Extending Modular Redundancy to NTV: Costs and Limits of Resiliency at Reduced Supply Voltage

Rizwan A. Ashraf
Department of Electrical and
Computer Engineering
University of Central Florida
Orlando, Florida 32816–2362
Email: rizwan.ashraf@knights.ucf.edu

Ahmad Alzahrani
Department of Electrical and
Computer Engineering
University of Central Florida
Orlando, Florida 32816–2362
Email: azahrani@knights.ucf.edu

Ronald F. DeMara
Department of Electrical and
Computer Engineering
University of Central Florida
Orlando, Florida 32816–2362
Email: demara@mail.ucf.edu

Abstract—Near Threshold Voltage (NTV) operation offers reduced energy consumption with an acceptable increase in delay for targeted applications. However, increased *Soft Error Rate (SER)* has been identified as a significant concern at NTV. In this work, tradeoffs are evaluated regarding use of spatial redundancy to mask increased SER at NTV. We show that conventional spatial redundancy techniques such as N Modular Redundancy (NMR) can exhibit higher mean delays than simplex arrangements due to increased sensitivity to *Process Variation (PV)* exacerbated at NTV. The implications on energy consumption and delay for benchmark circuits implemented using 45nm and 22nm Predictive Technology Models (PTM) are evaluated through simulation. Results indicate that the energy overheads of NMR systems in near-threshold region tend to be slightly higher than the expected N-fold increase in cost for nominal voltage operation.

I. INTRODUCTION

Improvement in power efficiency of CMOS circuits is desired due to thermal limitations associated with technology scaling. This *power wall* restricts the number of simultaneously activated components, e.g. the percentage of active cores within a many-core processor die. In this regard, supply voltage (V_{DD}) scaling is widely recognized as the most effective lever to reduce power dissipation / energy consumption for CMOS logic devices.

Total energy consumption is determined by dynamic and static energy components which are both dependent on V_{DD}. The dynamic energy has a quadratic relationship with V_{DD}, whereas the static energy has a linear dependence. Taken to an extreme, V_{DD} reduction would encounter theoretical limitations around a value that is twice the thermal voltage, i.e. about 36mV [1]. However, down-scaling the supply voltage below the transistor threshold voltage V_{th} specifies an operating region exhibiting a highly undesirable exponential increase in delay. Thus, leakage energy begins to supersede at a certain point, such that designs utilizing this *Sub-Threshold region* have much more limited applicability. Hence, operation in the *Near-Threshold region* is sought, as it provides an energy-efficient operating point. Here, the V_{DD} is set to be slightly above the V_{th} of the transistors to provide a 10X delay improvement compared to operation in the sub-threshold region with only a 50% reduction in energy savings [1]. Taking all of these factors into consideration, NTV can be preferred to provide up to 6X energy savings as compared to operating at nominal voltage [1], [2].

While NTV offers an attractive approach to balance energy consumption versus delay for power-constrained applications such as high-performance computing [1], it may also introduce reliability implications. In particular, radiation-induced *Single Event Upsets (SEUs)* which cause soft errors are expected to increase at this operating region [3],[4]. These errors may manifest as a random bit flip in a memory element or a transient charge within a logic path which is ultimately latched by a flip-flop. While soft errors in memory elements are feasible to detect and correct using Error Correcting Codes [1], their resolution in logic paths typically involves the use of spatial or temporal redundancy which allow area versus performance tradeoffs [5]. In this paper, we propose to utilize the well-accepted spatial redundancy technique of modular redundancy to mask soft errors in logic paths.

The contributions of this paper are:

- *Impacts of Variability on NMR Systems at NTV:* identifying the increased impact of PV at NTV for 45nm and 22nm based NMR Systems, and

- *Cost of Redundancy at NTV:* expressing the delay cost of NMR systems at NTV for common values of N within a given energy budget at nominal supply voltage.

The remainder of the paper is organized as follows. Section II establishes the need for mitigating soft errors in logic paths, and evaluates them for the NTV operating region. Section III develops a generic framework to quantify the level of protection established through error masking schemes. Section IV quantifies the effects of variability and provides relevant insights into the low-energy benefits of near-threshold operation for NMR systems. Section V evaluates voltage guard-banding to achieve these benefits while accommodating variations in NMR arrangements. The experimental results for these conditions are presented in Section VI in terms of NMR energy and delay performance. Finally, conclusions and future work place the results into the context of current considerations and future trends.

II. SOFT ERRORS IN LOGIC PATHS

The contribution of soft errors in logic paths as opposed to memory elements is becoming significant as the supply voltage is scaled down to the threshold region. It was predicted

in [6] that the SER of logic circuits per die will become comparable to the SER for unprotected memory elements, which was later verified through experimental data for a recent microprocessor [4]. Operation at NTV is expected to exaggerate these trends. For instance, [4] states that SER increases by approximately 30% per each $0.1V$ decade as V_{DD} is decreased from 1.25V to 0.5V. In the NTV region, it is shown through both simulation and experiment at the 40nm and 28nm nodes, that SER doubles when V_{DD} is decreased from 0.7V to 0.5V. Primarily, the critical charge needed to cause a failure decreases as V_{DD} is scaled and SER has an exponential dependence on critical charge [4]. Such trends are consistent with decreasing feature sizes due to technology scaling [7].

