Performance-Driven Energy-Efficient VLSI

by

Wei-Hsiang Ma

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering)
in The University of Michigan
2011

Doctoral Committee:

Professor Marios C. Papaefthymiou, Chair
Professor Dennis M. Sylvester
Associate Professor Jerome P. Lynch
Assistant Professor Zhengya Zhang

To my family and friends for their love and support

ACKNOWLEDGMENTS

First and foremost I offer my sincerest gratitude to my advisor, Prof. Marios Papaefthymiou, who has supported me throughout my thesis with his patience and knowledge while allowing me the room to work in my own way. I would also like to thank all the other committee members, Prof. Dennis Sylvester, Prof. Zhengya Zhang, and Prof. Jerome Lynch, providing valuable feedback and support.

I want to thank Jerry Kao, who as a good friend and group mate of mine as always willing to help and give his best suggestions. It would have been a lonely office without him. Many thanks to Visvesh, Carlos, Yu-Shiang, Juang-Ying, and Tai-Chuan for sharing their time and expertise. Thanks to the program coordinator Beth Stalnaker and the ACAL/DCO staffs: Denise, Bert, Lauri, Steve, Ed, and Joel for their immense help.

Finally, I would like to thank my parents. They are always supporting me and encouraging me with their best wishes.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

LIST OF ABBREVIATIONS

ADC Analog-to-Digital Converter

ADL Adiabatic Dynamic Logic

BIST Built-in Self-Test

CMOS Complementary Metal-Oxide Semiconductor

dB Decibel

DESL Dynamic Evaluation Static Latch

DNL Differential Non-Linearity

DSP Digital Signal Processor

ENOB Effective Number of Bits

ERBW Effective Resolution Bandwidth

FIR Finite Impulse Response

FOM Figure of Merit

I/O Input/Output

INL Integral Non-Linearity

LCC Leadless Chip Carrier

LSB Least Significant Bit

MOS Metal-Oxide Semiconductor

nTERF NMOS Energy Recovery Flip-Flop

PAL Pass-transistor Adiabatic Logic

PC Power Clock

PCB Printed Circuit Board

PFAL Positive Feedback Adiabatic Logic

pTERF PMOS Energy Recovery Flip-Flop

QFN Quad Flat No leads

QSERL Quasi-Static Energy Recovery Logic

RCK Resonant Clock

RERL Reversible Energy Recovery Logic

SAFF Sense-Amplifier Flip-Flop

SBL Subthreshold Boost Logic

SCAL Source-Coupled Adiabatic Logic

SCRL Split-level Charge Recovery Logic

SMA SubMiniature version A

SMD Surface Mount Devices

SNDR Signal to Noise and Distortion Ratio

SNFR Signal to Noise Floor Ratio

THA Track-and-Hold Amplifier

TSEL True Single-Phase Energy Recovery Logic

VLSI Vary Large Scale Integrated

ABSTRACT

Performance-Driven Energy-Efficient VLSI

by

Wei-Hsiang Ma

Chair: Marios C. Papaefthymiou

Today, there are two prevalent platforms in VLSI systems: high-performance and ultra-low power. High-speed designs, usually operating at GHz level, provide the required computation abilities to systems but also consume a large amount of power; microprocessors and signal processing units are examples of this type of designs. For ultra-low power designs, voltage scaling methods are usually used to reduce power consumption and extend battery life. However, circuit delay in ultra-low power designs increases exponentially, as voltage is scaled below Vth, and subthreshold leakage energy also increases in a near-exponential fashion.

Many methods have been proposed to address key design challenges on these two platforms, energy consumption in high-performance designs, and performance/reliability in ultra-low power designs. In this thesis, charge recovery design is explored as a solution targeting both platforms to achieve increased energy efficiency over conventional CMOS designs without compromising performance or reliability.

To improve performance while still achieving high energy efficiency for ultra-low power designs, we propose Subthreshold Boost Logic (SBL), a new circuit family that relies on charge recovery design techniques to achieve order-of-magnitude im-

provements in operating frequencies, and achieve high energy efficiency compared to conventional subthreshold designs. To demonstrate the performance and energy efficiency of SBL, we present a 14-tap 8-bit finite-impulse response (FIR) filter test-chip fabricated in a 0.13μm process. With a single 0.27V supply, the test-chip achieves its most energy efficient operating point at 20MHz, consuming 15.57pJ per cycle with a recovery rate of 89% and a figure of merit (FoM) equal to 17.37 nW/Tap/MHz/InBit/CoeffBit. In comparison with a static CMOS-based implementation derived by synthesis of the same FIR architecture and automatic place-and-route, the SBL-based FIR consumes 40% to 50% less energy per cycle in the 17MHz - 187MHz range.

To reduce energy consumption at multi-GHz level frequencies, we explore the application of resonant-clocking to the design of a 5-bit non-interleaved resonant-clock flash ADC with a sampling rate of 7GS/s. The ADC has been designed in a 65nm bulk CMOS process. An integrated 0.77nH inductor is used to resonate the entire clock distribution network to achieve energy-efficient operation. Operating at 5.5GHz, the ADC consumes 28mW, yielding 396fJ per conversion step. The clock network accounts for 10.7% of total power and consumes 54% less energy over CV^2. Operating at its maximum sampling frequency of 7.0GS/s with a 0.1V increase to each supply, the ADC dissipates 45mW. At this frequency, 11.1% of the total energy consumption is clock-related, and the ADC yields a FoM of 683fJ per conversion step. By comparison, in a typical flash ADC design, 30% of total power is clock-related.

CHAPTER 1

Introduction

Energy consumption has become a major design constraint in today's VLSI designs. For the last several decades, Moore's Law [3] has been the main driving force to reduce the size and energy consumption of silicon devices. However, this scaling does not reduce power consumption per unit area. As more devices can fit into a given area, and the heat generated increases. Heat removal at the package level limits further integration. Moreover, starting with 65nm, supply voltage no longer scales with device sizes, remaining essentially constant over past several years and foreseeable future. To make matters worse, leakage current increases due to the smaller device sizes. Therefore, one of the largest issues facing designers nowadays is energy and power dissipation. Their main challenge is to achieve energy efficient computing, extracting the maximum possible performance under a given power constraint.

Voltage scaling is one of the most effective methods for reducing energy consumption in digital circuits [4, 5, 6]. The energy consumption decreases quadratically when the supply voltage V_{DD} decreases, providing energy efficient operation. However, this energy-efficient operation comes at the expense of performance degradation. When voltage is scaled while remaining well above subthreshold ($V_{DD} >> V_{th}$), performance degradation is in approximately linear relationship with supply. When voltage is scaled deeper in the subthreshold regime ($V_{DD}<V_{th}$), circuit delay increases expo-

nentially with V_{DD}, and becomes more sensitive to process variation. Leakage current also increases in a near-exponential fashion in the subthreshold regime. This rise in leakage energy eventually dominates total power consumption and creates a minimum energy operating point [4]. Further scaling beyond this minimum energy point results in total energy consumption increasing and yields diminished energy efficiency. Moreover, voltage scaling has its limitation on different applications. For example, in applications with constantly high workloads, voltage scaling can only help to a limited extent.

Charge-recovery is an alternative design approach that can reduce energy consumption by gradually charging/discharging capacitance and recycling the charge at the end of each cycle [7, 8, 9, 10]. The energy dissipation of a traditional CMOS circuit that goes through a charge or a discharge cycle is governed by the equation $E_{conv} = \textit{switching activity} \times CV^2$, while the corresponding energy dissipation of a charge-recovery system is governed by $E_{charge-recovery} = (K/T)CV^2$, where T is the duration of the transition and K is a constant proportional to the RC constant of the system. Similar to voltage scaling, charge-recovery exhibits a trade-off relationship between energy consumption and computation delay, but this trade-off relation is linear. A large volume of previous work has been proposed based on this trade-off, improving the energy efficiency at the cost of system performance, focusing on achieving high energy efficiency in relatively low-performance designs. However, this trade-off dose not limit the design scope of charge-recovery techniques. In fact, charge-recovery design techniques can enable different design points compared to conventional static CMOS techniques. In principle, these design points could lead to better energy/performance trade-offs than conventional CMOS design.

This thesis argues that charge-recovery techniques can be used to design VLSI systems that achieve both high energy efficiency and high performance. To support this proposition, two charge-recovery systems operating at different frequency points are

demonstrated. Both designs provide high performance at their corresponding supply-levels and achieve higher energy efficiency than their conventional CMOS designs counterparts. For ultra-low energy consumption with high performance, we present a novel fine-grain charge-recovery circuitry that uses a single subthreshold-level supply. By amplifying the internal subthreshold-level signals with a two-phase power-clock, this charge-recovery circuitry relies on gate overdrive to enable fast operation while improving robustness to the variations. In addition, it can share the same supply with the clock generator, allowing operation with a single DC supply level. An 8×8 bit 14-tap finite impulse response (FIR) filter is used to demonstrate the high energy efficiency of this logic, across the 5MHz-187MHz frequency range with a subthreshold supply ranging from 0.16V to 0.36V. This design achieves a higher Figure of Merit (FoM) than previous implementations of the same architecture in charge-recovery logic and in static CMOS.

For high performance and high energy efficiency, we present a coarse-grain resonant-clock flash ADC structure. In this design, resonant-clocking is used to decrease the energy consumption of the network that distributes the clock signal to the analog and digital circuitry of the flash ADC. A 5-bit non-interleaved flash ADC that achieves a sampling frequency range of 4.5GS/s-7GS/s is used to demonstrate the high energy efficiency of resonant-clocking techniques. An integrated inductor is used to resonate the capacitance of the entire clock distribution network at a target operating frequency range. This work achieves a lower FoM than all previously published ADCs operating above 2.2GHz.

1.1 Principles of Charge-Recovery Design Techniques

This section describes the principles of charge-recovery techniques as applied to digital circuit designs. In digital circuits, a MOS transistor is usually used as a switch with one node connected to supply (V_{DD}) / ground (V_{SS}) and the other node

connected to a capacitive load C. By observing the voltage level of the load, we can determine the logic state of a gate. By turning the transistor on, we can charge (discharge) the voltage level of the load and thus change the logic state of a gate.

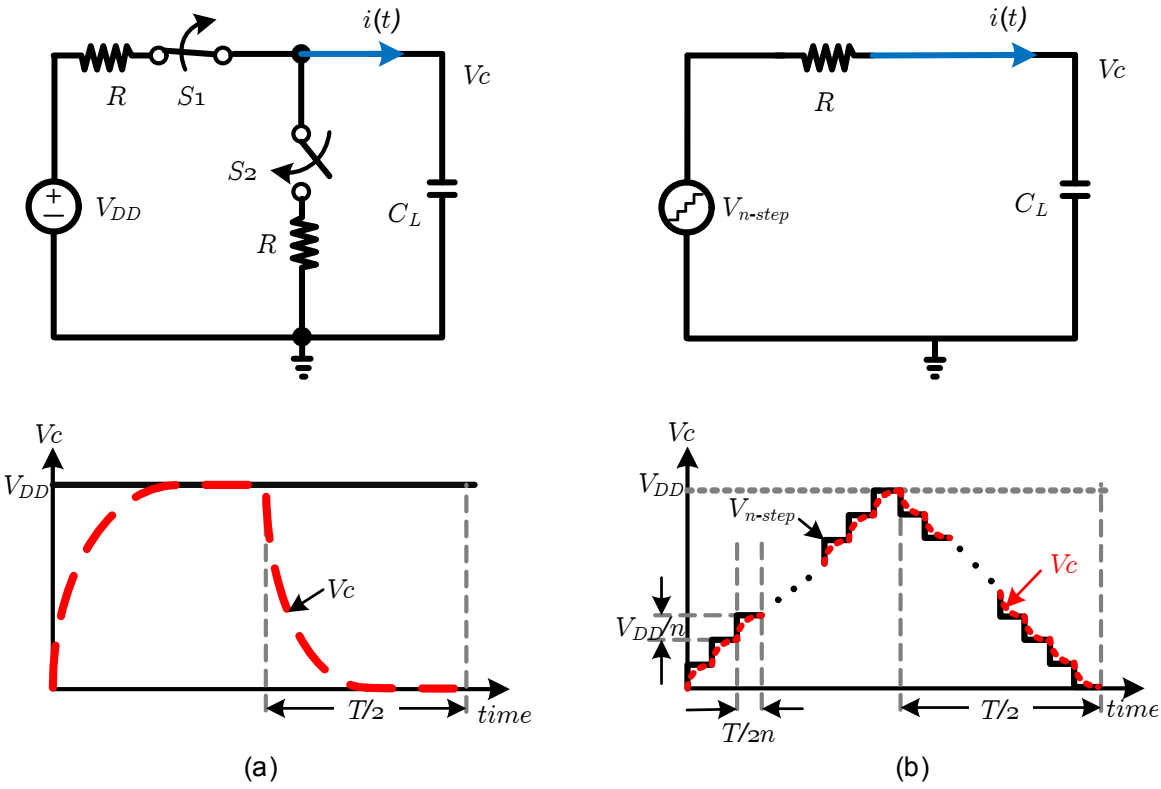


Figure 1.1: Charging and discharging load capacitance using conventional and charge-recovery techniques: (a) First-order RC network with DC supply. (b) First-order RC network with n-step supply.

For simplicity, we can model digital circuits as first-order RC networks, as shown in Figure 1.1(a). In such a network, a capacitive load C_L is charged and discharged through transistors which are modeled as a resistor R. By turning on switch $S1$ and turning off switch $S2$, supply V_{DD} starts to charge the load C_L from V_{SS} to V_{DD}, and the voltage level at node V_{out} ramps up as an function of $V_{DD}(1 - e^{-t/RC})$, where t is the transition time. The total energy dissipation in the circuit can be obtained by integrating the instantaneous power dissipation on the resistive device, I^2R, over the charging time of C_L:

$$
\begin{aligned}
E &= \int_0^\infty i_C(t) v_{out}(t)\, dt \\
&= \int_0^\infty C_L \frac{dv_{out}}{dt} v_{out}\, dt \\
&= C_L \int_0^V v_{out}\, dv_{out} \\
&= \frac{C_L V^2}{2}.
\end{aligned} \tag{1.1}
$$

The total energy drawn from the supply during charge is equal to the sum of the energy stored in C_L, $\frac{1}{2}C_L V^2$, and the energy loss on the resistor R, $\frac{1}{2}C_L V^2$.

Figure 1.1(b) illustrates how to charge and discharge C_L in a charge-recovery manner. In this figure, an n-step voltage source is used as an example. In practice, a supply could be a resonant source capable of reclaiming charge and re-using it for subsequent charging. This n-step source has a voltage of V_{DD}/n for each step and time interval of $T/2n$. While charging the capacitor, we assume that the time constant RC_L is much smaller than the time interval $T/2n$. This means that the output C_L can be charged to the same level of the supply during each time interval $T/2n$. To calculate the energy consumption in Figure 1.1(b) during charging, we apply Equation (1.1), and total energy dissipation is found to be

$$
\begin{aligned}
E &= \textit{Energy Dissipation per Step} \times n \\
&= \frac{C_L(\frac{V}{n})^2}{2} \cdot n \\
&= \frac{C_L V^2}{2} \cdot \frac{1}{n}.
\end{aligned} \tag{1.2}
$$

The total energy drawn from the supply during charging phase is equal to the sum

of the energy stored in C_L, $\frac{1}{2}C_LV^2$, and the energy loss on the resistor R, $\frac{1}{2}C_LV^2 \cdot \frac{1}{n}$. Unlike conventional CMOS circuits, energy loss for the circuit in Figure 1.1(b) decreases by a coefficient of $\frac{1}{n}$.

When discharging the capacitive load C_L with the same n-step source, the charge stored in the capacitive load C_L flows back to the source, and the amount of energy loss on resistor R is the same as when charging C_L. A typical voltage supply will shunt any returned energy to ground, dissipating it across some resistance, and rendering the charge-recovery discharging no more energy efficient than the conventional case. However, if the supply is, for example, a resonant source, it will be able to reclaim the returned charge and use it for subsequent charging. This discharging method is called charge recovery because the energy transferred to the capacitor is recovered and reused by the supply.

As the number of steps n approaches infinity, the dissipation approaches zero; however, large n also means small charge time $T/2n$. If $T/2n$ is small enough compared to the time constant RC_L of the circuit, the switching event may not be completed and the result in Equation (1.2) is no longer valid. Since n is related to the energy savings of the circuit, we conclude that there is a trade-off between time (T) and energy dissipation (E). This energy-time trade-off forms the basis of charge-recovery techniques.

Equation (1.2) implies that gradual transitioning is the key to achieve energy-efficient charge-recovery operation. Gradual transition reduces the potential difference across the resistive element, results in a low-level current flow, and thus minimizes energy dissipation.

It should be noticed that in most charge-recovery systems, the power supply performs a dual role, providing charge to internal circuit nodes, and synchronizing the computation of the gate. For this reason, such a power supply is usually referred to as a *Power-Clock* (PC).

1.2 LC Oscillation and Power-Clock

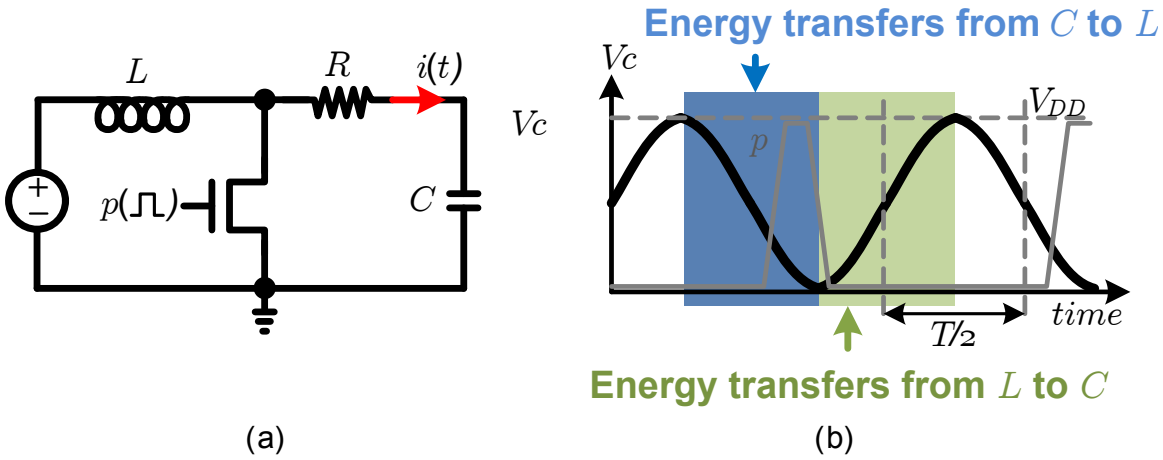


Figure 1.2: Practical implementation of power-clock using an inductor: (a) Schematic. (b) Waveform.

One realization of a power-clock generator with energy recovery ability is shown in Figure 1.2. Here, an inductor L is employed to store the energy of the charge returned from capacitor C in the form of a magnetic field. Periodic energy transfer between the inductor L and the capacitor C results in a sinusoidal waveform. The only energy losses in the system are due to the parasitic resistance of the circuit. To compensate for these losses, a shunt switch driven by pulse p is used to inject current into the inductor, replenishing energy in each cycle. The natural oscillation frequency f_n of the ideal LC system is

$$f_n = \frac{1}{2\pi}\sqrt{\frac{1}{LC}}, \tag{1.3}$$

where C is the total capacitance of LC oscillation system.

During practical design, the inductance value for L is chosen to meet a target frequency f_n with a given extracted load C. For example, with a load C_L and a target frequency f_n, the inductance is chosen as $L = 1/(2\pi f_n)^2 C_L = T^2/4\pi^2 C_L$. The

waveform of the forced LC oscillation is a sinusoid-like waveform, and the energy consumption can be derived as

$$\begin{aligned} E &= \frac{1}{2}\left|I^2\right| RT \\ &= \frac{1}{2}\left|\frac{V/2}{1/j\omega C}^2\right| RT \\ &= \frac{1}{2}(\frac{\omega V}{2}C)^2 RT \\ &= \frac{1}{2}(\pi f V C)^2 RT \\ &= \frac{1}{2}(\frac{\pi^2 CR}{T})CV^2. \end{aligned} \tag{1.4}$$

Using Equation (1.3), Equation (1.4) can be further simplified to obtain an expression of energy dissipation in terms of circuit parameters:

$$\begin{aligned} E &= \frac{1}{2}\cdot\frac{\pi^2 CR}{2\pi\sqrt{LC}}CV^2 \\ &= \frac{\pi}{4}R\sqrt{\frac{C}{L}}CV^2 \\ &= \frac{\pi}{4Q}CV^2. \end{aligned} \tag{1.5}$$

With the quality factor defined as $Q = \frac{\sqrt{LC}}{R}$, a direct energy dissipation comparison can be drawn between a charge-recovery system and a conventional CMOS one with the same load. As a result, Q is an important metric and is often used to evaluate the efficiency of a charge-recovery system.

1.3 Power-Clock Generator

Figure 1.3 shows a single-phase power-clock generator, which uses an inductor L and parasitic capacitance C_L to form an LC oscillation. This clock generator provides a single-phase sinusoidal clock waveform, and an inductor L is chosen to achieve a target resonant frequency for a given load C_L.

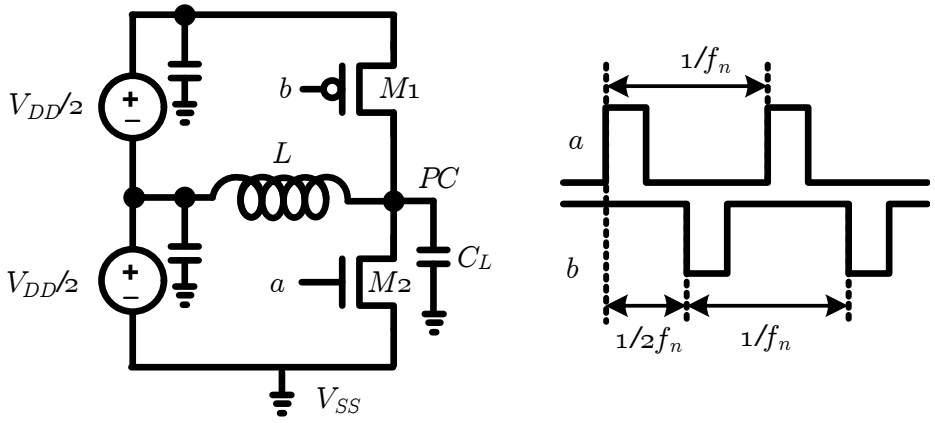


Figure 1.3: A single-phase power-clock generator with two supplies.

The clock drivers ($M1$ and $M2$), similar to the one used in [9], periodically replenishes the energy losses in the resonant system through current injection in the inductor. As the clock approaches its minimum, pulse a causes the pull-down switch to conduct, discharging the output clock voltage to 0V, and causing an RL current build-up in the inductor. At the falling edge of pulse a, the system continues oscillating freely with an initial condition $V(PC) = 0\text{V}$ and $I(L) = I_n$, where I_n is the current flowing in the inductor at that time. Similarly, when the clock reaches its peak, pulse b causes the pull-up switch to conduct, resulting in a similar RL current build-up in the inductor. At the rising edge of b, the system once again resumes a free oscillation, with an initial condition $V(PC) = V_{DD}$ and $I(L) = I_p$, where I_p is the current flowing in the inductor at that time.

The current build-up in the inductor at the crest and trough of $V(PC)$ enables the supply to provide energy to the system periodically, which is stored in the form

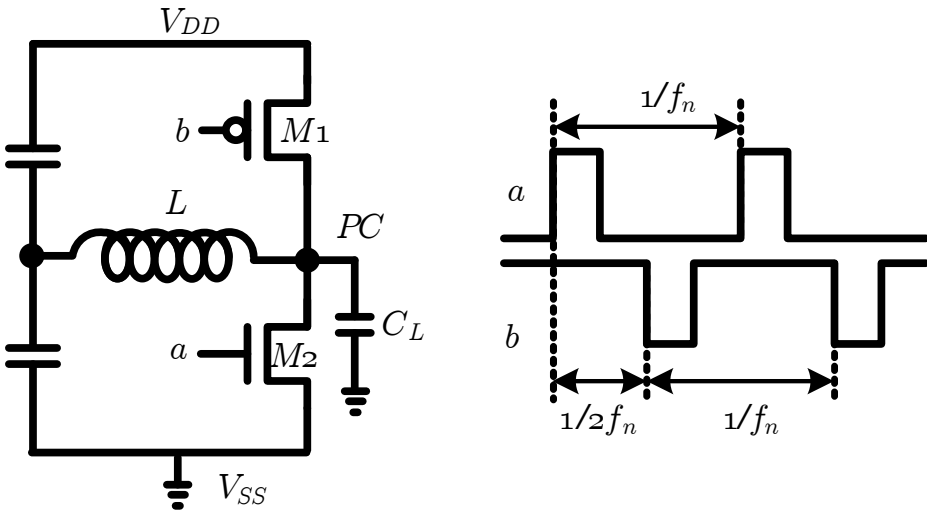


Figure 1.4: A single-phase power-clock generator with single supply and a voltage divider.

of a magnetic field in the inductor. The amount of replenished energy required to maintain stable oscillations is thus governed by the equation:

$$E_{replenished} = \frac{1}{2}LI_n^2 + \frac{1}{2}LI_p^2. \tag{1.6}$$

Notice that, in this driven power-clock generator, the natural frequency of the oscillation is determined by Equation (1.3), and the actual frequency of the generated sinusoid waveform is determined by the frequency of the pulses. If this frequency is too far away from the natural frequency of the oscillation, the generated waveforms will be distorted.

Instead of the two-supply scheme in Figure 1.3, Figure 1.4 uses a single supply with a capacitive voltage divider to achieve a similar functionality. Large capacitors are used in the divider to provide stable voltage sources at the cost of capacitor area.

Various circuit topologies for power-clock generators have been proposed for different charge-recovery logic styles and for different applications. Figure 1.5 shows an H-bridge clock generator that generates a two-phase power-clock with cross-coupled

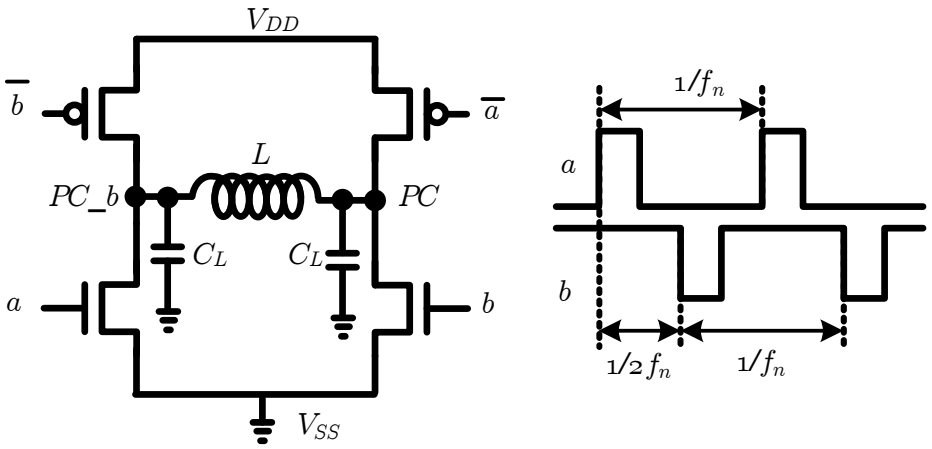


Figure 1.5: H-bridge two-phase power-clock generator.

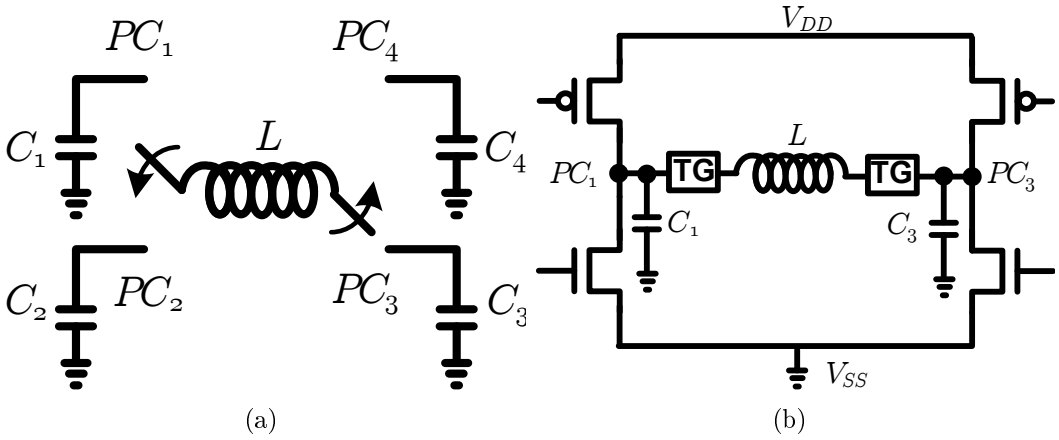


Figure 1.6: (a) Basic scheme for a four-phase power-clock generator. (b) Circuit schematic related to $C1 - C3$.

pairs of NMOS and PMOS transistors. It generates complementary power-clock waveforms using only one inductor.

A four-phase power clock generator has been presented in [11], and the schematic is shown in Figure 1.6. This circuit uses only one inductor to generate all four clock phases. The idea of using a single rotating inductor is shown in Figure 1.6(a). Rotation of inductor to transfer the energy between various clock phases is achieved by sequentially switching the transmission gates (TG), as shown in Figure 1.6(b) for

loads $C1$ and $C3$. One of the issues of this generator is that the power dissipation overhead of the controller circuit restricts its use to drive relatively smaller systems.

1.4 Charge-Recovery Systems

Charge recovery can be deployed various ways. The main two broad classifications are charge-recovery logic and resonant-clocked designs. Charge-recovery logic belongs to the class of fine-grain design, which employs charge-recovery techniques at the gate level to recover energy from load capacitance driven by all gates in a design. Resonant-clocked design is an example of coarse-grain design and recovers energy from the clock network in the design. Depending on the specific implementation of resonant-clocked designs, charge-recovery techniques may also be extended to the internal nodes of pipeline registers.

1.4.1 Fine-grain Systems

A fine-grain charge-recovery system is inherently gate-level pipelined with charge-recovery logic. This logic utilizes the idea of current-steering to conditionally charge or discharge load capacitance based on the outputs of its proceeding stage. To illustrate the structure, operation, and design of logic in a fine-grain charge-recovery system, we will use an early charge-recovery logic gate, 2N-2P [12] as an example. Even though the detailed implementations of various charge-recovery logic families differ , the underlying objectives, trade-offs and basic circuit topologies are quite similar.

Figure 1.7(a) shows an inverter implemented in 2N-2P with idealized power-clock waveforms, φ_1, φ_2, φ_3, and φ_4, shown in Figure 1.7(b). The gate utilizes cross-coupled PMOS transistors to steer the current from the power-clock generator to one of the (ideally) balanced output nodes, *out* or $\overline{out}$. The initial resolution at output nodes is determined by the complementary pull-down evaluation trees. The losses in such a charge-recovery system are due to steering devices and parasitic

wiring resistance through which the load current flows. For simplicity, we consider the operation of the gate with two non-overlapping idealized power-clock waveforms shown in Figure 1.7(c), and ignore the effect of the threshold voltage V_{th} on the operation of the gate.

The operation of a 2N-2P gate can be divided into four phases: evaluation, hold, reset, and wait. For correct operation, all gates are cascaded in a way that gates driven by φ_1 connect to gates driven by φ_2 and so on, until gates driven by φ_4 connect to gates driven by φ_1. At time t = 0, the gate connected to φ_1 is at the beginning of its evaluation phase. At this point, the input signal *in*, at its hold phase, provides full level inputs to transistor $M1$ and holds node *out* firmly to V_{SS}. Since $\overline{in}$ is low, transistor $M2$ is off, node $\overline{out}$ is floating during this time. As φ_1 begins to ramp up, transistor $M4$ conducts more strongly. Since *out* is held to V_{SS}, increasing V_{gs} of $M4$ causes $\overline{out}$ to track φ_1 closely. At the end of the evaluation phase, $\overline{out}$ reaches the full rail, while *out* remains at 0V. In the hold phase of the gate, *out* and $\overline{out}$ provide full level driving voltage to the next logic gate. Subsequently, as φ_1 ramps towards V_{SS}, the outputs of the gate reset. Note that when the gate is in reset phase, input nodes *in* and $\overline{in}$ are already in the reset phase, and $M1$ and $M2$ are both off. Consequently, $\overline{out}$ tracks the power-clock through $M4$ and discharges to V_{SS} gradually. Eventually, both outputs reach 0V and remain stable until the next evaluation phase.

In resonant systems, presenting a constant capacitive load C is important to provide a stable oscillation frequency, and the dual-rail property of 2N-2P logic can provide a near-constant load capacitance when looking into the φ node of the gate. As long as the effective capacitors at the 2N-2P output nodes are equal, the value of C in a resonant system is independent of the output state of gates, and stable LC oscillation is maintained. This is one of the salient advantages of using dual-rail logic in a fine-grain charge-recovery system. However, dual-rail logic has its limitations. One such limitation is the constant switching activity of each gate, which is

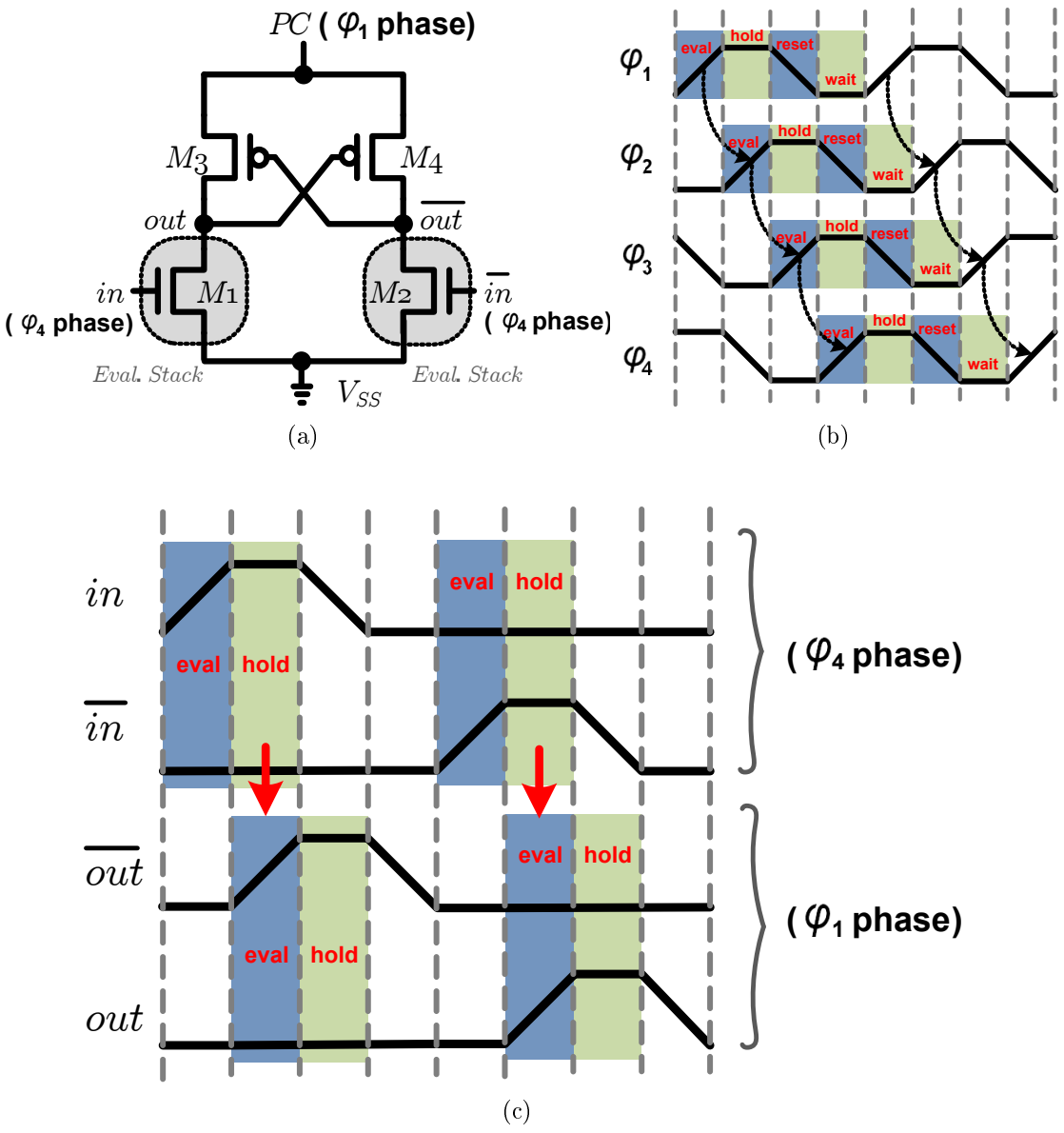


Figure 1.7: (a) Schematic of 2N-2P inverter. (b) Four-phase power-clock waveforms. (c) Operating waveforms of 2N-2P inverter.

50%, independent of the switching probability of the output state, since one of the two output nodes switches in every cycle. As a result, a fine-grain dual-rail charge-recovery system could cause low-switching activity gates to dissipate more compared to their conventional CMOS counterparts. A significant portion of energy savings is thus given up when applying fine-grain charge-recovery techniques to designs with

low switching activity.

The need of multiple clock phases for 2N-2P logic also increases the complexity and reduces the energy efficiency of the design. Use of multiple-clock phases requires more clock network broadcasting in physical design and thus increases design complexity. Gate cascading needs to follow a fixed phase order in 2N-2P operation, and only gates driven by specific phase-pairs can be connected. Furthermore, additional buffering is often required to phase-delay noncritical paths, balancing the paths so that they arrive at the same time as critical paths. This buffering results in additional power dissipation and affects the extent to which a design can be efficiently implemented using fine-grain charge-recovery techniques.

