29.8 SHARC: Self-Healing Analog with RRAM and CNFETs

Aya G Amer, Rebecca Ho, Gage Hills, Anantha P Chandrakasan, Max M Shulaker

Massachusetts Institute of Technology, Cambridge, MA

Carbon nanotube (CNT) field-effect transistors (CNFETs) are a promising emerging technology for energy-efficient electronics (Fig. 1). Despite this promise, CNTs are subject to substantial inherent imperfections; every ensemble of CNTs includes some percentage of metallic CNTs (m-CNTs). m-CNTs result in conductive shorts between CNFET source and drain, resulting in excessive leakage and degraded (potentially incorrect) circuit functionality (Fig. 1). Several techniques have been developed to remove the majority of m-CNTs (no technique today removes 100% of m-CNTs). While these techniques enabled the first digital CNFET circuits, it is still not possible to realize large-scale CNFET analog or mixed-signal CNFET circuits due to m-CNTs. As shown in Fig. 1, while a digital logic gate can still function correctly in the presence of a small fraction of m-CNTs (but with degraded resilience to noise) [1], a single m-CNT in an analog circuit can result in catastrophic failure (e.g., degrading amplifier gain resulting in functional failure of circuit blocks such as ADCs and DACs)[1]. This paper presents a circuit design technique, Self-Healing Analog with RRAM and CNFETs (SHARC), that leverages the programmability of non-volatile resistive RAM (RRAM) to automatically "self-heal" analog circuits in the presence of m-CNTs. Using SHARC, we experimentally demonstrate analog CNFET circuits robust to m-CNTs as well as the first mixed-signals CNFET sub-system (4-bit DAC and SAR ADC; these are the largest reported complementary (CMOS) CNFET circuit demonstrations to-date).

The working principle of SHARC is as follows (Fig. 2): (1) Fabrication: each CNFET is split into multiple minimum width FETs (*i.e., "sub-CNFETs"*), with a RRAM cell fabricated directly under (or over) the source or drain contact of each sub-CNFET. (2) Initialization: a positive voltage is applied across the RRAM electrodes (top-to-bottom) to program the RRAM into a low resistance state (LRS) (this is the RRAM "FORM" operation; the RRAM stack is chosen so this FORM voltage is above the operating voltage of the circuit so that during operation the RRAM is not inadvertently programmed). (3) Self-Healing: the gate terminals of the CNFETs are biased to turn the CNFETs "OFF" and a negative voltage (i.e., opposite polarity used during the previous FORM operation) is applied across the CNFET and RRAM. CNFETs with only semiconducting CNTs (s-CNTs) in the channel do not conduct significant current, and thus the RRAMs in series with those CNFETs remain in a low resistance state. However, CNFETs containing m-CNTs in the channel conduct current (m-CNTs are not controlled by the gate and conduct current regardless of gate bias), and with sufficient current under negative bias the RRAM is re-programmed into a high-resistance state (HRS) (the RRAM "RESET" operation). To ensure RRAM RESET, the RRAM stack should be designed to require less reset current than the off-state leakage current through an m-CNT (<100 µA for a scaled technology node). (4) Final circuit fabrication: Perform final chip fabrication (e.g., additional metal layers). After performing SHARC, sub-CNFETs containing m-CNTs are in series with the RESET high-resistance RRAM (effectively removing those sub-CNFETs from the circuit), while CNFETs containing only s-CNTs are in series with SET low-resistance RRAM. Fig. 2 shows measured current-voltage (IV) transfer curves for typical CNFET prior- and post-performing the SHARC self-healing. Prior to the self-healing, the CNFET exhibits substantial off-state leakage current due to the presence of m-CNTs (typical measured I_{ON}/I_{OFF} is <10 for a CNFET containing m-CNTs). Post self-healing, the CNFETs exhibit I_{ON}/I_{OFF} >1000 and I_{OFF} reduces by >800X due to the series RRAM resetting to a high-resistance state during the self-healing.

