Achieving Ultralow Standby Power With an Efficient SCCMOS Bias Generator

Yoonmyung Lee, *Member, IEEE*, Mingoo Seok, *Member, IEEE*, Scott Hanson, *Member, IEEE*, Dennis Sylvester, *Fellow, IEEE*, and David Blaauw, *Fellow, IEEE*

***Abstract*—Standby power frequently dominates the power budget of battery-operated ultralow power sensor nodes. Reducing standby power is therefore a key challenge for further power reduction. Applying known circuit techniques for standby power reduction is challenging since standby power of state-of-the-art sensor node systems is now on the order of nanowatts or less. Hence, the overhead of any leakage reduction technique quickly overshadows any gains. This brief proposes an efficient implementation method for super cutoff CMOS that exploits the unique conditions of power gating to enable a highly efficient charge pump design. The proposed techniques are applied to logic blocks and memory devices. For a very low initial standby power value of tens of picowatts, standby power reduction of up to 19.3× and 29% is achieved for logic blocks and memory, respectively.**

***Index Terms*—Charge pump, leakage current, sensor node, SRAM, standby power, ultralow power.**

I. Introduction

SIZE is a critical concern for ultralow power sensor systems, particularly for medical applications requiring implantation. An extremely small sensor platform, as small as a few cubic millimeters, is needed to provide intelligent controls and temporary storage for recently developed microsensors such as [1] and [2]. Since power source size is restricted in these applications, ultralow average power consumption, on the order of nanowatts (nW) and picowatts (pW), is required for these sensor nodes to sustain their activity for a reasonable lifetime without battery replacement. Similar power budgets are needed to enable energy autonomy with volume-limited energy harvesting, as demonstrated in [2].

Many circuit techniques such as supply voltage scaling [3] have been proposed to minimize active mode energy. However, many sensor systems spend much more time in standby mode than active mode. Therefore, it is critical that designers not neglect the power consumed in this standby mode since standby power can dominate the system budget. Recent work [4] has shown that a better balance between active mode power and standby mode power can be achieved by designing a system with standby power as a primary constraint. With architectural/circuit techniques shown in [4], standby power is reduced to tens of pW, giving ∼one-year lifetime with a 1-mm^3 system size, including the battery.

However, even with the sleep strategies presented in [4], standby power is still a dominant ($> 75\%$) source of total power consumption. The standby power consists of two components: the power consumed by circuits that are turned off (power gated) during standby mode and power consumed by circuits that must retain state and remain turned on (e.g., memory). Therefore, developing new techniques for reducing each type of standby power is the key challenge to extend the lifetime of ultralow power applications.

However, reducing the standby power for circuits that already consume only tens of pW is very challenging for several reasons.

1) The power overhead for using any leakage reduction technique itself must be a few pW in order to be beneficial.
2) Since these systems are typically battery operated, only a single supply voltage is available.
3) Any locally generated voltages for power reduction that are greater than the power supply voltage V_{DD} or less than the ground voltage should be controlled without level converters or other switches that introduce new leakage paths.

In this brief, we propose a standby power reduction technique that can be applied to ultralow power microsystems. First, we explore the use of super cutoff CMOS (SCCMOS) for reducing standby power in power-gated blocks. Using the unique requirements of the required bias voltages, a novel charge pump is proposed that can operate with ultralow power and extremely low frequency (< 10 Hz) for efficient SCCMOS bias generation. Next, we investigate leakage paths in memory and propose a leakage reduction strategy based on SCCMOS to reduce bit-line leakage as well.

II. Efficient Super Cutoff Power Gating

Multithreshold-voltage CMOS (MTCMOS) [5] is a power-gating technique that reduces leakage in standby mode with high-threshold-voltage $V_{\rm th}$ headers/footers, as shown in Fig. 1. SCCMOS [6] is a similar power-gating technique that further reduces the standby leakage current by underdriving headers or footers. Gate underdriving in SCCMOS can reduce subthreshold leakage current by approximately two orders of magnitude or more (see Fig. 2). Therefore, for a given leakage current budget, wider header/footer can be used to reduce the voltage drop across header/footer in active mode and corresponding performance penalty. Moreover, regular-$V_{\rm th}$ (RVT) devices can be used as header/footer instead of HVT devices to reduce performance penalty as well. Fig. 2 shows that, at low operating

Manuscript received July 5, 2013; accepted August 7, 2013. Date of publication October 9, 2013; date of current version December 24, 2013. This brief was recommended by Associate Editor K.-H. Chen.

