Energy Optimized Subthreshold VLSI Logic Family With Unbalanced Pull-Up/Down Network and Inverse Narrow-Width Techniques

Ming-Zhong Li, Chio-In Ieong, Man-Kay Law, Pui-In Mak, Mang-I Vai, Sio-Hang Pun, and Rui P. Martins

***Abstract*—Ultralow-energy biomedical applications have urged the development of a subthreshold VLSI logic family in standard CMOS. This brief proposes an unbalanced pull-up/down network, together with an inverse narrow-width technique, to improve the operating speed of the individual logic cell. Effective logical efforts save both power and die area in the process of device sizing and topology optimization. Three experimental 14-tap 8-bit finite impulse response filters optimized for ultralow-voltage operation were fabricated in 0.18-μm CMOS. Measurements show that the optimized 0.45 and 0.6 V libraries achieve minimum energy operations at 100 kHz, with a figure-of-merit of 0.365 (at 0.31 V) and 0.4632 (at 0.39 V), respectively. They correspond to 35.96% and 18.74% improvements, and the overall performances are well comparable with the state of the art.**

***Index Terms*—CMOS, device sizing, electrocardiography (ECG), finite impulse response (FIR) filter, inverse narrow width (INW), logical effort, process–voltage–temperature (PVT) variations, subthreshold standard logic library, ultralow energy, ultralow voltage.**

I. Introduction

With substantial energy reduction achieved in subthreshold operation as evidenced by the minimum energy point theory [1], VLSI logic family operating beneath the threshold voltage (V_T) is favored for wearable/implantable biomedical systems that require low-to-moderate computation speed with stringent power budget. However, the reduced overdrive voltage can dramatically worsen the device susceptibility in delay and noise margin due to process, voltage, and temperature (PVT) variations [2]. This inevitably leads to suboptimal performance in terms of power, delay, and area, and even logic failure in the worst case.

Traditionally, a balanced pull-up (PU) and pull-down (PD) network approach is preferred in logic cell design, which is important for subthreshold cell design to have a comparable PU/PD driving capability [3], [4]. Even though this can be readily achieved by either upsizing the pMOS in the PU network or stacking the nMOS in the PD network, the area overhead can lead to extra loading excessive leakage power, and also a suboptimal energy efficiency even with an identical total width as the unbalanced one [to be detailed in Fig. 1(a)]. In [5], a balanced PU/PD network is achieved using a body biasing scheme. This, however, requires extra monitoring blocks and can incur considerable power and area penalties. In [6], the statistical distribution of the drain–source current, rather than the current itself, is investigated to achieve the balanced networks. In [7], the reverse channel (RSC) effect is used for device optimization by increasing the channel length to have an optimal V_T and higher driving capability. Yet, the RSC effect may not be readily applicable to all the technology nodes.

This brief describes a subthreshold standard cell library targeting ultralow-energy biomedical applications. To improve the energy efficiency, the unbalanced PU/PD network, logical effort, and inverse-narrow-width (INW) techniques are exploited. The unbalanced PU/PD network achieves a better energy efficiency, and the enhanced influence from PVT variations is carefully verified with butterfly Monte Carlo simulations. Analytical delay models [8] for subthreshold circuits are utilized to predict the propagation delay with good accuracy. The logical effort [9] is utilized to qualitatively provide the delay spread estimate, as well as to distinguish dissimilar topologies of a particular logic and help determine an optimal architecture. The INW effect [10] is exploited for circuit level optimization. Instead of using the smallest width per finger for both pMOS and nMOS transistors as in [11], analysis and silicon implementation of the INW effect using the power-delay-product (PDP) metric is initiated for optimal gate performance. An entity of 56 power-optimized subthreshold logic cells is implemented in standard 0.18-μm CMOS. With them, three 14-tap 8-bit finite impulse response (FIR) filters optimized at different supplies for electrocardiography (ECG) signal transformation are demonstrated and compared. The achieved figure-of-merit (FoM) compares favorably with the state of the art [12].

