

# **Breaking the Power Delivery Walls using Digitally-controlled Integrated Regulation**

---

A Dissertation

Presented to the Faculty of  
The School of Engineering and Applied Science  
University of Virginia

---

In Partial Fulfillment  
of the Requirements for the Degree of  
Doctor of Philosophy in Electrical Engineering

**Kaushik Mazumdar**

August 2015

# APPROVAL SHEET

This dissertation  
Is submitted in partial fulfillment of the requirements  
for the degree of  
Doctor of Philosophy



This dissertation has been read and approved by the examining committee:

Mircea Stan, Ph.D.  
\_\_\_\_\_  
Advisor

John Lach, Ph.D.  
\_\_\_\_\_

Archie Holmes, Ph.D.  
\_\_\_\_\_

Ben Calhoun, Ph.D.  
\_\_\_\_\_

Kevin Skadron, Ph.D.  
\_\_\_\_\_

Accepted for the School of Engineering and Applied Science



Dean, School of Engineering and Applied Science

August 2015

© August 2015

Kaushik Mazumdar

All rights reserved

## Abstract

---

Integrated power regulation has become an essential tool in the arsenal of techniques that allow the semiconductor industry to continue Moore's law of exponential integration in sub-20nm CMOS technology. While form factor constrains the quality of the integrated passives, switching losses and quiescent current consumption of the regulators limit their usage to coarse-grained power management. Another significant design trend in these nano-scale nodes is the increasing use of digitally-assisted analog solutions to leverage the superior switching characteristics of CMOS technology. The introduction of 3D stacked memory has renewed interest in vertical integration in the form of 3D-IC design. However, the predicted 3D-IC scaling, from dual-layer to many-layers, is stalled by the more fundamental "3D versus 2D" power mismatch in the power delivery network (PDN). While load current increases with additional vertical layers, the 2D surface area for delivering power and the power-bump numbers do not scale proportionally, creating a power delivery mismatch or "wall".

This dissertation focuses on digitally-assisted circuit/architecture to improve integrated power regulation beyond state-of-the-art for system-on-chip design in nano-scale CMOS process nodes in both 2D and 3D-IC circuits. The primary contributions of this work are:

(1) Design and implementation of digitally-controlled low-dropout regulators with improved figure-of-merit (FOM), catering to a wide range of applications. The first “truly” hybrid IVR architecture is proposed, regulating a graphics core with 50% reduced voltage droop. Additionally, a digitally-adaptive LDO topology is proposed, using system power-modes information to regulate quiescent current loss in energy-harvesting architecture, achieving a FOM of 4.44ps.

(2) Cross-layer design explorations of multi-output switched-capacitor-assisted charge-recycled power regulation (also known as Voltage Stacking) to break the power delivery walls in 2D and 3D-IC. Voltage stacking (V-S), with its differential regulation, improves average power efficiency to more than 90% with superior power density. This claim has been validated with simulations using power-traces from architectural benchmarks (Parsec) and proof-of-concept experiments with FPGA chips and a fabricated switched-capacitor converter (SCVR). Another major focus of this dissertation involves an in-depth analysis of 3D PDN design with V-S. The first many-layered 3D PDN model (3D-Voltspot), with V-S and differential SCVR, has been developed in a collaborative effort to perform wide-range PDN tradeoff studies, characterizing V-S noise and system power efficiency with varying workloads.

## **Acknowledgement**

---

There are many people who have helped, supported and guided me to reach this juncture in my life and this thesis will be incomplete without offering my gratitude to them.

Firstly, I would like to express my gratitude to Prof. Mircea Stan for giving me the opportunity to work in his lab and guiding me through my PhD. I am extremely thankful for the freedom he provided me to pursue the research way I liked. His vast expanse of knowledge and his constant inquisitiveness to learn the unknown has always impressed and motivated me. He has contributed immensely in making grad school a great learning experience for me.

I will like to thank all my committee members for providing me with useful suggestions to improve my dissertation. Professor Ben Calhoun in particular, has immense contribution in shaping up my dissertation with his insightful comments during my proposal defense. I had the pleasure of collaborating with Prof. Kevin Skadron for the past one year and it has been an interesting venture combining architectural strategies with circuit techniques. I am grateful to Prof. Archie Holmes and Prof. John Lach for agreeing to be part of my defense committee and helping me out with the dissertation, especially the writing part.

During my PhD there were many instances where I felt something was not possible to achieve within the deadline and my lab mates were always there to pitch in and make the impossible possible. I am indebted to Xinfei for helping me out during the 130nm tapeout, staying up as late as 5 in the morning while debugging layout errors. My

collaboration with Runjie has been really fruitful, providing us with good publications and helping me understand the architectural design tradeoffs. I am thankful to Mehdi and Ben for the long discussions that we had on circuit modeling, voltage stacking or even on basic fundamentals of electronics. Beyond HPLP (my lab group), fellow colleagues from Prof. Calhoun's group, namely Alicia, Arijit, Divya, Aatmesh have helped me time to time on numerous things and in their own small ways have contributed to this final dissertation. I am especially grateful to Sudhanshu for his mentoring early in my grad school and later on during my internship at Texas Instruments.

I learnt a lot during my two internships, Intel in spring 2012 and TI during summer 2013. I was able to imbibe some of the design methodology and practices that ensure the success of design in my project as well. I am indebted to my managers James Tschanz (Intel) and Steven Bartling (TI) for giving me an opportunity to work in their group. I am also thankful to my mentor Muhammad Khellah at Intel whose guidance and motivation has been an inspiration for me all these years.

Life in Charlottesville would not have been this pleasant and eventful if not for my friends. I have been lucky enough to have a wide range of friends, whether the "bong group" or the pan-india group of "cho-junta" in Charlottesville. My best friends since school days (Krishnakali, Sreyoshi and Binit) who have witnessed and been part of this journey have been a major source of support and encouragement whenever the stress of grad school got to me. It is amazing to see how childhood friendship can blossom into such a strong bond even when thousands of miles apart.

Finally, I will be ever grateful to my parents for believing in me and supporting me in every endeavor that I undertook. They always ensured the best for me and gave me

all the freedom to do what I liked the most. My mom has probably dreamt more about my PhD than I have ever and it gives me a lot of joy to fulfil her dream. Bordi and Indrada have taken up the roles of parenting me here in USA and has been a big pillar of support while playing with rishi and ashmi has given me endless joy. My grad school story at Uva will be incomplete without acknowledging the role of my fiancé, Nivi. She is the best thing that has happened to me in a long, long time and I will always cherish our memories at Charlottesville. She had to bear through the worst of my tantrums, either during the stressful job-hunting days or during the chip tape-outs, yet made me smile at the end of the day.

Thank you

# Table of contents

---

<i>Abstract</i> .....	4
<i>Table of contents</i> .....	9
<i>Table of figures</i> .....	15
<i>List of tables</i> .....	21
<i>Chapter 1: Introduction</i> .....	22
<hr/>	
1.1. Introduction .....	22
1.1.1. Why integrated VRM.....	24
1.1.1.1. Types of integrated VRM .....	26
1.1.2. Why digitally-controlled techniques.....	27
1.1.3. Power delivery wall in 3D-IC .....	28
1.2. Organization.....	30
1.2.1. Background.....	30
1.2.2. Digitally-controlled low-dropout regulator.....	30

1.2.3.	Charge-recycled power regulation with integrated switched-capacitor.....	31
1.2.4.	Breaking 3D-IC power delivery walls using voltage-stacking .....	32
<i>Chapter 2: Overview of IVR .....</i>		<i>33</i>
<hr/>		
2.1.	Background .....	33
2.1.1.	Overview of IVR.....	33
2.1.2.	Low-dropout regulator (LDO) .....	35
2.1.2.1.	Working principle and figure-of-merits .....	35
2.1.3.	Fully-integrated switching converters.....	37
2.1.3.1.	Capacitive converter .....	37
2.1.3.2.	Inductive converter .....	39
2.1.4.	Overview of integrated passives .....	40
2.1.4.1.	Technology options for SC converter.....	40
2.1.4.2.	Technology options for integrated inductive converter.....	42
2.1.5.	Comparisons between integrated SC converter and inductive converter ...	43
<hr/>		
<i>Chapter 3: Digitally-Controlled Low-Dropout Regulator .....</i>		<i>48</i>
<hr/>		
3.1.	Digitally-controlled low-dropout regulator .....	48
3.2.	Background and comparisons with A-LDO.....	49

3.3.	Digitally-controlled LDO for reduced standby power-drain in ULP MCU.....	53
3.3.1.	Motivation.....	53
3.3.2.	Digitally-adaptive LDO architecture with ultra-low quiescent current .....	56
3.3.3.	D-LDO design components .....	58
3.3.4.	D-LDO simulation, implementation and measurements .....	66
3.3.5.	Comparison with state-of-art LDO .....	71
3.3.6.	Stability analysis .....	73
3.3.7.	Future scope of improvements.....	76
3.3.8.	Digitally-controlled LDO for hybrid IVR with fast droop mitigation.....	77
3.3.9.	Background.....	77
3.3.10.	Digitally-controlled dual-loop LDO with fast droop mitigation technique	79
3.3.11.	D-LDO implementation and measurement.....	82
3.3.11.1.	Adaptive bit-gain modulation.....	84
3.3.11.2.	Transient performance : simulation and measurement results .....	87
3.3.12.	State-of-art comparisons .....	88
3.4.	Summary .....	89

*Chapter 4: Charge-recycled Power Regulation with Integrated Switched-Capacitor .....* 91

---

4.1.	Charge-recycled power regulation with integrated switched-capacitor.....	91
------	--	----

4.1.1.	Background on voltage stacking – benefits and challenges.....	92
4.2.	Voltage stacking with stacked SC converters .....	94
4.2.1.	Switched-capacitor design and optimization .....	97
4.2.1.1.	Push-pull SC converter for V-S.....	97
4.2.2.	SC converter modeling and design optimization .....	100
4.2.2.1.	SC converter output impedance and power-loss modeling .....	100
4.2.2.2.	SC converter power optimization .....	103
4.2.3.	Higher efficiency in V-S.....	106
4.2.4.	SC versus linear regulator .....	108
4.2.5.	Open-loop versus closed-loop V-S regulation.....	111
4.3.	Voltage stacking measurements.....	118
4.4.	Summary and comparisons with state-of-art integrated converters.....	125

*Chapter 5: Breaking the 3D-IC Power Delivery Wall using Voltage Stacking*

.....	128
-------	-----

---

5.1.	Breaking the 3D-IC power delivery wall using voltage stacking .....	128
5.1.1.	Power delivery and heat removals walls for 3D-IC.....	129
5.1.2.	3D-IC literature review .....	131
5.1.3.	Multi-output switched-capacitor assisted V-S in 3D-IC.....	132
5.2.	System-level evaluation of V-S.....	134

5.2.1.	Modeling methodologies .....	135
5.2.1.1.	SC Converter : impedance model.....	135
5.2.2.	SCVR impedance model validation.....	137
5.2.3.	Transient model for SC converter.....	138
5.2.3.1	SCVR transient model validation .....	141
5.2.4.	Whole-system model .....	142
5.2.5.	Simulation setup.....	144
5.2.5.1.	Many-core 3D modeling.....	144
5.2.5.2.	PDN modeling with different TSV configurations.....	145
5.2.5.3.	Workload modeling.....	147
5.3.	Cross-layer design exploration in voltage stacked many-layer 3D-IC .....	147
5.3.1.	Load imbalance-induced voltage noise.....	147
5.3.2.	Cross-layer noise interference.....	150
5.3.3.	Circuit/architectural co-simulation using SPICE.....	153
5.3.4.	System power efficiency.....	155
5.3.4.1.	System power efficiency with workload imbalance.....	155
5.3.5.	V-S implementation in 3D-IC technology.....	156
5.3.5.1.	3D-IC Through-Silicon-Via .....	157
(a)	TSV Induced asymmetry in voltage stacked 3D-IC.....	157
5.4.	Future work and summary.....	159

*Chapter 6: Summary*.....162

---

6.1. Summary ..... 162

6.2. Future work ..... 164

    6.2.1. Digitally-controlled LDO..... 164

    6.2.2. Charge-recycled power regulation with integrated switched-capacitor.... 164

    6.2.3. Breaking 3D-IC power delivery wall using voltage stacking..... 164

*Glossary*.....165

*Bibliography* .....166

---

## Table of figures

---

Figure 1.1. Intel Haswell processor, moving off-chip VR to on-chip with passives-in-package (SiP) (8).....	24
Figure 1.2. Per-core voltage saves energy wasted in shared rail (34).....	26
Figure 1.3. Integration in the vertical direction or 3D-IC (98).....	29
Figure 2.1. Mixed-signal PMU with switching regulators providing global (off-chip) regulation and LDO providing local (on-chip) regulation.....	34
Figure 2.2. Low-dropout regulator popular for point-of-load regulation in SoC .....	35
Figure 2.3. 2:1 topology of switched-capacitor converter (SCVR) using fly-cap and switches.....	38
Figure 2.4. Synchronous buck converter using non-overlapped switches, inductor and filter capacitor .....	40
Figure 2.5. Comparative analysis between SCVR and inductive converters on efficiency-power density plane .....	46
Figure 2.6. SCVR for low-power market while inductive converter for high output power devices.....	47
Figure 3.1 (a) A-LDO            (b) Flash ADC based D-LDO (57) .....	51

Figure 3.2. D-LDO controller for steady-state response with single bit toggling once steady-state is reached. Bit granularity determines voltage ripple (34).....	51
Figure 3.3. Transient response to large di/dt limited by clock frequency due to single-bit response of the controller. Resulting in big droop on $V_{OUT}$ (34).....	52
Figure 3.4. Dropout voltage comparison between A-LDO and D-LDO .....	53
Figure 3.5. Energy harvesting system with MSP430 (TI) (46). Integrated LDO provides usable voltage to the MCU core.....	55
Figure 3.6. Proposed system power-aware D-LDO with integrated Fe-cap at output. Different LDO modes (active, standby) scale with corresponding MCU system modes. Turbo mode provides high-gain stages. ....	58
Figure 3.7. Up-down shift-register allowing single bit to toggle in a bidirectional manner, depending on $V_{OUT}$ with respect to $V_{REF}$ (57) .....	60
Figure 3.8. Single switch hierarchy for digitally-adaptive power management. Thermometric coding and additive enabling of power switches for non-linearly sized switches.....	62
Figure 3.9. Turbo-controller for enabling high-gain stages for large di/dt.....	64
Figure 3.10. Large di/dt droop improvements with turbo-controller enabled high gain stages.....	65
Figure 3.11. Fabricated D-LDO chip in 130nm low power CMOS process (TI).....	68
Figure 3.12. Regulated $V_{OUT}$ for different load current (simulation) .....	69
Figure 3.13. Measured plot showing D-LDO functionality. 13 $\mu$ s startup time achieved (on left) with ripple voltage (4.5%) on $V_{OUT}$ .....	70
Figure 3.14. More than 90% current efficiency measured in standby-mode.....	71

Figure 3.15. Monotonic change in $R_{DS_{ON}}$ with dynamic load current change indicates stability.....	75
Figure 3.16. Hybrid IVR with LDO (high $V_{OUT}$ ) and SCVR (low $V_{OUT}$ ) (34) .....	78
Figure 3.17. LDO mode, reusing SCVR topology for "truly" hybrid IVR (34).....	81
Figure 3.18. Proposed dual-loop D-LDO, with fine-grained slow inner loop for steady state and coarse-grained, fast outer loop for droop mitigation (34).....	81
Figure 3.19. Droop mitigation with dual-loop LDO (34).....	82
Figure 3.20. Hybrid IVR floorplan with less than 4% area overhead.....	84
Figure 3.21. Varying bit strength for different $V_{OUT}$ . Higher $V_{DS}$ , higher current per transistor.....	85
Figure 3.22 (a) Top half of IVR. 4 switches (enabled from each sub-module) comprise one bit for bit-gain of 1 (34) (b) 2 switches comprise one bit for bit-gain of 0.5.....	86
Figure 3.23. Measured voltage droop mitigation : with and without coarse-grained mode (34).....	87
Figure 4.1. Conventional load in parallel (left) versus stacked load in V-S (right).....	94
Figure 4.2. Intermediate voltage noise dependency on nature of load ; resistive or capacitive .....	95
Figure 4.3. Push-pull nature of regulation needed in V-S using capacitive converters....	98
Figure 4.4. Push-pull SC converter for differential V-S regulation.....	99
Figure 4.5. Stacked loads (three) with stacked SC converters (two), ideally providing $V_{dd}$ voltage headroom to each loads. Zoomed up single cell of 2:1 push-pull SC converter shown on left.....	100
Figure 4.6. Output impedance model of SCVR.....	103

Figure 4.7. Power-loss optimization in SCVR with $W_{sw}$ and $F_{sw}$ as design variables .	105
Figure 4.8. Conventional (parallel) vs V-S (series stacked) loads.....	108
Figure 4.9 (a) SC converter efficiency in delivering load current to conventional (non-stacked) loads (b) V-S system efficiency with varying workload between stacked loads.	
In worst-case, entire load current is provided through SC converter (similar to conventional loads in (a).....	109
Figure 4.10 (a) Power efficiency and power density (b) comparisons between stacked loads with linear and SC converter versus non-stacked loads with SCVR.....	111
Figure 4.11. Dual-boundary hysteretic PFM regulation scheme for V-S.....	113
Figure 4.12. Closed-loop versus open-loop regulation for conventional SCVR operation .....	114
Figure 4.13. V-S system efficiency with open and closed-loop regulation .....	117
Figure 4.14. P-P ripple reduction with open-loop regulation in V-S.....	117
Figure 4.15. Proof-of-concept V-S experiment showing the ability of the push-pull SC converter to recycle the current imbalance (“sink” additional charges) between stacked layers. $R_2 > R_1$ ( $R_2 \sim 2.5R_1$ ). The SC converter, fabricated in IBM130nm technology, is shown on left.....	118
Figure 4.16. Test setup for V-S using vertically connected FPGA boards and fabricated SC converter (left-bottom) with oscilloscopes measuring the stacked voltages and logic analyzer showing the outputs of the top and the bottom boards (bottommost). The connections of the vertically stacked FPGA boards are explained with colored wires (rightmost).....	120
Figure 4.17. Supply current reduction in stacked chips (compared to parallel chips)....	120

Figure 4.18. Measured results showing stacking without SC converter (top) and with SC converter regulating V-S (bottom) with load imbalance. $V_{IN} = 1.8V$ , Ideal $V_{OUT} = 0.9V$ for two loads stacked. ....	121
Figure 4.19. Three resistors stacked with dual-output SC converter for V-S measurement .....	124
Figure 4.20. Power efficiency performance of V-S with resistive loads with peak efficiency of ~98% measured with SC converter disabled in near-balance load condition ( $R1:R2=1$ ). ....	124
Figure 4.21. Comparison with state-of-art switching regulators shows how V-S lowers power overhead beyond the saturated limits. Zoomed-up figure shown below .....	126
Figure 5.1. Many-layer 3D-IC PDN including TSVs, C4s, Micro-connects.....	130
Figure 5.2. Impedance model of SCVR, capturing output voltage DC drop and power-losses (intrinsic and parasitic)(89) .....	137
Figure 5.3. Model versus simulation validation results (89) .....	138
Figure 5.4. 3 port RC model for 2-way interleaved SC converter (90) .....	140
Figure 5.5. Model versus simulation validation results (a) Output DC voltage (b) Output transient voltage trace (90).....	142
Figure 5.6. PDN structure for 3D-IC .....	144
Figure 5.7. Voltage noise evaluation with different TSV configurations, different workload distribution and different numbers of IVR in 8-layer 3D-IC with V-S (89)...	150
Figure 5.8. A plot of per-layer maximum noise amplitude over time. Only layer 3 has a noisy workload (90) .....	152

Figure 5.9. Three-layered stacked loads (architectural power traces) with stacked SC converter, simulated in 28nm FDSOI technology. ....	154
Figure 5.10. System power efficiency evaluation with different workload distribution and different numbers of IVR in 8-layer 3D-IC with V-S.....	156
Figure 5.11. Illustrates the tier and metal stack-up of the 5-tier 3D-IC version from NCSU EDA (104).....	158
Figure 5.12. Bidirectional buck-boost converter for differential power regulation in V-S .....	161

## List of tables

---

Table 3.1 Dynamic modulation of switch strength with system power-modes.....	62
Table 3.2 Performance comparisons among state-of-art low power LDOs.....	73
Table 3.3 Peak to peak ripple voltage with and without adaptive bit-gain modulation....	86
Table 3.4 Performance comparisons with state-of-art LDO (both analog and digital) ....	89
Table 4.1 Logic outputs modulating $F_{SW}$ with $V_{OUT}$ and $V_{REF}$ .....	112
Table 4.2 Dependence of V-S noise on ratio of current imbalance to aggregate current	122
Table 5.1 SCVR design specification (used in circuit design and modeling) .....	136
Table 5.2 Major 3D PDN modeling parameters .....	146
Table 5.3 TSV configurations used in this work .....	147
Table 5.4 Maximum voltage noise (%Vdd) per layer for different workloads on 3D-ICs with different PDN schemes. The “cross-layer mean” value averages all layers’ maximum noise amplitude (90) .....	153

# Chapter 1

---

## 1.1. Introduction

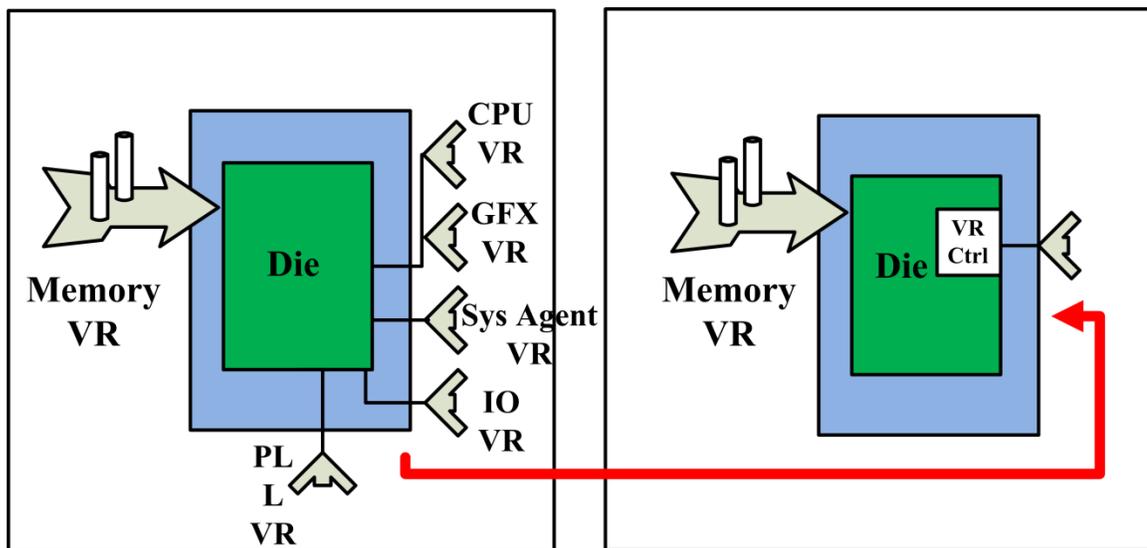
Moore's Law, or the "prophecy" made in 1965 by Gordon. E. Moore, co-founder of Intel Corporation, regarding the exponential increase in transistor count in integrated circuits, continues to inspire and guide the semiconductor industry to set roadmaps for future research and developments (94). Not just the transistor count, but the capabilities of many digital electronic devices, the cost-per-unit of microprocessors, memory capacity, and even the size of pixels in digital cameras have been increasing exponentially as well. The increased computational capability at lower energy-per-operation has led to the recent surge in market demand for portable devices like smartphones, tablet PCs and other handheld devices. The ability to integrate the entire system in a single silicon chip has made systems on chip (SoC) design flow extremely popular among semiconductor companies, especially those targeting the low power electronics market (95).

As CMOS scaling has gone past the 28nm process node, significant shifts in traditional design trends have been noticed, especially for SoC designs. Bulk CMOS have been replaced with Silicon-on-insulator (SOI) and FINFET technologies, mostly due to

their superior short-channel effects (SCE) (48, 95). While fully-depleted SOI (FDSOI) has gained popularity among low power SoC, FinFET technology has already scaled down to 14nm process node (97). Aggressive power management strategies like dynamic-voltage-frequency-scaling (DVFS) have become increasingly common for reducing active and standby power of CMOS load (35). However, unlike earlier implementations, focus has shifted more to integrated power regulation to benefit from fine-grained spatial and temporal power management techniques (34, 80). Integrated voltage regulation (IVR) for on-chip power management has been adopted commercially, as seen in the Haswell processor from Intel using an in-package inductor as shown in Figure 1.1 (8). Since 2012, more than five works on fully-integrated regulators, both capacitive and inductive, have been published by Intel Labs just on 22nm process node (25, 34). The highly power-optimized switching and the improved parasitic in these advanced nodes have led to significant improvements in performance.

Another interesting trend has been the increasing appearance of digitally-assisted analog solutions on chip, replacing their traditional analog counterparts (2, 29, 53). Apart from performance enhancement and scalable design-approach, a major motivation for digitally-assisted solutions arises from the challenges in conventional analog circuit design beyond 22nm with reduced transistor gain and lower voltage headroom available. Industry and academia have started actively exploring the option of a digitally-controlled power management-unit (PMU) with adaptive clock-modulation, and adaptive guard-band voltage scaling to monitor the workloads and dynamically enable different power modes (16).

While active and standby load-power are controlled through different power management techniques, chip-interconnect and clock-distribution network power consumption have become major contributors to power-walls (14). This has renewed interest in vertical integration of devices and functionalities with reduced latency, in the form of 3D-IC implementations (61). While different flavors of 3D-IC like 2.5D or multi-chip module (MCM) have already been commercially accepted, recent advancements with many-tiered memory (Hybrid memory cube by Micron) have generated significant interest in many-layered vertical integration once again (27, 84) .



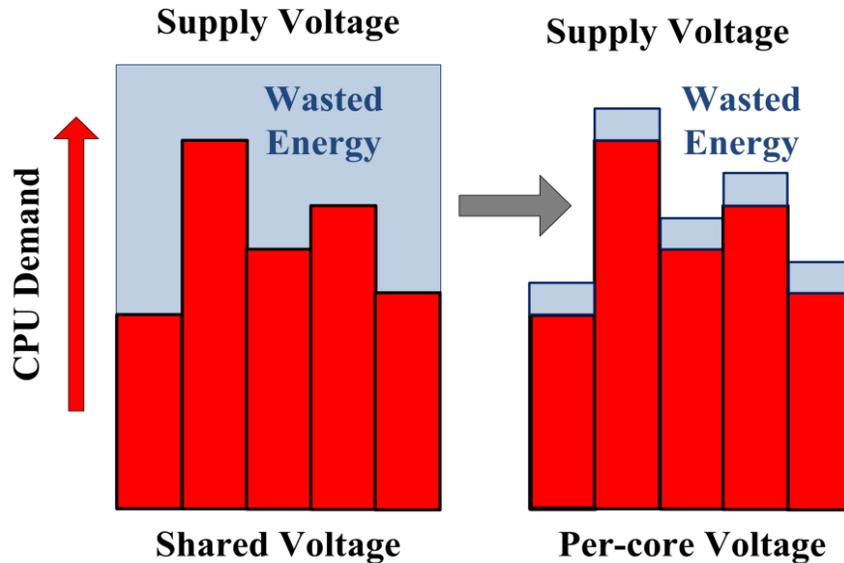
**Figure 1.1. Intel Haswell processor, moving off-chip VR to on-chip with passives-in-package (SiP) (8)**

### 1.1.1. Why integrated VRM

The most effective technique to reduce CMOS load power consumption is to dynamically change supply voltage and clock frequency (DVFS), depending on the changing

workload (80). Low voltage modes are used in conjunction with low clock frequency to minimize power consumption and only when significant computational power is needed, higher voltage/frequency modes are enabled. Due to the quadratic dependence of power on voltage, dynamic manipulation of voltage/frequency can significantly reduce the power consumption. State-of-the-art power management in a SoC revolves around the use of multiple voltage rails, dynamic regulation of the supply and adaptive voltage scaling techniques (83). However, generating these different voltages efficiently has proven to be challenging for practical implementation of DVFS.

Voltage regulator modules (VRM), which are used to deliver power from energy sources (e.g. battery) to integrated circuits at varying voltage levels, are the most vital components of the PMU. However, the slow voltage tuning capability, cost overhead and bulky appearance of the conventional VRMs make them less attractive for multiple power domain implementations. Therefore, the full promise of DVFS has been hindered by the slow off-chip VRMs, and most of the modern implementations are limited to temporal coarse-grained DVFS governed by runtime software (i.e. the operating system) (35).



**Figure 1.2. Per-core voltage saves energy wasted in shared rail (shown in blue) (34)**

IVR offers the potential to provide multiple on-chip power domains with faster voltage transitions (nanoseconds as opposed to hundreds of microseconds for off-chip VRMs), therefore allowing fine-grained DVFS and significant energy savings, especially for memory-intensive workloads, without degrading performance (62). As shown in Figure 1.2, per-core or per-block IVR has the potential to save large amount of energy otherwise wasted from shared rail. Unfortunately, these potential benefits of IVRs are tempered by their lower energy-conversion efficiencies resulting from higher switching frequencies and increased silicon area to accommodate on-die passives (56).

#### ***1.1.1.1. Types of IVR***

There are primarily two classes of IVR, linear and switching converters. Low-dropout regulator (LDO), the most common among integrated linear regulators, performs voltage conversion continuously by dissipating the difference between the input and regulated

voltage as excess heat. Therefore, LDOs are effective only over a small voltage range, suffering from poor efficiency otherwise. Nevertheless, they form a vital cog in SoC PMU (17). Switching converters regulate through switching of different passives and, based on the nature of these passives, they are either capacitive or inductive converters (38). While it is highly debatable whether one is superior over the other and is the subject of several research papers (further discussed in chapter 2), availability of highly dense capacitive technologies and improvements in switching performance with CMOS scaling, has made capacitive converters a compelling choice for IVR in the advanced nanoscale process nodes.

### 1.1.2. Why digitally-controlled techniques

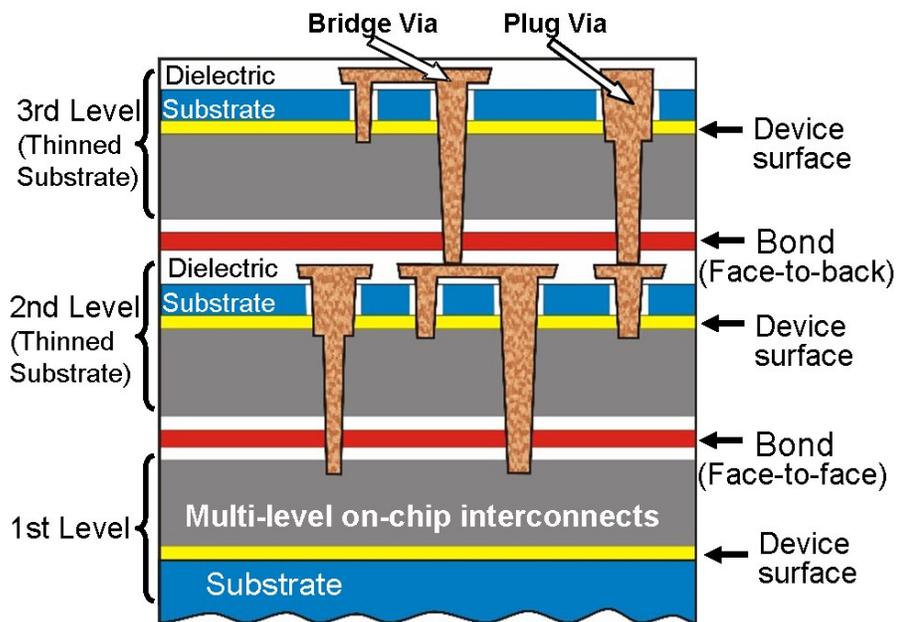
Traditional analog designs have not scaled optimally, as CMOS transistors are scaled to 28nm and below, particularly due to the low voltage headroom in these process nodes. Ensuring high gain and stability for feedback controlled circuits has become increasingly challenging. Therefore, digitally-assisted circuits with on-chip monitoring and extensive calibration capability have become popular for these ultra-nanoscale technologies. Digital control techniques with their simple circuit-implementation, fast response and wide-range stability can help save power wastage, especially in SoC-based designs. Easy configurability to load changes, higher immunity to process variations and scalability across technologies adds to the benefits (30).

### 1.1.3. Power delivery wall in 3D-IC

In two-dimensional (2D) ICs, even as minimum feature size decreases with scaling, silicon area stays constant due to the ever-increasing demand for functionality and higher performance. The impact of this miniaturization has affected the performance of interconnects across the chip. Not only has signal delay increased, interconnect power dissipation is dominating the total power consumption (50% - 60% of the total dynamic power dissipation) as technologies are entering the sub 20nm range (ITRS 2012). To keep up with Moore's law of integration and alleviate interconnect limited performance degradation in future SoC, 3D-IC design strategy is being explored as a viable alternative.

As shown in Figure 1.3, 3D-IC involves stacking silicon wafers and/or dies and interconnecting them vertically using through-silicon-vias (TSV) to achieve performance with lower power and smaller footprint, as compared to 2D-IC having the same functionalities. However, while the performance of the 3D-IC memory stacks has improved by leaps and bounds in recent years, logic or processor stacking in 3<sup>rd</sup> dimensions still has not been able to deliver the full promise of 3D-IC. There are three main possible issues that may conspire to make 3D processor stacking seem impractical: manufacturing issues (and associated cost), thermal issues, and power delivery issues. Manufacturing 3D stacks economically is not trivial, but is clearly not going to be a showstopper, as exemplified by the Micron hybrid memory cube and JEDEC Wide I/O parts that are already sampling with 4 and 8 TSV-stacked die - if 3D fabrication is economical for memories, it is even more so for logic, especially since it provides a truly heterogeneous platform. Thermal issues can also be daunting in 3D, with two possible

solutions: if the power is not too high, as in the case of memory and low power processing elements, conventional cooling techniques can still cope; for high performance cores more exotic cooling solutions will be needed, such as inter-tier liquid cooling microchannel . As the number of physical layers in a 3D-IC stack is predicted to increase in the future, from the present 2.5D multi-layer solutions, with only a couple of layers, to true 3D many-layer stacks, with tens of layers, the problem of delivering power to the 3D stack is the biggest obstacle. The main culprit is the fundamental mismatch between the volumetric (cubic) aspect of power consumption and dissipation in 3D-IC, and the fact that power delivery is limited to only a 2D surface (quadratic) (top or bottom of the die or stack). This 3D volume vs. 2D surface power wall is a fundamental obstacle to 3D-IC scaling.



**Figure 1.3. Integration in the vertical direction or 3D-IC (98)**

## 1.2. Organization

This dissertation has adopted a comprehensive circuits-architecture-technology approach for practical, efficient and scalable solutions in the form of digitally-controlled fully-integrated power regulation. The two major focuses of this dissertation are:

- I. Improve the figure-of-merits of power regulation architecture in 2D-IC through digitally-assisted novel techniques;
- II. Provide a fundamental solution to break the 3D-IC power delivery walls.

### 1.2.1. Background

Chapter 2 presents an in-depth survey and comparisons among existing VRM architectures. An overview of integrated switching converters, both inductive and capacitive, is presented, and comparison plots among state-of-the-art switching converters have been discussed to justify usage of different regulators for different specifications.

### 1.2.2. Digitally-controlled low-dropout regulator

Chapter 3 proposes digitally-controlled LDO topologies, catering to a wide-range of applications. The idea is to replace traditional analog-controlled LDOs with digitally-assisted ones to leverage the energy-efficient computing capabilities of digital CMOS. The first topology this chapter proposes is a system-power-aware digitally-controlled

LDO that can reduce the standby power-drain in an ultra-low-power (ULP) microcontroller unit (MCU) through a drastic reduction of its (LDO's) own quiescent current. The second topology this chapter proposes is a fast-response digitally-controlled LDO for designing the first “truly hybrid” integrated LDO/SC converter for highly efficient super-to-near-threshold voltage generation. The fundamental operating principles of both topologies are similar, but the circuit techniques applied and nature of the load are very different, ranging from energy-harvesting architecture to graphics core. Both of these converters have been implemented in silicon and simulation results match measurements with fairly good accuracy.

### 1.2.3. Charge-recycled power regulation with integrated switched-capacitor

Chapter 4 focuses on improving the figure-of-merits of conventional integrated power regulation by proposing an alternative: charge-recycled power regulation with integrated SC converter, also known as voltage stacking. By exploiting the differential nature of regulation and capability of SC converters to act as charge-equalizer, voltage stacking (V-S) claims an energy-efficient approach to integrated power regulation. Through its circuit/architectural simulation and proof-of-concept demonstration of V-S with fabricated SC converter and commercial FPGA chips, this chapter has shown the benefits of V-S.

#### 1.2.4. Breaking 3D-IC power delivery walls using voltage-stacking

Chapter 5 proposes a fundamental solution to 3D-IC power delivery walls in the form of multi-output SC converter-assisted voltage stacking (3D-MOSC). This chapter first discusses the physical mapping of charge-recycled regulation to through-silicon-via (TSV) supported many-layered 3D-IC. A cross-layer modeling approach, incorporating 3D-IC power-delivery-network (PDN), integrated multi-output SC converter models, and processor models, has been discussed in this chapter. Using this system-level 3D-IC voltage stacked PDN model, extensive simulations have been performed to analyze the various design tradeoffs in this unconventional power delivery approach. A significant increase in system efficiency, resolving 3D-Vs-2D power delivery mismatch, and improvements in voltage noise are among the many features of charge-recycled power regulation, making it attractive for future many-layered 3D-IC expansion.

## Chapter 2

---

### 2.1. Background

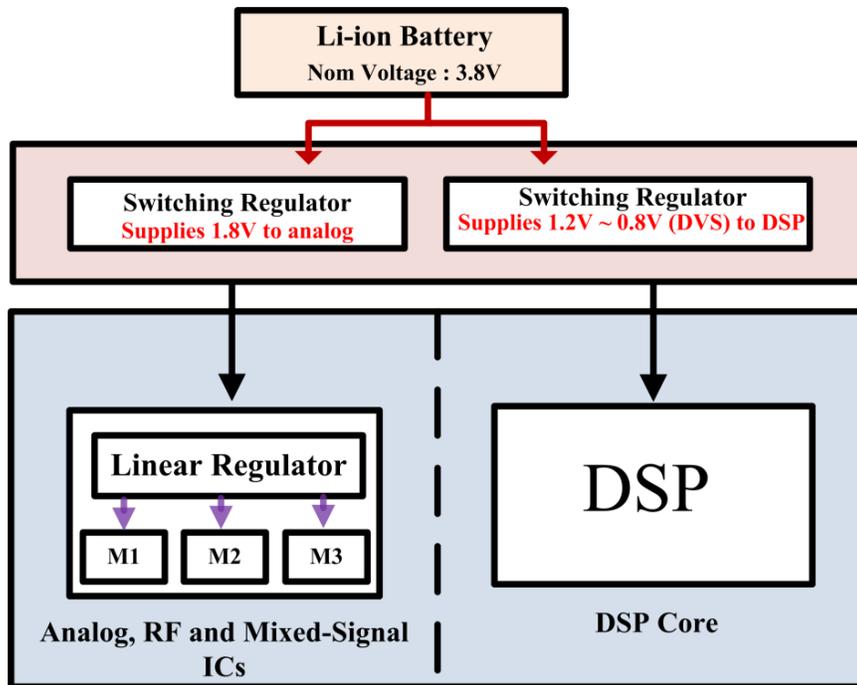
To provide a baseline for understanding the improvements in integrated power regulation that this work has achieved, this chapter summarizes the different figure-of-merits and comparisons among existing converter designs.

#### 2.1.1. Overview of IVR

As CMOS technology has continued scaling beyond 28nm, load characteristics have changed rapidly (lower voltage headroom, faster load transient) along with an increase in chip complexity with analog, digital and RF circuits co-existing in modern day SoC. Delivering power efficiently and reliably to such a dynamic workload directly from the supply is no longer a practical option. Battery technology, which powers most of these handheld devices, has not been able to keep pace with the exponential increase in integration. For example, the energy density of the Lithium-ion (Li-ion) battery has only doubled in two decades. With growing popularity of portable devices, there is an increasing demand for higher levels of integration to reduce both board space and cost.

The need for small size, low cost and extended battery life requires highly efficient power converters to deliver well-regulated voltage with widely varying load (97).

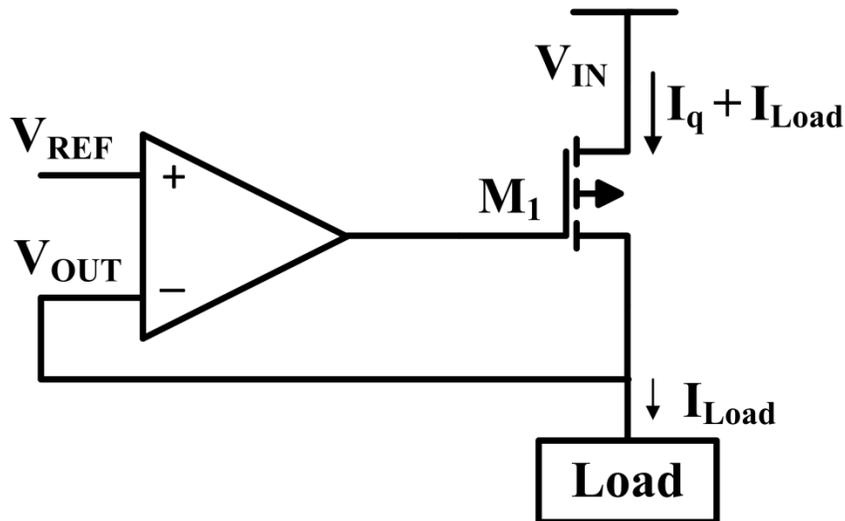
Principally two different approaches exist, linear regulators (mostly used in the form of LDO for integrated regulation), and switching regulators. While LDO is the popular choice where input ( $V_{IN}$ ) to output voltage ( $V_{OUT}$ ) ratio is low (i.e.  $V_{IN} - V_{OUT}$  is small), because of the higher efficiency and smaller area overhead, switching regulators are more commonly used to generate a wider range of output voltage (from super-threshold to sub-threshold voltages) at a higher power efficiency. A PMU may even combine several LDOs and switching regulators to support a wide range of load circuitry. Figure 2.1 shows a typical mixed-signal management IC, where both switching converters and LDOs are used to regulate voltages over a wide range.



