

Adaptive Resolution ADC Array for an Implantable Neural Sensor

Stephen O'Driscoll, *Member, IEEE*, Krishna V. Shenoy, *Senior Member, IEEE*, and Teresa H. Meng, *Fellow, IEEE*

Abstract—This paper describes an analog-to-digital converter (ADC) array for an implantable neural sensor which digitizes neural signals sensed by a microelectrode array. The ADC array consists of 96 variable resolution ADC base cells. The resolution of each ADC cell in the array is varied according to neural data content of the signal from the corresponding electrode. The resolution adaptation algorithm is essentially to periodically recalibrate the required resolution and this is done without requiring any additional ADC cells. The adaptation implementation and results are described. The base ADC cell is implemented using a successive approximation charge redistribution architecture. The choice of architecture and circuit design are presented. The base ADC has been implemented in $0.13\ \mu\text{m}$ CMOS as a 100 kS/s SAR ADC whose resolution can be varied from 3 to 8 bits with corresponding power consumption of $0.23\ \mu\text{W}$ to $0.90\ \mu\text{W}$ achieving an ENOB of 7.8 at the 8-bit setting. The energy per conversion step figure of merit is 48 fJ/step at the 8-bit setting. Resolution adaptation reduces power consumption by a factor of 2.3 for typical motor neuron signals while maintaining an effective 7.8-bit resolution across all channels.

Index Terms—Adaptive signal acquisition, analog–digital conversion, neural prosthesis, ultra low power.

I. INTRODUCTION

EACH YEAR, hundreds of thousands of people suffer from neurological injuries and disorders, resulting in the permanent loss of motor function. In a correctly functioning nervous system, signals are sent from the brain to muscles in order to control movement. In most cases of paralysis, both the driving neurons and the muscles, which control the limbs, are fully functional, but there is a disconnect in the nervous system's communication link between the two. Therefore, if the gap in the communication link is artificially bridged, the paralysis may be overcome. This concept is illustrated in Fig. 1 where a microelectrode array and integrated circuit (IC) are implanted above the motor cortex. Motor neuron signals are sensed by the electrodes and processed by an IC, which we call the implantable

Manuscript received October 15, 2010; revised February 20, 2011; accepted March 20, 2011. Date of current version May 18, 2011. This work was supported by the Focus Center for Circuit & System Solutions (C2S2), one of five research centers funded under the Focus Center Research Program, a Semiconductor Research Corporation Program. This paper was recommended by Associate Editor S. Chakrabarty.

S. O'Driscoll is with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95161 USA (e-mail: odriscoll@ucdavis.edu).

K. V. Shenoy is with the Departments of Electrical Engineering and Bioengineering and the Neurosciences Program, Stanford University, Stanford, CA 94305 USA.

T. H. Meng is with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA.

Digital Object Identifier 10.1109/TBCAS.2011.2145418

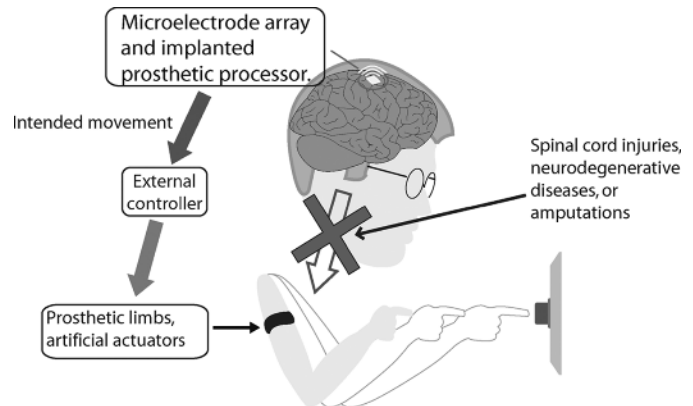


Fig. 1. Neuroprosthetic application.

prosthetic processor (IPP), to decode the intended movement. That intended movement is transmitted out of the body to an external controller, which may control prosthetic limbs or artificial actuators. A critical portion of this process is the brain interface (i.e., sensing and reading out the neural signals).

Neural signals are essentially voltage spike trains. Neural information is encoded in the variable timing between these spikes. Implanted electrodes are used to sense the neural signals. For a variety of fabrication and clinical reasons, electrodes cannot be oriented to sense individual neurons. Rather, each electrode typically senses signals from more than one neuron. In order to extract the underlying information, the signals sensed by each electrode must be decomposed into signals from individual neurons. To do this, the signals are first converted to the digital domain and then “spike sorted.” The sorted spikes are then decoded to identify intended movements. Our proposed implantable prosthetic processor (IPP) [1], therefore, comprises the following major building blocks: preamplification [2]; a variable-resolution analog-to-digital converter (ADC) array [3]; a digital spike sorter [4], [5]; a maximum-likelihood neural decoder [6]; a wireless data transceiver; and an adaptive millimeter-sized power receiver [7], [8]. A block diagram of the IPP, together with the electrode array and power receiving antenna, is illustrated in Fig. 2. The overall compression factor attained by employing the IPP is on the order of 10^6 , translating raw neural data at a rate of 80 Mb/s to less than 20 b/s, indicating the intended movement. Low power is essential for supply considerations and heat dissipation in the surrounding tissue, and it guides every aspect of the design. The total power budget of the IPP is limited to 1 mW, leaving a target of less than $1\ \mu\text{W}$ for each ADC cell.

This paper is organized as follows: Section II explains the motivation for, and algorithm to realize, ADC resolution adapta-

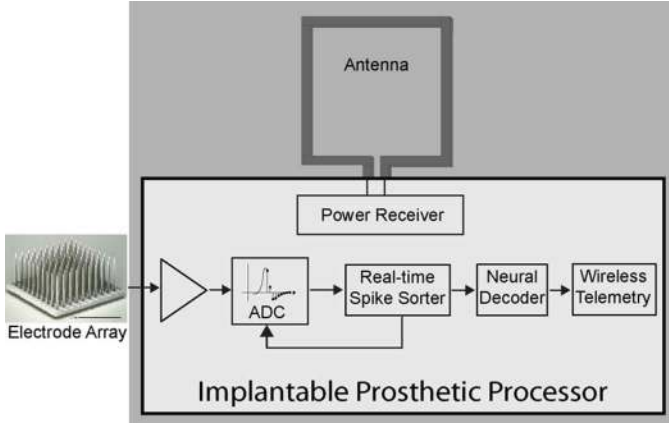


Fig. 2. Block diagram of implantable prosthetic processor.

tion. Next, Section III discusses lower bounds on how the power efficiency of various ADC architectures varies with resolution and uses this to determine the charge redistribution successive approximation (CR-SAR) ADC architecture as the most suitable for the implanted prosthetic processor. Section IV describes the circuit design of a variable resolution CR-SAR, and Section V presents the measured performance of the ADC implemented in $0.13\text{-}\mu\text{m}$ CMOS.

II. ADAPTIVE RESOLUTION ADC ARRAY

A. Variation of ADC Power With Resolution

Let us first consider how power consumption varies with resolution for some popular ADC architectures. Comparator power consumption dominates in flash ADC architectures and an n -bit flash ADC requires $2^n - 1$ comparators. Power dissipated in the digital-to-analog converter (DAC) dominates in successive approximation architectures and in an n -bit charge redistribution-based successive approximation ADC, the sub-DAC capacitance is proportional to 2^n . Similarly, for an n -bit sigma-delta ($\Sigma\Delta$) ADC, the required oversampling ratio is proportional to 2^{kn} where k varies based on the $\Sigma\Delta$ order and noise characteristics but is a constant for a given architecture. Therefore, in these cases, and in other ADC architectures, power consumption increases exponentially with increasing resolution, to the first order. There are usually necessary ancillary circuits, such as the control logic, whose power consumption is subexponential in resolution but nevertheless, the ADC power consumption is a very strong increasing function of resolution. This means that digitizing more bits than necessary wastes a lot of power.

