter allows both the low-frequency and high-frequency cutoffs to be controlled electronically by changing the appropriate bias currents.

2.2 Envelope Extraction and Averaging

An audio signal, $x_k(t)$, also can be represented as

$$x_k(t) = e_k(t)v_k(t), \quad (4)$$

where $v_k(t)$ is a band-limited signal with an approximately constant magnitude and $e_k(t)$ represents the envelope variation over time.

Figure 3 shows $e_k(t)$ and $x_k(t)$ for a single band from a sentence spoken by a female. Many different algorithms may be used to estimate $e_k(t)$ from $x_k(t)$ including

- half-wave or full-wave rectification of $x_k(t)$ followed by low-pass filtering

- using the magnitude of the output of Hilbert transform using FIR filters
- low-pass filtering of $x_k^2(t)$

An initial envelope estimate $\hat{e}_k(t)$ is found, for this system, by using a peak detector similar to a simple diode–capacitor AM demodulation circuit. This envelope is smoothed via low-pass filtering. This smoothing helps to make the final noise–reduced signal more free from artifacts like musical noise and missing consonants. The averaged noisy signal envelope estimate in the k -th subband, $\bar{e}_{x,k}(t)$ is described by

$$\alpha \frac{d\bar{e}_{x,k}(t)}{dt} + \bar{e}_{x,k}(t) = \hat{e}_k(t), \quad (5)$$

where α is a time constant for noisy signal envelope estimation, and $\bar{e}_{x,k}(t)$ is the (smoothed) envelope of the noisy signal. Note, if we average $\hat{e}_k(t)$ over too long of a time interval, it might degrade some accuracy of the level estimate of the noisy signal embodied in the envelope because the signal is assumed to be non-stationary. However, if we estimate noise level from this averaged noisy signal estimate, we can expect that averaging might be beneficial to noise level estimate which is assumed to be stationary. Therefore, the estimate of the noise level is obtained by further filtering of $\bar{e}_{x,k}(t)$.

2.3 Noise Level Estimation

The averaged noise envelope estimate, $\bar{e}_{n,k}(t)$, may be obtained in several ways. One way is to use single-pole recursive averaging with a time-varying pole. This method is similar to the soft-decision noise estimation approach used by Sohn and Sung in [16]. This noise estimation method can work well and variations of it are described in [8, 17].

Another effective method of noise estimation is the minimum statistics approach [18]—an approximation to this approach is to use an inverted peak detector or minimum detector with a long time constant on the averaged noisy signal envelope estimate, $\bar{e}_{x,k}(t)$. The inverted peak detector operates on the estimated envelope of the noisy signal, keeping an estimate of the minimum value which is assumed to be the noise floor.

2.4 Non-linear Gain Function

In the discrete-time domain, the Wiener gain is widely used in noise suppression for its superiority in performance. However, many other gain functions are possible (e.g., see [3]) and calculating the Wiener gain can be difficult in the target system. Another factor is that Wiener filtering for noise suppression is usually performed using block processing whereas our system must operate on the continuous signal in real-time. We simulated the Wiener gain and two other nonlinear gain functions to evaluate them for both performance and ease of implementation. The first of the other two non-linear gain functions is a simple bi-linear gain function, H_{bi} , as used in [9, 8].

$$H_{bi}(t) = \min \left[1, \left\{ \frac{\bar{e}_{x,k}(t)}{\gamma \bar{e}_{n,k}(t)} \right\} \right], \quad (6)$$

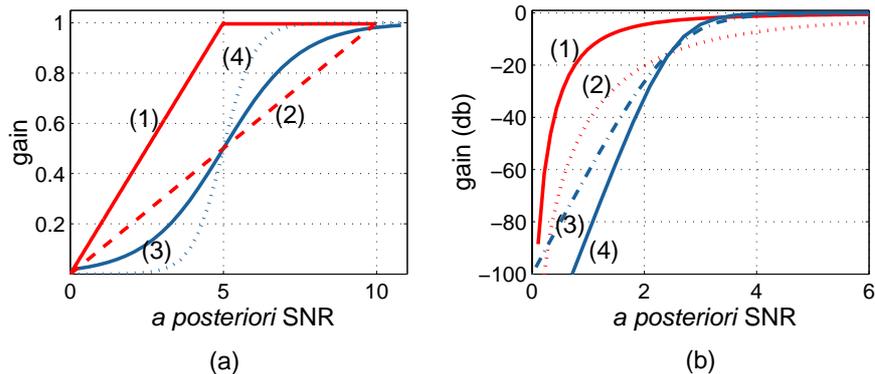


Figure 4: (a) Bi-linear gains and sigmoid gains. (1) bi-linear with $\gamma = 5$, (2) bi-linear with $\gamma = 10$, (3) sigmoid gain with $\phi = 5, \delta = 0.5$, and $\alpha = 0.4$, and (4) sigmoid gain with $\phi = 5, \delta = 0.5$, and $\alpha = 1$. (b) Wiener gains and sigmoid gains. (1) Wiener gain with oversubtraction factor = 1, (2) Wiener gain with oversubtraction factor = 7, (3) sigmoid gain with $\phi = 2.5, \delta = 0.5$, and $\alpha = 1$, and (4) sigmoid gain with $\phi = 2.5, \delta = 0.5$, and $\alpha = 1.4$.

