Process Variation Immune and Energy Aware Sense Amplifiers for Resistive Non-Volatile Memories

Soheil Salehi, Ronald F. DeMara
Department of Electrical and Computer Engineering
University of Central Florida
Orlando, Florida 32816-2362
Email: soheil.salehi@knights.ucf.edu, demara@mail.ucf.edu

***Abstract* — Spin-Transfer Torque Magnetic Random Access Memory (STT-MRAM) has been explored as a post-CMOS technology for embedded and data storage applications seeking non-volatility, near-zero standby energy, and high density. Towards attaining these objectives for practical implementations, various techniques to mitigate the specific reliability challenges associated with STT-MRAM elements are surveyed, classified, and assessed herein. Some solutions to the reliability issues identified are addressed to realize reliable STT-MRAM designs. In an attempt to further improve the process variation immunity of the Sense Amplifiers (SAs), two new SAs are introduced: Energy Aware Sense Amplifier (EASA) and Variation Immune Sense Amplifier (VISA). Results have shown that EASA and VISA achieve superior performance in most cases compared to two of the most common SAs, namely PCSA and SPCSA respectively, while reducing Bit Error Rate (BER) and increasing reliability.**

Keywords* — *Spin-Transfer Torque storage elements, STT-MRAM, Magnetic Tunnel Junction (MTJ), Reliability, Process Variation, Energy Efficient, Low Power, Sense Amplifier

I. INTRODUCTION

High reliability and energy efficient switching of STT-based devices is highly sought in the active area of emerging device research. In this paper, we examine improvements to the *Sense Amplifier (SA)* design that can achieve both of these objectives on a continuum of energy versus reliability trade-offs. Due to the increase in *Process Variation (PV)* as technology shrinks, *Spin-Transfer Torque Magnetic Random Access Memory (STT-MRAM)* cell reliability has become a significant concern in high density memory arrays and cache designs [1]. While a collection of innovative methods has been previously proposed to increase reliability and performance, each incurs costs and challenges that may lead to a sub-optimal performance profile. Of particular urgency is the need to reduce the effects of device mismatch and variation due to scaling of the devices, especially with respect to the use of different SAs. The most common SAs which have been studied are *Pre-Charge SA (PCSA)* [2] and *Separated Pre-Charge SA (SPCSA)* [3]. While, PCSA offers improved sense latency and power consumption compared to SPCSA, it suffers from increased *Bit Error Rate (BER)* [3]. SPCSA, on the other hand, offers increased reliability while incurring an acceptable increase in sense latency and power consumption with a negligible area overhead compared to PCSA [3]. In addition to addressing scalability to technologies beyond 10nm where traditional memory elements such as *Static Random Access Memory (SRAM)* and *Dynamic Random Access Memory (DRAM)* face significant scaling challenges [4], innovations to mitigate the power wall and reduce leakage power consumption occupy the forefront of on-chip memory design considerations [5, 6]. Power consumed by memory elements can become a significant portion of total power [7, 8] whereby the processing cores rely on these memory arrays that are significant contributors to standby mode power consumption. These concerns motivate the research into balancing energy and reliability effectively.

In this paper, our primary focus is performance improvement of the PCSA in terms of *Energy Delay Product (EDP)* and reliability improvement of SPCSA in terms of *Bit Error Rate Reduction (BERR)*. First, a variety of SAs and an overview of STT-MRAM functionality and technology aspects is provided are discussed as related works in Section 2. In Section 3, reliability issues of the STT-MRAM are identified and classified. In Section 4, new sensing circuits are introduced and simulation result and analysis for the proposed designs are provided in Section 5. In addition, a new metric is introduced as *Sense Error Energy Ratio (SEER)* that provides an insight on overall performance and reliability of SAs. Finally, this paper concludes in Section 6 with analysis of the proposed circuits. Performance comparison of the designs and metrics proposed in this paper is listed in Table 1.