The authors are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: sori@umich.edu).

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

Digital Object Identifier 10.1109/TCSII.2013.2281919

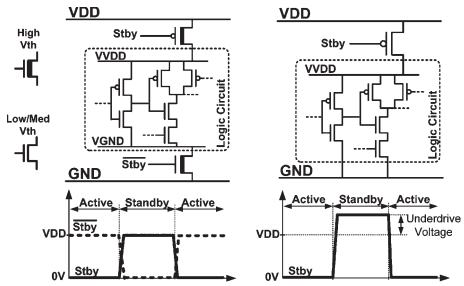

Fig. 1. Concept of (left) MTCMOS and (right) SCCMOS.

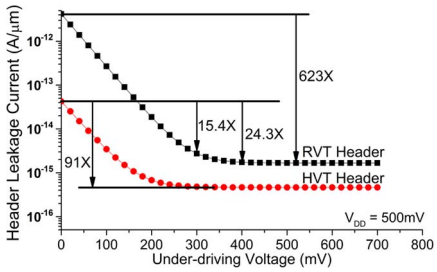

Fig. 2. SCCMOS header leakage current with RVT and HVT headers.

voltage of 500 mV, which is typical in ultralow power systems, an RVT header with sufficient underdriving (> 300 mV) incurs less leakage than an HVT header without underdriving.

Although SCCMOS is an efficient leakage reduction technique, efficient generation of underdriving voltage is challenging. For example, the gate bias generator in [6], consumed 50 nW, whereas the leakage power of power-gated logic circuits was on the order of pW. This imbalance between bias generator power and leakage power results in a system whose standby power is dominated by the bias generator. Therefore, to apply the SCCMOS scheme to a microsystem with sub-nW leakage power, an ultralow power charge pump whose power overhead is on the order of pW, or $1000\times$ lower than [6], is needed.

Fig. 3 shows the Dickson charge pump that is used for bias generation in [6]. Current flow in this charge pump can be represented as two alternating currents in each stage, i.e., charging and leakage currents. Fig. 3 shows the steady-state voltage of each node in two alternating phases. In phase 0, positive V_{GS} and V_{DS} of Δ are formed across transistor M_1, resulting in a charging current transferring charge in the desired direction. In Phase 1, clock Φ is raised, which couples up the V_1 node voltage, resulting in $V_{GS} = 0$ and $V_{DS} = 0.5 - \Delta$ for M_1. Therefore, charge stored in C_{stg1} will slowly leak away to V_{DD} through M_1 subthreshold leakage, resulting in the steady-state condition

$$I_{\mathrm{charged},i} - I_{\mathrm{leak},i} = I_{\mathrm{charged},i+1} - I_{\mathrm{leak},i+1} = \cdots = I_{\mathrm{load}}$$

where $I_{\mathrm{charge},i}$ and $I_{\mathrm{leak},i}$ denote the average charging and leakage current of ith stage during one clock cycle. In conventional charge pumps, I_{leak} is negligible compared with I_{charge}, and I_{charge} is of a comparable order as I_{load}; hence, most efforts to improve charge pumps have focused on making charging more efficient, i.e., maximizing I_{charge} in each cycle, such as by aiding charge transfer switch [8] or reducing the effect of reverse body bias (RBB) [9] rather than minimizing I_{leak}.

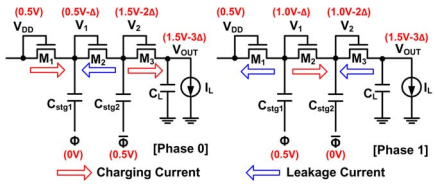

Fig. 3. Current flow in each phase of Dickson charge pump [9] where V_{DD} = clock swing $= 0.5$ V.

