Power Management for Multicore Processors via Heterogeneous Voltage Regulation and Machine Learning Enabled Adaptation

Xin Zhan, Jianhao Chen, Edgar Sánchez-Sinencio, *Fellow, IEEE*, and Peng Li, *Fellow, IEEE*

***Abstract*—This work is based on the vision that the ultimate power integrity and efficiency may be best achieved via a heterogeneous chain of voltage processing starting from onboard switching voltage regulators (VRs), to on-chip switching VRs, and finally to networks of distributed on-chip linear VRs. As such, we propose a heterogeneous voltage regulation (HVR) architecture encompassing regulators with complimentary characteristics in response time, size, and efficiency. By exploring the rich heterogeneity and tunability in HVR, we develop systematic workload-aware power management policies to adapt heterogeneous VRs with respect to workload change at multiple temporal scales to significantly improve system power efficiency while providing a guarantee for power integrity. The proposed techniques are further supported by hardware-accelerated machine learning (ML) prediction of nonuniform spatial workload distributions for more accurate HVR adaptation at fine time granularity. Our evaluations based on the PARSEC benchmark suite show that the proposed adaptive three-stage HVR reduces the total system energy dissipation by up to 23.9% and 15.7% on average compared with the conventional static two-stage voltage regulation using off-chip and on-chip switching VRs. Compared with the three-stage static HVR, our runtime control reduces system energy by up to 17.9% and 12.2% on average. Furthermore, the proposed ML prediction offers up to 4.1% reduction of system energy.**

***Index Terms*—Machine learning (ML), multicore processor, power delivery network (PDN), power management, voltage regulation.**

I. Introduction

SUPPLY voltage regulation serves the critical role of delivering power to on-die devices for high-performance VLSI systems such as in server and desktop applications [1]–[3]. Power shall be delivered with ensured power quality to prevent timing violations. On the other hand, achieving power efficiency has become a key challenge in the dark silicon age [4]. Power management must be employed to maximize power efficiency in every possible way [5], [6].

Power delivery networks (PDNs) and voltage regulators (VRs) significantly impact power efficiency and integrity. Switching VRs are more efficient than linear VRs such as low-dropout VRs [low dropout regulators (LDOs)] over wide output voltage and load ranges, while linear VRs are more area efficient and can achieve faster subnanosecond response times [7]–[9]. Distributed on-chip voltage regulation is an important ongoing design trend where multiple area-efficient linear VRs are distributed within a power domain to provide fast suppression of power supply noise in the vicinity of such linear VRs. For instance, the recent IBM POWER8 processor employs 1764 on-chip distributed linear VRs [2].

Tradeoffs between performance and power dissipation can be optimized using dynamic power management such as dynamic voltage and frequency scaling (DVFS) [3], [10]–[15]. However, dynamic workloads and power management may push the VRs away from their optimal operating points, degrading the efficiency of the entire system. Recent work has attempted to reconfigure the PDN based on the workload [16]–[19]. As an example, the workload-aware quantized power management (QPM) scheme in [17] adopts simple control policies to dynamically adjust the number of active on-chip and off-chip switching VRs. However, such schemes have only considered switching VRs and little work has been done toward holistic exploration of heterogeneous VRs and their systematic adaptation considering complex interdependencies between such regulators.

This work is based on the vision that the ultimate power quality and efficiency may be best achieved via a heterogeneous chain of voltage processing starting from on-board switching VRs, to on-chip switching VRs, and finally to networks of distributed on-chip linear VRs. As depicted in Fig. 1, we propose a heterogeneous voltage regulation (HVR) architecture encompassing regulators with complimentary characteristics in response time, size, and efficiency. This work aims to answer the following key question for the first time. Given a desired power supply voltage set by a higher level power management policy, e.g., one based on DVFS, for each power domain, how shall the VRs in the HVR system be adapted autonomously with respect to workload change at multiple temporal scales to significantly improve system power efficiency while providing a guarantee for power integrity? The contributions of this paper are severalfolds.

Manuscript received January 30, 2019; revised May 19, 2019; accepted June 8, 2019. Date of publication July 11, 2019; date of current version October 23, 2019. This work was supported in part by the National Science Foundation under Grant ECCS-1810125 and in part by the Qatar National Research Fund (a member of Qatar Foundation) through NPRP under Grant NPRP 8-274-2-107. *(Corresponding author: Xin Zhan.)*

The authors are with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843 USA (e-mail: zhanxin@tamu.edu; chenjh@tamu.edu; s-sanchez@tamu.edu; pli@tamu.edu).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVLSI.2019.2923911

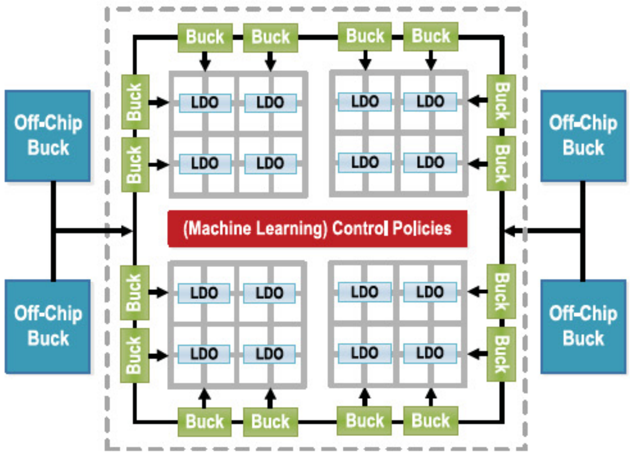


Fig. 1. Proposed HVR.

1) This is the first work that systematically explores HVR. The most general form of HVR consists of VRs with complimentary characteristics across three processing stages. In the first two stages, off-chip and on-chip switching (dc–dc) converters are employed to achieve high efficiency over a wide output voltage range, serving the major role of voltage conversion. Compared with single stage dc–dc conversion, two-stage dc–dc conversion allows for area reduction, improved power efficiency, and fine-gained DVFS, which is supported by the fast response time of on-chip dc–dc converters. Unlike conventional one or two-stage PDNs, HVR largely decouples *voltage conversion* from *voltage regulation*, the latter of which is optimally achieved by placing a large number of compact LDOs with subnanosecond response time in a distributed manner within each power domain, forming an interconnected active regulation network.
2) We propose systematic workload-aware control policies to jointly optimize power efficiencies of all voltage processing stages to maximize the overall system power efficiency. To best exploit the potential of energy efficiency of HVR, our control policies minimize system power losses by considering interdependencies across the entire voltage processing chain and adapt HVR at multiple time scales given the significantly different response times of the considered VRs.
3) Uncertainties caused by unknown nonuniform spatial distribution of the workload are hard to predict but can jeopardize power integrity. To minimize the extra voltage margin, hence power loss, needed for accounting for nonuniform spatial workload distribution, for the first time, we propose a novel machine learning (ML) solution that accurately sets the output voltage of the on-chip switching VRs to maximize the system power efficiency while effectively tracking the worst case voltage drop in each power domain to safeguard power integrity. Our ML solution consists of a few on-chip voltage-noise sensors that provide inputs to a low-overhead hardware-accelerated ML predictor, which fine tunes the output voltage of the on-chip switching VRs. This provides an autonomous end-to-end integrated ML solution whose low latency allows for fine-grained adaptation of HVR.

II. Motivation of Heterogeneous Voltage Regulation

A. Overview of Voltage Regulators

VRs are key components of a power delivery system and the characteristics of VRs have critical impacts on power efficiency and performance of the entire system. Generally, linear VRs such as LDOs are more area efficient and can achieve fast response time, while switching VRs are usually more energy efficient. The inductor-based buck converter and the switched capacitance (SC) converter are the two main categories of switching VRs. The integration of SC converter requires only capacitance, which have a significantly higher power density and can be integrated more easily than its inductance counterpart [20]. However, it only supports certain discrete voltage divide ratios and usually needs a large number of phases for ripple loss reduction [21]. On the other hand, the enabled continuous and wide range of output voltages with high efficiency makes inductor-based buck converter a natural choice for dynamic voltage scaling (DVS), and therefore it has been used for most switch VRs for past decades. In this work, we use inductor-based buck converter as the switch VR to demonstrate the benefit of multistage voltage regulation. However, the adaptive control scheme with multistage voltage regulation as proposed later may also be applied to the system with capacitance-based converters in a similar way.

In a PDN, off-chip inductor-based buck VR, switching at a rate of hundreds of kilohertz to tens of megahertz, can achieve excellent efficiency at the expense of bulky and costly off-chip *LC* components [22], [23]. Furthermore, off-chip VRs have slow response times and, hence, cannot support fine-grained DVS. There has been a great deal of progress on fully integrated buck VRs, thanks to on-die/in-package inductors and new magnetic materials [24]–[26]. Operating at a frequency of tens or hundreds of megahertz, fully integrated buck VRs come with fast response times and promises for efficient local power delivery and fine-grain DVFS. However, integrating high-Q power inductors to support high current density with low loss is still a major challenge [24]–[26]. Compared to their off-chip counterparts, on-chip buck VRs incur more conduction and switching losses, leading to lower efficiency, especially at light loads. On-chip linear (e.g., LDOs) are area efficient and can achieve subnanosecond response times [9]. Their efficiency drops with increasing dropout voltage, making them inefficient for wide-range voltage conversion. Clearly, those VRs have complimentary characteristics in response time, area and power efficiency and none of them can address the IC power delivery challenge alone.

Conversion Versus Regulation: Although conversion and regulation are used almost interchangeably, we shall note a fine distinction between them with respect to the best ways for realizing conversion and regulation. Switching VRs are well suited for wide-range voltage conversion for which linear VRs suffer from large loss. On the other hand, area-efficient

TABLE I
COMPARISON OF DIFFERENT VRs

	Settling time	Area	Efficiency	Function
Off-chip buck	10's of us [22]	Large	High	Conversion
On-chip buck	10's of ns [26]	Medium	Medium	Conversion
On-chip LDO	sub-ns [9]	Small	Low w/ large $V_{in} - V_{out}$	Regulation

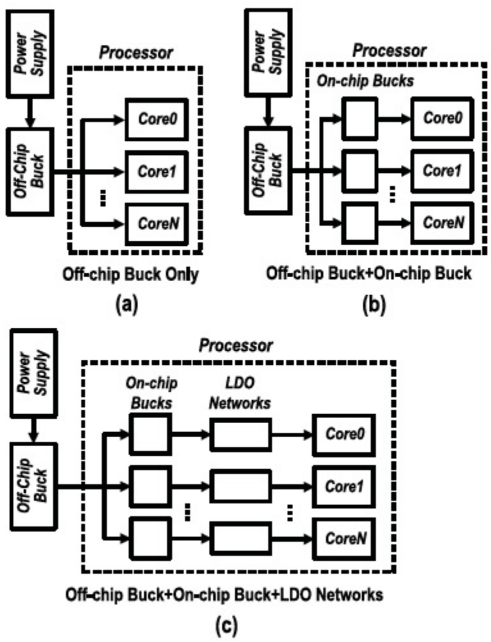


Fig. 2. PDN architectures. (a) Single-stage PDN using off-chip buck converters. (b) Two-stage PDN using both on-chip and off-chip buck converters. (c) Proposed three-stage HVR.

integrated linear VRs provide fast regulation. Table I summarizes the characteristics of different VRs.