Fine-grain pipelining enables greater throughput but limits the range of designs that can be efficiently implemented. Charge-recovery techniques involve a fundamental trade-off between energy dissipation and latency, as suggested by Equation (1.4); however, they do not affect system throughput, since fine-grain systems are inherently gate-level pipelined. Moreover, the maximum number of evaluation stack height in each charge-recovery logic limits the function that can be implemented in each gate. As a result, deeper pipelining is sometimes needed compared to a traditional CMOS datapath. Fine-grain charge-recovery is therefore advantageous in systems that are throughput intensive and can tolerate increased latencies.

Implementing fine-grain charge-recovery techniques in sequential circuits with feedback is also a challenge, since both latency and throughput are adversely affected. With feedback loop in a design, no further computation can occur until the previous one is completed, so throughput and latency become correlated.

Although the majority of work in charge-recovery design has focused on fine-grain systems, these techniques are not limited to such logic gates. Charge-recovery techniques are particular effective in applications which involve nets with large switching activity. Regular datapath structures without feedback loops are also well-suited

to fine-grain charge-recovery implementations. In the next section, we will discuss designs that implement charge-recovery techniques on specific nets, leading to coarse-grain charge-recovery design.

1.4.2 Coarse-grain Systems

Unlike fine-grain designs, which employ charge-recovery techniques throughout the design, coarse-grain designs employ charge-recovery selectively in part of of the design, where the application of these techniques is more effective. In this section, a resonant-clocked pipeline is discussed as an example of coarse-grain system. In particular, resonant-clocked pipeline designs apply charge-recovery techniques on the parasitic capacitance of clock distribution network and in some cases, may extend to the internal capacitance of pipeline registers.

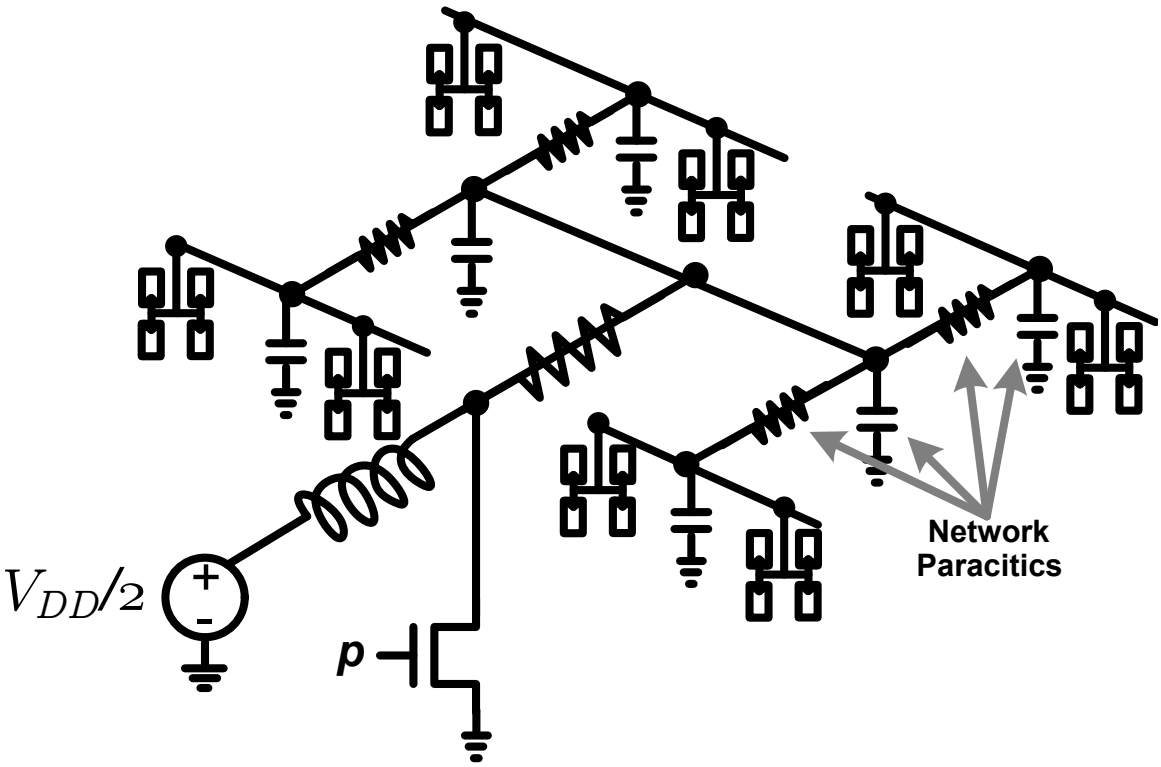


Figure 1.8: Resonant-clocked pipeline example.

Figure 1.8 shows an example of a resonant-clocked design where the timing ele-

ments (pipeline registers) are specially designed and clocked by a single-phase resonant clock, and combinational logic is identical to conventional CMOS design. An inductor L is used to resonate the capacitance of the entire clock network C in the system. Due to the high switching activity of the clock distribution network, substantial dynamic power reduction can be achieved with special pipeline register designs. In contrast to a fine-grain gate-level pipelined system, such designs are pipelined at a coarser level, which is similar to the conventional CMOS datapath with multi-stage pipelines. Consequently, coarse-grain systems can be designed to exhibit identical system-level timing properties with traditional clocked designs such as latency, throughput, and cycle time.

Unlike fine-grain charge-recovery systems, coarse-grain charge-recovery systems do not require additional buffers to balance delay, greatly increasing the range of designs that can be implemented in an energy manner. Furthermore, since logic gates in this system can be implemented with conventional CMOS logic, the design of resonant-clocked pipelines is more amenable to commercial tools, especially for synthesis and place-and-route.

1.5 Contributions

This section outlines the contributions of this thesis. The main motivation of this thesis is to apply charge-recovery techniques on designs and achieve better energy/performance trade-offs than conventional CMOS design. We demonstrate charge-recovery techniques on two systems with different frequency points. Both designs provide high performance at their corresponding supply-levels and achieve higher energy efficiency than their conventional CMOS designs counterparts.

1.5.1 SBL and SBL-Based FIR Filter Design

In this work, we present Subthreshold Boost Logic (SBL), a new circuit family that relies on charge-recovery design techniques to achieve order-of-magnitude improvements in operating frequencies while still achieving high energy efficiency using subthreshold DC supply levels. Specifically, SBL uses an inductor and a two-phase power-clock to boost subthreshold supply levels, overdriving devices and operating them in linear mode. Charge-recovery switching is used to implement this boosting in an energy-efficient manner.

To demonstrate the performance and energy efficiency of SBL, we also present a 14-tap 8-bit finite-impulse response (FIR) filter test-chip fabricated in a 0.13μm technology with $V_{th,nmos}$ = 400mV. The energy-efficient operation of the SBL-based FIR test-chip has been experimentally verified for clock frequencies in the 5MHz-187MHz range. With a single 0.27V supply, the test-chip achieves its most energy efficient operating point at 20MHz, consuming 15.57pJ per cycle with a recovery rate of 89% and a Figure of Merit (FoM) equal to 17.37nW/Tap/MHz/InBit/CoeffBit. With the introduction of a second subthreshold supply at 0.18V, energy consumption at 20MHz decreases further by 17.1%, yielding 14.40 nW/Tap/MHz/InBit/CoeffBit. At its maximum operating frequency of 187MHz, the test-chip achieves 35.31nW/Tap/MHz/InBit /CoeffBit and 34.47 nW/Tap/MHz/InBit/CoeffBit with one and two subthreshold supplies, respectively. To our knowledge, these figures of merit are the lowest published for FIR test-chips to date [13, 14]. In comparison with a static CMOS-based implementation derived by synthesis of the same FIR architecture and automatic place and route, the SBL-based FIR consumes 40% to 50% less energy per cycle in the 17MHz-187MHz range, based on device-level simulations, while incurring a 15% area overhead.

1.5.2 Flash ADC with Resonant Clock Distribution

Resonant clocking has been shown to be an effective approach to the reduction of power consumption in GHz clock speed distribution networks [14, 15, 16]. In the ADC presented in this thesis, resonant clocking is deployed to decrease the power consumption of the network that distributes the clock signal to the analog and digital circuitry of the ADC. Specifically, a fully integrated inductor is used to resonate the parasitic capacitance of the entire clock distribution network all the way to the clocked timing elements. This technique is thus compatible and can be used in conjunction with previous power optimization approaches at the circuit and architecture levels.

Operating in the vicinity of its resonant frequency with sampling rate 5.5GS/s, the ADC dissipates 28mW with only 10.7% of total power on clock distribution and a FoM equal to 396fJ per conversion step. This FoM is lower than all previously-published ADCs operating above 2.2GHz [2]. Correct operation has been validated at the clock rates up to 7GHz with 45mW of total power consumption and 11.1% of total power on the clock.

1.6 Thesis Outline

The remainder of this thesis is organized as follows: In Chapter 2, we give a brief introduction of reversible computing, which provides the main ideas and motivations of the charge-recovery design techniques. A summary of previous work is presented in the area of both charge-recovery logic and resonant-clocked designs.

In Chapter 3, we present our novel fine-grain charge-recovery circuit family, called SBL, that achieves high performance and high energy efficiency with a single subthreshold-level supply. We discuss its structure and operation, including high performance achievable through efficient signal boosting. We also give a detailed analysis of the energy consumption for SBL gates.

In Chapter 4, we describe an 8-bit 14-tap SBL FIR test-chip fabricated in 0.13μm technology. Results from device-level simulation of the SBL FIR filter and its static CMOS counterpart with identical architecture are given and compared. Measurement results from our SBL test-chip with both single- and two-supply schemes are presented and discussed. This work was published in [17, 18].

Chapter 5 presents the architecture of our resonant-clock flash ADC design and its main building blocks. It also gives detailed designs, and discusses operations of resonant-clocked dynamic comparators and sense-amplifier flip-flops used in our resonant-clock flash ADC.

Chapter 6 presents the design, evaluation, and testing of a 7GSample/s resonant-clock flash ADC test-chip. The ADC has been designed and fabricated in a 65nm bulk CMOS process. An on-chip inductor is used to resonate the entire clock distribution network, and a detailed analysis of the inductor from a commercial 3D full-wave electromagnetic field solver is given. Measurement results from our resonant-clock ADC test-chip with ADC performance characterization are also presented. This work was published in [19].

Chapter 7 summarizes the contributions of this thesis and presents directions for future research in this area.

CHAPTER 2

Background

In this chapter, we survey previous work on charge-recovery. In Section 2.1, we describe reversible computing, which inspired later charge-recovery techniques. Section 2.2 covers early research on charge-recovery logic. Section 2.3 discusses previous work in the area of resonant-clocked designs. Following these early work, we explore the techniques and challenges in charge-recovery area, which lead to the motivation of the work in this thesis.

2.1 Reversible Systems

2.1.1 Reversible Computing

Long before energy dissipation emerged as a matter of interest in VLSI design, physicists inquired into the fundamentals of energy dissipation and the loss of information in a computing system [20]. In the early papers, researchers largely focused on discussing the possibility of having physical machines which consume zero energy while computing and tried to find a lower bound on energy consumption. One of the conclusions drawn by Landauer is that the minimum possible amount of energy required to change one bit of information is equal to $kTln2$ [21], where k is the Boltzmann constant, and T is the absolute temperature of the environment. This loss of energy becomes heat expelled into the surroundings.

Another result of early work in the area is that to achieve zero energy computation, the operation must be reversible or be implemented in reversible logic [22]. If the devices in a computing system are designed to change state in a way that is logically reversible, in which no known bits are erased, then in principle, arbitrarily little free energy needs to be used. Of course, in practice, there are other sources of energy dissipation such as leakage, or resistive loss. However, unlike the $kTln2$ energy dissipation for changing a bit operation, there are no fundamental lower bounds for these sources of dissipation.

2.1.2 Reversible Logic

Reversible logic gates are digital circuits in which the number of inputs is equal to the number of outputs and there is a one-to-one mapping between vectors of inputs and outputs [23]. Therefore, in these gate, the input vectors can always be reconstructed from output vectors. A gate with k inputs and k outputs is called a $k \times k$ gate, and all gates in a reversible circuit must be reversible. Reversible logic realizes balanced functions on all outputs, i.e., half of all minterms are mapped to 1 and the other half to 0. Consequently, garbage outputs are necessary to realize non balanced functions (e.g., AND, OR, XOR etc).

Consider a reversible logic gate with two inputs. A conventional XOR gate takes two single-bit inputs A and B, and yields one single-bit output X. If A = B, then X = 1; otherwise X = 0. However, the XOR gate is not reversible, since we cannot uniquely determine what the input vector (A, B) is from the output. For instance, the output X = 1 could have come from either one of the two possible input vectors, (0, 1) and (1, 0). Although a reversible gate must be a $k \times k$ gate, not all gate with k-bit input and k-bit output are reversible. For example, consider a gate with input vector (A, B) and output vector (X, Y), where X= A XOR B, and Y = A OR B, as shown in Figure 2.1. Clearly, this is not a reversible gate, since there exists a 2-to-1

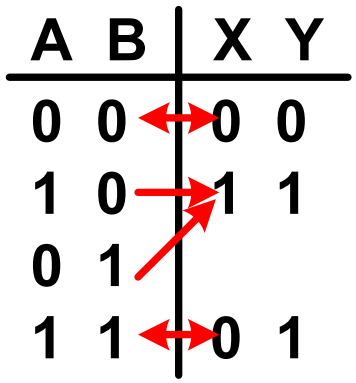
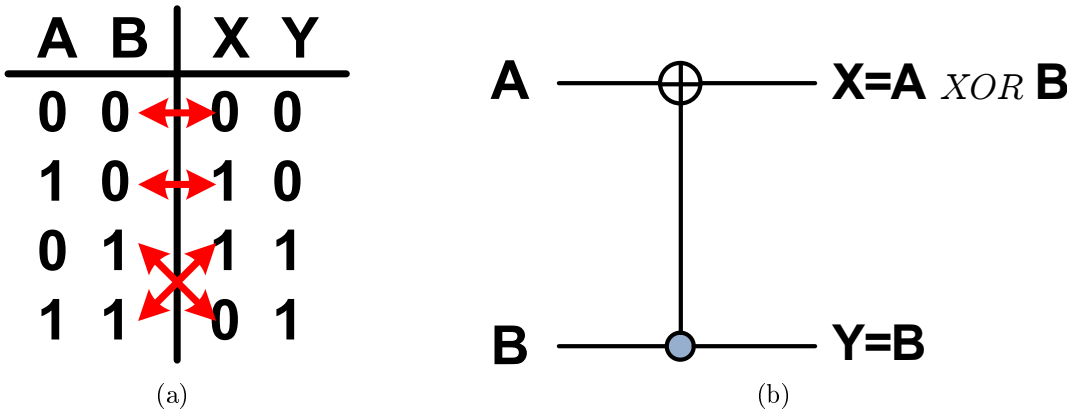


Figure 2.1: Truth table of an irreversible gate.

mapping from the input to the output.

For a two-bit input and two-bit output gate to be reversible, the mapping of its logic function should be such that the set of output vectors is a permutation of the set of input vectors, (00, 01, 10, 11). It then follows that there exist 4! = 24 possible reversible two-bit input and two-bit output logic gates. The Feynman gate is one of the most well known reversible two-bit input and two-bit output logic gates [24]. A Feynman gate with input vector (A, B) and output vector (X, Y) implements the logic functions: Y = B , and X = A XOR B. From input-to-output mappings, shown in Figure 2.2(a), it is evident that the Feynman gate is a reversible gate. In this example, one of the inputs B serves as a control signal. If B = 0, then the output X is simply duplicating the input A; if B = 1, then the output X = $\bar{A}$ (inverse of the input A). For this reason, the Feynman gate is also called the controlled NOT gate, or the quantum XOR gate (Figure 2.2(b)) due to its popularity in the field of quantum computing.

Implementing reversible logic in integrated circuits has many challenges and issues. As suggested in [25], reversible computation requires all logic operations to be carried out once in the forward direction and once in the backward direction, yielding additional latency and circuit overhead. More importantly, it requires a large amount of temporary storage to maintain intermediate results until computation in



(a) (b)

Figure 2.2: (a) Truth table and (b) symbol of the Feynman gate.

the backward direction is ready. Since storing these temporary values results in energy and circuitry overheads, implementing fully logically reversible logic in CMOS is not particularly attractive.

A practical alternative to full reversible logic is to use the idea of reversible computing in engineering systems with charge-recovery techniques, and try to approach the theoretical possibility of zero dissipation as closely as possible. If state information of a node in a circuit is available, and utilized when switching the state of that node, no information is lost. As a result, most of the free energy is conserved in the circuit and recycled for later reuse, rather than being dissipated.

2.2 Charge-Recovery Logic

Charge-recovery logic is a class of circuitry that recycles energy from the output load capacitance of logic gates to achieve ultra-low power operation. It utilizes the idea of current-steering to conditionally charge or discharge load capacitance based on the outputs of its proceeding stage. Each charge-recovery logic gate is a inherently gate-level pipeline stage, performing the role of both functional logic and timing elements in conventional CMOS designs.

Early charge-recovery logic, such as Split-level Charge-Recovery Logic (SCRL)

[26] and Reversible Energy Recovery Logic (RERL) [27], implemented fully logically reversible gates in CMOS. As pointed out in [25], the large number of temporary storage elements in a fully reversible circuit yields large circuit overheads. Later work in charge-recovery logic deviates from fully reversible circuits. These charge-recovery logic family retained the gradual transition of the charge between computation nodes but avoided design of reversible logic gates, keeping information around to reduce change in energy.

Figure 2.3 shows the structure of NMOS and PMOS inverters in Adiabatic Dynamic Logic(ADL). Proposed by Dickinson and Denker in 1995, ADL is a single rail adiabatic logic family [28]. ADL uses a diode to precharge the output node *out* to high when the power-clock rises, and the evaluation stack conditionally discharges the node *out* as the power-clock falls. The PMOS ADL inverter works in an opposite fashion. In an ADL system, ADL gates are cascaded by alternating the NMOS ADL and PMOS ADL gates. Both of them are synchronized by a two-phase power-clock with 180 degree phase difference to ensure that precharge and evaluation phases of all gates are synchronized. A chain of 64 inverters was successfully verified with an external 250MHz power-clock in 0.9μm process [29].

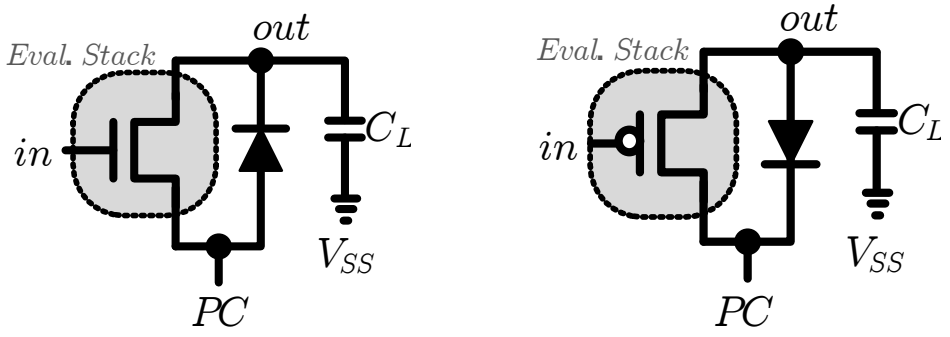


Figure 2.3: Schematic of (a) NMOS ADL inverter, and (b) PMOS ADL inverter.

Quasi-Static Energy Recovery Logic (QSERL) has been proposed by Ye and Roy [30] and has a similar diode structure to ADL, as shown in Figure 2.4. Instead

of precharging the output nodes with diodes, QSERL uses diodes to conditionally hold/discharge the voltage level at output nodes. This single-rail structure for both ADL and QSERL exhibits a data-dependent clock load, yielding high clock jitter and degrading system performance. Moreover, the use of diode in this logic may cause a substantial potential difference and generate large current flow, reducing energy efficiency.

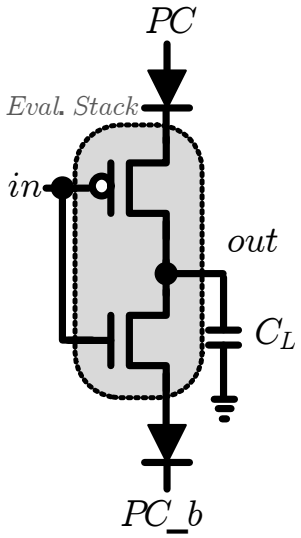


Figure 2.4: Schematic of QSERL inverter.

2N-2P is another early charge-recovery logic [12], and is has been discussed in detail in Chapter 1. Unlike ADL, the dual-rail topology of 2N-2P gates provides a data independent clock load for charge-recovery systems. However, requiring of four-phase power-clock increases design complexity and limits its applicability.

Various charge-recovery logic families have been proposed since 2N-2P. 2N-2N2P [12] is a variation of 2N-2P, as shown in Figure 2.5. By adding an additional pair of cross-coupled NMOS devices at the bottom, 2N-2N2P eliminates floating outputs during the hold phase.

Pass-transistor Adiabatic Logic (PAL), proposed by Oklobdzija et al.[31], is shown in Figure 2.6. PAL retained the cross-coupled PMOS structure in 2N-2P, and moved its evaluation stacks in parallel to the PMOS devices. Instead of a four-phase power-

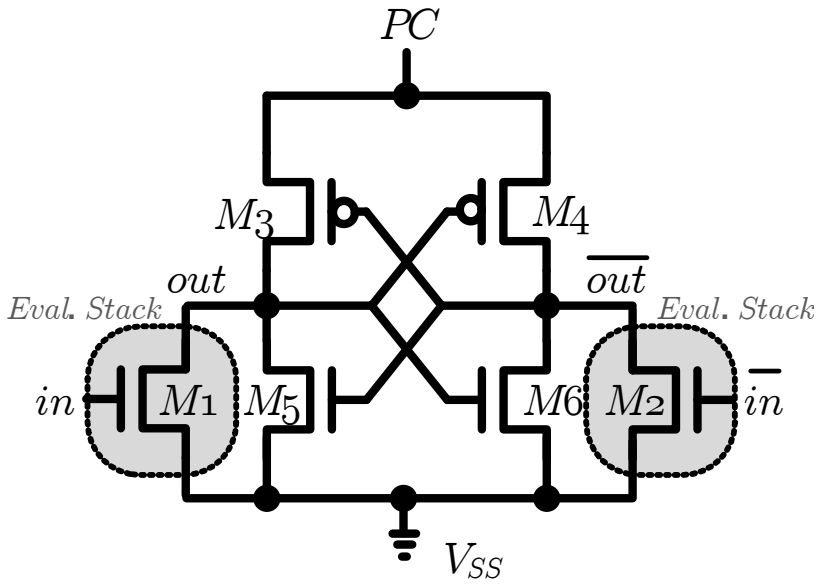


Figure 2.5: Schematic of 2N-2N2P inverter.

clock in 2N-2P, PAL can operate with a two-phase power-clock. A shift register with 1,600-stage PAL has been fabricated in a 1.2μm technology and correct operation has been verified at 10MHz.

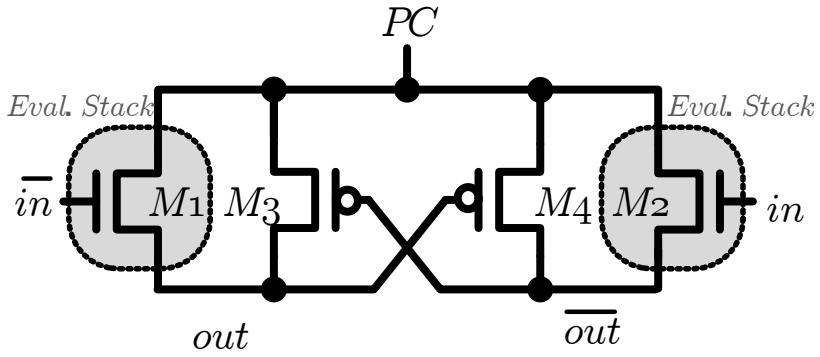


Figure 2.6: Schematic of PAL inverter.

Positive Feedback Adiabatic Logic (PFAL), shown in Figure 2.7, was proposed by Vetuli et al. [32]. PFAL is also a dual-rail charge recovery logic with a two-phase power-clock. Similar to 2N-2N2P, it eliminates floating outputs by using a pair if cross-coupled devices and its evaluation stacks are in parallel to PMOS devices. Compared to PAL, the cross-coupled NMOS and PMOS structure provides higher energy efficiency due to less leakage current. Compared to 2N-2P and 2N-2N2P,

PFAL has the potential to achieve higher operating frequency due to the full-rail input during evaluation phase.

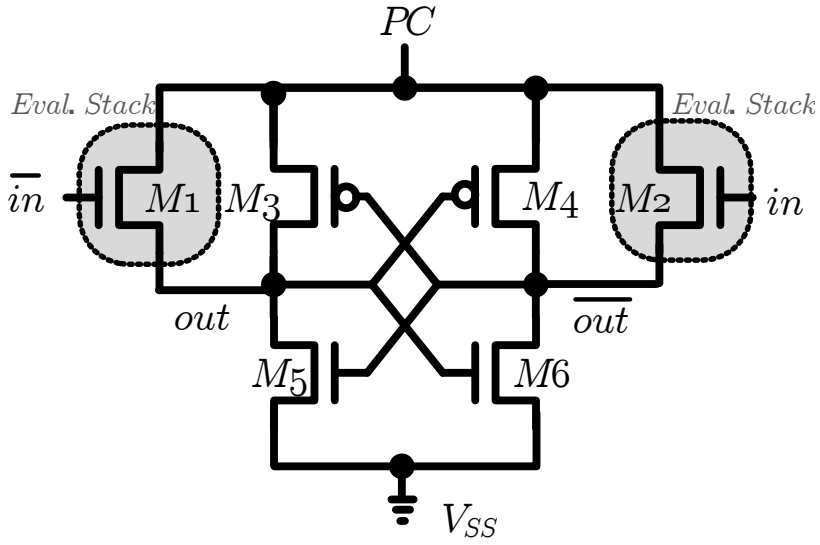


Figure 2.7: Schematic of PFAL inverter.

Figure 2.8 shows an inverter gate in True Single-Phase Energy-Recovering Logic (TSEL) [33]. TSEL cascades use alternating NMOS and PMOS stage and operate with a single-phase power-clock. A pair of current control switches ($M3$ and $M4$) and reference voltages (V_{PREF} and V_{NREF}) are used to improve its energy efficiency.

A Source-Coupled Adiabatic Logic (SCAL) [7] is derived from TSEL gates by replacing each DC reference voltage with a current source ($M7$), shown in Figure 2.9. Each current source can be individually tuned by transistor sizing and globally adjusted with PMOS and NMOS biasing voltage to optimal operating condition. An 8×8 multiplier test-chip has been fabricated in 0.5μm technology, and correct operation has been verified with operating frequency up to 130MHz [8].

The charge-recovery logic families discussed so far all face a common challenge of efficient operation at high frequency. To address this challenge, Sathe et al. proposed Boost Logic [34], which utilizes gate overdrive, reduced output swing, and charge-recovery techniques to achieve energy efficient operation at high operating frequency. Figure 2.10 shows the schematic of a buffer implemented in Boost Logic. Boost Logic

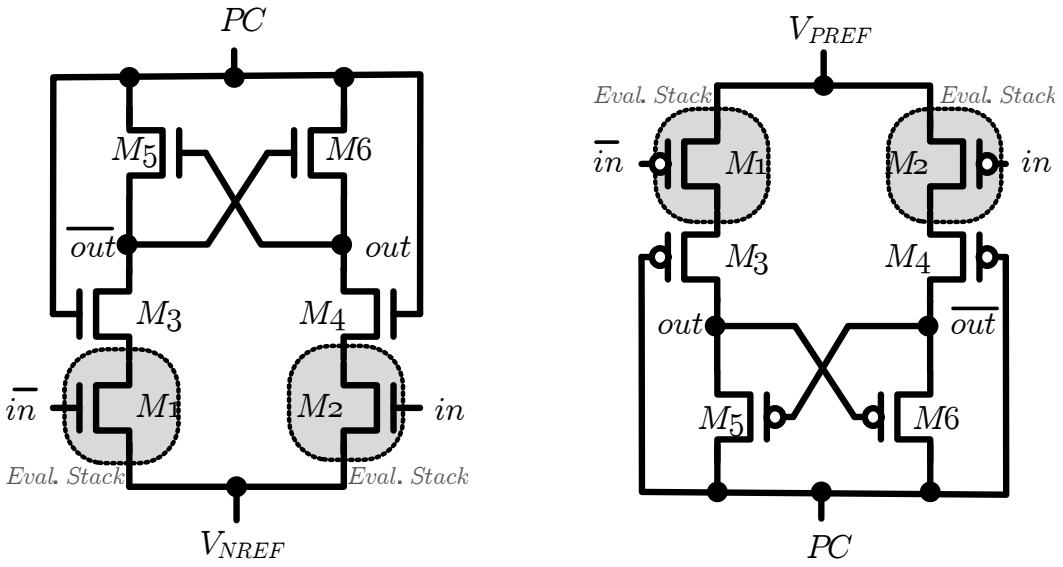


Figure 2.8: Schematic of (a) NMOS TSEL inverter, and (b) PMOS TSEL inverter.

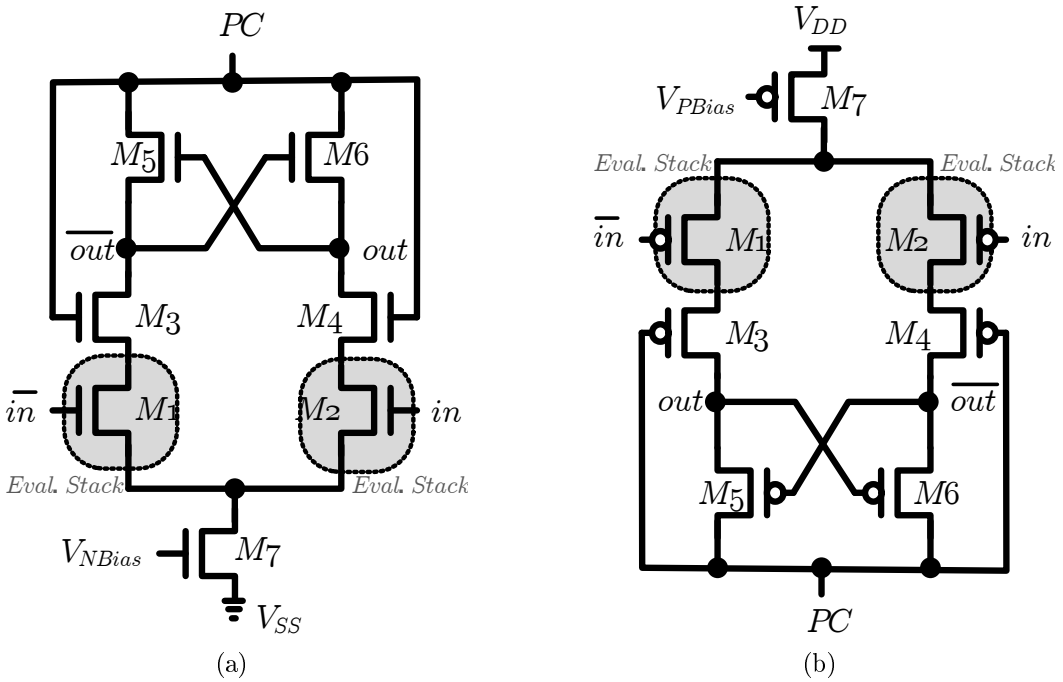


Figure 2.9: Schematic of (a) NMOS SCAL inverter, and (b) PMOS SCAL inverter.

is a two-phase, dual-rail, partially charge-recovering logic. The structure of a Boost gate can be divided into two parts − logical evaluation (Logic) and charge-recovery amplification (Boost). Logic performs functional evaluation when power-clock is low. When power-clock rises, Boost amplifies the potential different at output nodes to a

full-rail signal. The output nodes of Boost Logic are precharged to near $\frac{1}{2}V_{DD}$, which reduces the output swing of the gate when power-clock rises and thus reduces the energy dissipated in the Boost stage.

Another feature of Boost Logic that enables its efficient high-frequency operation is the fact that the logic stage provides the complementary output nodes with an initial voltage difference. Voltage difference ($\frac{1}{3}V_{DD}$) at outputs is pre-resolved at the onset of boost conversion, which precludes any 'fight' between cross-coupled inverters and results in efficient boost conversion. To demonstrate the high performance and energy efficiency of Boost Logic, a test-chip with eight chains of AND, OR, XOR, and INV gates was fabricated in 0.13μm technology, and correct operation was verified at operating frequencies exceeding 1GHz.

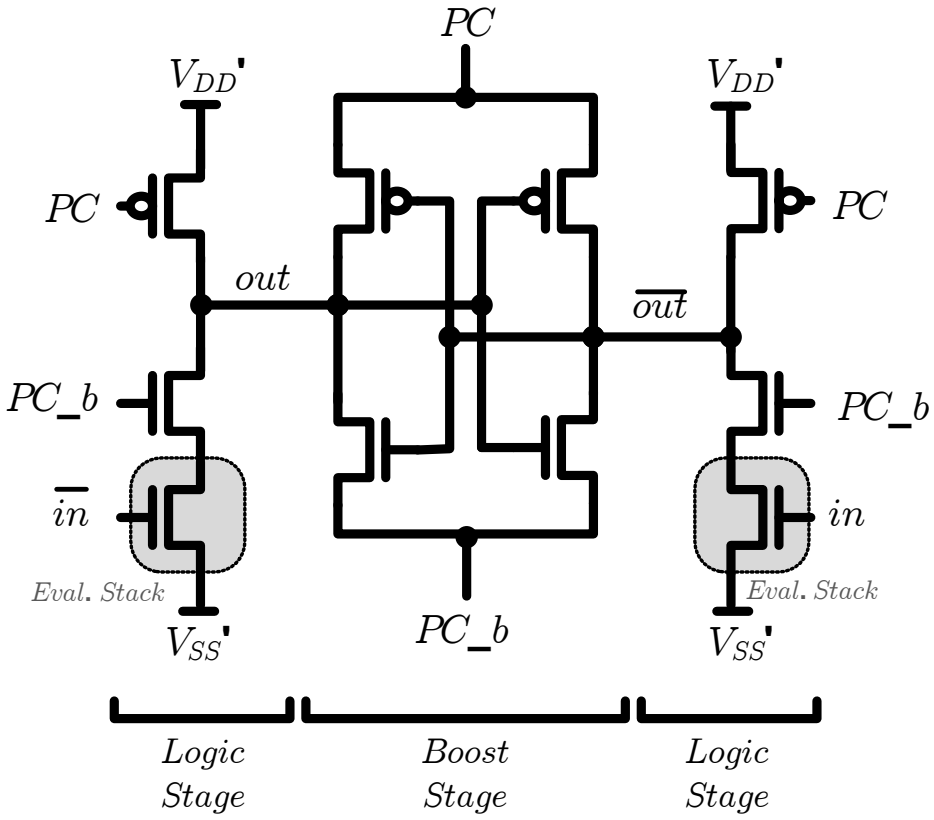


Figure 2.10: Schematic of Boost Logic inverter.

Looking back at the evolution of prior work, one observes that many similar traits

are shared among various charge-recovery logic families. The first common trait is the pair of cross-coupled PMOS devices which are used to steer the current of the power-clock and bring output to a full-rail signal. Another common feature of these gates is the fact that multiple clock phases, supplies, or clock devices are used to reduce the short current caused by the gradual transition of power-clock. These common traits provide us with good guidance in designing future charge-recovery logic.

2.3 Resonant-Clocked Designs

Resonant clocking is a charge-recovery design methodology that recovers energy from the clock distribution network. Due to the high switching activity and large capacitance of clock distribution networks, it becomes a good candidate for the application of charge-recovery techniques. Resonant-clock datapaths are similar to conventional-clock datapaths and combinational logic in resonant-clock designs can be implemented with conventional CMOS style logic.

Most of the previous work in resonant-clock designs focuses on developing new timing elements to reduce the leakage current and improve the performance due to the slow transition of the resonant clock. Athas et al. presented the E-R latch [35], shown in Figure 2.11. The E-R latch uses the bootstrapping technique at node bn, and the resonant clock recycles charge at both clock nodes and from the internal node n. Charge recycling at node n improves the total amount of recoverable capacitance but at the expense of reduced recovery efficiency due to the series resistance from $M3$. Correct operation of the E-R latch has been verified in AC-1, a 58.5MHz resonant-clocked microprocessor in 0.5μm technology [35].

Figure 2.12 shows the PMOS energy recovering flip-flop (pTERF) and NMOS energy recovering flip-flop (nTERF) proposed by Ziesler et al. [10]. A cross coupled NOR/NAND gate is added to convert the internal resonant signals to static ones, and an inverter is introduced at output node Q to increase drive strength.

