Dynamic Single and Dual Rail Spin Transfer Torque Look Up Tables with Enhanced Robustness under CMOS and MTJ Process Variations

Aliyar Attaran[1], Hassan Salmani[2], Houman Homayoun[3], and Hamid Mahmoodi[1]
[1]School of Engineering, San Francisco State University, San Francisco, CA
[2]Electrical and Commuter Engineering, Howard University, Washington, DC, USA
[3]Electrical and Computer Engineering, George Mason University, Fairfax, VA, USA
[1]aattaran@mail.sfsu.edu, [2]hassan.salmani@howard.edu, [3]hhomayou@gmu.edu, [1]mahmoodi@sfsu.edu

***Abstract*—In this paper, we investigate the limitation of existing STT-LUT designs and propose two new circuit styles of designing STT-LUTs that offer higher performance and robustness compared to the conventional STT-LUT design. The proposed styles include a Dynamic Single Rail (DSR) and a Dynamic Dual Rail (DDR) STT-LUT. The simulation results in a 16nm bulk CMOS technology shows that the proposed designs exhibits up to 3.3X read delay reduction, 2.4X active power reduction, and 441X sensing failure rate reduction compared to the best conventional STT-LUT design. The proposed DDR scheme offers the best overall performance even when considering the state of the art Separated Precharge Sensing Amplifier and Separated Decoding schemes.**

Keywords*—*Look up table; low power; magnetic tunnel junction; reconfigurable logic; spin transfer torque.

1. INTRODUCTION

The primary application of STT has been in memory application (STTRAM). STTRAM offers a CMOS compatible and non-volatile memory and hence a zero leakage alternative to SRAM. However, the main challenge remains to be the relatively high current and delay needed for the write operation [1]. STTRAM is particularly suited for realizing non-volatile and CMOS compatible programmable logic such as on-chip FPGA in which the write operation is infrequent. The basic programmable logic element is a look-up table that utilizes Magnetic Tunnel Junctions (MTJ) for storage (STT-LUT) [2] [3].

In this paper, we propose two new designs of STT-LUT, called Dynamic Single Rail (DSR) and Dynamic Dual Rail (DDR), which offer improved speed, power, and robustness compared to the existing designs. The contributions of the paper are as follows:

- Proposing two new STT-LUT designs based on dynamic voltage mode operation
- Improving the speed, power, and robustness of STT-LUT using the proposed designs
- Performing robustness comparisons by combined analysis of impact of CMOS and MTJ variations on read sensing failure rates of STT-LUTs
- Comparing the best proposed STT-LUT design against the state of the art Separated Precharge Sensing Amplifier (SPCSA) and Separated Decoding schemes.

The dominant existing STT-LUT design is the Dynamic Current Mode (DCM) logic based design [4].

The remainder of the paper is organized as follows. Section 2 presents the proposed DSR STT-LUT. Section 3 presents robustness analysis of the STT-LUTs under process variations. Section 4 presents the proposed DDR STT-LUT and its comparisons with DSR. Section 5 compares the proposed DDR and existing SPCSA and separated decoding schemes. Finally, conclusions are drawn in Section 6.

This research is sponsored by the Defense Advanced Research Projects Agency (DARPA-AirForce, FA8650-15-C-7569) of the USA. EDA tools are donated by Synopsys Inc.

2. PROPOSED STT-LUT DESIGN: DSR

Figure 1 shows the first proposed Dynamic Single Rail (DSR) STT-LUT that operates on a voltage rather than current mode logic. The design connects the common end nodes of all MTJs to ground in the read mode via a read enable (REN) transistor shared among multiple STT-LUTs. This transistor can also be used as a sleep transistor for power gating in the standby mode. Direct connection of this node to ground offers a voltage mode operation on the selection tree allowing for full voltage swing operation. Hence, the sense amplifier can be added directly on top of the selection tree as shown. This design is dynamic and uses a clock signal. When clock signal is low, the design is in the precharge phase: the sense amplifier is precharged to VDD and the outputs (Z and Z') are discharged to ground. Note that assuming inputs (A and B) and their differentials (A' and B') are coming from a similar LUT, these will be both at zero level and hence turning off the selection trees and therefore elimination any possible short-circuit path in the precharge phase. When the clock switches to high, the design is in the evaluation phase and as soon as all the inputs have been set by the previous stages, then a unique MTJ and the reference resistor get selected creating conductive paths to the ground. The resistance difference between the selected MTJ and the reference resistor will set the direction of the sense amplifier switching and hence the outputs.