In logic paths, there are three inherent masking mechanisms that prevent the propagation of a spurious transient pulse along a path towards the input of a flip-flop/latch, where it may be registered to cause an error [6]:

1) *Logical Masking:* occurs when a transient pulse does not affect the computation in other gates along the path towards the output for a given input vector,

2) *Electrical Masking:* due to the attenuation of the glitch as it passes through subsequent logic gates, and

3) *Latching-window Masking:* occurs when the generated glitch does not occur within the setup and hold time window of the flip-flop.

It is evident that logical masking effect is not impacted by operation at lower voltages. However, as operating frequencies at NTV are expected to be low, it has been suggested that pipeline stages consist of fewer gates to regain lost throughput. This will consequently lower the benefit of both logical and electrical masking. In addition, the electrical attenuation is lowered at low supply voltages as large pulse-width transients are created. However, there is a positive effect on masking due to latching-window masking as operating frequencies are lowered. The latching-window masking is also dependent on the design of flip-flop utilized [4], where some designs show more SER immunity as compared to others. Overall, reduced pipeline depths, technology scaling, and voltage reduction can be anticipated to have detrimental impact on logic SER. Thus, there is a need to develop effective soft error mitigation techniques for reliable NTV operation.

III. Soft Error Masking at NTV

The SER in logic paths can be reduced by schemes such as gate-sizing [8] or dual-domain supply voltage assignments [9] to harden components which are more susceptible to soft-errors. These techniques tradeoff increased area and/or power to reduce SER of the logic circuit, but may not be able to provide comprehensive coverage. For instance, SER reduction of only 33.45% is demonstrated in [9] using multiple voltage assignments. One option identified for masking soft errors is spatial redundancy, and in particular the readily-accepted use of *Triple Modular Redundancy (TMR)*, as being effective for mitigating soft errors. TMR is considered to be appropriate for applications which demand immunity to soft errors and also are able to accommodate its inherent overheads.

Spatial redundancy is often employed in mission-critical applications to ensure system operation even in unforseen
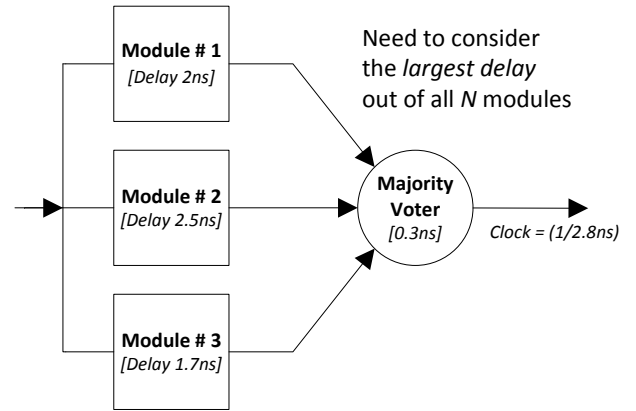

Fig. 1. Propagation delay of TMR system under increased Process Variations. Same functional modules have spread in their worst-case delays. Overall delay of the system is determined by Eq. 1.

circumstances, such as autonomous vehicles, satellites, and deep space systems [10],[11],[12]. It is also been employed in commercial systems such as High-Performance Computing applications [13] where significant increase in compute node availability is sought. Here, use of compute-node level redundancy at the processor, memory module, and network interface can improve reliability by a factor of 100-fold to 100,000-fold.

As depicted in Figure 1, spatial redundancy involves replicating N instances of a circuit module and obtaining the majority output via voting element. Hence, N is typically chosen to be an odd number to preclude outcomes which are ties. The output of a spatial redundancy arrangement can be considered to be correct whenever the majority of the instances produce identical and valid outputs. Identical yet invalid outputs in an NMR system with a multiple-bit word output require the transient(s) to impact distinct NMR instances at the corresponding functional locations to manifest identically incorrect outputs. In the case of an isolated Single Event Transient (SET) in an NMR system during a computation interval, the resultant soft error is masked. However, if more than one bit is upset then a Multi-Bit Upset (MBU) results. Spatial MBUs occur when a single particle upsets multiple bits which reside within the same physical neighborhood. Temporal MBUs occur when two or more particle strikes independently upset distinct NMR instances. MBUs may still generate a diagnosable error from an NMR word-wise voted output. In such scenarios, word-wise voting can be advantageous compared to bit-by-bit voting [14]. Finally, even though spatial MBU feasibility has increased due to technology scaling, non-planar devices offer a means to reduce SER. For example, 22nm Tri-Gate technology is shown to reduce neutron and alpha-particle induced SER at nominal voltage on the order of 1.5-fold to 4-fold and in excess of 10-fold respectively, compared to a 32nm planar process [15].