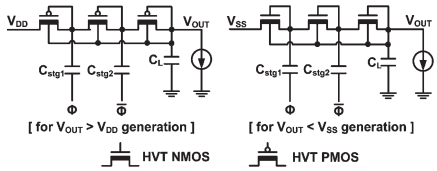

Fig. 4. Circuit diagram of proposed self-reverse body biasing charge pumps for (left) $V_{\mathrm{OUT}} > V_{DD}$ and (right) $V_{\mathrm{OUT}} < V_{SS}$.

Meanwhile, SCCMOS bias generation has two unique conditions, which we can take leverage for efficient bias-voltage generation.

1) *Load current for SCCMOS bias generation charge pump is near 0.*
   Unlike typical charge pumps, the SCCMOS bias generator only needs to drive the gate of footer/header through simple control logic, (e.g., inverter). Therefore, I_{load} is now comparable to I_{leak}, and minimizing I_{leak} is at least as important as maximizing I_{charge} for SCCMOS.
2) *Output voltage of bias generation charge pump does not need to exceed $V_{DD} + 300$ mV.*
   The reduction of leakage current with gate underdriving voltage saturates at $\sim$200 mV (HVT) or $\sim$300 mV (RVT), (Fig. 2). Therefore, techniques to reduce V_{th} drop in each stage and enhance the bias voltage [11], [12] can be avoided, allowing a reduction in power overhead.

The proposed charge pump is shown in Fig. 4, which exploits these two conditions. In the proposed charge pump, $V_{\mathrm{OUT}} > V_{DD}$ generation is obtained by using pMOS transistors with body terminals connected to the output V_{OUT}. This configuration creates a negative feedback of RBB, which applies weak RBB for pMOS transistors when V_{OUT} is low and applies strong RBB when sufficient V_{OUT} is developed in steady state. Therefore, fast charge pumping with high I_{charge} is achieved when the V_{OUT} is still low. However, efficient operation with low I_{leak} is achieved in steady state when the V_{OUT} is sufficiently high. The strong RBB in steady state hinders I_{charge}, resulting in lower steady-state V_{OUT}, which is undesirable for typical charge pumps. However, this is acceptable for SCCMOS bias generation since sufficient V_{OUT} (> 300 mV) is already developed at this point. To apply the technique of self-RBB to nMOS footers, nMOS transistors are used for $V_{\mathrm{OUT}} < V_{SS}$ generation. In addition, HVT transistors are used for both types of pumps to minimize overall leakage.

Fig. 5 shows the simulated power consumption of the proposed and Dickson charge pumps as a function of output voltage where $V_{DD} = 500$ mV and the number of stages is two. Simulated charge pumps are shown in Fig. 3 (Dickson) and in Fig. 4 (proposed), and identically sized transistors and capacitors are used for comparison. To approximate the SCCMOS scenario, only self-loading is considered ($I_{\mathrm{load}} = 0$).

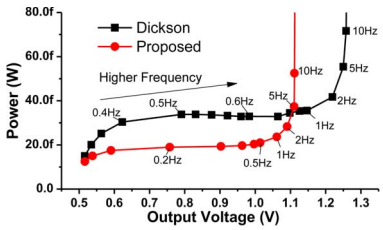

Fig. 5. Power versus output voltage at steady state for Dickson and proposed charge pumps simulated with zero load current.

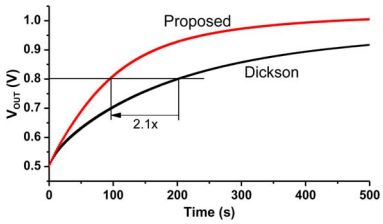

Fig. 6. Initial charging waveform of Dickson and proposed charge pumps simulated with same pumping frequency.

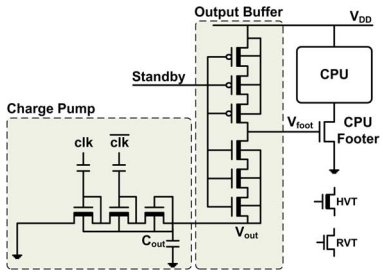

Fig. 7. Proposed circuit for reducing standby power of power-gated blocks.