**Figure 2.1. Mixed-signal PMU with switching regulators providing global (off-chip) regulation and LDO providing local (on-chip) regulation (17)**

### 2.1.2. Low-dropout regulator (LDO)

LDO works on the principle of resistive voltage division, where output voltage is regulated against varying load currents by continuous-time comparison with reference voltage via a feedback path (Figure 2.2). As the input-output voltage difference increases, large amount of power is dissipated as heat. Since no energy storage passives are involved in linear regulators, however, the area overhead for complete chip integration is lowest for this class of regulators. This kind of topology is especially preferred in low-power low-cost applications due to the lower standby (quiescent current) power or in multiple voltage island based power delivery networks (PDN) (46).



**Figure 2.2. Low-dropout regulator popular for point-of-load regulation in SoC**

#### 2.1.2.1. *Working principle and figure-of-merits*

In Figure 2.2 transistor  $M_1$  acts as a sense resistor and keeps the output voltage regulated against a varying workload. The control circuits monitor the output voltage ( $V_{OUT}$ ) and

regulate by modulating the gate voltage of the power MOSFET. Dropout ( $V_{\text{DROPOUT}}$ ) refers to the minimum voltage drop between the unregulated input voltage ( $V_{\text{IN}}$ ) and regulated  $V_{\text{OUT}}$ , which is a direct measure of the power efficiency. Another important FOM in LDO is the quiescent current ( $I_q$ ) i.e. the no-load current. The power MOSFET can be either NMOS (in common-drain configuration) or PMOS (in common-source configuration) as in Figure 2.2. While NMOS regulator is much more stable due to the lower output impedance, the PMOS regulator provides lower dropout voltage and, therefore, higher power efficiency. Maximum power drawn from the source and delivered to the load is given by Equations 2-1 and 2-2

$$P_{\text{OUT}} = (V_{\text{IN}} - V_{\text{DROPOUT}}) * I_{\text{Load}} \quad 2-1$$

$$P_{\text{IN}} = V_{\text{IN}} * (I_{\text{Load}} + I_q) \quad 2-2$$

Efficiency is given by Equation 2-3

$$\eta = \frac{P_{\text{OUT}}}{P_{\text{IN}}} = \frac{(V_{\text{IN}} - V_{\text{DROPOUT}}) * I_{\text{Load}}}{V_{\text{IN}} * (I_{\text{Load}} + I_q)} \quad 2-3$$

Therefore, dropout voltage and quiescent current needs to be minimized for improved efficiency. With the growth of digitally-controlled LDO using discrete control techniques, transient response ( $Tr$ ) has become another important FOM ( $FOM_1$ ) and is given by Equation 2-4. Equation 2-5 combines the impact of transient response with quiescent current ( $I_q$ ) for  $FOM_2$ , where  $C_{\text{OUT}}$  and  $\Delta V$  indicate output capacitance and worst-case transient voltage drop (20). While power efficiency ( $\eta$ ) is the most important FOM for switching regulators, for LDO current efficiency ( $\eta_{\text{current}}$ ) given by Equation 2-6, especially in standby-mode, dictates the battery lifetime (power efficiency is constrained by input to output voltage ratio)

$$FOM_1 (T_r) = \frac{C_{OUT} * \Delta V}{I_{Load}} \quad 2-4$$

$$FOM_2 = FOM_1 * \frac{I_Q}{I_{Load}} \quad 2-5$$

$$\eta_{current} = \frac{I_{OUT}}{I_{IN}} = \frac{I_{Load}}{(I_{Load} + I_Q)} \quad 2-6$$

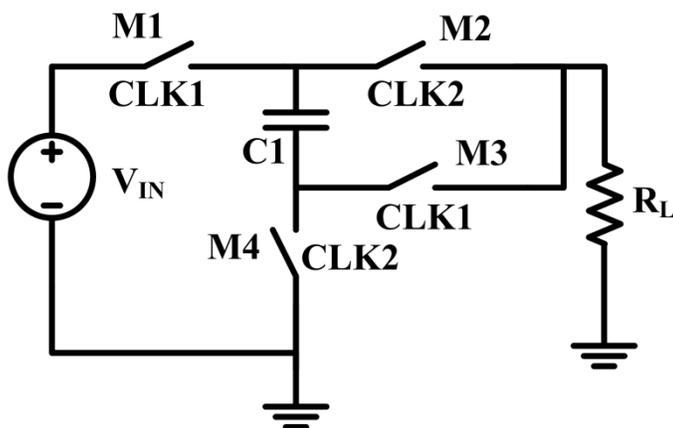
### 2.1.3. Fully-integrated switching converters

Switching power converters have the potential to support a wide range of voltages at high efficiency by exploiting the charge-storage capability of different passives, i.e. inductor and capacitor. However, the available volumes of these passives being small, the focus has been toward small-form-factor switching converters with integrated reactive components. While an inductive buck converter uses both inductor and capacitor along with MOSFET, as switches for voltage down-conversion, switched-capacitors use multiple ratios of switches and capacitors to provide the desired voltage ratios. A feedback control loop monitors the voltage response to varying load current and regulates the converters by modulating the switching frequency (pulse-frequency modulation or PFM), duty cycle of the pulse (pulse-width modulation of PWM), conductance of the switches, the amount of passives (capacitive modulation), or a combination of the above-mentioned techniques.

#### 2.1.3.1. *Capacitive converter*

The switched-capacitor (SC) converter accomplishes energy transfer and voltage conversion using switches and capacitors (25, 43, 65, 40, 3, 7, 74 66, 10, 60, 82, 19, 12).

A big reason for their popularity is the absence of bulky magnetic elements such as those used in inductive converters. Figure 2.3 shows the most energy efficient capacitive converter structure i.e. 2:1 converter (3). While the number of fly-caps, different configuration of switches, and polarity of the capacitor being charged determine the output voltage ratio, the fundamental principles of all SC converters are the same. During the first half of the clock cycle (CLK1), switches M1 and M3 turn on, charging capacitor C1, while M2, M4 are switched off, and vice versa, for the next half period of the clock signal, where C1 is connected to output. A non-overlapping clock is used to reduce shoot-through current loss. The duty cycle of the switches is set to 50% since that leads to the highest efficiency. The amount of charge transferred is a function of the switching frequency and the fly-caps. Therefore, higher switching frequency allows smaller on-chip capacitor for the same voltage drop on output. However, for real implementations, the maximum range of switching frequency is limited by the overall switching losses reducing the efficiency of the capacitive conversions (72).



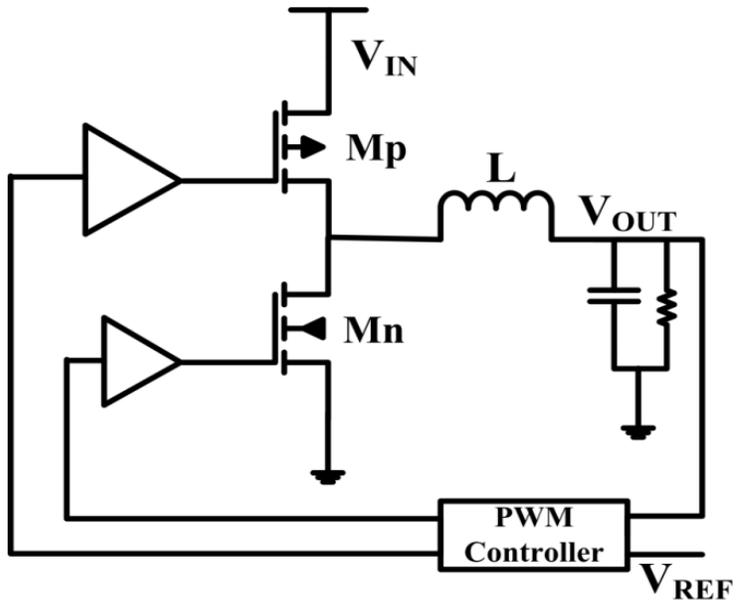
**Figure 2.3. 2:1 Topology of switched-capacitor converter (SCVR) using fly-cap and switches**

### 2.1.3.2. *Inductive converter*

The off-chip inductor-based buck converter in Figure 2.4 is widely used as a high power (current) converter with high efficiency (>90%), and is capable of generating lower, higher, even opposite polarity DC load voltages, with respect to input voltage (38, 4, 5, 87, 21, 86, 85, 71, 58, 56, 1, 39, 68 ). The relation between  $V_{OUT}$  and  $V_{IN}$  is given by

$$V_{OUT} = D * V_{IN} \quad 2-7$$

Based on this relation, the feedback controller changes the duty cycle when the reference voltage changes, allowing a continuous range of DC voltages to be generated by the buck converter at a higher efficiency than LDO. The sizing of the PMOS  $M_p$  and NMOS  $M_n$  power switches is determined by the maximum load current the converter needs to regulate. The control circuitry monitoring the feedback can be made completely digital, drawing very little power overhead. This makes the inductive converter a popular choice for DVS-based systems, where fine-grained power management can extract higher energy efficiency out of the digital loads. However, conventional approaches have used bulky off-chip filter components, making them costly. Recently, a number of research efforts have been invested in integrating the passives on-chip for buck converter (21, 37, 71). However, the integrated passives are constrained by the area overhead and poor quality factor owing to the large ESR. Higher switching frequency ( $F_{sw}$ ) to meet the output ripple voltage specifications comes at the cost of higher switching losses in the NMOS and PMOS power devices and their corresponding drivers, thereby reducing the power efficiency.



**Figure 2.4. Synchronous buck converter using non-overlapped switches, inductor and filter capacitor**

#### 2.1.4. Overview of integrated passives

Motivation for on-chip integration of voltage regulators is guided by the increasing demand for aggressive power-scavenging. However, to be economically beneficial, the many overheads of on-chip integration need to be addressed. The biggest challenge is to integrate high quality passives on-chip. Different technologies and techniques have been used, over the years, to integrate more and more passives on-chip, some of which are discussed in this section.

##### 2.1.4.1. *Technology options for SC converter*

Not just the capacitive density, but also the quality of passives in terms of bottom-plate loss and implementation costs, are deciding factors in choosing between different technology options.

- A. Bulk CMOS technology can provide a capacitive density of 4-12 nF/mm<sup>2</sup> when using the gate capacitance of MOS transistors (MOS-cap). Highest density is obtained using transistors with thin gate oxide. However, leakage and lower breakdown voltage are concerns here, requiring careful handling of voltages between the MOS terminals. Also, due to the proximity to substrate, the bottom-plate capacitance is largest (~10%). However, its seamless integration with CMOS load makes it very attractive (85).
- B. Metal-insulator-metal (MIM) capacitor is another option that has become attractive in recent years. Low capacitive density and extra cost of additional mask during the fabrication phase used to be detrimental factors preventing the use of this parasitic capacitor between the metal layers. Fortunately, with progress of CMOS process nodes, the lateral and vertical intervals between the metal interconnects have decreased and parasitic capacitance between the interconnects has increased. As Intel has shown, by hanging the MIM capacitor (between Metal 8 and Metal 9 layers) above the load, a high density IVR can be designed. MIM-cap has the additional advantage of lower parasitic (~1%) and higher breakdown voltages (25).
- C. Silicon-on-insulator (SOI) technology owing to the isolation from the bulk has a lower parasitic capacitance and a lower leakage current, favoring SC converter implementation. 14-17nF/mm<sup>2</sup> capacitive density is achieved using thin oxide MOS-cap (43).
- D. The deep trench capacitor is being considered as one of the future proponent of “More-than-Moore” paradigm of ITRS roadmap. With capacitive density being

100X times that of conventional MOS-cap and breakdown voltage being higher, trench cap has attracted a lot of attention for SC converter implementations. However, this technology is costly and is not compatible with conventional CMOS process flow (3).

#### **2.1.4.2. *Technology options for Integrated Inductive Converter***

Integrating inductive DC-DC converter along with the load is especially attractive, since the principle of buck converter is well-understood and a wide range of voltages can be generated through PWM control. However, integrating the inductor efficiently with small area overhead is challenging considering their large form factor. Two different methodologies are usually adopted for integration. First, a different and dedicated technology can be used to integrate the reactive components and provide an optimum solution. This approach is referred to as system-in-package (SiP). For example, off-chip surface mount devices (SMD), air-core inductors and capacitors are used (71). The reactive components may also be realized in a different IC technology or on a separate die. This is known as dual-die or multi-chip module (MCM), where the active and reactive dies are connected to each other via bumps. This way cheaper technology can be used for the reactive components, while active components will be in advanced process nodes (58).

The other way is to monolithically integrate the passives on the same die as the switches and load circuits to reduce area and cost of integration, as in (4, 86, 87). However, in standard CMOS, an inductor remains difficult to integrate within an acceptable area with the performance specifications. Another significant issue with the integrated inductor is its poor parasitic impedance, which can cause significant energy

dissipation. One option is to design on-chip inductors by connecting bond wires in a loop above the load (87). Bond wire inductors have a relatively low series resistance (approximately  $50\text{m}\Omega/\text{nH}$  at  $100\text{MHz}$ ). In addition, they show a low capacitive coupling to the substrate and can sustain high voltages. However, bond wire variations make them less reliable. A monolithic spiral inductor can be formed using a thick top metal, but owing to the high switching frequency that it demands and high ESR ( $250\text{m}\Omega/\text{nH}$  at  $1\text{GHz}$ ), efficiency is lower (85).

#### 2.1.5. Comparisons between integrated SC converter and inductive converter

In this section a comparative analysis among different integrated converter topologies is discussed. Since the primary FOM for switching converters are power density and power efficiency, the comparison plot is shown in an “efficiency versus power density” plane. While LDO is an equally important substitute for integrated converter and is, in fact, the more popular choice due to its lower cost, the efficiency and power density of this class of linear regulator is very much determined by the design specifications, unlike switching regulators, where these FOMs are determined by the nature and amount of passive technology and other design choices. What makes this comparison very difficult is the fact that different output voltages ( $V_{\text{OUT}}$ ), different power range ( $P_{\text{OUT}}$ ), and different technologies are being used for a wide variety of applications. However, based on numerous data from the literature and the conclusions drawn from similar work in the past (73), the following observations can be made.

- 1) A key difference between inductive and capacitive converters is that the inductive topology has the ability to provide variable output voltages by changing the duty cycle of the modulating signal, whereas the capacitive topology delivers a fixed output voltage depending on the switch/capacitor configurations. This any-range-voltage delivery capability of inductive converters makes them very popular for fine-grained DVS implementation. To circumvent this problem, a capacitive converter must change its switch and capacitor configuration into a different topology to provide a different voltage-conversion-ratio (VCR). With these kinds of multi-ratio SC converter and with the growth of different discrete DVS schemes (62,92), the switched-capacitor converter has been developed as a viable alternative to a fully-integrated inductive converter. While more switches are needed in SC converter, as compared to inductive converter, capacitive configurations have lower switch V-A stress than the buck converter (73).
- 2) Figure 2.5 shows the SC converter and the buck converter FOM comparisons for power density/efficiency plane. MOS and MIM-cap based SC converter implementations show the power and area overhead tradeoff with the bottom-plate parasitic loss of MOS-cap and lower energy-density of MIM-cap dictating their position on the plane. Transistor technology, as mentioned earlier, also plays a very strong role as the efficiency improves with better switching characteristics of advanced process nodes. Eventually, though, the highly-dense low parasitic capacitive technology determines the winner among the integrated converter topologies with highest peak power efficiency and power density (3, 12). Among the inductive converters, the monolithic and the bond wire based inductor

topologies show similar power density. It is hard to conclude much from the SiP converters, as they are all over the chart. However, the primary reason for choosing SiP will be to lower the cost, since it allows inexpensive technology to be used along with advanced nodes within the same package.

- 3) Finally, interesting conclusions can be drawn from the usage trend of inductive and SC converters, as shown in Figure 2.6. Around the boundary of 100mW, there is a clear demarcation between SCVR and inductive converters. SC converters dominate the low power market (10-50mW) while inductive converters are more attractive when load power ranges above 100mW to justify the larger form factor. However, recent works show that even at higher end, SC converters can be just as effective (3, 10). Some other notable works are missing from this plot due to lack of information regarding the area overhead of the converter. For instance Intel, in its latest Haswell processor, claims a significantly higher power density number for its IVR using package inductor (8). Similarly, Intel's work with the MIM-cap based SC converter shows higher density by hanging MIM-cap between higher metal layers over digital load (34) .

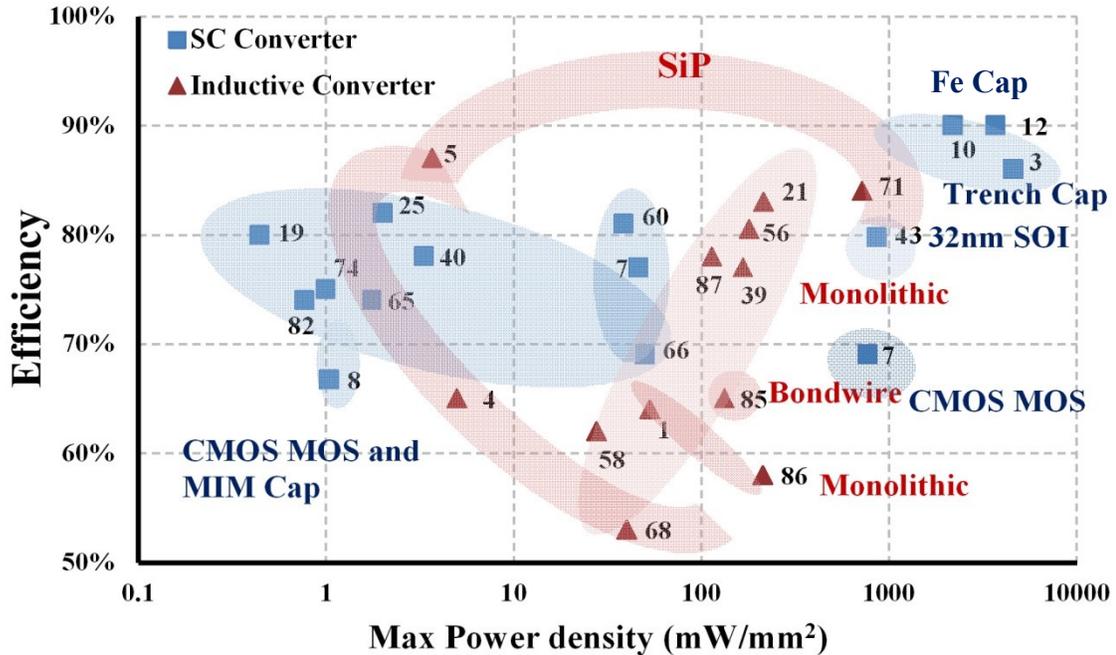


Figure 2.5. Comparative analysis between SCVR and inductive converters on efficiency-power density plane

Based on the above discussion, the conclusion that can be drawn is that, while inductive converters continue to remain the industry’s primary choice, the SC converter has created a niche market of its own as a fully-integrated DC-DC converter. However, power density and power efficiency of either kind of converters will be limited by the constraints on integrated passives. While the quality of CMOS compatible capacitors and monolithically integrated inductors has increased significantly in the past decade, it is unlikely that they can keep up with the urgent need for highly efficient (more than 90% on average) and highly dense (5-10W/mm<sup>2</sup>) on-chip VR for power management in power hungry applications such as graphical processing units (GPU) and mobile SoCs.

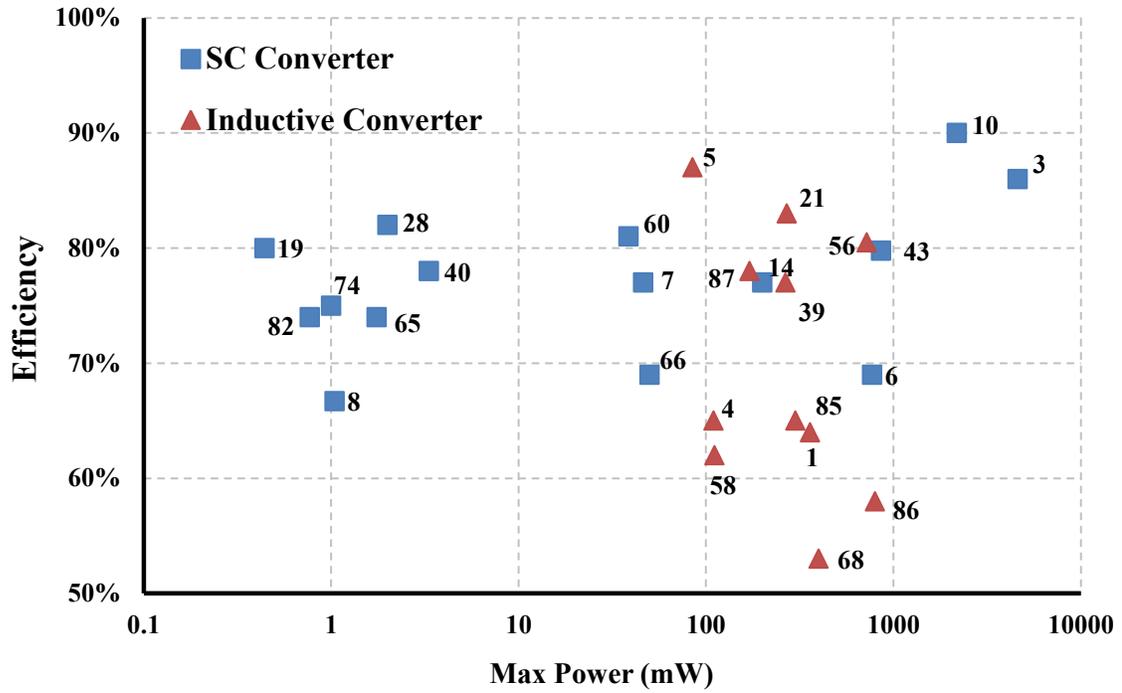


Figure 2.6. SCVR for low-power market while inductive converter for high output power devices

## Chapter 3

---

### 3.1. Digitally-controlled Low-dropout Regulator

Even as a fully-integrated switching regulator is actively researched for fine-grained spatio-temporal power management, LDO continues to be a popular choice as point-of-load regulator. While analog LDOs benefit from superior transient response (17, 20) and high power supply rejection ratio (PSRR), they face considerable design challenges as CMOS scaling continues beyond 28nm. Stability, linearity and gain of the analog designs suffer from low voltage headroom. Also, the analog LDOs do not integrate well with the digital design/process flow requiring custom integration and placement.

Digitally-assisted analog circuits have been suggested to leverage digital-computing capabilities to improve power and performance of analog electronics. Unlike their analog counterparts, digital circuits consume very little current in steady state, provide large output current when switching, and have the ability to operate under low supply voltage. Another opportunity of digital control comes from the power management concept. As more and more point-of-load regulators are being used, the most efficient way to coordinate all these regulators is to use a power management unit (PMU). A digitally-

controlled LDO can interact much more efficiently and adaptively with a PMU than an analog one (80).

Therefore, to overcome analog design-related challenges and pave the way toward digitally-adaptive PMU implementation, two novel digitally-controlled LDO (D-LDO) architectures are explored in this chapter. While the control structures of both architectures are fundamentally similar, they cater to very different applications. A truly hybrid LDO/SC converter, using the proposed D-LDO, is designed as part of a digitally-adaptive PMU regulating a graphics processor core in 22nm Intel tri-gate technology (34). Hysteretic-bound control scheme and fast droop mitigation techniques have been implemented, with measurement results showing improved FOM compared to state-of-the-art LDOs. The second D-LDO has been proposed to reduce the standby power-drain in ULP MCU used typically in energy-harvesting architecture (49). By modulating its drivability using system-power-aware modes, the proposed D-LDO drastically reduces its quiescent current consumption to improve standby-mode efficiency. With a fully-integrated Fe-cap acting as decoupling capacitor, the proposed D-LDO has been implemented in 0.13 $\mu$ m CMOS (Texas Instruments low-power) process node.

### **3.2. Background and comparisons with A-LDO**

Migration into advanced CMOS nodes with superior switching characteristics has inspired the recent works on digital implementation of LDOs targeted toward digital loads (16, 80). The rationale behind such design is to convert the control section into a digitally-controlled block which is easier to integrate scaled nodes (67). However, while

this opens up all the benefits associated with digital technologies, some of the superior performance of the fine-grained analog control is lost.

The D-LDO by Okuma et.al (57) is considered the pioneering work in the field of digitally-controlled LDO. As shown in Figure 3.1, the analog controlled power transistor of the conventional LDO is replaced with an array of switches and a comparator-based flash ADC regulation to replace analog op-amp. By monitoring the output voltage and enabling switch array on/off, regulation is ensured for this kind of LDO. The digital nature of the controller allows low voltage operations with ultra-low quiescent current ( $2.7\mu\text{A}$  at  $0.5\text{V } V_{\text{IN}}$ ). The switching action of D-LDO, however, causes higher ripple on output voltage, making analog LDO the popular choice for analog/RF loads. In order to reduce the ripple due to switching, a bidirectional up/down shift register is used to restrict single switching at each clock edge (Figure 3.2). While this improves steady state response ( $3\text{mV}$  ripple for  $V_{\text{OUT}}$  of  $0.45\text{V}$ ), it leads to poor response to sudden voltage droop. As Figure 3.3 shows, the transient response to handle the sudden  $di/dt$  is limited by the sampling frequency of the clock, due to the nature of this controller. Higher clock frequency designed for the worst  $di/dt$  will improve this voltage droop, but overdesign for the rest of the D-LDO operation will lower the power efficiency. To accommodate the additional droop, voltage guard-band is increased, leading to power wastage. Therefore, careful balance needs to be ensured between controller design, response time, and power overhead for D-LDO.

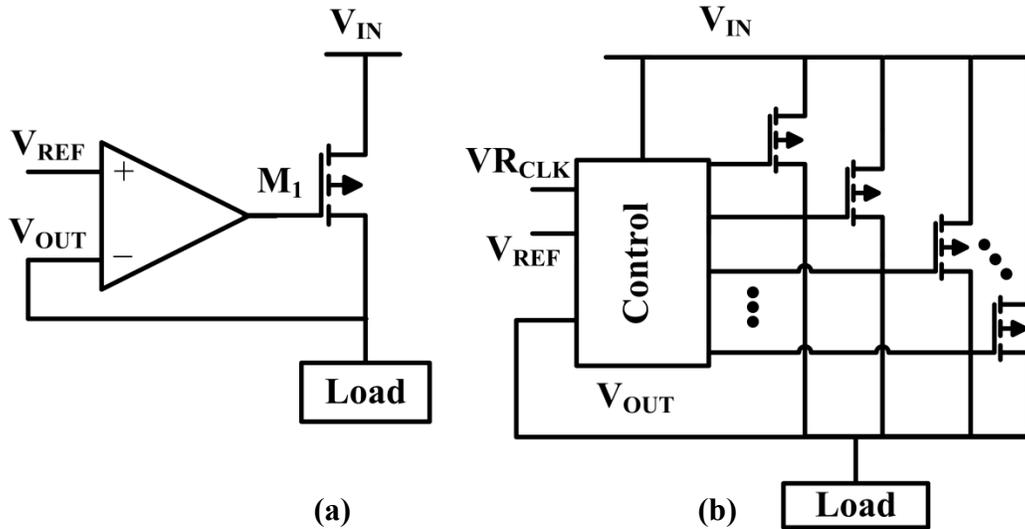


Figure 3.1 (a) A-LDO

(b) Flash ADC based D-LDO (57)

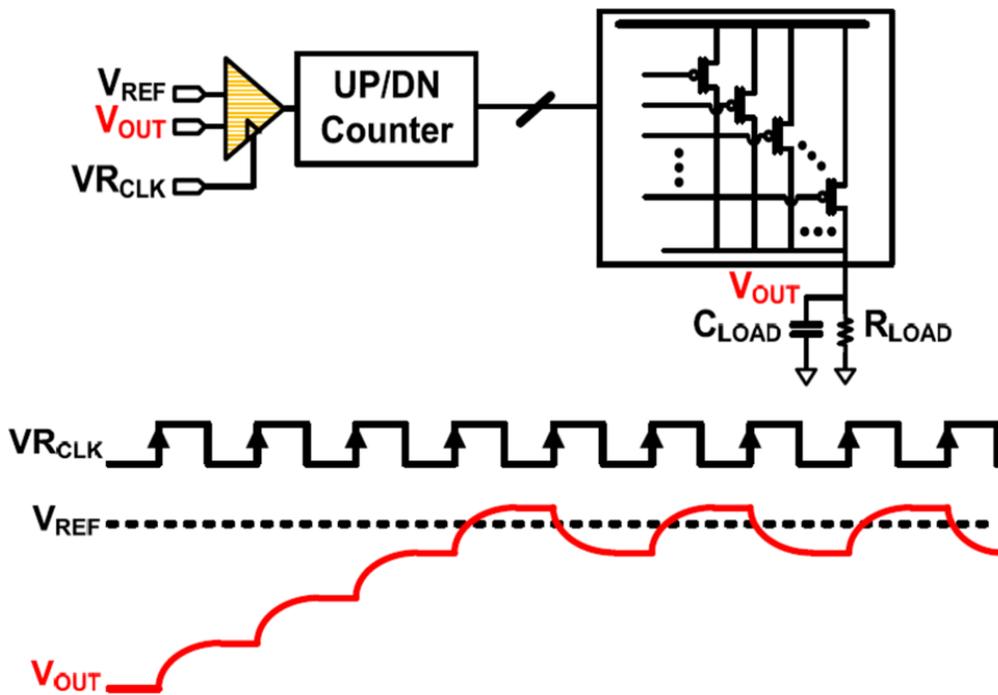
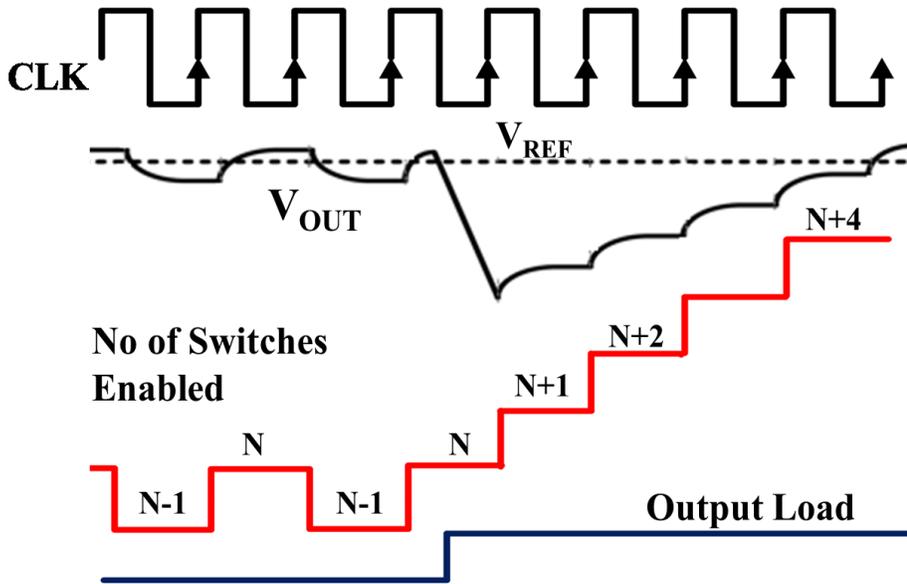


Figure 3.2. D-LDO controller for steady-state response with single bit toggling once steady-state is reached. Bit granularity determines voltage ripple (34)



**Figure 3.3. Transient response to large  $di/dt$  limited by clock frequency due to single-bit response of the controller. Resulting in big droop on  $V_{OUT}$  (34)**

With digital circuits being relatively immune to process induced variations, these D-LDOs are much more robust compared to A-LDO. A significant feature of any LDO is the  $V_{DROPOUT}$ . Reducing this dropout voltage improves power efficiency and allows LDO to be deeply embedded within the SoC (67). Figure 3.4 shows a plot with comparisons of different dropout voltages for different LDO circuits (A-LDO/D-LDO), and the process nodes for the D-LDO. Two interesting observations can be made from this comparative plot. First, as discussed earlier, D-LDO can benefit from the fast energy efficient switching of the advanced technologies, as evident from most of the implementations. Secondly, the reduction in dropout voltage in D-LDO makes a strong case for this class of regulators for energy-efficient integrated implementations. Unlike the A-LDO, where op-amp/controller design is strongly coupled with power switch sizing and any overdesign can potentially reduce stability of the design, D-LDO being digital in nature, decouples the switch array design from the controller block through properly

designed buffers. This leads to improved dropout performance in D-LDO (200-250mV in A-LDO as compared to 50-150mV in D-LDO).

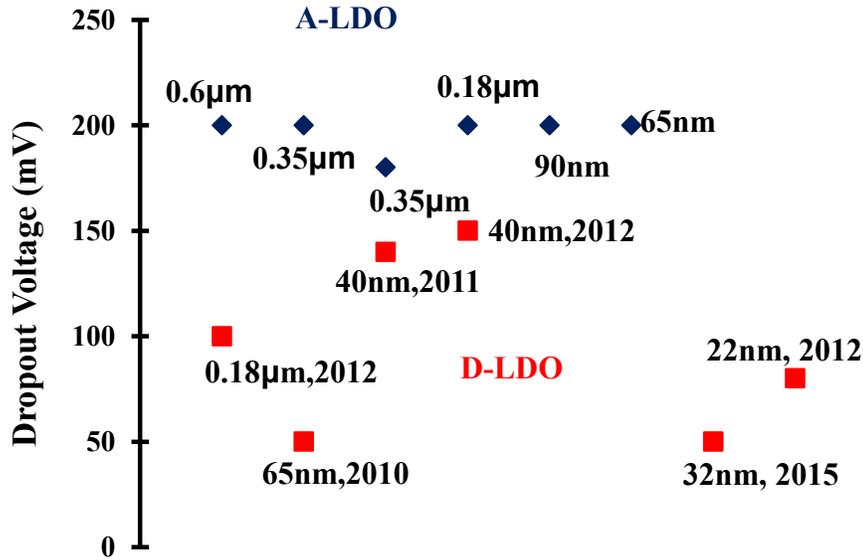


Figure 3.4. Dropout voltage comparison between A-LDO and D-LDO

### 3.3. Digitally-controlled LDO for reduced standby power-drain in ultra-low-power MCU

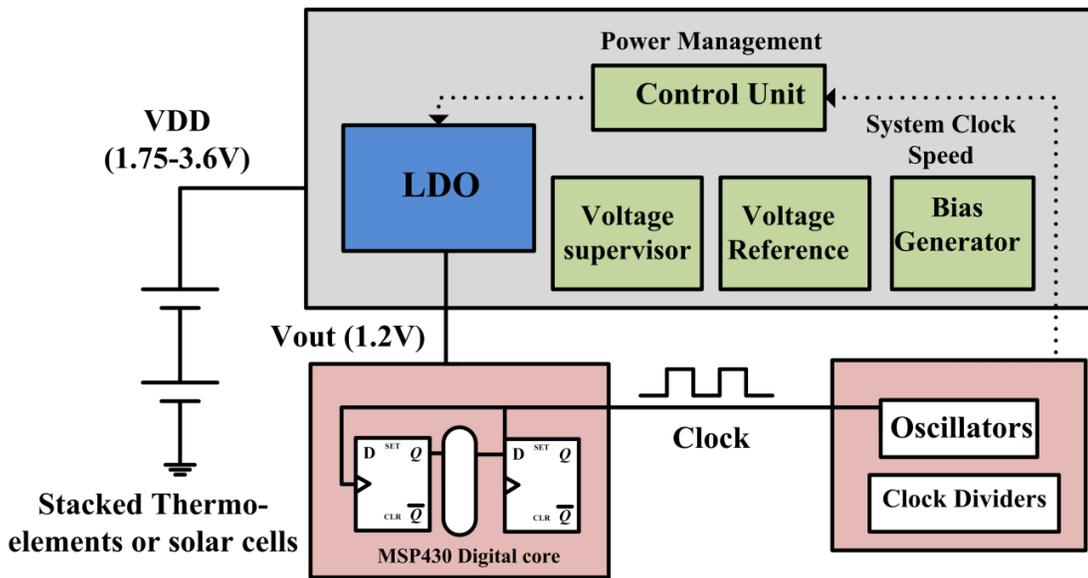
#### 3.3.1. Motivation

Energy harvesting is widely used in applications such as wearable electronics and wireless sensor nodes, to overcome the battery lifetime limitations of these energy-constrained devices. The input energy sources can be variable (solar power, thermal

energy, wind energy etc), while the output power level is of extremely small value. Therefore, adopting an efficient power management strategy is extremely crucial for maximizing battery lifetime. Figure 3.5 shows an energy-harvesting node with ULP-MCU (TI MSP430) microcontroller used for data processing (74, 46). LDO provides low voltage to the digital core while taking in a wide range of variable input voltage. Typically, energy harvested applications spend most of their time in sleep mode consuming only a few micro-amperes and waking up periodically to process the data. Therefore, the system needs to have an ultra-low-power sleep mode and the ability to ramp up very quickly into active-mode when it wakes up. To save power, the PMU disables most of the MCU digital-core during standby-mode (sleep) and LDO needs to power the bare-few minimum circuitries. Therefore, LDO quiescent current becomes a dominating factor in deciding the standby-mode battery lifetime, especially for an energy-harvested system.

Conventionally, A-LDO uses a big off-chip stabilization capacitor ( $C_{OUT}$ ) to provide the dominant pole for loop-stability. At each transition into sleep mode,  $C_{OUT}$  is discharged and, consequently, has to be recharged at wake-up, wasting energy. Therefore, while leakage energy of the digital core is saved during the standby-mode, energy wasted during the wake-up time prevents aggressive active-to-sleep mode transitions. For fast, power-efficient mode transition, integrated LDO is therefore preferred. However, the absence of large external capacitance presents stability issues for conventional LDOs. There are several LDO topologies, known as capless LDO, which use some form of miller compensation to establish an internal dominant pole and replace the external capacitor with a small on-chip capacitor (20, 79). They suffer from stability

issues, though, at low load currents when the non-dominant output pole starts shifting toward the origin, and need some minimum load or “dummy” current to guarantee stability. Now this “dummy” current (effectively wasted as LDO quiescent current) starts dominating the overall power consumption and determines the battery lifetime during the standby-mode, for example, (20, 79), where  $I_q$  currents are 6% and 28% of target load current. Quiescent current ( $I_q$ ) in A-LDO is further dictated by the different biasing currents, and reducing  $I_q$  will have a significant impact on transient response.



**Figure 3.5. Energy harvesting system with MSP430 (TI) (46). Integrated LDO provides usable voltage to the MCU core**

The digital controller removes the need for minimum load current, thereby providing an attractive alternative. Moreover, since digital load (here MSP430) provides a lot of information regarding system power-modes, a digital control can efficiently and adaptively integrate the power-modes’ information in a D-LDO. However, the switching activity of the controller draws considerable quiescent current even during standby-mode.

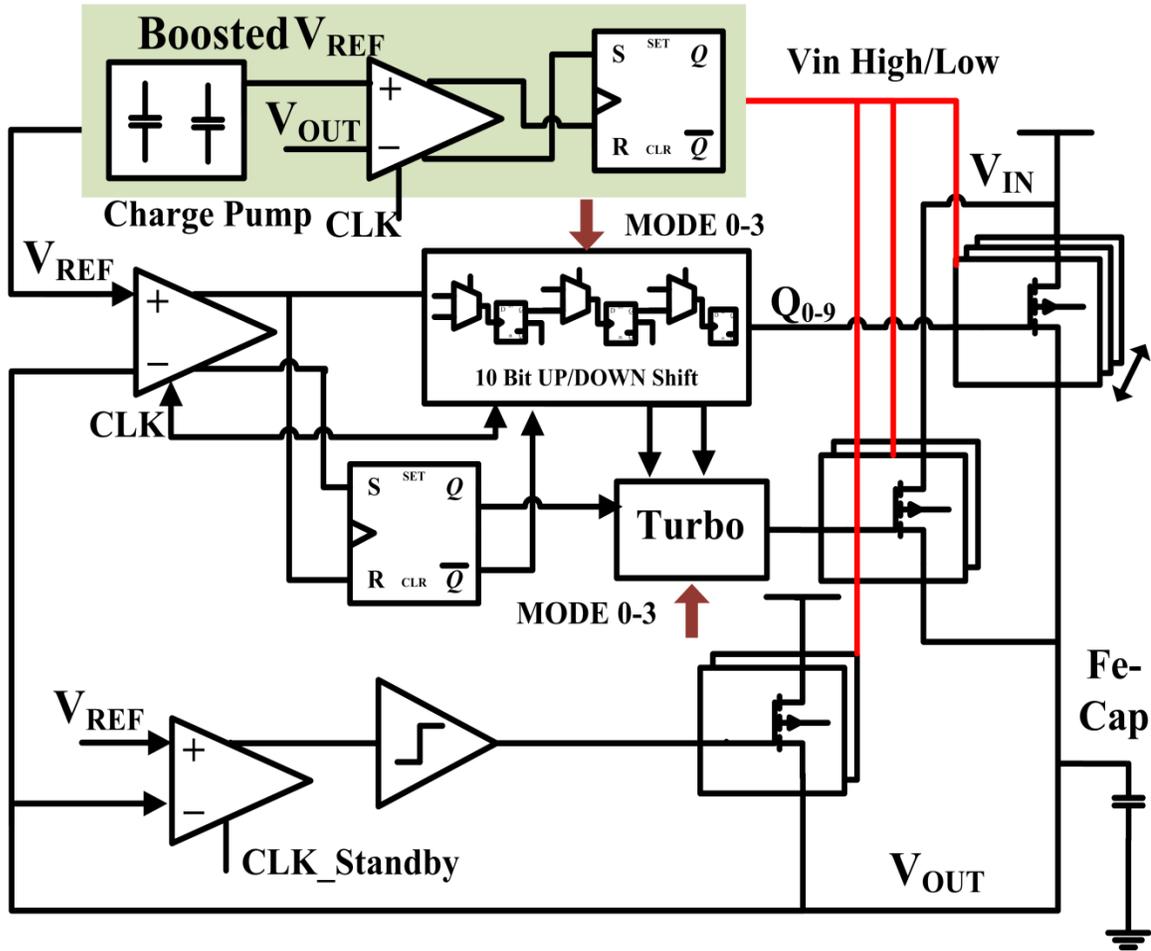
A D-LDO for near-threshold to sub-threshold voltage-conversion consuming  $30.8\mu\text{A}$  of quiescent current has been reported in (36). The D-LDO in (59) consumes  $130\mu\text{A}$  of quiescent current, and needs more than one supply voltage for stable regulation, making this LDO costly for deployment. While these D-LDOs meet the quiescent power specifications for regular low power systems, the energy-harvesting architecture consumes only tens of micro-amperes in standby-mode, therefore requiring a LDO with an ultra-low quiescent current consumption. In (57), quiescent current consumption has been drastically reduced to  $2.7\mu\text{A}$  through controlled clocking, making this D-LDO ideal for energy harvesting applications. However, severely degraded transient response, large area overhead, and poor FOM of the digital controller make it a less attractive choice for practical implementation.

