Adaptive Fault-Tolerant Architecture for Unreliable Technologies With Heterogeneous Variability

Nivard Aymerich, Sorin D. Cotofana, *Senior Member, IEEE*, and Antonio Rubio, *Senior Member, IEEE*

***Abstract*—This paper introduces an efficient adaptive redundant architecture, which makes use of the averaging cell (AVG) principle in order to improve the reliability of nanoscale circuits and systems. We propose an adaptive structure that is able to cope with nonhomogeneous variability and time-varying effects like degradation and external aggressions, which are expected to be key limiting factors in future technologies. First, we consider static heterogeneity of the input variability levels and derive a methodology to determine the weight values that maximize the reliability of the averaging system. The implementation of these optimal weights in the AVG gives place to the unbalanced AVG structure (U-AVG). Second, we take into consideration that circuits are exposed to time-dependent aggression factors, which can induce significant changes on the levels of variability, and introduce the adaptive AVG structure (AD-AVG). It embeds a learning mechanism based on a variability monitor that allows for the on-line input weight adaptation such that the actual weight configuration properly reflects the aging status. To evaluate the potential implications of our proposal, we compare the conventional AVG architecture with the unbalanced (U-AVG) and the adaptive (AD-AVG) approaches in terms of reliability and redundancy overhead by means of Monte Carlo simulations. Our results indicate that when AVG and U-AVG are exposed to the same static heterogeneous variability, U-AVG requires 4× less redundancy for the same reliability target. Subsequently, we include temporal variation of input drifts in the simulations to reproduce the effects of aging and external aggressions and compare the AVG structures. Our experiments suggest that AD-AVG always provides the maximum reliability and the highest tolerance against degradation. We also analyze the impact of nonideal variability monitor on the effectiveness of the AD-AVG behavior. Finally, specific reconfigurable hardware based on resistive switching crossbar structures is proposed for the implementation of AD-AVG.**

***Index Terms*—Averaging cell (AVG), degradation, fault tolerance, hardware redundancy, nanoscale technology, reliability.**

I. Introduction

COMPUTER architecture constitutes one of the key and strategic application fields for new emerging devices at nanoscale dimensions, potentially getting benefit from the expected high-component density and speed. However, these future technologies are expected to suffer from a reduced device quality, exhibiting a high level of process and environmental variations as well as performance degradation due to the high stress of materials [1]–[4]. This clearly indicates that if we are to make use of those novel devices, we have to rely on fault-tolerant architectures.

The necessity of producing reliable systems from unreliable devices was addressed for the first time by Neumann [5] in the 1950s. He proposed different circuit architectures based on hardware redundancy and demonstrated that it is possible to perform computation with an arbitrary level of reliability despite the use of unreliable devices. Since then, these fault-tolerant architectures have been refined and combined in many different ways in order to improve the reliability while minimizing the cost overhead. To date, three major groups of fault-tolerant architectures based on hardware redundancy have been proposed and studied: 1) multiplexing techniques, which rely on the replication of devices and interconnects [6]–[8], 2) reconfiguration techniques, which consist of redundant structures capable of detecting, locating, and avoiding manufacturing defects by using different hardware parts in different ways [9], [10], and 3) voting-based techniques, which implement the idea of modular redundancy and obtain the result by a voting process [11], [12].

Focusing on voting-based fault-tolerant systems, we observe that 1) most of the approaches make use of majority gates (MAJ) as decision operation [12], [13] and 2) the granularity at which the voting-based strategies are applied span from device and gate levels to more complex functional blocks. However, generally speaking, there are two possible ways to implement the voting process: 1) majority voting (digital) [14] and 2) averaging voting (analog) [15]. Given that the main parameter to be optimized after securing the reliability requirement is the overhead cost, the analog-based voting systems may have significant advantages over the digital one, since they can provide an effective means for absorbing faults [16]. A typical implementation for the analog voting is the averaging cell (AVG), which has been demonstrated to provide higher reliability than majority voting at a lower cost [16]–[18]. AVG underlying principle is to average several input replicas in order to compute the most probable output value. This approach is quite effective in case the AVG inputs are subject to independent drifts with the same/similar magnitude. In practice, however, input deviations can be nonhomogenous, the case in which the balanced

Manuscript received October 6, 2011; revised April 18, 2012; accepted May 7, 2012. Date of publication May 16, 2012; date of current version July 11, 2012. This work was supported by the Spanish Ministry of Science and Innovation through the Project TEC2008-01856 with additional participation of FEDER founds and the European TRAMS Project FP7 248789. The review of this paper was arranged by Associate Editor F. Lombardi.

N. Aymerich and A. Rubio are with the High Performance IC Design Group, Department of Electronic Engineering, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain (e-mail: nivard.aymerich@upc.edu; antonio.rubio@upc.edu).

S. D. Cotofana is with the Computer Engineering Group, Faculty of Electrical Engineering, Computer Science and Mathematics, Delft University of Technology, 2600 Delft, The Netherlands (e-mail: S.D.Cotofana@ewi.tudelft.nl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNANO.2012.2199513

average cannot provide a response that minimizes the output error probability.

In this paper, we focus on AVG structures and modify them adding a reconfiguration capability with the objective of improving their design such that they become more robust and tolerant to nonuniform variability scenarios typical for new-technology generations [19]. In this line of reasoning, we propose two consecutive steps to improve the AVG design and adapt it to time-varying heterogeneous variability scenarios. First, we consider the heterogeneity of the input variability levels. Given that different replicas present in the AVG structure may experience different levels of variability/reliability, we propose an AVG that makes use of a weighted average with unbalanced weights (U-AVG). Regarding this approach, we derive a methodology to determine the weight values that maximize the reliability of the AVG output. A description of this methodology can be found in [20]. Second, we take into consideration the fact that, when deployed in the field, circuits are exposed to time-dependent aggression factors, which induce temporal variations in the levels of variability. The aging and external aggressions can gradually change the input variability levels and this may result in significant differences between the input variabilities considered at unbalanced AVG (U-AVG) design time and the ones exhibited by its inputs at maturity. To be able to minimize the effects of such nonuniform aging, we introduce an adaptive AVG (AD-AVG), which embeds a learning mechanism that allows for the on-line input weight adaptation such that the actual weight configuration properly reflects the aging status. The proposed AVG adaptation mechanism relies on 1) obtaining the input levels of variability, by means of a variability monitor, 2) calculating the new weight values that maximize the reliability of the AVG output for the current variability scenario, and 3) updating the AVG structure often enough in order to keep the pace with the degradation dynamics.

To evaluate the potential practical implications of our proposal, we compare the basic AVG, the U-AVG, and the AD-AVG in terms of reliability and area overhead by means of Monte Carlo simulations as follows. We first expose AVG and U-AVG structures to static heterogeneous input variability and observe that for the same targeted reliability U-AVG requires $4\times$ less redundancy, thus by implication of substantially less area overhead. Subsequently, we include the temporal variation of input drifts in our simulations to reproduce the effects of aging and external aggressions and compare the static AVG (AVG and U-AVG) with the AD-AVG. Our experiments suggest that for the same redundancy overhead, the AD-AVG always provides the maximum reliability and tolerates larger amounts of accumulated degradation. Finally, we investigate the impact that a nonideal variability monitor may have on the effectiveness of the AD-AVG learning capability.

The paper is organized as follows: In Section II, we present the AVG architecture and give some details of its main features. In Section III, we introduce the idea of U-AVG, derive the analytic expression of the optimal set of weights, and simulate and compare the standard cell with the unbalanced weights approach. In Section IV, we propose the adaptive technique AD-AVG and evaluate its behavior in a scenario of changing heterogeneous input variability. Finally, in Section V, we propose a possible implementation of the adaptive technique AD-AVG and demonstrate how the weight reconfiguration and input variability measurement can be satisfied.

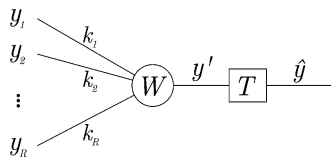

Fig. 1. AVG organization.