Section II presents the detailed implementation of the proposed library in standard 0.18-μm CMOS. Section III reports the measurement results of three FIR filters to validate the benefits of the proposed library. Section IV concludes this brief.

II. Standard Cell Sizing

A. Single-Stage Gates Design

Single-stage gates, such as INV, NOR, and NAND, are scaled with reference to the basic inverter [4]. As defined in [1], the total energy consumed by an arbitrary circuit is modeled as follows:

$$E_{\text{Total}} = C_{\text{eff}} V_{\text{DD}}^2 + W_{\text{eff}} I_{\text{leakage}} V_{\text{DD}} t_d L_{\text{DP}} \quad (1)$$

where E_{Total} is the total energy, C_{eff} and W_{eff} are the effective capacitance and width, I_{leakage} is the leakage current, t_d is the propagation delay, and L_{DP} is the logic depth. The smaller the device sizing, the smaller the C_{eff} and W_{eff}, and hence, the smaller the E_{Total}. Typically, a balanced PU and PD network approach is achieved by increasing the pMOS sizing for comparable driving capability. Nevertheless, an upsized nMOS increases the effective capacitance and also the leakage current, thus, will not

Manuscript received July 23, 2014; revised November 11, 2014; accepted December 31, 2014. Date of publication January 26, 2015; date of current version November 20, 2015. This work was supported in part by the Macao Science and Technology Development Fund under Grant 015/2012/A1, Grant 024/2009/A1, and Grant 047/2013/A2, and in part by the Research Committee through the University of Macau, Macau, China, under Grant MYRG079-FST12-VMI, Grant MYRG100-FST13-LMK, Grant MYRG103-FST13-VMI, and Grant MYRG115-FST12-LMK.

M. Z. Li, C. I. Ieong, M. I. Vai, and S. H. Pun are with the State Key Laboratory of Analog and Mixed-Signal VLSI and the Biomedical Engineering Laboratory, Department of Electrical and Computer Engineering, Faculty of Science and Technology, University of Macau, Macau, China.

M. K. Law and P. I. Mak are with the State Key Laboratory of Analog and Mixed-Signal VLSI, Department of Electrical and Computer Engineering, Faculty of Science and Technology, University of Macau, Macau, China (e-mail: mklaw@umac.mo).

R. P. Martins was with the Instituto Superior Técnico, Universidade de Lisboa, Lisbon 1649-004, Portugal. He is now with the State Key Laboratory of Analog and Mixed-Signal VLSI, Department of Electrical and Computer Engineering, Faculty of Science and Technology, University of Macau, Macau, China (e-mail: rmartins@umac.mo).

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

Digital Object Identifier 10.1109/TVLSI.2015.2388783

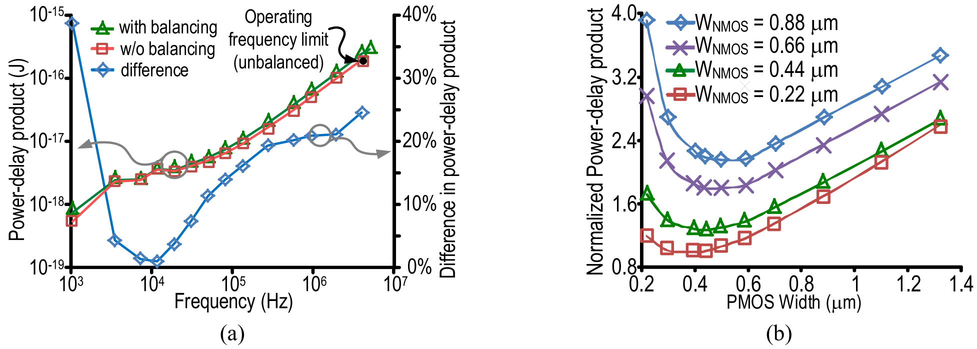

Fig. 1. (a) PDP of an inverter (FO4 loading) with balanced (P/N ratio = 5/1) and unbalanced (P/N ratio = 2/1) PU/PD network versus operating frequency at 0.3 V. (b) Normalized PDP of an FO4 inverter at various nMOS/pMOS widths at 0.3 V.

