AN ABSTRACT OF THE THESIS OF

Ryan Perigny for the degree of Master of Science in

Electrical and Computer Engineering presented on August 21, 2000.

Title: Area Efficiency Improvement of CMOS Charge Pump Circuits.

Redacted for Privacy

Abstract approved: _____

Un-Ku Moon

In this thesis, the literature relating to charge pump dc-dc converters and their uses is reviewed. Charge pumps are useful in many circuits, including low-voltage circuits, dynamic random access memory circuits, switched-capacitor circuits, EEPROM's and transceivers. The important issues relating to charge pump design are power efficiency, output voltage ripple and area efficiency.

This thesis describes the operation of three types of charge pump circuits. Power efficiency theory of charge pumps is discussed in detail. A method of estimating the output ripple of a charge pump from the size of the capacitors used is described. The optimal distribution of available capacitance for minimizing output ripple or maximizing power efficiency is derived. The tradeoffs between output ripple, power efficiency and total capacitance are discussed. The considerations involved in the design of charge pump circuits are described. A new charge pump circuit that uses two cascoded buffer transistors to improve the area efficiency is proposed. An implementation consisting of one of each of the three types of charge pumps was simulated for a 0.35-micron CMOS process. The simulation results verify the improved area efficiency of the double cascode charge pump.

Area Efficiency Improvement of CMOS Charge Pump Circuits

by

Ryan Perigny

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Completed August 21, 2000
Commencement June 2001

Master of Science thesis of Ryan Perigny presented on August 21, 2000

APPROVED:

Redacted for Privacy

Major Professor, representing Electrical and Computer Engineering

Redacted for Privacy

Chair of the Department of Electrical and Computer Engineering

Redacted for Privacy

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for Privacy

Ryan Perigny, Author

ACKNOWLEDGMENT

I would like to express my sincere and deep appreciation to my major professor and academic advisor, Dr. Un-Ku Moon, for his encouragement and guidance throughout the project. I am honored to have had the opportunity to work under his supervision. Without his help and support, this thesis would not have been possible.

I would like to express my appreciation to Dr. Gabor C. Temes for his helpful advice and enlightening discussions on my work during our research group meetings. I have benefited greatly from his technical expertise and intuition and his high technical and professional standards.

I also would like to thank my minor professor, Dr. Wojtek Kolodziej for serving on my committee on short notice.

I also would like to thank Dr. Thomas Dietterich for taking time out of his busy schedule to serve as the Graduate Council Representative of my committee.

I also extend my gratitude to Dr. John T. Stonick, for his valuable discussions and comments on my work, and for his excellent teaching of communications courses.

I wish to thank all of my colleagues at Oregon State University, especially José Silva for his generous help with all of the CAD tools and Mustafa Keskin for his kind help with my test chip.

I wish to thank all of the faculty and staff at Oregon State University for the excellent teaching and academic environment they have provided me.

I also wish to express my deepest gratitude to my family: to my parents, Ralph and Shirleen Perigny, and to my brother, Matthew, whose dedication, support and encouragement throughout my life have made me who I am today.

Finally, I thank God, for giving me a loving family and friends, and for providing me with the many blessings and opportunities that I have enjoyed. I pray that I will always continue to learn, and that I make use of the knowledge I have acquired for the good of human kind.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

Figure | Page

LIST OF APPENDIX FIGURES

DEDICATION

This thesis is dedicated to my parents.

AREA EFFICIENCY IMPROVEMENT OF CMOS CHARGE PUMP CIRCUITS

1. INTRODUCTION

1.1. BACKGROUND

The charge pump is a dc-dc converting circuit used to obtain a dc voltage higher or lower than the supply voltage or opposite in polarity to the supply voltage. Charge pump circuits use capacitors as energy storage devices. The capacitors are switched in such a way that the desired voltage conversion occurs. Charge pumps are useful in many different types of circuits, including low-voltage circuits, dynamic random access memory circuits, switched-capacitor circuits, EEPROM's and transceivers.

This thesis discusses the important issues that need to be considered in the design of on-chip charge pump circuits. One important issue is the output voltage ripple. For most applications, a low output ripple is desired. If the output ripple is too large, the performance of the circuit that the charge pump is powering is degraded. Another important consideration is power efficiency. Charge pumps with very low power efficiency waste too much power to be desirable for portable applications. Another issue is area efficiency. It is desirable to minimize the chip area taken up by the charge pump circuit because smaller chip areas are less expensive to fabricate.

The goal of this research was to develop techniques to use less chip area to achieve the same output ripple as existing charge pump circuits. Power efficiency issues and design tradeoffs are also discussed.

For low-noise applications, a very low output ripple is desirable to achieve good performance. This thesis describes an existing charge pump circuit that improves the area efficiency of a conventional charge pump, and introduces a new charge pump circuit that further improves the area efficiency. This new charge pump circuit provides a very low output ripple with a smaller amount of capacitance than existing charge pumps, at the cost of a reduced output voltage and a small amount of additional power loss.

1.2. LITERATURE REVIEW

The first widely used voltage boosting circuit was the Cockcroft-Walton voltage multiplier [1]. This circuit, shown in Figure 1.1, uses diodes and serially connected capacitors and can boost to several times the supply voltage. The Cockcroft-Walton charge pump provides efficient multiplication only if the coupling capacitors are much larger than the stray capacitance in the circuit, making it undesirable for use in integrated circuits.

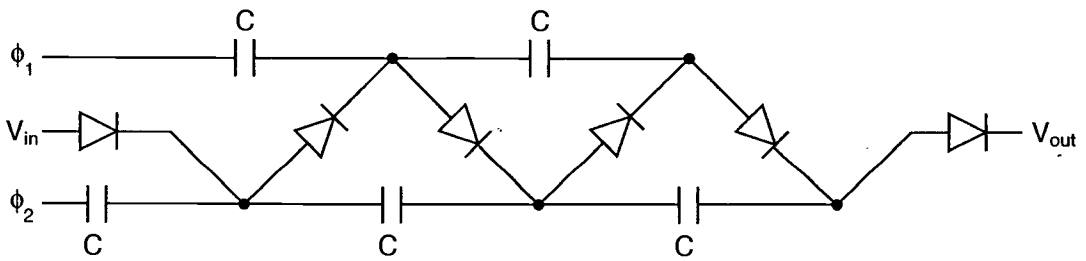

FIGURE 1.1. 5-stage Cockcroft-Walton charge pump [1].

In [2], the Dickson charge pump circuit is presented as an improvement of Cockcroft-Walton circuit. In the Cockcroft-Walton charge pump circuit, the coupling capacitors are connected in series. This results in a higher output impedance as the number of stages increases. In the Dickson charge pump circuit, shown in

Figure 1.2, the coupling capacitors are connected in parallel and must be able to withstand the full output voltage. This results in a lower output impedance as the number of stages increases. Both circuits require the same number of diodes and capacitors and can be shown to be equivalent. The drawback of the Dickson charge pump circuit is that the boosting ratio is degraded by the threshold drops across the diodes. The body effect makes this problem even worse at higher voltages.

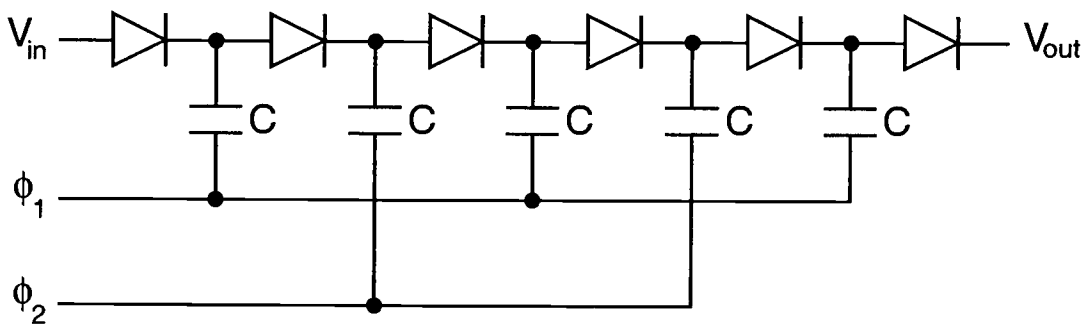

FIGURE 1.2. 5-stage Dickson charge pump [2].

Because of its importance in EEPROM circuits and other applications, some work has been done to describe in detail the operation of the Dickson charge pump circuit. In [3], a new model describing the operation of the Dickson charge pump circuit is presented where the diodes are implemented by MOS transistors. This model includes some deviations from the simple model proposed by Dickson. In [4], a dynamic analysis of the Dickson charge pump with an arbitrary number of stages is presented.

Charge pump circuits are commonly used in DRAM circuits to boost the word line signal to around 1.5 or 1.75 times the supply voltage. In [5], a feedback charge pump circuit that uses cross-coupled NMOS switches are used to achieve a high boost ratio for a low-voltage DRAM word-line driver. This circuit, shown in Figure 1.3, uses two capacitors that are switched in such a way that during every

clock cycle, one capacitor is charged to the supply voltage and the other capacitor is boosted to twice the supply voltage by the clock. The two capacitors reverse roles every clock cycle, causing the voltage at the output to be a square wave that switches between V_{dd} and $2V_{dd}$. Two of these cross-coupled NMOS pairs are used along with another type of charge pump and an inverter to make up the complete boosted voltage generator. An earlier circuit that switches between two networks of capacitors is described in [6] as an "inductance-less dc-dc converter." A lot of work has been done in recent years involving charge pumps for use in DRAM circuits. In [7], a high-efficiency word-line driver for a DRAM is presented. In [8], a charge pump circuit that provides a negative substrate bias for a DRAM is presented.

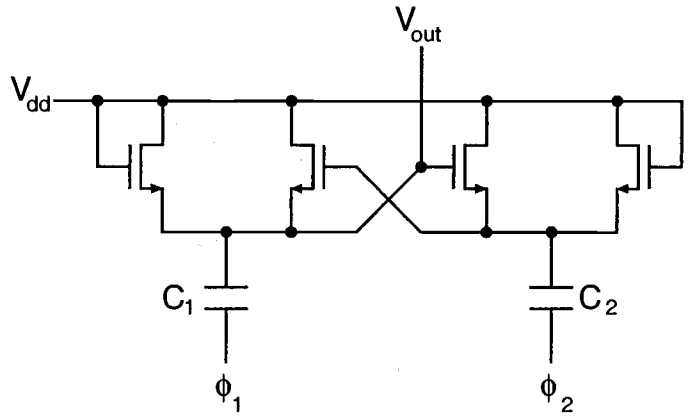

FIGURE 1.3. Cross-coupled NMOS switches from [5].

In [9], the cross-coupled NMOS charge pump introduced in [5] is used to improve the speed of a pipeline A/D converter by boosting the clock drive in order to reduce the on-resistance of transmission gates in the pipeline. This work also utilizes a bias voltage generator to bias the n-well to twice the supply voltage, preventing latchup from occuring during the initial startup transient.

Charge pumps are also widely used in the program circuits and word line drivers in an EEPROM. In [10], a circuit similar to the Dickson charge pump is used

in the program circuit of a flash EEPROM. In [11], a charge pump that uses four clock phases to create a negative high-voltage for the word line driver of an EEPROM is presented. A merged charge pump is presented in [12] that uses cascaded 2-capacitor voltage doublers to provide the read and write voltages for a flash memory. The Dickson charge pump is combined with a voltage doubler to drive a bandgap generator in a flash memory in [13]. In [14], the Dickson charge pump is implemented using diode-connected p-channel transistors in floating wells to eliminate the body effect. In [15], a charge pump for use in low-voltage EEPROM's is presented. This circuit is similar to the Dickson charge pump, but it uses a bootstrapped clock generator to eliminate the threshold drops across the pass transistors.

A different method for eliminating the threshold drops in the Dickson charge pump is presented in [16]. This circuit, shown in Figure 1.4, uses charge transfer switches in addition to the diode-connected transistors to eliminate the threshold drop.

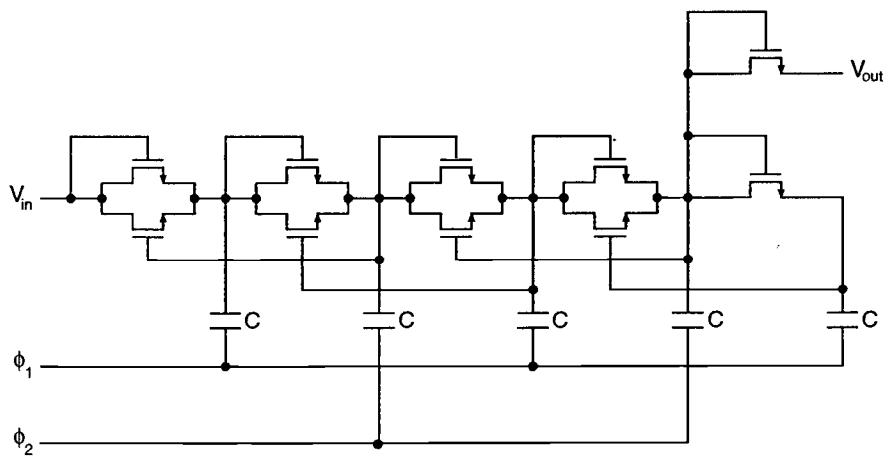

FIGURE 1.4. A method for eliminating the threshold drops across the pass transistors of a Dickson charge pump from [16] (4 stages).

The drawback of this circuit is that the charge transfer switches cannot be completely turned off, leading to a reverse charge sharing phenomenon which reduces

the voltage pumping gain. An improved design, which eliminates this problem by adding pass transistors to the previous circuit, is presented in [17]. This charge pump was designed for use in low-voltage circuits and is shown in Figure 1.5.

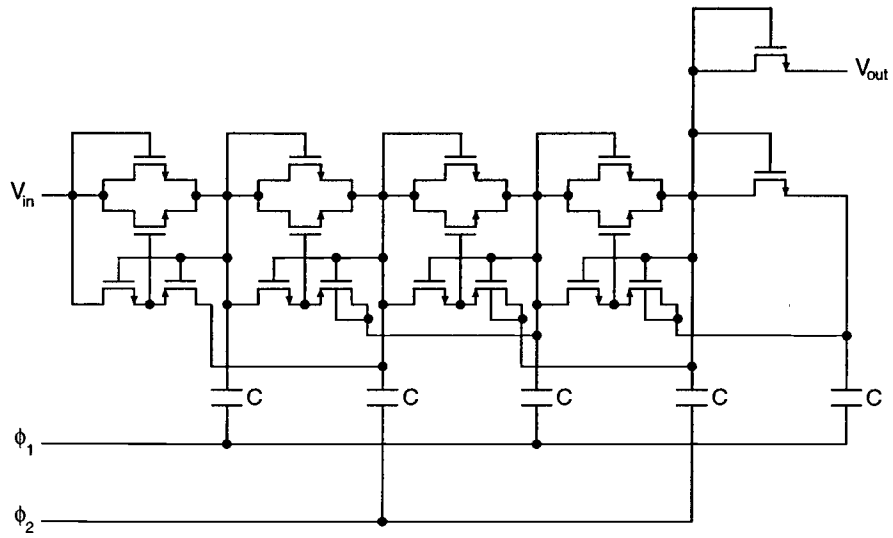

FIGURE 1.5. An improved method for eliminating the threshold drops across the pass transistors of a Dickson charge pump from [17] (4 stages).

Charge pump voltage boosters are also used in transceivers. A charge pump designed for an RS-232 transmitter/receiver is presented in [18].

Some work has been done recently on improving the power efficiency of charge pump circuits. In [19], the power efficiency of a voltage doubling circuit is discussed. A dual charge pump circuit that uses a frequency converter to vary the clock frequency according to the loading is presented. The power efficiency of a conventional 2-capacitor charge pump similar to the one presented in [5] is derived in [20]. Various methods for avoiding latchup during the initial startup transient are also described, and a low-voltage version of this charge pump is presented that uses a level-shifted clock to reduce the on resistance of the p-channel switches.