B. Neural Signal-Processing Considerations

1) *Sensing Limitations of Electrode Array:* In the IPP application a 100-electrode array is implanted and 96 electrodes are used to sense voltage signals generated by neurons, specifically action potentials, while the four remaining electrodes serve to generate a reference potential. We use 96 ADCs, one per electrode, in order not to discard the spatial resolution offered by the electrode array. For a variety of fabrication and clinical reasons, electrodes cannot be oriented to sense individual neurons, rather, the electrodes are spaced $400\ \mu\text{m}$ apart in a square grid. Fig. 3 illustrates four neurons close to the tip of

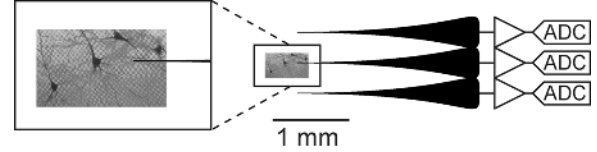


Fig. 3. Neurons and electrodes.

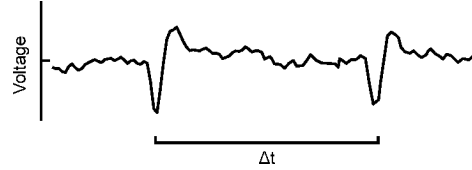


Fig. 4. Information is encoded in timing between spikes, not in spike shape.

an electrode, highlighting the fact that neurons do not occur in a square grid, and that the distance between them is typically considerably less than $400\ \mu\text{m}$. Correspondingly, the neurons whose potential an electrode senses can be at different distances, leading to different signal strengths and, thus, different signal-to-noise ratios (SNRs). Furthermore, each electrode typically senses signals from multiple neurons which need to be distinguished. Therefore, the resolution required in digitizing the signal from each electrode is not uniform across all electrodes. The standard analog approach would be to ask what the smallest voltage is that we must resolve, and to run each ADC at that resolution all of the time. This paper investigates whether ADC resolutions can be adapted to save power, without reducing information throughput. To answer this question, we need to look at the neural signals and the IPP system in greater detail and understand what dictates the resolution requirement.

2) *Characteristics of Sensed Neural Signals:* For the purposes of this paper, we are concerned with the neural signals known as action potentials, which are generated by individual neurons only, and not with signals which manifest at larger geometry scales, such as local field potential. Essentially, each neuron generates a train of voltage spikes. With the spikes from a particular neuron, the information is in the timing between those spikes, not in the shape or amplitude of the spike, as shown in Fig. 4.

Fig. 5 illustrates signal characteristics which are important to understand in order to develop methods which determine the required resolution. Fig. 5(a) shows a typical neural voltage spike. In applications where only one signal source contributes to the sensed signal, we can either use a 1-b ADC and fast automatic gain control in the preamplifier stage, or a fixed gain preamplifier and use SNR and dynamic range to determine the required ADC resolution. Frequently, as many as six neurons may contribute spikes to the signal sensed by a single electrode and the subsequent signal processing must be able to differentiate the spikes from each neuron, in these cases, a 1-b ADC would certainly discard information representing spikes from different neurons identically in the digital domain. Fig. 5(b) shows the overlay of many spikes from three different neurons sensed by a single electrode. This suggests that in order to allow differentiation of these spikes, ADC resolution should be determined by considering relative signal amplitudes or signal-to-signal ratios

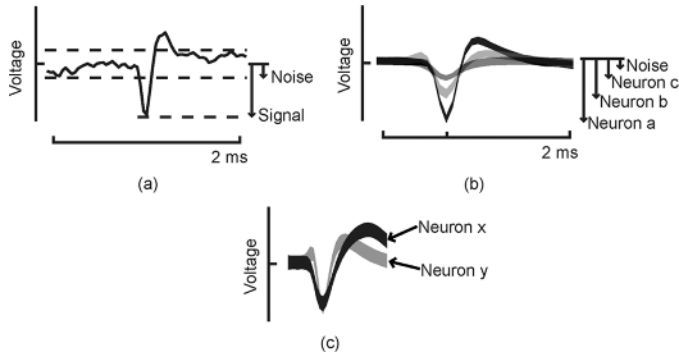


Fig. 5. (a) Signal and noise on a single spike signal (after HPF). (b) Multiple spikes with different magnitudes. (c) Spikes with a similar magnitude but different shape.

(SSRs) in addition to SNRs. If neural spikes were sorted solely according to signal amplitude, then SSRs and SNRs would be sufficient to determine the required ADC resolution.

However, Fig. 5(c) shows an overlay of spikes from two neurons which are very similar in amplitude, recorded by the same electrode. Required resolution estimation methods, based solely on SNR and/or SSR, would assume these spikes came from the same source, so that very low resolution would be deemed adequate and the subsequent spike sorting would be unable to distinguish these spikes. In order to allow differentiation of these spikes, the resolution assignment criteria must be equivalent to the sorting criteria in the downstream signal processing. This is most efficiently accomplished by using feedback from the real-time spike sorter to determine the resolution of each ADC cell.

3) *Spike Sorter Basics*: The spike sorter signal processing classifies each spike as originating from a specific neuron using principal component analysis (PCA). The spike sorter consists of a real-time spike sorter and a training block, as illustrated in Fig. 6. The parameters of the PCA sorter are retrained every 12 h to prevent errors due to drift of the electrodes, cell growth, etc. Since the subsequent signal processing assumes that the variation of the signal sources and characteristics over a 12-h period are negligible, and since the required ADC resolution is dependent on the PCA parameters estimated in the PCA training phase, the required ADC resolution is estimated every 12 h also. The real-time spike sorter power consumption is estimated to be $1.4 \mu\text{W}$ [5].

C. Resolution Adaptation Algorithm

Let n_i be the resolution for the i th ADC cell. During the ADC training phase, each spike that is received is digitized at 8-b resolution. The first spike sorter classifies the spike as having originated in a particular neuron using the 8-b representation of the spike which we know to be sufficiently accurate. Five additional real-time spike sorters are used to digitize 7 through 3 b representations of the signal. The resulting classification for each resolution is compared to the 8-b classification and if it is different, we say that a misclassification has occurred. The training phase is run for a very large number of spikes and at the end, misclassification rates are calculated for each possible resolution by dividing the total number of misclassifications by the number

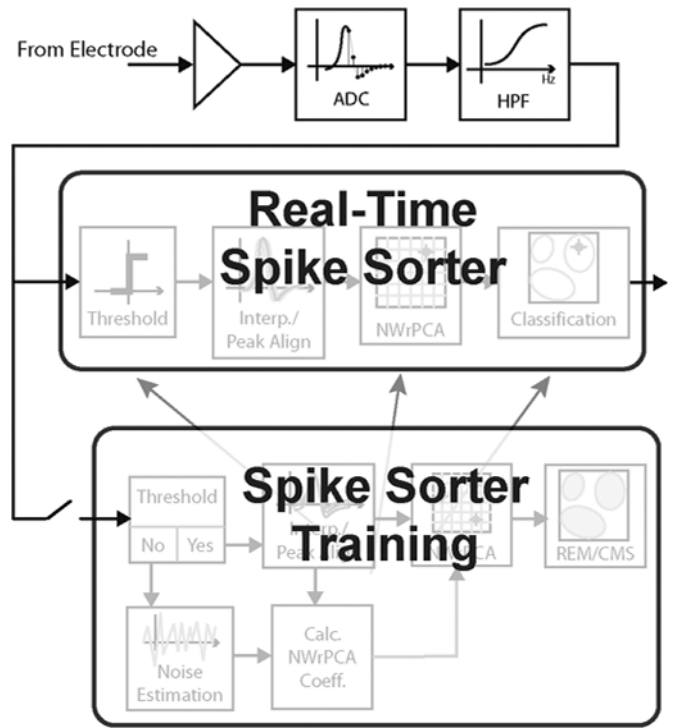


Fig. 6. Spike sorter consists of two major components: 1) real-time spike sorter, which runs continuously, and 2) spike sorter training, which must be run every 12 h.