B. Heterogeneous PDN Architecture

Three power delivery architectures are illustrated in Fig. 2. The single-stage PDN [Fig. 2(a)] is managed by only off-chip buck converters, achieving a high efficiency over a wide workload range. However, the board/package parasitics degrade the power quality delivered from the off-chip bucks to the on-chip power domains. Furthermore, the slow response time of off-chip buck converters limits the application of fine-grained DVS. Thanks to the progress of on-die/in-package inductors and new magnetic materials the buck converters can be integrated on chip. The two-stage PDN [Fig. 2(b)] consists of both off-chip and on-chip buck converters, improves the quality of power delivery by lowering the impedance from the power supply to the load circuits, and supports fine-grained per-core DVS since the integrated VRs can settle much faster. These benefits make this architecture widely used in modern SoCs such as the Intel's Haswell processors [27]. However, the response time of on-chip buck converters can still limit the PDN performance in the case of highly unpredictable load currents which may occur, for example, in server-class processors [14].

TABLE II
CONTROL VARIABLES IN HVR

Variable	Description
$N^{(i)}_{online,on}$	Number of online VRs in the i-th on-chip buck cluster
$V^{(i)}_{out,on}$	Output voltage of the i-th on-chip buck cluster
$N_{online,off}$	Number of online VRs in the off-chip buck cluster
$V_{out,off}$	Output voltage of the off-chip buck cluster

We argue that the ultimate quality and efficiency in supply voltage regulation may be only achieved by fully exploiting the heterogeneity in PDN architecture with heterogeneous regulators with complimentary characteristics in response time, power efficiency, and cost. As shown in Fig. 2(c), we propose an HVR architecture with three voltage processing stages: multiple off-chip buck VRs supplying power to multiple clusters of on-chip buck VRs with each cluster powering a network of distributed on-chip LDO driving a power/voltage domain. Fig. 3(a) depicts a more detailed view of the three-stage HVR. Clearly, the first stage enjoys high efficiencies of off-chip buck VRs over wide ranges of workloads. Their slow response is compensated by the second stage of on-chip buck VRs. Bypassing board/package parasitic impedances, on-chip buck VRs can settle much faster, enabling fine grained per core DVS otherwise impossible. Having two stages of buck converters gives the added benefit of lowering the step-down ratio for each stage, improving the efficiency of both off-chip and on-chip buck converters, and reducing sizes of the off-chip passives and power transistors [17]. Leaving most of the voltage conversion functionality to the first two stages, the on-chip LDO networks act as the last (main) stage of voltage regulation. Due to the small footprint of LDOs, a large number of compact LDOs with ultrafast response time can be placed on-chip in a distributed manner within a power domain, forming an interconnected active regulation network. In vicinity of on-chip hot spots, on-chip network can respond very quickly to local voltage droops, achieving good regulation performance.

C. Tuning Opportunities in HVR

Heterogeneity brings in a great deal of tunability at multiple HVR stages for workload-aware adaption. The power efficiency of a single VR stage is usually a function of its input–output voltages and current load. For a cluster of VRs, its power efficiency can be optimized according to runtime workload by either tuning its input–output voltages or modulating the number of online VRs, which changes the load per regulator. There are important interdependencies among different voltage processing stages which must be carefully considered in order to optimize the overall energy efficiency and regulation performance. For example, the output of the preceding VR stage is also the input of the subsequent VR stage. Fig. 4 summarizes the rich tunability and complicated energy and performance interdependencies in HVR system.

We define several important control variables in Table II, and will use them throughout this paper. Considering an HVR

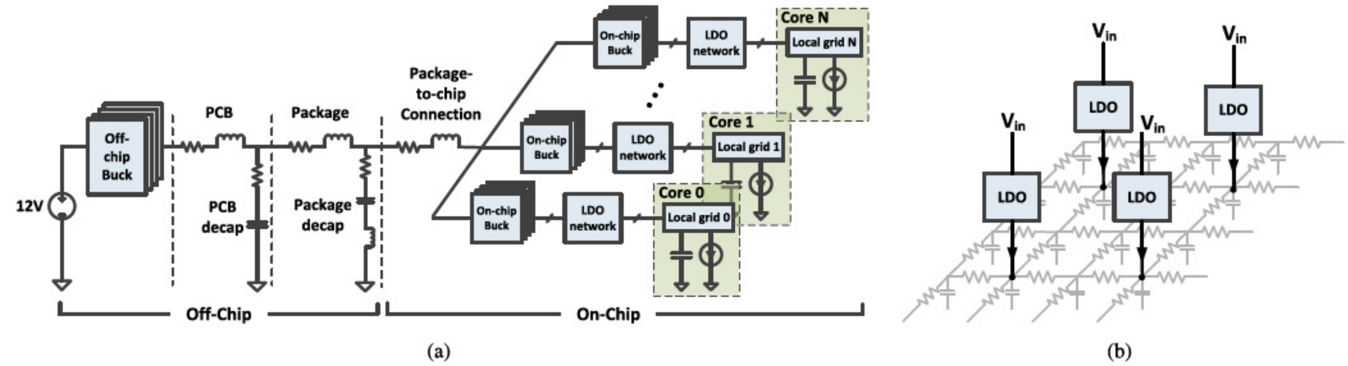


Fig. 3. (a) Modeling of three-stage HVR system. (b) Distributed LDO network.

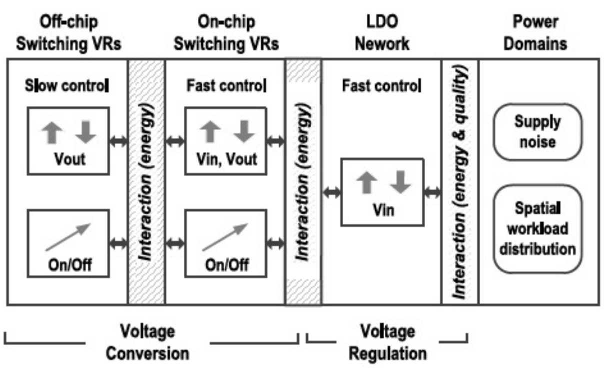


Fig. 4. Overview of tunability in HVR system.

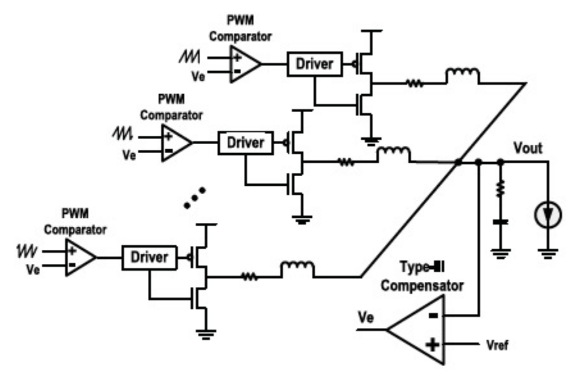


Fig. 5. Schematic of a multiphase PWM buck converter.

system consisting of N power domains as in Fig. 2(c), the control decision variables are: the number of online converters in each on-chip buck VR cluster $N_{\text{online,on}}^{(i)}$, and the cluster's output voltage $V_{\text{out,on}}^{(i)}, i = 1, \ldots, N$, which is the input voltage to the LDO network driven by the cluster; the number of online converters in the off-chip buck cluster $N_{\text{online,off}}$, and its output voltage $V_{\text{out,off}}$, which sets the input voltage to all on-chip buck VR clusters in the considered tree.

III. Modeling of HVR System

Clearly, the HVR voltage processing chain has a tree structure consisting of multiple voltage processing stages starting from a cluster of off-chip buck VRs and ending at the on-die loads in each local power domain. We look into the detailed energy and regulation characteristics at each individual stage first then consider the interdependencies across different stages in the HVR system.

A. Characteristics Per Stage

1) On/Off-Chip Buck Clusters: Fig. 5 shows a typically multiphase buck converter which is commonly used in modern processor systems. Each phase of the buck VR is implemented with fixed switching frequency and pulsewidth modulation (PWM). Each PWM comparator sets the duty cycle of its output voltage waveform which then drives power switches to produce the modulated final output voltage. The multiple parallel time-interleaved phases cancel out the high-frequency output noise and reduce the transient response time at the cost of increased overhead of inductors and control circuits [28].

The major power losses of a buck converter include two parts: the switching loss which is largely independent of load current and the resistive loss which is a function of the load current [29]. The switching loss dominates the power loss at light loads while the resistive loss grows quadratically with increasing load current. In addition, both parts of power loss are functions of the input–output voltages of the buck converter. In a cluster of buck VRs, its overall power efficiency can be further impacted by the number of online VRs N_{online} which varies the total switching loss under the same overall load current. As a result, the general form for the power efficiency of a buck cluster can be written as

$$\eta_{\text{buck}} = f(V_{\text{in}}, V_{\text{out}}, N_{\text{online}}, I_L). \quad (1)$$

For given input–output voltages, a single VR achieves the peak efficiency at an optimal load point I_{opt} where the ratio of total loss over the load power is minimized. Relying on the analytic power model as in [28] for a given set of design parameters such as switching frequency, filter inductance, and size of MOS switches, Fig. 6(a) demonstrates that the power efficiency curves of a buck cluster can be dramatically changed with a different number of online VRs N_{online}. The peak power efficiency for each curve can only be achieved at a certain optimal current load point, which is roughly $N_{\text{online}} I_{\text{opt}}$. Therefore, it is intuitive to bring online only a certain number

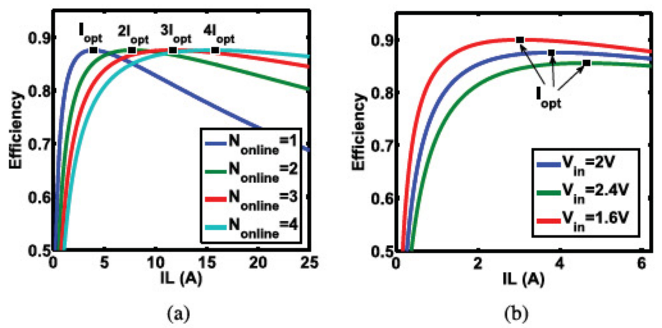


Fig. 6. (a) Impact of online buck VRs on power efficiency. (b) Impact of input voltage on I_{opt} for a single buck VR.