Some work has been done to reduce the output ripple of a high-voltage charge pump. In [21], the output voltage ripple of a charge pump circuit was reduced by driving the pump capacitors with a voltage-controlled current source, rather than a square-wave voltage source.

In [22], the single cascode charge pump is introduced. This charge pump has an improved area efficiency and is used in a high-performance rail-to-rail input audio amplifier. The area-efficient charge pump was used to bias the differential input pair to about 1V above the supply voltage. This was done so that only one p-channel input pair was needed to achieve a rail-to-rail input range. Previous rail-to-rail designs used an n-channel input pair and a p-channel input pair, which introduces a signal-dependent offset voltage leading to harmonic distortion. Using this new design, the authors were able to achieve -90 dB total harmonic distortion. The double cascode charge pump introduced in Chapter 2 is a modification of this circuit which further improves the area efficiency.

2. AREA EFFICIENCY IMPROVEMENT OF CMOS CHARGE PUMP CIRCUITS

2.1. INTRODUCTION

The charge pump [1] is a dc-dc converting circuit used to obtain a dc voltage higher or lower than the supply voltage, or opposite in polarity to the supply voltage. Charge pumps are widely used in EEPROM's [3], [10]- [15], low-voltage circuits [5], dynamic random access memory circuits [5], [7]- [8], switched-capacitor circuits [9] and transceivers [18].

Three important issues of on-chip charge pump circuits are output voltage ripple, power efficiency and area efficiency. For most applications, a low output ripple is desired. Large output ripple degrades the performance of the circuit that the charge pump is powering. Charge pumps with very low power efficiency cancel the benefit of scaling the supply voltage down and are not desirable for portable applications. Area efficiency is desirable for many applications because smaller chip areas are less expensive to fabricate.

For low-noise applications, a very low output ripple is desirable to achieve good performance. This chapter introduces a new charge pump circuit that provides very low output ripple with a smaller amount of capacitance than existing charge pumps, at the cost of a reduced output voltage and a small amount of additional power loss.

In this chapter, the basic operation of three types of charge pumps will be described. The design considerations and tradeoffs involved in charge pump design will be included. Comparisons between the three charge pumps with regard to area efficiency, power efficiency and output voltage ripple will be discussed.

In section 2.2, the conventional 2-capacitor charge pump will be described and analyzed. In section 2.3, an existing charge pump that uses a buffer transistor to improve the area efficiency will be discussed. In section 2.4, a proposed new charge pump that uses two cascoded buffer transistors to further improve the area efficiency will be described. In section 2.5, the simulation results of an implementation for a 0.35-micron CMOS process will be discussed.

2.2. CONVENTIONAL CHARGE PUMP

A conventional 2-capacitor charge pump that uses cross-coupled NMOS transistors [5] is shown in Figure 2.1. For simplicity, the body connections are not shown in the figure. The n-channel switches have thier body connections grounded. The p-channel switches are all in the same n-well, which is connected to the V_{well} node.

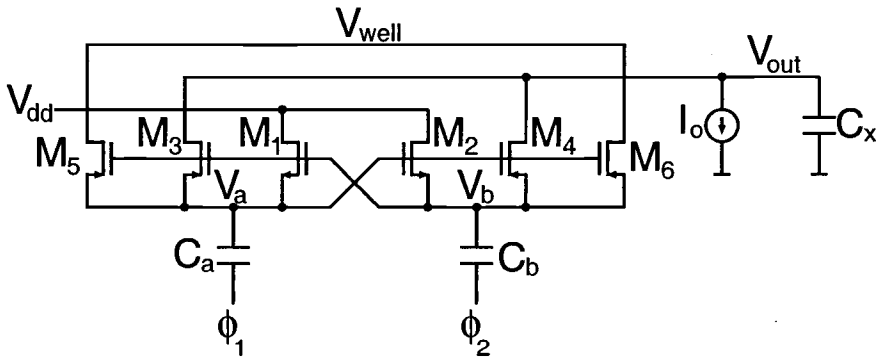

FIGURE 2.1. Conventional 2-capacitor charge pump.

This circuit uses two clock phases and operates as follows: during ϕ_2, switches M_1 and M_4 are turned on and capacitor C_a is charged to V_{dd}. During ϕ_1, switches M_2 and M_3 are turned on and capacitor C_b is charged to V_{dd}. Capacitor C_a, which was charged to V_{dd} during the previous clock phase, is now connected between V_{dd} and the load, lifting the voltage at the load to $2V_{dd}$. During every clock cycle,

one capacitor is being charged to V_{dd} and the other capacitor is providing the load current.

Switches M_5 and M_6 are operated in the same manner as M_3 and M_4, and are used to boost the voltage of the n-well to $2V_{dd}$. This ensures that the n-well-to-substrate pn junction is always reverse biased, preventing latchup from occuring during the initial startup transient. The well voltage is stored by parasitic capacitance from the node V_{well} to ground. Because the parasitic capacitance from node V_{well} to ground is much smaller than C_x, the well charges to $2V_{dd}$ before the output. This way of avoiding latchup is discussed in [20].

2.2.1. *Non-filtering Case*

If there is no load current, the output voltage will remain at $2V_{dd}$. If there is a significant load current and no output capacitor ($C_x = 0$), the voltage at the output will ramp down with a slope of $\frac{I_o}{C_a}$ for half of a clock period before it is boosted to $2V_{dd}$ again. Figure 2.2 shows the output of a conventional charge pump in this situation.

The peak-to-peak voltage ripple at the output, ΔV_{out}, can be calculated as

$$\Delta V_{out} = \frac{I_o}{2 f_{clk}\, C_a}\,, \tag{2.1}$$

where f_{clk} is the clock frequency.

If there is an output capacitor, which is the usual case when a significant current must be provided, the slope of the output is reduced to $\frac{I_o}{C_a+C_x}$ due to the added capacitance at the output node. This reduces the ripple by the same amount. The ripple can now be calculated to be

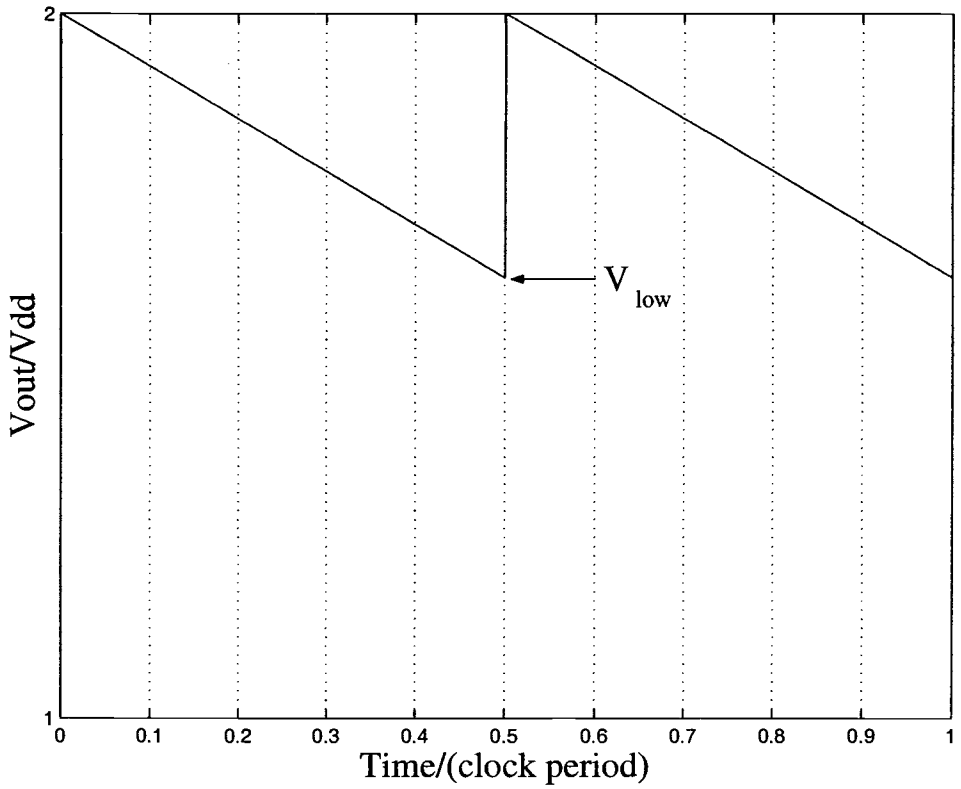

FIGURE 2.2. Output of conventional charge pump.

$$\Delta V_{out} = \frac{I_o}{2f_{clk}(C_a + C_x)} \ . \tag{2.2}$$

The output waveform of such a charge pump is shown in Figure 2.3. Notice that the maximum output voltage is not $2V_{dd}$. That is the case because the minimum output voltage, V_{low}, is determined only by the size of C_a and the load current I_o, and can be derived as

$$V_{low} = 2V_{dd} - \frac{I_o}{2f_{clk}\ C_a} \ . \tag{2.3}$$

The charge lost due to the output current during each clock phase is equal to $\frac{I_o}{2f_{clk}}$. This charge must be provided by C_a because the average charge provided by C_x during the clock phase must be equal to zero in steady-state operation. The lost charge accounts for the voltage drop of $\frac{I_o}{2f_{clk}\ C_a}$.

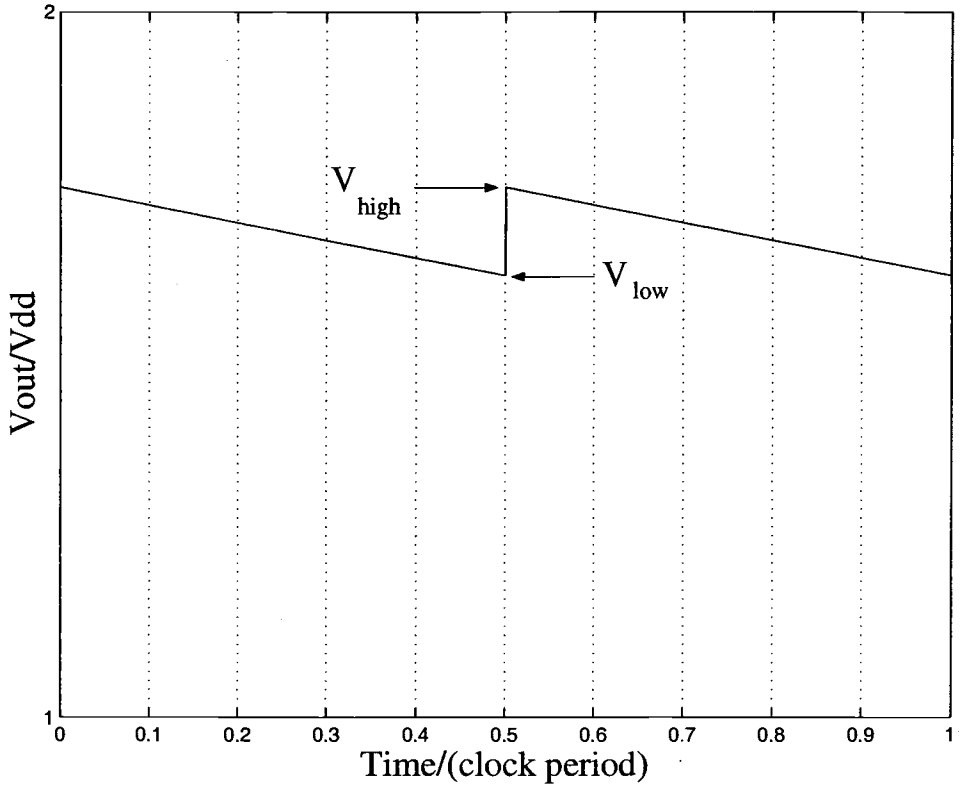

FIGURE 2.3. Output of conventional charge pump with output capacitor.

The addition of C_x reduces the ripple by lowering the maximum output voltage. The top of the ramp is reduced due to charge sharing between C_a and C_x. A more detailed derivation of this result is included in Appendix A.

A drawback of this circuit is that in order to achieve a very small output ripple, a large amount of capacitance is required. For example, in order to drive a load current of 50 μA with a clock frequency of 10 MHz and a desired output ripple of 5 mV, more than 500 pF of capacitance is required. Most of the chip area of a charge pump is due to the capacitor area, so the area efficiency of a charge pump can be significantly improved by reducing the total capacitance needed to achieve a desired output ripple.

Power efficiency is an important consideration in charge pump design. The efficiency is defined as the power delivered to the load divided by the total dissipated power. An important source of power loss in a charge pump is the dynamic power loss due to the charging and discharging of the bottom-plate parasitics of the pump capacitors (C_a and C_b in the previous example). These parasitics are charged to V_{dd} every clock cycle. Top-plate parasitics are generally negligible compared to the bottom-plate parasitics. For the conventional 2-capacitor charge pump shown in Figure 2.1, neglecting the resistance of the switches, and assuming that $\Delta V_{out} << V_{out}$, and $V_{out} \approx V_{low}$, the efficiency can be derived as

$$\eta = \frac{I_o \, V_{low}}{2I_o \, V_{dd} + 2f_{clk} \, p \, C_a \, V_{dd}^2} \, . \tag{2.4}$$

where p is the ratio of stray capacitance to desired capacitance ($C_p = p \, C_a$) and is determined by the type of capacitors used. For double poly capacitors, this value is usually between 10 and 20 percent. Thin oxide MOS capacitors provide 5-15% stray capacitance, with a higher capacitance per unit area than double poly capacitors. For poly-metal capacitors, p can be as much as 20-50%. The efficiency formula is derived in more detail in Appendix B. A good derivation of an equivalent expression for the power efficiency can be found in [20] as

$$\eta = \frac{1}{1 + p \, C_a \, f_{clk} \frac{(R_L + R_S)^2}{2R_L} + 2C_a \, f_{clk} \frac{R_S^2}{R_L}} \, , \tag{2.5}$$

where $R_L = \frac{V_{low}}{I_o}$ and $R_S = \frac{1}{2f_{clk}C_a}$.

The power efficiency maximum is reached for

$$R_L = R_S \sqrt{1 + \frac{4}{p}} \, . \tag{2.6}$$

Substituting for R_L and R_S, the maximum efficiency is reached for

$$C_a = \frac{I_o}{4f_{clk}V_{dd}} \left(1 + \sqrt{1 + \frac{4}{p}} \right) \, . \tag{2.7}$$

This is the value for C_a where the power efficiency is maximum. Unfortunately, the minimum output voltage, V_{low}, provided by this value for C_a may not be at the desired level. Usually V_{low} (may be given as the average output voltage if the ripple is small) and I_o are given as design constraints, therefore fixing the size of C_a by (2.3), and the efficiency by (2.4).

Figure 2.4 shows the relationship between efficiency and ripple for a fixed total capacitance ($C_{total} = 2C_a + C_x$). The top left part of the curve represents the condition where there is no output capacitor and C_a is equal to half of the total capacitance. Moving down the curve, C_x becomes larger at the expense of C_a while keeping the total capacitance constant. Because the ripple is inversely proportional to $C_a + C_x$, the ripple decreases as C_a gets smaller and C_x gets larger. The efficiency improves as well until the maximum efficiency is achieved and the efficiency falls off sharply. This occurs because as C_a becomes smaller, the degradation in efficiency due to the bottom-plate parasitics decreases, but the average output voltage also decreases, leading to a lower power delivered to the load. The efficiency drops sharply when C_a is small enough that the power delivered to the load decreases more rapidly than the bottom-plate losses. For a charge pump used in this configuration, C_a should be only large enough to provide the desired output voltage at the load as given by (2.3). If there is no constraint on the output voltage, C_a should be sized for maximum power efficiency according to (2.7). C_x should be as large as needed to achieve the desired ripple.

2.2.2. *Filtering Case*

One way to reduce the capacitance needed to achieve a given output voltage ripple is to use the on-resistance of transistor M_3 and capacitor C_x as a lowpass filter.