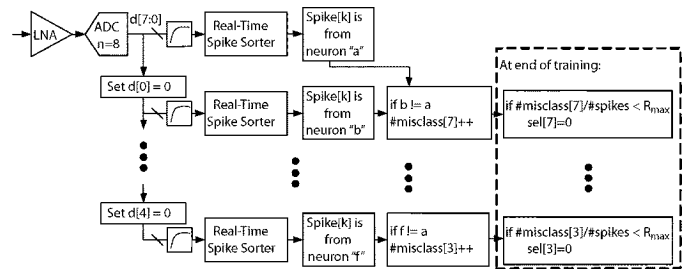


Fig. 7. Resolution determination and assignment method.

of spikes $R(n_i) = (\# \text{ misclassifications}(n_i)) / (\# \text{ spikes})$. The lowest resolution n_i for which the misclassification rate $R(n_i)$ is less than the maximum-allowable misclassification rate R_{\max} is chosen to be the ADC's resolution.

Resolution estimation uses five extra spike sorts for each electrode in a time window that is equal to the training period of the spike sorter, which is 120 s every 12 h. Since the real-time spike sorter consumes $1.4 \mu\text{W}$, this gives a power overhead for a resolution estimation of 20 nW per channel, or 5% of the ADC power with optimally assigned resolutions. The extra spike sorters are multiplexed across the channels for each training phase and so do not contribute a significant area overhead. This method does not interrupt the throughput as 8-b data are available throughout the calibration phase. Fig. 8 shows a sample resolution assignment for typical neural data from the motor cortex of a rhesus monkey with $R_{\max} = 1\%$. The spike sorter itself has a spike misclassification rate of approximately 5% based on measured neural data, so choosing $R_{\max} = 1\%$ does not materially compromise performance. We see that a

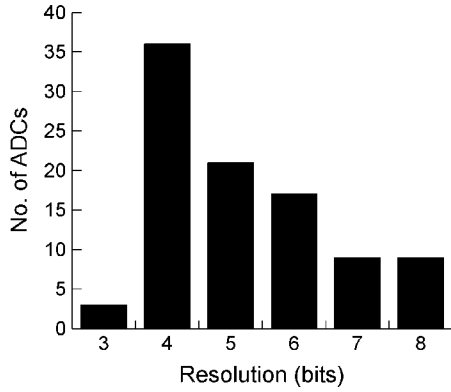


Fig. 8. Histogram of chosen ADC resolutions over 96 channels of measured neural data.

broad range of resolutions is assigned, clearly not all channels need full resolution all of the time. Therefore, an adaptive resolution ADC array has the potential to help reduce power consumption.

III. ADC CELL ARCHITECTURE

The potential efficiency improvement of the adaptive resolution ADC array will only be realized if the power overhead due to resolution adaptation is low compared to the power savings achieved. Regardless of how well the resolution adaptation performs, the results will be unconvincing and the ADC array will not be useful unless we build a state-of-the-art ADC, at least as power efficient as the previous work, with which to demonstrate the adaptive resolution technique. Therefore, choosing the ADC architecture, which is the most power efficient in this bandwidth and resolution space, is fundamental to this work. We do so by investigating theoretical lower bounds on ADC power consumption. The Appendix describes reported lower bounds on power consumption of matching-limited flash and pipeline ADC architectures while a theoretical power bound for charge redistribution successive approximation (CR-SAR) ADCs is derived in this section. All of these are plotted in Fig. 9.

A. Charge Redistribution Successive Approximation ADC

Successive approximation (SAR) ADCs are often used to realize low-to-moderate speed and medium-to-high resolution converters [9]. The fact that an SAR does not need any linear circuits, thus obviating the need for high bias currents makes it a very attractive architecture for ultra-low-power applications. An approximate lower bound for CR-SAR power consumption is derived here. Energy dissipation of the capacitor array (1) has been derived as a function of the minimum capacitance in the binary weighted charge redistribution sub-DAC, C_{unit} and the full-scale voltage V_{ref} [10].

$$E_{\text{switching},n\text{-bit}} = \sum_{i=1}^n 2^{n+1-2i} (2^i - 1) C_{\text{unit}} V_{\text{ref}}^2 \quad (1)$$

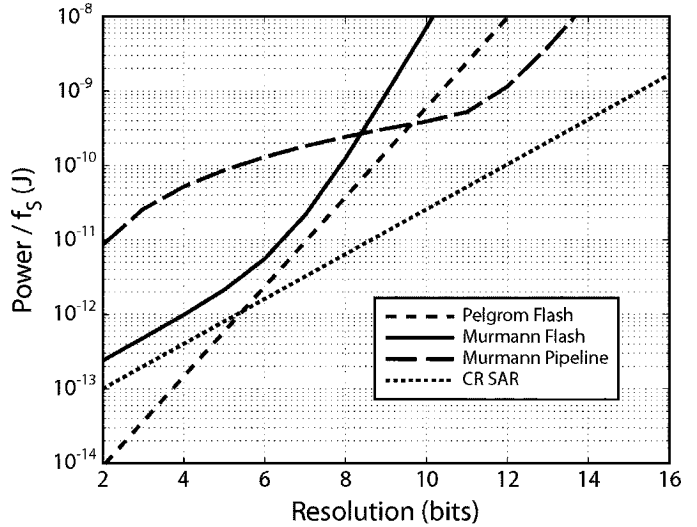


Fig. 9. Theoretical matching-limited ADC power bounds versus resolution in a typical 0.13- μm CMOS process.

which easily translates to power dissipation

$$\Rightarrow P_{\text{Cap Array}} = C_{\text{unit}} V_{\text{ref}}^2 f_s \sum_{i=1}^n 2^{n+1-2i} (2^i - 1), \quad (2)$$

The successive approximation register logic consists of two shift registers, each containing $n+1$ flip-flops. A rough estimate of the SAR power consumption can be found by assuming that all of the flip-flops are asynchronous settable and resettable and contain 2 NOR gates, 2 nand and 2 inverters [11], approximately equivalent to 5 nand gates. Further, assuming that all of these gates switch at frequency $n \cdot f_s$, where f_s is the signal bandwidth, then we have a close upper bound on the SAR dynamic power consumption as given

$$\Rightarrow P_{\text{SAR Logic}} \approx 10(n+1)n f_s E_{\text{gate}}. \quad (3)$$

The comparator power consumption can be estimated as

$$P_{\text{Comp}} = \pi \alpha A_{V_T} C_{\text{ox}} \sqrt{\frac{WL}{2}} n f_s \quad (4)$$

using the comparator power bound formula from (15) and recognizing that 1) the SAR has only one comparator and 2) the comparator switching rate increases by a factor n with respect to the comparator in a flash ADC. These power bounds are combined to give an estimate on the power consumption of a CR-SAR ADC

$$P_{\text{CR SAR}} = P_{\text{Cap Array}} + P_{\text{SAR Logic}} + P_{\text{Comp}}. \quad (5)$$

Equations (2)–(4), together, with process constants for a typical 0.13- μm process are substituted into (5) to generate curve CR SAR in Fig. 9.

B. Choice of Base-Cell Architecture

Fig. 9 shows that the CR-SAR architecture is more power efficient than pipeline or flash ADC architectures over a wide range of resolutions. This begs the question: Why are CR-SAR

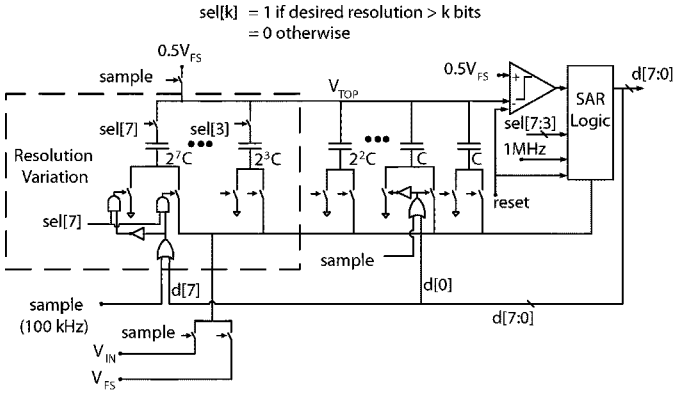


Fig. 10. Variable resolution SAR ADC cell.