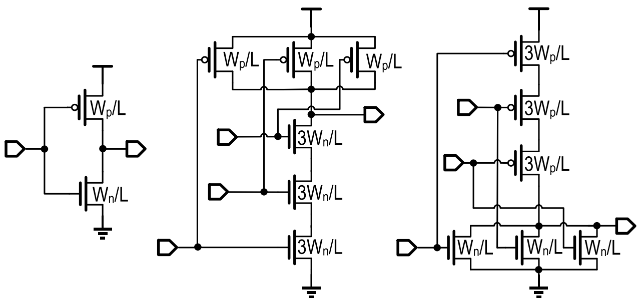

Fig. 2. Basic INVX1, NAND3, and NOR3 logic building blocks.

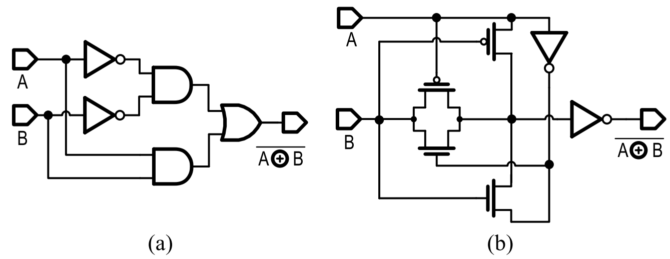

Fig. 3. XNOR gate. (a) Conventional. (b) Pass-gate based [14].

be adequate for biomedical applications with relaxed speed requirements. Fig. 1(a) shows the simulation results of a fan-out of 4 (FO4) inverter with and without balanced PU and PD networks, and the total width of the devices is kept identical for a fair comparison. The unbalanced approach can only operate up to ~4 MHz due to the reduced noise margin as expected. However, a better PDP merit can be achieved when compared with the balanced approach, demonstrating the benefits of using unbalanced implementation in low-to-moderate speed applications. One the other hand, simulation results show that a FO4 inverter with unbalanced network can operate up to 450 kHz in the slow-pMOS, slow-nMOS corner, which is sufficient for the designs with low-to-moderate speed requirement. Likewise, the PDP of the unbalanced NAND3, NOR3, and XNOR consistently shows improvement as in the case of the unbalanced inverter. Fig. 1(b) plots the PDP of an FO4 inverter at various nMOS/pMOS widths. The lowest PDP varies according to different nMOS/pMOS widths due to the INW effect (to be detailed in Section II-C), and should be selected with different PDP and driving strength considerations.

Fig. 2 shows the reference inverter INVX1 as well as NAND3 and NOR3 employing the resistor model to match the equivalent resistance, with the three-stack pMOS/nMOS device tripling the transistor width. As the three-input NAND (NAND3) and NOR (NOR3) gates provide the worst case PD and PU networks in the library [13] that have the same voting weight as INVX1, this worst case PD and PU networks are further adopted to provide stringent conditions to verify the metastability of the remaining gates. Though the stacked devices are prone to process variation in subthreshold operations, Monte Carlo simulation shows that they still meet the speed requirement for biomedical applications. Moreover, INVX1 is employed as a standard circuit to characterize the propagation delay and logical effort of each logic cell.

B. Multistage Gates Design

For multistage gates, such as XOR and XNOR, a logical effort similar to single-stage gates is utilized to capture the signal propagation delay as a result of different logic topologies. The logical effort g for a specific logic gate is defined as

$$g = \frac{\sum C_i}{C_{\text{inv}}} = \frac{C_b}{C_{\text{inv}}} \tag{2}$$

where C_b is the combined input capacitance C_i of every signal path i, and C_{inv} is the input capacitance of the reference inverter to have the same driving capability as the logic gate being characterized. By fixing the gate length L to be minimum as defined by the process, the input capacitance C_i is only a function of the transistor width W_i, and (2) becomes