Table 1: Performance comparison of Sense Amplifier designs.

Design	***Delay***	***Power***	***EDP***	***BERR***	***SEER (Iso-BERR)***	***SEER (Iso-EDP)***
PCSA [2]	✓✓✓	✓	✓✓	-	✓✓	-
EASA (Proposed herein)	✓	✓✓✓	✓✓✓	- -	✓✓✓	- -
SPCSA [3]	✓	- -	- -	✓✓	- -	✓✓
VISA (Proposed herein)	✓	-	-	✓✓✓	-	✓✓✓

Table 2: Sensing Schemes reviewed and their attributes.

Attribute	***Reference Numbers***
Process Variation Tolerant	[3, 10, 14]
Read Disturb Reduction	[3, 5, 11]
Wide Sense Margin	[2, 3, 4, 10, 12]
Write Polarization Asymmetry Reduction	[10, 13]
Yield Increase	[4, 11]

II. Related Work

Among promising devices, nanomagnetic devices such as STT-MRAM are considered as feasible candidates for post-CMOS devices [9]. STT-MRAM offers low read access time, near-zero standby power consumption, and small area requirements. STT-MRAM also offers integration with backend CMOS processes. To embrace their adoption in anticipated applications, a palette of cooperating reliability techniques is sought for comparison at the bit-cell level. *Magnetic Tunnel Junction (MTJ)* devices are constructed with layered pillars of ferromagnetic and insulating layers to leverage magnetic orientations that can be controlled and sensed in terms of electrical signal levels [4]. Identical magnetic orientation results in *Parallel (P)* configuration which introduces a low resistance that can be represented as logical "0" and opposite magnetic orientation results in *Anti-Parallel (AP)* configuration which introduces high resistance that can be represented as logical "1". In STT-MRAM, relative resistance values are used to determine the bit value stored in each MTJ. Research has shown that STT-MRAM SAs' performance span in three ranges across all proposed design strategies. Some suggestions are listed in Table 2. The highest performance strategies deliver a wide *Sense Margin (SM)* of approximately above 300mV while incurring low energy and power consumption in the order of pico-Joules and micro-Watts respectively with less than 5ns sense latency [2-5, 10-14]. The MTJ resistance in P (θ=0°), and AP (θ=180°) states is expressed by the following equations:

$$R(\theta) = 2R_{MTJ} \times \frac{1 + TMR}{2 + TMR + TMR.cos\theta} = \begin{cases} R_p = R_{MTJ}, & \theta = 0° \\ R_{ap} = R_{MTJ}(1 + TMR), & \theta = 180° \end{cases} \quad (1)$$

$$R_{MTJ} = \frac{t_{ox}}{Factor \times Area.\sqrt{\varphi}} \exp(1.025 \times t_{ox}.\sqrt{\varphi}) \quad (2)$$

$$TMR = TMR_0 / 1 + (\frac{V_b}{V_h})^2 \quad (3)$$

where V_b is the bias voltage, V_h = 0.5V is the bias voltage when *Tunnel Magneto-Resistance (TMR)* is half of the TMR_0, t_{ox} is the oxide thickness of MTJ, Factor is obtained from the resistance-area product value of the MTJ that relies on the material composition of its layers, Area is the surface of MTJ, and Φ is the oxide layer energy barrier height [15].

III. Reliablity Challenges Facing STT-MRAM

STT-MRAM has several advantages over other emerging memory technologies, however, it faces some distinct reliability challenges involving read and write failures [5] as listed in this Section. STT-MRAM bit errors can be significantly influenced due to PV, which precipitates another important issue that STT-MRAM suffers from as well as suffering from in addition to its unique intrinsic thermal randomness [1]. These variations include variation in the CMOS periphery circuits, MTJ geometric variation, and initial angle of the MTJ [1]. The difference between the sensed bit-line voltage and the reference voltage, which is known as the SM, will be small due to the wide distribution of MTJ resistance which can also result in a false detection scenario [10]. Errors due to the STT-MRAM's physical nature failures will be categorized into *Transient Faults* and *Permanent Faults.* Transient faults, also referred to as incorrect signal condition, are mostly caused by the parameters of the free layer such as current density and thermal stability factor. Permanent faults, also referred to as destructive device damage, are initially caused by susceptibility to the sensitive parameters of oxide barrier such as oxide thickness and TMR ratio [12]. As a result of these issues, demand for an advanced sensing circuit, which can provide the required SM along with low power operation has been increased.