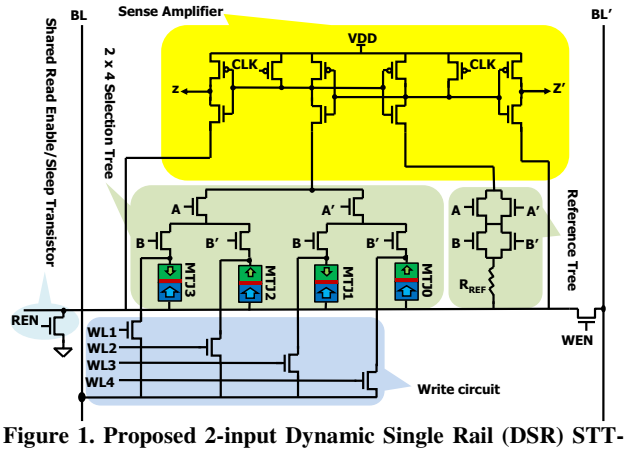

Figure 1. Proposed 2-input Dynamic Single Rail (DSR) STT-LUT

In the write mode, the read enable (REN) transistor is turned off and the write enable (WEN) transistor is turned on connecting the common node of the MTJs to one of the bitlines (BL'). The other bitline (BL) gets connected to a unique MTJ via a uniquely set wordline (WL) signal. The data present on the differential bitlines gets written to the selected MTJ.

3. ROBUSTNESS ANALYSIS

3.1 Simulation Setup

The designs are simulated in a predictive 16nm CMOS process [5] with nominal supply voltage of 0.7V. In our first simulation attempt, we keep all transistors minimum sized for both designs (NMOS: W/L=30nm/16nm and PMOS: W/L=60nm/16nm). Since the critical operation in programmable logic applications, where these LUTs are used, is the read mode, we simulated these designs in the read mode, where MTJs can be modeled as resistors. The MTJ resistance depends on the MTJ physical geometries and specifically the thickness of the insulator and the 2D cross-section area of the MTJ, and the voltage bias [6]. The parallel state resistance (R_P) can be modeled as [6]:

$$R_P = \frac{t_{ox}}{F\sqrt{\varphi}A} e^{\left(1.025 t_{ox}\sqrt{\varphi}\right)} \quad (1)$$

where, F and φ are the tunneling conductivity and potential barrier height of the insulator, respectively. Considering a circular shape 2D area of radius r, the area A is expressed as $A=\pi r^2$. The MTJ parallel resistance is estimated for the 16nm process. The resistance difference between the two MTJ states is quantified by the Tunneling Magneto Resistance (TMR). Using crystalline MgO as the insulator, TMR up to 6 is reported [6]. In this study, we choose a TMR value of 4.

3.2 MTJ Variations

The resistance of the MTJ is dependent on bias as well as MTJ geometries. Given the MTJ voltage bias is fixed in the read mode, we concentrate on the influence of MTJ geometrical variations. The two geometries of most influence are the insulator thickness (t_{ox}) and the cross-section area ($A=\pi r^2$) as expressed by Equation (1). Assume the t_{ox} and r exhibit variations represented by dt and dr from their nominal values, respectively.

The read failure occurs when the R_P is increased or R_{AP} is decreased. An accurate estimate of reliability against MTJ process variations can be obtained by applying statistical variations to t_{ox} and r of the MTJ and measuring the failure rate of the STT_LUT read operation.