where γ is a threshold that is related to the certainty of speech presence. A disadvantage of the bi-linear gain function is that if we increase γ in order to achieve high noise suppression rate at low SNR, then we also sacrifice audio signal magnitude throughout the region where the SNR equals γ . To minimize this signal reduction, we propose a sigmoid gain function, H_{sig} , which is quite easy to implement with current mirror circuits in the continuous-time domain.

$$H_{sig}(t) = \frac{1}{2} \left\{ \tanh \left(\alpha \left(\frac{\bar{e}_{x,k}(t)}{\bar{e}_{n,k}(t)} - \phi \right) \right) \right\} + \delta, \quad (7)$$

where, α , ϕ , and δ are parameters for slope, horizontal, and vertical shift, respectively.

The sigmoid gain function is superior to bi-linear gain function in that it not only reduces more noise when $\text{SNR} \leq \phi$, but also preserves more signal magnitude when $\phi \leq \text{SNR} \leq \gamma$. Figure 4(a) shows sigmoid gain functions with $\delta = 0.5$, $\phi = \gamma/2 = 5$ and bi-linear gain functions with the maximum noise suppression rate is 20dB ($\gamma = 10$) and 14dB ($\gamma = 5$). Figure 4(b) shows the distinction between Wiener gains with oversubtraction and sigmoid gains. We can see that sigmoid gains can achieve more noise suppression than Wiener gains when $\text{SNR} \approx 1$ (normalized SNR), while preserving as much signal as Wiener gain at higher SNR.

It is often desirable to leave some background noise to minimize the perceptual loss of unvoiced consonants that may be indistinguishable from the added noise. Typical δ should satisfy $\delta > 1/2$ to leave some residual background noise.

2.5 Normalized SNR

Estimating the SNR is simply a noisy process, especially when the added noise has a high variance. Since the noise suppression gain is a function of the SNR, when the estimate of SNR is too high, excessive amounts of noise are present in the output signal. Figure 5 shows

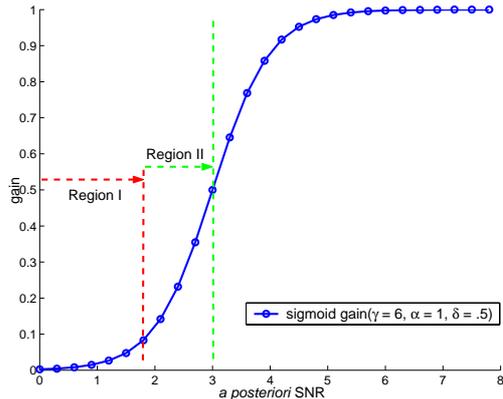


Figure 5: Dynamic range of gain function when the signal is absent. For consistent noise suppression when different input noise variances are applied, SNR, whenever signal is not present, should always be within the Region I. If it goes into the Region II, or, in worst case, above the Region II, the probability of false alarm increases and subsequently it generates artifacts.

two possible dynamic ranges of SNR when the signal is not present. If the estimated SNR lies within Region I, most of the background noise will be suppressed. If the SNR estimate goes into or beyond the Region II, the algorithm will sporadically preserve considerable background noise and generate noticeable artifacts. When the SNR is incorrectly estimated as high (in Region II or above), this can be called a *false alarm*. A false alarm shows that gain function optimized for specific input noise variance could fail to suppress the noise effectively for different noise variances. To avoid or severely reduce the false alarm rate, we should guarantee that the SNR, when the signal is not present, should be normalized to stay within the Region I.

Diethorn [8] estimated the noisy–signal envelope and the noise envelope in parallel from band-limited signal and we experimentally observed this could tend to generate false alarms, as it can be seen from Fig. 6(b). To overcome this problem as shown in Fig. 2, we estimate the noisy–signal envelope from the band-limited signal and estimate noise envelope from the averaged envelope of the noisy signal, instead of performing a parallel estimation from the band-limited signal. By estimating noise envelope from the estimated noisy–signal envelope, we can remove or significantly reduce the envelope separation during the period that signal is not active. This reduced envelope separation yields a better estimate of SNR than methods used by some other researchers and, therefore, provides reliable noise suppression independent of input noise variance.

Another potential problem is introduced by the noise estimation structure used herein, namely, that there is a bias introduced by the use of the minimum detector for estimating noise. (During noise–only periods the mean value of the noise envelope is greater than the estimated noise magnitude which is found from the minimum value.) The sigmoid gain function helps to alleviate this problem by providing a bias in the gain calculation relative to the other gain functions shown. This bias helps to compensate for the bias in our SNR estimate and provides a well matched system.