Under nominal operating conditions, the energy consumption of NMR systems is about N-fold as compared to simplex ($N = 1$) systems which lack soft error masking capability. This paper explores the tradeoffs of operating NMR systems at NTV beyond processor caches where a low-complexity means for improved resilience compared to error correcting schemes has been sought [16],[17]. For cache memories, *Orthogonal*

Latin Square Codes (OLSCs) have been employed to encode orthogonal groups of checking bits without syndrome generation, yet enable recovery with majority voting, and further extensions to *Variable-Strength Error Correcting Codes (VS-ECCs)* have been employed which combine the use of ECC and memory tests to ensure reliable cache operation under aggressive voltage scaling [18]. Herein, we concentrate on NMR for logic paths as opposed to memory elements.

IV. NMR SYSTEMS AT NEAR-THRESHOLD VOLTAGE

Nanoscale devices are susceptible to process variations created by precision limitations of the manufacturing process. Phenomena such as *Random Dopant Fluctuations (RDF)* and *Line-Edge Roughness (LER)* are major causes of such variations in CMOS devices [19]. The increased occurrence of PV results in a distribution of threshold voltage V_{th}. As the V_{th} increases, the increase in switching time affects the delay performance of the circuit. Such variability is observed to become magnified by continued scaling of process technology node [19]. For example, the effect of RDF is magnified as number of dopant atoms is fewer in scaled devices such that the addition or deletion of just a few dopant atoms significantly alters transistor properties. In addition, a large impact in circuit performance occurs as the transistor on-current is highly variable near the threshold region [3]. Recent approaches for dealing with increased PV at NTV in multicore devices include leveraging the application's inherent tolerance for faults through performance-aware task-to-core assignment based on problem size [20]. Variation impacts at NTV on cache reliability have also been developed to leverage adaptive methods to dynamically adjust error control strength [21]. In the remainder of the paper, we restrict our discussion to show how these PV effects combine in NMR systems of logic datapaths to exhibit a higher mean delay than simplex systems. Stated alternatively, NMR arrangements require a more-than-linear increase in energy in order to obtain a delay which is comparable to its component module.

While operation at NTV can be seen to increase PV by approximately 5-fold as quantified in [1] for simplex arrangements, the effect in NMR systems has not previously been investigated. In the case of an NMR arrangement, it can be expected that the worst-case delay will exceed that of any single module. For instance, the delay of TMR system shown in Figure 1 is 2.8ns, which is determined by the worst-case delay out of all instances. Generally, if the worst-case delay of an instance i of NMR system is τ_i, then the overall delay of NMR system τ_{NMR} is given by:

$$\tau_{NMR} = \max_{1 \leq i \leq N} (\tau_i) + \delta \qquad (1)$$

where δ represents the delay of the voting logic, which contributes directly to the critical delay. Furthermore, the chance of having an instance with higher than average delay increases with N, which has been validated through experimental results quantified in Section VI. Overall, these results are in agreement with distributions of 128-wide SIMD architectures demonstrated in [22], whereby the speed of the overall architecture is also determined by the slowest SIMD lane. Herein, we focus on the performance of NMR systems as compared to simplex systems, with $3 \leq N \leq 5$.
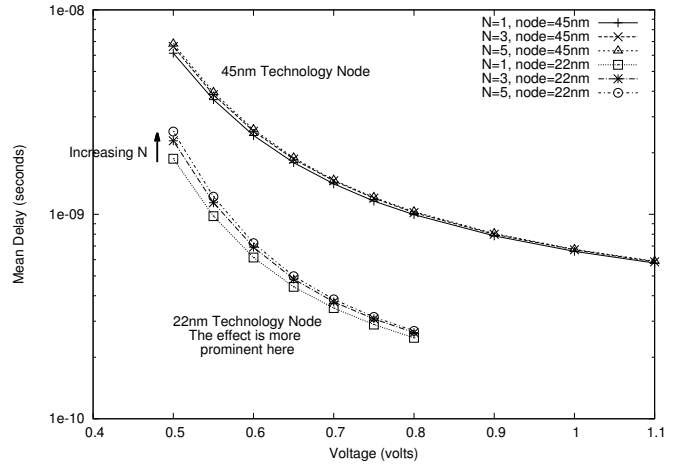

Fig. 2. Mean Delay (each point is obtained by averaging at least 1,000 samples) of NMR systems increases with scaling voltage down to the near-threshold region

Intra-die variations for both 22nm and 45nm technology nodes are simulated using the Monte-Carlo method in HSPICE. A viable alternative approach for simulating PV at NTV is also proposed in [23], which captures both the systematic effects due to lithographic irregularities and the localized variations due to RDF. In the current paper, each module in an NMR arrangement can be anticipated to exhibit comparable spatial variability due to the relative proximity of its module instances in the physical layout. Thus, the scope of this work focuses on random variation impacts while die-to-die variations would comprise future work.

The random effects are modeled through the variation in V_{th} caused due to RDF and LER effects. The standard deviation σV_{th} values are adopted from [19] which range from 25.9mV to 59.9mV for 45nm process and 22nm process, respectively. Figure 2 shows the mean delay for inverter chain for commonly-used values of N. It is observed that the performance impact for 45nm technology node is around $10.6X$ on average at $0.5V$ for simplex system. However, it reduces to $6.29X$ when the voltage is increased to $0.55V$. The performance impact for NMR systems with $N = 3$ and $N = 5$ tend to follow the same behavior. This is in agreement with the near-threshold performance results noted in [1].