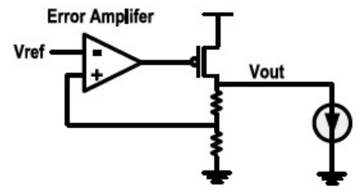


Fig. 7. Schematic of LDO.

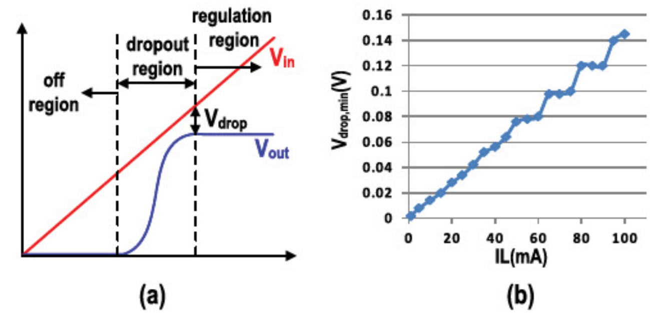


Fig. 8. Relationship between LDO's dropout voltage and load current.

of buck VRs in the cluster such that the load current per VR stays around I_{opt}. The number of required VRs N_{online} can be quantized as

$$N_{\text{online}} = \min\left\{N_{\max}, \left\lceil \frac{I_L}{I_{\text{opt}}} \right\rceil\right\} \tag{2}$$

where $N_{\max}$ is the maximum number of VRs in a cluster and I_L is the total load current.

Note that the chain structure of the HVR makes things much more complicated, because I_{opt} is a function of VR's input–output voltages which can be influenced by the preceding and subsequent stages. Fig. 6(b) illustrates the shift of I_{opt} for a single buck converter with varied input voltage. Such effect must be considered in (2). As will be discussed later, the adaptive control policy proposed in this work requires short processing latency to enable fine-grained temporal control resolutions. Therefore, the complex characteristics of buck converters are stored in two lookup tables (LUTs) for the ease of online use. For instance, LUT^{η} stores the power efficiency characteristics which are indexed by the input–output voltages and load current for each buck VR. As a function of the input and output voltages, LUT^{Iopt} stores the optimal load current under which the peak efficiency is achieved for a single buck VR.

Although the dc–dc buck converters are more suitable for voltage conversion as discussed earlier, the on-chip buck VRs, which is the final stage in the conventional two-stage PDN, have to be carefully designed with the consideration of supply noise. The power integrity will be largely determined by the transient response of the on-chip buck VRs. In general, increasing the switching frequency of the buck VRs will help reduce both the transient response time and output voltage ripples but at the price of increasing switching power loss. As a result, it is common to integrate on-chip buck converters operating at hundreds of megahertz in the two-stage PDN [27].

2) On-Chip LDO Networks: The proposed three-stage HVR system explores the fast voltage load regulation of an additional stage of distributed on-chip LDOs as discussed earlier. In addition, LDOs can be designed with a good power supply ripple rejection (PSRR) to suppress noise from the input voltage (i.e., line regulation) [30]. As a result, the on-chip buck converters in the three-stage HVR can be optimized to achieve better power efficiency, e.g., by operating them at a lower switching frequency.

To supply a specific output voltage, a linear LDO converts an input voltage using an error amplifier and feedback loop as depicted in Fig. 7. The power efficiency of an LDO is strongly limited by the input-to-output differential voltage $\Delta V = V_{\text{in}} - V_{\text{out}}$ for a given targeted output voltage V_{out}

$$\eta_{\text{ldo}} = \frac{V_{\text{out}}}{V_{\text{out}} + \Delta V}. \tag{3}$$

At a certain load point, the dropout voltage V_{drop} of an LDO is defined to be the minimum input-to-output differential voltage at which the LDO ceases to regulate the output voltage, i.e., entering the dropout region from the regulation region. Fig. 8(a) illustrates V_{out} as a function of V_{in}. Therefore, it is desirable to set V_{in} just V_{drop} above V_{out} to keep the LDO at the boundary between the dropout and regulation regions to maximize efficiency. However, setting V_{in} too low may jeopardize the regulation of LDOs and violate power integrity.

V_{drop} is a function of the load current I_L, which is shown in Fig. 8(b) for a realistic LDO design [30]. It can be seen that V_{drop} is approximately linear in I_L, hence $V_{\text{drop}} \approx (I_L/I_{L,\max})V_{\text{drop,max}}$, where $V_{\text{drop,max}}$ is the dropout voltage at the maximum current load $I_{\text{L,max}}$. Given a target output voltage V_{dd}, e.g., one set by DVS, the optimal LDO's input voltage (output voltage of the on-chip buck VRs), which leads to the highest of LDO power efficiency, is

$$V_{\text{in,opt}} \approx V_{\text{dd}} + \frac{I_L}{I_{L,\max}} V_{\text{drop,max}}. \tag{4}$$

B. Interdependencies Between Voltage Processing Stages

According to the above discussion, the power efficiency of a single VR stage largely depends on its input–output voltages and current load. Thus, there are important interdependencies among voltage processing stages which must be carefully considered in order to optimize the overall energy efficiency and regulation performance. Such interdependencies can be observed in (5), shown at the bottom of this page, which describes the overall power efficiency as the product of efficiencies at all stages. Since the input voltage of the off-chip buck VRs is assumed to be constant, it is not considered in the corresponding power efficiency $\eta_{\text{buck,off}}$. Under a certain workload $\vec{I}_{L,\text{on}} = \{I_{L,\text{on}}^{(1)}, I_{L,\text{on}}^{(2)}, \dots I_{L,\text{on}}^{(N)}\}$ and DVS setting $\vec{V}_{\text{dd}} = \{V_{\text{dd}}^{(1)}, V_{\text{dd}}^{(2)}, \dots V_{\text{dd}}^{(N)}\}$, where N is the number of power domains, the control variables listed in Table II can simultaneously influence the power efficiencies at multiple stages due to the interdependencies in the voltage regulation chain. For example, the output voltage of the off-chip buck VRs $V_{\text{out,off}}$ influences the efficiencies of both off-chip and on-chip buck VRs. Set by the output of corresponding on-chip buck cluster $V_{\text{out,on}}^{(i)}$, the input voltage to an LDO network significantly impacts the power efficiencies of the (preceding) on-chip buck cluster, and the final power quality for the loads observed on the power grids. As a result, such interdependencies have to be considered in the online adaption for maximal power efficiency and noise tradeoffs.

IV. HVR Control Policies

We present our proposed control policies for three-stage HVR, while these policies can be straightforwardly applied to adapt two-stage HVR consisting of only off-chip and on-chip switching VRs. Unlike most related work executing power management in the OS or software [16], [18], the proposed policies can be efficiently implemented in firmware based on simple arithmetics and precomputed LUTs supported by hardware accelerated ML prediction of workload.

The settling times of off-chip and on-chip switching VRs of the first two stages can differ by several orders of magnitude. Hence, they are adapted using two different control cycle times, denoted by T_{off} and T_{on}, respectively. Each T_{off} is split into a multiple of T_{on}. Accordingly, off-chip and on-chip switching VRs are adapted by two control procedures, which are shown in Fig. 9 for an HVR system with N power domains, one for each core. We estimate the core-level workloads $\vec{I}_{L,\text{Toff}} = \{I_{L,\text{Toff}}^{(1)}, I_{L,\text{Toff}}^{(2)}, ..I_{L,\text{Toff}}^{(N)}\}$ and $\vec{I}_{L,\text{Ton}} = \{I_{L,\text{Ton}}^{(1)}, I_{L,\text{Ton}}^{(2)}, ..I_{L,\text{Ton}}^{(N)}\}$, respectively, at the time granularities of T_{off} and T_{on} using power sensors [31] at the output of each on-chip switching VR (buck converter) cluster. At both time scales, we use the workload estimates obtained from the previous control cycle to generate control actions for the current cycle.

In each off-chip VR control cycle T_{off}, the off-chip VR control procedure VR_OFF_OPT is invoked to optimize the off-chip VR output voltage $V_{\text{out,off}}$ and the number of online off-chip VRs $N_{\text{online,off}}$ based on $\vec{I}_{L,\text{Toff}}$. Each T_{off} is divided into a multiple of much finer grained on-chip VR control cycles T_{on} as shown in Fig. 10. The on-chip control procedure VR_ON_OPT is invoked in each T_{on} cycle to adjust

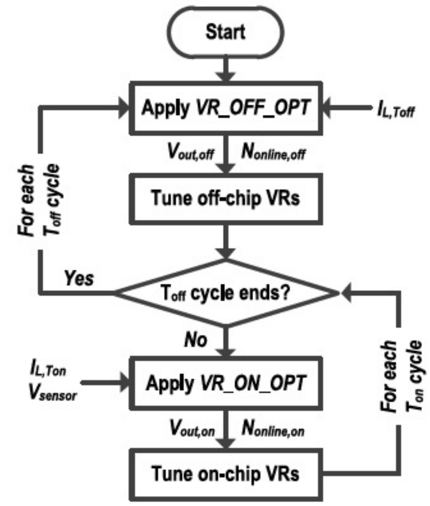


Fig. 9. Control of off-chip and on-chip switching VRs at two time scales.

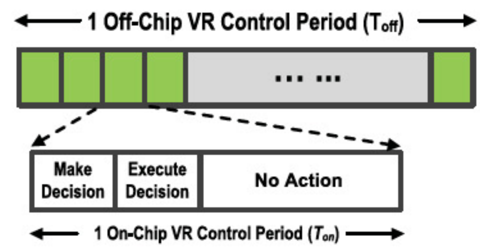


Fig. 10. Two control sequences.

the output voltage $V_{\text{out,on}}^{(i)}$ and the number of online VRs $N_{\text{online,on}}^{(i)}$ for each on-chip VR cluster, $i = 1, 2, \dots N$, based on the finer grained workload estimation $I_{L,\text{Ton}}^{(i)}$. As detailed in Section V, VR_ON_OPT relies on an ML module utilizing a small N_S number of voltage sensors to more precisely adjust $\vec{V}_{\text{out,on}} = \{V_{\text{out,on}}^{(1)}, V_{\text{out,on}}^{(2)}, \dots, V_{\text{out,on}}^{(N)}\}$, based on the spatial distribution of the workload in each power domain. The voltage sensor readings $\vec{V}_{\text{sensor}} = \{V_{\text{sensor}}^{(1)}, V_{\text{sensor}}^{(2)}, \dots, V_{\text{sensor}}^{(N_S)}\}$ are included as input to VR_ON_OPT.

Fig. 10 shows the timing of the control sequences. There are three steps involved in each T_{on} cycle. The first decision making step executes VR_ON_OPT procedure to compute $\vec{V}_{out,on}$ and $\vec{N}_{\text{online,on}}$ and the on-chip VRs are adjusted accordingly in the second decision execution step.

