UCLA
UCLA Electronic Theses and Dissertations

Title

A 0.6uW/Channel, Frequency Division Multiplexed Amplifier for Neural Recording Systems

Permalink

Author

Chandrakumar, Hariprasad

Publication Date

2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

A 0.6 μW/Channel, Frequency Division Multiplexed

Amplifier for Neural Recording Systems

A thesis submitted in partial satisfaction

of the requirements for the degree

Master of Science in Electrical Engineering

by

Hariprasad Chandrakumar

2012

ABSTRACT OF THE THESIS

A 0.6 µW/Channel, Frequency Division Multiplexed Amplifier for Neural Recording Systems

by

Hariprasad Chandrakumar

Master of Science in Electrical Engineering,

University of California, Los Angeles, 2012

Professor Dejan Markovic, Chair

Neural signal recording systems are vital for understanding the working of the brain. With the growing number of recording channels, it is imperative that their power consumption be kept as low as possible. The front-end amplifier used in these recording systems consumes a significant portion of the system power. This work presents a fundamentally different approach in the design of amplifiers by using concepts from signal processing to improve the achievable performance. Frequency division multiplexing is used to share the most power-hungry blocks of the amplifier, while maintaining all performance characteristics. Simulation results using a 65-nm CMOS technology show that the power consumed is lower than the theoretical limit for a single-channel amplifier by a factor of 2. The input-referred noise over a bandwidth of 0.2 Hz to 6 kHz is 4 μV_{rms}, while burning 500 nA per channel from a 1.2 V supply. The power consumed is a factor of 3.6 lower than the best designs published so far. This shows that the design presented in this work, when fabricated, could be the best design to date for a neural amplifier

The thesis of Hariprasad Chandrakumar is approved.

_____

Sudhakar Pamarti

_____

Asad Abidi

_____

Dejan Markovic, Committee Chair

University of California, Los Angeles

2009

To my parents, Lalitha and Madhavan Chandrakumar

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

CHAPTER 1

Introduction

In the last few years, there has been great interest in the neuroscience community in decoding the functioning of the brain. Among the various methods used, analysis of the recordings of the electrical activity of neurons has been among the most important tools available [1]. These recordings are indispensable for understanding and diagnosing neurological disorders like epileptic seizures, in the creation of brain-machine interfaces and for neuro-prosthetic technologies to aid paralyzed patients [2], [3]. Current solutions available for neural recordings are primarily wired or tethered setups (Fig 1.1), where the patient has wires that carry the signals from the implanted electrodes to the electronics outside that process these signals.
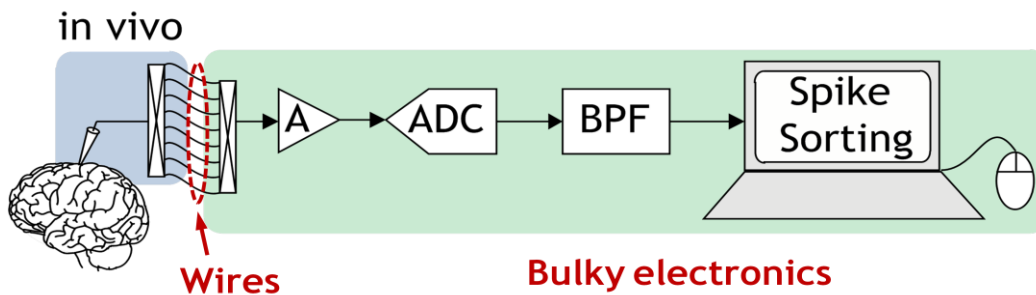
Fig 1.1: Conventional neural recording setup

This setup presents a significant risk of infection, since there is a break in the skin for the wires to come out. Also, the mobility of the patient is severely limited, which may present problems in studies that need the patient to have normal mobility. This is also a major issue in animal studies that need the test animals to be socially active, freely moving about in an enriched environment with other animals. During these studies, if the neural recording from specific regions of the brain is available in real time along with the animal's physical behavior, there could be

1

significant advances made in understanding the working of the brain. Thus a wireless neural recording system is needed, that can be fully implantable inside the animal, and that can transmit in real-time the recordings to an external terminal. This removes the need for wires, thereby eliminating the risk of infection and lets the animal roam freely.

1.1 Signals of Interest

The neural signals of interest that are recorded by the electrodes usually fall within the 1 Hz to 6 kHz bandwidth. Depending on the type of neural signal, their peak amplitudes can vary from 100 µV to 1 mV. The various types of biological signals, along with their bandwidth and amplitudes are shown in Fig 1.2. For the purpose of this work, the Local Field Potential (LFP) and action potential signals are required to be processed, which cover the complete range of frequencies. The LFP band is from 0.1 Hz to 300 Hz [4], and the action potential band is from 300 Hz to 6 kHz [5]. The peak amplitude of LFP signals is 1 mV, and the peak amplitude of action potentials is 100 µV.
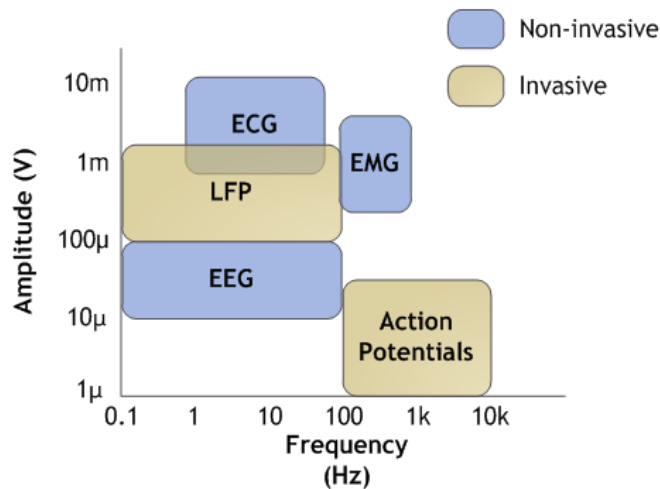
Fig 1.2: Various biological signals, their amplitudes and frequency range

2

1.2   Electrode Characteristics

The electrode used to record neural activity behaves like a capacitively coupled voltage source. Fig 1.3 shows an approximate electrical model for the electrode [6]. R_e and C_e model the metal-electrolyte surface coupling. R_m is the series resistance of the metal interconnect, and C_p is the stray capacitance to the substrate. The dominant components are R_e and C_e, and their values vary widely depending on the electrode's intended bandwidth, manufacturer etc.
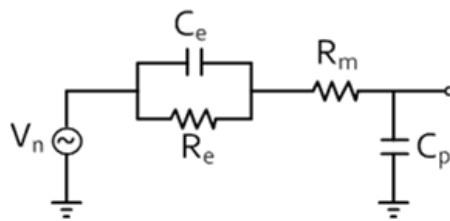
Fig 1.3: Electrical model of a typical neural recording electrode

To get reasonable estimates of the electrode impedance, two types of electrodes (surface electrodes and depth electrodes) that would be used in the actual system were characterized for their impedance. Depth electrodes are very thin and are made to record individual cell activity. Their capacitance (C_e) is about 10 nF and resistance (R_e) is about 200 MΩ. Surface electrodes are much bigger, and they are designed to record the activity of groups of neurons. Their capacitance is about 3µF, and their resistance is about 10 kΩ.

1.3   Noise and Linearity Requirements

From published work [7]-[8], and by interacting with the Neuroscience Department at UCLA, it was determined that an SNR of 7 bits is enough to accurately capture the required information in the neural recording. From previous work done by Professor Markovic's group at UCLA, it was

determined that spike-sorting algorithms that differentiate spikes (300 Hz – 6 kHz) from individual neurons (from a composite recording) are significantly immune to nonlinearity. The SNDR can be as low as 4 bits. The LFP band however, requires an accuracy of about 7-8 bits. This implies that for a 1 mV$_{\text{P-P}}$ signal swing at the electrode, the input-referred noise of the recording system should be lower than 4-5 µV$_{\text{rms}}$. To ensure an ENOB of 7 bits, the total harmonic distortion at the input of the ADC should be lower than 0.5%.

1.4 Power Consumption

The power density of the recording system should be much lower than 800 µW/mm^2. This is the limit at which neurons undergo necrosis and cell death [9]. Thus if there are 4 channels per mm^2, the per-channel power limit is 200 µW. To be safely away below this power limit, the maximum power is kept an order of magnitude lower than the absolute maximum. This brings the power requirement to less than 20 µW per channel. Lower power consumption is also necessary for longer battery life, which would prevent the need for frequent surgeries for battery replacement in an implanted system. Lower power consumption could also lead to alternatives to a conventional battery. Wireless inductive power transfer or thermal energy harvesting, which otherwise may be impractical, are being considered.

1.5 Input Impedance

Since the electrode is a non-ideal voltage source, the load presented to the electrode has to be such that there is no significant attenuation of the recorded signal. For the depth electrode, the electrode impedance is approximately a 10nF capacitor. Thus to ensure that the electrode is not

loaded significantly, and assuming that the input impedance of the amplifier is nearly constant in the signal band, the input impedance of the amplifier should be larger than 16 MΩ to ensure a corner of 1 Hz. For surface electrodes, the electrode impedance is approximately a 3 µF capacitor. This relaxes the input impedance requirement to a minimum of 50 kΩ for a 1 Hz corner.

Table 1.1: Summary of requirements for the recording system

Specification	Requirement
Power	< 20 µW per channel
Input-referred noise	< 5 μV_{rms}
Input Impedance	> 16 MΩ at 1 Hz
Bandwidth	1 Hz to 6 kHz
Peak input swing	1 mV
SNR	> 7 bits
THD	0.5%
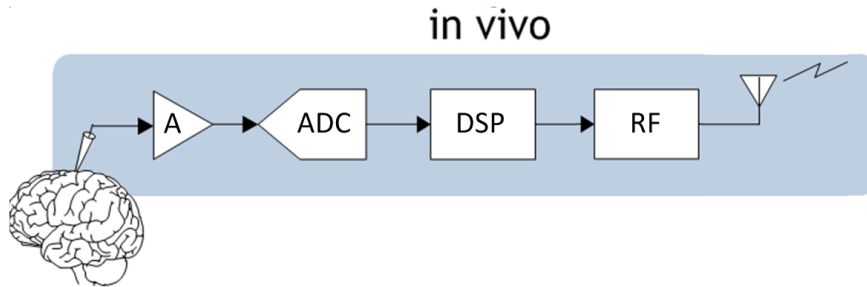
Fig 1.4: Block diagram of a neural recording system

A summary of the various requirements of a neural recording system is shown in Table 1.1. A typical neural recording system consists of a front-end amplifier, an ADC, a DSP unit and a transmitter. Front-end amplifiers are used in most sensor-based systems to amplify the signal to reasonable amplitudes before they are digitized. The larger swing available at the ADC input helps in reducing the required precision (size of LSB) of the ADC. Since the peak input signal swing is 1 mV, and the supply voltage is about 1 V, the output signal swing of the amplifier will

be 1 V with a gain of 60 dB. Anti-alias filters are also employed before the ADC to prevent out-of-band signals and noise from folding back in-band.

The DSP is used to provide additional filtering and data compression. For raw data transmission, the power of the system is dominated by the transmitter. But data compression reduces the output data rate, thus reducing the power dissipated in the transmitter. In [10], the DSP achieves a data reduction by a factor greater than 200, which can greatly reduce the transmitter power. Thus if the DSP is used for data compression, a large portion of the power consumed by the system is burnt in the amplifier because the front-end amplifier, being the first stage in the signal chain sets the noise figure of the recording system. Thus reducing the power consumed by the amplifier will lead to a significant reduction in the overall system power. This work addresses the problem of lowering the power consumption of the amplifier for the required set of performance parameters.

1.6   Thesis Outline

Current approaches to amplifier design are discussed in Chapter 2 along with their limitations. The idea of utilizing unused bandwidth of the 1st stage amplifier leads to the possibility of a multi-channel amplifier. Time Division Multiplexing and Frequency Division Multiplexing are two possible ways of sharing the fast amplifier. In Chapter 3, Time Division Multiplexing and its effect on output noise is discussed. In Chapter 4, Frequency Division Multiplexing is discussed as a possible solution, along with its noise analysis, implementation and overheads. Chapter 5 presents the simulated results, comparisons with other state-of-the-art designs and conclusions of the thesis.

CHAPTER 2

Prior Art

In the designs presented in [11]-[12], the capacitively coupled inverting amplifier is used, as shown in Fig. 2.1(a). The ratio of capacitors C_1/C_2 sets the closed loop gain, and the amplifier bandwidth along with the feedback network sets the closed loop bandwidth. If C_1, which is the input capacitance seen by the electrode can be kept much smaller than the electrode capacitance, then loading is minimized. The power (current) consumed by the amplifier is governed by the required noise figure and the bandwidth of the amplifier. In the application of neural recording, the bandwidth of the signals is very small, hence the minimum bias current requirement is set by the required noise performance of the amplifier. This results in excess bandwidth, which is reduced by increasing the size of the load capacitors, or Miller capacitors in a 2-stage amplifier. This prevents the out-of-band noise from folding in-band during Nyquist sampling.

The transconductor used in Fig 2.1(a) is realized using a differential amplifier. To minimize current consumption for a given noise requirement, the input transistors are biased in weak inversion to achieve the maximum transconductance for a given current. The transistor sizes are chosen to minimize noise contribution from other transistors, and the transistors are kept large to reduce flicker noise. The telescopic amplifier and the folded-cascode topologies are used, since they give large output impedance, and they can be designed such that the input transistors are the only major noise contributors. In [13], the chopping technique [14] is used along with the above design strategy to up-modulate the signal before it is processed by the amplifier. Subsequent down-modulation results in the signal being demodulated back to its original frequency, and the flicker noise is up-modulated to the clock frequency. The signal flow is shown in Fig 2.1(b).
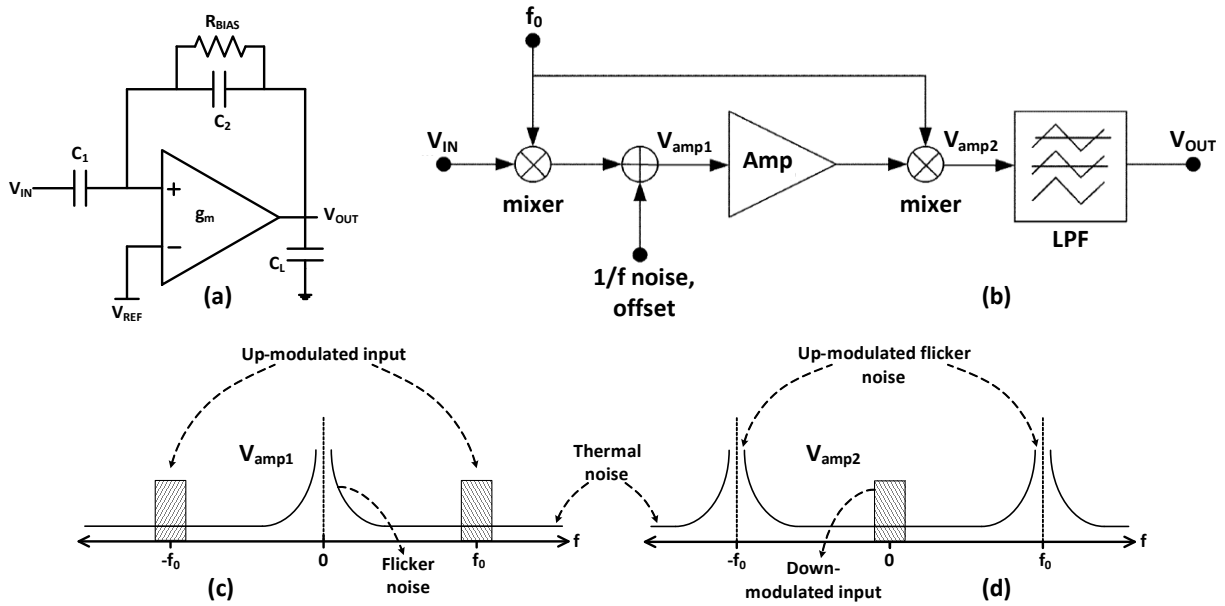
Fig 2.1 (a): Capacitively coupled inverting amplifier, (b) Signal flow diagram for Chopping, (c) Signal spectrum at amp input, (d) Signal spectrum at LPF input

Thus, the signal and flicker noise are kept separate in frequency as shown in Fig 2.1(c)-(d), and the flicker noise can be filtered out by an additional low-pass filter. In practice, it is easier to multiply a signal with a bipolar square wave (±1) than to multiply with a sinusoid, because multiplying with a square wave involves changing the sign of the signal, which is the same as swapping the paths of a differential signal. This can be implemented by 4 cross-coupled switches as shown in Fig. 2.2. Other advantages of using a square wave are that it is easier to generate, and there is no SNR degradation if there is no filtering between modulation and demodulation. In [12], the down-conversion is performed between the 1st and 2nd amplifier stages. This enables the modulation frequency to be significantly larger than the amplifier bandwidth, because before the 2nd amplifier stage the signal path is still fast, since the dominant pole has not yet arrived. This implementation also allows for the demodulation operation to extract signal energy from multiple harmonics, thus incurring minimal SNR degradation.

8
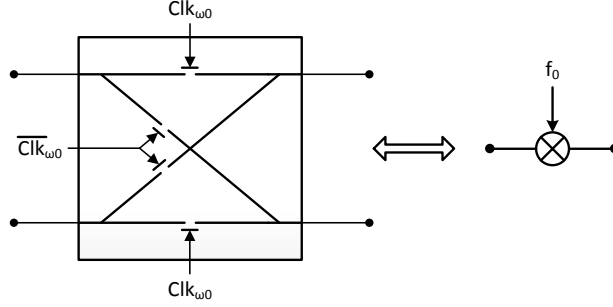
Fig 2.2: Implementing mixers using switches

In [15], a push-pull topology is used for the 1st stage g_m, which approximately doubles the trans-conductance of the amplifier, thus lowering the input-referred noise power by a factor of 2. In [16], one half of the differential amplifier is shared among multiple channels, thus reducing the effective power burnt per channel. But the reduction in power is not significant, and flicker noise limits the performance of the design. In [17], a second differential signal is processed by the amplifier, where the differential signal modulates the common mode voltages of 2 half-amplifiers. This design is also limited by flicker noise (for neural recording applications), and the number of channels is limited due to limited headroom. In [18], a digitally intensive mixed-signal design approach was used where the goal was to reduce area and supply voltage.

Table 2.1 summarizes the performance of existing designs. Among all designs published, these seem to be the designs that offer the best tradeoffs between power and noise. This tradeoff is captured by the Noise Efficiency Factor (NEF), which is defined as

$$NEF = v_{in,rms}\sqrt{\frac{2I_{tot}}{\pi U_T 4kT(BW)}}$$ (2.1)

where $v_{in,rms}$ is the input-referred noise voltage, I_{tot} is the total current consumed by the amplifier, BW is the bandwidth of the amplifier in Hz, and U_T is the thermal voltage. The NEF is widely

9

used to compare designs. The NEF represents the factor by which noise (measured in the rms sense) is greater as compared to an ideal system burning the same current and having the same bandwidth. Though originally introduced for BJT designs [22], it is still widely used to compare amplifiers. For a differential amplifier, the theoretical lower limit of the NEF is 2 [11]. As seen from Table 2.1, all designs except [13] are close to the theoretical limit.

Table 2.1: Summary of performance of current designs

Specification	[11]	[12]	[13]	[15]
Supply	2.8 V	3.0 V	1.8 – 3.3 V	1 V
Current	2.7 µA	5.2 µA	1.0 µA	12.1 µA
Gain	40.85 dB	71.5 dB	41 dB, 50.5 dB	40 dB
Bandwidth	45 Hz – 5.32 kHz	29 Hz – 22 kHz	0.05 Hz – 180Hz	10.5 kHz
Input ref. Noise	3.06 µV$_{rms}$	4 µV$_{rms}$	0.95 µV$_{rms}$	2 µV$_{rms}$
NEF	2.67	2.45	4.6/5.4	2.9

In all the aforementioned designs, the amplifier is designed to process a single channel, for which the minimum value of the NEF is 2. But it was seen that the 1st g$_m$ stage has a large bandwidth, which enables the chopping technique in [13] to use a chopping frequency much higher than the closed-loop bandwidth. If this large bandwidth could be used to process more channels, the overall system power could be reduced, since in a low-noise design the 1st stage g$_m$ burns most of the power to minimize the input-referred noise. In an application like neural recording, the number of channels that need to be recorded are around 32 to 64. Thus, this added degree of freedom of multiple channels can be used to realize a system burning lower power per channel by sharing the 1st g$_m$ stage. The most intuitive way of solving this problem is to time-share the fast amplifier among different channels, thus using the power of one amplifier but processing multiple channels. But the analysis in the next section reveals that the noise performance of this technique degrades with increasing number of channels.