II. AVG

The AVG, widely known for its application in the four-layer reliable hardware architecture [18] stems from the perceptron, the McCulloch–Pitts neuron model [21], [22]. Associated with fault-tolerant techniques based on redundancy, the AVG graphically depicted in Fig. 1 can calculate the most probable value of a binary variable from a set of error-prone physical replicas. While the MAJ-based voting technique operates in the digital domain, the AVG performs a weighted average of the replicated inputs in the analog domain, and thus is potentially more robust.

The AVG output $\hat{y}$ is an estimation of the ideal variable y according to

$$y' = W(y_1, \ldots, y_R) = \frac{1}{\sum_{i=1}^{R} k_i} \sum_{i=1}^{R} k_i y_i \tag{1}$$

$$\hat{y} = T(y') = \begin{cases} V_{cc} & \text{if } y' \geq V_{cc}/2 \\ 0 & \text{if } y' < V_{cc}/2 \end{cases} \tag{2}$$

where y_i are R replicas of the ideal input variable y:

$$y_i = y + \eta_i \qquad i = 1, \ldots, R \tag{3}$$

and each of the input replicas y_i has associated an independent drift η_i that modifies its ideal value y. As a consequence, input signals y_i are observed in the system as continuous voltage levels, where 0 and V_{cc} stand for ideal logical values "0" and "1," respectively.

Without loss of generality, in this work, we use $V_{cc} = 1\,\text{V}$ throughout all the simulations. We also assume positive magnitudes for all the averaging weights $k_i \geq 0$, $i = 1, \ldots, R$. In order to simplify the mathematical formulation, we use normalized weights $c_i = k_i / \sum_{j=1}^{R} k_j$ instead of k_i. We model drift magnitudes as the Gaussian random variables with null mean $\eta_i \sim N(0, \sigma_i)$. We consider them as mutually independent; thus, our model explores random effects between replicas. Systematic alterations, which affect all the replicas in the same direction, are evaluated in the Appendix. They are shown to cause independent effects on reliability; random and systematic effects can be studied and counteracted independently. Moreover, systematic effects are usually static and they can be mitigated by calibration in the initial stages of system operation.

When y' is processed by the threshold operation $T(y')$, an error is produced if and only if the deviation in the weighted

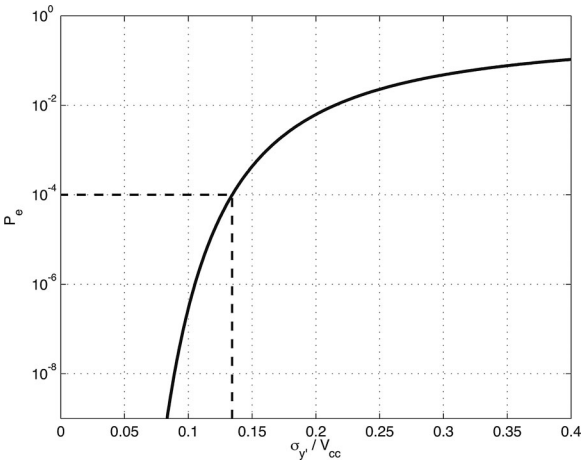

Fig. 2. Output error probability P_e against the ratio between the weighted average standard deviation $\sigma_{y'}$ and V_{cc}. The highlighted point corresponds to the reliability reference value $P_e^{\max} \equiv 10^{-4}$ and the maximum admissible variability level $\sigma_{\max}/V_{cc} = 0.1344$.

average $\epsilon = y' - y$ reaches $V_{cc}/2$ or $-V_{cc}/2$, depending on the logic value y. Since this deviation parameter ϵ can be expressed as a linear combination of normally distributed variables η_i, by the properties of the normal distribution, the probability density function $f_\epsilon(\epsilon)$ can be described as a normal distribution with parameters

$$\mu_\epsilon = E\left\{\sum_{i=1}^{R} c_i y_i\right\} - y = 0 \tag{4}$$

$$\sigma_\epsilon^2 = E\left\{(y' - y)^2\right\} = \sigma_{y'}^2. \tag{5}$$

And the variance of the weighted average $\sigma_{y'}^2$ can be expressed in terms of the input variances σ_i^2 and the averaging weights c_i, $i = 1, \ldots, R$

$$\sigma_{y'}^2 = \sum_{i=1}^{R} c_i^2 \sigma_i^2. \tag{6}$$

Thus, using the complementary Gauss error function, we can analytically formulate the output error probability as follows

$$P_e = \int_{V_{cc}/2}^{\infty} f_\epsilon(\epsilon) d\epsilon = \frac{1}{2} \times \operatorname{erfc}\left(\frac{V_{cc}}{\sqrt{8}\sigma_{y'}}\right). \tag{7}$$

Fig. 2 depicts the relationship between P_e and the ratio $\sigma_{y'}/V_{cc}$. It presents a monotonically increasing behavior. Thus, given a reliability requirement $P_e^{\max}$, there is a maximum admissible output standard deviation $\sigma_{\max}$ for any given V_{cc}. In this paper, we take $P_e^{\max} \equiv 10^{-4}$ as the reference value for the reliability requirement. Therefore, the maximum admissible weighted average standard deviation is $\sigma_{\max} = 0.1344\,\mathrm{V}$ with $V_{cc} = 1\,\mathrm{V}$.

Previous AVG-based work [17] assumes homogeneous input drifts (all the inputs have the same variability level σ_y), case in which a balanced weight set produces a weighted average y' with the minimum standard deviation $\sigma_{y'} = \sigma_y/\sqrt{R}$ and, therefore, the output error probability P_e is minimized. However, when the input drifts lose homogeneity due to aging and variability effects, the output standard deviation increases dramatically and so does the output error probability P_e.

In this paper, we study in detail the benefits that can be further obtained from the AVG structure by considering extra information related to the specific input variability levels. Accordingly, we propose the use of a variability monitor that provides the required measures. First, we assume ideal measurements from the monitor. In Section IV, we develop, further, the idea and outline a possible implementation together with the limitations that may result. The main objective of our proposal is to improve the fault-tolerant capabilities of the AVG structure by adjusting the values of the averaging weights in accordance with the monitor's information. This approach allows us to deal with scenarios of heterogeneous variability.

III. U-AVG

The AVG provides robustness when all inputs are under the same aggression factors, case in which balanced weights provide the best drift compensation. However, in practice, this may not be the case, as some replicas may have a larger drift with respect to others. As a consequence, due to this decompensation, the balanced weights approach becomes suboptimal. In this section, while preserving the AVG architecture, we consider the nonhomogeneity of aggression factors and degradation effects. We propose to adjust the AVG weighting scheme according to the following principle: assign greater weight to the less degraded and more reliable inputs, and lower weight to the ones that are more prone to be unreliable. Intuitively speaking, such an approach should improve the overall reliability.

In order to motivate the use of unbalanced averaging, we first analyze a simple case of heterogeneity when all inputs are exposed to homogeneous variability except one that has a higher standard deviation than the others. In this example, we demonstrate that a significant weighted average y' variability reduction, measured in terms of standard deviation $\sigma_{y'}$, can be achieved if we modify the weights according to our U-AVG proposal, when compared with the AVG with balanced weights. After this experiment, we analytically demonstrate the existence of optimal weights that minimize the output error probability P_e (or equivalently $\sigma_{y'}$, the standard deviation of the weighted average) in any possible heterogeneous variability scenario. Then, we find a general formula that allows us to calculate the optimal set of weights in any circumstance. We conclude this section with the theoretical redundancy level R savings that our U-AVG proposal can provide against the AVG.