ADCs not used for more applications? The answer is two-fold: 1) the capacitor array area increases exponentially with resolution, although there are ongoing attempts to overcome this obstacle [12] and 2) the CR-SAR is less suited to high-speed, high-resolution applications because the time to charge and discharge the capacitor array can be prohibitively large. The fundamental reason for the long charging times is that the minimum size capacitor in the charge redistribution DAC must be large enough to ensure that the mismatch between capacitors in the array is low enough to give the required resolution. Capacitor mismatch is technology dependent. Neither of these disadvantages of CR-SAR ADCs are an issue in this application which has low speed, 100 kSamples/s, and moderate resolution of 3 to 8 b. Furthermore, this capacitance mismatch is a function of the minimum linewidth and so CR-SAR ADC power efficiency improves with technology shrinkage. Fig. 9 shows that the CR-SAR architecture is most power efficient in the 6- to 8-b range, whereas flash ADCs may be more efficient at lower resolutions. The power consumption at the higher resolutions will dominate, so based on the theoretical analysis, we choose the CR-SAR architecture.

IV. VARIABLE RESOLUTION CR-SAR ADC

The operation and design of fixed resolution charge redistribution successive approximation (CR-SAR) ADCs were explained previously in the literature [13]. The fixed resolution CR-SAR ADC is modified to give a variable resolution converter as shown in Fig. 10. The larger capacitors are switched out for lower resolution, and the logic in the successive approximation register is reconfigured for lower resolution operations. Five select signals $sel[7], \dots, sel[3]$ are used to control the resolution. For each 1-b reduction in resolution, the total capacitance and, hence, the power dissipation in the capacitor array halves while the logic power scales linearly with resolution. In this section, we discuss the design of each subcircuit in Fig. 10.

A. Variable Resolution Capacitor Array

The largest capacitors are switched out to reduce resolution as illustrated in Fig. 10. The switches are placed on both sides of the capacitors. The parasitic capacitances at the top and bottom plates are appreciable and significant power would be dissipated in charging and discharging them without the double switches, degrading the efficiency at lower resolutions.

In [14], Lin derives the maximum-allowable capacitor mismatch as a function of the desired resolution n

$$\left(\frac{\Delta C}{C}\right)_{\max} = \frac{2^n}{2^{2n} - 2^n + 1} \quad (6)$$

where $(\Delta C)/C$ is the atch in capacitance of two nominally identical capacitors relative to their nominal value. So for 8-b resolution, we require

$$\left(\frac{\Delta C}{C}\right)_{8\text{-bit}} < 0.0039. \quad (7)$$

Equation (6) is approximately equal to $(1)/(2^n)$ for $n > 3$, which is the first-order estimate of the required accuracy.

The standard deviation of the ratio of capacitance mismatch between metal-insulator-metal (MiM) capacitors to their nominal capacitance value is given by

$$\sigma_{\Delta C/C} = \frac{A_{\Delta C/C}}{\sqrt{WL}} \quad (8)$$

where $A_{\Delta C/C}$ is approximately constant for the process and W and L are the width and length of the capacitor. Reference [15] reports that $A_{\Delta C/C} \approx 1\% \cdot \mu\text{m}$ is typical for 0.13- μm CMOS but this varies across processes. We restrict our design to using particular foundry recommended capacitor structures, overdesign for capacitor mismatch by a factor of two from this theoretical value, to accommodate variation and additional errors due to routing mismatch. Ultimately, we choose $4 \mu\text{m} \times 4 \mu\text{m}$ for the minimum size capacitor in the array which gives a capacitance of 20 fF. This is the unit capacitance C in Fig. 10.

MiM capacitors were used since they offer high capacitance per unit area and are inherently linear. In the variable resolution structure, the top plates of the capacitors are not connected to a single node; they must each connect to MOS devices. Technology layout constraints on interconnect and vias close to MiM caps and the use of top and bottom-plate switches prevent use of a single array capacitance as is usually employed in fixed resolution SAR ADCs. Instead, some separation is required between each capacitor which leads to a greater array area than in a fixed resolution converter. Nevertheless, a common centroid structure is achieved for the entire array which should improve the mismatch beyond the values mentioned before.

The power overhead for reconfigurability in the capacitor array is determined by 10 additional AND gates, whose state is changed, at most, once every 12 h (when the ADC and spike sorter training phases are run) and so consumes negligible power.

B. Reference Switch and Top Plate Switches

The switch connecting $0.5 V_{FS}$ to V_{Top} is called the reference switch. During the sample phase, the reference switch is closed, and the bottom plates of the capacitor array are connected to V_{IN} . The hold phase is reached by opening the reference switch and then switching the bottom plates to ground. The voltage at the top plate of the array capacitance at the end of the hold phase is given by

$$V_{Top} = 0.5 V_{FS} - V_{IN} - \frac{I_{Leak} t_{Leak}}{2^n C} \quad (9)$$

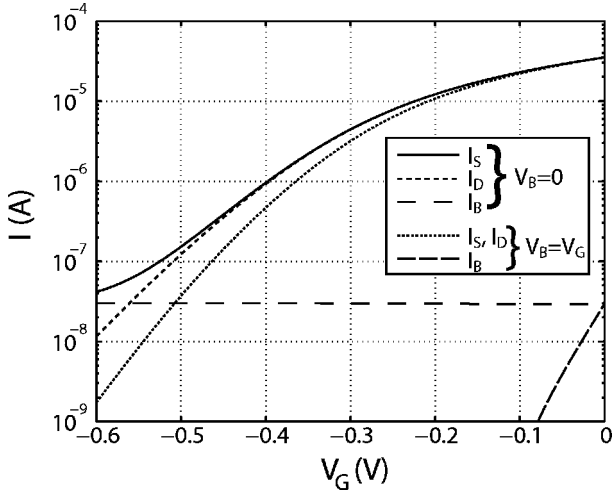


Fig. 11. Leakage currents of min size NMOS; $V_D = 0.6$ V, $V_S = -0.6$ V.

where I_{Leak} is the leakage current from the top plate and t_{Leak} is the time from the beginning of the hold phase to the first switch of the decision phase. Assuming that $V_{FS} = 1.2$ V, the top plate voltage varies from 0.6 V to -0.6 V as V_{IN} varies from 0 V to 1.2 V. If the reference switch is realized using a pass gate switch in which the applied gate and substrate voltages are 0 V or V_{DD} , then for $V_{IN} = 1.2$ V, $V_{GS,NMOS} = 0.6$ V during hold mode (i.e., the gate is partially ON when it should be OFF). This results in substantial leakage to the top plate $I_{Leak} = -35$ μ A, as shown by the rightmost point of the solid black curve in Fig. 11. This causes a large signal-dependent error. Therefore, the NMOS gate is driven with $V_{gate,Low} \leq -0.6$ V to turn the switch off.

I_{Leak} then becomes -41 nA, dominated by leakage current across the P -bulk/ N^+ diode which has a forward bias of 0.6 V, as shown by the leftmost point of the “ I_B at the $V_B = 0$ ” curve in Fig. 11. Since $t_{Leak} \approx 0.75$ μ s, the resulting error in the top plate voltage is therefore given by

$$\Delta V_{Top} = \frac{I_{Leak} t_{Leak}}{2^N C} \geq 6.0 \text{ mV} = 1.28 \text{ LSB} \quad (10)$$

which is still an order of magnitude or more too large. The cause of this error is that we use fine geometry CMOS to accommodate the DSP required by the IPP. One side effect of this is high substrate doping, giving a bulk-source diode with low junction potential and, thus, high leakage current. Furthermore, this is a slow speed application and, thus, a given leakage translates to larger voltage error. This error would not be noticed by those who use the SAR ADC architecture in older CMOS processes or those working at faster speeds in this process. One tactic at this point is to reduce the hold time, but that requires much finer clock phases and greater power consumption to generate them.