CHAPTER 3

Time Division Multiplexing
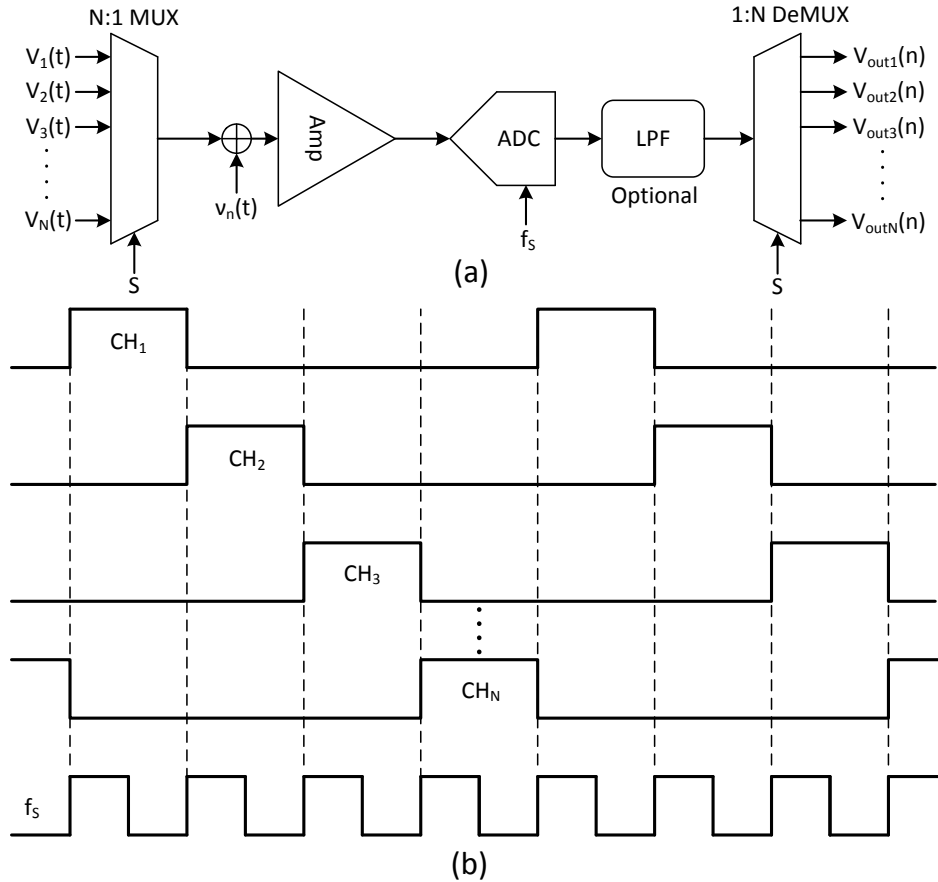
Fig 3.1: Block diagram (a) and clocking (b) for a time division multiplexing system

Since the amplifier bandwidth is larger than needed, one way of using this excess bandwidth is to use Time Division Multiplexing (TDM) to share the amplifier across multiple (N) channels. But the effect of muxing on SNR needs to be determined. The block-level diagram of the system is shown in Fig 3.1(a), and the clocking is shown in Fig 3.1(b). Each channel is sampled at f_S/N, where f_S is the sampling rate of the ADC. The input-referred noise spectral density of the amplifier is v_n and for 6τ settling, the amplifier bandwidth is also f_S. To find the SNR at the output, the output signal and noise powers are calculated.
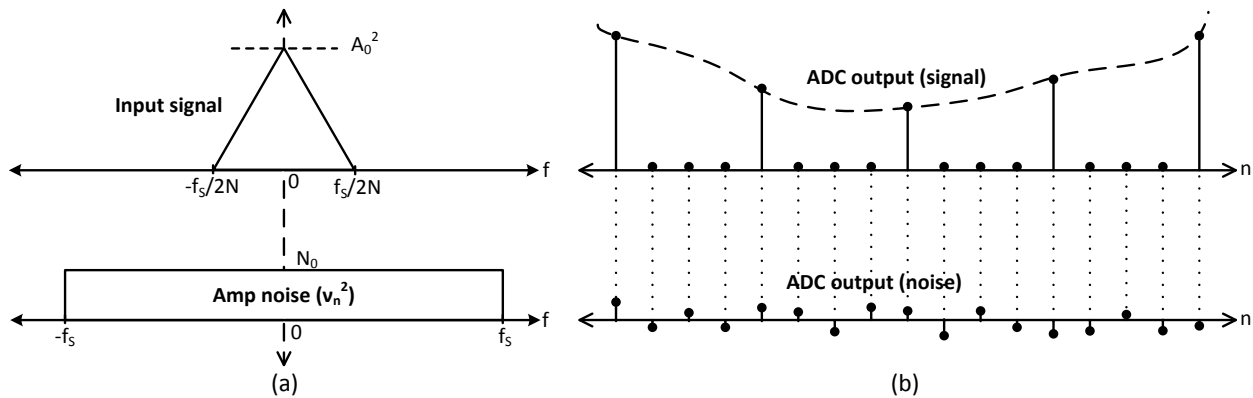
Fig 3.2 (a) Input signal and amp noise PSD, (b) Time domain ADC output

The significance of the anti-alias filter (labeled as LPF in Fig 3.1) will be revealed by the analysis. Assume that only channel 1 is operational, and all other inputs are grounded. The input signal and the amplifier noise spectral density are shown in Fig 3.2(a). The input SNR is A_0^2/N_0. Thus in the time domain, the ADC output will be the sum of the amplified signal and noise, as shown in Fig 3.2(b). To calculate the output spectrum of the signal and noise, the signal at the output of the ADC can be synthesized from the original analog signal by sampling it at Nyquist rate, and then up-sampling by N. The signal flow and the resulting output signal spectrum are shown in Fig 3.3. The noise spectrum at the output of the ADC can be obtained by sampling the amplifier noise at f_S. Since the noise bandwidth is f_S, the sampled noise floor would have increased by 2 (aliasing) because the Nyquist sampling frequency for the amplifier output noise is $2f_S$. But the actual floor would have increased by more than a factor of 2. This is because, since the amplifier noise is not brick-wall filtered, the rolling-off tail of the noise will get aliased in-band. Assuming a brick-wall filtered noise, the resulting noise spectrum is shown in Fig 3.3. Thus, at the ADC output, the signal has been replicated and there are N images, positioned at frequencies of $2\pi/N$ and its integer multiples.
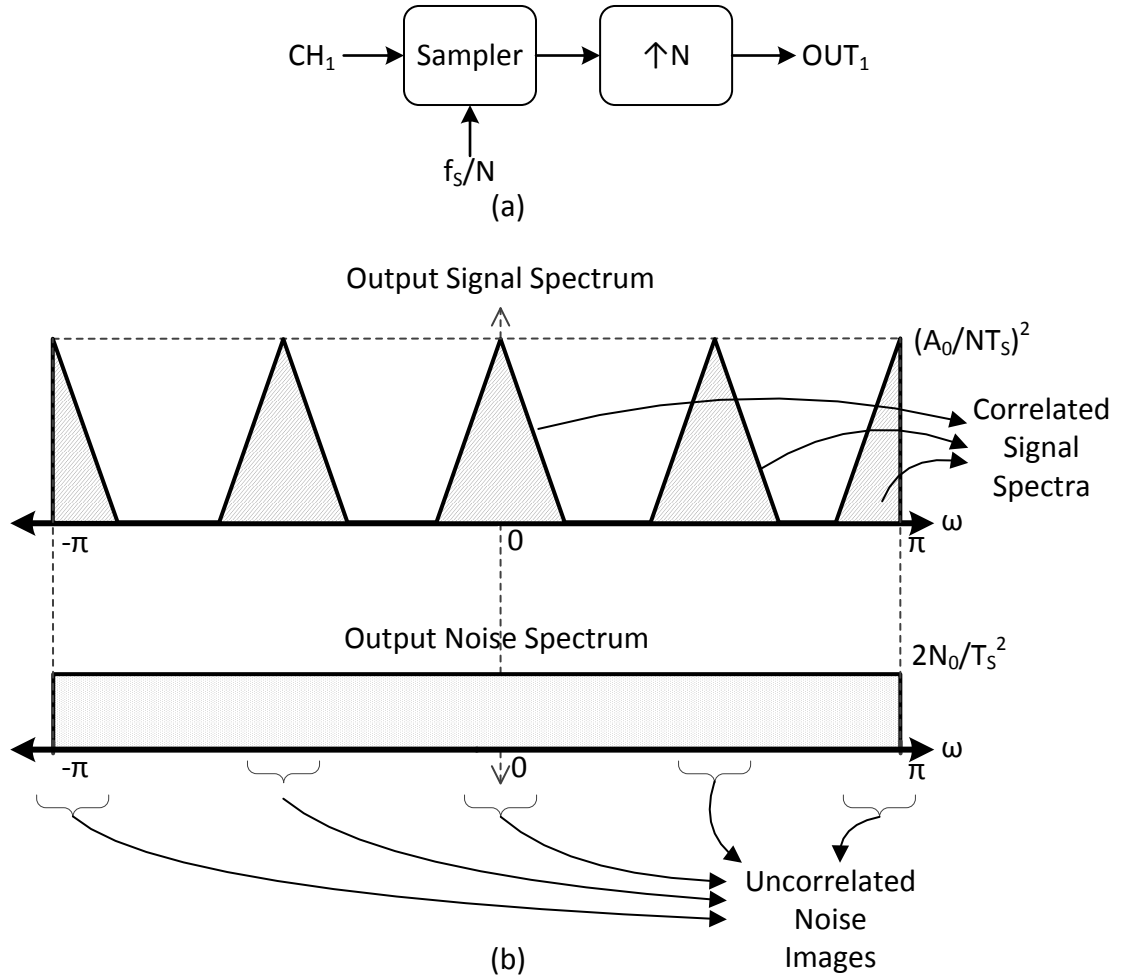
Fig 3.3 (a) Synthesizing the ADC output signal, (b) ADC output signal and noise spectra

If the baseband image is filtered out, then the resulting SNR is

$$SNR_{BB} = \left(\frac{A_0^2}{2N_0} \right) \frac{1}{N^2}$$
(3.1)

Thus as the channel count increases, the baseband SNR degrades by N^2. Hence, if the output of the ADC is low-pass filtered and then decimated, the resulting SNR degrades by N^2. This is because muxing the input causes the signal spectrum to spread, and copies are created at frequencies of $\omega = 2\pi/N$ and its integer multiples. So even though the original signal was narrow-

band, after muxing the signal is truly wideband. Thus the anti-alias filter removes signal content, reducing the signal amplitude by N. The noise PSD remains unchanged, except for the filtering effect of the filter. If the output of the ADC is decimated without using an anti-alias filter, the following signal and noise spectra are obtained, as shown in Fig 3.4.

Without anti-alias filter

With anti-alias filter

Output (Signal)

$(A_0/NT_S)^2$

$(A_0/NT_S)^2(1/N^2)$

$-\pi$ 0 π

$-\pi$ 0 π

Output (Noise)

$(2N_0/T_S^2)(1/N)$

$(2N_0/T_S^2)(1/N^2)$

$-\pi$ 0 π

$-\pi$ 0 π
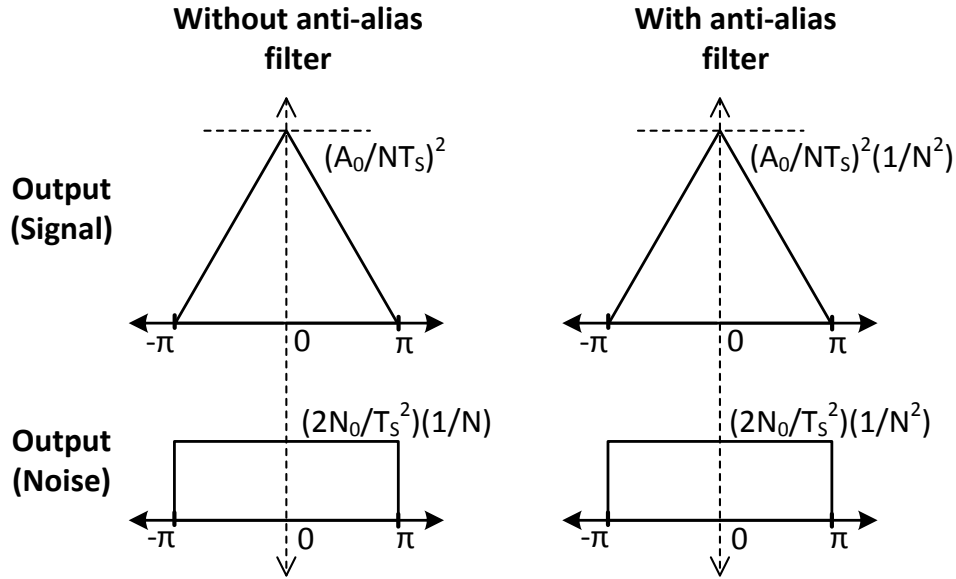
Fig 3.4: Output signal and noise PSD

Since the signal is first up-sampled by N and then down-sampled by N, the resulting spectrum is the same as the signal sampled at Nyquist rate. To obtain the noise spectrum, the amplifier noise can be directly sampled at f_S/N. This would result in noise folding and raise the in-band noise floor by a factor of 2N. The resulting SNR after decimation without using an anti-alias filter is

$$SNR_{OUT} = \left(\frac{A_0^2}{2N_0} \right) \frac{1}{N} \tag{3.2}$$

From the above analysis, it can be seen that with an anti-alias filter, the SNR is actually worse than not using the filter. While decimating without the anti-alias filter, the noise floor gets raised

by N. But the various (N) correlated signal images that were created by up-sampling are brought back down to baseband, and they add in phase. Thus the signal amplitude is restored to A_0/NT_S (Nyquist sampling). The resulting SNR is proportional to $1/N$ instead of $1/N^2$.

The fundamental limitation of this technique is the different bandwidths of signal and noise, and the different frequencies at which the signal and noise are getting sampled at. For TDM, the signal will always be sampled at a slower rate than the noise, and the amplifier bandwidth will have to be higher by a factor of N for processing N channels. Thus the noise power that comes with every signal sample is the integrated noise power over the entire amplifier bandwidth, which is N times larger than a single channel's bandwidth. Thus the best case SNR tradeoff is **SNR α 1/N**, which means that the SNR degrades as channel count increases for the same power. This is the same tradeoff as power vs noise in a single channel amplifier (I_{Bias} α $1/v_n^2$).

The above analysis was simulated to confirm its validity. The simulated results for 2 and 4 channels are shown in Fig 3.5 and Fig 3.6, and they agree with the analysis. The amplifier gain is 1, and the amplifier noise is modeled by a noisy resistor of 1 kΩ (noise PSD = 4 nV/√Hz) at the amplifier input. The signal bandwidth is 10 kHz. When the anti-alias filter is not used, the low-frequency gain is 1. The gain falls to 0.475 (N = 2) and 0.241 (N = 4) with the anti-alias filter, and the expected gain was 0.5 and 0.25 respectively. When the anti-alias filter is used, the expected noise PSD is √2 x 4 nV/√Hz = 5.65 nV/√Hz, since the ADC is sampling at half the Nyquist frequency of the amplifier noise. But the amplifier is not a brick-wall filter, hence the tail of the amplifier noise which was ignored in the above analysis is seen here as an increased PSD of 6.9 nV/√Hz (N = 2) and 6.5 nV/√Hz (N = 4). The tail can be accounted for by scaling up the noise PSD by $\pi/2$, since the equivalent noise bandwidth of a 1st-order RC circuit is higher than the signal bandwidth by $\pi/2$.
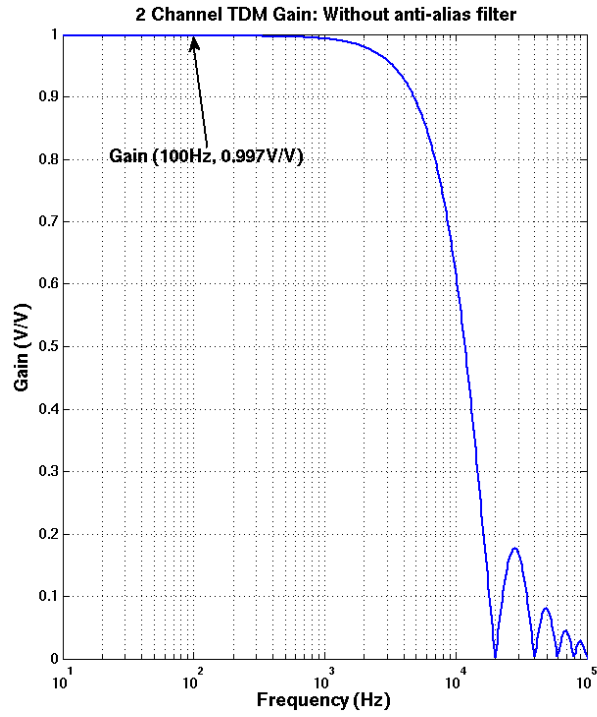
(a)
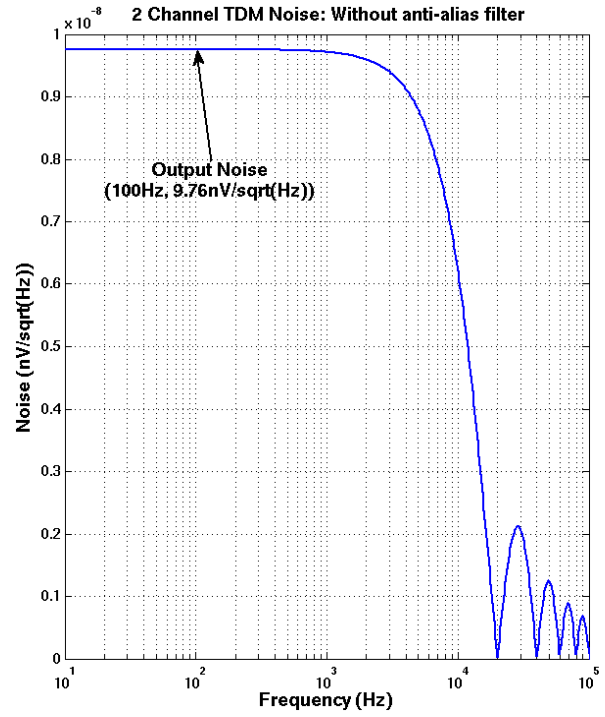
(b)
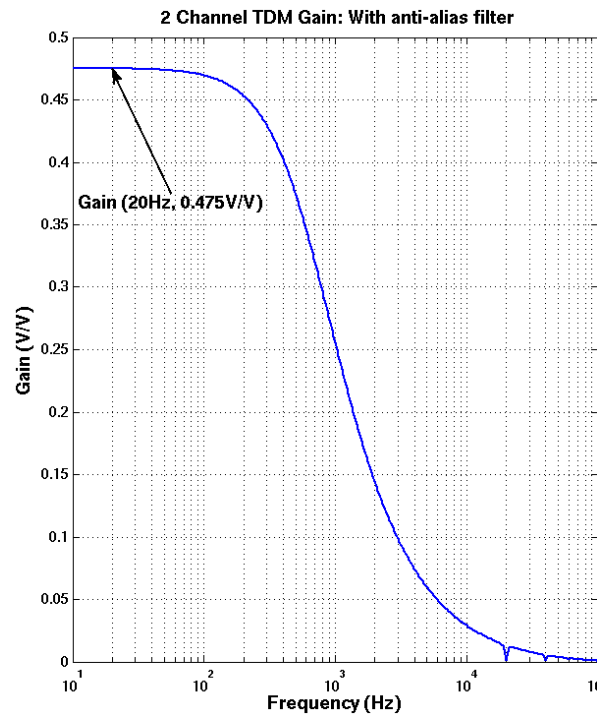
(c)
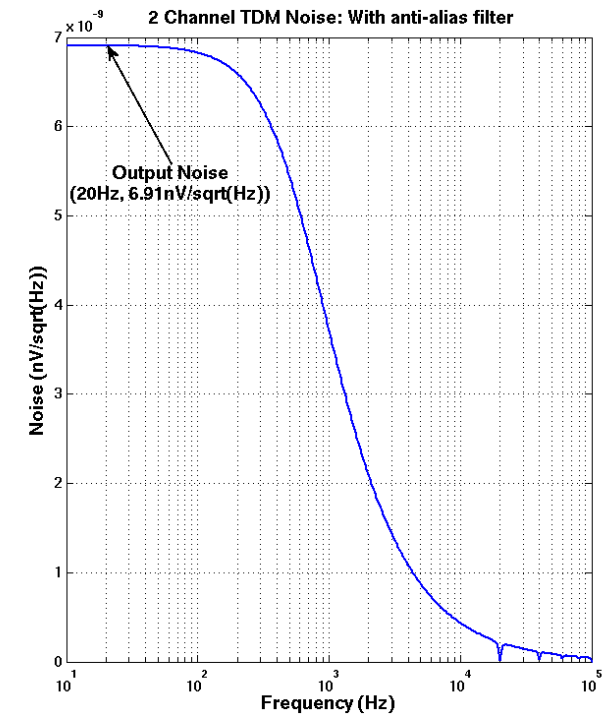
(d)