For ultralow power operation, charge pumps must be clocked at very low clock frequencies. To obtain sufficient SCCMOS bias voltage ($V_{\rm OUT} > 800$ mV), the proposed charge pump can run with a lower frequency (0.3 Hz) compared with the Dickson charge pump (0.5 Hz). At $V_{\rm OUT} = 800$ mV, the proposed charge pump consumes 43% less power than the Dickson charge pump. As expected, the proposed charge pump saturates at a lower output voltage at a higher frequency (> 5 Hz). This is due to stronger RBB and an HVT drop on each stage. Fig. 6 shows how quickly the proposed and the Dickson charge pumps can develop $V_{\rm OUT}$. With a fixed clock frequency of 0.6 Hz, the proposed charge pump reaches 800 mV twice as fast as the Dickson charge pump due to alleviated RBB during initial output development.

III. Standby Power Reduction for Logic Blocks

Standby power of large logic blocks can be reduced with efficient implementation of SCCMOS with the charge pump described in Section II. Fig. 7 shows the proposed standby power reduction scheme with the nMOS charge pump to generate a negative super cutoff voltage, coupled to an output driver to switch the gate voltage on the footer $V_{\rm foot}$ between the output voltage $V_{\rm out}$ in standby mode and V_{DD} in active mode. For the charge pump, metal–insulator–metal (MIM) capacitors are used to minimize parasitic capacitance values.

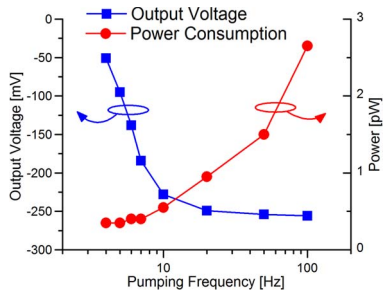

Fig. 8. Measured super cutoff voltage and charge pump power.

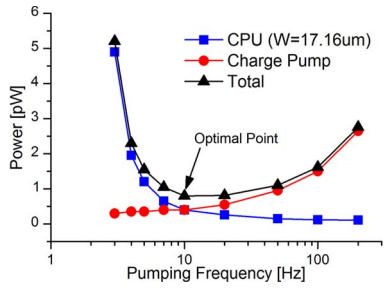

Fig. 9. Measured CPU and charge pump power in standby mode.

A triple stacked inverter is used for connecting $V_{\rm out}$ to the footer. The pMOS stack minimizes subthreshold leakage during standby mode, thereby lessening the pumping overhead and the required pumping frequency, whereas the nMOS stack plays a critical role when switching from standby mode to active mode. The long nMOS stack cuts the connection between $V_{\rm out}$ and the gate of the footer to eliminate contention between the pMOS stack and the charge pump. It is crucial to bias the bodies of the entire nMOS stack with $V_{\rm out}$ to ensure that the nMOS stack is not forward biased during active mode. The negative voltage developed at $V_{\rm out}$ is preserved during active mode, which is typically very short (on the order of milliseconds) [4], thus minimizing the time and power overhead for switching back to standby mode.

A CPU block with 23 472 transistors was fabricated and tested at room temperature (25 °C). Fig. 8 shows the generated super cutoff voltage and charge pump power consumption as charge pump clock frequency is swept. The charge pump clock was supplied externally in this specific experiment to give maximum tunability. However, for on-chip clock generation in this frequency regime, the sub-pW clock generator in [9] can be used to balance overhead of the charge pump power consumption. Strong super cutoff voltage of -230 mV is generated with pumping frequency of 10 Hz and sub-pW power. The required frequency and power of the charge pump were higher than simulation due to parasitic load and leakage but were still in pW order. The leakage reduction achieved using SCCMOS is shown in Fig. 9. With a footer width of 17.16 μm, the CPU block consumes 15.4 pW in standby mode without SCCMOS. At lower pumping frequencies (< 10 Hz), increasing the pumping frequency reduces total standby power since it improves the super cutoff voltage. However, as frequency exceeds 10 Hz, the charge pump overhead becomes dominant and increases total power consumption. Total standby power reaches a minimum of 0.8 pW at 10 Hz, i.e., a $19.3\times$ reduction over normal operation.

The size of the power-gating transistor is constrained by the standby mode leakage budget and active mode current demand.

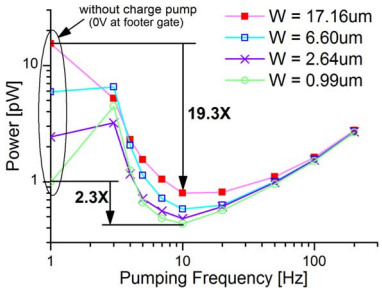

Fig. 10. Measured total standby power of CPU and charge pump with footer size.

