Energy Efficient Adiabatic FRAM with 0.99 pJ/bit Write for IoT Applications

Supreet Jeloka^{1, 2}, Zhehong Wang¹, Ruochen Xie³, Sudhanshu Khanna⁴, Steven Bartling⁴, Dennis Sylvester¹, and David Blaauw¹

¹University of Michigan, Ann Arbor, MI; ²ARM, Austin, TX; ³AMD, Boxborough, MA; ⁴Texas Instruments, Dallas, TX

Abstract

Low access energy embedded non-volatile memory is critical for low power sensing systems. This paper proposes a charge-recycling FRAM design that uses resonance between the FRAM array capacitance and an off-chip inductor to perform both write and read functions. In 130 nm, the 256x80 FRAM increases energy efficiency by 3× compared to standard operation, achieves 0.99pJ/bit write energy, 0.4pJ/bit read energy and 102Mbps at 1V.

Introduction

Battery-operated sensing systems have very low activity rates and power down their memories during inactive periods to save power, storing data in non-volatile memory (NVM). However, conventional NVMs such as flash use high voltages and long duration write operations, causing high energy per bit for writing (>100 pJ/b). Unlike flash, Ferroelectric RAM (FRAM) is significantly faster and uses nominal voltages, reducing write energy [1] and making it appealing for IoT applications.

This paper proposes a charge-recycling (i.e., adiabatic) FRAM design that further reduces the read and write energy. It leverages the observation that FRAM energy dissipation mainly arises from charging/discharging the large capacitance of the ferroelectric element (~10 s of fF), which can be recycled using a resonant clock. This approach increases energy efficiency by 3× compared to non-adiabatic operation, and the proposed 130-nm 256x80 FRAM memory achieves write energy of 0.99 pJ/bit.

The FRAM bit cell is structured similar to a DRAM bit cell (Fig. 1) except it uses a non-linear Ferroelectric capacitor (FeCap). When (dis)charged, the FeCap reaches a saturation charge level (-)Qs at voltage (-)Vdd, and when the voltage across the FeCap returns to 0 V, remnant charge (-)Or remains; the polarity of Or determines whether the bit cell state is a '1' or '0'. Fig. 1 shows the conventional FRAM write operation for writing a '1'; here the Bit-Line (BL) must be higher than the Plate-Line (PL). In this conventional design, for every bit, we must expend at least CV² energy to pull both the FeCap and the bit-line (BL) capacitance high. This limits conventional FRAM write energy efficiency, even when peripheral logic power is minimized [2]. Furthermore, most of this energy is wasted since the capacitance change is small, and only a fraction of it is used for the FeCap polarity switch. A second concern is that many IoT systems reduce core voltage to save energy, which places robustness constraints on memories and also increases access time. To address this issue, we use a 2T-2C bit cell [3][4] with differential read, which increases read margin and read speed but adversely affects access energy.

We address both of these issues synergistically by 1) using resonance between the FRAM array capacitance (FeCap + BL Capacitance) and an off-chip inductor to create a sinusoidal clock that is used to write and read the FRAM. This approach recycles energy from one cycle to the next and recaptures ~66% of the energy lost in conventional operation. 2) As the frequency is locked by resonance, the approach achieves better scaling of write speed at low voltages, increasing throughput and energy efficiency. The proposed FRAM is the first to apply the concept of resonating adiabatic design, used typically for logic, to NVMs.

Proposed adiabatic FRAM

Fig. 2 shows a simplified adiabatic FRAM array. The inductor has one input tied to Vdd/2 and resonates with the on-chip capacitive loading of PL. The transmission gates controlled by PL-enable (PLEN) and write-enable (WREN) connect the BL to PL

or write data, respectively. The pull-up (PU) and pull-down (PD) transistors restore the energy expended for write or read and maintain the amplitude of the resonating waveform.

Fig. 2 shows the write operation waveforms: At the peak (trough) of the resonating sine wave a PU (PD) pulse is generated. The PU and PD pulses are used to switch the WREN and PLEN transmission gate mux controls. While PL continuously oscillates, the BLs are transitioned adiabatically by connecting them to the resonating PL line and having them transition with the PL line. Since there is minimal voltage across the PLEN mux, the transition is "adiabatic," and energy is recycled. When the BLs reach their intended value (Vdd or 0), they are clamped and remain static. During a write '0', the BL is clamped to 0, while PL cycles from low–high–low. Conversely, for '1', BL is clamped high, while PL cycles from high–low–high (Fig. 2). Each row write requires 1.5 cycles of PL resonance, i.e., $f_{access} = 2/3 * f_{resonance}$.