A. Heterogeneous Variability Scenario (Simple Case)

In this experiment, we reproduce the most simple case of heterogeneous variability scenario. We simulate the AVG with homogeneous level of variability for all the inputs replicas except to the qth input; we assign standard deviation σ_q to the qth input while all the rest have the same standard deviation σ. In this experiment, we first consider the conventional AVG, which uses balanced weights $c_i = 1/R$, $i = 1, \ldots R$. In this case, the resulting output variance can be expressed as (8). This formula

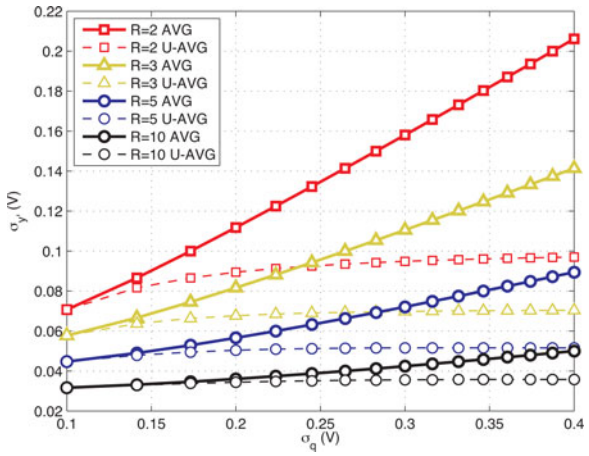

Fig. 3. Output standard deviation of AVG and U-AVG with 2, 3, 5, and 10 inputs subject to heterogeneous variability scenario. The scenario is modeled with a standard deviation $\sigma = 0.1$ V in all the inputs except from the qth that has σ_q ranging between 0.1 and 0.4 V.

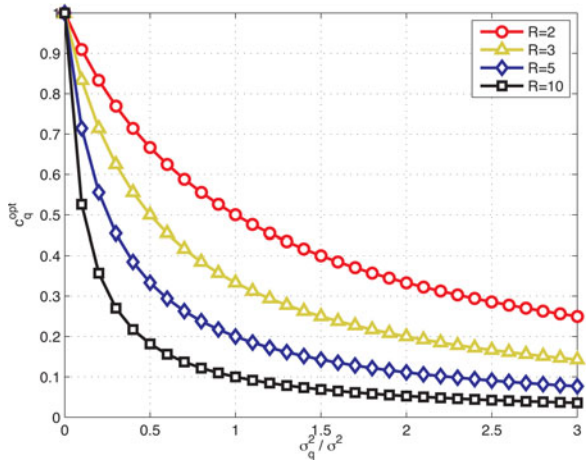

Fig. 4. Optimal weight c_q^{opt} against the ratio of input variances σ_q^2/σ^2 in a simple case of heterogeneous variability scenario. All the replicas have the same variance σ^2 except form the input qth that has variance σ_q^2.

is deduced from (6) and the σ_i model presented earlier

$$\sigma_{y'}^2 = \frac{R-1}{R^2}\sigma^2 + \frac{1}{R^2}\sigma_q^2. \tag{8}$$

Fig. 3 reproduces, in continuous lines, the analytic expression of the output standard deviation $\sigma_{y'}$ against different levels of variability in the qth input. One can observe that for $R = 3$, $\sigma = 0.1$ V, and $\sigma_q = 0.4$ V, the modeled heterogenous scenario implies more than doubling the output standard deviation with respect to the homogeneous case: from $\sigma_{y'} = 0.06$ V to $\sigma_{y'} = 0.14$ V.

Next we investigate how much better the overall performance can be when we properly configure the weight values c_i such that they reflect the fact that the qth input has a different variability. Using the general output variance expression in (6), an optimization problem can be derived. There are $R-1$ inputs with standard deviation σ and normalized weight c, while the remaining input has a higher standard deviation $\sigma_q > \sigma$ and a different normalized weight c_q. Note that $(R-1)c + c_q = 1$ must hold

$$\min(\sigma_{y'}^2)|_{c_q} \Rightarrow \frac{d}{dc_q}\left((R-1)c^2\sigma^2 + c_q^2\sigma_q^2\right) = 0. \tag{9}$$

Making the appropriate calculations to minimize the output variance $\sigma_{y'}^2$ by adjusting the value of weights c and c_q, the following expression for the optimum weight c_q^{opt} can be deduced

$$c_q^{\text{opt}} = \frac{1}{(R-1)\frac{\sigma_q^2}{\sigma^2} + 1}. \tag{10}$$

This formula, depicted in Fig. 4, clearly demonstrates that the optimal distribution of weights only depends on the ratio between the input variances, or equivalently on the relative reliability levels. In the particular case of $R = 2$, we can easily verify that the set of optimal weights is $(1/2, 1/2)$ when the variances or reliabilities are equal. If the variances ratio is higher than 1, then the input qth is less reliable and the optimal weight c_q^{opt} decreases. And vice versa, if the variances ratio is lower than 1, the optimal weight c_q^{opt} increases with respect to the equilibrated case. Fig. 4 also shows the impact of redundancy level R on the relationship between c_q^{opt} and the input variances. We observe that it corresponds to a compression to the origin with scale factor $R-1$.

The improvements achieved by the new weight configuration are depicted in Fig. 3 with dashed lines. One can observe in the figure that the U-AVG approach minimizes the output standard deviation compared to the AVG. Retrieving the previous example ($R = 3$ and $\sigma = 0.1$ V), but now considering the U-AVG approach, we can observe that an increase in the qth standard deviation σ_q from 0.1 to 0.4 V only increases the output standard deviation from 0.06 to 0.07 V. Thus, in this particular example, a 50% net reduction in the output standard deviation $\sigma_{y'}$ is achieved by the U-AVG with respect to the AVG. Moreover, one can deduce from the figure that the adjustment of weights is critical for low R (redundancy) values, which suggest that U-AVG may potentially require a lower redundancy than AVG for the same targeted reliability.

B. Existence of Optimal Weighted Averages

Continuing with the idea of making use of nonbalanced weights, we demonstrate in the following that there always exists a set of weights c_i^{opt}, $i = 1, \ldots, R$, that optimally minimizes the output error probability P_e, or equivalently the standard deviation of the weighted average $\sigma_{y'}$. First, we present a Lemma that displays this idea for individual input replicas.

Lemma 1: Given a weighted average

$$y' = \frac{1}{\sum_{i=1}^{R} k_i}\sum_{i=1}^{R} k_i y_i$$

with positive weights $k_i \geq 0$ and R input random variables y_i, $i = 1, \ldots R$, following statistically independent normal distributions. There always exists a unique value k_q^{opt} for the qth weight, $q \in (1, \ldots R)$, that optimally minimizes the standard deviation of the weighted average $\sigma_{y'}$.

Proof: Recalling the expression of the weighted average variance in (6), and substituting the definition of normalized weights ($c_i = k_i / \sum_{j=1}^{R} k_j$), we can express the dependence of $\sigma_{y'}^2$ on the weight k_q

$$\sigma_{y'}^2 = \frac{k_q^2\sigma_q^2 + \sum_{i=1,i\neq q}^{R} k_i^2\sigma_i^2}{\left(k_q + \sum_{i=1,i\neq q}^{R} k_i\right)^2}. \tag{11}$$

Differentiating with respect to k_q and matching to zero, we get the following optimal value for the qth weight:

$$\frac{d\sigma_{y'}^2}{dk_q} = 0 \Rightarrow k_q^{\text{opt}} = \frac{1}{\sum_{i=1,i\neq q}^{R} k_i} \sum_{i=1,i\neq q}^{R} k_i^2 \frac{\sigma_i^2}{\sigma_q^2}. \tag{12}$$

Differentiating two times with respect to k_q and substituting $k_q = k_q^{\text{opt}}$, we obtain a positive value. This confirms that indeed the value obtained k_q^{opt} corresponds to the unique weight value that optimally minimizes the $\sigma_{y'}^2$ value, and the standard deviation $\sigma_{y'}$ as well

$$\left.\frac{d^2\left(\sigma_{y'}^2\right)}{dk_q^2}\right|_{k_q=k_q^{\text{opt}}} = \frac{2\sigma_q^2 \sum_{i=1,i\neq q}^{R} k_i}{\left(k_q^{\text{opt}} + \sum_{i=1,i\neq q}^{R} k_i\right)^3} \geq 0.$$

□

Lemma 1 demonstrates that if we look at the value of a particular weight k_q, we can find a value that optimally minimizes the weighted average variance, and therefore minimizes the output error probability. In Fig. 5, we plot a sensitivity of the output error probability P_e to the modification of a specific weight in its normalized form c_q. For the analysis, we assume different levels of variability in the qth input modeled with the parameter σ_q ranging between 0.0 and 0.4 V. The rest of inputs are considered altogether with a fixed contribution to the weighted average variability when the value of weight c_q is null: $\sigma_{y'}|_{\{c_q=0\}} = 0.2$ V. One can observe in the figure, the different locations of the P_e minimum and the relation between optimal weights and different levels of variability. It clearly shows that it is possible to minimize the output error probability P_e if we properly tune the value of weight c_q.