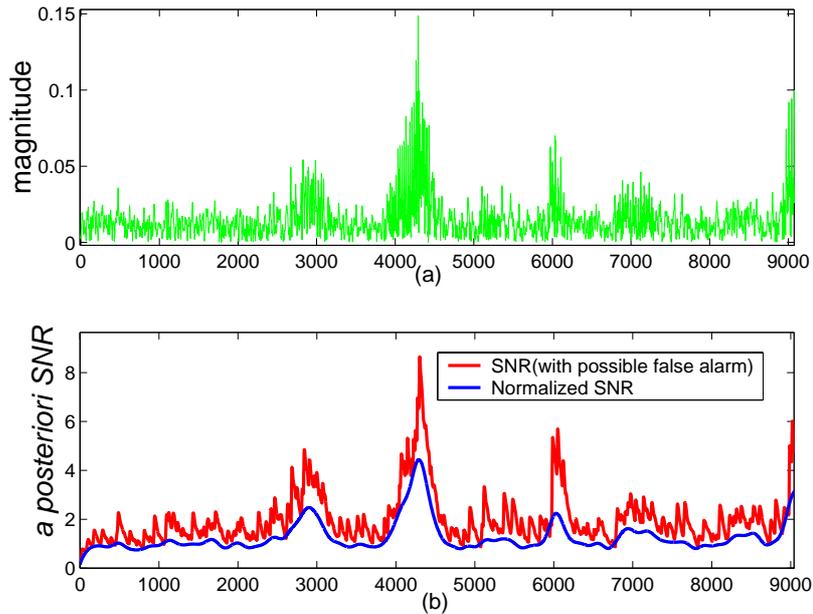


Figure 6: Normalized SNR. (a) Noisy signal envelope of a single band. (b) Normalized SNR (lower line) calculated from proposed method and SNR calculated from [8] (upper line). When the signal is assumed to be not present, normalized SNR have $\text{SNR} \approx 1$.

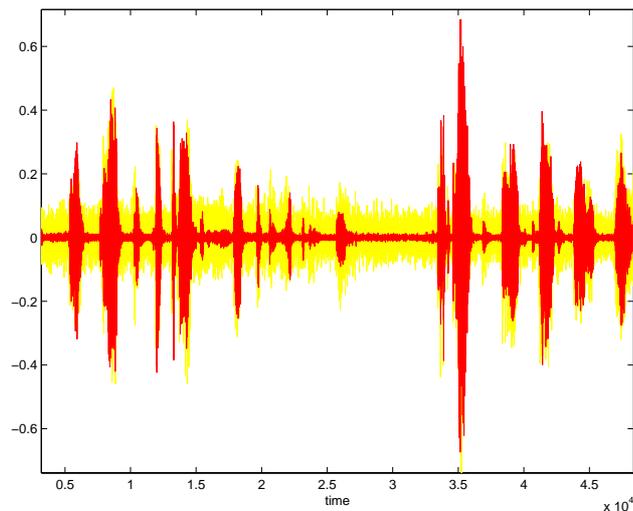


Figure 7: Time-domain waveform of original noisy signal (gray) and noise-suppressed signal (dark) from our functional circuit simulation.

3 Simulation Results

The continuous-time noise suppression algorithm has been functionally simulated in Matlab. The simulated system contained the same functional blocks as shown in Fig. 2 and as implemented in the analog system with continuous-time transfer functions converted to discrete-time functions using the bilinear transform. The sampling rate in the simulator was

chosen to be at least four times the required Nyquist rate to avoid the frequency warping that occurs near the Nyquist rate with the bilinear transform.¹ In order to anticipate accurate performance when implemented with circuits, we tried to include physical limitations as much as possible. For example, the filters were implemented as high Q second-order sections that conformed to previously measured circuit behavior for the C^4 filters.

In Fig. 7, original noisy signal, which was corrupted with background noise, is drawn overlapped with noise suppressed output signal. The perceived quality of the noise-suppressed signal is remarkably free from artifacts normally associated with Wiener filtering and spectral subtraction. We attribute this to the method of frequency decomposition that matches human auditory critical bands, and the proportional bandwidths of the subband envelopes.

From Fig. 7 it is evident that there is a slight amount of background noise remaining. While it is possible to further suppress the noise by changing parameters, we experimentally observed that severe noise suppression can deteriorate the intelligibility of signal by removing noise-like portions of the speech such as /s/ and /f/ sounds.