Importantly, SHARC is a broad technique that can be applied to a wide range of analog circuit blocks. Additional factors should also be taken into consideration to minimize impacting circuit performance. For example, the absolute values and ratios between the RRAM high and low resistance states impact the leakage power (e.g., increased RRAM high-resistance decreases leakage current through sub-CNFETs containing m-CNTs) and output signal swing (e.g., increased RRAM low-resistance decreases available voltage headroom). Moreover, the location of the RRAMs in the circuit may impact gain and input referred noise (e.g., for the case of input transistors to amplifiers, the RRAM should be placed in series with the drain terminal of the sub-CNFETs rather than the source terminal). Fig. 3 illustrates the benefit of SHARC on various amplifiers configurations. Simulations (based on experimentally calibrated CNFET and RRAM compact models for 10 nm technology nodes [2]) confirm that all of these circuits retain correct functionality despite an m-CNT (within any CNFET in the circuit). As an example, without SHARC, a single m-CNT results in a worst-case reduction of gain by 106 dB for a cascode amplifier, while with SHARC, a single m-CNT results in a worst-case reduction of gain of only 3 dB. The measured transfer characteristics of 2-stage op-amp is also shown indicating a gain of >800.

The key benefits of SHARC are: (1) the self-healing process (e.g., RRAM programming) is automatically performed by the m-CNTs themselves (as the m-CNTs provide the current path for resetting the RRAM during the self-healing process), (2) CNFETs and RRAMs are fabricated at low temperatures (<300 °C), and therefore can be fabricated directly vertically overlapping one-another in a monolithic three-dimensional fashion; this allows each RRAM cell to be fabricated either above or below the source or drain contact of the CNFET, minimizing area penalty associated with SHARC (the area overhead due to splitting the initial CNFET into *N* sub-CNFETs is ~(N-1)*S/W (subject to other considerations such as signal routing), where *S* is the minimum lithographic pitch and *W* is the original CNFET width), and (3) the circuit retains its healed configuration due to the non-volatility of the RRAM (though due to their programmability, the RRAM can potentially be reprogrammed repeatedly in the future).

Importantly, SHARC scales to larger-scale circuits since SHARC can be performed once for all CNFETs in parallel. An example of how RRAMs are formed and reset for the op-amp as well as the process flow is shown in Fig. 4. Bipolar RRAMs are initially fabricated on the first layer of the chip, and a sacrificial layer of metal can be defined to connect the RRAMs for the FORM operation. The sacrificial layer of metal is then removed, followed by the CNFETs fabrication (CNFET source/ drain contacts are patterned directly vertically overlapping the RRAM cells). The CNFETs are fabricated by depositing solution-based purified (~99.99% s-CNTs) CNTs on the wafer, followed by lithographically defining the source and drain regions. Another sacrificial layer of metal can be used to configure the circuit so all sub-CNFET gates are shorted, and all CNFET-RRAM pairs are connected in parallel. This enables all sub-CNFETs gates to be biased (e.g., $|V_{GS}|$ = 0V) simultaneously, as well as for all the RRAMs to undergo the self-healing RESET operation in parallel (importantly, all the CNFETs are PMOS as-fabricated, allowing a single gate bias to turn off all the CNFETs simultaneously). The sacrificial layer of metal can then be removed, and the CNFETs are doped to form the PMOS and NMOS followed by the final metal routing to complete the circuit (the NMOS CNFETs are formed through electrostatic doping, whereby a non-stoichiometric HfO_X dielectric is deposited over exposed CNTs in the channel of the FETs, converting PMOS CNFETS to NMOS CNFETs [3]). Importantly, all the experimental circuits fabricated and reported in this paper are complementary (using both PMOS and NMOS CNFETs).

As an experimental demonstration, we use SHARC to fabricate the following analog and mixed-signal circuits: 4-bit DAC (SHARC implemented in the 2-stage op-amp and switches), and a 4-bit SAR ADC (SHARC implemented in the strong-arm latch and switches). All circuits are fabricated with ~2 µm channel length CNFETs. Fig. 5 shows the schematic and die image of a 4-bit DAC. The measured transfer characteristics show monotonic behavior with non-linearity and gain error (due to large parasitic caps and routing resistance). Fig. 6 shows the SAR ADC schematic, its die image and its transfer characteristics. These circuits are the first mixed-signal CNFET circuits demonstrated, and the largest reported CMOS CNFET circuits to-date (see the comparison table in Fig. 4 [4], [5], [6]). Furthermore, since SHARC is implemented at the CNFET-level, it can be combined with additional existing circuit techniques to further improve performance, such as technology node scaling to improve energy efficiency and improved circuit topologies to improve linearity. Thus, SHARC offers a feasible path forward to enabling large-scale, energy-efficient analog and mixed signal circuits using CNFETs.