A. Off-Chip Switching VR Control

The output voltage $V_{\text{out,off}}$ of the off-chip switching VRs is the input voltage to all on-chip switching VR clusters. $V_{\text{out,off}}$ impacts the power efficiencies of both on-chip and off-chip buck VRs as well as the resistive power loss due to printed circuit board (PCB)/package parasitics. As in Algorithm 1, the off-chip control procedure VR_OFF_OPT uses the following iterative search to find the optimal $V_{\text{out,off}}$ among

$$\eta_{\text{HVR}} = \eta_{\text{buck,off}}(V_{\text{out,off}}, N_{\text{online,off}}, I_{L,\text{off}})\eta_{\text{buck,on}}(V_{\text{out,off}}, \vec{V}_{\text{out,on}}, \vec{N}_{\text{online,on}}, \vec{I}_{L,\text{on}})\eta_{\text{ldo}}(\vec{V}_{\text{out,on}}, \vec{V}_{dd}) \quad (5)$$

Algorithm 1 Off-Chip Control Algorithm VR_OFF_OPT

Inputs:
Workload current estimations $\vec{I}_L$ for each T_{off} cycle.

1: Maximize $\eta(V_{out,off})$, subject to
2: $V_{min\_on} \leq V_{out,off} \leq V_{max\_on}$
3: **for** each power domain i **do**
4: **if** $3\_stage\_HVR == True$ **then**
5: $V_{out,on}^{(i)} = V_{dd}^{(i)} + \alpha I_L^{(i)}/I_{max}^{(i)}$
6: **else**
7: $V_{out,on}^{(i)} = V_{dd}^{(i)}$
8: **end if**
9: $I_{opt,on}^{(i)} = LUT_{on}^{Iopt}(V_{out,off}, V_{out,on}^{(i)})$
10: $N_{online,on}^{(i)} = \left\lceil I_L^{(i)}/I_{opt,on}^{(i)} \right\rceil$
11: $\eta_{on\_sw}^{(i)} = LUT_{on}^{\eta}(V_{out,off}, V_{out,on}^{(i)}, I_L^{(i)}/N_{online,on}^{(i)})$
12: **end for**
13: $I_{pkg} = \sum_i (I_L^{(i)} V_{out,on}^{(i)}/V_{out,off}/\eta_{on\_sw}^{(i)})$
14: $\eta_{on\_chip} = (\sum_i V_{dd}^{(i)} I_L^{(i)})/I_{pkg}/V_{out,off}$
15: $\eta_{pkg} = V_{out,off}/(V_{out,off} + I_{pkg}R_{pkg})$
16: $I_{opt,off} = LUT_{off}^{Iopt}(V_{ext}, V_{out,off})$
17: $N_{online,off} = \lceil I_{pkg}/I_{opt,off} \rceil$
18: $\eta_{off\_chip} = LUT_{off}^{\eta}(V_{ext}, V_{out,off}, I_{pkg}/N_{online,off})$
19: $\eta = \eta_{on\_chip}\eta_{pkg}\eta_{off\_chip}$
20: Return $\{V_{out,off}, N_{online,off}\}$ with maximized η

Algorithm 2 On-Chip Control Algorithm VR_ON_OPT

Inputs:
Workload current estimations $\vec{I}_L$ for each T_{on} cycle.
Voltage sensor readings $\vec{V}_{sensor}$ for each T_{on} cycle.

1: **for** each power domain i **do**
2: **if** $3\_stage\_HVR == True$ **then**
3: $V_{out,on}^{(i)} = V_{dd}^{(i)} + \alpha I_L^{(i)}/I_{max}^{(i)}$
4: **else**
5: $V_{out,on}^{(i)} = V_{dd}^{(i)}$
6: **end if**
7: $I_{opt,on}^{(i)} = LUT_{on}^{Iopt}(V_{out,off}, V_{out,on}^{(i)})$
8: $N_{online,on}^{(i)} = \left\lceil I_L^{(i)}/I_{opt,on}^{(i)} \right\rceil$
9: **if** $MachineLearningOption == True$ **then**
10: $V_{out,on}^{(i)} = MachineLearning(\vec{S}_{PDN}, \vec{V}_{sensor}^{(i)})$
11: **end if**
12: **end for**
13: Return $\{\vec{V}_{out,on}, \vec{N}_{online,on}\}$

a set of discretized values of $V_{\text{out,off}}$ while considering the above interactions. At each iterative search step with a targeted $V_{\text{out,off}}$ value, we first estimate the input voltage to each LDO network $V_{\text{out,on}}^{(i)}$ in line 5 as a linear function of workload to maximize the power efficiency of the LDO's as in Section III-A for three-stage HVR. Otherwise, for two-stage HVR, $V_{\text{out,on}}^{(i)}$ is directly set by system's power management (e.g., DVFS) unit as shown in line 7. Then, the optimal load point for each on-chip buck VR $I_{\text{opt}}^{(i)}$ is determined in line 9 via an LUT with the known input–output voltages. $N_{\text{online,on}}^{(i)}$ is further determined in line 10. The power efficiency for each on-chip buck cluster is conveniently estimated through the use of another LUT in line 11. The total through-package current, which is the sum of the input currents of all on-chip buck clusters is computed in line 13 and used as the load current of the off-chip buck cluster. The power efficiency of the on-chip components of HVR is computed in line 14 considering both integrated buck VRs and LDOs. Our experimental study shows that the resistive loss caused by PCB/package parasitics may not be negligible, which is considered in line 15. Following a similar procedure, $N_{\text{online,off}}$ and the off-chip power efficiency are determined in lines 16–18. The overall system power efficiency at the current value of $V_{\text{out,off}}$ is the product of the efficiencies of all stages as in line 19. Finally, the combination of the value of $V_{\text{out,off}}$ and the corresponding $N_{\text{online,off}}$ that maximizes the system efficiency is chosen as the optimal control of the off-chip buck VRs for this T_{off} cycle.

B. On-Chip Switching VR Control

Once the slowly changing variables $V_{\text{out,off}}$ and $N_{\text{online,off}}$ are determined for each T_{off} cycle, $V_{\text{out,on}}^{(i)}$ and $N_{\text{online,on}}^{(i)}$ per domain are updated for each finer temporal cycle T_{on} by calling VR_ON_OPT shown in Algorithm 2. We follow a flow similar to VR_OFF_OPT to determine $N_{\text{online,on}}^{(i)}$ in lines 2–8. However, if the ML is enabled, the final $V_{\text{out,on}}^{(i)}$ is fine-tuned by the ML module with the consideration of fine-grained spatial workload distribution, described next.

V. Machine Learning Enabled Adaption

One key objective of voltage regulation is to deliver power to on-die devices with ensured power integrity, e.g., without dropping the worst case voltage from the on-chip power grids below a preset level. Power supply noise hotspots are created due to the nonuniform spatial distribution of workload on-chip. To make things even worse, the locations of hotspots can shift during runtime. Such effects can significantly impact the on-die supply noise. Thus, the output voltage of each on-chip switching VR cluster, which is the final point of two-stage voltage regulation, and also the input voltage to the distributed LDO network in the case of three-stage HVR, shall be adapted with the considerations of fine-grained spatial workload distribution. However, predicting such spatial workload distribution for the purpose of PDN adaptation is a challenging problem.

Recently, ML has been received a significant amount of interest for power system design. For instance, noise-sensor-based ML techniques [32], [33] have been developed to detect voltage emergencies within functional blocks. Different from these works, we leverage ML to directly learn the optimal control policy based on the fine-grained spatial workload distribution predicted from a small number of distributed voltage noise sensors. This enables a very desirable end-to-end ML solution that can lead to additional energy and power integrity benefits.

A. Machine Learning Problem Formulation

We first formulate the ML problem. For a power domain, denote the output voltage of the corresponding on-chip switching VR (buck) cluster $V_{\text{out,on}}$. By exploiting the correlation

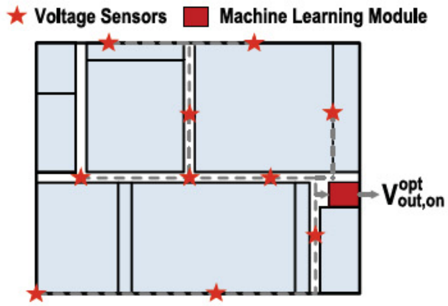
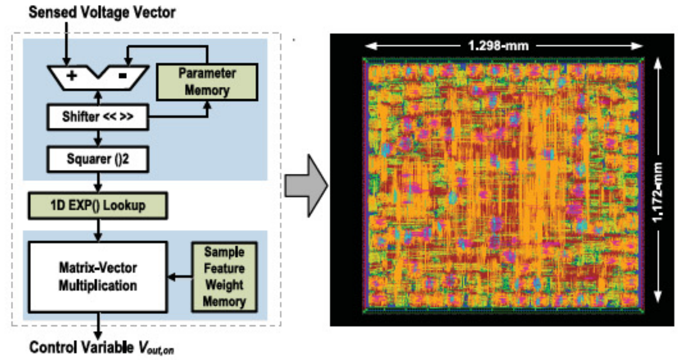


Fig. 11. Demonstration of ML module and voltage sensors.



(a) (b)

Fig. 12. (a) Proposed on-chip SRKM accelerator. (b) Layout of an SRKM accelerator with parallelism parameter equaling 8.

between voltage drops at different nodes in the power grids (including sensor locations) and the distribution of workload, an ML model can directly learn the optimal control variable $V_{\text{out,on}}^{\text{opt}}$ using the voltage sensor readings as input features. Here, $V_{\text{out,on}}^{\text{opt}}$ is defined as the minimum $V_{\text{out,on}}$ value such that the worst case supply voltage across the entire power grids does not fall below a preset safety voltage level. By leveraging the fine-grained spatial information of workload distribution, $V_{\text{out,on}}$ can be set in a more accurate way, achieving improved power efficiency and quality. An ML model is used to learn the following mapping:

$$\vec{S}_{\text{PDN}}, \vec{V}_{\text{sensor}} \rightarrow V_{\text{out,on}}^{\text{opt}} \tag{6}$$

where $\vec{V}_{\text{sensor}}$ is the worst case voltage values sensed by the voltage sensors during an on-chip VR control cycle T_{on}. $\vec{S}_{\text{PDN}}$ includes the PDN configurations such as control variables under which the voltage sensor values are measured. The training samples can be collected by circuit simulation by sweeping $V_{\text{out,on}}$ within a certain range to obtain the target $V_{\text{out,on}}^{\text{opt}}$ under the same workload. Fig. 11 illustrates the ML module and voltage sensors in a power domain.