IV. Proposed Designs

Reducing the amount of resistance in the MTJ devices' paths increases the SM and voltage headroom. This reduces the error rate in scaled technology nodes as supply voltage is reduced, which is the case with SPCSA versus PCSA [3]. To improve the performance and reliability of PCSA and SPCSA respectively, *Energy Aware SA (EASA)* and *Variation Immune SA (VISA)* are proposed herein as shown in Fig. 1 and Fig. 2 respectively. As it has been shown in [16] and [17], using *Transmission Gates (TGs)*, which provide near optimal full-swing switching, reduces either the circuits' vulnerability to reliability issues caused by PV or the

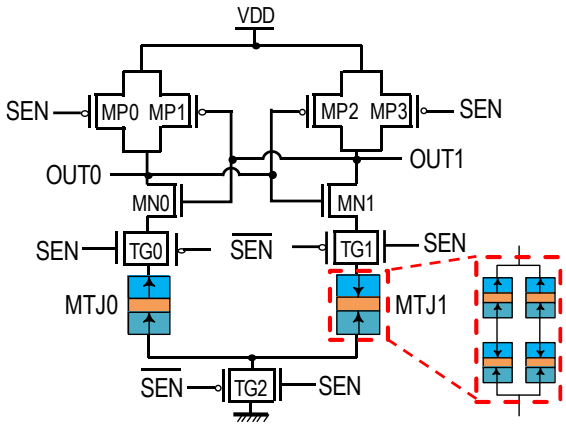

Fig. 1: Energy Aware Sense Amplifier (EASA).

energy consumption by reducing the leakage energy. Thus, **TG0**, **TG1**, and **TG2** are added to improve the performance of the PCSA. When **SEN** signal is low the pre-charge stage causes **MP0** and **MP3** to turn on, which then results in both **OUT0** and **OUT1** to be connected to V_{DD}. During the pre-charge stage, **TG0**, **TG1**, and **TG2** are off, which results in reduction of leakage energy through **MN0** and **MN1** that are turned on due to **OUT0** and **OUT1** being pre-charged to V_{DD}. As soon as the **SEN** signal becomes high, **MP0** and **MP3** to turn off and **TG0**, **TG1**, and **TG2** turn on and the sensing stage starts. Based on the resistance difference between the two branches with MTJs with regard to the MTJs' states, either **OUT0** or **OUT1** begin to discharge more rapidly. If **OUT0** discharges faster, then **MP2** turns on and **OUT1** will be connected to V_{DD}. On the other hand, if **OUT1** discharges faster **MP1** will be turned on and **OUT0** will be connected to V_{DD}. The novelty of this design is using TGs to reduce the leakage energy in the path of MTJ devices. However, adding TGs on the path of MTJs will add more resistance that will affect the SM as well as increasing vulnerability to PV, which may induce read errors.