$$g = \frac{\sum W_i}{W_{\text{inv}}} = \frac{W_b}{W_{\text{inv}}}. \tag{3}$$

Without loss of generality, Fig. 3 shows the XNOR gates implemented using the conventional and the pass-gate-based topologies [14], and the corresponding total logical efforts are 11 and 4, respectively. It is noteworthy that the AND/OR logic in Fig. 3(a) can be reduced to NAND/NAND logic. The pass-gate-based XNOR gate achieves less logical effort, less energy consumption a more compact size and also faster operation speed. In addition, the implemented XOR (XNOR) gate leads to a 9.25% (9.25%) and 58.5% (37.9%) reduction in area and power, respectively, when compared with the conventional NAND2/NOR2 followed by AOI22/OAI22 approach. The pass gate is applied to overcome the V_T loss. And it is also beneficial to low-voltage operation with the increase in gate delay readily tolerable for low-to-moderate speed applications, and is thus selected as a preferred topology for the subthreshold standard library. It is observed in Monte Carlo simulations that the induced timing variation is comparable with conventional implementations.

C. Multifinger Transistor Dimensioning

Typically, conventional transistor dimension of a specific logic gate is mainly focused on balancing the driving capability to improve

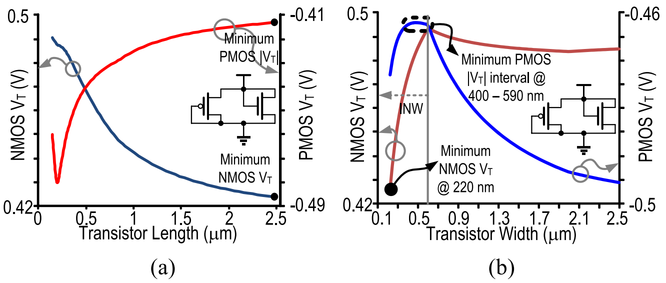

Fig. 4. nMOS/pMOS V_T versus (a) transistor length (RSC) and (b) transistor width (INW), at $V_{DD} = 0.3$ V.

circuit operating speed without much emphasis on the impact of the INW effect, and which is especially important for multifinger implementation. This section analyzes the INW effect and its influence on the circuit optimization. As described in [11], with a fixed transistor length L, the threshold voltage V_T is highly dependent on the transistor width W.

Fig. 4(a) shows the dc simulation result of the extracted threshold voltage versus device length using the selected 0.18-μm CMOS process, indicating that the minimum pMOS $|V_T|$ and nMOS V_T are located with the largest channel length and that the RSC effect is not beneficial for the ultralow-energy designs from the adopted technology process.

The small-geometry threshold voltage expression for an nMOS is expressed as follows:

$$V_T = V_{\text{fb}} + \psi_s + \left(\frac{Q_b}{\text{WLC}_{\text{ox}}}\right) \times \left[1 - \left(\sqrt{1 + \frac{2W_d}{r_j}} - 1\right)\left(\frac{r_j}{L}\right)\right]\left[1 - \frac{F}{W + F}\right] \quad (4)$$

where V_{fb} is the flat-band voltage, ψ_s is the surface potential, Q_b is the ionized impurity concentration, C_{ox} is the thin gate oxide capacitance, W_d is the depth of the gate-induced depletion region, r_j is junction depth, and F is the fringing factor. It can be observed that W can be optimized to achieve the smallest threshold voltage with a fixed length L.

Fig. 4(b) shows the simulation result of the extracted threshold voltage versus device width. Instead of keeping the smallest transistor width for both nMOS and pMOS, as suggested in [11], minimum V_T imposes a pMOS width in the range of 400–590 nm and an nMOS width of 220 nm. Moreover, although an upscaled pMOS width yields a reduced threshold voltage for enhanced driving capability, it also builds up the intrinsic and extrinsic loading that can affect the overall performance. Still, with reference to Fig. 1(b), the minimum PDP region is achieved at an pMOS width of ∼450 nm, and it is also used as a reference for pMOS width sizing for multifinger transistor design. The nMOS width should still be 220 nm to achieve the minimum PDP. On the other hand, the INW effect is highly process dependent and the optimum transistor width should be selected on a case-by-case basis, and five different technology nodes have been observed to exhibit PDP improvements after applying the INW and the unbalanced techniques in the kilohertz operating region required in biomedical applications.

