

DESIGN OF BUCK SWITCHING REGULATORS WITH INTEGRATED
FILTER FOR LOW-POWER INTEGRATED CIRCUITS

by

Patrick Webster Griffin

A thesis submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Electrical and Computer Engineering

The University of Utah

May 2011

Copyright © Patrick Webster Griffin 2011

All Rights Reserved

The University of Utah Graduate School

STATEMENT OF THESIS APPROVAL

The thesis of Patrick Webster Griffin
has been approved by the following supervisory committee members:

<u>Reid R. Harrison</u>	, Chair	<u>07/09/2010</u> Date Approved
<u>Cynthia M. Furse</u>	, Member	<u>07/09/2010</u> Date Approved
<u>Faisal H. Khan</u>	, Member	<u>07/09/2010</u> Date Approved

and by Gianluca Lazzi, Chair of
the Department of Electrical and Computer Engineering

and by Charles A. Wight, Dean of The Graduate School.

ABSTRACT

The purpose of this thesis was to determine if low-power switching power supplies can be made on-chip using integrated components. Integrated switching supplies are an emerging field that has followed the rise of systems-on-chip devices – especially in the biomedical field. Switching supply theory and implementation were examined systematically to determine the feasibility of such switching supplies.

Classical switching power supply theory was presented first, including fundamental principles of operation and essential analysis techniques. Due to the unique constraints placed on integrated power supplies as a result of the small component size, the classical treatment had to be updated and modified. The result was a new methodology for calculating ripple current and voltage, circuit losses, and efficiency of switching supplies in both continuous and discontinuous conduction modes. Integrated and micro-scale switching supply components were then examined. Most importantly, the design of integrated inductors was discussed. Double-layer coils were found to be the best choice for integrated inductors with a small number of coils as they offered four times the inductance and only twice the resistance of similar single-layer coils.

Six boards were tested using a variety of loads with manual switching cycle control. The test boards effectively modeled the behavior of integrated supplies and confirmed predictions about power loss and transfer. Using the test results and the equations previously derived, three test cases were simulated. The results were

efficiencies of 75.16%, 75.09%, and 75.10% using 2 and 5 turn double spirals, and an external 120 nH coil, respectively. With these results, it should be possible to build integrated switching power supplies that meet or exceed the efficiency of linear supplies.

TABLE OF CONTENTS

ABSTRACT.....	iii
LIST OF TABLES.....	vii
ACKNOWLEDGEMENTS.....	viii
Chapter	
1 INTRODUCTION.....	1
1.1 Linear Supplies.....	2
1.2 Switching Supplies.....	3
1.3 Purpose of Thesis.....	4
2 INDUCTIVE SWITCHING SUPPLY THEORY.....	5
2.1 Basic Switching Theory.....	5
2.1.1 The Small Ripple Approximation.....	8
2.1.2 Volt-Second Balance.....	9
2.1.3 Capacitor Charge Balance.....	13
2.2 Sources of Inefficiency.....	17
2.2.1 Nonideal Reactive Components.....	18
2.2.1.1 Inductors.....	18
2.2.1.2 Capacitors.....	25
2.2.2 Real Semiconductor Devices.....	27
2.2.2.1 MOSFETs.....	28
2.2.2.2 Diodes.....	33
2.2.3 Complete Loss Models.....	35
2.2.3.1 The Static Model.....	35
2.2.3.2 The Dynamic Model.....	40
2.3 Discontinuous Conduction Mode.....	44
2.3.1 Discontinuous Mode Conditions.....	46
2.3.2 Implications of Discontinuous Conduction.....	49
2.4 Contemporary Integrated Designs.....	56
3 SURVEY OF KEY COMPONENT TECHNOLOGIES.....	58

3.1	Components.....	58
3.1.1	Integrated Inductors.....	59
3.1.1.1	Spiral Inductors.....	60
3.1.1.2	Three-Dimensional Inductors.....	67
3.1.2	Chip Inductors.....	69
3.1.3	Integrated Capacitors.....	74
3.1.4	Chip Capacitors.....	79
3.1.4.1	Class I – C0G, NP0, C0H, C0J, C0K.....	80
3.1.4.2	Class II – XR5, X7R, XJ5, X75.....	81
3.1.4.3	Class III – Z5U, Y5V.....	81
3.1.5	Integrated Power Switches.....	82
3.1.6	Discrete Power Switches.....	88
4	SIMULATIONS, CALCULATIONS, AND OPTIMIZATIONS.....	94
4.1	Test Circuits.....	96
4.1.1	Initial Testing and Observations.....	97
4.2	Test Board Results.....	99
4.3	Extrapolation from Board Results.....	109
4.4	Integrating the Components.....	110
4.4.1	Power Switches.....	111
4.4.2	Inductors.....	112
4.4.3	Output Capacitor.....	114
4.5	Test Designs.....	115
4.5.1	Double Spiral 1 – 600 μm x 600 μm , two turns.....	116
4.5.2	Double Spiral 2 – 600 μm x 600 μm , five turns.....	117
4.5.3	External Inductor – Murata LQF1SHS, EIA 0402, 120nH... ..	117
5	CONCLUSIONS AND FUTURE WORK.....	118
5.1	Conclusions.....	118
5.2	Future Work.....	119
	APPENDIX: MATLAB CODE.....	121
	BIBLIOGRAPHY.....	123

LIST OF TABLES

Table	Page
1. Accuracy of linear current approximation for rising current.....	47
2. Coefficients for spiral coil inductance formula.....	54
3. Capacitance and dissipation factor (DF) for selected sizes and dielectrics.....	74
4. Comparison of SMT MOSFET package types and sizes.....	82
5. SOT323 P-channel transistors.....	84
6. Board assembly parts.....	88
7. MOSFET on-resistances and threshold voltages.	89

ACKNOWLEDGEMENTS

I would like to thank Dr. Harrison, the chair of my committee, for the use of his lab and for his patient and tireless guidance and advice. I would also like to thank the current and former members of my committee – Dr. Cynthia Furse, Dr. Faisal Khan, Dr. Cameron Charles, and Dr. Bradley Greger.

I would also like to thank my fellow students Ryan Kier, Brandon Thurgood, and Keith Tracey. Their support and advice, as well as their friendship, were key to my success.

I especially want to thank my supervisors and squadron commanders at the 84 Radar Evaluation Squadron for giving me the time and flexibility to complete my masters degree while continuing to work for the Air Force.

Lastly, I want to thank my wife, Erin, and my two sons, Timothy and Nathaniel, for giving up their husband and father these last few years. Without their love and constant support, I never would have been able to finish the long path I've followed.

CHAPTER 1

INTRODUCTION

Integrated circuit technology has seen an unprecedented advancement since the first integrated circuits were built in 1958. Process technologies have shrunk, transistor counts have increased, and power consumption has dropped. As a consequence, applications have been developed to take advantage of these advancements – applications which were unimaginable in 1958. One such application is the University of Utah Integrated Neural Interface (INI) – an implantable medical device for recording neural signals [1].

Implantable medical devices in general pose some unique problems to engineers. The devices must be small, reliable, biocompatible, and often must use very little power. Each of these concerns is an area of continued research and development for medical devices – and for integrated circuits in general – but this project specifically focuses on reducing power consumption by using a new type of power supply circuit.

A typical power supply for an implantable medical device consists of a power source, such as a battery or power receiving coil [2], and a voltage regulator. Previous versions of the INI chip used a linear voltage regulator to lower and regulate the incoming voltage from a power receiving coil [1]. While this system worked well, it might be possible to make an even better power supply.

All power supplies perform the same basic function: They take power from a source, perform some transformation of voltage or current, and output power at the output voltage or current. The quality of the power supply depends on how well the transformation is achieved.

1.1 Linear Supplies

Linear supplies are a class of power supply in which the output voltage is achieved using a variable resistance under linear feedback control. The variable resistance together with the load resistance forms a voltage divider with an output voltage determined by the ratio of the two resistances [see Figure 1(a)]. This type of supply is typically very easy to build, consisting of a variable resistor (typically a series pass transistor), an error-amplifier, and a voltage reference, as shown in Figure 1(b).

Linear supplies are capable of good load regulation, low noise, small size, and good transient response [3], but they have a serious flaw. Because all of the output current must flow through the transistor, the efficiency is always equal to the ratio of the output voltage to the input voltage: $\eta = V_{out}/V_{in}$. If the input voltage is very close to the

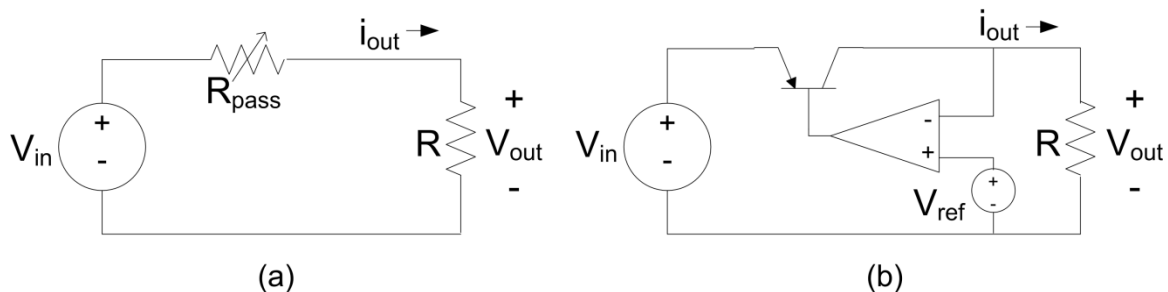


Figure 1: Linear power supply models.

output voltage, this is not a problem as the efficiency is high. However, if the input voltage is *not* close to the output voltage, the efficiency is low.

Obviously, high efficiency is important in an implantable device due to the low available power. Neural interfaces add a different dimension as well: heat. Heat generated by implantable neural devices is discussed at length in [4] where the American Association of Medical Instrumentation (AAMA) is quoted as allowing a chronic temperature increase of only 1 to 2 °C. Inefficient power supplies make this goal harder to meet due to the heat generated by the series pass transistor.

1.2 Switching Supplies

The other major class of power supplies is switching power supplies. There are many types of switching power supply such as step-down, step-up, inverting, isolation, step-up-down, and many others. All work on the same principle: using reactive components and AC signals to convert voltage and current [5], [6]. A basic switching power supply which performs the step-down voltage conversion function is shown in Figure 2. This circuit, in which the switch is varied between positions 1 and 2, will be discussed at length in the rest of the thesis.

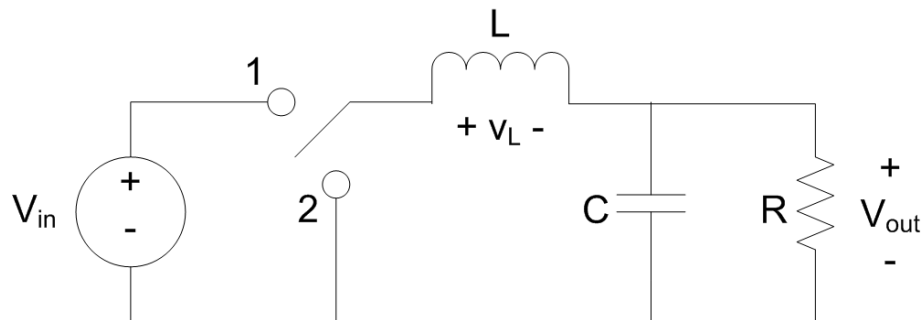


Figure 2: Ideal step down switching power supply

1.3 Purpose of Thesis

The main purpose of this thesis is to determine if a suitable step-down switching supply can be integrated into low-power chips such as the INI chip. As will be seen, low output power poses some interesting challenges in power supply design. It may be that all components of the switching supply can be integrated in silicon, or it could be that some components will need to be integrated just off-chip on top of the silicon or as part of the carrier package.

Previous attempts have been made to fully integrate a switching supply in silicon [7], [8], [9]. These previous efforts, all by the same group, were to build a high-power supply capable of delivering current greater than 100mA.

The theory of switching supplies will be reviewed and additional design equations will be derived as needed. The components of an integrated switching supply will be examined both for external and internal integration. Finally, the theory will be tested to determine the feasibility of integrating switching power supplies.

CHAPTER 2

INDUCTIVE SWITCHING SUPPLY THEORY

There are many different types of switching power supply. All rely on reactive components to regulate the output voltage and, depending on the circuit type, can step-down, step-up, invert, isolate, or perform almost any other imaginable conversion. The most diversity of supplies is in the type that use inductors and capacitors to regulate the voltage. These so-called inductive switching supplies (also called inductive switching converters), and in particular the step-down inductive supply, are what will be discussed in this chapter.

It should be noted that there is a class of switching converter that uses only capacitors. These are typically called charge pumps due to their mode of operation: moving charge from one voltage to another via the switching of capacitors. These circuits have uses in some very simple unregulated supplies, but will not be discussed further in this thesis. From this point forward, switching supplies will refer to the inductive switching type.

2.1 Basic Switching Theory

The most basic switching power supply that can be built is the step-down inductive supply, also known as a buck converter. The ideal model of a buck converter is

shown in Figure 3 and will be used to discuss basic switching theory that is applicable to all switching power supplies.

The switch in Figure 3 is alternated between the two contacts, labeled 1 and 2, to create the time varying switching waveform $v_s(t)$ as shown in Figure 4. Initially at time $t=0$, the switch is in position 1 and V_{in} is applied to the LC filter. At time $t=T_1$, the switch alternates to position 2 and the LC filter is grounded. At time $t=T_S$, the cycle repeats itself, generating a square wave with frequency $f_{sw} = 1/T_S$. The ratio of the time spent in position 1 divided by the total period is the duty cycle D .

The switching voltage $v_s(t)$ is connected to the LC filter and the equivalent load resistance R . The ideal switch in Figure 4 exhibits zero voltage drop and therefore dissipates no power. The reactive components, L and C , store energy proportional to their current and voltage, respectively, but do not dissipate any energy. As a result, the ideal switching regulator has efficiency equal to 100%.

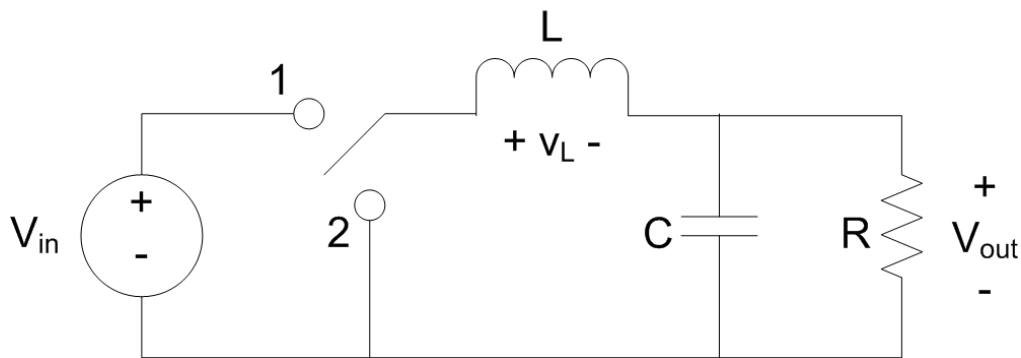


Figure 3: The ideal buck converter. The single pole, double throw switch is alternated between the two contacts to vary the output voltage. The load current is modeled by the load resistor R .

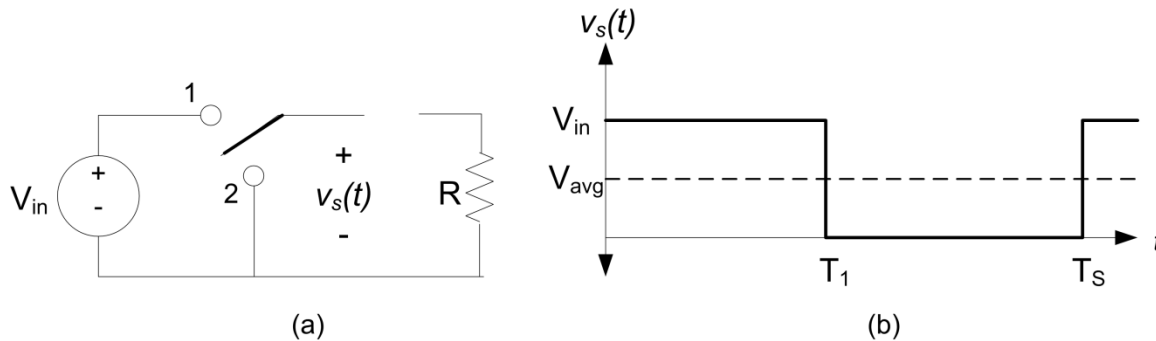


Figure 4: Switching voltage as seen by the output filter. View (a) has the LC filter omitted for clarity.

In reality, the components all have resistance and parasitic reactance that limit the efficiency of a switching regulator. These losses can generally be divided into frequency-independent, resistive losses (also referred to as static losses) and frequency-dependent, switching (also known as dynamic) losses. These two sources of loss and their effect on circuit performance are discussed after the ideal converter has been presented.

It should be noted that the LC network forms a low pass filter for the switching voltage. Obviously, it is desirable to have a perfectly flat output voltage, V_{out} , without any ripple induced by the switching action of the power supply. This can only be achieved by filtering out the harmonics present in the switching waveform. What is left after filtering is the DC component of the switching voltage plus the attenuated switching frequency and its harmonics. To put it mathematically, the output voltage is given by $v_{out}(t) = V_{out} + v_{ripple}(t)$, where V_{out} is equal to the DC component of the switching voltage and $v_{ripple}(t)$ is the attenuated AC portion of the switching voltage.

We know from Fourier analysis that the DC component of a signal is equal to its average value. Fortunately, this value is easy to calculate for a square wave with maximum value V_{in} and period T_S .

$$\langle v_s(t) \rangle = V_{in} \frac{T_1}{T_S} \quad \text{Equation 1}$$

If the duty cycle, D , is defined as the time spent in position 1, T_1 , divided by the total period T_S , then the output voltage of a buck converter with perfect low pass LC filter is given by equation 2 [5].

$$V_{out} = V_{in}D \quad \text{Equation 2}$$

Therefore, the output voltage of the buck converter can be controlled by changing the duty cycle of the switching waveform. Indeed, this is the way that the controller in a switching power supply alters the voltage to bring it to the desired value.

2.1.1 The Small Ripple Approximation

As was mentioned in the previous section, the output voltage of any real converter is equal to a DC component plus a superimposed AC component. For any well-designed converter, the ripple component is much smaller than the DC output voltage – typically less than 1%. As long as this assumption is true, the equations describing the switching voltages and currents can be greatly simplified, as described in the following sections about the volt-second balance and the capacitor charge balance.

One may wonder why the approximation is needed. For the case of the ideal buck converter, it is possible to find a closed solution for individual voltages and currents. However, when nonideal components are used as in Section 2.2 or when a more complicated switching topology is used, it quickly becomes cumbersome or even

impossible to find the closed form solution to these quantities. By contrast, if the circuit designer wishes to design a good converter, small ripple is not only a computational convenience, it is a design constraint that applies to all topologies.

2.1.2 Volt-Second Balance

Consider Figure 5(a) in which the buck converter has been redrawn for $0 < t < T_1$ and Figure 5(b) where it is drawn for $T_1 < t < T_s$. Because of the small ripple approximation, the output voltage is considered constant and therefore, the inductor voltage is also constant. In both cases, the inductor voltage $v_L(t)$ can be found by inspection, as shown in equation 3.

$$\begin{aligned} v_L(t) &= V_{in} - V_{out} & 0 < t < T_1 \\ v_L(t) &= -V_{out} & T_1 < t < T_s \end{aligned} \quad \text{Equation 3}$$

We know from any introductory circuits class that the voltage developed by an inductor is given by the equation $v(t) = L \frac{di}{dt}$, so if the voltage across an inductor is constant, the current changes at a constant rate equal to the voltage divided by the inductance [5]. This is called the linear slope approximation and it holds true as long as the inductor is nearly ideal and the output ripple is small.

Equations for both the ripple current and the DC current need to be derived for several reasons. First, the combination of the two currents needs to be found so that the physical components are sized to handle the maximum current. Second, the ripple

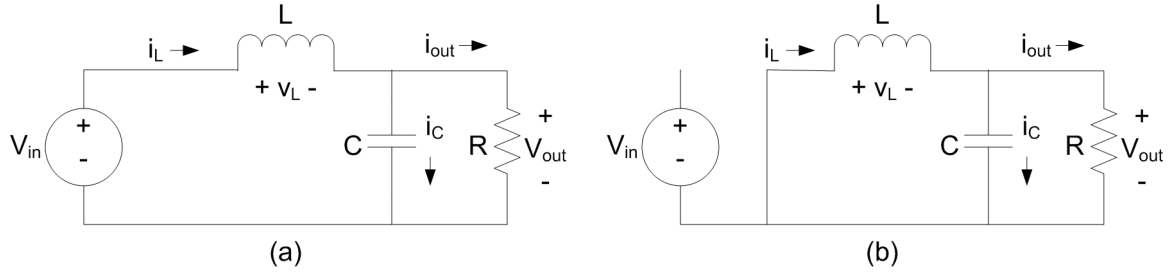


Figure 5: Buck converter when the switch is in position 1 (a) and in position 2 (b).

current should be kept as small as possible by proper choice of components so as not to stress them or to create additional EMI in nearby components.

The inductor voltage and current are shown in Figure 6. The current starts at some initial value $i_L(0)$ and increases linearly up to the point where the switch changes position at $t = DT_s$. Once the switch is in position 2, the current decreases linearly until the end of the switching period. The slopes of the current are equal to the inductor voltages divided by the inductance.

$$\frac{di}{dt} = \frac{V_{in} - V_{out}}{L}, \quad 0 < t < DT_s$$

$$\frac{di}{dt} = \frac{-V_{out}}{L}, \quad DT_s < t < T_s$$

Equation 4

The peak-to-peak ripple of the inductor current can be calculated very easily by noting that the inductor current starts at its minimum value and increases linearly to time DT_s , at which point it is at its maximum value. Since we know the slope and time, we can derive the current rise as shown in equation 5.

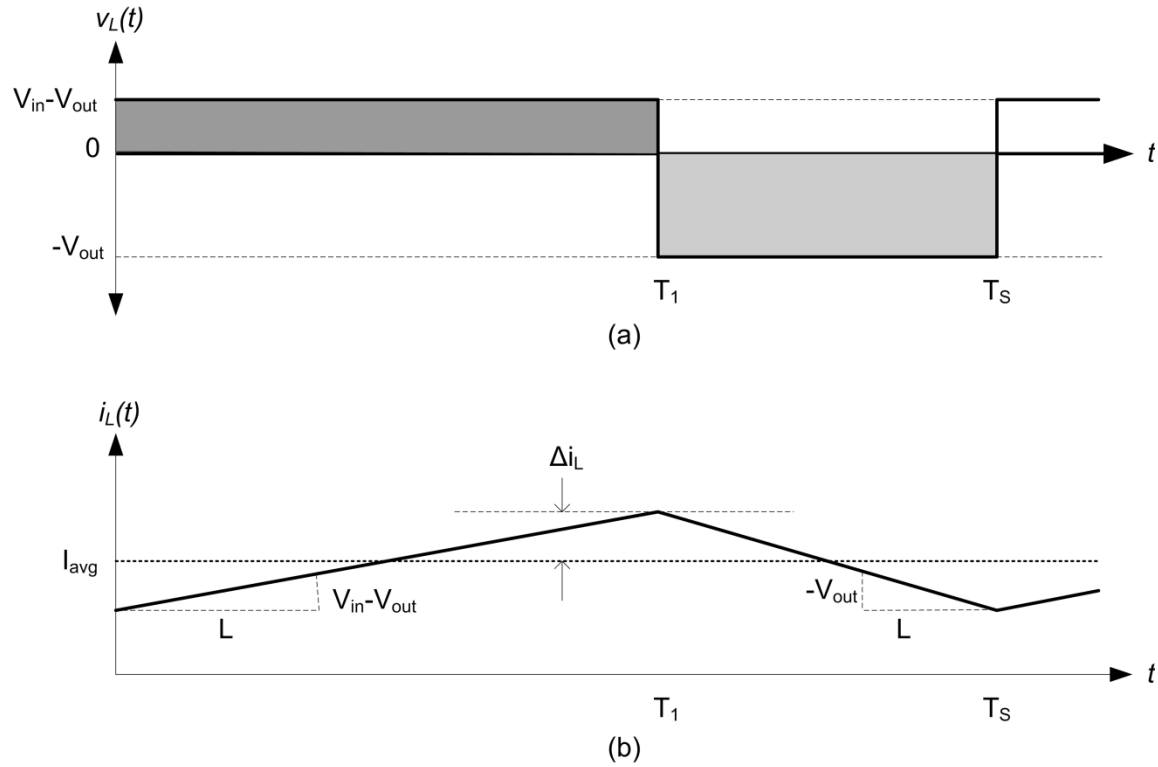


Figure 6: Inductor voltage and current. The shaded areas in subfigure (a) are equal, demonstrating the volt-second balance. The current in subfigure (b) is linear with slopes shown.

In equation 5, Δi_L is the maximum difference between the total current and the average current. Ideally, this value should be kept as small as possible for the aforementioned reasons. Equation 5 can be rewritten to solve for the inductor as shown in equation 6 which is useful for specifying the inductance needed for a given ripple current [5].

$$2\Delta i_L = \frac{V_{in} - V_{out}}{L} DT_S \quad \text{Equation 5}$$

$$L = \frac{V_{in} - V_{out}}{2\Delta i_L} DT_S \quad \text{Equation 6}$$

In steady-state, the output voltage and current should not change – this is what is meant by steady-state. Inspection of Figure 5 shows that if the output current does not change, then the average inductor current should not change either. This constraint leads to the principle of the volt-second balance.

To derive the volt-second balance, first remember that the inductor voltage is given by the differential equation, $v_L(t) = L \frac{di_L(t)}{dt}$. Separating and integrating over one switching cycle from $t = 0$ to $t = T_S$ yields equation 7.

$$i_L(T_S) - i_L(0) = \frac{1}{L} \int_0^{T_S} v_L(t) dt \quad \text{Equation 7}$$

Equation 7 states that the change in current for the switching period is proportional to the voltage-time product. Since we know that the total current change in the steady-state is equal to zero, the equation can be rewritten as equation 8.

$$0 = \frac{1}{T_S} \int_0^{T_S} v_L(t) dt = \langle v_L(t) \rangle \quad [5] \quad \text{Equation 8}$$

Both sides of the equation have also been divided by T_S so that the right side of the equation is equal to the average value of $v_L(t)$. The integral has units of volt-seconds, the sum of which balances to zero, hence the name “volt-second balance.” This can be seen graphically in Figure 6(a) where the area under the voltage curve from $t = 0$ to $t = DT_S$ is equal to the area above the voltage curve from $t = DT_S$ to $t = T_S$.

2.1.3 Capacitor-Charge Balance

Just as the steady-state assumptions led to the volt-second balance for inductor voltage, similar assumptions lead to the capacitor-charge balance. The capacitor-charge balance states that the net current flow in one switching cycle is zero in the steady-state. Qualitatively, this makes sense since a net current into or out of a capacitor will change its cycle-to-cycle voltage which violates the constraints of steady-state operation. Quantitatively, this is derived by starting with the capacitor's differential voltage and current relationship, separating, and integrating over one switching cycle as shown in equation 9.

$$i_C(t) = C \frac{dv_C}{dt}$$

$$v_C(T_S) - v_C(0) = \frac{1}{C} \int_0^{T_S} i_C(t) dt$$

$$0 = \int_0^{T_S} i_C(t) dt \quad \text{Equation 9}$$

Note that the integral term has units of Amp-seconds or simply Coulombs, hence the name capacitor-charge balance. As before, the left-hand side of the equation is zero because of the steady-state conditions. To restate this qualitatively, the total change of the charge on the capacitor is zero over one switching cycle [5].

$$0 = \frac{1}{T_S} \int_0^{T_S} i_C(t) dt = \langle i_C(t) \rangle \quad \text{Equation 10}$$

The charge balance equation can be rewritten in the alternate form of equation 10 where it is shown that the average current through the capacitor is zero. Both equations 10 and 11 are demonstrated by Figure 7 where the average current is equal to zero and the shaded areas representing charge are equal.

There are a few other key points to note about Figure 7. First, note that the capacitor current is equal to the inductor ripple current. This means that when the inductor current is at a minimum, the capacitor is supplying current to the load equal to the average current minus the inductor ripple current. Similarly, when the inductor current is at its maximum, the extra current is recharging the capacitor. Also note that the current and voltage waveforms are approximately 90 degrees out of phase such that the capacitor voltage is at its minima and maxima where the current crosses zero.

Note that while the voltage ripple appears large in Figure 7, it should in fact be small compared to V_{out} . This is assured by selecting a capacitor such that the ripple charges (the shaded areas) are small compared to the average charge stored in the capacitor. It is therefore important to find a way to estimate the value for a capacitor given a certain ripple current and desired ripple voltage.