The following conclusions can be drawn from Fig. 5.

1) There is always one and only one c_q value in the range from 0 to 1 that minimizes the error probability P_e in each possible variability environment.
2) The optimum c_q value is never exactly equal to 0. Even for large levels of deviation in the qth input with respect to the others, it is useful to have a contribution from the input q.
3) The optimum c_q value is never exactly equal to 1, except from the singular case with $\sigma_q = 0$ V, which in practice never occurs. Even for small levels of deviation in the qth input with respect to the others, it is useful to have a contribution from the rest of inputs.
4) To minimize the output error probability P_e, higher weights (close to 1) should be assigned to less deviating inputs and vice versa.

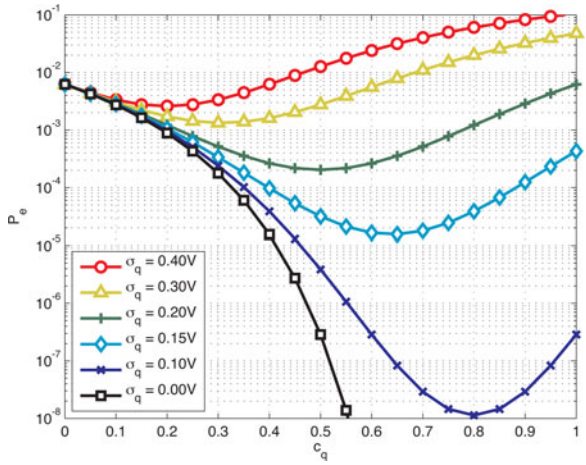

Fig. 5. Variation of the error probability P_e against the weight c_q. Standard deviation in the input q ranges from $\sigma_q = 0.0$ V to $\sigma_q = 0.4$ V and the standard deviation of the weighted average due the rest of inputs is $\sigma_{y'}|_{\{c_q=0\}} = 0.2$ V.

5) When the deviating effect in the qth input σ_q is equal to the deviating effect of the rest of inputs $\sigma_{y'}|_{\{c_q=0\}}$, the optimum c_q value is 0.5. Note in Fig. 5 that the curve with the minimum at $c_q = 0.5$ holds $\sigma_q = \sigma_{y'}|_{\{c_q=0\}}$.

So far we have proved the existence of an optimal value for a particular weight independently from the rest of the weights. However, the optimization of the entire set of weights requires an extra effort because the isolated optimization of averaging weights does not give rise to a global optimum. For example, if we optimize the values of the weights k_i from $i = 1$ to $i = R$ following the Lemma 1, we do not achieve a globally optimum. The optimal weight values found in the proof of Lemma 1 depend on the values of the other weights [see (12)]. Therefore, when we modify a weight the optimal value for the rest of weights changes. As we are looking for the globally minimum output error probability we need a further step.

In the following, we state and prove the Theorem that leads to the global optimization of weights. Before this, we want to note the existence of multiple sets of optimal weights if we use nonnormalized weights. Let us assume that the optimal set of weights that minimizes the variance $\sigma_{y'}^2$ of a particular AVG is $(k_1^*, k_2^*, \ldots, k_R^*)$. Then $(\alpha k_1^*, \alpha k_2^*, \ldots, \alpha k_R^*)$ is also an optimal set of weights. Indeed, as the weighted average variance only depends on the normalized weights [see (6)], any scaling operation in the set of weights does not affect its value. The following Theorem uses normalized weights c_i in order to avoid multiple solutions.

Theorem 1: Given a weighted average $y' = \sum_{i=1}^{R} c_i y_i$ of R variables normally distributed $y_i \sim N(y, \sigma_i)$, $\forall i \in (1 \ldots R)$, and averaging weights that satisfy the normalization condition $\sum_{i=1}^{R} c_i = 1$, there always exists a set of weights c_i^{opt}, $i = 1, \ldots, R$, that optimally minimizes the standard deviation of the weighted average $\sigma_{y'}$.

Proof: We use recursion to prove this Lemma. Let us express the variance of the weighted average as follows:

$$\sigma_{y'}^2 = \sum_{i=1}^{R} c_i^2\sigma_i^2 = c_R^2\sigma_R^2 + \sum_{i=1}^{R-1} c_i^2\sigma_i^2 = c_R^2\sigma_R^2 + (1-c_R)^2 \underbrace{\sum_{i=1}^{R-1}\left(\frac{c_i}{1-c_R}\right)^2\sigma_i^2}_{S_{R-1}^2}. \tag{13}$$

If we look at the S_{R-1}^2 definition, we observe that it corresponds to a weighted average with the first $R-1$ input replicas. The averaging weights defined as $c_i^* \equiv c_i/(1-c_R)$ add the unity, and therefore accomplish the normalization condition. Hence, S_{R-1}^2 can be regarded as a new minimization problem with $R-1$ normalized weights c_i^* independent from c_R.

Thereby, the $\sigma_{y'}^2$ minimization problem can be split into two independent subproblems (type 1 and type 2).

1) First, to minimize S_{R-1}^2 as a function of $R-1$ normalized weights c_i^*, $i=1,\ldots,R-1$.
2) Second, to minimize $\sigma_{y'}^2$ as a function of a single weight c_R ($\sigma_{y'}^2 = c_R^2\sigma_R^2 + (1-c_R)^2 S_{R-1}^2$). In this case, S_{R-1}^2 is regarded as a constant factor obtained from the first minimization subproblem.

The second subproblem presents a straightforward solution. The minimization of $\sigma_{y'}^2$ as a function of c_R, being S_{R-1}^2 independent from c_R, leads to the optimal value $c_R^{\text{opt}} = S_{R-1}^2/(S_{R-1}^2+\sigma_R^2)$. Substituting c_R^{opt} in the second derivative yields a positive value and confirms that c_R^{opt} corresponds to the unique minimum of the function.

In turn, the first subproblem can be split again into two independent subproblems as before (type 1 and type 2). If we keep splitting the problem this way, $R-2$ times in succession, we get a subproblem that cannot be split again. It corresponds to a function S_2^2 that only depends on two normalized weights and two input variances. Thus, it can be solved like the second subproblem type.

So far we have decomposed the problem into $R-2$ subproblems of type 2 plus one of type 1 that can be solved optimally. As a result, we have demonstrated that the whole problem can be solved optimally, and there exists an optimal set of weights that minimizes the weighted average variance. □

C. Optimal Unbalanced Weights

In the previous section, we demonstrated that there exists an optimal set of weights that minimizes the output error probability for the AVG structure. Therefore, there must be a way to express the optimal values of weights independently of the other weights and that is only function of the input variability levels. In order to find this formula, we perform the following analytic computation. We minimize the variance of the weighted average $\sigma_{y'}^2$, or equivalently its standard deviation that is directly related to P_e as (7) indicates. To perform this minimization considering all the weights c_i simultaneously, we have to use the Lagrange multipliers introducing the additional restriction of normalized weights. The target function is the variance $\sigma_{y'}^2$ and the variables to optimize are the magnitudes of the averaging weights c_i. The normalized weights condition ($\sum_{j=1}^{R} c_j = 1$) must hold. Therefore, the target function is

$$F(c_1, c_2, \ldots, c_R, \lambda) = \sigma_{y'}^2 - \lambda\left(\sum_{j=1}^{R} c_j - 1\right). \tag{14}$$

Differentiating with respect to the normalized weights c_i, $i=1,\ldots R$, recall (6), and the Lagrange multiplier λ, we obtain the following equations

$$\frac{d(F)}{dc_i} = 2c_i^{\text{opt}}\sigma_i^2 - \lambda, \qquad i=1,\ldots,R \tag{15}$$

$$\frac{d(F)}{d\lambda} = 1 - \sum_{j=1}^{R} c_j^{\text{opt}}. \tag{16}$$

Matching to zero (15), we obtain that the optimal weights are inversely proportional to the input variances

$$c_i^{\text{opt}} = \frac{\lambda}{2\sigma_i^2}. \tag{17}$$

Equation (16) equal to zero expresses the condition of normalized weights; combining both conditions, we deduce the value of λ

$$\lambda = \frac{2}{\sum_{j=1}^{R} 1/\sigma_j^2}. \tag{18}$$

Now we can calculate the explicit formula for the optimal weights

$$c_i^{\text{opt}} = \frac{1}{\sum_{j=1}^{R} \sigma_i^2/\sigma_j^2} \qquad i=1,\ldots,R. \tag{19}$$

This is the configuration of weights that optimally minimizes the error probability P_e. Observe that c_i^{opt} values are only dependent on the input variances σ_i^2. Depending on the input drifts distribution, each weight should be tuned according to (19) in order to achieve the lowest possible output error probability P_e. Now we can verify all the conclusions extracted from Fig. 5. We can see how the optimal weights are small when the input variance is large and vice versa.