4 Implementation

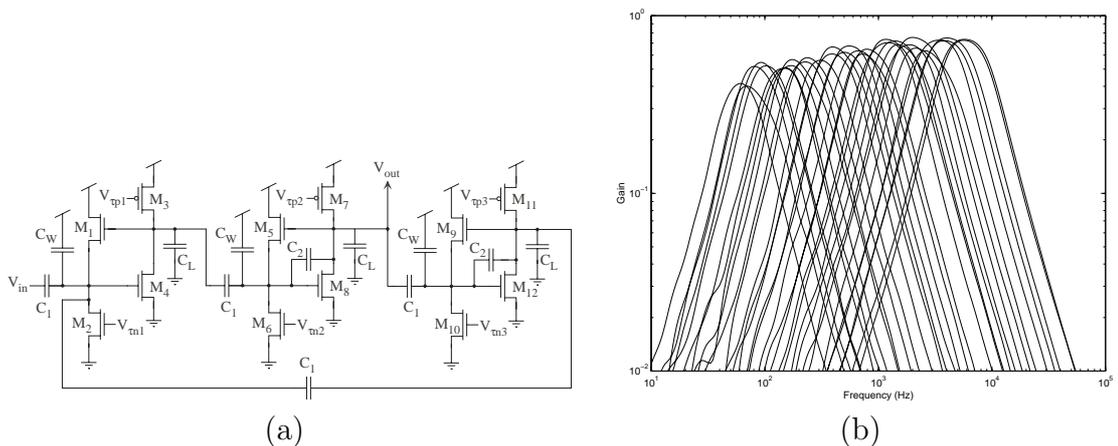


Figure 8: **(a)** The schematic showing the C^4 SOS band-pass filter. The bias voltages are set using a resistive-divider network, which creates a logarithmically spaced filter bank. Future versions will be implemented using floating gate elements instead of resistors to allow for programming and improved tuning. **(b)** This plot shows the frequency response of each filter in a 32-band filter bank. In this configuration, the band-pass rolloff is 40 dB. The noted bias voltages refer to the voltages set at each end of the resistive network. This data is from a circuit fabricated in a 0.5 μm process available through MOSIS.

This section describes the implementation of the noise suppression algorithm on a 0.5 micron CMOS VLSI chip. Note that the process of translating a DSP-based noise suppression algorithm to an analog system is an exercise in “meeting in the middle.” We have already described the algorithm and some of the changes to make it work well with anticipated constraints in the analog system. This section now describes the process of creating the analog

¹Multi-rate signal processing was used to improve efficiency but the high oversampling rule was always followed.

VLSI circuits implement each of the blocks shown in Fig. 2. Note that the entire process was non-trivial and was only possible through effective communication between signal processing and analog circuit engineers.

The analog systems described here are largely based on an analog floating gate circuit technology which allows each circuit to be programmable/tunable. The result is that the individual circuit elements can be very small since they do not require the extensive overhead of building perfectly matched circuits or of designing circuits tuned via external inputs. The input to the system is an analog signal such as the output of a microphone. The system then outputs an analog signal, making it transparent to the signal path except that it provides an enhanced signal.

4.1 Frequency Decomposition

The first structure on the left side of Fig. 2(a) is a filter bank that separates the noisy signal into 32 bands that are logarithmically spaced in frequency, similar to the human auditory system, for frequency domain processing. The filters used in the filter bank are the Capacitively-Coupled Current Conveyor Second-Order Sections (C^4 SOS), shown in Fig. 8(a). The C^4 SOS is composed of three C^4 s (described in [19]) in which the feedback capacitor of the first stage filter is removed in order to make that stage a high gain amplifier. This filter can have a frequency response of any defined bandwidth, and outside that bandwidth, slopes of ± 40 dB/decade or greater occur. By adjusting the voltage biases, the response at either corner can be tuned to have a sharp transition or even a Q peak. A high Q peak is useful because it helps isolate the respective center frequency. More information on the details of the C^4 SOS is available elsewhere [20].

After the incoming noisy signal has been band-limited by the filter bank, a gain factor is calculated based on the characteristics of each band-limited noisy signal.

4.2 Envelope Estimation

The first step in the gain calculation is to estimate the envelope of the noisy signal. Next, the noise envelope is estimated using a technique closely related to the minimum statistics method where the noise is found using a minimum detector on the envelope of the noisy signal.

4.2.1 Peak Detectors

The circuit in Fig. 9(a) is a peak detector which is used to estimate the envelope of the noisy signal. The bias voltage V_{τ_d} sets the time constant at which the output will decay after a peak. When a speech signal is input, the peak detector will follow the envelope of the signal, rising rapidly with the increasing signal amplitude and decaying slowly enough to result in a smooth envelope. The peak detector will also follow the level of the additive noise, particularly in times where the signal is absent. The circuit outputs both a voltage and current that are representative of the noisy-signal level (envelope).

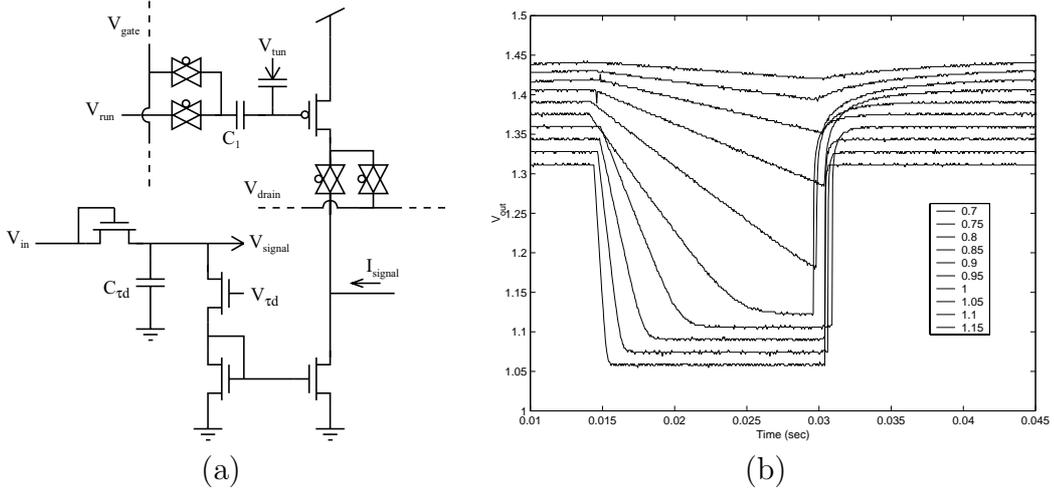


Figure 9: **(a)** The peak detector circuit shown in this schematic tracks the incoming band-limited noisy signal, giving the noisy signal level estimate. There are two outputs in this circuit: the current I_{signal} goes into the division circuit, while the voltage V_{signal} is the input to the minimum detector. The bias voltage V_{τ_d} sets the time constant at which the output will decay after a peak. **(b)** This plot shows the behavior of the peak detector at varying time constant biases with a test input signal of 1 kHz.