Results in Figure 2 indicate that the mean delay is slightly higher for NMR systems and tends to increase with N. The spread in mean delays between simplex and NMR systems also increases with increased variability effects notable at NTV as shown in Figure 4. However, more values are clustered closer to the mean when N is increased. This is observed through delay distribution in Figure 3 for $V_{DD} = 0.55V$ whereby the mean values are increased causing right-shifted peaks for $N = 3$ and $N = 5$ compared to simplex while the variances are decreased causing narrower spreads. In addition, the performance difference between simplex and NMR systems at 22nm technology node is magnified due to higher values for the coefficient of variation for 22nm technology node as compared to the 45nm technology node as noted in Figure 4. For instance, the mean delays for the 22nm node with $N = 3$ and $N = 5$ are $1.16X$ and $1.24X$ the mean
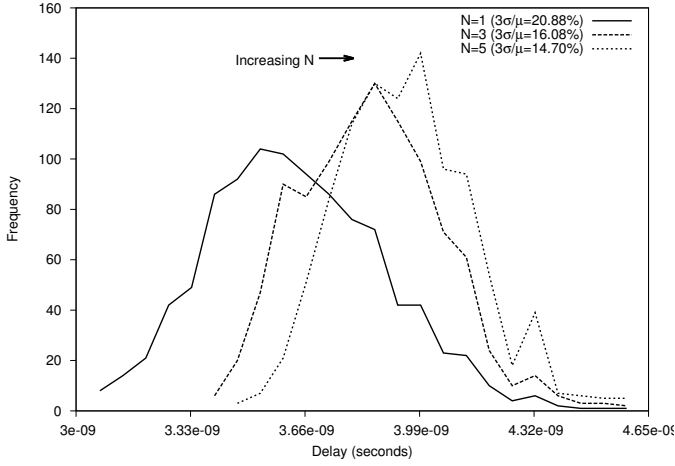
Fig. 3. Delay Distributions (1,000 samples) of NMR systems at Near-threshold Voltage of $0.55V$ with 45nm technology node

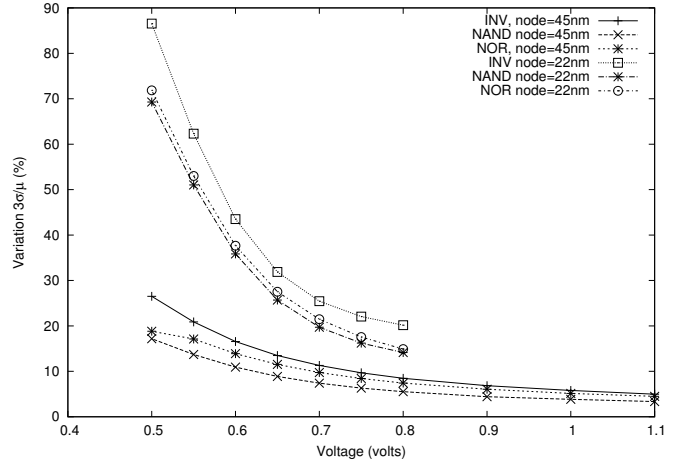Fig. 4. Delay variations (over 1,000 samples) decrease with increasing N for NMR systems

Fig. 5. The variation of simplex systems ($N = 1$) composed of different types of logic gates. Inverter chains based on NAND gates exhibit the least amount of variation at significant area and delay costs.
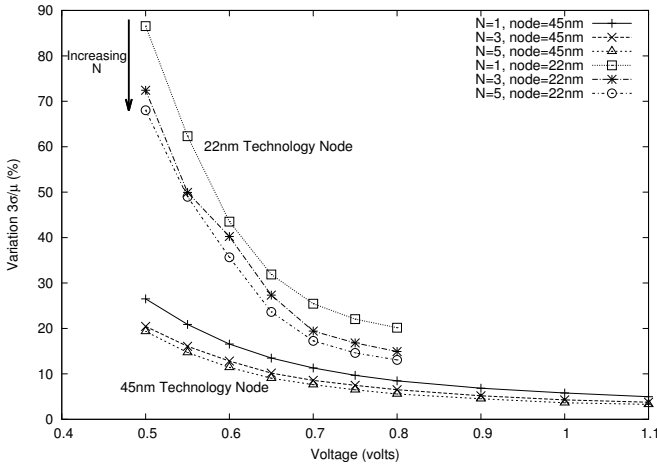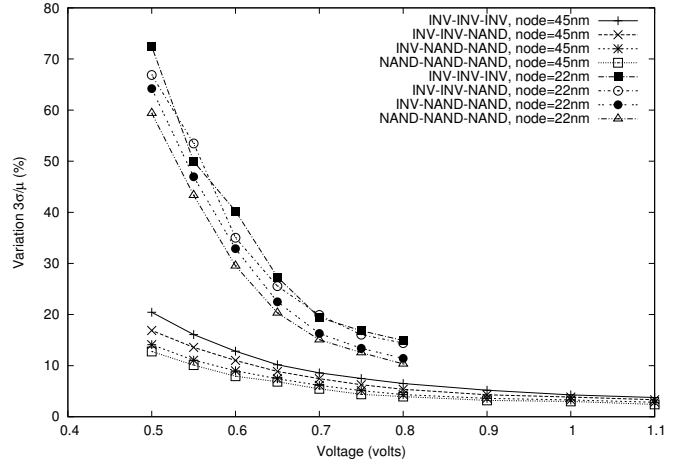