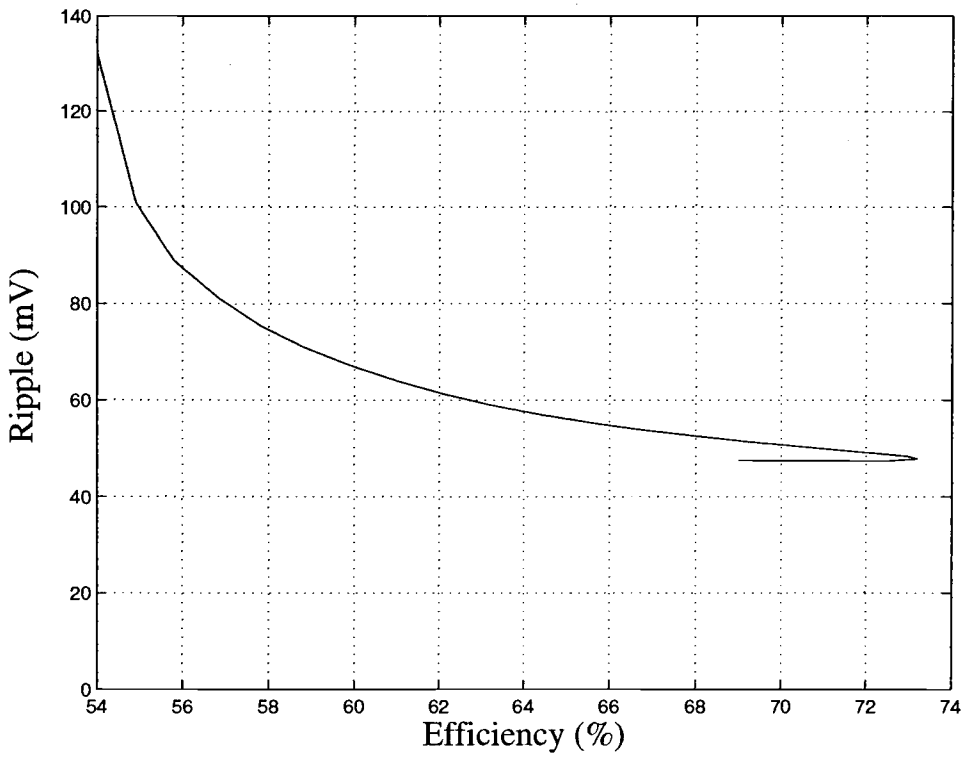

FIGURE 2.4. Efficiency vs. ripple for a conventional charge pump with no filtering (I_o=100 μA, C_{total}=100 pF).

This can be achieved by making transistors M_3 and M_4 narrow enough to provide enough resistance to bring the corner frequency of the RC branch below twice the clock frequency. Unfortunately, this also lowers the average output voltage due to the voltage drop across the switch, but for low current and high clock frequency applications the drop in output voltage may be small.

The ripple at nodes V_a and V_b in Figure 2.1 is a ramp with a fundamental of twice the clock frequency. The FFT of this waveform, shown in Figure 2.5, has a $1/f$ shape, with frequency components at all harmonics of $(2f_{clk})$. The fundamental is the largest component and each harmonic above the fundamental has a smaller magnitude than the previous one.

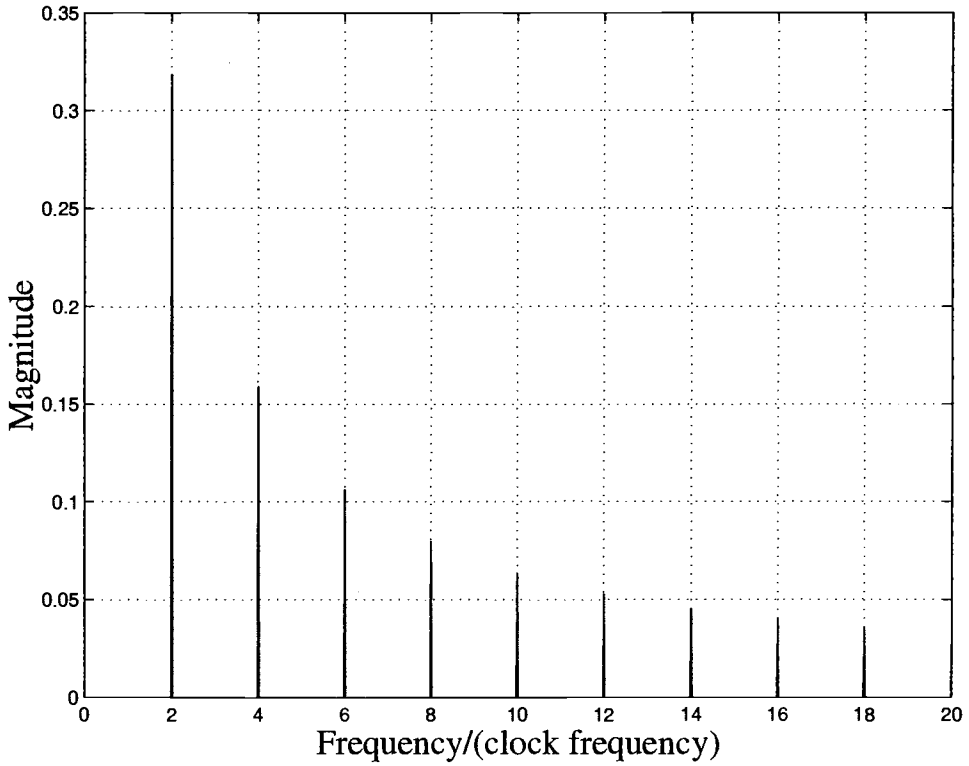

FIGURE 2.5. FFT of a ramp waveform with a peak-to-peak amplitude of 1.

The corner frequency of the RC filter is equal to $\frac{1}{2\pi\ r_{on(M_3)}\ C_x}$. If this corner frequency is less than twice the clock frequency, then an approximation for the ripple at the output, ΔV_{out} can be found by assuming, for simplicity, that the ripple at nodes V_a and V_b in Figure 2.1 is a sinusoid of frequency $2f_{clk}$ and amplitude $\frac{I_o}{2f_{clk}\ C_a}$. If this were the case, the output ripple would be calculated as

$$\Delta V_{out} = \frac{I_o}{8\pi\ f_{clk}^2\ r_{on(M_3)}\ C_a\ C_x}\ . \tag{2.8}$$

The higher frequency terms are attenuated more by the filtering, causing the actual ripple to be less than this single frequency approximation. Simulation has shown that the ripple should be about 0.83 times the single frequency approximation when the ramp waveform is filtered once. If the ramp waveform is filtered twice, this constant is about 0.69. If the ramp waveform is filtered many times, so that

only the fundamental is left, this constant becomes $\frac{2}{\pi}$. This is shown in more detail in Appendix C. Thus, for this charge pump, the approximate output ripple can be described as

$$\Delta V_{out} = \frac{0.83 I_o}{8\pi \ f_{clk}^2 \ r_{on(M_3)} \ C_a \ C_x} \ . \tag{2.9}$$

Figure 2.6 shows the relationship between efficiency and ripple for a fixed total capacitance when filtering is used. Notice that the ripple is now inversely proportional to the product of C_a and C_x, not the sum. It can be shown that for a given total capacitance, the condition for minimum output ripple is when $2C_a = C_x$. This is derived in detail in Appendix D. The shape of the curve in Figure 2.6 is similar to the non-filtering case, but the ripple reaches a minimum level and starts to increase before the efficiency falls off. Now there is a design tradeoff between ripple and efficiency. For this type of charge pump, C_x should be $\geq 2C_a$ for increased efficiency with a small amount of increase in ripple.

2.3. SINGLE CASCODE CHARGE PUMP

A charge pump circuit that greatly reduces the total capacitance needed to achieve a given output voltage ripple is reported by Duisters and Dijkmans in [22]. A similar circuit is shown in Figure 2.7 as a simplified schematic. The capacitors shown with boxes around them symbolize conventional 2-capacitor charge pumps of the type described in the previous section. This circuit uses a buffer transistor and the on resistance of the switch connecting Charge Pump 2 to the drain of M_{buf} to provide additional lowpass filtering of the ripple.

Charge pump 1 provides the gate voltage for the buffer transistor M_{buf}. Because there is no dc current through the gate of M_{buf}, the charge pump capac-

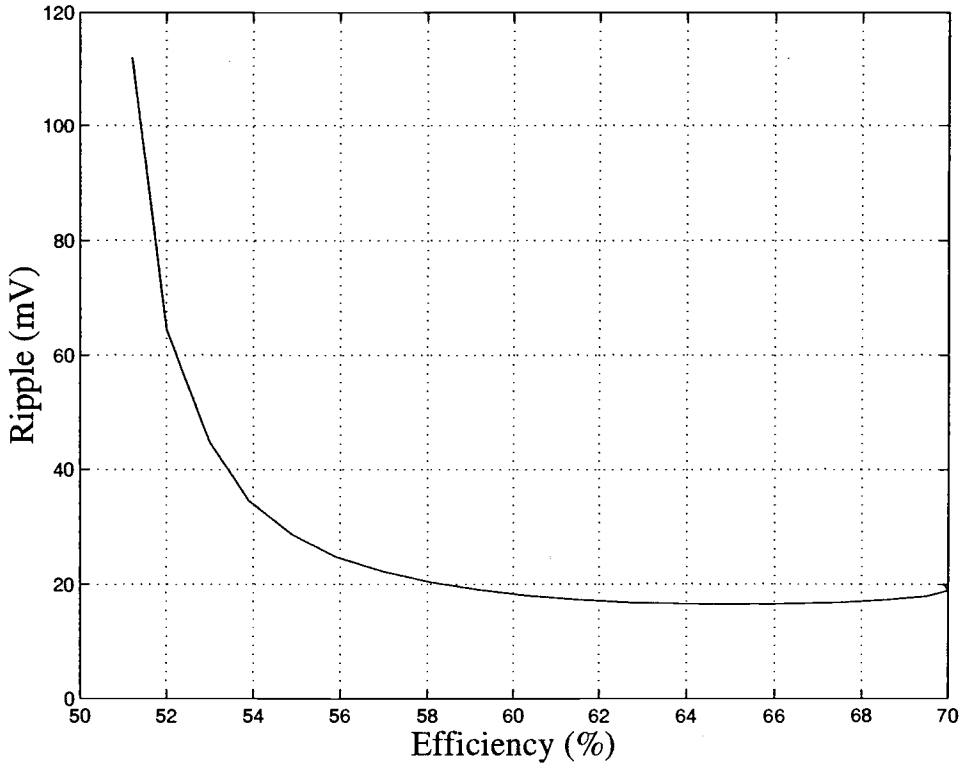

FIGURE 2.6. Efficiency vs. ripple for a conventional charge pump with filtering (I_o=100 μA, C_{total}=100 pF).

itors need only be large enough to overcome the leakage from the parasitic gate capacitance of the switches and provide around $2V_{dd}$ at the gate of M_{buf}. Charge sharing between the charge pump capacitors and the parasitic gate capacitance of the switches reduces the output voltage, but this can be neglected if the switches are small. Charge pump 1 also provides the n-well bias voltage for all of the p-channel switches. Charge pump 2 provides the load current.

Using the small-signal model, an equivalent circuit for approximating the ripple is derived in Appendix E. This equivalent circuit is shown in Figure 2.8. The on-resistance of M_3 and capacitor C_x act as a lowpass filter. The buffer transistor acts like a common-gate amplifier, with input and output reversed. A small signal at

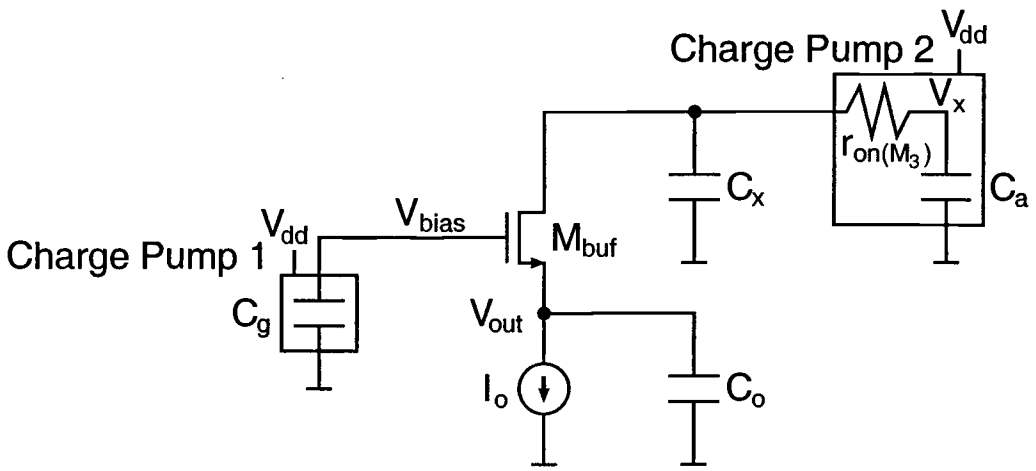

FIGURE 2.7. Single cascode charge pump.

the source is amplified by the intrinsic gain of the buffer transistor, and conversely, a signal entering the drain is attenuated by the same amount. The impedance seen looking into the drain of M_{buf} is $r_{o(M_{buf})} + \frac{C_o}{g_{m(M_{buf})}\, r_{o(M_{buf})}}$. This adds another lowpass filter branch to the equivalent circuit. The attenuation by the intrinsic gain of the buffer transistor can be moved to node V_x because it is a constant in the transfer function derived in Appendix C. The equivalent circuit has been drawn in this way to make it clear that there can be loading of the first RC branch by the second RC branch.

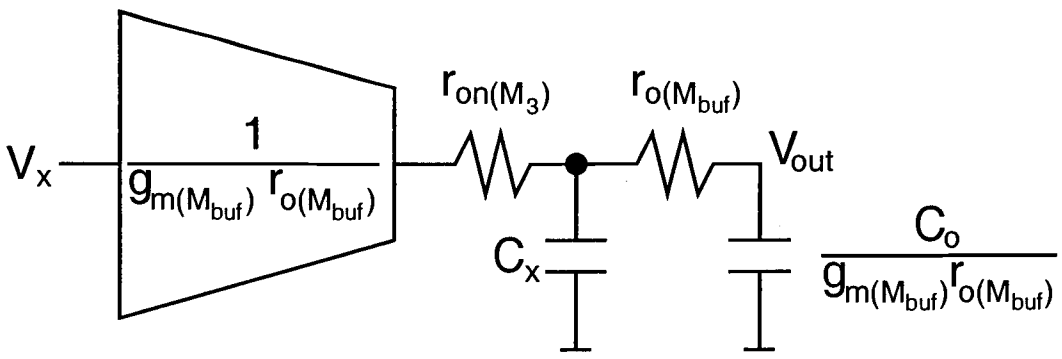

FIGURE 2.8. Small-signal equivalent circuit of single cascode charge pump.

The corner frequency of the first RC branch, ignoring the loading of the second RC branch, is $\frac{1}{2\pi\ r_{on(M_3)}\ C_x}$. The corner frequency of the second branch is $\frac{g_{m(M_{buf})}}{2\pi\ C_o}$.

2.3.1. *Non-filtering Case*

If C_x and C_o are small enough that both of the corner frequencies are greater than twice the clock frequency, then no filtering occurs, and the ripple at the drain of M_{buf} will be $\frac{I_o}{2f_{clk}\ (C_a+C_x)}$ as given by (2.2). The ripple at the output, ΔV_{out}, is reduced by the intrinsic gain of the buffer transistor and can be shown to be

$$\Delta V_{out} = \frac{I_o}{2f_{clk}\ g_{m(M_{buf})}\ r_{o(M_{buf})}(C_a + C_x)} \ . \tag{2.10}$$

Figure 2.9 shows the output of a charge pump used in this way. The upper plot shows the waveforms at the drain of M_{buf} and at the output. The lower plot is a zoomed-in view of the output ripple waveform. The output ripple has the same sawtooth shape as the ripple at the drain of M_{buf}, only it is much smaller. The two waveforms are the same shape because the attenuation is achieved only by the buffer transistor, not by filtering. This circuit trades some headroom loss for greatly reduced output ripple.