We can also calculate the magnitude of the minimum possible variance for the weighted average that we obtain when we apply the optimal set of weights. Using (6) and the expression of optimal weights $c_i^{\text{opt}} = \lambda/(2\sigma_i^2)$, we get a closed expression for the minimum weighted average variance

$$\sigma_{y'\min}^2 = \frac{\lambda}{2} = \frac{1}{\sum_{j=1}^{R} 1/\sigma_j^2}. \tag{20}$$

Therefore, it is possible to express each optimal weight in terms of its input variance and the minimum possible variance of the weighted average

$$c_i^{\text{opt}} = \frac{\sigma_{y'\min}^2}{\sigma_i^2} \qquad i=1,\ldots,R. \tag{21}$$

The optimal configuration has all the weights directly proportional to the constant $\sigma^2_{y'\min}$ and inversely proportional to the respective input variance σ_i^2. We note that the particular case in which one or more inputs have null variability $\sigma_i = 0$ has to be treated separately. If this situation happens, then the output error probability minimization is straightforward: it would suffice us to assign the maximum weight to the input with null variability $c_q = 1$ and we would obtain the lowest possible output error probability, $P_e = 0$, according to (6) and (7). However, this can never reflect a real case as, in practice, there is always at least a small noise contribution that affects all the inputs.

D. AVG Versus U-AVG

To assess the implications of our proposal, we carry on a reliability analysis for AVG with balanced and unbalanced weights. Given a nonuniform input drift scenario, we calculate the percentage of circuits that satisfies the reference reliability requirement stated in Section 1 ($P_e < 10^{-4}$). We use this condition as a criterion to define the yield of the AVG circuit. In order to simulate realistic environments with heterogeneous variability, we generate the per replica drift variances following the Gamma distribution function

$$\sigma_i^2 \sim \Gamma(x;k,\phi) = \begin{cases} \frac{1}{\phi^k \Gamma(k)} x^{k-1} e^{-x/\phi}, & \text{if } x \geq 0 \\ 0, & \text{otherwise.} \end{cases} \quad (22)$$

We use scale parameter $\phi = 2$ and different shape values k in order to reproduce increasing values of the input replica variances. This probability distribution function allows us to efficiently generate positive random values of variability. Additionally, it presents the infinite divisibility property, very useful in Section IV-A where we recursively generate consecutive stages of degradation.

In the following experiment, we simulate the AVG behavior in different technology scenarios. Each scenario is modeled with heterogeneous input drift variances σ_i^2 following the Gamma distribution function with different mean values. We basically perform 10 000 Monte Carlo simulations and estimate the yield for both architectures (AVG and U-AVG). Fig. 6 presents the simulation results against the redundancy factor. One can observe that U-AVG can deliver the same yield than AVG with a much lower redundancy level R. For example, if we required a 90% yield, given a variability scenario with $E\{\sigma_i\} = 3\sigma_{\max}$, we would need eight replicas with the AVG while only two with the U-AVG. This corresponds to a 4× redundancy saving.

Thus far we have demonstrated that if we know the distribution of deviations among the replicas at the design time, we can provide better reliability levels at lower cost by configuring the AVG weights accordingly. However, this improvement is optimal only for certain static variability conditions of the system. If the levels of variation gradually change in time due to degradation and external aggressions, the unbalanced design may become suboptimal. In the next section, we make a step further and propose a dynamically adapting structure that modifies the weight values at run-time in order to keep track with the possible variations of the input deviations.

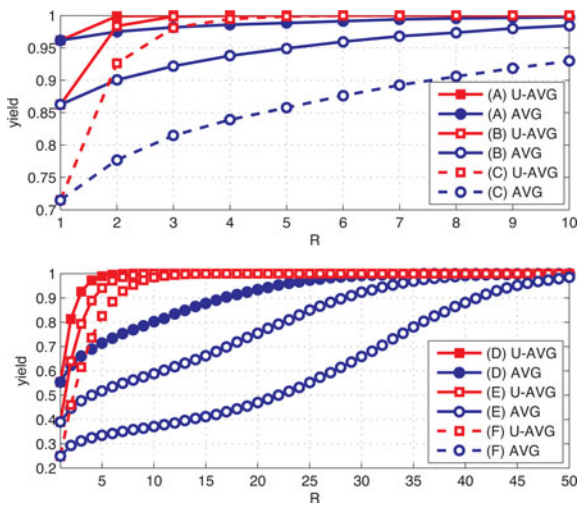

Fig. 6. Yield analysis of AVG and U-AVG against the redundancy factor for different technology scenarios. The simulated scenarios identified with letters (from A to F) have associated heterogeneous input variances following the Gamma distribution. The mean values for the standard deviation range from $E\{\sigma_i\} = \sigma_{\max}$ in scenario A to $E\{\sigma_i\} = 6\sigma_{\max}$ in scenario F. $\sigma_{\max}$ corresponds to the reference value 0.1344 V.

IV. AD-AVG

In the previous sections, we analyzed the reliability of AVGs in static environments of variability. We introduced the concept of heterogeneous variability scenario and we demonstrated the advantages of configuring the averaging weights according to the different input levels of variability. In this section, we also consider time variation of the input levels of variability in order to take into account the effects of aging and external aggressions, such as temperature, and dynamically adjust the AVG configuration such that we can tolerate the maximum possible amount of variation. In this line of reasoning, we add to the AVG the capability of dynamically reconfiguring the averaging weights according to the time-varying input variabilities measured at run-time. Our proposal is based on a reconfigurable crossbar of resistive switching devices and a disagreement detector like the one suggested in [23]. In Section V, we develop in detail the proposed technology implementation. We call this dynamic architecture the AD-AVG. In the following we analyze the main characteristics of the AD-AVG structure: redundancy overhead, reliability, and tolerance against degradation.

A. Degradation Model

In order to study the AD-AVG behavior in front of the degradation effects and external aggressions, we need a model to reproduce time-varying heterogeneous input drifts. The aging process and the dynamics of input variability levels depends on the particular technology used and the environmental conditions. However, based on the main properties of degradation and using a general expression for the probability distribution function, we can construct the required degradation model. With this purpose, we use the Gamma distribution function to generate positive random values of variability, see (22). Using the infinite divisibility property of this distribution, we can easily

simulate the influence of increasing amounts of degradation by simply adding random Gamma-distributed increments to the initial input variances.

For simulation purpose, we use the following methodology:

1) First, we generate the input variability levels using the Gamma distribution function [see (22)]. These levels of variability correspond to the initial imperfections already present in fresh devices (nonutilized) mainly associated to the manufacturing process. In the simulations, this initial stage of the circuits corresponds to the 0 in the axis of degradation in time (horizontal axis) and we assumed a mean value of $E\{\sigma_i^2\} = (2\sigma_{\max})^2$ $(= 0.0724\,\mathrm{V}^2)$.
2) Then, in order to simulate the AD-AVG at different moments of the replica's life, we need to calculate the input variabilities after the effect of increasing amounts of degradation. To do so, we estimate recursively the input variances at consecutive stages of degradation. We basically calculate the replicas' variance in the stage $n+1$ by adding positive random increment α_i to the variance in the previous degradation stage n:

$$\sigma_i^2[n+1] = \sigma_i^2[n] + \alpha_i[n]. \tag{23}$$

The increments used to update the input variances are also generated with the Gamma distribution function and reflect the effect of degradation occurred during the time between consecutive stages of degradation. In the simulations, we define the degradation in time unit so that it corresponds to a mean increase in the replicas' variance σ_i^2 of magnitude $E\{\alpha_i\} = (\sigma_{\max})^2$ $(= 0.0181\,\mathrm{V}^2)$. We use this normalized time scale because the exact relationship between degradation and time is too complicated and depends on the particular technology used, the stress applied to the system and other environmental conditions.