Fig. 6. The variation of TMR systems composed of modules with inverter chains synthesized with different types of logic gates. Here, 'INV-NAND-NAND' is a TMR arrangement with modules having same function, yet realized through diverse logic gates, such as INV and NAND gates.

delay for a simplex system respectively at the same voltage of $0.55V$. Whereas, in contrast, for 45nm technology node the difference is only $1.06X$ and $1.09X$ for $N = 3$ and $N = 5$ systems respectively. These numbers tend to be higher when operating very close to the V_{th} of the transistors due to increased variations.

Furthermore, the amount of variability is also dependent on the length of the logic datapath, i.e, the number of gates in the critical path. For instance, it is noted in [22] that the variability decreases as the length of inverter chain increases. However, as pointed out earlier, logic datapaths operating at NTV may be structured to have relatively smaller depths. Herein, alternative synthesis techniques are demonstrated which can lower the amount of variability at NTV. For instance, it is observed that the variation is also dependent on the type of logic gate utilized. Figure 5 shows that functionally-identical inverter chains built using NAND2 gates exhibit the least amount of variation, having inputs tied together to realize the same function as an INV gate. Furthermore, various TMR arrangements are considered utilizing diversity of NAND2 and INV gates in Figure 6. Again, all of these TMR arrangements

are functionally equivalent, yet exhibit different amount of variability. For example, 22nm TMR arrangements based on NAND2 gates exhibit about 13% less variation as compared to a TMR arrangement with INV gates. Now, for our experiments with the inverter chain, the mean delays for NAND-based systems are higher than INV-based systems which outweighs any benefit of reduced variation. Thus, a diversity-enabled NMR synthesis approach needs to be evaluated with more functionally-complex benchmark circuits which will be pursued as future work. Here, we restrict our PV analysis to uniform logic gates which achieve minimal delay.

Since the spread of delay distribution for NMR systems is narrower as compared to simplex systems, straightforward techniques to combat performance variability [1] can be used for NMR systems. For instance, smaller guard-bands are required to maintain same yield as compared to simplex systems as demonstrated in the results of this work.
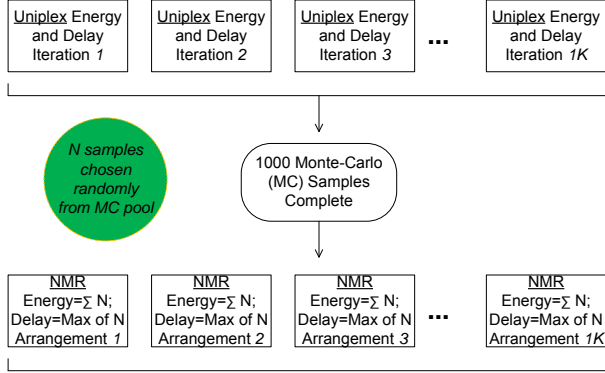
Fig. 7. Simulation framework developed to estimate the delay and energy for NMR systems. At least 1,000 samples are synthesized for each simplex, TMR, and 5MR system to conduct the statistical analysis.

V. Energy Cost of Mitigating Variability in NMR Arrangements

A fundamental approach to alleviate observed delay variations in the near-threshold region is to add one-time *timing guard-bands* [22]. Guard-bands can be realized by operating at reduced frequency via a longer clock period, or by operating at a slightly elevated voltage to compensate for the increased delay variations. This work assumes the latter for comparison against simplex systems. It is worth mentioning that the scope of this paper is to deal with soft errors at NTV using spatial redundancy, thus a straightforward voltage margining scheme is utilized to achieve same performance as a simplex system in presence of variations.

To achieve an expected yield of approximately 99% for NMR systems while achieving the same worst-case performance at a fixed NTV of simplex system; the voltage for NMR system is increased such that the respective delay distributions have the following statistical characteristics for $N \geq 3$:

$$(\mu_{NMR} + 3 * \sigma_{NMR}) \leq (\mu_{Simplex} + 3 * \sigma_{Simplex}) \quad (2)$$

where μ_{NMR}, $\mu_{Simplex}$ represent the mean delays for NMR and simplex systems, respectively, and σ_{NMR}, $\sigma_{Simplex}$ represent the respective standard deviations. The three sigma rule has the property that nearly all (99.7%) of the instances have the delay less than $(\mu_{NMR} + 3 * \sigma_{NMR})$. Thus, this results in a high expectation for an NMR system to have same worst-case delay as simplex system, i.e., the same throughput performance.