D. Driving Strength Design and Metastability Validation

Multiple driving capabilities for different cells having the same logic function are necessary to drive different output loading. Conventionally, logic gates with higher driving capabilities are upsized several integer times that of the weaker gates [11]. As a consequence, the design margin of the logic cell with higher driving capability is difficult to be clarified and characterized. Here, the FO4 structure is adopted as a standard setup for transistor scaling with increased driving capability to obtain a comparable PU and PD propagation delay with respect to the basic logic cell.

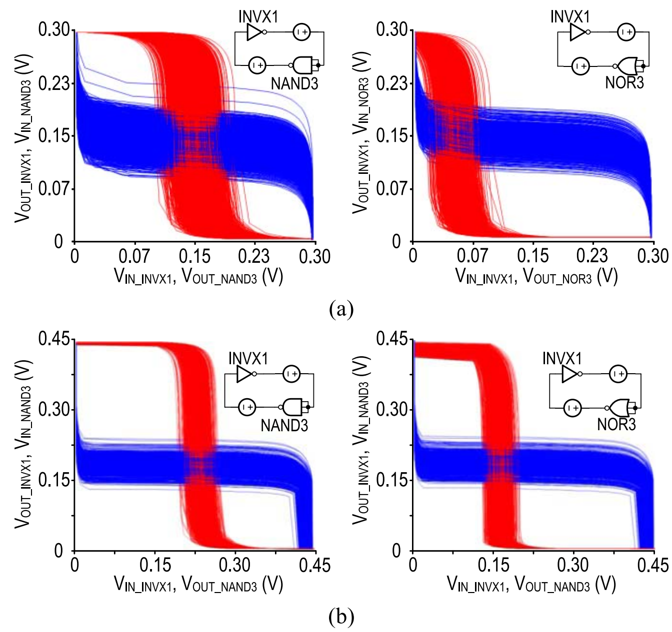

Fig. 5. Noise margin verification versus NAND3 and NOR3 from layout extractions with (a) 0.30 V and (b) 0.45 V operations.

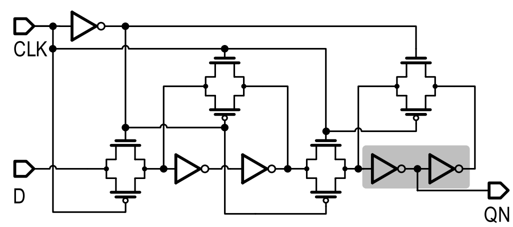

Fig. 6. Derived and optimized static flip-flop from [13].

As mentioned, NAND3 and NOR3 provide the most stringent PD and PU propagation delays, respectively, and their butterfly plots [13] are used for characterizing their noise margins, and serve as a reference for validating the metastability of the remaining gates. The static noise margin is defined by the largest inscribed square, which one is able to draw in the butterfly plots, and then the noise margin is the diagonal of those squares. Fig. 5 shows the corresponding 5-k Monte Carlo simulation results indicating that the gate under test has sufficient noise margin to tolerate the worst case transition slope. Fig. 5(a) shows that the worst case noise margin is stringent with NOR3 at 0.3 V, and the logic function is approaching the point of breakdown. In case of the occurrence of negative noise margin, the power supply voltage can be increased to guarantee an enough noise margin, as shown in Fig. 5(b).