It is important that the voltage at the drain of M_{buf} not drop more than V_T below the voltage at the gate in order to keep M_{buf} in saturation. This means that $\frac{I_o}{2f_{clk}\ C_a} + I_o\ r_{on(M_3)} < V_T$. Thus,

$$C_a > C_{a_{min}} = \frac{I_o}{2f_{clk}(V_T - I_o\ r_{on(M_3)})} \ . \tag{2.11}$$

If this condition is violated, M_{buf} becomes biased in the triode region and the output resistance of M_{buf} falls off sharply, causing the circuit to be much less effective. A similar constraint for $r_{on(M_3)}$ may be derived for fixed C_a.

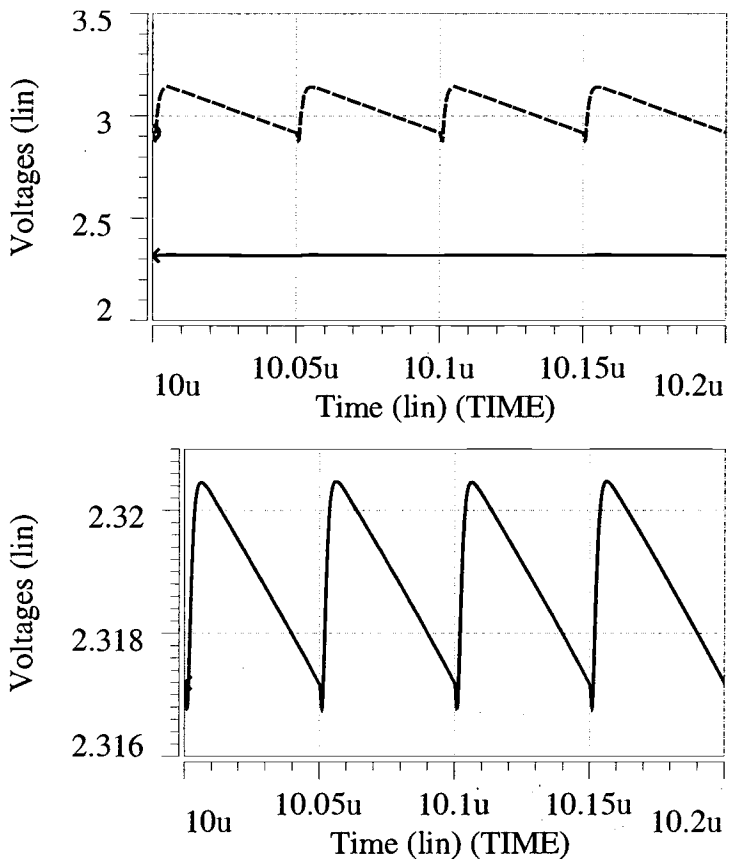

FIGURE 2.9. Output of single cascode charge pump with no filtering (small C_x and C_o).

The relationship between output ripple and efficiency for a fixed total capacitance is shown in Figure 2.10. The left end of the curve is the condition where C_a is equal to half of the total capacitance, and C_x and C_o are equal to zero. If C_o is not large enough to provide any filtering, putting capacitance there does not help, so C_o should only be large enough to filter out glitches coming from Charge Pump 1. Moving to the right on the curve in Figure 2.10, C_a gets smaller and C_x gets larger,

similar to the conventional charge pump ($C_o = 0$). As C_a decreases, the efficiency increases due to smaller bottom-plate parasitics and the ripple decreases until C_a gets too small and M_{buf} goes into the triode region. For a charge pump operating in this fashion, C_a should be large enough to keep M_{buf} in saturation, and C_x should be large enough to achieve the desired output ripple.

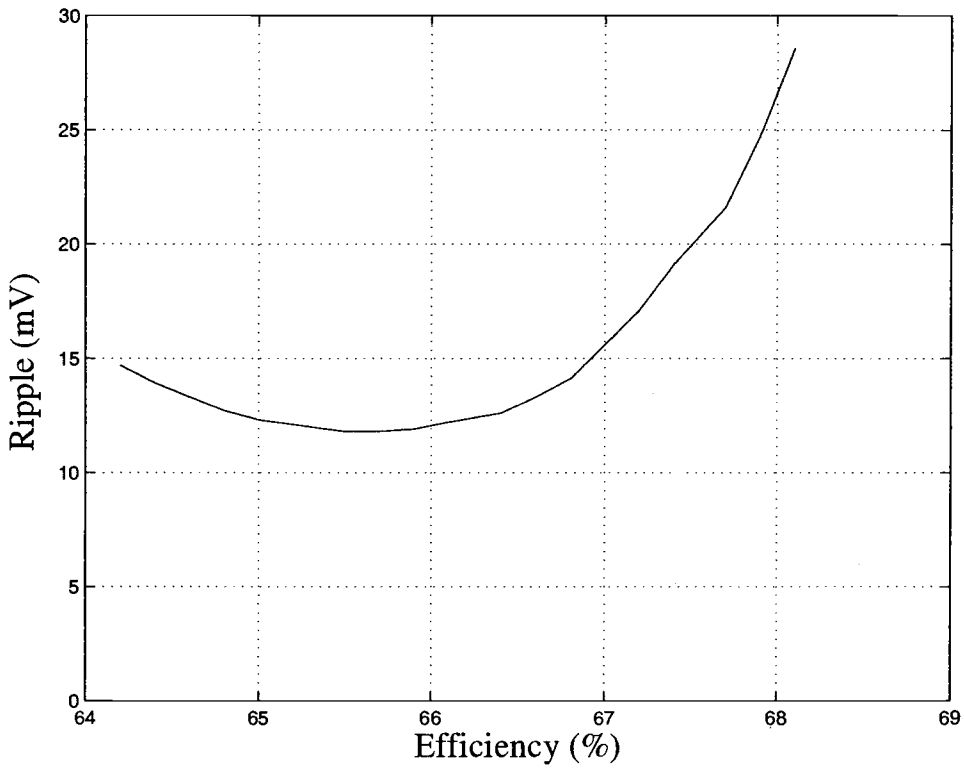

FIGURE 2.10. Efficiency vs. ripple for a single cascode charge pump with no filtering (I_o=100 μA, C_{total}=30 pF).

2.3.2. *Filtering Case*

If enough capacitance is added at the drain of M_{buf} or at the output, the ripple is further reduced due to additional lowpass filtering. The output waveforms of a

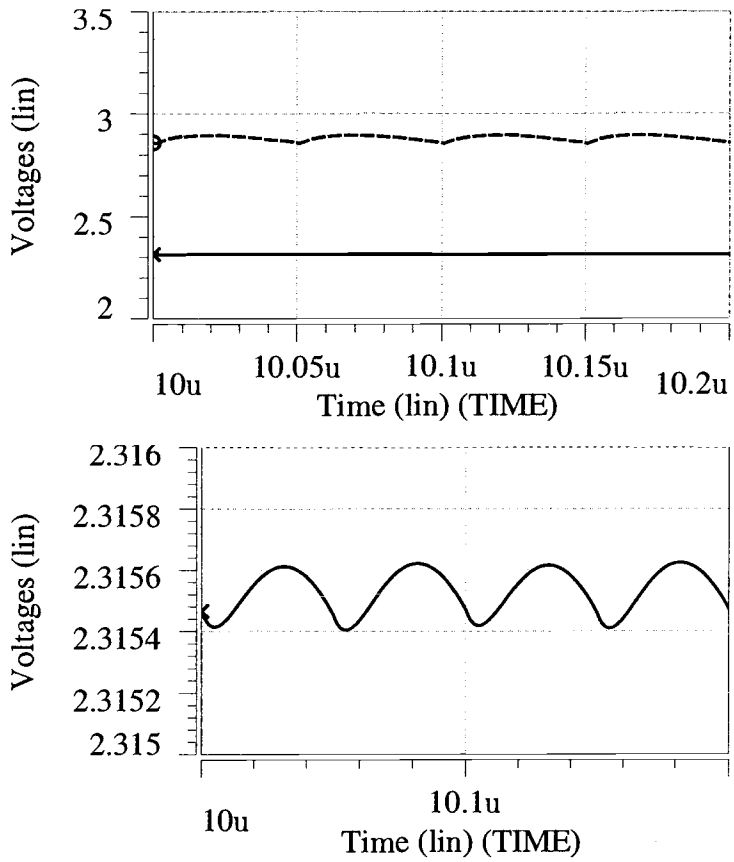

FIGURE 2.11. Output of single cascode charge pump with filtering (large C_x and C_o).

single cascode charge pump that uses filtering are shown in Figure 2.11. The waveforms shown are the same as those shown in Figure 2.9. Notice that the waveforms are smoother due to the lowpass filtering, and the output waveform more closely resembles a sinewave. This is because the higher harmonics of the ramp waveform have been attenuated much more than the fundamental due to the filtering from C_x and C_o. Because the impedance of the second RC branch is much larger than the

impedance of C_x (see Figure 2.8), assuming that C_x and C_o are of similar size, the second RC branch does not have much of a loading effect on the first branch, and we can treat them as two separate single-pole filters.

Assuming the on-resistance of M_3 is large, the ripple at node V_x, ΔV_x, can be found using (2.1) to be

$$\Delta V_x = \frac{I_o}{2 f_{clk}\ C_a}\ . \tag{2.12}$$

If the corner frequencies of the RC branches of this filter are all much less than twice the clock frequency, then an approximation for the ripple at the output, ΔV_{out} can be found to be

$$\Delta V_{out} = \frac{0.69 I_o}{32\pi^2\ f_{clk}^3\ r_{on(M_3)}\ r_{o(M_{buf})}\ C_a\ C_x\ C_o}\ . \tag{2.13}$$

The 0.69 constant is due to the additional filtering of the higher harmonics of the initial ramp waveform (see Appendix C).

The charge pump used to bias the gate of M_{buf} introduces more bottom-plate losses, which lowers the efficiency. The efficiency for a single cascode charge pump can be expressed as

$$\eta = \frac{I_o\ V_{out}}{2 I_o\ V_{dd} + 2p\ f_{clk}(C_a + C_g) V_{dd}^2}\ , \tag{2.14}$$

where V_{out} is now equal to $2V_{dd} - V_T - V_{eff(M_{buf})}$.

The relationship between output ripple and efficiency for a fixed total capacitance of the single cascode charge pump shown in Figure 2.7 is shown in Figure 2.12. Noticing that the ripple is inversely proportional to the product $C_a\ C_x\ C_o$, it can be shown that for a given total capacitance ($C_{total} = 2C_a + C_x + C_o$), the ripple is minimized when $2C_a = C_x = C_o$ (see Appendix D). On the left end of the curve, C_a is equal to half of the total capacitance and C_x and C_o are equal to zero.

Moving to the right, C_a gets smaller and C_x and C_o remain equal to each other and get larger.

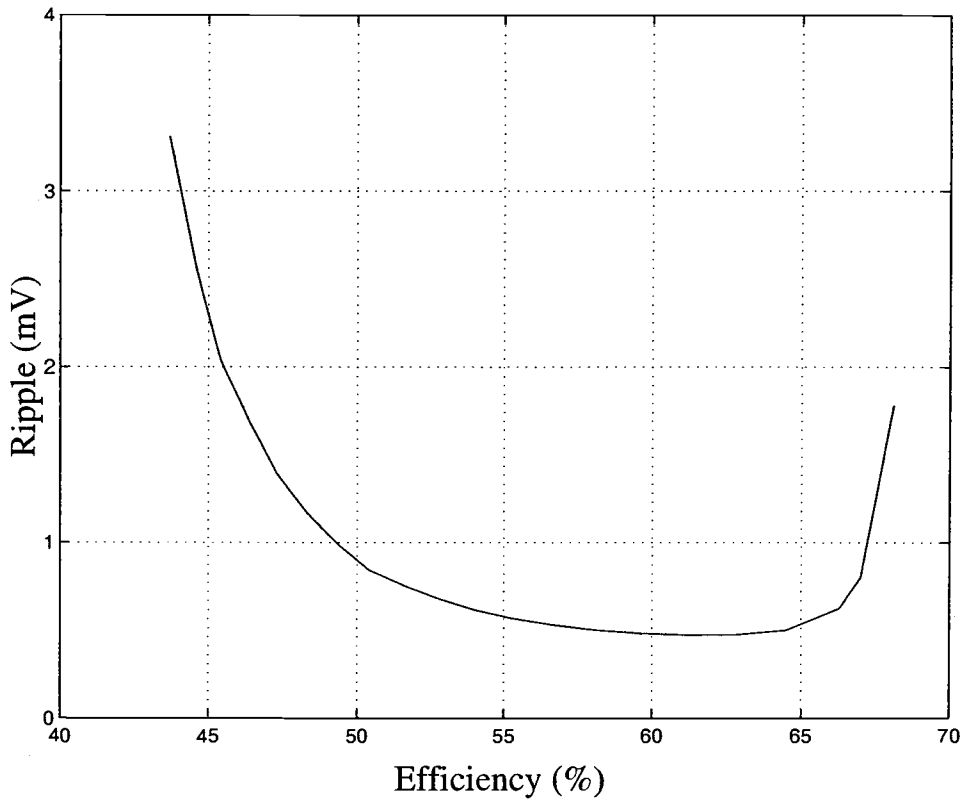

FIGURE 2.12. Efficiency vs. ripple for a single cascode charge pump with filtering (I_o=100 μA, C_{total}=100 pF).

From the minimum ripple point, decreasing C_a and increasing C_x and C_o will improve the efficiency but also increase the ripple. Increasing C_a and decreasing C_x and C_o has an adverse effect on both efficiency and ripple, so for this charge pump there is a tradeoff between efficiency and ripple similar to that of the conventional charge pump when filtering is used. A single cascode charge pump with filtering should have

$$\alpha\, C_a = C_x = C_o\ , \tag{2.15}$$

where $\alpha \geq 2$.

From this result,

$$C_{total} = 2C_a + C_x + C_o = 2(1+\alpha)C_a \ , \tag{2.16}$$

where $\alpha \geq 2$.

Because the minimum ripple for a given total capacitance is inversely proportional to the product of three capacitors and f_{clk}^3 as shown in (2.13), a 25% increase in either the clock frequency or the total capacitance will cut the minimum ripple in half.

2.3.3. *Gate Biasing Circuit*

Because M_{buf} acts as a source follower to the output of the single cascode charge pump, it is important that the gate bias voltage be as stable as possible in order to achieve very low ripple. Although charge pump 1 does not provide any current, there may be a significant ripple at the gate of M_{buf} caused by clocking glitches.

One possible way of ensuring that these glitches are small is to use overlapping clocks, where both clock phases are "high" for a brief period during the transition from one cycle to the next. This method is shown in the upper plot of Figure 2.13. The drawback of this type of clocking scheme is that during the time when both clock phases are "high", the n-channel transistors (M_1 and M_2 in Figure 2.1) are biased in the saturation region, and a significant current flows from nodes V_a and V_b to V_{dd}. This results in unnecessary power loss and also lowers the average value of the biasing voltage because the pumping capacitors lose charge during this transition period. This problem can only be helped by using clock phases with very fast rise and fall times, which require more power to generate and are more difficult

to design. If the clock phases overlap in such a way that both clock phases are "low" at the same time, the same problem occurs with the p-channel transistors.