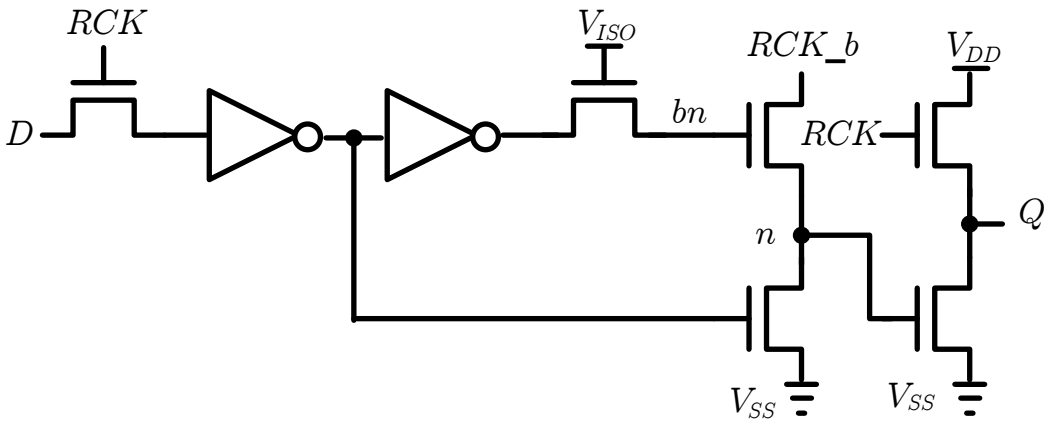
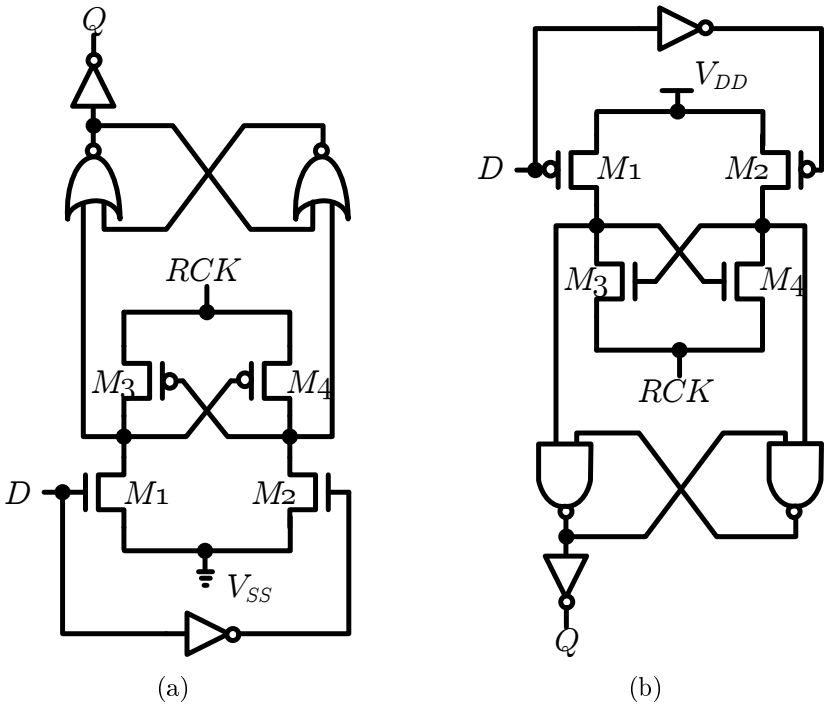


Figure 2.11: Schematic of Edge-Triggered (E-R) latch.

pTERF/nTERF has a similar issue to the E-R latch: the effective resistance of the cross-coupled devices dissipates energy during the charge and discharge cycles, limiting overall energy recovery efficiency. A 115MHz wavelet-transform test-chip has been fabricated in 0.25μm technology with pTERF and achieved 25% to 30% of energy saving compared to CV^2 [9].



(a) (b)

Figure 2.12: Schematic of (a) pTERF flip-flop, and (b) nTERF flip-flop.

Ishii et al. [1] deployed resonant clocking in conjunction with sense-amplifier flip-

flops, shown in Figure 2.13. They successfully demonstrated energy efficient operation on an ARM926EJ-S^{TM} microprocessor with operating frequency up to 200MHz. Clock-related power is reduced by 85%, and total power savings range from 20% to 35%, depending on application profile.

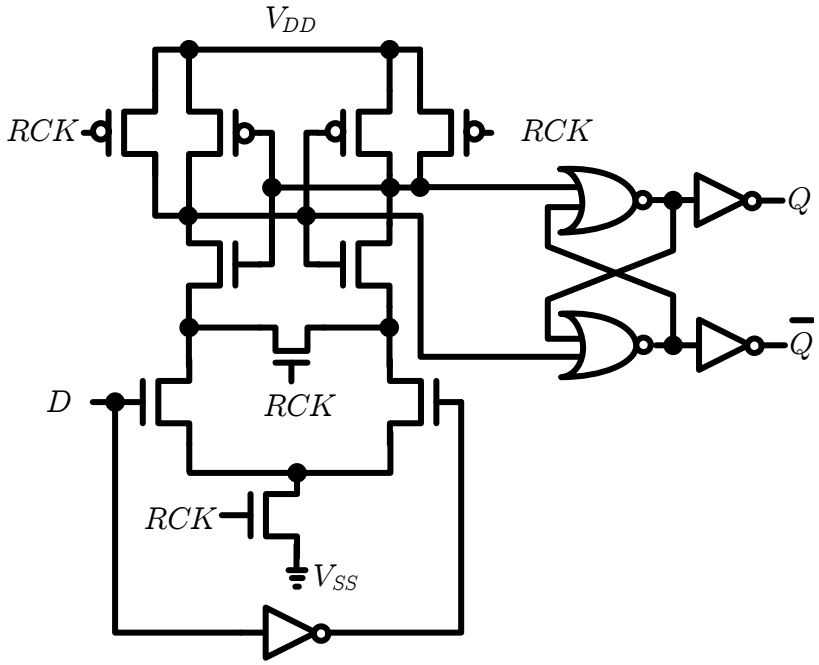


Figure 2.13: Schematic of sense-amplifier flip-flop used in [1].

Hansson et al. [36] have fabricated a 1.56GHz resonant-clock network in a 0.13 μm technology with an integrated 1.2nH inductor. A single-phase resonant clock directly drives 896 conventional master-slave flip flops without any intermediate buffers. The relatively slower edge-rate of the sinusoidal clock is reported to increase the power consumption in the flip-flops by 34%. Despite the higher power consumption in the flip-flops, there is still a 57% clock power saving, resulting in a total power reduction of 20% compared to the conventional clock network.

Chan et al. [15] later applied the global resonant clock distribution topology to a commercial microprocessor. Their processor has 830 on-chip spiral inductors with a natural resonant frequency of 3.2GHz. Unlike the work in [36], which drives flip-flops directly, Chan inserted local clock buffers between the resonant clock mesh and the

timing elements, resulting in limited energy recovery.

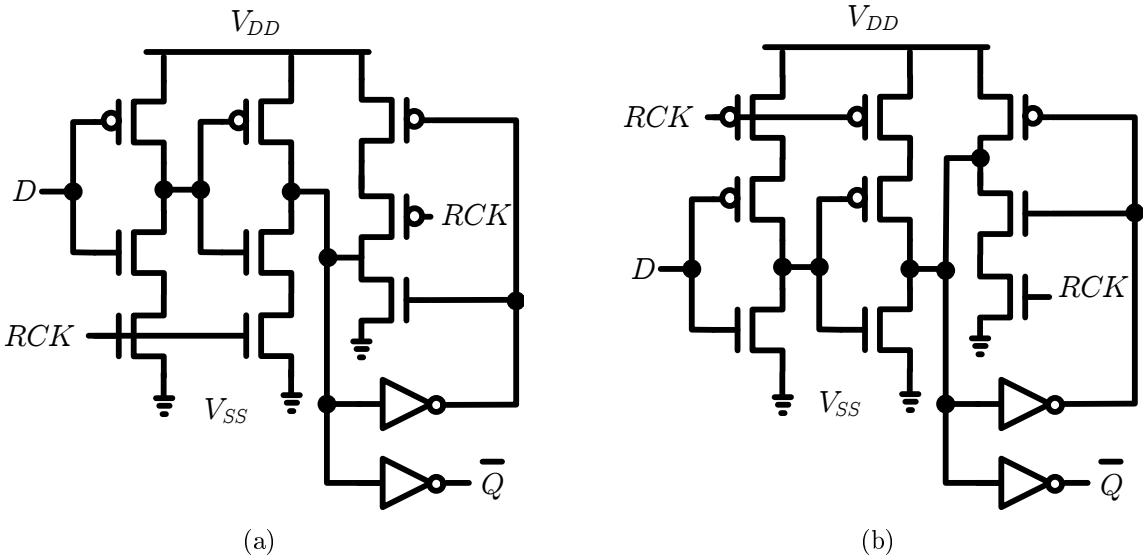


Figure 2.14: Schematics of resonant-clocked latches used in RF1: (a) H-LAT, and (b) L-LAT.

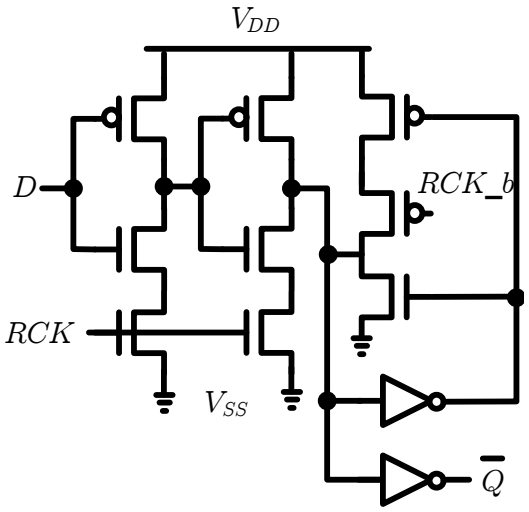


Figure 2.15: Schematic of resonant-clocked latch used in RF1, B-LAT.

Other work in this area has focused on evaluating and characterizing variability on resonant clocks. Chueh et al. [37] implemented a two-phase resonant clock network with programmable drivers and loading to evaluate the effect of imbalanced clock load on clock skew. The 2mm×2mm distribution network with on-chip inductors fabricated in 0.13μm technology performs a forced oscillation in the 900MHz to 1.2GHz

range. When running off-resonance by 10%, power dissipation increases by 3% and clock amplitude drops by 3%. Imbalanced loading impacts power and amplitude by less than 2%. When shifting from balanced to imbalanced loading, worst-case skew increases by 6% of cycle time.

More recent implementations of resonant-clocked designs focus on improving performance of timing elements. Sathe et al. have proposed a resonant-clock latch-based methodology and demonstrated its high performance and energy efficient operation with two FIR filter test-chips, RF1 and RF2, in a 0.13μm technology [38, 39]. Unlike flip-flops, which rely on sharp clock edges for effective operation, latch performance is primarily determined by the voltage level of the clock waveform. Moreover, latch-based designs have the potential to achieve higher performance than flip-flop-based designs, because the level-sensitive latches allow data to ripple through latch boundaries and enable time-borrowing across logic stages. Resonant-clocked latches used in RF1: H-LAT and L-LAT are shown in Figure 2.14 and B-LAT, used in RF2 is shown in Figure 2.15. RF1 and RF2 achieved 76% and 84% energy efficiency when operating their natural resonant frequencies of 1.03GHz and 1.01GHZ, respectively [14].

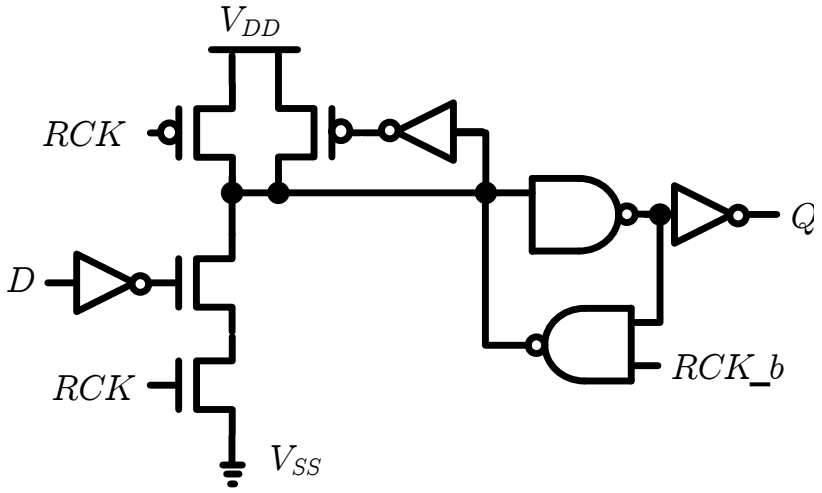


Figure 2.16: Schematic of DESL inverter.

Kao et al. introduced Dynamic Evaluation Static Latch (DESL) logic [16], a dy-

namic gate with a level-sensitive latch at the output stage, as shown in Figure 2.16. DESL mitigates performance degradation due to the resonant clock waveforms by relying on the voltage level of the clock. Moreover, a static latch provides a modest performance boost by relying on time borrowing and reduces dynamic power by reducing switching activity on large capacitance nets.

An 8-cycle reduced-latency fused-multiply-add single-precision FPU test-chip has been fabricated in a 90nm technology to demonstrate energy efficiency and performance of DESL. The resonant-clocked FPU operates with clock frequencies up to 2.07GHz, yielding 66.4% lower clock power and 31.5% lower total power over a conventionally-clocked version of the same architecture.

CHAPTER 3

Subthreshold Boost Logic

In this chapter, we present Subthreshold Boost Logic (SBL), a new circuit family that relies on charge-recovery design techniques to achieve order-of-magnitude improvements in operating frequencies while still achieving high energy efficiency using subthreshold DC supply levels.

The remainder of this chapter is organized as follows: Section 3.1 describes previous work on subthreshold designs. Section 3.2 presents the structure of SBL and the two-phase power-clock generator used for SBL. Section 3.3 describes the two-phase operation of SBL gates and the overdrive property of cascaded SBL gates. Section 3.4 analyzes the energy consumption of SBL gates. Summary is in Section 3.5.

3.1 Introduction

Early subthreshold circuit designs appeared in electronic watches in the 60s and 70s, driven by form factor limitations on battery size [40]. The recent emergence of untethered applications and energy scavenging devices has led to renewed interest in this field. A 1024-point FFT processor explored aggressive subthreshold designs for minimum energy operation, achieving a clock speed of 10KHz with a 350mV supply in a 0.18μm process with V_{th} = 450mV [40]. The Subliminal subthreshold processor achieved 833KHz with a 360mV supply using a 0.13μm process with V_{th} = 400mV

[41]. The Phoenix processor deployed leakage reduction techniques to achieve pW-level power consumption, targeting multi-year operation in sensor applications [5, 42]. Fabricated in a dual-threshold 0.18 μm process with $V_{th1} = 400$mV and $V_{th2} = 700$mV, it achieved 2.8pJ/cycle at 106KHz with a 385mV supply.

A common issue underlying all subthreshold circuit designs is that the significant energy advantages achieved through deep voltage scaling result in subthreshold currents, typically resulting in sub-MHz clock frequencies. Recent subthreshold designs have deployed circuit and architecture techniques to improve circuit robustness by improving gate overdrive. The 32-bit RISC core in [43] and the 8×8 FIR filter in [44] both deployed body biasing techniques to enable increased operating frequency, achieving 375KHz at 230mV, and 12KHz at 200mV respectively. A high-speed variation-tolerant interconnect technique relied on capacitive boosting to elevate the critical gate supply voltage and was demonstrated through a 6MHz clock distribution network with 400mV voltage supply [45]. A super-pipelining approach was demonstrated in [46], where the multipliers in a 1024-point FFT were heavily pipelined to reduce stage delay, achieving 30MHz with a supply of 270mV.

Subthreshold Boost Logic, introduced in this chapter, is a circuit family capable of operating at multi-MHz clock frequencies using subthreshold supplies. Unlike subthreshold circuitry, in which computations are performed using subthreshold currents and clock frequencies are typically limited to sub-MHz levels, SBL gates are overdriven to operate in the linear region, achieving order-of-magnitude improvements in operating speed over subthreshold logic. Energy efficient operation is ensured through the use of aggressively-scaled DC supplies at sub-threshold levels and by deploying charge recovery design techniques to boost these subthreshold supply levels by 3X to 4X.

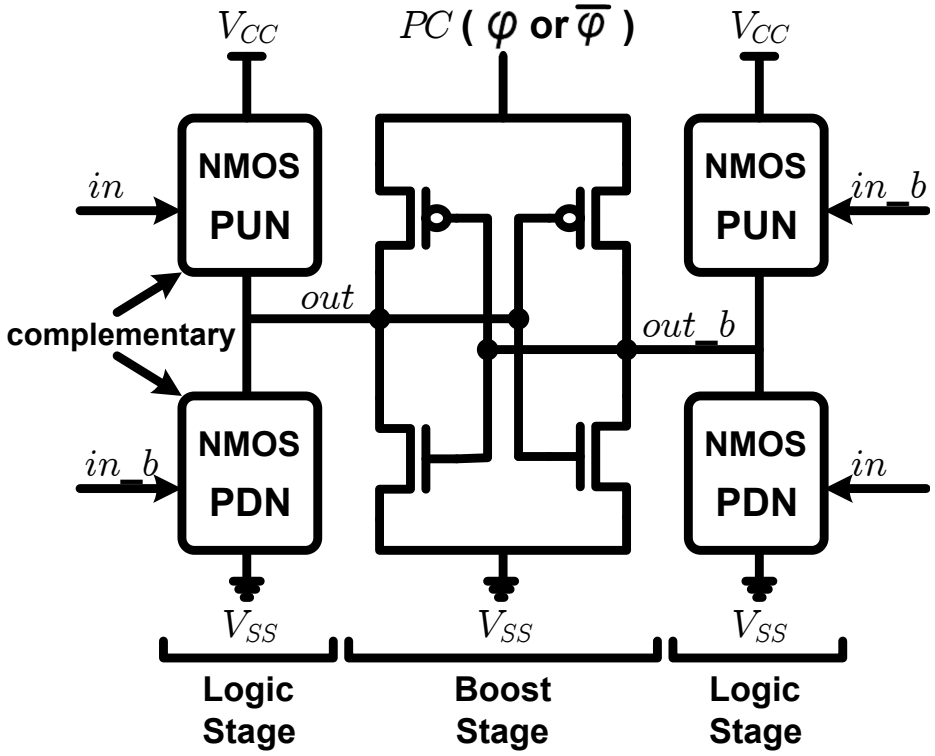


Figure 3.1: Schematic of an SBL gate.

3.2 SBL Overview and Blip Clock Generator

The structure of a SBL gate is shown in Figure 3.1. Each SBL gate consists of two stages: Logic and Boost. The Logic stage has differential outputs *out* and *out_b*. Each output is driven by a pull-up network (PUN) and a pull-down network (PDN), similar to static CMOS logic, except that an NMOS PUN is used instead of a PMOS one for increased gate overdrive ability. The Boost stage comprises a pair of cross-coupled inverters connected to ground (V_{SS}) and a charge-recovery power-waveform PC. From a functional standpoint, each SBL gate consists of a combinational logic block driving a transparent latch. Cascades of SBL gates are formed by clocking the gates on alternating power-clock phases φ and $\bar{\varphi}$.

The power-clock waveforms required by SBL can be generated using a clock generator circuit similar to the "blip" circuit in [47], as shown in Figure 3.2. This circuit is formed by connecting two RLC oscillators back-to-back, using the output wave-

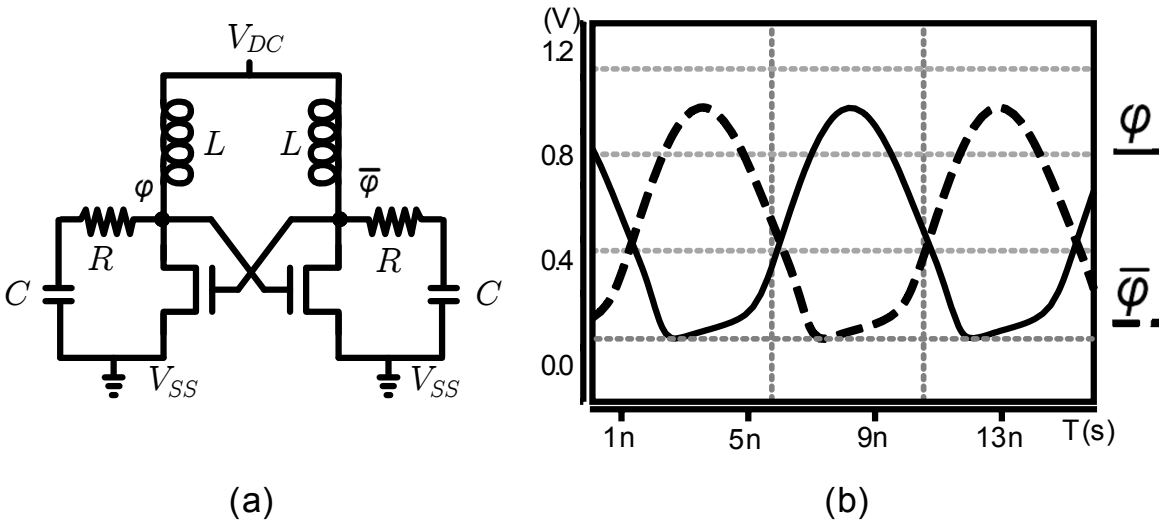


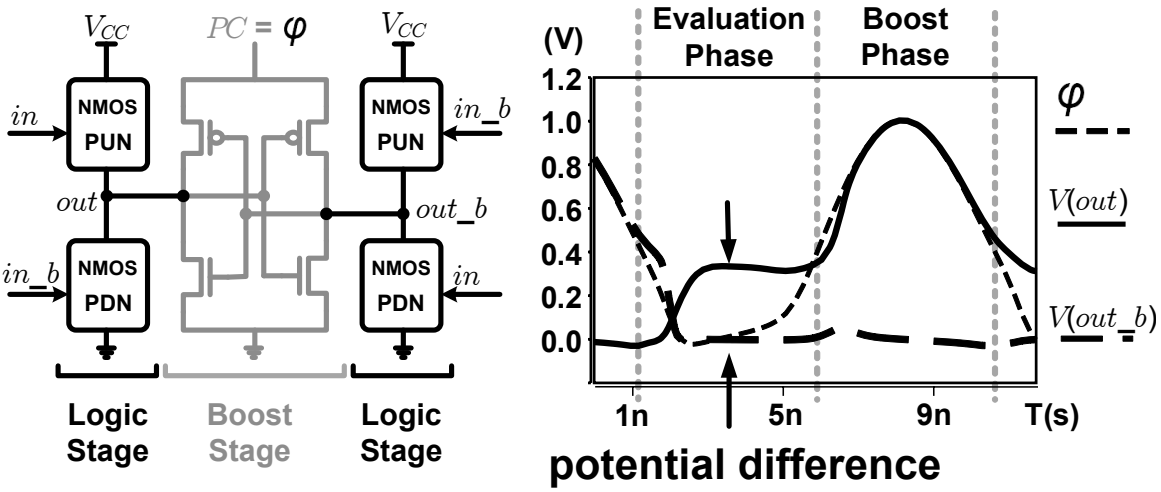
Figure 3.2: Simple "blip" clock generator: (a) Schematic. (b) Waveform.

form φ of one oscillator to drive the other, and vice versa. The two waveforms φ and $\bar{\varphi}$ are partially overlapping, since the NMOS devices are not fully on until their output voltages exceed the threshold voltage V_{th}. One of the advantages of this clock generator is that it can provide a larger clock amplitude beyond its DC supply V_{DC}. As a result, it can share the same supply with the SBL gates, allowing a single DC supply operation. In our FIR test-chip, we used a clock generator that is a distributed injection-locked version of this circuit and is described in Section 4.2.

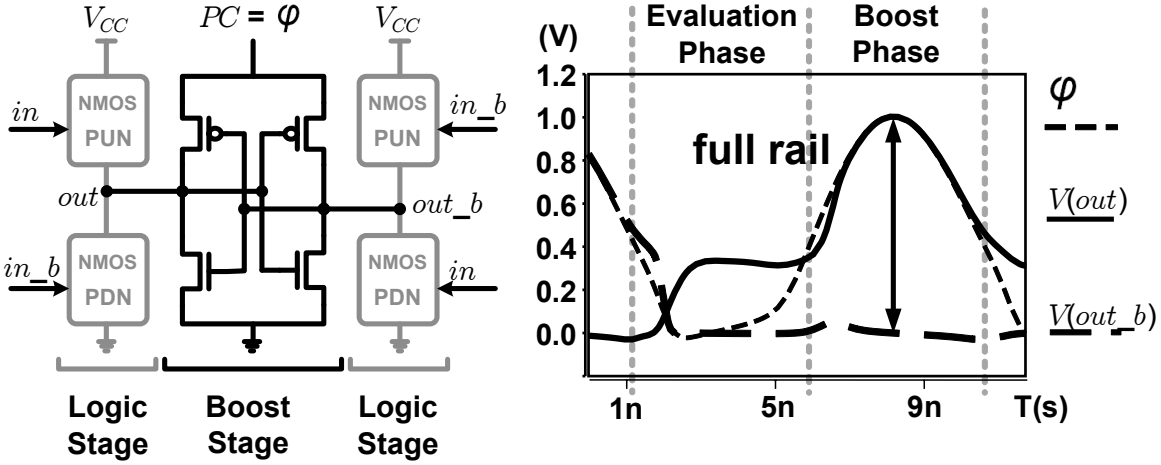
SBL improves upon Boost Logic [34], its closest charge-recovery logic relative, in a number of significant ways. Specifically, SBL can operate with a single DC supply, whereas Boost Logic requires three DC supply levels. Boost Logic uses two DC supply rails, V_{DD}' and V_{SS}', and develop a near $\frac{1}{3}V_{DD}$ potential difference at it output nodes. However, SBL shifts this voltage difference to V_{CC} and V_{SS} and only require a single supply V_{CC} in each gate. The third supply for Boost Logic is the supply of the clock generator, V_{DC}. For SBL, benefited by the blip generator, we explicitly share this supply with V_{CC} and achieve a single supply operation. Still, the energy efficiency of SBL improves when using different DC supply levels for logic and clock generation, as

demonstrated by the experimental results in Section 4.5. Moreover, the Logic stage in SBL is connected to ground, resulting in greater gate overdrive and thus higher performance than Boost Logic.

3.3 SBL Operation



(a)



(b)

Figure 3.3: SBL operation: (a) Evaluation Phase. (b) Boost Phase.

Each SBL gate operates in two phases, Evaluation and Boost, which are active

during mutually exclusive intervals. During Evaluation, as shown in Figure 3.3(a), the Boost stage is off, and there is no significant current flowing through any of the devices in the Boost stage, since the power-clock remains close to 0V. As φ transitions low, the drive strength of the Logic stage gradually weakens, since its inputs gradually ramp down. When its inputs reach the subthreshold supply level V_{CC}, the Logic stage is effectively off.

As the power-clock φ rises, as shown in Figure 3.3(b), the gate transitions into the Boost phase of its operation. During this phase, the Boost stage acts as an amplifier of the subthreshold voltage $V_{out} - V_{out\_b}$. The voltage V_{out} tracks φ, reaching approximately 1V as φ rises. As φ falls, the charge at the output node *out* is recovered by the power-clock φ, and the output voltage is brought back to approximately V_{th} levels. When φ falls below V_{th}, all transistors in the Boost stage are in cut-off, and the next logic evaluation phase begins. Throughout the Boost phase, the node *out_b* stays essentially at 0V.

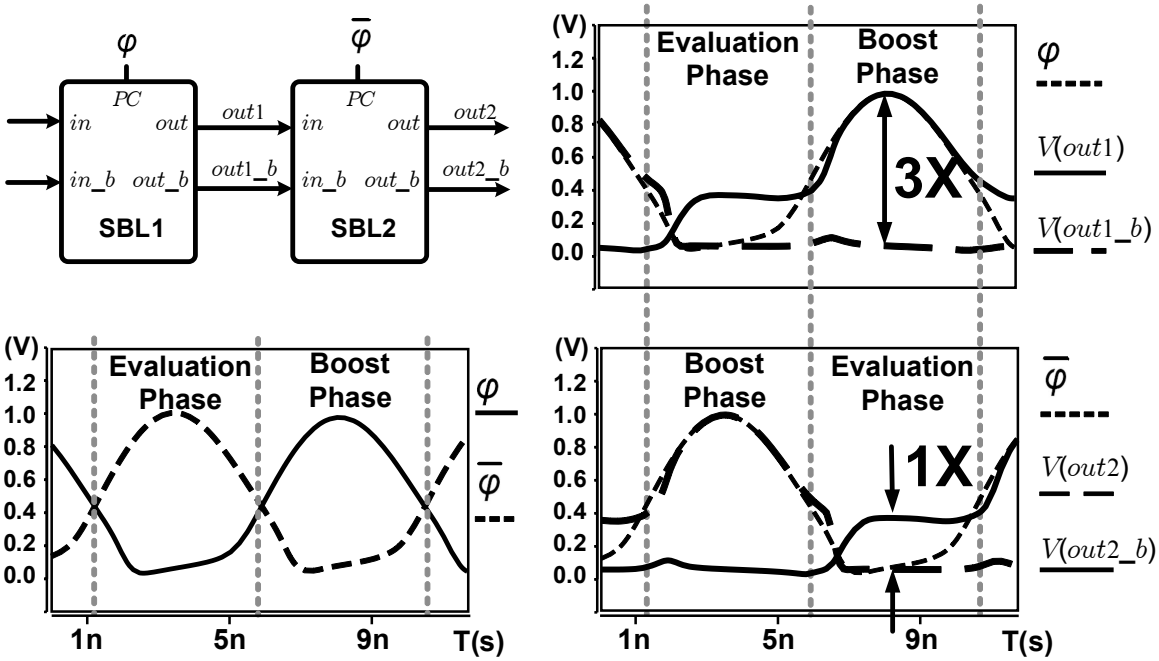


Figure 3.4: Cascade of SBL gates.

The graphs in Figure 3.4 show two cascaded SBL inverters, SBL1 and SBL2, and the waveforms at the output nodes $out2$ and $out2\_b$ with respect to the two power-clock waveforms φ and $\bar{\varphi}$. During Evaluation of the first SBL gate, φ remains effectively low, whereas $\bar{\varphi}$ transitions from low to high and then back to low. With their inputs boosted by the preceding SBL gate to be much higher than the supply voltage V_{CC}, the PUN and PDN network charge $out1$ to V_{CC} and discharge $out1\_b$ to V_{SS} in super-linear mode. Notice that even though PUNs are implemented in NMOS, the output node does not suffer a V_{th} drop when charged to V_{CC} , since the PUN inputs are boosted to a value significantly higher than V_{CC}. With their inputs boosted by the preceding SBL gate to be much higher than the supply voltage V_{CC}, PUN1 and PDN1 charge $out1$ to V_{CC} and discharge $out1\_b$ to V_{SS} in super-linear mode. Notice that even though PUNs are implemented in NMOS, the output node does not suffer a V_{th} drop when charged to V_{CC}, since the PUN inputs are boosted to be significantly higher than V_{CC}.

Due to the significant gate overdrive at the Logic stage, SBL can reach higher operating speeds than static CMOS operating with the same subthreshold supply. For example, when the Logic stage is evaluating, SBL can be designed so that the inputs to the Logic stage exceed 0.9V even with V_{CC} = 0.3V. Compared to static CMOS with a 0.3V supply level, the Logic stage has 3X the gate overdrive, allowing SBL implementations to operate at higher clock frequencies and drive larger output loads.

3.4 SBL Energetics

The energy consumed during each cycle in the operation of an SBL gate is given by the equation:

$$E_{SBL} = E_{Logic} + E_{Boost} + E_{Crowbar}, \tag{3.1}$$

where E_{Logic} and E_{Boost} denote the energy consumed in the two stages of SBL, and $E_{Crowbar}$ denotes the energy consumed by short-circuit currents during SBL operation.

The energy consumption of the Logic stage is given by the equation:

$$E_{Logic} = \frac{1}{2} C_L V_{CC}^2, \tag{3.2}$$

where C_L denotes the total switching capacitance at the SBL output. Compared to conventional switching, this energy consumption is significantly decreased due to the aggressively-scaled subthreshold supply level V_{CC}.

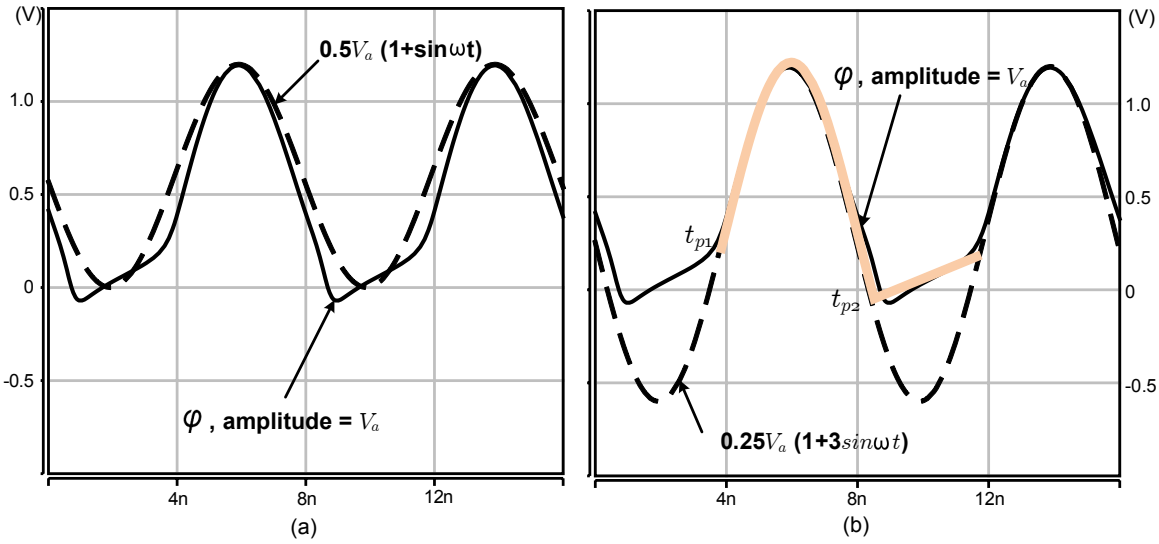


Figure 3.5: Clock waveform modeling: (a) Sine clock with equal peak-to-peak swing. (b) Sine clock with 1.5X peak-to-peak swing.

To derive an expression for E_{Boost}, we model the Boost stage as a simple RC series system with a "blip" voltage source that is modeled by two regions, sinusoidal and linear, as shown in Figure 3.5. Simulations suggest that in the sinusoidal region, a sinusoidal waveform with 1.5 times the peak-to-peak amplitude V_a of the clock waveform φ provides a good approximation. Moreover, in the linear region, they indicate that the clock waveform rises almost linearly to approximately 0.1V, independent of clock frequency and amplitude. Accordingly, the clock waveforms in the two regions

can be approximated as follows:

$$\varphi_{Sine} = 0.25V_a(1 + 3sin\omega t), \tag{3.3}$$

$$\varphi_{Linear} = 0.1 \cdot \frac{t}{T - |t_{P1} - t_{P2}|}, \tag{3.4}$$

where $\omega = 2\pi/T$, T is the period of the clock waveform φ, and t_{P1} and t_{P2} are the endpoints of the two regions, as shown in Figure 3.5.

Solving Equation (3.3) for 0.1V and 0V yields the following equation for the endpoints t_{P1} and t_{P2}, respectively, of the two regions:

$$t_{P1} = sin^{-1}\left[\frac{1}{3} \cdot (\frac{0.1}{0.25}V_a - 1)\right] \cdot \frac{1}{\omega}, \tag{3.5}$$

$$t_{P2} = sin^{-1}\left[\frac{-1}{3}\right] \cdot \frac{1}{\omega}. \tag{3.6}$$

The energy E_{Sine} consumed in the Boost stage of a SBL gate during operation in the sinusoidal region is given by integrating I^2R over time from t_{P2} to t_{P1}, where I is the AC component of the current resulting when φ_{Sine} drives the reactive load $1/j\omega C_B$, and R and C_B are the effective resistance and effective capacitance, respectively, when looking into the node PC of a SBL gate. (We assume that $R << 1/j\omega C_B$, as confirmed by our test-chip.) From Equation (3.3), we have:

$$\begin{aligned} I &= \left|\frac{V}{1/j\omega C_B}\right| \\ &= \left|\frac{0.25V_a \cdot 3sin\omega t}{1/j\omega C_B}\right|, \end{aligned} \tag{3.7}$$

and, therefore,

$$\begin{aligned} E_{sine} &= \int_{t_{P2}}^{t_{P1}} \left| \frac{0.75V_a \cdot sin\omega t}{1/j\omega C_B} \right|^2 Rdt \\ &= \frac{9V_a^2\omega^2 C_B^2 R}{32}(t - 2cos\omega t)|_{t=t_{P1}}^{t=t_{P2}} \\ &= \frac{9V_a^2\pi^2 C_B^2 R}{8T^2}(t - \frac{1}{\omega}cos2\omega t)|_{t=t_{P1}}^{t=t_{P2}} \\ &= \frac{K9V_a^2\pi^2 C_B^2 R}{8T}. \end{aligned} \tag{3.8}$$

Equation (3.8) has been simplified by including a coefficient K, $0.5 < K < 0.6$, which depends on the clock amplitude. Replacing the clock amplitude V_a by the effective voltage swing in the Boost stage, $V_a - V_{CC}$, we obtain

$$E_{Sine} = \frac{9K(V_a - V_{CC}^2)^2\pi^2 C_B^2 R}{16T}. \tag{3.9}$$

The energy E_{Linear} consumed in the Boost stage of a SBL gate during the linear region of the clock waveform is given by integrating IR over time, where I is derived from Equation (3.4):

$$\begin{aligned} E_{Linear} &= \int_0^{T-|t_{P1}-t_{p2}|} I^2 Rdt \\ &= \int_0^{T-|t_{P1}-t_{P2}|} \left| C_B \frac{dV}{dt} \right|^2 Rdt \\ &= \int_0^{T-|t_{P1}-t_{p2}|} \left| C_B \frac{d(\frac{0.1}{T\cdot|t_{P1}-t_{P2}|}t)}{dt} \right|^2 Rdt \\ &= \frac{0.01 \cdot C_B^2 R}{T - |t_{P1} - t_{P2}|}. \end{aligned} \tag{3.10}$$

From Equations (3.9) and (3.10), it follows that $E_{Linear}/E_{Sine} < 1\%$, and therefore the total energy consumption in the Boost stage can be approximated by E_{Sine}. From

Equations (3.1), (3.2), and (3.9), it follows that the total energy consumption of a SBL gate during a cycle is given by:

$$E_{SBL} = \frac{1}{2} C_L V_{CC}^2 + \frac{9K(V_a - V_{CC})^2 \pi^2 C_L^2 R}{16T} + E_{Crowbar}. \tag{3.11}$$

Note that C_B is replaced by C_L, since the Logic stage and the Boost stage are driving the same output loads. Based on Spice simulation results of SBL test-chip presented in Section 4.4, the effective resistance and capacitance seen from each clock phase are about 0.6 Ω and 57pF, respectively, confirming that $R << 1/j\omega C_B$.