4.2.2 Minimum Detector

The circuit in Fig. 10(a) is a minimum detector and is used to estimate the noise level in the signal (i.e. the noise envelope). The input to the minimum detector is the voltage output from the peak detector. In this way, we estimate the noise level by following the minimum values of the noisy-signal envelope. The bias voltage V_{τ_a} sets the attack time constant; it is set to run much slower than the peak detector in order to follow the slow changes found in the amplitude of relatively stationary noise. When the signal is present, the output will maintain a slowly rising level; when the signal is not present, the minimum detector will track the noise level.

In both of these circuits, the floating gate pFET device shown at the top of the schematics is used for offset cancellation. When the input is set to a particular bias voltage, around which the signal will modulate, a certain amount of current will be produced at the output node. The floating gate will be programmed in such a way as to cancel this offset current, causing zero current at the output node until a signal is present. The current outputs of the peak and minimum detectors are the inputs to the next component in the system.

4.3 Translinear Division

In terms of the earlier equation (Eq. 1), an estimation of the *a posteriori* signal-to-noise ratio (SNR) is defined by a ratio of the envelopes of the signals,

$$\widehat{SNR}(t) = \frac{\hat{e}_s(t)}{\hat{e}_n(t)} \approx \frac{e_x(t) - \hat{e}_n(t)}{\hat{e}_n(t)} = \frac{e_x(t)}{\hat{e}_n(t)} - 1 \quad (8)$$

where $\hat{e}_x(t)$ is the noisy signal envelope estimate that can be represented as the sum of the actual signal envelope $e_s(t)$ and the noise envelope estimate $\hat{e}_n(t)$.

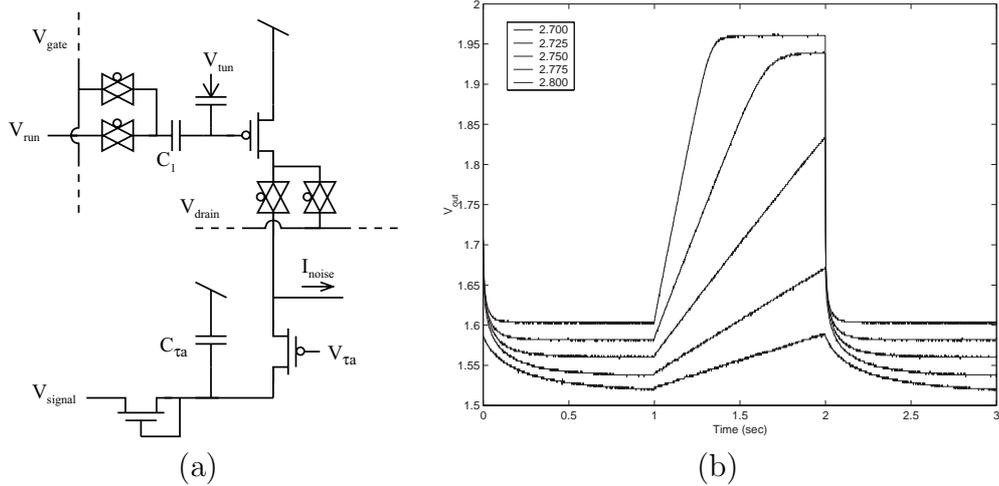


Figure 10: **(a)** The minimum detector is biased with V_{τ_a} to have a much slower time constant than the peak detector, so as to follow the slowly changing curve that results at the bottom of the noisy signal envelope. The output I_{noise} is an estimate of the noise level. **(b)** This plot shows the behavior of the min detector at varying time constant biases with a test input signal of 1 kHz.

Multiplication and division operations can be performed using the translinear principle [21]. The circuit shown in Fig. 11 performs a division operation where the output, I_{SNR} , can be represented by the following equation:

$$I_{SNR} = I_{scale} \frac{I_{signal}}{I_{noise}} - I_{scale}$$

and represents the estimated SNR. The current I_{scale} is set by the bias voltage V_{scale} and is used to put the output current into the proper range for the gain function.