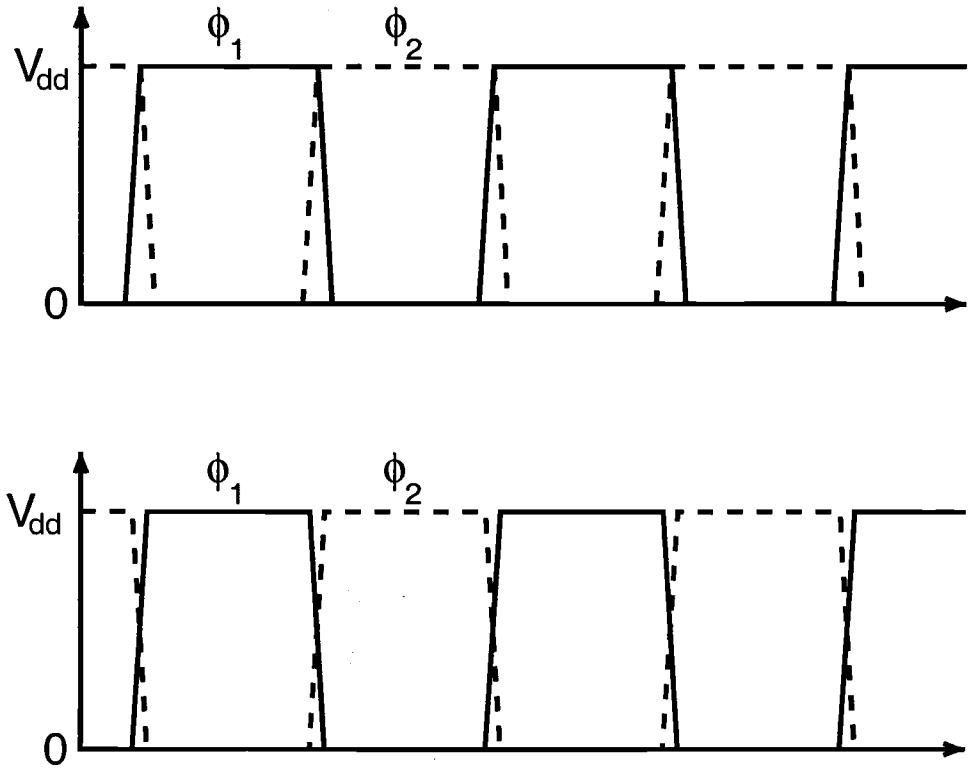

FIGURE 2.13. Timing diagrams for two possible clocking schemes.

From a power standpoint, the best clocking scheme is to have the two phases cross somewhere in the middle of the voltage range as shown in the lower plot of Figure 2.13. This minimizes the wasted power during the transition period, but also results in large glitches at the gate of M_{buf}. The ripple at this node occurs because when M_3 connects C_a to the gate of M_{buf}, the voltage at V_a has not yet reached $2V_{dd}$. The ripple at this node can be as much as a few millivolts, which can significantly increase the ripple at the output node if very low ripple is desired. Of course, filtering by C_o reduces the problem somewhat, but in many cases, it may not be enough.

A circuit which provides a much smaller bias voltage ripple with clock phases that overlap in the middle of the voltage range is shown in Figure 2.14. The circuit is similar to the conventional charge pump circuit, with the addition of two switches and two capacitors. During ϕ_2, switches M_1, M_4 and M_7 are turned on and capacitor C_g is charged to V_{dd}. During ϕ_1, switches M_2, M_3 and M_6 are turned on and capacitor C_h is charged to V_{dd}. Capacitor C_g, which was charged to V_{dd} during the previous clock phase, is now connected between V_{dd} and V_{gx}, lifting the voltage at V_{gx} to $2V_{dd}$. During the next clock phase (ϕ_2), Capacitor C_{gx} is connected between V_{gx} and the gate of M_{buf}, providing a bias voltage of $2V_{dd}$.

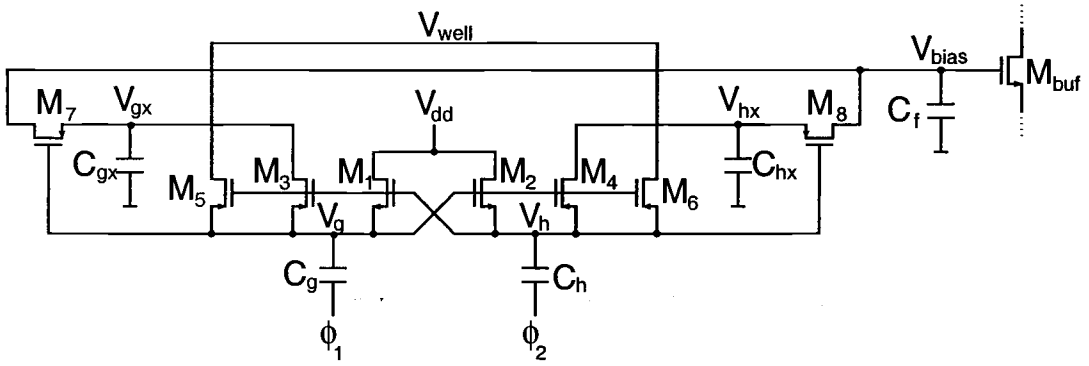

FIGURE 2.14. Modified conventional charge pump to provide the gate voltage for M_{buf}.

The ripple at nodes V_{gx} and V_{hx} is not easily determined, but depends on the rise and fall times of the clocks, the on-resistance of switches M_3 and M_4, and the size of C_{gx} and C_{hx}. If C_{gx} and C_{hx} are very small (less than 0.5 pF), the ripple at V_{gx} and V_{hx} is on the order of tens of millivolts. This is fairly large, but when these nodes are switched to the gate of M_{buf} through M_7 and M_8, the ripple at the output is very small (less than 0.1 mV). A small capacitor, C_f, at the gate of M_{buf} also helps to reduce the bias voltage ripple.

This biasing circuit provides very low ripple, even when clock phases with very slow rise and fall times are used.

2.3.4. *Design Considerations*

Assuming enough capacitance is used for filtering to occur, there are several ways of designing a charge pump of this structure. For example, if the desired ripple and efficiency are given, the capacitor sizes can be determined using the equations given in the previous section. If the desired efficiency, clock frequency, output current and output voltage are known, the size of C_a can be found from (2.14) to be

$$C_a = \frac{I_o\ V_{out} - 2I_o\ V_{dd}\ \eta}{2p\ f_{clk}\ V_{dd}^2\ \eta} - C_g\ , \tag{2.17}$$

where C_g needs to only be large enough to overcome the parasitics of the switches in charge pump 1 (1 pF is large enough for most applications). After C_a is known, the values for C_x and C_o can be determined from (2.13), using the constraint that C_x should be equal to C_o, to be

$$C_x = C_o = \sqrt{\frac{0.69 I_o}{32\pi^2\ f_{clk}^3\ r_{on(M_3)}\ r_{o(M_{buf})}\ C_a\ \Delta V_{out}}}\ . \tag{2.18}$$

For this case where the desired ripple and efficiency are given, there is only one solution because C_a is determined by the desired efficiency, and C_x and C_o are determined by the desired output ripple.

For another scenario where the desired ripple and total capacitance are given, the design can be determined as follows: from (2.15) and (2.16), the product $C_a C_x C_o$ is equal to $\alpha^2 C_a^3$ which is equal to $\frac{\alpha^2 C_{total}^3}{8(1+\alpha)^3}$. Combining this result with (2.13), the output ripple can be expressed as a function of the total capacitance and the variable α as

$$\Delta V_{out} = \frac{0.69 I_o}{4\pi^2 \, f_{clk}^3 \, r_{on(M_3)} \, r_{o(M_{buf})} \, C_{total}^3} \, \frac{(1+\alpha)^3}{\alpha^2} \, , \tag{2.19}$$

where $\alpha \geq 2$.

Rearranging (2.19), the total capacitance can be found in terms of α and ΔV_{out}. The result is shown in (2.20).

$$C_{total} = \left(\frac{0.69 I_o}{4\pi^2 \, f_{clk}^3 \, r_{on(M_3)} \, r_{o(M_{buf})} \, \Delta V_{out}} \, \frac{(1+\alpha)^3}{\alpha^2} \right)^{\frac{1}{3}} \, . \tag{2.20}$$

Solving (2.20) numerically for α gives the distribution of the capacitance C_{total} needed to achieve the desired output ripple. C_a, C_x, and C_o can then be easily determined. $C_a = \frac{C_{total}}{2+2\alpha}$, and $C_x = C_o = \frac{\alpha \, C_{total}}{2+2\alpha}$.

It is possible to create a number of different design scenarios with predetermined and undetermined design variables that can be solved by different mathematical derivations of the basic design equations.

In many cases, only the desired ripple is given, creating a design tradeoff between total capacitance and power efficiency, illustrated by the constant ripple curves in Figure 2.15. For each curve, C_{total} is varied and the value for α also varies in order to keep the output ripple the same. The lower left end of the curve is the minimum capacitance point (where $\alpha = 2$), which is equivalent to the minimum ripple point on the constant total capacitance curve. The other end of the curve is the maximum efficiency point, which is where C_a gets too small and M_{sat} goes into the triode region.

The efficiency is maximized when C_a is made as small as possible to minimize bottom-plate losses. The maximum efficiency that can be achieved within the constraints previously discussed can be found from (2.14) using the minimum allowed value of C_a from (2.11) to be

$$\eta_{max} = \frac{I_o \, V_{out}}{2 I_o \, V_{dd} + 2p \, f_{clk} (C_{a_{min}} + C_g) V_{dd}^2} \tag{2.21}$$

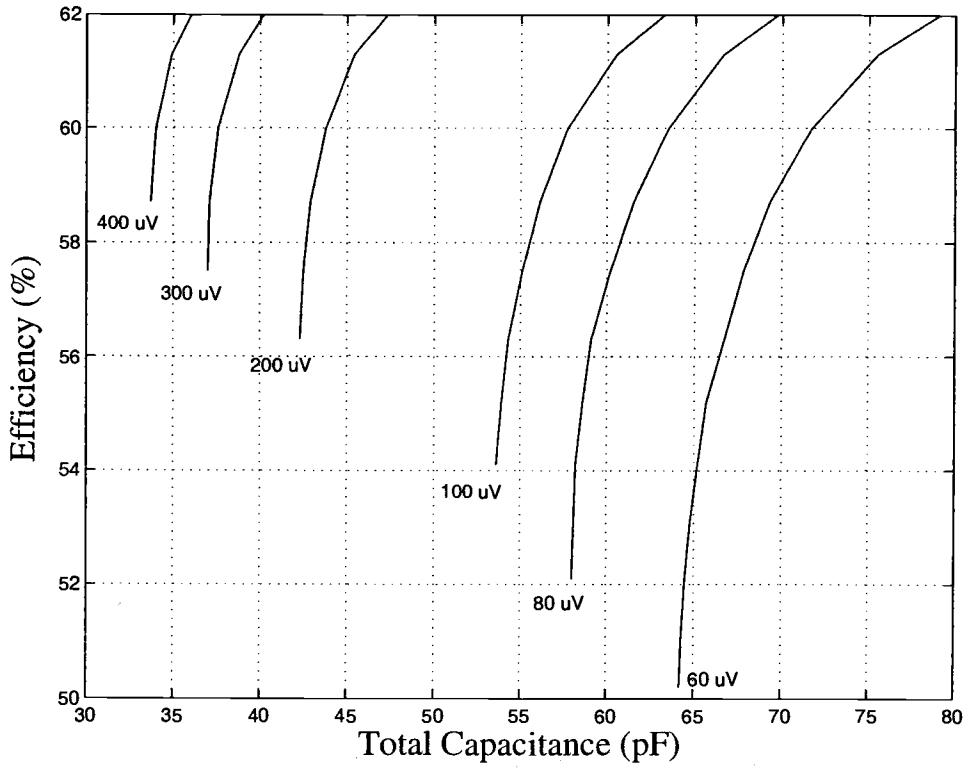

FIGURE 2.15. Total capacitance vs. efficiency for fixed output ripple (I_o=50 μA).

On the constant total capacitance curve of Figure 2.11, this is where the ripple rises sharply. On the constant ripple curve, this is where the curves bend sharply to the right, as opposed to the smooth parabolic curves before the breakpoint. This occurs because as you move up the curves in the figure, C_a is getting smaller and the minimum voltage at the drain of M_{buf} is getting lower. When this voltage gets low enough that M_{buf} goes into the triode region, the amount of capacitance needed to achieve the same low ripple greatly increases. In Figure 2.15, the top of the graph is approximately where this breakpoint occurs. Typically, some safety margin would be included in the design to ensure that C_a was not too small to avoid getting too close to the breakpoint.

Figure 2.15 shows that maximizing the efficiency unfortunately also maximizes the necessary total capacitance needed to achieve the desired ripple. The total capacitance needed to achieve this maximum efficiency and provide the desired output ripple, $C_{total_{max}}$ is equal to $2C_{a_{min}} + C_x + C_o$, and can be found using (2.18). The result is shown in (2.22), where $C_{a_{min}}$ is the minimum allowed value for C_a from (2.11).

$$C_{total_{max}} = 2C_{a_{min}} + 2\sqrt{\frac{0.69 I_o}{32\pi^2 \, f_{clk}^3 \, r_{on(M_3)} \, r_{o(M_{buf})} \, C_{a_{min}} \, \Delta V_{out}}} \; . \tag{2.22}$$

The total capacitance is minimized when α is equal to 2. Solving (2.20) with α equal to 2, the minimum total capacitance needed to achieve the desired ripple is found to be

$$C_{total_{min}} = \left(\frac{0.69 \times 27 I_o}{16\pi^2 \, f_{clk}^3 \, r_{on(M_3)} \, r_{o(M_{buf})} \, \Delta V_{out}} \right)^{\frac{1}{3}} \; . \tag{2.23}$$

Minimizing the total capacitance also minimizes the efficiency because it maximizes the size of C_a. The maximum value for C_a is equal to $\frac{C_{total}}{6}$. This sets the efficiency for the case where the minimum amount of capacitance is used to achieve a given ripple to be

$$\eta_{min} = \frac{I_o \, V_{out}}{2 I_o \, V_{dd} + 2p \, f_{clk} \left(\frac{C_{total_{min}}}{6} + C_g \right) V_{dd}^2} \; . \tag{2.24}$$

The constant ripple curves in Figure 2.15 are very steep near the minimum total capacitance point. This means that using a little more capacitance than necessary can significantly improve power efficiency for a given ripple, but the more capacitance is added, the less the efficiency improves. The best tradeoff between efficiency and total capacitance is not easily determined, and will vary according to the application the charge pump is designed for. Although there is no easy way to determine the "sweet spot" in the design curve, the end points of the curve (minimum and maximum capacitance and efficiency) can be easily determined. Between

the two extremes, the designer can trade total capacitance for power efficiency to find the best design for the needed application.

2.4. DOUBLE CASCODE CHARGE PUMP

A proposed new charge pump circuit that further improves the area efficiency of a charge pump is shown in Figure 2.16. This circuit includes two cascoded buffer transistors, a third charge pump and an additional output capacitor. Charge pump 1 provides a bias voltage of $2V_{dd}$ to M_{buf1}. Charge pump 2 provides a lower gate voltage for transistor M_{buf2} in order to bias it in the saturation region. Charge pump 3 provides the load current.

