IEEE Asian Solid-State Circuits Conference
November 9-11,2015 / Xiamen, Fujian, China

A 934MHz 9Gb/s 3.2pJ/b/iteration Charge-Recovery LDPC Decoder with In-Package Inductors

Tai-Chuan Ou, Zhengya Zhang, Marios C. Papaefthymiou
University of Michigan, Ann Arbor, MI
Email: terryou@umich.edu

***Abstract*—A 576-bit LDPC decoder is designed using a charge-recovery logic family and in-package inductors. The decoder test-chip is fabricated in a 65nm CMOS flip-chip process. Unlike all previously published high-performance charge-recovery chips, which use on-chip inductors to recover charge from parasitic capacitance, this charge-recovery design uses in-package inductors, avoiding the area overheads of on-chip inductors and achieving higher quality factors. Specifically, charge recovery is performed using 16 high-quality inductors that have been embedded in a custom-designed 6-layer FC-BGA package, significantly improving the area efficiency and energy consumption of the design compared to alternative implementations with on-chip inductors. When operating at 934MHz, the decoder consumes 3.2pJ/b/iteration to deliver a throughput of 9Gb/s at 10 decoding iterations.**

I. Introduction

A 576-bit, rate-5/6 low-density parity-check (LDPC) decoder has been designed using charge-recovery circuitry with in-package inductors. Charge-recovery circuits [1]–[3] rely on inductive elements to recover charge from gate fanouts and achieve energy-efficient operation. Previous GHz-speed high-performance charge-recovery chips have relied on integrated on-chip inductors at the cost of silicon area overhead. In addition to area overhead, on-chip inductors also suffer from relatively low quality factors (Q), typically less than 10 [4], [5], due to resistive on-chip metals. For charge-recovery designs operating at lower frequencies (hundreds of MHz or lower), high Q discrete inductors have been used, incurring extra board area and extra costs. This paper presents the first-ever charge-recovery test-chip with in-package inductors offering high Q inductors without area overhead.

The decoder described in this paper is shown in Figure 1. The LDPC test-chip is fabricated in a 65nm flip-chip process. The inductors required by charge-recovery logic are embedded in a custom-designed 6-layer flip-chip–ball-grid-array (FC-BGA) package substrate, achieving better Q than their on-chip counterparts thanks to the availability of thicker and therefore less resistive copper in the package substrate. Inductors in the package are connected to the charge-recovery network on the die through 48 flip-chip bumps, eliminating the need for bonding wires and their parasitics to enable efficient charge recovery. Correct operation has been verified from 624MHz to 1.08GHz. When operating at 934MHz, the chip consumes 286mW, yielding energy consumption of 3.19pJ/b/iteration and an area efficiency of 5.83Gbps/mm^2, improving on results in [5]–[8] by at least 2.3× in energy consumption with similar or better area efficiency.

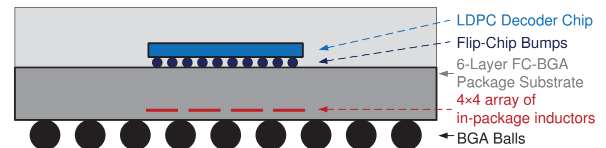

Fig. 1: Charge-recovery LDPC decoder: includes a LDPC decoder chip fabricated in 65nm flip-chip technology and a custom-designed FC-BGA package substrate with 16 in-package inductors.

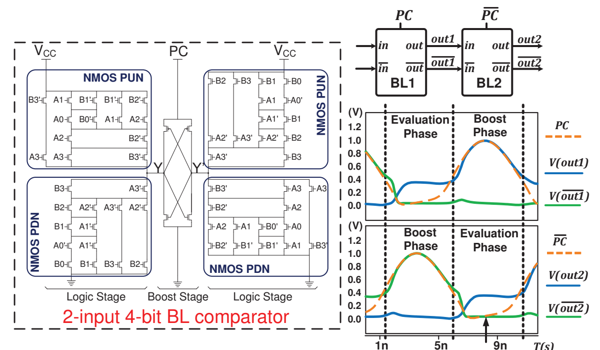

Fig. 2: Schematic of a BL two-input 4-bit comparator gate, cascade of BL gates, and the corresponding output waveforms.