VI. Experiments and Results

Experiments are conducted to quantify the energy overheads of NMR systems with increased variations due to operation in near-threshold region. For this case study, MCNC benchmark circuits of C880 and i5 are utilized. These circuits are synthesized using Synopsis Design Compiler based on the 45nm PTM-based NanGate open source library [24]. Then, the synthesized netlist is imported into Synopsis HSPICE tool for Monte-Carlo simulations. These simulations vary the V_{th} of the transistors in the netlist based on a gaussian distribution having a mean equal to the nominal model card for PTM and

	C880		i5	
N = 1, V_{DD} (NTV)	**N=3**	**N=5**	**N=3**	**N=5**
$0.55V$	3.03X	5.06X	3.02X	5.05X
$0.6V$	3.03X	5.05X	3.02X	5.04X
$0.65V$	3.02X	5.04X	3.02X	5.04X
$0.7V$	3.01X	5.03X	3.01X	5.02X

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.

If the overhead of the voter circuit is considered to be negligible, i.e. $\delta = 0$, then direct comparisons to simplex systems are possible. The Monte-Carlo simulations were conducted to utilize at least 1,000 experimental runs for a single module within the NMR circuit. Then, the NMR system is obtained by choosing N random samples from pool of module variants created by these runs. The module instance with maximum delay determines the delay of the NMR system indicated by Eq 1. Similarly, the energy consumption is computed by accumulating the energy requirement of the selected N samples operating at a frequency of $1/\tau_{NMR}$. This scenario is repeated 1,000 times to establish mean values for comparison to those obtained in an equivalent number of runs for a simplex system. This simulation framework is illustrated in Figure 7.

A. Iso-Performance Energy Consumption for NMR

Table I shows the mean energy consumptions at multiple near-threshold voltages for TMR and 5MR systems such that they have the same worst-case performance as a simplex system with yields of at least 99.7%. This is done by ensuring that the three sigma point for delay distribution of NMR system is less or equal to that of the simplex system as described earlier. For comparison purposes, the operating voltage of simplex system is used as a reference with the values shown in the Table I. Then, the operating voltage for the NMR system is elevated from the reference Near-Threshold Voltage (NTV) assumed for simplex system. Note, all the modules in this case are assumed to be operated at uniform voltage and are co-located within the same region of the chip. The elevated supply voltage results in a left-shift of the delay distribution towards that of the simplex system. It is seen that only a slight voltage increase is sufficient for this purpose. For example, on average $2mV$ increase is satisfactory to operate a TMR arrangement based on i5 circuit at comparable delay. This translates to a mean energy consumption of about $3.02X$ which is approximately 1% more than $3X$ assumed at nominal conditions.

The NMR energy consumptions as shown in Table I are not excessive because NMR systems tend to have lower variation (σ) as compared to simplex systems as demonstrated in Figure 4. Hence, even though NMR systems exhibit higher mean delays (μ), their reduced variance necessitates only a slight increase in reference voltage to meet the delay target of the simplex system. However, increasing the value of N to 5 increases the energy consumption slightly due to more increase in mean delay as compared to a TMR system.

	45nm		22nm	
N = 1, $\mathbf{V_{DD}}$ (NTV)	**N=3**	**N=5**	**N=3**	**N=5**
$0.5V$	3.04X	5.07X	3.17X	5.30X
$0.55V$	3.03X	5.05X	3.14X	5.27X
$0.6V$	3.03X	5.04X	3.13X	5.26X
$0.65V$	3.02X	5.03X	3.12X	5.23X
$0.7V$	3.01X	5.02X	3.10X	5.16X

B. Impact of Technology Scaling

To further analyze the impact of increased variability on energy overheads of NMR systems, experiments were conducted next with scaled technology node of 22nm PTM HP model. These experiments are conducted on inverter chains composed of 26 Fanout-of-4 inverters, which is a similar setup as adopted in [22]. The energy consumptions with same goals as defined earlier are listed in Table II as compared to 45nm-based technology node. The results indicate that the energy overheads are higher for both TMR and 5MR systems at more deeply scaled technology nodes. For instance, the 22nm node-based 5MR system requires 3.94% (on average) more energy consumption than a 45nm based 5MR system. This is consistent with the trend of increased variations as noted in Figure 4 beginning with the 22nm node.

C. Cost of Increased Reliability at NTV

Near-threshold operation allows improved energy efficiency. The energy savings can be utilized for either reduced power operation or to increase resilience via NMR. Thus, operation of NMR systems in the near-threshold region allows for consideration of interesting tradeoffs. For example, increasing N from 1 to 3 can be evaluated as a means to increase reliability within the same energy budget as a simplex system operating at nominal voltage. This is valid provided that the increase in delay, and thus corresponding drop in performance, is acceptable. Note that this pursuit of increased reliability is predicated upon the assumption that the source of variability in the near-threshold region is due to variation in V_{th} for which this work is restricted. Further study needs to be performed to determine the reliability levels provided by NMR systems in this operating region due to other noise sources such as variation in V_{DD} [25], inductive noise and temperature [26].

Based on these assumptions, we point out these tradeoffs in Figure 8. For instance, it shows the feasibility of TMR operation at approximately $0.69V$ (operating point B on plot) on average given an identical energy budget of a simplex system operating at nominal voltage of $1.1V$ (operating point A on plot) while incurring a delay difference of $2.58X$. Similarly, it is possible to achieve 5MR with approximately $0.545V$ (operating point C on plot) operation on average while incurring a performance impact of $7.15X$. This represents a greater-than-linear increase in delay as a function of N, as compared to TMR operation. Thus for mission-critical applications, this offers insights into tuning the degree of redundancy facilitated by a near-threshold computing paradigm.