3.3 Transistor Variations

The primary source of variations in bulk CMOS transistors is random threshold voltage (Vt) variations caused by Random Dopant Fluctuations (RDF) [7]. The RDF induced Vt variation from the nominal Vt (Vt_0) ($dvt=Vt-Vt_0$) for a transistor of size L (channel length) and W (channel width), follows a normal distribution with a standard deviation inversely proportional to the square root of the transistor area as follows [7]:

$$\sigma_{dVt} = \left[\frac{qT_{ox}}{\varepsilon_{ox}}\sqrt{\frac{\left(N_a W_d\right)}{3L_{min}W_{min}}}\right]\times\sqrt{\frac{L_{min}W_{min}}{LW}} = \sigma_{dVt0}\times\sqrt{\frac{L_{min}W_{min}}{LW}} \quad (2)$$

where all the process parameters are lumped into σ_{dvt0} representing the standard deviation of the threshold voltage variation of a minimum sized transistor with channel length and width of L_{min} and W_{min}. Increasing the transistor area provides a means to reduce the influence of variations at the cost of increase in area and power. To quantify the STT robustness against Vt variations, the read sensing failure rate can be measured by applying statistical variations to the threshold voltages of transistors in the STT-LUT.

3.4 Robustness Comparisons

There are two possible failures in the read mode: read sensing failure and read disturbance failure. The read sensing failure is related to the incorrect sensing by the sense amplifier due to the presence of mismatch caused by CMOS and MTJ variations resulting in low sensing margin. The read disturbance failure occurs in a read operation when the read current passing through the MTJ exceeds the critical write current resulting in the flipping of the state of the MTJ (i.e. write operation). In STT-LUTs the read sensing failure is the dominant type because of the stack of transistors in the read path (i.e. in the multiplexer tree) limits the read current. Hence, we focus on the read sensing failures. Random statistical variations are applied to MTJ parameters (t_{ox} and r) and threshold voltage (dvt) and read sensing failure rates of the STT-LUTs are measured. The standard deviations of t_{ox} and r variations are set at 10% of their nominal values. In the 16nm CMOS model used, the NMOS threshold voltage of a minimum sized transistor with the short channel effects applied was found to be 233 mV. The standard variation of Vt variation was set at 30mV which is 13% of the nominal value. The Monte Carlo simulation is setup to calculate the failed cases due to destructive process variations on the design. All the MTJ cells are read from the STT-LUT, and read failure is defined if a wrong value is read from any cells or if the delay of reading the correct values of a cell exceeds 500pS (half the evaluation cycle time). If reading from any of the MTJ cells fail, the entire STT-LUT is considered faulty.

The delay distributions and failure rates under intra-die variations of these process parameters for the 2-input STT-LUTs.

4. PROPOSED STT-LUT DESIGN: DDR

The proposed DSR style reduces the failure rates significantly compared to the conventional DCM STT-LUT style; however, the failure rates are still fairly high. We propose another new STT-LUT design style that further enhances the robustness by utilizing dual differential MTJs per bit and eliminating the reference resistor. The schematic of the proposed Dynamic Dual Rail (DDR) style for a 2-input STT-LUT is shown in Fig. 2. The inputs select two MTJs that are differentially programed and connect them to the sense amplifier. By utilizing two differentially programed MTJs per bit, the sensing margin for the sense

amplifier is enhanced resulting in shorter delay and lower failure rates.