### 3.3.2. Digitally-adaptive LDO architecture with ultra-low quiescent current

The proposed topology as shown in Figure 3.6 is based on existing flash-ADC based D-LDO topology (57). The major contribution of this dissertation lies in improving the FOM (i.e. transient response, area overhead, and quiescent current consumption) for this class of D-LDO through novel circuit and architectural techniques. A dual-loop D-LDO is proposed here, making use of known system power information to combine capless D-LDO with high current efficiency under all load conditions to enable a highly flexible and energy-efficient system operation.

The D-LDO consists of two parallel paths, the fine-grained loop (LDO active- mode corresponding to system active-mode), for handling the steady state regulation, and the

coarse-grained loop (LDO standby-mode corresponding to system standby-mode), with minimum power overhead during the standby-mode. Active-mode also uses an additional turbo mode for fast response to large  $di/dt$ . The power-switch array matrix consists of PMOS as the switch for lower dropout voltage. 1.2nF of ferro-electric capacitor (Fe-cap) (12) is used as explicit decoupling capacitor (decap), while 2nF capacitor is assumed from switching load capacitance. A charge-pump based voltage monitoring circuit is used to provide information about  $V_{IN}$  (supply voltage range), since energy-harvesting architecture can have a wide range of varying input voltage.



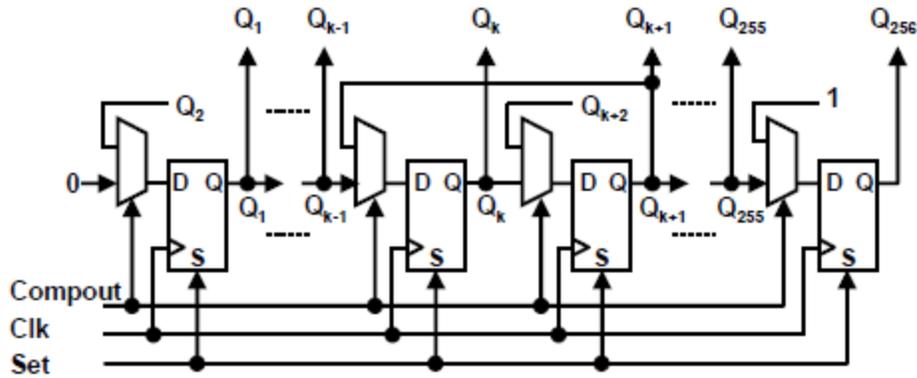
**Figure 3.6. Proposed system power-aware D-LDO with integrated Fe-cap at output. Different LDO modes (active, standby) scale with corresponding MCU system modes. Turbo mode provides high-gain stages**

### 3.3.3. D-LDO Design Components

Since a digital controller usually has a higher power overhead than analog blocks, all the components are optimized for lower power. Detailed descriptions of each of the design blocks and the working principle of this D-LDO architecture are discussed in this section.

**Comparator:** Due to fast-speed, low-power consumption, high-input impedance, and full-swing output, CMOS dynamic latched comparators are attractive alternatives to fully-differential comparators. Latched comparators commonly use a clock signal to operate on two different modes; precharge mode, where output is reset, and evaluation mode, where output is toggled by using a positive feedback. A simple latch-based sense amplifier has been used here as comparator (26). Being completely digital and clocked circuit, it removes the static power consumption problem in analog comparators. This operation is fast and input-referred offset voltage (arising from device parameter mismatch such as threshold voltage ( $V_t$ ), node capacitance, etc.) is relatively lower. Further reduction in offset is achieved through careful sizing (i.e using large devices) of the input transistor pairs. Input-referred latch offset voltage can be reduced further by using the pre-amplifier preceding the regenerative output-latch stage, but at the cost of large static power consumption.

**Bidirectional Shift Register:** Figure 3.7 shows the bidirectional serial-in-parallel-out shift register using D-FF and mux similar to the one used in (57). Figure 3.2 shows the circuit response to load current change to explain the feedback control of D-LDO. Initially, all the switches are disabled. Now, depending on the comparator response to  $V_{OUT}$  and  $V_{REF}$  (reference voltage), switches will be enabled or disabled on each clock cycle by shifting “0” or “1” left or right. In the steady state i.e. once  $V_{OUT}$  and  $V_{REF}$  are the same for a constant load current, only one bit will toggle, deciding the minimum output ripple voltage.

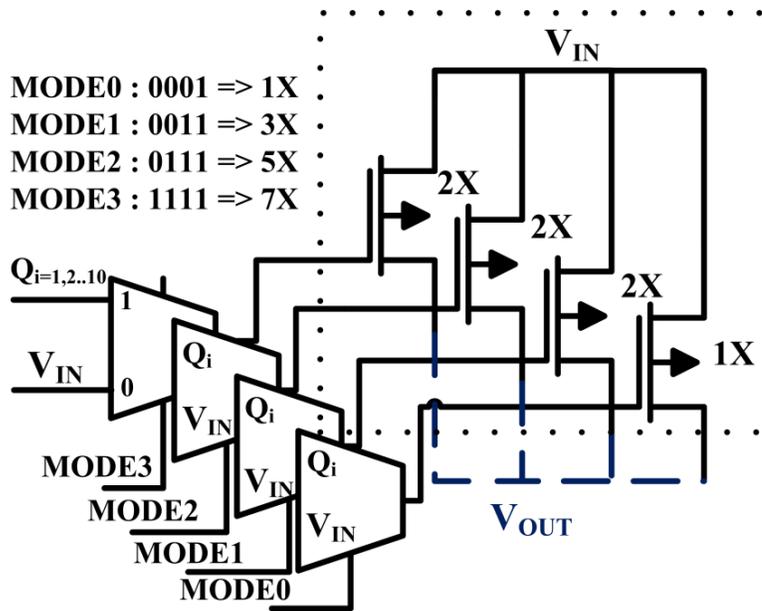


**Figure 3.7. Up-down shift-register allowing single bit to toggle in a bidirectional manner, depending on  $V_{OUT}$  with respect to  $V_{REF}$  (57)**

**Voltage Monitoring Circuit:** Ideally, LDO is used for low  $V_{IN}/V_{OUT}$  ratio for higher efficiency. However, energy harvesting architecture dictates that the regulator be compatible with a wide range of input voltage. For a fixed  $V_{REF}$ , higher the  $V_{IN}$ , higher is the drain-source voltage drop ( $V_{DS}$ ) across the switches and therefore higher the current delivery capabilities as the switches are operating in linear region. Therefore, designing for the worst-case  $V_{IN}$  will lead to poor D-LDO response for all other conditions, both in terms of efficiency and output ripple voltage. The charge pump circuit boosts up  $V_{REF}$  to twice its value (assuming there are no losses) and compares with  $V_{IN}$  (99). Depending on whether  $V_{IN}$  is greater or less than twice the reference voltage, the latched up comparison result is used to modulate the switch strength. This leads to an optimum ripple for a wide range of input voltage. By using a CMOS-compatible integrated capacitor instead of a resistive divider, area and leakage power overhead of resistors have been reduced. Since the voltage monitoring is a clocked circuit, it is enabled for a short duration only during the active-mode, and clock-gated for the remaining time, thereby saving power

**System-Power-Aware Switch Matrix:** The current consumption of a digital circuit (MCU core here) scales almost linearly with frequency. Therefore, the different clock settings information of the MCU-system-clock can be extracted and digitally integrated to modulate the D-LDO drivability; for example, in the setup shown in Figure 3.5, four different clock settings are used during MCU (MSP430) active-mode and 1 clock setting for MCU standby-mode.

This is illustrated in Figure 3.8, where the dual-loop architecture with three different LDO-modes (active, turbo and standby) dynamically modulates the LDO's drive strength in-sync with the four MCU power-modes (MODE0 to MODE3) and MCU standby-mode (here LDO active-mode, LDO power-mode and MCU power-mode are used interchangeably). One of the major limitations of the work in (34) lies in the identical sizing of the power switch-matrix resulting in large bit counts. However, in the proposed D-LDO, by adopting a non-uniform sizing strategy for the power-switch matrix, the required number of bits has been reduced to 10. While the control bits (indicating the system power modes) change linearly, thermometric style, switches have been sized nonlinearly to improve the transient response. This nonlinear sizing is explained in detail in Figure 3.8 and Table 3.1. Figure 3.8 shows one of the ten power-switches (each corresponding to one of the ten controller bits) with four unequally sized finger partitions. By enabling these fingers in an additive fashion, based on the digital bits representing MODE0-MODE3, the drive-strength of the switches increases (MODE0=>1X i.e. minimum drive-strength, while MODE3=>7X i.e. maximum drive-strength with all the fingers enabled).



**Figure 3.8. Single switch hierarchy for digitally-adaptive power management. Thermometric coding and additive enabling of power switches for non-linearly sized switches**

Table 3.1 shows that the modes (with varying switch strengths) and the number of switches enabled change with MCU load current requirements (switches being PMOS, ‘0’ indicates enabled and ‘1’ indicates disabled). This way, the LDO adaptively regulates the output voltage with fine-grained switching at lower active-mode load current, and coarse-grained switching at higher active-mode load current. Therefore, by scaling the quiescent current with load and adaptively changing the drivability and switching frequency of the regulator, this D-LDO architecture can ensure high efficiency and low ripple voltage.

**Table 3.1 Dynamic modulation of switch strength with system power-modes**

Active-mode		Mode Bits				Strength of a power switch					
0		0001				1X					
1		0011				3X					
2		0111				5X					
3		1111				7X					

$I_{LOAD}$ (mA)	Mode	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
0.2	0	0	0	0	1	1	1	1	1	1	1
0.8	1	0	0	0	1	1	1	1	1	1	1
1.6	2	0	0	0	0	1	1	1	1	1	1
3.2	2	0	0	0	0	0	0	0	1	1	1
6.4	3	0	0	0	0	0	0	0	0	0	0

**Turbo-Controller:** Since switches are enabled in a unary fashion, transient response will be limited by the LDO clock frequency, leading to significant voltage droop for large di/dt events. Therefore, a special turbo-controller (Figure 3.9), equipped with knowledge about the MCU power-modes, is designed to monitor the third and the fourth bit from the 10-bit controller and to enable high-gain power switches as soon as it predicts large current changes (since maximum-current/switch for a particular LDO mode is known, load current can be predicted based on the number of enabled bits). Figure 3.10 shows

how the turbo-controller reduces the voltage droop from 220mV to 80mV for worst-case LDO current change (0.2 $\mu$ A to 6.4mA with 1ns rise time). The fact that, by monitoring the controller bits, load current can be predicted at run-time is a very powerful feature of this LDO; this information can be used for adding more run-time configurability features in the future, and removes the need for additional current-monitoring apparatus during testing. Digital information from turbo-controller and MCU power-modes has also been used to automatically select one of the three clock frequencies generated by the on-chip clock, to lower active-mode power overheads.

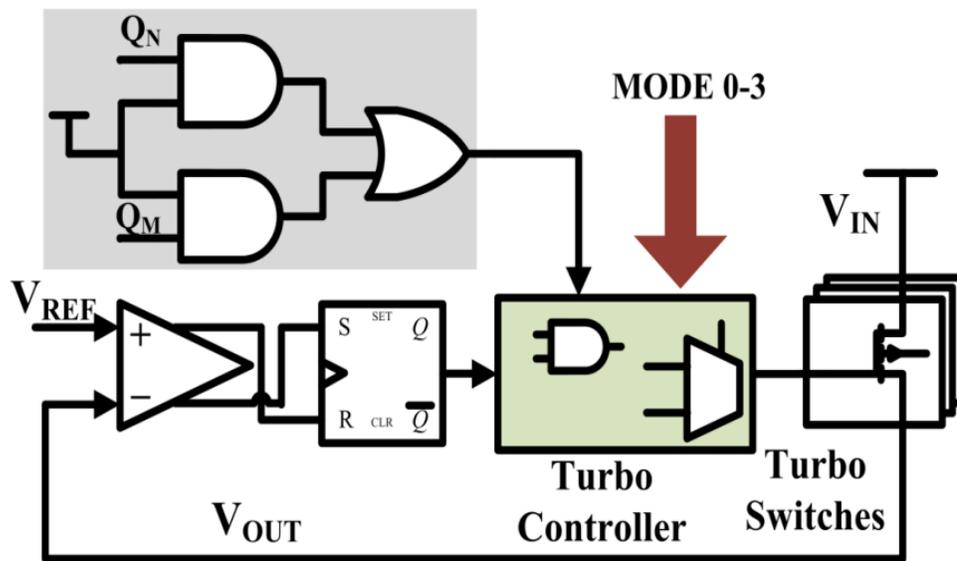
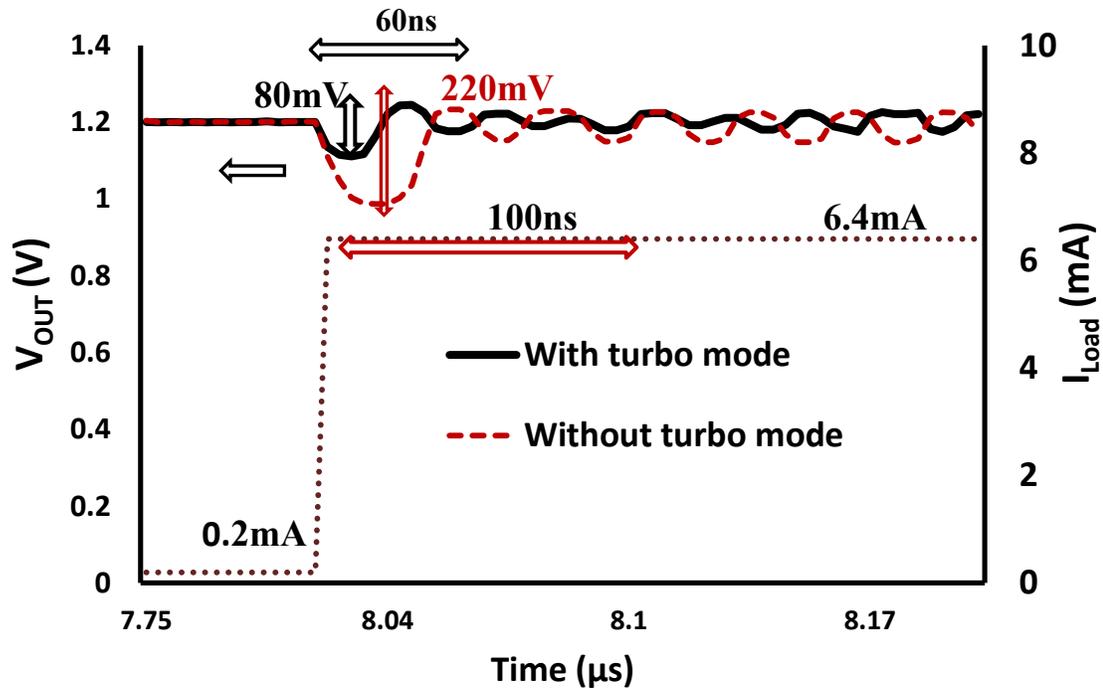


Figure 3.9. Turbo-controller for enabling high-gain stages for large di/dt



**Figure 3.10. Large di/dt droop improvements with turbo-controller enabled high gain stages**

**LDO standby-mode:** This coarse-grained loop of the proposed LDO is custom designed for ultra-low-power, with clocked comparator and current-starved ring-oscillator (RO) based clock, and is enabled only during MCU standby-mode (when the rest of the LDO and most of the MCU are disabled). With input voltage varying over a wide range, the starved clock can consume significant power. Therefore, the LDO regulated output voltage (1.2V) is used as the supply-rail for the RO clock, consuming as little as 240nW for 0.8MHz clock frequency with LDO output of 1.2V. A level shifter is inserted for the different voltage domains to interact. This feedback connection will sustain as long as the MCU does not wakeup in standby-mode. Most of the energy-harvesting/sensor driven applications using ULP MCUs will likely be waking up only when there is a need for data-processing i.e. active-mode, therefore making the above-mentioned assumption a

safe one. Even if an application wakes up in standby-mode, the LDO will still work, just with a larger droop which should be acceptable since standby-mode implies minimal data processing.

**Startup Controller:** The startup signal from the MCU is used to shut-off the D-LDO controller as soon as the output reaches the desired value ( $V_{REF}$ ). This way the typical overshoot during digital LDO startup can be avoided, as shown in Figure 3.12. Depending on the applications, and by making a range of clock frequencies available, startup time can be traded off with in-rush current.

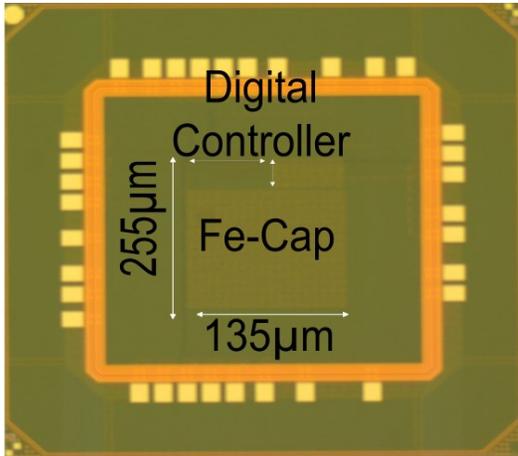
This way, the proposed D-LDO supports steady-state MCU power modes by adaptively reconfiguring the LDO active-mode switch matrix, improves droop response using turbo mode assists in large  $di/dt$  events, and drastically reduces quiescent current overhead during MCU standby-mode by reconfiguring LDO in its standby-mode.

#### 3.3.4. D-LDO simulation, implementation and measurements

The proposed D-LDO, implemented in  $0.13\mu\text{m}$  CMOS technology (low power process node from Texas Instruments), has the ability to regulate up to  $6.4\text{mA}$  of current in its active-mode and  $50\mu\text{A}$  of standby-mode current for a wide range of input voltage ( $1.75\text{V}$ - $3.3\text{V}$ ), providing nominal low voltage of  $1.2\text{V}$  to the digital MCU core.  $1.2\text{nF}$  of Fe-cap as decoupling capacitor (digital load contributes  $2\text{nF}$  further) and  $20\text{pF}$  Fe-cap in the voltage-monitoring charge pump have been used. Owing to the high-density of the Fe-cap, the chip occupies only  $0.034\text{mm}^2$ . The fabricated D-LDO is shown in Figure 3.11,

where the bulk of the area is occupied by the capacitor. Standard-cell based design allows automatic place-and-route for most parts of the LDO, improving design turn-around time. The MCU power mode, with its different current levels (0.2mA, 0.8mA, 3.2mA, 6.4mA) as in Table 3.1, corresponds to active-mode (MODE0-MODE3) of D-LDO, while standby-mode of D-LDO supports the MCU standby-state ( $\sim 40\mu\text{A}$ ). Two on-chip starved ring oscillators (RO) are used as clock, one for active-mode D-LDO (RO\_Active) with four frequency options (5MHz, 10MHz, 20MHz and 40MHz), and the other for standby-mode D-LDO (RO\_Standby) with 0.8MHz frequency. The MCU digital core is expected to run up to 16MHz and, therefore, the power mode settings can change anytime within one clock cycle (62.5ns) (46). During MCU active-mode, the D-LDO active-mode takes control of the regulation by adaptively configuring the regulator load-drivability, and turbo mode is enabled only for large di/dt event. In MCU standby/sleep mode, the D-LDO standby-mode regulates the few MCU circuitries which are enabled, while the rest of the D-LDO is disabled. The charge-pump based voltage monitoring circuit compares the boosted reference voltage against the supply voltage to ascertain the input voltage range (VIN HIGH/LOW in Figure 3.6), indicating whether the supply voltage is in the higher or in the lower range. Depending on this information from VIN HIGH/LOW, the power-switch drive strength is further modulated to reduce voltage ripple, for example from 7% to 4.4% for 3.2mA at 3.3V supply. Figure 3.12 shows the D-LDO response to a varying load current, incorporating both steady state and large di/dt current events. Typically worst-case droop will happen when mode changes from the lowest current level (MODE0) to the highest one (MODE3) within a clock cycle, as indicated in Figure 3.12. While voltage undershoot (or voltage droop) is understandable, voltage overshoot is

also of concern for D-LDO, especially when the mode changes from the highest current level to the lowest one. As  $V_{OUT}$  rises above  $V_{REF}$ , the controller starts disabling power switches; however, the response is limited by clock frequency. With very low current, it may take a couple of clock cycles for the overshoot to stabilize, as seen in Figure 3.12.



**Figure 3.11. Fabricated D-LDO chip in 130nm low power CMOS process (TI)**

Figure 3.13 shows the measured transient response of the LDO with a supply of 2V and off-chip load having multiple transients between 0.3mA and 5.4mA with 1ns rise-time. While the specified load current for the LDO was 6.4mA, the maximum supported load current as measured in silicon is 6mA. This can be speculated to bit-failure or variation-induced error. A startup time of 13µs without any overshoot, as shown in Figure 3.13, is made possible due to the small integrated capacitor and the startup controller. A settling time of 100ns with voltage undershoot of 90mV and ripple of 55mV has been measured for a load current change from 0.3mA to 3.2mA. Inherent immunity of digital logic gates to supply noise leads to simulated PSR of -25dB at 10MHz with 4mA load current, mostly due to the self-created ripple. While digital load can usually

handle up to 10% p-p ripple, the ripple voltage can be lowered through finer-granularity power-switch/modes for noise-sensitive applications.

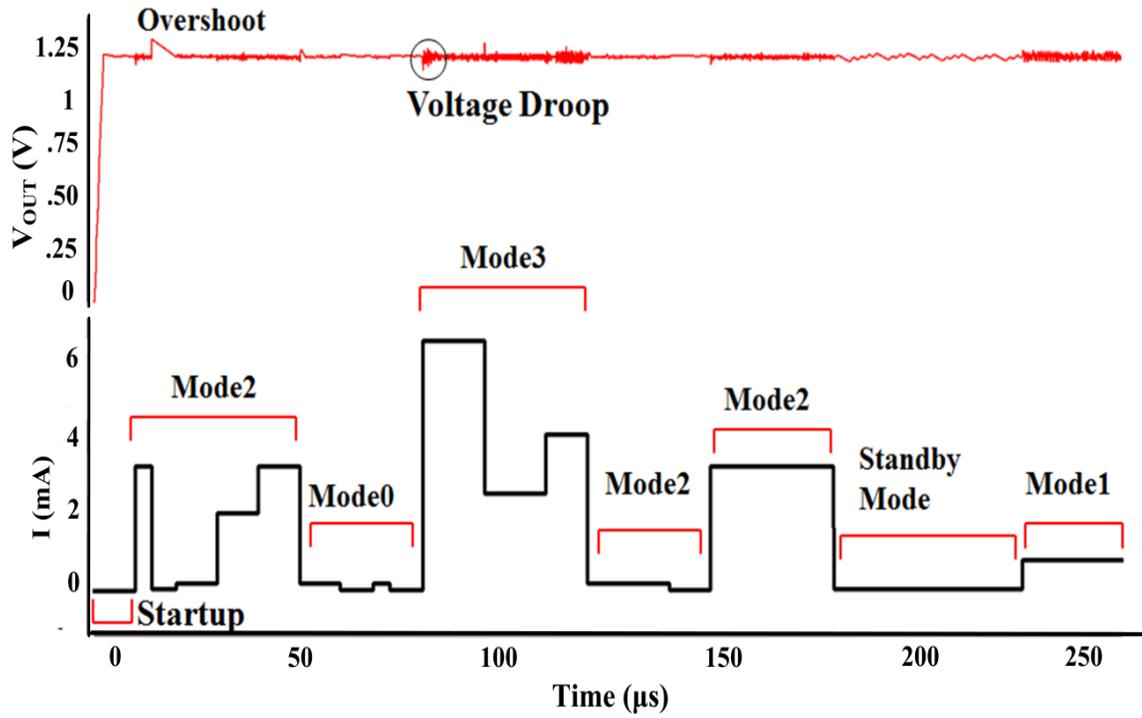
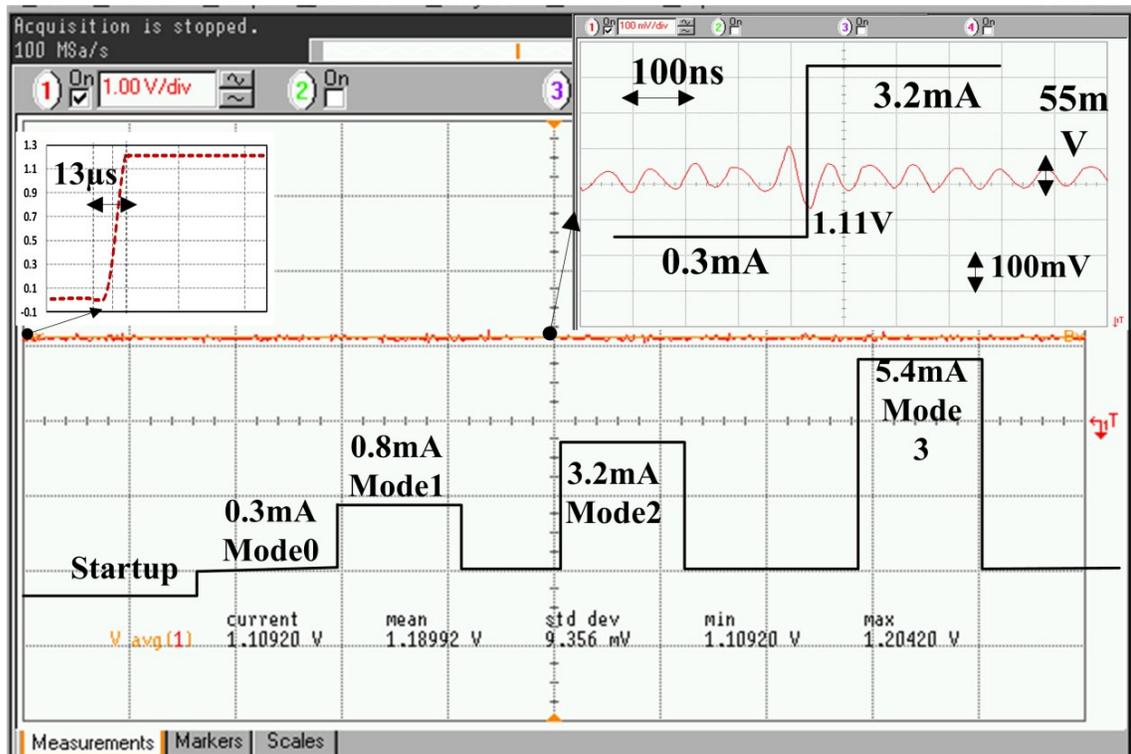
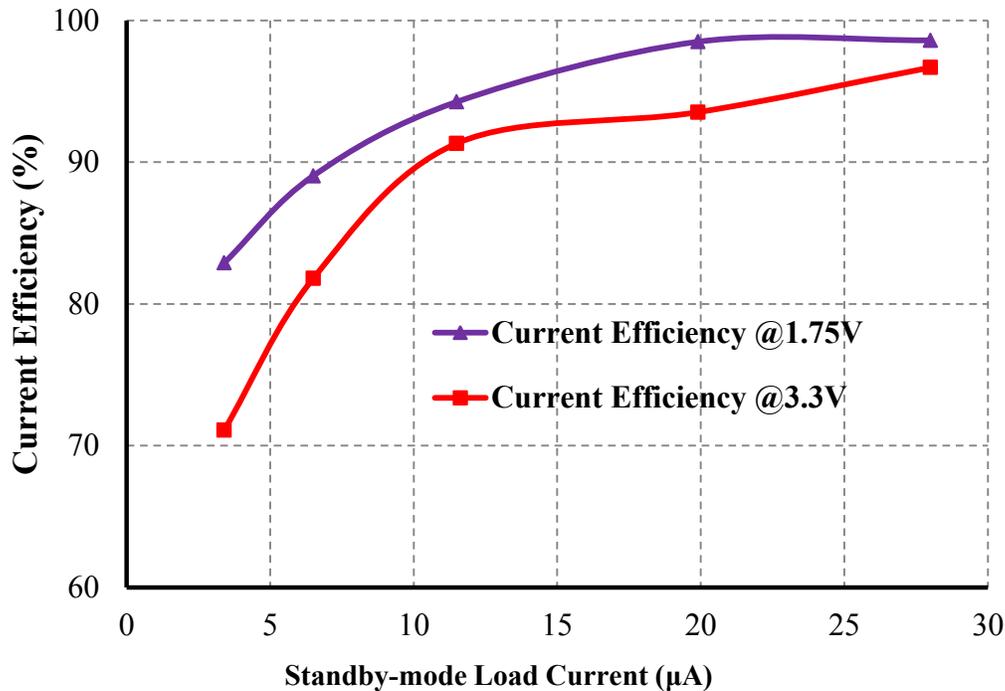


Figure 3.12. Regulated  $V_{\text{OUT}}$  for different load current (simulation)



**Figure 3.13. Measured plot showing D-LDO functionality. 13µs startup time achieved (on left) with ripple voltage (4.5%) on  $V_{OUT}$**

Figure 3.14 shows the measured standby-mode current efficiency of the LDO ( $\eta_{current}$  from Equation 2-6), while driving maximum load current of 40µA with different supply voltages. As load currents in the standby-mode drop to as low as few micro-amperes, the LDO quiescent current starts dominating the leakage current. By drastically scaling  $I_q$  in standby-mode, the proposed LDO has achieved an ultralow  $I_q$  of 500nA at 2V and 1.8µA at 3.3V, thereby allowing high current efficiency (more than 90% for load current as low as 10µA) for a wide-range of input voltages.



**Figure 3.14. More than 90% current efficiency measured in standby-mode**

### 3.3.5. Comparison with state-of-the-art LDO

In recent years, significant research has been invested in design of fully-integrated A-LDO and D-LDOs targeting a wide-range of applications. Therefore, the figure-of-merits of these designs varies a lot, making it hard to do a fair comparison. For example, LDOs in (17, 20) have targeted relatively higher current level and, therefore, the primary focus is on improving transient response. Similarly, D-LDOs of recent times have mostly targeted low input voltage applications (0.45V-0.7V) since digital implementations have a distinct advantage over analog LDOs there (57, 67). The D-LDO proposed here is targeted for powering ULP-MCU for a wide-range of  $V_{IN}$ , as applicable in energy-harvested systems. Since power efficiency of LDO is constrained by input to output

voltage ratio, it is the current efficiency that drives the battery lifetime, especially during system sleep-mode. Sensor based applications work in burst-mode i.e. they may have sudden very large load activity. Therefore, improving transient response is equally important. The two FOM ( $FOM_1$  and  $FOM_2$ ) proposed in (20) and given by Equation 2-4 and Equation 2-5, combine the impact of transient response and quiescent current ( $I_q$ ), and allow a fair comparison across works in literature.

A performance comparison with state-of-the-art LDOs is listed in Table 3.2, where only the LDOs providing similar output power have been considered. The transient response ( $FOM_1$ ) of the proposed D-LDO has improved upon the original D-LDO converter (57) by 60000x with 4-5x lower quiescent current in standby-mode. By making use of known system power information, the proposed D-LDO has achieved a measured FOM of 4.44ps, second best to 3.01ps of (47), but  $I_q$  of the proposed D-LDO shows 100x improvement over (47) ( $0.5\mu A$  compared to  $50\mu A$ ), thereby drastically reducing the standby current drain.

**Table 3.2 Performance comparisons among state-of-the-art low power LDOs**

Ref	Process (nm)	LDO Type	V <sub>IN</sub> (V)	V <sub>OUT</sub> (V)	I <sub>Load</sub> (mA)	I <sub>q</sub> (μA)	ΔV (mV)	C <sub>OUT</sub> (nF)	FOM <sub>1</sub> (ns)	FOM <sub>2</sub> (ps)
(13)	130	A-LDO	1.15	1	25	50	15	4000	2400	4800
(79)	45	A-LDO	1.8-1.62	0.9 - 1.1	42	1200		1.46	0.428	62.5
(36)	90	D-LDO	0.5	0.44	3	30.8	70	100	2333	23955
(57)	65	D-LDO	0.5	0.45	0.2	2.7	40	100	20000	27000
This Work	130	D-LDO	1.75-3.3	1.2	6	0.5	100	3.2	53.3	4.44

### 3.3.6. Stability analysis

Digital loads, as compared to analog loads, can undergo large dynamic changes in load current, depending on the nature of activity, for example standby-mode to active-mode transitions. Therefore, it is important to do an in-depth analysis of digitally-controlled LDO with the help of a control model. Obtaining phase margin (PM) through Bode plots is not feasible in a discrete control system like digital LDO. A z-domain control model of the LDO is needed to illustrate the relationship between key design parameters and transient response. While developing such a model is beyond the scope of this dissertation, stability of the proposed D-LDO architecture is justified through exhaustive circuit simulation and following a design strategy based on an existing discrete-domain model of digital LDO (54).

As explained in (52), the notion of phase margin can be approximated from time-domain response. Based on unity step response and from the peak-overshoot measured, the damping factor ( $\zeta$ ) can be calculated. Phase margin is given by

$$PM = 100 * \zeta \quad 3-3$$

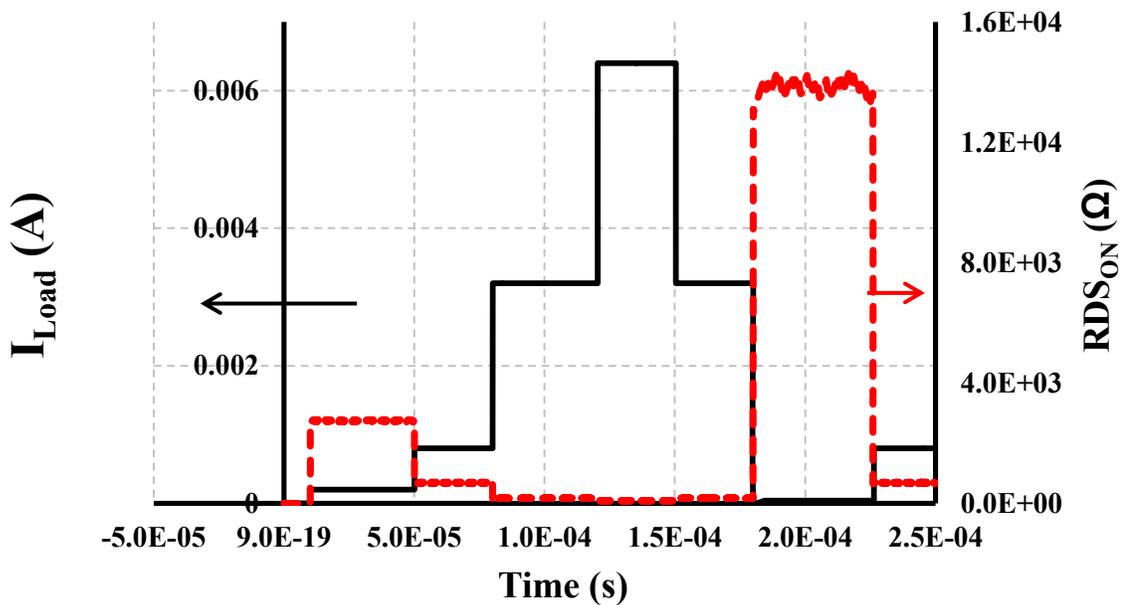
This is an approximation that works for a second-order system in finding out potential instability from transient response (54). Phase margin and stability are acutely related with the relative positioning of the two primary poles in this system,  $F_s$  (clock frequency) and  $F_1$  (output load pole) where  $F_1$  is given by

$$F_1 = \frac{1}{(R_L || (R_{PMOS}))} \quad 3-4$$

$R_L$  is the resistive load and  $R_{PMOS}$  is from the power switches.

Since the power switches are sized for the  $I_{Load}$  specification, the primary design knob here is  $F_s$  or the switching frequency. Following the design space explorations carried out in (54) for optimizing a digital LDO design, a ratio of 5 to 10 between  $F_s$  and  $F_1$  ( $F_s/F_1$ ) is found to provide the optimum tradeoff between transient performance and power efficiency. Another interesting observation is the monotonic dependency between switching frequency and voltage ripple. Typically, the ripple will be lower as the clock frequency is increased until a limit is reached. Therefore, choosing clock frequency is a tradeoff between optimum transient response, higher power efficiency, and lower voltage ripple. A constant sampling frequency can provide a stable LDO with stable output pole at higher frequency, but the same sampling frequency can lead to oscillatory behavior at lighter load. Therefore, to ensure stability across this dynamic nature of load, adaptive design techniques are needed so that  $F_s$  tracks  $F_1$  for a wide range of loads (54). Several adaptive techniques modulating gain (using turbo-controller, voltage monitoring circuit,

nonlinearly sized thermometric switching) dynamically have been discussed so far in this chapter. The power-modes information, available digitally from the MCU, has been used to select among the on-chip generated discrete clock frequencies (0.8MHz to 40MHz). An adaptive controller (similar to turbo-controller in Fig.1) can be added in the future to further tune the clock frequencies based on the location of the output pole, thereby providing stability across a wider range of load currents. In Figure 3.15, the  $R_{DS_{ON}}$  (dropout voltage divided by rated max load current) has been plotted for different load current transition and the monotonicity of the LDO output impedance (with change in load current) is a first order estimation of stability of this digitally-controlled regulator.



**Figure 3.15. Monotonic change in  $R_{DS_{ON}}$  with dynamic load current change indicates stability**

### 3.3.7. Future scope of improvements

A primary contribution of the proposed D-LDO lies in the drastic reduction in standby quiescent current and in the ability to adaptively reconfigure itself by using known system power information. However, as seen from Table 3.2, there is plenty of scope to improve transient response in order to make D-LDO more competitive with A-LDO.

1. Instead of using resistive switches (MOSFET working in linear region) in the power switch-matrix, current-controlled switches might provide better noise performance by isolating output node from input.
2. PMOS, as opposed to NMOS as switch, allows dropout voltage reduction ( $\sim 100\text{-}150\text{mV}$ ) in D-LDO. However, design specifications of this particular regulator with  $V_{\text{IN}}$  of  $1.75\text{V-}3.3\text{V}$  and  $1.2\text{V}$  as nominal  $V_{\text{OUT}}$  allows NMOS switches to be used for the LDO and being  $\sim 2\text{x}$  smaller than PMOS (for same drain current), it may provide better transient response due to reduced gate capacitance.
3. Since voltage ripple is a major concern for digitally-controlled LDO, bidirectional up/down shift-register has been used in the controller to ensure a single switching event every clock, thereby reducing steady-state ripple. However, this single-bit-switching nature of the controller makes the transient response worse, since it is limited by clock frequency. Therefore, it is important to include the ability to enable more-than-one bit per clock cycle in the controller (54).

### 3.3.8. Digitally-controlled LDO for hybrid IVR with fast droop mitigation

*This work has been done entirely at Intel CRL (circuit research lab) Portland, as part of a research internship. While the design of D-LDO is credited to the author, the silicon implementation and measurements (some of which have been included in this dissertation for understanding the topology and showing the improvements) have been performed at Intel Labs.*

### 3.3.9. Background

The semiconductor industry is increasingly applying wide-range DVFS, from a near-threshold voltage (NTV) region to super-threshold voltage, for improved energy efficiency of the digital core. When powered with a shared rail, some energy is wasted, as different blocks demand different voltages. Therefore, per-core IVR is a cost-effective solution to achieve autonomous DVFS (80). While conventional inductive converters suffer from scalability limits (73), the switched-capacitor voltage regulator (SCVR) has gained popularity among SoC designers due to the rapid improvement in capacitive technology. Prior works by Intel showed SCVR with four voltage-conversion-ratios (VCR) using a high density on-die MIM capacitor, residing between M8 and M9 layers above the load and delivering maximum load current of 88mA (25) and even 200mA (26) at high efficiency (~67%-84%) and across a wide-voltage range. In an area-constrained design, however, the limited size of the SCVR's fly-caps and the need for different configurable power stages sets an upper bound on the SCVR's maximum power delivery capability, restricting its use to lower  $V_{OUT}$ . LDO, on the other hand, benefits from high

power density with small area overhead as long as  $V_{IN}:V_{OUT}$  ratio is low, i.e. efficiency drops at lower  $V_{OUT}$ . This contrasting performance from these two regulators makes a strong case for a hybrid LDO/SCVR allowing a wide-range on-chip DVFS. Earlier work has combined A-LDO with SCVR stage for improved ripple performance from the switching regulator (6). However, because of the analog nature of conventional LDO, SCVR switches cannot be reused, leading to redundant hybrid design with significant area overhead for delivering high power. A truly hybrid LDO/SCVR structure has been proposed here where, owing to the digitally-controlled nature of LDO, all the design blocks from SCVR, including the power switches, have been reused, leading to a compact, highly efficient IVR supporting wide ranges of voltages on-chip (34). Figure 3.16 shows the maximum efficiency achievable (theoretically) using this kind of hybrid IVR.