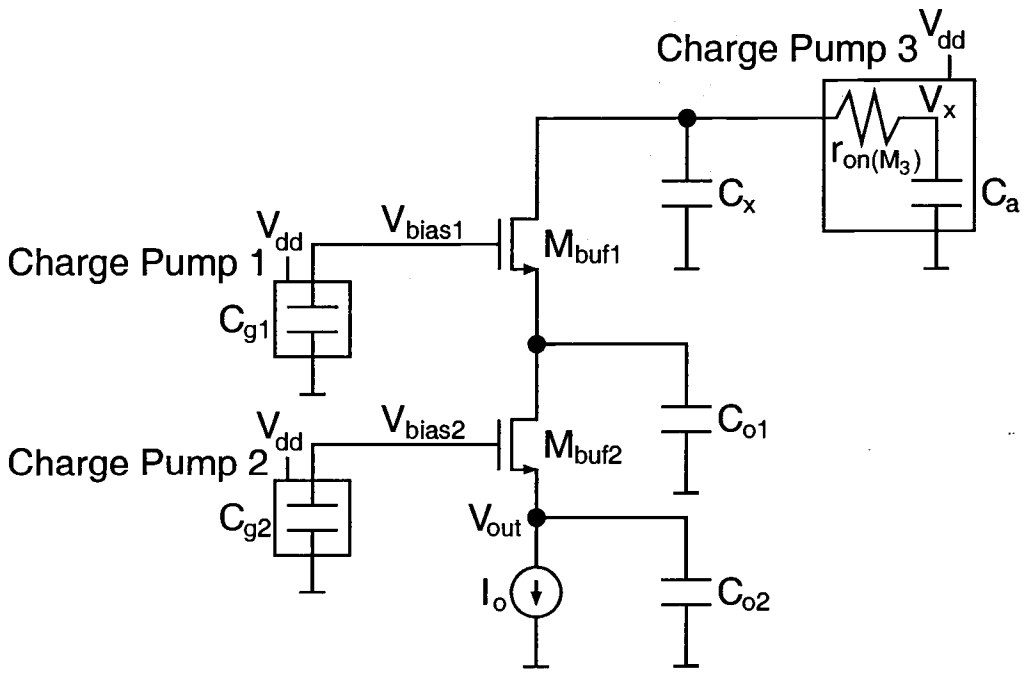

FIGURE 2.16. Double cascode charge pump.

2.4.1. *Second Gate Biasing Circuit*

The circuit used for biasing the gate of M_{buf2} is shown in Figure 2.17. This circuit is similar to the modified conventional 2-capacitor charge pump shown in Figure 2.10, with the addition of two capacitors, C_y and C_z from nodes V_{g2} and V_{h2} to ground. This circuit works as follows: when capacitor C_{g2} is charged to V_{dd}, C_y is also charged to V_{dd}. When the clock ϕ_1 rises to V_{dd}, some of the charge in C_{g2} is shared with C_y, causing the output voltage to be less than $2V_{dd}$. When ϕ_1 drops to zero, the additional charge stored in C_y goes back into C_{g2}. This is important because it allows a voltage lower than $2V_{dd}$ to be provided without any additional power loss due to the extra capacitors. The bias voltage V_{bias2} should be no higher than $2V_{dd} - V_{eff(M_{buf1})}$ in order to bias M_{buf2} in the saturation region.

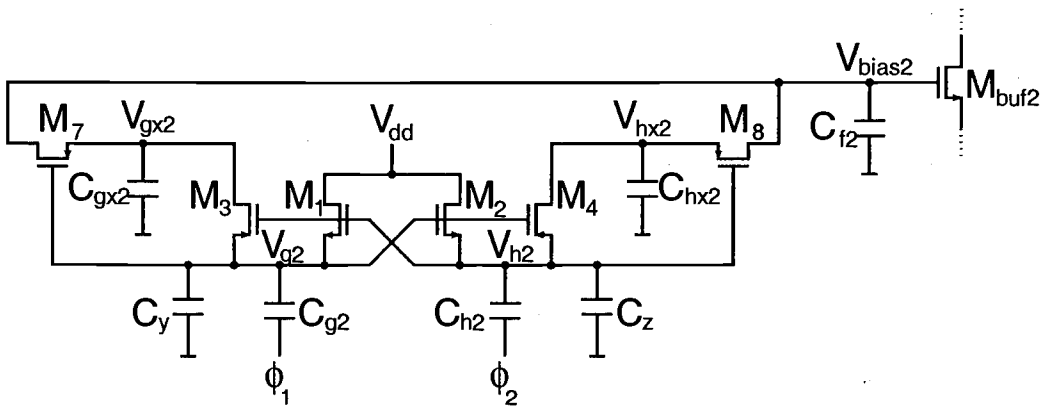

FIGURE 2.17. Conventional charge pump with added capacitors to obtain the second bias voltage.

The output voltage of this biasing charge pump is controlled by the size of C_y (and C_z) and is given by (2.25).

$$V_{bias2} = V_{dd}\left(1 + \frac{C_{g2}}{C_{g2} + C_y}\right) \tag{2.25}$$

The advantage of this type of bias circuit is that the only additional power loss is due to the bottom-plate parasitics of C_{g2} and C_{h2}.

2.4.2. *Non-filtering Case*

The small-signal equivalent circuit for this new charge pump can be found in the same manner as the single cascode charge pump circuit and is shown in Figure 2.18 (see Appendix E). It is similar to the single cascode equivalent circuit, but it includes an additional attenuation of the intrinsic gain of M_{buf2} and another filter branch consisting of a resistor, R, and a capacitor, C, where

$$R = r_{o(M_{buf2})}\ g_{m(M_{buf1})}\ r_{o(M_{buf1})}\ , and \tag{2.26}$$

$$C = \frac{C_{o2}}{g_{m(M_{buf1})}\ r_{o(M_{buf1})}\ g_{m(M_{buf2})}\ r_{o(M_{buf2})}}\ . \tag{2.27}$$

In order for the third branch to filter effectively, the corner frequency, $\frac{g_{m(M_{buf2})}}{2\pi\ C_{o2}}$, must be less than twice the clock frequency.

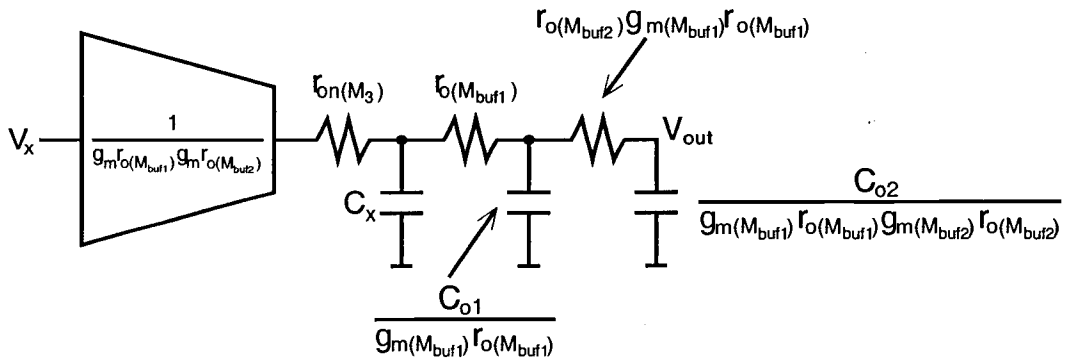

FIGURE 2.18. Small-signal equivalent circuit of new charge pump.

If the on-resistance of the p-channel switches in charge pump 3 is small, the ripple at the drain of M_{buf1} will be $\frac{I_o}{2f_{clk}(C_a+C_x)}$ as given by (2.2). If a small amount

of capacitance is used, so that no filtering occurs, the ripple at the output is reduced by the product of the intrinsic gains of M_{buf1} and M_{buf2}. The ripple at the output, ΔV_{out}, for a double cascode charge pump with no filtering is easily determined to be

$$\Delta V_{out} = \frac{I_o}{2f_{clk}\ g_{m(M_{buf1})}\ r_{o(M_{buf1})}\ g_{m(M_{buf2})}\ r_{o(M_{buf2})}(C_a + C_x)} . \tag{2.28}$$

In the design of a double cascode charge pump, if C_{o1} and C_{o2} are too small to provide any filtering, they should be only large enough to filter out glitches coming from the biasing charge pumps, C_a should be the minimum amount required to keep M_{buf1} in saturation given by (2.11), and C_x should be made large enough to achieve the desired output ripple. Making C_a as small as possible minimizes the loss due to the bottom-plate parasitics, and thus maximizes the power efficiency. Cascoding more buffer transistors would not change the optimum capacitor distribution in the non-filtering case, as any additional output capacitors should only be made large enough to filter out glitches.

2.4.3. *Filtering Case*

Treating the three filter branches as separate single-pole filters and assuming that the corner frequencies of all of the branches are well below $(2f_{clk})$, an approximation for the ripple at the output for the filtering case can be found. The result of this approximation is similar to that of the single cascode charge pump an is shown in (2.29).

$$\Delta V_{out} = \frac{\left(\frac{2}{\pi}\right)\Delta V_x}{64\pi^3\ f_{clk}^3\ r_{on(M_3)}\ r_{o(M_{buf1})}\ r_{o(M_{buf2})}\ C_a\ C_x\ C_{o1}\ C_{o2}} . \tag{2.29}$$

After replacing ΔV_x and simplifying, an equivalent expression for the ripple is derived, shown in (2.30).

$$\Delta V_{out} = \frac{I_o}{64\pi^4 \, f_{clk}^4 \, r_{on(M_3)} \, r_{o(M_{buf1})} \, r_{o(M_{buf2})} \, C_a \, C_x \, C_{o1} \, C_{o2}} \, . \tag{2.30}$$

The charge pump used to bias the gate of the second buffer transistor introduces additional bottom-plate losses, but because the capacitors used in this charge pump are small compared to C_a and C_b, this power loss may be tolerable. The power efficiency of the double cascode charge pump is described by

$$\eta = \frac{I_o \, V_{out}}{2I_o \, V_{dd} + 2p \, f_{clk}(C_a + C_{g1} + C_{g2})V_{dd}^2} \, , \tag{2.31}$$

where V_{out} is now equal to $2V_{dd} - V_T - V_{eff(M_{buf1})} - V_{eff(M_{buf2})}$.

The second buffer transistor lowers the average output voltage by another V_{eff}, which leads to a lower power efficiency.

If a large amount of capacitance is used, the minimum ripple that can be provided by a given total capacitance is achieved when $2C_a = C_x = C_{o1} = C_{o2}$, similar to the single cascode charge pump. If more stages are added, all of the capacitors should be equal, except for C_a, which should be half as large, because two are required but only one is connected at any given time.

The design tradeoffs of the double cascode charge pump are similar to the single cascode, where making $2C_a < C_x = C_{o1} = C_{o2}$ will improve the power efficiency of the charge pump, but also increase the amount of capacitance needed to achieve the desired output ripple. Notice that the minimum ripple point is now inversely proportional to the product $f_{clk}^4 \, C_{total}^4$, meaning that only a 20% increase in either f_{clk} or C_{total} will cut the minimum ripple in half.

2.5. IMPLEMENTATION

In order to compare the performance of the three charge pumps, an implementation was designed and simulated for a 0.35-micron CMOS process. The im-

plementation included a conventional charge pump, a single cascode charge pump, and a double cascode charge pump. The layout of this implementation is shown in Figure 2.19. The conventional charge pump is located at the top, with the single cascode charge pump and double cascode charge pump below. The bottom section contains the load resistors, some control logic and a source follower output buffer. The source follower is connected between the charge pump output and the output pad to ensure that the pad capacitance will not affect the operation of the circuit.

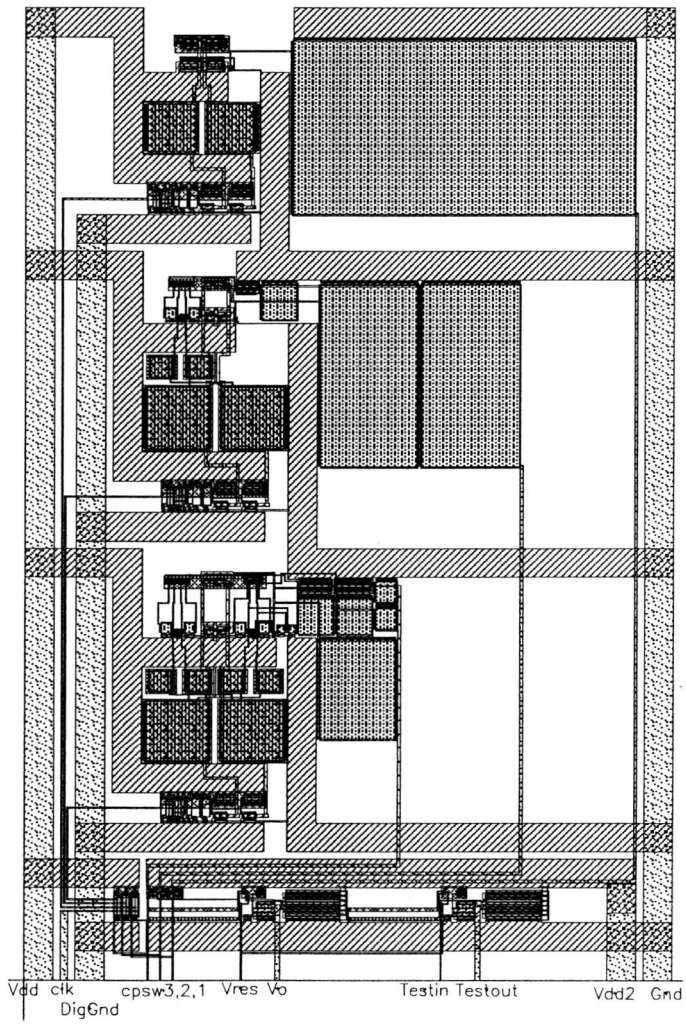

FIGURE 2.19. Layout of implementation.

All of the charge pumps were designed to boost a 1.65-volt supply to 2.3 volts and provide 100 μA of load current using a 10 MHz clock. The conventional charge pump was designed to maximize the power efficiency using 150 pF of capacitance. The single cascode charge pump was designed to have a slightly smaller chip area and provide a much smaller output ripple. The single cascode charge pump was also designed for maximum power efficiency. The chip areas of these two charge pumps were made to be similar in order to show that the single cascode charge pump could achieve a much smaller output ripple using about the same chip area as the conventional charge pump.

The double cascode charge pump was designed to provide about the same output ripple as the single cascode charge pump. This was done in order to verify that the double cascode charge pump could provide an output ripple similar to that of the single cascode charge pump using less capacitance, and therefore less chip area. The double cascode charge pump was also designed for maximum power efficiency.

The conventional charge pump had a chip area of 0.040 mm^2, including a total capacitance of 150 pF, and provided an output ripple of 20 mV. The single cascode charge pump had a chip area of 0.034 mm^2, including a total capacitance of 108 pF, and provided an output ripple of 140 μV. The double cascode charge pump had a chip area of 0.024 mm^2, including a total capacitance of 58 pF, and provided an output ripple of 180 μV.

To provide the two clock phases, each charge pump included a non-overlapping clock generator circuit which was connected to the reference clock by a switch. The chip areas given do not include the non-overlapping clock generator circuits. Each clock generator circuit had a chip area of about 0.002 mm^2. The total

capacitance numbers were calculated assuming a typical oxide capacitance of 5 fF per μm^2.

2.6. CONCLUSION

Figure 2.20 shows a plot of output ripple vs. efficiency for all three charge pumps where no filtering is used. This shows that as buffer transistors are added to the circuit, the maximum efficiency decreases, but the ripple gets much smaller by the intrinsic transistor gain for each stage that is added, regardless of the total capacitance or clock frequency that is used.