On the other hand, **TG0**, **TG1**, and **TG2** are added to improve the performance and reliability of the SPCSA. Similar to EASA, in the VISA circuit when **SEN** signal is low, the SA will be in pre-charge stage and **MP0**, **MP1**, **MP4**, and **MP5** will turn on and cause **OUT0**, **OUT1**, **Node0**, and **Node1** to be connected to V_{DD} and **TG2** will be off. Thus, the output of **INV0** and **INV1** will be low,

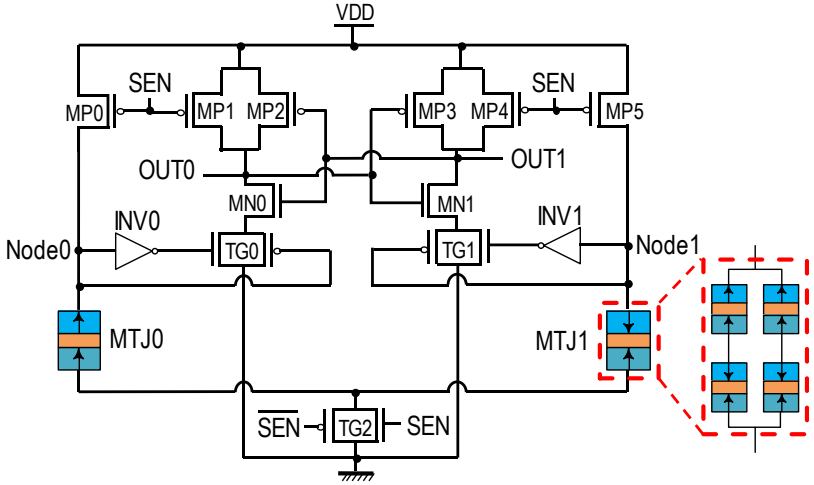

Fig. 2: Variation Immune Sense Amplifier (VISA).

turning **TG0** and **TG1** off. This results in reduction of leakage energy through **MN0** and **MN1** that are turned on due to **OUT0** and **OUT1** being pre-charged to V_{DD}. When the **SEN** signal becomes high, the SA will then be in the sensing stage and **TG2** will turn on, while **MP0**, **MP1**, **MP4**, and **MP5** are off. Based on the resistance difference of the MTJs, **Node0** and **Node1** start discharging. If **Node0** discharges faster, output of **INV0** will be high, which results in turning on **TG0** faster than **TG1**. This will result in discharging **OUT0** quicker than **OUT1**, causing **MP3** to turn on and connect **OUT1** to V_{DD}. On the other hand, if **Node1** discharges faster, **INV1** output will turn on **TG1** quicker and that results in **MP2** to turn on and cause **OUT0** to connect to V_{DD}. **INV0** and **INV1** are used to amplify the voltage difference of **Node0** and **Node1**. These inverters can become important elements for the circuit due to presence of PV. The novelty of this design is using TGs to reduce the effects of PV on the inverters to increase reliability and to reduce the leakage energy in the path of MTJ devices. Since only one TG is added on the main discharge paths of MTJs and

the other two TGs are not added to the main discharge paths of the MTJs, the chances of bit error occurrence due PV will be reduced. In order to further improve PV immunity of the reference cell in both SAs, we utilize **$(MTJ_P+MTJ_{AP})||(MTJ_P+MTJ_{AP})$** configuration for the reference MTJ, referred to as **MTJ1** in Fig. 1 and Fig. 2, that will result in an ideal reference cell in terms of resistance, which is $(MTJ_P+MTJ_{AP})/2$ and provides increased SM [14].