4.4 Gain Function and Multiplication

The final elements of the gain calculation algorithm are the gain function and multiplier. In the circuit implementation, the gain function and the multiplier comprise one circuit, shown in Fig. 12(a). The transistors at the top of the schematic are the differential pair where the actual multiplication takes place; the multiplier will be described shortly. The transistors below the differential pair create the behavior of the gain function. The output of the gain function circuit is I_{gain} , which can be approximated by

$$I_{gain} = \frac{(I_{SNR}^2 + I_{min1})I_{max}}{(I_{SNR}^2 + I_{min1}) + I_{max}}$$

I_{max} and I_{min1} are set by the voltage biases V_{max} and V_{min1} and effectively create the upper and lower bounds of the gain factor output. The voltage bias V_{min2} can be used to further adjust the range of input current that the circuit will accept. The diode-connected transistor on the input branch of the circuit causes the current that is mirrored onto the gain branch to

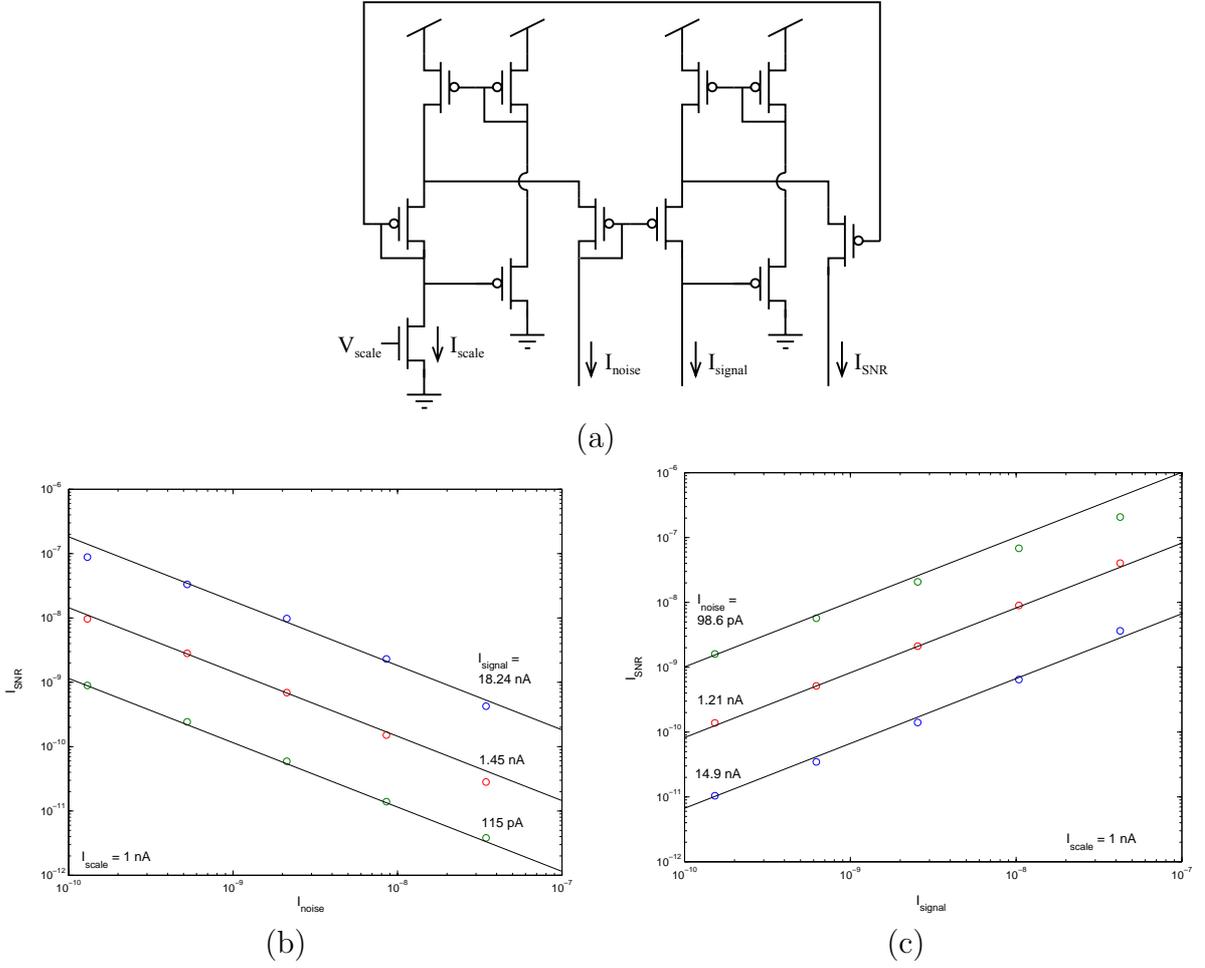


Figure 11: **(a)** The translinear circuit shown in this schematic implements the division of I_{noise} into I_{signal} , yielding a current representing the signal-to-noise ratio, I_{SNR} . We subtract an I_{scale} from the output current, I_{SNR} to mirror the function described in Equation 8. The voltage V_{scale} sets the bias current I_{scale} , which is used to put I_{SNR} into the proper range for the gain function circuit. **(b)** & **(c)** The theoretical and measured outputs for the divider circuit using fixed signal and noise currents respectively.

be approximately squared; this is important to ensure a quick transition from low to high gain as I_{SNR} increases.