II. Boost Logic (BL)

The LDPC test-chip is designed using a charge-recovery logic family called boost logic (BL) [1], [3]. Figure 2 shows the schematic of a BL gate. The structure of a BL gate can be divided into two parts, logic stage and boost stage. The logic stage is similar to static CMOS logic with a DC supply V_{CC}, except that NMOS devices are used in the pull-up network (PUN), instead of PMOS, to achieve gate-overdrive. From a functional standpoint, the logic stage performs functional evaluation and develops an initial voltage difference between the dual-rail outputs. The boost stage is composed by a pair of cross-coupled inverters with a charge-recovery supply PC. After the logic stage performs functional evaluation, the boost stage then amplifies this initial voltage difference by following the rising transition of the power-clock at pin PC. As the power-clock falls, charge is recovered from the output through PC, achieving energy-efficient operation. PC not only provides charge to internal circuit nodes but also synchronizes the computation of the gate, which is the reason it is called power-clock. Cascades of BL gates are formed by connecting the gates with alternating power-clock phases, PC and $\overline{PC}$.

978-1-4673-7191-9/15/$31.00 ©2015 IEEE

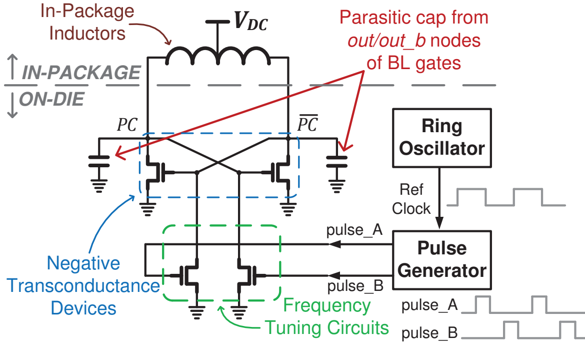

Fig. 3: Blip power-clock generator, formed by in-package inductors, cross-coupled internal drivers, and frequency tuning circuits with pulses generated by a ring oscillator and a pulse generator.

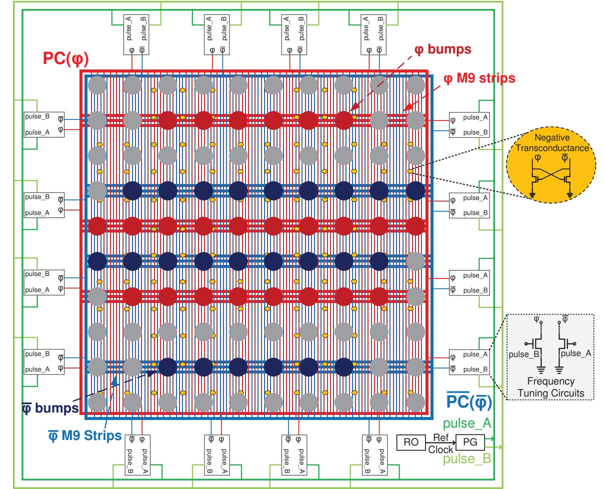

Fig. 4: Power-clock generation and distribution: clock mesh network for PC and $\overline{PC}$, 100 flip-chip bumps, 16 frequency tuning circuits, and 144 negative transconductance devices.

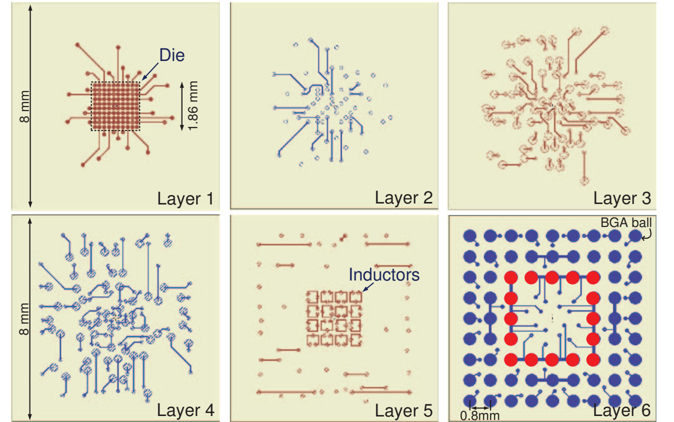

Fig. 5: Custom-designed 6-layer FC-BGA package substrate.