V. RESULTS

Simulation results are extracted using 22nm Predictive Technology Model (PTM) [18]. In this paper, we have utilized the approach proposed in [15] to model the behavior of STT-MRAM devices, in which a Verilog-AMS model is developed using the aforementioned equations. Then, the model is leveraged in a SPICE circuit simulator to validate the functionality of the designed circuits. The design parameters and PV values are provided in Table 2. All **PMOS** and **NMOS** transistors are considered minimum size except transistors used in **INV0** and **INV1**. Since **INV0** and **INV1** are vital to the reliability of the circuit, we have optimized the size of their transistors to maintain width to length ratio (W/L) of 4 to provide reliable functionality. Each design implemented and proposed herein was analyzed using *Monte Carlo (MC)* simulation in presence of PV elements such as *Threshold Voltage (V_{th})* variations of MOS transistors. In order to fully evaluate the proposed circuits, 10,000 MC simulations were performed considering different standard deviations for CMOS transistors' V_{th} and also MTJ's MgO thickness and shape area in order to have a variety of cases to analyze during the simulation. These simulations vary the V_{th}, width, and length of the transistors in the netlist based on a Gaussian distribution having a mean equal to the nominal model card for PTM and σV_{th} as provided in [19]. Ideally, the σV_{th} can be adapted to accommodate local and global variations, or their combined effects as considered in this work. Overall variation that has been taken into account here for the MTJs has an effect of σTMR=1%. As introduced in this paper, *Bit Error Rate (BER)* is calculated based on the number of incorrect output bits divided by all the input bits applied in both P and AP states. Due to structural limitations of MTJ devices, the TMR ratio of 100% is considered as the baseline design herein [2, 3, 20]. Table 3 lists MC Results with TMR=100%, σTMR=1%, σV_{th}=10%, and MTJ_{Ref}=5.7KΩ.

Results provided in Table 3, Table 4, and Table 5 indicate and as depicted in Fig. 3 and Fig. 4, EASA attains 2.1-fold EDP improvement on average over PCSA. However, EASA suffers from 2.6% increased BER on average considering TMR=100%, σTMR=1%, and σV_{th}=10% compared to PCSA. On the contrary, considering TMR=100%, σTMR=1%, and σV_{th}=10%, VISA offers 1.2% less BER on average while providing nearly 1.5-fold EDP improvement on average compared to SPCSA. Furthermore, our results indicate that by optimizing the reference MTJ and using $(MTJ_P+MTJ_{AP})/2$ configuration, the BER can be reduced by 8.9% on average considering TMR=100%, σTMR=1%, and σV_{th}=10%. In addition, our results indicate that the larger TMR reduces the impact of PV on sensing output by increasing SM, as anticipated. In order to be able to compare the reliability and performance of different SAs more comprehensive, SEER metric is introduced herein, which represents the ratio of BERR to average EDP as demonstrated in (4). BERR is calculated as shown in (5). SEER will enable the designers to effectively assess the most appropriate SA for their need based on whether they are seeking reliability or energy efficiency. Any increase in BERR or decrease in EDP will cause the SEER to increase. As a result, larger values of SEER imply increased reliability and performance and on the contrary, small values of SEER imply decreased reliability and performance. Based on physical layout design of PCSA, EASA, SPCSA, and VISA shown in Fig. 5a, 5b, 5c, and 5d respectively, it is clear that the proposed EASA and VISA designs offer small area overhead compared to their counterparts, considering overall size of memory.

Table 2: Simulation Parameters.

Parameter				***Value***	***Std. Dev.***
Pmos	*V_{th} (Threshold Voltage)*			460mV	50mV(10%)
	Width (=2×Length)			44nm	0.44nm(1%)
Nmos	*V_{th} (Threshold Voltage)*			500mV	50mV(10%)
	Width (=Length)			22nm	0.22nm(1%)
MTJ	*MgO Thickness*			0.85nm	Effects of variation are applied to TMR
	Shape Area	*main MTJ (MTJ0)*		$(\frac{\pi}{4})$x40x40nm^2	
		reference MTJ (MTJ1)	*MTJ_{AP}*	$(\frac{\pi}{4})$x30x30nm^2	
			(MTJ_P+MTJ_{AP}) || (MTJ_P+MTJ_{AP})	$(\frac{\pi}{4})$x40x40nm^2	
	ϕ (Potential Barrier Height)			0.4 V	N/A
	R·A (Resistance Area Product)			5Ω·μm^2	N/A
	α (Damping Factor)			0.01	N/A
	TMR (Tunnel Magneto Resistance)			100%	1%
Nominal Voltage (V_{DD})				1.0 V	N/A
SEN Signal Period (T)				1ns	N/A

Table 3: MC Results for σTMR=1%, σV_{th}=10%, and MTJ_{Ref}=5.7KΩ.