B. Machine Learning Algorithm

We integrate our ML module (accelerator) on-chip to enable fast real-time workload-aware adaption. Such ML module must come with sufficient accuracy, low area/power overhead, and should incur low processing latency to enable HVR adaptation at fine temporal granularity. In this work, we adopted a recently developed sparse Bayesian-based ML algorithm, namely, sparse relevance kernel machine (SRKM) [34], [35] as the ML algorithm. As a kernel machine, SRKM predicts the target value y of a new input vector $\boldsymbol{x}$ using N training samples X_i

$$y(\boldsymbol{x}; \boldsymbol{w}) = \sum_{i=1}^{N} w_i \cdot K(\boldsymbol{x}, X_i) \tag{7}$$

where $K(\boldsymbol{x}, X_i)$ is the kernel function, and $\boldsymbol{w} = [w_1, w_2, \ldots, w_N]^T$ is the vector of sample weights. It should be noted that for the system with analog/mixed-signal circuits, such as the PDN under this study containing a large number of active VRs, SRKM utilizes a nonlinear kernel function that can well capture the nonlinear mapping from voltage sensor readings and PDN configurations to the optimal control variable as illustrated earlier.

Applying the above nonlinear regression model for all N training samples gives $\boldsymbol{t} = \boldsymbol{\Phi} \cdot \boldsymbol{w} + \boldsymbol{e}$, where $\boldsymbol{\Phi}$ is an $N \times N$ matrix defined by $\boldsymbol{\Phi}(i, j) = K(\boldsymbol{x}_i, \boldsymbol{x}_j)$, $\boldsymbol{t}$ is the vector of the N target values, and $\boldsymbol{e}$ is the error vector. Similar to the relevance vector machine (RVM), the SRKM model is treated as probabilistic, whereby the model parameters $\boldsymbol{w}$ are considered as random variables, which are optimally inferred in the training process. It has been demonstrated the advantages of SRKM for a variety of applications in [34] and [35]. Unlike the widely adopted support vector machine (SVM) and RVM, SRKM can achieve sparsity in both the (training) sample and (parameter) feature space. For each sample and feature, a weight parameter is learned from the training process to signify the significance of the sample/feature with respect to the prediction of the target value. SRKM produces sparser models with improved accuracy compared to SVM and RVM, and offers a built-in mechanism to filter out redundant samples and features based upon quantitative weight information. The resulted sparse model is very appealing for achieving low processing latency and hardware overhead in our application. More details about SRKM can be found in [34] and [35].

C. SRKM Accelerator

Following (7), we propose an on-chip SRKM predictor design in Fig. 12(a). It only utilizes simple arithmetics such as ADD and MUL/DIV, and a 1-D LUT for the exponential calculation. The model parameters obtained from offline training are stored using a small amount of on-chip memory. The entire design is based on fixed-point 24-bit operation and only introduces a small quantization error of 2.2 mV evaluated under 4000 samples. By exploiting the rich parallelisms embedded in (7), we explore SRKM modules of different degree of parallelisms using a standard 45-nm CMOS technology. The main hardware results are summarized in Table III, demonstrating a good tradeoff between processing latency, power, and area overhead. The layout of the SRKM hardware with eight-way parallelism is shown in Fig. 12(b).

In our work, we train an SRKM model offline based on 2000 training samples collected from circuit simulation.

TABLE III
HARDWARE RESULT FOR SRKM ACCELERATOR WITH DIFFERENT PARALLELISM

Parallelism	Area (mm^2)	Power (mW)	Latency (ns)
1	0.410	17.94	506
4	0.860	35.77	176
8	1.522	66.18	121
12	2.179	96.45	103

TABLE IV
PROCESSOR CONFIGURATION

# Cores	4	Vdd	1V
Frequency	1.8GHz@45nm	Imax/Pmax	25A/25W
Branch Predictor	2K entries	Core area	40.4mm^2
ALU/MUL/FPU	6/2/6	I/D-TLB	48/64 entires
Load/Store buffer	32	ROB size	192
L1 I/D-Cache	32KB, 2-way, 2-cycle latency	L2 Cache	Shared 2MB, 20-cycle latency

It achieves a normalized mean square error (NMSE) of 4.3e-3, demonstrating excellent prediction accuracy. As mentioned earlier, the trained SRKM model is mapped to a hardware accelerator for efficient runtime application. It should be noted that the proposed overall control scheme does not need additional software support except the offline SRKM model training. At runtime, all the decision-making of control variables will be processed through light-weighted hardware such as simple arithmetics, precomputed LUTs storing VRs' characteristics, and the SRKM hardware accelerator for workload prediction.

VI. EXPERIMENTAL EVALUATIONS

A. Experimental Setup

1) Multicore Processor Model and Power Analysis: We use the full-system multicore simulator GEM5 [36] to generate runtime statistics with the granularity of 100 ns and then feed them into the power analysis tool McPAT [37] to produce realistic workload current traces. The 45-nm four-core processor model illustrated in Table IV is evaluated using the PARSEC benchmark suite [38]. The total die area including on-chip VRs is estimated as 286.4 mm^2 by McPAT and PowerSoC [28]. The peak workload current per core is 25 A. Each core is divided into 11 functional blocks. The current workload of each block, derived from McPAT, is evenly distributed within the block to load the PDN.

2) Power Delivery Network: To enable the comparison across different PDN architectures, we consider the widely used two-stage PDN shown in Fig. 2(b) with on-chip/off-chip buck VRs as the reference. The main structure of the reference system is similar to the three-stage HVR except that the centered on-chip buck converters are used as the last voltage processing stage instead of the distributed LDO network. We adopt a PCB/package/power grid model similar to [39], which is derived based on Pentium 4 processor, for both PDNs. The effects of packaging and long power routing are included in the power model of PDNs. Considering the feasibility of circuit-level simulation, the on-chip power grids of the PDN are modeled using an *RC* network with more than 3000 nodes. As the on-chip decoupling capacitance (decap) is highly correlated with the voltage noises, we scale the total amount of decap C_{decap} by keeping a similar $C_{\text{decap}}/I_{\max}$ ratio as in [39], where $I_{\max}$ is the maximum load current.

In the regulation chain of each PDN, a cluster of 5 off-chip buck VRs is used to drive 5 on-chip buck clusters with each cluster containing four identical on-chip VRs. In three-stage HVR, each on-chip VR cluster further drives a network of 250 on-chip LDOs for each core (power domain). The topology from [30] is adopted for on-chip LDOs with maximum 100-mA load capability. The off-chip and on-chip buck converters are designed using PowerSoC [28], which finds the key design parameters such as switching frequency, filter inductance, and size of MOS switches under a static nominal load condition. Considering the on-chip buck converters are the final regulation stage in the two-stage reference PDN, they are designed with more emphasis on regulation performance at the cost of more energy loss. As a result, the on-chip buck VRs of the two-stage PDN operate at 291 MHz, while those of the three-stage HVR operate at 107 MHz. The area of on-chip buck VRs for two-stage and three-stage PDNs are 15 mm^2 and 13.75 mm^2, respectively. The area of on-chip LDOs for three-stage PDN is 1.25 mm^2. Clearly, the total area budget of on-chip VRs (including LDOs) is set to 15 mm^2 for both PDNs for a fair comparison.

3) Control Scheme Setup: The on-chip and off-chip VR control periods T_{on} and T_{off} are set to 1 and 100 us, respectively, to suit the response times of the considered on-chip and off-chip switching VRs.

As shown in Algorithm 2, the ML-enabled control scheme takes the voltage sensor readings as input to predict the optimal output voltage of on-chip switching VRs. However, obtaining the voltage sensor readings for each PARSEC benchmark during runtime through the simulation of our complex PDN model is prohibitively computationally expensive. To speed up the evaluation process, we once again leverage ML but for the purpose of fast estimation of voltage sensor readings. We train another SRKM model offline which performs the following mapping:

$$\vec{S}_{\text{PDN}}, \vec{I}_{\text{block}}(n), \vec{I}_{\text{block}}(n-1) \rightarrow \vec{V}_{\text{sensor}}(n) \tag{8}$$

where $\vec{I}_{\text{block}}(n)$ and $\vec{I}_{\text{block}}(n-1)$ are the block-level workloads at the current and past 100ns time steps, representing the fine-grained workload transition, $\vec{V}_{\text{sensor}}(n)$ is the worst case voltage sensor readings caused by the corresponding transitions. Based on the traces of $\vec{V}_{\text{sensor}}(n)$, the worst case voltage sensor readings for each control cycle T_{on} can be computed as the input to the ML module. Similar to the online SRKM module in Section V, the PDN state variables $\vec{S}_{\text{PDN}}$ are included as part of the input features for this offline SRKM model to estimate $\vec{V}_{\text{sensor}}(n)$. Trained with 4000 samples, this offline SRKM model is very accurate and achieves an average NMSE of 1.52e−4.

B. Online Machine Learning Overhead

The area and power overhead of the proposed ML-enabled HVR adaptation comes from the voltage sensors and SRKM accelerators. The voltage sensors can be implemented

TABLE V

ADDITIONAL AREA AND POWER OVERHEAD (%). AREA IS NORMALIZED TO THE ORIGINAL ON-DIE AREA. POWER IS NORMALIZED TO HALF OF THE PEAK POWER

	Amount	Area overhead	Avg. power overhead
Voltage sensors	40	0.419%	0.166%
SRKM accelerators	4	2.126%	0.051%
Total	-	2.545%	0.217%

based on low-power high-speed analog-to-digital converters (ADCs) [40], [41]. The ADC design in [41] is considered to estimate the sensor cost. In our study, ten voltage sensors and a compact SRKM accelerator are placed in each core. The ML calculation is the major latency in the control loop. To avoid large performance degradation and achieve good responsiveness to workload change, the SRKM latency is expected to be much smaller than the on-chip control window size T_{on} (e.g., 1 μs). Considering the tradeoffs among processing latency and hardware overhead as in Table III, we choose the SRKM accelerator with parallelism equaling eight to achieve a satisfactory short latency with a moderate hardware overhead. As summarized in Table V, the proposed ML approach only incurs an overhead of 2.5% on area and 0.2% on power but comes with great benefits.

C. Power Integrity and Adaptive Control

1) Power Integrity: We examine the power quality of several adaptive PDNs through detailed circuit-level simulation. Verilog-A models with PWM control are used to model the on-chip buck converters based on design parameters obtained from PowerSoC. The ideal voltage source is used for the off-chip VRs since they have little impact on power supply noise. The complexity of our PDN model with a large number of VRs causes significant simulation challenge. It takes around 112 h to simulate a 100-μs segment of benchmark workload with four threads on an Intel Xeon E5-2697A processor at 2.60 GHz. We select a 100-μs workload segment from each PARSEC benchmark, forming a workload simulation set. This set contains a representative worst case workload segment from the fluidanimate benchmark and random segments from other benchmarks, serving as typical workload conditions. As described in Section IV, our control algorithm supports both two-stage and three-stage PDNs and also provides two options with and without ML module. This creates four adaptive PDNs and they are simulated based on the aforementioned workload simulation set.