III. Power-Clock Generation and Distribution

The two-phase power-clock required by BL gates is generated using a distributed version of the blip power-clock generator [9], shown in Figure 3. The power-clock generator consists of on-die and in-package components. 144 distributed on-die negative transconductance devices, 16 on-die frequency tuning circuits, an on-die ring oscillator (RO), and an on-die pulse generator (PG), along with 16 inductors in the package are used to generate the power-clock by resonating the parasitic capacitance of the power-clock distribution network and the BL gate fanouts. To enable frequency tuning, the RO generates a reference clock with a desired frequency feeding to the PG, and the PG then outputs a pair of 180-degree out-of-phase pulses with programmable duty cycle, achieving frequency scaling and forcing the power-clock to run at the desired frequency.

Figure 4 shows the on-chip power-clock distribution network. 48 flip-chip bumps (24 for each clock phase) are used to connect the two-phase power-clock, PC and $\overline{PC}$, from the 16 inductors in the package substrate to the on-chip power-clock distribution network. An additional 48 bumps are for supplies, V_{SS}, and signals for testing the decoder chip. PC and $\overline{PC}$ bumps are placed right on top of the top-level metal of the clock meshes, enabling efficient recovery. Clock meshes are employed to distribute the power-clock using top-level metals to minimize clock skew, allowing custom-designed BL dynamic cells to be easily connected to the two-phase power-clock using commercial EDA tools. To distribute the power-clock from the mesh to the PC pin of each BL gate, each standard-cell row has two metal-3 strips reserved for delivering PC and $\overline{PC}$ to the BL gates. These strips are tied to top-level clock meshes, allowing the routing of the power-clock network using an automatic place-and-route tool and avoiding large clock skew. 144 negative transconductance devices are distributed across the core to maintain the oscillation. To operate the design off-resonance, a pair of 180-degree out-of-phase pulses at the target frequency is distributed to the 16 frequency tuning circuits around the core using a tree structure with supply and ground shielding.

IV. Package Substrate Design

The FC-BGA substrate is manufactured through a 6-layer build-up 2-2-2 manufacturing process. The thickness of the copper is 15μm, which helps significantly with the Q of the in-package inductors compared to their on-chip counterparts (3.4μm for ultra-thick metal).

Figure 5 shows the layer-by-layer view of our package substrate design. The substrate occupies 8mm$\times$8mm, and the die occupies 1.86mm$\times$1.86mm with a 10$\times$10 array of flip-chip bumps. Layer 1 of the substrate is used for flip-chip bump connections. Layers 2 to 4 are mainly for routing. 16 horizontal in-package center-tap coil inductors are designed on Layer 5. Layer 6 is reserved for BGA ball connections. Note that the size of the substrate is dominated by the number of BGA balls (for supplies, V_{SS}, and I/Os) required for the decoder, not by the inductors. The amount of area taken by the inductors on Layer 5 is comparable with the area of the decoder chip.

Inductor design plays a key role in the efficiency of charge-recovery chips. The 16 in-package coil inductors have been carefully designed and characterized using a commercial 3D full-wave electromagnetic field solver tool. Figure 6 shows the dimensions and the specifications of one of the 16 inductors. When simulated at 1GHz, each inductor has 969.4pH inductance with a Q factor of 33.4, achieving 3.6$\times$ improvement in inductor Q compared to [5] with similar operating conditions.

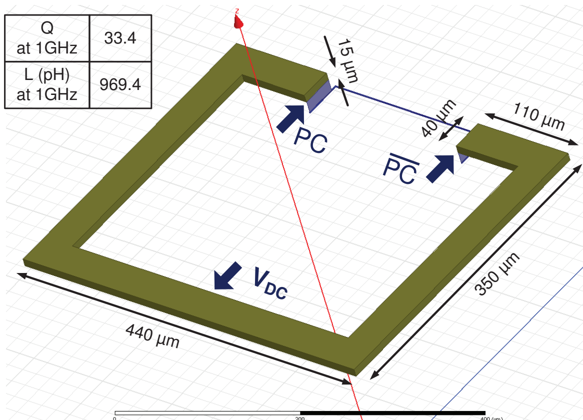

Fig. 6: One of the 16 inductors designed in the package substrate, and the HFSS simulation results.