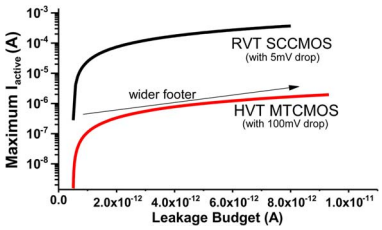

Fig. 11. Simulated maximum active current with RVT/HVT footer.

A wide power-gating transistor is preferred to minimize the voltage drop in active mode, whereas narrow width is preferred for minimum leakage in standby mode. The voltage drop across power-gating transistor effectively reduces V_{DD} for the logic, making the circuit slower, less robust, and less energy efficient.

Fig. 10 shows the standby power reduction for various footer sizes. Note that for all footer sizes, the standby power converges to $\sim$1 pW at an optimal pumping frequency of 10 Hz. Therefore, the power gain is largest $(19.3\times)$ with the widest footer and smallest $(2.3\times)$ with the narrowest.

Fig. 11 shows simulated maximum active current that RVT and HVT footers can support. Assuming a 250-mV underdrive voltage, RVT footers can support more than two orders of magnitude higher active current with 5-mV drop compared with HVT footers with 100-mV drop for a given leakage current budget. This implies that, with an SCCMOS RVT footer, V_{DD} can be reduced by more than 95 mV, resulting in 34% active mode energy reduction.

IV. Standby Power Reduction in Memory

In this section, we propose how the SCCMOS structure can be applied to SRAM leakage reduction. Various SRAM structures, such as the modified 6T [10], 8T [11], and 10T [12] topologies, have been explored for low-voltage applications. To demonstrate the standby power reduction for memory with the proposed SCCMOS technique, we selected the low-leakage memory cell proposed in [4]. However, many of the conclusions in this brief may be extended to other cells as well. As depicted in Fig. 12, the memory cell uses cross-coupled inverters with stacked HVT transistors to minimize the subthreshold leakage. A separate read buffer with RVT transistors is used to boost the read performance and improve cell stability.

A. Leakage Reduction for Read Circuits and Peripherals

Fig. 12 shows the most important leakage paths in the memory. Path 1 is the leakage path for the read circuit that

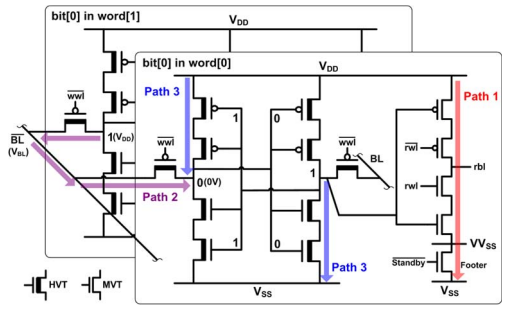

Fig. 12. Leakage paths in low leakage memory cell.

is power gated during standby mode. Only the read buffer is shown in Fig. 12, but this circuit block also includes memory control logic, such as row/column decoders, bit-line drivers. Since these circuits are all power gated, our analysis shows that Path 1 contributes only $\sim$2% of the total standby power.

B. Bit-Line Leakage Reduction

Path 2 in Fig. 12 shows the bit-line leakage path in the array structure of the memory. During standby mode, the bit lines (see BL and $\overline{BL}$ in Fig. 12) float to some intermediate voltage $V_{\rm BL}$ between 0 and V_{DD}. As a result, the access transistors will have a drain–source voltage of $V_{\rm BL}$ or $V_{DD} - V_{\rm BL}$. This drain–source voltage induces subthreshold leakage on the bit line, which contributes $\sim$50% of total standby leakage.

In order to reduce the bit-line leakage, a super cutoff voltage $(> V_{DD})$ can be applied to the gate of the pass transistors during standby mode. This is achieved by using a charge pump to boost the power supply for the word-line driver. The basic concept of this strategy is similar to the strategy used with power-gated logic blocks, but it raises the following new challenges: 1) A new power supply for the pass transistor control logic must be kept near V_{DD} or higher at all times since a low voltage at the gate of the pass transistors will result in catastrophic data loss; 2) the new power supply should be able to supply enough current to meet the demands of the pass transistor control logic during active mode; and 3) all these criteria should be met with a power budget on the order of pW.