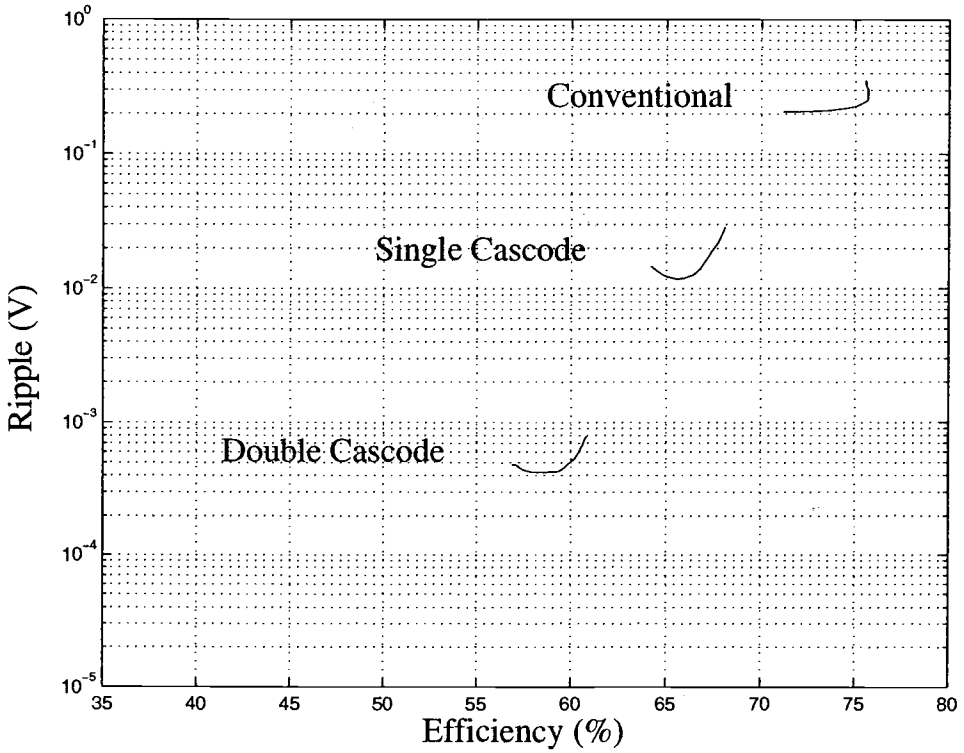

FIGURE 2.20. Efficiency vs. ripple for different charge pump circuits with no filtering (I_o=100 μA, C_{total}=30 pF).

The relationship between output ripple and efficiency for the case where filtering is used is shown in Figure 2.21. For the filtering case, the difference in minimum ripple for the three charge pumps is much larger, and is dependent on how much total capacitance is used. If more capacitance were used, the curves would be farther apart, and if less capacitance were used, the curves would be closer together.

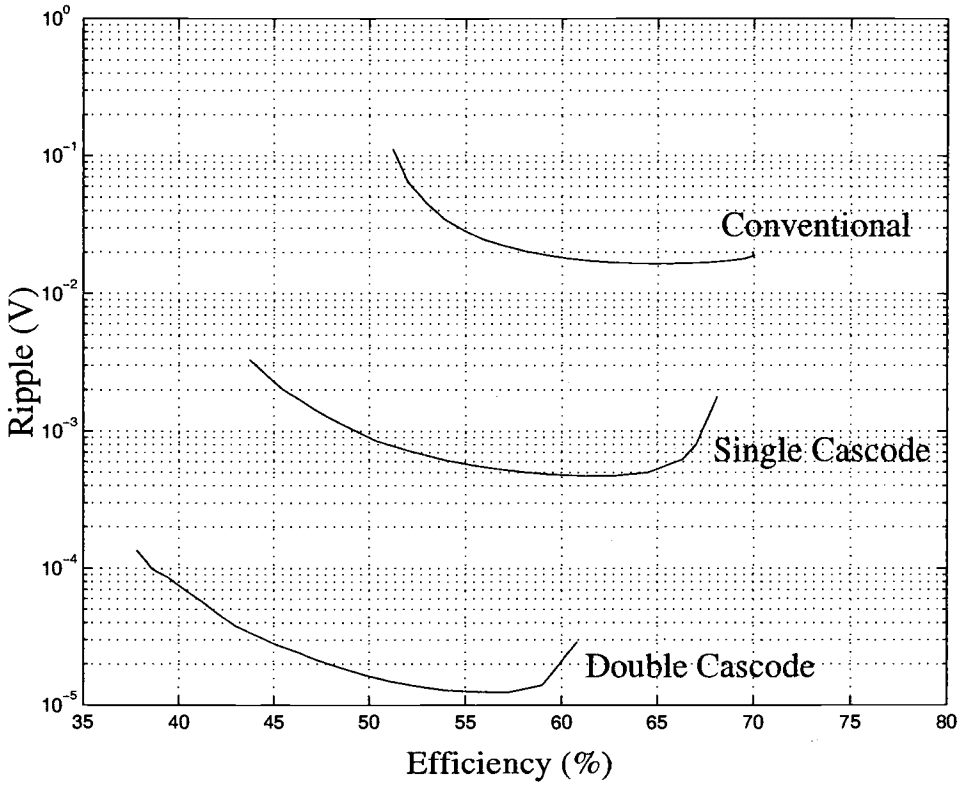

FIGURE 2.21. Efficiency vs. ripple for different charge pump circuits with filtering (I_o=100 μA, C_{total}=100 pF).

For the filtering case, a large output resistance is desired, which unfortunately also increases the headroom loss. Headroom losses limit the number of cascodes that can be used, but in the non-filtering case, the buffer transistors can be made wider to increase the intrinsic gain and reduce the headroom loss. Making the buffer

transistors wider helps as long as the transistors are not too wide. The gate-to-drain parasitic capacitances of the buffer transistors allow some of the ripple at the drain to couple into the gete, and then to the output through the source follower. This effect can significantly increase the output ripple from what is expected if the buffer transistors are very wide.

The simulation results show that for a similar chip area, the conventional charge pump has a much larger output ripple than the single cascode and double cascode charge pumps. Also, the double cascode charge pump uses less chip area than the single cascode charge pump to achieve a similar output ripple.

This new charge pump circuit is more area-efficient than the single cascode charge pump, at a price of reduced headroom and additional bottom-plate losses due to the second biasing circuit. The size of ripple that can be tolerated seems to be the most important factor in determining which charge pump to use for a given application. For example, if power efficiency is the biggest concern and a larger output ripple can be tolerated, a single cascode or even a conventional charge pump would be the best choice. This circuit would be especially useful for applications where power efficiency is not critical and a very low output ripple is desired.

3. SUMMARY

The basic operation of the conventional charge pump, the single cascode charge pump and the double cascode charge pump were described in detail. The important issues of output ripple, power efficiency and area efficiency were discussed. A method of estimating the output ripple of a charge pump given the size of the capacitors used was presented. Power efficiency theory for charge pumps was discussed.

The optimal distribution of available capacitance for minimizing output ripple or maximizing power efficiency was derived. The possible design tradeoffs between output ripple, power efficiency and area efficiency were outlined.

Techniques for improving the area efficiency of a charge pump were presented. A new area-efficient charge pump circuit was proposed that uses two cascoded buffer transistors to achieve a very low output ripple. The improvements involved in the double cascode charge pump included using the on resistance of the p-channel switches for additional filtering of the ripple, the development of a biasing charge pump circuit that is less sensitive to clocking glitches than the conventional charge pump circuit, and the introduction of a biasing charge pump circuit to provide a stable voltage of less than twice the supply voltage to the gate of the second buffer transistor. The double cascode charge pump is more area-efficient than the single cascode charge pump, at the cost of a reduced output voltage and a small amount of additional power loss.

The simulation results for an implementation that included one of each of the three types of charge pump circuits verified the improved area efficiency of the double cascode charge pump circuit. All of the charge pumps boosted a 1.65-volt

supply to around 2.3 volts, and provided 100 μA of load current with a 10 MHz clock frequency.

BIBLIOGRAPHY

[1] J. D. Cockcroft and E. T. Walton, "Production of high velocity positive ions," *Proceedings of the Royal Society of London, A*, vol. 136, pp. 619-630, 1932.

[2] J. Dickson, "On-chip high-voltage generation in MNOS integrated circuits using an improved voltage multiplier technique," *IEEE Journal of Solid-State Circuits*, vol. SC-11, pp. 374-378, June 1976.

[3] J. Witters, G. Groeseneken and H. Maes, "Analysis and modeling of on-chip high-voltage generator circuits for use in EEPROM circuits," *IEEE Journal of Solid-State Circuits*, vol. 24, pp. 1372-1381, October 1989.

[4] T. Tanzawa and T. Tanaka, "A dynamic analysis of the Dickson charge pump circuit," *IEEE Journal of Solid-State Circuits*, vol. 32, pp. 1231-1240, Aug. 1997.

[5] Y. Nakagome *et al.*, "An experimental 1.5-V 64 Mb DRAM," *IEEE Journal of Solid-State Circuits*, vol. 26, pp. 465-472, April 1991.

[6] S. Singer, "Inductance-less up dc-dc convertor," *IEEE Journal of Solid-State Circuits*, vol. SC-17, pp. 778-781, Aug. 1982.

[7] P. Gillingham *et al.*, "High-speed, high-reliability circuit design for megabit DRAM," *IEEE Journal of Solid-State Circuits*, vol. 26, pp. 1171-1175, August 1991.

[8] W. Martino *et al.*, "An on-chip back-bias generator for MOS dynamic memory," *IEEE Journal of Solid-State Circuits*, vol. SC-15, pp. 820-825, Oct. 1980.

[9] T. Byunghak and P. Gray, "A 10 b, 20 Msample/s, 35 mW pipeline A/D converter," *IEEE Journal of Solid-State Circuits*, vol. 30, pp. 166-172, March 1995.

[10] A. Umezawa *et al.*, "A 5-V-only operation 0.6-μ m flash EEPROM with row decoder scheme in triple-well structure," *IEEE Journal of Solid-State Circuits*, vol. 27, pp. 1540-1545, Nov. 1992.

[11] S. D'Arrigo *et al.*, "A 5V-only 256k bit CMOS flash EEPROM," *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 132-133, Feb. 1989.

[12] T. Tanzawa *et al.*, "Circuit technologies for a single-1.8V flash memory," *IEEE Symposium on VLSI Circuits Digest of Technical Papers*, pp. 63-64, 1997.

[13] T. Kawahara *et al.*, "Bit-line clamped sensing multiplex and accurate high voltage generator for quarter-micron flash memories," *IEEE Journal of Solid-State Circuits*, vol. 31, pp. 1590-1600, Nov. 1996.

[14] K. Choi *et al.*, "Floating-well charge pump circuits for sub-2.0V single power supply flash memories," *IEEE Symposium on VLSI Circuits Digest of Technical Papers*, pp. 61-62, 1997.

[15] K. Sawada, Y. Sugawara and S. Masui, "An on-chip high-voltage generator circuit for EEPROM's with a power supply voltage below 2V," *IEEE Symposium on VLSI Circuits Digest of Technical Papers*, pp. 75-76, 1995.

[16] J. Wu, Y. Chang and K. Chang, "1.2V CMOS switched-capacitor circuits," *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 388-389, Feb. 1996.

[17] J. Wu and K. Chang, "MOS charge pumps for low-voltage operation," *IEEE Journal of Solid-State Circuits*, vol. 33, pp. 592-597, Apr. 1998.

[18] D. Bingham *et al.*, "Integrated dual charge pump power supply and rs-232 transmitter/receiver," U.S. Patent 4 897 774, Jan. 1990.

[19] C. Wang and J. Wu, "Efficiency improvement in charge pump circuits," *IEEE Journal of Solid-State Circuits*, vol. 32, pp. 852-860, June 1997.

[20] P. Favrat, P. Deval and M. J. Declercq, "A high-efficiency CMOS voltage doubler," *IEEE Journal of Solid-State Circuits*, vol. 33, pp. 410-416, March 1998.

[21] M. Berkhout, G. van Steenwijk and A. J. M. van Tuijl, "A low-ripple charge pump circuit for high voltage applications," *Twenty-first European Solid-State Circuits Conference, Lille, France*, pp. 290-293, Sep. 1995.

[22] T. Duisters and E. C. Dijkmans, "A -90 dB THD rail-to-rail input opamp using a new local charge pump in CMOS," *IEEE Journal of Solid-State Circuits*, vol. 33, pp. 947-955, July 1998.

APPENDICES

APPENDIX A.

In this appendix, the minimum output voltage for a conventional 2-capacitor charge pump is derived using top-plate charge conservation.

The output of a conventional charge pump is shown in Figure 2.3. At the bottom of the ramp, the charge stored in C_x, Q_x, can be expressed as

$$Q_x = C_x \, V_{low} \; .$$

During the clock phase that is just ending, C_a was charged to V_{dd}. The charge added when the bottom plate of C_a is boosted from ground to V_{dd}, Q_a, can be expressed as

$$Q_a = C_a \, V_{dd} \; .$$

Immediately after C_a is connected to C_x, the charge stored in C_x, Q_{x2}, can be expressed as

$$Q_{x2} = C_x \, V_{high} \; .$$

The charge stored in C_a at this time, Q_{a2}, can be found to be

$$Q_{a2} = C_a(V_{high} - V_{dd}) \; .$$

Charge conservation requires that the total charge held by C_a and C_x must remain the same after C_a is connected to C_x. Thus,

$$Q_a + Q_x = Q_{a2} + Q_{x2} \; .$$

Substituting for Q_a, Q_{a2}, Q_x, and Q_{x2}, and simplifying the result leads to

$$C_x\ V_{low} = (C_a + C_x)V_{high}\ .$$

The slope of the output ramp is equal to $\frac{I_o}{(C_x+C_o)}$. The output slopes down for half of a clock period, so

$$V_{high} = V_{low} + \frac{I_o}{2f_{clk}(C_a + C_x)}\ .$$

Combining the previous two equations, it can be shown that

$$C_x\ V_{low} = (C_a + C_x)V_{low} + \frac{I_o}{2f_{clk}} - 2C_a\ V_{dd}\ .$$

Simplifying this equation leads to the result that

$$V_{low} = 2V_{dd} - \frac{I_o}{2f_{clk}\ C_a}\ .$$

Note that this result is independent of the size of capacitor C_x. This means that the reduction of the output ripple due to C_x is achieved by lowering the maximum output voltage, and not by raising the minimum output voltage.

APPENDIX B.

In this appendix, the power efficiency equations for the charge pumps are derived. The efficiency of a charge pump is defined as the power delivered to the load divided by the total power dissipated. The power delivered to the load is equal to the average output voltage, V_{out}, times the load current, I_o.

In order to calculate the total dissipated power, both the V_{dd} supply and the clock supplies must be considered. For this example, the clocks are assumed to be ideal voltage sources.

The power supplied by the clocks (ignoring parasitics) is equal to $I_o\ V_{dd}$ when the clock signal is "high" because the load current flows out of this source through the charge pump capacitor to the load. When the clock signal is "low", the power supplied is zero. There are two clocks, but only one is "high" at a time, the average power delivered by the clocks, $P_{clk} = I_o\ V_{dd}$.

To calculate the power supplied by V_{dd}, the charging of a capacitor through a resistance, shown in Figure B.1, must be considered, where $u(t)$ is a unit-step occuring at time t=0.

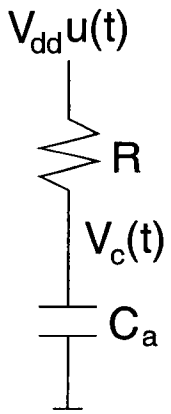

FIGURE B.1. Charging of capacitor C_a to V_{dd} through resistor R.

The voltage V_c is equal to $V_{dd} - \frac{I_o}{2f_{clk}\ C_a}$ at time $t = 0$ (where C_a is disconnected from the load and connected to V_{dd} through resistor R). This is the case because when C_a is disconnected from the load, the voltage across it has dropped by $\frac{I_o}{2f_{clk}\ C_a}$ (see Appendix A).

The energy needed by capacitor C_a to charge it from V_c to V_{dd}, E_{C_a}, can be found to be

$$E_{C_a} = \frac{1}{2} C_a\ V_{dd} \left(\frac{I_o}{2f_{clk}\ C_a} \right) .$$

This can be rewritten as

$$E_{C_a} = \frac{I_o\ V_{dd}}{4f_{clk}} .$$

Because half of the energy provided by V_{dd} is lost through the resistor R, the energy supplied per half clock period, $E_{V_{dd}}$, can be found to be

$$E_{V_{dd}} = \frac{I_o\ V_{dd}}{2f_{clk}} ,$$

making the average power, $P_{V_{dd}}$, equal to

$$P_{V_{dd}} = 2f_{clk}\ E_{V_{dd}} = I_o\ V_{dd} .$$

Notice that this power loss is independent of the resistance R.