This can be done by first assuming we have already picked a large enough capacitor such that the small ripple assumption holds true. We know from Figure 7 that the maximum voltage swing will occur between the zero crossing of the current and we know that the inductor current is approximately linear (because of the small ripple assumption). We can also see that the current zero crossings are separated by exactly one half cycle because the current waveform is symmetrical.

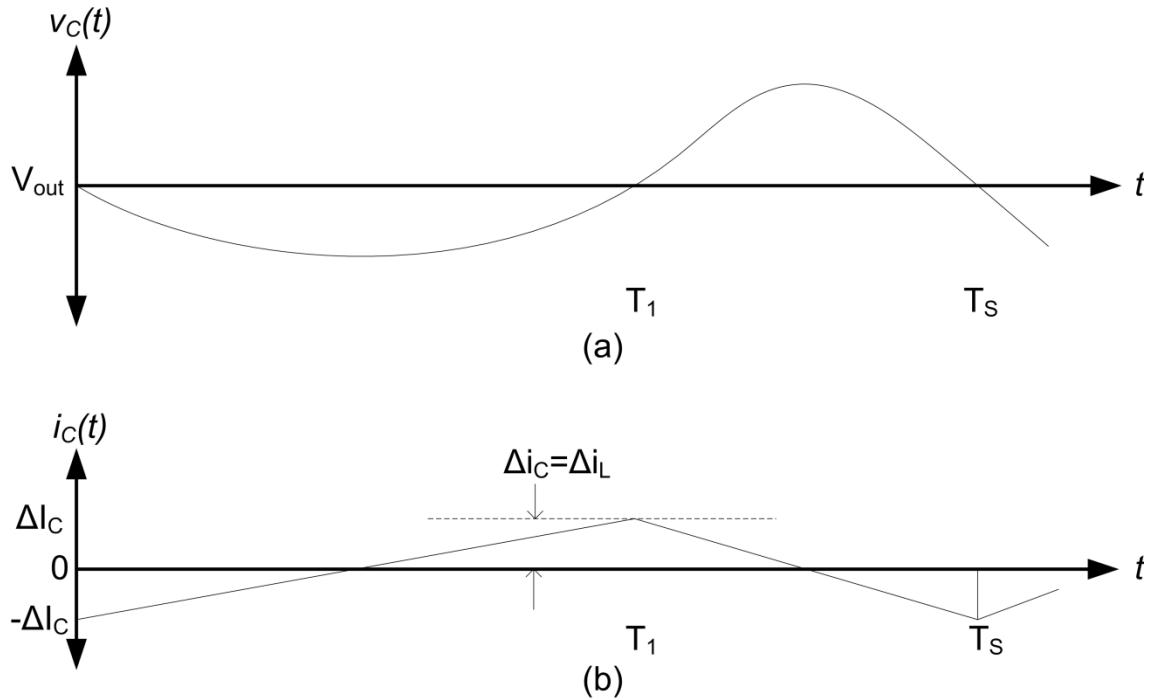


Figure 7: Capacitor voltage and current during one switching cycle. Note that the shaded areas are equal.

These conditions allow us to calculate the total charge transferred between zero crossings. This value is equal to the area in the triangle under the current curve with base equal to $0.5T_s$ and height of Δi_C . Therefore, the total charge transferred in either direction is given by equation 11 [5].

$$\Delta Q = \frac{1}{4} T_s \Delta i_C. \quad \text{Equation 11}$$

Assuming that the capacitor acts like an ideal capacitor, its voltage is linearly related to the stored charge by the factor of its capacitance. Therefore, the voltage change $v_{pp-ripple}$ is given by equation 12.

$$v_{pp-ripple} = \frac{1}{C} \Delta Q = \frac{1}{4C} T_S \Delta i_C. \quad \text{Equation 12}$$

Note that this is the peak-to-peak value for the voltage ripple, not the ripple amplitude. As was already stated, the capacitor ripple current is equal to the inductor ripple current. Substituting Δi_L for Δi_C , making use of the relationship between duty cycle and input/output voltage ratios, and converting period to frequency,

$$v_{pp-ripple} = \frac{D(V_{in}-V_{out})}{8LC} T_S^2, \quad D = \frac{V_{out}}{V_{in}}, \quad f_S = \frac{1}{T_S}.$$

$$v_{pp-ripple} = \frac{V_{out} - \frac{V_{out}^2}{V_{in}}}{8LC f_S^2} \quad \text{Equation 13}$$

So it can be seen that the output voltage ripple is reduced for larger reactance values and higher frequency, and is increased for higher output-to-input voltage ratios. This should intuitively make sense since the LC network of a buck converter is a low-pass filter for the switching voltage.

Generally speaking, one would use equation 6 to determine the inductor value such that the ripple current is small and then solve for a suitable capacitor using equation 13 to ensure that the ripple voltage is small. Note that these equations assume that the output current to the load is large enough that the inductor is constantly conducting a current much larger than the ripple current, thereby making the ripple current linear during the entire switching cycle. This assumption does not hold true in the discontinuous mode discussed in a later section.

2.2 Sources of Inefficiency

There are numerous sources of inefficiency in a switching supply. Every component is a potential source of one or more nonidealities that contribute to the overall inefficiency of a converter. Component resistance, stray reactance, semiconductor losses, and even the wiring between components steal energy from the circuit which would otherwise be transferred to the load.

Generally, these losses can be categorized into static losses and dynamic losses based on the mode of power loss. Static losses are directly related to the current through a device whereas dynamic losses are directly related to the switching frequency. For high-current, high-power devices, static losses will usually dominate the total loss in a converter. In low-current, low-power devices, the dynamic losses become a significant part of or even the majority of the total inefficiency.

Static losses are caused by the resistance of the materials that make up the inductor, capacitor, switches, and the wiring between the devices. The resistive power lost in some component is given by $P_j = I_j V_j$. The total power lost due to resistance in all the components is then: $P_{resistive} = \sum I_j V_j$. These losses can be mitigated by using higher quality inductors and capacitors or by using lower resistance semiconductor devices.

Dynamic losses are caused by switching the current, the voltage, or both the current and voltage across and through a device. These losses are seen anywhere a capacitance or inductance stores energy that is not transferred to the load. There are other places where these losses are seen which are discussed in the following section. Since a

major source of dynamic losses is stray capacitance, smaller components generally reduce the dynamic loss of a converter.

Higher quality and lower resistance devices are generally larger than their higher resistance counterparts. This is in direct contrast to components designed for lower dynamic losses which are generally smaller to reduce stray inductance and capacitance. Solving this dilemma between the two loss mechanisms is a major challenge, especially as switching frequencies increase. The major sources of inefficiency are presented in the following sections to help solve this problem.

2.2.1 Nonideal Reactive Components

Thus far, the reactive components were assumed to be ideal – capacitors were purely capacitive and inductors were purely inductive. In reality, reactive components always have some resistance and some small stray reactance. Inductors in particular have significant resistive losses and a nonnegligible parallel capacitance. Capacitors do have some small stray inductance, but this is usually negligible and is rarely even included in data sheets. Some of these nonidealities are now added to each component to obtain a more complete model.

2.2.1.1 Inductors

Resistance in the inductor is often referred to as copper loss since power inductors are usually constructed of low-resistivity copper wire [5]. As shown in Figure 8 where the two halves of the cycle are shown, this resistance is modeled by an equivalent series

resistance R_L . Depending on the wire gauge and number of turns, this value can be anywhere from a few milliohms to several tens of ohms.

Before numerically analyzing Figure 8, there are a few insights that should be made. The steady-state and small ripple assumptions still hold which means that the combined voltage across the inductor and resistor is still the same as for the ideal case. However, because of the current flowing through the resistor, the inductive voltage V_L will vary as the current changes through the resistor.

Inspection of Figure 8 leads to equation 14 for the inductor voltage. Note that the absolute voltages before and after time DT_S are lower due to the resistance of the inductor.

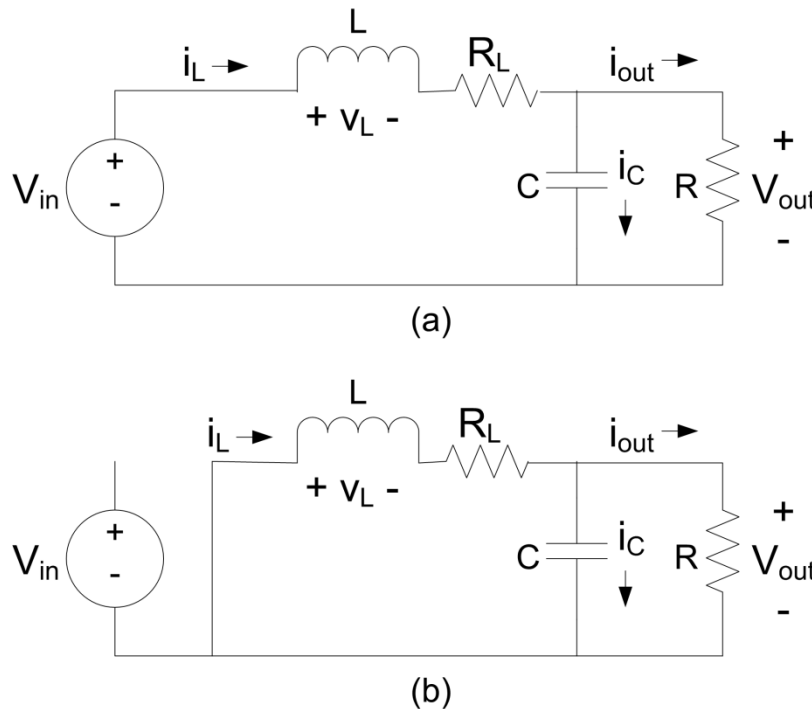


Figure 8: Buck converter with inductor equivalent series resistance R_L . In the first part of the cycle, R_L reduces the inductive voltage. In the second part of the cycle, it makes the inductive voltage more negative.

$$\begin{aligned}
v_L(t) &= V_{in} - i_L(t)R - V_{out} & 0 < t < DT_s \\
v_L(t) &= -i_L(t)R - V_{out} & DT_s < t < T_s
\end{aligned}
\tag{Equation 14}$$

Together, these two equations describe the inductor voltage for the entire switching period. Applying the volt-second balance to these equations yields:

$$0 = \frac{1}{T_s} \left[\int_0^{DT_s} V_{in} - R_L i_L(t) - V_{out} dt + \int_{DT_s}^{T_s} -R_L i_L(t) - V_{out} dt \right]
\tag{Equation 15}$$

If the inductor ripple current is assumed to be small and linear, $i_L(t)$ can be approximated by I_L , the average (DC) current. As will be later shown, this is not a good approximation under certain circumstances, but for now, it simplifies the equations immensely. Using this approximation, the above equation simplifies to:

$$0 = D(V_{in} - R_L I_L - V_{out}) + D'(-R_L I_L - V_{out})
\tag{Equation 16}$$

The quantity D' used above is defined as $D' = 1 - D$ such that the amount of time the switch is in position 2 is $D'T_s$. Combining terms and making use of the definition of D' yields equation 17 for the duty cycle of a nonideal converter with small ripple. The duty cycle will be higher due to the increased resistance. This is to be expected since additional resistance means more energy will need to be transferred from the source to make up for the resistive loss.

$$D = \frac{R_L I_L + V_{out}}{V_{in}} \quad \text{Equation 17}$$

The linear, small ripple current assumption is not a good assumption under all conditions. The inductor and its internal resistance form an RL circuit with a fixed step voltage applied at $t = 0$ and $t = DT_S$. If the inductor is high quality, the time constant of the RL circuit will be much larger than the switching period and the current ripple will be small and linear. However, if the inductor is low quality or the time constant is comparable to or smaller than the switching period, the ripple will be large and exponential.

The exact equation of the inductor current with a low quality inductor would depend on the initial currents, the input and output voltages, and the values of L and R_L . Finding this equation, plugging it into the volt-second balance equation and solving for duty-cycle would be difficult if not impossible and is of dubious value anyway since controllers always employ feedback to adjust the duty-cycle to the exact value required.

However, if the duty cycle is known or can be estimated, the ripple current can be found for both the linear ripple and exponential ripple cases. In the linear ripple case, the same approach can be used as for the ideal case. The slope of the current is equal to the voltage across the inductor and rises for the time DT_S . This yields equation 18 for the nonideal linear ripple current.

$$2\Delta i_L = \frac{V_{in} - I R_L - V_{out}}{L} DT_S \quad \text{Equation 18}$$

If the current curve is not linear due to either a large coil resistance or long switching period, the equation is much more complicated. To begin the derivation, recall that the step response of an RL circuit to a voltage V with initial current I_0 is given by the following equation [10]:

$$i_L(t) = \frac{V}{R} + \left(I_0 - \frac{V}{R}\right) e^{-(R/L)t} \quad \text{Equation 19}$$

The ripple current is the difference between the minimum current, which occurs at the beginning of a cycle, and the maximum current, which occurs at the transition point in the cycle at $t = DT_S$; or in mathematical terms:

$$2\Delta i_L = i_L(DT_S) - I_L(0) \quad \text{Equation 20}$$

Using the equation for the step response, we can substitute the equation for inductor current at $t = DT_S$.

$$2\Delta i_L = \frac{V_{LR}}{R_L} + \left(I_L(0) - \frac{V_{LR}}{R_L}\right) e^{-(R_L/L)DT_S} - I_L(0) \quad \text{Equation 21}$$

The quantity V_{LR} is the voltage across the real inductor which is equal to $V_{in} - V_{out}$. Substituting this in and rearranging terms results in equation 22.

$$2\Delta i_L = \left(\frac{V_{in} - V_{out}}{R_L} - I_L(0)\right) \left(1 - e^{-(R_L/L)DT_S}\right). \quad \text{Equation 22}$$

Using this equation, one could make a first order estimation of the ripple current by replacing $I_L(0)$ with the DC current. This estimate could further be refined by subtracting the estimate of the ripple current Δi_L from the DC current and using this number as the initial current.

In theory, one could reformulate the ripple formula for the second part of the switching cycle. If this formula was set equal to the above formula, it could be possible to find the duty cycle for an arbitrary converter. In reality, there are just too many variables that have to be known such as the initial current and switching frequency in addition to the input and output voltages. I was unable to find a closed solution via this method even after several hours of algebraic manipulations. Simulation is likely the best tool for finding the exact duty cycle of a converter with a nonideal inductor.

So far, the DC resistance of the inductor has been examined, but in reality, the resistance of the inductor increases with frequency due to the skin effect. The skin effect describes the tendency of conductors to conduct only on the outside (skin) of the conductor at high frequency. The depth of this skin is frequency and conductivity dependent. As the skin depth decreases, less of the conductor is able to conduct and the resistance increases. A full discussion of the skin effect will not be attempted here, but can be found in any book on electromagnetic theory such as [11]. Nevertheless, the equation for skin depth is presented below.

$$\delta = \sqrt{\frac{2}{\omega \mu_0 \sigma}}$$

Equation 23

In addition to the nonideal resistance in an inductor, they may also have a significant amount of parallel, parasitic capacitance between their terminals. The effects of this capacitance become more pronounced as inductors get smaller and as the frequencies get higher.

Resonance results when the inductor is operated at a frequency such that the inductance, resistance, and capacitance of the device form a tank circuit. As the switching frequency approaches the resonant frequency, the inductor's quality drops precipitously and the effective inductance drops to zero. This effectively limits the upper operating frequency of an inductor.

The other major effect of the parallel capacitance is energy loss. During the switching cycle, the inductor is first charged to $V_{in} - V_{out}$ and then discharges and charges to $-V_{out}$. Each time the capacitor is discharged, the energy that was stored in the parasitic capacitance is lost.

The energy stored in an ideal capacitor is $\frac{1}{2} CV^2$, so the energy lost in the parasitic capacitor each cycle is given by the equation

$$P_{C_L-one\ cycle} = \frac{1}{2} C_L (V_{in} - V_{out})^2 + \frac{1}{2} C_L V_{out}^2. \quad \text{Equation 24}$$

The power lost due to the parasitic capacitance is likely to be small compared to the resistive losses, at least for frequencies much lower than the resonant frequency. Nonetheless, if the capacitance is large, it can be a significant component of the total dynamic power loss in a converter.

2.2.1.2 Capacitors

Capacitors, just like inductors, are made of lossy materials. Some of this loss is due to the resistance of the electrodes and some of the loss is due to dielectric losses. The combination of these two mechanisms is modeled as the equivalent series resistance (ESR).

The capacitor's ESR is shown in Figure 9 where the circuit has been redrawn for each of the switch intervals. Adding the capacitor's resistance causes the voltage ripple to increase. The reason for this is because the ripple current from the inductor flows almost entirely through the capacitor. A large, ideal capacitor will absorb even large ripple current with only a small voltage increase, but the voltage developed by the ESR is always proportional to the current's magnitude.

At this point, there are two ways to analyze the capacitor and its ESR. If the small ripple assumption is made, then the capacitor voltage must fall when a positive current i_C is flowing. This is not possible since a positive capacitor current would increase the capacitor voltage. Therefore, the small ripple assumption cannot be used.

The other way to analyze the circuit is to remember that if the load current is assumed to be constant, then all of the inductor ripple current must flow through the capacitor. This leads to the first order estimate for the voltage ripple where the capacitive voltage is constant but the overall output voltage varies by $2\Delta i_L R_C$. This estimate works well if the capacitor is very large compared to the ripple current and the capacitive voltage change can be ignored, but that is often not the case.

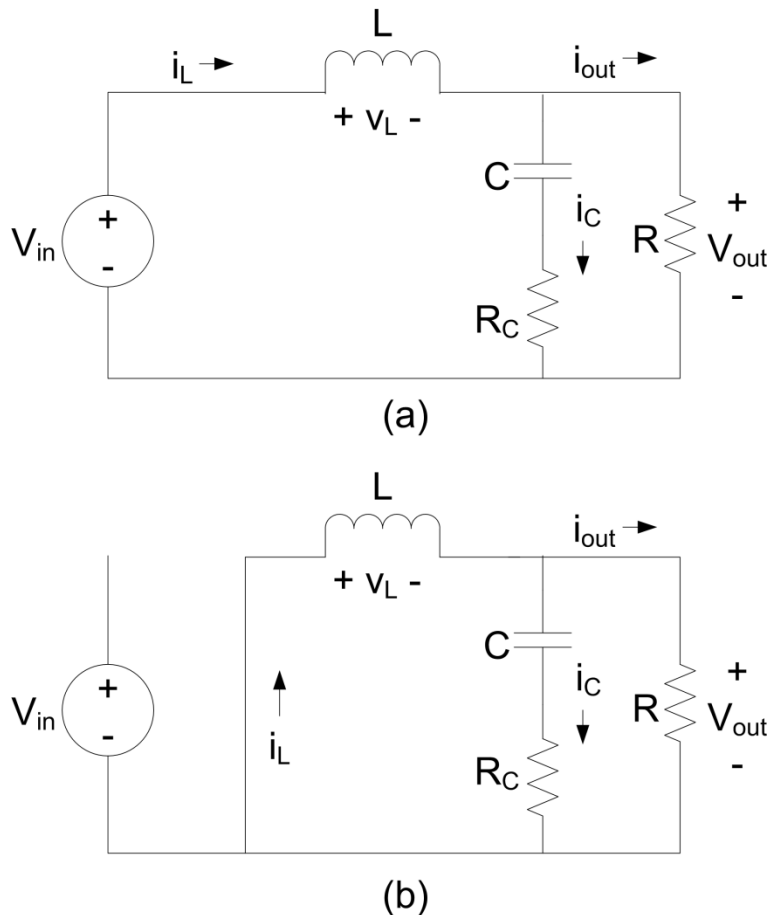


Figure 9: Buck model with nonideal capacitor.

If the ripple current is linear, we can easily adapt equation 12 for ripple voltage derived in Section 2.1.3. The previous equation does not include a term for the ESR of the capacitor. That term will now be added.

$$v_{pp-ripple} = \frac{1}{4C} T_S \Delta i_C + 2R_C \Delta i_C \quad \text{Equation 25}$$

As previously stated, the ripple current through the capacitor is equal to the ripple current through the inductor. Using real inductors, the linear approximation of the ripple

current is given by equation 18 derived in the previous section. Although equation 18 could be substituted into equation 25, the resulting mess of algebra does not reveal any significant insights into circuit design. Instead, the logical choice would be to first determine the current ripple and then plug this number into the output voltage ripple equation and either solve for the voltage ripple or solve for the capacitance given a certain desired maximum voltage ripple.

Keep in mind that one of the goals of converter design is to make the ripple small by properly choosing the components and operating conditions of the circuit. By ignoring the small ripple assumption (even though that is the goal), we have now derived a method to specify the resistance and capacitance of a capacitor such that the output ripple is small.

2.2.2 Real Semiconductor Devices

The two main types of semiconductor devices seen in small- to medium-scale converters are MOSFETs and diodes. Larger converters capable of handling many kilowatts of power use other devices such as IGBTs, SCRs, etc. where simple MOSFETs and diodes are incapable of handling the voltages and/or currents involved. These other devices are not covered here because they are not applicable to small, low-power devices such as integrated circuits.

Because every PN junction is just a parasitic, lossy capacitor waiting to happen, semiconductor devices often have significant dynamic power losses. In addition, semiconductor devices in general are charge controlled devices – a certain charge either present or absent at a certain place in a device determines its behavior [5]. Applying and

removing this charge, which can generally be modeled as a capacitor, is another significant source of power loss. All semiconductor devices exhibit resistive losses in addition to their dynamic losses.

For purposes of discussion, the standard diode-rectified buck converter shown in Figure 10 will be considered. This is easier to construct than the synchronous rectified converter, but is generally not as efficient. The p-channel MOSFET Q acts as a high-side switch controlled by the control voltage V_{cntl} . Switch position 1 is realized when Q is turned on and switch position 2 is realized by the diode D.

2.2.2.1 MOSFETs

MOSFETs suffer from both static losses and dynamic losses. The static losses are due to the resistance of the implants, channel, and, in some cases, the body of the device. The dynamic losses are due to several causes such as the switching transitions and the parasitic capacitances in the device.

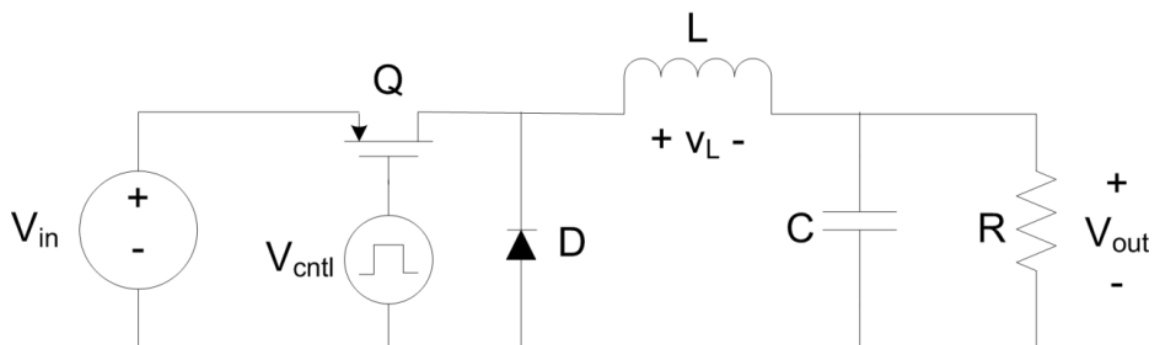


Figure 10: Buck converter with diode rectifier.

2.2.2.1.1 MOSFET Static Losses

When a MOSFET is used in a switching regulator as a switch, it should always be operating in the triode region. Recall that the triode region is the operating region where the device acts like a resistor for currents below the saturation current. This resistance can be made very small by increasing the width of the device – usually by using many copies of a basic cell. This resistance is fairly constant for currents much less than the saturation current, so it is modeled as a simple resistor, R_{on} .

Typically, a MOSFET is used as the high-side switch, although in synchronous rectifiers, they are used as both the high-side and low-side switches. The high-side switch conducts only during the first part of the converter cycle with a current equal to the inductor current. If the inductor ripple current is small, it can be approximated by the average current and the power dissipated by the MOSFET will be given by the equation 26. If the ripple is not small, the power can be expressed using equation 26 with the RMS current instead of the average current.

$$P = DR_{on}I_L^2 \qquad \text{Equation 26}$$

The additional resistance between the voltage source and the inductor further reduces the voltage across the inductor during the first part of the cycle. Note that the reduced voltage across the inductor further limits how quickly the inductor current will rise. Also note that because of the volt second balance, a lower voltage across the inductor has implications for the duty cycle. The effects of the additional resistance will

be discussed in Section 2.2.3 where the complete model of the diode-rectified buck converter is shown with all conduction losses.

2.2.2.1.2 MOSFET Dynamic Losses

The dynamic losses in MOSFETs are due to two separate causes. The first is the capacitances inherent in the construction of the device. The major capacitances are shown in Figure 11. These capacitances are alternately charged and discharged during the switching cycle, each time shunting energy to the AC ground.

The gate-to-source capacitance, C_{GS} , is essentially linear and therefore acts like an ideal capacitor [5]. This capacitance is primarily caused by the parallel plate capacitor between the gate and the body (to which the source is connected) in the channel region. Typically, the gate-to-source capacitance is the largest capacitance that must be driven in the MOSFET.

Unlike the gate-to-source capacitance, the drain-to-source and gate-to-drain capacitances are not linear and instead vary with the inverse square root of the applied

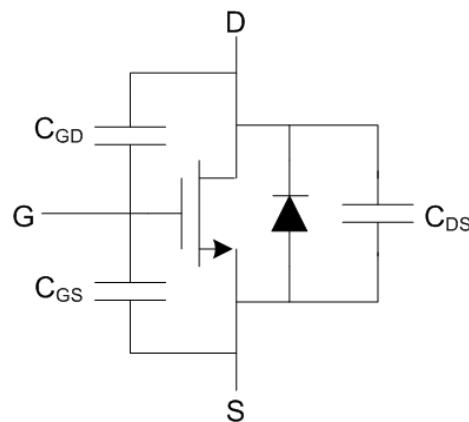


Figure 11: MOSFET model including body diode and parasitic capacitances.

voltage. For example, the drain-to-source capacitance varies according to equation 27, where V_0 and C_0 are constants intrinsic to the device. If the applied voltage is much larger than V_0 , then the simplification shown can be made, resulting in equation 28.

$$C_{DS}(v_{DS}) = \frac{C_0}{\sqrt{1 + \frac{v_{ds}}{V_0}}} \quad \text{Equation 27}$$

$$C_{DS}(v_{ds}) \approx \frac{C_0'}{\sqrt{v_{ds}}} \quad \text{Equation 28}$$

The second loss mechanism is caused by the brief time between the on and off states when the transistor is still conducting but the drain-to-source voltage is large. As shown in Figure 12, the on-to-off transition has two phases: when the transistor voltage is less than V_{in} and the current is constant, and when the voltage is equal to V_{in} and the current is decreasing to zero. In the first phase, the inductor current flows entirely through the transistor because the diode cannot turn on until the source voltage is blocked. In the second phase, the diode is starting to turn on and the inductor current flows more and more through the diode.

The entire process occurs over the course of a few tens to hundreds of nanoseconds. At this time scale, the inductor current is essentially constant. Although the transition is brief, the instantaneous power can be quite large. Reducing the on resistance of the transistor does not help since the transistor is not in the normal on-state during the transition. The best way to minimize this power is to maximize the speed with which the transistor is turned off and increase the speed at which the diode turns on [5].

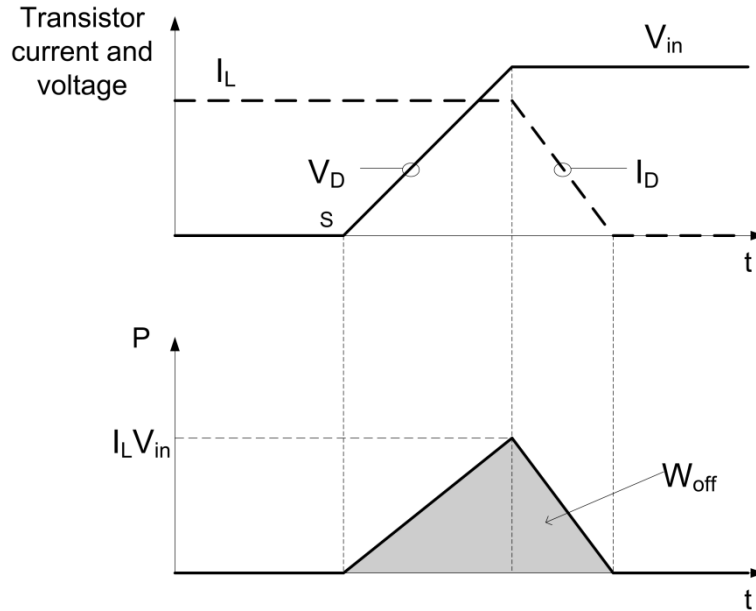


Figure 12: The MOSFET turn-off transition. Instantaneous power is shown in the lower subfigure with total power noted by the shaded region.