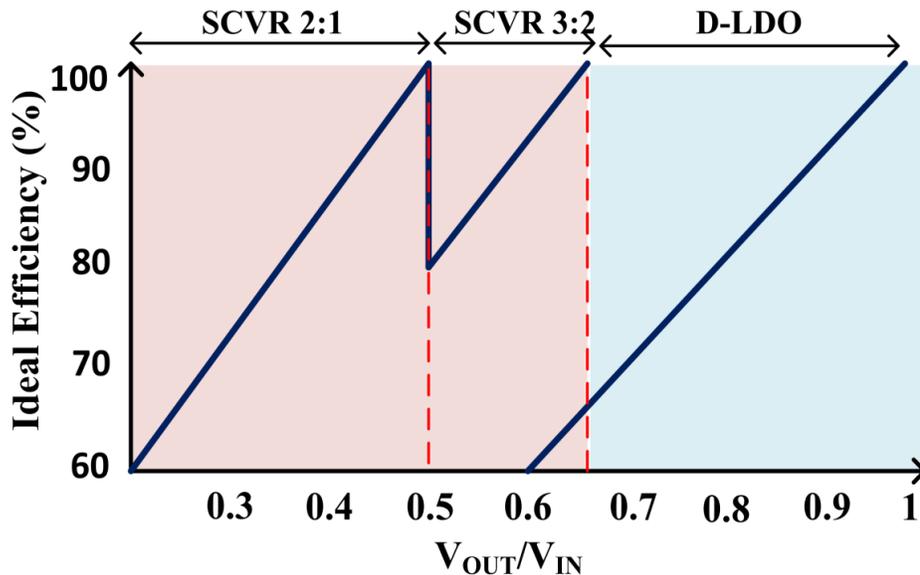


Figure 3.16. Hybrid IVR with LDO (high  $V_{OUT}$ ) and SCVR (low  $V_{OUT}$ ) (34)

As mentioned earlier, hybrid LDO/SCVR has been a collaborative project where this author's contribution lies mostly in designing the D-LDO part of the hybrid converter, developing hysteretic-based fast droop mitigation technique, reducing output voltage ripple using known system power information, and integrating all these techniques with the SCVR stages.

### 3.3.10. Digitally-controlled Dual-Loop LDO with fast droop mitigation technique

Compared to analog loads, digital loads undergo larger dynamic ranges resulting from different processor activities, for example the transition between different power modes. This sudden  $di/dt$  causes droop on the voltage rails. To compensate for these droops, large guard-band is allowed, leading to significant energy wastage in overdesign. Therefore, it is important to have fast droop mitigation technique in the point-of-load regulators.

Existing D-LDO architecture, discussed exhaustively in section 3.2, has been used in the proposed hybrid structure for providing higher  $V_{OUT}$  (super-threshold voltage domain), while the SCVR from (25) regulates the lower output voltage domains (near-threshold voltage domain). Figure 3.17 shows the SCVR power stage providing 2:1, 3:2, 3:1 VCR by reconfiguring the switches for the different ratios. In the D-LDO mode switches S3, S7, S9 are disabled, while S2, S4, S6, S8 are enabled, with switches S1, S5 regulating the output nodes. By enabling in this fashion, the high density MIM-cap used in SCVR mode are connected to the output as decap through S3 and S7 during D-LDO

mode, thereby improving the transient response and droop characteristics. A power management unit (PMU) decides the output voltage range for the loads and accordingly IVR is configured into SCVR or D-LDO mode using a  $V_{OUT}$ -lookup table. A clocked comparator-based lower-bound hysteretic control is used for fast frequency modulation-based feedback and the same comparator is reused in the D-LDO mode for controlling the number of enabled/disabled switches.

A major drawback of this D-LDO structure lies in its controller shifting single bit every clock cycle, as discussed earlier (Figure 3.2 and Figure 3.3), resulting in large droop to fast  $di/dt$  change. Therefore, a 1.5bit flash ADC based regulation with dual-loop D-LDO architecture has been proposed in this work for achieving fast droop mitigation efficiently, as shown in Figure 3.18. A “fine-grained” controller regulates in steady-state mode by enabling/disabling single bit, while an additional comparator, with a  $V_{REF} - \Delta V$  reference voltage, is used in parallel to kick in a “coarse-mode” counting mode (or CM) where 4 bits are turned on/off at a time in response to a droop. The CM comparator runs at a higher frequency than the “fine-grained” comparator for maximum droop sampling rate. A simple arbiter guarantees a smooth transition to “fine” counting once the droop event is over. The idea behind this dual-loop strategy is to have a minimum number of bits toggling (in this case only one) during steady state mode with fast recovery, i.e. multiple bits enabled (in this case four bits) simultaneously once a droop is detected. The response to sudden  $di/dt$  by the dual-loop D-LDO is explained in Figure 3.19.

CG mode kicks in when output droop crosses  $\Delta V$  voltage limit. Choosing this  $\Delta V$  limit carefully is necessary since while a smaller value can improve the droop, it can also

lead to increased ringing and ripple if set too low. In this proposed design, 20mV margin has been set for CG mode to reduce droop.

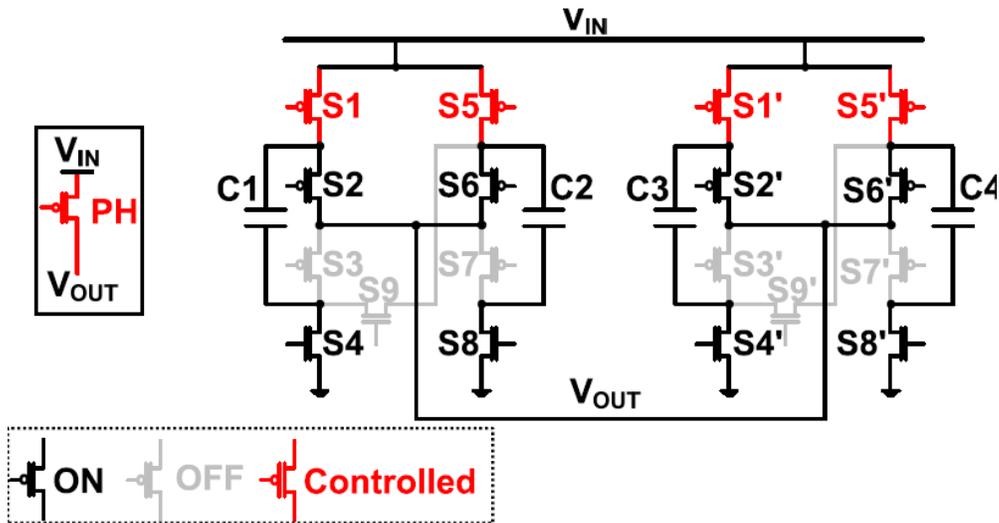


Figure 3.17. LDO mode, reusing SCVR topology for "truly" hybrid IVR (34)

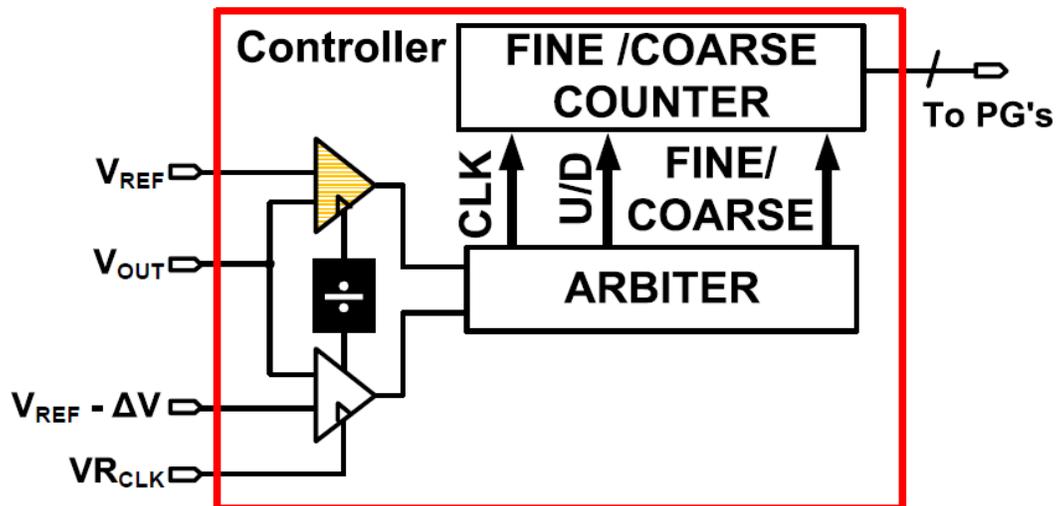
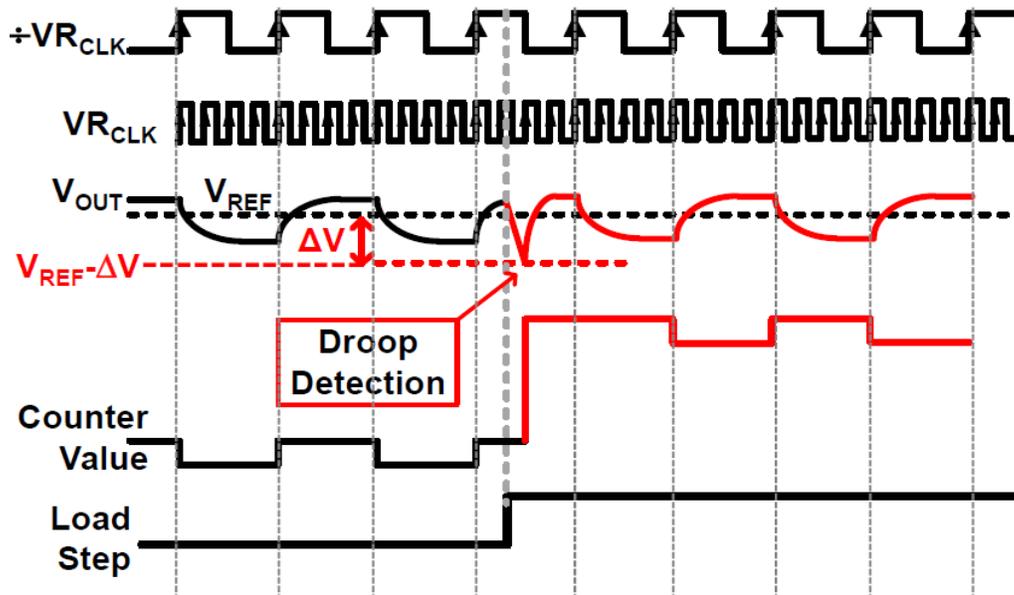


Figure 3.18. Proposed dual-loop D-LDO, with fine-grained slow inner loop for steady state and coarse-grained, fast outer loop for droop mitigation (34)



**Figure 3.19. Droop mitigation with dual-loop LDO (34)**

Clocked comparator-based hysteretic feedback can potentially cause instability, if not designed properly. This instability arises from the asynchronous nature of clocking of this comparator and the various latencies incurred in sampling and through the controller and switches. These latencies can either cause  $V_{OUT}$  to dip significantly before the regulator can respond (clock slower than latencies) or too many sub-harmonic oscillations (clock much faster compared to delay). Therefore, for stable regulation,  $T_{CLK}$  must be greater than all latencies ( $T_{controller} + T_{switch}$ ) but within bounds.

### 3.3.11. D-LDO Implementation and Measurement

The hybrid IVR structure with SCVR and D-LDO converter has been designed and implemented using 22nm Intel tri-gate technology. The 3.8mm<sup>2</sup> test-chip includes the

execution core for performing key operations for a graphics core (34). Fast energy-efficient switching characteristics and low voltage headroom of 1.05V of this advanced process node makes digitally-assisted analog solutions very attractive. Two clocked comparators have been used here for hysteretic feedback. A 0.2-to-9GHz programmable ring oscillator generates the comparator sampling frequency, while reference voltage ( $V_{REF}$ ) is externally supplied (34). The PMOS power header switches (PH) are sized to deliver the maximum current needed at the highest rated  $V_{OUT}$ . Since SCVR switches are reutilized in the LDO mode, same-strength switches are used. To reduce IR drop and for effective floorplan, the entire power switch hierarchy is split equally at the top and bottom of the core, spanning with width of the chip. The top and bottom D-LDO halves use identical controllers and can be programmed to run either in-phase or 180° out-of-phase. By running them out-of-phase, interleaved control similar to SCVR can be achieved, leading to improved transient response. Figure 3.20 shows the overall implementation of D-LDO, with the entire hybrid IVR structure having less than 4% area overhead.

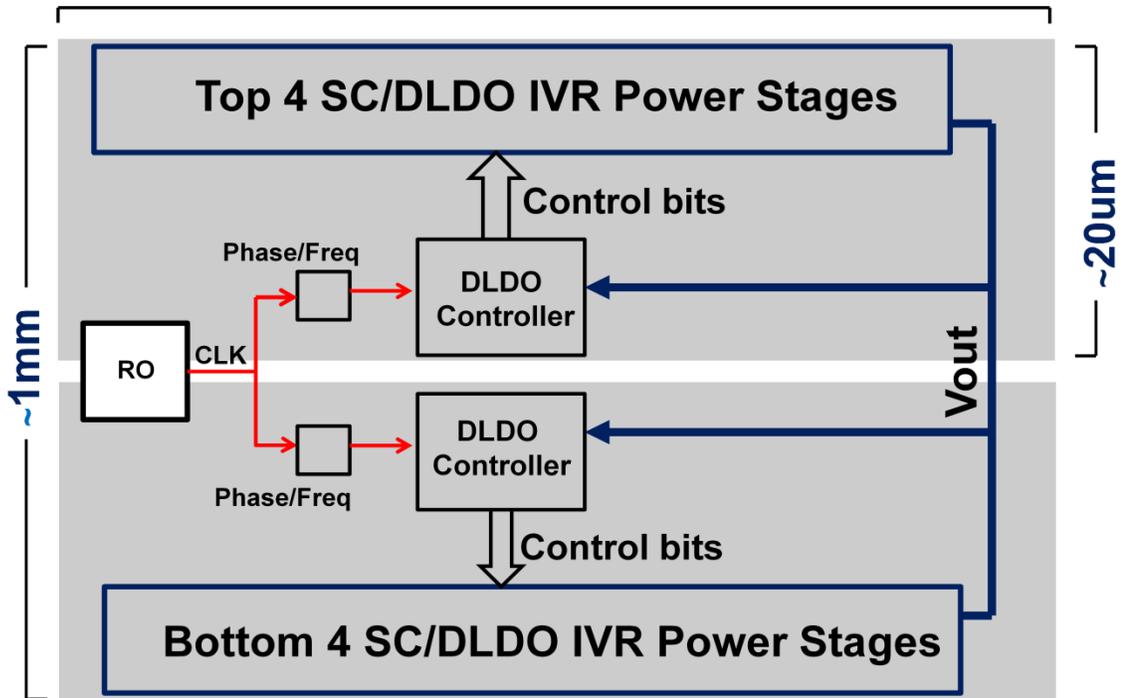
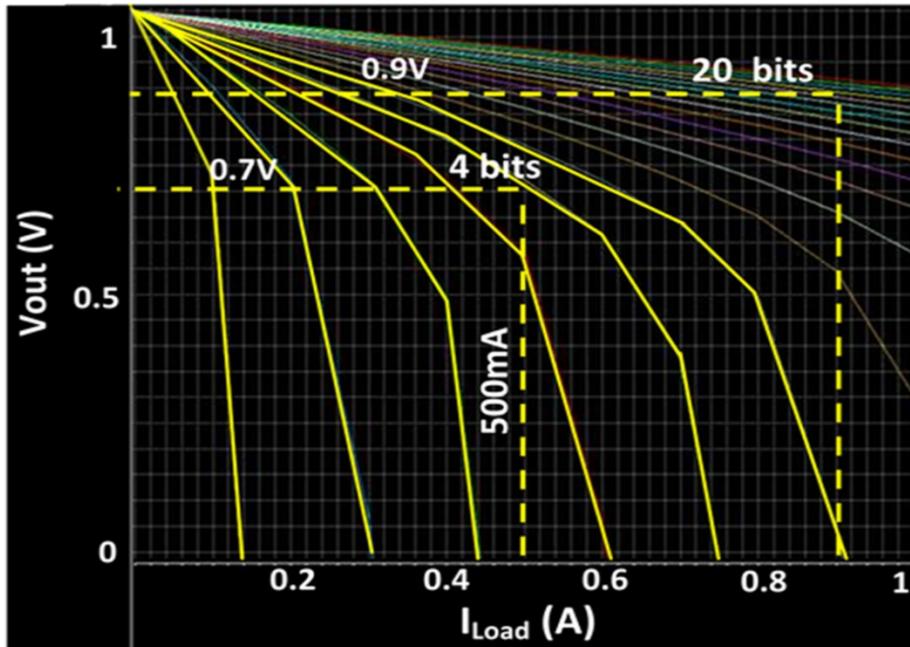


Figure 3.20. Hybrid IVR floorplan with less than 4% area overhead

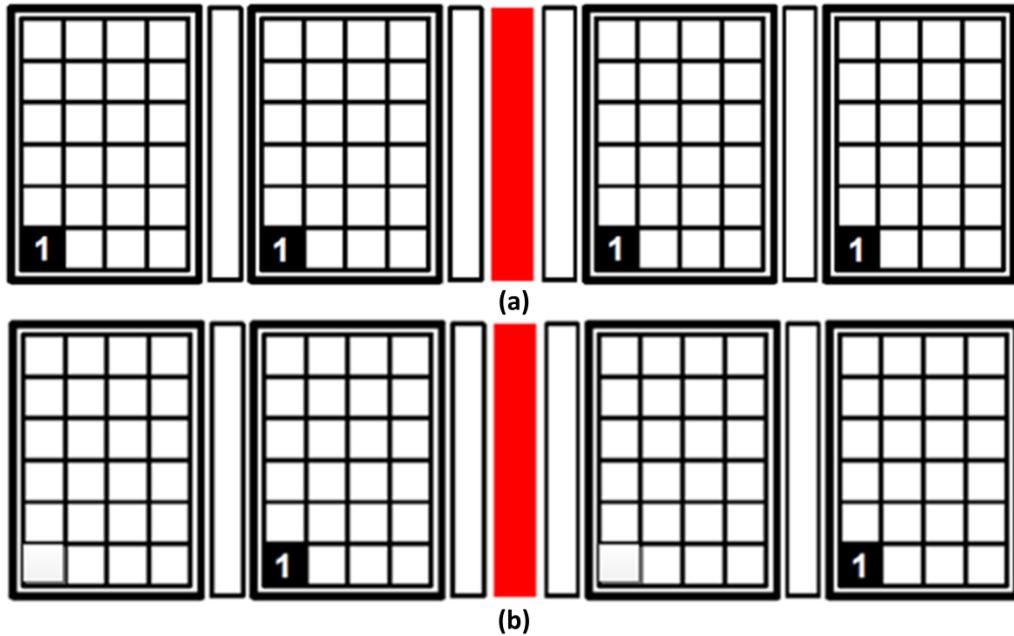
### 3.3.11.1. Adaptive Bit-Gain Modulation

Ripple voltage is a function of bit-granularity; the larger the number of bits, the smaller the incremental change in current. When  $V_{OUT}$  is low i.e. high drain-source voltage ( $V_{DS}$ ) for the MOSFETs, current/bit is significantly higher as compared to when  $V_{OUT}$  is high. For example, only 4 bits are needed to provide 500mA at 0.7V, while 20 bits can deliver 900mA as  $V_{DS}$  is swept for the switches in Figure 3.21. More current flow per bit can cause steady state ripple voltage due to the coarse-grained switching. Therefore, based on the  $V_{OUT}$ -lookup table, bit-gain needs to be optimized for different  $V_{OUT}$ . For equal current distribution from the header switches, in order to prevent stress effect, 4 switches are turned on/off at the same time, i.e. 1 bit is equal to the strength of 4 switches as shown in Figure 3.22 of power switch floorplan. Apart from a similar current gradient,

this floorplan allows effective bit gain modulation. While Figure 3.22(a) represents a bit gain of 1 (4 switches/bit) for higher  $V_{OUT}$ , Figure 3.22(b) shows a bit gain of 0.5 (2 switches/bit) for lower  $V_{OUT}$ . This way of bit-gain modulation using known system power information helps reduce peak-to-peak ripple voltage, as seen from Table 3.3.



**Figure 3.21. Varying bit strength for different  $V_{OUT}$ . Higher  $V_{DS}$ , higher current per transistor**



**Figure 3.22 (a) Top half of IVR. 4 switches (enabled from each sub-module) comprise one bit for bit-gain of 1 (34) (b) 2 switches comprise one bit for bit-gain of 0.5**

**Table 3.3 Peak to peak ripple voltage with and without adaptive bit-gain modulation**

$I_{Load}$ (mA)	P-P Ripple (mV)	
	Gain = 1	Gain = 0.5
5	15	12.5
50	15	13
100	17	15
200	30	25
300	32	20
500	35	7

### 3.3.11.2. Transient Performance: Simulation and Measurement results

While this dissertation focuses mostly on the design aspect of the D-LDO, some of the silicon validations performed at Intel lab are discussed below for better understanding.

From a 1.05V  $V_{IN}$ , the D-LDO achieves maximum power efficiency of 84%-56% for 0.92V-0.64V  $V_{OUT}$  (~3% worse than ideal LDO), while the SCVR extends the voltage range to 0.63V-0.38V with a sufficiently high efficiency of 73%-52%, improving over an ideal LDO by up to 44%. The D-LDO is fully functional till  $V_{IN}$  of 0.65V. At 1.05V the proposed D-LDO exhibits 130mV dropout voltage, improving power efficiency. As earlier mentioned the CM of D-LDO is triggered when  $V_{OUT}$  droops by 20mV and 50% droop reduction is attained, as exemplified in Figure 3.23. Reducing the droop guard-band allows to improve the core frequency by ~75% at maximum  $V_{OUT}$  versus if no fast mitigation is used.

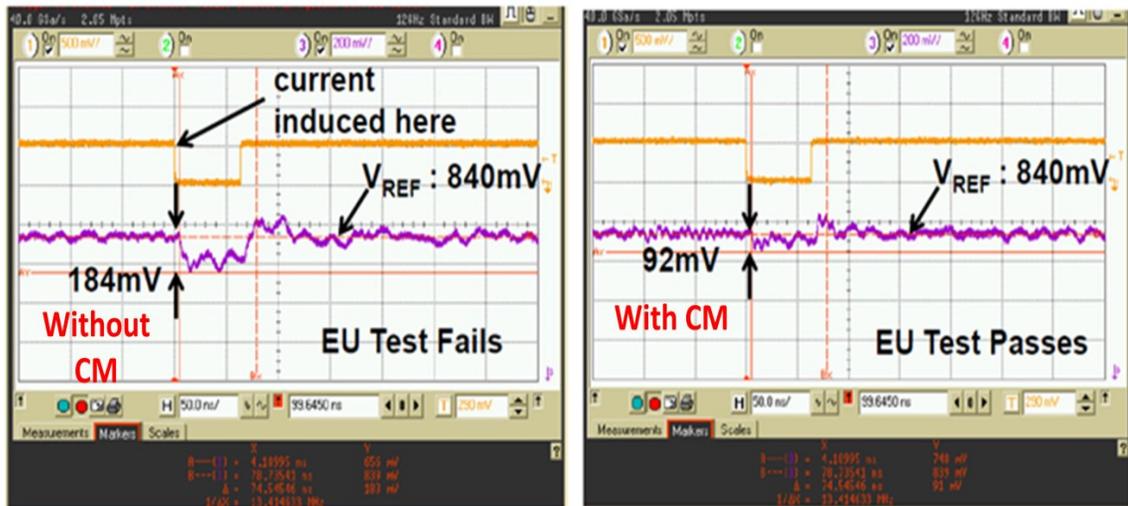


Figure 3.23. Measured voltage droop mitigation: with and without coarse-grained mode (34)

### 3.3.12. State-of-the-art comparisons

The idea of the proposed D-LDO here is to integrate seamlessly with an SCVR converter to provide a truly hybrid IVR with high efficiency across a wide range of  $V_{OUT}$ . However, since the author's contribution lies mostly in the design of the digitally-controlled LDO, comparative analysis has been performed only among the state-of-the-art LDOs, as shown in Table 3.4. Since data for measured quiescent current and worst-case output droop are not available, simulation results (including RC parameters extracted from chip layout to account for interconnect parasitic) have been used to perform the comparisons. The proposed D-LDO achieves superior transient response (transient response or  $FOM_1$ ) compared to any other D-LDO, while  $FOM_2$  of 0.82ps calculated from simulation betters any other LDO in literature. However, A-LDO (20) still achieves better transient response (1.37X faster) due to its linear regulation. One major reason for lower response speed in the proposed D-LDO is the fact that identically sized switches have been used. However, this restriction is due more to SCVR design restrictions. Using non-linearly sized switches with thermometric counter (as explained in section 3.3.3) can provide a linear modulation of conductance, improving ripple voltage and response time.

**Table 3.4 Performance comparisons with state-of-the-art LDO (both analog and digital)**

Ref	Process (nm)	LDO Type	V <sub>in</sub> (V)	V <sub>out</sub> (V)	I <sub>Load</sub> (mA)	I <sub>q</sub> (uA)	V <sub>Drop</sub> (mV)	FOM <sub>1</sub> (ns)	FOM <sub>2</sub> (ps)
(20)	90	ALDO	1.2	0.9	100	6000	90	0.54	32.4
(57)	65	DLDO	0.5	0.45	0.2	2.7	40	20000	270000
(22)	65	DLDO	1.1	1	100	128	80	1.2	1.53
(59)	40	DLDO	1.34	1.2	250	10000	50	0.114	4.56
(36)	90	DLDO	0.5	0.44	3	30.8	70	2333	23955
(11)	180	DLDO	0.9	0.8	200	750	70	350	1312.5
This Work	22	DLDO	0.65-1.05	0.38-0.95	1000	1100 (Sim)	45 (Sim)	0.74	0.82

### 3.4. Summary

An increasing number of power domains and power states per domain and a wide dynamic range of operation for digital load circuits necessitate the design of high-efficiency, compact on-die voltage regulators providing ultra-fine-grained spatio-temporal voltage distribution. Digitally implementable linear regulators operated in low-dropout (LDO) mode, exhibit process and voltage scalability, thus supplementing their

analog counterparts. In this chapter, two digitally-controlled architectures, catering to a wide range of applications, have been discussed. A truly hybrid IVR with SCVR/LDO modes has been designed in 22nm Intel tri-gate technology to provide wide-range DVFS (0.38V to 0.93V) capability. D-LDO mode achieves highest efficiency of 84%, while dual-loop strategy allows 50% droop mitigation. A system power-modes-aware adaptive control for D-LDO has been proposed to scale quiescent current consumption and reduce standby-power-drain in energy-harvesting architecture. Silicon measurements show current efficiency of more than 90% with standby-quiescent current of 500nA at  $V_{IN}$  of 1.75V, achieving a FOM of 4.44ps. To conclude, a digitally-adaptive design environment and the ability to work efficiently for wide  $V_{IN}$  range makes D-LDO an attractive alternative to the analog counterparts, especially for digital load.

## Chapter 4

---

### 4.1. Charge-Recycled Power Regulation with Integrated Switched-Capacitor

With battery technologies struggling to keep pace with increasing demands for lower core-voltage, small form-factor switching converters with integrated passives are being investigated to bridge the voltage gap. IVRs enable fast voltage transition with multiple on-chip voltage domains and allow aggressive power-scavenging. However, in trying to provide 100% power to the loads, these converters are limited by inherent parasitic losses which are not scaling as aggressively as the technology nodes. Therefore, alternative solutions for highly dense integrated power converters are needed to meet the future demands of power-hungry SoC designs.

Charge-recycling architecture, also known as *Voltage Stacking* (V-S), where part of the power delivered comes from load itself, is being actively explored as a viable alternative (18, 51, 63, 64) to integrated regulation. This chapter focuses on differential SC converter assisted V-S for superior integrated power regulation performance. A push-pull SC converter is designed and optimized in commercial CMOS technologies with the

help of the Matlab model. Design attributes that are unique to V-S (open-loop versus closed-loop regulation, differential regulation, efficiency characterization with workload imbalance, etc) are discussed in-depth in this chapter. A practical demonstration of V-S using commercial-off-the-shelf FPGA chips and fabricated dual-output SC converter validates the superior performance claims of V-S.

#### 4.1.1. Background on Voltage Stacking – Benefits and challenges

Voltage stacking refers to the power delivery arrangement with raised supply voltage ( $N$  times) and stacked loads ( $N$  loads) recycling supply current ( $1/N$  times) with implicit down-conversion of voltage (77). Figure 4.1 presents a conceptual block diagram of voltage stacking. While the load power consumption in this series-stacked architecture remains the same, high voltage instead of high current on-chip helps address some of the key issues in conventional power delivery. With reduction of the supply current, IR drop across the power delivery subsystem and  $I^2R$  loss in the package gets reduced by a factor of  $N^2$  compared to a non-stacked conventional approach (41). Moreover, in principle, implicit voltage conversion allows complete removal of on-chip regulators, therefore removing a big source of area and power overhead and, by down-converting high off-chip voltage using multiple stacked loads, device reliability is also ensured. The fundamentally simple and scalable nature of V-S has spurred wide-spread research activity on this charge-recycling technique. The earliest work proposed V-S as an efficient fine-grained power management technique where logic blocks are stacked with a raised supply voltage (63). As energy-efficiency is becoming a major design constraint,

even industry has started exploring this unconventional architecture in recent times (45, 81).

While voltage stacking has the potential to alleviate the inefficiencies related to power delivery, it introduces additional design challenges that must be addressed for practical implementation. Kirchoff's current law (KCL) dictates equal flow of current in the stacked cores. However, current consumption of cores can fluctuate as a function of workload, various power-saving schemes and data, leading to inter-layer current imbalances. The fact that V-S will try to compensate for any current mismatch between the stacked loads by distributing the intermediate voltage, gives rise to a voltage noise which can disrupt the functionality of this stacked architecture (41). For example, if a core consumes less current compared to others, its voltage headroom increases to compensate at the expense of reducing the voltage across the other cores. In the worst-case scenario, intermediate voltage rail noise can make this stacked PDN collapse. Therefore, balancing current consumption across the layers is a key challenge in V-S.

Most of the recent works on V-S have proposed different techniques for controlling intermediate voltage noise, either through software-controlled scheduling (31), where different threads can be controlled for effective balance, or hardware-controlled scheduling, by adding explicit local on-die linear regulator to flush out excess imbalance-related charges (64), or even through usage of inherently balanced stacked loads (81). However, most of the proposed solutions do not scale efficiently with more than two stacked loads. The push-pull linear regulator, as suggested in (64), has a very small area overhead. However, any imbalance will lead to poor efficiency, for example, if the top core has larger current requirements than the bottom core, then the linear

regulator will force the excess current to ground, wasting power. Hardware-driven scheduling by shuttling load between the stacked domains will have big area and energy overhead and will not scale efficiently with high-power loads (18). Stacking at finer-granularity, as proposed in these earlier works, also suffers from higher level-shifting overhead.

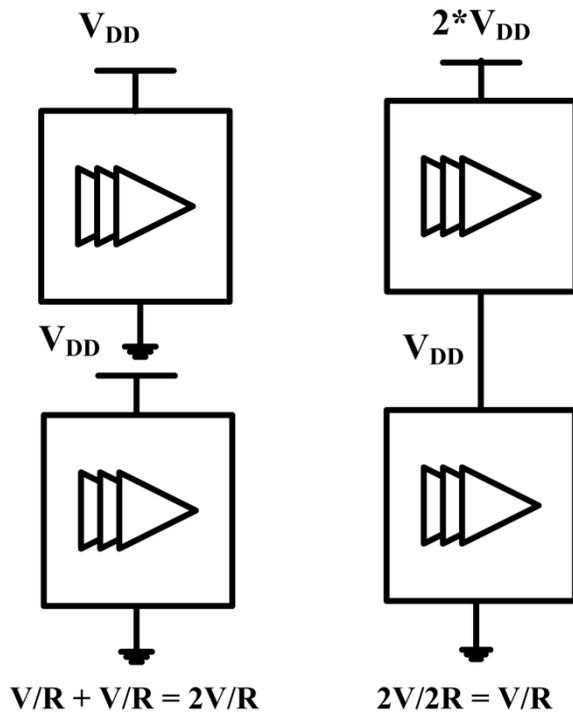


Figure 4.1. Conventional load in parallel (left) versus stacked load in V-S (right)

#### 4.2. Voltage Stacking with Stacked SC converters

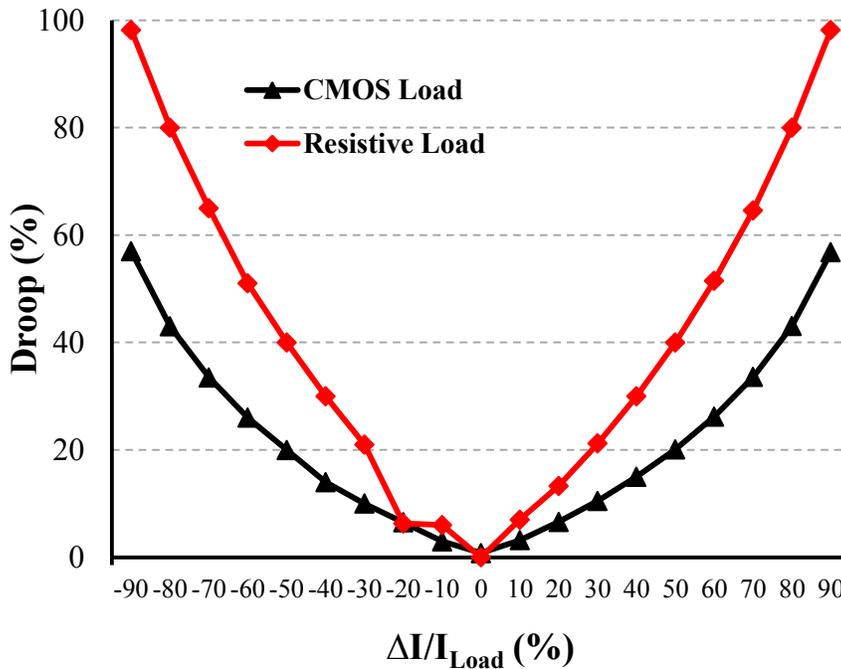
The idea of implicit down-conversion through voltage stacking and its benefit is very intuitive if the loads are imagined as resistors stacked upon each other. However, real load may behave differently from an ideal resistor. Even resistive loads of different

magnitudes will act as an imbalance, leading to internal voltage noise. Effectiveness of charge-recycling will depend not only on the workload differences, but also on the nature of the stacked loads. To demonstrate this load-dependency, two different kinds of load, CMOS load (multiple blocks of ring oscillators) and resistive load, drawing similar amounts of power are stacked. From first order analysis,

$$V_{\text{CMOS}} \propto \sqrt{I_{\text{CMOS}}} \quad 4-1$$

$$V_{\text{RESISTIVE}} \propto I_{\text{RESISTIVE}} \quad 4-2$$

Therefore, load current imbalance will cause a larger variation in resistive load than CMOS load. This quadratic versus linear dependency between voltage and current is also evident from Figure 4.2, where mid-voltage droop due to resistive load variation is larger than that due to CMOS load variation.



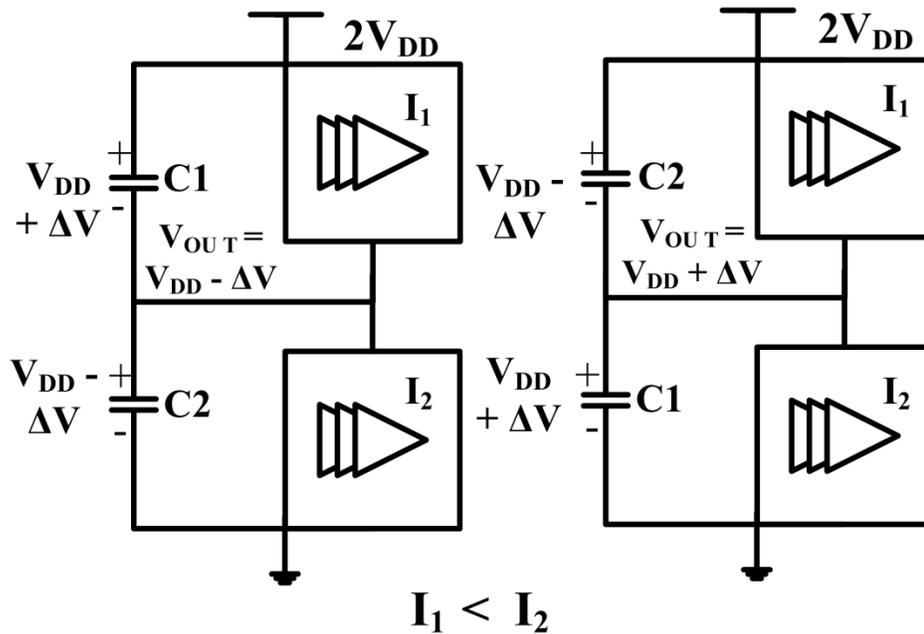
**Figure 4.2. Intermediate voltage noise dependency on nature of load; resistive and capacitive.**

In order to recycle the current imbalance between two voltage-stacked cores, current needs to be either sourced or sunk from the regulator, depending on which core consumes more current. When the bottom core consumes more, current needs to be sourced which is similar to a conventional regulator delivering current to load. However, when the top core consumes more, current needs to be sunk. Unlike conventional regulation schemes, where the regulators provide 100% power to the loads, V-S requires differential converters that only handle the current mismatch between the loads, and thereby converters with smaller passives can attain higher efficiency than conventional regulation. The differential and “push-pull” nature of the required converter strongly suggest use of a switching regulator. Unlike linear regulators, which are resistive in nature, switching regulators with their passives can store energy, thereby effectively recycling charges accumulated at the intermediate nodes and improving system efficiency. Even though passives can occupy significantly larger areas than resistive regulators, recent developments of exotic passive technologies have improved design density of these switching regulators. Therefore, this chapter has primarily focused on a push-pull SC converter (SC), recycling the charge imbalance between the stacked loads (77). Even off-chip bidirectional buck-boost converters have found usage as differential converters for other types of applications such as photovoltaic cells power management (75). However, the inductive converter is beyond the scope of this dissertation. Moreover, the fact that some of these capacitive technologies (MOS, MIM, Fe-cap) are completely integrated with the CMOS process makes the SC converter an attractive option.

## 4.2.1. Switched-capacitor Design and Optimization

### 4.2.1.1. *Push-Pull SC Converter for V-S*

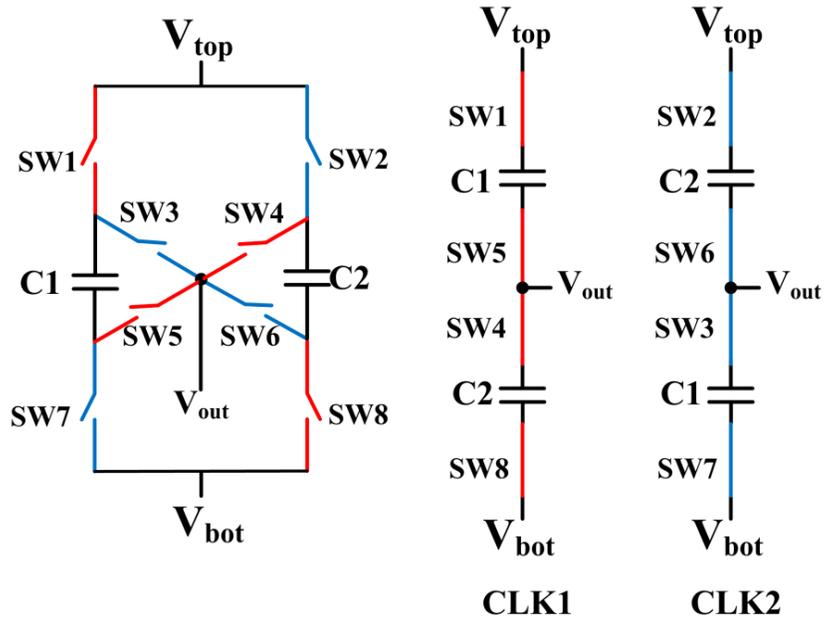
Since V-S implicitly down-converts high voltage to low voltage, SCVR for stacked loads (only two loads considered here for simplicity) will act more like a charge-equalizer, redistributing the imbalance between the loads and assisting the natural voltage division. Therefore, SCVR with 2:1 voltage conversion ratio (VCR) can be used as a point-of-load regulation for voltage stacked nodes. However, the nature of regulation demands a push-pull converter. Figure 4.3 illustrates the fundamental principle of push-pull regulation with switching capacitors (fly-caps). Consider an example of a slightly imbalanced workload with a supply voltage of  $2V_{DD}$ , where the current offset pulls  $V_{OUT}$  to droop  $V_{DD}$  by  $\Delta V$  as  $I_2$  (bottom load) is greater than  $I_1$  (bottom load). In the first phase, C1 begins charging to  $\Delta V$  voltage above  $V_{DD}$ , while the voltage across C2 falls below  $V_{DD}$  by  $\Delta V$ . In the second phase, through on-chip switches, C1 and C2 swap places. Since C1 was charged to a higher voltage, it redirects this charge back onto the  $V_{OUT}$  node. This redirection of charge helps pull the load voltage  $\Delta V$  above  $V_{DD}$ . The ripple voltage ( $2\Delta V$ ) is a manifest of the capacitor charging/discharging, and the faster the switching frequency, the lower the ripple.