Design	Area (Device Count)			Anti-Parallel (6.4 KΩ)				Parallel (3.2 KΩ)			
	Pmos	Nmos	MTJ	Delay (ps)	Power (μW)	EDP (fJ×ps)	σEDP (fJ×ps)	Delay (ps)	Power (μW)	EDP (fJ×ps)	σEDP (fJ×ps)
PCSA	4	3	2	16.10	0.72	11.63	2.49	17.12	0.71	12.15	2.23
EASA	7	5	2	22.53	0.23	5.31	1.18	27.20	0.22	6.12	1.12
SPCSA	8	5	2	25.72	2.32	59.79	8.38	25.85	2.31	59.71	7.06
VISA	11	7	2	24.01	1.87	45.06	7.13	23.74	1.84	43.72	5.76

Table 4: BER (%) for MTJ_{Ref}=5.7KΩ and MTJ_P=3.2KΩ.

Design	BER (%) (σTMR=1% and σVth=10%) for TMR=					
	100%	150%	200%	250%	300%	350%
PCSA	25.19	14.395	8.895	6.835	6.225	6.125
EASA	27.42	17.865	12.09	9.28	8.285	7.97
SPCSA	20.76	8.415	4.325	3.645	3.595	3.59
VISA	19.175	7.135	3.555	3.385	2.985	2.985

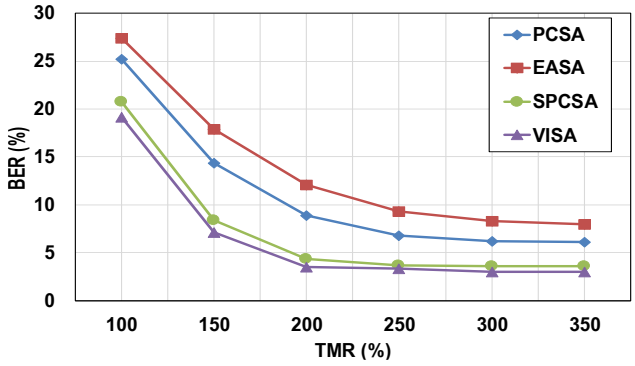

Fig. 3: BER (%) with σTMR=1%, σV_{th}=10%, MTJ_{Ref}=5.7KΩ, and MTJ_P=3.2KΩ.

Table 5: BER (%) for MTJ_{Ref}=(MTJ_P+ MTJ_{AP})/2 and MTJ_P=3.2KΩ.

Design	BER (%) (σTMR=1% and σVth=10%) for TMR=					
	100%	150%	200%	250%	300%	350%
PCSA	23.65	14.455	8.415	4.885	2.53	1.405
EASA	26.575	17.915	11.64	7.4	4.615	2.73
SPCSA	17.545	8.325	3.38	1.18	0.385	0.1
VISA	16.06	7.16	2.79	0.975	0.305	0.075

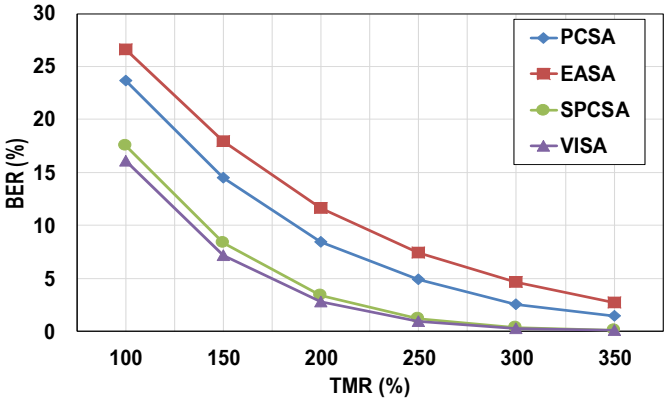

Fig. 4: BER (%) with σTMR=1%, σV_{th}=10%, MTJ_{Ref}=(MTJ_P+ MTJ_{AP})/2, and MTJ_P=3.2KΩ.