If the transition shown above is assumed to be piecewise linear, then the energy lost during the turn-off transition is given in equation 29, where Δt_{off} is the total time to turn off the device. The turn-on transition is almost identical except the time to turn the transistor on (and for the diode to turn off) is different. This leads to the expression for turn-on transition energy with Δt_{on} instead of Δt_{off} for the turn-on time.

$$W_{off} = \frac{1}{2} \Delta t_{off} V_{in} I_L \quad \text{Equation 29}$$

$$W_{on} = \frac{1}{2} \Delta t_{on} V_{in} I_L \quad \text{Equation 30}$$

The dynamic transistor loss will be quantified in Section 2.3.3 where all the loss models are presented.

2.2.2.2 Diodes

Diodes have significant static and dynamic losses. They are generally a poor choice for a low-power switching converter. However, their one major advantage is that they do not require active control and are therefore cheap and easy to use for rectification.

2.2.2.2.1 Diode Static Losses

The diode in Figure 10 functions as the switch for position 2. When the high-side switch is turned off, the current continues flowing through the inductor and the inductor voltage switches polarity. If the diode were ideal, the inductor node voltage would be zero. Of course, ideal diodes do not exist so the inductor node voltage must become negative so the diode is forward biased.

The easiest way to model both the turn-on voltage of the diode and the exponential current-voltage curve is with a voltage source and a resistor whose value is equal to something close to the incremental resistance at the inductor current. These elements are shown in the static loss model of Section 2.3.3.

Because of the turn-on voltage of the diode, a significant amount of power is lost in the diode. Reducing this turn-on voltage would therefore reduce the conduction loss of the diode. Furthermore, reducing the resistance of the diode by making it larger will help with the total power lost in the diode. An alternative is to use a Schottky diode which has a much lower turn-on voltage because of the different mechanism behind its operation.

2.2.2.2.2 Diode Dynamic Losses

While the steady-state performance of PN diodes is easily modeled with the basic exponential model, the dynamic performance of a diode is deceptively complicated. Like all semiconductors, the diode is a charge controlled device and the dynamic performance of the device is directly related to the processes involved in applying and removing charge.

PN diodes are especially prone to a phenomenon called reverse recovery. An excellent discussion of the reverse recovery process is presented in [5]. During the reverse recovery time, t_r , current continues to flow through a diode even after the junction has become reverse biased. The charge that flows during t_r is called the recovery charge, Q_r .

The recovered charge must flow through the MOSFET. Ordinarily, this extra current would not increase the power dissipation in the circuit, but as was discussed in Section 2.2.2.1.2, the MOSFET turn-on transition is a time when the voltage drop across the transistor is equal to the input voltage. If the transistor is assumed to turn on much faster than the reverse recovery time, then the MOSFET turn-on is dominated by the reverse recovery time. If the transistor voltage is also assumed to be equal to V_{in} during the entire transition time, then the energy lost in the MOSFET due to the diode recovery is given in equation 31.

$$W_{turn-on} = V_{in}I_L t_r + V_{in}Q_r \quad \text{Equation 31}$$

2.2.3 Complete Loss Models

Two models will now be presented: the static model which includes the frequency-independent losses due to resistance, and the dynamic model which includes the dynamic losses due to the semiconductors and the parasitic capacitance in the inductor.

2.2.3.1 The Static Model

One problem with deriving a simple model for the switching converter is the fact that the converter switches between two distinct states. Up to this point, this has been handled by redrawing the circuit as two circuits – one for each state. However, there is a better way: the DC transformer model.

The DC transformer model is presented in [5] as a way to explain the ability of switching converters to change DC currents and voltages, something that normal AC transformers cannot do. Just as in AC transformers, the power delivered to the transformer primary is equal to the power being delivered by the secondary – if the voltage is higher in the primary, the current is lower and vice versa.

The origins of the model are the volt-second balance and capacitor-charge balance equations derived from the converter diagrams in Figure 13. The equations, which have been simplified below in equations 32-34, contain source voltage and current quantities multiplied by the switching duty cycle which can be realized by ideal voltage and current sources as shown in Figure 14. These equations are identical to the current and voltage equations of a transformer with turn ratio equal to the duty cycle.

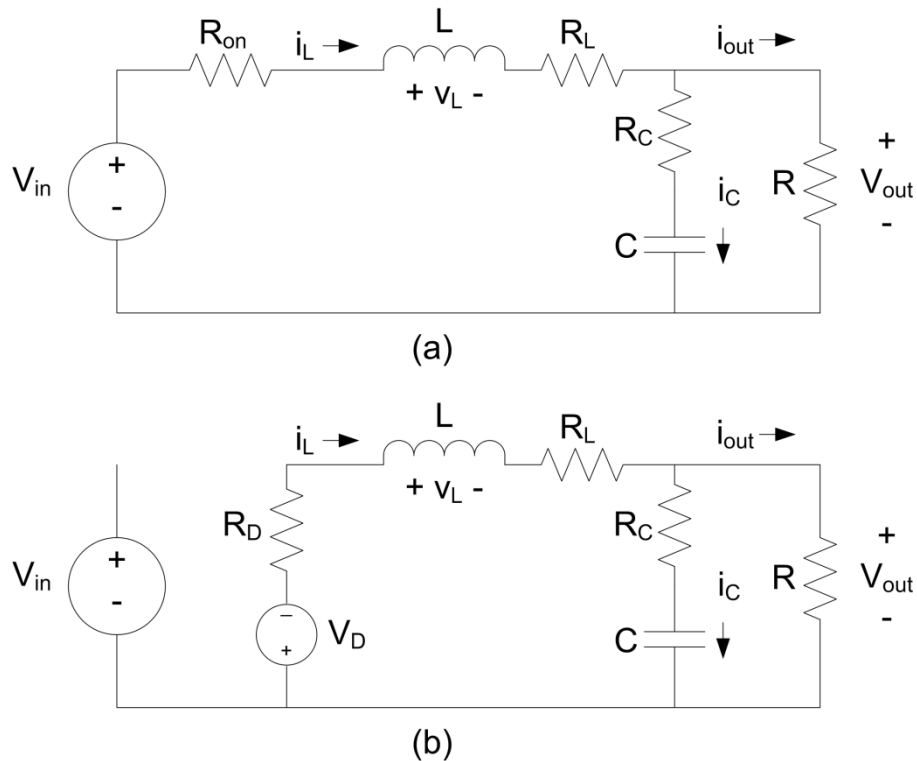


Figure 13: Buck converter with real device resistances included.

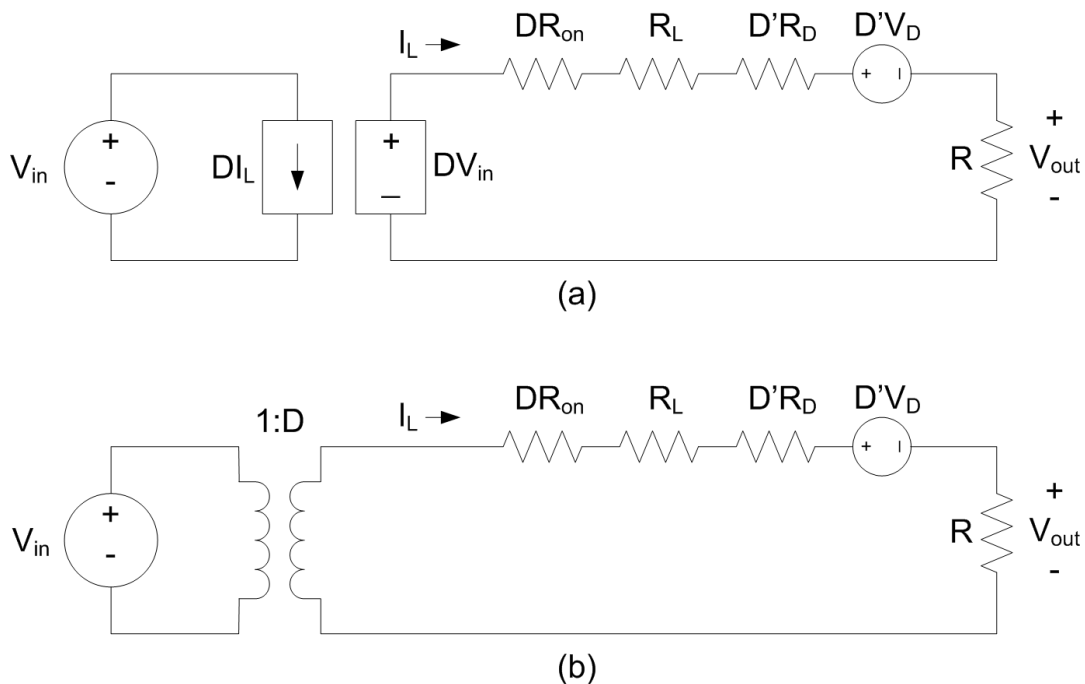


Figure 14: The DC Transformer Model. Note the equivalence between converter model with voltage and current sources (a) and DC transformer model (b). The actual conversion ratio, $M(D)$, is affected by the resistances and diode forward voltage.

$$\langle V_L \rangle = 0 = DV_{in} - DI_L R_{on} - I_L R_L - D'I_L R_D - D'I_L V_D - V_{out}$$

Equation 32

$$\langle I_C \rangle = 0 = I_L - \frac{V_{out}}{R}$$

Equation 33

$$\langle I_{in} \rangle = DI_L$$

Equation 34

In an AC transformer, the ratio of currents and voltages is determined by the turns ratio. In the DC transformer model, the turns ratio is approximated by the conversion ratio of the converter $M(D)$ which relates the output voltage to the input voltage. For an ideal buck converter, $M(D) = D$ since a buck converter acts like a low-pass filter for the pulse-width modulated switching signal.

In Figure 14, the converter current and voltage equations have been modeled as ideal current and voltage sources with conduction losses in Figure 14(a). The voltage source in the right circuit is equal to the duty cycle times the input voltage while the input current is equal to duty cycle times the inductor current. This is exactly the same situation as if the two circuits were connected with a transformer with turns ratio 1:D. This equivalence gives rise to Figure 14(b) where the circuit has in fact been redrawn with an ideal DC transformer.

The DC transformer model allows the losses to be drawn in a single circuit with the switches abstracted away, yielding a static model. Using this model, several key insights can be gained about the effect of nonidealities on the performance of buck converters.

A cursory examination of the circuit in Figure 14 would lead one to think that finding output power is as simple as multiplying the inductor current times the square of

the load resistance. The problem is that the inductor current is determined by the duty cycle which in turn is determined by output voltage and the losses in the current path. The first step in determining power output then is to determine the relationship between the duty cycle and the output voltage. By inspection, the output voltage is determined by the voltage divider formed by the parasitic losses and the load resistance.

$$V_{out} = (DV_{in} - D'V_D) \left(\frac{R}{R + DR_{on} + R_L + D'R_D} \right) \quad \text{Equation 35}$$

Dividing both sides by V_{in} and rearranging terms to get the voltage gain of the circuit results in equation 36.

$$\frac{V_{out}}{V_{in}} = \frac{R \left(D - D' \frac{V_D}{V_{in}} \right)}{R + DR_{on} + R_L + D'R_D} \quad \text{Equation 36}$$

Noting that $D' = 1 - D$, the numerator can be simplified in terms of D .

$$\frac{V_{out}}{V_{in}} = \frac{R \left(D - \frac{V_D}{V_{in}} + D \frac{V_D}{V_{in}} \right)}{R + DR_{on} + R_L + D'R_D} \quad \text{Equation 37}$$

At this point, it would be useful to know what the efficiency of the circuit is. Efficiency, η , is defined as the ratio of output power to input power. Electrical power is of course defined by current multiplied by the voltage, so electrical efficiency can be defined as equation 38.

$$\eta = \frac{V_{out}I_L}{V_{in}DI_L} \quad \text{Equation 38}$$

The currents cancel and the input-output voltage relationship can be substituted from equation 37. Furthermore, dividing the numerator by the duty cycle gives

$$\eta = \frac{R - R \frac{V_D}{V_{in}} \left(\frac{1}{D} - 1 \right)}{R + DR_{on} + R_L + D'R_D} \quad \text{Equation 39}$$

Several observations can be made based on equation 39. First, the efficiency is increased by decreasing the denominator and increasing the numerator. As is to be expected, decreasing the resistances R_{on} , R_L , and R_D increases the efficiency. If the on-resistance of the MOSFET is assumed to be smaller than resistance of the diode, which is usually the case, then the duty cycle should be as large as possible to minimize the denominator. This also has the effect of nullifying the effect of the diode forward-drop in the numerator which also increases efficiency.

The exact same process above can be used to determine the efficiency of a synchronously rectified converter. If the high-side and low-side switch resistances are equal, the efficiency of a synchronously rectified converter is shown in equation 40.

$$\eta = \frac{DR}{R + R_{on} + R_L} \quad \text{Equation 40}$$

There is a source of resistive loss that is not included in the static model of Figure 14. The equations presented earlier assumed that the capacitor was lossless. This is not the case as all capacitors have some ESR.

The problem with adding the ESR to a static model is that the ESR is in series with the output capacitor so only the ripple current flows through the resistor. The above model assumes the ripple current is negligible and linear. These assumptions are not always correct; quite often they are not.

The model cannot easily be adapted to include the ESR of the capacitor because the loss due to ESR comes from the ripple current only which does not affect the output voltage (except for a small ripple voltage). The easiest way to add the loss due to the capacitor is to use a three-step process to adjust the estimate obtained above.

First, the above equations should be used to find an estimate for the efficiency. The complement of this estimate is then multiplied by the output power, which should be known, to determine the power lost in the circuit. Next, calculate or measure the RMS ripple current and multiply this by the ESR to obtain the power lost due to ESR. This value is then added to the original estimate and added to the output power to recalculate the input power. Once the input and output powers are known, the efficiency can be calculated by dividing the output power by the input power.

2.2.3.2 The Dynamic Model

The dynamic losses are primarily due to two causes: switching capacitors and semiconductor switches in transition between states. As was seen in the previous

sections, most devices have some parasitic capacitance that is charged and discharged during the switching cycle. This energy is lost as it is shunted to an AC ground.

Generally, the parasitic capacitors only lose energy once per cycle as opposed to once at the beginning or end of the cycle and again at the transition at the middle of the cycle. The exceptions to this rule are the inductor shunt parallel capacitance, C_L , and the MOSFET gate-drain capacitance C_{GD} .

The MOSFET drain-source junction capacitance and the diode junction capacitance are both voltage dependent capacitances. According to [5], these capacitances can be modeled by linear capacitors of value $\frac{4}{3} C(V)$. Ignoring the small resistive drop across the MOSFET when it is on, both of these capacitors are charged to V_{in} so the loss due to these junction capacitors is given as

$$W_{JC} \approx \frac{1}{2} \left(\frac{4}{3} C_d(V_{in}) + \frac{4}{3} C_{ds}(V_{in}) \right) V_{in}^2 \quad \text{Equation 41}$$

Similarly, the gate capacitance is charged and discharged once per cycle. The gate capacitor voltage depends on the type of MOSFET and drive circuit used. For a simple P-MOS device where the gate is switched to ground, the gate-to-source voltage is V_{in} . The gate-to-drain voltage varies throughout the switching cycle. Just before the start of turn off, the gate-drain voltage is approximately V_{in} . Once the gate voltage is raised to V_{in} , the gate-drain voltage falls to zero, then is charged to negative V_{in} as the drain voltage drops to approximately ground (less the diode drop). When the MOSFET is turned on again, the cycle reverses. Taking into account the nonlinearity of the gate-drain capacitance, the total gate capacitance loss is therefore approximately:

$$W_G \approx \frac{1}{2} C_{GS} V_{in}^2 + \frac{4}{3} C_{GD} (V_{in}) V_{in}^2 \quad \text{Equation 42}$$

The energy lost due to the inductor shunt capacitance has already been derived in a previous section. The total energy loss due to capacitance is now shown.

$$W_C = \left[\frac{2}{3} C_d (V_{in}) + \frac{2}{3} C_{ds} (V_{in}) + \frac{1}{2} C_{GS} + \frac{4}{3} C_{GD} (V_{in}) \right] V_{in}^2 + \frac{1}{2} C_L [(V_{in} - V_{out})^2 + V_{out}^2] \quad \text{Equation 43}$$

The controller, which has not been discussed up to this point, is another source of dynamic power loss. The power loss of a generic IC circuit, such as a PWM feedback controller, is linearly dependent on the frequency of operation. According to [12], the dynamic power of a digital circuit can be expressed in terms of the capacitances being switched, the digital supply voltage, and the frequency of switching:

$$P_{digital} = C V_{DD}^2 f_{sw} \quad \text{Equation 44}$$

The total capacitance in the controller is likely to be between one-fourth and one-half of the capacitance of the gate-source capacitance. This assumption is based on the fact that the controller has to drive the gate capacitance at the switching speed with minimum delay. In order to do this, the output stage would be designed with the theory of logical effort in mind so as to minimize switching times.

The theory of logical effort states that using progressively larger drivers between the logic and the load results in the lowest path delay [12]. Assuming that each stage is four times larger than the previous stage, then the capacitance of the previous n stages is $\frac{1}{4} + \frac{1}{16} + \frac{1}{64} + \dots + 1/(4^n)$. The actual controller circuitry will add some additional capacitance, which together with the driver stages should be approximately one-half the capacitance of the gate-source capacitance. The dynamic power in the controller can therefore be estimated to be given by the equation 45.

$$P_{controller} \approx \frac{1}{2} C_{GS} V_{DD}^2 f_{sw} \quad \text{Equation 45}$$

The power lost due to the transistor and diode transitions has already been quantified in previous sections. All of the major sources of dynamic power loss have now been identified. Note that these sources of loss all depend directly on the switching frequency. Combining all of these losses into a single equation yields the somewhat unwieldy equation 47.

$$P_{dynamic} = P_{turn-on} + P_{turn-off} + P_{capacitors} + P_{controller} \quad \text{Equation 46}$$

$$P_{dynamic} \approx f_{sw} \left[(V_{in} I_L t_r + V_{in} Q_r) + \left(\frac{1}{2} \Delta t_{off} V_{in} I_L \right) + \left[\frac{2}{3} C_d (V_{in}) + \frac{2}{3} C_{ds} (V_{in}) + \frac{1}{2} C_{GS} + \frac{4}{3} C_{GD} (V_{in}) \right] V_{in}^2 + \frac{1}{2} C_L [(V_{in} - V_{out})^2 + V_{out}^2] + \left(\frac{1}{2} C_{GS} V_{DD}^2 \right) \right] \quad \text{Equation 47}$$

2.3 Discontinuous Conduction Mode

In the preceding discussions about buck converters, the inductor current has been assumed to be made up of a large DC component, equal to the average current, with a small ripple current superimposed upon it. While this assumption makes analysis easier, it often does not hold true.

Consider the inductor currents shown in Figure 15. The first inductor current is the familiar waveform where a small ripple is imposed on the large DC current equal to the load current. Next is the case where the load current is equal to the ripple current, Δi_L . At this load current, the inductor current just reaches zero before it begins ramping up at the start of the next cycle.

If the load current drops even further, the third waveform is the result. In this case, the inductor current drops to zero and stays at zero for some part of the switching cycle. Since current is not continuously flowing in the inductor, the converter is said to be operating in the discontinuous conduction mode (DCM) [5].

The DCM is only possible in converters where the low-side switch is either unidirectional or is under active control to turn it off when the current equals zero. A diode would fit the first criteria since it only allows current flow in one direction.

The concept of the discontinuous mode is very important to switching supply design. Designs which operate at low load current for some or all of their output current range will typically operate in discontinuous mode. Furthermore, some supplies are designed to always operate in the discontinuous mode.

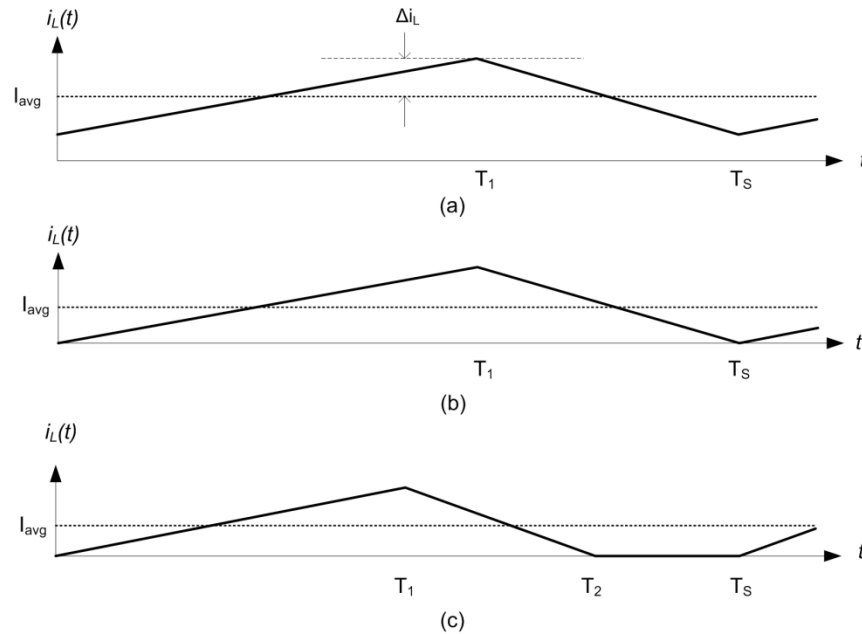


Figure 15: Inductor currents for CCM and DCM operation.

Although control methods will not be covered in depth, the Pulse Frequency Modulation (PFM) control method is a discontinuous control method so important to low power converters that it must be mentioned here.

Pulse frequency modulation is a DCM control technique frequently employed to control converters under light load. Pulse width modulators typically have great difficulty controlling the output voltage at low loads. In addition, the PWM controller typically dissipates as much power as it delivers to the load when the load is very light because of excessive switching activity [13]. Pulse frequency modulators, on the other hand, perform very well under light load.

Pulse frequency modulation typically uses a fixed on-time pulse on the high-side switch with a repetition frequency determined by the load current [7]. This on-time can be tuned to the LC filter such that the conduction loss and output ripple is not excessive. These controllers have the advantage of running in a low-current idle mode while the

power switches are not conducting. In addition, since the pulse time does not have to be dependent on a comparator circuit, like most PWM controllers, the pulse time can be made very small through RC timers or delay lines.

2.3.1 Discontinuous Mode Conditions

In the most basic of terms, discontinuous mode occurs when the output current is equal to or smaller than the ripple current. Under this condition, the area under the inductor current curve is equal to the area under the constant load current curve, even though the inductor current only flows for part of the cycle. This implies that the average positive ripple current is larger than the load current. Therefore, the conditions for discontinuous mode are the conditions for ripple current larger than the load current.

The most obvious case where the ripple current is larger than the load current is when the load current is zero. In this case, the ripple current will always be larger than the load current unless the duty cycle is zero – typically achieved through pulse-skipping or Pulse Frequency Modulation (PFM) as opposed to Pulse Width Modulation (PWM).

When the load current is not zero but still small, the dividing line between continuous and discontinuous mode is primarily determined by the inductor size and the conduction time. Assuming ideal components, the inequality in equation 48 is true for discontinuous mode [5].

$$\frac{2L}{RT_S} < D'$$

Equation 48

It is useful to notice that the input and output voltages do not appear in the above formula. Recall that in ideal buck converters, the positive duty cycle is equal to the ratio of the input and output voltages. Therefore, the negative duty cycle, D' , actually does contain the input and output voltages. The load current also does not appear in the above equation, although the equivalent load resistance does.

When real components are used, the inequality above does not work. Resistance in the power switches and inductor greatly affect the voltage conversion ratio so that it is no longer equal to the duty cycle. In addition, the inductor current is exponential instead of linear due to the resistance in series with the inductor and the voltage source.

To find a better formula for the boundary condition, start with the basic condition that in discontinuous mode, the ripple current is larger than the load current. The maximum current can be found by evaluating the step response of the parallel RL circuit at time $t = DT_S$, the high-side switch conduction time. This leads to the more complicated but general solution for the ripple current at the boundary condition,

$$\Delta i_L|_{DCM-boundary} = \frac{i_L(DT_S)}{2} = \frac{V_{drop}}{2R_S} \left(1 - e^{-DT_S R_S/L}\right) \quad \text{Equation 49}$$

The voltage V_{drop} is the voltage drop across the converter from V_{in} to V_{out} and R_S is the series combination of the inductor's ESR and the high-side switch on resistance. The current is also assumed to be symmetric about I_{out} such that the ripple current is half the maximum current (this assumption only holds at the boundary condition and above). A more accurate estimate of the ripple current would calculate the RMS value of the

current; however, if the switching time is comparable to the time constant L/R_S , this additional work is probably not justified.

As is to be expected, the ripple current in equation 49 is directly related to the voltage drop. The relationship between ripple current and both the inductance and resistance is more complicated. Larger inductance increases the time constant of the circuit which alone decreases the ripple current, especially if the switching period is shorter than the time constant, as shown in Figure 16. However, the final current value for the converter, as for all RL circuits, is determined by the resistance.

Increasing resistance decreases the current ripple simply by decreasing the maximum possible current through the circuit. However, increasing the resistance also causes the current to rise to the final value faster, as shown in Figure 17. A smaller time

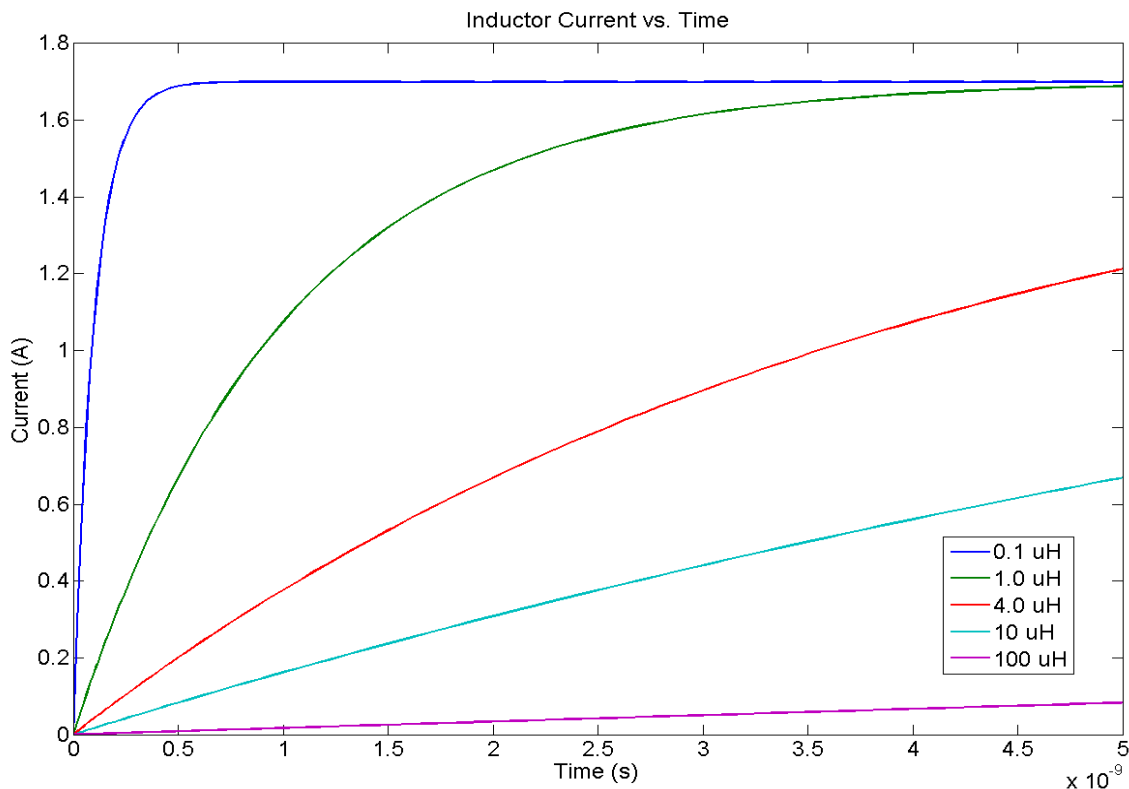


Figure 16: Inductor ripple current vs. time with fixed series resistance, $R_S=1$ Ohm.