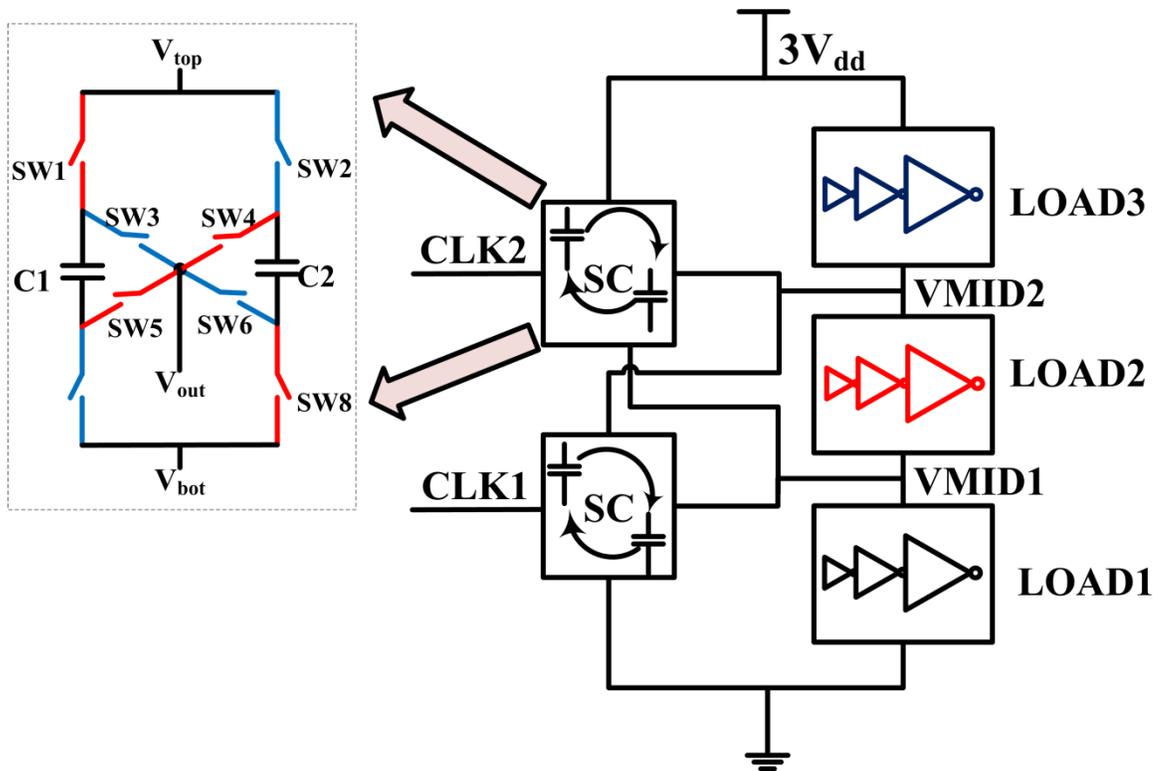
**Figure 4.3. Push-pull nature of regulation needed in V-S using capacitive converters**

This push-pull structure can be achieved by the proposed SC converter circuit shown in Figure 4.4. On the 1<sup>st</sup> phase of clock (CLK1), switches (i.e. SW1, SW5, SW4, SW8) are enabled connecting C1 between  $V_{IN}$  and  $V_{OUT}$  and C2 between  $V_{OUT}$  and ground, while the 2<sup>nd</sup> phase of clock (CLK2) enables switches (SW2, SW6, SW3, SW7), swapping C1 and C2 positions. Since PMOS passes better “1” and NMOS passes better “0”, SW1-SW6 are designed with PMOS, while NMOS are used for SW7-SW8. The 2:1 converter structure can be easily extended into a multi-output SC converter by stacking multiple cells one above the other, as shown in Figure 4.5 (92). Stacking the 2:1 SC cells ensures implicit level-conversion between the capacitive converters working at different voltage domains and allows a heterogeneous clocking scheme, i.e. different stacked nodes can be regulated at different clock frequencies (CLK1 and CLK2 in Figure 4.5). Higher voltage can be generated efficiently using stacked topology without suffering from high voltage

induced stress. However, stacked converters are typically switch-intensive design, requiring a careful sizing strategy for improved efficiency.



**Figure 4.4. Push-pull SC converter for differential V-S regulation**



**Figure 4.5. Stacked loads (three) with stacked SC converters (two), ideally providing  $V_{dd}$  voltage headroom to each load. Zoomed up single cell of 2:1 push-pull SC converter shown on left.**

#### 4.2.2. SC Converter Modeling and Design Optimization

##### 4.2.2.1. *SC Converter output Impedance and Power-loss Modeling*

Since multi-output converters for V-S regulation consist of multiple 2:1 SC converters in series stack, the assumption that optimizing individual cell will give optimum solution for the entire converter is valid here. A similar strategy was adopted earlier for multi-ratio SC converters (43).

M. Seeman had proposed a methodology for analyzing switched-capacitor (SC) DC-DC converter's steady-state performance through evaluation of its output impedance

(72). This resistive impedance is a function of switching frequency and has two asymptotic limits: one where charge transfers among idealized capacitors dominate the impedance, also known as *slow-switching limit* ( $R_{SSL}$ ), and one where resistive paths dominate the impedance, or *fast-switching limit* ( $R_{FSL}$ ). Figure 4.6 illustrates the simple equivalent circuit model for SCVR;  $N$ : indicates the voltage conversion ratio,  $R_{SERIES}$ : output resistance arising from SSL and FSL impedance,  $R_{PAR}$ : shunt losses resulting from switching the parasitic capacitances of the flying capacitors and power switches, and  $R_L$ : load resistance, which is  $V_{OUT}/I_L$  (28). Interestingly, the push-pull SC converter in Figure 4.4 can be rearranged as two 2:1 SC converters from Figure 2.3 running on opposite phases. Therefore, the SC converter characteristic equations derived in (72) for a basic 2:1 converter can be reused for the push-pull converter, given by Equation 4-3 to 4-5,

$$R_{SSL} = \frac{1}{4 \cdot C \cdot F_{SW}} \quad 4-3$$

$$R_{FSL} = 2 \cdot R_{ON} \quad 4-4$$

where  $F_{SW}$  is the switching frequency and  $C$  the fly-cap (half of total fly-capacitor) and the on-resistance of CMOS switch  $R_{ON}$ . As finding the exact output impedance might be nearly impossible with so many variables,  $R_{SERIES}$  is approximated by

$$R_{SERIES} = \sqrt{R_{SSL}^2 + R_{FSL}^2} \quad 4-5$$

With this knowledge of  $R_{SERIES}$ , the intrinsic and parasitic losses of the SC converter can be modeled. Power loss in switched-capacitor can be categorized as

$$P_{LOSS} = \sqrt{P_{SSL}^2 + P_{FSL}^2} + P_{SW} + P_{BP} \quad 4-6$$

Where  $P_{SSL}$  is the SSL impedance loss due to the component of the output impedance related to charge transfer ( $R_{SSL}$ ),  $P_{FSL}$  due to FSL output impedance  $R_{FSL}$  loss (also

known as switch conduction loss due to  $R_{ON}$  related losses in the switches),  $P_{SW}$  given by Equation 4-7 is the loss in switching gate/drain/source parasitic capacitances of the switches (gate-drive loss) and  $P_{BP}$  given by Equation 4-8 due to the bottom-plate parasitic capacitance of the capacitors. This parasitic capacitance is significant for integrated capacitors, and represents the capacitance between the physical bottom plate of a metal capacitor and the substrate, or for MOS capacitors, the junction capacitance between the source and drain and the substrate.

$$P_{SW} = V_{SW}^2 \cdot N \cdot W_{SW} \cdot C_{gate} \cdot F_{SW} \quad 4-7$$

$$P_{BP} = \alpha \cdot C \cdot V_{CAP}^2 \cdot F_{SW} \quad 4-8$$

where  $\alpha$  is a technology dependent parameter,  $V_{CAP}$  denotes the voltage swing of the parasitic capacitor with respect to substrate,  $V_{SW}$ ,  $C_{gate}$ ,  $N$  denote the voltage swing, gate capacitance and the number of switches.

While this IVR model has been used for multi-variable design optimization in this chapter, further validation against circuit simulation has been performed in chapter 5, where this model has been extended to a multi-output SC converter architecture to be included in a system-level PDN model.

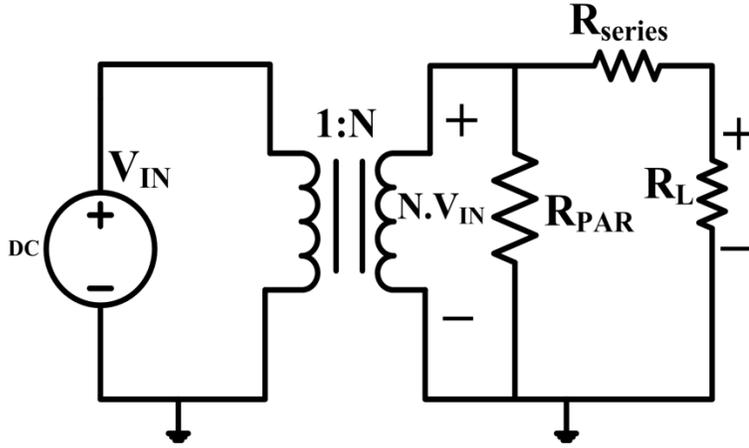


Figure 4.6. Output impedance model of SCVR

#### 4.2.2.2. SC Converter Power Optimization

Output current of a 2:1 SC converter is given by (3)

$$I_{OUT} = 2 * C * (V_{IN} - 2V_{OUT}) * K * F_{SW} \quad 4-9$$

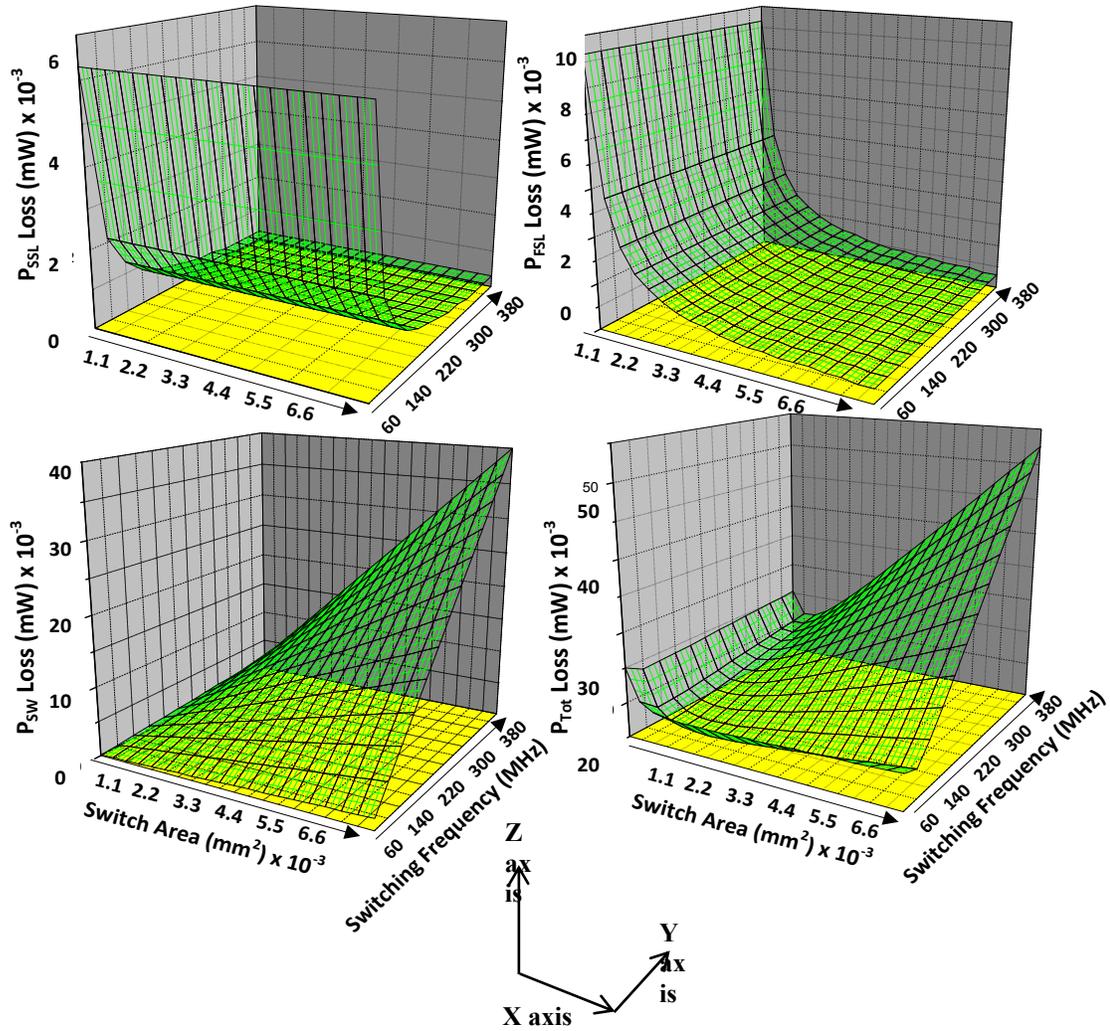
where K is a design variable ranging between 0 and 1 and given by

$$K = \frac{1 - e^{-1/(2 * F_{SW} * C * R_{ON})}}{1 + e^{-1/(2 * F_{SW} * C * R_{ON})}} \quad 4-10$$

Therefore, for a power specification containing  $V_{IN}$ ,  $V_{OUT}$  and  $I_{OUT}$ , the design parameters for optimization are C,  $R_{ON}$ , and  $F_{SW}$ . The bigger the capacitor, the lower the switching frequency needed for delivering a specified load-current, thus reducing losses. However, since integrated converters are constrained by area, higher switching frequency and bigger switch width (to modulate  $R_{ON}$ ) are the primary design knobs to optimize the converter, while the capacitor value is decided based on power density specification.

Based on Equation 4-3 to 4-10 and Figure 4.6, a model of the 2:1 push-pull SC converter has been designed in MATLAB using 45nm CMOS technology (gate

capacitance of  $2\text{fF}/\mu\text{m}^2$ ) parameters for 2V to 1V conversion for maximum 100mW output power. 8nF of ideal capacitor has been used as fly-cap in the modeling (bottom-plate parasitic loss ignored). Switch area and switching frequency have been swept over a wide-range of values to find the optimum range for minimizing the different loss components while restricting maximum voltage drop across IVR to be  $\sim 5\%$  of  $V_{\text{OUT}}$  (fly-cap is fixed at 8nF). Figure 4.7 illustrates three dimensional design-space optimization with switch area (x-axis), switching frequency (y-axis) and power loss (z-axis). Different loss components scale differently with  $F_{\text{SW}}$  and  $W_{\text{SW}}$ .  $P_{\text{SSL}}$  is mostly immune to  $W_{\text{SW}}$ , while  $P_{\text{FSL}}$  is independent of frequency. However, the role of  $P_{\text{SW}}$  makes it very interesting since switching loss scales with both  $F_{\text{SW}}$  and  $W_{\text{SW}}$ . As shown in Figure 4.7, there is an optimum value of the design knobs where the power loss is at its lowest ( $A_{\text{sw}} = 1.2\text{-}1.5\text{ mm}^2$ ,  $F_{\text{sw}} = 50\text{-}60\text{ MHz}$ ). However, all these simulations are done with respect to a fixed load current and, therefore, the optimum values of design variables change as output current changes. To dynamically adjust design variables, closed-loop feedback is used, more of which is discussed in section 4.2.5.



**Figure 4.7. Power-loss optimization in SCVR with  $W_{sw}$  and  $F_{sw}$  as design variables**

By adopting this kind of multi-variable optimization for a given power density requirement, the optimum switching frequency ( $F_{sw}$ ) and switch width ( $W_{sw}$ ) for any SC converter circuit can be determined. Using this modeling approach, the push-pull SC converter circuit has been ported to a commercial 28nm technology with similar ratio of switch sizing and switching frequency, and has been used for all circuit simulations/validation results henceforth.

### 4.2.3. Higher Efficiency in V-S

Unlike conventional regulation, where power efficiency is simply the ratio of power delivered by the converter to power drawn from the source, V-S system efficiency consists of two different components, explicit SC converter efficiency and implicit charge-recycled system efficiency. The higher the recycling, the higher the system efficiency (51). This can be better understood with Figure 4.8 showing a comparison between conventional and stacked loads. Since the stacked converter needs to regulate  $I_1$ - $I_2$  current (which is smaller than  $I_1+I_2$ ) for supporting the same loads as a conventional regulator, the differential nature of the stacked converter allows it to be designed with smaller passives, improving power density of the converter. While efficiency of the SC converter in V-S stays the same (Equation 4-11), stacked system efficiency includes both implicit and explicit regulation (Equation 4-12).

$$\eta_{SC} = \frac{VDD * |I_1 - I_2|}{2 * VDD * |I_{SC}|} \quad 4-11$$

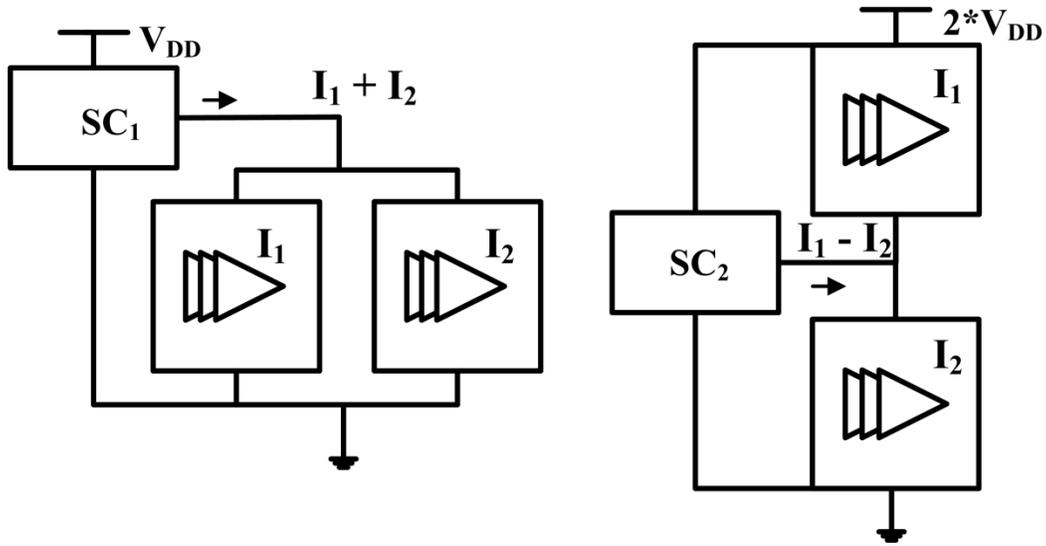
$$\eta_{V-S} = \frac{VDD * |I_1 - I_2| + 2 * VDD * I_1}{2 * VDD * (I_1 + |I_{SC}|)} \quad 4-12$$

Figure 4.8 shows an interesting observation; the worst-case voltage noise (i.e. IR drop across IVR) for a conventional SC converter ( $SC_1$ ) happens when it needs to deliver maximum load ( $I_1+I_2$ ), whereas the worst-case IR drop for V-S happens with one load at maximum power demand with the other at minimum power drive condition. Assuming that  $I_1$  and  $I_2$  are in similar range,  $SC_2$  converter can be made half the size of  $SC_1$  for similar voltage drop. Or the other way around, where the same SC converter is used for both stacked and non-stacked (conventional) loads as seen from the simulation in Figure 4.9 where SC converter can provide maximum of 200mA current to the load. Stacked

architecture can support much higher power as long as the maximum current-imbalance is limited to 200mA between the stacked loads. Here efficiency is plotted against current imbalance, measured as

$$\Delta I_{\text{Load}} = I_1 - I_2 \quad 4-13$$

i.e. positive  $\Delta I_{\text{Load}}$  indicates the top core consumes more current, while negative  $\Delta I_{\text{Load}}$  reflects a higher bottom core power consumption. As the efficiency plot shows (Figure 4.9), (assuming total load current stays the same) best-case efficiency for the SC converter in non-stacked loads (~80%) corresponds to worst-case system efficiency in a stacked loads scenario (where the entire load current of 200mA is provided by SC converter), when there is a 100% mismatch between the loads. However, best-case efficiency of V-S is when loads are implicitly balanced i.e. 0% mismatch, and as seen from Figure 4.9 is close to 95% (assuming 100% off-chip voltage conversion efficiency), leakage and SCVR loss accounting for the remaining 5%. Therefore, in the best and the average cases, V-S gains from higher regulation efficiency and higher power density, while in the worst-case it is almost as good as conventional regulation efficiency (additional switches in the proposed push-pull converter account for some more loss). Moreover, the push-pull SC converter in V-S can be disabled completely during the standby-mode (assuming all loads are in standby-mode with similar leakage profile), drastically reducing quiescent current and allowing close to 100% efficiency through implicit charge-recycling and stacked architecture, reducing battery-power drainage of portable applications like mobile-phones (that tend to stay in the sleep mode for a significant time).

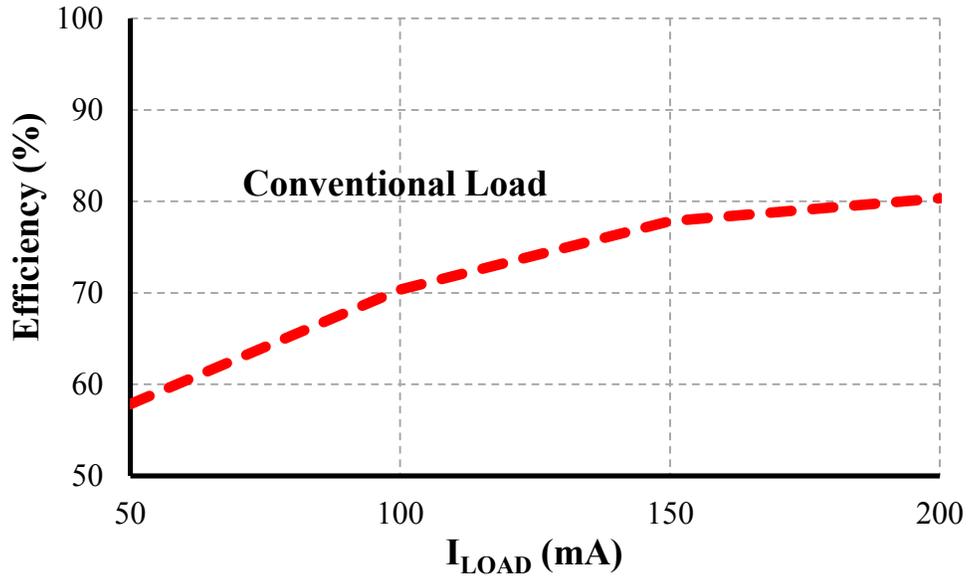


**Figure 4.8. Conventional (parallel) vs V-S (series stacked) loads**

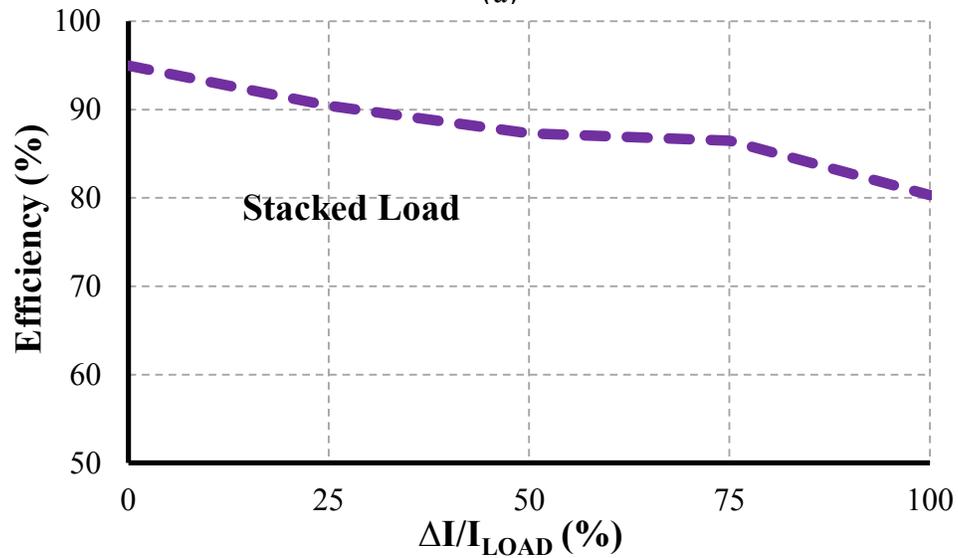
#### 4.2.4. SC Versus Linear Regulator

With its smaller area overhead, a push-pull linear regulator provides a low-cost solution compared to an SC converter for a V-S charge imbalance problem. However, this class of resistive converter will reduce the efficiency benefit of V-S whenever imbalance is high by pushing the excess charges to ground (i.e. top load is greater than bottom load). To gain a better perspective on this design tradeoff, Figure 4.10 shows a comparison plot among stacked loads with SC and linear regulator and non-stacked regulation. Several assumptions have been made to simplify this comparison. The total load power is kept constant at 2W (each load consuming 1W), while the push-pull SC converter (used in both stacked and non-stacked loads) is assumed to have a constant efficiency of 80% for all output power. This assumption may not be practical enough, as seen from Figure 4.9, but since it affects all PDN configurations equally and simplifies the analysis, it has been adopted. The two most important FOM of IVR i.e. power efficiency and power density

are plotted with respect to current imbalance for this comparative analysis. SCVR area overhead is normalized with respect to trench capacitor technology density ( $100\text{fF}/\mu\text{m}^2$ ).



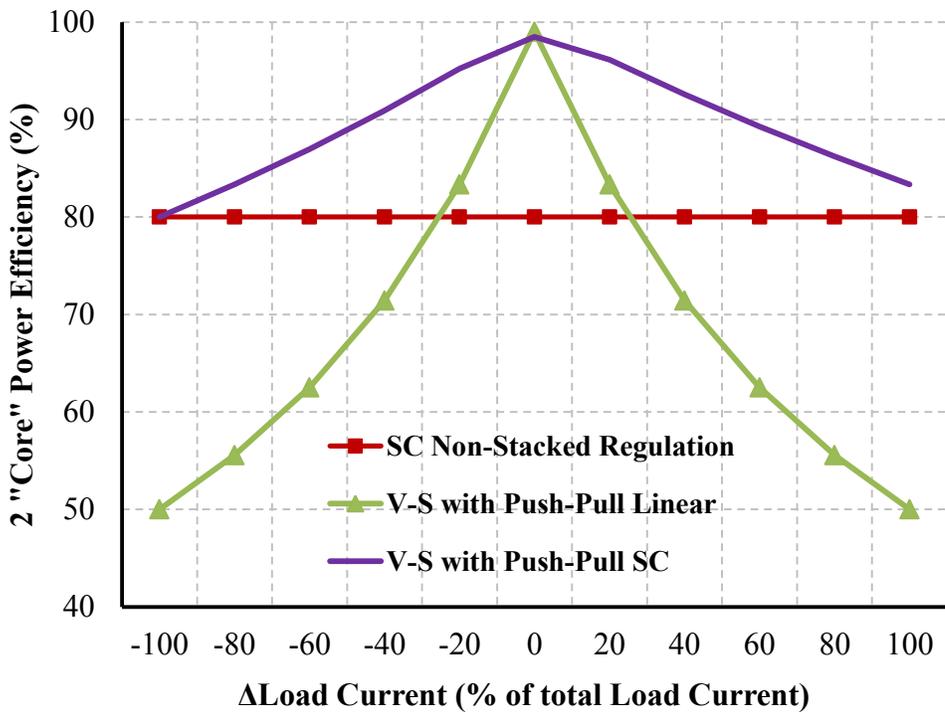
(a)



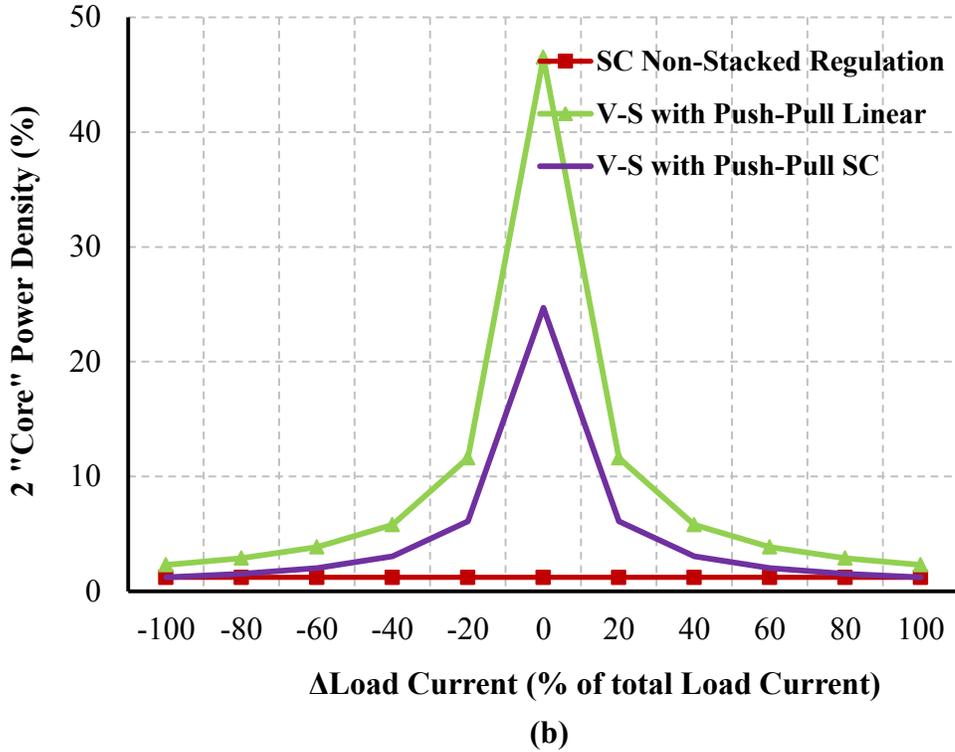
(b)

**Figure 4.9 (a) SC converter efficiency in delivering load current to conventional (non-stacked) loads (b) V-S system efficiency with varying workload between stacked loads. In worst-case, entire load current is provided through SC converter (similar to conventional loads in (a))**

As shown in Figure 2.5, the best power density and power efficiency performance among the SC converters are when a trench capacitor has been used, which makes it a compelling choice for future power regulation. The SC converter design specifications used for this plot are discussed in further detail in Chapter 5 and Table 5.1. As seen in Figure 4.10 (a), by recycling the charge accumulated due to imbalance, the push-pull SC converter can provide power at a higher efficiency than the linear regulator for both stacked and non-stacked loads. However, if the imbalance is within  $\pm 20\%$ , a linear regulator is preferred due to the high power density i.e. low area overhead as seen in 4.10 (b). This opens up a scope for future work where software scheduling can be accompanied with low-cost linear regulation for V-S.



(a)



**Figure 4.10 (a) Power efficiency and power density (b) comparisons between stacked loads with linear and SC converter versus non-stacked loads with SCVR.**

#### 4.2.5. Open-Loop versus Closed-Loop V-S regulation

The output of a SC converter is given by

$$V_{OUT} = nV_{IN} - I_{OUT} * R_{OUT}(F_{SW}, D_i, G_i) \quad 4-14$$

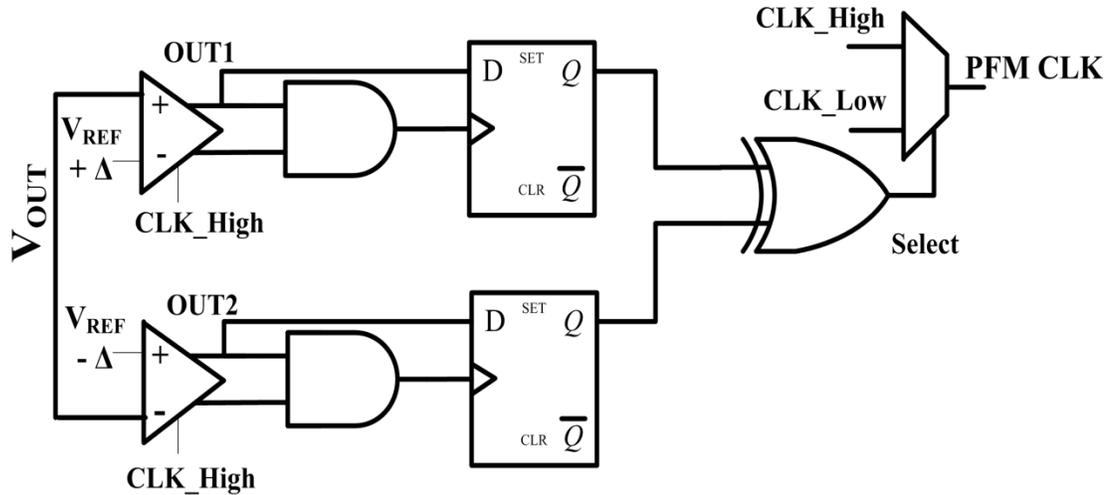
where  $n$  is the voltage conversion ratio (VCR);  $F_{sw}$  the switching frequency;  $D_i$  the duty cycle of switching; and  $G_i$  the conductance of the switches. The numerous control methodologies proposed in the literature modify one or more of these parameters (28, 38, 72). For practical implementation, SCVR can be configured to only a few conversion ratios. In V-S conversion ratio ( $n$ ) is fixed by the number of loads stacked. One of the main drawbacks of varying  $D_i$  or  $G_i$  of the switches at constant frequency is reduced

efficiency at lighter load. However, by keeping the duty cycle fixed at 50% and by modulating switching frequency with load current, higher efficiency can be achieved, especially at lighter load. Hybrid regulation, with two or three control variables together, will bring out the highest efficiency at increased design complexity. Depending on the application, a designer may choose to modulate more than one control variable to obtain optimum regulation.

For the proposed work, switching frequency has been used as the primary controlling variable. A traditional control method for a SC converter may include a linear feedback loop to regulate the switching frequency. However, obtaining stability and good transient response over varying load conditions with such a control technique is difficult. A nonlinear control can provide superior results; therefore, a hysteretic feedback scheme with lower and upper bounds has been used for regulation (7). However, the proposed feedback shown in the Figure 4.11 scheme is different from traditional hysteretic control due to the “push-pull” nature of the stacked loads i.e. outputs can either droop or overshoot. Therefore, switching frequency (CLK\_High) is needed whenever either of the boundaries is crossed by output voltage, while low frequency (CLK\_Low) can regulate the in-between state. Table 4.1 explains the states.

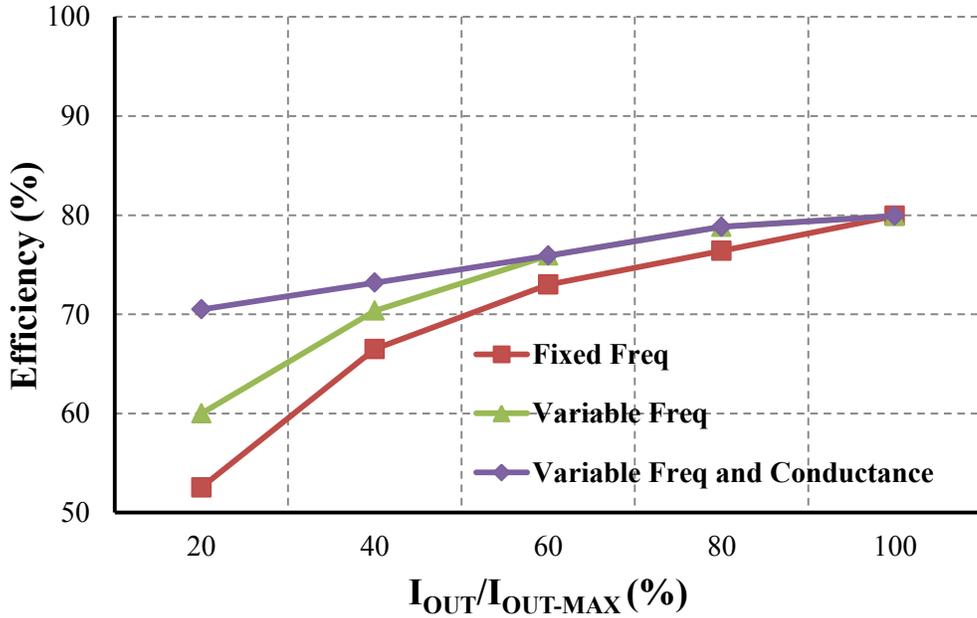
**Table 4.1 Logic outputs modulating  $F_{SW}$  with  $V_{OUT}$  and  $V_{REF}$**

State of O/P	OUT1	OUT2	Select	O/P Clock
$V_{out} > V_{ref} + \Delta$	Toggle	Low	1	CLK_High
$V_{ref} - \Delta < V_{out} < V_{ref} + \Delta$	Low	Low	0	CLK_Low
$V_{out} < V_{ref} - \Delta$	Low	Toggle	1	CLK_High



**Figure 4.11. Dual-boundary hysteretic PFM regulation scheme for V-S**

The feedback circuit consists of two comparators (along with latches, XOR gate and mux) that detect when the output voltage crosses the control boundary. By adding an edge-triggered latch, only rising edges are associated with a charge transfer (7). If there were no latches, then the falling edge, which appears because of the clocked comparator and not as a trigger for boundary crossing, will cause an unwanted charge transfer. In order to account for comparator response time, a clock is generated to the latches out of the comparator itself, implicitly tracking the delay across all PVT corners. Depending on the “Select” signal that triggers the output mux, low or high clock frequency is applied to the switches. For practical implementation, a counter-based approach with a wider range of clock frequencies needs to be used for aggressive power savings. Figure 4.12 shows the impact of feedback regulation on SC converter power efficiency performance for conventional loads (non-stacked). Overheads for both frequency modulation and switch-conductance modulation (not shown in the Figure 4.11) have been added to the efficiency plot. Current efficiency can be further improved by



**Figure 4.12. Closed-loop versus open-loop regulation for conventional SCVR operation**

using additional techniques (different  $V_t$  switches, low parasitic capacitor, shutting off some of the interleaved structures when load requirement is low) as proposed in literature. The downside of frequency-controlled regulation is the high output voltage ripple (38). An SC converter in the slow switching limit (SSL) is capacitive limited, i.e. the converter current consists of charge pulses that are transferred to the output causing the output voltage to rise. Output load current then causes  $V_{OUT}$  to fall down with an exponential RC ( $R_{ON}$  of switches and fly-cap) slope. This peak-to-peak voltage swing is known as ripple and is given by

$$V_{Ripple} = \frac{I_{Load}}{2 * C * F_{sw}} \quad 4-15$$

While scaling switching frequency with load current should ideally keep ripple voltage constant, practically it is not possible. As the switching frequency enters the SSL region of operation, the impulsive nature of charge transfer causes significant voltage noise. The

simplest way to reduce ripple voltage is interleaving, where the entire capacitive converter is divided into groups and clocked at different phases (26). This way, while over a clock cycle the total capacitor to be charged stays the same, through partial charging voltage magnitude is reduced. Another strategy that has been successfully used is increasing the switching frequency and managing the number of interleaved phases at low load current. In general, for conventional regulation, the SC converter is designed for the highest efficiency at the higher load current range and moderate efficiency and acceptable ripple voltage in the lower  $I_{out}$  range.

However, V-S regulation gives a unique opportunity to achieve both low ripple voltage and high efficiency. This can be explained from Equation 4-12 describing V-S system efficiency. While the stacked SC converter efficiency degrades in regulating low load current (in this case low  $|I_1-I_2|$ ), that invariably means that the loads are implicitly balanced through charge-recycling, masking the poor efficiency of the explicit converter and resulting in much higher stacked system efficiency. On the other hand, when the loads are running completely out of sync, more and more current needs to be regulated at the intermediate node, i.e. the SC converter load current will be high. In the absence of charge-recycling, stacked system efficiency will be dependent on the SC converter efficiency, which is designed to be higher at high load current range. In this way, V-S with stacked SC converter can achieve high efficiency for wider range of output load, providing an opportunity to run the regulators open-loop. This eases the design complexity, as feedback loop modulating switching frequency/switch-conductance can be removed, and improves stability. Figure 4.13 targeted stacked loads (0.5W as total load power) and open-loop SC converters regulating the intermediate node. Closed-loop

overheads have been added to compare against the open-loop counterpart of Figure 4.12. As expected, when SCVR delivers large current (i.e. less of charge-recycling), open-loop and closed-loop performances are similar. When SCVR delivers lower load current, however, open-loop V-S system efficiency is more than 85%, whereas the SC converter efficiency for similar load condition is closer to 50% (Figure 4.13 and Figure 4.12). However, adding a closed-loop can scale SC converter losses aggressively and boost efficiency to more than 90% at balanced-load scenario. Another interesting choice that V-S gives to the designer is to use feedforward regulation instead of feedback. Known-system-power information like different P-states, C-states and sleep mode can be used to modulate the converter drivability in conjunction with V-S and remove the need for closed-loop feedback. For faster response to output voltage droops and overshoot, voltage monitoring circuits can be distributed throughout the chip (16). Another benefit of open-loop V-S regulation will be the reduced ripple voltage since converters are now running at constant high switching frequency. Figure 4.14 exemplifies this feature, where maximum improvement of 78% ripple reduction is achieved compared to closed-loop regulation. However, in V-S, the noise on intermediate rail depends not just on the converter ripple voltage but also on the instantaneous load imbalance. While the latter contribution needs to be evaluated and quantified using statistical metrics, which is beyond the scope of this thesis, chapter 5 has included an in-depth exploration of noise in voltage stacked layers. To conclude, open-loop regulation in V-S allows the designer to replace a fine-grained feedback loop with coarse-grained voltage monitoring circuits and feedforward regulation using system power-modes information, thus simplifying SC converter design and improving stability performance.