The only other source of power loss is the charging of the bottom-plate parasitic capacitance to V_{dd} twice every clock period. The energy needed to charge the parasitic capacitor C_p to V_{dd} by an inverter is equal to $C_p\ V_{dd}^2$. Because the capacitor is charged to V_{dd} twice every clock cycle, the power dissipated, $P_{par} = 2f_{clk}\ C_p\ V_{dd}^2$.

The total power loss is then equal to $P_{clk} + P_{V_{dd}} + P_{par} = 2I_o\ V_{dd} + 2f_{clk}\ C_p\ V_{dd}^2$.

The power eficiency can be expressed as

$$\eta = \frac{I_o \, V_{out}}{2I_o \, V_{dd} + 2f_{clk} \, C_p \, V_{dd}^2} \, .$$

Adding an output capacitor does not change the power dissipated because the charge drained from C_a during every clock cycle remains the same. Adding additional stages also has no effect on the total power lost with the exception of additional bottom-plate parasitic losses, but the power provided to the load is lower for the single and double cascoded charge pumps.

APPENDIX C.

In this appendix, the factor used to estimate the difference between filtering a sawtooth waveform and a sinusoid is calculated. Determining the exact amplitude of the ripple of a charge pump that uses lowpass filtering can be difficult because the initial waveform at the input of the filter branches is a sawtooth function like the one shown in Figure 2.2. If the initial waveform was a sinusoid, each filter would have a linear effect, assuming the corner frequencies are small enough.

The sawtooth waveform has frequency components at all multiples of $(2f_{clk})$. If this waveform passes through a single-pole lowpass filter with a very low corner frequency, the higher harmonics are attenuated more than the lower ones. If the sawtooth function passes through many single-pole lowpass filters, only the fundamental will have a significant magnitude.

Using Fourier series, a sawtooth waveform with peak-to-peak amplitude of 1 can be expressed as

$$w(t) = \sum_{n}^{\infty} \frac{1}{n\,\pi} \sin(n(2\pi\ f_{clk})t)\ .$$

The fundamental is equal to $\frac{1}{\pi}\sin((2\pi\ f_{clk})t)$, and has a peak-to-peak amplitude of $\frac{2}{\pi}$. Thus, if this waveform was filtered by a large number of single-pole filters, so that only the fundamental was left, the peak-to-peak amplitude would be equal to $\frac{2}{\pi}$ times what it would be if the original waveform was a sinewave with peak-to-peak amplitude 1 and frequency $(2f_{clk})$.

If the sawtooth waveform is filtered by one single-pole lowpass filter, it is very difficult to find an expression for the difference in amplitude caused by the greater filtering of the higher harmonics. To estimate what the difference is, this situation was simulated in HSPICE. A sawtooth waveform of amplitude 1 was filtered by a

single-pole filter with a very low corner frequency. A sinusoid with a frequency equal to the fundamental of the sawtooth waveform was filtered by the same filter and the amplitudes of the filtered waveforms were compared. The sawtooth waveform had an amplitude that was about 0.83 times the amplitude of the sinusoid.

A similar simulation was done for two single-pole filters, and the result was a factor of about 0.69. For three or more single-pole filters, this factor gets close to $\frac{2}{\pi}$ which is about 0.64.

APPENDIX D.

In this appendix, the optimal capacitance ratios for maximizing the product of a group of capacitors with the total capacitance fixed is derived. The following proof is valid for all possible sum constraints for any number of capacitors.

For this example, C_{total} is fixed, and is equal to a weighted sum of n capacitors, meaning that

$$C_{total} = a_1C_1 + a_2C_2 + a_3C_3 + \cdots + a_nC_n \ .$$

It is desired to maximize the product function, $f = C_1C_2C_3 \cdots C_n$ for a fixed C_{total}.

From the first equation, it can be found that

$$C_n = \frac{C_{total} - a_1C_1 - a_2C_2 - a_3C_3 - \cdots - a_{n-1}C_{n-1}}{a_n} \ .$$

Using this result, the product function, f, can be expressed as shown in the first two equations at the top of the next page.

To find a maximum or minumum for this equation, all of the partial derivatives are found and set equal to zero. Taking the partial derivative of f with respect to C_1 leads to the third equation below.

$$f = (C_1C_2C_3 \cdots C_{n-1}) \times \left(\frac{C_{total} - a_1C_1 - a_2C_2 - \cdots - a_{n-1}C_{n-1}}{a_n} \right) \ .$$

$$f = \tfrac{1}{a_n}(C_1C_2C_3 \cdots C_{n-1})C_{total} - \tfrac{a_1}{a_n}(C_1^2C_2C_3 \cdots C_{n-1}) - \tfrac{a_2}{a_n}(C_1C_2^2C_3 \cdots C_{n-1}) - \cdots - \tfrac{a_{n-1}}{a_n}(C_1C_2C_3 \cdots C_{n-1}^2) \ .$$

$$\frac{\partial f}{\partial C_1} = \frac{1}{a_n}(C_2C_3C_4\cdots C_{n-1})C_{total} - \frac{2a_1}{a_n}(C_1C_2C_3\cdots C_{n-1}) - \frac{a_2}{a_n}(C_2^2C_3C_4\cdots C_{n-1}) - \cdots - \frac{a_{n-1}}{a_n}(C_2C_3C_4\cdots C_{n-1}^2) .$$

Setting that equal to zero and dividing by $\frac{1}{a_n}(C_2C_3C_4\cdots C_{n-1})$ leaves

$$C_{total} - 2a_1C_1 - a_2C_2 - a_3C_3 - \cdots - a_{n-1}C_{n-1} = 0 ,$$

or equivalently,

$$C_{total} = 2a_1C_1 + a_2C_2 + a_3C_3 + \cdots + a_{n-1}C_{n-1} .$$

Comparing this with the first equation leads to the result that $a_1C_1 = a_nC_n$.

Taking the partial derivatives of f with respect to the other capacitors is similar, and leads to the final result that the maximum (or minimum) of f lies where $a_1C_1 = a_2C_2 = a_3C_3 = \cdots = a_nC_n$.

To find out if this point is a maximum or a minimum, the second partial derivatives of f are taken, for example

$$\frac{\partial^2 f}{\partial C_1^2} = -\frac{2a_1}{a_n}(C_2C_3C_4\cdots C_{n-1}) < 0 .$$

Similarly, it can be shown that all of the second partial derivatives of f are negative, indicating that the previous result is a maximum.

Using this result, it is easy to determine the optimal capacitor ratios for any charge pump where the ripple is inversely proportional to the product of a number of capacitors. This is the case for any charge pump where all branches act as filters. For the conventional charge pump, the ripple is inversely proportional to the product C_aC_x, and the fixed sum is equal to $2C_a + C_x$, so the ripple is minimized when

$2C_a = C_x$. Similarly, for the single cascode charge pump, the ripple is minimized when $2C_a = C_x = C_o$, and for the double cascode charge pump, the ripple is minimized when $2C_a = C_x = C_{o1} = C_{o2}$.

APPENDIX E.

In this appendix, the small-signal equivalent circuits for single cascode and double cascode charge pumps are derived from the simple small-signal models of the buffer transistors. The derived circuits do not contain any dependent current sources, making the behavior of the circuits more clear.

Derivation of the small-signal equivalent circuit for the single cascode charge pump from the small-signal model shown in Figure E.1:

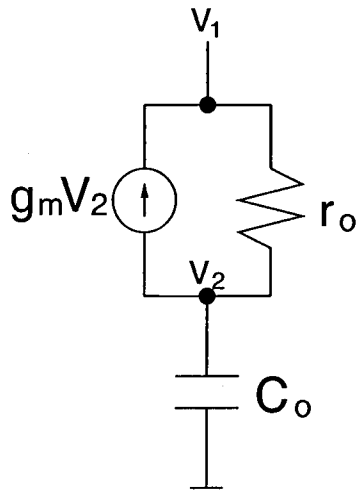

FIGURE E.1. Small-signal model for a single buffer transistor and output capacitor.

Summing the currents leaving V_2 (in s domain) gives

$$sC_o\ V_2 + g_m\ V_2 + \frac{V_2}{r_o} - \frac{V_1}{r_o} = 0\ .$$

Multiplying by r_o gives

$$sr_o\ C_o\ V_2 + (g_m\ r_o + 1)V_2 - V_1 = 0\ .$$

Thus,

$$\frac{V_2}{V_1} = \frac{1}{g_m\, r_o + 1 + s r_o\, C_o} = \frac{1}{g_m\, r_o + 1} \times \frac{1}{1 + \frac{s r_o\, C_o}{g_m\, r_o + 1}} .$$

Assuming that $g_m\, r_o >> 1$,

$$\frac{V_2}{V_1} \approx \frac{1}{g_m\, r_o} \times \frac{1}{1 + \frac{sC_o}{g_m}} .$$

This is just an attenuation by $g_m\, r_o$ followed by a first-order RC filter with time constant $\frac{C_o}{g_m}$. The impedance seen from V_1 is equal to the voltage at V_1 divided by the current through C_o, which is equal to $\frac{V_1}{sC_o\, V_2} \approx r_o + \frac{g_m\, r_o}{sC_o}$, so the resistor must be r_o and the capacitor $\frac{C_o}{g_m\, r_o}$. For the complete small-signal equivalent circuit shown in Figure 2.8, resistor $r_{on(M_3)}$ and capacitor C_x are added.

Derivation of the small-signal equivalent circuit for the double cascode charge pump From the small-signal model shown in Figure E.2:

Summing the currents leaving V_3 (in s domain) gives

$$sC_{o2}\, V_3 + g_{m2}\, V_3 + \frac{V_3}{r_{o2}} - \frac{V_2}{r_{o2}} = 0 .$$

Multiplying by r_{o2} gives

$$s r_{o2}\, C_{o2}\, V_3 + (g_{m2}\, r_{o2} + 1) V_3 - V_2 = 0 .$$

Thus,

$$V_2 = V_3 (g_{m2}\, r_{o2} + 1 + s r_{o2}\, C_{o2}) .$$

Summing the currents leaving V_2 (in s domain) gives

$$sC_{o1}\, V_2 + g_{m1}\, V_2 + \frac{V_2}{r_{o1}} - \frac{V_1}{r_{o1}} + sC_{o2}\, V_3 = 0 .$$

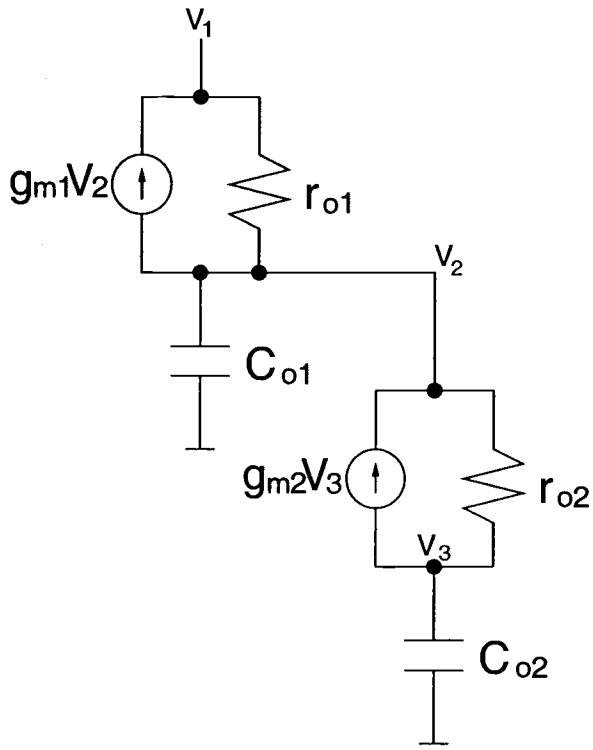

FIGURE E.2. Small-signal model of two cascoded buffer transistors with two output capacitors.

Multiplying by r_{o1} gives

$$sr_{o1}\ C_{o1}\ V_2 + (g_{m1}\ r_{o1} + 1)V_2 - V_1 + sr_{o1}\ C_{o2}\ V_3 = 0\ .$$

Substitution for V_2 leads to the equations shown at the top of the next page. These equations lead to the transfer function of the circuit shown in Figure E.2 (last equation below).

$$(g_{m1}\ r_{o1} + 1 + sr_{o1}\ C_{o1})(g_{m2}\ r_{o2} + 1 + sr_{o2}\ C_{o2})V_3 - V_1 + sr_{o1}\ C_{o2}\ V_3 = 0\ .$$

$$V_1 = ((g_{m1}\ r_{o1} + 1)(g_{m2}\ r_{o2} + 1) + s(r_{o1}\ C_{o1}(g_{m2}\ r_{o2} + 1) + r_{o1}\ C_{o2} + r_{o2}\ C_{o2}(g_{m1}\ r_{o1} + 1)) + s^2(r_{o1}\ r_{o2}\ C_{o1}\ C_{o2}))V_3\ .$$

Assuming $g_{m1}\ r_{o1} >> 1$ and $g_{m2}\ r_{o2} >> 1$,

$V_1 \approx (g_{m1}\ r_{o1})(g_{m2}\ r_{o2}) + s(r_{o1}\ C_{o1}(g_{m2}\ r_{o2}) + r_{o1}\ C_{o2} + r_{o2}\ C_{o2}(g_{m1}\ r_{o1})) + s^2(r_{o1}\ r_{o2}\ C_{o1}\ C_{o2}))V_3$.

$$V_1 \approx g_{m1}\ r_{o1}\ g_{m2}\ r_{o2}\left(1 + s\left(\frac{C_{o1}}{g_{m1}} + \frac{C_{o2}}{g_{m1}\ g_{m2}\ r_{o2}} + \frac{C_{o2}}{g_{m2}}\right) + s^2\left(\frac{C_{o1}\ C_{o2}}{g_{m1}\ g_{m2}}\right)\right)V_3 .$$

$$\frac{V_3}{V_1} \approx \frac{1}{g_{m1}\ r_{o1}\ g_{m2}\ r_{o2}} \times \frac{1}{\left(1 + s\left(\frac{C_{o1}}{g_{m1}} + \frac{C_{o2}}{g_{m1}\ g_{m2}\ r_{o2}} + \frac{C_{o2}}{g_{m2}}\right) + s^2\left(\frac{C_{o1}\ C_{o2}}{g_{m1}\ g_{m2}}\right)\right)} .$$

This is just an attenuation by $g_{m1}\ r_{o1}\ g_{m2}\ r_{o2}$ followed by two RC branches. For the filter shown in Figure E.3, the transfer function is

$$\frac{V_{out}}{V_{in}} = \frac{1}{1 + s(R_1\ C_1 + R_1\ C_2 + R_2\ C_2) + s^2(R_1\ R_2\ C_1\ C_2)} .$$

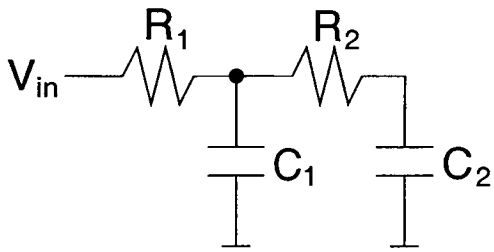

FIGURE E.3. Second order RC lowpass filter.

By inspection, and noting that R_1 must be r_{o1}, the other elements can be determined to be:

$$C_1 = \frac{C_{o1}}{g_{m1}\ r_{o1}} .$$
$$C_2 = \frac{C_{o2}}{g_{m1}\ r_{o1}\ g_{m2}\ r_{o2}} .$$
$$R_2 = r_{o2}(g_{m1}\ r_{o1}) .$$

For the complete small-signal equivalent circuit shown in Figure 2.18, resistor $r_{on(M_3)}$ and capacitor C_x are added.