High precision tuning of state for memristive devices by adaptable variation-tolerant algorithm

Fabien Alibart\textsuperscript{1}, Ligang Gao\textsuperscript{1}, Brian D Hoskins\textsuperscript{2} and Dmitri B Strukov\textsuperscript{1}

\textsuperscript{1} Department of Electrical and Computer Engineering, University of California Santa Barbara, Santa Barbara, CA 93106, USA
\textsuperscript{2} Materials Department, University of California Santa Barbara, Santa Barbara, CA 93106, USA

E-mail: alibart@ece.ucsb.edu and strukov@ece.ucsb.edu

Received 15 October 2011, in final form 22 December 2011
Published 20 January 2012
Online at stacks.iop.org/Nano/23/075201

Abstract

Using memristive properties common for titanium dioxide thin film devices, we designed a simple write algorithm to tune device conductance at a specific bias point to 1\% relative accuracy (which is roughly equivalent to seven-bit precision) within its dynamic range even in the presence of large variations in switching behavior. The high precision state is nonvolatile and the results are likely to be sustained for nanoscale memristive devices because of the inherent filamentary nature of the resistive switching. The proposed functionality of memristive devices is especially attractive for analog computing with low precision data. As one representative example we demonstrate hybrid circuitry consisting of an integrated circuit summing amplifier and two memristive devices to perform the analog multiply-and-add (dot-product) computation, which is a typical bottleneck operation in information processing.

1. Introduction

In order to fully realize the analog properties of resistive switching devices [Saw08, Was09, Per11, Kim11, Yan09, Lik08], which are also called ‘memristive devices’ [Chu11], one has to deal with significant variations in the switching behavior. In some applications such variations can be tolerated by the structure of the circuits, such as the case with some versions of artificial neural networks in which memristive devices are implementing synapses [Lik11, Sni07, Yu11b, Seo11, Cha11, Jo10]. However, for all other applications, e.g. multilevel memory [Bec00, Yu11a, Cho06], configurable filters [Dri10] and analog computing circuits (see, e.g., various theoretical proposals [Wei11, Shi11, Per10, Lai10, Pro10] and the extensive review in [Per11]), the circuit performance directly depends on how accurately the resistive state can be set.

A natural way to tackle variations in switching behavior is to utilize an active feedback scheme, e.g. applying iterative write and read (test) pulses to converge to a certain desired conductive state of the device. Such a scheme has been successfully applied to phase change memories to achieve multilevel memory operation [Pap11, Bed09]. A similar idea to use closed-loop circuitry has been proposed and theoretically simulated using a model for TiO\textsubscript{2} devices [Yi11]. In this paper, we experimentally demonstrate a simple feedback algorithm which takes into account specific memristive behavior in titanium dioxide devices to tune the resistance state of the device to within 1\% relative accuracy of the dynamic range. We then use our algorithm to demonstrate one of the most important operations in information processing—analog multiply-and-add computation (MAC).

2. Experimental results

Figure 1 shows typical resistive switching behavior of titanium dioxide devices. The full switching loop is obtained by a quasi-DC triangular voltage sweep between 0 and −1.5 V (RESET transition) followed by a quasi-DC triangular current sweep between 0 and 750 \textmu\text{A} (SET transition) with a sweeping period of about 4 s. In order to achieve analog operation (or multistate memory), the RESET and
SET transitions are implemented gradually. The incremental RESET switching is obtained by applying several voltage sweeps of increasing amplitude from $-0.8$ to $-1.5$ V with a step of 10 mV between each sweep. Similarly, an incremental SET transition is obtained with current sweeps of increasing amplitude from 100 to 750 $\mu$A with a step of 50 $\mu$A.

Instead of using sweeps of voltage and current, a more accurate analog control of the device is possible using a sequence of relatively large amplitude write pulses followed by smaller non-disturbing read pulses [Pic09]. The advantage of pulse measurement is to decorrelate the effect of time and voltage to provide a better insight into the switching dynamics. For example, varying the pulse duration with constant amplitude ‘freezes’ the effect of applied voltage and highlights the effect of stress duration. In this paper, we report different measurements where both voltage amplitude and pulse duration effects have been investigated.

In particular, each measurement is implemented by applying a sequence of two different pulses: (i) the read pulse of $-200$ mV and 1 ms width to probe the state of the device (which is represented by the resistance or current) and (ii) the write pulse with variable pulse duration and amplitude to change the state of the device. These two pulses (read and write) are alternated at a frequency of 0.5 Hz to prevent the accumulative Joule heating effect. Note that for all experiments described below we use voltage-controlled pulses for SET switching because current-controlled switching may not be readily compatible with large scale crossbar circuits.