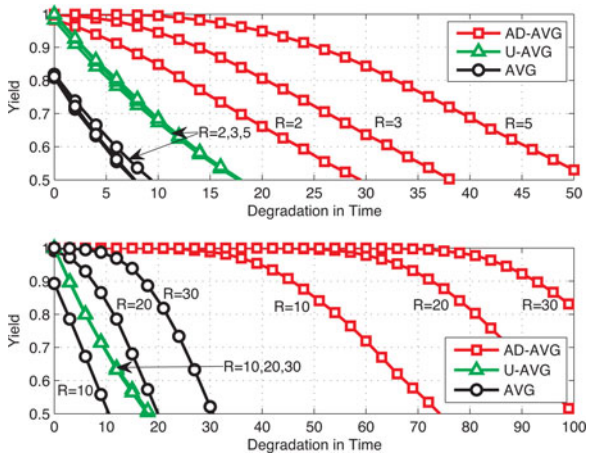

Fig. 7. Yield of different size AD-AVG, U-AVG, and AVG against the amount of accumulated degradation.

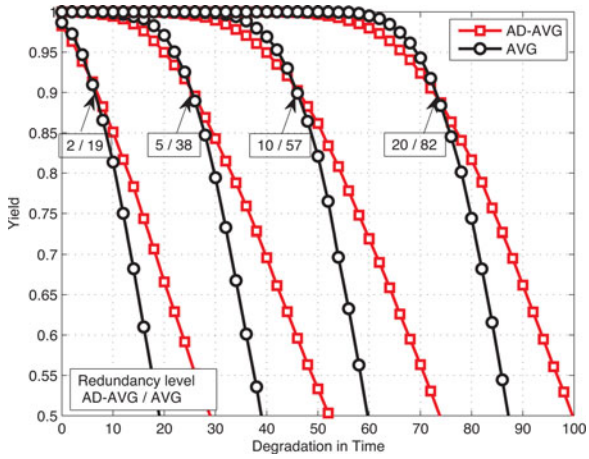

Fig. 8. Yield analysis of different size AD-AVG and AVG against degradation. Redundancy levels are chosen to meet the reliability requirements: 90% yield after 7, 25, 45, and 75 degradation in time units.

B. Learning Versus Static AVG

Using the described degradation model, we can simulate realistic situations of circuits that degrade in time. In the following, we analyze the difference between using adaptive weights (AD-AVG) and static weights (AVG and U-AVG). To do so, we perform 10 000 Monte Carlo simulations of the AVG structures with increasing amounts of degradation. Fig. 7 depicts the simulation results of yield for different size AVG circuits against the degradation.

We observe that at degradation in time 0, there is no difference between AD-AVG and U-AVG; both structures are configured with the optimal set of weights and the AVG yield is maximized. However, as the degradation increases, the U-AVG loses yield quickly regardless of the redundancy level. The U-AVG weights configuration is static and only optimal for the initial input variability levels; when the circuit degrades, the configuration of weights becomes suboptimal and after 18 degradation in time units the U-AVG yield drops below the 0.5. On the other hand, the AD-AVG technique is capable of tolerating much higher amounts of degradation and it improves notoriously with the redundancy level R. We can conclude that the AD-AVG always provides the maximum yield given any particular redundancy factor R.

Comparing the characteristic curves of U-AVG and AVG in Fig. 7, we also observe that the unbalanced approach is better than the AVG only for low redundancy levels (< 15). Since the U-AVG does not improve with redundancy but AVG does, in the conditions of this experiment, the AVG outperforms U-AVG for redundancy levels higher than 15. This is an interesting property of the static weight approaches. If we configure the weights according to the information of variability levels of fresh devices, the U-AVG will provide always better results at degradation in time 0. However, as we consider higher levels of redundancy, the AVG happens to outperform the U-AVG against degradation. For this reason, we dismiss the U-AVG technique and focus now on the comparison between AVG and AD-AVG.

Fig. 8 depicts the results of 10 000 Monte Carlo simulations comparing the AD-AVG versus AVG, but instead of using the same redundancy level, we impose the same reliability target under specific degradation conditions. The figure clearly demonstrates that the AD-AVG consumes less redundancy than

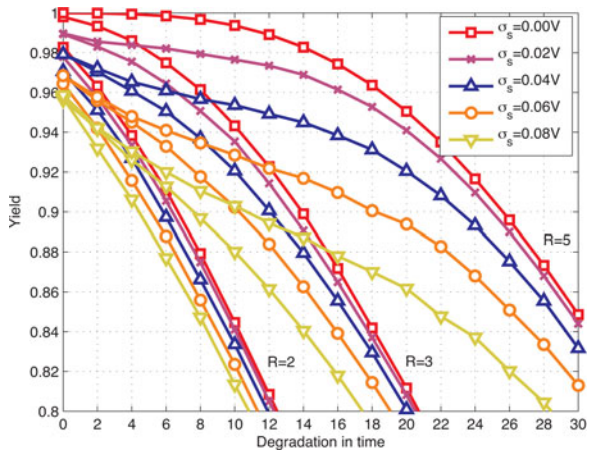

Fig. 9. Yield analysis of different size AD-AVG against degradation. Simulations include the effect of different noise levels in the variability monitor (from $\sigma_s = 0$ V to $\sigma_s = 0.08$ V).

AVG: from $9.5\times$ redundancy saving for a 90% yield after seven degradation in time units to $4.1\times$ redundancy saving for a 90% yield after 75 degradation in time units.

C. Noisy Variability Monitor in the AD-AVG

Thus far we assumed that the information which is driving the learning process is provided by an ideal variability monitor. In practice, this cannot be the case as there are many sources of noise in a real environment such as temperature, external radiation, interference within the same operating circuit, and discretization noise. We model these effects as an additive white Gaussian noise with standard deviation σ_s. In the following experiment, we analyze the influence of this noise in the effectiveness of the learning process. We perform 10 000 Monte Carlo simulations of the AD-AVG against degradation including the effect of different levels of noise in the variability monitor. Fig. 9 reproduces the obtained results for the yield of different size AD-AVG.

As expected, the noise added to the measures provided by the variability monitor worsens the characteristic reliability of the AD-AVG. We observe that the negative impact of noise increases with the redundancy level. For example, if we want a 90% yield in two-input AD-AVG with a noise affecting the monitor with standard deviation $\sigma_s = 0.06$ V, then the lifetime is reduced in two degradation in time units with respect to the noise-free case, whereas the same noise in the case of five-input AD-AVG reduces the lifetime in seven degradation in time units. From this experiment, we can conclude that the use of sensors to learn the variability changes over time is useful to improve the reliability of AVG structures as long as the noise in the variability monitor is small or comparable to the variability levels that it has to measure. The characteristic reliability of AD-AVG degrades with unreliable variability monitors, and therefore the particular tradeoff between reliability and overhead should be analyzed in each case.

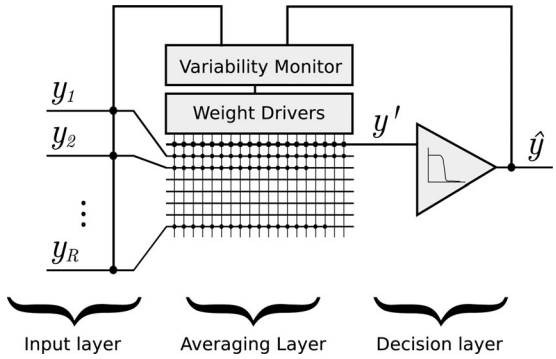

Fig. 10. AD-AVG implementation in switching resistive crossbar technology.