Fig. 3 shows the overall block diagram of the proposed FRAM design. A decoder is followed by Word Line (WL) drivers that operate at $(Vdd+V_t)$ to avoid the V_t drop across the access transistor of the selected bit cell. The write and read FSMs use PU and PD pulses as a 2-phased clock input, to generate the control signals for BL write and sense amplifiers, respectively. Hence, PU and PD pulse generation is critical to FSM timing. We use a continuous comparator with a reference near Vdd (Gnd) for PU (PD) generation (Fig. 3). As the comparator output pulse is wide, we use a pulse shaping and delay block to narrow and center the pulse at the peak of PL, saving power.

Fig. 4 shows the proposed FRAM read. In conventional read, we pre-discharge the BL, enable WL, and then let the BL float while we apply Vdd at PL. The stored state (0 or 1) of the bit cell determines the remnant charge stored in the FeCap, which leads to different voltages developing on the BL. For the 1T-1C bit-cell, the sense amplifier (SA) compares BL against a reference voltage. However, for improved robustness at low Vmin we use a 2T-2C bit cell that stores complementary data. The SA can be configured to support single-ended read of a 1T-1C cell for higher density at nominal Vdd or differential read of 2T-2C for low Vmin. Read also uses charge recycling for energy efficiency.

Measured Results

Fig. 5 provides measured results for the 130-nm test chip. The non-adiabatic baseline FRAM numbers are measured using a function generator at the PL input instead of the on-PCB resonating inductor. The board picture shows the small 22- μ H resonating inductor placed next to the chip. The figure also shows the 3.8 MHz resonance waveform captured on an oscilloscope. The measured write energy at 1 V is 0.99 pJ/bit, while the read energy is 0.4 pJ/bit at 1 V. Compared to baseline, the adiabatic technique reduces write energy by 3× and read energy by 2.5×, indicating that charge recycling is ~60–66% efficient.

The proposed FRAM shows >1.65× reduction in energy compared with the non-adiabatic baseline and other FRAM designs, even with the more robust 2T-2C bit cell (Fig. 6). It achieves the lowest read/write energy among those listed and maintains high throughput at low operating voltage, making it an attractive solution for ultra-low energy for sensor systems.

References

- [1] M. Qazi et al., JSSC, 2012 [2] H.P. McAdams et al., JSSC, 2004
- [3] A. Sheikholeslami et al., Proceedings of the IEEE, 2000
- [4] S.A. Qidwai et al., IMW, 2012 [5] Y.M. Kang et al., VLSI, 2006

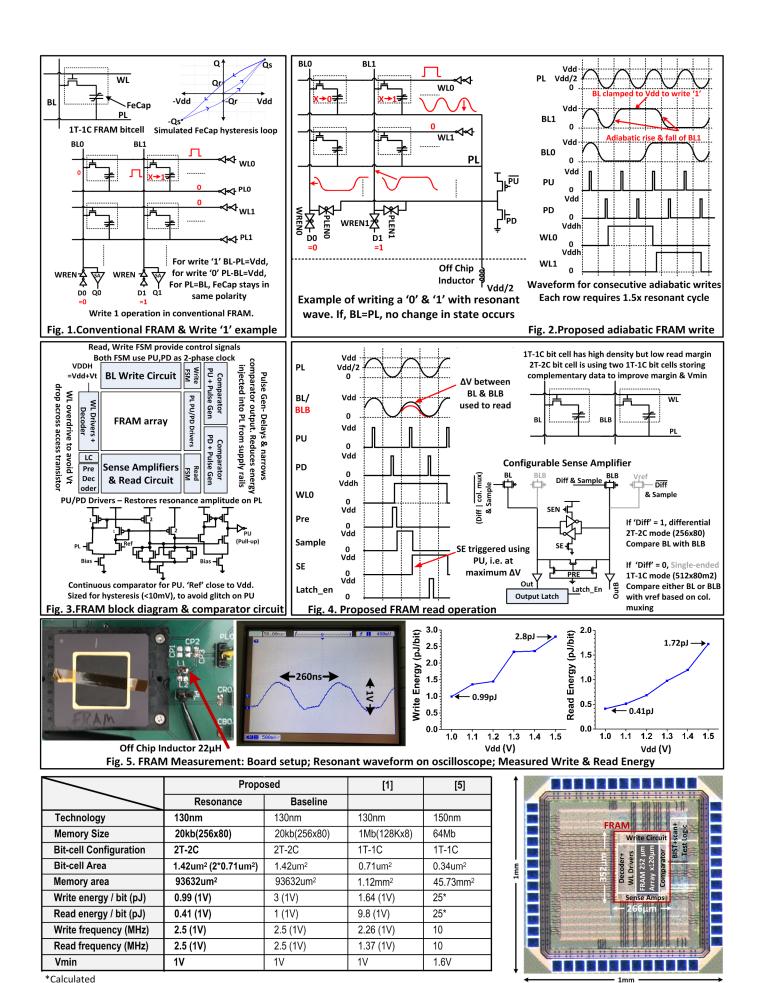


Fig. 6. Comparison table with previous works. Die picture of proposed FRAM