Our solution comes from the fact that fine geometry processes offer a deep N-well and we can take advantage of that to isolate and bias the bulk of the reference switch NMOS to below -0.6 V. This reduces the leakage to 1.7 nA as given by the leftmost point on the “ I_S at $V_B = V_G$ ” curve in Fig. 11. The source-bulk diode leakage is now negligible. $I_{Leak} = 1.7$ nA gives $\Delta V_{Top} = 0.25$ mV = 0.05 LSB, no longer impairing the ADC performance. Similarly, the bulks of the NMOS sides of

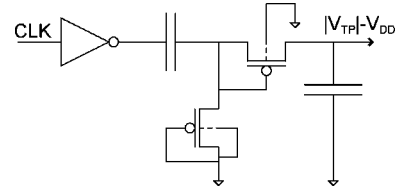


Fig. 12. Charge pump to generate the negative voltage.

the top plate switches are tied to ≤ -0.6 V to prevent leakage to the substrate when the sel_i for that switch is low. Since the drain of that NMOS is then floating, there is negligible drain-to-source leakage and no need to drive that gate with -0.6 V.

We must generate this negative voltage on-chip. Fig. 12 shows a simple single-stage charge pump used to generate a voltage of $|V_{TP}| - V_{DD} \approx -0.7$ V [16]. The generated voltage is applied to the bulk of the NMOS transistors in the top plate switches and the reference switch. The gate voltage for the NMOS in the reference switch is generated by passing the sample signal through a pair of inverters, the second of which has $V_{DD} = 1.2$ V and V_{SS} connected to the charge-pump output.

All of the switches shown in Fig. 10 are implemented as pass gates to reduce charge injection. This strategy, coupled with the fact that the gates are loaded on one side by low impedances (i.e., V_{FS} , V_{IN} and ground), ensures charge injection onto the capacitor array is negligible.

C. Comparator

The comparator is realized by using a simple resettable latch, as discussed in [17] and shown in Fig. 13. Low comparator power requires low tail current, but that tail current must be large enough to discharge the load capacitance in the time available. The load capacitance is dominated by the comparator’s intrinsic capacitance, not by the buffer. Therefore, minimizing comparator power dissipation demands small WL of the input and latch devices. Conversely, low V_T mismatch and low $1/f$ noise require large WL of the input devices. Noise simulations were used to find an acceptable tradeoff and suggested a minimum acceptable bias current of 100 nA. This current level results in a slew limited output and the “low” output of the latch stage only reaches 0.8 V in the available time (while the high output reaches $V_{DD} = 1.2$ V). Increasing the tail current would give close to full CMOS output swing, but at an unacceptable power cost. Instead, we accept an output swing of 0.8 V to 1.2 V and increase that to the full CMOS swing using two buffers with staggered thresholds. The first buffer in Fig. 13 uses a core PMOS and input/output (I/O) NMOS. Core devices have lower V_T than I/O devices so this, together with device sizing, moves the switching point close to the midrange of the signal swing of approximately 1.0 V. The second buffer is sized to switch at 0.7 V and sharpens the edge before passing the signal to the SAR logic.

D. Successive Approximation Register

The successive approximation register is based on the work of [18]. In order to accommodate resolution variation, a stage is powered down for each 1-b reduction in resolution, eliminating leakage in that stage. Custom digital design, optimized

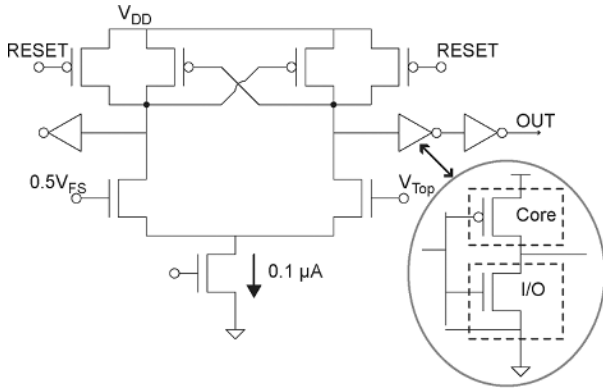


Fig. 13. Comparator schematic.

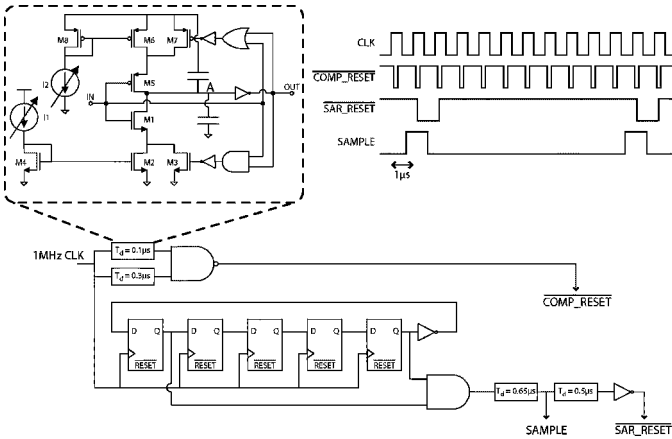


Fig. 14. Timing generation circuits.

for low-speed and low-power performance, is used for all logic circuits which are implemented using 2.5-V I/O transistors to minimize leakage. The set, reset, and input signals of the first flip-flop in the upper register must be reconfigured for each resolution setting and so ten 2-b muxes must be added. This results in negligible power and area overhead.

E. Timing Generation

All timing signals are derived on-chip from a single 1-MHz master clock signal as shown in Fig. 14. The comparator reset signal is a 0.2- μ s-wide pulse with a 1- μ s period. It is generated from the clock by using two variable delay stages and a NAND gate. The sample and SAR reset signals both have a 10- μ s period. A 5-b Johnson counter is used to divide the master clock frequency down to a 200-kHz clock signal, two phases of which are added and passed through variable delay stages to generate the sample and SAR reset signals.

A schematic of the variable delay stage is shown in the inset in Fig. 14. The delay between a positive input edge and the subsequent positive output edge is controlled by the RC time constant of the first half of discharging node A. That is, in turn, controlled by the current in M4. Similarly, the delay between a negative input edge and the subsequent negative output edge is controlled by the current in M8. The control currents are set once, during the initial testing of each device, prior to deployment as follows. The nominal value of I_1 and I_2 for each delay

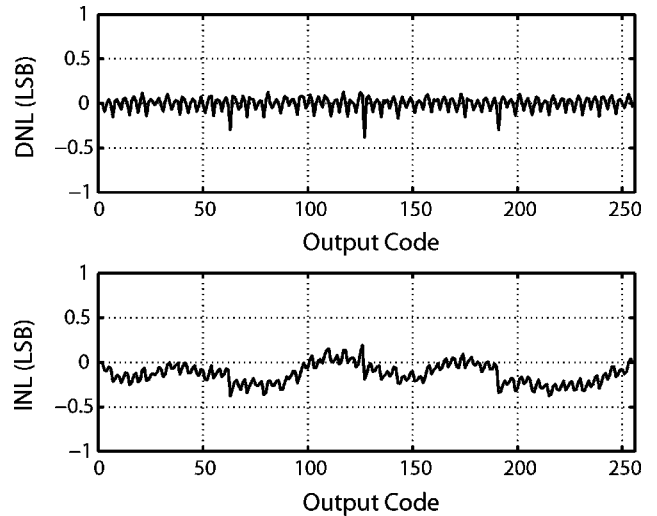


Fig. 15. Measured differential and integral nonlinearities versus output code at the 8-b resolution setting.

is designed with additional 3-b fine tuning of these currents which allows delays to be set accurately within $\pm 0.1 \mu$ s. The clock signals are probed, and the delay is tuned to the correct value by varying the external control bits for each delay element which control binary-weighted current sources which fine-tune the variable current sources in Fig. 14. Once the correct current has been determined in the initial testing phase, the current need not be retuned since the maximum possible variation in delay due to temperature and supply variations is much less than the required $\pm 0.25 \mu$ s accuracy.