VI. Conclusion

Results indicate that the proposed SAs have low power and delay overheads with acceptable area overhead since the SA is shared among all memory cells within an array. EASA and VISA provide improved performance in most cases compared to the

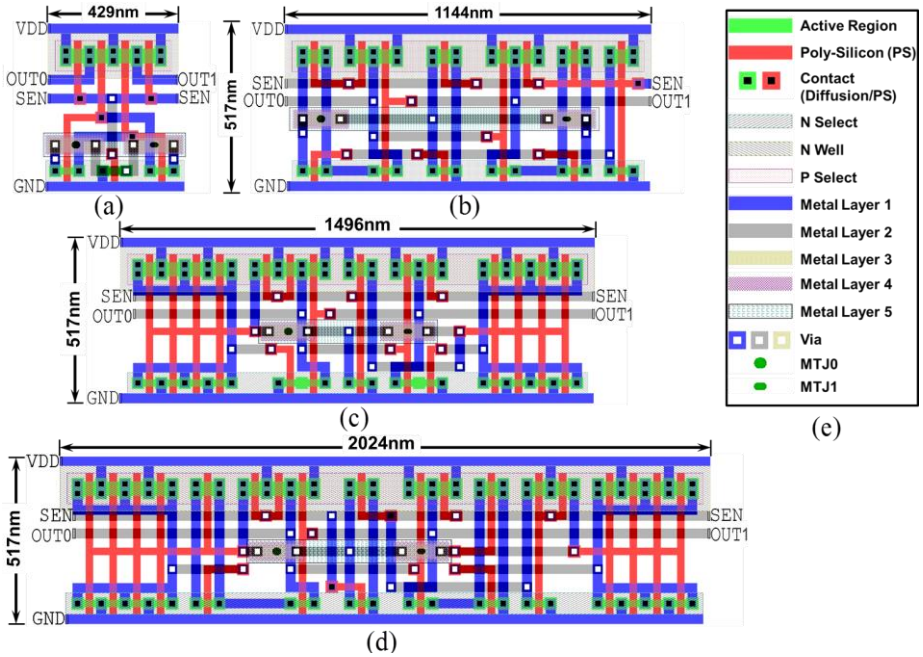

Fig. 5: a) PCSA, b) EASA, c) SPCSA, and d) VISA layouts, e) legend.

original PCSA and SPCSA respectively. Considering TMR=100%, σTMR=1%, and σV_{th}=10%, EASA delivers 2.1-fold EDP improvement on average compared to PCSA while VISA provides nearly 1.5-fold EDP improvement on average and 1.2% reduced BER on average compared to SPCSA. Our results also indicate that CMOS PV has a more significant impact on circuit's reliability compared to MTJ TMR variation and high values of TMR provide more reliable SMs.

REFERENCES

[1] S. Salehi, D. Fan, and R. F. DeMara, "Survey of STT-MRAM Cell Design Strategies: Taxonomy and Sense Amplifier Tradeoffs for Resiliency," *ACM J. on Emerging Technologies in Computing Syst. (JETC)*, p. 15, 2016.

[2] W. Zhao, C. Chappert, V. Javerliac, and J.-P. Nozière, "High speed, high stability and low power sensing amplifier for MTJ/CMOS hybrid logic circuits," *IEEE Trans. Magn.*, vol. 45, pp. 3784-3787, 2009.

[3] W. Kang, *et al.*, "Separated precharge sensing amplifier for deep submicrometer MTJ/CMOS hybrid logic circuits," *IEEE IEEE Trans. Magn.*, vol. 50, pp. 1-5, 2014.