Fig 3.5: Simulated results for 2 channel TDM

16
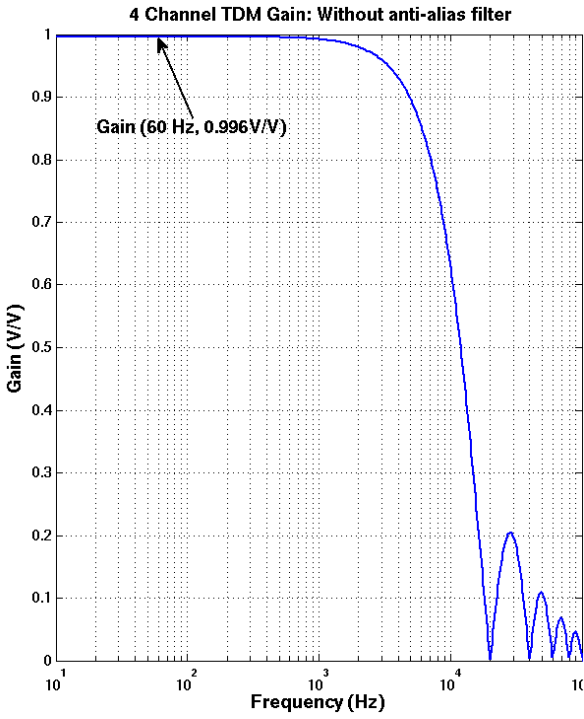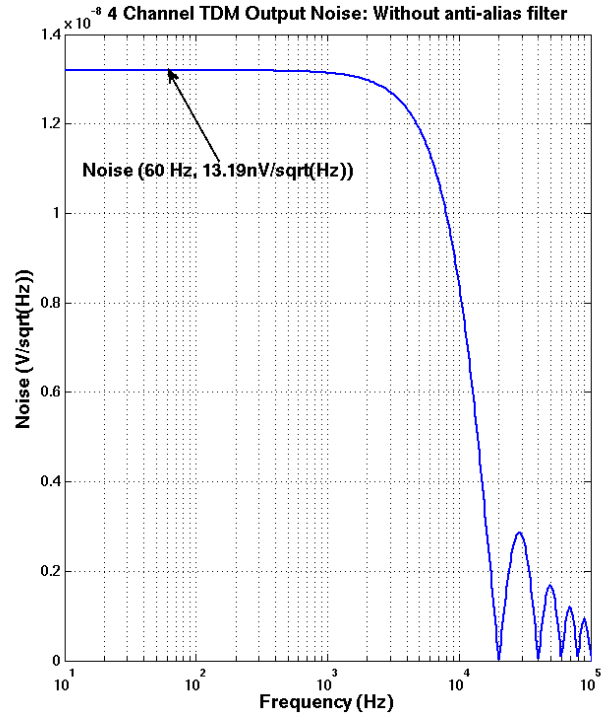
(a)   (b)
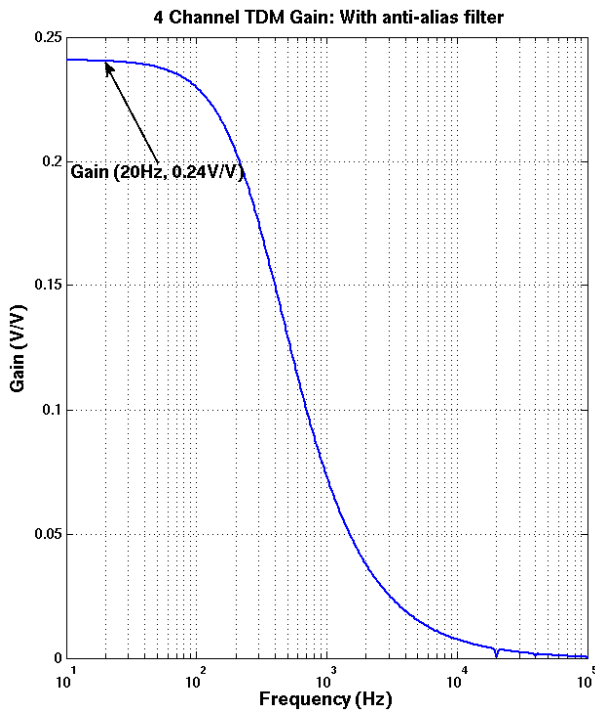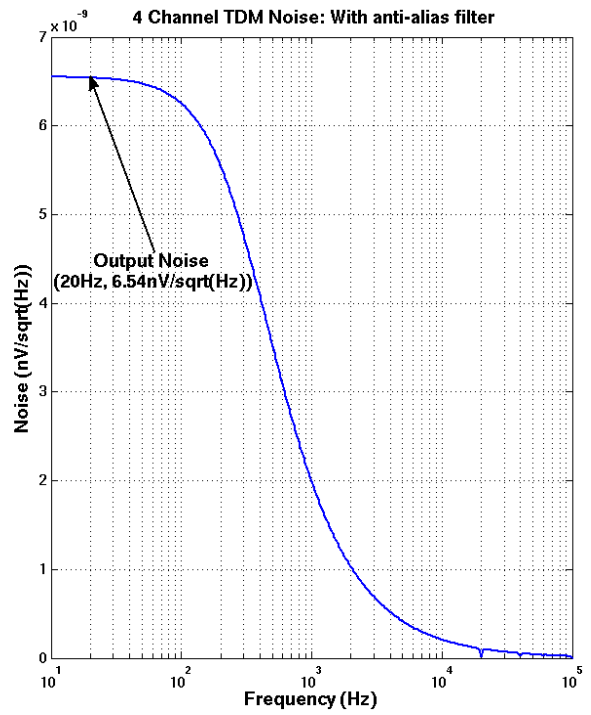
(c)   (d)

Fig 3.6: Simulated results for 4 channel TDM

When the anti-alias filter is not used, the output noise increases to 9.75 nV/√Hz (N = 2) and 13.19 nV/√Hz (N = 4), and the expected values are √2 x 6.9 nV/√Hz = 9.76 nV/√Hz and √4 x 6.547 nV/√Hz = 13.09 nV/√Hz, respectively. Thus the results of the above analysis for time division multiplexing are verified. The results are summarized in Table 3.1.

Apart from degradation in SNR, anti-alias filters are not desirable because the orthogonal property of the signals is lost. The signals are kept separate in the time domain by interleaving samples in different time slots. But if the interleaved samples pass through a low-pass filter, the samples from different channels will smear into each other. In the frequency domain, orthogonality is maintained due to the various images of the baseband data that are created with different phase shifts. By aligning the sampling clock to the correct channel, the images of the desired channel are demodulated back to baseband and add in phase, while the images of the other channels add up to zero. This property is lost if the images are filtered out.

Table 3.1: Summary of simulation results for TDM

Specification	2 Channel, No anti-alias filter	2 Channel, With anti-alias filter	4 Channel, No anti-alias filter	4 Channel, With anti-alias filter
Gain (Expected)	1	0.5	1	0.25
Gain (Simulated)	0.997	0.475	0.996	0.241
Noise (Expected)	9.76 nV/√Hz	7.08 nV/√Hz	13.09 nV/√Hz	7.08 nV/√Hz
Noise(Simulated)	9.756 nV/√Hz	6.9 nV/√Hz	13.19 nV/√Hz	6.5 nV/√Hz

Frequency division multiplexing is another technique by which multiple baseband channels can share the same spectrum, which in this case is the bandwidth of the fast amplifier. This technique is discussed in the next section.

CHAPTER 4

Frequency Division Multiplexing

In Frequency Division Multiplexing (FDM), channels are up-modulated to their respective carrier frequencies and then added together. The amplifier would process this composite signal, and on the receiver side, each channel is demodulated using their respective carriers (and a low-pass filter) to recover the original signal, as shown in Fig 4.1. Since the carriers have different frequencies, they remain orthogonal to each other, thus facilitating the extraction of channels at the receiver without inter-channel interference.
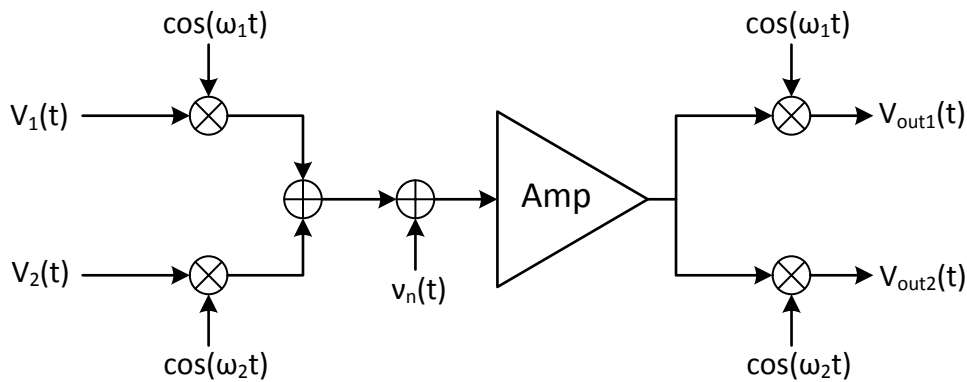
Fig 4.1: Signal flow for frequency division multiplexing

The mixer is realized using 4 cross-coupled switches, as shown earlier in Fig 2.2. The effect of this multiplexing on SNR is analyzed as follows. The signal paths of the input and noise are shown in Fig 4.2, with the amplifier gain set to 1. The wave $sq_1(t)$ represents a 50% duty-cycle square wave going from -1 to $+1$, with a fundamental frequency f_1. Thus, the signal gets reconstructed exactly as it was at the input, since $sq_1(t)$ multiplied with itself is 1. To analyze the noise at the received end, we use the following result. When a white noise process is multiplied with a periodic wave, the resulting PSD is the input PSD scaled by a factor equal to the power of

the periodic waveform (proof-Appendix A). Since the power of $sq_1(t)$ is 1, the output noise PSD is the same as the input PSD, provided the noise is white.
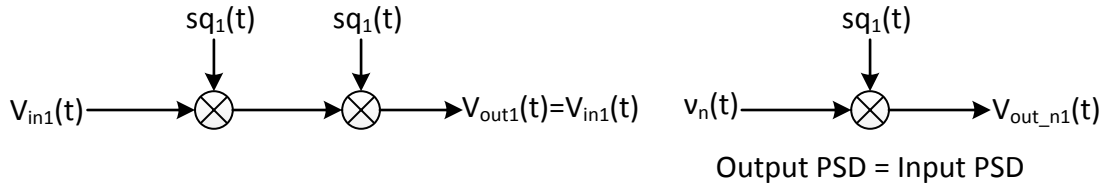
Fig 4.2: Signal paths of the input signal and noise

The important point to note here is that unlike TDM, the SNR at the receiver remains the same as a single channel system, irrespective of the number of channels being multiplexed. This can be inferred from the fact that the signal gets reconstructed perfectly at the output, and the noise PSD remains the same. Both are scaled by the gain of the amplifier. Thus, if FDM can be implemented using the fast amplifier, the overall power of the system can be reduced by the number of channels that share the fast amplifier. This assumes that other parts of the amplifier consume negligible power, which is not true in a realistic amplifier. The effect of these overheads is discussed later in section 4.3.

4.1 Architecture for Frequency Division Multiplexing

For FDM, the input signals need to be up-converted to their respective carrier frequencies, and then summed before the amplifier can process them. Adding signals is easily implemented by summing currents at a virtual ground, as in a conventional summing amplifier shown in Fig 4.3. Fig 4.4 shows a possible architecture for FDM. The chopper architecture has been extended to a

2-channel amplifier, where a second input is added to the virtual ground, and a different clock frequency f_2 is used to maintain orthogonality.
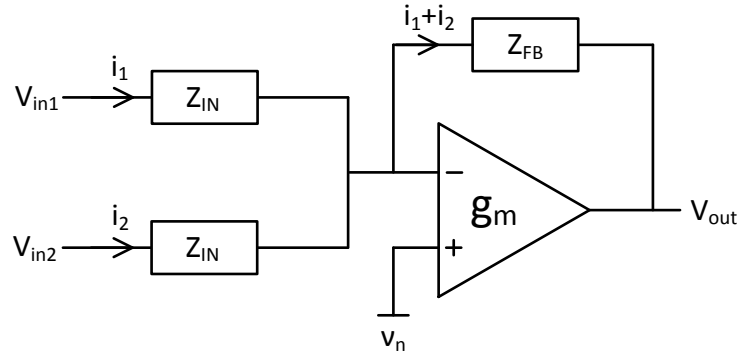
Fig 4.3: Summing amplifier
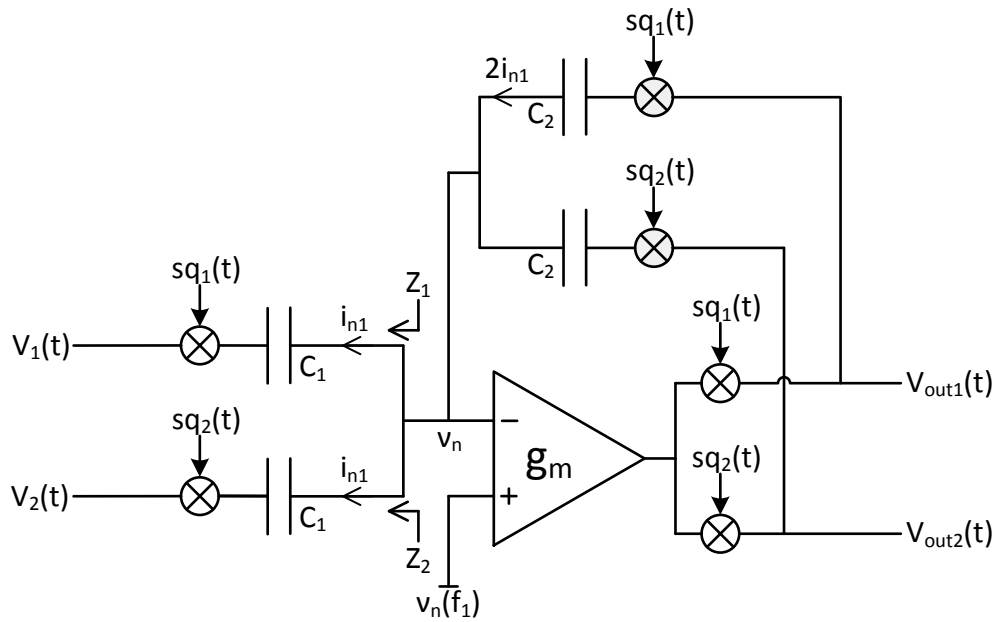
Fig 4.4: Extending the chopping technique as a summing amplifier

The problem with this architecture can be understood by referring to the conventional summing amplifier of Fig 4.3. The signal gain from any input to the output is $-Z_{FB}/Z_{IN}$, but the noise gain from the amplifier input to the output is $1 + 2Z_{FB}/Z_{IN}$. For an N-channel summing amplifier, the

signal gain remains the same, but the noise gain increases to $1 + NZ_{FB}/Z_{IN}$. Thus for large gains, the output noise power increases quadratically with increasing channel count. Thus the SNR degrades as

$$\text{SNR } \alpha \text{ } 1/N^2 \tag{4.1}$$

This is because, by introducing multiple channels, the feedback network gets loaded. Thus the noise gain increases with reducing feedback factor. This same effect is observed in the extended chopped architecture of Fig 4.4. The input arm of channel 2 (Z_2), i.e. the arm from the inverting terminal of the g_m stage to V_{in2}, reduces the feedback factor. This effect worsens with increasing channel count, giving the same relation between SNR and channel count as equation 4.1. The input arm impedances are shown in Fig 4.5(a).

The difference between the conventional summing amplifier and the chopped amplifier is that the signals of interest in the latter architecture are up-converted to different carrier frequencies. For a particular channel which uses the carrier frequency f_i, the noise of the amp around f_i gets down-converted to DC and is seen at the output.
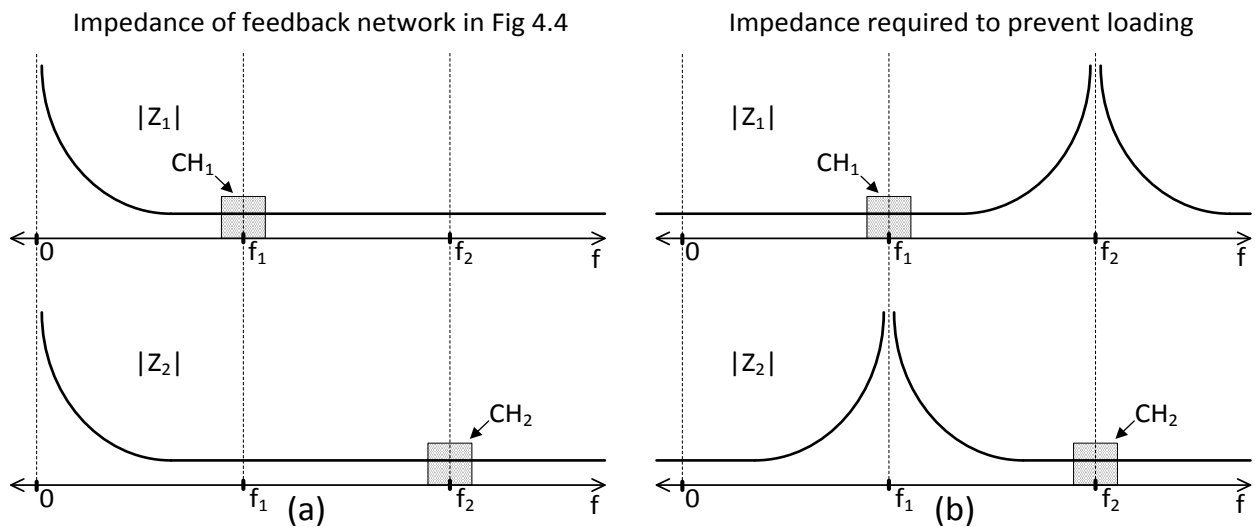
Fig 4.5: Impedance characteristics to reduce loading

Thus if the feedback network can be modified such that around any particular carrier frequency f_i, the loading effect due to the other N−1 channel arms can be reduced, then the noise gain from the input of the g_m stage to the output will be similar to that of a single channel amplifier. Fig 4.5(b) shows a possible set of impedances, which when used in the input arms of the feedback network (Z_1 and Z_2 in Fig 4.4) would reduce loading. The 2 carrier frequencies used for channel 1 and channel 2 are f_1 and f_2 respectively. Note that when Z_1 and Z_2 are in parallel, the impedance of the input arm of channel 1 dominates around f_1, and the impedance of the input arm of channel 2 dominates around f_2.

The impedance Z_1 is obtained when a capacitor is in series with a bilateral mixer, and the mixer operates on a clock with fundamental frequency f_1. Similarly, Z_2 is obtained with the same circuit, but the mixer is operated at a different frequency f_2.

(a)   (b)
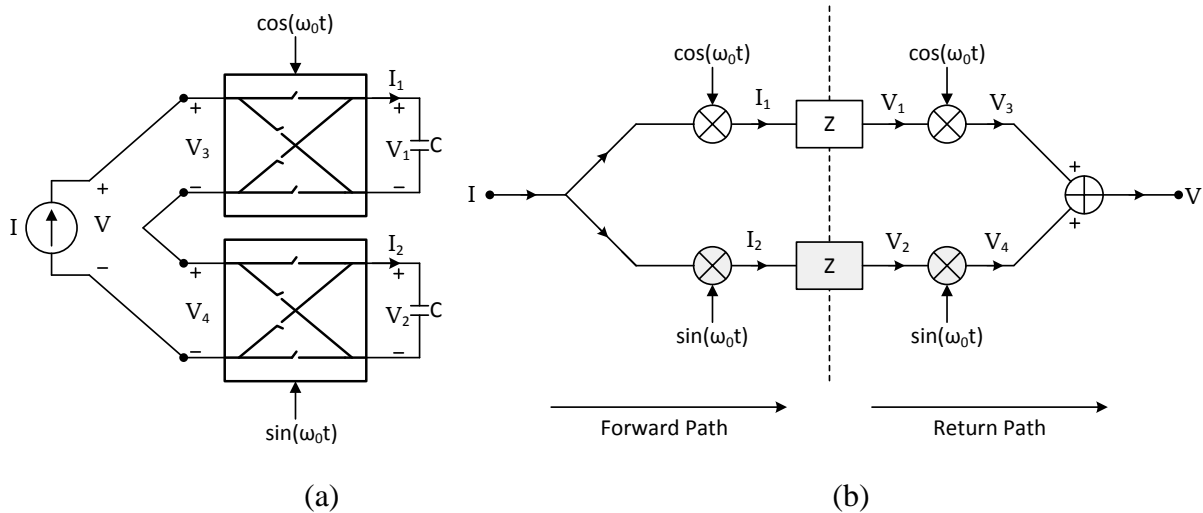
Fig 4.6 (a) Circuit for realizing a translated impedance, (b) Signal flow of circuit in (a)

Since the translated impedance is a key concept in this work, it is explained in detail. A translated impedance of a capacitor can be synthesized as shown in Fig 4.6(a). Note that the mixers used are bilateral. To find the impedance, the 1-port is driven by a current source denoted

23

as I in Fig 4.6(a), and a sinusoid at $\omega_o + \omega$ is applied. The corresponding frequency domain analysis is shown in Fig 4.7. Fig 4.7(a) shows the frequency content of the driving current. The 2 mixers are operated with clocks that are T/4 out of phase, where $T = 2\pi/\omega_o$. The carriers are shown in Fig 4.7(b) and Fig 4.7(c). Note that the passive mixer implements the multiplication of the input current with a square wave, but for ease of analysis only the fundamental frequencies are considered. A general analysis using all the harmonics can be found in [19]. From the signal flow graph in Fig 4.6(b), it can be seen that

$$i_1 = I\cos(\omega_0 t) , \quad i_2 = I\sin(\omega_0 t) \tag{4.2}$$

Define i_{C1} as
$$i_{C1}(t) = i_1 + ji_2 \tag{4.3}$$

Thus,
$$i_{C1}(t) = Ie^{j\omega_0 t} \tag{4.4}$$

The complex current i_{C1} is equivalent to the complex up-conversion of the input current I to the carrier ω_0. The frequency content of the currents i_1, i_2 and i_{C1} are shown in Fig 4.7(d)-(f) respectively. The current i_{C1} flows into impedance Z and produces voltages V_1 and V_2 as shown in Fig 4.6(b). V_{C1} is defined as

$$v_{C1}(t) = v_1(t) + jv_2(t) = (Ie^{j\omega_0 t}) \times Z(j\omega) \tag{4.5}$$

Thus, v_{C1} is the complex voltage generated when the complex current i_{C1} flows through the impedance Z. From Fig 4.6(b)

$$v_3 = v_1\cos(\omega_0 t) , \quad v_4 = v_2\sin(\omega_0 t) \tag{4.6}$$

Define v_{C2} as
$$v_{C2}(t) = v_{C1}(t)e^{-j\omega_0 t} \tag{4.7}$$

The complex voltage v_{C2} as defined in equation 4.7 is equivalent to the complex down-conversion of the voltage v_{C1}. The reason for defining v_{C2} this way will become clear in a moment. The frequency content of v_{C1} and v_{C2} are shown in Fig 4.7(h) and Fig 4.7(i) respectively. Note that the phase of the impedance Z has been considered, i.e. arc$\{Z(j\omega)\} = -\pi/2$ for $\omega>0$, and arc$\{Z(j\omega)\} = \pi/2$ for $\omega<0$. From Fig 4.6(b), v(t) can be found as

$$v(t) = v_3(t) + v_4(t) \tag{4.8}$$

From equations 4.6-4.8, v(t) can be written as

$$v(t) = v_1(t)\cos(\omega_0 t) + v_2(t)\sin(\omega_0 t) = \text{Re}\{v_{C2}(t)\} \tag{4.9}$$

By defining v_{C2} as in equation 4.7, the resulting driving point voltage v(t) can be easily obtained using equation 4.9. The frequency content of v(t) is shown in Fig 4.7(j).