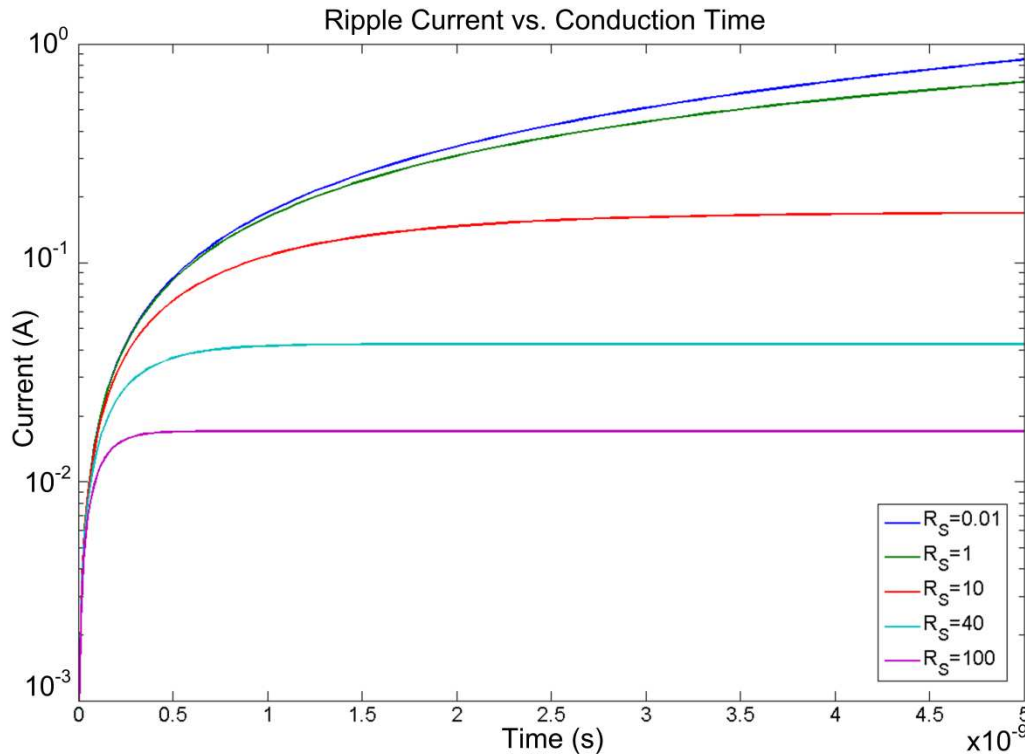


Figure 17: Ripple current vs conduction time showing rapid current saturation at high series resistance.

constant means the inductor's voltage drop is more likely to be purely resistive at the end of the on-time. This in turn reduces the efficiency of the converter since the energy dissipated by resistance is not recovered, unlike the energy stored as magnetic flux in the inductor.

2.3.2 Implications of Discontinuous Conduction

In some ways, discontinuous conduction does not change the way a converter behaves. The capacitor charge balance and volt-second balance still apply since these rules are based on cycle-to-cycle steady-state assumptions. The steady-state assumptions themselves still hold since the converter should not fundamentally change from cycle to cycle. Even the small ripple assumption is essentially valid, although it must be applied

with care since the very definition of discontinuous mode is that the ripple current is larger than the output current.

On the other hand, assumptions such as the equivalence between duty cycle and voltage conversion ratio no longer apply, even as a rule of thumb. Efficiency is also affected by the difference in the way that current flows throughout the conversion cycle. The discussion of the implications of discontinuous conduction will be facilitated by Figure 18 which assumes a MOSFET and diode are used for the high-side and low-side switches, respectively.

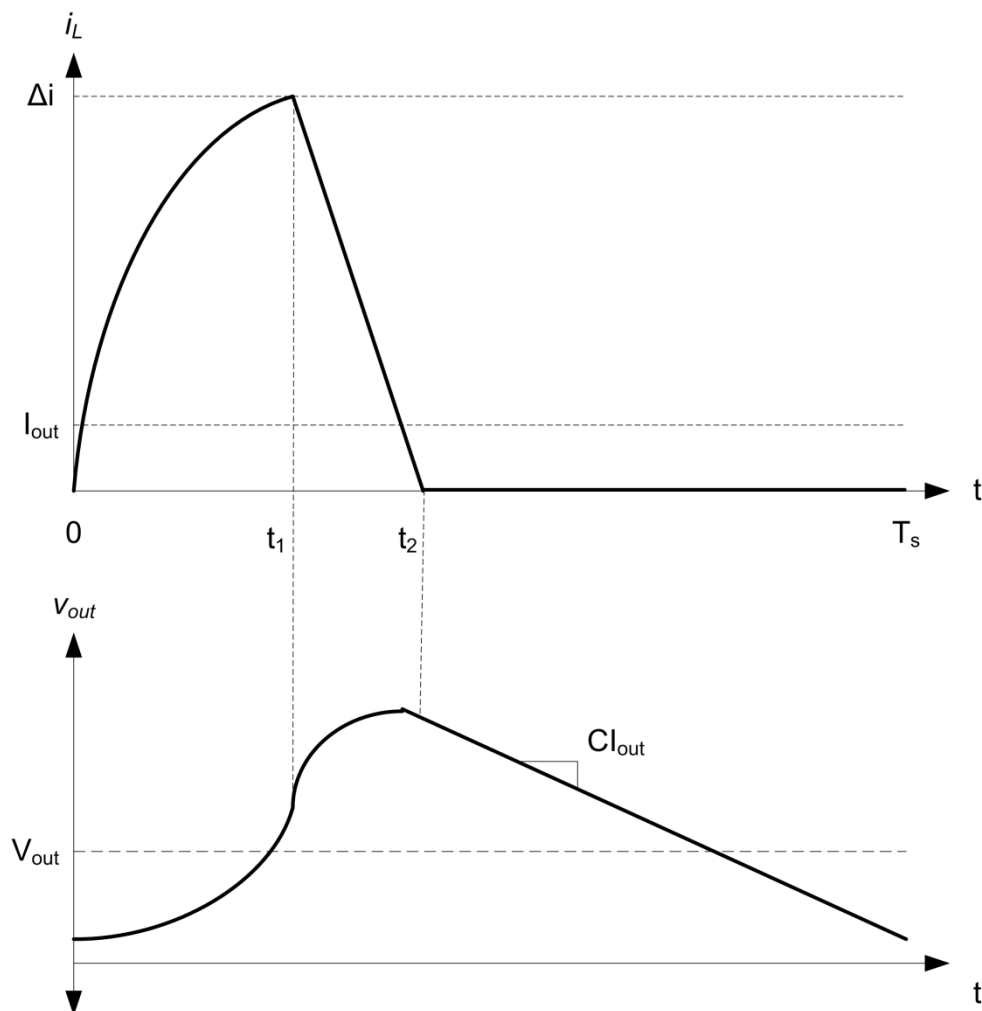


Figure 18: Exponential DCM signals.

Starting at time 0, the high-side switch turns on and the inductor current follows the step response of the series RL circuit formed by the series resistances and the coil inductance. At time t_1 , the high-side switch turns off and the diode begins conducting. The current drops at a rate determined by the output voltage, the resistive voltage, and the diode forward drop.

Since the inductor is reacting to a negative step voltage from t_1 to t_2 and has a positive initial voltage, the current drops almost linearly to zero. At time t_2 , the inductor current reaches zero and the diode stops conducting. From time t_2 until T_s , no inductor current flows and the output current is entirely supplied by the capacitor.

The output voltage ripple during the conversion cycle is affected by the inductive current relative to the output current. From time 0 until t_1 , there is a period during which the output voltage will continue to drop since the inductor current does not immediately exceed I_{out} . Once $i_L(t)$ is larger than I_{out} , the capacitor begins charging following an approximately exponential curve.

At time t_1 , the high-side switch turns off, but the inductor current is still larger than I_{out} , though it is dropping rapidly, and the output capacitor continues to charge. Eventually, the inductor current drops below I_{out} and the capacitor starts to discharge. Once the diode stops conducting at time t_2 , the capacitor discharges at a rate approximately equal to the capacitance times the output current.

The capacitor charge balance still holds, so the area under the inductor curve – that is, the charge that has flowed through the inductor – is equal to the area under the constant I_{out} line.

While the maximum inductor current is fairly easy to calculate given the resistance, inductance and on-time, the ripple voltage is not so straightforward. In qualitative terms, the voltage ripple is due to the integration by the capacitor of the excess inductor current, $i_L(t) - I_{out}$. When the inductor current was linear, the calculations were easy; however, the currents in a real converter are exponential. Even if the inductor currents are assumed to still be linear, the time t_2 must be known in order to determine the slope of the inductor current from t_1 to t_2 .

Solving for t_2 exactly would require the modeling of the inductor, diode, and component resistances in series with the output capacitor. This would be difficult and the result would likely be so complicated as to be unusable. Instead, t_2 can be estimated using a linear approximation of the falling inductor current.

The falling inductor current between t_1 and t_2 is an exponential curve with the starting and ending slopes shown in Figure 19. The average of these slopes should give a reasonable approximation of the time at which the inductor current reaches zero. Note that this approximation overestimates the amount of charge delivered to the capacitor.

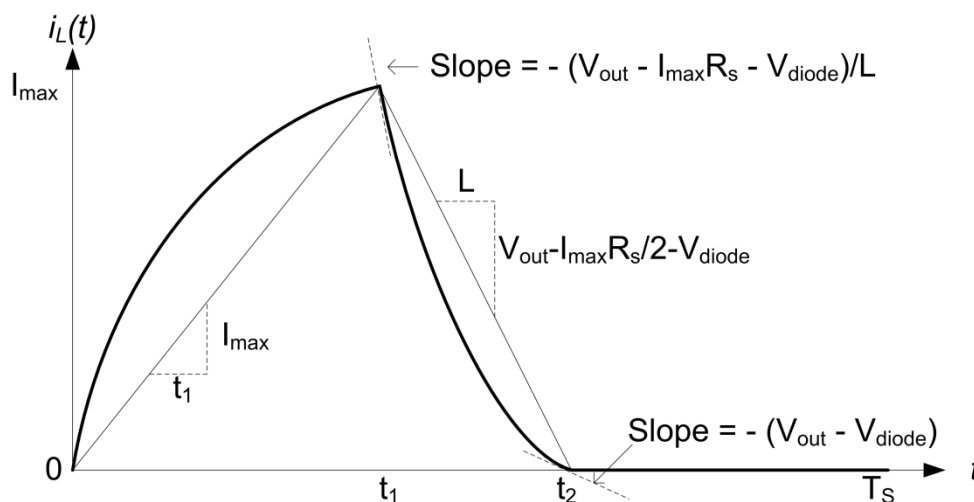


Figure 19: Discontinuous current ripple with linear approximations.

The rising current between time 0 and t_I is also linearized in Figure 19. The maximum value is easily calculated and the time t_I is equal to DT_S . This approximation underestimates the amount of charge delivered to the capacitor. If t_I is comparable to one or two time constants, the error should be fairly small. If t_I is longer than three time constants, the estimate will be much larger, as shown in Table 1.

Since the linear approximation understates the amount of charge transferred from time 0 to t_I , it will also tend to overestimate the time t_I if used for that purpose, since the charge transferred determines the length of the high-side on time. If the exact quantity of charge needs to be calculated, it is best to use the exponential formula for the inductor current instead of the linear formula.

Now that the ripple current has been approximated, the voltage ripple can be approximated more easily. The linear approximation of the ripple current is shown in Figure 20 along with the slopes of the various sections.

Table 1: Accuracy of linear current approximation for rising current. Constants used were 1 μH coil, 1 Ω series resistance, and 5V step voltage. The time constant is 1 μS . The percent error was calculated as the difference between actual and estimated charge divided by the actual charge.

Time	Charge	Current	Charge Approximation	% Error
1 τ	0.6254 μC	1.075 A	0.5373 μC	14.09 %
2 τ	1.930 μC	1.470 A	1.470 μC	23.83 %
3 τ	3.485 μC	1.615 A	2.423 μC	30.47 %
4 τ	5.131 μC	1.669 A	3.338 μC	34.94 %
5 τ	6.811 μC	1.689 A	4.221 μC	38.03 %

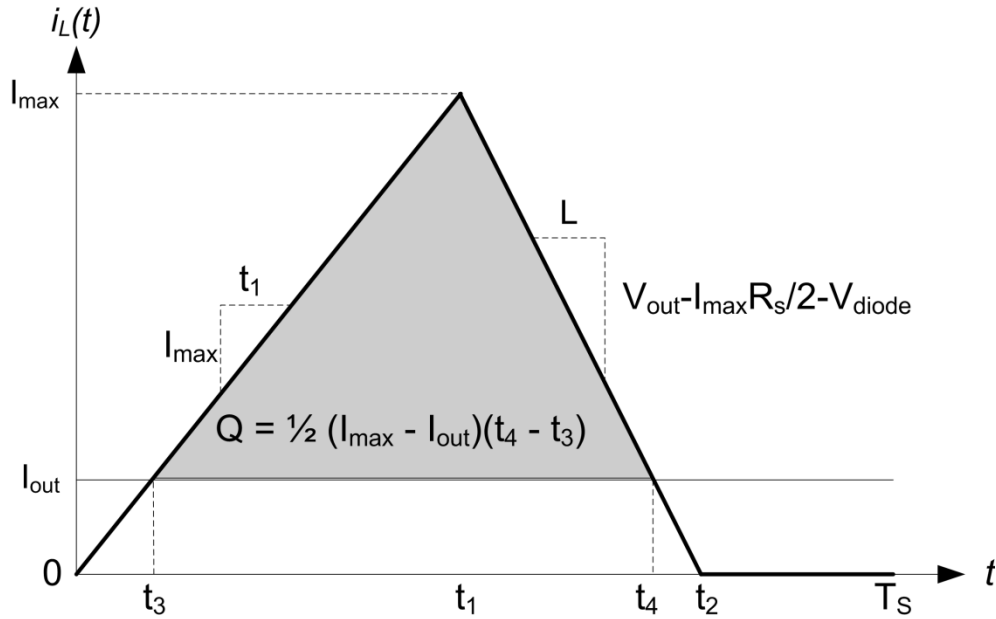


Figure 20: Linear approximation of ripple current showing the positive charge area (shaded triangle).

Through a series of algebraic manipulations, the area of the shaded triangle – the total charge transferred – can be found to be

$$Q = \frac{1}{2} I_{out} \left(DT_S + \frac{L}{-V_{out} - I_{max} R_S / 2 - V_{diode}} + DT_S \frac{I_{max}}{I_{out}} \right). \quad \text{Equation 50}$$

The voltage peak-to-peak ripple is then found by multiplying the Q transferred to the capacitor times the capacitance of the capacitor.

$$V_{ripple} = \frac{1}{2} I_{out} \frac{DT_S + \frac{L}{-V_{out} - I_{max} R_S / 2 - V_{diode}} + DT_S \frac{I_{max}}{I_{out}}}{C}. \quad \text{Equation 51}$$

The efficiency of the converter is slightly different than in the continuous conduction mode. There are two main reasons for this. First, the diode does not conduct the entire time from DT_S until T_S . Second, the capacitor is the only source of current from t_2 until T_S with the result that all the current during this time flows through the capacitor's ESR.

Because the ripple current is larger than the output current and because the average conduction current is larger than the output current, the transformer model cannot easily be modified to determine the efficiency of the converter. Instead, the efficiency will be calculated by determining the power lost in each element versus the power delivered to the load.

Since the current can be approximated from the linear model and the resistance can easily be measured, the energy formula used is,

$$W_x = T_S \int_0^{D_x} i(t)^2 R_x dt. \quad \text{Equation 52}$$

This is where the linear model makes computations much easier since the integral of a squared linear function is much easier to compute than an exponential. The power lost during a switching cycle is the energy lost in the cycle divided by the cycle time, T_S . The result of adding the resistive power lost in the MOSFET, diode, inductor, and capacitor is,

$$P_{loss} = R_C I_{out} + \frac{1}{3} I_{max}^2 [R_{ds} D_1 + V_{diode} D_2 + R_d D_2 + R_L (D_1 + D_2)].$$

$$\text{Equation 53}$$

The efficiency, η , is equal to the power output divided by the power input. Because of the conservation of energy, the power input is equal to the output power plus the power lost. If only resistive power is considered, the efficiency is then equal to,

$$\eta = \frac{1}{1 + \frac{R_C I_{out} + \frac{1}{3} I_{max}^2 [R_{ds} D_1 + V_{diode} D_2 + R_d D_2 + R_L (D_1 + D_2)]}{I_{out} V_{out}}}$$

Equation 54

2.4 Contemporary Integrated Designs

Building fully integrated switching supplies is not a new concept. A search through available literature turned up several attempts, all by the same group at the School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an China. All of these were based on simulations and not actual devices [7], [8], [9]. Some of the results from this group are suspect due to the extremely low ESR values calculated for their inductors. The values quoted are about one tenth of the value calculated from the resistance formula quoted by Lee in chapter 3.

A more recent approach is to build 3-D controllers where the power switches, controller, and sometimes the capacitor are fabricated in silicon with a thin film inductor mounted directly on top of the silicon die [14], [15], [16]. In addition, there have been numerous attempts to build small, high-speed controllers and power stages so that a switching controller can be accomplished with the addition of a miniature inductor and capacitor [17], [18].

One common feature of all these designs is the use of synchronous rectification instead of diode rectification. The main reason for this is the relatively low output

voltages, typically below 3 Volts. This implies that a significant portion of the inductor voltage will be across the diode (as the forward voltage drop and resistive loss) as opposed to voltage across the capacitor and load.

One other common feature is that all the designs designed to work at low output levels used PFM control. High-power designs universally used PWM control in CCM. There is a region where both PFM and PWM perform well, but at low load current, PFM typically performs well in the 70-80% efficiency range and at high current, PWM can perform as well as 95%.

One other common feature of the designs featuring integrated inductors was high frequency. This was explicitly stated to be a result of the low inductance of the flat spiral coils. Low inductance coils require high frequency for both efficiency and low ripple.

Although most of the research on integrated switching supplies focused on finding better methods of control, there were some interesting approaches to reducing the losses. Perhaps the most interesting was a controller which reduced switching losses by only turning on part of the power switches at very light load [15]. Although this increased the resistive losses, the resistive loss is proportional to the square of the current, whereas the switching loss is directly proportional to the capacitance being switched. This switching method would be very useful if the output current was expected to vary significantly.

CHAPTER 3

SURVEY OF KEY COMPONENT TECHNOLOGIES

A survey of existing commercial reactive and semiconductor components was conducted using a major online supplier, Digikey.com. These results will be compared with the on-chip devices available to determine if using miniature external components is a viable alternative to the all-integrated solution.

3.1 Components

The three components that might be placed off-chip using miniature surface mount devices are the inductor, capacitor, and power switch. These are also three of the biggest potential sources of inefficiency.

Inductors work by storing magnetic flux in a coil of conductive material. Unfortunately, the size of the coil has to be fairly large in order to enclose a sufficient amount of flux to be useful, so inductors do not scale down very well. As a result, both integrated and miniature chip inductors have limited inductances and consequently low qualities at low frequencies. This is made worse by the high resistance found in many of these coils [19].

In contrast to inductors, capacitors actually scale fairly well. For a parallel plate capacitor, capacitance decreases as the area of the plates decreases. Fortunately,

capacitance increases as the plate separation decreases, so it is possible to make small capacitors with good capacitance.

As was seen above, it is possible, even desirable, to use a small inductor with a large capacitor, but the opposite results in a very poorly regulated design. It is therefore a good idea to use the largest capacitor with a low ESR that one can find as the output capacitor.

Power switches at these currents and voltages are typically MOSFETs in the triode region. Discrete MOSFETs are fabricated using technologies specifically designed to yield good switching characteristics. Most of these devices use three-dimensional manufacturing techniques such as trench-FETs and HEXFETs. By contrast, power switches in CMOS IC technologies are typically scaled-up two-dimensional small-signal FETs with minimum channel length.

Inductors, capacitors, and switches, both integrated and discrete, are discussed in the following sections.

3.1.1 Integrated Inductors

Integrated circuit technologies provide very limited choices for fabricating integrated inductors. The easiest to design and most common integrated inductor type is the flat spiral type inductor which will be discussed at length in the following sections. Several three-dimensional inductor designs have also been fabricated and are presented at the end of this section.

3.1.1.1 Spiral Inductors

Flat spirals, as the name suggests, are fabricated using metal interconnect tracks that spiral inward toward a central cross-under. Multiple metal layers can be connected to reduce the series resistance of the coil. Spiral coils are limited to fairly low Q – around 5 to 10 – due to a number of first and second order effects.

Circular coils tend to perform better than square coils by up to 10%, but this is mostly due to second order effects [19]. Lee presents an equation for the inductance of circular, octagonal, hexagonal, and square coils along with a table of coefficients, shown in equation 55 and Table 2 for reference [19].

$$L \approx \frac{\mu_0 n^2 d_{avg} c_1}{2} \left[\ln \left(\frac{c_2}{\rho} \right) + c_3 \rho + c_4 \rho^2 \right], \quad \text{Equation 55}$$

where ρ is the fill factor defined as

$$\rho \equiv \frac{d_{out} - d_{in}}{d_{out} + d_{in}}. \quad \text{Equation 56}$$

Table 2: Coefficients for spiral coil inductance formula

Shape	c_1	c_2	c_3	c_4
Square	1.27	2.07	0.18	0.13
Hexagon	1.09	2.23	0.00	0.17
Octagon	1.07	2.29	0.00	0.19
Circle	1.00	2.46	0.00	0.20

The test results of the three coils on the test chip are shown in Figure 21. Each of the coils showed a high ESR in the tens of Ohms with only modest inductance at high frequency.

Figure 22 shows a relatively complete model for flat spiral coils which includes most of the first and second order effects [20]. The first order effects and second order effects are summarized briefly in the following two subsections.

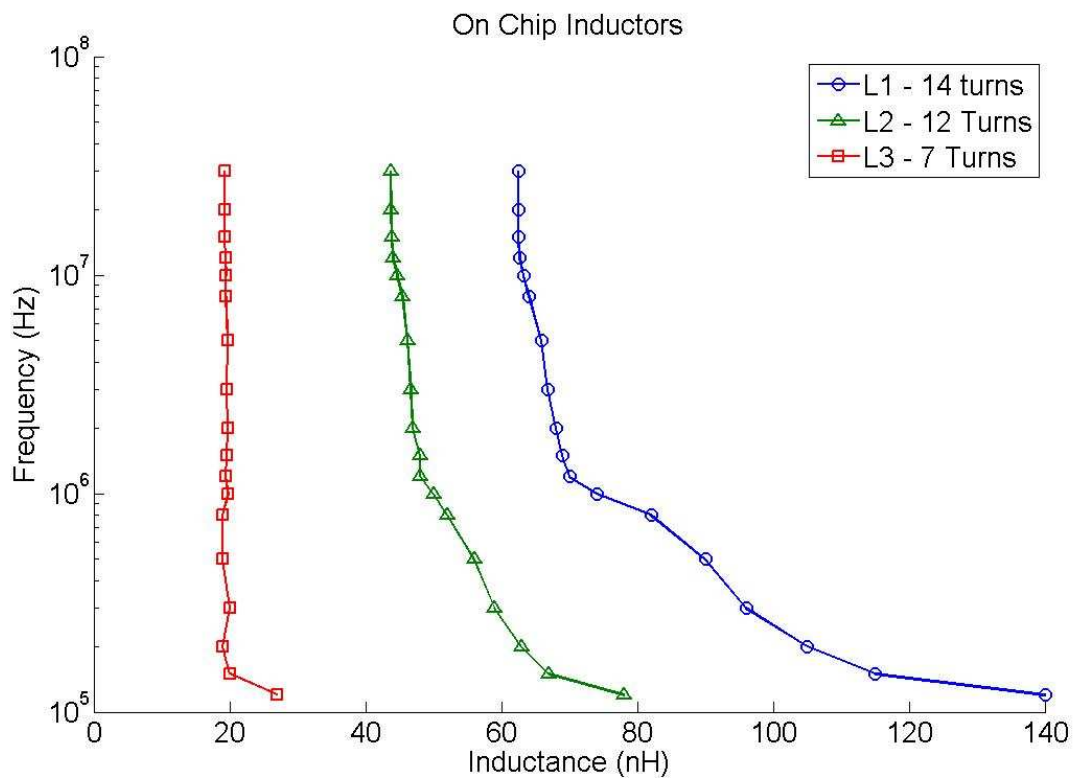


Figure 21: On-chip inductors showing inductance vs. frequency. The series resistance for coils L1, L2, and L3 were 51.5 Ohms, 35.5 Ohms, and 14 Ohms, respectively.

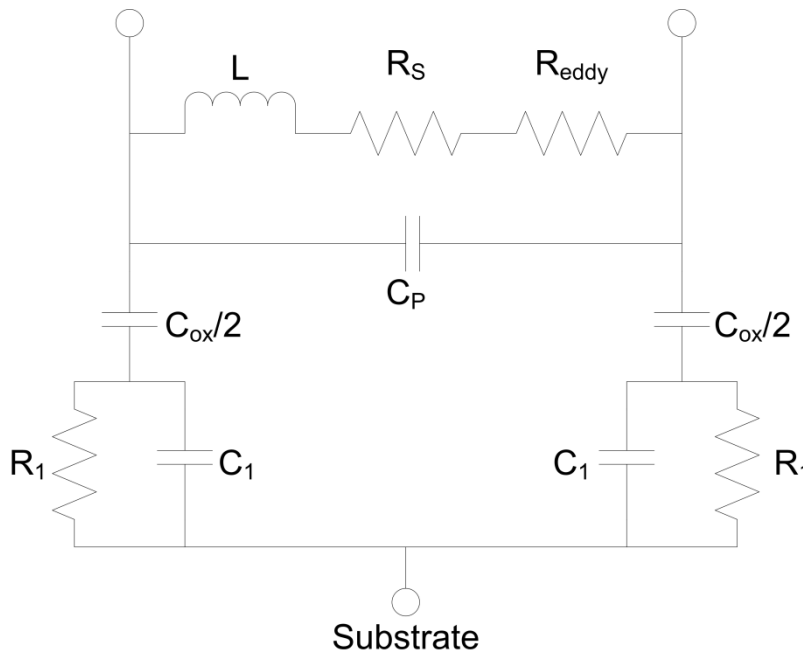


Figure 22: Model for a spiral inductor

3.1.1.1.1 First Order Limits on Q

The two biggest limiters of Q for spiral coils are low inductance and high series resistance of the coils, represented by L and R_S , respectively, in Figure 22. The inductance formula for a spiral inductor has already been presented in equation 55.

The inductance of an arbitrary coil is proportional to the amount of flux enclosed by the coil and hence the area enclosed by the turns of the coil. Since the area enclosed by an integrated coil is very, very small compared to discrete coils, the inductance is also very small. In addition to the limits on inductance due to available chip area, there are several second order effects which further limit the practical area of inductors. These factors are discussed in Section 3.1.1.1.2, “Second Order Limits on Q .”

The other first order limit on Q is the series resistance of the coil. The aluminum interconnect used in most fabrication technologies is not very thick which combined with the skin effect leads to resistance in the range of 50 m Ω per square. A large coil such as

the ones fabricated on the test chip can have lengths in the range of hundreds to thousands of squares leading to resistances of several tens of Ohms. If the total length of the winding can easily be determined, the series resistance can be calculated using the following equations [19]:

$$R_S \approx \frac{l}{w \cdot \sigma \cdot \delta (1 - e^{-t/\delta})}, \quad \text{Equation 57}$$

where l is the length, w is the width, and t is the thickness of the track. The conductivity of the metal is given by σ and the skin depth, δ , is defined by:

$$\delta = \sqrt{\frac{2}{\omega \mu_0 \sigma}}. \quad \text{Equation 58}$$

With such high resistance and low inductance, it is difficult to manufacture high quality inductors. The second order effects make this even more difficult.

3.1.1.1.2 Second Order Limits on Q

By far the biggest second order limits on Q are due to the parasitic capacitances. These capacitances are generally calculated by the formula $C = \epsilon \cdot A/h$, where A is the area of the plates, h is the separation, and ϵ is the dielectric constant. These parasitic capacitances are detrimental to performance in two ways. First, they form lossy paths across the terminals of the inductor and also to the substrate, which is a high-frequency ground. Second, they reduce the self resonant frequency (SRF) of the inductor by

making the inductor act like a tank circuit. At frequencies approaching the SRF, the inductance drops off sharply until the inductor is useless. The SRF of an inductor therefore limits the useful frequency range of a coil.

The turns of a spiral coil form a distributed capacitance between each turn and its neighboring turns. Even though the capacitance of a single turn can be large, the effect of the turn-to-turn capacitance turns out to be negligible because it is the series combination of this capacitance that appears across the terminals. Recall that the equivalent capacitance for n series connected identical capacitors is:

$$C_{eq} = \frac{1}{n} C, \quad \text{Equation 59}$$

which combined with the limited edge area and wide spacing leads to a very small capacitance compared to the major shunt capacitance due to the cross-under, C_P .

The formula for shunt capacitance given by Lee is based on the general formula for a parallel plate capacitor. The exact form presented is:

$$C_P = n \cdot w^2 \frac{\epsilon_{ox}}{t_{ox}}, \quad \text{Equation 60}$$

where w is the track width and t_{ox} is the oxide thickness between the coil and cross-under [19]. This formula is only a conservative approximation because the shunt capacitance is actually distributed across the inductor so the innermost portion of the capacitor shunts less voltage than the outermost portion.

However, to determine the real equivalent shunt capacitance, the capacitance at each intersection would need to be weight-averaged based on the voltage difference at each turn of the coil. Since each coil is slightly different in size and is not uniformly coupled with the other turns, calculating these voltages and finding the real capacitance would be difficult if not impossible. In practice, equation 60 gives a worst-case value for C_P that should be more than sufficient for all practical purposes.