The proposed circuit that meets these criteria is presented in Fig. 13. The charge pump presented in Section II with pMOS is used for boosting the power supply. The output of this charge pump is tied to the power rail of the word-line drivers. Charge is continuously pumped into the output capacitor $C_{\rm out}$ to develop $V_{\rm out}$. The word-line drivers are structured to always provide full $V_{\rm out}$ in standby mode while also enabling word-line control during the active mode. However, there can be no direct connection between power supply and the output node because a direct connection to V_{DD} would prevent $V_{\rm out}$ from rising higher than V_{DD} in standby mode. As a result, write operations during active mode will consume the charge stored in $C_{\rm out}$, thereby lowering $V_{\rm out}$. Therefore, consecutive write operations that occur between pumping cycles (due to the low pumping frequency) may bring $V_{\rm out}$ below V_{DD}, which can result in data loss.

To prevent this, a "holder" transistor is introduced, which indirectly connects V_{DD} with the output of the charge pump. When $V_{\rm OUT}$ drops below V_{DD}, the holder transistor is forward biased and effectively "hold" $V_{\rm OUT}$ near V_{DD}. With a HVT

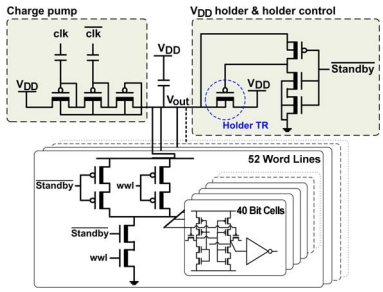

Fig. 13. Proposed circuit for bit-line leakage reduction.

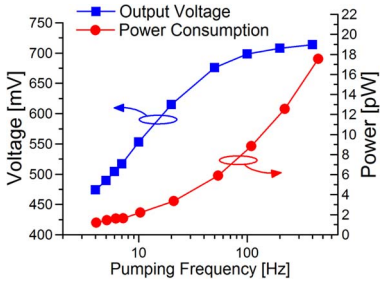

Fig. 14. Measured generated super cutoff voltage and charge pump power for memory leakage reduction.

holder transistor, worst case simulations show that this configuration maintains $V_{\rm out} > 489$ mV at $V_{DD} = 0.5$ V.

C. Intracell Leakage

Finally, Path 3 in Fig. 12 shows the intracell subthreshold leakage path. In each cell, the primary leakage paths include a single nMOS stack and a single pMOS stack. Our analysis shows that this leakage amounts to ~48% of total standby power.

In order to suppress intracell subthreshold leakage, an RBB can be applied to all transistors or HVT transistors can be used. However, according to our analysis, the overhead of generating enough well bias current to compensate for junction leakage was greater than the projected leakage improvement. Therefore, our memory structure uses HVT transistors as in [4].

To demonstrate the leakage reduction of the proposed techniques, a memory device with 2720 bit cells was fabricated and tested at room temperature (25 °C). Fig. 14 shows the generated super cutoff voltage and charge pump power consumption as functions of charge pump clock frequency. The power overhead for the charge pump is significantly higher than the previous section due to the larger number of leakage paths. This also results in higher optimal pumping frequency of 20 Hz, but the charge pump overhead is still below 5% of the original memory standby power. Total standby power is shown in Fig. 15. At a pumping frequency of 20 Hz, standby power is reduced by 29.1% compared with normal operation. Note that power actually increases at low frequencies since the output of the charge pump can fall below V_{DD} (0.5 V) in this region and cause increased leakage across pass transistors. Die micrograph of the test chip fabricated in 180-nm technology is shown in Fig. 16.

V. Conclusion

Super cutoff circuit techniques for reducing the standby power of ultralow power processors have been presented. A standby power reduction of 2.3–19.3× is achieved for power-gated logic blocks, whereas standby power is reduced by 29.1% for memory using the proposed techniques.