The crowbar $E_{Crowbar}$ in Equation (3.11) has three components: $E_{V_{CC}-V_{SS}}$, $E_{V_{CC}-PC}$, and $E_{PC-V_{SS}}$. The energy $E_{V_{CC}-V_{SS}}$ is associated with the Logic stage. Specifically, due to the relatively slow rise time of the input waveform, short current will flow from $V_{\rm CC}$ to $V_{\rm SS}$ during the evaluation phase. This component dominates $E_{Crowbar}$. At very low operating frequencies, it also dominates the total energy consumption E_{Total}, as we discuss the experimental results presented in Section 4.4. The energy $E_{V_{CC}-PC}$ is consumed during the Evaluation phase. As $V_{\rm CC}$ charges one of the output nodes, current flows from $V_{\rm CC}$ to the PC pin through the PMOS device in the Boost stage. Since $V_{\rm CC}$ is always at a subthreshold voltage level, this component is relatively small compared to $E_{V_{CC}-V_{SS}}$. The energy $E_{PC-V_{SS}}$ is consumed during the Boost phase. As φ rises, although the Logic stage is turned off, current still flows from the PC pin to $V_{\rm SS}$ through the evaluation NMOS. Similar to $E_{V_{CC}-PC}$, this component is significant only at very low operating frequencies.

Equation (3.11) provides guidance for device sizing and illustrates some of the energy trade-offs between E_{Logic}, E_{Boost}, and $E_{Crowbar}$ For example, in the Boost stage, up-sizing the PMOS devices reduces the effective resistance R, but increases the effective capacitance C_L. In the Logic stage, up-sizing the evaluation pull-up and pull-down networks yields a greater potential difference at the output nodes by

the end of the evaluation period, resulting in higher energy efficiency during the Boost stage. At low operating frequencies, however, such up-sized networks result in increased $E_{V_{CC}-V_{SS}}$.

3.5 Summary

SBL improves upon its closest charge-recovery logic, Boost Logic, in many significant ways. Specifically, SBL can operate with a single DC supply, whereas Boost Logic requires three DC supply levels. Moreover, SBL achieves greater gate overdrive and higher performance than Boost Logic through it supply configuration.

Compared to subthreshold logic, SBL accomplishes significant performance improvements through device overdriving. The NMOS-only PUN and PDN in the Logic stage are driven with inputs of approximately 1V, allowing SBL to operate at clock frequencies in the hundreds of MHz or, alternatively, to realize functions of significant complexity within a single clock cycle. In addition to enhanced performance, gate overdriving leads to improved variation tolerance. All transistors in the Logic stage conduct in super-threshold linear mode, and delay does not vary significantly with variations in the subthreshold supply V_{CC} or V_{th}.

CHAPTER 4

187MHz Charge-Recovery FIR Filter with Subthreshold Boost Logic

To demonstrate the fast and energy-efficient operation of SBL, we used it in the implementation of a transpose FIR filter. The relatively state-intensive nature of the transpose-type FIR filter, coupled with the relatively simple computation that is performed between state elements make it a natural fit for SBL. The latency of the SBL FIR is only 2 cycles longer than that of a similar-performance static CMOS design. This SBL FIR is the first design that demonstrates a near 200MHz operating frequency with a single 0.36V supply level.

The remainder of this chapter is organized as follows: Section 4.1 presents the architecture and SBL implementation of the FIR test-chip. Section 4.2 presents the two-phase clock generator the clock distribution network used in the SBL FIR. Section 4.3 describes the semi-custom design methodology used for the SBL circuitry in the FIR. Section 4.4 presents results from device-level simulations of the SBL FIR and its static CMOS counterpart with the same architecture. In Section 4.5, we present measurement results from our SBL FIR test-chip. Conclusions are given in Section 4.6.

4.1 FIR Filter Architecture

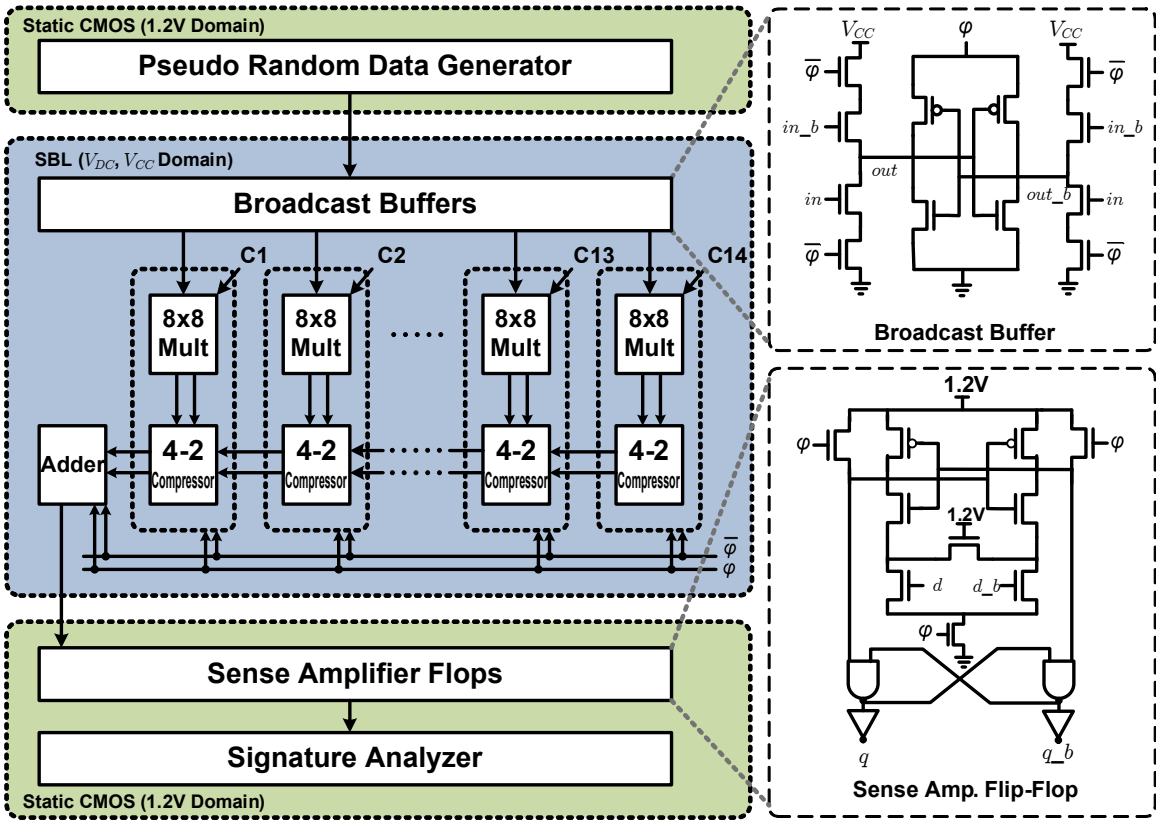


Figure 4.1: Block diagram of SBL FIR filter and BIST circuits.

We used SBL to design an 8-bit 14-tap FIR filter. A block diagram of the FIR chip is given in Figure 4.1. A static CMOS built-in self-test (BIST) circuit is used to generate and process the FIR input and output. The pseudo-random input sequence generated by BIST is broadcast to 14 modified 8×8 Booth multipliers. The products of these inputs with the 14 FIR coefficients are accumulated through 14 4-to-2 compressors. The final result is obtained from a hybrid adder, and then sent to a signature analyzer, generating a signature vector. 8×8 multipliers take 1.5 cycles to generate sum and carry vector pairs and each tap takes 1 cycle to merge the sum and carry vector pairs from the previous tap and from its 8×8 multiplier. The vector pairs are then merged in a 20-bit hybrid carry-look-ahead/carry-select adder with

2 cycles of latency. The longest path through the SBL-based FIR has a latency of 19 cycles, including 0.5 cycle latency of the broadcast buffers. Compared to static CMOS design with the same architecture, the latency overhead of the SBL FIR is 2 cycles: 1 cycle in the 8x8 multiplier, and 1 cycle in the 20-bit adder.

To enable SBL to communicate with the static CMOS BIST logic, two interface blocks are inserted before and after the FIR. On the FIR input side, broadcast buffers implemented in SBL are used to connect static CMOS signals from the BIST circuitry into dual-rail sinusoidal-like signals for the SBL datapath. On the output side of the FIR, sense-amplifier flip-flops that operate off the same clock as SBL gates latch the SBL signals from the FIR and make them available to the static CMOS signature analyzer.

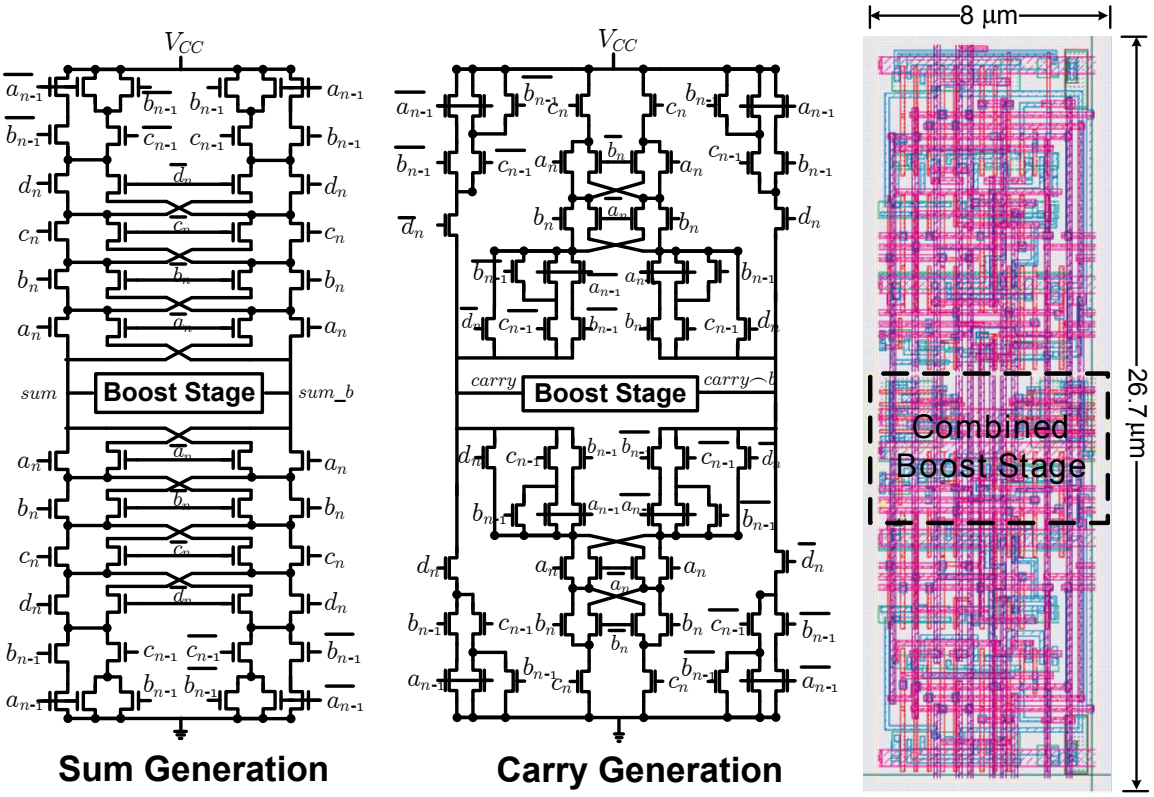


Figure 4.2: Schematic and layout of a 4-2 compressor.

Gate overdrive at the Logic stage of SBL gates allows the implementation of

functions with significant complexity within a single clock cycle. Figure 4.2 shows schematics and layout of the SBL-based 4-to-2 compressor used in the FIR. Designed in a 0.13μm bulk CMOS process Due to the dual-rail nature of the SBL gates, the SBL 4-to-2 compressor has 2.1X area overhead compared to a standard-cell implementation. Each SBL gate has a maximum transistor stack height of six and, as discussed in more detail in the experimental results presented in Section 4.5, can operate at 187MHz with $V_{CC} = 0.36$V.

4.2 Power-Clock Generator and Clock Network Design

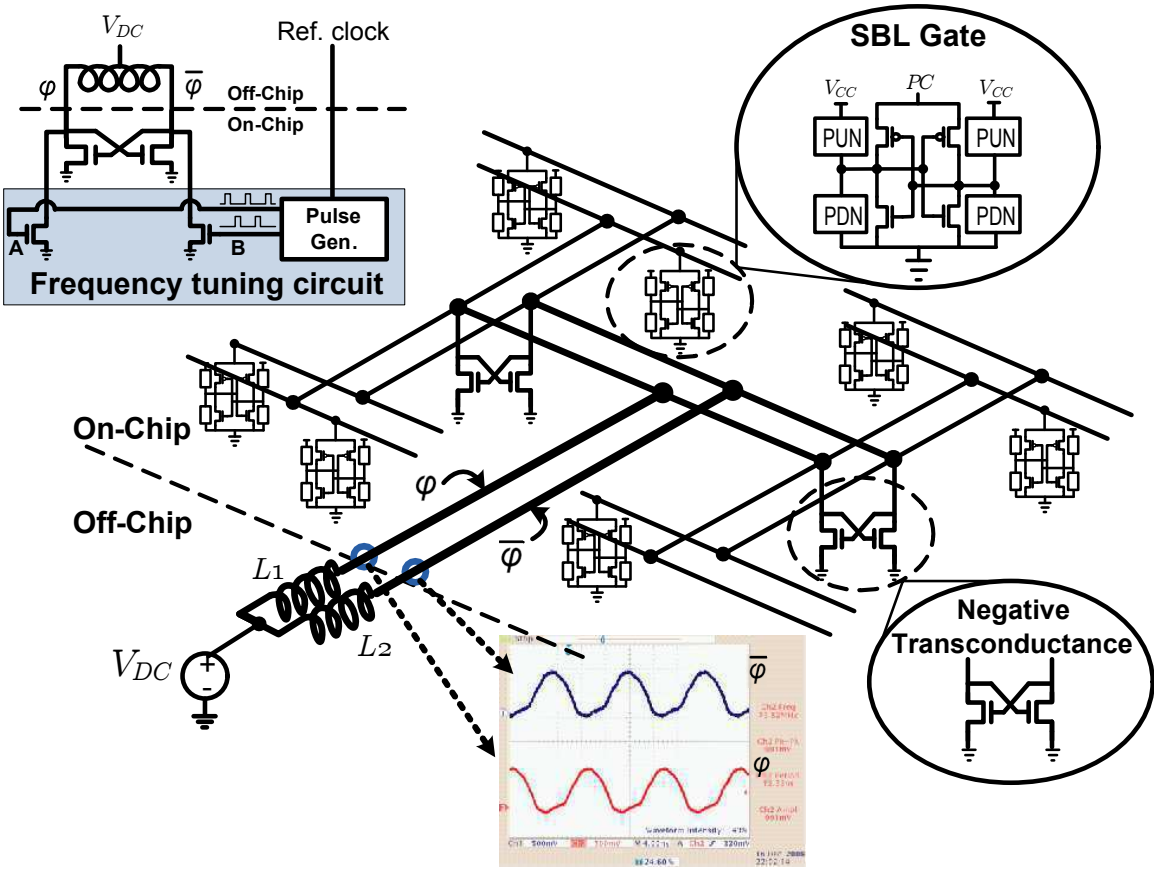


Figure 4.3: Distributed "blip" clock generator and measured clock waveform.

The SBL FIR uses two power-clock waveforms φ and $\bar{\varphi}$ that are generated by the clock circuit shown in Figure 4.3. The two-phase power-clocks are first distributed

through a H-tree structure made with M5-M8 and then connect to the SBL gates with a local clock mesh made with M2-M4. This local clock mesh has a rectangular distribution area of 792.8μm×346.8μm with 24 pairs of 0.6μm M2, 20 pairs of 0.6μm M3, and 65 pairs of 1.2μm M4 strips. In this clock generator, the basic "blip" generator circuit, introduced in Section 3.2, has been augmented to include a pair of weak drivers at the root of the tree that allow for the power-clock waveforms to be injection-locked to a target clock frequency. These drivers are pulsed by reference signals A and B that are generated by an on-chip pulse generator. In our test-chip, the drivers are 150μm wide and can tune the operating frequency by as much as ±3% off resonance. The tuning range can be increased by sizing up the injection-locked devices. To maintain the oscillation, fourteen pairs of cross-coupled NMOS switches with 2400μm total active width are distributed throughout a hierarchical two-phase distribution network, similar to [14]. Two off-chip inductors are used to resonate the parasitic capacitance of the clock distribution network and the SBL gates. In our test-chip, the load on each phase of the power-clock is approximately 57pF, as derived from layout extraction.

The clock circuitry is powered by a DC supply V_{DC} that can be controlled independently of the supply V_{CC} for the SBL gates. The level of V_{DC} determines the amount of energy re-introduced into the clock network each cycle, thus affecting the amplitude of the power-clock waveforms and controlling the overdrive level at the Logic stages. Although not required for correct operation, the independent control of V_{DC} and V_{CC} allows for increased energy efficiency. Specifically, by decreasing V_{CC} to limit crowbar current through the Logic stage while keeping V_{DC} sufficiently high to ensure the requisite overdrive, energy efficiency can be improved without sacrificing performance. As shown in Section 4.5, the FIR achieves energy-efficient operation with $V_{DC} = V_{CC}$, but its energy consumption per cycle decreases further by 17.1% when V_{DC} and V_{CC} are set to different subthreshold values.

4.3 SBL FIR Filter Design Methodology

The SBL FIR filter was designed using a semi-custom design flow that enables the use of industrial static timing tools for timing analysis, as well as the use of HDL languages for system-level functionality verification. The standard cell library used for the SBL filter consists of 65 different cells, most of which are simplified versions of a 4-to-2 compressor or a 3-to-2 compressor that deal with special cases.

The basic design flow is shown in Figure 4.4. Initially, a Spice-level characterization is performed on every standard cell. For each cell, pin capacitance is extracted, and a delay-to-capacitance matrix is generated by sweeping a range of output loads from 25fF to 75fF. Every SBL gate is modeled as a timing element, and the $D-Q$ delay is defined as the time that the gate needs to develop a V_{th}-level voltage difference at it output nodes. A HDL model for each cell is also established in Verilog and is used for the system-level functionality verification when system-level design is completed.

The next step in the design flow is the manual pipelining, placement, and routing of the cells. During this step, the design is optimized for minimum area and energy consumption. the power-clock pins of all cells are aligned, allowing for significant reduction in the resistance of the clock distribution network and, therefore, improvement in the efficiency of energy recovery.

After completion of layout, wire capacitance is extracted for all internal nets. Based on the extracted parasitics and the previously built timing models of the cells, a commercial static timing tool is used for post-layout timing analysis. From timing analysis results, we fix all timing violations by replacing the original cells with the ones that have larger driving ability and then repeat this place-and-route and timing check process to ensure timing closure. This process can also be used to separate out the critical paths in a design and improve the maximum operating frequency of a system.

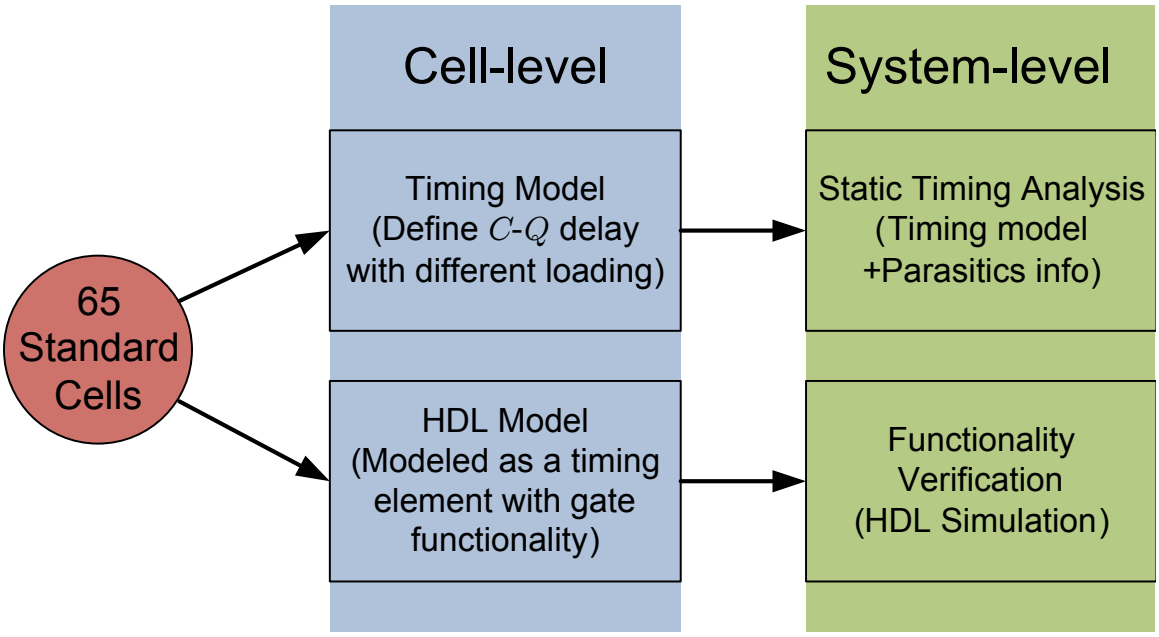


Figure 4.4: SBL design flow.

4.4 SBL FIR Filter Spice-Level Analysis

In this Section, we present results from Spice-level simulations of our SBL test-chip. For comparison purposes, we also present Spice-level simulation results of a conventional static CMOS version of the FIR, which was obtained by performing automatic synthesis, placement, and routing of the same FIR architecture that we used to derive the SBL FIR test-chip. Measurement results from our SBL FIR test-chip, along with a comparison of simulation and measurement results are given in Section 4.5.

4.4.1 Spice Simulation Results

Figure 4.5 gives a graph of energy consumption per cycle versus operating frequency for our SBL FIR design. This graph was obtained using Synopsis Hsim with the BSIM model on a netlist of our SBL FIR that was obtained from layout extraction. All data points were obtained with the minimum supply setting $V_{CC} = V_{DC}$ that

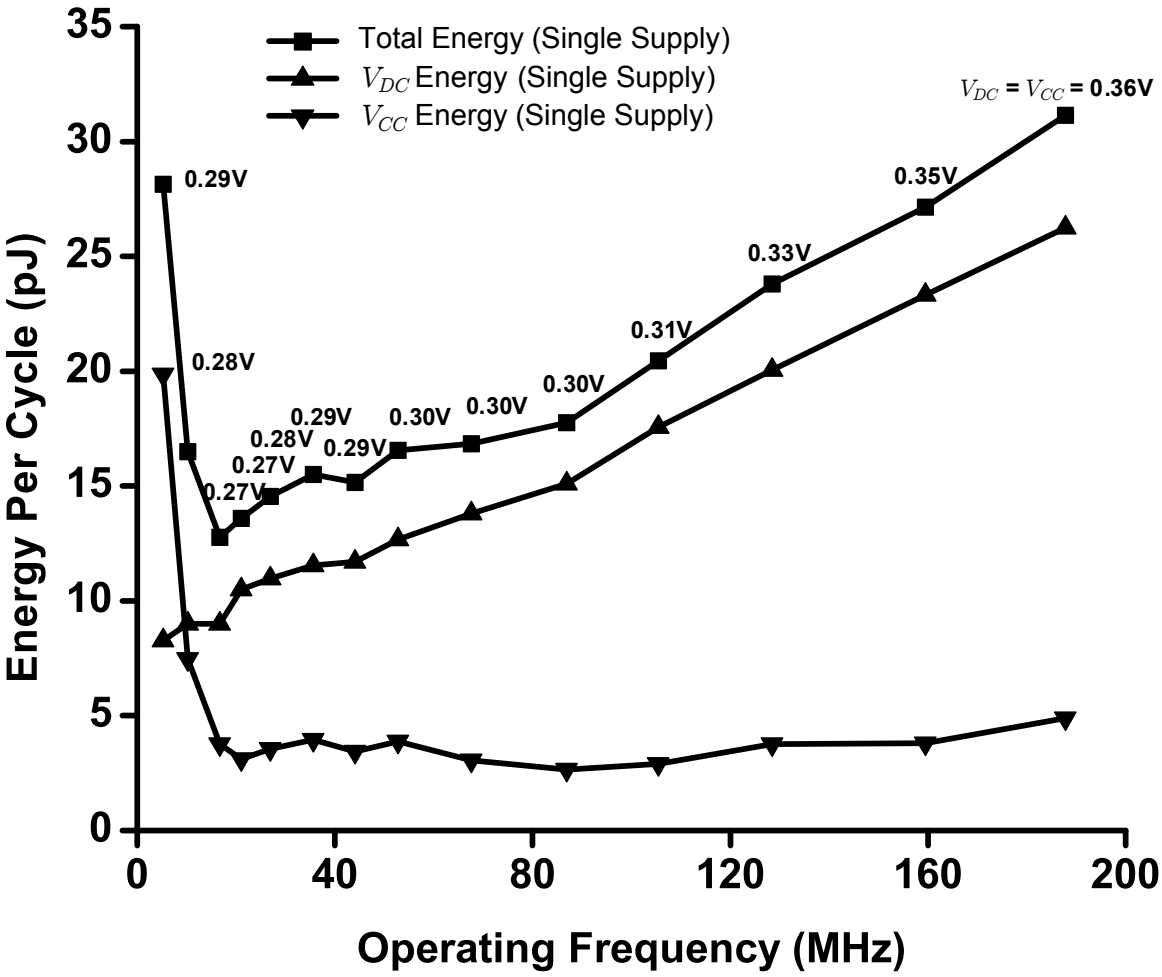


Figure 4.5: Simulated energy consumption of SBL FIR filter.

yielded correct operation at the corresponding operating frequency. The simulation result shows that the SBL FIR achieves a near 200MHz operating frequency with a single subthreshold level supply of 0.36V.

Notice that energy consumption is dominated by the component related to the power-clock generator, which corresponds to the power supply V_{DC}. Moreover, notice that at frequencies below 20MHz, the energy consumption of the Logic stage, which corresponds to the power supply V_{CC}, starts rising at an increasing rate, due to the increasing crowbar current from V_{CC} to V_{SS} caused by the slowly transitioning inputs of the Logic stage. Consequently, total energy consumption for the SBL FIR starts

increasing at operating frequencies below 17MHz.

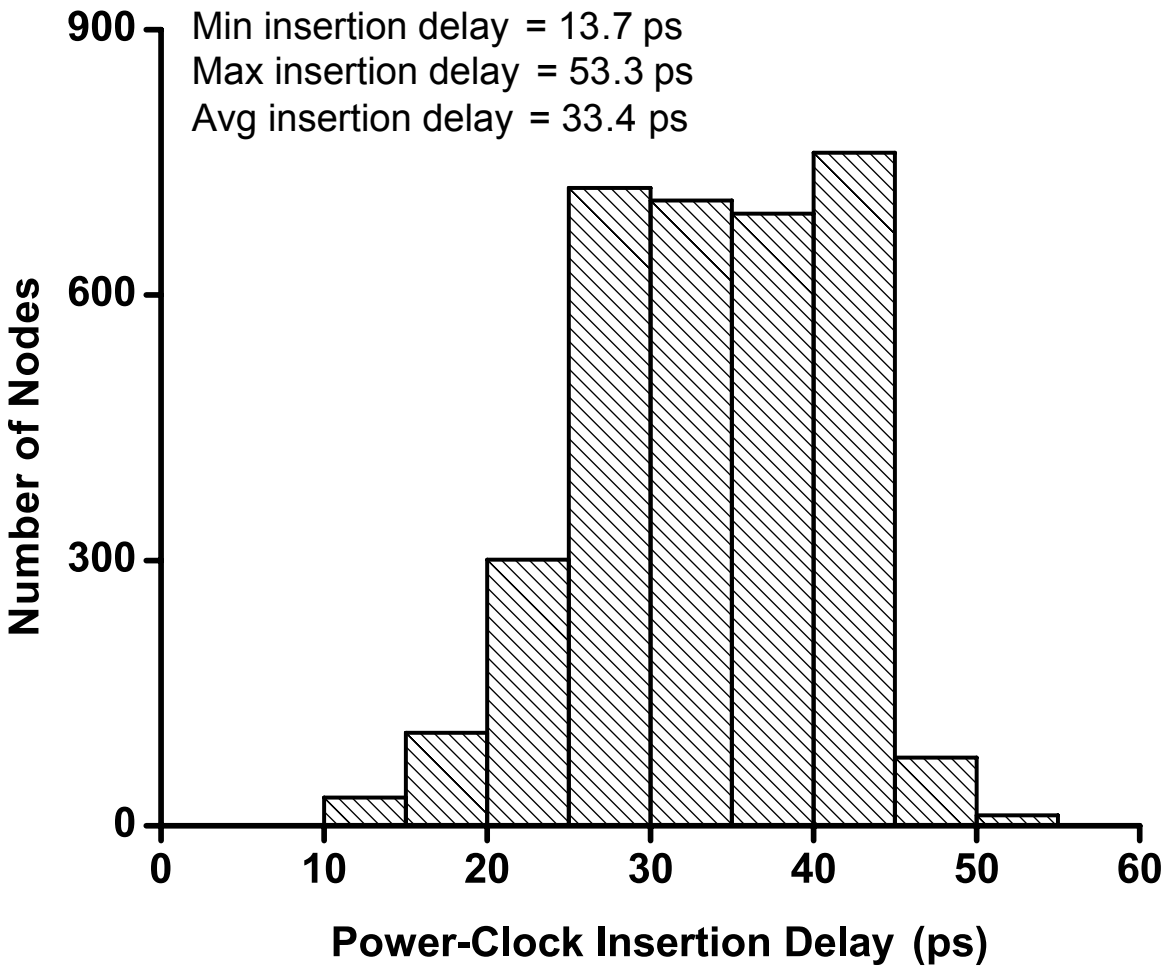


Figure 4.6: Histogram of simulated power-clock insertion delays at a resonant frequency of 53.7MHz.

Clock skew is introduced due to load variation across the chip. Figure 4.6 shows power-clock insertion delay data obtained from Spice-level simulations of the entire chip with extracted resistance, capacitance, and coupling capacitance. At the resonant frequency of 53.7MHz, the maximum and the minimum insertion delay of the power-clock is 53.3ps and 13.7ps, respectively. It follows that the difference of these two numbers yields a maximum possible skew of 39.6ps.

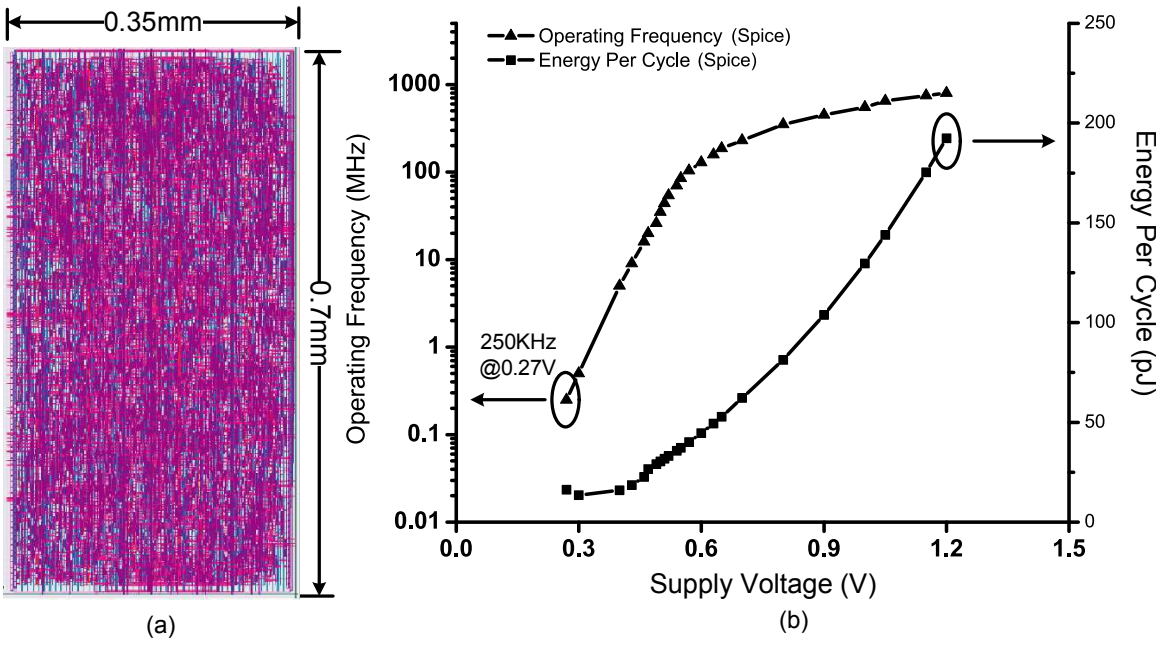


Figure 4.7: (a) Layout of conventional CMOS FIR. (b) Simulated operating frequency and energy per cycle vs. supply voltage for conventional CMOS FIR filter.

4.4.2 Performance Comparison with CMOS FIR Filter

To compare our SBL FIR with conventional CMOS design, we synthesized a standard-cell version of the same 19-cycle FIR architecture that we used to derive the SBL design in the same 0.13μm technology. Synthesis was performed by Synopsys Design Compiler, yielding a conventional FIR with the same latency as the SBL FIR. Placement and routing were performed in a fully automatic manner using Cadence SoC Encounter with 80% area utilization and a synthesized clock tree. The layout of the resulting design is shown in Figure 4.7(a). With a 0.35mm $\times$ 0.7mm footprint, the synthesized FIR occupies approximately 12.5% less area than its SBL counterpart.

Figure 4.7(b) gives Spice-derived graphs for the operating frequency and the per-cycle energy consumption of the static CMOS FIR as a function of the supply voltage. With 83% of its cells sized at X1 or X2 drive strength, this FIR achieves a clock frequency close to 800MHz with a nominal 1.2V supply. As expected, energy con-

sumption per cycle varies quadratically with supply voltage. Furthermore, operating frequency deteriorates exponentially fast, as supply voltage drops below 0.6V, barely exceeding 250KHz when the supply is set at 0.3V.

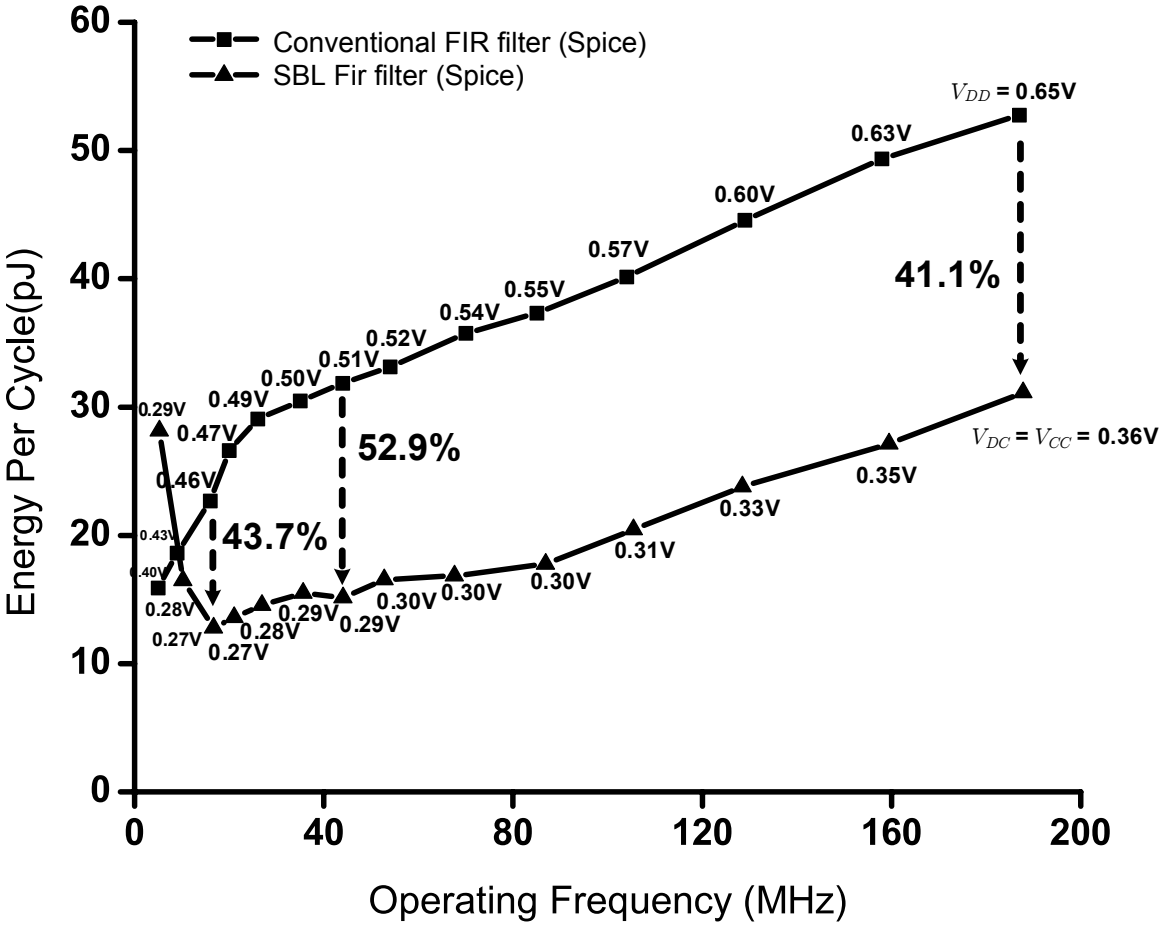


Figure 4.8: Simulated energy consumption of conventional and SBL FIR filters.