E. Sequential Cells Design

The sequential logic elements, such as latch and flip-flop, are indispensable to provide storage logic function. To ensure subthreshold operation with reduced power consumption, an 18-transistor flip-flop based on logical effort using (2) and (3) is derived and shown in Fig. 6. The butterfly plot is also used to verify the metastability and data retention capability. The proposed D flip-flop (DFF) is implemented with 18 transistors using the unbalanced technique. When compared with the conventional balanced 22-transistor imple-

TABLE I
COMPARISON OF FIR FILTERS DESIGNED USING SUBTHRESHOLD TECHNIQUE

	This Work				[12] TCAS-II'12	[5] VLSI'07	[16] JSSC'10	[17] JSSC'10	[18] CICC'10
	with 0.45-V .lib		with 0.6-V .lib						
FIR Type	14-tap, 8-bit				30-tap, 8-bit	8-tap, 8-bit	14-tap, 8-bit	8-tap, 8-bit	4th order
Technology	0.18-µm				0.13-µm	0.13-µm	0.13-µm	90-nm	0.13-µm
Optimum V_{DD} (V)	0.31		0.39		0.35	0.2	0.27	0.29	1.2
Frequency (Hz)	100k		100k		29k	12k	20M	148k	20k
Energy/Tap (pJ)	**0.02735**	**0.0234**	**0.03568**	**0.02964**	1.1	1.19	1.11	0.6275	39
Power (nW)	**38.29#**	**32.7**	**49.95#**	**41.5**	32	114	310,000	742.96	780
FoM*	**0.4273**	**0.3650**	**0.5575**	**0.4632**	0.57	18.55	17.37	9.80	N/A
Area/Channel (mm^2)	0.053		0.049		0.058	1.54	0.38	N/A	0.7

*FIR FoM = power(nW)/freq.(MHz)/# of taps/input bit length/coefficient bit length. [5]
#Multi-chip measurement results (mean value) from 15 chips.

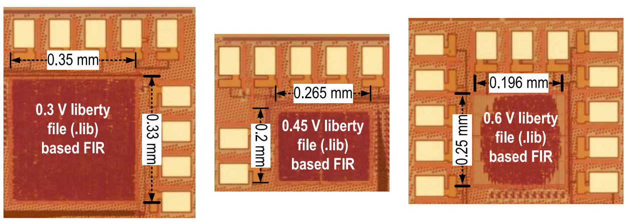

Fig. 7. Die micrographs of 0.18-μm subthreshold FIR test chips.

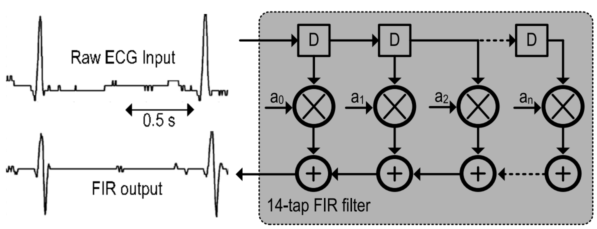

Fig. 8. Measured functional result of the FIR filters.

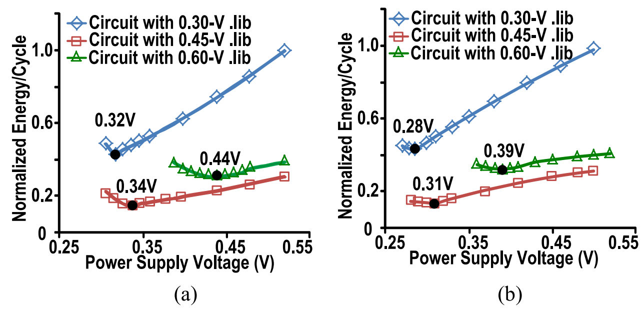

Fig. 9. Normalized energy/cycle with (a) random input signal and (b) ECG input signal (black dots indicate the optimum points).