Thus,
$$V(j\omega + j\omega_0) = I(j\omega + j\omega_0) \times \frac{1}{2}\left[Z(j\omega) + Z(j2\omega_0 + j\omega)\right] \tag{4.10}$$

Or
$$Z_{IN}(j\omega + j\omega_0) = \frac{1}{2}\left[Z(j\omega) + Z(j2\omega_0 + j\omega)\right] \approx \frac{1}{2}Z(j\omega) \tag{4.11}$$

Thus, the impedance observed at the input of the 1-port around the carrier frequency ω_0 is the baseband impedance Z. This assumes that Z is low-pass in nature, so that the higher frequency copies of Z that appear around ω_0 can be neglected. In the above analysis, the carriers used were $\cos(\omega_0 t)$ and $\sin(\omega_0 t)$, but the mixers implement multiplication with ±1 square waves, which have a fundamental component of $(4/\pi)\cos(\omega_0 t)$ or $(4/\pi)\sin(\omega_0 t)$. Thus the observed input impedance is

$$Z_{IN}(j\omega + j\omega_0) \approx \frac{2}{\pi}Z(j\omega) \tag{4.12}$$

25

(a) $i(t)=\cos((\omega_0+\omega)t)$

$\frac{1}{2}$ $\frac{1}{2}$

$-3\omega_0$ $-2\omega_0$ $-\omega_0$ 0 ω_0 $2\omega_0$ $3\omega_0$ ω

(b) $\cos(\omega_0 t)$

$\frac{1}{2}$ $\frac{1}{2}$

$-3\omega_0$ $-2\omega_0$ $-\omega_0$ 0 ω_0 $2\omega_0$ $3\omega_0$ ω

(c) $\sin(\omega_0 t)$

$-\frac{j}{2}$ $\frac{j}{2}$

$-3\omega_0$ $-2\omega_0$ $-\omega_0$ 0 ω_0 $2\omega_0$ $3\omega_0$ ω

(d) $i_1(t)$

$\frac{1}{4}$ $\frac{1}{4}$ $\frac{1}{4}$ $\frac{1}{4}$

$-3\omega_0$ $-2\omega_0$ $-\omega_0$ $-\omega$ ω ω_0 $2\omega_0$ $3\omega_0$ ω

(e) $i_2(t)$

$\frac{j}{4}$ $-\frac{j}{4}$ $\frac{j}{4}$ $-\frac{j}{4}$

$-3\omega_0$ $-2\omega_0$ $-\omega_0$ 0 ω_0 $3\omega_0$ ω

(f) $i_{C1}=i_1 + ji_2$

$\frac{1}{2}$ $\frac{1}{2}$

$-3\omega_0$ $-2\omega_0$ $-\omega_0$ $-\omega$ ω ω_0 $2\omega_0$ $3\omega_0$ ω

(g) $|Z|$

$-3\omega_0$ $-2\omega_0$ $-\omega_0$ 0 ω_0 $2\omega_0$ $3\omega_0$ ω

(h) $v_{C1}=v_1+jv_2$

$\frac{1}{2}Z(j2\omega_0 + j\omega)$

$\frac{1}{2}Z(-j\omega)$

$-3\omega_0$ $-2\omega_0$ $-\omega_0$ 0 ω_0 $3\omega_0$ ω

(i) $v_{C2}=v_{C1}e^{-j\omega_0 t}$

$\frac{1}{2}Z(j2\omega_0 + j\omega)$

$\frac{1}{2}Z(-j\omega)$

$-3\omega_0$ $-2\omega_0$ $-\omega_0$ 0 ω_0 $3\omega_0$ ω

(j) $v=Re\{v_1 e^{-j\omega_0 t}\}$

$\frac{1}{4}\left[Z(j\omega)+Z(j2\omega_0 + j\omega)\right]$

$\frac{1}{4}\left[Z(-j\omega)+Z(-j2\omega_0 - j\omega)\right]$

$-3\omega_0$ $-2\omega_0$ $-\omega_0$ 0 ω_0 $3\omega_0$ ω
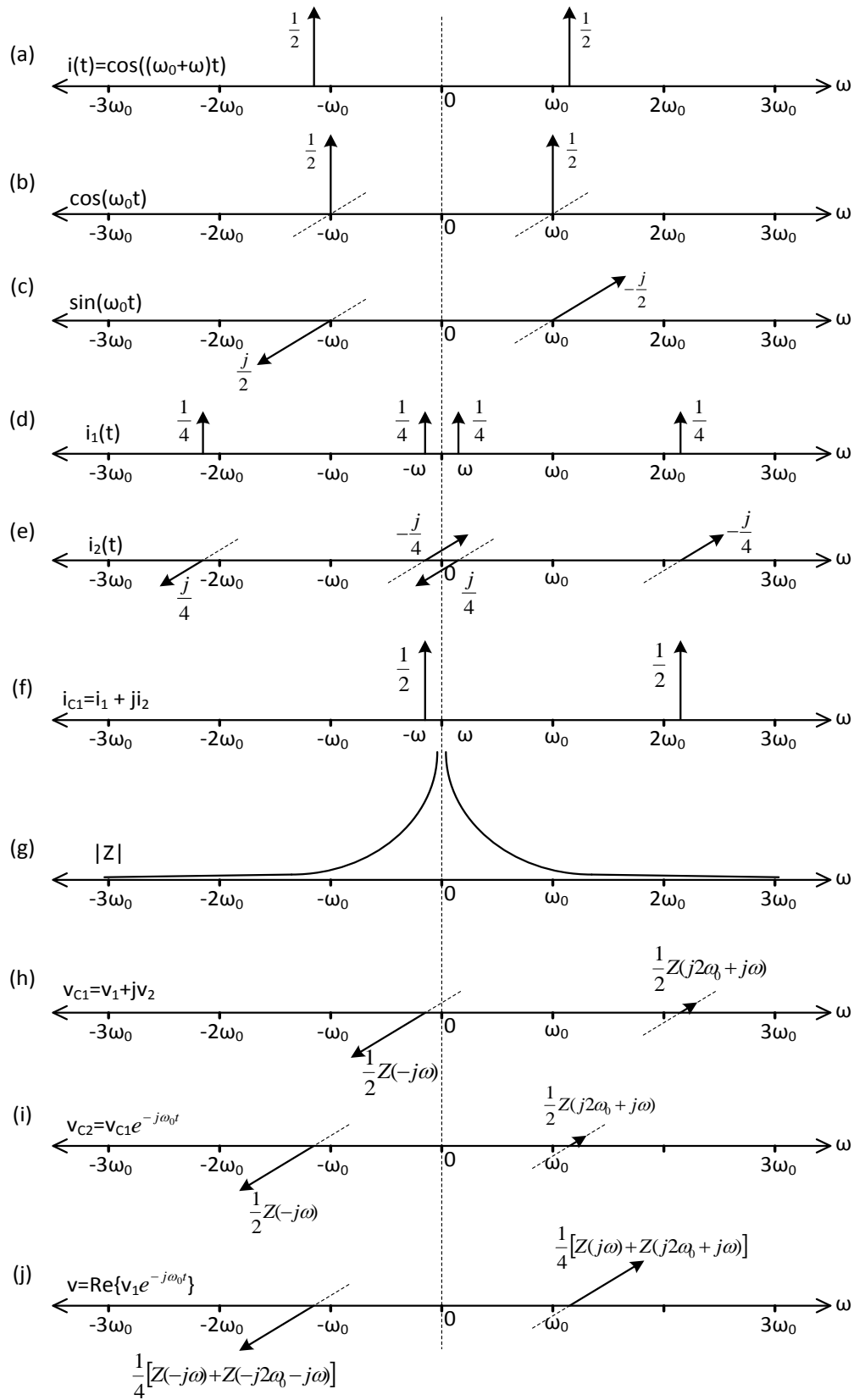
Fig 4.7: Frequency domain analysis of impedance translation

In the impedance translation implementation discussed above, the I-path (in-phase path) denotes the signal path driven by $\cos(\omega_0 t)$, and the Q-path (quadrature path) denotes the signal path driven by $\sin(\omega_0 t)$. If only one path is used, then an image will be created at $\omega_0-\omega$ for an input at $\omega_0+\omega$. This can be seen from Fig 4.6(b), where if only one path (say the I-path) is used, then the output v(t) is equal to $v_3(t)$. The frequency content of $v_3(t)$ is shown in Fig 4.8.
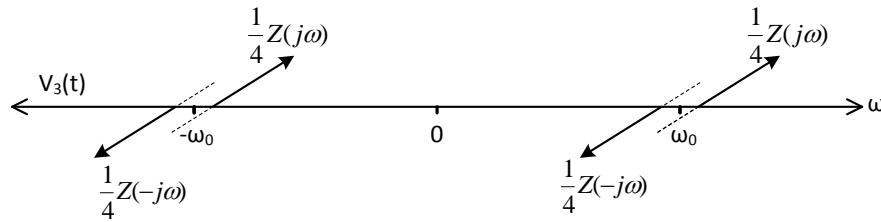
Fig 4.8: Frequency content of $v_3(t)$

Thus, any input with an offset around ω_0 will generate an image around ω_0 at the negative offset frequency. This will destroy any useful information of the applied input. But for noise, we don't require that the imaging problem be solved as long as the noise power remains the same, and the filtering effect is obtained. The input is a real signal at baseband (near DC), thus it never suffers the imaging problem in this signal flow. Hence, instead of both I and Q paths, one path will be enough to get the required filtering effect, as shown in Fig 4.9.
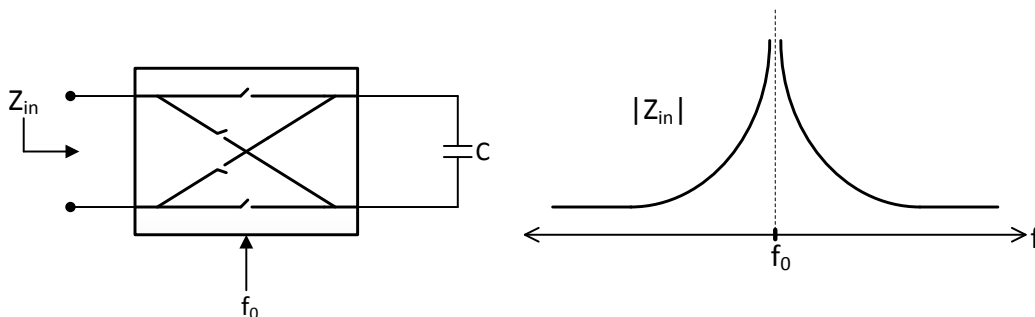
Fig 4.9: Single phase mixer based impedance translation

For channel 1, the impedance of the capacitor C_1 should be translated to f_2, while the input signal is translated to f_1 (and vice-versa for channel 2). The feedback network is modified accordingly (Fig 4.10), and the new impedances Z_1 and Z_2 obtained are as shown before in Fig 4.5(b). The circuit enclosed in the dashed-box in Fig 4.10 generates the required peak in the impedance characteristics, while the mixers MX_1 and MX_2 translate the input signal to the desired frequency. Note the positioning of channel 1 and channel 2 in Fig 4.5(b). The feedback around the g_m stage is completed by replicating a scaled version of the input arm between the output and the inverting terminal, as done in a conventional inverting amplifier as shown in Fig 2.1(a).
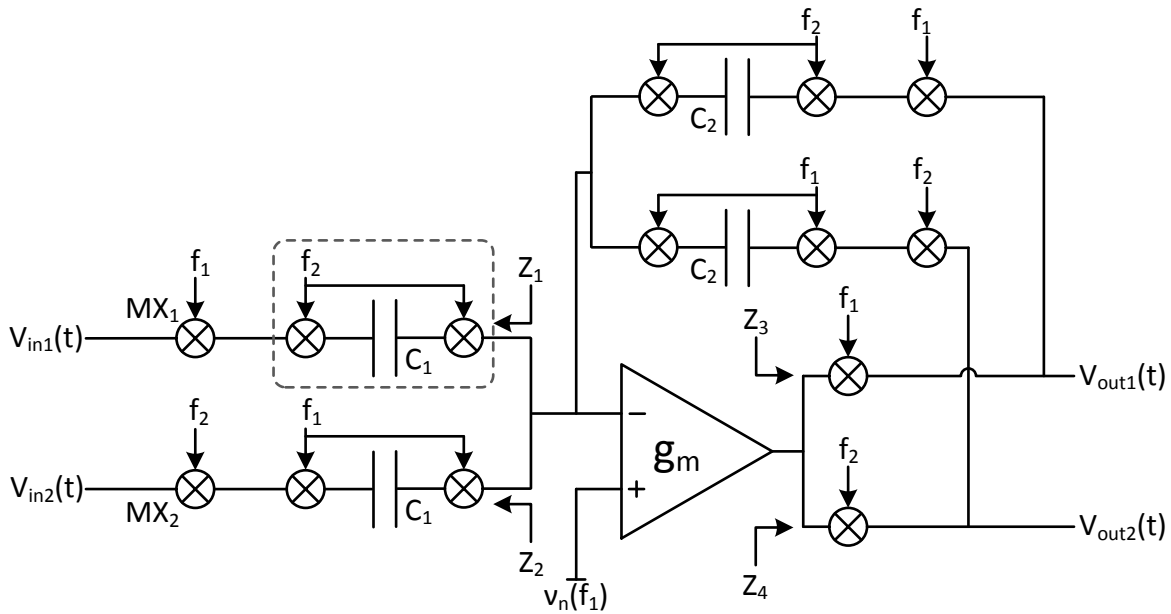
Fig 4.10: Amplifier with modified feedback network to prevent loading

Although the implementation in Fig 4.10 solves the loading problem of the feedback loop, the same problem is present at the output of the g_m stage. The outputs V_{out1} and V_{out2} will be connected to some reasonable load capacitance. Thus the impedance Z_3 as marked in Fig 4.9 is the load capacitance seen by V_{out1} (in parallel with the feedback arm) translated to f_1. Z_4 is similar to Z_3, but the translated frequency is f_2. Z_3 and Z_4 appear in parallel at the output of the 1st

28

g_m stage. As shown in Fig 4.11(a), Z_3 reaches a maximum at f_1, when the impedance of the other path (Z_4) is low. Thus the output current produced due to channel 1 would flow to V_{out2} instead
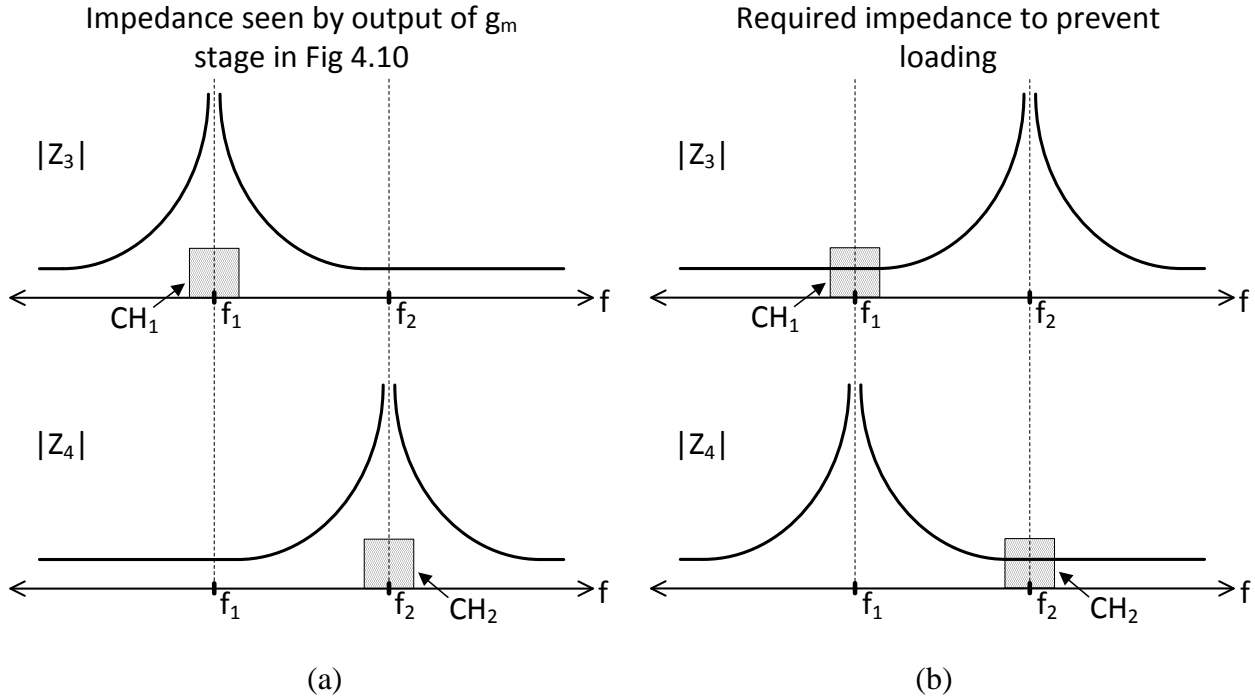
Fig 4.11: Loading observed by the 1st stage g_m

of V_{out1} (and vice-versa for channel 2). To prevent this, the modification shown in Fig 4.12 is made, and the impedance characteristics shown in Fig 4.11(b) are obtained. A second g_m stage, denoted by g_{m2} in Fig 4.12, is used to create a virtual ground at the input of g_{m2}. As done before, the circuit enclosed in the dashed-box in Fig 4.12 creates the required peaks in the impedance observed by the output of the 1st gm stage. Now, the current produced by channel 1 will flow to V_{out1} since the impedance Z_3 is much smaller than Z_4 around f_1. Similarly, the current produced by channel 2 will flow to V_{out2} since the impedance Z_4 is much smaller than Z_3 around f_2. The mixers MX_3 and MX_4 down-convert the signal current generated by the 1st stage g_m to baseband, by multiplying channel 1's current with frequency f_1 and channel 2's current with frequency f_2,
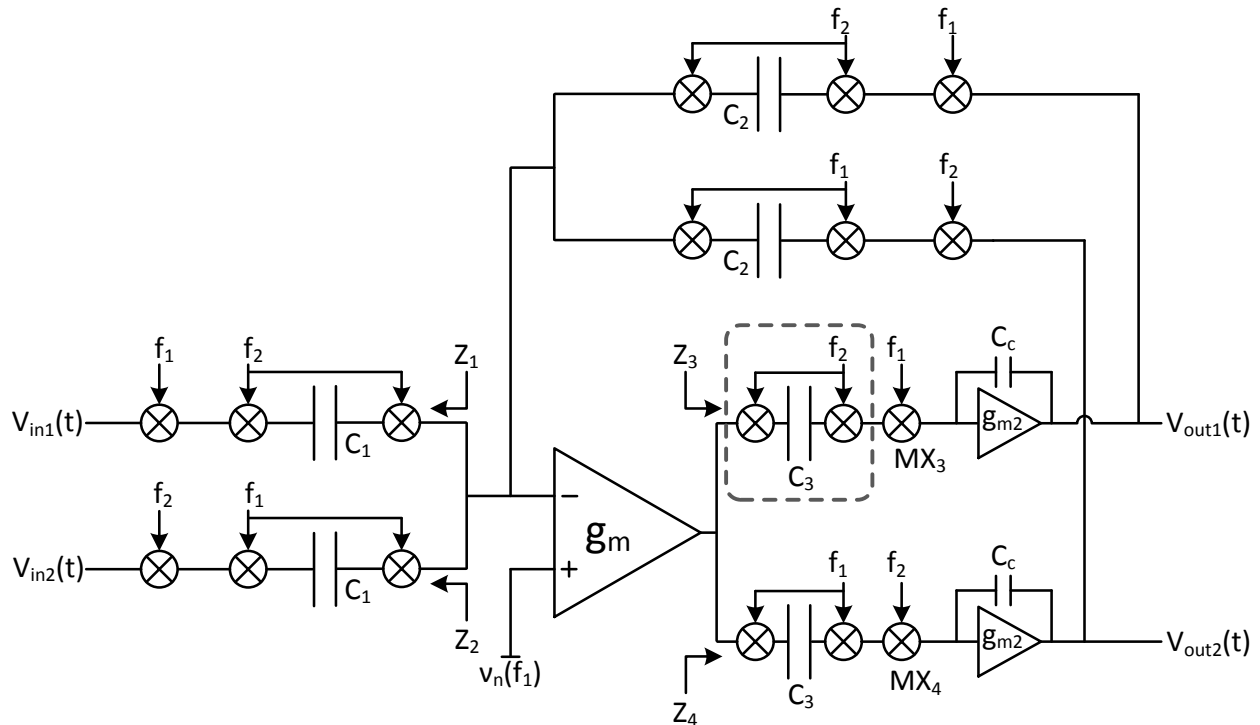
Fig 4.12: Amplifier with modified feedback network and reduced loading at 1st stage output

respectively. Thus, the idea is to use mixers and capacitors to create favorable impedances that give the required filtering properties. This enables the multi-channel FDM implementation to have the same noise performance as the single channel implementation. Another modification that can be made to the above architecture is to combine any 2 mixers in series, and drive them with the composite clock derived by multiplying the 2 individual clocks, as shown in Fig 4.13. This reduces the number of mixers in the signal path, thus reducing the thermal noise contribution due to the switches.
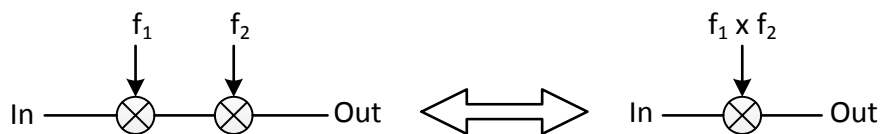
Fig 4.13: Combining series mixers
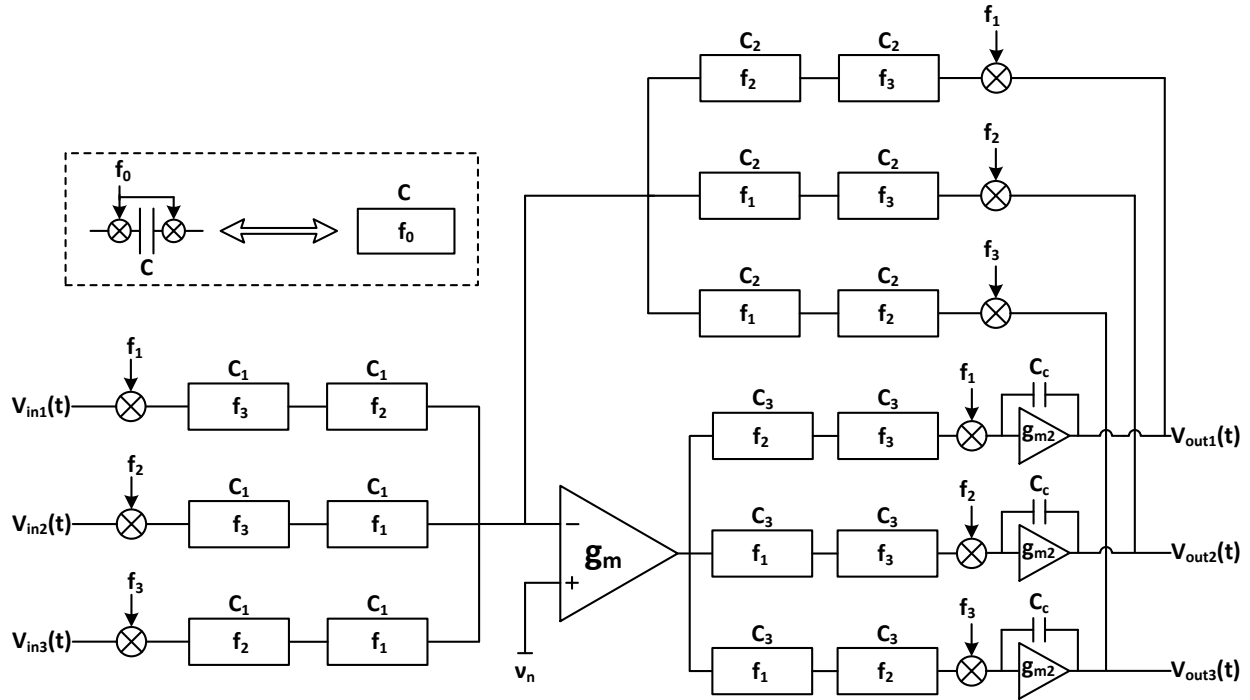
Fig 4.14: Extending FDM to multiple channels

The architecture in Fig 4.12 can be extended to multiple channels. The 3-channel amplifier is shown in Fig 4.14. For any particular channel with carrier f_k, there should be N−1 impedance blocks (dashed-box in Fig 4.14), which create peaks in the impedance characteristics at all other N−1 carrier frequencies, thus preventing loading.