For both the SBL and the conventional FIR, simulated per-cycle energy consumption versus operating frequency is given in Figure 4.8. For simplicity, the SBL FIR in this graph uses a single DC supply level, connecting to both V_{DC} and V_{CC}. In the frequency range from 17MHz to 187MHz, the SBL FIR achieves 40% to 50% lower energy consumption than its conventional counterpart. The SBL design yields minimum energy consumption at 17MHz, achieving 43.7% reduction over its conventional counterpart. The maximum relative energy reduction of 52.9% is achieved at 44MHz. At

187MHz, the maximum clock frequency at which the SBL design functions correctly, relative energy savings over the conventional FIR are 41.1%.

4.5 SBL FIR Filter Test-Chip Measurement

This section gives measurement results from the experimental evaluation of the SBL FIR test-chip, validating its energy-efficient operation with subthreshold supplies at clock frequencies up to 187MHz. It also presents a comparison of measurement and simulation results, showing good agreement between the two, with relative discrepancy between measurements and simulations staying within 12% for operating frequencies ranging from 20MHz to 187MHz.

Technology	**0.13 μm 8M CMOS (RVT)**	
Threshold Voltage	**NMOS: 0.40V** **PMOS: -0.42V**	
Taps, In / Coeff Bits / Out	**14, 8 / 8 / 20**	
Total Transistors Count (including BIST)	**PMOS: ~8000** **NMOS: ~33000**	
Total Area (including BIST)	**0.38 mm^2**	
Effective Cap. per Clock Phase	**~57pF**	
BIST Supply Voltage	**1.2V**	
Switching Activity	**0.5**	
Measured Frequency Range	**5MHz – 187MHz**	
	Single Supply Setting:	**Two Supplies Seeting:**
Supply Voltage	**V_{DC}=V_{CC}=0.27V @ 20MHz** **V_{DC}=V_{CC}=0.36V @ 187MHz**	**V_{DC}=0.27V, V_{CC}=0.18V @ 20MHz** **V_{DC}=0.36V, V_{CC}=0.28V @ 187MHz**
Energy per Cycle	**15.57 pJ @ 20MHz** **31.64 pJ @ 187MHz**	**12.90 pJ @ 20MHz** **30.88 pJ @ 187MHz**
Figure of Merit (nW/MHz/Tap/In-Bit/Coeff-Bit)	**17.37 @ 20MHz** **35.31 @ 187MHz**	**14.40 @ 20MHz** **34.47 @ 187MHz**

Table 4.1: SBL FIR filter statistics and performance measurements.

Two sets of measurements were obtained. In the first set, the supplies V_{CC} and V_{DC} were set equal to each other. In the second set, the two supplies were controlled

independently. As shown in the table of Table 4.1, for both sets of measurements, the FIR test-chip achieves a maximum operating frequency of 187MHz with all supplies set at levels below $V_{th,nmos} = 400$mV. With the two supply values tuned independently, the test-chip achieves higher energy efficiency than with a single-supply setting.

4.5.1 Single-Supply Configuration

Figure 4.9 shows the per-cycle energy consumption of our test-chip for operating frequencies ranging from 5MHz to 187MHz. Data points are given for both single-supply and dual-supply settings. At each frequency point, the energy drawn from each supply is given separately, along with the total energy consumed. The different operating points are obtained by selecting off-chip inductors that yield a resonant frequency at that clock frequency. In all cases, the off-chip inductors were 0612 discrete devices that were mounted on the printed circuit board in proximity to the test-chip. The maximum operating frequency of 187MHz was obtained with no external inductors, with only the bondwires and package traces related to the clock generator providing all the parasitic inductance.

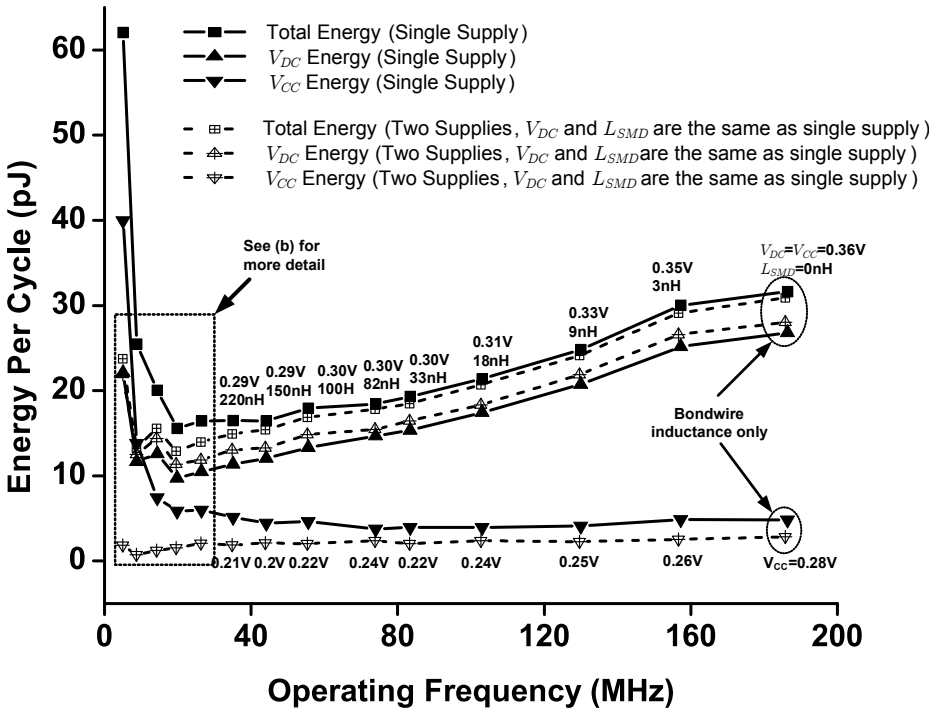
For each single-supply data point in Figure 4.9, the corresponding voltage and inductor value are given above the data point. The data show that energy consumption is dominated by the energy drawn from the clock generator, with V_{DC} accounting for more than 80% of total energy consumption. As operating frequency decreases from the maximum operating point of 187MHz, energy consumption decreases approximately linearly. The minimum energy point of 15.57pJ per cycle is obtained at 20MHz with $V_{CC} = V_{DC} = 0.27$V and two off-chip inductors of 680nH each. At this frequency, the recovery rate of the energy supplied through V_{DC} is approximately 89%, yielding a 17.37 nW/MHz/Tap/InBit/CoeffBit figure of merit. As operating frequency decreases below 20MHz, total energy consumption increases at an accelerating rate, due to increasing crowbar currents, with V_{CC} to V_{SS} crowbar currents in

the Logic stage quickly dominating, as evidenced by the cut-out that zooms on data in the 5MHz to 30MHz range.

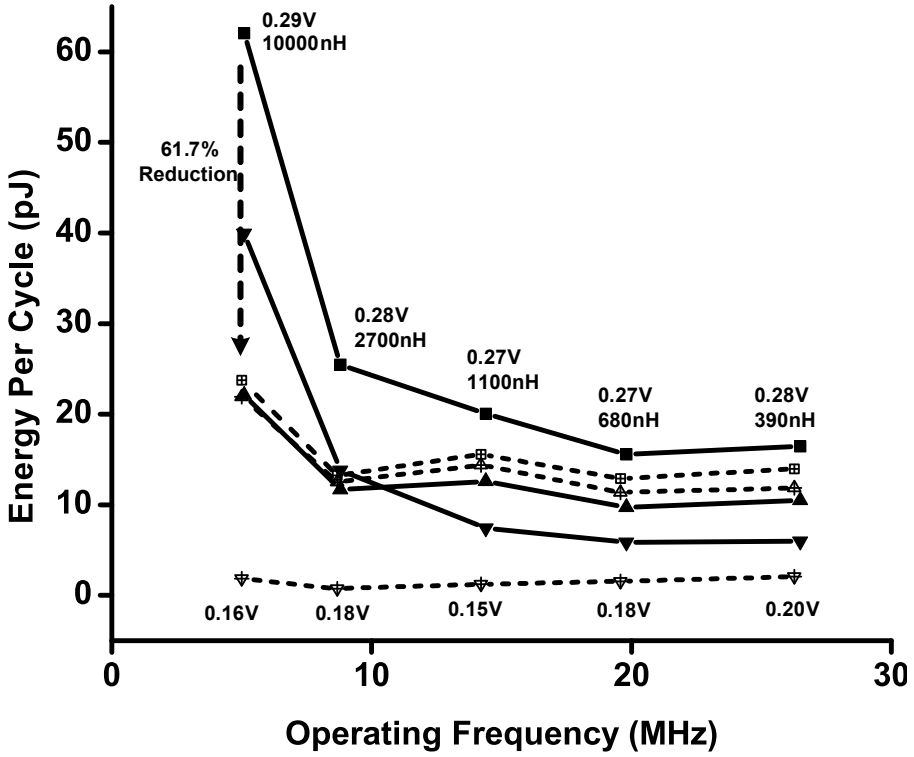
Beyond energy efficiency and performance, another question addressed by our experimental evaluation is the accuracy of the Spice simulation results presented in Section 4.4. Figure 4.10 gives simulation results under the conditions used to obtain measurements with a single supply. For operating frequencies in the 20MHz to 187MHz range, the discrepancy between simulations and measurements stays within 12%. At operating frequencies below 20MHz, the energy consumption of the Boost stage starts increasing. This increase is not reflected to the same extent in the simulations. With voltage supply below 0.27V, we conjecture that the increasing discrepancy between simulations and measurements is due to increasing model inaccuracies, due to the aggressively scaled voltage supply.

4.5.2 Two-Supply Configuration

The two-supply data points in Figure 4.9 have been obtained by keeping the same V_{DC} and inductor values as in the single-supply case, and by decreasing V_{CC} by as much as possible while still achieving correct function. The overall trends observed are similar to the single-supply case. With V_{CC} reduced, the energy drawn from V_{DC} increases, since the power-clock draws more energy to boost the smaller potential difference at the output of the Logic stage. As expected, however, energy consumption in the Logic stage is significantly decreased. The impact of reducing V_{CC} is particularly pronounced as operating frequencies decrease below 30MHz. Specifically, unlike the single-supply case where V_{CC}-related consumption starts increasing rapidly due to crowbar currents, with two separate supplies the energy consumption in the Logic stage remains relatively flat, even at frequencies as low as 5MHz. Notice that at 5MHz, where the crowbar current dominates, by separating V_{DC} and V_{CC}, we can reduce the energy consumption by 61.7%. The minimum energy point is obtained at 20MHz with



(a)



(b)

Figure 4.9: Measured energy consumption vs. operating frequency for SBL FIR filter (single supply and two supplies).

$V_{CC} = 0.18$V, yielding a figure of merit equal to 14.4 nW/MHz/Tap/InBit/CoeffBit, a 17.1% improvement over the single-supply case. At this frequency, the recovery rate of the energy supplied through V_{DC} is approximately 86%.

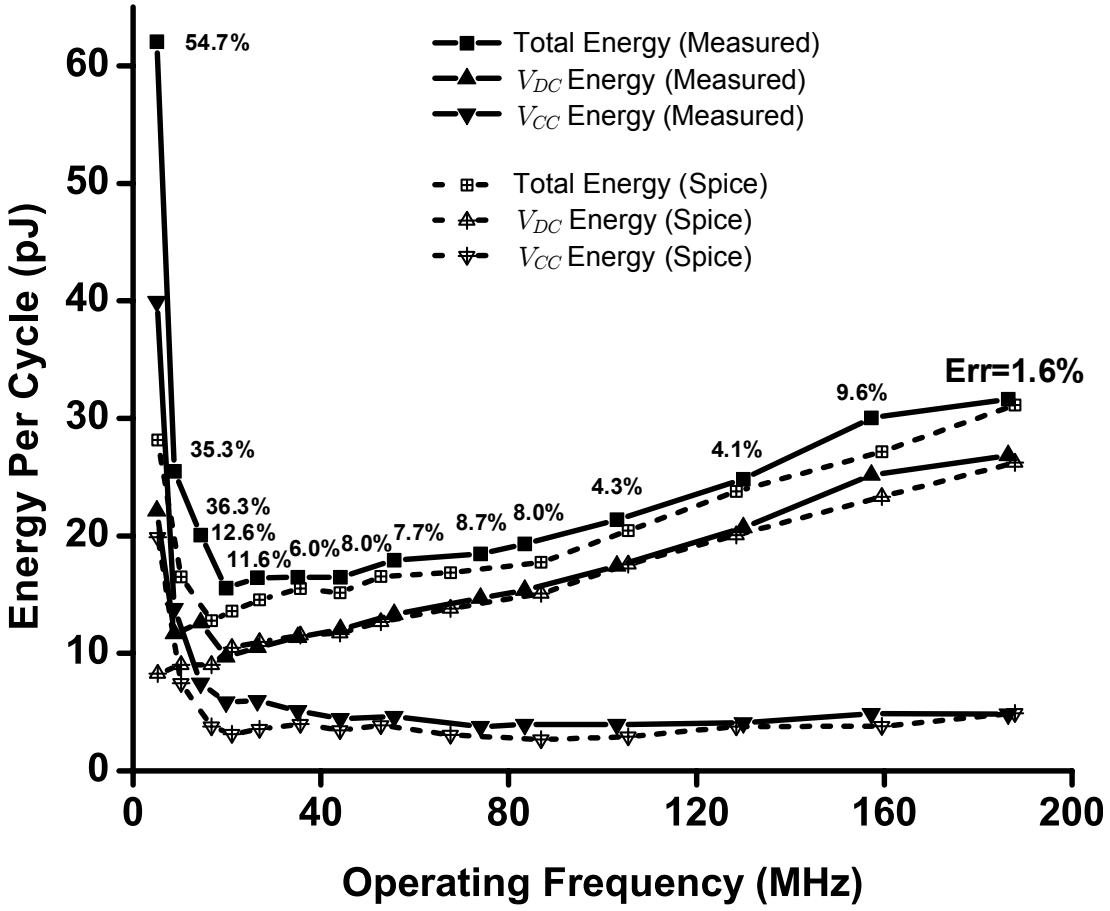


Figure 4.10: Comparison of measured and simulated energy consumption for SBL FIR filter (single supply).

4.5.3 Energy Trade-off and Robustness Analysis

Figure 4.11 gives a more detailed view of the trade-off between V_{CC}- and V_{DC}-related energy consumption. The rightmost data points inside the oval on the right-hand side give the energy consumption when a single supply is applied. By decreasing V_{CC}, V_{CC} energy decreases as expected, and V_{DC} energy increases gradually. Minimum total energy is obtained at $V_{CC} = 0.19$V. When V_{CC} decreases below 0.19V,

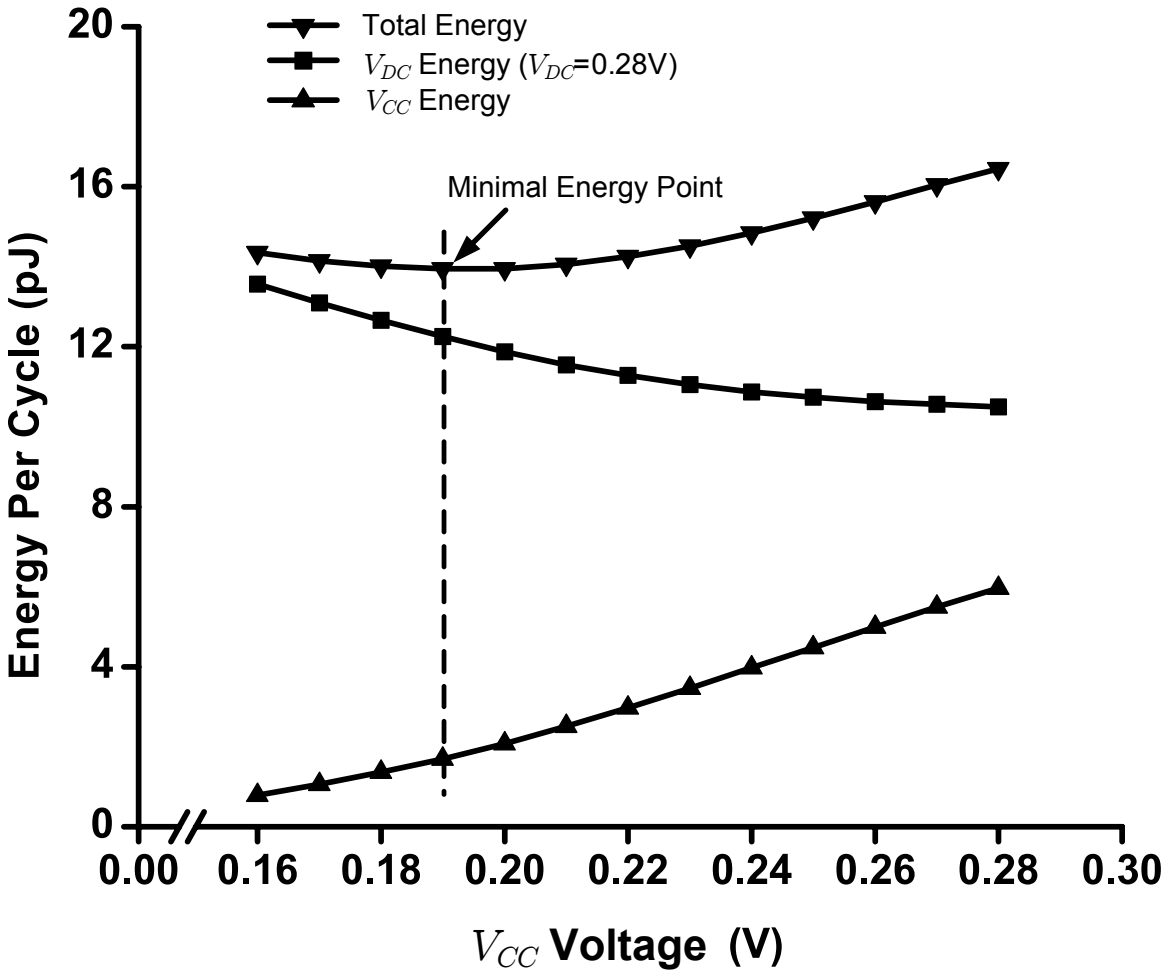


Figure 4.11: Measured total energy consumption vs. V_{CC} for the SBL FIR when operating at 26.4MHz with $V_{DC} = 0.28$V.

total energy consumption increases due to larger V_{DC}-related energy.

Another focus of our experiments was to determine the variability of resonant frequency across multiple test-chips. Figure 4.12 shows the resonant frequencies of 10 test-chips when running free with $V_{DC} = 0.36$V and fixed 3nH surface-mount inductors. Correct function has been validated for all 10 chips, with average resonant frequency $\mu = 160.33$MHz and standard deviation $\sigma = 0.83$MHz. The resonant frequency of these chips varies by $\pm 1\%$. Even with 3σ variation of 1.4MHz, it is still within the $\pm 3\%$ tuning range of the clock generator circuit.

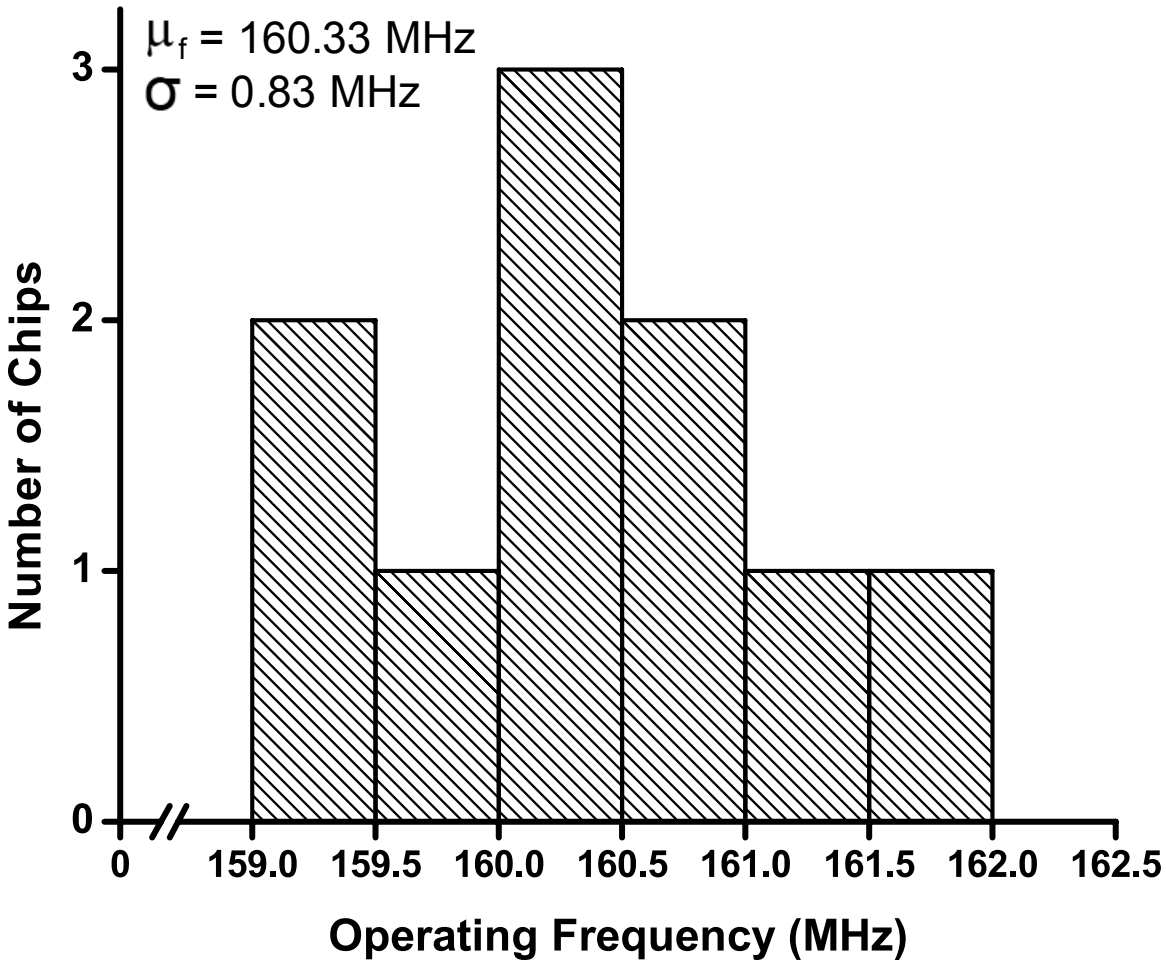


Figure 4.12: Measured resonant frequency distribution at $V_{DC} = V_{CC} = 0.36$V.

4.5.4 SBL FIR Filter Test-Chip Summary

A die photo of the SBL-based FIR is shown in Figure 4.13. Implemented in a 0.13μm bulk silicon regular-V_{th} process, the FIR test-chip comprises a total of approximately 41,000 devices. The FIR filter occupies 0.80mm $\times$ 0.35mm = 0.28mm^2. Including BIST, the entire test-chip occupies a total area of 0.38mm^2. To reduce the parasitic resistance of I/O pads and bondwires, two pads are used in parallel to connect each power-clock phase to one of the terminals of the corresponding off-chip inductor. With the exception of the inductors, which were discrete devices mounted off the die, all other test-chip circuitry was fully integrated on the die.

The table in Table 4.2 summarizes the performance data for our FIR test-chip. For comparison purposes, it also includes published results for other FIR chips. Depending on operating frequency and number of supplies used, our SBL-based FIR test-chip achieves figures of merit that improve upon previous designs [44, 13, 14] by a factor of at least 3X to 20X.

The results presented in this section suggest that SBL is a promising approach for the implementation of regular datapaths with low energy consumption. To assess the suitability and robustness of SBL for volume production, further evaluation would be required, including sensitivity to temperature and wafer-to-wafer process variation, device mismatch, and supply voltage variation.

Paper	**This Work**			[44]	[13]	[14]
Design Type	**14-tap 8x8 FIR**			8-tap 8x8 FIR	8-tap 6x6 FIR	14-tap 8x8 FIR
Technology	**0.13µm (RVT)**			0.13µm (LVT)	0.18µm (RVT)	0.13µm (RVT)
Supply Voltage	**0.27V**	**0.27V**	**0.36V**	0.085V	1.8V	1.08V
2nd Supply	**0.18V**	**N/A**	**N/A**	Chip Dependent Body Bias	N/A	0.59V
Operating Frequency	**20MHz**	**20MHz**	**187MHz**	240Hz	275MHz*	1.01GHz
Power Dissipation	**0.25mW**	**0.31mW**	**5.9mW**	40nW	36mW	124mW
Figure of Merit (nW/MHz/Tap/In-Bit/Coeff-Bit)	**14.40**	**17.37**	**35.31**	325.52**	227.27**	133

*** 550 MSamples/s with parallel architecture**
**** Derived from published results**

Table 4.2: Performance table.

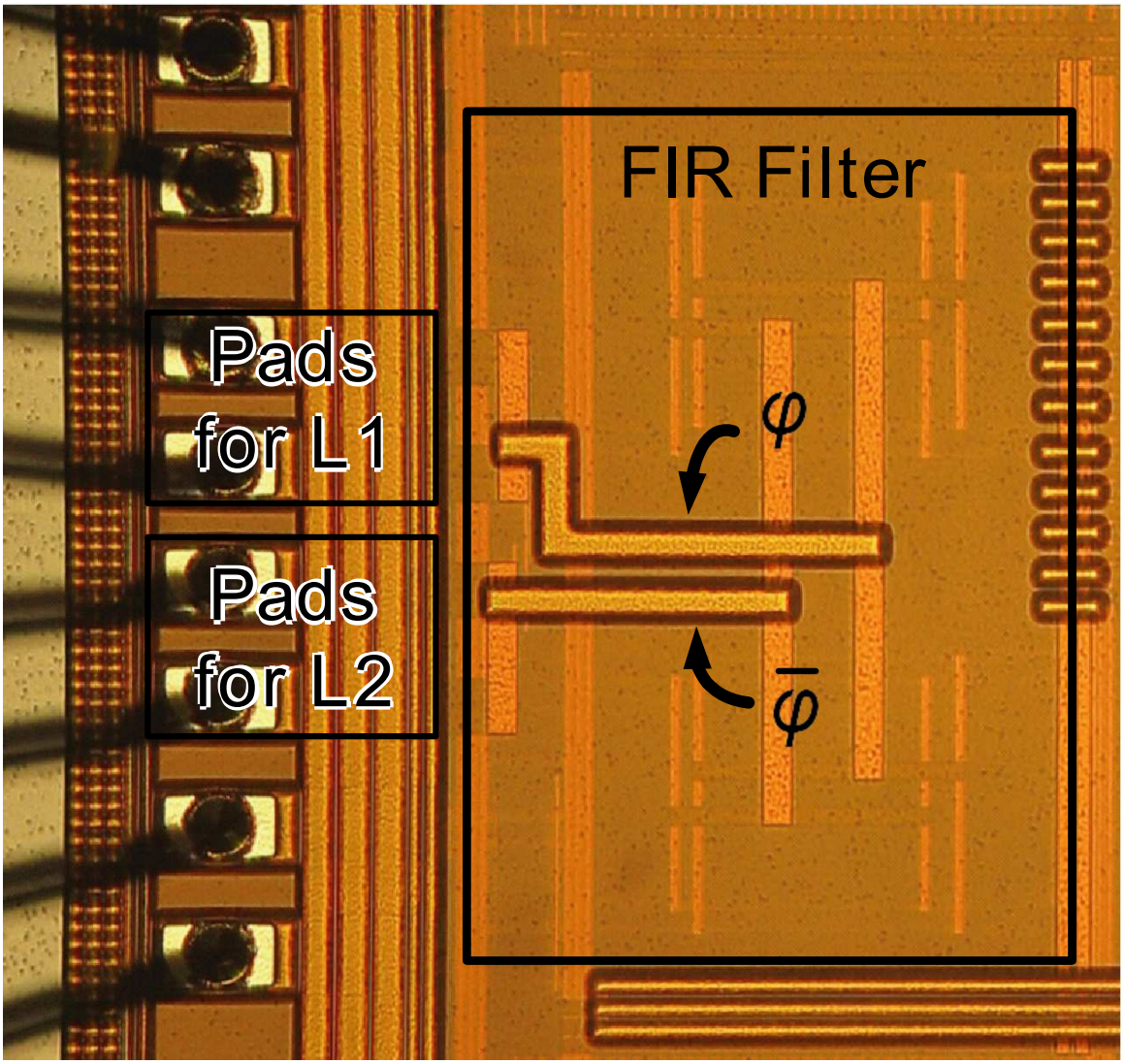


Figure 4.13: SBL FIR die microphotograph.

4.6 Summary

This chapter presents a FIR filter test-chip that relies on a charge-recovery logic, SBL, to achieve multi-MHz clock frequencies with subthreshold DC supply levels. The SBL FIR is the first design that demonstrates a near 200MHz operating frequency with a single 0.36V supply level. Fabricated in a 0.13μm CMOS process with $V_{th,nmos}$=0.40V, the FIR operates with a two-phase power-clock in the 5MHz-187MHz range and with DC supplies in the 0.16V-0.36V range. With a single 0.27V

supply, the test-chip achieves its most energy efficient operating point at 20MHz, consuming 15.57pJ per cycle with a recovery rate of 89% and a figure of merit equal to 17.37 nW/Tap/MHz/InBit/CoeffBit. With the introduction of a second subthreshold supply at 0.18 V, energy consumption at 20MHz decreases further by 17.1%, yielding 14.40 nW/Tap/MHz/InBit/CoeffBit. At its maximum operating frequency of 187MHz, the test-chip achieves 35.31 nW/Tap/MHz/InBit/CoeffBit and 34.47 nW/Tap/MHz/InBit/CoeffBit with one and two subthreshold supplies, respectively. In Spice simulations of extracted layouts, the SBL-based FIR consumes 40% to 50% less energy per cycle in the 17MHz-187MHz range, compared with a static CMOS-based implementation derived by synthesis of the same FIR architecture and automatic place-and-route.

CHAPTER 5

Architecture and Design of Resonant-Clock Flash ADC

This chapter presents the architecture and main building blocks of resonant-clock flash ADC that was designed in a 65nm bulk silicon process and achieved sampling frequency up to 7GS/s. The remainder of this chapter is organized as follows: In Section 5.1, we describe the previous work and challenges in high speed flash ADC designs. Section 5.2 presents details of the flash ADC architecture with single-phase resonant clocking techniques. Section 5.3 describes key building blocks of the ADC, including the track-and-hold amplifier, comparator, encoder, and sense-amplifier flip-flop circuitry. Section 5.4 presents the single-phase clock generator and its distribution network in resonant-clock ADC test-chip. Section 5.5 presents results from the analysis of the integrated inductor that were obtained using a 3D full-wave electromagnetic field solver.

5.1 Introduction

High-speed low-resolution ADCs are essential building blocks for wireless communication systems and data storage devices. Since flash ADC has the highest sampling rate among ADC architectures, it is a natural choice for such applications. How-

ever, power consumption in flash ADCs increases exponentially with resolution, and designing low-power flash ADCs is a challenging task.

Traditionally, ADC power reduction techniques have focused on the architecture and analog circuitry of the converter (e.g., folding [48, 49], interpolation [50], time-interleaving [51]). Based on previously published papers, however, the energy consumption on the clock network and digital circuitry in flash ADCs can be more than 50% of their total energy consumption [48, 52]. Clock-related power could be as high as 30% [48], depending on ADC architecture.

Inductor-based techniques have been used in ADC design in the past, but not in the context of reducing power consumption through resonant clocking. In [52, 53], inductors were used to improve sampling rate. In [54], an inductor was used to generate a low-jitter clock using an integrated LC-based VCO. In that design, the clock network was driven conventionally using clock buffers, yielding no power savings in clock distribution. Resonant clocking has been shown to be an effective approach to the reduction of power consumption in GHz clock speed distribution networks [14, 15, 16].

5.2 Resonant-Clock Flash ADC Architecture

The architecture of the resonant-clock ADC is shown in Figure 5.1. The differential input $In+/In-$ is captured by a track-and-hold amplifier (THA) similar to the one in [55]. 31 dynamic comparators compare the captured signal with reference voltages generated by a resistive ladder and store the resulting 31-bit thermometer code in high-speed SR latches. A 2-cycle Gray code encoder fixes any bubble or sparkle errors and converts the 31-bit thermometer code to a 5-bit Gray code. These blocks are described in more detail in Section 5.3.

All ADC components are synchronized using a single-phase resonant clock except for the THA, which requires sharp clock edges to capture data accurately. To pro-

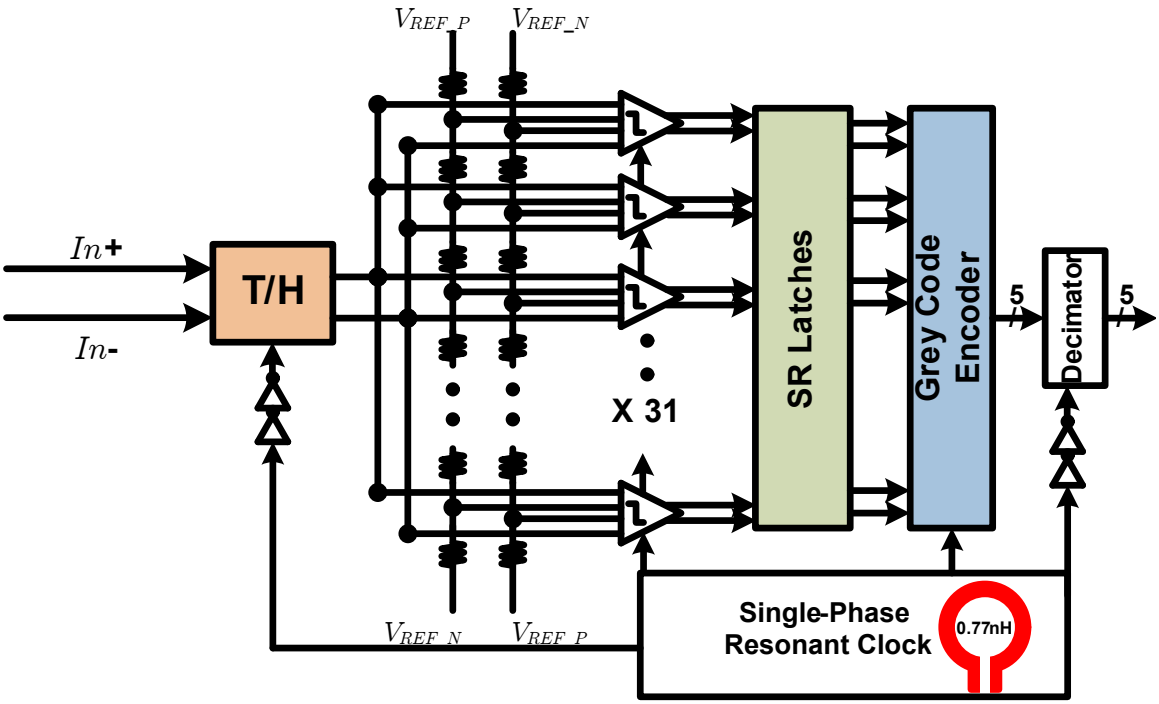


Figure 5.1: Architecture of ADC with single-phase resonant clock.

vide these sharp clock edges, multiple stages of CMOS buffers are used to convert the resonant clock waveform into a square-shape clock, as shown in Figure 5.1. To facilitate testing, the ADC output is captured by a decimator once every 64 cycles. A square-shape clock is generated using the same approach as in the THA, and a standard-cell based frequency divider is used to generate the divide-by-64 clock for the decimator.

5.3 Resonant-Clock Flash ADC Building Blocks

5.3.1 Track-and-Hold Amplifier

Figure 5.2 shows the THA circuitry that consists of a passive PMOS sampling switch and a sampling capacitor, followed by a source follower buffer. The passive switches connect to a sampling capacitor through a dummy switch which lowers the common-mode jump after the track-to-hold transition and improves the linearity of

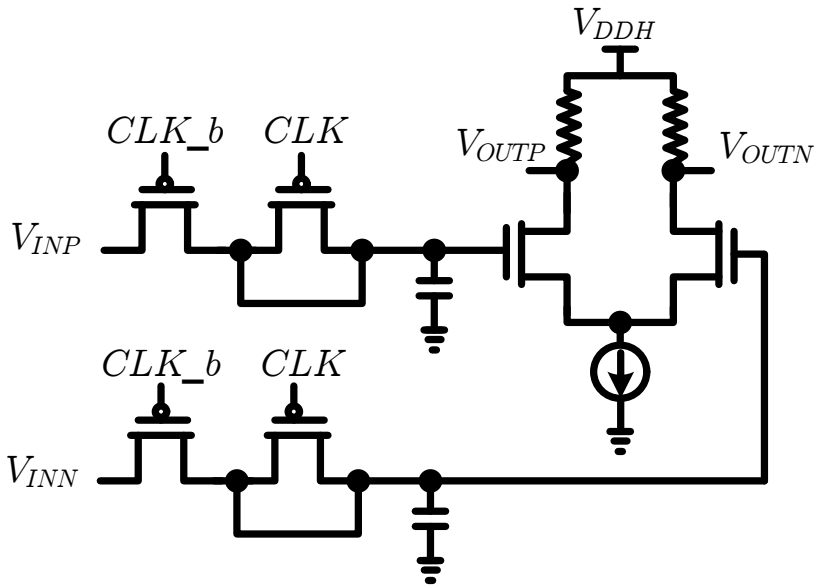


Figure 5.2: Track-and-hold amplifier (THA) circuit.

the switch outputs [56].