The multiplication portion of the circuit results from the interaction between the inputs to the differential pair V_{in1} and V_{in2} and the gain factor current I_{gain} from the gain function circuit in Fig. 12(a). The circuit is biased to operate in the linear range of the following equation:

$$I_{out1} - I_{out2} = I_{gain} \tanh\left(\frac{\kappa(V_{in1} - V_{in2})}{2U_T}\right),$$

where U_T is the thermal voltage of the transistors. The input to the multiplier is the band-limited noisy signal, i.e. the output of the C⁴ SOS band-pass filter. (Because the current implementation is single-ended rather than differential, the “negative” input voltage, V_{in2} , is

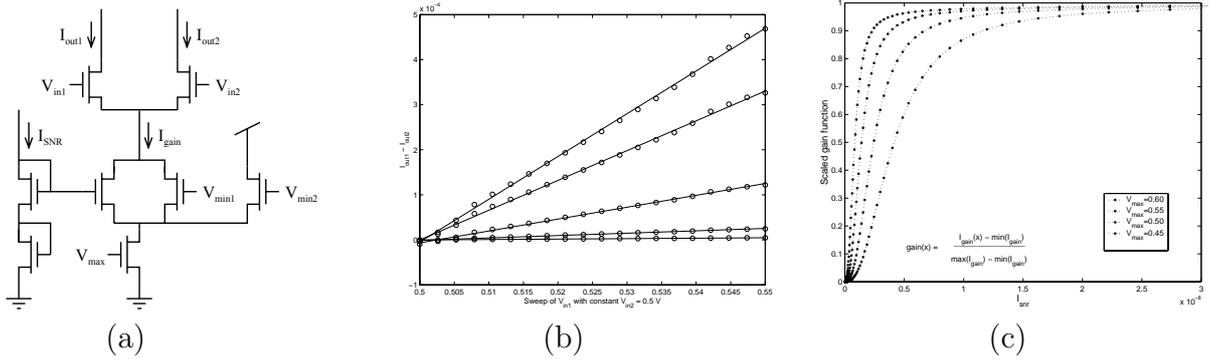


Figure 12: (a) The schematic of the multiplication and gain function circuit. The output current, $I_{out1} - I_{out2}$, is the product of the band-limited input signal, $V_{in1} - V_{in2}$, and the gain factor, I_{gain} . Each voltage bias sets a current that forms the overall behavior of the gain function circuit. (b) The output current, $I_{out1} - I_{out2}$, varies with its two factors: the gain factor and the differential input voltage. This plot shows sweeps of the input voltage for several values of I_{SNR} . (c) The gain function of the circuit depicted in (a) is plotted versus I_{SNR} . At the extremes of the signal-to-noise ratio, either a low gain factor or a unity gain factor is the result. Multiple curves show the bias voltages being adjusted.

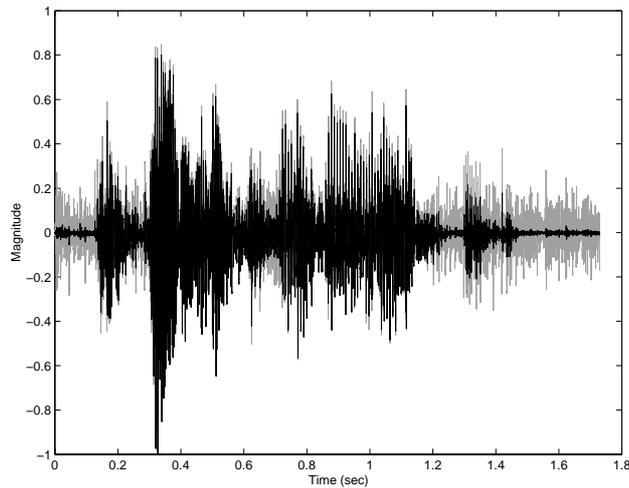


Figure 13: Noise suppression in one sub-band. The light gray data is the subband noisy speech input signal; the black waveform is the corresponding subband output, after the gain function has been applied.

held constant.) The data in Fig. 12(b) and (c) show the functionality of the multiplier and the form of the gain function.

5 Implementation Results

The experimental results presented in this paper are from tests on individual components that have not yet been integrated into a larger system. Figure 13 shows a noisy speech signal that has been processed by the components in our system. The system is effective

at adaptively reducing the amplitude of noise-only portions of the signal while leaving the desired portions relatively intact.

5.1 Circuit Noise and Distortion

Any noise or distortion created by the gain calculation circuits minimally affects the output signal because these circuits are not directly in the signal path. While the band-pass filters and the multipliers will inject a certain amount of noise into each frequency band, this noise will be averaged out by the summation of the signals at the output of the system. Distortion in the signal path will arise from the band-pass filters and the multiplier. In the band-pass filter array, the C^4 SOS structure is not cascaded as in cochlea models, therefore there is no distortion or noise accumulation. The distortion level in each frequency band for a 30 mV single-ended input signal is 2nd harmonic limited at -30 dB at peak. A differential filter-bank will eliminate 2nd harmonic distortion and reduce the 3rd harmonic level to -40 dB at peak. The distortion introduced by the multiplier is dependent on the output levels of the band-pass filters: if the signal is near 30 mV, 3rd harmonic distortion will be -20 dB down; however, if the signal is near 7.5 mV, the 3rd harmonic distortion will approach -46 dB. In speech, particularly in noisy environments, the signal is more evenly distributed across a broad frequency range than a simple tone, therefore distortion is significantly reduced.