The above architecture assumes that there are multiple channels that need to be amplified. But there are many applications where the channel count is just one or two. For such applications, the frequency division multiplexed amplifier can be adapted as follows. If only 1 channel needs to be amplified, then in Fig 4.14, the channel is applied to all 3 inputs V_{in1-3}. The outputs V_{out1-3} all contain the same data, i.e. the amplified version of the input. But the thermal noise present at each of these outputs is uncorrelated with each other, since the noise at each output is the noise of the amplifier, but frequency translated to the different carrier frequencies used. Hence, if the

31

outputs V_{out1-3} are summed, the SNR increases by the factor N. This is similar to diversity gain, where multiple copies of the signal of interest are available, with the noise in each signal being uncorrelated with each other. This architecture is shown in Fig 4.15.
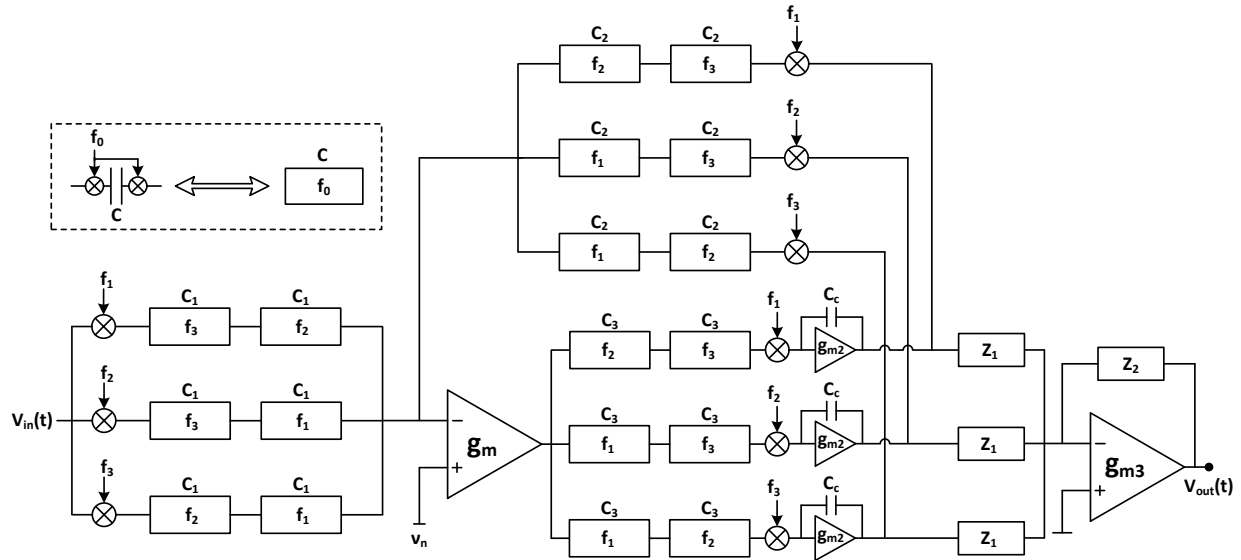
Fig 4.15: Modified FDM architecture for a single input channel

In the architecture of Fig 4.14 and Fig 4.15, the aim was to share the 1st g_m stage, so that the power consumed per channel is reduced. But one may argue that to accomplish this, extra power-consuming hardware has been used, like the passive mixers and the 2nd g_m stage. So would the savings be significant, or would there be savings at all? These overheads have been considered and analyzed later in Section 4.3.

4.2 Amplifier design

In the above discussion, it was shown that the loading problems inherent to a multi-channel implementation can be solved by using the techniques discussed. This simplifies the amplifier

design to a conventional single-channel design. A capacitively coupled 2-stage amplifier is used as shown in Fig 4.16(a). The closed-loop gain is given by

$$\frac{V_{out}}{V_{in}} = A_{CL} = -\frac{C_1}{C_2} \tag{4.13}$$

<table>
<tr><td>(a)</td><td>(b)</td></tr>
</table>
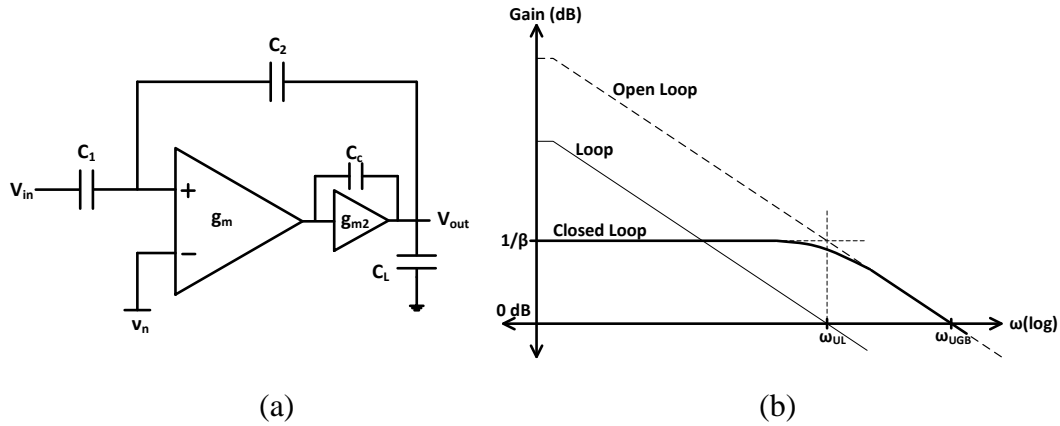
Fig 4.16(a) Capacitively coupled inverting amplifier, (b) Feedback analysis

This feedback can be analyzed as a voltage-sense, voltage-feedback circuit, with v_n as the input and V_{out} as the output. Note that although the feedback analysis and loading don't consider the mixers, the analysis remains the same because the effect of the mixers is to simply translate the impedance and the signal to the carrier frequency. As long as loading is prevented, the analysis of the feedback loop with the mixers is the same as that for a conventional feedback loop. Assuming a 1st-order roll-off, the open-loop gain and the unity-gain frequency are given by

$$A_{Open} = -\frac{g_m}{j\omega C_C} \quad , \quad \omega_{UGB} = \frac{g_m}{C_C} \tag{4.14}$$

The fraction of the output voltage that is fed back is given by

$$\beta = \frac{C_2}{C_1 + C_2} \approx \frac{C_2}{C_1} \tag{4.15}$$

33

The approximation is valid for large closed-loop gain. From the above expressions, the loop unity-gain frequency or the closed-loop −3dB bandwidth is given by

$$\omega_{UL} = \beta \omega_{UGB} = \beta \frac{g_m}{C_C} \qquad (4.16)$$

To ensure sufficient phase margin, the 2nd pole should be further away than the unity loop-gain frequency. For a closed-loop Butterworth response, the 2nd pole should be at

$$\omega_{P2} = \frac{g_{m2}}{C_C + C_L} \approx 2\omega_{UL} \qquad (4.17)$$

From the above expressions, g_m and g_{m2} can be related as follows

$$g_{m2} = 2\beta g_m \left(1 + \frac{C_L}{C_C}\right) \qquad (4.18)$$

A closed-loop gain of 40 sets $\beta=1/41$. C_C is chosen to be 2pF, and C_L is 1pF. Thus,

$$g_{m2} = g_m \left(\frac{3}{40}\right) \qquad (4.19)$$

The phase margin would be about 63°. This assumes that there are no other poles or zeros. But Miller compensation gives rise to a right-half plane zero given by

$$\omega_Z = \frac{g_{m2}}{C_C} \qquad (4.20)$$

This zero will reduce the phase margin of the loop. With the above choice of g_m and g_{m2}, the phase margin is >45°, which is acceptable. If the transconductance per unit bias current of the 1st and 2nd stage is the same, then the bias currents I_{S1} and I_{S2} are related as

$$I_{S2} \approx \frac{I_{S1}}{13} \tag{4.21}$$

If I_{S1} is chosen as 500nA, then I_{S2}=40nA. For a transistor in weak inversion

$$\frac{g_m}{I_D} = \frac{1}{\eta U_T} \approx 25 \tag{4.22}$$

where U_T is the thermal voltage (26mV at room temperature) and η is the slope factor (assumed to be 1.5). Thus if I_{S1}=500nA, this gives g_m = 12.5µA/V, and $g_{m2} \approx$ 1µA/V. With the above capacitor choices, the closed loop bandwidth is 24.3 kHz, which is more than sufficient for neural recording systems. The value of C_C can be increased to further reduce the bandwidth, or another stage of low pass filters can be used for bandwidth selectivity.
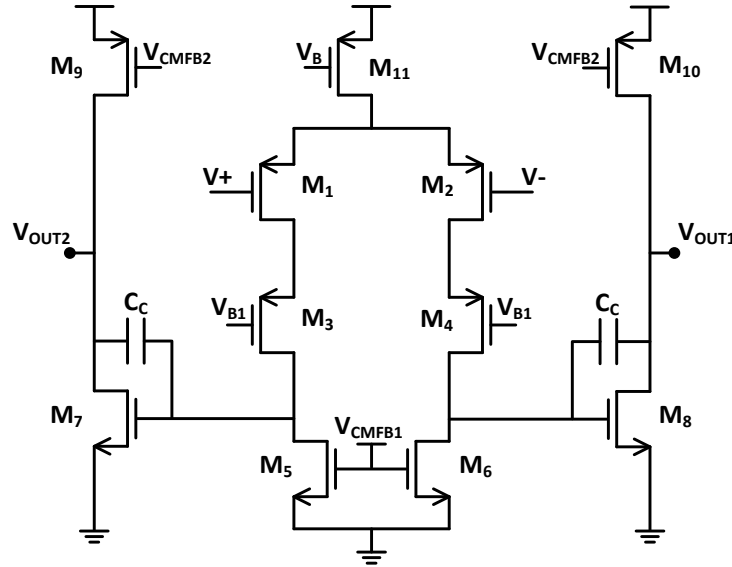

Fig 4.17: Conventional 2-stage fully differential amplifier

Using the above values, the input-referred noise can be found ($\gamma \approx$ 1.25, from simulations)

$$v_{n,IN} \approx \sqrt{\frac{8kT\gamma}{g_m}} = 57.4\,nV/\sqrt{Hz}, \quad v_{n,RMS} = \sqrt{v_{n,IN}^2 \times BW} = 4.45\,\mu V \tag{4.23}$$

35

Now that the above analysis is complete, any suitable amplifier architecture can be used to realize the transconductance stages. A possible amplifier architecture is shown in Fig 4.17. The 1st stage is realized by a PMOS differential amplifier with an active load. The 2nd stage is an NMOS Common Source Amplifier (CSA) with an active load. Both stages have individual common-mode feedback loops to control the bias of the active loads. The input-referred thermal noise density of the amplifier in Fig 4.17 is given by

$$v_{n,IN}^2 = \frac{8kT\gamma}{g_{m1,2}}\left(1 + \frac{g_{m5,6}}{g_{m1,2}}\right)$$

(4.24)

The noise contribution due to the 2nd stage is neglected since it will get divided by the gain of the 1st stage. The cascode transistors $M_{3,4}$ contribute little noise, since their noise currents never make it to the output (assuming infinite degeneration at the source). From the above expression, $g_{m1,2}$ should be maximized, and $g_{m5,6}$ should be minimized for a given bias current. For practical biasing, the lowest noise spectral density is found to be

$$v_{n,IN}^2 = \frac{8kT\gamma}{g_{m1,2}}\left(1 + \frac{1}{6}\right)$$

(4.25)

Where $g_{m5,6}/g_{m1,2} = 1/6$. Another possible architecture is shown in Fig 4.18, where the transconductance of the load transistors is also used for signal gain [15]. The input-referred noise spectral density is given by

$$v_{n,IN}^2 = \frac{8kT\gamma}{(g_{m1,2} + g_{m3,4})^2}\left(g_{m1,2} + g_{m3,4}\right) = \frac{8kT\gamma}{g_{m1,2} + g_{m3,4}}$$

(4.26)

The noise PSD of this amplifier can be atleast 2 times lower than the conventional differential amplifier. This is because the effective transconductance of the amplifier has doubled, and the output noise power of the 1st stage has also doubled. Thus, the signal gain has increased by a factor of 4 in power, but the noise power increased by 2, which is equivalent to a reduction in the input-referred noise.
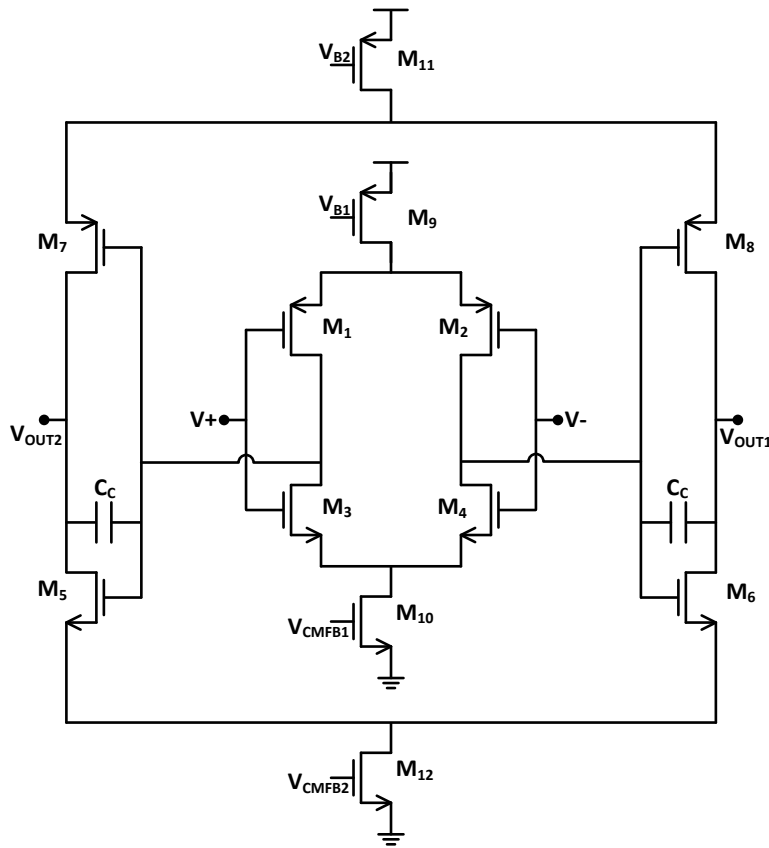
Fig 4.18: Inverter based amplifier for lower input-referred noise

Although the 2nd stage noise is suppressed, the inverter based amplifier is used here as well. This is because the transconductance of the 2nd stage has a minimum bound (equation 4.18), which is set by the required phase margin for stability. But by using the inverter based amplifier, the power burnt in the 2nd stage can be reduced by a factor of 2. The disadvantage of this topology as compared to a CSA in the 2nd stage is the limited output swing. But the expected output swing is

37

approximately 100 mV$_{p-p}$ or a peak swing of 25 mV at any output, which is small. Since the swing requirement is small, the topology in Fig 4.18 meets the requirements. The closed-loop gain could be made higher (which would increase the swing requirements), but it would increase the required capacitor area, because C_1 would have to be larger for higher gain. But using a second stage for gain considerably reduces the total required capacitance, thus saving area. This is illustrated in Fig 4.19. Hence, the 1st stage gain need only be high enough to reduce the noise contributions of subsequent stages, so that the power burnt in the subsequent stages is small. But the gain of the first stage should not be very large, as it would degrade the linearity of the signal chain. This led us to choose the first stage closed-loop gain as 40, which results in less than 100 mV$_{p-p}$ swings at the 1st stage output, as measured between V$_{out1}$ and V$_{out2}$ in Fig 4.18. The rest of the gain is achieved by a 2nd stage of low-pass filters.

$$A = \frac{C_1}{C_2} = 900 \qquad\qquad A_1 = \frac{C_1}{C_2} = 30 \quad A_2 = \frac{C_3}{C_4} = 30$$

V$_{in}$ — A — V$_{out}$ $\qquad$ V$_{in}$ — A$_1$ — V$_{o1}$ — A$_2$ — V$_{out}$

$$C_{Tot} = 901 \cdot C_2 \qquad\qquad\qquad C_{Tot} = 62 \cdot C_2$$
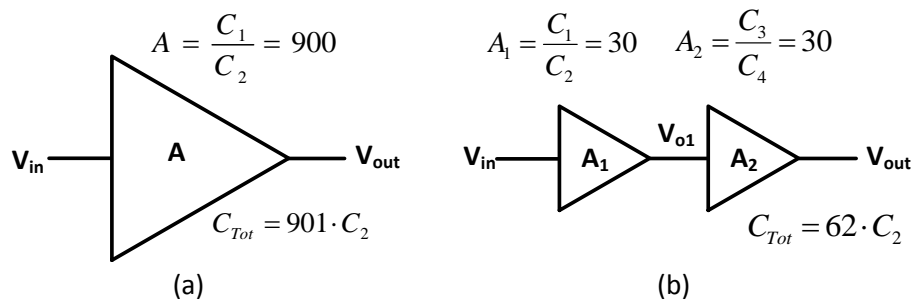
(a) $\qquad\qquad\qquad\qquad\qquad$ (b)

Fig 4.19: An example of area tradeoffs with number of stages

With these advantages over the conventional 2-stage differential amplifier, the amplifier topology of Fig 4.18 is used to realize the multi-channel frequency division multiplexed amplifier.

4.3 Overheads

As we've seen before, there is some cost in implementing FDM in an amplifier. For the multi-channel approach to be effective, this added cost needs to be analyzed and should be kept low, so that there is an overall saving in power by sharing the 1st g_m stage. There are 3 main overheads that could cause problems as the channel count is increased, which are the degrading noise performance of the amplifier, the increased power due to the 2nd g_m stage, and the increasing number of passive mixers. These are analyzed below.

Noise leakage: Although the techniques proposed above significantly reduce the loading effect, the loading is not completely eliminated. For any channel, the feedback factor is exactly the same as the single-channel implementation only if the conductance of the other channel paths (from the virtual ground) is zero. In Fig 4.17, an N-channel implementation is shown. The noise currents for channel 1 are shown. Ideally, the conductance of the other channel arms from the virtual ground should be zero. Hence, the noise currents flowing through the other channels should be zero. But the conductance is not zero, and it can be modeled as a fraction $(1/K)$ of the conductance of channel 1. Thus, the noise current through every other channel is i_{n1}/K.
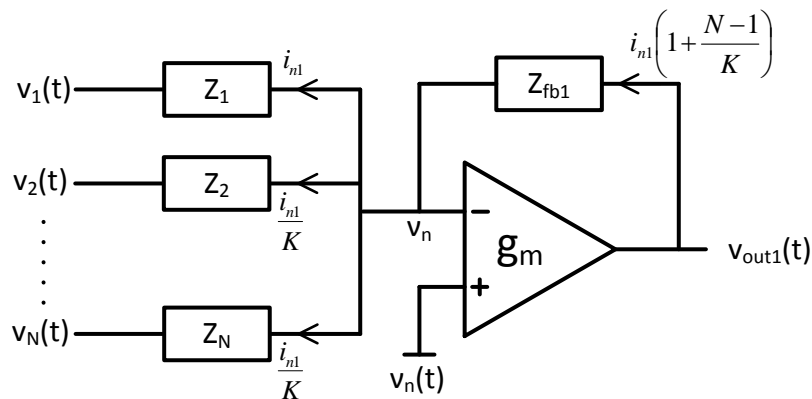

Fig 4.20: Noise degradation in an N-channel amplifier due to leakage

This gives a net noise gain given by

$$\frac{V_{out1}}{v_n} = 1 + \frac{Z_{fb1}}{Z_1}\left(1 + \frac{N-1}{K}\right) \approx \frac{Z_{fb1}}{Z_1}\left(1 + \frac{N-1}{K}\right) \qquad (4.27)$$

The term Z_{fb1}/Z_1 is the noise gain of a conventional single-channel amplifier. This shows that the output noise degrades as the channel count is increased, by the factor $1+(N-1)/K$. For the FDM amplifier, the factor K can be approximated as $\Delta f/f_{BW}$, where Δf is the frequency separation of the carriers used, and f_{BW} is the bandwidth of the input signal.