Figure 2 shows the evolution of the state as a function of both the write pulse amplitude and cumulative time (i.e. summation of the write pulse durations). Before the pulse measurements, the devices are set to a low (high) resistance state with a quasi-DC current (voltage) sweep in order to start the SET (RESET) transition with the same initial condition. In particular, in figure 2(a) each curve shows the dynamic evolution of the device’s state, as a result of the application of fixed amplitude pulses with an exponentially increasing duration from 200 ns to 1 ms. By exponentially increasing the pulse duration, we have studied switching dynamics over large time intervals—$\sim 10^5$ order of magnitude.

Figures 2(b) and (c) show in detail the dynamic evolution of the device’s state using 200 ns long pulses for RESET and SET transitions, respectively. Note that in this case we use a slightly different stress protocol (from that used to obtain figure 2(a)) in which both the amplitude and duration of pulses are fixed during each measurement. In general, both stress protocols can be used to study switching dynamics as a function of cumulative time of the applied pulses and their amplitude; however, there are several important distinctions. The most significant one is that running experiments across a large (cumulative) interval of time is challenging using fixed duration pulses because there is large delay in the parameter analyzer for every new pulse generated. Therefore, to study the dynamics over the large time interval (figure 2(a)) we used variable duration pulse stress. On the other hand, we use fixed duration pulse stress for our further experiments because of much better control over switching dynamics. Additionally, the undesirable slow transient heating of the substrate is likely to be less severe for the latter protocol.

Figures 2(a)–(c) clearly demonstrate that the device can be switched fast—of the order of 1 $\mu$s—for both transitions for the largest applied voltages and retain its state for at least 1 ms if the applied voltage is within the $-0.6 \, V < v < 0.5 \, V$ range. Moreover, figure 2 highlights that the switching dynamics is exponential with voltage for SET switching and roughly follows a power low for the RESET (at least before it saturates). This is similar to previously reported dynamics [Pic09]—though it should be noted that in our case the state is defined as a conductance (or resistance) at specific bias rather than a phenomenological parameter such as barrier width. The experimental results support that the retention (e.g. at DC bias $-0.2 \, V$) to write time ratio (at maximum amplitude) is larger than $10^5$. This ratio is probably much higher than that number. First of all, this is because the stress has not been applied for long enough time to see significant drift of the state so that the retention time is certainly longer. Second, the exponential nature of the switching dynamics is likely due to thermal effects.
Figure 2. Switching dynamics characterized by voltage pulse stress with variable amplitude and duration. Panel (a) shows general switching behavior and volatility characteristics over a large time interval. Panels (b) and (c) show the switching in detail (over shorter and more relevant for our case time intervals) for SET and RESET transitions, correspondingly, as well as corresponding phenomenological fitting with dashed black lines—see the appendix for more details on fitting. The top figure for each panel shows schematically a pulse sequence used for this particular measurement.

[ Stra11, Bor09, Shk09, Iel11, Str11] so that the switching time could be potentially much faster if larger write voltages are applied [ Iel11, Pic09, Str09]. We avoid checking this hypothesis because the experimental set-up does not allow applying short pulses, while relatively long high amplitude pulses lead to over-stressing and damaging of the device.

While the general dynamics for all our devices after forming is qualitatively similar there are significant variations—both from device to device and for the same device upon cycling—which prevents the use of a calculated stress voltage to drive the device to the desired state. In principle, such variations can be coped with by over-stressing the device when only two extreme resistive states in the device are needed. To tune the device to a specific intermediate state the over-stressing strategy is clearly not an option and instead we use the feedback algorithm.

Our algorithm is based on the fact that large amplitude pulses can be used to reach a desired state faster but also at much cruder precision (figures 2(a)–(c)) due to the fixed minimum pulse duration (which is limited by the experimental set-up). Once the device is driven close to the vicinity of the desired state smaller amplitude pulses can be used for fine tuning. (Note that using pulses of relatively small amplitude only is not an attractive option because it may require an exponentially long write time to reach the desired state.) To take advantage of such switching dynamics our algorithm (figure 3) is based on a sequence of increasing amplitude voltage ramps of appropriate polarity (which depends on the initial and the desired state of the devices). In particular, we apply 10 μs long voltage pulses with increasing amplitude starting from 0.5 V and −0.5 V for the SET and RESET sequences, respectively, and a step of 5 mV. Similar to previous measurements the device state is checked with the read pulse after each write pulse and the voltage ramp is applied until it reaches and overshoots the desired state. At this point the new voltage ramp of opposite polarity begins. (Note that the new ramp sequence is always started from the same initial voltage, i.e. 0.5 V for SET and −0.5 V for RESET. Such small initial values ensure starting the sequence with a non-disturbing pulse amplitude). Because this time the initial
state is closer to the desired one the maximum amplitude of the voltage pulse in the new ramp (i.e. for the last pulse before the ramp is stopped) is smaller, which in turn ensures that this ramp will drive the device to the desired state with even better precision. Using this algorithm we were able to tune the device with 1% accuracy with respect to the desired state within the dynamic range of the device (figure 3). In principle, even more accurate tuning should be possible by using more ramps (i.e. more time) though in this particular experiment we stopped our algorithm when 1% accuracy was achieved.