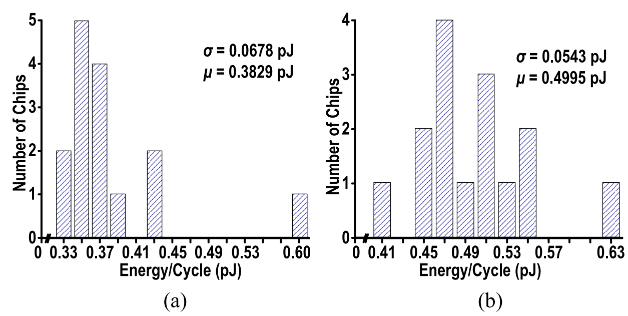

Fig. 10. Statistical data of the minimum energies measured from 15 chips ($T = 25$ °C). (a) Design with 0.45 V .lib. (b) Design with 0.60 V .lib.

mentation, the proposed DFF achieves 7%, 125%, and 48.4% in area, power (including clock power), and propagation delay reductions, respectively. For the flip-flops with higher driving capability, the output stage highlighted in Fig. 6 should be upsized accordingly.

III. MEASUREMENT RESULTS

To verify the proposed subthreshold standard cell library, a complete set of 56 subcells is implemented in 0.18-μm CMOS with a threshold voltage of approximately 0.42 V. The 56 proposed cells exhibit an averaged area saving of 7.13% when compared with the commercial standard cell library. By keeping identical design constraints for designing 12-bit FIR filters, the report from RTL Compiler shows that the power consumption of the FIR filter with the proposed library achieves approximately 30% improvement when compared with the one implemented using a commercial library. On the other hand, three 14-tap 8-bit FIR filters with 8-bit coefficients targeting on ECG signal transformation are designed. The FIR filters are synthesized using the previously characterized liberty (.lib) files (0.3, 0.45, and 0.6 V), while preserving identical design constraints. Furthermore, the clock gating technique is adopted to reduce dynamic power consumption. Fig. 7 shows the chip micrographs of the fabricated FIR filters, with active areas of 0.1155, 0.053, and 0.049 mm^2, respectively. Note that the 0.3 V design is suboptimal and is included only for comparison purposes. Its large area is due to the additional buffers to fulfill the timing specifications during the synthesis and layout stages. The combined operating supply voltage and frequency range coverage are 0.26–0.8 V and 500 Hz–1 MHz, respectively. Fig. 8 shows the FIR filter structure and the measured functional output. The normalized energy performance against different supplies with both random and ECG signals are shown in Fig. 9. It can be observed that the 0.45 V .lib achieves the lowest minimum energy point. As expected, due to the increased area, and hence parasitic capacitance, the circuit designed with the 0.3 V liberty file consumes the largest normalized energy per cycle. Furthermore, even though the area consumed by the circuit designed with 0.6 V liberty file is smaller than the one with 0.45 V liberty file, the normalized energy consumption is still larger owing to increased dynamic power consumption. Fig. 10 shows the statistical results of the designs at room temperature (25 °C). Table I benchmarks this brief with the state-of-the-art FIR filters that with advanced technology nodes. Using the FIR FoM described in [5], it can be observed that the proposed work achieves comparable performance to state-of-the-art designs.

IV. Conclusion

Effective circuit techniques and design methodology are proposed to realize a subthreshold VLSI logic family for biomedical applications. An entity of 56 standard cells was demonstrated in standard 0.18-μm CMOS with $V_T \approx 0.42$ V. The design framework involves unbalanced PU/DN network, logical effort, and INW-based multifinger topology for individual single-stage/multistage gate optimization. With them, three 14-tap 8-bit FIR filters are designed and measured according to different liberty timing files. The achieved FoMs at the minimum energy operating points for the 0.45 and 0.6 V library designs are 0.365 (at 0.31 V) and 0.4632 (at 0.39 V), respectively; both compare favorably with the state-of-the-art FIR filter designs.

Acknowledgment

The authors would like to thank T.-T. Zhang, T. Wu, Z.-Y. Chen, and C. Dong for the valuable discussions.

References

[1] B. H. Calhoun and A. Chandrakasan, "Characterizing and modeling minimum energy operation for subthreshold circuits," in *Proc. Int. Symp. Low Power Electron. Design (ISLPED)*, Aug. 2004, pp. 90–95.

[2] A. Tajalli and Y. Leblebici, "Design trade-offs in ultra-low-power digital nanoscale CMOS," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 58, no. 9, pp. 2189–2200, Sep. 2011.