References:

[1] Jie Zhang, et al. "Probabilistic analysis and design of metallic-carbon-nanotube-tolerant digital logic circuits." IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems28.9 (2009): 1307-1320.

[2] Chi-Shuen Lee, H.-S. Philip Wong (2015). Stanford Virtual-Source Carbon Nanotube Field-Effect Transistors Model. nanoHUB. doi:10.4231/D3BK16Q68.

[3] Christian Lau, et al. "Tunable n-Type Doping of Carbon Nanotubes Through Engineered Atomic Layer Deposition HfOX Films." ACS nano (2018).

[4] Yingjun Yang, et al. "High-performance complementary transistors and medium-scale integrated circuits based on carbon nanotube thin films." ACS nano 11.4 (2017): 4124-4132.

[5] Michael L. Geier, et al. "Solution-processed carbon nanotube thin-film complementary static random-access memory." Nature nanotechnology 10.11 (2015): 944.

[6] Shu-Jen Han, et al. "High-speed logic integrated circuits with solution-processed self-assembled carbon nanotubes." Nature nanotechnology 12.9 (2017): 861.

[1] Analog circuits sensitivity to m-CNTs is caused by reducing the circuit output impedance and changing the DC operating point which drives the CNFETs out of saturation regime.

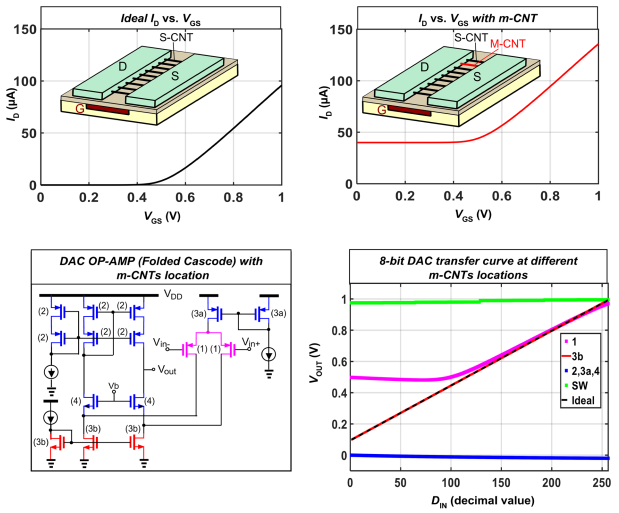


Figure 29.8.1: CNFET technology and impact of m-CNTs on device and circuit level. The 10-nm compact model used for simulations is calibrated to experimental data.

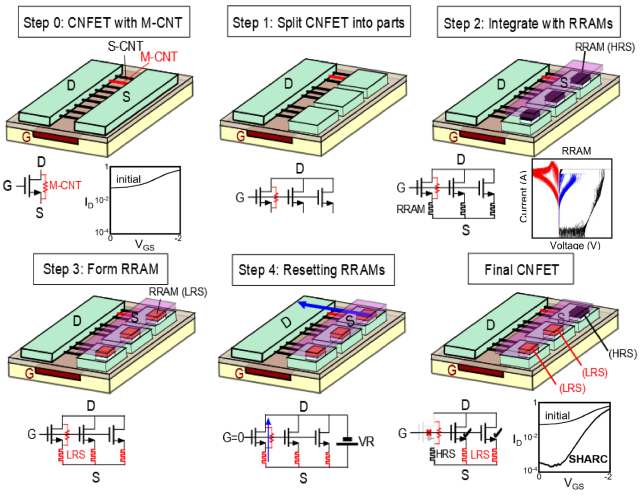
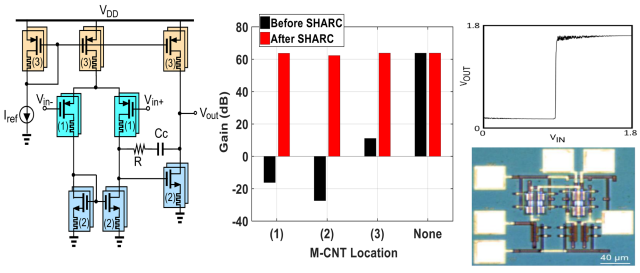


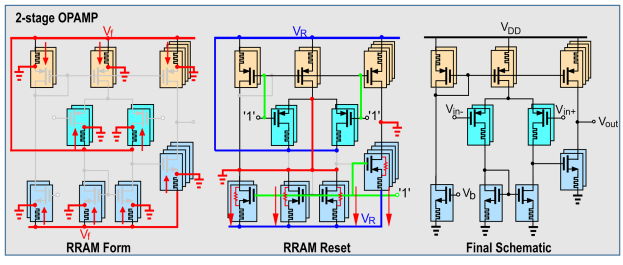
Figure 29.8.2: Principles of SHARC with programming steps.