To consider the feasibility of reducing energy consumption while simultaneously providing soft error masking, we again observe Figure 8. The energy requirement of a simplex arrangement at 1.1V is about 0.541 pJ, while the TMR curve
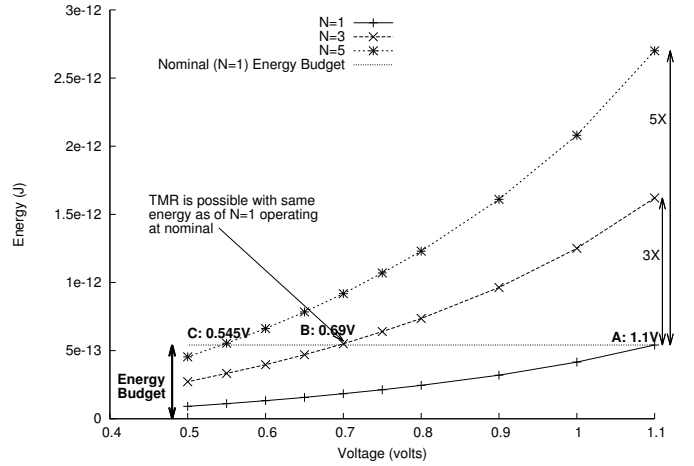

Fig. 8.   Mean Energy (each point is obtained by averaging at least 1,000 samples) of NMR systems at 45nm technology node. Operation at NTV corresponds to tradeoffs between performance and energy consumption required for soft error masking.

at 550mV is only about 0.330 pJ. Thus, a TMR arrangement at 550mV results in an energy savings of 38.4% as compared to simplex system at 1.1V. This means compared to a simplex arrangement at nominal voltage, selecting a supply voltage of 550mV allows for provision of TMR for soft error masking in the presence of technology scaling while still reducing the energy requirement significantly.

VII.   CONCLUSION

Operation of NMR systems in the near-threshold region allows for consideration of energy and delay tradeoffs in terms of N. Redundancy can be seen as a degree of freedom enabled by decreasing supply voltage. When doing so, it is essential to consider NMR arrangements' increased susceptibility to PV.

Further study is worthwhile to determine the reliability provided by NMR systems in this operating region due to other noise sources such as variation in supply voltage V_{DD}, temperature, and aging-induced variations. For instance, it is pointed out that a further variation of 2X is expected due to variation of V_{DD} and temperature. However, lower voltage helps to lower transistor junction temperatures and interconnect currents which can have beneficial effects on aging-induced defects such as Electromigration and Bias Temperature Instability. Furthermore, augmenting the use of temporal redundancy in the form of timing-speculation circuits to reduce variation effects and spatial redundancy would also be worthwhile to investigate.

REFERENCES

[1]   R. Dreslinski, M. Wieckowski, D. Blaauw, D. Sylvester, and T. Mudge, "Near-threshold computing: Reclaiming moore's law through energy efficient integrated circuits," *Proceedings of the IEEE*, vol. 98, no. 2, pp. 253–266, 2010.

[2]   E. Krimer, R. Pawlowski, M. Erez, and P. Chiang, "Synctium: a near-threshold stream processor for energy-constrained parallel applications," *IEEE Computer Architecture Letters*, vol. 9, no. 1, pp. 21–24, 2010.

[3]   H. Kaul, M. Anders, S. Hsu, A. Agarwal, R. Krishnamurthy, and S. Borkar, "Near-threshold voltage (NTV) design: Opportunities and challenges," in *Proceedings of the 49th ACM/EDAC/IEEE Annual Design Automation Conference (DAC)*, June 2012, pp. 1153–1158.

[4] A. Dixit and A. Wood, "The impact of new technology on soft error rates," in *IEEE International Reliability Physics Symposium (IRPS)*, April 2011, pp. 5B.4.1–5B.4.7.

[5] L. Szafaryn, B. Meyer, and K. Skadron, "Evaluating overheads of multibit soft-error protection in the processor core," *IEEE Micro*, vol. 33, no. 4, pp. 56–65, July 2013.

[6] P. Shivakumar, M. Kistler, S. Keckler, D. Burger, and L. Alvisi, "Modeling the effect of technology trends on the soft error rate of combinational logic," in *Proceedings of the International Conference on Dependable Systems and Networks (DSN)*, Jun 2002, pp. 389–398.

[7] M. Snir, R. W. Wisniewski, J. A. Abraham, S. V. Adve, S. Bagchi, P. Balaji, J. Belak, P. Bose, F. Cappello, B. Carlson, A. A. Chien, P. Coteus, N. A. Debardeleben, P. Diniz, C. Engelmann, M. Erez, S. Fazzari, A. Geist, R. Gupta, F. Johnson, S. Krishnamoorthy, S. Leyffer, D. Liberty, S. Mitra, T. S. Munson, R. Schreiber, J. Stearley, and E. V. Hensbergen, "Addressing failures in exascale computing," *International Journal of High Performance Computing*, vol. ANL/MCS-P5022-0913, Mar 2013.