5.3.2 Comparators

Figure 5.3 shows the dynamic regenerative comparator used in the ADC. The resonant clock RCK defines two operating phases: reset phase and comparison phase. During the reset phase, RCK is low, and the differential outputs $Out+$ and $Out-$ and nodes $T1$ and $T2$ are precharged high by devices $P3$, $P4$, $P5$, and $P6$. As the clock rises, devices $N4$ and $N5$ discharge the cross-coupled inverters with a slew rate dependent on the input voltage, creating a slight voltage difference across the outputs. During the comparison phase, the cross-coupled inverters regeneratively amplify the voltage difference across the differential outputs to full rail, which is captured in the SR latches.

Comparator offset is predominantly caused by transistor mismatches, especially MOS transistor threshold-voltage and current-factor mismatches [57]. In each comparator with two differential pairs, the standard deviation of the threshold voltage can be approximated as follows:

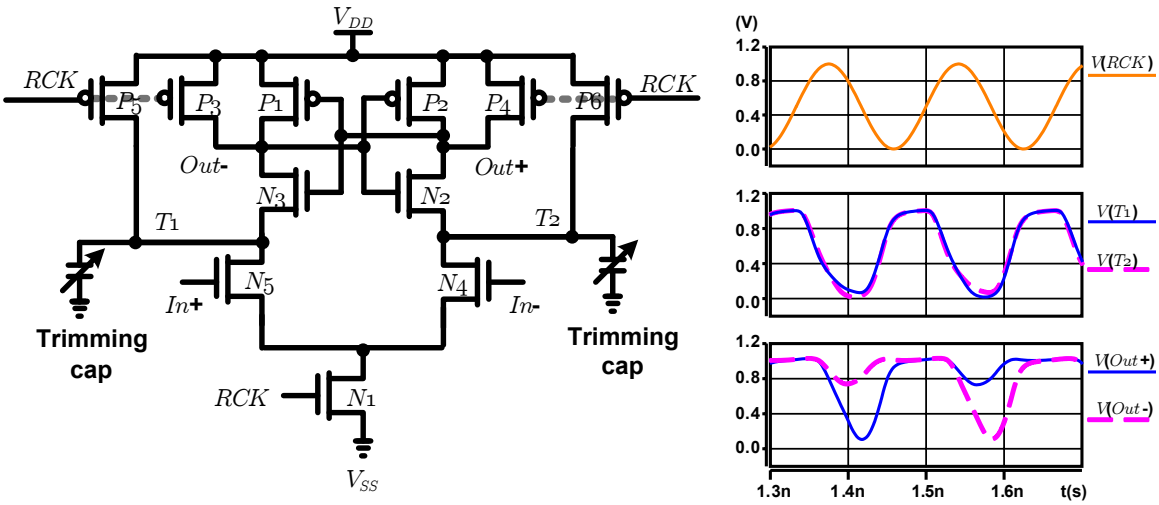


Figure 5.3: Schematic of the comparator and waveform.

$$\sigma(V_{th}) = \frac{A_{V_{th}}}{\sqrt{W \times L}} \times 2, \tag{5.1}$$

where $A_{V_{th}}$ is a process-specific parameter, and $W \times L$ is the transistor gate area. From Equation (5.1), for a 5-bit ADC to achieve an effective number of bits (ENOB) greater than 4.7, the comparator gate area needs to be sized up sufficiently to keep $\sigma(V_{th})$ below 0.1 LSB, where 1 LSB is equal to the full-scale input voltage divided by 2^5 [58]. Rewriting Equation (5.1), the transistor gate area $W \times L$ can thus be expressed as:

$$\begin{aligned} W \times L &= (\frac{2 \times A_{V_{th}}}{\sigma(V_{th})})^2 \\ &= (\frac{2 \times A_{V_{th}}}{0.1 \times LSB})^2 \\ &= (\frac{2 \times A_{V_{th}}}{0.1 \times \frac{V_{FS}}{2^5}})^2 \\ &= (\frac{2^6 \times A_{V_{th}}}{0.1 \times V_{FS}})^2, \end{aligned} \tag{5.2}$$

where V_{FS} is the full-scale input voltage of the ADC. Equation (5.2) yields transistor sizes that are prohibitively large in practice. Furthermore, these large comparators increase the output load capacitance of the THA and the load of the clock network, resulting in increased energy consumption.

To keep power dissipation low and transistor sizes practical, the differential pairs in the comparator are intentionally sized with standard deviation of threshold voltage variation, $\sigma(V_{th})$, of up to 0.45 LSB of ADC. Threshold calibration is then used to further compensate for threshold voltage mismatches in the differential pairs of the comparators. Threshold calibration is performed by digitally including additional PMOS capacitance. Using scannable controls to change the voltage applied to the gate of the PMOS, the effective capacitance presented at the source and drain junctions connected to nodes $T1$ and $T2$ can be varied. This capacitance tuning ability allows for the compensation of current differences caused by device mismatch. The calibration step size is 1/3 LSB with a 4-bit control signal. The reference voltage calibration range is about $\pm$2.5 LSB, which is enough to cover a 3 sigma variation in V_{th} mismatch.

To avoid edge effects in the comparator layout, the length of transistors in each comparator is sized up by 1.1X of the feature device length. The 31 comparators are placed in two rows (15 and 16 comparators in each row). Five dummy comparators are added to the two ends of the comparator array. The total number of comparators, including the 5 dummies, is thus 36.

5.3.3 Grey Code Encoder

The function of the Grey encoder is to convert the 31-bit wide thermometer code generated by the comparator array into a 5-bit Grey code. In a thermometer code, if input is high with respect to the reference level of a particular comparator, then the corresponding output of that comparator is high. Ideally, all comparator outputs

below the input level are 1s, and all comparator outputs above the input level are 0s. However, for fast input signals, small timing differences in the response times of the comparators, combined with threshold voltage offset can cause a situation where a 1 is found above 0. This phenomenon is called a bubble or sparkle error in the thermometer code. Metastability of the comparators and the presence of bubbles in the thermometer code are two classes of errors that must be addressed.

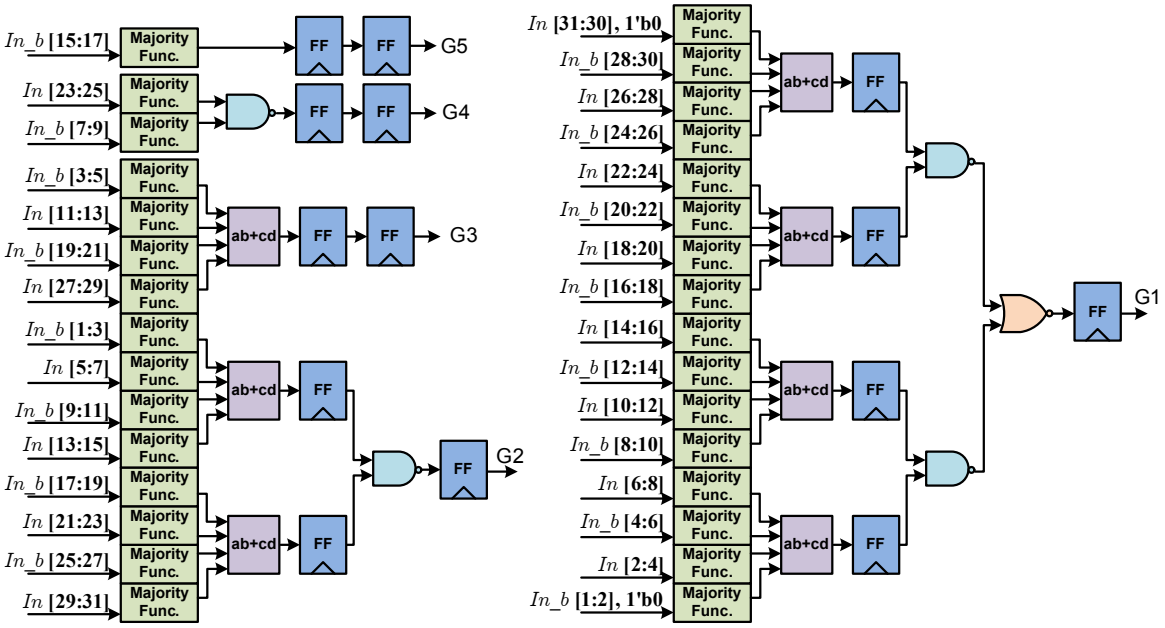


Figure 5.4: Schematic of the 2-cycle resonant-clocked Gray code encoder.

Figure 5.4 shows schematic of a 2-stage the digital encoder that is used in the ADC. In the first stage, a majority function corrects the bubble/sparkle errors in the 31-bit thermometer code. Then, in the second stage, the function $ab+cd$ is performed to reduce the intermediate data and thus reduce the number of flip-flops. Only 14 latches are used in our decoder, resulting in decreased power consumption.

5.3.4 Sense-Amplifier Flip-Flop

Figure 5.5 shows the schematic of the flip-flop used in the ADC. This sense-amplifier flip flop topology [1] has been chosen to ensure short $C-to-Q$ delay when

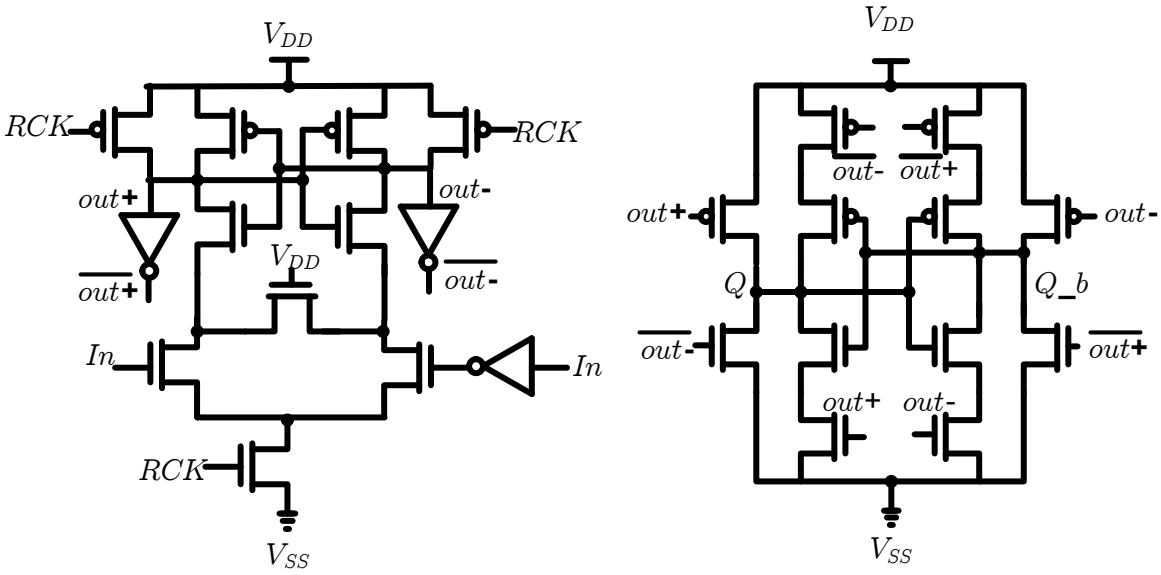


Figure 5.5: Schematic of the sense-amplifier flip-flop.

clocked by a 5.5GHz and 7GHz resonant clock waveform with 10% to 90% transition time equal to 54ps and 42ps, respectively. It also has the advantage of tracking the effect of variation in the comparator since their structures are very similar.

5.4 Clock Network Design

A single-phase resonant clock is generated by a single-ended clock generator, as shown in Figure 5.6. An on-chip inductor L is used to set up an LC tank, where C is the parasitic capacitance associated with the clock distribution network and the clock pins of the clocked elements. In this chip, L is estimated at 0.77nH using a commercial 3D electromagnetic field solver, and C is estimated at 1.2pF using a commercial RC extraction tool. Programmable switches (PMOS: 6μm to 60μm with step size of 6μm; NMOS: 3μm to 30μm with step size of 3μm) powered by supply V_{CK} are placed near the clock terminal of the inductor to replenish the energy lost due to the parasitic resistance of the clock network. A 12pF capacitive divider is placed on the other terminal of the inductor to stabilize the oscillation. A reference

clock generated by a programmable ring oscillator drives the programmable switches, enabling the operation of the LC tank at the frequency of that clock. The reference clock drives the switches either directly or through an on-chip pulse generator that varies its duty cycle D in the range $20\% \leq \mathrm{D} \leq 50\%$.

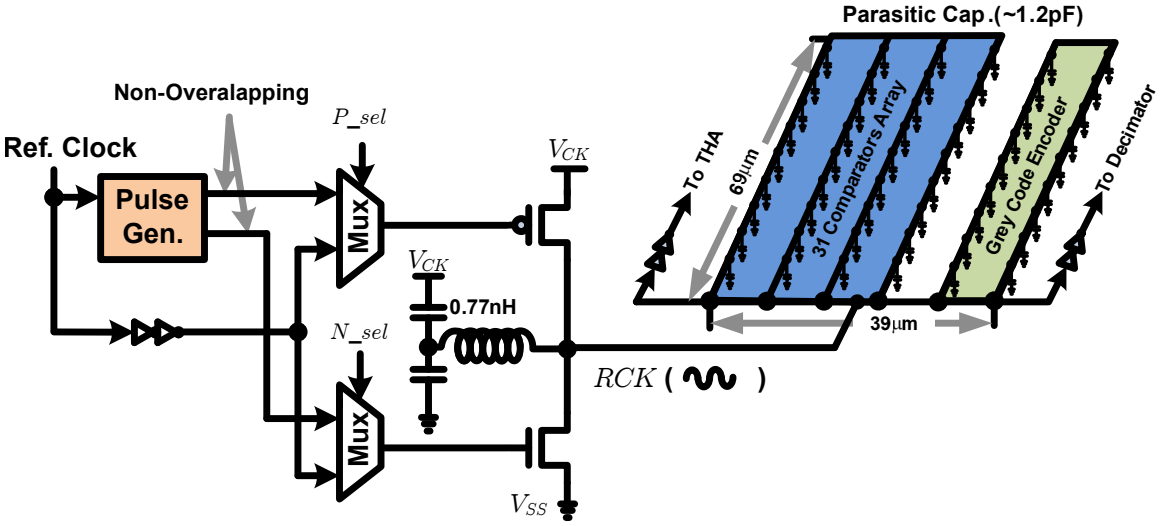


Figure 5.6: Resonant clock generator with variable pulse controls.

To achieve high energy efficiency, various inductor configurations have been evaluated using a commercial electromagnetic field solver. Based on this evaluation, a 1.75-turn 130μm $\times$ 130μm 2-metal-layer inductor has been chosen for the ADC. A ground shield consisting of patterned M1 has been added directly underneath the inductor to improve its quality factor (Q) and thus increase the efficiency of the LC tank. Detailed inductor analysis results are given in Section 5.5. Figure 5.6 also shows the clock distribution network topology of the ADC. The single-phase resonant clock is distributed over a 39μm $\times$ 69μm area. The clock terminal of the inductor is first connected to a 39-μm long, 2-metal-layer clock trunk with an effective width of 2.2μm. The clock is then distributed along six 69μm 4-metal-layer clock spines with an effective width of 2.6μm each. 45 clocked elements are placed directly under the six clock spines to achieve low distribution resistance and low coupling capacitance.

The maximum resistance between the clock terminal of the inductor and any clocked element is 1.15Ω, and the load for the corresponding path is 380fF, including wire and gate capacitance. The resulting maximal RC product for any clock element is quite small, yielding a worst-case clock skew below 1ps.

5.5 Inductor Design and Analysis

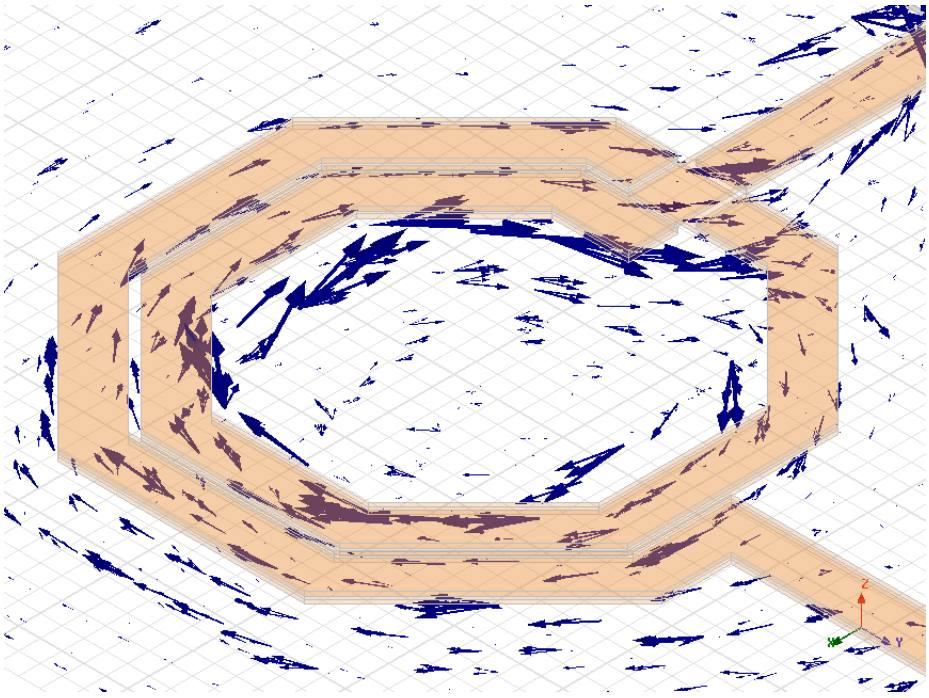
Inductor design and analysis were performed after extracting the clock network capacitance of the ADC design. Based on the parasitic capacitance of the entire clock network, the inductor value was selected to yield a resonant frequency near the target operating frequency. The relation between the inductance L, capacitance C, and resonant frequency f is as follows:

$$f = \frac{1}{2\pi\sqrt{LC}}. \tag{5.3}$$

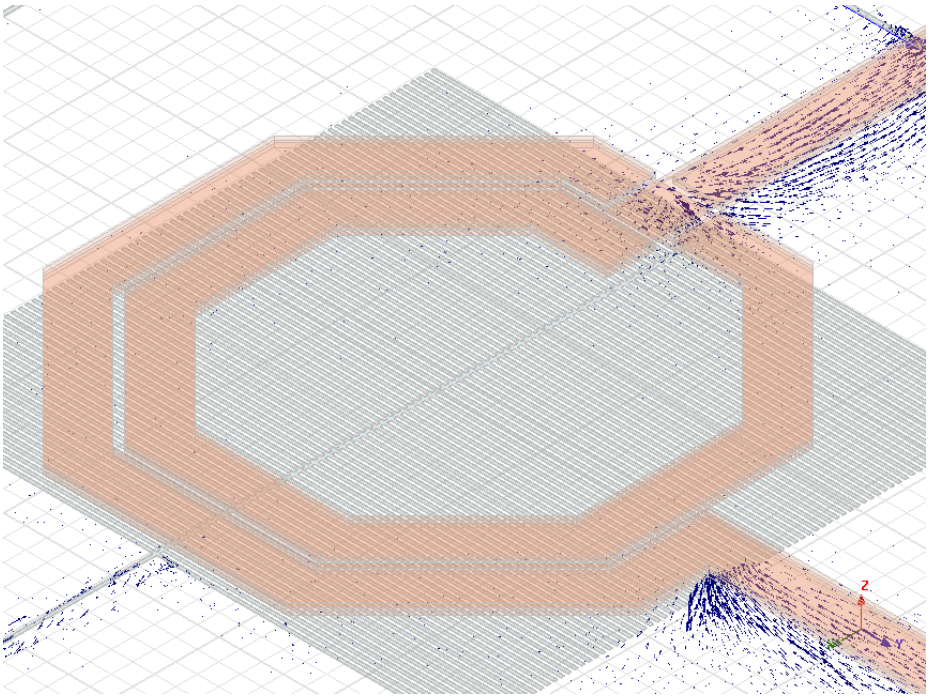
An inductor database supplied by the foundry gives inductor parameters (e.g., number of turns, diameter, metal width area, and quality factor) to achieve select inductance value. Since this database gives only a limited number of metal options and inductance values, to obtain an inductor that meets our exact specifications, we have used these database parameters as a starting point for designing the inductor used in our ADC.

The inductor layout was drawn in Cadence Virtuso. A pattern of M1 strips was added directly below the inductor to reduce the eddy current loss in the substrate and improve the inductor quality factor (Q) [59]. The layout was exported in stream format and imported to HFSS, a 3D full-wave electromagnetic simulator, for electromagnetic field analysis.

The final inductor design was a 1.75-turn coil with outer dimensions 130μm ×

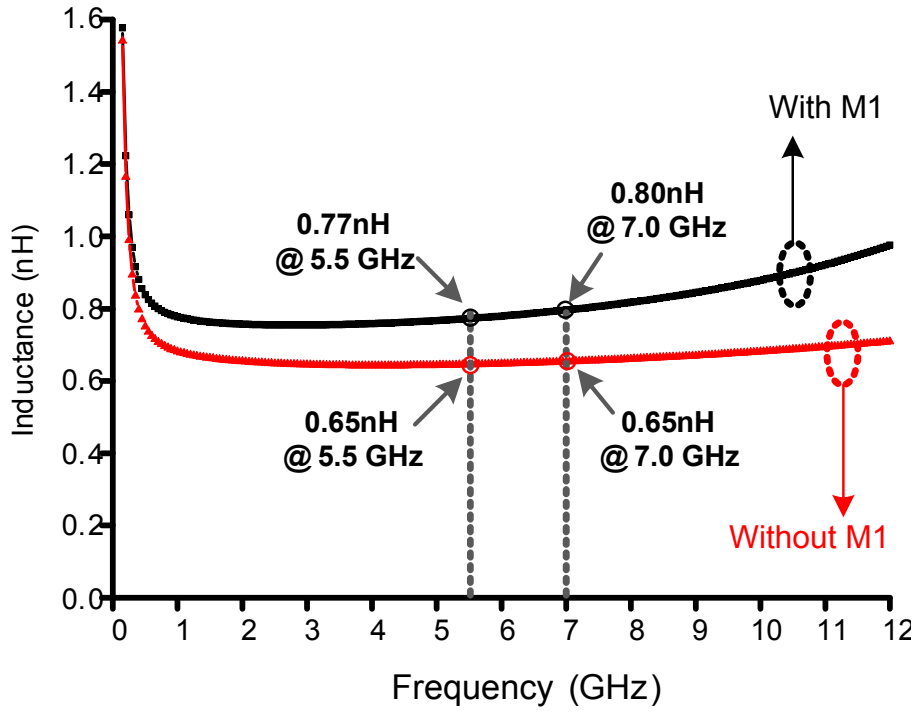


(a)

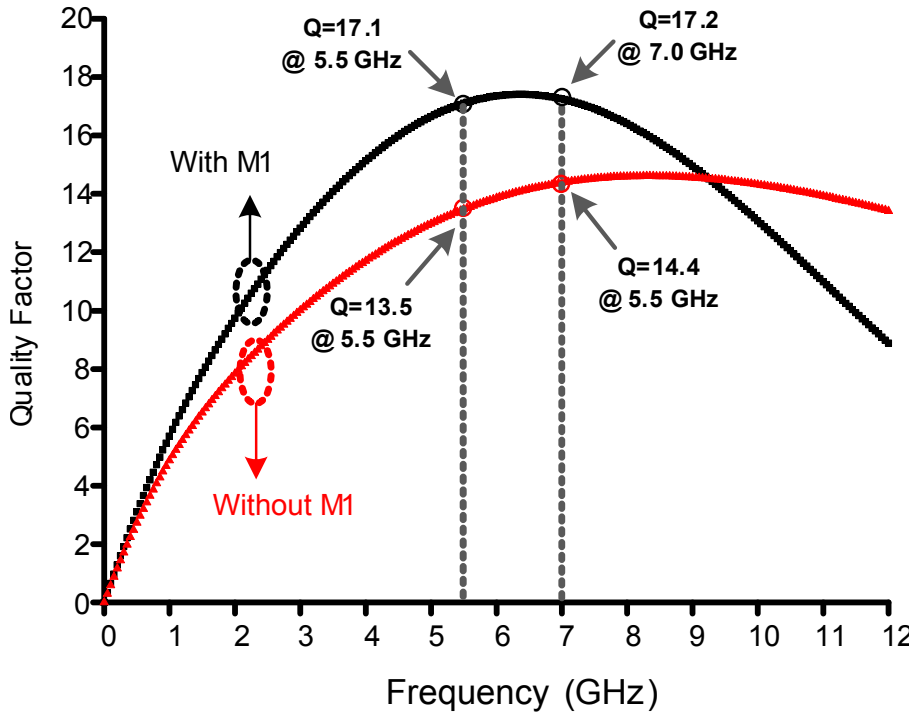


(b)

Figure 5.7: Current density in the substrate underneath inductor: (a) without M1 shield, and (b) with M1 shield.



(a)



(b)

Figure 5.8: (a) Inductance vs. resonant frequency with and without M1 shield. (b) Quality factor vs. resonant frequency with and without M1 shield.

130μm and inner diameter of 85μm. It consists of 2 metal layers (M8 and M9) with 11.5μm-wide wires each and an effective thickness of 1.8μm. The simulation results show that with M1 shield, the inductor provides 0.77nH of inductance and a quality factor of 17.1 at 5.5GHz, and 0.80nH of inductance and a quality factor of 17.2 at 7.0GHz.

Figure 5.7 shows the inductor layout with and without the M1 pattern shield and current density in the substrate, as obtained from HFSS. For the setup in Figure 5.7(a) without M1 shield, the image current induced by the magnetic field in the conductive substrate layer flows in a direction opposite to that of the current in the spiral. The negative self-inductance then leads to a significant drop in the total inductance and hence Q. In the simulation results shown in Figure 5.7(b), the induced eddy current in the substrate layer is greatly reduced by the M1 strips shield.

Figure 5.8(a) shows inductance versus frequency. With M1 strips added, eddy current loss in the substrate is reduced, and inductance increases by 15.4% at 5.5GHz and 23.1% at 7GHz. Figure 5.8(b) shows the quality factor Q versus frequency, where Q is defined as:

$$Q = -\frac{Im(Y(1,1))}{Re(Y(1,1))}, \tag{5.4}$$

The graph shows that the inductor with M1 shield achieves a higher peak Q, with the peak occurring at a lower frequency. For frequencies above the peak Q frequency, the qualify factor for the inductor with M1 shield drops faster than the one without M1. This drop is due to the increased coupling capacitance between the inductor and the M1 shield. By adding the M1 shield Q is improved by 26% at 5.5GHz and 19% at 7GHz, respectively.

CHAPTER 6

Evaluation and Testing of 7GS/s Resonant-Clock Flash ADC

In this chapter, we present experimental results from the evaluation of resonant-clock ADC test-chip, demonstrating its high performance and high energy efficiency. In this ADC, resonant clocking is deployed to decrease the power consumption of the network that distributes the clock signal to the analog and digital circuitry of the ADC. A fully integrated inductor is used to resonate the parasitic capacitance of the entire clock distribution network all the way to the clocked timing elements.

In Section 6.1 , we present measurement results and ADC performance characterization from the resonant ADC test-chip. Conclusions are given in Section 6.2.

6.1 Measurement Results

Multiple independent supplies are used in this ADC test-chip. V_{DDA} is the supply for the comparators, V_{DD} is the supply for digital components, V_{DDH} is the supply for THA, and V_{CK} is the supply for the clock generator. This setup allows for the independent monitoring of power consumption in various ADC components at different sampling frequencies. The calibration of the test-chip is performed by externally applying the DC voltages to the ADC inputs, at a sampling rate of 5.5GS/s. By setting

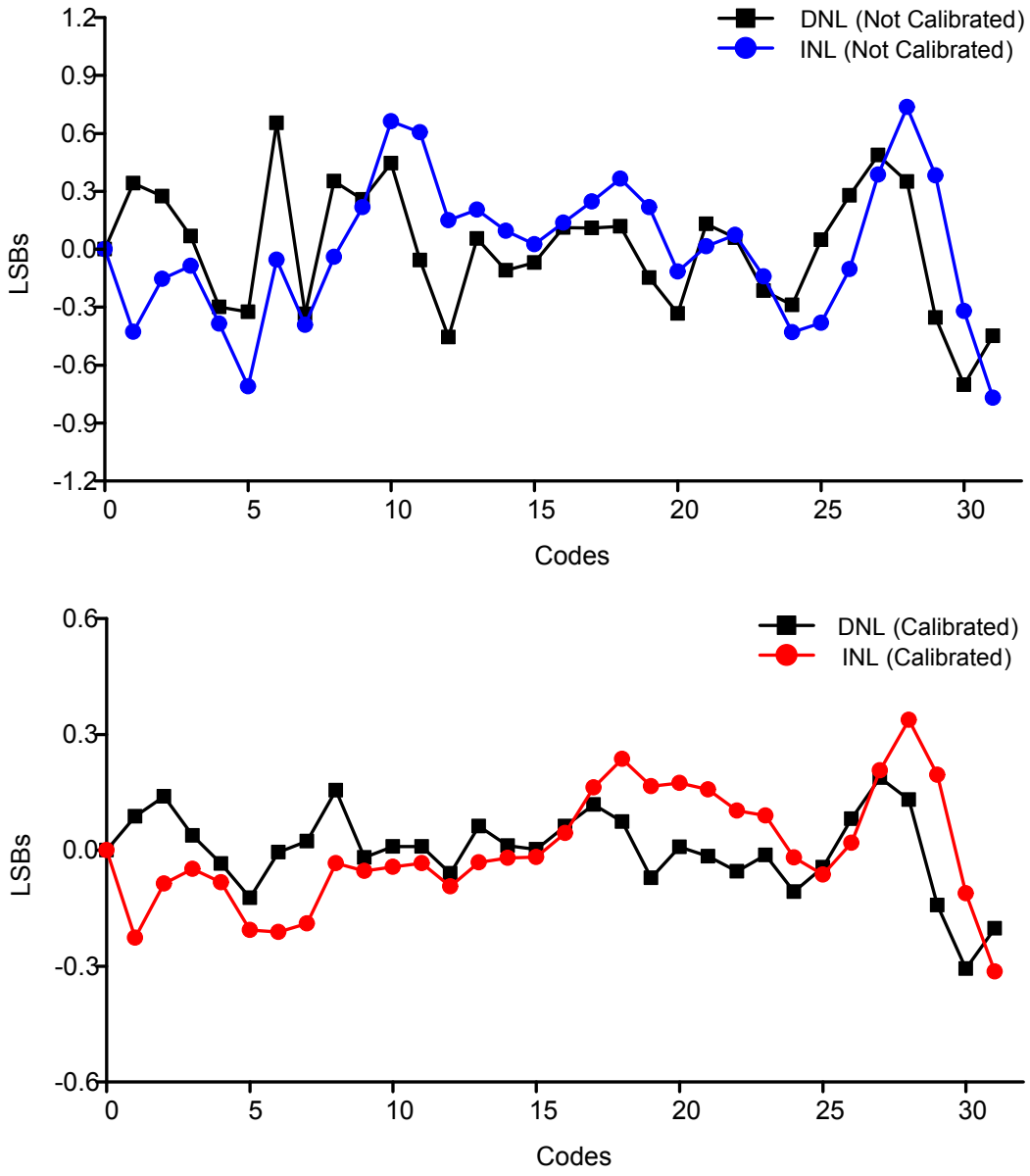


Figure 6.1: Measured DNL/INL.

threshold voltages at the ADC inputs, we adjust the offset of the corresponding comparators to a desired output code. The calibration process is done in a commercial software, Labview, and the optimal offset compensation value is chosen automatically for each comparator.

Figure 6.1 shows measured differential non-linearity (DNL) and integral non-linearity (INL) at 5.5GS/s. Measurements taken before and after calibration show that DNL improves from 0.65/-0.70 LSB to 0.21/-0.30 LSB, and that INL improves

from 0.74/-0.77 LSB to 0.33/-0.31 LSB.

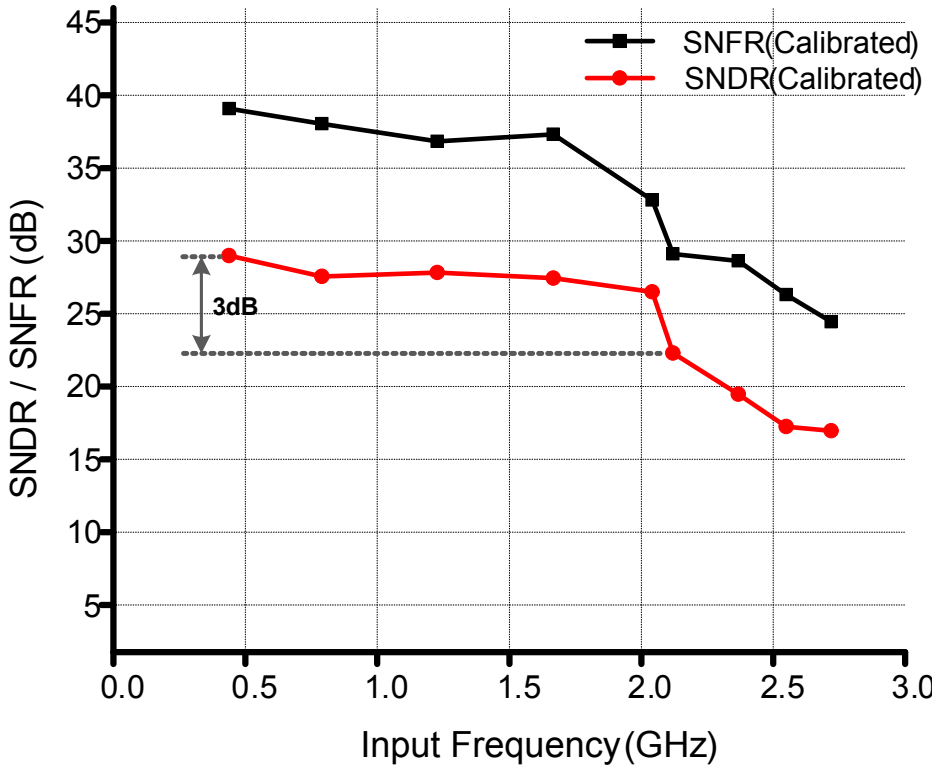


Figure 6.2: Measured SNDR/SNFR vs. input frequency.

Figure 6.2 shows signal-to-noise and distortion ratio (SNDR) and signal-to-noise floor ratio (SNFR) versus input frequency (f_{in}). Operating the chip at 5.5GS/s and applying a full-scale sine input yields a SNDR of 29.0dB and a SNFR of 39.1dB with f_{in} at 440MHz (4.56 ENOB) and a SNDR of 26.5dB with f_{in} at 2.04GHz (4.11 ENOB). The 3dB difference between the two frequency points indicates that the effective resolution bandwidth (ERBW) is 2.04GHz.

Figure 6.3 shows the breakdown of measured energy consumption versus sampling frequency. The ADC test-chip operates from 4.5GS/s to 7.0GS/s. The clock energy of the resonant clock ADC is obtained from measurements using a separate power supply V_{CK}. As shown in the figure, in the frequency range from 4.5GHz to 7.0GHz, V_{DDA} ranges from 0.96V to 1.10V, V_{DD} ranges from 0.99V to 1.10V, and V_{CK} ranges from 0.96V to 1.10V.

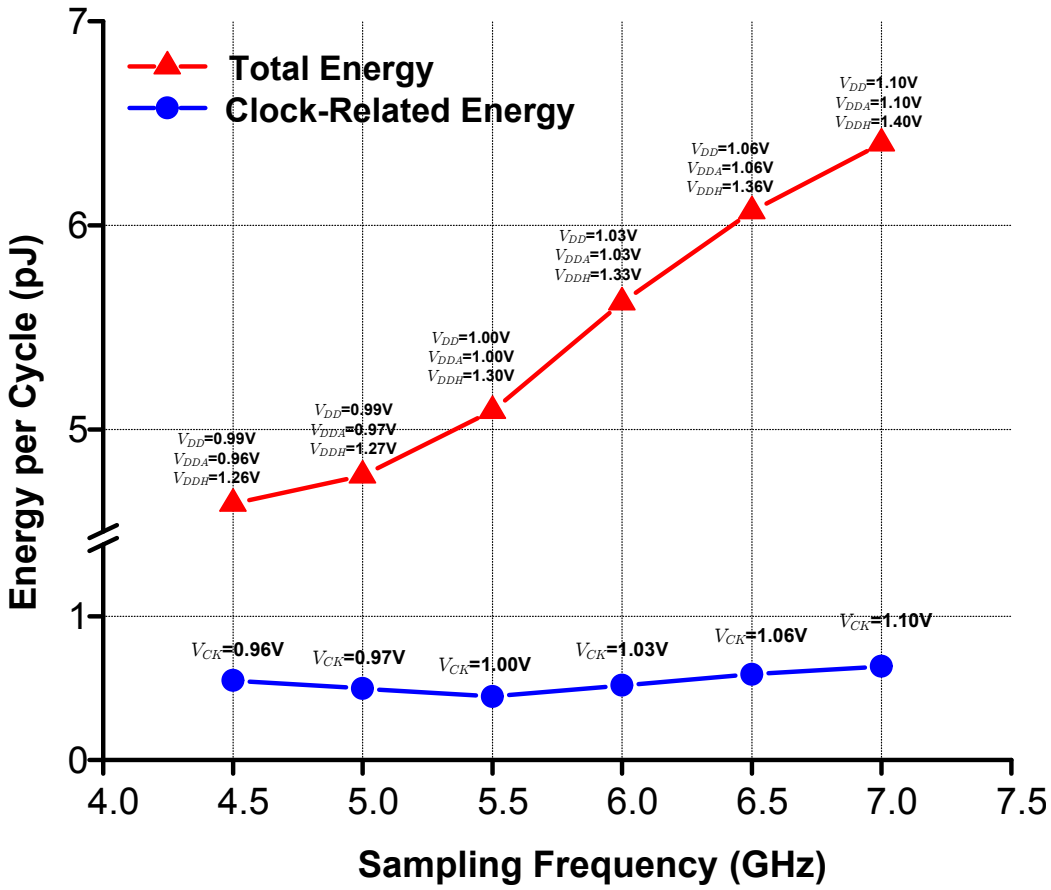


Figure 6.3: Measured energy per cycle vs. sampling frequency.