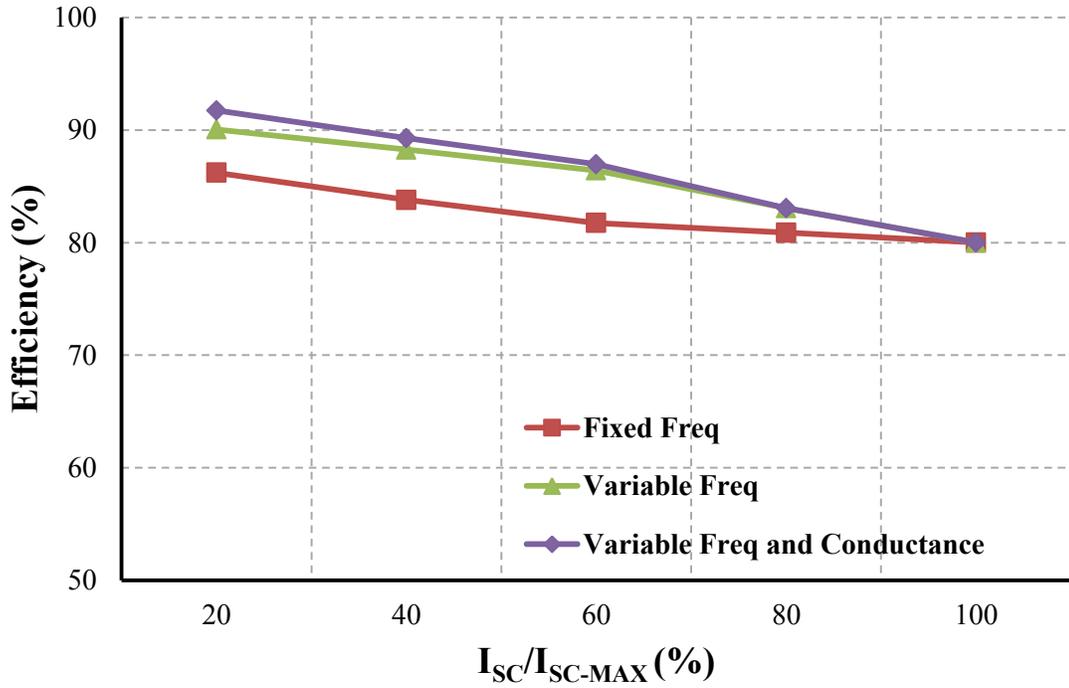


Figure 4.13. V-S system efficiency with open and closed-loop regulation

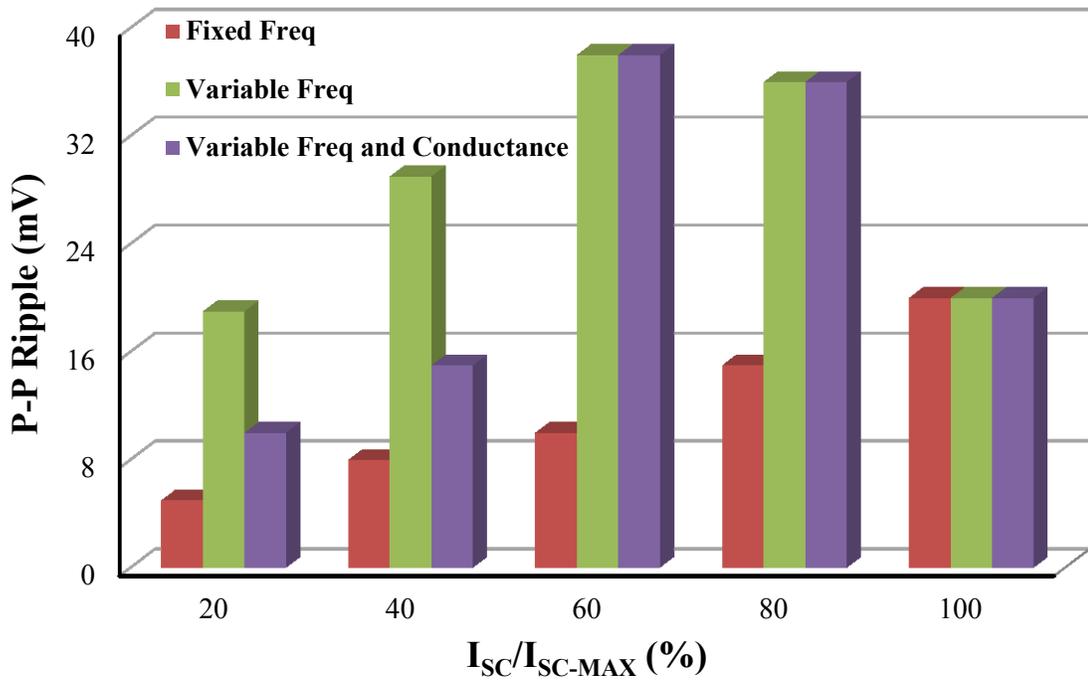
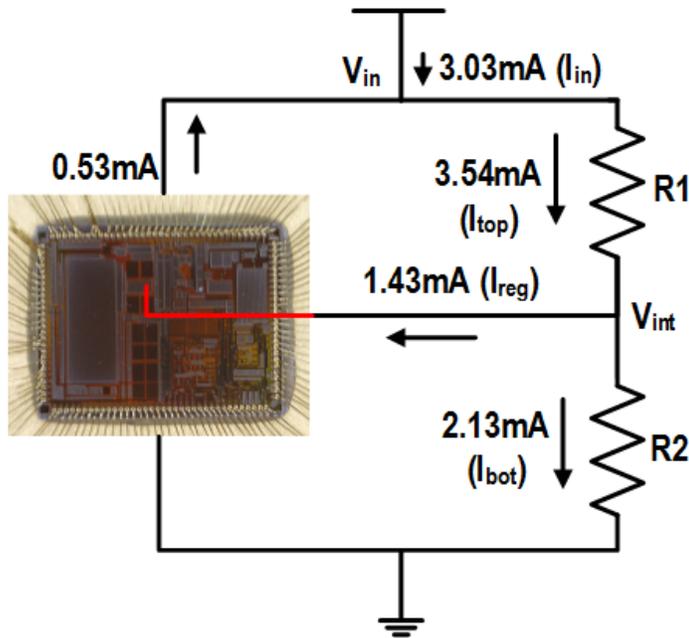


Figure 4.14. P-P ripple reduction with open-loop regulation in V-S

### 4.3. Voltage stacking Measurements



**Figure 4.15. Proof-of-concept V-S experiment showing the ability of the push-pull SC converter to recycle the current imbalance (“sink” additional charges) between stacked layers.  $R_2 > R_1$  ( $R_2 \sim 2.5R_1$ ). The SC converter, fabricated in IBM130nm technology, is shown on left.**

A prototype test setup has been designed using commercial off-the-shelf (COTS) FPGA chips and passive resistors as the vertically stacked loads, respectively. Look-up-tables (LUTs) in the FPGA are configured into multiple rows of ring oscillators (RO) with different control signals, emulating diverse activity among the stacked loads. The dual-output stacked push-pull SC converter (Figure 4.15) has been designed and fabricated in IBM130nm CMOS technology. Simulation of the standalone SC converter shows power efficiency of 77% while delivering output power of 3mW for  $V_{IN}$  of 1V and  $V_{OUT}$  of 0.3V and 0.6V with an area-overhead of  $0.6\text{mm}^2$ . However, measurement results from

the fabricated SC converter shows  $\sim 2\text{mW}$  being delivered to the load at a peak efficiency of 77%, but average efficiency drops down to 72-74%. This degraded chip performance can be speculated to result from a number of causes; undersized switches, limited chip area prompting use of low-efficiency MOS capacitors (instead of highly efficient MIM capacitors) and fabrication constraints forcing the fly-capacitors to be broken into smaller chunks and distributed throughout the chip, thereby adding a lot of RC losses.

Figure 4.16 shows the entire test setup for the V-S power regulation measurement. Bottom-left shows the zoomed-up fabricated die photo of the SC converter, while the stacked connections of the FPGAs are shown on right. In this setup the FPGA boards (and not just the chips) are stacked, for ease of measurement, to emulate the vertically-connected loads characteristics. The different colored wires (violet, red, green and black), as mentioned in Figure 4.16, are used for the different power/grounds connections of the stacked boards. The intermediate stacked voltages, measured using oscilloscopes, have average DC values of 2.27V and 1.15V, for a stacked supply of 3.3V (i.e. the bottommost board gets 1.15V, middlemost board 1.12V and the topmost board has 1.03V of headroom). Figure 4.18 shows the intermediate voltage for two stacked boards ( $V_{DD} = 1.8\text{V}$ ) with different load activity, with and without explicit SCVR. Push-pull mechanism of SCVR pulls up the node voltage from 0.628V to 0.914V (ideally 0.9V for  $V_{DD}$  of 1.8V). Similarly, Figure 4.17 demonstrates a 50%-60% current reduction using two stacked FPGA chips as compared to parallel FPGAs running similar load.

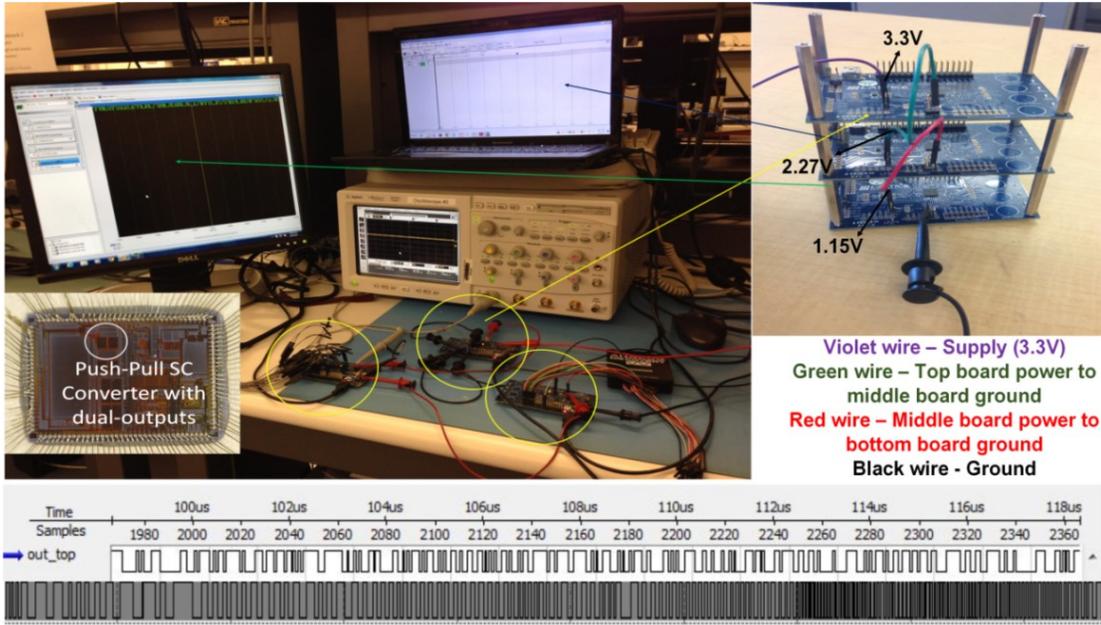


Figure 4.16. Test setup for V-S using vertically connected FPGA boards and fabricated SC converter (left-bottom) with oscilloscopes measuring the stacked voltages and logic analyzer showing the outputs of the top and the bottom boards (bottommost). The connections of the vertically stacked FPGA boards are explained with colored wires (rightmost).

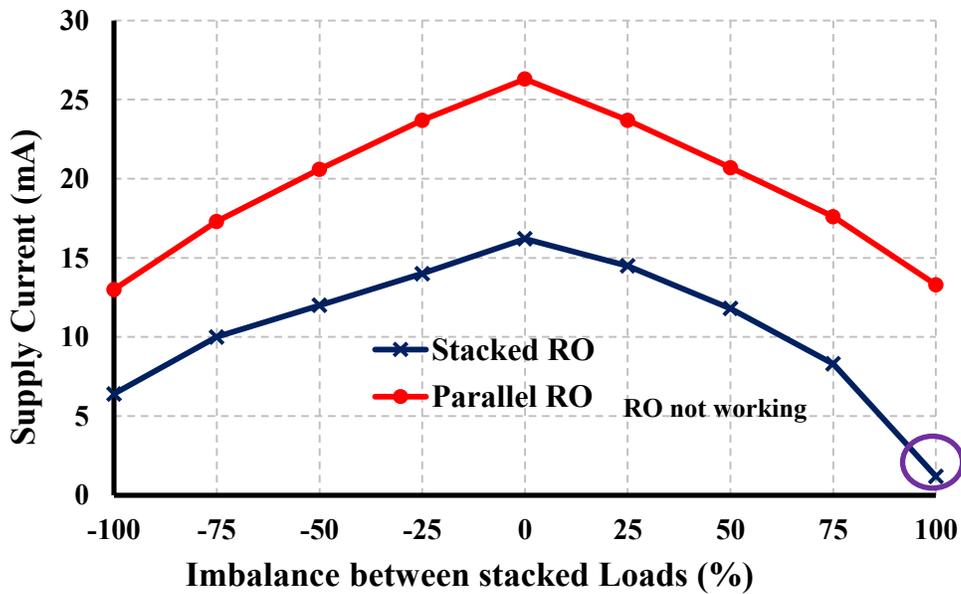
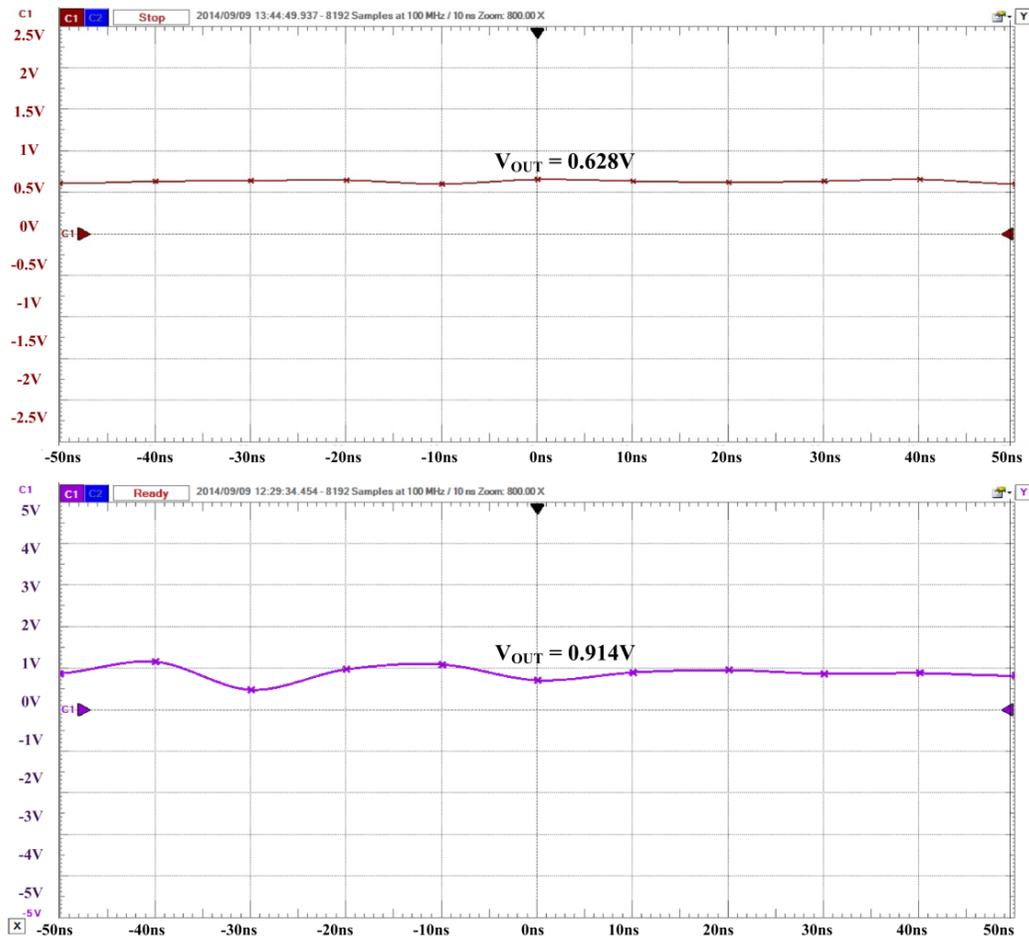


Figure 4.17. Supply current reduction in stacked chips (compared to parallel chips)



**Figure 4.18. Measured results showing stacking without SC converter (top) and with SC converter regulating V-S (bottom) with load imbalance.  $V_{IN} = 1.8V$ , Ideal  $V_{OUT} = 0.9V$  for two loads stacked.**

Now that the working of a V-S with multiple loads assisted by stacked multi-output SC converters is demonstrated, its performance will be evaluated using CMOS/resistive loads. Whenever there is a workload-induced activity difference between the stacks, intuitively it seems that absolute value of imbalance will determine the intermediate rail noise. However, as shown in Equation 4-12, the imbalance-induced voltage noise depends on the ratio of current mismatch to total load current, rather than on the absolute value of mismatched current. This is due to the fact that, unlike conventional circuits

which have low-impedance path to the power/ground, V-S can have a high impedance path to one or both rails, depending on the load activity. Measured values in Table 4.2 show how aggregate load power consumption (with and without mismatch) can make the internal voltage node relatively immune to current mismatches, which is consistent with the claim that V-S fits well in a many-core topology. However, in the measurement setup, the boards being asymmetric (unequal leakage and active currents for same load),  $V_{MID}$  voltage is always offset from the ideal  $V_{MID}$  (1V here).

**Table 4.2 Dependence of V-S noise on ratio of current imbalance to aggregate current**

No of RO enabled (Top)	No of RO enabled (Bottom)	Absolute Imbalance (No of RO)	Imbalance (% of Top Load)	VMID Noise (Implicit) (%)	VMID Noise with (SC) (%)
1	0	1	100	60	30
2	1	1	50	55	23
3	2	1	33	43	20
4	3	1	25	30	16
2	2	0	0	40	16
4	4	0	0	28	15
6	6	0	0	23	13
8	8	0	0	19	8

In evaluating the regulation efficiency of the voltage stacked system, impacts of both implicit charge-recycling and explicit SC conversions are included (Equation 4-12). V-S efficiency is measured for three stacked resistive loads (R1, R2 and R2), with the dual-output SC converter supporting the intermediate nodes, as shown in Figure 4.19. For ease of measurement, the same resistive load (R2) is used for bottom and middle load in series stack. Figure 4.20 shows the power efficiency and internal voltage noise for  $V_{MID2}$  (stacked voltage node between the top and the middle tier) of V-S architecture, measured with and without the SC converter (SCVR disabled). The power efficiency of a voltage-stacked system ranges from 77% to 98%, depending on the current mismatch, which is higher than any state-of-the-art DC-DC converters, especially considering the fact that the fabricated SC converter gives a relatively lower average efficiency of 72-74%. As mentioned in section II, for completely balanced loads during low current activity (standby-mode), the capacitive converter can be disabled or run at extremely low frequency to extract further power savings (accounting for the 96%-98% efficiency range in Figure 4.20). However, all these efficiency claims always assume a perfectly lossless off-chip power conversion (i.e. battery to usable lower voltage, here 1V), which is highly optimistic. Another important point to be noted from Figure 4.20 is that the noise performance improvement of the SC converter-assisted V-S over implicit stacking is less significant than predicted, due to the relatively poor power delivery capability of the fabricated SCVR limiting the maximum current mismatch between the stacked loads. While the standalone SCVR can regulate only up to 2mW of power, the voltage-stacked SC circuit can deliver up to 16mW of power to the loads for the same supply of 1V, improving power density by 8X through differential regulation.

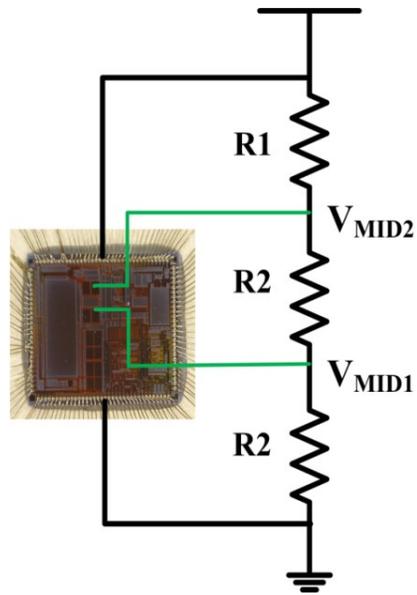


Figure 4.19. Three resistors stacked with dual-output SC converter for V-S measurement

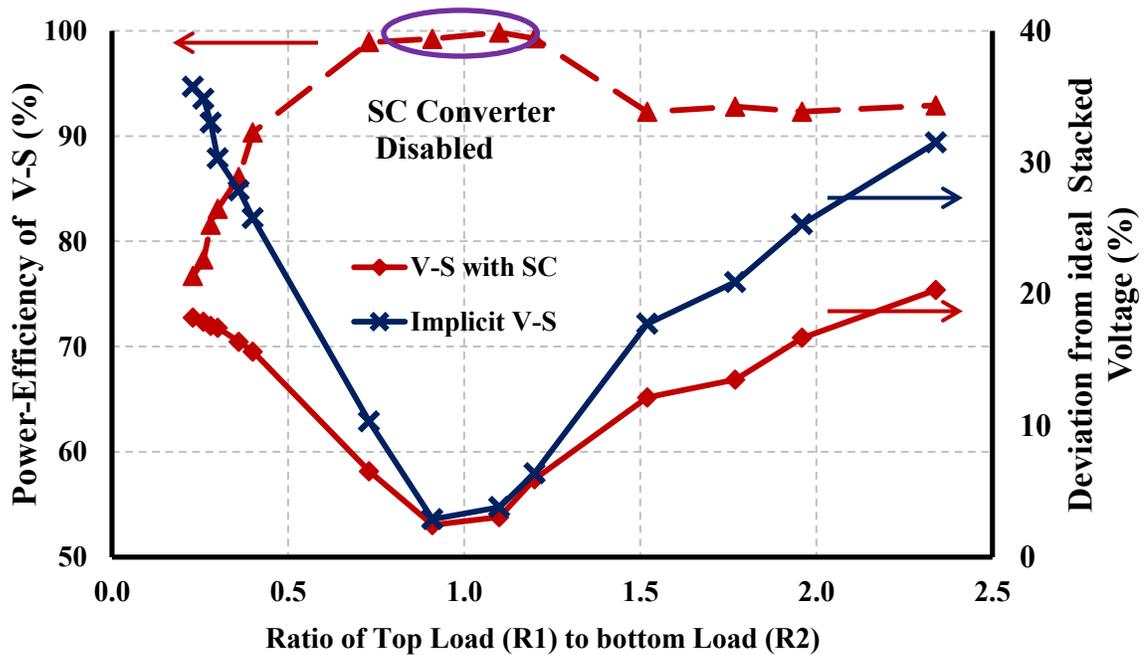


Figure 4.20. Power efficiency performance of V-S with resistive loads with peak efficiency of ~98% measured with SC converter disabled in near-balance load condition ( $R1:R2=1$ )

#### 4.4. Summary and comparisons with state-of-the-art integrated converters

Aggressive power management strategies are needed to sustain the growth of power-hungry SoC devices. Existing IVR architectures have almost saturated the limits of their performances in the present technologies. V-S can provide a simple yet elegant and cost-effective solution with its series-stacked architecture and differential regulation. Figure 4.21 shows a comparative analysis based on FOM (area overhead and power overhead) of different IVR and V-S with zoomed-up plot shown at the bottom. The best candidate will be the one closest to the origin with minimum loss and area overhead. It is difficult to compare V-S with traditional regulation, since charge-recycling performance is dependent on workload imbalance. For a fair comparison, the regular IVR performances are extracted based on output load of 1W, while the differential converter regulates stacked (two loads with total power consumption of 1W). That is, in the worst-case (WC), the differential converter is designed to handle 1W (similar to conventional regulator), while in the best-case (BC) with 0% mismatch and the average-case (AC) with 50% mismatch, the SC converter needs to regulate much lower current. Also area overhead of all of the converters is normalized with respect to trench-capacitor technology density. While worst-case design constrains V-S power density, very low power overhead can be achieved as seen from the zoomed-up section. A 100% power mismatch between the stacked loads is unlikely, even in the worst-case scenario, as leakage power will contribute. Moreover, as discussed previously, having some notion of

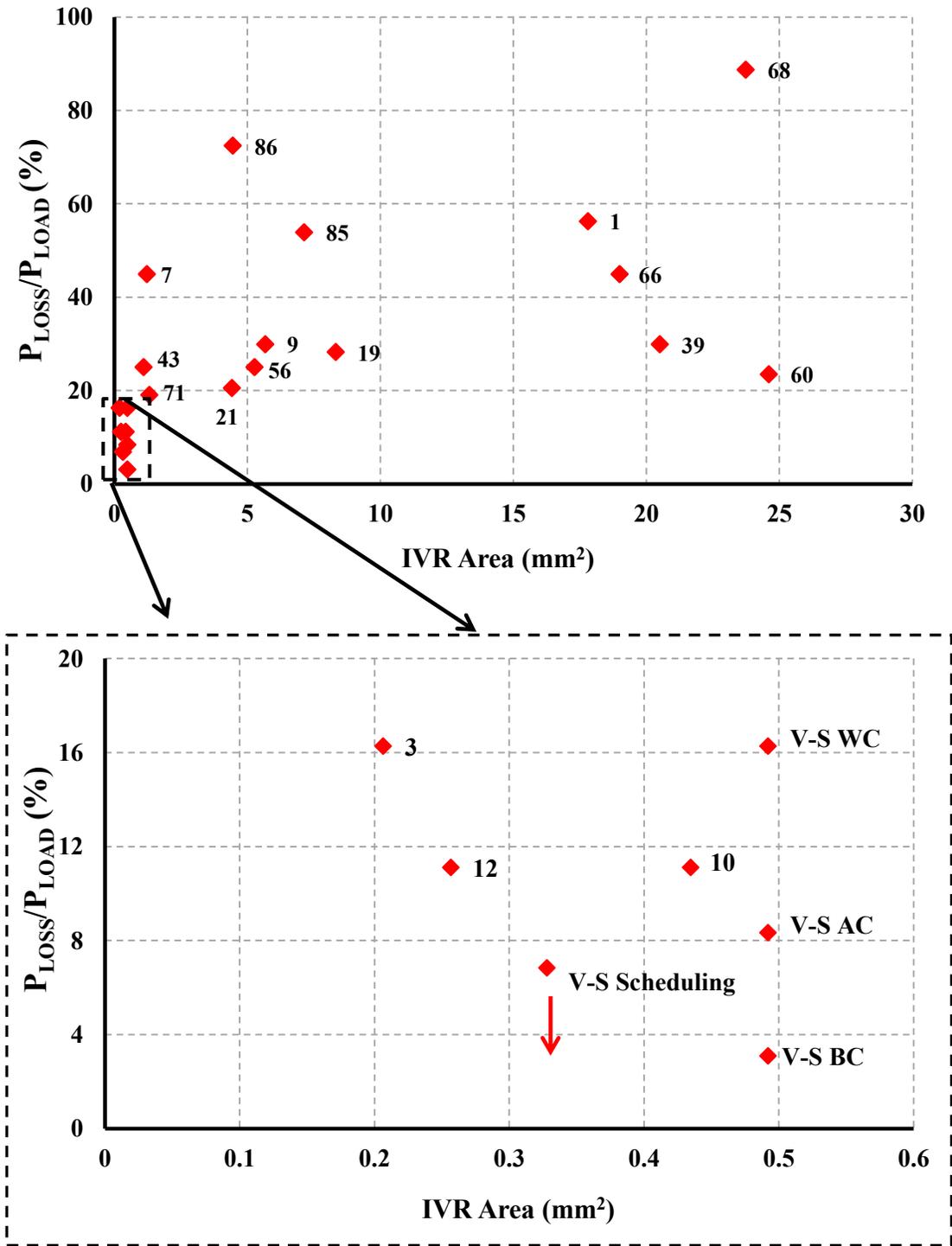


Figure 4.21. Comparison with state-of-the-art switching regulators shows how V-S lowers power overhead beyond the saturated limits. Zoomed-up figure shown below

scheduling in V-S, either software or hardware driven, will improve performance of this PDN and make it more practical. This is exemplified in Figure 4.21 by the V-S scheduling data (by scheduling it is assumed here that worst-case imbalance will be 50% through workload scheduling (31), therefore half the conventional converter can handle 1W load power). To enable the superior efficiency, the SCVRs ((3, 10, 12, 43) have used advanced capacitive technology (Trench-cap, Fe-cap), additional voltage rails to generate low voltage swing for SCVR switches and level shifters. In the proposed V-S scheme, push-pull SCVR has a moderately high efficiency of ~80% in standalone mode using MIM-cap and simple scalable circuit topology with single voltage rail, but the stacked system efficiency achieved is more than 90% for a wide range of load. Using additional techniques and advanced capacitor will lower the power overhead even further in V-S, whereas it has pretty much saturated the maximum limits of conventional SCVR. Similarly, a conventional regulator cannot sustain a load higher than its maximum drivable limit, but due to its differential nature, V-S can sustain much higher power and improve system power density significantly more than existing state-of-the-art IVRs.

## Chapter 5

---

### **5.1. Breaking the 3D-IC power delivery wall using voltage stacking**

The semiconductor industry seems poised to continue the historic Moore's Law trend of doubling the number of transistors every 1.5-2 years, even as the virtuous cycle benefits of Dennard's Scaling (devices that are simultaneously smaller, faster and lower power) are quickly vanishing. An essential tool in the arsenal of techniques the industry will use to continue the exponential increases in levels of integration is three-dimensional integrated circuits (3D-IC). Apart from a reduction in footprint leading to high device density, 3D-IC also provides a platform for heterogeneous integration; different nature of circuits (analog, digital, RF), functionalities (logic, memory, sensors), or even process nodes can be used for different layers. However, 3D-IC raises several fundamental technical difficulties in addition to the clear fabrication engineering challenges. As the number of physical layers in a 3D-IC stack will increase in the future, from the present 2.5D multi-layer solutions, with only a couple of layers, to true 3D many-layer stacks

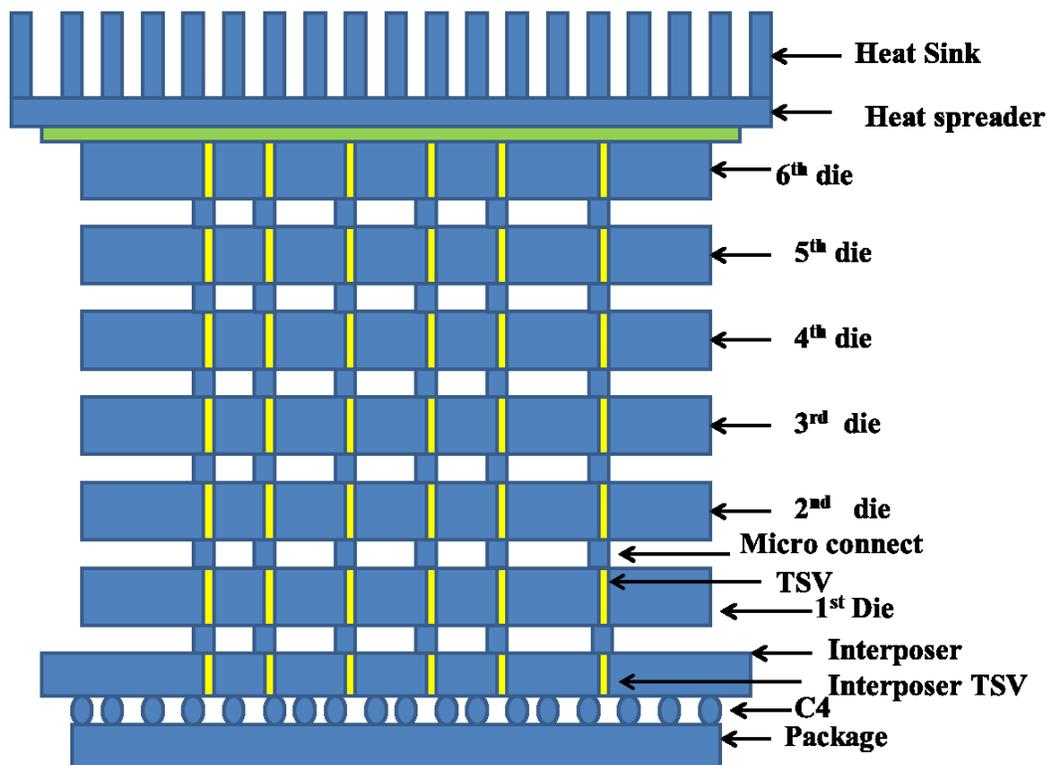
with many layers, the problems of delivering power to the 3D stack seem daunting (33, 55).

This chapter proposes an elegant solution in the form of 3D multi-output SC converter assisted voltage stacking (3D-MOSC) for breaking the power delivery walls in 3D-IC through charge-recycling (50). A literature review on 3D-IC power delivery shows that existing regulators will not scale efficiently for many-layer ICs. Starting with a circuit implementation of a charge-recycled voltage regulator, an architecture-level 3D-IC model, including pre-RTL PDN model with V-S, has been designed for cross-layer design explorations (90,91). An exhaustive trade-off study among voltage noise, differential IVR overheads, reliability improvements in PDN resources (example, C4, TSV), system efficiency, density and many other features, has been performed to justify V-S as an efficient and scalable solution to break the 3D-IC power delivery walls.

### 5.1.1. Power delivery and heat removal walls for 3D-IC

3D-IC is a general term for technologies that use several silicon layers on top of each other, such that more transistors can be integrated in the same lateral footprint. There are currently several different flavors of 3D-IC, with many of them only accommodating a couple of layers, thus more appropriately they are called 2.5D (33, 61). For the purpose of this dissertation though, focus is only on 3D-IC schemes that are scalable to many layers; such 3D-IC solutions will use a large number of thinned silicon layers that are connected to each other using through-silicon-vias (TSVs). Since these are true three-dimensional structures, power is usually delivered using C4s on the top of the topmost

layer in the stack (flip-chip), while heat is removed by a heat sink attached to the bottom of the bottom layer in the stack (shown in Figure 5.1). To stay on the historical Moore's law trend will require the number of layers in the 3D-IC stack to increase, which means the power consumption and dissipation will increase in a cubic fashion (with the volume of the stack), while the power delivery and heat removal will be limited by the quadratic area of the top and bottom layers. Even without 3D-IC, the number of power and ground pins and C4s on current chips is in the hundreds and can take more than half of the total number of pins [ITRS]; thus, it is clear that simply trying to keep up by increasing the number of pins and C4s is unsustainable for 3D-IC. Instead, there is an urgent need for practical, economic and scalable solutions for 3D-IC power delivery.



**Figure 5.1. Many-layer 3D-IC PDN including TSVs, C4s, Micro-connects**

### 5.1.2. 3D-IC literature review

In response to power delivery challenges caused by an unsustainable increase in 3D-IC current consumption, various research works (9, 70, 78) have explored the idea of using high on-chip voltage (instead of high current) by bringing in conventional off-chip DC-DC converters on-chip. Boosted voltage along with local regulation can alleviate some of the high current related issues. In (9) an IVR module is placed on a separate die within the 3D-IC packaging. Use of flip chip and through-hole-packaging technology has been discussed to vertically integrate regulator with processor (70). Making the DC-DC converter off-chip and using TSVs to deliver the power to each stratum is also explored in (70). This kind of stacked chip approach may work well for two tiers, but is not scalable across many tiers. A fully-integrated buck converter in 2D BiCMOS technology has been designed to be integrated in 3D-IC (78). To facilitate this 3-D integration, the footprint of the power supply must be comparable to the chip area below. This restricts the quality factor of the inductor, therefore increasing the switching frequency and reducing the efficiency to 62%-64%. Capacitors have significantly higher density than inductors and, while there have been numerous works on multi-output SC converters for 2D-IC, not much has been proposed for 3D-IC integrated regulation, significant energy losses of the capacitive converters being the primary reason. Brute force solutions i.e. doing whatever being done in 2D-IC, but quantitatively more, will not work for 3D-IC due to lack of scalability. Therefore this *3D volume vs. 2D surface* power wall mismatch needs a fundamental solution.

### 5.1.3. Multi-output switched-capacitor assisted V-S in 3D-IC (3D-MOSC)

Voltage stacking with charge-recycled power regulation (as discussed in chapter 4) can both solve the 3D-to-2D mismatch problem in a fundamental way, and at the same time provide a practical, efficient and scalable solution (50). V-S in 3D-IC involves different layers being electrically connected in series with a raised supply voltage (thus their voltages add up according to Kirchhoff's voltage law), instead of in parallel, as in a conventional approach (in which the currents add up), represents a fundamental solution keeping the total current constant (while the overall voltage increases) even as the number of layers scale up. This way power pins limitation in 3D-IC is overcome allowing significantly higher power density with little cost.

While V-S allows highly efficient power conversion (implicit down-conversion) and has, therefore, been proposed for breaking power regulation wall, as discussed in chapter 4, the number of loads that can be stacked will be limited (usually three) because of CMOS process-related constraints in 2D-IC technology. For bulk CMOS technology, either substrate is common for all stacked loads, therefore making the transistors in different stacks asymmetric owing to body-bias effect. Even if isolated biasing of individual p-well and n-well is possible (as in triple-well CMOS), there is a maximum voltage limit for well-to-substrate (for example 3V for 28nm SOI technology). However, physical layering of 3D-IC (i.e. physical isolation of layers with electrical connection through TSVs) naturally maps to V-S, providing an easily scalable power delivery solution for 3D-IC.

Compared with the conventional power delivery scheme for an N-layer 3D-IC, V-S reduces the off-chip and cross-layer current density by up to N times through recycling

charges between layers. This not only reduces the resistive noise (i.e., IR drop) across the PDN, but also significantly improves PDN's EM-induced reliability. However, a major design challenge arises from the fact that V-S will try to compensate for any current-consumption mismatch between the stacked loads by re-distributing the intermediate voltage, giving rise to voltage noise. To make the envisioned 3D-IC V-S practical, explicit voltage regulation is required for the general case when the currents of the various layers are not perfectly matched. This dissertation is the first to propose multi-output switched-capacitor (MOSC) regulation as an efficient method to complement and assist voltage stacking as a truly scalable and practical solution. By only having to regulate the difference in currents among the layers in the stack, the high conversion efficiency of a SC converter is effectively enhanced to higher numbers. The dual-output stacked SC converter, discussed in-depth in chapter 4, has been extended to a multi-output converter by stacking unit-cells for supporting many layers. Most of the circuit design and simulation in this dissertation has been restricted to single/dual-output SC converters for the sake of simplicity. However, modeling work on the capacitive converter has been extended to multi-output SC converter (MOSC), along with a many-layer 3D system-level PDN model. Due to the availability of dense integrated CMOS capacitive technology, the primary focus of this work has been on MOSC for explicit regulation. Alternatively, a bidirectional buck/boost converter has been used for differential power processing in high voltage photovoltaic cells and a similar philosophy can be adopted for V-S as well in future work (75).

## 5.2. System-level evaluation of V-S

Although researchers have previously identified V-S as a promising solution to alleviate the power delivery constraints in 3D-IC (24), a cross-layer study that examines the system-level impacts of voltage-stacking's current reduction, the area and power overheads of explicit voltage regulation, and the supply voltage noise under different workload conditions, is needed to justify an unconventional approach such as V-S for 3D-IC. This work has been a collaborative effort with Zhang (88) and the contributions of each of the authors are discussed at the end of the chapter in section 5.4.1.1.

A 3D PDN model incorporating V-S has been developed for system-level evaluation of charge-recycled power regulation in 3D-IC. Adopting a design methodology similar to (93), at first compact models of voltage regulators are designed and validated against circuit simulations. A resistive model capturing IVR output impedance with DC voltage drop (for faster system-level simulation allowing wide-range tradeoff study), and a compact RC model with both resistors and capacitors emulating SCVR (slower system-level simulation time but with accurate noise results) are designed and validated. An open-source, system-level PDN model, VoltSpot, is then extended into a 3D PDN and integrated with the IVR models, producing the first platform to enable whole-system, transient simulation for many-layer 3D-ICs' V-S PDN (3D-Voltspot) (89,90). Since the supply noise in V-S PDN is strongly correlated with the workload imbalance between the adjacent layers (50), a large range of workload imbalance is examined, using an example low-power, ARM-based many-core 3D processor and quantified the impact of charge-recycled power regulation on voltage noise, system power efficiency and PDN area

overhead. Valuable PDN resources such as on-chip wires, C4s, 3D metal stacks and TSVs are incorporated as distributed RLC in the 3D PDN model, also known as 3D-Voltspot.

### 5.2.1. Modeling methodologies

The power delivery networks of contemporary processors are usually large systems that contain up to several billion nodes, even in the context of 2D-IC. 3D integration and voltage stacking further increase the PDN's complexity with more device layers and new components like TSVs. For this reason, circuit-level simulations will be extremely computational-intensive and incapable of supporting whole-system design-space exploration studies. To enable a system-level study of V-S PDN's voltage noise, a compact RC model for the SC converters is designed, validated and integrated with a pre-RTL PDN model. This section discusses the different modeling methodologies and the validation results.