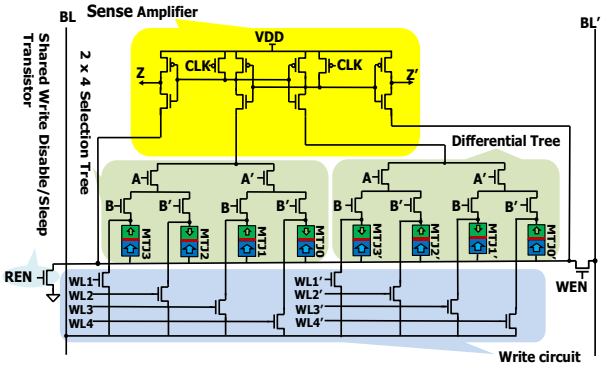

Figure 2. Proposed Dynamic Dual Rail (DDR) STT-LUT

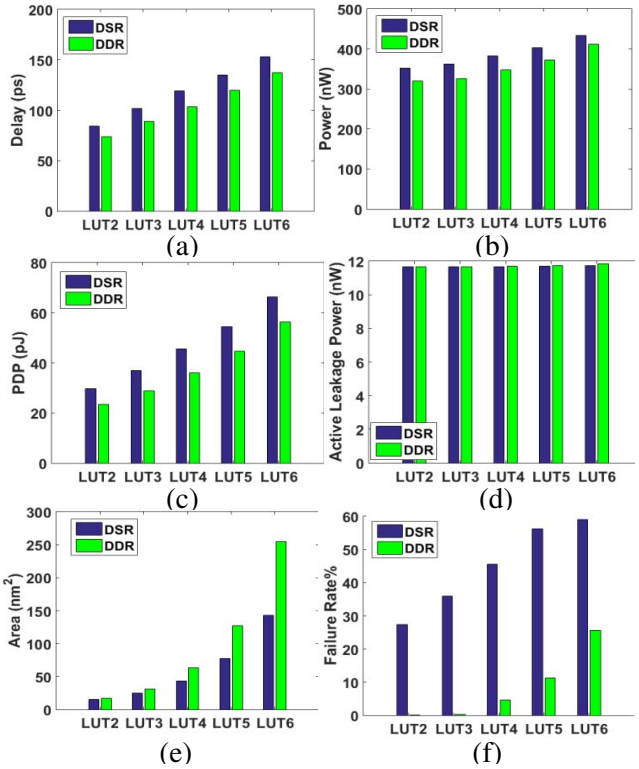

Figure 3. Comparisons of proposed DSR and DDR STT-LUTs

Fig. 3 shows the comparisons of the proposed DSR and DDR STT-LUTs. The DDR scheme shows lower delay, active power, and PDP, similar active leakage, and significantly lower read sensing failure rates. These benefits are at the cost of increase in area. Compared to the conventional DCM STT-LUT and depending on the fan-in, the proposed DDR STT-LUT reduces the read delay by 39% to 44%, the active power by 0% to 20%, and the read sensing failure by 9X to 441X.

5. SEPARATED DECODER and SPCSA

The separated decoder DDR STT-LUT is shown in Fig. 4 and the Precharge Sense Amplifier (PCSA) STT-LUTs presented in [8] use MTJs in an array with single access transistors per MTJ, and hence rely on separate decoders to provide the decoded MTJ select signals from primary LUT inputs

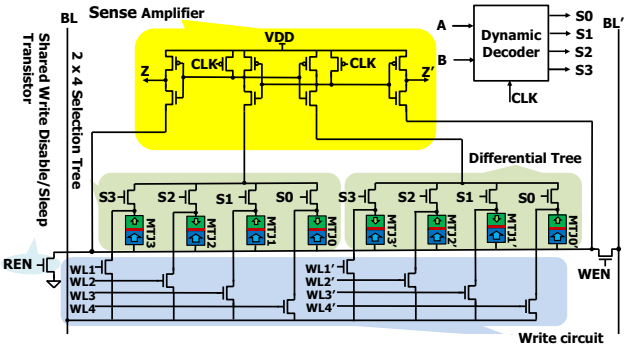

Figure 4. Separated decoder DDR STT-LUT

This separated decoding scheme reduces the number of series connected transistors above the MTJ and hence improves the speed, but may result in significant area increase for large fan-in LUTs due to the overhead of a separate dynamic decoder. The separated decoder needs to be dynamic style decoder to produce the monotonic inputs needed for the dynamic STT-LUT. Separated Precharge Sensing Amplifier (SPCSA) is presented in [6] [8] as a more reliable sensing scheme for reading the state of MTJs by separating the MTJ precharging from the sensing stage. Here, we also investigate effectiveness of SPCSA on DDR. In DDR STT-LUT with SPCSA, the MTJ prechareg stage is followed by inverters that amplify the voltage differential established by the differential MTJs during read sensing. The outputs of the inverter amplifiers are given to a voltage mode precharge sense amplifier in the second stage.