In the graph of Figure 6.3, the minimum clock energy consumption of 0.54pJ is reached at 5.5GHz, indicating that 5.5GHz is the closest frequency point to the natural frequency of the ADC. At this frequency, the resonant ADC consumes 5.1pJ per cycle with only 10.7% of the total energy consumption being clock related. At the same frequency, the ADC reaches the lowest FoM of 396fJ per conversion step. The clock system in the resonant ADC consumes 54% less energy than CV^2.

Since the energy used to drive the clock generator switches is included in the total clock-related energy consumption, the actual resonant LC recovery rate is higher than the resonant system recovery rate of 54%. Spice-level simulations show that clock distribution energy is about 69% of the total clock-related energy, with the remainder 31% consumed on the pulse generator and drivers to the clock generator

switches. By applying this number to the measurement result and including only energy consumption for clock distribution, the *LC* oscillation has an energy recovery rate of 77%.

The minimum clock-energy frequency of 5.5GHz is close to the estimated resonant frequency obtained from post-layout extraction. Inductance L has been extracted with a commercial 3D electromagnetic field solver from the layout of the inductor in the test-chip. The capacitance of the clock network has been extracted from the complete metal-filled design using a commercial *RC* extraction tool. These extracted inductance and capacitance values yield a natural resonant frequency of 5.2GHz, which is 6% away from the minimum clock-power frequency of 5.5GHz.

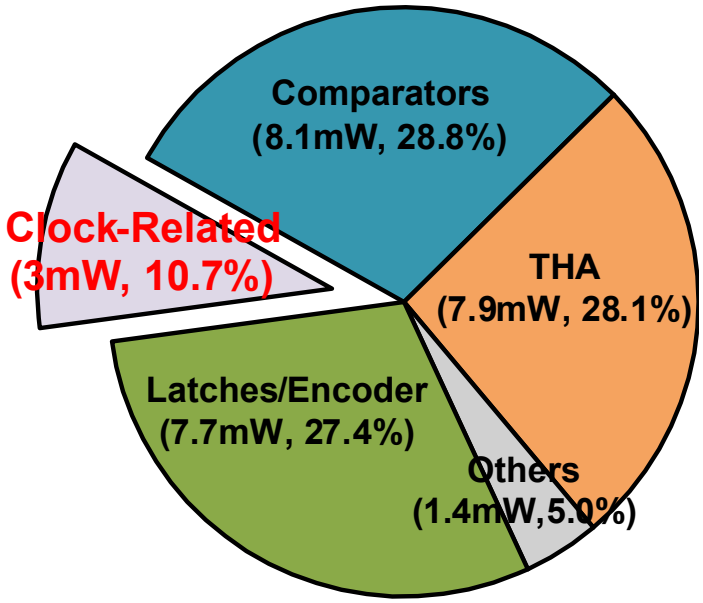


Figure 6.4: Measured power breakdown at 5.5GS/s.

Figure 6.4 shows the power breakdown at a sampling frequency of 5.5GHz, obtained by measuring current from the independent supplies powering each ADC component. The total power consumption at 5.5GS/s is 28mW, with 28.8% of this total used in the comparators, 28.1% in the track and hold circuits, and 27.4% in the digital components. Total clock power, which includes power consumption in the clock distribution network, programmable switches, and the pulse generator, is 3mW and

accounts for 10.7% of total power.

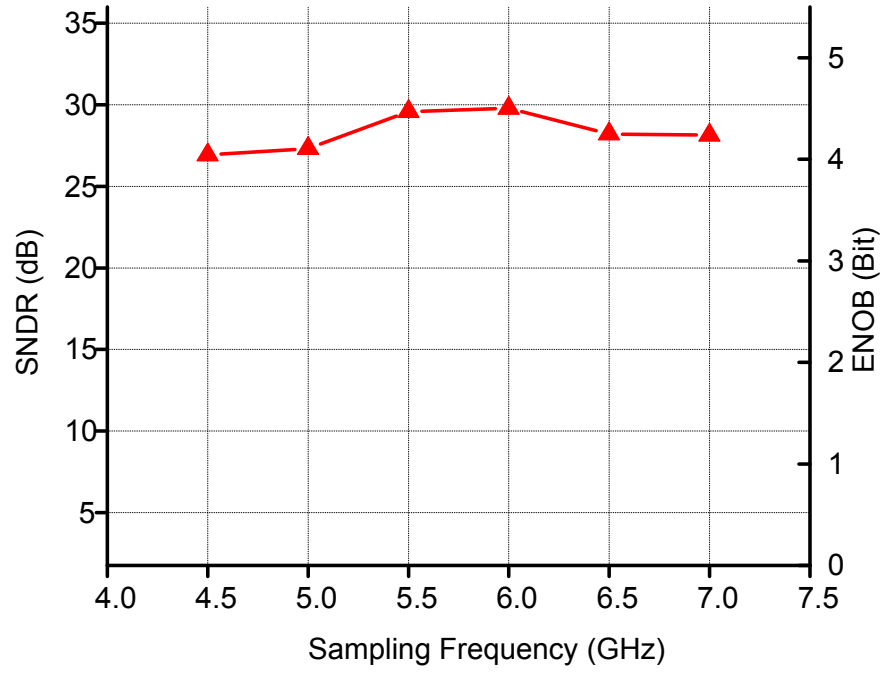
Figure 6.5(a) shows the measured SNDR versus sampling clock frequency (f_s) with a 400MHz full-power input. The measured ENOB is 4.18 and 4.62 at 4.5GS/s and 5.5GS/s, respectively. At 7.0GS/s, the ADC achieves more than 4.40 effective bits after calibration. Figure 6.5(b) shows FoM versus sampling frequency, where FoM is defined as follows:

$$FoM = \frac{power}{2^{ENOB} \times min(2 \times ERBW, f_s)}. \tag{6.1}$$

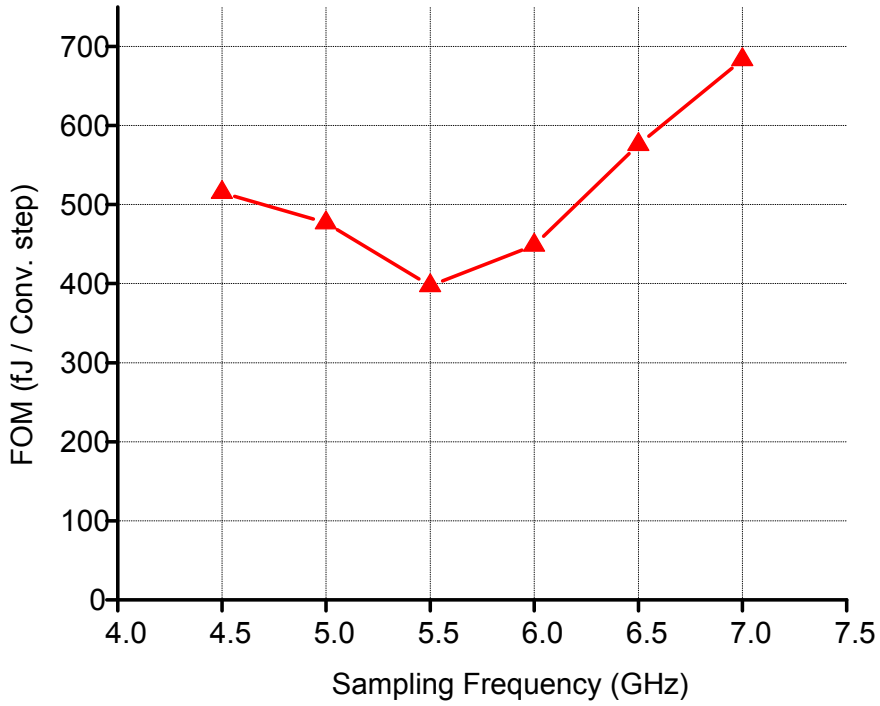
The lowest FoM of 396fJ per conversion step is achieved at the sampling frequency of 5.5GS/s. At sampling frequencies greater than 5.5GS/s, FoM increases due to the increasing supply voltages and thus higher energy consumption. At lower sampling frequencies, even though the ADC has lower energy consumption, ENOB decreases, and thus FoM increases.

Table 6.1 summarizes the performance of the ADC. Operating in the vicinity of its resonant frequency with sampling rate 5.5GS/s, THA supply at 1.3V, and other supplies at 1.0V, the ADC dissipates 28mW with f_{in} at 2.04GHz. FoM is 396fJ per conversion step. Operating at it maximum sampling frequency of 7.0GS/s, the ADC dissipates 45mW with a 0.1V increase to each supply. At that frequency, 11.1% of the total energy consumption is clock-related and the ADC yields a FoM of 683fJ per conversion step.

Figure 6.6 compares the ADC in this work with previously reported ADCs in [2]. Our resonant clock ADC achieves a lower FoM than all previously published ADCs operating above 2.2GHz. Specifically, it achieves 58% lower FoM than [60], which used interpolation and reduced-output-swing of front-end circuits. It also achieves 20% and 14% lower FoM than [61] and [51], respectively, which used clock duty cycle



(a)



(b)

Figure 6.5: (a) Measured SNDR vs. sampling frequency with input frequency of 400MHz. (b) Measured FoM vs. sampling frequency.

Technology	65nm	
Resolution	5 bit	
Input Range	800 mV_{pp}	
Effective Clock Load	1.2 pF	
DNL	0.21 /-0.30 LSB	
INL	0.33/ -0.31 LSB	
Active Area	0.018 mm^2 (without inductor) 0.035 mm^2 (with inductor)	
Package	QFN32	
Sampling Frequency	5.5 GHz	7.0 GHz
Power Supply	THA: 1.3V Other: 1.0V	THA: 1.4V Other: 1.1V
Power Consumption	28mW	45mW
Clock Power	3mW	5mW
ERBW	2.04GHz	2.05GHz
SNDR	29.0dB @ 440MHz 26.5dB @ 2.04GHz	28.1dB @ 400MHz 25.9dB @ 2.05GHz
ENOB (bits)	4.56 @ 440MHZ 4.11 @ 2.04GHz	4.38 @ 400MHz 4.00 @ 2.05GHz
Figure of Merit	396 fJ / Conv. Step	683 fJ / Conv. Step

Table 6.1: Performance summary.

control techniques and 8X time-interleaving to improve sampling rate. The techniques used in [51, 60, 61] are compatible with resonant clocking and can be used to further improve the FoM of resonant-clock ADCs, such as the one presented in this work.

Figure 6.7 shows the die microphotograph of the ADC chip. The device has been fabricated using a 65nm CMOS technology and packaged in a 32-pin QFN package. The ADC core occupies 277.6μm $\times$ 64μm = 0.018mm^2. Including the 130μm $\times$ 130μm inductor, the ADC occupies 0.035mm^2.

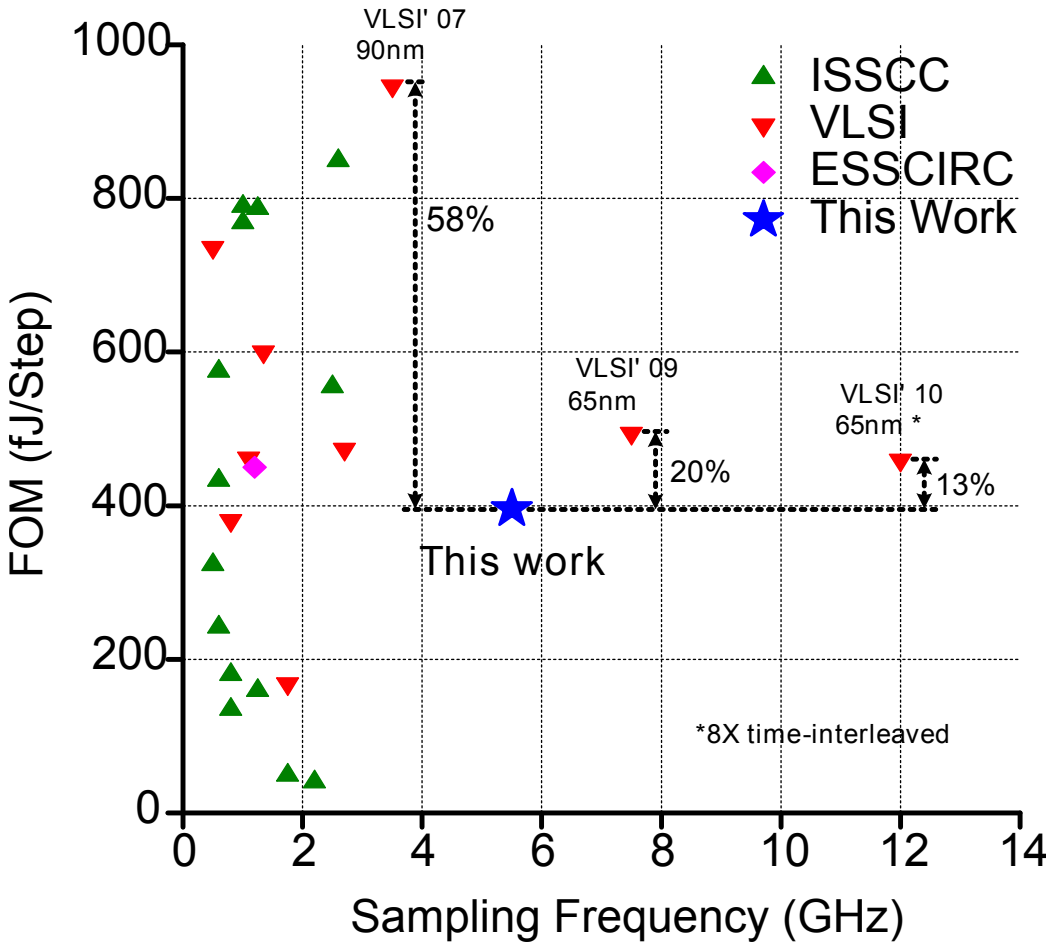


Figure 6.6: FoM vs. sampling frequency: comparison to prior work summarized in [2].

6.2 Summary

This chapter presented measurement results from the evaluation of a 5-bit non-interleaved resonant-clock flash ADC that achieves a sampling rate of 7GS/s. An on-chip 0.77nH inductor resonates the entire clock distribution network to achieve energy-efficient operation. The ADC has been designed in a 65nm bulk CMOS process and occupies 0.035mm^2, including the integrated inductor. Operating at 5.5GHz, it consumes 28mW, yielding 396fJ per conversion step, with its clock network accounting for 10.7% of total power and consuming 54% less energy over CV^2. Operating at it maximum sampling frequency of 7.0GS/s, the ADC dissipates 45mW with a 0.1V

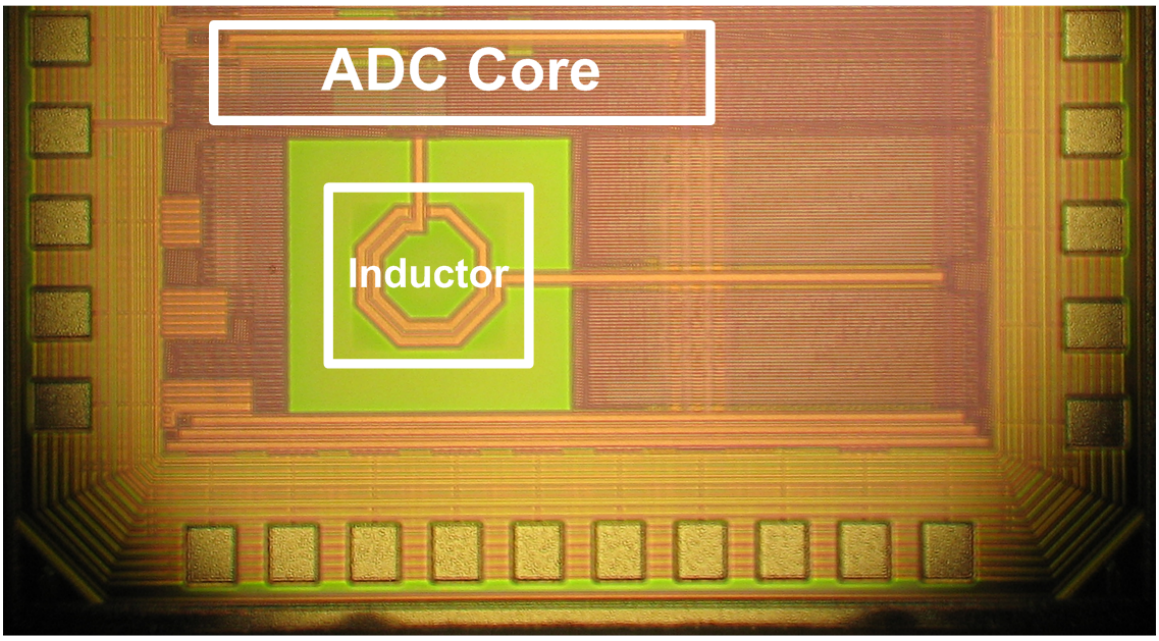


Figure 6.7: ADC die microphotograph.

increase to each supply. At that frequency, 11.1% of the total energy consumption is clock-related, and the ADC yields a FoM of 683fJ per conversion step. By comparison, in a typical flash ADC design, 30% of total power is clock-related.

CHAPTER 7

Conclusions and Future Directions

This chapter summarizes the contributions of this dissertation. Charge-recovery techniques are explored as a solution targeting both high-performance and ultra-low power platforms to achieve increased energy efficiency over conventional CMOS designs without compromising performance or reliability. Two charge-recovery systems operating at different frequency points are tested and evaluated. Both designs provide high performance at their corresponding supply-levels and achieve higher energy efficiency than their conventional CMOS designs counterparts.

7.1 Subthreshold Boost Logic

For ultra-low energy consumption with high performance, we present Subthreshold Boost Logic (SBL), a circuit family that is capable of operating at multi-MHz clock frequencies using subthreshold supplies. Unlike subthreshold circuitry, which uses subthreshold currents for computation and typically operates in sub-MHz range, SBL gates are overdriven to operate in the linear region, achieving order-of-magnitude improvements in operating speed over subthreshold logic. Energy efficient operation is ensured through the use of aggressively-scaled DC supplies at subthreshold levels and by deploying charge recovery design techniques to boost these subthreshold supply levels by 3X to 4X.

To demonstrate the performance and energy efficiency of SBL, we have designed a 14-tap 8-bit finite-impulse response (FIR) filter test-chip implemented using SBL. Fabricated in a 0.13μm bulk silicon process with regular thresholds, the test-chip functions correctly for clock frequencies ranging from 5MHz to 187MHz, relying on two discrete off-chip inductors to boost the subthreshold supply in an energy-efficient manner. Clock drivers are fully integrated and distributed across the entire clock network. With a single subthreshold supply set to 0.27V, it achieves its most energy efficient operating point at 20MHz, yielding a figure of merit equal to 17.37 nW/Tap/MHz/InBit/CoeffBit. With the introduction of a second subthreshold supply set to 0.18V, energy consumption due to crowbar currents at clock frequencies below 30MHz is significantly reduced. Maximum energy efficiency is improved by 17.1% and is achieved at 20MHz, yielding 14.40 nW/Tap/MHz/InBit/CoeffBit. At maximum energy efficiency, energy recovery rates range from 86% to 89%, depending on the number of supplies. Based on Spice simulations of the SBL FIR and a fully-automatic static CMOS implementation of the same FIR architecture, the SBL design consumes 40% to 50% less energy per cycle in the 17MHz-187MHz range while incurring a 15% area overhead.

7.2 Resonant-Clock Flash ADC Design

For high performance and high energy efficiency, we present a 5-bit 7GS/s non-interleaved flash ADC with resonant clocking techniques. The chip has been fabricated with a 65nm CMOS process. A 0.77nH on-chip inductor is used to generate a single-phase resonant clock and reduce power consumption of the ADC clock distribution network. From measurement results, the test chip achieves a DNL below 0.30 LSB and an INL below 0.33 LSB. With a 400MHz full-power input, ENOB is 4.62 bits and 4.40 bits at 5.5GS/s and 7.0GS/s, respectively. When operating at 5.5GHz, the resonant-clock ADC consumes 28mW, achieving a FoM equal to 396fJ

per conversion step. This FoM is lower than all previously published ADCs operating above 2.2GHz. The clock network dissipates only 10.7% of total power; in contrast, clock-related power in typical flash ADCs is roughly 30% of total power. Operating at its maximum sampling frequency of 7.0GS/s, the ADC dissipates 45mW with 11.1% of the total energy consumption attributed to the clock, yielding a FoM of 683fJ per conversion step.

7.3 Future Directions

In this section, we discuss some future possibilities in the implementation of charge-recovery logic and resonant-clocked designs.

Design Automation for Charge-Recovery Logic Designs

An important challenge for charge-recovery logic is the lack of automation tools and automated design flows. Unlike resonant-clocked designs, which are more amenable to the conventional design flow, charge recovery logic designs usually require a longer design process. Logic synthesis is done manually but the result may not be fully optimized for timing. For designs with regular structure, place-and-route may be relatively straight forward; however, for general non-regular designs, it can be a particularly challenge.

Establishing a standard cell library is the first step for automatic charge-recovery design An algorithm is needed to decide the size of a standard cell library and the complexity of functions implemented by logic gates. Since each charge-recovery logic is a pipeline stage, the function complexity of each gate has a strong relation to the number of pipeline stages and system latency.

When performing logic synthesis, another algorithmic technology is needed to automatically trade-off the number of pipeline stages and the number of buffers for clock phase alignment. More stages may reduce transistor sizes in each gate and

length of routing wires, reducing output loading for each cell. However, increasing the number of stages may require more buffers for aligning signals with the correct phase, resulting in increased power consumption and area overhead.

For place-and-route, a tool that balances the capacitive load for a dual-rail output pair would be highly desirable. Specifically, such a place-and-route tool should be capable of adjusting the wire length with the corresponding fan-out capacitance, achieving near-constant loading during switching and reducing clock jitter in the system.

A large inductor database with more configurable options is preferred to provide the target inductance and quality factor, achieving high energy efficient and reducing design effort.

AC-Powered Circuitry

Another promising area for research in charge-recovery is the design of AC-powered circuitry. AC-powered circuitry is a circuit family that can only rely on an AC source for its operation. No other DC voltages/ground is available in this AC-powered circuitry. Since charge-recovery logic makes good use of potential difference in a power-clock, it is potentially suitable for implementing AC-powered circuitry. AC-powered circuitry can improve energy efficiency in digital integrated systems, especially for wireless powering designs. In addition, the inherent use of inductors in wireless powering systems reduces the area overhead of the charge-recovery designs over conventional ones.

Wireless powering techniques based on RF electromagnetic wave propagation have captured the interest of many researchers in applications such as RFID and medical electronics. One of the major goals of these harvesting systems is to convert RF energy into usable DC source. AC-DC/DC-DC converters are usually used for this conversion, but their energy efficiency is not high. In order to improve harvesting

efficiency, many methods have been proposed to improve energy efficiency of embedded AC-DC/DC-DC converters. By designing charge-recovery AC-powered circuitry, these converters can be omitted, and the circuits can be directly powered by the received RF signal. As a result, overall design area decreases and energy efficiency improves, compared to conventional designs

Time-Interleaved ADC with Resonant-Clocking Techniques

A time-interleaved ADC (TIADC) utilizes multiple ADCs in parallel to increase the system sampling rate according to the number of used ADCs. One of the primary issues with time-interleaved structures is timing skew in clock signals, which results in non-uniform sampling of input. Clock buffers are one of the main sources of the clock skew. However, resonant-clocking is a different clocking technique from conventional clocking methods and requires no buffers for clock distribution. With resonant clocking, clock signals are distributed with metal wires, leading to minimum clock skew.

Precise timing control in resonant clock systems is significantly less challenging than in conventional clock systems. In resonant clock systems, pulses with high timing accuracy are needed to drive a resonant clock generator, and the area of a resonant clock generator is substantially smaller than that of an ADC core. Therefore, ensuring precise timing for clock generator pulses is a much more manageable task than distributing a high-accuracy/low-skew clock signal across the entire core with a conventional buffered clock network.

APPENDICES

APPENDIX A

Testing Setup for SBL FIR Test-Chip

This section presents the testing setup for SBL FIR test-chip. Figure A.1 shows the bonding diagram of the SBL FIR test-chip for an LCC84 package. The SBL FIR filter is on the same die with other cores. However, independent supplies are used for each core, and only the ground nets are connected together.

Table A.1 describes the function of each pad. As shown in the table, two pads are used in parallel to connect each power-clock phase and thus reduce the parasitic resistance when SBL FIR is in resonance.

Figure A.2 shows the schematic of the printed circuit board (PCB) used for testing the SBL FIR chip, and Table A.2 shows the parts list on this PCB. 0.1μF ceramic capacitors are used for near-package decoupling, and 10μF ceramic capacitors are used for decoupling near the connections between PCB and supplies. A 48-channel DIO PCI card is used to provide control and configuration signals to the test-chip. A 915 Ohm ferrite bead is used to connect the DIO ground and supply ground, and filters out high frequency noise generated from the DIO card. A Schmitt trigger is used as a level converter for the scan signals from DIO card and provides them with a sharp edge before entering the chip.

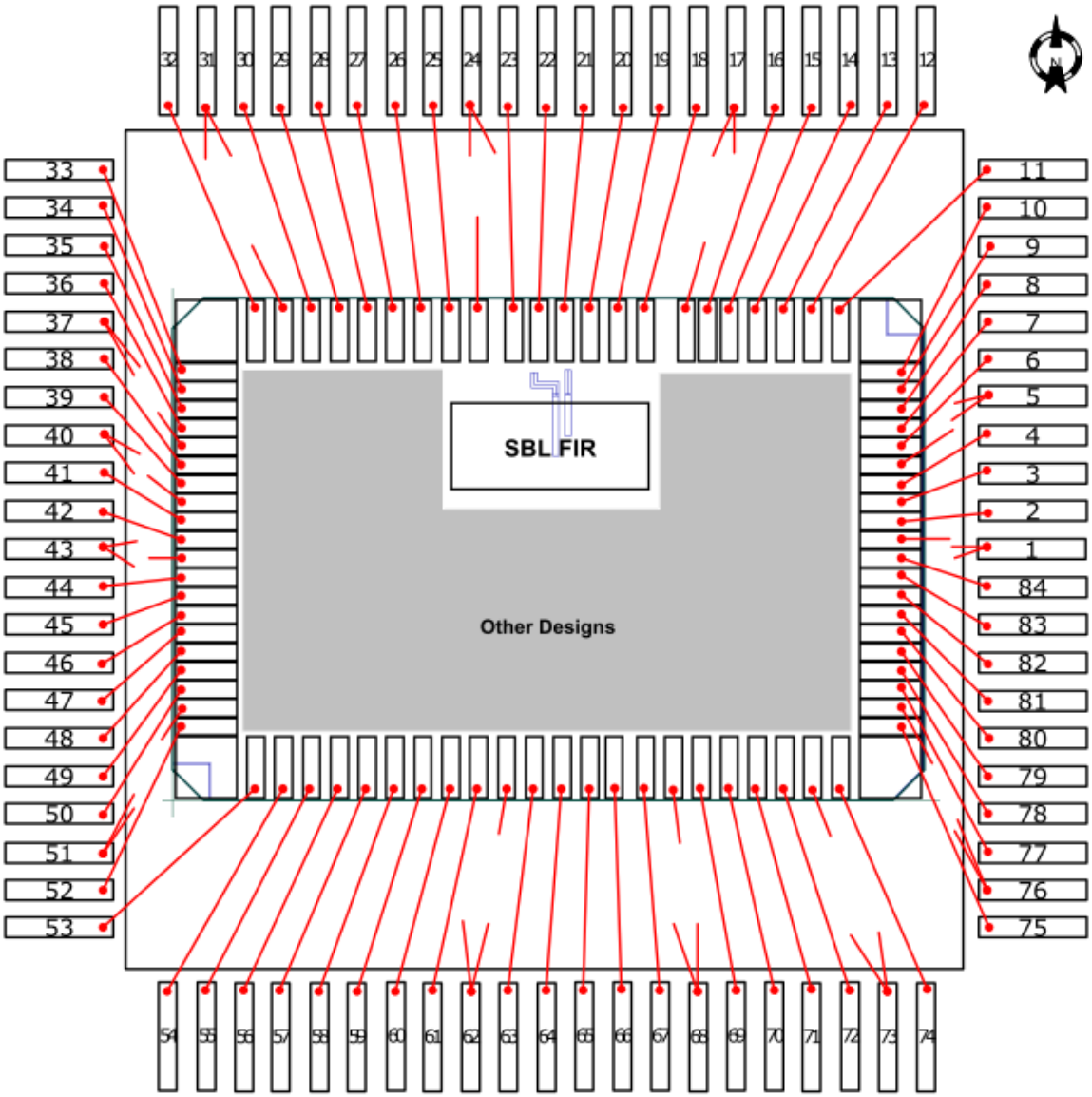


Figure A.1: Bonding diagram for SBL FIR test-chip.

Figure A.3 shows layout and photographs of the front and back sides of the PCB. In this testing, a four-layer PCB is used, and the top and bottom layers are signal layers. The second layer is used as a ground plane, and the third layer is used as a power plane with multiple voltage domains. The power/ground planes reduce coupling effects between signals on top and bottom layers. Figure A.4 shows the setup used to test the SBL FIR chip. Three voltage outputs are used: Dirty V_{DD} (left

Number	Pin Name	P/C/I/O	Comment	Number	Pin Name	P/I/O	Comment
1	VSS	Ground	Chip ground	43	VSS	Ground	Chip ground
2	EE_Clk0_I	Input	Scan clock0	44	VDD_ROA	Power	Other Designs
3	VDC_B1	Power	Other Designs	45	VDD_PG	Power	Pul.Gen. supply
4	EE_Clk1_I	Input	Scan clock1	46	CS_Reset	Input	Other Designs
5	VSS	Ground	Chip ground	47	VCC	Power	VCC to SBL
6	EE_Open_Gate	Input	Scan opengate	48	EBL0_Sig_O	Output	Other Designs
7	EE_Sel_0	Input	Scan sel0	49	VDD	Power	Scan supply
8	EE_Sel_1	Input	Scan sel1	50	EBL0_Fdiv	Output	Other Designs
9	VDD	Power	Scan supply	51	VSS	Ground	Chip ground
10	DVDD	Power	IO supply	52	DVSS	Ground	IO ground
11	DVDD	Power	IO supply	53	DVDD	Power	IO supply
12	VDD	Power	Scan supply	54	VDDA0	Power	Other Designs
13	VDC_B1	Power	Other Designs	55	VPEAK0	VPEAK0	Other Designs
14	CBL_Sig_O	Output	Signature out	56	VDD	Power	Scan supply
15	VDC_B1	Power	Other Designs	57	CS_Sel_0	Input	Other Designs
16	CBL_Fdiv	Output	Freq. Div. out	58	CS_Open_Gate	Input	Other Designs
17	VSS	Ground	Chip ground	59	CS_Clk0_I	Input	Other Designs
18	CBL_mgosc_in	Input	Ext. osc in	60	EBL2_Sig_O	Output	Other Designs
19	VCC	Power	VCC to SBL	61	VDC_B2	Power	Other Designs
20	PC	Clock	PC to SBL	62	VSS	Ground	Chip ground
21	PC	Clock	PC to SBL	63	VDC_B2	Power	Other Designs
22	PC_B	Clock	PC_b to SBL	64	CS_Clk1_I	Input	Other Designs
23	PC_B	Clock	PC_b to SBL	65	VDC_B2	Power	Other Designs
24	VSS	Ground	Chip ground	66	VDD_ROB	Power	Other Designs
25	VDD_PG	Power	Pul.Gen. supply	67	CS_Scan_I	Input	Other Designs
26	VDD	Power	Scan supply	68	VSS	Ground	Chip ground
27	VCC	Power	VCC to SBL	69	EE_Clk0_O	Output	Scan clk0 out
28	VDC_B0	Power	Other Designs	70	VCC	Power	VCC to SBL
29	VSS	Ground	Chip ground	71	VDD	Power	Scan supply
30	VDC_B0	Power	Other Designs	72	CS_Sel_1	Input	Other Designs
31	VSS	Ground	Chip ground	73	VSS	Ground	Chip ground
32	DVSS	Ground	IO ground	74	DVSS	Ground	IO ground
33	DVDD	Power	IO supply	75	DVSS	Ground	IO ground
34	VDD	Power	Scan supply	76	VSS	Ground	Chip ground
35	CS_Scan_O	Output	Other Designs	77	EBL1_Fdiv	Output	Other Designs
36	VDC_B0	Power	Other Designs	78	VCC	Power	VCC to SBL
37	VSS	Ground	Chip ground	79	EBL1_Sig_O	Output	Other Designs
38	VDC_B0	Power	Other Designs	80	VDD	Power	Scan supply
39	EBL2_Fdiv	Output	Other Designs	81	VDDA1	Power	Other Designs
40	VSS	Ground	Chip ground	82	VPEAK1	VPEAK1	Other Designs
41	EE_Reset	Input	Reset	83	EE_Scan_O	Output	Scan out
42	CS_Clk0_O	Output	Other Designs	84	EE_Scan_I	Input	Scan in

Table A.1: I/O information for SBL FIR test-chip.

output of the top supply) is used to supply the I/O pads, static 1.2V (right output of the top supply) is the supply for the BIST circuit, and the subthreshold supply (left

output of the second supply) is shared by V_{CC} and V_{DC} for the SBL FIR filter.

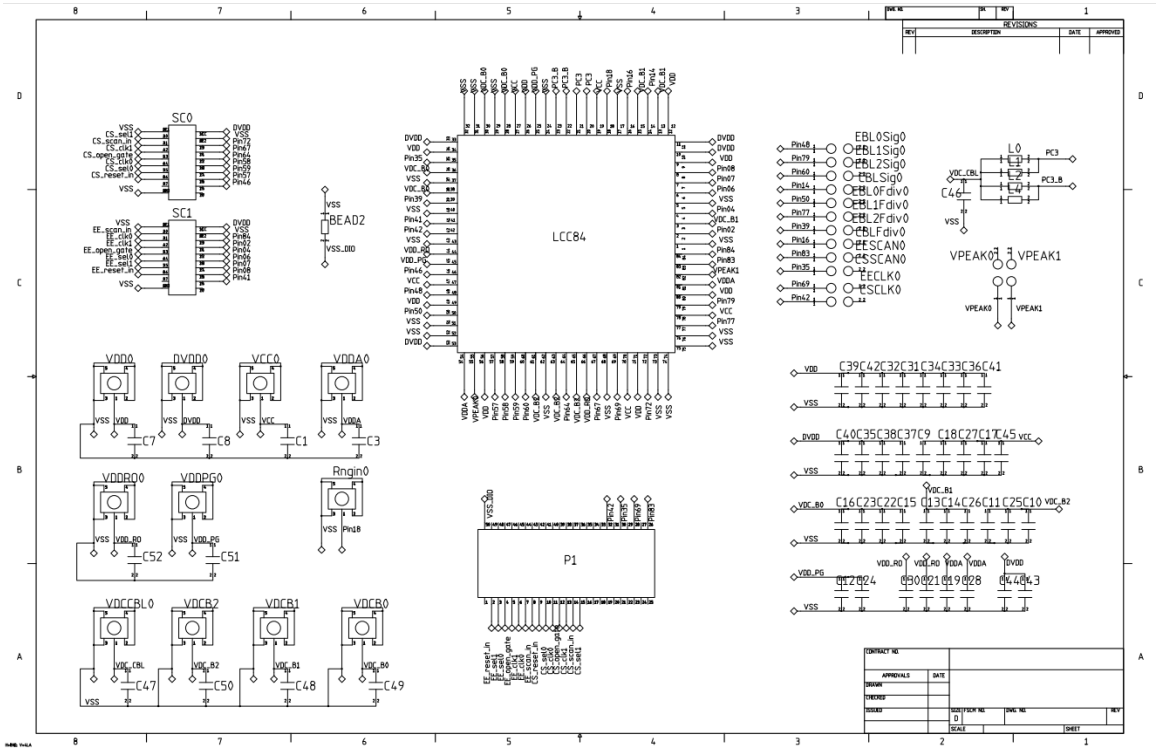


Figure A.2: Schematic of the printed circuit board for SBL FIR test-chip.

Part name	Description	Purpose
LCC84	LCC84 socket	Socket for SBL FIR chip
P1	50Pin header	Connection with DIO card
SC0, SC1	Schmitt trigger	Reset signal bounce suppression
L0, L1, L2, L3	Off-chip inductor	Provide inductance for SBL FIR chip
C1-C17	0.1µF ceramic capacitor	Near package decoupling
C18-C35	10µF ceramic capacitor	Near supply decoupling
BEAD2	915 Ohm ferrite bead	Noise Filter between DIO and PSU ground
VDD0, DVDD0, VCC0, VDDA0, VDDRO0, VDDPG0 RING0, VDCBL0, VDCB2, VDCB1, VDCB0,	SMA connector jack	Off-board PSU connection
EBL0SIG0, EBK1SIG0, EBL2SIG0, CBLSIG0, EBL0FDIV0, EBL1FDIV0, EBL2DFIV0, CBLFDIV0, EESCAN0, CSSCAN0, EECLK0, CSCLK0, VPEAK0, VPLEAK1	Test point	Probing signals

Table A.2: Parts list for SBL FIR test-chip.