[4] Z. Sun, H. Li, Y. Chen, and X. Wang, "Voltage driven nondestructive self-reference sensing scheme of spin-transfer torque memory," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 20, pp. 2020-2030, 2012.

[5] R. Bishnoi, M. Ebrahimi, F. Oboril, and M. B. Tahoori, "Read disturb fault detection in STT-MRAM," in *Proc. of Int. Test Conf. (ITC)*, pp. 1-7, 2014.

[6] R. A. Ashraf, *et al.*, "Reactive rejuvenation of CMOS logic paths using self-activating voltage domains," in *Proc. of Int. Symp. on Circuits and Syst. (ISCAS)*, pp. 2944-2947, 2015.

[7] E. Kultursay, M. Kandemir, A. Sivasubramaniam, and O. Mutlu, "Evaluating STT-RAM as an energy-efficient main memory alternative," in *Proc. of Int. Symp. on Performance Analysis of Syst. and Software (ISPASS)*, pp. 256-267, 2013.

[8] N. Khoshavi, X. Chen, J. Wang, and R. F. DeMara, "Bit-Upset Vulnerability Factor for eDRAM Last Level Cache Immunity Analysis," in *Proc. of 17th Int. Symp. on Quality Electronic Design (ISQED)*, pp. 6-11, 2016.

[9] *International Technology Roadmap for Semiconductors (ITRS)*. Available: http://www.itrs.net

[10] E. Eken, *et al.*, "A Novel Self-Reference Technique for STT-RAM Read and Write Reliability Enhancement," *IEEE IEEE Trans. Magn.*, vol. 50, pp. 1-4, 2014.

[11] J. P. Kim, *et al.*, "A 45nm 1Mb embedded STT-MRAM with design techniques to minimize read-disturbance," in *Proc. of Symp. on VLSI Circuits-Digest of Technical Papers*, 2011.

[12] W. Zhao, *et al.*, "Failure and reliability analysis of STT-MRAM," *Microelectronics Reliability*, vol. 52, pp. 1848-1852, 2012.

[13] H. Lee, *et al.*, "Design of a Fast and Low-Power Sense Amplifier and Writing Circuit for High-Speed MRAM," *IEEE IEEE Trans. Magn.*, vol. 51, pp. 1-7, 2015.

[14] W. Kang, *et al.*, "Variation-tolerant high-reliability sensing scheme for deep submicrometer STT-MRAM," *IEEE IEEE Trans. Magn.*, vol. 50, pp. 1-4, 2014.

[15] R. Zand, A. Roohi, D. Fan, and R. F. DeMara, "Energy-Efficient Nonvolatile Reconfigurable Logic using Spin Hall Effect-based Lookup Tables," *IEEE Trans. Nanotechnol.*, 2016.

[16] A. Alzahrani and R. F. DeMara, "Process variation immunity of alternative 16nm HK/MG-based FPGA logic blocks," in *Proc. of 58th Int. Midwest Symp. on Circuits and Syst. (MWSCAS)*, pp. 1-4, 2015.

[17] R. Zand, A. Roohi, S. Salehi, and R. F. DeMara, "Scalable Adaptive Spintronic Reconfigurable Logic using Area-Matched MTJ Design," *IEEE Trans. Circuits and Syst. II: Exp. Briefs*, vol. 63, p. 5, July 2016.

[18] *22nm Predictive Technology Model (PTM)*. Available: http://ptm.asu.edu/modelcard/HP/22nm_HP.pm

[19] Y. Ye, F. Liu, S. Nassif, and Y. Cao, "Statistical modeling and simulation of threshold variation under dopant fluctuations and line-edge roughness," in *Proc. of 45th Design Automation Conf. (DAC)*, pp. 900-905, 2008.

[20] E. Deng, *et al.*, "Design optimization and analysis of multicontext STT-MTJ/CMOS logic circuits," *IEEE Transactions on Nanotechnology*, vol. 14, pp. 169-177, 2015.