The coil as a whole also acts as a parallel plate capacitor with the substrate, modeled by C_{ox} . The area of the coil is equal to its length times its width which yields the exact form:

$$C_{ox} = w \cdot l \cdot \frac{\epsilon_{ox}}{t_{ox}}. \quad \text{Equation 61}$$

The substrate's dielectric loss is modeled by R_1 and has an approximate formula:

$$R_1 \approx \frac{2}{w \cdot l \cdot G_{sub}}, \quad \text{Equation 62}$$

where G_{sub} is a fitting parameter with a typical value of 10^{-7} S/ μm^2 .

The capacitance C_1 is a lumped element that models the capacitance of the substrate as well as the reactance related to image inductance. Its formula is given in terms of a fitting parameter C_{sub} which has typical values between 10^{-3} and 10^{-2} fF/ μm^2 .

$$C_1 \approx \frac{w \cdot l \cdot C_{sub}}{2} \quad \text{Equation 63}$$

The total effect of the substrate, that is the effect of C_{ox} , R_1 , and C_1 , can be optimized by several methods. Firstly, the spiral can be moved to the highest possible level of interconnect to maximize the t_{ox} term. In addition, a patterned ground shield (PGS) can be used to increase Q by reducing capacitance to the substrate and the associated dielectric losses [21].

A PGS is a layer of low-loss, conductive material placed between the coil and the substrate. This typically decreases the SRF because the PGS effectively replaces the lower plate of the C_{ox} capacitor at a reduced t_{ox} . However, since the R_1 and C_1 terms are reduced, the energy lost by the coil is also greatly reduced. Typically, the PGS is constructed of either a metal layer or polysilicon in a slotted square pattern to minimize the ESR of the shield while preventing the formation of eddy currents in the shield. A polysilicon PGS was used in the test chip as a compromise between a low resistance PGS and a higher SRF.

While eddy currents can be prevented from forming in the PGS, they still form in the substrate. These eddy currents reflect resistance back to the inductor, modeled above by R_{eddy} , by causing the substrate to act like the secondary winding of a transformer whose primary is the spiral inductor. A crude formula for the reflected resistance is given as:

$$R_{eddy} \approx \frac{\sigma_{sub}}{4e} (\mu n f)^2 d_{avg}^3 \rho^{0.7} z_{n,ins}^{-0.55} z_{n,sub}^{0.1}, \quad \text{Equation 64}$$

where σ_{sub} is the substrate conductivity, d_{avg} is the average of the inner and outer turn diameters, and ρ is the fill factor. The quantities $z_{n,ins}$ and $z_{n,sub}$ are the normalized

distance to the heavily doped region of the substrate and the skin depth of the substrate, respectively.

While the above equation is crude and difficult to calculate for a given fabrication technology, it still yields some useful insights into coil design. Note that eddy losses are proportional to the square of the number of turns; so removing an internal turn reduces eddy losses proportionally more than it reduces inductance. Removing internal turns also decreases the fill factor ρ and the series resistance. As a result, hollow inductors are generally better than “full” inductors.

Finally, note that eddy resistance is proportional to the cube of the average diameter. Practically, this means that as the coil is scaled up, DC resistance will go down linearly, but eddy losses go up geometrically. This effect is more pronounced in CMOS processes where the substrate is heavily doped.

3.1.1.2 Three-Dimensional Inductors

A design similar to the flat spiral coil is the multiple layer spiral which trades increased resistance (due to an overall longer track) for higher inductance. The simplest multilayer inductor is the two layer stacked spiral. In this case, one layer spirals in toward the center and the other layer spirals out in the same direction. Due to the small distances between layers, the spirals are strongly coupled, resulting in a combined inductance nearly four times that of a single layer [22], [9]. One of the other tradeoffs of a multilayer inductor is a lower SRF. However, if a large number of metal layers are available, the SRF can be raised by maximizing the separation of the metal layers used (e.g. using layers 1 and 3 instead of 1 and 2).

Although a multilayer inductor was not fabricated on the test chip, this represents a promising future approach for designing a practical integrated inductor for an inductive switching supply. The calculation below is Lee's formula for a single layer square coil with dimensions similar to the test coils and only 6 turns. The track width and separation are 20 μm and 10 μm , respectively, and the outer diameter is 470 μm .

$$L_{single} \approx \frac{4\pi \times 10^{-7} n^2 d_{avg}^{1.27}}{2} \left[\ln\left(\frac{2.07}{\rho}\right) + 0.18\rho + 0.14\rho^2 \right], \rho=0.4688,$$

$$d_{avg}=385\mu\text{m}, n=6;$$

$$L_{single} \approx 13.9nH;$$

$$R_{dc, single} = \frac{l}{w \cdot t \cdot \sigma}, l=7.8\text{mm}, w=20\mu\text{m}, t=675\text{nm}, \sigma=37.8\text{S/m};$$

$$R_{dc, single} = 15.29\Omega.$$

The resulting inductance and DC resistance from using this coil in a stacked inductor is calculated by multiplying the single layer inductance and resistance by 4 and 2, respectively.

$$L_{double} = 4 \cdot L_{single}$$

$$L_{double} = 55.6nH$$

$$R_{dc, double} = 2 \cdot R_{dc, single}$$

$$R_{dc, double} = 30.58\Omega$$

These computed results are marginally better than the single layer spiral coil, L2 while using the same total number of turns.

The equivalent parallel capacitance due to coil-to-coil capacitance cannot easily be determined, but a parallel plate calculation should give a very conservative estimate for the worst-case capacitance. This estimate is the same formula as equation 61 except in this case, t_{ox} is the coil separation instead of the height above the substrate.

An alternative to the stacked spiral is the design presented by [23]. Their design primarily differs from the traditional stacked spiral by using all available metal layers and only placing a single turn in each direction (spiraling up or spiraling down) on each metal layer. So for a 4 layer metal process, their design would have 7 turns. They claim to be able to achieve the same inductance as a dual spiral inductor in only 16% of the area with superior SRF.

3.1.2 Chip Inductors

Chip inductors were researched using Digi-key's catalog of SMT components in the EIA 0402, 0603, and 0805 standard sizes. These sizes were chosen because they are roughly comparable to the size of a high pin-count, VLSI chip and chip carrier. Within this size range, Digi-key carries more than 5,900 products. To further limit the number of products to be researched, inductors were chosen which had an ESR less than 1 Ohm, a maximum DC current greater than 100 mA, and inductance of at least 27 nH. Within these limits, several lines of chip inductors were chosen from MuRata, Panasonic, and TDK to represent current, commercially available, high-quality chip inductors.

Each company offers chip inductors using one or more of the following manufacturing technologies: etched, multilayer (also known as monolithic), and wire wound.

Wire wound inductors are constructed using small gauge wire wound around either a ceramic or ferrite core which also incorporates the end contacts. Figure 23 shows the mechanical drawing of a typical 0603 size wire wound inductor. As with their larger through-hole brethren, wire wound inductors with ferrite cores offer higher inductances in smaller areas, but suffer from hysteresis and saturation.

Etched inductors are constructed very similarly to carbon film resistors. An inductor blank is formed from a prism, usually square, uniformly coated with a metal such as copper or aluminum. A spiral pattern is then etched either chemically or with a laser from one end of the prism to the other forming a spiral conductive path as shown in Figure 24. These inductors can have very tightly controlled inductance tolerances and very high SRF because of low shunt capacitance.

Stacked multilayer inductors, commonly called monolithic inductors, are manufactured by laminating layers of insulating material covered with metal spirals, as shown in Figure 25. These inductors can pack a large number of turns in a small area and

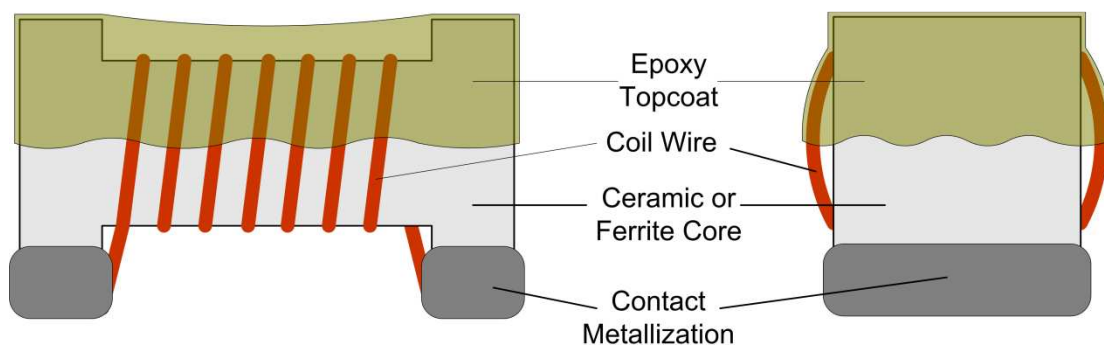


Figure 23: Mechanical drawing of a wire wound inductor.

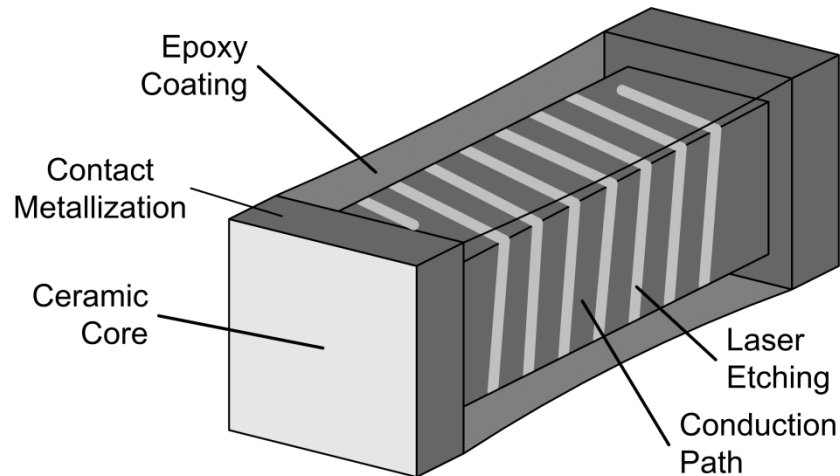


Figure 24: Laser-cut chip inductor.

are therefore capable of very high inductances, but at the cost of lower SRF due to layer-to-layer capacitance. Some monolithic inductors also feature either ferrite cores, magnetic shielding, or both.

A similar technology to the monolithic inductor is the film type. Instead of multiple layers, film type inductors are single layer spirals mounted on a ceramic carrier. These inductors typically feature tight tolerances, high Q, and high SRF due to the low stray capacitances.

Most of the chip inductors that were researched had SRF greater than 100 MHz. The chips with smaller inductance typically had much higher SRF – many were above 1 GHz. Conversely, chips with inductances higher than 1 μH had lower SRF with the very largest inductors having SRF in the tens of MHz. These inductors may still be useful since larger inductors do not need to be switched at such high frequencies in order to achieve sufficient energy transfer.

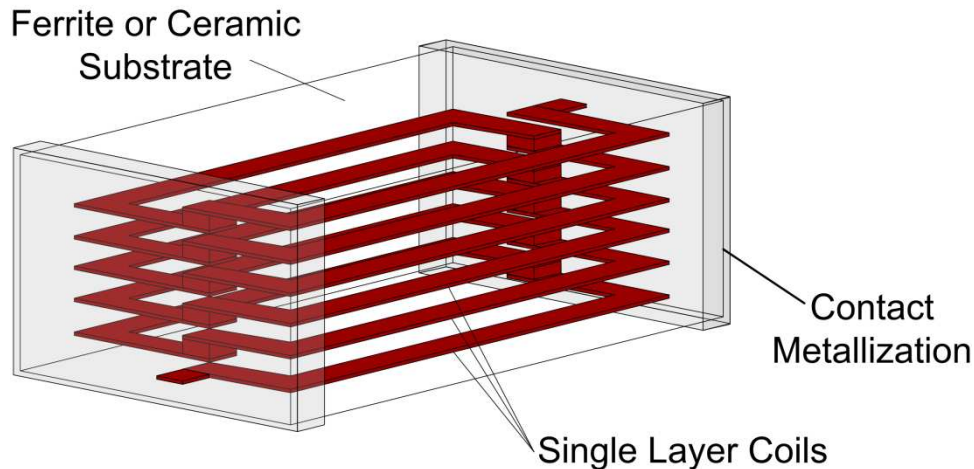


Figure 25: Multilayer inductor with ferrite shield.

The scatter plots for the various inductor series were grouped by physical size and were plotted in Figures 26, 27, and 28 which show inductance vs. series resistance. A logarithmic scale was used for inductance due to the wide range of values with series.

Generally, each series would be roughly linear on a normal-normal plot. These lines follow lines of constant Q on an L vs. R plot (Q lines are linear because $Q = \omega L/R$). This intuitively makes sense since L is increased by adding additional turns to an inductor which adds a proportional amount of both inductance and resistance due to the increased flux area and length, respectively. Note that this is not true of the integrated spiral coils because adding interior spirals adds considerable more resistance than inductance.

For ease of comparison, consider the maximum inductance available in each size for DC resistance less than 1 Ohm. For EIA 0402 size inductors, this is 68 nH – a fairly low value, but still much better than was achieved for the integrated coils. For EIA 0603 size inductors, the largest inductance under 1 Ohm was 4.7 uH, almost 100 times higher than that available in the 0402 size. For EIA 0805 size inductors, the largest inductance was 33 uH (although, this inductor has other undesirable characteristics).

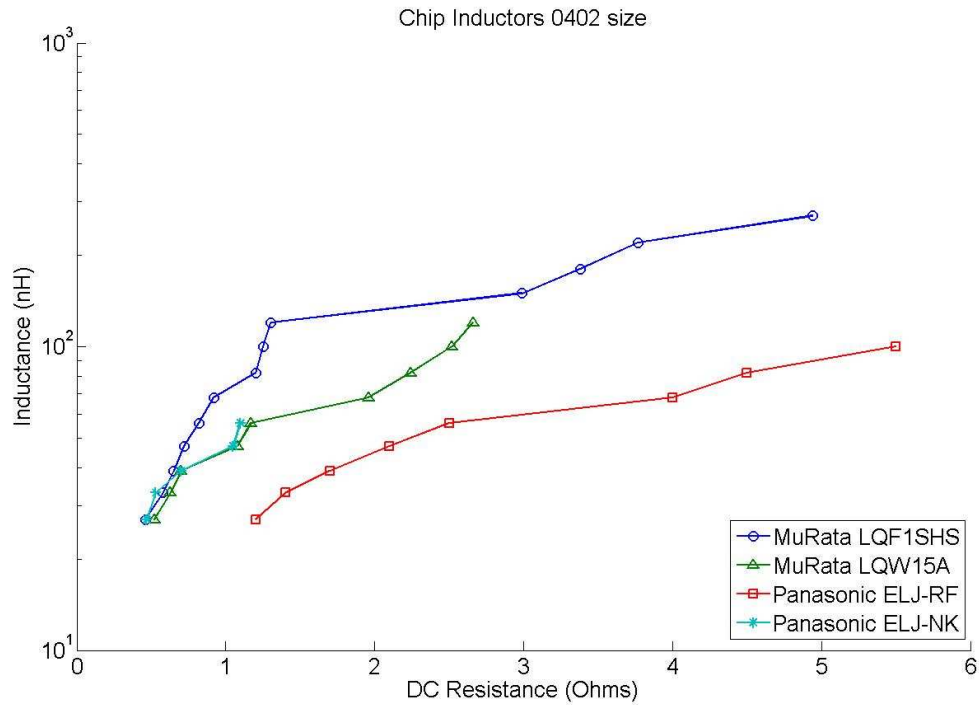


Figure 26: Chip inductors, 0402 size

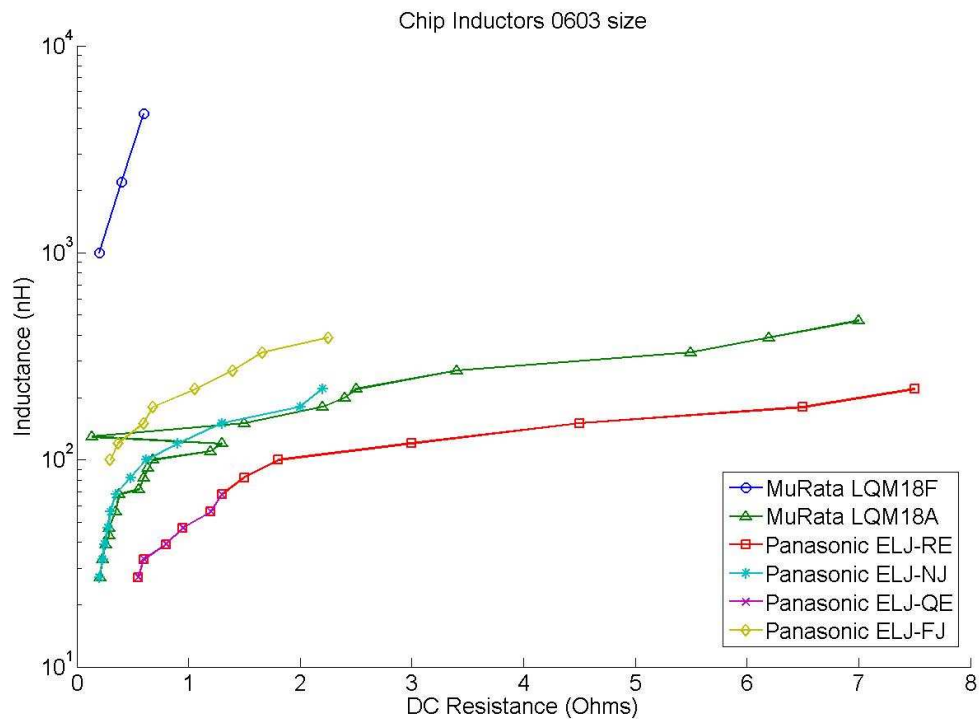


Figure 27: Chip inductors, 0603 size

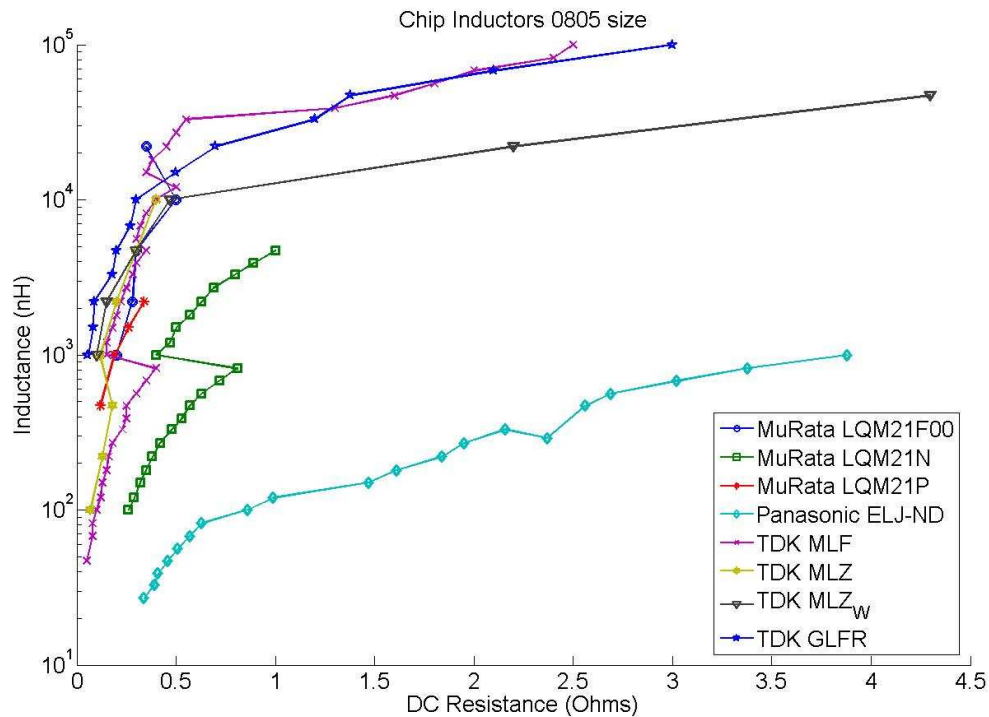


Figure 28: Chip inductors, 0805 size

So in summary, larger inductors are generally capable of higher inductances, lower resistances, and thus higher Q values, but generally at the expense of lower SRF.

3.1.3 Integrated Capacitors

There are several ways to build capacitors right on the die of a chip. Generally speaking, these all involve using different layers on the chip – metal, poly, and substrate or diffusion – to create the plates of a capacitor. The general capacitance formula, ignoring fringe capacitance, is given by the following formula,

$$C \approx \epsilon \frac{A}{H},$$

Equation 65

where A is the area of one of the plates (assumed to be identical), H is the distance between them, and ϵ is the dielectric constant of the dielectric material between the plates.

Perhaps the easiest way to make a capacitor is by laying out large polygons on adjacent metal layers or two poly layers if they are available. These conductive shapes, identical or nearly so in shape and size, form the plates of a simple parallel plate capacitor and the oxide acts as the dielectric. Although the dielectric constant of silicon dioxide is a relatively low 3.9, it is also a low-loss dielectric and therefore contributes very little to the ESR of the capacitor.

While parallel plate capacitors are easy to layout, they do have some disadvantages. First, they are not very area efficient. A large capacitance value requires a large area, wholly dedicated just to the capacitor. Second, the bottom plate also forms a capacitor with the substrate and anything else below it. If the bottom plate is tied to ground, this is not much of an issue; if the top plate is tied to ground, then this effect adds to the overall capacitance of the bottom plate to ground. Unless the bottom plate is tied to the same potential as the substrate, the substrate always adds to the ESR of the capacitor and lowers its Q factor.

An alternative to the simple parallel plate capacitor is the fingered capacitor where the total flux storage is a combination of vertical and lateral flux [19]. This design has several advantages over simple parallel plate capacitors. First, the lateral flux contributes a significant amount of capacitance above the capacitance of a simple plate capacitor of the same chip area. As a result, the capacitance per area is increased, so for a given capacitance, the area consumed is smaller. This decreases the amount of parasitic

capacitance to the substrate. Second, the horizontal flux steals some flux from substrate parasitic capacitor, lowering ESR and raising Q.

Taking the idea of lateral flux capacitors ever further, the perimeter – and the capacitance – of the plates can be maximized by using a fractal pattern. Presumably, these capacitors will suffer from the parasitic substrate capacitor. One possible way to alleviate this drawback would be to use a fractal edge on the fingered capacitor discussed above. In theory, this should nearly maximize the lateral flux and the vertical flux for a given area capacitor.

Yet another way to utilize lateral as well as vertical flux is the woven capacitor where fingered capacitors are laid on top of each, but rotated 90 degrees each layer. Since the distance that current flows is shorter, the interwoven capacitor has lower ESR than a simple interdigitated capacitor. The shorter current paths and their being orthogonal also decreases the series inductance of the capacitor. This is a potentially important side benefit because series inductance causes a capacitor to resonate and limits the usable frequency range of the capacitor.

The final type of capacitor, and the most space efficient, is the MOSCAP. MOSCAP's utilize the charge storage characteristics of the gate-channel capacitance of a MOS transistor, as shown in Figure 29(a). Because of the extremely small distance between the gate and substrate – only a few dozens of atoms thick in the deep, sub-micron technologies today – the capacitance is very high. Unfortunately, the capacitance value is voltage dependent, especially around the threshold voltage. However, as long as a sufficiently large D.C. voltage is applied to keep the channel strongly inverted, the capacitance should be fairly stable for small applied voltages and currents.

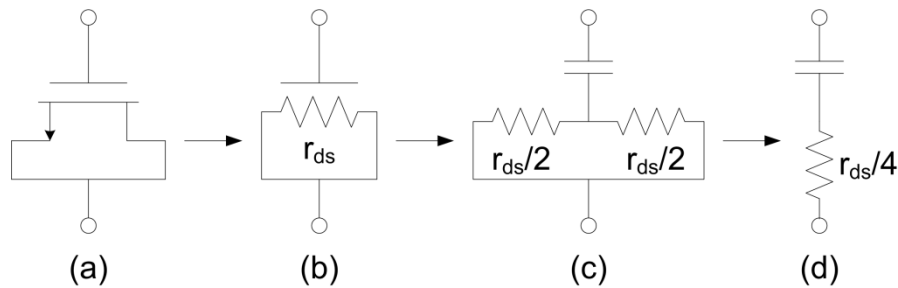


Figure 29: MOSCAP model showing derivation for the first order ESR.

One of the biggest drawbacks of the MOSCAP is the quality factor. The primary limiting factor of MOSCAP quality is the channel resistance. As a result, high-quality MOSFET's are also minimum channel length devices. A first order estimate of the ESR is shown in Figure 29(b), (c), and (d) where the resistance is first split between the drain and source with the capacitance in the middle and finally the parallel combination of the two half-resistances [19]. This estimate assumes that all the capacitance is located at the middle of the channel and passes through half of the total channel resistance.

A better model can be easily derived by assuming that the capacitance is evenly distributed across the gate. As long as the AC currents are low and the frequency is sufficiently low to allow charges to distribute themselves across the channel, this should be a good assumption. As seen in Figure 30, the resistance from a point in the channel to either drain or source implant, R_{CS} , increases linearly from zero at the ends to the middle point where it equals $r_{ds}/2$.

Considering only part of the channel between the midpoint and one of the implants, the resistance distribution is a right triangle with maximum at the midpoint and zero at the implant. Since the capacitance is evenly distributed, the average capacitance/resistance product is equal to half the total gate capacitance times the average resistance of the half channel.

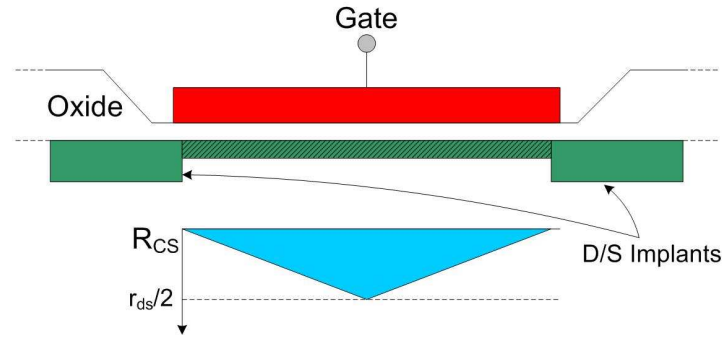


Figure 30: MOSCAP model (not to scale) showing the gate (red), drain and source implants (green), and induced channel (hatched green). The resistance seen in the channel with respect to either source or drain implant, R_{CS} , is show in light blue. The drain and source contacts are not shown for the sake of simplicity.

From basic calculus, the center of mass of a right triangle is located $1/3$ of the length from the 90 degree vertex on the two adjacent sides. Therefore, the value of the average resistance of the half channel is equal to the value of the resistance $2/3$ of the way from the implant to the midpoint of the channel – $r_{ds}/3$. If this value is used in Figure 29 (c) instead of $r_{ds}/2$, the equivalent parallel resistance is $r_{ds}/6$.

According to Lee, who states the real value is close to $r_{ds}/12$, this revised estimate still overestimates the ESR by up to 100%, although no explanation is offered for the source of the error. Several possibilities include a nonuniform channel thickness or resistance, nonuniform distribution of capacitance, or some second order effect involving the displacement currents in the channel.

The overlap and fringing capacitances between the gate and source/drain would also add to the overall capacitance while adding little ESR because of the greater conductivity of the implant regions. The relative impact of these two factors should increase with decreased channel length since they are the same value regardless of the

actual gate-to-channel capacitance. Finally, the junction capacitance between the inverted channel and the bulk substrate will add a bit of capacitance as well.

3.1.4 Chip Capacitors

Chip capacitors come in a variety of technologies, each with its own advantages and disadvantages.

Film type capacitors, both thick film and thin film, generally are good quality capacitors. They are formed by metalizing a dielectric film. The resulting capacitors have low dissipation factors – typically below 0.01. Unfortunately, they are not very space efficient and large value capacitors also require large physical dimensions.

Tantalum capacitors have by far the highest capacitance per volume of the common capacitor types. Unfortunately, tantalum capacitors have a number of drawbacks. First, they typically have relatively high ESR and ESL (although still very small compared to electrolytic types). Secondly, they are prone to two unique failure modes – burn through and ignition. Burn through causes the thin dielectric to break down and short out – sometimes even below the voltage rating. Ignition is a failure mode that follows either burn through or normal dielectric breakdown when the chemicals inside the capacitor undergo an exothermic reaction that continues even after voltage is removed from the capacitor.

Miniature aluminum electrolytic caps are manufactured, but not in the sizes being considered. In any case, aluminum capacitors typically have much higher dissipation factors than any of the other types discussed and in addition contain a liquid electrolyte which is toxic should it leak following a device failure.

The most common type of chip capacitor is the multilayer ceramic chip capacitor (MLCC). The MLCC offers high capacitance in very small packages - down to EIA-0201 size (0.020" x 0.010"). Ceramic capacitors also typically have low ESR – typically stated as dissipation factor (DF) or $\tan \delta$ instead of Ohms. The DF is the ratio of resistance to reactance at a certain test frequency.