Theoretically, i.e. from physical noise analysis, any computation at signal precision below eight-bit (for robotics, distributed sensor network applications, etc) could be done much more efficiently in analog or mixed signal circuits [Str07] with wires carrying multiple bits of information [Sar98]. However, even low precision analog processing in conventional CMOS technology is cumbersome, due to the lack of suitable hardware, e.g. to implement multilevel (analog) weights. On the other hand, low precision analog information processing can be implemented with hybrid circuits, which combine a single conventional CMOS chip and layer(s) of memristive devices [Lik08]. Figure 4 illustrates such a circuit consisting of an integrated circuit (IC) summing amplifier and two memristive devices programmed to a high precision state with the proposed algorithm. To realize MAC circuits, the memristive devices implement density-critical configurable weights, while conventional technology is used for the summing amplifier, which provides gain and signal restoration. As a result, individual voltages applied to memristive devices are multiplied by the unique weight (conductance) of the memristor and summed up by the IC amplifier—all in analog fashion.

3. Discussion and summary

In this paper we have demonstrated a write algorithm to tune device conductance at a specific bias point to 1% relative accuracy within its dynamic range even in the presence of large variations in switching behavior. We then use this algorithm to implement one of the most critical operations in information processing—the multiply-and-accumulate operation, using hybrid circuits consisting of an IC operational amplifier and memristive devices.

While the demonstrated accuracy should be sufficient for most of the analog operations [Sar98], the tuning time and memristive device density must be improved for the proposed MAC circuits to be practical. In this context we note that, though the results demonstrated in figures 1–4 are for microscale devices, there is strong evidence that it can be sustained in nanoscale devices based on previously reported experimental scaling analysis [Lee11] and the fact that the forming process produces nanoscale filaments with an active area localized to spots of a few nanometers [Stra10, Kwo10].

Tight integration with CMOS circuitry and further improvements of the memristive devices (e.g. by introducing artificial filaments into device structures [Ali11]) might reduce the algorithm time dramatically taking into account that write speed for metal oxide memristive devices could be as small as a few hundred picoseconds [Tor11]. The fact that the nanoscale memristive devices have typically less variations [Jo08] also helps because better tuning precision could be expected within a shorter amount of time using our algorithm. Moreover, one can envision having dedicated on-chip configuration circuitry implementing the high precision write algorithm and allowing for simultaneous tuning of large numbers of memristive devices in parallel (similarly to the scheme described for digital CMOL FPGA circuits [Str07, Lik08]) so that the configuration process would be shortened even further. The investigation of such design options (e.g. overheads of reconfiguration circuitry and relevant tradeoffs) will be a focus of future work.

It is also worth mentioning that in the passive crossbar structures semi-selection and leakage currents might present additional challenges. However, at least the former issue should not be a problem for our devices because of the strong nonlinearity in the switching dynamics (figure 2). Also, while there is a natural tradeoff between the tuning time and resulting accuracy one can expect a certain physical limit to
the precision, e.g. defined by the switching noise and temporal
instabilities [Lee09, Son10], which is certainly worthwhile
investigating.

4. Device fabrication and electrical characterization

The Pt/TiO2(30 nm)/Pt devices have been implemented in the
‘bone-structure’ geometry with an overlap area between two
electrodes of about 1 μm² (inset of figure 1). An evaporated
Ti/Pt bottom electrode (5 nm/25 nm) has been patterned by
the conventional optical lithography technique on an Si/SiO2
substrate (500 μm/200 nm, respectively). Then, a 30 nm TiO2
switching layer has been realized by atomic layer deposition at
200 °C using titanium isopropoxide (C12H28O4Ti) and water
as a precursor and reactant, respectively. A Pt/Au electrode
(15 nm/25 nm) was evaporated on top of the TiO2 blanket
layer. A rapid annealing of the device was finally applied
at 500 °C under N2 and N2 + O2 atmosphere for 5 min to
improve the crystallinity of the TiO2 material. Note that Ti
and Au metal layers were added to the device stack to provide
adhesion between the substrate and bottom electrode and to
improve switching properties of the devices [Yan10], and to
provide structural stability (i.e. to release strain) of the top
electrode, respectively.