Mixers: The number of mixers used per channel will increase approximately by N since the number of impedance blockers needed per channel is N−1. Also, every mixer will contribute to thermal noise at every channel output. This is because the mixer thermal noise is wideband, thus different frequency shifted versions of the noise will appear at the baseband of every channel output. Thus the resistance of the switches used in the mixer should scale down quadratically as the channel count is increased, so that the relative noise contribution of the mixers is kept low. This scaling would result in larger driving capacitance to the clocks, thereby increasing the power burnt in driving the mixers. Since the amplifier noise also increases with N, the switch size need not increase proportional to N^2. If the same relative noise performance is to be maintained, the switch size should be increased as

$$Size \propto \frac{N^2}{\left(1 + \frac{N-1}{K}\right)^2} \qquad (4.28)$$

Thus, the mixer driving power per channel will increase as

$$P_{mixer} \propto \frac{N^2}{\left(1 + \frac{N-1}{K}\right)^2} \qquad (4.29)$$

Second g_m stage: As the channel count is increased, the power burnt in the 1st g_m stage is effectively shared among the channels. But the 2nd g_m stage that was used to create a virtual ground will be replicated for every channel. Hence if N is increased indefinitely and if all other dissipation is ignored, the power per channel will asymptotically approach the power burnt in the 2nd stage. The 2nd stage power depends on the value of g_{m2}, which needs to be high enough to ensure good phase margin. This power can be kept low, because for a reasonably high closed-loop gain, g_{m2} needs to be a small fraction of g_m, as shown in equation 4.18. In this design, g_{m2} is 13 times smaller than g_m, as shown in equation 4.19.

Factoring the overheads discussed above, the optimum number of channels can be estimated. The single channel power P_1 can be written as

$$P_1 = I_{S1} + I_{S2} + I_{mix} \qquad (4.30)$$

where I_{S1} is the current consumed in the 1st stage, I_{S2} is the current consumed in the 2nd stage and I_{mix} is the current consumed by the mixers (per channel). For an N-channel implementation, while maintaining the same noise power as the single-channel implementation, the power consumed per channel (P_{NCH}) is given by

$$P_{NCH} = L^2 \left(\frac{I_{S1}}{N} + I_{S2} \right) + I_{mix} \cdot N^2 \qquad (4.31)$$

where L denotes the factor by which noise leakage causes the noise gain to increase, and is given by

$$L = 1 + \frac{N-1}{K} \tag{4.32}$$

Thus, the factor by which power is reduced can be written as

$$\frac{P_1}{P_{NCH}} = \frac{I_{S1} + I_{S2} + I_{mix}}{L^2\left(\dfrac{I_{S1}}{N} + I_{S2}\right) + I_{mix} \cdot N^2} \tag{4.33}$$
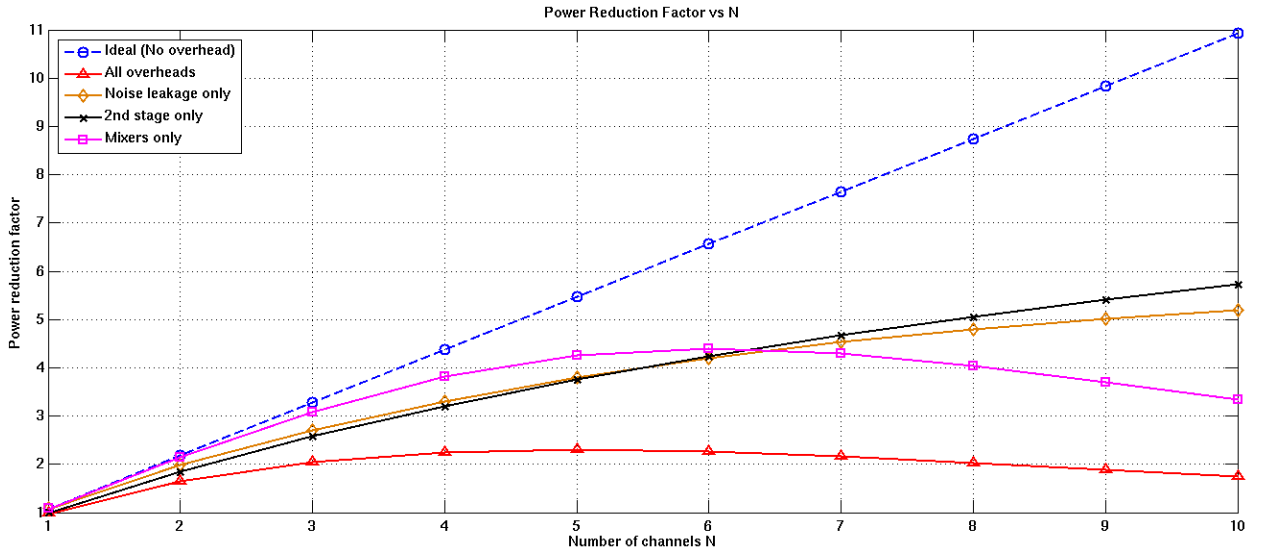

Fig 4.21: Power reduction vs Number of channels

Fig 4.21 shows the plot of P_1/P_{NCH} vs N. The factor K is assumed to be 20, since the clock frequencies used are spaced apart by 100 kHz, and the signal bandwidth is about 5 kHz. As seen from Fig 4.21, the noise leakage and 2nd stage power are the dominant overheads at low channel counts (N<6). But as the channel count increases beyond 6, the mixer power becomes the dominant overhead. The 2nd stage power has a minimum bound which is derived from the minimum transconductance of the 2nd stage for stability. This power can be lowered if the closed-loop gain is increased, but at the cost of increased area and lower linearity. Noise leakage

can be reduced if the factor K is increased, which would require the carriers to be spaced further apart. But spacing the carriers further apart would lead to higher clock frequencies, which would increase the mixer power and the clock generation power. The mixer power can be reduced if they are sized smaller so that the driving capacitance of the mixer switches are reduced, but at the cost of increase thermal noise contribution. Thus, the design parameters for the multi-channel amplifier are inter-dependent, and a more optimized solution can be found by considering all the parameters together.

When all overheads are considered in Fig 4.21, the optimum channel count is about 5, which would reduce the power per channel by a factor of 2.3. But a 3-channel implementation would reduce the power by a factor of 2, which is close to the optimum power reduction factor. Hence, a 3-channel implementation is simulated (for simplicity) to verify the functionality and performance of the FDM amplifier. Note that the power reduction could have been artificially increased if more power was burnt in the 1st stage g_m (thus realizing a lower noise floor) and increasing the closed-loop gain, so that g_{m2} need not be increased. This would work because the overheads would become a smaller fraction of the 1st stage power. Thus the power reduction would be more pronounced if the 1st stage burnt more power. But we would end up burning more power than necessary for the given specifications ($4\mu V_{rms}$ noise). So the 1st stage power is not kept higher than what is necessary to meet the noise requirements,

CHAPTER 5

Results and Conclusions

5.1   Simulated Results

A 65-nm CMOS technology was used to design the amplifiers. The single channel amplifier was simulated to verify the gain, bandwidth and noise specifications for which it was designed. The operating conditions of the transistors of the amplifier in Fig 4.16 are shown in Table 5.1
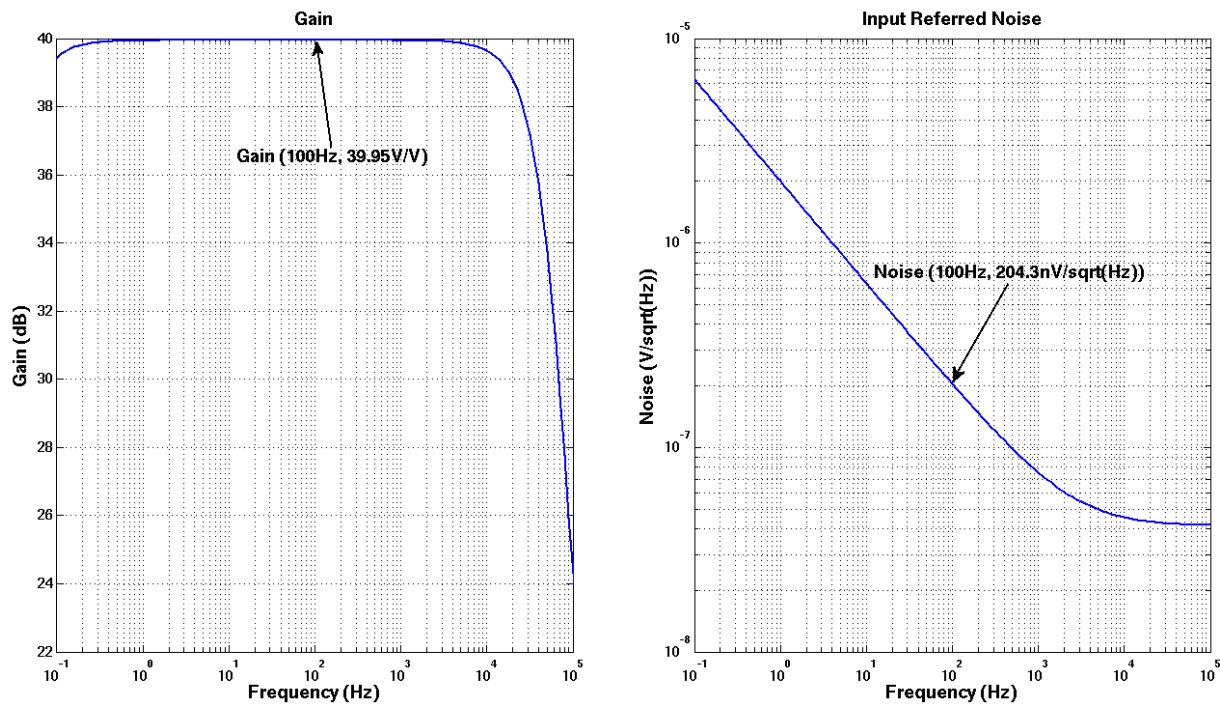

Fig 5.1: Gain and input-referred noise of the amplifier in Fig 4.16(a)

The gain and input-referred noise of the amplifier when used in closed loop is shown in Fig 5.1. The noise is dominated by flicker noise. The flicker noise corner is 2.5 kHz as shown in Fig 5.2, which is well into the signal band.

Table 5.1: Operating points of transistors used in inverter based amplifier in Fig 4.18

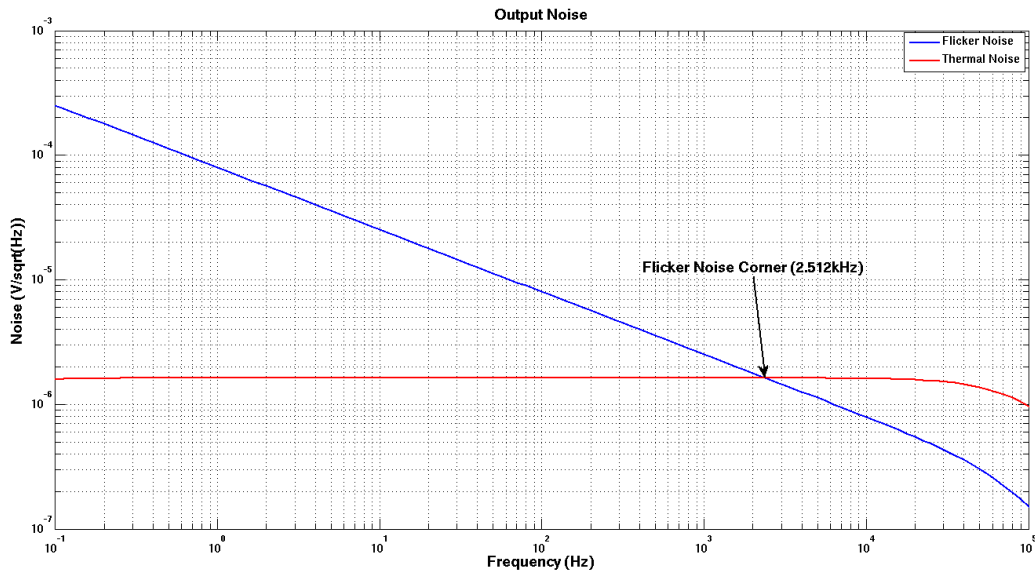| Device	Bias	g_m	g_m/Id
$M_{1,2}$	500 nA	13.77 µA/V	27.5
$M_{3,4}$	500 nA	12.86 µA/V	25.7
$M_{5,6}$	40 nA	1 µA/V	25
$M_{7,8}$	40 nA	1.07 µA/V	26.8
M_9	1 µA	20 µA/V	20
M_{10}	1 µA	19 µA/V	19
M_{11}	80 nA	1.76 µA/V	22
M_{12}	80 nA	1.72 µA/V	21.5

Fig 5.2: Output thermal and flicker noise of amplifier in Fig 4.16(a)

This amplifier is used in the chopper architecture shown in Fig 2.1(b), with a chopper frequency of 200 kHz. The corresponding gain and input-referred noise are shown in Fig 5.3. Note that the flicker noise corner has moved to 1.6 Hz, which is more than 3 decades lower than the previous amplifier. This is the single-channel amplifier, against which the multi-channel amplifier is to be compared with.
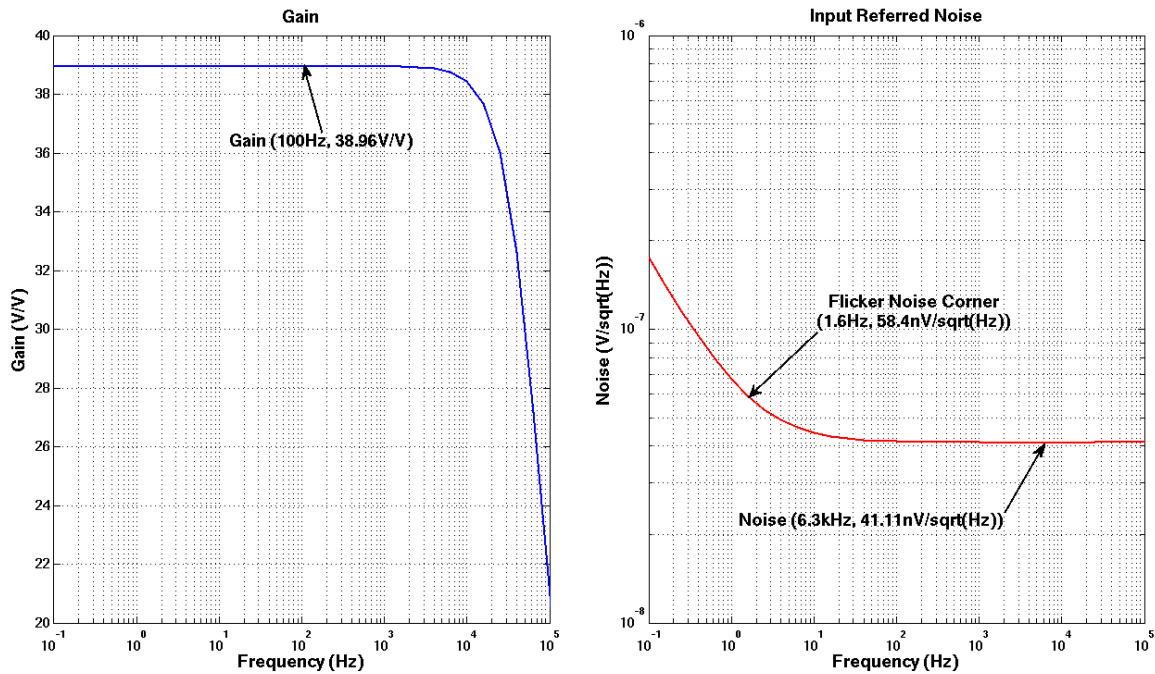
45

Fig 5.3: Gain and input-referred noise of a single channel chopped amplifier

The feedback impedances were simulated separately to verify their filtering properties. The impedance of the input arms for the 3-channel case is shown in Fig 5.4.
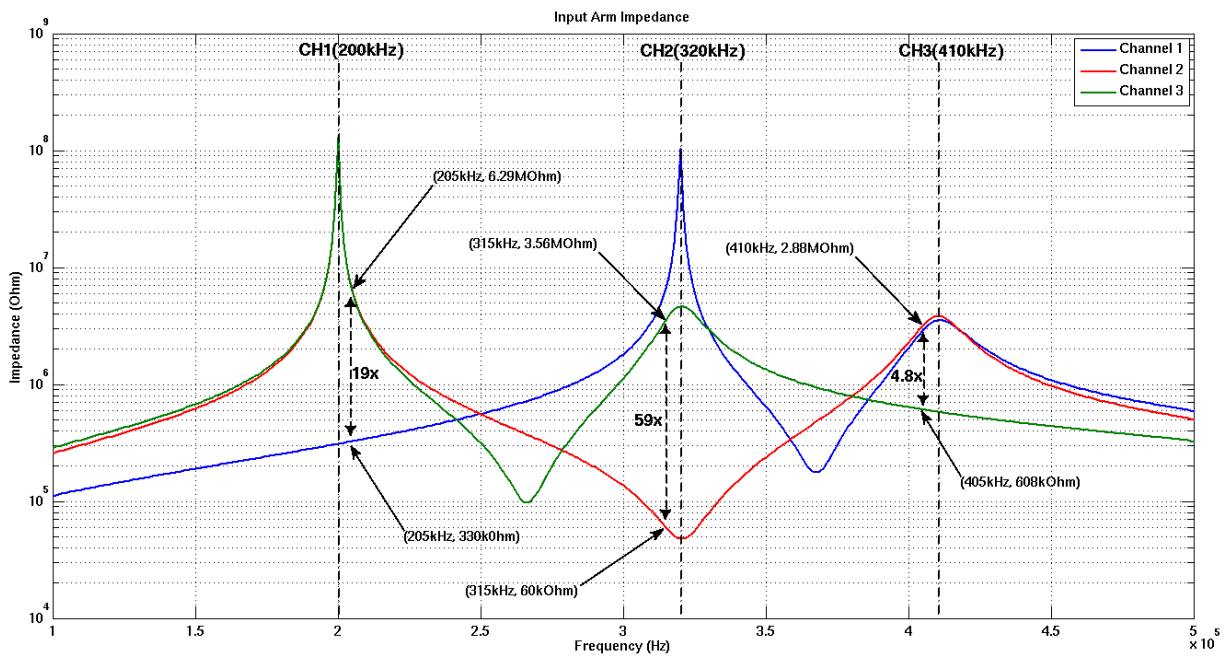
Fig 5.4: Impedance characteristics of input arms used in the feedback network

46

The multi-channel amplifier shown in Fig 4.14 was simulated with clock frequencies of 200 kHz, 320 kHz and 410 kHz. The results are shown below.

Fig 5.5: Gain from channel 1 input to all outputs

Fig 5.6: Gain from channel 2 input to all outputs

47

Fig 5.7: Gain from channel 3 input to all outputs

Fig 5.8: Input-referred noise for channel 1

Fig 5.9: Input-referred noise for channel 2
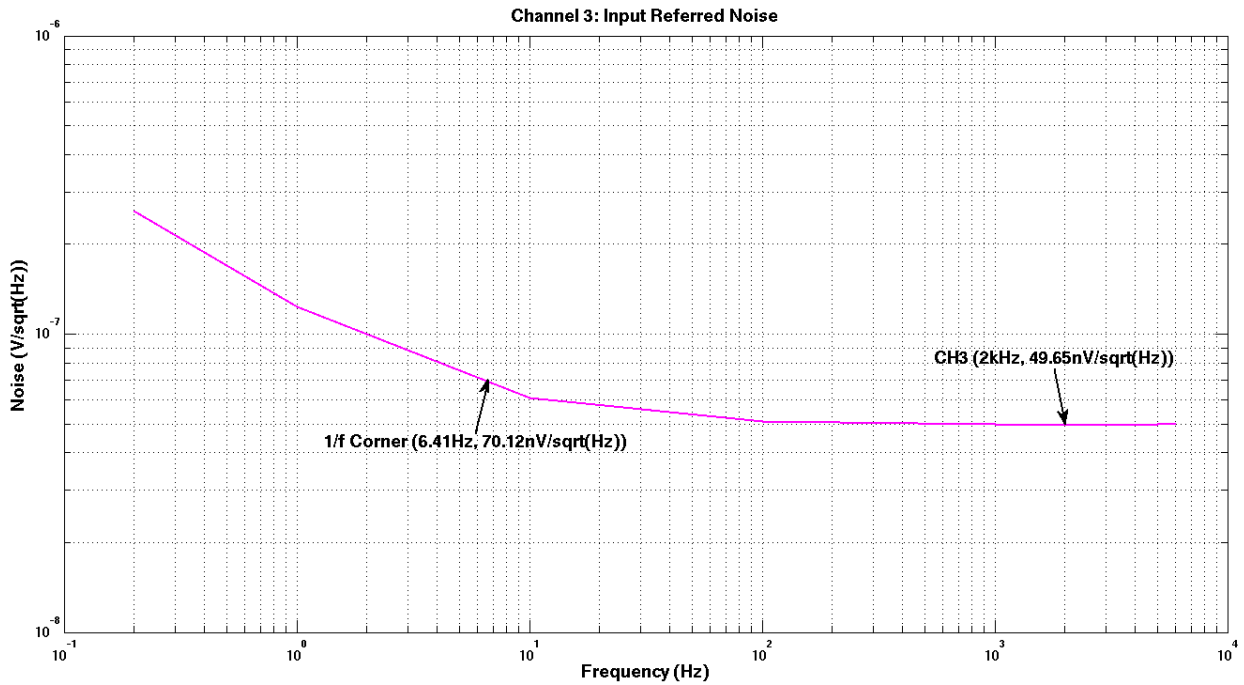
Fig 5.10: Input-referred noise for channel 3