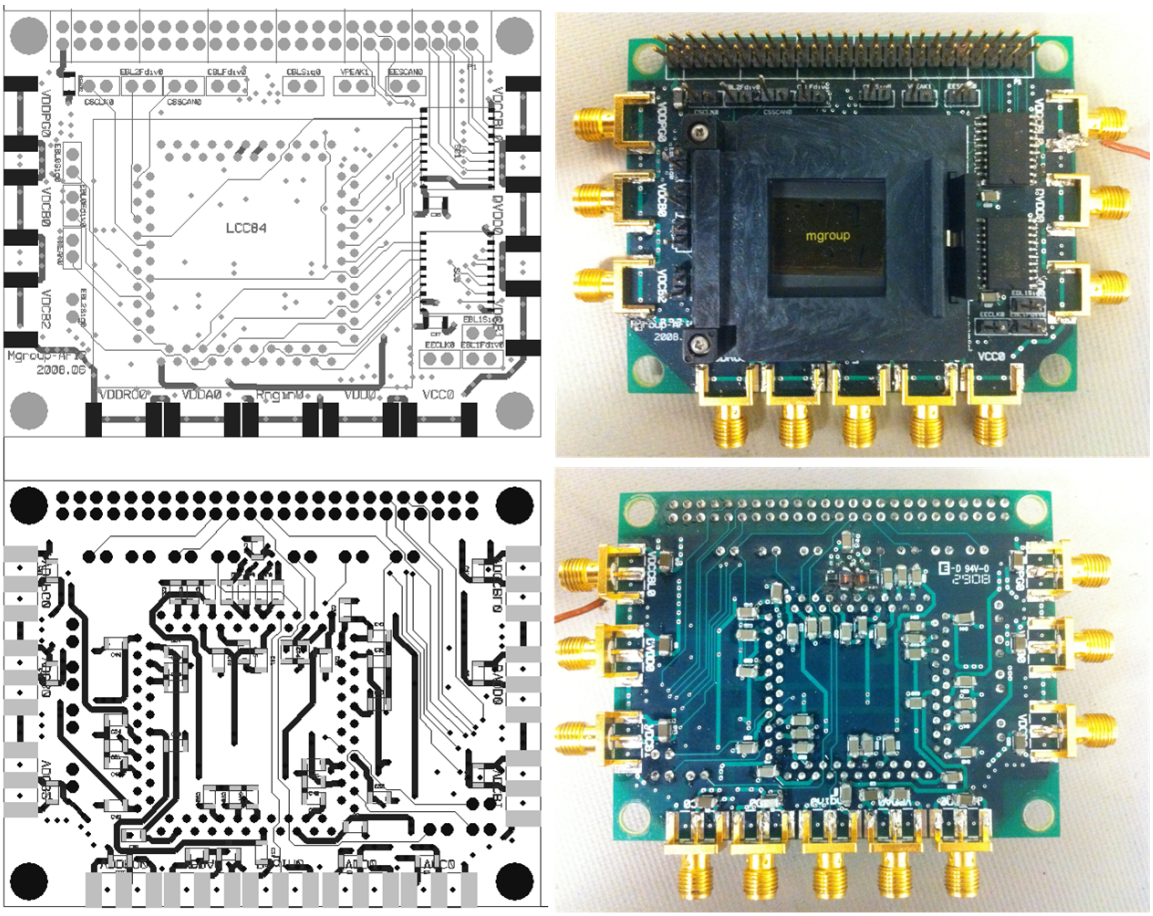


Figure A.3: Printed circuit board for SBL FIR test-chip.

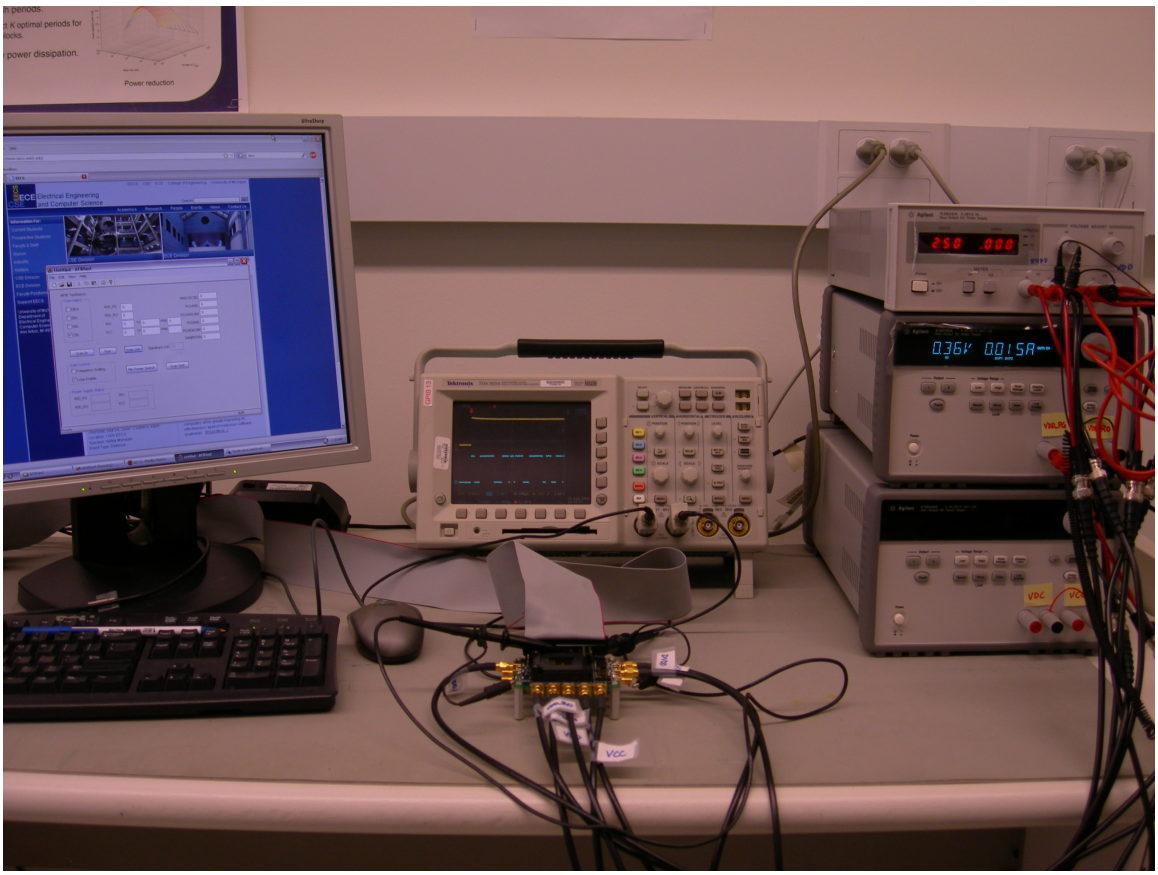


Figure A.4: Test setup for SBL FIR test-chip.

APPENDIX B

Testing Setup for Resonant-Clock Flash ADC Test-Chip

This section shows the test setup for the resonant-clock flash ADC test-chip. Figure B.1 shows the bonding diagram of the resonant-clock ADC chip. A QFN32 package with dimensions 5mm $\times$ 5mm is used to reduce parasitic capacitance and inductance of packages, and thus reduce the signal distortions especially for analog signals. Table B.1 describes functions of each pad. As shown in the table, analog power and ground are specially isolated from other supplies to reduce coupling noise from other power domains.

Figure B.2 shows the schematic of the PCB for the resonant-clock ADC test-chip, and Table B.2 shows the parts list on this PCB. A QFN package is directly soldered on the PCB to avoid parasitic resistance and inductance from a socket. A wide band balun transformer is used to convert single-ended signal to a differential signal, and two 50 Ohm resistors are used to provide needed resistance for impedance matching.

Figure B.3 shows the layout and photographs of the front and back sides of the two-layer PCB. In this testing, a two-layer PCB is used. The top layer is used for signals and different power domains, and the bottom layer is shared with multiple ground domains. These ground domains are connected with 915 Ohm ferrite beads

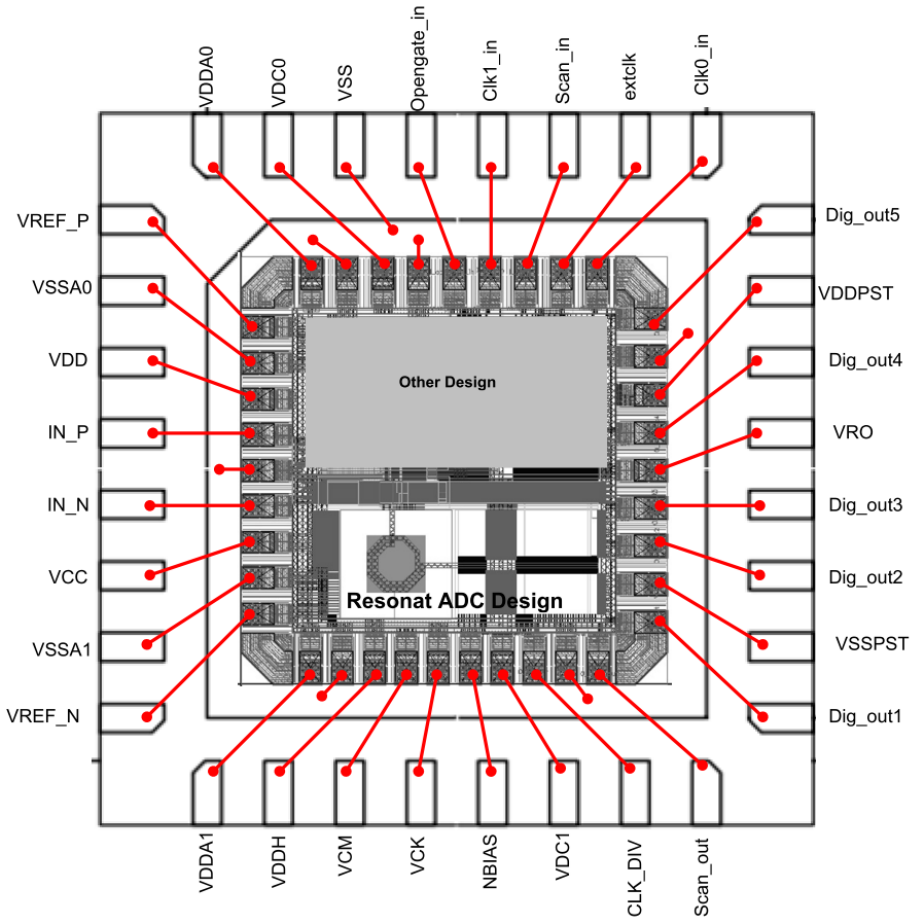


Figure B.1: Bonding diagram for resonant-clock ADC test-chip.

to reduce the coupling noise. The transformer and the matching resistors are placed closely to the QFN package to reduce PCB trace length and other parasitics effects. Figure B.4 shows the setup used to test the resonant-clock ADC chip. Two signal generators (top right) generate a sampling clock and an analog input signal. A logic analyzer (middle) is used to capture the decimated outputs from the resonant-clock ADC test-chip. The testing program is designed in Labview, and parameter sweeping is configurable through the program GUI interface. Frequency and voltage scaling are performed automatically through Labview .

Number	Pin Name	P/C/I/O	Comment
1	Clk0_in	Input	Scan clock0
2	Extclk	Clock	External clock
3	Scan_in	Input	Scan in
4	Clk1_in	Input	Scan clock1
5	Opengate_in	Input	Scan opengate
6	VSS	Ground	Chip ground
7	VDC0	Power	Clkgen bias
8	VDDA0	Power	Analog supply
9	VREF_P	Power	Pos. Ref.
10	VSSA0	Ground	Analog ground
11	VDD	Power	Digital supply
12	IN_P	Input	In+
13	IN_N	Input	In-
14	VCC	Power	Scan supply
15	VSSA1	Ground	Analog ground
16	VREF_N	Power	Neg. Ref.

Number	Pin Name	P/I/O	Comment
17	VDDA1	Power	Analog supply
18	VDDH	Power	THA supply
19	VCM	Power	Com. mode supply
20	VCK	Power	Clkgen supply
21	NBIAS	Power	RingOsc. bias
22	VDC1	Power	Clkgen bias
23	CLK_DIV	Output	Freq. Div. out
24	Scan_out	Output	Scan out
25	Dig_out1	Output	Decimator out
26	VSSPST	Ground	IO ground
27	Dig_out2	Output	Decimator out
28	Dig_out3	Output	Decimator out
29	VRO	Power	RingOsc. Supply
30	Dig_out4	Output	Decimator out
31	VDDPST	Power	IO supply
32	Dig_out5	Output	Decimator out

Table B.1: I/O information for resonant-clock ADC test-chip.

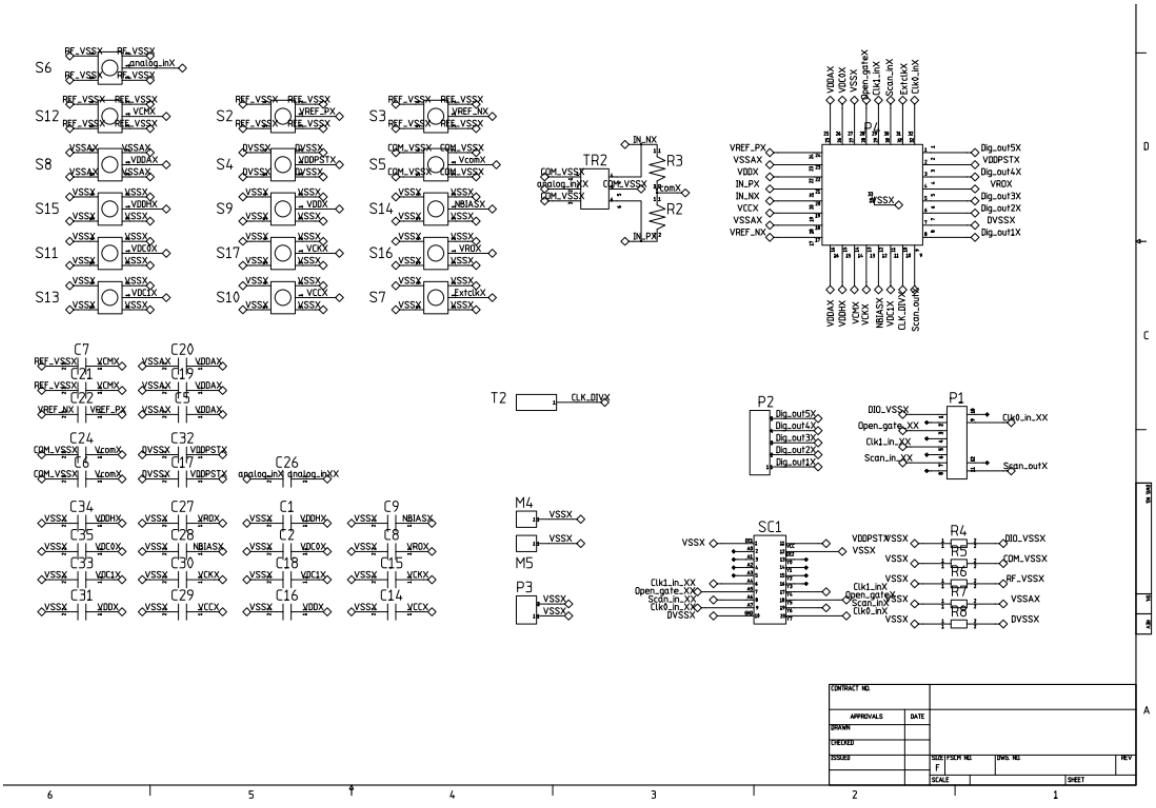


Figure B.2: Schematic of the printed circuit board for resonant-clock ADC test-chip.

Part name	Description	Purpose
P1	12Pin header	Connection with DIO card
P2	5Pin header	Connection with Logic Analyzer
SC1	Schmitt trigger	Rest signal bounce suppression
S1-S16	SMA connector jack	Off-board PSU connection
C1-C13	0.1µ ceramic capacitor	Near chip decoupling
C14-C27	10µ ceramic capacitor	Near supply decoupling
TR2	Transformer	Transform single ended signal to differential signals
R2, R3	50 Ohm resistor	Matching resistors
T2	Test point	Probing point for clock divided out
R4-R8	915 Ohm ferrite beads	Noise Filter between different ground domains

Table B.2: Parts list for resonant-clock ADC test-chip.

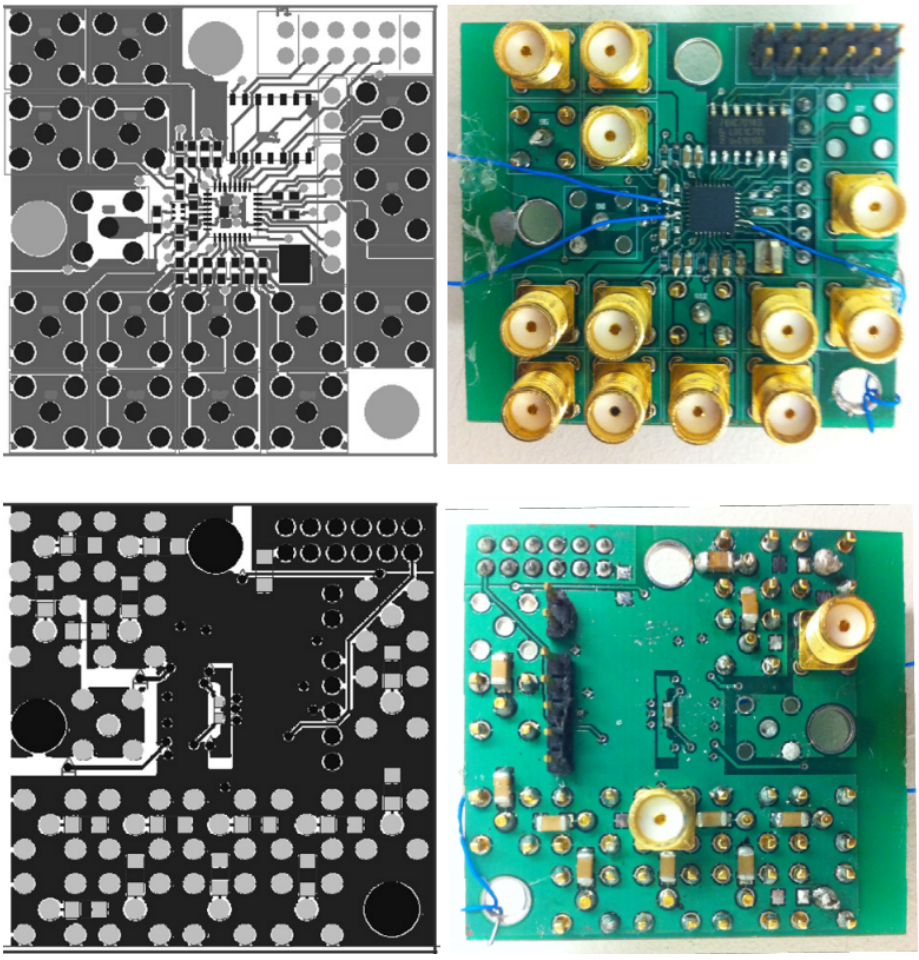


Figure B.3: Printed circuit board for resonant-clock ADC test-chip.

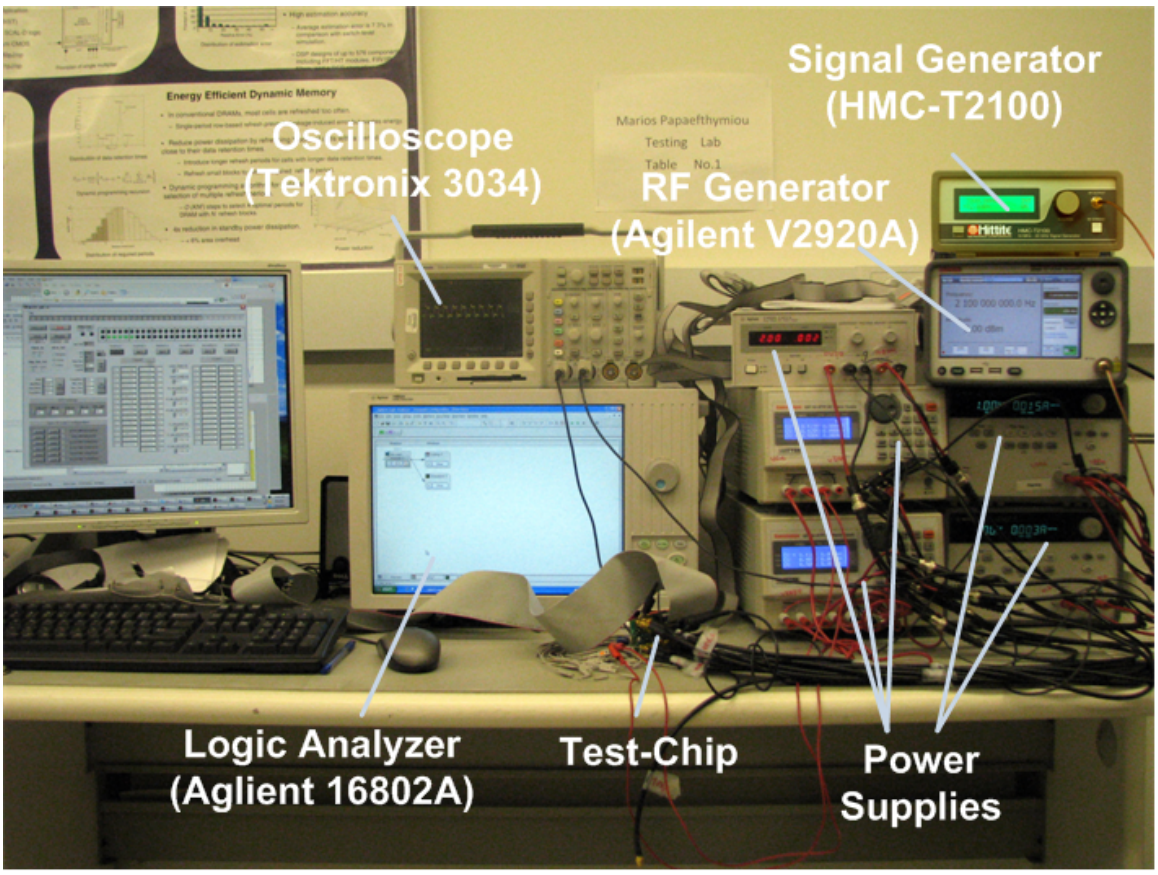


Figure B.4: Test setup for resonant-clock ADC test-chip.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] A. Ishii, J. Kao, V. Sathe, and M. Papaefthymiou, "A Resonant-Clock 200MHz ARM926EJ-S^{TM} Microcontroller," in *European Solid-State Circuits Conference, ESSCIRC*, pp. 356–359, September 2009.

[2] http://www.stanford.edu/∼murmann/adcsurvey.html.

[3] G. Moore, "Progress in Digital Integrated Electronics," in *International Electron Devices Meeting*, vol. 21, pp. 11–13, 1975.

[4] B. Zhai, S. Pant, L. Nazhandali, S. Hanson, J. Olson, A. Reeves, M. Minuth, R. Helfand, T. Austin, D. Sylvester, and D. Blaauw, "Energy-Efficient Subthreshold Processor Design," *IEEE Transactions on Very Large Scale Integration Systems, TVLSI*, vol. 17, pp. 1127–1137, August 2009.

[5] S. Hanson, M. Seok, Y.-S. Lin, Z. Y. Foo, D. Kim, Y. Lee, N. Liu, D. Sylvester, and D. Blaauw, "A Low-Voltage Processor for Sensing Applications with Picowatt Standby Mode," *Journal of Solid-State Circuits, JSSC*, vol. 44, pp. 1145–1155, April 2009.

[6] S. Hanson, Z. Y. Foo, D. Blaauw, and D. Sylvester, "A 0.5V Sub-Microwatt CMOS Image Sensor with Pulse-Width Modulation Read-Out," *Journal of Solid-State Circuits, JSSC*, vol. 45, pp. 759–767, April 2010.

[7] S. Kim and M. Papaefthymiou, "Single-Phase Source-Coupled Adiabatic Logic," in *International Symposium on Low Power Electronics and Design, ISLPED*, pp. 97–99, August 1999.

[8] K. S. and M. Papaefthymiou, "True Single-Phase Adiabatic Circuitry," *IEEE Transactions on Very Large Scale Integration Systems, TVLSI*, vol. 9, pp. 52–63, February 2001.

[9] C. Ziesler, J. Kim, V. Sathe, and M. Papaefthymiou, "A 225 MHz Resonant Clocked ASIC Chip," in *International Symposium on Low Power Electronics and Design, ISLPED*, pp. 48–53, August 2003.

[10] C. Ziesler, J. Kim, and M. Papaefthymiou, "Energy Recovering ASIC Design," in *IEEE Computer Society Annual Symposium on VLSI, ISVLSI*, pp. 133–138, February 2003.

[11] A. Bargagli-Stoffi, G. Iannaccone, S. Di Pascoli, E. Amirante, and D. Schmitt-Landsiedel, "Four-Phase Power Clock Generator for Adiabatic Logic Circuits," *Electronics Letters*, vol. 38, pp. 689–690, July 2002.

[12] A. Kramer, J. S. Denker, B. Flower, and J. Moroney, "2nd Order Adiabatic Computation with 2N-2P and 2N-2N2P Logic Circuits," in *International Symposium on Low Power Electronics and Design, ISLPED*, pp. 191–196, August 1995.

[13] R. Staszewski, K. Muhammad, and P. Balsara, "A 550-MSample/s 8-Tap FIR Digital Filter for Magnetic Recording Read Channels," *Journal of Solid-State Circuits, JSSC*, vol. 35, pp. 1205–1210, August 2000.

[14] V. Sathe, J. Kao, and M. Papaefthymiou, "Resonant-Clock Latch-Based Design," *Journal of Solid-State Circuits, JSSC*, vol. 43, pp. 864–873, April 2008.

[15] S. Chan, P. Restle, T. Bucelot, J. Liberty, S. Weitzel, J. Keaty, B. Flachs, R. Volant, P. Kapusta, and J. Zimmerman, "A Resonant Global Clock Distribution for the Cell Broadband Engine Processor," *Journal of Solid-State Circuits, JSSC*, vol. 44, pp. 64–72, January 2009.

[16] J. Kao, W.-H. Ma, S. Kim, and M. Papaefthymiou, "2.07 GHz Floating-Point Unit with Resonant-Clock Precharge Logic," in *Asian Solid State Circuits Conference, A-SSCC*, pp. 213–216, November 2010.

[17] W.-H. Ma, J. C. Kao, V. S. Sathe, and M. Papaefthymiou, "A 187MHz Subthreshold-Supply Robust FIR Filter with Charge-Recovery Logic," in *Symposium on VLSI Circuits*, pp. 202–203, June 2009.

[18] W.-H. Ma, J. Kao, V. Sathe, and M. Papaefthymiou, "187 MHz Subthreshold-Supply Charge-Recovery FIR," *Journal of Solid-State Circuits, JSSC*, vol. 45, pp. 793–803, April 2010.

[19] W.-H. Ma, J. Kao, and M. Papaefthymiou, "A 5.5GS/s 28mW 5-bit Flash ADC with Resonant Clock Distribution," in *European Solid-State Circuits Conference, ESSCIRC*, September 2011.

[20] R. P. Feynman, *Feynman Lectures on Computation*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1998.

[21] R. Landauer, "Irreversibility and Heat Generation in the Computing Process," *IBM Journal of Research and Development*, vol. 5, pp. 183–191, July 1961.

[22] C. H. Bennett, "Logical Reversibility of Computation," *IBM Journal of Research and Development*, vol. 17, pp. 525–532, November 1973.

[23] W. Pan and M. Nalasani, "Reversible Logic," *IEEE Potentials*, vol. 24, pp. 38–41, February 2005.

[24] R. P. Feynman, "Quantum Mechanical Computers," in *Foundations of Physics*, vol. 16, pp. 507–531, June 1986.

[25] W. Athas and L. Svensson, "Reversible Logic Issues in Adiabatic CMOS," in *Workshop on Physics and Computation*, pp. 111–118, November 1994.

[26] S. G. Younis, "Asymptotically Zero Energy Computing using Split-Level Charge Recovery Logic," in *Ph.D thesis, Massachusetts Institute of Technology*, (Cambridge, MA, USA), 1994.

[27] J. Lim, D.-G. Kim, and S.-I. Chae, "A 16-bit Carry-Lookahead Adder using Reversible Energy Recovery Logic for Ultra-Low-Energy Systems," *Journal of Solid-State Circuits, JSSC*, vol. 34, pp. 898–903, June 1999.

[28] A. Dickinson and J. Denker, "Adiabatic Dynamic Logic," in *IEEE Custom Integrated Circuits Conference, CICC*, pp. 282–285, May 1994.

[29] A. Dickinson and J. Denker, "Adiabatic Dynamic Logic," *Journal of Solid-State Circuits, JSSC*, vol. 30, pp. 311–315, March 1995.

[30] Y. Ye and K. Roy, "QSERL: Quasi-Static Energy Recovery Logic," *Journal of Solid-State Circuits, JSSC*, vol. 36, pp. 239–248, February 2001.

[31] V. Oklobdzija, D. Maksimovic, and F. Lin, "Pass-Transistor Adiabatic Logic using Single Power-Clock Supply," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 44, pp. 842–846, October 1997.

[32] A. Vetuli, S. Pascoli, and L. Reyneri, "Positive Feedback in Adiabatic Logic," *Electronics Letters*, vol. 32, pp. 1867–1869, September 1996.

[33] S. Kim and M. Papaefthymiou, "True Single-Phase Energy-Recovering Logic for Low-Power, High-Speed VLSI," in *International Symposium on Low Power Electronics and Design, ISLPED*, pp. 167–172, August 1998.

[34] V. S. Sathe, J.-Y. Chueh, and M. C. Papaefthymiou, "Energy-Efficient Ghz-Class Charge-Recovery Logic," *Journal of Solid-State Circuits, JSSC*, vol. 42, pp. 38–47, January 2007.

[35] W. Athas, N. Tzartzanis, L. Svensson, and L. Peterson, "A Low-Power Microprocessor Based on Resonant Energy," *Journal of Solid-State Circuits, JSSC*, vol. 32, pp. 1693–1701, November 1997.

[36] M. Hansson, B. Mesgarzadeh, and A. Alvandpour, "1.56 GHz On-Chip Resonant Clocking in 130nm CMOS," in *IEEE Custom Integrated Circuits Conference, CICC*, pp. 241–244, September 2006.

[37] C. J-Y, V. Sathe, and M. Papaefthymiou, "900MHz to 1.2GHz Two-Phase Resonant Clock Network with Programmable Driver and Loading," in *IEEE Custom Integrated Circuits Conference, CICC*, pp. 777–780, September 2006.

[38] V. Sathe, J. Kao, and M. Papaefthymiou, "RF2: A 1GHz FIR Filter with Distributed Resonant Clock Generator," in *Symposium on VLSI Circuits*, pp. 44–45, June 2007.

[39] V. Sathe, J. Kao, and M. Papaefthymiou, "A 0.8-1.2GHz Single-Phase Resonant-Clocked FIR Filter with Level-Sensitive Latches," in *IEEE Custom Integrated Circuits Conference, CICC*, pp. 583–586, September 2007.

[40] A. Wang and A. Chandrakasan, "A 180mV Subthreshold FFT Processor using a Minimum Energy Design Methodology," *Journal of Solid-State Circuits, JSSC*, vol. 40, pp. 310–319, January 2005.

[41] B. Zhai, L. Nazhandali, J. Olson, A. Reeves, M. Minuth, R. Helfand, S. Pant, D. Blaauw, and T. Austin, "A 2.60pJ/Inst Subthreshold Sensor Processor for Optimal Energy Efficiency," in *Symposium on VLSI Circuits*, pp. 154–155, June 2006.

[42] M. Seok, S. Hanson, Y.-S. Lin, Z. Foo, D. Kim, Y. Lee, N. Liu, D. Sylvester, and D. Blaauw, "The Phoenix Processor: A 30pW Platform for Sensor Applications," in *Symposium on VLSI Circuits*, pp. 188–189, June 2008.

[43] J.-S. Wang, J.-S. Cehn, Y.-M. Wang, and C. Yeh, "A 230mV-to-500mV 375KHz-to-16MHz 32b RISC Core in 0.18μm CMOS," in *International Solid-State Circuits Conference, ISSCC*, pp. 294–604, February 2007.

[44] M.-E. Hwang, A. Raychowdhury, K. Kim, and K. Roy, "A 85mV 40nW Process-Tolerant Subthreshold 8×8 FIR Filter in 130nm Technology," in *Symposium on VLSI Circuits*, pp. 154–155, June 2007.

[45] J. Kil, J. Gu, and C. Kim, "A High-speed Variation-Tolerant Interconnect Technique for Sub-Threshold Circuits Using Capacitive Boosting," in *International Symposium on Low Power Electronics and Design, ISLPED*, pp. 67–72, August 2006.

[46] M. Seok, D. Jeon, C. Chakrabarti, D. Blaauw, and D. Sylvester, "A 0.27V 30MHz 17.7nJ/Transform 1024-pt Complex FFT Core with Super-Pipelining," in *International Solid-State Circuits Conference, ISSCC*, pp. 342–344, February 2011.

[47] W. Athas, L. Svensson, and N. Tzartzanis, "A Resonant Signal Driver for Two-Phase, Almost-non-Overlapping Clocks," in *International Symposium on Circuits and Systems, ISCAS*, vol. 4, pp. 129–132 vol.4, May 1996.

[48] B. Verbruggen, J. Craninckx, M. Kuijk, P. Wambacq, and G. Van der Plas, "A 2.2 mW 1.75 GS/s 5 Bit Folding Flash ADC in 90 nm Digital CMOS," *Journal of Solid-State Circuits, JSSC*, vol. 44, pp. 874–882, March 2009.

[49] G. Geelen and E. Paulus, "An 8b 600MS/s 200mW CMOS Folding A/D Converter using an Amplifier Preset Technique," in *International Solid-State Circuits Conference, ISSCC*, pp. 254–526, February 2004.

[50] R. Taft, C. Menkus, M. Tursi, O. Hidri, and V. Pons, "A 1.8-V 1.6-GSample/s 8-b Self-Calibrating Folding ADC with 7.26 ENOB at Nyquist Frequency," *Journal of Solid-State Circuits, JSSC*, vol. 39, pp. 2107–2115, December 2004.

[51] M. El-Chammas and B. Murmann, "A 12-GS/s 81-mW 5-bit Time-Interleaved Flash ADC with Background Timing Skew Calibration," *Journal of Solid-State Circuits, JSSC*, vol. 46, pp. 838–847, April 2011.

[52] S. Park, Y. Palaskas, and M. Flynn, "A 4GS/s 4-bit Flash ADC in 0.18μm CMOS," *Journal of Solid-State Circuits, JSSC*, vol. 42, pp. 1865 –1872, September 2007.

[53] M. Choi, J. Lee, J. Lee, and H. Son, "A 6-bit 5-GSample/s Nyquist A/D Converter in 65nm CMOS," in *Symposium on VLSI Circuits*, pp. 16–17, June 2008.

[54] C. Paulus, H.-M. Bluthgen, M. Low, E. Sicheneder, N. Bruls, A. Courtois, M. Tiebout, and R. Thewes, "A 4GS/s 6b Flash ADC in 0.13μm CMOS," in *Symposium on VLSI Circuits*, pp. 420–423, June 2004.

[55] X. Jiang, Z. Wang, and M. Chang, "A 2 GS/s 6 b ADC in 0.18μm CMOS," in *International Solid-State Circuits Conference, ISSCC*, pp. 322–497, February 2003.

[56] D. Wei, D. Sun, and A. Abidi, "A 300MHz Mixed-Signal FDTS/DFE Disk Read Channel in 0.6μm CMOS," in *International Solid-State Circuits Conference, ISSCC*, pp. 186–187, February 2001.

[57] M. Pelgrom, A. Duinmaijer, and A. Welbers, "Matching Properties of MOS Transistors," *Journal of Solid-State Circuits, JSSC*, vol. 24, pp. 1433–1439, October 1989.

[58] K. Uyttenhove and M. Steyaert, "A 1.8-V 6-bit 1.3-GHz Flash ADC in 0.25-μm CMOS," *Journal of Solid-State Circuits, JSSC*, vol. 38, pp. 1115–1122, July 2003.

[59] S.-M. Yim, T. Chen, and K. O, "The Effects of a Ground Shield on the Characteristics and Performance of Spiral Inductors," *Journal of Solid-State Circuits, JSSC*, vol. 37, pp. 237–244, Feburary 2002.

[60] K. Deguchi, N. Suwa, M. Ito, T. Kumamoto, and T. Miki, "A 6-bit 3.5-GS/s 0.9-V 98-mW Flash ADC in 90-nm CMOS," *Journal of Solid-State Circuits, JSSC*, vol. 43, pp. 2303–2310, October 2008.

[61] H. Chung, A. Rylyakov, Z. T. Deniz, J. Bulzacchelli, G.-Y. Wei, and D. Friedman, "A 7.5-GS/s 3.8-ENOB 52-mW Flash ADC with Clock Duty Cycle Control in 65nm CMOS," in *Symposium on VLSI Circuits*, pp. 268–269, June 2009.