M-CNT location	Common Source		Source Follower		Common Gate		Diff. Pair		Folded Cascode	
	Before	After	Before	After	Before	After	Before	After	Before	After
None (All S.C)	43 dB	42.89 dB	-50 mdB	-48 mdB	40.5 dB	40.7 dB	33.8 dB	35.4 dB	67.7 dB	67.7 dB
Input (1)	-3.4 dB	9.3 dB *	-108 dB	-75 mdB	2.13 dB	7.8 dB *	-130 dB	32 dB **	40.5 dB	66 dB
Active Load (2)	2.8 dB	40.8 dB	-1.2 dB	-57 mdB	2.76 dB	38.6 dB	-3.6 dB	34.4 dB	1.2 dB	65 dB
Current Source (3)	-		-		-		31.9 dB	35.4 dB	-38.8 dB	65 dB
									45.8 dB	67.5 dB
Cascode (4)	-		-		-		-		22.7	65.5

**Gain reduction due to high impedance node → can be solved in –ve FB systems, ** 3dB reduction as gm reduced to half*

Figure 29.8.3: M-CNTs impact on analog circuits before and after SHARC (experimental OPAMP transfer curve (top right) and simulations (table and bar chart)).



Circuit	CNFETs per Circuit	Reference
4b-ADC	**306**	**This Work**
4-bit full adder	132	Yang et al. [4]
Inverter	2	
SRAM Cell	6	Geier et al. [5]
Inverter	2	Han et al. [6]
Ring Oscillator	12	

Figure 29.8.4: SHARC implementation (top) with fabrication process flow (bottom), as well as comparison to state-of-the-art CNFET CMOS demonstrations.

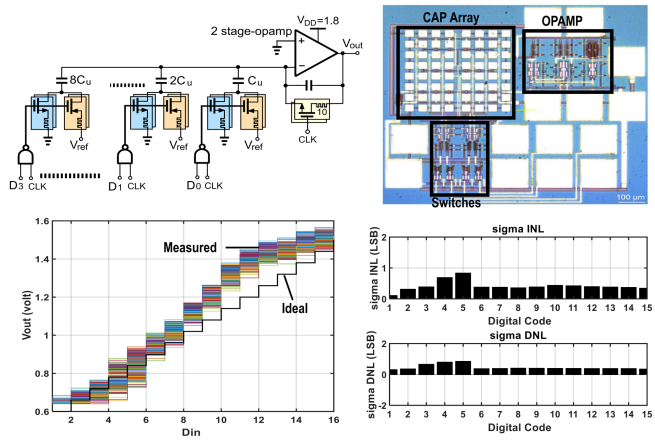


Figure 29.8.5: 4-Bit capacitive DAC with SHARC. Measured characteristics show the monotonic behavior with offset (50 mV) and non-linearity (σ_{INL}, σ_{DNL}< 0.8 LSB)

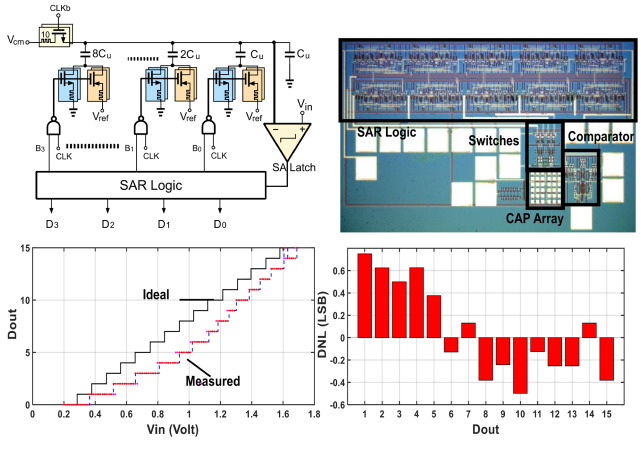


Figure 29.8.6: 4-Bit SAR DAC with SHARC. Measured characteristics show the ADC behavior with offset (35mV), gain error, and non-linearity (-0.5 LSB <DNL< 0.75 LSB).