It is also possible to combine separated decoder and SPCSA methods with the DDR STT-LUT. Here we investigate the effectiveness of these design methods applied to the proposed DDR style. For this experiment we use a 10nm high performance FinFET model [5]. The MTJ is scaled down to r=14nm with insulator thickness of 0.85nm and *TMR* of 2. The standard deviation of the transistor *Vt*, and MTJ t_{ox} and r are taken as 10% of their respective nominal values. Fig. 5 shows the comparisons of the basic DDR, and DDR with either or both of the SPCSA and separated decoding schemes applied. The separated decoding scheme improves the delay due to the reduction in transistor stacking above the MTJs. The SPCSA does not result in noticeable delay reduction, because although it enhances the sensing margin of the sense amplifier stage and hence faster sensing, it adds an additional stage. In terms of power consumption, the basic DDR offers the least power due to its compactness and area density. The basic DDR also offers the lowest PDP and well as active leakage power. The basic DDR also offers the least area. The separated decoder add significant area penalty especially at high fan-in LUTs. In terms of the failure rates several observations can be made. The SPCSA scheme shows some effectiveness at low fan-in LUTs (LUT2 and LUT3), but it shows no effectiveness for

medium fan-in (LUT4) and degrades the failure rate for high fan-in LUTs (LUT5 and LUT6).

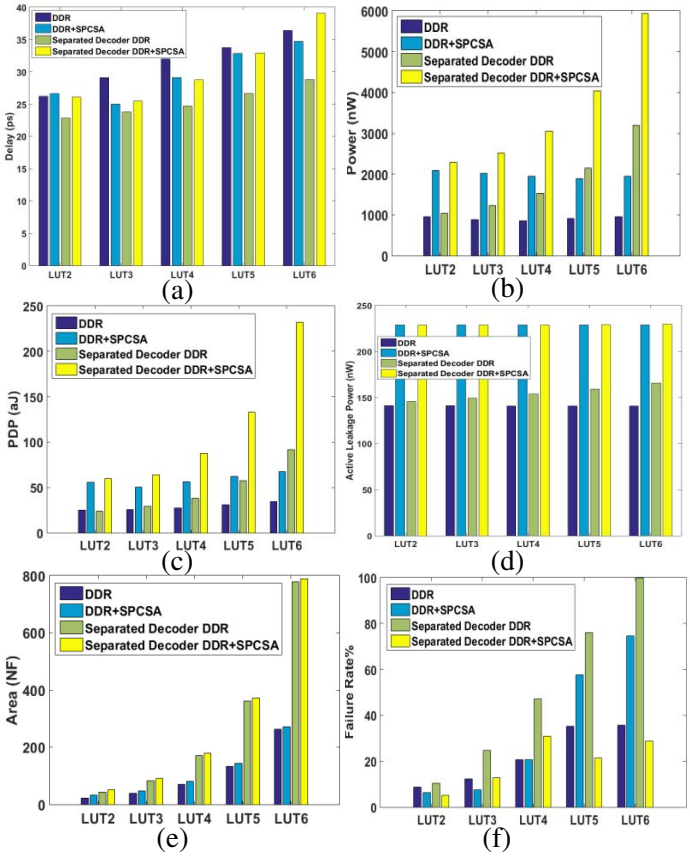

Figure 5. Comparisons of basic DDR and DDR with SPCSA and Separated Decoder schemes applied