[8] Q. Zhou and K. Mohanram, "Gate sizing to radiation harden combinational logic," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 25, no. 1, pp. 155–166, Jan 2006.

[9] K.-C. Wu and D. Marculescu, "Power-aware soft error hardening via selective voltage scaling," in *Proceedings of the IEEE International Conference on Computer Design (ICCD)*, Oct 2008, pp. 301–306.

[10] J. Celis, S. De La Rosa Nieves, C. Fuentes, S. Gutierrez, and A. Saenz-Otero, "Methodology for designing highly reliable fault tolerance space systems based on COTS devices," in *Proceedings of the IEEE International Systems Conference (SysCon)*, 2013, pp. 591–594.

[11] M. Pignol, "How to cope with SEU/SET at system level?" in *Proceedings of the 11th IEEE International On-Line Testing Symposium (IOLTS)*, Jul 2005, pp. 315–318.

[12] R. Al-Haddad, R. Oreifej, R. Ashraf, and R. F. DeMara, "Sustainable modular adaptive redundancy technique emphasizing partial reconfiguration for reduced power consumption," *International Journal of Reconfigurable Computing*, vol. 2011, no. 430808, 2011.

[13] C. Engelmann, H. Ong, and S. L. Scott, "The case for modular redundancy in large-scale high performance computing systems," in *Proceedings of the IASTED International Conference on Parallel and Distributed Computing and Networks (PDCN)*, vol. 641, Feb 2009, pp. 189–194.

[14] S. Mitra and E. McCluskey, "Word-voter: a new voter design for triple modular redundant systems," in *Proceedings of the 18th IEEE VLSI Test Symposium (VTS)*, May 2000, pp. 465–470.

[15] N. Seifert, B. Gill, S. Jahinuzzaman, J. Basile, V. Ambrose, Q. Shi, R. Allmon, and A. Bramnik, "Soft error susceptibilities of 22 nm tri-gate devices," *IEEE Transactions on Nuclear Science*, vol. 59, no. 6, pp. 2666–2673, Dec 2012.

[16] A. Seyedi, G. Yalcin, O. S. Unsal, and A. Cristal, "Circuit design of a novel adaptable and reliable L1 data cache," in *Proceedings of the 23rd ACM International Conference on Great Lakes Symposium on VLSI (GLSVLSI)*, 2013, pp. 333–334.

[17] Y. Choi, S. Yoo, S. Lee, J. H. Ahn, and K. Lee, "MAEPER: Matching access and error patterns with error-free resource for low vcc L1 cache," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 21, no. 6, pp. 1013–1026, June 2013.

[18] C. Wilkerson, A. Alameldeen, and Z. Chishti, "Scaling the memory reliability wall," *Intel Technology Journal*, vol. 17, no. 1, pp. 18–34, May 2013.

[19] Y. Ye, T. Liu, M. Chen, S. Nassif, and Y. Cao, "Statistical modeling and simulation of threshold variation under random dopant fluctuations and line-edge roughness," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 19, no. 6, pp. 987–996, Jun 2011.

[20] U. R. Karpuzcu, I. Akturk, and N. S. Kim, "Accordion: Toward Soft Near-Threshold Computing," in *Proceedings of the 20th IEEE International Symposium on High Performance Computer Architecture (HPCA)*, February 2014.

[21] C. Li, M. Zhang, and P. Ampadu, "Reliable ultra-low voltage cache with variation-tolerance," in *Proceedings of the IEEE 56th International Midwest Symposium on Circuits and Systems (MWSCAS)*, Aug 2013, pp. 121–124.

[22] S. Seo, R. Dreslinski, M. Woh, Y. Park, C. Charkrabari, S. Mahlke, D. Blaauw, and T. Mudge, "Process variation in near-threshold wide SIMD architectures," in *Proceedings of the 49th ACM/EDAC/IEEE Design Automation Conference (DAC)*, June 2012, pp. 980–987.

[23] U. R. Karpuzcu, K. B. Kolluru, N. S. Kim, and J. Torrellas, "Variusntv: A microarchitectural model to capture the increased sensitivity of manycores to process variations at near-threshold voltages," in *Proceedings of the 42nd Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, June 2012, pp. 1–11.

[24] W. Zhao and Y. Cao, "New generation of predictive technology model for sub-45nm design exploration," in *Proceedings of the 7th International Symposium on Quality Electronic Design (ISQED)*, March 2006, pp. 585–590.

[25] T. Miller, R. Thomas, X. Pan, and R. Teodorescu, "VRSync: Characterizing and eliminating synchronization-induced voltage emergencies in many-core processors," in *39th Annual International Symposium on Computer Architecture (ISCA)*, June 2012, pp. 249–260.

[26] Y. Kim, L. K. John, I. Paul, S. Manne, and M. Schulte, "Performance boosting under reliability and power constraints," in *Proceedings of the International Conference on Computer-Aided Design (ICCAD)*, Nov 2013, pp. 334–341.