6 Conclusions

We have presented a continuous-time audio noise suppression algorithm for implementation in floating-gate analog circuits. The system performs adaptive Wiener filtering on an analog input such as that provided by a microphone, and outputs an analog signal. We believe that this will provide the ability to significantly enhance audio inputs on portable devices without adding significant power drain or size.

For quality and practical issues, we proposed a normalized SNR and non-linear gain function along with band-dependent time constants for estimating the signal and noise respectively. Even though we use simple, low-complexity building blocks such as the low-complexity noise estimator and simple gain function, the simulation results show significant noise suppression while generating few artifacts. Initial tests of the circuit components indicate that a fully integrated analog system will perform as the simulated system. We believe that the choice of filters, noise estimator, and gain function yield a robust system for a variety of noise levels and conditions. Also, we demonstrated circuit implementations of each of the algorithmic components.

References

- [1] David V. Anderson and Mark A. Clements, "Audio signal noise reduction using multi-resolution sinusoidal modeling," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 1999, pp. 805–8.
- [2] Dionysis E. Tsoukalas, John N. Mourjopoulos, and George Kokkinakis, "Speech enhancement based on audible noise suppression," *IEEE Transactions on Speech and Audio Processing*, vol. 5, no. 6, pp. 497–514, Nov. 1997.

- [3] Levent Arslan, Alan McCree, and Vishu Viswanathan, “New methods for adaptive noise suppression,” in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 1995, vol. 1, pp. 812–815.
- [4] Olivier Cappé, “Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor,” *IEEE Transactions on Speech and Audio Processing*, vol. 2, no. 2, pp. 345–349, Apr. 1994.
- [5] Sharon Gannot, David Burshtein, and Ehud Weinstein, “Iterative and sequential Kalman filter-based speech enhancement algorithms,” *IEEE Transactions on Speech and Audio Processing*, vol. 6, no. 4, pp. 373–385, July 1998.
- [6] Yariv Ephraim and Harry L. Van Trees, “A spectrally-based signal subspace approach for speech enhancement,” in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1995.
- [7] Peter Händel, “Low-distortion spectral subtraction for speech enhancement,” in *4th European Conference on Speech Communication and Technology*, Madrid, Spain, September 1995, vol. 2, pp. 1549–1552.
- [8] Steven L. Gay and Jacob Benesty, *Acoustic Signal Processing for Telecommunication*, Kluwer Academic Publishers, 2000.
- [9] Eric J. Diethorn, “A subband noise-reduction method for enhancing speech in telephony & teleconferencing,” in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, Mohonk Mountain House, New Paltz, NY, Oct. 1997.
- [10] P. Hasler and T. S. Lande, “Overview of floating-gate devices, circuits, and systems,” *IEEE Transactions on Circuits and Systems II*, vol. 48, no. 1, pp. 1–3, Jan. 2001.
- [11] Paul Hasler and David V. Anderson, “Cooperative analog-digital signal processing,” in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Orlando, FL, May 2002, vol. IV, pp. 3972–5.
- [12] David V. Anderson and Paul Hasler, “Cooperative analog/digital signal processing,” in *World Conference on Systemics, Cybernetics, and Informatics*, Orlando, FL, July 2001.
- [13] Matt Kucic, AiChen Low, Paul Hasler, and Joe Neff, “A programmable continuous-time floating-gate fourier processor,” *IEEE Transactions on Circuits and Systems II*, vol. 48, no. 1, pp. 90–99, Jan. 2001.
- [14] David V. Anderson, “Model based development of a hearing aid,” M.S. thesis, Brigham Young University, Provo, Utah, 1994.
- [15] Rich Ellis, Heejong Yoo, David Graham, Paul Hasler, and David Anderson, “A continuous-time speech enhancement front-end for microphone inputs,” in *Proceedings of the IEEE International Symposium on Circuits and Systems*, Phoenix, AZ, 2002, vol. II, pp. 728–31.
- [16] Jongseo Sohn and Wonyong Sung, “A voice activity detector employing soft decision based noise spectrum adaptation,” in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 1998.
- [17] Heejong Yoo, David V. Anderson, and Paul Hasler, “Continuous-time audio noise suppression and real-time implementation,” in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Orlando, FL, May 2002, vol. IV, pp. 3980–3.
- [18] Ranier Martin, “Noise power spectral density estimation based on optimal smoothing and minimum statistics,” *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 5, pp. 504–512, July 2001.
- [19] Paul Hasler, Matt Kucic, and Bradley A. Minch, “A transistor-only circuit model of the autozeroing floating-gate amplifier,” in *Midwest Conference on Circuits and Systems*, Las Cruces, NM, 1999.
- [20] David Graham and Paul Hasler, “Capacitively-coupled current conveyer second-order section for continuous-time bandpass filtering and cochlea modeling,” in *Proceedings of the International Symposium on Circuits and Systems*, 2002, vol. to appear.
- [21] T. Serrano-Gotarredona, B. Linares-Barranco, and A. G. Andreou, “A general translinear principle for subthreshold MOS transistors,” *IEEE Transactions on Circuits and Systems*, vol. 46, no. 5, pp. 607–616, May 1999.

Acknowledgements

The authors wish to thank Dr. Bradley A. Minch for several useful circuit discussions.