In modern processors, multiple factors such as clock gating and workload variation can lead to unpredictable supply voltage noises. Once the worst case voltage droop exceeds an operating margin (10% of the nominal V_{dd} in this study), voltage emergency (VE) will happen and may cause timing violations. Although designing a static PDN based on worst case load scenario can guarantee the robustness with a large voltage safety margin, the power efficiency significantly degrades. Instead, more aggressive voltage margins can be used in modern designs to reduce the power consumption greatly and allow rare occurrences of VEs by fail-safe mechanisms such as the rolling-back recovery [42] or adaptive frequency tuning [43].

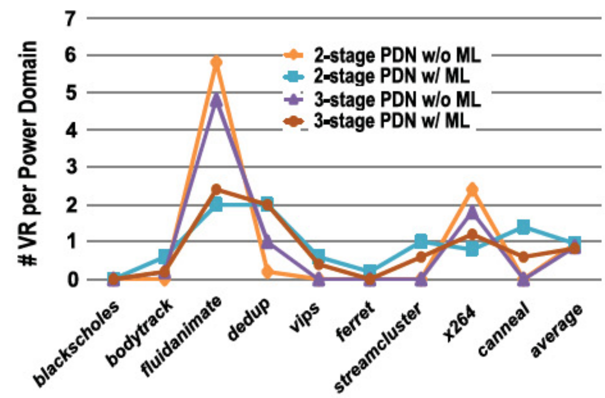


Fig. 13. Average number of VEs per power domain for each benchmark segment.

In our study, we assume the processor is equipped with such a mechanism to recover from the rare events of VEs. The count of VEs is used as a metric for power quality.

Fig. 13 plots the average count of VEs per power domain under the workload segment of different PARSEC benchmarks. On average VE only occurs about once in each power domain for all PDNs. In other words, all PDNs have the same power integrity level. Under this equal power quality condition, we will compare these PDNs in terms of energy efficiency in Section VI-D.

2) Case Study for Adaptive Control: Next, we use two simulation examples of the adaptive three-stage HVR systems with and without ML module to shed some light on how the proposed control policies adapt to the workload change and the benefits brought by ML. Fig. 14(a) shows the transient waveforms based on fluidanimate. Such workload segment represents a worst case scenario since the total load current suddenly increases to the maximum 25-A peak current from light-load condition. The fast and large load variations as such tend to cause a considerable amount of power supply noise, imposing a significant regulation challenge. The resulting worst voltage V_{PG} in the entire on-chip power grids is plotted. The dashed line indicates the supply voltage level under which VE is considered to happen. It can be seen that the system armored with ML can more accurately set the output voltage of the on-chip buck converters V_{out}. That is, V_{out} that is further fine-tuned by the proposed ML module becomes lower under lighter load conditions, reducing the energy loss of the LDO networks. On the other hand, V_{out} can be quickly increased in response to the arrival of heavier workloads. The number of online on-chip buck converters N_{online} is also well adapted to the workload variation for energy saving.

Fig. 14(b) shows a more typical workload example from the streamcluster benchmark. The corresponding power trace exhibits periodic behavior resulted from a for loop in the program. Although no VE happens in both PDNs, it is evident that the ML solution further improves energy efficiency due to lower values of V_{out}.

D. Overall Energy Evaluation

1) Energy Comparison: The overall energy efficiencies of different PDN architectures with various control schemes are

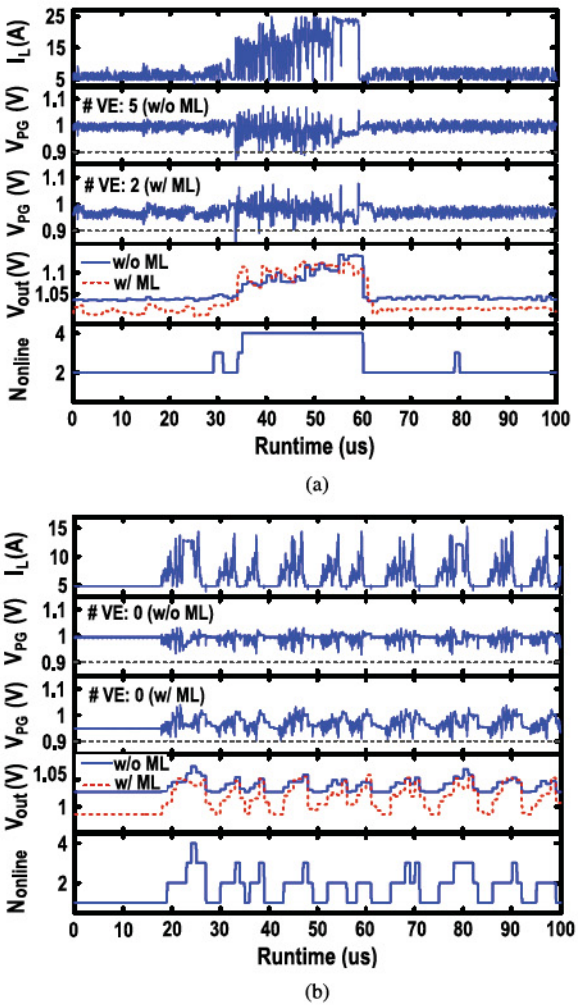


Fig. 14. Transient circuit simulation waveforms of the adaptive three-stage HVR running (a) fluidanimate benchmark and (b) streamcluster benchmark.

compared. We name all considered PDNs in Fig. 15 (top). There are four three-stage PDNs denoted by 3-S1–3-S4 with different control policies. Take the 3-S4 PDN with configuration three-stage $N_{\text{online,on/off}}$-$V_{\text{out,on/off}}$(ML) for example. The configuration means that the system utilizes a three-stage HVR PDN architecture, enables the tuning of the number of online VRs, output voltage of both on-chip and off-chip switching VRs, and it integrates the ML module. The first 3-S1 system indicates a static three-stage PDN with no runtime adaptation. Similarly, we have four different PDNs with two-stage architecture. We highlight several observations from Fig. 15.

1) Without any adaptive control, the static 3-S1 outperforms 2-S1 with an energy reduction of 4.0% on average. On the other hand, with a complete control scheme, 3-S4 shows up to 5.0% energy reduction over 2-S4. It demonstrates the potential of leveraging HVR to improve energy and performance tradeoffs.
2) The 2-S2 adopts a simple control scheme similar to [17] by tuning the number of online on-chip/off-chip buck VRs in the two-stage PDN. It is observed that it reduces the energy by 8.5% over the static 2-S1. However, adding $V_{\text{out,off}}$ into 2-S3 can bring in an additional 2.1% energy saving on average, since such a scheme captures more interdependency among the regulation chain. By comparing 2-S4 with 2-S3, the proposed ML module offers up to 4.1% reduction of system energy by utilizing the spatial workload distribution information.
3) The highest energy efficiency is achieved by the proposed ML enabled adaptive 3-S4 system. The 3-S4 system reduces the total system energy dissipation by up to 17.9% and 12.2% on average compared to the static 3-S1. Compared with the conventional static 2-S1, our 3-S4 with runtime control reduces system energy by up to 23.9% and 15.7% on average.

Fig. 16 further decomposes the energy consumption for 2-S1, 2-S4, 3-S1, and 3-S4 systems. It is observed that, in general, the processor in the two-stage HVR consumes less energy compared to that of the two-stage PDN. That is because the distributed LDO network enhances the supply noise suppression and thus enables lower supply voltage while maintaining the same power integrity, demonstrating the benefit of HVR in voltage regulation. By setting the output voltage of on-chip buck VRs in a more accurate way, the use of ML module significantly improves the LDO's power efficiency in the three-stage HVR system while reducing the processor's energy consumption in the two-stage system. With full consideration of the energy interdependency in the regulation chain, the proposed control policy achieves a near-optimal overall power efficiency by carefully trading off power loss at different stages.

2) Impact of Control Granularity: As discussed earlier, great benefits of adaptive control may be achieved at the finest possible temporal granularity by tracking the workload more closely. To demonstrate the same, Fig. 17(a) shows the corresponding power loss increments for the 3-S4 system by applying coarser on-chip control granularities. Enlarging T_{on} from 1 to 10 μs and 100 μs, the total power loss increases by up to 5% and 10%, respectively, demonstrating the benefits of fine-grained adaptive control. However, it is observed in Fig. 17(b) that even with a coarser T_{on}, significant power reduction can still be achieved over the static 3-S1 system, demonstrating the effectiveness of the proposed adaptive HVR over a wide range of control granularity.

VII. Related Work

Recently, various power management techniques [3], [10]–[13] have been proposed to save power and improve the overall processor's performance at the system and architecture level. For example, [3] explores the benefits of fast DVFS at submicrosecond time scale using on-chip switching regulators. And [13] proposes an adaptive guard-banding approach to dynamically adapt chip clock frequency and voltage based on timing-margin measurements at runtime. Different from these DVFS techniques which target the optimization of processor's power and performance, this work explores the energy reduction opportunity in the PDN which delivers energy to the processor.

At the circuit level, several works have investigated the benefits of workload-aware PDN designs. Optimizing toward the single-stage PDN, as shown in Fig. 3(a) [18], consolidates

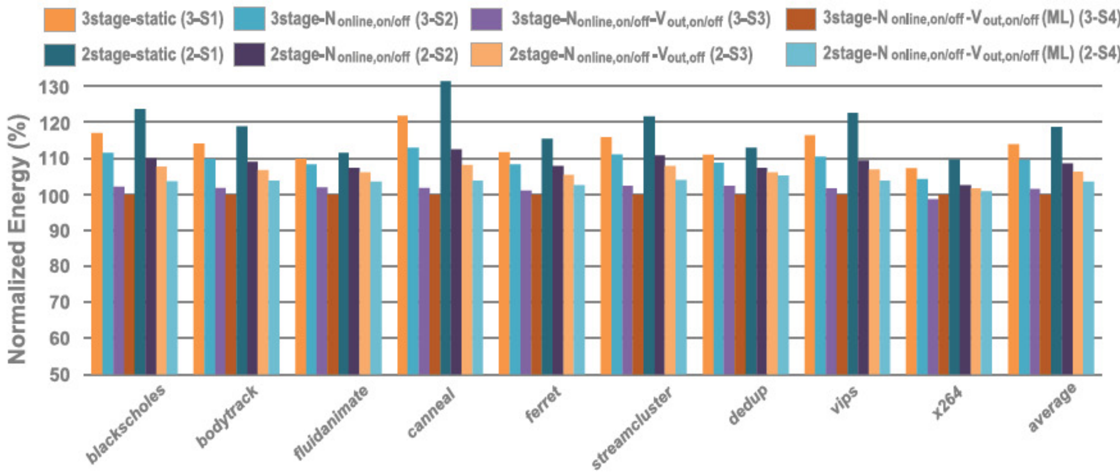


Fig. 15. Overall energy estimation for different PDN designs.