Designing the coil inductors in vertical orientation has also been considered, but process variation and misalignments between layers during the build-up manufacturing process make it difficult to design coils with a precise inductance value and high Q.

To avoid eddy currents, which are created by loops around the inductors and degrade efficiency, metal traces in the package are carefully designed to avoid any possible loop around these inductors in the package substrate. Bumps that connect the center-tap point of the inductors to the supply V_{DC} (highlighted in red, as shown in Figure 5) are routed to the supply on PCB board in a manner that encloses a much bigger loop, instead of connecting closely and forming a small loop, to reduce eddy currents. Loops in the power grid of the decoder have not been eliminated, as they are $320\mu m$ away from the inductors, which are located on the fifth layer of the package substrate. In simulations, the worst-case degradation of Q factor and inductance due to the power grid loops in the die is less than 5%.

V. LDPC Decoder Design

The 576-bit, rate-5/6 charge-recovery LDPC decoder adopts a block-parallel architecture [5]. Figure 7 shows the decoder architecture for the LDPC code specified by the IEEE 802.16e standard. The code matrix is partitioned into 4 blocks so that complex and long global interconnects are replaced with relay local interconnects between neighboring blocks. Two columns in the code matrix (shown in red) are swapped for regular partitioning. The min-sum decoding consists of check node operations and variable node operations. Starting from the check node operation on the first row of Block 1, the decoder then relays the results to Blocks 2, 3, and 4 in order. Following the check node operation on the first row of Block 1, the decoder performs the variable node operation on it while Block 1 begins the check node operation on the second row in parallel. For complete check node and variable node operations in all 4 blocks, one decoding iteration takes 24 cycles (48 phases). Due to this deeply-pipelined relay architecture, the decoder is able to process 4 streams in parallel without any pipeline stalls.