The MLCC is constructed of multiple layers of metal foil separated by layers of dielectric. The electrodes are placed at either end of the chip and connected to alternating layers of foil. In this way, each layer of foil forms a capacitor with the foil layer above and below it.

MLCC capacitors are available from many companies, but they typically use the same types of dielectric, referred to by a three digit alphanumeric identifier. The dielectrics vary in their tolerances, voltage ratings, temperature coefficients, and dissipation factors. The three classes are discussed below.

3.1.4.1 Class I – C0G, NP0, C0H, C0J, C0K

The class I dielectrics are the highest quality and highest cost dielectric available. The most common class I dielectric is C0G (also known as NP0, frequently listed under “C0G,NP0”). The dielectric is a nonferroelectric material which results in very low dielectric losses. Typically, but not always, capacitors that use class I dielectrics are manufactured to higher tolerances than other types of MLCC. In addition, the dielectric is temperature compensated so the capacitance varies very little with temperature, 0 ± 30 ppm/ $^{\circ}$ C by definition for C0G.

The major problem with class I dielectrics, other than cost, is that the dielectric is not very volumetrically efficient, i.e. for a given capacitance, they are a large physical size. However, when a design calls for either high-quality or stable temperature characteristics, class I dielectrics are the capacitor of choice.

3.1.4.2 Class II – X5R, X7R, X5S, X7S

The class II dielectrics are ferroelectric materials yielding higher volumetric efficiency but higher temperature variance than the class I dielectrics – typically 10% to 15% across the operating temperature range. The dielectric has higher dielectric losses which leads to a higher DF than the class I dielectrics.

The most common dielectrics in class II are X7R and X5R, both of which have a lower cost than COG,NP0 capacitors. These capacitors are a good choice when cost is a concern but good quality, stable capacitors need to be used.

3.1.4.3 Class III – Z5U, Y5V

The final class of dielectrics, class III, are best used when a large value capacitor is needed, but the exact value is not critical. These capacitors typically have a tolerance of -20,+80% and can vary by +22% to -56% over their temperature range. As if this were not bad enough, class III capacitors also have the highest DF of the three classes – typically as high as 20%.

Because of their high DF, they are not suitable for the output stage of a switching supply because they would induce high ripple voltage. However, they are very useful for

decoupling applications, such as the input stage of a switching supply, where their high capacitance, low cost, and nearly constant voltage load make them a good design choice.

The highest capacitance value and DF for each size and dielectric class is shown in Table 3. In general, class II capacitors were the largest capacitance size. Compared to class I, the class II capacitors had dissipation factors 10 times as high, but they offered more than 1000 times the capacitance in the same footprint.

3.1.5 Integrated Power Switches

The power transistor options available in CMOS processes are very limited. The simplest way to build a large-current carrying transistor is to layout strips of minimum length transistors with shared drain and source diffusions. Using gangs of wide, short transistors like this minimizes the drain-source resistance, R_{DS} , of the transistor. A secondary benefit of using minimum length transistors is that it also results in low gate capacitance.

It is important to note that the transistors should be kept deep in the triode region to minimize conduction losses. The reason for this is easily seen from a graph of drain current (I_D) vs. drain-to-source voltage (V_{DS}). The I_D vs. V_{DS} curve is steepest in the triode region and is nearly horizontal in the active region. Resistance is inversely

Table 3 - Capacitance and dissipation factor (DF) for selected sizes and dielectrics.

C,DF	EIA-0201	EIA-0402	EIA-0603	EIA-0805
Class 1	100pF, 0.001	1000pF, 0.001	10,000pF, 0.001	0.033uF, 0.001
Class 2	0.22uF, 0.1	4.7uF, 0.1	22.0uF, 0.1	47.0uF, 0.1
Class 3	0.047uF, 0.2	1.0uF, 0.2	10.0uF,	22.0uF, 0.2

proportional to the slope of the I_D , V_{DS} curve so it can be seen that in the triode region, the resistance is low and linear but is very high (ideally infinite) in the active region.

The drain current, drain-source voltage relationship in the triode region is given by equation 66.

$$I_D = \mu_n C_{ox} \left(\frac{W}{L}\right) \left[(V_{GS} - V_{th})V_{DS} - \frac{V_{DS}^2}{2} \right] \quad \text{Equation 66}$$

In this equation, C_{ox} is the gate capacitance per unit area and μ_n is the electron mobility. Equation 66 is valid until the channel reaches pinch-off and the transistor is in the active region. Pinch-off occurs when the voltage differential across the channel, V_{DS} , is equal to the effective voltage defined in equation 67.

$$V_{eff} = V_{GS} - V_{th} \quad \text{Equation 67}$$

Substituting this constraint into equation 66, the maximum triode current, and the ideally constant active region current, can be found using equation 68.

$$I_D = \frac{\mu_n C_{ox}}{2} \left(\frac{W}{L}\right) V_{eff}^2 \quad \text{Equation 68}$$

For the C5N process used in the test chip, the threshold voltages for minimum width devices are at most 0.78 V and -0.93V for N-channel and P-channel transistors, respectively. The K' value, which is the product of the carrier mobility and unit area

capacitance divided by 2, was $57.1 \mu\text{A}/\text{V}^2$ and $19.0 \mu\text{A}/\text{V}^2$ for N-channel and P-channel transistors, respectively. Using these values, the maximum drain current for $20 \times 0.6 \mu\text{m}$ transistors can be calculated assuming maximum inversion in a logic level design (0 to +5V). It is also very simple to derive the saturation current per μm of transistor width for minimum length devices ($L=0.6\mu\text{m}$).

N-Channel

$$I_{Dsat} = 57.1 \mu\text{A}/\text{V}^2 \cdot \left(\frac{20}{0.6}\right) (5 - 0.78\text{V})^2$$

$$I_{Dsat} = 33.895\text{mA}$$

$$I_{Dsat} = 1.694\text{mA}/\mu\text{m}$$

P-Channel

$$I_{Dsat} = 19.0 \mu\text{A}/\text{V}^2 \cdot \left(\frac{20}{0.6}\right) (-5 + 0.93\text{V})^2$$

$$I_{Dsat} = 10.491\text{mA}$$

$$I_{Dsat} = 0.525\text{mA}/\mu\text{m}$$

The minimum on-state resistance, which occurs at zero current and maximum gate drive, can be found by evaluating the derivative of equation 66 at $V_{DS} = 0$ and inverting it (since resistance is the inverse of the curve). The derivative of equation 66 is shown in equation 69.

$$\frac{dI_D}{dV_{DS}} = \mu_n C_{ox} \left(\frac{L}{W}\right) [(V_{GS} - V_{th}) - V_{DS}] \quad \text{Equation 69}$$

Substituting the known quantities and inverting yields R_{DS} values of 62.25 Ohms and 193.97 Ohms for $20 \times 0.6 \mu\text{m}$ N-channel and P-channel devices, respectively. As with the saturation currents, the channel conductance can be expressed in terms of Siemens per μm of width, with the result that N-channel transistor conductance is $0.803 \text{ mS}/\mu\text{m}$ and P-channel transistor conductance is $0.258 \text{ mS}/\mu\text{m}$. The graphs of drain current and channel resistance are shown in Figures 31 and 32.

As was discussed in the previous section on integrated capacitors, CMOS transistors act as voltage dependent capacitors. This can be desirable if one is trying to make high-value capacitors in a CMOS process, but if the transistor is being used in a high-frequency circuit such as a high-speed digital circuit or switched capacitor filter, this capacitance is a source of power loss.

Transistors, therefore, cannot be made arbitrarily large for a given application. At some point, there is a tradeoff between the dynamic power lost due to charging and discharging the gate of the transistor and the resistive power lost in the conducting channel. Above this point, decreasing the on-state resistance by increasing the size of the capacitor causes the overall power loss to increase rather than decrease.

Finding this point is crucial both for DC-DC converter design and also for general high-speed circuit design. The power lost due to switching a transistor at a certain frequency is found by multiplying the frequency by the power lost in a single cycle. An estimate for the single cycle power loss of a transistor is found by simply calculating the energy stored in the gate-channel capacitor and assuming that this is all lost when the capacitor is discharged. This is generally true, but since the capacitance is voltage dependent and is composed of junction capacitances, fringing capacitances, and the

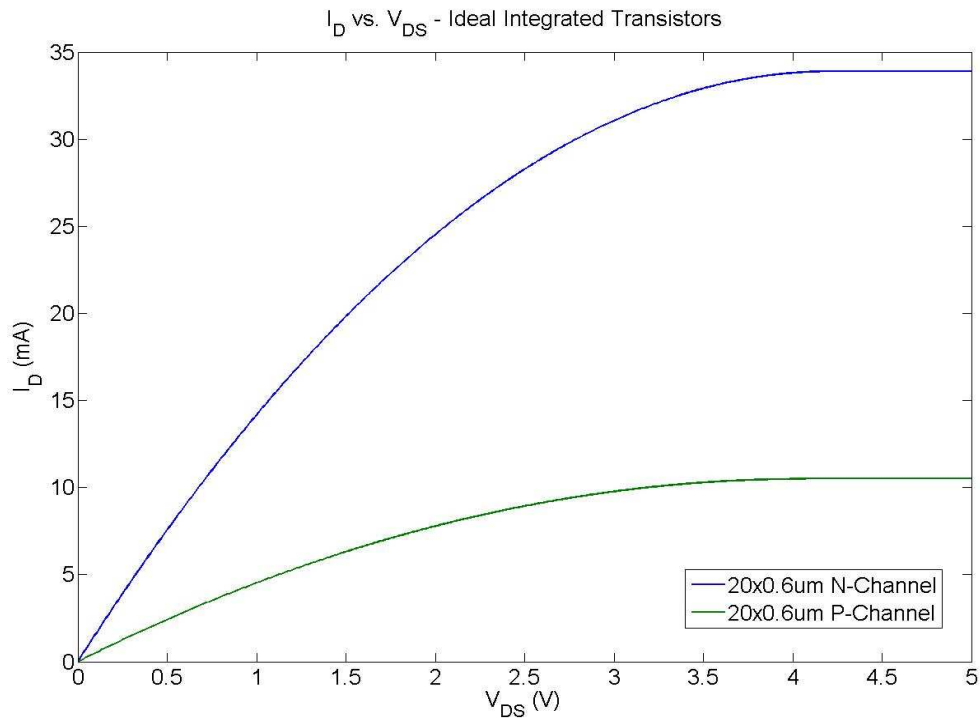


Figure 31: Ideal I_D vs V_{DS} relationship. When the curve goes flat, the transistor is in saturation and further increases in V_{DS} do not result in additional drain current.

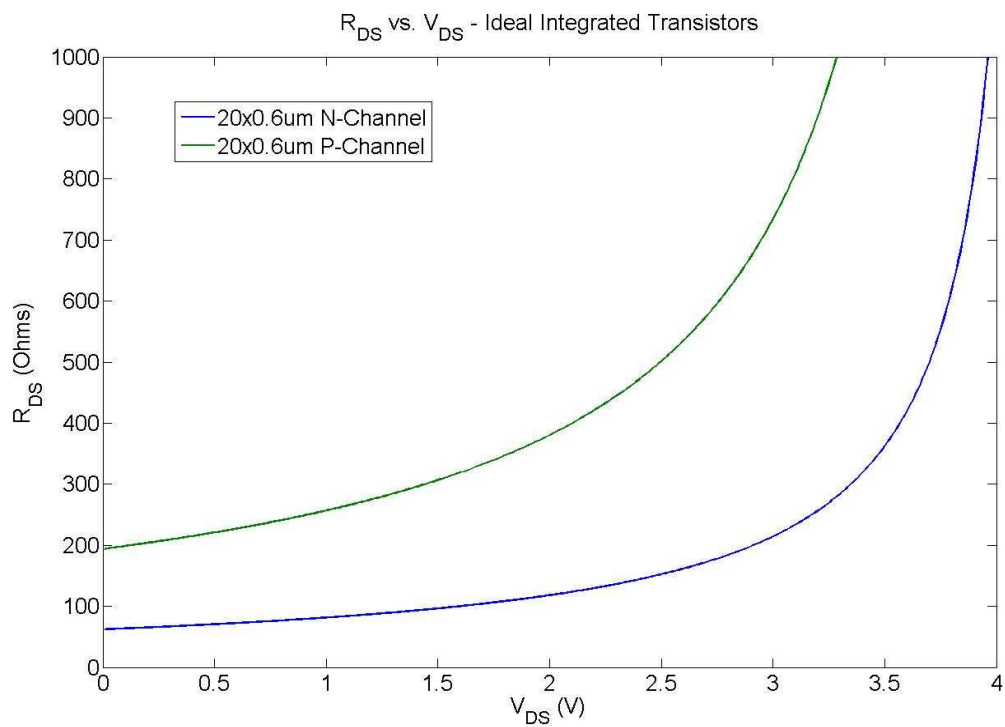


Figure 32: Drain-source resistance for two ideal transistors. Note that the resistance changes rapidly as the saturation voltage is reached.

parallel plate capacitance directly under the gate to the channel, the estimate for capacitor energy is just an estimate. The energy stored in a capacitor is given by equation 70 which yields the frequency dependent power transistor power loss in equation 71.

$$E = \frac{1}{2} CV^2 \quad \text{Equation 70}$$

$$P_{switching} = \frac{1}{2} f CV^2 \quad \text{Equation 71}$$

Further substituting the first order equation for transistor gate capacitance, $C_{ox}WL$, the switching loss of a transistor of given size at a certain frequency can be estimated, as shown in equation 72.

$$P_{switching} = \frac{1}{2} f C_{ox} WLV^2 \quad \text{Equation 72}$$

The resistive loss in the above circuit depends on the shape and amplitude of the current waveform. The loss can be estimated by multiplying the square of the average current by the R_{DS} of the transistor at the average current level. This estimate works well as long as the maximum current is well below the saturation current where the R_{DS} is fairly constant. Assuming that the channel resistance does not appreciably affect the current waveform, then it can be assumed that resistive losses are determined by the gate drive voltage and the dimensions of the transistor.

A key insight here is that the resistive losses decrease linearly with gate drive voltage but the switching losses increase with the square of the driving voltage.

Therefore, as the switching losses gain parity with the resistive losses, the resistive losses can be kept constant by widening the transistor and decreasing the drive voltage. At the same time, the area term of the equation for the switching losses goes up linearly with the width but the voltage squared term will drop due to the lower gate drive voltage.

3.1.6 Discrete Power Switches

There are several technologies that could theoretically be used for the power switch – BJTs, Darlington pairs, J-FETs, IGBTs, and MOSFETs.

Bipolar transistors such as the BJT and Darlington (which is just two cascaded BJTs to boost the total current gain) are best used when the current, saturation voltage product is lower than the current and R_{DS} product of a MOSFET. Typical saturation voltages are between 0.4 – 0.3 volts and a typical R_{DS} for a good MOSFET is well under 100 m Ω . For a worst-case MOSFET and a best-case BJT (or Darlington), the current where power dissipation is equal is about 3 A.

Power loss due to the base current means that the MOSFET is still a good choice at higher currents, especially with better quality MOSFETs. Even higher current MOSFET switching supplies can be designed, such as the 50+ Amp microprocessor supplies, using multiple phases. These designs are beyond the scope of this thesis, but are mentioned as an example of how far MOSFET switching supplies can be pushed.

The JFET works by varying the depth of the conduction channel in a long channel of semiconductor. It does this by varying the width of the depletion region surrounding the gate implant. Because of their relatively long channels, JFETs typically have fairly high R_{DS} , between 3 Ω and 300 Ω , which makes them unsuitable for a switching supply.

Even the low R_{DS} power JFETs offered by Vishay have an R_{DS} of 100mOhms which is ten to 100 times higher than power MOSFETs. The main advantage of JFETs in some power applications is their much lower gate capacitance (which is just the junction capacitance of the reverse biased PN junction). For very high frequency circuits, such as high frequency DC-DC converters and digital audio amplifiers, where the switching power losses are comparable to channel resistance losses, a transistor with a low gate capacitance but higher R_{DS} would be a good tradeoff.

Isolated gate bipolar transistors (IGBTs) are frequently used for very high current, very high input voltage power supplies. They can be thought of as a cross between a MOSFET and a BJT where the input is isolated like a MOSFET, but the power switch exhibits the low saturation voltage of a BJT. They are constructed as four-layer, three terminal devices in such a way that the MOSFET input transistor feeds current into the base region of the vertical BJT in much the same way that a Darlington transistor functions. This is actually an oversimplification of all the processes going on in an IGBT, but it is a workable conceptual model. Like BJTs, IGBTs are not typically used in low voltage, low current applications because MOSFETs perform very well in these applications.

MOSFETs clearly are the best technology for small switching supplies because of their very low power dissipation and zero gate current. Discrete power switches have several tradeoffs versus integrated power switches. Whereas integrated power switches are limited to relatively low gate and V_{DS} voltages, discrete MOSFETs are typically designed to handle larger voltages on the order of 20 volts or more for both gate and drain-to-source voltages. These switches are also capable of handling continuous

currents measured in amps instead of milliamps. This performance comes at the cost of increased gate capacitance and, in some cases, a lower maximum usable frequency.

There are several popular, standardized transistor sizes available with footprints smaller than 3mm x 3mm. The most common size is the SOT-23 and its derivatives. Some of the common sizes are listed in Table 4.

A full survey of all transistors in the 3mm x 3mm size range would require a careful examination of over 1,200 transistors listed by DigiKey. Instead, the SO-323, SC-70 size will be examined. This size transistor is smaller than the much more common SOT-23 but still offers a decent selection of transistor models to choose from (over 175 unique part numbers.)

Table 4: Comparison of SMT MOSFET package types and sizes.

Package name(s)	Width (mm)	Length (mm)
SOT-23 (roughly the size of EIA1206 package) – 844 items	0.11” (2.80)	0.047” (1.2)
SOT-416, SC-75 – 67 items	0.063” (1.6)	0.032” (0.8)
LGA (leadless grid array)	0.067” (1.7) varies	0.043” (1.1) Varies
SOT883, SC-101 – 5 items	(1.0)	(0.6)
SOT-346, SC-59 – 46 items	(3.0)	(1.6)
SOT-323, SC-70 – 175 items	(2.0)	(1.25)
SC-89 – 40 items	(1.6)	(0.88)
SOT-143 – 17 items	(2.9)	(1.3)
SOT-523 – 33 items	(1.6)	0.032” (0.8)
SOT-723 – 15 items	0.047” (1.2)	0.032” (0.8)

As the main switch will be a high-side switch, the survey was further refined to P-channel MOSFETs. Finally, MOSFETs with on-resistance higher than 1 Ohm were omitted. This left 14 parts to sort through. All of the remaining parts had maximum current ratings above 0.5A and maximum power ratings of approximately 300 mW. There are therefore three discriminating factors among these transistors: threshold voltage, gate charge, and drain-source resistance.

As was seen in previous sections, threshold voltage plays an important part in determining the resistance of the transistor. Beyond that, it is important that the transistor can be adequately turned on given a range of input voltages. For this reason, it is important to find a transistor with a low threshold voltage. Fortunately, all the transistors that were examined had maximum threshold voltages below 1.5 volts with typical values in the 0.7 to 0.9 volt range.

The gate capacitance of a transistor is specified by manufacturers in one of two ways: total gate charge or input capacitance, C_{iss} . The two measures are related but not directly comparable.

Total gate charge, Q_T , is the actual amount of charge in coulombs placed on the gate to achieve a certain gate voltage and V_{DS} or I_{DS} . Because the capacitance of the gate changes with the applied voltage, which changes as charge is added to the gate, the capacitance of the gate at a certain operating point is not exactly the same as would be calculated from the charge and voltage on the gate. To put it another way, the capacitance at the operating point, dQ/dV , is not equal to Q_T/V as is the case with ideal or near ideal capacitors. Gate charge is most useful when determining how much charge will be displaced in a switching circuit.

Input capacitance is the total instantaneous capacitance of the gate with regards to both source and drain at a specified operating point. Data sheets specify this capacitance at certain values of V_{GS} , V_{DS} , V_{GD} , and I_{DS} . Unfortunately, the operating points are not standardized even in data sheets from the same company. Nevertheless, this metric is a better measure than gate charge when determining the actual small signal capacitance seen at the gate of a transistor.

Drain-source resistance, also frequently referred to as on resistance or simply R_{DS} , is similarly described as a certain resistance at a specified gate voltage and load current. This figure is the most important in determining the resistive losses in a transistor. The drain-source resistance, as well as the gate charge and input capacitance, of several SOT-323 transistors is shown in Table 5.

Table 5: SOT323 P-channel transistors. Note the inverse relationship between R_{DS} and C_{iss} or gate charge.

Part Number	$R_{DS}@ V_{GS} = 2.5V$	C_{iss}	Gate Charge
DMG1013	0.7	59.76pF	622.4pC
SI1303DL	0.7		1.7nC
SI1307EDL	0.435		3.2nC
NTS2101P	0.078	640pF	6.4nC
AO7413	0.15	512pF	4.9nC
AO7407	0.137	540pF	6.2nC
AO7401	0.207	409pF	5.06nC
DMP2160	0.09	627pF	
SI1305DL	0.315		2.6nC
DMP2240	0.134	320pF	
NTS4101P	0.104	603pF	6.4nC
SI1305EDL	0.315		2.6nC
SI1303EDL	0.56		1.9nC
BSS209PW	0.563	89.9pF	0.92nC

One final difference to note about discrete transistors is that the body diode, the diode formed between the body and the drain due to the source-body connection, is usually designed to be quite robust. Often, the body diode is rated to handle as much current as the transistor itself. This feature is useful in the design of inductive circuits such as inductive voltage supplies and motor controllers since this diode rectifies the inductive current and prevents dangerous voltage spikes from damaging the rest of the circuit.

CHAPTER 4

SIMULATIONS, CALCULATIONS, AND OPTIMIZATIONS

Now that switching theory has been explained and derived and circuit elements, both integrated and discrete, have been examined, the central question of the thesis will be examined: Can a micropower switching regulator be designed on the scale of an integrated circuit? To answer this question, the criteria have to be defined: what performance must the design meet in order to be acceptable?

The first criterion is efficiency. If the switching converter does not perform at least as well as a linear regulator, there is no point in wasting the extra silicon to build the converter. Efficiency, as has already been seen, is affected by resistive losses and dynamic losses. Resistive losses are due to resistance in the power switches and reactive components. Dynamic losses are due to switching capacitances in the power switches and controller, and the high-current, high-voltage transitions in the switches during turn-on and turn-off.

The second criterion is output ripple. Output voltage ripple is a function of the inductor ripple current which in turn is a function of the inductance, series resistance of the coil, input and output voltage, and conduction time. Voltage ripple is also directly related to the size and quality of the output capacitor.

To summarize, a successful on-chip converter will have good efficiency and low output ripple. The question then becomes: can high-enough quality components be built in silicon or must they be placed off-chip?

Answering this question requires a combination of simulations, calculations, and physical devices. Although a test chip was built, the device was not well suited to testing anything other than the spiral inductors. The test boards are not designed to exactly mimic an integrated circuit, but to show similar operation and to gain further insight into the workings of the buck converter. The boards were designed to use devices similar in size to EIA 0805 which are easily soldered by hand. These circuits are examined first and data gathered will be extrapolated from these boards to more general controllers.

Simulations were run in Matlab to test the behavior of circuits with a range of device characteristics. The results of these simulations will be compared to the results of the actual circuits tested. This should give some confidence that the simulation results carry over to the smaller devices.

The results of the simulations and testing will further be generalized to integrated circuit technologies other than the C5N technology that was used to create the test chip. The idea is to determine if moving up or down in feature size helps the efficiency of an integrated switching converter.

At the end of this chapter, it should be obvious what fabrication technologies if any are feasible for designing integrated switching supplies. If it is feasible, the conditions under which it is possible will be stated. Additionally, if it is feasible to place some components off-chip and onto the chip carrier, those components will be identified.

4.1 Test Circuits

The test circuits were assembled on custom-designed circuit boards which included a high-side P-channel MOSFET, a diode, a chip inductor, and a chip capacitor. Provisions were made to add additional resistance to the coil as well as for testing the individual components using test structures. The circuit design is shown in Figure 33. Note that no controller is provided for – the controller is a frequency generator generating a square wave of manually variable duty cycle.

Six boards in total were ordered along with a variety of parts with which to assemble them. Two different types of SOD-123 diode were ordered: 1N4448 PN switching diodes and B0520 Schottky diodes. Three types of SOT-323 p-channel MOSFET were ordered: DMP2240, DMG1013, and NTS2102. EIA0805 inductors were ordered in both 100uH and 1uH sizes. Ceramic EIA0805 capacitors were ordered in 22uF ($R_C = 5.8\Omega$) and 4.7uF ($R_C = 3.39\Omega$). The sizes of inductor and capacitor correspond with the largest value components in EIA0603 and EIA0805 sizes.

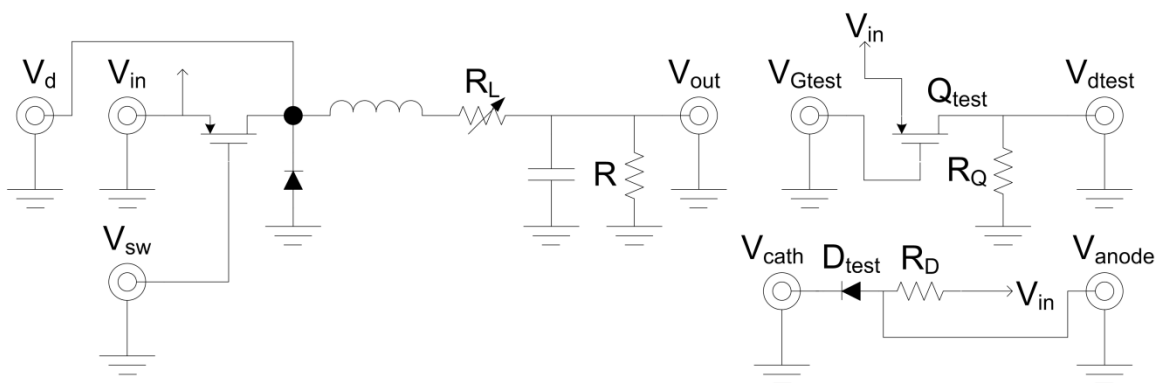


Figure 33: Switching supply test board with test structures.

Each board used a different combination of parts in such a way that each board varied from one other board by at most two parts so as to isolate the influence of the individual parts. The parts used for the boards are tabulated in Table 6.

4.1.1 Initial Testing and Observations

Initial testing of the circuit boards used a frequency generator at 100kHz, a load resistance of $1k\Omega$, and input voltage of 5 volts. The output voltage theoretically should have been approximately 2.5 Volts since the duty cycle was 50%. However, the output was five volts.

The reason for this is that there is nothing to stop the supply from completely charging the output capacitor up to the input voltage when the load current is small. The diode does not actively pull the common drain-diode-inductor node (hereafter referred to simply as “the common node”) down to zero since current cannot flow backwards through a diode. As a result, if the ripple current is much larger than the load current, the output capacitor will continue to charge due to the net positive current through it. In other words, the converter is in DCM mode.

Table 6: Board assembly parts

Board Number	Diode	Inductor	Capacitor	MOSFET
1	B0520 Schottky	100uH	22uF	DMP2240
2	B0520 Schottky	100uH	4.7uF	DMP2240
3	B0520 Schottky	100uH	22uF	DMG1013
4	1N4448 PN	1uH	22uF	DMG1013
5	B0520 Schottky	1uH	4.7uF	NTS2102
6	1N4448 PN	1uH	4.7uF	NTS2102

The open loop control used on the board only exacerbated this problem since the duty cycle remained at 50% regardless of the output voltage. Nevertheless, all switching supplies will have this problem at low load currents unless some method of drastically reducing the duty cycle, such as PFM or pulse skipping, is used.

Further testing of the boards partially fixed the output voltage problem by using 100 Ω resistors and a switching frequency of 1 MHz. The higher frequency leads to a smaller ripple current and the lower load resistance draws 10 times more load current. Initial testing using these new parameters showed that the output voltage was lower than 5 volts and could be varied with the duty cycle.

The transistors were tested using the test structure shown in Figure 33. Initial testing only tested the on-state resistance of the MOSFETs. Instead of a fixed resistor, R_Q , a source meter was used to pull a constant 90 mA from the drain of the transistors. Initially, smaller load currents were used, but the source-meter voltage varied almost imperceptibly right up until the threshold voltage was reached.

The voltage compliance for the load current source was set to +/- 5 volts so as not to exceed the drain-source voltage rating. Source meters were used to supply the input and gate voltages so as to ensure their accuracy.

The testing in this scenario showed that the transistors, though tiny in physical size, had reasonably low on-state resistance and low threshold voltage. The minimum on-state resistance for all transistors was found at $V_{GS} = -5V$ and increased slowly with gate voltage until approximately -3 volts, at which point it rapidly increased. The threshold voltage was determined as the gate voltage at which the source meter reached its compliance limit. On-state resistances and threshold voltages are shown in Table 7.