[3] N. Reynders and W. Dehaene, "Variation-resilient building blocks for ultra-low-energy sub-threshold design," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 59, no. 12, pp. 898–902, Dec. 2012.

[4] M. Alioto, "Understanding DC behavior of subthreshold CMOS logic through closed-form analysis," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 57, no. 7, pp. 1597–1607, Jul. 2010.

[5] M.-E. Hwang, A. Raychowdhury, K. Kim, and K. Roy, "A 85 mV 40 nW process-tolerant subthreshold 8 × 8 FIR filter in 130 nm technology," in *Proc. IEEE Symp. VLSI Circuits (VLSI)*, Jun. 2007, pp. 154–155.

[6] B. Liu, J. P. de Gyvez, and M. Ashouei, "Sub-threshold standard cell sizing methodology and library comparison," *J. Low Power Electron. Appl.*, vol. 3, no. 3, pp. 233–249, Jul. 2013.

[7] T.-H. Kim, H. Eom, J. Keane, and C. Kim, "Utilizing reverse short channel effect for optimal subthreshold circuit design," in *Proc. Int. Symp. Low Power Electron. (ISLPED)*, Oct. 2006, pp. 127–130.

[8] F. Frustaci, P. Corsonello, and S. Perri, "Analytical delay model considering variability effects in subthreshold domain," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 59, no. 3, pp. 168–172, Mar. 2012.

[9] I. S. Sutherland, R. Sproull, and D. Harris, *Logical Effort: Designing Fast CMOS Circuits*. San Francisco, CA, USA: Morgan Kaufmann, 1999.

[10] L. A. Akers, "The inverse-narrow-width effect," *IEEE Electron Device Lett.*, vol. EDL-7, no. 7, pp. 419–421, Jul. 1986.

[11] J. Zhou, S. Jayapal, B. Busze, L. Huang, and J. Stuyt, "A 40 nm dual-width standard cell library for near/sub-threshold operation," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 59, no. 11, pp. 2569–2577, Nov. 2012.

[12] A. Klinefelter, Y. Zhang, B. Otis, and B. H. Calhoun, "A programmable 34 nW/channel sub-threshold signal band power extractor on a body sensor node SoC," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 59, no. 12, pp. 937–941, Dec. 2012.

[13] J. Kwong and A. P. Chandrakasan, "Variation-driven device sizing for minimum energy sub-threshold circuits," in *Proc. Int. Symp. Low Power Electron. Design (ISLPED)*, Oct. 2006, pp. 8–13.

[14] J.-M. Wang, S.-C. Fang, and W.-S. Feng, "New efficient designs for XOR and XNOR functions on the transistor level," *IEEE J. Solid-State Circuits*, vol. 29, no. 7, pp. 780–786, Jul. 1994.

[15] M.-Z. Li, C.-I. Ieong, M.-K. Law, P.-I. Mak, M.-I. Vai, and R. P. Martins, "Sub-threshold standard cell library design for ultra-low power biomedical applications," in *Proc. 35th Annu. Int. Conf. Eng. Med. Biol. Soc. (EMBC)*, Jul. 2013, pp. 1454–1457.

[16] W.-H. Ma, J. C. Kao, V. S. Sathe, and M. C. Papaefthymiou, "187 MHz subthreshold-supply charge-recovery FIR," *IEEE J. Solid-State Circuits*, vol. 45, no. 4, pp. 793–803, Apr. 2010.

[17] I. J. Chang, S. P. Park, and K. Roy, "Exploring asynchronous design techniques for process-tolerant and energy-efficient subthreshold operation," *IEEE J. Solid-State Circuits*, vol. 45, no. 2, pp. 401–410, Feb. 2010.

[18] F. Zhang, A. Mishra, A. G. Richardson, S. Zanos, and B. P. Otis, "A low-power multi-band ECoG/EEG interface IC," in *Proc. IEEE CICC*, Sep. 2010, pp. 1–4.