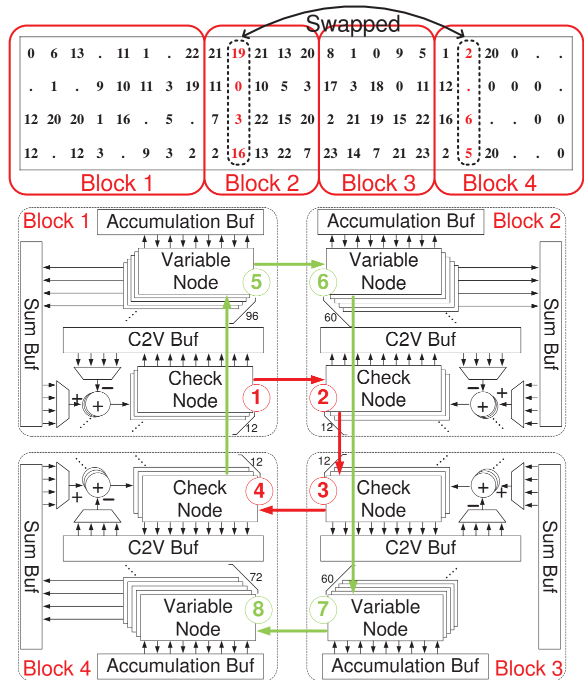

Fig. 7: 576-bit rate-5/6 IEEE 802.16e LDPC matrix, and decoder architecture

VI. Chip Implementation and Measurement Results

The chip has been fabricated in a 65nm CMOS flip-chip process. The charge-recovery LDPC decoder logic occupies 1.54mm^2. Correct function has been validated for clock frequencies ranging from 624MHz to 1.08GHz. Figure 8 shows measured energy per cycle versus operating frequency. Minimum energy consumption is 306.5pJ per cycle when the power-clock is operating at a frequency of 934MHz, dissipating 286.4mW of power at room temperature. With frequency tuning circuits turned off, the minimum energy consumption of the decoder in self-resonant mode is 398pJ per cycle at 875MHz. Note that, unlike [5], the minimum energy consumption operating point does not occur when the decoder chip is operating in self-resonant mode. We surmise that this might be the result of insufficiently many negative transconductance devices, resulting in reduced power-clock voltage swing and thus requiring higher V_{DC} and V_{CC} values. When the frequency tuning circuits are turned on to run the decoder off-resonance, they replenish the energy lost due to the resistance of the distribution network to maintain the oscillation, acting as extra negative transconductance devices and enabling the scaling of V_{DC} and V_{CC} to lower values. The energy efficiency of this chip is higher than that of the chip reported in [5] due to the improved Q of the in-package inductors, which increases the efficiency of charge recovery, and the lower resistance of the flip-chip connections (compared to bondwires), which allows for further supply voltage scaling and thus lower energy consumption.

Figure 9 gives the performance characteristics of the chip in this work and compares it with the most recently reported high-throughput LDPC decoders. The charge-recovery LDPC decoder chip in this paper outperforms state-of-the-art designs of

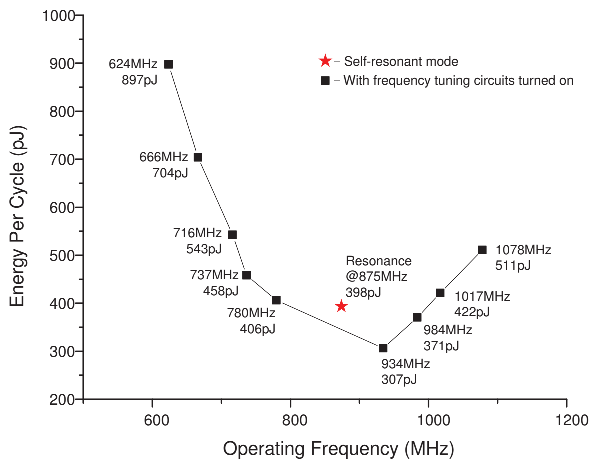

Fig. 8: Measured energy per cycle versus operating frequency.

	This Work	JSSC'12 [6]	VLSI'12 [7]	ISSCC'14 [5]	ISSCC'14 [8]
Technology	65nm bulk	65nm bulk	65nm bulk	65nm bulk	28nm FDSOI
Code Length	576	672	672	576	672
Code Rate	5/6	7/8	1/2	5/6	1/2
Core Area (mm^2)	1.54	1.56	1.6	1.54 (2.34 w/ inductors)	0.63
Frequency (MHz)	934	197	540	821	260
Iterations	10	5	10	10	3.75
Throughput (Gbps)	8.97	5.79	9	7.88	12
Power (mW)	286.4	361	782.9	576.8	180
Energy Consumption (pJ/bit/iteration)	3.19	12.48	8.95	7.32	8 [1]
Area Efficiency (Gbps/mm^2)	5.83	2.12 [3]	5.63	5.12 (3.35 w/ inductors)	7.14 [2, 3]

[1] Energy consumption for 28nm (not normalized to 65nm)
[2] Area efficiency for 28nm (not normalized to 65nm)
[3] Normalized to 10 iterations

Fig. 9: Chip summary and comparison with state-of-the-art designs.

comparable code length, complexity, and throughput [5]–[8], achieving at least 2.3× lower energy consumption compared to all designs, while having better area efficiency compared to the ones fabricated in the same technology. Even without any normalization for different process nodes, the area efficiency of this chip is still comparable with that of the chip in [8], which was fabricated in a more advanced 28nm technology.

this work still has a comparable area efficiency as [8], which was fabricated in a more advanced 28nm technology.

A die microphotograph is shown in Figure 10. A built-in-self-test (BIST) circuit that is used to generate and process the input and output of the decoder, along with RO, PG, and frequency-tuning circuits are implemented with static CMOS logic and are distributed around the decoder core.