Table 7: MOSFET on-resistances and threshold voltages.

Transistor	R_{ds} at $V_{GS}=0$	Threshold voltage (est)
DMP2240	1.2767 Ω	-0.9 volts
DMG1013	1.6333 Ω	-0.9 volts
NTS2102	1.2422 Ω	-0.7 volts

Further testing of the test board will include determinations of the diode forward voltage drop and differential resistance, estimates for total gate capacitance, and efficiency measurements for the switching circuit with a variety of load currents and inductor resistances.

4.2 Test Board Results

The boards were tested to determine the efficiency of a PWM switching regulator with diode rectification in both DCM and CCM modes. The first test performed was a constant load current, variable frequency, and duty cycle test. The results of this test are shown in Figures 34 and 35. Note that boards 5 and 6 could not be set to output 3.3 Volts except at high frequencies, so these boards were tested with 100mA current and 3.8 Volts output.

The boards that used the larger coil definitely had an efficiency advantage over the boards that used the smaller coils. The lower performance is in spite of the fact that the smaller coils had one-tenth of the resistance of the larger inductors. In addition, the MOSFET on-resistance is the same as or lower than the MOSFET resistance on the first four boards.

The difference is that the output voltage, and hence the output power, are considerably higher on boards 5 and 6 in this test. However, board 4, which is very

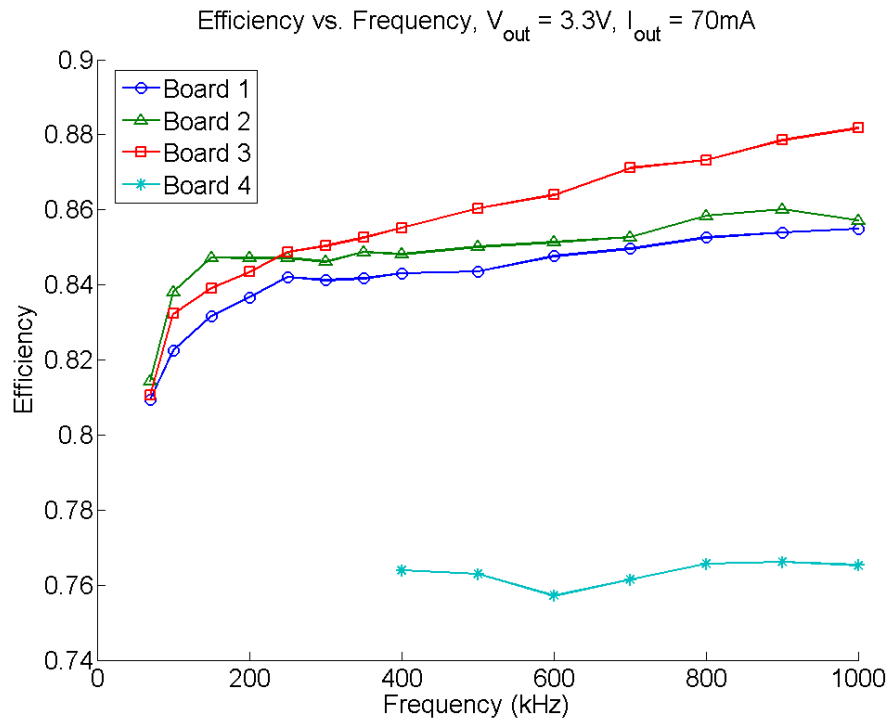


Figure 34: Constant power (230mW), variable frequency output. Boards 1-3 used 10uH coils, while board 4 used a 1uH coil.

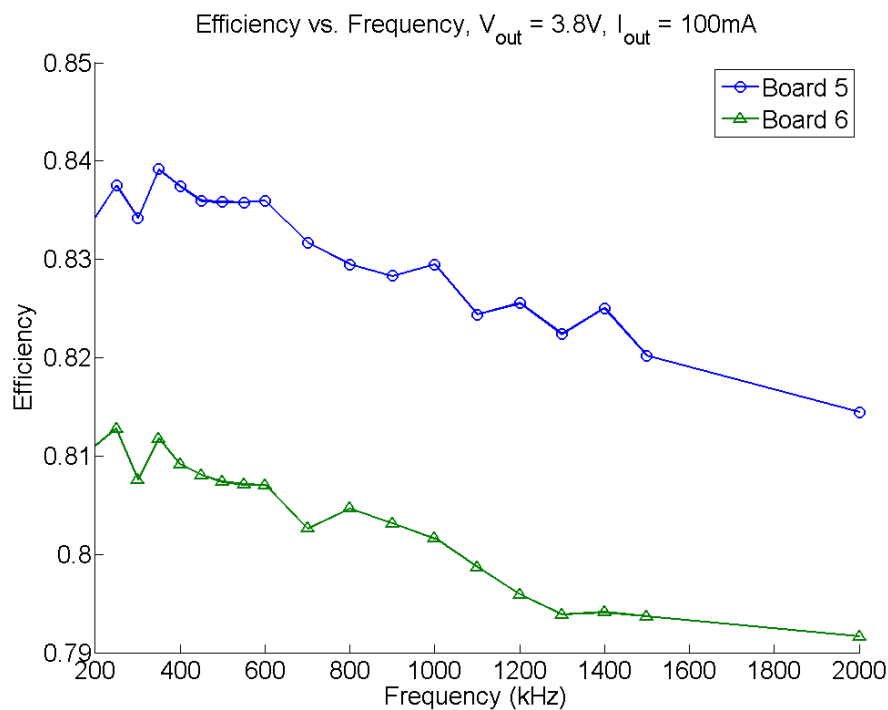


Figure 35: Constant power (380 mW), variable frequency and duty cycle. Boards 5 and 6 used 1uH coils.

similar to boards 5 and 6, shows the same decreased efficiency at the lower, 230 mW output level. As a result, this likely reflects a real, inherent difference in the efficiency across a range of frequencies.

One possible explanation is that since the power output is essentially fixed, each cycle at a certain frequency should transfer the same amount of energy to the load. Since the inductors are smaller in boards 4, 5, and 6, they must conduct for a longer portion of the cycle (or continuously) in order to transfer sufficient energy to the load. This of course raises the resistive losses in the inductor and the power switches. This effect is exacerbated by the fact that the current does not rise linearly with time but instead follows an exponential curve. This effect may be the reason for the decreased efficiency with the smaller inductors.

In general, the boards which used the PN diodes had at least 3% lower efficiency across the range of frequencies. This was expected due to the higher resistance and turn-on voltage of the PN diodes as well as the reverse recovery charge of the PN diodes from which the Schottkys do not suffer.

The efficiency of board 3 is initially lower than boards 1 and 2 at low frequency; however, above 200kHz, the efficiency is much better. Board 3 combined both the Schottky diode with a MOSFET (the DMG1013) that had one-fifth the input capacitance and one-tenth the reverse recovery capacitance. Clearly, it is better to sacrifice a little low-frequency performance for much better overall performance especially at high frequencies.

A second set of tests was run with a fixed output voltage, but varying current. These tests were run at several frequencies to add another dimension to the efficiency

measurements. Again, the output was regulated to 3.3 V by manually varying the duty cycle within the available range of 20% to 80%.

As can be seen in Figure 36, the efficiency of board 1 is a function of both frequency and output current. At 250kHz, the efficiency drops slightly with current, especially above 50mA. At higher frequencies, efficiency actually rises with output current. This rise in efficiency correlates with the power supply moving from the DCM region of operation into the CCM region. This was expected based on the writings of other authors [13], [8].

Finally, note that although testing was performed at load currents down to 1 mA, these results are not graphed. The reason for this is that because of the limited duty cycle range of the function generator, the duty cycle could not be made small enough to output

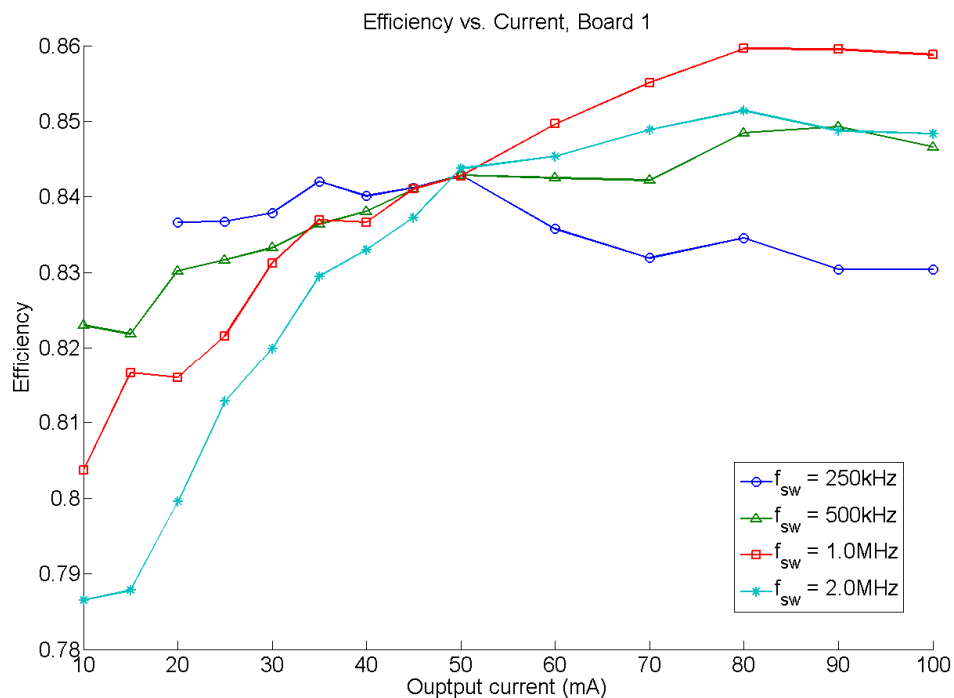


Figure 36: Board 1 efficiency vs. current for four frequencies in the controlled region. Nominal output voltage was 3.3 Volts.

the standard 3.3V at these load currents. Instead, the output voltage ranged from 4.79V at 1mA and 250kHz down to the controlled voltage of 3.3 volts, typically at or around 10mA load.

The data set for board 5 is limited, as seen in Figure 37, since the converter could not be set to 3.8 Volts due to the duty cycle limitation. The limiting factor is that smaller inductors tend to transfer more energy for a given pulse time due to their much higher peak current. An inductor with half the inductance of another inductor will have twice the peak current and four times the energy stored at peak current since energy is proportional to the square of the current. In addition, the energy passed by the inductor is twice what is passed in the larger inductor due to the higher current into the filter. All of this extra energy transfer comes at the price of increased ripple, as shown in Figure 38.

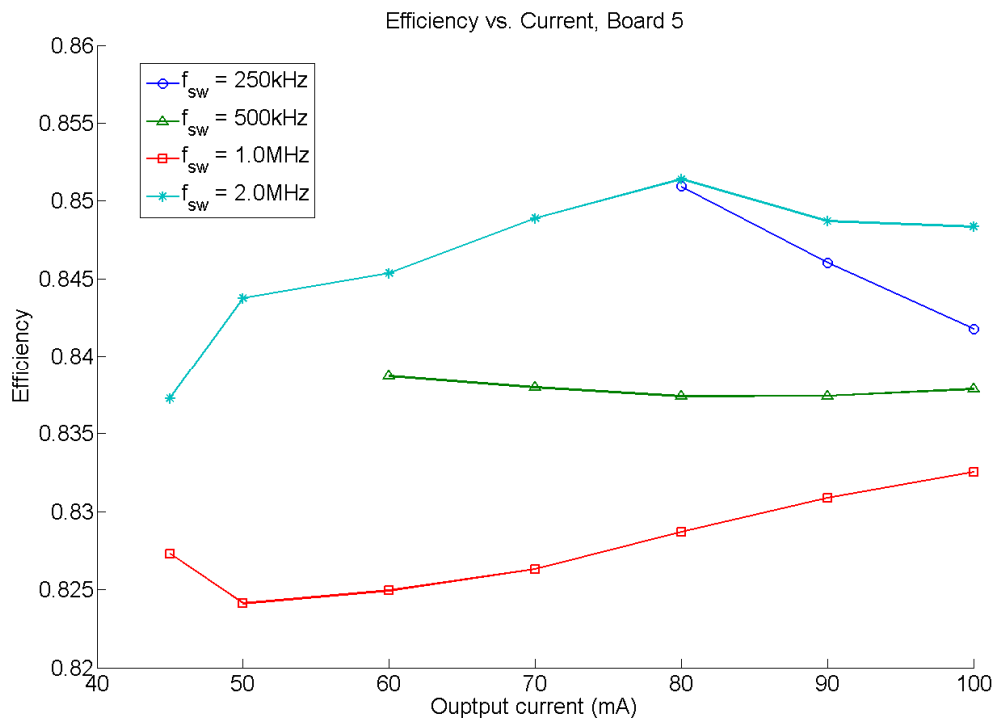


Figure 37: Efficiency vs. current for board 5 in the controlled region. Nominal output voltage was 3.8 Volts.

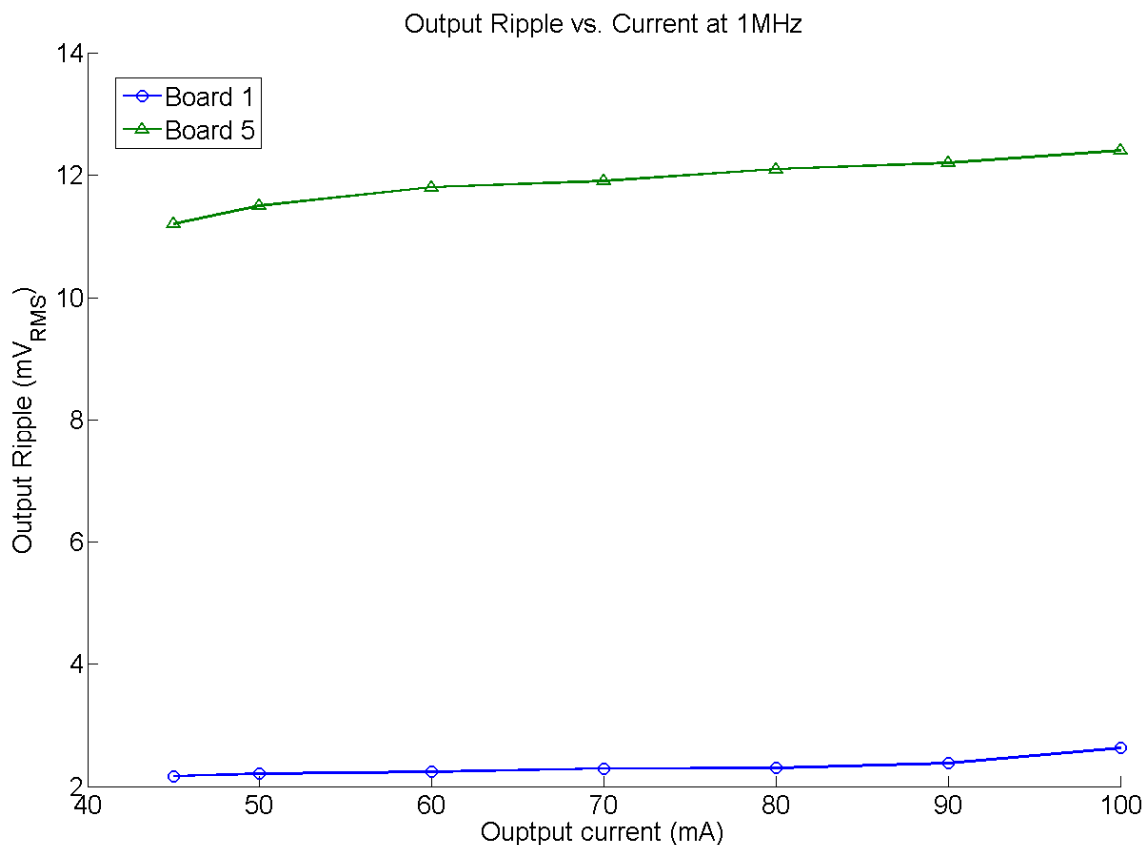


Figure 38: Output ripple at several output currents.

The graph of the output ripples of boards 1 and 5 in Figure 38 shows a slight upward trend with output current. This is expected since in CCM mode, the inductor ripple is only weakly related to the output current by the duty cycle.

Although a converter in PFM mode was not built, its operation can be inferred to some degree based on measurements of the converters at different frequencies and duty cycles and hence, pulse lengths. Figure 39 shows the pulse energy for board 1 at a variety of pulse lengths. Note that only DCM pulses are used in the figure since CCM pulses contain pulse energy as well as DC energy. The determination of which samples were DCM was based on visual inspection of the common node voltage.

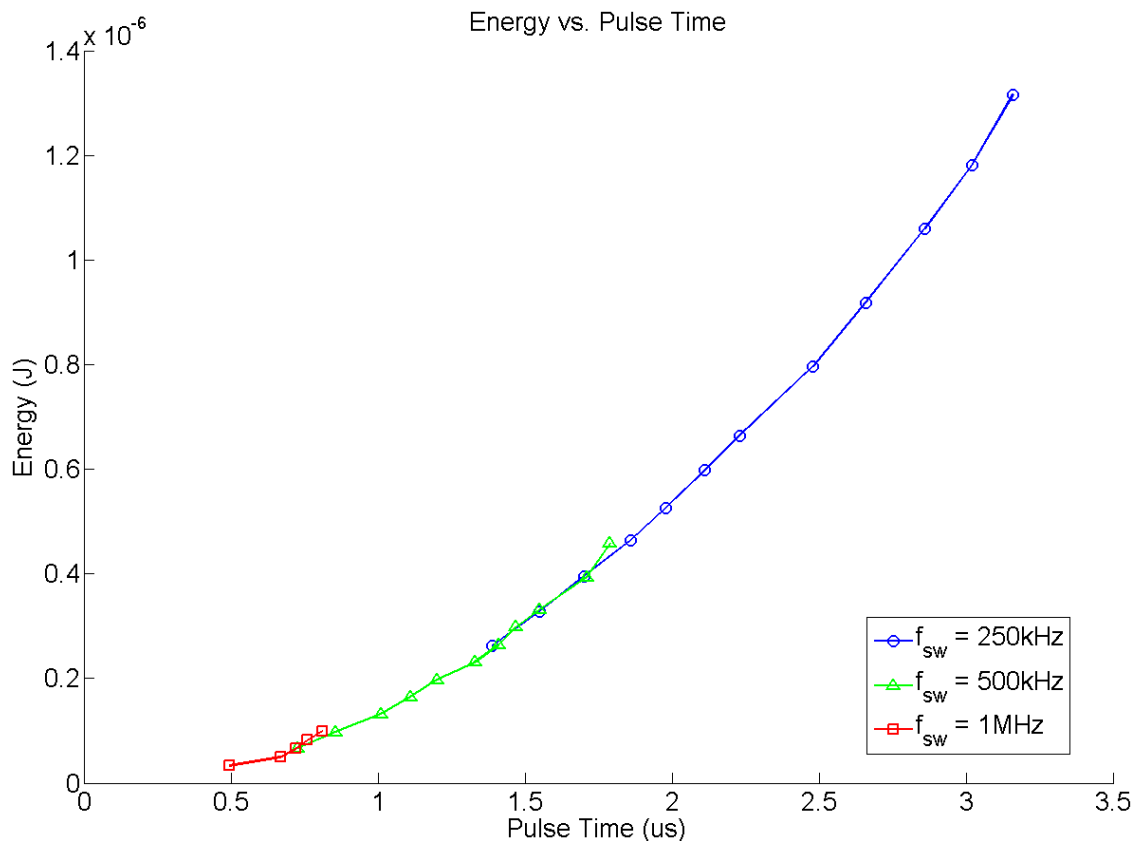


Figure 39: Board 1 energy transferred per pulse as a function of total pulse time (t_1 plus t_2).

As can be seen in Figure 39, the energy transferred follows a parabolic or exponential curve with increasing energy over time. Because the converter data above are collected from the DCM region, there is little or no change in the energy due to the load current. In other words, the rightmost ends of the data sets follow the general upward curve.

In Figure 40, the pulse energy is correlated with the high-side conduction time. The curve is essentially the same but slightly scaled to the left. The diode conduction time, or low-side MOSFET if synchronous rectification is used, should scale with the high-side conduction time by a factor roughly equal to $V_{out}/(V_{in} + V_{out})$.

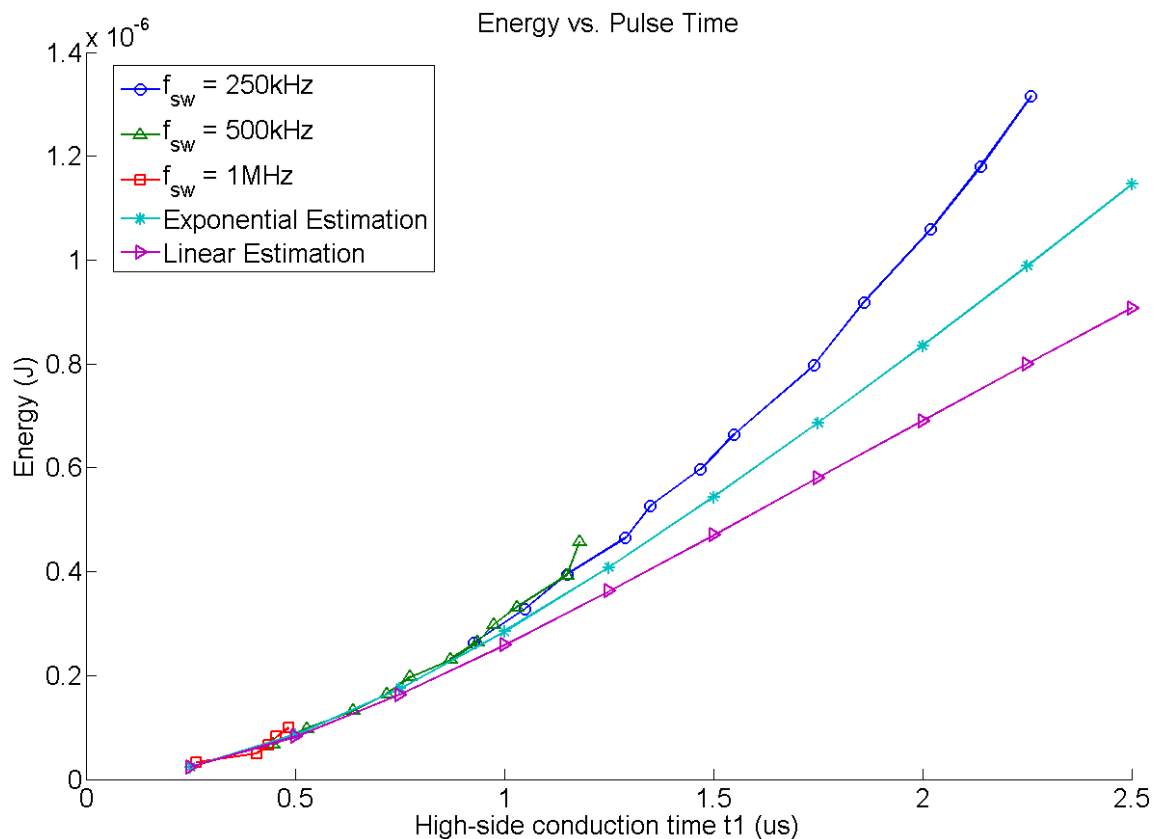


Figure 40: Pulse energy as a function of high-side conduction time (t_1). The exponential and linear estimations curve equal the charge transferred, Q , multiplied by the output voltage plus the stored energy.

The exponential estimation curve is based on a current simulation with measured values for switch and inductor resistance ($R_s = 8.21 \Omega$) and the inductance, $10 \mu\text{H}$. Once the current was estimated from these parameters, the charge was determined by numerical integration and multiplied by the output voltage. This gives the power output to the load up to time t_1 . The energy stored in the inductor is simply $\frac{1}{2}Li(t_1)^2$. The sum of these two quantities is the pulse energy delivered to the load per cycle. The fact that the exponential estimation underestimated the energy means that the current rose faster than expected and was larger than expected.

The linear estimation curve is based on a linear estimation of the inductor current (which is known to be exponential). As expected, this estimation underestimates the charge transfer, and therefore the energy transfer, by an amount that grows with the conduction time. At $2.5\mu\text{s}$, approximately two time constants have passed. According to Table 1, the error at this point should be 23.83%, but it is actually 39.24% below the measured value. The increase in the error is due in large part to the underestimation of the current at t_1 .

The data from board 1 were separated into DCM and CCM sets and plotted in Figure 41. An estimate for CCM efficiency was added to the plot based on the static efficiency equation, which overestimated the efficiency. If the remaining inefficiency is assumed to be a fixed dynamic loss, the second estimate shown is the result.

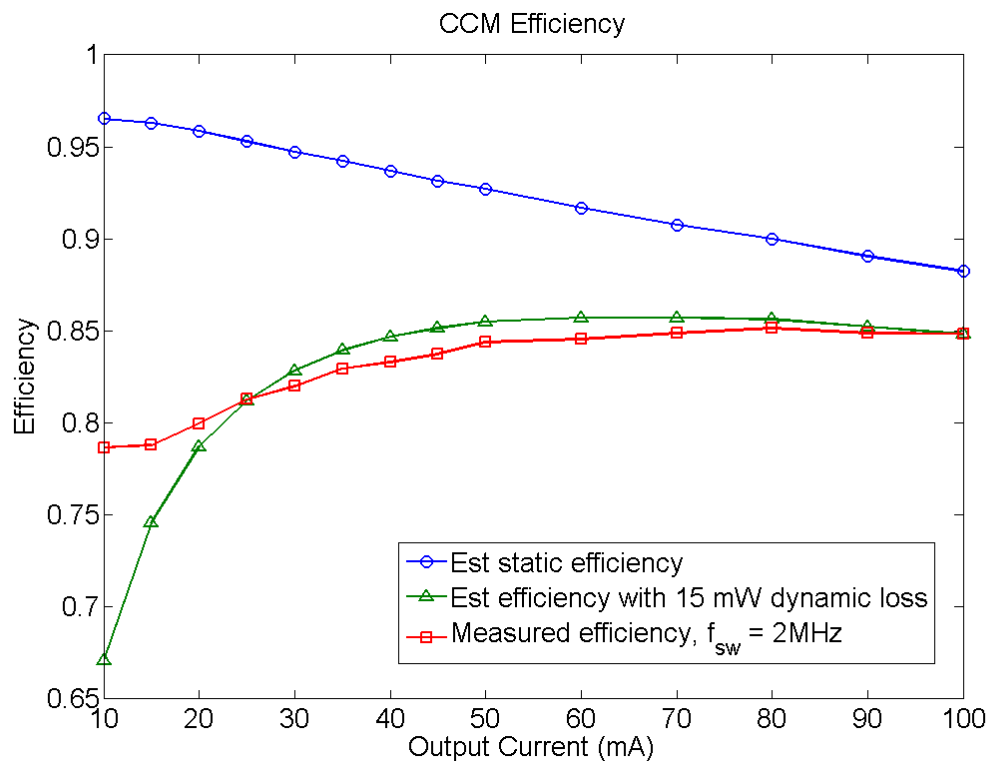


Figure 41: CCM efficiency at multiple loads with estimates.

The converter was observed to be coming out of CCM mode, as evidenced by diode conduction dropping rapidly, starting at 25mA. This may explain the inflection in the efficiency curve at this load in Figure 41 since dynamic losses should drop significantly under DCM mode.

Figure 42 shows the turn-on voltages and conduction resistances of the PN and Schottky diodes. This figure clearly shows the differences between the two types of diodes. As expected, the Schottky diode had a much lower turn-on voltage, but it also turned out to have a lower conduction resistance. These two factors, together with the Schottky diode's lack of a reverse recovery mechanism, affected the efficiency of boards 4 and 6, as discussed above.