Electrical measurements were carried out with a B1500
Agilent parameter analyzer in ambient conditions. We used
the conventional two-probe technique for the electrical
characterization of the device because the electrode resistance
is much smaller (−of the order of 80 Ω) compared to that
of the ON state of the device. Each device was formed by a
negative voltage sweep between 0 and −10 V (the forming
voltage was around −8 V) and a current compliance of
300 μA ensured by a transistor connected in series to the TiO2
device. After this forming process, the compliance transistor
was removed. A few sweeps between −1.5 and 1.5 V were
applied before obtaining a stable bipolar behavior of the
device. The extended waveform capability of the tool (B1530)
was used for the pulse measurement presented in this paper
(sampling rate of 10 ns and rise/fall time between 0 and 5 V
of 80 ns).

The demo circuit was implemented with a TL072
operational amplifier IC chip. During the programming,
the memristive devices were disconnected from the circuit
and programmed using the high precision algorithm set-up
(B1530 waveform generator and measurement unit). After
that the memristive devices were connected externally to the
IC chip and the output was measured with an oscilloscope
(Agilent 3000 series) while the input pulses are generated
with a waveform generator (Agilent 33520). For the circuit
experiment (figure 4) the memristive devices in the MAC
circuit were programmed to 15, 30 and 60 μA states (at a read
current of −200 mV).

Acknowledgments

We thank J Carruthers, D Hammerstrom, K Likharev, A Ramu
and C Teuscher for helpful discussions. This work was
supported via NSF grant CCF-1028336 and equipment
acquired through NSF grant DMR-1121053.

Appendix. Fitting of SET and RESET transitions

In this paper we study the phenomenological behavior and
trends of both transitions in order to implement an effective
algorithm using the two transitions in an analog way. The
physical interpretation of the fitting is beyond the scope of
this paper and will necessitate more detailed analysis and
measurements. The current during the RESET transition is
fitted with a power law:

\[
I_R = A \cdot V^B
\]

\[
I_S = A \cdot V^B
\]

Figure A.1. Fitting parameters of the (a) RESET and ((b), (c)) the
SET transitions.
\[ I = A_0 e^{-\beta v}. \]  
(A.1)
where \( \beta \) and \( A_0 \) are functions of voltage \( v \). Figure 2(b) shows the good agreement between measurement and fitting and figure A.1(a) presents the dependence of \( A_0 \) and \( \beta \) on voltage. The current during the SET transition has been fitted with a sigmoidal function:

\[ I = A_0 + \frac{A_1 \ln (Bt + 1)}{1 + e^{-\frac{t}{\tau_1} (t - \tau_2)}}. \]  
(A.2)
where \( A_0 \) and \( A_1 \) are the same constants (9 \( \times \) 10\(^{-6} \) A and 7.84 \( \times \) 10\(^{-5} \) A, respectively) for all the SET transitions, while \( B, \tau_1 \) and \( \tau_2 \) are parameters which depend on the applied voltage \( v \) (figure A.1). Figure A.1(c) shows the result of the fitting, highlighting the fact that the SET transition is well described by an exponential dependence with time (at least before the pseudo-saturation that follows a logarithmic dependence with respect to time). The fitting also emphasizes the exponential dependence of \( \tau_1 \) (equivalent to the time for switching) and \( \tau_2 \) (equivalent to the time for starting switching) with voltage that is consistent with the extrapolation of its switching time for TiO\(_2\) devices [Pic09].

References


[Bec09] Beck A, Bednorz J G, Gerber Ch, Rossel C and Waser R and Kohlstedt H 2010 Probing Cu doped Ge\(_{0.95}\)Se\(_{0.05}\) and its percolating nature Applied Phys. Lett. 95 122112


[Iel11] Ielmini D, Nardi F and Cagli C 2011 Physical models of size-dependent nano filament formation and rupture in NiO resistive switching memories Nanotechnology 22 254022


[Seo11] Seo K et al 2011 Analog memory and spike-timing-dependent plasticity characteristics of a nanoscale titanium oxide bilayer resistive switching device Nanotechnology 22 254023


[Sni07] Snider G S 2007 Self-organized computation with unreliable, memristive nanodevices Nanotechnology 18 365202