VII. Conclusion

This paper proposes and explores the use of in-package inductors to improve upon the quality and eliminate the area overhead of on-chip inductors in high-performance charge-recovery designs. The design of a 576-bit charge-recovery LDPC decoder with in-package inductors is presented, including a custom-designed 6-layer FC-BGA package substrate and

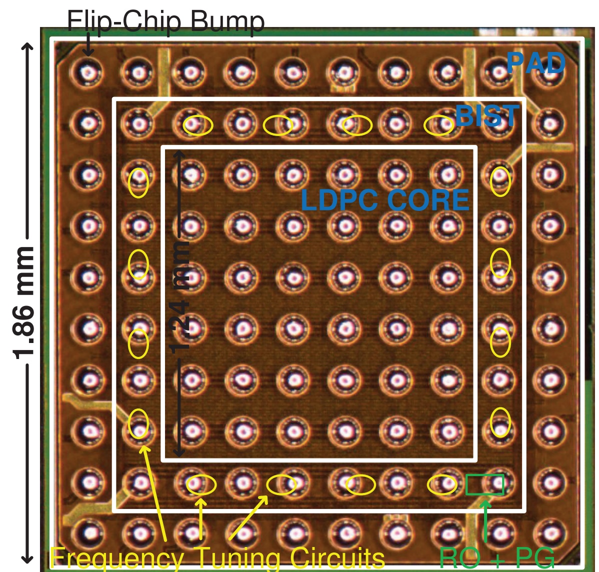

Fig. 10: Microphotograph of the charge-recovery LDPC test chip in 65nm CMOS.

a test-chip fabricated in a 65nm CMOS flip-chip process. 16 inductors are designed on the fifth layer of the 6-layer FC-BGA package substrate so that the need for on-chip inductors is eliminated. In-package inductors also provide better quality factors than on-chip ones, improving area efficiency and energy consumption. When operating at 934MHz, the decoder reaches a 9Gb/s throughput, consuming 286mW, or 3.2pJ/b/iteration, improving on the state-of-the-art published results by at least 2.3× in energy consumption with similar or even better area efficiency.

Acknowledgments

This work was supported in part by NSF under grant No. CCF-1320027. The authors thank Joohee Kim, Ye-Sheng Kuo, and Shih-Kai Wang for their valuable contributions to this work.

References

[1] V. Sathe *et al.*, "Energy-efficient GHz-class charge-recovery logic," *IEEE J. Solid-State Circuits*, vol. 42, no. 1, pp. 38–47, Jan 2007.

[2] J. Kao *et al.*, "A charge-recovery 600MHz FIR filter with 1.5-cycle latency overhead," *ESSCIRC*, pp. 160–163, Sept 2009.

[3] W.-H. Ma *et al.*, "187MHz subthreshold-supply charge-recovery FIR," *IEEE J. Solid-State Circuits*, vol. 45, no. 4, pp. 793–803, April 2010.

[4] V. Sathe *et al.*, "Resonant clock design for a power-efficient high-volume x86-64 microprocessor," *ISSCC*, pp. 68–70, Feb 2012.

[5] T.-C. Ou *et al.*, "An 821MHz 7.9Gb/s 7.3pJ/b/iteration charge-recovery LDPC decoder," *ISSCC*, pp. 462–463, Feb 2014.

[6] S.-W. Yen *et al.*, "A 5.79-Gb/s energy-efficient multirate LDPC codec chip for IEEE 802.15.3c applications," *IEEE J. Solid-State Circuits*, vol. 47, no. 9, pp. 2246–2257, Sept 2012.

[7] Y. S. Park *et al.*, "A 1.6-mm^2 38-mW 1.5-Gb/s LDPC decoder enabled by refresh-free embedded DRAM," *Symp. VLSI Circuits*, pp. 114–115, June 2012.

[8] M. Weiner *et al.*, "A scalable 1.5-to-6Gb/s 6.2-to-38.1mW LDPC decoder for 60GHz wireless networks in 28nm UTBB FDSOI," *ISSCC*, pp. 464–465, Feb 2014.

[9] W. Athas *et al.*, "A resonant signal driver for two-phase, almost-non-overlapping clocks," *IEEE International Symposium on Circuits and Systems*, vol. 4, pp. 129–132, May 1996.