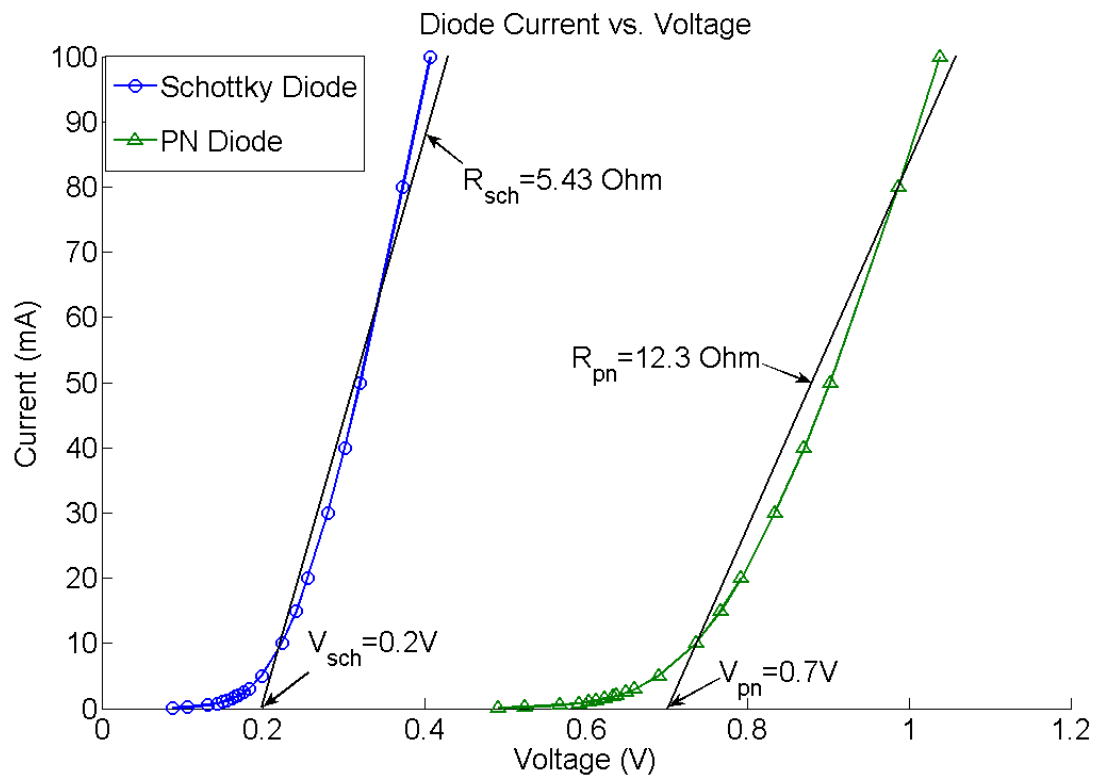


Figure 42: Linear fits for PN and Schottky diodes with turn-on voltages and conduction resistances noted.

4.3 Extrapolation from Board Results

As was seen in Chapter 2, the best choice for low-load current is pulse frequency modulation. This type of control was not possible with the test setup used with the test boards. However, there is an equivalence between PFM and PWM-DCM in that both control methods effectively adjust the duty cycle – the former through the use of variable interpulse time and fixed pulse duration, the latter by variable pulse duration and fixed interpulse time.

The difference in efficiency is evident when the dynamic power is comparable to the resistive power which typically occurs at low output power. In this case, a converter operates best when it operates least because dynamic power is greater than resistive losses. A PFM controller does this by idling when the output voltage is within tolerance, thereby reducing the effective switching frequency. A PWM controller cannot do this because the switches always operate at a fixed frequency regardless of how light the load is or, conversely, how short the pulse is.

The pulse energies of various length pulses were shown in Figure 40 and Figure 39. From these pulse lengths, the approximate frequency of operation can be inferred for several pulse lengths. For a 10mW load and pulse lengths of 0.484us, 0.976us, 1.47us, and 2.02us, the operating frequencies will be 101kHz, 33.6kHz, 16.7kHz, and 9.44kHz, respectively.

Several inferences can be made about the efficiency under PFM operation. First, the amount of energy lost during the active portion of the cycle (when the switches are on and conducting) will be equal for equal pulse lengths since the beginning and ending conditions are the same. In fact, because of the steady-state assumptions, there is no

operational difference between a PFM converter and a PWM-DCM converter with the same frequency and load. The steady-state assumptions guarantee that the both converters under feedback control will adjust their frequency and conduction time respectively to match each other and achieve the correct output energy.

Therefore, if the energy loss in the controllers can be ignored, either because they are small or similar to each other, the efficiency should be nearly identical. Of course, energy loss in the controller cannot be ignored, especially at low output currents where the converter power becomes a significant portion of the total power drawn from the input supply. Fortunately, the energy loss in a PFM controller should be lower than the energy loss in a PWM controller.

In general, the duty cycle of a PFM converter and a PWM-DCM converter will be similar, but not identical. Differences arise due to the fact that the input current has a very small DC component compared to the ripple current, which also happens to be exponential.

4.4 Integrating the Components

The main components that need to be integrated are the power switches, inductor, and output capacitor. The input capacitor has been assumed up to this point, and while it could be integrated, most designs should use an external or on-package bulk capacitance with at most a smaller, high-frequency bypass capacitor integrated on the front end. Finally, the controller including both the feedback circuit and the gate drivers will need to be integrated.

4.4.1 Power Switches

There are three main concerns with the power switches: on-resistance, capacitance, and physical size. Resistance and capacitance have a reciprocal relationship with efficiency, especially at the kinds of frequencies that will be encountered. Larger MOSFETs have lower resistance, but higher capacitance and dynamic loss. Smaller MOSFETs have higher resistance, but lower capacitance, resulting in lower dynamic loss and smaller gate driver chains.

The relationship between resistance and capacitance for N-MOS and P-MOS transistors is shown in Figure 43. According to [15] and [24], the best choice for transistor size is one in which the resistive loss equals the capacitive loss at a certain frequency. This characteristic is shown in Figure 43 where total transistor loss is plotted vs. width at a variety of loads and frequencies. Higher current favors wider transistors and higher frequency favors narrower transistors. At the frequencies plotted, the transistors are surprisingly small.

The correct size for the transistor depends on the expected frequency of the converter. This obviously depends on the size of the inductor since the inductor determines the per-cycle energy transfer. According to [9] and [14], the optimum width can be determined by equation 73 where R_0 and E are the resistance and gate energy of a 1 μm wide, minimum length transistor.

$$W = \sqrt{\frac{R_0 i_{rms}^2}{E f_{sw}}}$$

Equation 73

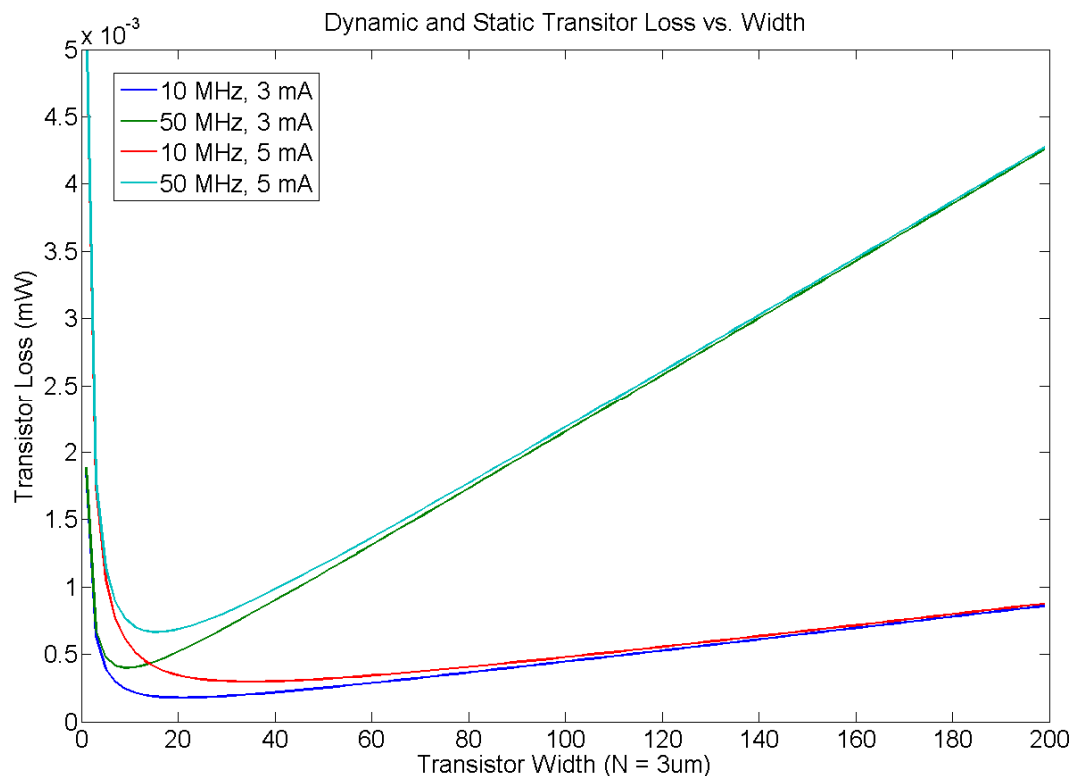


Figure 43: Transistor energy loss as a function of width. As expected, high-frequency and low-current favors narrow transistors.

Since the K' of CMOS processes is directly related to C_{ox} (and the carrier mobility), the on-resistance and gate capacitance scale at approximately the same rate. In other words, smaller processes have greater gate capacitance, but they also have lower on-resistance. Therefore, making transistors on-chip for low-power, high-frequency converters should not be a problem for most CMOS processes.

4.4.2 Inductors

From the research on inductors [22] and other monolithic converters [9], the dual layer spiral seems to be the best choice for high-inductance and low resistance in a small

area. Figure 44 shows the inductance and resistance for inductors of a given number of turns with increasing inner and outer diameters.

The inductance is a function of the physical size and number of turns of the inductors. The resistance is a function of the total length of the inductor divided by the length. In Figure 44, the width-to-diameter ratio was kept constant for all the coils tested.

If size were no object, the best choice for inductor would be as large as possible with the widest tracks possible, and the lowest resistance possible. In Figure 44, larger inductors were clearly better. However, physical size will almost always determine the size of the inductor simply because they are the biggest component to be fabricated. The question then is what balance should be struck between inductance and resistance.

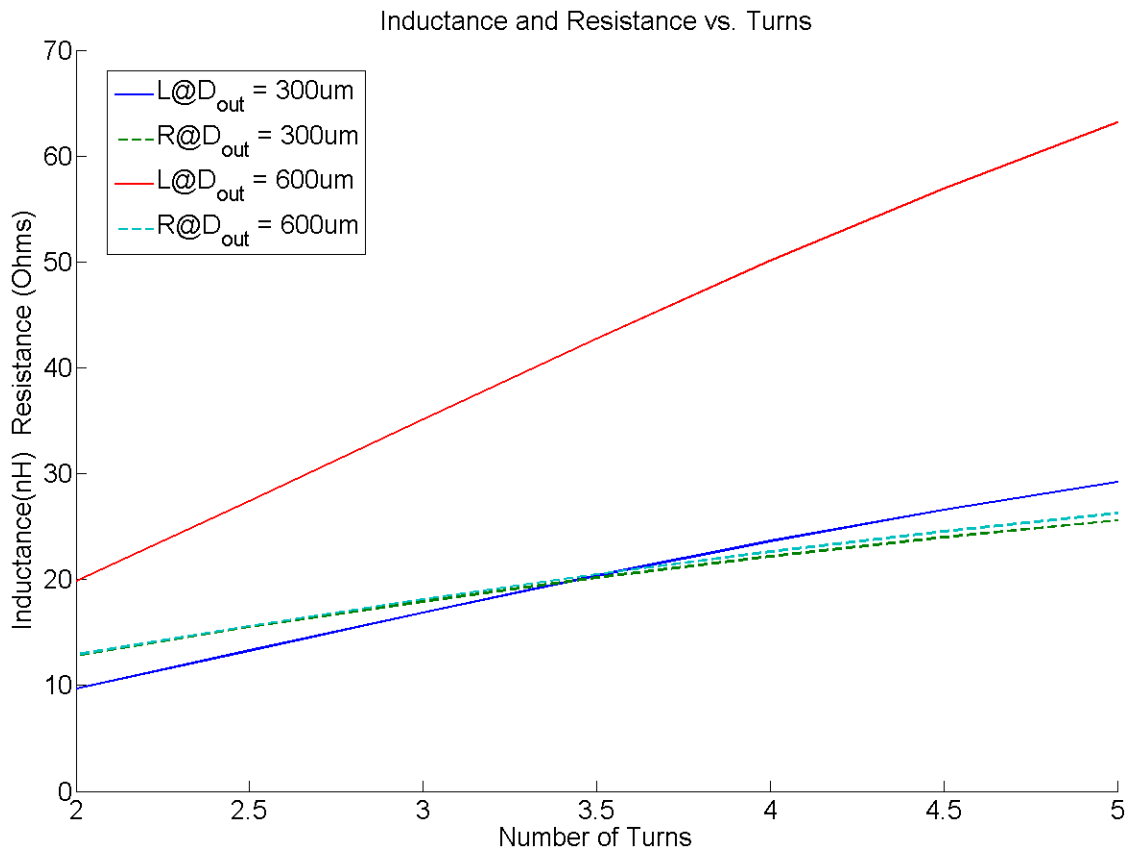


Figure 44: Double-layer spiral inductance and resistance as a function of the number of turns.

Decreasing the resistance will also decrease the inductance. However, resistance has a direct effect on efficiency, whereas inductance affects efficiency by changing the duty cycle. In general, it is better to go with a smaller inductance with smaller resistance.

As has been previously discussed, the inductance of a coil is dependent on the physical size of the inductor. Process parameters only change the quality of the inductor, not its inductance. Therefore, a coil with a certain inductance will be the same size in every fabrication technology. There is therefore a higher cost to be paid in lost active device area for smaller process sizes.

The conduction time of the inductor should be long enough that sufficient energy is transferred to the load that the repetition frequency does not result in excessive switching loss. However, it should also be short enough that the inductor does not start to act like a resistor. A good method for picking the conduction time in PFM mode is the time constant of the coil. At time $t = \tau = L/R$, the inductor is still continuing to store additional energy as flux, but most of the inductor voltage is due to Ohm's law. Further conduction is more and more wasteful as the self-resistance becomes the sole source of voltage drop across the inductor.

4.4.3 Output Capacitor

The output capacitor has two main considerations: it must maintain the output current flow between cycles, and it must suppress the output ripple. Constructing a capacitor of sufficient size will require multiple layers.

Starting with a MOSCAP, a poly-poly2 capacitor can be built on top of and parallel with the MOSCAP. Metal 1 can then be layered on top, followed by metal 2.

Metal 3 should be left as a bus layer, although interleaving the terminals will add a small amount of additional capacitance. Without counting the fringing capacitance obtained by fingering and alternating layers, the 100% fill capacitance with all these layers is 3397 aF/ μm^2 .

Due to device spacing requirements, the multilayer capacitor will not be 100% space efficient. Realistically, the best that can be achieved with the MOSCAP cells is approximately 30-50% depending on the length of the device. Placement of active contacts between polysilicon limits the amount of space that the polysilicon can cover. Upper layers can be spaced at the minimum distance both to maximize space and to capitalize on fringing capacitance. The MOSCAP dominates the achievable capacitance (it is 72% of the theoretical maximum) so the realizable capacitance per area is approximately 40% of the maximum, or 1360 aF/ μm^2 .

Since capacitance is inversely related to the separation distance between parallel plates, smaller processes will have higher capacitance per area than larger processes. In addition, smaller processes often have more metal layers available which can add even more capacitance per area.

4.5 Test Designs

Based on the test results using the test boards and the calculations above, several theoretical designs will be evaluated numerically. Since the inductor determines the conduction time and frequency, these designs start with an inductor design and extrapolate the circuit elements from there. The calculations are run iteratively in Matlab to obtain the final values quoted below.

4.5.1 Double Spiral 1 – 600 μm x 600 μm , two turns

The inductor used is a 600 μm diameter double-layer coil with 2 turns per layer. This coil has 12.88 Ω of series resistance and 19.84 nH of inductance. An initial guess of 10 Ohms was used for the MOSFET on-resistance. Five iterations were run in Matlab. The final pulse time was 0.906 ns with a pulse energy of 72.802 pJ per pulse. For an output voltage and current of 3.3V and 3mA, the switching frequency is 135.98MHz with an effective duty cycle of 18.67%. The RMS value of the inductor current was 12.2mA.

The Matlab program makes use of equation 73 to determine the optimal transistor width. The value that was calculated was a transistor width of 23 times the minimum width of 3 μm , or 69 μm wide. The resulting on-resistance was 8.953 Ω . Efficiency was calculated using a modified form of equation 54 in which the RMS current (which was already calculated) was multiplied by the MOSFET resistance and the inductor resistance to obtain the static power loss. Dynamic loss was estimated by multiplying the switching frequency by the per-cycle gate energy loss. The resulting efficiency was 75.16%, well above the minimum required value of 66%. The actual efficiency will be somewhat lower due to stray capacitances not taken into account, the ESR of the capacitor, and shoot-through current – all of which were not calculated.

Using the ripple voltage equation from Chapter 2, equation 51, and assuming 1% ripple, the output capacitor must be 0.4798 nF. A capacitor this size would take up an additional 0.3528 mm^2 . Not counting the area taken up by the power switches or controller, the converter takes up 0.7128 mm^2 .

4.5.2 Double Spiral 2 - 600 μm x 600 μm , five turns

The second test case uses the largest, widest inductor that was previously calculated. This inductor is a 5 turn inductor with trace width of 60 μm and outer dimension 600 μm . This inductor has inductance 63.16 nH with series resistance 26.28 Ω .

The Matlab calculations were run again with the new values for inductance and resistance. The resulting switching frequency was 125MHz. The MOSFET width was 54 μm for a resistance of 11.24 Ω . Overall efficiency was little changed from the previous test: 75.09%.

The capacitor required is 59.926pF – almost one-tenth the size of the previous test's output capacitor. The capacitor area is then 0.044mm². Therefore, the total area is reduced to 0.404mm² – much smaller than the previous converter example.

4.5.3 External Inductor – MuRata LQF1SHS, EIA 0402, 120nH

This design uses an external inductor with 120nH inductance and 1.3 Ω of resistance. Such an inductor could be mounted directly to the silicon or could be mounted just off-chip in the same mounting package. The results of the Matlab calculations are an efficiency of 75.10%. The switching frequency is a relatively slow 86kHz. The problem with this design is that the required transistor is 11.067mm long for an on-resistance of only 56.1m Ω . This transistor could be built in about the same space as one of the integrated inductors. This design requires a 1 μF output capacitor which cannot be built in silicon.

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

5.1 Conclusions

In this thesis, the theory of buck converters was presented. The concept of the ideal switching supply was introduced and, proceeding from general to specific, the intricate details of switching supplies were covered. Equations were either presented or derived to describe the operation and behavior of buck converters.

An in-depth discussion followed in which the main components of a switching supply were discussed. Power switches, inductors, and capacitors were introduced in a systematic way, highlighting the important features and distinctions between devices and technologies. Especially important was the discussion of integrated inductors.

Several test boards were built using a variety of parts. These boards reinforced the understanding of switching supplies while also adding new insight into the operation of switching supplies both in continuous and discontinuous conduction modes. These boards were found to work at efficiencies approaching 90% despite their use of diode rectification. The data taken from these test boards were generalized to pulse frequency modulation circuits.

Finally, the integrated devices were examined in more depth with the goal of determining if they were suitable for building monolithic switching supplies. A design

approach for fabricating both the power switches and the capacitors was presented. The inductors, however, proved to be a disappointment. Despite using a great deal of space, the inductors just could not be coaxed into operating efficiently.

In the test cases, the series resistance of the coils dictated the operating frequency of the converter and the size of the transistors and capacitors. The resistance of the coils meant that the converters had to operate at very high frequencies. Nevertheless, by correctly sizing the transistors, the efficiency was kept above 75%.

Based on the theory examined, the equations derived, and the test cases examined, it should be feasible to build fully integrated low-power buck converters in most CMOS processes provided that sufficient area is available for the inductor. The power switches and capacitor are both smaller than the integrated inductor.

In addition to the possibility of adding integrated inductors onto integrated circuits, it should also be possible to add a microscale inductor and capacitor to an IC. External transistors are not advantageous in this case since suitable transistors can be build on-chip. The external solution did not prove to be any more efficient, although it did allow more design freedom and lower switching frequencies.

5.2 Future Work

The lack of a proper signal source hampered the testing that was performed using the test boards. Any future testing using discrete, surface mount parts would benefit greatly from an on-board, variable pulse length trigger circuit. In addition, the frequencies involved were sufficiently high that transmission line ringing was observed throughout testing.

A real silicon integrated circuit would have been helpful in testing the real world efficiency of the test designs. It is possible that the physical devices will perform better or worse than predicted. The design equations used for the inductors were based on RF-coils operating at much higher frequencies than a converter would use.

A comprehensive way of determining inductor size would be useful for future designers. Such a methodology was not found in literature except for CCM converters under heavy load. Instead, the current approach seems to be to use as much area as is available. A better way should be found in the future.

Because the purpose of the thesis was to determine if integrating inductive supplies was even possible, no controller circuits were discussed except at very high levels. Future designs will need to address this shortcoming to determine the real efficiency of integrated switching supplies.

APPENDIX A

THE MATLAB CODE

```
V_in = 5;
V_out = 3.3;
I_out = 0.003;
R_L = 1.3;
L = 120e-9;

%initial guess for R_on
R_on0 = 1;

tau = L/(R_L+R_on0);
P_out = V_out*I_out;
E0_fets = 4.2678e-13; %Joules
R0 = 207;
I_pk = (V_in-V_out)/(R_L + R_on0)*(1-exp(-1));
W_pulse = 1/3*tau*V_out*I_pk + 0.5*L*I_pk.^2;
f_sw = P_out/W_pulse;
D = tau*5/V_out*f_sw;
I_rms = I_pk*(D/3).^0.5;
W = (R0*I_rms.^2/f_sw/E0_fets).^0.5;
R_on = R0/W;

%Second iteration
I_pk = (V_in-V_out)/(R_L + R_on)*(1-exp(-1));
tau = L/(R_L+R_on);
W_pulse = 1/3*tau*V_out*I_pk + 0.5*L*I_pk.^2;
f_sw = P_out/W_pulse;
D = tau*V_in/V_out*f_sw;
I_rms = I_pk*(D/3).^0.5;
W = (R0*I_rms.^2/f_sw/E0_fets).^0.5;
R_on = R0/W;

%Third iteration
I_pk = (V_in-V_out)/(R_L + R_on)*(1-exp(-1));
tau = L/(R_L+R_on);
W_pulse = 1/3*tau*V_out*I_pk + 0.5*L*I_pk.^2;
f_sw = P_out/W_pulse;
D = tau*V_in/V_out*f_sw;
I_rms = I_pk*(D/3).^0.5;
W = (R0*I_rms.^2/f_sw/E0_fets).^0.5;
R_on = R0/W;
```

```

%Fourth iteration
I_pk = (V_in-V_out)/(R_L + R_on)*(1-exp(-1));
tau = L/(R_L+R_on);
W_pulse = 1/3*tau*V_out*I_pk + 0.5*L*I_pk.^2;
f_sw = P_out/W_pulse;
D = tau*V_in/V_out*f_sw;
I_rms = I_pk*(D/3).^0.5;
W = (R0*I_rms.^2/f_sw/E0_fets).^0.5;
R_on = R0/W;

%Fifth iteration
I_pk = (V_in-V_out)/(R_L + R_on)*(1-exp(-1))
tau = L/(R_L+R_on)
W_pulse = 1/3*tau*V_out*I_pk + 0.5*L*I_pk.^2
f_sw = P_out/W_pulse
D = tau*V_in/V_out*f_sw
I_rms = I_pk*(D/3).^0.5
W = (R0*I_rms.^2/f_sw/E0_fets).^0.5
R_on = R0/W

%Efficiency Calculation
efficiency = V_out*I_out/(V_out*I_out + W*E0_fets +
I_rms.^2*(R_L+R_on))
C = 0.5*0.003*(tau + L/(-1*V_out-
0.5*I_pk*(R_on+R_L))+tau*I_pk/I_out)/0.033

```

BIBLIOGRAPHY

- [1] R. R. Harrison, P. T. Watkins, R. J. Kier, R. O. Lovejoy, D. J. Black, B. Greger, and F. Solzbacher, "A low-power integrated circuit for a wireless 100-electrode neural recording system," *Solid-State Circuits, IEEE Journal of* , vol. 42, no. 1, pp. 123-133, Jan. 2007
- [2] R.R. Harrison, "Designing efficient inductive power links for implantable devices," in *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on* , pp. 2080-2083, 27-30 May 2007
- [3] C. Shi, B. C. Walker, E. Zeisel, B. Hu, and G. H. McAllister, "A highly integrated power management IC for advanced mobile applications," *Solid-State Circuits, IEEE Journal of* , vol. 42, no. 8, pp. 1723-1731, Aug. 2007
- [4] P. D. Wolf, "Thermal considerations for the design of an implanted cortical brain-machine interface." *Indwelling Neural Implants: Strategies for Contending with the In Vivo Environment*, [Online] W. M. Reichert, Ed. Boca Raton, FL: CRC Press, 2008. Available:
<http://www.ncbi.nlm.nih.gov/books/NBK1977/>
- [5] R. W. Erickson and D. Maksimovic, *Fundamentals of Power Electronics*, 2nd ed. Boulder, CO: Springer Science+Business Media, Inc, 2001.
- [6] Maxim Application Notes (2001, Nov). "AN 2031: DC-DC Converter Tutorial." Maxim-IC, CA [Online]. Available:
<http://pdfserv.maxim-ic.com/en/an/AN2031.pdf>.
- [7] L. Geng, J. Li, W. Xu, and Q. Li, "A CMOS hybrid control mode DC-DC buck converter for embedded system," in *Solid-State and Integrated Circuits Technology, 2004. Proceedings. 7th International Conference on* , vol. 2, pp. 1472- 1475, 18-21 Oct. 2004
- [8] G. Li, L. Qinghua, and S. Zhibiao, "A new design strategy for the monolithic buck converters," *Power Electronics and Motion Control Conference, 2004. IPEMC 2004. The 4th International* , vol. 1, pp. 192-196, 14-16 Aug. 2004

- [9] Q. Li, L. Geng, and Z. Shao, "A monolithic low-power buck converter integrated in CMOS process," *Microwave Conference Proceedings, 2005. APMC 2005. Asia-Pacific Conference Proceedings*, vol. 2, pp. 4-7, Dec. 2005
- [10] J. W. Nilsson and S. A. Riedel, *Electric Circuits*, 6th ed. Upper Saddle River, NJ: Prentice-Hall, Inc, 2000.
- [11] D. M. Pozar, *Microwave Engineering*, 3rd ed. Hoboken, NJ: John Wiley & Sons, Inc, 2005.
- [12] N. H. E. Weste and D. Harris, *CMOS VLSI Design: A Circuits and Systems Perspective*, 3rd ed. Boston, MA: Pearson Education, Inc, 2005.
- [13] J. Xiao, A. V. Peterchev, J. Zhang, and S. R. Sanders, "A 4- μ a quiescent-current dual-mode digitally controlled buck converter IC for cellular phone applications," *Solid-State Circuits, IEEE Journal of*, vol. 39, no. 12, pp. 2342- 2348, Dec. 2004
- [14] J. Sun, D. Giuliano, S. Devarajan, J.-Q. Lu, T. P. Chow, and R. J. Gutmann, "Fully monolithic cellular buck converter design for 3-D power delivery," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 17, no. 3, pp. 447-451, March 2009
- [15] S. Musunuri and P. L. Chapman, "Improvement of light-load efficiency using width-switching scheme for CMOS transistors," *Power Electronics Letters, IEEE*, vol. 3, no. 3, pp. 105- 110, Sept. 2005
- [16] S. Musunuri, P. L. Chapman, J. Zou, and C. Liu, "Inductor design for monolithic DC-DC converters," in *Power Electronics Specialist Conference, 2003. PESC '03. 2003 IEEE 34th Annual*, vol. 1, pp. 227- 232, 15-19 June 2003
- [17] R. Strandberg, and J. Yuan, "Analysis and implementation of a semi-integrated Buck converter with static feedback control," in *Circuits and Systems, 2000. Proceedings of the 43rd IEEE Midwest Symposium on*, vol. 2, pp. 934-937, 2000
- [18] D. Ma, W.-H. Ki, and C.-Y. Tsui, "An integrated one-cycle control buck converter with adaptive output and dual loops for output error correction," *Solid-State Circuits, IEEE Journal of*, vol. 39, no. 1, pp. 140- 149, Jan. 2004
- [19] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, 2nd ed. New York, NY: Cambridge University Press, 2004.
- [20] C. P. Yue, C. Ryu, J. Lau, T. H. Lee, and S. S. Wong, "A physical model for planar spiral inductors on silicon," in *Electron Devices Meeting, 1996. IEDM '96., International*, pp. 155-158, 8-11 Dec. 1996

- [21] C. P. Yue and S. S. Wong, "On-chip spiral inductors with patterned ground shields for Si-based RF ICs," *Solid-State Circuits, IEEE Journal of*, vol. 33, no. 5, pp. 743-752, May 1998
- [22] A. Zolfaghari, A. Chan, and B. Razavi, "Stacked inductors and transformers in CMOS technology," *Solid-State Circuits, IEEE Journal of*, vol. 36, no. 4, pp. 620-628, Apr 2001
- [23] C.-C. Tang, C.-H. Wu, and S.-I. Liu, "Miniature 3-D inductors in standard CMOS process," *Solid-State Circuits, IEEE Journal of*, vol. 37, no. 4, pp. 471-480, Apr 2002
- [24] A. C. Jan, M. Rabaey, and B. Nikolic, *Digital Integrated Circuits*. Upper Saddle River, NJ: Prentice Hall, 2003.