V. AD-AVG Implementation

In this section, we suggest a possible implementation for the AD-AVG technique. The structure graphically depicted in Fig. 10 is based on the use of a variability monitor, a set of weight drivers and a crossbar of switching resistive devices, such as memristors [24], [25]. The dynamic adjustment of weights requires, on one hand, a technique for gathering the required information and, on the other hand, a technology with high-reconfiguration capabilities to implement the adjustable weights. The use of resistive switching crossbars provides this reconfiguration feature and it represents a good candidate for future technology with a high-integration capability.

The whole architecture can be decomposed in three layers. The first one corresponds to the input layer and receives the input signals from the replicas. The second layer performs the averaging operation and is composed by the resistive switching crossbar, the variability monitor and the weight drivers. Finally, the third layer is the decision layer, it restores the binary output value by means of a threshold function.

A. Variability Monitor

In this section, we propose a topology for the variability monitor implementation. It is based on a disagreement detector between the AD-AVG output $\hat{y}$ and the signal provided by each replica y_i, $i = 1, \ldots R$. This mechanism was already used by Mathur and Avizienis in [23]. Checking at run-time, the differences between the AD-AVG output, which is statistically more reliable than the inputs, and each of the input replicas, we can establish a criterion for evaluating the relative variability of each replica. For example, we can estimate the standard deviation associated to each replica $\hat{\sigma}_i$ by counting the number of times n_i that the difference $\hat{y} - y_i$ exceeds a certain level L in a given number of clock cycles N. With this number, we can access a look up table that stores the relationship between n_i and $\hat{\sigma}_i$:

$$\hat{\sigma}_i = \frac{L}{\sqrt{2}\,\mathrm{inverfc}\,(2n_i/N)}, \qquad i = 1, \ldots, N \tag{24}$$

This mechanism allows us to keep the adapting process running during the circuit normal operation, and therefore it does not require any test implementation to obtain the information of input variability levels.

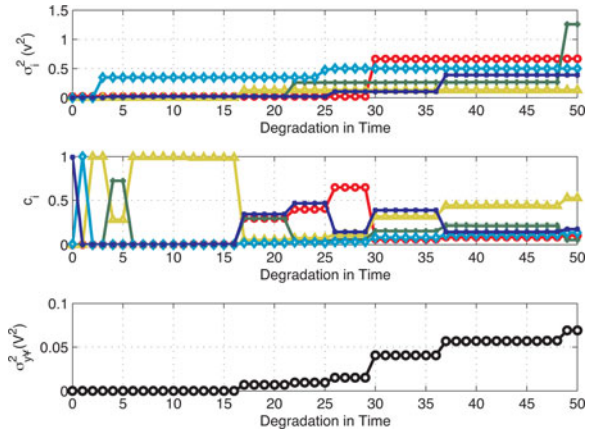

Fig. 11. Monte Carlo simulation of a five-input AD-AVG with the proposed topology. The subplots correspond to the five input replica variances σ_i^2, the five adaptive averaging weights c_i, and the variance of the weighted average $\sigma_{y'}^2$ against the degradation.

B. Weight Drivers and Switching Resistive Crossbar

In order to implement adjustable averaging weights, we propose the use of switching resistive crossbars. These structures consist of a grid of vertical and horizontal metal lines in which every cross point has a resistive switching element such as a memristor. These elements are capable of changing its characteristic resistance between two different values according to its state, R_{on} and R_{off}. In the case of memristors $R_{\text{on}} \simeq 1\,\text{M}\Omega$ and $R_{\text{off}} \simeq 1\,\text{G}\Omega$. This switching behavior can be easily controlled with the weight drivers applying specific configuring voltages to each of the vertical and horizontal lines. Using this feature, we can set up a network of interconnects that averages the input replicas with specific reconfigurable averaging weights.

Given a redundancy factor R from the initial stages of the AD-AVG design, this parameter is associated with the number of horizontal metal lines in Fig. 10. The number of vertical metal lines N determines the maximum number of interconnects per input, and therefore the maximum weight value. As a consequence, N parameter has a direct impact on the maximum accuracy when establishing the weight values. The total area in the crossbar structure is proportional to the number of vertical metal lines and the redundancy factor (Area $\propto R \times N$). Weight drivers receive the information regarding the relative replica reliabilities generated by the variability monitor. Based on these data, they connect or disconnect a different number of cross-points in the switching resistive crossbar. In order to increase the weight value of a particular input, they increase the connected cross-points and vice versa, to decrease the weight value they disconnect cross-points in the corresponding horizontal line. Apart from the configurable region of $R \times N$ metal lines, the switching resistive crossbar also has an additional horizontal metal line connected to all the vertical metal lines. This line corresponds to the average output y', see the top horizontal line in Fig. 10. This line is fed to a threshold function that amplifies the signal and restores the binary value.

In order to prove the feasibility of the proposed implementation, we perform a Monte Carlo simulation of a five-input AD-AVG with the described topology. The simulation uses the degradation model introduced in Section IV-A. Fig. 11 depicts the temporal evolution of the input variances σ_i^2, the adaptive averaging weights c_i, and the variance of the weighted average $\sigma_{y'}^2$. We can observe how the input variances keep increasing over time and how the averaging weights are modified in order to minimize the weighted average variance.

VI. Conclusion

This paper introduces an AD-AVG structure targeting the reliability improvement of nanoscale circuits and systems implemented in future technologies with heterogeneous variability. First, the potential advantages of adjusting the weights configuration according to the different input variability levels are analyzed and a method for determining the optimal averaging weight values that maximize the reliability of the AVG output is derived. This approach is called the U-AVG. Monte Carlo simulations of various U-AVG structures are performed in scenarios with heterogeneous variability and it is demonstrated that U-AVG substantially improves the output reliability at a lower cost than the classic balanced AVG, i.e., it requires $4\times$ less redundancy for the same specific reliability requirement. Then, the U-AVG scheme is further extended by proposing a methodology to on-line learn the temporal variations on the input drift levels induced by external aggressions and aging, when the circuit is deployed in the field. This learning augmented scheme is called the AD-AVG. Further, Monte Carlo simulations of different AVG, U-AVG, and AD-AVG instances in nonstatic heterogeneous environments are performed and provide meaningful information about the tolerance against degradation of the different AVG approaches. It is observed that the U-AVG rapidly loses performance with the circuits degradation and it is outperformed by the conventional AVG for redundancy factors higher than 15. Comparing the characteristic reliability of the AVG versus the AD-AVG against degradation, significant redundancy savings are obtained for the adaptive approach: from $9.5\times$ to $4.1\times$ redundancy reduction when imposing reliability requirements of 90% yield after increasing amounts of accumulated degradation. Apart from the significant advantages in terms of redundancy overhead savings and improved tolerance to degradation, the implications of using a nonideal variability monitor for the learning process are also studied. Based on this analysis, it is observed that monitor's noise has a greater negative effect on AD-AVGs with larger redundancy factors. Given a requirement of 90% yield with variability monitor with noise level $\sigma_s = 0.06\,\text{V}$, the lifespan of two-input AD-AVG reduces two degradation in time units while that of five-input AD-AVG reduces seven degradation in time units. Finally, a possible implementation of the AD-AVG structure based on resistive switching crossbars is sketched. The proposed topology is capable of reconfiguring the averaging weights at run-time and to measure the changes in the input variability levels. A simulation of the final AD-AVG implementation is provided. It is demonstrated that the AD-AVG proposal is potentially implementable with state-of-the-art technology.

APPENDIX

SYSTEMATIC EFFECTS

In this paper, no systematic effects have been considered when quantifying the reliability level. However, future scenarios targeted here will probably include this kind of effects. Fluctuations in temperature and voltage as well as systematic variations from manufacturing processes usually imply coherent deviations in many devices at the same time. In this appendix, we prove that even when taking into account these systematic effects, the results provided in this paper are correct.