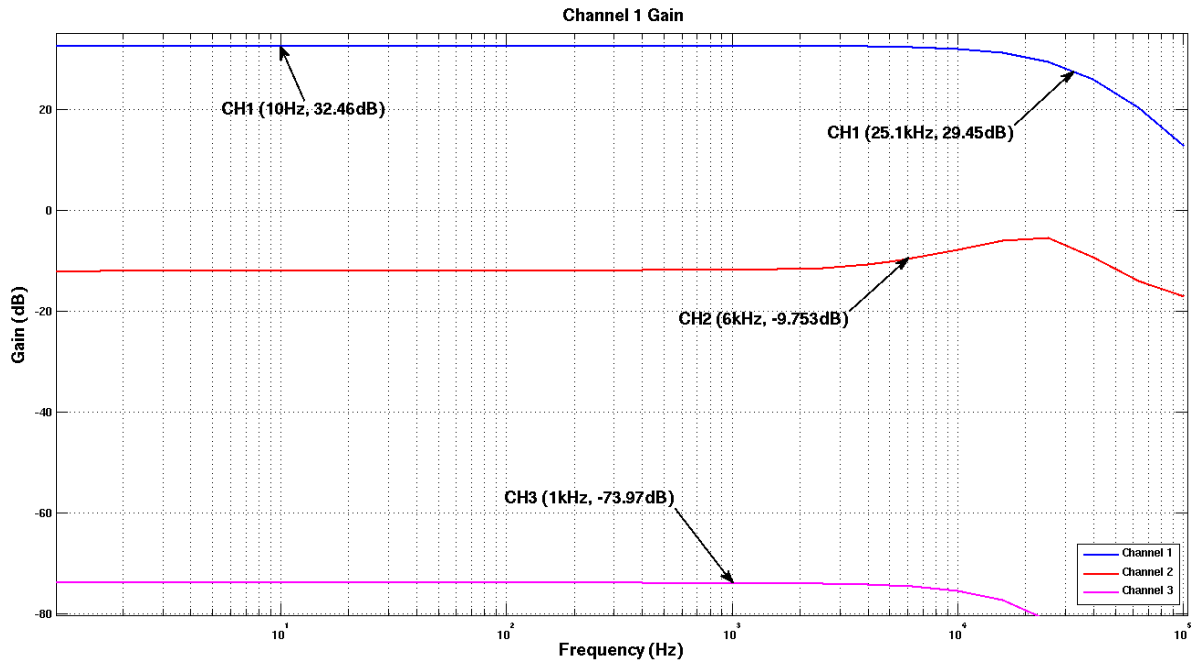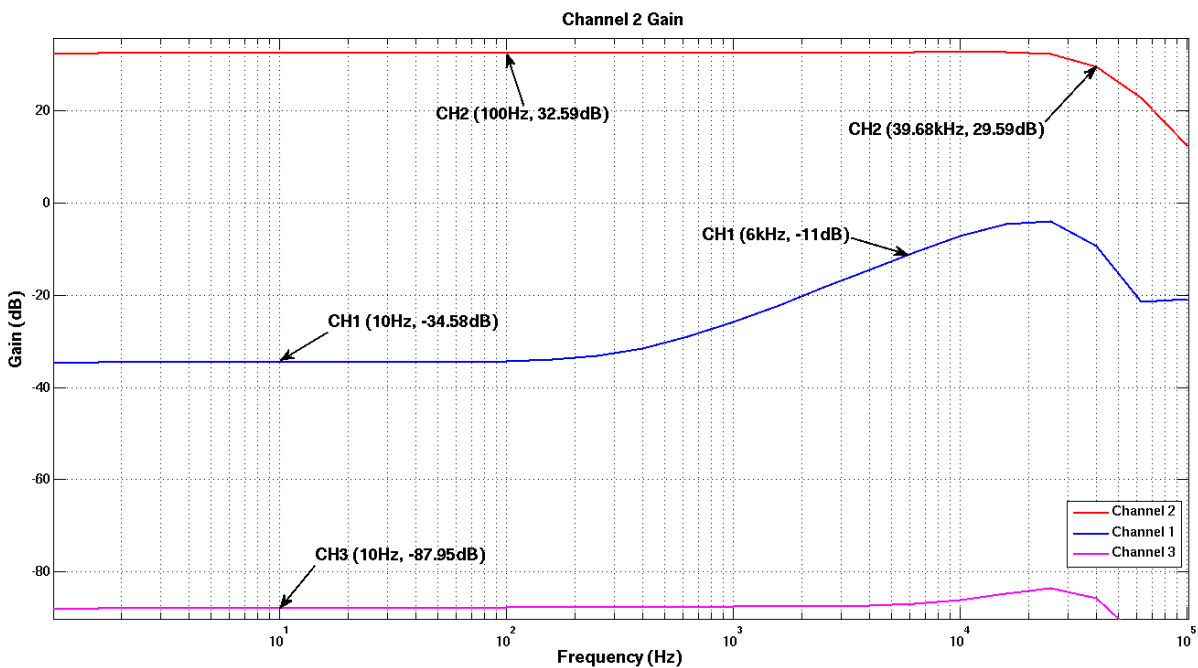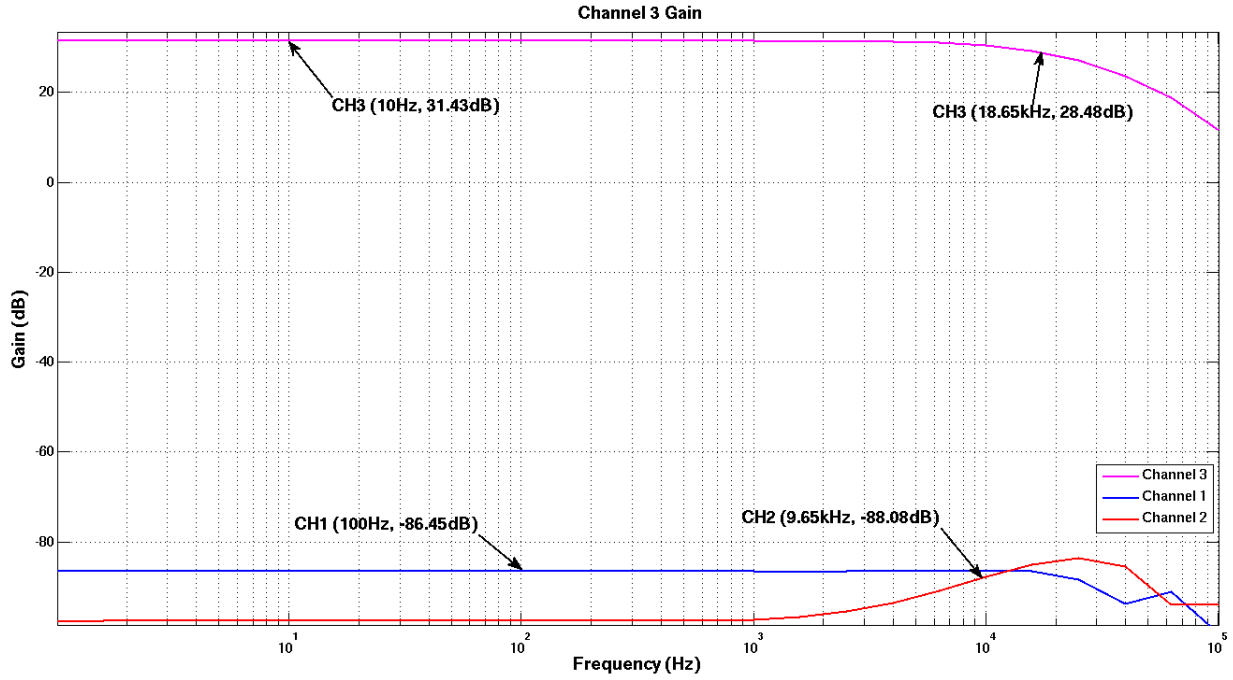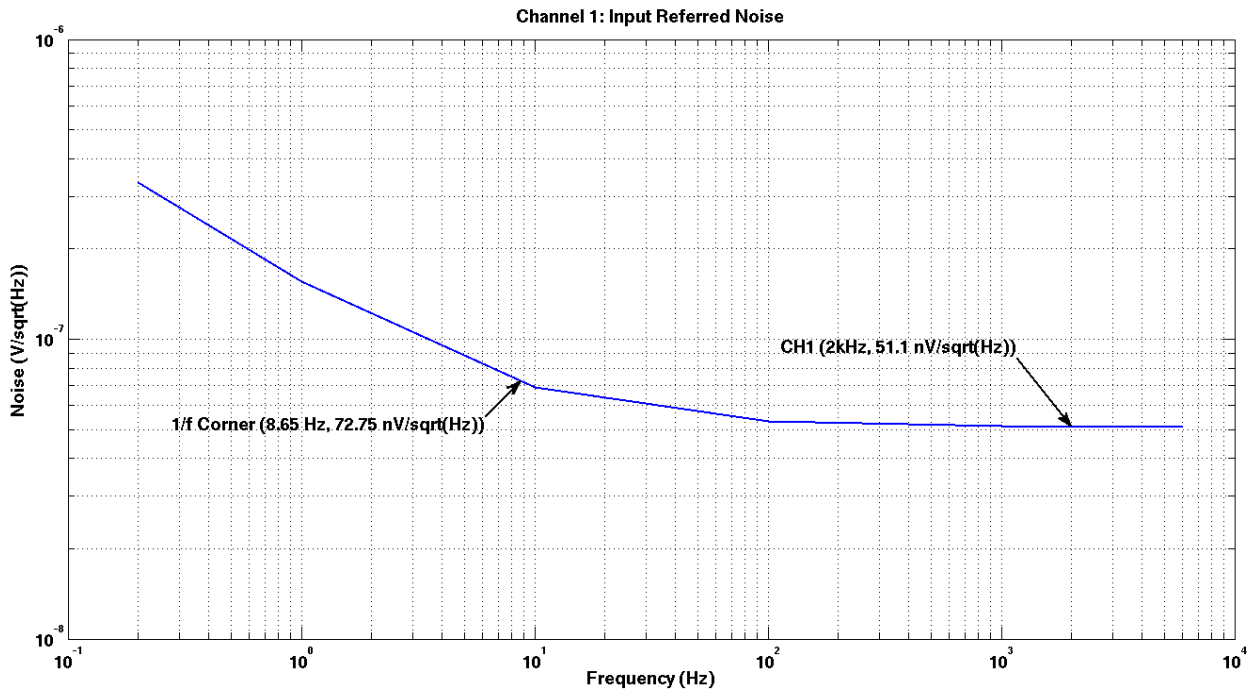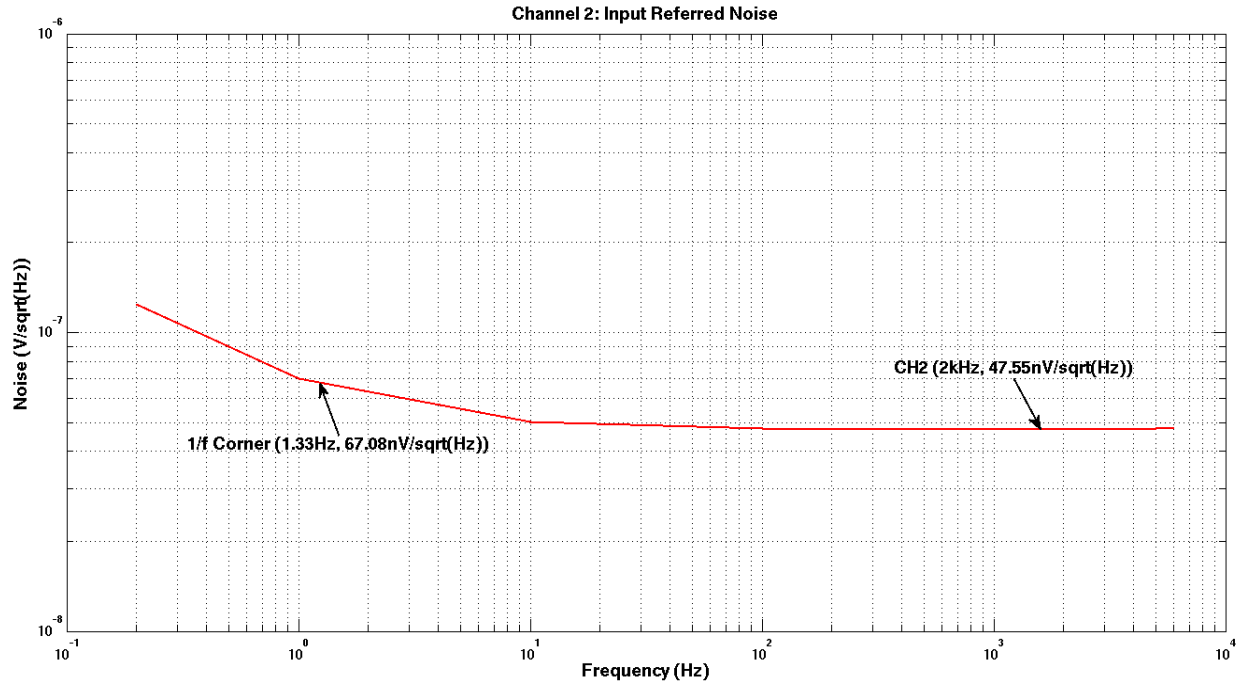The worst-case input-referred noise PSD is lower than 51.1 nV/√Hz, as shown in Fig 5.8. This is higher than the single-channel amplifier (Fig 5.3) because of leakage, as discussed in the previous section. Note that if the loading effect was not taken care of, then the input-referred noise would have been about 125 nV/√Hz. The simulated results for the single-channel FDM topology shown in Fig 4.15 are shown in Fig 5.11. The summing amplifier is designed for a net gain of 10, and with a bandwidth of 12 kHz. As seen in Fig 5.12, the input-referred noise PSD has reduced by a factor of 3, i.e. from $(51 \text{ nV}/\sqrt{Hz})^2$ to $(29 \text{ nV}/\sqrt{Hz})^2$. The net output gain is the gain of the 1st stage multiplied by the gain of the summing amplifier, which is 400 (or 52 dB).

Fig 5.11: Gain of the FDM amplifier architecture in Fig 4.15

A summary of the simulated results is shown in Table 5.2, and is compared with the current state-of-the-art designs. The amplifier consumes 1.1 µA in the 1st stage and 0.1 µA in each of the second stages. These include the CMFB loops. The total driving current for the mixers was found to be 65nA.
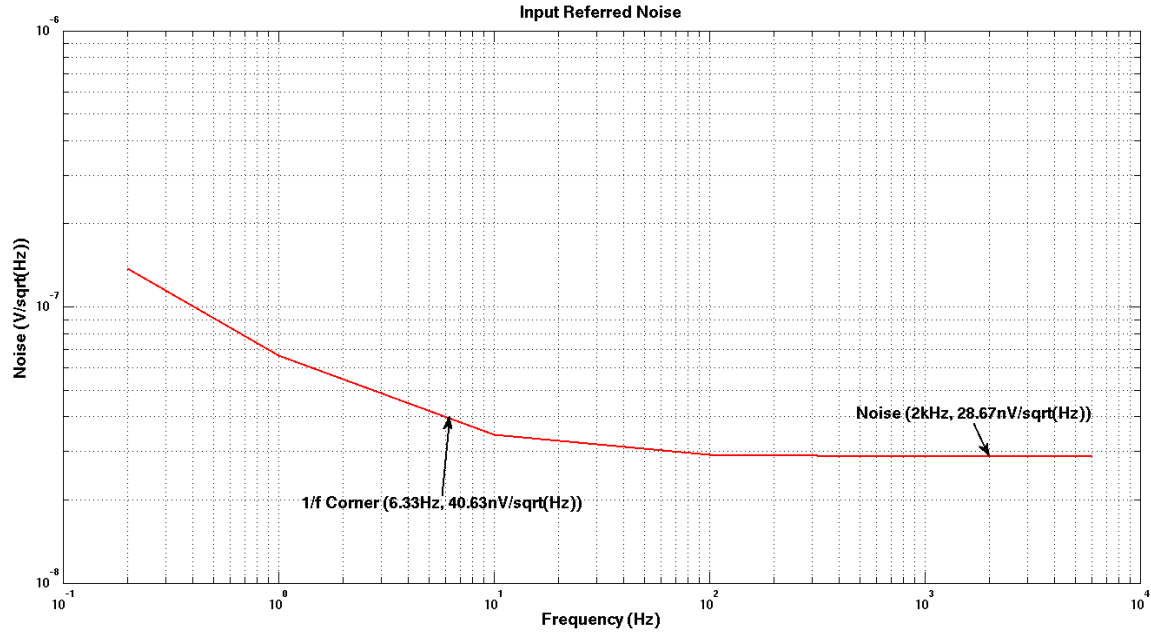
Fig 5.12: Input-referred noise of the FDM amplifier architecture in Fig 4.15

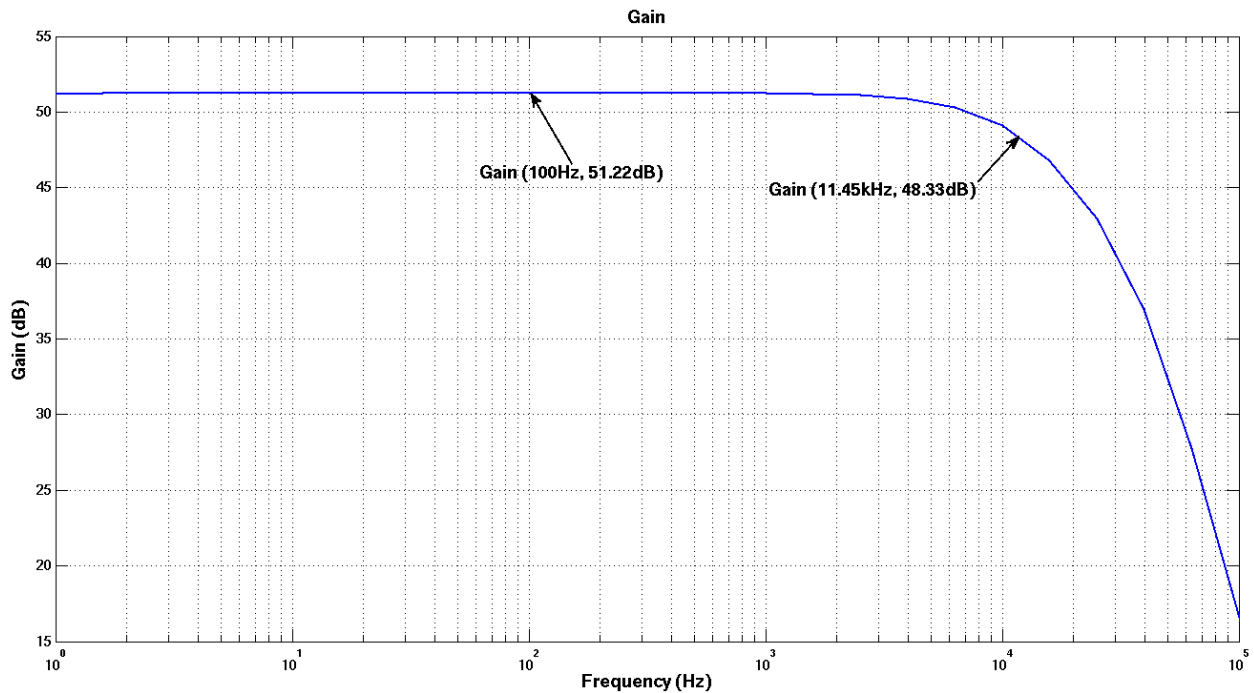The input-referred noise integrated from 0.2 Hz to 6 kHz is 4 μV_{rms}, which meets the requirements (Table 1.1). As seen from the Table 5.2, for similar noise performance, the power consumed per channel in this design is significantly lower than existing designs. The theoretical lower limit for the NEF of a single-channel design was calculated to be 2.02. The NEF that has been achieved here is 1.4, which is due to the departure from conventional single-channel designs. Although the technology node used here is 65 nm, the minimum channel length of the devices used is 0.5 μm, except for the mixer switches which use minimum channel length. There is no advantage in working at a lower technology node for such a low-frequency application.

In Table 5.2, the normalized current is the factor by which the current consumption of a particular design exceeds the current consumption of this work, while keeping the noise PSD and the integrating bandwidth the same. Similarly, the normalized power is the factor by which the power consumption of a particular design exceeds the power consumption of this work, while

51

keeping the bandwidth and the noise PSD the same. The Power Efficiency Factor (PEF) is defined as NEF^2 x V_{DD}. By comparing with the PEF of [15] which is the practical lower limit for single-channel implementations, it can be seen that **the design presented in this work achieves a power consumption that is 3.6 times lower than the best single-channel design. From Fig 5.13, it is observed that the design presented in this work achieves the lowest NEF (and PEF) and at the same time achieves the lowest current consumption.** It should be noted that the results of this work are simulated results and have not been verified through a prototype yet, whereas the designs used to compare with have been fabricated and tested, so the comparison is unfair. But since the margin between this work and the best single-channel design is more than a factor of 3, it is highly probable that the fabricated prototype will still give better performance than existing designs.