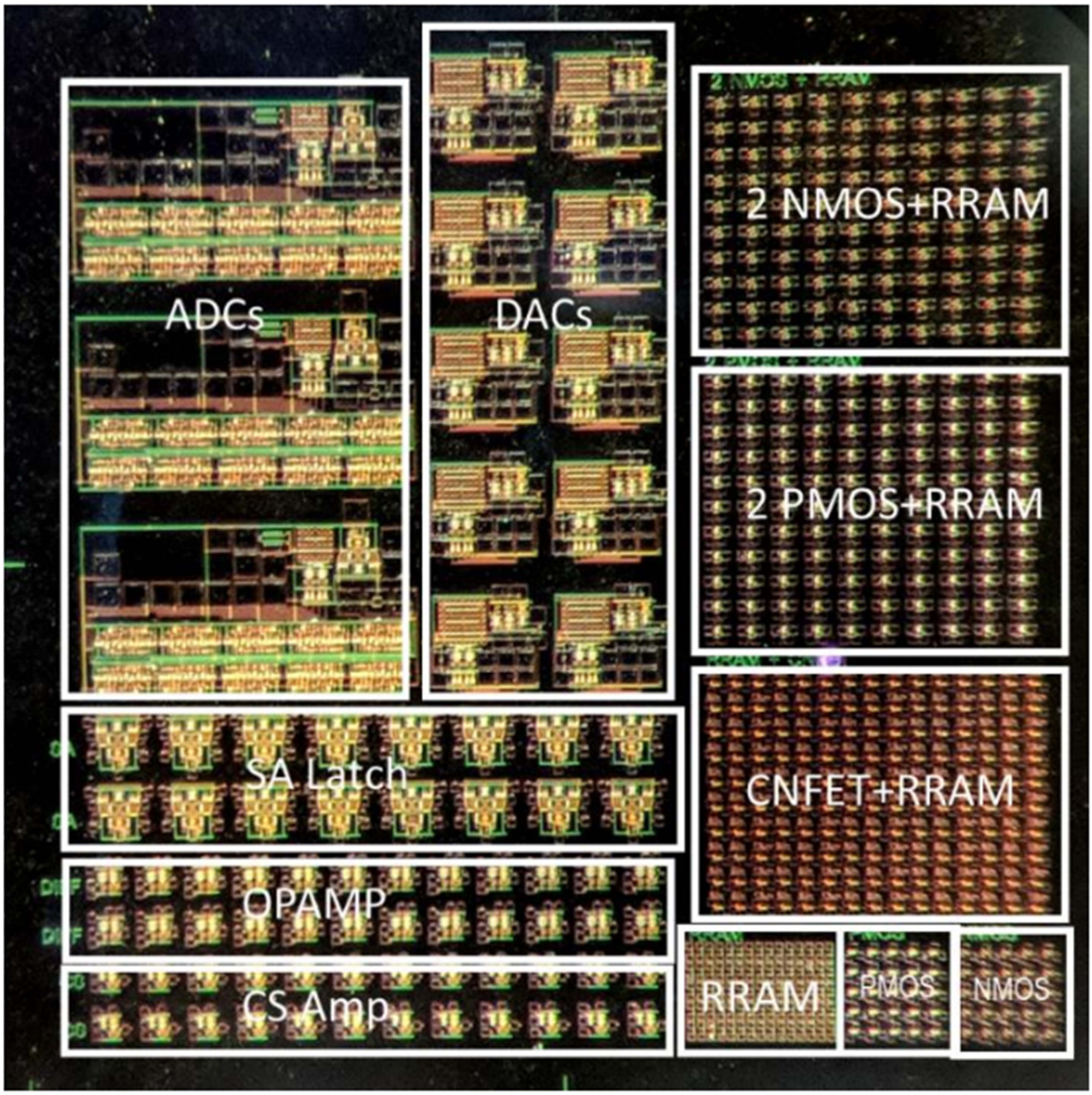


Figure 29.8.7: Die photo of the chip showing different circuits