In Section II, the statistical model of drifts was developed assuming no bias. The consideration of systematic effects can be taken into account by adding a nonzero mean to the drift variables $\eta_i \sim N(\delta, \sigma_{\eta_i})$. Applying this change in all the replicas, the statistical model of drift distribution becomes

$$f_{\eta_i}(\eta_i) = \frac{1}{\sqrt{2\pi\sigma_{\eta_i}^2}} e^{-\frac{1}{2}\frac{(\eta_i-\delta)^2}{\sigma_{\eta_i}^2}}. \tag{25}$$

As a consequence of this, the fault accumulation level ϵ will have also mean value δ and the error probability will have the following expression [see (7)]:

$$P_e = \int_{V_{cc}/2}^{\infty} f_\epsilon(\epsilon) d\epsilon = \frac{1}{2} \times \operatorname{erfc}\left(\frac{V_{cc}/2 - \delta}{\sqrt{2\sigma_\epsilon^2}}\right). \tag{26}$$

From (26), we can draw the conclusion that systematic effects and random effects imply a reliability worsening in independent ways. The effect of systematic deviation in devices is translated into a reduction in the margin of tolerable variation $V_{cc}/2 - \delta$. On the other hand, random fluctuation impacts on the standard deviation of the fault accumulation level σ_ϵ. When applying averaging technique with different weighting schemes, we can reduce the effect of random variations through the reduction of the standard deviation of the fault accumulation level, but this will not imply a change to the bias effect δ.

REFERENCES

[1] M. Mishra and S. Goldstein, "Scalable defect tolerance for molecular electronics," in *Proc. Workshop Non-Silicon Comput.*, Feb. 2002, pp. 75–85.

[2] S. Borkar, T. Karnik, S. Narendra, J. Tschanz, A. Keshavarzi, and V. De, "Parameter variations and impact on circuits and microarchitecture," in *Proc. 40th Annu. Design Autom. Conf.*, 2003, pp. 338–342.

[3] S. Borkar, "Electronics beyond nano-scale CMOS," in *Proc. 43rd Annu. Design Autom. Conf.*, 2006, pp. 807–808.

[4] S. Borkar, "Designing reliable systems from unreliable components: The challenges of transistor variability and degradation," *IEEE Micro*, vol. 25, no. 6, pp. 10–16, Nov. 2005.

[5] J. Von Neumann, "Probabilistic logics and the synthesis of reliable organisms from unreliable components," *Automata Studies*, vol. 34, pp. 43–98, 1956.

[6] J. Han and P. Jonker, "A system architecture solution for unreliable nanoelectronic devices," *IEEE Trans. Nanotechnol.*, vol. 1, no. 4, pp. 201–208, Dec. 2002.

[7] S. Roy and V. Beiu, "Multiplexing schemes for cost-effective fault-tolerance," in *Proc. 4th IEEE Conf. Nanotechnol.*, 2004, pp. 589–592.

[8] D. Bhaduri, S. Shukla, P. Graham, and M. Gokhale, "Comparing reliability-redundancy tradeoffs for two von neumann multiplexing architectures," *IEEE Trans. Nanotechnol.*, vol. 6, no. 3, pp. 265–279, May 2007.

[9] K. Nikolic, A. Sadek, and M. Forshaw, "Fault-tolerant techniques for nano-computers," *Nanotechnology*, vol. 13, pp. 357–362, 2002.

[10] J. Han and P. Jonker, "A defect-and fault-tolerant architecture for nanocomputers," *Nanotechnology*, vol. 14, no. 2, pp. 224–230, Feb. 2003.

[11] P. Depledge, "Fault-tolerant computer systems," *IEE Proc. A Phys. Sci. Meas. Instrum. Manage. Educ. Rev.*, vol. 128, no. 4, pp. 257–272, 1981.

[12] M. Stanisavljevic, A. Schmid, and Y. Leblebici, "Optimization of nanoelectronic systems' reliability under massive defect density using cascaded R-fold modular redundancy," *Nanotechnology*, vol. 19, no. 46, pp. 1–9, Nov. 2008.

[13] M. Stanisavljevic, A. Schmid, and Y. Leblebici, "Optimization of nanoelectronic systems reliability under massive defect density using distributed R-fold modular redundancy (DRMR)," in *Proc. 24th IEEE Int. Symp. Defect Fault Tolerance Very Large Scale Integr. Syst.*, 2009, pp. 340–348.

[14] J. Han, E. Boykin, H. Chen, J. Liang, and J. Fortes, "On the reliability of computational structures using majority logic," *IEEE Trans. Nanotechnol.*, vol. 10, no. 5, pp. 1099–1112, Sep. 2011.

[15] F. Martorell and A. Rubio, "Cell architecture for nanoelectronic design," *Microelectron. J.*, vol. 39, no. 8, pp. 1041–1050, Aug. 2008.

[16] A. Schmid and Y. Leblebici, "Robust circuit and system design methodologies for nanometer-scale devices and single-electron transistors," *IEEE Trans. Very Large Scale Integr. Syst.*, vol. 12, no. 11, pp. 1156–1166, Nov. 2004.

[17] F. Martorell, S. Cotofana, and A. Rubio, "An analysis of internal parameter variations effects on nanoscaled gates," *IEEE Trans. Nanotechnol.*, vol. 7, no. 1, pp. 24–33, Jan. 2008.

[18] M. Stanisavljevic, A. Schmid, and Y. Leblebici, "Optimization of the averaging reliability technique using low redundancy factors for nanoscale technologies," *IEEE Trans. Nanotechnol.*, vol. 8, no. 3, pp. 379–390, May 2009.

[19] (2010). International technology roadmap for semiconductors, 2010. [Online]. Available: http://www.itrs.net/Links/2010ITRS/Home2010.htm

[20] N. Aymerich, S. Cotofana, and A. Rubio, "Adaptive fault-tolerant architecture for unreliable device technologies," in *Proc. 11th IEEE Conf. Nanotechnol.*, 2011, pp. 1441–1444.

[21] W. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bull. Math. Biophys.*, vol. 5, pp. 115–133, 1943.

[22] W. Pitts and W. McCulloch, "How we know universals: The perception of auditory and visual forms," *Bull. Math. Biophys.*, vol. 9, pp. 127–147, 1947.

[23] F. Mathur and A. Avižienis, "Reliability analysis and architecture of a hybrid-redundant digital system: Generalized triple modular redundancy with self-repair," in *Proc. Spring Joint Comput. Conf.*, May 5–7, 1970, pp. 375–383.

[24] G. Snider, "Computing with hysteretic resistor crossbars," *Appl. Phys. A*, vol. 80, no. 6, pp. 1165–1172, Mar. 2005.

[25] J. Borghetti, Z. Li, J. Straznicky, X. Li, D. Ohlberg, W. Wu, D. Stewart, and R. Williams, "A hybrid nanomemristor/transistor logic circuit capable of self-programming," *Proc. Nat. Acad. Sci.*, vol. 106, no. 6, pp. 1699–1703, Feb. 2009.

Nivard Aymerich received the B.Sc. and M.Sc. degrees in electronic engineering from the Polytechnical University of Catalonia (UPC), Barcelona, Spain, where he is currently working toward the Ph.D. degree in the Department of Electronic Engineering.

He is currently a Research Assistant in the Department of Electronic Engineering, UPC. His research interests include fault-tolerant circuits, reliable computing, and degradation-aware architectures based on redundancy.

Sorin D. Cotofana (SM'00) received the M.S. degree in computer science from the "Politehnica" University of Bucharest, Bucharest, Romania, and the Ph.D. degree in electrical engineering from the Delft University of Technology, Delft, The Netherlands.

For a decade, he was with the Research and Development Institute for Electronic Components, Bucharest. He is currently an Associate Professor with the Computer Engineering Laboratory, Delft University of Technology. His research interests include various topics from nanoelectronics and nanodevice specific design methodologies and computational paradigms to fault-tolerant computing, embedded systems, reconfigurable computing, computer arithmetic, low-power hardware, and multimedia and vector architectures and processors.

Antonio Rubio (SM'96) received the M.S. and Ph.D. degrees from the Industrial Engineering Faculty, Polytechnical University of Catalonia (UPC), Barcelona, Spain.

He has been an Associate Professor in the Electronic Engineering Department, Industrial Engineering Faculty, UPC, and a Professor in the Physics Department, Balearic Islands University. He is currently a Professor of electronic technology at the Telecommunication Engineering Faculty, UPC. His research interests include very large scale integration design and test, device and circuit modeling, and high-speed circuit design.