Table 5.2: Comparison of performance of neural recording amplifiers

Specification	[11]	[12]	[13]	[15]	[18]	This work
Supply	2.8 V	3.0 V	1.8 − 3.3 V	1 V	0.5 V	1.2 V
Current	2.7 µA	5.2 µA	1.0 µA	12.1 µA	9.58 µA	0.5 µA
Gain	40.85 dB	71.5 dB	41/50.5 dB	40 dB	32 dB	32 dB
Bandwidth	45 Hz − 5.32 kHz	29 Hz − 22 kHz	0.05 Hz − 180 Hz	10.5 kHz	10 kHz	6 kHz
Inp. ref. Noise	3.06	4 µV$_{rms}$	0.95 µV$_{rms}$	2 µV$_{rms}$	6.52 µV$_{rms}$	4 µV$_{rms}$
NEF	2.67	2.45	4.6, 5.4	2.9	7.97	1.4
PEF=NEF^2 x V$_{DD}$	19.96	18	38.1, 52.5	8.41	31.76	2.35
Normalized Current	3.64	3.06	10.8, 14.88	4.29	32.4	1
Normalized Power	8.5	7.66	16.2, 22.3	3.6	13.5	1
Technology	0.5 µm	0.35 µm	0.8 µm	0.13 µm	65 nm	65 nm
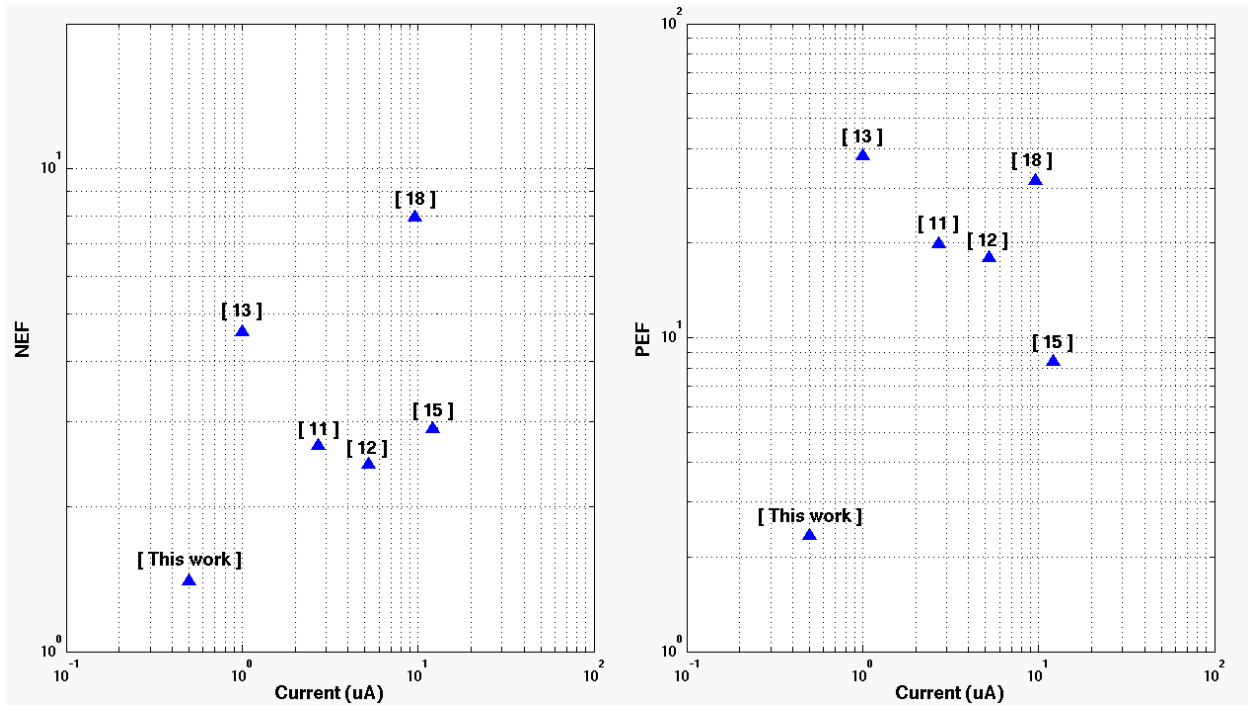
Fig 5.13: Comparison of NEF and PEF

5.2 Conclusions

A fundamentally different amplifier architecture has been presented, which uses frequency division multiplexing to share the 1st stage of the amplifier, thus reducing the effective power consumed per channel. The NEF achieved per channel is 1.4, which is lower than the theoretical limit of 2 for a single-channel design. Comparisons with existing state-of-the-art designs show that this amplifier achieves atleast a factor of reduction of 3.6 in power. Thus this amplifier can potentially achieve better performance than all previous designs presented to date. Hence this amplifier architecture could be used in a neural recording system, where achieving low system power is of primary importance.

CHAPTER 6

Future Work

6.1   Problems with FDM Architecture

Input impedance: The input impedance of the FDM architecture is low as compared to a conventional capacitive feedback amplifier of Fig 2.1(a) due to the action of the mixers present in the input arms of the feedback network. The equivalent impedance seen at the input of the FDM amplifier in Fig 4.14 is

$$Z_{in} \approx \frac{1}{2\pi(\Delta f)C_1} = 0.4M\Omega \tag{6.1}$$

where Δf is the difference in frequency between the channel's carrier frequency and the nearest adjacent carrier frequency. The value of Δf used in this design is 100 kHz, and the value of C_1 is 4pF. This impedance would give a high-pass corner of 40 Hz (considering a 10 nF electrode capacitance), but the required corner frequency is 1 Hz. Thus, the input impedance has to be increased, and one possible way to increase it is discussed in [20]. As a consequence of the lowered input impedance, the amplifier input requires a current proportional to the applied input voltage. Instead of the electrode providing this current (thus the low impedance), this current can be provided by a feedback loop that measures the output voltage of the amplifier, and feeds back a proportional current. This would be a positive feedback loop (labeled PFL in Fig 6.1), and by providing all the required current, the electrode need not provide any current to the amplifier input, thus in principle increasing the effective input impedance to infinity. The concept is shown

in Fig 6.1. The PFL provides the currents Ipf1 and Ipf2, which would equal the input current that would have been provided by the electrode in the absence of the PFL.
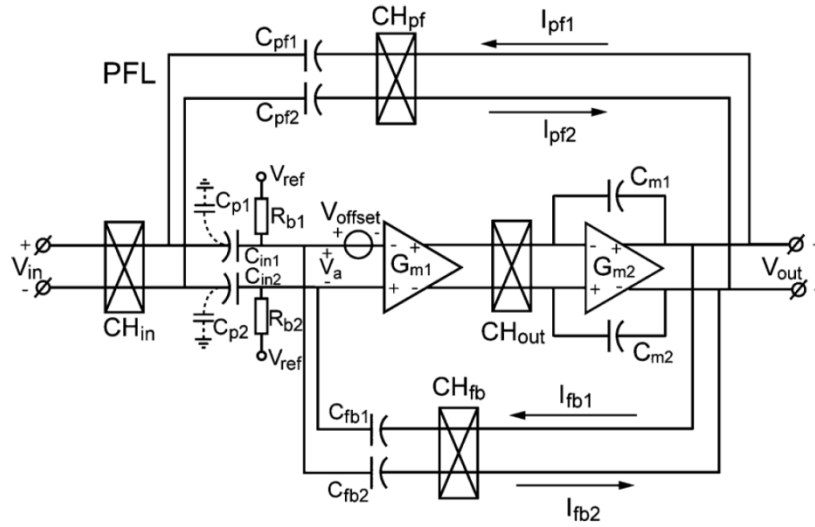

Fig 6.1: Positive feedback loop for boosting input impedance (from [20])

DC offset rejection: The neural signals of interest that are recorded by the electrodes are also accompanied by a large differential DC voltage. This offset can be as high as 100 mV. In the conventional inverting amplifier in Fig 2.1(a), the high-pass nature of the amplifier ensured significant rejection of this offset. But with the multi-channel FDM amplifier, the high-pass nature of the amplifier is lost, because the input is up-modulated to a carrier frequency. This converts the transfer function from a band-pass nature to a low-pass nature. To ensure sufficient offset rejection, a feedback loop can be used to suppress any output voltage below the required high-pass cutoff (1Hz). This is discussed in [20], and is shown in Fig 6.2(a). The bandwidth of the feedback loop has to be high enough only to suppress the signal content below 1 Hz. This requires a very low bandwidth feedback loop, and the ideas discussed in [21] can be used to realize very low bandwidth switched-capacitor low-pass filters using reasonable capacitor sizes, as shown in Fig 6.2(b).

(a)   (b)
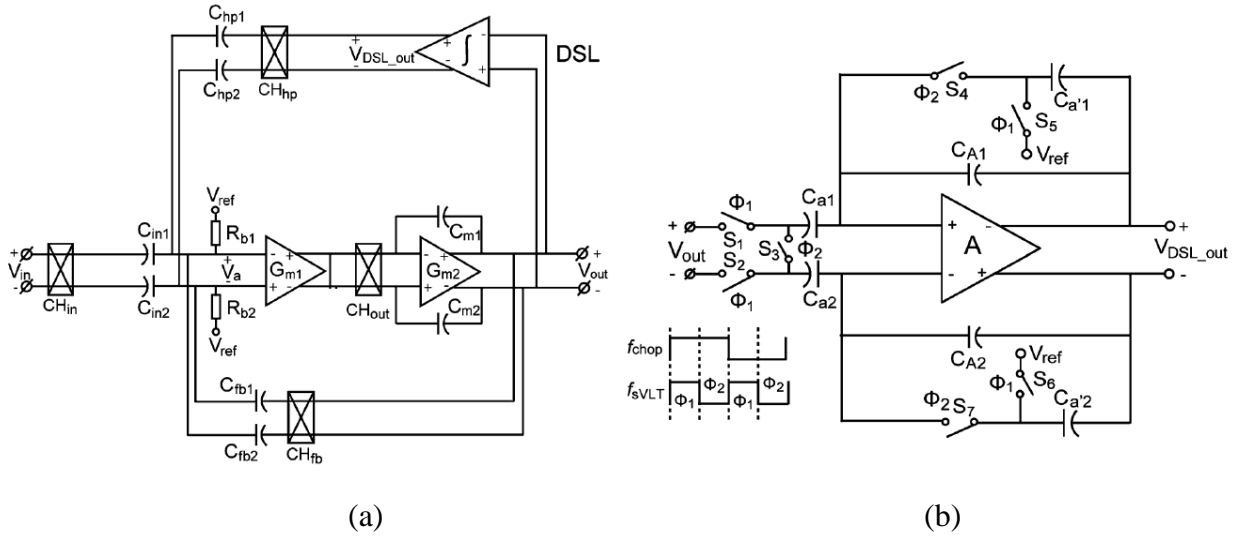
Fig 6.2 (a) Feedback loop to create a high-pass corner [20], (b) Low bandwidth switch-cap LPF

Chopper ripples: In the chopping amplifier, the chopping technique is used to reject amplifier DC offsets and low frequency noise. But the chopping action causes ripples in the output, due to the clocks coupling to intermediate nodes through the gate capacitance of the switch. This can cause large output swings at the chopping frequency, which can reduce the available output swing. To suppress these ripples, a feedback loop can be used to reject signals around the chopping frequency. This is discussed in [20], and is shown in Fig 6.3.

Fig 6.3: Chopper ripple rejection from [20]

56

The PFL requires negligible power, since the mixers used in the PFL loop can be minimum size. The ripple-rejection loop used in [20] consumed about 100nA, and the DC offset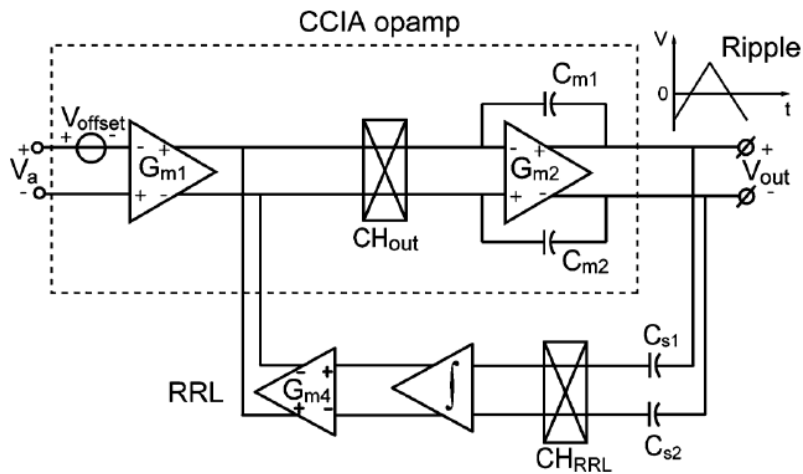 rejection loop consumed 300nA. The amplifiers used in the DC rejection loop could be designed using lower currents. Since these feedback loops have very low bandwidth and are implemented mainly using switched-capacitor filters, they lend themselves to be shared among different channels. Thus their power overhead per channel can be kept low enough to ensure that the amplifier is the dominant power consumer.

Clock Generation: The power required to generate the carrier frequency clocks has not been considered here. The power would depend on the jitter requirements of the clocks. After calculating the limits on the clock jitter, oscillators need to be designed and simulated to ensure on-chip clock generation. Other aspects, like drift between clocks need to be considered too. To ensure that the multi-channel implementation results in significant power reduction, the power burnt in generating the clocks has to be much lower than the power of the amplifier.

The 2nd stage anti-alias band-pass filter needs to be designed to limit the bandwidth to 6 kHz. The linearity of the amplifier will also be limited by the band-pass filter since the signal swings in the 2nd stage will be much larger than the 1st stage.

57

APPENDIX A

PSD of a White Noise Process Multiplied by a Periodic Wave

Let x be a white noise process (v_n) multiplied with a periodic wave, with fundamental period T. In the following analysis, 2 harmonics of the periodic wave are considered, but the analysis can be generalized by using more terms from the Fourier expansion of the periodic wave.

$$x(t) = v_n(t)\big[a_1 \cos(\omega_1(t+t_0)) + a_2 \cos(\omega_2(t+t_0))\big] \tag{A.1}$$

$$\text{where } t_0 \in [0,T)$$

The auto-correlation of x can be written as:

$$\mathfrak{R}_X(\tau) = E\{x(t) \cdot x(t+\tau)\} \tag{A.2}$$

or

$$\mathfrak{R}_X(\tau) = E\{v_n(t) \cdot v_n(t+\tau) \cdot \big[a_1 \cos(\omega_1(t+t_0)) + a_2 \cos(\omega_2(t+t_0))\big] \cdot \big[a_1 \cos(\omega_1(t+t_0+\tau)) + a_2 \cos(\omega_2(t+t_0+\tau))\big]\}$$

Thus,

$$\mathfrak{R}_X(\tau) = \mathfrak{R}_N(\tau)E\left\{\frac{a_1^2}{2}\big[\cos(\omega_1\tau) + \cos(2\omega_1(t+t_0)+\omega_1\tau)\big] + \frac{a_2^2}{2}\big[\cos(\omega_2\tau) + \cos(2\omega_2(t+t_0)+\omega_2\tau)\big]\right\}$$

$$+ \mathfrak{R}_N(\tau)E\left\{\frac{a_1a_2}{2}\big[\cos((\omega_1+\omega_2)(t+t_0)+\omega_2\tau) + \cos((\omega_1-\omega_2)(t+t_0)-\omega_2\tau)\big]\right\}$$

$$+ \mathfrak{R}_N(\tau)E\left\{\frac{a_1a_2}{2}\big[\cos((\omega_1+\omega_2)(t+t_0)+\omega_1\tau) + \cos((\omega_1-\omega_2)(t+t_0)-\omega_1\tau)\big]\right\} \tag{A.3}$$

58

If t_0 is uniformly distributed between 0 and T, then $E\{\cos(\omega_i t_0 + \phi)\} = 0$. Here, ω_i is an integer multiple of the fundamental frequency, and φ is a constant. Using this result, equation A.3 simplifies to

$$\Re_X(\tau) = \frac{\Re_N(\tau)}{2} E\{a_1^2 \cos(\omega_1\tau) + a_2^2 \cos(\omega_2\tau)\} \tag{A.4}$$

The PSD of x(t) is the Fourier transform of the autocorrelation function R_X

$$S_X(f) = \frac{S_N(f)}{2} \otimes \left\{ \frac{a_1^2}{2}[\delta(f - f_1) + \delta(f + f_1)] + \frac{a_2^2}{2}[\delta(f - f_2) + \delta(f + f_2)] \right\} \tag{A.5}$$

Thus, the output PSD is the same as a scaled version of the input PSD, where the scaling factor is equal to the average power of the multiplying periodic wave.
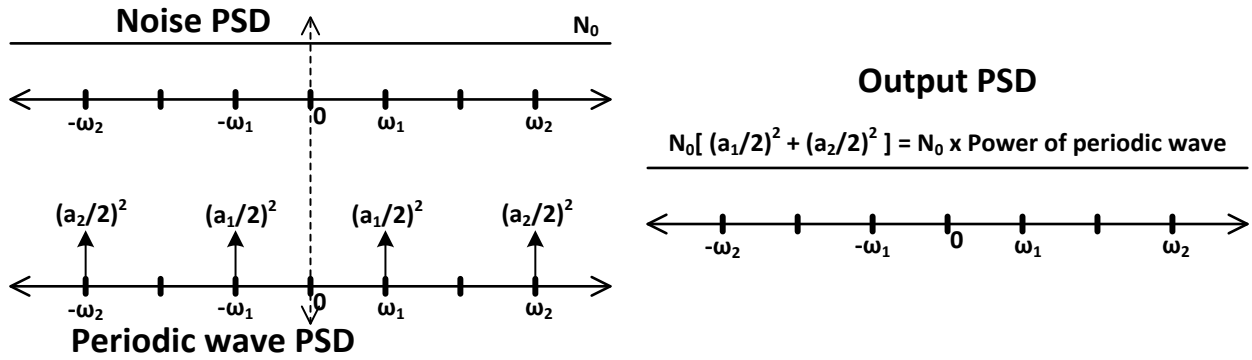

Fig A.1: Input noise PSD, multiplying wave PSD, and the resulting output PSD

REFERENCES

[1] R. R. Harrison, P. T. Watkins, R. Kier, R. Lovejoy, D. J. Black, B. Greger, and F. Solzbacher, "A low-power integrated circuit for a wireless 100-electrode neural recording system," *IEEE J. Solid-State Circuit*, no. 1, pp. 123–133, 2007

[2] A. B. Schwartz, X. T. Cui, D. J. Weber, and D. W. Moran, "Brain controlled interfaces: Movement restoration with neural prosthetics," *Neuron*, vol. 52, pp. 205–220, 2006

[3] J. Mavoori, A. Jackson, C. Diorio, and E. Fetz, "An autonomous implantable computer for neural recording and simulation in unrestrained primates," *J. Neurosci. Meth.*, vol. 148, pp. 71–77, 2005

[4] A. C. Metting van Rijn, A. Peper, and C. A. Grimbergen, "High-quality recording of bioelectric events," *Med. Biol. Eng. Comput.*, vol. 29, pp. 1035–1044, 1986

[5] K. Najafi and K. D. Wise, "An implantable multielectrode array with on-chip signal processing," *IEEE J. Solid-State Circuits*, vol. SC-21, pp. 1035–1044, Dec. 1986

[6] Najafi, K.;,"Solid-state microsensors for cortical nerve recordings," *Engineering in Medicine and Biology Magazine, IEEE* , vol.13, no.3, pp.375-387, June-July 1994

[7] K. Guillory and R. A. Normann, "A 100-channel system for real time detection and storage of extracellular spike waveforms," *J. Neurosci. Meth.*, vol. 91, pp. 21–29, 1999

[8] R. Harrison and C. Charles, "A low-power low-noise CMOS amplifier for neural recording applications," *IEEE J. Solid-State Circuits*, vol. 38, no. 5, pp. 958–965, Jun. 2003

[9] T. M. Seese, H. Harasaki, G. M. Saidel, and C. R. Davies, "Characterization of tissue morphology, angiogenesis, and temperature in the adaptive response of muscle tissue to chronic heating," *Lab. Invest.*, vol. 78, no. 12, pp. 1553–1562, 1998

[10] V. Karkare, S. Gibson, C.-H. Yang, H. Chen, D. Marković, "A 75µW, 16-Channel Neural Spike-Sorting Processor with Unsupervised Clustering," in *Proc. Int. Symposium on VLSI Circuits* (VLSI'11), June 2011, pp. 252-253

[11] Wattanapanitch, W.; Fee, M.; Sarpeshkar, R.; , "An Energy-Efficient Micropower Neural Recording Amplifier," *Biomedical Circuits and Systems, IEEE Transactions on* , vol.1, no.2, pp.136-147, June 2007

[12] Borghi, T.; Bonfanti, A.; Gusmeroli, R.; Zambra, G.; Spinelli, A. S.; , "A power-efficient analog integrated circuit for amplification and detection of neural signals," *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE* , vol., no., pp.4911-4915, 20-25 Aug. 2008

[13] Denison, T.; Consoer, K.; Santa, W.; Avestruz, A.-T.; Cooley, J.; Kelly, A.; , "A 2 µW 100 nV/rtHz Chopper-Stabilized Instrumentation Amplifier for Chronic Measurement of Neural Field Potentials," *Solid-State Circuits, IEEE Journal of* , vol.42, no.12, pp.2934-2945, Dec. 2007

[14] Enz, C.C.; Temes, G.C.; , "Circuit techniques for reducing the effects of op-amp imperfections: autozeroing, correlated double sampling, and chopper stabilization," *Proceedings of the IEEE* , vol.84, no.11, pp.1584-1614, Nov 1996

[15] Fan Zhang; Holleman, J.; Otis, B.P.; , "Design of Ultra-Low Power Biopotential Amplifiers for Biosignal Acquisition Applications," *Biomedical Circuits and Systems, IEEE Transactions on* , vol.6, no.4, pp.344-355, Aug. 2012

[16] Majidzadeh, V.; Schmid, A.; Leblebici, Y.; , "Energy Efficient Low-Noise Neural Recording Amplifier With Enhanced Noise Efficiency Factor," *Biomedical Circuits and Systems, IEEE Transactions on* , vol.5, no.3, pp.262-271, June 2011

[17] Johnson, B.; DeTomaso, D.; Molnar, A.; , "A low-power orthogonal current-reuse amplifier for parallel sensing applications," *ESSCIRC, 2010 Proceedings of the* , vol., no., pp.318-321, 14-16 Sept. 2010

[18] Muller, R.; Gambini, S.; Rabaey, J.M.; , "A 0.013 mm^2, 5 µW, DC-Coupled Neural Signal Acquisition IC With 0.5 V Supply," *Solid-State Circuits, IEEE Journal of* , vol.47, no.1, pp.232-243, Jan. 2012

[19] L. E. Franks and I. W. Sandberg, "An alternative approach to the realizations of network functions: N-path filter," *Bell Syst. Tech. J.*, pp. 1321–1350, 1960

[20] Qinwen Fan; Sebastiano, F.; Huijsing, J.H.; Makinwa, K.A.A.; , "A 1.8 μW 60 nV$/\sqrt{Hz}$ Capacitively-Coupled Chopper Instrumentation Amplifier in 65 nm CMOS for Wireless Sensor Nodes," *Solid-State Circuits, IEEE Journal of* , vol.46, no.7, pp.1534-1543, July 2011

[21] K. Nagaraj, "A parasitic-insensitive area-efficient approach to realizing very large time constants in switched-capacitor circuits," *IEEE Trans. Circuits Syst.*, vol. 36, no. 9, pp. 1210–1216, Sep. 1989

[22] Steyaert, M.S.J.; Sansen, W.M.C.; , "A micropower low-noise monolithic instrumentation amplifier for medical purposes," *Solid-State Circuits, IEEE Journal of* , vol.22, no.6, pp. 1163- 1168, Dec 1987