#### 5.2.1.1. *SC Converter: Impedance Model*

A cross-layer design exploration of the benefits and overheads of V-S in 3D-IC requires incorporating circuit-level insights with architecture-level study. To accurately capture power efficiency, output voltage drop and area overhead of SC converters, a 2:1 push-pull SC converter to be used in V-S has been designed in a commercial 28nm CMOS technology. The design details are provided in Table 5.1 and include area overhead assuming different capacitive technology. Using cadence ADE environment and Spectre simulator, the converter is simulated and the results are extracted to build a compact

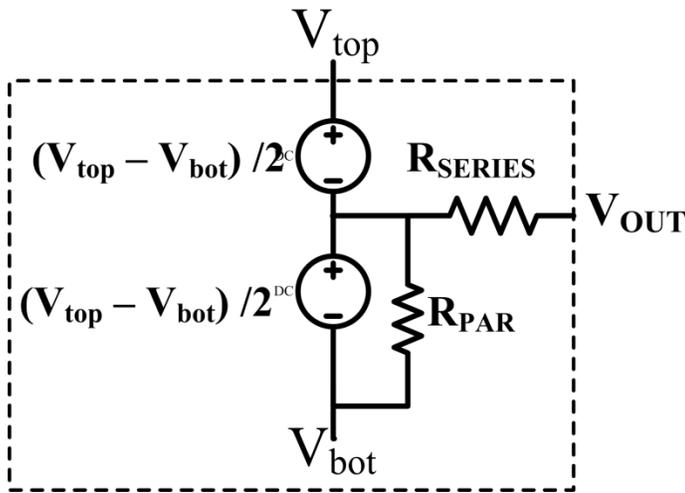
model for system-level exploration.

**Table 5.1 SCVR design specification (used in circuit design and modeling)**

SC Converter Design Specification	Numerical quantity
Max Load Current (mA)	100
Total Fly-cap (nF)	8
No of Interleaving	4
Optimum Fsw in open-loop(MHz)	50
Silicon Area (MIM-cap) (25) (mm <sup>2</sup> )	0.472
Silicon Area (Fe-cap) (12) (mm <sup>2</sup> )	0.102
Silicon Area (Trench-cap) (3) (mm <sup>2</sup> )	0.082

SC converter can be modeled as an ideal voltage source with desired ratio of conversion and the two resistors ( $R_{\text{SERIES}}$  and  $R_{\text{PAR}}$  in Figure 4.6) capturing the different intrinsic and parasitic losses. While  $R_{\text{SERIES}}$  models the switching and conductance losses,  $R_{\text{PAR}}$  captures the losses in driving switch parasitic capacitance and bottom-plate capacitance. This modeling is adopted from an analytical methodology introduced in (72) and further discussed in chapter 4. The slow ( $R_{\text{SSL}}$ ) and fast switching ( $R_{\text{FSL}}$ ) asymptotic limits of SC converter output impedance are given by Equation 4-3 and 4-4, while the resistive voltage drop i.e. IR drop of the SC converters is captured across  $R_{\text{SERIES}}$  given by Equation 4-5. For the SC converter designed with Table 5.1 specification,  $R_{\text{SERIES}}$  is calculated as  $0.6\Omega$ . As shown in Figure 5.2, the voltage-headroom (i.e., the potential difference between  $V_{\text{top}}$  and  $V_{\text{bot}}$ ) of the SC converter in the many-layer 3D-IC is

dependent on the adjoining layers' workload imbalance. In order to incorporate this dependency in the cross-layer study,  $V_{top}$  and  $V_{bot}$  and  $V_{OUT}$  are all included as inputs to the SC converter model and ideal output voltage  $V_{OUT}$  (i.e., without the IR drop on  $R_{SERIES}$ ) is calculated as  $(V_{top} + V_{bot})/2$ .

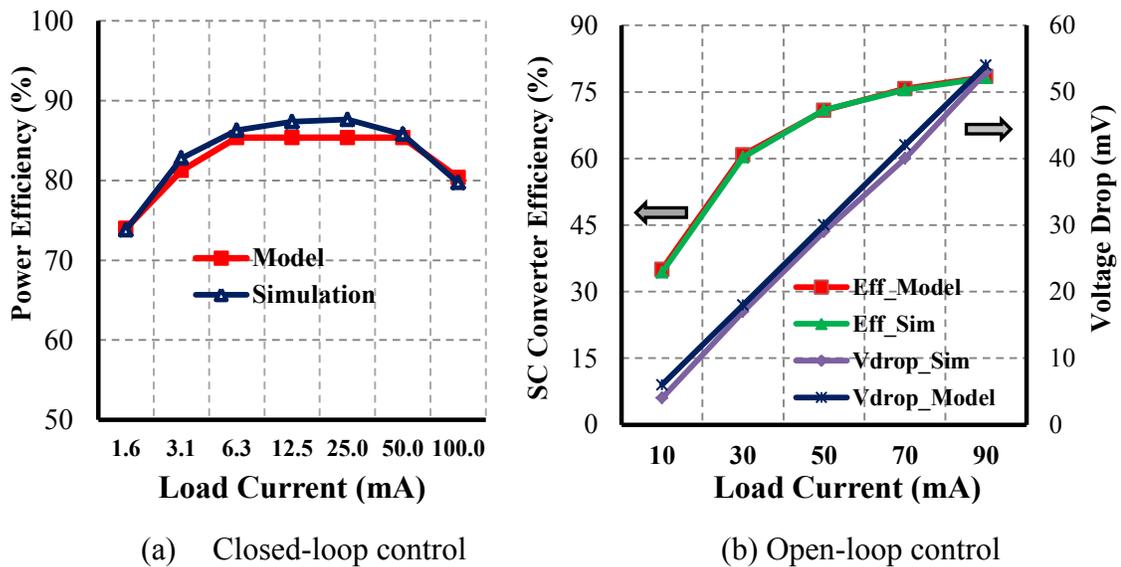


**Figure 5.2. Impedance model of SCVR, capturing output voltage DC drop and power-losses (intrinsic and parasitic) (89)**

### 5.2.2. SCVR impedance model validation

The SC converter can be configured for either open or closed-loop regulation. While closed-loop regulation is traditionally used for best power-optimization, V-S allows an interesting tradeoff with open-loop regulation, as explored in chapter 4. To verify the accuracy of the model, both open-loop (i.e. constant switching frequency) and closed-loop (i.e. varying switching frequency) regulation of the 2:1 SC converter are compared against circuit simulations under varying load currents.

Figure 5.3 shows the model accurately capturing the power efficiency and output voltage drop for both of the control schemes. According to Figure 5.3, closed-loop converters have higher power efficiency. However, implementation of the feedback loop makes a closed-loop scheme more difficult to model. For simplicity, open-loop SC converters are added to the 3D V-S PDN analysis and, henceforth, all analyses involve open-loop regulation, leaving closed-loop control modeling for future work.



**Figure 5.3. Model versus simulation validation results (89)**

### 5.2.3. Transient model for SC converter

SCVR with its switches and capacitors switching every phase of clock cycle interact with PDN parasitic (L, R and C) to give rise to significant  $Ldi/dt$  and LC resonance noise.

Voltage stacked SCVR with instantaneous charge mismatch between stacked loads adds

a further dimension to voltage noise i.e. workload imbalance. A resistive SCVR model will not accurately capture all these and a transient model with switches and capacitors is needed. Effect of interleaving, which is a common strategy to reduce output voltage ripples by running multiple SCVR blocks out-of-phase, also needs to be included in the modeling.

Figure 5.4 shows the compact RC circuit used to model the interleaved SC converters (90). Each pair of top and bottom RC branches represents a cell of the converter that is controlled by a separate clock signal. At each clock edge, the positions of the top and bottom fly-caps are exchanged, to model the switching activities of the converter cell. That is,

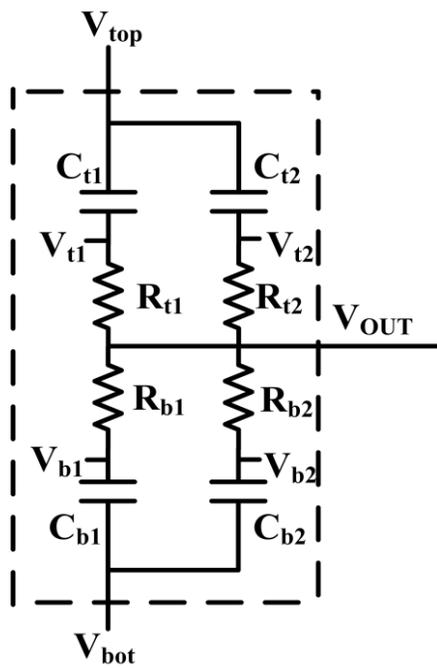
$$V_{\text{top}} - V_{t1} = V_{b1} - V_{\text{bot}} \quad 5-5$$

is calculated, where  $V_{t1}$  is the voltage value after the clock edge, while  $V_{b1}$  is the value before. Note that although the positions (i.e., electric charge) of the fly-caps at each clock edge are exchanged, the resistance of each top and bottom branch is kept unmodified. This is because each time the positions of the fly-caps are “flipped”, the set of switches to conduct the current are also changed (Figure 4.4). However, the switches are designed in a symmetric way such that both the top and bottom RC branch in the two different clock phases have the same equivalent resistance. Therefore, the eight switches can be collapsed into two resistors ( $R_t$  represents SW1&5 and SW2&6,  $R_b$  represent SW4&8 and SW3&7) to reduce the model’s complexity. From circuit simulations,  $R_t$  and  $R_b$  values are extracted as

$$R_t = 4.208\Omega, \text{ and } R_b = 4.68\Omega$$

( $R_t \neq R_b$  as NMOS and PMOS have different channel resistances). As common design

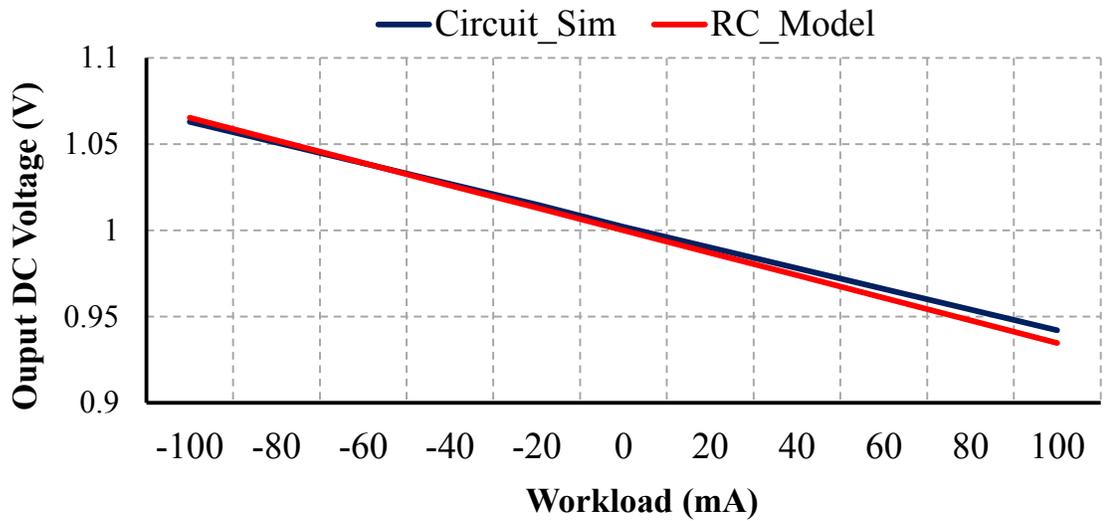
technique to smooth the output voltage ripple, designers divide the single-cell converters into multiple sub-cells and interleave their switching clocks (26). To model this interleaved structure, a pair of top/bottom RC branches are simply instantiated for each sub-cell, capacitance values are scaled according to the number of total sub-cells, and phase of each sub-cell's control clock are shifted. Figure 5.4 also illustrates two-way interleaving in the SC converter. Since the sub-cells are stacked for a multi-output converter, it is assumed that all the sub-cells have identical structure, and therefore the same RC values under ideal condition (i.e. identical voltage headroom on all layers).



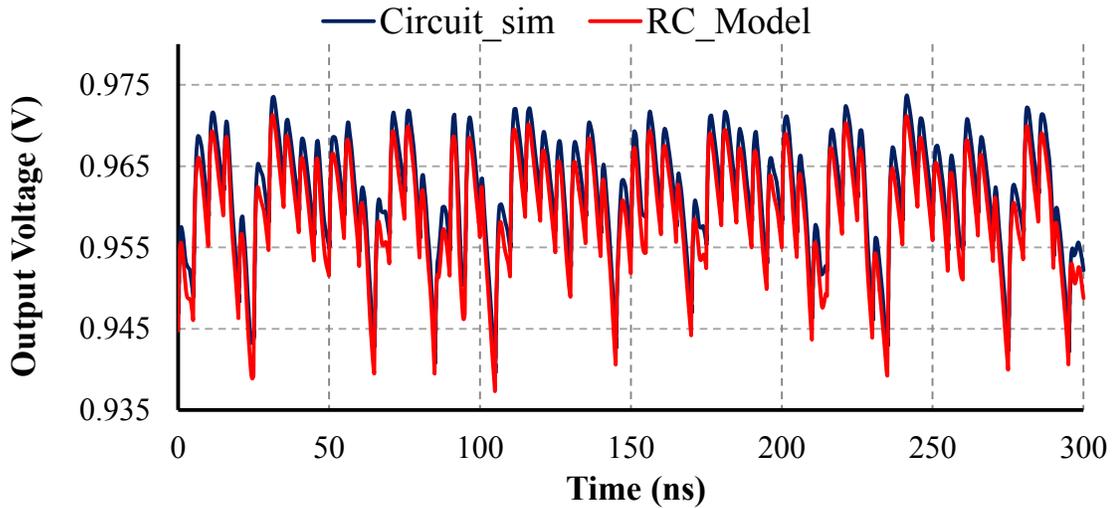
**Figure 5.4. 3 port RC model for 2-way interleaved SC converter (90)**

### 5.2.3.1. *SCVR transient model Validation*

Accuracy of the transient RC model is compared against simulation results for a two-layer, voltage-stacked system (i.e.,  $V_{\text{top}} = 2\text{V}$ ,  $V_{\text{bot}} = 0\text{V}$ ). Figure 5.5 (a) shows the DC results comparison under constant workload conditions. Since the SC converter's output voltage is directly related to its output current, an ideal current source is attached directly to the  $V_{\text{OUT}}$  port and the test cases are swept from maximum sourcing (positive 100mA) to maximum sinking (negative 100mA) current. Under a constant workload, the output voltage shows a periodic rippling behavior caused by the converters' switching activities. Validation result shows that the proposed RC model's maximum DC error is 75mV, or 0.75% Vdd. A time-varying load current is also used to validate this model to compare the accuracy in capturing steady-state transient response. Figure 5.5 (b) shows the output voltage trace over 300 ns. The load current is sampled from Parsec 2.0 benchmark raytrace (100); it induces an average current of 66.3mA in an ARM Cortex A9 core. Over the entire simulated time window, the output voltage trace of the proposed RC model matches well with circuit simulation in terms of DC component, AC amplitude, and slew rate. Overall, this model can capture the SC converter's transient output voltage with less than 72mV error at all times.



(a)



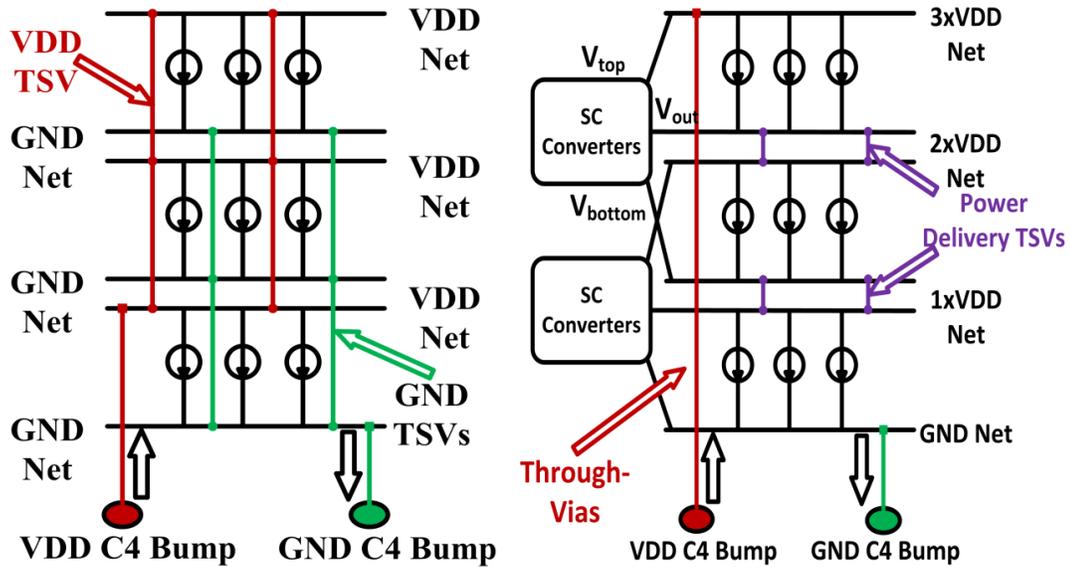
(b)

**Figure 5.5. Model versus simulation validation results (a) Output DC voltage (b) Output transient voltage trace (90)**

#### 5.2.4. Whole-system model

In order to quickly explore the multidimensional space of 3D-IC's PDN design and evaluate the cost and benefits of different design scenarios, Zhang et al. have proposed an early-stage PDN simulation using VoltSpot, a pre-RTL PDN model (91). VoltSpot

uses ideal current sources to model load (i.e., both dynamic and leakage power of the switching transistors), and RLC elements to model the on-chip PDN metal stack, C4 pads and chip package. Since VoltSpot only models 2D chips, it has been extended to support 3D-IC (89). Figure 5.6 illustrates this extension. To model the traditional PDN for 3D-IC, more layers of silicon are simply added on top of each other and all layers' Vdd nets and ground nets are connected with TSVs (Figure 5.6 (a)). To model V-S PDN, all layers' Vdd nets and ground nets are connected in series with regular TSVs and off-chip supply voltage (i.e., the single layers' Vdd multiplied by the number of layers) is provided to the top layer using TSVs (Fig. 4b). These TSVs are modeled as resistors. The resistive model and the transient model for SC converters have been described in Sec. 5.2.2 and 5.2.3. Depending on the tradeoff study (i.e. one that requires to capture IR drop versus one that needs accurate transient performance), these models are uniformly distributed within each core. For each SC converter, its three ports (i.e.,  $V_{top}$ ,  $V_{OUT}$ , and  $V_{bot}$ ) are connected to three consecutive layers in the voltage-stacked power grids. Ideally,  $V_{OUT} = (V_{top} + V_{bot})/2$ , which indicates that any change in either  $V_{top}$  or  $V_{bot}$  will also affect the regulator's output voltage. This way the model directly captures this inter-layer voltage dependency.



(a) Regular PDN

(b) V-S PDN

**Figure 5.6. PDN structure for 3D-IC (89)**

With the fine-grained pre-RTL modeling capability inherited from VoltSpot, 3D-Voltspot provides a detailed current profile for both the C4 pad and TSV arrays. It also captures on-chip IR drop for both regular PDN and V-S PDN under given workload behaviors. This model provides a key link to the tool chain that allows designers to explore the complex tradeoff space that involves power delivery architecture, C4 pad/TSV allocation, voltage regulation scheme, PDN noise/reliability, and workload characteristics. Figure 5.6 shows the structure of the whole-system model designed for many-layer V-S PDN.

### 5.2.5. Simulation Setup

#### 5.2.5.1. *Many-Core 3D Modeling*

For realistic voltage stacked 3D-IC design explorations, an example many-core, many-

layer 3D-IC based on a 40nm ARM Cortex A9 IP is modeled, using architecture-level power and area model McPat (44, 100). When running at 1GHz with 1V supply voltage, each core has a peak power density of  $172\text{mW}/\text{mm}^2$  (475 mW over  $2.76\text{ mm}^2$ ). Due to the power-efficient nature of these ARM processors, many-layer 3D-IC can be built without relying on aggressive, volumetric cooling solutions. With the help of pre-RTL floorplan tool ArchFP (15) and thermal model HotSpot (76), 3D stacks' maximum temperature is evaluated and it is found that, with a conventional air-cooling solution, a stack of up to eight layers of 16-core processors can be built without violating the typical upper limit of 100 degree Celsius. Apart from thermal and power delivery constraints, many-layer, especially many-logic-layer 3D-ICs, pose various fabrication challenges (69). However, the possibility of manufacturing 3D stacks economically has been exemplified in recent times by existing commercial products (e.g., the Micron hybrid memory cube with 4-8 layers (102)) which lend credibility to this many-layer 3D-IC model. To study the design tradeoffs for aggressively scaled voltage stacked 3D-IC and to evaluate how 3D scaling affects PDN design, a series of example 3D systems with 2 to 8 layers stacked together are designed. With 16 ARM cores per layer, the peak power consumption of these 3D processors ranges from 30.4W to 60.8W.

#### **5.2.5.2. *PDN Modeling with different TSV configurations***

A major extension made to existing Voltspot is adding an explicit resistor-inductor model for TSVs, based on specifications from prior work (32). Here the assumption is that all TSVs have equal size and resistance, and that they are uniformly distributed within each silicon layer. TSV capacitance in this work is usually orders-of magnitude smaller than

the on-chip and package decoupling capacitance and is therefore ignored. PDN modeling parameters, adopted from previous work (91), are listed in Table 5.2.

**Table 5.2 Major 3D PDN modeling parameters**

PDN Parameters	Numerical Quantity
C4 Pad pitch ( $\mu\text{m}$ )	200
C4 Pad resistance ( $\text{m}\Omega$ )	10
Minimum TSV pitch ( $\mu\text{m}$ )	10
TSV diameter ( $\mu\text{m}$ )	5
Single TSV's resistance ( $\text{m}\Omega$ )	44.539
TSV keep-out zone's side length ( $\mu\text{m}$ )	9.88
On-Chip PDN's pitch, width, thickness ( $\mu\text{m}$ )	810,400,720

Number of TSVs allocated for PDN is a design parameter; more TSVs provide more vertical current delivery channels, thereby reducing both average TSV current and the effective inter-layer PDN resistance at the cost of higher area overhead. To explore the tradeoff between power delivery quality and TSVs' area overhead, three TSV topologies are examined in this study; conservative (Dense), aggressive (Few), and average (Sparse), design scenarios. Table 5.3 gives more details about each configuration's TSV count and area overhead.

**Table 5.3 TSV configurations used in this work**

TSV Topology	Effective Pitch ( $\mu\text{m}$ )	Number of TSVs per core	Total Area overhead
Dense TSV	20	6650	24.2%
Sparse TSV	40	1675	6.1%
Few TSV	240	110	0.4%

### 5.2.5.3. *Workload Modeling*

Using an integrated tool flow that combines McPAT with performance simulator Gem5 (103), Parsec 2.0 benchmark suite has been simulated and dynamic power consumption traces are extracted to build realistic test cases. Each workload's average power consumption and maximum noise amplitude, when running alone (on a 2D-IC), is profiled to be used as load for various design explorations. Further details about workload modeling, Voltspot and 3D-Voltspot PDN model used in this dissertation are provided in (88).

## 5.3. **Cross-layer Design exploration in Voltage Stacked Many-layer 3D-IC**

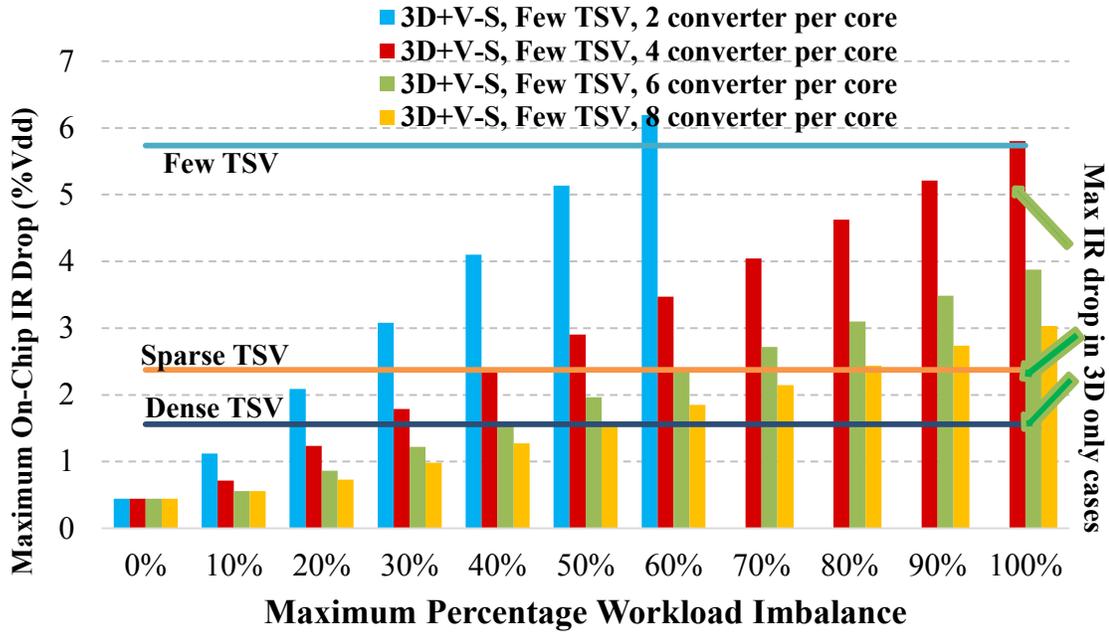
### 5.3.1. Load imbalance induced voltage noise

3D-IC technology brings many challenges for efficient and reliable power delivery. V-S

claims to break these power delivery walls through implicit down-conversion and differential SC converter regulation. However, load imbalance-induced voltage noise threatens to kill the benefits of V-S. 3D-Voltspot gives designers a platform to do an in-depth study of voltage noise overhead in V-S.

For 3D-ICs without V-S, worst-case IR drop happens when all layers are fully active. Therefore, the assumption about workload imbalance does not affect those evaluations. IVR is necessary in V-S PDN, because when the current consumptions of two adjacent layers do not match, the voltage regulators need to either provide or sink the difference. This introduces extra voltage noise due to the regulators' output voltage drop, Figure 5.5 (a), and the lateral impedance of the on-chip PDN. While larger workload imbalance increases noise with higher current demand regulated by the SC converters, having more regulators distributed across the silicon die reduces IR drop by amortizing the per-converter current load and reducing the average load-to-regulator distance. Figure 5.7 shows the noise levels of PDN for 8-layer 3D-IC under different regulator configurations and workload behavior conditions. Here the assumption is that the power consumption of the silicon layers has an interleaved “high-low” pattern, where the high-power layers are always fully active and the low-power layers consume  $X\%$  lower dynamic power (e.g., 100% imbalance means that the low-power layers are idle and only consume leakage power). This pattern serves as a good benchmark because it requires the converters on all layers to source/sink the same amount of current, therefore imposing the most stress on the PDN. This load-scheduling is used to study the worst-case noise of V-S PDNs. Since the SC converters are designed to have a maximum load of 100 mA, Figure 5.7 skips all data points that violate this limit. The lines in Figure 5.7 illustrate the maximum on-chip

IR drop of regulator PDNs with different TSV configurations. Regular PDN relies on TSVs to provide all current to all layers, and, therefore, the worst-case IR drop always happens when all layers are fully active. For this reason, regular PDNs' maximum IR drop results are irrelevant to the imbalance of workloads. Since adding one SC converter to an ARM core incurs around 3% area overhead (assuming the converters are implemented with high-density capacitors discussed in Sec. 2.1.5), a V-S PDN with 8 converters per core and "Few TSV" topology occupies the same area as a regular PDN with "Dense TSV" topology. If the voltage noise of these two cases is compared, the V-S PDN shows lower IR drop when the workload imbalance ratio is lower than 50%. When a larger imbalance exists, V-S PDN's IR drop surpasses regular PDN by up to 1.58% V<sub>dd</sub>. Therefore, the voltage noise in V-S PDN is dependent on the imbalance of different layers' power consumptions. With large workload imbalance, V-S PDN experiences more severe IR drop than regular PDN. Furthermore, unlike regular PDN, V-S PDN's voltage noise is insensitive to the number of layers. With the advance of cooling technologies that allows designers to stack more layers, V-S PDN will out-perform regular PDN, in terms of noise, even in the presence of a large workload imbalance.

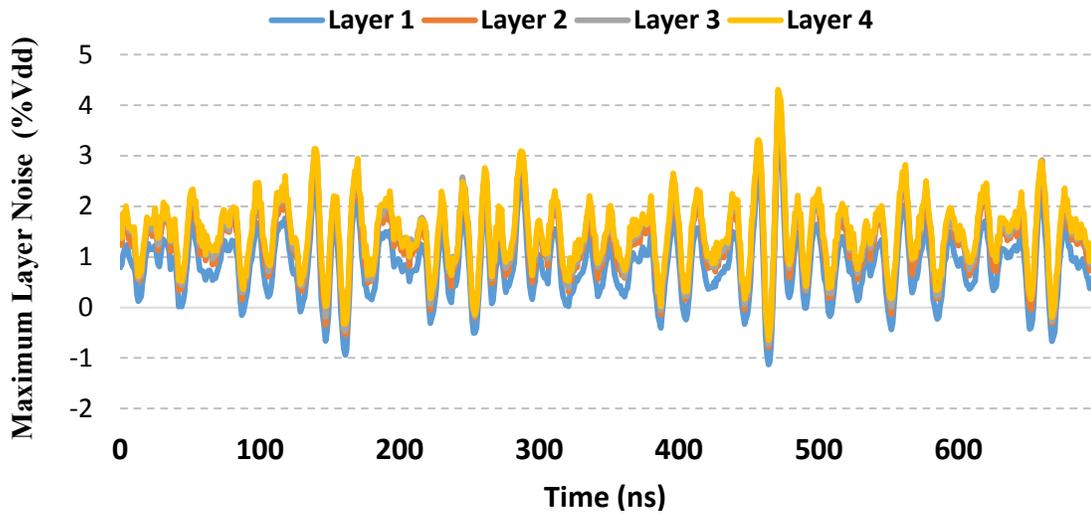


**Figure 5.7. Voltage noise evaluation with different TSV configurations, different workload distribution and different numbers of IVR in 8-layer 3D-IC with V-S (89)**

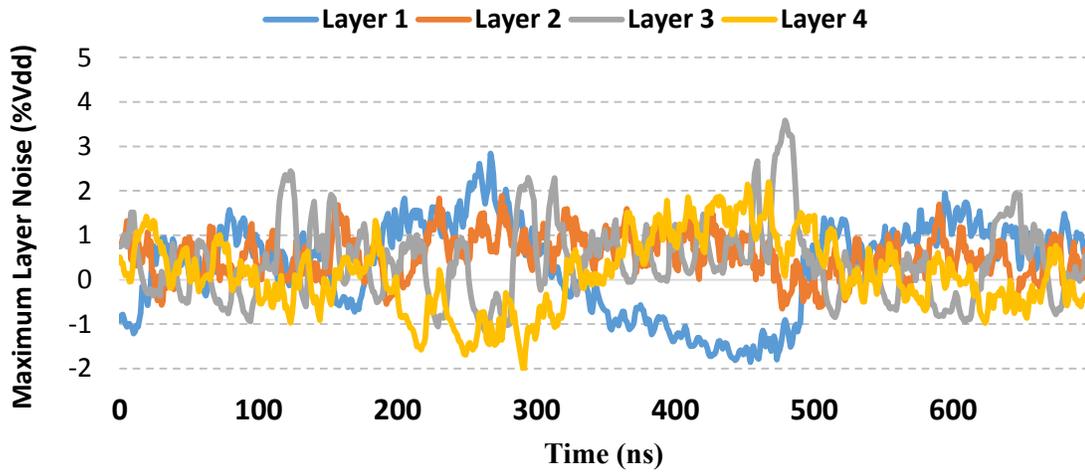
### 5.3.2. Cross-layer Noise Interference

To study transient behavior of the voltage stacked intermediate nodes and the dependency of voltage noise with other layers' voltage variations in V-S PDN, different workloads; one noisy and three less noisy workloads are assigned to a 4-layer example 3D processor. The first row in Table 5.4 shows each workload sample's maximum noise amplitude when running alone on a single-layer chip. Figure 5.8 shows each layer's maximum voltage noise (%Vdd) over time. In the traditional PDN (Figure 5.8 (a)), voltage noise in all layers is clearly highly correlated. Supply voltage fluctuations in one layer affect the entire 3D stack through the vertical connections (i.e., TSVs). Conversely, the V-S PDN connects layers in series and regulates voltage levels with SC converters. Consequently, it

breaks the inter-layer noise correlation, as in Figure 5.8 (b). Table 5.4 shows each layer's maximum noise amplitude over the entire simulated time window. Compared with a 2D PDN, the traditional 3D PDN significantly reduces Task3's noise because the decoupling capacitors (decap) on adjacent layers help to stabilize local voltage variation. However, other layers' voltage noise is also affected by Task3. V-S PDN isolates Task3's noise so that other layers have lower noise. Given the linear relationship between noise amplitude and transistor delay,  $x\%$  Vdd noise also requires an  $x\%$  decrease in clock frequency. This noise-shielding characteristic of V-S can be taken advantage of by using different architectural run-time adaptive strategy, for example, dynamic margin adaptation (42), to make less noisy layers run faster. The last column in Table 5.4 shows the arithmetic mean of all four layers' maximum noise amplitude. This cross-layer mean metric shows the whole-stack's average slowdown when per-layer margin adaptation is used. By isolating the cross-layer noise interaction, V-S PDN can improve system performance with less slowdown (88).



(a) Traditional PDN



(b) V-S PDN

**Figure 5.8. A plot of per-layer maximum noise amplitude over time. Only layer 3 has a noisy workload (90, 88)**

**Table 5.4. Maximum voltage noise (%Vdd) per layer for different workloads on 3D-ICs with different PDN schemes. The “cross-layer mean” value averages all layers’ maximum noise amplitude (90)**

PDN type	Task1	Task 2	Task 3	Task 4	Cross-Layer
	Layer 1	Layer 2	Layer 3	Layer 4	Mean
Single Layer	4.0	3.0	10.9	2.8	N.A
Traditional	3.7	4.2	4.2	4.3	4.1
V-S	2.8	1.9	3.6	2.3	2.7

### 5.3.3. Circuit/architectural co-simulation using SPICE

While the SCVR modeling has been validated against circuit simulations with good accuracy, circuit-architectural co-simulation of V-S is performed on SPICE to get better accuracy and confidence on this stacked regulation scheme. Different benchmarks from Parsec suite (100) - body tracking, raytrace (rendering application) and X264 (application for encoding video streams into the H.264/MPEG-4 AVC compression format), have been run on an architectural simulator (accurate full-system multicore simulator GEM5 and power model McPat) at 0.5V  $V_{IN}$  and clock-frequency of 100MHz, and the extracted power-traces (sampled at 5-cycle interval for a total of 100k execution cycles ) have been used to model resistive loads using Verilog A. Push-pull SC converters (with 80% power efficiency and 83.4mW/mm<sup>2</sup> power density) and the

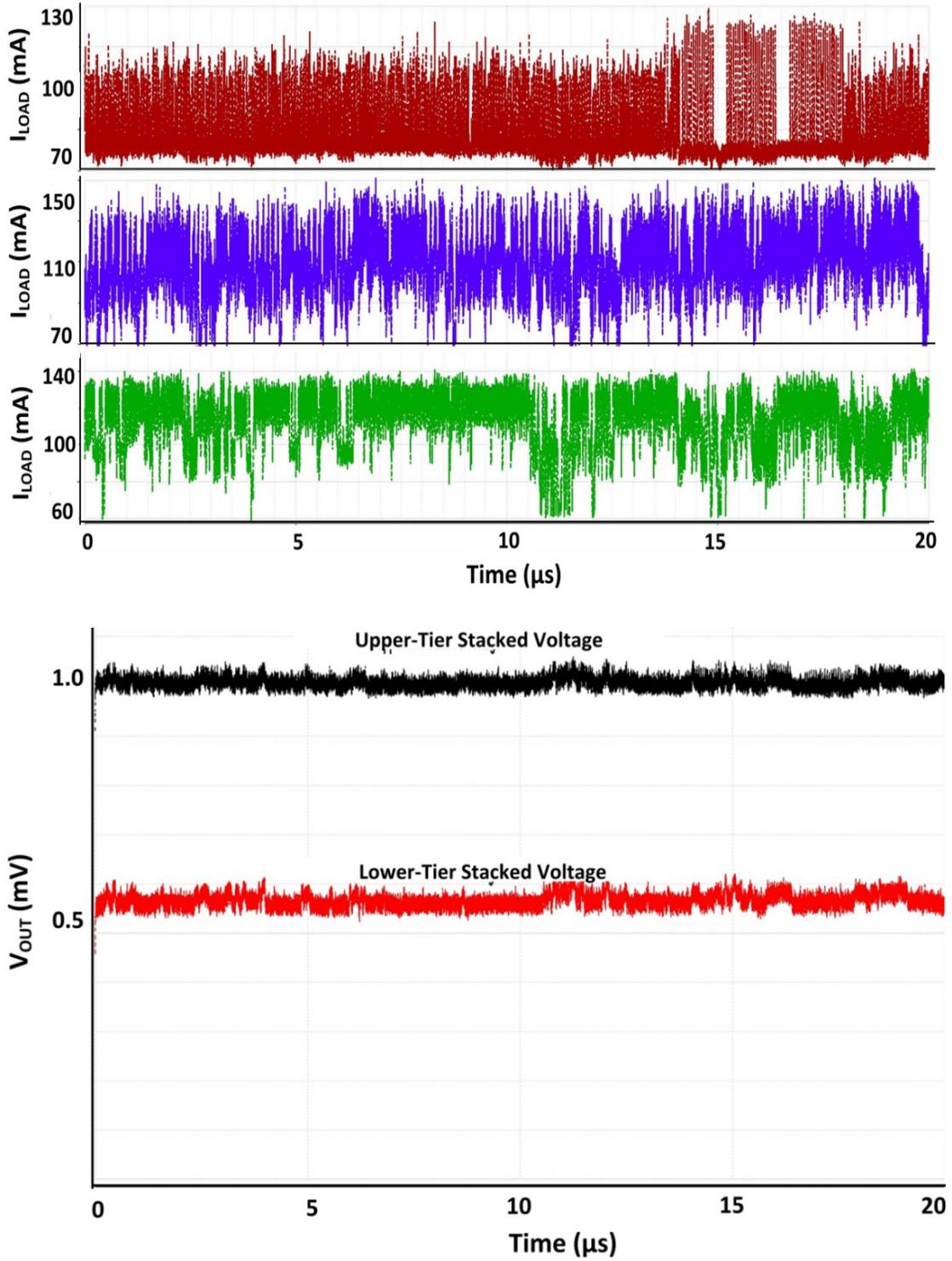


Figure 5.9. Three-layered stacked loads (architectural power traces) with stacked SC converter, simulated in 28nm FDSOI technology

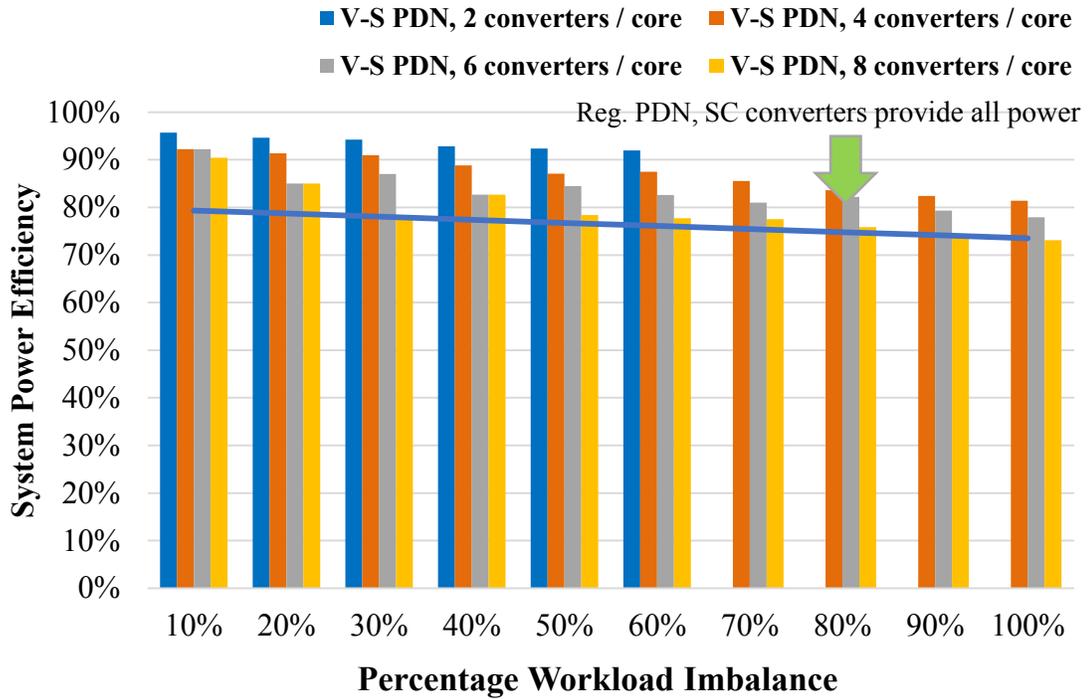
Verilog-A resistive loads are stacked, connected together (as shown in Figure 4.5) and simulated in Cadence ADE environment. Implicit charge-recycling provides the bulk of the current needed by the bottom stacked loads, thereby improving V-S regulation efficiency to 92.3% with output power density as high as  $185.3\text{mW/mm}^2$  and voltage noise of 9% and 2.8% for the two voltage-stacked nodes (Figure 5.9).