V. MEASURED PERFORMANCE

The testing of the base ADC was carried out according to [19] and [20]. Fig. 15 shows the measured integral nonlinearity (INL) and differential nonlinearity (DNL) at the 8-b resolution setting. Both are well within ± 0.5 LSB. Spikes in the plots are visible at codes 63, 127, and 191, corresponding to nonidealities due to the capacitance of the routing to the 2^6 C and 2^7 C capacitances. Of course, INL and DNL look even better at the lower resolution settings.

Fig. 16 shows a discrete Fourier transform (DFT) of measured digital output at 8-b resolution for a 1-kHz sinusoidal input based on 8192 samples. The 3rd, 7th, and 9th harmonics are clearly visible. The effective number of bits (ENOB) is 7.8 b; the signal-to-noise and distortion ratio (SNDR) is 48.6 dB, the spurious-free dynamic range (SFDR) is 61.0 dB, and the total harmonic distortion (THD) is -56.5 dB. The ENOB decreases to 7.55 b as the signal frequency is increased to 50 kHz, as can be seen in Fig. 17. The maximum signal bandwidth we expect is about 15 kHz.

The derivation of power bounds and particularly (15) suggest that an appropriate figure of merit (FOM) for matching limited converters is given by (11). This FOM corresponds to energy dissipation per conversion step and is a widely employed metric

$$\text{FOM} = \frac{\text{Power}}{2^n f_{\text{samp}}} = 48 \text{ fJ/conversion step} \quad (11)$$

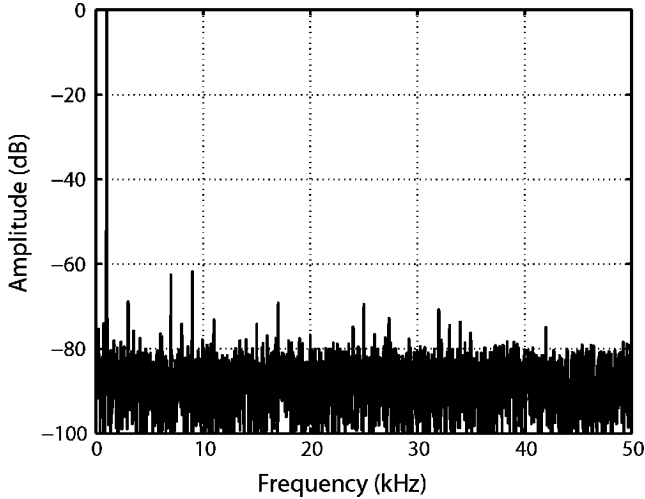


Fig. 16. FFT of the measured digital output for a 1-kHz input at an 8-b setting.

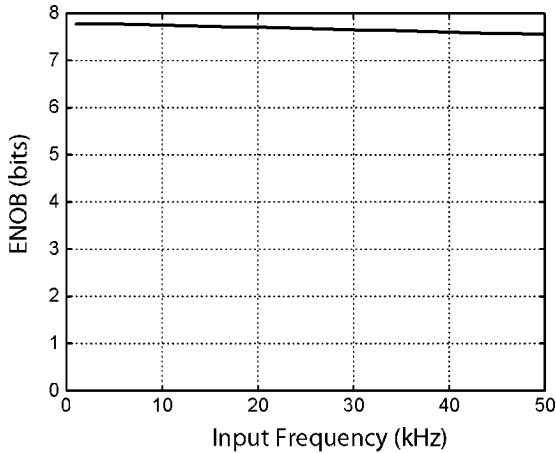


Fig. 17. ENOB versus frequency at an 8-b setting.

where f_{samp} is the sampling rate in the case of a Nyquist rate converter. This ADC achieves a $\text{FOM} = 48 \text{ fJ/conversion step}$ at the 8-b setting.

The solid curve in Fig. 18 plots the measured ADC power consumption at each resolution setting, averaged over many sampled words with the input set to a full-scale sinusoid at a frequency not harmonically related to the sample rate. The dashed lines are based on simulation and show the power consumption of the capacitor array, SAR logic, and comparator. We see that the ADC cell power consumption increases strongly with resolution from $0.23 \mu\text{W}$ at 3 b to $0.90 \mu\text{W}$ at 8-b resolution. This compares very favorably with the leading previous work for this application: $99 \mu\text{W}$ per electrode for 10 output bits at 15 kSamples/s [21]. At low resolutions, comparator power dominates. If the resolutions of 96 channels are assigned according to Fig. 8, then the total power dissipation is $37 \mu\text{W}$. The resolution adaptation reduces ADC power consumption by a factor of 2.3 for this device.

Table I shows a comparison of the performance of the ADC herein with some of the leading SAR ADCs previously reported in this resolution and bandwidth space. The ENOB

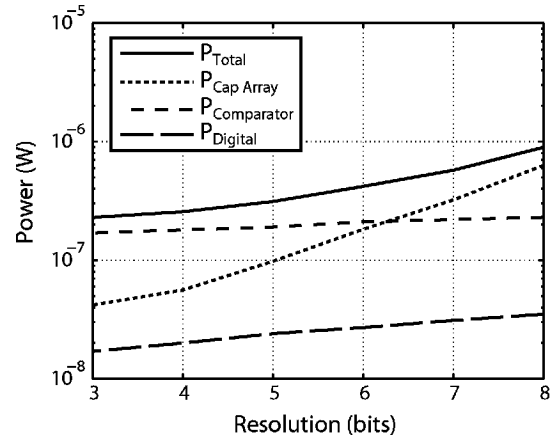


Fig. 18. Power consumption of the base ADC cell versus resolution.

TABLE I
SUCCESSIVE APPROXIMATION ADC PERFORMANCE COMPARISON

Ref.	ENOB (bits)	Sample Rate (<i>Samples/s</i>)	Power (μW)	FOM (fJ/step size)
[17]	7.5	100k	19	254
[23]	8	100k	19	742
[24]	7.6	200k	2.8	65
[10]	7.5	500k	7.75	0.09
[25]	9.4	100k	3.8	56
[22]	8.8	1M	19	4.4
[26]	12	100k	19	61
This work without resolution adaptation	7.55	100k	0.90	48
This work with resolution adaptation	7.55	100k	0.39	21

values quoted are those at an input frequency equal to half the sampling frequency. The FOM used is the energy per conversion step size, given in (11). We see that our underlying ADC circuit without resolution adaptation performs a little more efficiently than the best of its peers, bar one, achieving an energy per-step-size FOM of 48 fJ and that when resolution adaptation is included this work performs about three times more efficiently than the best of its peers, again bar one. That better performing device is the excellent work reported in [22]. While [22] does run at a ten times greater sampling rate, which reduces the contribution of the static power to the FOM, that is only a minor portion of the overall advantage. The primary reason why [22] performs so much better is through the use of adiabatic charging of the charge redistribution DAC. Furthermore, [10] uses a novel switching sequence and various other works use different configurations of split capacitor arrays in the DAC to reduce DAC power consumption. All three of these power reduction techniques are independent of the resolution adaptation technique presented here and so may be combined with our technique to further reduce power consumption. These other techniques all focus on reducing DAC power consumption and do not reduce comparator power. Future work could implement a comparator whose power consumption scales with resolution which would further increase the energy savings gained by resolution adaptation.

A microphotograph of the ADC implemented in $0.13\text{-}\mu\text{m}$ CMOS is shown in Fig. 19.

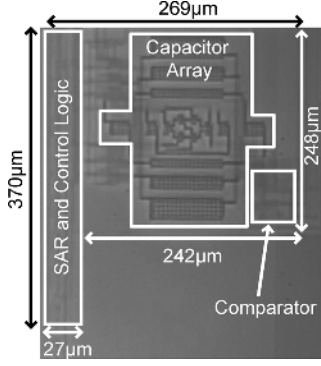


Fig. 19. Micrograph of the ADC cell on die area = 0.07 mm².