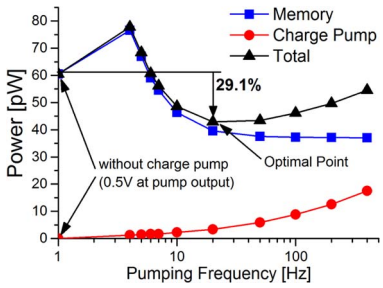

Fig. 15. Measured memory and charge pump power in standby mode.

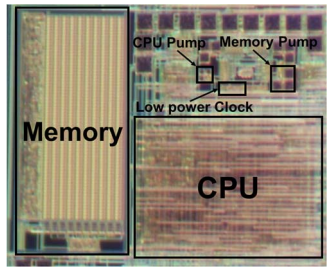

Fig. 16. Die micrograph of test chip.

References

[1] Y.-T. Liao, Y. Huanfen, B. Parviz, and B. Otis, "A 3 μW wirelessly powered CMOS glucose sensor for an active contact lens," in *Proc. IEEE ISSCC Dig. Tech. Papers*, Feb. 2011, pp. 38–40.

[2] Y. Lee, S. Bang, I. Lee, Y. Kim, G. Kim, M. H. Ghaed, P. Pannuto, P. Dutta, D. Sylvester, and D. Blaauw, "A modular 1 mm^3 die-stacked sensing platform with low power I^{2}C inter-die communication and multi-modal energy harvesting," *IEEE J. Solid-State Circuits*, vol. 48, no. 1, pp. 229–243, Jan. 2013.

[3] A. Wang and A. Chandrakasan, "A 180-mV subthreshold FFT processor using a minimum energy design methodology," *IEEE J. Solid-State Circuits*, vol. 40, no. 1, pp. 310–319, Jan. 2005.

[4] S. Hanson, M. Seok, Y.-S. Lin, Z. Foo, D. Kim, Y. Lee, N. Liu, D. Sylvester, and D. Blaauw, "A low-voltage processor for sensing applications with picowatt standby mode," *IEEE J. Solid-State Circuits*, vol. 44, no. 4, pp. 1145–1155, Apr. 2009.

[5] S. Mutoh, T. Douseki, Y. Matsuya, T. Aoki, S. Shigematsu, and J. Yamada, "1-V power supply high-speed digital circuit technology with multithreshold-voltage CMOS," *IEEE J. Solid-State Circuits*, vol. 30, no. 8, pp. 847–854, Aug 1995.

[6] H. Kawaguchi, K. Nose, and T. Sakurai, "A super cut-off CMOS (SCCMOS) scheme for 0.5-V supply voltage with picoampere stand-by current," *IEEE J. Solid-State Circuits*, vol. 35, no. 10, pp. 1498–1501, Oct. 2000.

[7] J.-T. Wu and K.-L. Chagn, "MOS charge pumps for low-voltage operation," *IEEE J. Solid-State Circuits*, vol. 33, no. 4, pp. 592–597, Apr. 1998.

[8] J. Shin, I.-Y. Chung, Y. J. Park, and H. S. Min, "A new charge pump without degradation in threshold voltage due to body effect," *IEEE J. Solid-State Circuits*, vol. 35, no. 8, pp. 1227–1230, Aug. 2000.

[9] Y. Lee, M. Seok, S. Hanson, D. Blaauw, and D. Sylvester, "Standby power reduction techniques for ultra-low power processors," in *Proc. IEEE ESSCIRC Dig. Tech. Papers*, Sep. 2008, pp. 186–198.

[10] B. Zhai, D. Blaauw, D. Sylvester, and S. Hanson, "A sub-200 mV 6T SRAM in 130 nm CMOS," in *Proc. IEEE ISSCC Dig. Tech. Papers*, Feb. 2007, pp. 332–606.

[11] L. Chang, D. M. Fried, J. Hergenrother, J. W. Sleight, R. H. Dennard, R. K. Montoye, L. Sekaric, S. J. McNab, A. W. Topol, C. D. Adams, K. W. Guarini, and W. Haensch, "Stable SRAM cell design for 32 nm node and beyond," in *Proc. Symp. VLSI Technol.*, Jun. 2005, pp. 128–129.

[12] B. Calhoun and A. Chandrakasan, "A 256kb sub-threshold SRAM in 65nm CMOS," in *Proc. IEEE ISSCC Dig. Tech. Papers*, Feb. 2006, pp. 2592–2601.