### 5.3.4. System power efficiency

#### 5.3.4.1. *System power efficiency with workload imbalance*

Figure 5.10 shows the power efficiency (i.e., the total power consumed by the processors divided by the total power drawn from the off-chip power source) results for 3D-ICs with V-S PDN. As the amount of workload imbalance increases, the SC converters need to compensate by delivering more power. Consequently, power overhead of voltage regulation increases. When V-S PDN with different numbers of SC converters is compared, it is observed that increasing the number of converters reduces power efficiency. This is primarily because SC converters running in open-loop suffer from lower efficiency as more converters are allocated to share the current load. Closed-loop control is an area for future work. Considering the fact that placing more converters can reduce on-chip IR drop, the allocation of SC converters in V-S PDN becomes a tradeoff between on-chip voltage noise and system-level power efficiency. 3D Voltspot can help designers to choose the optimal design point, based on their specific design objectives. Figure 5.10 also compares the power efficiency of using SC converters in 3D processors with regular PDN. Unlike V-S PDN, where the voltage regulators only need to

compensate for the differential power consumption between layers, SC converters in regular PDN have to provide current to all layers. As a result, V-S PDNs have higher power efficiency (close to 90% for 50% workload imbalance).



**Figure 5.10. System power efficiency evaluation with different workload distribution and different numbers of IVR in 8-layer 3D-IC with V-S**

### 5.3.5. V-S implementation in 3D-IC technology

Lack of availability of 3D-IC foundry has been a major challenge in academic research. However, with semiconductor industries such as Tezzaron, MOSIS, Xilinx, and Micron taking a leading role, 3D-IC industry is expected to progress in the near future. In the absence of actual technology, NCSU state EDA group provides a 3D platform for demonstrating and debugging new design tools for 3D-IC (104). It even includes layout

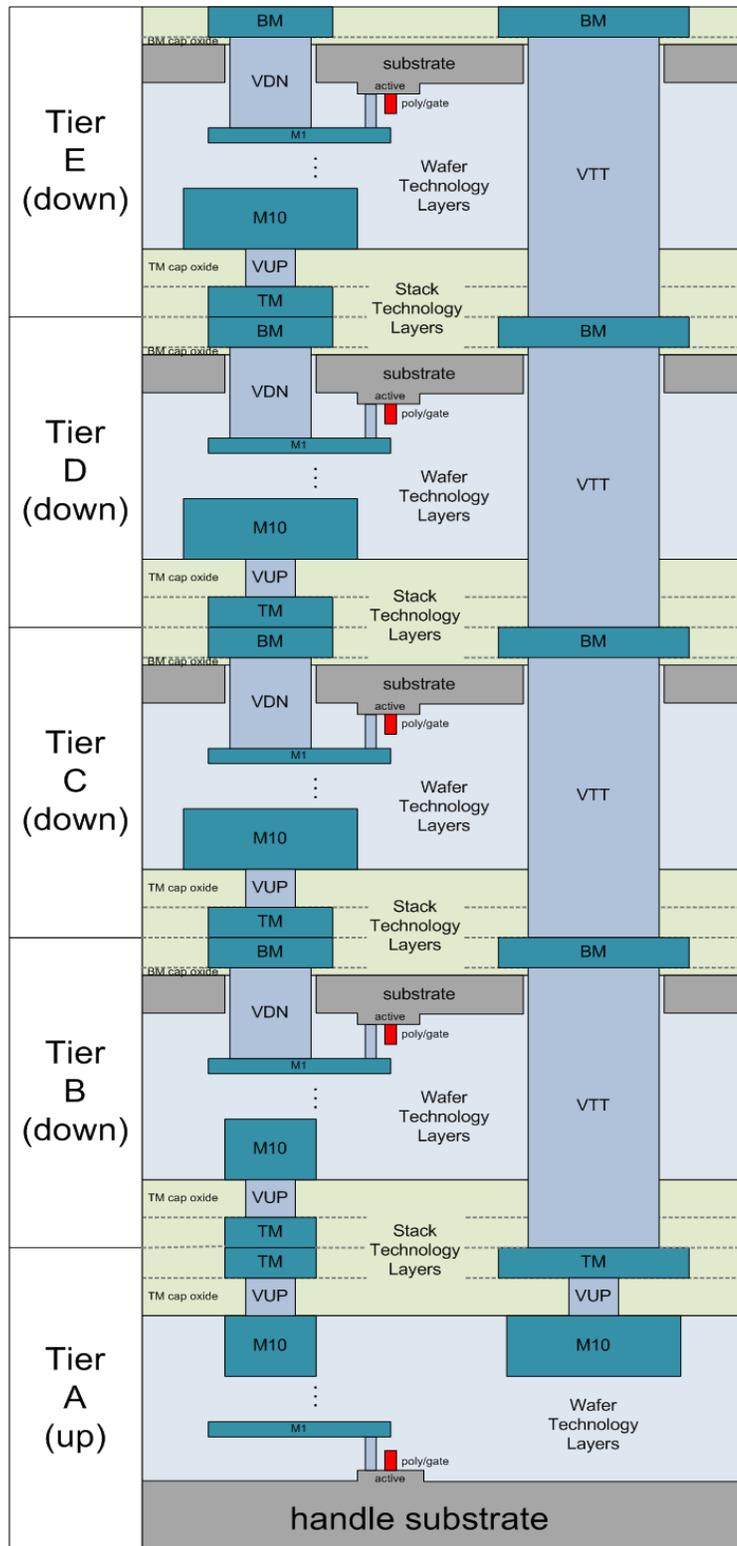
of transistors, regular vias and “special cut” vias used to create TSVs and emulates 3D design flow for a five-tier stack. While implementation of V-S in 3D-IC is beyond the scope of this dissertation, an insight into how TSV are used to bond the wafers and where to place the SCVR can be useful for any future implementation of V-S.

#### **5.3.5.1. 3D-IC Through-Silicon-Via**

Wafer or die bonding is one of the most critical modules to the 3D-IC industry. The bonded IC or wafer can be bonded with the IC side either “face down” or “face up” to a “face up” or “face down” lower die or wafer. The metal cross-section diagram (Figure 5.11) illustrates how a TSV from tier A (or tier level 1) to tier B (tier level 2) is created by assuming a “Top Metal” layer (TM) for each tier and a via from M10 to the top metal (using a special “up” via cut-layer called VUP). By bonding two top-metal shapes face-to-face, a connection is made. A back-metal (BM) is assumed to be patterned on the back of the substrate and special “down” via cut-layer (VDN) is used for back-metal connection. This way different layers in 3D-IC is connected through TSVs.

#### **(a) TSV Induced asymmetry in Voltage stacked 3D-IC**

Each of the push-pull SC converters can be partitioned into two (n) 2:1 switched-capacitor circuits, each placed on one of the 2 (n) layers. The fly-caps can swap their positions with the help of switches to maintain regulation. However, since TSV connect all the layers and each TSV adds considerable parasitic resistance, the additional IR drop between the top and bottom half will make this arrangement asymmetric. One way to solve this asymmetry is to further split each of the halves into two blocks and place one half of each on each of the layers. This will balance out the extra TSV loss on the lower layer.



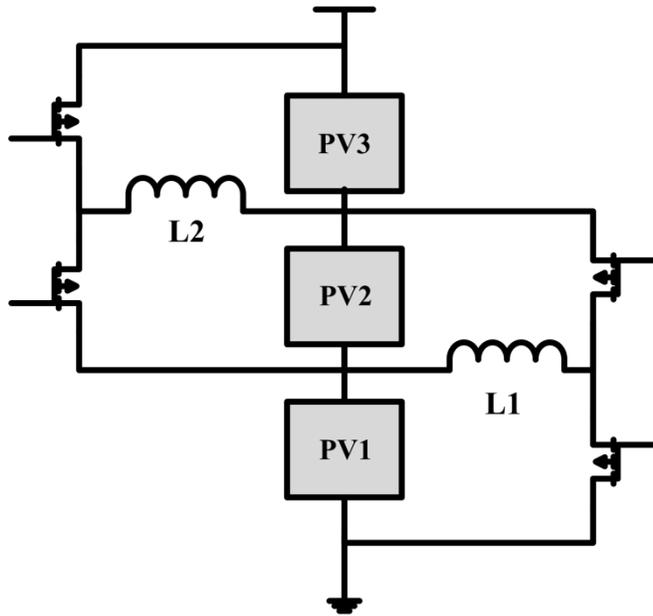
**Figure 5.11. Illustrates the tier and metal stack-up of the 5-tier 3D-IC version from NCSU EDA (104)**

## 5.4. Future work and summary

One of the great promises of 3D-IC is the possibility of having different layers with different technologies (e.g. CMOS for digital vs. GaAs for analog), different functionality (processor vs. memory), etc. This raises many issues about voltage stacking schemes since the heterogeneity in the stack means there are fewer opportunities for implicit regulation. There are fundamentally two types of solutions for this type of problem, using the present SCVR scheme: in case of extreme heterogeneity, voltage stacking can be used only on a subset of the layers in the stack; or hybrid series/parallel voltage stacking of layers can be used in order to improve the balance of currents and voltages. For example, several memory layers can still be connected conventionally in parallel since they are expected to consume less power, then they can be voltage stacked to a processor layer that would have natively a higher power consumption. Similarly, cache-memory and core can be connected using conventional high-density TSVs within the same voltage domain, while these voltage domains can be stacked together. This approach minimizes level-shifting overhead, as cache-to-core communication is within the same voltage domain and only core-to-core communication is across different voltage domains. Another possibility is to connect layers across more than just one rung in the switched-capacitor ladder, in case they need higher voltages – that might be the case for an analog circuit layer. However, since capacitors act here as charge-equalizers, i.e. they only redistribute the charge imbalance without changing the voltage conversion ratio, heterogeneity in terms of different voltages will not be possible. The inductive converter, on the other hand, offers to make V-S feasible in a truly heterogeneous 3D-IC system. A bidirectional

buck/boost converter can be used for differential power processing, as shown in Figure 5.12. However, integrated implementation of a many-output inductive converter may not be economically feasible.

3D-IC provides an essential mechanism for the industry to stay on the historical scaling trend of device integration, while raises fundamental challenges for reliable power delivery. Charge-recycled power regulation using series-connected architecture has been proposed and exhaustively studied as an alternative to conventional “parallel” PDN. A system-level PDN model for 3D-ICs has been designed to study charge-recycled, voltage-stacking PDN structure and compare it with the regular, non-voltage-stacked PDNs in the context of 3D-IC. Under the average workload imbalance ratio extracted from full applications (~65%), a V-S PDN's IR drop is no greater than 0.75% V<sub>dd</sub> beyond the noise level of a regular PDN, while system efficiency benefits from implicit recycling in more balanced cases. Combined with the observation that supply current, C4 bumps characteristics (count, EM-induced lifetime) and IR drop of V-S PDNs are insensitive to many-layer 3D-ICs' layer count, this study demonstrates that V-S provides a scalable and practical solution to the power delivery challenge in the era of the many-layer 3D-IC.



**Figure 5.12. Bidirectional buck-boost converter for differential power regulation in V-S**

#### **5.4.1.1. *Individual contributions in collaborative project***

This work is done in collaboration with Runjie Zhang, a Ph.D. student from the Computer Engineering department at the University of Virginia. Runjie’s major contributions include: 1) extending VoltSpot to support 3D-IC; 2) integrating the SC converter model with VoltSpot; and 3) whole-system PDN simulation and noise characterization

Author’s contributions include:

1) proposing 3D-MOSC to break the power delivery walls; 2) the circuit-level implementation of the SC converter; 3) the IVR models, resistive and transient and validation against circuit simulations; and 4) V-S system efficiency and voltage noise characterization with workload imbalance.

## Chapter 6

---

### 6.1. Dissertation Summary

In this dissertation, the focus has been on improving the performances of IVR, both linear and switching regulators, beyond state-of-the-art, with novel circuits and architectures targeted especially toward nano-scale CMOS nodes.

This work proposes and implements digitally-controlled LDOs catering to a wide range of applications, from graphics core to energy harvesting systems, with improved transient response and ultra-low quiescent current. Voltage droop mitigation by 50% allows 75% improvement in core frequency for the same voltage margin or else reduction in voltage guard-band leading to energy savings. Similarly, quiescent current reduction of LDO, especially in standby-mode, allows energy-harvested system to extend battery lifetime in idle-mode. Digitally-assisted solutions further lend easy scalability across technologies and allow low voltage operation.

Conventional switching regulators suffer from poor quality of integrated passives and parasitic losses, limiting the power efficiency and density of these IVRs. This work proposes charge-recycled power regulation with push-pull switched-capacitor to extend the performance metrics beyond the saturated limits of existing switching regulators. Implicit charge-recycling accompanied by differential regulation exhibits power efficiency of more than 90% for a wide range of load current with significantly higher

power density. While the primary focus of the work here has been on SCVR, the fact that bidirectional inductive converters fit perfectly with this differential regulation, makes V-S even more attractive

While IVR in 2D-IC has been extensively researched in literature, power regulation and power delivery in 3D-IC are limited by the more fundamental “3D versus 2D power wall”. This work explores the idea of multi-output switched-capacitor-assisted voltage stacking to provide an elegant, scalable and practical solution in bringing down this power wall. V-S, being an orthogonal approach to traditional PDN, involves an altogether new set of design tradeoffs which has never been explored before. In this work, a PDN model of many-layer 3D-IC is developed (under joint research effort with Zhang (88)) to study the impact of the series-stacked regulation on PDN quality, characterizing system efficiency and voltage noise with workload distribution. Transient simulation shows that, compared with the traditional PDN scheme, V-S provides stronger isolation for cross-layer noise interference in 3D-IC. With the same die area overhead for integrated capacitors, V-S PDNs provide up to 60% lower transient noise under the most noise-incurring workload behaviors. SC converter-assisted V-S coupled with workload scheduling can ensure high system efficiency (more than 90%) for a wide range of load distribution through implicit charge-recycling. V-S offers a unique tradeoff by significantly improving the EM-lifetime of C4 and TSV array (e.g., up to 5x), while only marginally increasing the average-case voltage noise (e.g., 0.75%  $V_{dd}$  IR drop), and thereby provides a scalable solution for many-layer 3D-IC's power delivery challenge (88).

## 6.2. Potential future work

### 6.2.1. Digitally-controlled LDO

- Develop a discrete domain model of controller to study stability under dynamic load conditions
- Replace PMOS switches with NMOS if  $V_{IN}$  to  $V_{OUT}$  ratio is large for better transient response
- Improve upon the single bit-shifting strategy of the controller
- Synthesizable D-LDO for ease of implementation and distribution in a digitally configurable power management-unit

### 6.2.2. Charge-recycled power regulation with integrated switched-capacitor

- Efficient level-shifter design for communication between voltage domains
- Extend V-S with bidirectional inductive regulation

### 6.2.3. Breaking 3D-IC power delivery wall using voltage stacking

- Fabrication of voltage stacked 3D-IC functional cores (MITLL, Terazzon foundries)
- Build truly heterogeneous system with inductive converters and V-S to provide variable voltage at different layers

- Develop efficient workload scheduling to maintain high system efficiency through implicit charge-recycling

## Glossary

---

PDN	Power-delivery-network
MOSC	Multi-output switched-capacitor
3D-MOSC	Multi-output switched-capacitor in 3D-IC
C4	Controlled-collapse-chip-connection
SCVR	Switched-capacitor voltage regulator
LDO	Low-dropout regulator
Decap	Decoupling capacitor
DVFS	Dynamic voltage frequency scaling
IVR	Integrated voltage regulation
V-S	Voltage stacking
VRM	Voltage regulator modules
SoC	System-on-chip

PMU	Power-management-unit
FOM	Figure-of-merit
PWM	Pulse width modulation
PFM	Pulse frequency modulation

## Bibliography

---

1. **Abedinpour S, Bakkaloglu B, Kiaei S.** A Multistage Interleaved Synchronous Buck Converter With Integrated Output Filter in 0.18 $\mu$ m SiGe Process. *IEEE Trans Power Electron* 22: 2164–2175, 2007.
2. **Amberg P, Liu F, Dayringer M, Lexau J, Patil D, Gainsley J, Moghadam HF, Alon E, Zheng X, Cunningham JE, Krishnamoorthy AV, Ho R.** Digitally-assisted analog circuits for a 10 Gbps, 395 fJ/b optical receiver in 40 nm CMOS. In: *Solid State Circuits Conference (A-SSCC)*. p. 29–32.
3. **Andersen TM, Krismer F, Kolar JW, Toifl T, Menolfi C, Kull L, Morf T, Kossel M, Brandli M, Buchmann P, Francese PA.** A 4.6W/mm<sup>2</sup> power density 86% efficiency on-chip switched capacitor DC-DC converter in 32 nm SOI CMOS. In: *2013 Twenty-Eighth Annual IEEE Applied Power Electronics Conference and Exposition (APEC)*. p. 692–699.
4. **Bergveld HJ, Karadi R, Nowak K.** An inductive down converter system-in-package for integrated power management in battery-powered applications. In: *IEEE Power Electronics Specialists Conference, 2008. PESC 2008*. p. 3335–3341.
5. **Bergveld HJ, Nowak K, Karadi R, Iochem S, Ferreira J, Ledain S, Pieraerts E, Pommier M.** A 65-nm-CMOS 100-MHz 87%-efficient DC-DC down converter based on dual-die system-in-package integration. In: *IEEE Energy Conversion Congress and Exposition, 2009. ECCE 2009*. p. 3698–3705.
6. **Bhattacharyya K, Mandal P.** A Low Voltage, Low Ripple, on Chip, Dual Switch-Capacitor Based Hybrid DC-DC Converter. In: *21st International Conference on VLSI Design, 2008. VLSID 2008*. p. 661–666.
7. **Van Breussegem TM, Steyaert MSJ.** Monolithic Capacitive DC-DC Converter With Single Boundary-Multiphase Control and Voltage Domain Stacking in 90 nm CMOS. *IEEE Journal of Solid-State Circuits* 46: 1715–1727, 2011.
8. **Burton EA, Schrom G, Paillet F, Douglas J, Lambert WJ, Radhakrishnan K, Hill MJ.** FIVR: Fully-integrated voltage regulators on 4th generation Intel ®

- Core<sup>TM</sup> SoCs. In: *2014 Twenty-Ninth Annual IEEE Applied Power Electronics Conference and Exposition (APEC)*. 2014, p. 432–439.
9. **Carlo S, Yueh W, Mukhopadhyay S.** On the potential of 3D integration of inductive DC-DC converter for high-performance power delivery. In: *2013 50th ACM/EDAC/IEEE Design Automation Conference (DAC)*. 2013, p. 1–8.
  10. **Chang L, Montoye RK, Ji BL, Weger AJ, Stawiasz KG, Dennard RH.** A fully-integrated switched-capacitor 2:1 voltage converter with regulation capability and 90% efficiency at 2.3A/mm<sup>2</sup>. In: *2010 IEEE Symposium on VLSI Circuits (VLSIC)*. p. 55–56.
  11. **Chu Y-C, Chang-Chien L-R.** Digitally Controlled Low-Dropout Regulator with Fast-Transient and Autotuning Algorithms. *IEEE Transactions on Power Electronics* 28: 4308–4317, 2013.
  12. **El-Damak D, Bandyopadhyay S, Chandrakasan AP.** A 93% efficiency reconfigurable switched-capacitor DC-DC converter using on-chip ferroelectric capacitors. In: *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2013*, p. 374–375.
  13. **El-Nozahi M, Amer A, Torres J, Entesari K, Sanchez-Sinencio E.** A 25mA 0.13um CMOS LDO regulator with power-supply rejection better than -56dB up to 10MHz using a feedforward ripple-cancellation technique. In: *Solid-State Circuits Conference - Digest of Technical Papers, 2009. ISSCC 2009*. p. 330–331,331a.
  14. **Esmaeili SE, Al-Kahlili AJ.** Integrated Power and Clock Distribution Network. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 21: 1941–1945, 2013.
  15. **Faust GG, Zhang R, Skadron K, Stan MR, Meyer BH.** ArchFP: Rapid prototyping of pre-RTL floorplans. In: *2012 IEEE/IFIP 20th International Conference on VLSI and System-on-Chip (VLSI-SoC)*. 2012, p. 183–188.
  16. **Fluhr E, Fluhr E, Polley M, Yang S-H, Erraguntla V, Noll T, van Berkel K.** F3: Adaptive design techniques for energy efficiency. In: *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014*, p. 514–515.
  17. **Gjanci J.** On-Chip Voltage Regulation for Power Management in System-on-Chip. University of illinois: 2006.
  18. **Gu J, Kim CH.** Multi-story power delivery for supply noise reduction and low voltage operation. In: *Proceedings of the 2005 International Symposium on Low Power Electronics and Design, ISLPED '05*, p. 192–197.
  19. **Hasib OA-T, Sawan M, Savaria Y.** Fully-integrated ultra-low-power asynchronously driven step-down DC-DC converter. In: *Proceedings of 2010*

- IEEE International Symposium on Circuits and Systems (ISCAS)*. 2010, p. 877–880.
20. **Hazucha P, Karnik T, Bloechel B, Parsons C, Finan D, Borkar S.** An area-efficient, integrated, linear regulator with ultra-fast load regulation. In: *2004 Symposium on VLSI Circuits, 2004. Digest of Technical Papers*. 2004, p. 218–221.
  21. **Hazucha P, Schrom G, Hahn J, Bloechel BA, Hack P, Dermer GE, Narendra S, Gardner D, Karnik T, De V, Borkar S.** A 233-MHz 80%-87% efficient four-phase DC-DC converter utilizing air-core inductors on package. *IEEE Journal of Solid-State Circuits* 40: 838–845, 2005.
  22. **Hsieh W-C, Hwang W.** Low quiescent current variable output digital controlled voltage regulator. In: *Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS)*. 2010, p. 609–612.
  23. **Hu J, Ismail M.** CMOS High Efficiency On-chip Power Management [Online]. Springer New York. <http://link.springer.com/10.1007/978-1-4419-9526-1> [11 May 2015].
  24. **Jain P, Kim T-H, Keane J, Kim CH.** A multi-story power delivery technique for 3D integrated circuits. In: *2008 ACM/IEEE International Symposium on Low Power Electronics and Design (ISLPED)*. 2008, p. 57–62.
  25. **Jain R, Geuskens BM, Kim ST, Khellah MM, Kulkarni J, Tschanz JW, De V.** A 0.45-1V Fully-Integrated Distributed Switched Capacitor DC-DC Converter With High Density MIM Capacitor in 22 nm Tri-Gate CMOS. *IEEE Journal of Solid-State Circuits* 49: 917–927, 2014.
  26. **Jain R, Sanders S.** A 200mA switched capacitor voltage regulator on 32nm CMOS and regulation schemes to enable DVFS. In: *Proceedings of the 2011-14th European Conference on Power Electronics and Applications (EPE 2011)*. 2011, p. 1–10.
  27. **Jeddeloh J, Keeth B.** Hybrid memory cube new DRAM architecture increases density and performance. In: *2012 Symposium on VLSI Technology (VLSIT)*. 2012, p. 87–88.
  28. **Jeon H.** Fully-integrated On-Chip Switched Capacitor DC-DC Converters for Battery-Powered Mixed-Signal SoCs. Department of Electrical and Computer Engineering, Northeastern University: 2012.
  29. **Jiang X, Jiang X, Malcovati P, Stojanovic V, Lizuka T.** F1: Digitally assisted analog and analog-assisted digital in high-performance scaled CMOS process. In: *Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, 2014, p. 510–511.

30. **Jiang X, Ramachandran NP, Kang DW, Chen CK, Rutherford M, Cong Y, Chang D.** Digitally-assisted analog and analog-assisted digital design techniques for a 28 nm mobile System-on-Chip. In: *European Solid State Circuits Conference (ESSCIRC), ESSCIRC 2014*, p. 475–478.
31. **Kanev S.** Motivating Software-Driven Current Balancing in Flexible Voltage-Stacked Multicore Processors. Computer Science, Harvard University: 2012.
32. **Katti G, Stucchi M, de Meyer K, Dehaene W.** Electrical Modeling and Characterization of Through Silicon via for Three-Dimensional ICs. *IEEE Transactions on Electron Devices* 57: 256–262, 2010.
33. **Khan NH, Alam SM, Hassoun S.** Power Delivery Design for 3-D ICs Using Different Through-Silicon Via (TSV) Technologies. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 19: 647–658, 2011.
34. **Kim ST, Shih Y-C, Mazumdar K, Jain R, Ryan JF, Tokunaga C, Augustine C, Kulkarni JP, Ravichandran K, Tschanz JW, Khellah MM, De V.** 8.6 Enabling wide autonomous DVFS in a 22nm graphics execution core using a digitally controlled hybrid LDO/switched-capacitor VR with fast droop mitigation. In: *Solid-State Circuits Conference - (ISSCC), 2015*, 2015, p. 1–3.
35. **Kim W, Gupta MS, Wei G-Y, Brooks D.** System level analysis of fast, per-core DVFS using on-chip switching regulators. In: *IEEE 14th International Symposium on High Performance Computer Architecture, 2008. HPCA 2008*. 2008, p. 123–134.
36. **Kim Y, Li P.** An ultra-low voltage digitally controlled low-dropout regulator with digital background calibration. In: *2012 13th International Symposium on Quality Electronic Design (ISQED)*. 2012, p. 151–158.
37. **Krishnamurthy HK, Vaidya VA, Kumar P, Matthew GE, Weng S, Thiruvengadam B, Proefrock W, Ravichandran K, De V.** A 500 MHz, 68% efficient, fully on-die digitally controlled buck Voltage Regulator on 22nm Tri-Gate CMOS. In: *2014 Symposium on VLSI Circuits Digest of Technical Papers*. 2014, p. 1–2.
38. **Kudva S.** High efficiency, low cost, fully-integrated DC-DC converter solution. university of minnesota: 2013.
39. **Kudva SS, Harjani R.** Fully-Integrated On-Chip DC-DC Converter With a 450X Output Range. *IEEE Journal of Solid-State Circuits* 46: 1940–1951, 2011.
40. **Kwong J, Ramadass YK, Verma N, Chandrakasan AP.** A 65 nm Sub-Microcontroller With Integrated SRAM and Switched Capacitor DC-DC Converter. *IEEE Journal of Solid-State Circuits* 44: 115–126, 2009.

41. **Lee SK, Brooks D, Wei G-Y.** Evaluation of Voltage Stacking for Near-threshold Multicore Computing. In: *Proceedings of the 2012 ACM/IEEE International Symposium on Low Power Electronics and Design*. ACM, p. 373–378.
42. **Lefurgy CR, Drake AJ, Floyd MS, Allen-Ware MS, Brock B, Tierno JA, Carter JB, Berry RW.** Active Guardband Management in Power7 to Save Energy and Maintain Reliability. *IEEE Micro* 33: 35–45, 2013.
43. **Le H-P, Sanders SR, Alon E.** Design Techniques for Fully-integrated Switched-Capacitor DC-DC Converters. *IEEE Journal of Solid-State Circuits* 46: 2120–2131, 2011.
44. **Li S, Ahn JH, Strong RD, Brockman JB, Tullsen DM, Jouppi NP.** The McPAT Framework for Multicore and Manycore Architectures: Simultaneously Modeling Power, Area, and Timing. *ACM Trans Archit Code Optim* 10: 5:1–5:29, 2013.
45. **Liu Y, Hsieh P-H, Kim S, Seo J, Montoye R, Chang L, Tierno J, Friedman D.** A 0.1pJ/b 5-to-10Gb/s charge-recycling stacked low-power I/O for on-chip signaling in 45nm CMOS SOI. In: *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2013*. 2013, p. 400–401.
46. **Luders M, Eversmann B, Gerber J, Huber K, Kuhn R, Schmitt-Landsiedel D, Brederlow R.** A fully-integrated system power aware LDO for energy harvesting applications. In: *2011 Symposium on VLSI Circuits (VLSIC)*. 2011, p. 244–245.
47. **Lu Y, Ki W-H, Yue CP.** 17.11 A 0.65ns-response-time 3.01ps FOM fully-integrated low-dropout regulator with full-spectrum power-supply-rejection for wideband communication systems. In: *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014*. 2014, p. 306–307.
48. **Makipaa J, Billoint O.** FDSOI versus BULK CMOS at 28 nm node which technology for ultra-low power design? In: *2013 IEEE International Symposium on Circuits and Systems (ISCAS)*. 2013, p. 554–557.
49. **Mazumdar K, Bartling S, Khanna S, Stan M.** A digitally-controlled power-aware low-dropout regulator to reduce standby current drain in ultra-low-power MCU. In: *2015 16th International Symposium on Quality Electronic Design (ISQED)*. 2015, p. 98–102.
50. **Mazumdar K, Stan M.** Breaking the 3D IC power delivery wall. In: *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. 2012, p. 741–746.
51. **Mazumdar K, Stan MR.** Charge recycling on-chip DC-DC conversion for near-threshold operation. In: *2012 IEEE Subthreshold Microelectronics Conference (SubVT)*. 2012, p. 1–3.

52. **Milev M, Burt R.** A Tool and Methodology for AC-Stability Analysis of Continuous-Time Closed-Loop Systems. In: *Proceedings of the Conference on Design, Automation and Test in Europe - Volume 3*. IEEE Computer Society, p. 204–208.
53. **Murmann B.** Digitally Assisted Analog Circuits. *IEEE Micro* 26: 38–47, 2006.
54. **Bin Nasir S, Lee Y, Raychowdhury A.** Modeling and analysis of system stability in a distributed power delivery network with embedded digital linear regulators. In: *2014 15th International Symposium on Quality Electronic Design (ISQED)*. 2014, p. 68–75.
55. **Nguyen HV, Ryu M, Kim Y.** Performance and power analysis of through silicon via based 3D IC integration. In: *2011 13th International Workshop on System Level Interconnect Prediction (SLIP)*. 2011, p. 1–1.
56. **Ni J, Hong Z, Liu BY.** Improved on-chip components for integrated DC-DC converters in 0.13 $\mu$ m CMOS. In: *Proceedings of ESSCIRC, 2009*. 2009, p. 448–451.
57. **Okuma Y, Ishida K, Ryu Y, Zhang X, Chen P-H, Watanabe K, Takamiya M, Sakurai T.** 0.5-V input digital LDO with 98.7% current efficiency and 2.7 $\mu$ A quiescent current in 65nm CMOS. In: *2010 IEEE Custom Integrated Circuits Conference (CICC)*. 2010. p. 1–4.
58. **Onizuka K, Inagaki K, Kawaguchi H, Takamiya M, Sakurai T.** Stacked-Chip Implementation of On-Chip Buck Converter for Distributed Power Supply System in SiPs. *IEEE Journal of Solid-State Circuits* 42: 2404–2410, 2007.
59. **Onouchi M, Otsuga K, Igarashi Y, Ikeya T, Morita S, Ishibashi K, Yanagisawa K.** A 1.39-V input fast-transient-response digital LDO composed of low-voltage MOS transistors in 40-nm CMOS process. In: *Solid State Circuits Conference (A-SSCC), 2011*, 2011. p. 37–40.
60. **Pique GV.** A 41-phase switched-capacitor power converter with 3.8mV output ripple and 81% efficiency in baseline 90nm CMOS. In: *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2012*. 2012, p. 98–100.
61. **Pizzagalli A, Buisson T, Beica R.** 3D technology applications market trends and key challenges. In: *Advanced Semiconductor Manufacturing Conference (ASMC), 2014 25th Annual SEMI*. Advanced Semiconductor Manufacturing Conference (ASMC), 2014, p. 78–81.
62. **Putic M, Di L, Calhoun BH, Lach J.** Panoptic DVS: A fine-grained dynamic voltage scaling framework for energy scalable CMOS design. In: *IEEE International Conference on Computer Design, 2009*. ICCD 2009. p. 491–497.

63. **Rajapandian S, Xu Z, Shepard KL.** Energy-efficient low-voltage operation of digital CMOS circuits through charge-recycling. In: *2004 Symposium on VLSI Circuits*, Digest of Technical Papers. 2004, p. 330–333.
64. **Rajapandian S, Xu Z, Shepard KL.** Implicit DC-DC downconversion through charge-recycling. *IEEE Journal of Solid-State Circuits* 40: 846–852, 2005.
65. **Ramadass YK, Chandrakasan AP.** Voltage Scalable Switched Capacitor DC-DC Converter for Ultra-Low-Power On-Chip Applications. In: *IEEE Power Electronics Specialists Conference, 2007. PESC 2007.* 2007, p. 2353–2359.
66. **Ramadass YK, Fayed AA, Chandrakasan AP.** A Fully-Integrated Switched-Capacitor Step-Down DC-DC Converter With Digital Capacitance Modulation in 45 nm CMOS. *IEEE Journal of Solid-State Circuits* 45: 2557–2565, 2010.
67. **Raychowdhury A, Somasekhar D, Tschanz J, De V.** A fully-digital phase-locked low dropout regulator in 32nm CMOS. In: *2012 Symposium on VLSI Circuits (VLSIC).* 2012, p. 148–149.
68. **Rojas-Gonzalez MA, Torres J, Sanchez-Sinencio E.** Design of a fully-integrated buck voltage regulator using standard CMOS technology. In: *2012 IEEE Third Latin American Symposium on Circuits and Systems (LASCAS).* 2012. p. 1–4.
69. **Sapatnekar SS.** Addressing thermal and power delivery bottlenecks in 3D circuits. In: *Design Automation Conference, Asia-Pacific, 2009. ASP-DAC 2009.* 2009, p. 423–428.
70. **Schrom G, Hazucha P, Hahn J-H, Kursun V, Gardner D, Narendra S, Karnik T, De V.** Feasibility of monolithic and 3D-stacked DC-DC converters for microprocessors in 90nm technology generation. In: *Proceedings of the 2004 International Symposium on Low Power Electronics and Design, 2004. ISLPED '04.* 2004, p. 263–268.
71. **Schrom G, Hazucha P, Paillet F, Rennie DJ, Moon ST, Gardner DS, Kamik T, Sun P, Nguyen TT, Hill MJ, Radhakrishnan K, Memioglu T.** A 100MHz Eight-Phase Buck Converter Delivering 12A in 25mm<sup>2</sup> Using Air-Core Inductors. In: *APEC 2007 - Twenty Second Annual IEEE Applied Power Electronics Conference.* APEC 2007. 2007, p. 727–730.
72. **Seeman MD.** A Design Methodology for Switched-Capacitor DC-DC Converters. EECS Department, University of California, Berkeley: 2009.
73. **Seeman MD, Ng VW, Le H-P, John M, Alon E, Sanders SR.** A comparative analysis of Switched-Capacitor and inductor-based DC-DC conversion technologies. In: *2010 IEEE 12th Workshop on Control and Modeling for Power Electronics (COMPEL).* 2010, 2010, p. 1–7.

74. **Seeman MD, Sanders SR, Rabaey JM.** An Ultra-Low-Power Power Management IC for Wireless Sensor Nodes. In: *IEEE Custom Integrated Circuits Conference, 2007. CICC '07.* 2007, p. 567–570.
75. **Shenoy PS, Kim KA, Johnson BB, Krein PT.** Differential Power Processing for Increased Energy Production and Reliability of Photovoltaic Systems. *IEEE Transactions on Power Electronics* 28: 2968–2979, 2013.
76. **Skadron K, Stan MR, Huang W, Velusamy S, Sankaranarayanan K, Tarjan D.** Temperature-aware microarchitecture. In: *30th Annual International Symposium on Computer Architecture, 2003.* Proceedings. 2003, p. 2–13.
77. **Stan M.** Breaking power delivery walls using voltage stacking. In: *2013 IEEE Computer Society Annual Symposium on VLSI (ISVLSI).* 2013, p. 212–212.
78. **Sun J, Giuliano D, Devarajan S, Lu J-Q, Chow TP, Gutmann RJ.** Fully Monolithic Cellular Buck Converter Design for 3-D Power Delivery. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 17: 447–451, 2009.
79. **Toprak-Deniz Z, Bulzacchelli J, Rasmus T, Iadanza J, Bucossi W, Kim S, Blanco R, Cox C, Chhabra M, Leblanc C, Trudeau C, Friedman D.** Dual-loop system of distributed microregulators with high DC accuracy, load response time below 500ps, and 85mV dropout voltage. In: *2011 Symposium on VLSI Circuits (VLSIC).* 2011, p. 274–275.
80. **Toprak-Deniz Z, Sperling M, Bulzacchelli J, Still G, Kruse R, Kim S, Boerstler D, Gloekler T, Robertazzi R, Stawiasz K, Diemoz T, English G, Hui D, Muench P, Friedrich J.** 5.2 Distributed system of digitally controlled microregulators enabling per-core DVFS for the POWER8™ microprocessor. In: *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014 IEEE International.* 2014, p. 98–99.
81. **Ueda K, Shunsuke O, Morishita F, Arimoto K, Okamura L, Yoshihara T.** Green semiconductor technology with ultra-low power on-chip charge-recycling power circuit and system. In: *Solid State Circuits Conference (A-SSCC), 2012 IEEE Asian.* 2012, p. 105–108.
82. **De Vos J, Flandre D, Bol D.** A dual-mode DC/DC converter for ultra-low-voltage microcontrollers. In: *2012 IEEE Subthreshold Microelectronics Conference (SubVT).* 2012, p. 1–3.
83. **Wang X, Xu J, Wang Z, Chen K, Wu X, Wang Z, Yang P, Duong H.** An Analytical Study of Power Delivery Systems for Many-Core Processors Using On-Chip and Off-Chip Voltage Regulators. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* PP: 1–1, 2015.
84. **Warnock J, Chan Y, Harrer H, Carey S, Salem G, Malone D, Puri R, Zitz JA, Jatkowski A, Strevig G, Datta A, Gattiker A, Bansal A, Mayer G, Chan Y-H,**

- Mayo M, Rude DL, Sigal L, Strach T, Smith HH, Wen H, Mak P-K, Shum C-LK, Plass D, Webb C.** Circuit and Physical Design of the zEnterprise #x2122; EC12 Microprocessor Chips and Multi-Chip Module. *IEEE Journal of Solid-State Circuits* 49: 9–18, 2014.
85. **Wens M, Steyaert M.** A fully-integrated 0.18 $\mu$ m CMOS DC-DC step-down converter, using a bondwire spiral inductor. In: *IEEE Custom Integrated Circuits Conference, 2008. CICC 2008*. p. 17–20.
86. **Wens M, Steyaert M.** An 800mW fully-integrated 130nm CMOS DC-DC step-down multi-phase converter, with on-chip spiral inductors and capacitors. In: *IEEE Energy Conversion Congress and Exposition, 2009*. 2009, p. 3706–3709.
87. **Wibben J, Harjani R.** A High-Efficiency DC #x2013;DC Converter Using 2 nH Integrated Inductors. *IEEE Journal of Solid-State Circuits* 43: 844–854, 2008.
88. **Zhang R.** Pre-RTL On-chip Power Delivery Modeling and Analysis. Dept of computer engineering, University of virginia: [date unknown].
89. **Zhang R, Mazumdar K.** A Cross-Layer Design Exploration of Charge-Recycled Power-Delivery in Many-Layer 3D-IC. DAC. 2015.
90. **Zhang R, Mazumdar K.** Transient Voltage Noise in Charge-Recycled Power Delivery Networks for Many-layer 3D-IC. ISLPED. 2015
91. **Zhang R, Wang K, Meyer BH, Stan MR, Skadron K.** Architecture implications of pads as a scarce resource. In: *2014 ACM/IEEE 41st International Symposium on Computer Architecture (ISCA)*. 2014, p. 373–384.
92. **Zhao Y, Yang Y, Mazumdar K, Guo X, Stan MR.** A multi-output on-chip switched-capacitor DC-DC converter for near- and sub-threshold power modes. In: *2014 IEEE International Symposium on Circuits and Systems (ISCAS)*. 2014, p. 1632–1635.
93. **Zhou P, Jiao D, Kim CH, Sapatnekar SS.** Exploration of on-chip switched-capacitor DC-DC converter for multicore processors using a distributed power delivery network. In: *2011 IEEE Custom Integrated Circuits Conference (CICC)*. 2011, p. 1–4.
94. Moore's law - Wikipedia, the free encyclopedia [Online]. [date unknown]. [http://en.wikipedia.org/wiki/Moore%27s\\_law](http://en.wikipedia.org/wiki/Moore%27s_law) [9 May 2015].
95. SoC vs. CPU – The battle for the future of computing | ExtremeTech [Online]. [date unknown]. <http://www.extremetech.com/computing/126235-soc-vs-cpu-the-battle-for-the-future-of-computing> [10 May 2015].

96. Intel® 22 nm Technology [Online]. [date unknown]. <http://www.intel.com/content/www/us/en/silicon-innovations/intel-22nm-technology.html> [9 May 2015].
97. Technology News | The big picture in technology [Online]. [date unknown]. <http://14nm.com/> [9 May 2015].
98. 3D-ICs [Online]. [date unknown]. [http://www.process-evolution.com/3d-ics\\_doe.html](http://www.process-evolution.com/3d-ics_doe.html) [13 May 2015].
99. Voltage doubler - Wikipedia, the free encyclopedia [Online]. [date unknown]. [http://en.wikipedia.org/wiki/Voltage\\_doubler](http://en.wikipedia.org/wiki/Voltage_doubler) [11 May 2015].
100. The PARSEC Benchmark Suite [Online]. [date unknown]. <http://parsec.cs.princeton.edu/> [11 May 2015].
101. Cortex-A9 Processor - ARM [Online]. [date unknown]. <http://www.arm.com/products/processors/cortex-a/cortex-a9.php> [11 May 2015].
102. Micron Technology, Inc. - Hybrid Memory Cube [Online]. [date unknown]. <http://www.micron.com/products/hybrid-memory-cube> [11 May 2015].
103. gem5 [Online]. [date unknown]. [http://gem5.org/Main\\_Page](http://gem5.org/Main_Page) [11 May 2015].
104. FreePDK3D45:Contents - NCSU EDA Wiki [Online]. [date unknown]. <http://www.eda.ncsu.edu/wiki/FreePDK3D45:Contents> [11 May 2015].