VI. CONCLUSION

This work has shown that a CR-SAR ADC is most appropriate for the IPP application, introduced algorithms, and circuit techniques to adapt the ADC resolution to minimize power consumption while maintaining maximum IPP accuracy; and optimized the CR-SAR ADC architecture for slow-speed applications in short-channel CMOS (as required for our SoC IPP). The ADC cell was demonstrated in 0.13- μm CMOS, and the measured performance shows an energy per conversion step of 48 fJ/conv step at 100 kSamples/s at the 8-b setting. Furthermore, the adaptive resolution technique, together with the variable resolution ADC, reduces power consumption by 2.3 times for typical neural data. The device demonstrates dramatically reduced power consumption for the digitization of neural signals compared to the leading previously reported work. The average ADC cell power consumption is 0.39 μW for effective 8-b resolution, giving a projected power consumption of 38 μW for the 96-cell ADC array.

The adaptive ADC performance could be improved by varying the comparator power consumption with resolution as mentioned in Section V.

APPENDIX

THEORETICAL LOWER BOUNDS ON POWER CONSUMPTION OF MATCHING-LIMITED ADCs

There has been a considerable amount of work done in estimating theoretical power bounds for process-limited flash and pipeline ADC architectures. The leading results are explained and summarized in this Appendix and plotted in Fig. 9.

A. Flash

For this application, which requires low speed and moderate resolution, performance is likely to be limited by component matching. It can be shown [27] that the standard deviation of the threshold voltage mismatch between a pair of nominally identical transistors with width W and length L is given by

$$\sigma(V_T) = \frac{A_{V_T}}{\sqrt{WL}} \quad (12)$$

where A_{V_T} is the threshold voltage mismatch coefficient, a constant for a given technology. In 0.13- μm CMOS, A_{V_T} is about

4.5 mV \cdot μm . From the observation that the energy required to switch a latch pair of transistors from the metastable state to a fixed state with one- σ certainty is

$$E_{\sigma(V_T)} = C_{\text{gate}}\sigma(V_T) = C_{\text{ox}}A_{V_T}^2 \quad (13)$$

where C_{ox} is the gate-oxide capacitance per-unit area, Pelgrom [27] derives the minimum ADC energy per step size per pair of matching critical transistors, to ensure that a least-significant bit (LSB) change in the input is detected when matching limited

$$\frac{\text{Power}}{2^n f_s} = \pi\alpha\sqrt{WLN_\alpha x_d} \quad (14)$$

where n is the number of bits to be decoded, f_s is the signal bandwidth, and α is chosen to be equal to 10 to allow giving a confidence interval assuming a Gaussian distribution in mismatch. Simple algebraic manipulation gives a power bound for matching-limited flash ADCs with more familiar process constants

$$P_{\text{Matching Limited Flash}} = N_C N_M \pi \alpha A_{V_T} C_{\text{ox}} \sqrt{\frac{WL}{2}} 2^n f_s \quad (15)$$

where N_M is the total number of pairs of matching critical transistors per comparator and N_C is the total number of comparators. $N_C = 2^n - 1$ for a flash ADC. Assuming $N_M = 1$, this gives a lower bound on the power consumption of an 8-b 50-kHz ADC of 0.3 μW for minimum-sized (0.2/0.13) devices and 2 μW for devices with a channel area of 1 μm^2 in 0.13- μm CMOS. Equation(15) is plotted as the curve Pelgrom Flash in Fig. 9 using typical process constants for 0.13- μm CMOS. Values obtained by using this formula are lower than those observed in modern low-speed moderate resolution flash ADCs. This discrepancy arises because 1) the derivation considers intrinsic capacitance only. Parasitic capacitors require roughly the same amount of charge as the intrinsic capacitance and 2) not only does the depletion charge contribute to the uncertainty of an LSB, so do W , L dependencies, mobility variations, and so on. Factoring in these considerations, the power is expected to be about 10 times the value predicted above. This lower bound assumes that flash comparators are simple regenerative latches. This is possible for low-to-moderate resolution if the transistors are sized large enough, but as shown, this sizing increases power consumption.

Another approach to finding a lower bound on flash ADC power consumption is given in [28]. This assumes that the comparators operate in a Class A manner and that 1/2 LSB matching with 3 σ -confidence is designed. It further assumes $2^n - 1$ components, partial supply usage represented by the factor α , and includes additional dynamic switching energy E_{dyn} per clock cycle. The derived lower bound on flash ADC energy consumption is given by

$$P = \left(12\pi \frac{1}{\alpha} C_{\text{ox}} A_{V_T}^2 2^{3n} + 2E_{\text{dyn}} 2^n \right) f_s \quad (16)$$

which corresponds to a power consumption of 7.3 μW for 8-b resolution and a 50-kHz signal bandwidth in 0.13- μm CMOS.

Class A operation may be required for high-resolution or high-speed flash ADCs wherein offsets must be canceled, but simple regenerative latches are feasible for moderate and low-resolution ADCs where large device sizes are more acceptable. So the lower bound on power consumption for a flash should be between (15) and (16) for this application. Equation (16) is plotted as a curve *Murmann Flash* in Fig. 9 using typical process constants for 0.13- μm CMOS.

B. Pipeline

A lower bound on power consumption of pipeline ADCs is derived in [28] which uses the power consumption of a switched-capacitor integrator (17) as a starting point

$$P_{\text{SC Integrator}} \approx \frac{16}{\alpha} N_{\text{set}} n_f k_B T \text{SNR} f_s \quad (17)$$

where k_B is Boltzmann's constant, T is the absolute temperature, SNR is the signal-to-noise ratio, f_{sig} is the signal frequency, N_{set} is the number of time constants required to achieve the desired settling,¹ n_f is a multiplier for $(k_B T)/(C)$ to account for excess circuit noise, α quantifies the fraction of supply voltage used for signal swing and is chosen to be 2/3.² Reference [28] presents an algorithm to estimate the power for each stage of the pipeline ADC relative to that switched-capacitor stage, taking into consideration the minimum feature size, noise, and mismatch constraints. The result, using process parameters for 0.13- μm CMOS, is plotted as a curve *Murmann Pipeline* in Fig. 9.

ACKNOWLEDGMENT

The first author was supported by the Lu Stanford Graduate Fellowship. The authors would like to thank Taiwan Semiconductor Manufacturing Corporation for their generous fabrication support.

REFERENCES

- [1] S. O'Driscoll, T. H. Meng, K. V. Shenoy, and C. T. Kemere, "Neurons to silicon: Implantable prosthesis processor," in *Proc. Int. Solid State Circuits Conf.*, 2006, pp. 552–553.
- [2] R. R. Harrison and C. Charles, "A low-power low-noise CMOS amplifier for neural recording applications," *IEEE J. Solid State Circuits*, vol. 38, no. 6, pp. 958–965, Jun. 2003.
- [3] S. O'Driscoll and T. H. Meng, "Adaptive resolution ADC array for neural implant," in *Proc. IEEE Eng. Med. Biol. Soc. Conf.*, 2009, pp. 1053–1056.
- [4] M. Sahani, "Latent variable models for neural data analysis," Ph.D. dissertation, Dept. Comput. Neural Syst., California Inst. Technol., Pasadena, CA, 1999.
- [5] Z. Zumsteg, C. Kemere, S. O'Driscoll, G. Santhanam, R. Ahmed, K. Shenoy, and T. Meng, "Power feasibility of implantable digital sorting circuits for neural prosthetic systems," *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 13, no. 3, pp. 272–279, Sep. 2005.
- [6] B. Yu, S. Ryu, G. Santhanam, M. Churchland, and K. V. Shenoy, "Improving neural prosthetic system performance by combining plan and peri-movement activity," in *Proc. IEEE Eng. Med. Biol. Soc. Conf.*, 2004, pp. 4516–4519.
- [7] S. O'Driscoll, A. S. Y. Poon, and T. H. Meng, "A mm-sized implantable power receiver with adaptive link compensation," in *Proc. Int. Solid State Circuits Conf.*, 2009, pp. 294–295.
- [8] A. S. Y. Poon, S. O'Driscoll, and T. H. Meng, "Optimal operating frequency in wireless power transmission for implantable devices," in *Proc. IEEE Eng. Med. Biol. Soc. Conf.*, 2007, pp. 5673–5678.