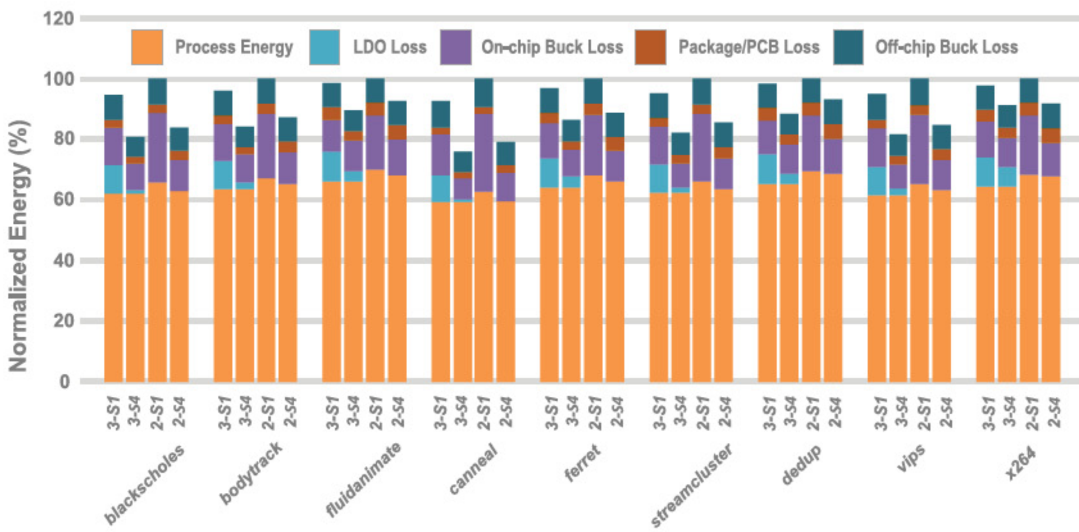
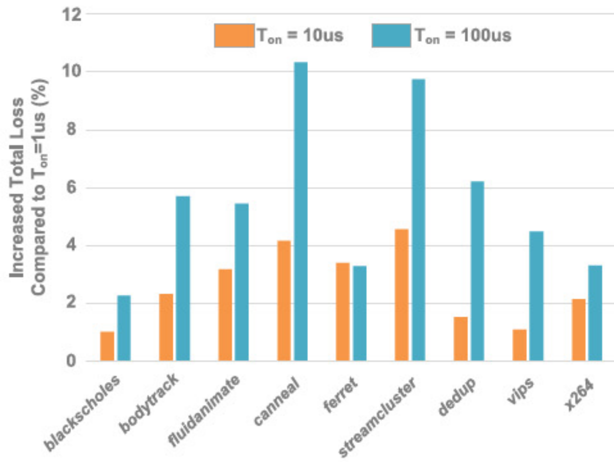


Fig. 16. Detailed energy breakdown for four different PDNs.

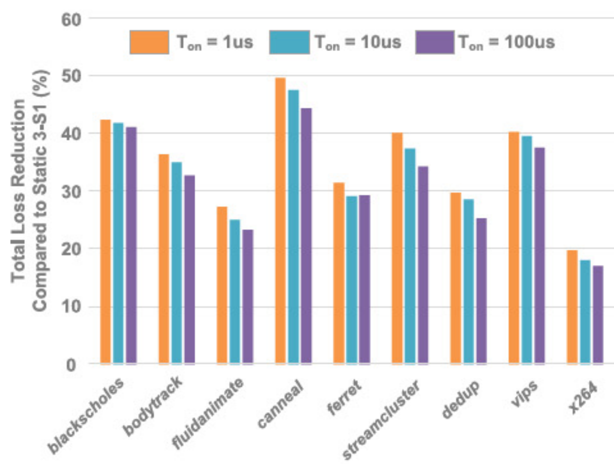
multiple power domains with the same supply voltage level to share a single off-chip inductor-based buck VR to avoid low conversion efficiency at light-load condition. In the same spirit, [16] proposes to reconfigure the PDN by combining the output of multiple VRs when the workload demand exceeds the peak current of a single VR. Targeting a two-stage PDN using both off-chip and on-chip switching VRs as shown in Fig. 3(b), a workload-aware QPM scheme is proposed in [17] to dynamically adjust the number of active on-chip and off-chip switching VRs at multiple granularities according to the chip-level runtime workload. However, these PDN reconfiguration techniques are all based on the core- or chip-level workload estimations without considering on-chip distributed LDOs and finer grained spatial workload distribution which can significantly impact on-chip supply noise. In addition, they do not consider the interdependencies among different power stages during power efficiency optimization. As discussed in this paper, the optimal tradeoff between power efficiency and quality can be best achieved with a systematic and joint consideration of all the related factors. For example, different from the adaptive control policy applied to the off-chip VRs as proposed in [17], this work sets the number of off-chip buck VRs along with other voltage processing stages to maximize the overall system power efficiency by considering interdependencies across the entire voltage processing chain. In addition, one additional control variable, i.e., the output voltage of off-chip VR stage, is considered in this work to make the proposed control scheme more comprehensive, bringing 2.1% overall system energy reduction on average. We also show the potential benefits of the three-stage heterogeneous PDN with multiple VR topologies with complementary characteristics over the conventional two-stage PDN with a single VR topology. Finally, the use of ML and voltage sensors to directly learn the control policy considering the spatial on-chip workload distribution presents excellent new opportunities. Our results demonstrate the great potential in leveraging the rich heterogeneity and optimization opportunities in multistage HVR systems for improved power efficiency and quality tradeoffs.

VIII. Conclusion

Targeting multistage HVR systems, this paper develops comprehensive workload-aware control policies acting at multiple temporal granularities based on complimentary characteristics of on-chip and off-chip VRs. The considered control



(a)



(b)

Fig. 17. Impact of different control granularities on the power loss of 3-S4 PDN. (a) Total loss increment compared to $T_{\mathrm{on}} = 1\ \mu$s. (b) Total loss reduction over static 3-S1 PDN.

variables are jointly optimized to improve the overall power efficiency according to important interdependencies existing in the regulation chain. Our control policies are further supported with an integrated machine-learning module to cope with fine-grained spatial distributions of workload, achieving further improved power quality and efficiency. We show that the proposed adaptive HVR and control policies reduce system energy by up to 17.9% and 23.9% over a static three-stage HVR and conventional two-stage PDN, respectively.

Acknowledgment

The statements made herein are solely the responsibility of the authors.

References

[1] S. Borkar, "Thousand core chipsa technology perspective," in *Proc. 44th Annu. Design Autom. Conf.*, Jul. 2007, pp. 746–749.

[2] V. Zyuban *et al.*, "IBM POWER8 circuit design and energy optimization," *IBM J. Res. Develop.*, vol. 59, no. 1, pp. 1–9, Feb. 2015.

[3] W. Kim, M. S. Gupta, G.-Y. Wei, and D. Brooks, "System level analysis of fast, per-core DVFS using on-chip switching regulators," in *Proc. IEEE 14th Int. Symp. High Perform. Comput. Archit.*, Feb. 2008, pp. 123–134.

[4] H. Esmaeilzadeh, E. Blem, R. S. Amant, K. Sankaralingam, and D. Burger, "Dark silicon and the end of multicore scaling," in *Proc. 38th Annu. Int. Symp. Comput. Archit. (ISCA)*, Aug. 2011, pp. 365–376.

[5] M. Pedram and J. M. Rabaey, *Power Aware Design Methodologies*. New York, NY, USA: Springer, 2002.

[6] L. Benini, A. Bogliolo, and G. De Micheli, "A survey of design techniques for system-level dynamic power management," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 8, no. 3, pp. 299–316, Jun. 2000.

[7] S. J. Kim *et al.*, "High frequency buck converter design using time-based control techniques," *IEEE J. Solid-State Circuits*, vol. 50, no. 4, pp. 990–1001, Apr. 2015.

[8] I. Vaisband and E. G. Friedman, "Heterogeneous methodology for energy efficient distribution of on-chip power supplies," *IEEE Trans. Power Electron.*, vol. 28, no. 9, pp. 4267–4280, Sep. 2013.

[9] J. F. Bulzacchelli *et al.*, "Dual-loop system of distributed microregulators with high DC accuracy, load response time below 500 ps, and 85-mV dropout voltage," *IEEE J. Solid-State Circuits*, vol. 47, no. 4, pp. 863–874, Apr. 2012.

[10] K. K. Rangan, G.-Y. Wei, and D. Brooks, "Thread motion: Fine-grained power management for multi-core systems," in *Proc. 36th Annu. Int. Symp. Comput. Archit.*, Aug. 2009, pp. 302–313.

[11] E. Rotem, A. Mendelson, R. Ginosar, and U. Weiser, "Multiple clock and voltage domains for chip multi processors," in *Proc. 42nd Annu. IEEE/ACM Int. Symp. Microarchitecture*, Dec. 2009, pp. 459–468.

[12] C. Isci, A. Buyuktosunoglu, C.-Y. Cher, P. Bose, and M. Martonosi, "An analysis of efficient multi-core global power management policies: Maximizing performance for a given power budget," in *Proc. 39th Annu. IEEE/ACM Int. Symp. Microarchitecture*, Dec. 2006, pp. 347–358.

[13] Y. Zu, C. R. Lefurgy, J. Leng, M. Halpern, M. S. Floyd, and V. J. Reddi, "Adaptive guardband scheduling to improve system-level efficiency of the POWER7+," in *Proc. 48th Annual IEEE/ACM Int. Symp. Microarchitecture (MICRO)*, Dec. 2015, pp. 308–321.

[14] V. Zyuban *et al.*, "IBM POWER8 circuit design and energy optimization," *IBM J. Res. Develop.*, vol. 59, no. 1, pp. 1–9, Jan. 2015.

[15] R. Teodorescu and J. Torrellas, "Variation-aware application scheduling and power management for chip multiprocessors," in *Proc. Int. Symp. Comput. Archit.*, vol. 36, no. 3, Jun. 2008, pp. 363–374.

[16] D. Pathak, H. Homayoun, and I. Savidis, "Smart grid on chip: Work load-balanced on-chip power delivery," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 25, no. 9, pp. 2538–2551, Sep. 2017.

[17] H. Li, J. Xu, Z. Wang, P. Yang, R. K. V. Maeda, and Z. Tian, "Adaptive power delivery system management for many-core processors with on/off-chip voltage regulators," in *Proc. Design, Automat. Test Eur. Conf. Exhib. (DATE)*, Mar. 2017, pp. 1265–1268.

[18] W. Lee, Y. Wang, and M. Pedram, "Optimizing a reconfigurable power distribution network in a multicore platform," *IEEE Trans. Comput.-Aided Design Integr.*, vol. 34, no. 7, pp. 1110–1123, Jul. 2015.