Separated decoding degrades the failure rate irrespective of fan-in. The reason is because having more transistor stack above MTJs as in the basic DDR will result in cancellation of intra-die transistor variations and hence less mismatch between the two selected MTJ paths. Fig. 5 shows the comparisons of the basic DDR, and DDR with either or both of the SPCSA and separated decoding schemes applied. The separated decoding scheme improves the delay due to the reduction in transistor stacking above the MTJs. The SPCSA does not result in noticeable delay reduction, because although it enhances the sensing margin of the sense amplifier stage and hence faster sensing, it adds an additional stage. In terms of power consumption, the basic DDR offers the least power due to its compactness and area density. The basic DDR also offers the lowest PDP and well as active leakage power. The basic DDR also offers the least area. The separated decoder add significant area penalty especially at high fan-in LUTs. In terms of the failure rates several observations can be made. The SPCSA scheme shows some effectiveness at low fan-in LUTs (LUT2 and LUT3), but it shows no effectiveness for medium fan-in (LUT4) and degrades the failure rate for high fan-in LUTs (LUT5 and LUT6).

Separated decoding degrades the failure rate irrespective of fan-in. The reason is because having more transistor stack above MTJs as in the basic DDR will result in cancellation of intra-die transistor variations and hence less mismatch between the two selected MTJ paths.

The combined separated decoding and SPCSA offers reduction in failure rate however due to the significant area overhead associated with the separated decoding, this improvement is too expensive and can be achieved for the basic DDR too by increasing its area. Overall, the basic DDR STT-LUT style seems to be most effective even when considering the SPCSA and separated decoding schemes.

The combined separated decoding and SPCSA offers reduction in failure rate however due to the significant area overhead associated with the separated decoding, this improvement is too expensive and can be achieved for the basic DDR too by increasing its area. Overall, the basic DDR STT-LUT style seems to be most effective even when considering the SPCSA and separated decoding schemes.

6. CONCLUSION

In this paper, two new designs for STT-LUTs are presented that operates using dynamic voltage mode logic and outperform the conventional dynamic current mode STT-LUT in all aspects of delay, active mode power, and robustness against process variations. The comprehensive analysis of the proposed designs shows that the proposed designs are able to offer higher fan-in STT-LUTs under a given power, performance, and robustness constraints. Moreover, the proposed DDR style remains to be the most effective designs style even when considering the SPCSA and separated decoding schemes.

REFERENCES

[1] M. Rasquinha, D. Choudhary, S. Chatterjee, S. Mukhopadhyay, and S. Yalamanchili, "An energy efficient cache design using Spin Torque Transfer (STT) RAM," in *Low-Power Electronics and Design (ISLPED), 2010 ACM/IEEE International Symposium on*, 2010, pp. 389-394.

[2] D. Suzuki, M. Natsui, S. Ikeda, H. Hasegawa, K. Miura, J. Hayakawa, *et al.*, "Fabrication of a nonvolatile lookup-table circuit chip using magneto/semiconductor-hybrid structure for an immediate-power-up field programmable gate array," in *2009 Symposium on VLSI Circuits*, 2009, pp. 80-81.

[3] W. Zhao, E. Belhaire, C. Chappert, and P. Mazoyer, "Spin transfer torque (STT)-MRAM--based runtime reconfiguration FPGA circuit," *ACM Trans. Embed. Comput. Syst.,* vol. 9, pp. 1-16, 2009.

[4] M. W. Allam and M. I. Elmasry, "Dynamic current mode logic (DyCML): a new low-power high-performance logic style," *IEEE Journal of Solid-State Circuits,* vol. 36, pp. 550-558, 2001.

[5] http://www.eas.asu.edu/~ptm. [Online].

[6] W. Kang, E. Deng, J. O. Klein, Y. Zhang, Y. Zhang, C. Chappert, *et al.*, "Separated Precharge Sensing Amplifier for Deep Submicrometer MTJ/CMOS Hybrid Logic Circuits," *IEEE Transactions on Magnetics,* vol. 50, pp. 1-5, 2014.

[7] Y. Taur and T. H. Ning, *Fundamentals of Modern VLSI Devices*: Cambridge University Press, 2009.

[8] E. Deng, W. Kang, Y. Zhang, J. O. Klein, C. Chappert, and W. Zhao, "Design Optimization and Analysis of Multicontext STT-MTJ/CMOS Logic Circuits," *IEEE Transactions on Nanotechnology,* vol. 14, pp. 169-177, 2015.