- [9] M. IC, "A simple ADC comparison matrix," Maxim Appl. Note 2094, 2003.
- [10] Y. K. Chang, C. Wang, and C. Wang, "A 8-bit 500-kS/s low power SAR ADC for bio-medical applications," in *Proc. Asian Solid State Circuits Conf.*, 2007, pp. 228–231.
- [11] N. Weste and D. Harris, *CMOS VLSI Design: A Circuits and Systems Perspective*. Reading, MA: Addison-Wesley, 2005.
- [12] J. Craninckx and G. Van der Plas, "A 65 fJ/conversion-step 0-to-50 MS/s 0-to-0.7 mW 9 b charge-sharing SAR ADC in 90 nm digital CMOS," in *Proc. Int. Solid State Circuits Conf.*, 2007, pp. 246–247.
- [13] J. McCreary and P. Gray, "All MOS charge-redistribution ADC techniques," *IEEE J. Solid State Circuits*, vol. SSC-10, no. 6, pp. 371–379, Dec. 1975.
- [14] Z. Lin, H. Yang, L. Zhong, J. Sun, and S. Xia, "Modeling of capacitor array mismatch effect in embedded CMOS CR SAR ADC," in *Proc. 6th Int. Conf. ASICs*, Oct. 2005, vol. 2, pp. 979–982.
- [15] C. Diaz, D. Tang, and J.-C. Sun, "CMOS technology for MS/RF SoC," *IEEE Trans. Electron Devices*, vol. 50, no. 3, pp. 81–84, Mar. 2003.
- [16] R. J. Baker, H. W. Li, and D. E. Boyce, *CMOS Circuit Design, Layout, and Simulation*. New York: Wiley, 1998.
- [17] M. Scott, B. Boser, and K. Pister, "An ultralow-energy ADC for smart dust," *IEEE J. Solid State Circuits*, vol. 38, no. 7, pp. 1123–1129, Jul. 2003.
- [18] T. O. Anderson, "Optimum control logic for successive approximation ADCs," *Computer Design*, vol. 11, pp. 81–84, 1972.
- [19] M. I. Products, "Defining and testing dynamic parameters in high-speed ADCs, Part 1," Maxim Appl. Note 728, 2001.
- [20] M. I. Products, "Dynamic testing of high-speed ADCs, Part 2," Maxim Appl. Note 729, 2002.
- [21] R. R. Harrison, P. T. Watkins, R. J. Kier, D. J. Black, R. O. Lovejoy, R. A. Normann, and F. Solzbacher, "Design and testing of an integrated circuit for multi-electrode neural recording," in *Proc. Int. Conf. VLSI Design*, 2007, pp. 907–912.
- [22] M. van Elzakker, E. van Tuijl, P. Geraedts, D. Schinkel, E. Klumperink, and B. Nauta, "A 1.9 μW 4.4 fJ/conversion-step 10 b 1 MS/s charge-redistribution ADC," in *Proc. ISSCC Dig. Tech. Papers*, 2008, pp. 244–245.
- [23] N. Verma and A. Chandrakasan, "An ultra low energy 12-bit rate-resolution scalable SAR ADC for wireless sensor nodes," *IEEE J. Solid-State Circuits*, vol. 42, no. 6, pp. 1196–1205, Jun. 2007.
- [24] H.-C. Hong and G.-M. Lee, "A 65-fJ/conversion-step 0.9-V 200-kS/s rail-to-rail 8-bit successive approximation ADC," *IEEE J. Solid-State Circuits*, vol. 42, no. 10, pp. 2161–2168, Oct. 2007.
- [25] A. Agnes, E. Bonizzoni, P. Malcovati, and F. Maloberti, "A 9.4-ENOB 1 V 3.8 μW 100 kS/s SAR ADC with time-domain comparator," in *Proc. ISSCC Dig. Tech. Papers*, 2008, pp. 246–247.
- [26] N. Verma, A. Shoeb, B. J. , D. J. , G. J. , and A. Chandrakasan, "A micro-power EEG acquisition SoC with integrated feature extraction processor for a chronic seizure detection system," *IEEE J. Solid-State Circuits*, vol. 45, no. 4, pp. 804–816, Apr. 2010.
- [27] M. Pelgrom, E. Sanchez-Sinencio, and A. Andreou, *Low-Voltage/Low-Power Integrated Circuits and Systems, Chapter 14 Low-Power CMOS Data Conversion*. New York: IEEE, 1999.
- [28] B. Murmann, "Limits on ADC power dissipation," presented at the 14th Workshop on Advances in Analog Circuit Design (AACD), Limerick, Ireland, 2005.



Stephen O'Driscoll (S'01–M'09) received the B.E. degree in electrical engineering from University College Cork (UCC), Ireland, in 2001, and the M.Sc. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 2005 and 2009, respectively, where he received the Lu Stanford Graduate Fellowship.

During 1999 and 2000, he designed and developed microwave circuits at Farran Technology, Ballincollig, Ireland. From 2001 to 2003, he was with Cypress Semiconductor, San José, CA, where

he designed analog circuits for clock and data recovery phase-locked loops for wireline communications. He joined the Department of Electrical and Computer Engineering, University of California, Davis, as an Assistant Professor in 2009. His research focuses on analog and radio-frequency integrated-circuit design for biomedical and other low-power applications, wireless power transfer, system-configured analog circuits, and analog optimization.

Dr. O'Driscoll received the Motorola and SCI Fellowships and the Institute of Electrical Engineers prize from UCC.

¹ $N_{\text{set}} \approx \ln(2^n)$ where the number of bits, $n \approx (\text{SNR}_{\text{dB}} - 1.76)/(6.02)$

² $\alpha = (2 * V_{\text{sig}})/(V_{\text{DD}})$, where V_{sig} is the signal amplitude



Krishna V. Shenoy (S'87–M'01–SM'06) received the B.S. degree in electrical engineering from the University of California, Irvine, in 1990, and the M.S. and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1992 and 1995, respectively.

He was a Neurobiology Postdoctoral Fellow at the California Institute of Technology, Pasadena, from 1995 to 2001 and then joined Stanford University, Stanford, CA, where he is an Associate Professor in the Departments of Electrical Engineering and

Bioengineering, and in the Neurosciences Program. His research interests include computational motor neurophysiology and neural prosthetic system design.

Dr. Shenoy received the 1996 Hertz Foundation Doctoral Thesis Prize, a Burroughs Wellcome Fund Career Award in the Biomedical Sciences, an Alfred P. Sloan Research Fellowship, a McKnight Endowment Fund in Neuroscience Technological Innovations in Neurosciences Award, and a 2009 National Institutes of Health Directors Pioneer Award.



Teresa H. Meng (S'82–M'83–SM'93–F'99) received the Ph.D. degree in electrical engineering and computer science from the University of California, Berkeley, in 1988.

She is the Reid Weaver Dennis Professor of Electrical Engineering at Stanford University, Stanford, CA. Her current research interests focus on neural signal processing and bioimplant technologies. In 1999, she left Stanford University and founded Atheros Communications (NASDAQ: ATHR), which is a leading developer of semiconductor system solutions for wireless communications products. She returned to Stanford University in 2000 to continue her research and teaching.

Dr. Meng received the 2009 IEEE Donald O. Pederson Award, the DEMO Lifetime Achievement Award, the McKnight Technological Innovations in Neurosciences Award in 2007, the Distinguished Lecturer Award from the IEEE Signal Processing Society in 2004, the Bosch Faculty Scholar Award in 2003, the Innovator of the Year Award by MIT Sloan School eBA in 2002, and the CIO 20/20 Vision Award, a Best Paper Award from the IEEE Signal Processing Society, a National Science Foundation Presidential Young Investigator Award, an ONR Young Investigator Award, and an IBM Faculty Development Award, all in 1989. In 2002, she was named one of the Top 10 Entrepreneurs by Red Herring for 2001. Dr. Meng is a member of the National Academy of Engineering.