[19] A. A. Sinkar, H. R. Ghasemi, M. J. Schulte, U. R. Karpuzcu, and N. S. Kim, "Low-cost per-core voltage domain support for power-constrained high-performance processors," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 22, no. 4, pp. 747–758, Apr. 2014.

[20] M. D. Seeman, V. W. Ng, H.-P. Le, M. John, E. Alon, and S. R. Sanders, "A comparative analysis of Switched-Capacitor and inductor-based DC-DC conversion technologies," in *Proc. IEEE 12th Workshop Control Modeling Power Electron. (COMPEL)*, Jun. 2010, pp. 1–7.

[21] W. Godycki, C. Torng, I. Bukreyev, A. Apsel, and C. Batten, "Enabling realistic fine-grain voltage scaling with reconfigurable power distribution networks," in *Proc. 47th Annu. IEEE/ACM Int. Symp. Microarchitecture*, Dec. 2014, pp. 381–393.

[22] L. Cheng, Y. Liu, and W.-H. Ki, "4.4 A 10/30MHz Wide-duty-cycle-range buck converter with DDA-based Type-III compensator and fast reference-tracking responses for DVS applications," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers (ISSCC)*, Feb. 2014, pp. 84–85.

[23] P. Y. Wu, S. Y. S. Tsui, and P. K. T. Mok, "Area-and power-efficient monolithic buck converters with pseudo-type III compensation," *IEEE J. Solid-State Circuits*, vol. 45, no. 8, pp. 1446–1455, Aug. 2010.

[24] H. K. Krishnamurthy *et al.*, "A digitally controlled fully integrated voltage regulator with on-die solenoid inductor with planar magnetic core in 14nm tri-gate CMOS," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers (ISSCC)*, Feb. 2017, pp. 336–337.

[25] C. Huang and P. K. T. Mok, "An 84.7% efficiency 100-MHz package bondwire-based fully integrated buck converter with precise DCM operation and enhanced light-load efficiency," *IEEE J. Solid-State Circuits*, vol. 48, no. 11, pp. 2595–2607, Nov. 2013.

[26] W. Kim, D. M. Brooks, and G.-Y. Wei, "A fully-integrated 3-level DC/DC converter for nanosecond-scale DVS with fast shunt regulation," in *Proc. IEEE Int. Solid-State Circuits Conf.*, Feb. 2011, pp. 268–270.

[27] E. A. Burton *et al.*, "FIVR—Fully integrated voltage regulators on 4th generation Intel Coreâ„¢ SoCs," in *Proc. IEEE Appl. Power Electron. Conf. Expo.*, Mar. 2014, pp. 432–439.

[28] X. Wang *et al.*, "An analytical study of power delivery systems for many-core processors using on-chip and off-chip voltage regulators," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 34, no. 9, pp. 1401–1414, Sep. 2015.
[29] G. Sizikov, A. Kolodny, E. G. Fridman, and M. Zelikson, "Efficiency optimization of integrated DC-DC buck converters," in *Proc. 17th IEEE Int. Conf. Electron., Circuits Syst.*, Dec. 2010, pp. 1208–1211.
[30] S. Lai and P. Li, "A fully on-chip area-efficient CMOS low-dropout regulator with fast load regulation," *Analog Integr. Circuits Signal Process.*, vol. 72, no. 2, pp. 433–450, Aug. 2012.
[31] M. Ware *et al.*, "Architecting for power management: The IBM POWER7 approach," in *Proc. 16th Int. Symp. High-Perform. Comput. Archit.*, Jan. 2010, pp. 1–11.
[32] X. Liu, S. Sun, X. Li, H. Qian, and P. Zhou, "Machine learning for noise sensor placement and full-chip voltage emergency detection," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 36, no. 3, pp. 421–434, Mar. 2017.
[33] T. Wang, C. Zhang, J. Xiong, and Y. Shi, "Eagle-Eye: A near-optimal statistical framework for noise sensor placement," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design (ICCAD)*, Nov. 2013, pp. 437–443.
[34] H. Lin and P. Li, "Relevance vector and feature machine for statistical analog circuit characterization and built-in self-test optimization," in *Proc. 53nd ACM/EDAC/IEEE Design Automat. Conf. (DAC)*, Jun. 2016, p. 11.
[35] H. Lin, A. M. Khan, and P. Li, "Statistical circuit performance dependency analysis via sparse relevance kernel machine," in *Proc. IEEE Int. Conf. IC Design Technol. (ICICDT)*, May 2017, pp. 1–4.
[36] N. Binkert *et al.*, "The gem5 simulator," *ACM SIGARCH Comput. Archit. News*, vol. 39, no. 2, pp. 1–7, 2011.
[37] S. Li, J. H. Ahn, R. D. Strong, J. B. Brockman, D. M. Tullsen, and N. P. Jouppi, "McPAT: An integrated power, area, and timing modeling framework for multicore and manycore architectures," in *Proc. 42nd Annu. IEEE/ACM Int. Symp. Microarchitecture*, Dec. 2009, pp. 469–480.
[38] C. Bienia and K. Li, *Benchmarking modern multiprocessors*. Princeton, NJ, USA: Univ. Princeton, 2011.
[39] M. S. Gupta, J. L. Oatley, R. Joseph, G.-Y. Wei, and D. M. Brooks, "Understanding voltage variations in chip multiprocessors using a distributed power-delivery network," in *Proc. Design, Automat. Test Eur. Conf. Exhib.*, Apr. 2007, pp. 1–6.
[40] C.-H. Chan, Y. Zhu, I.-M. Ho, W.-H. Zhang, U. Seng-Pan, and R. P. Martins, "16.4 A 5mW 7b 2.4GS/s 1-then-2b/cycle SAR ADC with background offset calibration," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC)*, Feb. 2017, pp. 282–283.
[41] B. Verbruggen, J. Craninckx, M. Kuijk, P. Wambacq, and G. Van-Der-Plas, "A 2.6 mW 6b 2.2GS/s 4-times interleaved fully dynamic pipelined ADC in 40 nm digital CMOS," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC)*, Feb. 2010, pp. 398–400.
[42] H. Akkary, R. Rajwar, and S. T. Srinivasan, "Checkpoint processing and recovery: Towards scalable large instruction window processors," in *Proc. 36th Annual IEEE/ACM Int. Symp. Microarchitecture*, Dec. 2003, p. 423.
[43] K. A. Bowman *et al.*, "A 16 nm all-digital auto-calibrating adaptive clock distribution for supply voltage droop tolerance across a wide operating range," *IEEE J. Solid-State Circuits*, vol. 51, no. 1, pp. 8–17, Jan. 2016.

Xin Zhan received the bachelor's degree in electronic science and technology and the master's degree in microelectronics and solid-state electronics from Huazhong University of Science and Technology, Wuhan, China, in 2011 and 2014, respectively, and the Ph.D. degree in computer engineering from Texas A&M University, College Station, TX, USA, in 2019.

His current research interests include computer-aided design of very large-scale integrated circuits and optimization of the on-chip regulated power delivery networks.

Dr. Zhan was a recipient of the Best Paper Award at the IEEE/ACM Design Automation Conference in 2016.

Jianhao Chen received the bachelor's degree in applied physics from the University of Science and Technology of China, Hefei, China, in 2018. He is currently working toward the Ph.D. degree in computer engineering at Texas A&M University, College Station, TX, USA.

His current research interest is the optimization of on-chip power delivery networks.

Edgar Sánchez-Sinencio (F'92–LF'10) was born in Mexico. He received the Professional degree in communications and electronic engineering from the National Polytechnic Institute of Mexico, Mexico, in 1966, the M.S.E.E. degree from Stanford University, Stanford, CA, USA, in 1970, and the Ph.D. degree from the University of Illinois at Champaign–Urbana, Champaign, IL, USA, in 1973.

He is currently the University Distinguished Professor, Texas Instruments Jack Kilby Chair Professor, and the Director with the Analog and Mixed-Signal Center, Texas A&M University, College Station, TX, USA. He has coauthored six books on different topics, such as RF circuits, low-voltage low-power analog circuits, and neural networks. He has graduated 61 M.Sc. and 56 Ph.D. students. His current interests include the area of ultralow-power analog circuits, RF circuits, harvesting techniques, power management, and medical electronics circuit.

Dr. Sánchez-Sinencio was a member of the IEEE Solid-State Circuits Society Fellow Award Committee from 2002 to 2004 and is a Fellow of the Institution of Engineering and Technology—the largest multidisciplinary professional engineering institution in the world. He was a recipient of the Texas Senate Proclamation #373 for Outstanding Accomplishments in 1996, the IEEE Circuits and Systems Society Golden Jubilee Medal in 1999, and the Prestigious IEEE Circuits and Systems Society 2008 Charles A. Desoer Technical Achievement Award. He received the Honoris Causa Doctorate (the first honorary degree awarded for microelectronic circuit-design contributions) from the National Institute for Astrophysics, Optics and Electronics, Puebla, Mexico, in 1995. He was a co-recipient of the 1995 Guillemin–Cauer Award for his work on cellular networks and the 1997 Darlington Award for his work on high-frequency filters. He was the IEEE Circuits and Systems Society's Representative to the IEEE Solid-State Circuits Society from 2000 to 2002. He served as a Guest Editor of the Analog Section of the IEEE JOURNAL OF SOLID-STATE CIRCUITS (JSSC) Special Issue in December 2016 and a Co-Guest Editor of the Special Issue on Circuits and Systems for the Internet of Things—From Sensing to Sensemaking in 2017. He served as the Editor-in-Chief for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—PART II: EXPRESS BRIEFS. He served as the Vice President of IEEE Circuits and Systems Society's Publications and as a Distinguished Lecturer of the IEEE Circuit and Systems Society from 2012 to 2013.

Peng Li (S'02–M'04–SM'09–F'16) received the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 2003.

He is currently a Professor with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA. His current research interests include integrated circuits and systems, computer-aided design, brain-inspired computing, computational brain modeling, machine learning, and its hardware realization in VLSI.

Dr. Li's work has been recognized by various distinctions including four IEEE/ACM Design Automation Conference Best Paper Awards, the ISCAS Honorary Mention Best Paper Award from the Neural Systems and Applications Technical Committee of IEEE Circuits and Systems Society, the IEEE/ACM William J. McCalla ICCAD Best Paper Award, the U.S. National Science Foundation CAREER Award, two Inventor Recognition Awards from Microelectronics Advanced Research Corporation, two Semiconductor Research Corporation Inventor Recognition Awards, the William and Montine P. Head Fellow Award, the TEES Fellow Award, and the Eugene E. Webb Fellow Award from the College of Engineering, Texas A&M University. He was an Associate Editor for the IEEE TRANSACTIONS ON COMPUTER-AIDED DESIGN OF INTEGRATED CIRCUITS AND SYSTEMS from 2008 to 2013 and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—PART II: EXPRESS BRIEFS from 2008 to 2016.